



PANIMALAR ENGINEERING COLLEGE



A Christian Minority Institution-Jaisakthi Educational Trust
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai
Accredited by National Board of Accreditation(NBA)
Approved by UGC for 2(f) & 12(B) Status
Bangalore Trunk Road, Poonamallee, Chennai-600123.



PECMACT PECTEAM 2K21

This is to certify that Dr./Mr./M **RAVIKUMAR N** of CEM from **MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE**. has presented paper entitled CONTRIVANCING 4D-BIM IN METRO RAIL PROJECT USING POWER PROJECT SOFTWARE in the 4th Phoenixes on Emerging Current Trends in Materials, Manufacturing, Management Practices And Construction Technologies (PECMACT) held on 26th & 27th March 2021



CONVENER

Dr.S.MALATHI,M.E.,Ph.D

CO-CONVENER

Dr.V.D.Ambeth Kumar,M.E.,Ph.D.,

CO-PATRON

DR.K.MANI,M.E.,Ph.D.,





ADITHYA INSTITUTE OF TECHNOLOGY

Accredited by NAAC

Sathy Road, Kurumbapalayam, Coimbatore, Tamil Nadu 641107



Certificate of Participation

This is to certify that

Mrs.M.MALINI GAYATHRI

Meenakshi Sundararajan Engineering College

presented a paper titled "ANALYSIS OF PROBLEMS IN PROCUREMENT OF MATERIAL AND EQUIPMENT IN CONSTRUCTION INDUSTRY" at the International Conference on "Sustainable Materials and its Green Technologies for Industry 4.0"-SMiGT-2021 held on, 07-09 April, 2021, Organized by Department of Civil Engineering, Adithya Institute of Technology, Coimbatore, India in association with The Institution of Engineers (India), Coimbatore Local Centre.

Organizing chair
(Dr.P. Magudeaswaran)

Principal
(Dr.D.Somasundareswari)

GEO
(Dr.Chitra Manohar)

SMiGT/CP-105



ADITHYA INSTITUTE OF TECHNOLOGY

Accredited by NAAC

Sathy Road, Kurumbapalayam, Coimbatore, Tamil Nadu 641107



Certificate of Participation

This is to certify that

Mr.N.RAVI KUMAR

Meenakshi Sundararajan Engineering College

presented a paper titled "CONTRIVANCING 4D-BIM IN METRO RAIL PROJECT USING POWER PROJECT SOFTWARE" at the International Conference on "Sustainable Materials and its Green Technologies for Industry 4.0"-SMiGT-2021 held on, 07-09 April, 2021, Organized by Department of Civil Engineering, Adithya Institute of Technology, Coimbatore, India in association with The Institution of Engineers (India), Coimbatore Local Centre.

Organizing chair
(Dr.P.Magudeaswaran)

Principal
(Dr.D.Somasundareswari)

CEO
(Dr.Chitra Manohar)

SMiGT/CP-099



SRM TRP ENGINEERING COLLEGE

SRM NAGAR, IRUNGALUR - 621105, MANNACHANALLUR TALUK,
TRICHIRAPALLI DISTRICT, TAMIL NADU



DEPARTMENT OF CIVIL ENGINEERING

Virtual National Conference on Sustainable Technologies and Smart Materials
(NCSSM - 2021)


Certificate of Presentation

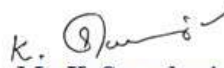
This is to certify that the following paper has been presented in the "Virtual National conference on Sustainable Technologies and Smart Materials (NCSSM - 2021)" held on 10th April 2021


Title of the paper : **GROUND WATER QUALITY ASSESSMENT ALONG COOUM
RIVER**

Author : **RAVI KUMAR N , Assistant Professor / Civil Engineering**
MEENAKSHI SUNDRARAJAN ENGINEERING COLLEGE.


Mr. S. Manimaran
Coordinator


Mr. R. Naveen Kumar
Coordinator


Mr. K. Sureshraj
Convenor


Dr. B. Ganesh Babu
Principal



RAMCO INSTITUTE OF TECHNOLOGY

Rajapalayam-626117

Approved by AICTE, New Delhi, Affiliated to Anna University, Chennai,
Accredited by NAAC & an ISO 9001:2015 Certified Institution



***NATIONAL WEB CONFERENCE
ON
CHALLENGES AND INNOVATION IN ENGINEERING & TECHNOLOGY 2021
(In Association with Professional Societies)***

Certificate of Presentation

This is to certify that Mrs.G.Anbuneema , Meenakshi Sundhararajan Engineering College, Chennai. has presented a paper entitled "Durability Study On Foam Concrete With The Effect of Silica Fume And Polypropylene Fiber" in the National Web Conference on Challenges and Innovation in Engineering & Technology 2021, during 19 - 20 March 2021.

Dr.S.Periyanayagi
Organising Secretary

Dr.S.Rajakarunakaran
Vice Principal

Dr.L.Ganesan
Principal

Certificate ID: NWCCIET_CIVIL_24_01



RAMCO INSTITUTE OF TECHNOLOGY

Rajapalayam-626117



Approved by AICTE, New Delhi, Affiliated to Anna University, Chennai,
Accredited by NAAC & an ISO 9001:2015 Certified Institution

***NATIONAL WEB CONFERENCE
ON
CHALLENGES AND INNOVATION IN ENGINEERING & TECHNOLOGY 2021
(In Association with Professional Societies)***

Certificate of Presentation

This is to certify that Mr.G.Amirtha Gokul, Meenakshi Sundhararajan Engineering College, Chennai. has presented a paper entitled "Durability Study On Foam Concrete With The Effect of Silica Fume And Polypropylene Fiber" in the National Web Conference on Challenges and Innovation in Engineering & Technology 2021, during 19 - 20 March 2021.

Dr.S.Periyamayagi
Organising Secretary

Dr.S.Rajakarunakaran
Vice Principal

Dr.L.Ganesan
Principal

Certificate ID: NWCCIET_CIVIL_24_02

SRI SIVASUBRAMANIYA NADAR COLLEGE OF ENGINEERING
(AN AUTONOMOUS INSTITUTION - AFFILIATED TO ANNA UNIVERSITY)
KALAVAKKAM - CHENNAI - 603110, TAMIL NADU, INDIA



DEPARTMENT OF MATHEMATICS
AND
DEPARTMENT OF CIVIL ENGINEERING

CERTIFICATE

This is to certify that Prof./Dr./Mr./Ms.
G. AMIRTHA GOKUL of Meenakshi Sundararajan Engineering College, presented a
paper titled A PLAXIS 2D DISPLACEMENT ANALYSIS OF GEOTEXTILE TUBE
TECHNOLOGY TO PREVENT COASTAL EROSION AT SANGUDURAI BEACH OF
KANNIYAKUMARI DISTRICT

at the INTERNATIONAL CONFERENCE ON APPLIED
MATHEMATICS & CIVIL STRUCTURES (ICAMCS 2021) jointly
organised by the Department of Mathematics and Department of Civil
Engineering, Sri Sivasubramaniya Nadar College of Engineering,
Kalavakkam, on February 12 -13, 2021.

Dr. P. Venugopal
HOD/Mathematics

Dr. N. Sivakumar
HOD/Civil

Dr. V.E. Annamalai
Principal



Certificate of Publication



This is to confirm that

Ajay. B, Udaya. T, Dr.B.Monica Jenefer

Published following article

Minimato Food Ordering Single Page Application with Json Web Tokens
Authentication Using Pern Stack

Volume 3, Issue 6, pp: 1054-1056

www.ijaem.net

A Peer Reviewed Journal

**International journal of Advances in Engineering
and Management (IJAEM)**

ISSN: 2395-5252

Publication Head

Minimato - Food Ordering Single Page Application with Json Web Tokens Authentication Using Pern Stack

Ajay. B, Udaya. T, Dr.B.Monica Jenefer

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai
Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai
Head of the department, Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Submitted: 01-06-2021

Revised: 14-06-2021

Accepted: 16-06-2021

ABSTRACT: Now a day, it is a mandate for each and every person to maintain social distancing and to reduce people crowds due to the global pandemic, COVID. Our project brings up a solution for maintaining social distance and also automating the manual food ordering system in college canteens. The aim of this project is to digitize ordering of foods in the college canteen and eliminate the waiting time (approximately 5 to 8 minutes) in queues and the human to human contact. An online food ordering application - Minimato is specifically developed for college canteens. The students can browse through multiple canteens and their menus and orders can be placed in few clicks. Canteen owners can easily navigate to the user interface to update the menu and manage the incoming orders. The order history is well maintained and can use for future calculation purpose. This application mainly focuses on authentication part using JSON Web Tokens on both canteen accounts and student accounts. The system also uses Auto Suggestion for displaying the canteens based on the user input. By implementing this system the food ordering system is digitalized and the waiting time and human to human interaction is reduced.

KEYWORDS: Online Food Ordering Application, JWT Authentication, Canteen, JSON.

I. INTRODUCTION

During this covid pandemic, it is necessary for us to minimize the human to human contact and stay safe. The aim of this project is to digitize ordering of foods in our college canteen. This project allows us to achieve this by eliminating the need to wait in queues and manually order food at a particular canteen. The system uses JWT Authentication for secure transactions and storing server data. Also, the

system uses Multer for image storage and Twilio for sending text messages for the user. The user can take-away the food from the canteen.

[1]. JSON Web Token (JWT) is an open standard (RFC 7519) that defines a compact and self-contained way for securely transmitting information between parties as a JSON object. This information can be verified and trusted because it is digitally signed. It is an Internet proposed standard for creating data with optional signature and/or optional encryption whose payload holds JSON that asserts some number of claims. The tokens are signed either using a private secret or a public/private key.

[2]. Multer is a popular Node.js middle ware used for handling multi-part / form-data requests. It makes use of busboy to parse any data that it received through an HTML form. This greatly enhances its performance because the busboy module is unmatched when it comes to analyzing form data. It provides us control and flexibility when handling requests - we get detailed information about each uploaded file, the ability to add a custom storage engine, validation of files according to our needs, the ability to set limits on uploaded files.

[3]. A Twilio Messaging Service is a messaging application, usually specific to a use-case or campaign that is powered by a group of phone numbers that you select based on the needs of that use-case. It is an API to send and receive SMS, MMS, and OTT messages globally. It uses intelligent sending features to ensure messages reliably reach end users wherever they are. We can enable the messaging service and select phone numbers all within the Twilio Console.

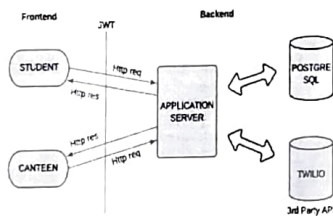
[4]. The aim is to digitalize the manual food ordering system in the college canteens. Keeping in mind the covid pandemic, bustling

canteen can lead to more human to human interaction. The students have to wait in long queues to get their food. Paper work is needed to maintain the order history. Also humans are prone to making mistakes or errors in managing and tracking the errors. Every canteen needs an employee for taking the order and processing the order and the labor rates are increasing every now and then.

II. METHOD OF STUDY

This paper surveys the journals of recent years. The main objective is to contemplate knowledge regarding efficient online food ordering systems. The surveyed paper focuses on several works contributed to online food ordering. We looked at several databases like IEEE, ResearchGate, IEEE Xplore, and other journals. We have also considered each article's references to get relevant papers.

III. SYSTEM ARCHITECTURE



SYSTEM ARCHITECTURE

The student can register and login into the system and view the available canteens. After the login, a JWT token is generated for each student, which is stored in the local system. The student can view the menu and order food from one of the listed canteen and he/she can view the cart and proceed for checkout. The canteens can be registered using the Canteen login and the menu can be updated and the incoming orders are maintained. The system also using Twilio for sending text messages to the students after the food is ready.

IV. SYSTEM IMPLEMENTATION

In the system construction, we use PERN stack for developing the system. The web application is created using react and for the backend node js. The database operations are performed using PostgreSQL. The following are the modules used in the system and their associated processes.

A. USER LOGIN/ SIGNUP:

The user can either perform the login or the signup. The signup option can be used for the new users. This module uses the JWT Authentication for storing the username and password in the database. It also ensures that the password is stored in the database in encrypted form. And also JWT is used for user validation purpose.

B. SEARCHING THE CANTEEN

The student can search for the canteen that are available in the list. This module uses Auto Suggestion. By enabling the auto suggestion, all the possible canteens are displayed based on the alphabet typed by the student.

C. ORDERING FOOD

The students can go through the food menu available in different canteens Upon selecting the canteen, the students can either add or delete food items to the cart and also view the cart.

D. CHECKOUT

The students can re-view the cart and can further add or delete food items to the list and finally proceed for placing the order.

E. CANTEEN LOGIN/ SIGNUP

The canteen manager can either perform the login or the signup. The signup option can be used for registering the new canteens. This module uses the JWT Authentication for storing the username and password in the database. It also ensures that the password is stored in the database in encrypted form. And also JWT is used for canteen validation purpose.

F. UPDATE MENU

The canteen managers can login and can update their food menu by adding food items to the list or deleting food items from the list.

G. INCOMING ORDER

The canteen manager can login into the business account and maintain the incoming food orders. The canteen managers can either accept the order or delete the food order.

H. NOTIFYING THE USER

After the food is processed, the canteen managers intimate the students that their food is ready via a text message. Finally, the students make the payment using cash on delivery method and collect their food from the respective canteens.



USER DASHBOARD

V. CONCLUSION

A food ordering system has been developed for college canteens that effectively reduce the human to human contact. The system automates the manual food ordering process by allowing the users to order the food without physically going to the canteen. Users can create orders through the web interface. The menus, canteens, users, and orders can be managed by the canteen managers. The system effectively reduces the waiting time, which can be used in a productive way. Hence this system works towards achieving to minimize the human contact and the waiting time.

VI. FUTURE ENHANCEMENT

Certain points of improvements on both technical and aesthetic fronts can be achieved in the future. An additional functionality of food recommendation system can be incorporated in the system. This would display the canteens based on the user's choice and interest. The payment system used in this system can be updated to an online mode payment. Lastly, based on the feedback from the user, the canteen ratings can be improved. Such enhancements and more are possibly to be implemented to make the system better and give a more intuitive experience to the user.

VII. ACKNOWLEDGEMENT

Our primary and sincere thanks go to our pillar of support Dr. B. Monica Jenefer, Head of the Department, Department of Computer Science and Engineering, Meenakshi Sundararajan Engineering College for her profound inspiration, kind cooperation and guidance. We would like to thank our Professor Dr. M. K. Sandhya for her invaluable words of encouragement and guidance.

REFERENCES

- [1]. Liyanage V, Ekanayake A, Premasiri H, Munasinghe P, Thelijjagoda, (2018), "Foody - Smart Restaurant Management and Ordering System", IEEE Region 10 Humanitarian Technology Conference, IEEE 2018, 10.1109/R10-HTC.2018.8629835.
- [2]. Varsha Chavan, Priya Jadhav, Snehal Korade, Priyanka Teli, (2015), "Implementing Customizable Online Food Ordering System Using Web Based Application", International Journal of Innovative Science, Engineering Technology, IJSET 2015, http://ijiset.com/vol2/v2s4/IJSET_V2_I4_112.pdf.
- [3]. Kirti Bhandge, Tejas Shinde, Dheeraj Ingale, Neeraj Solanki, Reshma Totare, (2015), "A Proposed System for Touchpad Based Food Ordering System Using Android Application", International Journal of Advanced Research in Computer Science Technology, IJARCSST 2015.
- [4]. Domokos C E, Sera B, Simon K, Kovacs L, Szakacs T B, (2018), "Netfood: A Software System for Food Ordering and Delivery", IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), IEEE 2018, 10.1109/SISY.2018.8524854.
- [5]. Trupthi B, Rakshitha Raj R, J B Akshaya, Srilaxmi C P, (2019), "Online Food Ordering System", International Journal of Recent Technology and Engineering, IRJET 2019, <https://www.irjet.net/archives/V5/i6/IRJET-V5I679.pdf>.
- [6]. Mayur D. Jakhete, Piyush C. Mankar, (2015), "Implementation of Smart Restaurant with e-menu Card", International Journal of Computer Applications, IJSER 2015, <https://www.ijser.org/researchpaper/IMPLEMENTATION-OF-SMART-RESTAURANT-WITH-E-MENU-CARD.pdf>.

ISSN 2229-5518

CERTIFICATE OF ACCEPTANCE

International Journal of Scientific & Engineering Research
(IJSER)



THIS IS TO CERTIFY THAT OUR REVIEW BOARD HAS ACCEPTED RESEARCH PAPER OF

Mrs. C JERIN MAHIBHA

FACE MASK DETECTION AND RECOGNITION BASED ATTENDANCE
SYSTEM USING DEEP LEARNING

May 26, 2021
Visit us at: www.ijser.org



Editor in Chief

FACE MASK DETECTION AND RECOGNITION BASED ATTENDANCE SYSTEM USING DEEP LEARNING

Mrs. C JERIN MAHIBHA
Assistant Professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India jerinmahibha.msec@gmail.com

RAJESH KUMAR C
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
kumarsep5@gmail.com

ARULBALAN A
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
arulbalan84@gmail.com

Abstract The face is one of the easiest ways to distinguish the individual identity of each other. Face recognition is a personal identification system that uses personal characteristics of a person to identify the person's identity. After the breakout of the worldwide pandemic COVID-19, face mask being the primary one to protect ourselves. The corona virus COVID-19 pandemic is causing a global health crisis so the effective protection method is wearing a face mask in public areas according to the World Health Organization (WHO). The COVID-19 pandemic forced governments across the world to impose lockdowns to prevent virus transmissions. Reports indicate that wearing facemasks while at work clearly reduces the risk of transmission. An efficient and economic approach of using AI to create a safe environment in a manufacturing setup. A hybrid model using deep and classical machine learning for face mask detection has been presented. A face mask detection dataset consists of images such as with mask and without mask, by using OpenCV for real-time face detection from a live stream via our webcam. The face mask recognition dataset consists of images of a person with a mask. The dataset to build a COVID-19 face mask detector with computer vision using Python, OpenCV, and TensorFlow and Keras. After detecting and recognizing the person, the attendance is generated.

Key Words: Deep Learning, OpenCV, Tensorflow, Keras, CNN.

I. INTRODUCTION

Face recognition is the task of identifying an already detected object as a known or unknown face. Often the problem of face recognition is confused with the problem of face detection. Face Recognition on the other hand is to decide if the "face" is someone known, or unknown, using for this purpose a database of faces in order to validate this input face.

The trend of wearing face masks in public is rising due to the COVID-19 corona virus epidemic all over the world. Before Covid-19, People used to wear masks to

protect their health from air pollution. While other people are self-conscious about their looks, they hide their emotions from the public by hiding their faces. Scientists proofed that wearing face masks works on impeding COVID-19 transmission. COVID19 (known as corona virus) is the latest epidemic virus that hit the human health in the last century. In 2020, the rapid spreading of COVID-19 has forced the World Health Organization to declare COVID-19 as a global pandemic. More than five million cases were infected by COVID-19 in less than 6 months across 188 countries. The virus spreads through close contact and in crowded and overcrowded areas. The corona virus epidemic has given rise to an extraordinary degree of worldwide scientific cooperation.

Artificial Intelligence (AI) based on Machine learning and Deep Learning can help to fight Covid-19 in many ways. The provision of healthcare needs funding for emerging technology such as artificial intelligence, IoT, big data and machine learning to tackle and predict new diseases. In order to better understand infection rates and to trace and quickly detect infections, the AI's power is being exploited to address the Covid-19 pandemic. People are forced by laws to wear face masks in public in many countries. These rules and laws were developed as an action to the exponential growth in cases and deaths in many areas. However, the process of monitoring large groups of people is becoming more difficult.

The monitoring process involves the detection of anyone who is not wearing a face mask. Here we introduce a mask face detection model that is based on computer vision and deep learning. The proposed model can be integrated with web camera to impede the COVID-19 transmission by allowing the detection and recognition of people who are wearing masks or not wearing face masks. After recognize the person, the attendance is generated for that person. The model is integration between deep learning and classical machine learning techniques with OpenCV, tensor flow and keras. We have used deep transfer learning for feature extractions and combined it with three classical machine learning algorithms.

II. RELATED WORKS

A. TENSORFLOW

TensorFlow is an end-to-end open-source platform for machine learning and software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.

B. KERAS:

Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear & actionable error messages. Keras contains numerous implementations of commonly used neural network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code.

C. OPENCV:

OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc.

D. CONVOLUTIONAL NEURAL NETWORK:

CNN is a Deep Learning Algorithm that takes images as an input, gives importance to certain features in image and creates differences among various objects in an image from each other. CNN is been inspired due to connectivity pattern of human nervous system. They are built of neurons along biases and learnable weights. Multiple inputs are accepted by each neuron, weighted sum of all inputs is directed for an activation function resulting in the output. Basically, convolution is sliding filter on an image and taking dot product along its way while sliding. Therefore, results a scalar quantity. We take input as tensor in CNN which is multidimensional matrix of a number. Convolution layers over tensor input are the major building blocks in CNN. Every layer depends on an unconstrained filter which is achieved by convolution of filter with image. Filters are

arbitrary initialized. we make them our parameters by training the network multiple times. In initial layers the filter depicts some basic level of features like colours, edges, gradients etc. While going deep in Convolution layers we sort out high level of features that gives us a system having complete knowledge of the image.

III. SYSTEM ARCHITECTURE

In this system, the dataset is loaded to check for cleanliness and validation, and visualize the dataset for better understanding of the patterns and to find anomalies that increase the data values in analysis. The splitting of dataset is an essential step in training any of the systems.

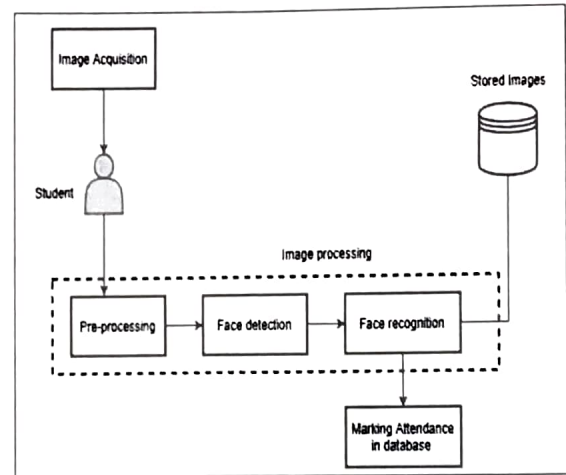


Fig 1. Architecture face recognition system

The architecture explains how the data and workflow throughout the system. The image of a student is sent to the system. The system starts the work of pre-processing the image that helps the input image to fit in our model. After pre-processing the image, the feature has been extracted using CNN. Using the trained model for face mask detection, it starts to detect the image where the mask is present or not. If the mask is present in the image, the person will be recognized using face recognition trained model. The pre-trained images of both face mask detection and recognition have been stored in the database. Finally, the attendance is generated to the face mask recognized person.

IV. SYSTEM IMPLEMENTATION

A. FEATURE EXTRACTION:

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. So, Feature extraction helps to get the best feature from those big data sets by select and combine variables into features, thus, effectively reducing the amount of data we have three Matrices or the channels Red, Green and Blue. So, in these

matrices, each of the matrix has values between 0-255 which represents the intensity of the colour of that pixel. The image is read using numpy library which is used to read the image and converted it to an array of an image. Now, we need to scale the image in the range 0 and 1, so the array of image is divided by 255. The array of an image is the extracted feature from this system.

B. PRE-PROCESSING:

Data Pre-processing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis. Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. In this system, the dimensions of the image are 28 x 28. But the input image can be different size. So, in this case, the image is read and can be resized into 28 x 28 dimension. Sometimes, the images can be rotated and can be dark or light image. In this case, the image is need to be straighten and normalization.

C. TRAINING:

Deep learning neural networks learn a mapping function from inputs to outputs. This is achieved by updating the weights of the network in response to the errors the model makes on the training dataset. Updates are made to continually reduce this error until either a good enough model is found or the learning process gets stuck and stops. The process of training neural networks is the most challenging part of using the technique in general and is by far the most time consuming, both in terms of effort required to configure the process and computational complexity required to execute the process. Unlike other machine learning algorithms, the parameters of a neural network must be found by solving a non-convex optimization problem with many good solutions and many misleadingly good solutions. H5 is a file format to store structured data, it's not a model by itself. Keras saves models in this format as it can easily store the weights and model configuration in a single file.

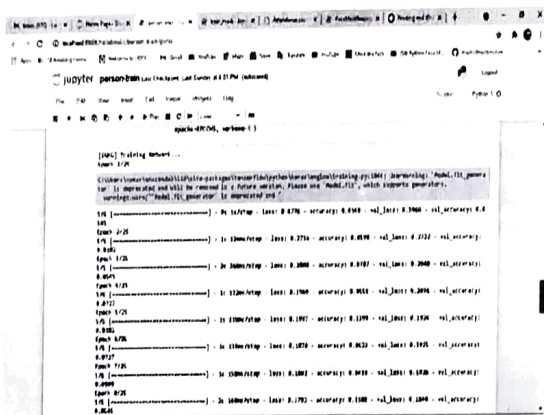


Fig 2. Training Model

D. FACE MASK DETECTION:

The system classifies the dataset in two classes (with/without mask). Using another neural net (MobileNetV2) to develop the model to detects the face with mask. Scores for the image from 0 to 1 signifying the probability of a face wearing a mask. The pre-trained face detection model seems to work great for his case, and it detects faces even when they are partially covered by masks. As there was no pre-trained classifier to distinguish faces with and without masks. The collection of datasets available in the internet (Kaggle) with nearly ~1600 images in a very clever and effective way. The dataset with regular faces in the presence of masks to the images. By using the h5trained model, the system will detect the image with mask or without mask.

E. FACE MASK RECOGNITION:

Face classifier is essential part of this paper and presents our contribution to masked face recognition. Deep learning is one of the most novel ways to improve face recognition technology. The idea is to extract face embeddings from images with faces. Such facial embeddings will be unique for different faces. The ultimate task is finding the unique features in added to the mask in face. So, this process works in the way of classification. It means classify the people as classes. A masked face classifier was built on feature extractor of CNN. In this paper, each person is considered as one class and each class having the images of that person.

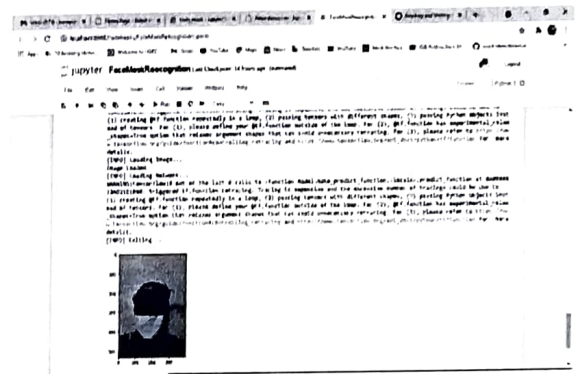


Fig 3. Face mask recognition

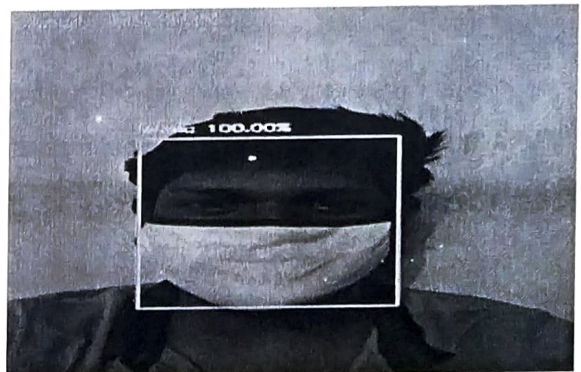


Fig 4. Final result

V. CONCLUSION AND FUTURE WORK

As the technology is blooming with emerging trends, the availability of technology which failed to recognize the person with mask. So, the novel face mask detectors and recognizer can possibly contribute to public healthcare. The architecture consists of Mobile Net as the backbone mask detector, it can be used for high and low computation scenarios. In this project OpenCV, tensor flow, keras, and CNN were used to recognize the people who worn face masks. The models were tested with images. After detecting the face mask, the algorithm contributes to recognize or identify the person. The backbone of this algorithm is CNN. It plays a vital role to classify the persons belonging to different classes. It aims to generate attendance to mask worn people.

However, in the future, further enhancements are expected to be made, such as live automate this process as web application or an android application with user friendly environment.

VI. REFERENCES

- [1] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Deep learning approaches for data augmentation and classification of breast masses using ultrasound images," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 1–11, 2019.
- [2] Y. Li, W. Song, and C. Cheng, "Attendance system of face recognitionbased on raspberry pi%" *Microcontrollers Embedded Syst. Appl.*, vol. 16, no. 11, pp. 28–30, 34, 2016
- [3] M. Parkhi, Andrea Vedaldi and Andrew Zisserman: "Deep Face Recognition", In Proc. CVPR, 2015.
- [4] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556, 2014.
- [5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. "Deep-Face: Closing the gap to human-level performance in face verification", In Proc. CVPR, 2014
- [6] Y. Sun, X. Wang, and X. Tang. "Deeply learned face representations are sparse, selective, and robust", *CoRR*, abs/1412.1265, 2014
- [7] S. Bashbaghi, E. Granger, R. Sabourin, and G.-A. Bilodeau, "Robust watch-list screening using dynamic ensembles of SVMs based on multiple face representations," *Mach. Vis. Appl.*, vol. 28, nos. 1–2, pp. 219–241, Feb.2017.
- [8] W. Abdalmageed, Y. Wu, S. Rawls, S. Harel, T. Hassner, I. Masi, J. Choi, J. Lekust, J. Kim, P. Natarajan, R. Nevatia, and G. Medioni, "Face recognition using deep multi-pose representations," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Mar. 2016, pp. 1–9.
- [9] F. Mokhayeri, E. Granger, and G.-A. Bilodeau, "Domain-specific face synthesis for video face recognition from a single sample per person," *IEEE*

Trans. Inf. Forensics Security, vol. 14, no. 3, pp. 757–772, Mar. 2019

- [10] W. Deng, J. Hu, and J. Guo, "Face recognition via collaborative representation: Its discriminant nature and superposed representation," *IEEE Trans Pattern Anal. Mach. Intell.*, vol. 40, no. 10, pp. 2513–2521, Oct. 2018.



Certificate of Publication



This is to confirm that

Aseema Parveen Az, Haripriya A, Rajeshwari C, Mrs.S.Yamuna

Published following article

Construction Artisan Inspection Android Application Using Olap (On
Line Analytical Processing) Cluster Algorithm

Volume 3, Issue 6, pp: 1076-1083

www.ijaem.net

A Peer Reviewed Journal

International Journal of Advances in Engineering
and Management (IJAEM)

ISSN: 2395-5252

A handwritten signature in black ink, appearing to be a stylized 'P' followed by some illegible characters.

Publication Head

Construction Artisan Inspection Android Application Using Olap (On-Line Analytical Processing) Cluster Algorithm

Aseema Parveen Az, Haripriya A, Rajeshwari C, Mrs.S.Yamuna,

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Professor, Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Submitted: 01-06-2021

Revised: 14-06-2021

Accepted: 16-06-2021

ABSTRACT. General contractors play an important role in the success of construction projects. Determining the most appropriate contractor for a construction project is a highly critical issue. Selecting the right contractor for the right job can significantly influence the overall project performance. Traditional selection process like bidding will lead to inappropriate selections. Therefore we propose an Android application for the user to find the contractor who can satisfy the budget constraints as well as other criteria. Top ranked contractor will be recommended to the user using data mining concepts. Offloading technique is done to store the user database and contractor database in the cloud database. OLAP(On-Line Analytical Processing) Cluster Algorithm is used for recommending the contractor in the ranking manner. Feedback from the user is also collected in this android application by ranking process. So that it will be useful in the future recommendation for the user to find his most suitable contractor who can satisfy the user with budget constraints and his other criteria.

Keywords: datamining, bigdata, OnlineAnalyticalClusterAlgorithm (OLAP), user, builder.

