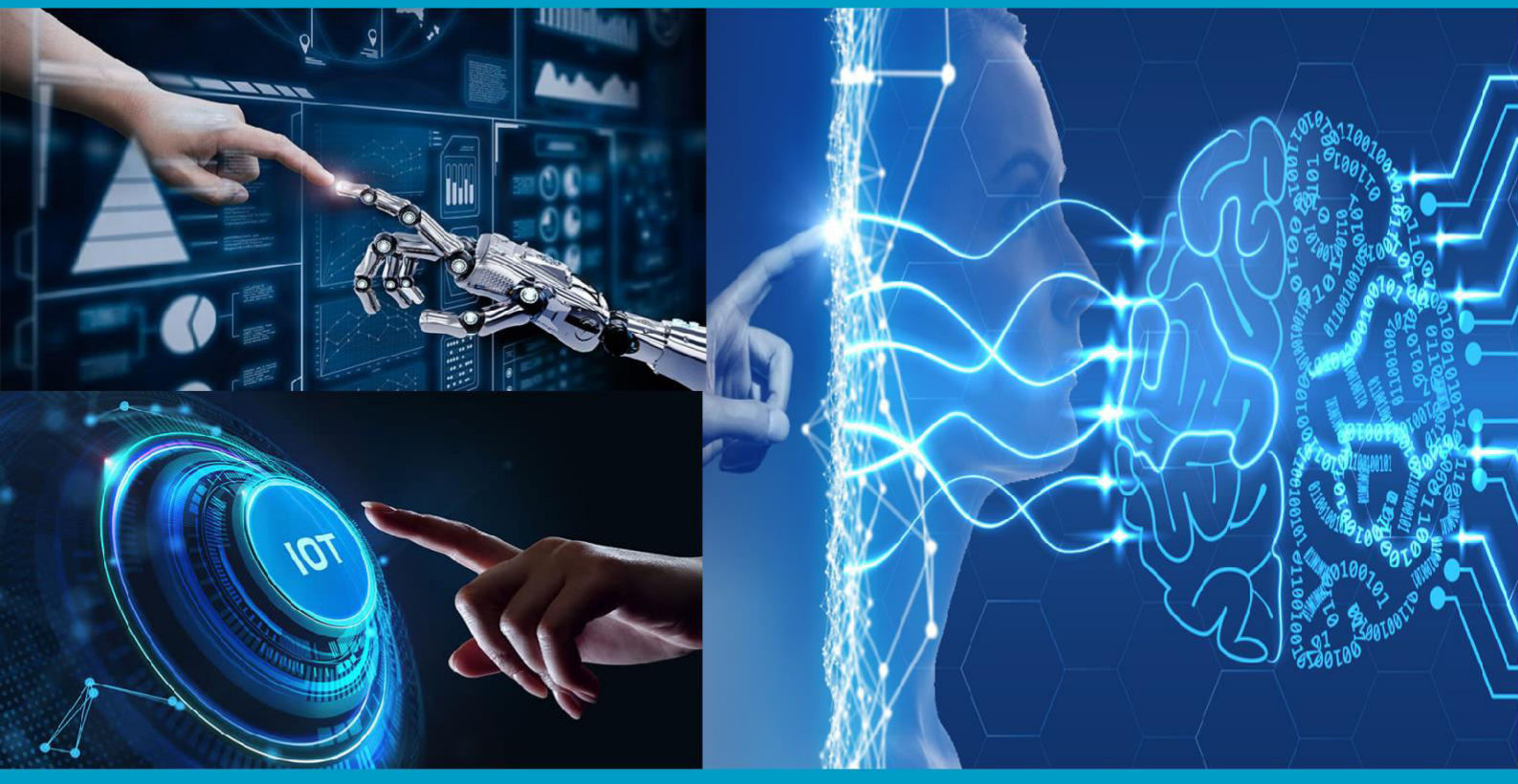
IJMRSETM

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 10, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.580

International Journal of Multidisciplinary Research in Science, Engineering, Technology &Management (IJMRSETM)

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Depression Detection Using Machine Learning

DR. SREEJA RAJESH, ALSHIFA, SAHANA, SAKSHA, SPOORTHI

Associate Professor, Dept. of Information Science & Engineering, MITE, Moodabidri, Karnataka, India

Student, Dept. of Information Science & Engineering, MITE, Moodabidri, Karnataka, India

