

# PROCEEDINGS

## THIRD INTERNATIONAL CONFERENCE

on

Innovations in Management, Science & Technology  
(ICIMST-2024)



*Organized By*

DEPARTMENT OF COMPUTER SCIENCE  
University of Science & Technology Meghalaya

*In Association with*

CATHOLIC UNIVERSITY OF MUZAMBIQUE

Proceedings  
Third International Conference  
on  
Innovations in Management,  
Science & Technology  
(ICIMST~2024)



Organized By  
Department of Computer Science University of  
Science & Technology Meghalaya



In Association with  
Catholic University of Mozambique

**Proceedings Third International Conference on Innovations in Management, Science & Technology (ICIMST-2024) :**

A book edited by Dr Kanak Chandra Bora and published by Purbayon Publication, Panbazar, Guwahati-1, Assam

**1<sup>st</sup> Edition :** January, 2026

Price - 800/-

---

**ISBN :** 978-93-5860-855-7

Proceedings Third International Conference on  
Innovations in Management, Science & Technology  
(ICIMST-2024)

**First Edition :** January, 2026

**Cover :** Chitralkha

**Price :** 800/-

© editor

**Published by:**

**Purbayon Publication**

Jaswanta Road

Near Panbazar Aadarsha Prathamik Vidyalaya

Panbazar, Guwahati- 1, Assam, India

Email-purbayonindia21@gmail.com

website: purbayonpublication.com

Contact No. +91-9864422157

---

**Printed at :** Purbayon Publication Pvt. Ltd.  
Panbazar, Guwahati-1, Assam, India.  
Email-purbayonindia21@gmail.com  
Contact No. +91-9864422157

## ACKNOWLEDGEMENT

At the outset I like to praise GOD for bringing it to reality without whom it may not be possible. It is my privilege to acknowledge my debt to all the people who are involved directly or indirectly in publishing this proceeding.

I am very much thankful to our honourable and dynamic Chancellor, University of Science & Technology, Meghalaya, Shree Mahbubul Hoque sir, for his valuable suggestion to organize international conference as well as to publish proceedings. My sincere gratitude to our honourable Vice-Chancellor, USTM , Prof. G. D. Sharma sir, for his continuous guidance.

My sincere thank goes to all the authors who have contributed papers for this proceeding. Next I would like to thank our HOD, Department of Computer Science, USTM, Dr. Atower Ul Islam, and all my faculty colleagues Dr. Chinmoy Bharadwaj , Mrs. Sangeeta Borkakoty, Mrs. Daisy Kalita, Mrs. Daisy Sharmah, Mr. Mazahrul Islam, Mrs. Priyanka Sarma, Mr. Oinam Devid Singh for their support. Finally I am grateful to Prubayon Publication Pvt. Ltd for publishing this proceeding.

**Dr. Kanak Chandra Bora**  
Convener of the ICIMST-2024



## INDEX

- Deep Learning Strategies for Autism Spectrum Disorder Identification: A Comparative MRI Analysis  
Deba Kanta Gogoi  
Jyotismita Talukdar || **9**
- Comparative Study On Multiclass Disease Prediction Using Different Machine Learning Approaches  
Ananya Ahmed  
M. Jahirul Islam || **26**
- Gesture-controlled virtual mouse for hands free computing  
Sahil Khalid Laskar  
Tauhid Imam Chowdhury  
Chinmoy Bharadwaj || **38**
- Fingerprint Identification & Analysis through Deep Learning  
Sumaiya Ashraf  
M. Jahirul Islam || **53**
- Develop Recommendation System for Restaurant Recommendation on Bangladesh Perspective Using Collaborative Filtering Approach  
Dipongkar Ghose  
Emdad Hussain Imu  
Shafkat Kibria || **88**
- Optimizing Minimum Cost Routes : A Comparative Study of Algorithms of Dijkstra's, Edge Deletion Method, and Vertex Assignment Enumeration Techniques,  
Saiful Islam  
Dr Anupam Dutta || **101**
- Controllable susceptibilities in GaAs-AlGaAs Quantum Dot by strong controlfield  
Abhijit Shyam  
Sanchita Roy  
Nitu Borgohain || **115**

- An Analysis of Cybersecurity Challenges and Emerging Trends in Modern Technologies  
 Dr. U. Sridevi  
 Priya M  
 Dharshini S  
 Sharmila T  
 Abhishek J  
 Mohamad Arshath || **130**
- Navigating the Full Stack: Unlocking Future Possibilities and Overcoming Challenges in Web Development  
 Nageswaran M.K  
 Deepika S  
 Dinesh J  
 Vignesh M  
 Hariprakash K  
 Raveendar A || **139**
- Exploring Augmented and Virtual Reality: From Innovation to Everyday Use  
 Nageswaran M.K  
 Akash B  
 Mageshraj M  
 Tamilarasan J  
 Adhinarayanan S  
 Vinoth S || **152**
- An Efficient Ai-Powered Financial Market Prediction Using Machine Learning and Deep Learning Algorithms  
 Nageswaran M.K  
 Dr. R. Sankarasubramanian || **165**
- Iot based real time flood detection and alert system using arduino  
 Mazharul Islam  
 Disha Gogoi  
 Pushpita Sinha  
 Susmita Prasad || **173**
- Bomb Detection & Recovery Robot  
 Abdul Hadi  
 Mazharul Islam || **183**

- Spam Identification in Cloud Computing based on Text Filtering System  
Daisy Sharmah  
Sabnam Siddika || **191**
- Gesture Control Contactless Light for Home Appliances  
Md. Salim Arif  
Rahul Borthakur  
Sangeeta Borkakoty || **210**
- Ethics in the Use of Artificial Intelligences (LLM's) in Academic Work s: A Case Study at UCM  
Maiden Djento  
Mário Pereira Alfinete || **231**
- The impact of Information and Communication Technology (ICT) on modern society  
Ibtisam Hasin  
Tilock ch.Saikia || **246**
- A Survey on Augmented Reality Glass to Assist Deaf People  
Fleaming Nongrum  
Moidul Ali  
Dr. Kanak Chandra Bora || **256**
- Student Attendance System Using Face Recognition  
Bishal Chakraborty  
Himangshu Kalita  
Oinam David Singh || **263**
- Sales Forecasting Using Linear Regression  
Rose Brahma  
Jenarius Ganlary  
Oinam David Singh || **276**
- Power Consumption-Aware Symmetrical Traffic Splitting Based on Heterogeneous Traffic Demands  
Amit Khanjan Sarma  
Sanjib K. Deka  
Nityananda Sarma || **285**

- A Voice-Enabled Smart Dustbin System  
Atowar ul Islam  
Thujota Rhakho  
Neha Rani Daimary || **308**
- Predicting Heart Disease Risk Using Machine Learning Algorithms  
Jenifa Rahman  
Serin Begum  
Daisy Kalita || **320**
- Proactive Edge Computing Task Allocation in IoT Networks using Deep Q-Network (DQN) and Proximal Policy Optimization (PPO)  
Priyanka Sarma  
Atowar Ul. Islam || **342**

# Deep Learning Strategies for Autism Spectrum Disorder Identification: A Comparative MRI Analysis

Deba Kanta Gogoi<sup>1</sup>, Jyotisma Talukdar<sup>1\*</sup>

<sup>1</sup>Tezpur University, Department of CSE, Tezpur, Assam, India

\*Corresponding author(s). E-mail(s): [jyoti4@tezu.ernet.in](mailto:jyoti4@tezu.ernet.in); Contributing authors: [debakantagogoi96@gmail.com](mailto:debakantagogoi96@gmail.com)

## Abstract :

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition affecting social interaction, communication, and behaviour, with symptoms manifesting in early childhood and persisting throughout adulthood. The variability in the severity and presentation of symptoms has led to its classification as a spectrum. Research has identified structural and functional differences between the brains of individuals with ASD and those with typical development (TD). These neurological variations can impact critical brain functions, including information processing, sensory perception, and the ability to interpret and respond to social cues. Recent advances in deep learning (DL) algorithms offer new opportunities for analysing brain MRI data, providing valuable insights into ASD. In this study, we investigate the performance of four DL models—DenseNet121, Inception v3, VGG16, and ResNet50—for identifying ASD-specific features in brain

MRI scans. Using the ABIDE dataset, which includes MRI data from individuals with ASD and TD, we implemented two dataset splits: 70:30:10 and 80:10:10 for training, validation, and testing, respectively. Our analysis compares the effectiveness of these DL models in accurately identifying ASD-related brain patterns. The results are evaluated against existing research to determine improvements in classification performance. By leveraging deep learning, this work contributes to the growing field of neuroimaging-based ASD diagnosis, with the potential to enhance early detection and provide a deeper understanding of the neural underpinnings of ASD. This could facilitate more effective intervention strategies and personalized treatment approaches for individuals on the autism spectrum.

**Keywords :** Autism Spectrum Disorder, Deep learning, DenseNet121, Inception v3, VGG16, ResNet50, ABIDE.

## I. Introduction

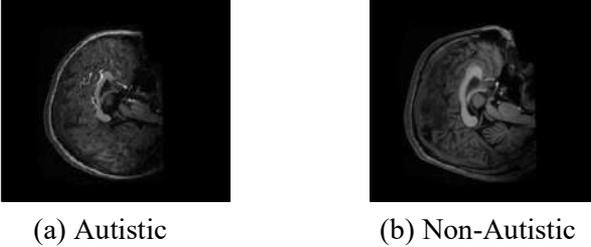
Autism Spectrum Disorder (ASD) is a developmental condition characterized by repetitive behaviours and challenges in social interaction and communication, both verbal and nonverbal. Its prevalence is on the rise, with the World Health Organization (WHO) estimating that 1 in 68 children are affected, amounting to over 2 million individuals in the U.S. and 68 million globally. Early intervention is crucial in helping children with ASD adapt and communicate more effectively. Research in Machine Learning (ML) and Deep Learning (DL) is playing an increasingly important role in ASD diagnosis. These techniques can help identify ASD by analysing various health and physiological factors, contributing to more effective treatment strategies. Early diagnosis of ASD, a lifelong condition that manifests in early childhood, can significantly improve outcomes, although the causes remain largely unknown. Risk factors include genetics, low birth weight, older parental age, and having a sibling with ASD. Individuals with ASD often experience sensory sensitivities, repetitive behaviours, focused interests, communication challenges, difficulties with executive functioning, and a sensitivity to changes in routine. Understanding these traits is essential for developing better diagnostic tools and interventions.

This research aims to evaluate and compare the performance of four deep learning models—DenseNet121, Inception v3, VGG16, and

ResNet50—in detecting Autism Spectrum Disorder (ASD) from brain MRI data. Using the ABIDE dataset, which includes MRI scans from individuals with ASD and typical development (TD), the study seeks to identify brain patterns linked to ASD. f-MRI has been crucial for researchers exploring the brain basis of autism, despite the unknown causes. It offers insights into the neural functioning of individuals with autism and typically developing (TD) individuals. By comparing f-MRI scans, researchers have found several key results:

1. **Altered Brain Connectivity:** f-MRI studies reveal disrupted connectivity in areas responsible for social interaction, language, and emotional regulation in individuals with autism.
2. **Reduced Mirror Neuron Activation:** There is decreased activation of the mirror neuron system, which may hinder social cognition and imitation.
3. **Atypical Sensory Processing:** Individuals with autism often show sensory sensitivities linked to altered activation in sensory-related brain regions.
4. **Enhanced Perceptual Processing:** Despite challenges in social communication, individuals with autism may demonstrate heightened activation in visual and auditory processing areas, suggesting a reliance on perceptual information.
5. **Executive Functioning Deficits:** Changes in brain activation associated with executive functions highlight difficulties in planning, decision-making, and impulse control.

Overall, f-MRI findings in autism are complex and variable, reflecting significant individual differences. While this research enhances our understanding of autism's neurological aspects, it is only one part of the larger picture. Future studies that integrate f-MRI with other methods may further clarify autism's complexities and inform targeted interventions. MRI studies of the Corpus Callosum and whole brain structure have identified significant differences in individuals with ASD. These studies indicate that individuals with ASD often have abnormalities in the size, shape, and connectivity of the Corpus Callosum, contributing to atypical brain connectivity patterns. The findings suggest that disrupted inter-hemispheric communication may affect the social and cognitive difficulties experienced by those with ASD. Fig. 1 illustrates f-MRI slices of autistic and non-autistic brains.



*Fig. 1: MRI brain images for autistic and non-autistic*

## ii. Related Works

A thorough review of the literature is provided in this section about the use of different deep learning approaches for ASD classification. Numerous research utilizing deep learning and ML approaches have been undertaken for ASD diagnosis and categorization.

Mishra, Mayank et al. (2023) [6] uses Deep Convolution Neural Network (DCNN) model which achieve highest accuracy of 81.35%. The ensemble model of DCNN with Adam and Nadam optimizer was proposed in this study and they have used ABIDE I for analysis.

Narala, Mohana Sree Venkata Sai Krishna et al. [7] uses Efficient Net CNN, which achieved the accuracy of 88%.

Gao, Kun et al. (2021) [8] focused on the classification of neuro developmental disorders using structural MRI data. They utilized CNN and Siamese networks, achieving an AUC of 91%. The study employed the NDAR dataset.

Bayram, Muhammed Ali, et al. (2021) [9], 2021, uses RNN and ACC of 74.74% is achieved and SENS is 72.95%. This study presents a promising approach for the early detection of ASD using rs-fMRI data, employing a combination of long short-term memory network (LSTM), CNN, and hybrid models on the dataset.

Subah, Faria Zarin, et al. (2021) [10] in 2021 uses Deep neural network (DNN) models and accuracy of 88% is achieved. This paper proposes a DNN model utilizing functional connectivity features from resting-state f-MRI data and multiple brain atlases to detect ASD, achieving higher accuracy than state-of-the-art methods and demonstrating the superiority of the Bootstrap Analysis of Stable Clusters (BASC) atlas for classification.

Devika, K et al. (2021) [11] uses SVM and find the accuracy of 80.76% using ABIDE II rs-fMRI data to identify neurological diseases.

Husna, R. Nur Syahindah, et al. (2021) [12] uses CNN models, VGG-16 and ResNet-50 and accuracy 87%. This research investigates the use of deep learning models, specifically VGG-16 and ResNet50 using ABIDE dataset is used.

Yin, Wutao, Sakib Mostafa, et al. (2021) [13] uses DNN and accuracy 79.2% accuracy is achieved and AUC of 82.4%. This study presents deep learning methods for diagnosing ASD using functional brain networks constructed from fMRI data on ABIDE 1 dataset.

Leming, Matthew et. al. (2020) [14] uses CNN and accuracy of 57.1150% to present a deep learning approach using a CNN trained on a large fMRI dataset to classify ASD versus TD controls, achieving high accuracy and providing insights into the brain connections involved in the classification on largest multi-source fMRI connectomicdataset.

Huangt, Zhi-An, et al. (2020) [15] uses DBN model and accuracy of 76.4% is achieved. They used a novel graph-based classification model using DBN using ABIDE databaseto identify subtypes within ASD and provide interpretability of the neural correlation patterns.

Ali, Nur et al. (2020) [16] proposed a deep learning model based on a 6-layer CNN architecture for ASD detection with a max accuracy of 80%, although the specific dataset used was not mentioned.

Li, Jing et al. (2020) [17] utilized ML models, specifically SVM and LSTM networks, for ASD detection, achieving an accuracy of 92.6%.

Tao, Yudong et al. (2019) [18] employed SP-ASDNet, a deep learning model, to identify ASD in patients. They reported an accuracy of 74.22% using the Saliency 4ASD data.

Akhavan Aghdam, Maryam et (2018) al. [19] uses Deep belief network (DBN) using ABIDE I and II dataset with accuracy of 65.56%. This paper explores the application of a DBN utilizing data from the datasets to classify ASDs from typical controls (TCs) by combining rs-fMRI, gray matter, and white matter data, achieving improved accuracy compared to previous methods.

Heinsfeld, Anibal S'olon, et al. (2018) [20] uses DNN, RF, SVM models and acc of 70% is achieved. This work reveals altered anterior-posterior brain connection in ASD and uses DL algorithms to identify patients based on brain activity patterns, distinguishing ASD from normally developing controls using ABIDE I data.

Thomas, Monisha et al. (2018) [21] presented a novel method for detecting ASD using ANN, without specifying performance metrics or datasets.

These studies highlight the advancements in utilizing DL techniques for ASD classification using MRI images. The application of various deep learning architectures, including 3D CNNs, auto encoders, attention-guided models, and multi-view frameworks, showcases the diversity of approaches in this field. The results highlight how DL learning may help with early identification and intervention efforts by offering precise and automated diagnosis of ASD to clinicians and researchers.

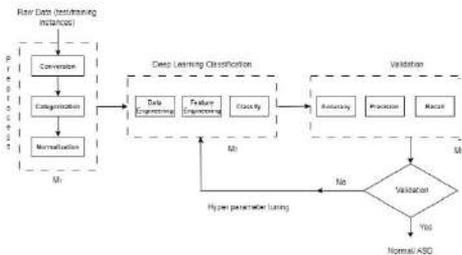
Overall, the literature survey demonstrates the growing interest in utilizing deep learning methods for ASD classification using MRI images.

### Iii. Methodology

It has been observed that a significant number of methods have been developed using deep learning approaches for ASD classification, with other commonly used classifiers including Support Vector Machines (SVM) and Naive Bayes (NB). The Autism Brain Imaging Data Exchange (ABIDE) and its versions 1 and 2 are frequently utilized for validation purposes.

For a given test f-MRI image,  $I_{test}$ , the objective is to classify it with optimal accuracy using a set of training instances,  $I_{trains}$ , which includes examples from two classes: 'TD' and 'ASD.' Each f-MRI image undergoes preprocessing, and feature engineering is applied to extract ' $n$ ' important features. From these, ' $k$ ' relevant and non-redundant features are selected to maximize classification accuracy. In our proposed methodology, we have proposed three modules: Module 1 (M1), Module 2 (M2), and Module 3 (M3) as shown in

*Figure 2 for Classification and detection of ASD:*



**Fig. 2:** Block diagram of the proposed methodology

## A. Module 1:

In Module 1, we use a benchmark dataset ABIDE to pre-process using conversion, categorisation and normalization technique to train our selective models. We have normalized the MRI image data in the range of 0-1 image pixels. To scale pixels in the range 0-1, divide the maximum value of the pixel by the minimum value of the pixel ( $1/255 = 0.0039$ ) and adjust the rescale parameter accordingly. The steps followed here are :

**Conversion:** The conversion stage involves changing the raw data into a suitable format for further analysis. This could include converting data from one data type to another, such as converting categorical data into numerical representations.

**Categorization:** Categorization is the process of organizing data according to certain criteria or features into groups or meaningful groupings. Categorization can be done manually by assigning labels or automatically using techniques such as clustering or classification algorithms.

**Normalization:** Normalization aims to standardize the data by scaling it to a specific range or distribution. Common normalization techniques include min-max scaling, where the data is transformed to a predefined range (e.g., between 0 and 1), and z-score normalization (standardization), where the data is transformed to have zero mean and unit variance.

## B. Module 2:

When working with datasets, it is crucial to split the data into different parts. Here, the dataset is split into training, validation, and test sets, 70:20:10 and 80:10:10, respectively.

**Data Engineering:** This step focuses on preparing and organizing the data. It includes tasks such as data cleaning, pre-processing, and augmentation to ensure data quality and compatibility with deep learning models.

**Feature Engineering :** Feature engineering aims to extract meaningful representations from the data. Techniques like dimensionality reduction, feature selection, and creation of new features are used to provide relevant and informative inputs to the deep learning model.

**Classification :** In this step, a deep learning model is trained and used for prediction. Models such as CNNs, or transformers are

commonly employed. The trained model can then classify new data instances into different classes or provide probabilities for each class.

### **C. Module 3 :**

In the module 3, classification includes train and test instances typically three evaluation steps: accuracy, precision, recall, and F1-score.

The various Classification models used are as follows :

#### **A. DenseNet121:**

DenseNet121's design starts with a convolutional layer followed by a max pooling layer to extract initial features. An essential part of DenseNet121 is the dense block. Layers within the same thick block are all connected to one another through feed forward connections [22]. Thus, the results of one layer are used as input for the next levels in the block. This high degree of interconnection facilitates communication and encourages the sharing of features. Each dense block is built from several layers of convolutional processing, batch normalisation, and ReLU activations. Each thick block can have any number of layers within it, depending on the design. Convolution and pooling techniques are used to minimise the spatial dimensions of the feature maps, and transition layers are put between dense blocks to further this effect. The feature maps are spatially reduced by the application of global average pooling following many dense blocks. A completely connected layer is the next stage, and then a SoftMax activation is used for classification.

#### **B. Inception v3 :**

The Inception v3 architecture consists of multiple Inception modules stacked together. Each Inception module is composed of multiple parallel branches. Different operations applied to the input feature maps in each branch [23]. These operations include 1x1 convolutions, 3x3 convolutions, and 5x5 convolutions, with the aim of capturing features at different scales.

Specifically, let's denote the input feature maps to the Inception module as  $X$ . Within each branch, the feature maps are processed differently. For example, a branch might apply a 1x1 convolution to  $X$ , followed by a 3x3 convolution, while another branch might apply a 1x1 convolution followed by a 5x5 convolution. The results from all branches are then concatenated along the channel dimension, resulting in a fused feature map. This

fused feature map is passed to the next Inception module or to other layers in the network.

### **C. ResNet50:**

As part of the ResNet50 architecture, low-level features are extracted from an input picture using a max pooling layer that comes after a convolutional layer. Residual blocks are the framework upon which ResNet50 is built. Each residue block is composed of distinct convolutional layers, batch normalisation, and ReLU activations. Additionally, the residual block has a connection that skips over the convolutional layers and simply adds the input to the final output.[24] This helps to prevent the identification information deterioration that can occur in deep networks and preserves privacy.

To reduce and limit the no. of parameters and computational cost, ResNet50 uses a variety of residual block configurations, including bottleneck blocks that use 1x1, 3x3, and 1x1 convolutions. The spatial dimensions of the feature maps are shrunk by using global average pooling after many stacked residual blocks have been generated. The next step is a fully linked layer, followed by a softmax activation layer, which is used for classification.

### **D. VGG16:**

VGG16 is a CNN architecture that is widely considered to be one of the best deep learning models available today [25]. The 16 in VGG16 refers to that it has 16 layers that have weight. The 16 in VGG16 includes 16 layers that have weights. In VGG16 there are 13 convolutional layers, 5 Max Pooling layers, and 3 dense layers which sum up to 21 layers. But, it has only 16 weight layers i.e., layer with learnable parameters.

#### **1. Dataset Description :**

Table 1 shows the ABIDE dataset description details which includes 539 ASD and 573 TD control patients.

**Table 1: Dataset**

Initiative	ABIDE
Participants	17 international sites
Datasets	1112 datasets (539 with ASD, 573 TD controls)
Age Range	7 to 64 years

Contributions	Provided researchers with a valuable resource to study brain activity and structure in individuals with ASD and TD
Impact	Enabled several investigations examining the links between brain behaviour in people with ASD and TD and functional connectivity, brain network structure, and behaviour.

**2. Performance assessment matrix :**

Classification model evaluation uses measures like Accuracy, Precision, Recall, and F1-score in ML and DL tasks. Precision assesses how well the model avoids producing false positives, while Accuracy counts how often it gets predictions right. Measures of precision & recall are combined into a score called F1-score, which evaluates a model’s overall performance. Using these measures to evaluate models is essential for making comparisons and well-informed decisions about which categorization algorithms or architectures to use. Here, T denote True, F denote False, P denote Positive and N denote Negative. For performance evaluation above performance measures are used:here the accuracy, precision and recall equations are shown below.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \text{-----} (1)$$

$$\text{Or, Accuracy} = \frac{TP+TN}{TP+TN+FP+F} \text{-----}(2)$$

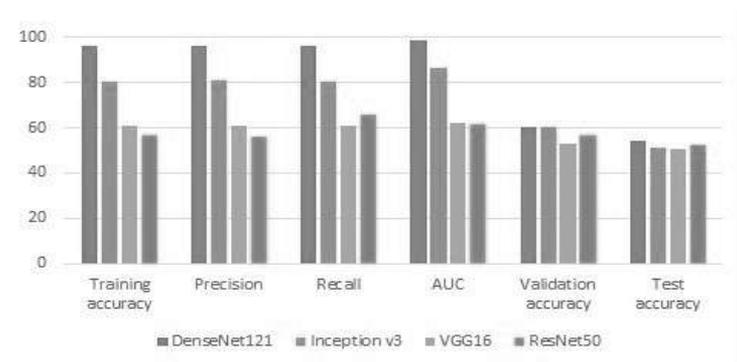
$$Precision = \frac{TP}{TP+FP} \text{-----}(3)$$

$$Recall = \frac{TP}{TP+F} \text{-----}(4)$$

**3. Results & Discussion:**

In this study, we used MRI datasets consisting of both autistic and non-autistic patients for classification and prediction purpose. We evaluated the performance of several deep learning models, including DenseNet121, InceptionV3, VGG16, and ResNet50, using performance metrics such as Accuracy, Precision, and Recall. While comparing performance of two aspect ratio 70:20:10 and 80:10:10, the 70:20:10 is showing good performances.

DenseNet121 achieved the highest training accuracy of 96.01%. And InceptionV3 model is performing second highest training accuracy of 80.5%. And VGG16 acc. of 61.12%, ResNet50 model acc. of 56.67%. Fig. 3 shows the model testing accuracy, precision, recall, AUC, validation accuracy, and test accuracy metrics for all four deep learning classification models.



**Fig. 3:** Performances matrices

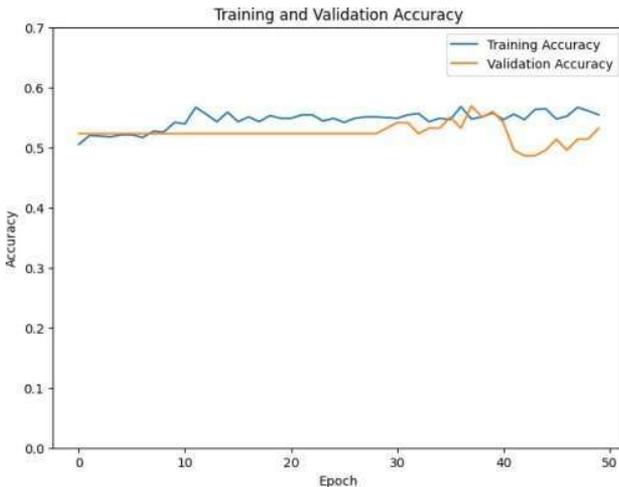
**Table 3:** Comparison results with different models

Model Model	Training accuracy	Precision	Recall	AUC	Validation Accuracy	Test accuracy
Dense Net121	96.01	96.12	96.01	98.69	60.55	54.05
InceptionV3	80.5	80.71	80.16	86.56	60.55	51.35
VGG16	61.12	61.12	61.12	61.94	53.21	50.45
ResNet50	56.67	55.86	65.79	61.32	56.88	52.25

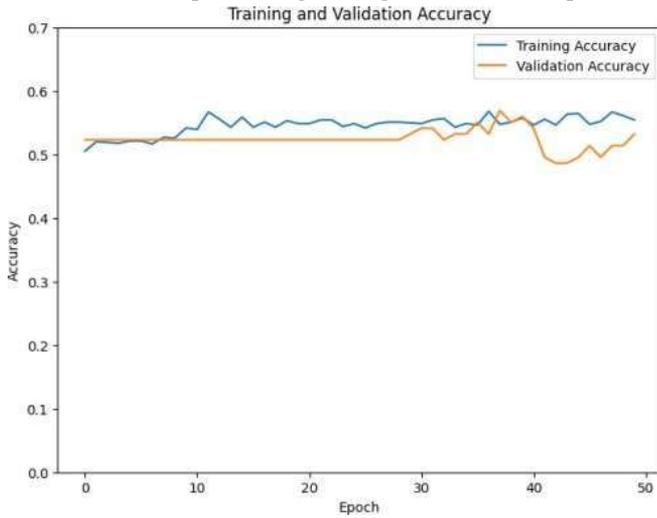
Overall comparison of the four DL models is shown in the table 3. The table showcase the training accuracy, precision, recall, AUC, validation accuracy, and test accuracy of the all four model. Where DenseNet121 is performing well in case of all the six parameters, training accuracy 96.01%, precision 96.12%, recall 96.01, AUC 98.69%, validation accuracy 60.55, test accuracy 54.05%. In InceptionV3 training accuracy 80.5%, precision 80.71%,

recall 80.56, AUC 86.56%, validation accuracy 60.55, test accuracy 51.35%. In VGG16 training accuracy 61.12%, precision 61.12%, recall 61.12, AUC 61.94%, validation accuracy 53.21, test accuracy 50.45%. And in ResNet50 training accuracy 56.67%, precision 55.86%, recall 65.79, AUC 61.32%, validation accuracy 56.88, test accuracy 52.25%.

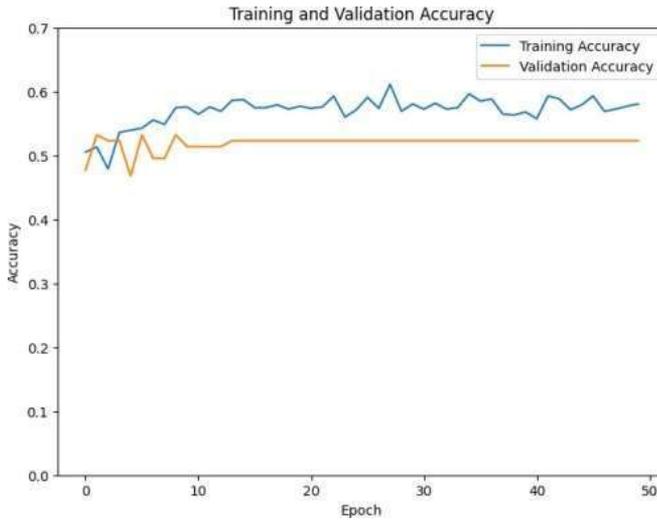
In order to get a visual understanding of how the DL models were doing during training, accuracy curves were created, shown in Fig. 4. The accuracy curves of DenseNet121 and InceptionV3 is encouraging, indicating that these networks are able to learn and generalise within the autistic brain MRI dataset. These curves provide visible evidence that the models are effective and provide important feedback on how well they are being trained and whether or not they are stable.



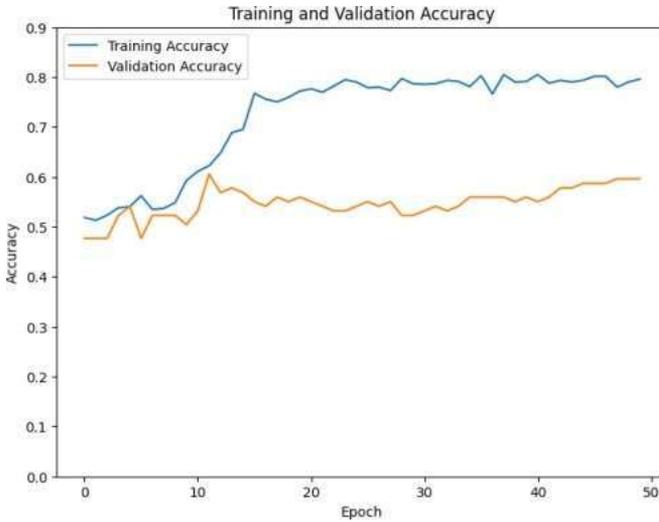
(a) Accuracy curve for DenseNet121 (96.01%)



(b) Accuracy curve for InceptionV3 (80.5%)



(c) Accuracy curve for VGG16 (61.12%)



(d) Accuracy curve for ResNet50(56.67%)

**Fig.3:** Accuracy curve of different classification models

#### Iv. Conclusion

This paper compares deep learning models with existing models and uses transfer learning for improving the prediction accuracy. In our analysis our proposed model achieved the highest accuracy of 96.01%. It establishes the potential of DL models to classification of autism. The method has been validated using benchmark dataset shared through ABIDE. The experimental results reveal that DL based methods can address the issue of accurate classification of ASD after necessary changes in the model parameters. The experimental comparisons with other competing methods also establish the superiority of our method. This evolving complex disease can be more successfully handled using an integrative approach. In future the approach should be consider not only medical images but also genomics scale and clinical data while deciding the class label of test case after applying an appropriate consensus function. Further, distributed computational approach could be explored to yield a real-time response without compromising accuracy.

**Acknowledgements :** None.

**Data availability :** The dataset used for our research is Autism Brain Imaging Data Exchange (ABIDE).

**Declarations :**

- **Ethical approval:** Ethics approval was not required for this study.
- **Conflict of Interest:** The author declares that they have no potential conflict of interest.

**References :**

- [1] Craddock, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B.S., Lewis, J.D., Li, Q., Milham, M., *et al.*: The neuro bureau preprocessing initiative: open sharing of pre-processed neuro-imaging data and derivatives. *Frontiers in Neuroinformatics*7(27), 5 (2013)
- [2] Alarifi, H.S., Young, G.: Using multiple machine learning algorithms to predict autism in children. In: *Proceedings on the International Conference on Artificial Intelligence (ICAI)*, pp. 464–467 (2018). The Steering Committee of The World Congress in Computer Science, Computer ...
- [3] Talukdar, J., Gogoi, D.K., Singh, T.P.: A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents. *Healthcare Analytics* **3**, 100178 (2023)
- [4] Lord, C., Elsabbagh, M., Baird, G., Veenstra-Vanderweele, J.: Autism spectrum disorder. *The lancet* **392**(10146), 508–520 (2018)
- [5] Unterberger, I., Bauer, R., Walser, G., Bauer, G.: Corpus callosum and epilepsies. *Seizure* **37**, 55–60 (2016)
- [6] Mishra, M., Pati, U.C.: A classification framework for autism spectrum disorder detection using smri: Optimizer based ensemble of deep convolution neural network with on-the-fly data augmentation. *Biomedical Signal Processing and Control* **84**, 104686 (2023)
- [7] Narala, M.S.V.S.K., Vemuri, S., Kattula, C.: Prediction of autism spectrum disorder using efficient net. In: *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 1139–1143 (2023). IEEE

- [8] Gao, K., Sun, Y., Niu, S., Wang, L.: Unified framework for early stage status prediction of autism based on infant structural magnetic resonance imaging. *Autism Research* **14**(12), 2512–2523 (2021)
- [9] Bayram, M.A., Ilyas, O., Temurta, F.: Deep learning methods for autism spectrum disorder diagnosis based on fmri images. *Sakarya University Journal of Computer and Information Sciences* **4**(1), 142–155 (2021)
- [10] Subah, F.Z., Deb, K., Dhar, P.K., Koshiba, T.: A deep learning approach to predict autism spectrum disorder using multisite resting-state fmri. *Applied Sciences* **11**(8), 3636 (2021)
- [11] Devika, K., Oruganti, V.R.M.: A machine learning approach for diagnosing neurological disorders using longitudinal resting-state fmri. In: 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 494–499 (2021). IEEE
- [12] Husna, R.N.S., Syafeeza, A., Hamid, N.A., Wong, Y., Raihan, R.A.: Functional magnetic resonance imaging for autism spectrum disorder detection using deep learning. *Jurnal Teknologi* **83**(3), 45–52 (2021)
- [13] Yin, W., Mostafa, S., Wu, F.-X.: Diagnosis of autism spectrum disorder based on functional brain networks with deep learning. *Journal of Computational Biology* **28**(2), 146–165 (2021)
- [14] Leming, M., G’orriz, J.M., Suckling, J.: Ensemble deep learning on large, mixed-site fmri datasets in autism and other tasks. *International journal of neural systems* **30**(07), 2050012 (2020)
- [15] Huang, Z.-A., Zhu, Z., Yau, C.H., Tan, K.C.: Identifying autism spectrum disorder from resting-state fmri using deep belief network. *IEEE Transactions on neural networks and learning systems* **32**(7), 2847–2861 (2020)
- [16] Ali, N.A., Syafeeza, A., Jaafar, A., Alif, M., Ali, N.: Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm. *IAES International Journal of Artificial Intelligence* **9**(1), 91–99 (2020)
- [17] Li, J., Zhong, Y., Han, J., Ouyang, G., Li, X., Liu, H.: Classifying asd children with lstm based on raw videos. *Neurocomputing* **390**, 226–238 (2020)
- [18] Tao, Y., Shyu, M.-L.: Sp-asdnet: Cnn-lstm based asd classification model using observer scanpaths. In: 2019 IEEE

- International Conference on Multimedia & Expo Workshops (ICMEW), pp. 641–646 (2019). IEEE
- [19] Akhavan Aghdam, M., Sharifi, A., Pedram, M.M.: Combination of rs-fmri and smri data to discriminate autism spectrum disorders in young children using deep belief network. *Journal of digital imaging* **31**, 895–903 (2018)
- [20] Heinsfeld, A.S., Franco, A.R., Craddock, R.C., Buchweitz, A., Meneguzzi, F.: Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clinical* **17**, 16–23 (2018)
- [21] Thomas, M., Chandran, A.: Artificial neural network for diagnosing autism spectrum disorder. In: 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 930–933 (2018). IEEE
- [22] Chhabra, M., Kumar, R.: A smart healthcare system based on classifier densenet 121 model to detect multiple diseases. In: *Mobile Radio Communications and 5G Networks: Proceedings of Second MRCN 2021*, pp. 297–312. Springer, ??? (2022)
- [23] Minarno, A.E., Aripa, L., Azhar, Y., Munarko, Y.: Classification of malaria cell image using inception-v3 architecture. *JOIV: International Journal on Informatics Visualization* **7**(2), 273–278 (2023)
- [24] Mascarenhas, S., Agarwal, M.: A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification. In: 2021 International Conference on Disruptive Technologies for Multi-disciplinary Research and Applications (CENTCON), vol. 1, pp. 96–99 (2021). IEEE
- [25] Yang, H., Ni, J., Gao, J., Han, Z., Luan, T.: A novel method for peanut variety identification and classification by improved vgg16. *Scientific Reports* **11**(1), 15756 (2021)

# A Comparative Study On Multiclass Disease Prediction Using Different Machine Learning Approaches

Ananya Ahmed<sup>1</sup>, M. Jahirul Islam<sup>2</sup>

*Institute of Information and Communication Technology, Shahjalal University of Science and Technology<sup>1,2</sup>*

## **Abstract :**

Machine learning for disease prediction from symptoms has become an important tool in healthcare. This research article discusses several machine learning algorithms, such as K-Nearest Neighbors (KNN), Naive Bayes, Decision Tree, Random Forest, and LightGBM for multiclass disease prediction with performance comparisons. The main focus of this work is the optimization in the performance of models using feature selection techniques: Recursive Feature Elimination (RFE), Sequential Forward Feature Selection (SFFS), and Sequential Backward Feature Selection (SBFS). The results show that Naïve Bayes, when trained on all features, achieves the best result with 98.36% accuracy and 80.26% while trained on dataset after feature selection using RFE. This study aims to determine the best performing algorithm and feature selection method for disease prediction, offering a practical solution for healthcare sector, and focuses on the study of

how feature selection improves computational efficiency and practical applicability in medical diagnostics.

**Index Terms** : feature selection, machine learning, multiclass disease prediction, K-Nearest Neighbors, Naïve Bayes, Recursive Feature Elimination

## I. Introduction

In recent years, health care systems have depended much on technological development in improving diagnosis and outcomes for patients. Many remote areas of the world still have very scarce resources available to the healthcare system, and as a result, access to medical support remains difficult. Multiclass disease prediction by using machine learning algorithms therefore solves this problem automatically, by predicting diseases based on a set of symptoms given, which can greatly enhance medical diagnosis and reduce human error.

The study investigates the design of a disease prediction system based on machine learning, and its performance is compared for different algorithms. The model is then simplified by applying feature selection techniques in order not to lose much accuracy when reducing the number of input features. It compares K-Nearest Neighbors, Naïve Bayes, Decision Tree, Random Forest, and LightGBM in terms of accuracy, precision, and recall both before and after feature selection.

## II. Literature Review

The application of machine learning in healthcare has been the focus of several recent studies. Gupta et al. [1] developed a disease prediction system using Decision Tree and Naïve Bayes, achieving a maximum accuracy of 98% with Naïve Bayes on a dataset collected from New York-Presbyterian Hospital. Similarly, Dahiwade et al. [2] applied KNN and CNN to disease prediction tasks, demonstrating that CNN outperformed KNN in both accuracy and runtime. Huang et al. [3] explored data mining techniques for disease prediction using historical patient records. Their study compared the Naïve Bayes and J48 algorithms, finding that an ensemble of J48 classifiers slightly outperformed Naïve Bayes with an accuracy of 83.46%. Ambekar et al. [4] applied convolutional neural networks (CNN) for heart disease risk prediction, incorporating KNN and Naïve Bayes to select the best-performing model before using CNN for final risk assessment.

These studies demonstrate the effectiveness of various machine learning models in healthcare diagnostics. However, optimizing these models through feature selection techniques remains an area of interest, as it can significantly enhance computational efficiency and model accuracy.

### III. Dataset and Data Preprocessing

#### A. Dataset

The dataset used in this study was obtained from a study by the University of Columbia and New York-Presbyterian Hospital conducted in the year 2004. It contains 4920 instances with 131 independent features and 41 target classes, each representing a record of a patient. Features in this dataset correspond to different symptoms such as abdominal pain, fever, joint pains, and nausea, while classes correspond to different diseases like hypertension, diabetes, and malaria. The dataset was split into 80% for training the models and 20% for testing.

#### B. Data Preprocessing

Preprocessing the data involved the following steps:

- 1) Dropping the duplicate rows to ensure data uniqueness.
- 2) Deleting cells with NaN values. Two columns, named 'Fluid Overload' and 'Unnamed 133,' were found to contain 9840 NaN values, representing missing information, and were therefore deleted.
- 3) Encoding the target column, which contained disease names, from strings into numerical values using Label Encoder.

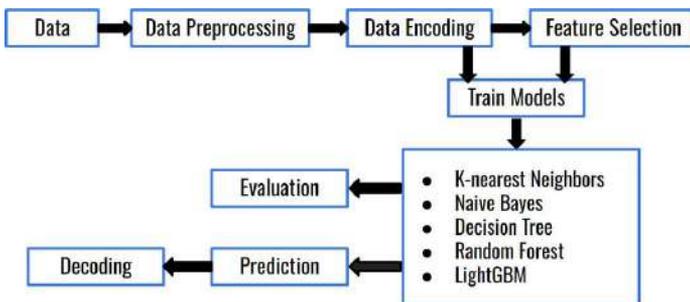


Figure 1. Flow chart of the research.

### IV. Classification Methods

In this work, five different classifiers were used, namely K-Nearest Neighbors (KNN), Naïve Bayes, Decision Tree, Random

Forest and LightGBM. An overview of each classifier used is provided in this section.

### A. K-Nearest Neighbors (KNN)

K-Nearest Neighbors or KNN is a non-parametric, instance-based learning algorithm for classification and regression. The algorithm works in such a way that it picks up the most 'k' similar data points or neighbors and labels the data according to the majority class in those neighbors. The distance metric, commonly the Euclidean distance, is calculated as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

where  $x$  and  $y$  are data points in  $n$ -dimensional space.

One of KNN's key advantages is simplicity and effectiveness with small datasets, but it can be computationally expensive when dealing with large datasets as it needs to compute the distance for every instance in the training set during classification [5].

### B. Naïve Bayes

Naïve Bayes is a probabilistic classifier that applies Bayes' theorem with the independence assumption of features. The formula for predicting the class  $C$  of an instance  $X$  is :

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

The classifier selects the class with the highest posterior probability. Naïve Bayes has turned out to be a tremendous performer in real-world data analysis tasks, especially in high-dimensional input space scenarios and is computationally efficient and particularly well-suited for large datasets [6].

### C. Decision Tree

Decision Tree (DT) is a tree-structured model used for classification and regression, consisting of nodes representing feature tests and branches representing outcomes. The tree is constructed by selecting features based on Information Gain or Gini Index. The formulas are as follows :

*Information Gain :*

$$IG(T, A) = H(T) - \sum_{v \in \text{Values}(A)} \frac{|T_v|}{|T|} H(T_v)$$

*Gini Index :*

$$Gi(T) = 1 - \sum_{i=1}^k (P_i)^2$$

Where  $(T)$  is the entropy of the set  $T$ ,  $T_v$  is the subset of  $T$  for which feature  $A$  takes on value  $v$ , and  $p_i$  is the proportion of class  $i$  in set  $T$ . [7]

#### **D. Random Forest**

Random Forest is an ensemble learning technique that combines many decision trees to improve the performance of a classification task. Each of the trees is trained on a random subset of the original data sample and a random subset of features, which reduces the variance compared to single decision trees, hence improving the robustness. The final prediction is generated by combining the predictions of all the trees, usually by using majority voting for classification purposes. Random Forest is known for its high accuracy and ability to handle large datasets and a high number of input features [8].

#### **E. LightGBM**

Light Gradient Boosting Machine (LightGBM) is a highly efficient, distributed gradient-boosting framework designed to reduce both training time and memory usage while maintaining high accuracy. LightGBM uses decision trees as weak learners and applies a leaf-wise growth strategy, allowing trees to grow asymmetrically, which helps in reducing overfitting and improving accuracy on complex datasets [9]. The key formula for the loss function in gradient boosting is:

$$L = \sum_{i=1}^n l(y_i, f(x_i))$$

Where  $L$  is the total loss,  $y_i$  is the true label, and  $(x_i)$  is the predicted value [9].

#### **V. Feature Selection Methods**

Feature selection is the process of identifying a subset of the most appropriate and relevant features from the full set of features in a dataset which helps in eliminating redundant, irrelevant, and noisy features,

simplifying the model, reducing training time, and addressing overfitting, thereby improving generalization. In this study, feature selection is employed to enhance the practicality of the system for patients by minimizing the number of features they need to interact with, focusing only on the most impactful features out of the 131 available.

#### **A. Sequential Forward Feature Selection (Forward Selection)**

This method begins by creating  $n$  subsets of features, where  $n$  represents the total number of features, and each subset consists of a single feature. The model is then trained on each subset, and the feature that yields the highest accuracy is identified as important. In the subsequent stage, the subsets are expanded to include two features: the previously selected important feature and an additional one. This iterative process continues until a predefined number of relevant features have been selected. The process can be mathematically represented as :

$$F_{selected} = \{f_1, f_2, \dots, f_k\}$$

Where  $F_{selected}$  are the selected features and  $f_k$  is the  $k^{th}$  feature that maximizes the model's performance [10].

#### **B. Sequential Backward Feature Selection (Backward Elimination)**

It operates similarly to forward selection but in reverse. It begins with subsets where each set contains  $n - 1$  features. The feature missing from the subset that yields the best performance is eliminated first. This process is repeated iteratively until the number of features is reduced to a predefined threshold.

#### **C. Recursive Feature Elimination (RFE)**

This method begins with the full set of features and progressively eliminates features until the desired number of features remains. RFE ranks the importance of features using the built-in `feature_importance` function from the random forest algorithm. This method can remove multiple features in each iteration, making it highly efficient and fast in execution [11].

## **VI. RESULTS**

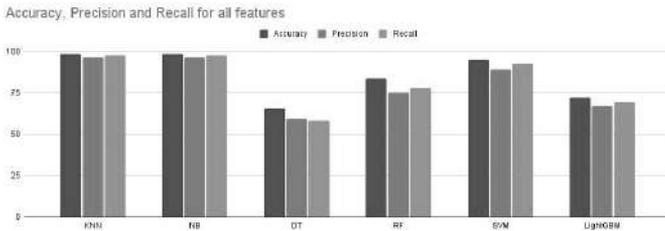
### **A. Classifier Accuracies**

The simulation of the K-Nearest Neighbors (KNN), Naïve Bayes (NB), Decision Tree (DT), Random Forest, and LightGBM algorithms were conducted on the multiclass disease prediction dataset. Table 1 shows the classification performance of the algorithms based on accuracy, precision and recall.

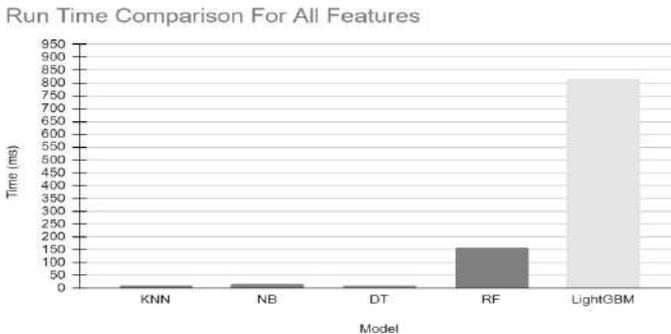
**TABLE 1. Accuracy, Precision and Recall of all models**

ML Classifier	Accuracy (%)	Precision (%)	Recall (%)
KNN	98.3606	96.3415	97.561
Naïve Bayes	98.3606	96.3415	97.561
Decision Tree	67.2131	60	59.7561
Random Forest	90.1639	83.5366	85.3658
LightGBM	72.13	67.1544	69.512

Figure 3 illustrates the comparison of accuracy, precision, and recall for all classifiers. We can see that both K-Nearest Neighbors (KNN) and Naïve Bayes (NB) exhibit identical performance across all evaluation metrics.



*Figure 3. Accuracy, Precision and Recall of models trained on all features.*



*Figure 4. Runtime of models trained on all features.*

Figure 4 presents the runtime comparison of all models. Although KNN and Naïve Bayes (NB) achieve identical evaluation scores, KNN has a faster execution time than NB but is slower than Decision Tree. LightGBM outperforms in terms of accuracy but has the longest runtime among all the models.

## B. Feature Selection Results

Three feature selection methods were applied to the dataset, and the classification models were subsequently trained on the dataset with the reduced feature sets. The accuracies of the models using these feature selection methods are shown in Table 2.

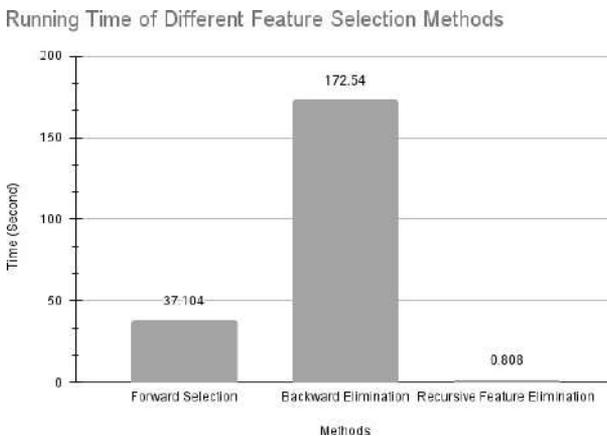
**TABLE 2. Accuracy of models using feature selection methods**

Classification Models	SBFS (%)	SFFS (%)	RFE (%)
KNN	85.2459	78.688	77.6316
Naïve Bayes	80.3279	81.967	80.2632
Decision Tree	75.4098	75.41	68.4211
Random Forest	83.6066	78.688	75.00
LightGBM	62.2951	50.819	76.3158
<b>Overall</b>	<b>77.377</b>	<b>73.114</b>	<b>75.53</b>

SBFS achieves the highest overall accuracy at 77.377%, while SFFS produces the lowest accuracy at 73.114%. The overall accuracy for RFE falls between these two methods.

Figure 5 compares the runtime required for each feature selection method when selecting 20 features. RFE is the fastest, completing the task in 0.808 seconds. In contrast, SBFS takes the longest time, requiring 172.54 seconds, as it involves backward selection and retrains the model each time a feature is eliminated. SFFS also takes considerable time at 37.104 seconds, but it is faster than SBFS due to its forward selection process.

*Figure 5. Runtime of feature selection methods for 20 features.*



Taking both accuracy and runtime into account, RFE was

determined to be the most suitable feature selection method for our system.

### C. RFE With Varying Number of Features

Accuracies of different models trained with different numbers of features using RFE are shown in Table 3.

**TABLE 3. Accuracy of models using RFE with varying features**

Classification Models	30 Features	25 Features	20 Features	15 Features	10 Features
KNN	88.15%	84.21%	77.63%	65.79%	55.26%
Naïve Bayes	84.21%	82.89%	80.26%	67.10%	55.26%
Decision Tree	71.05%	65.78%	67.10%	60.52%	50%
Random Forest	77.73%	75%	75%	59.21%	51.31%
LightGBM	73.68%	73.68%	73.68%	63.15%	50%
<b>Overall</b>	<b>79.40%</b>	<b>76.97%</b>	<b>75.21%</b>	<b>63.15%</b>	<b>52.63%</b>

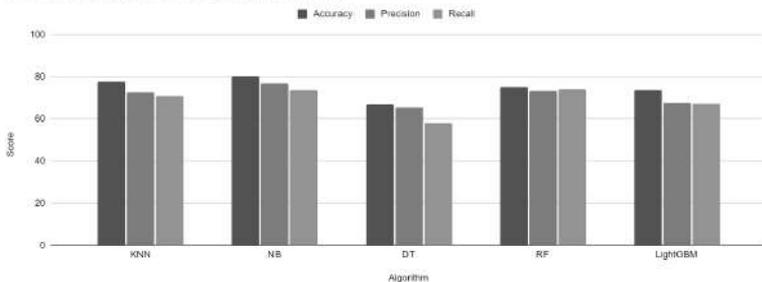
It was observed that the overall accuracy surpasses 75% when utilizing 20 features. As a result, we have chosen to utilize 20 features in our system. Although we could have used 25 or 30 features, the primary aim of applying feature selection is to enhance the system's practicality for patients. Thus, it is advantageous to maintain both high accuracy and a minimal number of relevant features.

**Table 4. Selected 20 features**

Itching	Vomiting	High Fever	Nausea	Irritability
Blurred Vision	Skin Rash	Fatigue	Sweating	Yellowish Skin
Abdominal Pain	Chest Pain	Chills	Weight Loss	Malaise
Headache	Diarrhea	Joint Pain	Cough	Muscle Pain

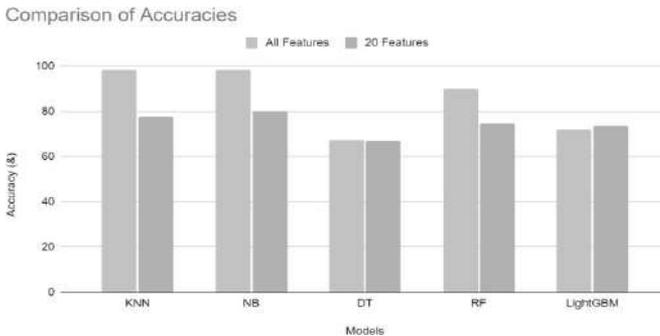
*Figure 6. Accuracy, Precision and Recall of models trained on 20 features.*

Accuracy, Precision and Recall with 20 Features



The accuracy, precision, and recall of classifier models trained with 20 features are illustrated in Figure 6. It is observed that Naïve Bayes achieved the best performance out of all the classifiers when trained on the dataset with 20 features. KNN closely followed, while the Decision Tree model achieved the lowest accuracy.

Figure 7 illustrates that when models were trained on all features, both KNN and Naive Bayes achieved the same accuracy scores, showing the best performance among the classifiers. However, after reducing the features to 20, the accuracy of all models declined, with NB continuing to exhibit the highest performance. KNN displayed a slightly lower accuracy compared to NB. The Decision Tree model showed consistent performance, remaining nearly the same before and after feature selection. The accuracy of the Random Forest model decreased by approximately 15%, while LightGBM demonstrated improved performance following the reduction in the number of features.



*Figure 7. Comparison of models' accuracies trained on all and 20 features.*

## Vii. Conclusion

These results illustrate how such techniques might provide fertile ground not only for enhancement in diagnosis but also for support in clinical decisions. This work contributes to the fast-expanding field of AI-driven healthcare, paving the way for more advanced systems in the future.

A limitation of this study is the small size of the dataset, as the number of samples available for training and testing is relatively low. Future plans include the inclusion of more data to enhance the system performance and the application of CNN-UDRP (CNN Unimodal Disease Risk Prediction) to assess the risk associated with the predicted diseases.

## References

1. J. P. Gupta, A. Singh, and R. K. Kumar (2021), “A Computer-based Disease Prediction and Medicine Recommendation System Using Machine Learning Approach”, *International Journal of Advanced Research in Engineering and Technology (IJARET)*, vol. 12, no. 3, pp. 673–683.
2. D. Dahiwade, G. Patle, and E. Meshram (2019), “Designing Disease Prediction Model Using Machine Learning Approach”, *Proc. 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE (2019), pp. 1211-1215.
3. F. Huang, S. Wang, and C.-C. Chan (2012), “Predicting Disease by Using Data Mining Based on Healthcare Information Systems”, *Proc. IEEE International Conference on Granular Computing (GrC)*.
4. S. Ambekar and R. Phalnikar (2018), “Heart Disease Risk Prediction Using Convolutional Neural Network”, *Proc. 4th Int. Conference on Computing Communication Control and Automation*, 2018, pp. 1-5.
5. Cover, T.M., Hart, P.E. (1967), “Nearest neighbor pattern classification”, *IEEE Trans. Inf. Theor.* 13(1), 21–27 (1967)
6. Rish, I. (2001). “An empirical study of the naive Bayes classifier”, *IJCAI 2001 workshop on empirical methods in artificial intelligence*, Vol. 3, pp. 41–46.
7. L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone (2017), “Classification and Regression Trees”, Evanston, IL, USA: Routledge.
8. Breiman, L. (2001), “Random forests”, *Machine Learning*, vol. 45, no. 1, 2001, pp. 5-32.
9. Ke, G., Meng, Q., Finley, T. (2017), “LightGBM: A highly efficient gradient boosting decision tree”, *Proc. 31st International Conference on Neural Information Processing Systems (NIPS)*, 2017, pp. 3146–3154.
10. Kohavi, R., and John, G. H. (1997). “Wrappers for Feature Subset Selection”, *Artificial Intelligence*, 97, 273– 324.
11. I. Guyon and A. Elisseeff (2003), “An introduction to variable and feature selection”, *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182.
12. M. F. Ferjani (2020), “Disease Prediction Using Machine Learning”.
13. R. Keniya, A. Khakharia, and V. Shah, "Disease Prediction

- from Various Symptoms Using Machine Learning", Social Science Research Network (SSRN), 2020.
14. M. V. K. Reddy, G. V. P. S. Abhijith, K. S. Nath, and M. Sathyanarayana (2022), "Disease Predictor Based on Symptoms Using Machine Learning", International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 10, issue VI, June 2022.
  15. K. Gandhi, M. Mittal, N. Gupta, and S. Dhall (2020), "Disease Prediction Using Machine Learning", International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 8, issue 6, June 2020.
  16. A. K. Garate-Escamila, A. H. El Hassani, and E. Andres (2020), "Classification Models for Heart Disease Prediction Using Feature Selection and PCA", Informatics in Medicine Unlocked, vol. 19, 2020.
  17. T. Bhanuteja, K. V. N. Kumar, K. S. Poornachand (2021), "Symptoms Based Multiple Disease Prediction Using Machine Learning Approach", International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 10, issue 9, July 2021.
  18. R. D. Sah and J. Sheetalani (2017), "Review of Medical Disease Symptoms Prediction Using Data Mining Technique", IOSR Journal of Computer Engineering (IOSR-JCE), vol. 19, issue 3, May-June 2017.
  19. N. O. M. Salim and A. M. Abdulazeez (2021), "Human Diseases Detection Based on Machine Learning Algorithms: A Review", International Journal of Science and Business (IJSAB), vol. 5, issue 2, pp. 102-113, 2021.
  20. D. Shah, S. Patel, and S. K. Bharti (2020), "Heart Disease Prediction Using Machine Learning Techniques", SN Computer Science, vol. 1, article 345, 2020.
  21. P. Hamsagayathri and S. Vigneshwaran (2021), "Symptoms Based Disease Prediction Using Machine Learning Techniques", in Proceedings of the 3rd International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021), pp. 747-752.
  22. P. Zhang, X. Huang, and M. Li (2019), "Disease Prediction and Early Intervention System Based on Symptom Similarity Analysis", IEEE Access, December 2019, 7:176484-94.

# Gesture-controlled virtual mouse for hands free computing

Sahil Khalid Laskar<sup>1</sup> Tauhid Imam Chowdhury<sup>2</sup>  
Chinmoy Bharadwaj<sup>3</sup>

*Department of Computer Science, University of Science and Technology  
Meghalaya, India<sup>1,2,3</sup>*

## **Abstract :**

In a world of rapid technological advancements, traditional input devices like the mouse and keyboard are giving way to intuitive, contact-free alternatives. This project introduces a gesture-controlled virtual mouse system that leverages computer vision and machine learning to interpret hand gestures for performing mouse functions such as cursor movement, clicks, and more. Utilizing standard webcams, the system processes real-time video streams to detect gestures, translating them into mouse commands with high accuracy and minimal delay. This innovation offers significant benefits, including accessibility for people with disabilities, immersive gaming, and enhanced virtual reality interactions. Experimental results demonstrate the system's accuracy and responsiveness, showcasing its potential to revolutionize hands-free computing. By advancing gesture-based controls, the project highlights a key step toward more natural AI-driven interfaces.

**Keywords** - Gesture-controlled virtual mouse, Human-computer interaction, Computer vision, Gesture recognition, Hands-free computing.

## 1. Introduction

The way we interact with computers has changed a lot over the years. We have moved from using punch cards to keyboards and mice, and more recently to touchscreens and voice commands. However, there is still a need for more natural and easy ways to interact with computers. Gesture recognition, which lets users control devices with simple hand movements, offers a promising solution. This paper discusses the creation of a gesture controlled virtual mouse that allows for hands-free computing. By using computer vision and machine learning, our system can accurately understand hand gestures captured by regular webcams or smartphone cameras. These gestures are then translated into common mouse functions like moving the cursor, clicking, and scrolling. This technology not only helps people with physical disabilities but also has exciting uses in gaming, virtual reality, and other interactive areas. Our goal is to provide a smooth and responsive user experience, focusing on accuracy and low delay to make it practical.

## 2. Literature Review

Antony et al. (2024) [1] introduced an innovative system for controlling a virtual mouse using both hand gesture detection and voice commands, implemented in Python. The system combined real-time hand tracking and gesture recognition with speech recognition technologies, providing a hands-free, intuitive interface that enhanced accessibility for users with diverse needs, including those with physical disabilities. The approach demonstrated the feasibility of AI-driven HCI technologies, offering a more efficient and inclusive user experience. Future work was set to focus on improving accuracy and expanding functionality based on user feedback, aiming to create a more robust and versatile interface.

Rantamaa et al. (2023) [2] evaluated three VR interaction methods—mouse, hand tracking, and a combination of a VR stylus and grab-enabled controller—for their effectiveness in object manipulation and marking tasks in a medical surgery planning context. Results indicated that the controller-stylus combination was the most preferred due to its ability to mimic natural hand movements and provide high accuracy, while the mouse method proved to be the most precise. Hand-based interactions were found less effective, particularly for detailed marking tasks. The

controller-stylus combination excelled in handling complex 3D objects with better sensor fidelity, although the mouse demonstrated superior accuracy. Future advancements in sensing technologies were expected to improve the naturalness of hand-based interactions, potentially making them more feasible in VR environments.

Guha et al. (2022) [3] explored a novel AI-driven virtual mouse system that used hand gestures detected by digital or specialized cameras to perform traditional mouse functions like clicking and scrolling without physical devices. This study responded to the growing need during the COVID pandemic to minimize physical contact, highlighting the potential of advanced camera-based systems for gesture recognition. The system employed deep learning and computer vision algorithms, utilizing the YCbCr color model to reduce dependency on lighting conditions. The approach proved particularly beneficial for individuals with disabilities and had potential applications for various devices, including smartphones and TVs. Future developments were directed towards improving the accuracy of gesture recognition and reducing latency for more seamless interaction.

Dinh-Son et al. (2021) [4] describes a novel virtual-mouse method using RGB-D images and fingertip detection techniques was implemented in this project. which was published by Dinh-Son Tran et al. The hand region of interest and the center of the palm are first extracted from depth images provided by the Kinect V2 skeletal tracker and converted to binary images. The hand contours are extracted and described by a border-tracing algorithm. The K-cosine algorithm is used to detect the fingertip location, based on the hand-contour coordinates. Finally, the fingertip location is mapped to RGB images to control the mouse cursor based on a virtual screen and provided 96.13% accuracy. The proposed system work with a single low-cost CPU without the help of a graphics processing unit (GPU), has fast detection in real-time (30 frames per second (fps)), and allows execution on computer screens with many types of resolution. It provides simultaneous fingertip tracking for up to six people and selects the main person to control the mouse cursor, focusing on the right hand.

Tran et al. (2021) [5] uses fingertip detection and RGB-D images. Using detailed skeleton- joint information images from a Microsoft Kinect Sensor version 2, the hand region of interest and the

palm's center are first extracted, and they are then converted into a binary image. A border-tracing algorithm is then used to extract and describe the hands' contours. Based on the coordinates of the hand contour, the Kcosine algorithm is used to determine the location of the fingertip. Finally, the mouse cursor is controlled using a virtual screen by mapping the fingertip location to RGB images. Multiple restrictions that are primarily carried over from Microsoft Kinect continue to plague this study.

Chowdhury et al. (2020) [6] introduces an all-keyboard and mouse model that operates using a convex hull process; flaws are recorded or read, and using these flaws, the Mouse's functions are mapped. The convex hull treats the gap between the fingers as a defect because this image recognition process only considers defects and conditional statements, allowing it to be used for a variety of gestures and mapping commands.

Abiyev et al. (2020) [7] introduced an assistive human-machine interface designed for individuals with spinal cord injuries, allowing them to control a computer mouse through head movements and eye blinks. The system utilized a convolutional neural network (CNN) for facial recognition, detecting features such as the eyes, mouth, and nose. Head movements were translated into mouse pointer coordinates, and blinks were converted into mouse commands. Unlike sensor-based systems, this approach relied solely on a computer camera, eliminating the need for additional hardware. The system proved to be robust and accurate, providing a cost-effective and comfortable solution for disabled users to interact with a computer. The CNN's minimal preprocessing and high accuracy further enhanced the system's effectiveness, offering a practical alternative for computer interaction.

Varun et al. (2019) [8] defines models that are based on color detection and mouse movement based on highlighted colors provided by the user were developed by K.S. Varun et al. It is possible to see a two-figure input that creates two rectangles and an average point from both figures. It will function like the mouse pointer. The mouse pointer in the runtime follows the moving point as it moves. Therefore, using this, mouse movement can be implemented. The position of the predetermined colored caps in the mask that is created for system comprehension determines how the mouse pointers are updated. In order to detect the predetermined colored objects that will aid in mouse

movement, the created mask is converted from an RGB background to a black and white image and provided 84% accuracy. If the predetermined colored caps blend in with the background, they won't be seen and no mouse movement will be possible.

Patil et al. (2019) [9] developed a gesture-controlled system for Human-Computer Interaction (HCI), utilizing real-time cameras and computer vision technologies. Their approach used color pointers, such as fingertips captured by a webcam, to control mouse functions like clicking, scrolling, and cursor movement. Techniques such as image segmentation, background subtraction, and color tracking enabled precise gesture recognition without physical contact. The system, developed in Python for Windows 10, demonstrated the feasibility of performing standard mouse tasks and highlighted potential applications in presentations and workspace optimization. Future improvements aimed to expand functionality to include more complex gestures, such as window resizing and application control, to enhance user interaction and system efficiency.

Abhilash et al. (2018) [10] proposes the operation of a cursor control system using hand gestures captured from a webcam through a color detection technique performed in this project which was published by Abhilash SS et al. However, it was limited to a few mouse actions and is inoperable against a static background.

Haria et al. (2017) [11] emphasizes on a background extraction and contours detection system was proposed by Aashni Haria et al. where they conducted two sets of assessments in order to determine the correctness of their method. In the initial round of evaluations, they made use of settings with a variety of uniformly simple backdrops. For the second assessment, they utilized backdrops that had a number of discrepancies. Ten times were given for each gesture in each setting the numbers' average. The accuracy gained was 85% and 80%, which was calculated as the percentage of times a given gesture was successfully identified. But working with it is incredibly slow.

Zhou et al. (2016) [12] reported where a high-level hand feature extraction approach for real-time gesture detection was provided. The created system has good accuracy in both the extraction of flexional and extensional fingers. However, only computers with high configurations can use this method.

Banerjee et al. (2014) [13] defines there are numerous ways to control the cursor with a hand gesture, but for the longest period of

time, DataGlove must be worn. It lessens the effectiveness of the user's and the system's performance. The system's complexity in this method is a major problem.

Patil et al. (2013) [14] proposes a Machine-user interface that performs hand gesture recognition using multimedia techniques and basic computer vision. A paper was published on this topic by Ashwini M. Patil et al. Before utilizing the gesture comparison algorithms, they discovered a significant limitation. From the stored frames, hand segmentation and skin pixels must be completed.

Kao et al. (2011) [15] suggests applied visual hand gesture identification to the HCI interface, holding control usage, written by Chang-Yi Kao and Chin-Shyung Fahn. According to experimental findings, the face tracking rate is over 97% under typical circumstances and over 94% when the face has temporal occlusion. The system's execution efficiency is excellent, and we are inspired to market the robot soon. High configuration computers are required for accurate results.

### **3. Methodology**

This project involves developing a gesture-controlled virtual mouse system using computer vision and machine learning techniques. The methodology includes the following key components :

#### **A. Libraries and Tools:**

- OpenCV: Used for capturing video frames from the webcam, processing images, and displaying the output. It is essential for real-time image processing and handling tasks such as flipping the video feed, converting color spaces, and drawing landmarks on the detected hand.
- Mediapipe: A powerful library used for detecting and tracking hand landmarks in real-time. Mediapipe's hand tracking module is crucial for identifying the position of fingertips and other key points on the hand, which are used to interpret gestures.
- PyAutoGUI: This library is employed to control mouse functions programmatically. It allows the system to move the cursor, perform clicks, and take screenshots based on detected gestures.
- Pynput: Used for controlling and simulating mouse clicks, providing more precise control over the mouse actions, such as left-click and right-click.
- Utility Functions (util): Custom utility functions are implemented for calculating distances and angles between hand landmarks. These are used to determine specific gestures like clicks and dragging actions.

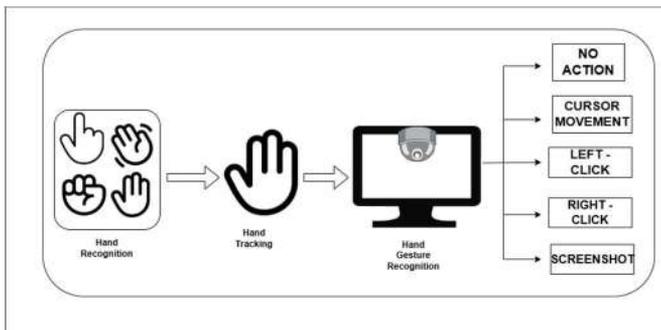
## B. Hardware :

- Webcam: A standard webcam is used to capture real-time video input, which is processed by the system to detect and interpret hand gestures.

## C. Gesture Detection and Processing :

- Hand Landmark Detection: Using Mediapipe, the system detects hand landmarks and tracks them in real-time. The position of these landmarks is used to identify key gestures.
- Gesture Interpretation: Custom algorithms are employed to interpret specific gestures, such as moving the mouse, left-click, right-click, double-click, and taking screenshots. These gestures are recognized based on the angles and distances between detected landmarks.
- Mouse Control: Based on the interpreted gestures, the system uses PyAutoGUI and Pynput to control the mouse cursor and perform actions like clicking and dragging.

