



Vol -1 Issue -1



BIIC-2022

Seminar
Proceedings

ALL INDIA SEMINAR ON
“BUSINESS INTELLIGENCE
WITH INTELLIGENT
COMPUTING”

March 4-5, 2022

Organized By

The Institution of Engineers (India)
Punjab & Chandigarh State Centre

in Association with

Department of Computer Science
and Applications

&

Department of Commerce and
Economics

Babbar Akali Memorial Khalsa College
Garhshankar, Hoshiarpur



Message



I'm glad to know that all India Seminar on "Business Intelligence with Intelligent Computing" is being organized by The Institution of Engineers (India) in association with B.A.M. Khalsa College, Garhshankar, Hoshiarpur on March 4-5, 2022

I congratulate The Institution of Engineers (India) to come up with this initiative of identifying new and innovative thoughts among the academicians and researchers related to the field of computer sciences, engineering, commerce and management from all around the world.

The flow need of combining education with research and industry is paramount and conceivable just through setting out such open doors for faculty, scholars, students and industrial experts having a place with different parts of the world to meet up and learn in a cooperative climate. I believe that the members will make the best out of this opportunity.

I wish B.A.M. Khalsa College, Garhshankar the very best for the seminar and for many other such drives that they will get ready for times to come.

**Prof. (Dr.) T.S. Kamal
Former Vice President, IE(I)**

Message



It gives me immense pleasure that IEI Punjab & Chandigarh State Centre is co-organizing the all India seminar from 4-5 March 2022 on the topic "Business Intelligence with Intelligent Computing" with the Department of Computer Science and applications and the Department of Commerce and Economics of Babbar Akali Memorial Khalsa College Garhshankar.

Businesses are moving at neck-breaking speeds and so is their competition. For these businesses to have an edge in the market, every decision they make must be informed. Irrespective of the industry, every business has access to a lot of data that they can leverage to their advantage. Business Intelligence (BI) helps these businesses to use their data to their advantage by presenting the otherwise unusable pile of data into an understandable and interpretable form. BI enables you to combine the power of technology and business expertise to make fully informed decisions and stay ahead of the competition.

Intelligent computing research aims at bringing intelligence, reasoning, perception, information gathering and analysis to computer systems. An intelligent decision rule has been used to generate the coefficients of the fuzzy constraints in the decision variables.

The recent research made by Professors/ Scholars/ Students would be a virtual feast for new thought in the respective domains. It will hopefully provide a fresh and incisive vision and add a new fillip to the process of 'Business Intelligence'. A new chapter of knowledge would set a promising trend to the ongoing research in the field of 'Business Intelligence' on such broad canvas and vast scales.

Eventually, I express my special thanks and appreciation to all.

Er. S.S. Mundi

Chairman, IE(I)

Punjab & Chandigarh State Centre

Message



At B.A.M. Khalsa College, we firmly believe that the students are the fate of the nation and understanding this we make sure that every one of our attempts is consistently coordinated towards the development of our students.

My Gratitude is due to the Chairman and Convenor of The Institution of Engineers (India) Punjab & Chandigarh State Centre for the support that we are receiving.

As the co-convenor of this seminar, I feel elated we at B.A.M. Khalsa College are taking effective steps towards our objective of conferring quality education which caters to expanding the skill set of the students. I firmly feel that the present powerful climate demands a student to be industry-ready from the very beginning with legitimate idea understanding which can be applied in reality.

This seminar is a continuous endeavour of ours to portray the need for research, dynamic idea, creativity and development, vital for the growth of any organization or economy. I would like to give my warm regards to every one of the academicians, researchers and participants for being a part of this national seminar and my best wishes to the whole group behind the seminar for taking such an endeavour.

Prof. (Dr.) Baljit Singh Khehra

Principal, B.A.M. Khalsa College

Honorary Secretary, IE(I)

Punjab & Chandigarh State Centre

INDEX

Sr. No	Name of Author	Title of Research Paper/Abstract	Page No.
<i>Research Papers</i>			
1	Baljit Singh Khehra Ramanjot Kaur	Analysis of research involved hybrid fuzzy logic approach used for disease detection through images	1-8
2	Dr Baljit Singh Khehra Dr Harmandeep Singh Gill Parvinder Kaur	Detection of motorcyclists without helmet using deep learning	9-15
3	Baljit S. Khehra Amarinder Singh Jagmohan Kaur	ANN for heart disease diagnosis	16-19
4	Kanwar Kulwant Singh	Social Media Analytical in The Perspective of Big Data	20-25
5	Dr Janki Aggarwal	Artificial Intelligence Reshaping the Automotive Industry	26-33
6	Dr Gurpreet Singh	Significance of business intelligence behind customer loyalty	34-38
7	Dr Arvinder Kaur	ICT channel - a profitable impact on banking sector	39-42
8	Gurpreet Singh Kalsi	Big data analytics for E-commerce: Amazon's impactful use of big data	43-46
9	Dr Jasminde Kaur	E-marketplaces minimize buyers' search costs: a critical review of reduced-price hypothesis	47-52
10	Manpreet Kaur Akriti	Big data analytics in e-commerce: understanding personalisation using social media platforms	53-56
11	Amandeep Kaur	Digital India: foundation for future India	57-62
12	Amandeep Kaur	E-commerce in India	63-67
13	Harpreet Kaur	E-commerce- swot analysis	68-72
14	Gurjinder Kaur	Role of big data in e-business	73-77
15	Mohit Kumar Dr Ajay Dutta Dr Rajneesh Kant Sachdeva	Implementation of cold store monitoring system (CSMS) software in cold store management system	78-87
16	Harmandeep Kaur	E-commerce trading technologies	88-92
17	Kamini	E-business services in India	93-97
18	Manpreet Singh Nikhil Chaudhary	Reviews on biometrics technologies	98-101

Sr. No	Name of Author	Title of Research Paper/Abstract	Page No.
<i>Research Papers</i>			
19	Dr Chetna Ms. Harpreet, Saloni	Blockchain technology to secure cyberspace	102-106
20	Shelly Garg Balkrishan Jindal	Segmentation techniques on melanoma skin cancer images	107-110
21	Kavita Verma Sunny Gulati	Digital India: an insight into new India	111-118
22	Shakti Kumar Anuj Kumar	Improved lesion detection using morphological operations	119-123
23	Silky Sachar Dr Anuj Kumar	Hybrid computer vision approach for identification of medicinal leaves	124-128
24	Muskan Kumari	A study of cloud computing	129-132
25	Harnoor Kaur Khehra	Parkinson disease detection using machine learning	133-139
26	Sandeep Kaur Parminder Kaur	A study of blockchain technology	140-144
27	Prabhjot kaur	Internet of things (IoT)	145-149
28	Dr Rakesh Kumar	Enterprise resource planning system for business management	150-154
29	Amrita Sekhon Dr Sukhwinder Singh	E-governance measures in Punjab post Covid - Quo Vadis?	155-160
30	Prof. Sandeep Kaur	Digital corporate governance in India	161-164
31	Navtej Singh Balkrishan	Weighted common spatial pattern method for feature extraction from EEG signals in BCI system	165-171
32	Dr Harmandeep Singh Gill	Various components and concepts related to fruit image enhancement using fuzzy	172-178
33	Parag Sharma	The role of artificial intelligence in cyber security	179-182
34	Aswin. P Dr Satwinder Singh Dilshad K	A systematic approach for speech to text translation and summarization	183-193
35	Dilshad Kaur Dr Satwinder Singh	Parallel data extraction from comparable data: a proposed architecture	194-202
36	Uttamjit Kaur	Role of big data for e-business –future opportunities, challenges and value creation	203-206

Sr. No	Name of Author	Title of Research Paper/Abstract	Page No.
<i>Research Papers</i>			
37	Harpreet Kaur	E-commerce: concepts and concerns	207-210
38	Rupinder Kaur	Classification techniques for crop disease detection	211-215
39	Neha Dhiman Kamini	Role of digitization in education	216-219
40	Manpreet Kaur	RFID technology and internet of things (IoT)	220-224
<i>Abstracts</i>			
41	Dr Sumeet Kaur Sukhvir Kaur	Digitalisation of rural Punjab	225
42	Jaswinder Kaur	Big data: possible revenue streams for social media platforms	226
43	Novedeep Jeerh	Artificial intelligence in education	227
44	Sarita Anuj Kumar	Pre-processing in automatic traffic light detection system for self-driving vehicles	228
45	Kumar Vaneet Jindal Balkrishan	Energy aware virtual machine allocation by fuzzy artificial bee colony algorithm in cloud data centers	229
46	Sukhjot Kaur	Applications of business intelligence	230

Analysis of Research Involved Hybrid Fuzzy Logic Approach used for Disease Detection through Images

Ramanjot Kaur

Research Scholar, Department of Computer Science & Engineering
I.K Gujral Punjab Technical University, Jalandhar
jot_klair@yahoo.co.in

Baljit Singh Khehra

Department of Computer Science & Engineering
Baba Banda Singh Bahadur Engineering College, Fatehgarh Sahib
baljitzkhera740@gmail.com

Abstract – Accurate detection plays an important role in the diagnostic process. In this paper, different hybrid methods are discussed and analysed, which are used to diagnose different diseases of different organisms. Soft computing algorithms provide precise and efficient output. The fuzzy-based process provides accurate results for various ailments like brain tumors, crop disease, breast tumors, dental issues, plant disease, and capsicum. Hybrid methods provide higher accuracy rate for classification of disease. In addition to disease detection, these methods are also provide better results for segmentation of remote sensing.

Index Terms – *Soft computing, Fuzzy logic, Hybrid, Genetic algorithm, Neuro-fuzzy.*

I. INTRODUCTION

With the advancement of technology, detection of accurate disease becomes possible. Lung cancer [1], breast tumor [2], brain tumor [3], liver tumor [4], COVID-19[5], etc. like diseases are diagnosed with the help of numerous image modality techniques like X-ray, MRI (Magnetic Resonance Imaging), and CT (Computer Tomography) scan[6]. Numerous diseases are detected with the use of Soft computing approaches related to human health, plant, leaf, crops, etc. Soft computing provides precision and handles uncertainty. Soft computing approaches like an artificial neural network, Genetic algorithm, fuzzy logic, and metaheuristic algorithms. Hybrid methods have been developed in recent years to improve the accuracy of results [7][8]. Some issues are like noise, incomplete information, and irrelevant data in the medical images during image modalities. Due to these

problems, the performance of classifiers goes down [9]. A segmentation process is required to assist the doctors for diagnosis purposes. Image segmentation is still a challenging task [10]. A number of methods have been developed till now to overcome this problem. In addition to human disease, a number of strategies were implemented for disease detection in plants [11]. Segmentation of plant diseases is required at early stages for better yield. A considerable amount of money is spent on the detection process in the early stages. More experts were needed to judge the diseases of plants in existing methods. Diseases in plants make the quality of crops very low. Automatic crop disease detection is done by using soft computing and image segmentation to overcome this drawback [12]. In this paper, fuzzy logic concepts, fuzzy-based techniques, and hybrid fuzzy models are discussed, which were used to diagnose different diseases related to humans and plants.

A. FUZZY LOGIC

Initially, the fuzzy term is introduced by Lotfi Zadeh [13]. Fuzzy logic is based upon fuzzy set theory. Classical set theory based upon the crisp values means either 0 or 1. Fuzzy logic gives precise results between 0 and 1 values. The fuzzy logic concept is used in various appliances of routine implements the human judgment process. Crisp input is provided for the fuzzification process. After that, fuzzy inference rules are applied to the fuzzification results to get the desired results. Resultant values are then converted into crisp values, which means the defuzzification process at this stage [14]. The flowchart for the fuzzy logic concept is shown in figure 1.

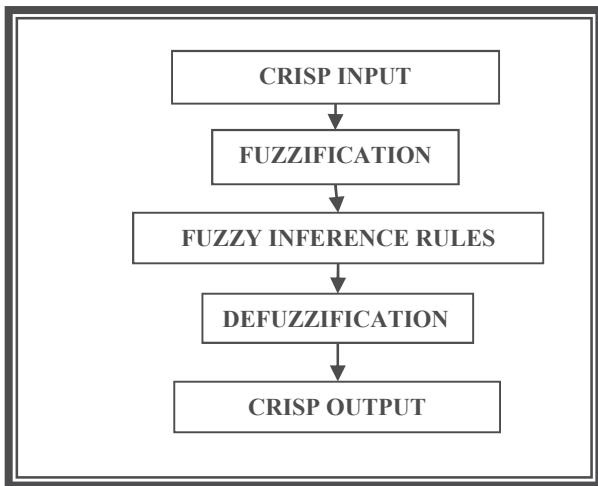


Fig. 1 Fuzzy logic Concept

B. FUZZY ADAPTIVE METHODS

Fuzzy logic is combined with other methods or cascades to make a hybrid fuzzy model. A Neuro-fuzzy model is a combination of artificial neural networks and fuzzy logic. Bayesian and fuzzy logic combined make a hybrid model of fuzzy Bayesian. A novel fuzzy-based method is developed for breast segmentation [15]. A convolution neural network does an accurate remote sensing image segmentation with the help of fuzzy logic [16]. The hybrid fuzzy-based method implemented for jaw lesions segmentation [17]. A hybrid fuzzy based optimization algorithm was deployed for classification of brain tumor [18]. An intelligent fuzzy set and particle swarm intelligence based approach was implemented for brain image segmentation [19].

II. RELATED WORK

Das et al. in [9] presents a hybrid Neuro-Fuzzy method. Due to the nature of data, which can contain partial, ambiguous, and imprecise information, medical disease classification using machine learning algorithms is a difficult process. The presence of information in the dataset influences the classification model of performance. A Linguistic Neuron-Fuzzy with the Feature Extraction model is used to analyze medical data for disease categorization in this paper. To begin, this

approach employs a linguistic fuzzy technique to generate membership values that address the problems of uncertainty. These membership values may not have a substantial impact on the model, but they will increase the dimensions, necessitating additional time to train the model. To overcome this problem, the Neuron-Fuzzy model uses Feature Extraction methods to extract only those features (a smaller feature set) that contribute significantly to the network. The Artificial Neural Network model is used to classify these reduced features once more. The performance of this suggested model is compared to that of other models after it has been evaluated and validated using eight benchmark datasets. For proof of correctness, the acquired results were evaluated using statistical techniques such as Friedman and Holm-Bonferroni. This experimental investigation demonstrates that proposed model outperforms other models.

Anjana et al. in [11] discussed that capsicum can be infected with a variety of bacterial, fungal, and viral diseases, and the signs of these ailments can be distinguished by analyzing the stem, leaf, or capsicum. The unhealthy region of the capsicum is detected using the k-means clustering technique that is further used the texture extraction along with various fungal diseases classified using support vector machine with the help of these features (SVM). For training and classification, many classifiers such as Tree, KNN, SVM and Linear Discriminant are utilized; however, KNN and SVM provide the best outcomes for our application. This system was tested on 62 photos of healthy or sick capsicum and its leaves, and SVM correctly categorized the images as healthy or diseased with a precision of 100. Accurate identification and classification of plant diseases, which can be accomplished using image-processing techniques, is critical for successful crop cultivation and increased agricultural production yield. In this study, the diseased impacted capsicum section is segmented first. Features of the infected part are extracted using a feature extraction technique, and then capsicum diseases are classified using an SVM classifier. The proposed method is evaluated for five diseases named anthracnose, powdery mildew, bacterial spot, grey, and Cercospora leaf-spot for

disease detection of its leaves. Healthy and affected capsicum can be distinguished with the help of precision of nearly 100% using the SVM classifier. The suggested method classifies healthy and infected capsicum.

Agricultural productivity are one of the most important factors in the economy. This could be the main reason to identify the maladies among plants is so crucial in the agriculture business because an ailment of plants is a natural occurrence. If proper look after is not provided in this area, it has a negative impact on plants, affecting the quality, quantity, and productivity of a variety of products. A disease known as a very little leaf illness or plant disease, for example, could be a dangerous disease present in trees such as pine fruit in the United States. Disease detection by some automated technology is beneficial since it reduces an enormous amount of labor in large farms of crops, and it detects disease symptoms at a very early stage. Singh et al. [12] proposed an algorithm mainly for image segmentation in order to use these to detect and categorize the diseases in the leaves automatically. Plant leaf diseases can be detected with several disease categorization systems that are analyzed in this study. The most commonly used genetic algorithm is implemented along with distinct classification techniques to complete the operation of image segmentation that plays a crucial role among plant leaf diseases.

Parkinson's disease (PD) is a devastating neurological disease that takes time to be diagnosed even at an early stage. The majority of practitioners such as radiologists and medical prefer to use the study of magnetic resonance imaging (MRIs) of PD patients to diagnose the condition. Pattern identification and visualization in MRIs were difficult due to the existence of gray scale characteristics and imprecise hereditary information. Huang et al. in [20] proposed a new technique to understand as well as visualize the distinct patterns in MRI images with this motivation. This study used FIG (fuzzy information gain) function and the Kmeans clustering technique to achieve this goal. The pixels information was

quantified using the FIG function, and the pixels information was clustered using the K-means clustering technique. Finally, MRI alterations were identified and divided in three different regions: an average changed region, the maximum changed region (MAXCR) as well as the minimum changed region (MINCR). The adaptive K-means clustering algorithm, the adaptive threshold approach, the grey threshold method, watershed method, K-means clustering algorithm, fuzzy based method as well as fuzzy c-means (FCM) algorithms. These were implemented to make a comparison of patients having PD in segmented MRIs. Among nine PD MRIs, the suggested technique had value of an average mean squared error by 63.49 while the peak signal-to-noise ratio was 30.14, and a Jaccard similarity coefficient up to 0.92. These entire coefficients were improved by 20.73 percent, -32.94 percent, and 3.54 percent -6.20 percent, and 6.98 percent -64.29 percent respectively.

Dental disease diagnosis is required because the majority of Indonesians have had dental disease at some point in their lives. According to Basic Health Research 2013, dental illness affects three areas: South Sulawesi, West Sulawesi, and South Kalimantan. Accurate detection is tough to achieve since it necessitates professional observations and interviews to improve perception. Dentists need accurate dental disease detection as a tool to help them enhance patient interaction and time efficiency. A method for obtaining a model capable of processing observation data is required for good and accurate detection. Parewe et al. in [21] presents a study that provides a solution based on a hybrid technique that incorporates both fuzzy logic and the evolution algorithm. For optimization, Evolution Strategies is utilized, which produces more accurate results than merely utilizing FIS Tsukamoto. The function of membership degree is the target of optimization. This can be used to classify the following dental conditions. Variance: pulpitis, gingivitis, periodontitis, and advanced periodontitis, calculated using the formula Root Mean Square Error (RMSE) 0.8.

The production of agricultural crops is one of the most important factors that contribute to an economy of any nation. Surprisingly, plant disease identification is also considered as the main aspect that is essential for the preservation of any country's economy. Kurmi et al. in [22] suggested algorithm's that contributes to optimize the information retrieved from different resources for improved results without adding any more complexity. Before performing the classification of any healthy or unhealthy image, the mentioned method can be used to look for the location of the leaf region that is affected. The main benefit of this study is that it helps to compare the outcomes collected from different sources with the motive of improvement in accuracy. The colour transformation is used to evaluate the leaf colours in order to identify the seed region. The spectral range of a low-dimensional RGB colour image is expanded by mapping it into color space. The early seeds are subjected to neighboring pixels-based leaf region growth. For appropriate curve fitting, we used a random sample consensus (RANSAC) to get the clear boundary of leaf as well as disease-affected areas. Following the extraction, a set of attributes related to visual words, handmade features and Fisher vectors, the multilayer perception model and classification using logistic regression along with the support vector machine is used. Plant Village datasets of bell pepper, apple, tomato, grape, cherry, potato and corn are used to evaluate the proposal's performance. The suggested contextualization-based image categorization technique performs out of expectation in the state of the art in a simulation-based examination [22].

Accurate, rapid, and automatic animal picture detection and classification is difficult, while it is critical for several real-world applications. Mohammed et al. in [23] provides a Mamdani model that is a hybrid type. Using separate datasets

containing roughly 27,307 photos, CNNs (convolution neural networks) and type-2 fuzzy rules were used to detect and distinguish diverse animals. The suggested system identifies the image with the assistance of fuzzy criteria before the

implementation of the CNN model to the predicate category of an object. More than 21,846 animal photos were used to train and test the CNN model. The proposed method studies revealed great speed and efficiency that may be a core factor in the system of image-processing linked to Type 2 fuzzy rules that describes the detection of the fix and moving images. Other investigations found that the proposed fuzzy technique has a 98 percent accuracy of identifying as well as recognizing moving objects with 0.1183464 value of mean square error. It also has a very high percentage of properly predicting dangerous items, with a recall of 0.98121 and a precision of 1. The F Score acquired a high percentage of 0.99052, was used to assess the test's correctness.

Salarika et al. in [24], The mechanical impact of field harvesting procedures and the stages of postharvest processing can impact pear fruit. Compared to non-damaged samples, bruised fruits had a lower postharvest shelf life and softened more quickly in cold storage. Understanding the dynamic reactivity of fruit to varied imposed loadings is crucial for developing techniques for preventing bruising during the supply chain. The FEM (Finite Element Method) is one of the most accurate and cost-effective methodologies for analyzing the components that cause impact-induced bruising. The simulation of the pear sample was performed by drop test. For this simulation, a 3D solid model of the pear was used to work under the technology of optical scanning. The primary purpose of the above-explained computer study was to evaluate distribution patterns of stress and strain within pear caused because of the fruit colliding with a flat surface constructed of various materials. It also looked at the contact force that exists between two forces those are produced after the collision of surfaces. Two adjacent drops oriented, as well as four distinct impacts on surfaces, were used in the simulations. The smallest and biggest stresses, contact forces and strains were created an impact on the surfaces such as steel and rubber, respectively, in both drop oriented. When fruit interacted with surfaces besides the horizontal axis, these parameters were generally lower than when fruit hit

with surfaces on the vertical axis. Consequently, strain and stress magnitude studies revealed that simulated stress and strain values were consistent with experimental data.

Hu et al. in [25] presents a study that aims to see how a healthy system based on fuzzy for medical image processing can predict brain illness. Because of the complex human brain tissues and their imaging technique like NMR (Nuclear Magnetic Resonance), brain MRI images include distinct weak borders, degrees of noise and artifacts. As a result, improvement is obtained with the fuzzy clustering algorithm. Hybrid Pyramid U-Net Model for Brain Tumour Segmentation (HPU-Net and improved fuzzy clustering is used for image processing and disease diagnosis prediction of brain. In simulation studies, brain MRI images taken from a hospital are utilized to validate the accuracy of the proposed method. The suggested algorithm has several nodes, more steady changes than previous models as well as less energy under similar conditions. The overall network performance suggested an algorithm that can finish data transmission jobs the quickest, averaging around 4.5 seconds, significantly faster than other models. Furthermore, the Jaccard coefficient of 0.845 demonstrates that it outperforms other models in terms of segmentation accuracy. In other words, the suggested algorithm may deliver a more noticeable denoising effect, perfect accuracy and more improved segmentation and recognition effect than existing models while consuming less energy. Even the findings could serve as the foundation of experiments for brain imaging feature recognition and predictive diagnosis.

In many application domains, image mining is critical in the decision-making process. Information management and processing include image mining. Productivity of the crop is reduced because of plant diseases, which influence the nation's social life and economics. The proper application of agriculture image mining can increase yield production while also benefiting the farmer and the country economically. The goal of the proposed approach deployed by Verma et al. was to use image mining

to automate the identification of rice illnesses for faster disease diagnosis. The disease-infected and disinfected plant photos are digitally taken and saved in a database with unique feature descriptors such as colour, texture. The collected five types of infected and one type of healthy photographs was recorded in the database in JPEG format in this study. Every type defines distinct image characteristics. In pre-processing processes, the captured images are accessible in RGB color space and cropped or scaled. For each image, the segmentation procedure creates three segments. Image analysis is used to extract 54 hybrid features, including 24 textures, 24 wavelets F-ratio of spatial-frequency components, and six color entropies. In wavelet features, the ANOVA analysis is used to determine the F-ratio. The retrieved features are fed into the CART, which uses the Gini index split point to choose important features. The CART developed a binary decision tree that has 54 attributes to just 13 that are relevant. For fuzzy filtering, the CART chose 13 attributes forwarded in FIS, reducing 13 attributes to 6 attributes. To construct a rice illness recognition model, fuzzy-filtered outcomes were utilized for training MLPNN using the Scaled Conjugate back-propagation training algorithm. The hybrid CART-FIS-MLPNN model has a training efficiency of 97.1 percent and a testing efficiency of 95.47 percent [26].

Climate variables mainly influence the growth of crops in greenhouses. It also has an impact quantity of water and fertilizer provided through irrigation. Control systems installed and the guidance of agricultural experts and farmers help manage these elements rapidly. Canadas et al. in [27] presents a study for decision support capabilities allow us to utilize priceless human expertise to make rapid decisions that make efficient crop growth certain. The supervisory stage look for climate sensor defects, the control phase manages climate variables at set points, and diagnoses diseases impacting the crop use strategic stage that alters climate variables accordingly. A real-time tool based on the rule was integrated into the control system. Experiments reveal that the technology improves climate management while also assisting in the prevention

of diseases that are tough to remove. Evaluating the manifestation of sickness and noticing the genuine system response were used to test the system. The key contribution has been to show how production rules can be used as a system-wide technique. It implies that elements like fault detection, temperature management, and illness monitoring are not handled separately. This article described a decision-making system that was used to monitor and control greenhouse crop climatic conditions. It has proven that a decision support system can be used to assist automated control systems in intense knowledge activities where the variables have a significant impact on crop development. The proposed system aids producers in making decisions by giving more data to climate control systems and a number of remedies to problems that are raised in the monitoring of climate conditions at the time of the growing season. In the system, three action steps were developed. The climate sensor failures were recognized at the alarm supervision stage, which provided a diagnosis that helped the farmer comprehend the cause of the error. Furthermore, reconfiguration was implemented, allowing the system to continue to function despite the discovered fault. By regulating window aperture ventilation, the temperature of the greenhouse was controlled. The graph reveals that the control system was in good operating order. Lastly, the disease management stage identified and diagnosed a variety of agricultural illnesses. In addition, the system was automatically modified for Botrytis. Greenhouse climate control, Fault detection and disease diagnosis have all been incorporated in this research. However, because they are enough complicated to be treated independently with new and more exact methodologies, further study will use multiple different routes in greater depth for each of the system's action levels.

III. CONCLUSION AND FUTURE SCOPE

In this paper, hybrid methods are discussed for detection of various ailments like brain tumor, breast tumor, Parkinson's disease, dental disease, crop disease etc. Capsicum disease classified with the help of support vector machine classifier. For

leaf disease detection, Genetic algorithm plays significant role. Artificial neural networks also helpful in classification process of disease. A fuzzy based hybrid approaches gives best and accurate results for classification and detection of disease. Type 2 fuzzy entropy based approach gives highest F score=0.99052. The main focus of this paper to provide best hybrid approaches for detection of diseases related to human, crops, leaves. These approaches provides early stage disease detection. In future, bayes classifier in hybrid models can be used for improvement in recognition rate. More feature extraction methods can be deployed in future.

ACKNOWLEDGEMENT

The authors are thankful for the support and guidance of I.K.Gujral Punjab Technical University, Jalandhar.

REFERENCES

- [1] K. S. Kumar, K. Venkatalakshmi, K. Karthikeyan, and M. E. Fantacci, "Lung Cancer Detection Using Image Segmentation by means of Various Evolutionary Algorithms," 2019, doi:10.1155/2019/490986.
- [2] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjoo, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020, doi:10.1080/00051144.2020.1785784.
- [3] S. Zhou, D. Nie, E. Adeli, J. Yin, J. Lian, and D. Shen, "High-Resolution Encoder-Decoder Networks for Low-Contrast Medical Image Segmentation," *IEEE Trans. Image Process.*, vol. 29, no. X, pp. 461–475, 2020, doi:10.1109/TIP.2019.2919937.
- [4] A. M. Anter and A. E. Hassenian, "Computational intelligence optimization approach based on particle swarm optimizer and neutrosophic set for abdominal CT liver tumor segmentation," *J. Comput. Sci.*, vol. 25, pp. 376–387, 2018, doi:

10.1016/j.jocs.2018.01.003.

- [5] R. Lokwani, A. Gaikwad, V. Kulkarni, A. Pant, and A. Kharat, "Automated detection of COVID-19 from CT scans using convolutional neural networks," *ICPRAM 2021 - Proc. 10th Int. Conf. Pattern Recognit. Appl. Methods*, vol. 41, no. 2, pp. 565570, 2021, doi:10.1016/j.bbe.2021.04.006.
- [6] A. Tiwari, S. Srivastava, and M. Pant, "Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019," *Pattern Recognit. Lett.*, vol. 131, pp. 244–260, 2020, doi: 10.1016/j.patrec.2019.11.020.
- [7] N. Devi *et al.*, "Categorizing diseases from leaf images using a hybrid learning model," *Symmetry (Basel.)*, vol. 13, no. 11, pp. 1–14, 2021, doi: 10.3390/sym13112073.
- [8] J. Kaur and B. S. Khehra, "Fuzzy Logic and Hybrid based Approaches for the Risk of Heart Disease Detection: State-of-the-Art Review," *J. Inst. Eng. Ser. B*, 2021, doi: 10.1007/s40031-021-00644-z.
- [9] H. Das, B. Naik, and H. S. Behera, "Medical disease analysis using Neuro-Fuzzy with Feature Extraction Model for classification," *Informatics Med. Unlocked*, vol. 18, p. 100288, 2020, doi:10.1016/j imu.2019.100288.
- [10] Y. Hamad, K. Simonov, and M. B. Naeem, "Breast cancer detection and classification using artificial neural networks," in *Proceedings - 2018 1st Annual International Conference on Information and Sciences, AiCIS 2018*, Feb. 2019, pp. 51–57, doi: 10.1109/AiCIS.2018.00022.
- [11] Anjna, M. Sood, and P. K. Singh, "Hybrid System for Detection and Classification of Plant Disease Using Qualitative Texture Features Analysis," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1056–1065, 2020, doi: 10.1016/j.procs.2020.03.404.
- [12] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Inf. Process. Agric.*, vol. 4, no. 1, pp. 41–49, 2017, doi: 10.1016/j.inpa.2016.10.005.
- [13] N. Senthilkumaran and R. Rajesh, "Image segmentation - A survey of soft computing approaches," *ARTCom 2009 - Int. Conf. Adv. Recent Technol. Commun. Comput.*, no. 1, pp.844846,2009,doi:10.1109/ARTCom.2009.219.
- [14] T. Zhao and H. Dai, "Tumor Region Location and Classification Based on Fuzzy Logic and Region Merging Image Segmentation Algorithm," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/1141619.
- [15] K. Huang, Y. Zhang, H. D. Cheng, and P. Xing, "MSF-GAN: Multi-Scale Fuzzy Generative Adversarial Network for Breast Ultrasound Image Segmentation," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2021, pp. 3193–3196, Nov. 2021, doi: 10.1109/EMBC46164.2021.9630108.
- [16] T. Zhao, J. Xu, R. Chen, and X. Ma, "Remote sensing image segmentation based on the fuzzy deep convolutional neural network," *Int. J. Remote Sens.*, vol. 42, no. 16, pp. 6267–6286, 2021, doi: 10.1080/01431161.2021.1938738.
- [17] M. K. Alsmadi, "A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 697–706, 2018, doi: 10.1016/j.asej.2016.03.016.
- [18] C. Narmatha, S. M. Eljack, A. A. R. M. Tuka, S. Manimurugan, and M. Mustafa, "A hybrid fuzzy brain-storm optimization algorithm for the classification of brain tumor MRI images," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2020, doi: 10.1007/s12652-020-02470-5.

- [19] R. Radha and R. Gopalakrishnan, "A medical analytical system using intelligent fuzzy level set brain image segmentation based on improved quantum particle swarm optimization," *Microprocess. Microsyst.*, vol. 79, no. September, p. 103283, 2020, doi: 10.1016/j.micpro.2020.103283.
- [20] Y. P. Huang, P. Singh, and H. C. Kuo, "A Hybrid Fuzzy Clustering Approach for the Recognition and Visualization of MRI Images of Parkinson's Disease," *IEEE Access*, vol. 8, pp. 25041–25051, 2020, doi: 10.1109/ACCESS.2020.2969806.
- [21] A. M. A. K. Parewe, W. F. Mahmudy, F. Ramdhani, and Y. P. Anggodo, "Dental disease detection using hybrid fuzzy logic and evolution strategies," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 1–8, pp. 27–33, 2018.
- [22] Y. Kurmi and S. Gangwar, "A leaf image localization based algorithm for different crops disease classification," *Inf. Process. Agric.*, no. xxxx, pp. 1–19, 2021, doi: 10.1016/j.inpa.2021.03.001.
- [23] H. R. Mohammed and Z. M. Hussain, "Hybrid mamdani fuzzy rules and convolutional neural networks for analysis and identification of animal images," *Computation*, vol. 9, no. 3, 2021, doi: 10.3390/computation9030035.
- [24] A. Salarikia, S. H. Miraei Ashtiani, M. R. Golzarian, and H. Mohammadinezhad, "Finite element analysis of the dynamic behavior of pear under impact loading," *Inf. Process. Agric.*, vol. 4, no. 1, pp. 64–77, 2017, doi: 10.1016/j.inpa.2016.12.003.
- [25] M. Hu, Y. Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy System Based Medical Image Processing for Brain Disease Prediction," *Front. Neurosci.*, vol. 15, no. July, pp. 1–13, 2021, doi: 10.3389/fnins.2021.714318.
- [26] T. Verma and S. Dubey, "Fuzzy-filtered neural network for rice disease diagnosis using image analysis," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 8 Special Issue 3, pp. 437–446, 2019.
- [27] J. Canadas, J. A. Sánchez-Molina, F. Rodríguez, and I. M. del Águila, "Improving automatic climate control with decision support techniques to minimize disease effects in greenhouse tomatoes," *Inf. Process. Agric.*, vol. 4, no. 1, pp. 50–63, 2017, doi: 10.1016/j.inpa.2016.12.002.

Detection of Motorcyclists without Helmet using Deep Learning

Parvinder Kaur
Department of CSE
BBSBEC
Fatehgarh Sahib, India
parvinderkaur1984@gmail.com

Dr. Baljit Singh Khehra
Principal
BAM Khalsa College
Garhshankar (Hoshiarpur), India
baljitzkhehra@rediffmail.com

Dr. Harmandeep Singh Gill
Principal
Bibi Sharan Kaur Khalsa
College, Chamkaur Sahib
profhdsgill@gmail.com

Abstract— Road traffic injuries are one of the highest public health hazards and in order to bring down the mishaps, one should be well aware of the road safety rules. Not wearing a suitable helmet while riding a motorized two-wheeler is something that is not just illegal but also something that puts one at risk of suffering from serious head injuries in an event of an accident. In order to enhance the enforcement of obeying traffic regulations such as wearing a suitable helmet for a 2-wheeler on the highway, an automated system that captures highway user's state and report defaulters to appropriate authorities is important. In this paper various feature extractors for this purpose are studied and evaluated that makes use of deep learning. Deep learning has been a game changer in the field of object detection in the last decade.

Keywords— convolution neural network, deep learning,

I. INTRODUCTION

Deep learning gave a gigantic increase to the already speedily developing arena of computer vision. With deep convolution neural networks, a vast variety of new applications of computer vision have been announced and are playing a significant role in our daily lives. These include object detection and recognition, face recognition, image content retrieval, self-driving cars, image search, image style transfer, image colorization, semantic segmentation, robotics, medical image analysis and disease detection, online shopping and many more. Although road safety rules are universal the world over, some of the traffic rules in India vary from state to state. The differences are mostly peculiar to signage, speed limits, and the punishments for breaking certain laws. It is therefore expedient to familiarize oneself with the traffic laws in one's society. Irrespective of one's intention toward driving, learning the rules of the road keeps highway users safe even in an unfamiliar terrain [1] [2] [3]. While there is sometimes an aura of relaxation and leisure about the driving experience, traffic laws in India are very strict and straying from these can cause serious trouble. Asides getting arrested by the police,

an accident is likely to happen which might lead to serious injury or loss of life.

Every state in India has mandated the use of a helmet while riding a two-wheeler. However, in December last year, the Gujarat high court made helmets temporarily optional. Not wearing a suitable helmet while riding a motorized two-wheeler is something that is not just illegal but also something that puts one at risk of suffering from serious head injuries in an event of an accident. Hence, one should wear an ISI-approved helmet of recommended specifications while riding a motorcycle or a scooter.

In order to enhance the enforcement of obeying traffic regulations such as wearing a suitable helmet for a two-wheeler on the highway, an automated system that captures highway user's state and report defaulters to appropriate authorities is important. A promising machine learning method to solve this problem is the Convolution Neural Network (CNN).

The remainder of this paper is organized into five sections. In Section II highlights the related work in this area. Section III represents the training and testing of various deep learning models and the effect of hyper tuning on performance. Finally, the whole paper is concluded in section IV along with significant findings. The aim of this research is to examine the possibility of using transfer learning by applying deep convolutional neural network pretrained models to classify the images of motorcyclists captured on the highway to either "With helmet" or "Without helmet".

The objectives of this paper are listed below:

- Obtain relevant highway images from Google images
- Annotate the images by grouping them into either "With helmet" or "Without helmet"
- Build a classification model that learns the patterns in each class
- Evaluate the model and then optimize it.

II. RELATED WORK

Tianyi Liu et al. [32] gave a detailed analysis of the process of the forward and back propagation of the

Business Intelligence with Intelligent Computing

convolutional neural network algorithm. They applied convolutional neural network algorithm to implement the typical face recognition problem in java. A parallel strategy was proposed. In addition, by measuring the actual time of forward and backward computing, they analyzed the maximal speed up and parallel efficiency theoretically.

Chen Wang [50] deduced the theory behind back-propagation neural network and implemented a back-propagation neural network from scratch in Java. He applied the neural network classifier to solve an image classification task using CIFAR-10 dataset. He also proposed a new approach for convolutional neural networks and made some experiments to test the functionality of dropout layer and rectified linear neuron.

Hijazi, S. et al. [19] presented the basics of convolutional neural networks including a description of the various layers in the architecture. Using traffic sign recognition as an example, they discussed the challenges of the general problem, and introduced algorithms and implementation software developed by Cadence that can trade off computational burden and energy for a modest degradation in sign recognition rates. They also outlined the challenges of using convolutional neural networks in embedded systems, and introduced the key characteristics of the Cadence® Tensilica® Vision P5 digital signal processor (DSP) for Imaging and Computer Vision and software that make it so suitable for convolutional neural network applications across many imaging and related recognition tasks.

Dang et al., 2016 [14] applied convolutional neural networks for the classification of vehicle pose and vehicle class on images of vehicles. He developed a convolutional neural network architecture and compared it to three state-of-the-art convolutional neural networks architectures. Considering the fact that the appearance of the vehicle class is dependent on the pose of the vehicle, he investigated two approaches. The first approach is the multi-task learning, which simultaneously learns the two tasks on the same model. For the second approach, he used an expert learner which uses the detected pose on the image to choose a specially trained model on this pose for the classification of the vehicle class. The error rate of the vehicle pose detection model was 1.12% while the model from the multi-task learning approach for the classification of the vehicle class produced an error rate of 26.04%. The ensemble learning improves the error to 24.87%. Yamashita et al., 2018 [53] offered a perspective on the basic concepts of convolutional neural network and its application to various radiological tasks, and discussed its challenges and future directions in the field of radiology. Two

challenges in applying convolutional neural network to radiological tasks; small dataset and overfitting, were addressed, as well as techniques to minimize them. Being familiar with the concepts and advantages, as well as limitations, of convolutional neural network is essential to leverage its potential in diagnostic radiology, with the goal of augmenting the performance of radiologists and improving patient care. Ramprasad et al., 2018 [38] used deep learning architecture, Convolutional Neural Network (CNN), for automatic classification of images. The proposed system uses the MNIST data set as a benchmark for classification of grayscale images. By training the images using convolutional neural network, they obtained a 98% accuracy on the validation set. Woolf, 2018 [52] demonstrated the viability of using deep convolutional neural networks (DCNNs) within the field of archaeology and cultural heritage preservation for the purpose of detecting looting. He utilized a pretrained deep convolutional neural network model (VGG-16) to show that convolutional neural networks can be highly efficient in expediting image recognition of looting pits. The new model obtained an overall accuracy of 96% with the Google Earth Pro (GEP) dataset and 90% accuracy with the Digital Globe Foundation (DGF) dataset. Mobilia, 2019 [36] proposed a convolutional neural network model for processing hyperspectral data for colon cancer detection. Using data taken from a limited number of colon tissue samples, he showed that the proposed CNN architecture can classify cancerous and non-cancerous tissue samples utilizing hyperspectral information. The obtained results were compared to the grayscale images of the same tissue samples of which the trained model had a 5.6% improvement in accuracy and a 0.037 improvement in F1 score over the individual band grayscale images, and a 21.7% improvement in accuracy and a 0.178 improvement in F1 score over the panchromatic grayscale images. The results were also compared to a k-Nearest Neighbor (KNN) classifier and a logistic regression classifier using the hyperspectral data, and the convolutional neural networks model outperformed both classifiers with a 17.9% improvement in accuracy and a 0.141 improvement in F1 score over the KNN classifier, and a 5% improvement in accuracy and a 0.061 improvement in F1 score over the logistic regression classifier. Some researchers [7], [8], [40], [41], [51], [12], [45] have proposed several methods to solve the problem of real-time helmet detection in traffic. Chiu et al., 2007 [7] proposed a system that detects motorcyclists in surveillance videos. This system segments the moving object and then tracks motorcycles and heads using a probability-based algorithm which handles the occlusion problem but it is unable to handle small variations due to noise and illumination effects. Also, it uses Canny edge detection

Business Intelligence with Intelligent Computing

with a search window of certain size in order to detect head. Chiverton et al., 2012 [8] used edge histogram-based features to detect motorcyclists. This method performs well even if there is low illumination in videos due to the use of edge histograms near the head but since the edge histograms used circular hough transforms to compare and classify helmets, it leads to a lot of misclassifications among motorcyclists with helmet as helmet-like objects were also classified as helmet. To overcome this misclassification problem, Silva et al., 2013, 2014 [40], [41] proposed a system which tracks the vehicles using Kalman filter [24]. This Kalman tracking system tracks objects even if they are lightly occluded but when there are two or more motorcyclists in the same frame, Kalman filter fails because it mostly works well for linear state transitions (that is, tracking single object at a time). In order to track multiple objects, non-linear functions are needed. Dahiya et al., 2016 [12] proposed a system which first uses Gaussian mixture model to detect moving objects. This model is robust to slight variations in the background. It uses two classifiers, one for the separating motorcyclist from moving objects, and another for separating without helmet from the upper one fourth part of the motorcyclists. However, it uses only hand engineered features such as SIFT [33], HOG [13], LBP [17] along with kernel SVM in both classifications. Their approach was promising as it accurately classifies motorcyclists and nonmotorcyclists but was not able to accurately classify between helmet and non-helmet riders under difficult conditions. Singh et al., 2016 [45] proposed a visual big data framework which scales the method in [12] to a city scale surveillance network. Experimental results show that the framework is able to detect a violator in less than 10 milliseconds. All these existing methods suffer from several challenges such as occlusion of objects, illumination effects, and poor localization on images with less pixels.

III. THE METHODOLOGY

Class Previous works performed on image recognition tasks such as the simple digit image recognition (LeCun, 1989) showed that the architecture of a network strongly influences the ability of the network to generalize. Good generalization on complex tasks can be achieved by designing a network architecture that has a certain amount of a priori knowledge about the task. The basic design principle is to minimize the number of free parameters in the network as much as possible without overly reducing the computational power of the network. This principle increases the probability of correct generalization because it results in a specialized network architecture that has reduced entropy.

Transfer learning is the reuse of a pretrained model on a new task. In transfer learning, a machine exploits

the knowledge gained from a previous task to improve generalization about another related task. Either the complete model is transferred or only part of the early layers. The architecture as well as the weights of the pretrained model are transferred.

There are two common approaches for developing models through transfer learning. In the first approach, the pretrained model is only used as a feature extractor. Some specific layers in the pretrained model are used for the generation of features. These features are then used as input to another machine learning algorithm to train the model on a new task. The weights for the entire network are frozen except that of the last fully connected layer. The last fully connected layer is replaced with a new one with random weights and only this layer is trained.

The second approach is known as fine-tuning. The weights of the new network are initialized with the weights of a pretrained model which are then used to train the network. In most cases, the pretrained models were initially trained on big datasets with millions of images which would have learnt useful features in the lower layers. These features could be useful for similar tasks and help to improve the performance of new models because the weights are well initialized to detect common features.

The source of the data used to train the classification model is Google images. Careful scraping of web images of motorcyclists was conducted and each relevant image was annotated or labelled as either "With Helmet" or "Without Helmet" in order to achieve the objectives of the research. Since with the use of transfer learning approach, a large amount of data is not necessarily required to achieve state-of-the-art performance, a total of 3000 images were used to train the model of which 1600 images were labelled as "With Helmet" and 1400 images were labelled as "Without Helmet".

The training set contains 2400 images while the validation set which was used to optimize the models contains 600 images. Figure 3.1 shows few samples of the images extracted from the dataset.



Fig. 3.1 Samples of images extracted from the dataset.

All the analyses in this study were done on Google Colaboratory, a free hosted Jupyter Notebook environment that runs on Google's cloud server with access to computing resources including GPU and TPU. With Google Colaboratory, one can write and execute Python code through the browser, and it is well suited for data analysis, machine learning, and deep learning. An open-source library known as "FastAi" was used for modelling. FastAi is a deep learning library which provides practitioners with high-level components that can quickly and easily provide state-of-the-art results in standard deep learning domains and provides researchers with low-level components that can be mixed and matched to build new approaches. It aims to do both things without substantial compromises in ease of use, flexibility, or performance. This is possible thanks to a carefully layered architecture, which expresses common underlying patterns of many deep learning and data processing techniques in terms of decoupled abstractions. These abstractions can be expressed concisely and clearly by leveraging the dynamism of the underlying Python language and the flexibility of the PyTorch library.

The FastAi library includes several pretrained models from torchvision namely:

- ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152
- SqueezeNet1_0, SqueezeNet1_1
- DenseNet-121, DenseNet-169, DenseNet-201, DenseNet-161
- VGG-16, VGG-19
- AlexNet

All these pretrained models were used to create convolution neural network learners which automatically handle whatever details passed to it such as the data and the evaluation metrics. Each learner downloads an ImageNet pretrained model, removes the classification head (that is, the last layer) of the model and replaces it with a layer with two output nodes, and set appropriate defaults for the optimizer, weight decay, learning rate e.t.c.

With the learner, the pretrained model is then fine-tuned. The results on the validation set are displayed in an HTML table showing the training loss, validation loss, and metrics after every epoch.

Table 4.1 Performance of the models without fine tuning

S/N	Model	No of Epochs	Training Loss	Validation Loss	Error-Rate	AUC_ROC Score	F1-Score
1	AlexNet	50	0.189940	0.450001	0.233333	0.908776	0.878990
2	SqueezeNet1_0	50	0.119692	0.599004	0.233333	0.878733	0.789994
3	SqueezeNet1_1	50	0.126517	0.446789	0.100000	0.909911	0.912900
4	VGG-16	50	0.079655	0.545678	0.200000	0.915453	0.828282
5	VGG-19	50	0.110300	0.400453	0.166667	0.950001	0.838710
6	ResNet-18	50	0.089143	0.4904562	0.133333	0.940891	0.870968
7	ResNet-34	50	0.096700	0.359210	0.100000	0.939993	0.862969
8	ResNet-50	50	0.095000	0.520120	0.150000	0.941123	0.842205
9	ResNet-101	40	0.051100	0.440004	0.133333	0.947720	0.883750
10	ResNet-152	40	0.059017	0.436745	0.073333	0.944383	0.928033
11	DenseNet-121	40	0.042390	0.5300321	0.133333	0.951057	0.876667
12	DenseNet-169	30	0.051000	0.6256431	0.133333	0.938821	0.872069
13	DenseNet-201	30	0.040023	0.3200056	0.100000	0.969897	0.993226
14	DenseNet-161	30	0.029967	0.2230012	0.033333	0.972353	0.943333

Table 4.2 Performance of the models after fine tuning

S/N	Model	No of Epochs	Training Loss	Validation Loss	Error-Rate	AUC_ROC Score	F1-Score
1	AlexNet	20	0.221956	0.396417	0.200000	0.903226	0.800000
2	SqueezeNet1_0	20	0.091015	0.483726	0.183333	0.895439	0.813559
3	SqueezeNet1_1	20	0.216620	0.065424	0.016667	0.995551	0.982456
4	VGG-16	20	0.093324	0.430497	0.166667	0.931034	0.848485
5	VGG-19	20	0.123758	0.426994	0.166667	0.943270	0.838710
6	ResNet-18	20	0.084704	0.352520	0.116667	0.948832	0.877193
7	ResNet-34	20	0.070485	0.336773	0.116667	0.952169	0.881356
8	ResNet-50	20	0.047834	0.464237	0.150000	0.939933	0.842105
9	ResNet-101	20	0.034877	0.442230	0.116667	0.947720	0.885246
10	ResNet-152	20	0.051333	0.506463	0.100000	0.945495	0.903226
11	DenseNet-121	20	0.036780	0.532718	0.133333	0.957731	0.866667
12	DenseNet-169	20	0.022426	0.648637	0.133333	0.937709	0.866667
13	DenseNet-201	20	0.034783	0.317461	0.100000	0.967742	0.913227
14	DenseNet-161	20	0.034585	0.281830	0.065567	0.978865	0.945666

The variation in the performance of these models can be attributed to properties such as the depth of the networks, the width of the networks, layer configurations, optimization algorithms etc. By comparison, the results gotten through the process of fine-tuning shown in table 4.2 are far better than the results of the models when used as feature extractors shown in table 4.1. Also, the training time of the fine-tuned models are far less.

		Predicted Label	
		With Helmet	Without Helmet
Actual Label	With Helmet	28	3
	Without Helmet	1	28

Fig. 4.1 Confusion Matrix of DenseNet-161 model as feature extractor

Business Intelligence with Intelligent Computing

Figure 4.1 shows the confusion matrix of the DenseNet-161 model as feature extractor. It shows that the model was able to correctly classify 56 out of 60 images in the validation set.

		Predicted Label	
		With Helmet	Without Helmet
Actual Label	With Helmet	31	0
	Without Helmet	1	28

Fig. 4.2 Confusion Matrix of fine-tuned SqueezeNet1_1 model

Figure 4.2 shows the confusion matrix of the fine-tuned SqueezeNet1_1 model. It shows that the model was able to correctly classify 59 out of 60 images in the validation set.

IV. CONCLUSION

This study utilized fourteen (14) pretrained deep convolutional neural network models using publicly available weights for training highway images of motorcyclists to classify them into “With Helmet” or “Without Helmet”. The two approaches for developing models through transfer learning were identified as feature extraction and fine-tuning. In feature extraction approach, the pretrained model is only used as a feature extractor, that is, some specific layers are used for the generation of features. These features are then used as input to another machine learning algorithm to train the model on a new task. In fine-tuning approach, the weights of the new network are initialized with the weights of a pretrained model which are then used to train the network allowing for slight adjustments to the more abstract representations of the model. Fine-tuning adapts some of the representations previously learned by an existing model to a new problem which helps the model to perform better. The applied methodology has shown that transfer learning with Convolutional Neural Networks can be highly efficient in expediting detection of helmet on the highway. The results appear to be at the state-of-the-art in image classification with the best performing model (fine-tuned SqueezeNet1_1). Also, the overall training time was relatively short as compared to training a convolutional neural network from scratch. Convolutional neural networks are proving to be the best type of deep learning algorithm when faced with image recognition and classification tasks yielding

higher and more accurate results than other previously used methods. They are particularly favourable when working with small datasets and are able to still yield decent results. However, one of the drawbacks with small datasets is overfitting, but this can be mitigated with data augmentation. Revolutionary recent advancements in deep learning have shown remarkable results in near-human level image classifications as well as in all perceptual problems. This research ultimately shows that intelligent systems can help capture highway user’s state as to whether a motorcyclist is with helmet or not.

REFERENCES

- [1] C. Vishnu, D. Singh, C. K. Mohan and S. Babu, "Detection of motorcyclists without helmet in videos using convolutional neural network," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 3036-3041, doi: 10.1109/IJCNN.2017.7966233.
- [2] Kaur, P., Taiwo, E., Detection of Motorcyclists without Helmet using Convolutional Neural Networks, International Journal of Advanced Trends in Computer Applications (IJATCA), 2021.
- [3] Bozinovski, S. (1972). Perceptrons: Training in Pattern Recognition. (original in Croatian: Perceptroni i obucavanje u prepoznavanju oblika) unpublished student scientific competition paper, University of Zagreb.
- [4] Bozinovski, S. (1985a). Adaptation and training: A viewpoint. Automatika 26 (3-4) 137- 144.
- [5] Bozinovski, S. (1985b). A representation theorem for linear pattern classifier training. IEEE transactions on systems, man, and cybernetics, (1), 159-161.
- [6] Bozinovski, S. (2020). Reminder of the first paper on Transfer Learning in Neural Networks, 1976. Informatica, 44(3).
- [7] Chiu, C. C., Ku, M. Y., & Chen, H. T. (2007, June). Motorcycle detection and tracking system with Occlusion Segmentation. In Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'07) (pp. 32-32). IEEE.
- [8] Chiyerton, J. (2012). Helmet presence classification with motorcycle detection and tracking. IET Intelligent Transport Systems, 6(3), 259-269.
- [9] Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column Deep Neural Networks for Image Classification. In 2012 IEEE Conference on Computer Vision and Pattern Recognition (pp. 3642-3649). IEEE.
- [10] Collobert, R., & Weston, J. (2008, July). A unified architecture for Natural Language Processing: Deep Neural Networks with Multitask learning. In Proceedings of the 25th International Conference on Machine Learning (pp. 160-167).
- [11] Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. Machine learning, 20(3), 273- 297.
- [12] Dahiya, K., Singh, D., & Mohan, C. K. (2016, July). Automatic detection of bike-riders without helmet using surveillance videos in real-time. In 2016 International Joint Conference on Neural Networks (IJCNN) (pp. 3046-3051).
- [13] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 886-893).

Business Intelligence with Intelligent Computing

- [14] Dang, H., & Luthardt, S. (2016). Using Convolutional Neural Networks to distinguish Vehicle Pose and Vehicle Class Erkennung von Fahrzeugpose und Fahrzeugklasse mit Convolutional Neural Networks.
- [15] Géron, A. (2017). Hands-on Machine Learning with Scikit-Learn and Tensorflow: Concepts. Tools, and Techniques to build intelligent systems.
- [16] Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks. URL <http://arxiv.org/abs/1312.6082>.
- [17] Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of Local Binary Pattern Operator for Texture Classification. IEEE transactions on image processing, 19(6), 1657-1663.
- [18] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [19] Hijazi, S., Kumar, R., & Rowen, C. (2015). Using Convolutional Neural Networks for Image Recognition. Cadence Design Systems Inc.: San Jose, CA, USA, 1-12.
- [20] Howard, J., & Gugger, S. (2020). FastAi: a layered API for Deep Learning. Information, 11(2), 108.
- [21] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7132-7141).
- [22] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected Convolutional Networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).
- [23] Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. The Journal of physiology, 195(1), 215-243.
- [24] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, vol. 82, no. 1, pp. 35-45.
- [25] Karpathy, A. (2011). Lessons learned from manually classifying CIFAR-10. <http://karpathy.github.io/2011/04/27/manually-classifying-cifar10/>.
- [26] Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.
- [27] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 25, 1097-1105.
- [28] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.
- [29] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to Handwritten Zip Code Recognition. Neural computation, 1(4), 541-551.
- [30] LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten Digit Recognition with a Back-Propagation Network. Advances in Neural Information Processing Systems, 2.
- [31] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
- [32] Liu, T., Fang, S., Zhao, Y., Wang, P., & Zhang, J. (2015). Implementation of training Convolutional Neural Networks. arXiv preprint arXiv:1506.01195.
- [33] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91-110.
- [34] McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In Psychology of learning and motivation (Vol. 24, pp. 109-165). Academic Press.
- [35] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), 115-133.
- [36] Mobilia, S., Sirkeci-Mergen, B., Deal, J., Rich, T. C., & Leavesley, S. J. (2019, June). Classification of Hyperspectral Colon Cancer Images using Convolutional Neural Networks. In 2019 IEEE Data Science Workshop (DSW) (pp. 232-236).
- [37] Naik, D., Mammone, R. (1993). Learning by learning in Neural Networks, In R.Mammone (ed.) Artificial Neural Networks for Speech and Vision, Chapman and Hall, London.
- [38] Ramprasad, M., Anand, M. V., & Hariharan, S. (2018). Image Classification using Convolutional Neural Networks. International Journal of Pure and Applied Mathematics, 119(17), 1307-1319.
- [39] Sharkey, N. E., & Sharkey, A. J. (1993). Adaptive generalisation. Artificial Intelligence Review, 7(5), 313-328.
- [40] Silva, R., Aires, K., Santos, T., Abdala, K., Veras, R., & Soares, A. (2013, October). Automatic detection of motorcyclists without helmet. In 2013 XXXIX Latin American Computing Conference (CLEI) (pp. 1-7).
- [41] Silva, R. R. V., Aires, K. R. T., & Veras, R. D. M. S. (2014, August). Helmet detection on motorcyclists using image descriptors and classifiers. In 2014 27th SIBGRAPI Conference on Graphics, Patterns and Images (pp. 141-148).
- [42] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.
- [43] Singh, D., Roy, D., & Mohan, C. K. (2016). DiP-SVM: Distribution Preserving Kernel Support Vector Machine for Big Data. IEEE Transactions on Big Data, 3(1), 79-90.
- [44] Singh, D., & Mohan, C. K. (2016, July). Distributed Quadratic Programming Solver for Kernel SVM using Genetic Algorithm. In 2016 IEEE Congress on Evolutionary Computation (CEC) (pp. 152-159).
- [45] Singh, D., Vishnu, C., & Mohan, C. K. (2016, December). Visual big data analytics for traffic monitoring in smart city. In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 886-891). IEEE.
- [46] Singh, Satya. (2019). Re: What is the pros and cons of Convolutional Neural Networks Retrieved from: <https://www.researchgate.net/post/What-is-the-pros-and-cons-of-Convolutional-neural-networks/5d1f6f37979fdc1be56e09a0/citation/download>.

- [47] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).
- [48] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for Convolutional Neural Networks. In International Conference on Machine Learning (pp. 6105-6114). PMLR.
- [49] Thrun, S., & Mitchell, T. M. (1995). Lifelong robot learning. *Robotics and autonomous systems*, 15(1-2), 25-46.
- [50] Wang, C., & Xi, Y. (2015). Convolutional Neural Network for Image Classification. Johns Hopkins University Baltimore, MD, 21218.
- [51] Waranusast, R., Bundon, N., Timtong, V., Tangnoi, C., & Pattanathaburt, P. (2013, November). Machine vision techniques for motorcycle safety helmet detection. In 2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013) (pp. 35-40). IEEE.
- [52] Woolf, T. (2018). Deep Convolutional Neural Networks for Remote Sensing Investigation of Looting of the Archeological Site of Al-Lisht, Egypt (Doctoral dissertation, University of Southern California).
- [53] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional Neural Networks: An Overview and Application in Radiology. *Insights into imaging*, 9(4), 611-629.
- [54] Zakhareuski, A. (2021, May). U.S. Rules of the Road. Website. <https://drivingtests.org/beginner-drivers/rules-of-the-road>.
- [55] Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In European Conference on Computer Vision (pp. 818-833). Springer, Cham.
- [56] Zhou, L., Li, Q., Huo, G., & Zhou, Y. (2017). Image Classification using Biomimetic Pattern Recognition with Convolutional Neural Networks features. *Computational intelligence and neuroscience*, 2017.

ANN for Heart Disease Diagnosis

Jagmohan Kaur

Research Scholar

IKG Punjab Technical University

Jalandhar, India

mohnideep1234@gmail.com

Baljit S. Khehra

Dept. of Computer Science & Engg.

BBSB Engg. College

Fatehgarh Sahib, India

baljitzkhehra74@gmail.com

Amarinder Singh

Department of Applied Sciences

BBSB Engg. College

Fatehgarh Sahib, India

amarinder77@gmail.com

Abstract—Diagnosis of heart disease is a critical and global challenge in healthcare and medical science. Nowadays, machine learning (ML) techniques are being widely used in the prediction of diseases due to their effective and significant results. There are many fields of ML techniques including ANN which is one of the extensively used supervised ML approach. An artificial neural network (ANN) is a computational non-linear model that is inspired from the brain i.e. neurons that can perform tasks like classification, prediction, visualization, decision-making and many others. It consists of large collection of artificial neurons or processing elements which operates in parallel. The aim of this paper is to offer a robust methodology of ANN using feed forward backpropagation method to predict heart disease. A dataset of 1000 patients from online Kaggle heart disease database is taken. Tansigmoid activation/transfer function is employed in the distinct hidden layers. The variation in hidden layers and neurons make possible the comparison analysis of accuracies during testing and training period. Mean Square Error (MSE) evaluate the system efficiently for the detection of heart disease. The highest accuracy of training 98.04% and testing 98.13% is achieved using four number of hidden layers. This technique can be effectively applied on the wide number of patients for early and accurate diagnosis than conventional ways.

Index Terms—machine learning, artificial neural network, mean square error, comparison analysis

I. INTRODUCTION

The heart is a vital organ of the human body whose essential role-play is to pump blood through the circulatory system. It pumps 2,000 gallons of blood and beats about one million times daily. Moreover, it provides oxygen-rich blood which is necessary for proper functioning of every body part. The reduction/blockage of blood flow through coronary arteries is caused by accumulation of fatty substance called plaque that further form blood clot that eventually results in complete blockage. Figure 1 presents the death rate of US citizens due to heart disease as released by AHA [1] in the year 2020. It clearly demonstrates the direct proportionality of the death rate with the age irrespective of the gender. High cholesterol, high blood pressure, smoking, diabetes and family history are the prime factors responsible for heart disease. Advanced computer technologies play immense role in the diagnosis and treatment of the diseases. In many fields of medicine, ANN [2] based approaches have been used due to its huge computation power in non-linear and complex real-life problems. Figure 2 depicts the model of an ANN consisting of an input layer with n input variables x_1, x_2, \dots, x_n , four hidden layers with neurons z^1, z^2, \dots, z^4 respectively while y_k and y_p are the outputs.

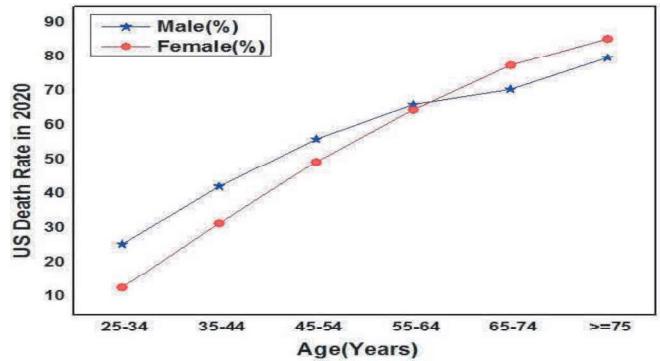


Fig. 1: Death Rate in US in 2020

1 Z^3, Z^3, \dots, Z^3 and Z^4, Z^4, \dots, Z^4 respectively while y_k and y_p are the outputs.

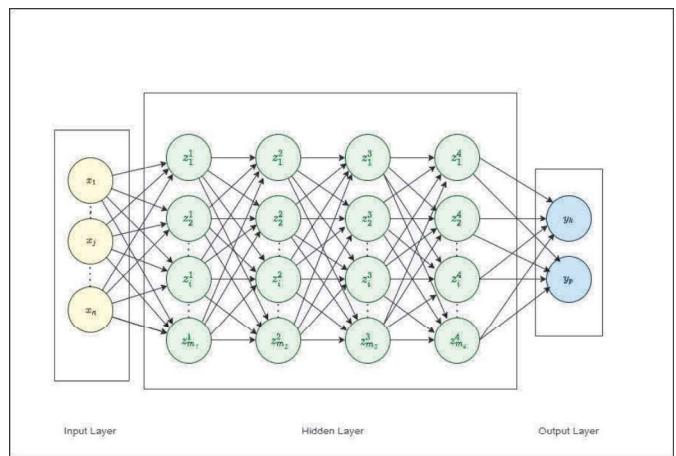


Fig. 2: The model of ANN

The aim of this paper is to employ ANN for developing diagnostic model of heart disease using genetic and non-genetic factors responsible for the disease. It is very helpful to solve many other mathematical models due to its simple, versatile and adaptable behavior. The paper is split into seven sections. The first section gives introduction while the literature survey and the proposed work are presented in the second and third sections respectively. The fourth section presents the ANN training technique and results are discussed in subsequent

section following which the conclusion and the future scope of the work are presented.

II. LITERATURE SURVEY

Hanan et al. [4] developed heart disease diagnostic model (2010) using Radial Basis Function (RBF NN) in an Intel Pentium® 4 system with MS Windows® XP. A data of a total of 325 patients, having ten input parameters, was procured from Aurangabad Sahara Hospital. It was preprocessed to machine understandable behavior i.e. binary numbers, 0 or 1. RBF NN was trained in MATLAB using data of 200 patients through different number of neurons and learning rate. The performance rate obtained from the testing data of 125 patients was 97% which was far better than other ML techniques. The medicines prescribed by the RBF NN were verified by the doctors/medical practitioners.

A prediction system evolved by Chen et al. [5] utilized ANN to improve the predication rate of heart disease diagnosis. The data set of 303 patients having thirteen clinical features was obtained from University of California (UCI) and was split into two sets: training and testing. The input of thirteen, six and two neurons was given to the respective layers of the ANN and accuracy was close to 80%. The other metrics like sensitivity and specificity were also provided along with receiver operating characteristic (ROC) curve in the study. The code was run in C and C# programming tools.

Sonawane et al. [6] employed Learning Vector Quantization NN using data base of thirteen input parameters from Cleveland Clinic Foundation (CCF). The training and testing sets consist of 164 and 139 patients respectively. The proposed NN was trained by changing number of neurons in the hidden layer that maximum accuracy of 85.55%, sensitivity of 84.15%, specificity of 89.58% was obtained for 200 epochs using eighteen neurons in the hidden layer in MATLAB R2012b software.

An impressive comparative and graphical study of training technique of NN using back propagation and extreme learning (EL) algorithm was revealed by Ismaeel et al. [7]. The study made emphasis on EL technique to train the system as it's accuracy rate was observed to be 86.5% which was more than the same for back propagation. A subset of thirteen attributes was extracted from 76 features of CCF dataset of ML repository. Weights and bias of input nodes were initialized before processing into the hidden layer and sigmoid function was used that revealed error percentage at each step.

Arabasadi et al. [8] represented computer-aided model to diagnose heart disease using integration of ANN with Genetic Algorithm. The study was validated through four datasets of various institutes extracted from UCI, ML repository having 303, 294, 123 and 200 samples and achieved the maximum accuracy of 93.85%, sensitivity of 97% and specificity of 92%. The performance was also depicted by ROC curve and RMSE which was almost near to zero.

An artificial model designed by Sivarajani et al. [9] was very effective to generate the highest accuracy in the prediction of heart disease. A record of 70,000 patients with thirteen

clinical parameters from cardiovascular disease dataset was obtained to validate the proposal using back propagation technique of ANN. The article presented analysis of distinct number of hidden layers with activation functions and their accuracies. Rectified linear (Relu) (1), tansig and log-sigmoid (logsig) (2) transfer functions [3] were used for training the NN. Default Python code was run to determine the model accuracy rate. A friendly graphical interface was presented to make it more comfortable for the end users.

$$\varphi(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x \leq 0 \\ 1 & \end{cases} \quad (1)$$

$$\varphi(x) = \frac{1}{1 + \exp(-x)}, x \in R \quad (2)$$

III. THE PROPOSED WORK

A dataset of 1000 heart patients is extracted from the Kaggle heart disease database [10] and is split in the proportion of 7:3 into two sets, larger for the training and other for the testing. The 'nntool' NN toolbox of Matlab R2019b is used for training the network by back propagation method using Levenberg-Marquardt technique. The Hyperbolic tangent sigmoid activation function popularly known as *tanh* (3), is used in hidden layers varying from two to five, using different combinations of neurons. The block diagram of the proposed work is represented in Figure 3.

$$\varphi(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}, x \in R \quad (3)$$

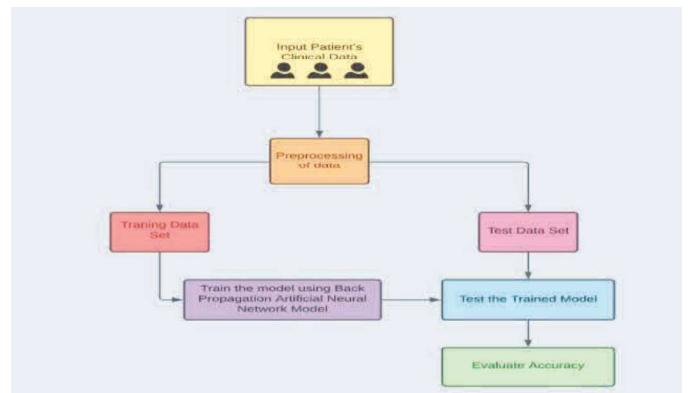


Fig. 3: The Block Diagram of the Proposed Study

IV. TRAINING OF ANN VIA *tansigmoid* ACTIVATION FUNCTION

Tables I - IV represents the performance of three-layered to six-layered NN. The experimental work of the distinct combinations of neurons varying from two to five hidden layers is recorded and the epochs, elapsed time, accuracy of training and testing, validation along with average is represented in these tables. The accuracies using six different combinations of neurons for two, three, four and five hidden layers are presented in figures 4 - 7.

TABLE I: Performance Evaluation Metrics in 3-Layer Architecture of ANN

Case No.	No. of neurons	No. of epochs	Time	Gradient	Best Validation Performance(epoch)	Accuracy(%)		Validation(%)	Average
						Training	Testing		
1	(10,9)	22	00:00:49	0.0000000328	0.062032(22)	99.20	92.76	87.57	93.18
2	(11,10)	20	00:01:14	0.000412	0.043496(14)	99.16	94.38	91.49	95.01
3	(9,8)	2	00:01:10	0.0182	0.053354(23)	97.32	86.13	88.92	90.79
4	(10,8)	21	00:00:37	0.0424	0.037464(15)	97.17	86.42	92.40	91.99
5	(11,9)	28	00:01:00	0.016	0.050256(22)	97.10	91.5	89.50	92.70
6	(9,7)	30	00:01:02	0.002	0.0324(24)	99.25	96.15	93.45	96.28
					Average	98.20	91.23	90.56	93.33

TABLE II: Performance Evaluation Metrics in 4-Layer Architecture of ANN

Case No.	No. of neurons	No. of epochs	Time	Gradient	Best Validation Performance(epoch)	Accuracy(%)		Validation(%)	Average
						Training	Testing		
1	(10,9,8)	29	00:02:10	0.112	0.027765(23)	98.39	94.41	94.57	95.80
2	(11,10,9)	20	00:00:43	0.000419	0.05353(14)	98.15	90.24	88.91	92.43
3	(9,8,7)	20	00:00:41	0.0163	0.06305(14)	95.12	88.60	86.93	90.22
4	(10,8,7)	30	00:01:01	0.0201	0.02124(24)	98.57	93.33	95.70	95.87
5	(11,9,8)	20	00:00:44	0.0239	0.027298(14)	94.24	85.32	94.42	91.33
6	(10,9,7)	36	00:01:20	0.000133	0.03744(30)	98.40	96.36	92.54	95.77
					Average	97.15	91.38	92.18	93.57

TABLE III: Performance Evaluation Metrics in 4-Layer Architecture of ANN

Case No.	No. of neurons	No. of epoches	Time	Gradient	Best Validation Performance(epoch)	Accuracy(%)		Validation(%)	Average
						Training	Testing		
1	(11,10,9,8)	17	00:00:08	0.0371	0.078378(11)	95.45	80.47	83.38	86.43
2	(10,9,8,7)	19	00:00:40	0.018	0.050163(13)	94.908	88.07	89.86	90.95
3	(9,8,7,6)	24	00:00:01	0.0000387	0.035785(18)	98.83	92.42	92.83	94.69
4	(10,8,6,5)	26	00:00:10	0.00006	0.017564(20)	98.65	93.31	96.45	96.14
5	(10,7,6,5)	31	00:00:01	0.00347	0.027053(25)	94.813	88.68	94.47	92.66
6	(10,8,7,6)	33	00:02:45	0.00706	0.045066(27)	98.04	98.13	90.09	95.42
					Average	96.78	90.18	91.18	92.71

TABLE IV: Performance Evaluation Metrics in 5-Layer Architecture of ANN

Case No.	No. of neurons	No. of epochs	Time	Gradient	Best Validation Performance(epoch)	Accuracy(%)		Validation(%)	Average
						Training	Testing		
1	(10,9,8,6,5)	25	00:00:10	0.00302	0.046474(19)	97.967	90.81	91.08	93.29
2	(9,8,6,5,4)	26	00:00:02	0.0144	0.071021(20)	97.667	96.11	85.81	93.20
3	(9,8,7,6,6)	19	00:00:01	0.0133	0.07952(13)	94.802	82.08	83.00	86.63
4	(9,7,6,5,4)	18	00:00:01	0.0025	0.079286(12)	96.222	84.14	83.55	87.97
5	(9,8,7,5,4)	16	00:00:01	0.041	0.11354(10)	90.80	90.53	75.37	85.57
6	(9,8,6,5,3)	30	00:00:027	0.0000957	0.035598(24)	96.81	94.55	92.82	94.73
					Average	95.71	89.70	85.27	90.23

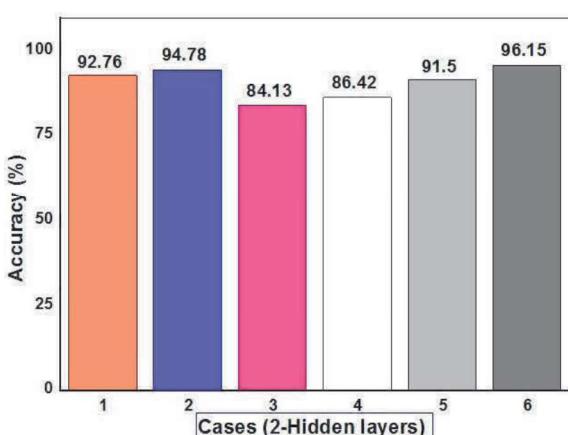


Fig. 4: Two hidden layers vs accuracies

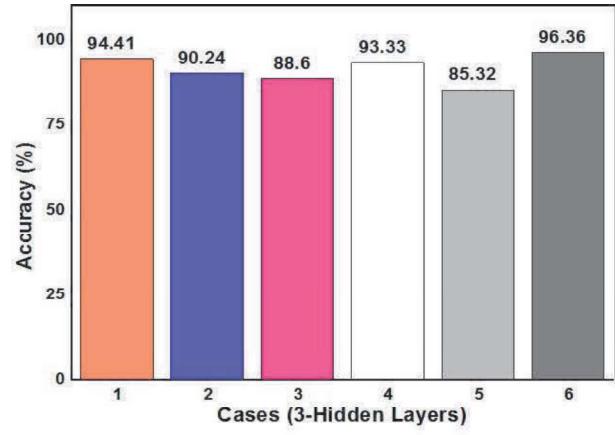


Fig. 5: Three hidden layers vs accuracies

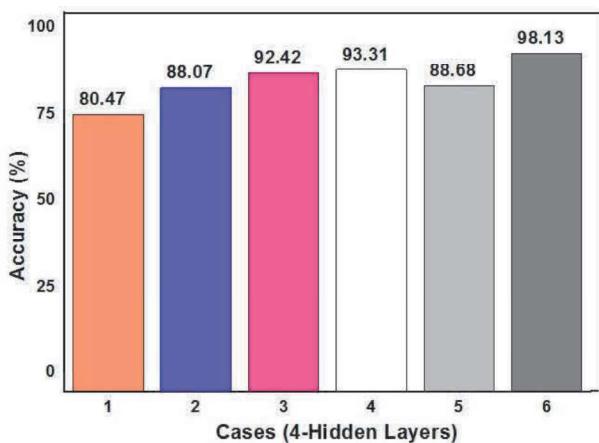


Fig. 6: Four hidden layers vs accuracies

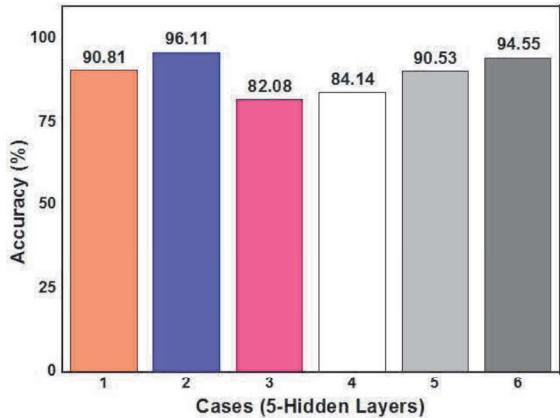


Fig. 7: Five hidden layers vs accuracies

V. RESULTS AND DISCUSSIONS

The above computations reveal that the best performance is achieved in 4-layered Neural Network architecture. The combination of ten, eight, seven and six neurons in respective four hidden layers yielded training and testing accuracy of 98.04% and 98.13%, respectively using *tansigmoid* activation function. The time elapsed during the task was just 2.45 mins. The Mean Squared Error (MSE) was computed to be 0.045121 at 17 epochs out of 23. A comparative analysis of hidden layers with accuracies is demonstrated in fig. 8 and MSE of corresponding layer and neurons is also shown in fig. 9.

VI. CONCLUSION AND FUTURE SCOPE

The research marks emphasis on various combinations of neurons in distinct layers to ascertain the accurate, early and low-cost diagnosis of heart disease using ANN. A dataset having thirteen clinical parameters of heart disease was trained using backpropagation algorithm and further by Marquardt-Levenberg method. In the paper, *tansigmoid* activation function is used and as a future study, the present methodology

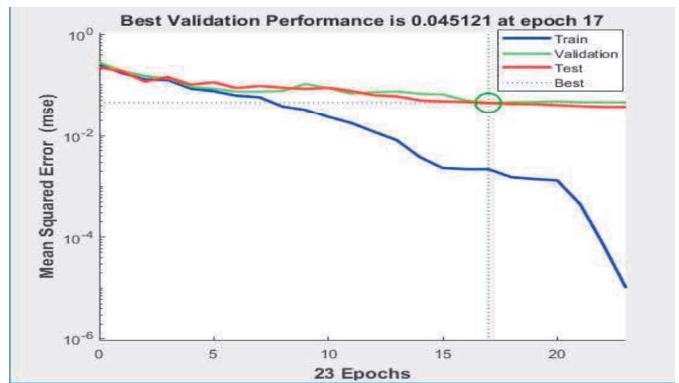


Fig. 8: Mean Squared Error

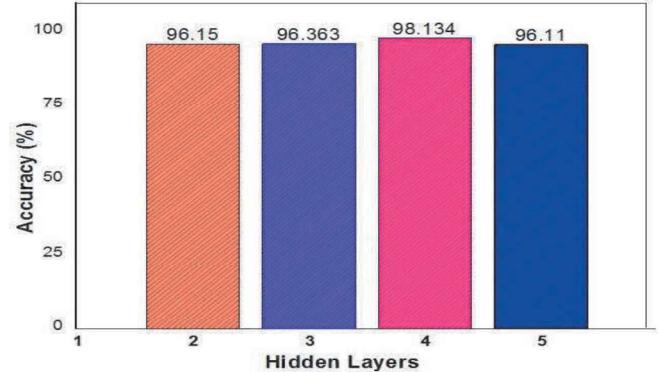


Fig. 9: Accuracy vs Neurons

can be further enhanced by using another activation function in more than four hidden layers to validate the output.

REFERENCES

- [1] American Heart Association, "Heart Disease and Stroke Statistics 2020 Update: A Report," *Circulation: An AHA Journal*, Vol. 141, No. 9, e139-e596, 2020.
- [2] M. Mishra and M. Srivastava, "A view of Artificial Neural Network," *International Conference on Advances in Engineering & Technology Research (ICAETR - 2014)*, pp. 1-3, 2014.
- [3] J. Lederer, "Activation Functions in Artificial Neural Networks: A Systematic Overview," *arXiv preprint arXiv:2101.09957*, 2021.
- [4] S. A. Hannan, A. V. Mane, R. R. Manza and R. J. Ramteke, "Prediction of heart disease medical prescription using radial basis function," *Proc. of 2nd IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India*, pp. 1-6, 2010.
- [5] A. Chen, S. Huang, P. Hong, C. Cheng and E. Lin, "HDPS: Heart Disease Prediction System," *Computing in Cardiology, Hangzhou, China*, vol. 8, pp. 557-560, 2011.
- [6] J. S. Sonawane and D. R. Patil, "Prediction of heart disease using learning vector quantization algorithm," *Proc. of Conference on IT in Business, Industry and Government, Indore, India*, pp. 1-5, 2014.
- [7] S. Ismaeel, A. Miri and D. Chourishi, "Using the Extreme Learning Machine (ELM) Technique for Heart Disease Diagnosis," *Proc. of International Conference on Humanitarian Technology, Ottawa, Canada*, pp. 1-3, 2015.
- [8] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei and A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," *Computer methods and programs in biomedicine*, vol. 141, pp. 19-26, 2017.
- [9] R. Sivarajanji and N. Yuvaraj, "Artificial Intelligence model for earlier prediction of cardiac functionalities using multilayer perceptron," *Journal of Physics: Conference Series*, vol. 1362, pp. 012062, 2019.
- [10] University of California, "Heart Disease Data Set," *Machine Learning Repository*, <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>, 2018.

Social Media Analytics in the Perspective of Big Data

Kanwar Kulwant Singh

Associate Professor

B.A.M. Khalsa College, Garhshankar

head_ce@bamkc.edu.in

Abstract

Since the arrival of ‘Big Data’ and social media’s meteoric rise in popularity, businesses have been forced to review, reinvent, and reallocate their marketing strategies. Social Media Analytics is a method of big data. Social media analytics needs collecting statistics from media platforms and evaluating the data gotten from the statistics. These data could be collected using the advocated social media analytics instruments to inform marketing assessments. Social media analytics is frequently used as a medium for accessing the feelings of your clients with regard to a particular product or service. Social media analytics is a research axis concentrated on extracting useful insights from social media data, with the motive of helping individuals and organizations take the most optimum decisions concerning several disciplines of life (business, marketing, politics, health, etc.). This paper throws light on the underlying phases of the social media analytics process. Also emphasis on how social media analytics framework use as a great tool in market research by E-Commerce Companies.

Keywords: *Big data, Marketing, Social media, Capture, Understand, Present, E-Commerce.*

I. INTRODUCTION

Big Data constructed on social media consists of text messages, songs, pictures, videos etc. People split information by way of text messages, videos, pictures, songs etc. They frequently squeeze out their intention to buy a product, appeal for feedback, share their service exposure or product reviews on social media. Social media has been a defining element of life in the 21st century, monetizing peer-to-peer splitting of information. Social media analytics is an emanating concept which is becoming a part of mainstream marketing strategy. It is based on social media data designed on sites such as Twitter, Facebook or WhatsApp. It is related with developing and evaluating informatics tools and frameworks to gather, monitor, analyze, summarize and visualize social media data (Zeng et al., 2010). [4]

Social media analytics is related to developing and evaluating informatics equipment and frameworks to gather, monitor, analyze, summarize,

and visualize social media data, to accelerate conversations and interactions, to root out convenient patterns and intelligence. Accor illustrates how social media analytics can assist businesses upgrade their reputations and resulting business execution. Ubiquitous smartphones and other mobile devices, Facebook and YouTube channels faithful in companies and products, and hashtags that build it freely and easily share experiences instantly combine to generate a social media landscape that is swiftly becoming part of the fabric of everyday business operations. As the numerous users on social media sites continues to grow, so does the necessity for businesses to monitor and use them to their benefit.

II. REVIEW OF LITERATURE

Piccialli and Jung (2017) examined how social media played a prominent role with consumer behavior, and many businesses capitalized on social media for financial gains (p. 605). Social media enhances communications across the various platforms. This study demonstrated that content analysis can help organizations distribute information through their marketing efforts. The authors noted that organizations have exploited social media marketing for advertising campaigns (p. 605). Piccialli and Jung asserted an innovation of ideas and data distribution were enhanced by social media (p. 611) [1]

Fischbach and Zarzosa (2018) explained social media analytics as “the practice of gathering data from blogs and social media websites and analyzing that data to make better business decisions” (p. 143). The authors advised that “consumers consume, create, and share content among each other and marketers should monitor and shape the conversation” (p. 145). New technologies allow the businesses and consumers to build marketing strategies together (p. 146). Atwong (2015) asserted social media marketing and analysis are key to understanding digital marketing’s process of “targeting audience, defining and executing strategy, managing contents, tracking metrics, and reporting analytics” (p. 27). [2]

Appel et al. (2019) anticipated social media will have a greater importance with marketing efforts in the future (p. 79). The authors projected the social media users will grow gradually, and the global marketers have acknowledged social media as a powerful marketing outlet (p. 79). The authors admitted the everchanging social media will lead to innovative technology to offer new services to the consumer (p. 79). They recognized the social media's influence on marketing, communications for businesses, organizations, institutions, and individuals, and politics (p. 79). The authors expressed social media was culturally significant and a major type of media for marketing and advertising (p. 91) [3].

III. OBJECTIVES OF THE STUDY

The present study has carried out with the following objectives:

1. To Examine Social Media Platforms as a tool in market research.
2. To Study the Process of Social Media Analytics.
3. To Study the Impact of Social Media Analytics on E-Commerce.

IV. RESEARCH METHODOLOGY

This study is based on secondary source of data. It covers the secondary sources of information collected through reference from books, Journals, Published Articles and Websites.

V. DISCUSSION

A. Social media platforms' role in market research

Social media networks are a great tool in market research. They are providing a direct, fast, and easier way to reach an audience through a person who is widely known. For example, an athlete who gets endorsed by a sporting goods company also brings their support base of millions of people who are interested in what they do or how they play and now they want to be a part of this athlete through their endorsements with that particular company. At one point consumers would visit stores to view their products with famous athletes, but now you can view a famous athlete's, such as Cristiano Ronaldo, latest apparel online without any delay. He endorses them to you directly through his Twitter, Instagram, and Facebook accounts.

Social media has demonstrate itself to be a great way of generating brand awareness, customer loyalty and satisfaction, and enhancing website traffic as means of reaching new consumers. With an increasing number of shoppers making their purchases online, social media has also go well with an essential channel for E-Commerce. According to Statista, 30% of users report having purchased something through social media and that's not all. 70% of consumers recite the reviews of products they plan to buy on social media platforms like Pinterest, Facebook, and Instagram.

In the initial days of social media the mid-2000s PR agencies would monitor customers' posts on a business's own website to try to recognize and manage unhappy customers. With the explosion in the number of social media sites and magnitude of users on them, monitoring alone is not adequate to render a complete picture of how a company is undertaking. Consider the pervasiveness of social media:

1. Increase in Number of Social Media Users



Figure 1 Social Media Users Over Time

Source: datareportal.com

Figure 1 throw light that there were 467.0 million social media users in India in January 2022. The number of social media users in India at the start

of 2022 was counterpart to 33.4 % of the total population. Kepios analysis discloses that social media users in India increased by 19 million (+4.2 percent) between 2021 and 2022. 95%

of adults online are consistent users of social media. Social media users on most-trafficked sites are:

- Facebook had 329.7 million users in India in early 2022.
- YouTube had 467 million users in India in early 2022.
- Facebook Messenger had 122.5 million users in India in early 2022.
- Instagram had 83 million users in India in early 2022.
- Snap chat had 126 million users in India in early 2022.



Figure 2 Time Spent Using Social Media Apps

Source: datareportal.com

3. Online shopping activities by social media users

Marketers have been utilizing social platforms to connect with ecommerce consumers for a while now. And for a good reason more and more internet users shop online and use social media to research brands. Take a look at these contemporary stats as shown in Figure 3:

- Twitter had 329.7 million users in India in early 2022.

2. Time spent using social media networking

Social media data has innumerable data types originating from a variety of data sources such as Facebook, Twitter, WhatsApp, Instagram, YouTube etc. Nonetheless, even these statistics in figure 1 do not fully account for the impact social media has on our lives. Users expand more than 30% of their time online on social media sites as shown in figure 2.



- 62.2% of Indian internet users purchased a product online in 2022.
- 29.4% of Indian internet users use social media to search for brand-related information.
- 17.3% of Indian internet users bought a second-hand item via an online store.
- 34.1% of Indian internet users ordered groceries via an online store.



Figure 3 Weekly Online Shopping Activities

Source: datareportal.com

B. Social media analytics process

This section outlines the SMA methodological skeleton which is adapted from the CUP SMA

framework, a three-stage procedure that was initiated by Fan and Gordon. The abbreviation of CUP contains the processes of capture, understanding and present. Capture is the procedure of obtaining applicable social media data by monitoring numerous

social media sources, archiving relevant data and extracting appropriate information (Fan & Gordon). [5] Acknowledge is the process of assessing the meaning from accumulated social media data and generating metrics useful for decision making. This is the heart of the entire SMA process. Assessing meaning may contain statistical methods, text and

data mining, natural language processing, machine translation and network analysis. The contemporary stage is when the results of various analytics are summarized, evaluated and shown to users in an easy-to-understand format comprising visualization techniques.

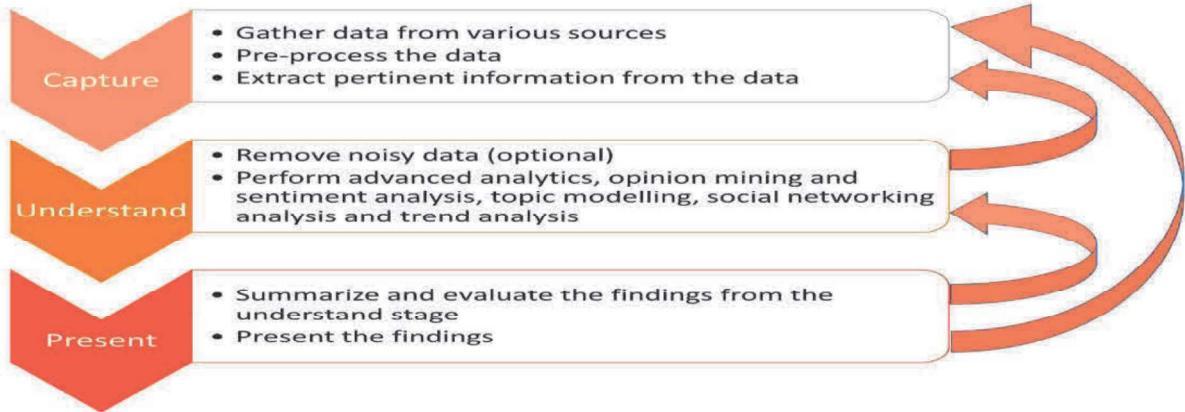


Figure:4 SMA Framework

Capture: This stage supports identify conversations on social media platforms related to its activities and interests. This is done by gathering enormous amounts of relevant data crosswise hundreds or thousands of social media cradles using news feeds and APIs or through crawling. The apprehend part covers popular platforms (such as Facebook, Foursquare, Google+, LinkedIn, Pinterest, Twitter, Tumblr, and YouTube), over and above smaller, more specialized sources (such as Internet forums, blogs, microblogs, wikis, news sites, picture-sharing sites, podcasts, and social-bookmarking sites). An extensive amount of data is archived and available to meet businesses' requirements. To get ready a dataset for the understanding stage, different preprocessing steps may be performed, involving data modeling, data and record linking from various sources, stemming, part-of-speech tagging, feature extraction, and other syntactic and semantic operations that support analysis. Information about businesses, users, and events, as well as user comments and feedback and other information, is also pulled for later analytical modeling and analysis.

Being tuned in to changing customer tastes and behavior, businesses can anticipate significant changes in demand and adjust accordingly by ramping production up or down.

The capture stage must balance the need to find information from all quarters (inclusivity) with a focus on sources that are most relevant and

Source: <https://www.google.com>
authoritative (exclusivity) to assist in more refined understanding.

Understand: When a business gathers the conversations related to its products and operations, it must then assess their meaning and create metrics useful for decision making them understand phase. Since the capture stage collects data from many users and sources, a considerable part may be noisy and thus have to be withdrawn prior to meaningful analysis. Simple, rule-based text classifiers or more sophisticated classifiers trained on labeled data may be utilized for this cleaning function. Evaluating meaning from the cleaned data can involve statistical methods and other methods derived from text and data mining, natural language processing, machine translation, and network analysis. The understand stage provides information about user sentiment how customers feel about a business and its products and their behavior, including the likelihood of, say, purchasing in response to an ad campaign. Many useful metrics and trends about users can be produced in this stage, enveloping their backgrounds, interests, concerns, and networks of relationships.

Present: In this last stage, the results from various analytics are outlined, estimated and shown to users in an easy-to-grasp format. Visualization tools may be used to describe valuable information; one commonly used interface structure is the visual dashboard, which accumulates and exhibits information from number of sources. Polished visual analytics go beyond the simple layout of information. By assisting customized ideas for various users, they

help make sense of huge quantity of information, including patterns that are extra clear to people than to machines. Data analysts and statisticians may add more help.

C. Impact of social media analytics on e-commerce

- Better position management and brand awareness (30%)
- Perception of new business opportunities (20%)
- Business process improvement (25%)
- Improved marketing strategies (75%)
- Better customer engagement (65%)
- Better customer service (35%)
- Product innovation (30%)

VI. CHALLENGES

Big information can still expedite the interference of social media enterprises into People's Privacy. As Facebook, Twitter, Instagram and Pinterest still legitimate their offerings, it might seem that the benefits that large amount of information can have for social media within the future can become even extra customized. A study printed by researchers from Cambridge and Stanford Universities describes that Facebook will use its information to foresee people's character with additional exactness than shut relatives and families. Every follow, like, share and comment is information that tells social media enterprises what you need or dislike, what your moves are going to be, that affect or finalize you'll support and what you're assured to shop for. Not to mention, any action you're taking on browsers and search engines these days can probably link back to your social media profile, giving behind a condensed way of digital footprint which will be used for disclose your next action. This case can be solely sharpened as individuals become extra dependent on social media platforms for collecting and sharing data.

