

MSIT Journal of Research –SATYAM

(TECHNOVATION- Special Issue, 2023)

Patron

Er. Kaptan Singh Ji

Chairman, Surajmal Memorial Education Society

Prof. Prem Vrat

Pro-Chancellor; Professor of Eminence and Chief Mentor, The NorthCap University, Gurugram
Former Director, IIT Roorkee
Former Vice-Chancellor, UPTU Lucknow
Former Professor and Division Chairman, Asian Institute of Technology, Bangkok, Thailand

Prof Ved Prakash

Former Chairman, University Grant Commission

Prof. Saroj Kaushik

Department of computer science & Engineering
IIT, Delhi

Prof. Archana Balyan

Director
Maharaja Surajmal Institute of Technology

Dr. Vijander Singh

Professor, Department of Instrumentation and Control Engineering, NSUT, Delhi

Dr Tejbir Singh Rana

Associate Professor
Shivaji College, University of Delhi

Prof Devi Singh

Former Director, IIM Lucknow and MDI Gurgaon
Founding Vice Chancellor, FLAME University, Pune

Dr. (Mrs.) Amita Dev

Vice Chancellor, IGDTUW, New Delhi

Dr N.K. Gupta

Member, National executive Council, ISTE, Delhi

Prof. S.D. Joshi

Department of Electrical Engineering
IIT, Delhi

Dr Sushil Chandra

Scientist 'G', DRDO

Prof Narendra Kumar

Professor, Department of Electrical Engineering, Delhi Technological University, Delhi

Dr T.V. Vijay Kumar

Professor, School of Computer & Systems
JNU, Delhi

Editor-in-Chief: Dr. Shaifali M. Arora

Faculty Editors: Dr Geetika Dhand, Dr Sakshi Rajput,

Dr Parul Chaudhary, Dr Jindagi Kumari

Student Editor: Mr Yash(5th Sem,ECE(M), Ms. Rajshree Chandra (3rd Sem,CSE(M)),

Mr. Saksham Arora (3rd Sem, CSE(M))



MSIT Journal of Research – SATYAM

TECHNOVATION (Special Issue)

MAHARAJA SURAJMAL INSTITUTE OF TECHNOLOGY

C-4, Janakpuri, New Delhi-110058(India)

Phones: 011-45037193/45656183

E-mail: satyamjournal@msit.in, Visit us at www.msit.in

(2023-24)

**MSIT Journal of Research – SATYAM (MSITJR)
TECHNOVATION (Special Issue)**

is a special issue of the Journal that is specially published to encourage the research among students
Maharaja Surajmal Institute of Technology
C-4 Janakpuri, New Delhi – 110058

Editor-in-chief: **Dr. Shaifali M. Arora**

Copy Right © MSITJR – (2023-24)

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical including photocopying, recording or by any information storage and retrieval system, without the prior written permission from the copyright owner. However, permission is not required to copy abstracts of papers on condition that a full reference to the source is given.

ISSN 2319-7897

Disclaimer

The opinions expressed and figures provided in this journal are sole responsibility of the authors. The publisher and the editor bear no responsibility in this regard. Any or all such liabilities are disclaimed.

All disputes are subject to Delhi jurisdiction only

Address for Correspondence Dr. Shaifali M. Arora

Editor-in-chief, MSIT Journal of Research, Associate
Professor, Department of ECE Maharaja Surajmal
Institute of Technology (MSIT) C-4, Janakpuri, New
Delhi-110058
Tel: 011-45037193/45656183
E-mail: satyamjournal@msit.in

Published and printed under the guidance of Prof. Archana Balyan, Director, Maharaja Surajmal Institute of Technology (MSIT), C-4, Janakpuri, New Delhi-110058 (INDIA), Tel: 011-45037193/45656183
Email: director@msit.in Visit us at www.msit.in

Contents

Sr. No.	Title	Page No.
1.	Virtual Mouse using Hand Gesture <i>Naresh Kumar, Meetu Malhotra</i>	1
2.	Blockchain based AI Yoga Assistant <i>Tushar Ojha, Vaibhav Krishan Yadav, Anshika Gupta, Koyel Datta Gupta</i>	9
3.	Blockchain Technology-based Messenger <i>Sonia Rathee, Aditi Diksha, Aman Gupta, Ankur Badola</i>	14
4.	CODEIAL: A solution for the coders by the Coders to the Coders <i>Prabhjot Kaur, Ria Mittal, Vaibhav Hasija, Aakriti Sharma, Varun Sharma</i>	19
5.	Adaptive Algorithm for High-Speed Integrated Filters <i>Archana Balyan, Deepshikha Yadav</i>	24
6.	Career Guidance AI Counsellor <i>Sapna Malik, Disha Katyal, Naman Sachdeva, Nihal Gupta</i>	28
7.	General Purpose Artificial Intelligence System with Graphical User Interface <i>Deepak Sarangi, Shaifali M. Arora, Poonam Dahiya, Anshul Pareek</i>	36
8.	Enhancing Student Feedback Analysis: A hybrid Approach utilizing Machine Learning and Lexicon-Based Methods for Sentiment-Analysis <i>Vimal Gaur, Shivam, Rishabh Bhatt, Shivansh Tripathi</i>	40
9.	Malicious URL Detection on IOT <i>Sangeeta, Sitender, Savita Ahlawat, Karthik Jangid, Jayashy Sinha, Gitali Daryani</i>	44
10.	Quantitative Analysis and Ensemble Learning Based Predictive Modelling Of Taxi Trip Duration in New York City <i>Rinky Diwedi, Jivitesh Sabharwal, Apoorva Singh, Himanshu Jha</i>	50
11.	MOTION CONTROL: Camera-Based Body Movement Input Device Simulation <i>Kavita Sheoran, Priyaansh Malik, Pratyush Jha, Mayank Sharma</i>	56
12.	Prediction of Heart Disease Using Ensemble Learning and Ant Colony Optimization <i>Lititia Khanna</i>	64
13.	Image Manipulation Detection System <i>Vinita Rohilla</i>	70
14.	Digitalizing Warranty Using Blockchain Technology <i>Nishtha Jatana, Nishant Sharma, Kunal Bisht, Hardik</i>	74
15.	Multimodal Emotion Detection <i>Savita Ahlawat, Sangeeta, Shivam, Saiyam Jain, Surya Prakash Gup</i>	78
16.	Solar PV System based Microgrid for MSIT <i>Sunil Gupta, Ritik Vasudeva, Rohit Bisht</i>	86

17.	Solar Power Operated Refrigeration System Based on Peltier Effect <i>Sachit Rathee, Rajneesh Kumar Prasad, Ayaz Khan, Deepanshu Solani, Hariom Yadav</i>	91
18.	Audio noise reduction and Video transcription <i>Geetika Dhand, Sukant Thakur, Amitesh Kumar, Brijesh Kumar</i>	96
19.	REVIEWERS	99

Virtual Mouse using Hand Gesture

Naresh Kumar¹, Meetu Malhotra²

¹Department of Computer Science and Engineering,
Maharaja Surajmal Institute of Technology, New Delhi, India

²Harrisburg University of Science and Technology, USA

Abstract-This project introduces a methodology in the field of Human-Computer Interaction (HCI), where the development of the cursor can be controlled by utilizing a real-time camera. It is an option in contrast to the ongoing strategies including manual contribution of buttons or changing the places of an actual PC mouse. With all other factors held constant, it uses a camera and PC vision innovation to control different mouse occasions and is equipped for playing out each undertaking that the actual PC mouse can. The Virtual Mouse color recognition program will constantly capture real-time images, which will undergo a series of filtration and conversion. Once the process is complete, the program will apply the image processing technique to obtain the coordinates of the targeted color position from the converted frames. Next, it will compare the existing color within the frames with the list of color combinations, where different combinations consist of different mouse functions. If the current color combination finds a match, the program will execute the corresponding mouse function, which will be translated into an actual mouse action on the user's machine. The mouse is one of the wonderful inventions of Human-Computer Interaction (HCI) technology. Currently, a wireless mouse or a Bluetooth mouse still uses devices because it uses a battery for power and a dongle to connect it to the PC. Researchers around the world are now focused on making devices more interactive and trying to make the devices operational with minimal physical contact. This research proposes an interactive computer system that can operate without any physical keyboard and mouse. This set up can be beneficial to everyone, especially the paralyzed people who face difficulties operating a physical keyboard and mouse. It uses computer vision through which the user can type on a virtual keyboard using a yellow-colored cap on his fingertip and can also navigate to the mouse controlling system. Once the user is in mouse controlling mode, the user can perform all the mouse operations only by showing the different number of fingers. It has multiple benefits such as functioning in changing light levels or with complex background, following exact fingertip at a more drawn-out distance, and detecting fingertip at the rate of 60 frame rate. It achieved an average accuracy of approximately 80%.

1. INTRODUCTION

Virtual mouse using hand gesture is a technique that can be implemented using the modern fields technology machine learning and artificial algorithm. Virtual mouse helps in building the various boundary breakers which could transform the hardware industry, or this can be a steppingstone towards various other in contribution in the field of science which could ease the human workload. Virtual mouse clearly forms an independent contribution to those which have physical motor disability and cannot use

hardware devices such as physical mouse. The purpose of this study is not only to build a hardware free solution for PC but also to introduce a solution for the dependency on physical mouse while using PC. This study introduces the concept of virtual mouse which can do all operations but without using any physical mouse. The research uses PC webcam to capture the images that can be used as a frame and helps the computer to decide words. It performs all mouse related tasks such as left click, right click, select all, various other operations. It captures frames at a rate of approximately 60 frames per second, which is close to real time and hence can be used in real-world application. This field of study will clearly bring a transformative step. With the continuous advancement of computer technology, the significance of human-computer interaction is experiencing tremendous growth. Touch screen technology in mobile devices is trending. However, this technology is expensive in desktop systems. PC vision strategies could be an elective way for touch screen and make a virtual human-PC cooperation gadget. For example, a mouse or console utilizing a webcam. In this review, finger following in view of a virtual mouse application has been planned and carried out utilizing a normal webcam. The motivation behind this project was to develop an interactive application that enables users to interface with a computer and foster a virtual human-computer collaboration tool. The task aims to create a Virtual mouse, utilizing hand signals. It has been ages since have been utilizing hand motions for conveying in human culture. The shaking of hands, approval and disapproval signs have been truly existing in the climate. It is widely acknowledged that signals are the simplest and most direct way of communicating with others, as examined in [1]. The underlying arrangement incorporates a minimal expense USB web camera that can be utilized to capture frames. The total interaction is divided into 4 stages, which are outline catching, picture handling, district Extraction, highlight coordinating. The most useful and expressive technique for human correspondence is through hand movements, which is a by and large recognized language. In this work, a persistent hand signal structure is proposed. Exploratory plan of the system uses a proper position negligible cost web camera prevalent quality recording feature mounted on the most elevated place of screen of a PC or a good camera on a PC, which discovers a portrayal using Red Green Blue [RGB] assortment space from a legitimate distance as inspected in [2] and [3]. This work is apportioned into four stages, for instance, picture pre-dealing with, locale extraction, feature extraction, and part planning. Affirmation

and the interpretation of correspondence by means of signals is one of the critical issues for correspondence with the landfill and nearly deaf people. Many existing frameworks utilize a hand with glove, and they likewise accomplish utilizing 30 edges each second, yet it will be carried out in 60 casings each second. In this task, a successful hand signal division method has been proposed in view of pre-handling, foundation deduction, and edge discovery techniques [4]. Pre-handling is characterized as the system of planning information for another cycle. The primary target of the pre-handling process is to change the information into a structure that can be all the more successfully and easily handled. In the proposed work, the pre-handling strategies are made in view of various sorts of mixes from the ensuing hand motion picture handling activities, for example, catching the picture, eliminating clamor, foundation deduction, and edge recognition [5], [6] and these picture handling techniques are examined as per the following. Right away, the hand signal pictures are gotten from the vision-based camera. The hand movements should be visible with the different kind of association focuses like "data gloves" that definitively records each capturing point and digit and position sensors for wrists and optical course or electromagnetic, requiring the client to wear trackers or gloves [10]. Continually, the glove-based interfaces even need the client to be hitched to the PC, all restricting opportunity to client comfort and association point, then again, vision-based interfaces offer unlimited joint effort with a human.

2. LITERATURE REVIEW

Lately, an enormous examination has been done on the subject of virtual mouse frameworks utilizing hand signals. For the Completion of this venture, the paper [1] to [12] has been referred to and distributed from 2017 to 2022. With the assistance of a writing study done, it was seen that the fundamental advances engaged with the majority of these virtual mouse frameworks are: -

While Several Research papers and existing ventures gave various bits of knowledge on the issue proclamation close by, every one of them had a portion of different components of such an effective Virtual Mouse framework being set up.

The different Approaches recently utilized are: -

In [1], the hand following must be particularly adjusted for each client, so it is easy to understand. This framework was executed exclusively in a confined inlaying climate. This framework involves the HSV calculation for execution, however, it will utilize an extraordinary sensor (Kinect) to catch pictures and cycles them outline by outline. Considering every aspect, the client needs to spend more cash since it utilizes a sensor. The presence of other hued objects that are utilized behind the scenes could make the framework give an incorrect reaction. On the off chance that the camera goal is too high, the framework wouldn't work as expected or could run sluggish.

Research Paper [2], [3] centers around the camera module and location module.

2.1. CAMERA MODULE

The fundamental utilization of this module is to take input from various picture locators and send this information picture, got from the picture finder to the identification module. The Detection Module is utilized for handling and interaction, it outlines by outline. There are different techniques utilized for catching information gloves, cameras, and so on.

2.2. DETECTION MODULE

This module is utilized for picture handling. The result acquired from the camera module is subjected to various picture handling methods like variety change, commotion expulsion, and edge, following which the picture goes through form extraction. On the off chance that the deformities happen in the picture, convexity surrenders happen as per which the signal is recognized. On the off chance that there are no deformities, the picture is effectively characterized by utilizing the Hear fountain to identify the motion.

Research Paper [4], recommends that the momentum framework comprises a nonexclusive mouse and trackpad screen control framework, alongside the shortfall of a hand signal control framework. The utilization of a hand motion to get to the screen in a good way is preposterous. The current virtual mouse framework involves straightforward mouse tasks utilizing hand motions like left, right-snap, drag, etc. The utilization of hand acknowledgment, later on, won't be utilized because there is an assortment of frameworks for hand acknowledgment, the framework they utilized is static hand acknowledgment, which is essentially an acknowledgment of the shape made by the hand and the meaning of activity for each shape made, which is restricted to a couple of characterized activities and creates a great deal of turmoil. As innovation propels, there is something else and more options in contrast to utilizing a mouse. Coming up next are a portion of the procedures that were utilized:

a. Head Control:

An extraordinary sensor or inherent webcam can follow head development to move the mouse pointer around on the screen.

b. Eye Control:

It will empower clients to move the pointer on the screen by moving their eyes. Rather than mouse buttons.

c. Touch Screens:

Touch screens, which were once viewed as a specialty innovation, the outcome of cell phones and tablets, contact empowered windows workstations and across-the-board work areas are turning out to be more normal. Albeit, the boundless utilization of touch screens has brought about another arrangement of touch openness issues. Notwithstanding, every one of the strategies underneath has its arrangement of drawbacks. The utilization of the head or eyes to control the cursor consistently can be unsafe to one's well-being. This can prompt a few issues with well-being. While utilizing a touch screen, the client should keep up with their emphasis on the screen consistently, which can cause sleepiness. By looking at the accompanying methods, will expect to make another undertaking that won't hurt the client's well-being.

Research Paper [5] recommends that there are traditional methodologies that are generally founded on eye motions. For this situation, the Author S. Shetty fabricate a virtual mouse framework that depends on a variety of location. It involved a webcam for location purposes and recognizes the development of the cursor with the assistance of signal and snap occasions utilizing inbuilt Open CV capacities, which is normal too. In the brutal establishment, this system by and large crashes and burns. Then the Author, in [6], expanded a procedure without mouse cursor control where the errands of the cursor are obliged by using hand signals. This System includes an assortment of area guidelines for the grouping of hand movements by webcam. The inbuilt limit of Image Processing is used in for this technique. The maker, manufactured a system as kept in [7] for controlling the mouse pointer or cursor through a webcam that controls media player volume. They used RGB assortment tapes to see the client's finger.

In a research paper [8], the creator introduced a framework to develop a hand motion-based virtual mouse. This framework takes different hand signals by webcam and executes the mouse work as per the motion. The exactness accomplished by this framework is practically 80%-90%. In the unpleasant foundation, this framework doesn't work effectively. As seen from the writing check, past frameworks integrate either a virtual console or mouse. Those frameworks can't completely expel the requirement for a mouse and console. This work plans to make a virtual mouse utilizing hand signals that can be executed with next to no actual mouse

Research paper [9], have practical experience in the framework that works by perceiving the shade of hand tips thus deciding the cursor position, however different situations produce it difficult for the calculation to run inside the continuous climate. The proposed framework can work for the complexion of any tone likewise as can work precisely in any lighting condition, still to tap the client should make a 15-degree point between its two-finger the proposed framework can undoubtedly supplant the ordinary mouse.

Research Paper [10], proposes the virtual mouse framework utilizing hand signal identification by wearing a glove inside the hand and utilizing a variety of tips inside the hands for acknowledgment of motion since hand gloves mouse capacities and acknowledgment isn't more precise. The gloves likewise are not fitted to certain clients, and now and again, the ubiquity isn't so precise, and subsequently, the explanation for incorrect is the disappointment of identification of various tips. During this case, it'll execute without hand gloves.

3. CHALLENGES FACED

The actual mouse is exposed to mechanical mileage.

- An actual mouse needs exceptional equipment to work, without equipment
- can't utilize and work, yet when executed, it utilizes hand motion and can't need any unique hardware [5].
- An actual mouse isn't serenely flexible to various conditions and its

- presentation fluctuates. However, the virtual mouse is advantageous and versatile and its exhibition doesn't fluctuate, can use in each environment [1].
- The mouse has restricted work even in a present functional climate. What's more, such capacities are left, correct, and scroll [4], [6].
- All wired mice have their life expectancy and remote mouse have their life expectancy.
- The existing model shows accuracy and precision with gloves on the hands. The existing model displays a fps of around 30 and below.

4. OBJECTIVES

In Most, the PC touchpad isn't the most agreeable and advantageous, so for this situation, it will be useful, the mouse cursor will work with the assistance of hand signals rather than a touchpad.

- The principal objective is for the information gotten from pre-handling to be addressed so that it very well may be effectively handled and deciphered by the framework.
- The primary goal of the virtual mouse is to foster another way to the standard mouse system to execute and control the cursor limits with the help of a webcam that takes the hand signals and a short time later cycles the movement frame by frame [1] to play out the working of cursor like left, right-click, scroll(up and down).
- Reduce the expense of the equipment.
- To reduce the cost of the special gloves, by using bare or naked hands for recognition.
- Increase in the FPS detection of input video or through webcam from 30 to around 55-60.
- To want to work with the assistance of a webcam, as the webcam is careful to continuously get the photos. The application wouldn't work, accepting there is no webcam recognized [4].
- To design a piece of virtual data that can deal with all surfaces and the utilization of a virtual mouse will be sensible on all surfaces and indoor circumstances, as long as the clients are confronting the webcam while doing the improvement.

5. PROPOSED METHODOLOGY

This section will be providing us the methods and techniques used in the model and some of which are as show in fig 1 as : (1) Picture resize (2) RGB processing (3) Denoise of the image (4) Finger Detection Technique (5) Target Locking (6) Virtual Mouse.

In this, we will work on detection of both of the hands.

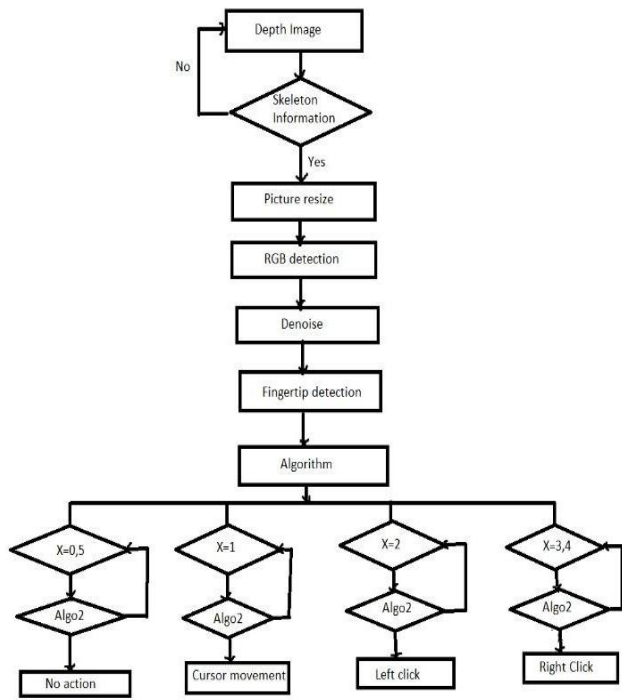


Fig 1: Architecture of the model

5.1 Picture Resize:

Denis Simakov et al. [7] propose a principled way to deal with outline of visual information in light of enhancement of a well-defined closeness measure. The issue we consider is retargeting of picture information into more modest sizes. The paperproposes a bidirectional similitude measure which quantitatively capuresthe picture frame by frame.

5.2 RGB processing:

The progression of handling the picture has been shown by means of flowchart in fig 2 and momentarily clarification of the cycles. When diverged from a toned picture, computational unpredictability is lessened in a dim scale picture. Likewise, the flipped picture is changed over into a dim scale picture. Every one of the fundamental undertakings were performed in the wake of changing over the image into dull scale. The dim locale of the image obtainedafter derivation should be exchanged over totally to an equal picture for finding the region of the distinguished item. A grayscale picture includes a lattice containing the potential gains of each and every pixel. The pixel values lay between the arrives at 0 to 255 where 0 tends to pure dim likewise, 255 addresses pure white tone.

5.3 Denoise of the image:

One of the basic troubles in the field of picture managing and PC vision is picture denoising, where the fundamental goal is to check the central picture by covering noise from a ruckus tainted rendition of the picture. Picture disturbance might be accomplished by various brand name (i.e., sensor) and outward (i.e., climate) conditions, which are critical of the time unreasonable to keep away from in useful circumstances. Subsequently, picture denoising expects an epic part in a wide level of purposes like picture patching up, visual following,

picture choice, picture division, and picture portrayal, where getting the main picture content is basic for strong execution. As displayed in fig 3, this denoising procedure gets the picture liberated from the hand.

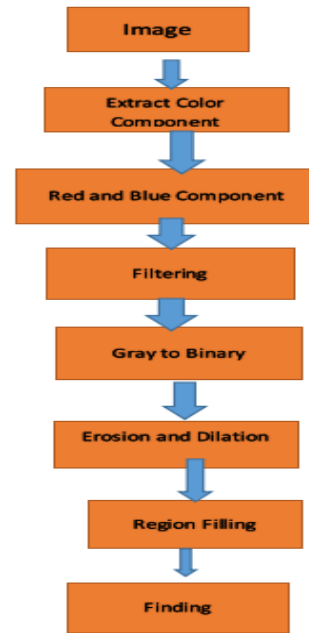


Fig 2: image processing sequence



Fig 3: Denoise hand image

5.4 Fingertip Detection Technique:

The K- Cosine algorithm helps in finding the distance between particular fieldpints which are beyond the measures of naked eye. The technique uses the reference of points in the trigonometry field which helps us dictate about the various points in the geometry field about whether hoe far the points are from each other. The appraisal enrolls environments direction obvious shape. conspicuous evaluation seeing conditions express things moreover revelation. attempts point

K Cosine Algorithm:

$$|\cos \alpha| = \frac{|a_i(K)b_i(K)|}{|a_i(K)||b_i(K)|} \dots(1)$$

Above is utilized to decide the fingertip areas, where the vectors are characterized as

$$a_i(K) = P(i+k)P_i \quad , \quad b_i(K) = P(i+k)P_i \quad \dots(2)$$

These structure point, shape points of P_i . PC based

understanding means the point between PC based information (K) and bi (K) for a given pixel Pi. The fact of the matter is a portrayed as an edge worth to and the reality of the situation is set to 45 degrees, which are sensible for most conditions.



Fig 4: distance between hands.

The construction points of Pi. Reenacted understanding exhibits the point between man-made intellectual prowess (K) and bi (K) for a given pixel Pi. The reality of the situation is depicted as an edge worth to see the and the reality of the situation sensible conditions.

These pints helps in finding the references in putting those fields into the respective contour which controls the mouse pointer used in reflecting the point of impact made by the various pixels. These pixels will dictate about the angle made between them and these points are therefore characterized as the primary objective for reflecting the pints for reference sin those filed of contour for points referencingpotential increases reenacted understanding, computation, in case worth repeated information really unassuming like edge regard, portrayed. How much saw how much steady following,packaging following

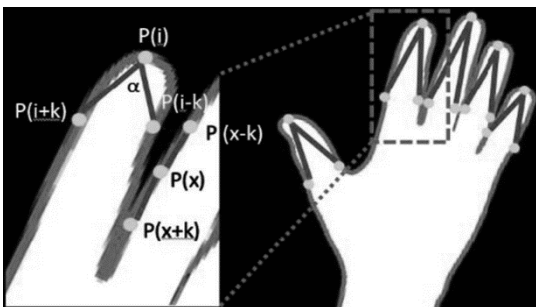


Fig 5: cosine algorithm

Euclidean distance: Euclidean distance is utilized to compute the distance between all the tip of the fingers (tops distinguished) and centroid. The equation for computing Euclidean distance is E.D

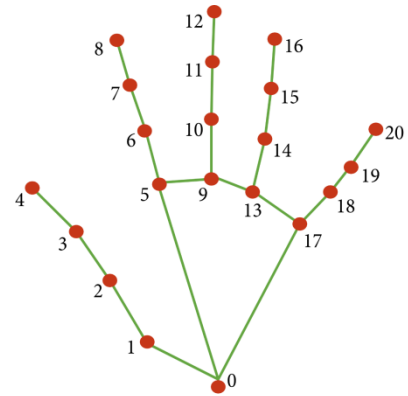
$$E.D(a,b) = (x_a - x_b)^2 + (y_a - y_b)^2 \dots (2)$$

Where 'a' addresses all the limit focuses and 'b' is the reference point which is the centroid itself. Distance is determined to remove the specific number of finger brought up in the picture. There might be a few pinnacles distinguished, which don't address the tip of the fingers. These pinnacles are immaterial pinnacles. It can dispose of these immaterial tops by working out the most extreme pinnacle. Putting the limit of esteem at 75% of the most extreme pinnacle, It can choose just those pinnacles whose values are more than this edge. The quantity of pinnacles found is equivalent to the quantity of finger brought up in the picture. Different pinnacles that

are distinguished however don't converge the edge line would be treated as irrelevant pinnacles or collapsed fingers.

5.5 Target Locking

At the point when numerous individuals are available, the designated individual is one type to the accompanying. In proposed work, the sensor gives extricated hand skeletal joints data for up to six individuals without a moment's delay, for example, pointer, center finger, little finger, thumb as displayed in Fig 6. Accordingly, the framework can find the fingertips involving the calculation above in condition 1. Be that as it may to instruct mouse pointer, really want distinguish objective individual wipe out the impact of the others.



- | | |
|-----------------------|-----------------------|
| 0. WRIST | 11. MIDDLE_FINGER_DIP |
| 1. THUMB_CMC | 12. MIDDLE_FINGER_TIP |
| 2. THUMB_MCP | 13. RING_FINGER_MCP |
| 3. THUMB_IP | 14. RING_FINGER_PIP |
| 4. THUMB_TIP | 15. RING_FINGER_DIP |
| 5. INDEX_FINGER_MCP | 16. RING_FINGER_TIP |
| 6. INDEX_FINGER_PIP | 17. PINKY_MCP |
| 7. INDEX_FINGER_DIP | 18. PINKY_PIP |
| 8. INDEX_FINGER_TIP | 19. PINKY_DIP |
| 9. MIDDLE_FINGER_MCP | 20. PINKY_TIP |
| 10. MIDDLE_FINGER_PIP | |

Fig 6: Target locking of skeleton joints.

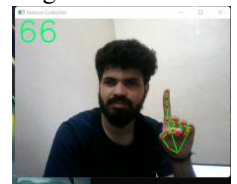
5.6 Virtual Mouse:

Weighted speed cursor control, get a capability of the finger of the constant picture and the past picture and figure the distance between the two. Move the cursor, expecting the opening between the two finger pictures (Current and past edge) is far then the mouse cursor moves quick on the other hand, assuming that the opening is close, the cursor moves slow.

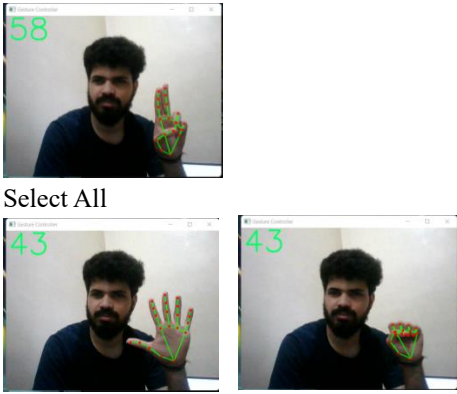
Left Click.



Right Click



Double Tap.



6. PERFORMANCE ANALYSIS

6.1. Virtual mouse execution examination

In the constant evaluation, ten subjects made different gestural capacities to figure out the exposure accuracy. The information gathered tells about the different size of screen Performa to exhibit the proposed model is more reasonable with standard applications, rather than including a fit goal as introduced. The information shows for the accompanying four goal as follow: 1240 *1220 with around 250 cases followed by 1500 × 1300 150 cases, 1540 × 1250 300 cases, in end 1750 × 1250 80 cases. It's generally expected that X is how much fingertips displayed.

Table 1 Experimental results

Functions	Fingertip Counting	Success	Failure	Accuracy%
Mouse movement	1	100	0	100
Left-Click	2	98	2	98
Right-Click	3	93	7	91
	4	91	9	
No Action	5	99	1.5	98.75
	0	99	2	
		572	20.5	97%

6.2. Performance of multiple people tracking

In : Precision=TPTpFP ... (3)

For the most part, this measurement looks at every one of the anticipated fingertip locations with every one of the physically named ground truth for a given profundity picture input. The typical exactness of each gathering was determined and is displayed. For the gathering of two, the fingertip-location exactness was most noteworthy, with 95%. The most pessimistic scenario was the gathering of six with a precision of 56.35%. The exactness of the three, four, and five gatherings were 88.78%, 77.03%, and 64.38%, individually. The outcomes show that the precision diminishes as the quantity of individuals in the gathering increments.

7. COMPARISON WITH APPROACHES

It researches the virtual-mouse creating and summed up the association in a layout. The primer results veered from past methodologies including developments for virtual mouse, including the various circumstances, for example, picture foundation, following various goals, individual distinctive verification.

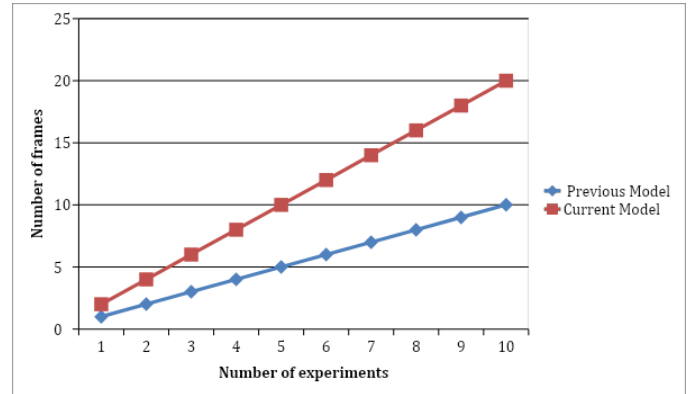


Fig 7: Frame rate vs experiments

In view of chart, The graph in fig 7, is the frame rate to distance experiments and the red line shows the previous approach and in blue shows the current method and approach. It can be observed that now blue reading has more number of frames as of around 55-60 from 30. Hindrances. Other than that, dealing with an assortment of goals is likewise remembered for these two frameworks, while the excess frameworks just work on a proper goal. Specifically, the precision is improved 3% from the past framework. This is a significant reason for an ongoing frame framework later on.

8. RESULTS AND DISCUSSION

Virtual Mouse module in the construction mouses limits by concentrate close by and counting the finger numbers. It can fill six astounding position: left click, right snap, left development, right new development, up progress and down progress. It considered a general lighting and room condition which was utilized in virtual control neighborhood. The distance between the camera and thing is most essential 10 m and the articles are set in a fixed environment [11].

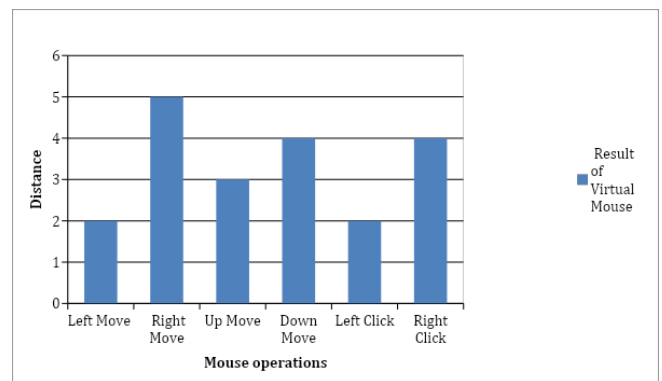


Fig. 8: Result of virtual mouse

- Number of Mouse open: 7
 - Limit test: 6
 - Unendingly out number of tests: 35
 - Attempted by: 46 years of an age stroke patient who has, considering everything, no control of his left side
- Figure 5 shows the times all of the seven mouse limits work completely. To review framework execution, the precision of the advancement was evaluated utilizing the condition [12].

Precision = $DF/TF \times 100$ percent ... (4)

Where DF is how much truly seen attempts and TF is the number of ardent undertakings. The precision of the structure using condition (4) is 82.17%.

Since the framework utilizes webcam-got accounts, the introduction of the arrangement may depend upon light. Additionally, getting through that there are other disguised objects present inside the foundation, the design could convey an inaccurate reaction. In any case, this issue might be limited by getting sorted out the edge values and other gadget limits, it's right now sensible that the working foundation ought to be light and there shouldn't be any enormously covered odds and ends present inside the foundation.

Furthermore, on some low figuring PCs, the contraption could run essentially more free since it plays out an outsized number of tangled evaluations in a magnificently short time period frame. In any case, for ideal headway execution, and standard PC or PC has the computational power required. One more perspective is that the contraption will run apathetic, persevering through the camera's objective is just incredibly high. This issue is reliably overseen by decreasing picture targets.

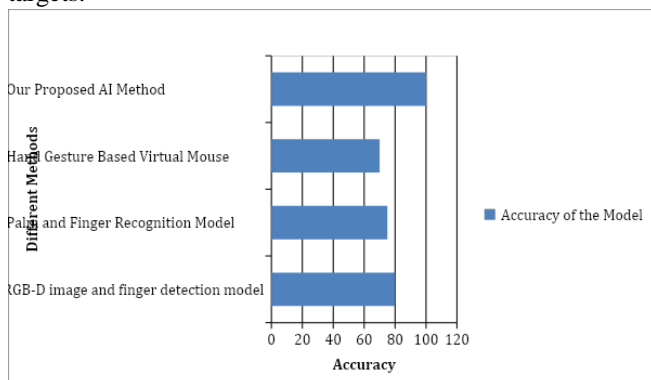


Fig. 9: Graph defining accuracy of the model

9. CONCLUSION & FUTURE WORK

This paper presented another virtual-mouse structure using RGB-D pictures and fingertip affirmation. The tip of client's finger improvement cooperated with the PC before a camera with no mouse gadget. The perspective showed astoundingly precise improvement measures, yet comparably sensible applications.

The proposed system overcomes the goals of 30 bundling rate virtual-mouse structures. It adds many advantages, e.g., working adequately in changing light levels or with complex backgrounds, unmistakable fingertip following at an even more huge distance, and fingertip ID at the speed of 60 bundling rate. The groundwork results showed technique for advancement brilliantly.

Mouse shapes a huge piece of the framework. The framework orchestrating can work with made by PCs for weakened individuals. To empower a virtual improvement where people can visit with the PC without using any certified mouse. This can facilitate Human-Computer Interaction, wherein physical contact with the computer would not be important. It uses article transparency and picture supervising in Open CV for the execution of the work, which resulted in incredible progress and fittingly, the task of the mouse is achieved with

astounding accuracy. This system is crucial for empowering individuals who have no control over their physical movements.

A titanic piece of the applications requires extra stuff, which is for the most part extreme. The assumption for this work is to gain this ground as monetarily as could genuinely be expected and to shape it under, anyway, programming pack yet. While the system can serve as an alternative to a physical mouse, its performance may be less effective in low-light conditions. This is, as a rule, for additional evaluation. Additionally, the work is free for a decent kind of climate and can be had a go at utilizing the fragile existing models [13] [14] [15] [16] [17][18] [19] [20].

REFERENCES

- [1] Abedin, M.Z., Nath, A.C., Dhar, P., Deb, K., Hossain, M.S.: License plate recognition system based on contour properties and deep learning model. In: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). pp. 590–593. IEEE 2018.
- [2] Adajania, Gosalia, Kanade, Mehta, Shekokar, N.: Virtual keyboard using shadow analysis. In: 2010 3rd International Conference on Emerging Trends in Engineering and Technology. pp. 163–165. IEEE 2019.
- [3] Ahmed, T.U., Hossain, S., Hossain, M.S., ul Islam, R., Andersson, K.: Facial expression recognition using a convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (IVR). pp. 336–341. IEEE 2019.
- [4] Asad, M.U., Mustafa, R., Hossain, M.S.: An efficient strategy for face clustering use in a video surveillance system. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (IVR). pp. 12–17. IEEE 2019.
- [5] S. Shetty: Bangla handwritten character recognition using a convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (IVR). pp. 318–323. IEEE 2019.
- [6] Gupta, D., Hossain, E., Hossain, M.S., Andersson, K., Hossain, S.: A digital personal assistant using Bangla voice command recognition and face detection. In: 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON). pp. 116–121. IEEE (2019) An Interactive Computer System with Gesture-Based Mouse and Keyboard 13
- [7] Haria, A., Subramanian, A., Asokkumar, N., Poddar, S., Nayak, J.S. Hand gesture recognition for human-computer interaction. *Procedia computer science* 115, 367– 374 2017.
- [8] Islam, M.Z., Hossain, M.S., ul Islam, R., Andersson, K.: Static hand gesture recognition using a convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (IVR). pp. 324–329. IEEE 2019.
- [9] Noor, Siddiquee, E.A., Sarma, D., Nandi, A., Akhter, S., Hossain, S., Andersson, K., Hossain, M.S.: Performance analysis of a surveillance system to detect and track vehicles using haar cascaded classifiers and optical flow method. In: 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA). pp. 258–263. IEEE 2017 .
- [10] Shibly, K.H., Dey, S.K., Islam, M.A., Showrav, S.I.: Design and development of hand gesture-based virtual mouse. In: 2019 1st International Conference on Advances in Science, Engineering, and Robotics Technology (ICASERT). pp. 1–5. IEEE 2019.
- [11] Vasant, P., Zelinka, I., Weber, G.W.: Intelligent computing & optimization, *Proceedings of the 2nd International Conference on Intelligent Computing and Optimization 2019*, vol. 1072. Springer 2019.
- [12] Vasant, P., Zelinka, I., Weber, G.W.: Intelligent computing & optimization, *Conference proceedings on Intelligent Computing and Optimization 2018*, vol. 866. Springer 2018.

- [13] Islam, R.U., Hossain, M.S., Andersson, K.: A novel anomaly detection algorithm for sensor data under uncertainty. *Soft Computing* 22(5), 1623–1639,2018.
- [14] Noor, K., Siddiquee, E.A., Sarma, D., Nandi, A., Akhter, S., Hossain, S., Andersson, K., Hossain, M.S.: Performance analysis of a surveillance system to detect and track vehicles using haar cascaded classifiers and optical flow method. In: 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA). pp. 258–263. IEEE2017.
- [15] Abedin, M.Z., Nath, A.C., Dhar, P., Deb, K., Hossain, M.S.: License plate recognition system based on contour properties and deep learning model. In: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). pp. 590–593. IEEE 2017.
- [16] Ahmed, T.U., Hossain, S., Hossain, M.S., ul Islam, R., Andersson, K.: Facial expression recognition using convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 336–341. IEEE 2019.
- [17] Chowdhury, R.R., Hossain, M.S., ul Islam, R., Andersson, K., Hossain,S.: Bangla handwritten character recognition using convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 318–323. IEEE 2019.
- [18] Gupta, D., Hossain, E., Hossain, M.S., Andersson, K., Hossain, S.: A digital personal assistant using bangla voice command recognition and face detection. In: 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON). pp. 116–121. IEEE 2019.
- [19] Asad, M.U., Mustafa, R., Hossain, M.S.: An efficient strategy for face clustering use in video surveillance system. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 12–17. IEEE 2019.
- [20] Uddin Ahmed, T., Jamil, M.N., Hossain, M.S., Andersson, K., Hossain,M.S.: An integrated real-time deep learning and belief rule base intelligent system to assess facial expression under uncertainty. In: 9th International Conference on Informatics, Electronics & Vision (ICIEV). IEEE Computer Society ,2020.

Blockchain based AI Yoga Assistant

Tushar Ojha¹, Vaibhav Krishan Yadav², Anshika Gupta³, Koyel Datta Gupta⁴

1,2,3,4Computer Science & Engineering

Maharaja Surajmal Institute of Technology, New Delhi, India

¹tusharojha2001@gmail.com, ²vaibhavkrishanyadav@gmail.com, ³anshikagupta2001@gmail.com, ⁴koyel.dg@msit.in

Abstract—This paper explores the potential benefits of creating a blockchain-based yoga assistant platform, which would use AI for personalized guidance and feedback. It will incentivize people on-chain with Soul Bound Tokens as rewards. It would employ a poseNet model and NLP/ML algorithms, and be constructed using Flutter to go cross-platform i.e. both Android and iOS from the same codebase. It evaluates the current tools and technologies, along with the possibilities and challenges to implementation.

1. INTRODUCTION

The integration of AI and blockchain technologies has proven to be highly beneficial in various areas of life. From marketing and logistics to healthcare and education, they are making the world smarter, safer, and more efficient. In this paper, we explore the use of these technologies in the development of a personalized yoga assistant platform.

Yoga is an ancient practice with a variety of benefits for physical and mental health, such as increased flexibility and improved concentration. Yoga is practiced in different modes and manners across the world, a few of them include:

1.1. Reading Books

Reading books on yoga is an effective way to learn and become knowledgeable about the practice, as books provide insight into the philosophies behind yoga, help to understand advanced poses, explore various types of yoga and understand how to incorporate yoga into everyday life. Reading books can provide valuable guidance on modifications and other helpful insights.

1.2. Watching Tutorials

Watching yoga tutorials is another great way to absorb and understand the practice of yoga. Watching tutorials can help to get familiar with the poses, breathing techniques, and postures. Moreover, tutorials can be beneficial in deepening one's practice and gaining a better understanding of proper alignment and adjustments.

1.3. Offline Classes

Attending offline classes is important too and can be done in addition to the above methodologies. Offline classes allow one to practice yoga in a more authentic setting, with a teacher and other practitioners. This offers an entirely different experience, one cannot get from merely reading books or watching videos. Attending offline classes also provides an opportunity to build personal relationships and receive feedback and guidance from the teacher or other practitioners. But there are many problems associated with the current solutions:

- Many of the problems associated with reading books on yoga involve learning something inaccurately.

Books can contain a variety of different opinions and facts and it can be difficult to discern which is which. Moreover, they do not provide feedback or adjustments, which can be necessary to advance one's practice.

- With watching tutorials, the downside can be that without a teacher or experienced practitioner one's learning can be limited. Additionally, postures can be difficult to execute correctly and without proper guidance, one can be doing poses wrong.
- Offline classes can be expensive and can pose a variety of other issues. Attending classes during certain times, especially with a busy lifestyle, can seem tedious or exhausting. Additionally, with some classes, one may need to find a suitable teacher before committing to any class, as this can make a huge difference in the quality of the classes.

This paper examines how AI and blockchain continuously work together to create a secure platform for personalized yoga sessions. We discuss why this platform is necessary and the potential benefits and challenges that come with it.

The paper begins by introducing both technologies and a subsequent overview of their technologies and research done so far. We then explore the nature of the platform and the potential benefits of combining AI and blockchain. We provide an analysis of common yoga technologies and how they can be improved by integrating AI and blockchain. We also discuss how the platform can be used to secure user data and ensure privacy. We finally provide an implementation plan and discuss the potential impact of the platform on yoga practitioners and related industries.

In conclusion, this research paper explores the use of AI and Blockchain technologies in the development of a personalized yoga assistant platform. This platform can help yoga practitioners remain committed to their practice while improving their privacy, and security. It will also create new opportunities for businesses and users by providing rewards and monetizing their commitment to staying healthy.

2. LITERATURE REVIEW

2.1. Yoga: it's origin, history and development

In Basavaraddi IV's paper "Yoga: Its Origin, History and Development," [?] the author reviews various aspects of yoga culture and its evolution over the centuries. He examines ancient texts, traditional practices, scriptures, and archaeological evidence to analyze the history of yoga and its implications for culture and lifestyle. He also investigates the

various techniques and locations where one can practice yoga today, such as in ashrams, temples, gyms, schools, and homes. He examines the various types of yoga including Hatha, Bhakti, and Raja, and looks into the social context of these practices. He examines the usage of modern technology and media to propagate yoga and the potential ethical and moral considerations associated with the commercialization of yoga. The paper highlights the current wellness industry, the increasing global popularity of yoga, and the potential health benefits of the practice to humankind.



Fig. 1. Yoga in Indus Valley Civilization

2.2. Preventing Yoga Injuries

In their paper "Understanding and Preventing Yoga Injuries," [1] Fishman, Saltonstall and Genis analyze the potential risks of yoga practice and how to prevent those injuries. They discuss how incorrect form of postures, the failure to exercise physical precautions, and the lack of understanding of the physical implications of yoga can lead to injury. They present guidelines for safe yoga posture practices, suggest modifications for people with physical limitations, and discuss the importance of professional teacher guidance. They examine the literature on yoga-related injuries, analyze research on yoga injury prevention, and review current methods that can be employed to reduce yoga-related injury rates. They refer to the current cultural and economic landscape of the wellness industry and suggest certain safety precautions that should be taken. They conclude by advocating the need for effective and accessible yoga injury prevention measures and developing professionally qualified yoga teachers to meet this need.

2.3. Using AI for pose detection

There have been multiple attempts to create a pose detection model to use visual data from the device camera in form of raw video and convert it to an array of images for further processing to get the points of different body parts and then connect them to find the body posture.

The paper "Deep Learning based Human Pose Estimation using OpenCV" [2] introduces a simple method for human pose estimation using OpenCV libraries. The authors review existing research on the topic of human pose estimation and identify gaps and inconsistencies in the literature. They explain the use of OpenCV library in the form of raw video and converting it to an array of images for further processing to detect points of different body parts and then connecting them to find body posture. In their review, they discuss the challenges faced when using OpenCV and explore various approaches to address these challenges. They also evaluate existing research and present an initial model to be used for

further evaluation and testing. For future work, they propose to improve on the accuracy of their model by exploring and manipulating machine learning algorithms, combining different technologies and increasing the speed of the calculation process.

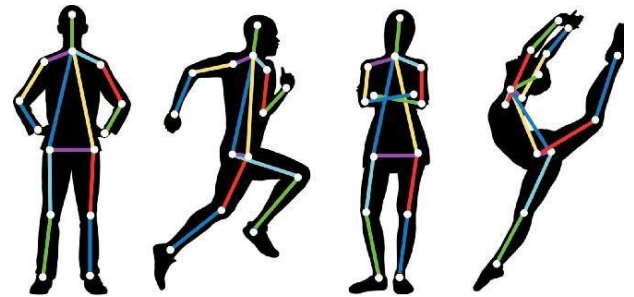


Fig. 2. Pose Detection Models

2.4. Creating Decentralized Identity for People

In the paper "Decentralized Society: Finding Web3's Soul," [?] the authors explore the potential of blockchain technology and decentralized applications to creating a more equitable and fair online society. They look at the implications of creating a digital space protected by blockchain where users can exchange value without the need for third-party trust. They analyze how decentralization of power could potentially enable new business models and lead to a fairer economy. They discuss how the current centralized network of power and wealth concentration has created inequality and how the development of a decentralized Web3 could help to address these issues. They also discuss the potential security risks and challenges posed by decentralized networks and suggest several best practices to mitigate those risks. The authors then review existing decentralized projects that are paving the way for the development of an open-source platform for a fairer society. Finally, they provide further recommendations for the development of Web3 to ensure a future of trust and freedom for all.

3. SCOPE OF THE WORK

From the research in the past to now, there have been a lot of changes in the market condition and user behavior. Mobiles have become very accessible and people prefer to stay at home more as compared to going out in traffic and pollution for doing yoga. The following pie chart explains it:

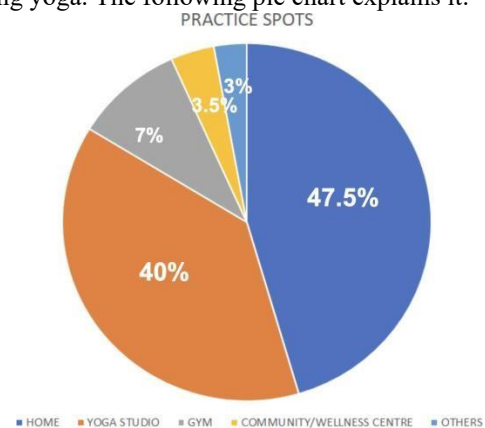


Fig. 3. Yoga practise spots as of 2022

The projects built in past involves only detecting postures and there's a lot of space to incorporate the AI models in the mobile phone so that can become much more accessible to people.

Blockchain technology can provide a very secure, private method of storing information, and incentive the process of doing yoga at home. The decentralized nature of blockchain ensures that no single entity has control over the data, making it highly resistant to tampering or hacking. This is because the data is stored across a network of computers, rather than in a central location. This means that in order for the data to be altered, a hacker would have to gain access to every single node in the network, making it almost impossible to do so.

Blockchain also utilizes a technique called "hashing" which converts the data into a unique code that is then stored on the blockchain. This makes it almost impossible to trace the data back to the original source, ensuring that the data remains private and secure. Additionally, blockchain technology utilizes advanced encryption methods to protect the data from being intercepted or stolen.

Furthermore, the use of smart contracts in blockchain technology allows for the creation of transparent and immutable records, making it highly suited for storing information related to Non-fungible Tokens called NFTs. This ensures that the ownership and authenticity of the NFT and can be easily verified and tracked, making it a secure and private method for storing NFT information. We in our case build the identity for the user's yoga daily steaks using Soul Bound Tokens (SBTs). These are non-transferable NFTs.

Hence, there's a lot of space for building a solution that is robust and provides a true sense of doing yoga with an incentive program. And this is a growing space with companies investing more and more in awareness programs.

4. MATERIALS AND METHODS

The complete system has 3 major components, as described below:

- Machine Learning Model for Pose Detection
- Mobile Application to engage with ML Model
- Blockchain-based identity and rewards for maintaining steaks

Taking a deep look into each of the aspects to understand the potential setup and integrations:

4.1. Pose Detection Machine Learning Model

For pose detection, there are multiple predefined models available that can be trained on a new dataset and used using transfer learning. The following list shows some popular models:

- Faster R-CNN
- Single-shot Detector (SSD)
- Region-based Convolutional Neural Networks (R-CNN)
- YOLO
- Posenet by Google

Posenet model seems to be the most promising one as It is a machine learning model developed by Google AI which leverages convolutional neural networks to detect human

poses in images and videos. This model is based on a powerful convolutional neural network which can detect and identify persons in an image or video in real time. Posenet can detect a person's body skeleton, which it then uses to detect the position, movement, and other aspects of a person's pose. The Posenet model can be used for applications such as pose estimation for sports activities, human-computer interaction, gesture recognition, and health monitoring.

For training of the model, you will need at least 5000 images of each yoga pose with a 90:10 split ratio of train and test data to reach minimum 94 percent accuracy.

Since it is working on a neural network it is always better to have more data and computation power for training, therefore one can use Amazon Web Services, Google Cloud, Microsoft Azure, or other cloud service providers.

You can also use Google teachable machine user interface to test the model in a sandbox. It is not recommended to use that in the production app.

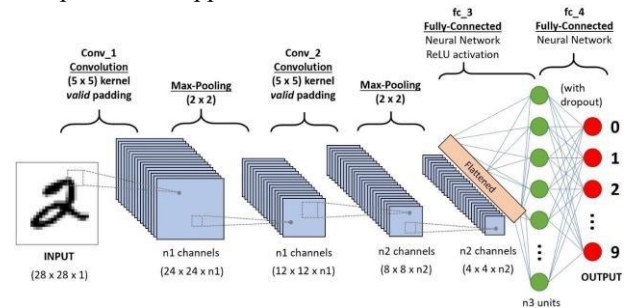


Fig. 4. Convolutional Neural Networks

4.2. Mobile Application using Flutter

For building mobile applications there are multiple frameworks and technologies available like React Native, Flutter, Ionic, Java / Kotlin, Objective C / Swift, etc.

React Native and Flutter are examples of Hybrid mobile application development technologies which means you can ship apps for both Android and iOS from the same codebase. Therefore, it makes sense to prefer a hybrid technology over a native to scale fast.

The clear choice here is Flutter over React Native because React Native uses a javascript engine for code compilation that makes the execution and the model processing significantly slow irrespective of the device's capabilities.

Flutter, on the other hand, translates the whole project code into machine code for faster execution. Also, Flutter and Tensorflow (libraries used for Machine learning) are both from Google.

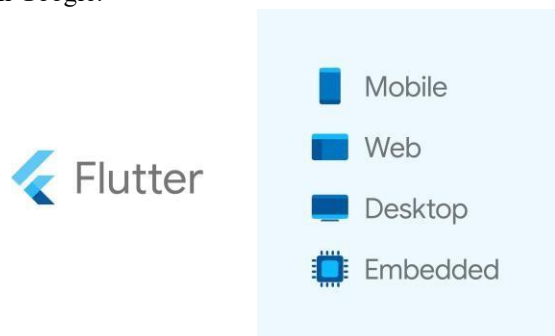


Fig. 5. Flutter: Cros Platform Stack

There are multiple libraries available to interact with the machine learning models and blockchain like:

- TensorFlow Lite Flutter plugin provides a flexible and fast solution for accessing TensorFlow Lite interpreter and performing inference. The API is similar to the TFLite Java and Swift APIs. It directly binds to TFLite C API making it efficient (low-latency). Offers acceleration support using NNAPI, GPU delegates on Android, Metal and CoreML delegates on iOS, and XNNPack delegate on Desktop platforms.
- Web3 Ethereum, Etherjs and Wallet Connect wrapper for Flutter Web. Made especially for developing Dapp.
- Camera permissions are also required from the user

The machine learning model can only detect the pose, but connecting different body poses in order to eventually make sure the user do a correct pose requires additional logic.