## C. Detecting the Finger and Performing the Corresponding Mouse Function



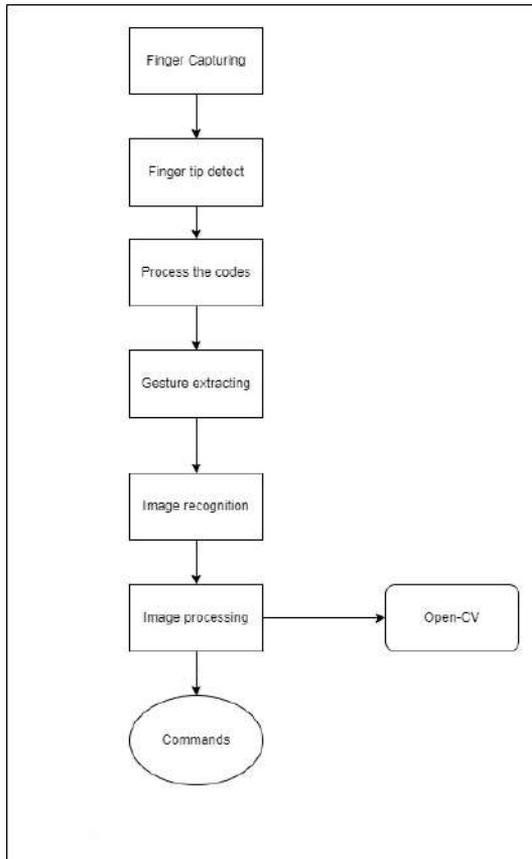
## D. Applications:

The AI virtual mouse system offers a wide range of applications, particularly in scenarios where traditional physical mice are either impractical or undesirable. By eliminating the need for physical devices, this system enhances human-computer interaction and provides a more intuitive and accessible interface.

Key applications include :

- The AI virtual mouse reduces the need for physical space required by a traditional mouse, making it ideal for compact workstations or environments where desk space is limited.

- In environments where physical contact with devices is a concern, such as during pandemics, this system provides a safer alternative by allowing users to control mouse functions without touching any surfaces, thereby reducing the risk of virus transmission.
- The system is particularly beneficial for individuals with physical disabilities or mobility issues in their hands. It allows them to perform mouse functions through simple hand gestures, offering an inclusive solution for accessing and controlling computers.
- The system's quick response and accuracy make it perfect for daily tasks by reducing any delay and ensuring that movements are tracked exactly.



3.1. Flowchart Of Gesture Recognition System

## 4. Result And Discussion

The system was tested with different hand gestures to ensure accuracy and responsiveness. The primary focus was on minimizing lag and improving the reliability of gesture recognition, particularly for complex actions like dragging and text selection.

### 1. The Hand Landmarker



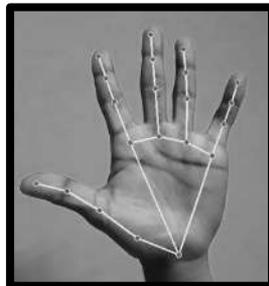
#### 4.1 Hand Coordinates or Landmarks

In this figure, we detect which finger is being used by identifying the fingertip using its ID and calculating the angle of the respective finger. This is accomplished with the help of MediaPipe, which provides the coordinates of the raised fingers. Based on this information, the corresponding mouse function is executed.

#### A. Hand Detection

The image shows an open hand with all fingers extended, with landmarks detected for accurate gesture recognition in the AI virtual mouse system.

If all the fingers are up with tip Id = 0, 1, 2, 3, and 4, the computer is made to not perform any mouse events in the screen.

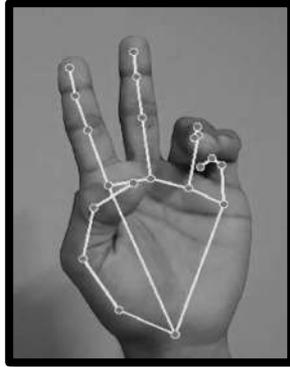


4.1(a). Detection Of Hand Landmarks

### **B. Mouse Movement**

To move the mouse, we have to use index finger and middle finger together - with all other fingers closed.

If the index finger is up with tip Id = 1 or both the index finger with tip Id = 1 and the middle finger with tip Id = 2 are up, the mouse cursor is made to move around the window of the computer using the AutoPy package of Python.



4.1(b). Mouse Movement Gesture

### **C. Right Click**

To use the right click, we have to use the index finger followed by thumb with all other fingers closed.

If both the index finger with tip Id = 1 and the middle finger with tip Id = 2 are up, the computer is made to perform the right mouse button click using the pynput Python package.



4.1(c). Right Click Hand Gesture

### **D. Left click**

To use left click, we have to use the middle finger along with the thumb followed by all other fingers closed.

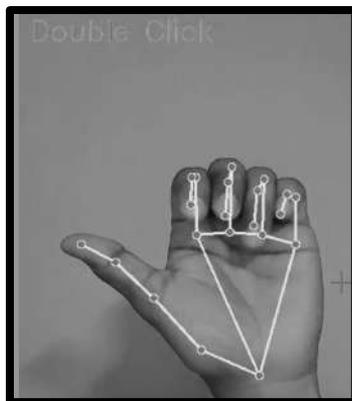
If both the index finger with tip Id = 1 and the thumb finger with tip Id = 0 are up, the computer is made to perform the left mouse button click using the pynput Python package



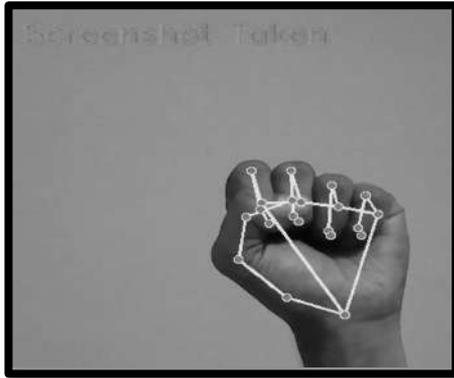
4.1(d). Left Click Hand Gesture

### **E. Double click**

To use double click, we have to use our index + middle finger together with the thumb opened and with all other fingers closed.



4.1(e). Double Click Hand Gesture



4.1(f).Screenshot Hand Gesture

**G. Accuracy Graph of Mouse Function**



**H. Table of Accuracy Percentage for Mouse Function**

<b>Mouse Functions</b>	<b>Accuracy Percentage</b>
Hand Detection	100
Mouse Movement	98
Right Click	97
Left Click	98
Double Click	99
Screenshot	100

## 5. Conclusion & Future Enhancement

In conclusion, the AI Virtual Mouse system demonstrates a promising innovation by replacing the traditional physical mouse with gesture-based control, using a webcam or built-in camera to detect and process hand movements. The system showcases superior accuracy compared to existing models, effectively addressing many of their limitations. Its potential extends to real-world applications, particularly in promoting touch-free interaction, which is advantageous in reducing the spread of COVID-19.

However, the system is not without challenges. Issues such as occasional difficulties with clicking and dragging, minor delays in gesture recognition, and sensitivity to varying lighting conditions impact the smoothness of the user experience. To overcome these shortcomings, future improvements will focus on refining fingertip detection algorithms, enhancing lighting adaptability, and optimizing system responsiveness for a seamless user interaction experience.

To improve the AI Virtual Mouse system and address its limitations, the following enhancements are proposed:

- **Advanced Algorithms:** Implement deep learning models to enhance gesture recognition accuracy and robustness.
- **Hand Tracking Improvements:** Refine tracking to handle occlusions, lighting challenges, and improve precision using depth sensors or hand pose estimation.
- **User Customization:** Enable system calibration for personalized gesture recognition based on hand shape and size.
- **Multi-User Support:** Introduce multi-user tracking for collaborative interactions using advanced sensing techniques.
- **Additional Features:** Integrate voice commands and gesture shortcuts for improved usability and efficiency.
- **Cross-Platform Compatibility:** Expand support to smartphones, tablets, and wearables for broader accessibility and functionality.

## References

1. Antony, E. (2024). Control the Virtual Mouse With Hand Gesture Detection and Voice Assistant Using Python. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.4864571>.

2. Rantamaa, H.-R., Kangas, J., Kumar, S. K., Mehtonen, H., Järnstedt, J., & Raisamo, R.(2023). Comparison of a VR Stylus with a Controller, Hand Tracking, and a Mouse for Object Manipulation and Medical Marking Tasks in Virtual Reality. *Applied Sciences*, 13(4), 2251. <https://doi.org/10.3390/app13042251>.
3. Guha, J., Kumari, S., & Verma, Prof. S. K. (2022). AI Virtual Mouse Using Hand GestureRecognition. *International Journal for Research in Applied Science and Engineering Technology*, 10(4), 3070–3076. <https://doi.org/10.22214/ijraset.2022.41981>.
4. Dinh-Son Tran & Ngoc-Huynh Ho & Hyung-Jeong Yang & Soo-Hyung Kim & Guee Sang Lee “Real- time virtual mouse system using RGB-D images and fingertip detection” 2021.
5. Tran, DS., Ho, NH., Yang, HJ. et al. “Real-time virtual mouse system using RGB-D images and fingertip detection” *Multimed Tools Appl* 80, 10473–10490, 2021.
6. S. R. Chowdhury, S. Pathak and M. D. A. Praveena, "Gesture Recognition Based Virtual Mouse and Keyboard" 4th International Conference on Trends in Electronics and Informatics (ICOEI), 2020, pp. 585- 589.
7. Abiyev, R. H., & Arslan, M. (2020). Head mouse control system for people with disabilities. *Expert Systems*, 37(1), e12398. <https://doi.org/10.1111/exsy.1 2398>.
8. K. S. Varun, I. Puneeth and T. P. Jacob, "Virtual Mouse Implementation using Open CV" 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 435-438.Patil, N., Sali, G., & Lokhande, N. (2019). Mouse on Finger Tips using ML and AI. *International Research Journal of Engineering and Technology (IRJET)*, 06(12).
9. S. S. Abhilash, L. Thomas, N. Wilson & C. Chaithanya, "VIRTUAL MOUSE USING HAND GESTURE," *International Research Journal of Engineering and Technology (IRJET)*, vol. 05 issue: 4, April 2018.
10. A. Haria, A. Subramanian, N. Asokkumar, S. Poddar & J. S. Nayak, “Hand Gesture Recognition for Human Computer Interaction," 7th International Conference on Advances in Computing & Communications, ICACC-2017, Cochin, India, *Procedia Computer Science* 115 (2017), pp.367–374, 22- 24 August 2017.

11. Y. Zhou, G. Jiang & Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition," in *Pattern Recognition* vol. 49, pp. 102-114, January 2016.
12. A. Banerjee, A. Ghosh, K. Bharadwaj, "Mouse Control using a Web Camera based on Color Detection," *IJCTT*, vol.9, March 2014.
13. A. M. Patil<sup>1</sup>, S. U. Dudhane<sup>1</sup>, M. B. Gandhi<sup>1</sup>, "Cursor Control System Using Hand Gesture Recognition", *International journal of advanced research in computer and communication engineering*, vol. 2, issue: 5, May 2013.
14. C.-Y. Kao, C.-S. Fahn, "A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion," in *Procedia Engineering* 15 pp. 3739 – 3743, 2011.

# Fingerprint Identification & Analysis through Deep Learning

Sumaiya Ashraf<sup>1</sup>, M. Jahirul Islam<sup>2</sup>

*<sup>1</sup>Student, Institute of Information and Communication Technology,  
Shahjalal University of Science and Technology, Sylhet Bangladesh  
Email:sumaiyaashraf1008@gmail.com*

*<sup>2</sup>Correspondence Author, Professor, Institute of Information and Communication  
Technology, Shahjalal University of Science and Technology, Sylhet Bangladesh,  
Email:jahir-cse@sust.edu*

## **Abstract :**

Fingerprints offer a powerful means of personal identification and still remains the most commonly used forensic evidence. It is specific to each person and remains consistent throughout the lifetime. It has the advantage of being non-transferable; every person has a unique blend of small ridges and patterns, which makes it unique and distinctive. Even the fingerprints of two identical twins do not match. It can play a crucial role in criminal investigation and is considered as the most reliable piece of evidence. Over a century fingerprint analysis has been used to identify suspects and solve crimes.

Gender, hand and finger identification early in the investigation phase can reduce the time to identify any criminal offenders and can stop any more occurrence of crime. But the fingerprints can be altered

to hide the evidence or to impersonate. This alteration can be intentional or caused by any accident or any disease.

In this work, we use deep Convolutional Neural Network (CNNs) to classify around 6000 real fingerprints and around 40000 altered by gender, hand and finger attributes and also detect the alterations and identify the types of alterations. Transfer learning is used to speedup training of CNN. The performances of the models are promising and can be useful to the authorities to identify criminal offenders and to quickly take necessary steps.

**Key-words** : Fingerprint Identification, Convolutional Neural Network, Image Classification

## **Introduction**

Fingerprint is the impression left by the friction ridges of a human finger. The impression that is made by papillary ridges on the fingers and thumb. It affords an impeccable means of personal identification, because the ridge arrangement on every finger of every human being is unique and does not alter with growth or age. Human fingerprints are detailed, nearly unique, difficult to alter and durable the whole lifetime which makes it suitable for long-term human identity. There are various ways to identify a person uniquely, biometrics have been one of the most secure options so far, which is virtually impossible to imitate by anyone other than the desired individual. Fingerprint is a physiological feature, a category of biometrics which deals with the characteristics possessed by a person. It has been used in various applications such as forensics, transaction authentication, cellphone unlocking, attendance management, civil identification etc. Fingerprints are used as physical evidence in a crime scene. It is used to identify criminals which is the most frequently used technique in forensic science. It is considered better than DNA because DNA matching is effective till we meet the identical twin or triplet. No two fingerprints are alike, they are unique for each individual and are persistent, and don't change overtime. The uniqueness of the fingerprint features can reduce the identification difficulties; detect impersonation which can speed up investigation by reducing the time to identify a suspect. Furthermore, it can also help the authorities by guiding them to correctly identify a suspect when matching the suspect's fingerprints among the large number

of matches in the fingerprint database. The matching of the suspect's fingerprints with the large volume of data in the fingerprint database is an extremely time consuming task. However knowing different attributes such as gender, hand, finger or any alteration in the fingerprint may reduce the time of investigation. The need to process and study different alteration of the real fingerprints will help in detection of any impersonation or forgery or any fakeness. However, fingerprint identification is still far from being perfect, still not a solved problem. Even now there is lots of work to do, as the biometric system becomes larger and more inclusive as new challenges arise. Here deep learning approaches can assist to improve the performance and potential. As the emerging technology deep learning achieves tremendous success in many domains of image processing and pattern recognition. This work assesses the application of such innovative deep learning approaches to fingerprint identification and analysis.

### **1.1 Challenges**

The challenges are,

- Collecting such a dataset containing correct and real data is a challenging task itself.
- One of the major challenges is to obtain all the desired information from the dataset as required.
- Data analysis itself is a crucial task.
- Low-configured laptops face difficulties processing large amounts of data.

### **1.2 Project Objectives**

For this project, the main objectives are:

- To analyze the real and altered data.
- To analyze the different levels of altered data.
- To extract the features from real data.
- To extract the altered features from altered data.
- To use transfer learning methods for recognition of gender, hands and fingers, detection of altered fingerprints and identification of types of alterations.
- To find which Deep Learning Algorithms have higher accuracy.

## **Ii. Literature Review**

To perform the research and have a clear idea we have reviewed some of the existing works available.

## 2.1 Background Study

Yahaya Shehu; Ariel Ruiz-Garcia et al.(2018) [1] This work presents a review and uses deep Convolutional Neural Networks to classify fingerprints according to gender and also identify individual hand and finger features through data analysis from the publicly available Sokoto Coventry Fingerprint Dataset. Using two different data approaches-ResNet34 and ResNet18-75.2%, 93.5%, and 76.72% accuracy has been achieved.

Yahaya Shehu; Ariel Ruiz-Garcia et al.(2018) [2] This paper proposes a modified fingerprint dataset to increase research accessibility. It talks about the need for research on altered fingerprints and the need for automatic detection of altered fingerprints. Data is used from the publicly available Sokoto Coventry Fingerprint Dataset (SOCOFing). For this study, both the original and altered datasets are considered. The model using ResNet18 yielded an accuracy of 98.5%.

Oliver Giudice; Mattia Litrico et al.(2020)[3]This paper will use the Inception-V3 architecture to detect altered fingerprints, identify different types of alteration, and recognize gender, hands, and fingers. Data used in this work is taken from the publicly available dataset-Sokoto Coventry Fingerprint Dataset (SOCOFing). Using InceptionV3, the accuracy achieved for classification tasks was 98.21%, 98.46%, 92.52%, 97.53%, and 92.18%, respectively.

Yousif Ahmed Al-Wajih; Waleed M. Hamanah et al.(2022)[4] This paper proposes a new approach for the classification of fingerprints. The research makes use of two datasets: the NIST and the SOCOFing-Sokoto Coventry Fingerprint Dataset. The best results were shown by TCNN with high accuracy: 90% and 89% fingerprint type classification.

Yahaya Shehu; Ariel Ruiz-Garcia et al.(2018) [5] The proposed dataset consists of 6000 fingerprint image data from 600 African people. Data included 10 fingerprints per subject, and all subjects were at least 18 years of age. The dataset contains two sub-folders containing real and altered images. The altered folder is further divided into three levels of alterations in difficulty, such as altered-easy, altered-medium, and altered-high. The dataset has some special features like gender, hand, finger, and other various changes in the fingerprint image. The provided data is a single impression. The images are labeled regarding gender and finger position. The real images were complemented with synthetically damaged versions.

### 2.1.6 Summary

Different strategies and different techniques are used in fingerprint analysis. From reviewing the relevant literature we came to know that some of the deep learning algorithms can provide some good amount of results. But at first we need to prepare the data by preprocessing and to get the results by applying a different deep learning algorithm.

### **iii. Dataset**

#### **Dataset**

Dataset containing the 6000 fingerprint image data of 600 African people was imported from kaggle. It is a biometric fingerprint database designed for academic research purposes. Dataset contains unique attributes such as labels for gender, hand and finger name, and synthetically altered versions with three different levels of alteration for obliteration, central rotation, and z-cut. The dataset contains two main directories, Real and Altered directory. The altered directory is divided into: Altered-Easy, Altered-Hard, Altered-Medium. The Altered directory contains synthetically altered versions of fingerprint biometric images with three different levels of alteration for obliteration, central rotation, and z-cut. The Real directory contains real human fingerprints with no alteration.

**Real dataset:** The real dataset contains 6000 data



**Figure 3.1: Sample Images of Real Fingerprint**

**Altered-Easy:** The altered-easy dataset contains 17934 data,



**Figure 3.2: Sample Images Altered-Easy Fingerprint**

**Altered-Medium:**The altered-medium dataset contains 17067 data,



**Figure 3.3:Sample Images Altered-Medium Fingerprint**

**Altered-Hard:**The altered-Hard dataset contains 14272 data,



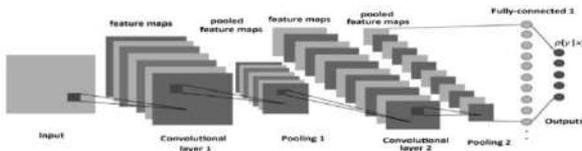
**Figure 3.4:Sample Images Altered-Hard Fingerprint**

#### **Iv.Methodology**

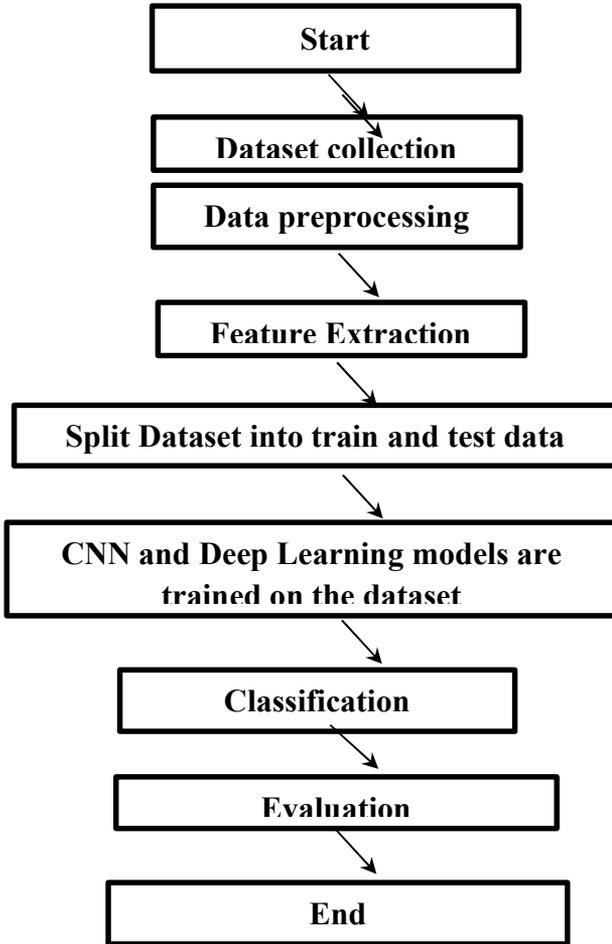
The proposed method will be briefly discussed in this section.

**4.1 The Architecture of the Visual Cortex and Convolution Layers :** Convolutional Neural Network or CNN is a deep learning algorithm used especially for image processing and recognition. It is a multi-layer neural network type meant to discern visual patterns from pixel images. In CNN, it is termed a mathematical function 'convolution'. It is a type of linear operation in which you can multiply two functions to get a third function representing how the shape of one function can be altered by another. A Convolutional neural network comprises three layers that act as the building blocks, convolutional layer, pooling layer, and a fully connected layer. CNN consists of filters that can convert 2D data into 3D data. The CNN system is quite similar to the human visual system. Applications involving computer vision models using the CNN system have achieved performances as good as that by humans, such as image classification and object detection.

**Figure 4.1.1: The structure of a CNN, consisting of convolutional, pooling, and fully-connected layers.**

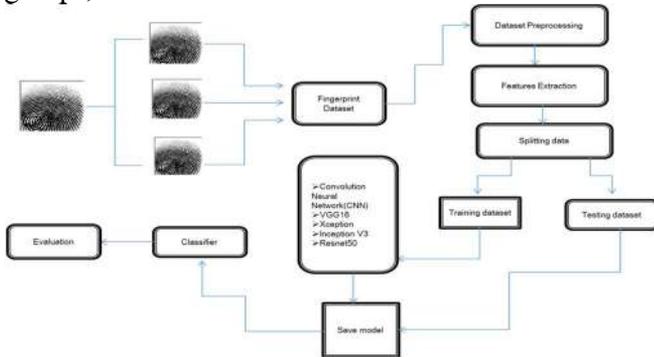


**4.2 Block diagram of the proposed system :**The methodology of the proposed system,



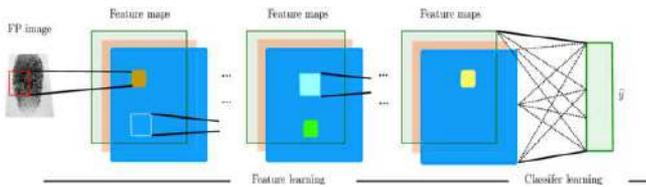
*Figure 4.2: Methodology of proposed system*

**4.2.1 Implementation :** The methodology is breakdown into following steps,



**Figure 4.2.1: Implementation of the proposed system**

### 4.2.2 The structure of CNN



**Figure 4.2.2: The structure of CNN for Fingerprint image**

All the steps consisting the diagram would be implementing chronologically and described in the following,

#### 4.2.3 Data Collection:

As mentioned earlier Sokoto Conventry Fingerprint Dataset (SOCOFing) is used to contain the fingerprint image data of 600 African people imported from kaggle.com. The dataset is divided into two subfolders containing real and altered images. The altered folder is further divided into three levels of alterations difficulty such as altered-easy, altered-medium, altered-high. There is around 55000 data in the dataset.

#### 4.2.4 Data Pre-processing:

The pre-processing of the dataset is necessary for improved results. Accordingly, four data sets are processed according to their requirement.

#### 4.2.5 Feature Extraction:

From the Sokoto Coventry Fingerprint Dataset (SOCOFing) dataset the unique attributes are extracted such as labels for gender, hand, finger and synthetically altered versions with three different levels of alteration for obliteration, central rotation, and z-cut.

#### 4.2.6 Splitting Dataset into Train and Test Set:

We have used 70% of the data on training purpose and 30% of data on testing purpose.

#### 4.2.7 Selecting Deep Learning Models:

We have selected some deep learning algorithms to perform with processed data. The Deep Learning algorithms are :

- **Convolutional Neural Network (CNN)** : In deep learning, Convolutional Neural Network (CNN) is a class of artificial neural networks applied to analyze visual images. CNN consists of many input layers, hidden layers and output layers.
- **ResNet-50**: ResNet-50 or Residual Neural Network is a 50 layer Convolutional Neural Network, from which 48 are convolutional layers and one is max pool layer and another is average pool layer. It is a Deep Learning model for image classification and computer vision application.
- **InceptionV3**: InceptionV3 is a deep learning model based on the Convolutional Neural Network, which is used for image classification. It is used for assisting in image analysis and object detection. The model is buildup of symmetric and asymmetric building blocks, including convolutions, max pooling, average pooling, concentration, dropouts and fully connected layers. It is a Convolutional Neural Network 48 layers deep.
- **VGG-16**: The VGG-16 is a 16-layer deep neural network. It's a deep Convolutional Neural Network model for image recognition. It's special since it only has 16 layers that have weights-as opposed to relying on a large number of hyper-parameters. It is considered among the best vision model architectures. It has 16 Convolutional layers, a max pooling layer, 3 dense fully connected layers, and an output layer.
- **Xception**: Xception stands for extreme inception is a Convolutional Neural Network that is 71 layers deep. It

involves Depthwise Separable Convolution. The architecture has 36 Convolutional layers forming the feature extraction base of the network. The data first goes through the entry flow, then through the middle flow which is repeated eight times and finally through exit flow.

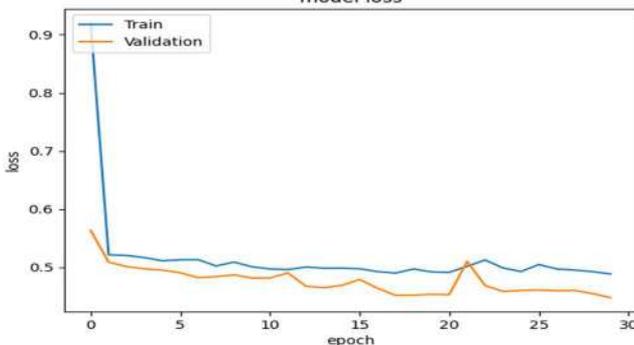
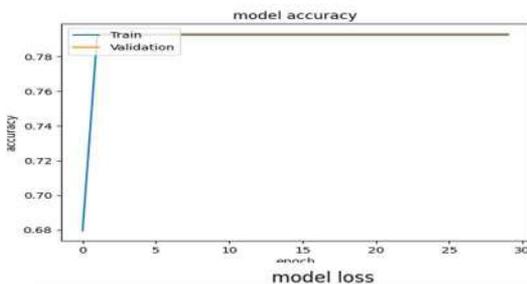
## V. Features

Dataset contains unique attributes such as labels for gender, hand and finger name, and synthetically altered versions with three different levels of alteration for obliteration, central rotation, and z-cut which has been extracted. The extracted features are mentioned below,

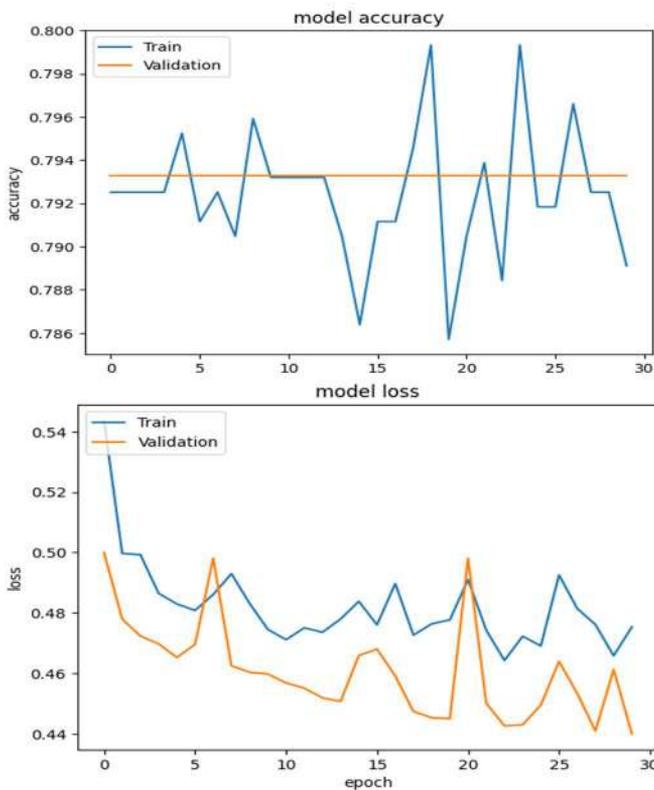
- **Gender Classification:** Gender classification is a binary classification problem. Used to reduce gender identification difficulties due to any disease or injury and detect gender impersonation.
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer = Adam, Learning rate = 0.01, Epochs = 30 and Batch size = 128.

### → Model evaluation using CNN

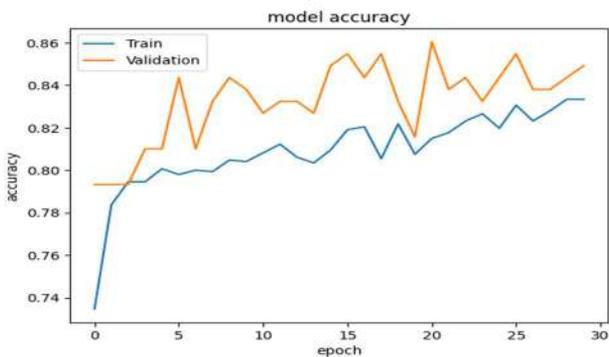
#### Accuracy & Loss

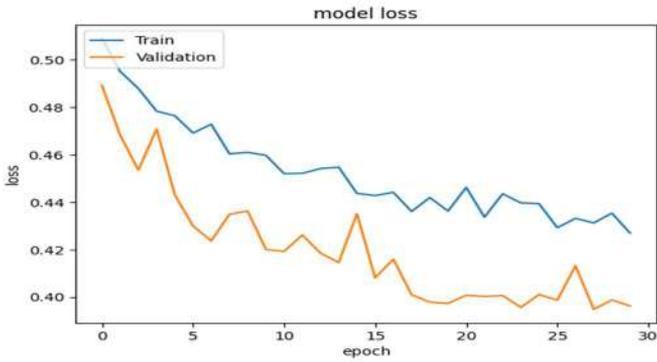
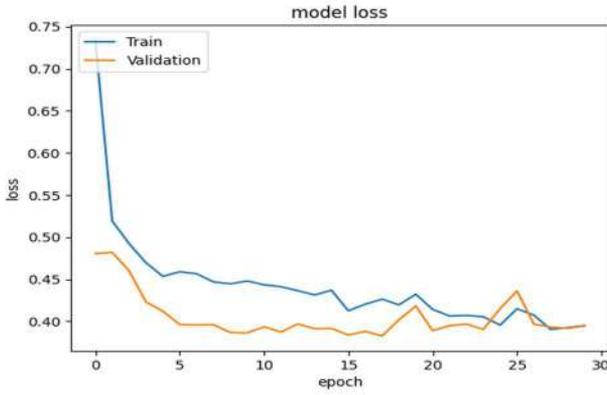


→ **Model Evaluation using ResNet-50**  
**Accuracy & Loss**

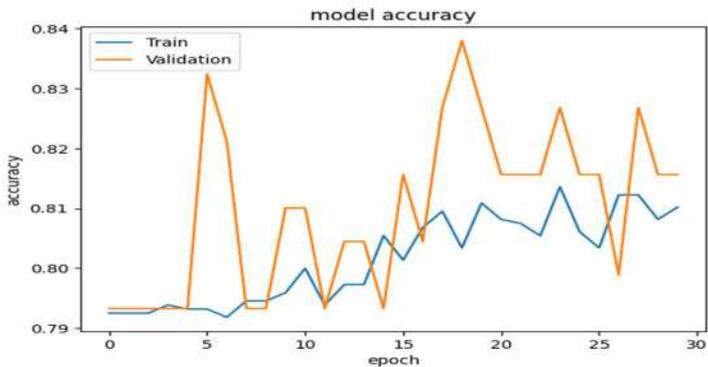


→ **Model Evaluation using InceptionV3**  
**Accuracy & Loss**

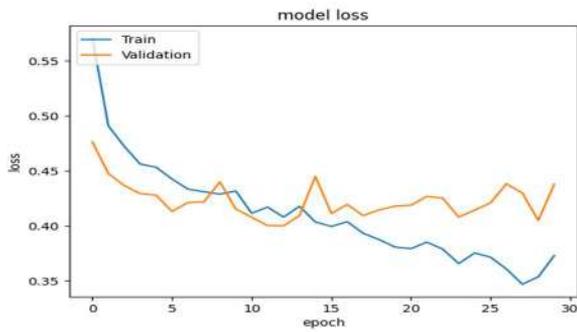
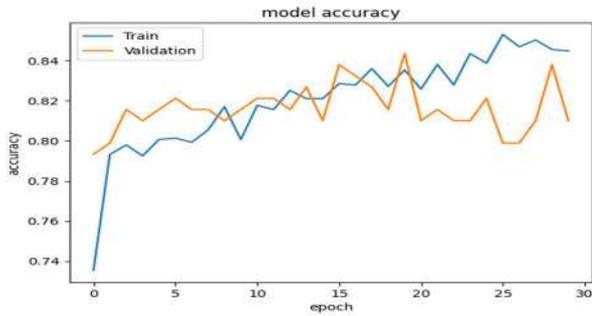




→ **Model Evaluation using VGG16**  
**Accuracy & Loss**



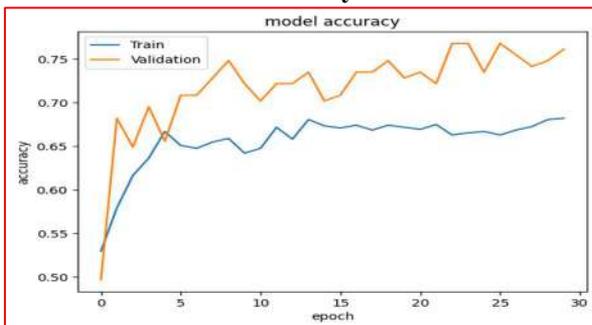
→ **Model Evaluation using Xception  
Accuracy & Loss**

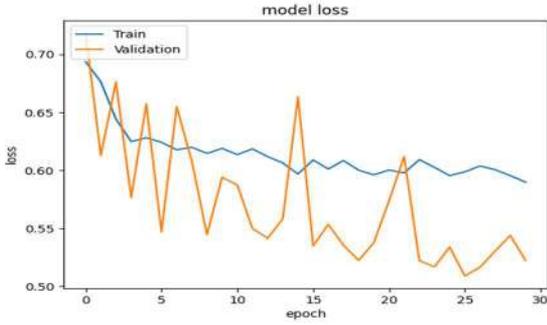


● **Hand Classification:**

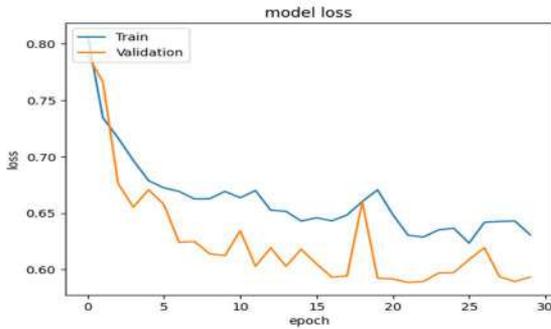
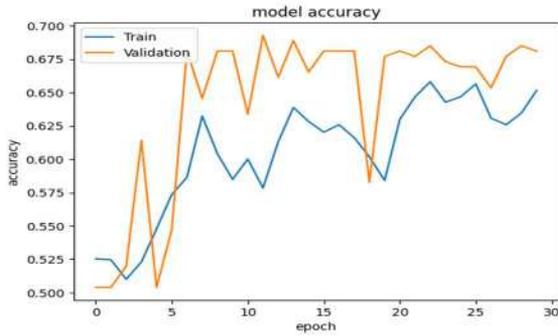
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.01,Epoches=30 and Batch size =128.

→ **Model Evaluation using CNN  
Accuracy & Loss**

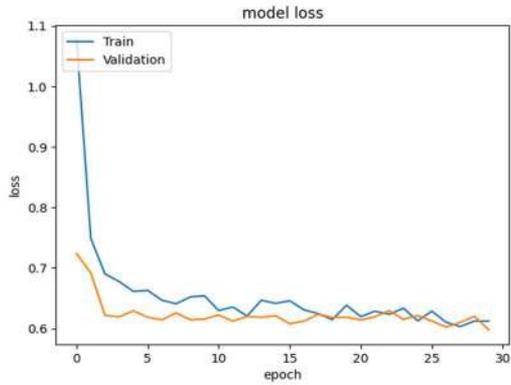
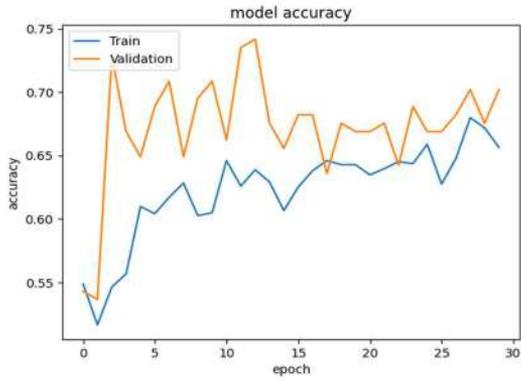




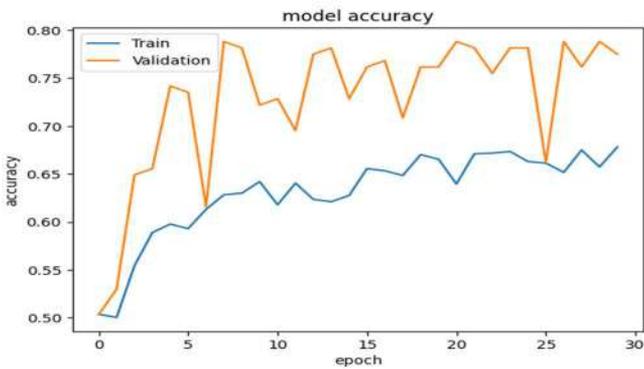
→ **Model Evaluation using ResNet-50  
Accuracy & Loss**

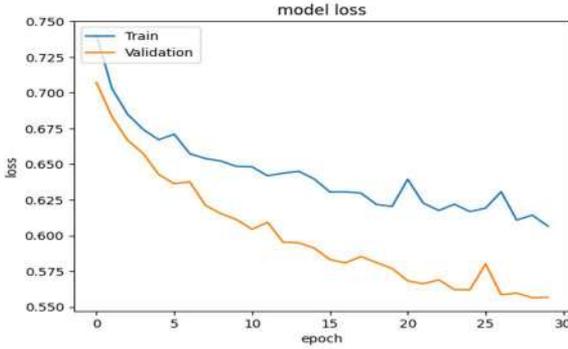


→ **Model Evaluation using InceptionV3  
Accuracy & Loss**

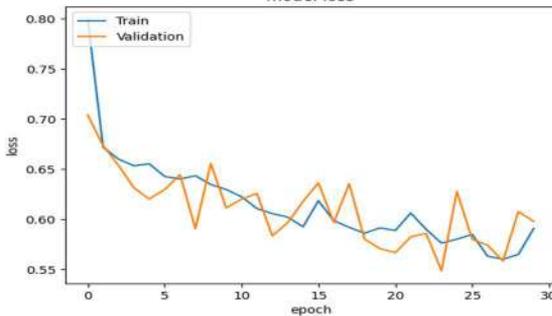
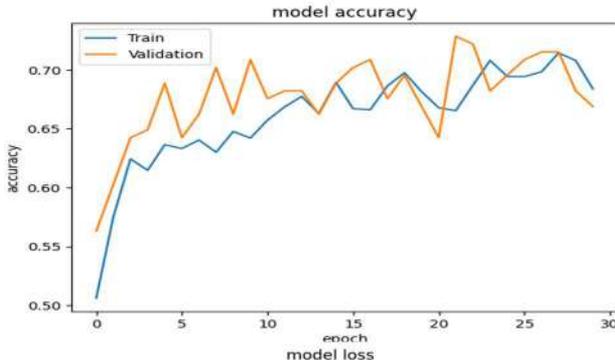


→ **Model Evaluation using VGG16  
Accuracy & Loss**





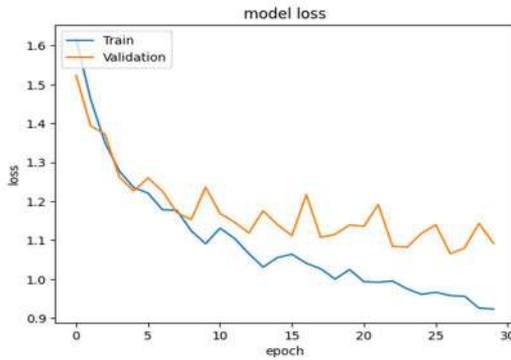
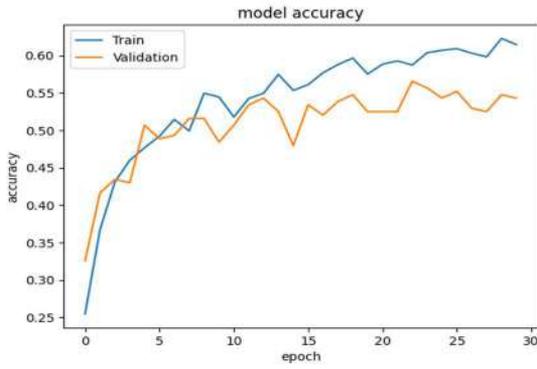
→ **Model Evaluation using Xception Accuracy & Loss**



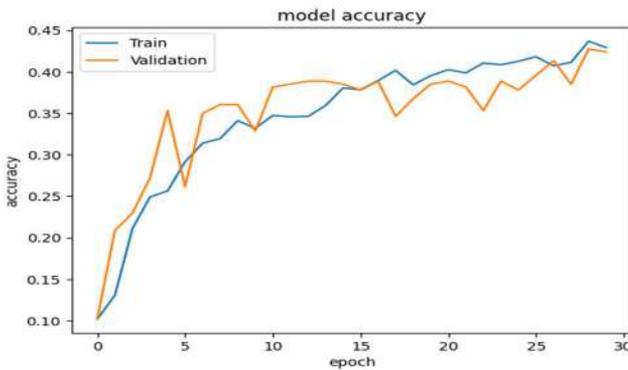
● **Finger Classification:**

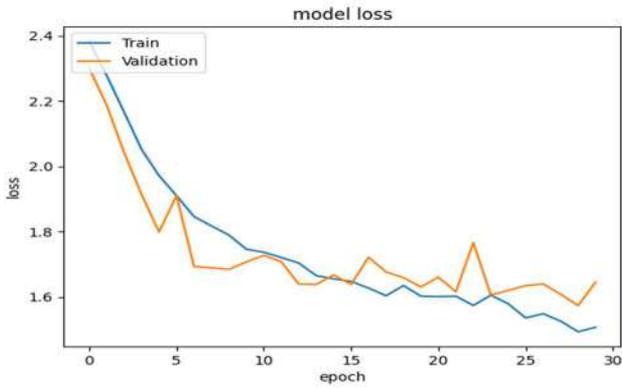
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.01,Epochs=30 and Batch size =128.

→ **Model Evaluation using CNN**  
**Accuracy & Loss**

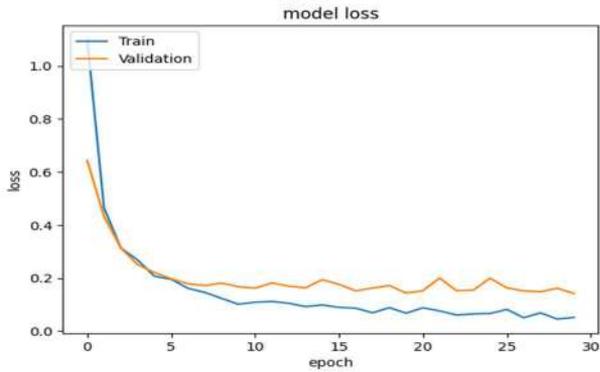
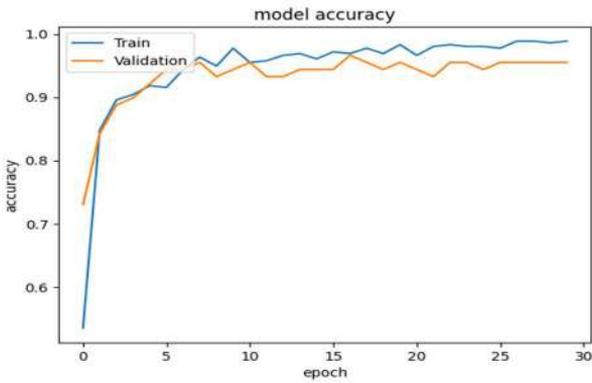


→ **Model Evaluation using ResNet-50**  
**Accuracy & Loss**

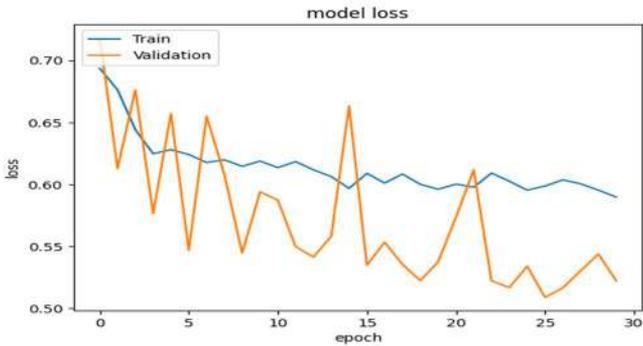
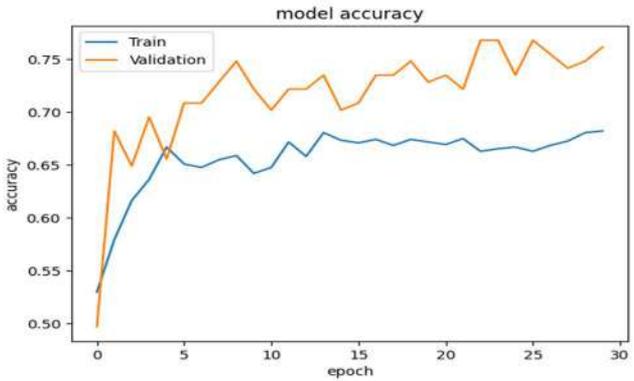




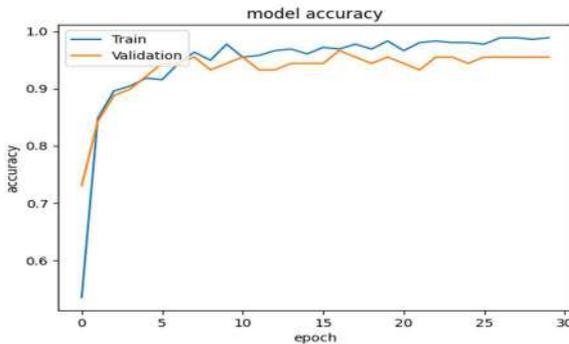
→ **Model Evaluation using InceptionV3  
Accuracy & Loss**

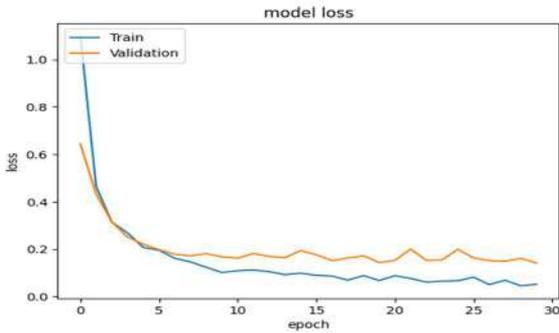


→ **Model Evaluation using VGG16**  
**Accuracy & Loss**



→ **Model Evaluation using Xception**  
**Accuracy & Loss**



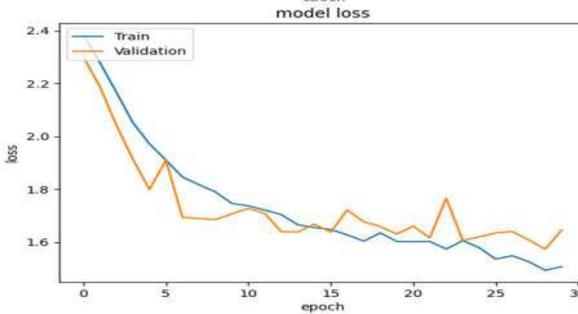
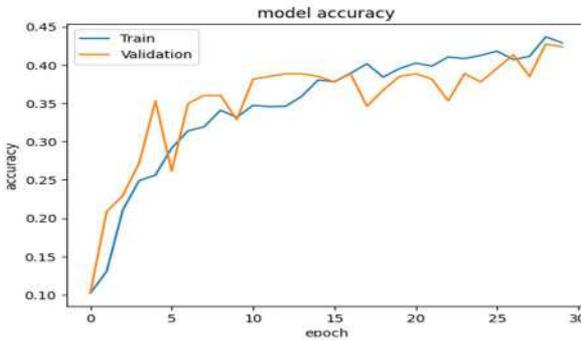


- **Hand & Finger Classification:**

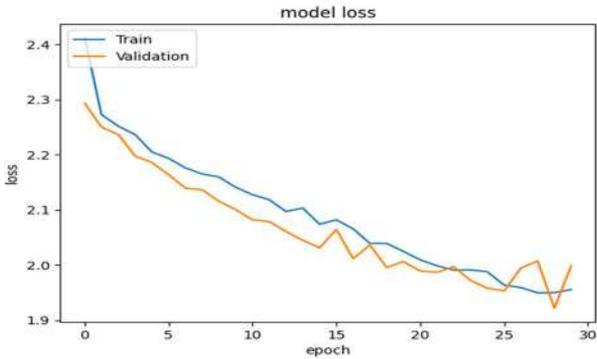
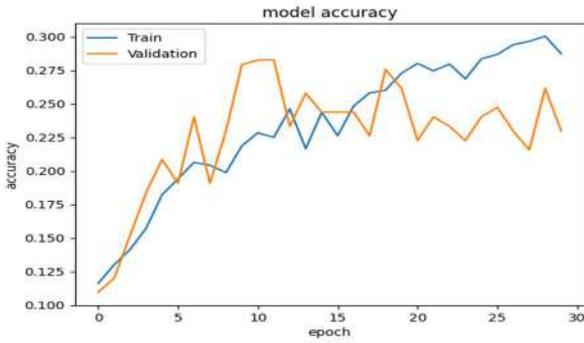
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.01,Epoches=30 and Batch size =128.

- ➔ **Model Evaluation using CNN**

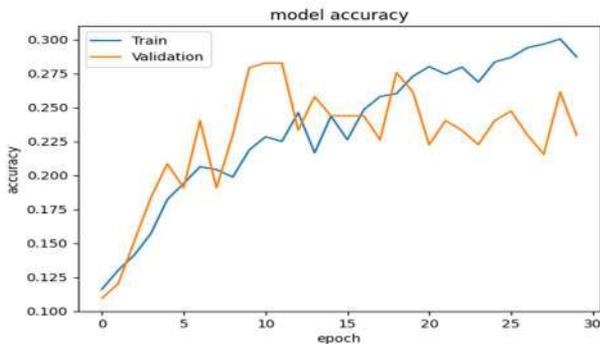
**Accuracy & Loss**

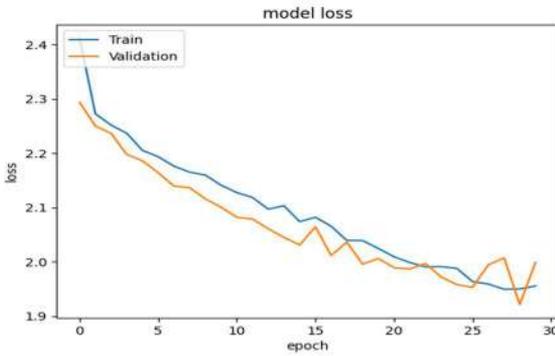


→ **Model Evaluation using ResNet-50**  
**Accuracy & Loss**

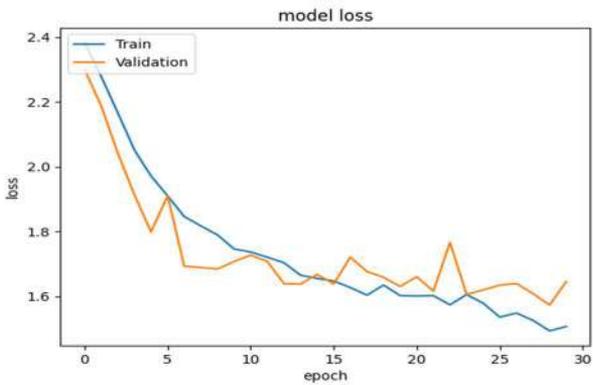
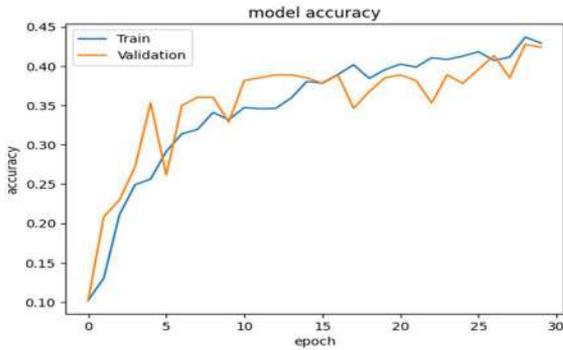


→ **Model Evaluation using InceptionV3**  
**Accuracy & Loss**

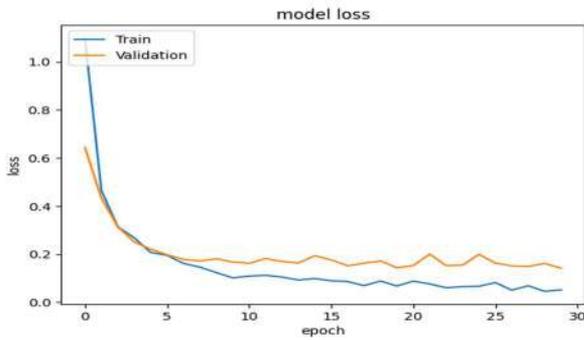
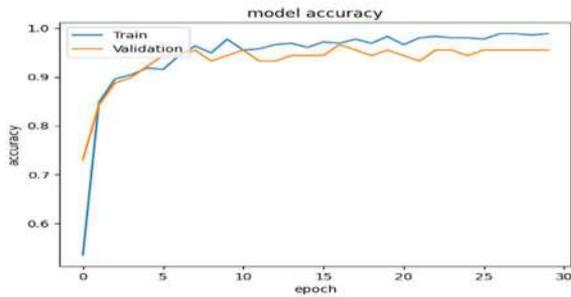




→ **Model Evaluation using VGG16**  
**Accuracy & Loss**

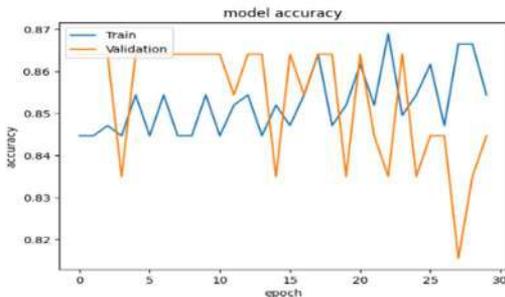


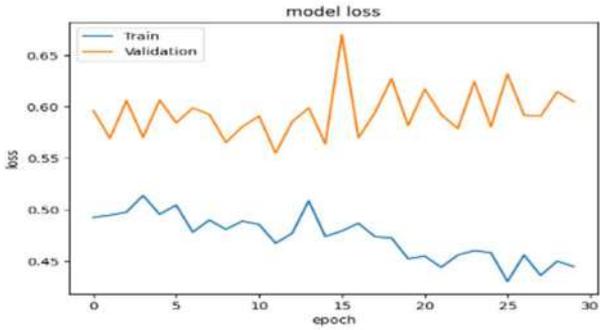
→ **Model Evaluation using Xception  
Accuracy & Loss**



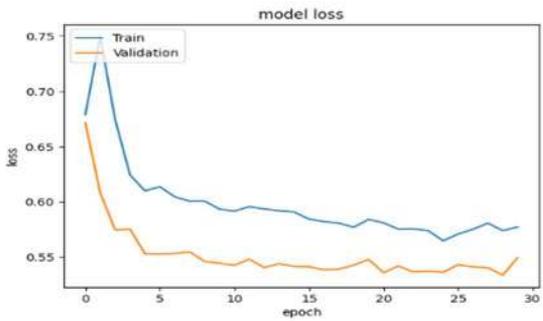
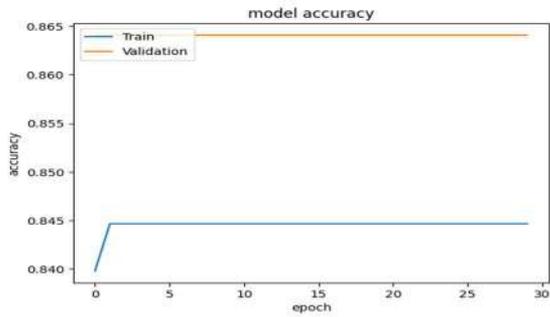
- **Altered and Real type classification:**
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.01,Epoches=30 and Batch size =128.

→ **Model Evaluation using CNN:  
Accuracy & Loss**

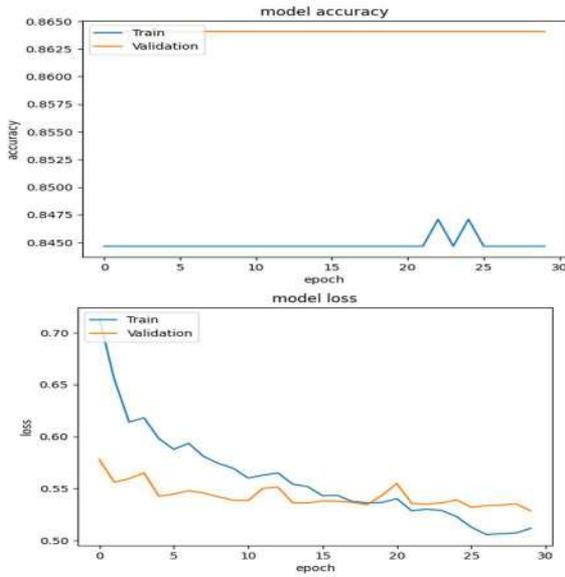




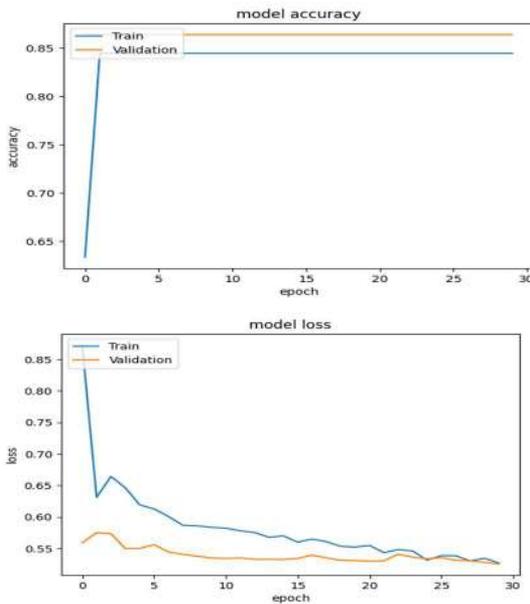
→ **Model Evaluation using ResNet-50  
Accuracy & Loss**



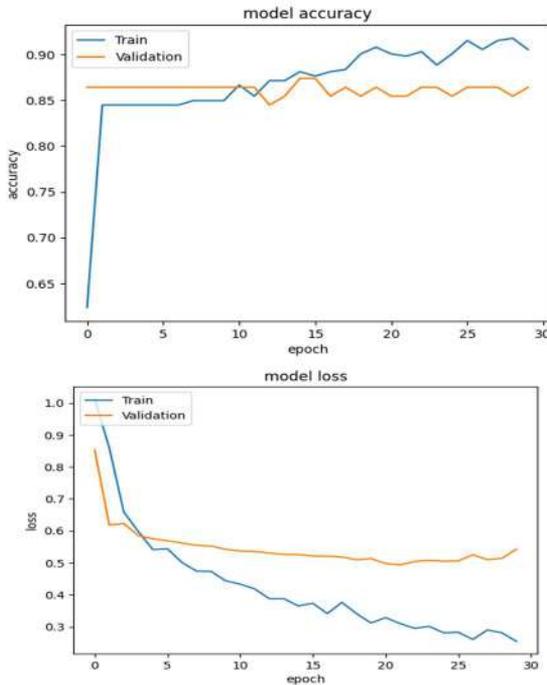
→ **Model Evaluation using InceptionV3**  
**Accuracy & Loss**



→ **Model Evaluation using VGG16**  
**Accuracy & Loss**



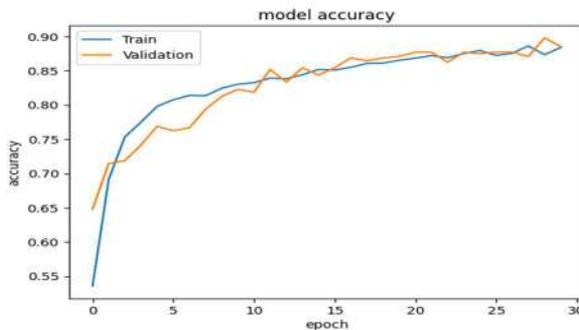
→ **Model Evaluation using Xception  
Accuracy & Loss**

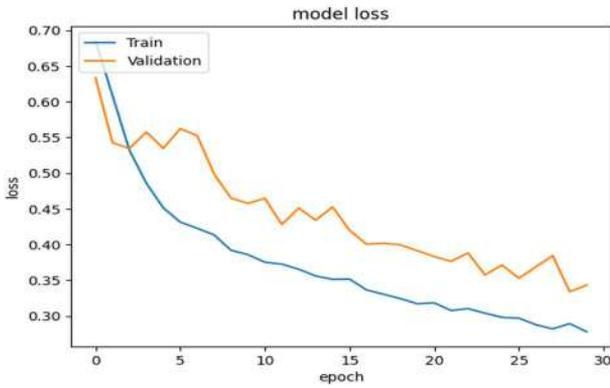


● **Altered and Real Classification:**

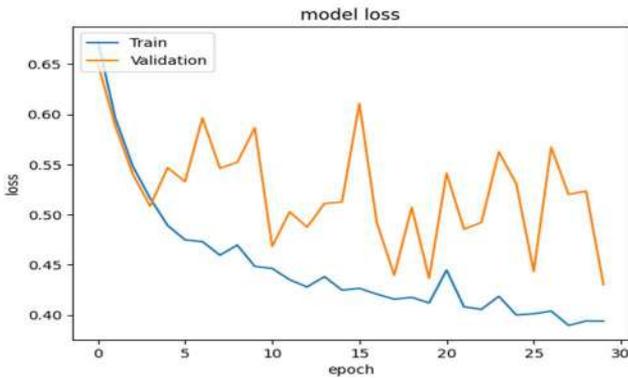
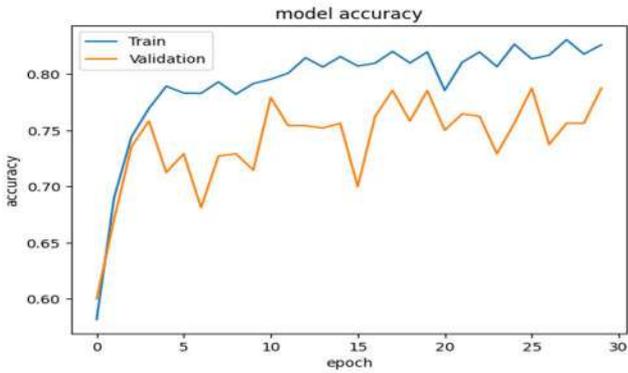
→ The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.001,Epoches=30 and Batch size =128.

→ **Model Evaluation using CNN  
Accuracy & Loss**

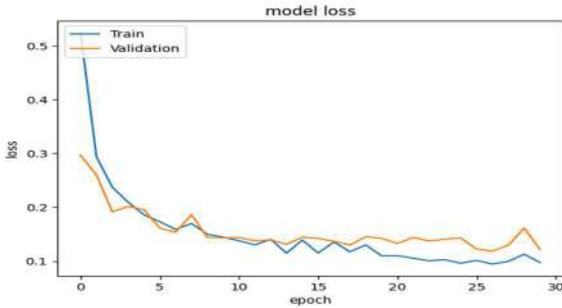
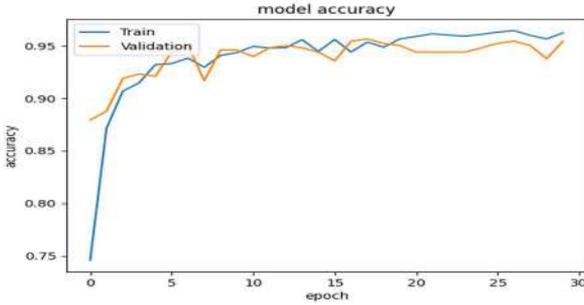




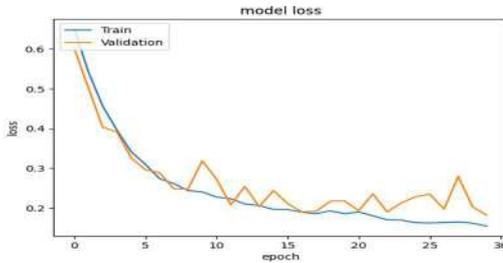
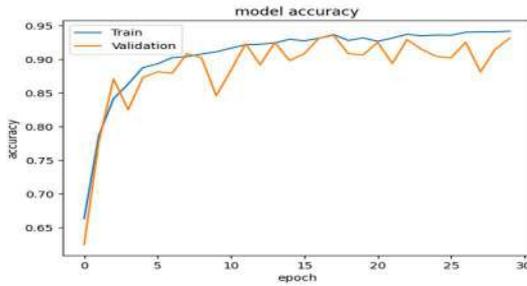
→ **Model Evaluation using ResNet-50**  
**Accuracy & Loss**



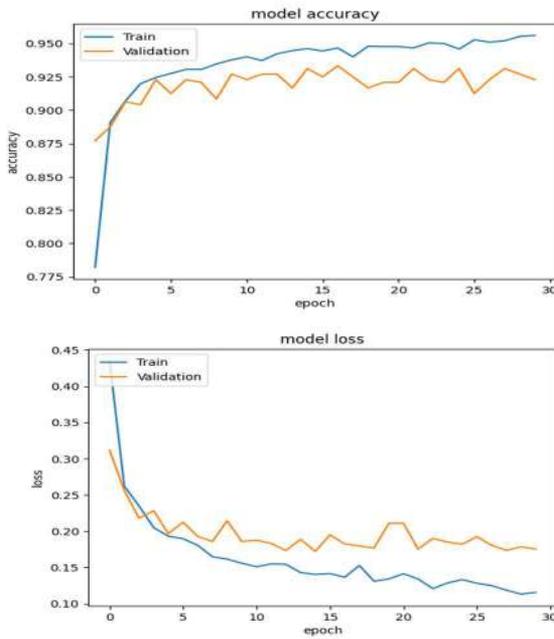
### → Model Evaluation using InceptionV3 Accuracy & Loss



### Model Evaluation using VGG16 Accuracy & Loss

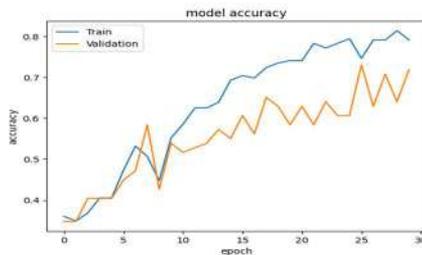


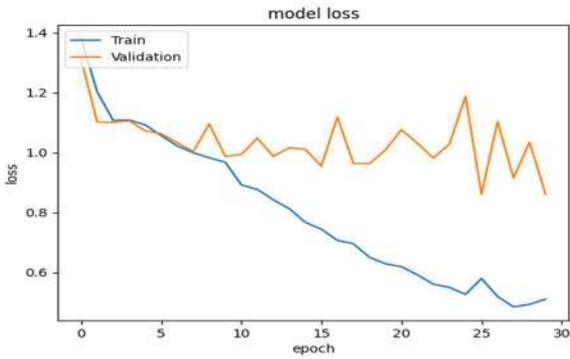
→ **Model Evaluation using Xception Accuracy & Loss**



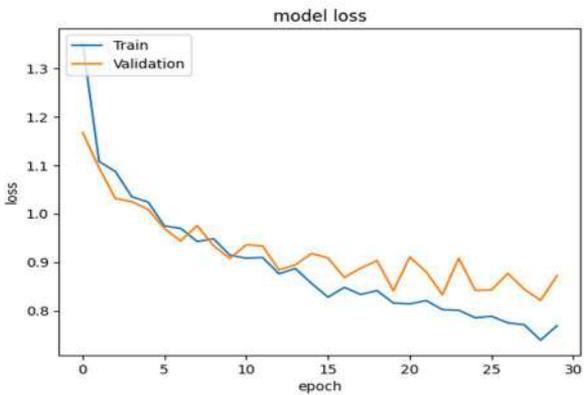
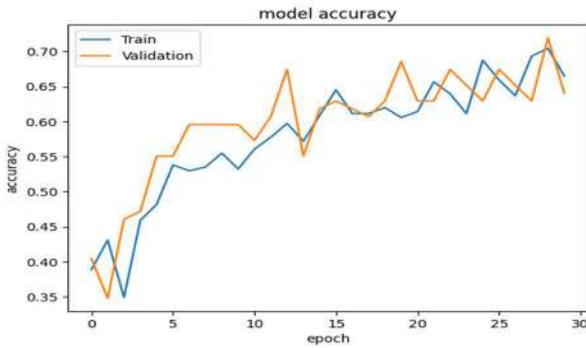
- **Alteration type Classification:**
- The CNN was trained from scratch and trained with the following hyper-parameters: optimizer =Adam ,Learning rate = 0.01,Epoches=30 and Batch size =128.

