IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

Disease Prediction Application Using Machine Learning : MEDSCAN

Sahil Sandhu
UG Student
Dronacharya College of Engineering, Gurgaon, Haryana, India

Abstract: *A Harvard study by Prof Jha shows that 5.2 million medical errors are happening in India annually. We can prevent this wrong diagnosis with the help of advanced technology of Machine Learning, Deep Leaning and mobile Applications. This application provides users the facility to predict the disease on the basis of their X-Ray and MRI Scan along with the precautionary measures. Even if user is not satisfied with the result, He/ She can do a live chat or Video call to the Experienced Doctors. The ambition of this paper is to evaluate the previous effort of disease prediction, exploratory evaluation of disease and implementation of technology in Medical. Correct Disease prediction at the starting stage is demanding challenge. There are numerous experiments done on disease prediction on the basis of complex datasets but due to the shortage of automation and technology in medical they aren't available directly to the public use.*

Keywords: Android Application, Machine Learning, Predictive analysis, Deep Learning, Disease detection

I. INTRODUCTION

Android Application, Machine Learning, Predictive analysis, Deep Learning, Disease detection The World Health Organization (WHO) warned that doctors errors hurt more than 138 million patients per year. The most developed nation USA itself encounters the highest record of medical errors and wrong diagnosis. Not even USA, India encounters 5.2 million medical errors occur annually.

We can't decreases this number with any magic but we can control it up to some extent and here comes the idea of MEDSCAN.

MEDSCAN is an android based application in which Machine Learning and Deep Learning Models are integrated. MEDSCAN is an advanced application that predict the disease on the basis of the XRAY and MRI Scan images. This application uses the advanced models for the prediction of disease. No Machine is 100% perfect. That's why this application shows the possibility of event to occur along with disease. Even if the user is not satisfied with the output which is an obvious as no person want to believe that he/ she has any disease. In such case, we provide the functionality of live chatting, calling and video calling to the doctors.

This application is based on the two Profiles 1. Patient and 2. Doctor.

Whoever register in this application get an option to choose one of the profile. Either he/ she can register as a Doctor or register as Patient. Registering as a patient means he/she can't help the other and can chat only to the doctors. Where as registering as doctor means he/she can help others in resolving their doubts and even can talk to other doctors.

1.1 Related Survey Articles Comparison

Yanase et al. offered a survey in medicine from the perspective of a computer-aided diagnostic (CAD) system. The essay delves into the history and development of CAD systems. The study also represents medical applications from the standpoint of data kinds, such as tabular, image, sound, and signal data. Caballe et al outlined the advantages and drawbacks of various machine learning algorithms for disease detection. The techniques of classification, regression, and clustering are all covered in this study. It does not, however, comprise a literature review or an in-depth analysis of the papers that have been evaluated. Jiang et al looked at research articles in the field of healthcare from an AI standpoint. The paper also discusses natural language processing approaches in healthcare, in addition to machine learning. Only three medical fields are covered in the paper: cancer, neurology, and cardiology. Schaefer gave an overview on the use of machine learning in rare diseases. It examines papers in the field of medicine that deal with diagnosis, prognosis, and

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

treatment. None of the articles provide a summary of previous work. They also only cover a few medical disciplines and do not provide an in-depth analysis of the publications they review.

1.2 Machine Learning in Disease Diagnosis: A Step-by-Step Guide

Medical diagnosis is a difficult undertaking that is often thought of as an empirical task but is little understood as a cognitive task. As a result, as complicated as it may appear, computer-assisted diagnosis, in our instance utilising ML, is divided into many parts. Data collection is the initial stage in disease diagnosis. This information could come in a variety of formats, including but not limited to medical interviews, clinical trials, and surveys. Patient historical data, demographics, imaging, voice, and even heart sound can all be used. Processing is the next stage. The data is prepared at this process, which includes missing values, dimensionality reduction, handling noisy data, and so on. After that, the predictors and the target variable are determined. This information is subsequently fed into one of the training models. The model is utilised for diagnosis after it has been trained.