I. INTRODUCTION

A budget constraint refers to all the combination of goods and services that can be purchased by a consumer with his or her income at their given prices. The concepts of a preference map and a budget constraint is used by the consumer theory for analyzing consumer choices. Consumer behavior is considered a maximization problem, which means that a consumer utilizes the most of his limited resources for maximizing his utility. Budget is the only thing that limits the

consumption of a consumer as the demand of consumer is insatiable and with quantity the utility function grows. An individual consumer should make the choice of consuming goods at the point in which the indifference curve which is most preferred on the preference map is tangent to the budget constraint. It means that the tangency of the indifference curve to the budget constraint shows the maximum utility that can be obtained by making use of the consumer's entire budget.

The point of tangency represents the combination of goods a consumer should purchase in order to utilize the budget fully to get maximum utility. A line that joins all tangent points between the indifference curve and budget constraint is known as expansion path.

II RELATED WORKS

A. ANDROID

Android is a mobile operating system (OS) based on the Linux kernel and currently developed by Google. With a user interface based on direct manipulation, Android is designed primarily for touchscreen mobile devices such as smartphones and tablet computers, with specialized user interfaces for televisions (Android TV), cars (Android Auto), and wrist watches (Android Wear). The OS uses touch inputs that loosely correspond to real-world actions, like swiping, tapping, pinching, and reverse pinching to manipulate on-screen objects and a virtual keyboard. Despite being primarily designed for touchscreen input, it also has been used in game consoles, digital cameras, and other electronics.

B. DATA MINING

Data mining is a process used by companies to turn raw data into useful information. By using software to look for patterns in large

batches of data, businesses can learn more about their customers to develop more effective marketing strategies, increase sales and decrease costs. Data mining depends on effective, data collection, warehousing, and computer processing.

III. LITERATURE SURVEY

Gokhan Arslan, Taylor, Francis(2012)Web-based contractor evaluation system for mass-housing projects in Turkey.In : Journal of Civil Engineering and Management.

The main objective of developing WEB-CONTEST is to facilitate the General contractors selection process in construction projects of Public Housing Development Administration. In addition, it aims to minimize the problems that may occur in traditional selection processes. The aim of this survey is to examine the relative weights of the criteria that have been using by this owner in the construction contractor selection process. Then, a web-based contractor evaluation system called WEB-CONTEST by which the contractors can be evaluated based on a combined criterion is presented.

Using this system, the construction owner can select the most appropriate General contractors for the relevant projects, speed up the selection process and gain the advantage of saving time and cost during the bidding process.

In this method, the lowest level of criteria in the hierarchy is named as attributes instead of sub-criteria. Decision-makers are asked to rank each of the attributes, assigning the first ranked attribute to a score of 100 and the others a value between 0 and 100 depicting their ranked relationships Then the performance values with relative weights for all attributes are determined and a utility value for each alternative is calculated.

M. C. B. Araújo, L. H. Alencar, C. M. M. Mota,Contractor Selection in Construction Industry: A Multicriteria Model.In : Proceedings of the 2015 IEEE IEEM

Contractor selection is a strategic question for the construction industry since the suppliers have an important role in projects performance. Two essential issues in this process are the choice of adequate criteria and methods of supplier's evaluation. Therefore, this paper applies a model for contractor selection in the construction industry. Normally, this selection is considering a group decision problem, since various departments are influenced for this choice and its managers participate of the process. Moreover, this paper focuses on the situation in which the company wants to select a set of contractors, maximizing the

use of resources according with the constraints imposed. In this context, the model utilized considers the interaction between a Group Decision and an Integer Programming methods. Afterward, it was made a numerical application of the model. The criteria used in this simulation were identified from a literature review in papers related to the supplier's selection in the construction industry.

R. M. Melol , D. D. Medeiros2 , A. T. de Almeida2,Selection and Ranking of Improvement Approaches in Construction Companies: SMARTS MethodIn : proceedings of 2011 IEEE IEEM

There are various ways, related to quality methods, tools and standards, to improve production processes. The aim of this paper is to present a multicriteria decision model for selecting and ranking these alternatives taking into account Quality, Environment and Sustainability, and Safety. The paper proposes a multicriteria decision model based in SMARTS (Simple Multi-attribute RatingTechnique) method. A study case in the context of construction companies illustrates the use of the model. The definition of the criteria and their evaluation was conducted based on interviews with experts in construction companies. One of the differentials of this model is the structure designed for a quality program planning decision making.Also, it lays on its use of an alternative set of fragmented improvements. These alternatives were combined together with some restrictions so that they became a new global set of alternatives.

The extensive research on quality management (QM) indicates that the ultimate goal of QM is to establish a management system and an organizational culture that ensures customer satisfaction and continuous improvement. Thus, the objective of this study is to build a model that facilitates the choice and sequencing of implementing improvement alternatives directed to Quality, Safety, Sustainability and Rationalization in construction companies. Thereafter, modeling was performed using a Multi criteria Decision Support approach. Two applications were conducted with two decision-makers to illustrate the use of the model. The first is the director of the construction company A and the second is a project manager for construction company C. These two applications were performed using SMARTS.

Sowndarya Sundar_ , Jaya Prakash Champatiy, and Ben Liang Multi-user Task Offloading to HeterogeneousProcessors with Communication Delayand Budget Constraints

Task scheduling and offloading in a cloud computing system with multiple users where tasks

have different processing times, release times, communication times, and weights. Each user may schedule a task locally or offload it to a shared cloud with heterogeneous processors by paying a price for the resource usage. We consider four different models in this paper:

- (i) zero task release and communication times,
- (ii) non-zero task release times and zero communication times,
- (iii) non-zero task release times and fixed communication times,
- (iv) non-zero task release times and sequence-dependent communication times.

Our work aims at identifying a task scheduling decision that minimizes the weighted sum completion time of all tasks, while satisfying the users' budget constraints. We propose an efficient solution framework for this NP-hard problem. As a first step, we use a relaxation and a rounding technique to obtain an integer solution that is a constant factor approximation to the minimum weighted sum completion time. This solution violates the budget constraints, but the average budget violation decreases as the number of users increases. Thus, we develop a scalable algorithm termed Single-Task Unload for Budget Resolution (STUBR), which resolves budget violations and orders the tasks to obtain robust solutions. We prove performance bounds for the rounded solution as well as for the budget-resolved solution, for all four models considered. Via extensive trace-driven simulation for both chess and compute-intensive applications, we observe that STUBR exhibits robust performance under practical scenarios and outperforms existing alternatives. We also use simulation to study the scalability of STUBR algorithm as the number of tasks and the number of users in the system increases.

Tong Liu¹, Anuradha Mathranil², Jasper Mbachu², Hunting the Popular Construction Apps, In : proceedings of 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering

Mobile computing offers several benefits which could help construction workers improve efficiency and productivity. However, the uptake of the technology is quite low in the industry, with little research on key constraint factors and the priority needs of the construction workforce. This study aimed to achieve three key objectives:

1) to identify the smartphone operating systems used by construction workers;

2) to identify barriers to greater uptake of mobile app technology in New Zealand Construction Industry;

3) to determine the popularity of construction apps and their key features. Through interview-based exploratory survey method, feedback received from members of the Registered Master Builders Association of New Zealand was analyzed using content analysis and descriptive statistics methods. Six factors were found to constrain the uptake of mobile apps in the industry, the most influential being cost of software and licensing. Findings in relation to the popular apps and their key features showed that PlanGrid – a cloud-based project collaboration and management system – was the most popular with a rating of 5/5. Others included JobFlex – an estimating and tendering software, Procore – used for project management, and SmartBidNet used for bid management. The findings could help app developers understand the specific needs of construction workers, and improve wider uptake of mobile apps in the industry.

PollaphatNitithamyong, Mirosław J. Skibniewski(2004), Web-based construction project management systems: how to make them Successful?, In : Automation in Construction

This paper describes research conducted at Purdue University on the identification of factors determining success or failure of web-based construction project management systems, particularly through the use of application service providers utilized by construction firms without in-house expertise to develop such systems for exclusive company use. The construction industry is fragmented due to the many stakeholders and phases involved in a construction project. This fragmentation has led to well-documented problems with communication and information processing and has contributed to the proliferation of adversarial relationships between the parties to a project. This fragmentation is also often seen as one of the major contributors to low productivity in construction. Information Technology (IT) is now routinely used in the construction industry as a tool to reduce some of the problems generated by fragmentation. The use of IT improves coordination and collaboration between firms participating in a construction project, leading to better communication practices.

Its benefits include an increase in the quality of documents and the speed of the work, better financial control and communications, and simpler and faster access to common data as well

as a decrease in documentation errors. IT spending in Architecture / Engineering / Construction (A/E/C) firms has increased significantly during the past few years, indicating that A/E/C firms are increasing their interests in IT applications to facilitate construction projects.

IV. PROPOSED SYSTEM

In the existing system, we have seen out online banking via personal computers and mobile apps on cell phones has made banking more convenient and accessible 24 hours a day. More preparation is necessary for the early stages of design. The architect is likely to be more restricted in his method of working. The designers may have some limitation in their work style. The owner has the sole power to decide what type of contract & delivery method should be used for a specific facility to be constructed and to set forth the terms in a contractual agreement. And as a principle to manage the project effectively we should first manage the contracts. While construction contracts serves as a means of pricing construction, they also structure the allocation of risk to the various parties involved. Critics say that the lack of competitive bidding by the design build process points to its greatest flaw in serving the interest of the client. By the mere fact of the design build method contracting with a single source appointment, by default eliminates the traditional competitive bidding component from the process. The nature of the design build method is a best value approach instead of a competitive bid solicitation approach. Owners face the same sacrifice decisions in the design-bid-build method, although at a greater expense when plans are completed before they discover these factors.

- It will handle with single source responsibility.
- Reduced cost and time Savings.
- Builders have maintained optimum project efficiency.
- User can know the value of builders depends upon the specification.
- Greater Owner Project Control.

V. MODULES DESCRIPTION

CHECKING INTERNET CONNECTIVITY

First, check that mobile data is turned on and you have a data connection. Open your Settings app "Wireless and Networks" or "Connections" Mobile data or Cellular data. On some devices, you may need to select "Data usage" before you see this. Turn mobile data on.

AUTHENTICATION AND AUTHORIZATION

• Authentication

Authentication is the process of determining whether someone or something is, in fact, who or what it declares itself to be. Authentication technology provides access control for systems by checking to see if a user's credentials match the credentials in a database of authorized users or in a data authentication server.

Users are usually identified with a user ID, and authentication is accomplished when the user provides a credential, for example, a password, that matches that user ID. Most users are most familiar with using a password, which, as a piece of information that should be known only to the user, is called a knowledge authentication factor.

• Authorization

Authorization is the process of giving someone permission to do or have something. In multi-user computer systems, a system administrator defines for the system which users are allowed access to the system and what privileges of use (such as access to which file directories, hours of access, amount of allocated storage space, and so forth).

Assuming that someone has logged in to a computer operating system or application, the system or application may want to identify what resources the user can be given during this session. Thus, authorization is sometimes seen as both the preliminary setting up of permissions by a system administrator and the actual checking of the permission values that have been set up when a user is getting access.

ADMIN

Admin can view all the user and user details they only have access to have seen the details. They will maintain the overall data with the most secure and immediate response will be passing to the end-users. They have to analyze the product confirmation once the user is given a request for the sale.

USER

The modules provide the user details. If the user new to the application means they want to register to this application after they can access this application easily. The user can register the details with proper validation and all the fields will be required for the registration process. Users can see the builders as per their requirements and it will be make more useful for the builders selection. They can see about builders details with the complete specification as per the user needs and they can

choose the builders and give the request for construction with the full details about its specification. Users can edit the profile with proper validation and they can update the profile then email id and name will be constant. They can see the order list, order status for their reference after that they will give feedback about the product as well as application.

BUILDERS

The modules provide the builder details. If the user new to the application means they want to register to this application after they can access this application easily. The user can register the details with proper validation and all the fields will be required for the registration process. Builders can add the design and their specification of budgets constraints for the user reference. Builders can edit the profile with proper validation and they can update the profile then email id and name will be constant. They can see the request list, construction status for their reference after that they will give status about the construction to the respective users for their reference.

BUDGET CONSTRAINTS

These modules provide information about the Budget constraints details for the individually users. Even before you enter the planning phase, address areas where your project could face changes at any point in the process. Since many project managers out there attribute their going over-budget to changes that occur along the way, hash out some potential scenarios first. The user can see the budget according to their user specification it get compare with all the builders.

RANKING RANGE

These modules provide the builders ranking information with the full specification of the projects. The builders ranking will be evaluated with rate, place, experience, specifications and projects. So the user can easily preferred the

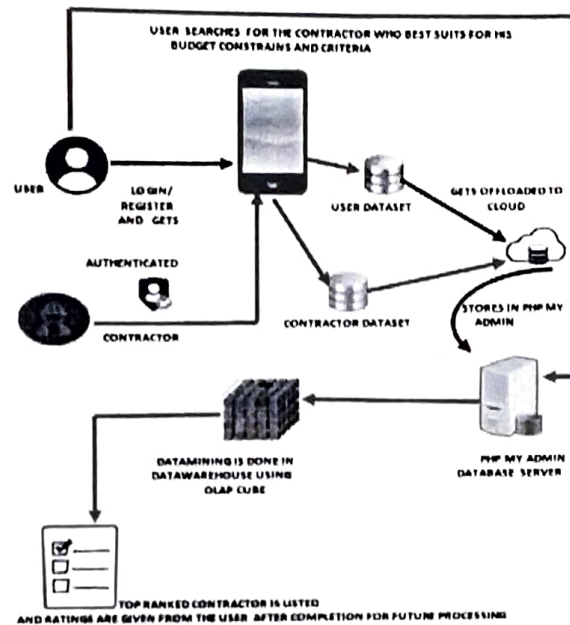
builders with the following details and it will be shown in the first its makes decision for user easily and it will be get clearance about builders value.

STORAGE

This module gives the information about the storage for the entire application. It will be get secured and maintained by the admin. Because nowadays in the current world data loss is happened frequently through the location or email id. So we secured the data with high credential and it makes more security levels for the user details. The data storage will have happened with split up and no chance for the data traffic also it will get a response immediately.

VI. SYSTEM ARCHITECTURE

This model of android application is used to find the most suitable contractor who will be able to satisfy the budget constraints of the user. In this Architecture diagram we understand that the user and the contractor will have separate login credentials, and they can also register accordingly. When the user and contractor registers the dataset of the user which consists of the budget constraints and the details of the user are gets offloaded to cloud database. Similarly, the details of the constructors are also gets offloaded to the cloud database. These databases are stored in phpmyadmin server and the admin also has the separate login to view the status of the application uses. When the user request and search for the builders who can satisfy the budget constraints the request will send to the phpmyadmin server. Then datamining is done in the datawarehouse using OLAP (Online Analytical Processing) Cluster Algorithms the top ranked and low budget builders are listed accordingly. After completion of the project, the user can rank the contractor for future processing.



VII. SYSTEM IMPLEMENTATION

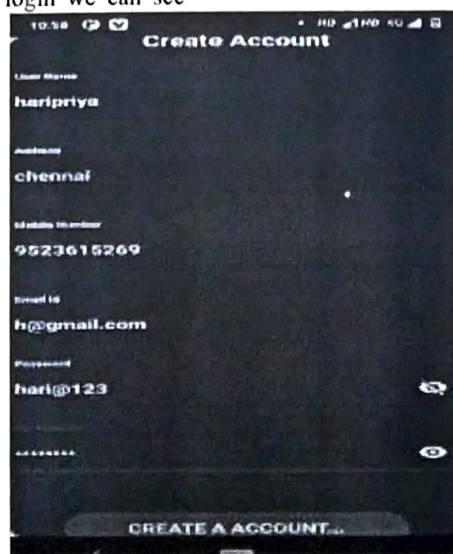
In the system construction we used java as front end , backend as mysql and servers as php .It has 3 module that is user,builder and admin .In user first he/she want to register themselves by creating new account after that login can be done . In user login we can see the builders details and can edit the details and also see the status of builders

.In builders login he can login and add the design architecture in how many day he would complete the project .In admin login we can see

both the user login and builder login all these information are stored in the database .It will retrieve and give the information when it is required.

A CREATE USER ACCOUNT

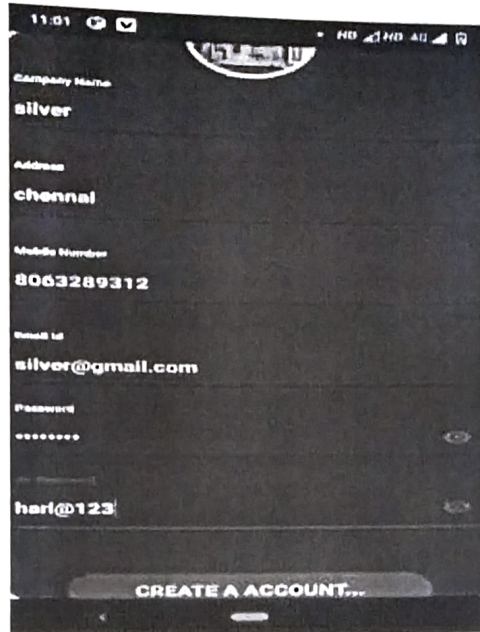
In user first he/she want to register themselves by creating new account after that login can be done . In user login we can see the builders details and can edit the details and also see the status of builders.



B .CREATE BUILDER ACCOUNT

In builders login he can login and add the design architecture in how many day he would complete the project with the help of how many

workers , how many days ,how much sq ft and also he can Update the status to the user about the completion work.



USER SEARCH PAGE:



VIII. CONCLUSION AND FUTURE WORK

In this application, we have multi-user computational offloading problem, for a system consisting of a finite-capacity cloud with heterogeneous processors. The offloaded tasks incur monetary cost for the cloud resource usage and each user has a budget constraint. We have formulated the problem of weighted-sum-completion-time minimization subject to the user budget constraints. We have formulated a problem to minimize the weighted sum completion time subject to the user budget constraints. The proposed algorithm relaxes, rounds, and resolves budget violations, and it sorts the tasks to obtain an effective solution. We have also obtained interesting performance bounds for both the underlying rounded solution and the budget-resolved solution for different release-time and communication-time models. Through simulation using realworld application traces, we have observed by scalable and substantially outperform

the existing alternatives especially for larger systems. A possible future research direction is to account for explicit task dependencies in the formulation. Additionally, the consideration of multiple types of computing resource in task scheduling is a challenging but important problem and additionally, we may consider different pricing and budget schemes as an extension to the linear scheme considered in this application with the better user interface and feels free to users.

REFERENCES

- [1]. Gokhan Arslan, Taylor, Francis(2012) Web-based contractor evaluation system for mass-housing projects in Turkey. In : Journal of Civil Engineering and Management
- [2]. M. C. B. Araújo, L. H. Alencar, C. M. M. Mota, Contractor Selection in Construction Industry: A Multicriteria Model. In : Proceedings of the 2015 IEEE IEEM.

- [3]. R. M. Meloi¹, D. D. Medeiros², A. T. de Almeida², Selection and Ranking of Improvement Approaches in Construction Companies: SMARTS Method, In : proceedings of 2011 IEEE IEEM.
- [4]. Sowndarya Sundar¹, Jaya Prakash Champatiy, and Ben Liang Multi-user Task Offloading to Heterogeneous Processors with Communication Delay and Budget Constraints, In : Proceedings of the 2019 IEEE IEEM.
- [5]. Tong Liu¹, Anuradha Mathrani¹, Jasper Mbachu², Hunting the Popular Construction Apps , In : proceedings of 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering
- [6]. PollaphatNitithamyong, Mirosław J. Skibniewski (2004), Web-based construction project management systems: how to make them Successful?, In : Automation in Construction



EASWARI
ENGINEERING COLLEGE
An Autonomous Institution
Approved by Anna University
RAMAPURAM CHENNAI



DEPARTMENT OF INFORMATION TECHNOLOGY

Paper id 366

Certificate of Participation

This is to certify that **Dr. /Mr. /Ms.** **V Sundari** of
Meenakshi Sundarajan Engineering College (Affiliated to Anna University) Chennai, India has participated in
2nd International e-Conference on Information, Communication and Networking (eICICN - 2021) organized by the
Department of Information Technology, Easwari Engineering College, held on 9th & 10th April 2021 and presented a paper titled
Segregation of Images of Waste Using Machine Learning

V. Balaji
 Coordinator
 Associate Professor

Dr. S. Kanaga Suba Raja
 Co-Convenor
 Associate Professor

Dr. N. Anandhi
 Convener
 Head / Information Technology



Segregation of Images of Waste Using Machine Learning

V SUDHIR

*Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
nacksadhira@gmail.com*

R BHARATH

*Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India
bharathroshanr@gmail.com*

S MOHAMED BASHEER

*Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India
mohamedbasheer10027@gmail.com*

Mrs. V SUNDARI

*Assistant Professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College (Affiliated to
Anna University)
Chennai, India
sundarikkt@gmail.com*

Abstract— Contamination is a huge problem in the recycling industry that can be mitigated with automated waste sorting and classification. The objective of this project is to capture images of a single waste material and effectively identify and segregate it into five classes viz. Metal, glass, paper, cardboard and plastic. The model used for the project is Convolutional Neural Network (CNN), a Machine Learning algorithm. This system will ensure effective automated waste management and will speed up the process of segregation without any human intervention. Image segmentation and classification is more and more being of interest for computer vision and machine learning researchers. Many systems on the rise need accurate and efficient segmentation and recognition mechanisms. This demand coincides with the increase of computational capabilities of modern computer architectures and more effective algorithms for image recognition. The use of Convolutional Neural Networks (CNN) for the image classification and recognition allows building systems that enable automation in many industries. This project presents a system for classifying Biodegradable and Non-Biodegradable waste, using CNN. The problem of segregation of renewable waste is a big challenge for many countries around the world. Apart from segregating waste using human hands, there are several methods for automatic segregation. The project proposes a system for classifying waste with the following classes: Metal, glass, paper, cardboard and plastic. The obtained results show that automatic waste classification, using image processing and artificial intelligence methods, allows building effective systems that operate in the real world.

Keywords—CNN, VGG16, Waste Segregation

I. INTRODUCTION

Sorting of materials is part of the municipal waste management necessary in the recycling process. The manual sorting process is tedious and expensive, which is why you need to create automated sorting techniques to improve recycling efficiency. An important element of the whole process is the preliminary division of waste into various groups of materials, followed by detailed segregation. The disposal of plastic waste is the most important problem in environmental protection. The project proposes a method for the automatic segregation of plastic as well as other Non-Biodegradable and Biodegradable wastes using single image input (**Convolutional Neural Networks**).

The content of this paper is divided into five sections. In the first section, the background is introduced along with the motivation and purpose of this paper. In the second section, the technology and documents that are related to this paper are mentioned. In the third section, the structure of the system built in this paper is explained. In the fourth section, the construction of the system is presented in detail. In the last section, conclusions are made along with future enhancement possibilities.

II. RELATED WORKS

A. VGG16:

VGG16 is a convolution neural net (CNN) architecture which was used to win ILSVR (Imagenet) competition in 2014. It is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG16 is that instead of having a large number of hyper-parameters they focused on having convolution layers of 3x3 filter with a stride 1 and always used same padding and maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC (fully connected layers) followed by a softmax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million (approx.) parameters.

B. CONVOLUTIONAL NEURAL NETWORK

A **Convolutional Neural Network (CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, CNNs have the ability to learn these filters/characteristics.

The architecture of a CNN is analogous to that of the connectivity pattern of Neurons in the Human Brain and was

inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area.

A CNN is able to successfully capture the Spatial and Temporal dependencies in an image through the application of relevant filters. The architecture performs a better fitting to the image dataset due to the reduction in the number of parameters involved and reusability of weights. In other words, the network can be trained to understand the sophistication of the image better.

The structure of CNNs consisted of three main substructures, which include: convolutional layers, pooling layers, fully connected layers.

Convolutional layers are made from several feature maps, and each unit of feature maps is made from convolving a small region in input data which is called the local receptive field. A new feature map is created by sliding a local receptive field over the input. The convolution can be used in various kinds of data such as image, text. For example, in the image, an area of pixels is convolved, and in the text, a group of characters or words are convolved. Unlike the standard neural network, each neuron in the layers is not connected to all of the nodes (neurons) in the previous layer but is just connected to nodes in a special region known as the local receptive field.

Pooling layers are commonly used immediately after convolutional layers. These layers were generated to simplify the information and reduce the scale of feature maps. In other words, pooling layers make a condensed feature map from each feature map in convolutional layers. In some references, these layers are called the subsampling layer. Pooling operation can be performed in various types such as geometric average, harmonic average, maximum pooling. Max-pooling and average-pooling are two of the most prevalent processes for pooling. The pooling layers are necessary to reduce the computational time and overfitting issues in the CNN.

Fully connected layers are the final layers in the CNN structure that can be one or more layers and placed after a sequence of convolution and pooling layers. This part of CNN comprises the composite and aggregates of the most important information from all procedures of CNN. Consequently, these layers provide the feature vector for the input data, which can be used for some machine learning tasks such as classification, prediction. The last layer of fully connected layers is known as softmax classifier and determines the probability of each class label over N classes.

C. EXISTING SYSTEM:

The prevailing garbage disposal system in India consists of unorganized waste collected from habitats which are then segregated at a station. The segregation is done by manual labour.

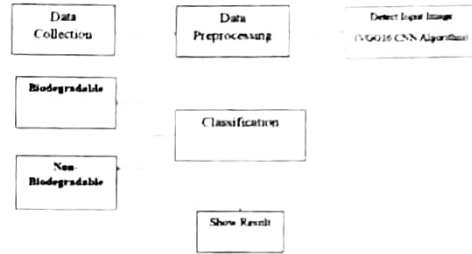
The segregation is usually done where all waste is dumped together (**Landfills**). Since it is done by manual labour, it opens up the possibility of the workers being exposed to hazardous health conditions and hence affects their overall health. It is also noted that the cost of manual labour is significantly higher as more workers are generally needed to manage a large amount of waste. To prevent this, we go for a programmed approach.

D. PROBLEM STATEMENT:

To maintain a database of images of waste products

and segregate them into biodegradable and non-biodegradable wastes. Dataset must be trained by applying CNN algorithm. Input image must be detected. With the CNN algorithm, image must be predicted and classified as biodegradable or non-biodegradable. To ensure a speedy automated segregation process.

III. SYSTEM ARCHITECTURE



A system architecture is the conceptual model that defines the structure, behaviour, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviours of the system. The dataset of the waste images is collected. The relevant data is then processed. After processing the data, an Input image is given to perform a test using the VGG16 CNN Algorithm. Once input is given, the image is checked with the trained database and is classified into **Biodegradable or Non-Biodegradable** along with certain specifications such as (cardboard, paper and organic or plastic, glass and metal). The output result shows the material type and the category (Biodegradable or Non-Biodegradable) of the given image

IV. PROPOSED SYSTEM

In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labelled data and neural network architectures that contain many layers.

One of the most popular types of deep neural networks is known as convolutional neural networks (CNN or ConvNet). A CNN convolves learned features with input data, and uses 2D convolutional layers, making this architecture well suited to processing 2D data, such as images.

CNNs eliminate the need for manual feature extraction, so you do not need to identify features used to classify images. The CNN works by extracting features directly from images. The relevant features are not retrained; they are learned while the network trains on a collection of images. This automated feature extraction makes deep learning models highly accurate for computer vision tasks such as object classification.

Maintain a database of images of waste products and segregate them into biodegradable and non-biodegradable wastes. In our system we train the dataset by applying VGG16 CNN algorithm with 3 dense layers and 2 dropouts. Input image will be detected. With the CNN algorithm, image will be predicted and classified as biodegradable or non-biodegradable. This system will ensure effective automated

waste management and will speed up the process of segregation without any human intervention. Advantages are easy segregation, consumes less time and ensures the sanitation of labourers.

V. SYSTEM IMPLEMENTATION

Systems implementation refers to the working of the application. It is better understood when bisected into modules. In this project, the application has 3 major modules,

Import Image from Dataset

The first step is to import the data set using keras pre-processing image data generator function also, create size, rescale, range, zoom range, horizontal flip. Then import the image dataset from folder through the data generator function. Here, set train, test, and validation also, set target size, batch size and class-mode from this function.

Train Dataset

To train the dataset using classifier and fit generator function also make training steps per epoch's then total number of epochs, validation data and validation steps using this data train the dataset.

Predicting and classifying the images

In this module the trained deep learning model is converted into hierarchical data format file (.h5 file) which is then deployed in Google Colab for providing better user interface and predicting the output whether the given input image is Biodegradable or Non-Biodegradable.

VI. RESULTS

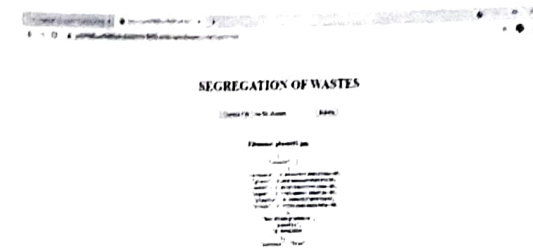
A. HOME PAGE:

This is the homepage of our website which contains the title as segregation of wastes and two buttons such as choose file and submit buttons.



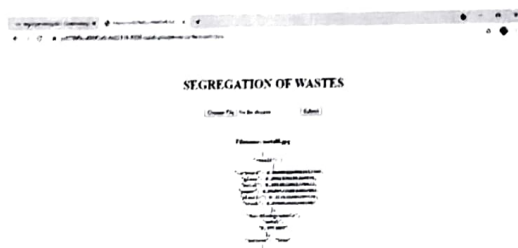
B. PLASTIC WASTE:

This page accepts image file and shows the type of waste whether it is biodegradable or non-biodegradable, the type of item, probability of that type and stage it as success or failure. In this case, it will show type of waste as non-biodegradable and type of item as plastic.



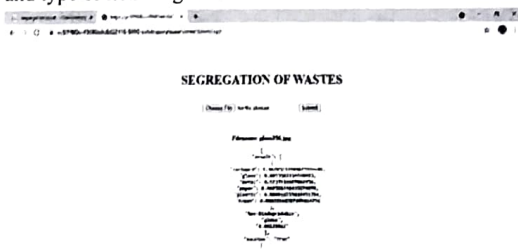
C. METAL WASTE:

This page accepts image file and shows the type of waste whether it is biodegradable or non-biodegradable, the type of item, probability of that type and stage it as success or failure. In this case, it will show type of waste as non-biodegradable and type of item as metal.



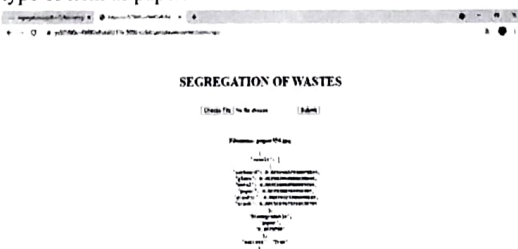
D. GLASS WASTE:

This page accepts image file and shows the type of waste whether it is biodegradable or non-biodegradable, the type of item, probability of that type and stage it as success or failure. In this case, it will show type of waste as non-biodegradable and type of item as glass.



E. PAPER WASTE

This page accepts image file and shows the type of waste whether it is biodegradable or non-biodegradable, the type of item, probability of that type and stage it as success or failure. In this case, it will show type of waste as biodegradable and type of item as paper.



F. CARDBOARD WASTE

This page accepts image file and shows the type of waste whether it is biodegradable or non-biodegradable, the type of item, probability of that type and stage it as success or failure. In this case, it will show type of waste as biodegradable and type of item as cardboard.

SEGREGATION OF WASTES



VII. CONCLUSION AND FUTURE WORK

Keeping our environment clean and eco-friendly is our main priority. The proposal for the management of wastes is efficient and time saving than the currently employed method where the municipality employees perform. Though this system is simple in concept, it is very valuable and affordable. Hence to ensure being automated, a system which takes lots of dataset as input without human intervention and also has the capacity to think by itself offers the best solution. Further advancements can be made in the system to increase its accuracy and effectiveness. More improvements can be made in differentiating plastic and glass since both of their images look similar. The input image dataset can be changed to a more application specific execution of this system. For instance, this system can be used in fast food chain restaurants where it would be able to classify the waste materials like soda cans, tissue papers, plastic cups, wrappers, etc. This prototype can be developed with Robots or other machineries for easy and safe collection and segregation of wastes.