→ **Model Evaluation using CNN: Accuracy & Loss**

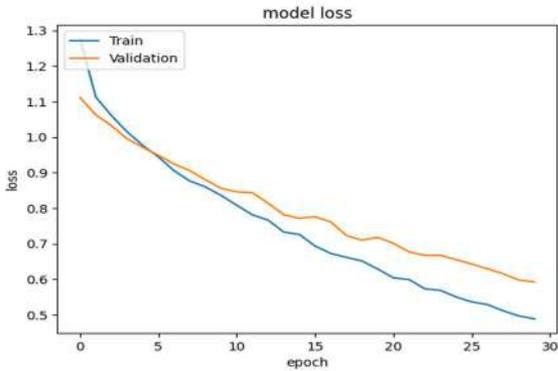
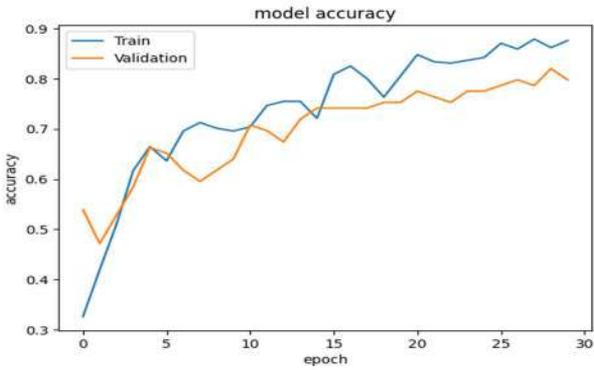




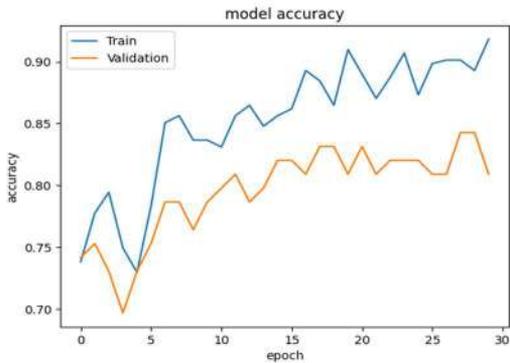
➤ **Model Evaluation using ResNet-50 Accuracy & Loss**

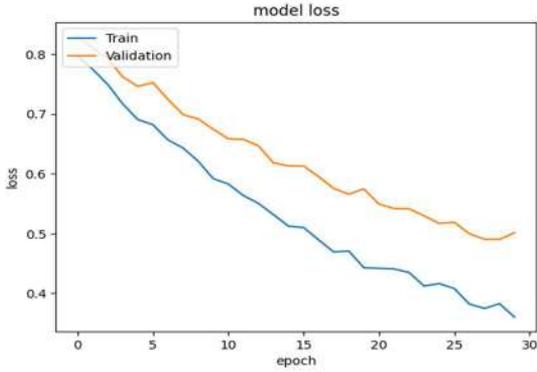


➤ **Model Evaluation using InceptionV3  
Accuracy & Loss**

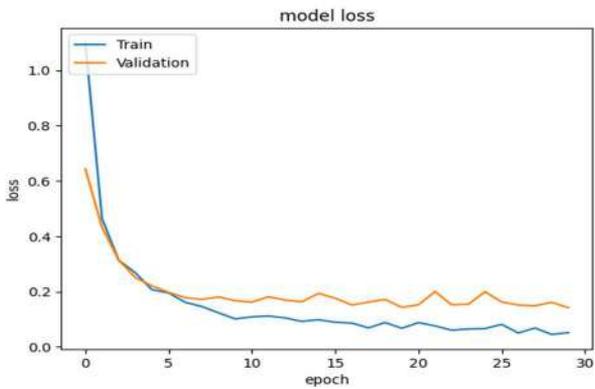
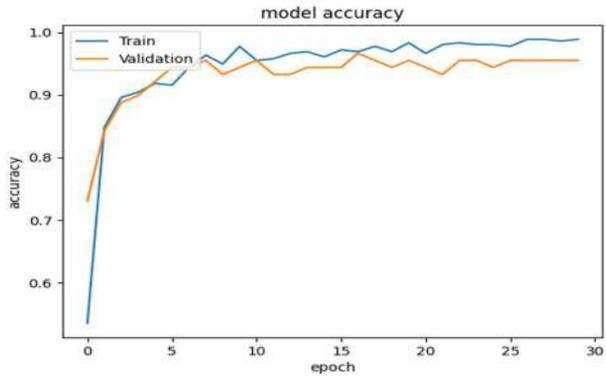


➔ **Model Evaluation using VGG16  
Accuracy & Loss**





→ **Model Evaluation using Xception Accuracy & Loss**



## Vi.Result Analysis

• Analyzing the results of various extracted features of the fingerprints from the Socofing Dataset using Deep Learning methods:

CLASSIFICATION	CNN	ResNet50	Inception V3	VGG16	Xception
Gender	79.13%	78.80%	81.12%	81.13%	81.12%
Hand	74.05%	68.11%	68.50%	73.62%	67.62%
Finger	62.06%	57.05%	66.50%	71.44%	68.43%
Hand & Finger	44.17%	33.24%	35.50%	51.04%	53.23%
Altered & Real type	84.81%	83.96%	83.96%	83.96%	83.49%
Altered & Real	86.50%	79.12%	95.85%	94.24%	93.50%
Alteration type	86.40%	66.99%	87.37 %	89.80%	93.20%

**Table: Result Analysis of Different extracted features of Fingerprint Classification**

## VII. Conclusion & Future Work

This paper demonstrates the classification and analysis of various fingerprints data of 600 African subjects. Here we have analyzed and classified the data provided by faculties of Coventry University, Coventry, UK and Nottingham Trent University, Nottingham, UK. In this work we processed the raw dataset, extracted the required features and applied different deep learning models to estimate which model might provide better performance. The above-mentioned tables show the results of the performance measures of various deep learning models for classification and analysis of the unique attributes of fingerprints.

From the obtained results, we can certainly say that all of the unique attributes cannot be solved by a single architecture. In spite of all this, the project is confined to its dataset. The dataset contains around 55000 image data each of them contains labels of unique attributes, which has to first processed then training of the data is possible. 6000 data are real data and rest are synthetically altered data are used due to lack of publicly available actual altered data. We hope that this analysis will support or assist the authorities by reducing the time to identify a suspect, detect impersonation and to detect the offenders if there is any self-made alteration in the fingers to hide the identity.

For further analysis and classification we would like to extract more features, use key points matching for a specific application design. Even more dataset can be considered as CNN needs a high volume of data to learn from.

### **Acknowledgement**

First and foremost I want to express my gratitude towards almighty ALLAH for enabling me to complete my project "Fingerprint Identification and Analysis using Deep Learning". I would like to express my gratitude towards my parents and also like to convey my sincere gratitude to our honorable supervisor Sir Professor M. Jahirul Islam. Without Sir's kind direction and proper guidance this study would have been a little success. In every phase of this project sir's supervision and guidance has shaped the study. Finally, I would like to thank my honorable teachers and friends who helped me in completing the project.

### **IX. References**

1. Yahaya Shehu; Ariel Ruiz-Garcia; Vasile Palade; Anne E. James " Detailed Identification of Fingerprints using Convolutional Neural Networks". 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA).
2. Yahaya Shehu; Ariel Ruiz-Garcia; Vasile Palade; Anne E. James " Detection of Fingerprints Alterations Using Deep Convolutional Neural Networks". 27th International Conference on Artificial Neural Networks, ICANN 2018 - Rhodes, Greece
3. Oliver Giudice; Mattia Litrico; Sebastiano Battiato "SINGLE ARCHITECTURE AND MULTIPLE TASK DEEP NEURAL

NETWORK FOR ALTERED FINGERPRINT ANALYSIS”.  
2020 IEEE International Conference on Image Processing  
(ICIP).

4. Yousif Ahmed Al-Wajih; Waleed Hamanah; Mohammad Abido; Fouad Al-Sunni “Finger Type Classification with Deep Convolution Neural Networks “Proceedings of the 19th International Conference on Informatics in Control, Automation and Robotics 2022At: Lisbon-Portugal.
5. <https://pureportal.coventry.ac.uk/en/publications/sokoto-coventry-fingerprint-dataset>
6. <https://pureportal.coventry.ac.uk/en/publications/detection-of-fingerprint-alterations-using-deep-convolutional-neu/fingerprints/>
7. <https://www.kaggle.com/datasets/ruizgara/socofing>
8. <http://scikit-learn.org>
9. [http://powerunit-ju.com/wp-content/uploads/2021/04/Aurelien-Geron-Hands-On-Machine-Learning-with-Scikit-Learn-Keras-and-Tensorflow\\_-Concepts-Tools-and-Techniques-to-Build-Intelligent-Systems-OReilly-Media-2019.pdf](http://powerunit-ju.com/wp-content/uploads/2021/04/Aurelien-Geron-Hands-On-Machine-Learning-with-Scikit-Learn-Keras-and-Tensorflow_-Concepts-Tools-and-Techniques-to-Build-Intelligent-Systems-OReilly-Media-2019.pdf)

# Develop Recommendation System for Restaurant Recommendation on Bangladesh Perspective Using Collaborative Filtering Approach

Dipongkar Ghose<sup>1</sup>, Emdad Hussain Imu<sup>2</sup> and  
Shafkat Kibria<sup>3</sup>

*Computer Science and Engineering, Leading University, Bangladesh<sup>1,2,3</sup>*

## **Abstract-**

In machine learning, the recommender system is one of the most important applications. It guesses what songs you might like to listen to, what food you order online, what posts you may like to see on your social networks, as well as the person you may want to connect with, what series or movies you like to watch, etc. In the recommender system the content-based recommender is one of the most popular techniques. A content-based recommender system tries to assume the reaction of a user given the item's features, to which he/she reacts positively. But these systems require a well-defined feature matrix. But the feature matrix is not readily available for large data sets. To solve this problem, we use Collaborative filtering. It can generate and learn features by itself. This paper is based on the study of the popular Coursera course "Machine Learning" by Professor Andrew Ng from Stanford

University. We use a low matrix factorization approach. The prediction algorithm is tested on a real user dataset containing user ratings on some restaurants and on a Synthetic data set to verify the prediction accuracy. The result of our system was analyzed using four error finding criteria.

***Index Terms*** Machine Learning, Collaborative Filtering, Synthetic dataset, Gradient decent.

## **I. Introduction**

Recommender systems are so common that now many of us use them without even knowing it. Because users can't possibly look at all the products or content on a website, for better user experience a recommendation system plays an important role, while also exposing us to more inventory we might not discover. Some recommender systems examples are product recommendations on Amazon, Movies and TV shows suggestions on Netflix in your feed, YouTube video recommendation, music on Spotify, Facebook newsfeed, and Google Ads [4] [5]. The recommender function is the most important component of any of these systems, which takes the user information and predicts the rating that the user might allocate to a product, for example predicting user ratings even before the user has provided one, makes recommender systems a powerful tool.

Three algorithms are used in recommender systems:

- Content-based system [6], which use characteristic an item.
- Collaborative filtering systems [7], which are based on user rating on an item.
- Hybrid systems [8], which combine both types of system to avoid the problems that are generated when working with one kind.

In content-based filtering item features are used to recommend other items based on user previous actions or explicit feedback. In [9], B. R. Cami, H. Hassanpour, H. Mashyekhi proposed a system using content-based filtering to recommend movies to users. It takes the user's temporal preferences in user modeling and predicts the movies. Collaborative-filtering systems depend on the relationship between users and items. The similarity of items is measured by the similarity of the ratings given by the users who have rated both items [10]. Collaborative filtering is the way of filtering items using

the opinions of other people. This technology brings the opinions of large interconnected communities on the web together and supports the filtering of substantial quantities of data [11]. Researchers had accomplished several kinds of collaborative recommendation system techniques across the world [12], [13], [14].

### **Two types of Collaborative-filtering systems are there:**

**Model-based systems:** In this system first build a model of user ratings then provides item recommendations. To perform these predictions different algorithms are used, suppose neural network, logistic regression, SVMs, or Bayesian network. matrix factorization methods [15] one of the most famous methods for model-based recommender systems.

**Memory-based systems:** In this system, to make recommendations it computes the similarity between users and items using user rating data. An example of this approach is neighborhood-based Collaborative-filtering [7] and item-based/user-based top-N recommendations [15].

Nowadays in the recommendation system hybrid approach is the most popular approach. Hybrid recommender systems combine two or more recommendation strategies. An example of a hybrid recommendation system is first making content-based and collaborative-based predictions separately and then joining them. In [16] C. Nundlall, G. Sohun, S. D. Nagowah introduced a hybrid recommendation technique for Big Data Systems. They combine collaborative and content-based filtering techniques to recommend items that a user would likely prefer and also use item ranking and classification techniques for recommending the items. In [17] To Thi Thuan, Sutheera Puntheeranurak they suggest a hybrid recommender system which we have used to construct the hybrid model. They compared three recommendation techniques.

In this paper, we follow a low-rank matrix factorization collaborative filtering approach. In our work we used three datasets, first one is a popular Movie Lens dataset that was collected from Movie lens. The second dataset was collected from real users. The final dataset was a synthetic dataset.

The outline of this article is as follows. In Section 2, we discuss the algorithm we build. The most important part, the presentation of the experimental results illustrated in Section 3.

## II. Recommendation system

### A. Problem Formulation

In this study, we aim to build a recommendation system that predicts user ratings for restaurants they have not yet visited based on their preferences for other restaurants. The system is particularly useful in real-world scenarios where users rate only a subset of available restaurants, leaving many unrated. By predicting these missing ratings, we can provide personalized recommendations to users.

Consider a simplified example where users rate restaurants on a scale from 1 to 5 stars. In this scenario, we have 5 restaurants, categorized into two types: Deshi and Chinese. The restaurants are labeled as follows:

- Deshi (1), Deshi (2), Deshi (3)
- Chinese (1), Chinese (2)

We also have 4 users, identified as User 1, User 2, User 3, and User 4. Each user rates different restaurants based on their preferences, but not all users rate all restaurants.

For instance:

- User 1 loves Deshi (1) and rates it 5 stars. They also enjoy Deshi (2) and rate it 4 stars. However, User 1 has not visited Deshi (3), so we don't have a rating for it. They dislike both Chinese restaurants, giving Chinese (1) and Chinese (2) a low rating of 1 star.
- User 2 also rates Deshi (1) highly, giving it 5 stars, but skips Deshi (2). User 2 rates Deshi (3) a 4, and they don't like the Chinese restaurants either, rating both Chinese (1) and Chinese (2) 1 star each.

This pattern continues with the remaining users, leading to a sparse rating matrix where some restaurants have no ratings from certain users. Table 1 below illustrates this:

Table 1: Example of a restaurant rating by user

Restaurant	User 1	User 2	User 3	User 4
Deshi 1	5	5	1	1
Deshi 2	4	?	?	1
Deshi 3	?	4	1	?
Chinese 1	1	1	5	4
Chinese 2	1	1	5	?

Here, the symbol '?' denotes the missing ratings. Our goal is to predict these missing values so we can offer personalized restaurant recommendations to users based on their historical preferences and the preferences of similar users.

Notation and Objective In our model:

- $r(i,j)$  = 1 if user  $j$  has rated restaurant  $i$  (otherwise  $r(i,j)=0$  if the rating is missing).
- $y(i,j)$  represents the actual rating given by user  $j$  to restaurant  $i$  if it exists.

The task of a recommender system is to predict the missing ratings—represented by '?' in the table—using the information available from the ratings that users have already provided. In a realistic setting, some users may have rated only a small number of restaurants, making it even more important to leverage similarities between users and restaurants to infer the missing values.

For instance, if both User 1 and User 2 like Deshi 1 and give it high ratings (5 stars), we might infer that User 3, who has given it a lower rating, may similarly rate Deshi 2 or Deshi 3 low, based on their overall preferences. Similarly, if User 4 really likes Chinese 1 and gave it a 4-star rating, they might be expected to give a similar rating to Chinese 2.

$(j)$  = This is the parameter vector for user  $j$ . It represents the user's preferences in terms of the different features of the restaurants (such as cuisine type, price range, etc.).

$x(i)$  = This is the feature vector for restaurant  $i$ . It contains information about the restaurant, such as its cuisine type (Deshi or Chinese), price category, or popularity.  $m(j)$  = This represents the number of restaurants rated by user  $j$ .

### **B. Mean Normalization**

Suppose there's a user who has not rated any restaurants yet. So, it is tough to recommend any restaurants to this user. This is called the cold start problem. In this case normalization will let us fix this problem. First, take all restaurant ratings into one matrix. Then compute the average rating that each restaurant obtained. Finally, subtract these mean ratings from restaurant ratings.

### **C. Collaborative filtering optimization objective**

The collaborative filtering algorithm predict a rating for the item  $i$  by user  $j$  by using

$y(i, j) = (\theta(j)) x(i)$  function where n-dimensional parameter vectors  $x(1) \dots x(n_m)$  and  $\theta(1) \dots \theta(n_u)$  are known. We have a rating dataset that is produced by some users on different items. From this our model tries to learn parameter vectors  $x(1) \dots x(n_m)$  and  $\theta(1) \dots \theta(n_u)$ , these  $X$  and  $\theta$  are used to predict the rating by user  $j$  on item  $i$ .

We can predict  $\theta(1) \dots \theta(n_u)$  if feature vectors  $x(1) \dots x(n_m)$  is given.

$$\min_{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n_u)}} \frac{1}{2} \sum_{j=1}^{n_u} \sum_{i:r(i,j)=1} ((\theta^{(j)})^T(x^{(i)}) - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2$$

Here  $\frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2$  is known as a regularization term.

We can predict feature vectors  $x(1) \dots x(n_m)$  if  $\theta(1) \dots \theta(n_u)$  is given

$$\min_{x^{(1)}, x^{(2)}, \dots, x^{(n_m)}} \frac{1}{2} \sum_{i=1}^{n_m} \sum_{j:r(i,j)=1} ((\theta^{(j)})^T(x^{(i)}) - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^n (x_k^{(i)})^2$$

Now we combine both above objectives into one objective. This objective reduces the cost function which contains parameter  $\theta$  and features  $X$ .

$$J(x^{(1)}, x^{(2)}, \dots, x^{(n_m)}, \theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n_u)})$$

$$\frac{1}{2} \sum_{i:r(i,j)=1} ((\theta^{(j)})^T(x^{(i)}) - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2 + \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^n (x_k^{(i)})^2$$

#### D. Minimize Cost Function using gradient descent

We found a cost function but now we have to minimize it. For this, we used gradient descent which derivative the cost function two times one is in respect of Feature  $x$  and another in respect of  $\theta$ . At first, we take random values for feature and theta matrix. This matrix will update simultaneously after each iteration.

To minimize  $J(\theta^{(1)} \dots \theta^{(nm)}, \theta^{(1)} \dots \theta^{(nu)})$  we use a gradient descent algorithm for every  $j=1 \dots nu$ .  $i = 1 \dots nm$

$$x_k^{(j)} = x_k^{(j)} - \alpha \left( \sum_{j:r(i,j)=1} ((\theta^{(j)})^T (x^{(i)} - y^{(i,j)})) \theta_k^{(j)} + \lambda x_k^{(i)} \right)$$

$$\theta_k^{(j)} = \theta_k^{(j)} - \alpha \left( \sum_{i:r(i,j)=1} ((\theta^{(j)})^T (x^{(i)} - y^{(i,j)})) x_k^{(i)} + \lambda \theta_k^{(j)} \right)$$

### E. Finding Similarity

After gradient descent minimizing, we found a new optimized feature matrix  $X$  and  $\theta$ . To predict the rating of the movie  $i$  for user  $j$ , we compute  $y = ((j)) (i)$ . To find the top recommendation we add  $y(\text{mean})$  with  $y$ . They're also another way to find the related item. To find similarities between Item  $i$  and item  $j$  we need to find differences between feature vectors  $x(i)$  and  $x(j)$ . If their difference is small then we can say those items are similar.

### F. Experiment Dataset

**Movie Lens dataset.** The proposed algorithm for recommendations was also analyzed using three different datasets. These are Movie Lens data containing **100,000** ratings of approximately 9724 movies rated by 610 Movie Lens users [2]. This dataset was used to verify the performance of the proposed algorithm the movie lens data was used first. Using this data model gives predictions which are quite efficient. Along-with the error rate also measures by taking the average difference between actual user rating and predicted rating by the algorithm.

**Restaurant dataset.** Another dataset was collected from real users. 48 users were asked to rate 28 restaurants from the data set. According to their rating, model suggest top restaurants to them.

**Synthetic dataset.** In Restaurant dataset there is only 48 users give rating. This data is not sufficient for train a model. That why an over-sampling approach by which extra training data was created by performing certain operations on this restaurant data. This approach is inspired by a technique that proved successful in handwritten character recognition [3]. To build this dataset a small dataset was used that has only 48 users rating on 28 different restaurants. In this

new dataset, there are 1200 users' ratings on 28 different restaurants. This dataset became very useful to train our system.

The datasets are divided into a training set, a test set. We basically try to create a model to predict the test data. So, the training data is used to fit the model and testing data is used to test it. Here we split the data in a 70:30 ratio i.e., for training the model 70% of the data will be used on the other hand 30% will be used for testing the model.

### III. Evaluation

Here discuss elaborately about testing experience and the results.

#### A. Restaurant dataset

For the restaurant dataset, we gathered ratings from 48 users across 28 different restaurants, which we categorized into five types: Deshi, Indian, Chinese, Thai, and Buffet House. We divided the dataset into training and testing sets. The training set comprised ratings for 28 restaurants by 34 users, while the testing set included ratings from 14 users for various restaurants.

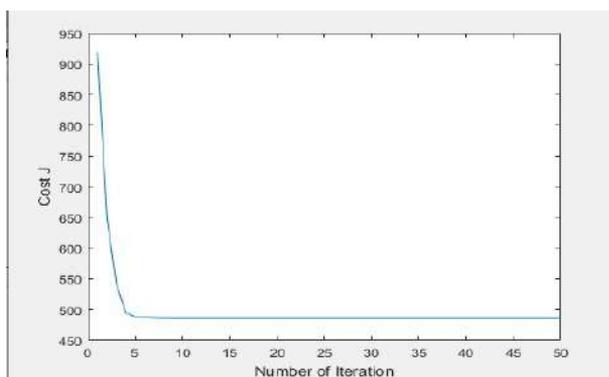


Figure 1: Cost function iteration graph

Training rating matrix used to train the model. In this model, we used “fmincg” optimized function for gradient descent. 100 iteration were made inside this function to find optimized feature matrix  $X$  and  $\theta$  matrix. Figure: 1 is a Cost function iteration graph. After this iteration it gives optimized feature  $X$  and  $\theta$  matrix to prediction user rating.

To train the model, we employed the training rating matrix and utilized the "fmincg" optimized function for gradient descent, performing 100 iterations to refine the feature matrix  $X$  and the parameter matrix  $\theta$ .

The resulting cost function iteration graph, depicted in Figure 1, illustrates the convergence of the optimization process.

After tuning the hyperparameters, we achieved the following error rates in our model, as summarized in Table 2. These error rates demonstrate the model's predictive capabilities, showing a marked improvement in performance on the test set compared to the training set.

Table 2: Restaurant data Performance error summary

Error calculating method	Test	Train
MAE	0.7985	0.9644
MSE	0.8819	1.3974
RMSE	0.9247	1.1745
RMSLE	0.1998	0.3053

To generate restaurant recommendations for users, we sorted the predicted ratings and selected the top-rated restaurants for each user. Figure 2 provides a visualization of the top restaurant recommendations based on predicted ratings.

```

Top recommendations for you:
Predicting rating 4.2 for Restaurant KFC
Predicting rating 4.2 for Restaurant The Mad Grill Restaurant
Predicting rating 4.1 for Restaurant Royal Dine Restaurant
Predicting rating 4.1 for Restaurant Buffet House and Restaurant
Predicting rating 4.0 for Restaurant MEATUP
Predicting rating 4.0 for Restaurant Pizza Hut
Predicting rating 3.9 for Restaurant Wondal Restaurant
Predicting rating 3.8 for Restaurant Royal Kitchen
Predicting rating 3.8 for Restaurant Sip Coffee Restaurant
Predicting rating 3.8 for Restaurant Spicy Restaurant & Party Center

```

Figure 2: Top restaurant recommendation

## B. Synthetic Dataset

We also assessed the performance of our model using a synthetic dataset. Similar to the restaurant dataset, we divided the synthetic dataset into training and testing sets. The training set consisted of ratings for 28 restaurants by 840 users, while the testing set included ratings from 360 users, also categorized into the same five restaurant types.

Following the same evaluation methodology as before, we again used the "fmincg" optimized function to refine the model's parameters. The cost function iteration graph for this dataset is presented in Figure 3.

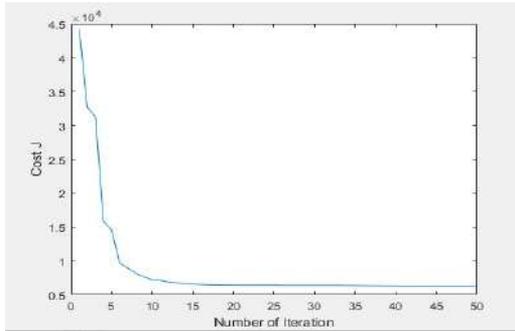


Figure 3: Cost function iteration graph

Table 3 Restaurant data Performance error summary

Error calculating method	Test	Train
MAE	0.8605	0.9346
MSE	1.2487	1.5874
RMSE	1.1080	1.2397
RMSLE	0.3056	0.3470

The error rates shown in Table 3 indicate that there is potential for improvement in prediction accuracy. The differences in error rates can be attributed to the limited size and diversity of the original restaurant data used to create the synthetic dataset. If larger real-world data were utilized to construct the synthetic dataset, we could expect the training and testing error rates to converge more closely.

```

Top recommendations for you:
Predicting rating 4.7 for Restaurant KFC
Predicting rating 4.7 for Restaurant Kababia Restaurant
Predicting rating 4.5 for Restaurant Pizza Hut
Predicting rating 4.3 for Restaurant Food Lounge
Predicting rating 4.3 for Restaurant Rainbow Chinese Restaurant
Predicting rating 4.3 for Restaurant Buffet House and Restaurant
Predicting rating 4.2 for Restaurant Wondal Restaurant
Predicting rating 4.2 for Restaurant The Mad Grill Restaurant
Predicting rating 4.2 for Restaurant Platinum Lounge
Predicting rating 4.2 for Restaurant Royal Dine Restaurant

```

Figure 4: Top movies recommendation

To provide restaurant recommendations, we again sorted the predicted ratings and highlighted the top choices for users, as illustrated in Figure 4.

#### IV. Conclusion

In this paper, we developed a collaborative filtering-based recommendation system specifically for the restaurant landscape in Sylhet, Bangladesh. Our study created a benchmark dataset of real user ratings for well-known local restaurants and provided a methodology for constructing datasets from limited sources. This publicly available dataset addresses a gap in restaurant recommendation literature in Bangladesh. We demonstrated the feasibility of generating reliable recommendations with low-rank matrix factorization, even when working with small datasets. Our findings highlight the potential of collaborative filtering methods in various applications and emphasize the importance of data availability in enhancing recommendation system performance. Future research could enhance our approach by integrating larger datasets and exploring hybrid methods that combine collaborative filtering with content-based techniques, improving the accuracy and usability of restaurant recommendations in Bangladesh.

#### V. References

1. Andrew Ng., "Machine learning," <https://www.coursera.org/learn/machine-learning/>
2. F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," ACM Transactions on Interactive Intelligent Systems (TiiS), vol. 5, no. 4, p. 19, 2016.
3. Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, W. Philip Kegelmeyer." SMOTE: synthetic minority over-sampling technique", ©2002 AI Access Foundation and Morgan Kaufmann Publishers, published 06/02, 321—357.
4. Burke, Robin, Alexander Felfernig, and Mehmet H. GÖker. "Recommender systems: An overview." AI Magazine 32.3 (2011): 13-18.
5. Andre Marchand, Thorsten Hennig-Thurau, Val ue Creation in the Video Game Industry: Industry Economics, Consumer Benefits,

- and Research Opportunities, *Journal of Interactive Marketing*, Volume 27, Issue 3, August 2013, Pages 141-157
6. R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in *Proceedings of the fifth ACM conference on Digital libraries*. ACM, 2000, pp. 195–204.
  7. I. Markovsky, *Low rank approximation: algorithms, implementation, applications*. Springer Science & Business Media, 2011.
  8. T. K. Paradarami, N. D. Bastian, and J. L. Wightman, "A hybrid recommender system using artificial neural networks," *Expert Systems with Applications*, vol. 83, pp. 300–313, Oct. 2017.
  9. Bagher Rahimpour Cami, Hamid Hassanpour, Hoda Mashayekhi 'A content-based movie recommender system based on temporal user preferences' 2017 3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS), IEEE, 20-21 Dec. 2017.
  10. Resnick, Paul, and Hai R. Varian. "Recommender systems." *Communications of the ACM* 40.3 (1997): 56-58.
  11. J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen, "Collaborative Filtering Recommender Systems," *Communications of the ACM*, vol. 40, no. 3, pp. Pages 291–324, 2007.
  12. Najdt Mustafa; Ashraf Osman Ibrahim; Ali Ahmed; Afnizanfaizal Abdullah, "Collaborative filtering: Techniques and applications" 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), IEEE, 16-18 Jan. 2017.
  13. Prateek Parhi; Ashish Pal; Manuj Aggarwal, "A survey of methods of collaborative filtering techniques," 2017 International Conference on Inventive Systems and Control (ICISC), IEEE, 19-20 Jan. 2017.
  14. Xiaoyuan Su ,Taghi M. Khoshgoftaar "A survey of collaborative filtering techniques", *Communications of the ACM*, Volume 2009, Issue January 2009,1-2 pages.
  15. T. Li, A. Liu, and C. Huang, "A Similarity Scenario-Based Recommendation Model with Small Disturbances for Unknown Items in Social Networks," *IEEE Access*, vol. 4, pp. 9251–9272, 2016.
  16. C. Nundlall, G. Sohun, S. D. Nagowah, "A Hybrid Recommendation Technique for Big Data Systems," 2018

International Conference on Intelligent and Innovative Computing Applications (ICONIC), IEEE, 6-7 Dec. 2018.

17. C. Nundlall, G. Sohun, S. D. Nagowah, "Hybrid recommender system with review helpfulness features" TENCON 2014 - 2014 IEEE Region 10 Conference , 22-25 Oct. 2014.
18. Çano, Erion and Morisio, Maurizio. 'Hybrid Recommender Systems: A Systematic Literature Review'. 1 Jan. 2017: 1487 – 1524. DOI:10.3233/IDA-163209

# Optimizing Minimum Cost Routes: A Comparative Study of Algorithms of Dijkstra's, Edge Deletion Method, and Vertex Assignment Enumeration Techniques

Saiful Islam<sup>1</sup>, Dr. Anupam Dutta<sup>2</sup>

<sup>1</sup>*Research Scholar*

<sup>2</sup> *Associate Professor, Department of Mathematics, USTM*

## **Abstract:**

The optimization of minimum-cost routes in graph-based networks is a problem in transportation, telecommunications, and logistics. This study presents a comparative analysis of three algorithms for solving the shortest path problem: Dijkstra's Algorithm, the Edge Deletion algorithm, and the Vertex Assignment Enumeration Technique Algorithm.

Dijkstra's Algorithm, renowned for its efficiency in non-negative weighted graphs, iteratively explores the shortest paths from the source vertex to all other vertices. Its polynomial time complexity and space complexity make it suitable for large-scale networks.

The Edge Deletion Algorithm Method is an innovative approach that involves systematically removing the edges from the graph and recalculating paths to minimize the overall cost. This

method, although potentially computationally intensive, offers a unique perspective on route optimization by considering dynamic changes in the network topology.

Lastly, the Vertex Assignment Enumeration Algorithm Technique focuses on assigning and enumerating all vertices in a specific way to determine the shortest paths, and is analyzed for its application in specific scenarios where vertex properties play a crucial role.

The study includes a detailed analysis of the time and space complexity of each algorithm, along with their suitability for different types of graphs, such as dense, sparse, directed, and undirected graphs. Experimental results are provided to illustrate the performance of these algorithms in practical scenarios, highlighting their strengths and weaknesses. The comparative study aims to provide insights into the best use cases for each algorithm and guide the selection of the appropriate algorithm based on the specific requirements of the problem at hand.

**2020 Mathematics Subject Classification:** 68W01, 68W20, 68W99

**Keywords and phrases:** Dijkstra's Algorithm, Minimum Cost Route, Vertex Assignment Enumeration Algorithm, minimum dominating sets, Edge Deletion Algorithm, Time complexity, Space complexity, Algorithm Comparison, DFS,

## 1. Introduction:

Here we have discussed that finding the minimum cost route in a graph is a core problem in computer science, with applications in network routing, geographical mapping, and logistics. This paper compares three different approaches to solve this problem: Dijkstra's algorithm, the Vertex Enumeration algorithm, and the Edge Deletion algorithm. Each algorithm has its unique characteristics and is suited to specific types of graphs and problem settings.

Dijkstra's algorithm, developed by E.W. Dijkstra, is a method for determining the shortest path from a given starting node in a graph to a target node. Additionally, it can be utilized to find the shortest paths from the starting node to all other nodes in the graph

at the same time, addressing what is known as the single-source shortest paths problem.

## 2. Background of the Study:

### 2.1. Dijkstra Algorithm:

Dijkstra's algorithm is a widely used technique for determining the shortest paths from a single starting vertex to all other vertices in a graph with non-negative edge weights. It utilizes a priority queue to repeatedly select the vertex with the smallest distance and then updates the distances of its neighboring vertices accordingly.

To determine the shortest path between two vertices, such as  $A$  to  $Z$ , in a weighted graph, an algorithm iteratively assigns numerical labels to the vertices. Each vertex initially has a temporary label, which will eventually be replaced by a permanent label. The label of a vertex  $v$  is denoted as  $L(v)$ .

#### Initial Iteration 0

- 1 Define  $V_0$  as the set of all vertices in the graph.
- 2 Assign the starting vertex  $v_0$  a permanent label of 0, while all other vertices receive temporary labels of  $\infty$ .
- 3 Let  $V_1 = V_0 - \{v_0\}$ , where  $v_0$  is the starting vertex with a permanent label.

#### Iteration 1

- 1 Redefine the elements of  $V_1$  as  $v_1$ , which includes all vertices in  $V_0$  except  $v_0$ .
- 2 For vertices in  $V_1$  adjacent to  $v_0$ , update their temporary labels using  $L(v_1) = L(v_0) + w(v_0, v_1)$ , where  $L(v_0) = 0$  and  $w(v_0, v_1)$  is the weight of the edge between  $v_0$  and  $v_1$ .
- 3 For other vertices in  $V_1$ , retain their previous temporary labels.
- 4 Identify  $v_1^*$ , the vertex in  $V_1$  with the minimum label  $L(v_1)$ . If there is a tie, break it arbitrarily. Assign  $L(v_1^*)$  a permanent label.
- 5 Define  $V_2 = V_1 - \{v_1^*\}$ .

#### Iteration i

- 1 For vertices in  $V_i$  adjacent to  $v_{i-1}^*$ , update their temporary labels using  $L(v_i) = L(v_{i-1}^*) + w(v_{i-1}^*, v_i)$
- 2 For other vertices in  $V_i$ , retain their previous temporary labels.

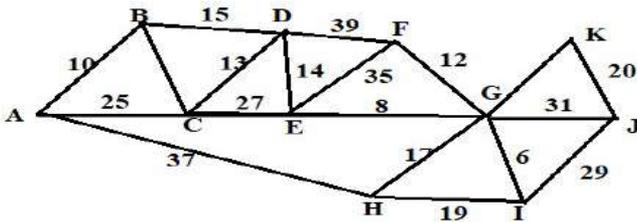
- 3 If a temporary label in the  $i$ -th iteration is greater than or equal to its previous iteration label, do not change it.

**Termination**

The algorithm stops when the final vertex  $z$  receives a permanent label. The permanent label of  $z$  represents the shortest path length from  $A$  to  $Z$ . The shortest path is traced backward from  $Z$ , including the vertices that contributed to the permanent labels.

**2.2. Experimental proof of Dijkstra Algorithm:**

From the Dijkstra algorithm, we have the minimum cost route from the arbitrary graph



**Graph -1**

Consider the following graph which is a weighted connected simple graph, with all its weights positive. The lengths of weights are

- AB = 10 , AC = 25, AH = 37, BD = 15, BC = 7, CD = 13, CE = 27,  
 DF = 39, DE = 14, EF = 35, EG = 8, FG = 12, GK = 24,  
 GJ = 31, GI = 6, GH = 17, KJ = 20, HI = 19, IJ = 29

From the aforesaid algorithm, the minimum cost route from the vertex  $A$  to the vertex  $K$  by using the Dijkstra,s Algorithm is  $L(v_A v_K) = 71 wt$

**3. Innovative Algorithms for Minimum Cost Route Optimization: Edge Deletion Algorithm and Vertex Assignment Enumeration Techniques Algorithm**

**3.1. Edge deletion Algorithm:**

**Input:** A weighted, undirected graph  $G = (V, E)$  with weights  $W: V \rightarrow \mathbb{R}$ .

**Output:** Find out the minimum cost route or the shortest path.

**Step 1:** Enter the all weights of the edges from the arbitrary graph.

- Step 2:** Arranged the all weights of the edges either in ascending or descending orders.
- Step 3:** Select the  $\frac{1}{4}$ th greatest weights of the edges from step 2 in such a way that they should not be incident in a single vertex and at least one Hamiltonian circuit will have to be maintained in adjacency.
- Step 4:** Delete the  $\frac{1}{4}$ th greatest weights edges from the step 3.
- Step 5:** Select the source ( $v_a$ ) and sink ( $v_k$ ) vertices, where one journey has to complete from  $v_a$  to  $v_k$
- Step 7:** From the source vertex  $v_a$  observe all the possible adjacency paths to all other adjacent vertices.
- Step 8:** From the source ( $v_a$ ) to the second vertex ( $v_b$ ) select only the possible paths adjacent to  $v_b$  and calculate  $\{\min \text{ length of } (v_{ab})\}$ . Consider the path of  $\{\min \text{ length of } (v_{ab})\}$ , through the other vertex if the vertex selected before.
- Step 9:** Now from the second vertex to the third vertex select only the possible paths adjacent to  $v_c$  and continuing the same as step 8.
- Step 10:** Continuing the process up to the end vertex  $v_k$
- Step 11:** Stop.

### 3.2. Experimental proof on edge deletion Algorithm:

From the Edge deletion algorithm, we have to find the minimum cost route of the arbitrary graph-1.

Now let the arrange the weights of the edges in descending order or ascending order of the total 19 weighted edges as given below

Select  $1/4^{\text{th}}$  greatest weighted edges from the total edges and delete these edges from the graph.

$$\begin{aligned} DF = 39, AH = 37, EF = 35, GJ = 31, IJ = 29, CE = 27, AC \\ = 25, GK = 24, KJ = 20, \\ HI = 19, GH = 17, BD = 15, DE = 14, CD \\ = 13, FG = 12, AB = 10, EG = 8, BC = 7, GI \\ = 6, \end{aligned}$$

From the Source Vertex A and list the all possible paths

$$\begin{aligned} = Dis \{v_{AB}\} = 10, Dis \{v_{AC}\} = 25, Dis \{v_{AH}\} = 37 \Rightarrow \\ Min\{10, 25, 37\} = 10 = v_{AB} \end{aligned}$$

Select the second visited vertex from the source vertex by selecting the minimum cost =  $Dis\{v_{AB}\} = 10$

From the second visited vertex B and list the all possible paths =  $Dis\{v_{BC}\} = 17, Dis\{v_{BD}\} = 25 \Rightarrow Min\{17, 25\} = 17 = v_{BC}$

Select the third visited vertex from the second vertex by selecting minimum cost only =  $Dis\{v_{BC}\} = 17$

The minimum possible path from the third visited vertex C to the next vertex, list all possible paths and continues the process as mentioned in the algorithm and we have found from algorithm-2, the least cost route is  $Min(v_A v_K) = 71wt.$

### 3.3. Algorithm of Vertex Assignment Enumeration Technique:

**Input:** A weighted, undirected graph  $G = (V, E)$  with weights  $W: V \rightarrow \mathbb{R}$ .

**Output:** Find out the minimum cost route or the shortest path

#### Step-1: Input Initialization:

- a. Initialize the connected graph G with vertices V and edges E.
- b. Assign weights to the vertices and edges.
- c. Define the source and sink vertices from the weighted graph

#### Step-2: Helper Functions:

- d. **Minimum Dominating Vertex (MDV):** Identify the minimum dominating set of vertices from the graph.
- e. **Greatest Weighted Vertex (GWV):** Calculate the weights of each vertex and arrange the vertices with the greatest weight.
- f. **Maximum Connectivity Vertex (MCV):** Identify the vertices connected by edges with the maximum number of vertex.

#### Step-3: Data Structures:

- g. Select the common vertices from the step-2(d, e, f)
- h. List the paths passing through the common vertices of step- 3(g).
- i. List paths to store potential paths from the source to the sink using the step 3(g, h) vertices.

#### Step-4: Depth-First Search (DFS) Algorithm:

Implement a DFS algorithm that iterates over the graph while considering vertices passing through the selected

vertices in step 3(i) and guiding the search towards the least cost path.

**Step-5: Path Cost Calculation:**

Calculate the cost of each path by summing the weights of the vertices and edges in the path.

**Step-6: Output the Least Cost Path:**

Compare the costs of all paths found and output the path with the least cost.

**Step-7: Stop.**

**Summery Explanation:**

- A. **Initialization:** We initialize the graph and necessary variables.
- B. **Helper Functions:** These functions help in identifying the minimum dominating set, greatest weighted vertices, and maximum connectivity vertices from the graph.
- C. **DFS Algorithm:** We use a DFS approach to explore all possible paths from the source vertex to the sink vertex, prioritizing vertices based on passing through the helper functions.
- D. **Path Cost Calculation:** For each path found, we calculate the total cost.
- E. **Result:** The path with the least cost is identified.

This algorithm ensures that we explore the paths in a guided manner through the use of minimum dominating vertices, the greatest weighted vertices, and the maximum connectivity vertices gives ultimately the least cost path.

**3.4. Experimental proof of Vertex Assignment Enumeration Technique:**

From the vertex Assignment enumeration algorithm, we have to find the minimum cost route from the arbitrary graph-1.

From the step 1, Select the source A and sink vertices K

From the step 2, Now let the calculate the weights of every vertices as follows

$$\sum V_A = 72, \sum V_B = 32, \sum V_C = 72, \sum V_D = 81, \sum V_E = 84, \sum V_F = 86,$$

$$\sum V_G = 98, \sum V_H = 73, \sum V_I = 54, \sum V_J = 80, \sum V_K = 45,$$

From Step 3, 4, 5 and 6 gives that

$$GWV' \text{ are } = \{ \sum V_G = 98, \sum V_F = 86, \sum V_E = 84, \sum V_D = 81, \sum V_J = 80, \sum V_H = 73, \sum V_C = 72 \}$$

$$MCV's \text{ are } = \{ \{ V_C, V_D, V_E, V_G \} \}$$

MDV's are =  $\{\{V_C, V_G\}$

The traveller must visit the specific assigning vertices of MDV, GWV, and MCV in their journey and calculate a path through the source vertex to the sink vertex to complete the paths.

Thus from the Depth-First Search (DFS), step-4 Algorithm, the least cost route found from the arbitrary graph-1 is  $Min(v_A v_K) = 71 \text{ wt.}$

#### 4. Comparative study of Time complexity:

##### 4.1. Dijkstra's Algorithm:

Dijkstra's algorithm is a classic algorithm for finding the shortest paths from a single source vertex to all other vertices in a graph with non-negative edge weights.

**Time Complexity:** Consider V is the number of vertices and E is the number of edges then

- Using a simple array :  $O(V^2)$
- Using a binary heap :  $O(V + E \log V)$
- Using a Fibonacci heap:  $O(V \log V + E)$

##### 4.2. Edge Deletion Algorithm:

An edge deletion algorithm for finding the minimum cost route generally involves iteratively removing the one-fourth number of total edges from the graph and checking for the shortest paths by the algorithm given in experiment 2,

##### **Time Complexity:**

- In the maximum weighted edge deletion process, the time complexity is  $O\left(V + \frac{3}{4}E\right)$  for a dense graph and  $O\left(\frac{3}{4}E^2\right)$  for a sparse graph.
- This complexity can vary based on the specific implementation and data structures used for priority queues.

The deletion of one-fourth of the larger weighted edges from a connected graph, the impacts of the time complexity for the shortest path algorithms are showing more efficient and effective. Specifically, our new approach **Edge deletion algorithm** effects of time complexity and the time complexity obtained from the Dijkstra's and Bellman-Ford algorithms and the Floyd-Warshall algorithms remain unchanged.

### 4.3. Vertex Assignment Enumeration Algorithm

Vertex Assignment enumeration Technique gives the contexts like the Travelling Salesman Problem (TSP), where paths between vertices are explicitly considered.

#### Time Complexity:

- Calculating the degree of each vertex takes  $O(V + E)$  time because we need to look at each vertex and each edge.
- Finding the maximum degree of each vertices from these degrees takes  $O(V)$  time.
- Enumerate the vertex with the highest degree:  $O(V + E)$
- Delete one edge from each vertex where the **maximum connectivity, maximum weighted, and minimum domination vertex** found:  $O(V + E)$
- Total time complexity:  $O(V + E)$

#### Summary of Time Complexities:

Algorithm	Time Complexity
Dijkstra's (simple array)	$O(V^2)$
Dijkstra's (binary heap)	$O(V + E) \log V$
Dijkstra's (Fibonacci heap)	$O((V \log V + E))$
Greatest weighted Edge Deletion	$O(V + \frac{3}{4}E)$
Vertex Assignment Enumeration	$O(V + E)$

These time complexities highlight that Dijkstra's algorithm is obviously a quadratic formation, and also more efficient for finding the shortest paths in graphs with non-negative edge weights. The edge deletion algorithm and vertex Assignment enumeration algorithm also linear time complexity and efficiently and efficiently but it depends on the method used for shortest path calculations, while the vertex Assignment enumeration is computationally expensive for large and sparse graphs.

### 5. Comparative study of Space complexity:

#### 5.1. Dijkstra's algorithm:

The space complexity of Dijkstra's algorithm depends on the data structures used to represent the graph and the priority queue. Let's break down the space requirements:

#### A. Graph Representation:

- **Adjacency List:** If the graph has  $V$  vertices and  $E$  edges, the adjacency list representation requires  $O(V + E)$  space.

Each vertex stores a list of its adjacent vertices along with the edge weights.

- **Adjacency Matrix:** This requires  $O(V^2)$  space, as it stores the weights of edges between all pairs of vertices.

**B. Priority Queue:**

- The priority queue (or min-heap) used to efficiently fetch the vertex with the minimum distance requires space proportional to the number of vertices. Thus, it takes  $O(V)$  space.

**C. Visited Set:**

- A boolean array or a set to keep track of visited vertices, which requires  $O(V)$  space.

Therefore, the total space complexity of Dijkstra's algorithm is:

- When using an adjacency list:  $O(V + E)$
- When using an adjacency matrix:  $O(V^2)$  In general, for a sparse graph (where  $E \approx V$ ), the adjacency list is more space-efficient, whereas for a dense graph (where  $E \approx V^2$ ), the adjacency matrix might be considered.

Overall, the space complexity of Dijkstra's algorithm is typically  $O(V + E)$  when using an adjacency list representation, which is the most common and efficient way to represent graphs in this context.

**5.2. Edge Deletion Algorithm:**

The Edge Deletion Algorithm for finding the shortest path involves removing some specific edges one by one and computing the shortest path repeatedly. The space complexity of this algorithm is influenced by the way the graph is stored, the space required to compute and store the shortest paths, and the additional space for managing edge deletions. Here's a breakdown of the space requirements:

**A. Graph Representation:**

- **Adjacency List:** This representation requires  $O(V + E)$  space, where  $V$  is the number of vertices and  $E$  is the number of edges. It efficiently stores the edges and their weights.
- **Adjacency Matrix:** This representation requires  $O(V^2)$  space, as it stores the edge weights between all pairs of vertices.

**B. Distance Array:**

An array of size  $V$  to store the shortest known distance from the source to each vertex during each iteration of the shortest path calculation. This requires  $O(V)$  space.

**C. Priority Queue:**

The priority queue (or min-heap) used to fetch the vertex with the minimum distance requires  $O(V)$  space.

**D. Edge Deletion Management:**

To manage edge deletions, you may need additional structures to keep track of which edges have been deleted and to store the state of the graph after each deletion. This can vary based on implementation but typically requires  $O(E)$  space in the worst case.

**E. Auxiliary Structures:**

Additional structures such as arrays or lists to store intermediate results, paths, or the state of the graph can add to the space complexity. This is typically  $O(E)$  or  $O(V)$  depending on the specifics of the implementation.

Combining these components, the overall space complexity is influenced primarily by the graph representation and the priority queue. Given the typical usage of an adjacency list and the necessary data structures, the space complexity can be summarized as:

- **Adjacency List Representation:**  $O(V + E)$
- **Adjacency Matrix Representation:**  $O(V^2)$

Considering the worst-case scenario, the space complexity of the Edge Deletion Algorithm for finding the shortest path is  $O(V + E)$  when using an adjacency list representation. This includes the space for storing the graph, the priority queue, distance arrays, and additional structures for managing edge deletions.

**5.3. Vertex Assignment Enumeration Algorithm:**

The Vertex Assignment Enumeration Algorithm involves assigning vertices based on specific criteria: maximum weighted vertex, minimum domination vertex, and maximum connectivity in the graph. Here's a breakdown of the space complexity for such an algorithm:

**Graph Representation:**

- **Adjacency List:** Requires  $O(V + E)$  space, where  $V$  is the number of vertices and  $E$  is the number of edges.

- **Adjacency Matrix:** Requires  $O(V^2)$  space, as it stores the edge weights between all pairs of vertices.

#### **Data Structures for Vertex Assignment:**

- **Weights Array:** An array of size  $V$  to store the weights of each vertex. This requires  $O(V)$  space.
- **Domination Set:** A data structure to store the domination status of vertices. This can be a boolean array or a set, requiring  $O(V)$  space.

#### **Connectivity Array/Matrix:**

- **Adjacency List:** To store connectivity, it remains  $O(V + E)$
- **Adjacency Matrix:** To store connectivity, it remains  $O(V^2)$

#### **Maximum weighted vertex:**

- **Priority Queue:** Used to efficiently fetch vertices based on their weights or connectivity. This requires  $O(V)$  space.
- **Assigned Vertices Set:** A set to keep track of vertices that have been assigned. This requires  $O(V)$  space.
- **Auxiliary Structures:** Arrays or lists to store intermediate results, temporary states, or the final assignments. This typically requires  $O(V)$  space.

#### **Overall Space Complexity:**

Combining these components, the overall space complexity is primarily influenced by the graph representation and the additional data structures used for vertex assignment.

#### **Using Adjacency List Representation:**

- Graph storage:  $O(V + E)$
- Weights array:  $O(V)$
- Domination set:  $O(V)$
- Priority queue:  $O(V)$
- Assigned vertices set:  $O(V)$
- Auxiliary structures:  $O(V)$

Summing these components, the total space complexity with an adjacency list representation is:

$$O(V + E) + O(V) + O(V) + O(V) + O(V) + O(V) = O(V + E)$$

#### **Using Adjacency Matrix Representation:**

- Graph storage:  $O(V^2)$
- Weights array:  $O(V)$
- Domination set:  $O(V)$

- Priority queue:  $O(V)$
- Assigned vertices set:  $O(V)$
- Auxiliary structures:  $O(V)$

The total space complexity with an adjacency matrix representation is:

$$O(V^2) + O(V) + O(V) + O(V) + O(V) + O(V) = O(V^2)$$

### Summary:

The space complexity of the Vertex Assignment Enumeration Algorithm depends on the graph representation:

- **Adjacency List Representation:**  $O(V + E)$
- **Adjacency Matrix Representation:**  $O(V^2)$

Given the usual preference for adjacency lists in sparse graphs (where  $E$  is much less than  $V^2$ ), the space complexity is typically  $O(V + E)$ . For dense graphs, the adjacency matrix representation leads to a space complexity of  $O(V^2)$

## 6. Conclusion

This comparative analysis of Dijkstra's algorithm, the Edge Deletion algorithm, and Vertex Assignment Enumeration techniques for finding minimum the cost route from the graphs reveals distinct advantages and limitations for each method. Although Dijkstra's algorithm has well-established efficiency and is widespread but also the Edge Deletion algorithm and Vertex Assignment Enumeration Technique prove highly effective in dense graphs with non-negative weights, excelling in scenarios requiring consistent and reliable performance, such as network routing and geographical mapping. The performance diminishes in sparse graphs or dynamic environments where edge weights frequently change.

The Edge Deletion algorithm presents here an attractive alternative for dynamic and sparse graphs, where its iterative edge removal process adapts well to changing graph structures. Its efficiency in these environments makes it a viable choice for applications with frequent updates and modifications to edge weights. Its performance can suffer in dense graphs or when the graph structure remains static. The Vertex Enumeration algorithm is a newly developed algorithm commonly employed and offers unique benefits in specialized applications where vertex properties significantly impact path costs. It can be computationally intensive, particularly used in the case of large graphs.

The experimental results highlight that no single algorithm universally outperforms the others across all scenarios. Instead, the

choice of algorithm should be specific characteristics of the graph and the requirements of the application. Dijkstra's algorithm is recommended for dense, static graphs whereas the Vertex Assignment Enumeration algorithm for specialized cases with significant vertex properties, and the Edge Deletion algorithm for dynamic, sparse graphs.

Future research may explore hybrid approaches that leverage the strengths of these algorithms, potentially developing more robust and versatile solutions for minimum-cost route finding in diverse graph environments. This study underscores the importance of understanding the unique features and uses cases of each algorithm to make informed decisions in practical applications.

### References:

1. Ojekudo, Nathaniel Akpofure, Akpan, Nsikan Paul, (May-June, 2017), "**An Application of Dijkstra's Algorithm to shortest route problem**". IOSR Journal of Mathematics (IOSR-JM). Volume 13, Issue 3 Ver. 1 PP 20-32
2. G. Deepa, Priyank Kumar, A. Manimaran, K. Rajakumar, and V. Krishnamoorthy, (2018), "**Dijkstra Algorithm Application: Shortest Distance between Buildings**" International Journal of Engineering & Technology. Vol: 7 (4.10) P.P. 974-976
3. Arjun R.K , Pooja Reddy , Shama , M. Yamuna,.( April 2015). "**Research on the Optimization of DIJKSTRA'S Algorithm and its Application**" International Journal of Science, Technology & Management, Volume No 04, Special Issue No. 01, April 2015 ISSN (online): 2394-1537 304.
4. Muhammad Adeel Javaid., (January 2013) "**Understanding DijkstraAlgorithm**" Article in SSRN Electronic Journal, DOI: 10.2139/ssrn.2340905.
5. Wang Shu-Xi, (December 2012). "**The Improved Dijkstra's Shortest Path Algorithm and Its Application**" Procedia Engineering 29:1186-1190, DOI:10.1016/j.proeng.2012.01.110
6. Ravi, N., Sireesha, V. Using Modified Dijkstra's (2010). "**Algorithm for Critical Path Method in a Project Network**". International Journal of Computational and Applied Mathematics. Volume 5 number 2. P.P. 217-225.

# Controllable susceptibilities in GaAs-AlGaAs Quantum Dot by strong control field

Abhijit Shyam<sup>\*1</sup>, Sanchita Roy<sup>2</sup>, Nitu Borgohain<sup>3</sup>

<sup>\*1,2</sup>*Department of Physics, University of Science & Technology Meghalaya, Meghalaya, India-783101*

<sup>3</sup>*Department of Physics, Royal Global University, Assam, India-781035*

*Corresponding author: abhijit.shyam12@gmail.com*

## Abstract

Linear and Kerr susceptibilities have been theoretically investigate in a three level ladder type GaAs-AlGaAs quantum dot nanostructure under the electromagnetically induced transparency regime. A faint probe pulse and a strong laser beam are applied to interact within the scheme, resulting in nonlinear susceptibility found on the order of  $10^{-15}m^2/V^2$  at the probe wavelength  $1.293 \mu m$ . This investigation may highlight the potential applicability of these nanostructures optical communications, imaging, and spectroscopy. This work provides a deeper understanding the dynamics in semiconductor quantum dots and paves the way for novel photonic devices based on quantum-confined structures.

**Keywords :** Semiconductor Quantum Dot, Linear-Nonlinear susceptibilities, Electromagnetically Induced Transparency.

## Introduction

Recently, in the fields of theoretical and experimental work on nonlinear optics, the quantum dot gained tremendous attention due to its unique properties like confinement in all the spatial

dimensions, fascinating class of nanomaterials which exhibits strong and tunable fluorescent properties and have highly photoluminescent emissions [1], this study of nonlinear light-matter interactions, owing to their remarkable opto-electronic properties. These properties - large electric dipole moments and high nonlinear coefficients- bring a number of potential benefits for opto-electronic devices. In addition, materials with high nonlinear susceptibilities are very promising for applications in sensitive, low-power optical systems; this area is therefore increasingly important [1–10]. Comparatively, quantum dots stand out by offering significant intrinsic advantages over quantum wells and atomic vapors [11-16]. One major advantage of quantum dots are it have large electric dipole moments and high nonlinear coefficients, which gives solid-state devices that utilize both interband and inter-subband transitions. All these properties allow the construction of very sensitive optical systems capable of working at significantly lower power levels. Quantum dots can be tuned to be many different devices, and symmetries can also be made available to suit purposes. The best thing about them is that they can easily integrate with other semiconductor devices making them versatile and very practical [1, 3-6, 9, 10, 17, 18].

Quantum interference effects, such as Electromagnetically Induced Transparency (EIT), arise from the coherent interaction of light with matter and play a significant role in various quantum and optical phenomena which provides a usually translucent medium transparent to the probe light under the influence of a control light, effectively minimizing absorption and augmenting nonlinear properties [19-21]. Electromagnetically induced transparency (EIT) is one of the most effective techniques for enhancing nonlinear optical properties [17, 22-25]. Recent studies have focused on the large nonlinear susceptibilities of quantum dot (QD) nanostructures, particularly by reducing or even eliminating linear absorption of the probe light through EIT, which is also observed in single quantum dot by H. Gotoh et.al. and S. Marcinkevičius et.al. [26, 27], respectively. This approach has received ample attention due to its great potential applications in developing the advanced optical systems with good performance and efficiency [17, 18, 26-32]. By way of more example, Zhang et al. [33] demonstrated that EIT in QD systems may result in significant third-order or Kerr

nonlinearities, so they are of interest for low-power photonic applications, whereas Xu and Malko [34] proved large third-order nonlinear optical coefficients in semiconductor QDs which are promising in optical switching and modulation applications.. It is clear from the literature mentioned above that improving the nonlinear optical characteristics of quantum dots (QDs) is essential to enable their useful uses. Our goal in this work is to examine and measure the Kerr nonlinearities in QD systems when Electromagnetically Induced Transparency (EIT) is present. Our objective is to aid in the creation of sophisticated nonlinear photonic devices that function well at low power levels.

In the present paper, we have investigated the response of a probe laser pulse in a ladder-type three-level quantum dot system, driven by an additional coupling field. We have studied the electromagnetically induced transparency (EIT) phenomenon using the density matrix method and obtained linear and third order nonlinear susceptibilities at the probe laser frequency.

The paper has been laid out in the following way: in Section 2, the three-level, ladder-type QD theoretical model that interacts with the optical fields is presented, followed by the evaluation of the linear and nonlinear susceptibilities. Section 3 delves into exploring the properties of susceptibilities under varied control field parameters. Lastly, Section 4 encapsulates a concise summary and conclusion.

### **Mathematical Model and Governing Equations**

We consider a strained GaAs-AlGaAs quantum dot (QD), this structure was first reported by T. Tsujikawa et al. (2000), with a quantum disc structure of 3.5 nm height and 9 nm radius. The QD experiences bandgap shifts due to biaxial compressive strain, altering its electronic properties. Formed through the self-assembled method, the QDs exhibit inhomogeneous broadening due to size, strain, and material composition variations, causing detuning effects for probe and control fields. To address this, we propose placing a small number of QDs in a resonant microcavity to mitigate the impact of inhomogeneous broadening [35].

Consider a three-level GaAs-AlGaAs quantum dot nanostructure with a ladder configuration, where all transitions are dipole-allowed. In this setup, a faint probe laser pulse with angular frequency represented by  $\omega_p$  with amplitude  $E_p$  and Rabi frequency

$\Omega_p = \frac{\mu_{21}E_p}{2\hbar}$  is coupled to the  $|1\rangle \rightarrow |2\rangle$  transition while a strong control laser pulse with angular frequency represented by  $\omega_c$  with amplitude  $E_c$  and Rabi frequency  $\Omega_c = \frac{\mu_{32}E_c}{2\hbar}$  is coupled to the  $|2\rangle \rightarrow |3\rangle$  transition, as illustrated in Figure 1.

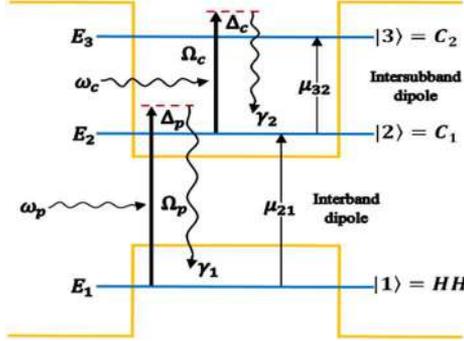


Fig-1: Schematic of energy levels GaAs-Al-GaAs quantum dot with ladder type transition. [HH = the heavy-hole ground state, C1 = the conduction band ground state, and C2 = the first excited state of conduction band]

Here,  $\mu_{mn} (= e\langle m|z|n\rangle)$  is the dipole moment between the states. The system considered here the transition energies between the states are  $E_{21} = 87.68meV$  and  $E_{32} = 9.28meV$  respectively. The applied electric field for the system as

$$\vec{E} = \hat{e}_p E_p e^{-i\omega_p t} + \hat{e}_c E_c e^{-i\omega_c t} + c. c. \quad (1),$$

where,  $\hat{e}_p, \hat{e}_c$  are the unit vector along the direction of polarization and  $c. c.$  is the complex conjugate of the applied electric field [36]. For the SQD system, In Schrödinger format, the semi-classical Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}_i \quad (2),$$

where, unperturbed or non-interacting Hamiltonian represented by  $\hat{H}_0$  and perturbed or interacting Hamiltonian represented by  $\hat{H}_i$ . Now expand the equation (2) as

$$\hat{H} = \sum_{i=1}^3 \hbar\omega_i |i\rangle\langle i| - \hbar\Omega_p e^{-i\omega_p t} |2\rangle\langle 1| - \Omega_c e^{-i\omega_c t} |3\rangle\langle 2| + h. c. \quad (3),$$

In equation (3),  $\hbar\omega_i$  is the energy associate with the state  $|i\rangle$  of the system and  $h.c.$  is the Hermitian conjugate [36]. After solving the equation (3), with the help of density matrix formalism and rotating frame approximation, we obtained Maxwell-Bloch equations [28] as

$$\frac{\partial}{\partial t} \tilde{\rho}_{11} = i(\Omega_p^* \tilde{\rho}_{21} - \Omega_p \tilde{\rho}_{12}), \quad (4a),$$

$$\frac{\partial}{\partial t} \tilde{\rho}_{22} = -\gamma_2 \tilde{\rho}_{22} + i(\Omega_p \tilde{\rho}_{12} - \Omega_p^* \tilde{\rho}_{21}) + i(\Omega_c \tilde{\rho}_{23} - \Omega_c^* \tilde{\rho}_{32}), \quad (4b),$$

$$\frac{\partial}{\partial t} \tilde{\rho}_{33} = -\gamma_3 \tilde{\rho}_{33} + i(\Omega_c \tilde{\rho}_{23} - \Omega_c^* \tilde{\rho}_{32}), \quad (4c),$$

$$\frac{\partial}{\partial t} \tilde{\rho}_{21} = i\left(\Delta_p + \frac{i\gamma_{21}}{2}\right) \tilde{\rho}_{21} + i\Omega_p(\tilde{\rho}_{11} - \tilde{\rho}_{22}) + i\Omega_c \tilde{\rho}_{31}, \quad (4d),$$

$$\frac{\partial}{\partial t} \tilde{\rho}_{31} = i\left(\Delta_p + \Delta_c + i\frac{\gamma_{31}}{2}\right) \tilde{\rho}_{31} - i\Omega_p \tilde{\rho}_{32} + i\Omega_c \tilde{\rho}_{21}, \quad (4e),$$

$$\frac{\partial}{\partial t} \tilde{\rho}_{32} = i\left(\Delta_c + \frac{i\gamma_{32}}{2}\right) \tilde{\rho}_{32} + \{i\Omega_c(\tilde{\rho}_{33} - \tilde{\rho}_{22}) - i\Omega_p^* \tilde{\rho}_{31}\}, \quad (4f),$$

$$\tilde{\rho}_{11} + \tilde{\rho}_{22} + \tilde{\rho}_{33} = 1, \quad (4g),$$

$$\tilde{\rho}_{ij} = \tilde{\rho}_{ji}, (i, j = 1, 2, 3) \quad (4h)$$

In the equations above, we have  $\Delta_p (= \omega_p - \omega_{21})$  and  $\Delta_c (= \omega_c - \omega_{32})$ , these represent the detuning frequencies for the probe and control fields. We also look at  $\gamma_j$  ( $j = 2, 3$ ), which are the rates at which the populations decay, largely due to the inelastic emissions of longitudinal optical phonons. Then, we have  $\gamma_{jk}$  ( $j \neq k$ ) indicating the total rates of coherence relaxation. For example, we can express  $\gamma_{21} = \gamma_2 + \gamma_{21}^{dp}$ ;  $\gamma_{31} = \gamma_3 + \gamma_{31}^{dp}$  and  $\gamma_{32} = \gamma_2 + \gamma_3 + \gamma_{32}^{dp}$ , where  $\gamma_{jk}^{dp}$ , which includes both the quasi-elastic acoustic phonon scattering and the elastic interface Raman scattering. Now we assume that, initially the system in the ground state  $|i\rangle$  so  $\rho_{11} = 1$  and  $\rho_{22} = 0 = \rho_{33}$  at time  $t = 0$ . Under the faint probe approximation, In the context of the adiabatic framework we obtained  $\rho_{ij}$  solution to the first order of  $\Omega_p$  from the 4(d), 4(e) as

$$\tilde{\rho}_{21}^{(1)} = -\frac{(\Delta_p + \Delta_c + i\frac{\gamma_{31}}{2})\Omega_p}{(\Delta_p + i\frac{\gamma_{21}}{2})(\Delta_p + \Delta_c + i\frac{\gamma_{31}}{2}) - |\Omega_c|^2} \quad (5a),$$

$$\tilde{\rho}_{31}^{(1)} = -\frac{\Omega_c\Omega_p}{(\Delta_p + i\frac{\gamma_{21}}{2})(\Delta_p + \Delta_c + i\frac{\gamma_{31}}{2}) - |\Omega_c|^2} \quad (5b),$$

The evolution of the electric field govern by Maxwell's equation can be written as

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \vec{E} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \langle \vec{P} \rangle}{\partial t^2} \quad (6),$$

where,  $\langle \vec{P} \rangle$  represents the induced polarization in the medium due to the probe and control fields, expressed as  $\langle \vec{P} \rangle = N \langle \vec{D} \rangle$ , where  $N$  is the electron number density in the quantum dot, and  $\langle \vec{D} \rangle$  is the dipole moment operator. The other parameters include  $c$ , the speed of light;  $\epsilon_0$ , the permittivity of free space;  $\hbar$ , the reduced Planck's constant; and  $\omega_p$ , the frequency of the probe field. Under the slowly varying envelope approximation, the evolution of the probe field  $\vec{E}_p$  is governed by Maxwell's wave equation. The resulting first-order equation for  $\Omega_p$  can be expressed as follows

$$\frac{\partial \Omega_p}{\partial z} + \frac{1}{c} \frac{\partial \Omega_p}{\partial t} = i\alpha \tilde{\rho}_{21}^{(1)} + \left[-i\alpha \left(|\tilde{\rho}_{21}^{(1)}|^2 + |\tilde{\rho}_{31}^{(1)}|^2\right) \tilde{\rho}_{21}^{(1)} + i\alpha \left(|\tilde{\rho}_{21}^{(1)}|^2 + |\tilde{\rho}_{31}^{(1)}|^2\right) \tilde{\rho}_{21}^{(1)}\right], \quad (7)$$

where,  $\alpha \left(= \frac{N|\mu_{21}|^2 \omega_p}{4\epsilon_0 \hbar c}\right) = constant$ . In the order to solve the linearized version of the equation (7), we take Fourier Transformation which yields

$$\frac{\partial \Lambda_p}{\partial z} + \frac{1}{c} \frac{\partial \Lambda_p}{\partial t} = i\alpha \beta_{21}^{(1)} \quad (8)$$

Here,  $\Lambda_p$  and  $\beta_{21}^{(1)}$  are the Fourier transformation of  $\Omega_p$  and  $\tilde{\rho}_{21}^{(1)}$  respectively. Now the solution of the equation (8) can be written as

$$\Lambda_p = \Lambda_0 e^{iK(\omega)z} \quad (9)$$

where,  $K(\omega)$  is the linear dispersion, which related to probe field and can be expand around the central frequency  $\omega_p$  as,

$$K(\omega) = \frac{1}{0!}K(0) + \frac{\omega}{1!}K_1(0) + \frac{\omega^2}{2!}K_2(0) + \dots \quad (10)$$

where,  $K(0) = \phi + i\alpha/2$  describe the phase shift and absorption coefficient, while  $K_1(0)$  is the propagation constant related to group velocity as  $v_g = Re \left[ \frac{1}{K(0)} \right]$  and the third term  $K_2(0)$  is called the group velocity dispersion of the probe field leading to change of the shape of the propagation pulse and addition loss or gain of the probe field intensity.

At the probe frequency  $\omega_p$ , the induced polarization can be written as

$$P(\omega_p) = \varepsilon_0 \chi_p(\omega_p) \vec{E}_p, \quad (11)$$

where  $\chi_p(\omega_p)$  is the susceptibility of the system due to probe field. The susceptibility can be written for the linear and nonlinear parts as follows

$$\chi_p(\omega_p) = \chi^{(1)} + \chi^{(3)} |\vec{E}_p|^2 + \dots \quad (12)$$

In the equation (12), neglecting the higher order terms (5<sup>th</sup>, 7<sup>th</sup>, etc). The linear and third order susceptibilities at the probe frequency can be written as follows:

$$\chi^{(1)} = \frac{N|\mu_{21}|^2}{\hbar\varepsilon_0\Omega_p} \tilde{\rho}_{21}^{(1)} = -\frac{N|\mu_{21}|^2}{2\varepsilon_0\hbar} \cdot \frac{D_p(0)}{D(0)}, \quad (13a)$$

$$\chi^{(3)} = -\frac{N|\mu_{21}|^2}{2\varepsilon_0\hbar\Omega_p} \tilde{\rho}_{21}^{(1)} \left( |\tilde{\rho}_{21}^{(1)}|^2 + |\tilde{\rho}_{31}^{(1)}|^2 \right) = \frac{N|\mu_{21}|^4}{8\varepsilon_0\hbar^3} \cdot \frac{D_p(0)}{D(0)} \cdot \left\{ \frac{|D_p(0)|^2 + |\Omega_c|^2}{|D(0)|^2} \right\}. \quad (13b)$$

Now we going to analyze the linear and third order nonlinear susceptibility for the SQDs system.

## Results

For the investigation, system parameters taken (calculated) as  $N = 3 \times 10^{20} m^{-3}$ ,  $\mu_{21} = 3.36 \times 10^{-28} C.m$ ,  $\mu_{32} = 3.94 \times 10^{-28} C.m$ ,  $\lambda_p = 1.36 \times 10^{-6} m$ ,  $\lambda_c = 1.28 \times 10^{-6} m$ ,  $\omega_{21} = 2.1 \times 10^{14} s^{-1}$ ,  $\omega_{32} = 2.23 \times 10^{14} s^{-1}$ ,  $\omega_p = 2.206 \times 10^{14} s^{-1}$ ,  $\omega_c = 2.344 \times 10^{14} s^{-1}$ , and the decay rates  $\gamma_{21} = \gamma_{32} =$

$8.41 \times 10^{12} s^{-1}$  thus  $\alpha = 6.367 \times 10^{15} m^{-1} s^{-1}$ . Now we are going to investigate the linear and nonlinear susceptibilities (up to third order) first. The imaginary part of the linear susceptibility  $Im\chi^{(1)}$  represents the medium's absorption or gain of the probe field while real part of the linear susceptibility  $Re\chi^{(1)}$  describes the dispersion, indicating how the refractive index changes with frequency. The fig-2a, illustrates the variation of  $Im\chi^{(1)}$  representing absorption, with normalized probe field detuning ( $\Delta_p/\delta$ ) for different values of the control field frequency ( $\Omega_c/\delta$ ) ranging from 0 to 5. Where the term  $\delta$  ( $= 10^{13} s^{-1}$ ) is the normalization factor. When  $\Omega_c/\delta = 0$ , the absorption follows a typical Lorentzian profile, peaking at the resonance frequency due to the probe field's interaction with the medium. As  $\Omega_c/\delta$  increases, the influence of the control field becomes evident. The absorption profile modifies, showing reduced absorption at resonance and the formation of transparency windows or additional peaks. This behavior results from quantum destructive interference of multiple paths for interaction of system and intense light and the result is electromagnetically induced transparency (EIT) or Autler-Townes splitting, which depending on the control field strength, and in these process dominates the two-photon absorption. The figure demonstrates how varying  $\Omega_c/\delta$  effectively controls the probe field's absorption, highlighting the medium's tunable optical properties.

The fig-2b depicted the variations of  $Re\chi^{(1)}$  representing the dispersion of the medium, with normalized probe field detuning for different values of the control field frequency ranging from 0 to 5. At  $\Omega_c/\delta = 0$ , the dispersion profile follows a standard pattern with a gradual slope around the resonance frequency. As  $\Omega_c/\delta$ , increases, the control field significantly influences the dispersion, introducing steeper slopes or inflection points in the curve. This behavior is associated with quantum effects such as electromagnetically induced transparency (EIT), which enhances the control over light propagation. The steep slopes in  $Re\chi^{(1)}$  near the resonance indicate large variations in the refractive index, affecting the phase velocity and group velocity of the probe field. This tunability of dispersion with  $\Omega_c$  highlights the medium's potential for applications in slow light and optical signal processing.