1.3 The Advantages of Using Machine Learning to Diagnose Disease

Because of the limits posed by a huge number of overlapping structures and cases, as well as distractions, fatigue, and limitations with the human visual system, a 'second opinion' can be beneficial. This has boosted the usage of computer-aided design (CAD) systems in diagnostic procedures. CAD is a notion in which computers and physicians play equal roles, i.e., it aids physicians in making the best clinical decisions/practices. Furthermore, EHR systems are being used to assist in making clinical decisions due to increased complexity among patients, high diagnostic errors, and the availability of a significant amount of data. With the advent of sophisticated data analysis tools, machine learning methods aid in the deciphering of relevant data linkages. It could support clinicians' findings as a second opinion. make a decision or refute it Integration of machine learning-based systems that monitor an ever-increasing volume of data streams for trends, aiding clinician decision-making, or automatically modifying settings of bedside equipment has improved patient treatment outcomes and significantly lowered overall treatment costs. On the negative, while ML promises to deliver the best therapeutic aid, the article claims that it has yet to prove beneficial, owing to opacity in ML algorithms and analytics. Furthermore, data quality and the generalizability of ML models are still issues.

II. REVIEW LITERATURE

In medical centers, a lot of studies have been conducted on disease prediction systems utilizing various machine learning algorithms.

Senthil Kumar Mohan et al suggested Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques, a strategy whose goal is to uncover essential features by using Machine Learning, resulting in increased precision in cardiovascular ailment prediction. The expectation model is made up of numerous combinations of features and a few well-known arrangement algorithms. They also educated about Different data mining approaches and expectation techniques, such as KNN, LR, SVM, NN, and Vote, have been fairly famous of late to distinguish and predict heart disease with a precision level of 88.7% through the prediction model for heart disease with hybrid random forest with a linear model (HRFLM).

Mr. Santhana Krishnan.J and Dr. Geetha.S, Mr. Santhana Krishnan.J and Dr. Geetha.S The use of a machine learning algorithm to predict heart disease This paper uses classification techniques to predict cardiac disease in male patients. This article contains thorough information about Coronary Heart Diseases, including Facts, Common Types, and Risk Factors. WEKA (Waikato Environment for Knowledge Analysis), an excellent Data Mining Tool for Bioinformatics Fields, was employed as the Data Mining tool. All three WEKA interfaces are used here; the main data mining techniques are Naive Bayes, Artificial Neural Networks, and Decision Tree, and these techniques are used to forecast heart disease in this system. Decision Trees such as CART, C4.5, CHAID, J48, ID3 Algorithms, and Naive Bayes Techniques are the most common methods for prediction.

Sharma and Shukla released a study on brain tumors in 2021. Prediction was the research's strength. Thresholding, as well as traditional supervised and unsupervised segmentation techniques, were briefly discussed. But that was only a brief overview of deep learning-based brain tumor classification and segmentation. It didn't even cover the performances of the surveyed literature.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

Similarly, Rao and Karunakara investigated various brain tumour segmentation strategies in 2021, including thresholding, region growth, atlas, deep learning, and traditional supervised and unsupervised machine learning. The results of tumour categorization approaches were clearly displayed here. The majority of the examined works on brain tumour classification are from 2019 or before in chronological order. Except for two pieces of literature that will be released in 2020. When displaying their performance metrics, the segmentation and classification algorithms are not clearly distinguished.

III. IMPLEMENTATION

3.1 One Way Communication

This application lays so much stress on the one way communication. That means one user can't directly communicate to the others. He/ She need to send the request to the other.

Even a request can only be send from the patient to the Doctor. A doctor can't send any type of message request to the Users (Patients) for any type of request or promotional messages.

But wait, What If a doctor want to consult to other doctor. This functionality is also provided to the user (Doctor). A doctor user can send message request to other doctor for consultation.

