A PRELIMINARY PROJECT REPORT ON

Depression Classification Model Using Machine Learning

SUBMITTED TO SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF ENGINEERING (COMPUTER ENGINEERING)

SUBMITTED BY

Harsimran Singh Dhillon	Exam No: B150134238
Himanshu Garud	Exam No: B150134242
Ojas Ahire	Exam No: B150134207
Aniruddha Kulkarni	Exam No: B150134270

DEPARTMENT OF COMPUTER ENGINEERING

K. K. Wagh Institute Of Engineering Education & Research

Hirabai Haridas Vidyanagari, Amrutdham, Panchavati, Nashik-422003

SAVITRIBAI PHULE PUNE UNIVERSITY

2020-2021

SAVITRIBAI PHULE PUNE UNIVERSITY

2020-2021

K. K. Wagh Institute Of Engineering Education & Research

CERTIFICATE

This is to certify that the Project Entitled

"Depression Classification Model Using Machine Learning"

Submitted by

Harsimran Singh Dhillon	Exam No: B150134238
Himanshu Garud	Exam No: B150134242
Ojas Ahire	Exam No: B150134207
Aniruddha Kulkarni	Exam No: B150134270

is a bonafide work carried out by Students under the supervision of Prof. Namrata Pagare and it is approved for the partial fulfilment of the requirement of Savitribai Phule Pune University, for the award of Bachelor of Engineering (Computer Engineering).

Prof. Namrata Pagare Internal Guide Dept. of Computer Engg.	Prof. Dr. S.S.Sane H.O.D Dept. of Computer Engg.

Dr. K. N. Nandurkar

Principal

Place: Nashik

Date:

ACKNOWLEDGEMENT

It gives us great pleasure in presenting the preliminary project report on ***'Depression Classification Model Using Machine Learning'****.*

We would like to take this opportunity to thank our internal guide ***Prof. Namrata Pagare*** *for giving us all the help and guidance we needed. We are really grateful to her for her kind support. Her valuable suggestions were very helpful.*

We are also grateful to ***Prof. Dr. S.S.Sane****, Head of Computer Engineering Department, K. K. Wagh Institute Of Engineering Education And Research for his indispensable support, suggestions.*

In the end our special thanks to ***Prof. Namrata Pagare*** *for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for our Project.*

Harsimran Singh Dhillon
Himanshu Garud
Ojas Ahire
Aniruddha Kulkarni
(B.E. Computer Engg.)

ABSTRACT

Depression is becoming major mental illness faced by humans of all ages and genders. The work culture, stressful life, emotional imbalance, family disturbances, and social life is leading to depression. So depression is predicted to rise to the second leading cause of illness as per the identifications of the World Health organization (WHO). Though well trained clinicians, medical and psychological treatments are available for depression treatment, but persons or families are unwilling to speak out or reach doctors about this disorder for various social reasons. Diagnosis of depression disorder includes numerous interviews with patient and family, clinical analysis, questionnaires which is time consuming and also demands well trained clinicians. A Machine learning approach is used for automation of depression detection. With the help of improvised machine learning techniques like MLP algorithm and DeepFace library, extracting acoustic features along with facial features becomes more efficient and increases the accuracy of the model.

INDEX

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

Globally, more than 264 million people of all ages suffer from depression. Depression is a leading cause of disability worldwide and is a major contributor to the overall global burden of disease [4]. Depression affect one in every 15 adults in a given area and risk in women is twice than men. People consider depression as an insensitive topic and often ignore it under the fact of being sad. This affect their lifestyle and could come along with many health related issues.

1.2 PROBLEM DEFINITION AND OBJECTIVES

1.2.1 Problem Definition

People often get confused between being sad or depressed this delays their diagnosis procedure affecting their health. Depression classification model is a system that takes input in the form of image, audio and questionnaire for predicting the severity of depression in individuals.