REFERENCES

- [1] João Sousa, Ana Rebelo, Jaime S. Cardoso, "Automation of Waste Sorting with Deep Learning", 2019 XV Workshop de Visão Computacional (WVC), IEEE, 9-11 Sept. 2019.
- [2] S. Sudha, M. Vidhyalakshmi, K. Pavithra, K. Sangeetha and V. Swaathi "An automatic classification method for environment: Friendly waste segregation using deep learning" in Technological Innovations in ICT for Agriculture and Rural Development (TIAR), 2016 IEEE, July 15-16 2016.
- [3] Dymitr Pietrow, MSc Military University of Technology, Jan Matuszewski, PhD Military University of Technology. "Objects detection and recognition system using artificial neural networks and drones".
- [4] George E Sakr, Maria Mokbel, Ahmad Darwich, Mia Nasr Khneisser and Ali Hadi, "Comparing Deep Learning and Support Vector Machines for Autonomous Waste Sorting" in Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International, November 2 - 4, 2016.
- [5] Yushi Chen, Hanlu Jiang, ChunyangLi, Xiuping Jia and Pedram Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks", IEEE Transactions on Geoscience and Remote Sensing, Vol. 54, No. 10, October 2016
- [6] Mingyuan Xin and Yong Wang, "Research on image classification modelbased on deep convolution neural network", EURASIP Journal on Imageand Video Processing, 2019

[7] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks", Neural Information Processing Systems 2016

[8] M.K. Pushpa, Aayushi Gupta, Shariqh Mohammed Shaikh, Stuti Jha, Suchitra V, "Microcontroller Based Automatic Waste Segregator". IJIREICE 2015

COVID-19 PNEUMONIA CLASSIFICATION USING DEEP LEARNING TECHNIQUE

Gayathri B^[1] , Shanmugappriya S^[2] , Vaishali R^[3] , Aarthi S^[4]
[1],[2],[3]-Students [4]-Assistant Professor
Meenakshi Sundararajan Engineering College

Abstract-- Segmentation of pneumonia lesions from CT scans or x-rays of COVID-19 patients is important for accurate diagnosis and follow-up. Deep learning has a potential to automate this task but requires a large set of high-quality annotations that are difficult to collect. The most common symptoms of COVID-19 patients include fever, cough and shortness of breath , and the patients typically suffer from pneumonia. automatic segmentation of COVID-19 pneumonia lesions from CT scans is challenging due to several reasons. First, the infection lesions have a variety of complex appearances such as Ground-Glass Opacity (GGO), reticulation, consolidation and others. Second, the sizes and positions of the pneumonia lesions vary largely at different stages of the infection and among different patients. In addition, the lesions have irregular shapes and ambiguous boundaries, and some lesion patterns such as GGO have a low contrast with surrounding regions.

Index Terms—AlexNet, LeNet, CNN, x-ray

I. INTRODUCTION

The presence of symptoms of COVID-19 pneumonia may be similar to other types of viral pneumonia. Because of this, it can be difficult to tell what’s causing one’s condition without being tested for COVID-19 or other respiratory infections. This is to determine how COVID-19 pneumonia differs from other types of pneumonia. Information from these studies can potentially help in diagnosis and in furthering our understanding of how SARS-CoV-2 affects the lungs. We present a Convolutional Neural Network in TensorFlow and Keras based Covid-19 pneumonia classification. The proposed system is based on CNN using Pneumonia images to classify the Covid-19, normal, pneumonia x-rays. It is predicted that the success of the obtained results will increase if the CNN method is supported by adding extra feature extraction methods and classify successfully covid-19 & pneumonia

II. DESIGN OF THE PROPOSED SYSTEM

The coronavirus disease (COVID-19) is rapidly spreading all over the world, and has infected more than 1,436,000 people in more than 200 countries and territories as of April 9, 2020. Detecting COVID-19 at early stage is essential to deliver proper healthcare to the patients and also to protect the uninfected population.

Also the presence of symptoms of COVID-19 pneumonia may be similar to other types of viral pneumonia. We present a Convolutional Neural Network in TensorFlow and Keras based Covid-19 pneumonia classification.

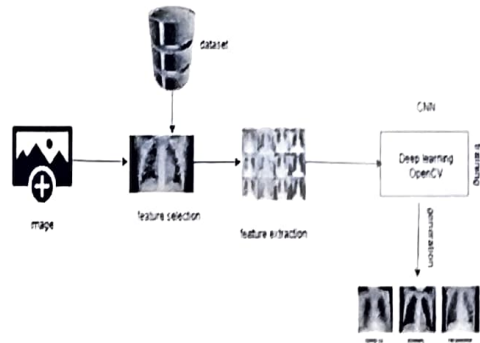


Fig 1. System architecture of the proposed system

TRAINING ALGORITHMS

i. Manual CNN

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area.

ii. AlexNet

AlexNet is the name of a convolutional neural network (CNN). AlexNet contained eight layers; the first five were convolutional layers, some of them followed by max-pooling layers, and the last three were fully connected layers. It used the non-saturating ReLU activation function, which showed improved training performance over tanh and sigmoid. AlexNet allows for multi-GPU training by putting half of the model's neurons on one GPU and the other half on another GPU. Not only does this mean that a bigger model can be trained, but it also cuts down on the training time.



iii Lenet

LeNet is a convolutional neural network structure proposed by Yann LeCun et al. in 1989. In general, LeNet refers to LeNet-5 and is a simple convolutional neural network. The LeNet architecture is *straightforward and small*, (in terms of memory footprint), making it *perfect for teaching the basics of CNNs* — it can even run on the CPU (if your system does not have a suitable GPU), making it a great “first CNN”.

iv Django framework

Django is a Python-based free and open-source web framework that follows the model-template-views architectural pattern. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It is free and open source, has a thriving and active community, great documentation, and many options for free and paid-for support. Django is a collection of **Python** libs allowing you to quickly and efficiently create a quality Web application, and is suitable for both frontend and backend.

III TRAINING

I. Import the images from the dataset

import data set using keras preprocessing image data generator function also, create size, rescale, range, zoom range, horizontal flip. Then import image dataset from folder through the data generator function. Now set train, test, and validation also we set target size, batch size and class-mode from this function, train using created network by adding layers of CNN.

Training the normal x-rays

Trained data for the name: normal

```

--- Images of: data/train/NORMAL
Images count: 471
min_width: 193
max_width: 254
min_height: 187
max_height: 254

```



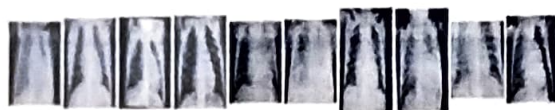
Training pneumonia x-rays

Trained data for pneumonia

```

--- Images of: data/train/PNEUMONIA
Images count: 17
min_width: 193
max_width: 254
min_height: 187
max_height: 254

```



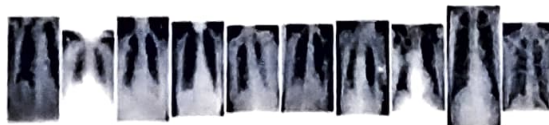
Training covid-19 x-rays

Trained data for COVID type disease

```

--- Images of: data/train/COVID
Images count: 488
min_width: 224
max_width: 477
min_height: 224
max_height: 477

```



II. Train the data using alexnet

To train the dataset using classifier and fit generator function also make training steps per epoch then total number of epochs, validation data and validation steps using this data train the dataset.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 75, 75, 32)	896
max_pooling2d (MaxPooling2D)	(None, 37, 37, 32)	0
conv2d_1 (Conv2D)	(None, 12, 12, 128)	36992
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 128)	0
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 256)	1179904
dense_1 (Dense)	(None, 3)	771

Total params: 1,218,563		
Trainable params: 1,218,563		
Non-trainable params: 0		

III. Train the data using lenet

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms.

While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

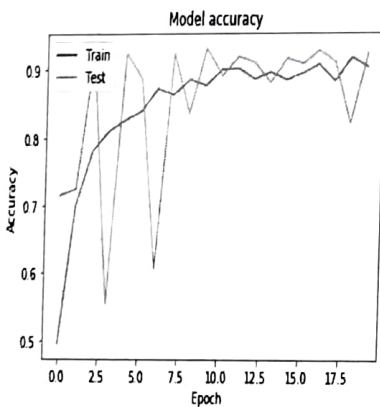
The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. Their network consists of four layers with 1,024 input units, 256 units in the first hidden layer, eight units in the second hidden layer, and two output units

Input Layer:

Input layer in CNN contain image data. Image data is represented by three dimensional matrixes. It needs to reshape it into a single column. Suppose you have image of dimension $28 \times 28 = 784$, it need to convert it into 784×1 before feeding into input.

Convo Layer:

Convo layer is sometimes called feature extractor layer because features of the image are get extracted within this layer. First of all, a part of image is connected to Convo layer to perform convolution operation as we saw earlier and calculating the dot product between receptive fields (it is a local region of the input image that has the same size as that of filter) and the filter. Result of the operation is single integer of the output volume. Then the filter over the next receptive field of the same input image by a Stride and do the same operation again. It will repeat the same process again and again until it goes through the whole image. The output will be the input for the next layer.



A. Pooling Layer:

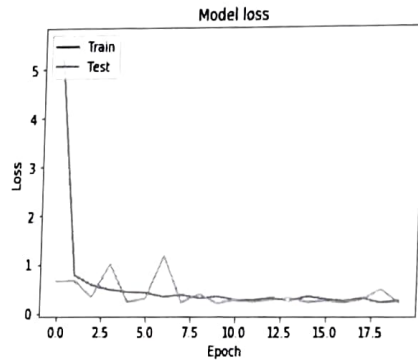
Pooling layer is used to reduce the spatial volume of input image after convolution. It is used between two convolution layers. If it applies FC after Convo layer without applying pooling or max pooling, then it will be computationally expensive. So, the max pooling is only way to reduce the spatial volume of input image. It has applied max pooling in single depth slice with Stride of 2. It can observe the 4×4 -dimension input is reducing to 2×2 dimensions.

B. Fully Connected Layer (FC):

Fully Connected Layer is simply, feed forward neural networks. **Fully Connected Layers** form the last few layers in the network. The input to the **fully connected layer** is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the **fully connected layer**. Fully connected layer involves weights, biases, and neurons. It connects neurons in one layer to neurons in another layer. It is used to classify images between different categories by training.

C. Softmax / Logistic Layer

Softmax or Logistic layer is the last layer of CNN. It resides at the end of FC layer. Logistic is used for binary classification and softmax is for multi-classification. The **softmax function** is a **function** that turns a vector of K real values into a vector of K real values that sum to 1.



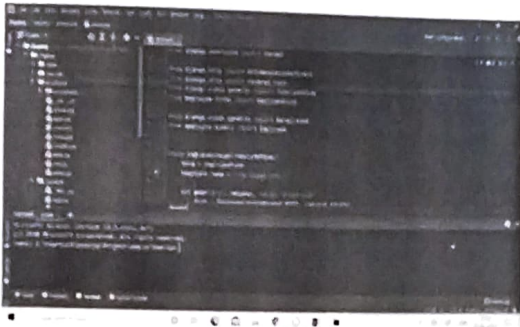
D. Output Layer

Output layer contains the label which is in the form of one-hot encoded. Now you have a good understanding of CNN.

IV. Deploy in Django framework

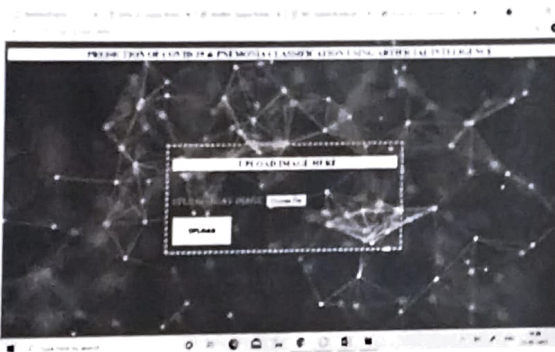
Django is a high-level Python Web Development framework that encourages rapid development and clean, pragmatic design. It has been built by experienced developers, and takes care of much of the hassle of Web development. It is also free and open source. Django REST Framework is a powerful and flexible toolkit for building Web APIs which can be used to Machine Learning model deployment. With the help of Django REST framework, complex machine learning models can be easily used just by calling an API endpoint.

In this module the trained deep learning model is converted into hierarchical data format file (.h5 file) which is then deployed in our django framework for providing better user interface and predicting the output whether the given chest X-ray is covid-19 pneumonia / normal.



IV RESULTS

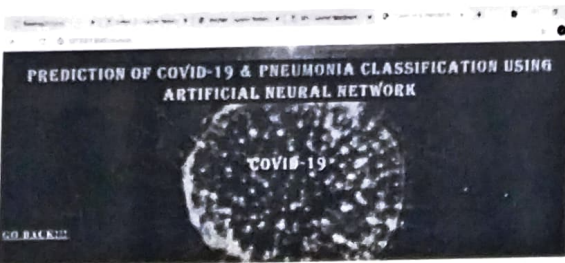
Primary screen



Uploading covid x-ray



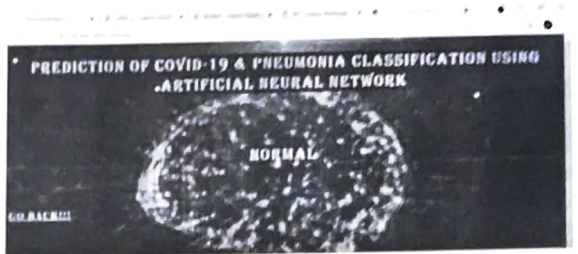
The model classifies the given input x-ray as covid 19



Uploading a normal x-ray



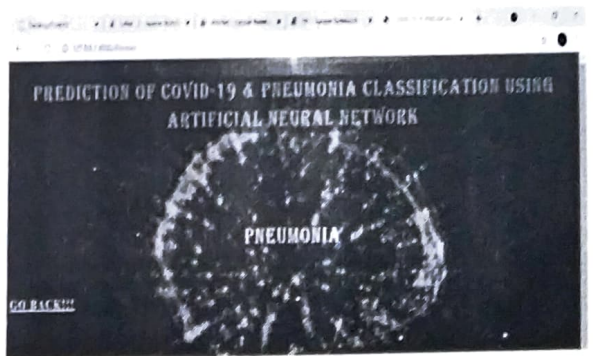
The model correctly classifies the input x-ray as normal x-ray



Uploading pneumonia x-ray(adenovirus)



Again, the model correctly classifies the input x-ray as pneumonia



It is known that covid 19 virus causes pneumonia which is similar to other types of viral pneumonia. Thus, it is very important to distinguish these accurately for proper treatment and follow up. Here we have proposed a model using Convolutional Neural Network, training the model and then testing the outputs based on the chest x-rays.

REFERENCES

- [1] Gengbin Ling, Congcong Cao (2019), "Automatic Detection and Diagnosis of Severe Viral Pneumonia CT Images Based on LDA SVM", IEEE Sensors Journal
- [2] Guotai Wang, Yinghong Liu, Chaoping Li, Zhiyong Xu, Jiegan Ruan, Haifeng Zhu, Tao Meng, Kang Li, Ning Huang, and Shaoting Zhang (2020), "A Noise-robust Framework for Automatic Segmentation of COVID-19 Pneumonia Lesions from CT Images", IEEE Transactions on Medical Imaging
- [3] Ling Jhuo, Mi-Yuen Hsieh, Ting-Chen Weng, Mei-Juan Chen and Chieh-Ming Yang (2019), "Trend Prediction of Influenza and the Associated Pneumonia in Taiwan Using Machine Learning-Song", International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)
- [4] Muhammad E. H. Chowdhury, Tawafur Rahman, Amith Khantakur (2020), "Can AI help in screening Viral and COVID-19 pneumonia?", IEEE Access
- [5] Qiyu Wang, Dun Yang, Zhihuan Li, Xiaohong Zhang, and Chen Liu (2020), "Deep Regression via Multi-Channel Multi-Modal Learning for Pneumonia Screening", IEEE ON IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY SECTION
- [6] Xi Ouyang, Juyu Hao, Liming Xia (2020), "Dual-Sampling Attention Network for Diagnosis of COVID-19 from Community Acquired Pneumonia", IEEE TRANSACTIONS ON MEDICAL IMAGING
- [7] Zakaria neeli, Mohamed Fezari, Redjati Abdeghani (2018), "Analysis of Acoustic Parameters from Respiratory Signal in COPD and Pneumonia patients", IEEE Access

Facial Emotion Recognition System Using Deep Learning

A WASIMAKRAM
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
wasimakram2720@gmail.com

M HANISHKAR
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
Hanishharish46@gmail.com

M SUGANTHAPRIYAN
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
msugan99@gmail.com

Mrs. M.R.NITHYA
Assistant Professor, Dept. of Computer Science and
Engineering Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
nithyamr06@gmail.com

Abstract—In our day to day life emotions or facial expression are the prime factor which are required for communication purpose. For humans it is quite easy to detect an emotion but difficult for a computer or a machine to do so. This project deals with 3 emotions they are Neutral, Happy and Sad. As we are developing the need and importance of automatic emotion recognition has increased which supports Human Computer Interaction applications. Facial expression defines the emotions of an individual which is required for Human Computer Interaction (HCI) in this project. In this work, user's emotion using its facial expressions will be detected. These expressions can be derived from the live feed via system's camera or any pre-existing image available in the memory. Emotions possessed by humans can be recognized and has a vast scope of study in the computer vision industry upon which several researches have already been done. The scanned image (testing dataset) is being compared to training dataset and thus emotion is predicted. The objective of this paper is to develop a system which can analyze the image and predict the expression of the person. The study proves that this procedure is workable and produces valid results.

Keywords—dataset, opencv, deep learning

I. INTRODUCTION

The project aims to integrate the concepts of convolutional neural networks (CNN), tensorflow, keras to recognize the facial emotions in humans. In this work, user's emotion using its facial expressions will be detected. These expressions can be derived from the live feed via system's camera or any pre-existing image available in the memory. Emotions possessed by humans can be recognized and has a vast scope

of study in the computer vision industry upon which several researches have already been done. The work has been implemented using Python (2.7), Open Source Computer Vision Library (OpenCV) and NumPy. The scanned image (testing dataset) is being compared to training dataset and thus emotion is predicted. The objective of this paper is to develop a system which can analyze the image and predict the expression of the person. The study proves that this procedure is workable and produces valid results.

II. RELATEDWORKS

A. DEEP LEARNING:

Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost. It is the key to voice control in consumer devices like phones, tablets, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It's achieving results that were not possible before. In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers.

B. CONVOLUTIONAL NEURAL NETWORKS:

A Convolutional Neural Network (CNN) is a Deep

quality of the processed images. It is an input image, assign weights, maintain weights and biases to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a CNN is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, CNN have the ability to learn these filters/characteristics.

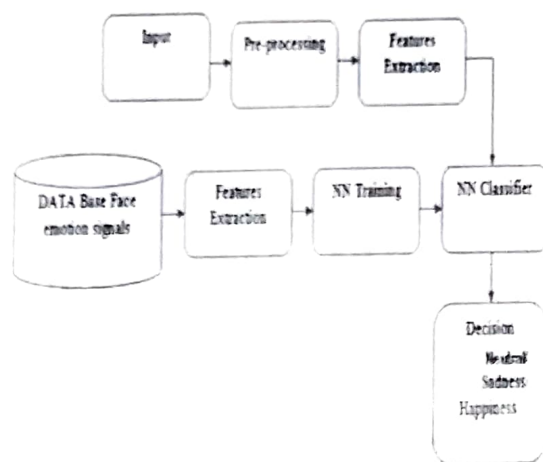
C. EXISTING SYSTEM:

The system detects and recognizes the human using image processing techniques. But the face recognitions provide only the user's name and details. But it fails to provide the actions performed by the user. Pose variation, illumination conditions, bad lighting etc. are still challenging factors faced by all algorithms. The existing system does not recognize the real time facial emotions which can be used in various fields for feedback purposes.

D. PROBLEM STATEMENT:

In present day technology human-machine interaction is growing in demand and machine needs to understand human gestures and emotions. Emotions and expressions are the natural physiological response of the human body which can be recognized by the facial expression. In our day to day life emotions or facial expression are the prime factors which are required for communication purpose. For humans it is quite easy to detect an emotion but difficult for a computer or a machine to do so. This can be done using the proposed system.

III. SYSTEM ARCHITECTURE



The human faces are collected as a dataset through camera

and converted to grayscale images, which is further used for feature extraction and stored. After which with the help of datasets, each emotions are trained. Then using camera the human images are fed as input to Pycharm and the emotion is recognized and displayed in the window.

IV. PROPOSED SYSTEM

The proposed system comprises of a high tech camera as input to detect various face expressions. The output of camera is fed into OpenCV processor for further dataset creation and processing. Based on a person's different face expressions, his/her emotions can be diagnosed or evaluated. The camera used here is of high quality, because every human emotion are reflected with minute changes in facial like changes in eyebrow shape, eyeball rotation, movement in cheek and chin. Then the dataset is trained using Convolutional Neural Network. In this system, both emotion analysis detection are done by using tensor flow algorithm which is more suited for face detection and recognition. After training the dataset the system prompts for camera access to get real-time facial expression and outputs the corresponding emotion and emoji in the window.

V. SYSTEM IMPLEMENTATION

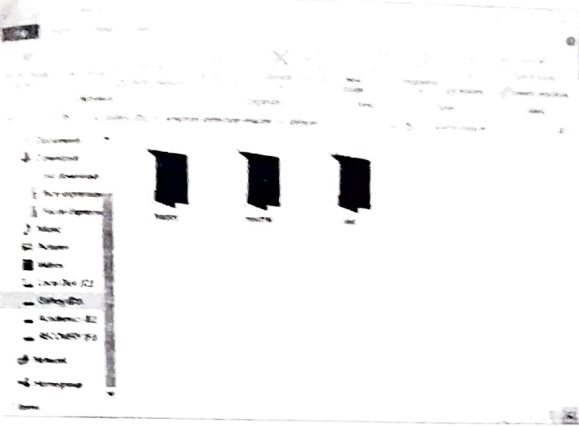
The system here is divided into three modules as listed below:

- i. Creating dataset
- ii. Learning using datasets
- iii. Detection of emotion

The system provides assistance to the admin by using the learning dataset. The dataset which contains the images of labelled emotions is fed and training using datasets is done. The accuracy is calculated based on the dataset provided. The system gets the web camera input and displays the predicted emotion on the output window spontaneously.

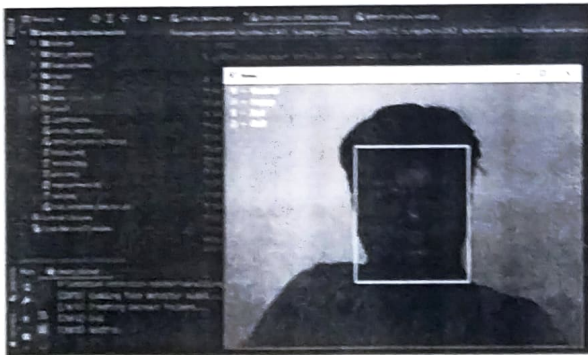
i. CREATING DATASET:

The admin creates a directory named 'dataset' to store the images of labelled emotions. Inside a unique folder is created for each emotion like Sad, Happy and Neutral. With the help of web camera the user adds the specific images in the corresponding folder based on the user inputs. For accessing the camera and performing functions like converting it into Grayscale image, cropping, displaying the cv2 python package is used. Thus the dataset is created. The user should provide ample amount of images labelled to get trained with accuracy.



ii. LEARNING USING DATASET:

The dataset is collected that consists of various significant parameters like eyebrow shape, eyeball rotation, movement in cheek and chin which are required for the decision and uses various techniques to quantize them. The learning of dataset is done using CNN with the help of libraries like tensorflow, keras, scikitlearn, scipy, numpy, matlab. A **Convolutional Neural Network (CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a CNN is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, CNN have the ability to learn these filters/characteristics. The based on the training from the dataset, the details are stored in a file.



iii. DETECTION OF EMOTIONS

The system prompts the user for web camera access and once the user allows it, the system capture the input of realtime and shows in the output window along with the correct predicted emotion. Based on the dataset learned, the system shows the detected emotion along with a emoji for every instance of the video's input. The accuracy depends on the distinction of each dataset and only the high accuracy obtained during the training is taken into account. The system

detects and displays the predicted emotion in accordance with the input from the web camera



VI. CONCLUSION AND FUTUREWORK

CONCLUSION:

It can be concluded that a reliable, secure, fast and an efficient Face Emotion detection system has been developed replacing a manual and unreliable system. This face detection and recognition system will save, reduce the amount of work done by the administration. There is no need for specialized hardware for installing the system as it only uses a computer and a camera. The camera plays a crucial role in the working of the system hence the image quality and performance of the camera in real time scenario must be tested especially if the system is operated from a live camera feed. The system can also be used in permission based systems and secure access authentication (restricted facilities) for access management, home video surveillance systems for personal security or law enforcement. The major threat to the system is Spoofing. For

future enhancements, anti-spoofing techniques like eye blink detection could be utilized to differentiate live from static images in the case where face detection is made from captured images from the classroom. From the overall efficiency of the system i.e. 83.1% human intervention could be called upon to make the system foolproof.

FUTURE ENHANCEMENT:

In the future, the next step would ideally be to make a complete and universal model that the organizations can use to help them reduce their human resource costs. Furthermore, can also be combined with other machine learning algorithms to create a highly robust and powerful system. Various other concepts of machine learning can be implemented to improve the efficiency of the system as a whole. The more the training data the more the accuracy. So, the importance should be given to select quality dataset and make the right partition to achieve more efficient system. Anti-spoofing techniques like eye blink detection could be utilized to differentiate live from static images from the classroom. To improve any business, feedback is very important. Example: hotels need feedback about their services and foods. But many customers or humans can't express their view (either written feedback or a word of mouth) exactly all the time because words either written or spoken are not that powerful as human expressions. Thus implementing suitable strategic ideas for different fields to make use of the detected emotions a very efficient manner.

VII. REFERENCES

- [1] H. Jung, S. Lee, J. Yim, and S. Park, "Joint fine-tuning in deep neural networks for facial expression recognition," in IEEE International Conference on Computer Vision, 2015.
- [2] X. Zhao, X. Liang, L. Liu, T. Li, Y. Han, N. Vasconcelos, and S. Yan, "Peak-piloted deep network for facial expression recognition," in European Conference on Computer Vision. Springer, pp. 425–442, 2016.
- [3] C. A. Corneanu, M. O. Sim'ón, J. F. Cohn, and S. E. Guerrero, "Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications," IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 8, 2016.
- [4] A. Mohammadian, H. Aghaeinia, and F. Towhidkhab, "Incorporating prior knowledge from the new person into recognition of facial expression," Signal, Image and Video Processing, vol. 10, no. 2, pp. 235–242, 2016.
- [5] S. Rifai, Y. Bengio, A. Courville, P. Vincent, and M. Mirza, "Disentangling factors of variation for facial expression recognition," in European Conference on Computer Vision, pp. 808–822, 2012.
- [6] P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain assessment with facial activity descriptors," IEEE Transactions on Affective Computing, vol. 8, no. 3, pp. 286–299, 2017.
- [7] K. Dobs, J. Schultz, I. B. ulthoff, and J. L. Gardner, "Task-dependent enhancement of facial expression and identity representations in human cortex," NeuroImage, vol. 172, 2018.

Managing Health Records Using Smart Contracts In Ethereum Blockchain And Performing Named Entity Recognition

Mrs. C JERIN MAHIBHA

Assistant Professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
jerinmahibha.msec@gmail.com

HARITHA R

Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
haritha1699@gmail.com

MEERAM

Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
meera2000.madhavan@gmail.com

SRUTIS

Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
sruti18898@gmail.com

Abstract— This paper proposes a system for managing health records on the blockchain environment with the help of smart contracts. The contracts are written in Solidity and are migrated to the Ethereum blockchain. Also a cryptocurrency wallet is used for the easier ethereum transactions. The health records uploaded by the patients are stored on the InterPlanetary File System which generates a hash value for the stored file and returns it. The returned hash value is then stored on the block. The patient can also analyse the medical record for understanding the drugs and diseases present in their report which is done by the Optical Character Recognition technique called, Named Entity Recognition(NER).

Keywords—Blockchain, Health records, smart contracts, IPFS

I. INTRODUCTION

With the advent of web technologies, the life of the human race is getting much easier from the way we perceive things previously. But it is also giving rise to a lot of issues like redundancy in data, tampering, counterfeiting, etc. The healthcare sector is showing tremendous improvements over the years but still there are many medical related crimes happening at the expense of patient's privacy. Hence, it is important to come up with a system that allows patients to digitize their medical files and store them securely.

Blockchain is becoming popular due to the high level of security it provides to its users. Smart Contract is the core of the Ethereum blockchain and it can perform decentralized transactions without relying on other third-party systems or databases. Also, the current health care system does not preserve the privacy of its patients, due to the widely centralized structure, the data once lost, is lost forever and the medical records can be easily hacked and are prone to many medical crimes such as Identity thefts, drug frauds, etc., To solve the above issues, the InterPlanetary File System(IPFS) can be used which is a content-based addressing network transport protocol for achieving decentralised storage. Further, to achieve Named Entity Recognition on the uploaded medical records, a Neural Named Entity Recognition and Multi-Type Normalization Tool for Biomedical Text Mining is applied.

The content of this paper is divided into five sections. In the first section, the background is introduced along with the motivation and purpose of this paper. In the second section, the technology and documents that are related to this paper are mentioned. In the third section, the structure of the system built in this paper is explained. In the fourth section, the construction of the system is presented in detail. In the last section, conclusions are made along with future enhancement possibilities.

II. RELATED WORKS

A. BLOCKCHAIN:

Blockchain is a technical solution for storing, verifying, transmitting and communicating network data based on cryptography. The core concept is to rely on cryptographic and mathematical decentralized algorithms. Blockchain can make participants reach a consensus without the intervention from the third party. This can solve the problem of lack of trust and unreliable value delivery. Through a public key, both of a private key, which are generated through asymmetric encryption, and the account address, we can perform a transaction or transmit data on the blockchain. The transaction on the blockchain is verified by every block on the blockchain instead of the third party.

After a block verifies the transaction addresses, transaction messages, and the source of cryptocurrency, it will push them to other nodes for verification. Once the transaction is confirmed by all nodes on the blockchain, the transaction data will be recorded in the block. The transaction record can't be changed and has full anonymity. Each transaction will produce a unique hash value, and each block header contains the previous block hash value, which connects all blocks and forms a chain.

Compared with the traditional financial transaction systems, which has to rely on trusted or guaranteed third-party organizations, the blockchain effectively implements the process of decentralization and combines

the network to ensure that transactions are collectively maintained and verified by each node and that transaction records are confirmed.

B. ETHEREUM

Ethereum is an open source platform and a public blockchain which functions on smart contracts. In addition to the usual cryptocurrency transactions, Ethereum features a decentralized Ethereum Virtual Machine (EVM) to deploy and apply smart contracts. In addition to mainly conducting cryptocurrency transactions, Ethereum also provides environmental constructions of private blockchains and multiple blockchains for testing, such as Rinkeby and Ropsten etc., which provide developers with different environments for testing and development.

C. SMART CONTRACTS

A smart contract is a piece of code which can be written in Solidity, Serpent, etc., and run on the blockchain. The smart contracts are helpful in eliminating the need of a third-party such as a bank which handles money transactions between two parties. Smart contracts permit trusted transactions and agreements to be carried out among disparate, anonymous parties without the need for a central authority.

D. INTERPLANETARY FILE SYSTEM

IPFS is a distributed system for storing and accessing files, websites, applications, and data. IPFS makes it possible to download a file from many locations that aren't managed by one organization.