VII. CONCLUSION

According to Measuring Success, over 80% of consumers still use the web to research products and services. Thereby social media marketing is also used by businesses in order to build relationships of trust with consumers. The revolutionary development of the Internet, accompanying the affirmation of social media manifesto has boosted the interconnectivity between people. At the present time, consumers

spend most of their time by sharing ratings, recommendations and past purchasing experiences through online channels and associations. These advancements have introduced the concept of social commerce, which comprise of empowering consumers to generate contents and persuade other users to buy from a specific e-shop. Furthermore, it has been authenticate that social media amplifies the level of shopper trust and at the same time stimulates them to buy through social networking sites. Correspondingly, all these interactions establish trust inside the network used, thus affecting the intention to purchase. All the statistics reported in this paper confirm the mounting growth in the popularity of various social media channels for marketing purpose not only among renowned companies, but also for small family businesses, like restaurants. In point of fact, the latter is investing money on this form of internet marketing with an objective of improving brand awareness, getting new purchaser, increasing loyalty and boosting revenues.

REFERENCES

- [1] Piccialli, F., & Jung, J. E. (2017). Understanding Customer Experience Diffusion on Social Networking Services by Big Data Analytics. *Mobile Networks and Applications*, 22(4), 605-612. <http://dx.doi.org/10.1007/s11036-016-0803-8>.
- [2] Fischbach, S., & Zarzosa, J. (2018). Big data on a smaller scale: A social media analytics assignment. *Journal of Education for Business*, 93(3), 142–148. <https://doi.org/10.1080/08832323.2018.1433123>
- [3] Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2019). The future of social media in marketing. *Journal of the Academy of Marketing Science*, 48(1), 79–95. <https://doi.org/10.1007/s11747-019-00695-1>
- [4] D. Y. Lin,D. Zeng, (2010) Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data, <https://doi.org/10.1002/gepi.20435>
- [5] Weiguo Fan , Michael D. Gordon (2014), The power of social media analytics, Volume 57,Issue 6, pp74-81, <https://doi.org/10.1145/2602574>
- [6] Stefan Stieglitz, MiladMirbabaie, Björn Ross, Christoph Neuberger, Social media analytics – Challenges in topic discovery, data collection, and data preparation, *International Journal of Information Management*, Volume 39, 2018, Pages 156-168, ISSN 0268-4012
- [7] Umar Ruhi Social Media Analytics as a Business Intelligence Practices Current Landscape & Future Prospects *Journal of Internet Social Networking & Virtual Communities* <http://www.ibimapublishing.com/journals/JISNVC/jisnvc.html> Vol. 2012 (2012), Article ID 920553, 12 pages DOI: 10.5171/2012.920553
- [8] Sheela Singh, PriyankaArya, Alpna Patel and Arvind Kumar Tiwari (2019) Social Media Analysis through Big Data Analytics: A Survey 2nd International Conference on Advanced Computing and Software Engineering (ICACSE-2019)
- [9] Adnan VeyselErtemel, (2019) Consumer Insight as Competitive Advantage using Big Data and Analytics *International Journal of Commerce and Finance*, Vol. 1, Issue 1, 2015, 45-51.

- [10] Andreas M.Kaplan Michael Haenlein (2010) "The challenges and opportunities of Social Media, Volume 53, Issue 1, Pages 59-68, <https://doi.org/10.1016/j.bushor.2009.09.003>.
- [11] Sinha, V., Subramanian, K. S., Bhattacharya, S., & Chaudhuri, K. (2012). The Contemporary Framework on Social Media Analytics as an Emerging Tool for Behavior Informatics, HR Analytics and Business Process Management, 17(2), 65–84.
- [12] Choi, T.M., Chan, H. K., & Yue, X. (2017). Recent Development in Big Data Analytics for Business Operations and Risk Management. IEEE Transactions on Cybernetics, 47(1), 81–92. <https://doi.org/10.1109/TCYB.2015.2507599>
- [13] https://hrnak.srce.hr/index.php?show=clanak&id_clanak_jezik=138312 Sivarajah, U., Irani, Z., Gupta, S., & Mahroof, K. (2020). Role of big data and social media analytics for business to business sustainability: A participatory web context. Industrial Marketing Management, 86, 163–179. <https://doi.org/10.1016/j.indmarman.2019.04.005>.
- [14] Valerio, Curzi; William, Lecoq; Noémier, Quéré (2019) : The Impact of Social Media on E-Commerce Decision Making Process, International Journal of Technology for Business (IJTB), ISSN 2644-5085, Springwish Publisher, Bratislava, Vol. 1, Iss. 1, pp. 1-9, <http://dx.doi.org/10.5281/zenodo.2591569>.
- [15] <https://cacm.acm.org/magazines/2014/6/175163-the-power-of-social-media-analytics/fulltext>
- [16] <https://www.frontiersin.org/articles/10.3389/fdata.2021.623794/full>
- [17] <https://www.websitemagazine.com/blog/big-data-analytics-shaping-the-e-commerce-industry>
- [18] <https://datareportal.com/reports/digital-2022-india>

Artificial Intelligence Reshaping the Automotive Industry

Janki Aggarwal

Associate Professor

Department of Commerce and Economics

B.A.M. Khalsa College, Garhshankar

Janki.aggarwal@bamkc.edu.in

Abstract

Artificial intelligence (AI) is intelligence exhibited by machines, with machines mimicking functions typically associated with human cognition. With the advent to industry 4.0 or the fourth industrial revolution characterized by cyber-physical systems, there is a focus on the innovative application of advanced robotics and artificial intelligence to bring about digital transformation in business. There will a great change in the work places in the manufacturing units as we see it today. We need to prepare us for the future period in which we have to face the technological changes brought about by AI. This research service examines the role artificial intelligence (AI) will play in the transformation of the automotive space. In this study, I have attempted to provide an overview of the changes in the automotive industry precipitated by the use of artificial intelligence. Technology implementation has increased, and the post-pandemic situation appears to be positive for all stakeholders; however, automakers have yet to fully harness AI's potential in their service offerings. The variety of possible applications of machine learning in the automotive industry is impressive. But despite its promising potential, the use of AI in the automotive industry is associated with several challenges like algorithm biases, data quality etc.

Key words: Artificial intelligence, Automotive industry, Machine Learning, Sustainability.

I. INTRODUCTION

AI is finding applications throughout the whole automotive value chain and, to a certain extent, is already reshaping the way that people and goods move across the world, from scanning traffic patterns to reducing road casualties and optimising routes to reduce emissions. Today, artificial intelligence is commonplace. Navigation systems in cars, fitness apps, Alexa and Siri, Amazon, Netflix, weather forecasting, and high-speed stock trading are among current must-have AI applications. This decade has witnessed paradigm change in automotive industry like changing customer needs, uncontrollable impact of technology, dynamic administrative environment

and uncertain face of vigorous infrastructure. As per the report furnished by Bloomberg New Energy Finance, faces of Indian vehicle industry on roads are going to change soon. Autonomous vehicles and Electric vehicles shall soon transform our cities to smart cities where electric two wheelers, buses and rickshaws shall run on roads by 2040. There will a great change in the work places in the manufacturing units as we see it today. We need to prepare us for the future period in which we have to face the technological changes brought about by AI. This research service examines the role artificial intelligence (AI) will play in the transformation of the automotive space. In this study, I have attempted to provide an overview of the changes in the automotive industry precipitated by the use of artificial intelligence. For decades, companies have been "digitizing" their plants with distributed and supervisory control systems. While this has greatly improved visualizations for operators, most companies with heavy assets have not kept up with the latest advances in analytics and in decision-support solutions that apply AI. Today's downsized teams of control-room operators are expected to manually monitor a multitude of signals on numerous screens and adjust settings as needed. This heavy reliance on experience makes it difficult to replace a highly skilled operator at retirement. Since variations in operators' qualifications can affect not only performance but also profits, AI's ability to preserve, improve, and standardize knowledge is all the more important. Moreover, since it can make complex operational set-point decisions on its own, AI is able to reliably deliver predictable and consistent output in markets that have difficulty attracting and retaining operator talent. With the help of AI, the system operates continuously and independently of any variations in experience, attention, or other negative influences. Asset-optimizer solutions have been developed and successfully deployed in chemicals, metals, mining, and other heavy-manufacturing environments, demonstrating that AI solutions are viable and economically attractive to a range of companies with heavy assets.

II. LITERATURE REVIEW

Currently, AI is being implemented in automotive manufacturing, including design, supply chain, production, and post-production. In addition, AI is being implemented in ‘driver assistance’, ‘driver monitoring’, and ‘driver risk assessment’ systems, which is transforming the transportation sector. Aftermarket services such as predictive maintenance and insurance are also transforming through the usage of AI.

Robots are simultaneously working with human and learning automotive manufacturing skills (design, part manufacturing, and assembly) with the help of AI. Currently, AI helps humans to develop cars using Exoskeletons; in the future, the entire plant can be operated by these AI-powered robots. The BMW group uses automated image recognition for quality checks, inspections, and to eliminate pseudo-defects (deviations from target despite no actual faults). As a result, they’ve achieved high levels of precision in manufacturing.

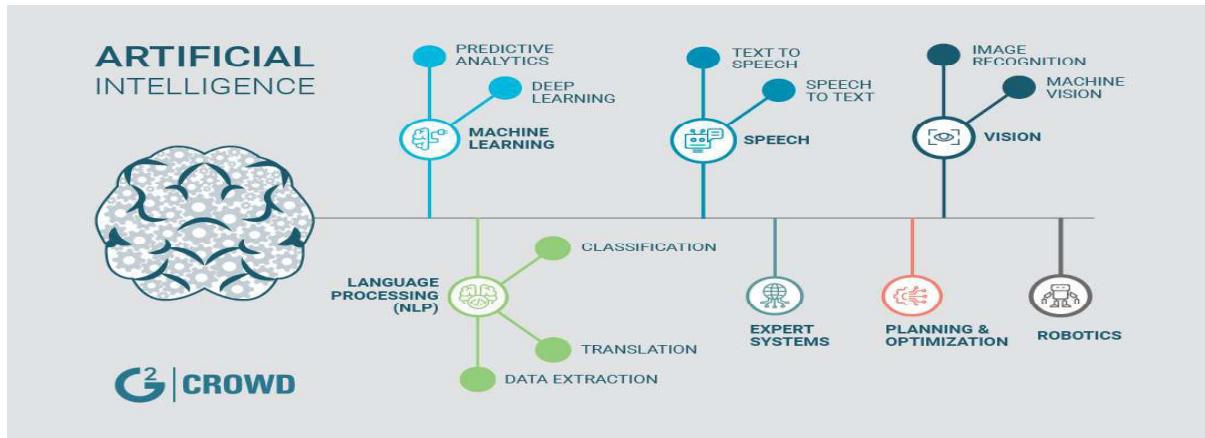
Another company that’s benefited from AI in manufacturing is Porsche. They use autonomous guided vehicles (AGVs) to automate significant portions of automotive manufacturing. The AGVs take vehicle body parts from one processing station to the next, eliminating the need for human intervention and making the facility resilient to disruptions like pandemics.

- *Machine Learning:* Applications make decisions based on data. A distinction is made between supervised learning and unsupervised learning. Supervised learning requires pairs of inputs and outputs. As soon as the system has been trained with enough data, it is able to create connections on its own. For example, Linear Regression etc. In unsupervised learning, the output labels are not provided and the application creates its own model with its own classifiers. For example, clustering etc. [1].

- *Deep Learning:* is a subset of Machine Learning. Also called the supreme discipline of AI. It is used to examine large amounts of data for patterns and trends — e.g. in face, object or speech recognition. So-called neural networks are used for this purpose, which are inspired by the biological neural networks of our human brain. In deep learning, a computer model learns to perform classification tasks directly from the images, text or sound without human intervention. In contrast to machine learning, which works with fixed model groups, deep learning algorithms develop the models independently [2].

- *Computer Vision:* Computer vision tasks include methods for acquiring, processing, analysing and understanding digital images or videos and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information. In the automotive industry, computer vision is used for inspection of every single detail during production. In simple words, it detects flaws in every single product produced, rejecting the ones which have one. This includes detecting surface detection (finding dents, dings, etc.) and functional flaws [3]. Computer vision could make a real difference in the automotive industry by adding important safety feature to our vehicles. If a car could detect danger, it could stop before an accident happens and save countless lives and property. It is used to detect simple objects such as road lanes, traffic signs and road boundaries.

- *Natural Language Processing (NLP):* Natural Language Processing attempts to process natural language on a computer. This should enable applications to be controlled by speech. Everyone who has ever talked to Alexa, Siri, Google Assistant or Cortana has been an NLP user.



Graphic: <https://media.g2crowd.com>

A. Transparency in the lower-level supply chains

Intelligent sustainability radar for the Supply Chain: The Porsche, Audi and Volkswagen brands are using Artificial Intelligence to identify sustainability risks such as environment pollution, human rights abuses and corruption at an early stage –not only among direct business partners but also at lower levels of supply chain. Artificial Intelligence simplifies the complex analysis of data, allowing them to address partners directly and request improvements in sustainability.

B. Key uses of AI in manufacturing of Automobiles

- *Defect Detection*

The timely detection of faults or defects and taking appropriate actions are essential to reduce operational and quality-related costs. Today, many assembly lines have no systems or technologies in place to identify defects across the production line. Even those which may be in place are very basic, requiring skilled engineers to build and hard-code algorithms to differentiate between functional and defective components. The majority of these systems cannot still learn or integrate new information, resulting in countless false-positives, which then have to be manually checked by an on-site employee.

By imbuing this system with artificial intelligence and self-learning capabilities manufacturers can save countless hours by drastically reducing false-positives and the hours required for quality control. AI-powered defect detection taps into

computer vision, which uses high-resolution cameras to monitor every aspect of the production process. Such a system can flag defects that the human eye might miss and trigger correcting measures automatically. This helps reduce product recalls and cut down on wastage.

- *Quality Assurance*

Manufacturing requires acute attention to detail, a necessity that's only exacerbated in the electronics space. Historically speaking, quality assurance has been a manual job, requiring a highly skilled engineer to ensure that electronics and microprocessors were being manufactured correctly and that all of its circuits were properly configured.

Today, image processing algorithms can automatically validate whether an item has been perfectly produced. By installing cameras at key points along the factory floor, this sorting can happen automatically and in real-time.

- *Assembly Line Integration*

Manufacturers use a variety of equipment and all these equipment send a wide array of data to the cloud. Unfortunately, this information tends to be siloed and doesn't play nicely together.

To get a holistic picture of your operation may require several different dashboards and a subject matter expert to make sense of it all. AI projects in assembly mainly focus on automated image recognition. Here, the technology is used to evaluate images of a component and compare them in milliseconds with hundreds of other images from the same sequence [4]. The system then identifies any

deviations from the norm, such as parts that are incorrectly positioned or fitted, or absent.

By creating an integrated application which can pull data from all the Internet of Things-connected equipment in your ecosystem, you can ensure that you are getting a bird-eye view of all the operations.

- *Assembly Line Optimization*

Moreover, by implementing artificial intelligence in your IOT ecosystem on top of all the data insights can help in automation of the assembly line. For example, when equipment operators are showing signs of fatigue, supervisors get notifications. When a piece of equipment fails, the system can automatically trigger contingency plans or other reorganization activities.

- *Generative Design*

Artificial intelligence is playing a major role in changing the way manufacturing companies design products using generative design. Here's how it works: a designer or an engineer inputs design goals into generative design algorithms[5]. Taking into consideration all these parameters, the algorithm will explore every possible permutation of a solution and generate design alternatives. Finally, it uses machine learning to test each of the iterations and improve upon it.

- *Predictive Maintenance*

Predictive maintenance aims at improving asset productivity by using data to anticipate machine breakdowns. A well-established and relatively simple method of recognizing failures early on is condition monitoring. The complexity of forecasting failure is often due to the enormous amount of possible influencing factors. Data sources can be manifold and depend on the scenario. E.g., in engines, gear boxes, or air conditioning, analysis of sound can detect an anomaly in device operation. In switches, machines, and robots, vibrations can be measured and used to detect errors. Since new sensors and IOT devices can be integrated in production processes and operations, the availability of data increases drastically[6]. AI-based algorithms are capable of recognizing errors and differentiating the noise from the important information to predict breakdowns and guide future decisions. Over the years, the manufacturing industry

has been using strategy to repair machines components after they fail. After that, the industry started to adopt preventive maintenance where the machine maintenance is done as per the planned schedule that is created while considering its previous failures. When artificial intelligence came into being, manufacturers started eliminating machine failures using predictive maintenance. In predictive maintenance, we can predict the time of potential component failures with the help of machine learning models.

Technology implementation has increased, and the post-pandemic situation appears to be positive for all stakeholders; however, automakers have yet to fully harness AI's potential in their service offerings[7]. Although AI is in the nascent stage of development, original equipment manufacturers (OEMs) are adopting it across the automotive value chain to improve manufacturing and to enhance customer experience, marketing, sales, and after-sales services.

III. METHODOLOGY

The world is about to be flooded with data, with the incubation of the Internet of Things and the overall trend of digitization generating data in places that we have never seen before. Analysing this type of data has become a recent area of focus in artificial intelligence, as accurate forecasting is becoming increasingly vital across all kinds of industries in order to make more informed decisions. In this study, an attempt is made to make a trend analysis of the adoption of Artificial Intelligence in the automotive industry. In order to obtain a fair overview over the use of AI methods and possible trends, a systematic review is used as method of choice in this paper.

A. Objectives of the Study:

1. To study the maturity level for AI adoption and AI developments in Automotive Industry.
2. To study the global markets trends of Autonomous Vehicles.

B. Significance of the study

Automobile industry is facing challenges due to globalization, acute competition and the ever changing demands of the customers. Indian industry is now a part of the global scenario and requires adapting the changes taking place all over the

world. The ubiquitous presence of the Information Technology has increased the pace of advancement all over the world. To sustain in the global competition Indian Automobile Industry has to embrace the new systems and keep updated as per the industry norms. So, it is the need of the hour to study the trends in the automotive industry so that the manufacturers can keep pace with the changing scenario.

C. Data Collection

For the purpose of this study, secondary data was collected from the Annual reports, investor presentations and press releases of companies operating in the market. Data was also collected from the Industry journal and studies published by

relevant association like International Organization of Motor Vehicle Manufacturer, Alliance of Automobile Manufacturers, National Automobile Dealers Association etc.

IV. DATA ANALYSIS

Data analysis and interpretation is the process of assigning meaning to the collected information and determining the conclusions, significance, and implications of the findings. Interpreting the analysed data from the appropriate perspective allows for determination of the significance and implications of the assessment. The data has been collected from various websites, reports etc. and analysed using charts and diagrams.

EXHIBIT 1: Industries based on their maturity level for Artificial Intelligence Adoption



Source: FutureBridge Analysis and Insights

It is very clear from Exhibit 1 that automotive industry is a new adopter of Artificial intelligence and it has to grow by leaps and bounds in area of Artificial intelligence. Artificial Intelligence (AI) is being used by the agriculture industry to help produce healthier crops, control pests, monitor soil and growing conditions, organise data for farmers, reduce effort, and improve a wide range of agriculture-related operations along the food supply chain. Smart sorting is already applied in the mining industry for

minerals and ores, and artificial intelligence algorithms powered by colour sensors and X-ray data are improving the quality and quantity of, for example, the diamond recovery process. Food and beverages organizations are rapidly digitizing their supply chain to differentiate, where AI is helping the organization to analyse this data and gain a better understanding of variables in the supply chain by anticipating future scenarios. Despite the baby steps these industries are taking, these can be grouped under **strugglers**.

The adoption of AI encourages chemical companies to make essential changes in machinery and processes to reduce pollutants being released into water bodies and the environment. AI helps oil and gas companies assess the value of specific reservoirs, customize drilling and completion plans according to the geology of the area, and assess risks of each individual well. AI is catching up with the chemical and oil industry but still these industries can be grouped as **beginners**. The developments are taking place in chemical and oil and natural gas as regards adoption of AI but still these industries can be grouped as beginners.

The automotive industry is regarded as a **new adopter** of AI.

The developments are taking place in life sciences' sector, mainly biomedicine and healthcare. AI is improving patient and clinical trials' data collection for drug development and repurposing. These industries can be categorized as **aspirational industries**.

AI can also lessen financial crime through advanced fraud detection and spot anomalous activity as company accountants, analysts, treasurers, and investors work toward long-term growth. ICT and Financial services industries are considered to be the **matured industries** in the race of AI adoption.

EXHIBIT 2: AI Development in Automotive Industry

Development → In March, 2021, Porsche, Audi and Volkswagen use Artificial Intelligence to minimise sustainability risks[8].

Description → Intelligent sustainability radar for the supply chain: the Porsche, Audi and Volkswagen brands are using Artificial Intelligence (AI) to identify sustainability risks such as environmental pollution, human rights abuses and corruption at an early stage – not only among direct business partners but also at the lower levels of their supply chain.

Development → In September, 2021, Maruti Suzuki India Limited launched an industry-first AI (Artificial Intelligence) based 24x7 virtual car assistant app, called “S-Assist” for its NEXA customers.

Description → The scan and voice-enabled virtual car assistant, S-Assist, optimizes the application of artificial intelligence and machine learning to provide an immersive online post-purchase experience to customers. S-Assist is a complimentary service which offers quick access to vehicle features, troubleshooting and driving tips on customers' smartphone. It brings real-time information to customers for their vehicle-related queries. Maruti Suzuki NEXA customers will be able to access S-Assist through the “Maruti Suzuki Rewards” mobile app available on both iOS and Android devices.

Development → In 2021, Mercedes-Benz Transformed Vehicle Cockpit with NVIDIA-Powered AI

Description → The AI cockpit has reached galactic proportions with the new Mercedes-Benz MBUX Hyperscreen.

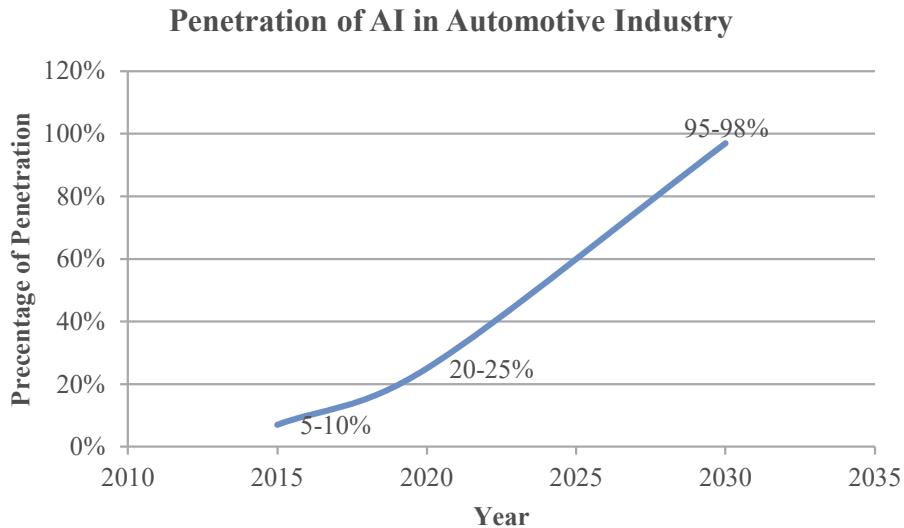
During a digital event, the luxury automaker unveiled the newest iteration of its intelligent infotainment system — a single surface extending from the cockpit to the passenger seat displaying all necessary functions at once. Dubbed the MBUX Hyperscreen, the system is powered by NVIDIA technology and shows how AI can create a truly intuitive and personalized experience for both the driver and passengers. The MBUX Hyperscreen reinvents how we interact with the car.

Development → BMW Group scaled artificial intelligence for data privacy in production – with innovative anonymisation algorithms and in April, 2021, BMW Group published AI anonymisation algorithms.

Description → The BMW Group published an anonymization solution based on artificial intelligence (AI) that can anonymise objects in photos and videos. Building on the BMW labelling tool Lite, these algorithms (github.com/BMW-InnovationLab) enable targeted protection of relevant information: The user-friendly software tool uses AI to block out or blur objects or people. The granularity and degree of anonymisation can be intuitively adjusted. Users with no programming skills can create AI applications, including data privacy, in a short time.

Source: Company website

EXHIBIT3: Chart showing penetration of AI in Automotive Industry



Source: FutureBridge Analysis and Insights

With AI making inroads in the automotive industry, there are some key aspects to look at. Some emerging trends like fully autonomous vehicles are expected to become a reality in the future. In 2015, penetration of AI was 5-10% which rose to 20-30% in the year 2020. In the year 2030, the use of AI is going to rise to 95-98%.

A. Global Autonomous Cars Market size

Autonomous vehicles on the production floor, like the ones used by Porsche can automate everything from assembly lines to conveyor belts. Self-driving cars and ships can optimize deliveries, operate 24/7, and speed up the overall delivery process. The demand for autonomous vehicles is rising steadily and is expected to make up 10-15% of global car sales by 2030.

Connected vehicles equipped with sensors can also track information about traffic jams, road conditions, accidents, and more in real-time to optimize delivery routes, reduce accidents, and even alert the authorities in case of emergencies. This improves delivery efficiency and road safety.

Self – driving cars have become the major revolution in automation and connectivity technology[9]. Computers, the internet, smartphones and advanced technologies merge with vehicles to assist and automate driving operations. Autonomous driving holds the promise of a smoother, safer, and more comfortable mobility experience. The automotive industry is on a continuous journey from assisted to autonomous driving.

EXHIBIT 4:Table showing Global Autonomous Cars Market Size

Year	Market size in USD (Billions)
2020	1.45

2021	1.64
2028	11.03 (projected figure)

The Global autonomous cars market size was USD 1.45 billion in 2020. The Global market exhibited a huge decline of 20.3% in 2020 as compared to average growth during 2017-2019. The market is projected to grow from USD 1.64 billion in 2021 to USD 11.03 billion in 2028 at CAGR of 31.3% in 2021-2028 period.

B. Advanced Driver Assistance System(ADAS) Market growth.

The advanced driver assistance system(ADAS) cover a wide range of active and passive system designed to assist the driver by providing safety, comfort and efficiency while driving and improves the driver, passenger and pedestrian security and safety. It has vital components such as sensors, cameras, radars and software that help the system function precisely. Autonomous emergency-braking, blind-spot detection, park assist, adaptive cruise control and lane warning system are a few systems included in ADAS[10]. The improved industrialization and living standards of people has impacted Advanced Driver Assistance System(ADAS) Market growth.

EXHIBIT 5:Table showing Advanced Driver Assistance System (ADAS) Market Growth

Year	Market size in USD (Billions)
2020	27.29
2021	27.52
2028	58.59 (projected figure)

The global Advanced Driving Assistance System (ADAS) Market size was USD 27.29 billion in 2020. The market is projected to grow to USD 58.59 billion in 2028 from USD 27.52 billions in 2021 at a compound annual growth rate (CAGR) of 11.4% in 2021-28 period.

V. CONCLUSION

It's time to realize that there is opportunity cost of not investing in AI adoption. Companies that realized the potential of AI early, such as Google and Amazon, have far outperformed their peers and grown aggressively, largely due to their superior ability to predict and continuously adapt to changing conditions and to generate higher margins. The automotive industry is emerging as a major source of AI and machine learning. The importance of artificial intelligence (AI) to the automotive industry over the coming decade cannot be overstated. The variety of possible applications of machine learning in the automotive industry is impressive. Manufacturers can deploy AI technologies for designing and building new prototypes, improving the efficiency of their supply chains and enabling predictive maintenance for both factory equipment and vehicles on the road. AI is increasingly being applied in automobiles: automated driving is the best-known example, but AI also finds a role in a wide variety of other applications, such as safety features for vehicles, comfort functions, Advanced Driver Assistance Systems, warning and 'driver risk assessment' systems, connectivity systems, infotainment systems, etc. Aftermarket services such as predictive maintenance and insurance are also being transformed through AI use. But despite its promising potential, the use of AI in the automotive industry is associated with several challenges like algorithm biases, data quality and understanding how a model came to a certain conclusion. Public trust and confidence will stimulate adoption and use of AI and automation across the entire mobility ecosystem.

REFERENCES

- [1] A. K. Agrawal, J.S. Gans, & A. Goldfarb, "What to expect from artificial intelligence," *MIT Sloan Management Review*, vol.58, no. 3, pp. 23–27, February 2017.
- [2] A. I. Canhoto, & F. Clear, "Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential," *Business Horizons*, vol. 63, no. 2, pp. 183–193, March-April 2020.
- [3] P. Choudhury, E. Starr, & R. Agarwal, "Machine learning and human capital complementarities: Experimental evidence on bias mitigation," *Strategic Management Journal*, vol. 41, no. 8, pp. 1381–1411, March 2020
- [4] B. Buchmeister, I. Palcic, & R. Ojstersek, "Artificial Intelligence in Manufacturing Companies and Broader: An Overview", in *DAAAM International Scientific Book 2019*, Ed. DAAAM International, 2019, pp. 81-98.
- [5] Birlasoft.com, '17 Remarkable Use Cases of AI in the Manufacturing Industry'. [Online]. Available: <https://www.birlasoft.com/articles/17-use-cases-of-ai-in-manufacturing>. [Accessed: 10-Feb- 2022].
- [6] P. Helo, & Y. Hao, "Artificial intelligence in operations management and supply chain management: an exploratory case study", *Production Planning & Control*, vol. 31, pp. 1-18. January 2021.

[7] Towardsdatascience.com, 'How can Artificial Intelligence Be Applied in Manufacturing?' 2020. [Online]. Available: <https://towardsdatascience.com/how-can-artificial-intelligence-be-applied-in-manufacturing-8662eaaea999>. [Accessed: 11-Feb- 2022].

[8] Volkswagenag.com, 'Skoda Auto Volkswagen India: The Three Highlights Projects 2020', 2020. [Online]. Available: <https://www.volksvagenag.com/en/sustainability/reporting/regions/india.html>. [Accessed: 11-Feb – 2022].

[9] Marutisuzuki.com, 'MarutiSuzuki unveils S-Assist: An industry-first AI-based 24x7 virtual car assistant for NEXA', 2021. [Online]. Available: <https://www.marutisuzuki.com/corporate/media/press-releases/2021/september/maruti-suzuki-unveils-s-assist-an-industry>. [Accessed: 11-Feb – 2022].

[10] Fortunebusinessinsights.com, 'The global autonomous cars market is projected to grow from \$1.64 billion in 2021 to \$11.03 billion in 2028 at a CAGR of 31.3% in forecast period.[Online]. Available:<https://www.fortunebusinessinsights.com/industry-reports/autonomous-cars-market-100141>. [Accessed: 11-Feb – 2022].

[11] Researchandmarkets.com, 'Global Autonomous Driving Industry Outlook, 2021 - Competitive Intensity Propelling the Market'. [Online]. Available: <https://www.researchandmarkets.com/reports/5323336/global-autonomous-driving-industry-outlook-2021>. [Accessed: 11-Feb – 2022].

[12] McKinsey.com, 'Smartening up with Artificial Intelligence (AI) – What's in it for Germany and its Industrial Sector?' [Online]. Available: <https://www.mckinsey.com/-/media/mckinsey/industries/semiconductors/our%20insights/smartening%20up%20with%20artificial%20intelligence/smarten-up-with-artificial-intelligence.ashx>. [Accessed: 11-Feb – 2022].

[13] Oica.net, 'Prospects for the future of the automotive industry by OICA members'. [Online]. Available:https://www.oica.net/prospects-for-the-future-of-the-automotive-industry-by-oica-members/file:///C:/Users/AVI/Downloads/AI_and_smart_use_of_data_at_BMW_Group_Plant_Munich.pdf. [Accessed: 11-Feb – 2022].

Significance of Business Intelligence behind Customer Loyalty

Gurpreet Singh

Department of Commerce & Economics

B.A.M. Khalsa College, Garhshankar

Panjab University, Chandigarh

Punjab, India

gurpreet.singh@bamkc.edu.in

Abstract-

The expense of holding existing customers through customer loyalty is much lower than the new customer acquisition cost, particularly in full grown and cutthroat business sectors. Along these lines, each organisation ought to have such strategy that prompts customer retention. In the competitive stage, some business houses step up and foster unique customer loyalty programs that give them an edge in contrast with comparative organizations. Business intelligence has made it simpler for organizations to comprehend their clients and learn new strategies to make them want more. The market is growing rapidly and the data becomes gigantic which suggests that analytics tools are growing too. From here, the requirement for business intelligence frameworks arises to assist with figuring out the various levelled data. This paper highlights the role of BI in client relationships strategy and proposed the model related to it.

Index Terms- Business Intelligence, Business Strategy, Big Data, Customer Loyalty, Customer Retention

I. INTRODUCTION

Customers are a significant part of any business. Nobody can profess to be maintaining a fruitful business without the requirement for a solid client base. In any case, there is dependably contest in business. Currently, every business needs to be very cautious regarding their goods and services offered to their clients because awful quality items and services experiences of clients will lead to gain to the opposition. Further, it may affect the future accomplishment of the business. You want to comprehend them to hold your current clients. Hence, will want to convey definitively what they need. The benefits coupled with customer loyalty are widely acknowledged in the literature. The cost of retaining existing customers through customer loyalty is much lower than the new customer acquisition cost, especially in mature and competitive markets [1]. Depending on the industry, an improvement of a mere five percent in customer retention leads to a boost of 25 to 85 percent in earnings [2]. A firm bear four to five-fold extra costs to obtain new customers rather than sticking with the present ones [3, 4]. An average

company loses half of its customers in five years [5]. Therefore, every company should have a successful customer loyalty strategy that leads to customer retention. In the competitive platform, some business houses take initiative to develop unique customer loyalty programs that give them an edge in comparison to similar organizations.

In this way, organizations ought to endeavour to hold their clients no matter what. That is not dependably straightforward, but rather the appearance of large information means big data and business intelligence has made it simpler for organizations to comprehend their clients and learn new strategies to make them want more. The market is developing quickly and the information becomes tremendous from various resources, and that implies that the interest for easy-to-understandanalytics tools is growing too. From here, the requirement for business intelligence frameworks emerges to assist with sorting out the hierarchical information. This paper will assist with distinguishing the job of BI in client relationships and how it becomes fundamental to involve such canny instruments in the insanely developing information age [6].

II. REVIEW OF LITERATURE

A. Business Intelligence

Traditionally, Business Intelligence initially arose during the 1960s as a way for sharing data across associations. During the 1980s alongsidecomputer models for transforming information into bits of knowledge before turning into a particular contribution from BI groups with IT-dependent assistance arrangements. Presently, BI arrangements focus on adaptable self-service analysis, represented information on trusted platforms, engaged business clients, and speed to understanding. Significantly more than a particular "thing," business intelligence is somewhat an umbrella term that covers the processes and techniques for gathering, storing, and analyzing information from business tasks or exercises to improve execution. These things meet up to make a thorough perspective on a business to assist with peopling improve,

significant choices [7]. A definitive objective of BI drives is to drive better business choices that empower associations to build an income, work on functional proficiency and gain upper hands over business rivals. To accomplish that objective, BI consolidates a blend of analytics, data management, and reporting tools.

Throughout recent years, business insight has advanced to incorporate more processes and exercises to assist with further developing performance. BI associates the business investigation, data mining & representation, data tools, and infrastructure, and assists business organizations to make data-driven choices [8]. BI (Business Intelligence) is a system of collection, designs, and advancements that transform crude information into significant data that drives beneficial business activities. It is a set-up of programming and administrations that convert information into significant intelligence and information. BI directly affects the methodology, strategic and functional organization choices of the organization. BI advances reality-based dynamics utilizing recorded information rather than presumptions. BI instruments attempt information, data analysis and make reports, synopses, dashboards, guides, diagrams, and graphs which give clients nitty-gritty knowledge about the idea of the business [9].

Business knowledge (BI) is the utilization of computing technologies (applications and programming) to gather business information, data from numerous resources and analyze it then, at that point, change it into valuable experiences that help administrators and proprietors to make the right moves to further develop the business performance and meet the objectives expected to business achievement [6]. Business Intelligence can help the organization in understanding its clients to work on its relationship with them, as the quicker transformation of potential into genuine clients, diminishing the number of outgoing clients and incrementing sales to existing clients, and that in its turn will expand sales and revenue [10]. This paper will talk about a portion of the procedures BI utilizes in working on the relationship with customer loyalty, on the whole, it will clarify the idea of the BI framework and the job of every part. A business intelligence design incorporates something other than BI programming. Business intelligence data is commonly stored in a data stockroom or warehouse worked for a whole association or in more modest information shops that hold subsets of business data for individual offices and specialty units, frequently with ties to an organizational data stockroom.

BI data can incorporate historical data and ongoing information assembled from source frameworks as it's created, empowering BI instruments to help both key and strategic decision-making processes. Before it's utilized in BI applications, crude information from various source frameworks for the most part should be coordinated, merged,

and cleansed utilizing data mix and information quality management tools to guarantee that BI groups and business clients are analyzing accurate and consistent information [11]. Business intelligence can assist organizations with settling on better choices by showing present and historical data inside their business setting. An analyst can use BI to give performance and contender benchmarks to make the association run smoother and all the more productively. An analyst can likewise more effectively spot market trends to build up sales or income. Utilized actually, the right information can assist with anything from consistency to recruiting endeavours. A couple of ways that business intelligence can assist organizations with making more smarter, information-driven choices: Recognize ways of expanding benefit, analyse customer behaviour, compare data with competitors, track performance, optimize operations, predict success, spot market trends, discover issues or problems [10].

B) Customer Loyalty And Its Drivers

Typically, literature discusses customer loyalty from two perspectives viz. attitudinal and behavioural. Primarily loyalty is treated as a customer attitude that leads to a relationship with the service provider and its brand or service. Secondly and mainly, it is treated as customer's revealed behaviour. Attitudinal loyalty is the favourable attitude towards the brand or the organization. It is the customer's psychological attachment toward the same brand or same brand set [12, 13]. Many researchers state that for true loyalty there must be strong "attitudinal commitment" and "deeply held attachment" to a brand and customers are not attracted to competitors' offerings [14, 15]. Behavioural loyalty is reflected as a particular brand's purchase frequency and the possibility of repurchasing [16]. Behavioural loyalty is reflected by "depth of repeat" i.e. repurchase of the items twice, thrice, and many times; which ultimately enhances their engagement with the same firm. Loyalty implies distinctive meanings to different individuals and will draw diverse responses and actions from customers [17].

Customer Satisfaction: Customer satisfaction is an assessment of feeling, mirroring how much clients trust their service provider, which in turn inspires positive emotions [18]. Customer satisfaction is one of the important drivers of customer loyalty. Many studies support this connection. "The relationship between satisfaction and loyalty is non-linear, meaning that in case satisfaction increases above a certain level, customer loyalty will increase rapidly" [19]. "Customer satisfaction is an important driver of the customer loyalty and it appears after providing active service and ready availability of the merchandise" [20]. **Commitment:** Commitment is characterized as an emotional or mental attachment to a brand inside an item class [13, 20, 21]. Commitment is a significant,

logical, useful, and key predecessor of firm loyalty [22]. Commitment and loyalty are two different but inter-connected concepts. In the purchaser and merchant relationship studies, commitment is utilized as one of the vital, common, and dependent variables [23]. Undoubtedly, commitment surpasses the system of the good state of mind towards the brand, it has a more grounded quality, heartiness, and steadiness than the general disposition or attitude towards the brand. Affective commitment creates a more grounded connection to loyalty. Affective commitment has an immediate constructive effect on client loyalty and mostly intervenes in the connection between service reliability and customer loyalty [24]. Trust: Trust is the conviction that an accomplice's or partner's guarantee is dependable and a stakeholder will satisfy his/her commitments in the relationship [25]. Trust is another variable that is instrumental in the building of customer loyalty. Trust is the main cause of relationship commitment and occurs when there is an assurance of a partner's trustworthiness and integrity [26]. Trust is the key to the determination of long-term orientation between buyer and seller. Trust is the most important determinant of customer loyalty [27]. Image: "Image is seen as the organization's presentation of itself to its various stakeholders and how it distinguishes itself from all other organizations" [28]. Corporate image is one of the major antecedents of customer loyalty [29]. Corporate image is impacted both by administration quality and customer satisfaction, which in turn impacts client loyalty. loyal to the firm. Hung (2008) suggests that a favorable brand image can enhance customer loyalty. Brand and corporate image are one of the major antecedents of customer loyalty [30, 31].

III PROPOSED CONCEPTUAL MODEL

Business intelligence is a process - a process of leveraging customer information to enhance corporate behaviour and improve your relationship with current and target customers for enhanced profitability and competitive advantage [32].

Business Intelligence, gives a holistic approach to dealing with customers which includes enhancements in customer profiling, less complex incentive findings for customers, estimating the success of the company in satisfying its customers, and making extensive customer relationship management. Business Intelligence can recognize different motivations to increase sales and revenue, such as the faster conversion of potential into actual clients, reducing the number of outgoing customers, and increasing sales to existing customers [13, 34, 35].

Drawing on insights acquired from BI and CRM literature, the accompanying research model is proposed and presented in Fig. 1. The proposed model portrays the following three variables: (i) organization's Business Intelligence, (ii) organization's Business Strategy, and (iii)

organization's CRM Strategy i.e., enhancement of customer



loyalty [ECL] through its drivers

Fig: 1 Proposed Conceptual Model

The incorporation of the organization's BI variable in the proposed model is reliable with the general management theory. This theory argues that BI systems enable "evidence-based management" in business [36]. More specially, the BI framework support decision-making by structuring, storage, and use of large amounts of high-quality data in data warehouses. Reporting, dashboard, and online analytical processing technologies and sophisticated statistical tools facilitate decision-makers' interpretation of organizational data and enhance their decision-making capabilities [35].

The business strategy focuses on a review and articulation of a company's vision as it relates to CRM. It also involves a review of the industry and competitive environment [37]. Most importantly, the business strategy determines how the customer strategy should be developed and how it should evolve over time. The customer strategy focuses primarily on examining the existing customer base. Besides, it involves identifying the most appropriate drivers that can enhance customer loyalty.

With respect to relationships, the model proposes three plausible associations among model variables. First of all, the model recognizes that an organization's BI plays an important role in achieving customer loyalty through strategic goals in CRM by allowing organizations to get valuable insights into their operations and customers. Typically, these insights are obtained through data mining & representation, data tools, and infrastructure, and assist business organizations to make data-driven choices. In this way, BI enables organizations to develop dynamic capabilities and apply well-informed business strategies [38].

Next, the model suggests that an organization's business strategy directly affects its customer strategy. Organizations can undertake a variety of business as well as competitive strategies to accomplish their customer relationship goals. Generally, well-informed competitive strategies that focus on customers for products, markets, or channels result in performance gains and create competitive advantage [38]. The proposed model anticipates that more goal-orientated business strategies will ultimately lead to superior customer strategies.

Finally, the model proposes the organization's CRM strategy. If an organization works on the important drivers of customer loyalty including customer satisfaction, commitment, trust, and image an organization can get advantages in the present competitive environment.

IV IMPLICATIONS, LIMITATIONS, AND FUTURE DIRECTIONS

The current proposed model has significant implications for research. For research, the current model uncovered the associations between the organization's BI, Business strategies with the Organisation's CRM strategy. It provides a superior and improved understanding of BI-enabled customer loyalty. Despite being relevant and interesting, the proposed model is restricted by the particular research context. Future exploration should be possible through empirical testing.

V CONCLUSION

This paper proposed a couple of chosen factors to work on the current understanding of the role of BI in customer loyalty. The conceptual model was proposed based on earlier literature evidence. The current research contributes through significant findings that BI can provide benefits in CRM by enabling an improved strategy development process and thus, organizational performance and competitive advantage. The present study provides valuable guidance to managers on how to leverage BI for greater benefits. However, the researchers need to be cautious due to a series of conceptual and contextual limitations. These limitations pose challenges and provide opportunities for future research.

REFERENCES

- [1] A. Ehrenberg, & G. Goodhardt, "New brands: Near-instant loyalty," *Journal of Marketing Management*, 16(6), 607-617, 2000.
- [2] F. P. Reichheld, & W. E. Sasser, "Zero defections: Quality comes to services" *Harvard Business Review*, 68(5), 105-111, 1990.
- [3] P. Kotler, & K. L. Keller, *Marketing Management* (12 ed.), Upper Saddle River, NJ:Prentice Hall, 2006.
- [4] V. A. Zeithaml, L. L. Berry, & A. Parasuraman, "The behavioral consequences of service quality," *The Journal of Marketing*, 60 (2), 31-46, 1996.
- [5] F. F. Reichheld, & T. Teal, *The Loyalty Effect*, Harvard Business School Press. Boston, MA, 1996.
- [6] https://www.researchgate.net/publication/338326425HowBusiness_Intelligence_Can_Help_You_to_Better_Understand_Your_Customers [Online] [Assessed February 10, 2022].
- [7] <https://www.tableau.com/learn/articles/business-intelligence> [Online] [Assessed February 11, 2022].
- [8] https://www.omnisci.com/technicalglossary/_business-intelligence [Online] [Assessed February 10, 2022].
- [9] <https://www.guru99.com/business-intelligence-definition-example.html> [Online] [Assessed February 11, 2022].
- [10] A. Habul, A. Pilav-Velic, "Business Intelligence and Customer Relationship Management" *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces*, Cavtat, 2010, pp. 169-174. [Online]. Available: <https://ieeexplore.ieee.org/document/5546381?arnumber=5546381/>. [Online] [Assessed February 15, 2022].
- [11] https://courses.lumenlearning.com/introduction_to_business/chapter/businessintelligence [Online] [Assessed February 15, 2022].
- [12] J. Jacoby, & D. B. Kyner, "Brand loyalty vs. repeat purchasing behaviour," *Journal of Marketing Research*, 10(1), 1-9, 1973.
- [13] H. Evanschitzky, B. Ramaseshan, D. M. Woisetschlag, V. Richelsen, M. Blut, & C. Backhaus, "Consequences of customer loyalty to the loyalty program and to the company," *Journal of the Academy of Marketing Science*, 40(5), 625-638, 2012.
- [14] J. Jacoby, & D. B. Kyner, "Brand loyalty vs. repeat purchasing behaviour," *Journal of Marketing Research*, 10(1), 1-9, 1973.
- [15] J. Kandampully, T. Zhang, & A. Bilgihan, "Customer loyalty: A review and future directions with a special focus on the hospitality industry," *International Journal of Contemporary Hospitality Management*, 27(3), 379-414, 2015.
- [16] Y. Yi, & H. Jeon, "Effects of loyalty programs on value perception, program loyalty, and brand loyalty," *Journal of the Academy of Marketing Science*, 31(3), 229-240, 2003.
- [17] G. Singh, I. Singh, & S. Vij, "Antecedents and consequences of customer loyalty: a conceptual model," *International Journal of Applied Business and Economic Research*, 15(23), 237-251, 2017.
- [18] E. W. Anderson, C. Fornell, & D. R. Lehmann, "Customer satisfaction, market share, and profitability: Findings from Sweden," *The Journal of Marketing*, 58 (July), 53-66, 1994.
- [19] R. L. Oliver, Satisfaction: A behavioral perspective on the customer, New York: McGraw-Hill, 1997
- [20] M. C. Cant, & M. Du Toit, "Identifying the factors that influence retail customer loyalty and capitalizing them," *International Business & Economics Research Journal (IBER)*, 11(11), 1223-1232, 2012.
- [21] S. E. Beatty, P. Homer, & L. R. Kahle, "The involvement—commitment model: Theory and implications," *Journal of Business Research*, 16(2), 149-167, 1988.
- [22] S. E. Beatty, & L. R. Kahle, "Alternative hierarchies of the attitude-behavior relationship: The impact of brand commitment and habit," *Journal of the Academy of Marketing Science*, 16(2), 1-10, 1988.
- [23] D. T. Wilson, "An integrated model of buyer-seller relationships," *Journal of the Academy of Marketing Science*, 23(4), 335-345, 1995.
- [24] E. E. Izogo, & E. E. Izogo, "Customer loyalty in telecom service sector: The role of service quality and customer commitment," *The TQM Journal*, 29(1), 19-36, 2017.
- [25] J. B. Rotter, "A new scale for the measurement of interpersonal trust," *Journal of Personality*, 35(4), 651-665, 1967.
- [26] R. M. Morgan, & S. D. Hunt, "The commitment-trust theory of relationship marketing," *The Journal of Marketing*, 58(3), 20-38, 1994.
- [27] N. O. Ndubisi, K. W. Chan, and G. C. Ndubisi, "Supplier-customer relationship management and customer loyalty: The banking industry perspective," *Journal of Enterprise Information Management*, 20(2), 222-236, 2007.
- [28] N. Markwick, & C. Fill, "Towards a framework for managing corporate identity," *European Journal of Marketing*, 31(5/6), 396-409, 1997.
- [29] S. Aydin, & G. Ozer, "The analysis of antecedents of customer loyalty in the Turkish mobile telecommunication market," *European Journal of Marketing*, 39(7/8), 910-925, 2005.
- [30] J. Kandampully, & H. H. Hu, "Do hoteliers need to manage image to retain loyal customers?," *International Journal of Contemporary Hospitality Management*, 19(6), 435-443, 2007.
- [31] J. Kandampully, T. Zhang, & A. Bilgihan, "Customer loyalty: A review and future directions with a special focus on the hospitality industry," *International Journal of Contemporary Hospitality Management*, 27(3), 379-414, 2015.
- [32] J. Hall, Business intelligence: The missing link in your CRM strategy. *Information Management*, 14(6), 36, 2004.
- [33] M. Handzic, K. Ozlen, & N. Durmic, "Improving customer relationship management through business intelligence," *Journal of Information & Knowledge Management*, 13(02), 1450015, 2014.
- [34] Y. Y. Helm, A. E. Khedr, S. Kolief, & E. Haggag, "An enhanced business intelligence approach for increasing customer satisfaction using mining techniques," *International Journal of Computer Science and Information Security*, 17(4), 2019.

Business Intelligence with Intelligent Computing

- [35] A. Habul and A. Pilav-Velic, "Business intelligence and customer relationship management," *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces*, pp. 169-174, 2010.
- [36] J. Pfeffer, and R Sutton, Evidence-based management, *Harvard Business Review*, 84(1), 62–68, 2006.
- [37] A. Payne and P. Frow, "A strategic framework for customer relationship management," *Journal of Marketing*, 69(October), 167–176, 2005.
- [38] R. Sharma, P. Reynolds, R. Scheepers, P. Seddon and G. Shanks, "Business Analytics and Competitive Advantage: A Review and a Research Agenda," In A Respicio, F Adam, G Phillips-Wren, C Teixeira, and J Telhada (eds.), In Proceedings of the 2010 Conference on Bridging the Socio-Technical Gap in Decision Support Systems: Challenges for the Next Decade, Amsterdam, Netherlands: IOS Press, pp. 187–198, 2010.

ICT Channel - A Profitable Impact on Banking Sector

Arvinder Kaur

Assistant Professor

B.A.M. Khalsa College, Garhshankar

Abstract

Information Technology has been one of the most important factors for the development of mankind. Information and communication technology (ICT) is the major advent in the field of technology which is used for access, process, storage and dissemination of information electronically. The growth and advancements in technology has led to a paradigm shift in the entire banking operations and systems. Banking industry is fast growing with the use of technology in the form of ATMs, on-line banking, Telephone banking, Mobile banking etc. Therefore, taking advantage of information technologies (IT) is an increasing challenge for developing countries like India. Hence, the present research paper has made an attempt to study the role of Information and .
Keywords: Information Technology (IT), Banking Industry, Customer Satisfaction, Network.

I. INTRODUCTIONS

Technological sophistication in the banks is aimed at not only providing better services to customers but also to attain competitive advantages among them. In the development of Indian Economy, Banking sector plays a very important and crucial role. With the use of technology there had been an increase in penetration, productivity and efficiency development of a sound and adequate ICT has become a necessity to meet the challenges of growth and diversification of the banking industry. It has given the banks an opportunity to offer a wide range of services to their customers. The usage of Information Technology (IT), broadly referring to computers and peripheral equipment, has seen tremendous growth in service industries in the recent past. The most obvious example is perhaps the banking industry, where through the introduction of IT related products in internet banking, electronic payments, security investments, information exchanges (Berger, 2003), banks now can provide more diverse services to customers with less manpower. Seeing this pattern of growth, it seems obvious that IT can bring about an equivalent contribution to profits. To get the benefits of enhanced technologies, Indian banks are continuously encouraging the investment in information technology (IT), i.e. ATMs, e-banking or net banking, mobile and tele-banking, CRM, computerisation in the banks, increasing use of

plastic money, establishment of call centres, etc. RBI has also adopted IT in endorsing the payment system's functionality and modernisation on an ongoing basis by the development of Electronic Clearing Services (ECS), Electronic Funds Transfer (EFT), Indian Financial Network (INFINET), a Real-Time Gross Settlement (RTGS) System, Centralised Funds Management System (CFMS), Negotiated Dealing System (NDS), Electronic Payment Systems with the 'Vision Document', the Structured Financial Messaging System (SFMS) and India Card – a domestic card initiative, implemented recently (2011).

II. MARKETING STRATEGIES

In order to improve operational efficiency, quality of customer service and to speed it up, the Committee on public Sector banks (1978) recommended a judicious use of computers for selected services of banks. Apart from an increase in efficiency, it will reduce the load of routine and repetitive work and leave sufficient time for staff to provide better customer service. In their study 'Services Marketing - Challenges and Strategies', Dr. Chidambaram and Ms. K. Alamelu (1996) suggested that banks should become technology friendly by investing in technology a bank can carve a niche for itself. Well-furnished premises are a must for the satisfaction of both employees and customers. Professionalised, Well-trained and motivated employees will improve the marketability of a bank. Gaston Leblanc (1990) studied customer motivations towards the use and misuse of an Automated Teller Machine (ATM) customer of a financial institution. An analysis of results based on demographic variables revealed significant differences between users and nonusers in terms of education only. Results also show that convenient accessibility of a financial institution and avoidance of waiting lines are the principal reasons for using the automated teller. Robert Rugimbana and Philip Iversen's study (1994) was to determine the association between consumer ATM usage patterns and their perceptions of ATM attributes by identifying those variables that distinguish users and nonusers. The results based on a survey of 630 retail banking consumers from two separate Australian banking institutions suggest that successful marketing strategies must focus on the most important attributes of ATMs as well as identify

different user groups and develop strategies to maximise their patronage.

III. SIGNIFICANCE OF ICT IN THE BANKING SECTORS

ICT revolution has distorted the conventional banking business model by making it possible for banks to break their comfort zones and value creation chain so as to allow customer service delivery to be separated into different businesses. Thus, for example, primarily Internet banks distribute insurance and securities as well as banking products, but not all the products they distribute are produced by their group (Delgado and Nieto, 2004). However, the main economic argument for the diffusion of adopting the Internet as a delivery channel is based on the expected reduction in overhead expenses made possible by reducing and ultimately eliminating physical branches and their associated costs. This specifically applies to and relevant in the Spanish banking system, which is one of the most “over branched” in Europe. As stated by DeYoung (2005) and Delgado et al (2006), the Internet delivery channel may generate scale economies in excess of those available to traditional distribution channels. Besides them, Haq (2005) also states that banks exist because of their ability to achieve economies of scale in minimising asymmetry of information between savers and borrowers. The unit costs of Internet banking fall more rapidly than those of traditional banks as output increases as a result of balance sheet growth. In this context, DeYoung et al (2007) refer to the Internet banking as a “process of innovation that functions mainly as a substitute for physical branches for delivering banking services”. In the case of the Spanish banks, there is some undependable evidence that shows that the Internet distribution channel has lower unit transaction costs than the two other distribution channels (branch and telephone) for a given type of transaction (money transfer, mortgage loan, brokerage or demand deposits). Today, any bank that doesn't offer the very latest in information and communication technology is bound to lag behind; customers are used to the pace of the “digital” business world, and they expect a certain standard of compatibility between their online banking services and their laptops, home PC's, Macs, iPhones and so on. For this reason, banks have had to step up and move into the hottest new information and communication technology. Most customers love the flexibility of online banking and other modern services; however, there may be some who prefer to never expose their banking information online, as they fear hackers and other violations of their privacy. For people like this, old-fashioned banking may feel safer and more secure.

IV. SATELLITE BANKING

Satellite banking is also an upcoming technological innovation in the Indian banking industry, which is expected to help in solving the problem of weak terrestrial communication links in many parts of the country. The use of satellites for establishing connectivity between branches will help banks to reach rural and hilly areas in a better way, and offer better facilities, particularly in relation to electronic funds transfers. However, this involves very high costs to the banks. Hence, under the proposal made by RBI, it would be bearing a part of the leased rentals for satellite connectivity, if the banks use it for connecting the north eastern states and the under-banked districts. The major and upcoming channels of distribution in the banking industry are ATMs, internet banking, mobile and telephone banking and card based delivery systems. Automatic Teller Machines: ATMs were introduced to the Indian banking industry in the early 1990s initiated by foreign banks. Most foreign banks and some private sector players suffered from a serious handicap at that time- lack of a strong branch network. ATM technology was used as a means to partially overcome this handicap by reaching out to the customers at a lower initial and transaction costs and offering hassle free services. Since then, innovations in ATM technology have come a long way and customer receptiveness has also increased manifold. Public sector banks have also now entered the race for expansion of ATM networks. Development of ATM networks is not only leveraged for lowering the transaction costs, but also as an effective marketing channel resource. The other payment and settlement systems deployed were mostly aimed at small value repetitive transactions, largely for the retail transactions. The introduction of RTGS in 2004 was instrumental in the development of infrastructure for Systemically Important Payment Systems (SIPS). The payment system in India largely followed a deferred net settlement regime, which meant that the net amount was settled between banks on a deferred basis, which may lead to settlement risks. Launched by RBI, RTGS system is used only for large value transactions and retail transactions take an alternate channel of electronic funds transfer, a minimum threshold of one lakh rupees was prescribed for customer transactions under RTGS on January 1, 2007.

E-Banking: Technology has created e-banking or electronic banking. E- Banking can be defined as the automated delivery of new and traditional banking services and products directly to customers through electronic, interactive communication channels. Technology has affected and changed banking with the many benefits and convenience e-banking has created. It includes a system that enables bank customers to access

accounts, transact business or obtain information on financial products and services. Customers can now quickly complete transactions such as 5-10 minute deposits/withdrawals to 30-60 secs, online checking accounts, online transfers and many e-banking transactions. The accessibility of e-banking has been possible due to the technological advancement in laptops or personal computers, kiosks, Touch Tone phones, personal digital assistants (PDA) and automated teller machines (ATM). According to industry analysts (BNET.com), electronic banking provides a variety of attractive possibilities for remote account access, including: Availability of inquiry and transaction services around the clock; worldwide connectivity; Easy access to transaction data, both recent and historical; and "Direct customer control of international movement of funds without intermediation of financial institutions in customer's jurisdiction. E-Banking is becoming increasingly popular among retail banking customers. E-Banking has helped in cutting costs by providing cheaper and faster ways of delivering products to customers. It also helps the customer to choose the time, place and method by which he wants to use the services and gives effect to multichannel delivery of service by the bank. This E-Banking is driven by twin engines of "customer-pull" and Bank-push".

V. IMPACT OF TECHNOLOGY ON BANKING SECTOR

Technology has influenced all aspects of banking activities including storage, processing, and collection of information. There are a few areas in banking that have been seriously influenced or impacted which includes; Tracking lending worthiness (Credit Scores): Technology has created or led to the creation of the credit bureau. The system mathematically tracks customer's payment records to provide data which help banks make decisions on the amount and who they should borrow money from. The advanced technology available has developed a scalable and resilient credit bureau platform that enables banks to track customer's necessary information. Technology has enabled software programs which have provided banks with input file preparation tools, validation tools and data entry tools. Collecting data allows banks to deliver credit reporting solutions in the form of credit reports, customer credit activity monitoring, fraud prevention systems and debtor tracking services. Today, the technologies in credit bureau infrastructures have enabled banks to collect, load, validate, store and disseminate both the positive and negative data as well as supplementary data. The development and management of the information technology for the Credit Bureau has provided control over all processes involving the credit bureau as it relates to the banking industry.

Some of the systems used in the credit bureau include; C++, C Sharp and Java, Oracle & SQL. All these programs contribute significantly to the whole operation of the credit bureau. The typical cycle and function of the credit bureau as it relates to banks includes; storing information-credit histories, observing fraudulent behaviour, previous enquiries, validating data and many more. Banks Contestability: Technology is affecting the degree of contestability in banks. Due to the advancement of technology, banks' superiority in information has deteriorated. New competitors have emerged and the many barriers provided by banks have been declining and security breach is more imminent today. Some financial products, services and commodities are becoming more transparent. Due to the lowered entry and deconstruction of some banks, contestability in banking is rising. The advancement of technology has influenced the methods banks use to deliver financial products to its customers. Technology has created alternative delivery mechanisms such as the internet, ATMs, and many others which all reduce the dependence on the network as a core delivery mechanism. Now, financial systems are substantially over-supplied with delivery systems through a duplication of networks which allows or encourages the banks to change their delivery strategy, rationalise their branch network strategy and develop a wide Variety of delivery options. Internet Banking: An internet bank can be defined as a bank that provides account balances and some transaction capabilities to retail customers over the World Wide Web. Technology has created internet banking, also called online banking. The creation of the internet through technology has lead to many banking transactions or activity options via the internet. Some of these activities includes; paying bills, 24 hour view of accounts, transferring money and many others. Customers access their banking information from a browser- software that runs the banking programs on the World Wide Web (www). Customers can personally and privately access their account information through the internet via a modem. Technology has allowed us to dial into the bank via the modem system which allows us to download data, and run programs that make us access a wide variety of banking information such as; account balances, number and types of banking transactions, bank statements, among others. On the downside, the internet has decreased operation and transaction at physical-brick and mortar banks as customers can basically conduct almost all the transactions possible in a real bank. Today, technology has helped create many banks which have no physical location or brick and mortar branches.

VI. CONCLUSION

Over the years, the banking sector in India has seen a number of changes. Most of the banks have begun to take an innovative approach towards banking with the objective of creating more value for customers and consequently, the banks. The banking industry which is the backbone of every economy is confronted with various challenges such as globalisation, deregulation, competition, significant high cost of installing ICT and maintenance. The usage of ICT can lead to lower costs, but the effect on profitability remains inconclusive, owing to the possibility of ICT effects that arise as a result of high demand of skilled workforce, issues of increasing demand to meet customer's expectation for customer service delivery, trustworthiness of the information system and competition in financial services. However, from the discussion whilst reviewing literature many researchers did not find ICT, for the delivery of customer service and profitable for a bank's financial performance. So there has been an arrow head among these findings on the perspective of profitability and customer service delivery. While, on the same vein, other researchers found ICT channel making profitable impact on the banks that are only internet start-ups than the conventional banks transforming into click and mortar. In addition, there are other studies that proclaim due to perceived security risk, lack of comfort with computer technology, either due to lack of awareness or age factor, and a host of other reasons that ICT did not appear to be significantly viable or accepted warmly or quickly by consumers. Other researchers also found that despite all these factors banks, themselves, have been unable to have provided efficient customer service delivery because of which the clients who were even ready to adopt this delivery channel did not turn up again to innovation, and banks couldn't successfully build the required contents of electronic banking environment for consumers. More research needs to be carrying out in different locations and different time frame may confirm or refute the previous findings by collecting the primary data to come up to a conclusion for the impact of ICT on customer service delivery and banks performance. It is quite evident from our study that enhancing ICT in the banking industry is a must in a rapidly changing market place, as the ICT revolution has set the stage for exceptional increase in financial activity across the globe.

- [4] Bhasin, T. M., E-Commerce in Indian banking, in Indian Banking Association Bulletin, Vol.
- Solanki, Virender Singh. Impact of ICT Banking on Customer Services: An empirical study. BVIMR Management Edge. 2011, Vol. 4 Issue 1, p50-56. 7p.
- [5] Bimil Jalan, "Strengthening Indian Banking and Finance-Progress and Prospects", The Bank Economist Conference, India, 2002.

REFERENCES

- [1] Arora, Kalpana, Indian banking managing transformation through IT, Indian Banking Association Bulletin ,Vol. 25(3), pp. 134-38. March 2006.
- [2] Avasthi, G. P. and Sharma , M, Information technology in banking : challenges for regulations, Prajnan Vol.29(4) ,pp. 17-22, 2004.
- [3] Janki, Unleashing Employee Productivity:needforaparadigmshift, Indian Banking Association Bulletin, Vol. 24(3) ,pp.7-9, 2003.

Big Data Analytics for E-Commerce: Amazon's Impactful Use of Big Data

Gurpreet Singh Kalsi

Assistant Professor: B.A.M. Khalsa College

Garhshankar

E-mail: gurpreet.kalsi@bamkc.edu.in

Abstract

Nowadays e-commerce is the latest universal method of doing business trading. E-Commerce organizations compose online marketing strategies based on real time data. This has guided to an ideal shift in E-Commerce, where data is recognized as a major asset to the firm in understanding certain requirements of customers, anticipate behavior, customize specific needs and contributing performance metrics to assess efficacy. A large volume of data will be captured in various aspects such as advertisement, information exchange, order and payment, delivery, customer support, and customer feedback when trading through the internet. The company's main threat is how to use the data to further improve the product. These companies are now creating a research department that conduct data analysis using different technologies and makes a good contribution to the profit of companies and helps the companies to make good decisions. Big data analytics seek to the arrangement of wider databases of sophisticated analytical techniques. Big data analytics is therefore regarding "big data" and "analytics" as well as how the two came together to generate business analytics. This paper examined the need for data analysis in e-commerce & how the uses of big data analytics is perceived as value creator that can guide E-Commerce companies attain competitive advantage.

Keywords: *Big data, E-Commerce, Amazon, Impactful use, BDA*

I. INTRODUCTION

Internet today, has become an essential component of business and its remarkable impact on organizations' arrangements is obvious. This leads to excellent growth in Indian E-Commerce. With the exponential growth in internet and online infrastructure in India, it was no surprise that the e-commerce market underwent a similar growth. Increasing rates of internet penetration, along with an increasing share of the organized sector within retail markets created the strong foundation that e-commerce requires.

For the E-Commerce organizations to grow and revive, a potential way to create business value is by using big data techniques which can be categorized in personalization, dynamic pricing, customer service, forecasting customer behavior, supply chain

viability, and managing frauds. Addition to this, a firm with strong customer orientation is therefore considered to outpace its competitors since it better recognizes customer needs, regulates products and services, and therefore meets customer needs. In the recent past, many organizations have moved to an online environment to achieve better growth, increase profits and know their customers in a better way with the help of public relations. E-Commerce firms are applying various online methods to enable online marketing strategies, strengthen customer involvement (using Web analytics tools), Social Media monitoring, Web Content Management systems and Customer Engagement Management.

According to Gartner Research, Big Data can have an impact on your firm's business in three ways. [1]

- a. Discover hidden patterns of data and offer meaningful insights.
- b. Improve your decisions, by stirring information for decision makers.
- c. Automate the business processes.

II. REVIEW OF LITERATURE

Lau and Chen (2016) describes the relationship between the five V's of big data and ecommerce. For future research direction the authors give the following suggestions • Automatic dynamic pricing mechanism for products. • Location based analytics for different life styles • Need a recommender that gives automatic promotional marketing. • Need to develop a real time product monitoring tool for handling mass orders. [2] Cosma and Acampora (2016) Developed computational intelligence predictor for analyzing customer reviews. This architecture framework is used to eliminate the noise and imprecise information in the data set received from customer review. The framework contains four modules namely the Natural Language Processing module (NLP) used to process the input data, Input Selection module removes the noise from the data, The existing methods are recording customer reviews without applying any validation mechanism. [3]

III. OBJECTIVES

Following are the objectives of the paper:

- To recognize the need for big data for companies.
- To analyze how e-commerce companies use big data?
- To analyze impactful use of big data by Amazon.

IV. METHODOLOGY

The proposed article uses the data review approach to collect data and information. All the data and information retrieved here are secondary data available on websites, research papers, books, journals and magazines.

V. NEED OF BIG DATA FOR COMPANIES

Big Data refers to massive amount of data produced by different sources like social media platforms, web logs, sensors, IoT devices and many more. It can be either structured (like tables in DBMS), semi-structured (like XML files), or unstructured (like audios, videos, images). Traditional database management systems are unable to manage this huge amount of data. Big Data helps companies to generate valuable understanding. The history of big data began many years before the present buzz about Big Data. Seventy years ago the first attempt to quantify the growth rate of data in the terms of volume of data came across. That has popularly been known as the “information explosion”. Many firms use big data methods such as Apache Hadoop, Spark, Hive, Pig, etc. to handle big data and gain insights from it.

Big Data initiatives were rated as “extremely important” to 93% of companies. Leveraging a Big Data analytics solution helps organizations to unlock the strategic values and take full advantage of their assets. It helps organizations:

- To understand Where, When and Why their customers buy
- Protect the company's client base with improved loyalty programs
- Provide targeted promotional information

- Seizing cross-selling and upselling opportunities
- Improve inefficiencies in the company's supply chain
- Optimize Workforce planning and operations
- Make companies more innovative and competitive
- It helps companies to discover new sources of revenue
- Predict market trends
- Predict future needs

VI. HOW E-COMMERCE COMPANIES USE BIG DATA?

Big Data collaboration helps E-Commerce Companies to take more benefits. Traffic behavior is monitored. Frequently visited online ads and data searched on search engines can be the source for monitoring traffic behavior. Users clicks on the links (links in facebook, tweeter etc.), their frequency in relation to these clicks, time spent on a particular page & information stored in the form of cookies in users computer is vouched for data collection.

Customer's preferences and products most viewed by them helps in marketing whether they are buying anything or not. Imprints of the data left by the customer on the internet in the forms of text, images etc. contains the customers' behavior on the internet and is being analyzed by the data analysts. This analysis helps the e-commerce companies to build their strategic plans, product pricing, sales, preferences for future.

Big Data is used to gather information from peers on various products to convey precise location, expected time of product delivery to their customers. It also helps E-commerce companies to predict and identify business events before it occurs. Figure-1: There were 658 million Internet users in India in January 2022. It is revealed from [Kepios](#) analysis that Internet users in India has been increased by 34 million between 2021 and 2022.

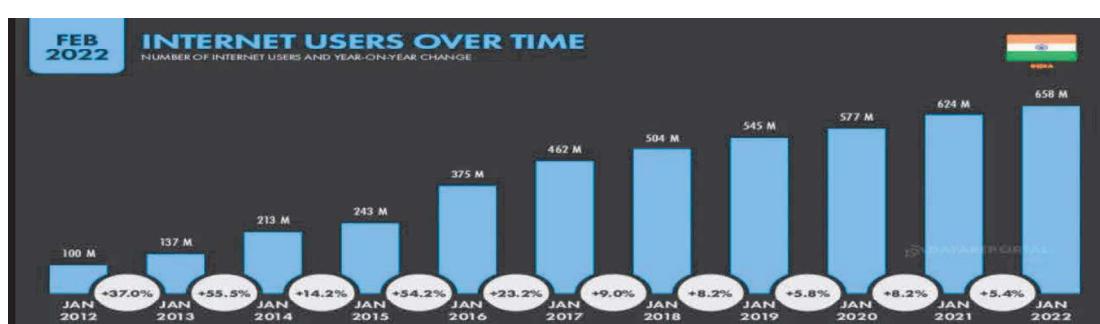


Figure 1: Internet Users Over Time

Source: [datareportal.com](#)

VII. WAYS IN WHICH AMAZON USES BIG DATA FOR MONITORING

According to Werner Vogels, C.T.O, Amazon, in the changing era of technology changes are reflected everywhere and in everything. So businesses are no exception either. The Giant in E-commerce, 'Amazon' & Reason being a giant is, wise use of technology.

"In retail, while things like the size of the catalogue, advertising and other stuff might play a role in success, at Amazon, I think success is largely technology driven"

Werner Vogels, C.T.O, Amazon

Amazon collects the data from every customer and analyze customer's buying behavior, which in return help them in target marketing, customer satisfaction & to build customer base. This data helps in building trust & loyalty among the customers.

Amazon's data collection sources:

A. Personalized Search:

Amazon follows behavioral analysis system and recommends its customers the products as per their previously purchased, liked, shortlisted, in the cart products. Customers review the purchase which in result helps the company to build recommendations. This information is further used for additional recommendations suitable with the products. It is also used for building product collaborations and then this information is used as recommendations for other customers with similar searches. For instance, a customer adds a T-shirt of a particular brand in the shopping cart, T-shirts of other brands are recommended for purchase option and further this search is used for similar products

to be bought together with it; for example Jeans, Trousers, Jackets, etc. Amazon's data collection system empowers it to uses suggestions to encourage impulsive purchases & enhances shopping experience. It helps the company to generate 35% of its annual sales.

B. Voice Recognition System / Assistance

'Alexa! Add tomato catchup in my shopping cart.' It has never been so easy to add things to your shopping cart. Amazon's Alexa, a virtual assistant, made it convenient for the user to access its shopping cart just by asking. It is used for various purpose ranging from weather updates, news, adding products in shopping cart. All this can be done just at a voice command. However, it is not even noticeable for a user that these voice commands are being uploaded with Amazon's server for big data. Amazon clarifies that this data from diverse group of customers helps 'Alexa' to

build its library and processing accurately. For some customers it may be a privacy issue and it cannot be neglected as well. So for them there is option to delete their recording through the application.

C. Placing orders - On fingertips

In this competitive environment, data shows that the customers shift quickly from one application to another if they feel inconvenience while interacting with a particular e-commerce website/app or delays in orders processing. Amazon has One-Click Ordering system. Customer's details and payment method get updated while placing first order. Order placing is one click away with this data stored.

D. Kindle – as social networking& discussion forum

Social networking service with Kindle Amazon has integrated the. With this service users can highlight words and put notes in a book to share & discuss with others. It also empowers the company to collect data about what readers are interested in. Further the data is used to recommend e-books to customers which tends to enhance their reading experience.

E. Amazon Web Services (AWS):

Amazon Web Services (AWS) is a subsidiary of Amazon. It provides cloud computing services and products on demand. It functions on 'pay-as-you-go' model which includes mobile development networking, developer tools, remote computing, security, servers etc. AWS provides solutions for data integration tools, storage, processing & visualization.

F. Anticipatory Shipping:

Amazon's developed and patented 'Anticipatory Shipping' where orders can be anticipated. It utilizes large information for anticipating that is the product customer is probably going to buy, estimated time of delivery and place of delivery. This helps amazon into 'predictive shopping', what a customer is probably going to buy. Companies are reducing distance between the customer and warehouses to reduce shipping time.

VIII. DISCUSSION& RECOMMENDATION

In order to get more of business outcome data analysis is to be made with the integration of dataset and technology to process the same. E-commerce businesses understand the elements like marketing, order processing, shipping, advertisement etc. for forecasting and improvements in business structures. Data analytics provides aid which helps

the companies in decision making. Now companies tend to simplify their processes to enhance customer base for which they use the database by integrating and aggregating available data. All this is done to predict the future for the company.

IX. CONCLUSION:

Data collection processing starts as soon as a person enters to a website. We see, when we search for a product, it starts recommending similar products or ads will appear in the webpages. Big data analysis helps e-commerce sector in providing personalized experience to customers, customer analysis, predictive shopping and in many more ways. Forecasting business scenarios and analyzing present and future aspects helps the e-businesses sustained growth.

REFERENCES

- [1] Gartner, Big Data Means Big Business Edited by Douglas Laney, <https://im.ft-static.com/content/images/e91a32d0-2bac-11e3-bfe2-00144feab7de.pdf>.
- [2] Raymond Y.K. Lau, J. Leon Zhao, Guoqing Chen, Xunhua Guo : Big data commerce. Information & Management à 2016 Elsevier B.V. All rights reserved.
- [3] Georgina Cosma, Giovanni Acampora.: A computational intelligence approach to efficiently predicting review ratings in e-commerce, Applied Soft Computing44 (2016) 153–162 © 2016 Elsevier B.V. All rights reserved.
- [4] Sinha, V., Subramanian, K. S., Bhattacharya, S., & Chaudhuri, K. (2012). The Contemporary Framework on Social Media Analytics as an Emerging Tool for Behavior Informatics, HR Analytics and Business Process Management, 17(2), 65–84.
- [5] Hurtado, Paula & Dorneles, Carina & Frazzon, Enzo. (2019). Big Data application for E-commerce's Logistics: A research assessment and conceptual model. IFAC-Papers OnLine. 52. 838-843. 10.1016/j.ifacol.2019.11.234.
- [6] Renata Thiebaut, "AI Revolution: How Data Can Identify and Shape Consumer Behavior in Ecommerce" In Entrepreneurship and Development in the 21st Century. Published online: 11 Apr 2019; 191-229.
- [7] Castro-Schez Raúl Miguel David Vallejo Vanesa Herrera, (2010)."A multi-agent architecture to support B2C e-Marketplaces: the e-ZOCO case study", Internet Research, Vol. 20 Iss 3 pp. 255 – 275
- [8] Sachin S. Deshmukh (2012) Big Data - an opportunity and challenge for E-commerce International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013 79 ISSN 2229-5518
- [9] Avinash B.M.1, Akarsha B.M.2Big Data Analytics for E-Commerce- Its Impact on Value Creation International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 6, Issue 12, December 2017
- [10] Ritesh Singh, Nisha Verma and Manali Gupta (2020) Role of E-Commerce in Big Data International Journal of Advanced Research in Engineering and Technology (IJARET) Volume 11, Issue 12, December 2020, pp. 1770-1777, Article ID: IJARET_11_12_166 ISSN Print: 0976-6480 and ISSN Online: 0976-6499
- [11] Alain Chong Boying Li EWT Ngai Eugene Ch'ng Filbert Lee, (2016),"Predicting online product sales via online reviews, sentiments, and promotion strategies: a big data architecture and neural network approach", International Journal of Operations & Production Management, Vol. 36 Iss 4 pp. –
- [12] Sarah S. Alrumiah and Mohammed Hadwan (Member, IEEE) (2021) Implementing Big Data Analytics in E-Commerce: Vendor and Customer View
- [13] Lai-Ying Leong, Noor Ismawati Jaafar, AININ SULAIMAN, "Understanding Impulse Purchase in Facebook Commerce: Does Big Five Matter?", Internet Research, <https://doi.org/10.1108/IntR-04-2016-0107>
- [14] Arief Rahman "Toward a Comprehensive Conceptualization of Digital Divide and its Impact on E-Government System Success" In E-Services Adoption: Processes by Firms in Developing Nations. Published online: 07 Oct 2015; 291-488.
- [15] Vikranth, K., & Krishna Prasad, K. (2020). The Sustained Improvements in E-commerce Business through Big Data and Data Analytics of Wal-Mart Company. International Journal of Case Studies in Business, IT, and Education (IJCSE), 4(2), 28-39.
- [16]https://www.researchgate.net/publication/264555968_Big_Data_Analytics_A_Literature_Review_Paper
- [17]<https://www.journals.elsevier.com/big-data-research/most-downloaded-articles>
- [18]<https://journalofbigdata.springeropen.com/articles/10.186/s40537-015-00>

E-Marketplaces Minimize Buyers' Search Costs: A Critical Review of Reduced-Price Hypothesis

Jasminder Kaur

Department of Commerce and Economics

B.A.M. Khalsa College, Garhshankar

Affiliated to Panjab University Chandigarh

Jasminder2597@gmail.com

Abstract:

This study typifies a critical review of reduced-price hypothesis advanced by Bakos (1997) in his ground-breaking research in the domain of e-marketplaces. After the detailed inquest of literature pertinent to the relevant theme, it is found that the postulation 'substantial plunge in the buyers' search costs spurred by e-commerce' is never empirically tested and buttressed in the Indian setting that entails a succinct and expository review first. Hence, the current study aims to provide an analytical review of erstwhile literature and offer new conclusions on reduced price hypothesis which would form the premise of formulating new models for the future research. For research to make headway and keep pace with these ambitions, significant evaluative involvement in conventional deliberations on electronic market adoptions by the sellers is essential along with considerable theoretical and methodological refinement like developing linear regression models assessing the impact of new market intermediaries in benefitting both sellers and buyers, thus ensuring win-win market. In nutshell this article suggests the existence of certain number of economic factors other than the search costs borne by consumers which defy the assertion that electronic market places make it tough for sellers to sustain whopping prices and upholds that the efficiency of e-commerce gets bucked by upsurge in the product cost.

Keywords: Reduced price hypothesis, e-marketplaces, traditional markets, buyers' search costs, e-commerce.

I. INTRODUCTION

A. Search Cost

Search cost borne by the consumers signifies money, laborious endeavours and time spent and devoted while trying to locate the correct sellers dealing in desired products and engages in value transactions. Search costs comprise the opportunity costs of efforts, energy and precious hours spent on searching products of utility coupled with possible explicit costs of money and other scarce resources with alternative applications absorbed by search. Hence, its kinds transaction cost induced before the occurrence of bankable transactions [5, 8, 10]. Because the crucial facts consumers or the buyers possess about the potential sellers or dealers are oftentimes imperfect and deficient, search cost is an obvious episode. It entails risks too; while on an outlook for the right article to buy, consumers always bear inevitable risk of flunking in getting what they are chasing. In congruence with the postulations of economic theory, a rational consumer essentially incurs

search costs up to the point of equivalence between the marginal cost of their quest and marginal advantage of successful probe considering the odds of ultimate success of their search [7, 15, 20]. In further elucidation, buyers conduct research on a product/service for consumption and bear pertinent costs in the form of expenses on commuting between the stores to survey several options, buying requisite data (of features, prices) or soliciting experts' advice on purchase. This is the energy and time that other more important activities could have imbibed, alternatively termed as opportunity costs [16].

Growing popularity of electronic marketplaces is gradually rendering the conventional form of commerce and brick and mortar stores unmarketable. With this spike in its popularity researchers predicted notable plummet in the market prices of products. Such assertion is styled as reduced-price hypothesis [16, 36, 70]. Consumers in the market full of intermediaries, bear search expenses while transacting to get information on prices and product portfolios of the sellers. Exorbitant search costs empower sellers to sustain prices considerably exceeding their marginal costs and trigger allocation inefficiencies in market deals. However, with the advent of e-commerce buyers incur drastically lower search costs while obtaining information regarding product pricing and offerings in the markets increasingly gaining e-platform. Diminishing search costs facilitate the buyers to browse through large number of product offerings which deter sellers in keeping the prices steeper. The reduced-price hypothesis postulates that consumers benefit from the abated product prices in consequence of thriving competition and rivalry among sellers in e-marketplaces [16, 82, 85]. In furtherance of the relevant subject, information systems are reckoned as intermediaries linking potential buyers and sellers in the vertical value chain, therefore making up an e-marketplace. Economic theory establishes that decline in search costs has pivotal role to play in ascertaining the implications of electronic systems for market effectiveness, productivity, and combative stance. Several economic models of search are formulated in the past and documented in the erstwhile research articles which examine the effect of reduced search costs in the markets on product prices, profit margins of sellers and welfare of the consumers. Such plunge kindles efficiency gains emanated from low intermediation costs. This reduction also augments productivity of inter-organizational engagements/dealings and sways the buyers' and sellers' market competency [55, 67, 72, 74, 78].

B. Electronic Marketplaces

An e-marketplace synonymously termed as e-market system is defined as an inter-organizational system of eminent information which facilitates an effective exchange of essential facts about product offerings and prices between prospective buyers and sellers [22, 34, 37, 38, 39, 43]. The firm that runs the system is called the intermediary which could be a consumer or seller, a self-reliant third party or an association of many companies. Such system extensively reduces customers' expenses incurred in acquiring fact-driven information about products/services they are desirous of buying or availing from alternative suppliers or sellers. Besides, it also brings down sellers' costs of disseminating the requisite information to their potential customers [11, 13, 14, 17, 31, 33, 35]. Owing to the existence of inter-organizational information mechanism, monopoly power of sellers and monopsony capacity of consumers in a vertical market gets significantly struck. With more firms joining the system, the advantages leveraged by e-market participants compound substantially [1, 2, 4, 6, 9, 21, 23, 24, 25]. But e-marketplaces entail inordinate switching costs for their participants for these systems warrant exorbitant investments in hardware, software, training and development of workforce and other organizational radical strategies. They necessitate colossal capital expenditure and generate prodigious economies of scope and scale. Mostly, intermediaries court staggeringly high development and maintenance costs which pay off later in the form of low incremental costs per transaction until the complete utilization of capacity that provokes massive economies of scale [3, 12, 18, 19, 28, 30]. Further, economies of scope are induced as a result of transferability of technical know-how and expertise obtained during the development of one system. In addition to capital investments, prospective participants of e-markets witness high degree of uncertainty with respect to privileges in becoming part of such system. Such uncertainty influences strategic posture of sellers, buyers and middlemen by stimulating them to embrace "wait and see" strategy in line of which they would hold up the introduction of or joining the e-system. They refrain in the hope of gleaning from the experiences of other companies competing in the e-marketplaces. This economic attribute of e-markets casts doubts on reduced-price hypothesis that demands comprehensive and deep inquest of literature in the sphere of e-commerce [26, 27, 32, 40, 50]. This article aims to produce interpretative appraisal of the data corralled from existing research and provide insights into the potential research domains that seek exploratory and empirical studies. Present research also draws novel inferences and deductions that would form grounds for future research.

Rest of the paper is organized in three sections. Section 2 documents detailed review of former literature and section 3 is devoted to its critical assessment. In section 4 unconventional conclusions are succinctly presented.

II. LITERATURE REVIEW

The conventional economic view theorizes that with the advancements in information and communication technology buyers' search expenses fall considerably further intriguing abatement in prices and assuring efficiency in the

markets. This implicates those sellers do not make economic benefits [41, 45, 48, 49, 60, 62]. Although information and communication technologies seem to have decreased the prices and resulted in consumer surplus, but the companies which formulate differentiation strategies and have created brand equity possess the potential to command the prices higher than usual. Price dispersion is discovered in online markets suggesting the existence of market inefficiencies. Price dispersion is a term that signifies charging different prices for identical product offerings at the same time [56, 75, 76, 83, 84, 86]. Market inefficiencies are instigated by intermediation, entry barriers and information uncertainty. Had consumers bypassed the intermediaries and accessed plethora of market suppliers, they could have bargained effectively, thus turning the markets efficient. Hence efficient markets are devoid of price dispersion [42, 47]. Previous literature cites two most eminent grounds of the assertion that e-marketplaces are more efficient than the brick-and-mortar stores viz. pervasive availability of information and notable decline in search costs for acquiring such information [51, 63, 64]. Consumers desirous of acquiring information on offerings and other product attributes like brand reputation, colour, texture, and most importantly after-sale services, must bear considerable amount of search expenses. Economic literature establishes that more the search costs borne while obtaining information, the higher the price dispersion and the more unlearned the buyers highly dispersed would be the prices, signalling the prevalence of inefficiencies [53, 57, 58, 77]. Subject to the association between easy access to information and market efficiencies, information search theory provides succinct insights into the impact of internet on prices. It proposes that price elasticity relies on product utility and benefits and information search costs. Higher anticipated benefits of information search reduce the focus of buyers on prices and result in lower price elasticity. On the other hand, large amount of search expenses borne for acquiring price information, lead to lower price elasticity which empower the sellers to set prices higher than their marginal costs leading to inefficiencies in the market [44, 46, 66, 69].

Search costs are ascribed to ascertaining an ideal product fit where products are imperfect substitutes. It is suggested that e-markets are found favourable by the buyers due to reduced search efforts [61, 68]. Costs in e-markets cannot be deemed as trivial even if information is easily accessible from the internet relative to traditional market. Search engines entail sizeable time span and efforts of consumers. It is documented that buyers pay increased prices as search becomes dearer and thus; expensive search bespeaks higher price dispersion. Consumers' search expenses are also associated with opportunistic stance of sellers reflected in their pricing. It is advocated that vendors' opportunistic pricing stance with a view to reap large profit margins from unlearned buyers, triggers price variability [52, 54, 81]. Present research proffers some peculiar arguments which challenge the low search-cost perspective and imply that other variables like brand recognition, trust relationships, customer loyalty make consumers willingly pay higher prices for online products.

Classic Bertrand model on price combats reeks of zero search expenditure and fully informed buyers which confirms equilibrium price equal to marginal costs. Bucking the above two mentioned assumptions is explained by price dispersion. There are three dimensions of non-zero search costs owing to incomplete information: i) buyers bear search costs in the form of time and efforts because of lack of information like pricing and vendors' location; ii) search costs to glean information on the properties and attributes of products; and iii) cost imputed to ascertaining a product mix or fit in the event of imperfect substitutes available in the market [71, 73, 81]. Former research suggests that from a buyer's viewpoint, e-markets favour consumers through steep fall in its search costs. The search-cost stance works out the reasons for price dispersion, emphasizing the variables like ill-informed buyers and resultant search costs and the replacements for excessive search expenditure such as trust, brand name and loyalty. Another perspective for discrepancies observed in the pricing structure is the differentiated services perspective. This perspective compounds the product value. It stresses over quality of service defined in terms of selection assistance, ease of placing order, timely and prompt delivery and order tracking. The surcharge which consumers pay happily becomes the cause of price dispersion. Third is the market-attributes perspective which puts forward the reasons behind price dispersion by laying clear focus on market parameters viz. mean price, number of rival firms and time. Spurt in the number of competing firms due to ease of entry or marginal entry barriers in e-marketplaces is affirmed to hit price dispersion. It is established that mean price of specific product impacts price dispersion and price increases are robust in case of superior products than low quality ones [62, 65, 79, 80]. Before this section is closed, it is necessary to take a note that reduced-price hypothesis concerns the gaps in information quality between buyers and sellers termed as information asymmetry forming the grounds for search-cost perspective. Nature of the information and search for authentic information hold prominence in the price dispersion of e-markets.

III. ARGUMENTS AGAINST REDUCED-PRICE HYPOTHESIS

Several critical points can be mentioned here after the comprehensive and thorough review of the past seminal research works which implicitly discredit the proposition by [16]. E-Markets or e-commerce firms are vulnerable to failures because consumers are unable to appraise the quality of products offered through e-platforms. Additionally, dealers operating in e-marketplaces do not encounter stringent inspection of stock which makes buyers doubt the quality embedded in their offerings. These loopholes proffer that the consumers face transaction risks in the event of becoming participants of e-markets. Unlike conventional markets, e-marketplaces impede consumers in conceiving personal evaluation of offerings. They have to form their decisions solely on the basis of electronic information without carrying out physical scrutiny and therefore confront the likelihood of flawed and incomplete information furnished by sellers. E-markets are entrenched

with asymmetric information for there are the odds that sellers keep the products' shortcomings and flaws under the veil. If e-markets would not guarantee standard of products customers would turn reticent in adopting e-trading services.

Aforesaid points indicate the stumbling blocks in e-marketplaces which further indicate the consumers' higher propensity to pay for same products in traditional markets because they might be unwilling to compromise on the quality front. With the dealers competing on e-platforms becoming more cautious of possible loss of customers, e-trading firms have started establishing rigorous institutional policies with ferocious inspection procedures in a bid to decrease the plausible transactional risks spawned by unreliable product quality. As a result, products offered online are of superlative quality in relation to those sold in brick-and-mortar stores which ultimately warrant higher prices.

Information uncertainty in e-trading cause rise in the buyers' search cost. In the event of inadequate information, onerous efforts are made to gain more information by the buyers. Additional information is obtained through increased search that entails cost too. Hence, information uncertainty indicates colossal search cost. Another challenge posed by e-marketplaces is that the buyers are prevented from inspecting products, their quality and standard in person. This augments the buyers' dependence on feedback from their acquaintances so as to decrease quality uncertainty. Such word of mouth serves as a product quality index and leads to firm and confident buying decisions. But antithetical feedback likely increases information ambiguity. Consumers unwinding such equivocality further incur search costs. Therefore, both absence of information and dubious knowledge ingrained in e-markets result in higher search costs. Arguments presented here need to be substantiated by developing some robust models in empirical analysis that can assess if online prices surpass what are offered in offline stores.

IV. SUGGESTIONS FOR FUTURE RESEARCH

Current study contributes significantly to the pricing behaviour in the electronic markets through an evaluative review of reduced-price hypothesis. To the best of author's knowledge this is the first work that puts forth interpretative statements on the postulations propounded by [16] and new conclusions. Future research must be pursued in studying the variables that induce higher prices in e-marketplaces than demanded by traditional markets. It is also suggested to do comparative analysis of efficiency of online suppliers and retailers and the traditional ones. However comparative studies are strongly recommended in the context of homogeneous product for such empirical evidence would generate coherent idea about the productivity of both the channels. Discrepancies in the price dispersion, posted prices of products, frequency of price changes and magnitude of such changes between the offline and online market places are suggested to be gauged deploying robust and sophisticated research methods. In future, researchers must study the differences in the pricing strategies of vendors for homogeneous products. They could also

empirically analyse pricing strategies of single supplier across several products. With this it would be determined if the prices vary across different products subject to information overload and ambiguity.

As suggested in the previous studies that buyers do not always purchase online, explanations buttressing this discrepancy need to be advanced. For this, vendors can be compared on the parameter of their brand equity. For instance, customers mostly shop at Amazon and Flipkart without giving a thought to prices offered by other stores including the traditional. This is Amazon's and Flipkart's brand equity. Such illustration upholds that e-commerce has not essentially standardized the companies as theorized. To further make worthy contributions to this realm, more studies are required to be pursued to inquire about the attributes of buyers which can sway the online prices. In addition to this, higher prices discovered on the e-platforms seem to be ascribed to exclusivity of the offerings. Therefore, empirical analysis of relationship between products attributes and price effect is required. All the potential studies suggested above would be constructive in apprehending e-marketplaces and their association with price dispersion and information facets.

V. IMPLICATIONS AND CONCLUSION

With the advent of e-commerce, consumers are swayed to believe that internet would be driven by price. Because buyers would glean adequate knowledge of the prices in the markets, they would solicit the lowest prices. This would result in drastic decline in prices and eliminate myriad of sweltering factors that generate opportunities for firms to charge premium viz. dearth of time, information, and imperfect product comparison. Consequently, prices of all product offerings would diminish to the equilibrium point as postulated by theory of perfect competition. Since the former researchers and consumers have discovered it, ironically this has not validated to be the case. Brands and their equity still hold significance and represent splendid quality and credibility. As consume manifest during their purchase, price is not the sole determining factor impacting the purchase irrespective of the medium (electronic or traditional) that enables the transaction and availability of or access to relevant information.

This article documenting evaluative review furnishes anomalous inferences that defy the assertions established in the erstwhile research in favour of such postulation. It's argued that e-marketplaces do not necessarily offer lower prices due to the effects of information uncertainty and equivocation. These information aspects, progressively prevalent in the contemporary buoyant e-marketplaces, facilitate the sellers to keep prices dearer and to impair efficiency of these markets.