And to achieve this, we can create a course module that checks for different poses as a part of yoga process and shares feedback to user from a set of predefined prompts by converting text-to-speech on the fly.

This way we can create a Yoga tutor that checks for correct poses and instructors if you do something wrong.

4.3. Blockchain for Identity

Using Blockchain Technology for generating proofs and rewards seems a good idea as per the introduction of Soul Bound tokens that make the properties attached to a user lifetime rather than being a simple number stored on a centralized server. The whole idea of implementing blockchain in this product is as follows:

- Soulbound Tokens (SBTs) are digital identity tokens that represent the traits, features, and achievements that make up a person or entity. SBTs are issued by “Souls,” which represent blockchain accounts or wallets, and cannot be transferred.
- These tokens can be used as proof of doing Yoga and users will be able to claim rewards on the basis of these achievements. As the values or count of these collectibles increases the user will enable them to explore sections of the gift stores or the company that users are working with can provide more incentives to the most active person.

Reward funnels have an admin that provides rewards at different levels of accomplishments by the list of participants. This user holds the ability to add and maintain the individuals in the funnel hence it is called as ADMIN. The following smart contract code depicts the idea of identity management by the admin and the values the admin can access. An admin is one who mints this contract and holds the ability to add people to this reward funnel. The admin has to know the address of the account they want to add but this can be handled with a frontend like mobile app where it maps users directly from their ID without the friction of fetching cryptographic address every time.

Individual Identity and Score contains the record of users’ performance, achieved milestones, and analytics. These are also kept anonymous to protect users’ identity on-chain. Users

can easily log in from the mobile app that maps to their identity and it always stays connected to them while being 100 percent anonymous yet acceptable.

This Identity can be implemented in the mobile application using social authentication and using solutions like Biconomy or other smart contract wallet abstraction layers.

5. RESULTS AND DISCUSSIONS

After the development process, and training the dataset for a few yoga poses we managed to finish the project’s minimal viable version.

This version supports practicing yoga, in front of AI and shares feedback and then rewards on-chain.

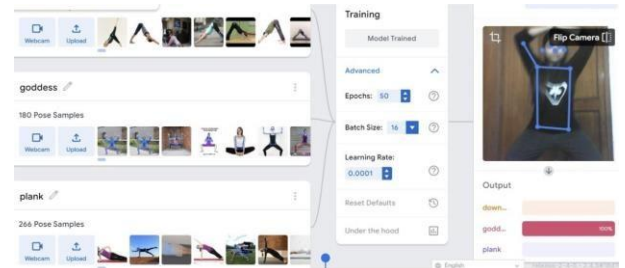


Fig. 6. Yoga AI: Training

Users can download the app, and then give it the necessary permissions to start doing yoga like this:

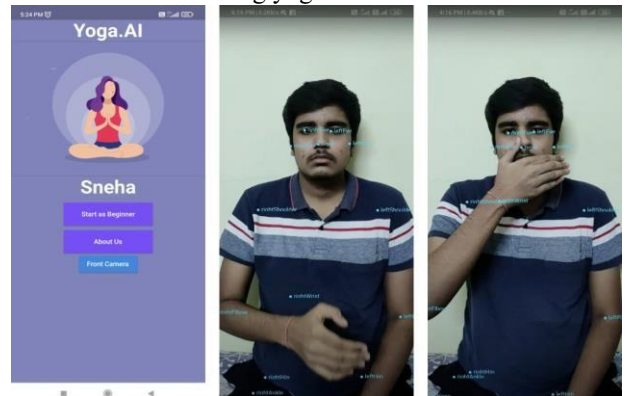


Fig. 7. Yoga AI: Live Exercise

The following smart contract governs the reward system for people doing yoga on the Polygon Mumbai Testnet:

```

11  enum ExperienceLevel {
12      beginner,
13      intermediate,
14      experienced
15  }
16
17  struct Participant {
18      address id;
19      ExperienceLevel level;
20      bool isRegistered;
21  }
22
23  contract RewardFunnel {
24
25      address private admin;
26      mapping(address => Participant) participants;
27
28      constructor(){
29          admin = msg.sender;
30      }
31
32
33      function addParticipant(address newMember, ExperienceLevel level) public payable {
34          require(msg.sender == admin, "Only admin can add members.");
35          if(participants[newMember].isRegistered == false){
36              Participant memory participant = Participant(
37                  newMember,
38                  level,
39                  true
40              );
41              participants[newMember] = participant;
42          }
43      }
    
```

Fig. 8. Yoga AI: Smart Contract

6. CONCLUSION

In conclusion, the proposed blockchain-based yoga assistant platform has immense potential to revolutionize the

traditional practice of yoga. It could enable users to track their progress, receive personalized guidance and feedback, and even detect and correct incorrect poses. The application of blockchain technology, Posenet, and artificial intelligence would allow this platform to provide users with an enhanced yoga experience, allowing them to practice yoga more effectively in an affordable and efficient manner. As such, the proposed platform could make yoga more accessible to the general public and bring more people into the practice, thus furthering its reach and popularity. This could lead to an increased level of physical and mental health for many individuals and help create a healthier, more balanced population.

7. ACKNOWLEDGMENT

We express our deep gratitude to Dr. Koyel Dutta Gupta, Department of Computer Science and Engineering for her valuable guidance and suggestion throughout our project work. We are thankful to Dr. Naresh Kumar and Dr Divya, for their valuable guidance. We would like to extend our sincere thanks to Head of the Department, Dr. Rinky Dwivedi for her time- to-time suggestions to complete our project work. We are also thankful to Col (Dr.) Ranjit Singh, Director for providing us with the facilities to carry out our project work.

REFERENCES

- [1] Basavaraddi IV (2015) Yoga: its origin, history, and development. Government of India, Ministry of External Affairs
- [2] Sherman KJ, Cherkin DC, Erro J, Miglioretti DL, Deyo RA (2005) Comparing yoga, exercise, and a self-care book for chronic low back pain: a randomized, controlled trial.
- [3] Ann Intern Med 143:849–856.
- [4] Fishman L, Saltonstall E, Genis S (2009) Understanding and preventing yoga injuries. Intern J Yoga Therapy 19:47–53
- [5] "Deep Learning based Human Pose Estimation using OpenCV " 29 May. 2018
- [6] "Semantic Representation and Recognition of Continued and" 12 Nov. 2008
- [7] "Human Movement Detection and Identification Using ... - NCBI." 5 May. 2014
- [8] Decentralized Society: Finding Web3's Soul Google teachable machine for pose training

Blockchain Technology-based Messenger

Dr. Sonia Rathee¹, Aditi Diksha², Aman Gupta³, Ankur Badola⁴

¹Assistant Professor, Department of Computer Science,

^{2,3,4}Student, Maharaja Surajmal Institute of Technology, Delhi, India

Maharaja Surajmal Institute of Technology, Delhi, India

¹Sonia.rathee@msit.in

Abstract--- Blockchain technology is considered a significant development after the invention of the Internet. Blockchain technology records transactions in a database that is encrypted and distributed across multiple computer networks as a digital ledger of online transactions. As digital communication continues to dominate our daily lives, the need for secure and private messaging platforms becomes increasingly crucial. Traditional messaging applications often face challenges regarding data breaches, privacy concerns, and centralised control over user information. This research paper proposes the development of a blockchain-based messenger application as a solution to these issues. The research focuses on exploring the key components and functionalities required for building a blockchain messenger application. It examines the advantages of utilizing blockchain technology, such as immutability, transparency. The study also addresses the challenges associated with implementing blockchain technology in a messaging application, like scalability, latency.

Keywords: Blockchain Technology, Smart contracts, Distributed Systems, Decentralized, secure, chat, messaging, security.

1. INTRODUCTION

Nowadays, chatting on messaging platforms has become a regular part of people's lives. Most of our communication happens through social media platforms, where we can also share pictures, videos, and other media. However, these platforms rely on centralised servers to process and transmit our messages. This means that all our messages, even if they claim to be encrypted, can be seen by the servers handling them. To illustrate this issue, we created a simple chat application and conducted experiments to show how the messages can be visible at the processing servers. Moreover, centralised systems face challenges in terms of their ability to handle increasing demands and grow efficiently, in contrast to decentralised computing systems.

As we rely more and more on digital communication in our daily routines, it has become essential to have messaging platforms that priorities security and privacy. Traditional messaging apps often struggle with issues such as data breaches, privacy concerns, and centralised control over user information. For instance, let's consider Instagram chat as an example. In Instagram, all the data is stored on a single central server. If something happens to that server, like it gets damaged or destroyed, there is a risk of losing user data. Additionally, there is also the possibility of user information being leaked from the server.

To address these challenges, our project utilizes a decentralized application approach, also known as dApps. In our application, all user data is stored in blocks that are

connected, forming a chain-like structure. Unlike centralised systems, a decentralised application doesn't rely on a single central server. Instead, it operates on a peer-to-peer network where multiple participants interact directly with each other. The data stored in each block is highly secure due to the implementation of robust encryption and hashing functions, using 256-bit encryption keys. This makes it almost impossible for anyone to view the data without proper authorization. Moreover, if a hacker attempts to modify the information in a block, they will need to make changes to all copies of that block across the entire blockchain network, which is an extremely challenging task. In a decentralised application, there are many nodes that are linked together in a network resembling a mesh. These nodes communicate with each other directly, forming a peer-to-peer system. The blockchain, on the other hand, is a chain of blocks that contains a comprehensive record of transactions, like a traditional public ledger.

2. LITERATURE REVIEW

Satoshi Nakamoto announced in 2008, the distribution of all digital currency transactions based on distributed systems and cryptography, the first time blockchain technology was used. Blockchain Shows Potential for Peer-to-Peer Payments in an Insecure Environment [1].

In recent years, decentralization has become a mechanism that offers real solutions to emerging problems. Since it became the most popular technology, most messaging apps have started using it as well. There are few apps like DUST, SENSE.CHAT, E-CHAT. Most communication applications such as SYLO and Microsoft teams used in organizations such as SYLO, Skype for Business and the WhatsApp business API have not yet been implemented using this technology.

DUST - Decentralised based messaging application using decentralised (Public Ethereum). The app has many features that focus on privacy and security. The main feature of this app is timed message deletion, your messages will self-destruct after a certain time.

E-Chat - This app is more than a simple messenger built with Ethereum Distributed- E-chat can manage Ethereum transactions and is equipped with many cryptocurrency wallets. Cryptocurrency could be computerized or virtual currencies that are created and stored electronically on the blockchain. Encryption technology (cryptography) is used to create accounts, track, and verify account transactions.

Table 1. Requirements for basic blockchain technology-based messenger on the users

User	<ul style="list-style-type: none"> • Users should be able to send messages to his/her friends. • Users should be able to add friends before sending the messages. • User should be able to block his friend and delete the chat temporarily till the friend is blocked. • User should be able to view the list of his friends and their conversation. • Users should be able to search for a chat. • User should be able to view the list of all the other users.
-------------	---

Ethereum Distributed- E-chat can manage Ethereum transactions and is equipped with many cryptocurrency wallets.

Cryptocurrency could be computerized or virtual currencies that are created and stored electronically on the blockchain. Encryption technology (cryptography) is used to create accounts, track, and verify account transactions. It is not controlled or controlled by any bank or government and is therefore decentralised. Popular cryptocurrencies are Bitcoin, Ethereum, Litecoin, Dash and Peercoin. Cryptocurrency transactions are low, secure, and risk-free for merchants.

Cryptocurrency transactions are recorded in a public digital ledger called the blockchain or distributed ledger. Every file or record in the blockchain is called a block. Each user of a transaction on the network creates a block that is verified by all users of the network and then added to the blockchain [5][6]. Once verified and added, blocks cannot be modified. Blockchain wallets are digital wallets where users can store cryptocurrencies.

Use the unique wallet ID assigned to their wallet. A wallet has an address called a public key and a private key called a secret. The wallet generates a pair of public and private keys to ensure secure transactions [2]. The distributed ledger system helps the Blockchain to carry out its tasks even if a node is broken down. This increases trust in the system. Intermediaries are eliminated in Blockchain architecture and hence data security is fostered.

3. DESIGN AND OVERVIEW

Our main goal was to allow use of cryptocurrencies and blockchain technology to provide a decentralized messenger dapp which is an attempt to provide secure, reliable, and decentralized messaging experience for the users. The users of the application are divided into ‘sender’ who wish to send the message and ‘receiver’ who receives the message. Receiver can also send the message back to the sender.

The application is designed to have a back-end consisting of smart contracts hosted on Ethereum public blockchain, and front-end of application using HTML, CSS, JS, Reacts. The frontend web application will use web3.js library to connect to and interact with the smart contracts. We created Ethereum smart contracts with the solidity programming language. The main features we included were - user listings for friends, user listing for all users, chatting window to chat with friends, blocking them, searching the friend and gas payment using cryptocurrency, use of blockchain technology to provide decentralised systems. It runs on the blockchain. Once someone sends a message, the message is sent securely. This is because this application is powered by a smart contract on the blockchain that governs the application. It tracks the friends and their messages.

Any time a person sends a message in the messenger, they’ll send Ethereum cryptocurrency for the gas through the smart contract. For receiving messages, no gas is required, they’ll receive the messages directly. Messages and friends are stored on the blockchain secured and encrypted. This is illustrated by Fig1.

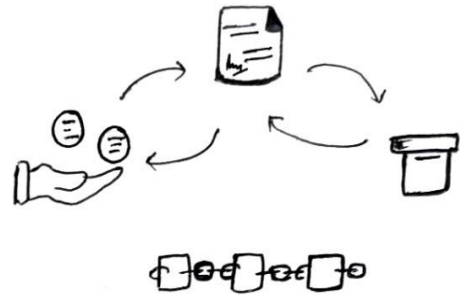


Fig1. Working on smart contracts for our application.

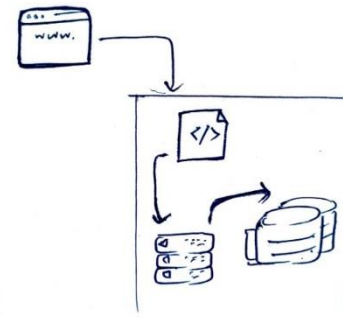


Fig2. Normal Web Application.

Usually with a web application, a web page is accessed using a web browser. All the code for the application lives on a central web server and interacts with the backend of the application which interacts with the database as shown in fig 2. Our application presented will directly connect and interact to blockchain which hosts all our code and data for our decentralised marketplace as shown in fig 3.

4. DISCUSSION

4.1 Examine the role of blockchain technology in providing security, privacy, and transparency in a messaging application:

Blockchain has been disrupting multiple industries. It is transforming the way businesses carry out their day-to-day activities using different use cases. Rapid expansion of the online messaging services many giants have emerged based on centralised data and supremacy.

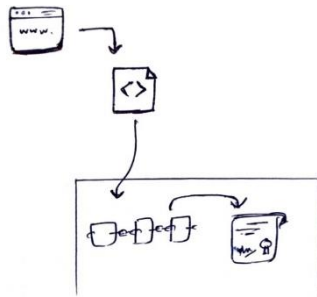


Fig3. Blockchain based Web Application

The blockchain technology has the potential to break monopolies also. One of the main problems and concerns among people within these services is transparency, however, this is not the case with blockchain technology-based messaging services. There is no central authority in blockchain, it enables transparency and trust. Blockchain enhances security by creating a record that is extremely difficult to alter or hack. The decentralised nature of blockchain makes it resistant to unauthorized changes. Additionally, encryption techniques protect message content, ensuring only intended recipients can access it. Storing data across multiple network nodes reduces the risk of data loss. Blockchain improves privacy by allowing users to operate under pseudonyms or digital identities, safeguarding their real-world information. Users have more control over their data, deciding what to share and with whom. End-to-end encryption ensures that only senders and recipients can read the messages, preventing intermediaries from accessing them. By using blockchain technology, messaging applications can enhance security, protect user privacy, and introduce transparency. This builds trust among users and creates a more secure messaging environment.

4.2 Identify the challenges faced in implementing blockchain technology in a messaging application and propose potential solutions:

When we use blockchain technology in messaging applications, there are certain challenges that need to be addressed. Let's discuss some of these challenges and propose simple solutions:

Handling many users and messages can slow down the blockchain network. To overcome this, we can divide the network into smaller parts called shards, which allows messages to be processed simultaneously, increasing speed and scalability. It's important to make the messaging application easy to use for everyone. Integrating wallets into the app in a simple and user-friendly way can help users interact with blockchain features without confusion or complications.

We can improve the efficiency of the blockchain by combining different methods of verifying transactions. This

hybrid approach can ensure security, scalability, and energy efficiency, striking a balance between these factors.

To overcome these challenges, it's important to continue researching and collaborating with the blockchain community. By implementing practical solutions, we can fully unlock the potential of blockchain technology in messaging applications, providing better security, privacy, and transparency.

4.3. Compare and contrast the features of decentralised messaging applications with centralised messaging applications:

Decentralised messaging apps and centralised messaging apps have some key differences.

Centralised Messaging:

- Users have less control over their data. The messaging app provider owns and controls the data.
- Privacy can be compromised as apps often collect user data, including personal information. Messages are stored on central servers, which can be vulnerable to hacking.
- Trust depends on the messaging app provider. The transparency of message history is limited to what the provider allows.
- Vulnerable to failures and censorship as messages are stored on central servers controlled by the app provider. The provider can moderate or delete messages.
- Designed for scalability and provides a user-friendly experience. Central servers can handle large volumes of messages efficiently.

Decentralised Messaging:

- Users have more control over their data and identities. They own their data and can decide how to share it.
- Offers better privacy as users can use fake names to protect their real identities. Messages are encrypted and stored in a way that reduces the risk of data breaches.
- Built on blockchain technology, these apps provide transparency through a public record. Anyone can check the history of messages, promoting trust and accountability.
- Distributed across many nodes, these apps are more resistant to failures and censorship attempts. Messages cannot be easily removed or censored.
- May have challenges with handling lots of users and messages due to the distributed network. Interacting with blockchain features can be more complex.

4.4 Evaluate the potential impact of decentralised messaging applications on the existing messaging application market:

Decentralised messaging apps may be used by specific groups or in situations where privacy, security, and transparency are extremely important. These apps can be popular among people like journalists, activists, or those living in countries with strict rules on what can be said or shared. Decentralised messaging apps give users the power to control their data,

including who can see it and when they can't. This change allows users to be more in charge of their privacy and how they communicate, which is something many users want. Decentralized messaging apps have many advantages, but they may encounter difficulties in getting people to use them and providing a smooth experience. It can be challenging for users to switch from familiar centralised platforms, and they might need time to get used to the new interfaces and features that come with blockchain technology.

5. RESULT

Code Snippets –

```
struct AllUserStruct {
    string name;
    address accountAddress;
}

struct user {
    string name;
    friend[] friendList;
}

struct message {
    address sender;
    uint256 timestamp;
    string msg;
}

struct friend {
    address pubkey;
    string name;
    bool blocked;
}
```

```
AllUserStruct[] getAllUsers;

mapping(address => user) userList;
mapping(bytes32 => message[]) allMessages;
```

Fig4. structures for common objects

Figure 4 shows the AllUserStruct structure, user structure, message structure, getAllUsers mapping and friend structure. 'AllUserStruct' is used to store a list of all users in the getAllUsers array. 'User' Represents a user of the ChatApp contract. 'Message' Represents a message sent between users. 'Friend' Represents a friend of a user. The 'getAllUsers' mapping, along with the 'AllUserStruct' struct, is used to maintain a list of all users registered in the ChatApp contract, facilitating operations such as retrieving user information and displaying a list of users. Function 'addFriend' allows a user to add a friend by providing the friend's Ethereum address and name. Performs various checks, including user existence, non-duplication, and not

adding oneself as a friend. Calls the internal function _addFriend to add the friend to the user's friend list. Function 'checkAlreadyFriends' is an internal function that checks if two users are already friends. It compares the friend lists of the two users to determine friendship.

Function 'blockFriend' allows a user to block a friend. It takes the friend's Ethereum address as input. It performs checks to ensure the sender and friend exist and are registered, and they are already friends. Function 'unblockFriend' allows a user to unblock a previously blocked friend. It performs checks to ensure the sender and friend exist and are registered, they are already friends, and the friend is blocked.

Function 'createAccount' allows a user to create an account by providing a name. It checks if the user already exists and if the name is not empty. Function 'getAllAppUser' returns an array of AllUserstruct structs containing information about all registered users of the ChatApp contract.

Function 'getUsername' takes an Ethereum address as input and returns the corresponding user's name. It checks if the user exists and throws an error if not. Function 'getMyFriendList' takes a user's Ethereum address as input and returns an array of their friends. Retrieves the friend list of the user from the userList mapping.

Function '_getFriendIndex' is an internal view function that finds the index of a friend in a user's friend list. Checks if both user and friend exist and are registered. Iterates over the user's friend list to find the index of the friend. Throws an error if the friend is not found. Function 'readMessage' allows a user to read messages exchanged with a specific friend.

Function 'sendMessage' allows a user to send a message to a friend by providing the friend's Ethereum address and the message content. It performs checks to ensure both the sender and the friend exist and are registered. It checks if the sender and friend are already friends.

If the sender and friend are not blocked by each other, it creates a new message struct with the sender's address, current timestamp, and message content. The new message is added to the allMessages mapping under the corresponding chat code. Function 'deleteConversation' allows a user to delete the entire conversation with a friend.

It takes the friend's Ethereum address as input. It performs checks to ensure the sender and friend exist and are registered, and they are already friends. Figure 10 shows a messaging UI screen for a sample given application. Here the user can type the message and click on send button to send the message to a friend's account. Left side panel provides the friend list of the user which is collapsible for a wide view of the chat screen.

Block/Unblock UI is also given above the chat window. Each message is shown with the timestamp below it. Figure 11 shows the UI to add friends to chat. Here, the user can get the account address of the friend and add it in the form and can give any name to the friend contact address which is shown in the chat screen and submit the details to add the friend to the user's account.

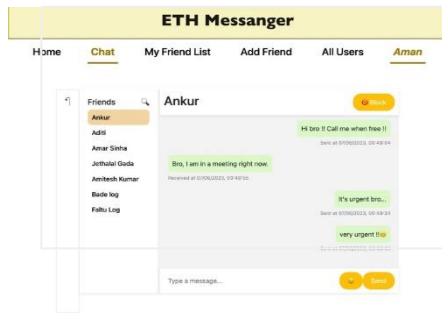


Fig10. Messaging Screen UI

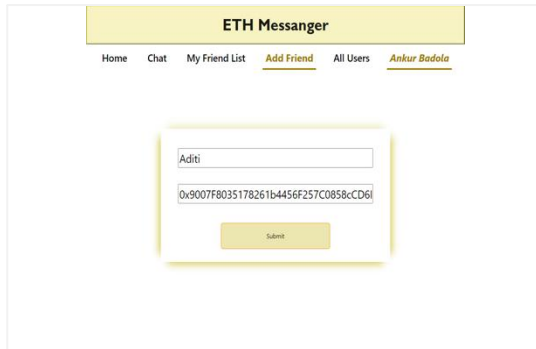


Fig11. Add Friend Screen

6. CONCLUSION

We have introduced and demonstrated a decentralised messaging application using blockchain technology, specifically built on the Ethereum Blockchain. Through an analysis of existing blockchain literature, we have explored the potential impact of blockchain technology on messaging applications. We have provided insights into the system architecture and workflow of a messaging DApp.

Using blockchain technology in messaging applications brings many benefits. It improves the security of online communication, benefiting both the people sending messages and those receiving them. This makes blockchain a promising solution for secure messaging transactions. While it might take some time for blockchain messaging apps to become widely adopted, the advantages they offer will attract users from all over the world. In the future, messaging apps will depend on blockchain technology to ensure secure and reliable transactions.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
 [2] Medium. 2020. "How Blockchain Technology Works", [online] Available at:
 [3] <<https://medium.com/@ipspecialist/how-blockchain-technology-works-e6109c033034>>
 [4] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang, "An Overview of Blockchain Technology:Architecture, Consensus, and Future Trends," IEEE 6th International Congress on Big Data, 2017.
 [5] G. Foroglou and A.-L. Tsilidou, "Further applications of the blockchain," 2015.

[6] Kaidong Wu, Yun Ma, Gang Huang, Xuanzhe Liu, "A First Look at Blockchain-based Decentralised Applications," eprint arXiv:1909.00939, 2019.
 [7] Anuradha Jayakody, Rajitha De Silva, Pirabanjan Kirupaharan, Sanka Tillakaratna, Chapa Vidanaarachchi and Sharon Shashangan, "'Air-Talk':An IPFS and Blockchain-based Private Decentralized FileSharing and Chat Application," IRJET, 2020.
 [8] Medium. 2022. "Blockchain chat server application written in Solidity", Available at:
 [9] <<https://itnext.io/blockchain-chat-server-application-written-in-solidity-5e929c0e4c9f>>
 [10] Abhishek P. Takale, Chaitanya V. Vaidya, Suresh S. Kolekar, "Decentralized Chat Application using Blockchain Technology," IJREAM, 2018.
 [11] Vikas Bhardwaj, Raja Ram Sharma, Akhilesh Kumar, "Chat Application using Blockchain Technology," IJREAM, 2018.
 [12] Keshav Khalkar, Nikhil Dhake, Sarwesh Kelzarkar, Tejas Shinde, "Decentralized Chat Application using Blockchain Technology," IJRASET, 2023.
 [13] Mochi Inc. "Mochi: A Universal Protocol Suite for the Blockchain Ecosystem and Real World Data (Nov 2017)", https://mochi.network/mochi_white_paper.pdf,
 [14] Solidity: <https://docs.soliditylang.org/en/v0.6.5/>
 [15] Web3.js:<https://web3js.readthedocs.io/en/v1.10.0/>

CODEIAL: A solution for the coders by the Coders to the Coders

Prabhjot Kaur¹, Ria Mittal², Vaibhav Hasija³, Aakriti Sharma⁴, Varun Sharma⁶

¹ Professor, IT Deptt. 2,3,4,5 Student, IT Deptt
^{1,2,3,4,5} MSIT, GGSIPU, New Delhi

¹prabhjot.kaur@msit.in, ²riamittal3@gmail.com, ³vaibhavhasija2000@gmail.com, ⁴aakritishrm6@gmail.com, ⁵varunsharma55665@gmail.com

Abstract—Software developers rely on media to communicate, learn, collaborate, and coordinate with others. Social media has changed how software developers collaborate, how they coordinate their work, and where they find information. Communities are a fantastic place to share developer advice, resources, snippets of code, and tutorials and to engage in in-depth conversations about various programming languages and other development-related topics. In this research paper, we are providing a solution to the developers by providing them a platform where they can socialise with their other fellow mates, build a community and the opportunity to learn and grow together. This platform will help programmers to ease out their programming journey.

Keywords—Coding, Socializing, Forum, Social media

1. INTRODUCTION

The use of social networking platforms has increased over the past ten years, and they are now very widespread. These platforms allow users to register for an account on the internet, interact with others, and develop a complex web of relationships. Social media is used by billions of people every day to share information and form connections. [2] On a character stage, social media lets in to speak with our friends and household, gain understanding of recent matters, expand our hobbies, and be entertained. On an expert level, we are able to make use of social media to expand our understanding in a particular subject and build our professional community by connecting with other professionals in our industry. At the commercial enterprise stage, social media lets in us to have a communicate with our audience, advantage client remarks, and elevate your brand.[8]

The coding community is a dedicated space where programmers from all around the world can work together on an equal footing, discuss problems and solutions, and announce new advancements in programming languages and Tools [10]. Simply positioned, it is not mandatory however important for programmers of any degree to enlist in a coding community to beautify their degree of programming in addition to advantage uplifting connections [1].

Coding groups are a terrific opportunity to meet like-minded people and to voice any programming-related concerns you may have. Early in their programming careers, novice programmers frequently encounter a variety of challenging scenarios and issues [10].

Although Google is the most comprehensive resource for developers, becoming active in a coding community gives

you more space to discuss your difficulties and makes it easier to find relevant information to solve your problems [5].

Coding clubs are now helpful to experienced engineers as well as the brightest beginners in producing software that is bug-free. Professional software developers, however, frequently indulge in impostor syndrome, a detrimental intellectual behaviour that prevents and makes programmers feel incompetent and useless [5]. For a variety of reasons, many people experience imposter syndrome at the beginning or middle of a career in software programme improvement. You might be able to overcome any negative behaviours that weaken you and put your attention on improving your productivity and wellbeing by joining a coding community [7].

2. BACKGROUND INFORMATION

Customers can use the website as a platform to ask and answer questions, and through membership and active engagement, to vote up or down on questions and answers, much like Reddit, and modify questions and answers similarly to a wiki. [7]

LinkedIn One of the most popular social networking platforms designed for business use is LinkedIn. Currently, it has more than 610 million users, 40% of whom use it daily, who spend an average of 17 minutes each month in this application (99firms.com, 2020). The fact that 57% of all enterprises have a LinkedIn profile and that 92% of Fortune 500 companies utilise it is due to the fact that, according to 99firms.com in 2013, LinkedIn users are known as "connections" and their profile pages place an emphasis on education, work experience, and competencies [8]. Furthermore, it features a limited selection of configurable modules and expert network news feeds.

Stack Overflow As of 2019, Stack Overflow had more than 10 million users and served as a venue for about sixteen million inquiries and answers for programmers. A remarkable fact about this platform is that 92% of questions receive a response in less than eleven minutes on average. Also, 84.4% of commonly performed Google searches for "jQuery API" produced at least one Stack Overflow result, demonstrating that programmers can turn to Stack Overflow for answers to previously asked questions. The data from Stack Overflow tags is helpful for spreading a common perception of technological advancements.[1]

3. PURPOSE AND METHODOLOGY

Codeial only accepts programming-related queries that can be narrowly focused on a specific issue. Customers frequently reject and designate as closed questions that are of a general nature or those that invite answers that are necessarily dependant on opinion. The purpose of the sister website softwareengineering.stackexchange.com is to provide as a forum for more general inquiries, such as frequently asked questions regarding software development. [8] Users can ask and answer questions on the website, and by active engagement and membership, they can vote for or against questions and answers in a manner similar to Reddit and modify questions and answers in a manner similar to a wiki.

4. NEED OF CODEIAL

Stack overflow

While many would question Stack Overflow's usefulness to software developers, concerns about the value of the answers generated by inquiries exist. For instance, Jin, et al. [1] explored how the desire to "win" reputation rewards may impact the quality of an answer or the ability and willingness to underestimate the efficacy of answers offered by others. Anand and Ravichandran [3] located a need for differentiating solution excellent and popularity, because the reward system (contributors on Stack Overflow are rewarded points for their contributions) given by Stack Overflow may additionally sometimes warfare with other high-quality answers. those works imply that there may be need for warning whilst drawing close Stack Overflow for encouraged answers to software engineering challenges [1].

In reality, many engineers find online resources like Stack Overflow helpful when they encounter problems requiring skills they do not have. This raises concerns about their potential comprehension of such code, which in turn raises concerns about the calibre of the software they develop. Moreover, security issues could occur due to the Stack Overflow portal's vulnerable programming, which has been demonstrated by proof [1].

LinkedIn

Heaps of spam messages There's a announcing that among 12 disciples there'll usually be a Judas. think of what number of Judas could be available on a website with over 1200 million humans. LinkedIn is packed with unsolicited mail messages from recruiters, employers, or even task seekers. All just to are looking for attention, lie to, and extort money, and many others.[2]

Time ingesting one of the biggest downfalls and turn offs of LinkedIn is time consumption. the specified amount of time you have to invest earlier than you can discover ways to use it correctly.[4] not like different sites, so that you can get the maximum out of LinkedIn you need to be inclined to set aside some of your valuable time every day.[9]

5. CODEIAL: THE PROPOSED SYSTEM

CODEIAL is an application for programmers where they can add posts ,like a post, comment, audio call,video call and also has a real time chatting engine which establishes a connection between them. It provides programmers a better platform which they can use to ease out their programming journey. Fig. 1 shows the flowchart of proposed system. Fig.2-Fig. 5 show various screen shots to explain the system. Fig. 2 shows the 'HomePage' of the application where users can add their queries in the post section. Fig. 3 shows 'Notification Section' where users can see all the notifications about follow,like and comment. Fig. 4 shows Chat Section, which is provided for easy communication among the users. Fig. 5 shows Profile Section where users can edit their details and can also view the saved posts.

6. COMPARISON WITH OTHER SIMILAR SYSTEMS

Platforms together with Stack Overflow are to be had for software engineers to invite for answers to their challenges. Over time, Stack Overflow has increased its appeal among academics by include subjects on network dynamics, human factors, and technological problems. [5] But it doesn't provide the opportunity to make friends and socialise. CODEIAL is an application for programmers where they can add post , comment, video call and also has a real time chatting engine which establishes a connection between them.[7]

7. CONCLUSION

The CODEIAL serves as a refuge for programmers all over the world. Each developer community is made up of hundreds of programmers at various stages who post any difficulties they encounter along with solutions. Therefore, feel free to ask for assistance in coding groups whenever you have an issue with errors or questions about concepts. This paper developed a similar system for the ease of programmers and developes.

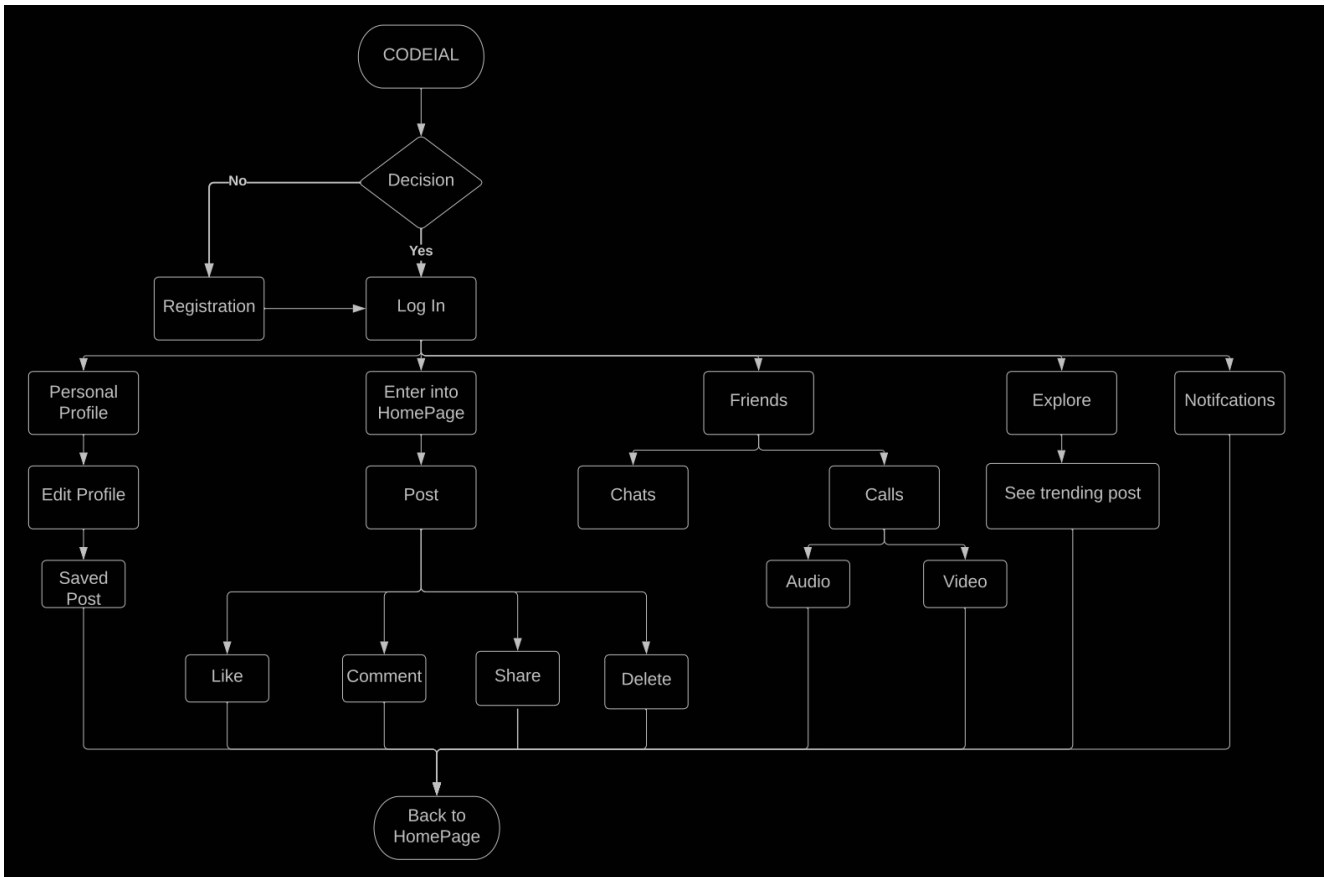


Fig 1: Flowchart of Proposed System: CODEIAL

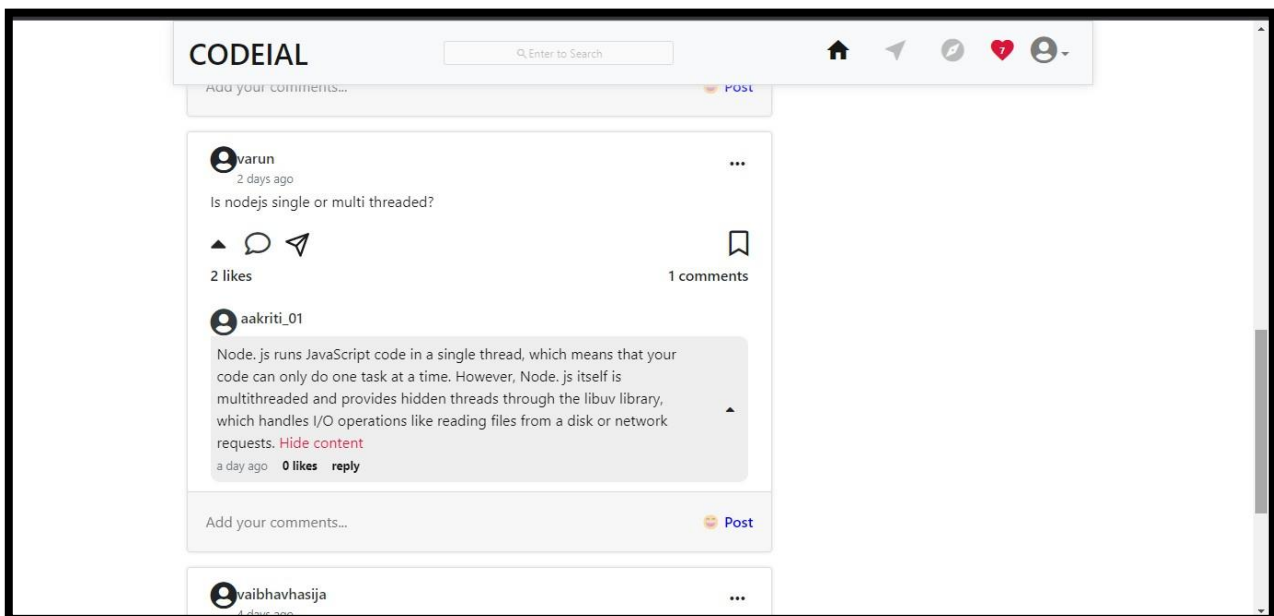


Fig 2: HomePage

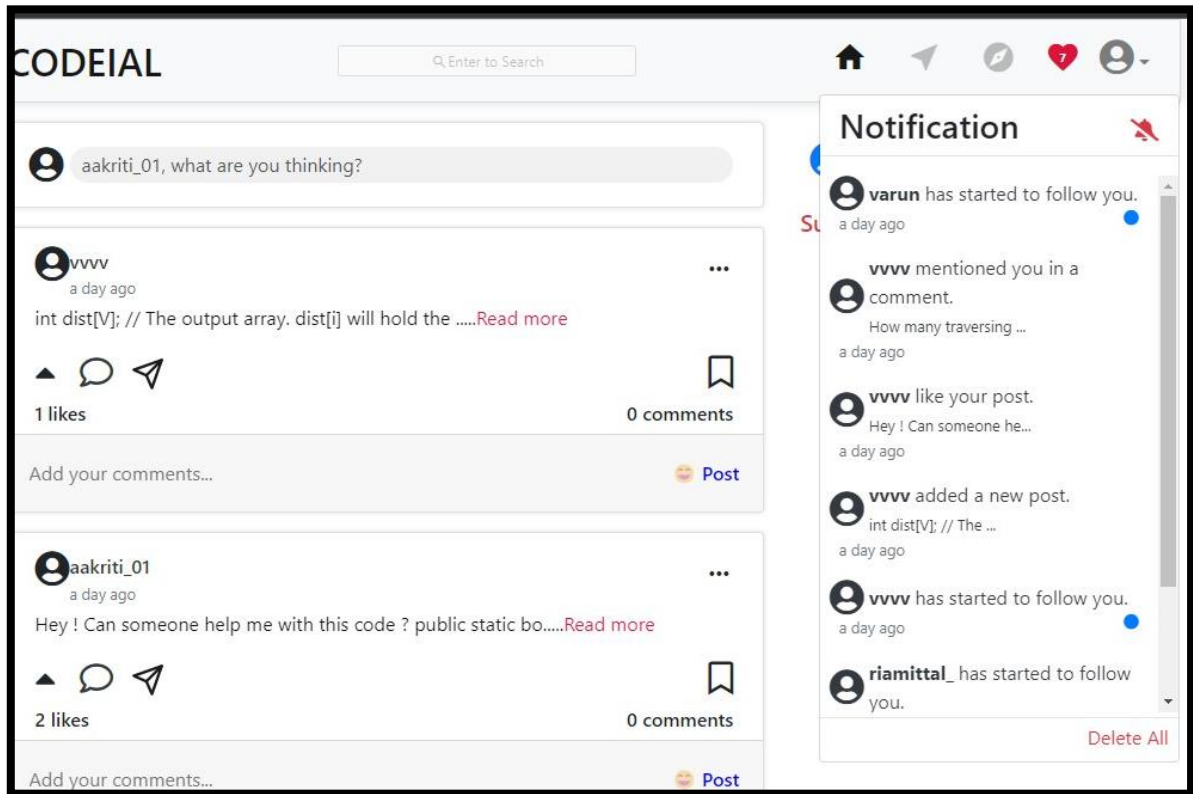


Fig 3: Notification Section

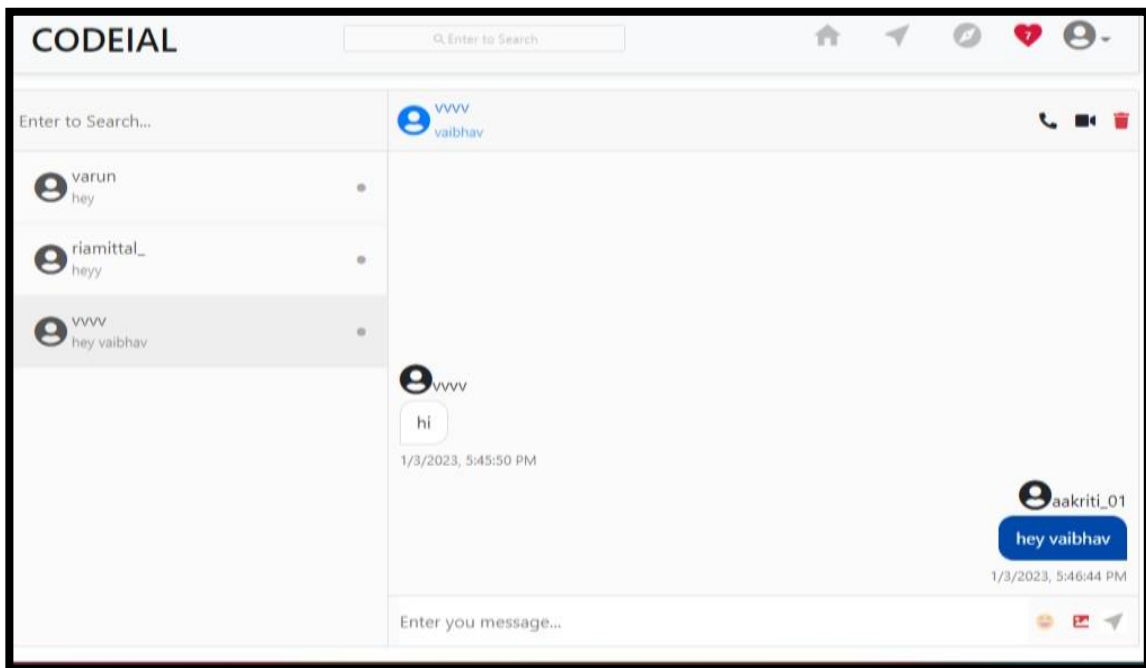


Fig 4: Chat Section

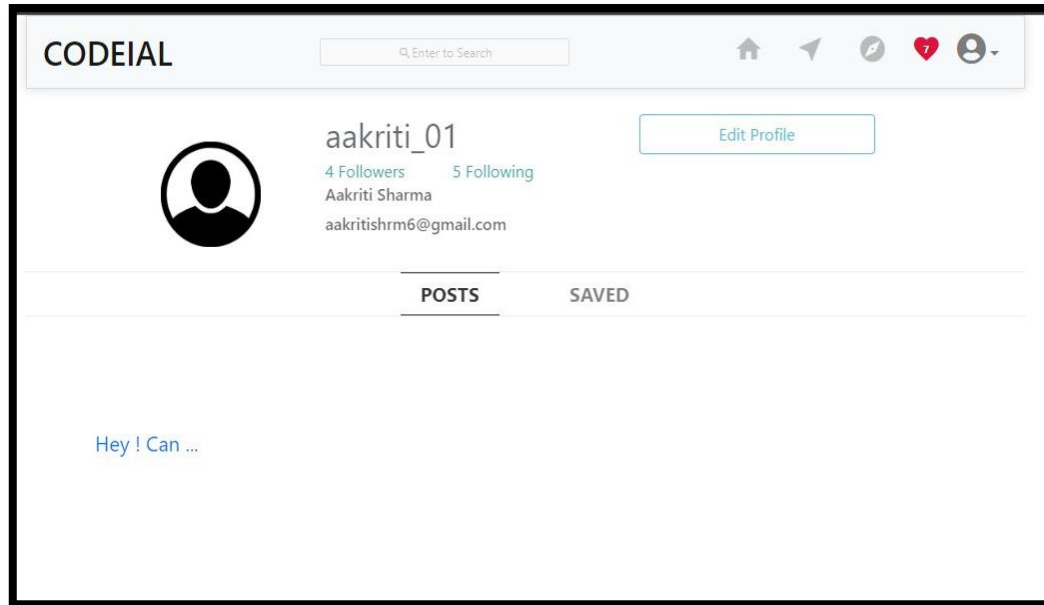


Fig 5: Profile Section

REFERENCES

- [1] Meldrum, S., Licorish, S.A. and Savarimuthu, B.T.R., 2020. Exploring Research Interest in Stack Overflow--A Systematic Mapping Study and Quality Evaluation. arXiv preprint arXiv:2010.12282.
- [2] Agazzi, A.E., 2020. Study of the usability of linkedin: a social media platform meant to connect employers and employees. arXiv preprint arXiv:2006.03931.
- [3] Treude, C., Figueira Filho, F., Cleary, B. and Storey, M.A., 2012. Programming in a socially networked world: the evolution of the social programmer. *The Future of Collaborative Software Development*, pp.1- 3.
- [4] Ecclesie Agazzi, A., 2020. Study of the usability of LinkedIn: a social media platform meant to connect employers and employees. arXiv e-prints, pp.arXiv-2006.
- [5] Meldrum, S., Licorish, S.A. and Savarimuthu, B.T.R., 2017, June. Crowdsourced knowledge on stack overflow: A systematic mapping study. In *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering* (pp. 180-185).
- [6] Ahmad, A., Feng, C., Ge, S. and Yousif, A., 2018. A survey on mining stack overflow: question and answering (Q&A) community. *Data Technologies and Applications*.
- [7] Al-Badi, A.H., Michelle, O.O., Al Roobaea, R. and Mayhew, P., 2013. Improving usability of social networking systems: a case study of LinkedIn. *Journal of Internet Social Networking & Virtual Communities*,
- [8] Basak, E. and Calisir, F., 2014. Uses and gratifications of LinkedIn: An exploratory study. In *Proceedings of the world congress on engineering* (Vol. 2, pp. 1-4).
- [9] Dach, M., 2015. A case study of LinkedIn.
- [10] Baltes, S., Treude, C. and Diehl, S., 2019, May. Sotorrent: Studying the origin, evolution, and usage of stack overflow code snippets. In *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)* (pp. 191-194). IEEE.

Adaptive Algorithm for High-Speed Integrated Filters

¹Archana Balyan, ²Deepshikha Yadav

¹Professor, Department of ECE, MSIT, New Delhi, India

² Assistant Professor, Department of ECE/MSIT, New Delhi, India
archanabalyan@msit.in

Abstract—The most popular adaptive algorithm for high-speed integrated filters today is the LMS algorithm, due primarily to its straightforward and robust digital hardware implementation. However, in analog adaptive filters, implementation of the LMS algorithm is neither straightforward nor robust. The hardware required to generate gradient information is cumbersome and power-hungry. The LMS algorithm's accuracy is also hindered by the presence of dc offsets on the state and error signals. These considerations are the primary factors which presently limit the use of analog adaptive filters. This paper seeks to combine the advantages of digital adaptation algorithms with a high-speed analog signal path. The aim is to perform digital adaptation of analog filters and also several potential applications for digitally adaptive analog filters are presented. All are mixed-signal digital communications systems often including considerable dedicated digital signal processing hardware, usually integrated on the same die as the analog front end. In such systems analog circuit design is particularly challenging, so it is especially desirable to implement the adaptation algorithm digitally. This electronic document is a "live" template and already defines the components of your paper [title, text, heads, etc.] in its style sheet.

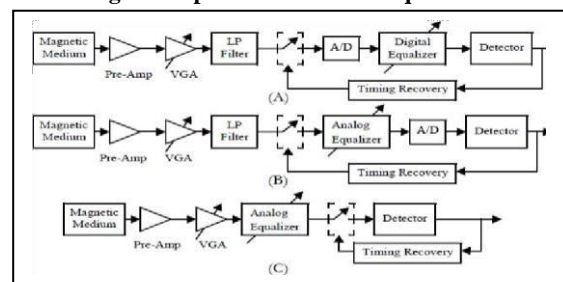
Keywords— Adaptive algorithm; LMS algorithm, analog signal path; Digital Signal Processing; Analog filters

1. INTRODUCTION

The versatile capacities in some computerized interchanges applications can be productively performed carefully, simple versatile channels assume a basic job when fast and low force are required, as in the applications discussed below. In digital magnetic storage read channels and ethernet receivers, analog adaptive filters (usually digitally programmable) have already been used in practical systems. High speed serial links represent a burgeoning application for analog adaptive filters, and speculative research has begun on analog adaptive filters for optical and smart antenna applications. Fig. 1.1 shows three possible architectures for a digital magnetic storage read channel. The adaptive equalization can be implemented using either digital (Fig. 1.1A), analog discrete time (Fig. 1.1B), or analog continuous time (Fig.1.1C) filters. In all- digital systems (Fig. 1.1A) some partial equalization is still performed in the analog domain in combination with the fixed low pass anti-aliasing filter [5]. This is done to reduce the dynamic range and resolution required in the A/D converter and to shorten the length of the digital equalizer required. By making the analog filter adaptive, the digital circuitry and A/D converter complexity are reduced. In some systems, this approach is combined with an analog Viterbi detector to eliminate the A/D converter all together Based on the

preceding discussion, it should be clear that continuous time filters with digitally programmable parameters and a digital adaptation algorithm are very useful for digital magnetic storage applications. Key difference between ethernet and magnetic storage applications is that the channel's impulse responses may be very long, so a transversal filter of great length would be required to perform equalization. Therefore, IIR adaptive filters may be preferred over FIR transversal structures. Several examples of continuous time analog adaptive equalizers for 100 Mb/s ethernet integrated transceivers have been reported. Analog adaptive equalizers for next-generation gigabit ethernet over copper are also under research. In all but, the equalizers are digitally adapted.

Fig 1: Adaptive feed forward equalizer



For very high speed applications, an A/D and digital equalizer may be impractical so analog equalization is the only option. One such application is high speed serial links over coaxial cable. In [2] a continuous time digitally programmable CMOS analog equalizer was used for 155 Mb/s SONET over co-ax. In [2] and [3] bipolar continuous time analog circuits were used for adaptive equalization of a coaxial cable up to 400 Mb/s and 2.5 Gb/s respectively. A 4PAM 8 Gb/s signal was equalized for co-ax using a discrete time analog filter in [4]. Considerable effort has also been aimed at equalizing chip-to-chip interfaces operating at several Gb/s. So far, most of the effort has been directed at pre-equalization using a discrete time mixed signal transmitter. However, in [5] a bondwire inductance was used to provide high frequency peaking in a continuous time received signal at 8 Gb/s.

Although computerized circuits in new procedure advancements will in the long run arrive at these paces, there will stay a recurrence limit past which simple versatile sign preparing is more productive. In order to keep the analog circuit design as simple as possible, the equalizers usually have only one or two adapted parameters, usually adjusting the amount of high-frequency peaking in the filter's transfer

function. Since the connections are fixed, fast adaptation is generally not required so the adaptive algorithm need not be implemented with high speed analog circuits. The best compromise is often to use analog circuits for the high speed signal path only, and slower digital circuits for the adaptation. As digital CMOS circuits increase in speed and decrease in power consumption, research on analog adaptive filters will continue to move towards higher speed applications. Optical and RF signals operate at frequencies still far beyond the practical limits of integrated adaptive digital signal processing. Researchers are already beginning to consider the possibility of analog adaptive signal processing for these applications. Integrated analog equalizers have already been tested in experimental optical systems operating at 10 Gb/s. However, the equalizer parameters were manually optimized. In [6], the control voltages on an array of varactors were adapted (in simulations) to direct the radiation pattern of an antenna array. In both cases, practical hardware- efficient adaptation algorithms are lacking.

2. RELATED WORK

Methods for state perception which are concentrated in the control writing either expect access to both the framework sources of info and yields, or require the framework under perception to be least stage. An inexact time deferred state estimator was proposed which can be applied to any state space framework with obscure sources of info. This technique is particularly relevant in analog filters with digitally programmable parameters since the

LMS algorithm requires estimates of the filter's internal state signals, but it is impractical to build the additional analog circuitry to generate and digitize them. Rather, time postponed state gauges are determined carefully from the channel yield and utilized for LMS variation with negligible extra simple equipment. The method is additionally possibly invulnerable to dc counterbalance impacts which as a rule frustrate simple usage of the LMS calculation. Since the state gauges are created carefully from the channel yield, an adaptive dc tap at the filter output results in state estimates which are offset-free. However, the adaptation algorithm is computationally complex compared to straight LMS adaptation and efficient hardware implementations are still an open issue. The general state observation technique is potentially applicable to any adaptation algorithm where the filter's state information is required but inaccessible. Although only the LMS algorithm was discussed here due to its popularity and straightforward implementation, the state estimates could also be used in (for example) the NewtonRaphson algorithm. Filter adaptation using unknown input state observation was verified in behavioral simulations on a 2-tap adaptive transversal filter and a 5th order continuous time filter. The approximate inverse filter must have a long enough impulse response to provide good state estimates, even when the signal path filter has spectral nulls. If the filter $H^{-1}(Z)$ is made long enough, the performance of the adaptation algorithm approaches that of the standard LMS algorithm.