1.2.2 Objectives

- To build a system using machine learning and deep neural network models for predicting the depression severity in people
- To reduce the time required for diagnosis by automating the diagnosis procedure

1.3 PROJECT SCOPE AND LIMITATION

1.3.1 Project Scope

The inputs to the system will be textual data, image and audio.

- The response to the PHQ-9 questionnaire set is collected
- The image data is collected and the required facial-emotion features are extracted

- The audio samples are collected using pyaudio and the required features are extracted

1.4 METHODOLOGIES OF PROBLEM SOLVING

- The prediction of the depression level is done using machine learning algorithm
- The image and audio samples needs to be compared for feature extraction step, which will require a considerable amount of time. So, the main challenge is to parallelize this comparison to reduce the time required to get the output
- To make the application as user friendly as possible we need to reduce the lag at each stage which could be achieved using parallel computing or by the use of GPU's

CHAPTER 2

LITERATURE SURVEY

One of the major mental illness faced by human of all ages and gender is the depression. Depression affect one in every 15 adults in a given area and risk in women is twice than men [4]. Symptoms include: change in food and habits, loss or gain in weight, loss in concentration, anxiety, hopelessness, feeling of less use, heart diseases, inflammation, sexual health problems, sleep disorder, etc [4]. Human beings are social animal, trained to exhibit only strength and hide the weakness so they are afraid to discussed these problems openly. The diagnosis of depression depression needs one's openness and honesty. The purpose of the application is to detect depression in an individual by using an android device at any place. This application will use diagnostic and assessment tools by using the machine learning techniques to detect depression which includes, interview style assessment, automatic detection using speech, extracting facial features from video. This application can detect the depression and then he can be referred to a trained and experience doctor.

- **The PHQ-8 as a measure of current depression in the general population (2014)[3]**:

  The Patient Health Questionnaire (PHQ) is a questionnaire that can be self-administered by the patient. The PHQ assesses 8 diagnoses, divided into threshold disorders (disorders that correspond to specific diagnoses: panic disorder, other anxiety disorder, and bulimia nervosa), and sub-threshold disorders (disorders whose criteria encompass fewer symptoms than are required for any specific diagnoses: other depressive disorder, probable alcohol abuse/dependence, somatoform, and binge eating disorder) .

- **Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression(2019)[7]:**

  The PHQ-9 is the 9 depression module questions. Major depression is diagnosed if 5 or more symptoms out of the 9 depressive symptoms crite-

ria have been present at least “more than half the days” in the past 2 weeks, and 1 of the symptoms is depressed mood or anhedonia [2]. Other depression is diagnosed if 2, 3, or 4 depressive symptoms have been present at least “more than half the days” in the past 2 weeks, and 1 of the symptoms is depressed mood or anhedonia [2]. One of the 9 symptom criteria (“thoughts that you would be better off dead or of hurting yourself in some way”) counts if present at all, regardless of duration. As with the original PRIME-MD, before making a final diagnosis, the doctor is expected to rule out physical causes of depression, normal bereavement, and history of a manic episode. As a severity measure, the PHQ-9 score can range from 0 to 27, since each of the 9 items can be scored from 0 (not at all) to 3 (nearly every day). An item was also added to the end of the diagnostic portion of the PHQ-9 asking patients who checked off any problems on the questionnaire: “How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?”

- **Tracking depression severity from audio and video based on speech articulatory coordination(2019)[2]:**

  To gain the high accuracy we combine the facial features and acoustic features. These extracted features are then cross verified with the score of PHQ-9 questionnaires and then evaluate the final result. Various acoustic features are extracted using COVAREP toolbox and are fuse to enhance the efficiency of classification. For feature selection Principal Component Analysis (PCA) is used. Machine learning algorithms like K Nearest neighbors(KNN), Gaussian Mixture Model(GMM), Support Vector Machine(SVM) are used to detect depression using different speech types and emotions. From the video and image data the facial features, expressions and posture are extracted. Feature engineering in speech includes proper extraction of features frame wise, calculating the statistical measures, feature fusion and their selection based on correlation. 73 baseline Low Level Descriptors(LLD) including Prosodic(2), Voice quality(8) and Spectral(63) features [15] are extracted using COVAREP

toolbox for each frame sampled at 10ms interval. Statistical measures- mean, median, variance, kurtosis, skewness, minimum, maximum, standard deviation are calculated on all the features resulting in 584 feature set. Prosodic, Spectral and Voice Quality features are fused and Principal Component Analysis (PCA) is applied to reduce feature space dimension.