Instead of being location-based, IPFS addresses a file by what's in it, or by its content. The content identifier is a cryptographic hash of the content at that address. The hash is unique to the content that it came from, even though it may look short compared to the original content. It also allows one to verify that you got what you asked for.

IPFS is based on the ideas of possession and participation, where many people possess each others' files and participate in making them available.

E. TRUFFLE

Truffle is a development environment and testing framework for blockchains using the Ethereum Virtual Machine (EVM). It has got built-in smart contract compilation and automated contract testing. Also provides package management with Node Package Manager (npm). And the truffle suite comes with an automated testing framework to test and run the smart contracts.

Truffle comes standard with npm integration, and is aware of the node_modules directory and therefore can be used to distribute contracts, dapps and Ethereum-enabled libraries via npm.

F. GANACHE

Ganache is a personal blockchain for rapid Ethereum and Corda distributed application development. Ganache can be used across the entire development cycle, thereby providing assistance to develop, deploy, and test the dApps in a safe and deterministic environment.

Ganache UI is a desktop application supporting both Ethereum and Corda technology. In addition, an Ethereum version of ganache is available as a command line tool ganache cli (formerly known as the testRPC). All versions of Ganache are available for Windows, Mac, and Linux.

G. METAMASK

MetaMask is a cryptocurrency wallet as well as a web browser extension that is available in Chrome, Firefox and Brave. It is used in order to store, send and receive Ethereum and ERC-20. In other words, it allows users to make Ethereum transactions through regular websites. MetaMask can be used to store keys for Ethereum cryptocurrencies. Only it is a wallet for the browser.

H. FLASK

Flask is a python web framework, basically a python library for developing web applications. It was developed by Armin Ronacher. Flask is based on Web Server Gateway Interface (WSGI) and Jinja2 template engine. A WSGI object can be created as follows:

```
from flask import Flask
app = Flask(__name__) # Flask constructor
@app.route('/')
def hello():
    return 'HELLO'
if __name__ == '__main__':
    app.run()
```

The run() method is used for starting the flask application.

I. REACT APP

React is an open source, front end, javascript library that is used for building user interfaces and it helps to manage the view layer. It helps set up the development environment so that the latest JavaScript features can be used. It also provides a great developer experience, and optimizes the application for production.

J. NAMED ENTITY RECOGNITION

Named Entity Recognition is used for extracting text from an image. For this paper, biomedical texts are involved and the a neural biomedical entity recognition and multi-type normalization tool (BERN) is used to identify those entities.

The following python code can be used for raw text

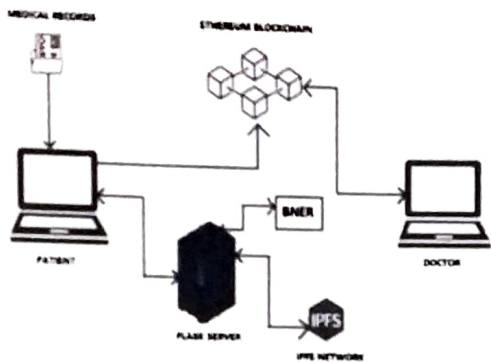
```
import requests
def query_raw(text,
url="https://bern.korea.ac.kr/plain"):
    return requests.post(url,
data={'sample_text':text}).json()
if __name__ == '__main__':
    print(query_raw("YOUR TEXT HERE"))
```

K. PYTESSERACT

Python-tesseract is an optical character recognition

(OCR) tool for python. That is, it will recognize and “read” the text embedded in images. Optical Character Recognition involves the detection of text content on images and translation of the images to encoded text that the computer can easily understand.

III. SYSTEM ARCHITECTURE



The User scans and uploads a medical record. The record is then encrypted from the client side and sent to the Flask Server. The Flask server sends this encrypted file to the IPFS network for storage. Once stored, it returns back a file hash. The file hash is then returned back to the client app and then it is safely stored in the ethereum blockchain. The user can then choose to perform NER (Named -Entity - Recognition) on the data in the medical record. This is to make the user comprehend the complex terms mentioned in the report. The record is sent to the Flask server. The Flask server runs the BNER(Biomedical NER) model on the received data after performing OCR on the report to get the text from the scanned medical record PDF. The medical keywords are then passed back to the Flask server. The Flask server sends this data back to the client app and the client can view the keywords and click on it for more information

IV. SYSTEM IMPLEMENTATION

In the system construction, we use ReactJS as the front end of the system, and Ganache is used as the blockchain environment for simulation. Besides, to write and test the smart contracts, we use Solidity. In the medical file upload section, we apply the IPFS API to uploading files to IPFS. The architecture of the proposed system is similar to that of MedRec. Both systems use smart contracts to register and obtain medical records, and utilize the Ethereum environment.

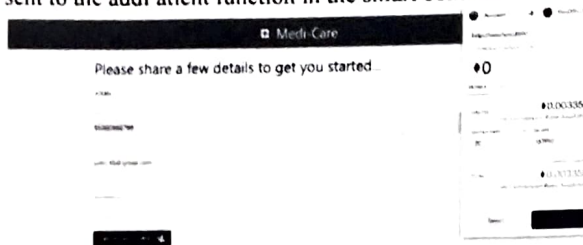
The first step is to stimulate the Ethereum blockchain using Ganache and the Truffle project is added. Then the migrations are performed using the “truffle migrate” command.

Secondly, the ganache and the metamask(cryptocurrency wallet) are to be connected by just importing the account by using its private key.

Thirdly, the React app has to be started using the command “npm start”. The react app starts running at <http://localhost:3000/>. The following are the modules present in the system and its associated processes.

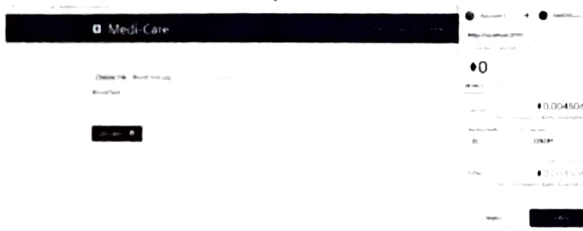
A CREATION OF USER ACCOUNT:

The user creates an account for themselves by entering a few personal details. A gas price needs to be paid for every transaction happening on the blockchain and thereby even while creating the account. The information collected during the registration process is stored using the web3 API and sent to the addPatient function in the smart contract.



B PROCESS OF UPLOADING FILES:

The user can now upload any necessary medical document to the portal by paying a gas price. The document uploaded is stored on the IPFS which in turn returns a hash value and this hash value is returned to the saveFile function in the smart contract and they are stored in the blockchain.



C. ANALYSING THE MEDICAL RECORDS:

The user can also perform an analysis on the medical report uploaded by them, and this analysis function returns keywords such as drugs and diseases present in the report, if any. For this purpose, the pytesseract tool is used. This will recognize and read the text embedded in images.

The user can click on the keywords which are mapped to Merriam Webster, to comprehend its meaning.

V. CONCLUSION AND FUTURE WORK

A management system for storing medical documents is created with React and IPFS on a blockchain environment. The decentralised structure is achieved with the help of IPFS. Also for the stimulation of the blockchain environment, we have used Ganache and also a cryptocurrency wallet, Metamask is used as an extension on the chrome browser which enables the user to clearly comprehend the gas price involved in each transaction.

However, in the future, further enhancements are expected to be made, such as a similar portal on the hospital end and also another portal for insurance claims. Also, simulations on private and public blockchains are expected to be made.

REFERENCES

- [1] Yeong-Sheng Chen, Wei-Kai Yang, and Jie-Si Chen, “An Electronic Medical Record Management System based on Smart Contracts”, Twelfth International Conference on Ubi-Media Computing, 2019

- [2] Ayesha Shanaz, Usman Qamari and Ayesha Khalid, "Using Blockchain for Electronic Health Records", IEEE ACCESS, October 9, 2019
- [3] Pei-Shan He, "Applying Blockchain Technology on an Outpatient Electronic Medical Record System," Master Thesis, National Taipei University of Technology, 2017
- [4] G. Wood, "Ethereum: A Secure Decentralised generalised transaction ledger EIP-150 revision," Tech Rep., Aug. 2017, p. 33
- [5] G. Jetley and H. Zhang, "Electronic health records in IS research: Quality issues, essential thresholds and remedial actions," Decis Support Syst., vol. 126, pp. 113–137, Nov. 2019
- [6] Dr. M. Humera Khanam, Md. A. Khudhuz and Prof. M. S. Prasad Babu, "Named Entity Recognition using Machine Learning Techniques for Telugu language", IEEE, 2016
- [7] Mei Ying, "Research on blockchain method for secure storage of medical records," Jiang Xi Normal University, Science Journal Vol. 41(5), pp. 481–487, 2017
- [8] Ethereum, <https://www.ethereum.org>
- [9] A Neural Named Entity Recognition and Multi-Type Normalization Tool for Biomedical Text Mining, <https://bern.korea.ac.kr/>
- [10] Truffle, <https://www.trufflesuite.com/docs/truffle/overview>
- [11] Ganache, <https://www.trufflesuite.com/docs/ganache/overview>
- [12] React, <https://reactjs.org/>
- [13] InterPlanetary File System, <https://docs.ipfs.io/>

A Survey of Local Data Storage over Cloud Storage

Sowmiya.V, Sandhya Nagarajan Iyer, Kavyaprabha S., Mrs. Aarthi S

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Assistant Professor, Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Submitted: 10-03-2021

Revised: 30-03-2021

Accepted: 01-04-2021

ABSTRACT: Storage Infrastructure is typically subdivided into Local and Cloud Storage Systems. On-Premise generally deals with local storage which deploys its servers taking all the risks which include maintenance and thereby makes the system be least cost-efficient. Cloud utilizes other's deployed servers on the internet and we have to pay for what we use rather than spending much on installing and maintaining servers on-premise. This paper focuses on surveys about efficient storage systems.

KEYWORDS: LocalStorage, Cloud Storage, EFS, Amazon S3

I. INTRODUCTION

Data Storage has become a task over the years due to a large amount of data being produced globally. Earlier we used to store data locally but now we see companies and firms all around us storing their data over the cloud.

[1] Local Storage as we know is storing data in physical devices like CDs, Floppy Disks, Hard Disks, USBs, etc. Here, you don't need the internet for accessing the stored data, and uploading data to local storage is fast. It kind stores data to about 2TB. Also, here you get to decide who accesses the data and how it is stored but still, it's vulnerable to attacks or the device may get lost. Data can easily be attacked or obtained by an unwanted user by hacking, spoofing, etc. It is comparatively expensive due to the hardware used. Also, if local backups get destroyed in unexpected events they can't be retrieved and data is lost. Nevertheless, the way of storing data has changed and evolved to the cloud.

[2] Cloud storage is just storing data over the internet servers unlike data stored locally on a computer's hard disk. Cloud makes accessing data and sharing it with anyone at any time and at any place viable unlike local storage where accessing and sharing the data is difficult. Though local

storage provides more space, the cloud works out cheaper per GB. It is much more secure and provides inbuilt mechanisms like firewalls, intrusion detection systems, event logs which make finding threats easier and keeping the files secure. It is much less vulnerable to attacks and threats. Cloud is much more scalable and cheaper in storing the large amount of data generated in companies. Also, automatic backups occur as and when you do any changes to your files in this way data isn't lost.

[3] Even Though the cloud is so efficient it has some minor drawbacks. If the provider faces a power or server issue, it is going to reflect on your cloud performance. Good internet plays a vital role in easy accessing, sharing, or uploading files. Also, the performance may sometimes be affected as the same servers are providing resources to many organizations. Yet, it is used extensively in the current world due to its high security, flexibility, very low maintenance, easy access, and backup.

II. METHOD OF STUDY

This paper surveys the journals of recent years. The main objective is to contemplate knowledge regarding efficient storage systems. The surveyed paper focuses on several works contributed to storage systems. We looked at several databases like IEEE, ResearchGate, IEEE Xplore, and other journals. We have also considered each article's references to get relevant papers.

III. CLOUD STORAGE SERVICES FOR COST-EFFICIENT FILE HOSTING

Typically today's cloud storage system deals with two distinct types of services offered, one is object storage like Amazon S3 and the other is filesystem storage like Amazon EFS. Amazon s3 is used for simple flat object storage with low unit storage price whereas EFS goes with hierarchical

complex filesystem operations with high unit storage price. A Cloud Storage user might look for both the services to be deployed at a low cost. So here is an intuitive approach to achieve this goal by combining both services. The technique is to host large files on s3 and small ones on EFS.

IV. DATA SECURITY FOR CLOUD STORAGE

The application of cloud storage is everywhere and it is now playing an indispensable role in all areas. Cloud is also vulnerable to risks, data tampering, unauthorized access, etc. So it is indeed a concern about security since it has got its wide range of deployment. Security can be enhanced with the help of encryption techniques. The Encryption Techniques comprises IBE(Identity-Based Encryption), ABE(Attribute-Based Encryption), Homomorphic Encryption, etc.

V. DETECTING INJECTIONS, ATTACK IN MYSQL

SQL queries are pretty much vulnerable to attacks. Here is a SEPTIC mechanism that is resistant to DBMS possible attacks. This SEPTIC technique is used to identify possible vulnerabilities in the applications. This technique is being implemented using MySQL and it has experimented with several test applications. These techniques lead to neither false positive nor false-negative results.

VI. BENEFITS OF AWS

The benefits mainly comply with Data Security, Regulatory Compliance, Cost-Effective, Flexible. On-premise storage requires us to set up servers of our own cost and also need to assist technicians in maintaining those servers. This may not lead to a cost-efficient system, therefore here comes the idea of cloud storage. In the cloud, the user can pay for what they use, Maintenance Free and it can be deployed anytime as needed.

VII. OPEN CHALLENGES FOR FUTURE RESEARCH

When it comes to security, whether on-premise or cloud is more secure than the answer will be the cloud is no more or less secure than the on-premise because people on both sides can commit mistakes. So security can be compromised irrespective of the storage system that we use.

VIII. CONCLUSIONS

Cloud is better than on-premise due to its reliability, security, and flexibility. Cloud removes the hassle of server maintenance thereby letting

you invest time, money, and resources into core business strategies. Real-time public access to systems and data from different devices irrespective of the location has guaranteed an uptime of 99%, cloud is now becoming the number one choice for UK businesses, with adoption rates reaching 88% in 2018.

IX. ACKNOWLEDGEMENT

Our primary and sincere thanks go to our pillar of support Dr. B. Monica Jenefer, Head of the Department, Department of Computer Science and Engineering, Meenakshi Sundararajan Engineering College for her profound inspiration, kind cooperation and guidance. We would like to thank our Professor Dr. M. K. Sandhya for her invaluable words of encouragement and guidance. We would also like to thank Assistant Professor Mrs. Aarthi.S for supporting this study.

REFERENCES

- [1]. Jinlong E Mingkang Ruan Ennan Zhai(2020)"HyCloud: Tweaking Hybrid Cloud Storage Services for Cost-Efficient Filesystem Hosting", IEEE Journal, Volume: 28, Issue: 6, Dec. 2020.
- [2]. Pan Yang; Naixue Xiong; Jingli Ren(2020),"Data Security and Privacy Protection for Cloud Storage", IEEE Access (Volume: 8), 16 July 2020.
- [3]. Iberia Medeiros, Miguel Beatriz, Nuno Neves, and Miguel Correia(2019),"SEPTIC: Detecting Injection Attacks and Vulnerabilities Inside the DBMS", IEEE, VOL. 68, NO. 3, SEPTEMBER 2019.
- [4]. B. Ahuja, A. Jana, A. Swarnkar, and R. Halder, "On preventing SQL injection attacks," Adv. Comput. Syst. Secur., vol. 395, pp. 49–64, 2015. Akamai Technologie, Cambridge, MA, USA, "Q1 2016 state of the Inter-net/security report," Tech. Rep., vol. 3, no. 1, Jun. 2016.
- [5]. S. Bandhakavi, P. Bisht, P. Madhusudan, and V. N. Venkatakrishnan, "CANDID: Preventing SQL injection attacks using dynamic candi-date evaluations," in Proc. 14th ACM Conf. Comput. Commun. Secur., Oct. 2007, pp. 12–24.
- [6]. C. A. Bell, Expert MySQL. New York, NY, USA: Apress, 2007.
- [7]. Amazon Web Services, <http://aws.amazon.com/>.
- [8]. Amazon EBS, <http://aws.amazon.com/ebs/>.
- [9]. Amazon EC2, <http://aws.amazon.com/ec2/>.
- [10]. Amazon S3, <http://aws.amazon.com/ebs/s3/>.
- [11]. Li, X. Yang, S. Kandula and M. Zhang,

- "CloudCmp: Comparing Public Cloud Providers", Proc. of ACM Internet Measurement Conference (IMC), pp. 1-14, 2010.
- [12]. Z. Lai, Y. Cui, M. Li, Z. Li, N. Dai and Y. Chen, "TailCutter: Wisely Cutting Tail Latency in Cloud CDN under Cost Constraints", Proc. of IEEE International Conference on Computer Communications (INFOCOM), pp. 1845-1853, 2016.
- [13]. Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, et al., "Towards Network-level Efficiency for Cloud Storage Services", Proc. of ACM Internet Measurement Conference (IMC), pp. 115-128, 2014.
- [14]. A. Akavia, S. Goldwasser and V. Vaikuntanathan, "Simultaneous hardcore bits and cryptography against memory attacks", Proc. TCC, pp. 474-495, 2009.
- [15]. T. Bhatia and A. K. Verma, "Data security in mobile cloud computing paradigm: A survey taxonomy and open research issues", J. Supercomput., vol. 73, no. 6, pp. 2558-2631, Jun. 2017.
- [16]. B. Casemore, "Network modernization: Essential for digital transformation and multicloud", Nov. 2019.



Certificate of Publication

This is to confirm that

Mrs. P Revathi, Jesintha Praveena A, Porkodi C D

Published following article

Breast Cancer Prediction Using Machine Learning Algorithms

Volume 3, Issue 7, pp: 4088-4093

www.jjaem.net

A Peer Reviewed Journal



International journal of Advances in Engineering
and Management (IJAEM)

ISSN: 2395-5252

A handwritten signature in black ink, appearing to be the name of the Publication Head.

Publication Head

Breast Cancer Prediction Using Machine Learning Algorithms

Mrs. P Revathi
Assistant Professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
revathisprajesh@gmail.com

JESINTHA PRAVEENA A
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
jesisandhiya7@gmail.com

PORKODI C D
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
porkodid@gmail.com

Abstract— Breast cancer is type of tumor that occurs in the tissues of the breast. It is most common type of cancer found in women around the world and it is among the leading causes of death in women. This becomes very handy, especially in the medical field where diagnosis and analysis are done through these techniques. Wisconsin Breast cancer data set is used to perform a comparison between Support vector machine, Logistic regression, Random forest classifier, K-Nearest neighbor. Based on the result of performed experiments, the Random Forest algorithm shows the highest accuracy (99.76%) with the least error rate.

Keywords—Random forest classifier, SVM, Machine Learning Algorithm.

I. INTRODUCTION

The second major cause of women's death is breast cancer (after lung cancer). 246,660 of women's new cases of invasive breast cancer are expected to be diagnosed in the US during 2016 and 40,450 of women's death is estimated. Breast cancer is a type of cancer that starts in the breast. Cancer starts when cells begin to grow out of control. Breast cancer cells usually form a tumour that can often be seen on an x-ray or felt as a lump. Breast cancer can spread when the cancer cells get into the blood or lymph system and are carried to other parts of the body. The cause of Breast Cancer includes changes and mutations in DNA. There are many different types of breast cancer and common ones include ductal carcinoma in situ (DCIS) and invasive carcinoma. Others, like phyllodes tumours and angiosarcoma are less common. There are many algorithms for classification of breast cancer outcomes. The side effects of Breast Cancer are – Fatigue, Headaches, Pain and numbness (peripheral neuropathy), Bone loss and osteoporosis. There are many algorithms for classification and prediction of breast cancer outcomes.

The present paper gives a comparison between the performance of four classifiers: SVM, Logistic Regression, Random Forest and kNN which are among the most influential data mining algorithms. It can be medically detected early during a screening examination through mammography or by portable cancer diagnostic tool. Cancerous breast tissues change with the progression of the disease, which can be directly linked to cancer staging. The stage of breast cancer (I-IV) describes how far a patient's cancer has proliferated. Statistical indicators such as tumour size, lymph node metastasis, distant metastasis and so on are used to determine stages. To prevent cancer from spreading, patients have to undergo breast cancer surgery, chemotherapy, radiotherapy and endocrine. The goal of the research is to identify and classify Malignant and Benign patients and intending how to parametrize our classification techniques hence to achieve high accuracy. We are looking into many datasets and how further Machine Learning algorithms can be used to characterize Breast Cancer. We want to reduce the error rates with maximum accuracy. 10-fold cross validation test which is a Machine Learning Technique is used in JUPYTER to evaluate the data and analyse data in terms of effectiveness and efficiency.

II. RELATED WORKS

A. SUPPORT VECTOR MACHINE:

Random forest is an algorithm that integrates multiple trees to form a forest through the idea of ensemble learning. Its basic unit is decision tree. Each decision tree is a classifier, so for an input sample, N trees will have N classification results. Random forest integrates all the classification voting results and specifies the category with the most voting times as the final output. The construction process of random forest is as follows:

1) Let N be the number of samples in the original training set, and M be the number of characteristic attributes. Bootstrap sampling technique is used to extract N samples from the original training set to form a training subset.

2) m features are randomly selected as candidate features ($m \leq M$) from M feature attributes. Each node of the decision tree selects the optimal attributes according to some rules (Gini impurity, information divergence, etc.) to split until all the training samples of the node belong to the same class, and they are completely split without pruning in the process.

3) Repeat the above two steps k times to build k decision trees and generate the random forest.

4) Using random forest to make decision, let x be the test sample, h_i be the single decision tree, Y be the output variable, which is the classification label, I be the indicative function, H be the random forest model, and the decision formula.

B. Advantages and disadvantages of random forest algorithm

Random forest algorithm has many advantages. As a combination algorithm of classifiers, it can optimize the overall performance of the classification system by synthesizing the capabilities of several weak classifiers, which is better than a single classifier. When generating random forest, each decision tree is independent of each other and generated at the same time. The training speed is fast, and it is easy to make parallel method.

At the same time, the random forest algorithm also has some disadvantages. Because the randomness of the decision tree added by the random forest almost only occurs in the feature selection when the decision tree is generated, the fixity of the decision tree generation rules will lead to a certain degree of over fitting. At the same time, in the face of data with high and unbalanced feature dimensions, performance of algorithm is seriously weakened because high-dimensional data usually contains a large number of irrelevant and redundant features.

C. K-Nearest Neighbour:

K-Nearest Neighbour (KNN) is said to be the simplest and the most straight forward classification algorithm. Like most machine learning algorithms, K-NN does not learn anything from the provided dataset and its attributes, but simply use the points from the training data and finds the K number of nearest neighbors to that data point using Euclidean Distance and classify it to the class which has the first K neighbors closest to it.

D. Support Vector Machine:

Support vector machine (SVM) is a quite simple classification algorithm. This classifier is named so because it takes the help of vectors in the feature space to classify the class of a new vector [9,10]. The Maximum Margin Hyper-plane (MMH) decides whether the new vector belongs to class one or class two. If the data point lies beyond the negative hyper-plane or to the left of MMH then it belongs to the class one, else it belongs to the class two, where class one and two are two different classes in a given situation. SVMs can also be used if there are more than two classes. It is a supervised learning algorithm which is used for both classification and regression problems. It consists of theoretical and numeric functions to solve the regression problem. It provides the highest accuracy rate while doing prediction of large dataset. It is a strong machine learning technique that is based on 3D and 2D modelling.

E. Logistic Regression:

Logistic regression was introduced by statistician DR Cox in 1958 and so predates the field of machine learning. It is a supervised machine learning technique, employed in classification

classification is based on training data. Logistic regression can be used to estimate the probability of a patient having breast cancer. However, it is a linear classifier and cannot be used for non-linearly separable data. Support Vector Machines (SVM) can be used to predict the probability of a patient having breast cancer from the independent variables.

The general workflow is:

1. get a dataset
2. Train a classifier
3. make a prediction using such classifier

2 Data Collection

To predict breast cancer, authors used Wisconsin Diagnostic data set collected from the UCI Machine Learning Repository. There were 699 instances with total of eleven features. Out of eleven features, ten features are input features and remaining one feature is treated as output feature. The whole data set is divided into training and testing instances in the ratio of 80:20. It means out of 699 instances, 560 instances were used as training data set and the remaining 140 instances were used as a testing data set.

2.1 Data Pre-processing

Wisconsin Diagnostic data set for breast cancer prediction has some missing values. To handle these missing values data pre processing was also used on the mentioned data set. The attribute like *Flare Number* column has missing features in the form of "-" string which need to be inputted. 16 such instances of missing values were found in this feature. These missing values were replaced by the average-mean values of the features. On the other hand, the attributes like *sample code number* have no relevance in predicting breast cancer so such types of attributes have been dropped from the data set.

2.2 Flask

Flask is a python web framework, basically a python library for developing web applications. It was developed by Armin Ronacher. Flask is based on Web Server Gateway Interface (WSGI) and Jinja2 template engine. A WSGI object can be created as follows:

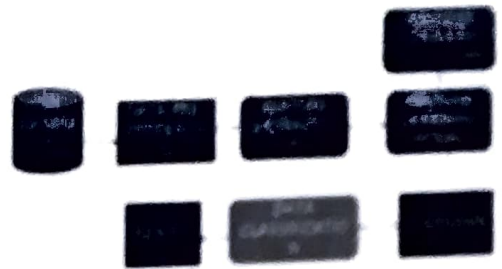
```
from flask import Flask
app = Flask(__name__) # Flask constructor
@app.route("/")
def hello():
    return "HELLO"
if __name__ == '__main__':
    app.run()
```

The `run()` method is used for starting the flask application.

2.3 PYTESERACT

Python-tesseract is an optical character recognition (OCR) tool for python. That is, it will recognize and "read" the text embedded in images. Optical Character Recognition involves the detection of text content on images and translation of the images to encoded text that the computer can easily understand.

3. SYSTEM ARCHITECTURE



The system explains how the user acquires the required information of the patients by means of collecting the data and after gathering the information the details has to be processed using data pre processing finally the pre processed data has been extracted according to their feature given and it has been applied with the machine learning algorithms finally it has been classified and analyzed according to the classification algorithms such as Support Vector Machine, K-Nearest Neighbor, Logistic Regression and Random Forest Classifier. After analyses using these algorithms it has been predicted that the breast cancer patient is in benign or in malignant stage.

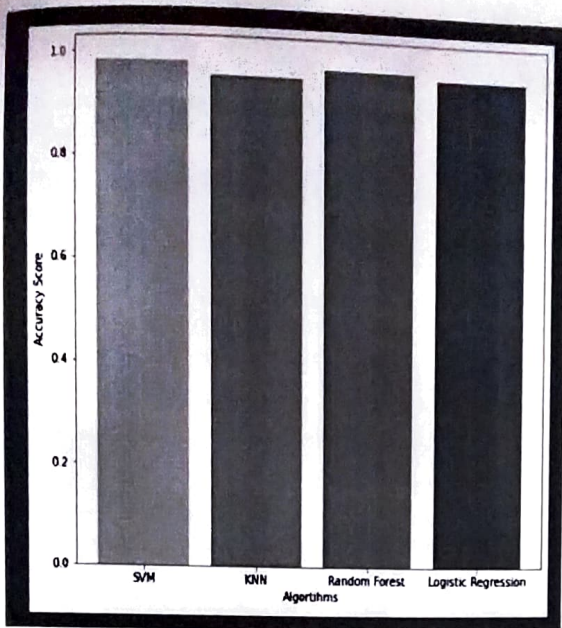
IV. SYSTEM IMPLEMENTATION

For the implementation of the ML algorithms, the data set was partitioned into the training set and testing set. A comparison between all the four algorithms will be made. The algorithm that gives the best results will be supplied as a model to the website. The website will be made from a python framework called Flask. And it will use the database on itself Python and Flask libraries. This data set is available on the UCI Machine Learning Repository. It consists of 30 real world attributes which are multi-variant. The total number of instances is 569 and there are no missing values in this data set.

The patient books an appointment through our website. The patient will then meet the doctor offline for the respective appointment. The doctor will first check the patient manually then perform a breast mammogram or an ultrasound. But after sound will show an image of the breast consisting the lumps or not. If the lumps are detected, a biopsy will be performed. The digitized image of the Fine Needle Aspirate (FNA) is what forms the features of the data set. These numbers will be provided to the system by the doctor and the model will detect if its a benign or a malignant cancer. Breast cancer if found at its early stage will help save lives of thousands of women a year. The report will be then forwarded to the patient as their respective account.

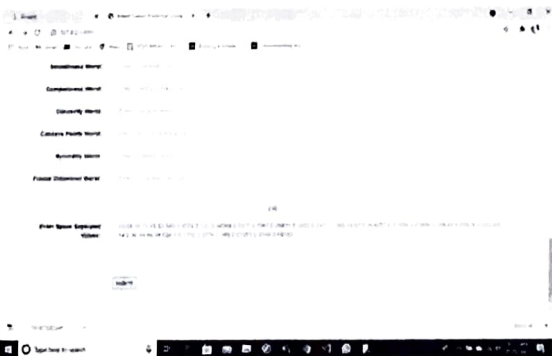
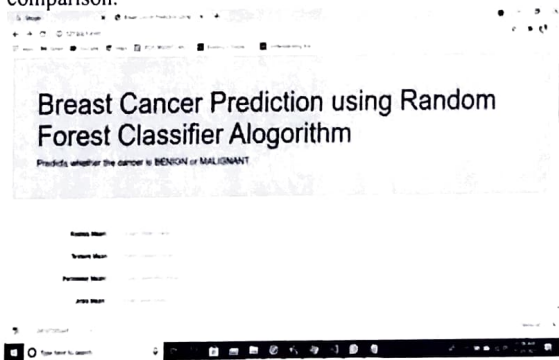
TABLE 1

This graph shows the comparability of all the four algorithms that has been analyzed for the prediction of breast cancer type. In this graph, the Machine Learning algorithms such as Support Vector Machine, Logistic Regression, Random Forest Classifier and K-Nearest Neighbor were implemented. The accuracy given by SVM is 97.55%, K-NN is 95.25%, LR is 97% and that by RF is 98.0%. This paper concludes that Random Forest Classifier (RFC) have better accuracy of 98.0%.



B. GIVE INPUT IN THE WEBPAGE

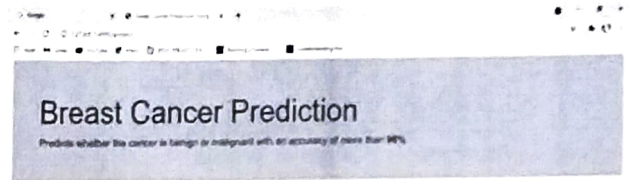
The user enters the featured details of him in the home page which includes features such as radius mean,perimeter mean,texture mean,area mean and 30 more features to identify whether the patient is in benign or malignant.This is the homepage of the website that is used to predict the breast cancer type using the algorithm that provided more accuracy on comparison with few other machine learning algorithms. Random forest classifier provided better solution on comparison.



C. BENIGN TYPE CANCER

Once when the features has been entered it checks whether the breast cancer type is benign or malignant.It checks with the

parameter of test data sets along with the trained data sets.The web page shows that the cancer is predicted to be benignBenign (non-cancerous) breast conditions are very common, and most women have them. In fact, most breast changes are benign. Unlike breast cancers, benign breast conditions are not life-threatening. But some are linked with a higher risk of getting breast cancer later on.