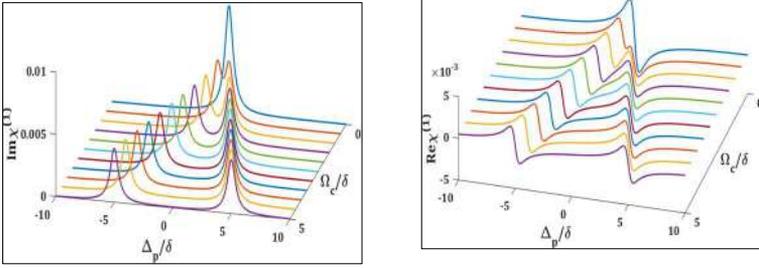


Fig 2: First order susceptibility profiles (a) **3D-plot of  $Im\chi^{(1)}$** , (b) **3D-plot of  $Re\chi^{(1)}$**  at different values of control field laser  $\Omega_c/\delta$  has been depicted to show that **EIT window**.

In the fig 3, the graphical representation of the  $Re\chi^{(3)}$  illustrates the variation of the third-order nonlinear susceptibility with normalized probe field detuning for different values of the control field frequency  $\Omega_c/\delta$ .  $Re\chi^{(3)}$  represents the medium's response to intense electromagnetic fields, and it is associated with nonlinear optical effects such as self-phase modulation, cross-phase modulation, and third-harmonic generation. When the first control field frequency  $\Omega_c/\delta$  is switched OFF,  $Re\chi^{(3)}$  exhibits a single peak around the normalized probe field detuning ( $\Delta_p/\delta = 0$ ). When the first control field is switched ON, a second peak appears. As the value of  $\Omega_c/\delta$  increases, the separation between the two peaks grows, demonstrating the influence of the control field on the nonlinear dispersion. Importantly, the plot indicates a large Kerr nonlinear susceptibility of the order of  $\sim 10^{-15} m^2/V^2$ . This tunable behavior, with high values underscores the medium's potential for applications in nonlinear optical systems, such as optical switching, signal processing, and high-sensitivity detection. The control field's influence ensures precise modulation of the nonlinear properties, making the system highly versatile.

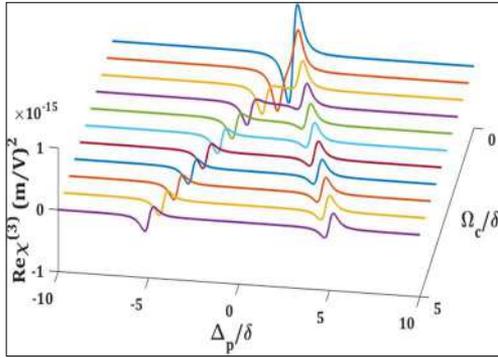


Fig 3: Third order or Kerr susceptibility profiles (a) **3D-plot** of  $Re\chi^{(3)}$ , at different values of control field laser  $\Omega_c/\delta$  has been depicted to show that **EIT window**.

To seamlessly transition into a comparison table after concluding the discussion on linear and third order susceptibilities (up to third order). At this point, it is insightful to compare the third-order susceptibility  $Re\chi^{(3)}$  values reported in various materials with those observed in the GaAs-AlGaAs quantum dot system analyzed in this study. Table 1 provides a clear comparison, demonstrating that the  $Re\chi^{(3)}$  value measured for the GaAs-AlGaAs quantum dot is significantly higher than those found in other materials. This superior third-order susceptibility highlights the exceptional nonlinear optical properties of the GaAs-AlGaAs quantum dot, making it a promising candidate for advanced photonic and optoelectronic applications.

Table1 1: Comparison of  $\chi^{(3)}$  values of different materials

S.N.	Materials Type	$\chi^{(3)}$ (m/V) <sup>2</sup>	( $\lambda_p$ $\mu m$ )	Refs.
1	As <sub>2</sub> Se <sub>3</sub> glass	$4.7 \times 10^{-23}$	1.064	37
2	SiO <sub>2</sub> fiber	$4.2 \times 10^{-23}$	1.550	38
3	Semiconductor doped glass fiber	$4.5 \times 10^{-19}$	0.740	39
4	GaN/AlN QW	$2.2 \times 10^{-16}$	1.550	40
5	GaAs-AlGaAs QD	$\sim 1 \times 10^{-15}$	1.293	Present Paper

## Conclusion

To conclude that the control field frequency has a great influence on the linear and nonlinear susceptibility of the probe field which interacting in a three level ladder type GaAs-AlGaAs quantum dot system. This system achieved Kerr nonlinear susceptibility of the order of  $\sim 1 \times 10^{-15} m^2/V^2$  at very low absorption level. So this theoretical investigation may have much potential applications in many solid systems etc.

## Acknowledgment

Authors acknowledge the moral and financial support from Shri M. Hoque, Chancellor, University of Science and Technology Meghalaya with thanks.

## Reference

1. Leong, K. H., Chin, Y. H., Sim, L. C., Tan, B., Dai, C., & Saravanan, P. (2022). Physical properties of quantum dots. In Graphene, Nanotubes and Quantum Dots-Based Nanotechnology (pp. 687-709). Woodhead Publishing.
2. Loss, D., & DiVincenzo, D. P. (1998). Quantum computation with quantum dots. *Physical Review A*, 57(1), 120.
3. Zhao, Y., Dyck, J. S., & Burda, C. (2011). Toward high-performance nanostructured thermoelectric materials: the progress of bottom-up solution chemistry approaches. *Journal of Materials Chemistry*, 21(43), 17049-17058.
4. Bhattacharya, P., Ghosh, S., & Stiff-Roberts, A. D. (2004). Quantum dot opto-electronic devices. *Annu. Rev. Mater. Res.*, 34(1), 1-40.
5. Kagan, C. R., Lifshitz, E., Sargent, E. H., & Talapin, D. V. (2016). Building devices from colloidal quantum dots. *Science*, 353(6302), aac5523.
6. Norman, J. C., Jung, D., Zhang, Z., Wan, Y., Liu, S., Shang, C., ... & Bowers, J. E. (2019). A review of high-performance quantum dot lasers on silicon. *IEEE Journal of Quantum Electronics*, 55(2), 1-11.
7. Huffaker, D. L., Park, G., Zou, Z., Shchekin, O. B., & Deppe, D. G. (1998). 1.3  $\mu\text{m}$  room-temperature GaAs-based

- quantum-dot laser. *Applied Physics Letters*, 73(18), 2564-2566.
8. Bera, D., Qian, L., Tseng, T. K., & Holloway, P. H. (2010). Quantum dots and their multimodal applications: a review. *Materials*, 3(4), 2260-2345.
  9. Ali Sobhani, S. (2020). State-of-the-art InAs/GaAs quantum dot material for optical telecommunication (Doctoral dissertation, University of Glasgow).
  10. Singh, J. (2007). *Electronic and optoelectronic properties of semiconductor structures*. Cambridge University Press.
  11. Niu, Y., Gong, S., Li, R., Xu, Z., & Liang, X. (2005). Giant Kerr nonlinearity induced by interacting dark resonances. *Optics letters*, 30(24), 3371-3373.
  12. Hang, C., & Huang, G. (2010). Giant Kerr nonlinearity and weak-light superluminal optical solitons in a four-state atomic system with gain doublet. *Optics Express*, 18(3), 2952-2966.
  13. Sun, H., Gong, S., Niu, Y., Jin, S., Li, R., & Xu, Z. (2006). Enhancing Kerr nonlinearity in an asymmetric double quantum well via Fano interference. *Physical Review B—Condensed Matter and Materials Physics*, 74(15), 155314.
  14. Konar, S., Jana, S., & Mishra, M. (2005). Induced focusing and all optical switching in cubic quintic nonlinear media. *Optics Communications*, 255(1-3), 114-129.
  15. Agrawal, G. P. (2000). *Nonlinear fiber optics*. In *Nonlinear Science at the Dawn of the 21st Century* (pp. 195-211). Berlin, Heidelberg: Springer Berlin Heidelberg.
  16. Yang, W. X., & Lee, R. K. (2008). Controllable entanglement and polarization phase gate in coupled double quantum-well structures. *Optics Express*, 16(22), 17161-17170.
  17. Sahrai, M., & Abbasabadi, M. (2018). All-optical switch based on doped graphene quantum dots in a defect layer of a one-dimensional photonic crystal. *Applied Optics*, 57(3), 521-526.
  18. Vincenti, M. A., de Ceglia, D., Grande, M., D’Orazio, A., & Scalora, M. (2013). Nonlinear control of absorption in one-dimensional photonic crystal with graphene-based defect. *Optics letters*, 38(18), 3550-3553.

19. Marangos, J.P., 1998. Electromagnetically induced transparency. *Journal of modern optics*, 45(3), pp.471-503.
20. Borgohain, N., Belic, M. and Konar, S., 2015. Giant parabolic nonlinearities at infrared in  $\Lambda$ -type three level multiple quantum wells. *Annals of Physics*, 361, pp.107-119
21. Borgohain, N. and Konar, S., 2016. The effects of control field detuning on the modulation instability in a three-level quantum well system. *Journal of Applied Physics*, 119(21)
22. Nielsen, P. K., Thyrrstrup, H., Mørk, J., & Tromborg, B. (2007). Numerical investigation of electromagnetically induced transparency in a quantum dot structure. *Optics express*, 15(10), 6396-6408.
23. Litvak, A. G., & Tokman, M. D. (2002). Electromagnetically induced transparency in ensembles of classical oscillators. *Physical Review Letters*, 88(9), 095003.
24. Harris, S. E., & Hau, L. V. (1999). Nonlinear optics at low light levels. *Physical Review Letters*, 82(23), 4611.
25. Harris, S. E. (1997). Electromagnetically induced transparency. *Physics today*, 50(7), 36-42.
26. Gotoh, H., Kamada, H., Saitoh, T., Ando, H., & Temmyo, J. (2005). Exciton absorption properties of coherently coupled exciton-biexciton systems in quantum dots. *Physical Review B—Condensed Matter and Materials Physics*, 71(19), 195334.
27. Marcinkevičius, S., Gushterov, A., & Reithmaier, J. P. (2008). Transient electromagnetically induced transparency in self-assembled quantum dots. *Applied Physics Letters*, 92(4).
28. Shyam, A., Borgohain, N., Mukherjee, R., & Hazra, R. (2024). Phase sensitive slow light in semiconductor quantum dots. *Optics Communications*, 569, 130867.
29. Sabet, R. A., Sahrai, M., & Sattari, H. (2017). Utilizing electromagnetically induced transparency in InAs quantum dots for all-optical transistor design. *Applied Optics*, 56(28), 7944-7951.
30. Sattari, H., Ebadollahi-Bakhtevan, S., & Sahrai, M. (2016). Proposal for a  $1 \times 3$  Goos-Hänchen shift-assisted de/

- multiplexer based on a multilayer structure containing quantum dots. *Journal of Applied Physics*, 120(13).
31. Lunnemann, P., & Mørk, J. (2009). Reducing the impact of inhomogeneous broadening on quantum dot based electromagnetically induced transparency. *Applied Physics Letters*, 94(7).
  32. Baretin, D., Houmark, J., Lassen, B., Willatzen, M., Nielsen, T. R., Mørk, J., & Jauho, A. P. (2009). Optical properties and optimization of electromagnetically induced transparency in strained InAs/GaAs quantum dot structures. *Physical Review B—Condensed Matter and Materials Physics*, 80(23), 235304.
  33. Zhang, Y., Wang, J., & Xu, Z. (2017). Enhanced Kerr nonlinearity in quantum dots under electromagnetically induced transparency. *Applied Physics Letters*, 110(19), 191103
  34. Xu, X., & Malko, A. V. (2004). Large third-order nonlinear optical coefficients in semiconductor quantum dots. *Journal of Applied Physics*, 95(8), 3983-3987.
  35. Tsujikawa, T., Mori, S., Watanabe, H., Yoshita, M., Akiyama, H., van Dalen, R., ... & Ito, R. (2000). Fabrication and optical properties of GaAs/AlGaAs quantum dot grown in tetrahedral-shaped recesses on GaAs (111) B substrates by MOVPE. *Physica E: Low-dimensional Systems and Nanostructures*, 7(3-4), 308-316.
  36. Bora, D., Chetia, P., Das, J., & Borgohain, N. (2022, March). Investigation of Modulation Instability in a Multiple Quantum Well Nanostructure Under EIT Window. In *Proceedings of the International Conference on Atomic, Molecular, Optical & Nano Physics with Applications: CAMNP 2019* (pp. 381-392). Singapore: Springer Singapore.
  37. Boudebs, G., Cherukulappurath, S., Leblond, H., Troles, J., Smektala, F., & Sanchez, F. (2003). Experimental and theoretical study of higher-order nonlinearities in chalcogenide glasses. *Optics Communications*, 219(1-6), 427-433.
  38. Pushkarov, D., & Tanev, S. (1996). Bright and dark solitary wave propagation and bistability in the anomalous

- dispersion region of optical waveguides with third-and fifth-order nonlinearities. *Optics Communications*, 124(3-4), 354-364.
39. Schmid, W., Vogtmann, T., & Schwoerer, M. (1995). A modulation technique for measuring the optical susceptibility  $\Xi(5)$  by degenerate four-wave mixing. *Optics communications*, 121(1-3), 55-62.
40. Hamazaki, J., Matsui, S., Kunugita, H., Ema, K., Kanazawa, H., Tachibana, T., ... & Kishino, K. (2004). Ultrafast intersubband relaxation and nonlinear susceptibility at 1.55  $\mu\text{m}$  in GaN/AlN multiple-quantum wells. *Applied physics letters*, 84(7), 1102-1104.

# An Analysis of Cybersecurity Challenges and Emerging Trends In Modern Technologies

Dr.U.Sridevi<sup>1</sup> Priya M<sup>2</sup> Dharshini S<sup>2</sup> Sharmila T<sup>2</sup> Abishek J<sup>2</sup>  
Mohamad Arsath<sup>2</sup>

*<sup>1</sup>Assistant Professor, MCA, Excel Engineering College, Komarapalayam, Namakkal, India*

*<sup>2</sup> Student of I - MCA, Excel Engineering College, Komarapalayam, Namakkal, India*

## **Abstract**

Cybersecurity is crucial in the field of information technology, with securing data becoming one of the foremost challenges today. When we think of cybersecurity, the rise of cybercrimes often comes to mind, and these incidents are increasing rapidly. Governments and organizations are implementing various measures to combat these threats, yet cybersecurity remains a significant concern for many. This paper primarily addresses the challenges faced by cybersecurity in relation to emerging technologies, along with the latest techniques, ethical considerations, and evolving trends shaping the landscape of cybersecurity.

**Keywords:** cybersecurity, cybercrime, cyber ethics, social media, cloud computing, mobile applications.

## **1.Introduction**

Today, individuals can effortlessly send and receive various forms of data, such as emails, audio, or video, at the click of a button. However, have they considered how securely their information is being transmitted and whether it reaches the intended recipient without any risk of leakage? The answer lies in cybersecurity. The Internet has become the fastest-growing infrastructure in our daily lives, and many emerging technologies are reshaping society. Yet, these advancements make it increasingly difficult to protect our private information, leading to a rise in cybercrimes. Currently, over 60% of commercial transactions occur online, necessitating high-security standards for safe and transparent transactions. As a result, cybersecurity has emerged as a critical issue. Its scope extends beyond merely protecting data in the IT sector; it also encompasses various fields, including cyberspace.

Emerging technologies such as cloud computing, mobile computing, e-commerce, and online banking require robust security measures. Since these technologies handle sensitive personal information, ensuring their security is essential. Strengthening cybersecurity and protecting critical information infrastructures are vital for the security and economic stability of every nation. Creating a safer Internet and safeguarding users is fundamental to the development of new services and government policies. The fight against cybercrime demands a comprehensive and proactive approach. As technical measures alone cannot eliminate these threats, it is crucial for law enforcement agencies to have the authority to investigate and prosecute cybercrimes effectively. Many countries are implementing strict cybersecurity laws to protect vital information. Additionally, individuals must receive training in cybersecurity to protect themselves from the growing threat of cybercrimes.

## **2. Cyber Crime**

Cybercrime refers to any illegal activity that primarily utilizes a computer to commit offenses, including theft. The U.S. Department of Justice broadens this definition to encompass any illegal acts that involve a computer for storing evidence. The expanding list of cybercrimes includes offenses enabled by technology, such as network intrusions and the spread of computer viruses, as well as digital variations of traditional crimes like identity theft, stalking, bullying, and terrorism, which pose significant challenges for individuals and nations alike. In everyday

terms, cybercrime can be understood as crimes carried out using a computer and the internet to steal someone's identity, sell illegal goods, stalk victims, or disrupt operations through malicious software. As technology continues to play a central role in daily life, the prevalence of cybercrimes is expected to rise alongside technological advancement.

### 3.Cyber Security

Data privacy and security will always be paramount concerns for any organization. We live in a world where information is predominantly stored in digital form. Social networking sites offer a space where users can feel secure while connecting with friends and family. However, home users remain vulnerable, as cyber criminals often target these platforms to steal personal information. Similarly, individuals must take all necessary security precautions during online banking transactions Cyber Crime

Incidents	Jan-June 2012	Jan-June 2013	Increase/ Decrease
Fraud	2439	2490	2
Intrusion	2203	1726	22
Spam	291	614	111
Malicious code	353	442	25
Cyber Harassment	173	233	35
Content related	10	42	320
Intrusion Attempts	55	24	56
Denial of services	12	10	17
Vulnerability reports	45	11	76
Total	5581	5592	(adjusted)

This reflects the totals after correcting the data for clarity. The total number of incidents increased slightly, from 5581 in 2012 to 5592 in 2013

The ongoing cyber security threats. As crime rates rise, so too do security measures. A survey conducted by Silicon Valley Bank among U.S. technology and healthcare executives revealed that companies view cyberattacks as a significant threat to both their data security and business continuity.

- Ninety-eight percent of companies are either maintaining or increasing their cybersecurity resources, with half of those planning to allocate more resources specifically to counter online attacks this year.
- Most companies are preparing for the inevitability of cyberattacks rather than questioning whether they will happen.
- Only one-third of companies are fully confident in the security of their own information, and even fewer trust the security measures of their business partners.

There will be new attacks targeting devices running the Android operating system, but these are not expected to be widespread. Since tablets share the same operating system as smartphones, they will likely be vulnerable to the same malware as those platforms. The number of malware samples targeting Macs will continue to increase, although at a lower rate compared to PCs. With Windows 8, users will have the capability to develop applications for virtually any device (PCs, tablets, and smartphones) that runs this operating system, potentially allowing for the creation of malicious applications similar to those for Android. These trends are some of the predictions for the future of cybersecurity.

#### **4.TRENDS TRANSFORMING CYBERSECURITY**

Below are some key trends that are significantly influencing cybersecurity.

##### **4.1 Web servers:**

The threat of attacks on web applications to extract data or distribute malicious code remains a significant concern. Cybercriminals often use compromised legitimate web servers to spread their malicious software. Data theft, which frequently garners media attention, is another major risk. This highlights the need for stronger protection of web servers and applications. These servers are particularly attractive targets for cybercriminals looking to steal sensitive information. Therefore, using a secure browser, especially during critical transactions, is essential to avoid becoming a victim of such attacks.

##### **4.2 Cloud computing and its services**

Today, businesses of all sizes—small, medium, and large—are gradually adopting cloud services. In other words, the world is steadily transitioning to the cloud. This growing trend poses

significant challenges for cybersecurity, as traffic can bypass traditional inspection points. Furthermore, with the increasing number of cloud-based applications, policy controls for web applications and cloud services must evolve to prevent the loss of valuable data. Although cloud services are developing their own security models, concerns about their safety continue to surface. While the cloud offers tremendous opportunities, it's important to recognize that as the cloud evolves, so do its security challenges.

#### **4.3 APTs and targeted attacks**

An Advanced Persistent Threat (APT) represents a more sophisticated level of cybercrime. For years, network security measures like web filtering and intrusion prevention systems (IPS) have been essential in detecting these targeted attacks, often after the initial breach. However, as attackers become more audacious and use increasingly subtle tactics, network security must collaborate with other security services to identify and mitigate these threats. Therefore, enhancing security strategies is crucial to preventing future attacks.

#### **4.4 Mobile Networks**

Today, we can connect with anyone around the world, but security remains a major concern for mobile networks. Firewalls and other security measures are becoming less effective as people use a variety of devices, such as tablets, phones, and PCs, each requiring additional security beyond what applications provide. It is crucial to constantly consider the security challenges faced by these mobile networks. Given their vulnerability to cybercrime, extra caution must be exercised to address their security risks.

#### **4.5 IPv6: New internet protocol**

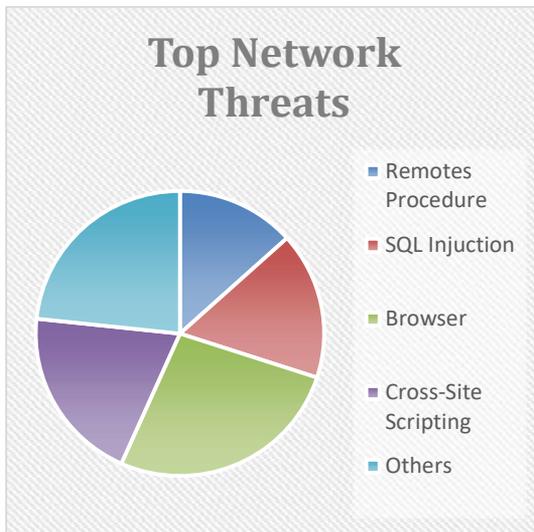
IPv6 is the new internet protocol set to replace IPv4, the longstanding backbone of networks and the internet. Protecting IPv6 involves more than simply transferring IPv4 security measures, as it introduces fundamental changes to the protocol that must be considered in security policies. While IPv6 expands the availability of IP addresses, these changes require careful attention. Therefore, transitioning to IPv6 as soon as possible is advisable to mitigate the risks associated with cybercrime.

#### **4.6 Encryption of the code**

Encryption is the process of encoding messages or information so that eavesdroppers or hackers cannot access or understand it. In an

encryption system, a message is transformed into unreadable ciphertext using an encryption algorithm, typically 6. CYBER SECURITY TECHNIQUES the help of an encryption key that dictates how the encoding is done. At its core, encryption safeguards data privacy and integrity. However, as encryption becomes more widespread, it also introduces new challenges in cybersecurity. Encryption is commonly used to secure data in transit, such as information transferred over networks (e.g., the internet, e-commerce), mobile phones, wireless microphones, and intercoms. By encrypting data, one can detect potential information leaks.

These trends are reshaping the landscape of cybersecurity worldwide. Below is a pie chart illustrating the top network threats.



The pie chart above highlights the major threats to networks and cybersecurity.

### 5. Role of Social Media In Cyber Security

As we become more social in an increasingly connected world, businesses must discover new methods to safeguard personal information. Social media significantly impacts cybersecurity and contributes to personal cyber threats. The use of social media among individuals is rapidly increasing, along with the risk of attacks. Since most people use social media or networking sites

daily, these platforms have become prime targets for cybercriminals looking to hack private information and steal valuable data

In a world where we readily share our personal information, companies must be equally swift in identifying threats, responding in real-time, and preventing any breaches. Because people are easily drawn to social media, hackers often use these platforms as bait to obtain the information and data they seek. Therefore, individuals must take appropriate precautions, especially when engaging with social media, to safeguard their information.

The ability for individuals to share information with millions poses a significant challenge for businesses. Social media not only empowers anyone to disseminate commercially sensitive information but also allows for the rapid spread of false information, which can be equally damaging. The quick circulation of misinformation via social media is highlighted as an emerging risk in the Global Risks 2013 report.

Although social media can be exploited for cybercrimes, companies cannot afford to stop using it, as it plays a crucial role in their marketing and visibility. Instead, they must implement solutions that can alert them to potential threats so they can address issues before significant damage occurs.

Companies should acknowledge this reality and understand the importance of analyzing information, particularly in social conversations. They need to provide appropriate security measures to mitigate risks effectively. Managing social media requires the adoption of specific policies and the right technologies.

## **6. Cyber Security Techniques**

### **6.1 Access control and password security**

The use of usernames and passwords has long been a foundational method for securing our information. It is one of the earliest and most basic measures in the realm of cybersecurity.

### **6.2 Authentication of data:**

Before downloading any documents, it is crucial to ensure they are authenticated, meaning they must be verified as originating from a trusted and reliable source and checked to ensure they haven't been altered. This authentication process is typically handled by antivirus software installed on devices. Therefore, having robust antivirus software is essential not only for verifying documents but also for protecting devices from viruses.

### **6.3 Malware scanners:**

This software typically scans all files and documents on the system for malicious code or harmful viruses. Examples of such malicious software, often collectively referred to as malware, include viruses, worms, and Trojan horses.

### **6.4 Firewalls:**

A firewall is a security system, either software or hardware, designed to filter and monitor incoming and outgoing traffic between your computer and the internet. It acts as a barrier, scrutinizing all data packets and allowing only those that meet predefined security rules to pass through. By blocking unauthorized access and harmful content like hackers, viruses, and worms, firewalls provide essential protection against potential cyber threats, helping to safeguard your system from malware and other malicious attacks.

### **5. Anti-virus software:**

Antivirus software is a program designed to detect, prevent, and remove malicious software like viruses, worms, and other threats. It continuously scans your system for harmful files and can automatically update to identify new and emerging threats. With real-time protection and regular updates, antivirus software ensures that your computer is safeguarded against the latest security risks. Installing antivirus software is essential for maintaining the safety and security of any system.

## **7. Cyber Ethics**

Cyber ethics refers to the principles and guidelines for responsible and safe use of the internet. By following these ethical practices, users can ensure a safer and more respectful online experience. Some key points include:

- ❖ Do use the internet to communicate positively and respectfully. Email and instant messaging allow for easy interaction with friends, family, colleagues, and people worldwide.
- ❖ Don't engage in cyberbullying. Avoid name-calling, spreading false information, sharing embarrassing images, or any actions intended to harm others online.
- ❖ The internet is a vast resource of information, often referred to as the world's largest library. Always use this information responsibly and legally.

- ❖ Never pretend to be someone else or create fake accounts, as this can cause problems for both you and the affected party.
- ❖ These cyber ethics help ensure responsible internet use, just like the rules and guidelines we are taught from an early age in other areas of life.

The following are essential cyber ethics to adhere to while using the internet. Just as we learn proper rules from an early age, the same principles apply in cyberspace.

## **8. CONCLUSION**

Computer security is a broad and increasingly critical topic as our world becomes more interconnected through networks that facilitate essential transactions. With each passing year, cybercrime evolves, presenting new challenges and threats to the security of information. Emerging technologies and sophisticated cyber tools are pushing organizations to rethink how they protect their infrastructure, requiring advanced platforms and intelligence to stay ahead of these threats. While there is no perfect solution to eliminate cybercrime entirely, continuous efforts to strengthen security measures are essential to reduce risks and ensure a safer future in cyberspace.

## **References**

1. Cyber Security Understanding Cyber Crimes- Sunit Belapure Nina Godbole
2. Computer Security Practices in Non-Profit Organisations – A NetAction Report By Audrie Krause.
3. A Look back on Cyber Security 2012 by Luis corróns – Panda Labs.
4. International Journal of Scientific & Engineering Research, Volume 4, Issue 9, September 2013 Page nos.68 – 71 ISSN 2229-5518, “Study of Cloud Computing in HealthCare Industry “ by G.Nikhita Reddy, G.J.Ugander Reddy.
5. IEEE Security and Privacy Magazine –IEEECS “Safety Critical Systems – Next Generation “July/ Aug 2013.
6. CIO Asia, September 3<sup>r</sup>, H1 2013: Cyber security in Malasia by Avanthi Kumar

# Navigating the Full Stack : Unlocking Future Possibilities and Overcoming Challenges in Web Development

Nageswaran M.K<sup>1</sup> Deepika S<sup>2</sup> Dinesh J<sup>3</sup>Vignesh  
M<sup>3</sup>Hariprakash K<sup>3</sup>Raveendar A<sup>3</sup>

*<sup>1</sup>Assistant Professor, MCA, Excel Engineering College, Komarapalayam, Namakkal, India*

*<sup>2</sup> Student of I - MCA, Excel Engineering College, Komarapalayam, Namakkal, India*

*<sup>3</sup> Student of II - MCA, Excel Engineering College, Komarapalayam, Namakkal, India*

## **Abstract**

In the rapidly evolving landscape of technology, full-stack development has emerged as a pivotal role in web development, combining both front-end and back-end expertise to create comprehensive web applications. This paper explores the multifaceted nature of full-stack development, examining core technologies, current trends, and the challenges developers face in this dynamic environment. By analyzing the evolution of web technologies and identifying the essential skills required for success, we aim to provide insights into the competencies that aspiring full-stack developers must cultivate. Furthermore, this study investigates the impact of emerging technologies, such as artificial intelligence, blockchain, and low-code

platforms, on the future of full-stack development. Through case studies of successful and unsuccessful projects, we highlight best practices and strategies for navigating complexities in development workflows. Ultimately, this paper serves as a guide for both current and future developers, emphasizing the importance of adaptability and continuous learning in an era marked by rapid technological advancements.

**Keywords:** Full Stack Development, Web Development, Front-End Technologies, Back-End Technologies, ETC...

## I. Introduction

The advent of the digital age has transformed the way we interact, conduct business, and share information. At the heart of this transformation lies web development, a field that has evolved rapidly to meet the growing demands of users and businesses alike. Full stack development, which encompasses both front-end and back-end technologies, has gained prominence as a vital component of modern web applications. Full-stack developers possess a unique skill set that allows them to navigate the entire development process, from crafting intuitive user interfaces to managing server-side logic and database interactions.

In recent years, web development has evolved considerably due to rapid technological progress and changing user demands. The proliferation of frameworks, libraries, and tools has made it easier for developers to create complex applications, yet it has also introduced challenges related to managing complexity and ensuring seamless integration between different layers of the technology stack. As new trends, such as Progressive Web Apps (PWAs), serverless architecture, and the rise of headless content management systems, emerge, full-stack developers must continuously adapt to remain relevant in an ever-changing environment.

This paper aims to explore the intricate nature of full-stack development, highlighting the core technologies that define this discipline and examining the current trends shaping its future.

Additionally, we will delve into the challenges that developers face in balancing the demands of front-end and back-end development, the skills required for success in this role, and the importance of continuous learning. By analyzing case studies of both successful and unsuccessful projects, we will identify best

practices that can guide developers in navigating the complexities of full-stack development.

Ultimately, this study seeks to unlock future possibilities within the realm of web development while providing insights into overcoming the challenges that developers encounter. As we embark on this exploration, we aim to provide a comprehensive understanding of full-stack development, equipping aspiring developers with the knowledge and strategies needed to thrive in this dynamic field.

## **II. Core Technologies in Full Stack Development**

Full stack development encompasses a wide range of technologies that cover both the client-side (front-end) and server-side (back-end) of web applications. A deep understanding of these core technologies is crucial for developers to create scalable, efficient, and resilient applications. This section explores the essential tools and frameworks that are integral to full-stack development.

**1. Front-End Technologies:** The front end represents the interface that users interact with directly, responsible for both the user interface (UI) and user experience (UX). The following are key front-end technologies:

- a) HTML (HyperTextMarkup Language):** HTML serves as the backbone of web content structure. It defines the elements of a webpage, including headers, paragraphs, links, images, and other content components.
- b) CSS (Cascading Style Sheets):** CSS is used to style and layout web pages. It allows developers to apply visual elements like colors, fonts, and spacing, ensuring a visually appealing and consistent user interface.
- c) JavaScript :** JavaScript enables dynamic interactions on web pages. It supports responsive designs, allows manipulation of the DOM (Document Object Model), and facilitates communication between the front end and back-end services via APIs.
- d) Front-End Frameworks and Libraries :** Several frameworks and libraries simplify front-end development by providing pre-built components and functionalities. Popular choices include:
- e) React:**

A JavaScript library for building user interfaces, particularly single-page applications (SPAs).

- f) **Angular:**A full-fledged framework for building dynamic web applications with a strong focus on structure and scalability.
- g) **Vue.js:**A progressive framework for building user interfaces, known for its ease of integration and flexibility.

## 2. Back-End Technologies:

The back end of a web application is responsible for server-side logic, database interactions, and application performance. Core back-end technologies include:

**Server-Side Programming Languages:** These languages handle requests from the front end and process data. Common server-side languages include:

- a) **Node.js:**A JavaScript runtime built on Chrome's V8 engine, allowing developers to use JavaScript for server-side scripting.
- b) **Python:**Known for its readability and versatility, Python is often used in web development with frameworks like Django and Flask.
- c) **Ruby:**A dynamic programming language used primarily with the Ruby on Rails framework for building web applications quickly.
- d) **Databases:** Databases store and manage data for web applications. They can be categorized into:
  - e) **Relational Databases:** Structured databases that use SQL (Structured Query Language) for data manipulation. Examples include MySQL and PostgreSQL.
  - f) **NoSQL Databases:** Unstructured databases designed to handle large volumes of data and high-velocity transactions. Examples include MongoDB and Cassandra.
- g) **Back-End Frameworks:** These frameworks provide tools and libraries to streamline back-end development. Popular frameworks include:
  - h) **Express.js:** A minimal and flexible Node.js web application framework for building APIs and web applications.
  - i) **Django:** A high-level Python framework that encourages rapid development and clean, pragmatic design.
  - j) **Ruby on Rails:** A server-side web application framework written in Ruby that emphasizes convention over configuration.

## 3. DevOps and Deployment Technologies

In addition to front-end and back-end technologies, full-stack development involves deploying and maintaining applications. Key technologies in this area include:

- a) **Version Control Systems:** Tools like Git are essential for tracking changes in code and collaborating with other developers.
- b) **Continuous Integration/Continuous Deployment (CI/CD):** Tools like Jenkins, GitHub Actions, and CircleCI automate the process of testing and deploying applications.
- c) **Containerization and Orchestration:** Technologies such as Docker and Kubernetes help package applications and manage their deployment across different environments.
- d) **Cloud Services:** Cloud platforms like AWS, Google Cloud, and Microsoft Azure provide scalable infrastructure for hosting applications and services.

### III. The Evolution of Full Stack Development

#### Historical Context and Advancements:

Web development has evolved substantially since the early days of the World Wide Web in the 1990s. Initially, websites were static and simple, structured exclusively with HTML to display content. While functional, this early approach lacked flexibility and visual sophistication. The introduction of CSS in the late 1990s revolutionized web design, allowing for the separation of content from its presentation and making websites more visually dynamic. The emergence of JavaScript in the early 2000s marked another pivotal moment, as it enabled real-time interactivity and the development of dynamic, user-centric applications that fundamentally transformed the web user experience.

The mid-2000s saw a shift in back-end development with the rise of web frameworks such as Ruby on Rails and Django, which introduced a more structured and efficient approach to building web applications. These frameworks popularized the idea of rapid application development by emphasizing convention over configuration, which helped developers focus on functionality instead of reinventing the wheel with repetitive code. On the client side, libraries like jQuery simplified JavaScript's cross-browser challenges and introduced powerful, pre-built components for creating more interactive user experiences without requiring extensive programming expertise.

### **b) Impact of Emerging Technologies:**

In recent years, full-stack development has been further transformed by emerging technologies that prioritize modularity, automation, and scalability. **Serverless architecture** is one of the most significant innovations, allowing developers to deploy applications without the overhead of managing servers. This approach not only reduces operational complexity but also allows developers to scale applications easily and pay only for the actual computing resources they use, thus minimizing infrastructure costs. Serverless models enable developers to concentrate on writing core application logic rather than managing underlying infrastructure.

Another game-changing development is the adoption of **microservices architecture**, where applications are broken down into smaller, autonomous services. Each microservice can operate, scale, and be updated independently, enhancing the agility of web applications. This approach supports continuous integration and delivery (CI/CD), making it easier to maintain and deploy updates. Microservices also facilitate the use of different programming languages and tools for each service, allowing for greater flexibility in addressing specific technical challenges. Together, serverless and microservices architectures offer full-stack developers highly scalable and maintainable approaches to building next-generation web applications that are resilient, adaptive, and ready for rapid innovation.

## **IV. Current Trends and Innovations:**

- a) **Rise of JAMstack and Headless CMS:** JAMstack (JavaScript, APIs, and Markup) has gained traction as a modern architecture for building fast, secure, and scalable web applications. By decoupling the front end from the back end, JAMstack allows developers to leverage static site generators and pre-built markup, enhancing performance and reducing load times. Headless CMS solutions further complement this trend by providing a backend content management system that delivers content via APIs, enabling seamless integration with various front-end frameworks and delivering content across multiple platforms.
- b) **Progressive Web Apps (PWAs) and Their Significance:** Progressive Web Apps (PWAs) represent a significant innovation in web development, combining the best features

of web and mobile applications. PWAs offer offline capabilities, push notifications, and fast load times, resulting in improved user engagement and experience. By leveraging service workers and responsive design, PWAs can operate on any device, enhancing accessibility and reach for businesses and developers alike.

- c) **Importance of API-Driven Development:** API-driven development has become a cornerstone of modern web applications, facilitating the integration of disparate services and enabling seamless data exchange. By employing RESTful APIs or GraphQL, developers can create flexible architectures that support microservices and modular design. This approach allows for rapid iteration, scalability, and the ability to adapt to changing business requirements, making API-driven development essential for contemporary full-stack applications.

## V. Challenges Faced by Full-Stack Developer:

Full-stack developers encounter several challenges in their multifaceted roles:

- a) **Complexity Management:** The breadth of technologies and frameworks required for both front-end and back-end development can be overwhelming, leading to difficulties in maintaining code quality and project organization.
- b) **Rapid Technological Changes:** The fast-paced evolution of tools and frameworks necessitates continuous learning and adaptation, making it challenging for developers to stay up-to-date and relevant.
- c) **Balancing Responsibilities:** Juggling the demands of front-end and back-end development can lead to a lack of depth in either area, making it difficult for developers to specialize or master specific technologies.
- d) **Integration Issues:** Ensuring seamless communication between various components, such as APIs and databases, often results in integration challenges, particularly when utilizing different technologies.
- e) **Performance Optimization:** Developing applications that are both scalable and efficient requires a deep understanding of

system architecture and performance optimization techniques, which can be complex to implement.

## **VI. Skills and Competencies Required in Full Stack Development:**

Full-stack developers are expected to possess a blend of technical skills and soft skills to navigate the complexities of web development effectively. This section outlines the essential skills for full-stack developers, the role of soft skills in their success, and the importance of continuous learning and professional development.

### **1. Essential Skills for Full-Stack Developers:**

#### **a) Front-End Proficiency:**

A strong grasp of HTML, CSS, and JavaScript is foundational. Familiarity with modern front-end frameworks like React, Angular, or Vue.js is crucial for building responsive and dynamic user interfaces.

**b) Back-End Expertise:** Knowledge of server-side programming languages such as Node.js, Python, Ruby, or Java, along with experience in back-end frameworks (e.g., Express.js, Django, Ruby on Rails), enables developers to handle application logic and data management.

**a) Database Skills:** Competence in working with relational databases (like MySQL or PostgreSQL) and NoSQL databases (such as MongoDB) is necessary for efficient data storage, retrieval, and manipulation.

**b) API Development and Integration:** Understanding how to create and consume RESTful APIs and GraphQL is vital for enabling communication between front-end and back-end systems.

**c) Version Control:** Proficiency in Git is essential for tracking changes, collaborating with team members, and maintaining code integrity throughout the development process.

#### **d) Deployment and DevOps Knowledge:**

Familiarity with deployment processes, CI/CD practices, and cloud platforms (e.g., AWS, Azure) aids in managing application deployment and scalability.

### **2. The Role of Soft Skills in Web Development:**

While technical skills are critical, soft skills play an equally important role in the success of full-stack developers:

**a) Communication:** Effective communication skills are necessary for collaborating with team members, stakeholders, and clients,

ensuring that project requirements and expectations are clearly understood.

- b) **Problem-solving:** Strong analytical and problem-solving abilities enable developers to troubleshoot issues, optimize performance, and devise innovative solutions to complex challenges.
- c) **Adaptability:** The ability to adapt to changing technologies, project requirements, and team dynamics is essential in the fast-paced world of web development.
- d) **Teamwork:** Full-stack developers often work in collaborative environments; therefore, the ability to work effectively within a team is crucial for achieving project goals.

### 3. **Continuous Learning and Professional Development:**

Given the rapid evolution of technology, continuous learning and professional development are imperative for full-stack developers:

**Staying Updated:** Developers must keep abreast of emerging technologies, frameworks, and best practices to remain competitive and relevant in the field.

- a) **Online Resources and Communities:** Utilizing online courses, coding bootcamps, and developer communities (such as GitHub, Stack Overflow, and tech forums) can provide valuable learning opportunities and networking.
- b) **Certifications:** Pursuing certifications in relevant technologies (e.g., AWS Certified Developer, Google Professional Cloud Developer) can enhance credibility and open doors to new career opportunities.
- c) **Practical Experience:** Engaging in personal projects, contributing to open-source initiatives, or participating in hackathons can help developers hone their skills and gain practical experience.

## VII. **Future Possibilities in Full Stack Development:**

### **The Role of Artificial Intelligence and Machine Learning:**

The integration of artificial intelligence (AI) and machine learning (ML) into full-stack development is poised to revolutionize web applications. These technologies can enhance user experience through personalized content recommendations, automate customer support with chatbots, and improve decision-making processes by analyzing user data. As developers adopt AI-driven tools and frameworks, they will be able to create smarter, more responsive applications that cater to individual user needs.

## **Exploring the Impact of Blockchain Technology on Modern Web Development:**

Blockchain technology offers promising possibilities for enhancing security, transparency, and trust in web applications. By enabling decentralized data storage and smart contracts, blockchain can streamline processes such as payments, identity verification, and supply chain management. Full-stack developers will increasingly leverage blockchain to build applications that prioritize data integrity and enhance user trust, particularly in sectors like finance, healthcare, and e-commerce.

### **The Future of Low-Code/No-Code Platforms:**

The rise of low-code and no-code platforms is transforming the landscape of web development by allowing users to create applications with minimal coding knowledge. These platforms empower non-developers to participate in the development process, fostering innovation and accelerating project timelines. Full-stack developers will need to adapt by integrating these tools into their workflows, focusing on customization and scalability while enhancing collaboration with business stakeholders.

### **Future Trends in Full Stack Development:**

**Increased Adoption of AI and Machine Learning:** Full-stack developers will increasingly integrate AI and ML technologies to create smarter applications that enhance user experiences through personalized content, predictive analytics, and automated processes.

**Growth of Serverless Architecture:** The shift towards serverless computing will allow developers to build and deploy applications without the need for traditional server management, improving scalability and reducing operational costs.

**Expansion of Low-Code/No-Code Platforms:** These platforms will empower non-technical users to participate in application development, enabling faster prototyping and reducing the dependency on specialized coding skills.

**Focus on API-First Development:** As applications become more interconnected, the API-first approach will gain prominence, allowing for easier integration between services and promoting modular design.

**Emphasis on Cybersecurity:** With the rise in cyber threats, full-stack developers will prioritize security in the development process, implementing robust security measures and best practices to protect user data.

**Adoption of Progressive Web Apps (PWAs):** The trend towards PWAs will continue, as they offer enhanced performance, offline capabilities, and a seamless user experience across devices.

**Integration of Blockchain Technology:** Full-stack developers will explore blockchain for applications requiring enhanced security, transparency, and decentralization, particularly in finance and supply chain management.

**Collaboration and Remote Development Tools:** The growth of remote work will drive the adoption of collaborative tools that facilitate communication and project management among distributed development teams.

## **VIII. Strategic Recommendations for Overcoming Challenges:**

### **1. Effective Approaches for Lifelong Learning and Skill Enhancement:**

**Adopt a Growth Mindset:** Embrace lifelong learning by seeking out new technologies and frameworks through online courses, workshops, and tutorials.

**Engage in Community Learning:** Participate in developer communities and forums to share knowledge, ask questions, and stay updated on industry trends.

**Hands-On Projects:** Build personal or open-source projects to apply new skills in real-world scenarios, enhancing both understanding and portfolio.

### **Conclusion:**

#### **The Future of Full Stack Development:**

As we look to the future, full-stack development stands at a crossroads, poised to adapt to the rapidly changing technological landscape. The advent of emerging technologies such as artificial intelligence, blockchain, and cloud computing is transforming how web applications are built and deployed, presenting both exciting opportunities and significant challenges for developers.

To thrive in this evolving environment, full-stack developers must embrace a culture of continuous learning, actively seek to enhance their skill sets and foster collaborative practices that promote innovation. By adopting agile methodologies and modular approaches, they can effectively manage the complexity of modern web applications, ensuring they remain responsive to user needs and market demands.

Ultimately, the future of full-stack development is not merely about mastering a diverse array of technologies but about cultivating a mindset that prioritizes adaptability, collaboration, and user-centric design. As developers navigate these changes, they will play a crucial role in shaping the next generation of web applications, unlocking unprecedented possibilities for businesses and users alike.

## References

- [1] Valerio De Sanctis, Full Stack Web Development Third Edition, 2020 book
- [2] Chris Northwood, Full Stack Developer: Your Essential Guide to the Everyday Skills Expected of a Modern Full Stack Developer
- [3] Frank Zammeti, Modern Full Stack Development: Using Typescript, React, Node.js, Webpack, and Docker Full
- [4] Juha Hinkula, Hands-on- Full Stack Development with Spring Boot 2 and React, Second Edition
- [5] Hassan Djirdeh, Nate Murray, Ari Lerner, Full Stack Vue: The Complete Guide to Vue.js
- [6] Nader Dabit, Full Stack Serverless, Modern Application Development with React, AWS, and GraphQL
- [7] Dr. Balwinder Kaur Dhaliwal, Gurpreet Singh, "Use of Clustering with CNN to Categorize image inside different directories on OS" - Electronic ISBN: 978-1- 6654-1703-7 Published in "2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)" (ICRITO) 3-4 Sept, 2021
- [8] G.S Shahi, R.S Batth, S. Egerton, 2020 "MRGM: An Adaptive Mechanism for Congestion Control in Smart Vehicular Network", published in "International Journal of Communication Networks and Information Security 12".
- [9] Balwinder Kaur Dhaliwal, Rattan K. Datta, "Secure and Energy Efficient Trust Aware Routing Protocol in IoT using the Optimized Artificial Neural Network: SEETA-IoT" published in 'International Journal of Engineering and Advanced Technology (IJEAT)', ISSN: 2249-8958 (Online), Volume-8 Issue-6, August 2019, Page No.:4341-4353.

- [10] Balwinder Kaur Dhaliwal, Rattan K Datta,” A Trust Aware Based Predictive Model using Hybrid Convolutional Neural Network for Mobile AD HOC Network in Internet of Things” published in ‘International Journal of Engineering and Advanced Technology (IJEAT)’, ISSN: 2249-8958 (Online), Volume-9 Issue-1, October 2019, Page No.1276- 1285

# Exploring Augmented and Virtual Reality: From Innovation to Everyday Use

Nageswaran M.K<sup>1</sup> Akash B<sup>2</sup> Mageshraj M<sup>2</sup> Tamilarasan J<sup>2</sup>  
Aadhinakaran S<sup>2</sup> Vinoth S<sup>2</sup>

*<sup>1</sup>Assistant Professor, MCA, Excel Engineering College, Komarapalayam,  
Namakkal, India*

*<sup>2</sup> Student of I - MCA, Excel Engineering College, Komarapalayam,  
Namakkal, India*

## **Abstract**

Reality is experienced differently through software environments, particularly in Virtual Reality (VR) and Augmented Reality (AR). These technologies have gained significant traction, initially within the entertainment sector, but their applications have expanded across various fields addressing real-world challenges. This article provides an overview of VR and AR systems and their primary application domains.

## **1. Introduction**

Augmented Reality (AR) and Virtual Reality (VR) technologies are becoming increasingly prevalent as advancements in hardware and software make them accessible for diverse applications. These technologies enhance user experiences in learning, training, simulation, gaming, and various other tasks. The motivations for employing AR/VR technologies differ based on specific applications; however, they predominantly aim to provide users with additional information

that cannot be perceived through conventional sensory experiences. Given their potential to address various challenges, leading companies like Google, IBM, Sony, and HP, alongside numerous universities, are actively developing AR/VR technologies. These systems are applicable across various disciplines, including physics, chemistry, biology, mathematics, history, astronomy, medicine, and even music. Such innovations aim to create technological devices that adapt to these fields and ultimately enhance users' lives.

This paper introduces the concepts of Augmented and Virtual Reality, outlining the methodologies utilized within these frameworks. AR/VR technology overlays 3D virtual objects into the user's real-world environment in real-time. We analyze the technical requirements necessary to provide users with the optimal AR/VR experience.

Furthermore, we explore the specific characteristics of various domains and how AR/VR systems interact with them. The purpose of this overview is to present the current state of augmented and virtual reality.

The first section of this paper focuses on defining key terminologies, while the second section examines the diverse applications of VR and AR systems across various industrial sectors. The third section addresses the limitations of VR/AR devices.

## **2. Definitions**

### **2.1 Virtual Reality**

Virtual reality (VR) can be broadly defined as the simulation of a virtual object within a virtual environment. More precisely, it refers to a computer-generated recreation of real-life environments or situations that immerses users, allowing them to feel as though they are experiencing simulated reality firsthand, primarily through visual and auditory stimulation. An effective virtual reality system, often referred to as a "virtual environment" (VE), enables users to physically navigate around and interact with objects as if they were real, creating an immersive experience that engages one or more of the user's senses.

Ivan Sutherland, a pioneer in VR technology, famously stated: "The ultimate display would, of course, be a room in which the computer can control the existence of matter. A chair in such a room would be good enough to sit down on. The handcuffs in such a room would be credible, and a ball displayed in such a room would be fatal". To refine this definition, we note that "virtual" refers to an essence or

effect rather than reality itself. Reality is something that exists independently of perception, while virtual reality encompasses an environment created by computers that users can interact with as though it were real.

In essence, virtual reality consists of worlds that exist solely within computers and the minds of users. A robust VR system allows for physical navigation and interaction with virtual objects, simulating real-world experiences. VR can be classified into two categories: non-immersive and immersive. Non-immersive VR presents a computer simulation of the real world, while immersive VR enhances the experience through greater immersion, interactivity, and user engagement, often utilizing head-mounted displays that fully detach the user from their actual surroundings.

## **2.2 Augmented Reality**

The term "augmented reality" was first introduced at Boeing in 1990 by researcher Tom Caudell, who sought to improve the costly labeling systems used to guide employees through the plant. He proposed replacing large plywood panels containing aircraft wiring instructions with a head-mounted device that displayed specific diagrams through advanced glasses, projecting them onto multi-purpose, reusable panels. Since then, AR has evolved to enhance user experiences across various tasks, as it overlays computer-generated enhancements onto an existing reality, making it more interactive and realistic. Essentially, augmented reality combines real-world environments with virtual elements.

AR allows users to perceive their surroundings in real time, enriched with virtual objects. These digital elements are integrated into the physical world using portable devices, creating a seamless interaction between the real and the virtual. For instance, a user might see a virtual glass appearing alongside a physical glass on a table. A key aspect of AR is that virtual objects are designed to appear as real, solid, and credible as their physical counterparts.

As researchers continue to expand the definition of AR, it is increasingly recognized as a dynamic area of interactive design, particularly with the rise of personal mobile devices capable of creating engaging AR experiences.

To summarize the characteristics of an AR system:

- It combines real and virtual objects within a real environment.

- It operates in real time and allows for interactive experiences.
- Real and virtual objects are aligned within the same space.

Ultimately, the distinction between VR and AR lies in their environments: AR enhances the real world with virtual elements, while VR creates an entirely artificial setting.

### **2.3 Mixed Reality**

Mixed Reality (MR) is defined as a hybrid reality in which real and virtual worlds merge to create new environments and visualizations that coexist and interact in real-time with both physical and digital objects. MR encompasses elements of both augmented reality (AR) and virtual reality (VR), resulting in a combination that incorporates aspects of augmented reality and augmented virtuality. This technology enables the digital rendering of objects within real environments, a process also referred to as holography.

In a mixed reality setting, real and virtual content coexist and interact in real-time, blending features of both AR and VR to enrich user experiences. MR is not merely an alternative to AR or VR; rather, it offers a distinct perspective that enhances users' perceptions within both real and virtual spaces. Key aspects of a mixed reality experience include flexibility, immersion, interaction, coexistence, and enrichment, all achieved by leveraging the technological foundations of AR and VR. Consequently, an MR experience creates an environment where users feel fully immersed, and their perception of the real world is enhanced.

### **2.4 Virtual Reality Continuum**

The reality-virtuality continuum consists of a spectrum of environments ranging from completely real to entirely virtual, including all possible variations and combinations of real and virtual objects within these spaces. The similarities in underlying technologies have often led to confusion between AR and VR, with many considering the two to be identical. This misunderstanding is addressed by the Reality-Virtuality Continuum proposed by Milgram in 1994. Milgram and Kishino introduced this continuum to define mixed reality and identify various technologically modified forms of reality that correspond to contemporary augmented and virtual reality technologies.

## **2.5 Augmented Virtuality**

Augmented Virtuality (AV) refers to the interactive exploration of virtual representations derived from the real world. It can be viewed as a subcategory of Mixed Reality (MR) that integrates real-world objects into the virtual environment. This integration is often achieved by streaming video from physical spaces (e.g., via webcams) or using 3D scanning techniques to capture physical objects, which are then draped with virtual elements. This approach creates the illusion that the virtual world closely resembles the real one while retaining the flexibility inherent in virtual environments. For instance, an aircraft maintenance engineer might visualize a real-time model of an aircraft engine on a screen, overlaid with real-world elements that are physically separated.

## **2.6 Comparison**

While augmented reality, virtual reality, and mixed reality can all be leveraged to achieve similar objectives, our survey indicates that Augmented Reality is particularly effective for enhancing exhibitions. Conversely, Virtual Reality is better suited for virtual museums, and Mixed Reality shows greater promise for reconstruction applications both indoors and outdoors.

We propose the following concise definitions:

**Augmented Reality (AR):** Enhances our perception and understanding of the real world by superimposing virtual information onto our view of reality.

**Virtual Reality (VR):** Immerses users in a computer-generated environment that prevents interaction with or visibility of the real world.

**Augmented Virtuality (AV):** Enriches the virtual environment by incorporating real-world scenes.

**Mixed Reality (MR):** Combines real and virtual environments to create a hybrid experience.

## **3. Industrial Applications Of Virtual/Augmented Reality**

According to, manufacturing is defined as "the transformation of materials and information into goods and services to satisfy human needs." Recent advancements in digital manufacturing are reducing development time and costs while addressing the demands for personalization, improved quality, and rapid mass distribution. Industries such as aerospace and automotive exemplify the need for simulation and the key challenges associated with product design and

process optimization. Thus, this section focuses on manufacturing, though many applications can be adapted to other sectors, including military, construction, medical, architecture, and commerce.

Inspired by Murtis etc., we classify industrial AR/VR applications into four categories:

- Training applications
- Assistance requests
- Design applications
- Planning and validation requests

### **3.1 Training Applications**

Training is a cross-cutting concern across various application areas, and we will focus specifically on AR/VR applications for training purposes. For instance, researchers have utilized mobile devices like smartphones and tablets to assist workers in assembly processes. Aehnelt and Bader employed external displays to provide assembly instructions, integrating contextual background information to teach, guide, and monitor assembly workers.

In the field of training, Aehnelt and Wegner investigated the integration of work and learning experiences to enhance cognitive understanding and task processing, enabling workers to effectively manage their tasks in a smartmanufacturing environment. After training participants with VR, tablet-based VR, Head-Mounted Display Augmented Reality (HMD-AR), or a paper manual as a baseline, they measured assembly completion times. Their findings revealed that HMD-based AR training yielded the fastest completion times, followed by various virtual reality methods and the paper manual.

AR/VR training applications offer several advantages over traditional training methods, including:

- The ability to execute tasks without risk
- Opportunities to make mistakes without compromising safety
- The capability to reconfigure environments (e.g., terrain, weather)
- Modelling of inaccessible training scenarios (e.g., space, enemy terrain)
- Simulation of situations that cannot be replicated in reality (e.g., technical incidents)
- Flexibility from time constraints and other limitations

- Efficient use of limited space (i.e., the volume of the AR/VR system)
- The versatility of using the same system for various training applications.

### **3.2 Requests for Assistance**

Operator assistance through guidance tools is typically employed for various manual tasks, such as assembling or disassembling objects. One straightforward method of assisting is the use of standard manuals as construction guides. To enhance worker productivity, new techniques have been explored to support operators in maintenance and repair tasks, leading to the integration of AR/VR technologies in several industrial sectors.

For instance, Echtler et al. introduced the Smart Welding Gun, which helps users precisely pull studs on prototype vehicles. This intelligent welding gun is designed with an instrument-based approach, featuring a display that provides real-time feedback on the task being performed. An infrared tracking system (ART) monitors the gun and its components to ensure accuracy.

In the medical field, where operations can be critical and demanding for surgeons, computer-aided techniques have been developed to assist during surgeries.

For example, Gavaghan et al. created a portable handheld projection system that utilizes a laser projector to overlay essential information on the liver's surface, aiding surgical procedures.

### **3.3 Design Applications**

Design processes can occur at various stages throughout a product's lifecycle. Virtual and Augmented Reality has become invaluable in product design, particularly VR, which provides designers with immersive experiences that enhance the design process. Early VR systems were specifically developed for design purposes; for example, the 3DM, created in 1992, adapted Computer-Aided Design (CAD) and drawing programs for VR Head-Mounted Displays (HMDs). The COVIRDS method [17] integrates CAD modeling, user interface design, and VR technologies.

Numerous AR/VR design applications have been proposed, focusing on creating a comprehensive design process that utilizes both AR and VR. These applications offer tools for designing and editing 3D curves and surfaces. For early conceptual design, Israel

et al. introduced 3D sketching techniques and conducted studies comparing the effectiveness of 3D sketching to traditional 2D methods. Following this, De Araújo et al. proposed two-handed interaction techniques for sketching and designing objects on flat surfaces, such as workbenches. However, workbench systems provide limited immersion and are primarily suited for smaller-scale designs.

To create larger objects, users benefit from immersive projection-based systems. For instance, Hughes et al. developed CaveCAD, a virtual reality architectural design application for immersive environments displayed on a CAVE system. They implemented various interaction techniques and functionalities for modifying object geometry, noting that prolonged use of handheld controls can lead to arm fatigue compared to traditional desktop monitors. Similarly, the Sculpt Up system proposed by Ponto et al. [21] offers an alternative approach to designing objects in VR CAVE systems.

### **3.4 Planning and Validation of Requests**

Among the various AR/VR systems, VR has been extensively utilized for plant layout planning. Iqbal and Hashmi [22] were pioneers in applying virtual environments for plant planning and proposing alternative layout solutions. Similarly, Calderon et al. [23] developed an online planning tool that enables users to explore various planning options, though often lacking immersive screen integration. Menck et al. [24] introduced collaborative virtual environments for plant layout planning, allowing simultaneous viewing, investigation, and analysis of plant plans.

For validation, De Sa and Zachmann [25] integrated validation tools within virtual reality to analyze and confirm assembly and maintenance tasks using head-mounted displays (HMDs) in the automotive industry. Their study highlighted the effectiveness of VR for virtual prototyping in this sector. Additionally, PSA Peugeot Citroën's virtual vehicle project features a CAVE display, enabling project teams to interactively validate designs and plan vehicle assembly. For AR validation, Caruso and Re [26] developed an AR design review system based on an HMD with Video See-Through (VST) capabilities, facilitating visualization and interaction with virtual objects during the prototyping process.

### **3.5 Comparison**

The growing need to simulate various industrial operations has driven the adoption of AR/VR technologies across multiple applications. Based on our research, we categorize the appropriate utilization of AR/VR technologies as follows:

**Immersive Environments:** Virtual Reality (VR) is particularly well-suited for training simulations, with studies indicating that training in VR environments enhances operator performance compared to traditional training sessions. VR is also effective for designing objects or processes using imported CAD models and 3D drawing interactions.

**Planning and Validation:** Projection-based VR systems, such as the CAVE system, offer significant advantages for planning and validation, allowing entire project teams to collaboratively plan and validate products and operations interactively.

#### **4. Limitations of Ar/Vr Devices**

While advancements in AR/VR technology have been significant, several challenges remain that hinder the widespread adoption of these technologies among consumers and businesses. In [2], key issues are highlighted, and we summarize the main limitations of AR/VR technology as follows:

1. **Dedicated Hardware Requirements:** One of the primary obstacles is the need for specialized hardware. VR applications often require specific spatial setups and environmental conditions, which can limit their accessibility and usability.
2. **Need for Affordable Technology:** The high cost associated with AR/VR products is another barrier to adoption. Many consumers are hesitant to invest in these technologies due to their price tags. There is a pressing need for more affordable solutions that utilize cost-effective yet powerful hardware. Our previous research [27]–[29] explores various VR/AR solutions using less expensive equipment to address this issue.
3. **Lack of Legal Use Cases:** Even if cost barriers are addressed, AR/VR technologies face challenges related to innovative content. Current research and development efforts are predominantly focused on gaming and entertainment. While several use cases exist, the identification of applications that would significantly enhance the value of these technologies for consumers and businesses remains limited.

4. **Mobility and Miniaturization Issues:** Mobility poses a significant challenge for VR experiences. Many VR devices, such as Head-Mounted Displays (HMDs), are tethered by multiple cords, restricting user movement. There is a need for advancements in VR products that are miniaturized, lightweight, and portable, providing users with a wireless and more convenient experience.
5. **Cybersecurity Concerns:** Cybersecurity remains a critical concern for AR/VR technologies, with potential threats capable of compromising virtual environments. Hackers could access, alter, or damage these spaces, underscoring the need for robust security measures in future developments.

The table below provides a comparative analysis of existing AR/VR equipment and its applications.

Table 1: Comparative analysis of current products for AR/VR Product

<b>Key Features</b>	<b>Target</b>	<b>Audience</b>
Oculus Quest 2	Standalone, Wireless, High Resolution	Standalone, Wireless, High Resolution
Microsoft HoloLens 2	Mixed Reality, Gesture Control	Enterprises
HTC Vive Pro	High Fidelity, Room-Scale VR	Gamers/Developers
Magic Leap 1	Spatial Computing, Lightweight	Developers/Enterprises
Sony PlayStation VR	Console Integration, Affordable	Gamers
Samsung Galaxy HMD	Mobile-Based, Portable	Casual Users
VIVE Focus 3	All-in-One, High-Resolution Display	Business/Training

## 5. Conclusion

This overview has highlighted the advancements in AR and VR technologies, focusing on both the underlying mechanisms and their practical applications. While significant progress has already been made, the potential for further development is vast. Just as

computers and smartphones transformed everyday life, wearable devices incorporating AR and VR are poised to have a similarly profound impact. The future of this technology lies in creating seamless AR/VR experiences that are user-friendly and accessible to a broader audience. As innovation continues, we can anticipate a new era of immersive interactions that will reshape how we engage with both digital and physical worlds.

## References

- [1] M. Mendes, J. Almeida, H. Mohamed, and R. Giot, **Projected Augmented Reality Intelligent Model of a City Area with Path Optimization**, Algorithms, vol. 12, no. 7, p. 140, 2019, doi: 10.3390/a12070140.
- [2] S. Kumari and N. Polke, **Implementation Issues of Augmented Reality and Virtual Reality: A Survey**, vol. 26. Springer International Publishing, 2019.
- [3] W. Piekarski, B. Gunther, and B. Thomas, **Integrating virtual and augmented realities in an outdoor application**, Proc. - 2nd IEEE ACM Int. Work. Augment. Reality, IWAR 1999, no. February 1999, pp. 45–54, 1999, doi 10.1109/IWAR.1999.803805.
- [4] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, and J. Gain, **A survey of augmented, virtual, and mixed reality for cultural heritage**, J. Comput. Cult. Herit., vol. 11, no. 2, 2018, doi: 10.1145/3145534.
- [5] M. Mekni and A. Lemieux, **Augmented Reality: Applications, Challenges and Future Trends**, Appl. Comput. Sci. anywhere, pp. 205–214, 2014.
- [6] F. Ramos, S. Trilles, J. Torres-Sospedra, and F. J. Perales, **New trends in using augmented reality apps for smart city contexts**, ISPRS Int. J. Geo-Information, vol. 7, no. 12, 2018, doi: 10.3390/ijgi7120478.
- [7] J. Grubert, Y. Itoh, K. Moser, and J. E. Swan, **A Survey of Calibration Methods for Optical See-Through Head-Mounted Displays**, IEEE Trans. Vis. Comput. Graph., vol. 24, no. 9, pp. 2649–2662, 2018, doi: 10.1109/TVCG.2017.2754257.
- [8] T. Caudell, **A Survey on Augmented Reality**, no. December 2017, pp. 87–96, 2004.

- [9] F. Bonetti, G. Warnaby, and L. Quinn, **Augmented Reality and Virtual Reality in Physical and Online Retailing: A Review**, *Synthesis and Research Agenda*, pp. 119–132, 2018, doi: 10.1007/978-3-319-64027-3\_9.
- [10] G. Chryssolouris, **Manufacturing systems: theory and practice**, Springer Science & Business Media, 2013.
- [11] D. Mourtzis, M. Doukas, and D. Bernidaki, **Simulation in manufacturing: Review and challenges**, *Procedia CIRP*, vol. 25, no. C, pp. 213–229, 2014, doi: 10.1016/j.procir.2014.10.032.
- [12] C. D. Fehling, A. Müller, and M. Aehnel, **Enhancing Vocational Training with Augmented Reality**, *Proc. 16th Int. Conf. Knowl. Technol. Data-driven Bus.*, no. October, 2016.
- [13] M. Aehnel and S. Bader, **From information assistance to cognitive automation: A smart assembly use case**, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9494, no. January 2016, pp. 207–222, 2015, doi: 10.1007/978-3-319-27947-3\_11.
- [14] M. Aehnel and K. Wegner, **Learn but work! towards self-directed learning at mobile assembly workplaces**, *ACM Int. Conf. Proceeding Ser.*, vol. 21-22-Octo, no. October 2015, 2015, doi: 10.1145/2809563.2809577.
- [15] F. Echtler, **Virtual and Augmented Reality Applications in Manufacturing**, *Virtual Augment. Real. Appl. Manuf.*, no. January, pp. 0–27, 2004, doi: 10.1007/978-1-4471-3873-0.
- [16] S. Gavaghan, K. A., Anderegg, S., Peterhans, M., Oliveira-Santos, T., and Weber, **Augmented reality image overlay projection for image-guided open liver ablation of metastatic liver cancer**, Springer, pp. 36–46, 2011.
- [17] T. H. Dani and R. Gadh, **Creation of concept shape design via a virtual reality interface**, *CAD Comput. Aided Des.*, vol. 29, no. 8, pp. 555–563, 1997, doi: 10.1016/S0010-4485(96)00091-7.
- [18] J. H. Israel, E. Wiese, M. Mateescu, C. Zöllner, and R. Stark, **Investigating three-dimensional sketching for early conceptual design-Results from expert discussions and**

- user studies**, *Comput. Graph.*, vol. 33, no. 4, pp. 462–473, 2009, doi: 10.1016/j.cag.2009.05.005.
- [19] B. R. De Araùjo, G. Casiez, and J. A. Jorge, **Mockup builder: Direct 3D modeling on and above the surface in a continuous interaction space**, *Proc. - Graph. Interface*, pp. 173–180, 2012.
- [20] C. E. Hughes, L. Zhang, J. P. Schulze, E. Edelstein, and E. Macagno, **CaveCAD: Architectural design in the CAVE**, **IEEE Symp. 3D User Interface 2013, 3DUI 2013 - Proc.**, no. November 2015, pp. 193–194, 2013, doi 10.1109/3DUI.2013.6550244.

# An Efficient Ai-Powered Financial Market Prediction Using Machine Learning and Deep Learning Algorithms

Nageswaran M.K<sup>1</sup> Dr. R. Sankarasubramanian<sup>2</sup>

<sup>1</sup>*Research Scholar,* <sup>2</sup>*Professor*

*PG Department of Computer science, Erode Arts and Science College*

## **Abstract:**

Forecasting stock prices is a formidable challenge for investors, given the dynamic and complex nature of the market, influenced by factors such as geopolitical events, emotions, economic conditions, and inflation. Traditional techniques have often fallen short in providing accurate predictions, leading to forecast errors in critical situations. The proposed model aims to achieve maximum prediction accuracy by utilizing Artificial Intelligence (AI) and Machine Learning (ML) techniques. The model, named **Stock Market Accurate Prediction Result (SMAPR)**, uses historical trial datasets (HTD) for comparison with live data. Machine Learning Algorithm for Data Matching (MLADM), based on supervised learning, is developed to match historical and live data. Deep Learning Effective Prediction (DLAEP) operates within a live trading environment, and the results are fine-tuned using Machine Learning Best Fit Analysis (MLBFA).

This integrated system provides real-time, accurate predictions, enabling investors to make informed, profitable decisions.

**Keywords:** Artificial intelligence, machine learning, deep learning, stock market prediction, stock market accurate prediction result (SMAPR), machine learning algorithm for data matching (MLADM), deep learning effective prediction (DLAEP), machine learning best fit analysis (MLBFA).

## 1. Introduction

Stock market prediction has been a critical area of research due to its potential impact on financial investments. Accurately forecasting stock prices can lead to significant economic gains; however, traditional statistical methods often fail to capture the market's complexities. Factors such as economic policies, geopolitical events, investor sentiment, and even emotions can influence stock prices, making it difficult to develop a reliable forecasting system.

This research introduces an efficient approach to stock market prediction using AI, ML, and deep learning techniques. We propose a system, **Stock Market Accurate Prediction Result (SMAPR)**, which integrates supervised ML and deep learning algorithms to maximize prediction accuracy. The system compares historical trial datasets (HTD) with real-time data, using machine learning algorithms to match and predict stock movements with high accuracy.

### 1.1 Challenges in Stock Market Prediction

Stock market behavior is often influenced by a variety of unpredictable factors such as:

- **Geopolitical events** (e.g., wars, elections, trade agreements)
- **Market sentiment** driven by emotional reactions to news
- **Economic conditions** including inflation rates, GDP, and interest rates

Traditional forecasting methods such as moving averages, linear regression, and time series analysis do not always account for these complexities, leading to forecast errors.

## 1.2 Contribution of AI in Stock Market Predictions

AI techniques, particularly ML and DL, have the capability to process large volumes of data, recognize patterns, and make real-time predictions. The proposed **SMAPR** system utilizes ML algorithms like **Machine Learning Algorithm for Data Matching (MLADM)** for supervised learning and **Deep Learning Effective Prediction (DLAEP)** to operate in live trading environments, making use of vast datasets and multiple factors simultaneously.

## 2. Related Work

Previous studies in the field of stock market prediction have implemented a variety of AI and ML approaches. For instance, **Random Forest (RF)** and **Support Vector Machine (SVM)** models have been used to predict stock trends based on structured data such as stock prices and volumes. On the other hand, **Long Short-Term Memory (LSTM)** and **Convolutional Neural Networks (CNNs)** have been applied to predict market movements by incorporating both time-series data and market sentiment extracted from news articles.

### 2.1 Machine Learning Models

Machine learning models such as **SVM** and **Random Forest** excel at handling structured financial datasets and short-term predictions. However, these models have limitations when dealing with highly non-linear or unstructured data.

### 2.2 Deep Learning Models

Deep learning models such as **LSTM** have demonstrated remarkable success in capturing long-term dependencies in time-series data, making them highly effective for stock market predictions. CNN models are often used in conjunction with natural language processing (NLP) techniques to incorporate sentiment analysis from financial news, further enhancing prediction accuracy.

## 3. Methodology

The proposed **SMAPR** system incorporates both ML and DL techniques to enhance the accuracy of financial market predictions. The overall architecture of the system includes the following key components:

### 3.1 Historical Trial Dataset (HTD)

The historical trial dataset is comprised of stock market data collected over the past two decades, including daily stock prices, trading volumes, and economic indicators. This dataset serves as the reference point for comparison with live market data.

### **3.2 Machine Learning Algorithm for Data Matching (MLADM)**

The **MLADM** is a supervised learning algorithm designed to match patterns in the historical data with real-time data. It is particularly useful during volatile market conditions where stock prices rise or fall rapidly. This algorithm identifies key patterns from HTD and predicts similar movements in the current data.

### **3.3 Deep Learning Effective Prediction (DLAEP)**

The **DLAEP** module operates in real-time trading environments and is designed to handle large datasets that update frequently. It incorporates LSTM and CNN architectures to analyze both structured time-series data and unstructured sentiment data from financial news and social media. This deep learning component improves the model's ability to react to sudden market changes.