Student, Dept. of Information Science & Engineering, MITE, Moodabidri, Karnataka, India

Student, Dept. of Information Science & Engineering, MITE, Moodabidri, Karnataka, India

Student, Dept. of Information Science & Engineering, MITE, Moodabidri, Karnataka, India

ABSTRACT: Depression is a severe mental health issue that can lead to early death, suicidal ideation, and reduced daily functioning. The model collects data through Patient Health Questionnaires (PHQ) and assesses depression levels based on the obtained scores. The data is analyzed using Random Forest algorithm to potentially detect depression in a more convenient and efficient manner. By utilizing PHQ data the model can analyze large amounts of data collected and identify individuals who may be at risk of depression. Annually, more than 15 adults suffer from depression affecting 300 million people around the world, According to WHO 3.8% of population is effected by depression including 5% among adults and 5.7% among adults older than 60 years. This app provides its users a platform to check whether a person is depressed or not. The depression is categorized into three levels i.e. Normal, Mild and Severe. The app allows the users to take the assessment and also check their progress. Based on the assesment and their progress test, the user can also check the IQ level. The project is aimed at being accessible to everyone. This project aims at motivating people and indirectly preventing people from taking drastic steps such as suicide.

KEYWORDS: Depression, Machine learning, Patient Health Questionare, Random Forest Algorithm**.**

.

I.INTRODUCTION

Mental health issues are a leading cause of disability worldwide, yet only a small percentage of those affected receive adequate treatment. Depression, in particular, can be a debilitating illness that disrupts daily life, impairs sleep and eating habits, and robs individuals of joy. The World Health Organization (WHO) reports that over 350 million people suffer from depression worldwide, with the illness now recognized as one of the most significant health problems facing the global population.While diagnosis of depression is typically based on patient interviews, it can take multiple sessions over a period of weeks to identify and treat the illness effectively. To address this issue, a proposed mobile app offers a platform for users to assess their own depression levels and track their progress over time. The app classifies depression into three levels: Normal, Mild, and Severe. By empowering individuals to identify their own mental health needs, this app aims to increase accessibility to treatment and motivate individuals to seek help before reaching a crisis point, such as considering suicide.

To maintain user privacy, the app ensures secure transmission of data and does not release any personal information. It also provides resources for individuals who may be experiencing domestic abuse or other sensitive issues and may not feel comfortable sharing this with others. The app also includes a feature that allows users to provide regular status reports of their mental health to a designated consultant, enabling users to seek professional help as needed.

Overall, the app's goal is to increase awareness about depression, stress symptoms, and their effects. It also enables users to choose the treatment options that best suit their needs and preferences, and helps to prevent depression by providing information on how it can be easily prevented. By providing a user-friendly platform for tracking and managing depression, this app aims to improve the quality of life for those who suffer from mental health issues.

II.LITERATURE REVIEW

In Nirmal Varghese Babu, E Kanaga, “Sentimental analysis in social media data for depression detection” [2022]; This paper shows a review of the sentiment analysis on Social media data for apprehensiveness or dejection detection utilizing various artificial intelligence techniques. In the survey, it was optically canvassed that social media data which consists of texts,emoticons and emojis were utilized for the sentiment identification utilizing various artificial

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

intelligence techniques. Multi Class Classification with Deep Learning Algorithm shows higher precision value during the sentiment analysis.

In a paper Sonam Gupta, Arjun Singh, Lipika Goel, Ajay Prasad, Mohammad Aman Ullar [2] , "Psychological Analysis for Depression Detection from Social Networking sites" [2022]; The study involved selecting and training tweet data with data preprocessing and removal of raw data. Training and cross-validation were carried out on two datasets - imbalanced and balanced - using differenttechniques such as SMOTE and RUS for oversampling and undersampling.

In a paper Bhanusree Yalamanchili, Nikhil Sai Kota, Maruthi Saketh Abbaraju, Venkata Sai Sathwik Nadella, Sandeep Varma Alluri [3], "Real-time acoustic based depression detection using machine learning techniques", [2020];In this paper, acoustic features are used to train a classification model to categorize a human as Depressed or not-Depressed. DIAC-WOZ database available with AVEC2016 challenge is considered for training the classifiers.SMOTE analysis is used for overcoming the class imbalance and 93% accuracy is obtained with the SVM algorithm resulting in Depression Classification Model (DCM). An android application cureD Deployed on Cloud is developed to self assess depression using DCM and PHQ-8 questionnaire. The application is tested on real time data of 50 subjects under the supervision of a qualified psychiatrist and an accuracy of 90% is obtained.