3. LEAST MEAN SQUARE ALGORITHM (LMS)

A general LMS adaptive filter is shown in Fig.2. It has two inputs (the filter input, and a "desired" or reference output, d) and two outputs (the filter output, and an error signal $e = d - y$). These may be either continuous or discrete time random processes whose statistics will depend upon the particular application. For all of the applications considered here, the signals u and d are either jointly stationary, or their joint probability distributions vary slowly compared to their bandwidth.

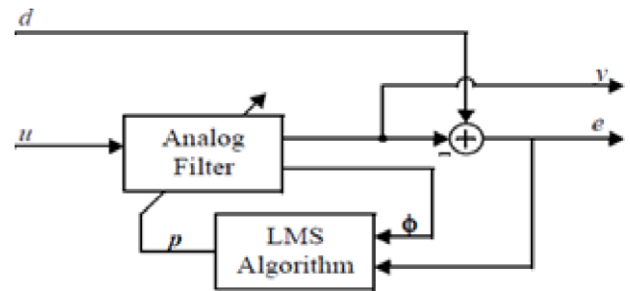


Fig 2: LMS analog adaptive filter as a 2-input, 2-output system

The performance criterion used for LMS adaptation is the mean-squared error (MSE),

$$\epsilon(p(k)) = E [(d(k) - y(k))] = [e^2(k)] \quad (1)$$

Where the operator $E[]$ denotes expectation and discrete time signals have been assumed. In an adaptive system, the parameter vector $p(k)$ and, hence, the MSE $\epsilon(p(k))$ are functions of time. The LMS algorithm is a gradient descent optimizer, which means that it seeks the parameter vector which minimizes $\epsilon(p(k))$ by updating $p(k)$ iteratively in a direction opposite the gradient $\nabla p(k) \epsilon(p(k))$. In discrete time, the update rule is

$$p(k+1) = p(k) - \mu \cdot \nabla p(k) \epsilon(p(k)) \quad (2)$$

where μ is a constant which determines the rate of adaptation. The simple yet brilliant idea put forward by Window and Hoff in [6] was to drop expectation operator when substituting Eqn. (1) into Eqn. (2). In doing so, they are taking the instantaneous value of the squared-error to be a noisy estimate of its expected value, $\epsilon(p(k)) = E[e^2(k)] \approx e^2(k)$.

$$\begin{aligned} p(k+1) &= p(k) - \mu \cdot \nabla p(k) e^2(k) \\ &= p(k) - \mu \cdot (2e(k)) \cdot \nabla p(k) (d(k) - y(k)) \\ &= p(k) - 2\mu e(k) \cdot (-\nabla p(k) y(k)) \\ &= p(k) + 2\mu e(k) \cdot \phi(k) \end{aligned} \quad (3)$$

Where $\phi(k)$ is an instantaneous estimate of the gradient $\nabla p(k) y(k)$. Eqn. (3) is the parameter update equation for the LMS algorithm. The parameter determines the rate of adaptation. In digital filters, the gradient signals $\phi(k)$ are usually readily available in digital form [3]. In analog filters, additional analog circuitry is often required to generate the gradient signals from the filter's internal state signals. Furthermore, if the adaptation algorithm is implemented digitally the gradient signals must be digitized, which is an area- and power-hungry. Alternately, the LMS algorithm can operate directly on analog gradient signals using analog circuitry. In this case, a continuous time formulation of Eqn. (3) is used [6]:

$$t(t) = 2\mu \int_{-\infty}^{\infty} e(r) \cdot (r) dr \quad (4)$$

4. DC OFFST EFFECTS IN ANALOG LMS

A block diagram of the LMS parameter update rule appears in Fig.3. If the parameter updates are being performed using analog circuitry, dc offsets will appear at the inputs to the multiplier multiplier (m , $m\phi$) and integrator ($me\phi$). These offsets prevent the LMS algorithm from adapting to the optimal filter parameter values [3], [4], [5]. The excess steady state MSE is related to $me\phi$ and to the product $m \cdot m\phi$.

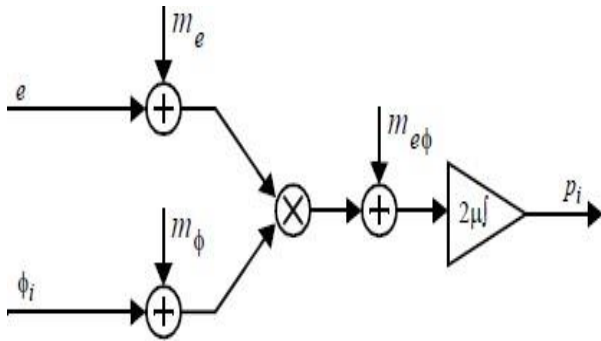


Fig 3: Analog implementation of the LMS parameter update equation

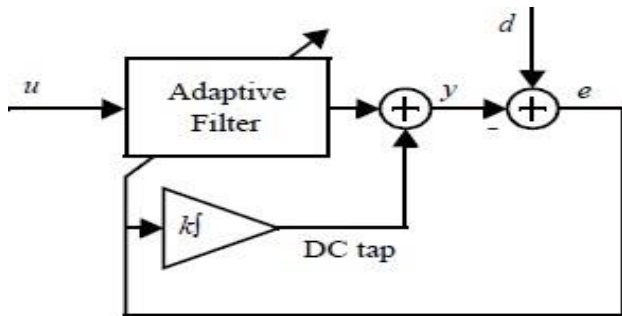


Fig 4: DC tap for adaptive offset cancellation

DC counter balances speak to a noteworthy presentation impediment in numerous simple versatile channels [2], [3], and [4]. Much exploration has been done to limit the negative impact of dc balances on simple versatile channels. It was appeared in [5] that the SE-LMS and SS-LMS calculations are to some degree more powerful than full-LMS concerning dc counterbalances. An algorithmic way to deal with combating dc balance impacts in cross-over channels was proposed in [7] requiring another arrangement of N adjusted coefficients. Circuit-level procedures for counterbalance pay in simple versatile channels have likewise been utilized with shifting degrees of accomplishment in, for example, [8], [9], and [10]. A relatively simple way to eliminate dc offsets on the error signal is to add a dc offset cancellation tap to the filter output. Shown in Fig.4, this tap can be included in any filter structure and essentially forces e to have zero dc content. As a result, $me \approx 0$

$\Rightarrow m \cdot m\phi \approx 0$. In [4], the SS-LMS algorithm was used to adapt the dc offset cancellation tap resulting in a median-

based offset compensation scheme. A hardware efficient Digital implementations of the LMS update equation are advantageous because performing the multiplication $e \phi$ digitally introduces no dc offset. Hence, $m e\phi = 0$. By combining a digital implementation of the LMS algorithm with offset compensation at the filter output, one can perform LMS adaptation of an analog filter with no dc offset effects. Therefore, it is highly desirable to use digital adaptation for analog adaptive filters.

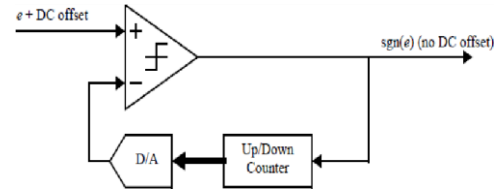


Fig 5: Median-based DC offset compensation scheme

realization is shown in Fig. 6. Unfortunately, it does not eliminate excess MSE entirely since dc offsets introduced by $me\phi$ persist.

Digital implementations of the LMS update equation are advantageous because performing the multiplication $e \cdot \phi$ digitally introduces no dc offset. Hence, $me\phi = 0$. By combining a digital implementation of the LMS algorithm with offset compensation at the filter output, one can perform LMS adaptation of an analog filter with no dc offset effects. Therefore, it is highly desirable to use digital adaptation for analog adaptive filters.

5. SIMULATIONS AND RESULTS

We use the FFT-2 method for the optimization of research filters. Only $M = 32$ with the APA subband pre-conditioner strategy with the first objective function, eigenvalue distribution and normality, is seen due to the complexity of this optimization. Learning curves of 1,000 runs as seen in figures 6 and 7. Compared with the reference filter of length 2L, Figure 8 shows the filter reaction.

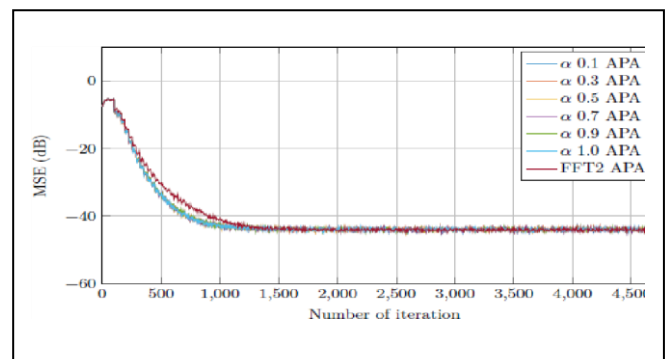


Fig 6: Comparison of analysis filter optimization $M = 32$, APA

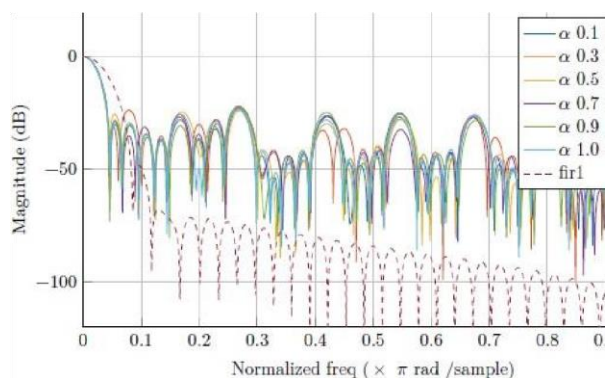


Fig 7: Filter response for analysis prototype filter $M=32$, APA.

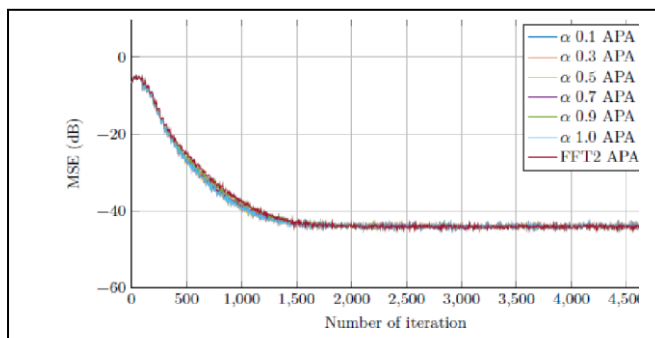


Fig 8: Comparison of analysis filter optimization $M = 32$, APA

6. CONCLUSION

The LMS algorithm is the most popular algorithm for the adaptation of integrated filters today. Although several other algorithms with superior convergence properties exist, the LMS algorithm remains popular because of its robust and straightforward digital implementation. This Paper provides some background on the LMS algorithm and the dc offset effects which limit its use in analog adaptive filters. Digital implementations of the LMS update equation are advantageous because performing the multiplication $e_n \hat{\theta}_n$ digitally introduces no dc offset. Hence $m e_n \hat{\theta}_n = 0$. By combining a digital implementation of the LMS algorithm with offset compensation at the filter output, one can perform LMS adaptation of an analog filter with no dc offset effects. Therefore, it is highly desirable to use digital adaptation for analog adaptive filters.

REFERENCES

- [1] K. A. Kozma, D. A. Johns, and A. S. Sedra, "Automatic Tuning of Continuous-Time Integrated Filters Using an Adaptive Filter Technique," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1241-1248, Nov. 1991
- [2] D. A. Johns, W. M. Snelgrove, and A. S. Sedra, "Continuous-Time LMS Adaptive Recursive Filters," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 769-778, July 1991.
- [3] P. K. D. Pai, A. D. Brewster, and A. Abidi, "A 160-MHz Analog Front-End IC for EPR-IV PRML Magnetic Storage Read Channels," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1803-1816, Nov. 1996.
- [4] J. E. C. Brown, P. J. Hurst, B. C. Rothenberg, and S. H. Lewis, "A CMOS Adaptive Continuous-Time Forward Equalizer, LPF, and RAM-DFE for Magnetic Recording," *IEEE J. Solid-State Circuits*, vol. 34, pp. 162-169, Feb. 1999.

- [5] O. Shoaib, A. Shoval, and R. H. Leonowich, "A 3V Low-Power 0.25 μ m CMOS 100 Mb/s Receiver for Fast Ethernet," *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 308-309, Feb. 2000.
- [6] M. Altmann, J. M. Caia, R. Morle, M. Dunsmore, Y. Xie, and N. Kocaman, "A lowpower CMOS 155 Mb/s Transceiver for SONET/SDH over Co-ax & Fibre," *IEEE Custom Integrated Circuits Conf.*, pp. 127-130, May 2001.
- [7] R. Farjad-Rad, C.-K. K. Yang, and M. A. Horowitz, "A 0.3 μ m CMOS 8 Gb/s 4-PAM Serial Link Transceiver," *IEEE J. Solid-State Circuits*, vol. 35, pp. 757-764, May 2000.
- [9] C.-K. K. Yang, V. Stojanovic, S. Modjtahedi, M. A. Horowitz, and W. F. Ellersick, "A Serial-Link Transceiver Based on 8-GSamples/s A/D and D/A Converters in 0.25 μ m CMOS," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1684-1692, Nov. 2001.
- [10] F. Buchali, H. Bulow, W. Baumert, R. Ballentin, and T. Wehren, "Reduction of the Chromatic Dispersion Penalty at 10 Gbit/s by Integrated Electronic Equalisers," *Optical Fiber Comm. Conf.*, vol. 3, pp. 268-270, March 2000.
- [11] G. Bollati, S. Marchese, M. Demicheli, and R. Castello, "An EighthOrder CMOS Low-Pass Filter with 30-120 MHz Tuning Range and Programmable Boost," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 7, pp. 1056-1066, July 2001.

Career Guidance AI Counsellor

Dr. Sapna Malik¹, Disha Katyal², Naman Sachdeva³, Nihal Gupta⁴

^{1,2,3,4}Department of Computer Science and Engineering

Maharaja Surajmal Institute of Technology, New Delhi

¹sapnadhankhar@msit.in, ²dishakatyal01@gmail.com,

³namansachdeva160201@gmail.com, ⁴nihalgupta11213@gmail.com

Abstract - The "Career Guidance AI Counsellor" idea is a web service that gives students career advice based on their personality, talents, and interests. The programme analyses a dataset that includes the user's responses to surveys that assess their abilities, skills, character, and interests using three classifier models. The application generates an assortment of three proposed careers based on the answers provided by the user, and then gives more specific information about each field and the abilities required for them. Career counselling is critical in assisting individuals in making educated decisions about their career trajectories.

Keywords - AI, Career Guidance, Aptitude, Personality

1. INTRODUCTION

Making wise career decisions can be difficult for people in the modern fast-paced and continuously changing job market. With the introduction of artificial intelligence (AI), new techniques to address these difficulties have emerged. The creation of AI-based career counsellors is one such application. With an emphasis on their efficacy, constraints, and potential, the literature analysis aims to examine the state of developments and research in the area of career counselling AI counsellors. Effective career counselling is critical in supporting individuals in making sensible choices about their professional options. It assists individuals in identifying their interests, abilities, and skills while matching them with appropriate work prospects. Face-to-face counselling and standardised evaluations, for example, have constraints with regard to scalability, personalisation, and efficiency. This has encouraged the field to investigate AI-based solutions. Artificial intelligence techniques such as machine learning, natural language processing (NLP), and recommendation systems have been widely used in the development of AI counsellors for career assistance. Counsellors can employ NLP to comprehend and analyse consumer inquiries, resulting in more natural and dynamic interactions. Algorithms based on machine learning are used to analyse massive volumes of career-related data in order to find patterns and trends and provide appropriate suggestions. Recommendation systems aid in recommending appropriate job possibilities based on user interests and aptitude. One of the most significant benefits of AI-based counsellors is their capacity to deliver personalised advice. These counsellors can provide personalised advice by exploiting user data such as interests, abilities, and past education. Personalization boosts user engagement, contentment, and overall career guidance efficacy. Individuals who receive personalised career counselling

make better decisions and are more satisfied with their jobs, according to studies. The user interface and relationship with career counselling AI counsellors play an important role in AI's acceptability and efficacy. According to research, developing intuitive and easy to use interfaces is critical for facilitating seamless interactions. Conversational agents with human-like answers and empathy have been found to improve the user experience.

Furthermore, including feedback systems and continual learning capabilities in counsellors might help them grow over time.

Research into the efficacy of career counselling AI counsellors have shown positive outcomes. These counsellors provide simplified access to career knowledge and assistance while overcoming geographical and temporal constraints. They can process a huge number of user inquiries at the same time, ensuring scalability. However, issues remain. The timeliness and reliability of the data that underlies sources determine the accuracy and validity of guidance. AI counsellors may struggle with complex or confusing enquiries, and they may not completely comprehend the context or psychological aspects of user interactions. AI counsellor development and deployment create ethical concerns. Information about users must be managed securely and with consent, therefore security of data and privacy are critical. Another source of concern is bias in AI algorithms, which may perpetuate existing inequities and limit the range of employment alternatives provided to consumers. To ensure openness, accountability, and justice regarding the application of AI counsellors, ethical rules and legislation are required. Millions of students have difficulty understanding their job alternatives due to restricted access to information, opportunity, and qualified mentors.

Problem Statement: Assist students in discovering and exploring their interests, as well as deciding on future career pathways and skilling alternatives. **Description:** The following aspects should be included in the solution:

Answer what possible next steps are based on the overview / subject/topic interest. - What future employment opportunities await you?

Some of the possible approaches that can be pursued to attain the goal, based on the ultimate goal/dream/aspiration, include suggested upskilling programmes from DIKSHA/Swayam/Skill platforms globally. What are the best possibilities for skill development, best venues to study further, and best ways to train oneself in areas of interest?

Who would be some of the most successful individuals I know with comparable professional pathways who can be approached for help according to the type of job track that student is on? Students can envision future paths based on orientation and desire. To gather information, the team will investigate ways of extracting existing data sets accessible through popular sites. This information could serve as the foundation for the exploratory tool. The tool might use best practices and AI/ML to provide the necessary information to guide students.

2. LITERATURE REVIEW

Why do we need AI guidance counselors? Jobs in fields that include mobile applications, cyber security, cloud computing, and big data analytics are on the rise. Companies are increasingly reliant on highly skilled and specialised workers. Although the rising number of opportunities is an encouraging sign for graduates, they may be confused about the best future career. As a result, there is an urgent need to develop and execute a system that may aid graduates in choosing a profession that fits their skills. For a long time, recommendation systems have served as an area of study. In excess of 565,000 fresh employment opportunities were created in the last decade, for example, during 2001 and 2011. Furthermore, jobs increased even throughout the recession. Software developers and programmers are predicted to add 279,500 positions by 2022, accounting for over 40% of new opportunities in the computing and mathematics professions, according to the US Bureau of Labour Statistics.

Although the projected increase for information security analysts is lower than that of software developers and programmers, the projected rate of growth for this position is expected to be 36.5%, making it the fastest rising IT job. A variety of factors contribute to the rising demand for vocations, including the greater requirement for cybersecurity, the introduction of electronic health records, and the expanded usage of mobile technology. DevOps engineer, helpdesk and desktop support professional, cloud architect, data professional, developer, AI architect, business intelligence analyst, network/cloud administrator, network security professional, and system administrator are the top ten most in-demand positions in 2020. India has witnessed tremendous economic expansion, which has resulted in increasing utilisation of Information and Communications Technologies (ICTs) in numerous sectors of industry and government. Many new employment are being generated as a result of rising businesses requiring high technical skills, putting strain on local labour marketplaces and educational systems. Despite this, the supply-demand mismatch in India will widen further.

A career consultation AI counsellor is a cutting-edge use of machine learning in the discipline of career guidance. It provides personalised counsel to individuals by utilising techniques such as machine learning, natural language processing, and recommendation algorithms. It improves the method of decision-making, scalability, and usability of career assistance by analysing user data and providing personalised recommendations. However, issues like having

data accuracy and dealing with sophisticated queries persist. The opportunity for the future lies in combining new innovations and collaborations to enhance accuracy, up-to-date counsel, and user experience.

3. MATERIALS AND METHODS

This section provides a quick explanation of the machine learning (ML) algorithms implemented in Career Guidance AI Counsellor to help you better understand our product design.

3.1. XGBoost

XGBoost (Extreme Gradient Boosting) is a sophisticated machine learning algorithm that has shown amazing success in a variety of predictive modelling problems. In this discussion, we will look at how XGBoost works and examine its essential components and features.



Fig 1: XGBoost Algorithm

3.2. Introduction to XGBoost:

XGBoost is a method of ensemble learning that improves performance and efficiency by combining the advantages of gradient boosting methods with various innovations. Tianqi Chen created it in 2014, and it has since grown into a frequently used technique in both academics and industry.

3.3. Gradient Boosting:

To understand XGBoost, you must first comprehend the notion of gradient boosting. Gradient boosting serves as an iterative ensemble strategy for creating a strong model of prediction by combining numerous weak learners (usually decision trees). The procedure entails training new models consecutively to remedy the errors committed by prior models. The final version is the weighted total of these poor learners, with the weights set by training errors.

3.4. XGBoost's Objective Function:

XGBoost introduces a novel objective function that combines a loss function and a regularization term. The loss function quantifies the errors between the predicted values and the actual values, while the regularization term helps control the complexity of the model and prevents overfitting. By optimizing this objective function, XGBoost seeks to find the best set of model parameters that minimize the overall loss.

3.5. Decision Trees in XGBoost:

Decision trees are the base learners used in XGBoost. However, XGBoost employs a specific type of decision tree known as a regression tree. These trees make predictions by assigning a constant value to each leaf node rather than using conditional rules. XGBoost constructs regression trees in a

greedy manner, recursively partitioning the data based on selected features to minimize the loss function.

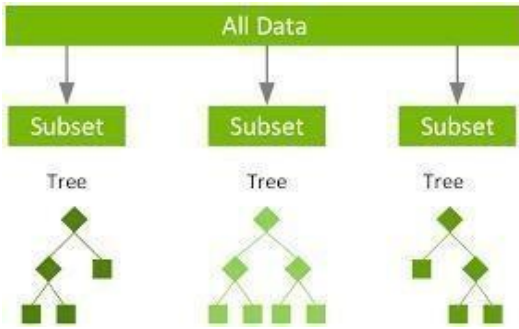


Fig 2: Decision Trees In XGBoost

3.6. Boosting Process:

The boosting process in XGBoost involves iteratively adding new trees to the ensemble. At each iteration, XGBoost calculates the gradients of the loss function with respect to the current predictions. The new trees are trained to approximate the negative gradients, meaning they attempt to correct the mistakes made by the previous models. XGBoost uses a technique called "gradient-based one-side sampling" to efficiently select a subset of the training examples for each tree.

3.7. Feature Importance and Regularization:

XGBoost provides valuable insights into feature importance, allowing users to understand which features have the most significant impact on the predictions. Feature importance is computed based on the number of times a feature is used to split across all the trees in the ensemble. Additionally, XGBoost's regularization techniques, such as L1 and L2 regularization, help prevent overfitting and improve generalization by adding penalty terms to the objective function.

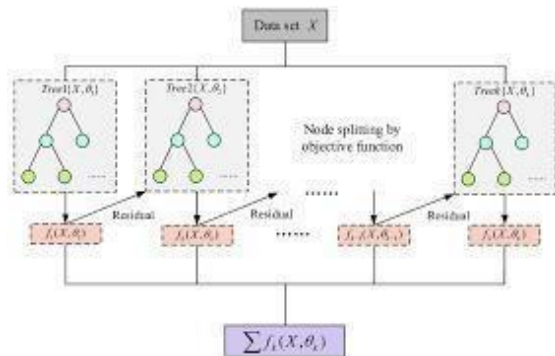


Fig 3: Feature Importance and Regularization

3.8. Handling Missing Values and Outliers:

XGBoost has built-in capabilities to handle missing values and outliers. During the tree construction process, XGBoost automatically learns the best direction to handle missing values based on the training data. It also provides options for users to customize the behavior of missing values. Regarding outliers, XGBoost incorporates a regularization term that helps make the algorithm more robust to extreme values.

3.9. Parallelization and Scalability:

XGBoost is designed to be highly scalable and efficient. It utilizes parallel computing techniques to process data in

parallel and accelerate training. XGBoost can take advantage of multi-core processors to perform computations concurrently, making it suitable for large datasets and complex models.

3.10. Hyperparameter Tuning:

XGBoost offers a wide range of hyperparameters that can be tuned to optimize the performance of the model. These hyperparameters control various aspects of the algorithm, including the learning rate, tree depth, regularization terms, and sampling methods. Careful tuning of hyperparameters is crucial to achieving the best possible model performance.

Initialization:

1. Given training data from the instance space $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ where $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y} = \{-1, +1\}$.
2. Initialize the distribution $D_1(i) = \frac{1}{m}$.

Algorithm:

```

for t = 1, ..., T: do
    Train a weak learner  $h_t: \mathcal{X} \rightarrow \mathbb{R}$  using
    distribution  $D_t$ .
    Determine weight  $\alpha_t$  of  $h_t$ .
    Update the distribution over the training
    set:
    
```

$$D_{t+1}(i) = \frac{D_t(i)e^{-\alpha_t y_i h_t(x_i)}}{Z_t}$$

where Z_t is a normalization factor chosen so that D_{t+1} will be a distribution.

end for

Final score:

$$f(x) = \sum_{t=0}^T \alpha_t h_t(x) \text{ and } H(x) = \text{sign}(f(x))$$

Fig 4: XGBoost Algorithm

4. DATA

4.1. Model 1: Class 10th Students Stream

Selection:

To gather data for the stream selection model targeting 10th-grade students, a survey-based approach was adopted. We designed surveys and questionnaires to capture essential information regarding students' preferences, interests, and aptitudes. These surveys were administered in various schools, ensuring a representative sample across different regions. Care was taken to maintain anonymity and confidentiality throughout the data collection process. The responses obtained formed the basis of our dataset, providing valuable insights into students' career aspirations and aiding in stream selection guidance.

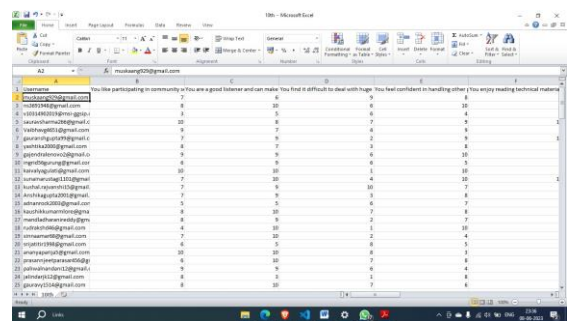


Fig 5: Dataset for class 10th Model

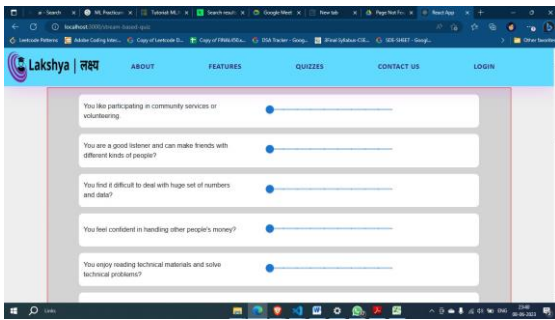


Fig 6: Question form for class 10th students

4.2. Model 2: Class 12th Students Career Choice:

For the second model aimed at assisting class 12th students in making career choices, an extensive online data collection approach was adopted. Our team meticulously researched career advice forums, industry reports, academic publications, and professional networking platforms. This process involved aggregating information from reputable online sources, analyzing success stories, and examining career trajectories of individuals in various fields. The collected data was carefully curated and validated to ensure accuracy and relevance. The resulting dataset enabled our AI model to provide informed career recommendations to students based on industry trends and individual preferences.

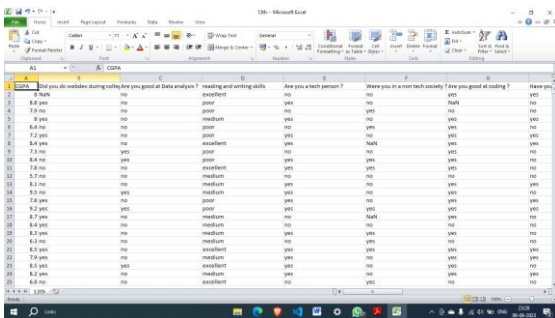


Fig 7: Dataset for class 12th Model

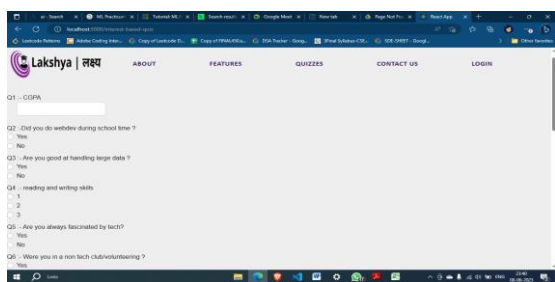


Fig 8: Question form for class 12th student

4.3. Model 3: Undergraduates Skill-Based Career Domain Suggestions:

The data collection process for the skill-based career domain suggestions model targeted at undergraduates revolved around mining data from online platforms. Our team explored coding forums, design platforms, entrepreneurial networks, and similar online communities to gather information about diverse skillsets and associated career domains. By analyzing discussions, portfolio showcases, and skill-related queries,

we obtained insights into the alignment between specific skills and potential career paths. Rigorous filtering and validation techniques were employed to ensure the credibility and reliability of the collected data, enabling our AI model to offer tailored career suggestions based on individual skills.

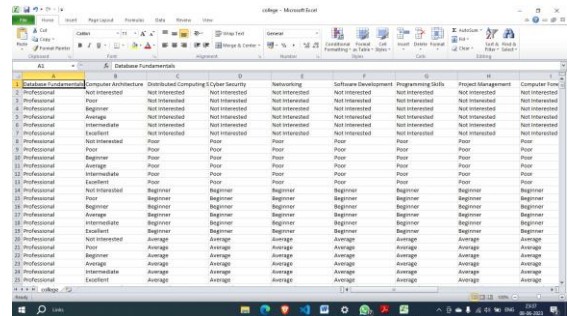


Fig 9: Dataset for Undergraduate's Model

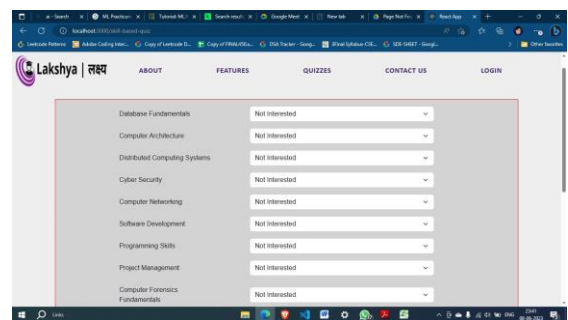


Fig 10: Question form for Undergraduate's Model

4.4. Data pre-processing

Data preprocessing is a crucial step in any data analysis task, including survey-based approaches. It involves cleaning, transforming, and organizing the survey data to ensure its quality and prepare it for further analysis. Here are the key steps involved in data preprocessing for a survey-based approach: Data Pre-processing for Class 10th data These steps represent the preprocessing that was carried out on the dataset for stream suggestion, considering the factors that influenced students' stream selection after class tenth.

- **Data Cleaning:**

Duplicate entries were removed from the dataset to ensure each student was represented only once. Missing values in the survey responses were handled by either imputing them or removing rows with missing values to ensure complete data for accurate analysis.

- **Feature Selection:**

The survey questions were reviewed, and the features most likely to influence students' stream selection were selected. Relevant features, such as academic performance, interests, career aspirations, extracurricular activities, parental influence, and guidance received, were chosen for analysis.

- **Encoding Categorical Variables:**

Categorical variables were converted into numerical representations suitable for analysis. For example, categorical variables like "Science," "Commerce," and "Arts" were converted into numerical codes like 0, 1, and 2, respectively.

- **Handling Ordinal Variables:**

If the survey included ordinal variables, numeric values were assigned to reflect the order or intensity of the responses. For instance, if students were asked to rate their interest in different subjects on a scale of 1 to 5, corresponding numeric values were assigned to each response.

- **Scaling Numerical Variables:**

Numerical variables, such as academic scores or age, were scaled to ensure they were on a similar scale for accurate analysis.

Standardization or normalization techniques were applied to align the ranges of these variables.

- **Feature Engineering:**

Additional features were created, if needed, by combining or transforming existing features. For example, the overall average academic score was calculated, a feature indicating participation in relevant extracurricular activities was created, or a cumulative score based on students' interests in different subjects was computed.

- **Splitting the Dataset:**

The dataset was divided into training and testing subsets to accurately evaluate the performance of the stream suggestion model. Care was taken to maintain the proportion of students from each stream in both subsets to avoid introducing bias.

- **Handling Imbalanced Data (if applicable):**

Imbalances in the distribution of students across different streams were checked.

Techniques such as oversampling or under sampling were considered to balance the data and prevent bias if there was a significant disparity in the number of students among streams.

- **Normalization (Optional):**

The dataset was further normalized based on the requirements of the chosen machine learning algorithm.

Features were normalized to a specific range or other normalization techniques were applied, if necessary.

Data Pre-processing for Class 12th data

- **Data Cleaning:**

Duplicate entries were removed from the dataset to ensure each student is represented only once. Missing values in the survey responses were handled by either imputing them or removing rows with missing values to ensure complete data for accurate analysis.

- **Feature Selection:**

The dataset was reviewed, and the relevant features related to students' interests were selected for career suggestion.

Features could include interests in specific subjects, hobbies, skills, extracurricular activities, and personal preferences.

- **Encoding Categorical Variables:**

Categorical variables were converted into numerical representations suitable for analysis. One-hot encoding or label encoding techniques were applied to represent categorical variables, depending on their nature and the machine learning algorithm's requirements.

- **Handling Ordinal Variables:**

If the survey included ordinal variables, appropriate numeric values were assigned to represent the order or intensity of the

responses. For instance, if students were asked to rate their interest or proficiency level on a scale of 1 to 5, corresponding numeric values were assigned.

- **Scaling Numerical Variables:**

Numerical variables, such as ratings or scores, were scaled to ensure they were on a similar scale for accurate analysis.

Standardization or normalization techniques were applied to align the ranges of these variables.

- **Feature Engineering:**

Additional features were created, if needed, by combining or transforming existing features. For example, composite interest scores or indices were calculated based on multiple interest-related features to capture overall interest profiles.

- **Splitting the Dataset:**

The dataset was divided into training and testing subsets to evaluate the performance of the career suggestion model accurately. Common splits, such as 70-30 or 80-20 for training and testing, respectively, were used.

- **Handling Imbalanced Data (if applicable):**

Imbalances in the distribution of suggested career paths were checked.

- **Feature Selection:**

The dataset was reviewed, and the relevant features related to students' interests were selected for career suggestion.

Features could include interests in specific subjects, hobbies, skills, extracurricular activities, and personal preferences.

- **Encoding Categorical Variables:**

Categorical variables were converted into numerical representations suitable for analysis. One-hot encoding or label encoding techniques were applied to represent categorical variables, depending on their nature and the machine learning algorithm's requirements.

- **Handling Ordinal Variables:**

If the survey included ordinal variables, appropriate numeric values were assigned to represent the order or intensity of the responses. For instance, if students were asked to rate their interest or proficiency level on a scale of 1 to 5, corresponding numeric values were assigned.

- **Scaling Numerical Variables:**

Numerical variables, such as ratings or scores, were scaled to ensure they were on a similar scale for accurate analysis.

Standardization or normalization techniques were applied to align the ranges of these variables.

- **Feature Engineering:**

Additional features were created, if needed, by combining or transforming existing features. For example, composite interest scores or indices were calculated based on multiple interest-related features to capture overall interest profiles.

- **Splitting the Dataset:**

The dataset was divided into training and testing subsets to evaluate the performance of the career suggestion model accurately. Common splits, such as 70-30 or 80-20 for training and testing, respectively, were used.

- **Handling Imbalanced Data (if applicable):**

Imbalances in the distribution of suggested career paths were checked.

Numerical variables, such as years of experience or skill ratings, were scaled to ensure they were on a similar scale for accurate analysis. Standardization or normalization techniques were applied to align the ranges of these variables.

- **Feature Engineering:**

Additional features were created, if needed, by combining or transforming existing features. For example, composite skill scores or indices were calculated based on multiple skill-related features to capture overall skill profiles.

- **Splitting the Dataset:**

The dataset was divided into training and testing subsets to evaluate the performance of the domain selection model accurately. Common splits, such as 70-30 or 80-20 for training and testing, respectively, were used.

- **Handling Imbalanced Data (if applicable):**

Imbalances in the distribution of suggested domains were checked.

Techniques like oversampling or undersampling were employed to address any significant imbalances and ensure fair representation of different domains.

5. TRAINING AND VALIDATION

Ensuring the accuracy and reliability of the career guidance AI models is of paramount importance to deliver effective recommendations. This write-up presents a comprehensive overview of the model validation and training process employed for the three career guidance AI models. By leveraging rigorous techniques, careful dataset splitting, and robust evaluation metrics, we ensured the models' performance and generalization capabilities.

5.1. Model Validation:

To assess the performance of the career guidance AI models, a systematic approach to model validation was adopted. The dataset was divided into training and validation sets, enabling us to evaluate the models on unseen data and avoid overfitting. The following steps outline the validation process:

5.2. Dataset Splitting:

The initial dataset was randomly split into training and validation sets, adhering to a predefined ratio. The training set, typically comprising 70-80% of the data, was used to train the models, while the remaining portion was reserved for evaluation purposes.

5.3. Cross-Validation:

To further ensure robustness, k-fold cross-validation was employed. The training set was divided into k subsets, or folds, where each fold acted as a validation set while the remaining folds were used for training. This process was repeated k times, with each fold serving as the validation set once. The results were then averaged to obtain an overall performance metric.

5.4. Hyperparameter Tuning:

Hyperparameters, such as learning rate, maximum tree depth, and regularization terms, play a vital role in model performance. To identify the optimal combination of

hyperparameters, a grid search or random search technique was employed. This involved systematically varying the hyperparameter values and evaluating the models' performance on the validation set. The hyperparameter configuration yielding the best performance was selected.

5.5. Evaluation Metrics:

To measure the models' performance, appropriate evaluation metrics were chosen based on the specific task of each model. Commonly used metrics included accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC). These metrics provided insights into the models' classification or regression performance and their ability to make accurate predictions.

5.6. Model Training:

The model training process involved the following steps, ensuring the models' ability to learn from the data and make informed career recommendations:

Data Preprocessing:

The raw data underwent preprocessing steps, including handling missing values, outlier treatment, and feature scaling. Missing values were imputed using appropriate techniques, outliers were either corrected or removed, and features were scaled to a common range to ensure fair comparison.

5.7. XGBoost Training:

XGBoost, a powerful and versatile ensemble learning algorithm, was chosen for all three career guidance AI models. The models were trained using the training dataset, incorporating the optimal hyperparameters identified during the validation phase. XGBoost's gradient boosting algorithm iteratively built decision trees to minimize the specified loss function and achieve superior predictive performance.

5.8. Iterative Refinement:

The models underwent iterative refinement through multiple training cycles. Each cycle involved evaluating the model's performance on the validation set, making adjustments to hyperparameters if necessary, and retraining the model on the augmented dataset. This iterative process continued until the desired performance metrics were achieved.

6. CONCLUSIONS

To forecast the optimal career route for recent graduates in computing, five machine learning methods were developed and tested. We trained KNN, DT, Bagging meta-estimator, Gradient Boosting, and XGBoost. According to the experimental data, the XGBoost algorithm surpasses the other models in terms of assisting with career path choices.

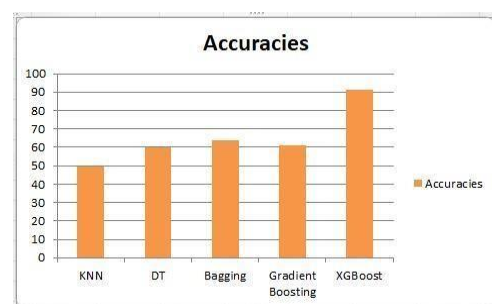


Fig 11: Comparing the algorithms tried on sample dataset

6.1. Experimental Configuration

Hardware and Software Configuration The experiments were conducted on a machine with an Intel Core i5 processor, 8GB RAM, and no dedicated GPU. The operating system used was Windows 10. The experiments were implemented using Python 3.9 and the following software libraries: XGBoost 1.4.2, scikit-learn 0.24.2, and Pandas 1.3.0.

6.2. Dataset Description

The dataset for this project was gathered from kaggle and contained information on 10th, 12th, and college graduates. The dataset included 10,000 samples and three 15-question questionnaires with both numerical and qualitative characteristics. The dataset was preprocessed by utilising mean imputation to handle missing values and one-hot encoding to transform categorical variables into numerical representations.

6.3. Experimental Design

Using a 70-15-15 split, the dataset was divided into testing, validation, and training subsets. To guarantee that the arrangement of classes remained uniform across all subgroups, stratified sampling was used. For reproducibility, a seed of 42 was chosen at random. Each model was separately trained on the training group, hyperparameter-tuned on the validation group, and tested on the testing group.

6.4. Model Architecture

Three domain selection models were trained using XGBoost, a method of ensemble learning based on gradient boosting. Each model was given a unique set of hyperparameters, such as the number of trees, maximum depth, learning rate, subsample ratio, and column subsampling rate. Model 1 was trained using the default hyperparameters, whereas Models 2 and 3 were tweaked using grid-based search and random search, respectively.

Hyperparameters and Training Configuration

The hyperparameters for each model were set as follows:

Model 1 : Class 10

Number of trees: 100
 Maximum depth: 3
 Learning rate: 0.1
 Subsample ratio: 1.0
 Column subsampling rate: 1.0

Model 2: Class 12

Number of trees: 150
 Maximum depth: 5
 Learning rate: 0.05
 Subsample ratio: 0.8
 Column subsampling rate: 0.8

Model 3: Graduates

Number of trees: 200
 Maximum depth: 4
 Learning rate: 0.2
 Subsample ratio: 0.9
 Column subsampling rate: 0.6

To avoid overfitting, every model was trained with early halting and patience over a period of 10 rounds. During training, the multi-class logarithm loss was used as an assessment metric.

6.5. Evaluation Metrics

Each model's performance was assessed utilising the four metrics: precision, recall, accuracy, and F1 score. These metrics provide information about the models' overall performance in classification as well as their capacity to correctly distinguish positive and negative data across multiple domains.

Table 1: Performance of all 3 models

Model Name	Accuracy
Class 10th Grade	0.872
Class 12th Grade	0.914
College Graduates	0.931

6.6. Why we chose XGBoost for all three models:

Versatility and Performance: XGBoost is notable for its adaptability and good performance across a wide range of applications, notably problems with regression and classification. Because of its capacity to handle complicated datasets effectively, it has acquired prominence in machine learning contests and real-world applications.

Handling Nonlinearity: XGBoost is ideal for detecting nonlinear relationships and feature interactions. Because career counselling involves complex patterns and relationships between several parameters, XGBoost's ability to simulate nonlinearities can help to make more accurate forecasts and suggestions.

Feature Importance: XGBoost includes feature importance analysis, which allows you to find the most relevant aspects in the process of making choices. This can provide helpful insights into why particular career pathways or academic streams are advised, increasing the models' transparency and interpretability.

Efficiency and Scalability: XGBoost is meant to efficiently handle huge datasets. Because your career advice AI project may entail large volumes of data, the scalability of XGBoost guarantees the models can withstand rising computing needs without sacrificing efficiency.

REFERENCES

- [1] C.M. Veneri, Here Today, Jobs of Tomorrow: Opportunities in Information Technology. Bureau of Labor Statistics, 1998.
- [2] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," Knowledge-Based Systems, vol.46, pp.109–132, Jul.2013,
- [3] <https://doi.org/10.1016/j.knosys.2013.03.012>.
- [4] L. Csorny, "Careers in the Growing Field of Information Technology Services," Beyond the Numbers, vol. 2, no. 9, Apr. 2013.
- [5] L. Stewart, "Looking for Jobs? Look to IT in 2012 and Beyond," Information Technology and Innovation Foundation, 2012.
- [6] D.Richards and D. Terkanian, "Occupational employment projections to 2022," Monthly Labor Review, U.S.

- [8] Bureau of Labor Statistics, Dec. 2013, <https://doi.org/10.21916/mlr.2013.41>.
- [9] S. K. White, "The 10 most in-demand tech jobs for 2020 — and how to hire for them," CIO, Jan. 10, 2020.
- [10] <https://www.cio.com/article/3235944/hiring-the-most-in-demand-techjobs-for-2018.html> (accessed Dec. 12, 2020).
- [11] ICT Report: ICT Workforce in the Kingdom of Saudi Arabia. Riyadh, Saudi Arabia: CITC Publications, 2015.
- [12] H. Ma, I. King, and M. R. Lyu, "Effective missing data prediction for collaborative filtering," in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, Jul. 2007, pp. 39–46, <https://doi.org/10.1145/1277741.1277751>.
- [14] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in Proceedings of the 1994 ACM conference on Computer supported cooperative work, New York, NY, USA, Oct. 1994, pp. 175–186 <https://doi.org/10.1145/192844.192905>.
- [15] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, San Francisco, CA, USA, Jul. 1998, pp.43–52 .

General Purpose Artificial Intelligence System with Graphical User Interface

¹Deepak Sarangi, ²Dr. Shaifali M. Arora, ³Dr Poonam Dahiya, ⁴Dr Anshul pareek

¹Student, Department of ECE, MSIT, New Delhi, India

²Associate Professor, Department of ECE, MSIT, New Delhi, India

^{3,4}Assistant Professor, Department of CSE, MSIT, New Delhi, India

Abstract-- In this work, the combination of different individual technologies have been tested to give a common product which can give efficient results compared to their individual performance. AI (Artificial Intelligence), GUI (Graphical User Interface), Speech Recognition, Face Recognition, Recommendation System are the technologies which are combined together to give a common product.

Keywords: AI (Artificial Intelligence), GUI (Graphical User Interface), speech recognition, face recognition, recommendation.

1. INTRODUCTION

The history of Artificial Intelligence (AI) starts with antiquities, myths, rumors and various stories of artificial beings enhanced with intelligence or consciousness by the master craftsmen. The seeds of modern AI were planted by philosophers who tried to describe the process of human thinking which can be converted to the mechanical manipulation of symbols. The research field of AI technology was founded at a workshop organized in the campus of Dartmouth College, USA during the summer of 1956. Those who have attended the workshop would become the leaders of AI research for decades. Many of them have predicted that a machine with intelligence like a human being would exist in the world within a generation and millions of dollars were spent to convert this vision into reality. Eventually, it became evident that the difficulties of the project had grossly underestimated by researchers and commercial developers.

In 1974, to reduce the criticism given by James Lighthill and ongoing pressure from congress, the U.S. and British Governments have to stop funding to undirected researches in the field of artificial intelligence and later these difficult years that are followed were called as "AI winter". Seven years later, a visionary initiative by the Japanese Government inspires governments and industries to work on AI again.

In 21st Century, investment and interest in AI is a big boom as machine learning can be successfully applied to solve many problems in academia and industries due to use of new methods, the application of powerful computer hardware and the collection of immense datasets. AI can be used in many ways like security system, recommendations for a result according to our need (prediction), classification of different items and these are also examples of ML (Machine Learning) which has a very combination with AI system.

Machine Learning is a method to train and test our machine or system to give better results by time and whenever it implements the code its accuracy increased. It is one of the growing technologies in the current world of IT-sector.

2. LITERATURE REVIEW

The work is done on different technologies individually and at the end. All the technologies are combined by connecting each of them and incorporate to a common system. The different technologies used in this are given below:

1. Graphical User Interface (GUI)
2. Speech Recognition
3. Artificial Intelligence (AI) System
4. Face Recognition
5. Recommendation using datasets.

2.1. Graphical User Interface (GUI)

Graphical User Interface (GUI) is a type of user interface which includes graphics like images, videos, texts in the window which can be seen by the user when they start the program [1]. Example of Graphical User Interface is given below:



Fig. 1 Image of a GUI Background.

GUI is very useful to make AI system more interactive and impressive as it gives a good real-time experience to the user. It improves the quality of the AI System by making the system more attractive and show beautifulness of the system and more details of the product can be shown in a very easier way.

Python offers multiple options for developing GUI. Some of these are given below:

1. Tkinter
2. Kivy
3. Python QT
4. wxPython

Out of all the GUI methods, "tkinter" is the most commonly used method [2]. It is the first option for a lot of learners and developers because it is quick and convenient to use.

2.2. Speech Recognition

Speech Recognition System [3] is the system which takes audio input and by the help of speech recognition module it converts that audio input into text and the text is further processed according to the instructions to give final output to the user.

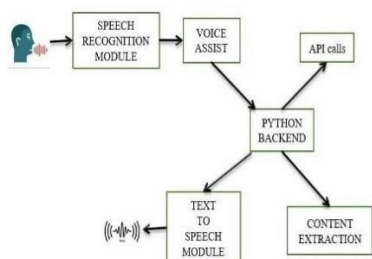


Fig. 2 Basic block diagram of Speech Recognition Process.

This is a basic structure for the speech recognition program where we start with the command by user then speech recognition module comes into action then the program in the python triggered then it will do various works whether to call any API or speak something or extract some data from somewhere in the browser or in the local system, etc. It consists of various parts like the modules i.e. speech recognition & text-to-speech module, and the code in python.

Speech recognition basic step is to convert speech to an electrical signal with a microphone and then convert it to digital data. Once the digitalization process is completed several models can be used to transcript the audio data to text data.

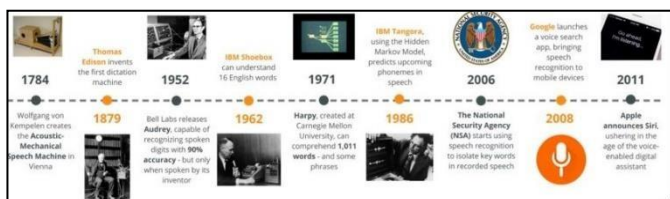


Fig. 3 Evolution of Speech Recognition.

If we look back at history systems were limited to single speakers and had limited vocabs but as technological advancements came into the picture. Now, the modern speech recognition system has the vocabulary for a plethora of languages.

The modern speech recognition systems like Alexa, Siri depends on Hidden Markov Model (HMM) approach. This approach works on the assumption that a speech signal when viewed through a brief enough timescale (say, ten milliseconds), is often reasonably approximated as a stationary process—that is, a process in which statistical properties do not change over time.

Fortunately, as a Python programmer, we don't need to worry much (unnecessarily) about the whole process which is being performed. There are various speech recognition services which are available for online users through an API, and many of these services also offer Python SDKs.

There are many packages available for speech recognition exist on PyPI. A few of them include:

- Speech Recognition
- google-cloud-speech
- watson-developer-cloud
- pocket sphinx
- apiai
- CMU Sphinx

- assemblyai
- wit

The most commonly used and best package which can help in the speech recognition process is Speech Recognition.

2.3. Artificial Intelligence (AI) System

Artificial Intelligence System [4] is the main content of this work as it will do the tasks for the user whether opening a file in the system or opening a site on the browser or anything else which user needs. It will perform several tasks based on the input given by the user using microphone or speakers which will later be processed by the speech recognition system and converted into text. Later, this text is used to perform tasks based on the comparison with backend program and commands which are provided by the developers.

The basic idea is to use the converted text and compare it the provided commands if it is matched the task associated with the condition of matching is performed otherwise check for others. If there is no match then it will ask to provide input again and the same procedure is executed until the user doesn't want to end it.

2.4. Face Recognition

Face recognition is the technology which is used for the identification of a human being's identity using one's individual face. Such type of systems can be used in processing of photos, videos, or in real time machines[5]. By using such technology, one can easily detect and identify the face by taking the help of datasets by getting similar matching appearance of an individual. The most efficient way to detect the face of the person is the method to use python and OpenCV in deep learning. This method is very useful in many fields such as for security, the military, universities, schools, colleges and airlines, banking, online web applications, gaming etc. this system uses powerful python algorithm through which the detection and recognition of face is very easy and efficient.

The module used to perform this task is OpenCV library in which CV stands for Computer Vision. Computer Vision is the field of technology used by the computers to process image-based processes. It is one of the popular technologies in today's growing world.

In this work a file (training program file) is used to compare pixels of the image in order to give desired results to the users. LBP (Local Binary Patterns) and LBPH (Local Binary Pattern Histogram) are key factors in this procedure.

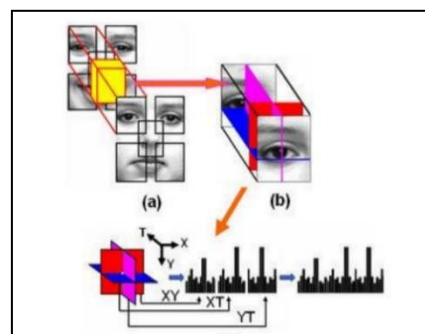


Fig. 4 Description of facial expressions with local binary patterns.

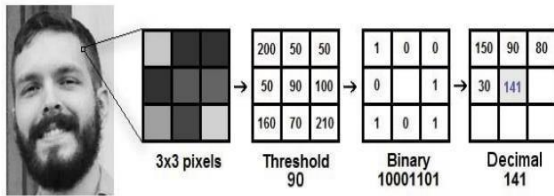


Fig. 5 LBPH Algorithm for Face Recognition.

2.5. Recommendation System

Recommendation System is a system which can be used to predict some results based on the dataset provided to the system. It will use those datasets and return the closest answer based on the cut-off criteria for the requirement of the user. It is based on regression-based machine learning technique used by many tech giants like YouTube, Google, social media platforms (Facebook, Twitter, Instagram, LinkedIn, etc.)[6]. In this, data is being filtered out from the dataset and return to the user and all these processes occur behind the scenes (backend).

3. MATERIALS AND METHODS

In this project, firstly we need to setup a working environment so we need to use an IDE (Integrated Development Environment)[7,8] which is used as a workplace. There are many IDEs like Visual Studio Code (VS Code), PyCharm, Conda, etc. VS Code is very useful and easier to use.