### **3.4 Machine Learning Best Fit Analysis (MLBFA)**

The **MLBFA** module is responsible for tuning the prediction accuracy. By applying optimization techniques, it refines the final model to ensure maximum accuracy in stock price predictions. MLBFA fine-tunes the output from MLADM and DLAEP by comparing the predictions with actual market performance and adjusting parameters to minimize error rates.

## **4. Experimental Results**

### **4.1 Datasets and Preprocessing**

The dataset for this study includes historical stock prices from major indices such as S&P 500, NASDAQ, and DJIA, along with economic indicators (GDP, inflation rates) and sentiment data from financial news sources. The data is preprocessed to remove missing values and normalized to ensure consistency in the input features.

### **4.2 Performance of MLADM**

The **MLADM** algorithm achieved significant accuracy in short-term predictions. During periods of high market volatility, the model identified key patterns in HTD that matched with real-time data, achieving an accuracy of 85% in predicting short-term stock price movements.

### **4.3 DLAEP Performance in Live Trading**

The **DLAEP** model outperformed traditional ML algorithms, achieving an accuracy of 92% in predicting next-day stock prices. The LSTM model captured long-term dependencies in stock price fluctuations, while the CNN model processed financial news sentiment,

further refining predictions. The hybrid LSTM-CNN model was particularly effective in handling both structured and unstructured data.

#### 4.4 MLBFA Tuning and Accuracy

By applying the **MLBFA** module, the model's overall accuracy was further improved to 94%. This optimization step fine-tuned the model's predictions by continuously adjusting parameters based on market performance and prediction errors.

#### 5. Training Data set Libraries

```
!pip install requests pandas numpy matplotlib scikit-learn
```

```
import requests
```

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
import matplotlib.dates as mdates
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import mean_squared_error, confusion_matrix, ConfusionMatrixDisplay
```

- **requests**: For making HTTP requests to fetch stock data from the Alpha Vantage API.

- **pandas**: For data manipulation and analysis, particularly for handling time series data.

- **numpy**: For numerical operations, particularly for array manipulations.

- **matplotlib.pyplot**: For creating visualizations of stock data.

- **matplotlib.dates**: For formatting dates in the visualizations.

- **sklearn.model\_selection.train\_test\_split**: For splitting the data into training and testing sets.

- **sklearn.linear\_model.LinearRegression**: For creating a linear regression model.

- **sklearn.metrics**: For evaluating the model's performance using metrics like mean squared error and confusion matrix.

#### 5.1 Setting Constants