Present study implicates for researchers, consumers, and vendors that internet does not drive the prices down. Instead, e-marketplaces command much higher prices than presumed due to ferocious search demeanour of consumers. Excessive information itself possibly restrains

consumers' decision and pushes them to adopt different pricing strategies. Buyers might get dependent on advice of experts to curb certain issues spawned by information attributes. This study issues a caveat for the consumers that they must not blindly rely on e-marketplaces created by e-commerce.

ACKNOWLEDGEMENT

I am immensely grateful to sages Dr. Baljeet Singh, Principal and Prof. Kanwar Kulwant Singh, Head of Department – Department of Commerce and Economics, B.A.M. Khalsa College, Garshankar, affiliated to Panjab University, Chandigarh for giving me such a progressive opportunity to use flair in making significant contributions to the research community.

REFERENCES

- [1] G. J. Stigler, "The economics of information" *Journal of political economy*, vol. 69, no. 3, pp. 213-225, June 1961.
- [2] R. W. Rosenthal, "A model in which an increase in the number of sellers leads to a higher price", *Econometrica: Journal of the Econometric Society*, vol. 48, no. 6, pp. 1575-1579, September 1980.
- [3] N. Capon, and D. Kuhn, "Can consumers calculate best buys?" *Journal of Consumer Research*, vol. 8, no. 4, pp. 449-453, March 1982.
- [4] G. J. Tellis, and G. J. Gaeth, "Best value, price-seeking, and price aversion: The impact of information and learning on consumer choices", *Journal of marketing*, vol. 54, no. 2, pp. 34-45, April 1990.
- [5] P. E. Gabriel, and J. R. Marsden, "An examination of market efficiency in British racetrack betting", *Journal of Political Economy*, vol. 98, no. 4, pp. 874-885, August 1990.
- [6] J. Y. Bakos, "A strategic analysis of electronic marketplaces", *MIS quarterly*, vol. 15, no. 3, pp. 295-310, September 1991.
- [7] B. P. Pashigian, and B. Bowen, "Why are products sold on sale? Explanations of pricing regularities", *The Quarterly Journal of Economics*, vol. 106, no. 4, pp. 1015-1038, November 1991.
- [8] R. Kohli, and V. Mahajan, "A reservation-price model for optimal pricing of multi-attribute products in conjoint analysis", *Journal of Marketing Research*, vol. 28, no. 3, pp. 374-354, August 1991.
- [9] L. Samuelson, and J. Zhang, "Search costs and prices", *Economics Letters*, vol. 38, no. 1, pp. 55-60, January 1992.
- [10] H. J. Miller, "Consumer search and retail analysis" *Journal of Retailing*, vol. 69, no. 2, pp. 160-192, Summer 1993.
- [11] J. Robert, and D. O. Stahl, "Informative price advertising in a sequential search model", *Econometrica: Journal of the Econometric Society*, vol. 61, no. 3, pp. 657-686, May 1993.
- [12] Z. Vladimir, "Electronic commerce: structures and issues", *International journal of electronic commerce*, vol. 1, no. 1, pp. 3-23, 1996.
- [13] F. Zettelmeyer, "The strategic use of consumer search cost", *Unpublished Ph. D. Thesis, Massachusetts Institute of Technology*.
- [14] T. Z. Chang, and A. R. Wildt, "Impact of product information on the use of price as a quality cue", *Psychology & Marketing*, vol. 13, no. 1, pp. 55-75, January 1996.
- [15] J. E. Urbany, P. R. Dickson, and R. Kalapurakal, "Price search in the retail grocery market", *Journal of Marketing*, vol. 60, no. 2, pp. 91-104, April 1996.
- [16] J. Y. Bakos, "Reducing buyer search costs: Implications for electronic marketplaces", *Management science*, vol. 43, no. 12, pp. 1676-1692, December 1997.
- [17] J. Alba, J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, and S. Wood, "Interactive home shopping: consumer, retailer, and manufacturer incentives to participate in electronic market places", *Journal of Marketing*, vol. 61, no. 3, pp. 38-53, July 1997.
- [18] K. Burdett, and M. G. Coles, "Steady state price distributions in a noisy search equilibrium", *Journal of Economic Theory*, vol. 72, no. 1, pp. 1-32, January 1997.
- [19] H. G. Lee, "Do electronic marketplaces lower the price of goods?"

- [20] *Communications of the ACM*, vol. 41, no. 1, pp. 73-80, January 1998.
- [21] J. P. Bailey, "Intermediation and electronic markets: Aggregation and pricing in internet commerce", *Unpublished Ph. D. Thesis, Massachusetts Institute of Technology*.
- [22] E. K. Clemons, I. H. Hann, and L. M. Hitt, "The nature of competition in electronic markets: An empirical investigation of online travel agent offerings", *The Wharton School of the University of Pennsylvania*, pp. 1-38, June 1998.
- [23] Y. Bakos, "The emerging role of electronic marketplaces on the Internet", *Communications of the ACM*, vol. 41, no. 8, pp. 35-42, August 1998.
- [24] C. Steinfield, A. Mahler, and J. Bauer, "Electronic commerce and the local merchant", *Electronic markets*, vol. 9, no. 1-2, pp. 51-57, 1999.
- [25] G. E. Smith, M. P. Venkatraman, and R. R. Dholaki, "Diagnosing the search cost effect: Waiting time and the moderating impact of prior category knowledge", *Journal of Economic Psychology*, vol. 20, no. 3, pp. 285-314, June 1999.
- [26] H. G. Lee, J. C. Westland, and S. Hong, "The impact of electronic marketplaces on product prices: An empirical study of Auction", *International Journal of Electronic Commerce*, vol. 4, no. 2, pp. 45-60, 1999.
- [27] A. Y. Hoque, and G. L. Lohse, "An information search cost perspective for designing interfaces for electronic commerce", *Journal of marketing research*, vol. 36, no. 3, pp. 387-394, August 1999.
- [28] M. A. Arnold, "Costly search, capacity constraints, and Bertrand equilibrium price dispersion", *International Economic Review*, vol. 41, no. 1, pp. 117-132, February 2000.
- [29] E. Turban, J. Lee, D. King, and H. M. Chung, *Electronic Commerce A Managerial Perspective*, Prentice-Hall, 2000.
- [30] Y. Bakos, H.C. Lucas, W. Oh, G. Simon, S. Viswanathan, and B. Weber, "The impact of electronic commerce on the retail brokerage industry" *Stern School*, 2000.
- [31] E. Brynjolfsson, and M. D. Smith, "Frictionless commerce? A comparison of Internet and conventional retailers", *Management science*, vol. 46, no. 4, pp. 563-585, April 2000.
- [32] J. G. Lynch Jr, and D. Ariely, "Wine online: Search costs affect competition on price, quality, and distribution", *Marketing science*, vol. 19, no. 1, pp. 83-103, February 2000.
- [33] S. S. Posavac, "Overbidding in value elicitation: When consumers report inflated reservation prices and what to do about it", *Journal of Consumer Psychology*, vol. 11, no. 2, pp. 87-97, January 2001.
- [34] D. W. Carlton, and J. A. Chevalier, "Free riding and sales strategies for the Internet", *The Journal of Industrial Economics*, vol. 49, no. 4, pp. 441-461, June 2001.
- [35] L. F. Pitt, P. Berthon, R. T. Watson, and M. Ewing, "Pricing strategy and the net", *Business Horizons*, vol. 44, no. 2, pp. 45, March 2001.
- [36] G. D. Bhatt, and A. F. Emdad, "An analysis of the virtual value chain in electronic commerce", *Logistics Information Management*, vol. 14, no. 1/2, pp. 78-85, March 2001.
- [37] W. Baker, M. Marn, and C. Zawada, "Price smarter on the net", *Harvard Business Review*, vol. 79, no. 2, pp. 122-127, February 2001.
- [38] N. Shin, "Strategies for competitive advantage in electronic commerce", *Journal of Electronic Commerce*, vol. 2, no. 4, pp. 164-171, 2001.
- [39] J. E. Harrington Jr, "Comment on "Reducing buyer search costs: Implications for electronic marketplaces"", *Management Science*, vol. 47, no. 12, pp. 1727-1732, December 2001.
- [40] P. Pereira, "Market power, cost reduction and consumer search", *International Journal of Industrial Organization*, vol. 19, no. 9, pp. 1457-1473, November 2001.
- [41] J. F. Rayport, and J. J. Sviokla, "Managing in the marketplace", *Harvard Business Review*, vol. 72, no. 6, pp. 141-150, 1994.
- [42] S. Borenstein, and G. Saloner, "Economics and electronic commerce", *Journal of Economic Perspectives*, vol. 15, no. 1, pp. 3-12, March 2001.
- [43] B. G. Silverman, M. Bachann, and K. Al-Akharas, "Implications of buyer decision theory for design of e-commerce websites", *International Journal of Human-Computer Studies*, vol. 55, no. 5, pp. 815-844, November 2001.
- [44] H. Riquelme, "An empirical review of price behaviour on the internet", *Electronic Markets*, vol. 11, no. 4, pp. 263-272, 2001.
- [45] M. Kung, K. B. Monroe, and J. L. Cox, "Pricing on the Internet", *Journal of Product & Brand Management*, vol. 11, no. 5, pp. 274-287, September 2002.
- [46] K. Baylis, and J. M. Perloff, "Price dispersion on the Internet: Good firms and bad firms", *Review of Industrial Organization*, vol. 21, no. 3, pp. 305-324, November 2002.
- [47] N. Singh, and S. Kundu, "Explaining the growth of e-commerce corporations (ECCs): An extension and application of the eclectic paradigm", *Journal of International Business Studies*, vol. 33, no. 4, pp. 679-697, December 2002.
- [48] S. W. Schmitz, and M. Latzer, "Competition in B2C e-Commerce: analytical issues and empirical evidence", *Electronic Markets*, vol. 12, no. 3, pp. 163-174, 2002.
- [49] C. Kocas, "Evolution of prices in electronic markets under diffusion of price-comparison shopping", *Journal of Management Information Systems*, vol. 19, no. 3, pp. 99-119, 2002.
- [50] X. Pan, B. T. Ratchford, and V. Shankar, "Can price dispersion in online markets be explained by differences in e-retailer service quality?", *Journal of the Academy of Marketing Science*, vol. 30, no. 4, pp. 433-445, October 2002.
- [51] P. Scholten, and S. A. Smith, "Price dispersion then and now: Evidence from retail and e-tail markets", In *The Economics of the Internet and E-commerce*. Emerald Group Publishing Limited, 2002.
- [52] J. R. Brown, and A. Goolsbee, "Does the Internet make markets more competitive? Evidence from the life insurance industry", *Journal of Political Economy*, vol. 110, no. 3, pp. 481-507, February 2002.
- [53] P. M. Lee, "Behavioural model of online purchasers in e-commerce environment", *Electronic Commerce Research*, vol. 2, no. 1, pp. 75-85, January 2002.
- [54] R. LaRose, and M. S. Eastin, "Is online buying out of control? Electronic commerce and consumer self-regulation", *Journal of Broadcasting & Electronic Media*, vol. 46, no. 4, pp. 549-564, 2002.
- [55] H. G. Lee, S. C. Lee, H. Y. Kim, and R. H. Lee, "Is the internet making retail transactions more efficient? Comparison of online and offline CD retail markets", *Electronic Commerce Research and Applications*, vol. 2, no. 3, pp. 266-277, Autumn 2003.
- [56] K. C. Chang, J. Jackson, and V. Grover, "E-commerce and corporate strategy: an executive perspective", *Information and Management*, vol. 40, no. 7, pp. 663-675, August 2003.
- [57] B. T. Ratchford, X. Pan, and V. Shankar, "On the efficiency of internet markets for consumer goods", *Journal of Public Policy & Marketing*, vol. 22, no. 1, pp. 4-16, April 2003.
- [58] B. K. Lee, and W. N. Lee, "The effect of information overload on consumer choice quality in an on-line environment", *Psychology & Marketing*, vol. 21, no. 3, pp. 159-183, February 2003.
- [59] R. Suri, M. Long, and K. B. Monroe, "The impact of the Internet and consumer motivation on evaluation of prices", *Journal of Business Research*, vol. 56, no. 5, pp. 379-390, May 2003.
- [60] Y. Cao, T. S. Gruca, and B. R. Klemz, "Internet pricing, price satisfaction, and customer satisfaction", *International Journal of Electronic Commerce*, vol. 8, no. 2, pp. 31-50, 2003.
- [61] R. Sen, and R. C. King, "Revisit the Debate on Intermediation, Disintermediation and Re-intermediation due to E-commerce", *Electronic Markets*, vol. 13, no. 2, pp. 153-162, 2003.
- [62] F. Ancarani, and V. Shankar, "Price levels and price dispersion within and across multiple retailer types: Further evidence and extension", *Journal of the Academy of Marketing Science*, vol. 32, no. 2, pp. 176-187, March 2004.
- [63] M. R. Baye, and J. Morgan, "Price Dispersion in the Lab and on the Internet: Theory and Evidence", *RAND Journal of Economics*, vol. 35, no. 3, pp. 449-466, Autumn 2004.
- [64] S. Moodley, and M. Morris, "Does e-commerce fulfil its promise for developing country (South African) garment export producers?", *Oxford Development Studies*, vol. 32, no. 2, pp. 155-178, 2004.
- [65] Y. Cao, and T. S. Gruca, "The influence of pre-and post-purchase service on prices in the online book market", *Journal of Interactive Marketing*, vol. 18, no. 4, pp. 51-62, September 2004.
- [66] M. J. Eppler, and J. Mengis, "The concept of information overload-a review of literature from organization science, accounting, marketing, MIS, and related disciplines", *The Information Society: An International Journal*, vol. 20, no. 5, pp. 1-20, 2004.
- [67] D. Kuksov, "Buyer search costs and endogenous product design", *Marketing Science*, vol. 23, no. 4, pp. 490-499, Fall 2004.
- [68] D. Wu, G. Ray, X. Geng, and A. Whinston, "Implications of reduced search cost and free riding in e-commerce", *Marketing*

- [67] *Science*, vol. 23, no. 2, pp. 255-262, May 2004.
- [68] J. L. Willis, "What impact will e-commerce have on the US economy?" *Economic Review-Federal Reserve Bank of Kansas City*, vol. 89, pp. 53-68, March 2004.
- [69] C. Kocas, "A model of Internet pricing under price-comparison shopping", *International Journal of Electronic Commerce*, vol. 10, no. 1, pp. 111-134, 2005.
- [70] P. Pereira, "Do lower search costs reduce prices and price dispersion?" *Information Economics and Policy*, vol. 17, no. 1, pp. 61-72, January 2005.
- [71] A. M. Chircu, and V. Mahajan, "Managing electronic commerce retail transaction costs for customer value", *Decision support systems*, vol. 42, no. 2, pp. 898-914, November 2006.
- [72] A. Molla, and R. Heeks, "Exploring e-commerce benefits for businesses in a developing country", *The Information Society*, vol. 23, no. 2, pp. 95-108, March 2007.
- [73] D. G. Soopramanien, R. Fildes, and A. Robertson, "Consumer decision making, E-commerce and perceived risks", *Applied Economics*, vol. 39, no. 17, pp. 2159-2166, 2007.
- [74] B. C. Su, "Characteristics of consumer search on-line: how much do we search?", *International Journal of Electronic Commerce*, vol. 13, no. 1, pp. 109-129, Fall 2008.
- [75] F. Sabaté i Garriga, A. Cañabate Carmona, E. Cobo Valeri, and D. Garcia, "Is the Internet making markets more efficient? The evidence according to price indicators in Spain", *Intangible Capital*, vol. 5, no. 2, pp. 125-15, March 2009.
- [76] O. Hinz, and J. Eckert, "The impact of search and recommendation systems on sales in electronic commerce", *Business & Information Systems Engineering*, vol. 2, no. 2, pp. 67-77, March 2010.
- [77] A. Ghose, A. Goldfarb, and S. P. Han, "How is the mobile Internet different? Search costs and local activities", *Information Systems Research*, vol. 24, no. 3, pp. 613-631, September 2013.
- [78] K. C. Laudon, and C. G. Traver, *E-commerce* (pp. 1-912). Boston, MA: Pearson, 2013.
- [79] Y. Liu, H. Li, G. Peng, B. Lv, and C. Zhang, "Online purchaser segmentation and promotion strategy selection: evidence from Chinese E-commerce market", *Annals of Operations Research*, vol. 233, no. 1, pp. 263-279, 2015.
- [80] A. Seetharaman, I. Niranjan, A. S. Saravanan, and D. Balaji, "A study of the moderate growth of online retailing (ecommerce) in the UAE", *The Journal of Developing Areas*, vol. 51, no. 4, pp. 397-412, 2017.
- [81] B. D. L. Santos, A. Hortacsu, and M. R. Wildenbeest, "Search with learning for differentiated products: Evidence from e-commerce", *Journal of Business & Economic Statistics*, vol. 35, no. 4, pp. 626-641, 2017.
- [82] Y. Wang, and S. H. Lee, "The effect of cross-border e-commerce on China's international trade: An empirical study based on transaction cost analysis", *Sustainability*, vol. 9, no. 11, pp. 1-13, November 2017.
- [83] L. Lu, and T. Reardon, "An economic model of the evolution of food retail and supply chains from traditional shops to supermarkets to E-commerce", vol. 100, no. 5, pp. 1320-1335, August 2018.
- [84] A. Ali, and J. Bhasin, "Understanding customer repurchase intention in e-commerce: role of perceived price, delivery quality, and perceived value", *Jindal Journal of Business Research*, vol. 8, no. 2, pp. 142-157, June 2019.
- [85] N. Vulkan, *The Economics of E-commerce*. Princeton University Press, 2020.
- [86] Y. Badra, "The impact of online low search costs on firms' incentives to invest in quality", *Managerial and Decision Economics*, vol. 41, no. 8, pp. 1629-1633, July 2020.
- [87] T. Malone, J. Yates, and R. Benjamin, "Electronic markets and electronic hierarchies", *Communications of the ACM*, vol. 30, no. 6, pp. 484-497, June 1987.

Big Data Analytics in E-Commerce: Understanding Personalization using Social Media Platforms

Manpreet Kaur

Assistant Professor: SGTB Khalsa College,
Anandpur Sahib
E-mail: karmanpreetminhas@yahoo.co.in

Akriti

Assistant Professor: B.A.M. Khalsa College
Garhshankar
E-mail: diwan29akriti@gmail.com

Abstract

The affirmation of social media platforms has radically changed customers purchasing habits & the way businesses develop their e-commerce sites. E-commerce has always been led by the tenet ‘know your customer’. The application of big data analytics in e-commerce has enabled the personalization of shopping to each customer. Big data derived from transaction or business activity, click stream, video and voice is processed to micro-segment customers and derive granular insights. This paper is an exploration into the mode in which such personalization is done through customer profiling, predictive analytics, targeted advertisements, upgrading user experience and price personalization. Furthermore, it inspects accompanying confrontations that come to light from both inside and outside the organization, like organizational culture, induction of business intelligence, data privacy and security, and data monopolies. The study attempts to examine the Impact of Big data on E-Commerce with special reference to Social Media Platforms.

Keywords: Social Influence, Big data analytics, SMA, E-Commerce, Personalization.

I. INTRODUCTION

The e-commerce boom is noticeable. The Indian e-commerce market will outstretch US\$ 99 billion by 2024, expanding, at a 27% CAGR over 2019-24. Some overall factors of growth are internet penetration, cheap smartphones and mobile data, social media shopping and m-commerce, innovations in supply chain management, improvements in logistics infrastructure and storage, new digital payment manifesto, and big data. The combination of e-commerce and big data has knocked off a virtuous cycle. On one hand, big data has been moderately answerable for the prosperity of e-commerce. On the other hand, its prosperity has meant that the newest implementations of big data analytics ('BDA') are quickly implemented in e-commerce. BDA has numerous uses in e-commerce, to the same degree personalization, logistics, consumer service, and fraud management. [1] Explained social media as “a group of Internet based applications that, built on the ideological and technological foundations of Web

2.0, permits the formation and exchange of user-generated content”. Retail and e-commerce entities are consumer facing by nature. The nature of their business stimulates them to recognize their consumer and feed in a manner that leads to competition. On the other hand, e-commerce is mentored by the cardinal tenet ‘know your customer’. Aberdeen found that 75% of shoppers favor brands who personalize their messages and offers. According to Smart Insights, 48% of consumers spend more when their experience is personalized. Janrain found that 74% become frustrated when brands provide irrelevant content to them. E-marketers enjoy personalization’s benefits. Invesp found that 59% of marketers experiencing good ROI after personalizing their online store. According to Marketing Week, providing a personalized experience can improve conversions by nearly 8 %.

II. REVIEW OF LITERATURE

Sinha et al. (2012) explained how social media has become innovative in its reach to the masses, and organizations can utilize the study of social media for various purposes (p. 66). Investigating the social media’s analytics will be a great way to help organizations understand their employees and customers. The authors detailed that organizations can capitalize on data from “blogs (Blogger, LiveJournal), micro-blogs (Twitter, FMyLife), social networking (Facebook, LinkedIn), wikis (Wikipedia, Wetpaint), social bookmarking (Delicious, CiteULike), social news (Digg, Mixx), reviews (ePinions, Yelp), and multimedia sharing (Flickr, Youtube)” (p. 67). Sinha et al. (2012) commented that recent research declared usage of social media was a great method to help an organization’s marketing efforts (p. 77). They noted the analysis of social media could have substantial benefits, including the study of employee behavior, highlighting their business goals and intentions, sharing information, and developing strategic focus (p. 79). Sinha et al. exclaimed there is a great necessity to integrate social media and big data with business intelligence (p. 80). [2] Choi et al. (2017)

noted that risk management will benefit from the developing field of big data analytics (p. 81). Today's data-driven world has created an abundant amount of data, and management of the data will be a challenge (p. 81). The authors claimed that the prevalence of cloud services or web-based systems will add security concerns and risk management issues (p. 82). Choi et al. stated analytics can systematically manage and secure data (p. 86). Big data analytics offers a more robust security system to protect organizations' data (p. 87). [3] Sivarajah et al. (2020) explored how using big data and social media analytics for marketing purposes can be lucrative for organizations. The authors elaborated that the "big data and analytical tools can create actionable insights for delivering sustainable value, improving business performance and providing competitive advantage" (p. 163). Sivarajah et al. exclaimed that social media marketing was key to competitive advantage (p. 163). The authors emphasized the use of analytics can "improve transparency, decision-making, as well as to improve collaboration" (p. 164). The authors indicated that sentimental analytics gives organizations the ability to gauge peoples' feelings on key issues (p. 167), which allow businesses to monitor and understand consumer trends. The team claimed social network analysis can be used to study groups of people by understanding their characteristics (p. 167). Sivarajah et al. concluded that analytics "allows the organization to know their customer base on a more emotional level which not only enhances the service proposition, but also enables a more focused and targeted collaboration" (p. 176). [4]

III. OBJECTIVE OF THE STUDY

The study attempts to examine the Impact of Big data on E-Commerce with special reference to Social Media Platforms.

IV. METHODOLOGY

The study is based on secondary sources of data. Different books, journals, newspapers and relevant websites have been consulted in order to make the study an effective one.

V. BIG DATA ANALYTICS IN E-COMMERCE

E-commerce companies have been one of the speedy groups to adopt BDA, due to their constitutional requirement to stay ahead of competition. Over that such competition is based on their data. Big data is obtained from transaction or business activity data, click stream data, video data and voice data.

Transaction or business activity data is collected from retail transactions, product consumption and consumer outranges sourced in the regular course of business. Clickstream data is gathered from clicks on social media and online advertisements. While video data is gathered from video in retail settings, while voice data is sourced from phone calls, call centers and customer service recordings. Due to the expediency of personalization as a commercial goal, e-commerce companies endeavor to tailor the shopping experience to each shopper. In order to derive business intelligence, e-commerce companies must endorse BDA.

VI. CURRENT SCENARIO OF SOCIAL MEDIA PLATFORMS

There were 467.0 million social media enjoyer in India in January 2022. The number of social media users in India at the start of 2022 was parallel to 33.4 percent of the total population. Kepios analysis disclose that social media users in India enlarged by 19 million (+4.2 percent) in the middle of 2021 and 2022.

Social networks offer free and fast communication with friends generally in multiple forms to the same degree tweets, pictures, videos and texts. Another attractive characteristic is creation of groups for multicast communication with friends, colleagues or family members. The most trendy social media platforms are given below.

Table 1
Social Media Applications

Sr. No.	Applications	Number of active users in India (million)
1	Facebook	329.7
2	Instagram	203.3
3	YouTube	467
4	Facebook Messenger	122.5
5	LinkedIn	83
6	Snapchat	126
7	Twitter	23.60

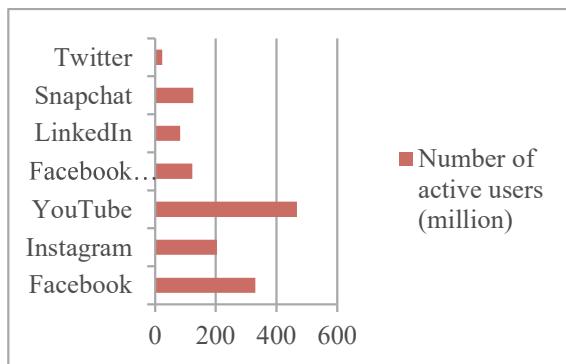


Figure 1: Social Media Users

VI. WAYS TO USE BIG DATA FOR E-COMMERCE PERSONALIZATION

Looking at the sales channel there are various courses of action. Big Data in increasing online sales and uplifting long-term consumer loyalty. Data can help answer questions correlated to future trends, stocking, pricing, generating a customized experience and eventually selling more.

A. 360 degrees view of the client

What if you knew your client as a friend? How would you sell to a person knowing their age, family status, income, likes, and fears? Big Data permits you to do this without taking them out for coffee on a consistent basis, just by gathering the cookies in their browser, social media interactions and transaction history.

B. Dynamic pages

Once you have constructed the complete profile of your customers, you can utilize it to tweak what each visitor sees and make it compatible to their situation. A simple example is the news feed demonstrated by Facebook, a tailor-made filtration of what is accessible online. Dynamic content can take other forms that are accessible to handle without changing the whole content of the site. These comprehend pop-ups, calls to action and retargeting in other state of affairs, like remarketing an abandoned product in the cart.

C. Recommendation engines

Studies reveal that 92% of clients are affected by recommendations, and half of them want to receive such information to assist them select wiser and faster. This is such an extensively known implementation that for some companies it is almost synonymous with Big Data for e-commerce. The

power of the recommendation engines depends on the fact that they uncover unnoticeable dependencies. A great advocacy tool is also integrated with inventory and logistics to be able to highlight only those products that are obtainable to be shipped to the location of the consumer. It wouldn't make any sense in recommending discontinued products that are not shipped to the mentioned address.

D. Customer care

Almost 9 out of 10 shoppers stop doing business with a company that has disconcerted them in the client service department. On top of that, 96% of unhappy customers don't complain, they just leave. Getting a new customer costs 6-7 times more than keeping an existing one. Having said that, not investing in a CRM based upon Big Data that acts as a supporter to your customer care team is similar to intentionally wasting money.

E. Case Study: How Facebook can drive personalization?

It's the world's substantial social network by an enormous margin, and most of us are habituated using it to allocate details of our everyday lives with our friends and families. It's not confidential now that we're also allocating it with their advertisers, but that hasn't put most of us off using it! So here's a short rundown of how Facebook has been one of the most successful companies in the world at collecting our data and converting it into profit and why some think its business operations sometimes overstep the mark. With every like we do, every comment, every status upgrade etc. users collectively create huge amounts of data every second. Every 60 seconds, 317,000 status updates; 400 new users; 147,000 photos uploaded; and 54,000 links are shared on facebook. Apache Hadoop is an open source framework that is utilized by Facebook to efficiently store and process large datasets fluctuating in size from gigabytes to petabytes of data. Instead of using one large computer to store and process the data, Hadoop permits clustering various computers to analyze massive datasets in parallel more quickly. Users who love social media have faith in the potential benefits far over balance the hazards. Putting aside how much easier it makes keeping in touch with our friends and family, there's certainly a lot to be comprehend from studying the data created during that communication. And collecting data from us is the foundation of Facebook's business template and can operate personalization for other E-commerce.

VII. CHALLENGES

Computerized personalization is beginning to be applied to customer relationship management, electronic commerce and information portal services. As these application areas are fairly new, there are some attendant challenges facing personalization. Let us examine them. A review of internal challenges indicate that it is wise to avoid big data hubris, as big data is suited more for information rather than explanation. Big data suffers from inherent limitations that may come in the way of deriving meaningful business intelligence. It is futile for an organization to heavily invest in data analytics and end up being data rich and insight poor. In external challenges, data privacy and security, and the danger of data dominance are noteworthy. With more facets of individual behavior being datafied, big data has the potential to tip the scales of power against the individual. Further, access to big data may create a divide between data rich and data poor players in online commerce. Effectively tackling these challenges will shape the future of e-commerce and the gains it can bring to consumers.

VIII. CONCLUSION

There is no denying the reality that in less than a decade, Big Data has become a multi-billion-dollar industry. Nowadays, the Big data revolution has appeared with the growth of the internet, wireless networks, smartphones, social media and other technologies. We can characterize social media analytics as a very large dataset that can be analyzed to disclose trends, patterns, and associations. It is useful for both big and small businesses. They are assembling data-driven decisions using BDA & SMA. Social media platforms aim at upgrading the interaction between the customers and ecommerce businesses instead of just educating them about profitable deals and new product launches. A segment of the social media platforms allows direct selling by utilizing a pair of social groups and communities. SMA assists them to have adequate information about the services and products. It also helps to learn about customers, suppliers, consumers' preferences that can be collected and analyzed. Overall it assists to understand and optimize the business processes. There is no denying the reality that Big Data is like a window to a new world. It gives a new vision to the world that helps to do things in a more efficient and productive way.

ACKNOWLEDGEMENT

We would like to thank Prof. Kanwar Kulwant Singh for his encouragement and support.

REFERENCES

- [1] Andreas M.KaplanMichael Haenlein (2010) "The challenges and opportunities of Social Media, Volume 53, Issue 1, Pages 59-68, <https://doi.org/10.1016/j.bushor.2009.09.003>.
- [2] Sinha, V., Subramanian, K. S., Bhattacharya, S., & Chaudhuri, K. (2012). The Contemporary Framework on Social Media Analytics as an Emerging Tool for Behavior Informatics, HR Analytics and Business Process Management, 17(2), 65–84.
- [3] Choi, T.M., Chan, H. K., & Yue, X. (2017). Recent Development in Big Data Analytics for Business Operations and Risk Management. IEEE Transactions on Cybernetics, 47(1), 81–92. <https://doi.org/10.1109/tcyb.2015.2507599>
- [4] https://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=138312 Sivarajah, U., Irani, Z., Gupta, S., & Mahroof, K. (2020). Role of big data and social media analytics for business to business sustainability: A participatory web context. Industrial Marketing Management, 86, 163–179. <https://doi.org/10.1016/j.indmarman.2019.04.005>.
- [5] Valerio, Curzi, William, Lecoq; Noémier, Quéré (2019) : The Impact of Social Media on E-Commerce Decision Making Process, International Journal of Technology for Business (IJTB), ISSN 2644-5085, Springwish Publisher, Bratislava, Vol. 1, Iss. 1, pp. 1-9, <http://dx.doi.org/10.5281/zenodo.2591569>.
- [6] Mir Saleem (2017) "E commerce personalization using Big Data and challenges", International Journal of Advanced Research and Engineering, Vol 6, Issue 1.
- [7] E. TaiebiJavida, M. Nazaria and M. R. Ghaelia (2019) "Social media and e-commerce: A scientometrics analysis" International Journal of Data and Network Science 3, 269–290.
- [8] Umar Ruhi Social Media Analytics as a Business Intelligence Practices Current Landscape & Future Prospects Journal of Internet Social Networking & Virtual Communities <http://www.ibimapublishing.com/journals/JISNVC/jisnvc.html> Vol. 2012 (2012), Article ID 920553, 12 pages DOI: 10.5171/2012.920553
- [9] Sheela Singh, PriyankaArya, Alpna Patel and Arvind Kumar Tiwari (2019) Social Media Analysis through Big Data Analytics: A Survey 2nd International Conference on Advanced Computing and Software Engineering (ICACSE-2019)
- [10] Adnan VeyselErtemel, (2019) Consumer Insight as Competitive Advantage using Big Data and Analytics International Journal of Commerce and Finance, Vol. 1, Issue 1, 2015, 45-51.
- [11] <https://towardsdatascience.com/big-data-in-e-commerce-personalization-explained-580efa0dec50>
- [12] <https://www.linkedin.com/pulse/case-study-big-data-facebook-viplove-bansal>
- [13] https://www.bernardmarr.com/img/bigdata-case-studybook_final.pdf
- [14] <https://techvidvan.com/tutorials/advantages-of-big-data/>
- [15] <https://data-flair.training/blogs/data-science-big-data-quotes/>
- [16] <https://datareportal.com/reports/digital-2022-india>

Digital India: Foundation for Future India

Amandeep Kaur

Department of Commerce and Economics

B.A.M. Khalsa College, Garhshankar

brar.anu18@gmail.com

Abstract

The Digital India project launched by our Prime Minister Narendra Modi on 1st of July in 2015. It is an effective scheme to transform India for better growth and development of people and country. It aims to make India digital push for good governance and more jobs. It is a Government of India's action aiming at enhancing the online infrastructure and broadening internet connectivity. The motto of Digital India is "Power to Empower". Digital India is a large umbrella program which will reconstruct and make-over some existing schemes to bring in a transformative footprint. The Digital India vision focuses on transforming our country into a digital economy with involvement from citizens and businesses. Today, we can't dream our life without technology. In the twenty-first century, one of the most crucial technologies is the impact of digitization. It allows every person to communicate anytime anywhere. The current study concentrated on nine pillars of digital India, distinct aspects of digitalization:- its impact on economy, society and environment. Benefits of digitalization and different challenges which lead to hurdles in the victorious implementation of this program are also talked through. The research methodology is descriptive cum analytical in nature and the data for this study is accumulated through secondary sources such as websites, research journals, newspapers, magazines etc. The study gives a positive thought of providing additional employment prospects for the youth which will improve the nation's economy if it comes into actuality.

Keywords: *Digitalization, Indian economy, Pillars, Technology, Internet, Future India*

I. INTRODUCTION

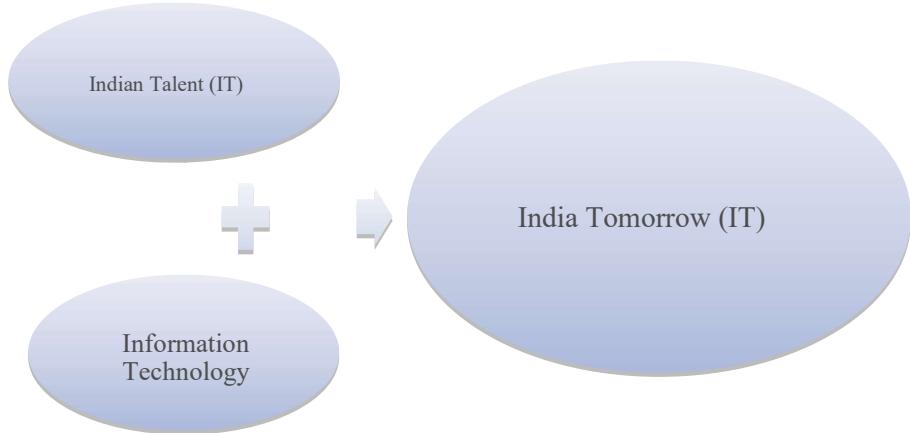
Digital India is a campaign launched by The Government of India on 1st July 2015. India today ranks amongst the largest economies of the world and has become the fastest growing large economy. However, it is still a lower middle income country in terms of per capita income with substantially high levels of poverty and deprivation and significant regional imbalances in development. The vision of a

future India must aim at holistic development in all sectors of the economy and society so that overall human development and quality of life is improved in the country.

Can technology led transformation lay the foundation of a future India? In this paper, I examine this central question with recommendation to the recently launched Digital India program of the Government of India that points at transforming the country into a digitally empowered society and knowledge economy. The program interlaces together a large number of ideas and thoughts into a single, broad vision so that each of them is seen as part of a larger goal. The focus of Digital India is on being transformative – to realize IT (Indian Talent) + IT (Information Technology) = IT (India Tomorrow) and on making technology central to enabling change. The idea is to transform the whole ecosystem of public services through the use of information and communication technologies (ICT) and build comprehensive capabilities across a wide range of sectors, e.g., ICT infrastructure, e-governance, software services and delivery platforms, electronics manufacturing, Internet of Things (IoT), IT skills and job creation, etc. The focus is on making ICT a key driver for transforming every sector of the economy and society. In order to attain this futuristic vision for a developed India, Digital India's center of attention on several key developmental strategies. These include promoting investments through improving 'ease of doing business', encouraging entrepreneurship across various sectors through leveraging IT, capacity building and creation of jobs in the IT sector, come up with easy access to public services anytime from anywhere, promoting financial inclusion through mobile banking and micro ATMs, promoting literacy through e-books and other digital contents, faster services and dissemination of information to promote growth in multiple sectors such as agriculture, education and healthcare, and encouraging more participation from women in diverse sectors of the economy and society. It is rightly said by the honorable Prime Minister of India, Narendra Modi that Information Technology plays

crucial role to make India a digital country, in his words:

Figure 1. Focus of Digital India: Transformative India



II. REVIEW OF LITERATURE

A number of research papers and articles from various disciplines provide a detailed insight about the great significance and influence of digital India on the economy as a whole and particularly on the technology sector. Gupta and Arora (2015) studied the impact of the digital India project on India's rural sector. The study found that many schemes have been launched in digital India to boost agriculture sector and entrepreneurship development in rural areas. The Digital India program has also set the stage for empowerment of rural Indian women. [1] Rani Suman (2016) concluded that the digital India project provides a huge opportunity to use the latest technology to redefine India the paradigms of the service industry. It also pointed out that many projects may require some transformational process, reengineering, refinements to achieve the desired service level objectives. [2] Midha Rahul(2016) concluded that digital India is a great plan to develop India for knowledge future but its improper implementation due to inaccessibility and inflexibility to requisite can lead to its failure. Though the digital India program is facing a number of challenges yet if properly implemented it can make the best future of every citizen. So we Indians should work together to shape the knowledge economy [3].

III. OBJECTIVES OF THE STUDY

Objectives of the study are:

1. To study the Vision of Digital India.

2. To study the Nine pillars of Digital India.
3. To find out the challenges faced in implementation of this program..
4. To find out practical solutions and innovative ideas to accomplish the vision of a digital India-a reality.

IV. RESEARCH METHODOLOGY

This is a descriptive research which helps to understand about the Digital India campaign and its effects so far and to develop a model to achieve Sustainable Development Goals with the pillars of Digital India. The type of data used is secondary data obtained from sources such as research papers, newspaper, journals and magazine articles, media reports, government official sites, etc.

IV. DISCUSSION

A. Current Status & Vision of Digital India

Digital India' is a central program to make India ready for a knowledge-based future. As per NASSCOM's (National Association of Software & Services Companies) Strategic Review 2021, India's technology sector was estimated to reach US\$ 194 billion in FY21, a 2.3% YoY increase on the back of rapid digital transformation and technology adoption in the country. In 2020, the Indian technology sector accounted for 8% of India's gross domestic product (GDP). In addition, the sector remained a net employer with emphasis on digital upskilling. NASSCOM also estimated that the country's digital talent pool is likely to exceed ~1.17 million employees in FY21, a 32% YoY surge.

In another report, McKinsey highlighted that the ‘Digital India’ initiative is expected to boost the country’s digital economy to US\$ 1 trillion by 2025, up from US\$ 200 billion in 2018. The Digital India program is centered on three key vision areas.

- Digital Infrastructure as a Utility to Every Citizen
- Governance & Services on Demand
- Digital Empowerment of Citizens

B. Nine Key Pillars of Digital India

In order to lay the foundation for future India, Digital India has identified nine key pillars of growth areas. These include broadband highways, universal access to mobile connectivity, public internet access program, reforming government through technology, electronic delivery of services, information for all, electronics manufacturing, IT for jobs and early harvest programmes. The pillars provide a number of specific targets and activities within those growth areas so that the concerned sectors can benefit from IT enablement. The figure 1 below depicts the nine pillars of the program.

Figure 2: Pillars of Digital India Programme



Source: visual.ly

1. Broadband Highways

The government with the vision of “Digital India” had granted ₹5 billion to build high speed broadband highways associating all the villages, government departments, universities, R&D institutes, etc. Internet access is essential in developing countries since it benefits to share information from one person to another. The number of internet subscribers in India stands at more than 500 million while Paytm has around 20 million live users. The government in union budget-2022 promised broadband access in all villages by 2022, as it launched the ambitious National Broadband Mission entailing stakeholder investment of Rs 7 lakh crore in the coming years. The number of towers in the country which is about 5.65 lakh will be increased to 10 lakh. For instance, the availability of the internet has

upgraded the learning process in higher education institutions. Students are capable of embracing digital learning from their suitable households thus cultivating a spirit of entrepreneurship in the country. High literacy rates assist to bring down poverty within the region thus boosting living standards and economic development at large.

2. Universal Access to Phone

The initiative is to spotlight network penetration and fill the space in connectivity in the country. All together 42,300 uncovered villages will be covered for contributing universal mobile connectivity in the country. Government is mainly preparing to connect unconnected areas and speedy utilization of technologies like network technologies like 3G, 4G and 5G etc. General public will access the online

government services with the guidance of handheld devices so that the nation will be well connected, efficient and more productive.

3. Public internet access

India's telecommunications market is nowadays the second-largest in the world due to the largest number of users. A recent report by the World Bank disclosed that over 41% of the world's population is connected through the use of the Internet. Year after year, India is regularly rising in terms of internet user base. In January 2022, there were 658.0 million internet users in India. The government of India has installed Wi-Fi hotspots in most hotels, railway stations and airports to generate digital cities. Through this initiative, citizens have been capable of receiving tickets online thus reducing queuing time. Particularly, higher education institutions have been enquired with fiber optic networks which are fast in browsing and sharing of information. Internet coverage in remote areas has enabled entrepreneurs to explore market trends in terms of prices thus enabling them to sell their products when market rates are advantageous. As a result, they are able to enhance sales revenue leading to more sophisticated living standards. Internet coverage has proceeded in e-commerce where businesses have derived to operate through websites and social media. By the way of online platforms, jobs are created including online support representatives and administrators. This further led the way to economic growth and development.

4. E-Governance

It mentions the implementation of information communication technology to revolutionize the effectiveness, transparency, efficiency and accountability of exchange of information in the middle of government agencies and citizens. This has been allocated by creation of online portals where the government publicizes and communicates essential information with citizens. India has boosted its economy since citizens are able to analyze and track progression of achievement thus closing loopholes for corruption in the distant future.

5. e-Kranti

The e-Kranti project came up with electronic delivery of services to the citizens. This pillar is appraised as vital for enhancing delivery of good governance, e-governance and easy governance in the country. The initiative was composed by head of Electronics and

Information Technology to speed up electronic delivery of services via through a selection of Mission Mode Projects.

6. Information for All

Open Data platform and online hosting of information & documents would clear the way open and easy access to information for citizens. Specifically, information technology amplifies the transfer of information from customers regarding satisfaction magnitude from the services and products furnished by manufacturers. Simultaneously, information technology entitles businesses to acquire data about the competition requirements and market slit that are newly established in the global market. With the exception of communication, information technologies are also essential in enhancing brand quality. In India, the government has set the seal on the public can access vital information through broadcasting, press conferences and government gateway.

7. Electronics Manufacturing

Presently, the electronics industry is among major contributors to India's merchandise imports after petroleum products. It is predicted that by 2030, Indians private customer market would have reached \$9.6 trillion which accounts for 47% of its GDP (Telecomlead.com, 2018). Indian government is for shadowing to outstretched 100% electronic manufacturing inside its territories in order to lower importing cost as well as generate employment to its juvenile generation.

8. IT for Jobs

This initiative targets at training youngster about online jobs to lower unemployment rates inside the country. The bulk targeted classes are entrepreneurs and graduate students who have completed their studies. This initiative was for shadowing to transform the way business is done in India. The Indian e-commerce market is predicted to grow to US\$ 200 billion by the year 2026. As per a McKinsey report, a digital economy is likely to create 60-65 million jobs by 2025.

9. Early Harvest Programmes

This program comprises of projects which the government is for shadowing to complete within

short time frames. They involve biometric attendance, equipping all universities with Wi-Fi, IT platform for message and government greetings to greetings. In keeping with the report by Agrawal and Sen (2017), almost 90% of India universities are furnished with internet facilities. [4] Over and above, the government has created a database to store citizen data online to keep away from distortion or data loss. This inventiveness ensures that the government has upgraded data for decision making.

C. Impact of Digital India

- Broadband in 2.5 lakh villages, universal phone connectivity
- Net Zero Imports by 2020
- 400,000 Public Internet Access Points
- Wi-fi in 2.5 lakh schools, all universities; Public wi-fi hotspots for citizens.
- Digital Inclusion: 1.7 Cr trained for IT, Telecom and Electronics Jobs
- Job creation: Direct 1.7 Cr. and Indirect at least 8.5 Cr.
- e-Governance & e Services: Across government
- India to be leader in IT use in services – health, education, banking
- Digitally empowered citizens – public cloud, internet access.

Table 1. Digital India by numbers

Initiatives	Statistics
Digital Identity(AADHAR)	131.68crore
Internet users	658.0 million
Mobile phone users	1.18 billion
Smartphone users	468 million connections
Social media users	250 million
Ration cards	100% digitized
Govt. e Marketplace	263,731 products by 105,889 sellers listed

V. CHALLENGES

More than Six years has passed since the Digital India mission has been proclaimed but it is facing multiple challenges in successful implementation. Some of the challenges are-

1. High level of digital illiteracy is the massive challenge in the success of digital India programme. Low digital literacy is a key obstacle in adaptation of technologies.
2. Making the Digital India programme known and creating an awareness between common masses about its benefits is also a considerable challenge.
3. A key element under this vision is the high speed of the internet as a core utility to smooth online

delivery of several services. India has ground level internet speed.

4. It is a giant piece of work to have connectivity with each and every village, town and city. Attaching 250000 Gram Panchayats through National Optical Fiber is not an unchallenging task. The biggest challenge is making certain that each panchayat point of broadband is fixed up and functional. It is observed that 67% of NOFN points are non functional even in the pilot phase.

5. The largest challenge faced by the Digital India programme is slow and delayed infrastructure progress. India's digital infrastructure is inclusively inadequate to tackle the growing increase in digital transactions. India requires over 80 lakh hotspots as against the availability of about 31000 hotspots at present to come to global level, according to the ASSOCHAM-Deloitte report.

6. According to the ASSOCHAM- Deloitte report, the issue of taxation and regulatory guidelines have confirmed to be a roadblock in realizing the vision of Digital India. Some of the common policies bottlenecks include lack of clarity in FDI policies have influenced the growth of ecommerce.

7. The private participation in government forecasts in India is low because of long and complex regulatory processes.

8. Many request proposals emerged by the government are not picked up by capable private sector organizations since they are not commercially viable. Currently over 55000 villages are deprived of mobile connectivity because furnishing mobile connectivity in such locations is not commercially viable for service providers, ASSOCHAM-Deloitte report pointed out.

9. There is an extensive digital divide among urban and rural India. Till now endowments have not been deployed effectively to encounter the cost of infrastructure creation in rural areas

10. India has 1600 languages and dialects. Unreachable digital services in local languages are a great hurdle in digital literacy.

11. Fear of cybercrime and breach of privacy has been an obstacle in adoption of digital technologies. Most of the technology including cyber security equipment is imported. We do not have indispensable skills to inspect these for hidden malwares. We have upper class experts for these high end jobs at present. According to NASSCOM, India requires 1 million trained cyber security professionals by 2025.

VI. RECOMMENDATIONS

1. The visions taken by the government can only be prosperous if people get elaborate in the transformation. The Schools and Colleges can

generate awareness among the people of their locality about the visions and transmitting knowledge to the people.

2. For doing so they also require to have the knowledge about the program and how to utilize the facilities layout by the government.

3. Community centers can be assembled where people who are more literate about the affairs can help the other people and experts can visit them to give guidance from time to time.

4. People should be transmitted with the knowledge of factors which influence their savings and how they can maximize their savings or the facilities available for them to help them to do so.

5. People should also be made conscious about the security of their personal information concerning their accounts and online frauds.

6. The banking system must also be made robust as people require to trust the system before they go with technological extension.

VII. CONCLUSION

Technological advancement is essential for developing nations. India is aspiring to be a leader in various global technology platforms in order to fulfill the requirement that it is crucial for digital technologies to be utilized to improve public services, deliver financial inclusion, etc. India has recognised this secret and they have encompassed a digital India campaign to enhance effective communication between citizens and government while giving essential services with ease. So far, the digital India campaign has been prosperous since it has helped India dramatically by creating job opportunities, refining literacy rates, eliminating corruption, technological advancements as well as uplifting gross domestic product. It also improved the social and economic surroundings of people living in rural areas through development of non-agricultural economic activities apart from providing an approach to education, health and financial services. This further activates economic development since finances to support social amenities and other public infrastructures are readily accessible. These initiatives by the government will support India to achieve the Sustainable Development Goals of the UN within 2030. E-governance is sharing more to achieve sustainability which avails transparency in every transaction, reliability, reducing paper works, etc. The Digital India vision is in the infant stage, it takes some time to behold the impact on the economy. "In this Digital age, we have an chance to transform the lives of people in ways that were hard to assume a couple of decades ago"- Narendra Modi.

REFERENCES:

- [1] Gupta Neeru and Arora Kirandeep (2015). Digital India: A Roadmap for the development of Rural India. International Journal of Business Management ,vol(2), pp1333-1342. Retrieved from www.ijbm. Co .in.
- [2] Suman Rani (2016)" Digital India: Unleashing Prosperity" Indian Journal of Applied Research, Volume, Issue 4, DOI: 10.36106/IJAR.
- [3] Rahul midha (2016) "DIGITAL INDIA: BARRIERS & REMEDIES"IJIRSE, Volume 2, Issue8.
- [4] VANDNA (2015) "DIGITAL INDIA" International Journal of Creative Research Thoughts (IJCRT), Volume 6, Issue 1, ISSN: 2320-2882.
- [5] Radheshyam Sen vs Shri Pankaj Agrawal on 11 September, 2017 <https://indiankanoon.org/doc/91364476/>
- [6] Digital India Programme: Importance and Impact .Retrieved from <http://iasscore.in/national-issues/digital-india-programme-importance-and-impact>.
- [7] Digital India. Unlocking the trillion Dollar Opportunity: ASSOCHAM –Deloitte report, November 2016.Retrieved from www.assocham.org.org.
- [8] Digital India Retrieved from <http://www.indiacelebrating.com/government/digital-india/9.www.digitalindia.gov.in>.
- [9] Rajeshtimane.com (n.d). The Nine Pillars of Digital India available at <https://www.rajeshtimane.com/the-nine-pillars-of-digital-india/>
- [10] GIS based Planning. See: <http://india.gov.in/gis-based-planning-atlas-lucknow-district>
- [11] Digital India plan could boost GDP up to \$1 trillion by 2025: McKinsey, December 2014. See:<http://economictimes.indiatimes.com/industry/telecom/digital-india-plan-couldboost-gdp-up-to-1-trillion-by-2025-mckinsey/articleshow/45536177.cms>
- [12] India is now the world's third largest internet. See: <http://www.thehindu.com/sci-tech/technology/internet/india-is-now-worlds-third-largest-internet-user-after-us-china/article5053115.ece>
- [13] Intel Digital Skills program. See: <http://www.livemint.com/Politics/tVi3qteBfYKkXOPdw1O4J/Intel-India-to-aid-govt-with-digital-literacy-programme.html>
- [14] E-Kranti scheme gets Rs 500 crore boost. See: <http://www.india.com/budget-2014/union-budget-2014-live-e-kranti-scheme-gets-rs-500-crore-boost-93234/>
- [15] Here is the biggest challenge to Modi's digital India vision, October 2014. See: <http://qz.com/274880/here-is-the-biggest-challenge-to-modis-digital-india-vision/>
- [16] https://en.wikipedia.org/wiki/Digital_India
- [17] <http://www.csi-india.org/>
- [18] http://deity.gov.in/sites/upload_files/dit/files/Digital%20India.pdf
- [19] <http://indianexpress.com/article/india/india-others/cabinet-clearsdigital-india-programme/>
- [20] <https://digitallocker.gov.in/>
- [21] <http://digitalindiainib.com/>
- [22] <http://vikaspedia.in/e-governance/digital-india/digital-india>.
- [23] <http://datareportal.com/>
- [24] Digital India Report-IBEF.
- [25] MEITY India's Trillion dollar digital opportunity Report.

E- Commerce in India

Amandeep Kaur

*Assistant Professor: B.A.M. Khalsa College,
Garhshankar*

E-mail: aman.kaurnsr18@gmail.com

Abstract

E-commerce is tremendously increasing market in developing economy of India. The concept of e-commerce provides easy availability of goods and services for the customer, saves time and its cost is very less. The process is not limited to only buying and selling of products and services on online platform. It is more than that and includes the activities of manufacturing, marketing, selling, delivering, servicing and making payment for goods and services. But on the other hand it is not easy to install such a huge setup, internet connectivity; hesitation among customers and other parties creates obstacles in the way of advancement. But in last some years, private investments and government initiatives help in the growth of the market. Customer trust in the market has been developed. The concept of e-commerce is very important for India that will help it to access global market and to become one of the major economies in the race of e-commerce. It would generate revenues, job opportunities and global player's investment in heavy projects. E-commerce is also being selected as a career option as per the interest. The concept has wider scope in the coming years where all the transactions are preferred to be contactless and paper less.

Keywords – E-commerce, Indian economy, investments and government Initiatives.

I. INTRODUCTION

Now a day's commercial transactions have become easy with the help of e-commerce that is growing rapidly. It has led to a growing e-commerce market in India with the increase in the internet and smartphones. During a pandemic, Covid-19 E-commerce has become the need of the hour and is expected to increase US\$200 billion by 2026 which was US\$ 48.5 billion in the year 2018. The scope of e-commerce in India is increasing at the high speed that it would be second at the global level by 2034. Maximum contributors to such growth are, increase in buying grocery, fashion habits, personal care, beauty, and wellness. Despite the challenges of Covid-19 the market is expected to grow at 5 %.

E-commerce has become a trend that fulfils the needs of modern business. E-Commerce is electronic commerce that provides a platform to buy and sell goods and services with the help of the internet. All the concerned parties interact with each other in paperless form. It includes

online activities for exchanging goods and services. The activities involved are manufacturing, marketing, selling, and paying for goods and services. It can also be an acronym for internet commerce.

In the words of Vladimir Zwass, 'Electronic commerce is sharing business information, maintaining business relationships and conducting business transactions using telecommunications networks'. The concept has the following features:

- Ecommerce is available anytime and everywhere.
- The size of the e-commerce market is equal to the online population at a global level which is very huge.
- The standards used in e-commerce are universal.
- The information and details about online transactions of e-commerce are easily available in visual and audios that help the parties to perform the activities on their part.
- It is a two-way communication where the buyer and seller share their feedback and experiences that guide them for future development and growth.
- It allows the personalization and customization of services to the customer to make their experience better.

II. LITERATURE REVIEW

In the reference [1], the author has initially explained the Mode of E-commerce where defined B2B, B2C, C2C, C2B, and B2G. The authors also explained the e-commerce retail supply chain. The situation of online retail in India is also elaborated. Further, vertical-specific E-Commerce in India was also presented. E-commerce-related challenges and risks have also been defined. The role of the Government in this field has also been elaborated. The economic and social impacts of e-commerce have also been described.

Authors in reference [2] have explained the importance of e-commerce in the Indian economy. Initially, E-commerce has been explained and Different Types of E-Commerce have been shown. Then the literature review up to the publishing year of the paper has been presented. Moreover, Snapshot of E-Commerce Industry in India, Digital penetration in the Indian economy, Internet Penetration in India, Retail E-Commerce Sales in India has

been commented on. Finally, Government Initiatives Supporting the E-Commerce Growth in India and findings have been explained.

In 2020 (reference [3]), the definition of E-Commerce and its types have been explained. Further authors have also explained the Impact of E-Commerce on India's Economy. Explanations related to Market Size and Investments/ Developments have also been described. Government initiatives in this area have also been elaborated. [4] have elaborated the E-commerce with Its Impact on Consumer Behavior. The author has explained the difference between online/offline Stores, the Impact of the internet on Consumer Behavior, Convenience, and Consumer Trust in Internet Shopping. There is a good elaboration that has been listed on the consumer behavior. Concerning [5], it explains the development of e-commerce in India and market size. The government and future road map are also listed. The various contributions of the investments could also be seen here.

III. BENEFITS OF E-COMMERCE

The concept has the following benefits:

1. E-commerce is a platform that helps customers to do their shopping fast without approaching many physical shops. Moreover, the customer has more options in online shopping with just a click.

2. The e-commerce concept allows the customers to view the product from all perspectives including images, a description, product category, price, shipping fee, and delivery date, and also customer reviews. The online platform reduces the cost of maintaining a physical business. Less cost of production helps the sellers in reducing the sale prices that attract more customers worldwide. It is easy for the business houses to advertise their products and services to advertise with very less investment. E-commerce is available anywhere and anytime.

3. It is very easy for the seller to compare the products alternatives in the market and the products competitors are selling. The seller can make his product accordingly to get more and more customer attention. It would increase revenues and profits.

4. With the help of e-commerce, this seller can reach diviners at distant places the cellar can cover a broad area even at the global level or so which is not possible in physical sales.

5. The interaction between different parties is very fast as they all are performing at the online platform it includes speedy delivery of product to the customer and easy returns that give the more satisfied customer.

IV. LIMITATIONS

1. Sometimes for enabling e-commerce activities, some information personal and financial and personal is to be

shared. The parties sometimes hesitate and do not want to share the same and prefer physical transactions.

2. There is no privacy at some websites through which personal and financial information can be traced by unauthorized persons.

3. The seller and buyer at different places have to pay sales tax that can be avoided in physical transactions.

4. There is still fear about online platforms. The people want to physically reach the party they are contacting which gives them safety and satisfaction.

5. Some customers face the problem that images shown online are different from the products they received which reduces the credibility of online transactions.

6. E-commerce has made the world very small but still, the cultural differences are there that serve as hindrances in the growth of e-commerce in India.

7. E-Commerce transactions require qualified and technically skilled staff organization or business houses have to appoint them by doing a lot of expenses after that to retain them they have to give extra facilities that can increase the cost of production to the company or business houses

8. Cyber security issues also stop some users for parties to enter the virtual business.

9. The parties of e-commerce have to face technical limitations like lack of proper software and insufficient bandwidth of Telecommunication that can stop their transactions of e-commerce.

10. The business houses that are Run as per traditional ways have to convert in electronic form requires a lot of expenses and costs that can be a problematic situation for the business houses.

V. INDIAN GOVERNMENT SCHEMES

Without the help of the government, no initiative can be succeeded. The same is the case with e-commerce. The government of India has started to take initiative to promote e-commerce in India since 2014 there are various schemes launched by the government of India to make e-Commerce progress fast. Some of them are startup India skill India and innovation fund digital India and make in India from time to time government of India has launched different schemes as given below:

- In October 2021, the government e-marketplace portal recorded 7.96 million orders and registered worth Rs. 145,583 crore transactions

- In November 2021, the government e-marketplace portal recorded 7.78 million orders and registered worth Rs. 152,315 crore transactions.

- The Department for Promotion of Industry and Internal Trade (DPIIT) has made Open Network for Digital Commerce (ONDC) for performing the transactions smoothly. It has given access to all the players with equality. It would help in the best use of e-commerce for the growth and development of India.
- The government of India has made National Retail Policy to give the facilities a convenient way of performing transactions, a good license system, a digital transaction in retail services, updating of the system, and better administration. All these facilities increase the credibility of e-commerce activities.
- There is an updating for all the e-commerce companies that it is required for all to show the country they belong to with the product details.
- In 2019 government e-Marketplace has signed a memorandum of understanding with a bank named Union Bank of India this memorandum was signed to give cashless paperless and transparent payment system for the customers. It helped a lot in increasing paperless payments on the end of customers.
- Under the Digital India scheme launched by the government, there were some sub-parts namely startup India portal beam that is Bharat interface for money and warms up among these schemes helped a lot in increasing digitalization in India.
- In the year 2020, Mr. Piyush Goyal started offering their products to two government organizations and public sector undertakings buy online procurement portal that was a public portal.
- In 2020 e-Government of India has made required for all the E-Commerce companies to have a Permanent Account Number that is pan as it is required for all foreign companies.
- To get more and more foreign direct investment in India by E-Commerce foreign players the limit of foreign direct investment was increased up to 100% in business to business models.
- Now Government of India is investing in 5G and fiber networks so that E-Commerce transactions can be boosted in India.
- There is a “Government e-Marketplace” made to buy government goods. It is available for all government employees. The government can get goods of Rs 50,000.
- There is an online platform launched for the farmer's so that they can access wholesale markets directly. In the beginning, only 12 different farm commodities were started to trade which was increased to 21 after some time. 18 states are covered under this scheme. Near about 7000 markets are covered under this platform.
- With the help of the Ministry of Women and Child Development, the government of India launched a portal named Mahila e-haat for women entrepreneurs. The women entrepreneurs can show their products on the portal free of cost.
- There is a start-up India initiative available for small businessmen. They can enjoy tax benefits, compliance easiness, tracking, etc. they will find it easy to make their traditional business to be online.
- National Skill Development Corporation, Skill India initiated to connect industries with trainers. There is a single-window named ‘Takshashila’ to get information and knowledge that can help them in making their activities more efficient.

VI. INVESTMENTS/INITIATIVES MADE IN E-COMMERCE

Following are the major investments made in the last year i.e. 2021 by the major players of India:

- In September 2021, Amazon launched a platform namely Prime Video in India. The platform is providing unlimited access to popular videos with some charges. The number of subscribers is increasing at a tremendous rate.
- Flipkart has raised funds of US\$ 3.6 billion for further development of E-commerce in India. The website has started a new center in West Bengal in June 2021 that has helped to create nearly 2200 job opportunities for employees. The e-commerce companies are also fulfilling their social responsibilities.
- Master card invested some amount for helping the digitalization process on stores and makes the payment systems more efficient in June 2021.
- In May 2021, Amazon launched a platform namely Mini TV a video streaming service providing app in India. The platform is providing unlimited access to popular videos, comedy shows, food, fashion, and other videos. With some charges. The number of subscribers is increasing at a tremendous rate.
- Flipkart after understanding the increasing online demands for grocery products tried to make the infrastructure of grocery stores and their collection more efficient invested huge money in Delhi, Kolkata, Chennai, Coimbatore, and Hyderabad.
- There was a tie-up In April 2021, between Flipkart and Adani Group to make better company's logistics and increase data center capabilities. They have a target to give 2,500 direct jobs to the unemployed.
- Flipkart purchased Cleartrip, an online travel technology firm in April 2021. It will lead to more online transactions and increase the revenues of the firms.

VII. E-COMMERCE PRESENT STATUS

From the last 2 years after Covid 19 e-commerce has become a need of the hour. The size of e-commerce is going to increase by each day. It is expected to be 25% of the total organized retail market now and will increase to 37 % after 8 years. The reason for such a drastic change is the increasing use of smartphones, the craze of online transactions AMONG generation, easy availability of internet facilities at every place. Daily there is an increase of 10 million active internet users as per research. Half of the transactions being performed in India Are on mobile phones.

VIII. JOB OPPORTUNITIES IN E-COMMERCE

1. *Web developer*

To perform a business online there should be a website that should be properly designed and functioning to design that websites there are web developers who develop websites as per the demands of business houses so web developer for E-Commerce companies is a good career option.

2. *Business analysts*

Every business require a business analyst who can give various analysis of various transactions in the same way E-Commerce transactions also require business analysts that can give or provide analysis on various transactions

3. *E-commerce Project Manager*

E-commerce also requires project managers who can lead their teams in the right direction assign roles to different persons to give responsibilities and have a look at the progress they have made to date if any business want satisfactory results in e-commerce this is the post that they have to create and fill with an efficient and effective person so it is also a trending career choice

4. E-commerce customer service representative

Customer satisfaction is key to success for every E-Commerce company CEO so there is an e-commerce customer service representative who solves complaints and problems of the customers and answers various queries of the customer and satisfied customer is an asset to the business

5. E-commerce order clerks

Every e-commerce company require a person who receives various orders for the products process customers information like the detail of shipping and payment mode the person is known as an order clerk who will manage all the routine activities of buying and selling

IX. CHALLENGES OF E-COMMERCE IN INDIA

Following are the major challenges that E-commerce companies are facing in India

- Most of the customers in India are new buyers they purchase goods from online platforms when the goods are delivered to them they do not find the product as was ordered they returned the goods this is a Universal problem that is

consumer remorse in a country like India. The digitalization process is initiated at a very high speed but still, many buyers prefer cash on delivery payment mode instead of electronic payment as defined online or prepayment as risky.

- One of the reasons for selecting cash on delivery payment mode is the failure of payment gateways where E-Commerce companies are using payment gateways that give that do not give satisfactory results and once a transaction is failed the customer does not reattempt payment again and prefer cash on delivery system.

- The internet connectivity is very poor in various states and areas of the country there are many Areas where internet facility is not easily available or if available then the quality of services is very poor but the government and some private companies are working on the issue and there will be a day when this connectivity issues and internet connectivity problem will be solved.

- Some users are still using feature phones instead of Smartphones on feature phones it is not possible to perform E-Commerce activities.

- When a buyer places an order on the online platform there is a problem they faced is finding the exact location there are many cities and rural areas that are not standardized in the way of Postal addresses which creates a problem for delivering an agent to find the exact location and it is also a problem for the delivery agents when they do not get their product delivered at the right place or sometimes delivered to a wrong.

- There are many areas or towns in India that are not easily accessible it can be due to the large size of the country huge population absence of the right postal codes and negligence of delivery agents sometimes but it all the reasons whatever the reason is it stops or works as an obstacle in the way of the growth of e-commerce.

X. CONCLUSION

The e-commerce industry has very bright future in a developing economy like India despite of many challenges. The concept of e-commerce has made for the buyer and seller easy to perform their roles on online platform which is fast more accurate and time saving. With initiatives of some big private players of the market and government schemes, this market will expand itself to drastic level of increase.

REFERENCES

- [1] International Journal of Commerce and Management Research, Volume 2; Issue 10; October 2016; Page No. 39-41.
- [2] Kumar Anuj1, Fahad Fayaz2, Ms Namita Kapoor3, Impact of E-Commerce in Indian Economy, IOSR Journal of Business and Management (IOSR-JBM) e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 20, Issue 5. Ver. VI. (May. 2018), PP 59-71.
- [3] Dr. R. S. Pawar, Kalyani Dattatraya Ausekar, Dr. R. S. Pawar Kalyani, Dattatraya Ausekar, "A Study of Impact of E-Commerce on Indian

Economy”, Aayushi International Interdisciplinary Research Journal (AIIRJ) VOL- VII ISSUE- X OCTOBER 2020

[4] Arjun Mittal, “E-commerce: Its Impact on Consumer Behavior”, Global Journal of Management and Business Studies, ISSN 2248-9878 Volume 3, Number 2 (2013), pp. 131-138.

[5] <https://www.ibef.org/industry/ecommerce.aspx>

[6] www.ibef.org/industry/ecommerce- presentation. Indian E-commerce Industry Analysis, 2021

[7] https://www.ibef.org as of December 2018

[8] [https://sell.amazon.in/seller-blog/advantages-](https://sell.amazon.in/seller-blog/advantages-of-ecommerce,) of-ecommerce, 2021

[9] [accountlearning.com/limitations-or-disadvantages-](https://accountlearning.com/limitations-or-disadvantages-of-electronic-commerce/) of-electronic-commerce/

[10] <https://www.mockingfish.com/blog/some-of-the-ecommerce-initiatives-taken-by-indian-government-in-various-sectors/>, some of the e-commerce initiatives taken by the Indian Government in various sectors

[11] <https://www.instamojo.com/blog/government-programs-to-accelerate-small-business-e-commerce-growth/>, 5 Government Programs to Accelerate Small Business E-Commerce Growth, Instamojo, March 22, 2021

[12] <https://www.investindia.gov.in/sector/retail-e-commerce/ecommerce#:~:text=After%20a%20surge%20in%20digital,to%20reach%2037%25%20by%202030>. E-Commerce

The fastest-growing e-commerce market in the world

[13] <https://ecommerceguide.com/guides/ecommerce-job-roles/> Types of Job Roles Within Ecommerce

[14] Ajeept Khurana, CEO, Society for Innovation and Entrepreneurship, 8 Challenges for e-commerce in India, 2013

[15] <https://www.pitneybowes.com/us/blog/8-challenges-ecommerce-india.html>,

E-Commerce- SWOT Analysis

Harpreet Kaur

Assistant Professor: B.A.M. Khalsa College,
Garhshankar

E-mail: hkaur96983@gmail.com

Abstract

This paper describes the innovations and accessibility of the internet and powerful instruments which are available online has created a new commerce region called E-commerce. It allocates the purchase and sale of goods on internet and computer networks. E-commerce gives a platform for upgrading economic extension, expanding business opportunities, and also enlarging the market share of an organization. E-commerce is a way which assists the business organization to enhance its market share and thus contributing to economic success. In this paper we will discuss the SWOT Analysis of E-commerce which explains the strengths, weaknesses, opportunities, threats faced by E-commerce in present market structure and we were planned to analyze SWOT of four companies; Amazon.com, eBay.com, Wal-Mart and Target then we discussed it and finally we talk about Potential SWOT of Online shopping e-companies. For collecting data, we searched and used current reports of companies and a few business report organizations. SWOT analysis components describe rivals also.

Keywords: *E-commerce, Strengths, Weaknesses, Opportunities, Threats.*

I. INTRODUCTION

Commerce may be explained as the transaction between the business organization. Business, on the other side, is explained as 'a commercial organization as a going concern'. In doing the business definitely one has to have to communicate with others. Before for Doing business one has to move through a long interval and bring wealth in this pocket all the time ready which every time be a threat for the transporter. It took a long time to grow-up even a sole transaction for example if a producer want to sale his generate he require to sale it to wholesaler first this is also with the assist of intermediary then the whole seller will sale the similar goods the distinct retail and the customer will purchase the goods this series will prefix few extra intermediary if the distance between producer and consumer is more. This process is very time utilizing and also adds little margin in cost at every stage which ultimately enhances the price of the goods and will put a more load on the pocket of the customer or buyer. To fasten the above process and to decrease the add in value in the cost of product

or to decrease the load on the pocket of customer manufacturers only need to use one miracle word that is "e". Here e stands for electronic where ever you add this word the speed will automatically enlarge with many fold then the traditional techniques acquired and on the other side the cost of supply or the price of goods will reduce manifold as contrast to before.

Likewise e-commerce concerns the interchange of money for soft or hard goods and facilities on the internet. Very fascinating definition was given by Kalakota and Whinston 1997. They defined e-commerce in four divergent methods.

1. E-commerce is the delivery of information, goods and services or payments over telecommunication medium, computer networks or any other electronic mode of communication.
2. E-commerce is the application of technology towards the automation of business transactions and work.
3. A tool that addresses the desire of firms, consumers and management to cut service costs while improving the quality of products and services and enhancing the speed of service delivery.
4. Capacity of selling and buying goods and information on the internet and other online services. [1] E-commerce is also known as click business where one uses all the technology of click operation to do a business. Click business as it sounds very easy is not only include click operation it also involves all those technologies which allow the click workable on the internet which involve domain address, server, Web Site, HTML, Electronic Data Interchange, XML, Protocols etc. It is a profitable way to conduct business which goes beyond the simple brick business. Click business can result in fast transactions, wide market coverage along with the number of advantages for instance speed, convenience, cost effectiveness, timeliness, high profit margins, instant consumer relations, no loss of customers and many more. A concern can do everything it can to run its business efficiently and profitably.

II. REVIEW OF LITERATURE

Gupta (2014) in her paper “E-Commerce: Role of e-commerce in today’s business”, presents a comprehensive definition of e-commerce while isolating it from e-business. The paper enlists the different ecommerce models i.e. B2B, B2C, B2G and C2C, narratively analysing the nitty gritties of each. [2]

Rina (2016) also elaborates the different applications of e-commerce in “Challenges and Future Scope of Ecommerce in India”, at the same time, defining the degree to which they are operational in the country. [3]

Gunasekaran, Marri, McGaughey, & Nebhwani (2002) give a broad outlook of electronic commerce within organisational systems in “E-commerce and its impact on operations management”, defining it with reference to e-trading and elaborating- how it has permeated every field of business. The paper identifies the revolutionary role played by earlier internet applications like e-mail and eletronic data interchange and details the revolutionary changes brought by the internet technologies in manufacturing, marketing, purchasing, design, production, selling and distribution, warehousing and human resource management. The paper studies in depth, the significance of web based technologies in different business operations, thus, improving their efficiency through effective B2B e-commerce.[4]

III. RESEARCH METHODOLOGY

In this study the main focus is given on the SWOT Analysis of E-commerce. For this purpose secondary data is collected from different resources like newspapers, journals and also visited on different websites.

IV. OBJECTIVES OF THE STUDY

1. To examine the SWOT Analysis of e-commerce
2. To study the greatest challenges faced by e-commerce.

IV. SWOT ANALYSIS

The E-Commerce market in India has the benefits of phenomenal extension of almost 50% in the last five years. Although the trend of E-Commerce has been making rounds in India for fifteen years, the suitable ecosystem has now started to fall in place. The considerable growth in the number of internet users, growing acceptability of online payments, the proliferation of Internet-enabled devices and

beneficial demographics are the key indicators driving the rise of Ecommerce in the country. Indian e -commerce will reach US \$ 99 billion by 2024, growing at a 27% CAGR over 2019-24, with grocery and fashion and appeal likely to be the key drivers of incremental extension.

Most of the time we see that the use of electronic methods for doing business add value either by decreasing transaction cost or by creating some type of network effect, or by a combination of both. In SWOT analysis (the acronym is short for Strengths, Weaknesses, Opportunities and Threats), here we try to find out the strengths and weaknesses of ecommerce in respect of Indian business environment. Then after we try to identify opportunities presented by that environment and the threats posed by that environment. As appear in the following number which reveals the questions that an analyst would ask in organizing a SWOT analysis:

A. Strengths

- *Global market:* E-commerce great strength is the boundary less access in other word no brick formation is compulsory to do business or no particular boundary is necessary. It enables all the companies to enlarge them to global level. The widening of geographic retail markets may facilitate the development of global retailers
- *Time saving:* Transactions through the internet are no doubt very fast. It saves time by decreasing physical movement.
- *No time constraints:* The concept of 25X7 reveals that online trans can be used anywhere at any time as there are no time constraints.
- *Price/Product comparison:* Information and to select are few of the rights which every consumer has. On the same footing e-commerce gives a platform to customers to contrast price and goods effectively and efficiently. It will tend to have far bigger bargaining effectively and efficiently. It will tend to have far bigger bargaining power with suppliers than traditional local or national retailers.
- *Cost effective:* Elimination of a long chain of intimidators, reducing the requirement of having brick infrastructure and outsource logistics are assisting a small business enterprise to stand at par with giants.
- *Flexible target market segmentation:* The success of business depends on the right

- option of segmentation. Target market segment here in e-commerce is changeable and can be improved *any time*.
- *Fast Exchange of information:* “e” will every time guarantee quick and accurate sharing of information among merchants and customers and enables prompt quick answer.
- *Faster buying procedure:* The purchasing is just a click away from the seller. No physical movement is needed, no hunting of right goods at right price is to done by the customer this create the buying process faster
- *Niche Market:* It is a notion of sub division where the goods of limited types are available without putting in any special effort by customers. Almost Everything can be sold on the internet. Even if goods targeted a smaller market the purchaser will be somewhere on the net.

B. Weaknesses

- *Security:* Security is a great challenge in the progress of e-commerce. Consumer every time found themselves insecure especially about the honesty of the payment process.
- *Fake websites:* Some fake websites are available on the net which promise better service and secure dealing. These websites can not only disgrace ecommerce but also put a bad name to ecommerce.
- *Fraud:* Personal and financial details provided for trading motives are mistreated by hackers for their personal undue interest.
- *Fewer bargaining and discounts:* Hardly online businesses provide bargaining and discounts cannot be possible.
- *Long delivery timing:* The task of Delivery is usually outsourced, who do not care about the timing of the seller. They provide their services as per their own convenience. Sometimes the delivery time may extend to days or weeks which one cannot wait for.
- *Impossibility of physical examination:* Products whose choice merely depend on the physical condition of the product will need personal touch before selection are not suitable for e-commerce business. Online products cannot be touched, worn or sit on the products.
- *Limitation of products:* Only a limited number of products can be available.
- *Lack of personal services:* Physical products can be available but lack interpersonal services which are intangible.

- *Limited exposure:* In developing areas where the internet is not accessible will have no or small submission to e-commerce.
- *Limited advertising:* Limited advertising chances are available because in e-commerce one cannot go for mass advertising. The advertising is required only to computer educated people. And out of them only those who are comfortable with e-commerce applications.
- *Customer's satisfaction:* There is no physical and personal or direct face to face interaction between consumer and the seller. Therefore the scope of convincing the consumer does not exist.

C. Opportunities

- *Changing trends:* People are very brand conscious. They are interested in purchasing branded goods rather than local ones. If such goods are available cross border they will not mind ordering it through e-commerce. E-Commerce is fast and effective even financial transactions can be construct from any component of the world. People of tomorrow will feel more elegant to buy goods through internet only.
- *Increasing number of users:* Daily number of internet users is increasing. People Feel more elegant shopping online.
- *Regular Global expansion:* E commerce can be operated anywhere any time without any interruption. It always has a scope of expansion. All new population and existing population who are not the users of e-commerce are the target expansion.
- *High availability (24 hour and seven days a week):* Along with each and every click of the mouse business is in operation. Those who are busy in day time and cannot spare time for themselves, have all the opportunity to shop as per their convenient time even during late night hours.
- *Wide business growth:* E business has wide scope and broader vision to grow. Business always took place in a gap. Gap filling is a never ending process hence the growth of business is also a never ending process.
- *Advertising:* Advertising is cost effective as compared to conventional offline systems.

D. Threats

- *Competitors:* Along with local competition, global competition also exists. Competition is expanding at an alarming rate. Big

organization have already entered this field. They are making people habitual at the cost of their companies.

- *Changes in environment, law and regulations:* Change in trends, fashion and fad can distress E Commerce side by side change in rule and regulations can also affect it.
- *Innovation:* Consumer nowadays are always in a search of innovative goods and techniques. Innovation will always work as an additional load on the pocket of the consumer, be it in product, place, promotion and even price.

- *Privacy concerns:* Fears that information can be misused lead to spam e-mail or identity fraud.
- *No direct interaction:* In e-commerce there is no direct interaction between customer and the seller. There is no scope of bargaining. People prefer to buy physically as compared to online to experience personal feelings.
- *Fraud:* Persons using unfair means to operate e-commerce can damage the confidence and faith of common people.
- *Risk:* Nature of fraud.

Here, we analyzed 3 online shopping e-companies.

Table 1. SWOT OF E-COMPANIES

Amazon SWOT Analysis	
Strengths	Weaknesses
<ul style="list-style-type: none"> • Cost strategy of leadership • Services and products of higher quality • Purchase Strategic • Chain logistic and effective distribution 	<ul style="list-style-type: none"> • Only attention to online. • Selling with no profit. • Amazon new categories are damaging its brand.
Opportunities	Threats
<ul style="list-style-type: none"> • System for paying online. • To release more its own products and services. • To increase the services and products • To open more international stores. 	<ul style="list-style-type: none"> • Safety of online. • Strategic alliances. • Legislation counters tax. • Retailers with lower regional cost.

eBay SWOT Analysis	
Strengths	Weaknesses
<ul style="list-style-type: none"> • Largest Internet market in the world. • Not a strong competition • Focusing on local markets • System of payment. • Reputation of the mark. 	<ul style="list-style-type: none"> • High expenses. • No other strategy of growth.
Opportunities	Threats
<ul style="list-style-type: none"> • Increasing of mobile customers. • Becoming a retailer market. • services and product growing • Opening more international e-stores. 	<ul style="list-style-type: none"> • Safety of being online. • Retailers with regional low cost. • Intensification of the competition of Amazon. • Rates of exchange.

Wal-Mart SWOT analysis	
Strengths	Weaknesses
<ul style="list-style-type: none"> Sometimes differential product strategy emphasizes on quality beyond discounted Huge market share in US Top gift card seller in US 	<ul style="list-style-type: none"> Neglecting international and focus is just on US Intensive involvement in resolving out the lawsuits. More expensive products than Wal-Mart
Opportunities	Threats
<ul style="list-style-type: none"> The company is expanding to Canada by acquiring the Canadian Zeller's chain. Increasing the private products commendably 	<ul style="list-style-type: none"> Hard competition with K-mart and Wal-Mart because both shares Us services and products.

After different analyses, we came to know that top online shopping companies focus and face more on these factors which are listed below.

- ❖ Strength:
 - Differential product strategy emphasizes on quality beyond discounted.
 - Huge market share in the US.
 - Efficient Distribution chain and logistics.
 - Cost leadership strategy.
- ❖ Weakness:
 - Only online presence.
 - Neglecting international and focus is just on the US.
 - High fees.
- ❖ Opportunities:
 - Release more of its own products and services.
 - Open more international e-stores.
- ❖ Threats:
 - Online safety.
 - Laws against tax avoidance.
 - Regional retailers with lower costs.
 - Very close competition.

The results of research have capacity for an application to the definition of the strategy for the E-companies centers. Using. The potential of internet services is essential in online business and e-entrepreneurship.

V. CHALLENGES FACED BY E-COMMERCE

The E-commerce industry in India is growing at a remarkable pace due to high penetration of the

internet. However, the recent growth rate of e-commerce in India is far lagging behind that of other developed countries. There are many big problems and challenges for e-commerce, these are given as below:

1. There is a lack of knowledge and awareness among people, especially in rural areas, about the use of the internet and various e-commerce websites.
2. The problem with security from threats and frauds is also a big challenge for e-commerce in the whole world.
3. Fear of making online payments is a universal psychological factor in people. With the spread of knowledge on online transactions and its reliability, some %age of customers have overlooked this fear and they are fearlessly engaging themselves in online shopping. But still, the majority of people are not aware about online transactions on the internet as it is very easy to perform tasks on the internet.
4. Majority of customers prefer to buy only those products which have physical existence in front of them. They tend to choose the product by touching the product directly. So those people do not buy products through online markets. This is also a big challenge for e-commerce.

VI. CONCLUSION

A developing country like India has a majority of rural population. There are various people who do not prefer online shopping on the internet. This becomes a major threat to e-commerce. Moreover there are various strengths like we can order products at any time, anywhere. By facing various threats the e-commerce even has a great capture over the whole

world. It also proves very beneficial from a business point of view as businessmen's cost of advertising reduces through online shopping etc.

ACKNOWLEDGEMENT

We would like to thank Prof. Kanwar Kulwant Singh for his encouragement and support.

REFERENCES:

- [1] Kalakota, Ravi, and Andrew B. (1997) Whinston. *Electronic commerce: a manager's guide*. Addison-Wesley Professional,
- [2] Anjali Gupta (2014), E-COMMERCE : ROLE OF E-COMMERCE IN TODAY'S BUSINESS, International Journal of Computing and Corporate Research ISSN (Online) : 2249-054X, Volume 4 Issue 1, International Manuscript ID : 2249054XV4I1012014-10.
- [3] Dr. Rina (2016)" Challenges and Future Scope of E-commerce in India " , International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 5, Issue 2, pp. 232-235 , ISSN 2278-6856.
- [4] Angappa Gunasekaran, Hussain Bux Marri, Ronald Mcgaughey(2002), E-commerce and Its Impact on Operations Management, International Journal of Production Economics 75(1):185-197, DOI:10.1016/S0925-5273(01)00191-8.
- [5] SWOT Analysis of E-Commerce: MsKiranYadav and Dr. Divya Sharma, ISSN 2231-1297, Volume 4, Number 6 (2014), pp. 663-668 © Research India Publications
- [6] Mohammad Hassan, Amir Jahanshahi Sistani and R. Satya (2014) Top Online Shopping E-companies and their Strength and Weakness (SWOT) Seyed, *Research Journal of Recent Sciences* ISSN 2277-2502Vol. 3(9), 102-107.
- [7] <http://alison.com/course/introduction-to-e-commerce>
- [8] <http://en.kioskea.net/contents/207-introduction-to-e-commerce-electronics-commerce>
<http://link.springer.com/articles/10.1007%2fs10660-010-9046-0>
- [9] <http://www.practicalecommerce.com/articles/3971-swot-analysis-for-ecommerce-companies>
- [10] <http://www.awepay.com/ecommerce-swot-analysis/>
- [11] <http://www.emavens.com/blog/ecommerce-2/major-challanges-faced-by-ecommerce-industry-in-india/>

Role of Big Data in E-Business

Gurjinder Kaur

Assistant Professor in Commerce

Department of Commerce and Economics

BAM Khalsa College, Gharshankar

gk41435@gmail.com

Abstract

There has been an increasing inflection on big data analytics (BDA) in e-commerce during recent years. However, it stays poorly-understood as a concept, which blocks its theoretical and practical development. This paper expresses BDA in e-commerce by drawing on a systematic review of the literature. This paper presents an explained framework that explores the definitional aspects, types, distinctive characteristics, business value and challenges of BDA in e-commerce typology. This paper also instigates broader discussions regarding future research challenges and opportunities in theory and practice. Overall, the findings of the study synthesize distinct BDA concepts that furnish great insights along the cross-cutting analytics applications in e-commerce. From the viewpoint of transaction cost theory in e-commerce, BDA can gain online firms by polishing up market transaction cost, managerial transaction cost efficiency and time cost efficiency. The period of big data is now coming. But conventional data analytics may not be able to manipulate such great quantities of data. The question that comes to light is how to develop a high performance platform to accurately analyze big data and how to configure an appropriate mining algorithm to find the useful things from big data? To thoroughly discuss this matter, this paper, set in motion with a quick introduction to data analytics, acts in accordance with the discussions of big data analytics.

Keywords- Big Data Analytics, Video data, voice data

I. INTRODUCTION

In the past few years, an increase of interest in big data has occurred from both education and the e-commerce industry. This explosion is driven by the very fact that e-commerce firms that inject big data analytics (BDA) into their value chain experience 5–6% higher productivity than their competitors (McAfee and Brynjolfsson, 2012)[1]. A recent study by BSA Software Alliance within the US (USA) indicates that BDA contributes to 10% or more of the expansion for 56% of firms(Columbus, 2014)[2]. Therefore, 91% of Fortune 1000 companies are investing in BDA projects, an 85% increase from the previous year(Kiron et al., 2014a)[3].

II. OBJECTIVES:

- To identify definitional perspectives of big data analytics

- To distinguish the characteristics of big data within e-commerce
- To explore the types of big data within e-commerce
- To provide guidelines for tackling the challenges of big data application within e-commerce.

A. Big data and their Features in e-commerce

The e-commerce landscape today is going through the roof with numerous big data that are being used to solve business problems. According to Kauffman et al. (2012, p.85)[4], the use of big data is skyrocketing in e-commerce “due to the internet, social networking, digital telephony and all kinds of new technologies that create and gather data”. With the help of convenient storage and processing capacity, and up-to-date analytical tools, big data now enables e-commerce firms to cut down costs and gains without any difficulty. However, analytics that record big data is different from traditional data in many aspects. Especially, owing to the elements of their distinct nature (i.e., voluminous, variety, velocity, and veracity), big data can be easily distinguished from the traditional form of data used in analytics. Now we will discuss these elements in turn, along with their inference for e-commerce.

Voluminous- With the arrival of web technologies, there is an ever-increasing growth in the amount of big data in the e-commerce environment (Beath et al., 2012)[5]. This huge quantity of data that e-commerce firms are trying to tackle to improve their decision-making process are defined as voluminous (McAfee and Brynjolfsson, 2012)[1]. BDA takes a large quantity of data that requires a huge amount of storage and requires a large number of records. In reality, BDA bound large volumes of data that are used by decision makers for making strategic decisions. Data collected in the big data environment are often unstructured and can incorporate video, image, or data generated from mobile technology. As such, it is unlikely that big data will be clean and free from any errors.

Variety- Variety of big data as it is collected from a wide variety of sources and formats including text, web, tweet, audio, video, click-stream, log files, etc (Russom, 2011)[6]. This variety of data requires the use of different analytical and predictive models which can enable information about different functional areas to be used. This analytic method used for collection of big data by e-commerce organizations can include a range of information for instance consumer profile, buying behavior, buying patterns, optimization of

supply chain operations, and fetch any type of internet data from social media to forecast buying a product, storage and promotional activities.

Velocity- Velocity refers to the density of data creation or the frequency of data confinement. As described by Gentile (2012)[7], the term ‘velocity’ is the frequency of change in big data and how rapidly big data should be used in business settlements in order to add value. In actual fact, that greater data velocity is assured, data has the potential to open up new resorts for organizations. The high velocity of BDA can be useful for analysts to conduct consumer behaviour analysis and give an idea to choose the product/brand.

Veracity- Veracity means the authenticity of big data. How much authorisation or security is being provided to your data. The best quality and security is the foremost requirement for big data that can provide better predictability in e-business (Schroock et al., 2012)[8]. These data demand rigorous verification, requiring full compliance with quality and security issues. Therefore, verification is necessary to generate authenticated and relevant data and to have the capability to screen out bad data (Beulke, 2011)[9] In actual sense, verification is essential in the data management process. Beulke (2011)[9] explained that the IT units of e-businesses are setting up an automatic verification system so that the big data used for business decisions have been authenticated and have passed through strict quality compliance procedures.

B. Types of big data used in e-business

E-business refers to the online transactions: selling goods and services on the internet, either in one transaction (e.g., Amazon, Zappos, eBay, Expedia) or through an ongoing transaction (e.g., Netflix, Match.com, LinkedIn etc.) (Frost and Strauss, 2013)[10]. E-commerce firms ranging from Amazon to Netflix capture various types of data (e.g., orders, baskets, visits, users, referring links, keywords, catalogues browsing, social data), which can be broadly classified into four categories: (a) transaction or business activity data (b) click-stream data (c) video data and (d) voice data. This section discusses different types of big data along with their implications for e-commerce.

- *Transaction or business activity data*

It includes the data related to customer and business relations over a long period of time. This data originated from so many sources like customer relationship programs through to sales transactions. A recent study by Chandrasekaran et al. (2013)[11] provided the example of an e-retailer that analyzes data from its loyalty program (i.e., its Clubcard loyalty program), entailing 1.6 billion data points, 10 million customers, 50,000 stock keeping units (SKUs), and 700 stores, which has resulted in the thorough coordination of big data with consumer insights. Overall, it is evident that e-retailers can derive numerous benefits across the value chain using transaction data.

- *Click-stream data*

Click-stream data is generated from online advertisements or social media content like facebook wall postings, blogs or tweets. All this type of data are very useful for promotional strategies of e-commerce operators, like the use of click-stream data in making informed, strategic, and tactical decisions. Earlier studies have found that many e-commerce firms worldwide (e.g., Amazon, eBay, Zappos, Alibaba etc.) rely on click-stream data in their efforts to capture data. Click-stream data can be applied to foresee customer requirements and tastes. The web data like the reviews of music that are liked, disliked, loved, etc. in order to understand customers' preferences. .

- *Video data*

Video data are live data. Nowadays, e-commerce firms are using not only click-stream data or transaction data but, in connection with image analysis software, they also tend to capture video data. As indicated by Schroock et al. (2012)[8], ecommerce firms have the necessary competencies to analyze extremely unstructured data, such as video or voice data. These data have the potential to add value for e-commerce firms. For example, Ramaswamy (2013)[12] reported that Netflix uses video data to predict viewing habits and evaluate the quality of experiences. In addition, the visualization and demand analytics tool based on the type of movie consumption helps Netflix understand tastes, which leads to success in their “House of Cards” program in the US. Thus, the use of video data is essential for firms in making better decisions than their rivals.

- *Voice Data*

Voice data is typically collected from phone calls, call centers, or customer service centers. Voice data are advantageous for analyzing consumer-buying behavior or spotting prospective customers. As described by Davenport et al. (2012)[13], credit card companies, for example American Express, use and track data related to call center activities so that personalized offers can be given in milliseconds. In Schroock et al. 's (2012)[8] survey, e-commerce firms were found to use advanced capabilities to analyze text and transcripts converted from call center transactions. Additionally, numerous delicacies of language, such as sentiment, slang and intentions, can be read and recognized by means of BDA in the context of e-commerce.

III. BIG DATA ANALYTICS

The problem of handling a vast quantity of data that the system is unable to process is not a brand-new research issue; in fact, it appeared in several early approaches[16, 17, 18] e.g., marketing analysis, network flow monitor, gene expression analysis, weather forecast, and even astronomy analysis. This problem still exists in big data analytics today; thus, preprocessing is an important task to make the computer, platform, and analysis algorithm be able to handle the input data. The traditional data preprocessing methods (e.g., compression, sampling, feature selection, and so on) are expected to be able to operate effectively in the big data age.

However, a portion of the studies still focus on how to reduce the complexity of the input data because even the most advanced computer technology cannot efficiently process the whole input data by using a single machine in most cases. By using domain knowledge to design the preprocessing operator is a possible solution for big data. In, Ham and Lee used the domain knowledge, B-tree, divide-and-conquer to filter the unrelated log information for the mobile web log analysis. A later study [19] considered that the computation cost of pre-processing will be quite high for massive logs, sensors, or marketing data analysis. Thus, Dawelbeit and McCrindle employed the bin packing partitioning method to divide the input data between the computing processors to handle this high computations of pre-processing on cloud systems. The cloud system is employed to preprocess the raw data and then output the refined data (e.g., data with uniform format) to make it easier for the data analysis method or system to perform the further analysis work. How to present the analysis results to a user is another important work in the output part of big data analytics because if the user cannot easily understand the meaning of the results, the results will be entirely useless. Business intelligence and network monitoring are the two common approaches because their user interface plays the vital role of making them workable. Zhang et al.(2012)[20] pointed out that the tasks of the visual analytics for commercial systems can be divided into four categories which are exploration, dashboards, reporting, and alerting. The study [21] showed that the interface for electroencephalography (EEG) interpretation is another noticeable research issue in big data analytics. The user interface for cloud system [22, 23] is the recent trend for big data analytics. This usually plays vital roles in big data analytics system, one of which is to simplify the explanation of the needed knowledge to the users while the other is to make it easier for the users to handle the data analytics system to work with their opinions. According to our observations, a flexible user interface is needed because although the big data analytics can help us to find some hidden information, the information found usually is not knowledge. This situation is just like the example we mentioned in “Output the result”. The mining or statistical techniques can be employed to know the flu situation of each region, but data scientists sometimes need additional ways to display the information to find out the knowledge they need or to prove their assumption. Thus, the user interface can be adjusted by the user to display the knowledge that is needed urgently for big data analytics.