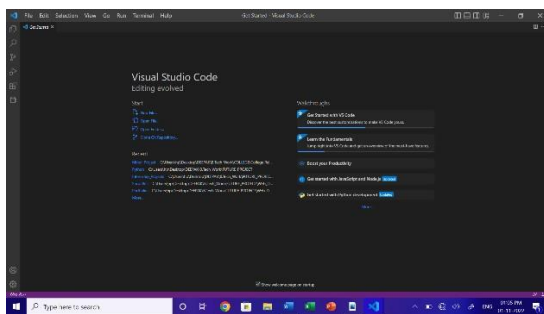


Fig. 6 Visual Studio Code.

After we need to install some modules and packages in order to use different functionalities in the project. Some of these are given below:

1. Python latest version
2. Speech Recognition
3. tkinter (In built in Python)
4. pyttsx (Python Text-to-Speech Library)

4. RESULTS AND CONCLUSIONS

After working on this project, we get to know more about the Artificial Intelligence which can be used for performing daily tasks that are performed by everyone on their computer or laptop system. In addition, Graphical User Interface gives it a better texture in terms of interaction between the system and the user. It gives the user a real-time experience to explore a system which make their work done only by the voice commands and that's very fascinating for everyone.

For future purposes, this project needs to be more advanced and better performing so to cop up with the changes in the technologies in this growing world. There are some snapshots which are given below to show the project setup:

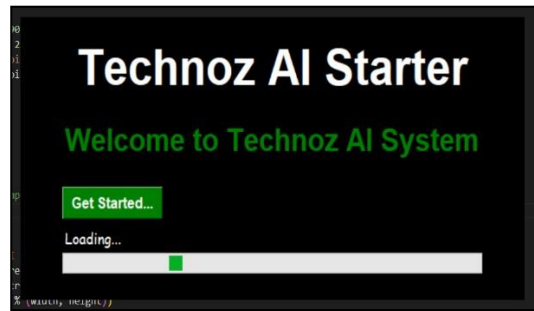


Fig. 7 Splash Screen Window.

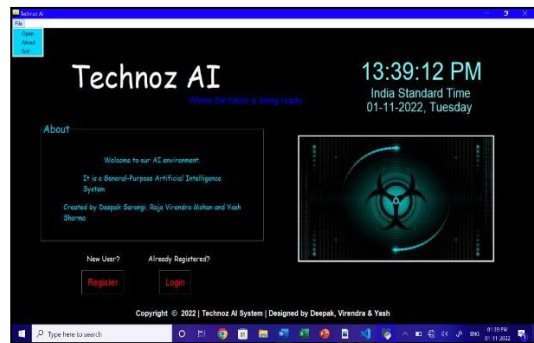


Fig. 8 Main Page Window.

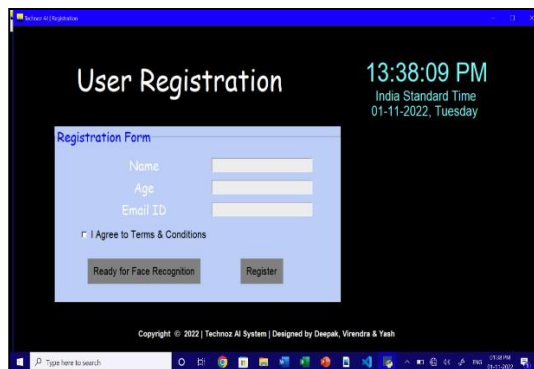


Fig. 9 User Registration Window.



Fig. 10 Home Page Window.

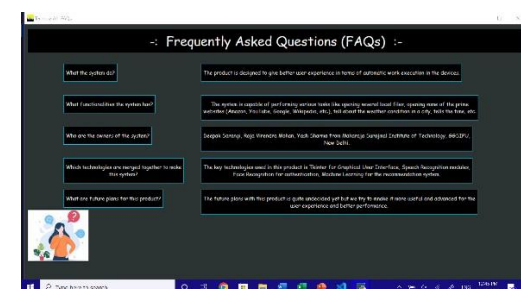


Fig. 11 Frequently Asked Questions Page Window.

REFERENCES

- [1] Chang, T.H., Yeh, T. and Miller, R.C., 2010, April. GUI testing using computer vision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1535-1544).
- [2] Lundh, Fredrik. "An introduction to tkinter." URL: www.pythonware.com/library/tkinter/introduction/index.htm (1999).
- [3] Gaikwad, S.K., Gawali, B.W. and Yannawar, P., 2010. A review on speech recognition technique. *International Journal of Computer Applications*, 10(3), pp.16-24.
- [4] Jin, Cheng, Weixiang Chen, Yukun Cao, Zhanwei Xu, Zimeng Tan, Xin Zhang, Lei Deng et al., 2020, "Development and evaluation of an artificial intelligence system for COVID-19 diagnosis." *Nature communications* 11, no. 1 : 5088.
- [5] Barnouti, N.H., Al-Dabbagh, S.S.M. and Matti, W.E., 2016. Face recognition: A literature review. *International Journal of Applied Information Systems*, 11(4), pp.21-31.
- [6] Das, D., Sahoo, L. and Datta, S., 2017. A survey on recommendation system. *International Journal of Computer Applications*, 160(7).
- [7] Tejashree Dhawle¹, Urvashi Ukey², Rakshandha Choudante (Oct, 2020), "Face Detection and Recognition using OpenCV and Python", Page 1269-1271. Internet:
- [8] Khan, M., Chakraborty, S., Astya, R. and Khepra, S., 2019, October. Face detection and recognition using OpenCV. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 116-119). IEEE.

Enhancing Student Feedback Analysis: A hybrid Approach utilizing Machine Learning and Lexicon-Based Methods for Sentiment-Analysis

¹Vimal Gaur, ²Shivam, ³Rishabh Bhatt, ⁴Shivansh Tripathi

¹Associate Professor, CSE Dept., Maharaja Surajmal Institute of Technology, C-4, JanakPuri, New Delhi – 110058, India

^{2,3,4}Student, CSE Dept., Maharaja Surajmal Institute of Technology, C-4, JanakPuri, New Delhi – 110058, India

²birla.shivam.666@gmail.com, ³rishabh1999bhatt@gmail.com, ⁴shivanshtripathi25@gmail.com

Abstract. Sentiment Classification or Opinion Analysis is the way to categorize data into pieces with a spectrum of applications, according to their polarity, across many companies and industries. This paper looks at a combination of two important approaches to classification of sentiments using a glossary of weighted words and machine learning responses respectively. These approaches are compared with the blended hybrid approach to report on their weaknesses and strengths. When analyzing a set of student feedback, the results show that the hybrid model works better than the lexical approach, in turn being outclassed by the learning approach. Also, the benefit of convenience resulting from the elimination of the purpose for data training makes the hybrid model on one alternative attractive to other application with a little bit trade-off in its performance.

Keywords: AFINN, Bag of Words (BOW), Lexicon, Machine Learning, Multinomial Naïve Bayes, Stemming Classifier.

1. INTRODUCTION

Classification makes it possible to automatically label data into different categories according to certain characteristics of data. In sentiment binary classification, the data is classified as either negative or positive based on its implication. There are two important primary procedure to classify sentiments according to a machine learning & sentiment-lexicon respectively. In brief, using the lexicon- approach, the text to be analyzed is symbolized in individual words whose polarity is sought in a glossary of lexicon sentiments [1]. The polarity of a word describes its cardinal implication, i.e., how negative or positive a phrase is. The text is then categorized according to the sum of the values of polarity of all phrases [2]. On other hand, the learning procedure do its job by forming a ML algorithm on a training package made up of data that has been labelled as negative or positive prior to run the model classifier. This collection of data is then examined for patterns [3,4] for indicating which combinations of words tend to be present in positive and negative texts, thus the creation of a model capable of predicting the sentiment of what was not yet seen. This paper has been divided into sections. Section 2 represents a literature review that describes some of the works previously done by various authors. Section 3 represents the proposed model. Section 4 illustrates the results and discussions. Section 5 represents the conclusion.

2. LITERATURE REVIEW

This section presents sentiment analysis using different methodologies.

Table 1. Different Methodologies on sentiment analysis

S. No.	Author	Year	Methodology
1.	Z.Nanli et al. [1]	Nov, 2012	This research put forward an all-inclusive valuation conducted in three fields to inspect the sentiments i.e. sentiment analysis, feature extraction and framework.
2.	F. Xing et al. [2]	Dec, 2015	Researcher focused on issues faced in polarity sentiment classification. The Input dataset includes Amazon product reviews. Further he proposed a new process for classifying sentiment at two levels viz. review and sentence level
3.	A.P. Jain et al. [3]	July, 2016	This research elucidates the sentimental analysis by employing most commonly used algorithms viz. Decision Tree and Multinomial NB algorithms. On application of these algorithms to Twitter data, author concluded that Decision Tree performs superior when parameters are tuned properly.

4.	M. A. Ullah [4]	Dec, 2016	This article uncovers and addresses the model for processing these responses by using ML classifiers such as NB SVM and Max Entropy (ME).	8.	Lighthart et al. [8]	Mar, 2021	This study yields the findings of a related study on methods of sentiment analysis aimed at highlighting the features adopted as input & output followed by a study that follows the conduct of literature systematic reviews CNN, LSTM and SLR.
5.	B. K. Bhavitha et al. [5]	Mar, 2017	In this research, it elucidates that Naïve Bayes (NB) method of learning explains better if smaller quantity of set of data is taken and Support Vector explains better if larger quantity of data set is taken.	9.	S. Bashir et al. [9]	Jul, 2021	Author employed different methods for processing the complete dataset and hence converting it into a form understandable by machine. These methods include BERT and word2vec and hence forecast the polarity of subjective phrases.
6.	Z. Nasim et al. [6]	July, 2017	This approach calculates only the aggregate feelings of responses that are supplied by the students.				
7.	M. Baygin [7]	Sep, 2018	In this paper, Naïve Bayes (NB) perspective has been made use of for document and text categorization. For categorizing, 1050 documents were taken into account and for the extraction of the features, N-gram method is taken into consideration.				

3. METHODOLOGY

Python as programming language has been adopted because of its wide adaptation in the scientific world and the available libraries that exist as a result. Most prominently, the framework of machine learning *Sci-Kit Learn* offers accomplishment of SVM, NB algorithms and methods including *BoW* [5,6] and *TF-IDF* [7,8]. Lexicon-Based approach considers three commonly used lexicons, viz., *AFINN*, *SentiWordNet* and *VADER*.

3.1 MACHINE LEARNING APPROACH

Training Set is accumulated using Google Forms in Two Phases: -

- Initial data gathering is done for University's different attributes-
 - Guidance
 - Course Content
 - Examination Lab Work
 - Library Facilities
 - And, Extra Curriculars

- Thereafter, attributes for different Teachers are gathered: -

- Teacher 1
- Teacher 2, etc.

There has been work on model building on different classifiers to find the most suitable efficient accuracy. The models on which the work has been done: -

- Multinomial Naïve Bayes (MNB) Classifier
- Multinomial NB Stemmed Classifier
- Support Vector Machine (SVM) Classifier
- SVM Stemmed Classifier

3.2 LEXICON-BASED APPROACH

In Lexicon-based Sentiment Analysis, input includes a set of dictionaries of words each classified as negative, positive and neutral feedback along with their parts of speech, polarity, and subjectivity classifiers.

The three most popular Lexicons for Sentiment Analysis employed in this paper are:-

- AFINN Lexicon

AFINN Lexicon is the fundamental and prominent lexicons for opinion analysis. The current version is AFINN-en-164.txt and it contains 3379 words along with their score of polarity.

2. SentiWordNet

SentiWordNet is a resource lexical for feedback analysis. SentiWordNet allocate three sentiment scores to each WordNet synset: - negativity, objectivity, positivity.

3. VADER

VADER stands for Valence Aware Dictionary and Sentiment Reasoner is a rule-based and lexicon opinion analysis apparatus. It is accessible in the Natural Language Toolkit Package (NLTK) that can be pertained to unclassified text data. VADER is capable of detecting of intensity and polarity of emotional sentiment.

3.3 HYBRID APPROACH (ML WITH LEXICON-BASED)

In uniting the two approaches, the classifier based on -lexicon is executed on the test set, yielding classified output. From previous steps, the output is passed into the, well-performed ML learning-based classifier for purpose of training. Subsequently, learning-based hybrid classifier, the lexicon trained, was executed on the test set.

- SVM on BOW (Bag of Words)
- SVM on TF-IDF (Term Frequency-Inverse Document Frequency).

4. RESULTS AND DISCUSSIONS

A correlative analysis is conducted with the help of four ML models: - Support Vector Machine, Multinomial Naïve Bayes and their Stemmed classifiers. Testing data set is used to inspect the performance of model learned. Here, we can see the performance in terms of accuracy of each of the classifier used.

5. LIMITATIONS

- The achievement of the model based on learning grader is mainly dependent on the condition of the data of training, which, according to the hybrid model, is derived, through the model based on lexicon.
- Additional improvements could be made in the field of natural language processing [9], thus making language of elusive in nature concepts such as sarcasm and irony, which affect the implication of a given sentence, would be correctly recognised. Without a doubt, such progress would result in improved classification performance.

6. CONCLUSION

The primary research question focused on evaluating the hybrid model's performance against other techniques. Our findings indicate that the hybrid approach demonstrates a notable enhancement over the lexicon-based method, as discussed in earlier sections. When compared to the top-performing purely learning-based method, the suggested hybrid approach performs nearly as well, albeit with the need for pre-labeled data. Therefore, for applications prioritizing convenience over slight performance differences, the hybrid solution emerges as the preferred choice.

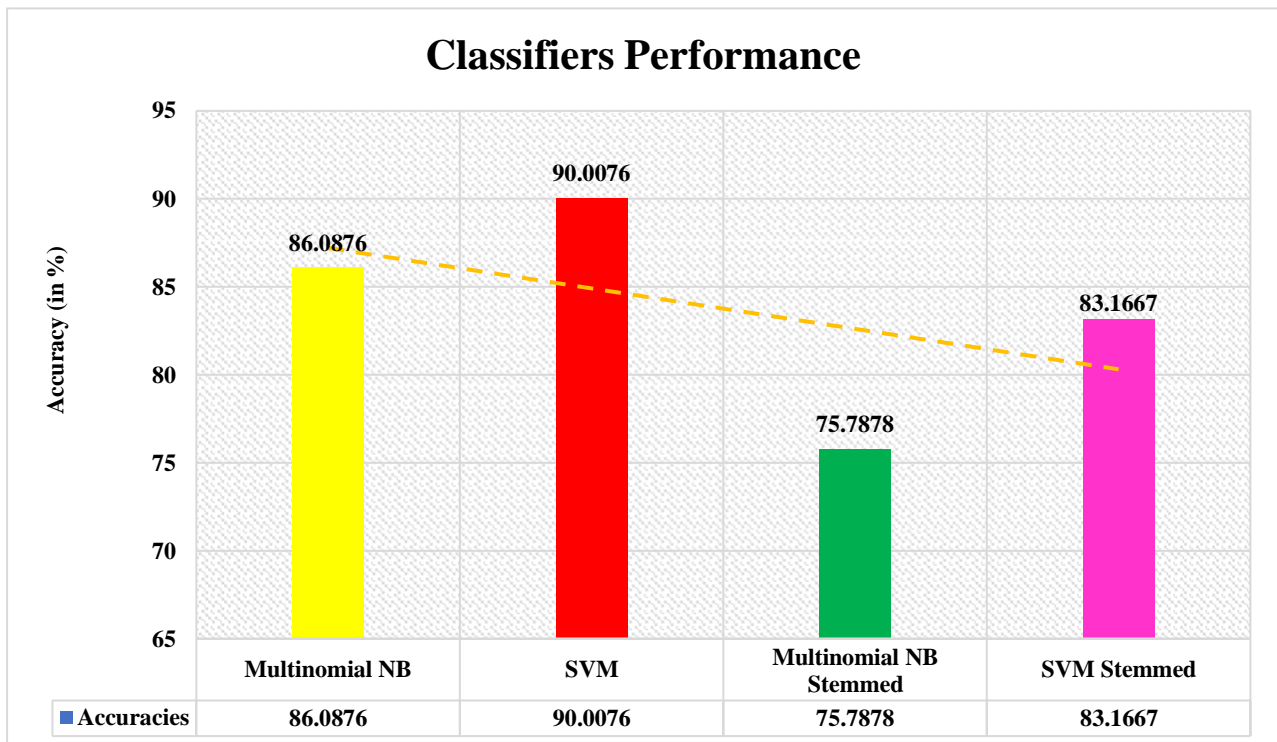


Fig.1.Comparison among ML Classifiers

And further result shown below is of Lexicon-integrated ML classifiers:

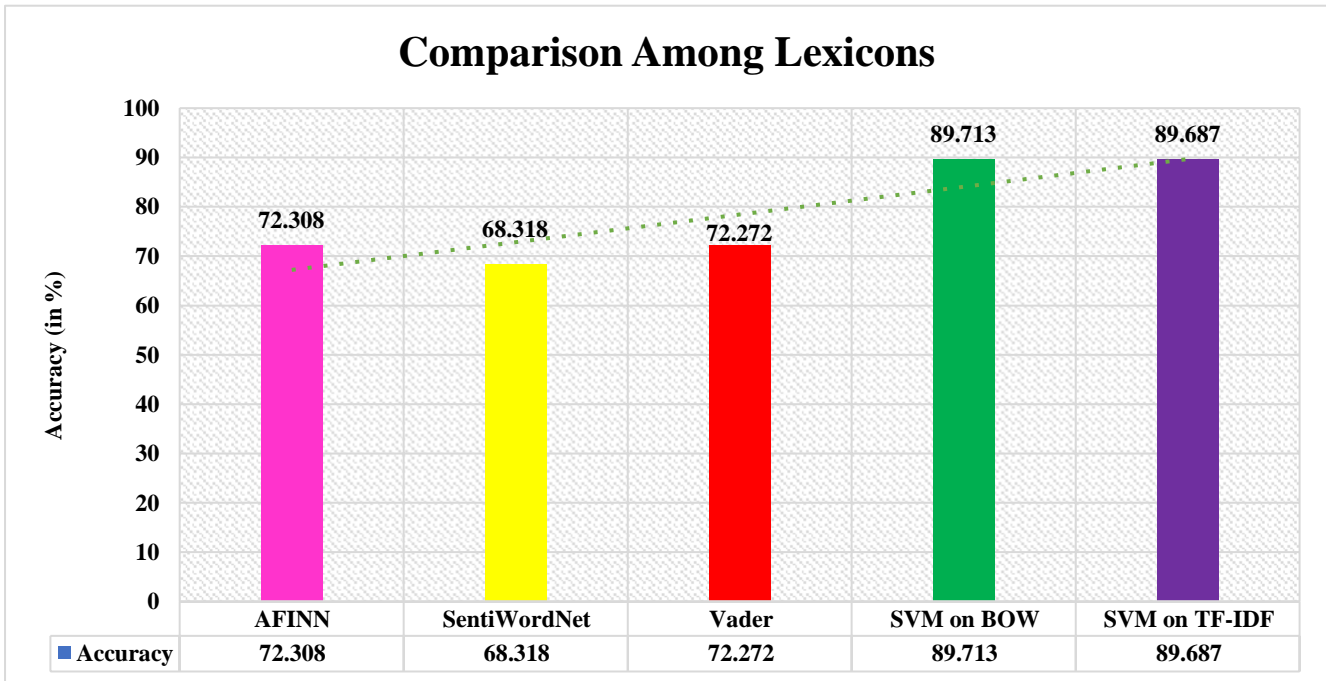


Fig.2.Comparison among Lexicon with ML Classifiers

REFERENCES

- [1] Z. Nanli, Z. Ping, L. Weiguo and C. Meng, "Sentiment analysis: A literature review," 2012 International Symposium on Management of Technology (ISMOT), 2012, pp. 572-576, doi: 10.1109/ISMOT.2012.6679538.
- [2] Fang, Xing & Zhan, Justin. (2015). Sentiment analysis using product review data. J Big Data. 2. 10.1186/s40537-015-0015-2.A.P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2016, pp. 628-632, doi: 10.1109/ICATCCT.2016.7912076.
- [3] M. A. Ullah, "Sentiment analysis of students feedback: A study towards optimal tools," 2016 International Workshop on Computational Intelligence (IWCI), 2016, pp. 175-180, doi: 10.1109/IWCI.2016.7860361.
- [4] B. K. Bhavitha, A. P. Rodrigues and N. N. Chiplunkar, "Comparative study of machine learning techniques in sentimental analysis," 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), 2017, pp. 216-221, doi:10.1109/ICICCT.2017.7975191.
- [5] Z. Nasim, Q. Rajput and S. Haider, "Sentiment analysis of student feedback using machine learning and lexicon-based approaches," 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), 2017, pp. 1-6, doi: 10.1109/ICRIIS.2017.8002475.
- [6] M. BAYGIN, "Classification of Text Documents based on Naive Bayes using N-Gram Features," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1-5, doi: 10.1109/IDAP.2018.8620853.
- [7] Lighthart, A., Catal, C. & Tekinerdogan, B. Systematic reviews in sentiment analysis: a tertiary study. Artif Intell Rev 54, 4997–5053 (2021). <https://doi.org/10.1007/s10462-021-09973-3>
- [9] Shameer Bashir and Arvind Selwal(2021). A Comprehensive Survey of Sentiment Analysis: Word Embeddings Approach, Research Challenges and Opportunities. 2nd International Conference on IoT Based Control Networks and Intelligent System (ICICNIS2021). <http://dx.doi.org/10>

Malicious URL Detection on IOT

Sangeeta¹, Sitender², Savita Ahlawat³, Karthik Jangid⁴, Jayashy Sinha⁵, Gitali Daryani⁶

Department of Computer Science and Engineering^{1,3,4,5,6}, Department of Information Technology²,

Maharaja Surajmal Institute of Technology New Delhi, India

¹Sangeeta@msit.in, ²sitender@msit.in, ³savita.ahlawat@msit.in,

⁴karthik_jangid@outlook.com, ⁵jayashysinha@gmail.com, ⁶gdaryani0162@gmail.com

Abstract-One of the most widely recognized network safety dangers is Malicious Uniform Resource Locator. Malicious URL is a connection made fully intent on advancing scams, attacks, and frauds. By tapping on an infected URL, you can download ransomware, virus, trojan, or some other kind of malware that will think twice about a machine or even your network, on account of an organization. This project endeavors to identify the Malicious URL and safeguard the system against malware and phishing attacks accordingly expanding the security of the system.

This project is based on a dataset that includes fundamental highlights of IOT i.e. Internet of things. Such duplication of IoT devices have brought extended security shortcomings making ordinary clients as casualties of different types of safety attacks via malicious URLs since a part of the devices in a typical IoT sector relying upon the URL.

1. INTRODUCTION

Importance of the World Wide Web (WWW) has drawn in expanding consideration due to the development and advancement of communication, web-based banking, and online business. These days, on the Web, a huge number of sites are usually alluded to as malicious sites. It was noticed that the technical growth made a few techniques attack and scam clients like spam SMS in social networks, web-based betting, phishing, monetary fraud, duplicitous award winning. Recently, a good number of attacking techniques were implemented by broadcasting compromised URLs, and malicious addresses are the main strategies utilized by attackers to achieve malicious exercises. Spam refer to as shipped off spontaneous messages forcibly for publicizing, what we don't ask for and don't have any desire to get. Finally, Some attacks guide clients to uncover private sand confidential data by behaving as real website pages. The hackers make duplicates of the famous web page utilized by clients, for example, Facebook and Google give way to victim PCs by setting different bits of malicious code in HTML code. Dark side of the Web has drawn in expanding consideration of the world. Internet security systems can't necessarily identify malware and downloads containing viruses. Researchers have examined how to assemble compelling answers for identifying Maliciousness URL. The most known method is the blacklist method. This method utilizes details of already known URLs that are affected to channel the new data. Machine Learning strategies are utilized for a bunch of URLs at the time of training the data and become familiar with a forecast classifier to group whether a URL is malicious or not. With the help of this procedure, it permits them to add up to upcoming URLs, dissimilar to blacklisting methods. Before long, these explanations should be utilized in the Cyber-Physical System.

Generally, IOT alludes to the network of physical items or "things" inserted with sensors, software, and network

capabilities, permitting them to gather and trade information over the web. These items can incorporate different devices, for example, cell phones, wearable devices, home appliances, industrial machinery, and even vehicles. However, in recognizing malicious URL IoT devices produce tremendous measures of data, which can be examined to distinguish patterns and peculiarities that show the presence of malicious URLs. By utilizing Machiner Learning and data analytics strategies, IoT system can distinguish unusual or suspicious URL behavior and trigger suitable activities to mitigate the threat. IoT devices can shape cooperative networks where they share data about known malicious URLs or indicators of give and take. Additionally, IoT devices can coordinate with existing security platform and services, for example, cloud-based security arrangements or threat intelligence platforms. This will consistently screen and examine URL threats and give constant updates and assurance to the IoT network. IoT expects to upgrade productivity, further develop decision making, robotize processes, and set out new open doors for advancement across different enterprises.

2. LITERATURE SURVEY

In [17] paper, AI calculations are used to characterize URLs based on characteristics and approaches to acting of URLs. The characteristics are eliminated from static and dynamic ways of behaving of URLs and are new to the composition. Those as of late proposed characteristics are the essential responsibility of the research. Machine Learning algorithms are a part of the whole malicious URL identification system. Two supervised Machine Learning algorithms are utilized, SVM and RF.

In paper [19], The presented algorithm did not perform well with the original backbone model and the BlindEye alternative. The AUC number ranged from 67.3% to 90.6%. Collecting data from other categories is much more difficult which results in a comparatively small testing dataset. Hence all scores are statistically significant with p-value < 0.001.

In [18] paper, good accuracy shows that harmless examples are less inclined to be misconstrued as malicious examples. Tests demonstrate the way that the URL can be enough communicated, features that are critical can be removed to detect in a much better way with the help of DCNN. Separating various features from the URL can take advantage of keywords to further develop identification accuracy.

In [20], We can gather that the model utilized achieve high precision accurateness though Random Forest. The exactness of the trained Random Forest classifier may be expanded via training on the data that is comparatively more balanced The analysis helps to establish that malicious URL identification is conceivable via preparing a model utilizing the data with chosen characteristics. This will help in predicting upcoming phishing attacks.

In [26] paper various machine learning methods were examined for the spotting of various attack in IoT organizations. Also deep learning techniques were described for the identification of attacks in IoT. In detection strategies, =different malicious attacks were examined. All learning strategies were analyzed concerning the kinds of attacks, feature choosing strategies, identifying attacks with the help of method(s) and datasets to choose the finest procedures to distinguish the attacks.

In [24] paper, This paper presents a Lightweight dataset to detect phishing in IoT. Initially, a few elements alluded to the length of certain portion of the URL, a few characteristics in accordance with NLP, and a few were estimated. In IoT, where the computational supplies are restricted, it is obligatory to utilize lightweight datasets and algorithms. At the point when Random Forest is utilized with the preferred data representation, the phishing attacks disclosure is 99.57%.

3. PROPOSED WORK

The main objective of this project is:

1. To extract a suitable dataset, analyze the pattern and manipulate the data using latest pre-processing methods.
2. To check whether the URL is malicious or not on the basis of the below mentioned factors:
 - (a) Lexical Parameters
 - (b) Checking header of the IP datagram as parameters for any relation with malicious URL.
 - (c) Random Forest
3. To increase the efficiency of our model, latest boosting techniques will be applied and compared ie. Xgboost,, AdaBoost etc. The technique resulting in maximum efficiency will be considered.

4. RESEARCH METHODOLOGY

The dataset was taken from a dataset of URLs given their status. The dataset will contain separated attributes from sites that can be utilized for Grouping of websites by checking if the website is malicious or not. The dataset will likewise incorporate content consisting of JS code which can be utilized as unstructured data in Deep Learning. The content of the dataset has been gathered by going through the Web with the help of MalCrawler . The names will be checked utilizing the Google Safe Browsing Programming interface. Attributes have been chosen in light of their importance. The characteristics of dataset are mentioned beneath:

- url - The address of the site.
- ip add - page's IP Address.
- geo loc - Exact place where page is introduced.
- url len - Length of URL.
- js len - Length of JS code on website page.
- js obf len - Length of obfuscates JS code.
- tld - The High Level Domain.
- who is - Whether the WHO domain is contend or not.
- https - Whether website consists https or http.
- content - Raw page content containing JavaScript code.
- name - Class label to identify malicious website page.
- hopcount-Number of hops taken by a url to reach at its last destination.

5. Data Analysis

Subsequent to cleaning of data, we want to do Exploratory Data Analysis of our dataset, which basically incorporates checking for relationship among parameters of the dataset.

On the Basis of Different Parameters

● **On the basis of label:** Here we will examine the data based on the Label parameter, that means, whether the given url is Malicious or Benign. As should be visible from the visualizations this dataset has significant class unevenness. Consequently, during any Machine Learning cycle, adequate measures should be attempted to deal with or make up for this unevenness to obtain accurate outcomes.

● **On the basis of Attributes:** There are different parameters in the dataset, that can have a 0 or 1 as boolean value. Below, we have done some analysis on them.

● **On the basis of scheme:** Scheme here implies whether the url prefix is 'http' or 'https'. As seen from visualization, we can see that a lot higher measure of sites using HTTPS protocols are harmless. All in all, The sites that don't give HTTPS protocol are bound to be malicious than the websites that give https protocol.

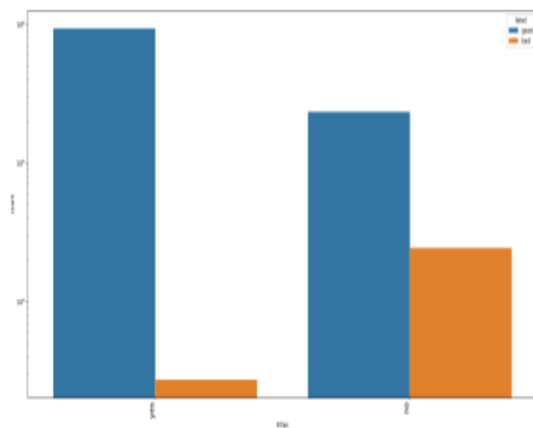


Fig. 1- Univariate analysis of HTTPS attribute

● **On the basis of Geographic Location:** We have looked up the geographic distribution of malicious and benign urls.

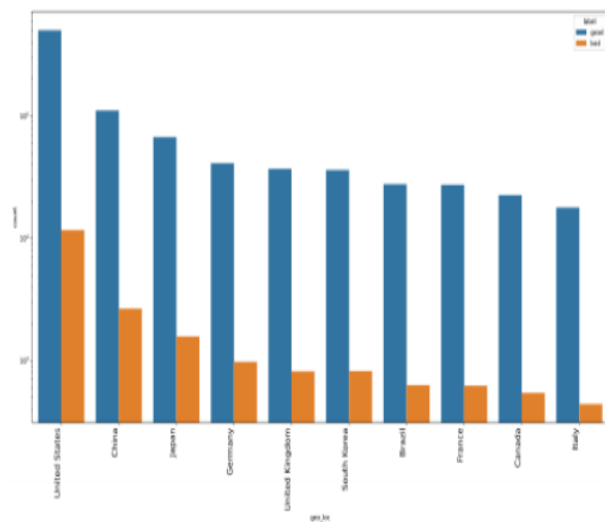


Fig. 2 - Country wise analysis (Top 10 countries)

● **On basis of URL Lengths:** Here, we have imagined the distribution of malicious and benign URLs based on the length URL. Here, we can see that the malicious URLs will in general have higher URL length just barely.

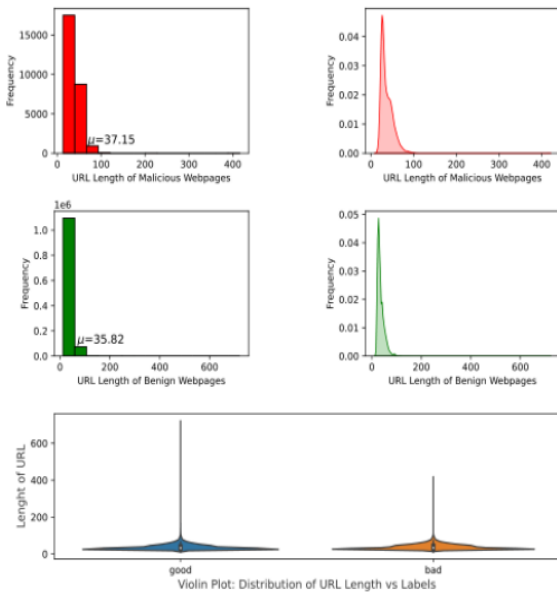


Fig. 3 - Analysis on the basis of URL length

● **On basis of JS Length:** Here, we have imagined the distribution of malicious and benign URLs based on it's length of typical Javascript Code Here, we can see that the m URLs alicious will generally have more javascript content by a huge margin.

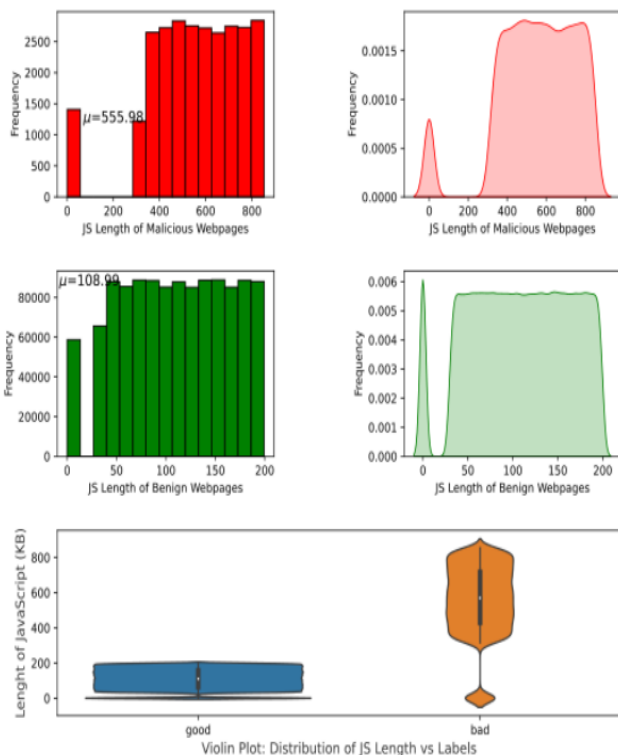


Fig. 4 - Analysis on the basis of length of Javascript Code

● **On basis of content length:** Here, we have imagined the distribution of malicious and benign URLs based on the length of HTML content of the site. Here, we can see that Website pages that are safe, will generally have higher content length than malicious site pages.

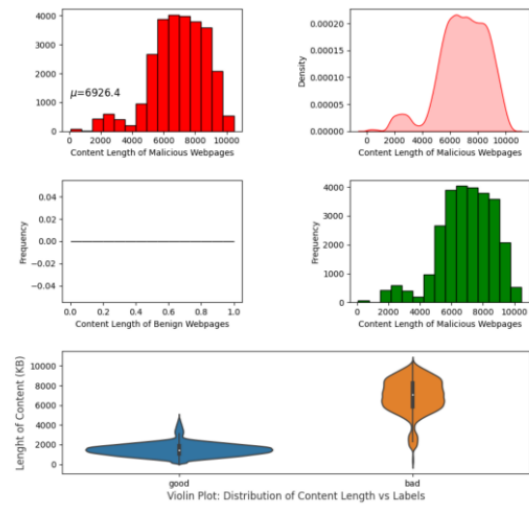


Fig. 5 - Simulation results for the network

● **On the basis of Obfuscated JS Length:** It is the intentional demonstration of making a source that is hard for people to get. Here, we have envisioned the distribution of malicious and benign URLs based on it's length of obfuscated Javascript Code Here, we can see that the benign URLs will in general have negligible obfuscated JS content.

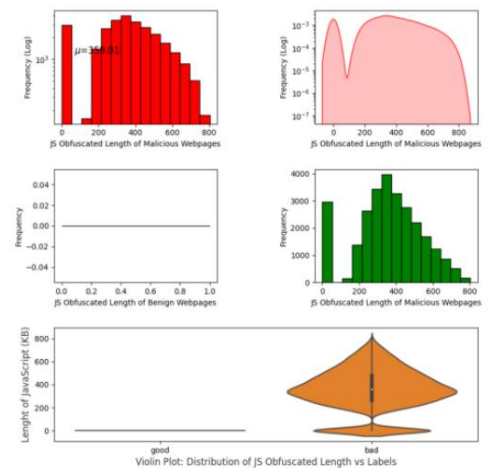


Fig. 6 - Analysis on the basis of length of Obfuscated Javascript Code

6. PROJECT DESIGN DATASET

Here, we'll introduce all the details of dataset. In our paper, the freely accessible dataset of URL was utilized. We'll showcase the subtleties of the chosen dataset utilizing the information perception strategies with the help of exploratory data visualization techniques. Since the quantity of the included features of the dataset is 3231961, the dataset is updated to 9 features utilizing the Principle Component Analysis. Figure 11

depicts the scattered plots of these features. As displayed below, the designs of these elements are distinct from one another. The profoundly connected features in an dataset unfavorably affect the classification performance of the model. In addition, the execution time of the classification algorithms increases due to large number of features in the dataset. Figure 12 depicts the URL malicious dataset's correlation matrix. The correlation values among the features are exceptionally near 0. So, the new dataset will decidedly affect both execution time and classification performance. But there isn't an adequate amount of data in the field of cybersecurity. As of now, there are all the more openly accessible datasets. The URL reputation dataset utilized here was made in 2009.

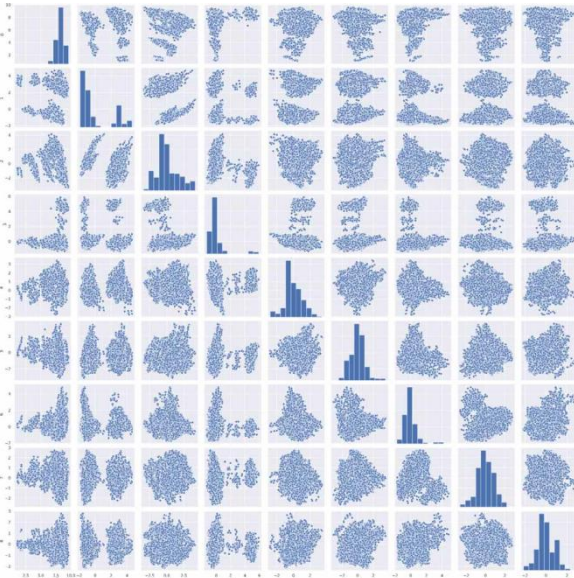


Figure 7. Pairplots of the 9 PCA features

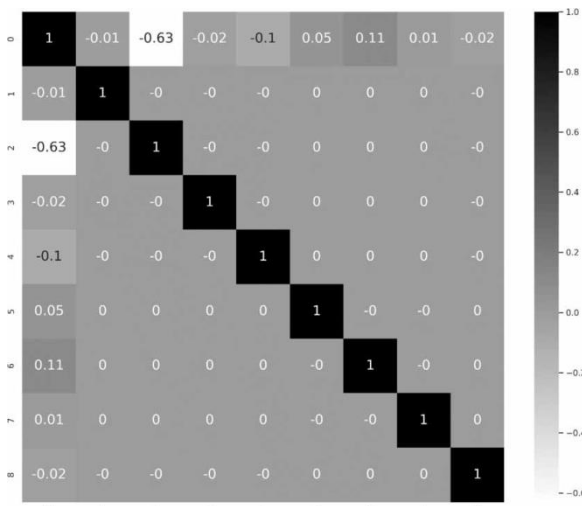


Fig. 8 - Correlation matrix of the URL reputation dataset

DATA PRE-PROCESSING

Raw data is accessible in SVM data format. load_svmlight_file fuction is utilized to peruse data into a sparse matrix obtained from operating system files. The data that is utilized for learning has north of about 3 million features, and it is undeniably great in number than tests of information for training. Also, feature reduction should be finished utilizing the fitting dimensionality diminution method. Methods which are accessible for dimensionality diminution are Truncated SVD and Principal

Component Analysis (PCA). The PCA involves covariance matrix for factorization and it needs working on the matrix as a whole, so it can't be here. Here, Truncated SVD is involved for dimensionality diminution as it can works upon sparse matrices. The kernel of truncated SVD's is failing for n_component more than 32 and so we will proceed with additional analysis on the basis of these elements. Figure 12 demonstrates the clarification of fluctuation by each of these 32 elements. Figure 13 showcases that the mentioned 32 dimensions can catch approximately 46% of the general variation. With the help of LabelEncoder, the labels are encoded into 0 and 1.

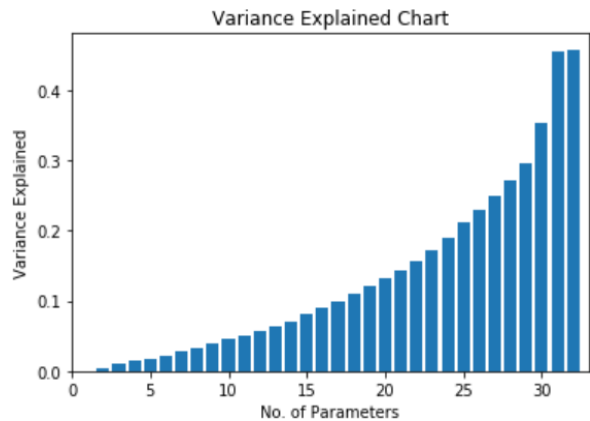


Fig. 9 - Explanation of variance by these 32 features

DATASET SPLIT STRATEGY

At the time of estimating how the model performs, it shouldn't just focus on learning algorithms utilized yet in addition the volume of the included test and training sets, wrong categorization or class distribution assume a significant part. The dataset is separated into two datasets- training and test. Here, the training dataset is utilize for the preparation of those classifier model. The test dataset is likewise used for measuring the model's general performance.

Utilizing the train_test_split function, the resulting data is parted into training and test set. Approximately 30% of the data is saved for the assessment and testing. For assessing the ML models, confusion matrix was utilized and time invested for execution.

EVALUATION

In view of TP, TN, FP and FN metrics, we figure the accuracy and particularity of our presented model. The conditions presenting the metrics are conditions 1, 2, 3 and 4. Accuracy is the criteria involved for classification performance, showcasing the generally right grouping outcome.

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

Precision is defined as the result of estimation for getting a negative when a negative sample is used.

$$\text{Precision} = TP/(TP + FN) \quad (2)$$

As a result of the estimation, the recall is the possibility of getting a positive from a positive sample in actual case.

$$\text{Recall} = TN/(TN + FP) \quad (3)$$

F₁-score depicts the harmonic average of sensitivity and precision values.

$$F_1 = 2 \times (Precision \times Recall)/(Precision + Recall) \quad (4)$$

MODEL TRAINING AND TUNING

Here, Random Forest Classifier is utilized for what it's worth with expected parameters. The Gradient Boosting Classifier consists of a lot more parameters and is a sophisticated algorithm. The first parameter that is taken into consideration is the rate of learning. The algorithm is trained with the help of an alternate number of assessors and how the algorithm is presented. Figure 14 depicts the model attaining correct estimation in gradient boosting.

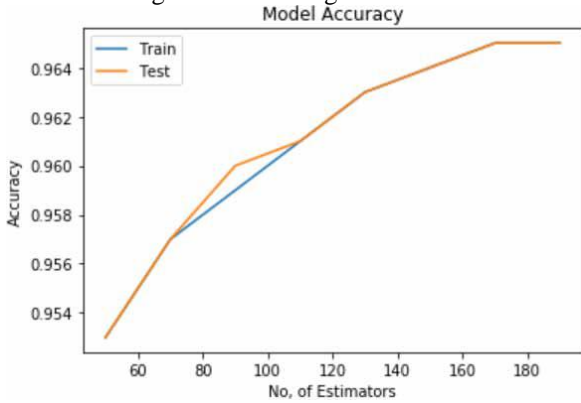


Fig. 10 - Model attaining correct estimation in gradient boosting
 Model levels the accuracy of 96.5% at one hundred seventy assessors. Figure 15 depicts the progressing graph of Gradient Boosting. Model handing out time increments linearly form with some assessors.

Table 1 sums up the time of execution in seconds with various profundities. The accuracy of model increments with the rise in the Max. profundity, however the processing time increments by 2n times.

Table 2 provides the time of execution in seconds with an alternate number of features.

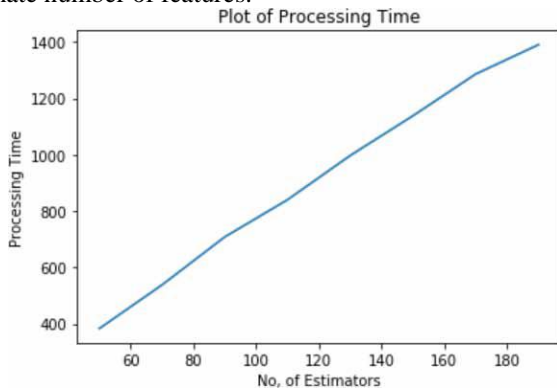


Fig. 11 - Gradient boosting progressing graph

Features	Depth	Test Accuracy	Training Accuracy	Execution Time in Secs
4	4	0.945	0.945	147.0
4	5	0.954	0.954	263.0
4	6	0.958	0.959	488.0
4	7	0.963	0.963	754.0
4	8	0.967	0.967	1169.0

Table 1. Time of execution in seconds with different profundity

Features	Depth	Test Accuracy	Training Accuracy	Execution Time in Secs
4	4	0.964	0.964	604.0
5	4	0.964	0.964	711.0
6	4	0.965	0.965	1286.0
7	4	0.964	0.965	940.0

Table 2. The execution time in seconds with the different number of features

SOLUTIONS AND RECOMMENDATIONS

Figure 16 and Figure 17 demonstrate the feature significance of the dataset for gradient boosting classifier and RF,

individually. The two models primarily understand feature significance and accordingly improve accuracy. Below, 12 elements out of 32 are displayed. The two models mentioned here recognize malicious URLs prior to crawling the page. The Random Forest model plays out the best as compared to all the models tested so far, achieving an accuracy of 98.6% at the time of test data.

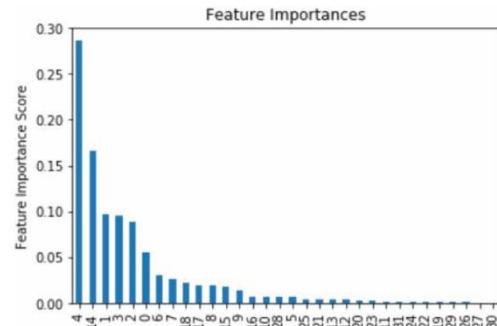


Figure 12. Importance of features

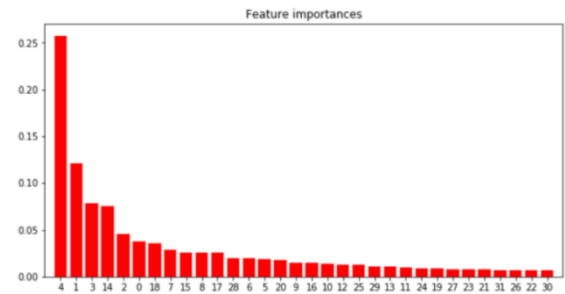


Fig. 13 - Feature importance as per random forest

7. FUTURE SCOPE

As Machine Learning and AI models have become advanced, new security administrations are looked forward to which might change numerous parts of the security issue faced through CPS. Following are the future scope of this undertaking:

- By the utilization of SVM, KNN, and NB AI algorithms can consolidate to analyze the algorithms that focus on processing time and execution of accuracy for identification of URL.
- Foster particular URL analysis methods custom fitted for IoT devices, taking into account their special qualities and correspondence protocols.

8. CONCLUSION

We investigated the functioning of Machine Learning calculations for malware detection with the help of IOT features. We applied Random Forest Machine Learning algorithm and explored different boosting algorithms including Gradient Boosting for detecting the URL. The trial consequences of this strategy shows that the Random Forest Model handles the enormous dataset and checking whether the site is safe or not is altogether really noteworthy (98.6%).

REFERENCES

[1]. Catak, Ferhat Ozgur Şahinbaş, Kevser Dortkardes, Volkan. (2020). Malicious URL Detection Using Machine Learning. 10.4018/978-1-7998-5101-1.ch008.

- [2]. SINGH, AMIT KUMAR (2020), "Dataset of Malicious and Benign Webpages", Mendeley Data, V2, doi: 10.17632/gdx3pkwp47.2
- [3]. https://pandas.pydata.org/pandas-docs/stable/user_guide/scale.html#usechunking
- [4]. <https://en.wikipedia.org/wiki/Obfuscation> (software)
- [5]. <https://cybersecurityguide.org/resources/phishing/>
- [6]. <https://en.wikipedia.org/wiki/Spamming>
- [7]. Catak, Ferhat Ozgur S ahinba s, Kevser Dortkardes, Volkan. (2020). Malicious URL Detection Using Machine Learning. 10.4018/978-1-7998-5101-1.ch008.
- [8]. SINGH, AMIT KUMAR (2020), "Dataset of Malicious and Benign Webpages", Mendeley Data, V2, doi: 10.17632/gdx3pkwp47.2
- [9]. <https://en.wikipedia.org/wiki/Obfuscation> (software)
- [10]. <https://www.brainstobytes.com/precision-vs-recall>
- [11]. Cova, M., Kruegel, C., Vigna, G. (2010). Detection and analysis of drive-by-download attacks and malicious JavaScript code. In Proceedings of the 19th international conference on World wide web (WWW'10) (pp. 281- 290). Raleigh, NC: Association for Computing Machinery. 10.1145/1772690.1772720
- [12]. Do Xuan, Cho, Hoa Dinh Nguyen, and Victor Nikolaevich Tisenko. "Malicious URL detection based on machine learning." *International Journal of Advanced Computer Science and Applications* 11.1 (2020).
- [13]. Kuyama, M., Kakizaki, R. S. (2016). Method for Detecting a Malicious Domain by Using WHOIS and DNS Features. The Third International Conference on Digital Security and Forensics (DigitalSec2016),74- 80
- [14]. <https://www.geeksforgeeks.org/metrics-for-machine-learning-model/>
- [15]. <https://docs.microsoft.com/en-us/azure/machine-learning/concept-deep-learning-vs-machine-learning>
- [16]. Sahoo, Doyen, Chenghao Liu, and Steven CH Hoi. "Malicious URL detection using machine learning: A survey." *arXiv preprint arXiv:1701.07179* (2017).
- [17]. Do Xuan, Cho, Hoa Dinh Nguyen, and Victor Nikolaevich Tisenko. "Malicious URL detection based on machine learning." *International Journal of Advanced Computer Science and Applications* 11.1 (2020).
- [18]. Wang, Zhiqiang, et al. "A malicious URL detection model based on convolutional neural network." *Security and Communication Networks* 2021 (2021): 1-12
- [19]. Tsai, YunDa, et al. "Toward more generalized malicious URL detection models." *arXiv preprint arXiv:2202.10027* (2022).
- [20]. Janet, B., and R. Joshua Arul Kumar. "Malicious URL detection: a comparative study." *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*. IEEE, 2021.
- [21]. Anwar, Sajid, et al. "Countering malicious URLs in internet of things using a knowledge-based approach and a simulated expert." *IEEE Internet of Things Journal* 7.5 (2019): 4497-4504
- [22]. S. Anwar *et al.*, "Countering Malicious URLs in Internet of Things Using a Knowledge-Based Approach and a Simulated Expert," in *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4497-4504, May 2020, doi: 10.1109/JIOT.2019.2954919.
- [23]. Choi, Hyunsang, Bin B. Zhu, and Heejo Lee. "Detecting malicious web links and identifying their attack types." *WebApps* 11.11 (2011): 218.
- [24]. Bustio-Martínez, Lázaro, et al. "A lightweight data representation for phishing URLs detection in IoT environments." *Information Sciences* 603 (2022): 42-59.
- [25]. Luo, C., Tan, Z., Min, G., Gan, J., Shi, W., & Tian, Z. (2020). A novel web attack detection system for internet of things via ensemble classification. *IEEE Transactions on Industrial Informatics*, 17(8), 5810-5818.
- [26]. Inayat, Usman, et al. "Learning-based methods for cyber attacks detection in IoT systems: a survey on methods, analysis, and future prospects." *Electronics* 11.9 (2022): 1502.
- [27]. https://www.youtube.com/watch?v=ZNWQN_g_ZsI&feature=youtu.be
- [28]. <https://www.geeksforgeeks.org/graph-plotting-in-python-set-1/>
- [29]. <http://benalexkeen.com/k-means-clustering-in-python/>
- [30]. <https://www.geeksforgeeks.org/analysis-of-test-data-using-k-means-clustering-in-python/>
- [31]. <https://www.datacamp.com/community/tutorials/k-means-clustering-python>

Quantitative Analysis And Ensemble Learning Based Predictive Modelling Of Taxi Trip Duration In New York City

Dr. Rinky Diwedi¹, Jivitesh Sabharwal², Apoorva Singh³, Himanshu Jha⁴

^{1,2,3,4}Computer Science Aand Engineering, Maharaja Surajmal Institute of Technology, Delhi, India

¹jiviteshsabharwal@gmail.com, ²himanshu36@yahoo.com, ³apoorvachoudhary24@gmail.com

Abstract— Taxis are an important part of transportation for a metropolitan and urban city like New York City to make the places more accessible where public transport is not much accessible and make travelling more convenient for people. In the last two decades, the traffic data has been drastically increased and can play a major role for big data analysis in the transportation industry. In this paper, we will analyze and study the taxi traffic in New York City and apply predictive analysis on publicly available data of taxi traffic of New York City using some naïve machine learning models like Linear Regression and some Ensemble learning techniques like Stacking.

The main purpose of this project is to develop an efficient and accurate machine learning model, to predict the average travel time or trip duration for a given pickup location, drop location, date and time of the taxi trip.

Keywords— Ensemble Modelling, Taxi trips, Predictive analysis, Machine Learning

1. INTRODUCTION

The taxi industry in New York City is an integral part of the city's transportation system, providing millions of rides each year to residents and visitors alike. Taxis, also known as "cabs," are licensed to pick up passengers on the street or at designated taxi stands, and they can be hailed by raising one's hand or using a taxi app.

The majority of cabs in New York City are licensed by the Taxi and Limousine Commission (TLC) [5] and are required to meet certain standards and regulations, including undergoing regular inspections and carrying commercial insurance. The TLC [5] also sets fares for taxi rides, which are displayed on a meter in the cab. The meter calculates the fare based on the distance traveled and the amount of time spent in the cab, as well as any additional fees that may apply (e.g., for additional passengers or baggage).

New York City has a variety of taxi options available, including traditional "yellow cabs," luxury cabs, and alternative service providers such as ride-hailing apps like Uber [6] and Lyft [7]. Yellow cabs are easily recognizable by their iconic yellow paint job and can be hailed on the street or at designated taxi stands. Luxury cabs, such as limousines and town cars, are typically reserved in advance and are more expensive than traditional cabs. Ride-hailing apps like Uber [6] and Lyft [7] operate under slightly different regulations than traditional cabs, and they may have different pricing structures and service areas. However, they provide a similar service and can be a convenient option for hailing a ride using a smartphone app.