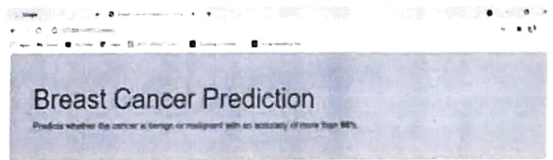


Cancer is Predicted to be Benign



C.MALIGNANT TYPE CANCER

The patient can check for the other type of cancer using the featured data sets available.The data for malignant type cancer has been entered and checked and it shows cancer is predicted to be malignant.If a tumor is found to be malignant, you have breast cancer or another form of cancer. Malignant tumors can be aggressive and may spread to other surrounding tissues.A biopsy may be done on a suspicious lump, which can identify whether it is a tumor, as well as whether it is benign or malignant.



Cancer is Predicted to be Malignant



V.CONCLUSION AND FUTURE WORK

In statics analysis of Breast Cancer Prediction, Machine Learning algorithms have been used to train classifiers with features.The Machine Learning algorithm such as Support Vector Machine,Logistic Regression,Random Forest Classifier and K-Nearest Neighbor were implemented.The accuracy given by SVM is 97.35% , KNN is 95.23%, LR is 95% and that by RFC is 98.6%.This paper concludes that Random Forest Classifier (RFC) have better accuracy of 98.6%.The Scope of our project is to detect the patient is in benign or in malignant stage with the help of Machine Learning Algorithms.In future these techniques may be implemented on data sets that consists of images.The system may also be integrated with an Android application.The accuracy of the model created may be increased in order to give better predictions.

REFERENCES

- [1] Zerina Masetic, [2016] "Congestive detection using random forest classifier", vol.130,ISSN 0167-2607.
- [2] S.Gokhale , [2018] "Ultrasound characterisation of breast masses",India.The indian journal of radiology imaging,Vol.19,pp.242-249.
- [3] Abien Fred M.Agarap [2019]"An Application of Machine Learning Algorithms" on the Wisconsin Diagnostic Dataset for Breast cancer Detection,7th February.
- [4] L.J .Grimm, J.R.Marks [2020] "prediction of upstaged ductal carcinoma in situ",IEEE trans, June,vol.67,no.6,pp.1565-1572.
- [5] Haifeng Wang and Sang Won Yoon [2017] "Breast Cancer Prediction Using Data Mining Method", Department of System Science and Industrial State University of New York at Binghamton, May.
- [6] Priyanka Gupta and Prof.Shalini L [2018] "Analysis of Machine Learning Techniques for Breast Cancer Prediction" VIT University, Vellore, 5 May.
- [7] Sep [2019] "Breast Cancer Wisconsin (Diagnostic) Data Set". [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic))
- [8] "Analysis of Machine Learning Techniques for Breast Cancer Prediction". International Journal of Engineering and Computer Science, <http://www.ijecs.in/index.php/ijecs/article/view/4071>.,May,Vol.7,no.05, [2018]
- [9] F.C.C.Garcia,I.Muga [2016] "Random forest for malware classification",aiXiv preprint:1609.07770

EMAIL SPAM DETECTION USING CONVOLUTIONAL NEURAL NETWORK

S. Kamalini^[1], M. Nandhini^[2], R.B. Nithya^[3], R.Venkatesh^[4],
[1],[2],[3]-Students [4]-Assistant Professor
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE

Unsolicited e-mail (Spam) has become a major issue nowadays for each e-mail user. Through email, companies and individuals send advertisements for various products, undesirable harmful news, contents, and fake proposals etc. The spam emails result in unnecessary consumption of network bandwidth resulting blocking email servers. Existing systems make it very difficult to detect spam as these emails are written or generated in a very special way so that anti-spam filters cannot detect such emails. Machine learning systems now a day's used to consequently filter the spam email in an exceptionally effective rate. Email filtering job relies upon data classification approach. While classify data, choose the most astounding performing classifier is a fundamental progress. In this project, we proposed Convolutional Neural Network (CNN) classifier to filter out the spam Email from the network. The performance of the proposed models was evaluated based on the three criteria, the prediction accuracy, learning time and false positive rate. CNN attain higher accuracy than other algorithms.

Keywords— Classify spam and ham messages,CNN

1. INTRODUCTION

Email Spam is an emerging issue of the web users. The recent enhancements in the spam rate had affected an excessive concern among the present internet community. Numerous resolutions to deal with the issue had been recommended for the non-technical and technical limits. This chapter presents a brief introduction of E-mail document classification issues in the perspective of spam email filtering, the issues that the web users have been recently facing because of spam mail. This study recommended or deployed solution for avoiding spams and it is also presents their related advantages and disadvantages. The different types of spam filtering mechanism are presented. Finally, brief introduction of medical website spam mail-filtering process using machine-learning algorithms and give the main drawbacks that the issue faces.

Various issues have arises from spam mails. Firstly, it wastes the organizations resources and network resources and a lot of bandwidth is wasted at the time of spam mail downloading from the inbox. The most of the organizations pay for the network and internet resources, so cost is an important factor for them. Secondly, spam emails can cause serious issues for Personal Computer (PC) users not installed in antivirus solutions. Thirdly, it is a waste of time for

organization works, resulting in decreases the company productivity and thus causes the overall system performance.

II. RELATED WORKS

In [],The project is about classifying the spam and ham message from the dataset.In this paper,accuracy level has been increased.

In [],This project consists of convolutional neural network which is used to classify the messages from the dataset.

In [],Here in the classification system, it train the dataset and identify the data is spam or ham.

III.WORKFLOW OF EMAIL SPAM DETECTION

Here we propose a system to identify the messages are spam or ham by using the convolutional neural network. The Email collection dataset we have taken from neural network repository. Data preprocessing is the most important phase in detection models as the data consists of ambiguities, errors, redundancy which needs to be cleaned beforehand. Data is transformed into lowercase and change the data types according to algorithm needs. The attributes are identified for classifying process and system performs feature extraction and then these classification systems classify the content into spam or ham.

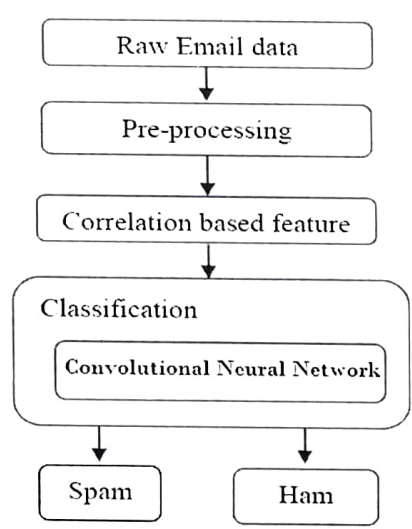


Fig 1. Architecture of email spam detection

).

The architecture, explains how the data and workflow throughout the system and how the data is classified. This is hence a basic conceptual picture of what is going on behind the screen and the components that are involved in classifying the data to the user. In this system we will have a dataset that consist of messages. In the architecture, first it get the message from the user. Then, in the preprocessing step it filter all the repeated and error messages from the dataset. In the third step, it convert messages into vector data types. In the fourth step, it train the dataset and classify the messages are spam or ham.

IV. CONVOLUTIONAL NEURAL NETWORK

A neural network is a system of interconnected artificial "neurons" that exchange messages between each other. The connections have numeric weights that are tuned during the training process, so that a properly trained network will respond correctly when presented with an image or pattern to recognize. The network consists of multiple layers of feature-detecting "neurons".

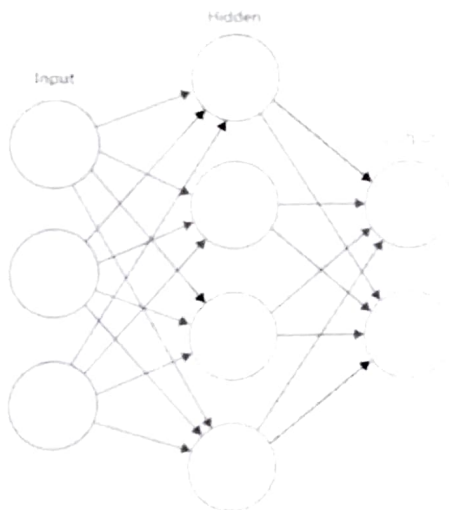


Fig 2. Artificial neural network.

Each layer has many neurons that respond to different combinations of inputs from the previous layers. As shown in Fig 1.3, the layers are built up so that the first layer detects a set of primitive patterns in the input, the second layer detects patterns of patterns, and the third layer detects patterns of those patterns, and so on. Typical CNNs use 5 to 25 distinct layers of pattern recognition.

This algorithm is processed under five layers. They are convolutional layers, pooling or sub sampling layers, non-linear layers, relu layers, fully connected layers

V. PREPROCESSING

This is the first stage that is executed whenever an incoming mail is received. This step consists of tokenization which is a process that removes the words in the body of an email. It also transforms a message to its meaningful parts. It takes the email and divides it into a sequence of representative symbols called tokens. Emphasised that these representative symbols are

extracted from the body of the email, the header and subject. It asserted that the process of replacing information with distinctive identification symbols will extricate all the characteristics and words from the email exclusive of taking into account the meaning.

VI. SPAM AND HAM

According to Wikipedia "the use of electronic messaging systems to send unsolicited bulk messages, especially mass advertisement, malicious links etc." are called as spam. "Unsolicited means that those things which you didn't asked for messages from the sources. So, if you do not know about the sender the mail can be spam. People generally don't realize they just signed in for those mailers when they download any free services, software or while updating the software. "Ham" this term was given by Spam Bayes around 2001 and it is defined as "Emails that are not generally desired and is not considered spam".

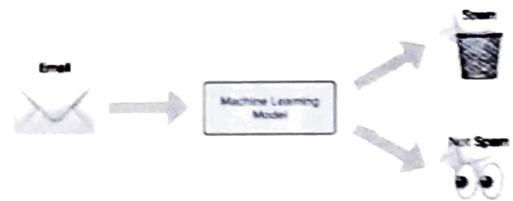


Fig 3. Classification into Spam and non-spam

Machine learning approaches are more efficient, a set of training data is used, and these samples are the set of email which are pre classified. Machine learning approaches have a lot of algorithms that can be used for email filtering.

VII. RESULTS AND DISCUSSIONS

With this result, it can be concluded that the CNN gives the best outcome but has limitation due to class-conditional independence which makes the machine to misclassify some tuples. Ensemble methods on the other hand proven to be useful as they using multiple classifiers for class prediction. Nowadays , lots of emails are sent and received and it is difficult as our project is only able to test emails using a limited amount of corpus. Our project, thus spam detection is proficient of filtering mails giving to the content of the email and not according to the domain names or any other criteria.

Therefore, at this it is an only limited body of the email. There is a wide possibility of improvement in our project. The subsequent improvements can be done: "Filtering of spams can be done on the basis of the trusted and verified domain names." "The spam email classification is very significant in categorizing e-mails and to distinct e-mails that are spam or non-spam." "This method can be used by the big body to differentiate decent mails that are only the emails they wish to obtain."

```

val_loss: 0.0593 - val_acc: 0.9839
Epoch 7/10
4457/4457 [=====] - 186s 42ms/step - loss: 0.0418 - acc: 0.9892 -
val_loss: 0.0652 - val_acc: 0.9794
Epoch 8/10
4457/4457 [=====] - 184s 41ms/step - loss: 0.0447 - acc: 0.9850 -
val_loss: 0.0624 - val_acc: 0.9857
Epoch 9/10
4457/4457 [=====] - 177s 40ms/step - loss: 0.0287 - acc: 0.9912 -
val_loss: 0.0418 - val_acc: 0.9865
Epoch 10/10
4457/4457 [=====] - 170s 38ms/step - loss: 0.0160 - acc: 0.9957 -
val_loss: 0.0454 - val_acc: 0.9874
wait a minute, this is a SPAM!
In [4]:

```

Fig 4. Spam message

```

Epoch: 0/10
4457/4457 [=====] - 184s 41ms/step - loss: 0.0420 - acc: 0.9868 -
val_loss: 0.0553 - val_acc: 0.9892
Epoch 7/10
4457/4457 [=====] - 191s 43ms/step - loss: 0.0368 - acc: 0.9877 -
val_loss: 0.0492 - val_acc: 0.9901
Epoch 8/10
4457/4457 [=====] - 189s 42ms/step - loss: 0.0234 - acc: 0.9933 -
val_loss: 0.0540 - val_acc: 0.9892
Epoch 9/10
4457/4457 [=====] - 194s 43ms/step - loss: 0.0155 - acc: 0.9951 -
val_loss: 0.0593 - val_acc: 0.9883
Epoch 10/10
4457/4457 [=====] - 154s 35ms/step - loss: 0.0127 - acc: 0.9962 -
val_loss: 0.0585 - val_acc: 0.9892
Ohh, this is a normal message.
In [5]:

```

Fig 5. Ham message

VIII. REFERENCES

1. Agarwal, K and T. Kumar,(2018) "Email Spam Detect ion Using Integrated Approach of Naïve Bayes and Particle Swarm Optimization," Second International Conference on Intelligent Comput ing and Control Systems (ICICCS), Madurai, India, 2018, pp. 685-690
2. Ameen A K and B. Kaya,(2018) "Spam detect ion in online social networks by deep learning," International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, pp. 1-4, 2018
3. Awad W A & S.M, ELseuofi. (2011). "Machine Learning Methods for Spam E-Mail Classification". International Journal of Computer Science & Information Technology. 3. 10.5121/ijcsit .2011.3112
4. Diren, D.D., Boran, S., Selvi, I.H., & Hatipoglu, T. (2019). "Root Cause Detect ion with an Ensemble Machine Learning Approach in the Multivariate Manufacturing Process"
5. Harisinghaney, Anirudh, Aman Dixit , Saurabh Gupta, and Anuja Arora.(2014) "Text and image-based spam email classification using KNN, Naïve Bayes and Reverse DBSCAN algorithm." In Opt imization, Reliabilty, and Information Technology (ICROIT), International Conference on, pp.153-155. IEEE, 2014
6. Karim, A., Azam, S., Shanmugam, B., Krishnan, K., & Alazab, M. (2019). "A Comprehensive Survey for Intelligent

Spam Email Detect ion". *IEEE Access*, 7, 168261-168295. [08907831].

7. Mohamad, Masurah, and Ali Selamat (2015). "An evaluation on t he efficiency of hybrid feature selection in spam email classification." In Computer, Communications, and Control Technology (14CT), International Conference on, pp. 227-231. IEEE, 2015

8. Shradhanjali, P rof. T oran Verma (2017)" E-Mail Spam Detect ion and Classification Using SVM and Feature Extraction" in International Journal Of Advance Research, Ideas and Innovat ion In Technology, ISSN: 2454-132X Impact factor: 4.295

9. Suryawanshi, Shubhangi & Goswami, Anurag & Pat il, Pramod. (2019). "Email Spam Detection: An Empirical Comparative Study of Different ML and Ensemble Classifiers". 69-74. 10.1109/IACC48062.2019.8971582

10. Tasnim Kabir, Abida Sanjana Shemont i, At if Hasan Rahman.(2018) "Notice of Violation of IEEE Publication Principles: Species Identification Using Partial DNA Sequence: A Machine Learning Approach", IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE), 2018.



Certificate of Publication



This is to confirm that

Kamalli V, Ramya Prabha Rk, Sakthe Priya P, Dr.M.K.Sandhya

Published following article

Deep Neural Network for Handwriting Recognition System.

Volume 3, Issue 5, pp: 222-225

www.ijaem.net

A Peer Reviewed Journal

**International Journal of Advances in Engineering
and Management (IJAEM)**

ISSN: 2395-5252

Publication Head

Deep Neural Network for Handwriting Recognition System

Kamalli V, Ramya Prabha Rk, Sakthe Priya P, Dr.M.K.Sandhya ,

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Professor, Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai

Submitted: 01-05-2021

Revised: 09-05-2021

Accepted: 10-05-2021

ABSTRACT: In a large-scale industry or organization, a number of documents will be sent and received. Due to variation in handwriting, interpretation of these texts will be difficult and manual automation of these would be nearly impossible. In our paper, we intend to translate the handwritten texts into digitalized text. Digitized text is useful as it makes information readily available. Some of the previously used approaches for HTR are, Deep Convolutional Neural Network (DCNN), Optical Character Recognition (OCR), Self-organising map and Projection approach. Problems faced here are over-fitting, identification of wrong text, limitations to human handwritten text, slow computation and reduced efficiency with smaller dataset. To overcome these hurdles in the existing system, we have proposed a system based on ANNs (Artificial Neural Networks). Recently, it was discovered that ANNs have an excellent capacity in sequence data analysis such as natural language processing. In our proposed system we are using both Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) to recognize the handwritten text efficiently. The classifier has CNN layers to extract features from the input image and RNN layers to propagate information through the image. This helps to optimize the data effectively and increases the accuracy.

Keywords: Convolutional neural networks, Artificial neural network, Recurrent neural network, convert into digital text, natural language processing

I. INTRODUCTION

Handwriting recognition is undoubtedly one of the most challenging areas of pattern recognition. It is extremely useful in a wide range of real-world practical problems, including documentation analysis, mailing address

interpretation, bank check processing, signature verification, document verification, and many others. Several pattern recognition approaches have been applied to both online and off-line handwriting recognition, including statistical methods, structural and syntactic methods, and neural networks. Some reading systems identify strokes; others try to identify characters, groups of characters, or entire words.

In the problem of handwritten text recognition, neural networks are heavily used for efficiently predicting handwritten symbols. Character recognition has long been a critical area of Artificial Intelligence. Recognition is a trivial task for humans, but to make a computer program that does character recognition is extremely difficult.

Recognizing patterns is just one of those things humans do well and computers don't. One of the primary means by which computers are endowed with human-like abilities is through the use of neural networks. Neural Networks are particularly useful for solving problems that cannot be expressed as a series of steps such as recognizing patterns, classifying them into groups, and series prediction and data mining. ANN approach for character recognition is now gaining importance because of ANN's highly parallel architecture and fault tolerance.

II. LITERATURE SURVEY

M. Ghanim et al. [5] proposed a Arabic Handwriting Recognition using DCNN. CNN algorithm is used to convert typewritten text into digitized text and HAC algorithm for clustering characters. Some advantages of this paper are optimization of the data with better accuracy and easy to recognize different handwriting styles whereas it may recognize words which are not in

dictionary when the written words are hard to read which is a major disadvantage.

Eltay et al. [2] proposed an approach to recognize Handwritten Arabic Texts using Deep learning. Review and investigation of different deep learning architectures and modeling choices for Arabic handwriting recognition is been done and an Adaptive data - augmentation algorithm is used to build this system. Some of the pros of this paper are more data can be used for recognition and it also reduces data over fitting while on the other hand ,this system is not efficient for small dataset .

Uddin et al. [1] proposed a literature review for handwritten Optical Character Recognition. Handwritten character Recognition and Printed character Recognition using OCR is used. Some advantages of this paper are processing information is fast and recognition engines which is used with imaging can capture highly specialized data sets whereas OCR systems equipments are expensive and workload to data collectors because OCR has severe limitations when it comes to human handwriting which serve as a major disadvantages.

R.Ezhilarasi et al. [3] proposed OCR for text recognition using recurrent neural network. OCR is used to give an impressive recognition exactness for manually written text using recurrent neural network which is used to improve accuracy. Some of the pros of this paper are it can process any length input and model size doesn't increase for longer input . while on the other hand ,the computation of this neural network is slow and therefore training is different which serve as major cons.

Aly et al. [4] proposed a robust digit recognition system using self organising map network. Development of this new efficient deep unsupervised network Deep Convolutional Self-organizing Maps (DCSOM) helps to learn invariant image representation from unlabeled visual data. Some pros of this paper include easy data interpretation and data visualization. On the other hand, if dataset is small then map size is also small. Map size depends on training data and some clusters end up in distant parts of map making it difficult to visualise the similarities between them serves as major cons.

Zin et al. [6] proposed a character recognition system by projection approach. Segmentation is carried out mainly on labelling and projection concept. Pros of this paper include better accuracy in recognition of cursive words compared to other techniques. But still it's maximum accuracy is only around 65% which is a drawback.

III. PROPOSED SYSTEM

Our proposed system uses ANN integrated with CTC loss and decoding algorithm to convert the handwritten text to digital or machine text. The fundamental steps that should be followed to accomplish our goal are:

Data Acquisition:

Dataset is the origin of knowledge for our Machine Learning models. Henceforth, feeding well-defined data and selecting the suitable dataset is crucial for obtaining profitable outcomes. Thus we have used the IAM dataset for our proposed system.

This dataset comprises about 1539 forms filled out by approximately 657 writers. Data is annotated on line level and word level. It has the most unique words out of all other handwriting recognition system datasets.

Data Preprocessing:

Data Preprocessing is that step in which the data gets transformed, or Encoded, to bring it to such a state that now the machine can easily parse it. In other words, the preprocessing is done to make things easier for the classifier.

Preprocessing techniques includes contrast normalization, and random modifications to the original dataset to enlarge the datasets

Classification:

Artificial Neural Network(ANN) serves as a classifier for our proposed system. Multiple Convolutional Neural Network(CNN) layers are trained to extract relevant features from the input image. These layers output a 1D or 2D feature map which is handed over to Recurrent Neural Network (RNN) layers.

The RNN propagates information through the sequence. Later on, the output of the RNN is mapped onto a matrix that contains a score for each character per character element. In this system, an implicit word segmentation method based on the decoding algorithm for the RNN output is proposed and evaluated. In this way, the ANN can be fed with line-text images.

Decoding:

Since ANN is trained using a specific coding scheme, a decoding algorithm must be applied to outputs of RNN to obtain the final text. Decoding can take advantage of the Language Model(LM). The loss calculation and decoding from the matrix are carried out by the Connectionist temporal classification (CTC)

operation of CTC serves two-fold (1) as a loss function,(2) as a decoder

In this system, we have to privilege to choose the decoding algorithm to be applied the options available are(1) Best Path Decoding(2)Beam Search (3) Word Beam Search. The default decoder set for training is best path decoding algorithm.

IV. SYSTEM ARCHITECTURE

Our proposed system makes use of ANNs to convert handwritten text to digital text. Multiple Convolutional Neural Network(CNN) layers are trained to extract relevant features from the input image. These layers output a 1D or 2D feature map which is handed over to Recurrent Neural Network (RNN) layers. The RNN propagates information through the sequence. Later on, the output of the RNN is mapped onto a matrix that contains a score for each character per character element.

Since the ANN is trained using a specific coding scheme, a decoding algorithm must be applied to the outputs of RNN to obtain the final text. The training, loss function calculation, and decoding from this matrix are carried out by the Connectionist Temporal Classification(CTC) operation. Decoding can take advantage of the language model(LM).

Preprocessing is done to make things easier for the classifier. Preprocessing operations include contrast normalization, and making random modifications to the original images to enlarge the size of the dataset. Our proposed system uses implicit segmentation, in which the ANN itself is trained to segment the words by learning the white space characters. Since the decoded text might contain misspelled words, a text postprocessing method is applied to them and the final output is displayed.



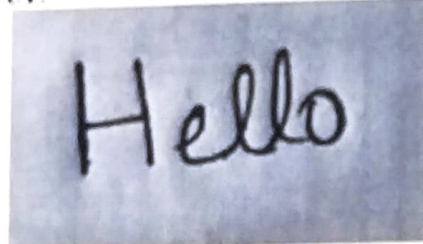
V. RESULTS TRAINING THE MODEL:

```
python ./train.py --data-dir ./data --num-epochs 100 --batch-size 32 --num-workers 4 --num-gpus 1 --log-dir ./logs
```

Epoch	Batch	Loss	Char Error Rate
1	1	28.0	0.000
1	2	28.0	0.000
1	3	28.0	0.000
1	4	28.0	0.000
1	5	28.0	0.000
1	6	28.0	0.000
1	7	28.0	0.000
1	8	28.0	0.000
1	9	28.0	0.000
1	10	28.0	0.000
1	11	28.0	0.000
1	12	28.0	0.000
1	13	28.0	0.000
1	14	28.0	0.000
1	15	28.0	0.000
1	16	28.0	0.000
1	17	28.0	0.000
1	18	28.0	0.000
1	19	28.0	0.000
1	20	28.0	0.000
1	21	28.0	0.000
1	22	28.0	0.000
1	23	28.0	0.000
1	24	28.0	0.000
1	25	28.0	0.000
1	26	28.0	0.000
1	27	28.0	0.000
1	28	28.0	0.000
1	29	28.0	0.000
1	30	28.0	0.000

Here, IAM data-set is trained. For each epoch CTC loss value is calculated individually for each batch. Training will stop when character error rate is not improved and the model which has lower error rate is saved.

INPUT:



This is our sample text image not exceeding 32-characters stored in the folder named data. It is then fetched, undergoes preprocessing and sent to the trained model.

OUTPUT:

```
python ./train.py --data-dir ./data --num-epochs 100 --batch-size 32 --num-workers 4 --num-gpus 1 --log-dir ./logs
```

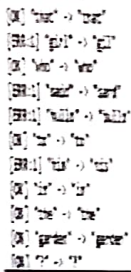
```
python ./infer.py --data-dir ./data --model-dir ./models --num-workers 4 --num-gpus 1 --log-dir ./logs
```

```
python ./postprocess.py --data-dir ./data --model-dir ./models --num-workers 4 --num-gpus 1 --log-dir ./logs
```

```
python ./display.py --data-dir ./data --model-dir ./models --num-workers 4 --num-gpus 1 --log-dir ./logs
```

Machine text from our input handwritten text is recognized and displayed.

ACCURACY:



Character error rate: 18.38294662807 word accuracy: 74.654177681

Character error rate(CER) is number of edit operation performed on the input text by total length of input text. For example, in the above image, “him” is recognized as “his”. Number of edit operation is 1(because m is recognized as s) and total length of original text is 3. Therefore, $CER=1/3$. This mismatch in recognition occurs due to many factors like quality of input image, similarities between letters and slanting.

Word accuracy is number of letters recognized correctly by total number of words in validation dataset. For example, consider there is 120 words in a validation dataset. Out of which 89 words are recognized correctly. Then accuracy for the same is calculated as,

$$Accuracy = 89/120 * 100 = 74.16\%$$

VI. CONCLUSION

Image recognition is an important aspect for the image processing. The collected dataset is trained using ANN which represents the current state-of-the-art for variety of applications. Thus, we extensively analyzed the model by carefully selecting their parameters and showing its robustness for handling our dataset. In this paper we perform five steps namely, image acquisition, pre-processing, classification, decoding. The input image is preprocessed and fed into Convolutional Neural Network (CNN) for feature extraction and the extracted feature map is handed over to Recurrent Neural Network (RNN). The RNN outputs a matrix that contains a probability distribution over the characters at each image position. The Connectionist Temporal Classifier (CTC) decodes this matrix and final machine text is displayed. Artificial Neural Network (ANN), has good accuracy for handwriting recognition because more training will result in more accurate writing recognition.

VII. FUTURE WORK

This work can be further extended for recognition of other languages. It can be used to convert the fax and newspapers into text format. Before these techniques, approach is mostly focused on efficient feature extraction from these images. Once you have good feature representation of the data, its very easy to build a model. In future, we need systems that can read a character array and modify it to the form that you want.

This project can be further extended to use multi-dimensional LSTM and deslanting technique can be performed at preprocessing level to obtain better accuracy. Eg: like a linguist or auto translation or detected text to voice conversion etc. In a long run, systems or phones would replace the jobs of scanners, dtp operators etc.

REFERENCES

- [1] Jamshed Memon, Maira Sami, Rizwan Ahmed Khan and Mueen Uddin (July 2020). “Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)”. IEEE Access, vol. 8, pp.142642-142668.
- [2] Mohamed Eltay, Abdelmalek Zidouri and Irfan Ahmad (May 2020). “Exploring Deep Learning Approaches to Recognize Handwritten Arabic Texts”. IEEE Access, vol. 8, pp.89882 -89898.
- [3] Parthiban.R, Ezhilarasi.R and D.Saravanan (December 2020). “Optical Character Recognition for English Handwritten Text Using Recurrent Neural Network”. IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), vol. 20, pp.1417-1419.
- [4] Saleh Aly and Sultan Almotairi (June 2020). “Deep Convolutional Self-Organizing Map Network for Robust Handwritten Digit Recognition”. IEEE Access, vol. 8, pp.107035-107045.
- [5] Taraggy M Ghanim, Mahmoud I Khalil, Hazem M.Abbas (2020). “Comparative Study on Deep Convolution Neural Networks DCNN-based Offline Arabic Handwriting Recognition”. IEEE Access, vol. 4, pp.1-20.
- [6] Thi Thi Zin, Shin Thant, Ye Htet and Pyke Tin (June 2020). “Handwritten Characters Segmentation using Projection Approach”. IEEE Global Conference on Life Sciences and Technologies, vol. 20, pp. 107-108.



ISSN No. : 2321-9653

IJRASET

**International Journal for Research in Applied
Science & Engineering Technology**

IJRASET is indexed with Crossref for DOI-DOI : 10.22214

Website : www.ijraset.com, E-mail : ijraset@gmail.com



ISRA Journal Impact
Factor: **7.429**



45.98
INDEX COPERNICUS



THOMSON REUTERS
Research ID: 10-9681-2018



TOGETHER WE REACH THE GOAL
SJIF 7.429

Certificate

*It is here by certified that the paper ID : IJRASET35237, entitled
Eye Blindness Detection(Diabetic Retinopathy) using ResNet
Convolutional Neural Network*

by

Mrs. P. Revathi

*after review is found suitable and has been published in
Volume 9, Issue VI, June 2021*

in

*International Journal for Research in Applied Science &
Engineering Technology*

Good luck for your future endeavors

By [Signature]

Editor in Chief, IJRASET

EYE BLINDNESS DETECTION(DIABETIC RETINOPATHY)USING RESNET CONVOLUTIONAL NEURAL NETWORK

Kiran T
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
kirant2212@gmail.com

Laxmi Prasad M
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
laxmiprasad.1505@gmail.com

Abstract:

This paper proposes a system to detect diabetic retinopathy faster using Resnet Convolutional neural network. Diabetic retinopathy is an eye condition that can cause vision loss and blindness in people who have diabetes. Diabetic retinopathy affects up to 80 percent of those who have had diabetes for 20 years or more. At least 90% of new cases could be reduced with proper treatment and monitoring of the eyes.
Keywords: Diabetic Retinopathy, Resnet CNN, Deep Learning.

I. INTRODUCTION

Diabetic retinopathy, also known as diabetic eye disease, is a medical condition in which damage occurs to the retina due to diabetes mellitus. It is a leading cause of blindness in developed countries. Diabetic retinopathy affects up to 80 percent of those who have had diabetes for 20 years or more. At least 90% of new cases could be reduced with proper treatment and monitoring of the eyes. Commonly this disease is detected at very later stage due to the unavailability of doctors and so it becomes untreatable. The diabetes pandemic requires new approaches to improve the detection and prevention. Thus our proposal is to train a big and varying dataset using a CNN deep learning

Mohamed Ejaz M
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
ejaz1999@gmail.com

Mrs P. Revathi
Asst professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
revathisprajesh@gmail.com

ResNet model. Deep learning models are best known for identifying the small features in an image.

The content of this paper is divided into five sections. In the first section, the background is introduced along with the motivation and purpose of this paper. In the second section, the technology and documents that are related to this paper are mentioned. In the third section, the structure of the system built in this paper is explained. In the fourth section, the construction of the system is presented in detail. In the last section, conclusions are made along with future enhancement possibilities.

II. TECHNOLOGIES USED

Deep Learning

Deep learning is a computer software that mimics the network of neurons in a brain. It is a subset of machine learning and is called deep learning because it makes use of deep neural networks. Deep learning algorithms are constructed with connected layers. The first layer is called the input layer and the last layer is called the output layer. All layers hidden between are called hidden layers. A deep neural network provides state of the art accuracy in many tasks, from object detection to speech recognition. They can learn automatically, without predefined knowledge explicitly coded by the programmers.

PYTHON

Python is a free, open-source programming language. Python is also a great visualization tool. It provides libraries such as Matplotlib, seaborn and bokeh to create stunning visualizations. Python is the most popular language for machine learning and deep learning. Python strives for a simpler, less-cluttered syntax and grammar while giving developers a choice. Python is meant to be an easily readable language. Its formatting is visually uncluttered, and it often uses English keywords where other languages use punctuation.