3.2 Modules

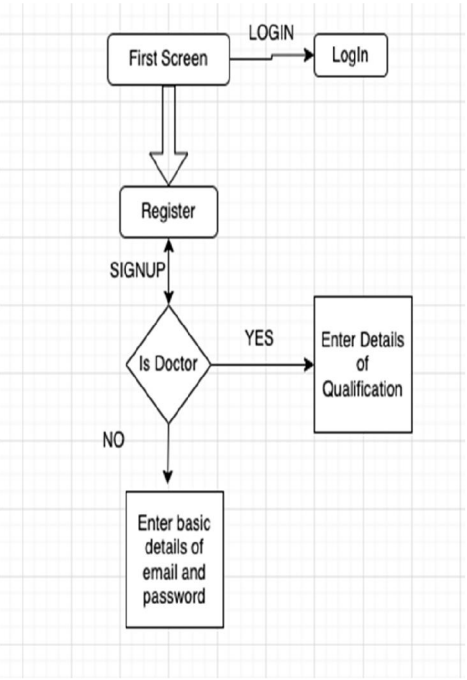
This application is divided into different modules which have different implementation, Flow chart and uses.

A. Registering Module

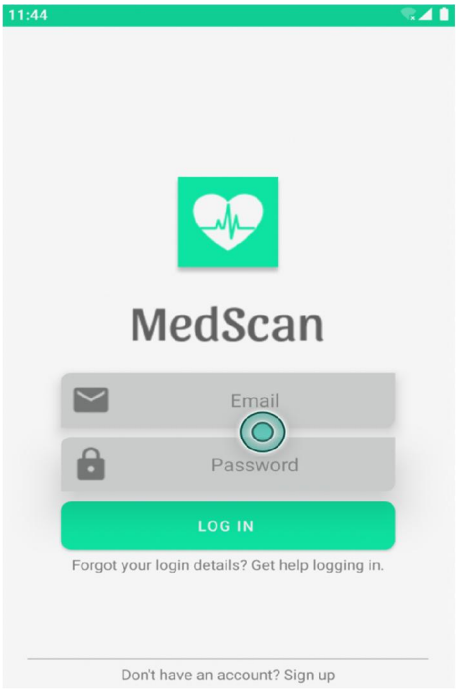
Through this module user can register to the application. Each user has 2 profiles to register with. Either he/she can register as a Doctor or register as patient.

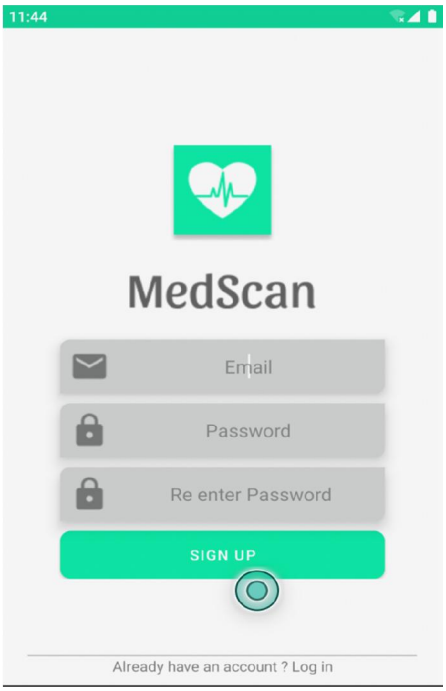
Registering as Patient is easy as here the user don't need to provide about his/ her carrier detail information whereas registering as Doctor is a long process where the user need to add his qualification, experience level, specialization and many more.

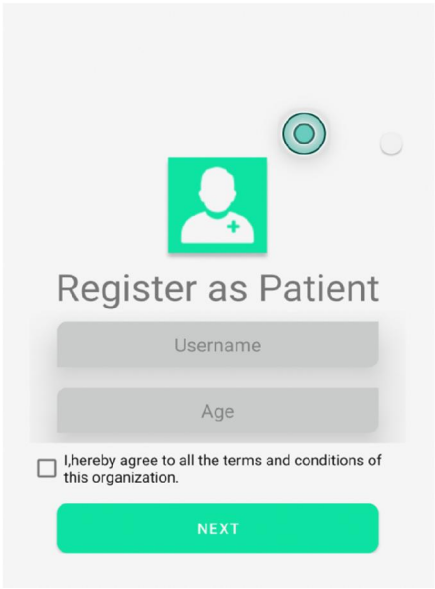
On the basis of this module their profile get created in the database.