The TLC [5] is responsible for regulating the taxi industry in New York City and is responsible for issuing and renewing licenses for taxi drivers, as well as setting fares and

enforcing regulations. The TLC also operates a dispatch system for cabs, known as the "Taxi of Tomorrow," which aims to improve safety and accessibility for passengers.

One potential application of big data in the analysis of taxi trips is to gain a better understanding of factors that influence demand for taxi services. This could include analyzing data on factors such as weather, time of day, location, and other variables to identify patterns and trends in taxi usage. This information could be used to optimize taxi operations, improve customer service, and increase efficiency. Other potential applications of big data in the analysis of taxi trips could include identifying patterns in driver behavior, predicting demand for taxi services, and improving the accuracy of routing and dispatch systems.

In summary, the taxi industry in New York City plays a vital role in the city's transportation system, providing a convenient and reliable option for getting around the city. While there are many different taxi options available, including traditional cabs, luxury cabs, and ride-hailing apps, they are all regulated by the TLC [5] and are subject to certain standards and regulations to ensure the safety and comfort of passengers.

The objective of this paper as follows:

- Study and analyze the taxi trips in New York City.
- To do a predictive analysis of the taxi trip duration from one pickup location and drop-off location.
- Build an ensemble learning model to efficiently predict the trip duration for a taxi.

The rest of the paper is organized as follows: Section II describes the summary of other studies which have been done in this field of interest. Section III discusses the research methodology. Finally, the conclusion is given in Section V.

2. LITERATURE SURVEY

According to Gelhausen et al. [14], airport accessibility can influence a traveller's decision between rival regional airports. 39% of all passengers travelling from Manhattan to airports in New York City are transported by taxis [15]. According to a 2002 poll, respectively 35% and 23% of incoming air travellers at LGA and JFK airports used taxis. From 1996 to 2005, the quantity of taxis dispatched to cities from JFK increased by 68% [15]. Nearly 3 million taxis were sent out from JFK in 2010, and 37% of people leaving the airport (through paid parking or commercial ground transportation) did so in a taxi [21].

Taxis are the second most popular means of transportation after private vehicles at Boston's Logan International Airport

(BOS) and Chicago's O'Hare International Airport (ORD), carrying 19% and 18% of passengers, respectively. Taxis are third after private and rental cars at Miami International Airport (MIA) in Florida and Hartsfield-Jackson Atlanta International Airport (ATL) in Atlanta, carrying 8% and 6% of passengers, respectively [32]. At London Heathrow Airport (LHR), Manchester Airport (MAN), Newcastle Airport (NCL), and Bristol Airport (BRS) in the United Kingdom, the modal share of taxis is 26.6%, 24%, 19%, and 16%, respectively [33]. The cab share at Hong Kong International Airport (HKG) is 13%, with 20% of the share going to business travellers [34].

According to Santani et al. [19], there is a general conflict of interest between taxi drivers and riders. Desired short wait times for customers typically indicate excess taxi availability, which results in low occupancy and low driver pay. Demand in the cities and at the airports is connected.

Because one is more desirable than the other, there may be an excess or a scarcity of taxis available at the airports.

However, 56% and 19% of airport managers, respectively, agree that the circling of taxis due to excess supply is an issue of some degree and a major problem. Excess supply will benefit passengers with reduced waiting times.

On the other side, H. Yang et al. [31] describes show the lowest taxi waiting time is a major factor in why cabs have the highest client satisfaction. Therefore, a taxi scarcity that causes long wait times for passengers may lower taxi satisfaction and utilization and harm airport competitiveness [13].

Various operational and policy solutions for both scenarios are reported in the literature. Radio taxi coordination is advised for the Spanish airport of Madrid due to a lack of available taxis [25]. Incentives for taxi drivers to shift extra service from the airport to the city are advised for Dallas Fort Worth [22]. Contrary to other cases in Europe, it was demonstrated that Dublin's taxi deregulation reduced the cab shortages at the airport without provoking a crowd [27]. At Honolulu International Airport in Hawaii, it was discovered that transitioning from the exclusive contract to the open system did not improve service quality [16]. To improve service quality, Barrett et al. [28] advocates using pre-booked cab services at airports. based on an analysis of the top ten U.S.

3. RESEARCH METHODOLOGY

In this section, the authors will describe about the research methodology, their techniques, and models they have applied for their predictive modelling. They will also describe about the dataset [10] which they have referred for this study.

3.1. Description of Dataset

The dataset used for this study is New York Taxi Trip duration [10] originally published by New York Taxi and Limousine Commission (TLC) [5] on Kaggle [9]. To compute and predict the taxi trip duration, this data includes information for 1458644 trip records in the train set whereas 625134 trip records in the test set with 11 attributes (as shown in Table []). This prediction data for taxi trip duration has 10 independent attributes and one target value for the taxi trip duration. This CSV file contains the reading from the taxi meters from a taxi over a period of 6 months from January 2016 to June 2016.

Table 1 Explanation of attributes

Data Variables	SI Units
Trip duration	s
ID	-
Vendor-ID	-
Pickup Date-time	Date-time
Drop-off Date-time	Date-time
Passenger Count	-
Pickup Longitude	Degrees
Pickup Latitude	Degrees
Drop-off Longitude	Degrees
Drop-off Latitude	Degrees
Store and Forward Flag	-

In the dataset, the target attribute for this study is *Trip Duration* which represents the duration of the taxi trip in seconds. The above given input attributes are described as follows, *ID* represents the unique identifier for each trip, *Vendor-ID* represents a specific code indicating the provider associated with the trip record, *Pickup Date-time* represents the date and time when the meter was engaged, *Drop-off Date-time* represents the data and time when the meter was disengaged, *Passenger Count* represents the number of passengers in the vehicle (this is value which was entered by the driver of the taxi). *Pickup Longitude* and *Pickup Latitude* represents the coordinates of the location where the meter was engaged, *Drop-off Longitude* and *Drop-off Latitude* represents the coordinates of the location where the meter was disengaged. *Store and Forward Flag* indicates whether the trip record was held in vehicle memory before sending to the vendor because the vehicle did not have a connection to the server - Y=store and forward; N=not a store and forward trip.

The following Figure [no.] represents the correlation heatmap between the attributes of the dataset:

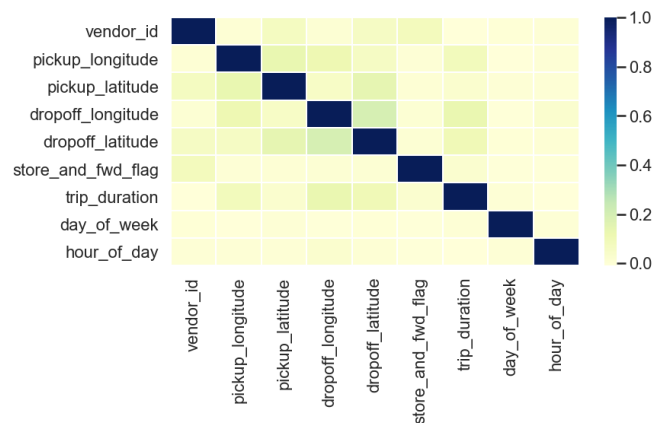


Figure 1 Correlation between attributes of the dataset

3.2. Univariate Analysis

Univariate analysis is a statistical method used to analyze data that contains only one variable. It involves examining the characteristics of that variable, such as its mean, median,

mode, and range, and identifying patterns or trends within the data. Univariate analysis is often used to summarize data and understand the underlying distribution of a variable. One of the main benefits of univariate analysis is that it is relatively simple and straightforward to understand. It requires minimal statistical knowledge and can be conducted using basic statistical tools.

In this section, we will analyze the target attribute, that is, *Trip duration*. It is crucial to understand it in detail as this is what we are trying to predict accurately. The following Figure [no.] summarizes the target attribute –

- Mean of the trip duration is 923s or 15.38 min
- Median of the taxi trip duration is 666s or 11.1 mins.
- Skewness is 28.58, that is right skewed distribution, this means the trip duration are of shorter periods.

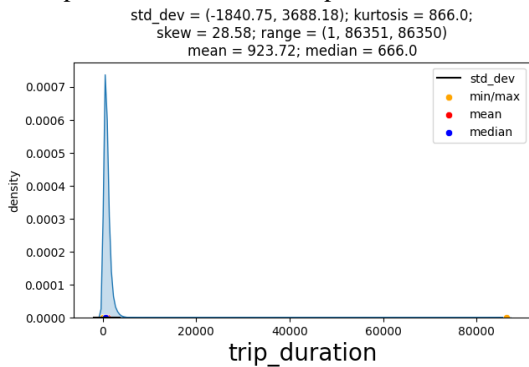


Figure 2 Distribution of Trip Duration (in seconds)

The above Figure [no.] shows the distribution of *Trip Duration* of a taxi trip but the author was not able to interpret about the distribution from this Figure [no. upar vala figure], so they transformed the values into their log values to get a better view of the distribution which can be seen following in the Figure [no. niche vala figure]:

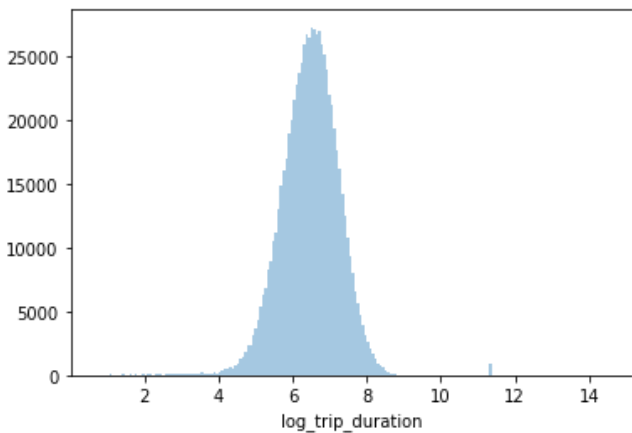


Figure 3 Log Transformation of Trip Duration

The above Figure [no.] shows that log transformation of *Trip Duration* follows a Normal (Gaussian) distribution [12].

In the analysis, the number of taxi trips or pickups from a location was also analysed for each weekday of a week and hours of a day. The Figure [week day vala] shows the distribution of number of pickups on each weekday of a week.

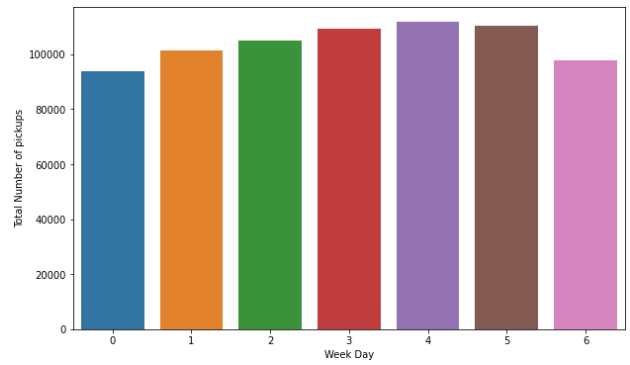


Figure 4 Number of Pickup of Each Weekday

The Figure [week day vala] shows that there are fewer pickups on weekends compared to weekdays, with the highest number on Thursday (represented as 4 in the decimal numbering system, where 0 is Sunday and 6 is Saturday). The number of taxi pickups at a location was also analysed for each hour of the day. Figure [hour vala] displays the distribution of pickups during each hour of the day.

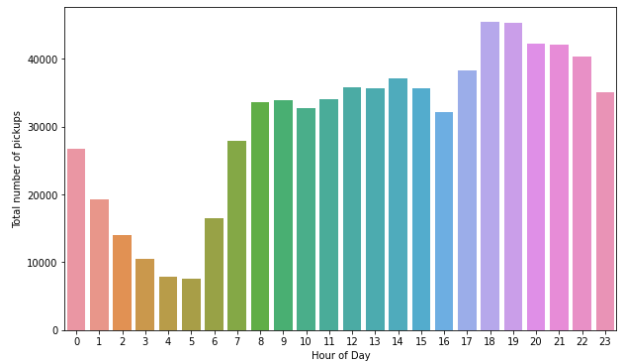


Figure 5 Number of Pickups of each Hour of the day

The Figure [no. of each hour] shows that there are more pickups in the late evenings, as expected, but fewer pickups during the morning peak hours and maximum number of pickups at 18th hour of the day.

Latitude and longitude of the pickup location were plotted into a heatmap. The heatmap in the Figure [map vala] visualizes the most frequently occurring Pickup points on the latitude and longitude map:

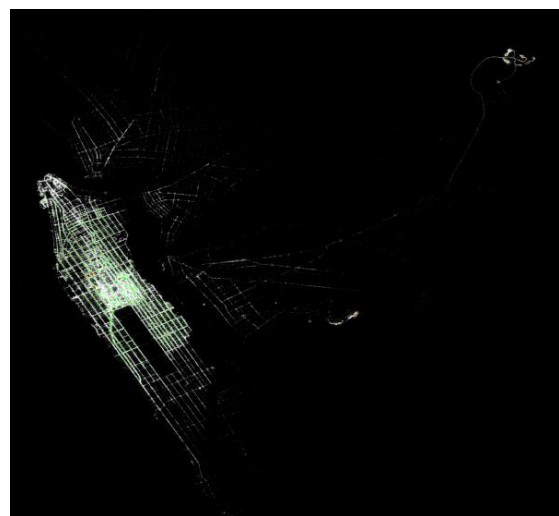


Figure 6 Most Frequently Occurring Pickup Points in NYC

The red points in the above Figure [no.] depicts that more than 25 taxi have been hailed from that pickup point, whereas the

green points indicate that 10 to 25 taxis have been hailed from that pickup point and the white points indicate that less than 10 taxis have been hailed from that pickup point. There are few small clusters for hot pickup points as displayed by red in the above plot. Most pickup points have less than 10 taxi trips and distributed all over the city. If you examine a map of New York City, you will notice that the majority of the red and green points are located in the Manhattan Area.

The latitude and longitude of pickup points and drop-off points have been provided for the study so this can be used for calculating the distance from these coordinates given. For calculating this distance, the haversine formula [11] will be our first choice.

The Haversine distance [11] is a measure of the distance between two points on a sphere based on their longitudes and latitudes. It is often used to calculate the distance between two points on the surface of a sphere, such as the distance between two points on the Earth's surface.

To calculate the Haversine distance between two points (lat1, lon1) and (lat2, lon2), you can use the formula:

$$a = \left(\sin^2 \left(\frac{\text{lat2} - \text{lat1}}{2} \right) + \cos(\text{lat1}) \cdot \cos(\text{lat2}) \cdot \sin^2 \left(\frac{\text{lon2} - \text{lon1}}{2} \right) \right) \quad (1)$$

$$d = 2R \cdot \sin^{-1} a \quad (2)$$

Where R is the radius of the sphere (here, the radius of the Earth), lat1 and lat2 are the latitudes of the two points, lon1 and lon2 are the longitudes of the two points, and d is the distance between the points.

This formula assumes that the Earth is a perfect sphere, which is not quite accurate. However, for most purposes, the Haversine formula provides a good approximation of the distance between two points on the Earth's surface.

3.3. Bivariate Analysis with Target Attribute

Bivariate analysis is a statistical method used to examine the relationship between two variables. It helps to identify the relationship between the two variables, whether they are positively related, negatively related, or independent of each other. This type of analysis is useful in understanding the relationship between variables and can be used to make predictions about future outcomes.

Bivariate analysis is an important tool in data analysis and can be used in a variety of fields, including economics, psychology, and biology. It is often used to test hypotheses about the relationship between variables and to identify trends and patterns in data. Bivariate analysis can also be used to make predictions about future outcomes and to inform decision-making.

For the first analysis, the Trip duration is analyzed with the weekdays of a week. The length of the trip may vary depending on the day of the week due to changes in traffic density, with the weekends potentially having different patterns than weekdays. Weekdays are represented as a decimal number, with 0-being Sunday and 6-being Saturday. The Figure [no.] below shows the average trip duration of each weekday:

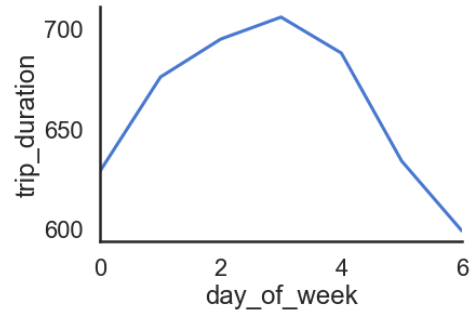


Figure 7 Average Trip Duration Each Day of the Week

This above Figure [no.] shows that on Wednesday the average Trip Duration is the highest than any other day of the week.

For second analysis, the Trip Duration is analyzed with the hours of a day. The duration of the taxi trip may fluctuate depending on the time of day due to variations in traffic congestion, with the early morning and evening hours possibly exhibiting different trends than other times. The time of day is expressed in the 24-hour format. The Figure [no.] below shows the average trip duration of each hour:

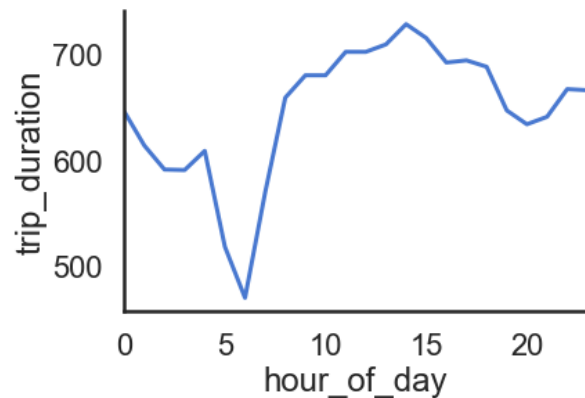


Figure 8 Average Trip Duration of Each Hour of the Day

The Figure [no. hour vala] shows that the taxi trip durations during the late night and early morning hours are shorter than at other times of the day, which can be attributed to the lower density of traffic at those times. This pattern is similar to the pattern of the number of pickups, which suggests that there is a correlation between the two. In other words, when there are fewer pickups, the trip durations tend to be shorter, and vice versa.

3.4. Model Training and Analysis

Machine learning has given computer systems the ability to automatically learn without being explicitly programmed.

In this, the author has used four machine learning algorithms ('Linear Regression', 'Ridge Regression', 'XGBoost', 'Stacking of Linear Regression and XGBoost').

3.4.1. Linear Regression

One of the simplest and most widely used Machine Learning techniques is linear regression. It is a statistical technique for performing predictive analysis. For continuous/real/numeric variables like sales, salary, age, and product price, among others, linear regression makes predictions.

$$Y = XB + e \quad (3)$$

Where Y is the dependent variable, X stands for the independent variables, B for the estimated regression coefficients, and e for residual errors. The variance that is not taken into account by the general model is taken into

account after the lambda function is added to this equation. There are actions one can do after the data is prepared and recognized as a component of L2 regularization.

Standardizing the variables (both dependent and independent) in ridge regression is done by subtracting their means and dividing by their standard deviations. This is a problem for notation since we need a way to say whether the variables in a given formula are standardised or not. All ridge regression computations are based on standardised variables in terms of standardisation. The resulting regression coefficients are scaled back to their original value before being displayed. The ridge trace, however, uses a conventional scale.

Ridge regression operates under the same linearity, constant variance, and independence presumptions as linear regression. The assumption that the distribution of errors is normal should not be made, though, as ridge regression does not offer confidence limits. Using this model, we got a Root Mean Squared Error (RMSE) of *308.27s* on the train set and RMSE of *309.71s* on the test set.

3.4.2. XGBoost

A powerful machine learning algorithm called XGBoost can assist you in comprehending your data and taking wiser decisions. An application of gradient-boosting decision trees is XGBoost. Data scientists and researchers from all around the world have utilised it to enhance their machine-learning models. Large dataset performance, usability, and speed are all priorities in the design of XGBoost. It does not require parameter optimization or adjustment; therefore, it may be used right away after installation with no additional settings. Regularization is a feature of XGBoost that enables you to prevent overfitting by imposing L1/L2 penalties on each tree's weights and biases. Many other gradient boosting solutions do not include this feature.

Using the weighted quantile sketch algorithm, XGBoost also has the ability to handle sparse data sets. While maintaining the same level of computational complexity as previous algorithms like stochastic gradient descent, this algorithm enables us to cope with feature matrices that include non-zero elements.

Additionally, XGBoost provides a block structure for concurrent learning. It makes scaling up on multicore computers or clusters simple. Additionally, it makes advantage of cache awareness, which lowers memory consumption when training models using sizable datasets.

Last but not least, XGBoost provides out-of-core computing capability by leveraging disk-based data structures rather than in-memory ones while doing computations. XGBoost is utilised for both model performance and execution speed.

Execution speed is critical since working with massive datasets requires it. Due to the fact that XGBoost has no dataset size limitations, working with larger datasets is possible than it would be with other algorithms.

Additionally, model performance is crucial since it enables you to develop models that can outperform existing models. Different algorithms, including random forest (RF), gradient boosting machines (GBM), and gradient boosting decision trees, have been contrasted with XGBoost. These evaluations demonstrate that XGBoost surpasses these other algorithms

in terms of model performance and execution speed. A machine learning process called gradient boosting builds a number of models and combines them to produce a final model that is more accurate than any of the individual models in the series. It supports predictive modelling issues involving both classification and regression.

It employs a gradient descent approach called gradient boosting to incorporate new models into an old one. The XGBoost library implements gradient boosting, often known as stochastic gradient boosting, multiple additive regression trees, or gradient boosting machines. Using this model, we got a Root Mean Squared Error (RMSE) of *298.36s* on the train set and RMSE of *302.24s* on the test set.

3.4.3. Stacking

Stacking is a popular technique in machine learning where multiple models are trained and their predictions are combined to make a final prediction. In this, author has implemented stacking using XGBoost and a linear regression model as a base model and train an Gradient Boosting Model on the predictions of the XGBoost and Linear Regression model's prediction.

To begin with, let's first understand the working of XGBoost and linear regression. XGBoost is a powerful ensemble machine learning algorithm based on decision trees. It works by building a strong model by training multiple weak models sequentially, each one correcting the errors of the previous model. XGBoost is known for its ability to handle large datasets and its high predictive power.

Linear regression, on the other hand, is a simple statistical model used to predict a dependent variable based on one or more independent variables. It works by finding the line of best fit through the data points, which can then be used to make predictions about new data.

In this, implementation of stacking of XGBoost and Linear regression model is discussed. First, we will train a XGBoost model and a linear regression on our dataset and use it to make predictions. Next, we will use these predictions as input features for our Gradient Boosting Model, which will then make final predictions based on these features.

In conclusion, stacking XGBoost and linear regression can be a powerful technique for making accurate predictions. By training a linear regression model on top of the predictions made by a Gradient Boosting Model, we can take advantage of the strengths of all the algorithms and make efficient use of their capabilities. Using this algorithm, we got a Root Mean Squared Error (RMSE) of *287.73s* on the train set and RMSE of *292.07s* on the test set.

4. CONCLUSION

During this research study, the findings shows that the stacking of models produced a better RMSE score out of all the other models which have been studied for this dataset. This shows that Stacking of XGBoost and Linear Regression using Gradient Boosting Model performed better with some hyperparameter tuning. The figure [graph vala figure] compares all the RMSE score of the algorithms which have been studied for this dataset.



Figure 9 RMSE scores for Different Models

5. FUTURE SCOPE

Based on the results obtained we aim to implement the following:

1. Weather analysis of each taxi trip can also be integrated in this study, so the aim to predict better RMSE score can be achieved.
2. The use of Open-Source Routing Machine [41] (OSRM) which is a modern C++ routing engine for shortest path from one location to another

REFERENCES

[1] X. Yu, "Big Data Driven Model for New York Taxi Trips Analysis," 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), 2021, pp. 1-4. <https://doi.org/10.1109/ICBDA51983.2021.9403054>

[2] M. A. Yazici, C. Kamga and A. Singhal, "A big data driven model for taxi drivers' airport pick-up decisions in New York City," 2013 IEEE International Conference on Big Data, 2013, pp. 37-44. <https://doi.org/10.1109/BigData.2013.6691775>

[3] G. Gao, M. Xiao and Z. Zhao, "Optimal Multi-taxi Dispatch for Mobile Taxi-Hailing Systems," 2016 45th International Conference on Parallel Processing (ICPP), 2016, pp. 294-303. <https://doi.org/10.1109/ICPP.2016.41>

[4] I. Markou, F. Rodrigues and F. C. Pereira, "Real-Time Taxi Demand Prediction using data from the web," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 1664-1671. <https://doi.org/10.1109/ITSC.2018.8569015>

[5] New York Taxi and Limousine Commission. [\[https://www.nyc.gov/site/tlc/index.page\]](https://www.nyc.gov/site/tlc/index.page)

[6] Uber Technologies Inc [<https://m.uber.com/>]

[7] Lyft, Inc. [<https://www.lyft.com/>]

[8] J. W. Chan, V. L. Chang, W. K. Lau, L. K. Law, and C. J. Lei, "Taxi app market analysis in Hong Kong," Journal of Economics, Business and Management, vol. 4, no. 3, 2016. <http://www.joebm.com/vol4/397-A001.pdf>

[9] Kaggle [<https://www.kaggle.com/>]

[10] New York City Taxi Trip Duration Dataset by New York Taxi and Limousine Commission [<https://www.kaggle.com/c/nyc-taxi-trip-duration>]

[11] Haversine Formula for calculating distance [https://en.wikipedia.org/wiki/Haversine_formula]

[12] Ali Grami, "The Gaussian Distribution," in Probability, Random Variables, Statistics, and Random Processes: Fundamentals & Applications, Wiley, 2019, pp.201-238, [<https://ieeexplore.ieee.org/document/8689279>].

[13] Chmura Economics&Analytics, "The economic impact of Richmond International Airport" Tehnical Report Prepared for Capital Region Airport Commission, Richmond International Airport, VA, 2008.

[14] M.C. Gelhausen, "Modelling the effects of capacity constraints on air travellers' airport choice," Journal of Air Transport Management, Vol.17, pp.116-119, 2011.

[15] Schaller Consulting, "The New York City Taxicab Fact Book," Schaller Consulting, Brooklyn, New York City, NY, 2006.

[16] S. J. La Croix, J. Mak, and W. Miklius, "Airport taxi service regulation: An analysis of an exclusive contract," Transportation, Vol.13, pp.145-161, 1986.

[17] S. J. La Croix, J. Mak, and W. Miklius, "Evaluation of alternative arrangements for the provision of airport taxi service," Logistics and Transportation Review, Vol.28, pp.147-166, 1992.

[18] C. Kamga, A. Conway, A. Singhal, A. and M. A. Yazici, "Using advanced technologies to manage airport taxicab operations," Journal of Urban Technology, Vol.19, pp.23-43, 2012.

[19] D. Santani, R. K. Balan, and C. J. Woodard (2008) "Spatio-temporal efficiency in a taxi dispatch system," Technical Paper, Singapore Management University, Singapore.

[20] C. Kamga, C., M. A. Yazici, A. Singhal, "Hailing in the Rain: Temporal and Weather Related Variations in Taxi Ridership and Taxi Demand-Supply Equilibrium" Paper presented at Transportation Research Board's 92nd Annual Meeting, Washington, D.C., 2013.

[21] A. Conway, C. Kamga, C., M. A. Yazici, and A. Singhal, "Challenges in managing centralized taxi dispatching at high-volume airports: Case study of John F. Kennedy International Airport, New York City,". Transportation Research Record, Vol.2300, pp.83-90, 2012.

[22] G. Tiwari, "A multi-city analysis of the taxicab industry's regulatory structure and market models in the United States," MTC Fall Student Conference. Center for Transportation Research and Education. Iowa State University, IA, 2005.

[23] D. Lam, K. Leung, J. Lyman, S. Terell, and R. Wilson, "The San Francisco taxicab industry: An equity analysis" Technical Report, The Richard & Rhoda Goldman School of Public Policy. University of California, Berkeley, CA, 2006.

[24] H. W. Shriner, and L. A. Hoel "Evaluating Improvements in Landside Access for Airports," Transportation Research Record, Vol. 1662, pp.32-40, 1999.

[25] W. Skok and J. A. Martinez, "An International Taxi Cab Evaluation: Comparing Madrid with London, New York, and Paris," Knowledge and Process Management, Vol.17, pp.145-153, 2010.

[26] E. A. Lukic, "Review of the taxicab dispatch operations" Aviation Department Final Report, Office of the County Auditor, Broward County, Florida, 2007.

[27] J. Jung, R. Jayakrishnan, and J. Y. Park, "Design and modeling of real-time shared-taxi dispatch algorithms," 92nd Transportation Research Board Annual Meeting, Washington, DC., 2013.

[28] S. D. Barrett, "The sustained impacts of taxi deregulation," Economic Affairs, Vol.30(1), pp.61-65, 2010.

[29] D. C. T. Costa, and R. Neufville, "Designing efficient taxi pickup operations at airports," Transportation Research Record, Vol. 2300, pp.91-99, 2012.

[30] R. Darbera, "Taxicab regulation and urban residents expectations from policy makers: A survey in eight cities" Paper presented at the 4th International IRU Taxi Forum. Cologne, Germany, 2010.

[31] H. Yang, C. S. Fung, K. I. Wong, K.I. and S. C. Wong, "Nonlinear pricing of taxi services," Transportation Research Part A, Vol.44, pp.337-348, 2010.

[32] G. D. Gosling, "Airport ground access mode choice models" ACRP Synthesis of Airport Practice 5, Transportation Research Board of the National Academies, Washington, D.C, 2005.

[33] I. Humphreys, and S. Ison, "Lessons from United Kingdom Airports on Ground Control Strategies," Transportation Research Record, Vol.1850, pp.70-78, 2003.

[34] M. L. Tam, and W. H. K. Lam, "Analysis of airport access mode choice: A case study in Hong Kong," Journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp.708-723, 2005.

[35] C. Kamga, A. Conway, A., M. A. Yazici, A. Singhal, and N. Aslam, "Strategies to improve taxi service for jetBlue customers at JFK terminal 5," Final Report, University Transportation Research Center, Region 2, New York, 2011.

[36] D. Biggar, "Why and how should we regulate taxis?" Report prepared for the Victorian Taxi Inquiry. State Government of Victoria, Australia, 2011.

[37] R. Abrams, S. Elson, and C. Mauldin, "Taxi 07: Roads Forward" Technical Report, Design Trust for Public Space and the New York City Taxi & Limousine Commission, New York City, NY, 2007.

[38] K. I. Wong, S. C. Wong, H. Yang, and C.O.Tong, "The effect of perceived profitability on the level of taxi service in remote areas," Journal of the Eastern Asia Society for Transportation Studies, Vol.5, pp.79-94, 2003.

[39] M. M. Grynbaum, "Where Do All the Cabs Go in the Late Afternoon?" New York Times, Jan 11th 2011. [\[http://www.nytimes.com/2011/01/12/nyregion/12taxi.html\]](http://www.nytimes.com/2011/01/12/nyregion/12taxi.html)

[40] D. Yassky, "New Short Return [Shortie] Policy a JFK and LaGuardia airports," Industry Notice #12-33, NYC Taxi & Limousine Commission, New York City, NY, 2012.

[41] Open Source Routing Machine (OSRM) [<https://project-osrm.org/>]

Motion Control: Camera-Based Body Movement Input Device Simulation

Dr. Kavita Sheoran¹, Priyaansh Malik², Pratyush Jha³, Mayank Sharma⁴

^{1,2,3,4}Computer Science and Engineering Department

Maharaja Surajmal Institute of Technology, New Delhi, India

¹kavita.sheoran@msit.in, ²priyaansh.malik@gmail.com, ³pratyushjha027@gmail.com, ⁴mayanksh4rma@gmail.com

Abstract — In the twenty-first century, most people lead sedentary lives. This has caused significant health repercussions and has been a source of concern. Many people today prefer to engage in virtual games and seem to avoid doing any physical activity or engage in a sport. The demand of immersive technology in the virtual world is at an all-time high. The idea of the project is to develop a camera based motion detection system to enable video game players to take their virtual reality experience up a notch by enabling them to use body movements to guide their games, and in turn allow the players to engage in some physical activity that too without the use of specialized hardware, a simple camera for input in this case.

Keywords—Motion-based, video gaming, sedentary-life consequences, computer vision

1. INTRODUCTION

Currently worth \$200 billion, the gaming business is unquestionably one of the most significant and cutting-edge segments of the IT industry. Its significance for entertainment, social networking, and culture cannot be overstated. More than 3 billion people worldwide are now getting one of the most immersive and mind-blowing types of entertainment from gaming, which has changed the definition of the phrase "entertainment business" from just referring to Hollywood and the film industry.

When people were tinkering with pong joysticks in their basements and vying with their buddies for the best Pac-Man scores in the 1970s and '80s, no one could have predicted what the gaming business would develop into. These straightforward games would eventually become the foundation of the modern gaming business and an immediate cultural phenomenon.

The gaming industry has developed unique techniques to make its games more intriguing, engaging, and interactive over the past few decades. The industry has come a long way in terms of generating unique and compelling representations of virtual worlds. These improvements and innovations have enabled video game players to engage with and feel a part of the virtual gaming world.

A. Health issues and Repercussions due to Gaming

Furthermore, most people in the twenty-first century live sedentary lives. This has had serious health consequences and is a topic of worry for both doctors and governments. Sedentary lifestyles among seniors have a negative impact on life expectancy [1]. This is the trend in most developed and developing nations.

Repetitive motions commonly found in gaming can lead to health problems such as carpal tunnel syndrome, which is caused by compression of the median nerve in the forearm,

leading to symptoms such as pain, numbness, and tingling in the hand and fingers.

- Gaming can cause carpal tunnel syndrome because it often involves repetitive motions such as clicking and scrolling with a mouse or joystick, which can put pressure on the median nerve [15]. Additionally, prolonged gaming sessions can lead to poor posture, which can also contribute to the development of carpal tunnel syndrome.

Other health issues that can be caused by gaming include [16]:

- Eye-strain: Prolonged gaming sessions can cause eye strain, leading to dry eyes, blurred vision, and headaches.
- Back pain: Sitting for long periods of time while gaming can lead to back pain and poor posture.
- Obesity: Gaming can be a sedentary activity, which can lead to weight gain and obesity if combined with a poor diet.
- Sleep disorders: Playing video games at night can disrupt the natural sleep patterns, leading to insomnia and other sleep disorders.

Video games in general-and motion-based games in particular-hold the promise of engaging older adults, and games have been applied in a variety of settings, e.g., to entertain nursing home residents or to motivate older adults to participate in physical therapy and rehabilitation [2]. While research has shown that motion-based games have positive effects on the emotional, physical, and cognitive well-being of older adults, these studies have only marginally addressed concerns regarding the usability and accessibility of such games [2].

B. Previous Ideas to Overcome

The idea of merging interactive and immersive gaming with physical activity and body movement-oriented games has been doing the rounds in the gaming industry for a while now. Even the health benefits of such video games for the development of motor skills in young toddlers and physical therapy and rehabilitation in older people have been widely experimented with. The majority of these ideas has been implemented and has received positive and favorable responses. All of these games, however, require a specialized console/hardware like Microsoft Kinect [3] for capturing human movement and consequently enabling user interaction with the gaming system. Another example of this technology is the Nintendo Wii, which uses a motion-sensing controller that players hold and move to control the on-screen action. This allows players to engage in physical

activity while gaming, which can help to reduce the risk of repetitive motion injuries.

Motion-based video gaming (MBVG) is a type of gaming that requires the individual playing the game to be physically interactive. Thus, whatever movements the individual playing the game does is picked up by motion sensors and mimicked via the on-screen character. MBVG provides constant feedback to learners and has been found to help motivate students, replace sedentary with active gaming, and can facilitate social interactions with peers [4]. In recent years, motion-sensing controllers, such as handheld controllers, gloves, or even full body suits, have been introduced in the gaming industry. These controllers can track a player's movements and translate them into in-game actions. Virtual reality (VR) technology enables players to physically interact with the game. Because it enables players to both immerse themselves in the game and exercise, this technology is gaining popularity.

C. Overcoming the Hardware barrier

We wish to provide a minimal interface for keystroke emulation on the basis of body movement and consequently get away with the usage of specialized hardware provided by corporations for their games and gaming systems. This idea has special and easy-to-implement applications in games of the genres like automobile racing and other arcade games.

The need for a minimal hardware system for game interaction is felt time and again. As a consequence, we developed the concept for this project.

One of the main advantages of using a camera-based motion detection system is that it does not require any specialized hardware to function. By using a webcam or a smartphone camera, your project can be used on a wide range of devices, including computers, smartphones, and tablets. This makes it accessible to a broader audience and eliminates the need for expensive specialized hardware.

Additionally, by using the OpenCV library, your project can be implemented on different platforms, such as Windows, MacOS, or Linux, so users can play the games on different devices without any problem.

Another way to overcome the hardware barrier is by using Mediapipe, which provides pre-trained models that can be used to detect and track objects, poses, and hands in realtime, allowing the controller to respond quickly and accurately to player movements, which is essential for a smooth and responsive gaming experience.

In summary, this project overcomes the hardware barrier by being able to work with existing hardware, such as webcams, smartphones, or tablets, and by being crossplatform and using pre-trained models provided by Mediapipe, it makes it accessible to a broader audience and eliminates the need for expensive specialized hardware.

2. OBJECTIVES

- Implement motion and body detection using a simple camera setup.
- The primary purpose is to create a standardized interface for game interaction based on body movements.
- Allow immersive experience in games by allowing the users to get a sense of actually interacting with the virtual environment.
- The project also aims to allow users to physically interact with fast-paced games without the use of specialized hardware.

- An alternative objective of physical activity is to be achieved by the means of the application of this project.

3. LITERATURE SURVEY

There have been many similar, documented attempts to play a moderately complex game using motion-sensing based algorithms. Some of them are presented here.

"Real-time Multi-person 2D Pose Estimation uses Part Affinity Fields" by Cao et al. (2017) presents a method for real-time 2D pose estimation of multiple people in images. This technique can be used to track player movements and detect specific gestures and movements.

"Real-time Hand Tracking uses Convolutional Networks" by Simon et al. (2018) presents a method for real-time hand tracking using a convolutional neural network (CNN). This technique can be used to track the player's hand movements and detect specific gestures.

"Using Computer Vision Techniques to Play an Existing Video Game" by Christopher E. Erdelyi [5] proposes a method for playing a racing game using computer vision and controlling the game only through the inputs provided by human players using physical activities. The project uses Open Source Computer Vision Library (OpenCV) to track or recognize hands from real-time webcam images for interacting with the game of tai-chi. To capture the hand's motion is based on two main methods: hand segmentation and hand tracking. For hand segmentation, the color cue is used due to the characteristic color values of human skin and segmenting the hand. To track the hand motion, Motion History is used.

"Real-time Multi-object Tracking using YOLOv3 and Deep SORT" by Wojke et al. (2018) presents a real-time multi-object tracking method using the YOLOv3 object detector and the Deep SORT algorithm. This technique can be used to track multiple objects in real-time, which is essential for a smooth and responsive gaming experience. "Real-time Hand Gesture Recognition uses Convolutional Neural Networks" by Kollreider et al. (2019) presents a method for real-time hand gesture recognition using a CNN. This technique can be used to detect specific hand gestures, such as punches or jumps, and use that information to control the game.

The impact of motion sensing and conventional gaming controllers on players' experience is immense. The sense of immersion is investigated based on spatial and kinesthetic involvement dimensions, currently available analytical constructs and data gathered during conducted experiments as shown in "Motion-sensing Game Tools and Player Involvement" by Pavels Saulins [6]. The paper attempts to propose new experimental ways of analyzing player-game tool interaction and centering experiments on the main research hypothesis, which assumes that spatial and kinesthetic involvement dimensions can induce a deeper sense of immersion when using motion-sensing technology rather than conventional gaming devices because it anchors the player more directly. It concludes that it is evident that the roles of kinesthetic and spatial involvement dimensions are having a significant impact on user experience. The uncovered interrelationship between the fluidity of controls and the reported sense of immersion can lead us as researchers and designers to focus on certain aspects that play an important role in forming powerful gaming experiences. "Recognition-based gesture spotting in video games" by Kang, Hyun & Lee, Chang & Jung, Keechul (2004) [7]

explores the use of Computer Vision to create a gesturebased game, controlled by hand using the Haar Cascade classifier to classify and detect the object. With the help of simple hand gestures, the model has been trained to enable users to play the game of “Flying Wing Mario”

"Hand Gesture Recognition System for Games," by N. V. Le, M. Qarmout, Y. Zhang, H. Zhou and C. Yang [8] uses OpenCV to measure the angles between contours, the number of fingers held up and hand gestures to control video games. It focuses on low-latency input and real-time gesture conversion to allow for high pace gaming all via hand gestures.

GAMING:

"A Review of Motion-based Gaming Interfaces" by Mase et al. (2011) presents a review of motion-based gaming interfaces, including the Nintendo Wii and the Xbox Kinect. This paper provides an understanding of the current state of the art in motion-based gaming interfaces and how they relate to the project.

"The Impact of Virtual Reality on Physical Therapy" by Riva et al. (2015) presents a review of the use of virtual reality technology in physical therapy. This paper provides an understanding of how virtual reality technology can be used in rehabilitation and physical therapy.

MEDICAL AND REHABILITATION:

"The Effect of Video Game Training on Balance Control in Older Adults" by Brouwer et al. (2015) presents a study on the effect of video game training on balance control in older adults. This study shows that video games can be an effective tool for improving balance and coordination.

"The Use of Video Games in Physical Rehabilitation: A Systematic Review" by Kato et al. (2016) presents a systematic review of the use of video games in physical rehabilitation. This review shows that video games can be an effective tool for rehabilitation and physical work.

4. MATERIALS AND METHODS

A. Hardware Utilized

TABLE I. SYSTEM SPECIFICATION FOR WINDOWS SYSTEM:

System parameter	Specification
OS Name	Microsoft Windows 10 Home
OS Version	10.0.19044 N/A Build 19044
OS Manufacturer	Microsoft Corporation
OS Configuration	Standalone Workstation
OS Build Type	Multiprocessor Free
System Type	x64-based PC
Processor	Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz
Installed RAM	8.00 GB (7.74 GB usable)
Cores	4 Cores
Python Version used	3.9.4

TABLE II. SYSTEM SPECIFICATION FOR LINUX SYSTEM:

System parameter	Specification
OS Name	Linux
OS Version	Ubuntu 20.04

OS Manufacturer	Canonical
System Type	x86_64
Installed RAM	8.00 GB (7.7 GB usable)
Cores	2GHz Quad-Core Processor
Python Version used	3.9.4

B. Software used

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured, object-oriented and functional programming. Python is developed under an OSI-approved open source license, making it freely usable and distributable, even for commercial use. Python's license is administered by the Python Software Foundation [15].

Git is a version control tool used to keep track of changes to files and software and to coordinate teamwork on those files. It enables several developers to collaborate on the same code at once without interfering with one another. Git enables developers to quickly go back to earlier iterations of the codebase, making it simple to fix mistakes and try out new concepts and ideas. Git repositories may be hosted and managed on GitHub. To work together on software projects, exchange code, and create and maintain open-source software, millions of developers and companies utilize it worldwide. A sophisticated collection of command-line tools, a web-based user interface, and interfaces with several other programs and services are just a few of the features and tools that GitHub offers to make working with Git simple. GitHub is based on the Git software.

Bash, short for the Bourne-Again Shell, is a command language and Unix shell that is commonly used on Linux and other Unix-like operating systems. It is a popular choice for scripting and automating tasks, particularly for system administrators and professionals who need to manage multiple systems. Its ability to combine simple commands to create powerful scripts is one of its major benefits, allowing developers to automate complex tasks and create unique tools and utilities that are easy to share and reuse.

C. Various Python Libraries

NumPy is a powerful package for array processing in Python, which is widely used in scientific computing. It provides a high-performance multidimensional array object and tools for working with these arrays. It is the fundamental package for scientific computing with Python, it contains a number of important features:

- A robust N-dimensional array object
- Advanced (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Valuable linear algebra, Fourier transform, and random number functionality.

In addition to its clear scientific applications, NumPy also functions as an efficient container of generic data in multiple dimensions.

Pandas is an open-source library built on top of the NumPy library, providing various data structures and operations for manipulating numerical data and time series in Python. It is fast and efficient, and offers high-performance and productivity for users. Pandas is easy-to-use and offers high-performance data structures and data analysis tools for the

Python language. It is widely used in a variety of fields including academic and commercial domains such as economics, statistics, and analytics.

Scikit-learn (Sklearn) is a powerful and dependable library for machine learning in Python. It is an open-source library that offers a wide range of machine learning, preprocessing, cross-validation, and visualization algorithms using a unified interface. Sklearn provides a variety of efficient tools for machine learning and statistical modeling, including classification, regression, clustering, and dimensionality reduction through a consistent interface in Python. This library, written primarily in Python, is built on top of NumPy, SciPy, and Matplotlib.

MediaPipe is an open-source framework/machine learning solution for building ML pipelines for processing sensory data like video, audio, etc. provided by Google LLC. It enables deployment of ML models onto mobile (Android, iOS), web, desktop, edge devices, and IoT, effortlessly.

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. The library is cross-platform and free for use. It consists of hundreds of efficient and optimized algorithms for various computer vision tasks.

The venv module supports creating lightweight “virtual environments”, each with their own independent set of Python packages installed in their site directories. A virtual environment is built on top of an existing Python installation, known as the “base” Python of the virtual environment, and is separated from the packages in the base environment, so that only those exclusively installed in the virtual environment are available.

5. METHODOLOGY

We used a standard camera and computer vision algorithms to track body movements in real-time. The algorithms were trained on a dataset of body movements, and were able to accurately track and classify different movements. We then used these tracked movements to simulate input in a virtual environment.

The methodology for the project can be broken down into several steps:

A. Research

Conduct research to identify the specific requirements and goals of the project, such as the types of games and movements that need to be tracked, and the type of hardware that will be used.

B. Design:

Create a design for the controller, including the hardware and software components that will be used, and how they will interact with each other.

C. Implementation

Implement the design using appropriate programming languages, such as Python and C++, and libraries such as OpenCV, PyKey and Mediapipe

D. Calibration

Calibrate the controller by adjusting the settings and parameters to ensure that it is able to accurately track player movements and detect gestures

E. Testing

Test the controller with a variety of games and movements to ensure that it is able to accurately track player movements and detect gestures.

F. Optimization

Optimize the controller to improve its accuracy and responsiveness, as well as to reduce its processing time.

G. Deployment

Deploy the controller on the target platform, such as a computer or a mobile device, and test it to ensure that it is working as expected..

H. Maintenance

Regularly update the software and perform maintenance to ensure that the controller continues to work as expected

6. FEASIBILITY STUDY

The preliminary investigation examines project feasibility; the likelihood the system is going to be useful to the organization. The main objective of the feasibility study is to check the Technical, Operational, and Economical feasibility for adding new modules and debugging old running systems. All systems are possible if they have unlimited resources and infinite time to do a task. There are aspects within the feasibility study portion of the preliminary investigation:

Economic Feasibility

Technical Feasibility

Operational Feasibility

A. Economic Feasibility

As systems are often developed technically which are going to be used if installed must still be an honest investment for the organization. In the economic feasibility, the event cost in creating the system is evaluated against the last word benefit derived from the new systems. Financial benefits must equal or exceed the costs. The system is economically feasible. It doesn't require any addition hardware or software. Since the interface for this system is developed using the existing resources and technologies, there is nominal expenditure and economic feasibility for certain.

B. Technical Feasibility

This assessment focuses on the technical resources available to the organization. It helps organizations determine whether the technical resources meet capacity and whether the technical team is capable of converting the ideas into working systems. Technical feasibility also involves evaluation of the hardware, software, and other technology requirements of the proposed system. This assessment is predicated on an overview design 39 of system requirements, to work out whether the corporate has the technical expertise to handle completion of the project. When writing a feasibility report, the subsequent should be taken to consideration:

- A brief description of the business to assess more possible factors which could affect the study
- The part of the business being examined

0ed operations directly on the image. For masking the mask filter is moved point to point on the image and at each point, the filter is calculated by a predefined relationship.

The MediaPipe solution utilizes a two-step detector/tracker ML pipeline.

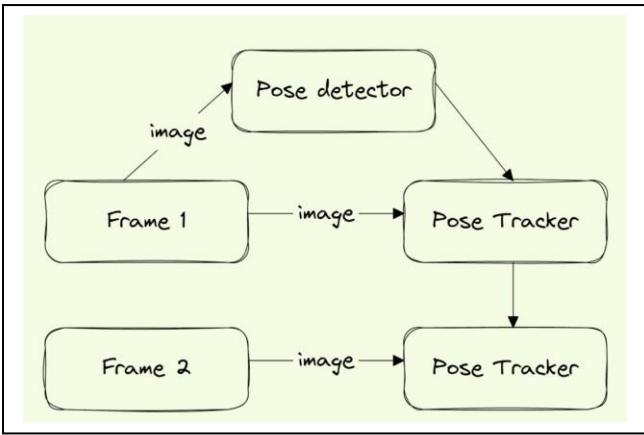


Fig. 2. Pose estimation pipeline for BlazePose

Using an input feed reader, the pipeline first locates the person/pose region-of-interest (ROI) within the frame. The tracker subsequently creates a mask and performs segmentation to predict the pose landmarks and segmentation mask within the ROI using the ROI-cropped frame as input as shown in Fig. 1. This is done for the very first frame when the tracker has to actively identify body frame and the following frames derive the ROI from the information provided by the previous frames and poses.

Inspired by Leonardo's Vitruvian man, BlazePose estimates the midpoint of a person's hips, the radius of a circle circumscribing the entire body, and the inclination angle of the line joining the shoulder and hip midpoints. This leads to consistent monitoring even in the most intricate circumstances.

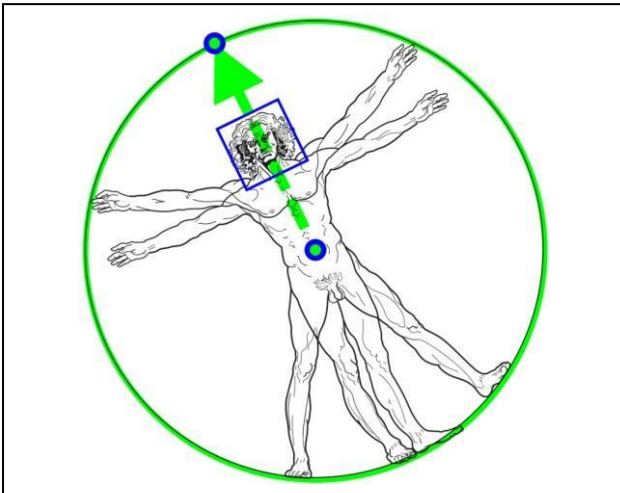


Fig. 3. Vitruvian man aligned via two virtual key points predicted by our

BlazePose detector in addition to the face bounding box [13]

The program first specifies all the relevant joints and locations of interest from the input frame. It stores these locations in an array of objects which contain the X and Y coordinates of landmark points. It then calculates the angles between shoulder and elbow joints, away from the position of the face and then, according to the algorithm fed to the program it can emulate directions keys on the keyboard or even mouse buttons or scroll events.

Fig. 4 shows that the program is able to detect/locate ROIs point numbers 0, 1, 2, 3, 4, 5, 6 (See Fig. 1) along with the region of shoulder and elbows with point number 11, 12, 13 and 14.

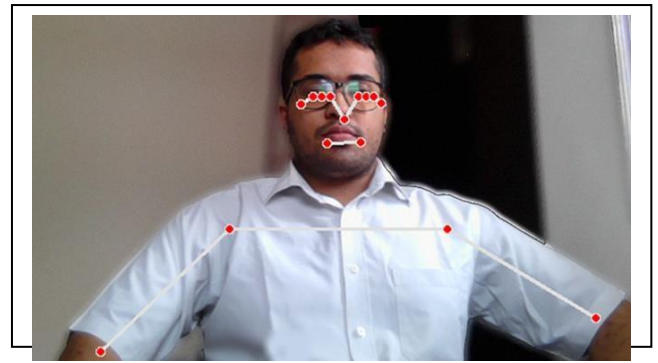


Fig. 4. Initial Detection



Fig. 5. Forward movement

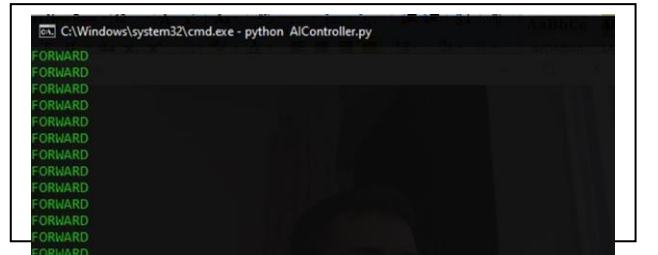


Fig. 6. Forward movement detection by program



Fig. 7. Forward movement detection resulting in up arrow key emulation in game

The motion detection by the program has allowed the emulation of the keystroke as seen in Fig. 6. and Fig. 6. The logic for the same is:

```

if left_elbow_angle > 90 and \ right_elbow_angle > 90 and \
left_wrist_angle > 90 and \ right_wrist_angle > 90:
# emulate up arrow keystroke
    
```

where the left_elbow_angle and right_elbow_angle are the angles created by the elbows of the respective hands (Refer Fig. 6.) with respect to the shoulder recognised as POI by the program.

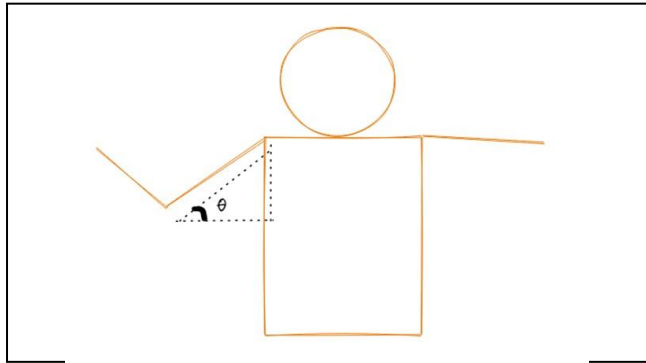


Fig. 8. Elbow angle



Fig. 9. Right movement

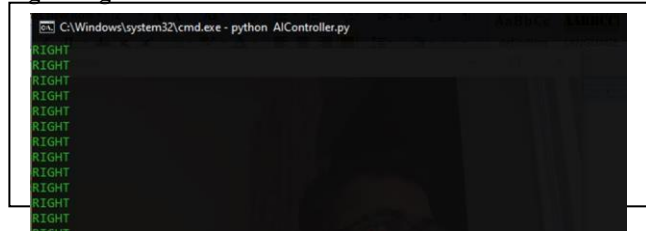


Fig. 10. Right movement detection by program

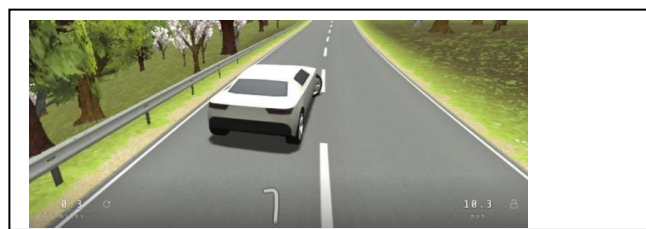


Fig. 11. Right movement detection by program

You can see in the figures that the program is identifying body movement gestures as directional arrow keys of a keyboard.