Pandas

Pandas is a popular Python package for data science, and with good reason: it offers powerful, expressive and flexible data structures that make data manipulation and analysis easy, among many other things. The DataFrame is one of these structures. Pandas is built on top of the NumPy package, meaning a lot of the structure of NumPy is used or replicated in Pandas. Data in pandas is often used to feed statistical analysis in SciPy, plotting functions from Matplotlib, and machine learning algorithms in Scikit-learn.

Numpy

Numpy is the core library for scientific computing in Python. It provides a high-performance multidimensional array object, and tools for working with these arrays. If you are already familiar with MATLAB, you might find this tutorial useful to get started with numpy. NumPy is a Python library that is the core library for scientific computing in Python. It contains a collection of tools and techniques that can be used to solve computer mathematical models of problems in Science and Engineering.

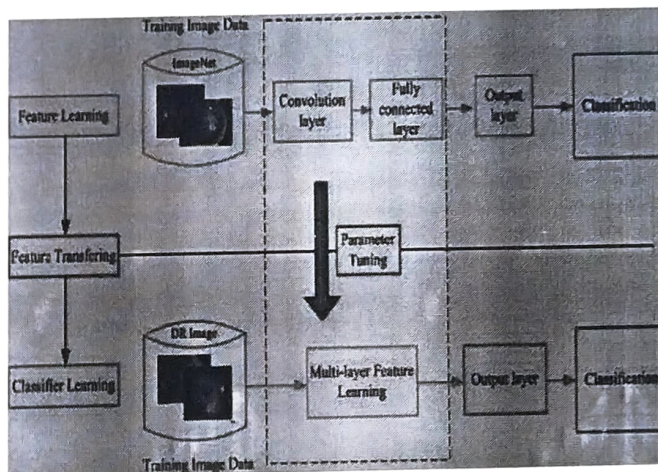
Matplotlib

Plotting of data can be extensively made possible in an interactive way by Matplotlib, which is a plotting library that can be demonstrated in Python scripts. Plotting of graphs is a part of data visualization, and this property can be achieved by making use of Matplotlib. One of the free and open-source Python libraries which is basically used for technical and scientific computing is Python

SciPy. Matplotlib is widely used in SciPy as most scientific calculations require plotting of graphs and diagrams.

III. SYSTEM ARCHITECTURE

The Architecture explains how the patient's images are uploaded and with the help of our large dataset, the data is preprocessed and our model is trained to classify whether the patient has Diabetic retinopathy, if yes it classifies the severity of diabetic retinopathy.



IV. SYSTEM IMPLEMENTATION

The proposed system consists of four modules. The image data is analyzed using exploratory visualization. Then using that information, required preprocessing of the images is done. Pre processing of images extracts the features that causes the disease and thus increases accuracy. A ResNet convolutional neural network model is created and the images are fed into the model for training. After training, the model is optimized by hyper parameter tuning. Finally the test data is used to predict the accuracy of the model in detecting thoracic disease or not. Performance analysis is done using loss & accuracy graph, ROC curve and confusion matrix.

Image Analysis and Visualization:

first step in your data analysis process. Here, you make sense of the data you have and then figure out what questions you want to ask and how to frame them, as well as how best to manipulate your available data sources to get the answers you need. You do this by taking a broad look at patterns, trends, outliers, unexpected results and so on in your existing data, using visual and quantitative methods to get a sense of the story this tells.

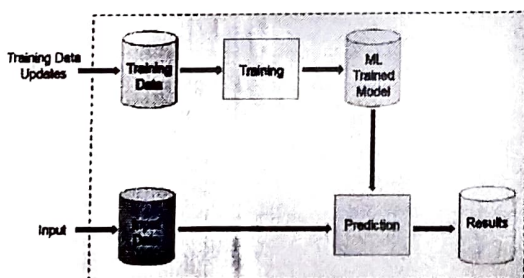
Image Pre-processing:

Sometimes you may find some data missing in the dataset. We need to be equipped to handle the problem when we come across them. Obviously you could remove the entire line of data but it may lead to removing crucial information unknowingly which will cause a lot of errors in the prediction phase. One of the most common ideas to handle the problem is to take a mean of all the values of the same column and have it to replace the missing data.

The library that we are going to use for the task is called Scikit Learn preprocessing. It contains a class called Imputer which will help us take care of the missing data.

Model Creation & Training:

In the proposed method the network will be implemented for feature and classification though ResNet proved success in solving problems in medical imaging. Training very large medical data having many parameters but limited numbers of samples has a problem of overfitting can be well adapted by ResNet. It has reduced the number of parameters to learn by a simple network to speed up training on a selective typical value. From the input image pixels are used directly to feed the network for image classification.



V. CONCLUSION AND FUTURE WORK

In this work, through a collection of different eye retina images, we demonstrated how to classify eye blindness level accurately. The proposed system showed us that ResNet could aid in the detection of the diabetic retinopathy disease. We constructed our model using various techniques such as data augmentation and windowing. In rural places, due to lack of medical infrastructure diabetic retinopathy disease is not analysed by expert doctors. Clinical implementation of this algorithm can help decrease financial costs since the algorithm currently utilizes only easily obtainable clinical data such as fundal images. The high value of sensitivity shows how dependable this technology can be for implementation in real-world scenarios. In future, we can create a website where anyone can upload the eye image and get the output in the form of eye blindness level. This website can be developed for the public which can be accessed by everyone.

REFERENCES

1. "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms" by Lifeng Qiao, Ying Zhu, Hui Zhou at IEEE Access (Volume 8)
2. "Detecting diabetic retinopathy using deep learning" <https://www.researchgate.net/publication/3177308>
3. <https://www.longdom.org/open-access/diagnosis-of-diabetic-retinopathy-using-machine-learning-jrd-1000127.pdf>
4. <https://www.scribd.com/document/391482030/20695>
5. Kaggle. (Jul. 27, 2015). Diabetic Retinopathy Detection. Accessed: May 7, 2018. [Online]. Available: <https://www.kaggle.com/c/diabeticretinopathy-detection/>
6. V. Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *Jama*, vol. 316, no. 22, pp. 2402–2410, 2016.
7. R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.
8. B. Graham. (Aug. 6, 2015). Kaggle Diabetic Retinopathy Detection Competition Report. Accessed: May 20, 2018. [Online]. Available: <https://kaggle2.blob.core.windows.net/forum-messageattachments/88655/2795/competitionreport.pdf/>
9. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.
10. M. Lin, Q. Chen, and S. Yan. (2013). "Network in

PERSONALITY PREDICTION USING RANDOM FOREST CLASSIFIER AND GRADIENT BOOSTING CLASSIFIER

Dr.B.Monica Jenefer
Head of the Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai,India
bmonica20@gmail.com

MADHUMITHA P
Dept. of Computer Science and
Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India
madhumithap19@gmail.com

MIRA SOWNDARYA R
Dept. of Computer Science and
Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India
mirasowndarya1@gmail.com

PRIYADHARSHINI S
Dept. of Computer Science and
Engineering
Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India
priyasridharan123@gmail.com

Abstract— Personality tests are very important to gauge how a person impacts his team. Many organizations find it important to know the personality of the person before recruiting. The system introduced here uses a web application to predict the personality of the candidate using machine learning algorithms applied on the recent tweets and through a psychometric test to classify a person's personality among the 16 different personality types as described by Myers Briggs Indicator. This gives the employer speculation about how the person will respond to different situations.

Myer Briggs Personality Indicator scheme is used which is composed of 16 personality types that are generated by the combination of the following: Extraversion/Introversion, Sensing/Intuition, Thinking/Feeling and Judging/Perceiving. The algorithms used for classification in this system are Random forest classifier and Gradient Boosting classifier. These two algorithms are ensembled together using the Voting classifier. The ensembled model is used for predicting the MBTI personality type.

Keywords—Personality prediction, Myers Briggs Type Indicator (MBTI), Random Forest Classifier, Gradient Boosting Classifier, Tweet analysis, Psychometric test.

I. INTRODUCTION

Personality is defined by the how a person reacts under different circumstances and is based on their characteristic features. As the process of recruitment is changing, organizations are concerned about the personality of their employees. While recruiting people, one of the factors to consider is the personality of the candidate and their influence on the team. As a member of a team, recognizing one's colleagues' personality type may improve the understanding and appreciation of one another's differences – and can show one how to get along better with them.

Psychologists propose that the personality of a candidate is closely related to their performance, their job tenure and how much involvement they show to their tasks. Hence, finding the right people with the right nature helps put right people on the right positions and as a result benefits both the organization and the candidate.

In the recruitment process, there is usually a round where the nature of the candidate is judged through a face-to-face round. This hence varies according to the interviewer. In this system, we use the Myers-Briggs Type Indicator (MBTI) to classify the personality of the candidates. The system uses machine learning algorithms Random Forest Classifier and Gradient Boosting Classifier to classify the candidate into one of the 16 personality types, each of which is represented by a four-letter code - ISTJ, ISTP, ISFJ, ISFP, INFJ, INFP, INTJ, INTP, ESTP, ESTJ, ESFP, ESFJ, ENFP, ENFJ, ENTP and ENTJ.

The methods used in this system target forming a structured personality prediction method that functions in a uniform way. The proposed system has two methods of detecting the personality – through a psychometric test and through the analysis of Tweets. This system, thus aims to automate the process of filtering the candidates based on the required personality.

II. LITERATURE SURVEY

Madhura Jayaratne et. al. [1] used HEXACO personality model that summarizes human personality characteristics in terms of six dimensions, or factors: Honesty-Humility (H), Emotionality (E), Extraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). HEXACO trait values from textual content were inferred using natural language processing (NLP) and machine learning methods through a regression model. They used Random forest classifier used to classify the degree to which each characteristic was present. The candidates were given a psychometric test and based on their answers, the results were given. The accuracy of the model was calculated based on whether the candidate accepted their results to be fair or not.

Sandhya Katiar et. al.[2] used the Big Five personality model which represents the five dimensions of personality that can be used to analyze and predict the behavior of a person. The big five factors are Openness to experience, Agreeableness, Extraversion, Neuroticism and

onscientiousness. The system uses three different algorithms to predict user behaviors and form a base from received pattern. These three algorithms are the first is Big Five Model along with advanced data mining, second is Support Vector Machine and third is Naïve Bayes theorem. They identify user's personality based on past user history and obtained traits pattern. The main advantage of this system is that Naïve Bayes algorithm combines efficiency with reasonable accuracy, though it is to be noted that support vector machine had lower accuracy than Naïve Bayes algorithm.

Yunan Li et. al. [3] proposed a system presents a deep Classification-Regression Network (CR-Net) for analyzing personality. The steps included data preprocessing, feature extraction using CR-Net and then ETR Regression and this was used for analyzing big five personality problem. In this CR-Net takes the ResNet-34 as the backbone network. It uses the classification features for the regression process towards the end of the network. Finally, features of different modalities are fused. The final prediction is obtained through ETR Regression. The noted advantage of this system is that the divergence between the prediction and label is small. Also, the system has more robust regression results. However, it was computationally expensive, time consuming and sometimes caused over fitting because of too much dependence on training data.

Aditi V. et. al. [4] Kunte used Adaboost algorithm, Multinodal Naïve Bayes and LDA Algorithm to predict the personality of the candidate. The dataset was created from the social media platform Facebook and then the text was tokenized. Pre-processed data is then visualized and applied to the above-mentioned classification algorithms to the dataset, result of which classify personality in any one of the five class labels provided by Big five test of psychology. The system also provided graphical representation of the output for easier interpretation. Also, more focus can be thrown on real-time data which can have significance with real world. Combining Machine Learning algorithms can be useful in improving accuracy.

Yash Mehta et al. [5] published a paper where in they described the different prospects of applying machine learning algorithms to determine the personality of a person. They explored the computational datasets, and their applications in various industries. Also, state-of-the-art machine learning models for personality detection. This survey concluded that, most of the current datasets focus on the Big-Five personality model.

Michael Tadesse et. al. [6] attempted to investigate the predictability of the personality traits of Facebook users based on different features and measures of the Big 5 model. The system used XGBoost algorithm. It improves the computing power for boosted trees algorithms. This is because it was developed for achieving highest computational power and efficiency. Also, LIWC (Linguistic Enquiry and Word Count Tool) and SPLICE (Structured Programming for Linguistic Cue Extraction) linguistic dictionaries were used to improve results.

III. METHODOLOGY

A. MBTI PERSONALITY TYPES

Myer Briggs Personality Indicator reflects the different psychological preferences of a person based on the following categories: Introversion/Extraversion; Intuition/Sensing; Thinking/Feeling; Judging/Perceiving.

Extraversion/Introversion: Extroverts are action-oriented people. They generally have excellent outgoing or social skills. The extroverts feel energized when they are a part of social gatherings. Introverts are people who are thought-oriented. They enjoy being alone. They are drawn towards deep and emotional connections than casual ones.

Intuition/Sensing: People with a Sensing personality focus on facts and details. They rely on hands-on experience. People with Intuition personality focus more on possibilities. They usually spend time imagining the future and believe in abstract theories.

Thinking/Feeling: People with a Thinking personality make decisions that are impersonal and logical based on facts and objective data. People with a Feeling personality consider others and their emotions before making decisions.

Judging/Perceiving: People with Judging personalities make strong and firm decisions. People with Perceiving personalities make open, adaptable, and flexible decisions.

B. RANDOM FOREST CLASSIFIER

The Random forest Classifier is a supervised Machine learning algorithm used for classification and regression by using decision trees. The Random forest classifier spawns a number of decision trees for randomly selected samples from the training dataset. It votes among different decision trees to make final conclusions

- (I) Select random samples from training dataset and generate a decision tree for that sample
- (II) Predict results from each decision tree
- (III) Vote for decision tree using mode in case of classification and mean in case of regression
- (IV) Select most voted decision tree as final prediction.

C. GRADIENT BOOSTING CLASSIFIER

Boosting is an ensemble technique used to improve the accuracy of an algorithm. There are so many algorithms available for boosting accuracy. One such algorithm is Gradient Boosting Classifier which is also based on decision trees. This classifier tries to make improvements on the predecessor by reducing errors. It makes use of iteration technique and improves the predecessor in each iteration. In simple words, this classifier ensembles predictions from decision tree in the previous iteration with decision tree predictions of current iteration. In each iteration the discrepancies are reduced to produce accurate results.

D. SYSTEM ARCHITECTURE

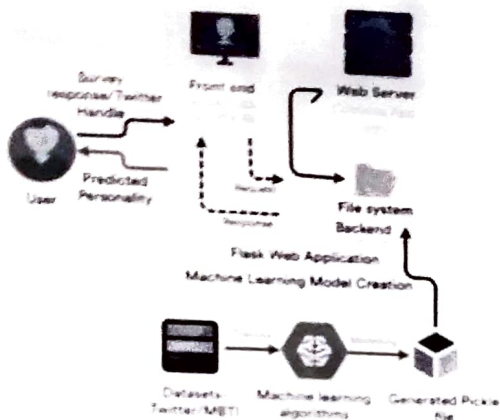


Fig 1. Architecture of the System

The system consists of two modules: Tweet based personality prediction and Psychometric test-based personality prediction. Tweet based prediction gets the twitter handle of a person as input. This page is designed using HTML, CSS. The Flask application uses the necessary packages to extract the recent tweets of that person and predicts their MBTI personality. A model ensemble from Random forest and Gradient Boosting classifier models is used for prediction.

The other module- Psychometric test-based personality prediction aims to collect the way a person would respond to certain scenarios. Those answers are converted to textual statements and the MBTI personality type is predicted using the model trained with Random Forest Classifier and Gradient Boosting Classifier.

The trained models are generated using textual dataset. The dataset contains a collection of tweets and their respective MBTI personality type.

IV. SYSTEM IMPLEMENTATION

A. DATA PROCESSING

Dataset vectorization is the process of transforming the data into numerical values that can be understood by the machine learning model. The textual data is converted into numerical entries in the matrix form. In this system, count vectorization technique is used to vectorize the data. In count vectorization, a document term matrix is generated. The term matrix contains a set of variables that indicates if a particular word appears in the data. For each word, a column is assigned. As a result, a sparse matrix is generated. In this system, the vectorization module is the same for both the psychometric test and tweet analysis modules. The vectorized data is stored in the form of pickle files so that it can be used while making predictions.

B. ALGORITHMIC MODEL

The steps involved in generating the machine learning model is as follows -

- (I) The vectorized data is split into test data and training data.
- (II) These are then passed into the Random Forest Classifier. The model is fitted with labels and features.
- (III) The same process is repeated again for Gradient Boosting Classifier.
- (IV) Using Voting Classifier, both the fitted models are ensemble. Prediction is done using the test data and the generated model is stored as pickle files.
- (V) The trained model can be then used for making predictions.

C. WEB APPLICATION

The machine learning model is deployed using a Flask based web application. On the front end, HTML, CSS, JavaScript and W3CSS are used. On the backend, Flask framework is used. The logic of the system is coded in Python. The web application includes a home page where the details of the personality classification are given. Then the option to take the psychometric test and the tweet analysis is displayed. Upon selecting the psychometric test, the user is given a form and the response is passed to the server. The response is then used to predict the personality using the machine learning model generated and the result is displayed. Upon the selection of tweet analysis, the user is prompted to enter the twitter handle, which is then sent to the server. Once the result is predicted, it is displayed to the user.

The twitter handle submitted by the user is used to uniquely identify the twitter account. The tweet extractor uses a twitter API with consumer key, consumer secret key, access token and secret access token which is assigned to the developer at the time of creation of the API. The tweet extractor scrapes up to 100 tweets from the user's account, normalizes it and the cleaned data is used by the predictor.

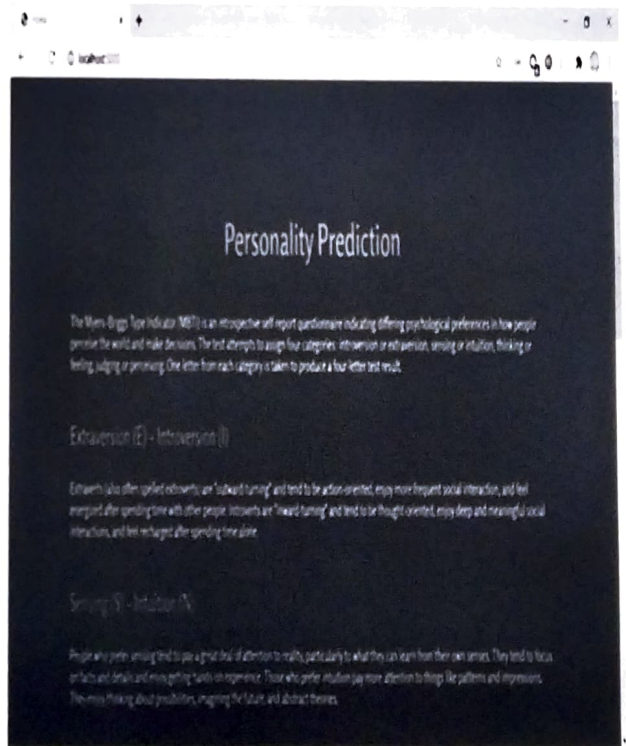


Fig 2 Home page of the web application

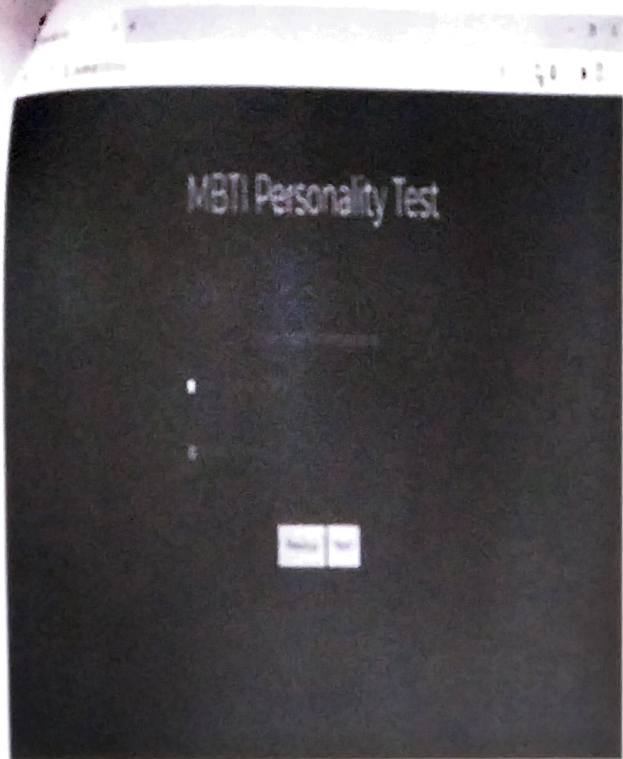


Fig 3 Psychometric test form

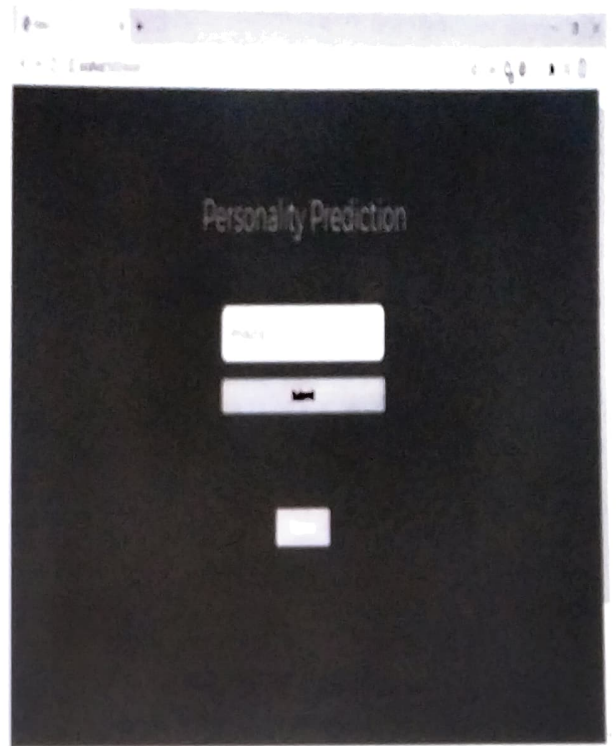


Fig 5 Twitter Handle Input form



Fig 4 Psychometric test prediction results

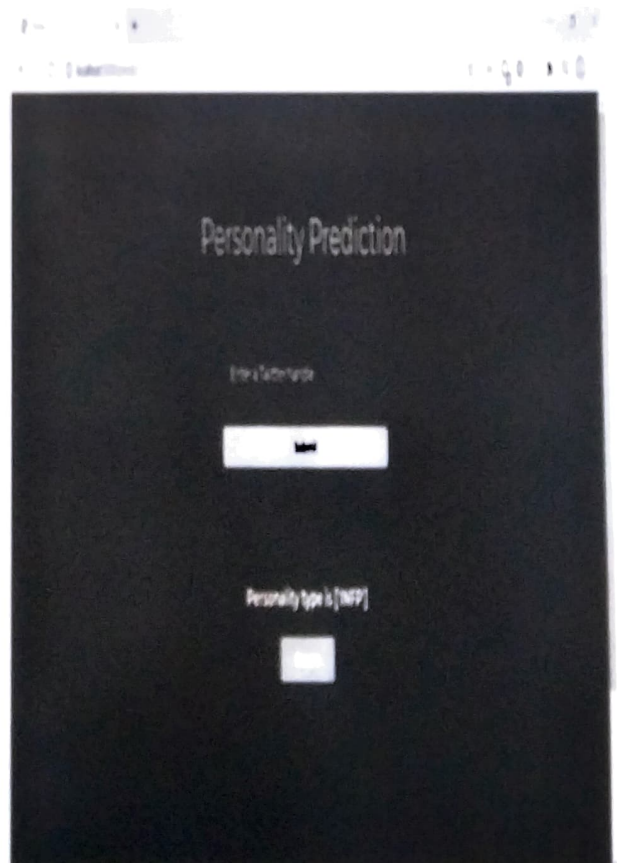


Fig 6 Twitter Handle prediction results

V. RESULTS

The Random Forest Classifier used in this system gave an accuracy of 64%. The Gradient Boosting Classifier used in the system gave an accuracy of 66%. When the two models were ensemble together using the Voting Classifier, the accuracy the resultant model yielded was around 65%.

```

=====
RandomForestClassifier
accuracy of train set: 1.0
accuracy of test set: 0.6409221902017291
Gradient Boosting Classifier
accuracy of train set: 0.9707492795389049
accuracy of test set: 0.6605187319884727
Ensembled Classifier
accuracy of train set: 0.9857348703170029
accuracy of test set: 0.652449567723343

```

Fig 7 System results

VI. CONCLUSION AND FUTURE WORK

A system has been developed that effectively predicts the personality of the user according to the Myers Briggs Type Indicator using random forest classifier and gradient boosting classifier. The classification was done in two ways – through a psychometric test or through the analysis of tweets. This system can be used in the recruitment process to identify the personality of the candidates and select them based on the requirements of the team.

Certain points of improvements on both technical and aesthetic fronts can be achieved in the future. The algorithms used in this project can be updated to better and more accurate ones which may come up as technology advances. Combinations of various algorithms can be studied to achieve highest accuracy. Furthermore, the dataset can be extended to improve the accuracy. The additional functionality of scraping Facebook comments or statuses to predict the personality can be added. Such features and improvements can make the system more efficient.

REFERENCES

[1] Madhura Jayaratne, Buddhi Jayatilleke, "Predicting Personality using Answers to Open Ended Interview Questions", IEEE ACCESS, June 22, 2020.

[2] Sandhya Katiar, Himdweep Walia, Sanjay Kumar, "Personality classification system using data mining", 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Amity University, Noida, India. June 4-5, 2020

[3] Yunan Li, Jun Wan, Qiguang Miao, Sergio Escalera, "CR-Net: A Deep Classification-Regression Network for Multimodal Apparent Personality Analysis", Springer 2018

[4] Aditi V. Kunte, Suja Panicker, "Using textual data for Personality Prediction: A Machine Learning Approach", (2019, IEEE)

[5] Yash Mehta, Navonil Majumder, Alexander Gelbukh, Eric Cambria, "Recent trends in deep learning based personality detection", Springer Nature B.V. 2019

[6] Michael M. Tadesse, Hongfei Lin, "Personality Predictions Based on User Behavior on the Facebook Social Media Platform", IEEE ACCESS, October 17, 2018.

[7] <https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76>

[8] <https://towardsdatascience.com/hyperparameter-tuning-the-random-forest-in-python-using-scikit-learn-28d2aa77dd74>

[9] <https://builtin.com/data-science/random-forest-algorithm>

[10] https://www.analyticsvidhya.com/blog/2016/02/complete-guide-parameter-tuning-gradient-boosting-gbm-python/?utm_source=blog&utm_medium=4-boosting-algorithms-machine-learning

[11] <https://analyticsindiamag.com/a-hands-on-guide-to-hybrid-ensemble-learning-models-with-python-code/>

[12] <https://towardsdatascience.com/optimizing-hyperparameters-in-random-forest-classification-ec7741f9d3f6>

[13] <https://towardsdatascience.com/how-voting-classifiers-work-flc8e41d30ff>

[14] <https://realpython.com/python-web-applications-with-flask-part-iii/>

[15] <https://betterprogramming.pub/build-test-and-deploy-an-interactive-flask-application-part-i-templates-53a7b0cbe760>

[16] <https://flask.palletsprojects.com/en/1.1.x/testing/>

[17] <https://docs.python.org/3/library/unittest.html#unittest.TestCase.assertIn>

[18] <https://towardsdatascience.com/vectorization-implementation-in-machine-learning-ca652920c55d>

[19] https://www.w3schools.com/w3css/w3css_colors.asp

[20] <https://stackabuse.com/gradient-boosting-classifiers-in-python-with-scikit-learn/>

[21] https://www.kaggle.com/datasnaek/mbti-type?select=mbti_1.csv

LOST CHILD RECOVERY SYSTEM USING FACE RECOGNITION

MADHUVAN.K^[1],NANDHINI.R^[2] ,NANDHINI.R.G^[3],SUMITHRA.M^[4].

[1],[2],[3]-STUDENTS [4]-ASSISTANT PROFESSOR MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE

Abstract-In India a countless number of children are reported missing every year. Among the missing child cases a large percentage of children remain untraced. This project is used to recover the lost child to their parents by displaying their information. Its been tedious to manually search the documents and get the information of the child. This project aims to reduce the time to search the details of the child whenever the child is been reported to the police station by a stranger.

This project avoid the traditional method of training the dataset .Normally,training of images needs hundreds of picture for single child which ends up in occupying large space.This project optimises the space by using special face recognition extraction feature which normally requires one or two images to classify a child.

Keywords— *Missing child identification, face recognition,dlib,cmake,svm*

1.INTRODUCTION

When a child is found, the photograph at that time is matched against the images uploaded by the Police/guardian at the time of missing. Sometimes the child has been missing for a long time. This age gap reflects in the images since aging affects the shape of the face and texture of the skin. The feature discriminator invariant to aging effects has to be derived.

This is the challenge in missing child identification compared to the other face recognition systems. Also facial appearance of child can vary due to changes in pose, orientation, illumination, occlusions, noise in background etc.

The image taken by public may not be of good quality ,as some of them may be captured from a distance without the knowledge of the child. Each face image corresponds to a child and child face recognition is considered as an image category classification problem. The task is to classify input image uploaded by the public into one of the given category based on the image representation. Basically CNN architecture consists of computational layers for feature extraction and a classifier layer at the final stage. The VGG-face CNN model employs the softmax activation function for labeled class prediction, suggesting the class each image belongs to. The softmax in the CNN layers is replaced with a multi class SVM trained with feature vector array from each image. One-versus-rest linear SVM classifier is used and is trained on the dataset. Extracted feature vector array is used to train this classifier

II.RELATED WORKS

In[2],The project is about identifying a specific person by detecting their faces from images or videos.In this paper the accuracy and performance of three conventional neural network such as VCC.

In [7],Here the problem of classification within a medical image dataset based on a feature vector extracted. from the deepest layer of preprocessed CNN .This uses feature vector from several trained structures.

In[8],This system consist of two parts entry and exit monitoring .The basic idea introduced each child using real time face detection algorithm.

III WORKFLOW OF FACE RECOGNITION

Here we propose a system to identify the missing child by using the facial features. To easily identify the face in a given image, we are using dlib to locate the face on a given image. While the 68-point detector localizes regions along the eyes, eyebrows, nose, mouth, and jaw line, the 5-point facial landmark detector reduces this information to 2 points for the left eye,2 points for the right eye and 1 point for the nose. The dlib is used to find the face in the given image.

For the smooth working of dlib,we used cmake for correctly identifying the face. CMake is used to control the software compilation process using simple platform and compiler independent configuration files, and generate native make files and workspaces that can be used in the compiler environment of your choice.

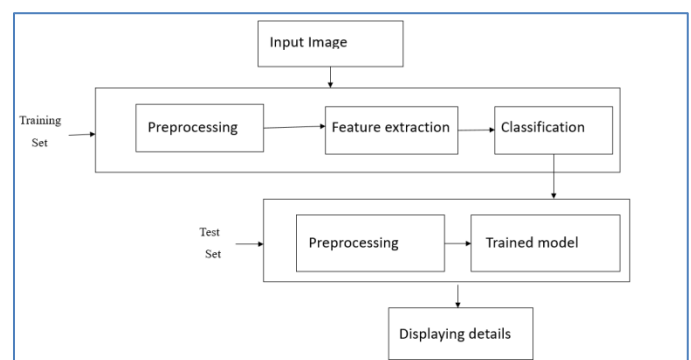


Fig 1. Architecture of lost child recovery system

In this system we will have a database that consist of missing's children images and information reported by the parents at the time of missing. Face recognition has been helpful in identifying the child with less number of data than the conventional method of training images. This system requires one or two images for each child which reduces space at a greater extent. When the child is been found by a stranger, they would report the child to the police station. Here the software enables the user to upload the picture of the child. If a matching is found in the repository, the system displays the most matched photo. Along with the image of the child, it displays the details of the child's parents like phone number and address which would be further used to contact the child's parent.