In a paper Shubham Dham, Anirudh Sharma, Abhinav Dhall, "Depression scale recognition from audio, visual and text analysis", [3][2017]; This report outlines our work during a summer internship on the AVEC 2017 challenge, where we aimed to automatically detect depression using machine learning algorithms. We utilized multimodal feature extraction and decision level fusion methods, processing data from the DAIC-WOZ database. Our approach involved using GMM clustering and Fisher vector analysis on visual data, statistical descriptors on gaze and pose, low-level audio features, head pose, and text features to extract relevant information.

In a paper Paula Lopez-Otero, Laura Docío Fernández, Alberto Abad, Carmen Garcia-Mateo, "Depression Detection Using Automatic Transcriptions of De-Identified Speech"[4],[2017];This paper presents a study on the influence of speech de-identification when using transcription-based approaches for depression detection. To this effect, a system based on the global vectors method for natural language processing is proposed. In contrast to previous works, two main sources of nuisance have been considered: the de-identification process itself and the transcription errors introduced by the automatic recognition of the patients' speech. Experimental validation on the DAIC-WOZ corpus reveals very promising results, obtaining only a slight performance degradation with respect to the use of manual transcriptions.

In a paper Lang He, Jonathan Cheung-Wai Chan, Zhongmin Wang, "Automatic depression recognition using CNN with attention mechanism from videos",[6] [2021];In the present paper we propose an integrated framework – Deep Local Global Attention Convolutional Neural Network (DLGA-CNN) for depression recognition, which adopts CNN with attention mechanism as well as weighted spatial pyramid pooling (WSPP) to learn a deep and global representation. Two branches are introduced: Local Attention based CNN (LA-CNN) focuses on the local patches, while Global Attention based CNN (GA-CNN)learns the global patterns from the entire facial region. To capture the complementary information between the two branches, Local–Global Attention-based CNN (LGA-CNN) is proposed.

In a paper Johannes C. Eichstaedt, Robert J. Smith, Raina M. Merchant, and H. Andrew Schwartz, "Facebook language predicts depression in medical records",[7] [2018]; In this paper we use language from Facebook posts of consenting individuals to predict depression recorded in electronic medical records. We accessed the history of Facebook statuses posted by 683 patients visiting a large urban academic emergency department, 114 of whom had a diagnosis of depression in their medical records. Using only the language preceding their first documentation of a diagnosis of depression, we could identify depressed patients with fair accuracy.

In a paper Manju Lata Joshi, Nehal Kanoongo, "Depression detection using emotional artificial intelligence and machine learning: A closer review"[8] [2022]; The literature review of this paper is divided into three sub sections as per the detection of emotions with respect to different sources. The first sub-section discusses about the studies con ducted to detect depression through sentiment analysis of twitter tweets. The second sub-section converse about detection of depression using facial expression (image and video processing). The last sub-section deals with the use of chatbots, emotional Al and combined inputs (text, audio, image and video) for detecting depression.

In a paper Gourav Kumar Gupta, Dilip Kumar Sharma, "Depression detection on social media with aid of machine learning platform"[9] [2021]; The objective of this paper is to identify the different machine learning algorithm methods, techniques, and approaches used by various studies related to depression detection on social media platforms

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

by conducting a comprehensive review. Various studies of from year 2013 to 2020 are reviewed to explore the research gaps and future directions.

In a paper Wheidima Carneiro De Melo, Eric Granges, Miguel Bordallo Lopez, "MDN: A deep maximisation differentiation network for spatio temperal depression detection", [10][2021]; The MDN operates without 3D convolutions, by exploring multiscale temporal information using a maximization block that captures smooth facial variations, and a difference block to encode sudden facial variations. Extensive experiments using our proposed MDN result in improved performance while reducing the number of parameters by more than 3 when compared with 3D-ResNet models.

III.PROPOSED METHODOLOGY

This app is based on the Patient Health Questionnaire dataset and is designed to help individuals assess their depression levels. The dataset has undergone data pre-processing, and raw data has been removed to ensure accuracy. The training dataset has been cross-validated, and a sampling process has been implemented to ensure the data is sampled according to psychology.