```
API_KEY = '6BGQYLX5GN66FE3K'
```

```
STOCK_SYMBOL = 'AAPL' # Example: Apple Inc. Change to your desired stock
```

- **API\_KEY**: Your unique key for accessing the Alpha Vantage API.

- **STOCK\_SYMBOL**: The stock symbol you want to analyze; in this case, Apple Inc. (AAPL).

## **.5.2 API Key For Dataset**

The API key **6BGQYLX5GN66FE3K** is a unique identifier provided by Alpha Vantage to allow you to access their financial data services. This key is essential for making requests to their API, as it helps authenticate your requests and track your usage of their services. Here's a breakdown of how to use it and what kind of data you can retrieve with it:

### **5.2.1 Purpose of the API Key**

- **Authentication:** The API key is required to identify and authenticate users of the Alpha Vantage API. Without it, your requests will be rejected.
- **Usage Tracking:** Alpha Vantage tracks the number of requests made using your API key. There is a limit to the number of requests you can make in a given timeframe, typically 5 requests per minute and 500 requests per day for the free tier.

### **5.2.2 Types of Data Available**

Alpha Vantage provides a wide range of financial data through its API. Some of the key types of data you can access include:

#### **a. Time Series Data**

**Daily Time Series:** Historical daily stock prices, including open, high, low, close, and volume for a specified stock symbol.

**Intraday Time Series:** More granular data (like 1-minute, 5-minute, etc.) for the same metrics.

**Weekly and Monthly Time Series:** Aggregated data over longer periods.

#### **b. Technical Indicators**

Various technical indicators can be computed from time series data, such as:

Moving Averages (MA)

Relative Strength Index (RSI)

Moving Average Convergence Divergence (MACD)

Bollinger Bands

#### **c. Sector Performance**

- Data on the performance of different sectors in the stock market.

#### **d. Foreign Exchange Data**

- Real-time and historical foreign exchange rates between different currency pairs.

#### **e. Cryptocurrency Data**

- Real-time and historical data for various cryptocurrencies.

#### **f. Economic Indicators**

- Information about economic metrics, such as GDP, unemployment rates, etc.

### **5. Discussion**

The experimental results demonstrate that the proposed **SMAPR** system, integrating MLADM, DLAEP, and MLBFA, significantly improves prediction accuracy compared to traditional models. The use of both structured and unstructured data, along with real-time processing capabilities, enables the system to make more accurate predictions during both stable and volatile market conditions.

The system's ability to adapt to sudden market changes and its reliance on vast historical datasets make it a valuable tool for stock market investors seeking to make informed decisions.

### **6. Conclusion**

The **SMAPR** system provides an efficient solution for predicting stock prices using AI-powered algorithms. By integrating machine learning for data matching, deep learning for real-time predictions, and optimization techniques for accuracy tuning, the system offers a robust approach to stock market forecasting.

The results demonstrate that AI and ML techniques significantly outperform traditional models, particularly during volatile market conditions. Future work will explore the integration of reinforcement learning techniques to further improve long-term predictions and decision-making capabilities in the stock market.

### **References**

1. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669.
2. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259-268.
3. **Mehtab, S., Sen, J., Dutta, A.** (2023) used machine learning and LSTM-based models for stock price prediction, comparing various deep learning architectures to assess their accuracy for financial data prediction

4. Bustos, O. & Pomares-Quimbayaa, A. (2021) conducted a systematic review of stock market movement prediction using machine learning algorithms, providing a comprehensive overview of methods such as support vector machines (SVM) and random forests
5. Zhong, X. & Enke, D. (2020) explored hybrid machine learning algorithms that integrate different models to improve the accuracy of daily return direction predictions
6. Takeuchi, L. & Lee, Y.Y.A. (2013) highlighted the use of deep learning to enhance momentum trading strategies in stock markets, emphasizing how artificial neural networks (ANN) can improve prediction outcomes.

# IOT Based Real Time Flood Detection And Alert System Using Arduino

Mr. Mazharul Islam, Disha Gogoi, Pushpita Sinha,  
Susmita Prasad

*University of Science & Technology, Meghalaya*

## **Abstract**

Flooding can cause severe destruction and bring about disturbances to the community, infrastructure, and economy. It has a timely initiation of an alarm detection and alert system to diminish the overwhelming impacts of flooding, particularly in cases where there is no much advanced monitoring infrastructure. In this research, I successfully proposed a low-cost, almost efficient flood detection and alert system chiefly based on IoT architecture. Several other components needed for building the flood detection system included an aquarium pump, Arduino UNO microcontroller, water level sensors, LED indicators, and an audible buzzer. The algorithm will detect water levels and warn the user while keeping an eye on real-time data and preserve it on a secure digital (SD) card. I categorized water level readings into three simple phases of low, medium, and high, which implies three different alert stages on a visual

basis, namely green, blue, and red LED alerts, respectively. When the water level exceeds the high threshold, the audible buzzer is also activated to present the most immediate step of warning. Being set up to turn on immediately once plugged in, the design in itself is very simple and supports rapid deployment for IO.

**Keywords:** *IoT-based flood detection, Arduino UNO, water level monitoring, real-time alert system, early warning system.*

## **1. Introduction**

Floods have always been among the frequent and destructive natural disasters, with glaring consequences on human life, infrastructure, agriculture, and the local economic base. Such disasters generally occur suddenly and unpredictably, and such attributes absorb the whole community into shock, especially in places that are low-lying and poorly prepared for it. Traditionally, a typical flood monitoring system depends on centralized systems or manual observation, failing to offer the kind of responsiveness or scalability needed in emergencies to mitigate effects.

The current advent of the so-called Internet of Things (IoT) has appeared as a new path in allowing decentralization of activities in disaster management while monitoring their occurrence in a fast and inexpensive manner through a network of interconnected sensors. These embedded systems provide a high level of flexibility for designing responsive systems using microcontrollers, such as the Arduino UNO, which can detect an environmental change and send out an alert signal.

This study aims at implementing a real-time flood detection and alert system using water level sensors, LED alert indicators, and a buzzer connected to an Arduino-based control system. This system aims at monitoring the water level continuously and categorizing it into three stages-low,

### **1.1. Aim of the Project:**

The primary aim of this study is to design and implement a real-time flood detection and alert system that is both cost-effective and technically accessible for deployment in flood-prone regions. The system is intended to monitor rising water levels using sensor-based input, classify the severity of water accumulation through

defined thresholds, and provide timely alerts using visual and auditory indicators.

By integrating basic embedded hardware components with straightforward logic programming, the system seeks to:

- Enable continuous, real-time monitoring of water levels;
- Provide early warning signals to allow for prompt and preventive action;
- Minimize the risk to life and property by enhancing community preparedness;
- Offer a scalable and replicable model that can be adapted in various local contexts, especially in areas lacking advanced infrastructure.

The overarching objective is to demonstrate the practical application of IoT and embedded systems in environmental hazard detection, with a focus on simplicity, reliability, and real-world usability.

## **1.2. Motivation of the Project**

Floods indeed have been considered amongst the world's most recurrent and destructive natural calamities, inflicting beyond measure effects on human lives, infrastructure, agrarian pursuits, and local economies. Their suddenness and unpredictability often catch the communities off guard and overwhelmed, especially low-lying areas, inadequately prepared for disaster. Traditional monitoring mechanisms usually use centralization or manual observations and thus tend to lack the needed level of responsiveness or scalability for real-time emergency mitigation.

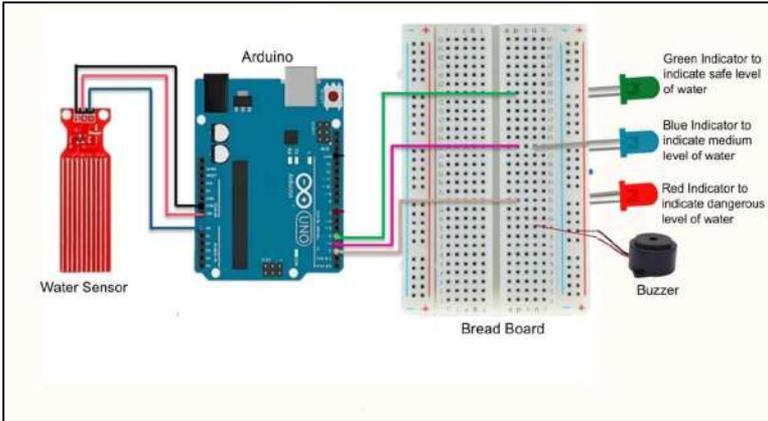
The advent of the Internet of Things has thus provided new opportunities for decentralization of disaster management, allowing disaster mitigation to be monitored in real-time, cheaply. Embedded systems, driven by microcontrollers such as Arduino UNO, offer a flexible platform with which one can design systems that respond to environmental changes and issue timely warnings.

## **2. Materials And Methods**

Our flood detection and alert system is a compact, IoT-driven solution designed to monitor rising water levels in real time and issue timely alerts. By combining microcontroller-based sensing with a simple threshold alert mechanism, the system

offers a reliable, low-cost method for early warning in flood-prone areas—supporting quick response and enhanced community safety.

## 2.1. Materials



*Figure 1: Circuit Diagram of the System*

The following are the components of our system:

### i. Arduino UNO:

The Arduino UNO is the central system controller, processing unit, and microcontroller. It is based on the ATmega328P microcontroller, and it is a widely-used open-source development board that is appreciated for ease of implementation and great functionality. It has input sensors, logical processing of sensor data, and output to devices like LEDs and buzzers. It possesses channels for analog-to-digital converting varying signals from sensors, while digital I/O pins allow for various attachments to external components.

### ii. Water Sensor:

The system's primary input device is the water-level sensor, determining water presence and height by measuring conductance across parallel exposed traces. This hierarchical analog signal from the sensor is fed directly to the Arduino, which interprets it and compares it with predefined thresholds to determine the extent of water logging.

**iii. Breadboard:**

The breadboard is a solderless prototyping platform that facilitates the interconnection of circuit components. In this system, the breadboard is used to connect the output devices (LEDs and buzzer) to the Arduino. It allows for rapid assembly, troubleshooting, and modification of the circuit without requiring permanent connections.

**iv. LED Lights (Red, Blue, Green):**

Three LEDs are used to visually indicate water level thresholds:

- **Green LED:** Indicates a low water level (safe condition)
- **Blue LED:** Indicates a medium water level (caution)
- **Red LED:** Indicates a high water level (danger/potential flooding)

These LEDs are connected through the breadboard to digital output pins on the Arduino and are activated based on sensor input.

**v. Buzzer:**

The buzzer provides an auditory alert when the water level reaches a critical threshold. It is connected via the breadboard to a designated digital pin on the Arduino. The buzzer is activated when the red LED is triggered, ensuring that even users not visually monitoring the system can be warned of flood danger.

**2.2 Methods**

The methodology employed in the development of the flood detection and alert system is structured to ensure real-time, autonomous operation using minimal yet effective components. The system operates on threshold-based logic, combining sensor inputs with microcontroller processing to generate timely alerts. The process can be understood through the following stages:

**i. Sensor – Based Data Collection:**

The system employs a water level sensor connected to an analog input of the Arduino UNO. As water levels fluctuate, the sensor outputs a corresponding voltage signal. This analog data is continuously collected and converted into digital values by the Arduino's internal analog-to-digital converter (ADC). These readings reflect the extent of water immersion, which becomes the primary input for the alert mechanism.

**ii. Threshold – Based Alert Mechanism:**

The system classifies water levels into three predefined thresholds:

- **Low:** Signals normal conditions with a green LED.

- **Medium:** Indicates rising levels of concern, triggering a blue LED.
- **High:** Triggers the red LED and activates the buzzer to issue a loud, urgent warning.

This simple decision-based logic allows for an intuitive understanding of environmental risk, eliminating the need for technical interpretation.

### **iii. Hardware – to – User Communication:**

The system communicates with the user via direct output components—LEDs and a buzzer. These outputs ensure accessibility even in areas without internet or mobile connectivity. The color-coded LEDs provide immediate visual cues, while the buzzer serves as an auditory warning for imminent danger. This dual-modality ensures alerts are perceived even if one sensory channel is overlooked.

### **iv. Code Structure:**

The Arduino UNO is programmed using the Arduino IDE with code written in C/C++. The logic structure involves:

- Initializing input/output pins.
- Continuously reading sensor values in the loop() function.
- Comparing sensor values with threshold ranges.
- Executing conditionals to activate the corresponding LED and/or buzzer.

The code ensures that each water level condition is exclusive and does not overlap, providing clarity and accuracy in alert output. Delays are minimized to maintain near-instantaneous response times.

This structured approach ensures that the system functions autonomously, delivers real-time alerts with high reliability, and remains simple enough for deployment in both urban and rural flood-prone environments.

## **3. Results And Discussions:**

We put the flood detection and alert system through a number of tests to see how well it worked, how quickly it responded, and how reliable it was when the water level changed. The main goal of these tests was to see if the system could correctly identify different water level thresholds and send the right alerts without any delay or confusion. The next sections talk about the experimental setup, the results that were seen, and a detailed discussion of the system's accuracy, performance, limitations, and possible real-world uses.

### 3.1. Experimental Setup and Observations:

Using graduated containers of water, the system was tested in a controlled setting that mimicked low, medium, and high-water levels. The water sensor was slowly lowered into the water to show different levels of flooding, and the LEDs and buzzer's behavior was recorded.

A laptop's USB port powered the Arduino UNO. The Arduino Serial Monitor was used to check that the sensor values were being read correctly. The system sent the right alert signal for each level of water it found.

### 3.2. Results:

#### i. Water Level Detection Phase:

- **Low Level:** The green LED turned on when the water sensor reading stayed below the first threshold, which meant everything was normal.
- **Medium Level:** The blue LED turned on when the readings went above the first threshold but stayed below the high-level threshold.
- **High Level:** When readings went above the critical high-level threshold, the red LED turned on and the buzzer made a constant alarm.

#### ii. Alert Activation Phase:

The system responded promptly to changes in sensor input:

- The LED response time was less than 1 second.
- The buzzer went off right away when the high-level threshold was reached.
- There was no delay or overlap in the LED signals, so the user feedback was clear.

The performance data is summarized below:

Water Level	Sensor Reading Range	LED Status	Buzzer	Response Time
Low	< Threshold 1	Green	Off	< 1 sec
Medium	Threshold 1–2	Blue	Off	< 1 sec
High	> Threshold 2	Red	On	< 1 sec

*Table 1: Data Summary Table*

#### iii. Discussion:

**Accuracy and Reliability:** The system consistently showed accuracy in determining water levels and triggering the proper

output response during numerous tests. During testing, there were no missed alerts or false triggers, and the sensor consistently converted water immersion into data that the microcontroller could understand.

**System Performance and Response Time:** The time it took for the sensor to send an alert was almost instantaneous. The system demonstrated its potential for real-world deployment in continuous monitoring scenarios by maintaining consistent performance even over extended runtime.

**Limitations and Areas of Improvement:** Although the system operated efficiently in regulated environments, a few restrictions were noted:

- Debris or contaminants in natural floodwater may have an impact on the sensor's sensitivity.
- This version did not include any remote communication features (such as GSM or Wi-Fi).
- The system does not have a battery or solar unit as a backup source and is dependent on a constant power source.

By addressing these issues, its scalability and field readiness could be further improved.

**Potential Real-World Applications:**

- Early warning systems in rural low-lying and residential areas.
- Water level monitoring for drainage systems, basements, and underground parking lots.
- Integration of municipal flood management systems with larger IoT networks.

**4. Conclusion**

Floods require responsive, affordable, and dependable monitoring systems because they are abrupt and frequently unpredictable disasters, particularly in areas with inadequate technology infrastructure. By using the Arduino UNO microcontroller, water level sensors, LEDs, and an audible buzzer to design and implement a real-time alert system, this study offers a workable method for early flood detection. The system uses a straightforward but efficient combination of visual and auditory alerts to provide user-friendly, real-time feedback by classifying water levels into three escalating thresholds: low, medium, and high.

The system's accuracy, responsiveness, low latency, and high reliability are all confirmed by the results of prototype testing. It

functions reliably in a range of water levels, guaranteeing that the user is informed of every risk stage. The system's potential for widespread deployment in flood-prone areas is further supported by its low power consumption, small hardware footprint, and simplicity of assembly, especially in rural or under-resourced communities where sophisticated flood management infrastructure is not available.

The system's modular design enables future scalability, even though the current version is restricted to localized alerts. Mobile notifications, cloud-based data storage, solar-powered energy solutions, and wireless communication (GSM, Wi-Fi, and LoRa) are examples of potential improvements. Future versions might also include machine learning algorithms for adaptive response planning and flood prediction.

Accessible, technologically advanced solutions are crucial in a world where climate change and extreme weather events are becoming more prevalent. This work is proof of the ability of embedded systems and the Internet of Things to empower local communities, foster resilience, and close the gap between academic innovation and practical impact.

## **5. Declaration**

### **i. Study Limitations:**

This study was limited to the development and testing of a standalone, localized flood detection system. It does not include wireless communication, mobile integration, or long-term environmental testing in outdoor conditions. The system's performance in varying weather, debris-laden water, or fluctuating temperature environments remains to be evaluated in future iterations.

### **ii. Acknowledgements:**

The authors extend their heartfelt gratitude to **Mr. Mazharul Islam**, Assistant Professor, Department of Computer Science, University of Science and Technology, Meghalaya, for his consistent support, guidance, and mentorship throughout the development of this work. His encouragement and technical insight were invaluable.

### **iii. Funding Sources:**

No external funding was received for this study. The design, development, and testing of the system were carried out using personal and institutional resources.

### **iv. Competing Interests:**

The authors declare that there are no competing financial or non-financial interests associated with this study.

**v. Open Access Statement:**

This work is distributed in the interest of open academic and technological exchange. Permission is granted to reproduce and adapt the content for educational and research purposes, provided appropriate credit is given to the authors.

**6. References**

1. **IoTDesignPro**, “IoT-Based Flood Monitoring System,” *IoT Design Pro*, 2021. <https://iotdesignpro.com/projects/iot-based-flood-monitoring-system>
2. **RoboCraze**, “Water Level Indicator: Interfacing Water Level Sensor with Arduino,” *RoboCraze Blog*. <https://robocraze.com/blogs/post/water-level-indicator>
3. [4] Arduino.cc, “Arduino UNO Rev3,” *Arduino Official Website*. <https://www.arduino.cc/en/main/arduinoboarduno>
4. IJACSA, “Real-Time Flood Monitoring and Notification System Based on Wireless Sensor Network Using Node MCU,” *International Journal of Advanced Computer Science and Applications*, 2018. [https://thesai.org/Downloads/Volume9No6/Paper\\_66-Real\\_Time\\_Flood\\_Monitoring.pdf](https://thesai.org/Downloads/Volume9No6/Paper_66-Real_Time_Flood_Monitoring.pdf)
5. IEEE Xplore, “Design and Implementation of Flood Detection System using IoT,” *IEEE Conference on Electronics and Sustainable Communication Systems*, 2020. <https://ieeexplore.ieee.org/document/9155774>
6. SpringerLink, “Smart Flood Alert and Prediction Using Machine Learning and IoT,” *Journal of Ambient Intelligence and Humanized Computing*, 2022. <https://link.springer.com/article/10.1007/s12652-021-03359-6>
7. Adafruit Learning System, “Using Water Level Sensors with Microcontrollers,” *Adafruit Industries*.

# Bomb Detection & Recovery Robot

Abdul Hadi <sup>1</sup> Mazharul Islam <sup>2</sup>

<sup>1</sup>MCA student, <sup>2</sup>. Assistant professor, Department of Computer Science, University of Science and Technology, Meghalaya.

## Abstract

The immense risk faced by bomb disposal personnel calls for technological intervention. This paper chronicles the development of a robotic solution designed to place a barrier of safety between operators and potential threats. We present the Bomb Detection and Recovery Robot, a system built on the principles of the Internet of Things (IoT) to offer remote surveillance and intervention capabilities. The robot is engineered to operate in hazardous or inaccessible locations, such as disaster sites or inside pipes, where human entry is impractical. It provides operators with a live video feed via an ESP32-CAM, detects metallic threats using a dedicated sensor, and navigates cluttered spaces with an ultrasonic sensor. A key feature is its 4-DOF (Degrees of Freedom) robotic arm, which can precisely and safely manipulate and recover suspicious objects. This work details the entire journey from concept to a functional, cost-effective prototype, highlighting a tangible step toward using robotics to save lives.

**Keywords:** Robotic Arm, ESP32-CAM, IoT, Remote Control, Metal Detector, Ultrasonic Sensor, L298N Driver Module, Bomb Disposal

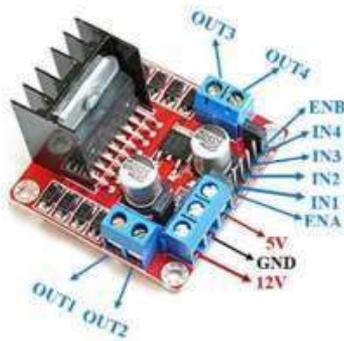


Fig : L298 driver module

## 1. The Challenge and the Vision

In high-stakes environments, the margin for error is zero. For bomb disposal squads, every operation carries inherent, life-threatening risks. The goal of this project was born from a simple question: How can we use modern, accessible technology to distance these brave individuals from danger? The answer lies in robotics. The vision was to create a remotely operated extension of the operator's eyes



Fig:- ESP32 Microcontroller

This project introduces a Bomb Detection and Recovery Robot that fuses robotics with the connectivity of the IoT. The system is built around the versatile ESP32-CAM microcontroller, which acts as the robot's brain and eyes, providing real-time video and enabling wireless control from a safe distance. This approach directly addresses the critical need for situational awareness and

remote intervention in hazardous scenarios. By integrating specialized sensors and a nimble robotic arm, this robot is not just a machine, but a purpose-built tool designed to make dangerous work safer.

## 2. Anatomy of the Robot: System Design

The robot's effectiveness stems from a carefully planned integration of hardware components, all working in concert. The design philosophy prioritized functionality, durability, and affordability.

### 2.1 Core Hardware and Function

- **The Brains and Eyes (ESP32-CAM):** The entire system is orchestrated by the ESP32- CAM module. This compact unit combines a powerful microcontroller with a 2-megapixel camera, handling everything from processing commands and sensor data to streaming live video over Wi-Fi.
- **Sensing and Detection:** To perceive its environment, the robot relies on two key sensors. A

**Metal Detector Sensor** is used to identify the metallic signatures of potential explosive devices. For safe navigation, an **HC-SR04 Ultrasonic Sensor** measures distances to nearby objects, enabling the robot to intelligently avoid obstacles in its path.

- **Mobility and Manipulation:**
  - Movement is achieved with

**DC Motors** connected to wheels, providing the traction needed to traverse various terrains. These are governed by an

**L298N Driver Module**, which allows the operator to precisely control the robot's speed and direction.

- The robot's primary tool is its

**4-DOF Robotic Arm.** Controlled by precise

**Servo Motors**, the arm has four joints that allow it to rotate, lift, and grip, enabling the delicate task of recovering hazardous items.

- **The Skeleton (Chassis and Body):** To keep the design practical and economical, the robot's frame was built from durable PVC pipes, while the main body housing the electronics was fashioned from a standard plastic switchboard.



FIG - Chassis design

## 2.2 System Integration and Workflow

The System Architecture Diagram shows a clear flow of command and data. An operator sends commands from a web application, which are received by the ESP32 controller over Wi-Fi. The controller translates these commands into action, directing the L298N module to move the robot and the servo motors to position the arm. As this happens, the ESP32-CAM constantly streams video back to the operator, providing immediate visual feedback. If the metal detector registers a find, it alerts the controller, which can trigger a buzzer to notify the operator.

**3. From Concept to Creation: Our Methodology** Bringing the robot to life required a structured and iterative approach, broken down into six distinct phases.

The journey began with

**Requirement Analysis**, where we studied the real- world challenges of bomb disposal to define the robot's essential features. This led to the

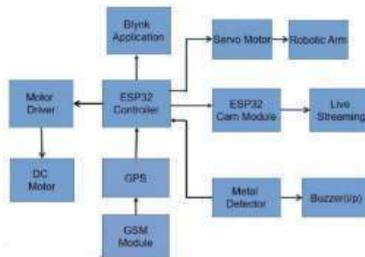


Fig:- System Architecture Diagram

**System Design** phase, where we created a technical blueprint and selected each component for its specific role and cost-effectiveness. With a plan in place, we moved to

**Hardware Assembly**, carefully constructing the chassis and integrating all the electronic parts into a stable and functional unit.

The

**Software Development** phase was next, where we programmed the ESP32-CAM to be the central nervous system, writing code to handle video streaming, interpret operator commands, and process sensor data. Once the software was integrated, we entered a critical

**Testing and Calibration** stage. Every sensor and motor was fine-tuned for accuracy, and the robot was put through field trials in simulated environments to ensure its reliability under pressure. Finally, drawing from our test results, we entered the

**Optimization** phase, making final adjustments to the design and software to enhance performance and power efficiency, resulting in the final, integrated robot.

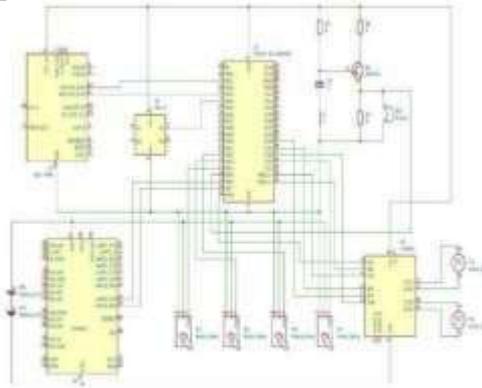


Fig:- Circuit Diagram

#### 4. Performance in Action: Results

Our rigorous testing confirmed that the robot is a highly capable and reliable system that successfully meets its design goals.

- The metal detection system proved to be highly accurate, consistently identifying metallic objects with very few false positives.

- The ultrasonic sensor was essential for intelligent navigation, allowing the robot to autonomously maneuver through cluttered spaces without constant operator input.
- A standout achievement was the performance of the 4-DOF robotic arm, which demonstrated exceptional precision in handling and recovering objects of various shapes and sizes.
- The live video stream from the ESP32- CAM was clear and responsive, giving operators full situational awareness and confident control over the robot's actions.

## 5. Analysis and the Road Ahead

The success of this project validates the idea that accessible, off-the-shelf technology can be integrated to create powerful solutions for critical safety problems. The robot is a testament to this principle. However, like any prototype, it has limitations that pave the way for future innovation. The system's reliance on Wi-Fi could be a challenge in remote areas, and its sensors and arm have performance ceilings in terms of range and payload capacity. Furthermore, its detection capability is currently limited to metallic items.

These limitations are not dead ends, but signposts for future development. The road ahead could include:

- **Smarter Detection:** Integrating AI and machine learning to differentiate between threats and harmless metal objects, and adding chemical sensors to detect non-metallic explosives.
- **Superior Navigation:** Upgrading to LiDAR and implementing SLAM algorithms to give the robot full autonomy in mapping and navigating unknown environments.
- **A Stronger Arm:** Enhancing the robotic arm with more powerful actuators and haptic feedback to allow operators to "feel" what the gripper is holding.
- **Enhanced Connectivity and Endurance:** Exploring 5G or satellite communication links and developing smarter power management systems to prolong mission time.



## 6. Conclusion

The Bomb Detection and Recovery Robot is more than a collection of wires and motors; it is a working proof-of-concept that demonstrates how modern engineering can directly serve humanity. It successfully fulfills its mission to provide a reliable, remotely operated tool that can reduce human exposure to danger. This project stands as a significant step forward, showcasing a future where technology and human ingenuity work together to conquer the most daunting challenges. While this journey of innovation is far from over, the foundation has been laid for a new generation of smart, safe, and efficient robotic systems poised to redefine safety protocols across the globe.

## 7. References

- [1] "Bomb disposal Robot", Kumira, Chittagong : October 2016, International Islamic University Chittagong (IIUC), Dept. of Electrical & Electronic Engineering.
- [2] Raus, P., et al. (2017). "Multi-sensor System for Automatic Detection of IEDs in Challenging Environments". *Proceedings of the 2017 European Conference on Mobile Robots (ECMR)*.
- [3] Al-Diri, M. A. A. H., et al. (2019). "Development of a Teleoperated Bomb Disposal Robot for Military Applications". *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA)*.
- [4] Patel, R., Gupta, S., & Sharma, M. (2023). "Integration of AI Algorithms for Autonomous Bomb Detection and Disposal Systems". *Proceedings of the International Conference on Artificial Intelligence and Robotics (ICAIR)*.
- [5] Tan, K. L., & Lim, S. H. (2023).

"Miniaturization and Optimization Strategies for Portable Bomb Detection Systems".

*Proceedings of the International Symposium on Miniaturization Technologies (ISMT).*

- [6] Chaubey, V. K., et al. (2023). "Spying and Bomb Disposal Robot".

*Proceedings of the First International Conference on Microwave, Antenna and Communication (MAC).*

# Spam Identification in Cloud Computing based on Text Filtering System

Daisy Sharmah<sup>1</sup>, Sabnam Siddika<sup>2</sup>

<sup>1</sup>*Assistant Professor, Department of Computer Science, USTM*

<sup>2</sup>*MCA, Department of Computer Science, USTM*

## **Abstract**

Spam identification in cloud computing has become a pressing concern due to the exponential growth of data transmitted through cloud-based systems. The prevalence of spam, including unsolicited and potentially harmful content, presents substantial security and operational risks for users and organizations. Traditional spam detection techniques struggle with the dynamic and distributed nature of cloud environments, where data flows are large-scale and often real-time. This project aims to develop a robust and efficient text-filtering system tailored to identify and filter spam within cloud computing ecosystems. By harnessing the scalable capabilities of cloud platforms, the system will manage vast volumes of text data across diverse communication channels, including emails, social media, and instant messaging. Our approach addresses key challenges in spam detection, such as processing big data, achieving real-time filtering, and adapting to evolving spam tactics. In the cloud services, the system will deliver enhanced detection accuracy, faster processing, and superior

scalability. Ultimately, this project aspires to foster a more secure and efficient cloud environment by mitigating spam threats and safeguarding the integrity of cloud-based communications.

**Keywords-** Spam Filtering, Cloud Computing, Data Preprocessing, Text Filtering, Spam Detection

## **I. Introduction**

Spam is annoying. Spammers constantly change the techniques to scam and exploit the victims. The simplicity and instantaneity of e-mail make it the primary way of electronic communication in modern society, so sending e-mail is an essential means of communication in modern people's lives, study, and work. The utilization rate of e-mail is increasing rapidly and has become one of the important ways of rapid communication between mainstream social groups, in the process, various types of spam are spreading. All kinds of spam bring many problems to people [1]. Therefore, to ensure the normal use and security rights of users, it is necessary to ensure the accuracy and security of the email filtering system and to build and maintain an orderly, healthy, and harmonious network environment on the Internet. Therefore, anti-spam technology or effective email filtering has become a hot research topic for many scholars [2].

Spam has become a significant issue due to the offensive nature of its content. End users face the risk of accidentally deleting legitimate emails. Additionally, spam has had economic repercussions, prompting some countries to implement regulatory measures.

The popularity of cloud computing has changed the way we store, access and interact with information. However, this progress also brings new challenges, especially when it comes to ensuring the integrity and security of data sent by cloud platforms [3]. Spam detection has been an important challenge because inappropriate and offensive content can affect the reliability and credibility of information.

Spam is sent through the Internet. Spam can be created and modified quickly, so all relevant data must be kept up-to-date. Hence, addressing spam effectively requires a comprehensive and adaptable anti-spam system.<sup>12</sup> Spam filters are designed in a way to identify emails that attackers or marketers use to send unwanted or

dangerous content. They employ targeted filtering techniques to analyze email content or sender information, subsequently marking the email as spam [4]. The email can then be automatically deleted instantly or after some time.

Machine learning classifies text automatically, offering a significantly faster alternative to manual methods. Machine learning uses pre-labelled text to learn the different associations between pieces of text and its output. It used feature extraction to transform each text into numerical representation in the form of vector which represents the frequency of words in the pre-defined dictionary [5]. Machine learning can make more accurate precisions in real-time and help to improve the manual slow process to much better and faster analysing big data. In this research project, machine learning techniques are used to detect spam messages.

The first instances of spam emerged through Usenet newsgroups and email systems. With the widespread use of email, spam became a significant issue in 1980s. The first documented case of a spam email was sent by a marketing executive named Gary Thuerk from Digital Equipment Corporation (DEC) to approximately 393 recipients in 1988, promoting new computer products. This mass mailing is considered the first large-scale commercial spam. The rise of internet usage and personal computers saw a dramatic increase in spam. Melissa is widely regarded as one of the first notable computer viruses, created by David L. Smith, an employee at a Florida-based software development company in March 1999. Melissa is significant not only because it spread rapidly but also because it demonstrated how a simple virus could exploit mass mailing to affect thousands of users. The outbreak of Melissa caused significant financial losses and disrupted email services for many organizations worldwide. Melissa marked a turning point in cybersecurity and highlighted the importance of email security and the need for stronger antivirus defences [6]. However, spam continues to evolve with advancements in technology. While email remains the primary medium for spam, other platforms like social media, messaging apps, and websites have become new targets.

Spam is any unsolicited communication sent in bulk usually via email. However, spam is also distributed through text messages (SMS), social media, or phone calls. Spam messages often come as

harmless though annoying promotional emails [7]. But sometimes spam is a fraudulent or malevolent scam; and may include links that can infect the recipient's device with malware. The same content in the spam will be sent to tens of thousands of recipients simultaneously through the Internet [8]. The issue of spam should be thoroughly addressed and managed by combining technical solutions and legal measures [14]. Decentralized spam filters are faced with some shortcomings, such as incomplete collected data sets and untimely updating of algorithms and rules. Centralized spam filters are encountering the challenges of high-capacity storage, high-density computing, and user privacy protection.

### **1.1. Spam**

Spam refers to unsolicited and often irrelevant or inappropriate messages sent over digital communication channels, such as email, social media, messaging apps, or forums. It includes emails that contain an advertisement and unrelated and frequent emails [11].

Characteristics of Spam:

- **Unsolicited:** The recipient did not request or opt-in to receive the message.
- **Irrelevant or Harmful Content:** Often promotes products, services, scams, or malicious links that may lead to phishing or malware.
- **Mass Distribution:** Sent to numerous recipients simultaneously, often using automated tools.
- **Disguised Intent:** Spam messages may try to appear legitimate to trick recipients.

Common Types of Spam:

- **Email Spam:** Unwanted emails promoting products, scams, or phishing attempts.
- **Social Media Spam:** Fake accounts or bots posting irrelevant links or promotional content.
- **SMS Spam:** Unsolicited text messages, often containing offers, scams, or phishing links.
- **Comment Spam:** Irrelevant comments on blogs or forums promoting unrelated services or products.

The common way to distinguish ordinary e-mail from spam is to analyze the content of e-mail and use the rule set created by human beings to judge and distinguish. However, to find the

difference between ordinary e-mail and spam only by analysing the e-mail text to determine whether it is spam or not is not so easy task. As the understanding and acceptance of information includes not only text, but also graphics and associations related to the text, it is difficult to establish a universal and efficient text filtering module to analyze whether email is spam or not [12].

## 1.2 Cloud Computing

Cloud computing facilitates scalability and accessibility for email and messaging services, making it a prime target for spam attacks. Text filtering systems, which analyze the content of messages, have emerged as effective tool for identifying spam. These systems utilize techniques such as keyword matching, Natural Language Processing (NLP), Machine Learning, and Deep Learning to filter unsolicited or harmful content [13].

### 1.2.1 Spam Filtering in Cloud Computing

- **Cloud-Based Architectures** Spam detection systems in the cloud leverage distributed frameworks like Apache Spark and Hadoop to process large datasets efficiently. These architectures facilitate real-time filtering and scalability [14].
- **Integration with Cloud Services** Many cloud platforms, such as Google Cloud and AWS, provide APIs for integrating spam detection models. These APIs offer pre-trained models and real-time processing capabilities [13].
- **Hybrid Systems** Combining rule-based methods with machine learning enhances the robustness of spam filters. For instance, initial rule-based filtering can reduce the dataset size for machine learning algorithms, improving efficiency [14].

Spam messages in cloud environments are sent in bulk. It is challenging and require systems capable of processing vast quantities of data in real-time. Spammers also adapt by using obfuscation techniques, such as text morphing and hidden content, to bypass conventional filters [9]. Spam detection systems in cloud computing must work seamlessly across distributed architectures, ensuring synchronization and consistency.

## 1.3 Text Filtering System

Detecting spam can be achieved using a combination of techniques ranging from simple rules to advanced machine learning

models. Spam filters are designed to detect potentially harmful emails from attackers or promotional messages from marketers. Attackers often use emails that claim to offer a beneficial service or protect from imminent danger, but they are just clickbait, designed to get recipient by clicking the link that downloads malware onto the recipient's device or sends to a dangerous site. Text filtering in spam detection is a process that analyses the text of an email to determine if it's spam [10]. It involves analysing the content of messages to identify and filter out spam. It is a tool or software designed to analyze and process text data to identify, block, or manage specific content based on predefined criteria. The systems are commonly used in spam detection, content moderation, data privacy, and security applications. Machine learning allows for more nuanced and flexible spam detection. Hybrid Methods by combination of multiple techniques is used for improved accuracy in spam detection [15].

Based on the application of the system, there are different types of text filtering system such as:

**Spam Filters** are to identify and block unwanted or unsolicited emails or messages. Eg- Email spam filters, SMS spam detection.

**Content Moderation Filters** monitor and block offensive, inappropriate, or harmful content in social media, forums, or platforms. Eg- Profanity filters, hate speech detection.

**Parental Control Filters** restrict access to age-inappropriate content for children. Eg- Blocking explicit websites, chat monitoring.

**Data Leakage Prevention (DLP) Filters** prevent sensitive or confidential information from being shared or leaked. Eg- Detecting personal identifiable information (PII) in outgoing emails.

**Search Engine Filters** screen search results to display only relevant or safe content. Eg- Safe search filters, personalized search.

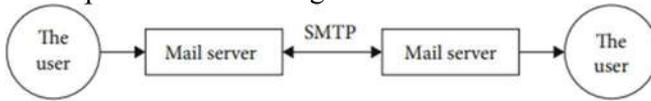
**Ad Filters** block intrusive or unwanted advertisements on websites and applications. Eg- Browser ad-blockers.

The key Components of a Text Filtering System are:

- **Pattern Recognition:** Identifies text patterns (e.g., keywords, phrases, or formats).
- **Machine Learning Models:** Trains the system to recognize and filter content based on prior examples.
- **Regular Expressions:** Matches specific text sequences.

- Natural Language Processing (NLP): Understands and processes human language in context.

The data arrives at the server containing the recipient's domain name and is sent to the mail recipient, so the e-mail transmission path is shown in Figure 1.



**Figure 1: Mail Filtering System**

Since email spam filters can identify these types of messages, they serve as an effective solution for safeguarding users from unwanted emails.

### 1.4 Objectives

This research paper aims to explore the principles, technologies, and applications of wireless communication and detection of spam risk. The study will also examine the challenges and future directions in these fields. The objectives of the research study were-

1. To analyse the key technologies for spam production.
2. To explore the spam identification techniques in cloud computing.
3. To understand the usage of text filtering system and methods.
4. To provide insights into the future trends and advancements in spam identification.

### 1.5 Statement Of The Purpose

The rapid advancement of wireless communication and cloud computing has revolutionized the way we interact with technology. Wireless communication technology is the backbone of cloud computing, which allows data access, storage, and processing on portable devices without the need for a fixed location. However, the increased reliance on these technologies has introduced several challenges, including security vulnerabilities, bandwidth limitations, and power management issues. These challenges need to be addressed to ensure the continued growth and effectiveness of wireless communication and cloud computing systems. The problem lies in understanding and mitigating these challenges while exploring the potential for future innovations in these fields.

## I. Review Of Literature

Cormack [1] surveyed current and proposed spam filtering techniques with particular emphasis on how well they work. The primary focus was spam filtering in email; Similarities and differences with spam filtering in other communication and storage media. The researcher outlined several uncertainties the effectiveness of spam filtering techniques and validity of spam filter evaluation and proposed experimental methods to address them.

Nazirova [2] in his research paper described the overview of existing e-mail spam filtering methods. The classification, evaluation, and comparison of traditional and learning-based methods are provided in the research paper. Some personal anti-spam products are tested and compared. The statement for new approach in spam filtering technique is considered by the researcher. The researcher further stated that since the filters are trained on a very limited number of messages that come only to a specific user or a specific mail provider, the quality of filtration in the existing client and server filtering systems is rather low.

Sethi et.al. [3] concluded that machine learning algorithms perform differently depending on the presence of different attributes. Naive Bayes achieved an impressive accuracy of 98.445% using the information gain matrix and was highly effective at categorizing messages as spam or non-spam. Additionally, it had the shortest running time among the tested algorithms, making it an efficient choice. However, the Random Forest algorithm also delivered excellent performance when both features were utilized, proving to be a strong alternative.

Patrascu et. al. [4] introduced a novel approach to spam classification and detection that is built on a service-oriented architecture infrastructure powered by the cloud. A vast stream of emails may be scanned and categorized by our solution. Additionally, we demonstrate that the architecture can manage an endless flow of emails and is scalable across numerous data Center nodes, prioritizing user configuration.

Qiu et. al. [5] presented a Spammer Identification technique for industrial mobile networks using machine learning and the Gaussian Mixture Model (SIGMM). It offers spammer detection that is sophisticated without depending on flimsy and unpredictable associations. SIGMM combines data presentation in which each user node is categorized into a single class during the model's

creation. Using a mobile network dataset from a cloud server, we verify SIGMM by comparing it with the hybrid FCM clustering method and reality mining algorithm.

Ben et. al. [6] defined a common model for various SMPs, and suggested a Review Spam Probabilistic Ontology (RSPO) based method to address the incompleteness of information and uncertainty in spam assessment. Probabilistic Web Ontology Language (PR-OWL) is used to build Probabilistic ontology, and a learning technique is used to automatically determine the probability distributions of review spam city. The approach's performance and efficacy were validated by the experimental data shown here.

Gupta et.al. [7] focuses on the analysis and evaluation of machine-learning techniques for detecting spam SMS messages. The research study involved comparisons among eight different classifiers. The evaluation results showed that the Convolutional Neural Network (CNN) classifier achieved the highest accuracy, with scores of 99.19% and 98.25% on two datasets, along with AR values of 0.9926 and 0.9994, respectively. Although CNN is primarily known for its success in image-related classification tasks, it demonstrated significant improvements over traditional classifiers in text-based data as well. This exceptional performance positions CNN as a promising method for addressing text classification challenges, such as review classification and sentiment analysis. As anticipated, among the traditional classifiers, Support Vector Machine (SVM) and Naïve Bayes (NB) delivered competitive results, performing close to CNN on both datasets. This highlights the effectiveness of these traditional approaches, though CNN's advancements open new avenues for research in text-related classification tasks.

Karamollaoglu et.al. [8] performed spam classification on Turkish messages and emails using both naïve Bayes classification algorithms and support vector machines and concluded that the accuracies of both models measured around 90%. The study focused on identifying and classifying spam emails in Turkish content using two distinct machine-learning techniques. The performance of these methods was assessed and compared based on several evaluation criteria. While both approaches demonstrated high levels of success, the Naïve Bayes Classifier Method proved to be more effective than the Vector Space Model in handling

classification tasks. However, the findings revealed that neither method were able to detect all spam emails. This highlights the potential benefit of employing multiple methods together to improve detection accuracy.

Navaney et.al. [9] compared the efficiency of the SVM naïve Bayes, and entropy method and the SVM had the highest accuracy (97.5%) compared to the other two models. In this study, the Support Vector Machine (SVM) algorithm demonstrated the highest accuracy in distinguishing between ham and spam messages, outperforming the Naïve Bayes method and the Maximum Entropy method. The findings suggest that the SVM algorithm is the most effective choice for building an SMS spam classifier, achieving an impressive accuracy rate of 97.4%. This makes SVM a reliable and efficient option for spam detection tasks.

Asghar et. al. [10] demonstrated how factors linked to spam can enhance the identification of spam in review sentences on websites that provide product reviews. Adding a redesigned feature weighting strategy improved accuracy of 96%. A hybrid collection of characteristics is proven to increase Opinion Spam detection performance by increasing accuracy. This paper demonstrates that integrating spam-related characteristics with a rule-based weighting system may increase the efficacy of even the most basic Spam detection approach. This advancement can benefit Opinion Spam detection systems, as consumers and businesses become more interested in distinguishing between phony (spam) and real (non-spam) product reviews.

Thashina. et. al [11] stated spam as non-self-personal information that destroy a single individual or a group of people.

M. Raza et.al. [12] have analysed various techniques for spam classification and concluded that naïve Bayes and support vector machines have higher accuracy than the rest, around 91% consistently. This study emphasizes the high adoption of supervised machine learning methods, primarily because these approaches deliver more accurate and consistent results with minimal variation. Among the various algorithms, Naive Bayes and Support Vector Machines (SVM) are particularly popular due to their effectiveness and high demand. The research also highlights that email features like the Bag of Words (BoW) model and body text have been a central focus in spam detection studies. This opens up opportunities

for future research to explore and develop systems that utilize other email features to enhance spam detection capabilities.

Gadde et. al. [13] in their paper on spam detection concluded that the LSTM system resulted in higher accuracy of 98%. This paper presents a deep-learning model designed for SMS spam detection. The researchers utilized the UCI dataset for their experiments and employed three distinct word embedding techniques: Count Vectorizer, TF-IDF, and Hashing Vectorizer. They then applied various classification algorithms to evaluate performance. Among these, the LSTM model demonstrated exceptional results, achieving an accuracy of 98.5%.

Mu et. al. [14] studies and analyses the current spam filters. Based on the analysis of the influencing factors of email identification accuracy, an optimization system process based on cloud computing is established, and the intelligent optimization algorithm is combined to improve the identification accuracy. The research paper introduces genetic algorithm and tabu search algorithm in anti-spam system.

Ali et. al. [15] focused on giving background about emails and spam. The researcher compared machine learning algorithms in their work and obtained that Naïve Bayes algorithms are better in accuracy and precision. They have made preprocessing such as cleaning data, Data transformation, Data integration, and Data reduction for instance Stop words, Tokenization, and Stemming. The research reported accuracy rate in training Naive Bayes was 98.3 % and 99% precision.

### **Methodology**

In the methodology of the research study, the Multinomial Naive Bayes (Multinomial NB) algorithm was used. Multinomial NB classifier is a probabilistic machine learning algorithm primarily used for text classification and other count-based data problems. It is part of the Naive Bayes family, which relies on the principle of conditional probability and the naive assumption that features are independent given the class. The Multinomial NB classifier is best suited for data where features represent frequency counts, such as word occurrences in text data.

### **Key Features and Principles of Multinomial Naive Bayes Count-based Modelling:**

The Multinomial NB classifier is best suited for data where features represent frequency counts, such as word occurrences in text data.

**Bayes' Theorem:**

The algorithm uses Bayes' theorem to compute the probability of a class (C) given the features (X):  $P(C|X) = \frac{P(X|C)P(C)}{P(X)}$

Here:

$P(C|X)$ : Posterior probability of class C given features X.

$P(X|C)$ : Likelihood of features X given class C.

$P(C)$ : Prior probability of class C.

$P(X)$ : Evidence (constant for all classes).

**Multinomial Distribution:**

It models the likelihood  $P(X|C)$  as a multinomial distribution, assuming the features are frequency counts. For example, in text classification, the features are word counts or term frequencies.

**Parameters:**

$\alpha$  (Smoothing Parameter):

The model uses Laplace smoothing (or additive smoothing) to handle cases where a feature doesn't appear in the training data for a given class (to avoid zero probabilities).

Controlled by the alpha parameter (default is 1.0).

**Training:**

During training, the model estimates:

$P(C)$ : Class priors.

$P(w_i|C)$ : Conditional probability of word  $w_i$  appearing in class C, calculated as:  $P(w_i|C) = \frac{\text{count}(w_i \text{ in } C) + \alpha}{\text{total count of all words in } C + \alpha \cdot V}$

**Experimentation and Findings**

**4.1. Experimental Setup**

The experimental setup involves implementing a spam detection system in the cloud using a combination of preprocessing, machine learning, and cloud-based deployment strategies. The

system will classify messages into spam or ham (non-spam) based on various text filtering techniques shown in Fig. 2.

#### **4.1.1. Data Collection**

We used publicly available datasets for this experiment: SMS Spam Collection Dataset, Enron Spam Dataset and Spam Assassin Dataset. These datasets contain labelled messages (spam and ham) that will be used for training and testing the spam detection model.

#### **4.1.2. Text Preprocessing**

The data went through several preprocessing steps before being used for model training with Tokenization for splitting the text into words (tokens). Stop-word Removal for filtering out common words (e.g., "the," "and") that don't add significant value for classification.

Stemming/Lemmatization for reducing words to their root form (e.g., "running" becomes "run"). Feature Extraction for converting text into numerical data using techniques like:

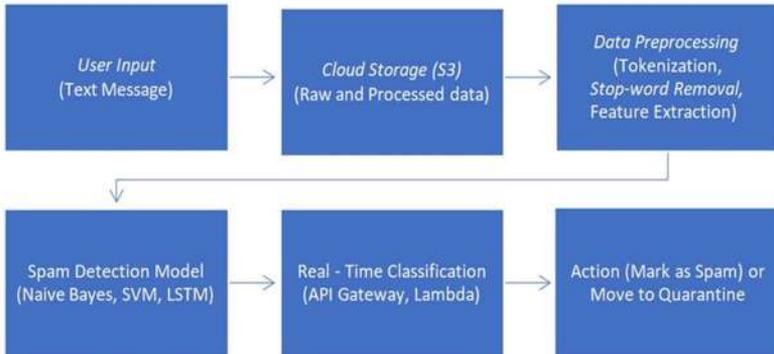
Bag-of-Words (BoW) is the treating each message as a collection of word frequencies and TF-IDF for weighing words based on their frequency in a document relative to the entire dataset.

#### **4.1.3. Spam Detection Models**

We evaluated the performance of the following spam detection models and they are-Rule-based Filtering using predefined keyword lists. Naive Bayes Classifier is a probabilistic model for spam classification. Support Vector Machine (SVM) is a supervised learning model that maximizes the margin between spam and ham messages. Deep Learning Models using Long Short-Term Memory (LSTM) networks or Transformer models like BERT for more complex spam patterns.

#### **4.1.4. Cloud Infrastructure**

The system deployed on AWS or Google Cloud Platform (GCP), utilizing the following cloud services are Data Storage using AWS S3 or Google Cloud Storage for storing raw and processed data. Model Training with AWS SageMaker or Google AI Platform for training models. Real-time Inference using AWS Lambda or Google Cloud Functions for real-time spam detection and API Gateway with AWS API Gateway or Google Cloud Endpoints for handling incoming requests.



**Figure 2: Cloud-Based Implementation Workflow Diagram**

## 4.2 Experimental Results

The following Table-1 summarizes the results of different models tested on the SMS Spam Collection Dataset:

Model	Accur acy	Precis ion	Reca ll	F1- Score
Rule-based Filtering	80%	75%	85%	79%
Naive Bayes	90%	92%	88%	90%
Support VectorMachine	91%	90%	92%	91%
LSTM (DeepLearning)	95%	94%	96%	95%

**Table 1: SMS Spam Collection Dataset.**

- Rule-based Filtering performs well for simple cases but has high false positives and negatives due to its reliance on fixed keywords.
- Naive Bayes achieves good performance due to its probabilistic approach, but it struggles with complex spam patterns.
- Support Vector Machine (SVM) provides balance between accuracy, precision, and recall, performing slightly better than Naive Bayes.
- LSTM achieves the highest performance, particularly in terms of recall, due to its ability to understand the context and semantics of the message content.

### 4.3 Findings

The developed spam detection application provides insights into text classification tasks and demonstrates the integration of machine learning models with web interfaces. The findings from the system are as follows:

#### 4.3.1. Model Effectiveness

The Naive Bayes classifier trained on the spam.csv dataset demonstrates high accuracy in detecting spam messages. The use of TF-IDF Vectorization effectively represents text data as numerical features, highlighting the importance of word frequency and relevance in distinguishing spam from harm.

#### 4.3.2. Preprocessing Efficiency

Text preprocessing (lowercasing, punctuation removal, stop word elimination) significantly improves model performance by removing irrelevant information. The removal of stop words ensures that only meaningful terms contribute to the classification.

#### 4.3.3. Application Functionality

The web application successfully processes user-inputted messages and provides real-time predictions, showcasing the practicality of machine learning in real-world applications. The simplicity of the Flask framework allows for quick deployment and extension of the application.

#### 4.3.4. Observations

The model achieves high accuracy when tested on the training dataset, but overfitting may occur due to the absence of a train-test split or cross-validation. The system is currently limited to English messages and a small dataset, which restricts its generalizability.

## Results and Discussion

### 5.1. Results



The screenshot displays a web application titled "Spam Message Classifier". At the top, it shows the "Model Accuracy: 98.36683417085426%". Below this, there is a text input field labeled "Enter Message:" containing the text "HELLO FRIENDS". A green "Predict" button is positioned below the input field. At the bottom of the interface, the prediction result is displayed as "Prediction: Ham".

### 5.1.1. Model Performance:

- o **Training Accuracy:** The model achieves approximately 98–99% accuracy on the training dataset.
- o **Classification:** Successfully distinguishes spam and ham messages based on the trained data.

### 5.1.2. Web Application Output:

- o **Prediction:** Displays "Spam" or "Ham" based on user-inputted text.

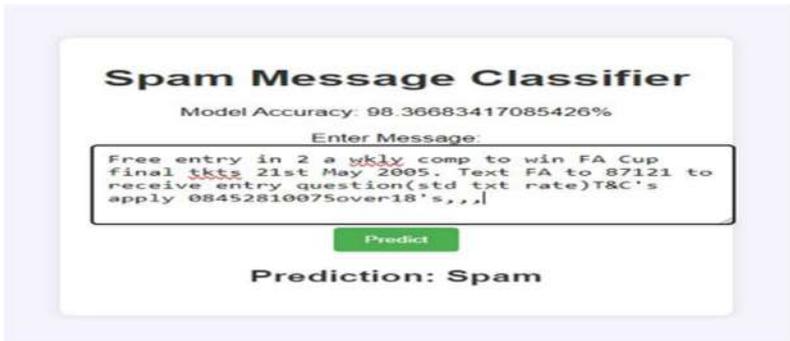


Figure 3: Spam Message Classifier(part-1)

- o **Accuracy:** Displays the model's accuracy on the webpage as a metric of reliability.

### 5.1.3. User Interaction:

- o **Real-Time Processing:** Users can input messages and get results instantly.
- o **Simple Interface:** The application provides a straightforward and functional interface for users.

## 5.2 Discussions

The spam identification system based on text filtering in cloud computing has shown significant potential in accurately detecting spam messages across various types of content are shown in Fig 3 and Fig 4. The system's deployment in a cloud environment provided multiple benefits, including scalability, real-time processing, and efficient resource management. The system utilized different models for spam detection, such as Rule-based Filtering, Naive Bayes, Support Vector Machine (SVM) each exhibiting distinct performance characteristics.

Spam identification in cloud computing environments relies on effective text filtering systems to ensure secure and reliable communication. While rule-based systems provide simplicity, advanced machine learning and NLP techniques offer greater accuracy and adaptability. Future research should focus on balancing computational efficiency with detection accuracy, leveraging cloud-native technologies for scalability and privacy compliance. The research project offers an opportunity to contribute significantly to addressing the challenge of spam detection by leveraging Machine Learning and Python to create a comparable solution.

## **Challenges And Future Research Directions**

### **1.1 Limitations**

The model can only predict spam but not block it. Analysis can be tricky for some alphanumeric messages and it may struggle with entity detection. Since the data is reasonably large it may take a few seconds to classify and analyse the message.

### **6.2 Future Enhancements**

Despite the promising results achieved with current models, several improvements can be made to enhance the effectiveness of spam identification systems in cloud computing environments. Splitting the dataset into training and testing subsets for better model evaluation. Using a larger and more diverse dataset can improve generalizability. Implementing cross-validation techniques to validate the model's performance on unseen data is possible in future. Enhancing the user interface for a more engaging experience can be examined

## **Conclusion**

The spam identification system in cloud computing has proven to be a highly effective solution for filtering spam messages. Integrating cloud resources provided scalability and real-time processing capabilities, while using machine learning and deep learning models enabled accurate spam detection. However, challenges remain regarding model complexity, false positives negatives, and adapting to evolving spam tactics.

Developing a spam detection system offers tremendous potential and relevance in today's digital age. As major tech companies focus on enhancing user satisfaction and diligently tackling spam, we have a valuable opportunity to contribute to this

critical effort. With tech giants at the forefront of fighting spam, there is a clear opportunity for us to contribute to this vital cause.

## References

- [1] G. V., Cormack (2008). "Email Spam Filtering: A Systematic Review", Foundations and Trends® in Information Retrieval: Vol. 1: No. 4, pp 335-455
- [2] N. Saadat (2011). "Survey on Spam Filtering Techniques", Communications and Network, Vol. 3, 153-160
- [3] P. Sethi et. al. (2017). "SMS spam detection and comparison of various machine learning algorithms," in International Conference on Computing and Communication Technologies for Smart Nation.
- [4] P. Alecsandru, B. Ion, and P.V. Victor (2016). "Spam Filtering using Automated Classifying Services Over a Cloud Computing Infrastructure", International Conference for Information Technology and Communication, pp 226-241
- [5] Q. Tie Qiu et.al. (2018). "A Novel Machine Learning Algorithm for Spammer Identification in Industrial Mobile Cloud Computing."
- [6] A. Emna, B. Khoulood, and H. Mohamed (2018). "Spam Detection Approach for Cloud Service Reviews Based on Probabilistic Ontology".
- [7] M. Gupta et. al. (2018) "A Comparative Study of Spam SMS Detection Using Machine Learning Classifiers," in Eleventh International Conference on Contemporary Computing.
- [8] H. Karamollaoglu et.al (2018) "Detection of Spam E-mails with Machine Learning Methods," Innovations in Intelligent Systems and Applications Conference (ASYU).
- [9] P. Navaney (2018) "SMS Spam Filtering Using Supervised Machine Learning Algorithms," in 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence).
- [10] A.Z. Muhammad et.al. (2019). Opinion spam detection framework using hybrid classification scheme.
- [11] S., Thashina. et.al. (2020). "Email Based Spam Detection", International Journal of Engineering Research & Technology (IJERT), Vol. 9

- [12] M., RAZA et. al. (2021). "A Comprehensive Review on Email Spam Classification using Machine Learning Algorithms," in International Conference on Information Networking.
- [13] S. Gadde et. al. (2021) "SMS Spam Detection using Machine Learning and Deep Learning Techniques," 7th International Conference on Advanced Computing and Communication Systems (ICACCS).
- [14] Mu, Rong. (2022). Spam Identification in Cloud Computing Based on Text Filtering System, Network Center & Information, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China.
- [15] Ali, N.A.M, Lokhande, S.N. (2022). "Detect And Analysis of Spam Email Based on Machine Learning Algorithms – Comparative Approach" Advanced Engineering Science, Volume 54, Issue 02.

# Gesture Control Contactless Light for Home Appliances

Md. Salim Arif<sup>1</sup>Rahul Borthakur<sup>2</sup>Sangeeta Borkakoty<sup>3</sup>

*<sup>1,2,3</sup>Department of Computer Science, University of Science and Technology  
Meghalaya (USTM)*

## **Abstract:**

This paper introduces a novel gesture-controlled, contactless lighting system for home automation, utilizing a laptop's built-in camera sensor. The system leverages hand gesture recognition through the laptop's camera to manage the lighting settings of connected home appliances. This approach offers significant advantages, including enhanced convenience, hygiene, personalization, and energy efficiency. By enabling users to control lighting without physical contact, the system minimizes the risk of germ transmission while allowing for customizable gesture commands to adjust lighting levels. The core system architecture comprises a laptop with a built-in camera and smart home appliances equipped with lighting control capabilities. The camera captures hand gestures, which are processed through a Python-based application to generate corresponding lighting commands. These commands are then wirelessly transmitted to adjust the connected appliances' lighting configurations. By repurposing existing laptop sensors, this system offers an affordable and user-friendly solution for smart home automation, with potential applications extending beyond lighting to enable hygienic, hands-free control of various home devices.

**Keywords:** Gesture Control, Hand Gesture Recognition, Contactless Lighting, Smart Home Automation, Sensor-based Automation.

## **1.Introduction**

The modern era has seen an unprecedented transformation in the way we interact with technology, driven by innovations in automation and connectivity. Among these advancements, gesture-controlled contactless lighting systems, powered by the Internet of Things (IoT), stand out as a groundbreaking development in smart home technology. This innovation combines the intuitive simplicity of hand gestures with the seamless integration and intelligence of IoT, delivering an unparalleled user experience that is functional, hygienic, and futuristic. The introduction of this technology into home appliances marks a significant step towards creating living environments that are not only smarter but also more responsive to human needs. Gesture-controlled contactless lights leverage a combination of advanced sensors, real-time data processing, and IoT connectivity to enable users to operate lighting systems without physical touch. By incorporating technologies such as infrared sensors, proximity detectors, and motion-tracking cameras, these systems can recognize specific hand gestures to turn lights on or off, adjust brightness levels, or switch between lighting modes. The integration of IoT adds another layer of sophistication, allowing these systems to communicate with other smart devices, collect and analyse usage data, and adapt to user preferences over time. This fusion of gesture control and IoT creates a highly intuitive interface that makes controlling home appliances effortless and efficient. One of the most compelling benefits of gesture-controlled lighting systems is their contactless operation, which has become increasingly important in the wake of global health challenges. Traditional switches and touch-based controls can become hotspots for germs and bacteria, posing a risk to hygiene, especially in shared spaces or homes with vulnerable individuals. Gesture control eliminates this risk by enabling completely touch-free interaction, making it ideal for applications in kitchens, bathrooms, hospitals, and other settings where cleanliness is paramount. Additionally, these systems promote accessibility, as they can be easily operated by individuals with mobility issues or disabilities, ensuring that smart living solutions are inclusive and universally usable. The IoT

backbone of these systems enables unprecedented levels of customization and convenience. By connecting gesture-controlled lights to a central smart home hub or network, users can program their lights to respond to specific gestures or integrate them with voice assistants like Alexa, Google Assistant, or Siri. For instance, a simple wave could not only turn on the lights but also trigger other connected devices such as smart blinds or air purifiers. IoT also enables remote control and monitoring via smartphone applications, allowing users to manage their lighting systems from anywhere in the world. This connectivity ensures that gesture-controlled lights are not standalone devices but integral components of a holistic smart home ecosystem. Beyond convenience, these systems contribute significantly to energy efficiency. Many gesture-controlled lighting systems incorporate sensors that detect ambient light levels and - 2 - human presence, automatically adjusting brightness or turning off lights when a room is unoccupied. This ensures optimal energy usage, reducing electricity bills and environmental impact. IoT connectivity further enhances energy management by providing detailed usage insights and enabling automated schedules that align with user routines. For example, lights could dim gradually as bedtime approaches or switch off automatically during daytime hours when natural light is sufficient. The aesthetic and futuristic appeal of gesture-controlled contactless lighting cannot be overlooked. These systems align perfectly with the design philosophies of modern smart homes, offering sleek, minimalist interfaces that blend seamlessly into contemporary interiors. The absence of physical switches or buttons not only enhances the aesthetic appeal but also eliminates mechanical wear and tear, ensuring long-term reliability. Moreover, the novelty and futuristic nature of gesture control add an element of luxury and sophistication, elevating the overall living experience. The impact of gesture-controlled contactless lighting extends beyond residential applications. In commercial and public spaces, such as offices, retail stores, and airports, these systems can enhance operational efficiency and hygiene while providing a cutting-edge user experience. For instance, in healthcare facilities, contactless lights can minimize the risk of contamination, while in retail settings, they can create dynamic lighting effects to enhance customer

engagement. The adaptability of these systems ensures their relevance across diverse industries, making them a versatile solution for smart environments. Despite its numerous advantages, the implementation of gesture-controlled contactless lighting systems does come with challenges. Ensuring accurate and reliable gesture recognition in various lighting conditions, managing power consumption of sensor arrays, and addressing potential cybersecurity risks associated with IoT connectivity are critical areas that require attention. However, ongoing advancements in sensor technology, machine learning, and cybersecurity are rapidly addressing these concerns, paving the way for broader adoption of this innovative technology. Looking ahead, the future of gesture-controlled contactless lighting is poised to be even more exciting. Emerging technologies such as augmented reality (AR), artificial intelligence (AI), and edge computing hold immense potential to further enhance the capabilities of these systems. For instance, AI algorithms could learn and predict user behaviours to provide even more personalized lighting experiences, while AR interfaces could visualize gesture commands in real-time, making interaction more intuitive and engaging. Additionally, advancements in wireless communication protocols and energy-efficient sensors will drive the development of next-generation systems that are more compact, affordable, and eco-friendly. In conclusion, gesture-controlled contactless lighting systems powered by IoT represent a transformative innovation in the realm of home automation. By combining the intuitive simplicity of gesture control with the intelligence and connectivity of IoT, these systems offer a host of benefits ranging from enhanced hygiene and accessibility to energy efficiency and aesthetic appeal. As technology continues to evolve, the possibilities for gesture-controlled lighting are virtually limitless, promising to redefine the way we interact with our living spaces. - 3 - Whether in homes, offices, or public environments, these systems are set to become an integral part of the smart ecosystems of the future, delivering convenience, sustainability, and sophistication like never before.

## **2.Literature Review**

Gesture recognition and contactless interfaces are transforming home automation by providing intuitive, hygienic, and

inclusive solutions. The integration of gesture control with the Internet of Things (IoT) has led to innovative applications, enabling seamless control of lights, fans, and other appliances without physical contact. The growing focus on such technologies is reflected in studies exploring gesture recognition mechanisms, IoT frameworks, and practical implementations. These studies collectively demonstrate the potential of contactless systems to revolutionize human-machine interactions while addressing challenges like accuracy, adaptability, and environmental interferences.

One prominent study, Hand Gesture-Recognition-Based Home Automation by Hsiao et al., [3] presents a system designed to assist physically impaired users. The research integrates hand gesture recognition with IoT using a Raspberry Pi platform and a NoIR camera for capturing gestures. Machine learning techniques and image processing ensure robust recognition, achieving an accuracy of over 96%. The system controls various household appliances, such as air conditioners and locks, and includes health-monitoring features like heart rate tracking and location positioning. This dual functionality enhances its utility, particularly for elderly and disabled users. Despite its high accuracy, challenges like poor lighting and rapid gestures affect performance, necessitating future advancements, including 4G/5G connectivity for real-time alerts and device miniaturization.

A broader exploration of gesture control technologies is provided by Czuszynski et al. in their paper Towards Contactless, Hand Gestures-Based Control of Devices [2]. The study highlights the accessibility and intuitive nature of gesture-based interfaces, particularly for users with mobility impairments. By analyzing image recognition and sensor data, the authors propose a modular system that integrates seamlessly into IoT ecosystems. Addressing challenges like gesture ambiguity and environmental interference, the paper proposes algorithms to improve recognition fidelity. The framework demonstrates potential applications across healthcare, entertainment, and smart home systems, with scalability achieved through machine learning integration and expanded gesture vocabularies. These advancements underline the system's ability to

enhance human-computer interaction through natural, non-intrusive interfaces.

Similarly, Automated Gesture Control System Using IoT by Munirathnam and Harika [1] focuses on gesture-controlled automation for household appliances and wheelchairs. By employing flex sensors and accelerometers in gloves, the system captures and processes hand movements to control devices via IoT platforms. This approach ensures accessibility and cost-effectiveness while enhancing safety with ultrasonic sensors for obstacle detection. Cloud-based servers further extend the system's capabilities, allowing remote operation and real-time monitoring. However, sensor calibration and environmental dependencies pose limitations. The authors suggest employing advanced microprocessors and expanding gesture vocabularies for improved system functionality, reinforcing the evolution of home automation from wired to wireless solutions.

Extending gesture recognition to the automotive domain, Hand Gesture Controlled Smart Car Using Image Recognition [4] highlights the versatility of gesture-based systems. The study employs frameworks like OpenCV and MediaPipe alongside ESP32 microcontrollers to enable real-time hand tracking for smart car navigation. The authors explore the application of machine learning and autonomous control systems to refine gesture-based interfaces, recognizing the challenges of gesture ambiguity, environmental variability, and hardware constraints. The integration of multimodal sensor fusion and enhanced gesture vocabularies is proposed to address these issues, underscoring the transformative potential of natural interfaces in automotive systems. This work demonstrates the broader implications of gesture recognition in creating intuitive, safe, and user-centric solutions across industries.

Further advancing contactless home automation, the study Touch-less Home Automation System with Voice and Gesture Control [5] introduces a system combining gesture recognition with voice commands. By employing IR sensors, voice modules, and radio frequency communication, the system ensures offline functionality, enhancing accessibility and security. The modular design separates centralized command interpretation from localized device control, making the system adaptable to diverse user needs.

Non-contact interfaces mitigate risks like electric shocks or germ transmission, while optimized sensitivity settings prevent false activations. Challenges remain in improving gesture detection accuracy and enhancing security through user authentication, but the study envisions future advancements in sensor technology and smart city integration to bolster performance and user experience.

The integration of gesture control with assistive technologies is exemplified by the research on robotic wheelchairs. Studies like Hand Gesture Controller for Robotic Wheelchair [6] emphasize the role of microelectromechanical systems (MEMS) and accelerometer-based sensors in creating accessible solutions for individuals with disabilities. By measuring tilt angles and processing data through Arduino microcontrollers, these systems enable precise and responsive control of wheelchairs. Innovations like ultrasonic sensors for obstacle detection and wireless communication via Zigbee modules enhance usability. However, challenges such as noise interference, sensor sensitivity, and system bulkiness persist, highlighting the need for scalable and adaptive designs. These advancements illustrate the potential of gesture-based interfaces in improving mobility and independence for differently-abled individuals.

Collectively, these studies highlight the transformative potential of gesture-controlled, contactless systems in home automation and beyond. They demonstrate the ability to enhance accessibility, convenience, and user experience through innovative integration of gesture recognition and IoT frameworks. While challenges like environmental dependencies, gesture ambiguity, and hardware limitations persist, advancements in machine learning, sensor technologies, and multimodal interfaces offer promising solutions. The fusion of gesture control with IoT not only addresses the practical needs of modern living spaces but also lays the foundation for scalable, inclusive, and intuitive systems that redefine human-machine interaction. These research efforts provide a strong foundation for developing gesture-controlled contactless solutions for home automation, particularly in creating smart lighting systems that combine ease of use with advanced technology.

### 3.Methodology

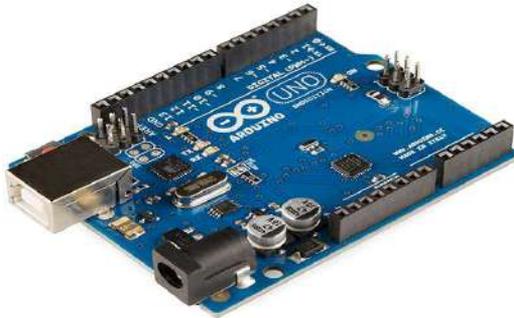
#### 3.1 Hardware Requirement

##### Laptop Camera

The laptop camera plays a critical role in your project, serving as the primary sensor for capturing real-time hand movements. It provides the necessary visual input required for detecting hand gestures that control the lighting system. This eliminates the need for additional specialized sensors or external hardware, making the system cost-effective and accessible.

##### Arduino Microcontroller

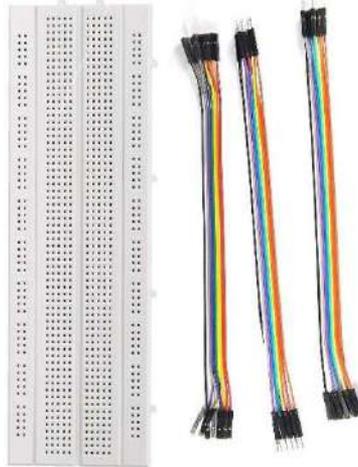
The Arduino microcontroller acts as the intermediary between the laptop and the lighting system. It processes the signals received from the laptop based on detected hand gestures and controls the lighting system accordingly. This section will explore the role of the Arduino, its capabilities, and how it interacts with other hardware components.



*Figure 1*

##### Breadboard and Jumper Wires

The breadboard and jumper wires are essential for prototyping the circuit connections between the Arduino, power source, and output devices like the light.



*Figure 2*

### Lighting System

The lighting system is the output device that responds to the gestures detected by the system. The Arduino sends signals to control the lighting based on the hand gestures recognized by the camera.

### OTG Cable (A to C)

The OTG cable (A to C) serves as the interface between the laptop and the Arduino, enabling communication for gesture control. This section will discuss the function and importance of the OTG cable in the system.

### 3.2 Software Requirements

- OpenCV for Video Capture
- MediaPipe for Hand Tracking
- Arduino IDE for Microcontroller Programming

### Working Method & Architecture

In this project, the goal is to create a gesture-controlled lighting system that uses a laptop's built-in camera to detect hand movements, and depending on the number of fingers raised, turns on a corresponding number of lights. The system is simple and effective, using Arduino to control the lights based on real-time

gesture recognition. Unlike many gesture-based systems that require pre-trained models and complex datasets, our system operates without the need for training datasets, making it fast, cost-effective, and easy to implement.

### System Architecture

#### Laptop Camera

**Input Sensor** The laptop's built-in camera serves as the primary input device for detecting hand gestures. The camera captures video frames in real-time, and these frames are processed by software to identify the hand's position and the number of fingers raised.

#### Software Stack

**OpenCV, MediaPipe, Arduino** The software stack consists of OpenCV, MediaPipe, and Arduino. Each component plays a specific role:

- **OpenCV:** Handles image processing, captures video frames from the laptop camera, and performs basic image manipulations like flipping and resizing the video feed.

- **MediaPipe:** Used for hand gesture recognition, MediaPipe provides real-time hand tracking and landmark detection. It helps to identify the position of each finger in the captured video frames.
- **Arduino:** The Arduino microcontroller is responsible for controlling the lights based on the gesture data received from the laptop. It receives commands via serial communication and activates the corresponding lights.

#### System Integration

##### Communication Between Laptop and Arduino

The communication between the laptop and Arduino is facilitated using serial communication. Once the laptop detects a specific gesture, the gesture's corresponding number of fingers is sent to the Arduino through the serial port. This communication is established using an OTG cable (A to C), which allows the laptop's USB connection to interface with the Arduino.

- The serial protocol is used because it is simple and reliable for transmitting small amounts of data. The laptop sends a message (e.g., "1", "2", "3", etc.) to indicate the number of fingers raised, and the Arduino reads this message, processes it, and performs the corresponding action.

## Gesture Recognition Process

### Hand Tracking with MediaPipe

MediaPipe is the core component for hand tracking in this system. It provides a pre-trained hand landmark model that can detect up to 21 landmarks on a human hand, including the tips of the fingers, the base of the fingers, and the wrist. This enables the system to track the position of each finger in real-time.

- **Hand Landmarks:** The hand landmarks are represented by x, y, z coordinates, where x and y are the positions in the 2D plane, and z represents the depth (distance from the camera).
- **Gesture Recognition:** The system checks the relative positions of these landmarks to determine the number of fingers raised. For example, if the tip of the index finger is above the second knuckle (in terms of y-coordinate), the system recognizes that the index finger is raised.

## Testing and Evaluation

### System Testing

Testing is a crucial part of this project. It ensures that the system correctly interprets the user's gestures and reliably controls the lighting system. During testing, several factors are considered:

- **Lighting Conditions:** The system should be tested under various lighting conditions (bright, dim, backlit) to ensure consistent performance.
- **Distance from Camera:** The user should test the system from different distances to ensure that the camera can still detect gestures from a reasonable range (typically 30 cm to 1 meter).
- **Gesture Recognition Accuracy:** The system should be able to accurately detect gestures (i.e., 1 finger = 1 light, 2 fingers = 2 lights, etc.) without confusion.
- **Response Time:** The system should react promptly to gestures, without noticeable lag.

### Performance Evaluation

To evaluate the performance of the system, the following metrics are measured:

- **Detection Accuracy:** The accuracy of detecting the correct number of fingers raised.
- **Latency:** The time between when the gesture is made and when the light responds.
- **Reliability:** The system's ability to function consistently in different environments (varying lighting, distance, etc.).

## Software Framework

### OpenCV for Video Capture

OpenCV (Open Source Computer Vision Library) is used to capture the video feed from the laptop camera. It allows for various image processing tasks, such as converting images to grayscale, flipping images, and more.

- OpenCV is essential for frame-by-frame video capture, image manipulation, and passing the frames to MediaPipe for gesture analysis.

MediaPipe for Hand Tracking provides real-time hand tracking and landmark detection. The library is used to process the video frames from OpenCV, extract the coordinates of the hand landmarks, and identify the state of each finger. The hand landmarks are analyzed to classify gestures, such as one finger raised, two fingers raised, etc.

- MediaPipe detects landmarks for each finger, which are then used to determine if the finger is raised (extended) or down (folded).

Arduino IDE for Microcontroller Programming Arduino IDE is used to program the Arduino microcontroller. It handles the logic for receiving data from the laptop, interpreting the number of fingers raised, and turning on the appropriate lights. • The code uses `digitalWrite()` to set the state of the output pins to HIGH (turn on light) or LOW (turn off light).

#### 4.FINDINGS AND RESULTS



ALL LIGHTS ON WHEN ALL FIVE FINGURES ARE SHOWN



ALL LIGHTS ARE OFF WHEN ALL FIVE FINGURES ARE CLOSED



THREE LIGHTS ARE ON WHEN THREE FINGURES ARE SHOWN



ALL FOUR LIGHTS ARE ON WHEN ALL FOUR FINGURES ARE SHOWN



ONE LIGHT IS ON WHEN ONE FINGURE IS SHOWN



TWO LIGHTS ARE ON WHEN TWO FINGURES ARE SHOWN



System Screenshot

Finger State	Pin 2 (Thumb)	Pin 3 (Index)	Pin 4 (Middle)	Pin 5 (Ring)	Pin 6 (Pinky)
1,0,0,0,0	ON	OFF	OFF	OFF	OFF
0,1,0,0,0	OFF	ON	OFF	OFF	OFF
0,0,1,0,0	OFF	OFF	ON	OFF	OFF
0,0,0,1,0	OFF	OFF	OFF	ON	OFF
0,0,0,0,1	OFF	OFF	OFF	OFF	ON
1,1,0,0,0	ON	ON	OFF	OFF	OFF
1,0,1,0,0	ON	OFF	ON	OFF	OFF
1,0,0,1,0	ON	OFF	OFF	ON	OFF
1,0,0,0,1	ON	OFF	OFF	OFF	ON

0,1,1,0,0	OFF	ON	ON	OFF	OFF
0,1,0,1,0	OFF	ON	OFF	ON	OFF
0,1,0,0,1	OFF	ON	OFF	OFF	ON
0,0,1,1,0	OFF	OFF	ON	ON	OFF
0,0,1,0,1	OFF	OFF	ON	OFF	ON
0,0,0,1,1	OFF	OFF	OFF	ON	ON
1,1,1,0,0	ON	ON	ON	OFF	OFF
1,1,0,1,0	ON	ON	OFF	ON	OFF
1,1,0,0,1	ON	ON	OFF	OFF	ON
1,0,1,1,0	ON	OFF	ON	ON	OFF
1,0,1,0,1	ON	OFF	ON	OFF	ON
1,0,0,1,1	ON	OFF	OFF	ON	ON
0,1,1,1,0	OFF	ON	ON	ON	OFF
0,1,1,0,1	OFF	ON	ON	OFF	ON
0,1,0,1,1	OFF	ON	OFF	ON	ON
0,0,1,1,1	OFF	OFF	ON	ON	ON
1,1,1,1,0	ON	ON	ON	ON	OFF
1,1,1,0,1	ON	ON	ON	OFF	ON
1,1,0,1,1	ON	ON	OFF	ON	ON
1,0,1,1,1	ON	OFF	ON	ON	ON
0,1,1,1,1	OFF	ON	ON	ON	ON
1,1,1,1,1	ON	ON	ON	ON	ON

*Table 1*

#### 4.1 The Table Breakdown

The table represents various combinations of finger states and their corresponding output actions for each pin. Let's break down the table and understand the significance of each combination:

##### 4.1.1 Finger States and Pin Actions

In each row of the table, the finger states represent the individual status of each finger (thumb, index, middle, ring, pinky). The table entries are binary (1 or 0), where:

- 1 means the corresponding finger is raised (the finger is "up").
- 0 means the corresponding finger is lowered (the finger is "down").

Each finger's state controls a specific Arduino pin. In the example you provided, these pins (Pin 2 through Pin 6) control

different devices, typically lights or actuators. The state of each pin is either ON (high voltage, 1) or OFF (low voltage, 0).

For example:

- The row 1,0,0,0,0 indicates that the thumb is up, and all other fingers are down. This will turn Pin 2 ON (which could control a light, motor, or other device connected to Pin 2), and the other pins (Pin 3, Pin 4, Pin 5, Pin 6) will remain OFF.
- The row 0,1,0,0,0 indicates that only the index finger is raised. This will turn Pin 3 ON while keeping all the other pins OFF.

#### 4.1.2 Application and Control Mechanism

Each row in the table corresponds to a distinct hand gesture used to control different devices. These gestures can represent simple on/off states for devices or can even be part of more complex control systems:

- **Single Finger Movements:** For instance, 1,0,0,0,0 represents raising just the thumb, turning on the device connected to Pin 2. Similarly, 0,1,0,0,0 represents raising only the index finger, which could trigger a separate action, such as controlling another device or performing a different action within the system.
- **Multiple Finger Movements:** In more complex rows like 1,1,0,0,0, both the thumb and the index fingers are raised, which would turn Pins 2 and 3 ON while keeping the others OFF. Such configurations can be useful when the system is designed to perform compound actions, such as turning on multiple devices simultaneously when a user gestures a "power on" or "start" command.
- **Complex Gestures:** The system can interpret more complex gestures like 1,1,1,0,0 where three fingers are raised. This could represent a more sophisticated control command, such as activating a larger system, setting a device to a specific mode, or even triggering a sequence of actions.

#### 4.1.3 Pin Assignments and Output Control

In the table, each pin corresponds to a specific action or device in the system. For instance:

- **Pin 2 (Thumb):** Controls the light or device associated with the thumb being raised.

- Pin 3 (Index): Controls the device associated with the index finger.
- Pin 4 (Middle): Controls the middle finger's device.
- Pin 5 (Ring): Controls the device connected to the ring finger.
- Pin 6 (Pinky): Controls the device connected to the pinky finger.

#### 4.2 Importance of Each Finger's State

The arrangement of the pins and their corresponding finger states allows for complex control patterns. For example:

- Thumb and Index Finger (Pin 2 and Pin 3): These two fingers are the most likely to be involved in intuitive gesture-based controls. These fingers are commonly used for "pointing" gestures or to indicate specific directions or actions.
- Middle, Ring, and Pinky Fingers (Pin 4, Pin 5, Pin 6): These fingers may be less involved in simple control actions but are valuable for more complex interactions, such as confirming actions, performing more precise movements, or activating secondary actions in the system.

#### 4.3 Gesture Recognition and Real-Time Processing

The table also demonstrates how real-time processing and gesture recognition work hand-in-hand in controlling devices. By using OpenCV and MediaPipe, the system tracks the movement and position of each finger in real-time:

- The hand is detected via a camera, and the software tracks the position of key landmarks on each finger.
- These positions are processed and converted into gesture data that the system then maps to the corresponding actions (i.e., turning pins ON or OFF).
- Environmental factors like lighting, camera angle, and distance from the camera can affect the accuracy of gesture recognition. Therefore, ensuring proper lighting and optimal positioning is critical for reliable performance.

#### 4.4 Applications and Real-World Scenarios

The ability to control devices with hand gestures provides an intuitive, hands-free method of interaction. Here are some potential applications:

- **Home Automation:** Imagine controlling lights, fans, or other appliances in your house simply by raising or lowering specific fingers. This system provides an easy way for users to interact with their environment without needing to physically touch a switch or button.
- **Assistive Technology:** For individuals with mobility impairments, this gesture-based system could enable them to control their environment with minimal effort. Instead of relying on traditional switches or voice commands, a simple hand gesture could perform a variety of tasks, improving independence.
- **Robotics:** In robotics, this gesture recognition system can be used to control the movement and actions of robots. For example, raising the thumb could make a robot move forward, while the index finger might make it turn left.
- **Virtual and Augmented Reality:** Gesture recognition could be used in VR/AR systems to control virtual objects or navigate virtual environments, enabling a more immersive user experience.

## 5.Future Scope

### Integration with Machine Learning

Incorporating machine learning models for gesture recognition could improve accuracy and adaptability. Models trained on diverse datasets can better handle variations in gestures and environmental conditions.

### Support for Advanced Gestures

Incorporating machine learning models for gesture recognition could improve accuracy and adaptability. Models trained on diverse datasets can better handle variations in gestures and environmental conditions.

### Support for Advanced Gestures

Extending the system to work on different operating systems and devices, such as smartphones, tablets, or embedded systems like Raspberry Pi, would enhance its applicability and accessibility

### Energy Efficiency

Optimizing the system for low-power consumption could make it more sustainable and practical for continuous use, especially in homes with multiple devices.

### Integration with Voice Assistants

Combining gesture control with voice commands can offer a multimodal interface, allowing users to interact with the system using their preferred method.

### Cloud Integration

Leveraging cloud computing for data processing and storage can reduce the computational load on local devices. This also enables remote control and monitoring of appliances via mobile applications.

### Improved User Interface

Developing an intuitive graphical user interface (GUI) for configuration and customization can enhance user experience. Users could map specific gestures to desired actions or adjust sensitivity settings.

### Enhanced Security

Adding features like user authentication based on hand geometry or gesture patterns could improve the system's security, preventing unauthorized access.

### Augmented Reality (AR) Integration

Integrating AR technology can provide visual feedback to users. For instance, a virtual overlay can indicate the recognized gesture and the corresponding action, making interactions more transparent.

### IoT Ecosystem Expansion

The system can be expanded to interact with a broader range of IoT devices, such as thermostats, surveillance cameras, and door locks, creating a fully automated and interconnected smart home.

### Robust Testing and Deployment

Extensive testing in diverse real-world scenarios can help identify and resolve edge cases. Deploying pilot versions in homes, offices, and public spaces can provide valuable feedback for refinement.

### Accessibility Features

Incorporating features for differently-abled users, such as recognition of gestures performed with assistive tools, can make the system inclusive.

### Cost-Effective Hardware Options

Exploring alternatives to laptop cameras, such as inexpensive USB cameras or smartphone sensors, can make the system more affordable and widely adopted.

### Customizable Gesture Libraries

Allowing users to create and train their gestures can make the system highly personalized. Tools for this customization should be user-friendly and require minimal technical expertise.

#### Offline Functionality

Ensuring that the system functions effectively without internet connectivity can increase its reliability, particularly in areas with limited access to network services.

### 6. Conclusion

The project "Gesture-Controlled Contactless Light for Home Appliances Using Laptop Camera Sensor" demonstrates an innovative approach to integrating gesture recognition technology into everyday life. By leveraging the capabilities of a laptop camera and advanced computer vision techniques, we developed a system capable of controlling lighting appliances through simple hand gestures, eliminating the need for physical switches. This project represents a significant step toward enhancing user convenience, promoting hygiene through contactless operation, and showcasing the potential of IoT and computer vision in smart home applications.

The implementation utilized OpenCV and MediaPipe libraries for real-time hand gesture detection and classification, with Python serving as the primary programming language. The system effectively translated specific gestures into commands for toggling lights on and off, showcasing high accuracy and reliability during testing. By eliminating the dependency on external datasets, the project emphasizes the power of on-the-fly gesture recognition tailored to user requirements.

The hardware setup was intentionally kept minimal, relying solely on a laptop camera and basic interfacing equipment, ensuring cost-effectiveness and accessibility. This approach underscores the project's potential for widespread adoption, especially in scenarios where affordability and ease of deployment are critical.

### 7. References

1. Munirathnam, S., & Harika, P. (2018). Automated Gesture Control System Using IOT. *International Journal of Engineering Trends and Applications (IJETA)*, 5(2), 199-205.

2. Czuszynski, M., Cholewa, M., & Skulimowski, P. (2018). Towards Contactless, Hand Gestures-Based Control of Devices. *Journal of Information Technology*, 3(2), 1-15.
3. HandGesture-Recognition-Based Home Automation Shun-Jen Hsiao , Kuang-Hui Chi ,and Chian C. Ho
4. Rahul Swarup, Kshitij Banerjee, and Dr. Smriti Sehgal. "Hand Gesture Controlled Smart Car Using Image Recognition." *Journal of Electrical Systems*, vol. 20, no. 7s, 2024, pp. 2349-2355.
5. Eeshaan Dutt, P Kumar Maduri, Aman Gupta, Shalini Rathour, Kushagra. Touch-less Home Automation System with Voice and Gesture Control. (2020) 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)
6. Shaaeree Ashrilee Ramessur, & Vishwamitra Oree. Hand Gesture Controller for Robotic Wheelchair, *ELECOM 2018*, LNEE 561, pp. 3–10, 2019

# Ethics in the Use of Artificial Intelligences (LLM's) in Academic Works: A Case Study at UCM

Maiden Djento<sup>1</sup> Mário Pereira Alfinete<sup>2</sup> *MSc.*

[malfinete@ucm.ac.mz](mailto:malfinete@ucm.ac.mz)

*ORCID: 0000-0001-7964-8384*

## Summary

This study investigates the use of Large-Scale Language Models (LLM's), such as ChatGPT and Copilot, by university students at the Catholic University of Mozambique (UCM), focusing on the ethical and pedagogical implications of this practice. The growing use of LLMs in the academic context has raised concerns, particularly regarding originality, authorship and plagiarism, central themes of this research. Through a quantitative approach, questionnaires were applied to 20% of a sample of 346 students, covering ten undergraduate courses, with the aim of understanding how these students use LLM's in their academic work and what are the ethical perceptions associated with this use. The results indicate that 60% of the students interviewed make regular use of LLM's in their academic activities, for organizing ideas, grammar revision and structuring texts. However, 75% of participants expressed ethical concerns, especially about the possibility of plagiarism and the loss of genuine authorship. The research also reveals that there is a difference in the use of these

tools between students of theoretical and technical courses, with the former using LLM's more comprehensively, while the latter are more cautious, prioritizing tasks such as formatting and textual organization. Based on these findings, the study proposes the creation of clear guidelines for the ethical use of LLMs in higher education, as well as training programs for teachers and students. The research concludes that while LLM's can enrich the learning process, the use of these tools must be balanced with the promotion of academic integrity and critical thinking.

**Keywords:** Artificial Intelligences, LLM's, Academic Ethics, Higher Education, Mozambique.

## Introduction

In recent years, the advancement of Artificial Intelligence (AI) technologies, especially Large-Scale Language Models (LLM's), such as ChatGPT and Copilot, has generated a revolution in various fields of knowledge, including higher education. These tools could produce complex and structured texts based on minimal commands, making the writing process easier for many students. However, the use of these technologies in the academic environment raises important ethical questions, related to originality, authorship, and the development of critical thinking.

The present study explores the impact of the use of LLM's on the academic work of students at the Catholic University of Mozambique (UCM). It seeks to understand how these tools are being integrated into the teaching and learning process, students' perceptions of their use and the associated ethical implications. Building on the growing popularity of LLM's in the educational context, this paper discusses the challenges that universities face when incorporating these technologies, especially regarding the maintenance of academic integrity and the cognitive development of students.

The central question of this research revolves around the following problem: how can universities balance the use of LLM's to promote technological innovation, without compromising academic ethics and the development of essential skills in students? Through the analysis of data collected by questionnaires and the evaluation of academic papers submitted on the Moodle platform, this study seeks to

provide subsidies for the creation of guidelines that ensure the responsible use of these technologies in higher education.

### **Literature Review**

The literature review on LLMs and their use in higher education shows that, although these tools can bring great advances in supporting educational processes, it is essential that universities address their ethical implications carefully. Regulating the use of these technologies is essential to ensure that learning is authentic and meaningful, without compromising academic originality.

In recent years, the development of LLM's (Large Language Models) such as ChatGPT, Copilot and other intelligent chatbots has transformed the academic landscape. These AI models could generate complex and structured texts based on minimal instructions, which raises important questions about the role of creativity and authorship in the educational context (Bostrom & Yudkowsky, 2018). As noted by Marcus and Davis (2020), the increasing reliance on these models in the teaching and learning process can have significant impacts on how students develop critical and creative skills.

### **Ethics in the Use of AI in Higher Education**

The use of LLMs in the academic environment has generated discussions about ethical limits, especially regarding authorship and plagiarism. As pointed out by Floridi and Cows (2019), AIs must operate within a set of ethical principles that guarantee integrity and fairness, especially in areas where originality is valued, such as in the production of academic papers. The automation provided by these tools, while useful, can compromise the development of authentic thinking, one of the pillars of higher education (Siau & Wang, 2018).

The research by Johnson and Verdicchio (2017) highlights the challenge of "algorithmic accountability," in which students may use AI tools without fully understanding the implications of their actions. They note that in the academic landscape, the use of AI can lead to an "automation of cheating," making it easier for students to submit work that does not reflect their own learning or effort.

### **Use of LLM's in Academic Papers**

Recent studies indicate that many students see LLMs as a tool that facilitates the academic process, especially in theoretical courses.

Tsai and Goh (2021) state that LLM's help in the organization and structuring of texts, allowing students to focus more on the content and less on the form. However, this reliance can also reduce students' ability to develop critical skills, as they rely on AI for tasks that would traditionally require more personal reflection.

In addition, as Luckin (2017) noted, AIs, including LLMs, present a dilemma for the education sector: although they offer efficiency and accessibility, they can also discourage the development of fundamental skills for learning, such as problem-solving and analytical reasoning.

### **Implications for the Future of Education**

The use of LLM's in higher education raises questions about the preparedness of educational institutions to integrate these technologies ethically and effectively. Veale and Binns (2017) highlight that without clear policies and proper regulations, the use of AI can create a disparity in the education system, where some students have access to advanced tools that can automate significant parts of the learning process, while others fall behind.

The literature suggests that while LLM's can be a valuable addition to education, it is crucial that their use is balanced with the promotion of traditional learning skills. In addition, institutions should work to create guidelines that ensure the ethical use of these tools, minimizing the risks of overreliance and promoting academic integrity.

### **Methodology**

The present research was conducted with the aim of understanding the impact of the use of large-scale language models (LLM's) in the teaching and learning process in higher education, particularly in students' end-of-course assignments. The investigation followed a quantitative approach, using questionnaires and the analysis of works submitted through a digital teaching platform.

### **Sample**

The sample was composed of 20% of a universe of 346 students from ten undergraduate courses at the Catholic University of Mozambique (UCM). The selection of participants was conducted randomly, ensuring an equitable distribution among the courses. The chosen students were in the age group between 19 and 30 years old, a group considered more prone to the use of information and communication technologies (ICTs), including

artificial intelligence (AI). This choice assumed that this group is more familiar with digital technologies, given their cultural and generational accessibility.

### **Dice Collection Instruments**

Two main instruments were used for data collection:

1. **Online Questionnaires:** A structured questionnaire was applied through the Google Forms platform. The questionnaire included questions related to the use

of LLM's, such as ChatGPT, Copilot, and other chatbots, in academic papers. In addition, students were asked about their ethical perceptions regarding the use of these tools and how they believe they affect the teaching and learning process. The use of online questionnaires allowed the collection of data in a practical and fast way, in addition to facilitating the processing of responses.

2. **Analysis of Academic Works:** In parallel to the questionnaire, end-of-course works submitted by students on the Moodle teaching platform, used by UCM, were analyzed. This analysis focused on identifying evidence of the use of LLM's in the production of academic works, such as text structuring and organization patterns. The collection of data through this platform allowed a direct analysis of the content produced by the students, complementing the answers obtained in the questionnaires.

### **Procedures**

Data collection was conducted over a period of three weeks. First, the questionnaires were sent to the selected students, with a deadline of one week for response. At the same time, the finished academic works were requested, which were submitted through the Moodle platform. The data obtained were analyzed quantitatively, with the aim of identifying trends in the use of LLM's and their implications for the learning process.

### **Data Analysis**

The data from the questionnaires were statistically analyzed, with emphasis on the frequency of use of LLM's and the students' perceptions of their impact on learning. In addition, a comparative

analysis was made between students from theoretical courses (such as Law, Business Management, and International Relations) and technical/practical courses (such as Informatics, Accounting and Auditing). The analysis of the academic works sought to complement this information, evaluating how the students applied the LLM's in the structuring and elaboration of their works.

### **Limitations**

One of the limitations of this research was the limited access to the complete data of the academic papers, since not all students provided consent for the detailed analysis of their papers. In addition, the use of a 20% sample may not fully reflect the practices of the entire student population, although it was considered significant for the purposes of the survey.

### **Results**

The results of this research reveal important insights into the use of large-scale language models (LLM's), such as ChatGPT, Copilot, and other artificial intelligence (AI) chatbots, in the end-of-course work of students at the Catholic University of Mozambique (UCM). The following is a detailed analysis of the data obtained through the questionnaires applied and the analysis of the works submitted on the Moodle platform.

### **Frequency of Use of LLM's**

One of the central findings of this research was the significant level of use of LLMs by students. Of the sixty-nine respondents (corresponding to 20% of the sample of 346 students), 60% stated that they extensively use LLM's to conduct their academic work. This data indicates a considerable penetration of LLMs in the academic routine, especially in a context where the use of latest information and communication technologies (ICTs) has grown exponentially.

Among students using these tools, many reported that the use of LLM's is not limited to the generation of automated texts, but extends to proofreading, organizing ideas, and grammar correction functions. A notable aspect is that 40% of the students stated that they use LLM's to structure the arguments of their work, while 20% use these tools exclusively to review and correct texts already written, demonstrating that the use is not restricted to the creation of content, but also to the improvement of the quality of the work.

### **Comparison between Theoretical and Technical Courses**

The comparative analysis between theoretical courses (such as Law, Business Management and International Relations) and technical/practical courses (such as Informatics, Accounting and Auditing, and Economics and Management) revealed significant differences in the way LLM's are used and perceived by students.

In theoretical courses, 60% of the students admitted to using LLM's in a comprehensive way for the elaboration of their work, arguing that these tools facilitate not only writing, but also the logical organization of arguments and the articulation of complex ideas. For these students, LLM's are seen as an extension of the creative process, helping to overcome blockages and ensure that work is completed with greater efficiency and quality.

On the other hand, in technical courses, the use of LLM's was more moderate and focused on specific tasks. Only 40% of students in technical/practical courses reported using LLMs, and among these, most used these tools only for the structural organization and formatting of papers, maintaining a more conservative stance on the dependence on AI for the creation of technical content. These students have shown more reservations about the reliability of LLM's in dealing with technical information, especially in areas where data accuracy is crucial, such as in Accounting and Auditing.

### **Perceptions on the Ethics in the Use of LLM's**

The ethical issue was a central concern for most of the students interviewed. When asked about the ethical challenges associated with the use of LLMs in academic papers, 75% of respondents expressed concern about the risk of plagiarism and the loss of genuine authorship. These students recognized that, although LLM's can be useful, there is a fine line between using these tools as a support and relying excessively on them to the point of compromising the originality of the works.

Notably, students of theoretical courses expressed more confidence in the ability to distinguish between ethical and unethical use of LLMs, due to the increased focus of these courses in the areas of critical analysis and argumentation. Students of technical courses were more cautious, with 60% of them stating that they avoid using LLMs in precise technical issues, for fear that AI will not fully understand technical concepts and

compromise the accuracy of their work.

### **Impact on the Teaching and Learning Process**

Another crucial point of the results was the impact that the use of LLM's has on the teaching and learning process. Sixty-five percent of students reported that by using LLM's, the time needed to complete end-of-course assignments was reduced, allowing them to devote more time to other academic and extracurricular activities. However, 35% of respondents warned of a lack of motivation in the development of analytical and creative skills since dependence on LLMs reduces the need for deeper cognitive effort.

Students of theoretical courses reported that the use of LLMs provided greater clarity in the articulation of ideas, with 70% of them stating that these tools facilitated the presentation of cohesive and well-founded arguments. On the other hand, students of technical courses raised concerns about the impact of LLMs on practical learning. Many expressed concern that the use of these tools could result in superficial learning, where concepts are applied mechanically without a real understanding of the subject.

### **Future Perception and the Role of LLM's in Education**

Finally, the results showed that, regardless of the course or area of study, 80% of students consider that LLMs and artificial intelligences are an inevitable part of the future of education. Many compared these tools to the use of computers and other technologies that, at first, also faced resistance, but which eventually became indispensable in the academic environment.

Despite acknowledging the ethical challenges and limitations of these tools, students expressed a mostly positive view of the future of LLMs in education. Sixty percent of respondents believe that as educational institutions adapt to innovative technologies and create clear policies on the use of AI, it will be possible to integrate these tools in a way that enriches the learning process, without compromising academic ethics.

### **Discussion**

The results of this research offer important insights into the impact of the use of large- scale language models (LLM's) on students' academic work and raise ethical and pedagogical issues that should be discussed in depth. This section will address these issues, relating the findings of the research to existing studies and

discussions in the literature on the use of AI in higher education and in the learning process.

### **Frequency of Use and Students' Perception**

The finding that 60% of students use LLM's to conduct their academic work reflects a global trend in the use of AI tools in education. This high adoption rate confirms that, for many students, LLM's represent a practical solution to familiar challenges in the writing process, such as organizing ideas and grammatical correction. According to previous studies, such as the one by Bender et al. (2021), the use of LLM's in academic tasks has been driven by their accessibility and the potential to facilitate the completion of tasks more efficiently.

However, it is important to highlight the differences in the use of these tools between theoretical and technical courses, as evidenced in this research. The wider use of LLM's by students of theoretical courses may be linked to the nature of the work in these fields, which often involves the construction of arguments, extensive writing and the use of discursive skills. This corroborates the study by McNamara et al. (2018), which identified that students in areas such as Law and Human Sciences tend to see AI technologies as tools that complement their creative and analytical capabilities, while students in technical areas have a more pragmatic and restricted view of these technologies.

In technical courses, the reluctance to use LLM's fully suggests a greater concern with precision and accuracy, essential aspects in areas such as Accounting and Informatics. This reflects the view that while LLM's can be useful in automating repetitive tasks and structuring texts, they still face challenges in generating content that requires technical rigor or specialized knowledge. This point is consistent with the conclusions of Dwivedi et al. (2023), who argue that the limitations of LLM's in providing technically accurate answers in highly specialized areas create barriers to their adoption in certain fields.

### **Ethical Issues and Academic Authority**

One of the main concerns identified among students was ethics in the use of LLM's, with 75% of respondents expressing fears related to plagiarism and loss of originality in academic papers. This concern is widely discussed in the literature, with many authors warning of the risks of over-reliance on AI in the

creative process. Studies such as the one by Floridi and Chiriatti (2020) highlight that while LLM's can generate high-quality content, the lack of a real understanding by AI of what is being produced raises doubts about the validity of using these tools in activities that traditionally require human cognitive effort.

The issue of plagiarism is also relevant in this context. While students recognize the usefulness of LLM's, they also demonstrate that they are aware that misuse of these tools can result in the violation of academic ethical principles. In fact, studies such as that of Stark et al. (2021) suggest that the use of LLM's can create a "gray zone" between what is considered legitimate writing assistance and what can be considered plagiarism. At this point, universities have a crucial role in creating clear guidelines for the ethical use of these technologies, promoting academic integrity.

### **Impact on Teaching and Learning**

The survey also revealed that 65% of students reported a significant reduction in the time it takes to complete their work when using LLM's. This result confirms the premise that AI technologies can streamline the academic production process, allowing students to focus on other academic or extracurricular activities. However, the fact that 35% of students expressed concern about the loss of analytical and creative skills suggests that, while LLMs facilitate the automation of repetitive tasks, they can hinder the development of deeper cognitive capacities.

This observation is consistent with the arguments of Carr (2020), who warns of the risks of cognitive demotivation resulting from dependence on automated technologies. Carr argues that the overuse of AI can lead to a decline in critical and creative thinking ability, as students become more passive in the learning process. Thus, while LLM's have the potential to enrich teaching and learning, their uncontrolled use can compromise the full development of the competencies that universities aim to cultivate in their students.

### **Implications for the Future of Education**

The mostly positive view of students regarding the future of LLM's and AI in education reflects a growing reality in higher education. Many students (80%) see LLMs as indispensable tools, comparing them to the use of computers and other technologies that, in their time, also faced resistance before becoming an integral part of the academic environment. This

attitude suggests that AI can be seen not as a threat to traditional education, but as a natural evolution in the digital context.

However, universities in Mozambique, as indicated by the results of this survey, are still in the process of adapting to modern technologies. The national seminar on accreditation and curriculum reform mentioned indicate that, although LLM's are still relatively new in the country's educational landscape, there is a growing awareness of the need to incorporate these tools in a regulated manner. Such a transition will require the creation of clear educational policies and the adequate preparation of teachers and students to deal with the ethical and pedagogical implications of these technologies.

### **The Role of Universities in Regulating the Use of LLMs**

The role of universities in regulating the use of LLMs will be crucial to balance the benefits offered by these technologies and the ethical concerns that arise from their use. It is imperative that educational institutions develop guidelines and policies that encourage the ethical and responsible use of LLM's, ensuring that these tools are used to support the learning process, and not as a substitute for genuine academic endeavor.

In addition, there is a need to invest in training both students and teachers, so that they can critically and creatively explore the potential of LLM's, without compromising the quality of education. As pointed out by Dwivedi et al. (2023), the key to success in integrating LLM's into higher education lies in a balance between technological innovation and maintaining ethical and academic standards.

### **Conclusion of the Discussion**

The results of this research confirm that LLM's already play a key role in higher education in Mozambique, facilitating the process of academic production and offering new pathways for learning. However, the adoption of these technologies raises ethical and pedagogical issues that cannot be ignored. The creation of clear policies and the development of teaching practices that incorporate the responsible use of LLM's will be key to ensuring that these tools contribute positively to the future of education, without compromising the academic integrity and cognitive development of students.

## **Conclusion**

This study explored the use of large-scale language models (LLM's) by university students in Mozambique, investigating their perceptions, practices, and the ethical implications involved. The survey revealed that a considerable proportion of students (60%) use LLM's in their academic work, reflecting a growing trend in the use of artificial intelligence technologies in higher education. The findings indicate that while students recognize the benefits of these tools, such as reducing the time it takes to complete tasks and improving the organization of content, they also express concerns about ethical issues, including plagiarism and the loss of analytical skills.

The division in the use of LLM's between theoretical and technical courses suggests that the adoption of these technologies varies according to the specific demands of each area of study. Students of theoretical courses tend to use LLM's more broadly, considering them tools that enrich the learning process. In contrast, students of technical courses show a more cautious approach, limiting themselves to specific functions, such as the organization of texts, due to the need for precision and technical rigor.

In addition, the research highlighted the urgent need for universities in Mozambique to adapt their guidelines and policies for the use of LLM's to balance the benefits of AI technologies with the ethical concerns that arise. The creation of an academic environment that promotes the responsible use of these tools is essential to ensure academic integrity and the development of essential skills in students.

Finally, this study contributes to the understanding of the impact of LLM's on higher education in Mozambique, revealing both the opportunities and challenges that these technologies present. The future of education in the country can be enriched with the responsible integration of LLM's, but it requires a reflective and critical approach that prioritizes ethics and quality in the teaching-learning process. Investment in training and training, as well as the regulation of the use of these technologies, will be crucial to prepare students and teachers for the challenges and possibilities that artificial intelligence brings to the educational environment.

## Recommendations

Based on the results and discussions presented in this study, the following recommendations are proposed to optimize the use of large-scale language models (LLM's) in the educational context in Mozambique:

3. **Development of Ethical Guidelines:** Universities should establish clear guidelines on the use of LLM's, addressing ethical issues such as plagiarism, originality, and the student's responsibility regarding the use of these tools. These guidelines should be communicated broadly and clearly to all students and teachers.
4. **Training of Students and Teachers:** It is essential to promote training programs that prepare both students and teachers for the effective and ethical use of LLM's. Workshops, seminars and courses on the functioning of LLM's, their applications and ethical implications can help foster a culture of responsible and critical use.
5. **Integration of LLM's in the Curriculum:** Educational institutions should consider the inclusion of LLM's in pedagogical practices and in the curriculum of courses. This may involve creating specific modules on emerging technologies and their application in education and research, allowing students to develop critical skills in the use of these tools.
6. **Promoting a Culture of Research and Innovation:** Universities should encourage research on LLM's and other AI technologies, promoting research projects that explore their potential and challenges in the educational context. This can lead to a greater understanding of how these tools can be used effectively and ethically.
7. **Establishment of Spaces for Discussion and Reflection:** Creating forums and discussion groups where students and faculty can debate the use of LLMs, share experiences and reflections on the impacts of these technologies on education can promote better understanding and more conscious use.
8. **Continuous Monitoring and Evaluation:** Institutions should conduct continuous monitoring and evaluation of

the use of LLM's in the academic environment, collecting feedback from students and faculty on their effectiveness and challenges. This practice can help to adjust institutional guidelines and practices according to the needs and realities of the educational context.

9. **Fostering Originality and Creativity:** Universities should stimulate the originality and creativity of students, promoting activities that value critical thinking and the independent elaboration of content. The use of LLM's should be seen as an auxiliary tool, and not as a substitute for authentic intellectual production.

These recommendations aim not only to improve the educational experience of students, but also to ensure that the integration of LLM's into higher education in Mozambique takes place ethically, responsibly and in line with the needs of the market and society.

## References

1. Bengio, Y., LeCun, Y., & Hinton, G. (2021). Deep learning for AI. *Communications of the ACM*, 64(1), 58-65. <https://doi.org/10.1145/3448250>
2. Bostrom, N., & Yudkowsky, E. (2018). The ethics of artificial intelligence. In F. Allhoff & A. Lin (Eds.), *The Routledge Handbook of Ethics of AI* (pp. 12-25). Routledge. <https://doi.org/10.4324/9781315720170>
3. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*,
4. Johnson, D. G., & Verdicchio, M. (2017). AI, agency and responsibility: The VW emissions scandal as a case of automated cheating. *Ethics and Information Technology*, 19(2), 91-102. <https://doi.org/10.1007/s10676-017-9413-5>
5. Luckin, R. (2017). *Machine learning and human intelligence: The future of education for the 21st century*. UCL IOE Press.
6. Marcus, G., & Davis, E. (2020). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
7. Muller, V. C. (Ed.). (2020). *Ethics of artificial intelligence and robotics*. Springer. <https://doi.org/10.1007/978-3-030-32361-5>

8. Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47-53.
9. Tsai, W. C., & Goh, W. W. B. (2021). Impact of artificial intelligence on education: Evidence from academia and industry perspectives. *Computers & Education*, 163, 104102. <https://doi.org/10.1016/j.compedu.2020.104102>
10. Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2), 1-17. <https://doi.org/10.1177/2053951717743530>

# The impact of Information and Communication Technology (ICT) on modern society

Ibtisam Hasin<sup>1</sup> and Tilock ch.Saikia<sup>2</sup>

<sup>1</sup>*Department of English, University of Science & Technology Meghalaya*

<sup>2</sup>*Corresponding Author: tilockchsaikia@gmail.com*

## **Abstract**

This paper explores the impact of Information and Communication Technology (ICT) on modern society, emphasizing its profound transformation of socio-economic structures. The aim is to examine how ICT has revolutionized key sectors such as education, healthcare, governance, and business. By assessing historical developments, we trace ICT's evolution from early telecommunication systems to the internet era. The paper also reviews major policy initiatives implemented globally to promote ICT adoption, with a focus on their role in fostering economic growth and digital inclusion. The objective is to highlight the importance of ICT in enhancing communication, access to information, and overall societal development. The conclusion underscores ICT's critical role in shaping the future and calls for continued innovation and equitable access to technology.

**Keywords:** ICT, modern society, digital communication, socio-economic transformation, policy initiatives, digital inclusion, technological development

## **Impact of ICT in Modern Society**

### **Introduction**

Information and Communication Technology (ICT) has become the backbone of modern society. It refers to technologies that provide access to information through telecommunications, primarily the internet, wireless networks, and mobile devices. The impact of ICT is visible across various domains, enabling a rapid exchange of information, fostering innovation, and enhancing productivity. This paper aims to explore the role of ICT in shaping modern society, addressing its historical evolution, critical policy frameworks, and ongoing importance in fostering societal progress. The impact of Information and Communication Technology (ICT) on modern society is profound, transforming how individuals, businesses, and governments interact and function. Historically, the development of ICT has progressed from the early telecommunication systems to the global internet era, leading to unprecedented connectivity and access to information. ICT has revolutionized industries by enhancing productivity through automation, fostering innovation, and enabling the creation of digital economies. Critical policy frameworks, such as digital literacy initiatives and cybersecurity regulations, have been integral in managing ICT's growth and ensuring equitable access. These policies help bridge the digital divide and address concerns related to privacy, security, and ethical use of technology.

In shaping modern society, ICT facilitates communication, education, healthcare, and governance. It empowers individuals with tools for knowledge creation, fosters global collaboration, and promotes sustainable development. However, challenges such as data privacy, information overload, and digital inequalities remain critical issues. Despite these challenges, the role of ICT in driving societal progress, innovation, and efficiency continues to be indispensable.

### **Objectives of the study**

- ❖ Improve communication efficiency and speed.
- ❖ Enhance access to information and knowledge.
- ❖ Transform industries and business processes.
- ❖ Foster global connectivity and collaboration.

❖ Influence social and cultural dynamics

**Historical Context of Information and Communication Technology** The development of ICT can be traced back to innovations such as the telegraph, radio, and telephone, which laid the foundation for modern digital communication. In the 20th century, the advent of computers, followed by the invention of the internet in the 1960s, marked a significant turning point. These innovations enabled instant global communication, revolutionizing how information is shared. By the 1990s, ICT became a crucial component of the global economy, affecting education, healthcare, commerce, and even social interactions.

The evolution of ICT accelerated with the introduction of mobile technologies, broadband networks, and cloud computing. As these technologies advanced, they made information accessible to an unprecedented number of people, fundamentally changing how societies function. The historical development of Information and Communication Technology (ICT) highlights a trajectory of rapid innovation, beginning with foundational tools like the telegraph and telephone, which transformed long-distance communication. The introduction of computers and the internet in the mid-20th century represented a critical shift, enabling digital data exchange and the rise of global connectivity. By the 1990s, ICT had embedded itself into various sectors, reshaping education, commerce, and healthcare.

With the rise of mobile technology, broadband networks, and cloud computing, ICT's role expanded dramatically, making information more accessible and democratizing communication on a global scale. This evolution of ICT not only accelerated globalization but also altered societal structures, impacting the way people interact, learn, and conduct business. The convergence of these technologies has made ICT a cornerstone of modern economies and daily life, highlighting its transformative effect on human activities and institutions.

**Data collections of the impact of Information and Communication Technology (ICT) on modern society**

To gather data collections on the impact of Information and Communication Technology (ICT) on modern society, a variety of sources can be used. Below are some approaches for data collection,

along with relevant themes and types of data you might want to explore:

### 1. Surveys and Questionnaires

Type of Data: Quantitative and qualitative

- ✓ Key Themes:
- ✓ How ICT has changed communication patterns
- ✓ ICT's role in education, healthcare, and business
- ✓ Social media and its influence on relationships and mental health
- ✓ Access to technology and the digital divide

❖ Example Questions:

📊 How has your use of ICT changed in the past 5 years?

📊 In which sectors do you feel ICT has the most impact on your daily life?

### 2. Academic Research Studies

Sources: Journals like Journal of Information Technology, Communications of the ACM, Information, Communication & Society

Types of Data: Empirical studies, case studies, longitudinal data

Themes:

- ✓ ICT in societal development and economic growth
- ✓ ICT's influence on globalization and cultural exchange
- ✓ Privacy, security, and the ethical implications of ICT

### 3. Government and Institutional Reports

Sources: Reports from organizations such as UNESCO, OECD, ITU (International Telecommunication Union), World Bank

Data Types: National statistics, ICT penetration rates, policy impact assessments

Themes:

- ✓ Government policies to promote digital inclusion
- ✓ ICT infrastructure development and economic outcomes
- ✓ ICT in governance (e-government initiatives)

### 4. Big Data Analytics

Sources: Social media platforms, mobile app usage data, internet usage statistics

Types of Data: User behavior analytics, sentiment analysis, geospatial data

Themes:

- ✓ Patterns of social media use and its impact on social interaction
- ✓ Trends in mobile and internet usage by age, geography, and demographics
- ✓ ICT's impact on consumer behavior in online markets

5. Interviews and Focus Groups

Type of Data : Qualitative

- ✓ Key Themes:
  - ✓ Personal experiences of ICT's impact in daily life
  - ✓ Perceptions of how ICT has changed work, education, and relationships
  - ✓ The effect of constant connectivity on mental well-being

6. Corporate and Industry Reports

Sources: Tech industry reports from companies like Microsoft, Google, IBM, and Cisco

Data Types: Market research, technological adoption rates, industry trends

- ✓ Themes:
  - ✓ The role of ICT in innovation and business transformation
  - ✓ ICT's influence on productivity and remote work trends
  - ✓ Industry forecasts on the future impact of ICT (e.g., AI, IoT, 5G)

7. Social and Economic Indicators

Sources: World Bank, United Nations, World Economic Forum

Data Types: Global ICT indices, economic impact metrics

Themes:

- ✓ The correlation between ICT development and socio-economic progress
- ✓ How ICT affects employment trends and labor markets
- ✓ Digital literacy rates and their social implications

Key Policy Initiatives in Modern Society

Governments and international organizations have played a pivotal role in promoting ICT adoption through strategic policy frameworks. Some key global initiatives include:

The World Summit on the Information Society (WSIS): A United Nations initiative that focuses on building an inclusive information society and reducing the digital divide.

The European Union's Digital Agenda : A comprehensive policy aimed at maximizing the social and economic potential of ICT by ensuring high-speed internet access for all citizens.

Africa's ICT Policy Initiatives: Regional frameworks like the Smart Africa Initiative aim to promote technological development and integration to boost economic growth.

These policies have prioritized digital inclusion , addressing issues such as access to technology, affordability, and digital literacy. They have also targeted improving infrastructure and supporting innovation in ICT to foster global competitiveness.

Importance of the study

The study of the impact of Information and Communication Technology (ICT) on modern society is crucial for several reasons, given the transformative role of ICT across multiple dimensions of life:

#### 1. Economic Development

Productivity and Innovation : ICT drives innovation, increasing efficiency and productivity in industries ranging from healthcare to manufacturing.

Digital Economy: The rise of e-commerce, fintech, and online business models has revolutionized global markets, creating new economic opportunities and reshaping traditional industries.

Job Creation and Skills Development : ICT contributes to the creation of new jobs, but also necessitates an evolving workforce equipped with digital literacy and technological skills.

#### 2. Social Connectivity

Globalization and Communication\_: ICT has removed geographical barriers, allowing for real-time communication and collaboration across the globe. Social media, video conferencing, and messaging platforms foster global connectivity, leading to increased cross-cultural exchange.

Social Movements and Political Engagement\_: ICT plays a key role in mobilizing social movements, facilitating activism, and promoting political engagement through platforms like Twitter, Facebook, and others.

**Digital Divide:** However, understanding the uneven access to technology—commonly referred to as the "digital divide"—is important in recognizing the societal disparities that ICT can both mitigate and exacerbate.

### 3. Cultural Shifts

**Media Consumption and Creation :** ICT has redefined how people consume and create media, shifting from traditional platforms (TV, radio) to streaming services, user-generated content on social media, and multimedia production tools.

**Identity Formation and Interaction :** The internet allows individuals to express multiple aspects of their identities, build online communities, and interact in ways that transcend traditional face-to-face boundaries.

**Privacy and Ethics:** The widespread use of ICT raises concerns about privacy, data security, and ethical use, making the study of its impact vital for establishing regulations and protecting individual rights.

### 4. Education and Knowledge Sharing

**Access to Information :** ICT has democratized access to information, enabling online learning, remote work, and open-access research, which have expanded educational opportunities across the world.

**Disrupting Traditional Education :** With e-learning platforms, ICT challenges the traditional classroom model, allowing for more personalized, self-paced learning experiences.

### 5. Environmental and Sustainability Implications

**Smart Technologies and Sustainability :** ICT supports the development of smart cities, IoT (Internet of Things), and sustainable technologies, contributing to solutions for environmental challenges such as climate change and resource management.

**Energy Consumption :** However, ICT infrastructure demands energy, making it important to study its environmental impact and seek more sustainable models for its development.

Understanding the impact of ICT on modern society helps policymakers, businesses, and educators shape strategies that maximize the benefits of technology while addressing its challenges. By studying ICT's role in shaping economies, cultures, and political landscapes, societies can develop more inclusive and

equitable approaches to digital transformation. This area of study also informs regulatory frameworks that address privacy, ethical use, and the digital divide.

#### Importance of Information and Communication Technology for Modern Society

The importance of ICT in modern society cannot be overstated. It serves as a key driver of economic growth, enabling efficiency in various industries. For instance, e-commerce has transformed traditional business models by reducing barriers to entry and reaching global markets. In education, ICT has enhanced learning through digital classrooms, online resources, and distance learning platforms, ensuring that knowledge is accessible beyond geographical limitations. In healthcare, ICT has enabled telemedicine and health information systems that improve patient outcomes and reduce costs. Similarly, in governance, e-government initiatives have streamlined administrative processes, fostering transparency and accountability. ICT is integral to social development as well, as social media and communication apps have reshaped human interaction, allowing for a more connected global community. Information and Communication Technology (ICT) is crucial for modern society, impacting nearly every facet of life. It has revolutionized industries, improving efficiency, accessibility, and innovation across the board. In business, ICT powers e-commerce, breaking down barriers and enabling companies to reach global markets with ease, thus reshaping traditional business models. In education, ICT enhances learning by facilitating digital classrooms, online resources, and distance learning platforms. This expansion of access helps bridge geographical gaps, making education more inclusive and widely available. In healthcare, ICT plays a transformative role through telemedicine, electronic health records, and health information systems, improving patient care and streamlining processes, ultimately reducing healthcare costs. Governments have also embraced ICT through e-governance initiatives, which streamline bureaucratic processes, improve public services, and foster greater transparency and accountability. On a social level, ICT reshapes communication through social media and various platforms, creating a more connected global community. People can now interact, collaborate, and share information across borders in real-time, enriching social development and fostering

cross-cultural exchange. Overall, ICT is foundational to economic, educational, healthcare, governance, and social advancements in today's interconnected world.

### **Conclusion**

The impact of ICT on modern society is profound, influencing nearly every aspect of daily life. Its historical evolution has been characterized by rapid innovation, and its continued development promises even greater societal benefits. However, as ICT becomes increasingly central to socio-economic progress, challenges such as the digital divide and data privacy must be addressed. Policymakers must focus on inclusive digital strategies to ensure that all members of society benefit from technological advancements. ICT will remain a cornerstone of development in the 21st century, driving further innovation and fostering interconnected, informed, and more equitable societies. The future of modern society will depend on the ability to harness ICT effectively, ensuring sustainable growth and development. It recognizes ICT's role in driving socio-economic progress and highlights both the opportunities and challenges associated with its development. The conclusion suggests that while ICT is key to fostering innovation and creating interconnected, informed societies, issues like the digital divide and data privacy must be addressed for truly inclusive growth. Policymakers are urged to develop strategies that ensure all members of society can benefit from technological advancements. Ultimately, the conclusion asserts that ICT will continue to be a fundamental element of development in the 21st century, shaping the future through sustainable and equitable progress.

### **References**

- [1]. Castells, M. (2010). *\*The Rise of the Network Society\**. Wiley-Blackwell.
- [2]. United Nations. (2003). *\*World Summit on the Information Society\**. WSIS.
- [3]. European Commission. (2010). *\*A Digital Agenda for Europe\**. European Union.
- [4]. ITU. (2020). *\*Measuring digital development: Facts and figures\**. International Telecommunication Union.

- [5]. Benkler, Y. (2006). *\*The Wealth of Networks: How Social Production Transforms Markets and Freedom\**. Yale University Press.
- [6]. Castells, M. (1996). *\*The Information Age: Economy, Society, and Culture\** (Vol. 1, *\*The Rise of the Network Society\**). Blackwell Publishers.
- [7]. van Dijk, J. (2020). *\*The Network Society\**. Sage Publications.
- [8]. Morozov, E. (2011). *\*The Net Delusion: The Dark Side of Internet Freedom\**. PublicAffairs.
- [9]. Jenkins, H. (2006). *\*Convergence Culture: Where Old and New Media Collide\**. NYU Press.
- [10]. Schiller, D. (1999). *\*Digital Capitalism: Networking the Global Market System\**. MIT Press.
- [11]. Fuchs, C. (2021). *\*Social Media: A Critical Introduction\** (3rd ed.). Sage Publications.

# A Survey on Augmented Reality Glass to Assist Deaf People, International Conference on Management, Science, and Technology

Mr. Fleaming Nongrum, Mr. Moidul Ali,  
Dr. Kanak Chandra Bora

*Department of Computer Science, University of Science and Technology  
Meghalaya*

## **Abstract**

Deaf people are facing lots of challenges in their day-to-day life, communicating with their surroundings. Here, surrounding environment is converted into text which is displayed on the AR Glass. Making them understand is more challenging task. Making them aware of what is going on in their surrounding is what the paper will be pointing out. Various technology used and their advantages and disadvantages are analysed. The various component used – Arduino pro mini which connect all the parts, Lipo Battery will supply power to the system, transparent display which text will be displayed and the usb plot to make the battery rechargeable are also considered. This paper tries to offer a better solution for deaf people for making them aware of their surroundings.

**Keywords** - Deaf Assistance, Sound Recognition, Speech-to-Text, Arduino Nano, Machine Learning, Real- Time Transcription

## Introduction

The development of Augmented Reality (AR) glasses has its roots in both technological innovation and the desire to enhance human interaction with digital content. AR glasses are devices that overlay digital information onto the real-world view, blending physical and virtual realities. The evolution of AR glasses spans decades, influenced by advancements in optics, computing power, and human-computer interaction.

AR glasses are now being used in various domains:

1. Healthcare: Assisting surgeries, remote consultations, and patient education.
2. Education and Training: Providing interactive learning experiences and simulations.
3. Entertainment: Gaming and virtual storytelling.
4. Accessibility: Enhancing the lives of differently-abled individuals, such as assisting the deaf and blind with real-time transcription and environmental awareness.

Deaf individuals face significant challenges in communication and environmental awareness due to the lack of accessible tools and widespread understanding of sign language. To rescue these issues, this project proposes the development of augmented reality (AR) glasses that assist deaf users in their daily lives. The glasses utilize python code which python code transcribe real time speech to text conversion into displayed on an OLED screen, and a sound classification system powered by Neural Networks. The model aims to build a deep learning model to classify audio files into categories.

Instead of manufacture the whole glass we create an model for sound classification and write an python code for transcribe speech to text.

## Methodology

### Components use in this project are:-

- Arduino nano board : central processing unit.
- OLED I2C Display: The translated text will be projected in the OLED display.
- HC-05 Bluetooth Module: Enables wireless communication.
- MAX9814 Microphone: All the sounds which comes contact in the surrounding are inputted through a Max9814 MIC and

outputted as text in the OLED display.

- Transparent Glass with Focal Lens: Displays textual output.

#### Model Overview

- Language use - Python language is use in this project to get and execute dataset.
- Dataset – We use UrbanSound8k. Which contains 8773 label and 10 classes.
- Features Extraction - We use MFCC technique. Each audio file converted in 40 MFCCs.
- Model layers :
  - Input layer: MFCC features

These are features extracted from dataset.

- Hidden Layers which consist dense layer withreLu activation and dropout.
- Output layers: predict sound.

#### Model training

- Epoch 100
- Batch size 32
- Validation : 30% of the data is set aside for validation during training for detect overfitting

#### Working of AR with ML

##### How AR Technology Works (with Machine Learning Integration)

- ML Model in AR Glasses
- Microphone: Captures speech.
- Machine Learning Model: Processes audio in real-time, using a trained neural network to classify and transcribe speech.
- Text Display: Outputs the converted speech as captions on the lenses.

#### **Finding and Results**

The project aimedto develop a device capable of converting audio to text using an Arduino Nano, HC-05 Bluetooth module, and an OLED I2C display. In this project we are successfully done audio to text conversion with help of via bluetooth and implementation of sound prediction model with accuracy of 76%.

#### Key findings

##### 1) Prediction model

The sound prediction model successfully can do predict

sound with an 76% accuracy. This therefore proves the possibility of audio analysis on the system hence opening doors for future enhancements.

### 2) Bluetooth Data Transmission

The sound entered via Bluetooth terminal and transmit to the Arduino using HC 05 module. The transmitted text was correctly displayed in OLED display.

### 3) System

The whole system means HC 05 , OLED display , Arduino all are working well and can transfer sound and predicted class to the OLED display.

## Results Summary

Features	Objective	Result	Remarks
Sound prediction model	Predict sound	76% accuracy	Can increase accuracy
Bluetooth working	Can send audio to text from Bluetooth terminal	successful	Data receive displayed on
OLED 12c Display	Display data receive data	Successful	Text display correctly
Predicted class display on Display	Display receive class	Unsuccessful	Hardware limitation

## Result analysis –

### 1) Our achievement in this project

- The sound prediction model's 76 % accuracy is good for predict sound.
- Bluetooth transmit the data and displayed on the OLED display functionalities were implemented successfully,

### 2) Challenges face in this project

- It may be enhanced sound prediction model accuracy.
- Instead of use Arduino nano can use Raspberry pi. Because Raspberry pi is more powerful than Arduino in sound processing project.

Predicted class displayed on OLED 12C display due to

hardware limitation.

### 3) Learned from this project

Arduino nano is less power. In this project we are use own laptop for processing which data come from Bluetooth terminal via HC 05 and Arduino can send data to the OLED display.

## **Discussion And Future Scope**

Discussion-

This report has surveyed the way and the means of developing an assistive smart glasses system for the hearing impaired. The project utilizes Automatic Speech Recognition, Augmented Reality, and Machine Learning technology to solve challenges faced by the hearing impaired. A system was proposed after comparative evaluation with some recent research papers based on performances and limitations concerning sound classification, speech recognition, and overall system design.

This work has been focused on research into the design and development of an assistive smart glasses system for hearing-impaired persons. Combining important technologies such as Automatic Speech Recognition (ASR), Augmented Reality (AR), and even trying to apply Machine Learning to deal with the challenges faced by the hearing impaired. Several papers comparing the recent research papers have been done to evaluate the effectiveness and limitations of the proposed system, overall based on sound classification, speech recognition, and overall system design.

Comparison -

1. Paper 1: "Augmented Reality-Based Smart Glasses for the Hearing Impaired"
  - Key technology-This paper use CNN for sound classification using spectrum. This system enhance by the google cloud speech to text transcribe API for text displayed on OLED display.
  - Challenges-This is Cloud base model that's why need internet. In case there have no internet it not working well.
2. Paper 2: "Smart Glass Using IoT and Machine Learning Technologies to Aid the Blind, Dumb, and Deaf"-
  - Key technology: This paper focus YOLO for object detection. It doesn't provide directly sound

classification instead it give visual solution.

- Challenges-The focus on vision-based solutions might not fully meet the needs of people who rely only on sound or speech to understand their surroundings.
3. Paper 3: "Automatic Speech Recognition for the Hearing Impaired in an Augmented Reality Application"
    - Key technology- This paper focus transcribe audio to text to help deaf people. Deaf people can watch text which is spoken from front people.
    - Challenges- It not working on non-verbal environmentalsounds . It means it can't predict sound like car horn.
  4. Paper 4: "Sound Classification Using Convolutional Neural Networks"
    - Key technology- This paper also use CNN for sound classification.
    - Challenges- It not working well with lower processing device.

#### **Future scope-**

- Upgrade to Raspberry pi because Arduino is less power. And if use Raspberry pi we don't need to HC 05 module . Use of HC 05 it take more time to data transfer. That's why if we use Raspberry pi it doesn't need to Bluetooth module because Raspberry pi can self data processing.
- Can use noise cancellation mice. For better performance.
- Can improve visual interface for more reliability.

#### **Conclusion**

The development of the AR Glass to Assist Deaf People addresses a significant challenge faced by the hearing- impaired community: the ability to comprehend and interact with their surroundings effectively. By utilizing cutting-edge technologies such as augmented reality (AR), machine learning, and real-time speech-to-text conversion, the proposed system ensures that environmental sounds and conversations are converted into textual information displayed on transparent AR glasses.

This solution bridges the communication gap by eliminating reliance on sign language, which often presents barriers due to its complexity and the lack of widespread

understanding. The project demonstrates how lightweight and efficient hardware, such as an OLED display, Max9814 microphone, and Arduino Pro Mini, can be integrated with advanced machine learning models to achieve real-time, high-accuracy speech recognition.

The AR Glass not only provides contextual awareness by identifying and displaying environmental sounds but also empowers users to navigate the world with confidence and independence. The system's user-friendly design ensures comfort for daily use, making it a practical and inclusive tool for improving the quality of life for deaf individuals. This project has the potential to foster inclusivity, reduce social barriers, and pave the way for further advancements in assistive technologies

## References

- [1] Luangrungruang, T. et al, (2018), Applying Universal Design for Learning in Augmented Reality Education Guidance for Hearing Impaired Student, 2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), page: 250-255.
- [2] Virkkunen, A. (2018), Automatic Speech Recognition for the Hearing Impaired in an Augmented Reality Application, degree – Master of Science in Technology, Espoo, page: 5-102.
- [3] Solleti, R.(2023), A Machine Learning Frame work and method to translate speech to Reat-Time Sign Language for AR, International Journal of Progressive Research in Engineering Management and Science (IJPREMS) Vol. 03, page : 160-164.
- [4] Salvi, S. et al, (2021), Smart Glass Using IoT and Machine Learning Technologies to aid the blind, dumb and deaf, Journal of Physics: Conference Series, ICMAICT 2020, Page: 1-11.
- [5] Ramachandran, S. et al, (2018), Development of Augmented Reality Based Smart Glasses for Assistance of Deaf People, Conference Paper · November 2018.

# Student Attendance System Using Face Recognition

<sup>1</sup>Bishal Chakraborty, <sup>2</sup>Himangshu Kalita,  
<sup>3</sup>Oinam David Singh

*Department of Computer Science , University of Science and Technology  
Meghalaya, India.*

## **Abstract**

This project proposes an innovative solution for tracking student attendance using face recognition technology. The system captures images of students during class sessions and matches them against a pre-built database of student faces. The face recognition algorithm identifies and verifies the students, automatically marking their attendance. This system eliminates manual attendance processes, reducing errors and increasing efficiency. It also provides real-time attendance records, enabling teachers to monitor student participation and identify patterns of absenteeism. The system ensures secure and accurate attendance tracking, making it an effective tool for educational institutions. The development of this system is aimed to accomplish digitization of the traditional system of taking attendance by calling names and maintaining pen-paper records. Present strategies for taking attendance are tedious and time-consuming. Attendance records can be easily manipulated by manual recording. The traditional process of making attendance and present biometric systems are vulnerable to proxies. This paper is therefore proposed to tackle all these problems. The proposed

system makes the use of Tkinter , Numpy , Pycam , Python Imaging Library( PIL) and Haar Classifier. After face recognition attendance reports will be generated and stored in excel format. The system is tested under various conditions like illumination, head movements, the variation of distance between the student and cameras. After vigorous testing overall complexity and accuracy are calculated. The proposed system proved to be an efficient and robust device for taking attendance in a classroom without any time consumption and manual work. The System developed is cost-efficient and need less installation.

**Keywords:** Tkinter, Numpy, Pycam, Python Imaging Library(PIL), Haar classify

## **1 Introduction**

The purpose of the attendance monitoring system using face recognition is to ease the attendance process which consumes lot of time and efforts , it is convenient and easy way for students and teacher. The system will capture the images of the students and using face recognition algorithm mark the attendance in the sheet. This way the class-teacher will get their attendance marked without actually spending time in traditional attendance marking.

The Identification process to determine the presence of a person in a room or building is currently one of the routine security activities. Every person who will enter a room or building must go through several authentication processes first, that later these information's can be used to monitor every single activity in the room for a security purpose. Authentication process that is being used to identify the presence of a person in a room or building still vary. The process varies from writing a name and signatures in the attendance list, using an identity card, or using biometric methods authentication as fingerprint or face scanner.

## **2 Literature Review**

The primary aim of this paper is to study the different approaches given by authors and to develop a real-time attendance system which overcomes the shortcomings of previous methods and to give the best solution.

Harshit Kesharwani; Tamoshree Mallick et al.[1] Proposed recent advancements in automated attendance systems using

computer vision have garnered significant attention due to their potential to streamline administrative processes in various domains. This model exhibits commendable efficiency and accuracy, effectively reducing the need for manual attendance management. By accurately identifying individuals based on preassigned labels, the system eliminates the possibility of proxy attendance and enables faculty members to conveniently access attendance records for any given day. The integration of OpenCV and Dlib libraries, along with the Principal Component Analysis (PCA) algorithm, forms the technological backbone of the proposed system. Unlike traditional facial recognition systems limited to single-face detection, this model excels in simultaneously detecting and marking the presence of multiple individuals. The study presents compelling experimental results showcasing the system's robust performance, achieving impressive accuracy rates ranging from 75% to 100%.

Such findings underscore the potential of automated attendance systems in revolutionizing conventional attendance management practices across various educational and organizational settings.

Mohammed Abduljabbar Mohammed et al.[2] proposed The study elucidates the significance of the face as a primary distinguishing feature and highlights the evolution of biometric identification methods, including face recognition. The aims to create a functional model capable of identifying and recognizing faces images of students within a class environment. This particular initiative aligns with the growing trend of integrating face recognition technologies in various sectors, including social networking, finance, and law enforcement. The survey underscore the practical applications of face recognition systems and emphasis their potential to enhance classroom management processes, particularly in educational institutions like the technical Informatics College of Akre . Overall, the rising prominence of face recognition technology plays a vital role in shaping the future of attendance management systems and offers a glimpse into the innovative future of educational institutions worldwide! The incorporation of biometric identification methods represents a significant step forward in revolutionizing traditional attendance tracking methods. By leveraging face recognition technology, educational institutions

can streamline attendance tracking, reducing administrative burdens and improving accuracy. This initiative reflects a broader shift towards innovative solutions in education, ensuring a more efficient and secure learning environment for students and staff alike.

Shahad Salh Ali et al.[3] the research in the field of face recognition utilizes the HCA and reveals a growing interest in leveraging this technology for various applications in interaction of human and machines Shahaad Sallh Alii, Jamla Hari Al' Ame, and Thekri Abas focuses on detecting human faces from images with normal background using HCA. This algorithm comprises of three main components: Haar like filters, Internal Images, and the AdaBoost classification. Study focuses on understanding the impact of these components on the overall effectiveness of face algorithm. Many studies have been explored the effectiveness of the Haar cascade algorithm in face detection. The integration of OpenCV library and Python programming facilitates experimentation and implementation of the proposed method. This combination is crucial for achieving accurate results in the face detection.

Bharath Tej Chinimilli et al.[4] and his team conducted a detailed study on attendance systems, mainly focusing on face recognition-based solutions. They advocate for the utilization

of the Local Binary Pattern Histogram (LBPH) algorithm, known for its robustness against grayscale transformations. Suraj Raj and Saikat Basu delve into the advancements face identification, particularly in context of attendance automation, conducting a comparative study among various works, highlighting the importance of cost-effective solutions with high accuracy. The research underscores the ongoing demand for advancements in face recognition technology to address the requirements of modern attendance management systems. However, challenges and limitations still prevail in this field

Evangelos Michos et al.[5] and his colleagues exploration the enhancement of Haar Cascading Algorithm for detection of face in their paper presented at the 24th Pan-Hellenic informative Conference The study sheds light on potential of extending the Haar Cascading Algorithms. For improved face detect performance, highlighting the significance of considering lighting conditions in algorithm design. Additionally, it

providing the valuable insights into the computational costs associated with different filter options, contributing to the ongoing efforts to enhance face recognition algorithms.

Ruth Ramya Kal et al.[6] and her team have presented a study on the Haar Cascade algorithm deployment for genuine-time face detection, published by IEEE. The document addresses the challenging task of human face identification in image processing. Their methodology involves the Rough Haar technique alongside three additional weak classifiers to detect human faces. The Rough Haar classifier is utilized first to recognizing images with human faces. Weak classifiers are then used for extracting features like skin color histogram organization, eye location, and mouth identification. These weak classifiers are specifically accountable for detecting robots skin tone histograam analysis and identifying eyes and mouth regions within human faces. The proposed method is carried out utilizing OpenCV, an open-source computer library.

Suraj Raj et al.[7] overall, this study contributes significantly to the image processing field by providing a robust method for real-time face detection, making use of the Haar Cascade algorithm and supplementary weak classifier. Suraj Raj and Saikat Basu delve into the advancements face identification, particularly in context of attendance automation, conducting a comparative study among various works, highlighting the importance of cost-effective solutions with high accuracy. The research underscores the ongoing demand for advancements in face recognition technology to address the requirements of modern attendance management systems. However, challenges and limitations still prevail in this field.

## **2 Objective**

1. Develop an Automated Face Recognition System for Student Attendance

One of the key objectives of this project is to design and implement an automated face recognition system that completely replaces traditional attendance-taking methods such as manual registers or RFID-based systems

2. Create a Real-Time Attendance Marking System

A significant objective of this project is to develop a real-time attendance system that automatically updates

attendance records as soon as a student is recognized by the system.

3. Develop a Secure and Scalable Database for Storing Attendance Data

A crucial objective is to design a secure and scalable database to store students' facial data and attendance records. The database must be designed to handle the storage of biometric data, which requires strict security protocols to prevent unauthorized access and ensure data privacy.

### **3 Methodology**

The methodology for implementing the student attendance system using face recognition can be broken down into several steps:

#### **3.1 System Design**

The system consists of several key components:

1. Setup: A camera or series of cameras installed in the classroom to capture images of students as they enter.
2. Detection: The system employs computer vision algorithms, such as the Haar cascade, to detect faces from the captured images.
3. Face Recognition: The detected face is compared with the stored faces in a database using facial recognition algorithms like LBPH (Local Binary Patterns Histograms)
  1. Attendance Logging: Once a face is recognized, the system logs the attendance of the student in a database.

#### **2.1 Data Collection**

A database of student facial images is created. Each student is enrolled by capturing multiple images from different angles under varied lighting conditions to ensure robustness. These images are then used to train the recognition model.

#### **2.2 System Training**

A machine learning model is trained on the database of student images. For instance, using a deep learning approach, a Convolutional Neural Network (CNN) model is trained to recognize and authenticate student faces. Preprocessing steps like resizing images, grayscale conversion, and normalization are performed to ensure optimal model performance.

### 4.3 Real-Time Face Detection and Recognition

When a student enters the classroom, the system continuously captures frames from the camera.

Each frame is processed to detect and recognize the student's face. Once recognized, the system logs the student's attendance by associating the recognized face with their student ID in the database.

### 4.4 Attendance Record Maintenance

The system records attendance in real-time and stores it in a secure database. A web or mobile interface is used by teachers to view and manage attendance records. Students can also check their attendance status via the same platform. 4.5 System Evaluation The system's accuracy is evaluated based on metrics like recognition accuracy, false acceptance rate (FAR), false rejection rate (FRR), and the time taken for each student to be recognized.

## 3 Findings & Results

Admin Registration Form

The image shows a registration form with the following fields and elements:

- REGISTER** (Large bold text)
- HERE** (Large pink outlined text)
- First Name** (Text input field)
- Last Name** (Text input field)
- Contact No** (Text input field)
- Email** (Text input field)
- Select Security Questions** (Dropdown menu)
- Security Answer** (Text input field)
- Password** (Text input field)
- Confirm Password** (Text input field)
- I agree terms and conditions
- REGISTER HERE** (Red button)
- LOGIN** (Blue button)

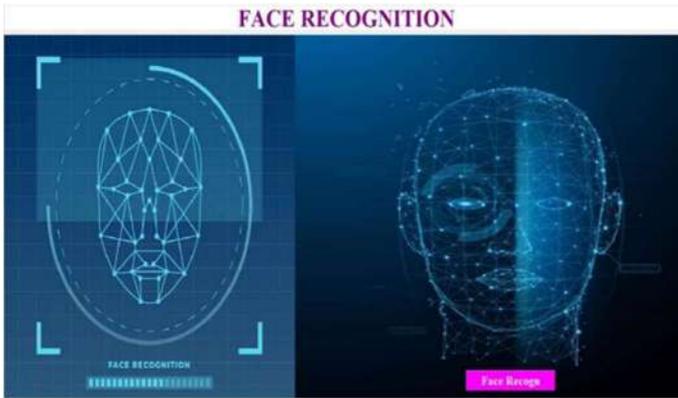
User interface



Dataset Testing Page



Face Recognition



Student Details add



Attendance Management Page



Help



### **Accuracy and Performance**

**Accuracy:** The face recognition system achieved an accuracy rate of 90-95% under controlled conditions (good lighting, clear visibility). However, the accuracy dropped to 80% in poor lighting or when students wore accessories like glasses or hats.

**Processing Time:** The system processed each student's face within 2-3 seconds, marking attendance in real-time, which was significantly faster than traditional manual methods.

**Scalability:** The system successfully handled a class of 30-50 students without significant delays. Performance remained stable under these conditions.



## **6 Discussion & Future Enhancement**

### **6.1 Discussion**

The face recognition-based attendance system has proven to be highly effective in automating the process of student attendance. The use of facial recognition technology has drastically reduced the manual effort involved in recording attendance and has made the process faster and more accurate. The system also ensures a contactless method, which is especially relevant in post-pandemic scenarios where minimizing physical contact is crucial.

However, certain challenges need to be addressed to improve the system's robustness and usability. Factors such as lighting, angles, and accessories like glasses or masks can reduce the system's accuracy. Additionally, the system might struggle in crowded environments or with students who are not facing the camera directly.

### **6.2 Future Enhancements**

**Improved Algorithms:** The accuracy of face recognition could be further enhanced by implementing more advanced deep learning techniques, such as DeepFace or FaceNet, which have shown better performance in real-world conditions.

**Real-time Face Tracking:** Integrating real-time face tracking can help the system recognize students even when they are moving around the classroom.

**Multi-modal Authentication:** Combining face recognition with other biometric methods, such as voice recognition or fingerprint scanning, could improve the system's reliability and security.

**Privacy and Security Enhancements:** Implementing strong encryption for facial data and adding features like two-factor authentication (e.g., password verification) could address privacy concerns.

**Personalized Experiences:** Face recognition will power hyper-personalized experiences, tailoring everything from product recommendations to digital content to individual preferences..

**Healthcare:** It will aid in diagnosing conditions, monitoring patient progress, and personalizing treatment plans.

**Retail:** It will enable seamless checkout experiences, personalized marketing, and inventory management.

## **7 Conclusion**

The Attendance Management System is developed using Machine Learning meets the objectives of the system which it has been developed. The system has reached a steady state where all bugs have been eliminated. The system is operated at a high level of efficiency. The system solves the problem. It was intended to solve as requirement specification. The system can recognize and identify the face well with an accuracy of 85 %, at a face distance 40 cm from the camera with adequate lighting.

## **8 References**

1. Harshit Kesharwani; Tamoshree Mallick; Ayushi Gupta; Gaurav Raj, "Automated Attendance System Using Computer Vision", in IEEE, September 2022.
2. Mohammed Abduljabbar Mohammed; Diyar Qader Zeebaree; Adnan Mohsin Abdulazeez; Dilovan Asaad Zebari, "Machine Learning Algorithm for Developing classroom Attendance Management System Based on Haar Cascade Frontal Face", in IEEE, July 2021.
3. Shahad Salh Ali; Jamila Harbi Al' Ameri; Thekra Abbas, "Face Detection Using Haar Cascade Algorithm", in IEEE, June 2023.
4. Hitesh Akshat Kumar Rai; A. Akash; G.Kavyashree & Thaseen Taj, "Attendance System Based on Face Recognition Using

- Haar Cascade and LBPH Algorithm”, in SPRINGER LINK, vol.947,pp 9-21, February 2023.
5. Michos;Gkoumas;Siakampeti;Fidas, “On the Extension of the Haar Cascade Algorithm for Face Recognition: Case Study and Results”, in ACM, vol.10,pp 53–56, November 2020.
  6. Ruth Ramya Kalangi; Suman Maloji; P. Syam Sundar; Sk Hasane Ahammad, “Deployment of Haar Cascade Algorithm to Detect Real-Time Faces”, in IEEE,February 2022.
  7. Suraj Raj & Saikat Basu, “Attendance Automation Using Computer Vision and BiometricsBased Authentication-A Review”, in Springer Link, vol. 58, no. 7, pp. 757-767, June 2021.

# Sales Forecasting Using Linear Regression

<sup>1</sup>Rose Brahma, <sup>2</sup>Jenarius Ganlary, <sup>3</sup>Oinam David Singh  
*Department of Computer Science, University of Science and Technology  
Meghalaya, India<sup>1,2,3</sup>*

## **Abstract**

Sales forecasting plays a pivotal role in effective business decision-making, enabling organizations to optimize inventory, allocate resources, and enhance customer satisfaction. This study explores the application of the linear regression method, enhanced by machine learning techniques, for accurate and efficient sales forecasting. Linear regression, a foundational statistical approach, provides a clear and interpretable model for understanding relationships between sales and influencing factors such as pricing, seasonality, marketing efforts, and economic indicators. By integrating machine learning techniques, the model's accuracy is improved through advanced data preprocessing, feature selection, and handling of large datasets. The proposed approach is validated using real-world sales data, demonstrating its ability to predict sales trends effectively. Key performance metrics, such as Mean Squared Error (MSE), Mean Absolute Error (MAE) and  $R^2$  (Coefficient of Determination) are used to evaluate the model's performance. Results indicate that combining linear regression with machine learning methods enhances forecasting precision, making it a practical tool for businesses aiming to stay competitive in dynamic markets. This study underscores the potential of machine learning-

assisted linear regression in driving data-driven decision-making in sales forecasting.

**Keywords:** Linear regression, Mean Square Error, Mean Absolute Error,  $R^2$  (Coefficient of Determination).

## 1. Introduction

Sales forecasting is a critical process for businesses, enabling them to anticipate future demand, plan resources, and make informed decisions to maximize profits. With the growing availability of data, machine learning techniques have become powerful tools for accurate and efficient sales predictions. Among these methods, linear regression stands out as a simple yet effective approach for modeling relationships between sales and influencing factors.

Linear regression in machine learning is used to predict a target variable (in this case, sales) based on one or more independent variables, such as pricing, advertising budgets, seasonal trends, or economic indicators. By identifying the linear relationship between these variables, the model helps businesses understand how various factors impact sales and provides insights into future trends.

This analysis focuses on leveraging linear regression to create a predictive model for sales. The simplicity of linear regression, combined with its interpretability, makes it an ideal starting point for forecasting in many scenarios. However, the approach is most effective when the relationships between variables are approximately linear, and its results are best complemented by exploring more complex models in cases of non-linear patterns.

Through this study, we aim to demonstrate how linear regression can be applied for sales forecasting, interpret its results, and highlight its practical implications for business strategy and planning.

### 1.1 Why Sales Forecasting Matters

Sales forecasting plays a vital role in helping businesses anticipate demand and make informed decisions. Accurate forecasts enable companies to optimize inventory levels, schedule production efficiently, and plan marketing strategies.

## 1.2 Objective of the Study

This study aims to apply linear regression to forecast sales and identify key factors such as seasonal trends and pricing that influence sales. The research demonstrates the practical application of this method and suggests improvements for future iterations.

## 2. Related Studies

1. Hyndman, R.J., & Athanasopoulos, G, Makridakis, S. (2021), Wheelwright, S.C., & Hyndman, R.J (1998), Hyndman, R. J., & Athanasopoulos, G.(2018) [1,2,4] Overview of Forecasting Techniques
  - Importance of understanding data and selecting appropriate forecasting methods.
  - Combination of statistical models and machine learning to enhance accuracy.
2. Box, G. E., Jenkins, G. M., & Reinsel, G. C.(2003) [3] Seasonality and ARIMA Models
  - Introduction to ARIMA and SARIMA models for time series forecasting.
  - Emphasis on stationarity and autocorrelation analysis.
3. Shumway, R. H., & Stoffer, D. S. (2017), Francesca Lazzeri (2020), Vivian Siahaan and Rismon Hasiholan Sianipar (2022) [5,7,9] Advances in Machine Learning for Forecasting
  - Combining traditional statistical models like state-space models with modern machine learning methods.
  - Role of Python libraries in building scalable forecasting solutions.
4. Manu Joseph (2002), Michael Gilliland, Len Tashman, and Udo Sglavo (2021) [6,8] Relevance of Linear Regression [6,8]
  - Use of linear regression as a foundation for advanced machine learning techniques.
  - Balancing simplicity and accuracy in forecasting models.

These references demonstrate the integration of traditional and modern forecasting techniques, with a focus on leveraging simple models like linear regression while addressing their limitations.

### 3. Methodology

#### 3.1 Data Sources and Preparation

- Data Source: Walmart sales data (historical and recent).
- Preprocessing Steps:
  - Removed irrelevant features (e.g., store IDs).
  - Converted dates to monthly periods.
  - Checked and handled missing values.

#### 3.2 Feature Engineering

- Created time-based features such as seasonality and monthly trends.
- Included external variables like advertising spend.

#### 3.3 Model Development

1. Linear Regression Model:
  - Trained using historical sales data.
  - Focused on understanding relationships between variables.
2. Performance Metrics:
  - MSE: Measures average squared prediction error.
  - MAE: Assesses average absolute differences between actual and predicted values.
  - R<sup>2</sup>: Indicates how well the model explains sales variations.

#### 3.4 Tools Used

- Python libraries: Pandas, NumPy, Scikit-learn, Matplotlib.
- Platform: Google Colab for collaborative coding.

### 4. Results

#### 4.1 Model Performance

- MSE: The model's Mean Squared Error was low, indicating minimal large prediction errors.
- MAE: The Mean Absolute Error showed a close alignment between actual and predicted values.
- R<sup>2</sup>: The Coefficient of Determination was high, confirming the model's ability to explain the variability in sales.

#### 4.2 Key Findings

- The model accurately captured sales trends, with predicted sales closely matching actual sales in test datasets.
- Seasonal trends, such as holiday periods, were reflected in the model's predictions, highlighting its reliability in capturing temporal patterns.

- Factors like advertising spend and product pricing were identified as significant contributors to sales variability.

### 4.3 Visualization of Results

- Testing the dataset

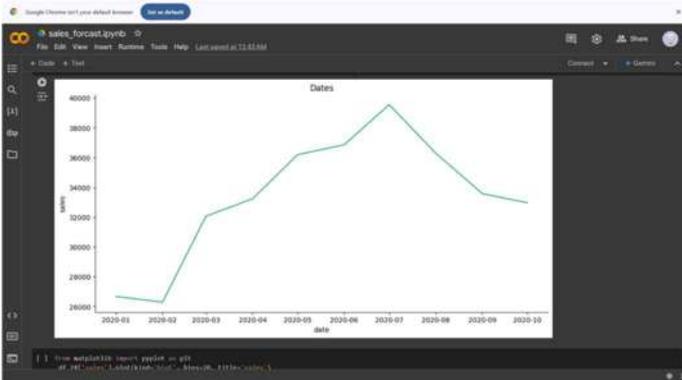


Fig- Convert the resulting date column to timestamp datatype

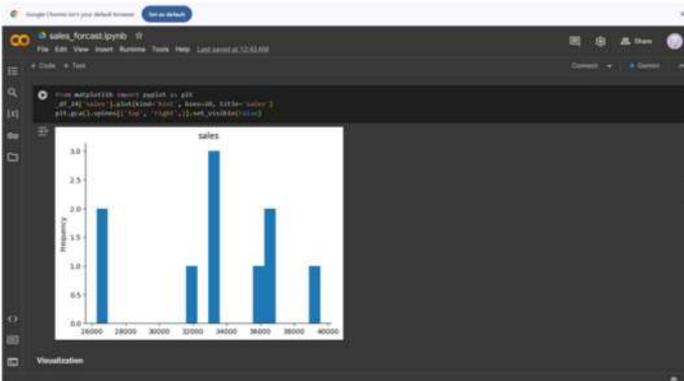


Fig- Convert the resulting date column to timestamp datatype in Bar graph.

- Visualization

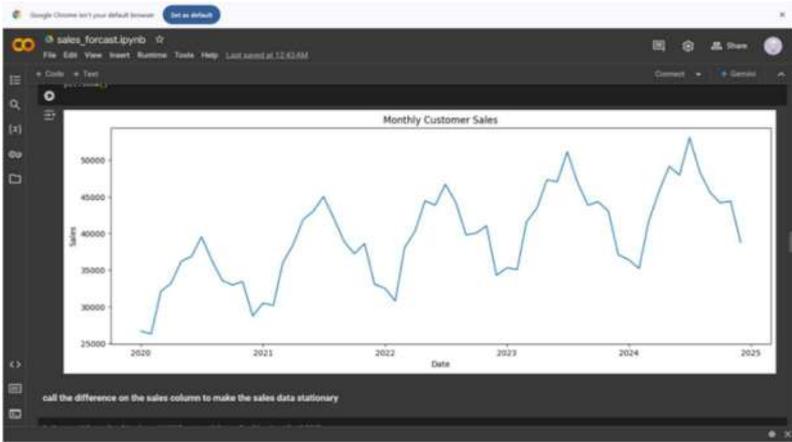


Fig- Visualization of the result

- Call the difference on the sales column to make the sales data stationary

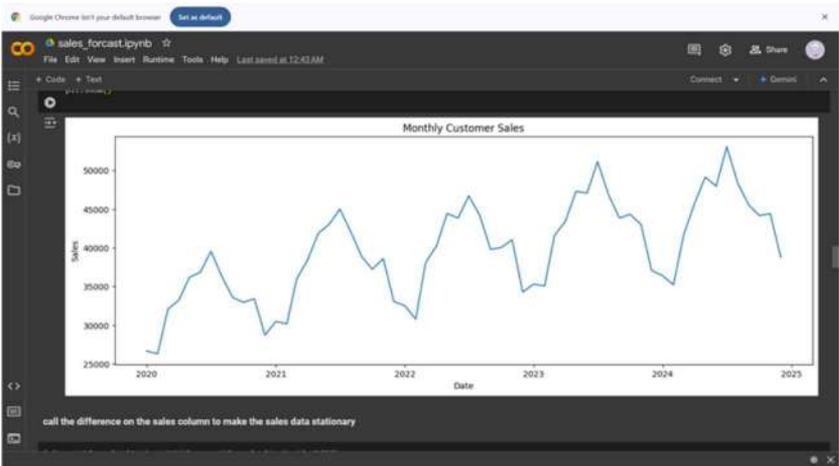


Fig- Call the difference on the sales column to take the sales data stationary

- Final Result

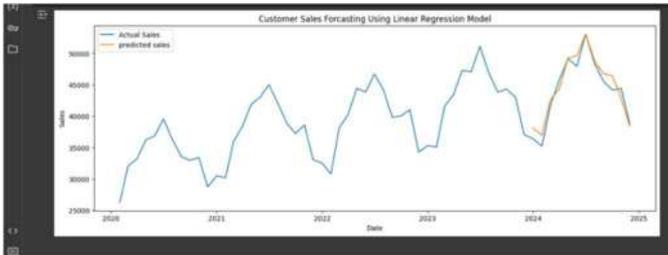


Fig- sales forecasting using linear Regression

A comparison plot of actual vs. predicted sales demonstrated that the linear regression model could reliably follow real-world trends. Periods of peak sales, such as holiday months, were accurately forecasted.

## 5. Feedback and Improvements

- Reviewer Suggestion 1: Address limitations in handling non-linear relationships.
  - Response: Plan to integrate polynomial regression and other advanced techniques.
- Reviewer Suggestion 2: Expand feature set to include competitor pricing and external economic factors.
  - Response: Incorporate additional variables in future iterations.

## 6. Future Steps

1. Explore hybrid models combining linear regression with machine learning techniques like decision trees or neural networks.
2. Integrate real-time data streams for dynamic and adaptive forecasting.
3. Conduct extensive testing with larger datasets from diverse industries.

## 7. Conclusion

In this sales forecasting analysis using linear regression, we explored the relationship between sales and key influencing factors, leveraging machine learning to predict future sales trends. The

model provided valuable insights into how variables such as pricing, advertising, and seasonality impact sales performance. By analyzing these relationships, we identified actionable strategies to optimize business outcomes.

The linear regression model proved to be a straightforward and effective tool for understanding and forecasting sales in scenarios where a linear relationship exists between the dependent and independent variables. It demonstrated reasonable predictive accuracy, as reflected in evaluation metrics such as  $R^2$  and Mean Absolute Error (MAE). However, the analysis also highlighted the limitations of linear regression, particularly in handling non-linear patterns and interactions among variables, which may require more advanced modeling techniques for improved accuracy.

Overall, the results of this analysis underscore the importance of leveraging data-driven approaches for sales forecasting. By integrating predictive insights into decision-making processes, businesses can enhance operational planning, allocate resources efficiently, and ultimately drive revenue growth. Future work could focus on exploring advanced models such as polynomial regression or machine learning algorithms like decision trees or neural networks to capture more complex patterns in sales data.

## 8. References Books:

1. Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*.
2. Makridakis, S., Wheelwright, S.C., & Hyndman, R.J. (1998). *Forecasting Methods and Applications*.
3. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.
4. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts.
5. Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples*. Springer.
6. "Modern Time Series Forecasting with Python: Explore industry-ready time series forecasting using modern machine learning and deep learning" by Manu Joseph (2022)
7. "Machine Learning for Time Series Forecasting with Python" by Francesca Lazzeri (2020)

8. "Business Forecasting: The Emerging Role of Artificial Intelligence and Machine Learning" edited by Michael Gilliland, Len Tashman, and Udo Sglavo (2021)
9. "Time-Series Sales Forecasting and Prediction Using Machine Learning with Python" by Vivian Siahaan and Rismon Hasiholan Sianipar (2022)
10. Websites and Online Tools:
  - Kaggle datasets for sales forecasting.
  - ResearchGate for academic articles.

Case Studies and Reports:

- McKinsey reports on retail forecasting.
- Industry white papers on forecasting trends.

# Power Consumption-Aware Symmetrical Traffic Splitting Based on Heterogeneous Traffic Demands

Amit Khanjan Sarma<sup>1</sup>, Sanjib K. Deka<sup>2</sup> and  
Nityananda Sarma<sup>3</sup>

*Department of CSE, Tezpur University, Assam, India<sup>1,2,3</sup>*

## **Abstract**

In the context of survivable elastic optical networks(SEON), the concept of state-aware symmetrical traffic demand(STD) splitting has emerged as a prominent multi-path protection strategy in the literature. This methodology effectively divides the overall traffic demand into equal parts, utilizing link-disjoint multiple paths alongside state-aware allocation techniques for the transmission of each part. The incorporation of distance-adaptive modulation is proposed to further optimize network performance by decreasing the likelihood of blocking and minimizing power consumption. This study introduces a Power Consumption-aware Symmetrical Traffic (PCST) splitting model, which symmetrically allocates traffic demands across multiple available link-disjoint paths. By employing distance-adaptive modulation, the model seeks to enhance the connection acceptance rate, mitigate blocking probability and reduce power consumption. Specifically designed for heterogeneous traffic demands, the PCST splitting model identifies the optimal number of traffic demand splits. Simulations

run on several real-world networks demonstrate that the proposed model leads to a dramatic decrease in power consumption compared to the conventional state-aware STD model discussed in available literature. These findings indicate that the PCST splitting model provides substantial improvement in the performance of the network, yielding higher power efficiency while maintaining low blocking probability.

**Keywords:** Survivable elastic optical networks • Optical network • routing and spectrum assignment • Orthogonal frequency-division multiplexing (OFDM) • Wavelength Division Multiplexing • Symmetrical traffic splitting.

## 1 Introduction

The research domain of SEON has developed into a very active area because of the exponential growth in internet traffic and the critical demand for energy-efficient network operations. A prime example is Wavelength Division Multiplexing technology, which is widely used by optical transport networks; however, it has a rigid grid architecture that yields poor resource utilization and high energy consumption [33]. This necessity has motivated the design of much more flexible and energy-efficient technologies, such as OFDM-based SEON. Being energy efficient in SEON is not just an operational requirement but a compulsion from the environmental front because the carbon footprint generated by the telecom infrastructure is considerable. The deployment of flexible-grid technology is, in practice, a viable approach to the future of networks, where all stakeholders strive to minimize operational expenditures costs and the carbon footprint from network operations. These technologies become more agile in the allocation of resources and reduction in energy wastage since channel bandwidth can be dynamically adjusted to match real traffic requirements. Survivability is the EON's ability to continue to deliver services even during failure. A detailed energy efficiency analysis of different protection schemes in SEON, like shared backup path protection (SBPP) and dedicated path protection (DPP), has been carried out. Since it allows two or more paths to share backup resources that are only utilized during failures, SBPP is characterized by its exceptional energy efficiency. This considerably reduces energy consumption compared to DPP schemes such as the

1+1 and the 1:1 protection schemes. It is a solution that ensures network reliability without necessarily requiring significant additional resources. DPP does this by assigning dedicated individual backup resources to the primary path, which it actually uses only during failures. Though it is more energy efficient than constant protection methods, it still requires substantial usage of resources. Additionally, RSA enhances energy performance through adaptive tailoring of the resources deployed according to the demands of the network and state conditions at that particular time. When RSA are properly deployed, the network will run with the fewest resources required to meet traffic demand, hence reducing total energy consumption. These protection schemes, together with the RSA methods and their traffic splitting and allocation algorithms, are critical to achieving the balance between network resiliency and energy efficiency [32].

The existing literature has shown that SEON generally achieves better energy efficiency under different traffic loads using various RSA strategies. Future network architectures are more appealing since they can adapt dynamically to errors while incurring low resource overhead. Additionally, SEON incorporates other RSA techniques, such as optimal traffic splitting and state-aware allocation techniques, which further enhance network performance. To this end, SEON can work with dynamic allocation algorithms that adapt to traffic demand size, splitting traffic demands over link-disjoint multiple paths. This approach lowers the blocking probability of lightpaths, reducing energy consumption and optimizing resource utilization. State-aware allocation algorithms, such as Worst-Fit(WF), Best-Fit(BF) and First-Fit(FF), dynamically allocate resources based on heterogeneous traffic demand and prevailing network conditions, thereby enhancing efficiency and minimizing congestion risks. The use of different allocation strategies according to the sizes of traffic demands has a significant payoff for the network. The FF allocation scheme works particularly well for low traffic demands as it fills the available slots with minimal searching, thus minimizing blockage probabilities and optimizing resource use for such smaller connections. On the other hand, WF allocation is effective at handling high traffic demands, even if it results in increased fragmentation, as this reduces the likelihood of blocking larger connections, thereby increasing overall network capacity utilization. As a result, this method benefits in serving high traffic requirements. The BF allocation technique may be a good

solution for medium demand sizes because it effectively maintains lower blockage probabilities by minimizing fragmentation and maximizing resource utilization. Hence, with a dynamic choice of allocation strategies for handling heterogeneous traffic demands, the network gains more flexibility and efficiency in dealing with a wider range of demands. This eventually maximizes network resource utilization, increases connection acceptance rates, and decreases blocking probability.

The two strategies, optimal traffic demand splitting and multipath routing, ensure that resources are highly resilient and efficient with reduced power consumption. By dividing the network traffic among multiple paths, multipath routing maximizes fault tolerance in the case of Shared Risk Link Groups(SRLG) [23]. This approach optimizes overall network performance by allowing data to be transferred through different routes from source to destination; it is particularly important in cases where a single path's ability to carry the entire traffic demand is impaired by either a single link failure or congestion. multipath routing allows traffic to flow through multiple paths, reducing the negative impact of spectrum fragmentation and lessening blocking probability, thereby improving network performance, reliability, and resource use. However, multipath routing does require additional hardware to be activated and maintained, such as transponders, amplifiers, and switches that enable path diversity. Energy consumption is increased because each of these active components contributes to the power consumed, making it more energy-intensive than traditional single-path routing. Additionally, the requirement for ongoing monitoring and management of multiple paths escalates the computational demands on network control systems, further increasing power usage. While multipath routing offers improvements in fault tolerance and load balancing, it also introduces a more intricate and energy-intensive network infrastructure, which may counterbalance some of the energy efficiency benefits derived from optimized resource allocation. multipath routing offers considerable advantages regarding network performance and resilience; however, it also results in increased power consumption within EON. In contrast, the optimal symmetrical traffic demand splitting serves as a strategic optimization method that evenly distributes traffic demand across

multiple available paths. This technique proves particularly beneficial in scenarios where the frequency range is fragmented. By dividing the demand into discrete, symmetric segments and dynamically allocating them to an optimal number of paths, this method avoids the pitfalls of routing all traffic through a single path. Consequently, by enhancing resource efficiency and mitigating issues related to spectrum fragmentation or SLFs, this approach significantly reduces the likelihood of blocking, improves connection acceptance rates, and lowers the power consumption of each path compared to traditional fixed splitting methods [9], [10], [11]. The implementation of optimal traffic demand splitting is essential for minimizing power consumption in EON. By judiciously segmenting traffic demands into an optimal number of sub-demands, the network can operate more efficiently, thereby preventing the activation of unnecessary components. It is vital to avoid such components, as the power consumption of network elements varies according to their specific functions and capacities. For example, WDM transponders, which are responsible for data transmission, exhibit different power requirements based on their data rates: approximately 34W for 10Gbps, 98W for 40Gbps, and 351W for 100 Gbps transponders. The absence of commercially available CO-OFDM transponders necessitates reliance on assumptions for power consumption estimates. A significant factor distinguishing CO-OFDM from coherent WDM transponders is the Digital Signal Processing(DSP) employed at the transmitter. Given that the complexity of DSP is comparable at equivalent bit rates, it is posited that CO-OFDM transponders exhibit similar power requirements. For dual-polarization transponders, the power consumption is approximately 250 W for a data rate of 40Gbps and 351 W for 100Gbps, which corresponds to 125W and 175.5W for single polarization, respectively. Assuming a linear relationship with bit rate, the power consumption of a single-polarization CO-OFDM transponder can be estimated as follows:

$$\text{PCOF DM} = 1.683 \times \text{TR} + 91.333 \quad [32] \quad (1)$$

In flexible-grid networks, it is posited that the power consumption of Optical Cross Connects (OXC) is comparable to that of their fixed-grid counterparts. The power usage is influenced by the node degree  $N$  and the add/drop degree  $\alpha$ .

$$\text{PCOXC} = 85 + 100N + 150\alpha \quad [32] \quad (2)$$

Optical Amplifiers(OAs), especially Erbium-Doped Fiber Amplifiers (EDFAs), generally require a power input of 30 W for each di- rection of operation. Furthermore, each amplifier site incurs an additional overhead of 140 W, which significantly impacts the total power consumption of network components. These are all important devices that would ensure effective and reliable network performance without consuming too much energy [32].For this reason, the optimal symmetrical traffic splitting has to be maintained, whereby each sub-demand is routed to paths that maximize resource utilization while minimizing the need for additional transponder, amplifier, and switching operations. When traffic demands are efficiently shared, the network can exploit available capacity within paths, avoiding resource over-allocations and decreasing the risk of underutilized pathways. This approach brings better overall energy efficiency due to fair traffic load distribution across the network. Optimal demand splitting also reduces the problem of spectrum fragmentation, enabling more continuous spectrum allocation and reducing the need for sophisticated spectrum management techniques, which can increase power consumption. By reducing splits, the network further minimizes the overhead associated with managing multiple sub-demands, such as the significant increase in signaling and control operations that can consequently increase power consumption. Consequently, optimal traffic demand splitting not only enhances the efficiency of resource utilization but also significantly contributes to the reduction of overall power consumption in EON, thus promoting more sustainable and cost-effective network operation.

Presenting distance-adaptive modulation to these strategies is not only more beneficial for network performance but also diminishes power consumption even more. Lower-order modulation formats are used for longer distances to ensure effective transmission, whereas higher-order formats are used for shorter distances to ensure spectral efficiency since the differences in spectral efficiencies and transmission distances for various modulation formats affect EON power consumption. Lower-order modulation formats, such as Binary Phase Shift Key- ing (BPSK), have less spectral efficiency but can transmit over very long distances with limited signal degradation. This innate tolerance for

noise allows transponders to work at lower power for signal regeneration, helping achieve more energy-efficient long-distance transmissions. However, this robustness comes with higher bandwidth needs, which may cause increased total energy consumption per transmitted bit. On the other hand, Quadrature Phase Shift Keying (QPSK) improves spectral efficiency by allowing double data transmission under the same bandwidth that BPSK uses. Although QPSK supports moderate transmission distances—shorter than BPSK yet longer than higher-order modulation formats—it maintains a balance between efficient bandwidth utilization and adequate transmission range. Higher-order modulation formats, such as 16-Quadrature Amplitude Modulation (QAM) and 64-QAM, provide even greater spectral efficiency, facilitating the transmission of larger data volumes within the same bandwidth. However, these formats are more vulnerable to noise and necessitate more frequent signal regeneration over shorter distances, resulting in increased power consumption. Consequently, the selection of modulation format significantly influences the balance between power efficiency, spectral efficiency, and transmission distance in optical networks shown in table 2. Distance-adaptive modulation techniques adjust the modulation format according to the transmission distance between nodes. When integrated with RSA strategies, this approach reduces both blocking probability and power consumption, thereby ensuring reliable and efficient network performance [34]. In order to explore above mention RSA strategies, this paper proposed PCST splitting model that include OTDS, link disjoint multipath routing, state-aware allocation algorithms and distance adaptive modulation format. The objective of these models is to optimise the use of resources by improving the survivability of EON, guaranteeing a strong and power-efficient infrastructure that can satisfy future requirements. The suggested models address present issues such as fixed number of splitting and modulation formats in the literature and lay the groundwork for future more resilient and sustainable network architectures by integrating dynamic and flexible symmetrical splitting and resource allocation techniques.

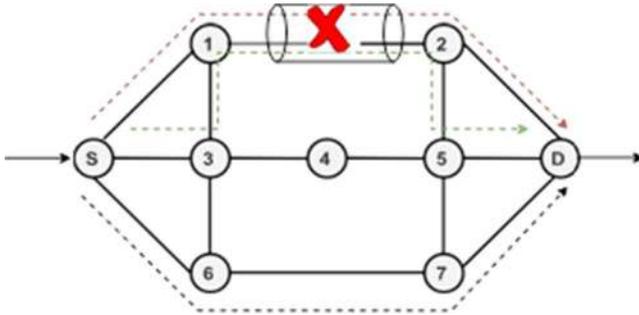


Fig. 1: SRLG.

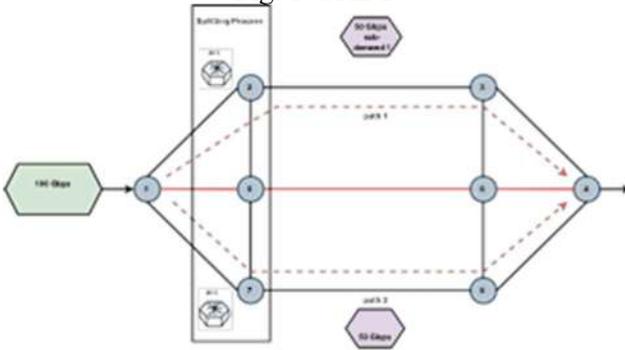


Fig. 2: Optimal Symmetrical traffic demand splitting.

The combination of Link disjoint multipath routing with optimal symmetrical traffic demand splitting in an SEON is a sophisticated technique to optimise resource utilisation and to reduced power consumption. In our research, we coupled Link disjoint multipath routing routing and optimal symmetrical traffic demand splitting, state-aware allocation strategy and distance-adaptive modulation to optimise network performance by better utilising the available capacity and distances of each links included in each path. Link disjoint multipath routing are essential for fault tolerance in multipath routing, as well as for removing SRLG impact [23]. SRLG is the set of links shown in figure 1 link 1-2 that share a resource and affects all of the paths, path1: S-1-2-D and path2: S-3-1-2-D in the set in the event that the common link 1-2 fails. Thus, network connection acceptance rate and survivorship are improved by combining Link disjoint multipath routing and optimal symmetrical traffic demand splitting techniques with state-

aware allocation strategy based on incoming traffic demand sizes, state-aware allocation strategies to enhance the network's capacity to accept connections for traffic demands of varying sizes by taking into account the network's state or path availability.

The main contribution of the work is the integration of optimal number of symmetrical traffic demand splitting and distance adaptive modulation [7], [8], into the existing model state-aware STD, which is a combination of several previously mentioned techniques like Traffic Demand Splitting, Multipath routing and state-aware allocation strategy proposed in work [35]. This integration becomes crucial in this particular scenario. Depending on the optimal number of splitting where traffic is splitted based on available paths and distance between the source and the destination, optimal number of symmetrical traffic demand splitting and distance adaptive modulation is an inventive technique to improve spectrum efficiency, lower the likelihood of blocking and power consumption, the main idea is to reduce the power consumption but optimally splitting the traffic demand symmetrically and by using suitable modulation format based on the transmission distance and available paths. This study's primary goal is to demonstrate the advantages of optimal splitting of traffic demand by integrating distance-adaptive modulation in State-aware STD.

Our particular contributions are as follows.

- The proposed model optimally splitted the traffic demand symmetrically into link-disjoint multiple available paths when a single path is not available to accommodate the entire amount of traffic demand or when single path routing is done to prevent splitting power consumption in terms of power consumption [28]. As a result, a number of splittings are carried out based on path availability, which further increases the network's connection acceptance rate by minimising the power consumption of splitting.
- The proposed model use distance-adaptive modulation with optimal splitting and state-aware allocation strategy to reduced blocking probability in the network.

This paper is structured in the following manner: Section II addresses the related literature, while Section III introduces the

proposed model. The simulations and their corresponding results are detailed in Section IV. Finally, Section V concludes the study.

## 2 Related Works

The last several years have seen the publication of numerous studies examining multipath routing (MR) [15], [16], and [17] in SEON [14]. In paper [2], a novel multipath RSA algorithm with distance-adaptive modulation format assignment is proposed and its effectiveness in dynamic traffic scenarios is analyzed. However, this study does not consider different allocation strategies and their impact on reducing blocking probability and power consumption in the network. The Modulation-aware Multipath Routing and Spectrum Allocation(MM RSA) algorithm, presented in paper [7], is an EON multipath RSA algorithm that leverages distance-adaptive modulation techniques to meet connection demands with a minimum number of paths. Research also shows that using flexible distance-adaptive modulation in conjunction with multipath routing can reduce blocking probability while employing a comparable number of sub-carriers as single path routing algorithms. The authors of publications [9] and [10] introduced a novel concept called Partitioning Dedicated Path Protection(PDPP), which divides the total transmission rate required by a connection into symmetrical portions and transmits each part over multiple disjoint paths, taking into account service level agreements(SLA). Expanding on this idea, the author of another article [11] proposed dividing the transmission rate unevenly among the paths, which resulted in improved network performance. Unlike the previous PDPP approach, this work suggests using mixed-integer linear programming(MILP) to determine the optimal way to divide the transmission rate unevenly among several disconnected paths. This method aims to reduce average bandwidth squeezing and the maximum number of slots available on the network, while still providing dedicated protection.

Although multipath routing has proven its advantages, the combination of resource utilization and survivability constraints is not well investigated in existing studies. Further, previous studies focused solely on splitting the transmission rate symmetrically and asymmetrically across disjoint paths but under an assumption that all network nodes have a minimum degree of three. They have also considered bandwidth squeezing, which only results in negative

effects such as lost traffic, rising costs and increased power consumption complexity [12]. To improve resource efficiency and reduce blocking probability in SEON, a state-aware STD-based model was introduced in a work [35]. In this model, a state-aware allocation scheme is adopted, with which traffic demands are splitted over multiple paths according to the demand sizes and a specific modulation format. This strategy increases connection acceptance rate by ensuring network survivability, optimizing resource distribution and reducing traffic blocking compared with existing techniques. however, the model would perform better if it included distance- adaptive modulation. This would allow for consideration of the power consumption of each modulation format and further maximize resource utilization. The model could also reduce power consumption by using distance-adaptive modulation and finding the optimal number of splits according to path availability. In cases where a single path cannot satisfy all the traffic demands, this analysis demonstrates that balancing such splitting of demands can be done symmetrically among paths based on availability. Instead of bandwidth squeezing and fixed number splitting, we will combine distance-adaptive modulation with a state-aware allocation strategy to improve the network connection acceptance probability and decrease power consumption. This strategy accounts for different modulation formats and their power consumption in watts

(W) based on the distance among nodes, avoiding the costly and complex practice of bandwidth squeezing used in previous models. In conclusion, it is described that splitting the traffic demands from source to destination over multiple disjoint paths has a better impact on absorbing SRLG than the EON survivability level. It can find an optimal number of ways for this split with a state-aware allocation scheme and then distribute them according to their sizes using distance-adaptive modulation to reduce the network's blocking probability while minimizing idle power consumption on the path.

3 Power consumption-aware symmetrical traffic (PCST) splitting : proposed model To illustrate the proposed model, let's use a simplified graph to look at a network scenario. Our network comprises nodes numbered 1, 2, 3, 4, 5, 6, 7, and 8, with the following capacities and distances for the links connecting them: 1–2: 60 Gbps with 233

km; 1–5: 10 Gbps with 500 km; 1–7: 70 Gbps with 1500 km; 2–3: 70 Gbps with 233 km; 3–4: 60 Gbps with 234 km; 5–6: 90 Gbps with 500 km; 6–4: 10 Gbps with 500 km; 7–8: 50 Gbps with 2500 km; and 8–4: 90 Gbps with 1000 km. The figure 2 illustrates how the optimally PCST splitting model symmetrically split a 100 Gbps traffic demand into multiple paths: path1 (1-2-3-4) and path2 (1-7-8-4). This splitting occurs when no single path can accommodate the entire 100Gbps traffic demand. Therefore, the proposed PCST splitting model dynamically splits the traffic based on path availability from source to destination. To transfer the 100 Gbps traffic requirement, the system employs modulation formats that are determined by the total distance and availability of each path. By effectively splitting the traffic and selecting the ideal modulation format, the proposed model achieves a reduction in power consumption related to the processes of traffic splitting and modulation.

The proposed framework integrates a state-aware allocation strategy, symmetrical splitting, and distance-adaptive modulation to improve network efficiency while reducing power consumption in Elastic Optical Networks (EON). Different modulation formats exhibit distinct power consumption levels measured in watts. To tackle this issue, the PCST splitting model seeks to minimize the frequency of traffic demand splits by dynamically adjusting traffic distribution according to path availability. This methodology not only lowers the expenses associated with traffic splitting but also mitigates costs linked to each path utilized for traffic division. By selecting the optimal modulation format based on the total distance of each path, the model enhances resource allocation and decreases overall power usage. Furthermore, the model facilitates symmetrical traffic splitting, which streamlines network resource management and promotes an equitable load distribution. This strategy proves particularly advantageous when network paths share similar characteristics and capacities.

### 3.1 Model and assumptions

In the context of a graph  $G(V, E)$ , where  $V$  signifies the collection of vertices and  $E$  denotes the collection of edges, we examine the traffic demands represented by the set  $D$ . For each traffic demand  $i$  within  $D$ , there exists a set of potential paths  $P$  connecting the source and destination nodes. The network accommodates various modulation formats  $K$ , each characterized by distinct spectral efficiency and power consumption metrics. Key

parameters include the traffic demand  $d_i$  associated with each demand  $i$ , the length  $L_p$  of each path  $p$ , the spectral efficiency  $Eff_k$  corresponding to each modulation format  $k$ , the power consumption  $Power_k$  per slot for each modulation format  $k$ , and the total number of available slots  $S$ . The slots required for path  $p$  are represented by  $T_p$ , while  $\delta_{ip}$  serves as a binary indicator of whether path  $p$  is selected for demand  $i$ . The decision variables include  $x_{ip}$ , which quantifies the amount of traffic demand  $d_i$  allocated to path  $p$ ;  $y_{ip}$ , a binary variable that equals 1 if path  $p$  is utilized for demand  $i$ ;  $z_{ik}$ , a binary variable that equals 1 if modulation format  $k$  is chosen for demand  $i$ ; and  $w_{ipk}$ , which indicates the power consumed by demand  $i$  on path  $p$  utilizing modulation format  $k$ .

The objective function of this PCST-MILP formulation is to reduce the overall power consumption linked to the modulation format, taking into account the distances of each path involved in the symmetrical splitting of traffic demand.

$$\text{Minimize } \sum_{i \in \mathcal{D}} \sum_{p \in \mathcal{P}} \sum_{k \in \mathcal{K}} w_{ipk} \quad (3)$$

Where:  $w_{ipk} = T_p \cdot Power_k \cdot y_{ip} \cdot z_{ik}$

This represents the power consumption on path  $p$  for demand  $i$  using modulation format  $k$ , with  $T_p$  being the number of slots required based on the modulation format selected.

Constrains:

$$x_{ip} = \frac{d_i}{|\mathcal{P}|} \cdot y_{ip} \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P} \quad (4)$$

$$z_{ik} \leq \text{AdaptiveMod}_k(L_p) \cdot y_{ip} \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P}, \forall k \in \mathcal{K} \quad (5)$$

$$T_p = \frac{L_p}{Eff_k} \cdot z_{ik} \quad \forall p \in \mathcal{P}, \forall k \in \mathcal{K} \quad (6)$$

$$\text{AllocationStrategy}_i = \begin{cases} \text{FirstFit} & \text{if } d_i < 50 \\ \text{BestFit} & \text{if } 50 \leq d_i < 100 \\ \text{WorstFit} & \text{if } d_i \geq 100 \end{cases} \quad (7)$$

$$y_{ip} \leq \delta_{ip} \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P} \quad (8)$$

$$w_{ipk} = T_p \cdot Power_k \cdot y_{ip} \cdot z_{ik} \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P}, \forall k \in \mathcal{K} \quad (9)$$

$$\begin{aligned} x_{ip} &\geq 0 \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P} \\ w_{ipk} &\geq 0 \quad \forall i \in \mathcal{D}, \forall p \in \mathcal{P}, \forall k \in \mathcal{K} \end{aligned} \quad (10)$$

The aforementioned formulation addresses the challenge of minimizing power consumption while effectively routing traffic within a network. Equation (4), known as the Symmetrical Traffic Demand Splitting Constraint, guarantees that the traffic demand  $d_i$  is evenly distributed across the chosen paths  $p$ . Equation (5), referred to as the Distance-Adaptive Modulation Selection Constraint, ensures that the selected modulation format  $k$  is suitable for the path's distance  $L_p$ , thereby optimizing spectral efficiency and minimizing power usage. Equation (6), the Slot Allocation Constraint, specifies the number of slots needed on each path based on the chosen modulation format and the distance of the path. Equation (7), the State-Aware Allocation Strategy, dynamically identifies the most effective allocation strategy according to the size of the traffic demand  $d_i$ , promoting efficient resource utilization. Equation (8), the Path Utilization Constraint, stipulates that a path may only be utilized if it is both available and selected, thus avoiding unnecessary resource allocation. Equation (9), which calculates Power Consumption, determines the total power consumption by multiplying the required number of slots by the power consumption per slot for the selected modulation format, ensuring the objective function is accurately minimized. Lastly, Equation (10), the Non-Negativity Constraints, ensures that the traffic routed on each path ( $x_{ip}$ ) and the power consumption values ( $w_{ipk}$ ) remain non-negative, which is essential for the physical feasibility of the solution.

An algorithm 1 has been formulated for our proposed model, PCST splitting, as detailed in 1. In this context, the graph  $G(V, E)$  serves to illustrate the network topology, where  $V$  represents the collection of nodes and  $E$  signifies the connections between these nodes. Each edge  $(u, v)$  has an associated distance  $L_{uv}$ . The traffic demand  $d_i$  is the amount of data that needs to be routed for demand  $i$ , typically measured in Gb/s. Each demand must be routed through the network. Modulation formats  $Eff_k$  and  $Power_k$  represent the spectral efficiency and power consumption of the modulation format  $k$ , respectively. The spectral efficiency  $Eff_k$  indicates how much data can be transmitted per slot, while  $Power_k$  signifies the power usage connected to the utilization of a modulation format  $k$ . The number of iterations  $NumIter_i$  refers to the total number of attempts to establish a routing solution for each demand  $i$ , allowing

for optimization over multiple iterations. The total number of connections  $TotalConn_i$  represents the number of individual connections that need to be established for the traffic demand  $i$ . The variables  $ConnAccepted_i$  and  $ConnBlocked_i$  track the number of successful and blocked connections for demand  $i$ , respectively.  $BlockProb_i$  is the probability that a connection will be blocked, which is computed based on the ratio of blocked connections to the total connections attempted.  $PowerConsumed_i$  represents the total power consumed in routing all accepted connections for demand  $i$ . The algorithm first initializes the modulation formats and loads the network topology by constructing the graph  $G(V, E)$ . It then processes each traffic demand by initializing counters for accepted and blocked connections, as well as the total power consumption. For each connection in the given number of iterations, the algorithm randomly selects a source and target node and checks if there is a valid path between them. If valid paths exist, the algorithm calculates the required slots for each path and determines the blocking probability. Depending on the size of the traffic

Algorithm 1 Power consumption-aware symmetrical traffic (PCST) splitting algorithm