IV. SUPPORT VECTOR MACHINE

Support Vector Machine is a supervised classification algorithm where we draw a line between two different categories to differentiate between them. SVM is also known as the support vector network. In that case, the hyperplane dimension needs to be changed from 1 dimension to the Nth dimension. This is called Kernel. To be more simple, its the functional relationship between the two observations. It will add more dimensions to the data so we can easily differentiate among them.

Support vector machines (SVMs) are formulated to solve a classical two class pattern recognition problem. We adapt SVM to face recognition by modifying the interpretation of the output of a SVM classifier and devising a representation of facial images that is concordant with a two class problem. Traditional SVM returns a binary value, the class of the object. To train our SVM algorithm, we formulate the problem in a difference space, which explicitly captures the dissimilarities between two facial images. This is a departure from traditional face space or view-based approaches, which encodes each facial image as a separate view of a face.

For non-linearly separable plane, data are input in an input space which cannot be separated with a linear hyper plane. So, we map all the points to feature space using 25 specific type of kernel, in order to separate the non-linear data on a linear plane. After separating the points in the feature space we can map the points back to the input space with a curvy hyper plane. The following figure demonstrates the data flow of SVM

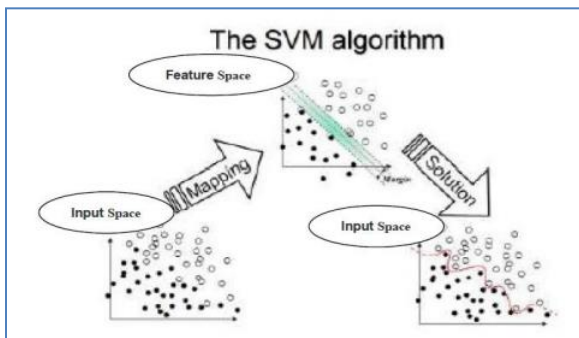


Fig 2. Support Vector Machine algorithm.

V. PREPROCESSING

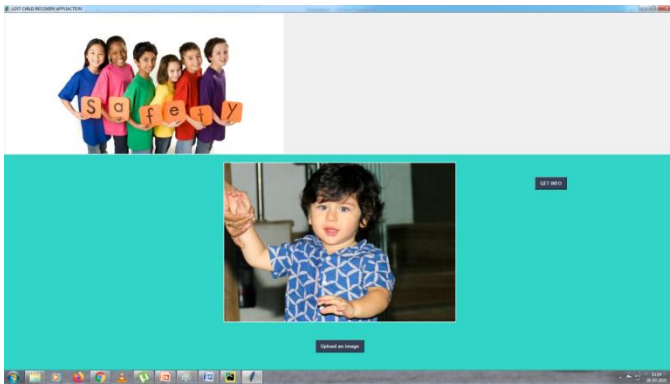
Detecting facial landmarks is a *subset* of the shape prediction problem. Given an input image (and normally an ROI that specifies the object of interest), a shape predictor attempts to localize points of interest along the shape. In the context of facial landmarks, goal is detect important facial structures on the face using shape prediction methods. A training set of labelled facial landmarks on an image. These images are manually labelled, specifying **specific** (x, y) -coordinates of regions surrounding each facial structure. Priors, of more specifically, the probability on distance between pairs of input pixels. Given this training data, an ensemble of regression trees are trained to estimate the facial landmark positions directly from the pixel intensities themselves (i.e., no "feature extraction" is taking place).

VI. DLIB'S FACIAL LANDMARK DETECTOR

The pre-trained facial landmark detector inside the dlib library is used to estimate the location of 68 (x, y) -coordinates that map to facial structures on the face. These annotations are part of the 68 point iBUG 300-W dataset which the dlib facial landmark predictor was trained on. It's important to note that other flavors of facial landmark detectors exist, including the 194 point model that can be trained on the HELEN dataset. Regardless of which dataset is used, the same dlib framework can be leveraged to train a shape predictor on the input training data — this is useful if you would like to train facial landmark detectors or custom shape predictors of your own.

VII. RESULTS AND DISCUSSIONS

The face identification algorithm is implemented using PYCHARM platform. The experiments are carried on Microsoft Windows 7, 64 bit Operating System with Intel core i7, 3.60GHz processors having 32GB RAM. For dealing with CNN architectures additional processing capability is needed. The user defined database includes child face images with unique children cases. Training and test set is prepared by splitting the database images. The training set and validation set consists of images of each child in the earlier days and testing is done with images of children after an age gap to evaluate the system in all conditions. CNN implemented with deep integration of CNN building blocks. The training set images are preprocessed to the size specified by the CNN architecture before passing to the CNN model. The face region is cropped from every image of the acquired input database. The images are fed. The activations to the input image produced by the first fully connected layer of the VGG-Face network architecture is taken as the CNN Feature descriptor. The normalized feature vector is used for training the SVM classifier for classifying the image of face and recognizes the child. Face identification accuracy is computed as the ratio of correctly identified face images to the total number of child face images in the test set. The computed recognition accuracy of the multi class SVM using learned features from CNN is 99.41%



VIII. REFERENCES

1. Deepak, R.U. and Pournami, S. (2018) 'Missing Child Identification System using Deep Learning and Multiclass SVM', IEEE Recent Advances in Intelligent Computational Systems (RAICS).
2. Deniz, O. and Torre, F.D. (2011) 'Face recognition using histograms of oriented gradients', Pattern Recognition Letters, 32(12):1598–1603.
3. Geng, C. and Jiang, X. (2009) 'Face recognition using sift features', IEEE International Conference on Image Processing (ICIP).
4. LeCun, Y. and Hinton, G. (2015) 'Deep learning', Nature, 521(7553):436–444.
5. Mei Wang, and Weihong Deng, (2019) 'Deep Face Recognition: A Survey', arXiv, pp. 1-3.
6. Parkhi, O.M. and Zisserman, A. (2015) 'Deep Face Recognition', in British Machine Vision Conference, vol. 1, no. 3, pp. 1-12.
7. Rohit Satle, Vishnuprasad Poojary, John Abraham, Shilpa Wakode, (2016) 'Missing child identification using face recognition system', International Journal of Advanced Engineering and Innovative Technology (IJAEIT), Volume 3.
8. Simonyan, Karen and Andrew Zisserman, (2015) 'Very deep convolutional networks for large-scale image recognition', International Conference on Learning Representations (ICLR).
9. Vedaldi, A. and Lenc, K. (2015) 'MatConvNet: Convolutional Neural Networks for MATLAB', ACM International Conference on Multimedia, Brisbane.

Self-diagnosis Medical Chat-bot Using Artificial Intelligence

PRAVEENKUMAR K
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
praveenkumar1106464@gmail.com

PRADEEPKUMAR A
Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
(Affiliated to Anna University)
Chennai, India
pradeepsgf04@gmail.com

Mrs. S YAMUNA
Assistant Professor, Dept. of Computer Science and
Engineering Meenakshi Sundararajan Engineering
College (Affiliated to Anna University)
Chennai, India

Abstract—By all the aspects, 2016 was the year of chat bots. Some of the critics argued that it will replace the use of websites and apps. But it still has a long way to go. Chat bot or Intelligent Conversational Agent development using Artificial Intelligence or Machine Learning technique is an interesting problem in the field of Natural Language Processing. In many research and development projects, they are using Artificial Intelligence, Machine Learning algorithms and Natural Language Processing techniques for developing conversation/dialogue agent. Their research and development is still under progress and under experimentation. These virtual agents are adopted by businesses ranging from very small start-ups to large corporations. There are many chatbot development frameworks available in market both code based and interface based. But they lack the flexibility and usefulness in developing real dialogues. The proposed idea is to create a medical chatbot using Artificial Intelligence that can diagnose the disease and provide basic details about the disease before consulting a doctor. To reduce the healthcare costs and improve accessibility to medical knowledge the medical chatbot is built. Certain chatbots acts as a medical reference books, which helps the patient know more about their disease and helps to improve their health.

Keywords-dataset, natural language processing, artificial intelligence

I. INTRODUCTION

A chatbot is a piece of software that conducts a conversation with users via auditory or textual methods. A medical chatbot facilitates the job of a healthcare provider and helps improve their performance by interacting with users in a human-like way. There are

countless cases where intelligent medical chatbots could help physicians, nurses, therapists, patients, or their families.

II. RELATEDWORKS

A. ARTIFICIAL INTELLIGENCE:

Artificial intelligence (AI) is the ability of a computer program or a machine to think and learn. It is also a field of study which tries to make computers "smart". As machines become increasingly capable, mental facilities once thought to require intelligence are removed from the definition. AI is an area of computer sciences that emphasizes the creation of intelligent machines that work and reacts like humans. Some of the activities computers with artificial intelligence are designed for include: Face recognition, Learning, Planning, Decision making etc. Artificial intelligence is the use of computer science programming to imitate human thought and action by analyzing data and surroundings, solving or anticipating problems and learning or self-teaching to adapt to a variety of tasks.

B. NATURAL LANGUAGE PROCESSING:

Natural language processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language. NLP draws from many disciplines, including computer science and computational linguistics, in its pursuit to fill the gap between human communication and computer understanding. While natural language processing isn't a new science, the technology is rapidly advancing thanks to an increased interest in human-to-machine communications, plus an availability of big data, powerful computing and enhanced algorithms.

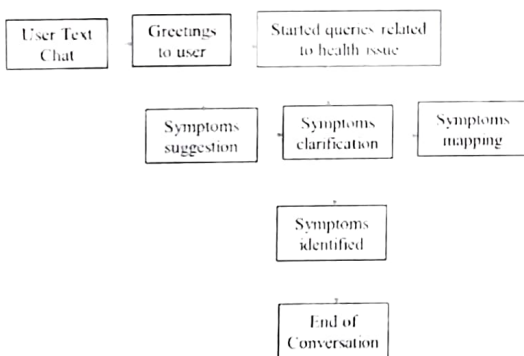
C. EXISTING SYSTEM:

MeS (Health On – Line Medical Suggestions), a novel recommendation system that leverages a chatbot to emulate a human physician in a clinical environment, in order to overcome the mentioned limitation of biased interaction between the user and the software. The cluster – computing facility is provided by Databricks where a cluster of servers allows cluster-computing over the Spark framework. The chat-bot application is implemented by the Watson Conversation Service, designed and trained via the Bluemix platform. It is designed to autonomously interact with the user by understanding natural language in a chat and acting as a human physician. It is made of different modules to provide several advanced eHealth services through an intuitive chat application.

D. PROBLEM STATEMENT:

Most of the government hospitals in india do not have doctors working at regular hours to attend patients. so we can use this chat bot that diagnose the disease and provide basic details before consulting a doctor.

III. SYSTEM ARCHITECTURE



This flow diagram shows that once user text chat is open, system will start the conversation through greetings. After that starting sequence of queries related to user's answers. System will give the symptoms suggestion to user for system clarification and automatically system will map the accurate symptoms. Then finally identified the disease whether it is severe or mild disease and given solution to that disease to user also given the doctor's contact details

IV .PROPOSED SYSTEM

In the proposed system the user dialogue is a linear design that proceeds from symptom extraction, to symptom mapping, where it identifies the corresponding symptom, then diagnosis the patient whether it's a major or minor disease and if it's a major one an appropriate doctor will be referred to the patient, the doctor details will be extracted from the database, the user will be identified by the login details which is stored in the database. Chatbot's dialogue design is represented using finite state graph. In order to achieve an accurate diagnosis, the logic for state transitions are made, natural language generation templates were used, and system initiative to the user and get responses from the user. Besides its greetings and

goodbye states, our agent has three main conversational phases: acquisition of basic information, symptom extraction, and diagnosis. Our bot starts off by asking about the user's email and password for login and then enters a loop of symptom extraction states until it acquires sufficient information for a diagnosis. Users have the option of entering the loop again to talk to the doctor about another set of symptoms after receiving their first diagnosis and another option is that the user can view their history of chats about what they have discussed.

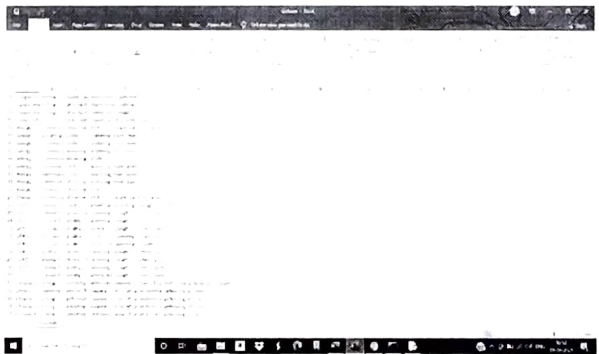
V. SYSTEM IMPLEMENTATION

The system here is divided into three modules as listed below:

- i. Dataset
- ii. Users
- iii Webchat
- iv Suggestion and clarification
- v Prediction

MODULE 1: DATASET COLLECTION

Datasets are collected from medical field which include kind of disease and related symptoms. And also, more added Doctors names, contact number recommended for severe stage of health problems. Dataset are collected and save in .xml file, which is stored in local drive.



MODULE 2: USERS

These are the users of the system. Users can be both patients and doctors. Patients can interact with the system to know about their health conditions, to book an appointment with any specialist, etc. Doctors can interact with the system to view their appointments for a particular date.

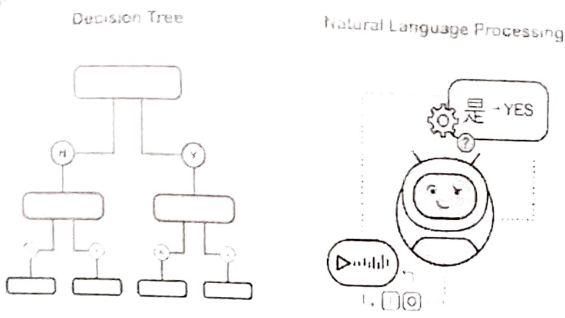
MODULE 3: WEB TEXT CHAT

This is the module where the actual chat takes place. This is the module through which the users interact with the system. All the information which are retrieved from the database are displayed here.

MODULE 4: SUGGESTION AND CLARIFICATION

This is the module which will process the queries of the user using string search algorithm. This module will analyse all the symptoms mentioned by the patient and classify whether it is a major

a minor disease. If it is a major disease, it will suggest a...
 for the patient to visit a. If it is a minor disease, it will
 suggest a suitable remedy.



has some cons such as sometimes it may give some irrelevant reply to the patient or sometimes it may happen that the bot may provide some solution that the user couldn't believe. Adding more quality data will further improve performance. Also, the training model should be trained with other hyper - parameters and different dataset for further experimentation. This was an attempt to experiment with Deep Neural Network for dialogue generation in order to develop intelligent chatbot. The efficiency of the chatbot can be improved by adding more combination of words and increasing the use of database so that of the medical chatbot could handle all types of diseases

MODULE 5: PREDICTION

The system will predict, what kind of disease you might have and given Doctors contact details for severe health problems otherwise suggest take rest.

- Person face recognition from given input image
- To capture the faces from given video for training purpose
- To build a model for training process
- Implementation of person faces re-identification by haar cascade classifier

FUTURE ENHANCEMENT:

In the future, the bot's symptom recognition and diagnosis performance could be greatly improved by adding support for more medical features, such as location, duration, and intensity of symptoms, and more detailed symptom description. The implementation of Personalized Medical assistant heavily relies on AI algorithms as well as the training data. At last, the implementation of personalized medicine would successfully save many lives and create a medical awareness among the people. As said before, the future era is the era of messaging app because people going to spend more time in messaging app than any other apps. Thus medical chatbot has wide and vast future scope. No matter how far people are, they can have this medical conversation. The only requirement they need is a simple desktop or smartphone with internet connection. The efficient of the chatbot can be improved by adding more combination of words and increasing the use of database so that of the medical chabot could handle all type of diseases. Even voice conversation can be added in the system to make it more easy to use.

VI. CONCLUSION AND FUTUREWORK

CONCLUSION:

From the review of various journals, it is concluded that, the usage of Chatbot is user friendly and can be used by any person who knows how to type in their own language in mobile app or desktop version. A medical chatbot provides personalized diagnoses based on symptoms. In the future, the bot's symptom recognition and diagnosis performance could be greatly improved by adding support for more medical features, such as location, duration, and intensity of symptoms, and more detailed symptom description. The implementation of Personalized Medical assistant heavily relies on AI algorithms as well as the training data. At last, the implementation of personalized medicine would successfully save many lives and create a medical awareness among the people. As said before, the future era is the era of messaging app because people going to spend more time in messaging app than any other apps. Thus medical chatbot has wide and vast future scope. No matter how far people are, they can have this medical conversation. Although it

VII. REFERENCES

1. Divya, S., V. Indumathi, S. Ishwarya, M. Priyasankari, and S. Kalpana Devi. "A self-diagnosis medical chatbot using artificial intelligence." *Journal of Web Development and Web Designing* 3, no. 1 (2018): 1-7. Chen, G. (2018). Nurturing the Companion Chat Bot.
2. Madhu, D., Jain, C. N., Sebastain, E., Shaji, S., &Ajayakumar, A. (2017, March). A novel approach for medical assistance using trained chatbot.
3. Honghao, W. E. I., Zhao, Y., &Ke, J. Building Chatbot with Emotions.
4. Mishra, S. K., Bharti, D., & Mishra, N. (2018). Dr.Vdoc: A Medical Chatbot that Acts as a Virtual Doctor. *Research & Reviews: Journal of Medical Science and Technology*, 6(3), 16-20.
5. Bulla, Chetan, Chinmay Parushetti, Akshata Teli, Samiksha Aski, and Sachin Koppad. "A Review of AI Based Medical Assistant Chatbot." *Research and Applications of Web Development and Design* 3, no. 2 (2020).

- Ayanouz, Soufyane, Boudhir Anouar Abdelhakim, and Mohammed Benhmed. "A smart chatbot architecture based NLP and machine learning for health care assistance." In *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*, pp. 1-6. 2020.
7. Vineeth, R., R. Rithish, D. V. S. N. Sai Varma, and B. V. Ajay Prakash. "Smart Health Care Chatbot for Prognosis of Treatments and Disease Diagnosis Using Machine Learning." *Journal of Computational and Theoretical Nanoscience* 17, no. 9-10 (2020): 3947-3951.
 8. Kalla, Dinesh, and Vatsalya Samiuddin. "Chatbot for medical treatment using NLTK Lib." *IOSR J. Comput. Eng* 22 (2020).
 9. Bumrungsena, Parichaya. "The Development of Traditional Thai Massage Chatbot Application for Hypertension Treatment." In *Advances in Human Factors and Ergonomics in Healthcare and Medical Devices: Proceedings of the AHFE 2020 Virtual Conference on Human Factors and Ergonomics in Healthcare and Medical Devices, July 16-20, 2020, USA*, vol. 1205, p. 246. Springer Nature, 2020.
 10. Kavitha, B. R., and Chethana R. Murthy. "Chatbot for healthcare system using Artificial Intelligence." *International Journal of Advance Research, Ideas and Innovations in Technology* ISSN 5, no. 3 (2019): 1304-1307.

E-VOTING OVER A PEER TO PEER NETWORK USING BLOCKCHAIN TECHNOLOGY

Dr.M.K.SANDHYA

Professor, Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
sandhyamk2007@gmail.com

S. PRATHIKSHA SRI

Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
prathikshasri20@gmail.com

T. RAMADEVI

Dept. of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
ramadevi0699@gmail.com

Abstract— This paper elaborates the implementation of blockchain technology in developing an electronic voting system. The usage of blockchain will ensure utmost transparency and security as compared to the traditional methods. The distributed ledger system is used to design a synchronized model for maintaining the voting records. This paper discusses the benefits of using blockchain to eliminate the forgery in votes in the voting process. Blockchain expresses its effectiveness by achieving a peer to peer verifiable voting system. An admin will contain all the details of the voting process. The voters will be able to cast their votes in a secure way. The votes will then be counted and displayed by the admin. The electronic voting system using blockchain will ensure security and prevent the tampering of votes.

Keywords— Blockchain, Distributed ledger system, Electronic voting system

I. INTRODUCTION

Voting is very much important for every citizen as it exercises his or her right. The entire voting process should be conducted with utmost security and the tampering of votes should be avoided. The goal is to minimise the costs and to provide full security during the voting procedure. Replacing the traditional methods with a brand new election system is important to make the voting process traceable and verifiable.

To overcome all the limitations in the traditional methods blockchain technology is used to store the records and to provide full efficiency and accuracy. Blockchain is one of the emerging technologies with strong cryptographic foundations enabling applications to leverage these abilities to achieve fullest security solutions. A Blockchain resembles a data structure which maintains and shares all the transactions being executed through its genesis. It is primarily a distributed decentralized database that maintains a complete list of constantly germinating and growing data records secured from unauthorized manipulating, tampering and revision.

Repository allows every user to connect to the network, send new transactions to it, verify transactions and create new blocks. Each block is assigned a cryptographic hash (which may also be treated as a finger print of the block) that remains valid as long as the data in the block is not altered. If any changes are made in the block, the cryptographic hash would change immediately indicating the change in the data which may be due to a malicious activity. Therefore, due to its strong foundations in cryptography, blockchain has been

increasingly used to mitigate against unauthorized transactions across various domains.

The content of this paper is divided into five sections. In the first section, the background is introduced along with the motivation and purpose of this paper. In the second section, the technology and documents that are related to this paper are mentioned. In the third section, the structure of the system built in this paper is explained. In the fourth section, the construction of the system is presented in detail. In the last section, conclusions are made along with future enhancement possibilities.

II. RELATED WORKS

A. BLOCKCHAIN:

Blockchain is a technical solution for storing, verifying, transmitting and communicating network data based on cryptography[3]. The core concept is to rely on cryptographic and mathematical decentralized algorithms. Blockchain can make participants reach a consensus without the intervention from the third party. This can solve the problem of lack of trust and unreliable value delivery. Through a public key, both of a private key, which are generated through asymmetric encryption, and the account address, we can perform a transaction or transmit data on the blockchain. The transaction on the blockchain is verified by every block on the blockchain instead of the third party.

After a block verifies the transaction addresses, transaction messages, and the source of cryptocurrency, it will push them to other nodes for verification. Once the transaction is confirmed by all nodes on the blockchain, the transaction data will be recorded in the block. The transaction record can't be changed and has full anonymity. Each transaction will produce a unique hash value, and each block header contains the previous block hash value, which connects all blocks and forms a chain.

Compared with the traditional financial transaction systems, which has to rely on trusted or guaranteed third-party organizations, the blockchain effectively implements the process of decentralization and combines the network to ensure that transactions are collectively maintained and verified by each node and that transaction records are confirmed[1].

B. ETHEREUM:

Ethereum is an open source platform and a public blockchain which functions on smart contracts. In addition to the usual cryptocurrency transactions[4].Ethereum features a decentralized Ethereum Virtual Machine (EVM) to deploy and apply smart contracts. In addition to mainly conducting cryptocurrency transactions, Ethereum also provides environmental constructions of private blockchains and multiple blockchains for testing, such as Rinkeby and Ropsten etc., which provide developers with different environments for testing and development.

C. SMART CONTRACTS:

A smart contract is a piece of code which can be written in Solidity, Serpent, etc., and run on the blockchain[7]. The smart contracts are helpful in eliminating the need of a third-party such as a bank which handles money transactions between two parties. Smart contracts permit trusted transactions and agreements to be carried out among disparate, anonymous parties without the need for a central authority.

D. REMIX IDE:

Remix IDE allows developing, deploying and administering smart contracts for Ethereum like blockchains[2][6]. It can also be used as a learning platform. Remix is a Solidity IDE that's used to write, compile and debug Solidity code. IDE stands for Integrated Development Environment and is an application with a set of tools designed to help programmers execute different tasks related to software development such as writing, compiling, executing and debugging code.

E. NODE JS:

Node js is a platform built on Chrome's JavaScript runtime for easily building fast and scalable network applications. Node. js uses an event-driven, non-blocking I/O model that makes it lightweight and efficient, perfect for data-intensive real-time applications that run across distributed devices[5]. All APIs of Node.js library are asynchronous, that is, non-blocking. It essentially means a Node.js based server never waits for an API to return data. The server moves to the next API after calling it and a notification mechanism of Events of Node.js helps the server to get a response from the previous API call. Being built on Google Chrome's V8 JavaScript Engine, Node.js library is very fast in code execution. No Buffering – Node.js applications never buffer any data. These applications simply output the data in chunks.

F. JQUERY:

JQuery is a fast and concise JavaScript Library created by John Resig in 2006 with a nice motto: Write less, do more. jQuery simplifies HTML document traversing, event handling, animating, and Ajax interactions for rapid web development[8]. jQuery is a JavaScript toolkit designed to simplify various tasks by writing less code. The jQuery made it easy to select DOM elements, negotiate them and modifying their content by using cross-browser open source selector engine called Sizzle. The jQuery offers an elegant way to capture a wide variety of events, such as a user clicking on a link, without the need to clutter the HTML code itself with event handlers. The jQuery helps you a lot to develop a responsive and feature rich site using AJAX technology. The jQuery comes with plenty of built-in animation effects which you can use in your websites. The jQuery is very lightweight library - about 19KB in size. The jQuery supports CSS3

selectors and basic XPath syntax.

III. SYSTEM ARCHITECTURE

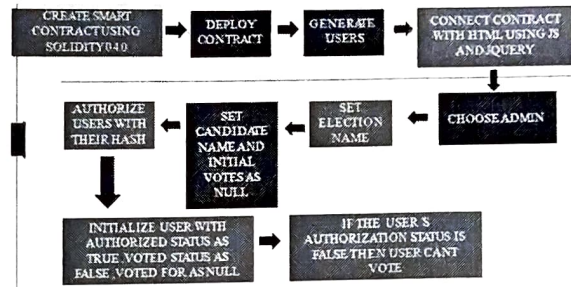


Fig 1. System Architecture

The system illustrates the implementation of blockchain in the electronic voting system. It is an efficient and cost-effective way for conducting a voting procedure, which has a characteristic of accepting unanimous data in real time and requesting high safety. We design a synchronized model of voting records based on distributed ledger technology to avoid forgery of votes. Firstly, a smart contract is created using Solidity 0.4.0. A contract is deployed and the requires users are generated. The smart contract created is connected with html using js and jquery to develop a portal. An admin is chosen and he/she will be having all the details of the entire election process. The required modules are set namely adding the name of the election and getting the information of the candidates and the voters. The initial votes are set as null depicting that the voting process has not started. The voters are authorized with a hash address and they can enter the portal only by giving their valid hash address. The voting status of the voters will be specified. Every voter can vote only once. If the voter's authorization process is false the voter cannot vote. The admin will then login into the portal and get the results of the election.

IV. SYSTEM IMPLEMENTATION

In the system construction, we use solidity to write and test the smart contracts. The smart contract is connected with html using js and jquery. We incorporate the Ethereum platform to implement this voting system by running the code in Remix IDE. The following are the modules used in the system and their associated processes.

A. CREATE ELECTION:

A module called create election is developed to create an election that will get the votes and help in the contesting of the election.

B. ADD CANDIDATE:

This module will add the candidates who are contesting the elections. Each candidate will be added only once.

C. ADD ELECTION NAME:

The election name will be displayed i.e, the candidates and voters will be part of the specified election.

D. GET AUTHORIZED VOTERS:

Only authorized voters can cast their votes and tampering of the votes will not take place. Voters will not be able to cast their votes more than once.

E. GET AUTHORIZED CANDIDATES

Only authorized candidates will be able to contest the elections.

F.DISPLAY LEADING CANDIDATE

The leading candidate will be displayed once everyone has cast their votes. This can be accessed by the admin.

The admin will first enter into the portal through his/her unique hash address. The following modules will be filled accordingly. The voters will then cast their votes for the required candidates through their unique hash address. The admin will then be able to see the results of the voting process.

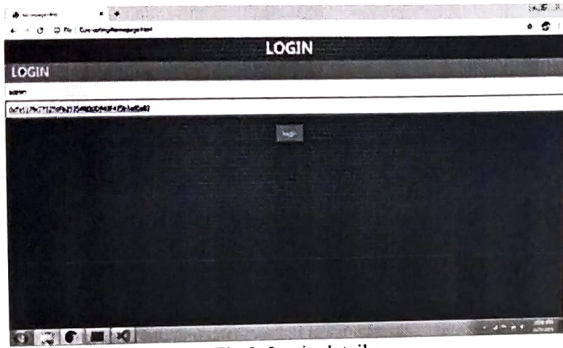


Fig 2. Login details

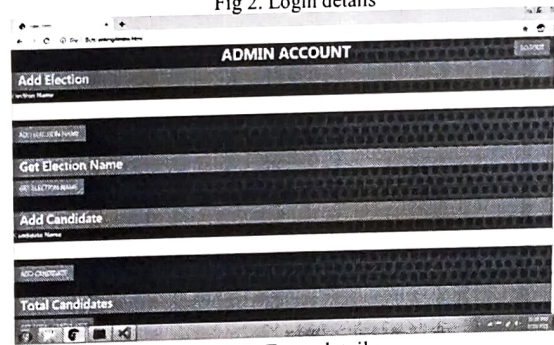


Fig 3. Enter details

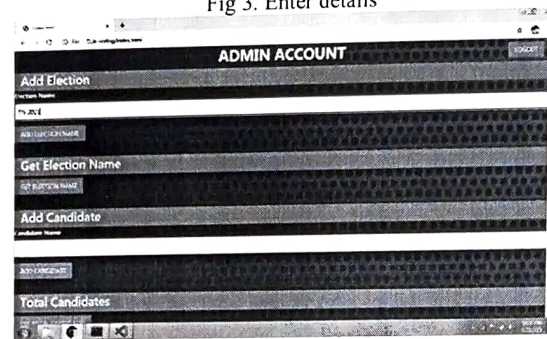


Fig 4. Add election name

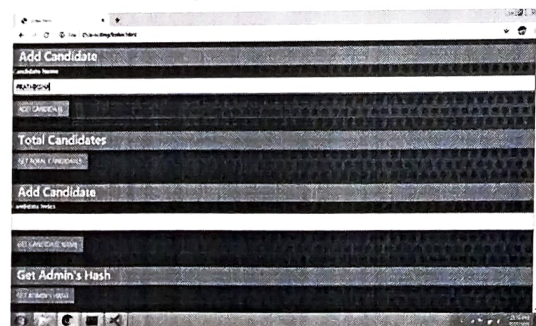


Fig 5. Add candidates

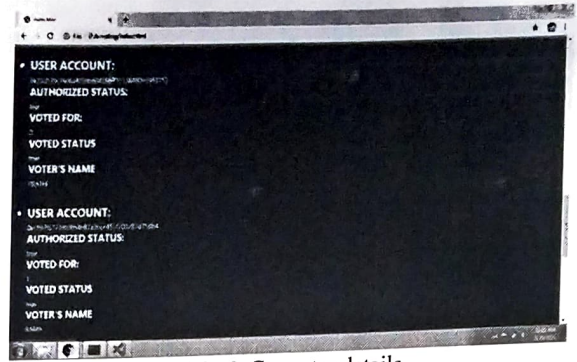


Fig 6. Get voter details

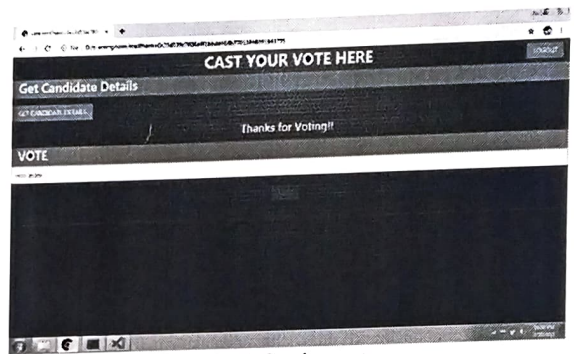


Fig 7. Casting votes

V. CONCLUSION AND FUTURE WORK

An electronic voting system is created by using blockchain technology that will ensure high levels of security and transparency. The blockchain technology will maintain the records that will be accessed by the admin easily and efficiently. The tampering of votes will be avoided and the entire process is faster and cost effective.