IV. THE OPEN ISSUES

Although the data analytics today may be inefficient for big data caused by the environment, devices, systems, and even problems that are quite different from traditional mining problems, because several characteristics of big data also exist in the traditional data analytics. Several open issues caused by the big data will be addressed as the platform/framework and data mining perspectives in this section to explain what dilemmas we may confront because of big data.

A. Input and output ratio of platform

A large number of reports and researches mentioned that we will enter the big data age in the near future. Some of them insinuated to us that these fruitful results of big data will lead us to a whole new world where “everything” is possible; therefore, the big data analytics will be an omniscient and omnipotent system. From the pragmatic perspective, the big data analytics is indeed useful and has many possibilities which can help us more accurately understand the so-called “things.” However, the situation in most studies of big data analytics is that they argue that the results of big data are valuable, but the business models of most big data analytics are not clear. The fact is that assuming we have infinite computing resources for big data analytics is a thoroughly impracticable plan, the input and output ratio (e.g., return on investment) will need to be taken into account before an organization constructs the big data analytics center.

B. Communication between systems

Since most big data analytics systems will be designed for parallel computing, and they typically will work on other systems (e.g., cloud platform) or work with other systems (e.g., search engine or knowledge base), the communication between the big data analytics and other systems will strongly impact the performance of the whole process of KDD. The first research issue for the communication is that the communication cost will incur between systems of data analytics. How to reduce the communication cost will be the very first thing that the data scientists need to care. Another research issue for the communication is how the big data analytics communicates with other systems. The consistency of data between different systems, modules, and operators is also an important open issue on the communication between systems. Because the communication will appear more frequently between systems of big data analytics, how to reduce the cost of communication and how to make the communication between these systems as reliable as possible will be the two important open issues for big data analytics.

C. Bottlenecks on data analytics system

The bottlenecks will be appeared in different places of the data analytics for big data because the environments, systems, and input data have changed which are different from the traditional data analytics. The data deluge of big data will fill up the “input” system of data analytics, and it will also increase the computation load of the data “analysis” system. This situation is just like the torrent of water (i.e., data deluge) rushed down the mountain (i.e., data analytics), how to split it and how to avoid it flowing into a narrow place (e.g., the operator is not able to handle the input data) will be the most important things to avoid the bottlenecks in data analytics system. One of the current solutions to the avoidance of bottlenecks on a data analytics system is to add more computation resources while the other is to split the analysis

works to different computation nodes. A complete consideration for the whole data analytics to avoid the bottlenecks of that kind of analytics system is still needed for big data.

D. Security issues

Since much more environmental data and human behavior will be gathered to the big data analytics, how to protect them will also be an open issue because without a secure way to handle the collected data, the big data analytics cannot be a reliable system. In spite of the security that we have to tighten for big data analytics before it can gather more data from everywhere, the fact is that until now, there are still not many studies focusing on the security issues of the big data analytics. According to our observation, the security issues of big data analytics can be divided into fourfold: input, data analysis, output, and communication with other systems. For the input, it can be regarded as the data gathering which is relevant to the sensor, the handheld devices, and even the devices of internet of things. One of the important security issues on the input part of big data analytics is to make sure that the sensors will not be compromised by the attacks. For the analysis and input, it can be regarded as the security problem of such a system. For communication with other system, the security problem is on the communications between big data analytics and other external systems. Because of these latent problems, security has become one of the open issues of big data analytics.

E. Privacy issues

The privacy concern typically will make most people uncomfortable, especially if systems cannot guarantee that their personal information will not be accessed by other people and organizations. Different from the concern of security, the privacy issue is about if it is possible for the system to restore or infer personal information from the results of big data analytics, even though the input data are anonymous. The privacy issue has become a very important issue because the data mining and other analysis technologies will be widely used in big data analytics, the private information may be exposed to the other people after the analysis process. For example, although all the gathered data for shop behavior are anonymous (e.g., buying a pistol), because the data can be easily collected by different devices and systems (e.g., location of the shop and age of the buyer), a data mining algorithm can easily infer who bought this pistol. More precisely, the data analytics is able to reduce the scope of the database because location of the shop and age of the buyer provide the information to help the system find out possible persons. For this reason, any sensitive information needs to be carefully protected and used. The anonymous, temporary identification, and encryption are the representative technologies for privacy of data analytics, but the critical factor is how to use, what to use, and why to use the collected data on big data analytics.

ACKNOWLEDGMENTS

I would like to thank Prof. Kanwar Kulwant Singh for his encouragement and support. Also would like to thank the anonymous reviewers for their valuable comments and suggestions on the paper.

REFERENCES

- [1] McAfee, A., Brynjolfsson, E., 2012. Big data: the management revolution. Harvard business review.
- [2] Columbus, L., 2014. Making Analytics Accountable: 56% Of Executives Expect Analytics to Contribute To 10% Or More Growth in 2014, Forbes.
- [3] Kiron, D., Prentice, P.K., Ferguson, R.B., 2014a. The analytics mandate. MIT Sloan management review 55, 1- 25.
- [4] Kauffman, R.J., Srivastava, J., Vayghan, J., 2012. Business and data analytics: New innovations for the management of e-commerce. Electronic Commerce Research and Applications 11, 85-88.
- [5] Beath, C., Becerra-Fernandez, I., Ross, J., Short, J., 2012. Finding Value in the Information Explosion. MIT Sloan Management Review 53, 18-20.
- [6] Russon, P., 2011. The Three Vs of Big Data Analytics. TDWI.
- [7] Gentile, B., 2012. Top 5 Myths About Big Data.
- [8] Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., Tufano, P.P., 2012. Analytics: The real-world use of big data. IBM Institute for Business Value, NY, USA.
- [9] Beulke, D., 2011. Big Data Impacts Data Management: The 5 Vs of Big Data.
- [10] Frost, R., Strauss, J., 2013. e-Marketing. Pearson Higher Ed.
- [11] Chandrasekaran, S., Levin, R., Patel, H., Roberts, R., 2013. Winning with IT in consumer packaged goods: Seven trends transforming the role of the CIO. McKinsey & Company, pp. 1-8
- [12] Ramaswamy, S., 2013. What the Companies Winning at Big Data Do Differently. Bloomberg, June: <http://www.bloomberg.com/news/2013-06-25/what-the-companies-winning-at-big-data-do-differently.html>.
- [13] Davenport, T.H., Barth, P., Bean, R., 2012. How ‘Big Data’ is Different. MIT Sloan Management Review 54, 43- 46.
- [14] Tsai, C.W., Lai, C.F., Chao, H.C. et al. Big data analytics: a survey. Journal of Big Data 2, 21 (2015).<https://doi.org/10.1186/s40537-015-0030-3>
- [15] Akter, S., Wamba, S.F. Big data analytics in E-commerce: a systematic review and agenda for future research. Electron Markets 26, 173–194 (2016). <https://doi.org/10.1007/s12525-016-0219-0>
- [16] Xu R, Wunsch D. Clustering. Hoboken: Wiley-IEEE Press; 2009
- [17] Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases. Proc ACM SIGMOD Int Conf Manag Data. 1993;22(2):207-16.
- [18] Lee J, Hong S, Lee JH. An efficient prediction for heavy rain from big weather data using genetic algorithm. In: Proceedings of the International Conference on Ubiquitous Information Management and Communication, 2014, pp 25:1–25:7.
- [19] Ham YJ, Lee H-W. International journal of advances in soft computing and its applications. Calc Paralleles Reseauxet Syst Repar. 2014;6(1):1–18.
- [20] Zhang L, Stoffel A, Behrisch M, Mittelstadt S, Schreck T, Pompl R, Weber S, Last H, Keim D. Visual analytics for the big data era—a comparative review of state-of-the-art commercial systems. In: Proceedings of the IEEE Conference on Visual Analytics Science and Technology, 2012. pp 173–182.
- [21] Harati A, Lopez S, Obeid I, Picone J, Jacobson M, Tobochnik S. The TUH EEG CORPUS: A big data resource for automated eeg interpretation. In: Proceeding of the IEEE Signal Processing in Medicine and Biology Symposium, 2014. pp 1–5.
- [22] Thusoo A, Sarma JS, Jain N, Shao Z, Chakka P, Anthony S, Liu H, Wyckoff P, Murthy R. Hive: a warehousing solution over a map-reduce framework. Proc VLDB Endowment. 2009;2(2):1626–9.

- [23] Beckmann M, Ebecken NFF, de Lima BSLP, Costa MA. A user interface for big data with rapidminer. RapidMiner World, Boston, MA, Tech. Rep., 2014. [Online]. Available: <http://www.slideshare.net/RapidMiner/a-user-interface-for-big-data-with-rapidminer-marcelo-beckmann>.

Implementation of Cold Store Monitoring System (CSMS) Software in Cold Store Management System

Mohit Kumar^{1,a}, Dr. Ajay Dutta^{2,b} Dr. Rajneesh Kant Sachdeva^{3,c}

¹ Post Graduate Student, Department of Food Science & Technology, I.K Gujral Punjab Technical University, Kapurthala, Punjab, India

² Assistant Professor, Department of Computer Science & Engineering, Babbar Akali Memorial Khalsa College, Garhshankar, Punjab, India

³ Head Of Department, Department of Food Science & Technology, I.K Gujral Punjab Technical University, Kapurthala, Punjab, India

- a) Author: - Mohitajoria5@gmail.com
- b) Drajaydutta13@gmail.com
- c) pturajneesh@ptu.ac.in

Abstract "A Cold Store is a Specialised Area to Minimize the respiration rate and increase the life span of the product for future use."

This paper includes the cold storage capacity of India and the condition of cold storage and implementation of the Cold Store management system (CSMS) Software. This software helps to remove all the technological gaps that are present in the cold store system. Loss of the Agricultural waste is major issue due to the improper management of the cold store management. "Cold storage is an integral component of the postharvest management system for mainly Fruits and Vegetables. "According to a United Report International Development Program, about 40% of India's total production is lost on consumer-oriented travel or it's broken." [10] The report provides an overview of the cold storage areas in India including current storage strengths, styles, and Government programs such as initiation, growth potential, challenges, and key achievements of the industry. The role of the Cold Store Monitoring system (CSMS)

Software is very helpful in cold storage management. This paper is focused on implementation of Our Software in cold storage sector in India. The Cold Store Monitoring System (CSMS) provides greater visibility and control by improving the entire Cold Chain network running daily across India. This literature review is created by reviewing a good number of papers published in peer-reviewed journals and online sources using second-hand data

obtained. The purpose of the study is to Implementation of the Software to the Cold Chain System and to obtain research opportunities.

I. INTRODUCTION

India ranks number 2 in the list of countries by population. "India's arable land area of 159.7 million hectares (394.6 million acres) is the second largest in the world, after the United States (B.P Sharma 2016). Its gross irrigated crop area of 82.6 million hectares (215.6 million acres) (B.P Sharma 2016) is the largest in the world. India is the second-largest producer of vegetables and fruits after China. India on average produces 85 million tons of vegetables and 45 million tons of fruit annually. But India's overall contribution to world trade is far from ideal as the country loses 30 to 40 percent of its fruit and vegetables due to improper cold storage facilities. Also, the cold cell area in India is very concentrated in a few provinces. Uttar Pradesh, West Bengal, Punjab, Gujarat, etc., is one of the coldest places in India. Like Punjab, with its "extremely cold" storage capacity" [1].

According to a 2017 report by the Associated Chambers of Commerce and Industry of India (ASSOCHAM), India's 150 billion-square-foot [150 billion m] storage capacity is the largest in the world, followed by the USA and China (V. Paul 2016).

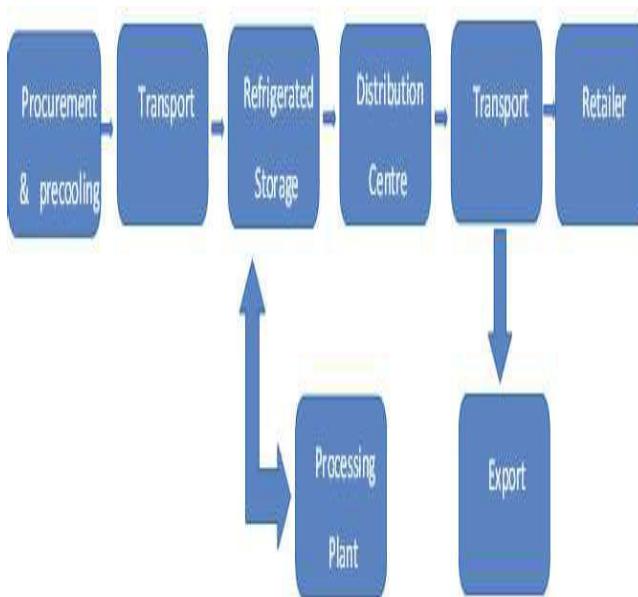
1.1 Cold chain

Cold chain refers to the management of temperature-sensitive product because it travels through a series. It's a term accustomed to describing a series of activities that depend upon the assembly, distribution, storage, and re-sale of the decayable product.

Cold chain supplying could be a system of materials that give an honest atmosphere for a warm and decaying hot waste product from the purpose of production to the extent of distribution by refrigeration reefers and security arrangements to safeguard the integrity of those shipments. Simply put, a chilly chain could be a provide chain controlled by temperatures. It's accustomed to facilitating extend and guarantee period of a product like a new agricultural product, seafood, frozen foods, photographic material, chemicals, and prescription drugs.

The main objectives of Cold Chain Management are to improve product quality and product safety and reduce waste. Operation Cold Waste Management Means the constant temperature management in each step within the production, storage, and transport at intermediate and intermediate levels. With the growing demand for fast food, ready-to-eat foods and the organization of frozen products seeks better cold solutions. Many temperatures suit a variety of factors. "Frozen (-18 ° C to 23 ° C), Cold (0 ° C to 4 ° C) and Cold (10 ° C to 18 ° C) are some of the most frequently used clips with a specified product range, depending on the product, the weather is meat, or ice cream or potatoes" [28]. The main feature of the cold chain is that when there are no links the whole system fails.

1.2 Process Flow of Cold Chain



A typical cold chain network comprises of following stages.

1.3 Information & Communication Technology (ICT)

Information and Communication Technologies (ICT) referring to all communications technologies, including the Internet, middleware, software, computers, video-conferencing, wireless networks, social media, and other media applications and services that allow users to access, retrieve, store, transmit, and manipulate data in digital form [30].

Through the use of ICTs, there are already "a wide variety of new developments in the agricultural sector, including data and stock market analysis, data collection, extension services for agricultural extension services, early disaster risk management and control programs, financial services, agricultural product tracking, agricultural statistical data collection, etc. [30] (Sustainable Agriculture ICT, FAO, 2013).

In 2007, FAO and a group of founding partners started the e-Agriculture Community of Practice - sharing knowledge and experiences of projects that use ICT for agriculture and rural development (e-agriculture 10-year review report). - There is an online location to facilitate delivery. Implementation of World Summit on Information Society / WSIS, 2015) [30].

1.4.1 Components of ICT

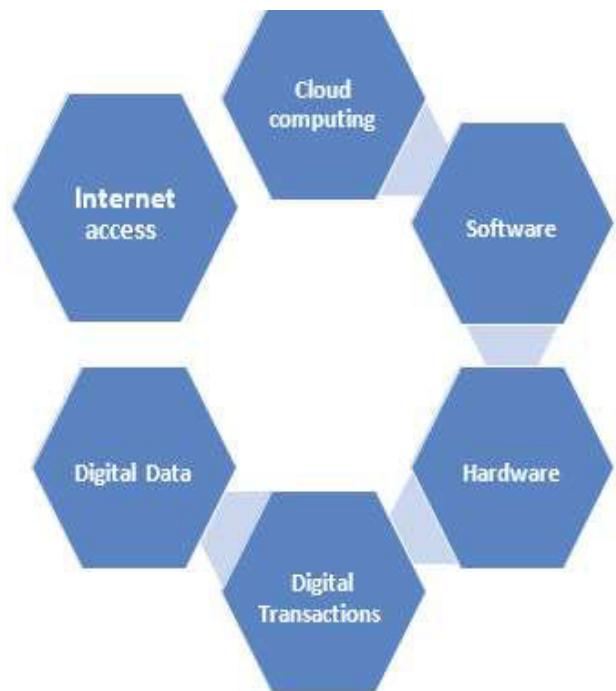


Fig 1.1 Components of ICT(M.K Pratt 2021)

2.1 Objective

The intention of research is to apply and develop innovative techniques and methods for the enhancement of the monitoring and tracking the cold chain perishable food products.

1. Use of information and communication technologies in the cold chain.
2. Real time monitoring of the temperature inside the cold store using temperature sensors.
3. Optimization of cold store space, due to real-time control of inventory management.
4. Step by step processes followed after installation of information & communication technology which may reduce wastage.

2.2 Need for Cold Storage

One of the biggest problems Indian farmers face in selling their produce is storing and shipping it on time. Farmers/food processors are having difficulty finding a place near their farms or a convenient place to deliver cold to get their products from the farm to the market in the right conditions and at the right time. When there are a lot of products and food on the market and prices are plummeting. To get better prices, farmers have to move the product to larger areas that use more.

Large cities that waste major cities and food processing industries are often thousands of miles away. However, exporting agricultural products throughout the length and breadth of the country or exporting through Cold Chain Logistics with a task is not an easy task. Unripe products tend to rot and must be stored until time to get better prices or move them over long distances. In addition, changing lifestyles and demand for processed or integrated foods creates the need for global cold solutions.

The initial investment is very high in this business. Typically, the investment includes acquiring land, make buildings, obtaining permits, obtaining licenses, arranging services such as water, electricity, etc. In addition, you will need to invest a high initial investment in cooling equipment. In general, modern and updated equipment ensures better performance and higher durability.

3 Literature Review

The literature on the economic aspect of the cold storage industry had a profound effect on the availability of relevant information in this area. However, a serious effort has been made to dig out the necessary details of the relevant literature in the framework of this analysis from all sources available in the country as well as from abroad.

3.1 Cold storage Status in India

The Indian government is promoting the building of cold-blooded infrastructure. Progress so far has been seen in the last post, mainly based on earlier successes of storing potatoes every year. In some horticultural crops, the point of production infrastructure in the form of pre-coolers and pack-houses must act as initiators of cold movements. It is well known that food processing units as a food processing facility have been used successfully (Hou, Y-F, Xie, D & Wang, J-B 2015). Without such manufacturing units, goods could not be frozen and because of domestic production passing through our cold climates, we have imported a product that is successfully supplied.

If a cold retention, the two main types continue. The first is the farm gate infrastructure planted near production districts to sustain long-term farm production. The latter types are reserved for the storage of a single asset when the product is subject to later sale for an extended period of its sale. These shops are mainly available in potato shops and those for spices and certain vegetables such as carrots, apples, oranges, onions, etc. CA stores, due to their long-term shelf life, fall under this category (Rossetti, MD 2015). The product stored on that cold street is not designed for duplicate handling and these stores can serve as starting points for later market links.

The second type of cold store, it goes slower than the type of service they offer. These are the distribution hubs near the market, the ones that are most likely to be eaten (grocery stores and home refrigerators), the storage areas (E.g. grapes and those installed in packing houses), etc. The integrity of cold materials depends largely on the transport infrastructure including equipment. The goods in these stores are those that are healthy on the short shelf, those that are out of stock, and those that are on the way to marketing or end use. Most horticultural plants are healthy on a limited shelf even if they are in cold conditions and

cannot be stored during the growing season. In addition, all products by the end of their life, need to move to the market with those cold resources. Such final infrastructure has recently been built in India.

Among the unfamiliar people all cold stores look the same, but the cold storage gate infrastructure should not be confused with other forms of cold, especially when it comes to the cold-water network that connects the farm to the markets.

According to the latest reports, there are about 6,000 cold stores in India (ASSOCHAM Report 2020), most notably the long-standing farm gate, which was originally built to store just one item. Over the past decade, a larger market connected to the type of transportation has improved. Without these resources, the chain is cut short as the product exits the production gates of the farm gates if there is no proximity to market areas.

Packing houses with pre-coolers with insufficient storage in India. Reefer transport that connects advanced processing units with cold resources is there for commercialization that is in the category we should have (ice cream, milk, meat and imported into the cold chain) (NCCDB 2019).

3.2 Activities During Storage

Almost all of the cold storage is private in which some are designed under some government schemes like NCCD, MoFPI, NABARD etc. are some government bodies which gives the guidelines and pension to the establishment of the cold store. The owner of the cold store is itself have some land and they grow their own fruits and vegetables on it. so they store their own commodities that are grown by the owner. Let's take a example of the potato in India the almost all of the cold store are meant to store only potato because India is the third largest producer of the potato after RUSSIA and CHINA.

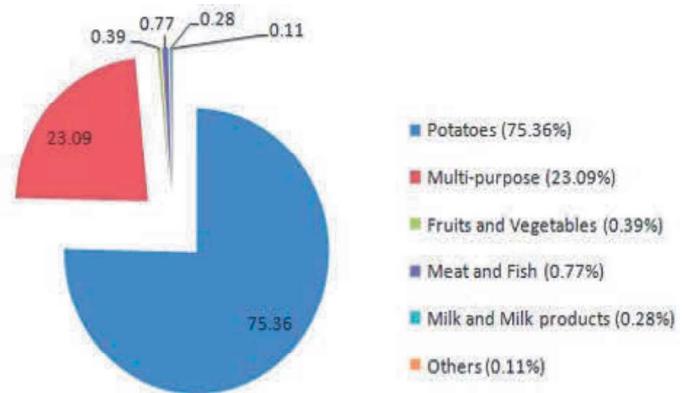


Figure 3.1.: Capacity of Cold Storage based on Type of Commodity

3.3 Technologies Used by U.S. Cold Store

3.3.1 eUSCOLD: Customer Online Portal

EUSCOLD is our safe, secure customer management environment for setting up and tracking order across all our businesses.

EUSCOLD fully integrates with our WMS and TMS systems, giving customers 24/7 access to inventory details, shipping status, order placement and custom claim reports (ICT Infrastructure of Warehouse US 2019).

3.3.2 Warehouse Management System

Our WMS is an intermediate customer service program designed to accommodate any type of customer business needs.

USCS has started releasing its brand-new WMS which will be called eWM. EWM is built on cutting edge, stand-alone technology and an ultramodern database to provide a wide range of advanced features, support USCS business growth and keep our warehouses competitive and driven by technology (ICT Infrastructure of Warehouse US 2019).

3.3.3 Transportation Management System

USCS uses Mercury-Gate, a world-class system, a single system, and a multidisciplinary transport system that enables our business to make quick decisions with confidence. The TMS software provides the ability to purchase, organize, use, deploy and resolve cargo movements in full view of the tower. The design of its system makes it easy to integrate with our ERP financial system and the WMS patent system. Add to the carrier company's experience by riding fast using its embedded EDI solutions (ICT Infrastructure of Warehouse US 2019).

3.3.4 Electronic Data Exchange

Electronic Data Exchange (EDI) computer-to-computer document exchange in a standard format among business partners (ICT Infrastructure of Warehouse US 2019). From the paperbased exchange of business document to the electronic version, businesses enjoy significant benefits such as reduced costs, increased repair speed, reduced errors and improved relationships with business partners.

Electronic Data Exchange allows us to share and retrieve information in a variety of ways so that you can get information quickly, efficiently, and securely whenever you need it.

USCS Internal EDI solutions and business-saving talent offers great flexibility and expertise to integrate seamlessly with its clients' ERP programs. USCS supports more than 30 final documents and transit documents (ICT Infrastructure of Warehouse US 2019).

3.3.5 COLD-Key

All pilot is provided with a Cold-Key that they can enter through the gate with this key card. This helps to keep their entry times faster in the warehouse (ICT Infrastructure of Warehouse US 2019).

Cold-Key allows the use of smartphone before reaching the warehouse gate and generates a barcode that can be scanned at

the gate and speed up the installation. this is a very unique approach followed by the US cold store.

4. DEVELOPMENT OF THE BASE MODEL OF OUR SOFTWARE

This chapter describes the approach used in this study to create the base model for our software named Cold store monitoring system (CSMS) with following the procedures. These procedures include system definition, model building, and model verification and validation [26]. In addition, this chapter presents the outputs of the base model and the validation results to verify that the base model is Representative of the actual system [26]. The Output of the base model in throughput, processing time, shelf life, operating cost, and wastage, using the mean values of 100 replications, are compared against the actual values of the system [26].

4.1 Tools Used

Tools	Purpose
Microsoft Visual Studio framework	System Development
Microsoft SQL Server	Database Management
C#	Programming Language
Java Script	Data Validation
Internet Information System (IIS)	Web Server
Windows 10	Operating System
Microsoft Visio	System Modelling
Browser (Any)	User Interface
Cascading Style Sheets (CSS)	Designing Interface

Table: Tools used for developing the proposed system

4.2 Process flow diagram

In this step, the actual-world system is converted into a logical representation of the processes and activities involved. The conceptual model represents the basic concepts of actual system of a potato supply chain with farms, processors (Cold Store), transporter, and middleman firm [26]. Figure 4.4 shows the Diagrammatic Representation of the simulation model and the description of the business mapping diagram for the potato supply chain activities, which are presented in Table 4.2.

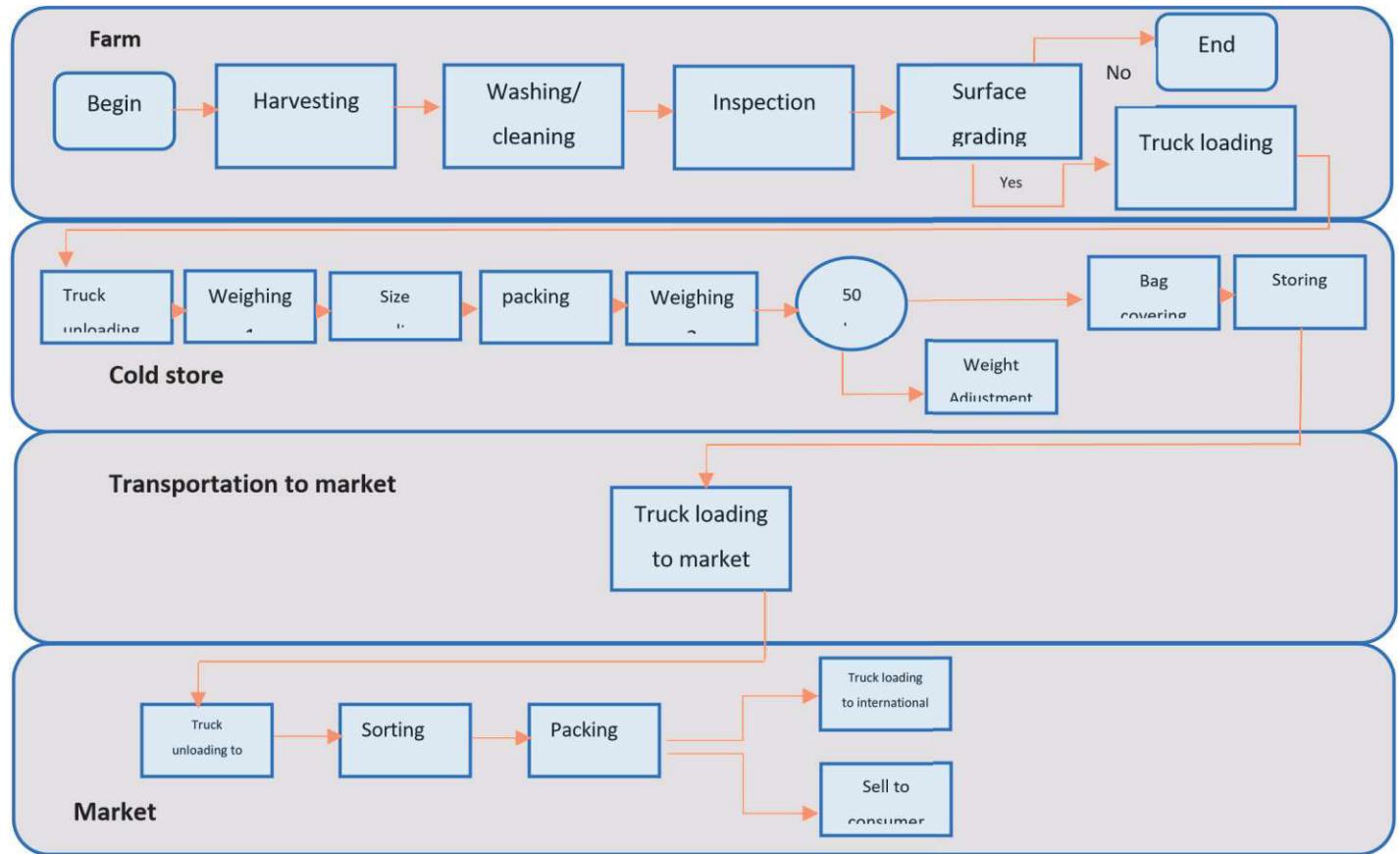
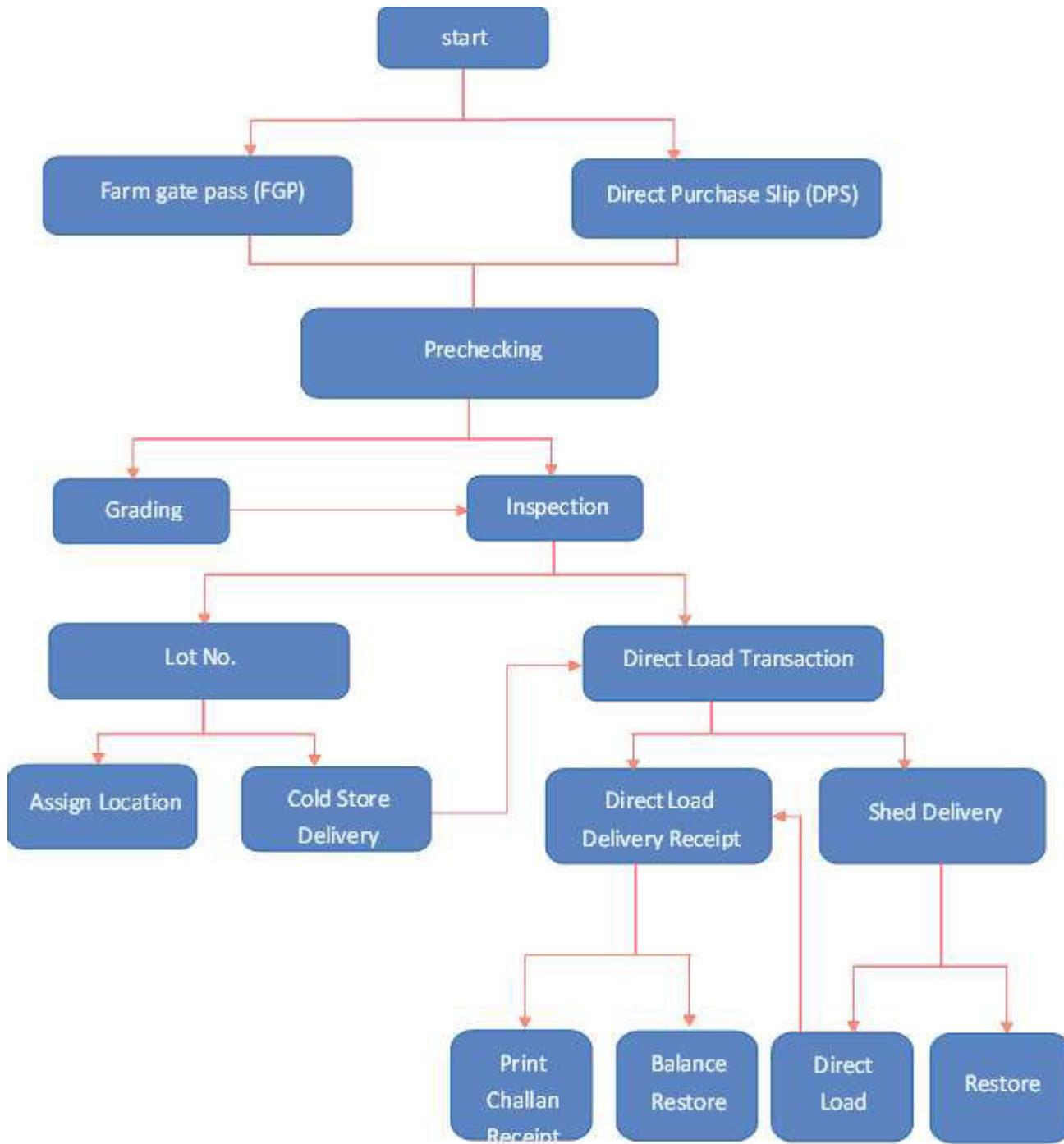


Figure 5.1 – Diagrammatic Representation of Potato Process Flow.

4.3 Overview of the CSMS (cold store monitoring system) Software

During the research there are many technological and non- technological gaps are present so implementation of Information & communication technology in the cold store by making a software CSMS (cold store monitoring system). This is an online software in which all processes of cold store from post-harvest to transportation of potato from cold store to mandis are monitored. There is very easy to keep record due to this system. This system can access by farmer login or customer login and second for cold store. Farmer can check the status of their productivity that are stored in the cold store.

Diagrammatic view of the software



5. Application of the Cold Store Monitoring System (CSMS)

With the continued growth of the world economy, cold storage has become more widely used to store fruits and vegetables. Therefore, the use of the Cold Store Monitoring System

(CSMS) plays an important role in maintained cold store, not only ensuring that the system gets a safe and stable operation, improves cold chain performance, but also saves resources and reduces costs. Therefore, the benefits of a cold store are as follows:-

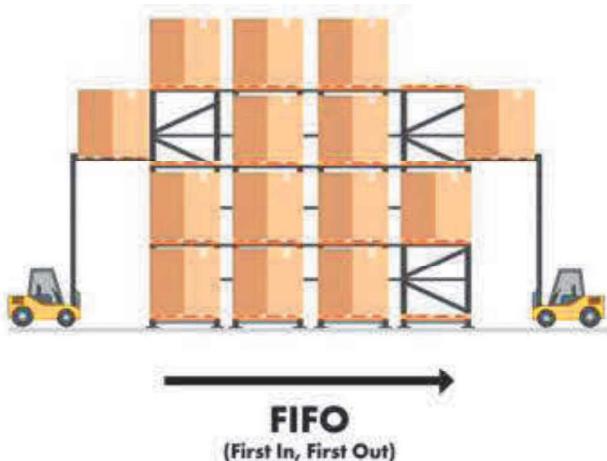
5.1 Temperature monitoring: - To keep the exact temperature inside the chamber is very important for the storage of the potato. So, installation of the temperature sensors inside the cold chambers then the temperature is watched outside the chamber with the help of the cold store monitoring system. Cold store owner can check the temperature from anywhere. It also eliminates the human interventions from going inside the cold chambers to check the temperature. Sensors can detect the weather, rise in temperature or fall in temperature to the actual temperature needed due to any type of breakdown like in power or in the cooling system.

5.2 Inventory accuracy: - due to this cold store monitoring system will help in support the inventory accuracy rate. We can prompt check the how much space is required or how much space is filled in each of the chambers. When bags are ready after the grading the location is assigned to that batch of the farmer and place it where it assigned.

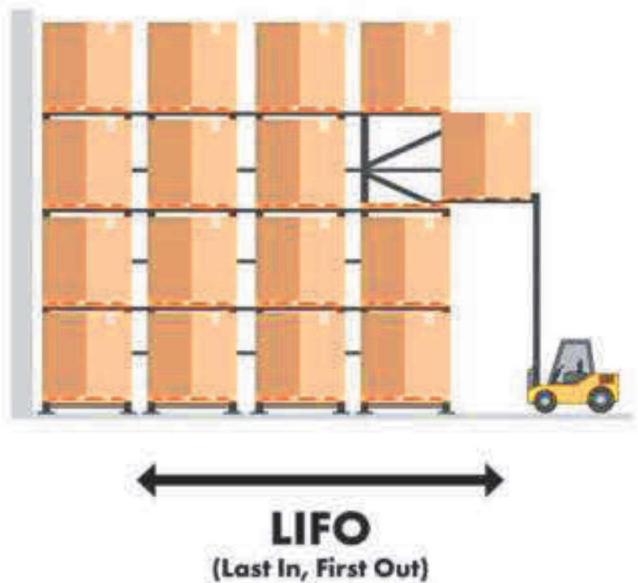
5.3 Record keeping: - After installation of cold store monitoring system in the cold store the record should be kept in this CSMS. The list of all the farmers who stored their potato in that cold store should be stored. The detail of each farmer like no. of bags of, amount paid or not etc. kept in the system. All the data of the cold store are kept in this online system.

5.4 Expiration date control: - We can also control the expiration date whether product is fresh or not through this software by checking such as FIFO (first in first out), LIFO (last in first out) and LEFO etc.

FIFO (First In, First Out) - The flow of goods where the first thing is the first thing to leave the warehouse or solution. A stockpile box / box is entered, and the first one inside is the first to leave the warehouse.



LIFO (Last In, First Out) - The flow of goods operating under the assumption that the bag / storage box loaded into the racking system is the first to be distributed to the warehouse.



Conclusion

Cold chain adoption is a well-recognized practice in developed countries, such as the USA and EU countries, for fresh potato supply chains, to lengthen shelf life and reduce wastage" (K. Mongkut & T. Thonburi 2018). "However, due to hot weather and lack of cold chain management, fresh potato supply chains in India is suffer from issues such as short shelf life, high wastage, and low quality. Implementation of cold chains in India is limited because of the high cost of cold chain technology, lack of equipment and infrastructure, inadequate knowledge and training, lack of collaboration, and absence of government support. To promote cold chain adoption in developing countries, this study examined an alternative cold chain design for fresh potato supply chains by borrowing insights from the Information and Communication Technology (ICT) (K. Mongkut & T. Thonburi 2018).

The goal of the research become to construct a device for preserving the tune of all Cold Storage activities. The integrated device can meet all of the simple needs. It will offer a facility for the bloodless shop to be capable of hold tune to all of the shares offered. Asset Management can even enjoy the proposed device, because it will streamline the whole delivery chain, with a purpose to lessen the burden. System safety and decrease waste are likewise certainly considered one among the most important concerns.

There is usually room for development in any software, regardless of how the device can also additionally work. The vital part is that the device must be bendy sufficient for destiny repairs. The device has been incorporated into diverse modules to make the device well suited with different changes. Every attempt has been made to cater to all the wishes of the person and make it clean to use.

References

- [1] B. P. Sharma, "Present Position of Agriculture in India," *Int. J. Sci. Res.*, vol. 5, no.4, pp. 240–243, 2016, doi: 10.21275/v5i4.nov162488.
- [2] V. Paul, "~ oII ~ J PR," no. May 2016.
- [3] M. K. Pratt, "ICT (information and communications technology, or technologies)," pp. 1–7, 2021.
- [4] T. Ford, *A Technology Challenge*, vol. 62, no. 11. 1990.
- [5] L. Ruiz-Garcia and L. Lunadei, "The role of RFID in agriculture: Applications, limitations and challenges," *Computer. Electron. Agric.*, vol. 79, no. 1, pp. 42–50, 2011, doi: 10.1016/j.compag.2011.08.010.
- [6] R. Badia-Melis, P. Mishra, and L. Ruiz-García, "Food traceability: New trends and recent advances. A review," *Food Control*, vol. 57, pp. 393–401, 2015, doi: 10.1016/j.foodcont.2015.05.005.
- [7] Y. Zou, J. D. Clark, and A. A. May, "A systematic investigation on the effects of temperature and relative humidity on the performance of eight low-cost particle sensors and devices," *J. Aerosol Sci.*, vol. 152, p. 105715, 2021, doi: 10.1016/j.jaerosci.2020.105715.
- [8] A. Dada and F. Thiesse, "Sensor applications in the supply chain: The example of quality-based issuing of perishables," *Lect. Notes Computer. Sci. (including Subset. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4952 LNCS, pp. 140–154, 2008, doi: 10.1007/978-3-540-78731-0_9.
- [9] N. America, C. Storage, and M. Size, "North America Cold Storage Market Size, Share & Trends COVID-19 Impact Analysis," pp. 2021–2028, 2021.
- [10] C. T. Infrastructure, "UKDiss.com," pp. 1–15, 2021.
- [11] Rossetti, MD 2015, *Simulation modelling and Arena*, John Wiley & Sons, USA.
- [12] Steins, K 2010, 'Discrete-event simulation for hospital resource planning: Possibilities and requirements', PhD thesis, Linköping University, Sweden.
- [13] Kelton, WD, Sadowski, RP & Sturrock, DT 2004, *Simulation with ARENA*, McGraw-Hill, New York.
- [14] Banks, J 1998, *Handbook of simulation: principles, methodology, advances, applications, and practice*, John Wiley & Sons, New York.
- [15] Banks, J, Carson, JS, Nelson, BL & Nicol, DM 2010, *Discrete-event system simulation*, Prentice Hall, New Jersey.
- [16] Altıok, T & Melamed, B 2007, *Simulation modelling and analysis with Arena*, Academic press, San Diego.
- [17] Balci, O 1998, 'Verification, validation, and testing', in J Banks (ed.), *Handbook of simulation*, Wiley, New York, pp. 335–93.
- [18] Alzahrani, MJ 2011, 'Using modelling and simulation to improve oral health services delivery in Hampton Roads, Virginia', PhD thesis, Old Dominion University, Virginia.
- [19] Sargent, SA, Talbot, MT & Brecht, JK 1988, 'Evaluating precooling methods for vegetable packinghouse operations', *Eng. Sci.*, vol. 23, pp. 1147–55.
- [20] <https://www.engineeringforchange.org/solutions-library/product/eco-frost-micro-cold-storage>.
- [21] P. Food and P. Export, "Government Incentives and Schemes for the Cold Chain Industry," pp. 1–6, 2021.
- [22] Hodges, R, Buzby, J & Bennett, B 2011, 'Postharvest losses and waste in developed and less developed countries: Opportunities to improve resource use', *The Journal of Agricultural Science*, vol. 149, no. 1, pp. 37–45.
- [23] Hou, Y-F, Xie, D & Wang, J-B 2015, 'The cold chain logistics for perishable agricultural products in China', *Advance Journal of Food Science and Technology*, vol. 8, no. 11, pp. 802–7.
- [24] Hsieh, L-F, Wang, L-H & Su, C-P 2011, 'Developing green supply chain management strategy of hotel-centered in vegetable and fruit purchasing and recycling', paper presented to 2011 IEEE 18Th International Conference on Industrial Engineering and Engineering Management (IE&EM), Changchun, China, 3–5 September 2011.

- [25] Huang, W-C, Kuo, T-C & Wu, S-C 2007, 'A comparison of analytical methods and simulation for container terminal planning', Journal of the Chinese Institute of Industrial Engineers, vol. 24, no. 3, pp. 200-9. Iannoni, AP & Morabito, R 2006, 'A discrete simulation analysis of a logistics supply system', Transportation Research Part E: Logistics and Transportation Review, vol. 42, no. 3, pp. 191-210.
- [26] International Trade Administration 2016, 2016 top markets report cold supply chain, United States Department of Commerce, USA. Jacxsens, L, Luning, P, Van der Vorst, JGAJ, Devlieghere, F, Leemans, R & Uyttendaele, M 2010, Simulation modelling and risk assessment as tools to identify the impact of climate change on microbiological food safety—The case study of fresh produce supply chain'.
- [27] Mongkut, T. Thonburi, B. B. A. I. Business, and B. It, "Consolidated Cold Chain Design for Fresh Fruit Supply Chains in Developing Countries: A Simulation Study," no. March 2018.
- [28] Joint Ventures, Franchising, India Entry Program, Setting up of Company- ILO Consulting.
- [29] Agribusiness Consulting | Sathguru Management Consultants.
- [30] O. R. Penetration, "Retail penetration across countries (2010-11) Market Share – Retail Segment," 2018.
- [31] T. Ford, A Technology Challenge, vol. 62, no. 11. 1990.

e-Commerce Trading Technologies

Harmandeep Kaur

Assistant Professor, Department of Computer Science & Applications

B.A.M Khalsa College, Garhshankar, Distt. Hoshiarpur

E-mail-kaurharman195@gmail.com

Abstract - Electronic commerce is a commercial activity through computer system and networks. e-commerce isn't restricted to merchandising however extends to includes business operations, electronic information interchange and purpose of sale systems. Moreover, the connected expansion of e-commerce is dependent on growth tactics and financial transaction security. This study will demonstrate how e-commerce has influenced the trading system by highlighting some of its benefits and issues.

Keywords—E-Commerce, Security issues, Security Tools, e-Commerce Portals.

I. INTRODUCTION

e-Commerce known as electronic commerce. We can sale or purchase any product through E-Commerce. If you are businessman, you can sale any product or if you are consumer then you can purchase any product. Buying or Selling goods through Internet is known as E-Commerce. E-commerce took place after a wide use of Internet. It can be using any type of technology and communication such as Cell Phones, TV, IT and Computers. It is good alternative for many companies and businessman who are depending on Computers. e-Commerce is not a technology itself; it is correlated with doing business with use of technology. e-Commerce provides best opportunities for making profits throughout business. It changes traditional commerce to electronic commerce, paper cheque or money to E-payment system. e-commerce covers many areas in which electronic market, grocery, health care, tourism, telecommunications assets are there. Top e-commerce websites in India are Amazon, Flipkart, eBay India, Myntra, Health kart ,Zomato ,OLX, Paytm etc.

“It's approximately commercial enterprise interactions wherein companies talk with clients

electronically as opposed to through physical contact.”

e-commerce is a vital aspect of daily life in today's Indian civilisation. According to a study conducted by IAMAI and Kantar Research, India's internet users would grow by 45 percent to 900 million by 2025, up from 622 million in 2020.

By 2026, the online grocery industry is estimated to have grown at a CAGR of 28.99 percent, reaching INR 1310.93 billion (2021-2026). Online grocery sales are expected to expand at a 60 percent annual rate by 2020.

The pandemic of COVID-19 is one of the major causes of this rise. It compelled consumers, particularly in urban locations, to conduct daily grocery shopping through online platforms to adhere to lockdown-related constraints and preserve social distance.

II. REVIEW OF LITERATURE

“Emerging Trends of E-Commerce in India: An Empirical Study” [1], as the title of the article suggests. The purpose of this article is to examine recent E-commerce trends as well as the numerous obstacles that the E-commerce business faces, as well as the success of E-commerce in India.

“E-Commerce Security Issues and Solutions” [2]. This research aims to investigate E-commerce security, as well as numerous security challenges in E-Commerce and suggestions for safe and secure online shopping through shopping websites.

“Advantages and Challenges of E-Commerce Customers and Businesses: In Indian Perspective” [3]. The purpose of this article is to examine factors like safety and security of online money transaction and major challenges faced by e-commerce in India.

"E-Commerce Technology Made Easy" [4]. According to this study, the cost of operating an online business is less than that of operating a physical store, and there are no additional charges in an E-Commerce firm.

"Future of E-commerce in India,"[5] according to the article. The purpose of this article is to examine the future of E-commerce in India and to evaluate the future growth segments in the Indian E-commerce market. Various aspects were discovered to be important for the future growth of Indian E-commerce, according to the study. According to the report, E-commerce will develop at an exponential rate in the next years in India's fast-growing industry.

III. TYPES OF E-COMMERCE

1. Business to Consumer (B2C): Company sell product/goods or services directly to the consumer online.**Fig.1** a company wants to sell products or goods to the customers by using a medium like websites or mobile apps. Example: **Amazon, Flipkart.**

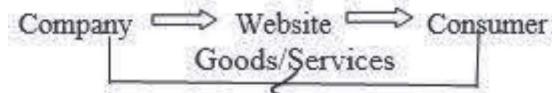


Fig.1

2. Business to Business (B2B): A business transaction involving a manufacturer and a wholesaler, or a wholesaler and a retailer. For example, Business-to-business refers to transactions that take place between businesses rather than between businesses and consumers. **Fig.2** two company (Manufacturer & Wholesaler) do business through online trade medium. It encourages business online. Example: **India Mart, Alibaba**



Fig.2

3. Consumer to Business (C2B): In this consumer provides a service or product to the business. Consumers post its services /products on or and company can bid on the posts or purchase services.**Fig.3** Customer post some of the products or services on website, the business organization place order hires the services through websites, once deal is done customers receives money & business organization receives products. Example: **Freelancer, Fiverr**

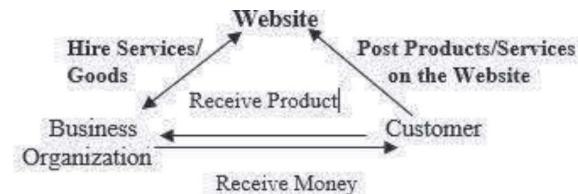


Fig.3

4. Consumer to Consumer (C2C): In this one consumer sell product or services to another consumer through internet. It means that a consumer would contact a business in search for a suitable customer{C2B2C} **Customer to Customer but through Business Organization.****Fig.4** Customer1 posts products on website for sale, customer2 buying products by placing order through website once order placed customer2 receives the product and customer1 receives the money. Example: **OLX, eBay**

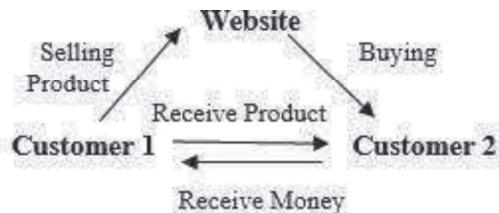


Fig.4

IV. PROCESS OF E-COMMERCE

1. A consumer connects to the home page of a merchant's Web site on the Internet using a Web browser.
2. The consumer browses the catalog of products featured on the site and selected items are placed in the electronic equivalent of a shopping cart.

3. Customer pays bills by payment methods like credit\debit card or any other method.
4. Once order is completed email or text message send to the customer and merchant.
5. When the consumer ready to complete the purchase selected item then it will be ready for shipping.

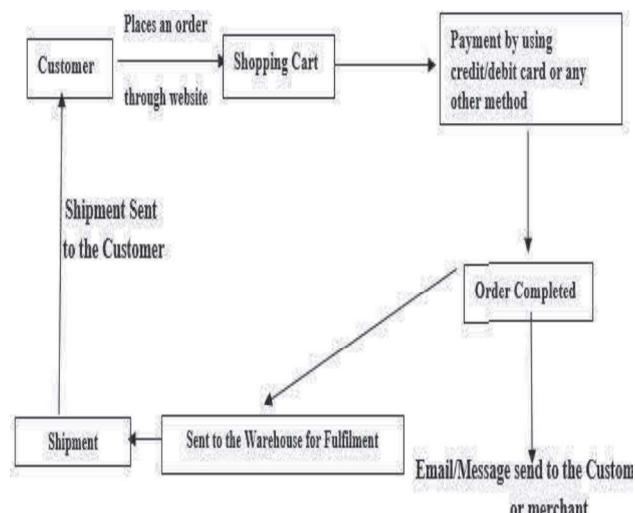


Fig.5

V. E-COMMERCE IN INDIA

NASSCOM reports that, despite COVID-19 issues and disruptions, India's e-commerce business is predicted to increase by 5% in 2021, with sales of US\$ 56.6 billion. The Indian e-commerce market is expected to grow to US\$ 200 billion by 2026, up from US\$ 38.5 billion in 2017. By September 2020, the 'Digital India' initiative is estimated to have increased India's Internet connections to 776.45 million.

Gem (Government e- Marketplace) is a one-stop shop allowing government departments, organizations, and public-sector enterprises to purchase common-use goods and services online.

The goal of Gem is to improve public procurement transparency, efficiency, and speed. The purchases through Gem by Government users have been authorized and made mandatory by Minister of Finance by adding a new Rule No 149 in the General Financial Rules, 2017. Gem is created in a record time of 5 Months to

facilitate online procurement of Goods & Services. The platform is owned by Gem SPV (Special Purpose Vehicle), a government-owned, non-profit organization under the Ministry of Commerce and Industry of India.

In 2016, the Directorate General of Supplies and Disposals was replaced by Gem, a contactless, paperless, and cashless online marketplace. Gem has improved public procurement transparency and visibility since its start. By focusing on its three pillars of inclusion, usability, and openness, as well as efficiency and cost savings, the portal has revolutionized public procurement in India. The cost of the marketplace portal is around 9.75 percent of the cost of the media. The Indian government has mandated that vendors display the country of origin on products sold through the Gem site.

e-RUPI: The National Payments Corporation of India (NPCI) has introduced a new digital solution called e-RUPI in collaboration with the Department of Financial Services (DFS), the National Health Authority (NHA), the Ministry of Health and Family Welfare (MoHFW), and partner banks.

e-RUPI is a digital voucher that a beneficiary receives by SMS or QR code on his phone. It's a pre-paid coupon that he or she can use at any location that takes them. This contactless e-RUPI is simple, safe, and secure because it keeps the beneficiaries' information private. The entire transaction process with this voucher is substantially faster and more trustworthy because the required amount is already contained in the voucher.

VI. E-COMMERCE SECURITY PURPOSE

e-commerce security is nothing but preventing loss and protecting loss the areas financially and informational from unauthorized access, use or destruction. Security issues means any situation, threat, vulnerability, act, or omission posing a risk of giving to a Security Incident or any breach of supplier's representation or covenants in the agreement and or Order regarding safeguarding of UTC Information Maximum security target five areas:

1. **Confidentiality:** - Information should not be accessible to an unauthorized person. During transmission, it should not be intercepted.
2. **Integrity:** - Information should not be tampered with while being transmitted over the Internet.
3. **Availability:** - Information should be available wherever and whenever required within a time limit specified.
4. **Authenticity:** - There should be a mechanism to authenticate a user before giving him / her access to the required information.
5. **Non-Reputability:** - It is the protection against the denial or order or denial of payment. A sender should not be able to deny sending a message once it has been sent.

VII. SECURITY THREATS

1. Hacking: It accesses computer system without the knowledge of owner of the system. The person who does such types of unlawful activities/acts known as hackers. Whenever, they get access of another system they can alter or steal the important information like bank or credit card details. Mostly hacker's attacks E-commerce websites because these websites offer the facilities of online banking.

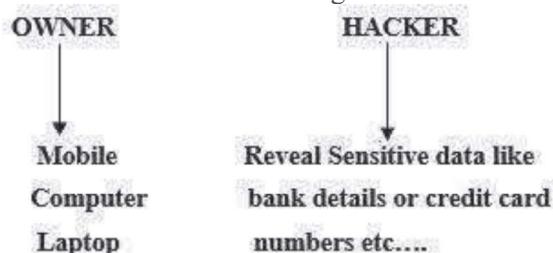


Fig.6

2. **Virus & Worm:** Virus is self-replicating computer program that perform unwanted events and interrupts the working of the system. **Worm** is a virus that spread during internet connections like fraud mails or spam mails etc.

3. Denial of Services (DoS): Denial of Service is a security attack where hackers take action that stops users to use electronic devices correctly. It makes a network unavailable for real users by disrupting the services of host unavailable. DOS attacks include slow performance of system and unavailability of websites.

VIII. SECURITY SOLUTION TOOLS

Security-related solutions Protect your data from the inside out, ensuring that only those who need to know have access.

- 1. Firewalls – (Software and Hardware)** A firewall is a network security device that monitors and controls network traffic both inbound and outbound.
- 2. Public Key infrastructure** – It uses a pair of keys to achieve the underlying security service.
- 3. Digital certificates** – It is electronic file that typically contains identification information about the holders.
- 4. Digital Signatures** – A digital signature is a mathematical mechanism for verifying the validity and integrity of a messaging software or digital document.
- 5. Password** – A password is a string of characters used to verify the identity of a user during the authentication process.

IX. ADVANTAGES / DISADVANTAGES

Advantages of E-Commerce

1. Faster buying /selling product, as well as easy to find.
2. Buying and selling products 24/7.
3. There are no theoretical geographic boundaries, allowing for greater client reach.
4. Low operating costs and higher service quality.
5. There is no requirement for actual business setups.
6. It is simple to start and run a business.

7. Customers can simply choose products from a variety of vendors without having to physically move around.

Disadvantages of E-Commerce

1. Customers are unable to inspect products for themselves.
2. Internet access isn't available to everyone.
3. Theft of credit card numbers is a risk.
4. Mechanical failures can have a wide range of consequences for the entire operation.

X. FINDINGS OF THE STUDY

1. In the era of e-commerce, wholesalers face the greatest threat because producers can easily disregard them and sell their products to retailers and consumers. Wholesalers can use E-commerce to form contracts with reputable producers and connect their firm to the internet.
2. The retailer can save his firm by integrating it with online distribution. The shop may give customers more information about many topics, fulfil electronic orders, and always stay in touch with them. As a result, E-commerce is a good option.
3. Producers can benefit from E-commerce by connecting to the internet, providing information about their products to other nodes in the supply chain, and establishing a brand identity.
4. As more people become involved in e-commerce, the demand for centers that provide internet access is growing. As a result, those who seek to benefit from it can set up cyber and reap the rewards.
5. Wholesalers, retailers, producers, and the public all have different types of opportunities because to e-commerce. Retailers must fulfil electronic orders and maintain constant contact with customers.

XI. CONCLUSION

Today e-Commerce is one of the greatest innovations which have come with the industrial and technology revolution. It entails utilizing technology to facilitate the selling and purchasing of goods. The word "high-tech" refers to computers, the internet, and other smart gadgets such as cell phones, vending machines,

and so on. Ecommerce is divided into four

categories (B2B, B2C, C2B, and C2C) E-commerce not only expands business chances, but it also expands educational and intellectual opportunities. It looks that delivering E-business education has a lot of potential. Thus, E-Commerce is one of the best ways to have a great successful in a people's lives. The time of the covid-19 pandemic has seen a lot of booms in online businesses, because the customers could not go out and shop due to social distancing, so they preferred online shopping due to which there has been a boom in many areas like health care, electronic market, grocery market etc.

REFERENCES

- [1] Rajeshwari M.Shetter (2016)"Emerging Trends of E-Commerce in India: An Empirical Study," International Journal of Business and Management Invention,ISSN (Online): 2319 – 8028, ISSN (Print): 2319 – 801X, Volume 5, Issue 9
- [2] Niranjana Murthy M, DR. Dharmendra Chahar(2013)"The study of E-Commerce Security Issues and Solutions" International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 7.
- [3] C. Eugine Franco , Bulomine Regi. S (2016)"Advantages and Challenges of E-Commerce Customers and Businesses: In Indian Perspective", International Journal of Research-Granthaalayah, Vol.4 (Iss.3: SE, ISSN- 2350-0530(O) ISSN- 2394-3629(P)
- [4] S.Sridhar(2017) "E-Commerce Technology Made Easy", International Journal of Innovative Technology and Research, Volume No.5, Issue No.3, 6183-6198.
- [5] Nisha Chanana and Sangeeta Goele(2012) "Future of E-commerce in India", International Journal of Computing & Business Research ,ISSN (Online): 2229-6166

E-Business Services in India

Kamini

Assistant Professor, Department of Computer Science & Applications

B.A.M Khalsa College, Garhshankar, Distt:Hoshiarpur

Email: sharmakamini253@gmail.com

Abstract

E-Business is rapidly increasing day to day. Government plays the important role in enhancement of the business. These services are saving lot of time and money. It also creates transparency. The initiative taken by government is very helpful in growth of E-business in India. These services also provide online registration and license to small and medium scale industries which save their time. This paper covers all the aspect of status of Business services in India ,benefits and issues of G2B services like- lack of online identity verification is also mentioned .Some points of Jordan Report,E-Commerce services in India has also discussed in this. Moreover, this paper also include the scope of future work done on the E-business.

I.INTRODUCTION

E-business (electronic business) is the use of the internet to conduct business processes. These e-business processes include purchasing and selling goods and services, providing customer service, processing payments, managing production control, collaborating with business partners, sharing information, running automated employee services, and recruiting, among other things.

In India Digitization plays an important role in growing the business. Many startup business are established under this. It has provided new opportunities to people and businessman to economically grow their business. Digital India Program is launched by the government of India with the vision to transfer India into digitally. This program is launched with the aim to provide services to the people online and ensure that all the services are available in the rural areas. In India 37.90 billion digital transaction have done in 2021. In this Government to Business services plays

an vital role in establishing the good connection between the entrepreneur.

ICT Information Computer Technology is also an part of E-business. It is based on internet it impacts in various fields like education, business, science and health care. As ICT can enhanced the teaching and learning process. Artificial Intelligence is a part of it and it is very helpful in the health care which helps to easily diagnosis of the disease. Many software are developing in today times as in COVID-19 Pandemic many software are developed to verifying the vaccination status, knowing about lockdown. So there is a rapid development in ICT services and internet services are also increased. E-Business provides us many services from which business can easily run and it can enhance. The workload and time is also reduce by this. Many services provided by Government to Business are E-Classification, Approval and Occupancy of Hotels, FSSAI, E-Filing Tax Return, E-Mandi. These services allows the entrepreneur to get their Hotel registered by following out some steps. This application allows to fill the form online with some details and apply for Hotel Project Approval or Hotel Classification which is chosen. Another is FSSAI Food Safety and Standards Authority of India. It is established by government of India. Its function is to provide the good and healthy food to the public. Businessman who involved in it must have compulsorily to obtain FSSAI Registration or License. It is of 14 digits number that is printed on all food packages. E-Filing Tax Return helps to fill online tax and it provides information related to filling tax return. E-Mandi portal is useful for the entrepreneur to registered and it is beneficial for farmers too.

II. REVIEW OF LITERATURE

This paper reviews the services of E-businesses in India, how the business is growing and the issues of E-business. As of June 2012, India had approximately 137 million internet users. When compared to markets such as the United States and the United Kingdom, e-business access is limited, but it is developing at a far faster rate, with a huge number of new entrants. Cash on delivery is a distinctive feature of India and is a popular mode of payment. As a result of the cash economy, over 80% of Indian e-commerce is Cash on Delivery. In terms of e-business, India is still in its infancy, but it has a lot of potential in emerging countries like India. A growing installed base for Internet use, as well as a more competitive Internet Service Provider (ISP) market, have fuelled e-commerce growth in Asia's second most populous country.[1] E-Business can be done online. It can be done by using IT information technology. In this buying and selling can be done online with the help of E-commerce refers to the electronic exchange of goods and services. Due to the great innovation of the Internet, which is rapidly spreading over the world, it transforms the entire commercial environment. With the debut of the World Wide Web (WWW) in 1994, the Internet's power as a global access was felt. This worldwide network facilitates international business relationships.[2] The percentage of Indian e-commerce space is increasing as more and more online retailers enter the market. Although this degree of entrance into the e-commerce business is favourable in the long run, most entrepreneurs lack the resources or funds to wait years for returns. Instead of relying on advertising revenue, many Indian portal sites have moved their focus to e-commerce. The content-rich web communities that have sprung up around these portal sites have been successfully used to sell everything from event and movie tickets to groceries and computers.[3]

Government plays a vital role in the business this is showed in Jordan Report(2000) government play the role of customer, tax collector and the supporter of the business. Organization can directly deal with the government and get the immediate access to the information. All this things can be achieved with the help of electronic services. Further, according to the

CH2(2000) study introducing e-government in report, the three business components are administration, trade and finance. As compared to G2C services G2B services work more rapidly according to the Jordan Report[4].

India's e-commerce industry has grown at an exponential rate over the last decade. Many factors contribute to this growth, including Indian consumers' rapid adoption of technology, significant increases in internet users, new enabling technologies, innovative business models, and alternative payment options provided by e-commerce companies.[2]

Business services have many benefits but this can be achieved only if the people have trust on the government and internet. If citizens have trust only then the E-government services can be achieved. While using any E-government portal citizens always have an issue either their data is protected nor any hackers can attack their data. This type of trust issues are mostly came in mind while the online transaction. Either the portal and technology is safe or not. So government must take necessary steps to protect and safe the data of every citizens. If the government can win the trust of citizens only then the E-Business can become successful.[5]

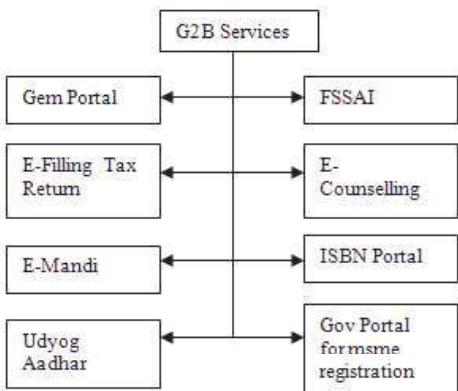
Government to Business helps a lot to enterprise and the customer. It reduces the gap between the seller and buyer. Entrepreneur can directly connect with the government inspite of visiting government offices again and again. The data can easily accessible by the company which can helps of increasing the more consumption and economic prosperity. Many initiatives are taken to use the ICT tools like internet and mobile devices to grow the E-business. The internet is the main tool which is used in it, so this is the duty of government to ensure that everyone should facilities with it. This can be helpful for business organization in auctions, tenders and some other submission functionalities. This helps in updating corporate information and filling out the electronic forms i.e. social insurance forms, taxforms. Government can provide all the services to the business organization. Information can be provided through various website of government and this information

is very useful in starting up the new business. Business organization can download, fill the form and submit on the website in spite of visiting government again and again. This is opposite to the B2G e-commerce model. Government has launched www.makeinindia.com website from where the new initiative design to facilitate investment has launched.[4]

Business can improve if the raw materials to factories are provided on time and the finished goods placed in the market quickly. This can be done if the maintenance of roads rails lines, bridge, telecommunications systems and seaport has done in proper way. Governments fund free elementary and secondary education, grants and loans for higher education, and worker training programmes to guarantee that firms have access to trained people. Free worker training is frequently provided by governments in collaboration with trade schools, community colleges, and universities.[6]

III. BUSINESS SERVICES IN INDIA

E-business service providers (eBSPs) are businesses that assist other businesses in using e-business technology to improve operations. This could include tasks like designing, building, and operating a company's website or streamlining a company's procurement processes. These services are Gem Portal, E-Filing Tax Return, E -Mandi, Udyog Aadhar, FSSAI, E-Counselling, ISBN Portal, Gov. Portal for msme registration. Some business services are shown in Fig.1



A. Gem Portal

Gem Portal is the Government e-marketplace where the new and small scale industries can register themselves on GEM Portal to start their business. It is launched in 2016 to maintain the transparency in the system. Firstly e-tender is filling by going through various government offices so there are more chance of corruption but by this portal corruption time and work load has reduced.

B.E-Filling Tax Return

E-Filing Tax Return portal provides the information related to the e-filing to the tax return. On the 13th of October 2021, the Income Tax Department's e-filing system (www.incometax.gov.in) reported receiving more than 2 crore Income Tax returns. This portal is beneficial to the government employee, businessman and the common people.

C.E-Mandi

E-Mandi Portal is very useful for the farmers, and the person who want to do the business in Mandi. This portal is quite beneficial to them. With this mandi process can become digital. In this portal license for mandi is easily applied.

D.UdyogAadhar

UdyogAadhar portal is used in the registration of the small, micro and medium enterprise. The motive behind this portal is to provide benefits to the small, micro and medium enterprise which are registered by their adhar card Number.

E.FSSAI

Food Safety and Standards Authority of India (FSSAI) has been established under food safety and standards act , 2006. The act is related to the food safety and regulations in India. The Food Safety and Standards Authority of India (FSSAI) is in charge of safeguarding and improving public health through food safety regulation and oversight.

F.E-Counselling

From the application form to online fee payment, e-admit card, answer key display, e-score card, online selection of courses / institutes, seat allocation, document verification, online admission, seat withdrawal, reports and dashboards, e-Counselling is a one-stop solution for all examination / counselling needs.

G.ISBN Portal

ISBN stands for International Standard Book Number, and it is a thirteen-digit number that has replaced the management of large bibliographic descriptive information. ISBN stands for International Standard Book Number, which is a brief and easy-to-understand machine-readable identification number that uniquely identifies any book. ISBN stands for "International Standard Book Number," which is a 13-digit code that can be read by machines.

H.Gov Portal for msme registration

The Micro, Small, and Medium Enterprises (MSME) sector of the Indian economy has evolved as a lively and dynamic industry. MSMEs not only serve an important role in creating large numbers of jobs at a cheaper cost of capital than large enterprises, but they also aid in the industrialization of rural and backward areas, reducing regional imbalances and ensuring a more fair distribution of national revenue and wealth.

IV. TOOLS AND TECHNIQUES

Tools and Techniques means technology used in E-Business services. Tools are the infrastructure which are required for the making E-business services. NeGP National e-governance Plan is a government agency which is used to implementing E-business services to the citizens and the industries. ICT plays a vital role in it. Information and communications technology (ICT) is a term that emphasises the role of unified communications and the integration of telecommunications (telephone lines and wireless signals) and computers, as well as necessary enterprise software, middleware, storage, and audiovisual, to enable users to access, store, transmit, understand, and manipulate information.

V.ISSUES IN BUSINESS SERVICES

Although G2B services provides us many features and services regarding it. Many issues in business service like that all the work done is with internet and sometimes server go down due to technology. Some of its issues are as follows:

A. The policy on product returns and refund

This is the challenge in G2B services in the product refund and return policy, business has to

suffered a great loss, either the reason is what, wheather the product is destroyed/damaged. There is great loss to the shipping companies also.

B.Data Security Problems

Another Challenge is in this data security. Hackers can attack the website and infect it with virus. All the essential and confidential data is attacked and there is huge loss to the business.

C. A Lack of Online Identity Verification

There is a lack in online verification of a person who visit the website. When the user enter the user-id password, portal is unaware of that either the user is authenticate or questionable. Some people fill the wrong data on the website so it is difficult for the companies to deliver the product. It makes a heavy loss to the business.

D.Barrier in Language

Language Barrier is also an issue in the E-business because most of the website are in English language which cannot be understand by all so it should be available in the regional language so that it can be accept by all.

E.Infrastructure

Infrastructure constraints such as rural telecommunications networks, power outages in various states, and communication methods affect implementation speed, causing delays. An efficient e-Government programme cannot be established until these areas are improved.

VI. SCOPE FOR THE FUTURE WORK

Scope of the Business services cannot predict it in efficient way but can be estimate from the current situation. As the E-Business is increasing rapidly in India. The internet customer in India is growing day to day. In 2021 India rank 2nd after China in using internet. As the Block Chain technology is increasing it can helpful in building the business. This is a decentralized system which is reliable in the case of security. The unique number called Hash is assigned to the Bitcoin. It can helps to know from where chain has started. It is difficult to do fraud because record of transaction is stored in bitcoin. It becomes difficult for hackers to hack the system. Block Chain can enhanced the business because data stored in it is secure. It has reduce the cost by eliminating the middleman—vendors. This

also increased the speed of the work by eliminating middleman. So Block chain can help in enhancing the business in future and it can provide many features like data security.

VII. CONCLUSION

This paper reviews the various government services and the technology provided for the benefit of the E-business in India. G2B provides many services to the business which helps to low corruption, transparency and reduce the time and cost. Various Benefits also provided by the G2B is that any businessman can registered on government portal like Gem Portal. This is the e-marketplace where the product & services are available. G2B services are very beneficial but this can be possible only if the infrastructure like internet, telecommunications services are available to each and every citizen of India.

VIII. REFERENCES

- [1] Ashok Panigrahi, Ranjan Upadhyaya, Dr. P. P. Raichurkar, "E-Commerce Services in India: Prospects and Problems," International Journal on Textile Engineering and Processes ISSN 2395-3578 Vol 2, Issue 1 January 2016.
- [2] Basheer K.T, "An Analytical Study Of E-Commerce In India" International Journel of Creative Research Thoughts(IJCRT).
- [3] Sarbapriya Ray, "Emerging Trend of E-Commerce in India: Some Crucial Issues, Prospects and Challenges," Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol 2, No.5, 2011.
- [4] Mohammad Taisir Masa'deh, "The Economic impact of e-government initiative (G2B) as a tool of ICT upon SMEs in Jordan"
- [5] Zunawanis Mustafa Mohd Sahrulnizam Abu Hassan Mohamad Dhiyauddin Abd Aziz, "E-Commerce Challenges and Solutions"
- [6] Henrik Florén, Johan Frishammar, Carmen Lee, Magnus Ericsson & Stefan Gustafsson, "A framework for raw materials management in process industries"

Reviews on Biometrics Technologies

Manpreet Singh

Assistant Professor

B.A.M. Khalsa College Garhshankar

Manpreets98h@gmail.com

Nikhil Chaudhary

Student(B.Sc Non-Medical)

B.A.M. Khalsa College, Garhshankar

Abstract:

In this era of technology, there are a lot of cases of cybercrime, data leak and loss of data, manipulation of data by unauthorized users, hacking of personal accounts which are growing at unprecedented rate due to common password based security systems which could be hacked. Therefore this has created a secure system requirement which could eliminate these security related issues, the alternate solution for this grave problem could be seen in biometric authentication technology as this technology could not be copied as this system consists of

several software technology which identifies or validates the user by matching the data being fed with the digital images of the unique characteristics of the every user. This data cannot be copied or hacked, so it makes the recognition more reliable than the outmoded way of authentication that usually consists of ID Cards or PIN statistics(password).

Furthermore authentication on the basis of biometric methods eradicates the essential to recall a password or carry a token.

Keywords:-

Biometrics, Authentication, Hacking, Iris, retina and Gait Scan (Body Movement) .

Introduction - The word “Biometric” is derived from the Greek words "Bios" which means Life and "Metric" means “to measure”. This simply means recognition of humans on the basis of intrinsic physical or behavioural traits. Biometric systems are automated methods of verifying or recognizing the identity of living person on the basis of some of physiological characteristics, like a fingerprint or face pattern ,or some aspects of behaviour or nature and even gait scan can be a reliable aspect for biometrics.

The two major ways of biometrics are **Identification and Verification**.

Identification determines who the person is. It involves taking the measured characteristics and trying to find a match in a database containing records of people and those characteristics.

Verification involves taking the measured characteristic and comparing it to the previously recorded data for that person. This method requires less processing power and time , and is **Why Biometrics?**

This era of innovation and technology has significantly transformed the lives of individuals

often used for accessing places or information. In other words , a biometric System is a technological system that uses information about a person to identify that person.

History- Authentication of a person is not a new concept. There are numerous examples in the past showing that identification was done in past as well. One such example for supporting this fact is of Joao de Barros who was an explorer and writer. He wrote that Chinese merchants used a form of biometric by stamping children's palm prints and footprints on another .In the 1890s, 'Bertillonage', created a method/technique of bodily measurement with the criminals who repeated the crimes with different names etc. In the future, finger printing, which was developed by Richard Edward Henry of Scotland Yard.paper with ink. Chinese solved a way to differentiate young children from one another.

of this planet. Technological innovations has improved modern living for many and has provided access to opportunities in human

history. Combined with mobile technology, in-memory computing and prospective analysis, cloud computing has become a driving force and as more of our data has become digitized, and we share more information online, securing data has always been crucial. Customers share their data with the companies that they deal with. They trust the company with their data. A single company may possess the personal information of its billions of customers and the data is supposed to be kept private so that identity of its customer must stay as safe and protected as possible, and the reputation of the company remains impeccable. This thing is not only just limited to business but also as an individual have a lot at stake when it comes to securing data. The more secure our data is the lesser risks of data breach.

When data that should be kept private gets in the wrong hands catastrophe and misfortunes can happen. A data breach at a government agency can, for instance, put top secret information in the hands of hostile state. A breach at an institute can put unshared data in the hands of other contender and a breach at a school could put personal information accessible to delinquent who could commit cyber identity theft. Therefore with more number of users in cyberspace there is growing demand of data security. It is true that PINs may be forgotten and tokens may be forged and strained or they may get lost or embezzled.

To tackle this risk of security breach of data we use biometrics which seems to be more reliable than antiquated security patterns involving tokens or PIN statistics. In order to solve this stumbling block of remembering the keywords or passwords and ducking the issue of carrying tokens Biometric Authentication can be a great perk.

General Scheme of a biometric recognition system:-

Biometrics can use physical features like your face, fingerprints, irises or veins, or behavioral characteristics like your voice, handwriting or typing rhythm which are unique in nature.

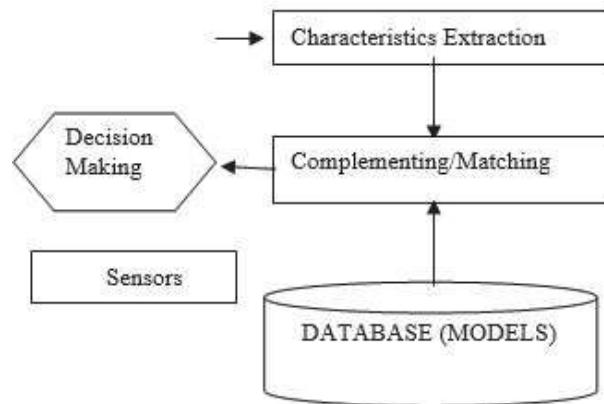


Fig1: Anatomy of biometric recognition

Different biometric devices operate on different hardware and software but they all use the same three steps mentioned below.

Enrollment: The first time a user uses a biometric system, it records basic information about the user, like name or an identification number. It then captures an image or recording of the user's specific trait.

Storage: Contrary to what you may see in movies, most systems don't store the complete image or recording. They instead analyze the trait of the user and translate it into a code or graph. Some systems also record this data onto a smart card that a user can carry with him/her.

Comparison: The next time user uses the system, it compares the traits of the user present to the information on file. Then, it either accepts or rejects that you are who you claim to be.

Systems also use the same three components:

A **sensor** that detects the characteristic being used for identification.

A **computer** that reads and stores the information.

A **Software** that analyzes the characteristic, translates it into a graph or code and performs the actual comparisons and authenticates the user.

Fig 1. shows the basic anatomy of biometric recognition system.

Kinds of Biometrics:-

Biometric strategies are widely used in a great many areas like computer and network safety, government agencies, prisons and in business platforms too.

Biometrics can be broadly categorized on the basis of the physical/physiological and the behavioral or personal traits.

Physical/Physiological Characteristics:-

Biometrics on the basis of physiological characteristics of a person involves characteristics that are physical in nature. Fingerprint, Handprint, Palmprint, Footprint, Facial Recognition, Thermal Image and Iris or Retinal Pattern are some of the examples of biometrics used in daily lives that senses physical features of the user.

Fingerprints Authentication:- It is one of the oldest Biometric characteristics used. This technology involves the process of carrying the digital images of fingerprints. It scans the friction ridge skin impression of the human fingers. The sensor senses the unique curves, bifurcations of the skin of fingers. Same principle is used in the palm scanning.

Eye Scanning:-

There are two methods for eye recognition:

Retinal scanning:- This technology involves the analysis of the configuration of blood vessels of the user in view of the fact that the configuration of blood vessels of a person is unique. The person has to look in a device that performs laser-scanning of his retina.

Iris Scanning:- In contrast to the Retinal scanning the person needs not to be close to the device. This technology necessitate the imaging by a camera. The iris patterns are obtained through a virtual based imaging system. The image is acquired and then analyzed by the device. The image contains 266 different spots, these spots are based on the characteristics of iris like furrows

and rings. Since the iris is stable throughout the life therefore no timely update of the image is essential.

Face Recognition:

In face recognition, a good resolution simple camera or a web camera is required. Facial recognition in visible light acquires features from the central portion of face image. Usually these characteristics do not change over the period of time. Superficial features such as facial expressions, hairs are avoided as these superficial features can be easily duplicated . The representation is compared with the existing database.

Handprint Imaging

This is based upon the principle of the hand geometry .In this method, the illustration of a user's hand is being scanned. Features like distance between the fingers, length of fingers, length of the hand are extricated and saved with the help of digital signal processing algorithms are generated.

DNA Analysis

This modus operandi of verification is predominantly used in criminal cases. DNA of the user in the form of blood, tissue, hair or nails is collected for confirmation. Usually DNA analysis occurs at snail's pace . Though DNA of a person is unique characteristic yet it comes with its own limitation since the hair or nail of a person can be stolen.

Voice Verification:- In voice verification, user is supposed to speak a phrase that could be a secret code as well . His vocal characteristics are measured. Physiological features like shape and size of vocal cords and behavioural features like pitch of the voice are examined. The verification process is different from the process of voice identification. In verification sample of speaking style pattern is saved and is juxtapose with the same person's speech.

Signature scanning:-

It is the dynamic analysis of the shape, size of signature, writing speed, time taken for signing, pressure applied by user's hand on the screen while signing etc. Though

Signature may be copied yet the traits while signing may not be.

Benefits of biometric technology:-

Biometrics was once something limited to the world of science fiction, but over the past few decades, we have seen the extensive emergence of biometrics in our daily lives. The COVID-19 pandemic has also contributed to this exponential growth and the requirement for biometric technology. Gaining contactless access to buildings and services, contactless payments and ATM interactions is more important than ever before.

As biometric technology continues to advance, the most common uses of biometrics in everyday life are mentioned below.

Airport Security:- Using biometric technology to verify passenger identity has been used in some of the major airports around the world for

several years and the use of the technology is now becoming more popular. Delta Air Lines, in partnership with U.S. Customs and Border Protection, Transportation Security Administration and Hartsfield-Jackson Atlanta International Airport, launched the first curb-to-gate biometric terminal in the U.S. at Atlanta's Terminal F.

Law Enforcement:- Law enforcement agencies like FBI and Interpol are widely using biometrics for criminal investigations. Fingerprints, iris and facial recognition, gait, and voice recognition are some of the the most commonly used biometric technologies among them.

Mobile Access and Authentication:-

One of the principal uses of biometric technology is in the field of smartphone security. Apple was the pioneer in introducing touch fingerprint in smartphones. The iPhone 5S was the first model with a Touch ID sensor integrated with its home button. There are hardly any smartphones that do not have any of the biometric authentications.

Banking:- The banking sector is another embracing biometrics across a range of services in order to deliver a more seamless experience for its customers. Japan is rolling out a trial of Facial Recognition at ATMs. Facial Recognition will be used as an additional level of security to authenticate that the owner of the card is the person using the card.

Home Assistants:- Alexa ,Siri, Amazon and Apple's digital voice assistants, are much more than a convenient tools .They are very real applications of artificial intelligence that is becoming an integral part of our daily life. These are already be accustomed to use voice recognition as a biometric identifier. Google Assistant that powers Google Home as well as the assistant on Android devices is compatible with a wide range of IoT (Internet of Things) devices including light bulbs, door locks, security cameras, security lights and more.

Schools :- The US, UK, Belgium, and other countries already implement biometric

technology on school premises. It is also a growing technology in U.S. education sector. Biometrics enhances the security and makes the enrolment process more efficient.

Building Access:- Biometrics brings a great many pros when it comes to access control. The technology holds the potential to provide a frictionless entry experience when using facial or iris recognition to control access secure areas within a building.

Conclusion- The advances in accurateness and Service ability and reducing cost have made the biometric technology a secure, reasonable and cost effective way of identifying individuals. Biometric strictures such as fingerprint perusing, retinal scanning, iris scanning, signature verification and others are all well The promises of e-commerce and e-government can be achieved through the utilization of strong persona authentication procedure. Each biometric techniques performance can be widely, depending on how is used and its environment in which it is used.

Refernces :-

1. <https://www.nec.co.nz/market-leadership/publications-media/the-top-9-common-uses-of-biometrics-in-everyday-life/>
2. <https://en.wikipedia.org/wiki/Biometrics>
3. <https://www.ijert.org/research/review-paper-on-biometric-authentication-IJERTCONV5IS23026.pdf>
4. <https://english.alarabiya.net/business/property/2015/06/25/Gaza-rebuilding-moving-at-snail-s-pace->
5. <https://www.ijert.org/research/review-paper-on-biometric-authentication-IJERTCONV5IS23026.pdf>
6. <https://www.scribd.com/document/101533774/Biometric-Authentication-System>
7. <https://www.engati.com/glossary/biometric-recognition>

Blockchain technology to secure Cyberspace

Dr. Chetna

Chitkara University Institute of
Engineering and Technology, Chitkara
University, Punjab
chetna@chitkara.edu.in

Harpreet

B.A.M. Khalsa College,
Garhshankar,
Hoshiarpur

Saloni

B.A.M. Khalsa College,
Garhshankar,
Hoshiarpur

Abstract - Through the usage of cryptocurrencies, blockchain technology has been adopted in a variety of industries, the most prominent of which is finance. However, in terms of cybersecurity, the technique is practical. This article examined 30 researchers' visions of Blockchain application cases in the cybersecurity business.. It was found that the majority of researchers are interested in making use of Blockchain for securing sensitive information ,webwork and IoT devices. The report looked at the procedures that preceding academics have suggested that Blockchain can provide security to the three sectors of IT that have been identified as problematic. At last, the article suggested that future studies focus on creating network safety applications on a solitary Blockchain to take into account unification and consistency among arrangements.

Keywords -IoT, Network Security, Blockchain.

I. INTRODUCTION

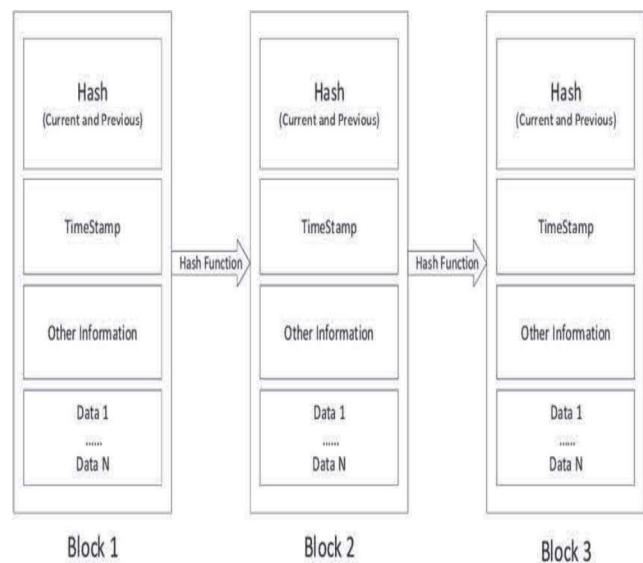
Blockchain is a game-changing innovation that has the capability to convert the technology in the coming future and challenge various areas by introducing prominent creative solutions. It is accessible, persistent, and shared, thus making it useful for various purposes.. The technology became popular because of the increasing demand of cryptocurrencies. The name "blockchain" can be generally characterized as a succession of blocks that are related to cryptography¹. A block can be defined as data structure that comprises mainly of three sections: data, the previous block's hash, and the hash of the data and the prior hash². To sum up, there is a series of interdependence between blocks that can be used to keep the Blockchain's integrity³. If there is any change in the block, the hash will also change. The hashes of the blocks will not be valid, resulting in a twisting effect. This is the reason why blockchain transactions are irreversible⁴. This type of framework could be extremely useful in solving cybersecurity issues.

II.THEORY

A Blockchain's blocks are impossible to change because if we did it that would jeopardize the quality

of all following blocks.

Due to the severe Blockchain design, care should be taken while adding other blocks to the chain to ensure that it won't be changed later. The following diagram illustrates a block diagram:



Instant access is provided to the whole Blockchain, this is done when a node produce another block by transmitting it to other nodes. All other nodes must agree after a block has been constructed and is ready to be added to the Blockchain⁵. Proof of work (PoW) and Proof of Stake (PoS) are the two most common consensus algorithms. The process of nodes validating a block by indicating that the work is accomplished and has come to a settlement on the results⁶. The agreement of the nodes must be on right solution⁷. Miners are responsible for doing this, and a lot of computation power is used. In the confirmation of stake agreement, a stake is owned by nodes in the Blockchain which support the expansion of another block to Blockchain .A Blockchain investment is done by individuals and in order to do this, and it isn't usually computationally intensive⁸. An example of a PoW consensus algorithm is as follows:

1. Users' transactions are grouped together in a

2. memory pool by the network.
 3. Each transaction in the pool is verified by miners in a race.
 - a. Solve a **complicated** puzzle to **confirm** transaction
 - b. Verify your solution by sending it to the network.
- If Correct then
- Distribute the right answer to the other miners.
 - ii. Else
 - Make a new computation.
4. The first miner who receives the perfect solution is remunerated.
 5. The block is added to the existing Blockchain and the memory pool is updated with new information.

The following flowchart for this algorithm is:

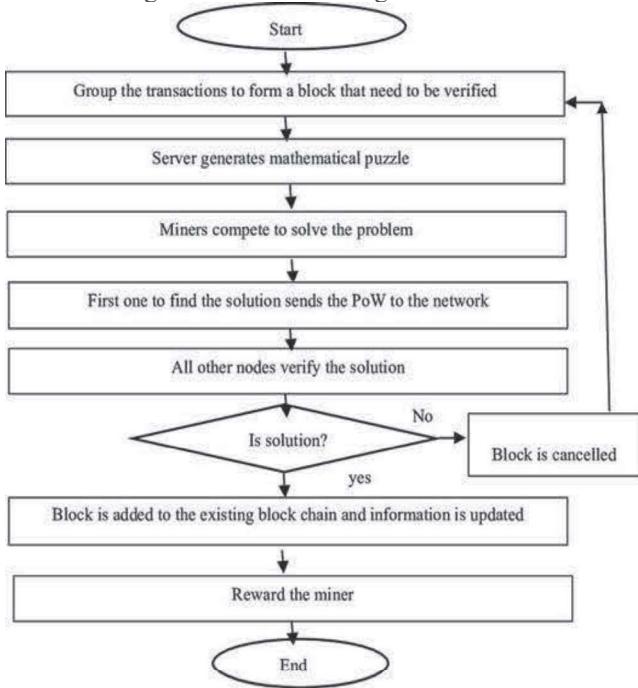


Fig 2: The PoW consensus algorithm is depicted as a flowchart.

The distributed model of recordkeeping is another essential element of Blockchain. In a blockchain network, nodes are present and all the network's information can be stored entirely. Many nodes do that due to the fact it's far required for both consensus or reference. This gets rid of the want for centralized facts storage¹⁰. Any rival endeavoring to hack the Blockchain will have to understand a critical number of the nodes that store the distributed information. In effect, the network compares blocks of data stored in decentralized locations to see which is unique. The

majority of the time, the data is correct or uncorrupted. Blockchain's improved elements set it appropriate for the present network safety expectations.

The aim of cybersecurity is to offer tools that may keep away from fraud and identification theft. Security breach to and manipulation of data is a persistent concern to users. Since many clients have centralized information storage, that is the case. As a result, a hacker can without difficulty benefit get admission to the data storage web pages and modify the information with dangerous intent. Such scenarios are prevented by Blockchains, with the help of its distributed data storage. A lot of sensitive information i.e. banking transactions, are probably stored in thousands of computers, resulting in a copy of the information. The intruder can invade an extensive number of systems containing duplicate of data, changing only a few machines would not impact the other information in the network. Blockchain technology also can be used to save you identification theft. Today's data theft events, according to NASDAQ, are the result of inadequate data management¹¹.

Customers are regularly forced to disclose greater records than is essential so one can achieve offerings from many businesses. The organizations claim to protect information, yet hackers frequently identify exploitable security flaws and infiltrate data storage places. Identity theft can be stopped by blockchain, through a decentralized identity system by validating that consumer is what they say they are. This may be completed with an uniform approach that permits all companies to confirm the identification of customers without requiring them to reveal touchy details. It reduces the risk of compromise and makes stealing user data more expensive for hackers.

III.METHODOLOGY

In the current network protection business, use of Blockchain innovation will be assessed involving subjective analysis of auxiliary information in this review. It will focus on Taylor et al 2019.'s study, which checked out 30 latest studies of research on Blockchain cybersecurity application cases. Basically this article will focus on two elements of each of the publications that have been highlighted. To start, it'll observe the maximum current cybersecurity packages of the rising Blockchain technology. Next, it will check the various deployment strategies for Blockchain cybersecurity solutions. Many studies and evidence can be used to arrange a dialogue on the topic of Blockchain offering protection in the IT consumer domain.

IV.DISCSSION AND RESULTS

On the basis of the results of the 30 research, it is concluded that Blockchain has been discovered to be greater possible in Internet of Technology, network security, and data storage protection. The reference chart underneath represent present executions of Blockchain security such as public key infrastructure, information, IoT:

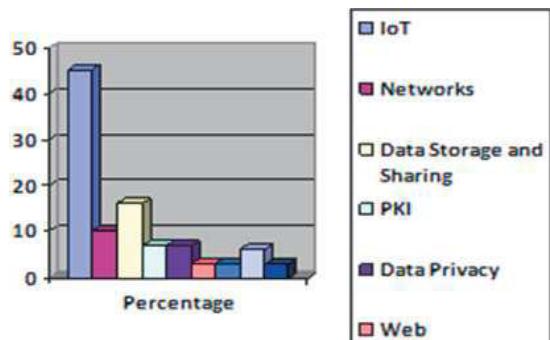


Fig 3: Areas of Blockchain security research that have received the most attention

Because there are currently 9 billion IoT devices, the focus on Blockchain IoT security is understandable. Many of these devices have insecure security configurations, and many are hacked. The Mirai Botnet, which has been around for a while, is made up of IoT devices. Employed with high success rates against large targets like DynDNS is one of the largest domain name resolving services on the Internet¹².

It can also be seen that researchers are exploring the application of Blockchain in securing networks. Most of the research in these areas is around authentication since current network security measures such as security through WPA encryption can be compromised. Lastly, there is significant attention towards Blockchain security solutions for data privacy. Most of the studies are evaluating ways of protecting personally identifiable information through a universal Blockchain authentication scheme. This will negate the need for users to send their information to organizations, instead, the organizations will authenticate users via the Blockchain.

For instance, a digital interruption on Yahoo in 2014

brought about the burglary of information from three billion users¹³. Subsequently, security analysts are searching for Blockchain security answers for information stockpiling areas, like cloud stages. Obviously specialists are investigating the utilization of Blockchain in network security. Since momentum network security instruments, like security through WPA encryption, can be broken, the majority of the examination there centers around confirmation. At long last, Blockchain security answers for information protection are getting a ton of consideration. Most of the exploration is centered around deciding how to safeguard by and by recognizing data utilizing a widespread Blockchain validation framework. People will presently not be expected to give their data to organizations; all things considered, ventures will utilize the Blockchain to validate clients. The examination's subsequent spotlight is on how Blockchain innovation may be utilized to upgrade network protection. Despite the fact that current security arrangements give outstanding degrees of insurance to IT resources, they are in any case defenseless to disappointment. This is on the grounds that most security items are set up to chip away at their own with regards to getting an IT asset. Programmers can point to a solitary security arrangement, take it disconnected, and afterward assault the now-uncovered IT asset, similarly as they can with DDoS (Distributed Denial of Service) attacks. Analysts checking out what Blockchain can emphatically mean for present degrees of safety base their contentions on conveyed security apparatuses' expanded potential to offer preferred insurance over a solitary instrument. According to the data in the bar graph above, many researchers are interested in how Blockchain can enhance the security of IoT devices, data, and networks. Unauthorized access and control of devices is the most serious security vulnerability in IoT networks. Access control and data sharing for all IoT devices may be more effectively managed using blockchain security solutions. To provide accurate user identification, authentication, and data transfer, a Blockchain security solution could be implemented.