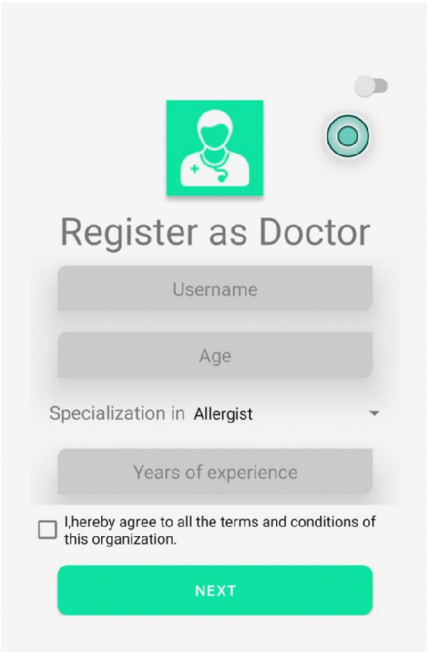


IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252









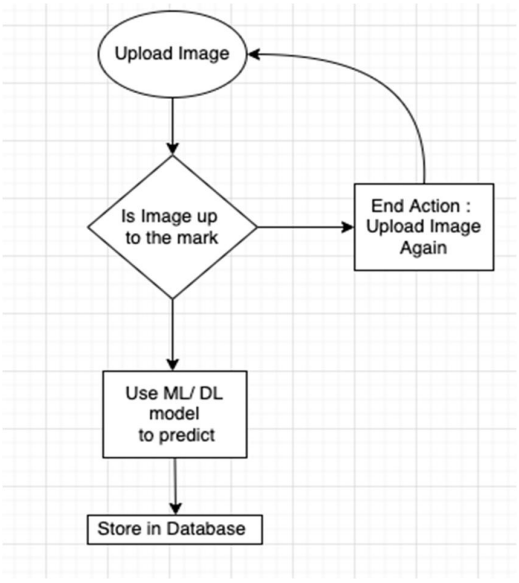
B. Prediction Module

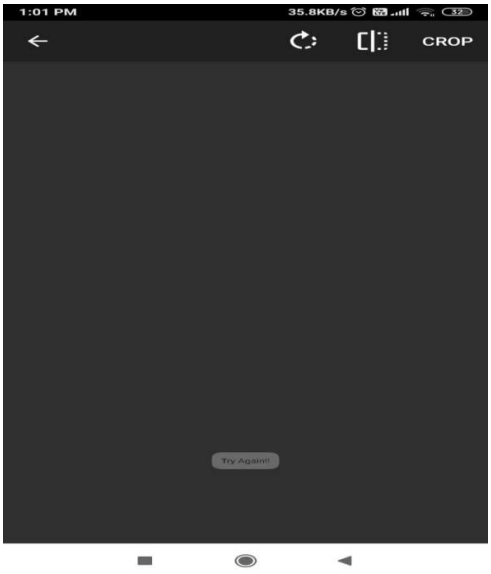
This module deals with the prediction of disease on the basis of the user MRI and XRAY scan. This module involves the uploading of document predicting it and uploading it in the database.

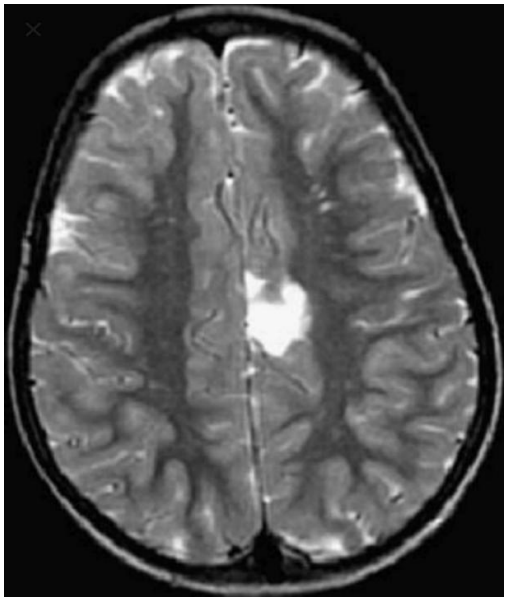
When the user try to upload the image, then he will be suggested with some images of "How the image should be."