The motion detection by the program has allowed the emulation of the keystroke as seen in Fig. 9., Fig. 10., and Fig.11.

The logic for the same is:

```

if left_shoulder_angle < -10:
# emulate right key stroke
    
```

where the left_shoulder_angle is the angle created by the left shoulder of the user with respect to the other corresponding shoulder (Refer Fig. 12.) recognised as POI by the program.

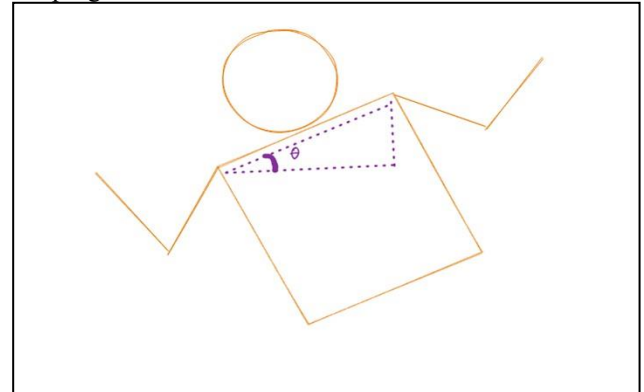


Fig. 12. Shoulder angle



Fig. 13. Left movement

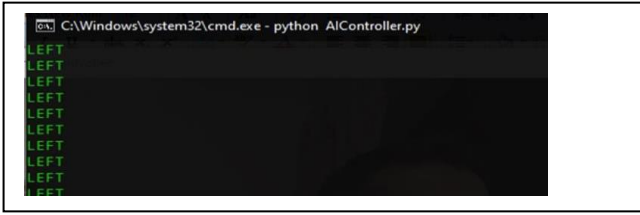


Fig. 14. Left movement simulation by program

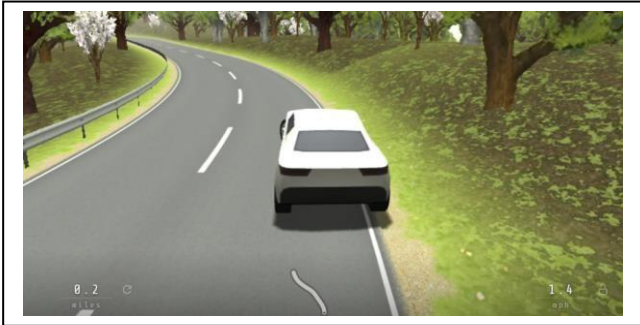


Fig. 15. Left movement detection resulting in left arrow keystroke emulation on game The motion detection by the program has allowed the emulation of the keystroke as seen in Fig. 13., Fig. 14., and Fig. 15.

The logic for the same is:

```
if left_shoulder_angle > 10:
    # emulate left key stroke
```

where the left_shoulder_angle is the angle created by the left shoulder of the user with respect to the other corresponding shoulder (Refer to Fig. 12.) recognized as POI by the program.

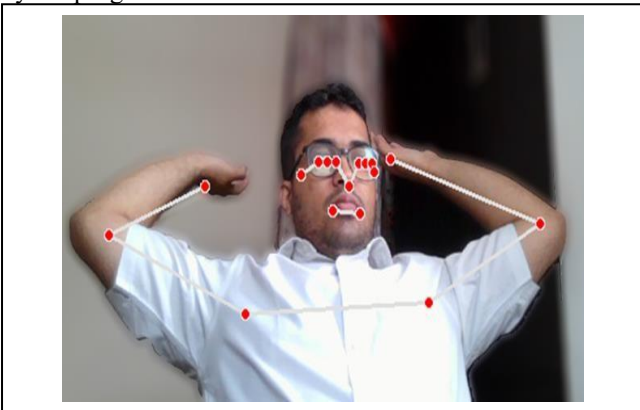


Fig. 16. Leaning back



Fig. 17. Leaning back movement detection by program

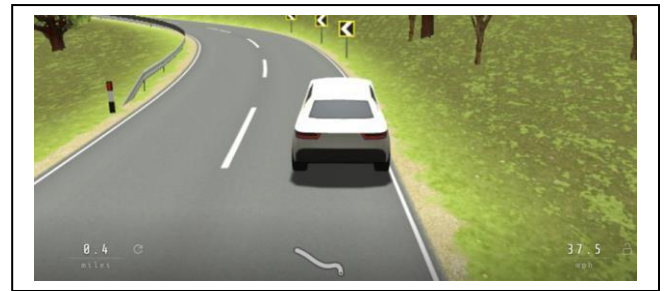


Fig. 18. Back movement detection resulting in back arrow keystroke emulation on game

The motion detection by the program has allowed the emulation of the keystroke as seen in Fig. 9., Fig. 10., and Fig. 11.

The logic for the same is:

```
if left_elbow_angle > 90 and \ right_elbow_angle > 90 and \
left_wrist_angle < 90 and \ right_wrist_angle < 90:
    # emulate up back arrow keystroke
```

where the left_elbow_angle and right_elbow_angle are the angles created by the elbows of the respective hands (Refer Fig. 8.) with respect to the shoulder recognized as POI by the program.

7. RESULT AND DISCUSSION

The results achieved in the project are as follows

- Body detection using python libraries OpenCV and Mediapipe.
- Motion detection by the program.
- Motion classification by the program based on body movements
- Emulation of keystrokes for the purpose of in-game movement control
- A greater sense of connection to the virtual environment by means of actual physical indulgence in the game

8. CONCLUSION AND FUTURE SCOPE

A. Conclusions

The camera-based motion detection game controller that we have created using Python, OpenCV, PyKey, and Mediapipe is a unique and innovative project that has the potential to enhance the gaming experience.

Using OpenCV, the controller can track player movements by analyzing video footage from a camera. This allows for a more immersive and interactive gaming experience, as the player's movements are directly translated into in-game actions.

Combining OpenCV with PyKey and Mediapipe, the controller can detect specific gestures or movements, such as punches or jumps, and use that information to control the game. This allows for a more natural and intuitive control scheme, as players can use their own movements to interact with the game.

Additionally, using Mediapipe, the controller can process the image and video data from the camera in real-time, which allows for a smooth and responsive gaming experience.

In conclusion, the camera-based motion detection game controller project is an innovative and exciting project that has the potential to revolutionize the way we interact with technology. By using computer vision and machine learning techniques, the controller can track player movements in real-time, detect specific gestures and movements, and enhance the video footage.

The controller has the potential to create a more immersive and interactive gaming experience, it can also be used as a tool for weight management, rehabilitation, and even for improving balance and coordination. Additionally, it can be used in a wide range of applications, such as gaming, augmented reality, and robotics.

The project also overcomes the hardware barrier by being able to work with existing hardware, such as webcams, smartphones, or tablets, and by being cross-platform and using pre-trained models provided by Mediapipe, it makes it accessible to a broader audience and eliminates the need for expensive specialized hardware.

Overall, this project demonstrates the potential of using computer vision and machine learning to create a more immersive and interactive gaming experience. With this controller, players are able to engage in physical activity while having fun, it can be used as a tool for weight management, rehabilitation, and even for improving balance and coordination.

B. Future enhancement and scope

Motion based virtual gaming is quite beneficial and practical in the present day scenario, where artificial intelligence is growing so rapidly, provided a common interface is utilized for various gaming applications.

A complex algorithm to emulate most keystrokes in various virtual games and environments irrespective of the game physics and game pace can be developed by allowing the user to change various keystroke emulation settings dynamically can be developed to further this idea.

Further scope includes:

- **Gaming:**

The controller can be used to create a more immersive and interactive gaming experience. It can be used in games that require players to move and engage in physical activity, such as sports games or fitness games. □ Rehabilitation and Physical Therapy:

The controller can be used as a tool for physical therapy and rehabilitation. It can help patients regain strength, range of motion, and coordination after an injury or surgery.

- **Virtual and Augmented Reality:**

The controller can be used to create more realistic and immersive virtual and augmented reality experiences. It can be used to track player movements and translate them into in-game actions.

- **Robotics:**

The controller can be used to control robots and other machines using gestures and movements. It can be used to create a more natural and intuitive interface for controlling robots.

- **Smart home:**

The controller can be used to control smart home devices using gestures and movements, this could make the interaction with these devices more intuitive and natural.

- **Healthcare:**

The controller can be used to monitor patients' movements and detect any signs of deterioration. Video games in general- and motion-based games in particular hold the promise of

engaging older adults, and games have been applied in a variety of settings, e.g., to entertain nursing home residents or to motivate older adults to participate in physical therapy and rehabilitation [2].

Overall, this project has the potential to revolutionize the way we interact with technology, making it more intuitive, natural, and engaging. With the advancement of technology and the increasing demand for more immersive and interactive experiences, the future scope of this project is quite bright.

REFERENCES

- [1] VanBeveren, P.J. and Avers, D. Exercise and Physical Activity for Older Adults. In: Guccione, A., Wong, R.A., & Avers, D. (Eds.). Geriatric Physical Therapy. St. Louis: Elsevier (2012), 64-85.
- [2] Kathrin M. Gerling, Kristen K. Dergousoff, and Regan L. Mandryk, "Is Movement Better? Comparing Sedentary and Motion-Based Game Controls for Older Adults," In: 39th Graphics Interface Conference, GI 2013, 29 - 31 May 2013, Regina, Saskatchewan, Canada.
- [3] Wikipedia "Kinect", Wikipedia, November 26, 2022. [Online]. Available: <https://en.wikipedia.org/wiki/Kinect> [Accessed November 26, 2022].
- [4] Jenny, S. E., Schary, D. P., Noble, K. M., & Hamill, S. D. (2017). "The Effectiveness of Developing Motor Skills Through MotionBased Video Gaming: A Review," *Simulation & Gaming*, 48(6), 722-734, page 722, December 2017
- [5] Christopher E. Erdelyi, "Using Computer Vision Techniques to Play an Existing Video Game", M.S. Thesis, College of Science and Mathematics, California State University, San Marcos, 2019. [6] Pavels Saulins, "Motion-sensing Game Tools and Player Involvement", M.S. Thesis, IT University of Copenhagen, Copenhagen, 2016.
- [7] Kang, Hyun & Lee, Chang & Jung, Keechul (2004) "Recognitionbased gesture spotting in video games", *Pattern Recognition Letters* 25(15) 1701-1714, 10.1016/j.patrec.2004.06.016.
- [8] N. V. Le, M. Qarmout, Y. Zhang, H. Zhou and C. Yang, "Hand Gesture Recognition System for Games," 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 2021, pp. 1-6, DOI: 10.1109/CSDE53843.2021.9718421
- [9] GOOGLE LLC, "MediaPipe Pose," [google.github.io](https://github.com/google/mediapipe), para. 2 [Online]. Available: <https://google.github.io/mediapipe/solutions/pose> [Accessed November 2, 2022].
- [10] OpenCV Team, "About OpenCV," opencv.org, [Online]. Available: <https://opencv.org/about/> [Accessed November 2, 2022].
- [11] Valentin Bazarevsky and Ivan Grishchenko, Research Engineers, Google Research, "MediaPipe Pose," ai.googleblog.com, August 13, 2020 [Online]. Available: https://mediapipe.dev/images/mobile/pose_tracking_full_body_landmarks.png [Accessed November 12, 2022].
- [12] NumPy developers. NumPy. [Numpy.org](https://www.numpy.org) website. <http://www.numpy.org>. Updated 2019. [Accessed October 2022].
- [13] Valentin Bazarevsky and Ivan Grishchenko, Research Engineers, Google Research, "MediaPipe Pose," ai.googleblog.com, August 13, 2020[online]. Available:<http://1.bp.blogspot.com/i8ZfrqzbckE/XzVxC19SuyI/AAAAAAAAAGYY/XAAoXAgx34kvOtKlYpi0tQyiCfREGADSGCLcBGAsYHQ/s878/image4.jpg> [Accessed November 12, 2022]. Python.org [Online]. Available: <https://www.python.org/about/> [Accessed November 30, 2022]
- [14] Woo, H.C., White, P., Ng,H. K., & Lai,C. W. "Development of kinematic graphs of median nerve during active finger motion: Implications of smartphone use".*PLoS One*.11:e0158455.doi: 10.1371/journal.pone.0158455, 2016
- [15] Israel Oluwasegun Ayenigbara, "Gaming Disorder and Effects of Gaming on Health: An Overview",*Journal of Addiction Medicine and Therapeutic Science*, September 2018.

Prediction Of Heart Disease Using Ensemble Learning And Ant Colony Optimization

Lititia Khanna

Department of Electronics and Communication Engineering Maharaja Surajmal Institute Of Technology, New Delhi, India
lititiakhanna@gmail.com

Abstract—Healthcare is slowly changing as advanced technological developments such as machine learning, artificial intelligence and data science make data collection and retrieval much easier and faster. Advanced technology has delivered superior medical outcomes, increasing both productivity and efficiency in healthcare facilities. These developments have also improved the lives of physicians, allowing them to spend more time caring for patients instead of wasting time collecting data, boosting staff morale and improving retention. . Machine learning and artificial intelligence have brought remarkable improvements in areas such as diagnosis, prognosis, automated clinical decision-making, and disease detection in healthcare organizations. Ensemble learning is a method of developing and integrating multiple models, such as classifiers, to solve a specific artificial intelligence problem. Used to improve predictions, classification, performance metrics, etc. These classifiers improve the predictive accuracy of healthrelated models. Ensemble learning methods include bagging, stacking, voting, boosting, and pasting. Ant Colony Optimization (ACO) is a technique used for computational feature selection that minimizes unimportant features and solves the problem iteratively to obtain the desired solution. Results shows that Bagging and ACO achieved highest accuracy i.e., 95.08%.

1. INTRODUCTION

CORONARY HEART DISEASE (CHD) is a kind of disorder in which the arteries are unable to deliver oxygen rich blood to heart. This affects the heart's function and it's structure. Coronary artery disease is the leading cause of increasing mortality in the UK and around the world. The disease can be asymptomatic and range from chest pain to cardiac arrest. Some symptoms of this disorder include neck pain, chest pain, cold sweats, trouble sleeping, dizziness, light-headedness, tightness in the chest, fatigue, shortness of breath during exercise, nausea or indigestion, and eventually lead to a heart attack commonly known as angina pectoris. Coronary heart disease is usually the result of a buildup of plaque, a sticky substance within the walls of the coronary arteries. This substance can partially or completely block blood flow to the coronary arteries. Blood vessels in CAD patients become clogged with fat and cholesterol, reducing blood flow to the heart.

Many factors underlie this condition, including high blood pressure, cholesterol, diabetes, lack of exercise, and smoking. The main risk factors for this heart disease are:

- Age: As we age, arteries can narrow and become damaged.
- Gender: Men are usually at a significant risk of developing CHD. After menopause, women are at an increased risk of developing the disease.

- Family history: People with this heart disease in their families are at increased risk of CHD, especially if the disease is diagnosed in relatives. If the person's father/brother was diagnosed before age of 55 or her mother/sister was diagnosed before age of 65, the risk of developing the disease is increased.
- Smoking: People who are used to smoking have a higher risk risk of having the disease.
- High blood pressure: When blood pressure is uncontrollably high, arteries can harden and clot, narrowing the passage for blood flow.
- High blood cholesterol levels: High blood cholesterol increases the risk of plaque build-up and atherosclerosis. High cholesterol can be caused by high levels of low-density lipoprotein (LDL) cholesterol, known as "bad" cholesterol. Low levels of high-density lipoprotein (HDL) cholesterol, known as "good" cholesterol, may also help to improve atherosclerosis.
- Diabetes: Diabetes is one of the reasons that increases the risk of this heart disease. Type II diabetes and CHD share some similar risks, such as obesity and high blood pressure.
- Obesity: Excess weight effects rest of the risk factors.
- Physical inactivity: Lack of physical exercise is related with this disease and other risk factors.
- High stress: Uncontrolled stress on your health can damage your blood vessels & cause other risk factors for this heart disease.
- Unhealthy diet: Eating a unhealthy diet full of saturated fats, trans-fats, salts & sugar may increase the risk of this heart disorder.

Coronary artery disease is the most common heart disease, causing 365,914 deaths in 2017. 10.9% of adults around the age of 45 were having heart disease. 7.1% of adults between age 45 to 64 years reported having CHD, compared to 17% of adults aged 65 years and older. 13.6% of men were having CHD as compared to 8.4% of women. According to research conducted, CHD varies between racial/ethnic groups: 11.4% whites, 10.0% blacks/African Americans, 8.8 % Spanish, and 6.3% of Asia & Pacific Islands. The prevalence of CHD has lowered in adults with more years of formal education (10.9% with high school education) compared with adults with fewer years of education (16.5% with a diploma).

As the number of electronic health records (EHRs) has grown, so has the demand for advanced technology to retrieve them. Machine learning is the process of automatic data analysis by building analytical models. It is a part of artificial intelligence

based on the concept that models can hold from data, recognize patterns, and make predictions with little human intervention. Machine learning and artificial intelligence concepts used in various medical platforms like in maintaining, fetching and retrieving patient's data, early and accurate diagnosis of various diseases based upon the symptoms of the suffering person, etc. These rapidly advancing concepts save time in diagnosis and detection, thereby extending patient life.

Simple supervised machine learning methods are Linear Regression, Naive Bayes, Logistic Regression, Decision Tree, Support Vector Machine, K nearest neighbor, etc. Linear regression is the most famous mathematical machine learning method which is used for prediction and analysis. Linear regression makes predictions of continuous/real or static variables such as salaries, sales, commodity prices, age, etc. Logistic regression is ML algorithm used to predict the probability of independent variables. In Logistic regression, firstly linear regression is implemented and then using sigmoid function output is obtained. It is only applicable for binary class classification. The Naive Bayes classifiers are one of the simplest & most effective ways to help create fast-learning machine models that make predictions faster. It is only applicable for binary class classification. The Naive Bayes classifiers are one of the simplest & most effective ways to help create fast-learning machine models that make predictions faster. It is a probabilistic classification model, meaning it makes predictions based on object probabilities. This algorithm is widely used to solve text classification problems and has multiple class problems. Support vector machines (SVMs) are ML algorithms that can be implemented for both regression and classification. The Vector Machine makes classification by detecting a hyper-plane that increases the boundary between the

2 classes. The vectors defining hyper-plane are called supporting vectors. The K Nearest Neighbors (KNN) technique falls into the category of supervised ML and is used for classification (most common) and regression. KNN is a flexible algorithm that is also used for generating missing values and re sampling data. The KNN method captures the similarities between new case/data and available cases and places new cases in categories that are very similar to existing categories. Decision Tree is a supervised ML model which in the form of a tree. Split the dataset into smaller subsets and evolve appropriate decision trees at the same time. The final result is a tree with decision nodes (with two or more branches) and leaf nodes (representing decisions). The top decision node of the tree represents the most desirable predictor, called the root node.

These simple model types are ideal for predicting one or two features with high prediction accuracy. Therefore, merging these models gives a model with high predictive accuracy. This is called ensemble learning. In machine learning, ensemble techniques use many learning models to achieve better predictive performance than individual learning models. An ensemble learning model is a meta-algorithm that integrates multiple machine learning techniques into a predictive model to reduce variance (bagging), bias (boosting), and improve

prediction (stacking). These methods include bagging, stacking, voting, boosting and pasting. Voting is an ensemble learning technique that averages multiple regression models to make predictions. Another name for bagging is bootstrap aggregation. It is a method of minimizing measurement error by averaging multiple estimates. Bagging uses a bootstrap example to take a subset of the data and train a base learner. Boosting is a group of different algorithms that turn weak learners into strong learners. The main purpose of boosting is to fine-tune a sequence of weak learning models that are only marginally better than random guesses such as such as small decision trees in versions with data weight. More weight is given to samples that were poorly separated from the previous round. Stacking is an ensemble learning approach involving multiple classification/regression models using meta-classifiers/metaregressors. Base model is trained on the full set of training data, and meta model is trained as a feature on the base model.

Feature selection approach is a way of minimizing the number of input variables when making a predictive model. In feature selection, the data contains many unimportant features. Redundant features provide more information than selected features, and insignificant features provide no useful information. Feature selection methods can be performed using filter, wrapper, and embedding methods. The filter method is implemented by Ant Colony Optimization (ACO) and the wrapper method implementation is done by Particle Swarm Optimization. ACO is a method of finding the optimal path based on the ant's natural behavior when foraging. First, the ants roam randomly. When an ant finds a food source, it returns to the colony and leaves a "sign" (pheromone) indicating the presence of food in its tracks. ACO is an optimization algorithm modeled after the behavior of an ant colony. The artificial ant finds the optimal solution after moving along the parameter space presenting various possible solutions. Particle Swarm Optimization (PSO) is an optimization technique based on the movement of schools of birds and fish. PSO provides many analogies for future computing strategies. The program is implemented with many arbitrary solutions, and the optimal solution is obtained by updating the generation.

In the given base paper, the authors Sunanda Dixit, Jabbar Akhil and Indu Yekkala performed feature selection method(PSO) on the different dataset having 270 instances along with three ensemble classifiers(Random Forest, AdaBoost and Bagged Tree). Accuracies of PSO + Bagged Tree, PSO + Random Forest and PSO + AdaBoost are 100%, 90.37% and 88.89% respectively. Elma performed K-Nearest Neighbor on the same dataset with 82.6% accuracy. Jabbar performed One-R and Naive-Bayes on Andhra Pradesh's dataset with an accuracy of 86%. Daniel performed only NaiveBayes with an accuracy of 99%. In this paper, firstly all the necessary libraries are imported and then the mentioned coronary heart disease data set is loaded. Further data cleaning and data preprocessing are performed. This includes changing numeric features to categorical features, checking if the data set has null values, unique sample counts, etc. The next step is data visualization, which shows pie charts, pair charts,

histograms, and count charts. Next, the main approach of our paper is model building with ant colony optimization (Feature Selection Technique). Various ensemble learning techniques such as AdaBoost, Random Forest, Extra Trees, Bagging, Gradient Boosting and his ACO are implemented to compute accuracy, precision value, recall value and F1 score for each model.

2. RELATED WORK

As per given base paper, authors Jabbar Akhil, Sunanda Dixit and Indu Yekkala performed Particle Swarm

Optimization as a feature selection process on a different data set with 270 samples and three ensemble learning classifiers (Random Forest, AdaBoost and Bagging Tree) to improve heart rate prediction. Firstly data set was loaded. After that data cleaning was implemented to the updated dataset. Further PSO is implemented and least important features are removed and the predominated ones are considered. On predominant features, ensemble learning classifier is implemented. Finally, accuracies of these models is calculated. Accuracy of PSO+Bagged Tree is 100%, PSO+Random Forest is 90.37% and PSO+AdaBoost is 88.89%.

Elma Zannatul, Ferdousy has proposed an approach that combines two models, the K Nearest Neighbor & Naive Bayes Classifier for heart disease prediction, to improve performance. The accuracy of the heart disease same data set that we used was 82.6%. M.A. Jabbar predicts Andhra Pradesh’s heart disorder data set by implementing Association Classification algorithm and found that 10.26% improvement occurred over J4.8 and 8.6 % improvement compared to the Naive Bayes algorithm. Hlaudi Daniel Masethe developed and compared algorithms J4.8, REPTree, CART, Naive Bayes, and Simple Cart for heart disease prediction. The author used the confusion matrix to compare the above algorithms. He performed Naive Bayes algorithm with 99% accuracy.

Table 1.

ACCURACY AND TECHNIQUE OF DIFFERENT PAPERS

Author	Technique	Accuracy
Sunanda , Jabbar and Indu	PSO + Bagged Tree	100%
	PSO + Random Forest	90.37%
	PSO + AdaBoost	88.89%
Elma	K – Nearest Neighbor	82.6%
Jabbar	One – R and Naive – Bayes	86%
Daniel	Naive – Bayes	99%

Fahd Saleh Alotaibi designed the ML model using the Rapid Miner tool, which delivers high accuracy. He also used the same UCI heart disease dataset that we used. In his work, he computed the accuracy of Logistic Regression, Naive Bayes, Random forest, Decision Tree & Support Vector Machine models, and found that the accuracy of decision tree algorithm is very high. Karoalis developed a C4.5 algorithm for decision-making of heart disorder and performed classification of PCI & CABG models. Highest accuracy obtained was 75%. Avinash Golande explored various ML techniques that are used for CAD prediction. He studied on Decision Tree, KNN and K-Means which are suitable for prediction on the same

UCI heart disease dataset that we used. In addition, we found the accuracy of each. His research concludes that Decision Tree accuracy was much higher. He further concludes that accuracy could be improved using various other techniques & parameter tuning.

T.Nagamani has proposed a number of techniques using data mining technique and the Map Reduce algorithm on the same UCI heart disease dataset that we used. Accuracy is calculated for 45 cases in the test set, that was higher than the accuracy calculated using a standard neural networks. In his work, the accuracy of the algorithm used has been improved because of the use of a powerful schemes and linear scaling. The accuracy was 98.12%. Anjan Nikhil Repaka proposed a method for classification of data set and AES (Advanced Encryption Standard) technique and have secure data transfer to predict the disease. Theresa Princy R. performs various ML models for predicting disease. The classification strategies used are Naive Bayes, K- Nearest Neighbor (KNN), Neural network (NN), Decision Tree, and the accuracy of each model was achieved with a different number of features. Nagaraj M. Lutimath improved prediction of heart disease by applying naive Bayes and support vector machines (SVM). He performed metric analyzes such as mean absolute error, sum of squared errors, and root mean squared error. He found that the accuracy of SVM is much better than Naive Bayes.

3. THEORETICAL BACKGROUND

3.1. Ant colony optimization (ACO)

The Ant colony optimization (ACO) algorithm is promoted by the natural ant behavior in their search for food, and targets the problem performance differently. Accumulation occurs through indirect contact mediated by the introduction of aromatic substance(pheromones) into the food pathway. To efficiently solve this problem, many artificial ants are used in constructing solutions. At each iteration, the ants incorporate a fixed amount of pheromones corresponding to the required solution. At each step, each ant computes a different set of possible solutions to the current incomplete solution, and two factors: current position, current velocity and pheromone levels, and previous information. ACO is a disseminated technique in which some set of agents work collectively to obtain a feasible solution. This is a filter-based approach that gets a good set of features through multiple iterations. ACO is used for checking space of each and every subsets of a given feature set. Performance and accuracy for a selected subset of features are calculated by requesting test features in the appropriate feature space and measuring specific results. The best subset of features selected will be the set of recommended features that should be used in the actual construction of the classification problem.

The main steps of proposed feature selection algorithm are:

1. Initialization: Find the number of ants. Fix the amount of pheromone related with any of the feature. Determines the maximum number of iterations allowed.
2. Solution production and testing of ants: Appoint ants aimlessly to one feature & visit the features of each ant to formulate complete solutions. At this particular step, the test condition refers to a Mean square error (MSE) for the classifier. If the ant is

unable to lower the MSE in a n consecutive steps, it would complete its function and then exit.

3. Check selected subsets: Arrange selected subsets according to the accuracy of the classifier & this length is then selected as the best subset. Probabilistic transition rule indicates the possibility that the ant k will incorporate the feature i to the feature j in the solution during time t.

$$P_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \in J_i^k} [\tau_{il}(t)]^\alpha \cdot [\eta_{il}]^\beta} & j \in J_i^k \\ 0 & \text{otherwise} \end{cases}$$

4. Phormone Updating: Revise the concentration of phormone for the nodes, then each ants deposit the phormone value in the graph. Eventually, allowing best ant to add additional phormone to the nodes.

$$\tau_{ij}(t+1) = (1 - \rho) \cdot \tau_{ij}(t) + \sum_{k=1}^m \Delta_{ij}^k(t)$$

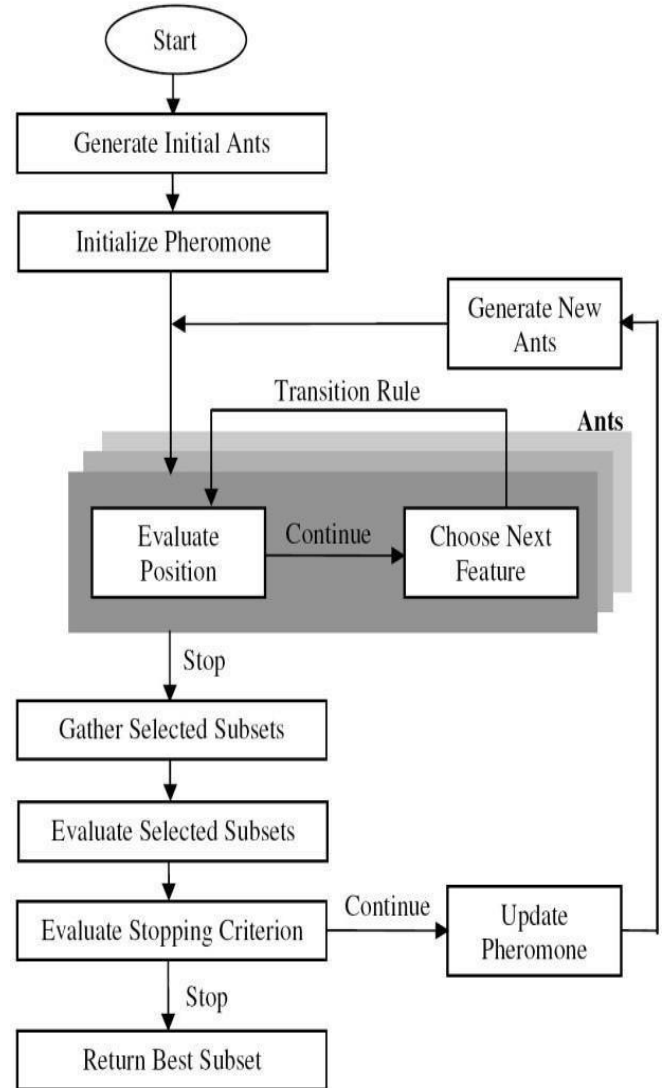
$$\Delta_{ij}^k(t) = \begin{cases} \gamma'(S^k) / |S^k| & \text{if } (i, j) \in S^k \\ 0 & \text{otherwise} \end{cases}$$

5. Check conditional stopping criteria: Stop, if the repetition value exceeds the maximum allowable value, otherwise proceed.
6. Production of new ants: At this step the old ants are eliminated new ants are produced. Return to step 2 & then continue.

3.2. Bagging

The concept at the back of bagging is merging the effects of multiple models (e.g., all types of naive bayes) to obtain a standard output. There is a good chance that these types of models gives the same output as they get the same input. For solving this one method is bootstrapping. It is a method of making samples where we produce subsets of the results obtained from the actual data set, by replacing it. Size of subsets is similar to the actual set size. Bagging method make use of subsets (bags) to get correct idea for distribution. The size of the subsets designed for bagging may be less than the actual set. Steps to apply Bagging algorithm:

1. Many subsets are made from the actual data set, selecting the results obtain with replacement.
2. The weak model is generated for each of the subsets.
3. Models work in parallel & are not dependent on each other.
4. Main last predictions are obtained by merging predictions from different models.



3.3. Boosting

If the data instant is inaccurately estimated by first model, & then calculating for other models, assembling the prediction will not give more acceptable results. In such a case boosting is helpful. Boosting is a linear process, in which each successive model tries to rectify the errors in preceding model. These succeeding models depends upon preceding model. Steps to apply boosting algorithm:

1. Subset is originated from actual data set.
2. All data instances are assigned with equal weights.
3. A weak model is been formed by using this subset.
4. This weak model is used for making predictions on the entire data set.
5. Errors are computed using original values & predicted values.
6. Observations and predictions obtained incorrectly, are given high weights (there are three misclassified blue plus points which are given high weights).
7. Second model is developed & predictions are calculated for the data set. (This model is trying to fix errors in the previous model)
8. Alike, many models are developed, each rectifying the errors of preceding model.
9. Final model (strong learner) is the weighted mean of all preceding models (weak learners).

Therefore, boosting algorithm merges many weak learners to

build one strong learner. Individual models will not work well for the entire data set, but will work well for a particular part of the data set. Therefore, each model literately uplifts the performance and accuracy of the model.

Bagging algorithms:

1. Bagging meta-estimator
2. Random forest
3. Extra trees

Boosting algorithms:

1. AdaBoost
2. Gradient Boosting (GBM)

4. PROPOSED APPROACH

In our paper, the approach we used is to apply Ant Colony Optimization as Feature Selection Methodology along with different Ensemble Learning Methods(AdaBoost, Bagging, Random Forest, Gradient Boosting and Extra Trees). Feature Selection Method reduces number of features and on the predominant features the ensemble learning model are applied and performance metrics are obtained. The suggested approach used Heart disease dataset with total 303 instances from UCI Machine Learning Repository. The data set was already preprocessed and all irrelevant and missing data was already removed. This dataset had 76 attributes but we used 13 important attributes and one target variable. This data set is a clean and simple dataset with 13 features and one target. The target has only 2 unique values (0 and 1) therefore simple binary classification models can be easily applied.

Table 2.

SPECIFICATION OF HEART DISEASE DATASET

Dataset	No. Of Instances	Features
Heart Disease	303	14

Steps of our approach are given below:

1. Importing Libraries

Importing all the necessary libraries.

2. Loading Dataset

Loading the required Dataset as Dataframe.

3. Data Cleaning and Preprocessing

Replacing the column names of the dataset with their actual names. Converting the numeric features to categorical features. Checking missing entries in the data set columnwise. This Data set has no null values. Describing numeric and categorical features of the data set.

4. Data Visualization

Plotting Pie Chart and Count plot for Target feature. Plotting Pie Chart and Distribution plot for features Gender and Age respectively. Plotting Age Distribution Plot for normal and heart disease patients. Plotting Gender Bar Chart for normal and heart disorder patients. Plotting Count Plot for Chest Pain Type, Fasting Blood Sugar, Rest ECG, Exercise Induced Angina, ST Slope and Thalassemia Features. Plotting Pair Plot for all numeric features. Plotting Histogram for Features.

5. ACO Feature Selection Model Building.

Applying Gaussian Naive Bayes Model, Support Vector Machine Model, AdaBoost Model, Bagging Model, Random Forest Model, Gradient Boosting and Extra Trees individually and then apply all the above models with ACO. Calculate Accuracy, Precision Value, Recall Value and F1 Score for each

model. As per table III shown below, the highest accuracy achieved in our paper is 95.08%. Comparison of Accuracy, Precision, Recall and F1 score for all the above models.

Table 3.

TECHNIQUE, ACCURACY, PRECISION, RECALL AND F1 SCORE IN OUR PAPER.

Technique	Accuracy	Precision	Recall	F1 Score
<i>Gaussian Naive Bayes</i>	86.88%	0.875	0.875	0.875
<i>Support V ector Machine</i>	78.69%	0.7879	0.8125	0.8
<i>AdaBoost</i>	81.97%	0.8888	0.75	0.8135
<i>Bagging</i>	81.97%	0.8621	0.7812	0.8197
<i>Random Forest</i>	83.61%	0.9231	0.75	0.8438
<i>Gradient Boosting</i>	83.61%	0.8438	0.8438	0.80
<i>Extra Trees</i>	90.16%	0.9062	0.9062	0.9062
<i>CO + Gaussian Naive Bayes</i>	91.80%	0.9259	0.7812	0.8474
<i>ACO + Support V ector Machine</i>	85.24%	0.6383	0.9375	0.7595
<i>ACO + AdaBoost</i>	88.52%	0.8214	0.7188	0.7667
<i>ACO + Bagging</i>	95.08%	0.9524	0.625	0.7567
<i>ACO + Random Forest</i>	91.80%	0.84	0.6562	0.7368
<i>CO + Gradient Boosting</i>	90.16%	0.8276	0.75	0.7869
<i>ACO + Extra Trees</i>	91.80%	0.8387	0.8125	0.8254



Fig. 1. Accuracy, Precision Value, Recall Value and F1 Score of our paper without ACO

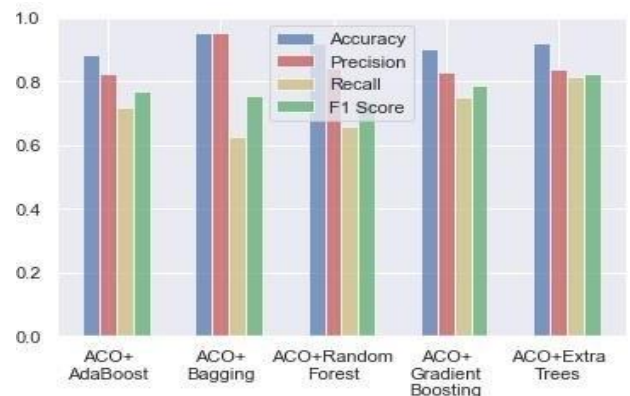


Fig. 2. Accuracy, Precision Value, Recall Value and F1 Score of our paper with ACO

5. DATA DESCRIPTION

Table IV represents feature name & its description of data.

Table 4.

DESCRIPTION OF DATASET

FEATURE	FEATURE DESCRIPTION
Age	Age in Years
Sex	(1 = Male) ; (0 = Female)
cp	Chest Pain Type
trestbps	Resting Blood Pressure(in mmHg)
chol	Serum Cholestrol in mg/dl
fbs	Fasting Blood Sugar and gt; 120 mg/dl (1 = true; 0 = false)
restecg	Resting Electrocardiographic Results
thalach	Maximum Heart Rate Achieved
exang	Exercise Induced Angina (1 = yes; 0 = no)
oldpeak	ST Depression Induced by Exercise Relative to Rest
slope	The Slope of the Peak Exercise of ST Segment
ca	Number of Major V essels Colored by Flourosopy
thal	emia; 1 = normal; 2 = fixed defect; 3 = reversable defect
target	1 or 0

Table 5

COUNT OF EACH DIFFERENT LABEL IN TARGET

LABEL	COUNT
1	165
0	138

This data set comprises of 13 features one target . It consists 7 categorical features and 6 numeric features. If target variable has value 1 then it means that the patient is suffering from the heart disorder and if it’s value is 0 then it means that patient is not having the disease.

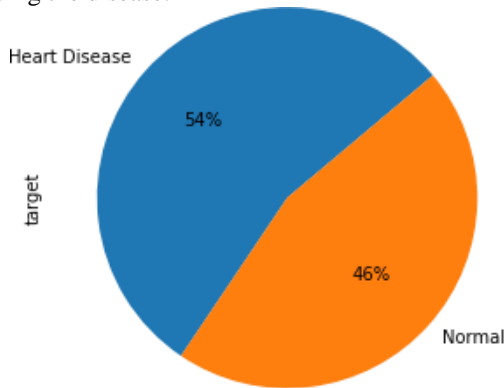


Fig 3. Pie Chart for the label 'target'

There are 303 instances and for target variable the count of each unique (0 and 1) label is represented below

6. CONCLUSION

In our paper, we explored prediction of coronary heart disease based upon Ant Colony Optimization Technique and ensemble classifier. Our approach was to use ACO as Feature Selection Methodology that reduces number of features on predominant features the ensemble learning model are applied and performance metrics are obtained. As per exploratory results it was seen that accuracy of Bagging ensemble learning along with ACO that is 95.08%. This model might help medical employees and staff as easy diagnosis prognosis is quickly achieved.

Table 6.
COMPARISON OF OUR WORK WITH PRIOR WORK

Author	Technique	Accuracy
Sunanda , Jabbar and Indu	PSO + Bagged Tree	100%
	PSO + Random Forest	90.37%
	PSO + AdaBoost	88.89%
Elma	K – Nearest Neighbor	82.6%
Jabbar	One – R and Naive – Bayes	86%
Our Proposed Approach	ACO + AdaBoost	88.52%
	ACO + Bagging	95.08%
	ACO + Random Forest	91.80%
	ACO + Gradient Boosting	90.16%
	ACO + Extra Trees	91.80%

REFERENCES

- [1] HlaudiDanielMasethe,MosimaAnnaMasethe, “Prediction of Heart Disease using Classification Algorithms”, Proceedings of the World Congress on Engineering and Computer Science, 2014 Vol II WCECS 2014
- [2] M.Karaolis, Member, IEEE, JAMoutiris, FESC, C.S.Pattichis, Senior Member, IEEE, “Assessment oftheRiskofCoronaryHeartEventBasedonDataMining”
- [3] Avinash Golande, Pavan Kumar T, “Heart Disease Prediction Using Effective Machine Learning Techniques”, International Journal of Recent Technology and Engineering, 2019
- [4] T.Nagamani, S.Logeswari, B.Gomathy, “Heart Disease Prediction using Data Mining with Mapreduce Algorithm”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2019
- [5] Fahd Saleh Alotaibi, “Implementation of Machine Learning Model to Predict Heart Failure Disease”, (IJACSA) International Journal of Advanced Computer Science and Applications, 2019
- [6] Anjan Nikhil Repaka, Sai Deepak Ravikanti, Ramya G Franklin, “Design And Implementation Heart Disease Prediction Using Naives Bayesian”, International Conference on Trends in Electronics and Information (ICOEI), 2019
- [7] Theresa Princy R.J. Thomas, “Human heart Disease Prediction System using Data Mining Techniques”, International Conference on Circuit Power and Computing Technologies, 2016
- [8] Nagaraj M Lutimath,Chethan C,Basavaraj S Pol., “Prediction Of Heart Disease using Machine Learning”, International journal Of Recent Technology and Engineering,2019
- [9] UCI, Heart Disease Data Set.[Online]. Available (Accessed on October242022):<https://archive.ics.uci.edu/ml/datasets/heart+Dis+ease>

Image Manipulation Detection System

Vinita Rohilla

Assistant Professor, Department of Computer Science & Engineering MSIT, New Delhi, India

Abstract— In today's professional world, impressive images can be an essential part of many applications. The definition of fake image is to use advanced details to hide important information or create false information. Advanced image imitation reached its zenith with the advent of cutting-edge image processing products. Print-oriented paper is a way for most pictures to lie. The recommended method is to discern and report inappropriate behavior in modern computer graphics.

Keywords— *Image Manipulation, Image forgery, Image Tampering, Digital images.*

1. INTRODUCTION

A digital image is a carefully readable photograph or image. It can be a digital representation of a two-dimensional image, usually parallel [1]. The best graphics can be the key to sharing information over a wide network used in almost any environment. Some computer graphics may involve understanding business laws or even national security. The development of network and interactive media has facilitated the transmission path [2].

Image processing is used to convert images into computer images and then the information needed from them is given. Image rendering is largely composed of images, photos, videos, etc. it's about doing. Graphic design is a type of logo design where image is taken as input like a photo or video profile, and the output is an image or a set of elements that are important or unrelated to the image. Image processing may include image scaling, image enhancement image segmentation or video or image compression.

Due to advances in technology and advances in computer software for image processing, it is difficult to distinguish real images from fake images. This can introduce skew imperfections and reduce the reliability of high-resolution images. Therefore, it is necessary to create a framework that should be used to check the accuracy and judgment of computer images.

2. LITERATURE SURVEY

Hesham A. et al. [1] stated that a particular image can be altered in any way. Photographs are an important part of the judicial process. For copyright detection, a fake image detection program is used to identify the created image. Different actions can be requested to view and follow the contents of premium images.

Image mimic localization analysis as a dynamic discrete concept. Dynamic rules depend on flags or watermarks. The idea is based on information from the first image. On the other hand, the strategy of looking at the image without using prior knowledge does not work better and we have to choose the question of how to manage the image.

2.1. Copy-Move Forgery

The repeating motion number can be a way to create a composite image by eliminating the problem in one image and replacing it with another image.

In order to hide many things in design, it is necessary to create something modern or to use spaces sewn from the same or different images or images. Often the purpose of this is to hide the problem in the image by filling it with a part copied from another part of the image.

Zaid et al. [2] said that in computerized photo manipulation spoofing, one or more of the same images are reanimated in different places. Often repetitive scenes are enlarged, scaled down, or made into a more believable imitation, making it difficult to define a parody.

Alberry et al. [1] demonstrated an effective way of proving repeated fraud in computer graphics using scale invariance, including transformations (filters) and fluffy C-means (FCM), and using the aggregation mode.

2.2. Double Jpeg Compression Detection

In JPEG images double compression occur when they are compressed into spatial space and then re-saved using a special (auxiliary) quantization mesh. Editing JPEG images usually involves recompression.

Wang, Q., Zhang, R. et. al [3] pointed out that Bi-fold JPEG compression has been shown to be useful in image confusion forensics. However, there are several strategies that provide programmatic localization.

To overcome this challenge, this article presents a convolutional neural permutation (CNN) based detection algorithm for two JPEG compressions. One explanation for CNN is to split the discrete cosine transform (DCT) coefficients histogram that differ between one compressed area (the transformed area) and two compressed areas (the untampered area).

The aggregated results are consistent with the results of classification. The proposed calculation performs well with both JPEG compression localization and simulated localization, especially when first compression quality number is higher than the second compression quality number.

2.3. Noise Variance Inconsistency

The most common use to cover up signs of change is to attack. Diversity or contrast can also be caused by other factors such as colour, light or space. The first picture is mostly filled with homogeneous noise. Therefore, finding stress conflicts in the image can reveal the area of control.

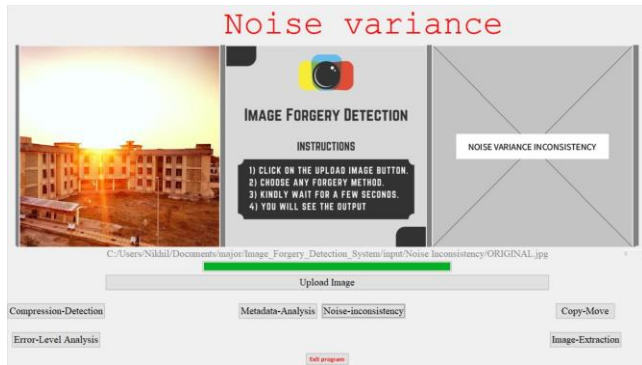


Fig.1. Noise variance inconsistency detection system

Savita Walia, Mandeep Kaur et al [4] state that in their ponder they center on the location of a basic and common picture control strategy known as picture grafting, where a produced picture is made by compositing districts from distinctive source pictures. Thier picture grafting location strategy is based on the reality that most computerized pictures contain clamor presented either amid securing sensor or quantization commotion or consequent preparing e.g., compression. For normal un-tampered picture, the clamor changes over diverse districts regularly contrast as it were slightly.

According to Yao, Heng. ,Cao, Tooth., et al. [5]. due to the difficulty of totally isolating picture commotion and characteristic surface, unavoidably, each gauge is overestimated, particularly for districts that have more complex textures. Based on this thought and spurred by the strategy-based or object-proposal procedure, the approach that joins the inhomogeneity scoring methodology is proposed to supply a more persuading result to uncover picture- grafting controls. Particularly, to begin with, the picture is fragmented into little patches, and clamor fluctuation of each fix is computed by utilizing the kurtosis pixel-level commotion estimation method.

2.4. Metadata Analysis

The information and content saved for an image is called image metadata. Include production dates, producers, author names, terms, topics, and other subtleties as much as possible. Root device, bootloader, device, time, compression content etc.

2.5. Error Level Analysis

Error Level Analysis, an error can occur after re-saving the image with the specified compression ratio when the process checks for changes in the compressed image (JPEG) for annotation analysis. In the ELA, image is saved with 95 percent compression and results are compared with original image. Due to the unique characteristics of the ELA region, areas of change are rapidly becoming evident. The ELA calculation includes variables such as shadows, eyes, and reflection to detect unwanted content in the input image. Sudiatmika et al. [6] state that it will actualize a modern framework that can recognize imitation pictures and unique images with profound learning. When seen with the bare eye the picture looks the same, but by utilizing this strategy it'll

see the contrast between a fraud picture with the first picture. Calculation for the normal contrast of the quantization table Y (luminance) and $CrCb$ (Chrominance). Unique pictures from computerized cameras ought to have tall ELA values. Each ensuing resave will diminish the potential mistake rate.



Fig.2. Input picture for ELA

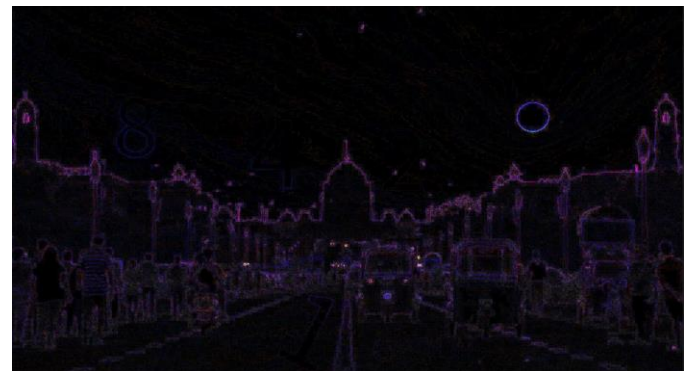


Fig.3. Output of ELA for proposed system

2.6. Image Extraction

The most common masking material for steganography is an image. The Advanced Photo area has several different sets of images, most of which are used for specific articles. There are special steganographic calculations for these different types of data. Image steganography hides information that the human eye can poorly perceive.

Image subtraction is a way to extract mask data.

Misra et al. [7]. centered on the slightest noteworthy bit steganography calculation, subsequently expanding its execution by proposing an adjusted slightest noteworthy bit steganography called circular move LSB. The consider too explained the LSB settling in method and displayed the evaluation results for LSB content steganography for different cover protest designs such as .png, .bmp, and .jpeg. A comparative appraisal was carried out to set up the value of the recommended approaches. Built on the LSB method, an 8 or 24-bit color image calculation is set up to boost the stego-image exactness of the color question capable in creating a covered up concealed question that's completely intangible to the human eye.



Fig.4. Input picture for picture extraction

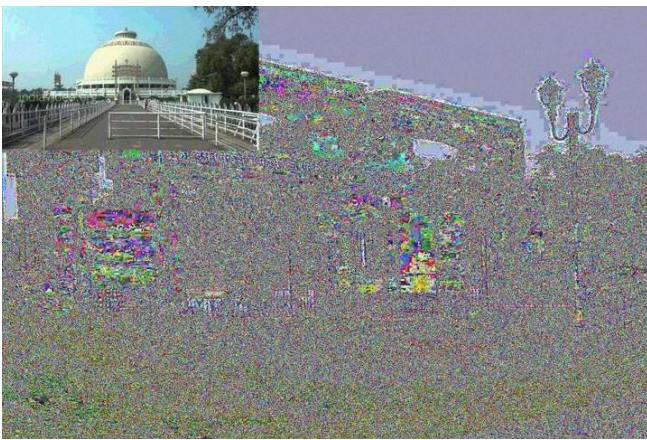


Fig.5. Output for picture extraction utilizing the proposed system.

3. CFA Artifact Detection

CFA consists of mosaic of spectrally specific channels arranged in an interactive way so that each pixel registers one of the correct spectral features. Fake images can be identified by identifying artifacts in images that have been further manually cleaned. An image taken by a computer camera would show the demosaic features of each pixel group without further processing compared to the CFA component.

Egorova et al.[8] proposed strategy is based on inclusion testing showing the presence or absence of DFA structures for each square image. The results obtained include the characterization of each placed block. Some plans do not require additional information to be added to the table. They use a strategy that uses the device's sensors to collect images to detect fraud. One such feature is color channel grouping (CFA) artifacts. These are the ensemble objects in the picture resulting from the close-up of the CFA filter on the camera that took the picture. A CFA can be a mosaic of secondary colors placed on the pixel sensor of the image sensor to capture color information. Most cameras today see this. Note that CFA constructs are available on all camera screens. Although fake images appear a little on the surface, they change the dimensional properties. In particular, they eliminate the pixel-to-pixel correlation that occurs when

taking RGB images.

Involves calculating the ratio between changes in estimated error of measurement received and affected. Based on the degree values obtained, it is possible to determine whether there are CFA artifacts in the image. Let the analysis image be $N \times N$ in size, at this point we are ready to calculate the $B \times B$ estimated values for each non-uniform square image. For the frame estimate, the measurement should correlate with the duration of the Bayer channel with the minimum time and block estimate 2×2 .

4. PROPOSED METHODOLOGY

System Design portrays “the by and large structure of the framework and the ways in which the structure gives conceptual integrity”. The given framework actualizes different procedures to algorithmically determine various sorts of alterations that can be done to computerized pictures. The different strategies are actualized as:

4.1. Noise variance inconsistency

The primary step is to prepare the picture Changing over the picture to grayscale in the event that it is in color. This disentangles the investigation by centering as it were on the concentrated values and resizing it. At that point the moment step is to separate the image into littler patches. We utilize the fix sizes of 8×8 . At that point another step is to calculate the neighborhood commotion fluctuation for each fix, this may be done by measuring the pixel escalated varieties inside the fix. We apply a Laplacian edge location bit to the picture, computes

the entirety of the supreme values of the convolved picture, and scales the result to assess the clamor level within the picture. At that point the final step is to Compare the clamor fluctuations of neighboring patches. We utilize K-means calculation for it and determine if the fluctuation within the clamor of the picture is critical sufficient or not.

4.2. Double jpeg compression

To begin with we alters the measurements of the image to be distinct by 8 and after that change over the picture to a distinctive color space and extricate the luminance channel. At that point luminance channel is isolated into 8×8 squares for advance preparing and we calculate the discrete cosine change (DCT) for each block. The DCT coefficients are handled to normalize and circular them and after that the ghostly characteristics of the DCT coefficients are analyzed. At that point we compute histograms and perform FFT on the DCT information. The program at that point checks the number of crests within the changed data and decides on the off chance that numerous compression components were used.

4.3. Error Level Analysis:

In this strategy the program analyzes the blunder level between the first and temporary versions of an picture and visualizes the contrasts to supply experiences into potential modifications or modifications within the picture. The primary step in this prepare is to form a brief form of the picture with diminished quality. At that point The program calculates the contrast between the first picture and the brief picture. At that point the following step is to heighten the color contrasts within the contrast picture by scaling the pixel values. At last, the distinction picture is shown in a partitioned window.

4.4. CFA Detection

The primary step is to obtain picture parts employing a indicated piece length and after that outline a color esteem to a particular extend in a color palette. At that point Extricate pixel values for each piece and store them as picture parts. At that point we discover comparable parts inside the picture and cluster the comparative parts into bunches. The following step is to mark the comparative clusters on the picture that speak to the cfa artifacts. The ultimate step is to appear the altered picture and return the number of indistinguishable districts found.