In the future, this voting system can be further enhanced by incorporating cryptographic algorithms that will provide more security.

REFERENCES

- [1]Amna Qureshi, David Megias, Helena Rifa Pous(2019),SeVEP:Secure and Verifiable Electronic Polling System, IEEE Access, Volume 7, pp. 19266-19290.
- [2]Feng Hao, Samiran Bag, Shen Wang, Rob Proctor(2020),'End-to-End Verifiable E-voting Trail for Polling Station Voting', IEEE Security and Privacy, Volume 8,pp. 2-9.
- [3] Kazi Md. Rokibul Alam, Shinsuke Tamura, S. M. Saifur Rahman, and Yasuhiko Morimoto(2020), 'An Electronic Voting Scheme Based on Revised-SVRM and Confirmation Numbers', IEEE Transactions on Dependable and Secure Computing, Volume 18,pp. 1-11.
- [4] Linh Vo-Cao- Thuy, Khoi Cao-Minh, Chuong Dang-Le-Bao, Tuan A. Nguyen(2019), 'Votereum: An Ethereum-Based E-Voting System' IEEE-RIVF International Conference on Computing and Communication Technologies, IEEE 2019, pp. 101-104.
- [5] Shiyao Gao, Dong Zheng, Rui Guo, Chunming Jing, Chencheng Hu(2019), An Anti-Quantum E- Voting Protocol in Blockchain with Audit Function, IEEE Access, Volume 7, pp. 1-13.

[6] Sudeepthi Komatenini, Gowtham Lingala(2020), Secured E-voting System Using Two- factor Biometric Authentication, Fourth International Conference on Computing Methodologies and Communication, IEEE 2020, Volume 20, pp. 245-248.

[7] Syada Tasmia Alvi, Mohammed Nasir Uddin, Linta Islam(2020), ' Digital Voting: A Blockchain-based E-Voting System using Biohash and Smart Contract', Third International Conference on Smart Systems and Inventive Technology , IEEE 2020, pp. 228-233.

[8] S.K.Vivek, R.S.Yashank, Yashas Prashanth, N. Yashas, M. Namratha(2020) 'E-voting systems using blockchain', Second International Conference on Inventive Research in Computing Applications, IEEE 2020, pp. 890-895.

FUTURE PREDICTION OF COVID - 19 USING MACHINE LEARNING MODELS

SELINI MAGNIDA

Dept of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
seini.magnida@gmail.com

JITHUWYA FERNANDEZ

Dept of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
jithuwyafernandez@gmail.com

R SAKTHI VIGNESHWARI

Dept of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
sakthi.r1619@gmail.com

Mrs P Revathi

Assistant Professor, Dept of Computer Science and Engineering
Meenakshi Sundararajan Engineering College
Chennai, India
revathisprajothi@gmail.com

Abstract— In this project, we plan to conjecture the conceivable number of positive instances of Coronavirus utilizing different AI calculations. The expectation will be finished utilizing the accompanying boundaries such as the quantity of recuperated cases, every day positive cases, also the number of perished cases. An information driven anticipating technique has been utilized to assess the conceivable number of positive instances of COVID-19.

Keywords: Linear regression, Least absolute shrinkage and selection operator, Support vector machine, Exponential smoothing

I. INTRODUCTION

AI (ML) based determining components have demonstrated their importance to expect in perioperative results to improve the dynamic on the future course of activities. The ML models have for quite some time been utilized in numerous application areas which required the recognizable proof and prioritization of unfriendly factors for a danger. AI (ML) based determining components have demonstrated their importance to expect in perioperative results to improve the dynamic on the future course of activities. The ML models have for quite some time been utilized in numerous application areas which required the recognizable proof and prioritization of unfriendly factors for a danger.

This examination plans to give an early estimate model to spread of novel Covid, World Health Organization (WHO) Coronavirus is as of now an intense danger to was human existence everywhere on the world. Toward the finish of 2019, the infection was first distinguished in a city of China considered Wuhan, when an enormous number of individuals created side effects. It diversely affects the human body, including extreme intense respiratory condition and fever which can at last prompt passing in an exceptionally brief length. Hundreds of thousands of individuals are influenced by this pandemic all through the world with a large number of death serious coming day. To with contributing with the current human emergency our endeavor in this examination is to build up an anticipating framework for COVID-19. The estimating is accomplished for the three significant factors of the illness for the coming days taking into account the number of positive, death and recovered cases. This issue of estimating has been considered as a relapse issue in this examination, so the investigation is based on ML models such as linear regression (LR), least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and exponential smoothing (ES).

The substance of this paper is isolated into five sections. In the first section, the foundation is presented alongside the inspiration and motivation behind this paper. In the second section, the innovation and reports that are identified with this paper are mentioned. In the third section, the design of the framework worked in this paper is explained. In the fourth section, the development of the framework is introduced in detail. In the last section, conclusions are made alongside future upgrade prospects.

II. RELATED WORKS

MATERIALS AND METHODS:

Dataset:

The goal of this research is to forecast COVID-19 'potential distribution with an emphasis the number of positive new events, mortality and recoveries. Data from Kaggle was collected.

Supervised Machine learning models:

For the situation that a vague information occasion is given, a regulated learning model can foresee

The regressor is used for the regression model. A forecast for unpredictable entrants or test data is then generated in the qualified model. For the development of predictive models, Regression techniques and classification algorithms study method used here. In this COVID 19 prediction analysis, regression models of four are used.

- Linear Regression
- LASSO Regression
- Support Vector Machine
- Exponential Smoothing

Linear Regression

The aim class concentrates on individual regression simulation characteristics. It should even be went to define and model the connection between independent variables and dependent. The foremost useful computer method for mathematical analysis of the machine learn is linear regression type regression simulation. A rectilinear regression observation relies on two values, one on the dependence and one on the isolation. Linear regression defines a linear relation between these variables' dependency and independence.

Equations (1) and (2) are necessary for the linear regression search. Equation (1) indicates how y is associated with x, which is named regression.

$$y = \beta_0 + \beta_1 x + \epsilon \quad (1)$$

$$E(y) = \beta_0 + \beta_1 x \quad (2)$$

This is the linear term for error regression. This error term takes into consideration the variability between x and y, β_0 is that the y-intercept and β_1 is that the pitch.

A class mark is per the computer file set for the aim of the model x training of the linear regression within the context of machine study. The aim is to search out the optimum values for β_0 (intercept) and β_1 (coefficient) to induce the most effective curve. The difference between the particular values and also the values predicted should be minimum to form sure that this minimising problem is presented:

$$\text{Minimize } \frac{1}{n} \sum_{i=1}^n (\text{pred}_i - y_i)^2$$

here, g, which is the mean root square of the expected value for y (pred_i) and y (y_i), n is the cumulative number of data points. g is called the cost function.

Lasso:

LASSO may be a regression model that's a part of a linear regression method that uses shrinkage. Shrinking means reducing the acute data sample values to the key values during this case. This strengthens and stabilises LASSO and reduces the error by the shrinking process. For multi-linear situations, LASSO is taken into account a more fitting model. LASSO thus makes the regression smoother in terms of the quantity of functions it uses. It uses a style of regularisation to penalise additional tasks automatically.

However, the LASSO regression tries one at a time, because it doesn't add importance of zero if the new function wouldn't boost the penalty term's fit therewith function. the facility of regularisation is therefore to automatically pick for us by adding the penalty for extra functions. Therefore, during this case of regularisation the models become sparse with few coefficients because the method removes values are zero. This regression LASSO acts to scale back the coefficient, which may be known by the square residual β slope), where, β slope could be a concept of penalty.

Support Vector Machine:

The SVM may be a sort of ML managed algorithm for reverse and regression classification. The SVM regression depends on a range of statistical functions as a non parametric technician. The set of the kernel function converts data input into the shape you wish. so as to beat regression problems employing a linear function, SVM maps the vector(x) input(s) within the n-dimensional space called the function space(z) when coping with non-linear regression problems. After linear regression is implemented within the space, non-linear mapping techniques are used for this mapping. Put the concept into an ML context employing a number of observations with y from a multivariate training dataset (x) to N. The goal is therefore that the worth of f(x) with (β_0, β_1) because the minimum standard values is found as flat as possible. The dilemma then blends in with the minimization function. If the worth of all residues isn't greater than p, as within the following equation:

Predictions are rendered supported data from previous times of exponential family smoothing techniques. As previous data findings grow old, their effect declines exponentially [11, 12]. The weights are therefore geometrically reduced to the various lag values. Particularly for uni-variate may be a statistic provision. Ft 1 is that the previous

The prediction is as follows. Ft 1 is that the real value within the preceding time-frame, during which Ft 1 is that the first moment within the prediction.

Evaluation Parameters:

The performance of every learning model is evaluated within the R2 scoring, Modified R-Square (R2), MSE, Mean Absolute Error (MAE) and Root Means Square Error (RMSE). during this analysis, the performance of each learning model is evaluated.

R-squared score:

The R-square (R2) score for effective regression models may be a statistical measure. These figures show the share of difference between the dependent and also the variable quantity. it's important to simply quantify 0-100 percent between the dependent differ and therefore the regression model. We may verify the fitness of the trained models with the R2 scoring after the regression models. The R2 scoring reveals that the info points round the regression are spread, often mentioned because the dissuasive component. Its ranking also varies between 0% and 100%.

The response variable is shown by 0 per cent as a results of the mean explained by the model, and 100 per cent as a median uncertainty of the solution variable. The high R2 value represents the consistency of the model learned. R2 may be a linear formula describing the proportion of disparities.

Adjusted r-squared score:

A transformed R2 shape is that the Adjusted R-square (R2). This shows, like R2, the quantity of features within the forecast that the latter will adjust. The minimum normal value of f(x) with (β_0, β_1) . But this can be the trend. For R2, the amount of latest features has grown. this can cause an improvement if the new functions for the prediction model are helpful. However, once the new characteristics are added, their value would plummet. The sample size is n and also the sum of independent regression parameters is k.

Mean Absolute error (mae):

This is the common of the test results between the model projections and also the observed statistics of equal weight for all variations. The matrix range ranges from zero to endlessness, and fewer scores demonstrate the goodness of learning models, so it's also brought up as negative scores

Mean square error (mse):

Another way of calculating regressive models output could be a medium-square error. MSE takes and squares the regression line data points. Squared is critical if the negative sign of the worth is omitted and greater weight is given to greater variations. The lower the medium defect, the closer you discover the higher match.

Root mean square error (rmse):

A standard deviation from the expected error is also set to RMSE. The source applies to a square failure. forecast error are notorious as residue because the space among finest match lines and real data points. Thus, RMSE tests the simplest fitness of the particular data points. RMSE is For the MSE root, this can be the subsequent error rate.

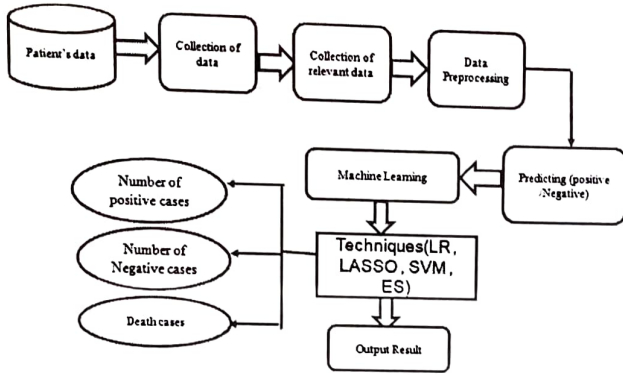
III.SYSTEM ARCHITECTURE

The patient's data set is taken from kaggle. Only the relevant data is processed further. These data sets are processed using the following machine language techniques such as Linear Regression(LR) Algorithm, Least absolute shrinkage and selection operator(LASSO) Algorithm, Support vector machine(SVM) Algorithm and Exponential smoothing(ES) Algorithm. Based on the three important factors that

number of positive cases, the number of negative cases and the number of death cases, the output is generated predicting the future of Covid-19 cases.

IV. SYSTEM IMPLEMENTATION

In the framework development, we use java as the front end of the system. The software requirement specification is created at the end of the analysis task. The function and performance allocated to software as part of system engineering are developed by establishing a complete information report as functional representation, a representation of system behavior, an indication of performance requirements and design constraints, appropriate validation criteria.



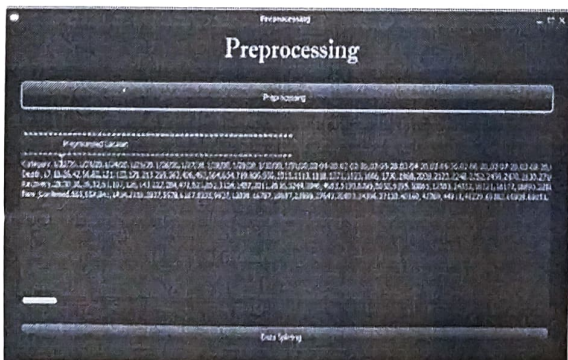
The first step is to analyse the dataset to find out whether they are relevant and the correct input for the system. Then they are processed by the four machine learning techniques such as Linear Regression(LR) Algorithm, Least absolute shrinkage and selection operator(LASSO) Algorithm, Support vector machine(SVM) Algorithm and Exponential smoothing(ES) Algorithm.

Secondly, based on the number of positive cases, the number of negative cases and the number of death cases, the output is generated predicting the future of Covid-19 cases.

Finally, an estimated comparison is made to find the accuracy from the four machine learning techniques. These are the modules present in the framework and its related cycles.

A. ESTIMATION PROCESS:

In different control stages, the Basic reproduction number changes greatly and it affects the intensity of control directly. In addition, the incubation period of the virus affects the speed of transmission directly. These two parameters need to be estimated. Current literature shows that the uncontrolled Basic reproduction. Therefore, we chose the valuation range in the corresponding range. For the controlled Basic reproduction number, the range of valuation was selected in the range of [0, 1.5].

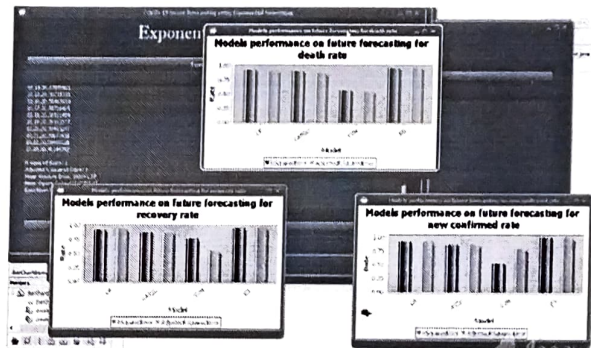


B. DATA-DRIVEN METHODS TO PREDICT COVID-19:

The data has been used (when the first case of COVID-19 was reported in India) with 80% data is used for training and rest 20% for forecasting and validation purposes. The resulting plot showing the total number of confirmed cases, the observed data is the data used for training purposes, official data (green line) indicates the official data available and forecasted data indicates the forecast of a total number of confirmed cases. From this graph, it is observed that the forecasted number of total confirmed positive cases closely matches with the available official data.

C. PREDICTION OF ACCURACY:

This technique is suitable to use predictive neural networks or characteristic data such as infection event or non-event binomial effects. The prediction accuracy of various measurements can be used for different purposes. They include the rate at which normal (non-predicted prediction correctly predicts sensitivity (non-infectious disease), accuracy (predicted percentage of predicted trend), positive predictive value, negative predictive value (correctly predicted infection rate is), the ratio is Expected predictions are a measure of the likelihood that the increase in the entire process exceeds the accuracy of the individual)



V. CONCLUSION AND FUTURE WORK

A data-driven forecasting/estimation method has been used to estimate the possible number of positive cases of COVID-19. The number of recovered cases, daily positive cases, and deceased cases has also been estimated by using and curve fitting. The effect of preventing measures as social isolation and lockdown has also been observed which shows that by these preventive measures, the spread of the virus can be reduced significantly.

Generally speaking we infer that model expectations as per the current situation are right which might be useful to comprehend the forthcoming circumstance. The examination figures in this manner can likewise be of extraordinary assistance for the specialists to take opportune activities what's more, settle on choices to contain the COVID-19 emergency. This study will be upgraded persistently later on course, next we intend to investigate the forecast philosophy utilizing the refreshed dataset and utilize the most exact and suitable ML strategies for estimating. Constant live estimating will be one of the essential concentrations in our future work.

Reference:

- [1] 1. Alsaecdy, A. A. R., & Chong, E. (2020). Detecting Regions at Risk for Spreading COVID-19 Using Existing Cellular Wireless Network Functionalities. *IEEE Open Journal of Engineering in Medicine and Biology*, 1-1.
- [2] 2. Sear, R. F., Velasquez, N., Leahy, R., Restrepo, N. J., El Oud, S., Gabriel, N., ... Johnson, N. F. (2020). Quantifying COVID-19 content in the online health opinion war using machine learning. *IEEE Access*, 1-1.
- [3] 3. Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X. ... Yang, G. (2020). Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification from CT Images. *IEEE Access*, 1-1.
- [4] 4. Zhang, Y., Li, Y., Yang, B., Zheng, X., & Chen, M. (2020). Risk Assessment of COVID-19 Based On

- [6] Abdel-Basset, M., Mohamed, R., Elhoseny, M., Chakraborty, R. K., & Ryan, M. (2020). A hybrid COVID-19 detection model using an improved marine predator's algorithm and a ranking-based diversity reduction strategy. *IEEE Access*, 1-1.
- [7] F. Petropoulos and S. Makridakis, "Forecasting the novel coronavirus covid-19," *Plos one*, vol. 15, no. 3, p. e0231236, 2020.
- [8] G. Grasselli, A. Pesenti, and M. Cecconi, "Critical care utilization for the covid-19 outbreak in lombardy, italy: early experience and forecast during an emergency response," *Jama*, 2020.
- [9] C. P. E. R. E. Novel et al., "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) in china," *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi*, vol. 41, no. 2, p. 145, 2020.
- [10] Y. Grushka-Cockayne and V. R. R. Jose, "Combining prediction intervals in the m4 competition," *International Journal of Forecasting*, vol. 36, no. 1, pp. 178-185, 2020.
- [11] N. C. Mediaite. Harvard professor sounds alarm on 'likely' coronavirus pandemic: 40% to 70% of world could be infected this year. Accessed on 2020.02.18. Available: <https://www.mediaite.com/news/harvardprofessor-sounds-alarm-on-likely-coronavirus-pandemic-40-to-70-ofworld-could-be-infected-this-year/>



ISSN No. : 2321-9653

IJRASET

**International Journal for Research in Applied
Science & Engineering Technology**

IJRASET is indexed with Crossref for DOI-DOI : 10.22214

Website : www.ijraset.com, E-mail : ijraset@gmail.com

Certificate

*It is here by certified that the paper ID : IJRASET34666, entitled
An IoT based Weather Monitor System for Drying Clothes*

*by
Sundari*

*after review is found suitable and has been published in
Volume 9, Issue V, May 2021*

*in
International Journal for Research in Applied Science &
Engineering Technology*

Good luck for your future endeavors

By

Editor in Chief, IJRASET



ISRA Journal Impact
Factor: 7.429



INDEX COPERNICUS
45.98



THOMSON REUTERS
Researcher ID: N568172016



TOGETHER WE REACH THE GOAL
SJIF 7.429





SRM TRP
ENGINEERING COLLEGE



CERTIFICATE OF PARTICIPATION

This is to certify that **Dr.A.KANIMOZHI** of **Meenakshi Sundararajan Engineering College** has presented a Paper titled **Vehicle Collision Avoidance System using ESP8266 and ThingSpeak** in the International Conference on Artificial Intelligence for IOT and Sustainable Electrical Networks (ICAIISEN'21) conducted by the SRM TRP Engineering College in association with IEEE SB, IIC, IETE and CSI on 29.04.2021.

B. Ramasubramanian

Dr.B.RAMASUBRAMANIAN
VICE PRINCIPAL &HOD/ECE

P. Sudhakaran

Dr.P.SUDHAKARAN
HOD/CSE

P. Elango

Dr.P.ELANGOVAN
HOD/EEE

B. Ganesh Babu

Dr.B.GANESH BABU
PRINCIPAL



LOYOLA INSTITUTE OF TECHNOLOGY

(Approved by AICTE Affiliated to Anna University, Chennai)

INTERNATIONAL VIRTUAL CONFERENCE ON EMERGING TRENDS ON ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0



CERTIFICATE OF APPRECIATION

This is to certify that DEVI LAKSHMI G S of
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE has presented a paper on AUTOMATED STAFF ATTENDANCE
MONITORING SYSTEM USING BIOMETRIC AUTHENTICATION

at *“International Virtual Conference on Emerging Trends on Artificial Intelligence In Industry 4.0”* in Association with Open POWER & IBM at *Loyola Institute of Technology*, Chennai on 21.06.2021.

Convenor
Mrs.G.Indhumathi

Principal
Dr. Sujatha Anand

Correspondent
Rev.Sr.M.K.Teresa

Dr.SIVANTHI ADITANAR COLLEGE OF ENGINEERING TIRUCHEDUR-628215



(Accredited by NBA(CSE,EEE), New Delhi & ISO 9001:2015 Certified Institution)



NATIONAL e-CONFERENCE ON ADVANCED TRENDS IN COMPUTER SCIENCE AND ENGINEERING "NCATCSE 2K21"

Organized By

Department of Computer Science and Engineering
Computer Society of India (CSI) Students' Chapter
(Institute Membership No.: I00859)



CERTIFICATE OF PRESENTATION

This is to certify that ~~Mr./Ms./Dr.~~ DEVI LAKSHMI G S
of Meenakshi Sundararajan Engineering College, Chennai has presented
a paper titled A Web Based College Online Forum

_____ in the "National
e-Conference on Advanced Trends in Computer Science and Engineering (NCATCSE'2K21)" on
31st March, 2021 organized by the Department of Computer Science and Engineering and
Computer Society of India (CSI) Students' Chapter (Institute Membership No.: I00859).

CO-ORDINATORS

HOD/CSE

PRINCIPAL



KG COLLEGE OF ARTS AND SCIENCE
Affiliated to Bharathiar University, Accredited by NAAC
KGiSL Campus, Saravanampatti, Thudiyalur Road,
Coimbatore-641035
in association with
Universiti Malaya



Organizes
TWO-DAY VIRTUAL INTERNATIONAL CONFERENCE
ON

INTERDISCIPLINARY RESEARCH IN SOCIAL SCIENCES, SCIENCES AND TECHNOLOGY

Certificate of Presentation

This is to certify that **Ms. R. Selvi** of Meenakshi Sundararajan Engineering College has participated and presented a paper entitled **Person Face Re-identification Using Deep Learning** in Two-Day Virtual ICIRSSST 2021 in association with Universiti Malaya held on 21.05.2021 and 22.05.2021 organized by KG College of Arts and Science, Coimbatore, Tamilnadu, India.

Dr. B. Adalarasu
Convener
KGCAS

Dr. Shivakumara Palaiahnakote
Convener
Universiti Malaya

Dr. Mohd. Yamani Idna Idris
Deputy Dean (R&D), FSKTM
Universiti Malaya

Dr. J. Rathinamala
Principal 89
KGCAS



KG COLLEGE OF ARTS AND SCIENCE
Affiliated to Bharathiar University, Accredited by NAAC
KGiSL Campus, Saravanampatti, Thudiyalur Road,
Coimbatore-641035
in association with
Universiti Malaya



**UNIVERSITI
MALAYA**
Fakulti Sains Komputer
dan Teknologi Maklumat

Organizes
TWO-DAY VIRTUAL INTERNATIONAL CONFERENCE
ON
INTERDISCIPLINARY RESEARCH IN SOCIAL SCIENCES, SCIENCES AND TECHNOLOGY
Certificate of Presentation

This is to certify that **Mr. SRIRAM K.P** of Meenakshi Sundararajan Engineering College has participated and presented a paper entitled **Impacting Online Classes With AR Based EDUCATION** in Two-Day Virtual ICIRSSST 2021 in association with Universiti Malaya held on 21.05.2021 and 22.05.2021 organized by KG College of Arts and Science, Coimbatore, Tamilnadu, India.

Dr. B. Adalarasu
Convener
KGCAS

Dr. Shivakumara Palaiahnakote
Convener
Universiti Malaya

Dr. Mohd. Yamani Idna Idris
Deputy Dean (R&D), FSKTM
Universiti Malaya

Dr. J. Rathinamala
Principal
KGCAS



LOYOLA INSTITUTE OF TECHNOLOGY

(Approved by AICTE Affiliated to Anna University, Chennai)

INTERNATIONAL VIRTUAL CONFERENCE ON EMERGING TRENDS ON ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0



CERTIFICATE OF APPRECIATION

This is to certify that _____ K.P.SRIRAM _____ of
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE *has presented a paper on* AUTOMATIC VEHICLE

INFORMATION RETRIEVAL SYSTEM USING NUMBER PLATE DETECTION

at *“International Virtual Conference on Emerging Trends on Artificial Intelligence In Industry 4.0”* in Association with Open POWER & IBM at *Loyola Institute of Technology*, Chennai on 21.06.2021.

Convenor
Mrs.G.Indhumathi

Principal
Dr. Sujatha Anand

Correspondent
Rev.Sr.M.K.Teresa



LOYOLA INSTITUTE OF TECHNOLOGY

(Approved by AICTE Affiliated to Anna University, Chennai)

INTERNATIONAL VIRTUAL CONFERENCE ON EMERGING TRENDS ON ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0



CERTIFICATE OF APPRECIATION

This is to certify that MATHANGI N of
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE has presented a paper on NEURAL NETWORK
APPROACH FOR ENHANCED AUTOMATIC DETECTION OF GLAUCOMA

at “*International Virtual Conference on Emerging Trends on Artificial Intelligence In Industry 4.0*” in Association with Open POWER & IBM at *Loyola Institute of Technology* , Chennai on 21.06.2021.

Convenor
Mrs.G.Indhumathi

Principal
Dr. Sujatha Anand

Correspondent
Rev.Sr.M.K.Teresa



SRM TRP
ENGINEERING COLLEGE



CERTIFICATE OF PARTICIPATION

S. PRISKILLA MANONMANI

This is to certify that

of **MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE** has presented a Paper titled **SECURE QUERY BY IMAGE CONTENT AND COPY DETERRENCE USING INVISIBLE DIGITAL WATERMARKING IN CLOUD SERVER**

in the International Conference on Artificial Intelligence for IOT and Sustainable Electrical Networks (ICAIISEN'21) conducted by the Department of ECE,EEE,CSE on 29.04.2021

DR.P.SUDHAKARAN
HOD/CSE
SIGNATURE

DR.B.RAMASUBRAMANIAN
VICE PRINCIPAL & HOD/ECE
SIGNATURE

DR.P.ELANGOIVAN
HOD/EEE
SIGNATURE

Dr.B.Ganesh Babu
PRINCIPAL
SRM TRP Engineering College.



LOYOLA INSTITUTE OF TECHNOLOGY

(Approved by AICTE Affiliated to Anna University, Chennai)

INTERNATIONAL VIRTUAL CONFERENCE ON EMERGING TRENDS ON ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0



CERTIFICATE OF APPRECIATION

This is to certify that GAYATHRI R of
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE has presented a paper on IDENTIFYING REAL AND FAKE
JOB POSTING - MACHINE LEARNING APPROACH

at “*International Virtual Conference on Emerging Trends on Artificial Intelligence In Industry 4.0*” in Association with Open POWER & IBM at *Loyola Institute of Technology* , Chennai on 21.06.2021.

Convenor
Mrs.G.Indhumathi

Principal
Dr. Sujatha Anand

Correspondent
Rev.Sr.M.K.Teresa



LOYOLA INSTITUTE OF TECHNOLOGY

(Approved by AICTE Affiliated to Anna University, Chennai)

INTERNATIONAL VIRTUAL CONFERENCE ON EMERGING TRENDS ON ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0



CERTIFICATE OF APPRECIATION

This is to certify that VIJAYALAKSHMI S of
MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE has presented a paper on RECOGNITION OF EMOTION
USING FACIAL EXPRESSION BY DEEP LEARNING APPROACH

at “*International Virtual Conference on Emerging Trends on Artificial Intelligence In Industry 4.0*” in Association with Open POWER & IBM at *Loyola Institute of Technology* , Chennai on 21.06.2021.

Convenor
Mrs.G.Indhumathi

Principal
Dr. Sujatha Anand

Correspondent
Rev.Sr.M.K.Teresa

Dr.SIVANTHI ADITANAR COLLEGE OF ENGINEERING TIRUCHEDUR-628215

(Accredited by NBA(CSE,EEE), New Delhi & ISO 9001:2015 Certified Institution)



NATIONAL e-CONFERENCE ON ADVANCED TRENDS IN COMPUTER SCIENCE AND ENGINEERING "NCATCSE 2K21"

Organized By

Department of Computer Science and Engineering
Computer Society of India (CSI) Students' Chapter
(Institute Membership No.: I00859)



CERTIFICATE OF PRESENTATION

This is to certify that Mr./Ms./Dr. THARESA VINOETHINI A
of Meenakshi Sundararajan Engineering College, Chennai has presented
a paper titled Open Source Lms/Lcms Mobile Application For Adaptive And
Collaborative Learning in the "National
e-Conference on Advanced Trends in Computer Science and Engineering (NCATCSE'2K21)" on
31st March, 2021 organized by the Department of Computer Science and Engineering and
Computer Society of India (CSI) Students' Chapter (Institute Membership No.: I00859).

CO-ORDINATORS

HOD/CSE

PRINCIPAL



Jyothi Engineering College

NAAAC Accredited College with NBA Accredited Programmes*

Approved by AICTE & affiliated to APJ Abdul Kalam Technological University

A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR

JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777



NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.



International Conference on Energy and Environment (ICEE 2021)

April 09-10, 2021.

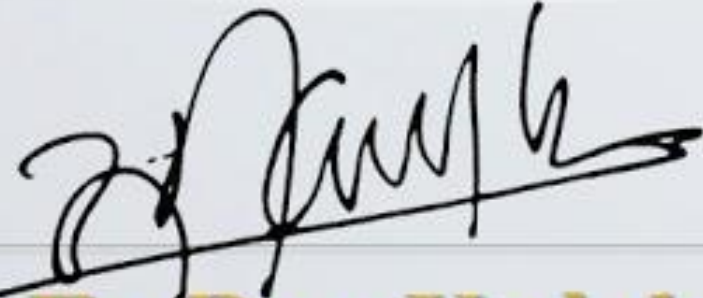
CERTIFICATE

This is to certify that the following paper has been presented and published in the International Conference on Energy and Environment (ICEE 2021 - Virtual Conference) during April 09-10, 2021 organized by Jyothi Engineering College, Thrissur, Kerala, India in association with Indian Society for Technical Education (ISTE).

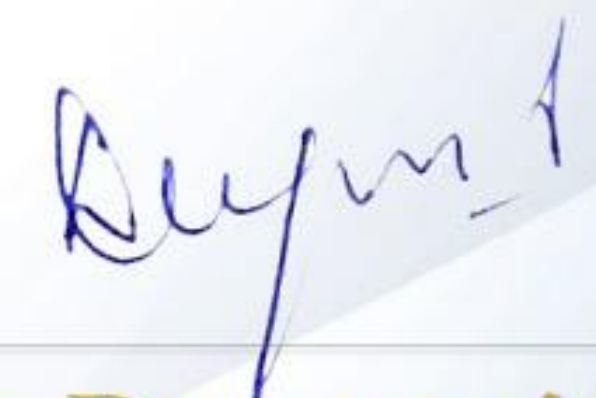
Title of the Paper : A review on recent trends in the microbial production of biodiesel

Author(s) : P. Shalini, L. Anantha Raman, S. Santhanakrishnan, S. Vijayalakshmi, J. Ranjitha

Presented by : S. Santhanakrishnan


Rev. Fr. Roy Vadakkan
Chairman ICEE 2021


Dr. Sunny Joseph Kalayathankal
General Chair ICEE 2021


Dr. Deepanraj B.
Organizing Secretary ICEE 2021