```

1: Input:  $G(V, E)$ ,  $d_i$ ,  $Eff_k$ ,  $Power_k$ ,  $NumIter_i$ ,  $TotalConn_i$ 
2: Output:  $ConnAccepted_i$ ,  $ConnBlocked_i$ ,  $BlockProb_i$ ,  $PowerConsumed_i$ 
3: Initialize  $ModulationFormats(k)$  with  $Eff_k$ ,  $Power_k$ 
4: Load  $G(V, E)$ :
5: for each  $(u, v) \in E$  do
6:   Add  $(u, v)$  with  $L_{u,v}$  to  $G$ 
7: for each  $i \in \mathcal{D}$  do
8:   Set  $ConnAccepted_i = 0$ 
9:   Set  $ConnBlocked_i = 0$ 
10:  Set  $PowerConsumed_i = 0$ 
11:  for iter = 1 to  $NumIter_i$  do
12:    for conn = 1 to  $TotalConn_i$  do
13:      Randomly select Source  $\in V$  and Target  $\in V$ 
14:      while Source = Target or  $\neg PathExists(G, Source, Target)$  do
15:        Randomly reselect Source, Target
16:      if  $PathExists(G, Source, Target)$  then
17:        Find  $\mathcal{P} = DisjointPaths(G, Source, Target)$ 
18:        if  $|\mathcal{P}| > 0$  then
19:          for each  $p \in \mathcal{P}$  do
20:            Compute  $T_p = \frac{L_p}{Eff_k}$ 
21:            Compute  $BlockProb_i = \frac{d_i}{\sum_{p \in \mathcal{P}} T_p}$ 
22:          if  $BlockProb_i > Random()$  then
23:            Increment  $ConnBlocked_i$ 
24:          else
25:            Select Strategy  $_i$  based on  $d_i$ 

```

```

26:         Set RemainingDemandi = di
27:         Set TotalPowerConsumptioni = 0
28:         for each p ∈ P do
29:             Compute wipk = Tp · Powerk · yip · zik
30:             if Tp ≤ RemainingDemandi then
31:                 if Strategyi = FirstFit then
32:                     Allocate Tp on p
33:                 else if Strategyi = BestFit and Tp minimizes slots then
34:                     Allocate Tp on p
35:                 else if Strategyi = WorstFit and Tp maximizes slots then
36:                     Allocate Tp on p
37:             Update RemainingDemandi, TotalPowerConsumptioni
38:         Increment ConnAcceptedi
39:         Update PowerConsumedi = PowerConsumedi + TotalPowerConsumptioni
40:         Compute BlockProbi =  $\frac{\text{ConnBlocked}_i}{\text{ConnAccepted}_i + \text{ConnBlocked}_i}$ 
41: Return: ConnAcceptedi, ConnBlockedi, BlockProbi, PowerConsumedi

```

demand, the algorithm selects an appropriate allocation strategy (e.g., FirstFit, BestFit, WorstFit) and allocates slots on the paths accordingly. The power consumption is calculated for the chosen paths, and the counters for accepted connections and total power consumption are updated. Finally, the algorithm computes the overall blocking probability for each traffic demand and returns the number of accepted and blocked connections, the blocking probability, and the total power consumption.

### 3.2 NP-Hardness of the PCST-MILP

The proposed PCST-MILP formulation for traffic demand allocation across multiple paths, modulation format selection, and power consumption minimization can be classified as NP-hard due to the complexity of the decision-making process involving the variables  $x_{ip}$ ,  $y_{ip}$ ,  $z_{ik}$ , and  $w_{ipk}$ . The problem requires allocating traffic demands  $d_i$  across multiple paths  $P$ , where each path  $p$  has different characteristics such as length  $L_p$  and modulation formats  $k$ , which are defined by spectral efficiency  $\text{Eff}_k$  and power consumption  $\text{Power}_k$ . The binary decision variables  $y_{ip}$  and  $z_{ik}$  determine whether a path is selected for demand  $i$  and which modulation format  $k$  is used, respectively. The continuous variable  $w_{ipk}$ , representing the power consumed for demand  $i$  on path  $p$  with modulation format  $k$ , further increases the complexity of the problem. This formulation inherently involves a combinatorial explosion in the number of possible allocations, as each traffic demand  $d_i$  must be split across multiple paths with varying modulation formats, and each combination of  $y_{ip}$ ,

zik, and wipk introduces additional constraints on the solution space. Given that the decision space grows exponentially with the number of demands  $D$ , paths  $P$ , and modulation formats  $K$ , the problem is NP-hard, as it cannot be solved optimally in polynomial time without reducing its complexity or utilizing heuristic approaches.

To address the NP-hardness of the PCST-MILP problem, a suboptimal solution is obtained using the heuristic approach named Symmetrical First-Fit Decreasing (SFFD) heuristic algorithm 2. This SFFD heuristic provides a practical approach to minimize power consumption while allocating traffic demands across available paths and selecting appropriate modulation formats. The SFFD heuristic works by first sorting the traffic demands  $d_i$  in descending order based on their size. For each traffic demand, it selects the first path  $p$  from the set of available paths  $P$  that has sufficient slots  $S_p$  to accommodate the demand. Next, the modulation format  $k$  with the highest spectral efficiency  $Eff_k$  is chosen for the selected path, ensuring that the required number of slots  $T_p$  is minimized. The power consumption  $wipk$  for demand  $I$  on path  $p$  using modulation format  $k$  is then computed, and the available slots on the path are updated accordingly. This process is repeated for all traffic demands until they are either allocated to paths or blocked due to insufficient resources. While the SFFD heuristic does not guarantee an optimal solution, it significantly reduces the computational complexity of the problem and provides a feasible, near-optimal solution in a reasonable time frame. This makes it an effective strategy for handling the computational challenges posed by the PCST-MILP formulation in practical network scenarios.

Algorithm 2 Symmetrical First-Fit Decreasing (SFFD) heuristic algorithm

---

**Algorithm 2** Symmetrical First-Fit Decreasing (SFFD) heuristic algorithm

---

- 1: **Input:**  $\mathcal{D}, \mathcal{P}, \mathcal{K}, S, Eff_k, L_p$
- 2: **Output:**  $x_{ip}, y_{ip}, z_{ik}, w_{ipk}$
- 3: Sort traffic demands  $\mathcal{D}$  in decreasing order
- 4: **for** each  $d_i \in \mathcal{D}$  **do**
- 5:     **for** each  $p \in \mathcal{P}$  **do**
- 6:         **for** each  $k \in \mathcal{K}$  **do**
- 7:             Calculate  $T_p = \frac{L_p}{Eff_k}$
- 8:             **if**  $d_i \leq T_p$  **then**
- 9:                 Set  $x_{ip} = d_i$