To forestall unlawful access, it could work by safeguarding dispersed records about trusted earlier associations and meetings. New associations could be allowed provided that an enormous number of existing associations vote or check the new user. As an outcome, an IoT gadget in a home, for example, an IP camera, will just give admittance to confided in family gadgets. On the off chance that a programmer attempts to get sufficiently close to the camera, the Blockchain

arrangement will disallow them from doing as such until a greater part of the dependable gadgets votes to empower them to do as such. The presence of a weak link or compromise is the greatest shortcoming in information security, as indicated by scholars. Subsequently, information can be taken, adjusted, or lost. The security scientists depicted how Blockchain's severe foundation might be utilized to get information security. Outsiders will not be able to alter the information since each square will be hashed and associated with the following square. Since just the two people associated with the correspondence will actually want to get to and control the information, any taken data will be pointless, and outsiders will not be able to transform it. As far as organizations, security specialists found that Blockchain innovation might be used to give bunched network security, forestalling undesirable associations and correspondence. The utilization cases featured showed that Blockchain innovation is turning out to be more feasible in the field of cybersecurity¹⁴. In spite of the way that different subjects were examined, the three featured regions are the most significant in the present IT world. They show that Blockchain can close troublesome security holes that are past the compass of conventional security innovations.

I. CONCLUSION

Blockchain technologies continue to evolve and find new applications in today's environment. Cybersecurity is one of the major areas where it has been explored and implemented. The Blockchain infrastructure makes it extremely feasible for addressing existing security issues in domains like IoT devices, networks, and data transfer and storage. Taylor et al. investigated the application of Blockchain technology from the standpoint of 30 researchers in the paper. Most Blockchain security experts appear to be focusing their efforts on the adoption of Blockchain security for IoT devices. Networks and data security are also important aspects of Blockchain security. As mentioned earlier in the discussion, Blockchain technology can be utilized to safeguard IoT devices by providing more secure authentication and data transfer techniques. These can keep hackers out of these gadgets, which often come with inadequate security settings. The technology can also be used to protect networks by utilizing a strict infrastructure to prevent unauthorized connections and communication¹⁵. Finally, Blockchain can protect data during transmission and storage by encrypting blocks that can only be opened by communicating parties and are

resistant to manipulation. More use cases are being investigated, but these three are the most important. Future academics should investigate the feasibility of a single Blockchain that can be utilized to construct security solutions, as most present solutions utilize many Blockchains, making integration difficult.

REFERENCES

- [1] M. King, B. Zhu, and S. Tang, "Optimal path planning," *Mobile Robots*, vol. 8, no. 2, pp. 520-531, March 2001.
 - [2] H. Simpson, *Dumb Robots*, 3rd ed., Springfield: UOS Press, 2004, pp.6-9.
 - [3] M. King and B. Zhu, "Gaming strategies," in *Path Planning to the West*, vol. II, S. Tang and M. King, Eds. Xian: Jiaoda Press, 1998, pp. 158-176.
 - [4] B. Simpson, et al, "Title of paper goes here if known," unpublished.
 - [5] J.-G. Lu, "Title of paper with only the first word capitalized," *J. Name Stand. Abbrev.*, in press.
 - [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Translated J. Magn. Japan*, vol. 2, pp. 740-741, August 1987 [Digest 9th Annual Conf. Magnetics Japan, p. 301, 1982].
 - [7] M. Young, *The Technical Writer's Handbook*, Mill Valley, CA: University Science, 1989.
 - [8] Mengelkamp, Esther, et al. "A blockchain-based smart grid: towards sustainable local energy markets." *Computer Science- Research and Development*, 2018, 33.1, pp. 207-214.
 - [9] Gao Y, Nobuhara H. A proof of stake sharding protocol for scalable blockchains. *Proceedings of the Asia-Pacific Advanced Network*. 2017;44:13-6.
 - [10] Taylor PJ, Dargahi T, Dehghantanha A, Parizi RM, Choo KK. A systematic literature review of blockchain cyber security. *Digital Communications and Networks*. 2019, 12(5), pp. 1-14.
 - [11] Sharma PK, Moon SY, Park JH. Block-VN: A distributed blockchain based vehicular network architecture in smart City. *JIPS*. 2017, 13(1), pp. 184-95.
 - [12] "How Blockchain Can Fight Fraud Based on Know-Your- Customer Data", Nasdaq.com, 2019. [Online].
- Available: <https://www.nasdaq.com/articles/how-blockchain-can-fight-fraud-base-d-know-your-customer-data-2019-02-11>. [Accessed: 19- Sep-

2019].

[13] Kolias C, Kambourakis G, Stavrou A, Voas J. DDoS in the IoT: Mirai and other botnets. Computer. 2017, 50(7), pp. 80-4.

[14] Trautman LJ, Ormerod PC. Corporate Directors' and Officers' Cybersecurity Standard of Care: The Yahoo Data Breach. Am. UL Rev.. 2016, 66(1), pp. 1231.

[15] Kshetri N. Can blockchain strengthen the internet of things?. IT professional, 2017, 19(4), pp. 68-72.

[16] Yeoh P. Regulatory issues in blockchain technology. Journal of Financial Regulation and Compliance. 2017, 25(2), pp. 196-208.

SEGMENTATION TECHNIQUES ON MELANOMA SKIN CANCER IMAGES

Shelly Garg¹

¹Research Scholar, Department of Electronics and Communication, Punjabi University, Patiala, 147002, Punjab, India

¹Assistant Professor, Department of Electronics and Communication, SGT University, Gurgaon, India

¹shellygarg96@gmail.com

Balkrishan Jindal²

²Associate Professor, Yadavindra College of Engineering, Punjabi University campus, Talwandi sabo, 151302, Dist. Bathinda, Punjab, India

²balkrishan_76@rediffmail.com

Abstract: Melanoma is now considered one of the most common types of skin cancer and a leading public health concern. There is a focus on assembling models for the prediction of melanoma utilizing some evolutionary techniques. Using K-mean, an evolutionary algorithm referred to as the Particle Swarm Algorithm (PSO) is used first to enhance K-mean by using the Particle Swarm Algorithm. Both Enhanced Algorithms are used separately to detect and segment skin cancer images. In this study, the results of K-mean and PSO algorithms are compared and also used to achieve a high accuracy rate.

Keywords—Melanoma; K-mean; Particle Swarm Optimization; Segmentation

I. INTRODUCTION

The world's most serious form of skin cancer, melanoma, is among the main general wellbeing concerns today and becomes the fastest widening skin cancer type. If diagnosed and treated early, the prognosis of melanomas can be greatly improved and surgically cured with a 100% cure rate [1, 2]. Traditionally, dermoscopy (or "Skin Surface Microscopy") has been used as an early detection and survival tool for melanoma patients with a significant improvement in the detection and survival rates of the disease. Despite these arguments, many researchers have maintained that clinical review of melanoma detection is the uttermost reliable method and that the latest technology like advanced artificial intelligence could be an effective alternative to clinical examination [1], [2]. Melanoma is more easily detected with dermoscopy, a method that involves two separate processes, including optical magnifying and also liquid immersion methods. By using this method, the skin contact area becomes luminous, which allows viewing of the skin's design beneath the epidermis. However, relying on dermoscopic diagnosis alone can be very difficult, resulting in inaccurate results. Thusly, numerous analysts are as of late becoming keen on creating and utilizing "automatic advanced

dermoscopic image analysis" strategies to upgrade the demonstrative precision of malignant melanoma all around the world.

In like manner, a further itemized examination is frantically required [2]. Dermoscopic image investigation commonly comprises of four primary advances: a) image acquisition; b) lesion segmentation; c) feature extraction process and d) classification method. From these steps, segmentation is more significant to examine the cancer image as the exactness of different advances is exceptionally reliant upon it [1]. In the analysis of dermoscopy image, executing the segmentation process becomes more arduous for some demonstration. Instances of some elements that might affect the precision of this progression with little differentiation between the skin lesion and the encompassing skin; some unpredictable sore boundaries and skin surface; existence of air pockets and hair; and finally the existence of various tones in the skin lesion [2]. The main motivation behind this research paper is to location and division of Melanoma in skin images utilizing transformative techniques.

This research paper follows four sections. In section 1, Introduction has been described. Section 2 audits a portion of literature work. Section 3 talks about the description of the methodology. The test details and the results acquired are examined in Section 4. At long last, the Conclusion and future work are described in Section 5.

II. LITERATURE RELATED WORK

Xie et al. described in [1] using a model named Self-generating neural networks (SGNN) along with a Genetic algorithm for the automatic segmentation of dermoscopic skin cancer images. SGNN is summed up from individuality producing the process of neural tree to the process of Self-Generating Neural Forest with gathering of ideal seed tests is chosen by Genetic Algorithm at that point. Then, these ideal seeds are being utilized to create an ideal grouping apportioning of the skin cancer images. This model conveyed

to focus on precise segmentation and also other programmed techniques. Sood et al. introduced an algorithm in [2] for the identification of skin sores from given digital images utilizing the Genetic Algorithm technique. The skin lesion segmentation technique has been contrasted with the consequences of different methods. This proposed method produced high sensitivity, specificity, and also accuracy rate thought about to other algorithms. Situ et al. proposed a developmental procedure [3] based on the segmentation algorithm used on the skin lesion image. It helped to distinguish the sore naturally without proper setting of boundaries physically. This segmentation strategy was adaptable to take on other wellness capacities. Guerra-Rosas et al. developed an algorithm in [4] to isolate the skin cancer image into their separate RGB channels for acquiring the different spectral properties of every channel. It utilized the green color channel as indicated by the truth that contains more data. The authors recognized skin malignant growth in light of an examination of frequencies presented in the green color channel along with the k-law nonlinear channel. They examined the various sorts of skin diseases presenting different types of the scope of an order for every type. Esteva et al. introduced automated AI in [5] for the skin malignant growth by utilizing convolutional neural organizations strategy joined with distributing calculation on dermoscopy skin lesion dataset images that contains 2,032 number of images with various mentioned disease titles to accomplish 72% for the three-way grouping process and 55.4% for the nine-way order step. Mishra et al. [6] presented an outline framework of dermoscopy skin cancer images by sore segmentation process after that feature extraction step then applied an algorithm based on machine learning for the disease skin recognition and then classification of benign or malignant skin cancer type for the early detection. Gutman et al. [7] developed an automated algorithm to carry out and planned for the finding of melanoma lesions from dermoscopy skin medical images. To assess their outcomes, they utilized a confusion matrix additionally utilized other estimation methods. Jaleel et al. [8] proposed a mechanized model for the prediction and characterization of skin disease into two different kinds mainly benign and malignant melanoma utilizing neural organizations methods. Their model relies upon image preprocessing, extracting features and afterward they applied neural organizations algorithms. They accomplished 84% exactness. Lam et al. [9] described the PSO algorithm to look for the best centroid point with minimum mean error and to find out the nearest distance level. Kkklahli et al. [10] utilized multi-approach techniques for the detection of brain melanoma division in MR images such as k-mean, genetic segmentation, and particle swarm optimization (PSO) algorithms. From these results, particle swarm optimization methods accomplish better outcomes as compared to others. The authors assessed their proposed approaches by utilizing the rand list, and Global consistency error. Li et al. [11]

developed a hybrid algorithm by using dynamic particle swarm optimization and also the K-mean method to further develop the quality and proficiency in image segmentation. These hybrid-developed method outcomes are further compared to the K-mean method. Kaur et al.[12] assessed PSO algorithm given various methodologies like neural organizations, clustering, genetic algorithm, fuzzy framework , wavelets, and these methods applied it on a few skin cancer images for the process of segmentation. This developed approach supported that PSO given various algorithms as hybrid technique will be more effective as compared to others. Xu et al. [13], recommended an advanced programmed strategy for segmentation to start from diminished the color image change into an intensity image by utilizing a power thresholding method. At long last, they utilized a versatile bend model to address the final process of segmentation. This strategy relies upon three boundaries: such as image gradient to determine the threshold value, standard deviation value by Gaussian Smoother method and high sharpness of different color types. Then the outcomes demonstrated to calculate the average error rate for 20 irregular images that are considered equivalent to four specialists physically fragmented the images.

III. METHODOLOGY

3.1 DataSet: The Data Set from the "ISIC 2017: Skin Lesion Analysis Towards Melanoma Detection" challenge dataset [14]. Melanoma data can be collected from International Skin Imaging Collaboration. This dataset is online available. Pre-processing and image segmentation methods are applied to these dataset images. Melanoma samples have been shown in Fig.1

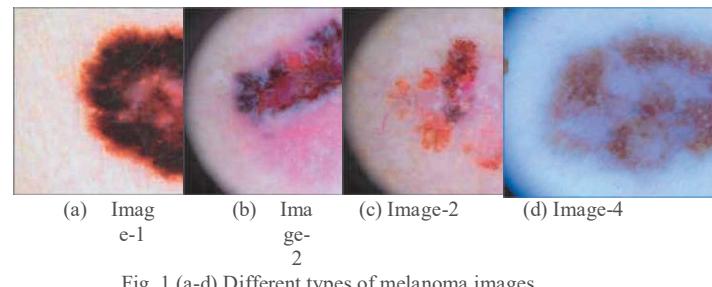


Fig. 1 (a-d) Different types of melanoma images

3.2 Preprocessing: Dataset images are minimized to 25 % scale-down by using the interpolation method [15], to resize the height or width of the original skin cancer images. From these dataset images, a few images contain unwanted hair in the preparation put have inside hair which makes the separating into parts process hard. The presence of hair additionally makes the point extraction process for further image perceptions mentioned [16].

3.3 Image Segmentation: It is a fundamental stage to analyze the skin cancer image and it tends to be characterized as the most common way of isolating and separating the locale of premium from the skin cancer image. This process is the most considered and intriguing area of computer vision. A division strategy is normally constructed and planned thinking about the specific properties of different classes of particular images [13]. If the segmentation process is precise then the lesion border can be accessed easily. This strategy utilized here relies upon the K-mean clustering process. For this cycle, the k-implies method is to be used to segment the cancer images. And another PSO method is utilized to find out the input centers for k-mean.

3.3.1 K-mean Clustering Algorithm

It is an unsupervised machine learning technique that doesn't have any name or class name which is used to take care of different issues of clustering. It is a strategy that depends on the centroid gravity of the fragment components to address the cluster. This algorithm is used to compute the distance between the cluster point and the center of the cluster for each item. This algorithm has a fundamental issue underlying the centroid haphazardly [10]. There are numerous techniques for clustering produced for a wide scope of purposes. K-implies algorithm used to find out the groupings ghastly by regularly. This is the easiest method and is utilized to make the cluster according to color pixels based on intensive values. This clustering algorithm depends upon the number of K elements [17]. There are some advantages to using the K-mean Technique such as with countless factors, K-Means might be computationally quicker than various leveled grouping (on the off chance that K is little). K-Means might deliver more tight groups than progressive clustering, particularly on the off chance that the groups are globular. Some disadvantages of using the K Mean Technique are several number of fixed clusters that can make it difficult to get out whatever will happen later on. It doesn't function admirably with non-globular gatherings of things.

3.3.2 Particle Swarm Optimization (PSO)

This is one of the most famous transformative computation procedure structures motivated by naturally friendly conduct and dynamic developments with the correspondence of birds and fish for looking through conquer food source element [10]. In this, there are a colossal number of particles that look for the ideal arrangement and speak with different particles simultaneously. Every molecule refreshes its situation as indicated by best location, global best location, and speed, by

the accompanying equation [10]. This optimization technique is used to enhance the image features and to resolve a complex problem. In this technique, a threshold value is characterized based on which the background space is replaced with the foreground to recognized skin cancer appropriately. The number of cycles that consistent while including k-implies for PSO for the different number of skin cancer images.

IV. RESULTS

This experiment is done on different types of melanoma images. In this, the results for both models are shown in Fig. 2 &3. In Figure 2, the First, input melanoma image is being pre-processed to remove unwanted noise. In the next step, this pre-processed image is segmented by using K-Mean with an accuracy of 87.7% and an error rate is 12.27. Then again this pre-processed image is segmented by using PSO with an accuracy rate is 88.3% and an error rate is 11.687. In Figure 3, the First, input melanoma image is being pre-processed to remove unwanted noise. In the next step, this pre-processed image is segmented by using K-Mean with an accuracy of 94% and an error rate is 5.43. Then again this pre-processed image is segmented by using PSO with an accuracy rate is 95% and an error rate is 4.86. From the comparison of these results, it has been shown that the accuracy rate of PSO is higher than K-mean.

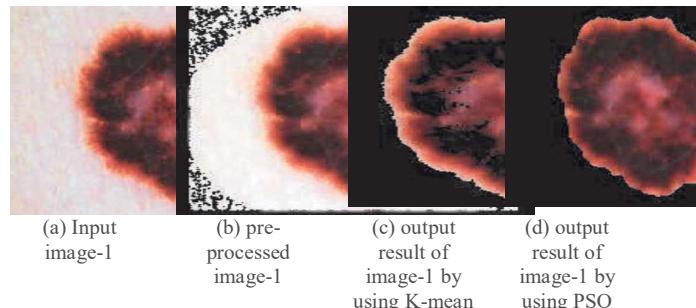


Fig.2 Experimental results on melanoma image-1 (a) Input image-1 (b) pre-processed image-1 (c) output result of image-1 by using K-mean (d) output result of image-1 by using PSO

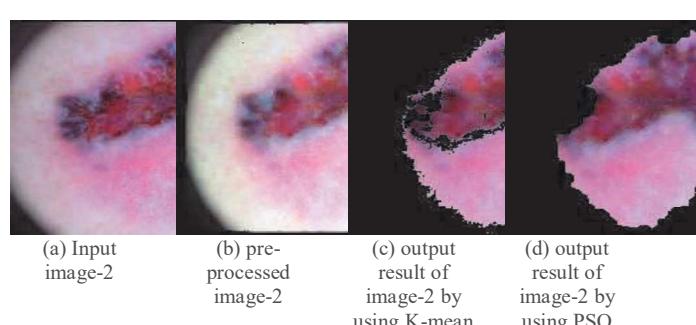


Fig.3 Experimental results on melanoma image-2 (a) Input image-2 (b) pre-processed image-2 (c) output result of image-2 by using K-mean

(d) output result of
image-2 by using PSO

V.CONCLUSION AND FUTURE WORK

In this paper, a model is designed to detect melanoma in skin images utilizing developmental algorithms. K Mean has been upgraded by utilizing an evolutionary algorithm named Particle Swarm Algorithm (PSO). The enhanced algorithm and k-mean independently are utilized to divide skin images. From the experimental results, it has been concluded PSO Algorithm

is used to achieve a high accuracy rate as compared to the K-mean algorithm. In future work, some other evolutionary algorithms to instate k-mean will be applied to the other skin types. This work can be carried out and also apply to different applications.

REFERENCES

- [1] Xie, F., & Bovik, A. C. (2013). Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. *Pattern Recognition*, 46(3), 1012-1019.
- [2] Sood, H., & Shukla, M. (2014). Segmentation of Skin Lesions from Digital Images using an Optimized Approach: Genetic Algorithm. *International Journal of Computer Science and Information Technologies*, 5 (5) , 6831-6837.
- [3] Situ, N., Yuan, X., Zouridakis, G., & Mullani, N. (2007). Automatic segmentation of skin lesion images using evolutionary strategy. In *Image Processing IEEE International Conference*, Vol. 6, pp. VI-277.
- [4] Guerra-Rosas, E., lvarez-Borrego, J., & Coronel-Beltrn, (2014). Diagnosis of skin cancer using image processing. In *AIP Conference Proceedings*, 1618(1), 155-158.
- [5] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- [6] Mishra, N. K., & Celebi, M. E. (2016). An overview of melanoma detection in dermoscopy images using image processing and machine learning. *arXiv preprint arXiv:1601.07843*.
- [7] Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., & Halpern, A. (2016). Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). *arXiv preprint arXiv:1605.01397*.
- [8] Jaleel, J. A., & Salim, S. (2013). Implementation of ANN Classifier using MATLAB for Skin Cancer Detection. *International Journal of Computer Science and Mobile Computing*, 87-94.
- [9] Lam, Y. K., Tsang, P. W. M.& Leung, C.S. (2013). PSO-based K-Means clustering with enhanced cluster matching for gene expression data. *Neural Computing and Applications*, 22(7-8), 1349-1355.
- [10] Kkklahli, E., Erdomu, P., & Polat, K. (2016). Brain MRI Segmentation based on Different Clustering Algorithms *Brain*, 155 (3).
- [11] Li, H., He, H., & Wen, Y. (2015). Dynamic particle swarm optimization and K-means clustering algorithm for image segmentation. *Optik- International Journal for Light and Electron Optics*, 126(24), 4817-4822.
- [12] Kaur, A., & Singh, M. D. (2012). An overview of pso-based approaches in image segmentation. *International Journal of Engineering and Technology*, 2(8), 1349-1357.
- [13] Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C.,& Huntley, A. (1999). Segmentation of skin cancer images. *Image and Vision Computing*, 17(1), 65-74.
- [14] Codella N, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza S, Kalloo A, Liopyris K, Mishra N, Kittler H, Halpern A. "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)". *arXiv: 1710.05006 [cs.CV]*
- [15] R. Gonzalez, *Digital Image Processing*. Pearson Education, 2009. [Online]. available: <https://books.google.co.in/books?id=a62xQ2rf8wC>
- [16] Frangi, A.F. Niessen, W. J., Vincken, K. L, and Viergever, M. A., (1998) "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 130– 137.
- [17] Jain, A. K., (2010). Data clustering: 50 years beyond kmeans. *Pattern Recognition Letters*, Elsevier, 31(8), 651 – 66

Digital India: An insight into new India

Kavita Verma*

School of Management

Apeejay Institute of Management & Engineering

Technical Campus,

Jalandhar, Punjab

E-mail: vermakavita770@gmail.com

*First Author

Sunny Gulati**

School of Management

Apeejay Institute of Management & Engineering

Technical Campus,

Jalandhar, Punjab

E-mail: sunnygulati1234@gmail.com

** Corresponding Author

Abstract

The purpose of the study is to provide an insight into evolution of digital growth in India. Digital India has been vision of the country's leaders from way back. The journey of the digital India along with vision of our leaders is discussed in the study. Further the light has been shed on 9 pillars of digital India which had helped the country to grow and become digitally independent. The concluding remarks involve that how as a nation we have implemented digital India in every possible sphere and gained success for the same.

Introduction

India being a developing country has been a land of trying new things for making the nation a sustainable economy. Increasing literacy rates highlighted the need for advance technology. Knowledge development can be dated back to 1980s when Vijay Bhatkar the father of Indian supercomputer quoted that "*Great nations are not built on borrowed technology*". The Superpowers stopped the export of a

Cray supercomputer fearing the idea of developing countries competing with superpowers like America and Europe. However, this bane proved to be an opportunity for India. India took advantage of this opportunity and set up a Centre for development of advanced computing (C-DAC) in March 1988. The idea behind the centre was to become independent not only in terms of technology but also in research and development. The, then Prime Minister of the country Mr. Rajiv Gandhi had his vision clear about the future of technology in India.

Recently Sam Pitroda, in one of his speech, said that the former Prime Minister, Mr. Rajiv Gandhi always believed in the idea of using technology as a stimulus in the process of modern India. The legacy was carried forward by then Finance Minister of India, Dr. Manmohan Singh in 1991, when he bought one of the biggest reforms of the country into action. The policy of Liberalization, Privatization and Globalization commonly known as LPG policy made our economy the biggest

economy for the world business houses to work in. The investment was welcomed with open hearts and hands, making India a hub for businesses of all sectors.

Same vision was shared by another former Prime Minister Mr Atal Bihari Vajpayee when he saw the vision of making country a nuclear power. He not only made Pokhran mission possible in 1998 but made sure that the country made a huge sound in the entire world. The same sense of vision is being shared by the present Prime Minister of India Mr Narendra Modi who launched the flagship program name Digital India in the year 2015 giving the name and shape to the dreams of all the previous leaders of the country. He has laid down the 9 pillars for pacing up the Digital India Campaign.

9 Pillars of Digital India



Broadband Highways

Broadband Highway is a broader aspect for: Broadband for all Rural, Broadband for all Urban and National Information Infrastructure. Project cost for Broadband for all Rural is approximately ₹32,000 crores. The nodal department for the same will be department of telecommunications. This project will cover 2 50,000 village panchayats by the end of December 2016.

Similarly, the priority under broadband for all urban would be developing a communication infrastructure in all old and new urban areas along with providing virtual network operators that will leverage the service delivery.

For another aspect of Broadband highway which is National Information Infrastructure, the very first step would be integrating the networks like SWAN, NKN and NOFN with cloud enabled national and state data centres. It will also focus on providing horizontal connectivity of 250, 20, and 5 government offices or service centres at state, district, block, and panchayat levels. Nodal department for the third aspect will be DeitY and the project will cost roundabout 15686 crore and the implementation time would be of 2 years and maintenance and support for 5 years.

Universal Access to Mobile Connectivity

The universal access to mobile connectivity is an initiative which will help in network penetration and covering the gaps of connectivity in the country. The estimated 42,300 villages are not yet covered for universal mobile connectivity in the country. The Department of Telecommunication will serve as a nodal department for the project. The estimated cost of the project is ₹16,000 crore and must be completed within the financial year 2014-18.

Public Internet Access Programme

The Public Internet Access Program is further bifurcated into 2 subcomponents namely common service centres and post office as multi service centres. The first component being common service centre must be strengthened and its number is to be increased from 1,35,000 at present to 2,50,000 which means that there will be one common service centre in each gram panchayat. Common Service Centres would be multi-functional and feasible endpoints for delivery of government business and services. DeitY would serve as nodal department for implementing the scheme.

Proposal has been given to convert 1,50,000 post offices into multi service

centre. Department of post will serve as nodal department to implement the scheme.

E-Governance: Reforming Government through Technology

Government Business Process Re-engineering using IT aims to improve the transactions, it stands out to be the most significant change across the government sector and can be successful if implementation of the programme is done by all ministries and departments. Government has given the guiding principles for this technological reform which are:

- The first principle being simplifying the forms and field reduction. This principle aims in making the forms user friendly and simple and the focus is to collect only the necessary information which is required.
- The application tracking and interface should be provided online between the departments to make it more accessible.
- Online repository is to be used for obtaining proofs like school certificates, voter ID card, and all other kinds of identification proofs, instead of showing the physical form of these documents.

- For facilitating the integration and interoperable service delivery system for citizens the service platforms like UIDAI, Payment Gateways, Mobile Platforms, Electronic Data Interchange should be integrated, and this should be mandated.

Moreover, the entire database should be electronic instead of manual. To make system more transparent the entire workflow conducted inside government departments and agencies should be automated and this will not only ensure efficiency in the government processes but will also build the faith between the citizens and the government. Information technology can also be used for public grievance redressal it will not only help in timely response but more efficient analysis of the data to resolve the persistent problems this step will largely improve the processes.

E-Kranti - Electronic Delivery of Services

There are 31 Mission Mode Projects under different stages of e-governance project lifecycle. Further, 10 new MMPs have been added to e-Kranti by the Apex Committee on National e-Governance Plan (NeGP) headed by the Cabinet Secretary in its meeting held on 18th March 2014.

The government has started 31 Mission Mode Projects for different stages of e-governance project lifecycle. The apex committee on National e-Governance Plan being headed by the Cabinet Secretary in its meeting held on 18 March 2014, has added 10 new MMP's to the project. The government has specified steps to be taken in different sectors like education, healthcare, farming, security, financial inclusion, justice, planning and cyber security.

- **Technology for Education: E-Education:** The vision is to make every school independent in terms of connectivity. The first step is to cover all secondary and higher secondary schools with the approximate number of 250,000 schools. To leverage the e-education, development of MOOCs-Massive Online Open Courses will be paced up, and a programme on digital literacy will be taken up at national level.
- **Technology for Health:E-Healthcare:** This will focus on providing better health assistance to the entire country by going online. The services like consultation, medical records, medicine supply will be provided online, pan-India. The project

started in 2015 and was likely to completed within a span of 3 years.

- **Technology for Farmers:** Providing the facilities like price information, ordering of inputs, loan, and relief payments online to farmers will not only save time of the farmers but of government officials also.
- **Technology for Security:** This will also provide relief to citizen in times of disasters. Mobile based emergency services will provide citizen with information on real time basis and will help them to take precautionary measures in time to avoid loss of lives and properties.
- **Technology for Financial Inclusion:** Financial Inclusion will strengthen by increasing the usage of banking services like Mobile Banking, Micro-ATM program and CSCs/Post Offices.
- **Technology for Justice:** Criminal Justice will strengthen by investing more in E-Courts, E-Police, E-Jails, and E-Prosecutions.
- **Technology for Planning:** For project planning, conceptualization, design and development, National GIS Mission Mode project will be implemented, and this will

facilitate decision making also in the respective fields.

- **Technology for Cyber Security:** For safe and secure cyber space in the country, National Cyber Security Co-ordination Centre will be setup.

Information for All

- Information for all aims on providing information and documents on online data platforms, providing open and easy access to information.
- Government will proactively interact with citizens via social media and other web-based platforms for this MyGov.in has already been launched where government invites ideas and suggestions from the citizens. This will facilitate the 2-way communication between government and citizens.
- The citizens on the special occasions or programs will be addressed by the government via online messaging through emails and SMSs.

Electronics Manufacturing This pillar targets on NET ZERO Imports

which requires coordinated efforts and actions in many fronts.

- a. Taxation, incentives
- b. Economies of scale, eliminate cost disadvantages
- c. Focus areas – Big Ticket Items FABS, Fab-less design, Set top boxes, VSATs, Mobiles, Consumer & Medical Electronics, Smart Energy meters, Smart cards, micro-ATMs
- d. Incubators, clusters
- e. Skill development
- f. Government procurement

IT for Jobs

- The first initiative under this will be training around 1Cr. students from smaller towns and villages for IT sector jobs over the period of 5 years. The nodal department for this scheme will be served by DeitY.
- For increasing the connectivity in north-eastern states, BPOs would be setup in every state of northeast for facilitating the growth and the nodal department for the scheme will be DietY.
- For successfully running businesses for delivering IT services 3,00,000 service delivery agents will be trained. Nodal

department for the scheme will be DeitY.

- Telecom service providers will train 5,00,000 rural workforces for providing the employment and catering their own needs. Department of telecom will serve as nodal department for the scheme.

Early Harvest Programmes

- An IT platform for mass messaging will be developed by DeitY, who's members will be elected representatives and all the government employees. This database will connect 1.36 crore mobiles and 22 Lakh emails.
- The greetings by government named as E-Greetings has been made live on 14th August 2014, a basket has been made with templates of E-Greetings and is available on MyGOV platform for ensuring crowd sourcing.
- Biometric attendance is another step under this pillar. All the central government offices in Delhi are already under this operation headed by DeitY. This has been initiated in department of Urban Development also. Other departments are also on board with this step.

- Wi-Fi in all universities is ensured by National Knowledge Network, shall be covered in this scheme. The nodal ministry for the same will be Ministry of Human Resource Development.
- For interaction with employees' email has been made primary mode of communication. Phase I will focus on updating the email IDs of approximately 10,00,000 employees. In Phase II the target is of 50 lakh employees by March 2015 which will cost approximately ₹98 Crore. Nodal Department for this scheme will be DeitY.
- For the same purpose government is making standardised templates for email which will be ready by October 2014.
- Cities which have population more than 1 million and are tourist hubs will be provided with public Wi-Fi to promote digital cities. The scheme will be implemented by department of telecommunication and MoUD.
- The next initiative is to convert all the books into eBooks. Nodal agencies for this scheme will be Ministry of Human Resource Development/ DeitY.
- Weather updates and disaster alerts will be provided via SMS mobile seva platform will be provided by deity and is already ready and available for the purpose. MoES (IMD)/MHA(NDMA) will act as nodal organizations for implementing the scheme.
- National portal will be designed for lost and found children which will provide real time information gathering and sharing about these children. This will also help in quick response and will help in avoiding the crime against children. DeitY/DoWCD will be the nodal departments for the project.

Conclusion

India Today is seeing a big revolution in technology sector. The country started with STD PCO and we are now heading towards 5G spectrum today, with the mobile in every house. The journey started with BSNL and a major chunk of responsibility is being shared by Jio today. This evolutionary growth was not only limited to telecom sector but all other sectors also saw the same growth, Odisha has been prone to destruction lead by the cyclones for many years. The first cyclone experienced by Odissa was in 1737 and the latest in 2021. But over the years the states destruction management against cyclone

has improved and it has minimalised the loss occurred by cyclone this has been only possible because of the development of Indian Metrological Department. Timely warnings by Indian Metrological Department has helped the state to not only save the property but also many human lives. The technological development in satellite forecasting has been a gift for the entire country.

Next big digital reform was seen in banking sector with introduction to Core Banking System, this system has not only improved the connectivity among the branches of the banks but is the single biggest power working behind the effective change of transactions and insightful handling of banking business in the country.

Today India has the highest demographic dividend and to make the best out of this demographic index the country has to take up with all these opportunities to channelize the youth in the right direction. Major chunk of the population is in between the age of 30 to 40 years making us the youngest country in the entire world. This demographic dividend is a cherry on the cake for our digital India campaign.

References

1. <https://www.thebetterindia.com/82076/india-first-supercomputer-param-cdac-vijay-bhatkar/>
2. <https://economictimes.indiatimes.com/news/politics-and-nation/rajiv-gandhi-contributed-lot-to-nation-building-sam-pitroda/articleshow/78748627.cms>
3. <https://www.financialexpress.com/india-news/the-story-of-pokhran-tests-that-established-india-as-nuclear-power-became-cornerstone-of-atal-bihari-vajpayees-tenure-as-pm/1283660/>
4. <https://odishabytes.com/odishastryst-with-cyclones-1737-2021/>
5. <https://www.indiatoday.in/india/story/odisha-encounters-10-cyclones-two-decades-jawad-1884225-2021-12-05>
6. <https://www.downtoearth.org.in/news/historically-prone-to-cyclones-42421>
7. <https://www.rbi.org.in/scripts/bsviewcontent.aspx?Id=982>
8. <https://www.forbesindia.com/blog/enterprise/does-indias-telecom-sector-still-need-restrictive-policies/>
9. <https://vikaspedia.in/e-governance/digital-india/nine-pillars-of-digital-india>
10. <https://www.digitalindia.gov.in/content/programme-pillars>

Improved Lesion Detection using Morphological Operations

Shakti Kumar

*Research Scholar, Department of Computer Science and Applications
Panjab University
Chandigarh, India
shaktibajpai@pu.ac.in*

Anuj Kumar

*Assistant professor, Department of Computer Science and Applications
Panjab University
Chandigarh, India
anuj_gupta@pu.ac.in*

Abstract – To study the segmentation of Lesions from infected skin images the experimental study was performed using edge detection and morphological filters. For the study PH2, the publicly available dataset was used. The dataset holds the ground truth information of lesion which was compared using SSIM with the Threshold opening-closing by reconstruction. The result of the study emphasizes an alternative method for displaying the segmented object. The clear structural similarity is observed using a structural similarity image map of ground truth and computed results. The proposed approach helps to remove the unwanted background information and provides information pertinent for the diagnosis of melanoma.

Index Terms – *Preprocessing, Markers and boundaries, Lesion Segmentation, Gradient Magnitude, SSIM.*

I. INTRODUCTION

Melanoma is a type of skin cancer. As compared to Basal cell carcinoma (BCC) and Squamous cell carcinoma (SCC) malignant melanoma is the most dangerous form of skin cancer as it has the highest mortality rate. Melanocytes are responsible for skin pigmentation. The uncontrollable growth of these melanocytes leads to melanoma skin cancer. UV rays exposure is one factor that leads to this melanoma skin cancer. It must be detected at the early stages because early detection leads to less death rates. In automatic detection of melanoma for given infected skin area image the automatic system produces results in the form of simple benign or malignant melanoma. In Automatic Melanoma Detection System (AMDS) four stages play a vital role [1]. First is the preprocessing stage for noise removal including the unnecessary and unwanted components like hairs, and scar marks. The second is the segmentation stage used to extract the lesion as a region of interest from infected skin.

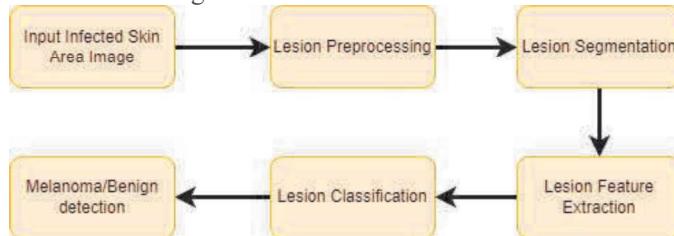


Fig 1. Stages of Automatic Melanoma Detection System

The third is feature extraction to extract features like Asymmetry, Border, Colour, and Shape and the Fourth is the Classification to classify the results as benign or melanoma. In the proposed approach the strategy followed is based on pre-processing step so that better segmentation results can be obtained. The morphological operation performed on the lesion image gives the result which is close to the ground truth for the given image taken from the PH2 Dataset [6] and can be used in the segmentation sub-stage of the Automatic Melanoma Detection System. The ensuing section covers related work in literature, methods, and proposed methodology, followed by experiments, results and conclusion.

II. RELATED WORK

To design the proposed approach we have studied selective literature close to our work. Such types of vision systems include Pre-processing, Segmentation, Feature Extraction, and Classification as a common phase [1]. According to Abdallah, Y. et. al the edge detection and morphological analysis are applicable for the segmentation of Thyroid scintigraphy. Thresholding, clustering, and morphology are the vital operations to carry out the segmentation [3]. Different morphological operations help to distinguish the foreground and background properties of an object hence making the segmentation task easier and more accurate. Halder, A et. al, proposes an adaptive morphology-based approach for segmentation. According to authors [5], the segmentation of MRI brain images is done using Thresholding and Morphological operations. “Authors [11]” provided a unique 3-D segmentation technique using morphological operations. “Authors [12]” performed brain tumour segmentation using hybrid clustering and morphological operations. From the study carried out for the state-of-the-Art techniques, it is observed that segmentation will produce more accurate results if the object to be segmented goes through morphological operations.

III. PROPOSED METHODOLOGY

Removal of unwanted artifacts from the infected area of the skin image is a challenging task. Segmentation techniques like watersheds perform well if carried out using the detection of foreground lesions and background skin. The proposed approach uses the following procedure:

- A. Segmentation function computation focuses on dark regions as an object to be segmented.
- B. Foreground Marker Computation which finds out connected blobs of pixels.
- C. Background Marker Computation to identify pixels that are not part of any object.
- D. Segmentation function Modification for proper identification of foreground and background marker.
- E. Computation of Structural Similarity Index for observed and actual lesion mask.

The proposed algorithm is given below.

Proposed Algorithm:

- Step1. Read the colored Lesion Image. And Convert it to Gray Scale.
- Step 2. Define Segmentation function as Gradient Magnitude.
- Step 3. Perform foreground step marking.
- Step4. Compute Threshold Opening Closing by Reconstruction and Complement the results.
- Step 5. Find SSIM.
- Step 6. Visualize the Results

The proposed algorithm listed above is carried out on dermoscopic image PH2 Dataset [6] which is publically available for researchers. The dataset cover manual segmentation and clinical diagnosis carried out by dermatological experts. It contains a total of 200 images. Out of which 80 belong to common nevi, 80 belong to atypical nevi and the remaining 40 are from the category of malignant melanoma.

IV. EXPERIMENTS AND RESULTS

The experimental observation was performed to study accurate lesion detection methods. The results obtained are compared with ground truth using SSIM to find out the structural similarity between thresholds-based opening-closing operation and ground truth-based mask.

Step 1 The conversion of the RGB Lesion image to its Gray Scale as shown in Figure 2.



Fig. 2 RGB to Gray Scale Converted Image (Original Image Source: PH2)

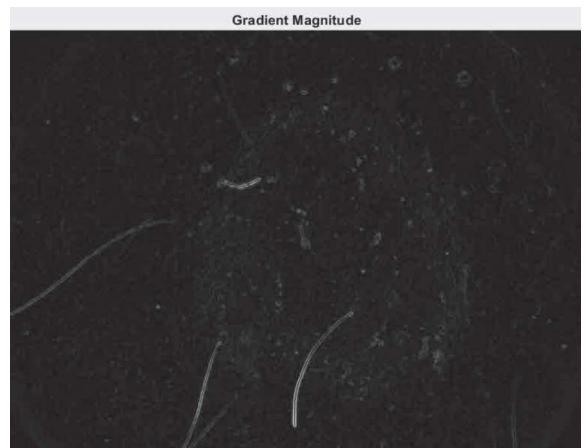


Fig. 3 Gradient Magnitude High at Border and Low a remaining places

Step 2 Application of Gradient Magnitude for Segmentation Function as shown in figure 3 it can be observed that the gradient is having a high value at the lesion border and low inside the lesion boundaries.

Step 3 Foreground Object Marking

For marking the foreground object that is Lesion different techniques are applicable. Here the applicability of morphological operations is emphasized. Opening by construction and closing by construction operations are used to clean up the image. This operation helps to find out maxima.

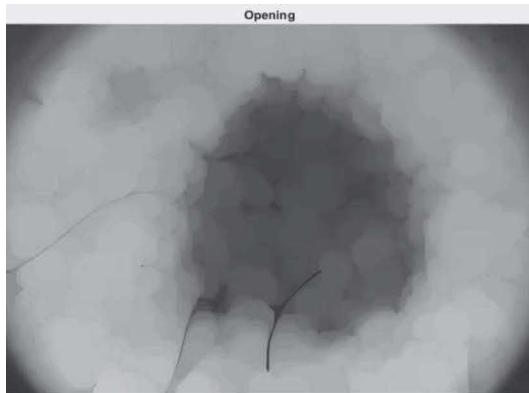


Fig. 4 Opening Operation

In opening operation erosion is followed by dilation.



Fig. 5 Opening by Reconstruction Operation

In opening by reconstruction, erosion is followed by morphological reconstruction.

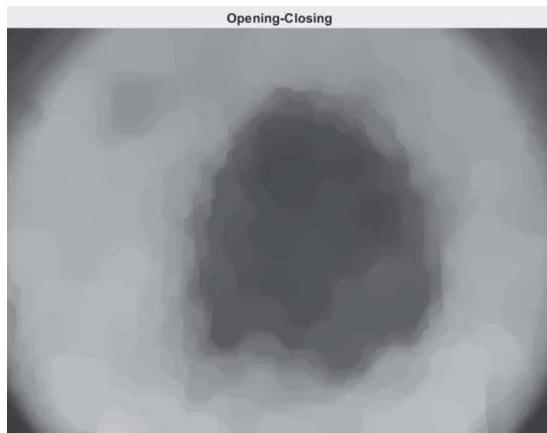


Fig. 6 Opening Closing Operation

To remove dark spots in the lesion image further Opening and closing operation is applied.



Fig. 7 Opening Closing by Reconstruction Operation

More Accurate results are obtained using Opening and closing by reconstruction. For the Foreground marker, Regional Maxima is calculated and superimposed with the original image. The computation of regional maxima and results of lesion image with a superimposed value of regional maxima is shown in figure 9.

Step 4 Background Marker Computation

As dark pixels belong to the background Threshholding based Opening-closing by reconstruction is used and results are obtained as shown in Figure 7.

Step 5 Computation of structural similarity

Structural similarity is computed on Ground Truth available in a dataset with a comparison of with complemented value of Threshold opening-closing by reconstruction. The SSIM value close to value 1 shows symmetry in structure for the computed values while a value close to 0 shows dissimilarity.

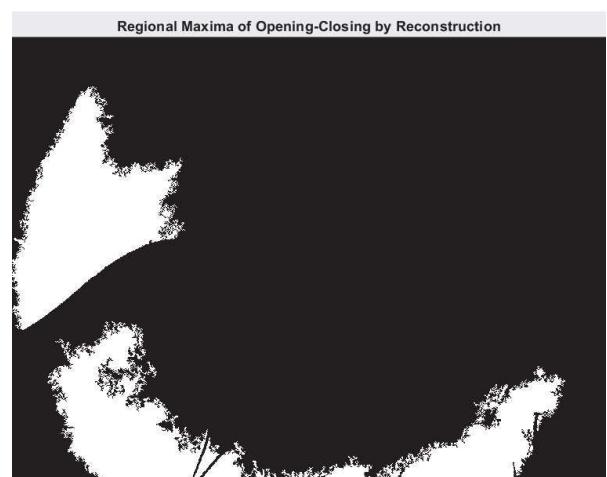


Fig. 8 Regional Maxima of Opening Closing by Reconstruction Operation

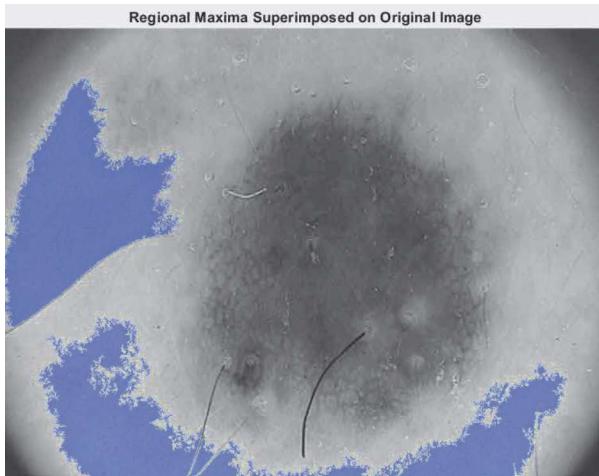


Fig. 9 Regional Maxima superimposed on Lesion Image



Fig. 10 Threshold opening-closing Vs. Ground Truth

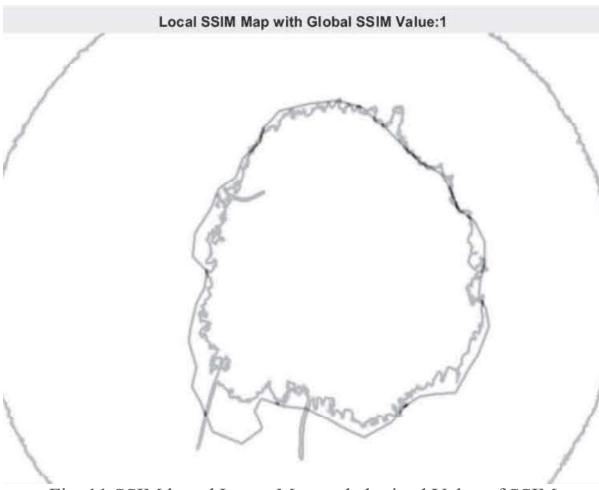


Fig. 11 SSIM based Image Map and obtained Value of SSIM

Step 6 Then SSIM is calculated for the ground truth and obtained results and the value obtained close to 1 indicated that the segmented results after the threshold-based opening-closing morphological operation are close in structure to ground truth.

The output obtained in the present stage can be forwarded to the remaining sub-stages of the AMDS. The AMDS system

will produce more accurate results if preprocessing on lesions is carried out appropriately [1].

V. CONCLUSION

The experimental study was conducted on PH2 publicly available dataset to find out an alternate approach for lesion detection. The ground truth values given in the dataset were compared with the segmented results obtained from the threshold-based opening-closing morphological operation. The SSIM value is close to 1 in the result showing that the obtained results were close to the ground truth in the dataset. Due to various artifacts available in the lesion images, the preprocessing step is vital. The output of this preprocessing sub-stage of AMDS can be forwarded to further sub-stages of the system. The study of behavior and the accuracy of the complete automatic melanoma detection system using the given approach can be considered as a future scope for the researcher.

REFERENCES

- [1] Kumar, S., & Kumar, A. (2020, February). Automatic Melanoma Detection System (AMDS): A State-of-the-Art Review. In *International Conference on Emerging Technology Trends in Electronics Communication and Networking* (pp. 201-212). Springer, Singapore.
- [2] Abdallah, Y., & Wagiallah, E. (2014). Segmentation of Thyroid Scintigraphy Using Edge Detection and Morphology Filters. *International Journal of Science and Research (IJSR)*, 3(11), 2768-2772.
- [3] Atikah, L., Hasanah, N. A., Sarno, R., Fajar, A., & Rahmawati, D. (2020, September). Brain Segmentation using Adaptive Thresholding, K-Means Clustering and Mathematical Morphology in MRI Data. In *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)* (pp. 161-167). IEEE.
- [4] Halder, A., Chatterjee, S., Dey, D., Kole, S., & Munshi, S. (2020). An adaptive morphology-based segmentation technique for lung nodule detection in thoracic CT image. *Computer Methods and Programs in Biomedicine*, 197, 105720.
- [5] Myint, H. H., & Aung, S. L. (2020, February). An efficient tumor segmentation of MRI brain images using thresholding and morphology operation. In *Proceedings of the eighteenth international conference on computer applications (ICCA 2020)*.
- [6] Mendonça, T., Celebi, M., Mendonca, T., & Marques, J. (2015). Ph2: A public database for the analysis of dermoscopic images. *Dermoscopy image analysis*.
- [7] Rout, R., Parida, P., Alotaibi, Y., Alghamdi, S., & Khalaf, O. I. (2021). Skin lesion extraction using multiscale

- morphological local variance reconstruction-based watershed transform and fast fuzzy C-means clustering. *Symmetry*, 13(11), 2085.
- [8] Wang, X., Jiang, X., Ding, H., Zhao, Y., & Liu, J. (2021). Knowledge-aware deep framework for collaborative skin lesion segmentation and melanoma recognition. *Pattern Recognition*, 120, 108075.
- [9] Ceballos-Arroyo, A. M., Robles-Serrano, S., & Sanchez-Torres, G. (2020). A morphological convolutional autoencoder for segmenting pigmented skin lesions. *Eng. Lett*, 28(3).
- [10] Irfan, R., Almazroi, A. A., Rauf, H. T., Damaševičius, R., Nasr, E. A., & Abdalgawad, A. E. (2021). Dilated semantic segmentation for breast ultrasonic lesion detection using parallel feature fusion. *Diagnostics*, 11(7), 1212.
- [11] Höhne, K. H., & Hanson, W. A. (1992). Interactive 3D segmentation of MRI and CT volumes using morphological operations. *Journal of computer assisted tomography*, 16(2), 285-294.
- [12] Zhang, C., Shen, X., Cheng, H., & Qian, Q. (2019). Brain tumor segmentation based on hybrid clustering and morphological operations. *International journal of biomedical imaging*, 2019.

Hybrid Computer Vision Approach for Identification of Medicinal Leaves

Silky Sachar

Department of Computer Science and Applications
Panjab University
Chandigarh, 160014, India
silky.sachar@pu.ac.in

Dr. Anuj Kumar

Department of Computer Science and Applications
Panjab University
Chandigarh, India
anuj_gupta@pu.ac.in

Abstract – Plants have numerous benefits for the human beings. The study of Ayurveda suggests that the medicines availed from the plants can be obtained at cheaper costs and have proven to be more affective with minimum to none side effects. For the same reason, there is extensive research going on in the field of automatic detection of medicinal plants using the images of their leaves. In this paper, we have explored the field of transfer learning as it can be applied on smaller datasets when the resources are limited. We propose a hybrid approach where the pre-trained deep learning models like VGG16, VGG19, ResNet and MobileNetV2 were used to extract the features from the leaf images and traditional machine learning models namely SVM, Random Forest, Decision Trees and KNN were used for the classification. The medicinal leaf dataset consisting of 30 species was taken from the Mendeley. Each species has about 60 images. Validation accuracy of 96.18% was observed using MobileNetV2 and KNN classifier.

Index Terms – Medicinal Plants, Deep learning, CNN, classifiers.

I. INTRODUCTION

Extensive research is going on in the field of Ayurveda which is a study aiming at obtaining medicines from the plants. In certain cases, the parts of plants can themselves be used to cure diseases like obesity [1]. The most important benefit of these medicines is that they don't cause harsh side effects like their allopathic counterparts and are also not as expensive. This draws the attention of people all over the world to automatically detect plants and use them for minor or major ailments.

Computer vision techniques have been applied in many areas including automatic segmentation of plants from their natural habitats, segmentation of weeds from the crop, detection of plant diseases, etc. The use of these techniques reduces the chance of human bias and also allow the botanists to be less burdened. In this mobile era, by the use of hand-held devices, different species can be easily recognized even by the person who doesn't have much botanical knowledge. Machine learning is the science of making computers learn how to make decisions based on the previous learnings or experiences. Traditional machine learning techniques in image processing initiate with pre-processing the acquired input image, segmenting the image to obtain the required region of interest (leaf in this case), applying feature extraction methods and inputting the feature vector thus obtained into the classifier which outputs the final species name. The features acquired from a leaf image are the shape, color, texture and veins. Since 2018, the trend is more towards deep learning for image

processing. Deep Learning involves the use of Convolutional Neural Networks (CNN) which consists of layers that learnt the features of the input image followed by softmax classification layer that outputs the name of the species.

As deep learning involves a huge number of resources like GPUs and high processing powers, the most commonly used concept to apply the deep learning for smaller datasets is the Transfer Learning [2]. Transfer learning aims to apply the knowledge learnt by the network in its previous image identification task to learn the features of the new set of input images. These deep learning models have been trained on the ImageNet database which consists of 1000 classes of images, in high end environments. In this paper, we have proposed a hybrid approach that uses transfer learning to extract the features of the leaf image. The models used are VGG16[3], VGG19, ResNet[4] and MobileNetV2[5]. For the better feature selection PCA was applied and the feature vector thus obtained is input into the traditional machine learning classifiers namely Support Vector Machines, Random Forest, Decision Tree and K-nearest Neighbors to output the species of the input leaf image. The overall working of the system is depicted in Fig. 2. Section II presents the related work. Section III shows the Materials and Methods. Section IV presents the System design. Section V shows experiments and results and Section VI concludes the paper.

II. RELATED WORK

This section depicts the contribution of authors in the field of computer vision in identification of plant species using leaf images. Traditional techniques used in the identification of Ayurvedic medicinal plants were reported in [6] where the researchers extracted the features of leaf image using Speeded Up Robust Features(SURF) and the Histogram of Oriented Gradients(HOG). The feature vector obtained was input into the KNN classifier to get the final output. The algorithm was tested on the plants acquired from the Western ghats of India. An average accuracy of 99.9 was reported. A comparison among different techniques was presented in [7]. The authors studied three texture extracting techniques – HOG, Local Binary Pattern (LBP), and SURF with SVM classifier. The experiments were conducted on their own created dataset and freely available dataset Flavia and it was observed that HOG and LBP produce results similar to each other and were better than SURF in texture feature extraction. For identification of fragmented leaf images, authors devised a technique based on the bag of features and combination of edge textured histogram

with fuzzy color [8]. Multi-Layer perceptron was used to categorize the classes. The variations of scale and orientation angle were tested. High identification rate was reported when fragmentation was low due to more information in the image. A Convolutional neural Network based system called D-Leaf was proposed in [9]. After pre-processing the leaf image features were extracted using three different CNN models pretrained AlexNet, fine-tuned AlexNet and D-Leaf. For the classification different set of classifiers were utilized and compared. the classifiers included Support vector machines(SVM), Artificial neural network, k nearest Neighbors, Naive Bayes and CNN. For studying the results, benchmarked results were obtained using the Sobel operator for vein segmentation. The proposed D-Leaf model could achieve a testing accuracy comparable to the accuracy obtained on AlexNet and Fine tuned AlexNet. It was also observed that the CNN models performed better than traditional technique. The features extracted from the CNN also were classified better with ANN.

A deep learning model CNN was proposed in [10] in which the authors created their own leaf image dataset containing 4 classes with total of 8000 images. They achieved an accuracy of 85% accuracy on images taken in fields. To identify leaves in natural environments a 50-layers deep network was proposed in [11]. The experiments were based on 185 classes of leaves taken from Columbia University.

Fine-tuned GoogleNet CNN model was used in [12] to identify leaves based on the leaf venation. GoogleNet and VGG16 models were trained and tested for three different leaf datasets namely DLeaf, Flavia and Leafl datasets using CNN and the SVM. It was observed that the fine-tuned GoogleNet outperformed fine-tuned VGG16 and Dleaf CNN.

Transfer learning technique was adopted in [13] in which deep learning model was used to learn the feature representation and logistic regression was used in the classification process. When tested on the Leaf Snap and Flavia dataset which has 184 classes and 32 classes respectively, an accuracy of 90.6% and 99.54% was obtained. In another contribution shown in [14] it was observed that GoogleNet, AlexNet and VGGNet were fine-tuned and trained on the LifeCLEF 2015 dataset. Various parameters were adjusted to get high performance. The model was able to achieve a comparable accuracy as obtained in the LifeCLEF 2015 plant identification competition. While analyzing different transfer learning techniques, authors in [15] experimented with end to end CNNs, Fine-tuning, Cross dataset Fine tuning, Deep feature Learning and CNN based feature extraction followed by RNN classification on 4 freely available datasets Flavia, Swedish, UCI leaf and PlantVillage.

Another technique adopted in identification of plants in complex environments was based on Faster RCNN[16]. The convolutional layers were replaced by the InceptionV2 and Batch Normalization which in turn could provide multiscale images to the region proposal network. The proposed technique outperformed the Faster RCNN in identification of medicinal plants.

III. MATERIALS AND METHODS

The experiments were performed on 11th Gen Intel Core i7 @3.30 GHz equipped with NVIDIA GeForce RTX 3070 GPU GDDR6@8GB.

A. Dataset

The medicinal Leaf dataset has been taken from the Mendeley which consists of 30 species of healthy medicinal leaves. Each species consists of 60 to 100 images. Few samples are shown in Fig. 1. The leave images were obtained from the local gardens and the images were acquired from Mobile Camera and printer.

B. Convolutional Neural Network Models used for feature Extraction

1) *VGG16 and VGG19*: VGG stands for Visual Geometry Group was the first runner up in the ImageNet large Scale Visual Recognition Challenge 2014 which invites and evaluates techniques for object detection and image classification at large scale. It motivates researchers to come up with new techniques in the field of machine learning. VGG 16 and VGG 19 represent the 16 and 19 layers in the network respectively. The use of 3x3 filters reduces the number of parameters and lower number of parameters means that loaded model takes less space in memory. The architecture of Convolution layers followed by softmax layer.

2) *ResNet*: In Residual Networks the layers receive the input from the previous layer and also from the residual units. The network architecture contains 34 layers beginning with maxpooling and ending with average pooling. The last fully connected layer has 1000 neurons to classify objects into 1000 classes. The fourth layer receives output from 3rd and 1st layer. This arrangement helps to reduce the training errors.

3) *MobileNetV2*: This architecture was developed with vision to replace standard convolutions with depth-wise separable convolutions which helps to improve model performance by having less parameters and low computational complexity. This architecture mostly finds applications when incorporated in mobile devices. The idea behind depth-wise separable convolution was decompose the standard convolution into depthwise and pointwise convolution. This decomposition helps to learn channel and spatial features separately.

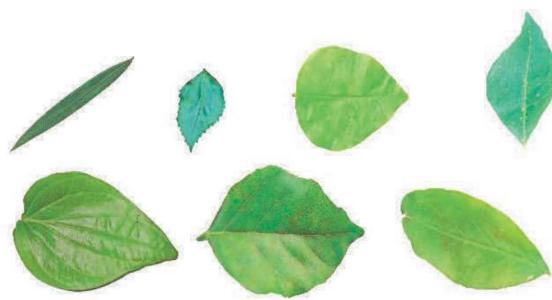


Fig. 1 Sample Images from the Medicinal leaf dataset (From left Alpinia galanga(Rasna), Hibiscus Rosa-sinensis, Jasminum, Murraya Koenigii(Curry), Piper Betle(Betel), Pongamia Pinnata(Indian Beech), Punica Granatum (Pomegranate))

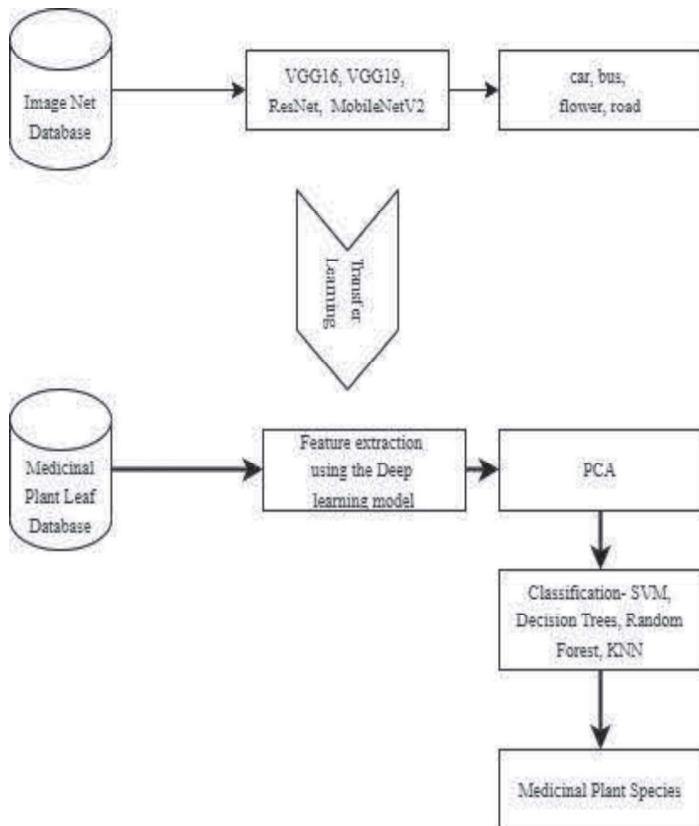


Fig. 2 Working of the proposed medicinal leaf Identification process

C. Dimensionality Reduction using Principal Component Analysis

Principal Component Analysis is a statistical technique to reduce the dimensions of data. The main purpose of incorporating PCA is to reduce the dimensionality of the data so as to simplify the structure of the feature space.

D. Classifiers used for classification of leaves into respective species

Support Vector Machines (SVM), Random Forest, Decision Trees and K-nearest Neighbors are the popular classification algorithms. They have been chosen based on their diversity in methods used to obtain the final classification output. In all the classifiers, grid search was used to obtain the best parameters for each algorithm. In supervised learning, the objective of the SVM is to find a hyperplane that can efficiently classify the datapoints in N-dimensional feature space. The Kernel takes the input from low dimensional space and transforms it into high dimensional space so that non-separable problem can be converted into separable problem. Support vector classifier particularly implements “one-versus-one” approach in multi-class classification. The parameter C in SVM is the penalty term that defines the penalty in case of misclassifications.

Decision Tree infers simple decision rules from the input samples provided and builds a hierarchical partitioned

features space. However, Due to their overfitting ability they can lead to inaccurate generalization. The parameters considered were criterion as entropy and max depth.

Random Forests is an ensemble learning technique which uses several decision trees or various subsets of the input data. A random choice of feature subspace can also be provided for each classifier. Final classification of species is made by the averaging technique. Thus, reliability and accuracy can be improved. Different parameters considered were max depth, bootstrap, max features, minimum samples per leaf, minimum samples split, and number of estimators were considered.

K-nearest Neighbors is the simplest of all classifiers as it compares the input image to the closest samples in training feature space. This approach needs only the feature information and the training data of input images. The grouping of images is done by the KNN label using majority of votes technique.

IV. SYSTEM DESIGN

The leaf image identification process as shown in Fig 2 is followed. The hybrid approach here refers to the combination of deep learning models for feature extraction and machine learning classifiers for classification. The concept of transfer learning has been used which has produced worthy results in image processing. As it eliminates the need of high-end resources like high processing speeds and GPU support, transfer learning allows researchers to carry out experiments even with a smaller number of resources. It can be applied in image segmentation, classification and feature extraction. In this case, we have used the weights specified while creation of these models. These models have the last dense layer which contains softmax for image classification into 1000 classes. As in this case only the feature extraction layers are needed, hence the models are loaded without their final layers.

The flow of the process is as follows.

- a) Load the Deep learning Model (VGG16/VGG19/ResNet/MobileNetV2) without the top layer.
- b) Load the medicinal leaf dataset. Resize to 224x224x3.
- c) Divide dataset into training (1468 images) and testing (367 images).
- d) Feature vector obtained by using transfer learning. 25088 features in VGG16 and VGG19, 100352 for ResNet50, 62720 for MobileNetV2.
- e) Input feature vector into PCA.
- f) Select best hyperparameters for the classifiers using grid search. Classifiers SVM, Decision trees, Random Forest and KNN.
- g) For the best hyperparameters selected, perform 3-Fold cross validation to get the scores.
- h) Calculate mean accuracy and mean standard deviation across all folds.(Level -1 validation)
- i) The accuracy value of the model is also checked for the held-out dataset (test set)(Level -2 validation)

V. EXPERIMENTS AND RESULTS

As shown in Table 1, for each CNN model, 4 different classifiers were tested and the results are reported. As, the validation is done on 2 levels, the Mean Accuracy represents the mean value of the accuracies obtained using 3-Fold cross validation and the Mean Std represents the mean of the standard deviations obtained in the 3-Folds. The “accuracy on the test set” shows the accuracy obtained on the unseen samples of the dataset. It can be concluded that the combination of MobileNetV2 and KNN outperformed other combinations of feature extractor models and classifiers. It was also observed that from all the classifiers listed, KNN gives the maximum accuracy in all the feature extracting models.

VI. CONCLUSION

An automatic identification approach for the identification of medicinal plants is presented. The medicinal leaf database consists of 30 species of medicinal plant leaves with a total of 1500 images. The dataset was divided into training and test sets. Transfer learning has been applied to extract the features of the leaf images using deep learning CNN models – VGG16, VGG19, ResNet and MobileNetV2. The feature vector obtained from the models was input into the PCA as PCA helps in better feature selection. The feature vector was then input into the SVM, Decision Trees, random Forest and KNN classifiers and the results were reported. For the validation, 3-fold cross validation on the training set and hold out validation was used. It can be observed from Fig. 3 that out of all the feature extractors, MobileNet V2 performed better followed by VGG16, VGG19 and ResNet. While comparing the classifiers, it was observed that KNN performed better of all, followed by Random Forest, SVM and lastly decision trees. An accuracy of 96.18% and standard deviation of 0.0059 was obtained using the MobileNetV2 feature extractor and KNN as classifier which shows that the combination of light weight model and simple classifier can make leaf identification task easy.

Table 1. Experimental Outcome: Mean Accuracy and Mean Std (standard deviation) and Testing Accuracy

Model	Classifier	Mean-Accuracy (3Fold)(%)	Mean Std	Accuracy on the test set(%)
vgg16	SVM	72.68	0.008	85.03
	Random Forest	80.92	0.013	85.01
	Decision Tree	58.31	0.029	63.76
	KNN	93.34	0.007	95.36
vgg19	SVM	76.22	0.03	82.01
	Random Forest	81.19	0.01	83.65
	Decision Tree	60.08	0.0055	64.03
	KNN	92.98	0.012	94.82
Resnet	SVM	50.34	0.0127	65.94
	Random Forest	56.74	0.0084	63.48
	Decision Tree	51.72	0.018	56.67
	KNN	76.63	0.02	83.1
MobileNet V2	SVM	68.86	0.0335	74.38
	Random Forest	85.08	0.001	88.55
	Decision Tree	63.07	0.022	60.49
	KNN	95.91	0.0059	96.18

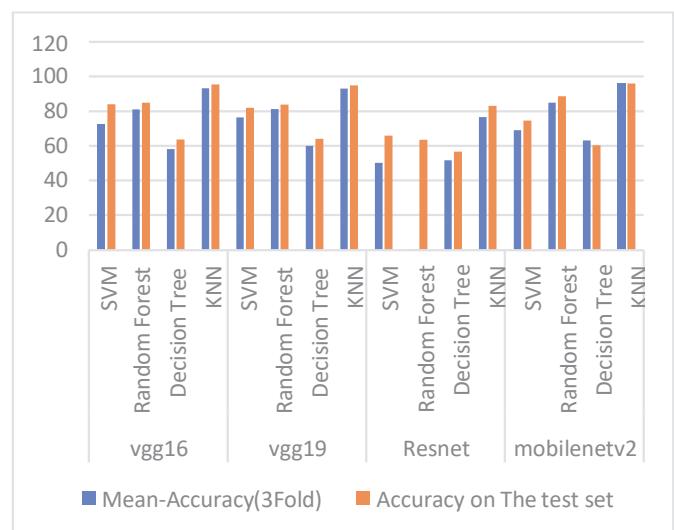


Fig. 3 Comparison of results

VI. REFERENCES

- [1] L. M. de Freitas Junior and E. B. de Almeida, “Medicinal plants for the treatment of obesity: Ethnopharmacological approach and chemical and biological studies,” *Am. J. Transl. Res.*, vol. 9, no. 5, pp. 2050–2064, 2017.
- [2] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.
- [3] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016–December, doi: 10.1109/CVPR.2016.90.
- [5] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” 2018, doi: 10.1109/CVPR.2018.00474.
- [6] A. Sabu, K. Sreekumar, and R. R. Nair, “Recognition of ayurvedic medicinal plants from leaves: A computer vision approach,” *2017 4th Int. Conf. Image Inf. Process. ICIIP 2017*, vol. 2018-Janua, pp. 574–578, 2018, doi: 10.1109/ICIIP.2017.8313782.
- [7] Z. Ibrahim, N. Sabri, and N. N. A. Mangshor, “Leaf recognition using texture features for herbal plant identification,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 9, no. 1, pp. 152–156, 2018, doi: 10.11591/ijecs.v9.i1.pp152-156.
- [8] J. Chaki, N. Dey, L. Moraru, and F. Shi, “Fragmented plant leaf recognition: Bag-of-features, fuzzy-color and edge-texture histogram descriptors with multi-layer perceptron,” *Optik (Stuttg.)*, vol. 181, no. December 2018, pp. 639–650, 2019, doi: 10.1016/j.ijleo.2018.12.107.
- [9] J. W. Tan, S. W. Chang, S. Abdul-Kareem, H. J. Yap, and K. T. Yong, “Deep Learning for Plant Species Classification Using Leaf Vein Morphometric,” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 17, no. 1, pp. 82–90, 2020, doi: 10.1109/TCBB.2018.2848653.
- [10] C. Amuthalingeswaran, M. Sivakumar, P. Renuga, S. Alexpandi, J.

Elamathi, and S. S. Hari, "Identification of medicinal plant's and their usage by using deep learning," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, no. Icoei, pp. 886–890, 2019, doi: 10.1109/ICOEI.2019.8862765.

- [11] V. Bodhwani, D. P. Acharjya, and U. Bodhwani, "Deep residual networks for plant identification," *Procedia Comput. Sci.*, vol. 152, pp. 186–194, 2019, doi: 10.1016/j.procs.2019.05.042.
- [12] P. Jasitha, M. R. Dilleep, and M. Dilvya, "Venation Based Plant Leaves Classification Using GoogLeNet and VGG," *2019 4th IEEE Int. Conf. Recent Trends Electron. Information, Commun. Technol. RTEICT 2019 - Proc.*, pp. 715–719, 2019, doi: 10.1109/RTEICT46194.2019.9016966.
- [13] A. Beikmohammadi and K. Faez, "Leaf Classification for Plant Recognition with Deep Transfer Learning," *Proc. - 2018 4th Iran. Conf. Signal Process. Intell. Syst. ICSPIS 2018*, pp. 21–26, 2018, doi: 10.1109/ICSPIS.2018.8700547.
- [14] M. Mehdipour Ghazi, B. Yanikoglu, and E. Aptoula, "Plant identification using deep neural networks via optimization of transfer learning parameters," *Neurocomputing*, vol. 235, no. April 2016, pp. 228–235, 2017, doi: 10.1016/j.neucom.2017.01.018.
- [15] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Comput. Electron. Agric.*, vol. 158, no. October 2018, pp. 20–29, 2019, doi: 10.1016/j.compag.2019.01.041.
- [16] X. Zhu, M. Zhu, and H. Ren, "Method of plant leaf recognition based on improved deep convolutional neural network," *Cogn. Syst. Res.*, vol. 52, pp. 223–233, 2018, doi: 10.1016/j.cogsys.2018.06.008.

A Study of Cloud Computing

Muskan Kumari

*Department of Computer Science (4th Semester),
Baba Banda Singh Bahadur Engineering College, Fatehgarh Sahib, Punjab-140407, India.*

ABSTRACT: Cloud Computing has come of agedness Amazons introduce the first of its kind of cloud services in 2006. It is particularly suitable to Hong Kong of the unbelievable amounts of the data are being processed here daily in several sectors, and there are signs that subscription to cloud services by the local companies will soon be on a skyrocket course, despite a slow start in beginning years. As a research theme, cloud computing now easily tops any schedule of topics in a computer science because of its far-reaching suggestion in many sectorin computing, especially a big data which without cloud computing is at the great concept. The current creation of a main cloud R&D centre in Hong Kong by Lenovo (January 2015) certify to this fact.

Cloud computing has become the most enticing field of computing as proposed by Google in 2007. Cloud computing provides the computing infrastructure, storage and any type of applications from audio to word processors what even we can imagine. Cloud computing is capable of assigning the required resources to any client on demand. The clients do require neither any cost for computing infrastructure building nor maintenance cost of any hardware or software that removes the huge financial burden and mental stress on them. The clients only need to connect their computers or networks to the cloud computing servers. In this paper we focused on the introduction and need of cloud computing, Evolution of cloud computing, types of cloud, services provided by clouds and drawbacks of cloud computing.

Keywords: Cloud, Public Cloud, Private Cloud, Hybrid Cloud, Community Cloud, IaaS, PaaS, SaaS, ARPSNET (Advanced Research Projects Agency Network)

INTRODUCTION: Joseph Carl Robnett Licklider in the 1960s developed Cloud Computing with his work on ARPSNET to interconnect with people and data from in any place at any time. In 1983, CompuServe presented its users as a little amount of disk space that could be used to assemble any files they choose to upload.

Every organization either business or education needs IT infrastructure (hardware and software) for computing and storage for carrying out their regular work. An organization may need different hardware or software for different types of applications which means new software must have to buy each time when there may be changes in applications. Buying only the computers for the employees in an organization is not enough as the new applications may need new software which may be of very high cost, or the software license may not permit more than one user to install. Besides, the most important factor is the maintenance

of both the hardware as well as software is always very high as compared to the installation cost of the same. So there is a lot of load in terms of cost and stress on the IT personal responsible for managing the IT infrastructure in an organization.

Cloud computing or cloud technology is an innovation that redefined the way data and information are stored, retrieved, and processed. The use of cloud technology is the use of cloud storage. A cloud is a storage space that is virtually on the interest (Dynamix Solution, 2017). Cloud has become jargon, but a cloud is just a server located in the data center of vendors/providers which is marked and sold to clients.

Thus, the concept of cloud computing is emerged to overcome the above problems. On demand delivery of computer resources - everything like computing, storage and software etc. is known simply as cloud computing. It is called cloud computing as data and applications exist on a cloud of web server. If an organization's computers are connected to the cloud computing server, the organization does not need to install a suite of software for each computer rather to load one application which would allow clients to log into a web-based server which hosts all the software that is enough for the client's need.

EVOLUTION OF CLOUD COMPUTING:

One day in a speech at MIT around in 1960 Jhon McCarthy indicated that like water and electricity, computing can also be solid like a utility. And in 1999,

the Salesforce Company started distributing the applications to the customers through a convenient and computation. In around 2009 big companies like Google, Microsoft, HP, Oracle had started to provide cloud computing services. Now a days each and every person is using the services of cloud computing in their daily life. For example Google Drive, and iCloud etc. In future cloud computing will become the basic need of ITIndustries.

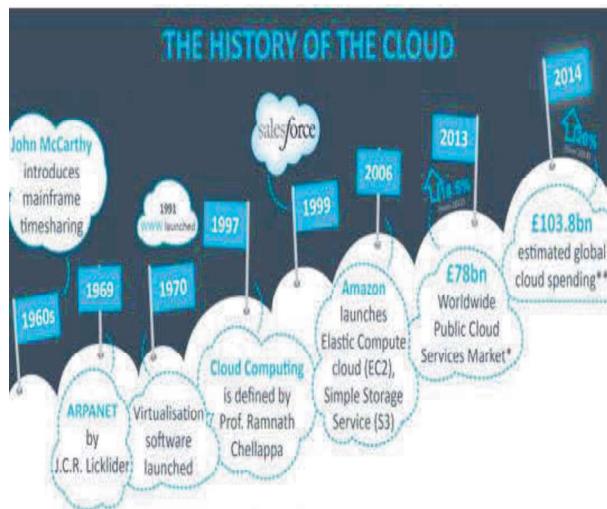


Fig: Evolution of Cloud Computing

CLASSIFICATION OF CLOUD COMPUTING:

Cloud computing is usually classified on the basis of location of cloud in four ways-

- 1) **Public Cloud:** Public cloud is owned by the cloud service provider companies. The clients have no authorization to control the cloud rather they just enjoy the services provided by the cloud although it seems to the clients that everything is running evenly at their own computer due to the virtualization of the cloud infrastructure. Users don't need to buy hardware, software or supporting infrastructure, which is owned and managed by providers with public cloud services. The cloud computing in general is shared by several organizations. Google, Amazon, Yahoo, SUN etc. are the example of public serviceprovider.
- 2) **Private Cloud:** This cloud distribution model is a modified infrastructure maintained by a single business. It offers a precise environment in which contact to IT resources is additionally centralized within the business. The present exemplary perhaps visibly introduced obtainable handled internal. Even though secluded cloud introducing obtainable valuable,

as largest productions it could be action a developed equal of safety and extra self-sufficiency to modify the storing, interacting and calculate mechanisms toward ensemble their ITnecessities.

- 3)
- 4) **3) Hybrid Cloud:** Hybrid cloud is a combinationof public and private clouds. **Hybrid cloud = public cloud + private cloud.** The main aim to merge this cloud (Public and Private) is to create a unified, automated, and well-managed computing environment. In the Hybrid cloud, non-critical activities are performed by the public cloud and critical activities are performed by the private cloud. Mainly, a hybrid cloud is used in finance, healthcare, and Universities The best hybrid cloud provider companies are Amazon, Microsoft, Google, Cisco, andNetApp.
- 5)
- 6) **4) Community Cloud:** Community Cloud is a hybrid form of private cloud. They are cross functional platforms that enable different organizations to work on a shared platform. Community Cloud may be hosted in a data center, owned by one of the occupants, or by a third-party cloud services provider and can be either on-site or off-site. In community cloud, computing infrastructure is being shared by the organizations belonging to the same community. For example all Government organizations may share the same cloud to fulfill the requirements of astate.

CLASSIFICATION BASED ON SERVICE PROVIDED BY CLOUDS:

- 1) **Infrastructure as a service (IaaS):** Infrastructure as a service (IaaS) is a type of cloud computing service that offers crucial compute, storage and networking resources on demand, on a pay-as-you-go basis. These resources are dynamic in nature and they can be break, assigned and resized to make available to the clients as their requirements. Vendors like Amazon Elastic Compute Cloud, Amazon S3, IBM'sBlue Cloud, Rackspace Cloud Servers and Flexiscale provides this type services to the clients.
- 2) **Platform as a service (PaaS):** Platform as a service offers a cloud-based environment that supports the complete life process of building and delivering web-based applications without buying and managing the essential hardware, software and

hosting cost. The platform enables the organization to develop, run, and manage business applications without the need to build and maintain the infrastructure such software development processes require.

- 3) Software as a service (SaaS):** It provides the software as a service on the cloud. Clients may access any software applications as they need or may run their applications or programs in the cloud via internet that may be hosted and maintained. SaaS is also known as "**On-Demand Software**". It is a software distribution model in which services are hosted by a cloud service provider. These services are available to end-users over the internet so, the end-users do not need to install any software on their devices to access these services.

SaaS examples: BigCommerce, Google Apps, Salesforce, Dropbox, MailChimp, ZenDesk, DocuSign, Slack, Hubspot.

DRAWBACKS OF CLOUD COMPUTING:

Cloud security and data

Most cloud service providers implement relevant security standards and industry certifications to ensure that their cloud environment remains safe. However, storing data and business-critical files in virtual data centers can potentially open you up to risks.

Common risks are:

- data loss or theft
- data leakage
- account or service hijacking
- insecure interfaces and APIs
- denial of service attacks
- technology vulnerabilities, especially on shared environments

The levels of **data protection** and security achieved and maintained by different cloud providers can vary. Choose your provider carefully and make sure that the provider is stable, reliable, reputable and offers reasonable terms and conditions of service.

Cloud downtime

The cloud, like any other IT set-up, can experience technical problems such as reboots, network outages and downtime. These events can incapacitate business operations and processes and can be damaging to business.

You should plan for **cloud downtime and business continuity**. Try to minimize the impact and the number of outages and ensure the maximum level of service availability for your customers and staff.

Limited control

The cloud service provider owns, manages and monitors the cloud infrastructure. You, as the customer, will have minimal control over it. You will be able to manage the applications, data and services operated on the cloud, but you won't normally have access to key administrative tasks, such as updating and managing firmware or accessing server shell.

In order to reduce risks, it may help to carry out a risk assessment before you hand over any control to a service provider. Once you have a clear idea of the risks, you will be able to think about them against the **advantages of cloud computing**.

CONCLUSION:

Cloud computing marks the beginning of a new stage in the area of data and communication technology as it carries with a development pattern which has the possibility to change the way in which computing was done. Users are still getting aware through this skill and a change from conformist subtracting to cloud computing will ensue but progressively. Owed to this technology, developers with novel ideas about internet services will no longer need to spend large amounts of currency in structuring their programs and tools substructure abilities. It has been noticed that the offerings of internet services on demand has been highly increased today and the trend will be more and more in future. On demand delivery of computer resources – everything like computing, storage and software etc. is known simply as cloud computing.

The clients do not require neither any cost for computing infrastructure building nor maintenance cost of any hardware or software that removes the huge financial burden and mental stress on them. The clients only need to connect their computers or networks to the cloud computing servers and to pay service charges by simple payment methods like pay-per-use or subscription method. In this paper we focused on the need of cloud computing, architecture of cloud

computing, types of cloud, services provided by clouds and the benefits or offerings of cloudcomputing.

REFERENCES:

- [1] Mrs. Ashwini Sheth Research Paper on Cloud Computing.pdf
- [2] 711-1636-1-PB.pdf
- [3] 15.MYJCS10099.pdf
- [4] Cloud_Computing_A_Detailed_Relationship_to_Grid_a_n.pdf
- [5] Venters, W., Whitley, E.A.: A Critical Review of Cloud Computing: Researching Desires and Realities. *J. Inf. Technol.* 27, 179–197 (2012).
- [6] Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., Ghalsasi, A.: Cloud computing — The Business Perspective. *Decis. Support Syst.* 51, 176–189 (2011).
- [7] Weinhardt et al. Cloud Computing-A Classification, Business Models, and Research Directions. *Business and Information Systems Engineering*, 2009.
- [8] community.spiceworks.com/cloud/articles/2504
- [9] Garrison, G., Kim, S., Wakefield, R.L.: Success Factors for Deploying Cloud Computing. *Commun. ACM.* 55, 62–68 (2012).
- [10] Venters, W., Whitley, E.A.: A Critical Review of Cloud Computing: Researching Desires and Realities. *J. Inf. Technol.* 27, 179–197 (2012).
- [11] N. Sadashiv and S. D. Kumar, "Cluster, grid and cloud computing: A detailed comparison," in Proc. the 2011 6th International Conference on Computer Science and Education (ICCSE), 2011, pp. 477-482.
- [12] Welcome to TCloud. (2014). [Online]. Available: <http://www.Tclouds-project.eu/>

Parkinson Disease detection using Machine Learning

Harnoor Kaur Khehra

*Department of Computer Science & Engineering
Baba Banda Singh Bahadur Engineering College
Fatehgarh Sahib, Punjab - 140407, India*

harnoorkhehra28@gmail.com

Abstract - Parkinson's disease is caused by the disordering of the brain cells that produce substances to allow brain cells to communicate with each other, called dopamine. The cells that produce dopamine in the brain are responsible for the control, adaptation and fluency of movements. When some percentage of these cells are lost, then enough dopamine is not produced and Parkinson's motor symptoms appear. It is thought that the disease begins many centuries before the motor symptoms and therefore, researchers are looking for ways to recognize the nonmotor symptoms that appear early in the disease as early as possible, thereby stalling the progression of the disease [7]. By using machine learning techniques, the problem can be solved with minimal error rate. Machine learning also allows for combining different modalities, such as magnetic resonance imaging (MRI) and single-photon emission computed tomography (SPECT) data, in the diagnosis of PD. It is possible to detect Parkinson's disease using the drawings alone. In this paper we have to quantify the visual appearance (using HOG method) of these drawings and then train a machine learning model to classify them and using Histogram of Oriented Gradients (HOG) image descriptor along with a Random Forest classifier to automatically detect Parkinson's disease in hand-drawn images of spirals and waves.

Index Terms - Machine Learning, Parkinson Disease, Histogram of Oriented Gradients, Random Forest classifier.

I. INTRODUCTION

Parkinson's disease is a chronic neurodegenerative brain disease that affects the ability of those who are engaged in regular activities. More than 10 million people are living with Parkinson's Disease worldwide, according to the Parkinson's Foundation, which is considered to be the second most common neurodegenerative disorder next to Alzheimer's disease [1]. The disease can be affected by several factors, including the social environment, drug use, and complications of neurological conditions. The number one signs may be categorized into motor signs and non-motor signs. Motor signs take place as tremor, bradykinesia, tension muscles, and extraordinary gait, while non-motor signs display intellectual disorders, sleep problems, autonomic dysfunction, and sensory disturbance [2]. It has been identified that the main cause of PD is a lack of a chemical messenger called dopamine in the specific area of the brain

called the substantia nigra, which is a neurotransmitter that not only controls motoractivity but also regulates muscle function [1]. Figure 1 gives a clear picture that Parkinson's disease is primarily associated with the gradual loss of cells in the substantia nigra of the brain. This area is responsible for the production of dopamine. Dopamine is a chemical messenger that transmits signals between two regions of the brain to coordinate activity. For example, it connects the substantia nigra and the corpus striatum to regulate muscle activity [3].

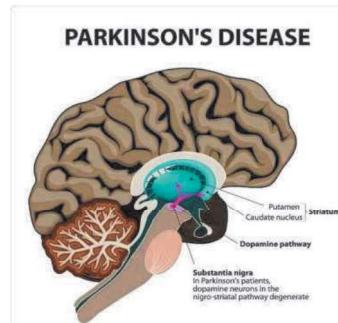


Fig 1: PD pathophysiology

Loss of dopaminergic neurons from the basal ganglia-leading to biochemical abnormalities at low dopamine levels and Lewy bodies are a pathological feature of PD as shown in figure 2.

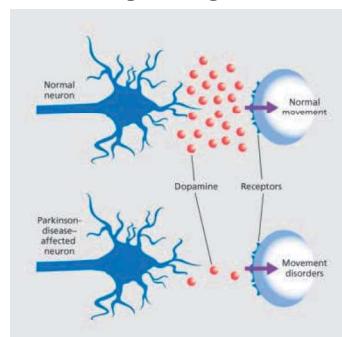


Figure 2: Dopamine levels in a normal and Parkinson disease-affected neuron [4]

Over the last decade, researchers have associated mutations in certain genes with familial Parkinson's disease. It accounts for nearly 15% of all cases of Parkinson's disease worldwide. Familial PD is also known as early-onset PD because the age

of onset is less than 40 years. Recessive familial PD is associated with parkin, DJ1, and PINK1 mutations, while alpha-synuclein and LRRK2 mutations are associated with dominant familial PD [4]. Furthermore, During its development, exercise-related symptoms such as tremor, stiffness, and early difficulty can be observed before cognitive and behavioral changes, including dementia, appear. PD severely affects patients' quality of life (QoL), social functions and family relationships, and places heavy economic burdens at individual and society levels. Machine learning techniques are increasingly being used in healthcare. As the name implies, machine learning allows computer programs to

semi-automatically learn and extract meaningful expressions from data [5]. Parkinson's disease cannot be cured, but early detection and proper dosing can significantly improve symptoms and quality of life. Researchers have found that patients with Parkinson's disease have slower drawing speeds and lower pen pressure. One of the signs of Parkinson's disease is muscle tremors and stiffness, making it difficult to draw smooth spirals and waves. Instead of measuring the speed and pressure of the pen on paper, it is possible to diagnose Parkinson's disease from the drawings alone. Our goal is to quantify the appearance of these drawings (using the HOG method) and train and classify machine learning models. This paper uses the Histogram of Oriented Gradients (HOG) image descriptor and random forest classifier to automatically detect Parkinson's disease in hand-drawn images of spirals and waves. The figure 3 indicates the algorithm for Parkinson Disease diagnosis.

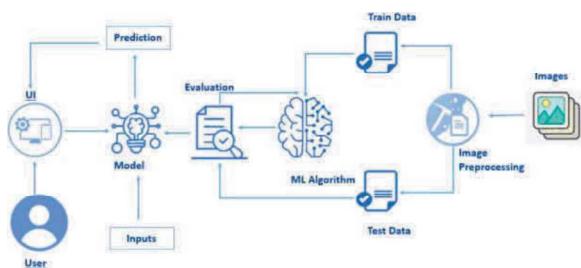


Figure 3: Algorithm for PD Diagnosis[6]

The main contributions of the proposed ML-based early PD diagnostic method are listed below [7].

- 1) Recursive feature removal (RFE) and feature importance (FI) methods were used to determine the most relevant features to use in a classification task.
- 2) The significance of the use of FS strategies within the preprocessing segment of class of PD patients become proven. The overall performance of

the SVM classifier has progressed approximately 13% through FS.

- 3) The fewest voice functions in the literature are used for PD diagnosis and are very large. Recognition accuracy (93.84%) was achieved with little effort.
- 4) The performance of various popular classifiers was evaluated and the best classifier was identified and detected as an SVM for PD diagnostic problems.
- 5) Various optimizers were evaluated against the dataset to determine the best optimizer.
- 6) The proposed technique is higher than the opposite techniques with recognition to computational fee since few varieties of voice functions have been used rather than heavy function extraction tactics such as MRI, movement sensors or handwriting assessments.

A. OBJECTIVES

The objectives of the paper are as follows

- 1) To understand the problem to classify if it is a regression or a classification kind of problem.
- 2) To know how to pre-process the image by using different data pre-processing techniques.
- 3) To use OpenCV and machine learning to automatically detect Parkinson's disease in hand-drawn images of spirals and waves.
- 4) To know how to find the accuracy of the model.
- 5) To Build web applications using the Flask framework.

II. PAPER FLOW AND THEORETICAL ANALYSIS

In the paper flow, the user interacts with the UI (User Interface) to upload the image as input then the uploaded image is analyzed by the model which is integrated. Moreover, it is considered that once the model analyzes the uploaded image, the prediction is showcased on the UI and OpenCV window.

To accomplish this, we have to complete all the activities and tasks listed below:

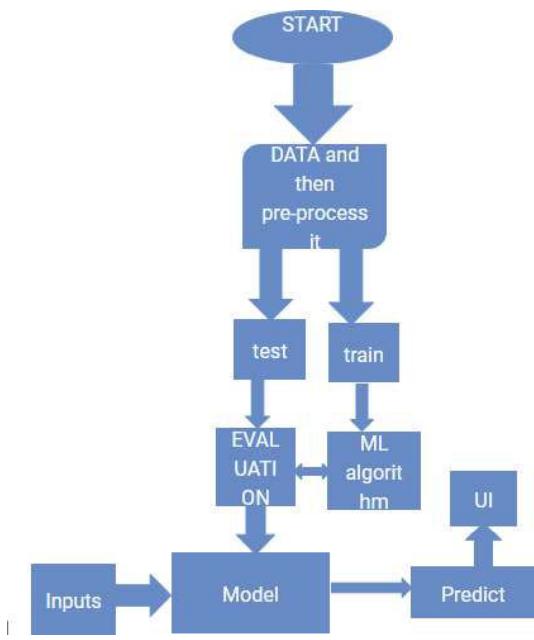


Figure 4: Flow diagram of process

A. Data collection

ML is highly data-dependent, and without it it is impossible for machines to learn. This is the most important aspect that enables algorithm training. Machine learning projects require training datasets. The actual dataset is used to train the model to perform different actions. In order to use the data we collect to develop practical artificial intelligence (AI) and machine learning solutions, it must be collected and stored in a way that makes sense for the business problem at hand. Moreover, we illustrate that data sets are very important as Collecting data allows you to keep a record of past events so that you can use data analysis to find repetitive patterns. From these patterns, we use machine learning algorithms to look for trends and create predictive models that predict future changes.

In this paper, waves and spirals are used as datasets which further are divided into training and testing datasets. Then, data for the person who has Parkinson disease or is healthy[8]. The image datasets are being evaluated for healthy people and people with Parkinson's disease. The figure 5 shows the image training dataset for spiral and the person with the disease[9].

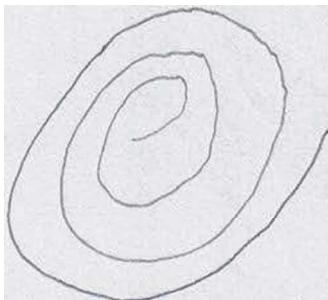
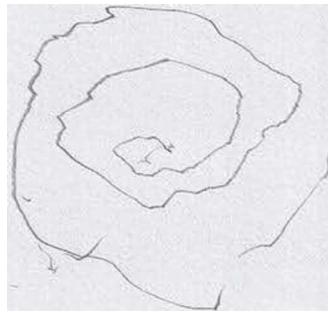


Figure 5: Spiral images of a person with PD [9]

The figure 6 shows the image training dataset for the spiral of the healthy person.

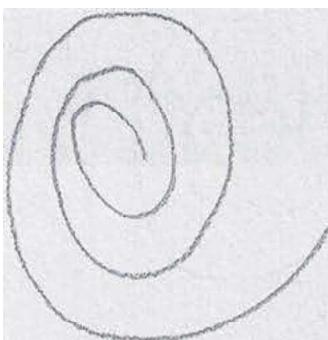
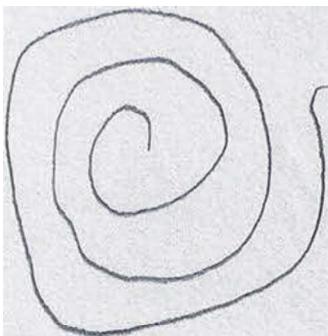


Figure 6: Spiral images of a healthy person [9]

Similarly, for the testing datasets we have different spiral images for healthy people and people with a disease.

The figure 7 shows the image training dataset for the wave of a person with Parkinson disease.