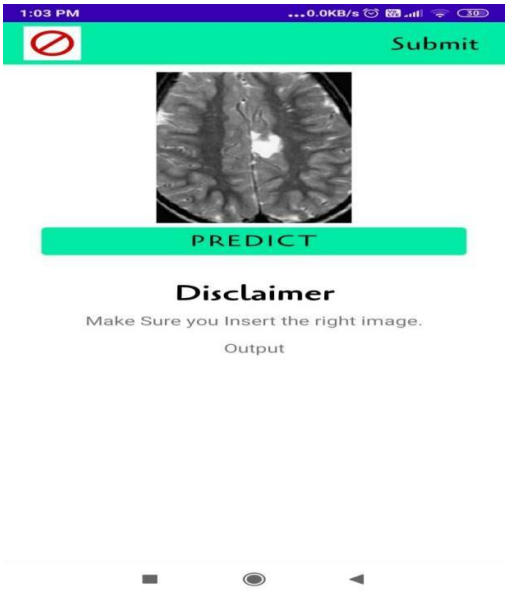
The Prediction model mainly deals with the complex algorithms of machine learning and deep learning. Each model are trained on different complex algorithm on the basis of different requirements. Efficient NET algorithm is used in Brain Tumor Prediction and Corona Prediction, ResNet is used in Retino Pathy, CNN is used in Malaria prediction, ImageNet is used in Normal Chest X - Ray

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

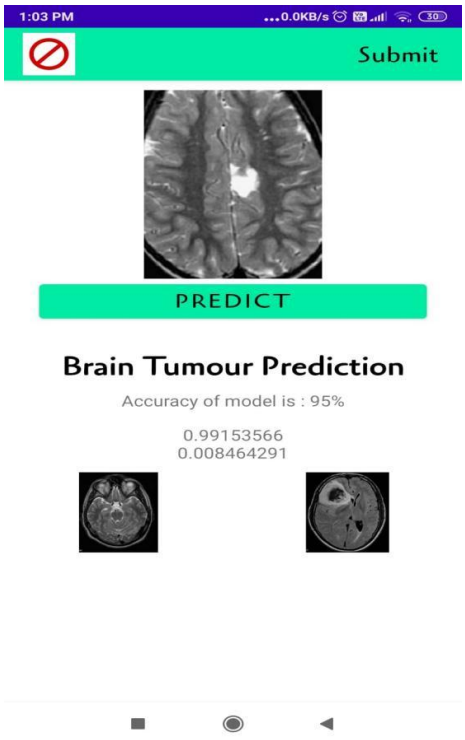








IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

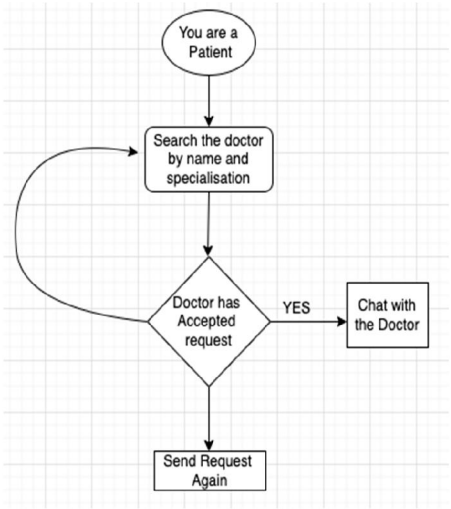


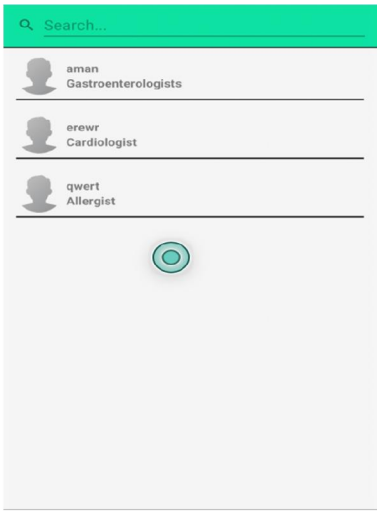
C. Chat Module

This module deals with the permissions and access level of the users. This involves the functionality of searching doctor on the basis of Specialization and name of doctor. After getting the doctor, a user can send the request for chatting and he/ she will only be able to access the chat if doctor approves his/her request. A sender and the receiver i.e. a doctor and a patient can both revoke their request.