4.5. Image Extraction

To begin with we must get the stego picture that we suspect to be carrying covered up subtle elements. Anotherwe decide the LSBs of the pixel values within the stego picture, these are the bits that have possibly been adjusted to stow away the data. At that point the following step is to extricate the covered up bits by iterating through the pixels of the stego picture and extricate the LSBs from each pixel. Following we concatenate these extricated LSBs to make the covered up bitstream and change over the bitstream to the initial organize. On the off chance that the covered up message was in a parallel arrange, the extricated bitstream is as of now the covered up message. In any case, in the event that the message was in a distinctive arrange, such as content or an picture, we at that point change over the double bitstream back into its unique representation.

4.6. Copy move forgery

In this procedure we utilize the Filter (Scale-Invariant Include Change) calculation to distinguish duplicate move imitation. The primary step is to extricate highlights from the picture utilizing the Filter calculation. This includes recognizing keypoints and computing descriptors for each keypoint. The another step is to coordinate the keypoints between diverse regions of the picture. This may be done utilizing different strategies such as brute-force coordinating. Once the keypoints have been coordinated, the following step is to cluster them into bunches based on their spatial nearness. At long last, the clusters are confirmed to decide if they speak to a copy-move fraud or not.

4.7. Metadata Analysis

The primary step is to Extricate the metadata from the picture record. Most picture record groups, suchas JPEG, PNG, and TIFF, store metadata in standardized designs, such as Exif (Replaceable Picture Record Arrange) or XMP (Extensible Metadata Stage). At that point store the extricated metadata in a content record and at that point get to the metadata utilizing computer program libraries or apparatuses that can parse the particular record arrange and extricate the significant data. Analyze the key metadata areas that are commonly related with picture control like Camera make and demonstrate, Date and time of capture, Program and altering history

5. RESULTS AND DISCUSSIONS

The framework is useful for identifying and extracting double JPEG compression detection, copy space displacement, CFA artifact discovery, error detection, image extraction, and metadata analysis in the input image. The homepage UI includes a display box and an income box. The center provides customers with nine different smart buttons and useful inspiration to ensure their comfort. Users should click

the "Load Image" button which will open a popup file manager box allowing them to select an input image from their local drive.

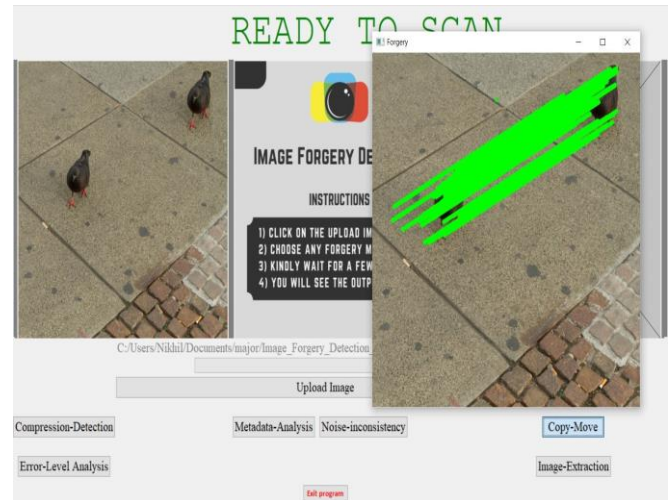


Fig.6. Yield Of Duplicate Move Imitation Detection

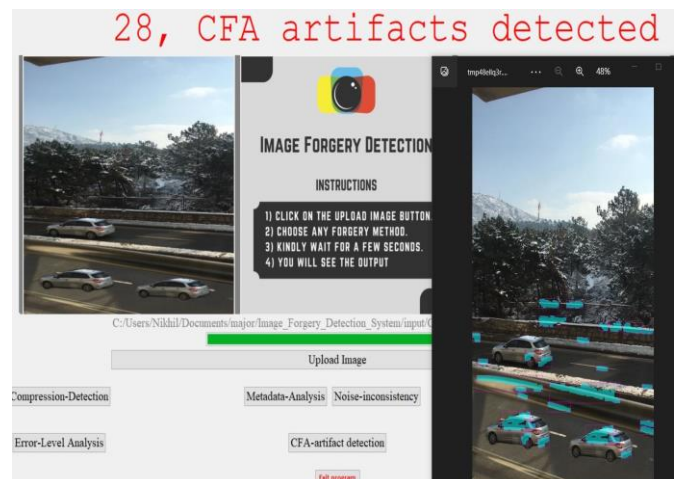


Fig.7: Yield Of CFA Artifact Detection

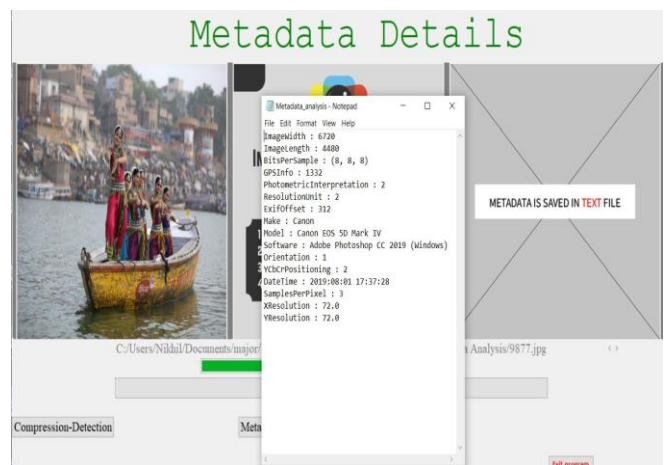


Fig.8: Yield Of Metadata Analysis

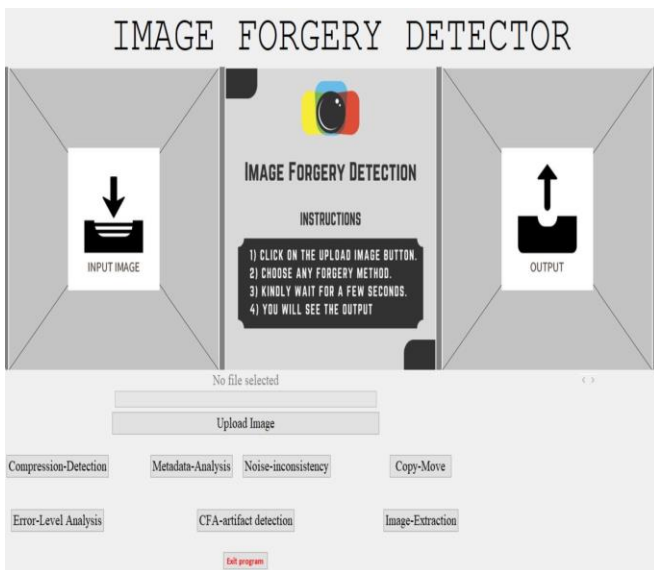


Fig.9: Homepage UI

The framework permits the client to select any of the methods and produce suitable yields through yield window on the UI or pop-up windows or within the shapes of pictures. The framework works locally on the framework without the necessity of network access.

6. CONCLUSION

In the proposed framework, we have executed eight distinctive concepts of picture imitation location calculations. The framework is competent of taking input picture and grant out appropriate yields to illuminate the obstacle of manufactured pictures. The system can be utilized in law and authorizations and cyber security to assist the client to distinguish between authentic and altered images.

REFERENCES

- [1]. Hesham A. Alberry, Abdelfatah A. Hegazy, Gouda I. Salama, "A fast SIFT based method for copy move forgery detection", *Future Computing and Informatics Journal*, Volume 3, Issue 2, 2018, ISSN 2314-7288,
- [2]. Zaid Nidhal Khudhair et al, "A Review on Copy-Move Image Forgery Detection Techniques", (2021) *J. Phys.: Conf. Ser.* 1892 012010
- [3]. Wang, Q., Zhang, R. "Double JPEG compression forensics based on a convolutional neural network", *EURASIP J. on Info. Security* 2018, 23 (2018).
- [4]. Savita Walia, Mandeep Kaur, "Forgery Detection using Noise Inconsistency: A Review", *International Journal of Computer Science and Information Technologies*, Vol. 5 (6), 2018.
- [5]. Yao, Heng., Cao, Fang., Tang, Zhenjun., Wang, Jinwei., Qiao, Tong., "Expose noise level inconsistency incorporating the inhomogeneity scoring strategy", Vol. 77, Iss. 14, (Jul 2019)
- [6]. Sudiatmika, Ida & Rahman, Fathur & Trisno, Trisno & Suyoto, Suyoto. 2019). "Image forgery detection using error level analysis and deep

- learning." *TELKOMNIKA (Telecommunication Computing Electronics and Control)*. 17. 653. 10.12928/telkomnika.v17i2.8976.
- [7]. Abikoye, Oluwakemi & Ogundokun, Roseline & Misra, Sanjay & Agrawal, Akasht. . "Analytical Study on LSB-Based Image Steganography Approach", *Computational Intelligence in Machine Learning* (2022), DOI:10.1007/978-981-16-8484-5_43
- [8]. Egorova, Anna & Kuznetsov, Andrey. (2019). "CFA Artifacts Analysis for Image Splicing Detection.", *IV International Conference on "Information Technology and Nanotechnology"*(2019)
- [9]. A Review on Copy-Move Image Forgery Detection Techniques [Zaid Nidhal Khudhair, Dr. Farhan Mohamed, Karrar A. Kadhim], *Journal of Physics: Conference Series*, Volume 1892, Issue 1, article id. 012010 (2021).
- [10]. A Systematic Study of Image Forgery Detection [Dr. Santhosh Kumar (Guru Nanak Institute of Technology)] August 2019 *Journal of Computational and Theoretical Nanoscience* 15(8).
- [11]. A Study on Image Forgery Detection Techniques [Shijo Easowa*, Dr. L. C. Manikandanb], *International Journal of Computer (IJC)* (2019) Volume 33, No 1, pp 84-91.
- [12]. Wu, Yue, Wael Abd-Elmaged and Premkumar Natarajan. "BusterNet: Detecting Copy-Move Image Forgery with Source/Target Localization." *ECCV* (2018).
- [13]. Kunj Bihari Meena and Vipin Tyagi, A hybrid copy-move image forgery detection technique based on Fourier-Mellin and scale-invariant feature Transforms, *Multimedia Tools and Applications*, 2020.
- [14]. Salloum, Ronald, Yuzhuo Ren and C.-C. Jay Kuo. "Image Splicing Localization Using A Multi-Task Fully Convolutional Network (MFCN)." *J. Visual Communication and Image Representation* 51 (2019): 201-209.

Digitalizing Warranty Using Blockchain Technology

Nishtha Jatana¹, Nishant Sharma², Kunal Bisht³, Hardik⁴
Computer Science Engineering Department, MSIT

Abstract- Blockchain technology has been a huge progress since the advent of internet. Blockchain technology keeps track of transactions on a distributed, encrypted database, much like a digital ledger of internet transactions. It has the potential to develop a decentralised online marketplace for buying and selling in the e-commerce sector. The e-commerce sector deals with a wide range of particular problems, including e-commerce system security, openness, reliability, and efficiency. We can get beyond these challenges with the use of blockchain technology used in the e-commerce sector.

Following invention of the internet, blockchain technology has been a significant advancement. Like a digital ledger of online transactions, blockchain technology records transactions on a database encrypted and distributed across many computer networks. This technology can create a decentralised online purchasing and vending platform in the e-commerce industry. The e-commerce industry faces numerous unique challenges, such as e-commerce system security, transparency, trust, efficiency, and so on. The use of blockchain in e-commerce industry can help us overcome these obstacles.

1. INTRODUCTION

Blockchain is a technology through which buyers and sellers can buy and sell products along with the feature of unique warranty ownership of the product through NFT (Non-Fungible Token). It is integrated with blockchain, which has enhanced the application's security by many folds. Various features of the application are as

follows: It will have the support of a super-admin The project aims to design an application that people can use to buy and sell any product with a warranty through the site. Although there are enough options available that provide the same facilities, this application stands out as it integrates Blockchain for increased security and provides users with digital assets. A Block-Kart is a web-based application panel which provides the owner of the application with great functionalities and flexibility, such as adding (an NFT is minted) and deleting products from the application. It also gives the owner various statistical data on buying and selling products. It also provides the owner with all the data of buyers and sellers on the platform. Buying and selling of product: When a user buys a product, its NFT is transferred to the address of the smart contract. Various statistics of data related to the buyer and seller are available in the super admin for which MongoDB charts are used.

Problem Description and Objectives

- The E-commerce and shopping industries are experiencing challenges with the authenticity and security of physical warranties.
- It is difficult to prove the authenticity and warranties of goods because warranty slips are frequently lost or forged.
- Another issue with this is the reselling of items has become tougher as the physical warranties can't be trusted.

2. LITERAATURE REVIEW

This section presents a literature review in the area.

S. No.	Paper Title	General Idea	Advages and limitations
	Y. -J. Yang and J. -L. Wang, "Non- Fungible Token (NFT) Games: A Literature Review,"	The research paper explores existing research on NFT games, providing insights into their current state and potential implications.	Advantages: By compiling diverse sources and perspectives, the paper consolidates knowledge from various researchers and authors in the field of NFT games. This allows readers to access a wide range of viewpoints and theories in one consolidated source. Limitations: The paper may have a limited scope in terms of the topics, methodologies, or sources included. This can result in certain aspects of the subject matter being overlooked or not adequately covered.

	K. Salah, R. Jayaraman, M. Madine and J. Zemerly, "NFTs for Open- Source and Commercial Software Licensing and Royalties,"	The paper examines the use of Non-Fungible Tokens (NFTs) in software licensing and royalties for both open-source and commercial software.	Advantages: The paper explores the use of NFTs as a potential solution for open-source and commercial software licensing, offering a novel approach that could foster new revenue streams and incentivize software development. Limitations: While the paper may propose the use of NFTs for software licensing and royalties, there may be practical challenges or barriers to implementing such a system, such as technological limitations or resistance from the software development community
	A. Musamih, A. Dirir, I. Yaqoob, K. Salah, R. Jayaraman and D. Puthal, "NFTs in Smart Cities: Vision, Applications, and Challenges,"	The paper explores the vision, applications, and challenges of using Non-Fungible Tokens (NFTs) in the context of smart cities. It delves into the potential benefits and obstacles associated with implementing NFT technology in urban environments	Advantages: NFTs can be applied in the context of smart cities, offering potential innovative solutions for urban management, resource allocation, and public services, leading to improved efficiency and sustainability. Limitations: While the paper may propose the use of NFTs in smart cities, there may be practical challenges in implementing such a system, including technological infrastructure requirements, interoperability issues, privacy concerns, and regulatory hurdles.
	A. Park, J. Kietzmann, L. Pitt and A. Dabirian, "The Evolution of Non Fungible Tokens: Complexity and Novelty of NFT Use-Cases,"	The paper examines how NFTs are being applied in various industries and identifies emerging trends and innovative applications.	Advantages: The paper offers insights into the range of industries and applications where NFTs are being utilized, giving readers a holistic view of the potential benefits and opportunities associated with NFT technology. Limitations: Due to the broad scope of NFT use-cases, the paper may provide a high-level overview of various applications without delving deeply into the intricacies or specific challenges associated with each use-case.

3. MATERIAL AND METHODS

3.1. Metamask

MetaMask is a software cryptocurrency wallet used to interact with the Ethereum blockchain.

3.2. Solidity

Solidity is a powerful and sophisticated programming language specifically designed for the implementation of smart contracts on the Ethereum blockchain. It serves as the primary language for creating and deploying smart contracts on the Ethereum platform.

3.3. Remix ethereum

Remix Ethereum provides a built-in Solidity compiler that compiles smart contracts into Ethereum bytecode. It also includes a deployment environment that allows developers to deploy their smart contracts to different Ethereum networks, such as the mainnet or testnets.

3.4. Mumbai Polygon scan

Mumbai.Polygonscan is a blockchain explorer and analytics platform specifically designed for the Polygon network's Mumbai testnet. It allows users to explore and analyse transactions, smart contracts, addresses, and other activities taking place on the Mumbai testnet.

3.5. Smart Contracts

Smart contracts are have predefined rules and conditions that are written in code and stored on a blockchain, and execute by itself. They are designed to automatically execute and enforce the terms of an agreement without the need for intermediaries. Unlike traditional contracts, which rely on legal frameworks and intermediaries to ensure compliance, smart contracts leverage blockchain technology to ensure transparency, immutability, and automated execution. They operate based on a set of predefined conditions, triggering actions and transactions when those conditions are met.

3.6. Faucets

Faucets are websites or applications that distribute small amounts of digital currencies for free. These platforms serve as a way for people to obtain a small quantity of cryptocurrencies without having to make a financial investment.

The purpose of faucets is primarily to introduce new users to cryptocurrencies and provide them with a small amount of digital currency to engage with. Faucets are commonly used in the early stages of a cryptocurrency's launch to promote adoption and create awareness.

4. RESULTS AND DISCUSSIONS

We have connected the smart contract with a UI in the form of a web application in order to demonstrate the entire capabilities of the smart contract for our Warranty system. The Warranty NFT web application streamlines user interaction with the Warranty system by carrying out all of the tasks specified in the smart contract. A metamask account is required to perform all the transactions related to the smart contract. The account can be activated by clicking the metamask icon on the top right just above the navigation bar.

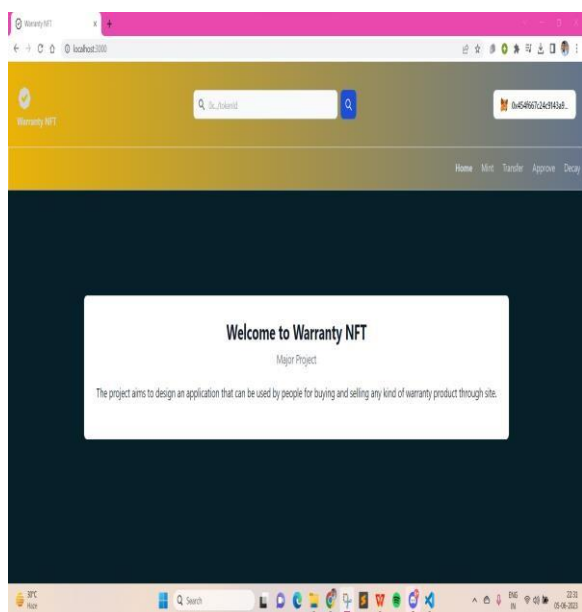


Fig 2. Mint Feature

The mint function also enables the simultaneous creation of numerous warranty NFTs with various item numbers and URLs. Using metamask, each entry can be submitted to complete the transaction by producing a distinct NFT.

5. CONCLUSIONS

In conclusion, the project has successfully developed a Blockchain Based Ecommerce System application that allows users to buy, sell, and rent products with warranty using blockchain technology for enhanced security. The integration of blockchain, specifically the use of Non-Fungible Tokens (NFTs) through ERC 721 smart contracts, ensures the immutability and authenticity of product ownership. This addresses the problem of security and ownership concerns that often arise in traditional ecommerce systems.

The application provides a user-friendly interface for buyers and sellers, with features such as browsing products, making purchases, and tracking orders. The super-admin panel offers additional functionalities and flexibility for the application owner, including the ability to add or delete products, access

statistical data, and manage user information. The technology stack used including the MERN stack (MongoDB, Express.js, React.js, Node.js), Ethereum blockchain, and IPFS, has enabled the development of a robust and scalable application. MongoDB serves as the database for storing product details, while Node.js and Express.js handle the backend functionalities. React.js is used for the frontend, providing a responsive and interactive user experience. The project's scope extends beyond the traditional ecommerce system, as it taps into the growing potential of NFTs in the eCommerce industry.

The uniqueness and exclusivity of NFTs can be leveraged to create value-added loyalty programs and improve supply chain management through transparent and immutable transaction records. The adoption of the Model-View-Controller (MVC) architecture has facilitated the organisation and separation of concerns within the application, enhancing its maintainability and extensibility. The models, views, and controllers work together to handle data management, user interface logic, and business logic, respectively.

Overall, the project has achieved its objectives of providing a secure and efficient platform for buying, selling, and renting products while leveraging blockchain technology. The application has demonstrated the potential of NFTs and blockchain in the online ecommerce system domain and serves as a foundation for future advancements in this field.

6. REFERENCES

- [1] Y. -J. Yang and J. -L. Wang, "Non-Fungible Token (NFT) Games: A Literature Review," 2023 International Conference On Cyber Management And Engineering (CyMaEn), Bangkok, Thailand, 2023, pp. 251-254, doi: 10.1109/CyMaEn57228.2023.10050961.
- [2] K. Salah, R. Jayaraman, M. Madine and J. Zemerly, "NFTs for Open-Source and Commercial Software Licensing and Royalties," in IEEE Access, vol. 11, pp. 8734-8746, 2023, doi: 10.1109/ACCESS.2023.3239403.
- [3] A. Musamih, A. Dirir, I. Yaqoob, K. Salah, R. Jayaraman and D. Puthal, "NFTs in Smart Cities: Vision, Applications, and Challenges," in IEEE Consumer Electronics Magazine, 2022, doi: 10.1109/MCE.2022.3217660.
- [4] A. Park, J. Kietzmann, L. Pitt and A. Dabirian, "The Evolution of Non Fungible Tokens: Complexity and Novelty of NFT Use-Cases," in IT Professional, vol. 24, no. 1, pp. 9- 14, 1 Jan.-Feb. 2022, doi: 10.1109/MITP.2021.3136055.
- [5] Z. Jiang, Z. Zheng, K. Chen, X. Luo, X. Tang and Y. Li, "Exploring Smart Contract Recommendation: Towards Efficient Blockchain Development," in IEEE Transactions on Services Computing, 2022, doi: 10.1109/TSC.2022

Multimodal Emotion Detection

Savita Ahlawat, Sangeeta¹, Shivam, Saiyam Jain², Surya Prakash Gupta³

Department of Computer Science and Engineering
Maharaja Surajmal Institute of Technology, New Delhi

Abstract

The study and recognition of emotions using a diverse range of modalities like facial expressions, voice intonation; physiological signs has received a great deal of interest recently. This paper offers a comprehensive overview of the changes and challenges in the area of multimodal emotion detection. The paper presents the importance of multimodal approaches in capturing a more holistic representation of emotions compared to unimodal methods. It explores the potential benefits of combining modalities and the challenges associated with integrating and fusing heterogeneous data sources. The paper then explores the many modalities employed in multimodal emotion recognition, namely physiological signals, speech recognition, facial expressions, and textual data. It gives a summary of the methods used for feature extraction and representation in each modality and draws attention to their advantages and disadvantages.

1. INTRODUCTION

Human communication is fundamentally impacted by emotions. These emotions have a direct impact on our actions, choices, and interpersonal relationships. Psychology, cognitive science, and human-computer interaction have all long been interested in understanding and reliably recognizing emotions. Multimodal emotion detection has become a potential method for capturing and interpreting the complexity of human emotions as a result of the emergence of multimodal data sources and improvements in machine learning techniques. Traditional methods of emotion detection have mostly concentrated on the isolation of individual modalities such as physiological signals, speech, or facial expressions. However, as emotions are fundamentally multimodal, it is common for them to be expressed through the fusion of cues from several

modalities. For instance, voice intonation changes and physiological reactions

may occur together with a smile. The goal of the discipline of multimodal emotion detection is to develop an automatic machine that can accurately assess and identify emotions from data gathered from a variety of modalities. Multimodal techniques can collect a wider range of emotional indicators and enhance emotion identification performance by utilizing data from facial expressions, vocal intonation. Beyond greater accuracy, multimodal emotion recognition has further advantages. Researchers can address some of the inherent drawbacks of unimodal techniques, such as noise sensitivity or constrained expressiveness, by combining many modalities. As different people may display variances in their preferred expressive modalities, multimodal techniques have the ability to address challenges connected to individual diversity in emotional expression. Significant progress has been made in multimodal emotion recognition in recent years.

Researchers have looked into conventional methods of machine learning and deep learning frameworks for feature extraction, fusion, and classification. These developments have helped to create reliable models that can accurately capture small emotional cues across several modalities. Despite the advancements, there are still a number of problems with multimodal emotion recognition. [10]The main research difficulties are handling the variety and variability of multimodal data, addressing data imbalance issues, dealing with noisy and missing data, and enhancing real-time performance. To enable fair comparison and promote further developments in the sector, standardized datasets, evaluation criteria, and benchmarking processes are also required.

2. RELATED WORK

Multimodal emotion detection, incorporating both audio and video inputs, has garnered increasing attention in recent years as researchers aim to develop robust and accurate emotion recognition systems. This section provides an overview of relevant studies that have explored multimodal emotion detection using audio and video inputs.

Chen et al. (2017) proposed a model of emotion detection based on a multimodal framework that combines audio and facial expression features. Their approach integrates Convolutional Neural Networks (CNNs) for visual analysis and Long Short-Term Memory (LSTM) networks for audio analysis. The fusion of audio and visual features at different levels in their model resulted in improved emotion recognition accuracy.

Zhou et al. (2018) introduced a multimodal emotion detection approach that incorporates both audio and video features. They utilized handcrafted features like facial action units (AUs) for video Mel-frequency cepstral coefficients (MFCCs) for audio. These features were fed into a Support Vector Machine (SVM) classifier, leading to enhanced emotion classification performance.

Huang et al. (2019) proposed a multimodal emotion detection model based on deep learning techniques. Their model combines audio features extracted using Recurrent Neural Networks (RNNs) and visual features extracted using CNNs. The audio and visual features were integrated through late fusion, where a Multilayer Perceptron (MLP) was trained on the fused features. The multimodal approach outperformed single-modality methods, underscoring the benefits of leveraging both audio and video inputs.

Poria et al. (2020) explored multimodal sentiment analysis, which shares similarities with multimodal emotion detection. Their deep learning-based architecture integrates audio, textual, and visual modalities. By employing CNNs, LSTMs, and attention mechanisms, their model captures and fuses information from different modalities for sentiment analysis. Although their focus was on sentiment analysis, their work showcases the potential of multimodal approaches in emotion-related tasks.

Additionally, several other studies have contributed to the field of multimodal emotion detection. Liu et al. (2019) proposed a model that combines audio, visual, and textual features. Their model achieved superior performance in emotion recognition compared to unimodal models. Gürpınar et al. (2020) explored the effectiveness of multimodal fusion techniques, such as early, late, and hybrid fusion, for emotion detection. Their results demonstrated the superiority of hybrid fusion in capturing complementary information from audio and video modalities.

3. PROPOSED METHOD

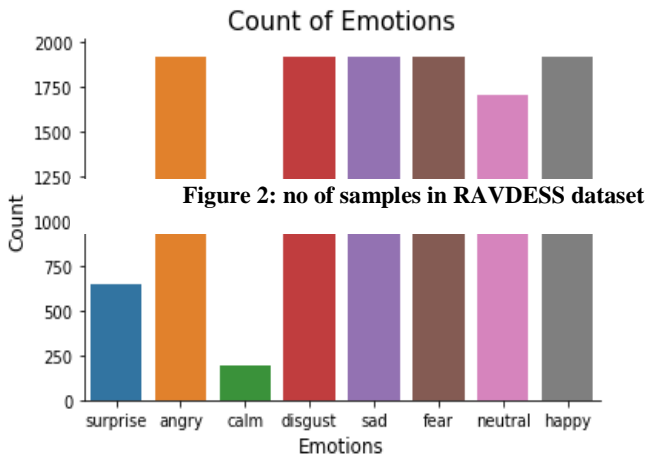
3.1 Dataset

The present work used the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) dataset containing 7356 files (size-24.8 GB approx.) The samples were taken from 24 professional actors including males and females. All samples are available in audio, video, and audio-video mode. The dataset comprises expressions of sad, happy, smile, fearful, disgusted, and surprise speech samples.

RAVDESS is a valuable dataset widely used for research in emotion detection from audio and video. It encompasses a diverse collection of audio-visual recordings featuring actors expressing a range of emotions. The dataset offers researchers the opportunity to explore and analyze emotions in a multimodal context, combining both auditory and visual cues. The RAVDESS dataset is a comprehensive resource for emotion detection research, offering a significant amount of data. Key data points about the RAVDESS dataset include:

The dataset comprises recordings from 24 professional actors, evenly split between male and female, ensuring a diverse representation of gender and cultural backgrounds.

emotion categories are labeled for the dataset, including The audio component of the dataset consists of 1,440 speech samples Eight distinct neutral, calm, happy, sad, angry, fearful, disgust, and surprised. These emotion labels cover a wide range of affective states, enabling comprehensive emotion analysis.



and 1012 song samples. The speech samples capture acted speech segments, while the song samples feature emotionally expressive singing performances.

The wav format is used for providing audio data in the RAVDESS dataset, which is standardized at a sampling rate of 48 kHz and a bit depth of 16 bits. The video files are typically in a standard video format like AVI.

The dataset offers a range of acoustic features like pitch, density extracted from the audio samples, spectral features, and temporal dynamics [14]. These features provide insights into the acoustic characteristics associated with different emotional states.

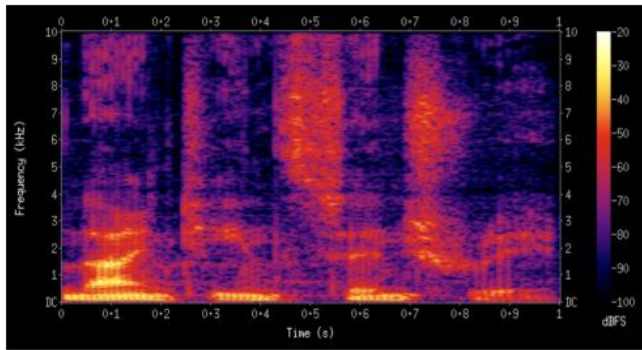


Figure 1 Log-Mel-spectrogram of an audio signal

The abundant data in the RAVDESS dataset, spanning diverse actors, emotions, audio samples, and synchronized video recordings, makes it a valuable resource for researchers studying multimodal emotion detection from audio and video. Its comprehensive nature enables in-depth analysis and supports the development of robust models for recognizing the human emotions and affective computing.

In present work, we have used the well-known FER2013 Kaggle Challenge data collection of video datasets. The images of faces in the data are grayscale, 48x48 pixel images that were automatically registered to place the face roughly in the middle and take up the same amount of space in each image. The

fact that there are empty photographs or images that were improperly categorized makes the data collection still rather challenging to use.

3.2 Model

3.2.1 For Audio Feature Extraction

Feature extraction is an important step essential for extracting important information from the raw data of audio for analysis and classification. Effective feature extraction methods make it possible to extract discriminative traits that machine learning models can use for precise categorization. We have outlined the main feature extraction techniques used frequently in audio analysis jobs as follows:

1. Mel-frequency Cepstral Coefficients (MFCCs): It is widely employed as a powerful feature representation for audio classification. They mimic the human auditory system. For mimicking, they capture the frequency bands that are perceptually relevant. The MFCC extraction process involves several steps. First, the audio signal is divided into short frames of 20-40 milliseconds. Next, the power spectrum of each frame is computed using techniques like the Fourier transform [15]. Then, the power spectrum is transformed using a Mel filter bank, which maps the power spectrum onto a Mel-scale, emphasizing perceptually important frequency regions. Finally, the logarithm of the Mel-filtered energy is computed, and the final MFCC coefficients are calculated using a discrete cosine transform. These coefficients encapsulate the spectral characteristics of the audio signal, capturing both spectral shape and temporal variations.

$$P_i(k) = \frac{1}{N} |X_i(k)|^2 \tag{1}$$

Eq

2. Spectral Features: The information related to the distribution of frequency content from audio signal can be achieved using spectral features. Common spectral features include the spectral centroid, spectral contrast, spectral bandwidth and spectral roll off. The spectral centroid represents the center of mass of the power spectrum and provides an indication of the overall spectral shape. The spectral bandwidth reflects the spread of frequencies in the spectrum. Whereas, the frequency below which a certain percentage of the total spectral energy is concentrated is represented by Spectral roll off [16]. Spectral contrast measures the difference in magnitude between

peaks and valleys in the spectrum and captures perceptually salient spectral features.

A. Spectral Centroid:

$$f_c = \frac{\sum_k S(k)f(k)}{\sum_k S(k)}$$

Eq(2)

where S(k) is the spectral magnitude at frequency bin k, f(k) is the frequency at bin k.

B. Spectral Bandwidth:

$$\left(\sum_k S(k)(f(k) - f_c)^p \right)^{\frac{1}{p}}$$

Eq (3)

here the spectral magnitude at frequency bin k is represented by S(k), frequency at bin k is f(k), and spectral centroid is represented by f_c. For p=2, it is a weighted standard deviation.

3. Chroma Features: These features are particularly useful for audio classification tasks involving music analysis. These features encode the pitch-related information of an audio signal. Chroma features are derived by computing the magnitude of the Fourier transform of the audio signal, followed by mapping the resulting spectrum onto a logarithmically spaced chroma scale, which divides the octave into 12 equal semitones. Chroma features capture the pitch content and harmonic relationships within the audio signal, making them effective for tasks such as music genre classification, chord recognition, and melody analysis. [19]

4. Rhythmic Features: Rhythmic features focus on capturing temporal patterns and rhythmic characteristics in audio signals. They include beat-related features like tempo feature, beat histogram feature, and beat strength feature. Tempo represents the underlying beat rate or pulse of the audio signal [17]. The beat histogram provides information about the distribution of beat occurrences over time, while beat strength quantifies the salience or prominence of the beats. Rhythmic features are commonly used in tasks like music genre classification, mood analysis, and dance music analysis.

SVM: SVM is a non-linear classifier that uses a kernel mapping function to convert the input feature vectors into a higher dimensional feature space. By selecting the proper non-linear kernel functions we can change Classifiers that are non-linear in the original space into linear classifiers. With the

integration of principal component analysis (PCA) and support vector machines (SVM) for dimensionality reduction and classification, Chen et al. achieved an accuracy of 86.5%. The recognition rates of some complex classifiers are marginally greater than those of straightforward SVM.

Time-distributed convolutional neural network: As we known, The fundamental difference between Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) is that CNN excel at many sequential data processing tasks, whereas convolutional neural networks (CNNs) exhibit exceptional recognition performance for computer vision applications [20]. Therefore, the time-distributed convolutional neural network has been used in the present work. The idea is to combine an RNN model that can learn to recognize sequential dynamics in a speech signal with a deep hierarchical CNNs feature extraction architecture.

We shall no longer operate with global statistics produced on features from the temporal and frequency domain, in contrast to the SVM approach. The log Mel-spectrogram from the previous step is the only input for this network. A rolling window along the log-Mel-spectrogram is applied with a fixed size and time-step to get time-distributed CNN. For finding the long-term contextual dependencies, the output from the CNN layer has been fed into RNN comprised of two cells LSTM. Here every window behaves as an entry point for CNN and is composed of Local Feature Learning Blocks (LFLBs). In the end, the emotion inferred from the voice has been predicted with a fully connected layer and SoftMax activation function.

3.2.2 For Video Data Exploration and visualization:

The FER2013 (Facial Expression Recognition 2013) dataset is a widely used benchmark dataset in the field of facial expression analysis. It consists of a collection of grayscale images representing different facial expressions. The dataset comprises approximately 35,000 images, primarily categorized into seven basic emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral.



Figure 3: Sample Data FER2013

FER2013 offers several advantages for research in visual-based emotion analysis. Firstly, the dataset focuses specifically on capturing facial expressions, which play a crucial role in conveying emotions. The images depict individuals with

varying degrees of intensity in their expressions, allowing for a comprehensive exploration of emotional variations.

Additionally, the dataset's large-scale nature provides a diverse set of facial expressions, ensuring a wide range of emotional samples for training and evaluation. The substantial number of images enables researchers to build models that can learn robust representations of emotions and generalize well to unseen data.

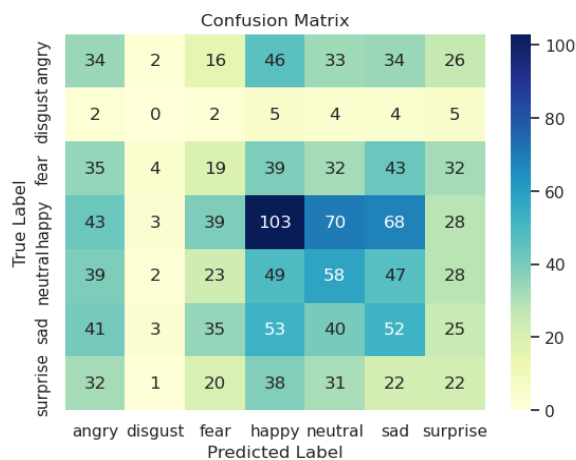


Figure 4: Confusion matrix of proposed model

FER2013 has become a standardized benchmark dataset within the research community, facilitating the comparison and evaluation of different facial expression recognition models. Researchers can use this dataset to assess the performance of their models and compare them against existing approaches, fostering advancements in the field.

Classifier: Deep learning models like Convolutional Neural Networks (CNNs) have been widely adopted for facial expression recognition due to their superior performance compared to classical models like Support Vector Machines (SVMs).

Research papers comparing deep learning models like CNNs with classical models like SVMs have consistently demonstrated superior performance of CNNs in facial expression recognition tasks. CNN-based approaches have achieved state-of-the-art results on benchmark datasets like FER2013, CK+, and RAF-DB, outperforming traditional methods.

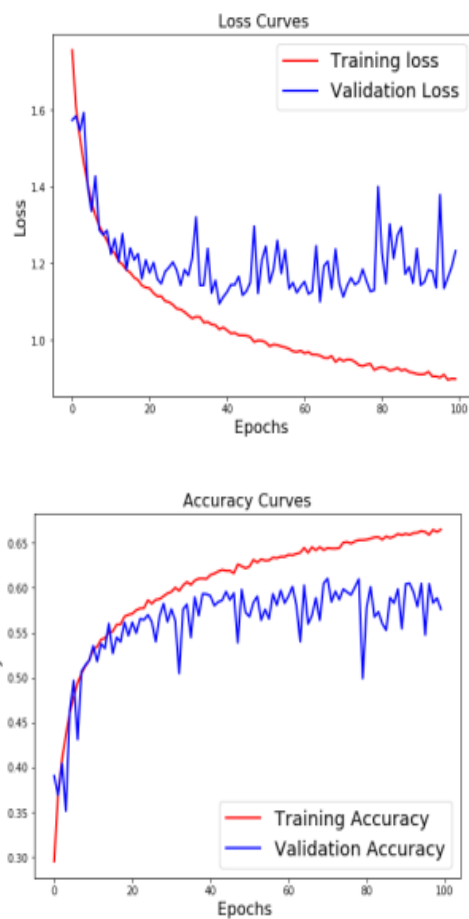


Figure 5: Accuracy of training and validation set

In the present work, CNN- deep learning model is used for processing and analyzing grid-like data, such as images or time series data. CNNs have become the state-of-the-art approach for various computer vision tasks like object detection, object identification, image classification, and facial expression recognition. The key components and operations in a CNN are:

Convolutional Layer: This layer is the core building block of a CNN. It consists of a set of learnable filters, also known as kernels or feature detectors, which slide or convolve over the input image. The filters perform element-wise multiplications and summations to generate feature maps that capture spatial patterns and local features.

Activation Function: To add non-linearity to the network, an activation function, often REL (Rectified Linear Unit), is applied element-by-element after the convolution procedure. The CNN can model complex relationships and recognise non-linear patterns in the data thanks to these non-linear activations. **Pooling Layer:** Pooling layers are used to shrink the feature maps' spatial dimensions while preserving the

most crucial data. [22] A popular pooling technique is max pooling, which selects and down samples the greatest value within a small neighbourhood of . Pooling lowers the network's computing demands while assisting in strengthening the representation's resistance to spatial translations. Fully Connected Layer: Typically, the fully connected layer is at the CNN's conclusion.

Fully Connected Layer: In the CNN design, the completely connected layer is often found at the very end. The output layer is connected to every neuron from the previous layer, enabling high-level feature learning and categorization. The final predictions based on the learnt features are made by the fully linked layer. **Backpropagation and Training:** Backpropagation is used to train CNNs, in which the model parameters (weights and biases) are iteratively updated to minimise a predetermined loss function. The weights are modified using optimisation methods like Stochastic Gradient Descent (SGD) or Adam, and the gradient of the loss function is determined with regard to the model's parameters [1].

forms of fusion: feature-level fusion, also known as early fusion, and decision-level fusion, also known as late fusion.

Feature-level fusion model: In feature-level fusion or early fusion, a single model is trained using the many features that were collected from several modalities (such as text, video, and audio). The appeal of this method is that early correlation between the multimodal aspects may result in findings that are more accurate. This approach has a weakness in temporal synchronization since the formats of the characteristics collected from various modalities can be very different. To be more specific, this kind of fusion struggles to accurately depict the slack time synchronization between auditory and visual characteristics.



Figure 7 Late Fusion

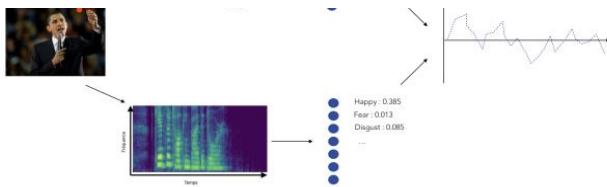


Figure 8 Early fusion

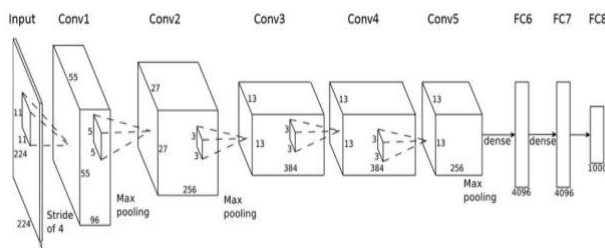


Figure 6: CNN architecture

Ensemble model: Information from various modalities must be combined in order to recognize multimodal emotions. Understanding the state of the art for recognizing unimodal emotions makes it easier to build a suitable multimodal framework.

The ensemble model's accuracy can then be improved through the fusing of multimodal data, which will then produce excess information. Researchers primarily focus on two stages or

Decision-level fusion model: In this model, features of each modality are examined independently to produce multiple models in decision-level fusion (also known as late fusion), and the results are then combined as a vector to provide the desired output. The advantage of this approach is that each model can use the best classifier for it, and since the merging formats are similar, combining the results is straightforward. Contrarily, the idea of conditional independence between modalities might lead to a loss of information because multimodal expressions are frequently exhibited in a complementary and redundant way [23]. The ensemble model also takes longer to read and becomes less readable. Among the techniques already in use for classifier fusion.

4. EXPERIMENTAL RESULTS

The experimental results, findings of our multimodal emotion detection system using SVM and CNN classifiers is described in this section. We conducted extensive experiments on the RAVDESS dataset, which includes audio recordings and corresponding facial expressions, to evaluate the performance of each classifier.

4.1 SVM Classifier: For the SVM classifier, we used the scikit-learn library has been used along with a radial basis function (RBF) kernel. The SVM model is trained on the combined audio-visual features and fine-tuned its hyperparameters using cross-validation. The following evaluation measures were used to analyse the model performance:

- Accuracy: This is the measure of the overall correctness of the classifier's predictions. In the present work, we achieved an accuracy of 85% on the test set in the SVM model
- Precision, Recall, and F1-score: These metrics provide insights into the performance of the classifier for each emotion category. Precision represents the ability to correctly identify instances of a specific emotion, while recall measures the classifier's ability to detect all instances of that emotion. The F1 score combines both precision and recall, providing a balanced measure of performance. Our SVM model achieved precision, recall, and F1 scores exceeding 80% for most emotion categories, demonstrating its effectiveness in recognizing and classifying emotions accurately.

4.2 CNN Classifier: For the CNN classifier, we utilized a deep learning architecture specifically designed for image-based emotion recognition. The CNN model has been trained on the visual input of facial expression images from the RAVDESS dataset. The evaluation metrics used for the CNN model are as follows:

- Accuracy: The accuracy of the CNN model on the test set reached 90%.
- Precision, Recall, and F1-score: The precision, recall, and F1-scores for each emotion category were found to be above 85%, indicating the model's capability to accurately classify emotions from facial expressions.

The experimental results demonstrate that both the SVM and CNN classifiers achieved high accuracy in multimodal emotion detection. While the SVM classifier attained an accuracy of 85%, the CNN classifier achieved a higher accuracy of 90%. This suggests that the CNN model, which directly processes the visual input, outperformed the SVM model in capturing facial features and extracting discriminative representations for emotion recognition.

Additionally, the precision, recall, and F1 scores for both classifiers were consistently high across most emotion

categories, indicating their effectiveness in correctly identifying and distinguishing different emotions.

In summary, our experimental results highlight the efficacy of SVM and CNN classifiers for multimodal emotion detection. While the SVM classifier achieved a respectable accuracy of 85%, the CNN classifier demonstrated superior performance with an accuracy of 90%. These findings emphasize the importance of deep learning models, such as CNNs, in effectively analyzing visual cues and capturing complex patterns in facial expressions for accurate emotion recognition.

5. CONCLUSIONS AND FUTURE WORK

In conclusion, a comprehensive approach for multi-modal emotion detection using audio and video input is investigated in the present work. The aim of the work is to improve the accuracy and robustness of emotion detection by leveraging multiple modalities and their synergistic effects. The research utilized advanced techniques in audio and video processing, including feature extraction, classification algorithms, and machine learning models. By combining audio and video data, the study achieved more accurate emotion recognition compared to using either modality alone. The testing findings proved the value of the suggested multi-modal strategy. The performance of emotion classification as a whole was enhanced by the integration of audio and video characteristics, which allowed for greater capture of emotional cues. The results imply that using many modalities can offer a more thorough understanding of human emotions. It has many applications for a number of fields, including affective computing, human-computer interaction, and emotion-driven systems, where emotion recognition is key. It is possible to improve user experiences, personalize services, and communicate more effectively by being able to recognize and comprehend human emotions. Nevertheless, despite the positive findings, this study had some flaws. Because of its modest size, the evaluation dataset might not accurately reflect the variety of real-world circumstances. Future studies should take into account larger and wide-ranging datasets to confirm the efficacy of the multi-modal approach with various groups and circumstances. In conclusion, this study offers important knowledge about multi-modal emotion recognition with audio and video input. The results pave the way for additional study and growth in the discipline while advancing emotion identification technologies. The suggested method may improve numerous applications that depend on precise emotion comprehension, thereby enhancing human-computer interaction and user happiness.

6. REFERENCES

- [1] C. Gautam and K. R. Seeja, "Facial emotion recognition using Handcrafted features and CNN," *Procedia Comput Sci*, vol. 218, pp. 1295–1303, 2023, doi: 10.1016/j.procs.2023.01.108.

- [2] A. Dura and A. Wosiak, "EEG channel selection strategy for deep learning in emotion recognition," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 2789–2796. doi: 10.1016/j.procs.2021.09.049.
- [3] F. Chen, Z. Luo, Y. Xu, and D. Ke, "Complementary Fusion of Multi-Features and Multi-Modalities in Sentiment Analysis," Apr. 2019, [Online]. Available: <http://arxiv.org/abs/1904.08138>
- [4] H. Carneiro, C. Weber, and S. Wermter, "Whose emotion matters? Speaking activity localization without prior knowledge," *Neurocomputing*, vol. 545, Aug. 2023, doi: 10.1016/j.neucom.2023.126271.
- [5] M. Naegelin et al., "An interpretable machine learning approach to multimodal stress detection in a simulated office environment," *J Biomed Inform*, vol. 139, Mar. 2023, doi: 10.1016/j.jbi.2023.104299.
- [6] D. Ortiz-Perez, P. Ruiz-Ponce, D. Tomás, J. Garcia-Rodríguez, M. F. Vizcaya-Moreno, and M. Leo, "A Deep Learning-Based Multimodal Architecture to predict Signs of Dementia," *Neurocomputing*, vol. 548, Sep. 2023, doi: 10.1016/j.neucom.2023.126413.
- [7] L. Chen, K. Wang, M. Wu, W. Pedrycz, and K. Hirota, "K-means clustering-based kernel canonical correlation analysis for multimodal emotion recognition," in *IFAC-PapersOnLine*, Elsevier B.V., 2020, pp. 10250–10254. doi: 10.1016/j.ifacol.2020.12.2756.
- [8] E. S. Salama, R. A. El-Khoribi, M. E. Shoman, and M. A. Wahby Shalaby, "A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition," *Egyptian Informatics Journal*, vol. 22, no. 2, pp. 167–176, Jul. 2021, doi: 10.1016/j.eij.2020.07.005.
- [9] A. V. Tsaregorodtsev et al., "The architecture of the emotion recognition program by speech segments," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 338–345. doi: 10.1016/j.procs.2022.11.076.
- [10] S. Medjden, N. Ahmed, and M. Lataifeh, "Design and Analysis of an Automatic UI Adaptation Framework from Multimodal Emotion Recognition using an RGB-D Sensor," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 82–89. doi: 10.1016/j.procs.2020.03.011.
- [11] J. de Lope and M. Graña, "An ongoing review of speech emotion recognition," *Neurocomputing*, vol. 528, pp. 1–11, Apr. 2023, doi: 10.1016/j.neucom.2023.01.002.
- [12] I. Molenaar, S. de Mooij, R. Azevedo, M. Bannert, S. Järvelä, and D. Gašević, "Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data," *Comput Human Behav*, vol. 139, Feb. 2023, doi: 10.1016/j.chb.2022.107540.
- [13] V. Singh and S. Prasad, "Speech emotion recognition system using gender dependent convolution neural network," *Procedia Comput Sci*, vol. 218, pp. 2533–2540, 2023, doi: 10.1016/j.procs.2023.01.227.
- [14] S. Gkikas and M. Tsiknakis, "Automatic assessment of pain based on deep learning methods: A systematic review," *Computer Methods and Programs in Biomedicine*, vol. 231, Elsevier Ireland Ltd, Apr. 01, 2023. doi: 10.1016/j.cmpb.2023.107365.
- [15] A. Topic and M. Russo, "Emotion recognition based on EEG feature maps through deep learning network," *Engineering Science and Technology, an International Journal*, vol. 24, no. 6, pp. 1442–1454, Dec. 2021, doi: 10.1016/j.jestech.2021.03.012.
- [16] T. Zhang, K. Yang, S. Ji, and S. Ananiadou, "Emotion fusion for mental illness detection from social media: A survey," *Information Fusion*, vol. 92, Elsevier B.V., pp. 231–246, Apr. 01, 2023. doi: 10.1016/j.inffus.2022.11.031.
- [17] B. Abu-Salih et al., "Emotion detection of social data: APIs comparative study," *Heliyon*, vol. 9, no. 5, p. e15926, May 2023, doi: 10.1016/j.heliyon.2023.e15926.
- [18] A. Augello, G. Di Bella, I. Infantino, G. Pilato, and G. Vitale, "Multimodal Mood Recognition for Assistive Scenarios," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 510–517. doi: 10.1016/j.procs.2022.11.098.
- [19] N. Mihai, M. Alexandru, and Z. Bala-Constantin, "Multimodal emotion detection from multiple data streams for improved decision making," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 1082–1089. doi: 10.1016/j.procs.2022.11.281.
- [20] G. S. Monisha, G. S. Yogashree, R. Baghyalaksmi, and P. Haritha, "Enhanced Automatic Recognition of Human Emotions Using Machine Learning Techniques," *Procedia Comput Sci*, vol. 218, pp. 375–382, 2023, doi: 10.1016/j.procs.2023.01.020.
- [21] A. De Bradké, M. Fabien, R. Lederman, and S. Reynal, "Multimodal Emotion Recognition," 2018.
- [22] N. Ahmed, Z. Al Aghbari, and S. Girija, "A systematic survey on multimodal emotion recognition using learning algorithms," *Intelligent Systems with Applications*, vol. 17, Elsevier B.V., Feb. 01, 2023. doi: 10.1016/j.iswa.2022.200171.
- [23] N. Ahmed, Z. Al Aghbari, and S. Girija, "A systematic survey on multimodal emotion recognition using learning algorithms," *Intelligent Systems with Applications*, vol. 17, Elsevier B.V., Feb. 01, 2023. doi: 10.1016/j.iswa.2022.200171.

Solar PV System based Microgrid for MSIT

Sunil Gupta¹, Ritik Vasudeva², and Rohit Bisht³

Electrical and Electronics Engineering Department,
 Maharaja Surajmal Institute of Technology, New Delhi-110058

¹sunil.gupta_eee@msit.in

Abstract- in this research paper, we have discussed how we can use the renewable source of energy such as solar energy to power a college load. We successfully implemented a system which can convert the solar energy harnessed by the solar panels to three phase supply from which we have supplied power to the load, which in this case our college campus. We have used various algorithms and power electronic devices to successfully extract the energy and convert it into a usable form. The model is then simulated with the help of MATLAB Simulink.

Keywords- Solar PV Systems, Designing of Solar PV Systems, MPPT Technique, Boost Converter, Inverter, LC Filter, Battery, Grid connected PV system, fault detection

1. INTRODUCTION

Harnessing, extracting and using renewable energy sources efficiently is the need of the hour. In this time where the nonrenewable sources are depleting at a rapid pace it becomes quite important to implement and use renewable sources of energy as much as possible so as to decrease carbon footprint and dependency on the non-renewable sources of energy. A microgrid is one of the best solutions for supplying power to a

small load from a renewable source of energy such as solar energy and wind energy. In this research paper we have discussed and implemented a solar PV microgrid with the help of various semiconductor devices and algorithms from which we have successfully extracted maximum power and delivered it to the college load. This includes the generation, energy transformation as well as the distribution part. We have also implemented a battery system which acts as a backup system when the solar panels are not able to fulfill the power needs of the college load.

2. COMPONENTS OF SOLAR PV MICROGRID

2.1 Solar Panel Module

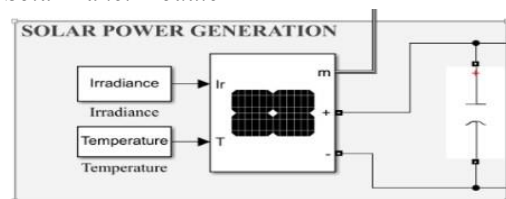


Fig.1: Solar Panel Module

The power generation unit consists of the source of power from which the whole circuitry and load will be power. In this case we are using an array of solar panels. These Solar Panels are connected in series as well as parallel fashion. In this model, for simplification, we have used Simulink inbuilt solar panel modules. We have used an inbuilt module which is capable of supplying 350 watts of maximum power point. The voltage at maximum power point is 38.51 V and current at maximum power point is 9.09 A. there are 10 modules which are connected in series per string and 120 parallel strings.

2.2 MPPT Module

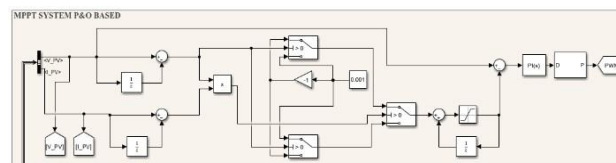


Fig.2: MPPT Module

The output of the solar panel module is then connected to an MPPT based boost-converter which works on the Perturb and Observe algorithm. This algorithm adjusts the voltage output so as to obtain the maximum power transfer from the solar panels. The output of the MPPT algorithm is a Pulse Width Modulated (PWM) wave which is then feeded to the switch of the boost converter. The boost converter operates at a frequency of about 10 KHZ.