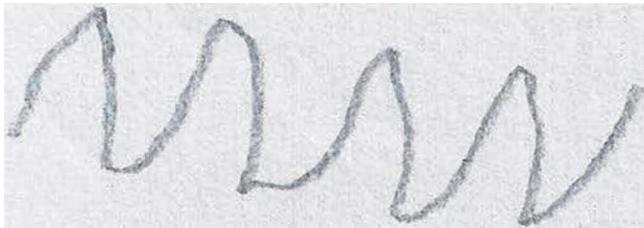
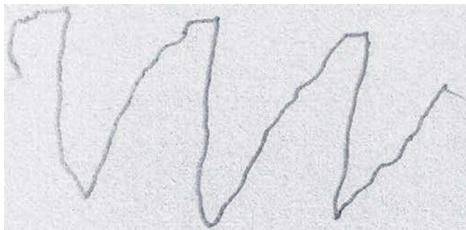


Figure 7: Wave images of a person with disease [9]

The figure 8 shows the image training dataset for the wave of a person who is healthy.

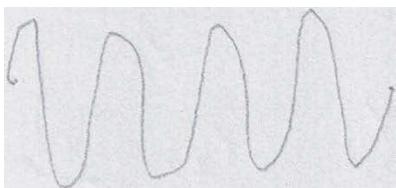
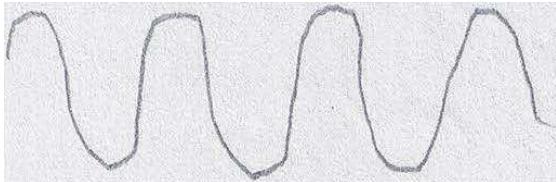


Figure 8: Wave images of a healthy person [9]

A similar trend is seen for the testing datasets. We have different wave images for healthy people and people with a disease.

B. Image pre-processing

The image data are redundant and require processing for presentations facilitating their interpretation by a physician and enabling efficient utilization of machine learning algorithms in the next study stage. Computer Vision technique is applied for this [10]. For Importing the necessary libraries, The important step for image pre-processing is to import the libraries. OpenCV is a prebuilt, open source CPUonly library (package) that is widely used for computer vision, machine learning, and image processing applications. It supports a good variety of programming languages including Python [11]. Imutils is a series of convenience functions to make basic image processing functions such as translation, rotation, resizing, skeletonization, and displaying Matplotlib images easier with OpenCV. We will build_montages for

visualization. Our paths import will help us to extract the filepaths to each of the images in our dataset [12]. sklearn. Metrics are the module that implements several loss, score, and utility functions to measure classification performance. The sklearn.preprocessing package provides some common utility functions and transformer classes for transforming raw feature vectors into better representations for downstream estimators. This package contains RandomForestClassifier and many more inbuilt algorithms [13]. Scikit-image, or skimage, is an open-source Python package designed for image preprocessing. Histogram of Oriented Gradients (HOG) will come from the feature import of scikit-image. It is a collection of algorithms for image processing. It is available free of charge and free of restriction [14]. Python pickle module is used for serializing and de-serializing python object structures. The process to convert any kind of python objects (list, dict, etc.) into byte streams.

C. Loading Train Data and Test Data

The observations in the training set form the experience that the algorithm uses to learn. In supervised learning problems, each observation consists of an observed output variable and one or more observed input variables. The test set is a set of observations used to evaluate the performance of the model using some performance metric. It is important that no observations from the training set are included in the test set. If the test set does contain examples from the training set, it will be difficult to assess whether the algorithm has learned to generalize from the training set or has simply memorized it. Our dataset contains both hand-drawn spiral and wave patterns. Here we are taking spiral patterns into consideration and training the model. We split the data into train and test. Using the training dataset we train the model and the testing dataset is used to predict the results. Same with the wave pattern dataset [15]. The load_split function accepts a dataset path and returns all feature data and associated class labels. From there we grab input image paths by making use of imutils. Then, Initialize the data and label lists, loop over all image paths which we have grabbed in the previous step and each label is extracted from the os.path.split() method in Python which is used to split the pathname into a pair head and tail. Here, the tail is the last pathname component and the head is everything leading up to that. Each image is loaded and preprocessed. Next, read the input image, convert image to grayscale, we convert the image to grayscale to reduce the processing time, resizing the image, threshold image. The thresholding step segments the drawing from the input image, making the drawing appear as a white foreground on a black background. Features are extracted via our quantify_image function. The features and labels are appended to the data and labels lists respectively. Finally, data and labels are converted to NumPy arrays and returned conveniently in a tuple.

D. Quantifying Images

Quantization is the process of determining the value of each pixel in an image. First, the basic gray level, the range of values a pixel can have, is determined. Then, according to the value from the continuous signal, the pixel value of the digital image is calculated. We will extract features from each input image with the quantify_image function. HOG is a structural descriptor that will capture and quantify changes in local gradient in the input image. HOG will naturally be able to quantify how the directions of both spirals and waves change. It will be able to capture if these drawings have more of a “shake” to them, as we expect from a Parkinson’s patient. The most important parameters for the HOG descriptor are the orientations, pixels_per_cell, and the cells_per_block. These three parameters (along with the size of the input image) effectively control the dimensionality of the resulting feature vector. The resulting features are a 12,996-dim feature vector (list of numbers) quantifying the wave or spiral. We’ll train a Random Forest classifier on top of the features from all images in the dataset[16].

E. Label Encoding

This approach is very simple and you need to convert each value in the column to a number. In this, we replace the categorical value with a numeric value between 0 and the number of classes minus 1. If the categorical variable value contains 5 distinct classes, we use (0, 1, 2, 3, and 4). Also, in this categorical variables are converted into a form that could be provided to ML algorithms to do a better job in prediction. We apply Encoding in order to convert the values into 0’s and 1’s. As we have observed that we have labels in our dataset we need to convert them into binary values by using Label encoding. 0:healthy,1:Parkinson. Create an object le and fit the y_train,y_test using fit_transform[17].

III. MODEL BUILDING

A machine learning model is built by learning and generalizing from training data, then applying that acquired knowledge to new data it has never seen before to make predictions and fulfill its purpose. Lack of data will prevent you from building the model, and access to data isn’t enough. There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms can be chosen according to the objective. As the dataset which we are using is a classification so you can use the algorithms like Random Forest Classification and Decision Tree Classification. You will need to train the datasets to run smoothly and see an incremental improvement in the prediction rate[18].

A Training the model

To train an ML model, you need to specify the Input training datasource, name of the data attribute, required data transformation instructions, training parameters to control the learning algorithm. Once after splitting the data into train and test, the data should be fed to an algorithm to build a model. There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms that you can choose according to the objective that you might have may be Classification algorithms are Regression algorithms like Logistic Regression, Decision Tree Classifier, Random Forest Classifier, KNN. Initialize our Random Forest classifier and train the model using a number of estimators as 100 [19].

A. Testing the model

To train an ML model, you need to specify the Input training datasource, name of the data attribute, required data transformation instructions, training parameters to control the learning algorithm. Once after splitting the data into train and test, the data should be fed to an algorithm to build a model. There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms that you can choose according to the objective that you might have may be Classification algorithms are Regression algorithms like Logistic Regression, Decision Tree Classifier, Random Forest Classifier, KNN. Initialize our Random Forest classifier and train the model using a number of estimators as 100 [19].

B. Testing the model

In machine learning, model testing is called the process of assessing the performance of a fully trained model in a test set. A test set consisting of a set of test examples should be separated from both the training set and the validation set, but should follow the same probability distribution as the training set. After training the model, the model should be tested by using the test data which has been separated while splitting the data for checking the functionality of the model. We randomly took images from the testing dataset. Our images list will hold each spiral image along with annotations added via OpenCV drawing functions. We proceed to loop over the random image indices. Inside the loop, each image is processed in the same manner as during training (convert to grayscale, resize, threshold). From there we’ll automatically classify the image using our new HOG + Random Forest based classifier and add color-coded annotations. Each image is quantified with HOG features. Then the image is classified bypassing those features to model.predict. We can then clearly differentiate between the geometric patterns drawn by healthy and Parkinson patients [20].

C. Model Evaluation

Evaluation is a process during the development of the model to check whether the model is the best fit for the given problem and corresponding data. Classification Evaluation Metrics used to find out the accuracy of models built in the classification type of machine learning models. We have accuracy_score, confusion matrix, roc- Auc Curve.

IV.RESULTS

In this we detected the presence of Parkinson's Disease in individuals using various factors. We used an XGBClassifier for this and made use of the sklearn library to prepare the dataset. This gives us an accuracy of 96.66%, which is great considering the number of lines of code in this.

Figure 9 describes the page on which details are given for the disease.

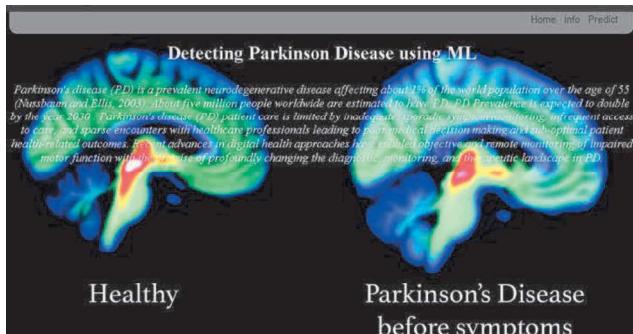


Fig 9: Introduction about PD

Figure 10 describes the second web page telling more details about the disease.

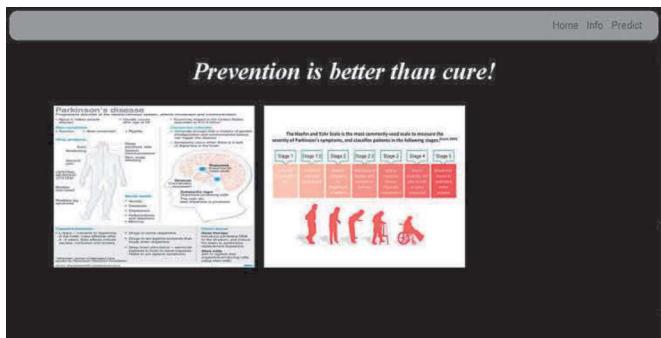


Fig 10: Detail explanation of Parkinson Disease

The figure 11 describes when the computer machine is given the image data set for the prediction to know whether the person is healthy or with disease and out comes to be healthy.



Figure 11: The computer detects through image dataset that person is healthy

IV.CONCLUSION

The researchers found that the drawing speed was slower and the pen pressure is lower among Parkinson's patients. One of the indications of Parkinson's is tremors and rigidity in the muscles, making it difficult to draw smooth spirals and waves. It is possible to detect Parkinson's disease using the drawings alone instead of measuring the speed and pressure of the pen on paper[21]. Our goal is to quantify the visual appearance(using HOG method) of these drawings and then train a machine learning model to classify them. In this project, We are using, Histogram of Oriented Gradients (HOG) image descriptor along with a Random Forest classifier to automatically detect Parkinson's disease in hand-drawn images of spirals and waves [22]. To the best of our knowledge, the present study is the first review which included results from all studies that applied machine learning methods to the diagnosis of PD[23]. Here, we presented included studies in a high-level summary, providing access to information including (a) machine learning methods that have been used in the diagnosis of PD and associated outcomes, (b) types of clinical, behavioral and biometric data that could be used for rendering more accurate diagnoses, (c) potential biomarkers for assisting clinical decision making, and (d) other highly relevant information, including databases that could be used to enlarge and enrich smaller datasets. In summary, realization of machine learning-assisted diagnosis of PD yields high potential for a more systematic clinical decision-making system, while adaptation of novel biomarkers may give rise to easier access to PD diagnosis at an earlier stage. Machine learning approaches therefore have the potential to provide clinicians with additional tools to screen, detect or diagnose PD [24].

REFERENCE

S

- [1] Balaji E., Brindha D., Balakrishnan R. "Supervised Machine Learning based Gait Classification System for Early Detection and Stage Classification of Parkinson's Disease", 10.1016/j.asoc.2020.106494, 16 June 2020.
- [2] Bo Peng, Suhong Wang, Zhiyong Zhou, Yan Liu, Baotong Tong, Tao Zhang, Yakang Dai, "A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinson's disease" 3 rd ed., 18-4-2017.
- [3] <https://www.news-medical.net/health/Parkinsons-Disease-Pathophysiology.aspx>
- [4] <https://europepmc.org/article/pmc/5442366>
- [5] <https://www.frontiersin.org/articles/10.3389/fnagi.2021.633752/full>
- [6] Timothy J. Wroge¹, Yasin Ozkanca², Cenk Demiroglu², Dong Si³, David C. Atkins⁴ and Reza Hosseini Ghomi⁴, "Parkinson's Disease Diagnosis Using Machine Learning and Voice," DOI: 10.1109/SPMB.2018.8615607 , November 2018.
- [7] Dr. Zehra KARAPINAR SENTURK, Early Diagnosis of Parkinson's Disease Using Machine Learning Algorithms, 10.1016/j.mehy.2020.109603, 24 January 2020.
- [8] <https://www.datarobot.com/wiki/data-collection/#:~:text=Data%20collection%20is%20the%20process,the%20business%20problem%20at%20hand>
- [9] <https://www.kaggle.com/kmader/parkinsons-drawings>
- [10] <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11442/114420G/Selected-problems-of-image-data-preprocessing-used-to-performance-examination/10.1117/12.2565138.full>
- [11] <https://www.analyticsvidhya.com/blog/2021/09/a-beginners-guide-to-image-processing-with-opencv-and-python/>
- [12] <https://github.com/PyImageSearch/imutils#:~:text=imutils,A%20series%20of%20convenience%20functions%20to%20make%20basic%20image%20processing.Python%202.7%20and%20Python%203>
- [13] <https://scikit-learn.org/stable/modules/preprocessing.html>
- [14] <https://scikit-image.org/>
- [15] https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_training_test_data.htm
- [16] <https://towardsdatascience.com/image-processing-4391c5bcef78>
- [17] <https://www.mygreatlearning.com/blog/label-encoding-in-python/>
- [18] <https://www.techtarget.com/searchenterpriseai/feature/How-to-build-a-machine-learning-model-in-7-steps#:~:text=A%20machine%20learning%20model%20is,to%20data%20isn't%20enough>
- [19] <https://docs.aws.amazon.com/machine-learning/latest/dg/training-ml-models.html>
- [20] https://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_231#:~:text=In%20machine%20learning%2C%20model%20testing,evaluated%20on%20a%20testing%20set
- [21] https://www.researchgate.net/publication/336513175_A_Systematic_review_on_Application_based_Parkinson's_disease_Detection_Systems#:~:text=The%20various%20parameters%20like%20Electroencephalograph,y,disease%20progression%20in%20later%20stages
- [22] <https://www.frontiersin.org/articles/10.3389/fnagi.2021.633752/full>
- [23] <https://ieeexplore.ieee.org/document/8615607>
- [24] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8134676/>

A Study of Blockchain technology

Sandeep Kaur

BBSBE College , Department of Computer Science and Engineering , Fatehgarh sahib , Punjab ,
India , IK Gujral Punjab Technical University , Jalandhar, Punjab , India.

Abstract—Blockchain, the foundation of Bitcoin, has received extensive attentions recently. Blockchain serves as an immutable ledger which allows transactions take place in a decentralized manner. Blockchain-based applications are developed , covering infinite fields including financial services, reputation system and Internet of Things (IoT), and so on. However, there are still many challenges of blockchain technology such as scalability and security problems waiting to be conquered. This paper presents a comprehensive overview on blockchain technology. We provide an overview of blockchain concept firstly and compare some typical consensus algorithms used in different blockchains.

Keywords :Blockchain , leader , transaction , block , bitcoin , cryptography ,Token.

Introduction

The Blockchain was generated in 2008 and is proposed by Nakamoto in the Bitcoin white paper. It is an intelligent peer to peer network that uses distributed databases to determine broadcasting and record information also known as value internet. In 2009 , Nakamoto founded the bitcoin social network and developed the first block , “ Creation block ” [3].

Blockchain has developed rapidly in recent years and has become another technical innovation after cloud computing , big data , mobile internet and other new generation information technology . It has received attention from international organizations such as the united nation and the international monetary fund and the government of developed countries such as

the United Kingdom , the United States , Russia and Singapore in recent years , and

Parminder Kaur

has slowly been applied to financial services and supply chain management , culture , entertainment , intelligent manufacturing , social welfare , education and employment and other industries .

At the same time , the generation , identification , protection and trading of intellectual property rights are facing unprecedented challenges in the information technology era . The blockchain technology can solve the problem above . The correct use of blockchain technology will contribute to the protection and trading of intellectual property [3] [6].



Fig :1 Blockchain data structure and how Blockchain work

The Concept of Blockchain Technology

Blockchain Technology is a continuously growing list of records, called blocks, which are linked and secured using cryptography. Each block typically contains a 1. Blockchain or Distributed Ledger Technology (DLT) is a distributed ledger recording technology, which contains information about transactions or events. It can record transactions in a transparent, safe, decentralized, efficient, and low-cost way [2]. Hence, the Blockchain Technology has the following attribute : a distributed ledger, decentralized data management, data security, transparency and integrity,

anti-tampering and anti-forgery, high efficiency, low cost, programmable features that increase flexibility and reliability and no risk of a centralized database failure.

There are many types of Blockchains, some of the most important are: Public Blockchain, Private Blockchain and Consortium Blockchain (hybrid Blockchain). Each type has its advantages and disadvantages, allowing them to meet the needs of various applications .

Specifically, using a Public Blockchain, anyone can transact on the network transactions which are transparent and are identified . A Public Blockchain, such as bitcoin, is completely decentralized. The system operates based on users's consensus; there is no central point of failure. However, Public Blockchain is susceptible to system attacks. For instance, an attacker could recreate and properly chain all the blocks that had been modified, without being detected by the participants; b) Private Blockchain, the transactions are secret, the data is not available for public view, but the members are known. In a private Blockchain network, a participant cannot read or write the Blockchain unless the participant has a permission or an invitation to join the network. Private Blockchain is usually used by large companies with permissions defined between various stakeholders of the enterprise Blockchain. For instance, a bank can have its own Blockchain network for its private use with restricted access to its various shareholders such as customers, employees and suppliers; c) Consortium Blockchain is a hybrid model of both Public and Private Blockchain. Choosing this model, enterprises or institutions can have their own Private Blockchain network to share the data among the consortium participants (such as banks, institutions and other enterprises or firms) [2] [10].

Three Generations of Blockchain

The scope of blockchain applications has

increased from virtual currencies to financial applications to the entire socialdomain . Based on its applications, blockchain is defined to Blockchain 1.0, 2.0, and 3.0.

Blockchain 1.0

Blockchain 1.0 was related to virtual currencies, such as bitcoin, which was not only the first and most widely used digital currency but it was also the first application of blockchain technology. Digital currencies can reduce many of the costs associated with traditional material currencies, such as the costs of circulation. Blockchain 1.0 produced a great many applications, one of which was Bitcoin. Most of these applications were digital currencies and tended to be used commercially for small-value payments, foreign exchange, betting, and money laundering. At this stage, blockchain technology was generally used as a cryptocurrency and for payment systems that relied on cryptocurrency ecosystems [1] [4].

Blockchain 2.0

Broadly speaking, Blockchain 2.0 includes Bitcoin 2.0, smart-contracts, smart-property, decentralized applications (Dapps), decentralized autonomous organizations (DAOs), and decentralized autonomous corporations (DACs) (Swan 2015). However, most people understand Blockchain 2.0 as applications in other areas of finance, where it is Xu et al. The financial sector requires high levels of security and data integrity, and thus blockchain applications have some inherent advantages. The greatest contribution of Blockchain 2.0 was the idea of using smart- contracts to discontinued traditional currency and payment systems. Recently, the integration of blockchain and smart contract technology has become a popular research topic in problem settlement. For example, Ethereum, Codius, and Hyperledger have established programmable contract language and executable infrastructure to implement smart contracts[1] [4].

Blockchain 3.0

In ‘Blockchain: Blueprint for a New Economy’, Blockchain 3.0 is described as the application of blockchain in areas other than currency and finance, such as in government, health, science, culture, and the arts (Swan 2015). Blockchain 3.0 aims to popularize the technology, and it focuses on the regulation and governance of its decentralization in society. The scope of this type of blockchain and its applications suggests that blockchain technology is a moving target (Crosby et al. 2016). Blockchain 3.0 envisions a more advanced form of “smart contracts” to establish a distributed organizational unit that makes and is subject to its own laws and which operates with a high degree of autonomy. The integration of blockchain with tokens is an important combination of Blockchain

3.0. Tokens are proofs of digital rights, and blockchain tokens are widely recognized thanks to Ethereum and its ERC20 standard. Based on this standard, anyone can issue a custom token on Ethereum and this token can signify any right or value. Tokens refer to economic activities generated through the creation of encrypted tokens, which are principally but not exclusively based on the ERC20 standard. Tokens can serve as a form of validation of any right, including personal identity, academic diplomas, currency, receipts, keys, event tickets, rebate points, coupons, stocks, and bonds. Consequently, tokens can validate virtually any right that exists within a society. Blockchain is the back-end technology of the new era, while tokens are its front-end economic face. The combination of the two will bring about major societal conversion. Meanwhile, Blockchain 3.0 and its token economy continue to develop[1] [4].

Consensus Algorithms

In blockchain, how to reach consensus among the unreliable nodes is a transformation of the Byzantine Generals (BG) Problem, which was raised in . In BG problem, a group of generals who command a portion of Byzantine army circle the city

Some generals prefer to attack while other generals prefer to retreat. However, the attack would fail if only part of the generals attack the city. Thus, they have to reach an agreement to attack or retreat. How to reach a consensus in distributed environment is a challenge. It is also a challenge for blockchain as the blockchain network is distributed. In blockchain, there is no central node that ensures ledgers on distributed nodes are all the identical. Some protocols are needed to ensure ledgers in different nodes are consistent. We next present several common approaches to reach a consensus in blockchain.

A. Approaches to consensus

PoW (Proof of work) is a consensus strategy used in the Bitcoin network . In a decentralized network, someone has to be selected to record the transactions. The simple way is random selection. However, random selection is vulnerable to attacks. So if a node wants to publish a block of transactions, a lot of work has to be done to prove that the node is not likely to attack the network. Generally the work means computer calculations. In PoW, each node of the network is calculating a hash value of the block header. The block header contains a nonce and miners would change the nonce frequently to get different hash values. The consensus requires that the calculated value must be equal to or smaller than a certain given value. When one node reaches the target value, it would broadcast the block to other nodes and all other nodes must mutually confirm the correctness of the hash value. If the block is validated, other miners would append this new block to their own blockchains. Nodes that calculate the hash values are called miners and the PoW procedure is called mining in Bitcoin.

In the decentralized network, valid blocks might be generated coeval when multiple nodes find the suitable nonce nearly at the same time. However, it is unlikely that two competing forks will generate next block simultaneously. In PoW protocol, a chain that becomes longer thereafter is judged as the authentic one. Consider two forks created by simultaneously validated blocks U4 and B4

Miners keep mining their blocks until a longer branch is found. B4,B5 forms a longer chain, so the miners on U4 would switch to the longer branch. Miners have to do a lot of computer calculations in PoW, yet these works waste too much resources. To reduce the loss, some PoW protocols in which works could have some side-applications have been designed. For example, Primecoin searches for special prime number chains which can be used for mathematical research.

PoS (Proof of stake) is an energy-saving alternative to PoW. Miners in PoS have to prove the ownership of the amount of currency. It is believed that people with more currencies would be less likely to attack the network. The selection based on account balance is quite unjust because the single richest person is bound to be dominant in the network. As a result, many solutions are proposed with the combination of the stake size to decide which one to forge the next block. In particular, Blackcoin uses randomization to expect the next generator.

It uses a formula that looks for the lowest hash value in combination with the size of the stake. Peercoin favors coin age based selection. In Peercoin, older and larger sets of coins have a greater probability of mining the next block. Compared to PoW, PoS saves more energy and is more effective. Unfortunately, as the mining cost is nearly zero, attacks might come as a consequence. Many blockchains adopt PoW at the beginning and transform to PoS gradually. For instance, Ethereum is planning to move from Ethash (a kind of PoW) to Casper (a kind of PoS)[7].

PBFT (Practical byzantine fault tolerance) is a duplicate algorithm to tolerate byzantine faults . Hyperledger Fabric utilizes the PBFT as its consensus algorithm since PBFT could handle up to 1/3 malicious byzantine replicas. A new block is determined in a round. In each round, a primary would be selected according to some rules. And it is responsible for ordering the transaction. The whole process could be divided into three phase: pre-prepared, prepared and commit. And it is responsible for ordering the transaction.

The whole process could be divided into three

phase: pre-prepared, prepared and commit. In each phase, a node would enter next phase if it has received votes from over 2/3 of all nodes. So PBFT requires that every node is known to the network. Like PBFT, Stellar Consensus Protocol(SCP) is also a Byzantine agreement protocol. In PBFT, each node has to query other nodes while SCP gives participants the right to choose which set of other participants to believe. Based on PBFT, Antshares has implemented their dBFT (delegated byzantine fault tolerance). In dBFT, some professional nodes are voted to record the transactions [8].

DPOS (Delegated proof of stake) The major disparateness between PoS and DPOS is that PoS is direct democratic while DPOS is representative democratic. Share holders elect their delegates to generate and validate blocks. With significantly fewer nodes to validate the block, the block could be confirmed quickly, leading to the quick confirmation of transactions. Meanwhile, the parameters of the network such as block size and block intervals could be tuned by delegates. Additionally, users need not to worry about the dishonest delegates as they could be voted out easily. DPOS is the backbone of Bitshares.

Ripple is a consensus algorithm that utilizes collectively-trusted subnetworks within the biggest network. In the network, nodes are divided into two types: server for participating consensus process and client for only transferring funds. Each server has an Unique Node List (UNL). UNL is important to the server. When determining whether to put a transaction into the ledger, the server would query the nodes in UNL and if the received agreements have reached 80%, the transaction would be packed into the ledger. For a node, the ledger will remain correct as long as the percentage of faulty nodes in UNL is less than 20% [9].

Tendermint is a byzantine consensus algorithm. A new block is determined in a round. A proposer would be selected to broadcast an unconfirmed block in this round. It could be divided into three steps: 1) Prevote step Validators choose whether to broadcast

a prevote for the proposed block. 2) Precommit step. If the node has received more than 2/3 of prevotes on the proposed block, it broadcasts a precommit for that block. If the node has received over 2/3 of precommits, it penetrates the commit step. 3) Commit step. The node validates the block and broadcasts a commit for that block. If the node has received 2/3 of the commits, it accepts the block. Contrast to PBFT, nodes have to lock their coins to become validators. Once a validator is found to be dishonest, it would be penalized [6].

Conclusion :

From a theoretical perspective, based on the literature review, Blockchain Technology has high value and good prospects in resolving problems of data integrity, improving transparency, enhance security, preventing fraud, and establish trust and privacy. Blockchain Technology can bring revolution in the areas of Finance, Accounting, e-Government, BPM, insurance, entertainment, trading platforms, healthcare, internet-of-things, as well as law firms and others. Hence, Blockchain Technology has a huge potential in introducing innovative solutions, depending on the area or the sector of its implementation, since economic efficiency and social benefits can be achieved through technical innovation and applications. However, implementing Blockchain Technology at organizations in different industries, could prove to be very costly. Migrating or moving legacy systems require a significant amount of investment from organizations. Adopting the Blockchain Technology, at this early stage, organizations will have to deploy a 13 unified platform to support such hybrid application architecture, incorporating Blockchain and legacy systems. Thus, they need to deepen their understanding of Blockchain Technology, its value, its opportunities, and its risks. As a result, there are only a small number of instances in which the technology has been applied with these systems. Therefore, Blockchain Technology may not replace legacy systems or old applications soon. However, Blockchain can certainly be a complementary application to legacy systems and may even lead to the development of new systems in the near

future. In conclusion, more intensive research in this area of Blockchain Technology is necessary to advance the maturity of this field, since it is still in the exploratory stage and there are many legal and technical issues to be resolved. Therefore, this review offers a useful starting point for future research themes for the development of Blockchain application, and assist practitioners and researchers.

References :

- [1] Xu2019 Article A Systematic Review Of Blockchain.pdf
- [2] 1996a557.pdf
- [3] 1-s2.0-S187705091930239X-main.pdf
- [4] 0_AnReviewStudyofBlockchainTechnologyandItsApplicationsinBusiness_13F3Y1.pdf
- [5] “State of blockchain q1 2016: Blockchain funding overtakes bitcoin,” 2016. [Online]. Available: <http://www.coindesk.com/state-of-blockchain-q1-2016/>
- [6] J. Kwon, “Tendermint: Consensus without mining,” URL <http://tendermint.com/docs/tendermint{ } v04. pdf>, 2014.
- [7] S. King and S. Nadal, “Ppcoint: Peer-to-peer crypto-currency with proof-of-stake,” Self-Published Paper, August, vol. 19, 2012.
- [8] C. Miguel and L. Barbara, “Practical byzantine fault tolerance,” in Proceedings of the Third Symposium on Operating Systems Design and Implementation, vol. 99, New Orleans, USA, 1999, pp. 173–186.
- [9] D. Schwartz, N. Youngs, and A. Britto, “The ripple protocol consensus algorithm,” Ripple Labs Inc White Paper, vol. 5, 2014
- [10] J. Barcelo, “User privacy in the public bitcoin blockchain,” 2014

Internet of Things(IOT)

Prabhjot kaur

BBSBE College , fatehgarh sahib , Punjab , India , department of computer science and engineering ,

I.K. Gujral P. T.University , Jalandhar , Punjab , india .

Abstract: The word “internet” can be regarded as one of those constant living objects that have been developed over the past few years and has continuously been gaining the newer heights in the modern era. In other words, it has always been upgradable and changeable along with its evolution at the same time. Different various technologies with the newest features have been something very attractive that has successively gained people’s attention. And with no doubt, this can be stated as the major reason behind the milestones that internet has received in very less time. With an evolution of an internet, various things related to the internet have been rapidly increased, which can relatively be termed as the Internet of Things (IoT). Rightfully so, the combination of technology and fun creates IoT for us. Since the use of IoT is a huge number in this short span of time. In this paper we will discuss about its brief introduction , its history the growth of IoT, characteristics of IoT , advantages and disadvantages of IoT.

Keywords : Internet of things (IoT), Business, Development, Evolution, Existence, , Market Study, Research, Things, Upcoming .

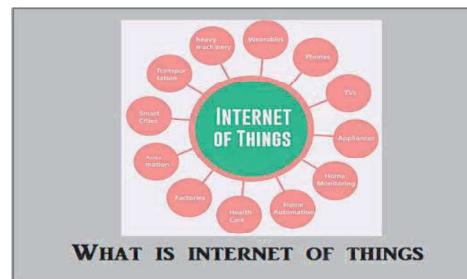
Introduction:The internet of things, or IoT,

is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers (UIDs) and the

ability to transfer data over a network without requiring human-to-human or human-to computer interaction A thing in the internet of things can be a person with a heart monitor implant, a farm animal with a biochip transponder, an automobile that has built in sensors to alert the driver when tire pressure

is low or any other natural or man-made object that can be assigned an IP address and is able to transfer data over a network. Increasingly, organizations in a variety of industries are using IoT to operate more efficiently, better understand customers to deliver enhanced customer service, improve decision-making and increase the value of the business .The internet of things (IoT) is a computing concept that characterize the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices. The term is closely identified with RFID as the method of communication, although it also may include other sensor technologies, wireless technologies or QR codes.

Fig :1 Uses of internet of things in surrounding [10]



History:In 1989, the "Internet" was first introduced and quickly spread throughout the world. Since the birth of the Internet, the trend for connecting various things to the Internet has become very large. Coffee Pot Trojan is the first application of its kind. In 1990, John Romkey designed the first Internet device, a toaster that could be activated and deactivated with the Internet. Steve Mann developed Wearcam in 1994. This happens almost in real time with a 64-processor system. In 1997, Paul Saffo gave first short description of censorship and his future actions.

In 1999, Kevin Ashton created Internet of Things, Managing Director of the Auto-ID Center at MIT. They also formed RFID-based systems for identification of objects throughout the world. As a major step in IoT commercialization, LG Electronics Giant made an announcement about its plan in 2000 to supply refrigerators that would now identify whether the stored food would be refilled or not. In 2003, the US Army used RFID in a huge number as a part of the Savi Program. In 2008, a group of companies had established the IPSO Alliance to increase the use of Internet Protocol (IP) in the smart object network and to enable the Internet of Things. In 2008, the FCC accepted the use of the "white space spectrum". The introduction of IPv6 in 2011 has generated tremendous growth and interest in this area. IT giants such as Cisco, IBM and Ericsson are continuing many educational and commercial initiatives with IoT. IoT technology can also be interpreted as the most important link between people, the Internet and various things. Everything we use today in our daily lives can be monitored and monitored via the Internet (IoT). Most of the processes in IoT are carried out with the help of sensors. This sensor is used in almost all devices. The sensor first receives physical raw data or data (data), converts it into a digital signal, and sends it back to the control center. We can say that IoT has brought great comfort to our lives and made it very easy for all of us. In this way, we can also monitor changes in the environment in other parts of the world by using the internet remotely in the world. So, IoT has a big impact and plays an important role in our daily lives[5].

Growth of IOT: Internet has been part and parcel of the social animal's life. It's a huge space of information and people. The internet first evolved as "internet of computers". It is a global platform where many services like the World Wide Web could be implemented on top of it. It was an era of information exchange. As the days passed by, people started emerging into the internet- "Internet of people". Many social websites came into picture which kept people connected all the time. This has led to internet being filled with people rather than information. On the other hand, technology has been advancing day by day and simultaneously an era of "MobiComp" (mobile

computing) had begun. Mobile helped man to be always connected to the internet on the move. Nowadays 3G and 4G mobile internet connections have led to faster internet access and deliver better quality in video calls. Wireless technologies and mobile computing have become cheap and have gained more popularity. Hence a new computing had emerged- Ubiquitous computing.

Nowadays 3G and 4G mobile internet connections have led to faster internet access and deliver better quality in video calls. Wireless technologies and mobile computing have become cheap and have gained more popularity. Hence a new computing had emerged- Ubiquitous computing. This computing focuses on smart, intelligent space and minimal user involvement. Advancement in technology led to mobile and other hand-held devices to diminish in size. Smart phones, Ipads, tablets and notebooks replaced ordinary mobiles and PCs. Hence there was a change in the device with which people access the internet. This in turn resulted in sophisticated features being configured in devices such as sensors, Global Positioning system (GPS) and actuators. In such a scenario devices were not only connected to the internet but also sense, compute and perform intelligent tasks. Later physical objects were configured with identification tags such as bar code and RFID so that they could be scanned by devices like smart phones and upload their information into the internet. This way of connecting the physical world with cyberspace with the help of a smart device led to internet being called as "Internet of Things". Hence IOT has its roots from Mobile computing, ubiquitous computing and information technology. IOT connects the objects in an intelligent way. The "thing" here refers to the physical object's information read through sensors and RFID reader and uploaded into the internet. The physical object can be anything from smart phones to objects at home. The International telecommunications Union (ITU) has pointed out four dimensions of IOT : object identification ("tagging things"), sensors and wireless sensor networks ("feeling things"), embedded systems ("thinking things") and nanotechnology ("shrinking things"). Hence from the above, IOT changes the connectivity view from "any-

time , anywhere” for “any-one” into “any-time , anywhere” for “any-thing”. These things once connected to the internet provide smart services beneficial to the environment and society. They play a major role in supply chain, energy, defence, health care and other useful applications[1] [6].

Characteristics of IOT:

Some most popular characteristics of Internet of things are:

- (a) Intelligence (b) Connectivity (c) Dynamic Nature (d) Enormous scale (e) Sensing (f) Heterogeneity (g) Security

(a) Intelligence IoT comes with the combination of algorithms and computation, software & hardware that makes it smart. Ambient intelligence in IoT enhances its capabilities which facilitate the things to respond in an intelligent way to a particular situation and supports them in carrying out specific tasks. In spite of all the popularity of smart technologies, intelligence in IoT is only concerned as means of interaction between devices, while user and device interaction is achieved by standard input methods and graphical user interface [4].

Together algorithms and compute (i.e. software & hardware) provide the “intelligent spark” that makes a product experience smart. Consider Misfit Shine, a fitness tracker, compared to Nest’s intelligent thermostat. The Shine experience distributes compute tasks between a smartphone and the cloud. The Nest thermostat has more compute horsepower for the AI that make them smart[4].

(b) Connectivity Connectivity empowers Internet of Things by bringing together everyday objects. Connectivity of these objects is pivotal because simple object level interactions contribute towards collective intelligence in IoT network. It enables network Convenience and compatibility in the things. With this connectivity, new market opportunities for Internet of things can be created by the networking of smart things and applications. Connectivity in the IoT is more than slapping on a WiFi module and calling it a day. Connectivity enables network

accessibility and compatibility. Accessibility is getting on a network while compatibility provides the common ability to consume and produce data. If this sounds familiar, that’s because it is Metcalfe’s Law and it rings true for IoT .

(c) Dynamic Nature The primary activity of Internet of Things is to collect data from its environment, this is achieved with the dynamic changes that take place around the devices. The state of these devices change dynamically, example sleeping and waking up, connected and/or disconnected as well as the context of devices including temperature, location and speed. In addition to the state of the device, the number of devices also changes dynamically with a person, place and time.

The state of devices change dynamically, e.g., sleeping and waking up, connected and/or disconnected as well as the context of devices including locality and speed. Moreover, the number of devices can change dynamically[4] [7].

(d) Enormous scale The number of devices that need to be managed and that communicate with each other will be much larger than the devices connected to the current Internet. The management of data generated from these devices and their interpretation for application purposes becomes more critical. Gartner (2015) confirms the enormous scale of IoT in the estimated report where it stated that 5.5 million new things will get connected every day and

6.4 billion connected things will be in use worldwide in 2016, which is up by 30 percent from 2015. The report also forecasts that the number of connected devices will reach 20.8 billion by 2020.

The number of devices that need to be managed and that communicate with each other will be at least an order of magnitude larger than the devices connected to the current Internet. Even more critical will be the management of the data generated and their interpretation for application purposes. This relates to semantics of data, as well as efficient data handling[7].

Difference between IOT and Traditional Network:

In the beginning, the IoT technology has broken a lot of the traditional ideas of network and started a new era of telecommunication technology. Can be

considered IoT as an extension and expansion network based on the Internet; but it is dissimilar from either traditional network or the socalled

. Internet of people and WSN although considered as backbone to build any IoT block. The major equation to represent the IoT environment is "IoT environment= Internet + WSN", it is a common statement that uses to express the IoT environment. To analyze and judge the correctness of this statement, must be determined the similarities and differences between IoT, Internet, and WSN. From the previous knowledge about the IoT environment can be judged on this view, it's a wrong; because there are two basic reasons for rejecting this view. First; IoT may not necessarily use IP in all cases for addressing things, because nature of IoT needs lightweight communication protocols, the complexity of the TCP/IP protocol is not suitable in particular, when works with the smart little things. Second, the IoT environment is mainly based on the connected smart objects unlike traditional network. That's what makes them move from only a mere extension of the Internet, also the behavior of IoT depends on the creation of the interoperable systems [10], based on these arguments, can be corrected the previous statement: IoT= Internet + WSN+ Smart Items surrounded by Intelligent environment[10].

Advantages of IOT:

Access of Data

The more information is available, the easier it is to make an appropriate decision. You have access to real-time data and information that is far away from your location . Knowing what you get from the supermarket by going out without checking yourself not only saves time but also remains practical. This is only possible because a device network gives a person access to all information in the world. This makes it very easy for people to do their work even when they are not physically present[2].

Connectivity On the network of directly connected devices, better communication is possible, making device communication more transparent and

reducing inefficiencies [16]. Processes in which machines have to work with each other become more effective and produce better, faster results. The machines in the production or production unit are the perfect examples [18]. Another example is in our home, the computers follow both the quality and durability of things. The knowledge that the product ends before consumption increases safety and quality of life. Moreover, you will never run out of anything when you need it at the last minute[8].

INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 12, DECEMBER 2019 ISSN 2277-8616

2993

IJSTR©2019 www.ijstr.org

Time Saving By programming the work, whenever needed or required will be completed and doing this will save human valuable time and energy[2] .

Expenditure Effective As mentioned above, communication with electronic devices through Internet networks can be easily facilitated. It helps people in their daily work. The transfer of data packets to a connected network saves time and money . The same information that can be transmitted faster can be done less than ever, just by Internet of Things[8].

Disadvantages of IOT: 7.1 Privacy/Security Privacy is a big issue with IoT. All information must be encrypted so that you can back up your financial status data . Nowadays, every device is connected worldwide via the internet. This increases the risk of data loss, which can be important. This is a major disadvantage when exchanging information, as confidential information may not be secure and can easily be damaged by third parties. Complexity A diversified network that connects different devices together is what we call internet things.

A hole in the system can affect the entire system. It is by far the most complex aspect of IoT that can have a major impact. There are various destruction options in complex systems. For example, you and your family can receive news that the milk is ready and you can all buy the same . This means that you need

different quantities. Or a software error has occurred that allows the printer to order ink multiple times if it only needs one cartridge[8].

Dependency We may not notice it, but we see a large technology transfer and its implementation in daily life. There is no doubt that technology dominates our way of life and reflects human reliability in technology . If there is a malfunction in the system, there is a risk of damaging a connected device. It will affect our daily lives, as we become increasingly dependent on it[2].

Business Model In the following you will find a great impetus to start a business, to invest and to run a business. Without a good and solid business model for the Internet of Things, we have another bubble. This model should fulfill everything that is required for all types of e-commerce. standing markets, tight markets and hypermarkets. However, this category is always a victim of regulatory and legal investigation [8].

Society Understanding the IoT from the consumer's point of view is not an easy thing to do because their needs or requirements change over time, they want new features in existing devices as well as new ones. And that happens at a fast pace; before the solution to the previous problem is found, a new problem appears. And solving problems takes time and resources, leaving a bad picture of the IoT as a whole

Regulatory Standards Regulatory standards for the information market are inadequate for information brokers leading to companies that sell data grasped from different sources .

Even if the data seems to be the Internet's motto, there is a lack of transparency in who has access to data and how it is used to improve products or services and to sell it to people, announcers and third parties. Clear guidance on data retention, use and security, including metadata, is required [2]. The European Union has adopted data and data-sharing regulations.

- [1].[A Study on Internet of Things based Applications.pdf](#)
- [2].[A-Detailed-Study-Of-An-Internet-Of- Things-iot.pdf](#)
- [3]. [ali-2015-ijca-906430.pdf](#)
- [4].[internet-of-things-iot-applications-and-security-challenges-a-review-IJERTCONV7IS12028.pdf](#)
- [5].[3 The Future of Internet of Things IoT and Its Empowering Technology](https://www.researchgate.net/publication/33206055)
- [6].[22 Internet of Things IoT In a Way of Smart World](https://www.researchgate.net/publication/30380922)
- [7].[56 INTERNET OF THINGS IOT ADVANTAGES ON E-LEARNING IN THE SMART CITIES](https://www.researchgate.net/publication/322249356)
- [8].Internet of Things, 2014
<http://postscapes.com/internetof-things-history>.
- [9].InternetofThings,2015
<http://www.rfidjournal.com/articles/view?4986>.
- [10].E. Fleisch, Auto-ID Labs: What is the Internet of Things? - An Economic Perspective, 2010.

References:

Enterprise Resource Planning System for Business Management

Dr. Rakesh Kumar

Asstt. Professor, P.G. Department of Commerce, S.G.G.S. Khalsa College, Mahilpur

Introduction

An Enterprise resource planning system is a fully integrated business management system covering functional areas of an enterprise like Logistics, Production, Finance, Accounting and Human Resources. It organizes and integrates operation processes and information flows to make optimum use of resources such as men, material, money and machine. Enterprise resource planning promises one database, one application, one user interface for the entire enterprise, where once disparate systems ruled manufacturing, distribution, finance and sales

Enterprise resource planning helpful for a business concern to collect, store, manage and interpret data from many business activities such as: Product planning, cost and development, Manufacturing or service delivery, Marketing and sales management, Inventory management and for Shipping and payment

purpose etc. It is business process management software that allows an organization to use a system of integrated applications to manage the business and automate back office functions. ERP software integrates all facets of an operation, including product planning, development, manufacturing processes, sales and marketing

ERP allows companies to integrate various departmental information. It has evolved from a human resource management application to a tool that spans IT management. For many users, an ERP is a “do it all” system that performs everything from entry of sales orders to customer service. It attempts to integrate the suppliers and customers with the manufacturing environment of the organisation. For example, a purchase entered in the order module passes the order to a manufacturing application, which in turn sends a materials request to the supply-chain module, which gets the necessary parts from suppliers and

uses a logistics module to get them to the factory. At the same time the purchase transaction shows in general – a ledger module as revenue. The traditional application systems, which organisations generally employ, treat each transaction separately. They are built around the strong boundaries of specific functions that a specific application is meant to cater for.

Features of ERP System

- (i) Modular design comprising many distinct business modules such as financial, manufacturing, accounting, distribution, etc.
- (ii) Centralized common database management system (DBMS)
- (iii) The modules are integrated and provide seamless data flow among the modules, increasing operational transparency through standard interfaces
- (iv) High Cost but flexible system.
- (v) The modules work in real time with online and batch processing capabilities
- (vi) Internet-enabled system.

An ERP system is designed keeping in mind the financial management, customer relationship management, human resource management, supply chain management and manufacturing resource planning of the

business enterprise. These aspects are also shown in figure 1 given below:-



Figure 1: An Overview of ERP System in Business Organisation

Components of an ERP System

- 1. ERP Software** - Module based ERP software is the core of an ERP system. Each software module automates business activities of a functional area within an organization. Common ERP software modules include product planning, parts purchasing, inventory control, product distribution, order tracking, finance, accounting and human resources aspects of an organization.
- 2. Business Processes** - Business processes within an organization falls into three levels – strategic planning, management control and operational control.
- 3. ERP Users** - The users of ERP systems are employees of the organization at all levels, from workers, supervisors, mid-level managers to executives.

4. Hardware and Operating Systems -

Many large ERP systems are UNIX based. Windows NT and Linux are other popular operating systems to run ERP software. Legacy ERP systems may use other operating systems. The boundary of an ERP system is usually small than the boundary of the organization that implements the ERP system. In contrast, the boundary of supply chain systems and e-commerce systems extends to the organization's suppliers, distributors, partners and customers. In practice, however, many ERP implementations involve the integration of ERP with external information systems.

Advantages of ERP

- (i) Gives Accounts Payable personnel increased control of invoicing and payment processing and thereby boosting their productivity and eliminating their reliance on computer personnel for these operations.
- (ii) Reduce paper documents by providing on-line formats for quickly entering and retrieving information.
- (iii) Provides timely information on daily basis instead of monthly.
- (iv) Maintain accuracy of information with detailed content presentation which will also helpful for the auditors.

- (v) Rapid response and follow-up on customers.
- (vi) Easily monitoring and quicker resolution of queries with regard to changed business conditions.
- (vii) Helps to achieve competitive advantage by improving its business process.
- (viii) Provides a unified customer database usable by all applications.

Challenges

The ERP business segment is accompanied by a lot of challenges. One of the challenges concerned with its cost. Vendors are offering ERP at an introductory cost of Rs 4-6 lakh. In fact, the cost of ERP software should not be viewed as an expense but as an investment, towards an ability that provides better profitability and customer service. The costs will remain high but the advantages far outweigh the expense. At present educating about the system to small and medium companies about their benefits is not an easy task. To conquer this problem, solution providers and implementers are spreading awareness in their target audience. There are also technical challenges regarding training, integration and implementation, which can take from two to 12 months.

Despite of above challenges, some problems also exist in the company's perspectives. These are includes Top Management attitude towards IT adoption, Lack of IT Knowledge, Lack of Business Process Reengineering culture, prolonged implementation of the system, Insufficient IT Resources, lesser competency of employees, Poor Infrastructure etc.

Conclusion

ERP systems are sets of integrated applications that can provide a total solution to an organization's information system needs by addressing a large proportion of business functions including financial, accounting, human resources, supply chain and customer information. Indian context, the following issues can assume immense importance namely, clarity in goals and objectives behind the implementation, adequacy of user training, competency of the project implementation team, acceptance of changes brought about by the implementation and adequate vendor support and external consultant participation. It concludes that SMEs make use of ERP solutions mostly for contingency, exogenous reasons rather than as a result of accurate analyses of their needs

and opportunities. Use of ERP is a new issue and there is a need to develop these systems in such a way so that every business enterprise can adopt it without any restriction.

References:-

- [1].Loh T.C. and Koh S.C.L. (2004), "Critical elements for a successful enterprise resource planning implementation in small-and medium-sized enterprises", International Journal of Production Research, vol. 42, no. 17, 3433-3455.
- [2].Sharma K.M. and Bhagat R.,(2006) "Practice of information systems Evidence from select Indian SMEs", Journal of Manufacturing Technology Management, Vol. 17 No. 2, pp. 199-223.
- [3].Sledgianowski D. and Tafti, Mohammed,H.A. (2008),"SME ERP system sourcing strategies: a case study", Industrial Management & Data Systems, Vol. 108 No. 4, pp. 421-436.
- [4].Suraweera T., Pulakanam V., Guler O.(2006), "Managing the Implementation of IT Projects in SMEs: An Exploratory Investigation", Digital Information Management

Davenport, T. H. (1998). Putting the enterprise into the enterprise system.
Harvard Business Review, 76(4), 121-131.

E-GOVERNANCE MEASURES IN PUNJAB POST COVID - QUO VADIS?

Amrita Sekhon

*Department of Computer Science
Punjabi University
Patiala*
amritakaur@gmail.com

Dr Sukhwinder Singh

*Department of Computer Science
Guru Hargobind Sahib Khalsa Girls College
Karhali Sahib, Patiala*
oberoimca@gmail.com

Abstract - COVID-19 or Corona Virus was reported in Wuhan for the first time in December 2019, which soon spread in other parts of China, and slowly transformed into a global pandemic by March 2020. With the death toll constantly increasing, its epicenter of target gradually shifted to Europe from Asia, while that of affected people shifted to the USA, India and Brazil, the most. Although different countries used different governance mechanisms to tackle with the pandemic, some governance decisions made a difference like extensive usage of emerging technologies besides efficient medical treatment. This in addition to community behavior and solidarity not only made the response more efficient but also controlled the spread of the disease. Although it was a global pandemic but its responses were local and relied on cultural, socio-economic context and local governance. In the ongoing corona virus disease 2019 (COVID-19) crisis, uncertain and experimental forms of governance have emerged. Regular administrative routines and traditional management techniques were abruptly dissolved amid crisis events and management by the state of exception. This paper analyzes the responses in several countries across the world like China, South Korea, Japan, Netherlands, Taiwan, Switzerland, USA and Australia and finally provides some e-initiatives taken by the Government of Punjab (India).

Index Terms – Covid-19, E Governance, Punjab, Crisis Response, Governance reforms.

I. INTRODUCTION-BRIEF HISTORY OF THE PANDEMIC

It is widely accepted that Corona virus (or COVID-19 virus or SARS-CoV-2 virus strain) was reported for the first time in December 2019 in Wuhan, China, and was acknowledged as a new virus by Chinese authorities in January 2020. The World Health Organization's Beijing office received the first information about an unknown virus on December 31, 2019. The WHO declared it as a Public Health Emergency of International Concern by the end of January 2020. Although the Chinese authorities made efforts to check the spread of this disease, but it swiftly spread not only to other parts of China but to a number of other countries too, by the end of January. The WHO Director General used the term 'infodemic' at the early stage of spreading of the disease [1], which seems to be applicable till date. The WHO warned the public against the tsunami of information, particularly on the social media, which often calls for a panic-like situation. This has been observed in several countries like the spreading of fake news through social media. The WHO named this virus as COVID-19 virus

on February 11, 2020. The global epicenter of the disease shifted from China to other countries officially on February 24, 2020 when the number of infected people outside China crossed those within China. Apart from Korea and Japan, Iran and northern Italy were observed as the two major clusters. There was a steady growth of infected people globally during early March. The WHO confirmed that the virus had affected more than 100,000 people in 100 countries by March 7, 2020. Finally the WHO declared it as a global pandemic on March 11, 2020. The USA declared it as a national emergency on March 13, 2020. From that point on, it has continued to affect different parts of the world with USA being the one affected the most. By October 18, 2020, the virus had affected more than 215 countries and territories worldwide [2]. The virus affected a huge number of people, took a significant number of lives and brought down the economies of many countries under lockdown. Several countries made travel bans, locked down cities and provinces, which significantly impacted not only the local economy but the global economy as well.

II. LITERATURE ON GOVERNANCE IN CRISIS RESPONSE

COVID-19 virus and its gloomy repercussions spread promptly across the globe. The whole world was caught by surprise and all nations had to act without having time to reflect.

A lot research on public governance and its ability to cope with complex problems can be found, but there is a dearth of research on the issue of turbulence and its impact on public governance. In connection with handling turbulent problems, research literature on crisis management by governance can be found in [3][4][5] and for emergency response in [6][7][8][9]. This literature is focused on dedicated crisis response organizations instead of public agencies in particular, and is more concerned about communication and coordination rather than finding out ways to produce adequate governance solutions. Research related to problems concerning robust governance in turbulent situations remains at an infancy stage, and the range of governance strategies to ensure robustness needs further exploration.

It is not possible to avoid turbulent situations in this global world where freight, information and people keep crossing borders at an ever-increasing pace. It is tough for the public to bounce back from such situations since restoring the old

equilibrium may not be feasible in most cases. In other words, any nation cannot handle turbulent problems by recruiting well trained, dedicated workforce or by having warehouses equipped with emergency apparatus ready to gear up for the next uncertain, unpredictable emergency situation that may take the public sector by storm. Resilience, foresight and protection are not sufficient to handle turbulent situations. Alternatively, turbulent situations must be met with robust strategies devised by the public sector where agile, creative public agencies adapt to the disruptive problems by developing partnership and creating networks with civil society and the private sector [10].

It is important for the public programs and agencies to become more agile and flexible so that they may adapt and transform themselves to handle turbulence scale up and down their problem solving efforts [11][12]. Modularized, flat and easily integrated agencies would adapt to the new, emerging requirements better than insulated, big compartmentalized hierarchies. Along with the new organizational design, there is a need to change the mindset, management vocabulary and routine setup of employees and managers from standard service production to scalable, innovative and creative solutions that enhance robustness.

III- GOVERNANCE REFORMS POST COVID ACROSS THE WORLD

The pandemic spread across all countries of the world, where some countries and regions were more severely affected as compared to others. Since the spread of COVID-19 has been uneven, similarly the policy response has also been uneven. Different styles of governance have been adopted by different countries. This section attempts to summarize some key lessons learnt on governance at varied levels. The effective response systems may be based on the double-loop learning process developed by the experience of SARS (Severe Acute Respiratory Syndrome) of 2003, H1N1 (novel swine-origin Influenza-A) of 2009 and MERS of 2015.

A. Strong government control:

China presented an extremely strong government control right from the beginning, which was since the third week of January when there was an official confirmation of the first COVID-19 case. In addition to the lockdown in Wuhan of the Hubei province, which was gradually spread across the whole country, strict measures were taken to avoid promotion of fake news and panic right from the initial stages. Strict Supreme Court advisories were issued on fake news all over the country. Different type of resources and supplies were provided by the different provincial governments to help the most affected province – Hubei and city – Wuhan to compact the situation [13].

B. Transparency and democracy:

South Korea was successful in responding to COVID-19 as they transparently disclosed accurate information and held on to democracy of the entire society [14]. Since the first COVID-19 case was confirmed in Korea on January 20, 2020, their government in coordination with the Disease Control and Prevention Agency of Korean (previously known as KCDC-Korean Center for Disease Control and Prevention), exchanged necessary information with various Chinese agencies and the WHO and made public their government responses, which lead to their citizen's voluntary participation without protest. The Korean government traced the confirmed cases using big data analysis, CCTV and credit card usage data analysis and reveled them freely using government website and CBS (Cell Broadcasting System) mobile service [15].

The wide circulation of the confirmed cases data helped citizens to determine if they had contacted those people, so they could report to the public health centers. If they carried any sign of the virus infection, then their diagnostic test was carried out. Since the Korean government was very well prepared for performing diagnostic analyses and testing, they promptly analyzed all the potentially infected citizens, which resulted in prevention of the spread of the infectious diseases.

C. Expert based advices:

Japan did not call for any national emergency or lockdown as their legislation does not approve a forced lockdown, but permits an advisory or request for the lockdown. Japan undertook this cautious approach based on the outcome of close interaction with expert groups from diverse fields like medicine, economics, social and politics. Their government officials held regular meetings with these experts. Japan's governance approach aimed to flatten the viral infection growth curve, such that their health response mechanism got enough time to respond to the emergency situation, which would probably provide enough time to develop the much needed vaccine and devise other preventive measures.[13]

D. Agile and adaptive governance in crisis response:

Countries across the globe had to use adaptive governance within a short time span in response to the COVID-19 situation. Agile governance deals with the working practices and methods used to facilitate quick response to any situation. [16] The COVID-19 pandemic needed adaptability predominantly in the domains like prescription drug supply, testing and contact tracing, medical equipment supply, hospital capacity, food supply assurance and the financial support to keep the economy going. The **Netherlands** adopted a 'smart lockdown' policy to combat the COVID-19. The general idea behind the 'smart lockdown' was to encourage various anti-virus measures rather than make them mandatory. Thus the responsibility was shifted to the individual level. It also gave

the individuals a certain degree of freedom to adapt to their local conditions and their situation as they saw fit. Using this approach, the Netherlands was successful to quickly scale up the intensive care capacity and kept their economy going. The **Dutch** response to COVID-19 suggested that adaptively and agility can work together, but there may be some conflict when put to practice. Thus adaptive and agile governance have different purpose, origin and implications, so they must not be mixed together. Agility is essentially applied to organizations and is a way of working used in software development. Adaptive governance is method-agnostic and takes an emergency approach.[17]

E. Digital governance infrastructure:

The effective COVID-19 response of **Taiwan** depends on three factors. Firstly, they properly implemented the mask policy as a preventive measure and the quickly enhanced their mask production ability. Secondly, they used technology and big data for increasing the various disease detection and prevention measure implementation successfully. Thirdly, their well-built state-society relationship enhanced collaboration, communication and transparency in Taiwan. Its developmental state model policy enhances the first two factors. The strong democratic relation of the state and civil society intensified the citizens' voluntary obedience and its crisis governance authenticity. The nation provided transparent communication about the nature of the disease, which reduced panic situation among public, thereby increasing their compliance level as well as their trust in government. The crisis measures were further refined by their technology savvy society. Their winning strategy of crisis management won applaud for Taiwan at the international level.[18]

F. Transformative governance as a new emerging governance form:

Transformative governance emerged around the globe as most nations adopted measures for substantial, immediate relief to economies, civil societies and health systems. According to Bosomworth [19] four important requirements need to be satisfied for transformative governance, namely experimentation, spontaneous learning, addition of diverse perspectives and decision-making under uncertainty. **Switzerland** observed a change in governance forms which could be termed as transformative one. This transformation relied on a number of factors like financial resources, societal preferences, political will and power relations. Transformative governance promotes digital working of schools, restructuring supply chain system, open discussion on data privacy issues, better public security in public transport, global production networks and above all increased preparedness for pandemic like crisis situations. Observing Transformative governance in

Switzerland, four characteristic issues can be observed namely, paradigms driving policy change, evidence base of policy-making, potential of experimental governance and role of the state in such transformative governance. [20]

G. Individualistic, Voluntary, non-government approach:

The **USA** chose an individualistic, voluntary and non-government type of approach as the most feasible governance option. The public sector enterprise, Centre for Disease Control advised the states to exploit the private sector enterprise, Google-Apple Exposure Notification API. Separate apps were developed by more than 20 US states. [21] The public had more trust on this platform provider than the US government, especially IT governance. The US population poll also suggested that the public trusted in technology at large. Also the US companies mandated their employees to use proximity tracing apps, which the US government could not mandate.

H. Centralized proximity tracing solution:

Australia employed a varied, balanced approach for governance. A Centralized proximity tracing solution, COVIDSafe was implemented by the Australian government, which was based on Singapore's TraceTogether architecture. The public was encouraged to adopt this app by a number of entities. Not only did the Australian government strongly promote this app on TV, messaging and other health campaigns, but also many private companies joined these initiatives. They all actively advertised the use of contact tracing apps by employees for a speedy recovery period. [22] Thus the successful implementations of a number of different COVID-19 responses were observed that can be credited to the able leadership of various government heads. Precise decisions were taken based on the advice of expert committees which tried to understand the upcoming risk factors of the unpredictable situations faced globally.

III- IT- INITIATIVES OF PUNJAB [23]

COVA Punjab (Corona Virus Alert) was the first Mobile App to be launched in India, on 9th March 2020, bearing geotagging and geo-fencing features. It was developed by the Government of Punjab for updating its citizens with various government advisories, Covid-19 Hospitals with Nodal officers contact numbers, Check Health Status Chat-bot based services, FAQs and preventive care information. Eleven other states of India requested Punjab to get access to this App and its features. By mid-August 2020 as many as 51 Lac+ downloads had been done of this App which was available on both Android PlayStore and iOS AppStore. Some key features of the app include finding the nearest COVID patient, contact tracing of positive cases, geo-fencing of various quarantined

individuals, reporting of mass gathering, real time dashboard and analytics, order placing facility for grocery and essential items, plasma donor registration and booking of train tickets, to name a few. This app in addition with Aarogya Setu app of Government of India use contract tracing technique to benefit the citizens.

Harvest Season Management To ensure smooth operations in Mandis during the procurement weeks under COVID-19, Punjab developed a unique centralized automated, logic-based tech platform to issue e-Passes to farmers. All movement in the Mandis was managed online.

A dedicated dashboard - **corona.punjab.gov.in** was prepared to analyze the trends and patterns of cases including prediction of future hotspots to help the authorities to take timely decisions.

A platform - **covidhelp.punjab.gov.in** was setup exclusively for those people who wanted to go to their home states from Punjab and those wanting to return.

Punjab launched **tele-medicine and tele-counseling** initiatives for the general patients other than COVID-19 to get consultation from home with the empanelled doctors/counselors through video call.

Punjab deployed **UiPath RPA BOT** to extract district-specific information from the email, coming from more than 20 district authorities. The BOT saves the information in respective district folder.

Punjab has initiated a **mass awareness campaign** to educate and inform its citizens across all the channels (Radio, TV, Social Media, Print Media) about this pandemic.

eSanjeevani (Online OPD), a CDAC Mohali's flagship integrated telemedicine solution was implemented across the state to connect with a network of senior doctors over video conferencing.

Video Conferencing Solutions were leveraged to facilitate Work-from-Home for the staff along with important webinars and VCs by the CM and other administrative staff.

Plasma donor registration collaborated with Department of Governance Reforms for testing and on boarding of portal/software through COVA APP for registration of plasma volunteers.

Online COVID-19 Trainings Numerous physical trainings were provided by the Government of Punjab to different departments – Police, Social Security, Women and Child Development, Local Bodies and Agriculture. By August 2020, at least 6812 Medical Officers, Resident Medical Officers, AYUSH Doctors, Staff nurses, Pharmacists and lab technicians, etc. were provided online training. As many as 343 National Urban Health Mission (NHUM) Staff, 546 Special Newborn Care Unit Staff and Labor Room Staff were trained online. Also 125 Sanitation workers, 125 House Keeping staff and 91 people from the Prison department were given online training.

Operation FATEH was a public awareness campaign of the government of Punjab to control the spread of Covid-19 in the state. The aim of Mission Fateh was to make the citizens aware of the precautionary measures which were taken by the government along with the health department to check the spread of corona virus. The campaign was carried out in both rural and urban areas of the state. A total of 4409 participants were trained as on 15 August 2020 (India's Independence Day) under the campaign. It was a one month long campaign under which the State Government created mass awareness about personal health and hygiene to be maintained to prevent spread of the disease.

GOT Trainings Apart from the medical staff online trainings, even the government employees were provided integrated online training on the DIKSHA app for capacity building as frontline workers. The Course was completed by 2,91,917 participants as on 14 August 2020, thus making Punjab the second state in the country to provide such mass training campaign.

IV- CONCLUSION

The COVID-19 pandemic had dreadful consequences, causing illness and fatality around the world. On the other hand, it gave an opportunity to governments to showcase their inherent potential to adapt. Governments along with semi-public organizations in collaboration with private organizations toiled hard to cope up with qualms from strategic to operational level. The paper studied the successful responses of governments of different countries and then analyzed the response of government of Punjab, India to the COVID-19 pandemic. It also discussed how adaptive governance was put in practice. Specifically, role of E Governance was investigated.

Nations globally had to act in response to the COVID-19 outbreak with inadequate information and many uncertainties. Their capability to be agile and adaptive was stressed, predominantly with regard to the policy measures implementation timings, autonomy of decision making, level of decision centralization and equilibrium between change and stability.

The most common instant government response to this pandemic was to impose a near-complete lockdown of society which impacted the economy of most under developed nations and forced poor people lacking any welfare support to pick between illness and hunger. Other common solution like social distancing and wearing a facemask were more experimental solutions that gave evidences of reduced risk of infection. Thus it was comprehensible that private, public and third-sector actors like citizens themselves, mobilized in their effort to control the health crisis and the consequential social and economic crisis situation.

When COVID-19 came up, nations globally had to respond at different levels of government in collaboration with the society. Different response mechanisms appeared from diverse

domains and the issue was approached differently in each domain. Governments took the responsibility to take the lead along with decision making based on expert committees' advice and inform the public to finally impose many restrictions on them for the betterment of everyone. The Corona virus outbreak imposed a sense of emergency to adjust to fight the pandemic together.

Digital contact tracing emerged as a significant measure for restricting the spread of SARS CoV-2, the virus causing COVID-19 respiratory disease [24]. Contact tracing is the successful public health practice to recognize and intimate about those people who personally contacted the contaminated person during their contagious time slot.

While the world is now struggling with the third wave of the pandemic, there are a few important lessons learnt which are as follow:

Although global implications are there, but the response to the Pandemic is local:

True as the saying goes, the world is just a global village; the world is so well interconnected, such that people are constantly on the move from one country to another. This aggravated the pandemic situation, making it a global affair within days. But different countries used different response approaches. Although medical treatment measures were universal, but the emergency health response measures of nations were not universal. The response approaches were an amalgamation of a nation's regulation policies based on science-based decision making, different governance mechanisms at national and local level along with community behavior.

Appropriate use of information technology:

The pandemic response was simply not just medical response. Different types of technologies were linked to each other in a suitable way. COVID-19 response in East Asian countries like South Korea China and Japan showed widespread use of promising technologies viz. 5G, block chain, artificial intelligence, robotics, big data, drone, automated vehicle, etc. worked in collaboration to various medical technologies.

Risk assessment: Djlante et al. [25] in a quick analysis has pointed out the need of converging the health response, emergency response and disaster risk reduction in the viewpoint of the Sendai Framework. They analyzed and concluded that current mechanisms and strategies for disaster resilience, as outlined in the SFDRR, can enhance responses to epidemics or global pandemics such as COVID-19. Some of the recommendations are as follow: recommendations concern knowledge and science provision in understanding disaster and health-related emergency risks, the extension of disaster risk governance to manage both disaster risks and potential health-emergencies, particularly for humanitarian coordination aspects; and the strengthening of community-level preparedness and response. A proper risk assessment is

required taking into consideration of health risk, exposures, behaviors and policy framework.

Use of social media and sensitization on fake news: In different countries, with different level of social media penetration, the importance of distinction of proper news and fake news becomes more relevant. Importance of negative consequences of fake news is well understood in longer run, not only to fight this pandemic, but also for the longer-term recovery process.

Economic implications: The global economic impacts of the pandemic are yet to be understood, but there is a unanimous agreement of a global recession due to the pandemic. However, in different countries, sectorial impacts are already prominent, especially in tourism and hospitality sectors. MSMEs (Micro, small and medium enterprises) are possibly the hardest hit in all the countries need special economic revitalization package.

Socio-psychological impacts and lifestyle changes: Country wide or partial local down in cities have initiated a different work culture in East Asian countries, as well as in most of the other countries. Tele-work is becoming popular, online meetings, online classes in the universities are getting common, online education for school children becoming obvious. Thus, there has been a life-style change in many countries and communities, which may have relatively longer socio-psychological and behavioral implications.

REFERENCES

- [1] Zaroncostas, J. 2020. How to Fight an Infodemic in www.thelancet.com, Vol 395, Accessed on: 2nd October, 2020.
- [2] <https://www.worldometers.info/coronavirus/> (Accessed: 15 October 2020)
- [3] Bland, M. 1995. Strategic crisis management. In *Strategic public relations* (pp. 276-298). Palgrave Macmillan, London.
- [4] Groh, M. 2014. "Strategic Management in Times of Crisis." *American Journal of Economics and Business Administration* 6 (2): 49-57. doi:10.3844/ajebasp.2014.49.57.
- [5] Ansell, C., and A. Boin. 2019. "Taming Deep Uncertainty: The Potential of Pragmatist Principles for Understanding and Improving Strategic Crisis Management." *Administration & Society* 51 (7): 1079-1112. doi:10.1177/0095399717747655.
- [6] Drabek, T. E. 1985. "Managing the Emergency Response." *Public Administration Review* 45 (1): 85-92. doi:10.2307/3135002.
- [7] Perry, R. W., and M. K. Lindell. 2003. "Preparedness for Emergency Response: Guidelines for the Emergency Planning Process." *Disasters* 27 (4): 336-350. doi:10.1111/j.0361-3666.2003.00237.x.
- [8] Manoj, B. S., and A. H. Baker. 2007. "Communication Challenges in Emergency Response." *Communications of the ACM* 50 (3): 51-53. doi:10.1145/1226736.1226765.
- [9] Aven, T., and O. Renn. 2010. *Risk Management and Governance: Concepts, Guidelines and Applications*. New York: Springer.
- [10] Ansell, C., Sørensen, E. and Torfing, J., 2020. The COVID-19 pandemic as a game changer for public administration and leadership? The need for robust governance responses to turbulent problems. *Public Management Review*, pp.1-12.
- [11] Overbeek, S., B. Klievink, and M. Janssen. 2009. "A Flexible, Event-Driven, Service-Oriented Architecture for Orchestrating Service Delivery." *IEEE Intelligent Systems* 24 (5): 31-41. doi:10.1109/MIS.2009.90.

- [12] McCann, J., and J. W. Selsky. 2012. Mastering Turbulence: The Essential Capabilities of Agile and Resilient Individuals, Teams and Organizations. New York: John Wiley & Sons.
- [13] Shaw, R., Kim, Y.K. and Hua, J., 2020. Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. *Progress in disaster science*, p.100090.
- [14] Normile, D., 2020. Coronavirus cases have dropped sharply in South Korea. What's the secret to its success. *Science*, 17(03).
- [15] Park, O., Park, Y.J., Park, S.Y., Kim, Y.M., Kim, J., Lee, J., Park, E., Kim, D., Jeon, B.H., Ryu, B. and Ko, D., 2020. Contact transmission of Covid-19 in South Korea: Novel investigation techniques for tracing contacts. *Oson Public Health and Research Perspectives*, (1).
- [16] Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R. and Kern, J., 2001. Manifesto for agile software development.
- [17] Janssen, M. and van der Voort, H., 2020. Agile and adaptive governance in crisis response: Lessons from the COVID-19 pandemic. *International Journal of Information Management*, p.102180.
- [18] Yen, W.T., 2020. Taiwan's COVID-19 Management: Developmental State, Digital Governance, and State-Society Synergy. *Asian Politics & Policy*, 12(3), pp.455-468.
- [19] Bosomworth, K., 2018. A discursive-institutional perspective on transformative governance: A case from a fire management policy sector. *Environmental Policy and Governance*, 28(6), pp.415-425.
- [20] Willi, Y., Nischik, G., Braunschweiger, D. and Pütz, M., 2020. Responding to the COVID-19 Crisis: Transformative Governance in Switzerland. *Tijdschrift voor economische en sociale geografie*, 111(3), pp.302-317.
- [21] Dave, P. 2020. Google says 20 US states, territories ‘exploring’ contact tracing apps. itnews. <https://www.itnews.com.au/news/google-says-20-us-states-territories-exploring-contact-tracing-apps-551132> (Accessed: 15 October 2020)
- [22] Riemer, K., Ciriello, R., Peter, S. and Schlagwein, D., 2020. Digital contact-tracing adoption in the COVID-19 pandemic: IT governance for collective action at the societal level. *European Journal of Information Systems*, pp.1-15.
- [23] Punjab COVID 19 Response Initiatives and Policy Measures, Prepared by Governance Fellows, August 2020
- [24] Sun, K. and Viboud, C., 2020. Impact of contact tracing on SARS-CoV-2 transmission. *The Lancet Infectious Diseases*. (pp 876–877).
- [25] Djalante, R., Shaw, R. and DeWit, A., 2020. Building resilience against biological hazards and pandemics: COVID-19 and its implications for the Sendai Framework. *Progress in Disaster Science*, p.100080.

DIGITAL CORPORATE GOVERNANCE IN INDIA

Prof. Sandeep Kaur

Assistant Professor

(Post Graduate Department of Commerce)

SGGS Khalsa College, Mahilpur.

E-mail: skrandhawa688@gmail.com

Mob. No. 7696235221

ABSTRACT

Corporate governance has emerged as an issue of worldwide importance. Corporate governance is an exercise of power for steering an organisation, as well as a process by which organizations are directed, controlled, and held to account to their stakeholders. It is a set of the systems and processes concerned with ensuring the overall direction, effectiveness, supervision and accountability of an organization. Electronic corporate governance is a relatively new discipline and subject of significance for both public policy and marketers. It is useful to recognise that it is a dynamic concept, in terms of scope, thrust and relevance. The application of ICT for the purpose of governance is commonly known as e- corporate governance. In view of the large number of corporate scams and scandals shocking the nation, e-governance tools have to be employed at a large scale to improve corporate governance. Through e- corporate governance, information can be distributed to the public in a transparent manner. So this research paper highlights the concepts of e corporate governance in India.

Keywords: E-governance, ICT, MCA-21, Corporate Governance, Digitalisation

INTRODUCTION

Corporate governance has become the major focus and matter of concern for the corporation in the 21st century. Various corporate scandals and collapses,

especially Enron debacle in 2002 in the US, witnessed the drastic change of corporate governance systems across the world. Corporate governance requirements in India were only in law books. But after Harshad Mahta case, an independent capital markets regulator, Securities and Exchange Board of India (SEBI) was set up by an Act of Parliament in 1992. Government also tried through SEBI and by amending the existing Companies Act 1956 to harmonise the practices & procedures followed by different corporate. The Companies Act 2013 incorporates various provisions which are going to bring a revolution in the field of corporate governance. The Act clearly and strongly accepts and adopts the use of technology for corporate governance. Corporate governance is an important framework for ensuring efficient, profitable, and ethical advancement of a business organisation. Good corporate governance is not just about compliance with the legal requirements. It is about the commitment on the part of management to ensure 'fair play' in decision making and ensuring that the broad objectives of the organisation are taken care of in a balanced and transparent manner.

E-governance is the use of electronic means, to promote good governance. It connotes the implementation of information technology in the processes and functions so as to cause simple, moral, accountable and transparent governance.

DIGITAL CORPORATE GOVERNANCE

Corporate governance technologies enable organizations to move a large part of their manual and paper-based workflows to online processes. Board portals and e-Meeting systems, which enable virtual meetings, are some of the ways to implement digital corporate governance. Use of integrated solutions provided by a single service-provider at an enterprise level will ensure seamless data sharing, reduced security risks, and improved user experience.

Digital corporate governance includes:

- Transfer of paper based board meeting processes to a secure digital platform.
- Instant cloud-based access to board materials .
- Virtual work groups for committee collaborations, commenting and discussion tools, project boards.
- Online proxy appointment and online voting.
- Geographical mobility, multi-device support, and real-time data sharing abilities.
- Meeting management solutions to simplify administration of large-scale meetings like AGMs.
- Board pack creation, agenda building, attendance recorder, minutes publication, and action items management capabilities.
- Virtual/hybrid meeting capabilities.

Ministry of Corporate Affairs (MCA) has started an ambitious e-governance project called the MCA-21 e-governance programme and spending huge amount for e-filing and e-registration. It also aims at continuously repositioning MCA as an organisation capable of fulfilling the aspirations of its stakeholders in a globally competitive business environment.

In addition to the conventional challan-based off-line payment system in the pre-

MCA 21 system, online payment systems have been introduced, including use of digital signatures based on a Director Identification Number (DIN) database. These services are now available on a 24 X 7 X 365 time frame. MCA or the Ministry of Corporate Affairs has introduced many initiatives in order to facilitate the life of corporates and businesses in India. One of these initiatives was inducting One-day Company Incorporation. Services offered by MCA21 are: obtain digital signature, apply for DIN, LLP e-filing, e-stamping, e-taxation integration, integration with banking institutions and introduction of refund process etc.

Digital governance initiatives taken under Companies Act, 2013

Electronic book maintenance:

Section 120 of the companies' act 2013 provides that companies may maintain and present books and accounts in electronic form. By this record production cost, time and paper will save.

Sending annual reports electronically:
Income statements and financial statements can be sent to members in electronic form.

Notice of meeting:

Notice of AGM/ EGM (annual/extraordinary general meeting) and board meeting can be sent by companies electronically to their shareholders at their registered e-mail addresses.

E-voting:

E-Voting is an web based system through which participation of shareholders in larger numbers in the decision making process of companies is made easy through login and registration of their votes on company resolutions. To encourage huge participation of shareholders in the general meeting; provision of e-voting has been made in the new act. All most every big

company adopted this feature to make transparent decision making.

Board meeting through video conferencing:

Any interested board of directors may join board meeting through video conferencing from remote place if he is unable to join it physically. This participation will also be counted in quorum formation. Speed, accuracy and ease to the Board's functioning will increase with this provision.

E-filing:

The Companies Act, 2013 contains provisions for e-filing and online payment of fees by companies. Section 20 (1) provides that – “A document may be served on a company or an officer thereof by sending it to the company or the officer at the registered office of the company by registered post or by speed post or by courier service or by leaving it at its registered office or by means of such electronic or other mode as may be prescribed: Provided that where securities are held with a depository, the records of the beneficial ownership may be served by such depository on the company by means of electronic or other mode.

GOOD GOVERNANCE BEFORE DIGITAL CORPORATE GOVERNANCE

Corporate working should be transparent, reliable, complying all rules and regulations and satisfying all stakeholders' interest. Main three pillars in corporate governance i.e. directors, management and shareholders must work ethically. Good governance becomes best with the application of digitalisation. So, good governance is the key essential for the success of digital corporate governance.

RISK INVOLVED IN DIGITAL CORPORATE GOVERNANCE

Major risk involved is cyber security. How to secure very sensitive data? How to prevent hacking? Security is the main concern for the stakeholders and citizen while redefining rules and procedures, information transparency, legal issues, infrastructure, skill, awareness, access to right information, inter-departmental collaboration, tendency to resist the change in work culture etc. are some other risk dimensions.

CONCLUSION

E-governance is definitely essential for Indian corporate especially in present-day world of cutthroat competition, and ever-increasing need for greater transparency and accountability. Considering these highly significant facts, the Government of India has rightly promulgated the provisions for e-governance in the corporate sector of the country, in its latest Companies Act of 2013. Provisions like E-voting towards e-governance are not only a move which will largely support green initiatives but will increase transparency and enhance organizational effectiveness. This is a more efficient, convenient and cost effective method. The top priority of corporate governance policies of the companies need to focus on the interests of the shareholders and protecting their rights. The regulatory frameworks are in place, the regulatory bodies and other responsible authorities are actively searching for the best way to strengthen its regulations to enhance corporate governance practices.

REFERENCES:

1. Diana Moșneanu “corporate governance in the digital world” 14th international conference on business pp. 333-342, issn 2558-9652
2. Veeramani Siv “digital transformation and corporate

- governance in india: a conceptual analysis “8th international conference on management and economics – isbn 978-955-1507-66-4
3. Dr. shruti singh, Dr. Rubee Singh, “impact of e-governance in india: opportunities & challenges” vol 07 issue07, jun 2018 issn 2456 – 5083
 4. <https://economictimes.indiatimes.com/news/economy/policy/e-governance-initiative-mca21-for-companies-extended-till-july-2021/articleshow/16531997.cms?from=mdr>
 5. egalservicesindia.com/article/1454 /Implementation-of-Corporate-E-Governance:-Implementation-of-MCA-21-Model.html

Weighted Common Spatial Pattern method for feature extraction from EEG signals in BCI system

Navtej Singh

Punjabi University, Patiala,

Punjab, India

Email: navtejghuman@sbsstc.ac.in

Balkrishan

Yadavindra College of Engineering,

Talwandi Sabo, Punjab, India

Email: balkrishan@pbi.ac.in

Abstract—Very weak signals emanate from the human scalp, whenever a motor imagery (MI) activity happens in the brain. These signals can be fetched using electroencephalogram (EEG) recording. Hidden patterns in these signals can be analyzed to predict the MI task intended by the subject. Although brain-computer interface (BCI) applications extend to various fields including navigation, educational, security, and virtual gaming, the main motivation for BCI research is to enable disabled individuals to control their artificial limbs and communicate with external systems. However, its use is restricted by lack of robustness. Machine learning can be used to implement preprocessing, feature extraction and classification stages of a BCI system. The Common Spatial Pattern (CSP) algorithm is a popularly used for extracting features from EEG signals. This paper presents an approach by improving its performance by assigning appropriate weights to its spectral parameters. It leads to improvement in classification accuracy, when tested on benchmark dataset.

Keywords—EEG, motor imagery, machine learning, brain-computer interface

1. Introduction

Brain-computer Interface (BCI) is a technique of establishing a direct communication channel between the brain of user and external devices, without using normal nerve pathways [1]. It deals with computer-aided control by exclusively analyzing signals from the brain activity. It has found applications not only in engineering but also in medical fields. Most commonly used BCI systems uses the technology of electroencephalogram (EEG) because of its affordable recording equipment of noninvasive nature which is conducive for its real-time operation [2]. The BCI transforms the imaging of intended task directly into machine-implementable commands by bypassing the conventional neural muscular conduction pathway in the human body [3]. It is basically a communication mechanism which does not require any movement of peripheral muscular. Applications of BCI are found in areas of robotics, mobility devices, environmental control, and device communication. It has the capability to provide an alternative communication as well

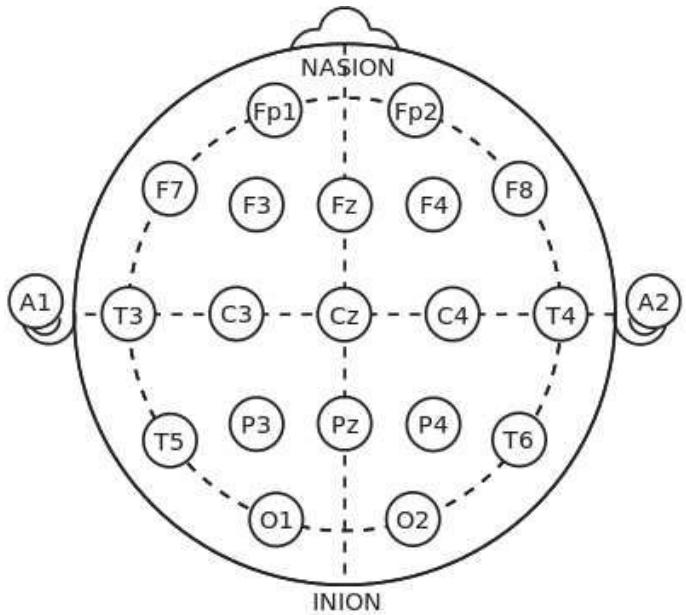


Figure 1. International 10-20 system of electrode placement [5]

as control technology, to patients suffering from severe neuromuscular disorders [4]. Most of the common BCIs are based on the EEG, in preference over other methods of data fetching from brain waves. The EEG based BCIs have an affordable and easy-to-use recording equipment, which can be used to provide real-time operation [2]. The human brain controls and coordinates different body parts by the transmission of neural signals. Electroencephalography (EEG) can be used for recording functional brain activity as it is non-invasive, safe and easy to use, as compared to invasive techniques. An EEG-based BCI system analyzes signals fetched by electrodes placed as per standard international 10-20 system of EEG as shown in figure 1, on designated parts of the scalp [6].

EEG signals are fetched from the scalp of the brain, which are generated as an associated result of motor imagery (MI). MI is a mere mental intention of a motor act, without its actual execution by the corresponding muscle. MI is a commonly performed mental task in which subjects are

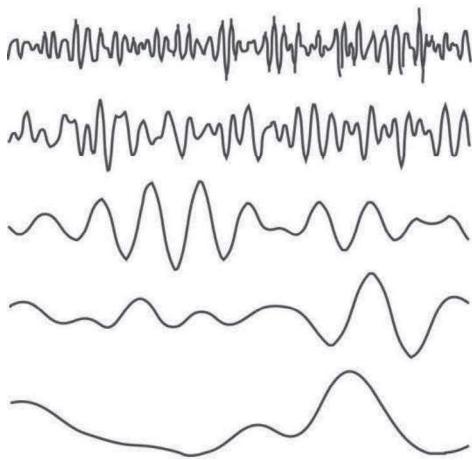


Figure 2. EEG Sub-bands

asked to imagine themselves executing a particular movement such as that of a hand or foot but without actually performing it physically. It is observed by electrical recordings from the scalp surface of the subject person, as it is accompanied by a continuous activity of the brain waves in various parts of the brain [7]. Immense growth has been witnessed in the field of bio-informatics where techniques have been investigated for extraction, interpretation and analysis of biological data. Every BCI system is required to be vigorous, portable, safe, affordable and easy to use, for communication or device control. The equipment required for signal acquisition should be minimalist and easy to setup so that it can be used without extensive setup and training [4]. The brain waves lack regularity, and thus it is difficult to establish a specific pattern in the EEG [8]. The signals emanating from motor imagery activity, have to be recognized by the BCI system. It is a daunting task as the EEG signal is non-stationary in nature and is adversely affected by undesirable electrical interference, involuntary muscular movement artifacts and is also subject specific. The precise and accurate classification of MI tasks requires that the most distinguishing features are extracted from the brain signals acquired in the data fetching phase of BCI. Many algorithms, which are collectively called as classifiers are used for this purpose. EEG is a test that involves measurement of electrical signals emanated to the scalp by the brain, by placing electrodes on the outer surface of scalp. EEG is a typical non-invasive method which records brain's electrical activity along the scalp surface, however invasive electrodes are sometime used in specific applications. In such a BCI implementation, electrical signals are recorded from scalp surface of head, as there is a continuous electrical activity in the brain, which is directly in correspondence to intended tasks [7]. The EEG signals are further categorized into sub-categories of EEG sub-bands - delta, theta, alpha, beta and gamma based on frequency of signal. The standardized 10-20 electrode placement system of EEG demarcates the position of electrodes on various parts of the scalp. It is used to record signals emanating as byproduct of a mental

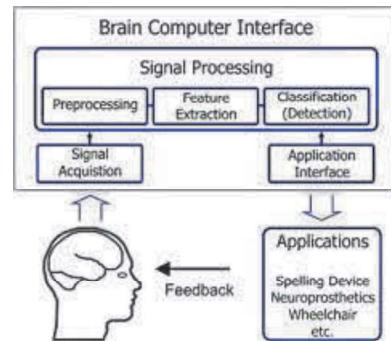


Figure 3. BCI model

activity, from somatosensory and the motor areas of brain. The signals recorded from these electrodes, reflect the motor activity of the person like hand movements, foot movements and tongue movements etc. [9]. The EEG-based BCIs can manipulate MI of users, and thus can cater the requirement of robust communication performance between brain and computer. There are many other techniques of acquiring such signals, such as that of magnetic resonance imaging (MRI) and positrons emission tomography (PET). However EEG has emerged as preferred choice for implementation of noninvasive BCI system, as it is inexpensive, has higher temporal resolution (approximately 0.05s) and is easy to use [10]. The foremost application of BCI is to aid patients suffering from neurological disabilities, by providing them with an alternative method of communication with some external devices. The other BCI applications include transmitting instructions in entertainment applications besides communicating with personal computers by use of thoughts' translation into appropriate corresponding control signals. BCI is a integrative research field, which encompasses domains of psychology, neurology, human-computer interaction (HCI), rehabilitation engineering, signal processing and that of machine learning. The main objective of BCI applications, is to meticulously translate the brain wave patterns automatically extracted from emanating EEG signals into desired machine control commands. The aim of many researches is to improve the accuracy of interpretation of the harvested EEG signals [11]. Many BCI systems, which are based on EEG for signal fetching, were recently developed. In these systems, patterns of EEG emanating from different motor imagery tasks were analyzed and classified by using different algorithms in multiple stages of a BCI system. It has been concluded in the researches that efficiency and effectiveness of a BCI system depend on the quality of algorithms used for signal acquisition, pre-processing, feature selection and that of classification phase [12]. A BCI is implemented using a particular pattern recognition approach which is characterized by extracting features from EEG signals and then using an appropriate algorithm for classifying it into a mental task.

Properties of EEG signals

Various features of EEG signals which can be analyzed for implementation of BCI system are amplitude of EEG signal, value of power spectral density (PSD), band powers (BP), autoregressive parameter, besides some time-frequency features [13]. Some of the critical properties of EEG signals are:

- Noise and Outliers: EEG signals fetched from scalp of human brain are highly prone to noise, due to sheer complexity of brain. The human body is continuous host to various activities of physiological and neural nature, which are closely intertwined, and are often undistinguishable from analysis of the waves emanating from the skin.
- High Dimensionality: In implementation of a BCI, the feature vector extracted from EEG signals, is multi-dimensional, as multiple features are fetched from several channels, and can span to multiple time segments. These multiple features which are extracted, are preprocessed and then combined and concatenated to a single feature vector.
- Temporal Information: The BCI features represented by EEG signals should have particular time-stamped information engrossed in them, to represent the particular time variations of the signals.
- Volatility: The BCI features are highly volatile in nature, and are stationary for a very short time and vary at a great rate. Hence the signals have to be fetched at a very high speed, which leads to an increase in the complexity. The short life span of the signals complicates the signal fetching stage of the BCI.
- Small Training Set: The BCI system has to be trained, which is a time consuming exercise. This also has adverse effect on subjects, who have to spend considerable time. Usually the subjects are not very cooperative while training the proposed BCI.

2. Data Acquisition

The data is acquired from analysis of measurement of brain activity, using many methods, which are broadly classified as invasive and non-invasive. However, the EEG has gained most acceptance, mainly because of its non-invasive character, low setup cost and relatively higher resolution. The electric potential in a single dipole EEG system a function of value of its charge, the distance separating the poles and distance to the poles and is represented as following:-

$$V_P = \frac{qd.r}{4\pi\epsilon_0 r^3} = \frac{qd\cos\theta}{4\pi\epsilon_0 r^2}. \quad (1)$$

Many BCI systems have also been developed using other alternative methods like functional magnetic resonance imaging (fMRI). It detects variation in local cerebral blood volume, its rate of flow besides levels of oxygenation. These parameters change during neural activation by means of

electromagnetic fields and magneto-encephalogram (MEG). It registers the brains magnetic activity emanating from scalp, by measuring magnetic induction. The data fetching in BCI systems can also be performed by invasive interfacing like electrocorticograms (ECOG). It requires surgical procedure, so as to place electrode directly underneath the scalp, to measure pattern of electrical activity taking place in area of cerebral cortex. The surgically installed electrodes directly over the surface of the brain, which requires a surgical operation of craniotomy to create an electrode grid [14]. This can be a serious health hazard, and thus non-invasive BCIs are preferred over it. The BCI systems are usually based data fetching by recording of the EEG signals, associated with brain activity to recognize associated patterns of various mental tasks. It has been established that various motor imagery tasks like episode of only imagination of moving the left and right hands, are correlated with characteristic pattern of EEG activity observed from the electrodes placed on the scalp of respective right and left side portion of the motor cortex. Such characteristic patterns are associated with various variations in EEG activity. For instance, it has been established that during movement or even preparation for the movement, the beta rhythms (18-26 Hz) and mu rhythm (8-12 Hz) decrease which is called as event-related desynchronizing. They have been observed to increase immediately following execution of this movement, which is called as event related synchronization [4]. Many researchers have made use of the data prepared and provided by BCI competition IV, which is a continuation in the series of previous BCI competitions, which are organized to provide neuroscientific data, which has a high quality and can be openly accessed by the scientific community [5]. The data set has signals recorded from scalp of nine human subjects, while they perform motor imagery tasks. The EEG recording include a total of twenty five channels that mostly were distributed over and around sensory motor areas of the scalp. Each participating subject contributed in a total of two sessions consisting of a calibration session and an evaluation session. A calibration session includes performance of four of the motor imagery tasks, by each of the subject. These are categorized on basis of imagining movement of left hand as class-1, right hand as that of class-2, both of the feet as of class-3, and that of the tongue as class-4. A visual cue is shown on the computer screen as a signal to the subject, for performing a specific motor imagery task for 4 seconds. Data set was recorded from 9 subjects, labeled as A01, A02, A03, ..., A09), while they performed four-class motor imagery tasks. The nine subjects were seated in front of a computer screen in an arm chair. A signal appears on the screen which is a fixation cross appearing on black screen, which signals start of a trial ($t=0s$). Another cue appears on the screen after two seconds ($t=2s$), in the form of an arrow pointing in the direction of right, left, right, up or down. Each of this is corresponding to four classes of right hand, left hand, tongue or foot movement, which appears on the screen for 1.25 seconds. These arrows displayed in one of the four directions give signal to the subjects undertake the intended motor imagery task, without

providing any feedback. The participating subjects were instructed to execute the motor imagery task as long as the fixation cross continues to appear on the screen ($t=6s$). A brief break of blank screen was incorporated between each cycle of above data recording. Two such recording sessions were executed for each of the protracting subject. Each of the above session comprised of six executions, which were separated by short breaks in between. Thus each of the execution consisted of 48 such recording trials, which made up of 12 runs for each of the four possible MI classes, generating data for a total of 228 different trials.

3. Preprocessing

This is a first phase of signal processing, which as the role of preparing the recorded electric signal by enhancing it, so as to enhance the features for their improved distinguishability for detection. Pre-processing techniques further improve the ratio of signal to noise (SNR), by rejecting unwanted artifacts. The efficiency of the preprocessing phase has a great role in the efficiency attainment of overall BCI system [15]. Technique of independent component analysis (ICA) is usually applied for the removal of power line noise as well as EOG, EMG and ECG artifacts.