- The app works by requiring users to create an account and log in.
- Upon logging in, users will be directed to the homepage, where they can access the depression assessment through a menu list.
- The assessment contains a set of questions based on the user's age category.
- After completing the PHQ assessment, the user's depression level is determined, and the results are displayed.
- If the result shows the user is normal, they are not considered to be depressed.
- If the result shows mild depression, the app recommends books, meditation, and songs.
- If the result shows severe depression, the user is directed to a page where they can access a list of consultants whom they can contact via call or chat.
- After some time has passed, the user can take a progress test with a different set of questions to assess their depression levels once again.
- Additionally, users can check their IQ levels on the app.

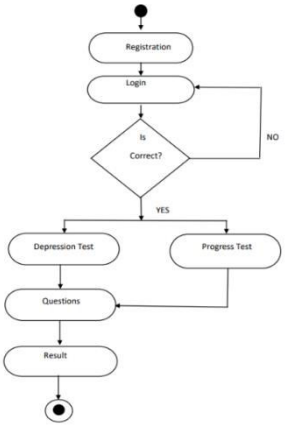


Fig 1: Activity Diagram

IV.SYSTEM IMPLEMENTATION

A crucial phase in the system development life cycle is successful implementation of new system design. Implementations simply mean converting new system design into operation. The term implementation has different meanings, ranging from the conversion of a basic application to a complete replacement of computer system. Implementation is used here to mean the process of converting a new or revised system design into an operational one.

Algorithm Used

The algorithms used in this is:

- Random Forest

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Random Forest Classifier

Random forest is a widely used machine learning algorithm for classification and regression problems. It combines the output of multiple decision trees to arrive at a single result. Random forest utilizes both bagging and feature randomness to create an uncorrelated forest of decision trees, making it an extension of the bagging method. By generating a random subset of features, feature randomness ensures low correlation among decision trees, distinguishing it from decision trees, which consider all possible feature splits. Random forests only select a subset of features, making it a flexible and easy-to-use algorithm

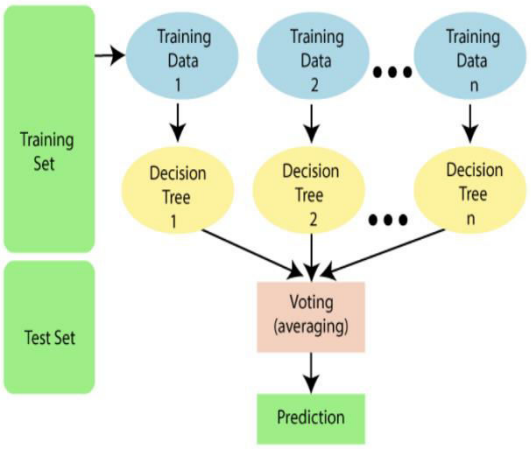


Fig 2: Random Forest Classifier

V. WORKFLOW

The following steps are carried out to implement the proposed system

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.

Dataset is pre-processed to remove redundancy words and null values from the dataset. Obtained dataset from Internet has redundant and null values which are removed during pre-processing.

Train Test Slpit

The classification of data involves organizing it into groups or classes based on specific characteristics. This allows for comparisons between observations and can be done according to numerical characteristics or attributes. To ensure accuracy, it is important to visualize the prepared data and confirm that the training data contains the correct target attribute or label. In this particular project, the dataset is split into 90 samples for training and 10 for testing. This division allows for the creation and evaluation of a machine learning model that can accurately classify new data based on the patterns and characteristics observed in the training data.

Training

The process of training an ML model involves providing an ML algorithm (that is, the learning algorithm) with training data to learn from. The term ML model refers to the model artifact that is created by the training process. The training data must contain the correct answer, which is known as a target or target attribute. The learning algorithm finds patterns in the training data that map the input data attributes to the target (the answer that you want to predict), and it outputs an ML model that captures these patterns.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Picking up the model

Pickling is a process in which model is stored in a file for future use. Before pickling model must be trained well and optimized to maximum extent possible because after picking model, data will not be trained. By pickling there is no need of training the model each time the user makes a classification.

Code implemented

The following are the code snippets used in the project.