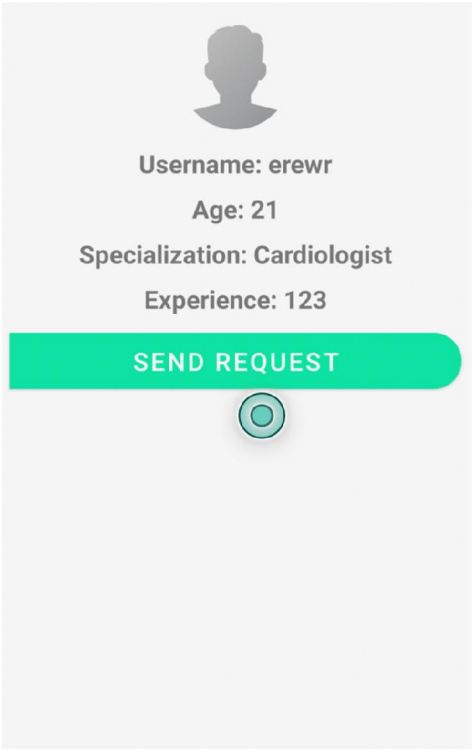
When the doctor approve the request and both users starts chatting now that person get added in their chatting list.

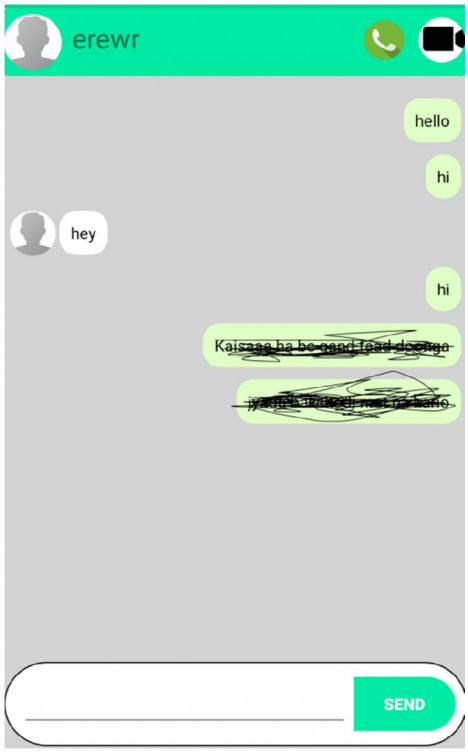
Even on the chatting screen there are signs of video call and normal call, where the user can perform the calling operation.





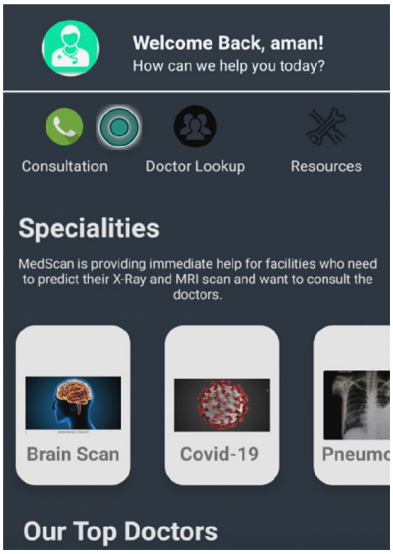
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