4. Feature Extraction

Once the brain signals have been converted into digital format, they are then processed by one or more feature extraction methods. The feature extraction methods extract specific characteristics from the fetched signals, which contains messages or commands emanating from subject's scalp of brain by evoked or spontaneous inputs. Pre-processing and feature extraction are two important steps in EEG signal processing. Pre-processing techniques are employed to isolate unintended artifacts from the EEG signal. This leads to improvement in the signal to noise ratio, which leads to a better performance of the system. Subsequently, a feature extraction process retrieves the relevant features from the input signal. These features are required for decision making mechanism in generating the required output. EEG signals are very volatile in nature, and sustain for a very short time. There is a change in spectrum with time hence different feature extraction methods are required for this variation. There are many feature extraction methods, which are available. However, methods of autoregressive modeling, wavelet transform and fast fourier transform, are popularly employed to accomplish feature extraction in BCI. These methods used for feature extraction are explained below, so as to analyze and identify the most effective feature extraction method.

Feature Extraction

Feature extraction is a process of reducing the raw data to accomplish pattern recognition. The most distinguishing features from EEG signals have to be extracted for their

classification in next stage. Large data sets require a lot of memory and computing power to process. The EEG signals from MI tasks are applied upon by different feature extraction techniques, as described below, before these can be input to classifiers for classification.

4.1.1. Common Spatial Pattern. The EEG signals can be represented by certain set of features, which are described by few numerical values called as features [16]. CSP is a popularly used feature extraction technique in BCI [17]. It creates a new feature vector which has better variance and discriminatory power than the original EEG signal. Employing simultaneous diagonalization of two covariance matrices, this method focus on activities with particular spatial distributions through combining data from multiple channels. However, it is applicable only in two-class classification BCIs. Hence multi-class BCI problems have to be converted into multiple two-class classification problems. CSP establishes spatial filters which maximize the variance of one class and minimize the variance of the other class.

SWCSP

Signals from multi-channel EEG are correlated and contain varying quantity of discriminative information. A reduction in number of channels may not provide sufficient information for satisfactory performance of the BCI. However, an increase in number of EEG channels generates more noise and redundancy leading to degradation in BCI performance [18]. The CSP also has a tendency to overfit with a large number of channels due to inaccurate estimation of the covariance matrix [19]. A regularized CSP address channel selection problem for common spatial pattern algorithm by reducing coefficient of a few unimportant channels to zero values [20]. To address the problem of optimizing time segment, channel configuration and temporal frequency band-pass filtering for EEG signals jointly, the spectrally weighted common spatial pattern (SWCSP) is calculated on different time segments and channel configurations. These SWCSP features are highly correlated and feature selection techniques are used to select the best time segment and channel configuration for each individual subject. Feature selection techniques can be organized into three categories, depending on how they combine the feature selection search with the construction of the classification model: filter methods, wrapper methods and embedded methods [21]. Filter methods based on mutual information and Fisher ratio, wrapper method based on SVM classifier are used in this study. The main contribution of this paper includes firstly the optimization of spatial spectral pattern, channel and time configuration is arranged in an optimization framework for each subject with few parameters to tune. Secondly, we show that the channel configuration is equivalently important to the BCI performance as temporal or frequency information optimization

5. Feature Selection

Measurement of brain activity, by analysis of EEG generates a huge quantity of data. The EEG signals can be recorded with a sampling frequency which is in range of 100 Hz to 1000 Hz. These signals have to be converted into a fewer number of values, which characterize specific features of signals. For example, these features could be the representation on the scale of power of the EEG signals belonging to various frequency bands. Such features, then are normally further combined into a form of vector, which is called as a feature vector. Feature extraction, thus may be defined as a process of transformation of one or multiple signals to a single feature vector. However, the selection and extraction of good features, improves the performance of the system, which eases the task of further stage of classification algorithm [21]. The designer of BCI system can choose from a variety of many available feature selection algorithms. The performance can be validated, by categorizing the input data into a set of training data, and another into a set of validation data. Feature selection is performed on set of training data, by selection of the discriminative CSP features, which is based on the relation between each of the feature and the respective motor imagery class [8]. Next stage in FBCSP is to select a combination of discriminative CSP features from the subject task. It uses mutual information-based best individual feature algorithm for feature selection. For verifying the efficiency attained in the BCI, cross-validation is executed, wherein input data is randomly categorized into two sets of the training data and the validation data. This method performs operation of feature selection only upon the training data by selecting the discriminative CSP features, which calculates the degree of correct correlation established between each of the occurring feature, with the appropriate intended motor imagery classes.

6. Classification

The classification stage of BCI, which executes a classification algorithm to model the selected CSP features and classify it into MI classes. The purpose of the classification stage is to perform automatical association of a MI class to an appropriate feature vector, which is extracted in the previous stage. This class identifies the type of mental task performed by the BCI user. To accomplish above task of correlating various EEG signals, the classification of the selected features obtained from feature extraction, the following classification algorithms are often used:

Linear Discriminant Analysis (LDA)

The classification in many BCI applications, is executed by using linear discriminant analysis (LDA) as it offers higher efficiency. LDA classifies a recorded set of observations into pre-configured classes. It executes linear transformation of data from that of a high dimensionality space to that in a lower dimensional space. Subsequently, it decides the classification based on set of observations

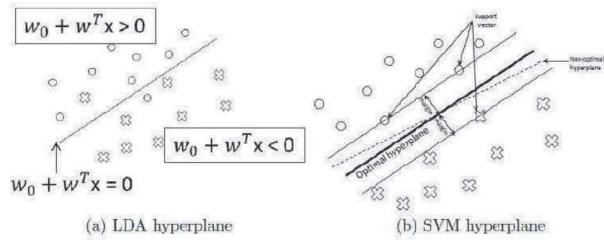


Figure 4. LDA and SVM hyperplane

in a low dimensional space, in conjunction with a defined decision boundary. Hence choice of deciding an appropriate boundary, has a direct influence in the outcome efficiency of classification. The distributions of class, includes some degree of finite variance are included in the selected space during this process. Hence the value of mean and the standard variance of the targeted data is taken into consideration, to establish dimensions of above decision boundary. LDA is used to establish a hyperplane and linear combination of features, which can successfully demarcate multiple classes. LDA is applicable, when there is a normal distribution of raw data, which has equal covariance matrices for each of the classes. The separating hyperplane is established by calculating the value of projection that produces a maximum distance between means of the two classes [22]. It also has to take into account that the interclass variance, is also minimum. However, multiple hyperplanes are required, if the number of classes is more than two, and these hyperplanes are required to be calculated. LDA has a lower computational requirements, which makes it as a proffered choice, whenever the available hardware is small, and for establishing an online BCI systems. This classifier is also preferred due to its ease of use, and offers higher success for classification. Hence, the LDA has been successfully employed in many motor imagery based BCIs [21].

7. Results and discussion

Many researchers have used various combinations of algorithms in different phases of implementation of BCI.

Subject	CSP	SWCSP
1	82.5%	87.5%
2	63.3%	60.8%
3	58.3%	60.0%
4	86.7%	74.2%
5	62.5%	77.5%
6	75.0%	70.8%
7	61.7%	68.3%
8	61.2%	56.9%
9	65.0%	60.8%

The comparison of the percentage age accuracy attained and various factors effecting it is clearly seen from the graph

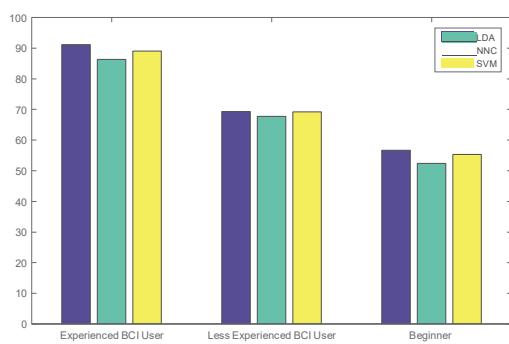


Figure 5. Classification accuracy of LDA, NNC and SVM with CSP method

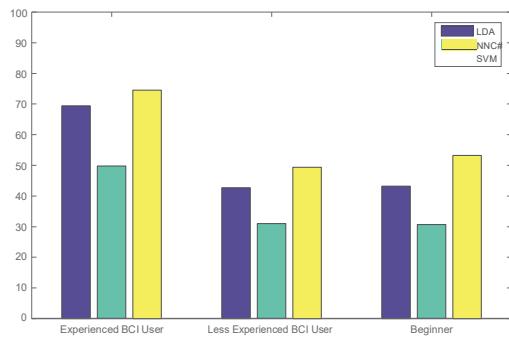


Figure 6. Classification accuracy of LDA, NNC and SVM with AAR method

8. Applications

The BCI has a potential to offer a wide range of applications ranging from entertainment to that of medical rehabilitation. One of the interesting area of BCI research is design and development of peripheral devices that can be controlled by thoughts, without relying on muscle activity. To accomplish this, the basic mechanism of converting thoughts into computerized or robotic action is evolving, thus offering limitless potential uses of this technology. The applications can be broadly categorized into following

Communication

The primary objective of the BCI is to provide alternative to functionality of muscles, to accomplish the intended task, for rehabilitation of people suffering from severe disabilities such as locked-in syndrome. Such patients loose partial or full capability to use their muscles or limbs and may be completely paralyzed and are unable to perform daily chores or even speak. The BCI researchers have proposed and experimented with several alternative methods of assisted communication, which widely vary from simpleton simple binary (yes/no) capabilities [24] to state-of art applications such as TalkAssist [25]. They have also designed virtual keyboards that supports spelling, with partial input

from user. Several alternative approaches to spelling have been proposed, which have been used by a patient suffering from severe neurological disorder, to compose words in a realistic home environment, for effective communication with outside world.

8.2. Environmental Control and Virtual Worlds

The applications of BCI, also include the field of creation of virtual reality, which is used for gaming, entertainment and environmental control. The researchers have described a virtual driving environment, that has evaluated motor imagery signal responses when subjects saw a stop-light [26]. Subsequent works includes design of a virtual apartment which has allowed the residents to interact with virtual people and devices, by establishing alternative communication channels directly with their thoughts process. The researchers have proposed a BCI system that allows users to transcend a complex puzzle of maze, by taking turn decisions at crossroads [27]. Virtual reality has a potential to offer a safe environment and opportunity for training and configuring the BCI controlled interfaces to practically used devices, such as a powered wheelchair for physically challenged people.

9. Conclusion

This paper gives glimpses of various methods used in different phases of BCI, which are used for wide range of applications ranging from medical to entertainment. Researchers are employing different techniques to improve the robustness, efficiency and accuracy, and thus choose the appropriate methods to be used in implementation of different phases of BCI. The researchers in [23] have demonstrated that the accuracy attained in successful classification, hugely depend on choice of the preprocessing as well as the classification method used for implementing different stages of BCI. The signals emanating from scalp are identified by EEG data acquisition, which after fetching is subjected to particular feature extraction and classification methods. The researchers make an appropriate selection of feature extraction, feature selection and the classification methods, which are important building blocks in the development of a robust and efficient EEG-based BCI system. Preprocessing of such signals, prior to above steps, attempts to reduce unwanted artifacts from the EEG signal. The accuracy of overall BCI system, is thus dependent on choice of technique employed in different phases of its implementation. The suitable combination of different extraction techniques with appropriate classification method, provides the desired efficiency and robustness in a BCI system. The improved accuracy attained and lower error rate observed is a guiding objective to choose appropriate methods for implementation of a BCI system. The improving understanding of human brain and its cognitive skills, along with rapid advances in hardware and software technologies, offers huge scope to applications of BCI in wide range of applications in near future.

References

- [1] M. Dai, S. Wang, D. Zheng, R. Na, and S. Zhang, "Domain transfer multiple kernel boosting for classification of eeg motor imagery signals," *IEEE Access*, vol. 7, pp. 49951–49960, 2019.
- [2] S.-A. Park, H.-J. Hwang, J.-H. Lim, J.-H. Choi, H.-K. Jung, and C.-H. Im, "Evaluation of feature extraction methods for eeg-based brain-computer interfaces in terms of robustness to slight changes in electrode locations," *Medical & biological engineering & computing*, vol. 51, no. 5, pp. 571–579, 2013.
- [3] D. Huang, K. Qian, D.-Y. Fei, W. Jia, X. Chen, and O. Bai, "Electroencephalography (eeg)-based brain-computer interface (bci): A 2-d virtual wheelchair control based on event-related desynchronization/synchronization and state control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 379–388, 2012.
- [4] J. Wolpow, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [5] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. Mueller-Putz, et al., "Review of the bci competition iv," *Frontiers in neuroscience*, vol. 6, p. 55, 2012.
- [6] G. Costantini, M. Todisco, D. Casali, M. Carota, G. Saggio, L. Bianchi, M. Abbafati, and L. Quitadamo, "Svm classification of eeg signals for brain computer interface," in *Proceedings of the 2009 conference on Neural Nets WIRN09: Proceedings of the 19th Italian Workshop on Neural Nets, Vietri sul Mare, Salerno, Italy, May 28–30 2009*, pp. 229–233, IOS Press, 2009.
- [7] P. JESSY, "Analysis of eeg signals for eeg-based brain-computer interface," 2009.
- [8] P. Sarma, P. Tripathi, M. P. Sarma, and K. K. Sarma, "Pre-processing and feature extraction techniques for eegbci applications-a review of recent research," *ADBU Journal of Engineering Technology*, vol. 5, no. 1, 2016.
- [9] J. C. Lee and D. S. Tan, "Using a low-cost electroencephalograph for task classification in hci research," in *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pp. 81–90, ACM, 2006.
- [10] S. Siuly and Y. Li, "Improving the separability of motor imagery eeg signals using a cross correlation-based least square support vector machine for brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 526–538, 2012.
- [11] M. H. Alomari, A. Samaha, and K. AlKamha, "Automated classification of l/r hand movement eeg signals using advanced feature extraction and machine learning," *arXiv preprint arXiv:1312.2877*, 2013.
- [12] K. Li, G. Sun, B. Zhang, S. Wu, and G. Wu, "Correlation between forehead eeg and sensorimotor area eeg in motor imagery task," in *2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, pp. 430–435, IEEE, 2009.
- [13] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [14] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213–230, 2015.
- [15] S. G. Mason and G. E. Birch, "A general framework for brain-computer interface design," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 11, no. 1, pp. 70–85, 2003.
- [16] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of Neural engineering*, vol. 4, no. 2, p. R32, 2007.
- [17] S. Selim, M. M. Tantawi, H. A. Shedeed, and A. Badr, "A csp\amba-svm approach for motor imagery bci system," *IEEE Access*, vol. 6, pp. 49192–49208, 2018.
- [18] X. Lei, P. Yang, P. Xu, T.-J. Liu, and D.-Z. Yao, "Common spatial pattern ensemble classifier and its application in brain-computer interface," *Journal of Electronic Science and Technology*, vol. 7, no. 1, pp. 17–21, 2009.
- [19] G. Huang, G. Liu, J. Meng, D. Zhang, and X. Zhu, "Model based generalization analysis of common spatial pattern in brain computer interfaces," *Cognitive neurodynamics*, vol. 4, no. 3, pp. 217–223, 2010.
- [20] J. Farquhar, N. J. Hill, T. N. Lal, and B. Schölkopf, *Regularised CSP for sensor selection in BCI*. na, 2006.
- [21] G. Pfurtscheller and F. L. Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [22] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller, "Spatial filtering and selection of optimized components in four class motor imagery eeg data using independent components analysis," *Pattern recognition letters*, vol. 28, no. 8, pp. 957–964, 2007.
- [23] F. Lee, R. Scherer, R. Leeb, C. Neuper, H. Bischof, and G. Pfurtscheller, "A comparative analysis of multi-class eeg classification for brain computer interface," in *Proceedings of the 10th Computer Vision Winter Workshop*, pp. 195–204, 2005.
- [24] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the wadsworth center," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 222–226, 2000.
- [25] P. R. Kennedy, R. A. Bakay, M. M. Moore, K. Adams, and J. Goldwainhe, "Direct control of a computer from the human central nervous system," *IEEE Transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 198–202, 2000.
- [26] J. D. Bayliss and D. H. Ballard, "A virtual reality testbed for brain-computer interface research," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 188–190, 2000.
- [27] G. E. Birch and S. G. Mason, "Brain-computer interface research at the neil squire foundation," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 193–195, 2000.

Various Components and Concepts Related to Fruit Image enhancement using Fuzzy

Harmandeep Singh Gill

*Principal (officiating) Bibi Sharan Kaur Khalsa college,
Chamkaur Sahib(Roopnagar), Punjab-140112*

Abstract

Fruit image contrast improvement is found one of most challenging filed in digital image processing applications. Enhancement depends on the type of features an image contains. Images capture during improper shutter speed, poor weather and haze conditions may reduce the quality. Due to this, uncertainty added with the image information. Fuzzy logic is the best model to handle ambiguity and vagueness. In this paper, fuzzy logic is involved to handle the uncertainty associated with the fruit images. Empirical study proves that fuzzy can handle uncertainty effectively and obtained results are better as compared to existing competitors.

I. INTRODUCTION TO IMAGE ENHANCEMENT

Image Enhancement is essentially the simplest and most attractive area of digital image processing. Image enhancement is a method used to enhance the overall superiority of the corrupted images that can be attained by using enhancement mechanisms [1]. So that the human eye can effortlessly detect the key features of the pictures. It is used to eliminate the inappropriate artifacts from the pictures like noise or brighten the photograph and it simply to identify main features and then it looks improved [2]. It is an individual area of digital image processing. To create a

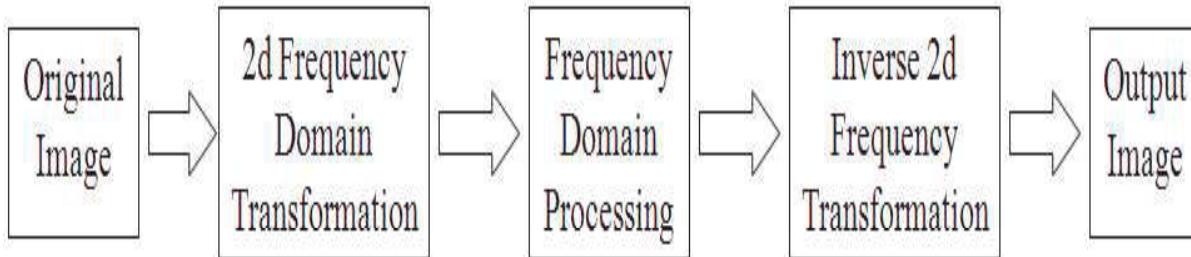


Figure I. Frequency Domain Method Process

Here we have used the word transformation.

B. Transformation

A signal can be converted from time domain into frequency domain using mathematical operators called transforms. There is a variety of conversion that does this. Various of them are given as follows [30].

1. Fourier Series
2. Fourier transformation

3. Laplace transform
4. Z transform

Frequency components: In a frequency domain can be illustrated any picture that are in spatial domain. While what this frequency basically specifies. We will segregate frequency method into two most important mechanisms [30].

- a. High frequency components - High frequency components correspond to edges in an image.

- b. Low frequency components - low frequency components in an image correspond to smooth regions.

Techniques used by spatial and frequency domain methods are:

- a. Point operations: Where every pixel is customized according to a mainly equation that is not based on further pixel values.
- b. Mask operations: Where each pixel is customized according to the values of pixel's neighbors (Using convolution theory).



(a) Before Enhancement

- c. Global operations: When all the pixels value in a picture are taken into concern.

II. IMAGE ENHANCEMENT METHODOLOGIES

In this paper, a fuzzy-based image enhancement scheme is employed to enhance the fruit images. The following figure shows the enhancement results obtained by existing and proposed schemes.

Figure 2 is used to show the difference between the original images i.e. input image and the enhanced image.



(b) After Enhancement

Figure 2. Results of enhancement (a) before enhancement (b) after enhancement (adapted from [34])

III. APPLICATIONS OF IMAGE ENHANCEMENT

1. **Industrial inspection/quality control:** Quality control is a procedure by which operations analysis the superiority of every factors concerned in construction.
2. **Surveillance and security:** Surveillance is the observing of the performance, actions, or further varying data, generally of public for the reason of affecting, managing, directing, or defending them like CCTV cameras.
3. **Face recognition:** It is a computer appliance for mechanically specifying or justifying an individual from a digital photograph or a video frame from a video source.
4. **Gesture recognition:** It is an area in computer science and language technology with the objective of explaining person gestures via mathematical operations. Gestures can begin from any physical motion of body or state but are usually invent from the face or hand.
5. **Medical image analysis:** It is the method, procedure, and ability to make visual depictions of the internal of a body for clinical examination and medical interference.
6. **Autonomous vehicles:** It involves the use of mechatronics, artificial intelligence, and multi-agent system to support a vehicle's operator.
7. **Virtual reality:** It is referred to as immersive multimedia, is a computer-simulated atmosphere that

can replicate bodily existence in places in the real world. Virtual reality can reconstruct sensory experiences, including virtual flavor, aroma, etc.

Contrast Enhancement:

The Visibility or the whole quality of an image can be improved without introducing unrealistic visual façade and/or irrelevant artifacts. The normal global contrast enhancement method basically increases the luminance for bright pixel and decreases the luminance for the dim pixels. So the neighborhood dependent contrast enhancement is desirable to get enough contrast for image enhancement without losing the dynamic range compression. It can be classified into different categories are as follow:

- Linear Contrast Enhancement

Linear contrast enhancement is also known as contrast stretching, the original image can be linearly expanded into a new distribution. The total range of sensitivity of the digital device can be utilized by expanding the original value of an image. This method of enhancement can mostly be used in remote sensing images.

- Non-linear Contrast Enhancement

Through the use of an algorithm, the nonlinear contrast enhancement involves the histogram equalization method. The limitation of nonlinear contrast enhancement is that in which each value of the input image has several values in the output image due to this the original object loses their

accurate brightness. Figure 3 shows an input image and the

output image. It clears the difference between both images.

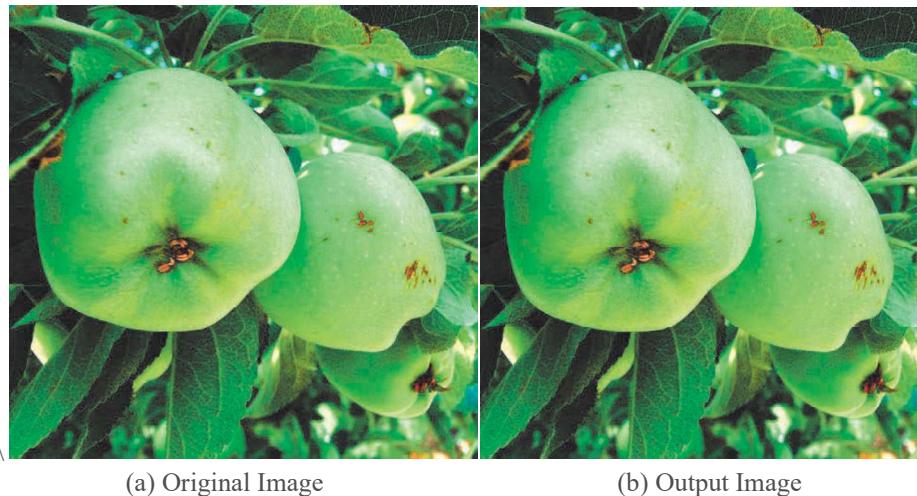


Figure 3 Contrast Enhancement Images (a) Original Image (b) Output Image

Histogram: Histogram is a graphical illustration of distribution of data. It is similar to bar chart. It shows that how many times a particular grey level appears in an image.

Histogram of an image: Histogram of an image, like other histograms also shows frequency. The histogram of an

image shows frequency of pixels value. In photograph histogram, the X-axis displays the gray level intensities and the Y-axis shows the frequency of these intensities.



Figure 4 Original Image

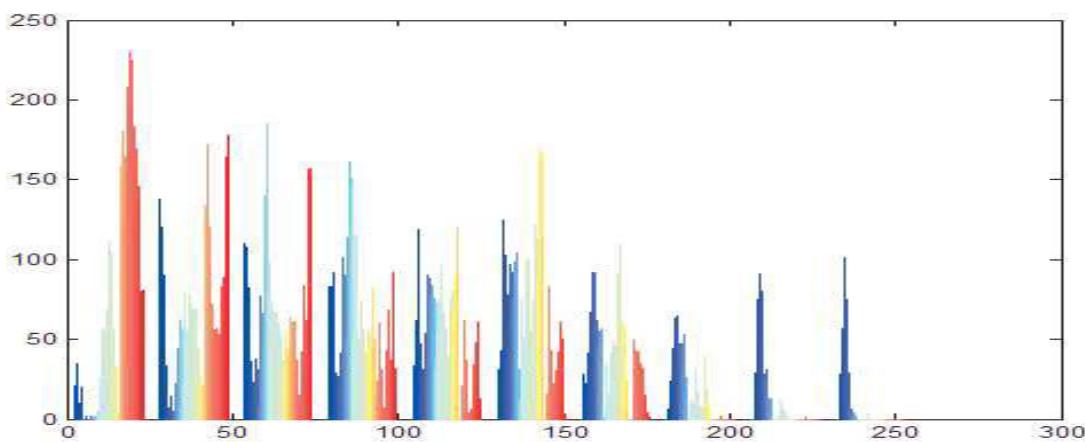


Figure 5 Result of original image

Figure 1.8 shows the histogram of an original image. The X-axis of the histogram shows the range of pixel values. It begins from 0 and stops at 250 with a difference of 50, whereas on the y axis, is the count up of these intensities.

Histogram Equalization: Histogram equalization is the method of image enhancement that is used to enhance the contrast of images. In histogram equalization, it is not compulsory that the contrast of an image will always be raised. Sometimes it shows that it can be not as good as the contrast of an image reduced. Before working with histogram equalization it's necessary to recognize the two main concepts of histogram equalization that are known as PMF (probability mass function) and CDF (cumulative distributive function). First of all estimate the PMF and CDF for all pixels in an image then work further. The transformation $T(r)$ needed to be obtain by using Eq. (2)

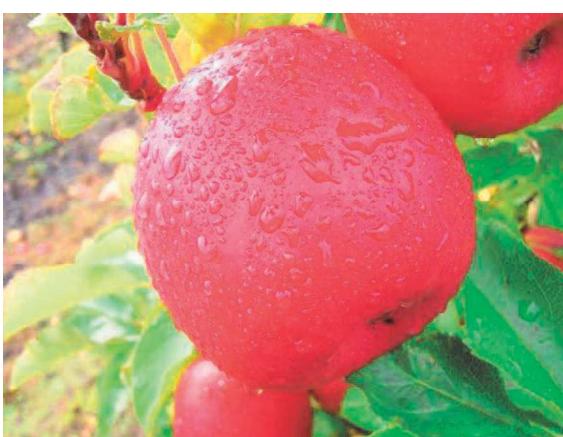
$$s = T(r) = p_r(r)dw = cdf_i \quad (2)$$

Histogram equalization is further divided into two broad categories:

- Local Histogram Equalization: The overall contrast of an image can be improved efficiently.
- Global Histogram Equalization: Based on grey level content of an image, the pixels are modified by transformation function.

Histogram equalization is a point process. In order to obtain a uniform histogram for an image the point process redistributes the image's intensity distribution. Figure 1.9 is used to show the results of histogram equalization. Histogram equalization can be done in three main steps:

- Compute the histogram of an image.
- Calculate the normalized sum of histogram.
- Transform the input image to an output image.



(a) Input Image



(b) Output Image

Figure 6 Results of histogram equalization (a) original image (b) output result of histogram equalization

IV. ADVANTAGE AND DISADVANTAGE OF HISTOGRAM EQUALIZATION

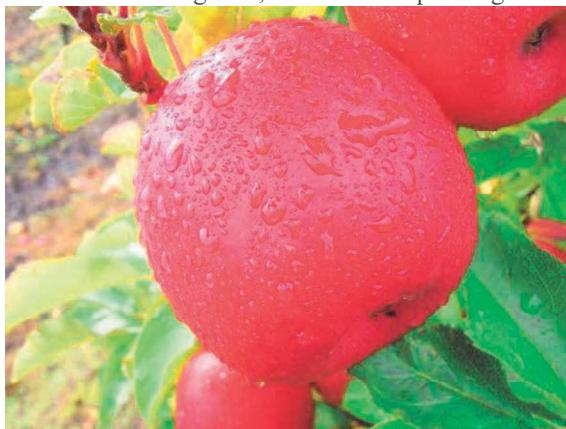
1. This method is finest for visual perception mostly when image has close contrast data.

2. The purpose of this method developed best outcomes for radiographic and thermal images.

The main limitation of this method is the noise amplification when the image has major low intensity area.

V. ADAPTIVE HISTOGRAM EQUALIZATION [AHE]

Adaptive histogram equalization [AHE] is a computer image processing technique used to recover contrast of the pictures. It is a sparkling contrast enhancement for both natural images and medical images and other initially non visual images. It differs from ordinary histogram equalization [HE] in the respect that the adaptive method computes several histograms, each corresponding to a



(a) Original image

distinct section of the image, and uses them to redistribute lightness value of the image. In image fusion process, fusion process may degrade the sharpness of the fused image so to overcome this problem of poor brightness adaptive histogram equalization will be used to enhance the results further. We can say that adaptive histogram equalization will come in action to preserve the brightness of the fused image. The main point of AHE is that in which at smaller scales contrast of an image is enhanced; while at larger scales contrast of an image is reduced or decreased. The advantage of adaptive histogram equalization [AHE] is that it is automatic, reducible, and locally adaptive and usually produces superior images. Figure 1.10 is an example of adaptive histogram equalization [AHE].



(b) Output AHE

Figure 7 The results of adaptive histogram equalization (a) original image (b) output AHE

Advantage of Adaptive Histogram Equalization: This is the most excellent method to apply when global histogram equalization cannot produce best outcomes.

VI. DOMINANT BRIGHTNESS LEVEL ANALYSIS

Dominant Brightness means that is effective or impressive technique for the images. Contrast enhanced images may contain intensity distortion and lose image information in various regions. To overcoming the problems of contrast enhanced images, to decompose the input image into several layers of single dominant brightness levels [3].

The image can be equally decomposed into different levels so that it can be easily handled. After that to execute the discrete wavelet transform on remote sensing images and then calculate the dominant brightness level by using the log-average luminance in the low-low sub band, to use the low frequency luminance components. In view of the fact that high-intensity values are dominant in the bright region, and vice versa, the dominant brightness at the position (x, y) is computed. Dominant Brightness level can be computed by using Eq. (1.3) [3].

$$D(x, y) = \exp\left(\frac{1}{NL} \sum_{(x,y) \in S} \{\log l(x, y) + \varepsilon\}\right) \quad (3)$$

Where S represents a rectangular region encompassing (x, y) , $l(x, y)$ represents the pixel intensity at (x, y) , NL represents the total number of pixels in S , and ε represents a sufficiently small constant that prevents the log function from diverging to negative infinity. Figure 1.11 is used to show the results of dominant brightness level analysis.

The low-intensity layer has the dominant brightness lower than the pre specified low bound. The high intensity layer is determined in the similar manner with the pre specified high bound, and the middle-intensity layer has the dominant brightness in between low and high bounds. The normalized dominant brightness varies from zero to one, and it is practically in the range between 0.5 and 0.6 in most images. For safely including the practical range of dominant brightness, we used 0.4 and 0.7 for the low and high bounds, respectively [3].



Figure 8 The results of dominant brightness level (a) original image (b) output DBLA

VII. ADAPTIVE INTENSITY TRANSFORMATION

1. The function which able to improve the intensity of bright region and the dark region at a same time.
2. Using the dominant brightness level, the knee transfer function, and the gamma adjustment function, the adaptive intensity transfer function is divided into three decomposed layers [3].
3. Then, this is applied for color-preserving high-quality contrast enhancement. By the inverse discrete wavelet transform the result is in the form of enhanced image.

Discrete Wavelet Transform: The wavelet transform decomposes the image into low-low (LL), low-high (LH), high-low (HL), high-high (HH) spatial frequency bands at different scales. The LL band contains the approximation coefficients whereas the other bands contain directional information due to spatial orientation. LH band contains the horizontal detail coefficients. HL band contains the vertical detail coefficients; HH contains the diagonal detail coefficients and also contain the higher absolute values of wavelet coefficients correspond to salient features such as edges or lines.

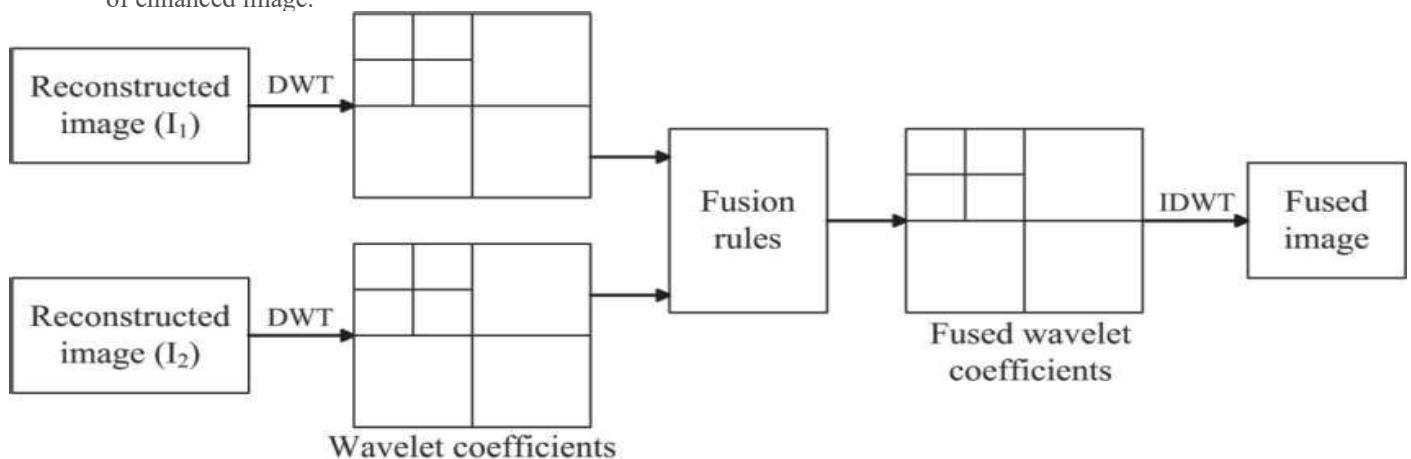


Figure 9 Discrete Wavelet Transform based Image Fusion

The wavelets-based approach performs the following tasks:

1. It is a multi-scale (multi resolution) approach well suited to manage the different image resolutions. It is Useful in a number of image processing applications including the image fusion.
2. The discrete wavelets transform (DWT) allows the image decomposition in different kinds of coefficients preserving the image information.

3. Such coefficients approaching from different images can be appropriately combined to obtain new coefficients so that the information in the original images is collected appropriately.
4. After the coefficients are merged then the final fused image is achieved by applying the inverse discrete wavelets transform (IDWT), where the information in the merged coefficients is also preserved.

VIII. FUZZY IMAGE ENHANCEMENT

The concept of fuzzy was coined by Prof Lotfi A. Zadeh. Fuzzy handles vagueness efficiently but fails to handle uncertainty and ambiguity. Type-II fuzzy controls ambiguity related to data and information effectively [4].

REFERENCES

[1] Singh, Harmandeep, and Baljit Singh Khehra. "Visibility enhancement of color images using Type-II fuzzy membership function." *Modern Physics Letters B* 32.11 (2018): 1850130.

[2] Gill, Harmandeep Singh, and Baljit Singh Khehra. "Efficient image classification technique for weather degraded fruit images." *IET Image Process.* 14.14 (2020): 3463-3470.

[3] Gill, Harmandeep Singh, and Baljit Singh Khehra. "A Novel Type-II Fuzzy based Fruit Image Enhancement Technique Using Gaussian S-shaped and Z-Shaped Membership Functions." *Proceedings of International Conference on Communication and Computational Technologies.* Springer, Singapore, 2021.

[4] Gill, Harmandeep Singh, Baljit Singh Khehra, and Bhupinder Singh Mavi. "Fruit images Visibility enhancement using Type-II Fuzzy." *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS).* IEEE, 2021.

The Role of Artificial Intelligence in Cyber Security

Parag Sharma

Baba Banda Singh Bahadur Engineering College,

Fatehgarh Sahib Sahnewal, Ludhiana

Sharmaparag076@gmail.com

ABSTRACT

Artificial Intelligence(AI)is rapidly developing. It is predicted that over the coming decade people will find AI technology in the home, office, business and general community on a large scale. It will pervade nearly every aspect of our lives. The world is seeing the transformation of AI technology being used by governments and the broader community for security. Today, AI can be found in many devices in our homes that we call "smart" because they operate in a more intelligent way, for example, smartphones, smart watches, robot vacuum cleaners and lawn mowers, self-driving cars, drones, etc.s. AI evolves a more comprehensive study will be needed to better understand whether the developers of the technology have been successful in building adequate safeguards into the systems and platforms to protect personal data from illegal collection and use. Hence it is very much needed to use AI technology for the security purpose.

KEY WORDS

Artificial Intelligence (AI), Cyber Security, General Data Protection Regulation (GDPR), Personal Identifiable Information (PII), Privacy, Trade-off, Deep Learning.

1. INTRODUCTION

TODAY word is not all about living and for the human beings meaning of living is all about development. As we know that "Intelligence" is only the property that distinguishes human from anything else on this planet. The idea of having that Intelligence in man-made machines is quite fascinating although the machines can't have that inherited intelligence.

Yes, it is completely understandable that the AI technology may take over the world from our

imperfect hands because the learning skills and working of the AI tech or we say humanoids are very satisfactory as compared to the humans.

But the satisfying think is that we can control this faze by some precaution techniques. As we can use the AI not only for working purpose but for security purpose as well. Due to these scenarios, we came up with the idea of "Artificial intelligence in cyber security".

Movie References: As this is the time to future, we have also seen many movies showing the impact of AI in near future, for me best reference movies related to this topic is "Avengers: The age of Ultron" and "Robot".

So, for all us there are some threads to privacy which can be done or overcome only by the means of AI. Related content adds are one of the biggest things which is only be done by tracking or using of private data of the user.

Any person in this digital world is not safe from these data snatching.

2.BACKGROUNDS

AI are now becoming essential to information security, as these techs are capable of shifting and analyzing trillions of data sets and tracking down a wide variety of cyber threats — from malware menaces of suspicious that might result in a cyber-attack.

AI can be used to spot cyber threats and possibly malicious activities. Traditional software systems simply cannot keep pace with the sheer number of new malwares created every week, so this is an area AI can really help with. By using productive algorithms, AI operated systems are being trained to detect malware, run pattern recognition, and detect even the minutest behaviors of malware or ransomware attacks before it enters the system.

Business Intelligence with Intelligent Computing

AI allows for great predictive intelligence with natural language processing which provides data on its own by scraping through articles, news, and studies on cyber threats.

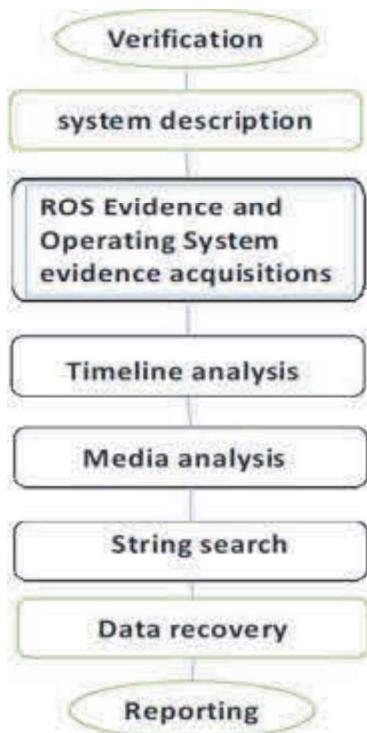


Figure 1: A security framework of ROS forensics procedure

Figure 1 represents an AI security example during a robot forensics investigation, i.e., a robot ROS forensics investigation procedure framework. Based on ROS features, these steps are: 1, Verification systems, (in other words, verify authentication); 2, Find out system description (for example, the vendor, version, serial number and so on); 3, Evidence acquisitions from the ROS and the (targeted system's) operating systems; 4, Timeline analysis; 5, Media analysis; 6, Data string or byte search; 7, Data recovery (and evidence analysis); 8, Reporting forensics investigation results

3 Understanding Basics AI refers to technologies that can understand, learn, and act based on acquired and derived information. Today, AI works in three ways:

- Assisted intelligence, widely available today, improves what people and organizations are already doing.

- Augmented intelligence, emerging today, enables people and organizations to do things they couldn't otherwise do.

Autonomous intelligence, being developed for the future, features machines that act on their own. An example of this will be self-driving vehicles, when they come into widespread use.

AI can be said to possess some degree of human intelligence: a store of domain-specific knowledge; mechanisms to acquire new knowledge; and mechanisms to put that knowledge to use. Machine learning, expert systems, neural networks, and deep learning are all examples or subsets of AI technology today.

- Machine learning uses statistical techniques to give computer systems the ability to “learn” (e.g., progressively improve performance) using data rather than being explicitly programmed. Machine learning works best when aimed at a specific task rather than a wide-ranging mission.
- Expert systems are programs designed to solve problems within specialized domains. By mimicking the thinking of human experts, they solve problems and make decisions using fuzzy rules-based reasoning through carefully curated bodies of knowledge.
- Neural networks use a biologically inspired programming paradigm which enables a computer to learn from observational data. In a neural network, each node assigns a weight to its input representing how correct or incorrect it is relative to the operation being performed. The final output is then determined by the sum of such weights.
- Deep learning is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Today, image recognition via deep learning is often better than humans, with a variety of applications such as autonomous vehicles, scan analyses, and medical diagnoses

4 Role of Cybersecurity

Nowadays the individual is more frequently falling into cyber-attacks due to the nature of risks in the cyberspace. Pathways are constructed through malignant and offensive activities, which give

unauthorized access to hackers and trackers on computer systems or networks. These activities are called cyber threats. Predators work on the bugs and faults in the system or network to establish these pathways. There are cyber threats like ransomware, virus, worms, Trojans, spyware/adware, attack vectors, social engineering, Man in The Middle (MITM) and many more (Panimalar, A., Giri, P.U. & Khan, S., 2018). Some valuable assets and confidential data which are under their authority and when an outsider gets access to those assets and data, they can do extreme harms. Taking cyberspace into consideration, these accesses without the consent of the owner can be the results of one or more cyber threats. Here cybersecurity entered into game. It ensures the availability and integrity of your system or network and helps it to work efficiently without compromising with the security of your privacy.

5. Principles of Cyber Security

To ensure the three important goals of cybersecurity, i.e., availability, confidentiality and integrity, some simple but effective principles can be followed.

To ensure the three important goals of cybersecurity, i.e., availability, confidentiality and integrity, some simple but effective principles can be followed.

1. Focus on Prior Systems: Stabilizing the degree of availability, confidentiality, and integrity of resources comes under biggest challenges and hence it can only be achieved by focusing on the vital systems and providing the best cyber shield to it whereas other methods are applied for the protection of less-prior systems.

2. Different Users, Different Level of Accessibility: What data should be accessible by whom should be based on what type of authority or position he is, and not every individual should get access to all the data and information. This means minimum particular responsibility. Hence the change in responsibility is directly proportional to change in privileges.

3. Provision of Independent Defense (Protocols): Several supervisions for a single job are a far better idea than a single one. It highly reduces the risk of successful

cyber-attacks and the basic principle is increasing the work of the attacker as he has to perform numerous tasks to break through different fire wall and several protections layers.

4. Backups: Precaution is always better than the cure so it is better to take the backup of the system of the data at any point of the cyber threat

6. Modern Challenges in Cybersecurity

“Cyber Security is a shared responsibility, and it boils down to this: In Cyber Security the more systems we secure, the more secure we all are (Johnson, J., 2014)”.

The more we talk about these problems we are seems to be more keen to solve these ones. In today technical world, where we need more automations, we are helpless to supervise everything or every operation done by any means of AI or computer programs we need badly

need a highly satisfactory supervision on all the function of it, so it is much better for humans to program an AI not just for labor also for the supervising to the other AI operated Robots.

But the thing in these situations is that only one AI is sufficient for multi-tasking like a same AI system can control different working systems at no time and there is possibly no chance of a mistake as operation is done on behalf of the operation by the programmer

But as we know that only an AI can face a security threat from an AI, SO it is good for the humans to use AI against AI in the field of cyber security.

As AI matures and moves increasingly into the cybersecurity space, companies will need to guard against the potential downsides of this exciting new technology:

- Machine learning and artificial intelligence can help us to guard against cyber-attacks, but hackers can target the data they train on and the warning flags they look for
- Hackers can also use AI to break through defenses that can also perform changes in its structure to avoid detection
- AI is man-made and also not so perfect; it delivers the incomplete results without sufficient amount of data.
- If it goes undetected, they may also harm the AI performing the security and backing up the data. In that condition it is very difficult for us to recover that data.

7.CONCLUSION

In recent years, AI has emerged as a required technology for augmenting the efforts of human information security teams. Since humans can no longer scale to adequately protect the dynamic enterprise attack surface, AI provides much needed analysis and threat identification that can be acted upon by cybersecurity professionals to reduce breach risk and improve security posture. In security, AI can identify and prioritize risk, instantly spot any malware on a network, guide incident response, and detect intrusions before they start.

AI allows cybersecurity teams to form powerful human-machine partnerships that push the boundaries of our knowledge, enrich our lives, and drive cybersecurity in a way that seems greater than the sum of its parts.

8. References

1.“The Role of Artificial Intelligence in Cyber Security”by Kirti Raj Battelle Harsh Shrivastava Neha Kumari RJIT, India RJIT, India RJIT, India

2. Google
article:<https://www.balbix.com/insights/artificial-intelligence-in-cybersecurity/>

3. Article : “Artificial Intelligence in Cyber Security”
by Rammanohar Das and Raghav Sandhane*
Symbiosis Centre for Information
Technology, Symbiosis International(Deemed

University), Pune, Maharashtra, India Email:
*raghav@scit.edu

4. “Artificial Intelligence Cyber Security Strategy” by
Xiaohua Feng School of Computer Science & Tech.
University of Bedfordshire,
Luton, UK

Xiaohua.feng@beds.ac.uk Tel.
+441234400400

Yunzhong Feng Hebei Normal University
Shijiazhuang, Hebei,
P.R. China fyz02817@sina.com

Edward Swarat Dawam School of CST
University of Bedfordshire Luton, Bedfordshire, UK

Edward.Dawam@study.beds.ac.uk

7. “A decade of research on patterns and architectures for
IoT security” by Tanusan Rajmohan

8. “Artificial Intelligence and Law” Book by Robert
Walters, Marko Novak
https://link.springer.com/chapter/10.1007/978-981-16-1665-5_3

9. “Security with AI And Machine Learning” by Laurent
Gill & Allan Liska

10. “Hands-On Artificial Intelligence for
Cybersecurity” by Alessandro Parisi

A Systematic Approach for Speech to Text Translation and Summarization

*Aswin.P¹,
MTech. Scholar,

Department of C.S.T., Central University of Punjab-151401
aswinp2610@gmail.com

Dr.Satwinder Singh²
Associate Professor, Department of C.S.T., Central University of Punjab-151401
²satwindercse@gmail.com

Dilshad K³
PhD. Scholar, Department of C.S.T., Central University of Punjab-151401
dilshadkaur@outlook.com

Abstract

Speech is considered the most usual mode of interaction between humans. Nowadays, speech is playing a significant role in the field of speech recognition. Speech recognition is seen as a method that is used for extracting meaningful information from speech signals. The technologies related to languages can solve many problems like it can encourage the people who speak a different kind of languages for communication and data transfer. The survey and study did not find an efficient technology to translate speech to text in multiple languages without human interventions. The speech recognition technology helps the machine to track the human voice and accordingly understand languages spoken by human beings. The current study has proposed the framework for understanding speech in native languages. Mann Ki Baat (Indian Prime Minister monthly speech) dataset is used to test the framework. Proposed framework utilizes the recorded speech as input and output a text in Hindi and summarizes it in the text format. The framework extracts the speech's core message with a Similarity Score precision value of 0.0355 for the translated text. On comparing the translated text with actual data, the maximum precision value of similarity score achieved by the frame work is 0.1905.

Keywords: Automatic Speech Recognition (ASR), speech-to-text translation system (STT), multi-lingual, bilingual, speech-to-text translation, text summarization, similarity score

1. Introduction

Speech is known to be the most common form of human contact. Speech synthesis is an integral aspect of signal processing science [1]. Expression has become increasingly important in the area of speech recognition. Speech recognition is described as a technique for extracting useful information from speech signals. Speech recognition requires a variety of data, such as the speaker's records, linguistic data, and so on. This type of knowledge has prompted the advancement of technologies that automatically process speech, such as speech

amplification, speech synthesis, speech compression, and so on. The work on topic modelling is growing by the day, and scholars have published numerous topic modelling articles and used them in various fields, including software engineering, political science, and medicine [66]. Speech recognition requires a variety of data, such as the speaker's records, linguistic data, and so on. There are two forms of speech recognition: speaker-dependent and speaker-independent.

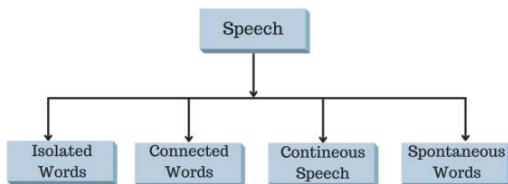
Furthermore, speech recognition is mainly designed for foreign languages. Language-related technology can handle all types of problems, such as encouraging people to speak various languages. It also addresses many topics, such as encouraging people who speak a variety of languages to communicate and share data [49].

When listening to someone's speech in everyday life, a person may be required to write a report on it and outline the main points. If the speech is lengthy, this can be a time-consuming process. As a result, the machine was intended to create a Speech-To-Text conversion device. In English, the expression must be oriented, and the language will be used for content interpretation after the translation process. The goal of this approach is to develop and introduce a Speech-To-Text Translation method for the required languages and summarize speech in textual format for text analysis.

Sorts of Speech

The utterances in the signal of speech, defined in various classes as follows, make the types of speech understandable (Fig. 1.).

- **Isolated Words:** The words would be constant, as even the title indicates. In this form of voice, humans speak naturally [52].
- **Spontaneous Words:** This is the kind of speech that is spontaneous and used in everyday situations [53].
- **Connected Words:** This sort of speech commonly contains two or three words. Between every word, a connective will be included [51].
- **Continuous Speech:** The words would be constant, as even the title indicates. In this form of voice, humans speak naturally [52].

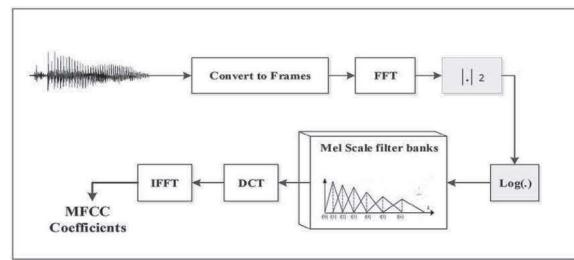
**Fig. 1 Different sorts of speech**

The approach aims to build a system that can translate speech from English to Hindi. It extracts vital sentences and does summarization in English and Hindi to resolve to translate speech to text. This approach comes forward with a technology that can perform tasks parallelly and make the process faster. This research paper is divided into different sections. The first section gives us a brief introduction to the speech and its related terms. The Speech to Text translation and Text Analysis techniques used by various researchers are mentioned in the literature survey section. The third section explains the research methodology used for the proposed framework. Next, the paper proceeds towards the results and discussion area, giving the facts and figures involved for testing the proposed framework. Finally, the fifth section gives the conclusion and future work.

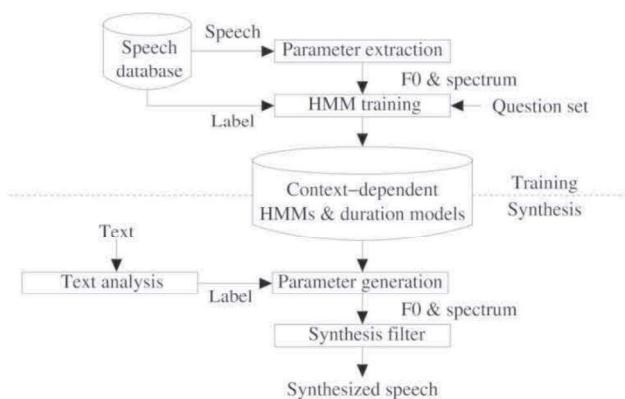
2. Related Works

Many studies on speech-to-text conversion and automated speech recognition have been published. Here, we concentrate on compiling the most recent works in Speech Translation(ST). Various methodologies/models and structures used by various researchers are described below.

In [5] suggested a system for speech recognition technology that is fixed across different languages. Features such as acoustic multiple language modelling, dialect recognition, and pronunciation modelling help accomplish this aim by utilizing Mel-Frequency-Cepstral (MFC), a method to represent a speech and power continuum. The frequency groups in the MFC were similarly separated based on the mel scale. Fig. 2 illustrates the MFCCs from the audio recorded signals [61]. Here, the speech input is converted to frames, and the output MFCC coefficients are utilized for further processing.

**Fig. 2 MFCCs from the audio recording signals [61]**

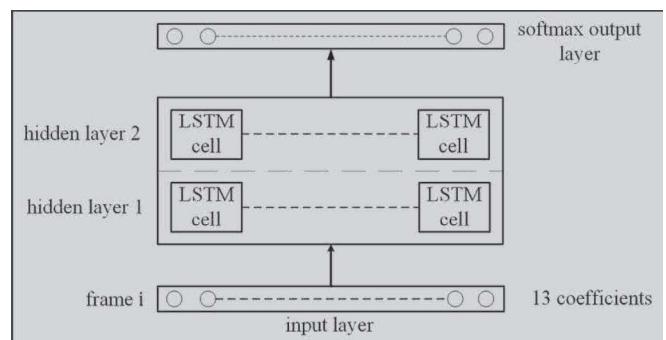
Later, M. Yuan and C. Zou [1] used MFCCs for multilingual speech to text translation. The Speech to Text system takes a human speech expression as input and outputs a sequence of sentences. The framework's purpose was to extract, explain, and perceive speech data. The Mel-Frequency-Cepstral-Coefficients (MFCCs) and the Minimum Distance Classifier and Support-Vector-Machines were used to complete this project (SVMs). Then X. You [3] introduced a system that collects, separates, and reads online newspapers in various languages: Malayalam, Hindi, Tamil, and English. The effective use of technology as an aid to the blind or handicapped, the elderly, and even the uneducated. They utilized MBROLA, which is a diphone-link-based speech synthesizer. The input is phonemes, and the output is a 16-bit linear speech sample at the sampling rate of the diphone server. In the construction of the polyglot proposed by[4], the Hidden Markov model was used. Synthesizer was able to translate the content knowledge into any of the appropriate languages by using the model. Test output was assessed using ABX Listening, ranging from 73 to 86 per cent for the analysis[4]. Then [48] looked into the problem of English speech recognition to find the best word grouping from the given segment of an English voice. HMM-model demonstrates spectral vector sequences of variable time using simple, efficient technology. Fig. 3 describes the speech synthesis using Hidden Markov Model [63]. From the figure, one will understand the speech synthesis process and its various stages in Hidden Markov Model.

**Fig. 3 Speech Synthesis Using HMM [63]**

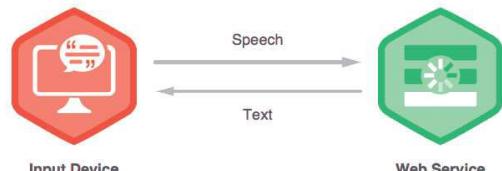
[14] designed a system that looks through web indexes. This application was intended to perceive the speech by an individual and enhanced information that best showed the significance of the speech structure. Plan of speech program with utilizing raspberry pi module and on Raspbian stage is utilized for voice-based Google search.

Later [18] improved the speech synthesizer, allowing it to produce speech from several speakers. The trial results show that the A2W platform, when combined with a multi-speaker system, provided a significant improvement over the baseline. The speech and text encoders were created using six stacked BLSTM layers fitted using the 1024 invisible measurements. Then [26] suggested another end-to-end technology for speech translation with two decoders intended to deal with the deeper relations between the source language: sound and the target language: text. The architecture has shown the impact of incorporating an extra layer in the centre for the pre-training system.

In speech recognition, LSTM is the machine learning model which is used. Since the defects disappeared when backpropagation was present in deep layers, it advanced in the problems of neural-network architecture. In Fig. 4, the Long-Short Term Memory recurrent neural network representation is illustrated. The input layer, hidden layers, and the output layer are shown in the figure, respectively.

**Fig. 4 LSTM RNN for speech recognition [55]**

Then in [45], [33], and [38] introduced approaches utilizing Web Speech API, which makes it easier to integrate sound into web applications. Speech Synthesis and Recognition are the two sections of the Web Speech API. Many researchers in the field of speech-to-text translation utilise this API. The Mammo Class module will take a sound source, convert it into composing sentences, and then determine the variable qualities. The Web Speech API was used in this interface to provide a connection between the written report and the audio. After the results of spoken mammography reports were eliminated, it was discovered that identical influences were removed for both types of information: typed in and guided text. Fig. 5 describes a simple overview of how the speech API is working. When the web service receives speech input from a device, it converts it to text online and sends it back to the input device.

**Fig. 5 Way of Web speech API working [60]**

Later [39] developed a method that uses Deep Recurrent Neural Networks to convert speech to text in the Bangla language. They proposed, to cut down the preparation time, the fragmented language format is valid. Deep recurrent neural networks are also used for speech recognition.

The research conducted at the LIMSI laboratory [6] defined the spoken language, which dealt with various languages. The corpora development and enhancement are a focus of the Spoken Language Processing team. This spoken language community is interested in ARPA-sponsored tests, the LE-SQALE scheme, and the AUPELF-UREF method for linguistic data collection. It

has been shown that they are involved in voice processing for reading speech. A similar recognizer may be customized to work in a variety of languages. Later [46] presented a method that portrayed sequence architecture for speech translation. The words spoken in the source language are directly translated into the target language.

Sphinx is a term used to refer to the speech recognition frameworks developed at Carnegie Mellon University (CMU) as a whole [19]. These include several acoustic model trainer series and Sphinx structures for proper signal training and monitoring. The numerous collections of Kannada sentences ranging in length from four to ten words were examined in depth. It also proposed a Voice to Text system for medical care associations [34]. Counsellors, including NGOs, used it to document discussions during assessments.

The KALMAN Filter Algorithm is a method for estimating variables that are undefined at the time of measurement. Kalman channels are accessible and

functional in a low-power computing environment. Then created a continuous speech recognition system that was tested in a noisy environment regularly [22]. To increase the capability of this real-time speech recognition system, they used the design of a bidirectional nonstationary Kalman channel. The system immediately transforms the spoken words when they have been articulated. This method presented a more straightforward and more powerful speech recognition platform than the HMM-based speech recognition framework.

Speech recognition testing was carried out in [55] with sound recording instruments held safely. In comparison to overhead microphones, shotgun recording mics were traditionally thought to be the best choice. In signal processing, beamforming is used in sensors to exhibit directional sign transmission. Any of the principles were investigated in addition to improving efficiency for speech recognition at a distance. Table. 1 shows a comparison of the Language pairs and languages in similar research works and their tools and technologies.

Table 1 Some Related Works in Speech to Text Translation

Paper Details	Language sets used	Technologies and Tools Used
Pantazoglou, Kladis and Papadakis [28]	English-Spanish	Bidirectional LSTM encoder
Noel [29]	Turkish	Mel Frequency Cepstral Coefficients (MFCC) LSTM encoder
Nicolao, Drioli and Cosi [30]	English-German	ESnet tool
Nasib, Kabir, Ahmed and Uddin [31]	Bengali	CMUSphinx framework
Nahid, Purkaystha and Islam [33]	Bahasa Indonesia	Google Cloud Speech API, CMUSphinx API
Murray, Kinnison, Nguyen, Scheirer and Chiang [34]	multi-language	CMUSphinx
Lero, Exton and Gear [35]	multi-language	BEAMFORMING ALGORITHMS
Lee and Kim [36]	Sinhala	HMM algorithm
Lamel, Adda-Deckes, Gauvain and Adda [37]	Multi-lingual	Fast Fourier Transform, Hidden Markov Model
Lakdawala, Khan, Khan, Tomar, Gupta and Shaikh [39]	Bangla	TensorFlow
Kusumah, Hartanto and Hidayat [40]	Chinese-English	LSTM, HMM, MFCC
Kunchukuttan [41]	Bahasa Indonesia-Javanese	Fast Fourier Transform, MFCC
Kano, Sakti and Nakamura [43]	English-German, English-French	LSTM
Stephen, Anjali and Bhadran [16]	Hindi, Manipuri and Urdu	GOLDWAVE software tool
You X. [3]	Malayalam, Hindi, Tamil and English	MBROLA Based TTS Engine

3. Research Methodology

Speech-to-text technology effectively transcribes audio content to actual words in a document or any other output device. Anyone who needs to generate written content without

much manual typing will find this speech recognition technology instrumental. It is also handy for individuals who cannot use a keyboard because of a disability. Fig. 6 represents the proposed system framework.

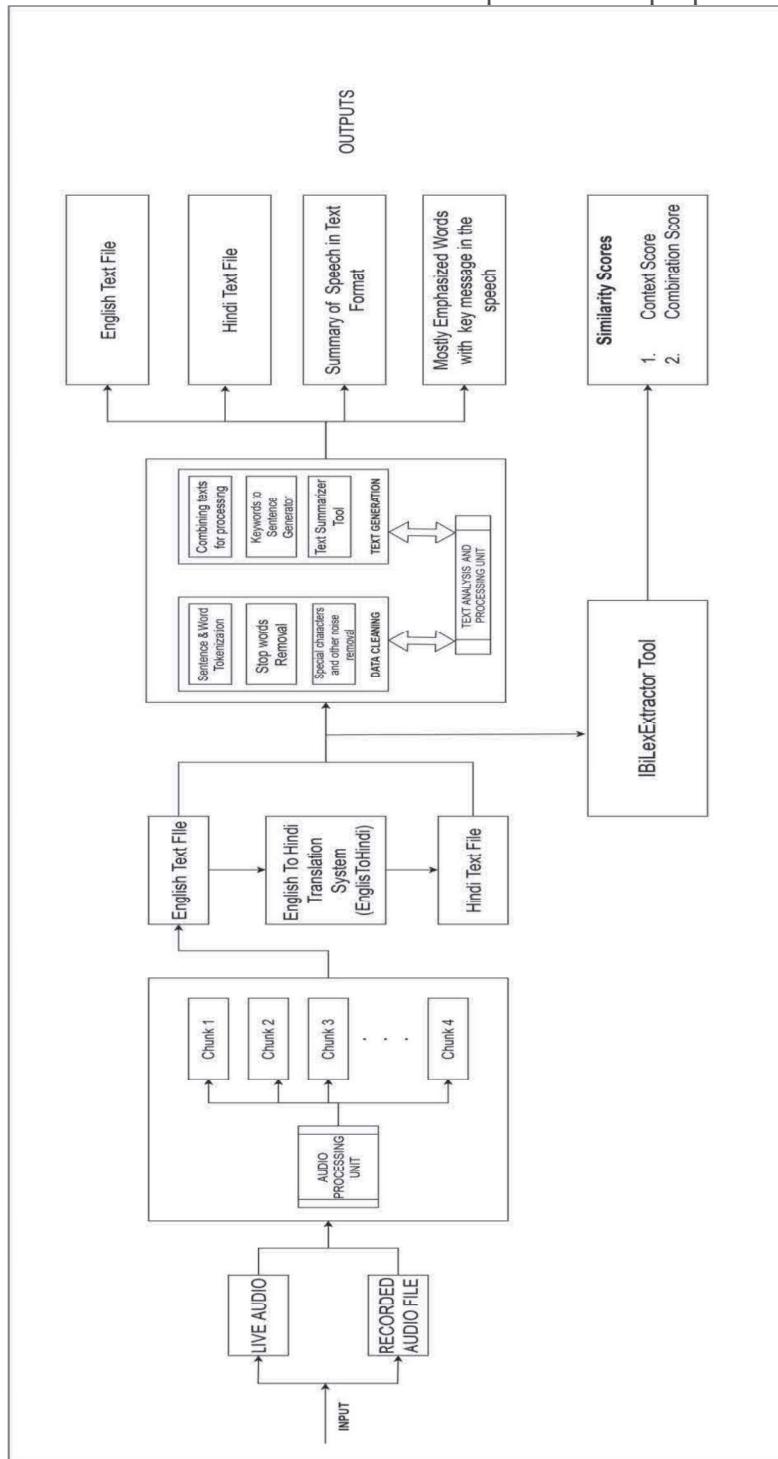


Fig. 6 Proposed Framework of Speech to Text Translation and Summarization

3.1. Overview and Algorithm

In the current work, the approach introduced a method for Translating speech data to another language in textual

form and find its vital message with a summary. EnglishToHindi Python library used for translation and t5-base transformers used for the key message extraction and summarization. The IBILexextracor tool is used to find the similarities between translated Hindi text and English text. The steps of the methodology that were used to perform speech to text translation and summarization were as follows, and Fig. 6 also illustrates the algorithmic steps included in the approach:

- 1) The study used a speech dataset as "Mann ki Baat" in English, and it is publicly available.
 - 2) Converted the audio file to .wav format for better performance.
 - 3) Divided the audio file into several chunks to differentiate the sentences.
 - 4) Translated each sentence to Hindi language Using the EnglishToHindi package in Python.
 - 5) Data cleaning of text in Hindi and English languages using python Libraries
 - 6) Finding the similarity scores based on topic and Context using the iBiLexExtractor tool.
 - 7) Python-based code to Find the key message in the speech.
Developed a method to summarise the text using the transformers.

3.2. Dataset

The proposed framework can accept two kinds of inputs. Either Live Audio or Already Recorded Audio. The dataset entitled "Mann ki Baat" available in the public domain, will be used as a recorded version in the current study. In this program of Man Ki Baat, every month, Indian Prime Minister Narendra Modi delivers a speech to the citizen of India. This dataset is available in several Indian languages. This study will use the English language recorded audio version as input data and processed the system in two phases the Speech processing phase and the text Analysis phase.

3.2.1 Speech Processing

The Audio Processing Unit is responsible for speech processing, as shown in Fig. 6. The Audio file is initially converted to the ".wav" format for ease of usage and support in Python. The converted audio file is then sliced into several chunk files. The chunk file is created when 1 second of silence occurs in audio. At the same time as chunk formation, the chunks of audio files are converted to textual format and saved in a text file. The conversion of Audio format and slicing into chunks are done using the pydub python library. Python speech recognition library used for speech to text conversion (Fig 7.) and the EnglishtoHindi python library for English to Hindi translation.

my dear countrymen Bains, on the finer festival of Deepavali greetings to all of you, we are familiar with the lines of the sword Shapoor to radical Naam sarmi and grandson Pita. Rehnia says it is my that washes in your health and prosperity into our lives this time in negative phone bringing in positive always playing respectful, what could be a better way than to have a positive phone in our life. The Indian diaspora in the world is growing day by day. The Indian culture and Dussehra is celebrated across many country, it is not Indian government organisations or hardly celebrate dussehra with Aryan PTA, but in the diaspora across the world Indian people are more active, more involved in their own activities. In India, the number of festivals presence is limited to specificities in the realm of festive occasions. It would be our endeavour to ensure that we should welcome include people of other states and other countries. When we celebrate Holi in India, we do not care about the ethnicity of the people, we are open to all. The Indian diaspora consists of a large number of interesting people from other countries. The Indian diaspora has a significant role to play in promoting festival tourism in India. My dear countrymen in the previous episode of Han Ki Baat we had decided to do something different this to wall.

Fig. 7 English text file created from speech input

3.2.2 Text Analysis Module

In current work, the `EnglishtoHindi` python library is utilized for translating English text to the appropriate Hindi text. This library utilizes the web scrapping method for translation. The sample output is shown in Fig 8.

Fig. 8 Hindi text file which is translated from the English data

3.3 Pre-processing and Data Cleaning

In the pre-processing stage, data was cleaned for better textual analysis. Stop-words, special characters, and parsing letters, punctuation marks, and digits are eliminated in this cleaning process. To determine the most emphasized terms, generate a key message, and summarize the text, the sentences and words are tokenized using the NLTK library in Python. Fig 9 to Fig 11 shows the output after pre-processing of data to perform the textual analysis.

['my dear countrymen Naina.', 'on the finest festival of Deepavali greetings to all of you.',

Fig. 9 Sentence Tokenization

['my', 'dear', 'countrymen', 'Naina', '.', 'on', 'the', 'finest', 'festival', 'of',

Fig. 10 Word Tokenization

['dear', 'countrymen', 'Naina', '.', 'finest', 'festival', 'Deepavali', 'greetings',

Fig. 11 Stop words Removal

3.4 Similarity Scores

In the next step, one needs to calculate the similarity function. It is a real-valued function that determines how similar two objects are [65]. In this approach, the IBiLexExtractor tool is used to find the Similarity Scores based on Topic and Context modelling. The IBiLexExtractor tool has cleaned English and Hindi text files to produce the Context score and Combination score based on Topic and Context [64] based modeling approach. The current study result table of similarities will be measured as Mean Reciprocal Rank (MRR). The Bilingual lexicon extraction system working is depicted in Fig. 12.

$$\text{Similarity}(A, B) = \frac{1}{1 + \text{Distance}(A, B)}$$

$$\text{MRR} = \sum \frac{1}{\text{Rank}}$$

The rank is the position of the relevant document, and the Mean Reciprocal Rank (MRR) is the harmonic mean of the rank. Topic score calculates a single topic's score by comparing the semantic similarity of high-scoring words in the topic. These metrics help in distinguishing between subjects that are semantically interpretable and topics that are statistical inference artefacts. The context score measures the similarity of contextual distributions between languages. Combining these two methods may measure the distributional similarity of two terms using both topical and contextual knowledge, making bilingual lexicon extraction more trustworthy and accurate than merely using one knowledge source and which is given by combination score.

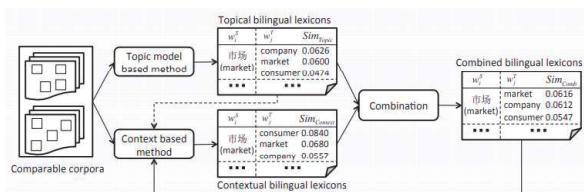


Fig. 12 Bilingual lexicon extraction system [56]

3.5 Key Message Generation

The key message in the speech is considered the core sentence, representing the whole speech. This framework tried to extract some mainly occurred words in the speech and utilizing those words as a key message with the help of the keytotext module of the Python library. As per the library working, the keywords are extracted and processed by utilizing pipelines [65]. Fig. 13 demonstrates the generation of the key message with the library keytotext Python module.

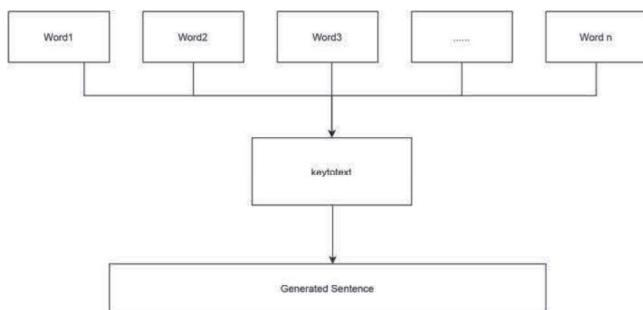


Fig. 13 Sentence generation using keytotext module

3.6 Summarization

The significance of the summary arises where the speech is too long. Also, nobody is ready to listen to long speeches or read long paragraphs in the fast-moving world. Abstractive summarization is an NLP task that aims to provide a succinct description of a source text. Abstractive summarization, unlike extractive summarization, does not

merely copy essential phrases from the source text but can also generate new specific phrases, known as paraphrasing. Abstractive summarization has a wide variety of applications in various fields, including books and literature, science and R & D, financial studies, and the study of legal documents. Here, the approach used T5 abstractive summarize for the task. As Fig. 14 shows, the converted text file from the speech is used as input and is given to the T5-base-abstractive-summarizer model. One abstractive summary is given as the output from the model.

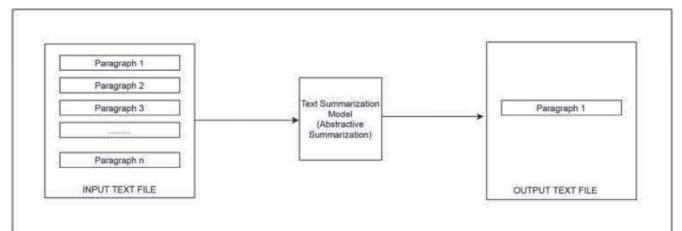


Fig. 124 Representation of How Summarization Happens

4 Results and Discussions

The proposed framework is implemented using Python 3.9.0(64-bit). For the experiment, the Mann ki Baat dataset (English Version) is used. The framework was tested with this dataset. In the first step, speech to text translation was performed, then English to Hindi translation was done, next, most emphasized words were extracted to find a key message. At last, summarized the speech data to measures and compared the various similarity score. The similarity scores are calculated based on topic, Context and combined approach.

4.2 Comparative Analysis

The experiment was carried out using a pair of text data in English, and Hindi derived from speech (Man Ki Baat) data to examine the performance of the proposed framework on multiple similarity measures. The IBIlexExtractor tool is used to do the operation. Similarities are extracted based on the Topic, Context and Combined Score. In similarity measure, Precision for single data, MRR value, and Precision for all are calculated. Table 2 and Fig. 17 shows the similarities obtained with 100 iterations. As per the experimental results, Precision for Topic and Combined approach is similar, and the Context-based score in Precision@1, MRR, and Precision@all is higher than the Topic and Combined based approaches. These results represent that framework works well with Context-based modelling.

Table 2 Similarity Score for Converted Hindi and English Text

	Topic score	Context score	Combination score
--	-------------	---------------	-------------------

Precision @1	0.0018	0.0053	0.0018
MRR	0.0055	0.0087	0.0054
Precision @all	0.0355	0.0355	0.0355

Table 3 Similarity Score for Mann ki Baat Dataset Hindi and English Text

	Topic score	Context score	Combination score
Precision @1	0.0317	0.0159	0.0317
MRR	0.0482	0.0310	0.0482
Precision @all	0.1905	0.1905	0.1905

(Actual dataset)

Table 2 and Table 3 illustrates the comparison between the similarity scores basis on Topic modelling, Context modelling and Combined Score before processing and after processing with the dataset. As per the findings, the similarity score is higher when we find the similarity based on the Topic and Context of the already available text dataset of Mann Ki Baat. This variation is found because the data after processing have some noise. At the same time, the already available dataset or the actual dataset text data is clean and noise-free, which can be used to compare with machine-translated data. Were, already available dataset, or the actual dataset is the Mann ki Baat dataset by PM Narendra Modi

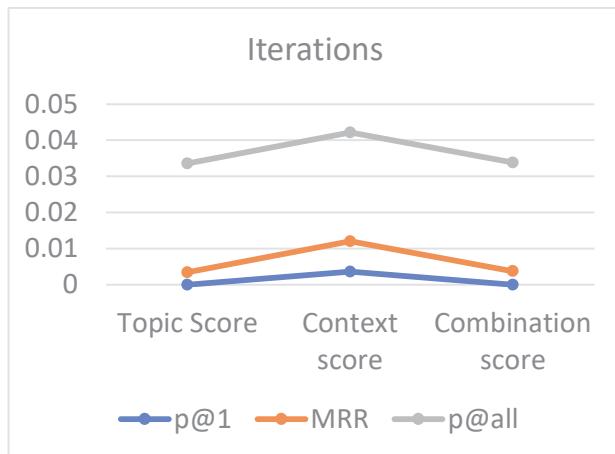


Fig. 135 Similarity Scores

The speech to text translation and summarization approach converted the speech in the English Language to Text, as shown in Fig. 7. Then translated it into the required language (Hindi), shown in Fig. 8. The proposed framework found the mostly emphasized words in the

speech, as shown in Fig. 16. The above results are then analyzed for similarity score based on Topic modelling, Context Modelling and Combined modelling score.

As per the above steps, the similarity scores for the Context and topic are illustrated in tabular form in Table 2 and Table 3. The same was represented graphically in Fig. 15 and Fig. 17. The next step of message extraction with the Python library was performed, and sample message extraction was shown in fig 18. Also, the framework Summarized the speech with T5-base-abstractive-summarizer and sample output of the speech summarization is shown in Fig. 18.

[('I', 77), ('India', 36), ('people', 36), ('country', 32), ('time', 30)]

Fig. 146 Mostly Emphasised Words in speech

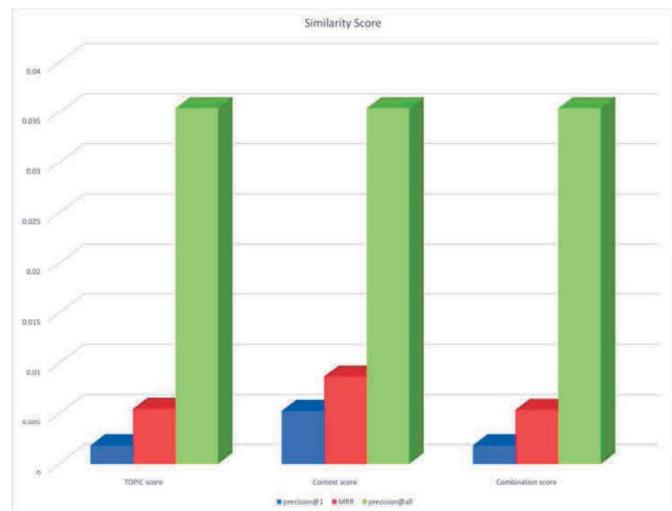


Fig. 157 Similarity Score basis of Topic and Context

'The country of origin is india.'

Fig. 16 Key Message in the speech

The day is celebrated as national unity day in many around the world . It is a time when people from all life come together to make a difference . This is t n the nation is united in a new way of life . The w ing calamities like the earthquake in pakistan in 2

Fig. 17 Summary of Speech

4.3 Contribution and application of the proposed framework

The technique of breaking into modules is the most significant reason for the framework to get a better result quickly. This framework is divided into Input Unit, Audio/Speech Processing Unit, Translation Module, Text Analysis and Processing unit and Output Unit. Each processing modules work parallelly to produce results and

combine them as fast as possible. This system makes it fruitful for the uneducated rural groups or the academically poor people to understand the speech in native languages. This approach utilizes the speech as input, gives output as text in the Hindi language, and summarizes it in text format. With the help of summarization, the framework is capable of extracting the core message in the speech. The proposed framework will have the following application in real life:

- **News and media:** The live speech can be recorded, or live speech instantly been used to produce news articles with this framework.
- **Education Field:** The education field can be said to be utilizing speech and text at a higher rate. The system can be used to collect the speech sample of the person and make a summary of the speech. So, a lecture, conference, seminars, or some other related activities happening on campus can be recorded and provide a brief.
- **For Differently abled Persons:** In the world, many persons are suffering from many kinds of disabilities. At the same time, considering a hard of hearing person, who cannot hear, but cannot see. This framework can convert the speech into a summarized textual format for understanding.

5 Conclusion and Future Work

The most critical area of machine intelligence is speech recognition, which must be taken into account. This method attempted to show how much speech recognition and speech-to-text translation have progressed in recent years. This method further elucidates how various methodologies of speech-to-text translation are applied to different language pairs. Speech is used as input in this process, and it is then translated into text. Data editing is also done for the text. Were, Man Ki Baat by PM Narendra Modi dataset have taken for comparison between text. As a result, we got the summarized text and the translated text with a precision of 0.0355. This system translated English text into Hindi and produced a text description of the input expression. The most emphasized terms with the most important message are also effectively retrieved. Based on Context and subject, a similarity score is calculated. In general, the method has met all of the work's goals.

Multi-lingual speech recognition, the fastest training, and a working environment that includes smartphone

applications and improved interfaces can be marked for future work.

Conflict of Interest

We the authors declare that we have no conflict of interest.

References

- [1] M. Yuan and C. Zou, "Text Keyword Extraction Based on Meta-Learning Strategy," IEEE 2018 International Conference on Big Data and Artificial Intelligence (BDAI) - Beijing, China, 2018.
- [2] D. Yuan, X. Yang, X. Nong and H. Fu, "A new algorithm: Extracting text information from Webpage based on block and tag-function," *IET International Conference on Information Science and Control Engineering 2012 (ICISCE 2012)*, 2012.
- [3] X. You, "Automatic Summarisation and Keyword Extraction from Web Page or Text File," *IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET)*, 2019.
- [4] K. Wolk and K. Marasek, "Real-Time Statistical Speech Translation," in *New Perspectives in Information Systems and Technologies, Volume 1*, Cham, 2014.
- [5] C. Vimala, "A Review on Speech Recognition Challenges and Approaches," in *World of Computer Science and Information Technology Journal (WCSIT)*, 2012.
- [6] I. Viikki, I. Kiss and J. Tian, "Speaker- and language-independent speech recognition in mobile communication systems," 2001.
- [7] S. K. Upadhyay and V. N. Chavda, "INTELLIGENT SYSTEM BASED ON SPEECH RECOGNITION WITH CAPABILITY OF SELF-LEARNING," in *2014 International Journal For Technological Research In Engineering*, 2014.
- [8] A. H. Unnibhavi and D. S. Jangamshetti, "A survey of speech recognition on south Indian Languages," in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi*, 2016.
- [9] S. Ueno, M. Mimura, S. Sakai and T. Kawahara, "Multi-speaker Sequence-to-sequence Speech Synthesis for Data Augmentation in Acoustic-to-word Speech Recognition," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [10] M. Tu, F. Zhang and W. Liu, "End-to-End Speech Translation with Self-Contained Vocabulary Manipulation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.
- [11] M. Trinkle and A. Hashemi-Sakhtsari, "Untethered Distant Speech-to-Text Transcription Using Arrays of Directional Microphones," in *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2019.
- [12] B. TombaloGlu and H. Erdem, "A SVM based speech to text converter for Turkish language," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017.
- [13] M. T. Tausif, S. Chowdhury, M. S. Hawlader, M. Hasanuzzaman and H. Heickal, "Deep Learning-Based Bangla Speech-to-Text Conversion," in *2018 5th International Conference on Computational Science/ Intelligence and Applied Informatics (CSI)*, 2018.

- [14] T. Sung, J. Liu, H. Lee and L. Lee, "Towards End-to-end Speech-to-text Translation with Two-pass Decoding," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [15] M. C. Stoian, S. Bansal and S. Goldwater, "Analysing ASR Pretraining for Low-Resource Speech-to-Text Translation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.
- [16] J. Stephen, M. Anjali and V. K. Bhadran, "Voice-enabled multi-lingual newspaper reading system," *IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS)*, 2013.
- [17] K. M. Shivakumar, K. G. Aravind, T. V. Anoop and D. Gupta, "Kannada speech to text conversion using CMU Sphinx," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, 2016.
- [18] K. M. Shivakumar, V. V. Jain and P. K. Priya, "A study on impact of language model in improving the accuracy of speech to text conversion system," in *2017 International Conference on Communication and Signal Processing (ICCP)*, 2017.
- [19] K. M. Shivakumar, V. V. Jain and P. K. Priya, "A study on impact of language model in improving the accuracy of speech to text conversion system," in *2017 International Conference on Communication and Signal Processing (ICCP)*, 2017.
- [20] N. Sharma and S. Sardana, "A real-time speech to text conversion system using bidirectional Kalman filter in Matlab," in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2016.
- [21] R. Shadiev, B. L. Reynolds, Y. Huang, N. Shadiev, W. Wang, R. Laxmisha and W. Wannapipat, "Applying Speech-to-Text Recognition and Computer-Aided Translation for Supporting Multi-lingual Communications in Cross-Cultural Learning Project," 2017.
- [22] E. Saranya, B. B. Sam and R. Sethuraman, "Speech to text user assistive agent system for impaired person," in *2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)*, 2017.
- [23] R. S. Rocha, P. Ferreira, I. Dutra, R. Correia, R. Salvini and E. Burnside, "A Speech-to-Text Interface for MammoClass," in *2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, 2016.
- [24] B. R. Reddy and E. Mahender, "Speech to Text Conversion using Android Platform," in *International Journal of Engineering Research and Applications (IJERA)*, 2013.
- [25] B. Ramani, M. P. A. Jeeva, P. Vijayalakshmi and T. Nagarajan, "Voice conversion-based multi-lingual to polyglot speech synthesiser for Indian languages," *IEEE International Conference of IEEE Region 10 (TENCON 2013)*, 2013.
- [26] N. Prasangini and H. Nagahamulla, "Sinhala Speech to Sinhala Unicode Text Conversion for Disaster Relief Facilitation in Sri Lanka," in *2018 IEEE International Conference on Information and Automation for Sustainability (ICIAS)*, 2018.
- [27] S. Pathak, "<http://shikhapathak6.blogspot.com/2017/05/how-to-add-speech-recognition-to-your.html>," 2017.
- [28] F. Pantazoglou, G. Kladis and N. Papadakis, "A Greek voice recognition interface for ROV applications, using machine learning technologies and the CMU Sphinx platform," *WSEAS Transactions on Systems and Control*, vol. 13, pp. 550-560, 3 2019.
- [29] S. Noel, "Human-computer interaction(HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU)," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020.
- [30] M. Nicolao, C. Drioli and P. Cosi, "Voice GMM modelling for FESTIVAL/MBROLA emotive TTS synthesis," in *INTERSPEECH*, 2006.
- [31] A. U. Nasib, H. Kabir, R. Ahmed and J. Uddin, "A Real-Time Speech to Text Conversion Technique for Bengali Language," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (JC4ME2)*, 2018.
- [32] A. U. Nasib, H. Kabir, R. Ahmed and J. Uddin, "A Real-Time Speech to Text Conversion Technique for Bengali Language," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (JC4ME2)*, 2018.
- [33] M. M. Nahid, B. Purkaystha and M. S. Islam, "Bengali speech recognition: A double-layered LSTM-RNN approach," *20th International Conference of Computer and Information Technology (ICCIT)*, 2017.
- [34] K. Murray, J. Kinnison, T. Nguyen, W. Scheirer and D. Chiang, "Auto-Sizing the Transformer Network: Improving Speed, Efficiency, and Performance for Low-Resource Machine Translation," 2019.
- [35] R. D. Lero, C. Exton and A. L. Gear, "Communications using a speech-to-text-to-speech pipeline," in *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2019.
- [36] S. Lee and H. Kim, "News Keyword Extraction for Topic Tracking," in *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, 2008.
- [37] L. Lamel, M. Adda-Decks, J. L. Gauvain and G. Adda, "Spoken language processing in a multi-lingual context," *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP*, 1996.
- [38] T. P. Laksono, A. F. Hidayatullah and C. I. Ratnasari, "Speech to Text of Patient Complaints for Bahasa Indonesia," in *2018 International Conference on Asian Language Processing (IALP)*, 2018.
- [39] B. Lakdawala, F. Khan, A. Khan, Y. Tomar, R. Gupta and A. Shaikh, "Voice to Text transcription using CMU Sphinx A mobile application for healthcare organisation," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018.
- [40] R. Kusumah, R. Hartanto and R. Hidayat, "Hybrid automatic speech recognition model for speech-to-text application in smartphones," in *2019 International Conference on ICT for Smart Society (ICISS)*, 2019.
- [41] A. Kunchukuttan, *The IndicNLP Library*, 2020.
- [42] Y. Kumar and N. Singh, "An automatic speech recognition system for spontaneous Punjabi speech corpus," in *Int. J. Speech Technol.*, 2017.
- [43] T. Kano, S. Sakti and S. Nakamura, "End-to-End Speech Translation With Transcoding by Multi-Task Learning for Distant Language Pairs," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1342-1355, 2020.
- [44] S. K. Gaikwad, B. W. Gawali and P. Yannawar, "Article: A Review on Speech Recognition Technique," *International Journal of Computer Applications*, vol. 10, p. 16–24, 11 2010.
- [45] F. Johnson, S. Garza and K. Gutierrez, "Research on the use of voice to text applications for professional writing," in *2016*

- IEEE International Professional Communication Conference (IPCC), 2016.
- [46] Y. Jia, M. Johnson, W. Macherey, R. J. Weiss, Y. Cao, C. Chiu, N. Ari, S. Laurenzo and Y. Wu, "Leveraging Weakly Supervised Data to Improve End-to-end Speech-to-text Translation," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [47] S. Indurthi, H. Han, N. K. Lakumarapu, B. Lee, I. Chung, S. Kim and C. Kim, "End-end Speech-to-Text Translation with Modality Agnostic Meta-Learning," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.
- [48] X. Huang and L. Deng, "An Overview of Modern Speech Recognition," *Handbook of Natural Language Processing, Second Edition, Chapter 15, Chapman & Hall/CRC*, vol. , pp. 339-366, 2010.
- [49] C.-Y. Hsu and C.-P. Chen, "Speaker-dependent model interpolation for statistical emotional speech synthesis," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012, 8 2012.
- [50] I. Gogul and S. Kumar, "Flower species recognition system using convolution neural networks and transfer learning," *Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, 2017.
- [51] Y. H. Ghadage and S. D. Shelke, "Speech to text conversion for multi-lingual languages," *International Conference on Communication and Signal Processing (ICCP)*, 2016.
- [52] B. W. Gawali, "A Review on Speech Recognition Technique," in *International Journal of Computer Applications (0975 – 8887)*, 2010.
- [53] M. A. D. Gangi, M. Negri and M. Turchi, "One-to-Many Multi-lingual End-to-End Speech Translation," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2019.
- [54] L. Curling, "English Speech Recognition Method Based on Hidden Markov Model," in *2016 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, 2016.
- [55] Y. Chung, W. Weng, S. Tong and J. Glass, "Towards Unsupervised Speech-to-text Translation," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
- [56] C. Chu, T. Nakazawa and S. Kurohashi, "Iterative Bilingual Lexicon Extraction from Comparable Corpora with Topical and Contextual Knowledge," in *Computational Linguistics and Intelligent Text Processing*, Berlin, 2014.
- [57] J. Choi, H. Gill, S. Ou, Y. Song and J. Lee, "Design of Voice to Text Conversion and Management Program Based on Google Cloud Speech API," in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018.
- [58] M. Chen, X. Xu, X. Yu and X. Zhu, "The SRC-B speech-to-text systems for OC16 Chinese-English mix-ASR challenge," in *2016 Conference of The Oriental Chapter of International Committee for Coordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, 2016.
- [59] S. Chandrakala and N. Rajeswari, "Representation Learning-Based Speech Assistive System for Persons With Dysarthria," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, pp. 1510-1517, 2017.
- [60] S. Bansal and S. S. Agrawal, "Development of Text and Speech Corpus for Designing the Multi-lingual Recognition System," International Conference on Speech Database and Assessments, pp. 1-8, 2018.
- [61] S. Bansal and S. S. Agrawal, "Development of Text and Speech Corpus for Designing the Multi-lingual Recognition System," in *2018 Oriental COCOSDA - International Conference on Speech Database and Assessments*, 2018.
- [62] P. Bahar, T. Bieschke and H. Ney, "A Comparative Study on End-to-End Speech to Text Translation," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2019.
- [63] L. Abou-Abbas, C. Tadj, C. Gargour and L. Montazeri, "Expiratory and Inspiratory Caries Detection Using Different Signals' Decomposition Techniques," *Journal of voice : official journal of the Voice Foundation*, vol. 30, 8 2016.
- [64] D. S. Dilshad Kaur, "A systematic literature review on extraction of Parallel Corpora from Comparable corpora," *Journal of Computer Science, Science Publications*, 2021.
- [65] R. S. Satwinder Singh, "Text Similarity Measures in News Articles by Vector Space Model Using NLP," *Journal of The Institution of Engineers (India): Series B*, 2021.
- [66] S. S. Anukriti Srivastav, "Proposed Model for Context Topic Identification of English and Hindi News Article Through LDA Approach with NLP Technique," *Journal of The Institution of Engineers (India): Series B*, 2021.

Parallel Data Extraction from Comparable Data: A Proposed Architecture

Dilshad Kaur, PhD Research Scholar, Central University
of Punjab, Bathinda,
dilshadkaur@outlook.com

Dr. Satwinder Singh, Associate Professor CST, Central
University of Punjab, Bathinda,
satwinderese@gmail.com

Abstract: In today's era of globalisation, translation is holding an utter importance in every field. Manual translation is a slow process so the research society started to concentrate towards the branch of machine translation. Machine translation requires the set of data called corpora for the translation process to take place from one language to another. Parallel corpora in any two language sets can be used for Machine translation. Comparable Corpora can be used as an alternate to parallel corpora as the latter is scarcely available. In this article, a proposed architecture is projected for extraction of parallel corpora from the comparable corpora. It is also anticipated that what requirements are techniques will be used in the proposed architecture. Along with that a brief literature is given that depicts the related work done by the research society in the field of machine translation for parallel data extraction.

Keywords

Comparable Corpora, Parallel Corpora, Machine Translation

1. Introduction

In the present scenario, the demand of translation is having very high priority. Translation of data from one language to another is growing rapidly. Translation is holding a great importance in various fields of natural language processing, linguistic studies, social sciences etc. Manual translation with human interference is a slow and time-consuming process. Machines on the other hand can do the work faster and efficiently. Machine Translation (MT) is a medium for achieving the target of translation amongst two or more languages. Machine Translation (MT) employs data driven methodologies. The data required for the purpose of translation by machines is given a linguistic term called corpus. Corpus is the collection of textual data which is used to analyse the usage of words/tokens, phrases/ sentences and language. Corpus is broadly classified into Comparable Corpora and Parallel Corpora. Parallel corpus consists of two different language corpora where one is the translation of other. Parallel corpus is sentence aligned bilingual texts. Whereas a

comparable corpus is set of two or more different language corpora which is not the exact translation of each other and hence are not aligned. There are two main approaches to create a corpus namely rule-based and statistical analysis. In early days of MT research, rule based played a keen role. In this, all the translation rules are written manually and then encoded into the MT system. However, languages are very vast and complex, it is quite impossible to write the rules manually and in a short span of time. Over the last decade or two, MT research has taken a turn towards statistical machine translation (Koehn, 2010) (Och and Ney, 2003) and it has emerged as a main approach in both the research community and the commercial sector. In SMT, translation knowledge is automatically acquired from parallel corpora (sentence-aligned bilingual texts), and making the rapid development in MT systems for different language pairs and domains. It can be achieved once parallel corpora are available. Nowadays, most machine translation research is conducted in this

approach. In SMT, because of the high dependence on parallel corpora, the quality and quantity of parallel corpora are crucial. However, except for a few language pairs and some specialized domains, high quality parallel corpora of sufficient size remain a scarce resource. This insufficiency of parallel corpora has become the main obstacle for SMT. Exploiting comparable corpora is an effective way to address the scarceness of parallel corpora for SMT. The main reason behind this is firstly, comparable corpora are far more available for various domains than parallel corpora, such as Wikipedia, patent documents, news articles and academic papers which are easily available than parallel corpora. Secondly, when using comparable corpora there is one monolingual corpus per language suffices, and monolingual corpora are easy to obtain. Thirdly, there are a large amount of parallel data in comparable corpora, such as bilingual lexicons, parallel sentences and parallel fragments. These three types of parallel data can be extracted from comparable data.