1. Importing required packages

```
import pandas as pd
import numpy as np
from sklearn.model_selection import  train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix, accuracy_score
```

2.Splitting Dataset and Training

```
vect = CountVectorizer(ngram_range=(0, 1), max_df=.30, min_df=0)
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.25,random_state=42)
print('Started..')
vect.fit(x_train)
```

3.Loading Model and model weight

```
dataset=pd.read_csv('../dataset.csv',delimiter=',')
def  remove_pattern(input_txt, pattern):
        r = re.findall(pattern, input_txt)
        for i in r:
            input_txt = re.sub(i, '', input_txt)
        return input_txt
```

VI.SYSTEM TESTING

When it comes to testing software, there are many different approaches that can be taken to ensure its correctness, completeness, and quality. Testing is a crucial process that helps verify and validate if the software is working as intended. This involves using both static and dynamic methodologies to test the application. When designing test cases, there are two main methods that can be used: black box testing and white box testing. Black box testing involves testing the software without knowledge of its internal workings, while white box testing involves testing the software with full knowledge of its internal workings. Both methods are important in ensuring comprehensive testing of software.

White Box testing: White box testing strategy deals with the internal logic and structure of the code. It is also called as glass, structural, open and clear box testing. The test that are written based on the white box testing strategy incorporate coverage of the code written, branches, statements and internal logic of the code etc. In order to implement white box testing the tester has to deal with the code and hence it is required possess knowledge of the coding and logic i.e. Internal working of the code.

Black Box testing: Black box testingtakes the internal perspective of the test object to derived test cases. These tests can be functional or non-functional though usually functional. The test designer selects valid and invalid inputs and determines the correct input. There is no knowledge of the test object's internal structure. This method of test design is applicable to all levels of software testing: unit, internal, functional and system and acceptance.

Testing Methodology:

- Unit Testing
- Integration Testing
- System Testing

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Unit Testing

In computer programming, unit testing is a method by which individual units of source code, sets of one or more computer program modules together with associated control data, usage producers, are tested to determine ifthey are fit to use. Intuitively, one can view a unit as the smallest testable part of an application, In procedural programming a unit could be an entire module but is more commonly an individual function or procedure. In object oriented programming a unit is often an entire interface, such as class, but could be an individual method. Unit tests are created by programmers or occasionally by white box testers during the development process.

Integration Testing

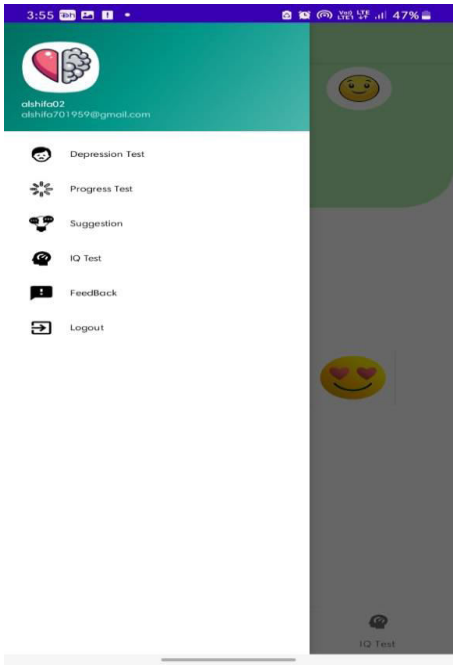
The purpose of integration testing is to verify functional, performance, and reliability requirements placed on major design items. These design items, i.e. assemblages (or group of units), are exercised through their interfaces using black box testing, success and error cases being simulated via appropriate parameter and data inputs. Simulated usage of shared data areas and inter process communication is tested and individual subsystems are exercised through their input interface. Test cases are constructed to test that all components with in assemblages interact correctively, for example across producers call of procedures activation, and this is done after testing individual modules, i.e. unit testing.

System Testing

A system testing of software or hardware is testing conducted on a complete, integrated system to evaluate system's compliance with its specified requirements. System testing falls within the scope of black box testing, and such as, should require no knowledge of the inner design of the integrated software components that have successfully passed integration testing and also the software components itself integrated with any applicable hardware system(s). System is more limited type of testing, it seeks to detect defects both within the inter-assemblages and also within the system as whole.