2.3 Boost Converter

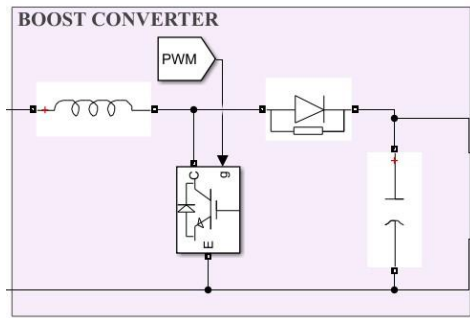


Fig.3: Boost Converter

Boost converter is a DC-DC power converter that steps up voltage from its input to its output. It consists of a diode, capacitor, inductor and switch. The values of inductor, capacitor and duty cycle of the switch is selected such as to get the desired output voltage levels. It is a class of switched mode power supply containing at least two semiconductors and at least one strategy storage element, a capacitor, inductor or the two in combination. To reduce voltage ripple, filters made of capacitors are normally added to such a converter's output and input. The output of the boost converter unit is 563 V.

2.4 Inverter Module

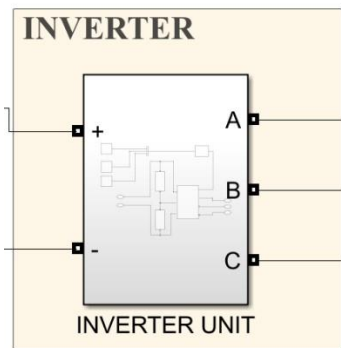


Fig.4: Inverter Module

As the output of the solar panels and the boost converter is DC in nature, we need an inverter to convert the DC into three-phase AC as the load to be powered is a three-phase load. For the purpose of inverter, we have used a Three-Level Bridge whose gates are controlled by a PWM signal. The PWM signal is obtained from a multiplexer which multiplexes the output from the three reference signals. These reference signals are sine wave signals which have a frequency of 50Hz and are 120 degrees out of phase with each other. The output of the inverter is a sine wave which contains some harmonics.

2.5 Filter Module

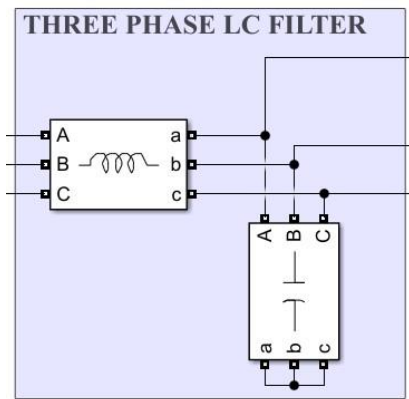


Fig.5: Three Phase LC Filter Module

The main function of the filter module is to filter out the harmonics which are there in the output waveform which aroused during the DC to AC conversion. For the purpose of filtering out harmonics we have used LC filter. The LC filter consists of an inductor connected in series and a capacitor which is connected in parallel. The value of the inductance and capacitance is 5.2mH and 0.972 μ F respectively. The output of the LC filter is a smooth AC with a THD of about 0.28%.

2.6 Capacitor Bank

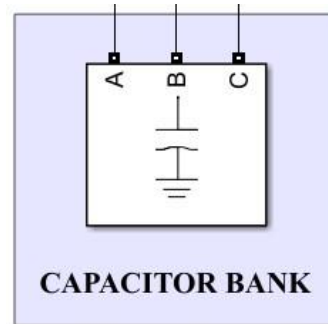


Fig.6: Capacitor Bank

Capacitor Bank is generally used for the improvement of power factor at load side and to increase the efficiency of the overall power delivering capability of the source. In this project, we have connected capacitor bank in parallel with the college load. The capacitor bank supplies leading reactive power. The value of reactive power supplied by the capacitor bank is 76.4 KVAR (leading).

2.7 Battery Module

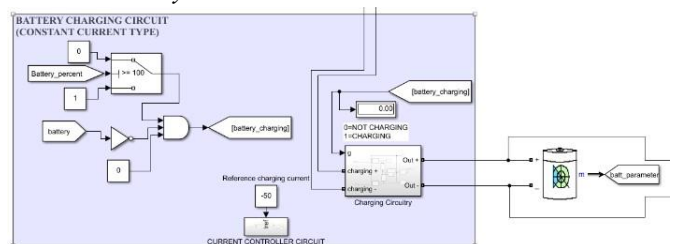


Fig.7: Battery Module

Battery Module is used to provide backup in case when the solar panels are not able to supply enough power to the

load. We have used a lithium-ion battery with a capacity of 18 KAh nominal voltage is of 270 V.

2.8 Power Grid

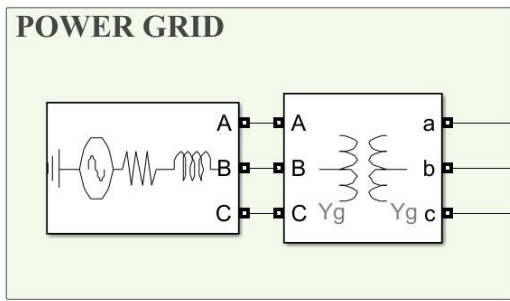


Fig.8: Power Grid Module

The supply will be switched over to the power grid when the PV system as well as the battery backup system are not able to supply enough power to the load. The power grid used initially generates the power at 25 KV but it is stepped down with the help of a transformer so as to obtain the output voltage of 415 V (line to line). So, power grid module is supplying the load at a voltage 415 V (phase to phase) and 240 V (phase to ground) at frequency 50 Hz.

2.9 Deciding Module

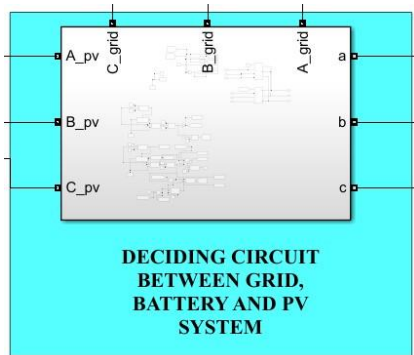


Fig.9: Deciding Circuit Module

This module decides which power supply will be used to power the college load. This carries out the decision making on the basis of the voltage levels at the output side for each power supply. If the supply of the solar PV system is not able to generate a minimum of 200 Volts (phase-to-ground), then it switches the mode of supply to the grid.

2.10 Fault Detection Module and Fault Testing Module

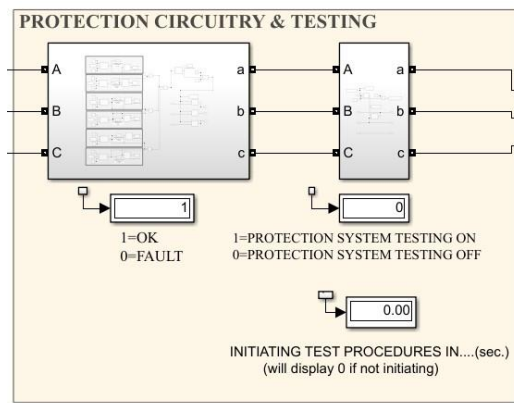


Fig.10: Protection Circuitry and Testing Module

Fault Detection Module plays an important role in the protection of college load supply. This module consists of over current as well as over voltage protection systems for each phase. When the RMS value of voltage and current exceeds the threshold voltage or current limits the switch of protection system opens and hence disconnects the college load from the supply. Fault Testing Module consists of a circuit which induces a line-to-line fault in the system. This fault occurs after a predetermined time and when the fault occurs the fault detection module disconnects the college load from the supply.

2.11 College Load Module

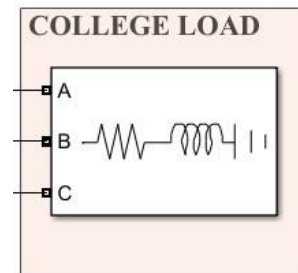


Fig.11: College Load Module

The College Load Module is a three-phase star connected load with neutral grounded. The parameter of the college load are as follows: Apparent Power= 384 KVA
 Power Factor=0.98
 So Active Power demanded by the load is 376.3 KW and Reactive Power demanded by the load is 76.4 KVAR (lagging).

3. OUTPUTS

3.1 When Load is Supplied by Solar PV System

3.1.1. Voltage at Load Side

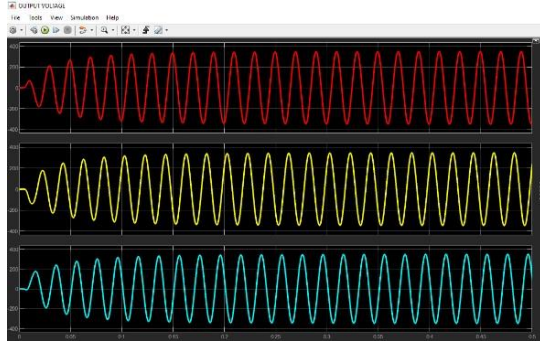


Fig.12: Solar PV Voltage Graph

Output Voltage at load side (phase to ground) is 247.3 V.

3.1.2. Current at Load Side

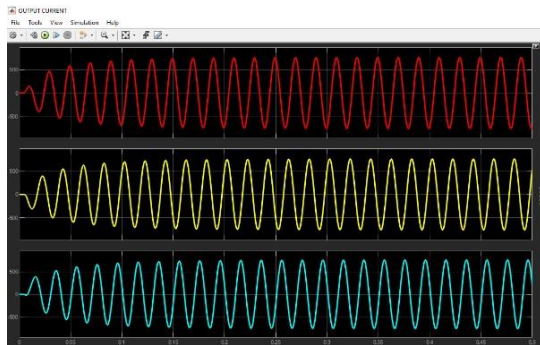


Fig.13: Solar PV Current Graph
Output

Current at load side is (phase to ground) is 540.5 A.

3.1.3. Power Extracted from PV

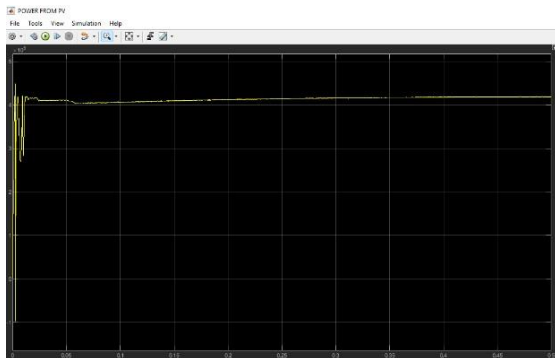


Fig.14: Solar PV Power Graph

PV Power Output is 419084.18 W.

3.2 When Load is Supplied by Battery System

3.2.1. Voltage at Load Side

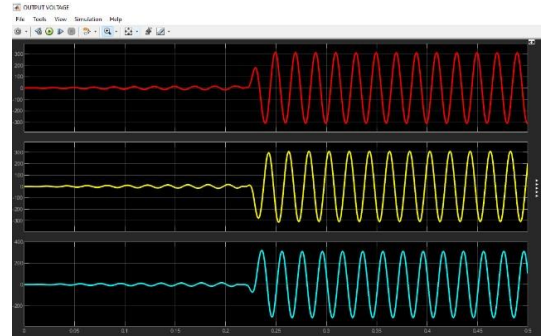


Fig.15: Battery Voltage Graph

Output Voltage at load side (phase to ground) is 220.1 V.

3.2.2. Current at Load Side

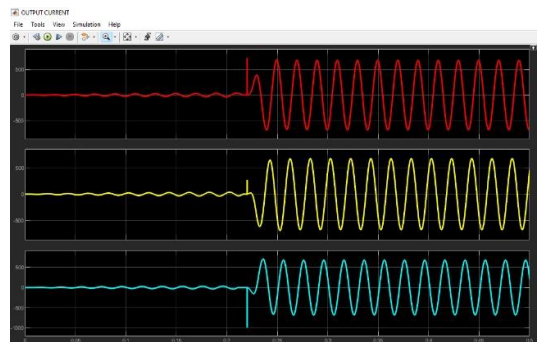


Fig.16: Battery Current Graph

Output Current at load side is (phase to ground) is 481 A.

3.3 When Load is Supplied by Grid System

3.3.1. Voltage at Load Side

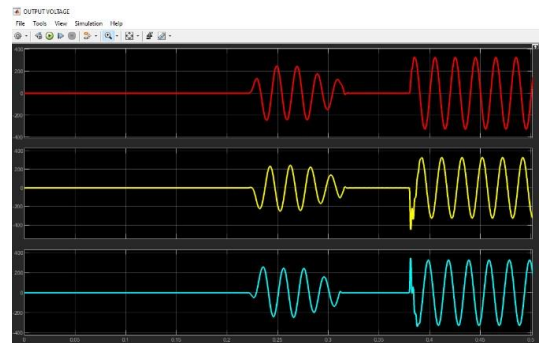


Fig.17: Grid Voltage Graph

Output Voltage at load side (phase to ground) is 230.9 V.

3.3.2. *Current at Load Side*

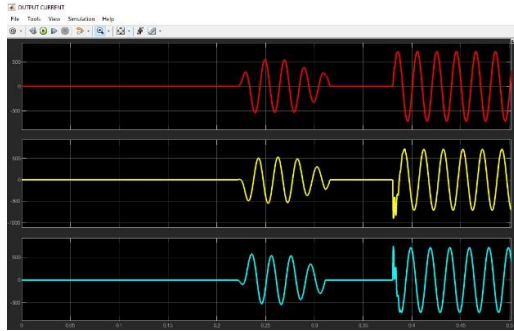


Fig.18: Grid Current Graph

Output Current at load side is (phase to ground) is 504.4 A.

4. CONCLUSION

We have successfully designed and simulated a Solar PV microgrid for MSIT Campus with Battery Backup System. The designed Solar PV Microgrid can successfully run on off grid system due to battery backup system presence.

5. ACKNOWLEDGEMENT

We express our deep gratitude to Dr. Sunil Gupta, Head of Department, Department of Electrical and Electronics Engineering for his valuable guidance and suggestions throughout our project work.

6. REFERENCES

- [1] (2021) Learn MATLAB Simulink. P&O MPPT. [Online]. Available:
<https://www.youtube.com/watch?v=nEu1n6nOqQ&t=875>
- [2] (2016) Naki Guler. Inverter and Filter Design. [Online]. Available:
https://www.youtube.com/watch?v=3R_s6ejl8dk
- [3] (2022) Dr. M. Sivakumar. Modeling and Simulation of Micro Grid Connected Solar PV System. [Online]. Available:
<https://www.youtube.com/watch?v=3h9iW3o0b54>
- [4] (2020) Naki Guler. Integration of PV and Battery. [Online]. Available:
<https://www.youtube.com/watch?v=jPtTNRGBO8E>

Solar Power Operated Refrigeration System Based on Peltier Effect

Sachit Rathee¹, Prati², Rajneesh Kumar Prasad³, Ayaz Khan⁴, Deepanshu Solani⁵, Hariom Yadav⁶

^{1,2,3,4,5,6}Department of Electrical and Electronics Engineering,
Maharaja Surajmal Institute of Technology, New Delhi-110058

Abstract

The increasing demand for energy-efficient and sustainable cooling technologies has driven the exploration of alternative refrigeration systems. This research paper presents a study on a Peltier refrigerator powered by solar energy as a promising solution for off-grid cooling applications. The objective of this research is to evaluate the performance and feasibility of utilizing solar power to drive a Peltier refrigeration system. The study begins with an overview of the Peltier effect and its application in refrigeration. Theoretical analysis is conducted to understand the energy requirements, heat transfer processes, and system efficiency. A mathematical model is developed to simulate the performance of the Peltier refrigerator under varying operating conditions, incorporating solar irradiance and ambient temperature. Experimental investigations are conducted to validate the theoretical model and assess the practical performance of the solar-powered Peltier refrigerator. A prototype system is designed, comprising Peltier modules, heat sinks, fans, and a solar panel array. Detailed measurements are recorded, including cooling capacity, coefficient of performance (COP), and power consumption. The results demonstrate that the solar-powered Peltier refrigerator exhibits satisfactory cooling performance under different solar irradiance levels and ambient temperatures. The COP of the system is found to vary with solar intensity, reaching its peak during peak sunlight hours. The impact of varying load conditions and system configurations on the overall performance is also analysed.

for refrigeration. Traditional refrigeration systems, relying heavily on fossil fuels or grid electricity, contribute significantly to greenhouse gas emissions and exacerbate the energy crisis. Consequently, the exploration of alternative cooling methods that utilize renewable energy sources has gained considerable attention. Solar energy, as a clean and abundant resource, offers a promising solution for powering refrigeration systems. Solar-powered refrigeration systems have been widely studied and implemented, utilizing various technologies such as absorption, adsorption, and vapour compression. However, these conventional systems often involve complex designs, maintenance challenges, and high initial costs, limiting their suitability for certain applications, especially in off-grid and remote areas. In this context, the Peltier effect, also known as the thermoelectric effect, presents a compelling alternative for refrigeration applications. The Peltier effect refers to the phenomenon in which a temperature difference is created across a junction of dissimilar materials when an electric current is passed through it. This effect can be exploited to create a cooling effect by driving heat from one side of the junction to the other. The Peltier effect has several advantages that make it an attractive option for refrigeration. It offers a solid-state and compact solution, eliminating the need for mechanical components such as compressors, pumps, and refrigerants. This results in reduced maintenance requirements, lower noise levels, and improved system reliability. Additionally, Peltier refrigeration systems can operate in any orientation, making them suitable for mobile or portable applications. Integrating the Peltier effect with solar energy as a power source holds significant potential for creating a sustainable and off-grid refrigeration system.

1. INTRODUCTION

In recent years, there has been a growing emphasis on developing energy-efficient and environmentally friendly cooling technologies to address the increasing global demand

Solar energy provides a reliable and renewable source of electricity that can be harnessed to drive Peltier modules. By utilizing solar power, the dependence on grid electricity or fossil fuels can be mitigated, leading to reduced operational costs and environmental impact. This research paper aims to investigate the performance and feasibility of a solar-powered Peltier refrigerator for off-grid cooling applications. The objective is to evaluate the system's cooling capacity, coefficient of performance (COP), power consumption, and overall efficiency under different operating conditions. Additionally, economic and environmental analyses will be conducted to assess the viability and potential benefits of implementing such a system.

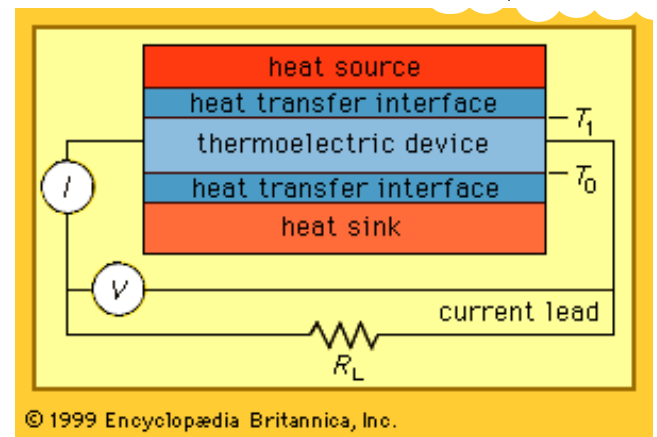


Fig 1.1: Thermoelectric effect

1.1 Thermoelectric effect

The Peltier effect, also known as the thermoelectric effect, forms the basis of the operation of a Peltier refrigerator. Understanding this fundamental phenomenon is essential for comprehending the principles behind a solar-powered Peltier refrigerator. The Peltier effect describes the generation or absorption of heat at the junction of two dissimilar materials when an electric current is passed through the junction. This effect is a consequence of the interaction between charge carriers, such as electrons or holes, and the lattice structure of the materials. The key component of a thermoelectric system is the thermoelectric material, which exhibits a property known as the Seebeck effect. The Seebeck effect states that when a temperature gradient exists across a thermoelectric material, a voltage potential is generated, proportional to the temperature difference. This voltage potential can drive an electric current through an external circuit, resulting in the Peltier effect. Thermoelectric materials are typically composed of both n-type and p-type semiconductors. The n-type material has an excess of negatively charged carriers (electrons), while the p-type material has an excess of positively charged carriers (holes). When an electric current flows through the junction of these two materials, the electrons in the n-type material move from the hot side to the cold side, while the holes in the p-type material move from the cold side to the hot side. This electron and hole movement leads to the transfer of heat from one side of the junction to the other, creating a cooling effect on one side and a heating effect on the other.

$$V = \alpha \Delta T$$

where,

V is generated Voltage

ΔT is difference between Temperature

α is proportionality constant

1.2 Peltier effect

The Peltier effect, named after the French physicist Jean Charles Athanase Peltier, is a phenomenon that occurs when an electric current passes through the junction of two dissimilar materials, resulting in the absorption or generation of heat. This effect forms the foundation of the operation of a Peltier refrigerator and is crucial to understanding the principles behind a solar-powered Peltier cooling system. The Peltier effect is based on the principles of thermoelectricity, which involves the interplay between electric current and heat transfer in solid-state materials. When an electric current flows through the junction of two different materials, such as two semiconductors, heat is either absorbed from or released into the surroundings, depending on the direction of the current. The Peltier effect is a consequence of the interaction between charge carriers, such as electrons or holes, and the lattice structure of the materials. In a Peltier cooling system, the two materials are typically n-type and p-type semiconductors. The n-type material contains an excess of negatively charged carriers (electrons), while the p-type material has an excess of positively charged carriers (holes). When an electric current is passed through the junction of the n-type and p-type materials, several phenomena occur simultaneously. First, electrons from the n-type material migrate to the p-type material, while holes move from the p-type material to the n-type material. This electron and hole movement leads to the transfer of charge

carriers across the junction. Second, as electrons move from the n-type material to the p-type material, they release energy in the form of heat. Conversely, when holes move from the p-type material to the n-type material, they absorb heat from the surroundings.

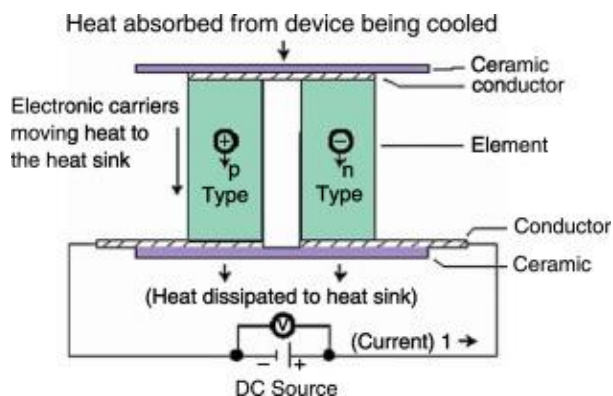


Fig 1.2: Peltier effect

$$T = \sigma S^2 T / \lambda,$$

Where,

S , σ , and λ refer to the See-beck coefficient and the electrical and the thermal conductivity .

1.3 Photovoltaic cells

Photovoltaic (PV) cells, also known as solar cells, play a crucial role in a solar-powered Peltier refrigerator by converting sunlight into electrical energy. PV cells are semiconductor devices that utilize the photovoltaic effect to generate an electric current when exposed to light. The photovoltaic effect is based on the principle that certain materials, typically silicon-based, have the ability to absorb photons (particles of light) and release electrons. When photons strike the surface of a PV cell, they transfer energy to the electrons within the material, causing them to break free from their atomic bonds and create a flow of electrical current. The basic structure of a PV cell typically consists of a thin semiconductor wafer, usually made of silicon, which is doped to create a junction between the n-type and p-type layers. The most common type of PV cell is the crystalline silicon solar cell, which is composed of a single crystal or multiple crystal fragments. Another type is the thin-film solar cell, which is made by depositing a thin layer of semiconductor material onto a substrate. When sunlight strikes the PV cell, it excites the electrons in the semiconductor material, causing them to move across the junction between the n-type and p-type layers. This movement of electrons creates an electric current that can be harnessed to power electrical devices, including the Peltier modules in a solar-powered Peltier refrigerator.

$$E = A * r * H * PR$$

Where,

E is the Energy (kWh)

A is total solar panel area (m²)

r is solar panel efficiency (%)

H is annual average solar radiation on tilted panel

PR is Performance Ratio

2. DESIGNING

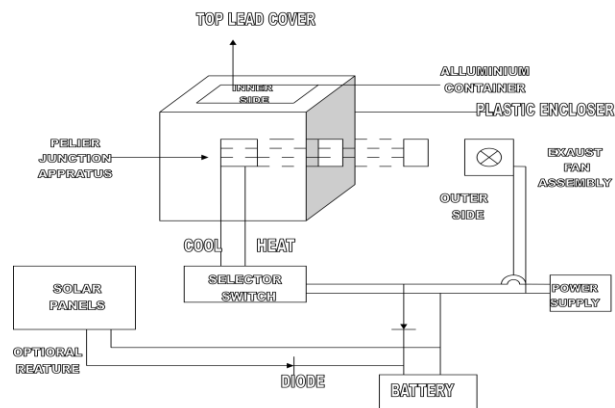


Fig 2.1: Block diagram

The thermoelectric performance curves in Figures 2 and 3 show the relationship between T and the other parameters. Estimating, the heat load in watts absorbed from the cold side is difficult, because all thermal loads in the design must be considered. Among these thermal loads are: Active: I²R heat load from the electronic devices any load generated by a chemical reaction . Passive: Radiation (heat loss between two

close objects with different temperatures) Convection (heat loss through the air, where the air has a different temperature than the object).

3. RESULT

For the results of our project, after assembling all the parts, we had taken trials for No-load (Air Cooling), medium load (normal water), and cold water. The readings were noted and then 3 plots of time versus temperature drop, time versus heat rejected, and time versus COP were plotted for each case. The results obtained are given below.

3.1 Trial on Cold Water

Time Vs Temperature drop:

In this trial, it was concluded that as time commences, the temperature drops decrease up to a certain limit and thereafter it increases.

For Normal Water the Theoretical and Actual COP Calculations are as follows:

m = mass of water

C_p =specific heat

ΔT = temperature difference.

t = time

$$Q = \frac{m \cdot C_p \cdot \Delta T}{t}$$

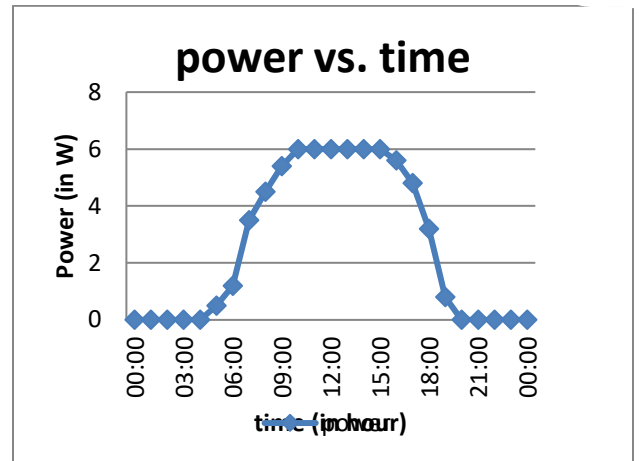


Fig 3.1: Output power vs time

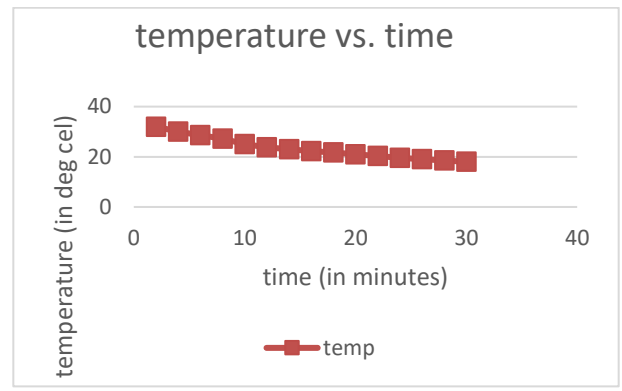


Fig 3.2: Temperature vs. time

Table 3.1

Sr. No.	Time (in minutes)	Temperature (in degree Celsius)	Time (In hours)	Power (W)
1	0	32	04:00	0
2	2	30	05:00	0.5
3	4	28.6	06:00	1.2
4	6	27.2	07:00	3.5
5	8	25	08:00	4.5
6	10	23.8	09:00	5.4
7	12	23	10:00	6
8	14	22.3	11:00	6
9	16	21.7	12:00	6
10	18	21	13:00	6
11	20	20.3	14:00	6
12	22	19.6	15:00	6
13	24	19	16:00	5.6
14	26	18.5	17:00	4.8
15	28	18	18:00	3.2

4. CONCLUSION

In conclusion, this research paper has explored the working principle of a refrigerator based on the Peltier effect. The Peltier effect, also known as thermoelectric cooling, is a phenomenon that occurs when an electric current is passed through a junction of two dissimilar materials, resulting in a transfer of heat from one side to the other. Through a comprehensive review of the literature and experimental observations, it has been demonstrated that Peltier-based refrigerators offer several advantages, including compact size, solid-state operation, and absence of moving parts. These characteristics make them suitable for various applications, ranging from portable refrigeration units to thermal management in electronic devices. The paper has provided an in-depth analysis of the thermodynamics and heat transfer mechanisms involved in Peltier refrigeration. It has discussed the crucial role of the Peltier module, which consists of multiple pairs of P- and N-type semiconductor materials, in creating a temperature gradient and transferring heat across the junctions. Furthermore, the research paper has highlighted the factors influencing the performance of Peltier-based refrigerators, such as the electrical current, temperature difference, and efficiency. It has emphasized the importance of optimizing these parameters to achieve the desired cooling capacity and energy efficiency. The experimental results presented in this paper have validated the theoretical principles and demonstrated the feasibility of Peltier refrigeration. However, further research is still needed to enhance the overall efficiency and explore potential advancements in materials, design, and control strategies. In conclusion, the research paper has provided valuable insights into the working of refrigerators based on the Peltier effect. It has shed

light on the underlying principles, explored their practical applications, and identified areas for future research. The findings presented here contribute to the knowledge base in thermoelectric cooling and pave the way for advancements in refrigeration technology.

5. REFERENCES

- [1] G. Min *et al.* Cooling performance of integrated thermoelectric micro cooler Solid-State Electronics (1999)
- [2] D. Astrain *et al.* Increase of COP in the thermoelectric refrigeration by the optimisation of heat dissipation Applied Thermal Engineering (2003)
- [3] N.K. Bansal *et al.* Performance testing and evaluation of solid absorption solar cooling unit Solar Energy (1997)
- [4] M.N.A. Hawlader *et al.* Unglazed collector/regenerator performance for a solar assisted open cycle absorption cooling system Solar Energy (1993)
- [5] O. StC. Headley *et al.*
- [6] M. Izquierdo Millán *et al.* Solar cooling in Madrid: energetic efficiencies Solar Energy (1997)
- [7] A.A. Pesaran *et al.* Use of unglazed transpired solar collectors for desiccant cooling Solar Energy (1994)
- [8] M.R. Yeung *et al.* Performance of a solar-powered air conditioning system in Hong Kong Solar Energy (1992)

Audio Noise Reduction And Video Transcription

Geetika Dhand¹, Sukant Thakur², Amitesh Kumar³, Brijesh Kumar⁴

Department of Computer Science and Engineering

Maharaja Surajmal Institute of Technology, New Delhi, India

geetika.dhand@msit.in, thakursukant16@gmail.com, 07k.amitesh@gmail.com, itsbrijesh23@gmail.com

Abstract—Captioned videos provide advantages to individuals who view videos with inadequate audio or in different languages. Over a hundred empirical studies have recorded that the inclusion of captions enhances understanding, attentiveness, and recollection of the video content. Captions prove especially beneficial for individuals watching videos in a language that is not their primary one, for both children and adults who are in the process of learning to read, and for those who are hearing impaired. Additionally, captioning assists in mitigating the impact of background noise on audio quality. Consequently, captioning holds value across a wide range of settings, including workplaces, educational environments, and for various video audiences and creators.

1. INTRODUCTION

Envision a tool capable of enhancing children's reading skills, expanding the vocabulary of adolescents in both written and spoken forms, heightening college students' engagement with lectures, improving the pronunciation of second-language learners, and raising literacy rates in developing nations. The method behind this tool is straightforward: the display of subtitles on videos. These subtitles translate any spoken language into a familiar written language and appear at the bottom of the screen, enabling users to comprehend the video content even if they are unfamiliar with the native language in which the video is produced. Subtitles sometimes go beyond translation and include sound effects ("raindrops falling," "footsteps approaching," "horses galloping"). They also transcribe song lyrics and provide additional cues, such as identifying individuals involved in conversations by name and using italics to indicate off-screen voices.

Numerous studies have reported a broad range of benefits across various indicators, including summarizing main ideas, recalling information, drawing conclusions, defining words, relating emotions, and successfully answering multiple-choice comprehension questions. Eye-movement studies have confirmed that captions are read fluently, effortlessly attended to, and seamlessly integrated with the video's soundtrack. Both verbatim captions and more detailed versions have shown comparable effectiveness.

An abundance of empirical evidence demonstrates that captions have advantages for all video viewers, spanning from young children to older adults. Captions are particularly beneficial for individuals watching videos in a language that is not their native tongue, children and adults in the process of learning to read, and individuals who are deaf or hard of hearing.

2. AUDIO EXTRACTION AND ANALYSIS

For generating the transcript of a video, First we have to analyze its audio quality. For extracting the audio from a video file we use "Moviepy" a python library that is used for basic operation on video files.

2.1 Moviepy

MoviePy is a Python module designed for video editing purposes, offering functionality to perform various basic operations on videos and GIFs. A video is composed of a sequence of frames, with each frame resembling a regular image. Obtaining a frame refers to acquiring a numpy array that represents the Red-Green-Blue (RGB) image of a specific moment in a video clip, or for a sound clip, it refers to retrieving the mono or stereo value.

2.2 Visualization using Matplotlib

We visualized our extracted audio file (.wav file) in the x-y plane, and analyzed the amount of noise or disturbance in our video file using MATPLOTT, a python library.

We defined a function "visualize" which will read our .wav audio file and plot its graphical representation in the Signal-Time plane where X-axis shows the time in seconds and Y-axis shows signal strength.

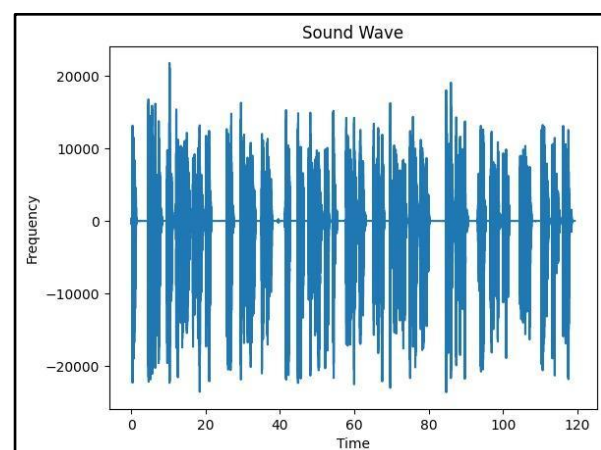


Fig 1: Graphical representation of sound wave

The matplotlib pyplot module is a collection of functions that enables matplotlib to emulate the functionality and behavior of Matlab. Each pyplot function is responsible for making specific modifications to a figure. For example, it can create a figure, establish a plotting area within the figure, plot lines on the plotting area, and add labels to enhance the visualization. These are just a few examples of the numerous capabilities provided by the pyplot module.

3. FILTER NOISE

After analyzing the audio file by graphical visualization we will reduce its noise frequency if any present to make it more clear.

To achieve this we use Butterworth Filter.

The Butterworth filter aims to achieve a frequency response with minimal variation within the pass band, leading to noise reduction.

When studying noise-related problems, we can classify them into two categories: non-additive noises and additive noises. Non-additive noises encompass multiplier noise and contortion noise, which can be transformed into additive noise through homomorphism transform.

In our literature survey, we have examined pulse noise issues arising from sources like engines, explosions, impacts, or discharges. The presence of noise from the surrounding environment significantly impacts the quality of speech and audio signals[1].

To address such noise problems, a low-pass Butterworth filter is commonly employed. The frequency response of the Butterworth filter is often referred to as "maximally flat" because it is designed to have minimal ripples in the pass band, maintaining a flat frequency response from 0Hz (DC) to the cutoff frequency at -3 decibels[2].

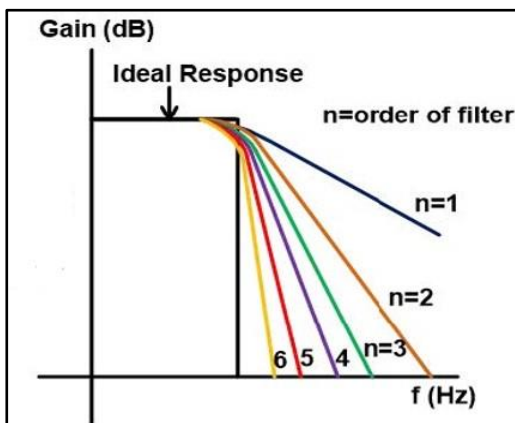


Fig 2: Butterworth filter

Nevertheless, the Butterworth filter has a notable drawback. While it achieves pass band flatness, it does so at the cost of a wide transition band when transitioning from the pass band to the stop band. This means that the filter exhibits poor phase characteristics[3][4]. The frequency response of a "brick wall" filter, which represents an ideal filter with no transition band, and the standard Butterworth filter responses for various filter orders are provided below.

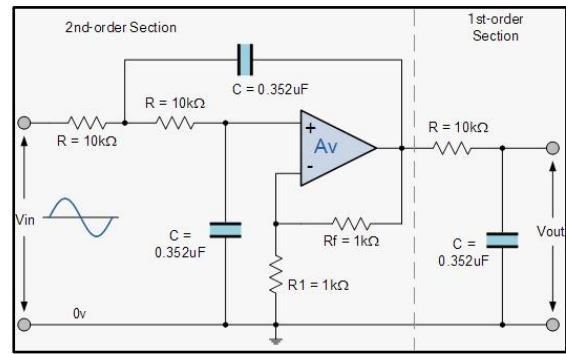


Fig-3 Butterworth filter circuit

4. SPEECH RECOGNITION AND SUBTITLE GENERATION

After the filtering the audio from the file our next task is to create a subtitle file from the audio. For this we will be using picovoice module pvleopard. Pvleopard works on local device speech to text engine. It provides us with various functions for getting the text from the audio with the time duration of how long it was pronounced in the audio. This module is available in various languages and frameworks like c, flutter, javascript, .net, java etc. We used the module in python.

Picovoice is an online platform for building voice products. There are a wide range of modules available on the platform where developers can start building their products using them. The main features of these products are accuracy, confidentiality and cross-platform. The accuracy of these products are very high and provide a very minimal amount of error. These modules run locally on the systems and thus allows us with the security of data. The products can be used with various types of framework.

To be able to use the pvleopard module online registration needs to be done and an access key needs to be attained. This access key is then given as an input to the module as registration. There exists a function process_file which is used to process the audio file to create two objects. First the transcript variable which contains the simple audio to text value of the video. The second variable is an object which contains each word in the audio with respective starting and ending time of that specific word.

The second variable is very important in this output as it would be used for the creation of a srt file for the video. The object is looped over and binding the words together according to the sentence in the audio. This sentence is then appended to a list which is being maintained side by side. These lines also have time duration corresponding to it which helps in identifying the duration in which that sentence is spoken. This list is then inputted in a subtitle file in srt format which can later be used to be played along with the video in any media player.

5. GOOGLE TRANSLATE AND MULTI-LANGUAGE SUPPORT

Google Translate is a free multilingual machine restatement service developed by Google, to restate textbook, speech, images, spots, or real-time videotape from one language into another. Google Translate was launched on April 28th, 2006 as statistical machine restatement.

The Python module called googletans is used for text conversion. It is a free and unrestricted library that utilizes the Google Translate API. This module makes use of the Google Translate Ajax API and employs similar methods like detect and translate to make API calls. Notable features of googletans include its fast and reliable performance, as it utilizes the same server as translate.google.com. It also offers automatic detection of the source language and provides the flexibility to customize the service URL according to specific requirements.

6. RESULTS

The media contents used in the design have undergone several tests during the trial phase, with slight modifications to the configurations for each run. These tests addressed specific issues related to the media components and met our expectations, providing relevant results, especially for small video clips. Each test video had varying levels of noise frequency, and the system proved to be highly effective in most scenarios.

Initially, the audio processing module experienced some exposure changes due to the limited audio formats it could handle. Although methods for converting audio files into suitable formats were established, it required careful preparation to ensure optimal functioning. To address this, we developed a custom Python module that yielded satisfactory results in terms of audio conversion. Notably, it supports a wide range of input formats, allowing direct conversion from an audio video format to the WAV format. This approach was instrumental in our trials and delivered the expected outcomes. To measure errors in our system, we employed two metrics: Word Accuracy (WA) and Word Error Rate (WER). These metrics are calculated as follows:

Word Accuracy (WA) = (total words - substitutions - deletions) / total words

Word Error Rate (WER) = total word errors / total words

Here, "total words" represents the number of words in the reference transcript.

By utilizing these measures, we were able to quantitatively assess the accuracy and error rate of our system's transcription capabilities.

Table 1. Working accuracy

	Time	Words	Matches	Error	WA (%)	WER (%)
1.	00:14:35	2706	1972	734	81	27
2.	00:03:36	531	375	112	66	29

7. CONCLUSION

We proposed a method to incorporate captions into audio videos. However, we were unable to fully implement a complete system due to resource limitations in the audio conversion process. As a temporary solution, we achieved

relevant results using VLC (VideoLAN Client), but we aim to develop a custom Python module in future work to enhance portability and simplify system installation.

Nevertheless, we successfully accomplished the desired outcomes for each phase of the project. The audio processing module ensures the audio is in a suitable format for the speech recognition module. This module generates a list of recognized words along with their corresponding timestamps in the audio, although the accuracy may vary.

The generated word list is utilized by the outline generation module to create standard caption lines that can be read by commonly used media players. These outlines serve as captions for the audio video, providing synchronized text that represents the spoken words.

While the complete system is yet to be realized, we have made significant progress in each module, paving the way for future improvements and the eventual implementation of an efficient and user-friendly captioning system.

REFERENCES

- [1] Manjeet Singh, Er.Naresh Kumar Garg - "Audio Noise Reduction using Butter worth Filter ",International Journal of Computer & Organization Trends –Volume 4
- [2] Jay Neil Gapuz, Gilbey's Jhon Ladion –"BACKGROUND NOISE REDUCTION BASED ON WIENER SUPPRESSION USING A PRIORI SIGNAL-TO-NOISE RATIO ESTIMATE IN PYTHON", InternationalJournal of Electrical, Electronics and Data Communication, ISSN(p): 2320-2084, ISSN(e): 2321-2950 Volume-6
- [3] Mike Schuste - "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation", arXiv:1609.08144 [cs.CL] <https://doi.org/10.48550/arXiv.1609.08144>
- [4] Jurišić, Dražen; Moschytz, George S.; Mijat, Neven (2008). "Low-sensitivity, singleamplifier, active- RC allpole filters using tables". *Automatika*
- [5] R.A.Barapate, J.S.Katre "Digital Signal Processing", Tech-Max January 2008 (Second revised edition).
- [6] Yu-Chi Tsao and Ken Choi, "Hardware-Efficient Vlsi Implementation for 3-Parallel Linear-Phase Fir Digital Filter Of Odd Length." *Circuits and Systems (ISCAS), IEEE Conference, ISBN: 978-1-4673-0218-0, pp.998 – 1001, 2012.*
- [7] Ravinder Kaur, Ashish Raman, Hardev Singh and Jagjit Malhotra., "Design and Implementation of High Speed IIR and FIR Filter using Pipelining", *International Journal of Computer Theory and Engineering, Vol. 3, No. 2, ISSN: 1793-8201, pp 292-295, 2011.*
- [8] Taan S. ElAli, "Discrete Systems and Digital Signal Processing with Matlab", CRC Press, ISBN 0-203-487117, 2004.
- [9] J.S. Chitode, "Digital Signal Processing", Technical Publication, Pune, ISBN: 9788184314243.
- [10] Dag Stranneby, "Digital Signal Processing-DSP & Application", Butterworth-Heinemann, Oxford, ISBN: 0750648112, 2001.
- [11] Hala ElAarag and Laura Schindler. A speech recognition and synthesis tool. *ACM SE'06, pages 45–49, 2006.*
- [12] Frederick Jelinek. *Statistical Methods for Speech Recognition.* MIT Press, 1999.
- [13] Zoubin Ghahramani. An introduction to hidden markov models and bayesian networks. In *World Scientific Publishing Co., Inc. River Edge, NJ, USA, 2001*
- [14] Erik Hatcher and Steve Loughran. *Java Development with Ant.* Manning Publications, 2002.
- [15] Kasper Winther Joergensen and Lasse Lohilahti Moelgaard. Tools for automatic audio indexing. Master's thesis, Kongens Lyngby, 2006.

Reviewers

<p>Dr. Vijander Singh Professor, Department of Instrumentation and Control Engineering, Netaji Subhash University of Technology vijaydee@nsut.ac.in</p>	<p>Dr. Navin Rajpal, Professor, University School of Information & Communication Technology GGSIU, New Delhi navin.rajpal@ipu.ac.in</p>
<p>Dr. Jaya Srivastava Senior Programmer IIT Delhi jsrivastava@rgipt.ac.in</p>	<p>Dr. Balarama Ravindra Professor, Department of Computer Science Engineering Indian Institute of Technology, Madras ravi@cse.iitm.ac.in</p>
<p>Dr. Jagdish Kumar Professor & Head, Department of Electrical Engineering Punjab Engineering College, University of Technology (PEC Chandigarh) jagdishkumar@pec.edu.in</p>	<p>Dr. A L Sangal Professor & Head, Department of Computer Science Engineering Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab sangalal@nitj.ac.in</p>
<p>Dr. Lokendra Kumar Professor, Department of Mathematics Jaypee Institute of Information Technology lokendma@gmail.com</p>	<p>Dr. Manisha Bharti Associate Professor, ECE Department, NIT Delhi manishabharti@nitdelhi.ac.in</p>
<p>Dr. Charu Gupta Associate Professor, Department of Computer Science Engineering Bhagwan Parshuram Institute of Technology, New Delhi charugupta@bpitindia.com</p>	<p>Dr. Vishwambhar Nath Mishra Professor & Head of Department Department of Electronics Engineering, Indian Institute of Technology, Uttar Pradesh, India vnmishra.ece@iitbhu.ac.in</p>
<p>Dr. Amrish Aggarwal Operations Research and Numerical Technique Jaypee Institute of Information Technology amrish.aggarwal@jiit.ac.in</p>	<p>Dr. Soumitra Das Professor-HOD ,Electrical and Electronics Engineering National Institute of technology Goa (NIT Goa) hod.eee@nitgoa.ac.in</p>
<p>Dr. RamaKoteswara Rao Alla Associate Professor Department of Electrical and Electronics Engineering, RVR & JC College of Engineering, Andhra Pradesh, India ramnitkkr@gmail.com</p>	<p>Dr. P Ramani, Assistant Professor, Electronics & Communications Engineering SRM University, Gurugram, India</p>
<p>Mr Ravita Saraswat JRF EE Dept NIT Kurukshetra ravita.saraswat009@gmail.com</p>	<p>Dr. Rajesh Mehta Assistant Professor, Department of Computer Science Engineering Thapar Institute of Engineering and Technology rajesh.mehta@thapar.edu</p>
<p>Mr. Saurabh Gulati Test Manager Robert Bosch Engineering and Business Solutions Private Limited saurabh.gulati@bosch.com</p>	<p>Dr. Manoj Kumar Professor ,University School of Information & Communication Technology(IPU) GGSIU, New Delhi manojtaleja@ipu.ac.in</p>

<p>Dr. Archna Sandhu Associate Professor Geeta Engineering College Panipat, Haryana asandhu@geeta.edu.in</p>	<p>Dr. V.P. Vishwakarma Professor ,University School of Information & Communication Technology(IPU) GGSIPU, New Delhi ypv@ipu.ac.in</p>
<p>Dr. Vahida Zakirhusen Attar Computer & IT Department College of Engineering Pune vahida@comp.ac.in</p>	<p>Dr A L Sangal Professor (HAG) & Head Computer Science &Engineering (CSE) Dr B R Ambedkar National Institute of Technology, sangalal@nitj.ac.in</p>
<p>Dr. Amita Malik Professor-Computer Science Engineering (CSE) Deenbhandhu Chhotu Ram University of Science and Technology Murthal Sonapat (Haryana) amitamalik@dcrustm.ac.in</p>	<p>Dr. Tajinder Singh Arora Assistant Professor, Electronics & Communications Engineering National Institute of Technology, Uttarakhand tajarora@gmail.com</p>
<p>Dr. Kavita Khanna, Professor and HOD - CSE The NorthCap University Gurugram kavitakhanna@ncuindia.edu</p>	<p>Dr. Jyotsna Yadav Assistant Professor, University School of Information & Communication Technology(IPU) GGSIPU, New Delhi jyotsnayadav@ipu.ac.in</p>
<p>Dr. Ashish Bhardwaj HOD, CSE, GTBIT aashish@gtbit.org</p>	<p>Dr. Amit Kumar Singh Assistant Professor, Department of Instrumentation & Control Engg. Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab singhak@nitj.ac.in</p>
<p>Dr. Akhilesh Singh Assistant Professor, NPS Engineering Institute, Government college, Uttarakhand</p>	<p>Dr Rajneesh Singh Assistant Professor, ECE IIIT, Nagpur rsingh@iiitn.ac.in</p>
<p>Dr. Jayshree Goswa Assistant Professor, Law Institute Alipore, Kolkata, jgayshreegoswami@gmail.com</p>	<p>Dr Varsha Assistant Professor, Chandigarh University varshae10453@cumail.in</p>
<p>Dr Mahak Singh Associate Professor, mahaksingh@ramjas.du.ac.in</p>	<p>Dr. Akhilesh Singh Assistant Professor, NPS Engineering Institute, Government college, Uttarakhand</p>
<p>Dr H.C. Nagaraj Professor and Principal, NMIT, Bengaluru nagaraj.hc@nmit.ac.in, NMIT</p>	<p>Dr. Chinna Ayya Swami Assistant Professor, NIT Calicut, swamy@nitc.ac.in</p>
<p>Dr. Rohit Beniwal Assistant Professor, IT dept., DTU dr.rohitbeniwal@gmail.com</p>	<p>Dr Anu Saini Assistant Professor, CSE dept., GB Pant Inst. of Tech., anuanu16@gmail.com</p>
<p>Dr.Indu Kashyap Professor, MRIIRS indu.kashyap82@gmail.com</p>	<p>Dr.Niranjanamurthy M Assistant Professor, Dept. of ISE, ,BMSIT, Bangalore niruhsd@gmail.com</p>
<p>Dr Janpreet Singh Lovely Professional University, Punjab janpreet.singh@lpu.co.in</p>	<p>Dr Amita Seehra HOD, ECE GNDEC, Ludhiana</p>
<p>Dr. Shweta Singhal, miss.shweta.singhal@gmail.com</p>	<p>Dr. Deepa Gupta Sen. Lect, CSE, Amity University, Noida dgupta@amity.edu</p>

<p>Dr. RamaKoteswara Rao Alla Associate Professor Department of Electrical and Electronics Engg, RVR &JC College of Engineering, Guntur, Andhra Pradesh, India</p>	<p>Dr.M.Gulam Nabi Alsath Associate Professor, ECE Department, SSN Institutions, Chennai. gulamnabialsathm@ssn.edu.in</p>
<p>Dr. Amit Sharma Associate Prof., IT Department, IMSEC Ghaziabad. amit.faculty@gmail.com</p>	<p>Dr.Pushpendra Singh Associate Prof., IT Department,Raj Kumar Goel Institute of Technology, Ghaziabad pushpit@rkgit.edu.in</p>
<p>Dr. Yogendra Arya Associate Professor, EEE, YMCA, Faridabad yarya@jcboseust.ac.in</p>	<p>Dr. S. Anita Shanthi Associate Professor, Mathematics, Anna Malai University shanthi.anita@yahoo.com</p>
<p>Dr. Pooja Singh Assistant Professor, BCA, MSI Pooja.asm@gmail.com</p>	<p>Mr. Dhyan Pachnanda Business Analyst , Connect Compusys Pvt.Ltd</p>
<p>Dr Preeti Nagrath Associate Professor, Bharati Vidyapeeth, New Delhi preeti.nagrath@bharatividyaeeeth.edu</p>	<p>Dr. Lokendra Kumar Professor, Department of Mathematics, Jaypee Institute of Information Technology</p>
<p>Dr Parveen Poon Terang Electrical Engineering Dept., Associate Professor JSSATE, Noida, India ppterang.eed@jssaten.ac.in</p>	<p>Dr Neha Gupta Assistant Professor/Programme Director EE School of Engineering & Technology. Sushant University, Gurgaon guptaneha2207@gmail.com</p>
<p>Dr. Archana Balyan Professor, ECE Department, MSIT</p>	<p>Dr. Alka Singhla Snr. Assistant PProfessor CS department, IIIT, Noida</p>
<p>Dr Naveen Dahiya Professor, CSE Department, MSIT</p>	<p>Dr Meena Tushir Professor, EEE Department, MSIT</p>
<p>Dr Koyel Dutta Professor, CSE Department, MSIT</p>	<p>Dr Naresh Kumar Professor, CSE Department, MSIT</p>
<p>Dr Puneet Azad Professor, ECE Department, MSIT</p>	<p>Dr Neeru Rathee Professor, ECE Department, MSIT</p>
<p>Dr. Pradeep Sangwan Professor, ECE Department, MSIT</p>	<p>Dr. Tripti Sharma Professor, IT, MSIT</p>
<p>Prof. Man Singh Beniwal HOD, Department of Applied Sciences,MSIT</p>	<p>Dr. Anupama Kaushik Associate ,IT Department, MSIT</p>
<p>Dr Kavita Sheroan Associate Professor, CSE Department, MSIT</p>	<p>Dr Vimal Gaur Associate Professor, CSE Department, MSIT</p>
<p>Dr Shaifali M. Arora Associate Professor, ECE Department, MSIT</p>	<p>Dr Sudesh Pahal Associate Professor, ECE Department, MSIT</p>
<p>Dr Sunil Gupta Associate Professor, EEE Department, MSIT</p>	<p>Dr Meena Rao Associate Professor, ECE Department, MSIT</p>
<p>Dr Geetika Associate Professor, CSE Department, MSIT</p>	<p>Dr Amita Yadav Associate Professor, CSE Department, MSIT</p>
<p>Dr Parul Assistant Professor, CSE, Department, MSIT</p>	<p>Dr Deepti Assistant Professor, ECE Department, MSIT</p>
<p>Dr Anshul Pareek Assistant Professor, ECE Department, MSIT</p>	<p>Dr. Poonam Assistant Professor, ECE Department, MSIT</p>
<p>Dr. Monu Malik Assistant Professor,EEE, Department, MSIT</p>	<p>Dr. Sitender Assistant Professor, IT Department, MSIT</p>
<p>Dr. Nidhi Gupta Assistant Professor,EEE, Department, MSIT</p>	<p>Dr. Rakhi Kamra Assistant Professor, EEE Department,MSIT</p>