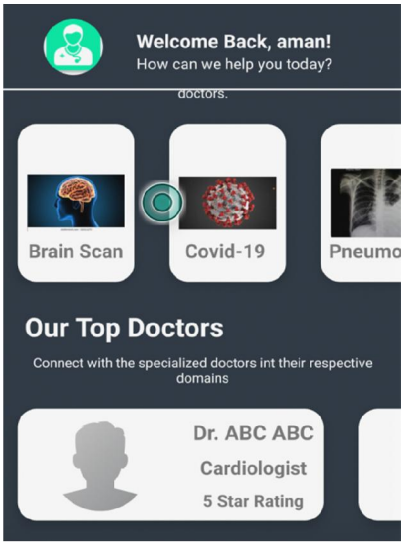




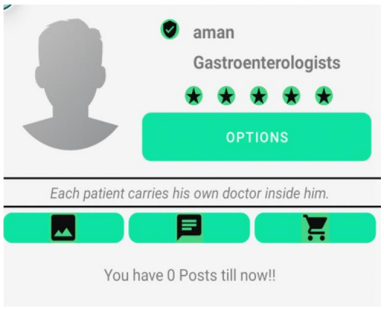
D. User Module

This is the high level module which only deals how the application will look to the user. It consists of the main profile where different options are provided from there the user get redirected to the other functions and profiles.





IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252



IV. TECHNOLOGY USED

A. Java

Java is an object-oriented programming language with a high level of abstraction and as few implementation dependencies as possible. It is a general-purpose programming language that allows programmers to write once and run anywhere (WORA), which means that compiled Java code can run on any platform that supports Java without the need to recompile it.

B. Python

Python is a high-level, general-purpose programming language that is interpreted. The use of considerable indentation in its design philosophy emphasis code readability. Its language elements and object-oriented approach are aimed at assisting programmers in writing clear, logical code for both small and large-scale projects.

V. TOOLS USED

A. Android Studio

Android Studio is Google's official[8] integrated development environment (IDE), based on JetBrains' IntelliJ IDEA software and intended exclusively for Android programming. In 2020, it will be accessible for download on Windows, macOS, and Linux operating systems, as well as as a subscription-based service. It takes the position of the Eclipse Android Development Tools (E-ADT) as the primary IDE for developing native Android apps.

B. TensorFlow

TensorFlow is a machine learning and artificial intelligence software library that is free and open-source. It can be used for a variety of applications, but it focuses on deep neural network training and inference.

C. Scikit Learn

Scikit-learn (previously scikits.learn, and also known as sklearn) is a free Python machine learning package. It includes support vector machines, random forests, gradient boosting, k-means, and DBSCAN, among other classification, regression, and clustering techniques, and is designed to work with the Python numerical and scientific libraries NumPy and SciPy.

D. Tkinter

Tkinter is Python's standard GUI library. When Python is used in conjunction with Tkinter, creating graphical user interfaces is quick and simple. The Tk GUI toolkit has a sophisticated object-oriented interface called Tkinter.

E. Keras

Keras is an open-source software library for artificial neural networks that includes a Python interface. Keras serves as a user interface for TensorFlow. Keras supported a variety of backends up until version 2.3, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

F. Matplotlib

Matplotlib is a graphing package for Python with NumPy, the Python numerical mathematics extension. It provides an object-oriented API for embedding charts into applications utilising GUI toolkits such as Tkinter, wxPython, Qt, or GTK.Up until version 2.3, Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML.

G. Seaborn

Seaborn is a matplotlib-based Python data visualisation package. It has a high-level interface for creating visually appealing and instructive statistics visuals.

VI. CONCLUSION AND FUTURE WORK

This application can help the millions of users to diagnose itself on the basis of technology of tomorrow and complex algorithms. The scope of this project is not confined to the Medical users only, it can also help the normal people to get direct consultation with live doctors and help the pharmacist to sell the quality drugs to the needy ones.

This application is an innovative application that can help the users to diagnose them self at home. If this application can be implemented for real or a business model this can change the future. As an advancement we can improve its model accuracy by keep on trying to improve its prediction procedure. Even we can convert it into a Med - E commerce application by providing the user a capability to online sell or buy the medicine.

VIII. REFERENCES

[1]. https://ieeexplore.ieee.org/document/9123590
[2]. https://ieeexplore.ieee.org/document/5206848
[3]. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
[4]. https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/world-health-statistics
[5]. https://developer.android.com/docs

BIOGRAPHY

Hello Readers, This is Sahil Sandhu. An, undergraduate student of Computer Science and Engineering Background having deep interest in the Machine Learning, Depp Learning and Application Development.