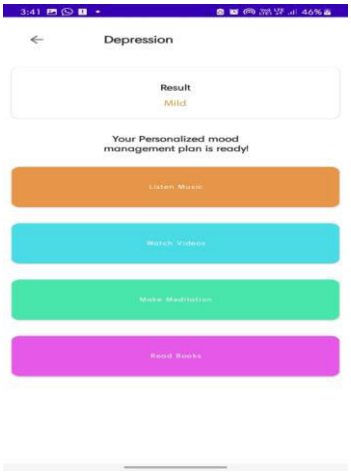
VII.RESULT

The final result of this project is an application in which a person can check his levels of depression and also his/her IQ level.The section discussed the results obtained from the input of datasets into the machine algorithm to determine the accuracy depression level.



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |



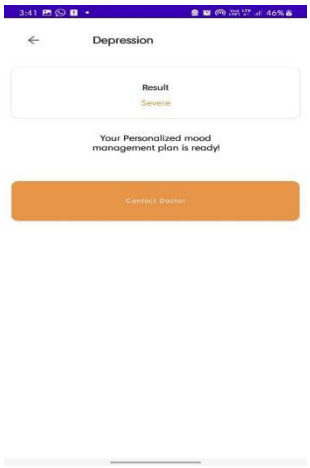


Fig 6a:Mild depression **Fig 6b:Severe deprression**

After completing the depression test,if the user gets the result as mild then the user will be recommended with some activities that he/she can can choose from.. The user can choose any of the activities where he/she are comfortable of, as shown in fig 6a.

Next, if the user gets the result as Severe then the user can get help from the consultants that are recommended by the system.The user can comfortably choose any consultants of their choice and can interact with them using chats or audio sessions, as shown in fig 6b.

After going through the depression test and making use of the solutions given by the systemThe user can go with the progress test where he /she can be able to know if they are any improvements in their conditions.

VIII. CONCLUSIONS

While apps for mental health should not be used as a replacement for professional mental healthcare, they can offer support and assistance for general mental wellness. They are a convenient way to access help between therapy sessions or office visits, and can provide ongoing support even after therapy is completed. These apps offer daily tasks, encouragement, and various strategies that can help improve overall mental health. Mobile applications have shown great potential in assisting people with mental health challenges to adopt healthier habits using persuasive tactics.There is much room for improvement in this field, and future work can refine current projects. Mental health apps are useful for both tracking overall health and monitoring mental health issues. Users can keep track of sleeping habits, mood patterns, symptoms, mealsphysical activity, thoughts, and emotions. The data collected can be examined by the user or delegated to a professional for further analysis.Overall, mental health apps offer a convenient and accessible way to support and improve general mental wellness. They can supplement professional mental healthcare and offer ongoing support, making them a valuable tool for individuals seeking to improve their mental health.

REFERENCES

1. Nirmal Varghese Babu, E Kanaga, “Sentimental analysis in social media data for depression detection” [2022].
2. Sonam Gupta, Arjun Singh, Lipika Goel, Ajay Prasad, Mohammad Aman Ullar, “Psychological Analysis for Depression Detection from Social Networking sites” [2022].
3. Hessa Almutairi, Kholound Alharbi, “Postpartum Depression Detection: Concept Analysis” [2022].
4. Xiaowei Li, Xin Zhang, Jing Zhu, Wandeng Mao, Shuting Sun, Zihan Wang, Chen Xia, Bin Hu, “Depression recognition using machine learning methods with different feature generation strategies”, [2019].
5. Aastik Malviya, Rahul Meharkure, Rohan Narsinghani, Viraj Sheth, “Depression Detection Through Speech Analysis: A Survey” [2019].
6. Qingxiang Wang, Huanxin Yang, Yanhong Yu, “Facial expression video analysis for depression in Chinese patient”, [2018].
7. J Koo LB Marangell, M Nakamura, A Armstrong, C Jeon, T Bhuthani, JJ Wu, “Depression and suicidality in psoriasis review of literature including the cytokine theory of depression”, [2017].

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

8. Paula Lopez-Otero, Laura Docío Fernández, Alberto Abad, Carmen Garcia-Mateo, "Depression Detection Using Automatic Transcriptions of De-Identified Speech",[2017].
9. Vikramjit Mitra, Elizabeth Shriberg, "Effects of feature type, learning algorithm and speaking style for depression detection from speech",[2015].
10. Manju Lata Joshi, Nehal Kanoongo, "Depression detection using emotional artificial intelligence and machine learning: A closer review" [2022].

Impact Factor: 7.580

INTERNATIONAL JOURNAL
OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com