The main goal of BLE is the construction of bilingual dictionaries, which are important for both SMT and Cross-Lingual Information Retrieval (CLIR) (Pirkola et al., 2001). BLE from comparable corpora is based on the distributional hypothesis (Harris, 1954), stating that words with similar meaning appear in similar distributions across languages. Contextual similarity is mostly used in BLE. There are two main categories of methods for Bilingual Lexicon Extraction (BLE) from comparable corpora, namely Topic Model Based Method (TMBM) and Context Based Method (CBM). Both methods are based on the distributional hypothesis. TMBM measures the similarity of two words on cross-lingual topical distributions, while CBM measures the similarity on contextual distributions across languages. Parallel

sentence extraction identifies parallel sentences from comparable corpora and automatically constructs parallel corpora for SMT. Parallel sentences can be identified based on classification (Munteanu and Marcu, 2005) or by using translation similarity measures (Utiyama and Isahara, 2003). There are very less or no parallel sentences in comparable corpora with low comparability but there could be parallel fragments in comparable sentences. These parallel fragments are very useful in statistical machine translation (SMT). Parallel fragment extraction depends on bilingual lexicons (Munteanu and Marcu, 2006) or alignment models (Quirk et al., 2007).

In this global era, the demand of translation is growing at a fast rate and manual translation of everything is near to impossible and is very time consuming. So, there is a requirement of an automation system. With the coming of automation in this field, there can be more efficiency in the work and eventually cost of translation would also be less. Mostly the work has been done in European languages but our focus would be on the regional languages i.e., Punjabi and Hindi. New parallel corpora would be created of English – Punjabi and English-Hindi. The availability of parallel corpora will make the development of machine translation systems fast for the required language pair and domains. In SMT, the translation knowledge is acquired from 3 parallel corpora, so the quality and quantity of parallel corpora is very important. High quality of parallel corpora in regional language of sufficient size is scarce resource. To remove this scarcity would also be the focus of this research. The comparable corpora would be exploited to address the problem of scarceness of parallel corpora. Bilingual lexicons, parallel sentences and parallel fragments

would be extracted from comparable corpora. These extracted sentences and fragments can prove to be helpful in addressing the accuracy and scarcity of SMT.

2. Related Work

Comparable corpora can be huge in size so it is quite difficult to examine every sentence in the corpora. So the concentration is made only on those documents and sentences which have similar 6 kind of content. For finding similar or comparable documents, techniques like topic alignment, content alignment, text alignment, and cosine similarity can be employed. Cosine similarity is used to measure the similarity in the two documents which are represented as vectors of the words they contain. Cosine similarity is the Dot product of the vectors (Fung and Cheung, 2004a). Lehal et al. (2019) compared the similarity and distance measures. Their research has analyzed cosine similarity, jaccard coefficient, Hamming distance and Euclidean distance. The data was taken from Wikipedia in Punjabi and English language. The content similarity was found and compared in the form of recall, precision, f1 score and accuracy. Their work concluded that cosine similarity gives better accuracy and f1 score. Resnik and Smith (2003) also revealed word to word translation. This method uses translation similarity based on word-byword translation lexicon. It is also known as content-based alignment. Utiyama and Isahara (2003) proposed two steps to obtain accurate alignments of article and sentence. The article alignment test uses similarities in sentences associated with Dynamic Programming (DP) matching, and similarities in articles aligned with Cross-Language Information Retrieval (CLIR) for sentence alignment. The experiments involved improvement of

one another and allowed for accurate extraction of relevant article and sentence alignments from an extremely noisy parallel Japanese-English corpus. Goyal et al. (2020) generated a comparable corpus from the data taken from Wikipedia. The data was aligned according to the topics. The corpus was tagged only for nouns and a php script was used to align the articles according to the topics.

Extraction of bilingual lexicons is the oldest way of using comparable corpora. Liu et al. (2013) developed a method by converting document aligned comparable corpora into a parallel topic-aligned corpus using BiLDA topic models, and identify word translations with the help of word alignment. The findings from the study proposed an efficient way of extracting bilingual dictionaries through a novel combination of topic modeling and word alignment method. The study focused on the use of conversion of a comparable document-aligned corpus into a parallel topic aligned corpus. Bouamor et al. (2013) introduced a novel approach to the creation of specific domain bilingual lexicon that depends on Wikipedia. This massively multilingual encyclopedia made it possible to make lexicons for an oversized number of language pairs. A year later Chu et al. (2014) proposed a new method that was the combination of topic based and context based which extracted bilingual lexicons. Experiments of this research conducted on Chinese-English and Japanese-English Wikipedia information showed that their planned methodology performs considerably higher than a state-of-the-art technique that solely uses topical information.

Parallel corpora are sentence- aligned bilingual texts. The translation knowledge is gained mostly from parallel corpora but the quality and quantity of parallel corpora

is scarce. Lu et al. (2010) presented a large parallel corpus that is derived from a wide corpus of comparable English-Chinese patents obtained from the Internet. Firstly parallel sentence pairs were created with Champollion, a publicly available aligner of sentences, and then filtered the candidates with another sentence aligner, namely MS Aligner. Post et al. (2012) gathered and refined document level parallel corpora between English and six verb final languages which are Bengali, Hindi, Malayalam, Tamil, Telugu and Urdu. Their study has described the collection of six parallel corpora containing four-way redundant translations of the source-language text. They revealed that the Indian languages of these corpora are low-resource and understudied, and exhibit markedly different linguistic properties compared to English. Deep et al. (2018) provided in the research different sources to collect the English Punjabi corpus and English Hindi Corpus. They used human translation approach, online translation through web and already available data for

converting the data into required language. The corpus was created from resources like TDIL, EMILLE, Brills Bilingual newspaper, Gyan Nidhi and Wikipedia. Jindal et. al. (2018) worked upon English to Punjabi machine translation using free translation software called moses. Moses require a parallel sentence pair corpus for performing the translation work. In their research they created a corpus of 20000 sentences which were of different domains. GIZA++ alignment tool was used to align the sentences. Moses software was trained and translations were made. The accuracy was checked using BLEU scripts. Lakshmi et al. (2020) revealed that one of the promising resources to extract dictionaries are said to be parallel corpora. Their study found that comparable corpora could be an alternative to extract dictionary. Proposed approach was to extract dictionary for a low resource language pair English and Kannada using comparable corpora obtained from Wikipedia dumps and corpus received from Indian Language Corpus Initiative (ILCI).

Table 1: Research performed with different language sets under parallel data extraction

Author Name	Language Used	Work Performed
Utiyama and Isahara (2003)	Japanese English	Obtaining alignment of articles and sentences using CLIR
Cao et al. (2007)	English Chinese	Transliterations and translations of large Chinese web content
Lu et al. (2010)	English-Chinese	Largest corpora of parallel sentences
Qian et al. (2012)	English Chinese	Bilingual dependency mapping model
Post et al. (2012)	English – Bengali, Hindi, Malayalam, Tamil, Telugu and Urdu	Gathered and refined document level parallel corpora amongst six languages
Bouamor et al. (2013)	French English	Increase the adequacy of context vector

Gupta et al. (2013)	English Bengali	Similarity calculation and parallel fragment extraction using Wikipedia documents
Stefanescy and Ion (2013)	English- German, Romanian and Spanish	Parallel sentence extraction from Wikipedia data
Chu et al. (2014)	Chinese English and Japanese English	Data from Wikipedia was taken to construct corpora and extract bilingual lexicons
Deep et al. (2018)	English-Punjabi and English -Hindi	Collection of data from different resources and Creation of manual corpora
Jindal et al. (2018)	English Punjabi	Machine translation using moses (corpus based)
Lehal et al. (2019)	English Punjabi	Analysis of similarity and distance measure
Lakshmi et al. (2020)	English Kannada	Dictionary from comparable corpora from Wikipedia dump
Goyal et al. (2020)	English -Punjabi	Wikipedia data was taken and aligned according to topics

Parallel corpus is quite difficult to build but parallel fragments, sentences and lexicons can be easily achieved through comparable corpora. Gupta et al. (2013) translated a source fragment with an existing SMT system, and identified the target fragment by calculating the similarity between the translated source and target fragments. The study revealed the usage of an automatic approach of extracting English-Bengali parallel fragments of text from comparable corpora created by using Wikipedia documents. Chu et al. (2013) also proposed a system for sentence extraction from the quasi-comparable corpus. The system trained and tested a unique classifier that stimulates the parallel sentence extraction. The study used linguistic information of Chinese character for extraction. Stefanescu and Ion. (2013) described a set of parallel sentences extracted from Wikipedia for three pairs of languages that

is English-German, English-Romanian and English-Spanish. To do this, a tool LEXACC was utilized for extracting parallel sentences from comparable corpora which was developed throughout their project. Table 1 depicts the research carried under different language sets for comparable and parallel data extraction.

3. Proposed Methodology:

The extraction of parallel data involves the number of tasks which are elaborated in Fig. 1. Firstly, there is the requirement of data resources for performing the task of extraction. But as mentioned earlier also in the survey, parallel data is not easily available in desired languages. To overcome this problem of scarcity, Comparable Corpora which are available in huge amount but in raw form can be used for the extraction of parallel data. There are three types of parallel data in Comparable

Corpora i.e., “parallel sentences”, “parallel fragments” and “bilingual lexicons”.

A parallel data extraction will consist of the following steps:

- Potential resources, like comparable corpora in the desired language pair, Extraction of Bilingual Lexicons and a seed parallel dictionary.

Initially, collection of comparable corpora will be made in these two languages i.e., English and Punjabi from different available resources like bilingual newspapers, articles, e-books, bilingual websites, Wikipedia etc. Along with the Comparable data, a dictionary is required to be built that is to be given to system for training and testing purpose. The dictionary would contain of words translation from English language to Punjabi.

For bilingual lexicon extraction, first the TMBM (Topic Model Based Method) will be applied to gather the bilingual lexicons from the given set of data in two languages. Then the obtained topical bilingual lexicons obtained are to be used as initial seed dictionary. This seed dictionary is then used with CBM (Context based method) to obtain the contextual bilingual lexicons. Later, we will combine the topical bilingual lexicons with contextual bilingual lexicons to obtain combined bilingual lexicons. The combination is performed using the similarity scores of both topical and contextual bilingual lexicons.

- Document Alignment model, to get similar document pairs.

After the collection of data in two different textual files, alignment of the data will be

done using Python Scripts. Different alignment software's are also available online for the alignment such as GIZA ++, moses or Lingtrain aligner. The aligners can be modified for Punjabi language. Once the alignment of the articles is done, the text cleaning is to be performed by removing special characters, stop words, html tags, phonetic symbols while preserving the actual sentences. All the cleaning can be executed using Python toolkit NLTK.

- Parallel Sentence and Fragment Extraction, to get parallel sentences and fragments from the aligned documents.

Using the combined dictionary, CLIR will be applied to generate parallel sentence candidates from comparable corpora. The parallel sentence extraction will be created that will comprise of parallel sentence classifier and a classifier to identify the parallel sentences. Next, parallel sentence extraction will be applied that will classify the parallel sentences and comparable sentences. Bilingual Lexicon Extraction is to be used here for performing parallel sentence extraction. Additionally, parallel fragment extraction is applied to extract parallel fragments from the comparable sentences. The extracted parallel sentences can be used to support parallel fragment extraction. An alignment model can be used to find the parallel fragments in the textual data. The extracted parallel sentences and parallel fragments can be used as training data for statistical machine translation (SMT). Finally, BLE is employed to further test the accuracy of statistical machine translation (SMT).

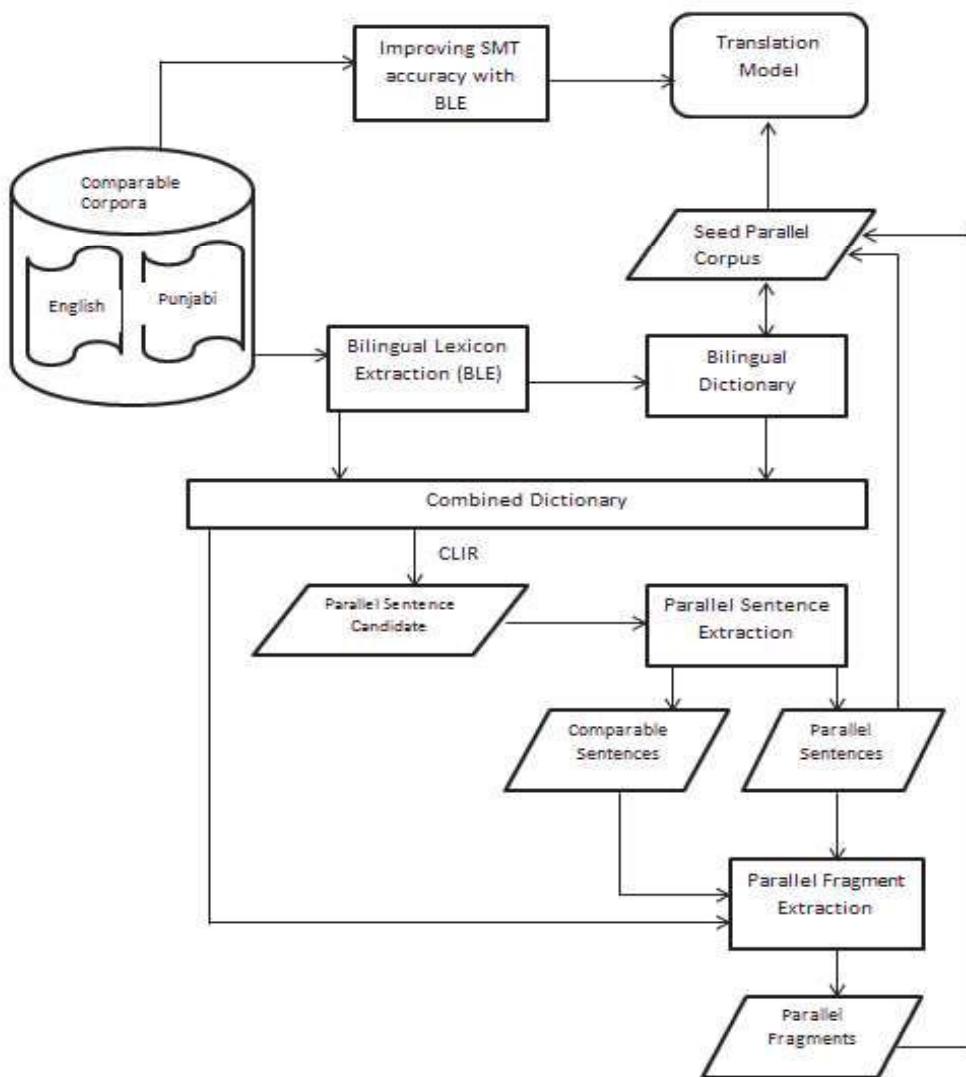


Fig.1. Proposed Model for parallel data extraction

4. Conclusion

In this paper, we presented a proposed architecture for mining the parallel data from the comparable data. The parallel data is available in the form of lexicons, fragments and sentences. We presented the methodology for extracting the bilingual lexicons by exploring the topic-based models and context-based models. The architecture proposed here will focus on

two languages i.e., English and Punjabi. The comparable corpora will be created for the two language sets. The corpora created will be cleaned and aligned before the extraction process. Also, in proposed model, sentence and fragment extraction for parallel data will be performed. This paper gives a brief insight about the mining process of parallel data from the available comparable data.

References

- [1] Bouamor, D., Popescu, A., Semmar, N. and Zweigenbaum, P. (2013). Building specialized bilingual lexicons using large scale background knowledge. Conference on Empirical Methods in Natural Language Processing: 479-489.
- [2] Bouamor, D., Semmar, N. and Zweigenbaum, P. (2013). Context vector disambiguation for bilingual lexicon extraction from comparable corpora. Association for Computational Linguistics (2):759-764.
- [3] Cao, G., Gao, J. and Nie, J. -Y. (2007). A system to mine large-scale bilingual dictionaries from monolingual web pages. MT Summit 2007:57-64.
- [4] Chu, C., Nakazawa, T. and Kurohashi, S. (2013). Chinese-Japanese parallel sentence extraction from quasi-comparable corpora. Sixth Workshop on Building and Using Comparable Corpora: 34-42.
- [5] Chu, C., Nakazawa, T. and Kurohashi, S. (2014). Iterative bilingual lexicon extraction from comparable corpora with topical and contextual knowledge. CICLing2014: 296-309.
- [6] Deep, K., Kumar, A. and Goyal, V. (2018). Development of Punjabi-English(PunEng) Parallel Corpus for Machine Translation System. International Journal of Engineering and Technology 7(2):690-693.
- [7] Fung, P. and Cheung, P. (2004). Multi-level bootstrapping for extracting parallel sentences from a quasi-comparable corpus. Coling 2004:1051-1057.
- [8] Goyal, V., Kumar, A. and Lehal, M.S.. (2020) Document Alignment for Generation of English Punjabi Comparable Corpora from Wikipedia. International Journal of E-Adoption (12) :42- 51.
- [9] Gupta, R., Pal, S and Bandyopadhyay, S. (2013). Improving mt system using extracted parallel fragments of text from comparable corpora. Workshop on Building and Using Comparable Corpora:69-76.
- [10] Jindal,S., Goyal, Vishal. and Singh, Jaskaran.(2018) English to Punjabi statistical machine translation using moses(corporus based). Journal of Statistics and management Systems 21(4):553-560.
- [11] Koehn, P. (2010). Statistical Machine Translation. Cambridge University Press, 1st edition.
- [12] Lehal, M.S., Kumar, A. and Goyal, V. (2019) Comparative Analysis of Similarity Measures for extraction of Parallel Data. International Journal of Control and Automation (12):408-417.
- [13] Lu, B., Jiang, T., Chow, K. and Tsou, B. K. (2010). Building a large english-chinese parallel corpus from comparable patents and its experimental application to smt. Workshop on Building and Using Comparable Corpora, LREC 2010: 42-49.
- [14] Munteanu, D. S. and Marcu, D. (2005). Improving machine translation performance by

- exploiting non-parallel corpora. Computational Linguistics 31(4):477-504.
- [15] Munteanu, D. S. and Marcu, D. (2006). Extracting parallel sub-sentential fragments from nonparallel corpora. International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics: 81-88.
- [16] Matt Post, Callison-Burch, C. and Osborne, M. (2012) Constructing Parallel Corpora for Six Indian Languages via crowdsourcing. Workshop on Statistical Machine Translation: 401- 409.
- [17] Och, F. J. and Ney, H. (2003). A systematic comparison of various statistical alignment models. Computational Linguistics 29(1):19-51.
- [18] Pirkola, A., Hedlund, T., Keskustalo, H. and Järvelin, K. (2001). Dictionary-based crosslanguage information retrieval: Problems, methods, and research findings. Information Retrieval (4):209-230.
- [19] P. Resnik and N. A. Smith. (2003). The web as a parallel corpus. Comput. Linguist ,29(3): 349- 380.
- [20] M. Post, C. Callison-Burch, and M. Osborne. (2012)
- Constructing parallel corpora for six indian languages via crowdsourcing. Workshop on Statistical Machine Translation.401-409
- [21] Qian, L., Wang, H., Zhou, G. and Zhu, Q. (2012). Bilingual lexicon construction from comparable corpora via dependency mapping. COLING2012: 2275- 2290.
- [22] Quirk, C., U, R. U. and Menezes, A. (2007). Generative models of noisy translations with applications to parallel fragment extraction. MT Summit XI, European Association for Machine Translation.
- [23] Stefanescu, D. and Ion, R. (2013). Parallel-wiki: A collection of parallel sentences extracted from wikipedia. CICLing 2013: 117-128.
- [24] Sowmya Lakshmi, B.S., Shambhavi,B.R. (2020) Extraction of Bilingual Dictionary from Comparable Corpora for Resource Scarce Languages. Journal of Computational and Theoretical Nanoscience 17 (7):54-60.
- [25] Utiyama, M. and Isahara, H. (2003). Reliable measures for aligning japanese-english news articles and sentences. Annual Meeting of the Association for Computational Linguistics: 72- 79.

Role of Big Data for E-Business –Future opportunities, Challenges and Value Creation

Prof.Uttamjit Kaur

Department of Computer Science

Bibi Sharan Kaur Khalsa college

Sri Chamkaur Sahib

Er.uttamkaur@gmail.com

Abstract - E-Business today is considered as one of the most important and popular trends in business .As E-Commerce which is limited to Online Transactions i.e. buying and selling services over the internet Analytics needs to be adapted by big data. As e-business, which includes with Clients in order to improve performance..To improve the enterprise process in electronic business. The paper focuses on recognizing the need for big data in E-business issues that need to be tackled. One such breakthrough technology largely used in big data that can handle Volume, Velocity, Variety and Veracity of data that is generated in the real time. The paper forced on understanding that need for big data in E-business companies, what opportunities can be harnessed by the use of it and Challenges that need to be addressed.

Index Terms - *Big data for E-Business, Analytics for E-Business, Big Data opportunities, Big Data Challenges*

I. INTRODUCTION

Big Data: In the current business scenario, data is getting rapidly generated and is considered as an asset that creates value. To extract meaningful value from such data, there is a need for technologies and built-in architectures. Handling a large amount of data that is growing at a great speed is a challenge to performing effective analysis using traditional means.

The term Big Data is referred to data in Size and surpasses the handling capability of traditional database systems. The data are created so quickly and enormous that it does not fit into traditional databases.

TABLE 1
Traditional Data Vs Big Data

Traditional Data	Big Data
Documents	Photos
Records	Graphics, Audio and Video
Files	GPS data



Fig 1. Small Data Vs Big Data

E-Businesses-Business (Electronic Business) is the conducting business on the Internet most often using web technologies. E-Business facilitates companies to collaborate their internal and external processes more efficiently and effectively, and work more closely with suppliers and partners leading to improvements in overall business performance

II.CHARACTERISTICS OF BIG DATA



Fig. 1 Four Vs of Big Data

A. Data Volume. Volume expresses the measures of the amount of data available to an organization. It represents a

challenge because in order to manipulate and analyse a big volume of data requires a lot of resources.

B. Data Velocity. Velocity refers to the speed at which the data gets generated; streamlined and arithmetic operations are performed. Data Velocity determines the speed at which data is created, streamlined and aggregated. E-Commerce firms have drastically increased the pace and fertility of data used for business transactions.

C. Data Variety. Variety refers to the different types of data that are generated and that are stored in the database. It is used to measure the relevance of data representation. The data that is generated through E-Commerce business may be in the form of text, audio, video, and images.

D. Data Veracity. Veracity refers to the conformity to truth or factual details. The data that is getting generated can vary largely. The correctness of analysis depends on the veracity of data source.

E. Data Value: The main aim of use of big data is to create business value. Firms should be able to develop meaningful insights that can derive economic value. It measures the usefulness of data in making decisions.

3. AIMS OF THE RESEARCH PAPER

- A. To understand the need for Big Data Analytics in E-Business
- B. To explore the opportunities for E-Business by use of Big Data Analytics
- C. To address the Big Data challenges faced by E-Business
- D. To foresee the trends in future.

A. To understand the need for Big Data Analytics in E-Business

Big Data Analytics for E-Business. "You can't manage what you don't measure." [10] E-Business companies are gathering more data than they know what to do with and how to deal with it. To turn all this information into a competitive advantage, they'll need new skills and techniques. Large datasets containing a variety of data types are processed by big data analytics that aims to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful business information. Akter, S., and Wamba, S.F, 2016 in their research gave a detailed explanation of the important applications of big data. The main areas of applications included personalization, customer service, predicting customer behavior, dynamic pricing, visibility in supply chain and detecting & managing fraud. [5] Data in E-Commerce may be in Structured, SemiStructured and UnStructure forms and are generated from different sources.

Such sources include: a. Machine generated data that are generated from real-time sensors b. Human-generated data that are generated through social media, tweets, photos, videos and status updates. c. Organization generated data that are of the traditional type comprising information related to business transactions. The analysis of such data leads to data-driven decision making for implementing effective marketing strategies, opportunities for new revenue generation, improve customer service, better operational efficiency and gain the competitive advantage against competitor organizations. [11] By use of Big Data Analytics E-Businesses can derive timely insights from the vast amounts of data that are generated from external third-party sources, the Internet, social media and remote sensors. It is possible to do Real-time monitoring and forecasting of events that impact either business performance or operation. By using Big Data Analytics it is possible to find, acquire, extract, manipulate, analyze, connect and visualize data [1].



Fig 2. Ways to increase business productivity with Big Data

B. To explore the opportunities for e-business by use of big data analytics

Unlike before where big data used to be a technical problem, now it's a business opportunity. Enterprises today, are exploring big data to discover facts they didn't know before. Using advanced analytics, businesses can study big data to understand the current state of the business and track still evolving aspects such as customer behavior. Visualization and advanced analytics are poised for aggressive adoption [2]. Big Data Leads to Insights, Improvements & Automation [3]. As mentioned in the Gartner Research, Big Data can significantly impact on your business in three different ways that include a. Discovering the hidden insights from the available data b. Improve decision making by enriching information c. Automating the business processes. A research firm in association with MIT, in 2012, mentioned that the firms that make use of Big Data Analytics outshined their competitors by 5% in terms of productivity and 6% in terms

of profitability. By coupling voluminous information, big data analytics helps organizations to devise consumer-centric insights for improving business operations. [11]

C. To address the Big Data Challenges faced by EBusiness

1. Challenges faced: Use of Big Data Analytics calls for people with specialized skills, sponsors and robust technologies and architectures. E-Business houses need to have the ability to ingest data in quick turnaround time given design, cleansing, and distillation required to align to semantics and syntax of enterprise canonical data model. To understand that there is a High latency between the time data is generated to the time when data is available for consumption. Data privacy, security and piracy and legal aspects [11]. Challenges in difficulty in verification of authenticity and precision of data. Complexity in integrating the non-traditional kinds of data with traditional data from other enterprise systems for analysis. [11]

2. Other challenges include:

- a. Data Acquisition and Recording and Cleaning[3]
- b. Data Integration, Aggregation, and Representation
- c. Query Processing, Data Modelling, and Analysis[3]
- d. Interpretation [5]
- e. Complex and Evolving Relationships
- f. Data complexity, Computational complexity, System complexity [11]

3. Addressing the challenges: To overcome the challenges, E-Businesses need to adopt big data analytics that is driven by a perfect blend of technologies, management, and economics. Real-time processing of streaming big data is crucial to finding meaningful data and reacting to it. [16]

4. Future challenges of big data applications

It is estimated over 80% of an enterprise data consists of data generated from supply chain partners, market surveys, emails, secondary research, real time sensors and consumers' social media data. The data may be coming in different shapes and sizes through flat file, relations database systems, XML or JSON etc. With the voluminous data, a firm faces several challenges for the future that needs to be addressed in a sophisticated manner.

Some of the key challenges for the future include:

4.1 Data integration As most of the data is either unstructured or semi structured, the ability to combine them to derive at a meaningful information, leading to better business decisions are more important. Due to different variety, the challenge is how to handle and control the quality of data. Data integration should focus on reducing data complexity, increased data availability through increased collaboration that adds value to big data.

4.2 Data Volume One of the key characteristics of big data is Volume. The question is, whether just the volume can give better insights about the business? And for this data needs to be acquired, stored, accessed and processed which is a challenging task for any systems. While processing larger set of data, cost considerations should be taken into account in addition to performance elements.

4.3 Availability of required skill set There are several big data tools available in the market place that needs to be looked into. In addition, skilled people are required who know how to make use of it in an efficient manner. There is definitely shortage of people (like data scientists, Machine learning developers, data analysts, statisticians and mathematicians). 2016 Predictions: Chief data officers will become the new „sought after“ personnel of IT. Self-service big data (BDaaS) portals will bond the gap between data scientists and business analysts, - Current estimates state that around one-third of chief data scientists spend up to 90 percent of their time "cleaning" raw data

4.4 Data Privacy When the data is collected, it has lot of personal information about the customers. People are/should be worried as to how information about them is being used that affects them. As their data gets monitored quite often during real time, the privacy is lost.

4.5 Data Security and Piracy: With the huge amount of data that the companies have today about their customers, they must scale up the capability to make sure that this data/information is secure. In addition, several anti-theft elements have to be addressed in order to restore the piracy issues.

4.6 Legal aspects: Once the companies collect and have the customer data, the question raises is “who is the owner of that data”. Is it the person we are referring to or the person who generated? In case of negative consequences how situations needs to be handled and who will be held responsible

5. FINDINGS

Use of big data technology is considered as an opportunity and not a problem. [11]. Bringing in big data for E-Business brings along with it valuable associate called Agility. Making

the best use of big data technologies and harnessing benefits requires trailing and far-reaching search for data. By using big data analytics technology in e-business brings benefits of having better customer relations, enhanced business performance, better productivity and modern capabilities.[10]. But there is a resilience about profit potential where experts opine that use of big data does not automatically bring profits.[22]. There are several business models depicting interdependent ecosystem of data. Deriving the best out of use of big data means having a clear-cut business model and making its prime importance to the business. [11]

6. CONCLUSIONS

Big Data Analytics has gained notable attention due to its ability to transform business practices and the possibility for use of wide range of effective decision-making tools and services [11]. Use of Big Data technologies has made it easier to capture, store and analyse unstructured data in the enterprise. The future looks very promising as it analyzes all kinds of unstructured data, with a goal to make better decision making.[10] However, there are several challenges that need to be addressed before Big Data technologies will attain its full potential. Using big data empowers decision makers to decide on the basis of evidence rather than intuition. For that reason, it has the potential to revolutionize E-Business.

REFERENCES

- [1] Monica Bulger, Greg Taylor, Ralph Schroeder, Data-Driven Business Models: Challenges and Opportunities of Big Data, Oxford Internet Institute September 2014
- [2] Walunj Swapnil K. et.al, Big Data: Characteristics, Challenges and Data Mining, International Journal of Computer Applications (0975 – 8887), ICAIM – 2016
- [3] Amir Gandomi, MurtazaHaider, International Journal of Information Management Volume 35, Issue 2, April 2015, Pages 137 – 144
- [4] Eaton et al. 2012, Understanding Big Data, McGraw-Hill 2012, Link:<http://www.ebooksdirectory.com/details.php?ebook=7122>
- [5] Chen, J.S., Tsou, H.T. (2012). Performance effects of IT capability, service process innovation, and the mediating role of customer service. Journal of Engineering and Technology Management, 29(1), 71-94.
- [6] Lutz, A. 2012, "FORRESTER: Here's Proof That 'Showrooming' Panic Is Blown Way Out Of Proportion", Business Insider – Retail [internet].

[7] Cuzzocrea, A., Song, I.-Y., and Davis, K.C. (2011). Analytics over large- scale multidimensional data: the big data revolution! In Proceedings of the ACM 14th International Workshop on Data Warehousing and OLAP, (ACM), pp. 101–104.

[8] Boyd, D., & Crawford, K. (2012). Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15, 662–679.

[9] Sam Madden (2012). Massachusetts Institute of Technology, "From Databases to Big Data", Published by the IEEE Computer Society.

[10] Marr, B. (2013, November 13). The Awesome Ways Big Data is used today to Change Our World. Retrieved November 14, 2013

[11] Minelli, M., Chambers, M., & Dhiraj, A. (2013). Business analytics (pp.99–125). Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses.

E-Commerce: Concepts and Concerns

Harpreet Kaur

Assistant Professor

Post Graduate Department of Commerce

SGGS Khalsa College, Mahilpur

(Hoshiarpur) Phone 94640 32847

Email : hksekhon01@gmail.com

Abstract

E-commerce is the process of doing business through computer and internet services. A person sitting in front of computer or using mobile devices can access to sell or purchase the products or services. E-commerce which was started in early 1990's has taken a big leap in the world of technology but including security there are many challenges facing by E-commerce in contemporary times. The major merit of E-commerce over traditional commerce is the consumer can browse online shopping, compare the products as well as prices on the simple clicks on his PC or mobile. The concept and concerns of E-commerce has changed the entire system of traditional business. There are of course, many disadvantages of E-commerce that can be eliminated with the business acumen and skills.

Keywords : E-commerce, Business, Technology, Internet, Online, Shopping, E-business

I. DEFINITION OF E-COMMERCE

In the Contemporary times E-commerce is a popular term which is also known as Electronic commerce or internet commerce. The term E-commerce is self defined, explicit and communicative which is used to buying and selling of goods/services or transmitting of funds and data, over on electronic network, especially on the internet. So as comprehensive understanding E-commerce is the activity of buying and selling the products on online services or over the internet.

So the general assumption is that E-commerce is the selling or purchasing the product in online mode but E-commerce also includes the sale and purchase of non physical goods, such as services and digital products. So the transaction of money, funds and data are also considered as Electronic Commerce or Internet Commerce. So the term E-commerce can describe any kind of commercial transaction that is facilitated through the internet Services.



II. BRIEF HISTORY OF E-COMMERCE:

Online shopping was invented in 1979 by entrepreneur Michael Aldrich in the U.K. Michael Aldrich was able to connect a modified domestic television to a real time multi-user transaction processing computer via a telephone line. The system was marketed in 1980 and offered as business-to-business system that were then sold in the U.K. Ireland and Spain.

With the title of 'Internet is open' On August 11, 1994 the issue of New York Times chronicled the sale between two friends of a sting CD. The Times said, "The team of young entrepreneurs celebrated what was apparently the first retail transaction on the Internet using a readily available version of powerful data encryption software designed to guarantee privacy." This is the first example of a consumer purchasing a product from a business through the World Wide Web or "e-commerce" as we generally know it today.

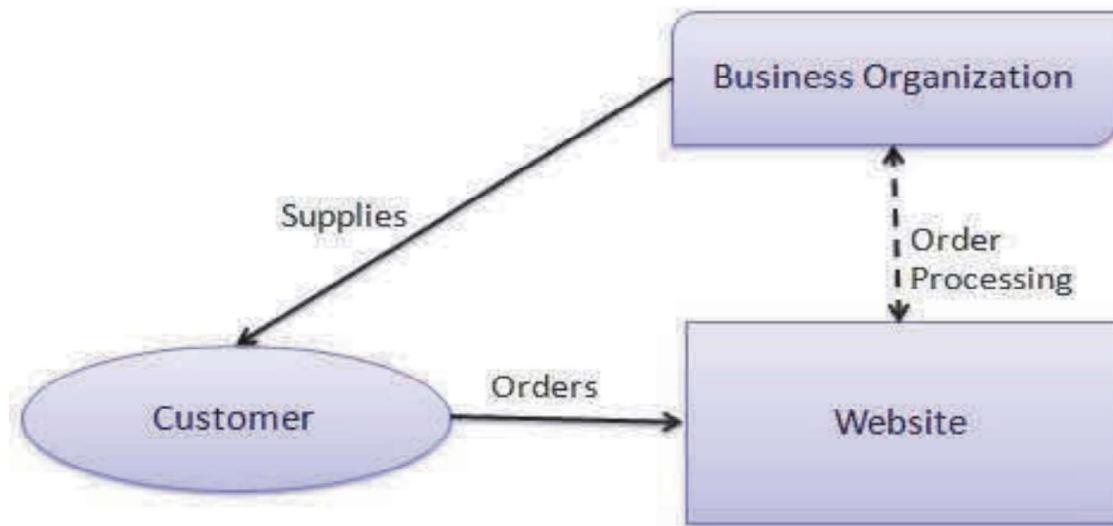
In the middle of 90s, there were major advancements in the commercial use of the Internet. Since then, e-commerce has evolved to make product easier to discover and purchase through Internet services. Freelancers, large scale corporations, small businesses and all benefited from e-commerce, which provides opportunity them to sell their goods and services at a global scale that was not possible with fixed and conventional offline retail. With the rapid development of information and communication media, the business of e-commerce has expanded very fast. Global retail e-commerce sales are projected to reach \$27 trillion by 2020. Websites created for the purpose of e-business on the internet have become

popular. One of the best example is Amazon, the first e-commerce site, which started in 1995s as an online bookstore but grew to become the largest online retailer in the world. Amazon's range has expanded over the years and now it includes almost everything related with human life.

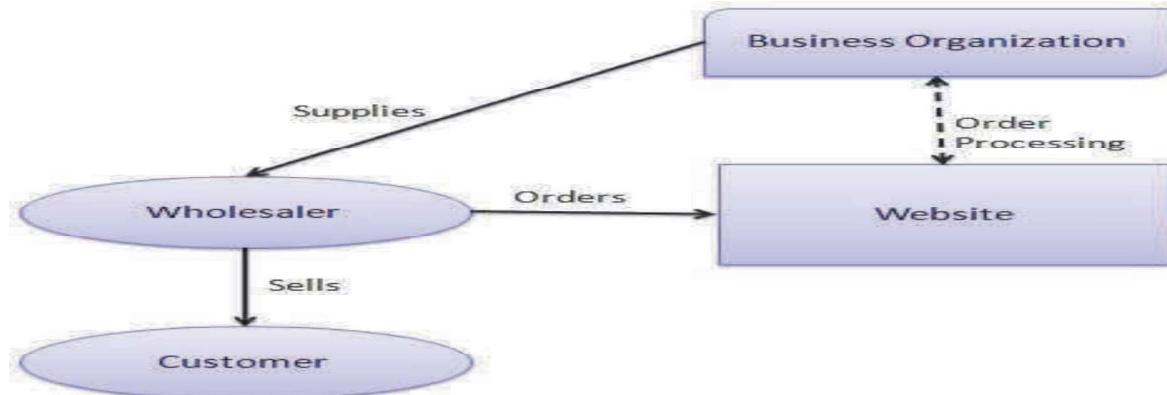
III. TYPES OF E-COMMERCE

There are four major types of e-commerce that are concerned with every transaction that take place between consumers and businesses.

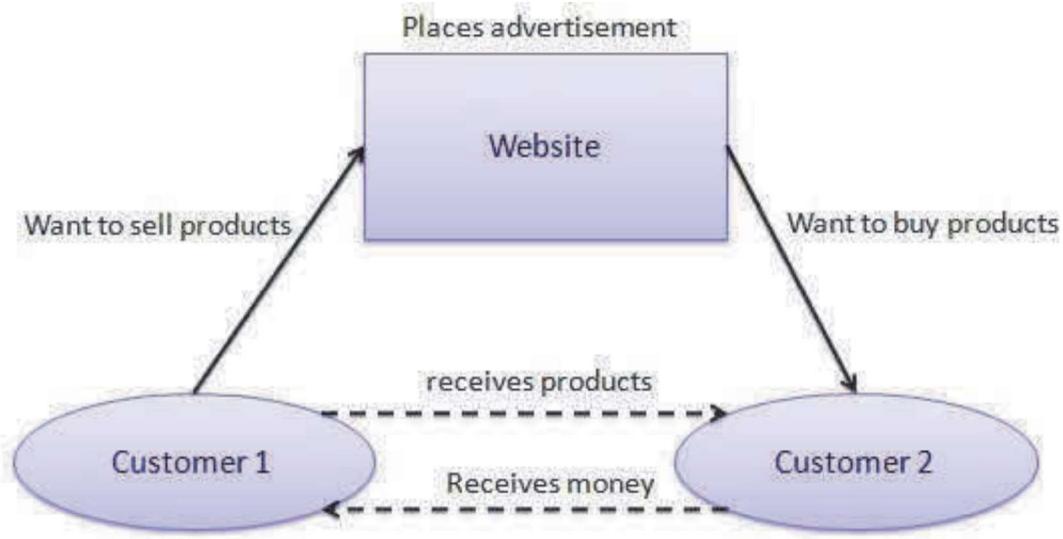
1. Business to Consumer – When a business sells a good or service to an individual consumer. For example, you buy a book from an online retailer.



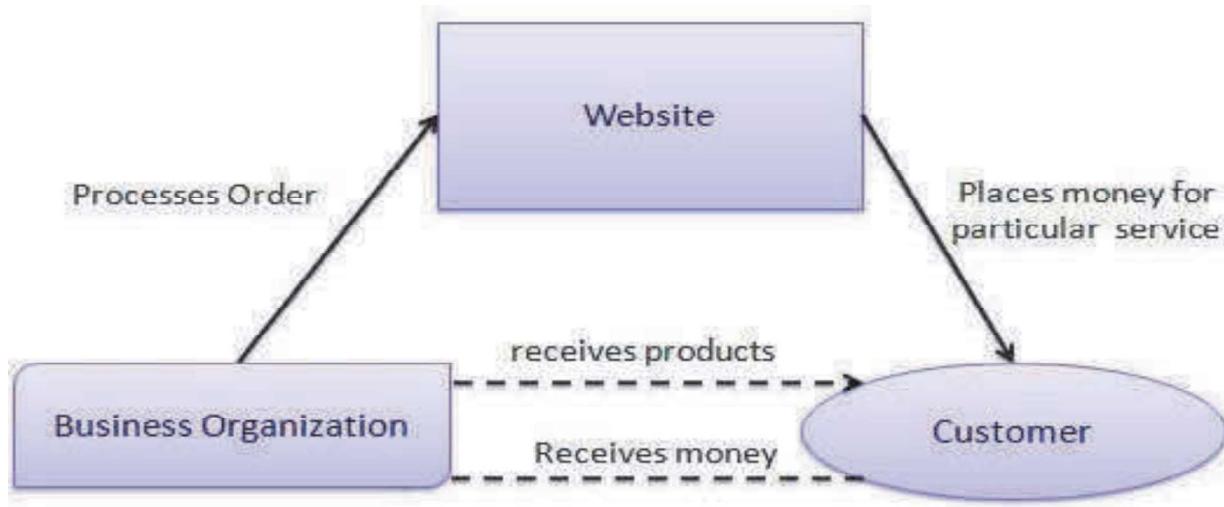
2. Business to Business – When a businesses sells a good or service to another businesses.



3. Consumer to Consumer – When a consumer sells a good or service to another Consumer. For example a person sells his old bike on Ebay to another consumer.



4. Consumer to business – When a consumer sells their goods or services to a business or organization.



IV. MAJOR IMPACTS ON MARKET, RETAILORS AND CUSTOMERS

E-commerce markets are blooming at outstanding rates. The online market is grow by more than 50 percent between the year of 2015 to 2020 and e-retail revenues are projected to grow to 4.88 trillion US Dollars in 2021. Traditional markets are expected 2 to 3 percent growth in the same period.

E-commerce allows customers not only to overcome the geographical hurdles but also allows them to purchase product anytime and from anywhere. There are different strategies of conducting business between the online and traditional markets. The pricing strategies are

also different. Traditional retailer base their prices on store traffic and the cost to keep inventory. Online retailers base prices on the speed of delivery. But there are limitations of online shopping. Online retailer cannot offer the quality of a product without the physical experience. Security of online transactions is also the major issue regarding the online market. So E-commerce websites used various tools to turn away the security threats. E-commerce has a more sophisticated level of impact on the three major flows (physical flow, financial flow and information flow) of the supply chain management. Finally E-commerce supplied companies a more intelligent, effective, efficient and aesthetical way to get together within the supply chain management.

E-commerce has also a key role to provide new opportunities to new generation of digital entrepreneurs. It helps to create fresh job possibilities due to information related services, software app and digital products. E-commerce is most benefited for the customers as they do not have to leave home and only need to browse website online, especially for purchasing the products which are not available in nearby shops. Online shopping help customers to purchase wider range of products and save the time and energy of customers. With many advantages there are also some demerits of E-commerce. It lacks the human interaction for customers, especially who prefers face-to-face connectivity. E-commerce is also not free from the errors when there are afraid of file corruption, spying, fear of card theft etc. website interruption and crashing problems, delayed in shipping and the problems of online transactions tend to remain loyal to well-known retailers. With such problems most of the customers would prefer shopping by traditional methods alternately online.

V. CONCLUSION

In this world of Globalization and information technology businesses need to accommodate to the new types of consumer needs and new concerns because it will prove to be essential for the success and survival of business. However, just like anything else, E-commerce has its demerits as well as consumer uncertainties but nothing that can not be settle or elude by brilliant decision taking and business strategies.

REFERENCES

- [1] David Whiteley, E-Commerce, Strategy, Technologies and Application, Mc Graw Hill Education, New Delhi, 2017.
- [2] Dr. Nidhi Dhawan, A Handbook of E-Commerce, Sun Indis Publication, New Delhi, 2017.
- [3] Sushma Malik, Charul Nigam, JSR Publication House, New Delhi, 2019.
- [4] P.T. Joseph, S.J. E-Commerce An Indian Perspective, PHI Learning Pvt. Ltd. New Delhi, 2019.

Classification Techniques for Crop Disease Detection

Rupinder Kaur

Department of Computer Science
Sri Guru Granth Sahib World University
Fatehgarh Sahib, India
Rkaur340@gmail.com

Abstract - Production of crops with better quality is the necessary attribute for the economic growth of any country. The agricultural sector provides employment to many people and accounts for major portion of gross domestic product in many countries around the world. Therefore, for enhanced agricultural productivity the detection of diseases in crops at an early stage is quite significant. The traditional approaches for disease detection in crops required considerable amount of time, intense research, and constant monitoring of the farm. However, optimized solutions have been obtained over the past few years due to technological advances that have resulted in better yields for the farmers. Machine learning and image processing are used to detect the disease on the agricultural harvest. The image processing steps for crop disease identification include acquiring of images, pre-processing, segmentation and feature extraction. In this review paper, we focused mainly on the most utilized classification mechanisms in disease detection of crops such as Convolutional Neural Network, Support Vector Machine, K-Nearest Neighbor, and Artificial Neural Network. It has been observed from the analysis that Convolutional Neural Network approach provides better accuracy compared to the traditional approaches.

Index Terms - Image Processing, Machine Learning, Convolutional Neural Network, Support Vector Machine, K-Nearest Neighbor, and Artificial Neural Network

I. INTRODUCTION

Agriculture is an important source of livelihood in India. Majority of the country's population is directly or indirectly associated with the agricultural sector. Hence, producing high-quality agricultural yield is necessary to sustain the country's economic development. In order to obtain crops with better quality and productivity the farmers decide upon the right products by monitoring and controlling the necessary temperature, light and humidity requirements [1]. Furthermore, the agricultural industry has started to look for new ways to increase food production due to population growth, weather changes, and political instability. This has attracted the researchers to look for unique, resourceful, and reliable technologies that would help to enhance the productivity of agriculture. Still, there are challenges such as early identification of diseases in crops that the farmers are struggling with. To observe the type of disease on the crop's leaf through naked eyes is not possible all the time, so an

automated expert system that will help detect the disease timely would be quite useful. The advancement in technology, specifically the use of image processing in combination with the machine learning approach would help the farmers in terms of discovering the crop disease in the initial stages [2].

The input images are required to be of better quality so that images can be analyzed properly and correct classification and detection can be made. IoT sensor network has become popular in industry usage and normal day to day life to increase the crop yield production and its quality. Conventional techniques for detecting and plant diseases include direct visual analysis by visual identity of disease symptoms appearing on crop or by chemical strategies. These methods are time consuming and require a lot of people. AI approach helps to find the crop disease in the underlying stages. The researchers developed techniques to capture good quality images with professional cameras.

The work flow is detailed as follows. A brief introduction of the significance of crop disease detection is provided in the first section. The second section, the methodology for crop disease detection has been discussed.

The third section addresses the related work in the area and the classification techniques that have been used by the researchers. Finally, in last section conclusion is given which is followed by the references.

II. LITERATURE REVIEW

Many researchers have focused their work to enhance the accuracy of an automatic detection system for crop leaf diseases. This section discusses different techniques used for the classification of crop disease using various classifiers such as Convolutional Neural Network, Support Vector Machine, K-Nearest Neighbors and Artificial Neural Network.

A. Convolutional Neural Network (CNN)

Convolution Neural Networks are a class of deep feed forward neural networks that have the ability of processing multidimensional data. The purpose of CNN is to reduce images into an easier-to-process form, without compromising the features that are essential for getting a

good prediction.

There are different available architecture's for CNN such as AlexNet, GoogLeNet, VGGNet etc. Its growth has generated a lot of interest among researchers in various fields of computer science [9]. In agriculture, it has been used for the classification of diseases in crops.

The following work has been done in crop leaf detection using the CNN approach:

Mohanty et al.(2016) [10] have focused mainly on two CNN architectures i.e. AlexNet and GoogLeNet and the training mechanisms that have been used are transfer learning and training from scratch. The images used are from CropVillage dataset and about 54,306 images of different crops with 38 classes of diseases were taken. All the images in the dataset were downscaled into 256×256 pixels, and model optimization and predictions were performed on the resized images.

Lu et al. (2017) [11] have proposed a model for disease detection of rice using deep convolution neural network. 500 images captured with a camera from a field and 10 common diseases in rice crops have been detected. The results have been compared for different pooling (mean, max and stochastic pooling), different convolution filter sizes (5X5, 9X9, 16X16, 32X32) and for different algorithms (CNN, BP, SVM, PSO).

Gandhi et al.(2018) [12] proposed a system based on two different CNN architectures i.e. Inception v3 and MobileNets. 56,000 images with 38 classes of crops from the CropVillage dataset. A deep convolutional generative adversarial network (DCGAN) has been used for the augmentation of limited images in the dataset.

Ferentinos, K. P. (2018) [13] have used different architectures of CNN, such as AlexNetOWTBn, Overfeat, AlexNet, VGG, and GoogLeNet that trained using various parameters. The training and testing of these models were implemented using Torch7, which is a computational framework for machine learning. Around 87,848 images of CropVillage dataset having 25 crop species in 58 distinct classes of disease were used in this work.

Khamparia et al.(2019) [14] have integrated convolutional neural networks (CNN) for detection of diseases in crops. The authors have utilized a dataset with 900 images of three crops with five different types of diseases such as early blight and late blight for potato, leaf mould and yellow leaf curl for tomato and rust disease for maize crop. The convolution filters of size 2 X 2 and 3 X 3 have been used and analyzed accuracy varies for different convolution filters for different number of epochs. For loss reduction and improved accuracy while training, Adam optimizer has been used.

The Summary of work done for crop disease detection using the CNN technique is presented in below Table 1

Table 1: Summary of work done using CNN technique

Title	Technique(s)	Crops Used	Outcome/ Accuracy
Using Deep Learning forImage-Based Crop Disease Detection [10]	CNN	Apple, Blueberry, Cherry, Corn, Grapes, Orange, Peach, Potato, Strawberry, Tomato (Dicots)	The overall accuracy is obtained in the range of 85.53% to 99.34%. It has been observed that GoogLeNet performed better than AlexNet with an accuracy of 98.21%.
Identification of rice diseases using deep convolutional neural networks [11]	Deep-CNN	Rice (Monocot)	Maximum accuracy was achieved using stochastic pooling and a 16x16 filter size which was 95.48% and 93.29% respectively.
Crop Disease Detection Using CNNs and GANs as an Augmentative Approach [12]	CNN with GANs	Multiple Crops	In this the inception v3 model achieves an accuracy of 88.6% and MobileNets achieved an accuracy of 92%.
Deep learning models for crop disease detection and diagnosis [13]	CNN with different architecture	Apple, Blueberry, Corn, Cabbage, Cassava, Cherry, Peach, Strawberry, Tomato (Dicots)	The proposed model has an accuracy that varies from 97.26%- 99.47%. The highest accuracy was achieved by, Visual Geometry Group (VGG) convolutional neural network.

Seasonal crops disease prediction and classification using deep convolutional encoder network [14]	CNN and auto encoders	Potato, tomato (Dicots) maize (monocots)	This work attained 97.50% accuracy in 2×2 filter size of convolution in 100 epochs, whereas 100% accuracy corresponds to 3×3 filter size.
--	-----------------------	--	---

B. Support Vector Machine (SVM)

Support vector machine (SVM) is a type of learning algorithm that is based on structural risk minimization and is also used for classification and regression problems. It is designed in such a way as to maximize the classification boundaries so that two classes are separated as widely as possible.

The work done using SVM for the classification of diseases in crops is discussed below:

Badol et al. (2016) [18] have used a Linear Support Vector Machine for the classification of two types of diseases i.e. Downey and Powdery Mildew in grape leaves. The researchers used 137 images of grapes that they captured using a camera. Images are preprocessed to remove the noise using gaussian filter and thresholding is done to remove unwanted components. Then, segmentation of images is done using k-means and later features are extracted and the images are then fed to the SVM classifier.

Singh et al. in 2017 [19] have identified and classified pea rust disease that is caused by a fungus known as Uromyces fabae (Pers.) de Bary. About 500 images of pea crops were used that have been collected from Hill Agricultural Research and Extension Centre in Himachal Pradesh, India. Various steps of image processing were used and at the classification phase, SVM was utilized for disease detection.

Bhimte et al.(2018) [20] presented a model which detected Bacterial blight and Magnesium Deficiency in cotton crops. The dataset consisted of 130 images that were captured with a camera. Quality of images are improved using preprocessing techniques and then k- means clustering is used for segmentation. The images are then classified using SVM classifier after the features have been extracted using Gray Level Co-occurrence Matrix (GLCM).

Kumar et al.(2018) [21] introduced a new exponential spider monkey optimization (ESMO) technique of feature selection for identification of disease in crop leaves. For feature extraction subtractive pixel adjacency model (SPAM) technique is used. In this work 1000 images from CropVillage dataset have been used for detection of disease. To make classification among healthy and diseased leaves KNN, SVM, ZeroR, and LDA classifiers were used. After analysis, SVM

classifier performed better than the other classifiers.

Hossain et al.(2018) [22] proposed a system to identify two classes of diseases namely brown blight and algal in the tea crop. About 300 images of healthy and diseased tea crop were captured using a camera from Bangladesh Tea Research Institute. After preprocessing and feature extraction the data was fed to the SVM classifier for the accurate prediction of the disease..

In below Table 2, discussed the work that makes use of the SVM technique to detect crop disease.

Table 2: Summary of work done using SVM technique

Title	Technique(s)	Crops Used (Dicots or Monocots)	Outcome/ Accuracy
SVM Classifier Based Grape Leaf Disease Detection [18]	SVM	Grape (Dicot)	The system provided an average accuracy of 88.89 per cent for both Downey and Powdery grape leaf disease.
Support vector machine classifier based detection of fungal rust disease in Pea Crop (<i>Pisum sativum</i>) [19]	SVM	Pea (Dicot)	Proposed method can successfully diagnose and analyze disease with 89.60% accuracy.
Diseases Detection of Cotton Leaf Spot using Image Processing and SVM Classifier [20]	SVM	Cotton (Dicot)	The proposed system detected the disease in cotton crop with 98.46% accuracy.
Crop leaf disease identification using exponential spider monkey optimization [21]	SVM	Potato and Apple (Dicot)	An accuracy of 92.12% has been obtained using this method.
Recognition and Detection of Tea Leaf's Diseases Using Support Vector Machine [22]	SVM	Tea (Dicot)	The system achieved an overall accuracy of 93.33%.
Detection and Classification of Diseases of Banana Crop Using Local Binary Pattern and Support Vector Machine [23]	SVM	Banana (Monocot)	The maximum accuracy that the proposed work attained was 89.1% and 90.9%.

C. K-Nearest Neighbor (K-NN)

This classification scheme is statistical along with non-parametric and the weight is given corresponding to neighbors. Here, the classification is done based on the computed Euclidean distance metric.

This technique is used widely in areas such as text mining, pattern recognition, forecasting the trends in stock market and in agriculture for classifying various diseases in crops. The following work has been done using the K-NN approach for detecting crop diseases.

Parikh et al.(2016) [24] proposed a system that uses cascades of KNN classifiers and multiple training sets to successfully detect Grey Mildew, a fungal disease in cotton crops from unconstrained images. 130 images collected from Sardarkrushinagar Dantwada Agriculture University. The images are segmented, and the features extracted and KNN classifier is then used.

Ramcharan et al.(2017) [25] utilized transfer learning mechanism on the basis of SVM and KNN, by utilizing the convolutional layers of existing trained Inception v3 model. The classification of the disease has been done in three specific ways; SVM, KNN and original softmax layer of inception v3 model.

Suresha et al.(2017) [26] presented a system for detecting diseases such as Blast and Brown Spot in paddy crop. About 300 images taken from a camera in paddy fields of Shivamogga district in Karnataka state have been used in this paper. For segmentation Otsu technique have been

Business Intelligence with Intelligent Computing

used along with global threshold. Connected component has been used for feature extraction and classification is done using the KNN technique.

Hossain et al.(2019) [27] have considered crop diseases such as alternaria alternata, anthracnose, bacterial blight, leaf spot, and canker of crops and used K-NN to classify them. The dataset comprises of 237 leaf images acquired from the Arkansas crop disease database. The features of crops have been extracted using the GLCM technique. To prevent overfitting, the 5-fold cross validation was applied on the training dataset.

Abdulridha et al.(2019) [28] developed an automatic early identification of diseases such as laurel wilt, phytophthora root rot (Prr), and deficiency of iron and nitrogen in avocado crops. The images were acquired using cameras and Multilayer perceptron (MLP) and K-nearest neighbor (K-NN) classification techniques were used. The table 3 given below, summarizes the work that uses KNN technique for various crop disease detection.

Table 3: Summary of work done using KNN technique

Title	Technique(s)	Crops Used (Dicots or Monocots)	Outcome/ Accuracy
Disease Detection and Severity Estimation in Cotton Crop from Unconstrained Images [24]	KNN	Cotton (Dicot)	Accuracy of 82.5% was achieved using this system.
Deep learning for image-based cassava disease detection [25]	KNN and SVM	Cassava (Dicot)	The obtained accuracy is 73% by applying KNN, and by using SVM 91% accuracy was achieved.
Recognition of Diseases in Paddy Leaves Using kNN Classifier [26]	KNN	Paddy (Monocot)	With this method the authors achieved an overall accuracy of 76.59%.
A Color and Texture Based Approach for the Detection and Classification of Crop Leaf Disease Using KNN Classifier [27]	KNN	Various Crop leaves	The classification performance of KNN on crop leaf disease provides 96.76% accuracy. This approach provides better results compared to some existing methods.
A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses [28]	KNN and MLP	Avocado (Dicot)	In all cases, the MLP approach obtained higher classification values than the KNN and sometimes reached up to 98%.

REFERENCES

1. A. Verma, "Precision Agriculture using Internet of Things: An Overview," in *International Journal of Management, IT and Engineering*, vol. 9, no. 7, pp. 298-303, July 2019.
2. M. Jhuria, A. Kumar and R. Borse, "Image processing for smart farming: Detection of disease and fruit grading," *2013 IEEE Second International Conference on Image Information Processing (ICIIP- 2013)*, Shimla, 2013, pp. 521-526. DOI: 10.1109/ICIIP.2013.6707647
3. N. N. P., C. K. A. and A. N. P., "Automatic classification of monocots and dicots using SVM classifier," *International Conference on Information Science (ICIS)*, Kochi, 2016, pp.28-31. DOI: 10.1109/INFOSCI.2016.7845295
4. S. Patil and A. Chandavale, (2015), "A survey on methods of crop disease detection," in *International journal of Science and Research (IJSR)*, 4(2), pp.1392- 1396.
5. B. S. Prajapati, V. K. Dabhi and H. B. Prajapati, "A survey on detection and classification of cotton leaf diseases," *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, 2016, pp. 2499-2506. DOI: 10.1109/ICEEOT.2016.7755143
6. Anup Vibhute and SK Bodhe (2012). "Applications of image Processing in Agriculture: A Survey," in *International journal of Computer Applications (095-8887) Volume 52- No 2, August 2012*. DOI:10.5120/8176-1495
7. Shashank Chaudhary, Upendra Kumar, Abhishek Pandey (2015), "A Review: Crop Crop disease detection using image processing," in *International Journal of Innovative Technology and Exploring Engineering. Volume-8, Issue-7 May, 2019*.
8. S. D. Khirade and A. B. Patil, "Crop Disease Detection Using Image Processing," in *International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 768-771. DOI: 10.1109/ICCUBEIA.2015.153
9. Aditya Khamparia, Aman Singh, Ashish Kr. Luhach, Babita Pandey, Devendra K. Pandey (2019), "Classification and Identification of Primitive Kharif Crops using Supervised Deep Convolutional Networks," in *Sustainable Computing: Informatics and Systems*. 2210-5379. <https://doi.org/10.1016/j.suscom.2019.07.003>
10. S. Mohanty, D. Hughes, M. Salathe, (2016), "Using Deep Learning for Image-Based Crop Disease Detection," *Frontiers in Crop Science*, Volume 7, <https://doi.org/10.3389/fpls.2016.01419>
11. Y. Lu, S. Yi, N. Zeng, Y. Liu, Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing* 267 (2017) 378–384. <https://doi.org/10.1016/j.neucom.2017.06.023>
12. R. Gandhi, S. Nimbalkar, N. Yelamanchili and S. Ponkshe, "Crop disease detection using CNNs and GANs as an augmentative approach," *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, Bangkok, 2018, pp. 1-5. DOI: 10.1109/ICIRD.2018.8376321
13. Ferentinos, K. P. (2018), "Deep learning models for crop disease detection and diagnosis," in *Computers and Electronics in Agriculture*. Vol.145, pp.311-318, <https://doi.org/10.1016/j.compag.2018.01.009>
14. A. Khamparia, G. Saini, D. Gupta, A. Khanna, S. Tiwari, V.H.C. Albuquerque, (2019), "Seasonal crops disease prediction and classification using deep convolutional encoder network," *Circuits, Systems, and Signal Processing*, 39, 818–836. <https://doi.org/10.1007/s00034-019-01041-0>
15. Kamal, K.C., Zhendong Yin, Mingyang Wu, and Zhilu Wu, (2019), "Depthwise Separable Convolution Architectures for Crop Disease Classification" *Computers and electronics in agriculture* <https://doi.org/10.1016/j.compag.2019.104948>
16. G. Geetharamani, A. Pandian, (2019), "Identification of crop leaf diseases using a nine-layer deep convolutional neural network," in *Computers & Electrical Engineering*. Vol. 76, pp. 323-338, <https://doi.org/10.1016/j.compeleceng.2019.08.010>
17. R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, R. Menaka, (2020), "Attention embedded residual CNN for disease detection in tomato leaves" in *Applied Soft Computing*, Vol. 86. <https://doi.org/10.1016/j.apsoftec.2020.105440>

- <https://doi.org/10.1016/j.asoc.2019.105933> 549-557.
18. P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," *2016 Conference on Advances in Signal Processing (CASP)*, pp. 175-179. DOI: 10.1109/CASP.2016.7746160
19. K. Singh, S. Kumar, P. Kaur, "Support vector machine classifier based detection of fungal rust disease in Pea Crop (*Pisum sativum*)," *Int. j. inf. tecnol.* 11, 485–492 (2019). <https://doi.org/10.1007/s41870-018-0134-z>
20. N. R. Bhimte and V. R. Thool, "Diseases Detection of Cotton Leaf Spot Using Image Processing and SVM Classifier," *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, pp. 340-344. DOI: 10.1109/ICCONS.2018.8662906
21. S. Kumar, B. Sharma, V.K. Sharma, H. Sharma, J.C. Bansal, (2018), "Crop leaf disease identification using exponential spider monkey optimization," *Sustainable computing: Informatics and systems*. <https://doi.org/10.1016/j.suscom.2018.10.004>
22. S. Hossain, R. M. Mou, M. M. Hasan, S. Chakraborty and M. A. Razzak, "Recognition and detection of tea leaf's diseases using support vector machine," *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, Batu Ferringhi, 2018, pp. 150-154. DOI: 10.1109/CSPA.2018.8368703
23. A. Aruraj, A. Alex, M. S. P. Subathra, N. J. Sairamya, S. T. George and S. E. V. Edwards, "Detection and Classification of Diseases of Banana Crop Using Local Binary Pattern and Support Vector Machine," *2019 2nd International Conference on Signal Processing and Communication (ICSPC)*, Coimbatore, India, 2019, pp. 231-235. DOI: 10.1109/ICSPC46172.2019.8976582
24. A. Parikh, M. S. Raval, C. Parmar and S. Chaudhary, "Disease Detection and Severity Estimation in Cotton Crop from Unconstrained Images," *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Montreal, QC, 2016, pp. 594-601. DOI: 10.1109/DSAA.2016.81
25. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, D.P. Hughes, (2017), "Deep learning for image-based cassava disease
26. M. Suresha, K. N. Shreekanth and B. V. Thirumalesh, "Recognition of diseases in paddy leaves using knn classifier," *2017 2nd International Conference for Convergence in Technology (I2CT)*, Mumbai, 2017, pp. 663-666. DOI: 10.1109/I2CT.2017.8226213
27. E. Hossain, M. F. Hossain and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Crop Leaf Disease Using KNN Classifier," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh, 2019, pp. 1-6. DOI: 10.1109/ECACE.2019.8679247
28. J. Abdulridha, R. Ehsani, A. Abd-Elrahman, Y. Ampatzidis, (2019), "A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses," *Computers and electronics in agriculture*, 156,
- <https://doi.org/10.1016/j.compag.2018.12.018>

Title : Role of Digitization in Education

Neha Dhiman¹

Assistant Professor

Department of Computer Science and Applications

dhiman.neha2296@gmail.com

Kamini²

Assistant Professor

Department of Computer Science and Applications

sharmakamini253@gmail.com

Abstract

Digitization impact a great influence on education. It replaced the traditional system of education such as chalk and text book. Smart lessons with audio and video are replacing traditional classrooms with blackboards. Teachers are started giving their lessons in ppt. This technology made a information more accessible from anywhere in the world. It refers to use the technology like internet, mobile devices, software for teaching the students. Digital education is in current trend for providing compulsory education to the students. In today's time students have not need to depend upon the text books to learn they can easily learn from digital educational tools. Due to COVID 19 digital education has increased among the students. Now we are moving from physical system to virtual mode of teaching. With the change in education system students are broadening their knowledge and their confidence is also increasing.

I. INTRODUCTION

The process of transforming information into a digital (i.e. computer-readable) format is known as digitization. In today world education is very important and it is our basic rights, it also leads to development of economy. Now days teaching style has changed , if we go back education was given in gurukul where the students lived there in discipline and in strict rules. After that education has changed , it comes to the classrooms teaching on blackboard. Students can use text books for further reading. As the time is changing education system

has also changed .Blackboard has replaced with the smart board. They are connected with the projector. Teachers started giving their lectures with the help of ppt.

We are living in twenty-first century, a time when technology has no limitations. This is a period of rapid development, with technology is increasing. Smartphones, laptops, and tablets are now common place. During this phase, the educational system is growing for the betterment of students, as this generation's students were not born to be limited by the simple learning; their curiosity is enormous and cannot be satisfied by educational systems that were designed previously. We would deprive our students of their tomorrow if we continued to teach them as we did yesterday. Our outdated educational system is unable to compete in the twenty-first century. As a result, we feel forced to use digitization in our work.

The main objectives of digitization in education is to improve the skills and confidence levels in educators. Digital learning increase your critical thinking and also which helps in the growth of reasoning skills. Access to education and information is improved through digital learning. Digitization in classroom is growing day to day. In today's time teachers and students are using technology in aid of learning. Learning digital skills are necessary for making professional skills.

II.REVIEW OF LITERATURE

Digitization in education helps the students and teachers in effective learning with the technology.

Wireless Communication is commenced over the 100 years ago. Radio is among from the technology and the messages is passed in the air, on radio waves. Now the technology is moved to the television, computer, CDs, CDROM. Urgent messages are send quickly with the help of this technology.[1]

With the change of time technology has plays a great role in education. Today's technology taking place of traditional systems. Text books are taking place of e-books. Blackboards are replacing with smart classrooms. Many softwares are used to converting the syllabus in digital form. There is a big revolution in the education from past to present time. In past time there is gurukul system, students lived there, in strict rules. After that gurukul system is changed to classroom where teachers taught the students on blackboard. As the improvement of technology smart rooms with projector are taking place of blackboard.[2]

Online education is much increased in the covid time. Offline education is shift to online platform. But this is only achieved if technology is achieved to every person. Many students are facing internet connectivity issues during their papers. Infrastructure should be available to all. Some people are against of it because there is lack of face to face relationship.[3]

There is much growth in education sector of India. Cloud learning is very beneficial for the students. They can learn from anywhere and any time. It can provide a scalability and it has reduce the cost.[4] Cisco's India has developed a Live Remote Education Platform for Indian Classrooms. It helps in skills development. Dwara is a platform which is provided by cisco education enabled.[5].

Distance learning is a part of digitization. It is possible only if there are smart classrooms. This is achieved with the combination of internet. Any one can learn online, in free of cost. Many online platforms are available for free learning. There are millions of courses are provided by the educators. If you want to do preparation for any entrance test you can learn from online platform in free of cost.[6]

In education India has reached at highest level. The government of India has introduced PM e-VIDYA programme in May 2020 for e-

learning. This can make e-learning easy for teachers and students. Government of India has provides many channels for learning.[6]

The graphical user interface provides an attractive platform for learning. The study material is provided in text and graphics.[1]

III. INITIATIVE OF DIGITAL INDIA

The Indian administration additionally released the 'Digital India' initiative in July 2015, to bolster online infrastructure and make bigger internet accessibility amongst citizens (for example, connecting rural regions to high-pace internet networks). As component of 'Digital India' initiative, the administration additionally commenced e-Education initiative to offer online education in faraway and concrete regions the usage of smartphones, apps and internet services.

COVID-19 pandemic had in quick disrupted the university training system in India. Since the primary lockdown in March 2020, all the states have responded to the COVID-19 situation with speed and innovation. The - modern day digital reassets have been leveraged to keep continuity in students' training. The various initiatives under this programme are as under:

A. SWAYAM:

The 'Study Webs of Active Learning for Young Aspiring Minds' (SWAYAM) is an integrated platform for presenting on line guides and cover-up college (9th to 12th) to Post Graduate Level. Till now, 2769 MOOCs (Massive Open Online Courses) were presented on SWAYAM, in which approximately 1.02 crore college students have enrolled to numerous guides until date. The online guides are getting used now no longer most effective through the scholars however additionally through the academics and non-pupil learners, with inside the shape of lifelong learning. It can be accessed on swayam.gov.in NCERT (National Council of Educational Research and Training) has been growing route modules for MOOCs for college schooling machine in 12 concern areas (Accountancy, enterprise studies, biology, chemistry, economic, history, geography, mathematics, physics, political science, psychology

and sociology) for training IX-XII. Twelve (12) guides have been released with inside the first cycle. Nearly 22,000 college students have been registered on numerous guides. Twenty (20) guides have been released with inside the 2nd cycle. Nearly 33,000 college students have been registered.

B. National Digital Library (NDL):

The National Digital Library of India (NDL) is a project to develop a virtual storage framework for learning resources with a single search bar. More than 3 million digital resources are available through NDL. The content covers almost all major areas of education, including lifelong learners, and learners at all major levels. More than 50,000 students have enrolled in NDL, of which approximately 20,000 are active users. NDL is also available through the mobile app. It can be accessed at ndl.gov.in.

C. Spoken Tutorials:

10-minute audio-video tutorials on open source software designed to improve student employability. Designed for self-study with audio and online versions dubbed in 22 languages. C, C++, Java, PHP, Python, PERL, Scilab, OpenFOAM, OpenModelica, DWSIM, LibreO, and more. Spoken Tutorials are effectively designed to teach novice users without the support of a Physical teacher.

D. Free and Open Source Software for Education (FOSSEE):

FOSSEE is a project that promotes the use of open source software in educational institutions (<http://fossee.in>). This is achieved through textbooks such as oral tutorials, documentation such as textbook companions, and awareness programs such as conferences, educational workshops and internships. A Textbook Companion (TBC) is a set of codes for a standard textbook example solved. About 2,000 college students and teachers participated in the event, and about 1,000 TBCs were created in Scilab and made available for free download.

E. E-Yantra:

eYantra is a project that provides effective education in embedded systems and robotics to Indian Institutes of Technology. Teacher and student training is conducted through seminars and participants learn the basics of embedded systems and programming. More than 275 universities across India have leveraged this initiative. All projects and code are available as open source content on the eYantra website www.eyantra.org.

IV. ADVANTAGES OF DIGITAL EDUCATION IN INDIA

A. Individual Learning Experience:

A major drawback of traditional education systems is that many students experience a lack of interest when they cannot keep up with the rest of the class. Modern digital formats allow teachers to customize learning materials according to the pace of learning and individual abilities. With the digitization of the education system, the impact of educational programs is growing.

B. Unlimited Information:

The world of the Internet is vast, full of information, and most of it is free. The advent of digital education has made it possible for students to explore and use this knowledge base. Previously, students relied on limited sources of information, but now with the popularization of digital education systems, the lack of necessary information is no longer a barrier to knowledge exploration.

C. Highly participatory learning:

While traditional education systems offer limited opportunities for participation because they include limited elements such as textbooks, instructors, and handwritten notes as strengths in the workplace, digital education systems provide a range of learning opportunities. Unlimited availability of resources makes every session very innovative and fun. Interactive and fun learning sessions attract more students.

D. Ease of Sharing:

Whereas traditional education systems relied heavily on students to keep in thick

notebooks handwritten notes with information provided by teachers in the classroom or through extensive research in libraries, modern digital education systems are changing everything. . Information can now be stored and shared with a single click, saving students a lot of time and manual labor.

V.DISADVANTAGES OF DIGITAL EDUCATION IN INDIA

A. Lack of social interaction:

E-learning is one of the causes of social isolation as it is no longer face-to-face with teachers and classmates.

B. Deception Is Inevitable:

E-learning includes assessment just like a regular classroom. However, you will not be supervised by a teacher or proctor during the exam. Online students can easily share answers knowing that no one is watching.

C. Technology sometimes fails:

Many innovations, especially gadgets, tend to experience some bugs. For example, many servers suffer from problems and downtime, making learning difficult. Connection issues are also pretty common. These issues take time to resolve. So learning is delayed.

D. Focuses more on theory:

Most of the time you will listen to podcasts, watch videos, and watch presentations. There is no manual experience like experimentation. The most modern way to learn, all you need is a gadget and an internet connection. However, this is not for everyone. Knowing the pros and cons of e-learning can help you decide if it's right for you or not.

VI.CONCLUSION

This paper reviews the impact of digitization in education. It tells how the technology is taking place of traditional system. Any one from anywhere can learn through online system. Students are learning from laptops, computer and tablet. But

this all can be achieved with the help of good internet connection. Many initiative has also taken by government of India. TV channels have launched by government in free of cost which is very beneficial for students.

VII. REFERENCES

- [1] G.N. Wikramanayake “Impact of Digital Technology on Education”
- [2] Dr. Himanshu Rastogi, “Digitalization of education in India – An analysis”
- [3] Tudorel Toader,Marieta Safta,Cristina Titirisca and Bogdan Firtescu, “Effects of Digitalisation on Higher Education in a Sustainable Development Framework-Online Learning Challenge during the COVID-19 Pandemic.”
- [4] Nivedita Jha,Veena Shenoy, “Digitization of Indian Education Process:A Hope or Hype”
- [5]<https://www.indiatechonline.com/it-happened-in-india.php?id=996>
- [6] Tanya Chatwal, “Digitalization of Higher Education in India:A Boom or A Bane”

RFID Technology And Internet of Things (IOT)

Manpreet Kaur¹

¹ Student, BBSBEC

¹ dhalivalmanpreet209@gmail.com

Abstract— Radiofrequency identification system (RFID) is an automatic technology and aids machines or computers to identify objects, record metadata or control individual target through radio waves. Connecting RFID reader to the terminal of Internet, the readers can identify, track and monitor the objects attached with tags globally, automatically, and in real time, if needed. This is the so-called Internet of Things (IOT). RFID is often seen as a prerequisite for the IOT. This paper introduces the technologies of RFID and IOT, discusses the applications and challenges of RFID technology used in IOT.

Keywords- *RFID technology; application; Internet of Things*

I. INTRODUCTION

Radio frequency identification system (RFID) is an automatic technology and aids machines or computers to identify objects, record metadata or control individual target through radio waves. The RFID technology was first appeared in 1945, as an espionage tool for the Soviet Union, which retransmitted incident radio waves with audio information. A typically RFID system is consisted of tags (transmitters/responders) and readers (transmitters/receivers).

RFID is one of the big opportunities in information technology, which will change the world broadly and deeply. When the RFID readers abide by appropriate communication protocols are connected to the terminal of Internet, the readers distributed throughout the world can identify, track and monitor the objects attached with tags globally, automatically, and in real time, if needed. This is the so-called Internet of Things (IOT). The IOT refers to uniquely identifiable objects (things) and their virtual representations in

an Internet-like structure. RFID is often seen as a prerequisite for the IOT. If all objects of daily life were equipped with radio tags, they could be identified and inventoried by computers.

This paper introduces the primary concepts and technologies of RFID and IOT, discusses the applications and challenges of RFID technology used in IOT.

II. IOT AND RFID TECHNOLOGY

A. Internet of Things

Internet of Things (IOT) is a global network infrastructure, linking physical and virtual objects through the exploitation of data capture and communication capabilities. It will offer specific object identification, sensor and connection capability as the basis for the development of independent cooperative services and applications. These will be characterized by a high degree of autonomous data capture, event transfer, network connectivity and interoperability. The IOT system architecture is generally divided into three layers: the perception layer, the network layer, and the service layer.

Perception layer: It is the information origin and the core layer of IOT. All kinds of information of the physical world used in IOT are perceived and collected in this layer, by the technologies of sensors, wireless sensors network (WSN), tags and reader-writers, RFID system, camera, global position system (GPS), intelligent terminals, electronic data interface (EDI), objects, and so like.

Network layer: This layer, also called transport layer, including access network and core network, provides transparent data transmission capability.

Information from perception layer can be sent to the upper layer. At the same time, this

layer provides an efficient, reliable, trusted network infrastructure platform to upper layer and large scale industry application.

Service layer: This layer, also called application layer, includes data management sub-layer and application service sub-layer. The data management sub-layer provides processing complex data and uncertain information, such as restructuring, cleaning and combining, and provides directory service, market to market (M2M) service, Quality of Service (QoS), facility management, geomatics, etc. by service oriented architecture (SOA), cloud computing technologies, and so on. The application service sublayer transforms information to content and provides good user interface for upper level enterprise application and end users, such as logistics and supply, disaster warning, environmental monitoring, agricultural management, production management, and so forth.

B. RFID System

RFID systems are composed of three main components: RFID tags, reader, application system.

RFID tags: also known as transponders (transmitter/ responder), are attached to the objects to count or identify. Tags could be either active or passive. Active tags are those that have partly or fully battery powered, have the capability to communicate with other tags, and can initiate a dialogue of their own with the tag reader. Passive tags, on the other hand, do not need any internal power source but are powered up by the tag reader. Tags consist mainly of a coiled antenna and a microchip, with the main purpose of storing data.

Reader: also known as transceiver(transmitter/receiver) made up of a radio frequency interface (RFI) module and control unit. Its main functions are to activate the tags, structure the communication sequence with the tag, and transfer data between the application software and tags. **Application system:** also called data processing system, which can be an application or database, depending on the application. The application software

initiates all readers and tags activities. RFID provides a quick, flexible, and reliable way for electronically detecting, tracking and controlling a variety of items. RFID systems use radio transmissions to send energy to a RFID tag while the tag emits a unique identification code back to a data collection reader linked to an information management system. The data collected from the tag can then be sent either directly to a host computer, or stored in a portable reader and up-loaded later to the host computer.

C. RFID Tags

RFID tags come in many different shapes, sizes, and capabilities. When an RFID solution is designed, the solution's architect must take into account both business and technology requirements before choosing the type of RFID tag to use. All RFID tags have the following essential components in common: antenna, integrated circuit, printed circuit board (or substrate). The main responsibility of antenna of RFID tag is to transmit and receive radio waves for the purpose of communication. The antenna is also known as the coupling mechanism, which can transform the energy in the form of electromagnetic radiation. This is the way the tag and reader communicating each other. In a suitable environment and proximity to an RFID reader, the antenna can collect enough energy to power the tag's other components without a battery. The integrated circuit (IC) is a packaged collection of discrete components that provide the brains for the tag. The IC in a RFID tag is much like a microprocessor found in any cellular phone or computer, but it is usually not very sophisticated. For many RFID tags, the IC component has only a single purpose, to transmit the tag's unique identifier (ID). If the tag has any peripheral components, the IC is also the master controller that is responsible for gathering any extra information and transmitting it along with the tag's ID. The printed circuit board (PCB) is the material that holds the tag together. The circuit board may be rigid or flexible, and is composed of many different types of materials, depending on the type and

purpose of the tag. For example, tags that are used for tracking components on an assembly line where extremely high temperatures may be encountered would tend to be much more rigid and are usually placed inside a protective enclosure.

D. RFID Reader

RFID readers are also referred to as interrogators because they query tags as the tags enter their read range. The reader is responsible for orchestrating the communication with any tags in its read range and then presenting the tags' data to an application that can make use of the data. Readers in all systems can be reduced to two fundamental functional blocks: the control system and the high frequency (HF) interface, consisting of a transmitter and receiver, as shown in. The entire system is controlled by an external application via control commands. The reader's HF interface performs the following functions: (a) generation of high frequency transmission power to activate the transponder and supply it with power; (b) modulation of the transmission signal to send data to the transponder; (c) reception and demodulation of HF signals transmitted by a transponder. The reader's control unit performs the following functions: (a) communication with the application software and the execution of commands from the application software; (b) control of the communication with a transponder (master-slave principle); (c) signal coding and decoding. In more complex systems the following additional functions are available:

(d) execution of an anti-collision algorithm; (e) encryption and decryption of the data to be transferred between transponder and reader; (f) performance of authentication between transponder and reader. The control unit is usually based upon a microprocessor to perform these complex functions.

Cryptological procedures, such as stream ciphering between transponder and reader, and also signal coding, are often performed in an additional ASIC (application specific

integrated circuit) module to relieve the processor of calculation intensive processes.

III. APPLICATIONS OF RFID TECHNOLOGY

The functions of RFID system generally include three aspects:

monitoring, tracking supervising. Monitoring generally means to be aware of the state of a system, by repeated observing the particular conditions, especially to detect them and give warning of change. Tracking is the observing of persons or objects on the move and supplying atimely ordered sequence of respective location data to a model. Supervising is the monitoring of the behaviors, activities, or other changing information, usually of people. It is sometimes done in a secret or inconspicuous manner. The

RFID applications are numerous and far reaching. The most interesting and successful applications include those for supply chain management, production process control, and objects tracking management.

^¾ Logistics and Supply ^¾ Manufacturing
^¾ Agriculture Management ^¾ Health Care and Medicine
^¾ Marine Terminal Operation ^¾ Military and Defense ^¾ Payment Transactions
^¾ Environment Monitor and Disaster Warning ^¾ Transportation and Retailing ^¾ Warehousing and Distribution Systems
^¾ Other applications in many walks of

life businesses For instance, in Manufacturing, RFID technology offers a number of applications in the automotive industry. A RFID-based antitheft vehicle immobilizer is a protective device installed in many cars. RFID also holds great promise for the assembly and manufacturing processes of automobiles, in particular, for flexible and agile production planning, spare parts, and inventory management. RFID technology not only helps to automate the whole assembly process in which a significant reduction in cost and shrinkage can be achieved, but it also offers improved services to automobile users

that include more efficient replacement part ordering and automated generation of maintenance reminders. The benefits that RFID offers to the automotive

industry, both to the production process as well as to end users, are visibility, traceability, flexibility, and added security.

IV. CHALLENGES OF RFID TECHNOLOGY

Although promising, RFID is not without its challenges, which arise from both a technological and usage point of view.

A. Collision Problems

Communication between tags and readers are inherently susceptible to electromagnetic interference. Simultaneous transmissions in RFID lead to collisions as readers and tags typically operate on a same wireless channel. Therefore, efficient anti-collision protocols for identifying multi-tags simultaneously are of great importance for the development of large-scale RFID applications. Many anti-collision protocols for RFID tag identification have been proposed, such as query tree protocol (QT), binary tree protocol (BT), frame slotted ALOHA protocol (FSA), etc., but almost all known protocols exhibit an overall identification efficiency small than 50%. Besides, uniform IDs distribution has always been assumed in the past. Furthermore, it is very useful for pointing out the best performing features of RFID tag identification protocols, and for designing new and better protocols. We present a novel and efficient anti-collision protocol for RFID tag identification, i.e., collision tree protocol (CT), which outperforms all the other anti-collision protocols proposed so far.

B. Security and Privacy

Concerns

Security and privacy issues of RFID tags can effect both organizations and individuals. Unprotected tags may be vulnerable to eavesdropping, traffic analysis, spoofing or denial of service and many more. Even unauthorized readers can affect the privacy by accessing tags without enough access control. Even if the tag content is secure then also it can be tracked by the predictable tag responses; “location privacy” can be affected by a traffic analysis attack. Attacker can also

threaten the security of systems, which depends on RFID technology through the denial of service attack. Due to its cost and resource constraint limitations, RFID system does not have a sufficient security and privacy support. Many researcher and scientist work to implement low cost security and privacy protocol to increase the applicability. Lots of lightweight solutions have been proposed for RFID, but they are still expensive and vulnerable to the security and do not fully resolve the security issues. So there is a good research scope in the field of designing an ultralightweight cryptographic protocol for low-cost RFID system.

C. Other Challenges

Three other issues also are mainly holding back RFID's widespread adoption. The first one is the cost. The RFID tags are still more expensive than printed labels. The second issue is design. It is still needed to engineer tags and readers so that they guarantee highly reliable identification. Another challenge to RFID is its integration into existing systems. It is significant to develop effective RFID middleware which used to link new RFID systems into existing back-end infrastructures. Despite these challenges, it's only a matter of time before these issues could be solved. RFID's potential benefits are large, and many novel applications will be seen in the future, even some of which can not begin to imagine.

V. CONCLUSION

The IOT uses a variety of information sensing identification device and information processing equipment, such as RFID, WSN, GPRS, etc. combining with the Internet to form an extensive network in order to informationize and intelligentize the entities or objects. This paper analyzes the applications and challenges of RFID technology, which is the important and foundational component of IOT.

References

- [1].Xiaolin Jia, Quanyuan Feng, Taihua Fan, Quanshui Lei, "RFID Technology and its Applications In Internet Of Things", Southwest University of Science and Technology, July 2018.
- [2]. X.L. Jia, Q.Y. Feng, C.Z. Ma, "An efficient anti-collision protocol for RFID tag identification," IEEE Communications Letters, vol.14, no.11, pp.1014-1016, 2010. [3].K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication, New York: Wiley, 2010.
- [4].D.Culler, "10 Emerging Technologies That Will Change the World," Technology Review, pp.33-49, Feb., 2003. [5].R.K.Pateriya, S.Sharma, "The Evolution of RFID Security and Privacy: A Research Survey," 2011 International Conference on Communication Systems and Network Technologies (CSNT), pp.115-119, June 2011.

DIGITALISATION OF RURAL PUNJAB

Dr. Sumeet Kaur¹, Sukhvir Kaur²

Abstract

Indian economy is an agrarian economy with majority of its population living in rural areas. Within the rural sector, the population is mainly engaged in agriculture and allied activities. The standard of living, education level, internet self efficacy level, adoption of new technology level, economic growth level is different and low in rural areas as compared to urban areas. The people in rural areas are less updated as far as all the above mentioned things are concerned.

Although in the last few decades, India has faced continuous innovations and information and communication technologies revolution, their penetration levels in rural areas has been quite low. Various efforts are being made by the government to make rural India more digitally active. ‘Digital India’ programme is one such initiative of Government of India for making rural India digitally literate.

In this paper, an attempt has been made to study the concept of Digital India, the changes made by the scheme in rural Punjab. An attempt has been made to understand the need and benefits of this scheme and how this scheme is working in changing the face of rural Punjab. To study the real impact of this programme, this paper brings out an empirical study in district Roopnagar of Punjab. By emphasizing on the services provided by this scheme, this paper elaborates the extent to which rural Punjab is really benefitted by this scheme and also brings out the shortcomings of this programme.

Big Data: Possible Revenue Streams For Social Media Platforms

Jaswinder Kaur

Assistant Professor, B.A.M. Khalsa College Garhshankar

Abstract

Social media initially devoted for connecting people has evolved into completely new avatar. The major social media platforms like Facebook, Instagram, Twitter, whatsapp etc have entirely transformed in the last decade and taken the world by storm. The extensive usage of social media naturally includes large amounts of information regarding users' unadulterated feelings and thoughts documented in a publicly visible platform. The collection of this information provides naturally-occurring and publicly-visible big data, that can potentially be analysed for insights that improve decisions and give confidence for making strategic business moves and applied to improve global business environment. Kapis analysis shows that there are more than 4.62.billion social media users around the globe in January 2022 which is 58.4 percent of the total world population. With more than 2 billion Facebook users ,1.3 billion Instagram ,1 billion engaged on Twitter or LinkedIn, above 540 million connected on Google+ define the massive expansion, growth and demand of these social media sites. In this paper we identify new revenue streams for social media service providers. We identify three possibly ways to extend existing social media business models. This study will lead the way to further research especially in terms of Big Data analytics with new e-commerce trends.

Keywords: *Big Data, big data analytics, Social Media, Advertising, Business Models.*

Artificial Intelligence in Education

Novedeep Jeerh

*Student, Department of Computer Science and Applications,
B.A.M. Khalsa College, Affiliated to Panjab University, Chandigarh*

Abstract

Artificial Intelligence is the potential of a computer, or a robot supervised by a computer to do various tasks that are typically done by individuals since they require human intelligence and judgment. Artificial Intelligence is universally implemented to provide personalized recommendations to humans such as optimizing resources, planning inventory, logistics, etc. Artificial Intelligence plays a crucial role in the education sector in the matter of teaching and learning. In the schools and colleges educators use artificial intelligence to ameliorate the teaching techniques for best consequences. This paper mainly focusses on the applications used in three education sectors via artificial intelligence. In Primary Education most common used applications are Inferkit, Petalica Paint, Pix2Pix, GauGaN, Art breeder whereas, in secondary education Brainly, Thinkster Math, Querium, Cargenie Learning, Quizlet are used. In Higher Education Knewton, Knowji, Cognii, TED, Evernote are widely used. The issues faced by three education sectors such as costs and scalability, ethics and privacy, lack of AI experts as well as lack of actionable guidelines for educators are also mentioned. It is concluded that artificial intelligence in education enables that AI technologies are not only for smart learning but also for the tutoring systems and social robots; moreover, there are plenty of other technologies which also contribute to this sector such as virtual facilitators, open learning environments, learning management systems, and learning analytics as well.

Preprocessing in Automatic Traffic Light Detection System for Self-Driving Vehicles

Sarita¹

Research Scholar,

Panjab University, Chandigarh.

Anuj Kumar²

Assistant Professor,

Panjab University, Chandigarh.

Abstract

Self-Driving cars are the need of the future. Google, Tesla, Toyota, Mercedes Benz and motor industrialist around the globe are in a race to built more and more efficient self-driving vehicles. The self-driving vehicles are required to detect traffic light accurately to avoid any mishap on the road. Traffic Light Recognition system (TLRS) helps detecting traffic lights in varying illumination and lightening conditions. A TLR system contains various stages like preprocessing, feature extraction, post-processing. However, for accurate detection better quality images are required and thus preprocessing is done on raw images to provide system with high-quality images. In this paper we focused on various preprocessing techniques and implemented some of them on our local dataset collected from the roads of Chandigarh.

Energy Aware Virtual Machine Allocation by Fuzzy Artificial Bee Colony algorithm in cloud data centers.

KUMAR Vaneet¹, JINDAL Balkrishan²

¹Research scholar , Computer Science and Engineering Section, Punjabi University, Patiala 147002, India;

²Associate Professor, Computer Engineering Section, Yadvindra College of Engineering, Punjabi University Campus, Talwandi Sabo 151302, India

¹aggarwal.vkg@gmail.com

²balkrishan_76@rediffmail.com

Abstract

Cloud computing is based on the concept of distributed computing that has transformed the dream of “computing as a utility” into reality. The increasing demand of cloud resources leads to design huge data centers globally. The large size data centers consume huge energy and most of the energy consumption occurs due to lack of resource utilization. With the proper virtual machine allocation resources can be optimization. This paper proposes virtual machine placement algorithm called fuzzy based artificial bee colony (FABC) algorithm. In FABC virtual machines allocate to host according to ABC where fuzzy logic is applied as a fitness function which makes placement based on multi criteria. The efficacy of our algorithm is exhibited by comparing results using MATLAB simulator. Simulation results shows that the overall energy savings achieved by proposed method is 41% and 24% as compared to FFD and ACO approach respectively.

Keywords: -Cloud Computing, VM allocation, VM Migration, MBFD, ABC, Fuzzy Logic, Energy Consumption, and SLA violation

Applications of Business intelligence

Sukhjot Kaur

Student, UIET, Panjab University, Chandigarh

Abstract

Business intelligence is a technology driven process for analyzing data and delivering the necessary information to the assigned roles. The business organizations collect data from relevant resources, prepare the data for analysis which is then used by the organization as per its requirement. The various applications that are being used in Business intelligence are Online Analytical Processing (OLAP), mobile BI, real-time BI, operational BI, cloud BI, open-source BI etc. These BI applications help the business organizations to improve and speed up the decision making process, identify the market trends and consumer demands, recognize various business opportunities, improve working efficiency, make reports for the Business Organizations and to increase revenue.

Keywords: *Business intelligence, OLAP, BI*

EDITOR'S PROFILE



Dr Gurpreet Singh is an Assistant Professor at the Department of Commerce & Economics and Associate NCC officer of the college NCC unit Babbar Akali Memorial Khalsa College Garhshankar (Hoshiarpur). He has more than 15 years of experience in academic research and industry. He has presented papers in reputed National and International conferences and published various research papers including UGC and COPUS indexed journals.



Dr Ajay Dutta is working as an Head & Assistant Professor at Department of Computer Science and Applications, Babbar Akali Memorial Khalsa College Garhshankar (Hoshiarpur). He is Member, Computer Society of India, Internet Society of Computers (ISOC) and Inclusive Policy Lab (IPL) Expert(UNESCO) etc. He has presented research paper at 10th International Conference on Theory and Practice of Electronic Governance (ICEGOV) coordinated by the United Nations University Operating Unit on Policy-Driven Electronic Governance (UNU-EGOV).



Prof. Gurpreet Singh Kalsi is working as an Assistant Professor at the Department of Commerce & Economics Babbar Akali Memorial Khalsa College Garhshankar (Hoshiarpur). He coordinates the NPTEL Local Chapter & Learner Support Centre - Jagat Guru Nanak Dev Punjab State Open University, Patiala. He is having more than 10 years of teaching experience.

ISBN 978-81-955643-2-3