CHAPTER 3

SOFTWARE REQUIREMENT SPECIFICATION

3.1 ASSUMPTION AND DEPENDENCIES

- The user should have camera and mic
- The user should record themselves in a less noisy place having good lighting conditions
- The user should have good internet connectivity

3.2 FUNCTIONAL REQUIREMENT

3.2.1 System Feature 1(Functional Requirement)

The user needs to respond to a questionnaire having a set of 9 questions in it. The system then shall predict the severity of depression in the user. This method has a sensitivity of around 88 percent.

3.2.2 System Feature 2(Functional Requirement)

The system then takes input from camera and mic, in this step the user needs to read a paragraph with their mic and camera permissions allowed. The facial features and audio samples which were collected are preprocessed and the model will classify the severity of depression.

3.3 EXTERNAL INTERFACE REQUIREMENTS (IF ANY)

3.3.1 User Interfaces

- Front-end software: Vue.js
- Back-end software: Python, JSON and MongoDB

3.3.2 Hardware Interfaces

- Windows or Mac
- A browser which supports Vuejs, HTML and Javascript

3.3.3 Software Interfaces

- Operating system: Windows for its best support and user friendliness
- Database: The patient details are saved in MongoDB database
- Datasets: To datasets required for training the model are RAVDESS, TESS, SFC and TFW

3.3.4 Communication Interfaces

This system is supported by all types of web browsers.

3.4 NON FUNCTIONAL REQUIREMENTS

3.4.1 Performance Requirements

- System will perform with more efficiency if proper dataset is provided
- System performs better with good internet connectivity
- If the system has the recommended hardware specifications then the system will perform smoothly

3.4.2 Safety Requirements

Data is stored in mongoDB database as well as in json file format. If there is extensive damage to a wide portion of the database due to catastrophic failure, such as a disk crash then the data from json file is used to recover the data back.

3.4.3 Security Requirements

- Integrity: Avoid the improper (unauthorized) information modification or destruction. Here it is included to ensure the non-repudiation and information authenticity
- Availability: The information must be available to access whenever it is required and can be used with reliable authorization. Definitely it must be true for those who have right access to the data

- Identification: Allow users tell to application who they are
- Authenticate: verify the credentials of users

3.4.4 Software Quality Attributes

There are few software quality attributes which describes the quality of proposed system and they are as mentioned below.

- Availability
- Correctness
- Usability

3.5 SYSTEM REQUIREMENTS

3.5.1 Database Requirements

MongoDB database is used to store the user details.

3.5.2 Software Requirements (Platform Choice)

- Front-end: Vue.js because it is interactive and user friendly
- Back-end: Python as it makes visualization and representation of data easy
- Operating system: Windows for its best support
- Database: MongoDB to store unstructured data

3.5.3 Hardware Requirements

- CPU: i3 8th gen
- RAM: 4 GB

CHAPTER 4

SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

The user would have to insert their details like their name, email-id, gender and age while signing up.

After signing up the user should respond to the PHQ-9 Questionnaire which comprises of 9 questions based on the symptoms experienced in the past 14 days. The responses collected from these 9 questions are given to the rule based model for feature classification.

The facial and audio samples of the user is collected in a model where the user needs to read a paragraph after allowing the access to their mic and camera and keeping them on. The camera would be capturing frames after every 3 secs and that frame is stored in the database. This process is very important to understand the sentiments of the user. Along with the facial features audio sample is also collected from the mic. The length of the audio sample would be around 45 seconds

These captured used data is given to pre-processing/feature extraction block so that required normalization and standardisation of the data could be done. Pre-processing step is very important as it makes the feature extraction more reliable. There are different features which are required by the image and the audio model. These features required for classification are extracted separately for the image and audio model.