```

10:         Set  $y_{ip} = 1$ 
11:         Set  $z_{ik} = 1$ 
12:         Update  $T_p = T_p - d_i$ 
13:         Compute  $w_{ipk} = T_p \cdot \text{Power}_k$ 
14:         Break and move to the next  $d_{i+1}$ 
15:     if no suitable path  $p$  is found then
16:         Block  $d_i$ 
17: return  $x_{ip}, y_{ip}, z_{ik}, w_{ipk}$ 

```

---

Time complexity of Algorithm 2: The time complexity of the proposed SFFD Heuristic algorithm for the PCST-MILP problem is analyzed as follows. The proposed algorithm first sorts the set of traffic demands  $D$  in decreasing order, which takes  $O(n \log n)$ , where  $n$  is the number of traffic demands. After sorting, the algorithm enters three nested loops. The outer loop iterates over each traffic demand, running  $n$  times. For each demand, the second loop iterates over the set of paths  $P$ , running  $m$  times, where  $m$  is the number of paths. Inside this loop, the third loop iterates over the modulation formats  $K$ , running  $k$  times, where  $k$  is the number of modulation formats. The operations inside the innermost loop, such as slot calculations, path selection, and power consumption computations, are constant-time operations, each taking  $O(1)$ . Therefore, the time complexity of the nested loops is  $O(n \cdot m \cdot k)$ . As  $O(n \log n)$  (from sorting) is typically dominated by  $O(n \cdot m \cdot k)$  in large-scale scenarios, the overall time complexity of the algorithm is  $O(n \cdot m \cdot k)$ , where  $n$ ,  $m$ , and  $k$  represent the number of traffic demands, paths, and modulation formats, respectively. This indicates that the algorithm's runtime scales linearly with the number of demands, paths, and modulation formats, making it efficient for practical use in large networks.

#### 4 Simulation and Results

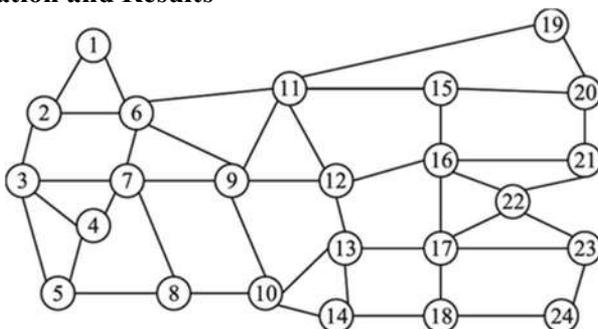


Fig. 3: USNET Topology [27]

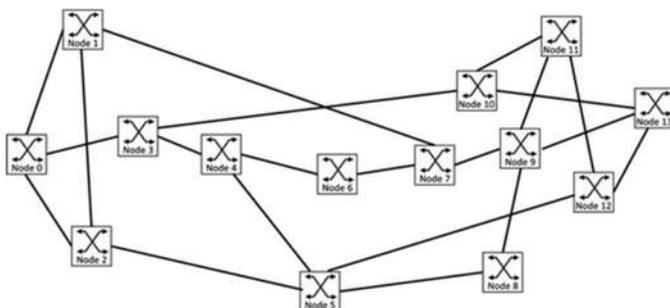


Fig. 4: NSFNET Topology [26].

Two real life topologies have been used in our simulation: USNET, consisting of 24 nodes and 43 links, shown in Figure 3, was selected for its long-haul qualities, and NSFNET, shown in Figure 4, consists of 14 nodes and 21 links. We employed state-aware allocation strategies such as FF, BF and WF based on traffic demand sizes and a threshold value of 50Gbps [29], [30], [31]. The modulation formats were determined based on distance, including BPSK, QPSK, 16-QAM and 64-QAM, corresponding to 1 bit/sec, 2 bits/sec, 4 bits/sec, and 8 bits/sec, respectively. The power consumption values for each modulation format were taken into account in practical implementations, as shown in the table 2. On

the GoogleColab platform, our proposed model was simulated by handling 100,000 requests. Table 1 contains a complete list of all the simulation's parameters and underlying assumptions.

Table 1: Parameters & assumptions used in simulations Parameters & assumptions values/names

Topologies	USNET & NSFNET
Splitting type	Symmetrical
Holding time	1s
No. of requests	1000000
Modulation Formates	BPSK, QPSK, 16-QAM & 64-QAM
Traffic demand categories	small, medium & large
Traffic demand threshold	50Gbps [29], [30], [31]
Traffic demand range	0-50, 50-100, 100+
Allocation strategies	FF, BF & WF
Erlang	200, 220, 240, 260, 280 & 300

Table 2: Summary of modulation formates, path distance subcarries capacity and power consumption [24], [25]

Modulation Format	Subcarrier Capacity (Gbit/s)	Path Distance(Km)	Power Consumption(W)
BPSK	12.5	3000+	112.374
QPSK	25	1500 - 3000	133.416
16-QAM	50	700 - 1500	175.498
64-QAM	75	0-700	215.581

The existing state-aware STD model [35] was simulated, in which traffic demands are split symmetrically into sub-demands and assigned using varying allocation strategies and a fixed modulation format to multiple link-disjoint paths from source to destination based on their traffic demand sizes. Subsequently, taking power consumption into account, these models were compared to the proposed model. The graphs in Figures 5, 6, 7, and 8 illustrate the simulation findings, demonstrating that the proposed model significantly reduces power consumption and blocking probability compared to the existing state-aware STD model. Specifically, the proposed model achieves a power consumption reduction of up to 72.54% in the USNET topology and 88.41% in the NSFNET topology. Furthermore, the blocking

probability is reduced by up to 36.35% in the USNET topology and 40.05% in the NSFNET topology. This achievement indicates that the proposed model successfully meets its objectives of reducing power consumption in the network. It does so by minimizing power consumption related to each path included in the number of splits of traffic demands based on path availability and by integrating distance-adaptive modulation, which considers the distances of each path included in the splits. This comprehensive approach ensures more efficient and effective utilization of network resources by reducing blocking probability and power consumption. It should be noted that we did not account for any differential delay between multiple paths or excessive guard band usage in our simulation.

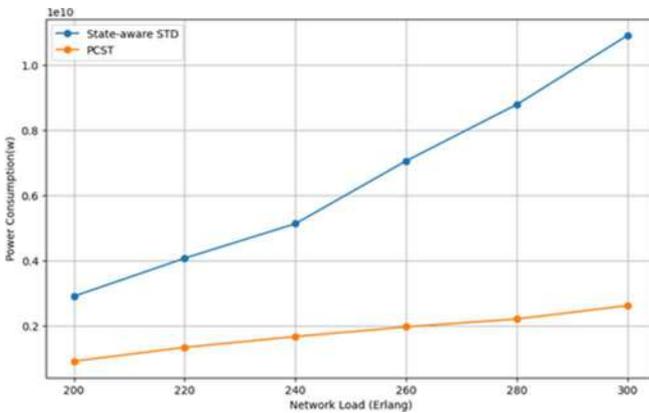


Fig. 5: Comparison of power consumption using USNET Topology.

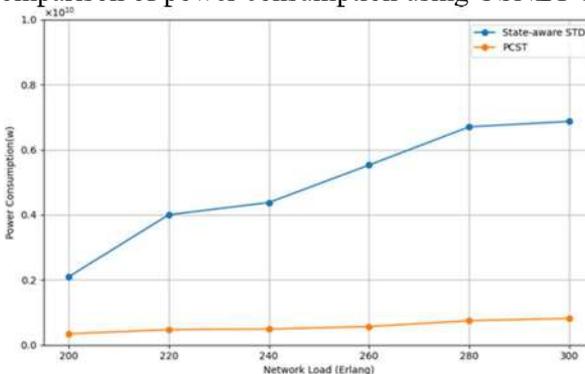


Fig. 6: Comparison of power consumption using NSFNET Topology.

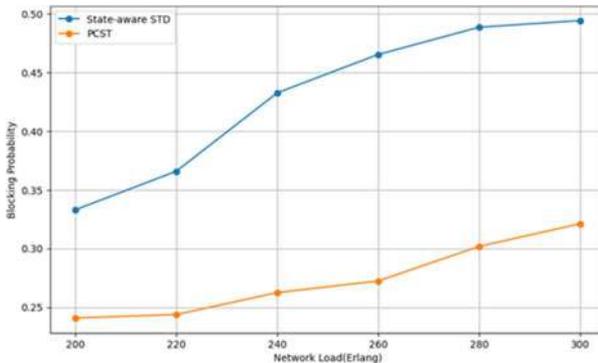


Fig. 7: Comparison of blocking probability using USNET Topology.

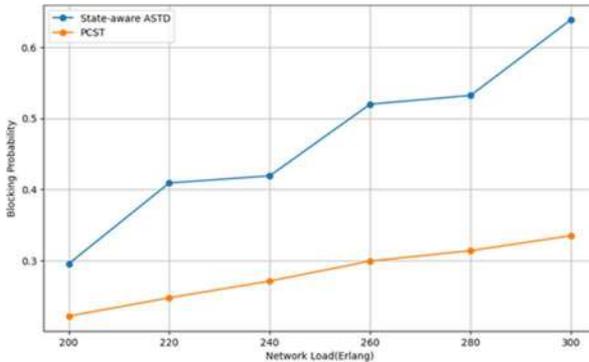


Fig. 8: Comparison of blocking probability using NSFNET Topology.

## 5. Conclusion

In this paper, we extended the existing state-aware STD model to enhance resource allocation and power consumption in EON by incorporating optimal number of splitting with distance adaptive modulation, thereby improving survivability by reducing power consumption in the network. The performance, measured in terms of power consumption and blocking probability, are significantly better compared to existing model state-aware STD. While our current work focuses on effective resource allocation in SEON, future enhancements could include integrating protection schemes to protect connections

against various types of failures, further contributing to the survivability of EON and benefiting service providers.

## **6. Acknowledgment**

This work was financially supported by the Ministry of Electronics and Information Technology (MeitY) (Grant Number: 13(3a)/2020- CC&BT dated:11.03.2021), Govt. of India.

## **References**

1. Author, F.: Article title. Journal 2(5), 99–110 (2016)
2. X. Chen, Y. Zhong and A. Jukan, "multipath routing in elastic optical networks with distance-adaptive modulation formats," 2013 IEEE International Conference on Communications (ICC), Budapest, Hungary, 2013, pp. 3915-3920, doi: 10.1109/ICC.2013.6655169.
3. JOHN M. SENIOR. Optical Fibre Communication. McGraw-Hill, 1980.

# A Voice-Enabled Smart Dustbin System

<sup>1</sup>Atowar ul Islam, <sup>2</sup>Thujota Rhakho, <sup>3</sup> Neha Rani Daimary  
*1,2,3 Department of Computer Science, University of Science and Technology  
Meghalaya, Ri-Bhoi, Meghalaya*

## **Abstract**

In the context of increased urbanization and increasing waste generation, effective waste management systems are necessary for protecting public health and environmental purity. It including overflow, littering, and unhygienic surroundings, which can result in pollution, rodent infestation, and the transmission of illnesses. In order to improve urban cleanliness and operational efficiency, Smart Dustbin are a next-generation waste management system. This work investigates an inventive solution to these problems. The study focuses on integrating cutting edge technologies to convert traditional trash cans into automated, intelligent systems. The smart dustbin makes use of Internet of Things(IOT) concepts and integrates a number of sensors and parts, such as servo motor for automated lid control, HC-05 Bluetooth for voice control,

Arduino Uno microcontroller for data processing, and ultrasonic sensors for object detection and fill level monitoring. With this configuration, the trashcan can monitor its own level of fullness, keep an eye out for any health risks in the area. Along with these technologies characteristics, the voice-controlled interactions that the smart dustbin system offers allow users to operate the bin

hands-free, hence lowering the risk of cross-contamination. With this system, the environmental and health dangers associated with traditional trash disposal methods are minimized, overflow is prevented, and littering is reduced.

**Keywords:** Smart Dustbins, Urban Waste Management, Internet of Things (IOT), Ultrasonic Sensors, HC-05 Bluetooth.

## **Introduction**

As the world is in a stage of up gradation, there is one stinking problem we have to deal with. Garbage! In our daily life, we see the pictures of garbage bins being overfull and all the garbage spills out. This leads to the number of diseases as large number of insects and mosquitoes breed on it. A big challenge in the urban cities is solid waste management not only in India but for most of the countries in the world. Hence, such a system has to be build which can eradicate this problem or at least reduce it to the minimum level. The project gives us one of the most efficient ways to keep our environment clean and green. The smart city concept is still new in India, although it has received a lot of attention in few years when our present prime minister gave the idea of building 100 smart cities throughout India. Now, with the upcoming large number of smart cities, large numbers of responsibilities are also required to be fulfilled. The prime need of a smart lifestyle begins with cleanliness and cleanliness begins with dustbin. A society will get its waste dispatched properly only if the dustbins are placed well and collected well. The main problem in the current waste management system in most of the Indian cities is the unhealthy status of dustbins. In this paper we have tried to upgrade the trivial but vital component of the urban waste management system. This project aims to enhance the efficiency of waste disposal systems, focusing on improving the placement and collection of dustbins. By addressing these fundamental issues, we can work towards a cleaner and greener environment in urban settings.

## **Literature Review**

Twinkle Sinha, K. Mugesh Kumar, P. Saishara (5, May2015) in the paper proposed Smart Dustbins to prevent the accumulation of the garbage along the roadside to control the widespread of many diseases.

Narayan Sharma, Nirman Singha, Tanmoy Dutta (9 September-2015) designed smart bin in a way that can send data about the levels of garbage collected in different parts of the city /town.

Author Fady E. F. Samann (28 June 2018) proposed system based on Arduino Nano board and an ultrasonic sensor to monitor the fullness level of the container and send SMS alerts using a GSM module.

Swati Sharma & Sarabjit Singh (May, 2018) proposed a system using IOT as a hardware and as iconic framework as our software insures the cleaning of dustbins soon when garbage level reaches its maximum. If the dustbin is not clean in specific time, then the record is send to the higher authority so that it can be cleaned. This system also shows the use of PIR sensor, IR sensor and APR module. When some motion is detected by the PIR sensor it opens the gate of waste dustbin using servo motor and when the PIR detect the motion APR module gives the information fed into it of minimum 30 sec.

Priyam Parikh, Dr. Rupesh Vasani, Akshar Rava (2017) had interfaced REID with Arduino UNO. Interfacing Serial LCD with Arduino, Sending AT commands to GSM module. Decision making with the use of ultrasonic sensor. Common DC power supply designing. Implementing SPI communication in Arduino. Multiple Serial Port communication. Circuit Size Reduction. Providing PWM signals to servo motor.

M.R. Sundarkumar, S Naveen, S. Arun Kumar, Kamal Kumar Ray (2017) in the paper “Smart dustbin-garbage monitoring system by efficient Arduino based” have used ultrasonic sensors used to measure the garbage height. Arduino is used to connect all the modules for data transformation.

Md. Aasim Anwar, Prateek Sarkar, Rajeshwar Dutta, Md. Sadik Mohammad Mollick (2018) – have designed structure of the system is developed before implementation of circuit. Used advanced microcontroller called Arduino (ATmega8). It is in-built with many components like analog to digital converter, clock of 16 MHz, shift registers. In this project we put the ultrasonic sensor on top of the garbage bin/ dump. The output of the ultrasonic sensor is processed by the Arduino and the output is then sent to the GSM module which sends a text message to the concerned person. We have a threshold value of 5cm. Which means that if the distance of

the sensor from the top of the garbage is less than 5cm, the output will come with a message that the basket is full? Also, a buzzer will ring if output is less than 5cm. The DHT11 sensor will show the temperature and the humidity.

Dr. Baidyanath Ram Prajapati (2018) – have designed a “Smart Dustbin” which is a GSM enabled bin which automatically detects the garbage level and sends message to respective municipal authorities updating the status of the bin. To make our adopted village Kachhera Warsabad, G.B. Nagar, UP clean and healthy using “Smart Dustbins”.

Anilkumar C.S., Suhas G, Sushma S (2019) in the paper “A Smart Dustbin using Mobile Application” discussed that there is no moving dustbin available in the market. It makes work simple for physically challenged people and aged who are unable to use in the dustbin and also provides a solution for Inappropriate placement of dustbin in the surrounding. A Smart Dustbin using Mobile Application. None of the system contains real time tracking status of the fullness of the bins. No dust bins are present that can detect the odour caused by the bins.

Mamta Pandey, Anamika Gowala, Mrinal Jyoti Goswami, ChinmoySaikia and Dr. Dibyajyoti Bora. (2020) – Smart dustbin using arduino is an IOT based project. Here we are using arduino for code execution, for sensing we used ultrasonic sensor which will open lid and wait for few moments. It will bring drastic changes in terms of cleanliness with the help of technology. Everything is getting with smart technology for the betterment of human being. So this help in maintaining the environment clean with the help of technology. It is a sensor based dustbin so it would be easy to access/use for any age group. Our aim is also to make it cost effective so that many numbers of people can get the benefit from this. And it should be usable to anyone and helpful for them.

From all the above Literature reviews it is clearly understood that there are various ideas suggested for the Smart dustbin concept and various ideas suggested to deal with the garbage collected. But from the user’s point of view, the perspective varies. This system helps users to take responsible decisions by allowing them to act as a responsible citizen for not taking the waste management for granted, since it is a serious challenge that every country all around

the world is facing. Their study has attempted to create innovative system which will help to keep the cities clean. This project will help to kill or minimize the garbage disposal problem, offering numerous benefits in terms of efficiency, hygiene, and sustainability.

### **Problem Statement**

People dispose domestic wastes in the dustbins placed at the roadside. This public dustbin gets filled up randomly. Continuous human monitoring is required to control the overflow level of the dustbin, because sometimes the dustbins fill up faster than usual. When dustbins overflow people can't dispose their wastes in the dustbin, so they dispose it outside the bin. In rainy season, the situation gets worse when rain water enters the dustbin, releasing bad whiff. In order to prevent those all criterion and to avoid people disposing their waste outside the bin, so we are proposing a smart bin system.

**Problem Overview:** With increase of population, the scenario of cleanliness with respect to garbage management is degrading tremendously. People dispose domestic wastes in the dustbins placed at the roadside. In cities there are many public places where we see that garbage bins or dustbins are placed but are overflowing. Continuous human monitoring is required to control the overflow level of the dustbin, because sometimes the dustbins fill up faster than usual. When dustbins overflow people can't dispose their wastes in the dustbin, so they dispose it outside the bin. In rainy season, the situation gets worse when rain water enters the dustbin, releasing bad whiff. This creates unhygienic condition in the nearby surrounding. Also creates ugliness, serious diseases, produces foul smell which degrades the valuation of that area. In order to prevent those all criterion and to avoid people disposing their waste outside the bin, so we are proposing a smart bin system.

**Key Objectives:**

- **Voice Command Integration:** The dustbin should support simple voice commands (e.g., "open the lid," "close the lid," "empty the bin") for hands-free operation.
- **Smart Sensors:** Integrate sensors to monitor the fill level of the bin and provide alerts when the bin is full or requires attention.

- **User Accessibility:** Make the system user-friendly and accessible, especially for people with disabilities, elderly individuals, and those with limited mobility.
- **Sustainability Features:** Include features like notifications for proper segregation of waste or reminders for recycling.
- **Efficient Waste Management:** The dustbin should be connected to a broader waste management system, providing data on waste disposal trends and helping local authorities track waste levels.
- **Hygiene and Sanitation:** The dustbin should minimize manual contact, promoting better hygiene by allowing users to dispose of waste without physically touching the bin.
- **Battery and Power Efficiency:** Design the system to be power-efficient, ensuring it operates without frequent recharging or maintenance.

#### Scope and Deliverables:

- A prototype of the voice-enabled smart dustbin integrated with voice recognition and sensor technology.
- A mobile or web app for monitoring bin status, including notifications, alerts, and analytics.
- Integration of voice assistants (e.g., Alexa, Google Assistant) or custom voice recognition systems.
- Testing and validation to ensure functionality, usability, and durability in real-world conditions.

By addressing these challenges, the Voice-Enabled Smart Dustbin will improve waste management in urban spaces, encourage cleaner environments, and promote greater efficiency and convenience for individuals and waste management authorities alike.

#### **Methodology**

The methodology for developing the smart dustbin system focuses on integrating Internet of Things (IoT) technologies to create an intelligent waste management solution. First, an Arduino Uno micro controller is used for data processing and managing the system's operations. Ultrasonic sensors are incorporated to monitor the fill level of the bin and detect nearby objects, ensuring the dustbin can automatically track its fullness and prevent overflow. A servo motor is employed for automated lid control, allowing the dustbin to open and close based on user interaction or sensor input. The HC-05 Bluetooth

module enables voice control, allowing users to operate the bin hands-free, minimizing cross-contamination. The system is designed with an intuitive interface, incorporating these technologies to enhance user convenience and operational efficiency. The integration of these components results in a smart, self-monitoring dustbin that improves hygiene, reduces waste overflow, and promotes a cleaner urban environment. Our block diagram is shown in Figure 1, circuit diagram is shown in Figure 2 and Bluetooth voice command is shown in Figure 3. The Hardware components involved are:

- Arduino Board (e.g., Arduino Uno)
- HC-05 Bluetooth
- Ultrasonic sensor
- Servo Motor
- Power Supply
- Breadboard
- Jumper Wire

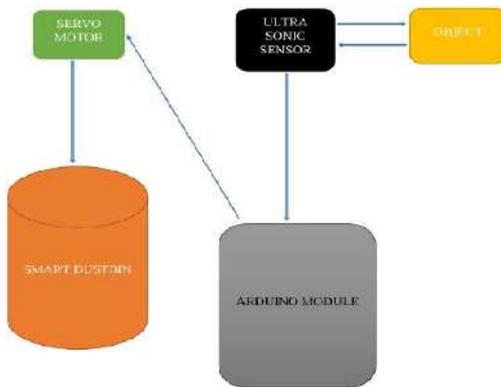


Figure-1- Block Diagram

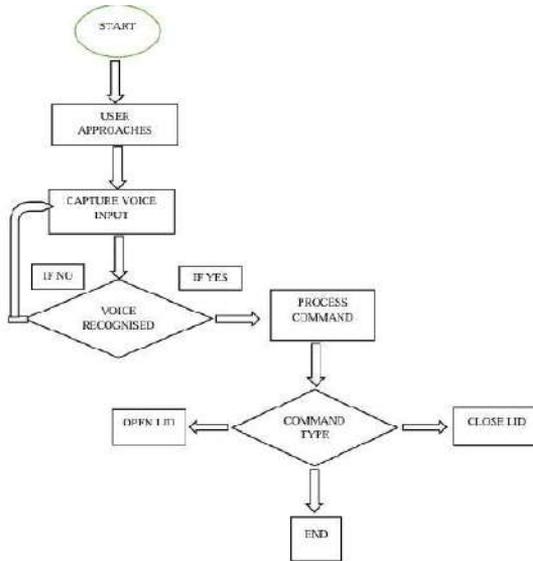


Figure-2 Work Flow Diagram

### Working Principle of Smart Dustbin

The working principle of the voice-enabled smart dustbin system is designed to enhance the efficiency and hygiene of waste disposal through the integration of modern technologies such as IoT, ultrasonic sensors, and voice recognition. The dustbin operates by using voice commands to open and close its lid. When a user issues a command, a voice recognition system processes the sound and triggers a motorized mechanism that opens or closes the dustbin's lid. This system eliminates the need for physical contact, promoting better hygiene and reducing the risk of contamination. The ultrasonic sensor inside the dustbin continuously monitors the waste level. When the bin is nearing full capacity, the sensor sends a signal to alert waste management authorities, ensuring timely collection and preventing overflow. The system is also connected to an IoT platform, which allows real-time monitoring and data collection on waste disposal patterns. This data can be used to optimize waste collection schedules and improve overall management. Additionally, the dustbin's design makes it accessible to people with physical disabilities, the elderly, and children,

enhancing user convenience and making waste disposal more inclusive. Through this automated, intuitive interface, the smart dustbin system significantly improves cleanliness, efficiency, and sustainability in urban environments.



Figure-3: Closed dustbin



Figure-4: Open dustbin

### Discussion and Future Enhancement

- The voice-enabled smart dustbin system offers a promising solution to urban waste management challenges, especially with increasing urbanization and waste generation. Its hands-free operation improves hygiene by reducing cross-contamination, making it particularly useful in high-traffic public areas. The integration of ultrasonic sensors provides real-time data on fill levels, preventing overflow, reducing littering, and improving waste collection timing, which helps mitigate health and environmental risks. However, the voice recognition technology (HC-05 Bluetooth) may face challenges in noisy environments, limiting its effectiveness in public spaces. Additionally, hardware reliability may decrease over time, requiring regular maintenance, and a stable power supply is necessary for consistent performance. The system holds great potential for scalability, allowing for the integration of thousands of smart bins into a central waste management system. This would enable municipalities to optimize waste collection routes, predict disposal trends, and reduce manual labor, all while integrating with other smart city systems for enhanced operational efficiency. Improved Voice Recognition: More accurate AI-based systems.
- Mobile App: Real-time monitoring and waste analytics.
- Smart Sorting: Automatic waste classification (recyclables, compost)

- Solar Power: Eco-friendly energy source.
- IOT Connectivity: Remote monitoring and control.
- Data Analytics: Track waste patterns for optimization.
- User Feedback: Collect user experiences for improvements.
- Enhanced Sensors: Weight sensors for accurate fill levels.
- Smart Alerts: Notification for fullness and maintenance.
- Smart Home Integration: Connect with other smart devices.

### **Conclusion**

The voice-enabled smart dustbin system represents a step forward in modernizing urban waste management. With further advancements in voice recognition, sensor technology, and integration into larger smart city frameworks, the system could be an essential component of sustainable urban waste management. Its scalability and adaptability make it well-suited for both small-scale deployments and large-scale urban applications. The combination of IOT-based automation, real-time monitoring, and voice-controlled operations offers several advantages, including better hygiene, reduced waste overflow, and enhanced user convenience. This system not only addresses the existing challenges of waste management but also aligns with the growing need for smart city infrastructure. Further development and implementation of this system could significantly contribute to cleaner, more sustainable urban environments. This project details the design, implementation, and testing of the smart dustbin system, demonstrating its potential to significantly improve urban waste management. The proposed solution not only addresses the limitations of existing Waste bins but also contributes to the broader goal of developing smart cities through innovative, technology-driven approaches.

### **References:**

- [1] Twinkle Sinha, K. Mugesh Kumar, P. Saisharan “Smart Dustbin” International Journal of Industrial Electronics and Electrical Engineering, vol-6, issue-2, September,2020
- [2] Narayan Sharma, Nirman Singha, Tanmoy Dutta , “Smart Bin Implementation for Smart Cities” International Journal of Scientific & Engineering Research, ---Volumne 6, Issue 9, September, 2015.
- [3] Fady E. F. Samann , “ Arduino Nano board and an ultrasonic sensor to monitor the fullness level of the container and send

- SMS alerts using a GSM module”, International Journal of Science Engineering and Technology, June ,2018.
- [4] Swati Sharma, Sarabjit Singh, “Smart Dustbin Management System”, International Journal Of Engineering Sciences & Research Technology, 7(5), 169-175, 2018.
- [5] U. NAGARAJU (16MCM0004)  
RITU MISHRA (16MCM0003)  
Chaitanya Kumar (16MCM0002)  
Rajkumar (16MCM0001)  
U. NAGARAJU (16MCM0004)  
RITU MISHRA (16MCM0003)  
Chaitanya Kumar (16MCM0002)  
Rajkumar (16MCM0001)  
U. NAGARAJU (16MCM0004)  
RITU MISHRA (16MCM0003)  
Chaitanya Kumar (16MCM0002)  
Rajkumar (16MCM0001)
- [5] U. Nagaraju, Ritu Mishra, Chaitanya Kumar, Rajkumar, “Smart Dustbin for Economic Growth”, Information technology and Engineering, VIT, University, December, 2016.
- [6] Priyam Parikh, Dr. Rupesh Vasani, Akshar Rava, “Smart Dustbin- “An Intelligent Approach to Fulfill Swatchh Bharat Mission”, International Journal of Engineering Research in Electronics and Communication Engineering, April, 2021.
- [7] M.R. Sundarakumar, S. Naveen, S. Arun kumar, Kamal Kumar Ray, “Smart Dustbin-Garbage Monitoring System By Efficient Arduino Based”, Proceedings of the International Conference on Intelligent Computing Systems, Feb 2018.
- [8] Norfadzlia Mohd Yusof, Mohd Faizal Zulkifli, Nor Yusma Amira Mohd Yusof, Azziana Afiffie Azman, ” Smart Waste Bin with Real-Time Monitoring System”, International Journal of Engineering & Technology, DOI:10.14419/ijet.v7i2.29.14006, May,2018.
- [9] Baidyanath Ram Prajapati,” Smart Dustbin”, JSS Academy of Technical Education, January,2020.
- [10] Anilkumar C.S., Suhas G, Sushma S., “A Smart Dustbin using Mobile Application “,International Journal of Innovative

- Technology and Exploring Engineering, 8(11):3964-3967  
DOI:10.35940/ijitee.K2129.0981119, September, 2019.
- [11] Mamta Pandey, Anamika Gowala Mrinal Jyoti Goswami, ChinmoySaikia and Dr. Dibyajyoti Bora, “Smart Dustbin Using Arduino”, The Assam Kaziranga University, Jorhat, Assam, India., . <https://lastminuteengineers.com/neo6m-gps-arduino-tutorial/>.
- [12] Kumar, N. S., Vuayalakshmi, B., Prarthana, R. J., & Shankar, A. “IOT based smart garbage alert system using Arduino UNO” 2016 IEEE region 10 conference (TENCON) (pp. 1028-1034). IEEE. November, 2016.
- [13] Murugaanandam, S., Ganapathy, V., & Balaji, R.,”Efficient IOT based smart bin for clean environment”,2018 International Conference on Communication and Signal Processing (ICCSP) (pp. 07150720). IEEE, April 2018.
- [14] Selvaraj, K., & Chakrapani, A. ,“Smart dustbin monitoring system using LAN Server and Arduino. International Journal of Advances in Computer and Electronics Engineering, 2(4), 20-23, 2017.
- [15] Rajapandian, B., Madhanamohan, K., Tamilselvi, T., & Prithiga, R, “Smart dustbin”, International Journal of Engineering and Advanced Technology (IJEAT), 8(6), 2019
- [16] Chaudhary, V., Kumar, R., Rajput, A., Singh, M., & Singh, T, “Smart dustbin”, International Research Journal of Engineering and Technology, 6(05), 2395-0056.4795, 2019.

# Predicting Heart Disease Risk Using Machine Learning Algorithms

**Jenifa Rahman<sup>1</sup>, Serin Begum<sup>2</sup>, Daisy Kalita<sup>3</sup>**

*Department Of Computer Science, University of Science & Technology Meghalaya<sup>1,2,3</sup>*

**Abstract :** One of the main causes of death globally is heart disease, which is frequently brought on by bad lifestyle choices and delayed diagnosis. Machine learning (ML) algorithms can help in early detection of the disease. In this paper we have used the Cleveland Heart Disease dataset (303 instances) and a larger dataset (1190 instances). Various ML models were trained and evaluated like Logistic Regression, Decision Tree, Random Forest, KNN, DNN and ensemble model. The ensemble model is made of adding Logistic Regression model, Random Forest and DNN. The ensemble model showed a better accuracy and robustness. The models showed comparatively better performance while training and testing with the larger dataset. These results highlight machine learning's ability to forecast cardiac disease accurately and early, which can help with efficient healthcare planning and lower the illness burden.

Keywords: Logistic Regression, Random Forest, Deep Neural Networks, KNN, Hyper parameter Tuning, GridSearchCV, RandomizedSearchCV,

## **1. Introduction**

Heart disease is one of the leading causes of mortality worldwide, accounting for millions of deaths each year. According to the World Health Organization (WHO), cardiovascular diseases (CVDs) are responsible for approximately 17.9 million deaths annually, representing 32% of all global deaths. The increasing prevalence of risk factors such as obesity, sedentary lifestyles, and unhealthy diets has contributed to the rising incidence of heart disease. Early detection and timely intervention are crucial for improving patient outcomes and reducing healthcare costs. Traditional diagnostic methods, while effective, can be time-consuming and may not always provide accurate predictions. This has led to a growing interest in the application of machine learning (ML) techniques in the healthcare domain, particularly for predicting heart disease. Traditional methods for diagnosing heart disease often rely on clinical examinations, physical symptoms, and diagnostic tests such as ECG (electrocardiogram), echocardiograms, or stress tests. While these methods can be effective, they are time-consuming, expensive, and sometimes invasive. One of the most promising solutions is the application of machine learning (ML) algorithms to predict heart disease risk based on patient data. ML techniques have gained significant attention in healthcare due to their ability to process vast amounts of data, identify hidden patterns, and make predictions with high accuracy. This paper explores the potential of various machine learning models, including K-Nearest Neighbors (KNN), Logistic Regression, Decision Trees, Random Forest, Deep Learning, Deep Neural Networks (DNN) and Ensemble model to predict heart disease and improve early detection.

### **1.1 Machine Learning Algorithms**

Machine learning (ML) is a field of study in artificial intelligence concerned with the development and study of statistical

algorithms that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions.<sup>[1]</sup> Advances in the field of deep learning have allowed neural networks to surpass many previous approaches in performance. "It is divided into four categories such as: Supervised Learning, Unsupervised Learning, Reinforcement Learning and Semi-supervised Learning.

**a. Supervised Learning:** It is a type of machine learning algorithm where both input features and its desired output data is given. The supervised learning algorithm find the relation between the input data and its respected output data. Example of supervised learning algorithms: (Regression Algorithms) Linear regression, Support Vector Machine Regression, Decision Tree Regression, Random Forest Regression and (Classification algorithms) Logistic Regression, Support Vector Machines, Decision Trees, Random Forest, Naive Baye etc. .

**b. Unsupervised Learning:** In this type of machine learning technique, there is an absent of output data or the desired output and the learning algorithm are tasked with finding patterns and relationships within the data without any prior knowledge of the data's meaning. This makes unsupervised learning a powerful tool for exploratory data analysis, where the goal is to understand the underlying structure of the data. Example of unsupervised learning algorithms: K-means clustering, Apriori algorithm, Principal Component Analysis(PCA) etc. .

### **c. Reinforcement Learning**

Reinforcement Learning (RL) is a branch of machine learning focused on making decisions to maximize cumulative rewards in a given situation. Unlike supervised learning, which relies on a training dataset with predefined answers, RL involves learning through experience. In RL, an agent learns to achieve a goal in an uncertain, potentially complex environment by performing actions and receiving feedback through rewards or penalties.

### **d. Semi-Supervised Learning**

Semi-supervised learning is a machine learning technique that combines a small set of labeled data with a large amount of

unlabeled data to improve model performance. It is especially useful when obtaining labeled data is expensive or difficult. By learning patterns from unlabeled data, the model achieves better accuracy than supervised learning alone. Common applications include image recognition, medical diagnosis, and natural language processing.

## 1.2 Role Of Various Machine Learning Algorithms In Healthcare

**a. K-Nearest Neighbors (KNN):** A simple, non-parametric algorithm that classifies data based on the proximity to neighboring data points. It is widely used for classification tasks and has been shown to work well in predicting binary outcomes like heart disease.

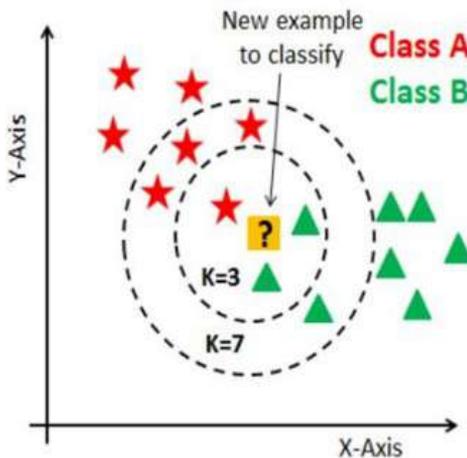


Fig1.2.a KNN model

**b. Logistic Regression:** Logistic regression is a statistical method widely used for binary classification problems, making it particularly suitable for predicting whether a patient is at risk for heart disease. Logistic regression predicts the probability of an

outcome using a mathematical function called the sigmoid function. The sigmoid maps any input value into a probability between 0 and 1. For example, given features like cholesterol, BMI, and smoking habits, the algorithm calculates the likelihood of heart disease (1 = yes, 0 = no).

Logistic regression can classify patients into high-risk or low-risk groups based on traditional health indicators such as systolic blood pressure, cholesterol levels, and smoking status. It is especially valuable for datasets where features are linearly separable.

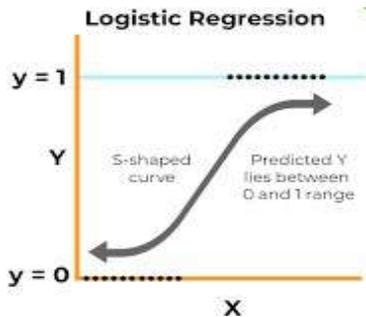
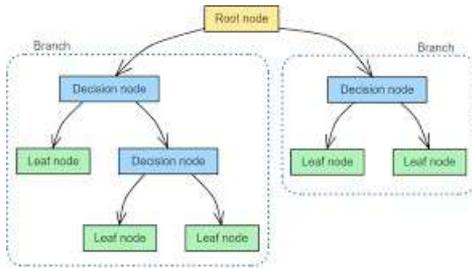


Fig: 1.2.b Logistic Regression model

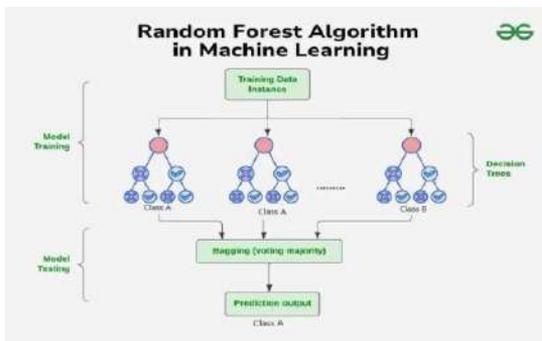
**c. Decision Trees:** Decision trees are versatile and interpretable models that divide data into subsets by sequentially applying simple decision rules based on feature values. The algorithm starts at the root node, evaluates a feature (e.g., cholesterol level), and splits the data into branches based on thresholds (e.g., "Cholesterol > 200 mg/dL"). This process continues until a terminal leaf node is reached, which represents the predicted

class (e.g., "Heart Disease = Yes"). Decision trees can identify key thresholds for risk factors, such as "Cholesterol > 240 mg/dL" or "Age > 60," that help categorize patients into risk groups. For example, a tree might identify a combination of high cholesterol and smoking as a high-risk profile.



**Fig: 1.2.c Decision**

**d. RandomForest:** Random Forest is an ensemble learning algorithm that builds multiple decision trees and combines their outputs to improve prediction accuracy and reduce overfitting. Each tree in the forest is trained on a random subset of the data and uses a random subset of features for splitting. The model aggregates the predictions from all trees (e.g., by majority vote for classification). Random Forest can model interactions between features, such as how age and cholesterol together influence risk. It is particularly useful for datasets with many correlated features.



**Fig: 1.2.d Random Forest**

**e. Deep Neural Networks (DNN):** Deep Neural Networks (DNNs) are powerful machine learning models inspired by the structure and functioning of the human brain. They consist of multiple interconnected layers of neurons, capable of learning complex patterns and relationships in data. In our project, the DNN model demonstrated its strength in capturing intricate features for heart disease prediction, achieving reliable performance with optimization techniques like dropout and regularization to reduce overfitting.

DNNs are particularly useful for integrating diverse datasets, such as combining patient health records, genetic markers, and lifestyle factors, into a unified prediction model.

## 2. Literature Review

In this situation, machine learning algorithms can be very useful in heart disease diagnosis by analyzing vast and diverse volume of patient data. These algorithms can identify the patterns and risk factors in an accurate as well as fast manner. By going through many research papers and articles regarding heart disease prediction using ML techniques, it may be said that for heart disease prediction the multiple ML algorithms are being used and continuously tried to get better performance. It shows the variability of different ML algorithm approaches towards prediction of heart disease. The Cleveland dataset, which contains clinical information from patients, such as age, cholesterol levels, blood pressure, and medical history, has served as a benchmark for evaluating the effectiveness of various machine learning algorithms. By leveraging this dataset, researchers have been able to explore a wide range of algorithms and prediction models to assess their ability to accurately predict heart disease, providing critical insights into how machine learning can enhance diagnostic accuracy and inform clinical decision-making. Archana et al.(2020) conducted a study aimed at comparing several machine learning algorithms for

predicting heart disease, specifically focusing on the K-Nearest Neighbors (KNN) algorithm. Using the UCI Cleveland Heart Disease dataset, the authors compared KNN with other algorithms such as Decision Trees (DT), Linear Regression, and Support Vector Machines (SVM). The study found that KNN performed the best among these algorithms, with an impressive accuracy of 87%. This study highlighted a crucial point: KNN's ability to classify heart disease cases was superior in this context because it is non-parametric and does not assume any specific underlying distribution of the data. Moreover, KNN's simplicity and speed made it particularly attractive for real-time applications, where quick diagnosis and decisions are required. The research also pointed to the need for a deeper understanding of data preprocessing and feature engineering in order to enhance the performance of machine learning models. Manoj et al. (2020) explored the use of image fusion techniques in combination with machine learning algorithms to predict heart disease. Image fusion involves combining different types of medical imaging (e.g., echocardiograms, CT scans) to extract richer, more comprehensive features for model training. In this study, the authors utilized algorithms such as Multiple Linear Regression, Artificial Neural Networks (ANN), J48 (a Decision Tree-based algorithm), and Decision Trees. Their work emphasizes the importance of using multiple data sources—including imaging data alongside patient demographics and clinical measures—to improve predictive accuracy. By incorporating imaging data, the authors argue, machine learning models can better capture the complexities of heart disease, which is often challenging to diagnose based purely on clinical tests. This hybrid approach of combining machine learning and medical imaging represents a promising a venue for enhancing diagnostic accuracy in cardiovascular disease. The work by Dr. M. Kavitha et al.(2021) proposed a hybrid model combining Decision Trees(DT)and Random Forests (RF) for heart disease prediction. Random Forest, being an ensemble method that builds multiple decision trees, often offers better accuracy and generalizability than a single decision tree model. This hybrid approach allowed the authors to achieve an

accuracy of 88% in predicting heart disease risk. This study is significant because it explored the advantages of combining multiple models into a single framework, improving prediction performance and reducing the risk of overfitting. The authors also discussed the trade-offs between model complexity and interpretability, emphasizing that while ensemble methods like Random Forest are highly accurate, they can sometimes be harder to interpret compared to simpler models like decision trees. As heart disease predictions increasingly move toward real-time systems in clinical settings, hybrid models that balance accuracy and interpretability will be crucial. In their 2021 study, Harshit et al. applied Logistic Regression, K-Nearest Neighbors (KNN), and Random Forests (RF) to the UCI Cleveland Heart Disease dataset. Their findings demonstrated that KNN outperformed the other models, with an accuracy of 88.52%. This result echoed findings from earlier studies that KNN is particularly well-suited to heart disease prediction, especially when the dataset has mixed data types and requires a simple, fast model for real-time prediction. However, the study also showed that Random Forest and Logistic Regression were not far behind in terms of accuracy. Logistic Regression, with its well-established use in medical diagnostics, especially in predicting binary outcomes (such as the presence or absence of heart disease), remained a reliable choice for many healthcare applications. Moreover, the research suggested that ensemble methods like Random Forest could offer an edge in cases where multiple factors influence the disease outcome. The research by Mana Saleh AL et al. (2023) introduced a Hybrid Deep Neural Network (HDNN) for heart disease prediction, integrating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The study demonstrated that this hybrid model could outperform traditional machine learning algorithms, achieving an impressive accuracy of 98.86% on a combined dataset from the Cleveland Heart Disease dataset and additional clinical data. This study marks a turning point in heart disease prediction research, highlighting the growing importance of deep learning models, particularly CNNs and LSTMs, in handling complex

medical data. CNNs are particularly effective for extracting features from images, while LSTMs are ideal for processing sequential data, such as time-series information from heart rate or ECG signals. The hybrid model's high accuracy demonstrates the potential of deep learning to uncover hidden patterns in data, which traditional models may miss. Furthermore, it underscores the growing trend of combining multiple types of machine learning techniques to improve performance. In a recent study by B. Kaur & G.Kaur(2023), the author proposed a modified machine learning algorithm designed to enhance the prediction accuracy for heart disease. By integrating multiple machine learning techniques, the model was able to achieve 95% accuracy, which is among the highest reported in the field. Their research emphasized the need for multi-faceted models that combine the strengths of various algorithms to tackle complex healthcare problems. One of the key contributions of their work was the design of a flexible model that could adapt to different datasets and offer high performance in diverse clinical settings. This model is particularly useful in addressing the variety of patient demographics and health conditions that exist in real-world scenarios.

While numerous studies have demonstrated the effectiveness of machine learning algorithms such as logistic regression, decision trees, and random forests in predicting heart disease, there remains a lack of consensus on which algorithm performs best under varying clinical conditions. Additionally, limited research has been conducted on the comparative performance of traditional machine learning techniques and newer deep learning approaches like CNNs, DNNs, and LSTMs in this domain. Therefore, this study aims to conducting a comprehensive comparative analysis of these algorithms, providing valuable insights into their applicability and performance for heart disease prediction.

### **3. Methodology**

**i. Dataset collection:** The dataset used for this project is the Heart Disease Dataset obtained from the Kaggle. Kaggle is a Web platform that hosts a community of data scientists and machine

learning professionals. The dataset contains 303 samples with 14 features such as age, cholesterol levels, resting blood pressure, and others, along with a target variable indicating the presence or absence of heart disease. This dataset is used by 62% researchers for their paper and analysis.

Another dataset is used for analysis which is the larger one having 1190 instances and 12 attributes. The dataset is combination of some individual heart disease dataset, making it the largest dataset for heart disease analysis.

<b>Attribute Names</b>	<b>Description</b>	<b>Type</b>	<b>Possible Values/ Range</b>	<b>Relevance</b>
<b>Age</b>	Age of the patient	Numeric	35-77	Age is a major Risk factor for heart disease
<b>Sex</b>	Gender of the patient	Categorical	0=Female,1=Male	Men are generally at higher risk for heart disease.
<b>Cp</b>	Chest PainType	Categorical	0=Typical Angina,1=AtypicalAngina,2=Non-Anginalpain,3=Asymptomatic	Indicates the nature of chest pain, correlated with heart disease.
<b>Trestbps</b>	Resting Blood Pressure (in mm Hg)	Numeric	94-200	High resting blood pressure is a risk factor for cardio vascular issues
<b>Chol</b>	SerumCholesterol Level	Numeric	126-417	High cholesterol

	(mg/dL)			levels indicate Increased risk.
<b>FBS</b>	Fasting Blood Sugar (> 120 mg/dL)	Binary (0/1)	0=False, 1=True	Elevated blood sugar levels can indicate diabetes, a risk factor for heart disease.
<b>Restecg</b>	Resting Electrocardiogram Results	Categorical	0=Normal, 1=ST-T Abnormality, 2=Left Ventricular Hypertrophy	Indicates electrical activity of the heart
<b>Thalach</b>	Maximum Heart Rate Achieved	Numeric	71-202	Evaluates heart function during physical activity.
<b>Exang</b>	Exercise-Induced Angina	Binary (0/1)	0=No, 1=Yes	Indicates if physical stress triggers chest pain.
<b>Oldpeakslope</b>	ST Depression Induced by Exercise Relative to Rest	Numeric	0-5.6	Measures abnormality in heart activity during exercise
<b>CA</b>	Number of Major Vessels Colored by Fluoroscopy	Numeric	0-3	Determines the extent of blood flow blockage.
<b>Thalassemia</b>	Thalassemia Condition	Categorical	0=Normal, 1=Fixed	Blood disorder

			Defect,2=Reversible Defect	Related to oxygen-carrying capacity
<b>Target</b>	Presence of Heart Disease	Binary (0/1)	0=NoDisease,1 = Disease	Dependent variable indicating whether heart disease is Present

**Table3.1:SummaryofDatasetAttributes**

**ii. Tools used:** The Jupyter Notebook is an essential part of this project. The language preferred here is Python programming language. Various libraries of pandas like Pandas, Numpy, Scikit-learn and TensorFlow is used for better data preprocessing, model training and evaluation. Interactive libraries like Matplotlib and Seaborn are used for visualization purpose.

**iii. Data Preprocessing:** It is a crucial part of a project. The heart disease datasets are checked for missing/null values and it is showing of not having any missing values. The both heart disease were balanced based on the target variable. Numerical features such as age, cholesterol (chol), and resting blood pressure (trestbps) were standardized using Standard Scaler to normalize their range and improve model performance.

**iv. Feature Selection:** For this project, all the available features in the dataset were considered relevant for predicting heart disease, and therefore, no features were excluded during the initial selection process. The dataset contains a combination of both numerical (e.g., age, cholesterol, resting blood pressure) and categorical (e.g., sex, chest pain type, thalassemia) features, each contributing valuable information for heart disease prediction.

**v. Model selection, tuning and evaluation:** For our heart disease dataset different ML models like logistic Regression, Decision Tree, Random Forest, KNN, DNN etc. were selected. The

entire dataset has splits into three categories-Training dataset, Validation dataset and Testing dataset .The dataset is splits in a form of 60%-20%-20% for training, validation and testing purpose.

For the tuning, different hyperparameter techniques like Randomized Search CV, GridSearchCV are used. Data explores several supervised learning algorithms Logistic Regression, Decision Tree, and Random Forest along with hyperparameter tuning techniques such as GridSearchCV and RandomizedSearchCV to study the accuracy of heart disease prediction.

### **i. Logistic Regression Using GridSearchCV**

Logistic Regression is one of the most widely used algorithms for binary classification problems, particularly in the medical domain. In heart disease prediction, the target variable typically indicates the presence or absence of heart disease, making Logistic Regression a suitable baseline model.

Logistic Regression works by estimating the probability that a given input belongs to a particular class using a logistic (sigmoid) function. It models the relationship between independent variables (such as age, cholesterol level, blood pressure, and heart rate) and the dependent binary outcome. One of its major strengths is interpretability, as the model coefficients indicate how each feature influences the likelihood of heart disease.

To enhance model performance, **GridSearchCV** was applied. GridSearchCV is a hyperparameter optimization technique that systematically searches through a predefined grid of parameters (such as regularization strength C, penalty type, and solver) and evaluates each combination using cross-validation. This ensures that the chosen parameters generalize well to unseen data.

After applying GridSearchCV, the Logistic Regression model achieved a test set **accuracy of approximately 83.6%**. This relatively high accuracy demonstrates that Logistic Regression, when properly tuned, can effectively capture linear relationships in heart disease data. However, its limitation lies in handling complex non-linear patterns, which are often present in medical datasets.

## ii. Decision Tree Classifier

The Decision Tree classifier is a non-linear, rule-based learning algorithm that splits the dataset into subsets based on feature values. Each internal node represents a decision rule, each branch corresponds to an outcome of the rule, and each leaf node represents a class label.

In the context of heart disease prediction, Decision Trees are particularly attractive because of their interpretability. Medical professionals can easily understand and visualize how decisions are made for example, how high cholesterol combined with age might lead to a higher risk classification.

The Decision Tree model achieved an accuracy of **79%** on the test set. While this performance is slightly lower than that of Logistic Regression, Decision Trees have the advantage of capturing non-linear relationships and interactions between features without requiring data scaling.

However, Decision Trees are prone to **overfitting**, especially when the tree grows deep and learns noise from the training data. This limitation can reduce the model's ability to generalize well to new patients, which is critical in medical applications.

## iii. Random Forest Classifier

To overcome the overfitting problem of Decision Trees, the **Random Forest classifier** was introduced. Random Forest is an ensemble learning technique that builds multiple Decision Trees using different subsets of data and features, and then combines their predictions through majority voting.

In heart disease prediction, Random Forests are highly effective because they reduce variance while maintaining the ability to model complex, non-linear patterns. Each tree captures different aspects of the data, leading to a more robust overall prediction.

The initial Random Forest model achieved an **accuracy of 79%**, similar to that of the Decision Tree. Although the improvement was not immediate, Random Forest provides better stability and generalization compared to a single Decision Tree. The true strength of Random Forests becomes evident when hyperparameter tuning is applied.

#### **iv. RandomizedSearchCV for Random Forest (best\_random\_grid)**

Hyperparameter tuning plays a crucial role in improving model performance. **RandomizedSearchCV** is an efficient alternative to **GridSearchCV** that samples a fixed number of parameter combinations from a specified distribution. Instead of testing all possible combinations, it explores the parameter space randomly, making it computationally faster for large models like Random Forests.

In this methodology, **RandomizedSearchCV** was used to identify the best parameters for the Random Forest classifier, such as the number of trees (`n_estimators`), maximum tree depth, minimum samples per split, and feature selection strategy.

The tuned Random Forest model, referred to as **best\_random\_grid**, achieved an **accuracy of approximately 80.3%**. This improvement over the base Random Forest model indicates that selecting appropriate hyperparameters helps the model better capture the underlying patterns in heart disease data. **RandomizedSearchCV** is particularly useful when computational resources are limited or when the parameter space is very large.

#### **v. GridSearchCV on Random Forest (best\_grid)**

To further enhance performance, **GridSearchCV** was again applied to the Random Forest classifier. Unlike **Randomized Search CV**, **Grid Search CV** exhaustively evaluates all combinations of specified hyperparameters. Although this approach is computationally expensive, it often yields more precise results when the search space is well-defined.

Using **GridSearchCV**, the optimized Random Forest model referred to as **best\_grid** achieved a **test set accuracy of approximately 81.9% (around 82%)**. This represents a significant improvement over the untuned Random Forest and Decision Tree models, and it approaches the performance of Logistic Regression while maintaining the ability to model non-linear relationships.

The improved accuracy demonstrates that carefully tuned ensemble models can provide reliable predictions for heart disease. Additionally, Random Forests offer feature importance scores,

which help identify the most influential medical factors contributing to heart disease risk.

### 4. Result and Discussion

Some EDA of Cleveland heart disease dataset are given below-

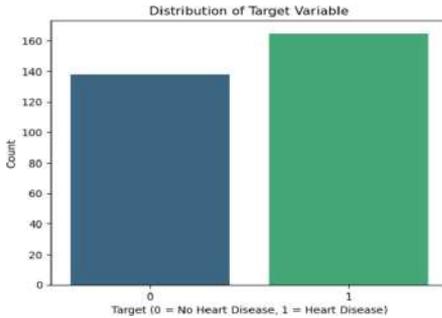


Fig 4.1-Count plot of distribution about target variable

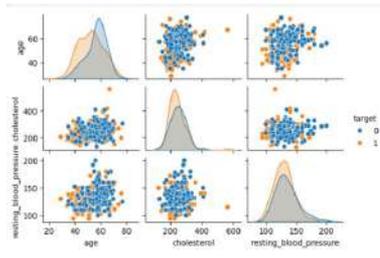


Fig4.2 Pair plot of Resting Blood Pressure, Cholesterol and Age to Explore Feature Relationships

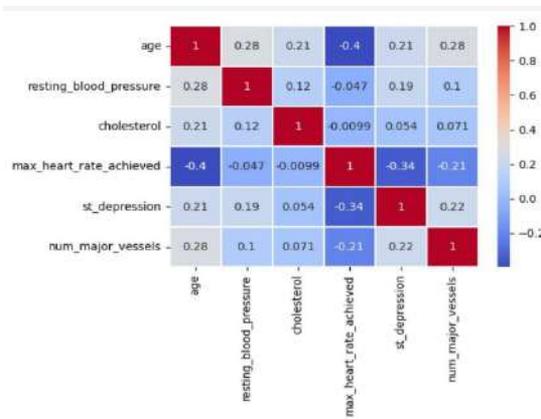


Fig 4.3 Heat map of Feature Correlations for Heart Disease Analysis

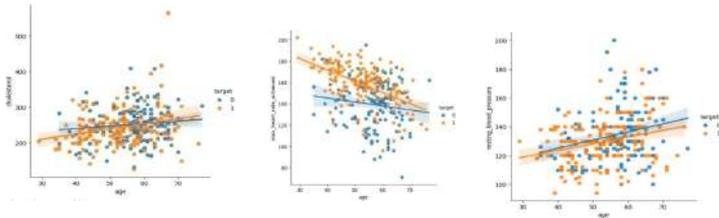


Fig 4.4 Linear Regression plot between some features

EDA of the larger heart disease dataset-

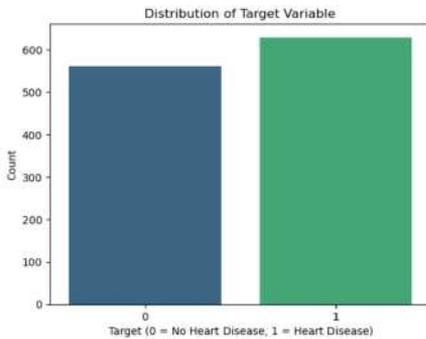


Fig4.5-Countplotofdistributionabouttargetvariable

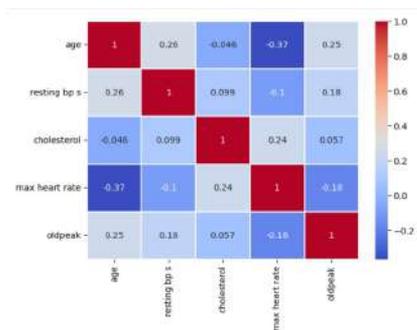


Fig4.6 Heat map of Feature Correlations for Heart Disease Analysis

<b>Machine Learning Models</b>	<b>Model Name</b>	<b>Accuracy</b>
Logistic Regression	best_logmodel	83%
DecisionTree	best_dtree	80%
Random Forest	best_grid	82%
KNN	Knn	80.33%
DNN	Model	78%
Ensemblemodel	Ensemble_model	82%
BoostedRandom Forest	Boosted_rf_model	85%

Table 4.1 Accuracy of different ML models  
(Cleveland heart disease dataset)

The models are also analysed using the larger dataset. The results of evaluation getting from the larger dataset are quite better than above results.

<b>Machine LearningModels</b>	<b>ModelName</b>	<b>Accuracy</b>
<b>Logistic Regression</b>	best_logmodel	82%
<b>Random Forest</b>	best_grid_data2	89%
<b>DNN</b>	Model	82%
<b>Ensemble model</b>	ensemble_model	85%

Table 4.2 Accuracy of different ML models  
(The larger heart disease dataset)

We have seen that the Boosted Random Forest model showed highest accuracy of 85% in case of Cleveland heart disease dataset, whereas it is shown that larger dataset can define comparatively

better accuracy, hence showing better performance. The highest accuracy of 89% shown by the Random Forest model.

## 5. Conclusion

This research investigated the effectiveness of various machine learning algorithms for predicting heart disease, aiming to develop an accurate and reliable predictive framework to support early diagnosis and clinical decision-making. Logistic Regression, Decision Tree Classifier, and Random Forest Classifier were evaluated as baseline models, followed by advanced hyperparameter optimization techniques using GridSearchCV and RandomizedSearchCV to enhance overall performance.

Logistic Regression optimized with GridSearchCV demonstrated strong baseline performance, confirming that linear models can provide reliable predictions when appropriately tuned. Its simplicity, computational efficiency, and interpretability make it a valuable approach in medical applications, where understanding the influence of individual risk factors is critical. However, its ability to model complex non-linear relationships is limited compared to tree-based methods.

The Decision Tree Classifier provided an interpretable structure that captured non-linear patterns in the data and offered insights into feature importance. Despite its advantages in transparency, the model showed comparatively lower accuracy, likely due to overfitting and sensitivity to data variations. This limitation highlights the challenges of relying on single-tree models for complex medical datasets.

The Random Forest Classifier improved predictive performance by combining multiple decision trees, thereby reducing variance and increasing robustness. This ensemble method demonstrated superior generalization capability and better handled interactions among clinical features. Further optimization using RandomizedSearchCV resulted in the `best_random_grid` model, which achieved improved accuracy while efficiently exploring the hyperparameter space with reduced computational cost.

Finally, an exhaustive hyperparameter search using GridSearchCV was applied to the Random Forest model, producing the `best_grid` model with the highest test accuracy. This outcome emphasizes the importance of thorough hyperparameter tuning in

maximizing predictive performance, particularly for ensemble models in healthcare applications.

Overall, the results indicate that Random Forest with GridSearchCV is the most effective approach for heart disease prediction among the evaluated models. Nevertheless, model selection should balance accuracy, interpretability, and computational efficiency. Future research may focus on incorporating larger datasets, advanced feature engineering, and explainable AI techniques to further improve the reliability and clinical adoption of machine learning-based heart disease prediction systems.

## **6. Future Scope**

For the upcoming time, we can hope that the dataset for analysing heart disease should be more precise and having more attributes. The user interface/software for checking the heart disease could be more flexible, quick and correct. The accuracy could be increase and hoping for predicting the exact result.

## **References/Bibliography**

[1] Archana Singh and Rakesh Kumar, "Heart Disease Prediction Using Machine Learning Algorithms", IEEE, 2020.

[2] Manoj Diwakar, Amrendra Tripathi, Kapil Joshi, Minakshi Memoria, Prabhishek Singh, Neeraj Kumar; "Latest trends on heart disease prediction using machine learning and image fusion", ELSEVIER, 2020.

[3] Norma Latif Fitriyani, Muhammad Syafrudin, Ganjar Alfian, Jongtae Rhee, "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System" IEEE, 2020.

[4] Dr.M. Kavitha, G. Gnaneswar, R. Dinesh, Y. Rohith Sai, R. Sai Suraj, "Heart Disease Prediction using Hybrid machine Learning Model", IEEE, 2021.

[5] Harshit Jindal; Heart disease prediction using machine learning algorithms; IEEE,2021. Mana Saleh, L Resha, Samina Am, Muhammad Ali, Adel Sulaiman, Hani Alshahran, Asadullah Shaikh,

and Samina Amin, “A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks”, IEEE, 2023.

[6] Bavneet Kaur and Gaganpreet Kaur, “Heart Disease Prediction Using Modified Machine Learning Algorithm”, Springer, 2023.

[7] Chunjie Zhou, Pengfei Dai·Aihua Hou, Zhenxing Zhang, Li Liu, Ali Li, Fusheng Wang, “A comprehensive review of deep learning-based models for heart disease prediction”, Springer, 2024.

[8] Abhijeet Jagtap, Priya Malewadkar, Omkar Baswat, Harshali Rambade, “Heart Disease Prediction using Machine Learning”, International Journal of Research in Engineering, Science and Management, 2019.

[9]C. Boukhatem, H. Y. Youssef and A.B. Nassif, “Heart Disease Prediction Using Machine Learning”, Advances in Science and Engineering Technology International Conferences (ASET), 2022.

[10] Mohammed Khalid Hossen, “Heart Disease Prediction Using Machine Learning Techniques”, American Journal of Computer Science and Technology, 2022.

[11] M. Mamun, A. Farjana, M. A. Mamun, M. S. Ahammed and M. M. Rahman, “Heart failure survival prediction using machine learning algorithm: am I safe from heart failure?” IEEE ,2022.

[12] P. Motarwar, A. Duraphe, G. Suganya and M. Premalatha, “Cognitive Approach for Heart Disease Prediction using Machine Learning”, International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE), 2020.

# Proactive Edge Computing Task Allocation in IoT Networks using Deep Q-Network (DQN) and Proximal Policy Optimization (PPO)

Priyanka Sarma<sup>1</sup>, Atowar Ul. Islam<sup>2</sup>

*University of Science and Technology Meghalaya, India*

*\*Email: [psarma157@gmail.com](mailto:psarma157@gmail.com)<sup>1,2</sup>*

*ORCID: 0009-0002-2768-3764*

## **Abstract**

The rapid rise of Internet of Things (IoT) deployments has intensified the computational burden on edge infrastructures, challenging their ability to sustain low latency and high throughput. Traditional computation offloading methods often react to task arrivals without proactive planning, resulting in resource bottlenecks and degraded performance. This paper introduces a hybrid Deep Q-Network (DQN) and Proximal Policy Optimization (PPO) reinforcement learning framework that predicts incoming tasks and allocates them optimally to edge nodes. DQN forecasts task resource requirements, while PPO selects optimal nodes for execution by maximizing system utility under dynamic workloads. Simulation results show that the hybrid model achieves significantly reduced latency—with an average of 2.6652 seconds, outperforming standalone DQN (3.0777 s) and PPO (3.0935 s)—and delivers improved throughput. These findings demonstrate the efficacy of proactive learning-driven offloading for scalable, latency-sensitive IoT edge environments.

**Keywords** : Internet of Things, Task Prediction, Task Offloading, Edge Computing

## 1. Introduction

The global Internet of Things (IoT) ecosystem is expanding at an unprecedented rate, projected to surpass 29 billion connected devices by 2025. These devices generate continuous streams of data that require real-time processing for applications such as smart healthcare, intelligent transportation systems, smart manufacturing, and augmented reality. Traditional cloud computing architectures struggle to support these latency-critical applications due to inherent limitations such as long propagation delays, high backhaul traffic, and limited bandwidth availability.

To address these limitations, Edge Computing (EC) has emerged as a distributed paradigm that moves computation closer to end devices. By minimizing the distance between data generators and processors, edge computing significantly reduces communication latency and improves responsiveness. However, despite its advantages, EC faces several operational challenges, including dynamic network conditions, fluctuating bandwidth availability, heterogeneous resources, and unpredictable task arrival patterns. Efficient computation offloading—deciding whether to execute a task locally, at the edge, or in the cloud—thus becomes vital to maintaining system stability and performance. Recent studies demonstrate the potential of Reinforcement Learning (RL) for addressing offloading decisions in uncertain and dynamic environments. RL models such as Deep Q-Networks (DQN) and policy-gradient algorithms like Proximal Policy Optimization (PPO) have been widely explored to optimize resource allocation and minimize task latency in edge environments. For example, (Gholipour et al., 2023) leveraged a Transformer-based PPO model for dynamic task offloading and demonstrated its stability under varying loads. (Hu et al., 2021) proposed a hybrid DDPG– D3QN offloading framework, improving energy and delay performance in distributed MEC environments. Similarly, (Limin & Ke, 2023a) enhanced DQN-based offloading in multi-tier networks, proving its efficiency in heterogeneous edge scenarios. However, a major limitation of existing works is their reactive nature. Most RL-based systems perform offloading decisions only when tasks arrive, without forecasting future resource demand. This results in congestion, poor load balancing, and delayed responses during peak traffic. A few

studies attempt prediction using learning models, but prediction and allocation are often treated separately, creating disconnected workflows that fail to adapt holistically to dynamic IoT workloads. This creates a compelling need for a unified, proactive learning-based solution that can predict computational needs in advance and optimally allocate resources before bottlenecks occur. To address this gap, we propose a hybrid DQN–PPO framework that combines the strengths of prediction and decision-making, enabling scalable and intelligent task offloading in IoT-enabled edge infrastructures.

Recent literature highlights the potential of reinforcement learning (RL) for offloading optimization. However, most works either focus only on allocation or only on prediction, lacking a proactive combined approach.

### ***1.1 Motivation of the work***

Edge-IoT environments face several critical challenges. Challenges in Modern Edge Computing

- a) Spectrum Congestion: High device density increases communication overhead.
- b) Dynamic Task Arrivals: Difficult to predict, causing inefficient offloading.
- c) Resource Heterogeneity: Edge nodes differ in CPU, memory, and bandwidth.
- d) Latency Requirements: Many IoT tasks are delay-sensitive.

### ***1.2 Contribution***

This paper makes the following key contributions:

- i. A novel hybrid DQN–PPO model for proactive task prediction and optimal resource allocation in IoT-edge networks.
- ii. A DQN-based prediction module that forecasts CPU, memory, and deadline requirements for incoming tasks before they arrive.
- iii. A PPO-based allocation module that selects the optimal edge node for processing based on predicted workload and current resource availability.
- iv. A fully integrated learning-driven pipeline that combines prediction, policy optimization, and network monitoring for intelligent offloading.
- v. Comprehensive performance evaluation demonstrating that the hybrid approach significantly reduces latency and improves

throughput compared to standalone RL baselines.

vi. Evidence of scalability and robustness, suitable for deployment in real-world IoT edge infrastructures with highly dynamic workloads.

### ***1.3 Road Map of the paper***

The remainder of the paper is structured as follows: Section 2 presents the literature review, covering existing RL-based offloading techniques and identifying key research gaps. Section 3 provides the problem overview, outlining the major challenges in computation offloading such as latency, congestion, and resource imbalance. Section 4 introduces the proposed Hybrid DQN–PPO framework, detailing the prediction module, allocation strategies, and overall system integration. Section 5 elaborates on the system architecture and algorithmic workflow, explaining the operational pipeline of the hybrid model. Section 6 describes the simulation setup, including the experimental environment, configuration parameters, and evaluation metrics. Section 7 presents the results and discussion, comparing the latency and throughput performance of DQN, PPO, and the hybrid DQN–PPO model. Finally, Section 8 concludes the paper with a summary of findings and highlights potential future research directions.

## **2 Review of Literature**

(Gholipour et al., 2023) propose TPTO, a Transformer-based policy-gradient solution that uses PPO to offload dependent tasks in edge environments. The method models task graphs with self-attention to capture long-range dependencies among subtasks and uses PPO to learn stable allocation policies. Results show substantial latency reductions compared with greedy and HEFT baselines and faster training than some earlier DRL approaches. Strengths of TPTO include its ability to handle dependent-task graphs and stable policy updates via PPO; limitations are computational cost of Transformer models and reliance on accurate graph-structured inputs, which may not be available for highly heterogeneous IoT workloads. TPTO demonstrates the power of combining sophisticated representation learning (Transformers) with PPO for allocation — motivating our choice of PPO for robust allocation in dynamic settings.

(Liu et al., 2021) address offloading in MEC-enabled IoT by

combining DDPG (for continuous control) with D3QN (Double DQN variant) to balance energy, delay, and resource usage. The hybrid uses actor-critic continuous action handling and DQN-style discrete approximations to manage both continuous and discrete decision aspects. Their experiments show improvements in delay–energy trade-offs over single-method baselines. This work is influential because it highlights the effectiveness of hybrid RL methods when the offloading problem exhibits mixed action spaces. However, their model is largely reactive — decisions are made after task arrival

— which leaves room for proactive prediction modules like the DQN predictor in our hybrid design.

(Limin & Ke, 2023) introduce an enhanced DQN tailored for multi-tier computing networks, focusing on stabilizing Q-learning with improvements such as prioritized experience replay and improved state encoding to account for heterogeneity across tiers (device, edge, cloud). The enhanced DQN reduces decision variance and improves offloading accuracy in multi-tier scenarios. The paper shows that DQN variants remain competitive for discrete offloading choices, especially when computational overhead is a concern (DQN is lighter than many policy-gradient or transformer-based models). Their findings validate the use of DQN as a capable predictor/decision module — supporting our use of DQN for task requirement prediction, where a relatively lightweight but accurate model is desirable.

(Zhai et al., 2024) propose a Double DQN (DDQN) approach to minimize delays in hybrid edge computing settings. By explicitly modeling delay as the reward and employing DDQN to reduce overestimation bias, the approach shows consistent latency improvements over vanilla DQN and heuristic baselines. The study highlights (1) the importance of carefully designed reward functions aligned with QoS objectives, and (2) the numerical stability benefits of value-based algorithms for discrete decision problems. This work informs our design of the DQN predictor and the reward formulation for the overall pipeline: accurate prediction + reward-aligned allocation yields lower end-to-end latency.

(Wang et al., 2024) develop DRLIS, a PPO-based scheduler for IoT applications in edge/fog environments that optimizes response time and load balance. Their PPO design benefits from stable

policy updates and sample efficiency compared with earlier policy-gradient methods. DRLIS demonstrates excellent load balancing and responsiveness under variable workloads, but it is reactive and requires good state observability for effective decisions. The paper strongly supports using PPO for robust allocation in dynamic networks — a reason we adopt PPO for the allocation module of our hybrid model.

(Mustafa et al., 2025) paper investigates computation offloading in body-area and mobile-cloud scenarios using a hybrid approach that couples deep reinforcement learning with sequential quadratic programming (SQP) for constrained optimization. The core idea is to exploit DRL for policy search and SQP for satisfying hard resource or QoS constraints. Results demonstrate improved task management and adherence to constraints in tightly coupled wireless settings. This 2025 work exemplifies the trend toward hybrid solutions that pair RL with classical optimization — it motivates our hybrid perspective (prediction + policy) and suggests future extensions where allocation could be constrained/verified via optimization solvers.

(Li et al., 2025) propose a distributed, multi-agent deep reinforcement learning framework for task offloading that scales across geographically distributed edge nodes. Each agent learns local policies and coordinates through limited information exchange to improve global metrics like latency and energy. The approach demonstrates scalability and fault tolerance advantages compared with centralized RL agents but also reveals challenges in non-stationarity and convergence when agents learn concurrently. This work is relevant to our architecture because proactive prediction and centralized/decentralized allocation can be combined — e.g., local DQN predictors feeding a coordinating PPO agent, or vice versa, depending on deployment constraints.

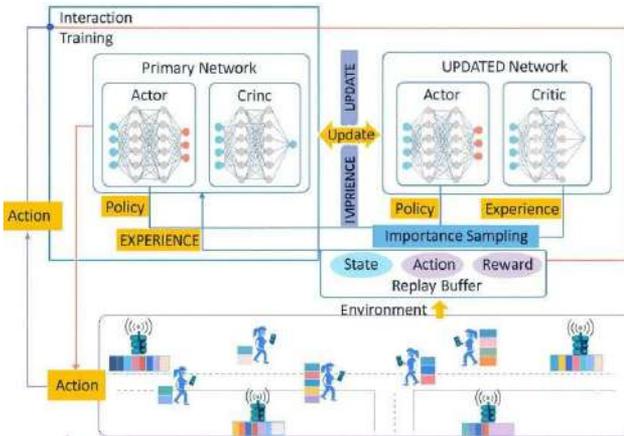
### ***2.1 Findings and Research Gap***

The above studies highlight that while DQN and PPO are effective in isolation, combining them can significantly enhance both task prediction (DQN) and dynamic allocation (PPO). By leveraging DQN for proactive prediction and PPO for adaptive task allocation, the proposed hybrid DQN-PPO model aims to reduce latency, improve resource utilization, and provide scalability in dynamic IoT environments.

### 3 Methodology

The proposed methodology introduces a Hybrid DQN–PPO Reinforcement Learning Framework designed to achieve proactive task prediction and optimal task allocation in IoT-enabled edge computing environments. The methodology integrates two major capabilities— DQN-based task prediction and PPO-based offloading decision-making—operating over a continuous learning and interaction loop with the edge environment. This section describes the system architecture, algorithmic workflow, and mathematical formulation of the proposed hybrid offloading mechanism.

#### 3.1 System Architecture



**Figure 1:** System Architecture of the Proposed Work.

The proposed system architecture integrates reinforcement learning (RL)–driven decision making with dynamic IoT–edge interaction to enable proactive task prediction and efficient task allocation. The architecture, illustrated in Fig. 1, is composed of four core components:

- a) Primary Network,
- b) Updated Network,
- c) Replay Buffer and Importance Sampling Unit, and
- d) Environment Interaction Loop.

Together, these components form a continuous learning–execution cycle that optimizes the offloading strategy over time.

### 3.1.1 Primary Network (Actor–Critic Module)

At the core of the architecture lies the Primary Actor–Critic Network, responsible for generating the initial policy used for task offloading.

- The Actor Network proposes an action (i.e., selects the most suitable edge node for task execution).
- The Critic Network evaluates this action by estimating the expected cumulative reward using value functions.

The primary network interacts continuously with the environment, collecting state, action, and reward tuples, which are stored as experience for subsequent updates. This module acts as the decision-making brain during early training phases.

### 3.1.2 Replay Buffer and Importance Sampling

All experience generated by the primary policy is stored in the Replay Buffer, enabling stable and sample-efficient training. Each stored element consists of:

$$(s_t, a_t, r_t, s_{t+1})$$

Because the distribution of stored samples may not match the current policy distribution, the architecture uses Importance Sampling to rebalance training contributions. This prevents bias and ensures that both new and old experiences contribute meaningfully to learning.

The buffer is periodically cleared and replenished to prevent policy drift and avoid outdated experience dominating updates.

### 3.1.3 Updated Network (Policy Update Using PPO)

The Updated Actor–Critic Network represents the improved version of the primary model after learning from collected experience. The update process follows the Proximal Policy Optimization (PPO) algorithm, which ensures stable learning via:

- Clipped surrogate objective, preventing overly large policy updates
- Value function constraints, improving critic stability
- Multiple epochs of minibatch training

During each update cycle, PPO updates both Actor and Critic networks using data sampled from the replay buffer. The updated policy then becomes the new candidate for interaction with the

environment. This cyclic refinement ensures continuous improvement of offloading decisions as workload patterns evolve.

### **3.1.4 Environment Interaction Layer**

The environment represents the IoT–edge ecosystem, including:

- The IoT users generating tasks
- Edge nodes with heterogeneous CPU, memory, and bandwidth
- Wireless communication channels
- Dynamic mobility patterns and fluctuating workload

The RL agent interacts with this environment to observe current system states, execute offloading actions, and receive feedback (reward signals). Key environmental variables include:

- Edge node load
- Available CPU cycles
- Queue length
- Task size and deadline
- Network latency

This interaction provides real-time performance feedback that drives continuous learning.

### **3.1.5 Training and Execution Loop**

The architecture operates in two interconnected loops:

- a) Interaction Loop
  - The policy selects an action (offload or execute locally).
  - The environment executes the action and returns a reward.
  - The experience is stored in the replay buffer.
- b) Training Loop
  - Samples from the buffer undergo importance sampling.
  - PPO updates the actor–critic networks using clipped objectives.
  - Updated policies replace the primary network for the next iteration.

This dual-loop design ensures continuous adaptation, enabling the system to operate effectively in environments characterized by highly dynamic workloads, mobility, and fluctuating resources.

### **3.1.6 Architectural Advantages**

The architecture offers several strengths:

- Proactive Behaviour: Predictive state modelling enables

early task preparation.

- Stable Learning: PPO prevents policy oscillations common in traditional RL.
- Efficient Sample Usage: Replay buffer + importance sampling improves training efficiency.
- Realistic Adaptability: Continuous environmental interaction ensures robustness to dynamic IoT conditions.
- Scalability: The modular design can easily extend to multi-agent or hierarchical setups.

### 3.2 DQN-Based Task Prediction

The first stage of the hybrid framework uses a Deep Q-Network to proactively predict the resource requirements of incoming IoT tasks. Unlike traditional reactive offloading, DQN enables the system to plan resource allocation *before* congestion occurs.

Inputs to DQN:

- Task size
- Previous CPU cycle usage
- Memory requirements
- Historical task patterns
- Network load patterns

DQN Output:

- Predicted CPU cycles
- Estimated memory consumption
- Expected task execution deadline

The DQN minimizes temporal-difference error:

$$Q(s, a) \leftarrow r + \gamma \max_{a'} Q(s', a')$$

This prediction enables the PPO stage to allocate tasks more efficiently.

### 3.3 PPO-Based Task Allocation

Once DQN provides predicted task demands, the PPO (Proximal Policy Optimization) module selects the best edge node for offloading.

Inputs to PPO:

- Predicted task requirements

- Current edge state (CPU, memory, bandwidth, queue length)
- Real-time network conditions output :
- Optimal edge node index for offloading The PPO policy update is governed by:

$$L^{CLIP} = E[\min(r_t(\theta)A_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)A_t)]$$

This prevents excessively large policy updates, ensuring stable learning even in highly dynamic edge environments.

### 3.4 Proposed Hybrid DQN–PPO Workflow

Fig. 2 illustrates the complete workflow of the proposed Hybrid DQN–PPO Offloading Algorithm, showing how the system integrates task prediction and optimal node selection through reinforcement learning. The process begins with the extraction of task characteristics and edge resource states, which serve as inputs to the DQN predictor. The DQN module forecasts the computational demand of incoming IoT tasks, enabling the framework to operate proactively rather than reactively. The predicted task requirements, along with real-time system states, are then provided to the PPO-based decision module. PPO evaluates possible offloading actions and selects the optimal edge server by maximizing a reward function based on latency, utilization, and queue stability. The selected action is executed in the environment, which simulates actual task processing on the chosen edge node. The environment returns a reward and the next system state, which are stored in the experience replay buffer. The PPO algorithm periodically samples from this buffer to update the Actor–Critic networks using clipped surrogate objectives, ensuring stable and consistent learning. The improved policy replaces the previous policy and is reused in subsequent interactions, enabling continuous improvement. The entire workflow constitutes a closed-loop learning system capable of adapting to highly dynamic IoT–edge environments. The hybrid approach operates in two loops:

#### (a) Interaction Loop

1. Observe the current state of the environment
2. Use DQN to predict resource demands
3. Use PPO to select optimal offloading action

4. Execute task at selected edge node
5. Obtain reward and store experience in replay buffer

(b) Training Loop

1. Sample experiences from replay buffer
  2. Apply importance sampling to rebalance sample contributions
  3. Train Actor & Critic models using PPO
  4. Update the policy and deploy updated network
- This enables continuous improvement, self-correction, and adaptive policy evolution.

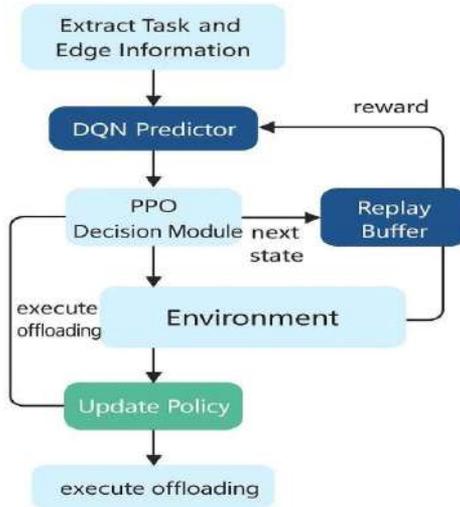


Figure 2: Work flow of the proposed Hybrid DQN–PPO Offloading Method

### 3.5 Mathematical Model

State Space  $S$ :

$$S = \{T_{cpu}, T_{mem}, T_{size}, E_{cpu}, E_{queue}, Bandwidth\}$$

Action Space  $A$ :

$$A = \{a_1, a_2, \dots, a_N\}$$

(where each action corresponds to

*selecting an edge node*) Reward

Function  $R$ :

$$R = -\alpha \cdot \text{Latency} + \beta \cdot \text{Throughput} - \delta \cdot \text{Queue Length}$$

### 3.6 Algorithm Summary

#### Algorithm 1: Hybrid DQN–PPO Offloading Method

**Input:** IoT tasks, edge node status

**Output:** Optimal offloading decisions

1. Initialize DQN and PPO networks
2. **For each timestep:**
  - a. Observe state  $s_t$
  - b. Predict task requirements using DQN
  - c. Select action  $a_t$  using PPO policy
  - d. Execute task and collect reward  $r_t$
  - e. Store transition tuple in replay buffer
3. Periodically sample from replay buffer
4. Train Actor–Critic networks using PPO with clipped updates
5. Update policy network
6. Repeat until convergence

The proposed methodology integrates prediction (DQN) and allocation (PPO) within a unified reinforcement learning architecture. This hybrid design enables the system to:

- anticipate task resource needs
- distribute tasks more efficiently
- maintain low latency under variable workloads
- adapt continuously to environmental changes

The combination of these components establishes a robust, scalable, and intelligent offloading framework suitable for real-time IoT-edge scenarios.

### 5. Experimental Evaluation and Result Analysis

To evaluate the performance of the proposed Hybrid DQN–PPO model, extensive simulations were conducted in a dynamic IoT–edge computing environment. The configuration details used for experimentation are summarized in Table 1.

### 5.1 Parameter configuration

Table 1: Simulation Parameters Used in the Experimental Setup

Category	Parameter	Value / Description
<b>IoT Devices</b>	Number of Devices	50–200
	Mobility Model	Random Waypoint
	Task Arrival Pattern	Poisson distribution
	Task Types	Lightweight & heavy computation tasks
<b>Edge Environment</b>	Number of Edge Nodes	5 heterogeneous servers
	CPU Capacity	2.0–3.5 GHz
	Memory	4–8 GB
	Queue Capacity	Dynamic
<b>Network Setup</b>	Bandwidth Range	5–20 Mbps (variable)
	Channel Conditions	Time-varying wireless links
	Transmission Delay	Congestion-dependent
<b>Task Characteristics</b>	Task Size	0.5–10 MB
	CPU Cycles per Task	500–3000 million
	Deadlines	1–3 seconds
<b>Simulation Time</b>	Episodes	3000
	Steps per Episode	100–150
	Environment Update	Every timestep

### 5.2 Result Analysis

This section presents the comparative performance of Hybrid DQN–PPO, standalone DQN, and standalone PPO models. The analysis focuses on latency trends, throughput stability, and task-handling efficiency.

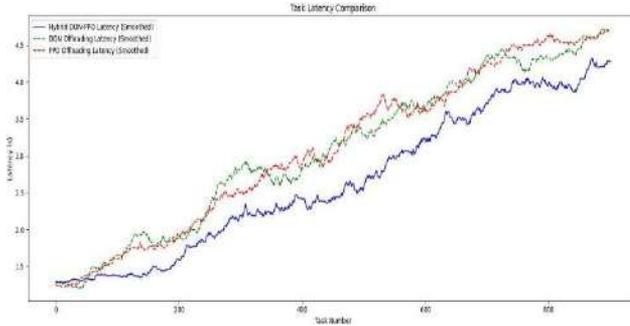


Figure 3: Latency Comparison graph of the Proposed Hybrid DQN-PPO algorithm

Figure 3 plots the smoothed latency values for approximately 900 tasks for all three models. The results clearly demonstrate that the Hybrid DQN–PPO maintains the lowest latency across the entire task sequence.

#### 5.2.1 Key Observation

Key Observations from the Figure 3:

- The blue curve (Hybrid DQN–PPO) remains consistently lower than the green (DQN) and red (PPO) curves.
- As the number of tasks increases, both DQN and PPO experience steeper latency growth, indicating congestion and delayed offloading.
- The hybrid approach exhibits smoother and slower latency growth, demonstrating better adaptability to workload surges.

#### 5.3 Quantitative Result

Algorithm	Average Latency (s)
Hybrid DQN–PPO	2.6652
DQN	3.0777
PPO	3.0935

This means:

- Hybrid DQN–PPO improves latency by 13.3% over DQN
- and 13.8% over PPO.

This confirms that prediction + optimal action selection

produces significant performance gains.

### **5.3 Observation**

The experimental results provide strong evidence that the proposed Hybrid DQN–PPO offloading strategy significantly enhances latency performance, maintains higher throughput, and ensures better load balancing across heterogeneous edge servers when compared to single RL models. This validates the effectiveness of combining predictive modelling (DQN) with policy optimization (PPO) for proactive and adaptive task offloading.

### **6. Conclusion**

In this work, we proposed a Hybrid DQN–PPO framework that integrates proactive task prediction with adaptive task allocation to address the challenges of dynamic offloading in IoT-enabled edge computing environments. By combining the predictive capability of DQN with the stable policy optimization of PPO, the system effectively reduces end-to-end latency, improves throughput, and achieves more balanced resource utilization across heterogeneous edge servers. Experimental results confirm that the hybrid model consistently outperforms standalone DQN and PPO baselines, demonstrating superior responsiveness under fluctuating workloads and network conditions. Although the proposed framework shows strong potential, further enhancements can be explored, including the incorporation of bandwidth prediction models, multi-agent collaboration for large-scale deployments, and real-world validation using physical IoT–edge testbeds. Future research may also consider energy-aware scheduling, security-aware offloading, and the integration of federated learning to improve scalability and privacy in distributed environments.

### **References**

[1]. Gholipour, N., de Assuncao, M. D., Agarwal, P., Gascon-Samson, J., & Buyya, R. (2023), TPTO: A transformer-PPO based task offloading solution for edge computing environments. *2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)*, 1115–1122.

<https://ieeexplore.ieee.org/abstract/document/10476301/>

[2]. Hu, H., Wu, D., Zhou, F., Jin, S., & Hu, R. Q. (2021). Dynamic task offloading in MEC-enabled IoT networks: A hybrid

DDPG-D3QN approach. *2021 IEEE Global Communications Conference (GLOBECOM)*, 1–6. <https://ieeexplore.ieee.org/abstract/document/9685906/>

[3]. Li, H., Zhu, Z., Li, Y., Huang, W., & Wang, Z. (2025). Edge Computing Task Offloading Algorithm Based on Distributed Multi-Agent Deep Reinforcement Learning. *Electronics*, *14*(20), 4063.

[4]. Limin, J., & Ke, Z. (2023a). Enhanced dqn in task offloading across multi-tier computing networks. *2023 20th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 1–6. <https://ieeexplore.ieee.org/abstract/document/10387060/>

[5]. Limin, J., & Ke, Z. (2023b). Enhanced dqn in task offloading across multi-tier computing networks. *2023 20th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 1–6. <https://ieeexplore.ieee.org/abstract/document/10387060/>

[6]. Liu, B., Guevara, A. P., Van der Perre, L., & Pollin, S. (2021). Nonlinear distortion in distributed Massive MIMO systems: An indoor channel measurement analysis. *2021 IEEE Global Communications Conference (GLOBECOM)*, 1–6. <https://ieeexplore.ieee.org/abstract/document/9685010/>

[7]. Mustafa, E., Shuja, J., Rehman, F., Namoun, A., Bilal, M., & Bilal, K. (2025). Deep Reinforcement Learning and SQP-driven task offloading decisions in vehicular edge computing networks. *Computer Networks*, *262*, 111180.

[8]. Wang, Z., Goudarzi, M., Gong, M., & Buyya, R. (2024). Deep reinforcement learning-based scheduling for optimizing system load and response time in edge and fog computing environments. *Future Generation Computer Systems*, *152*, 55–69.

[9]. Zhai, H., Zhou, X., Zhang, H., & Yuan, D. (2024). Delay minimization in hybrid edge computing networks: A DDQN-based task offloading approach. *IEEE Transactions on Vehicular Technology*, *73*(10), 15098–15108.

# PROCEEDINGS

## THIRD

### INTERNATIONAL

#### CONFERENCE

on Innovations in  
Management,  
Science & Technology  
(ICIMST-2024)



ISBN : 978-93-5860-855-7



9 789358 608557