The extracted features are given to the DeepFace and MLP classifier model for classifying the depression severity of the user. The DeepFace model is used to classify the facial emotion of the user and the MLP classifier is used to classify the acoustic sentiment of the user from the audio samples.

This model classifies the output in one of the 5 possibilities of the depression severity. This would help the user to know that what level of depression they are facing so that they could get treatment accordingly

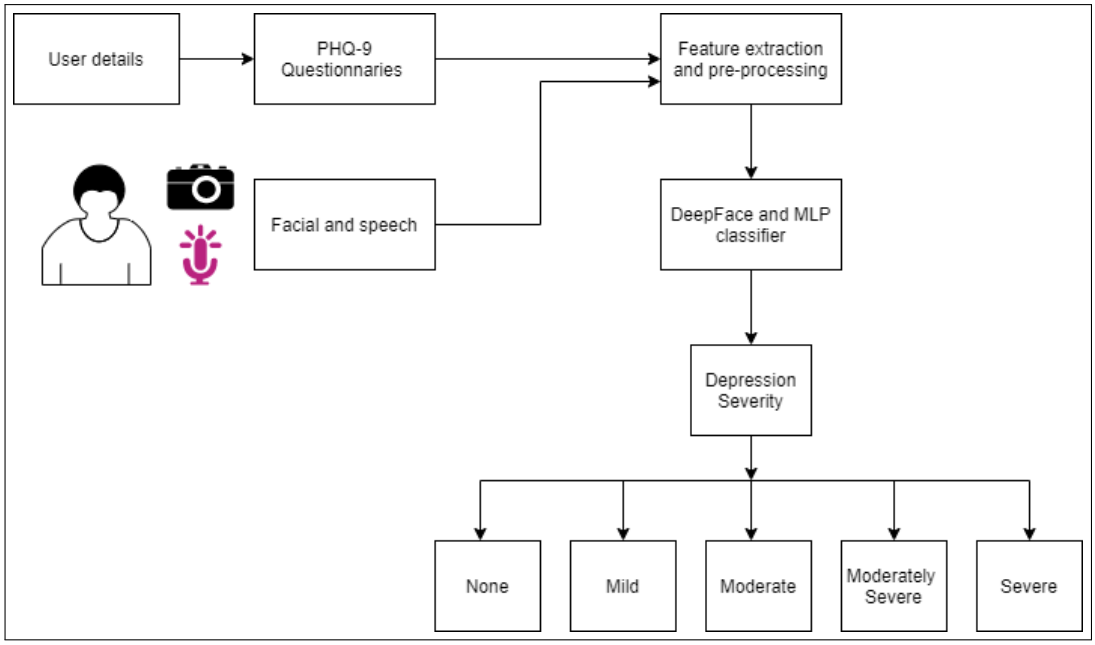


Figure 4.1: Architecture diagram

4.2 DATA FLOW DIAGRAM

4.2.1 Level 0 Data Flow Diagram

The user needs to enter their details such as their name, age, gender and email-id while signing up. From this the age and gender will be taken into the consideration, as depression level are different for different age group and different gender.

4.2.2 Level 1 Data Flow Diagram

In this level the user will answer the PHQ-9 questionnaire and even they would be told to read a passage while keeping their camera and mic on. This will allow the system to capture frames of camera and audio samples of the user.

4.2.3 Level 2 Data Flow Diagram

The data collected from level 1 is preprocessed and then given to our model for predicting the output.

4.3 ENTITY RELATIONSHIP DIAGRAM

Shown below is the entity relationship diagram model having two entities doctor which could also be the admin of the system and the patient which would be the user. The doctor would have the access to the patients details such as name, email, phone, gender, age and depression severity of the patient. There are multiple doctors consulting multiple patients.

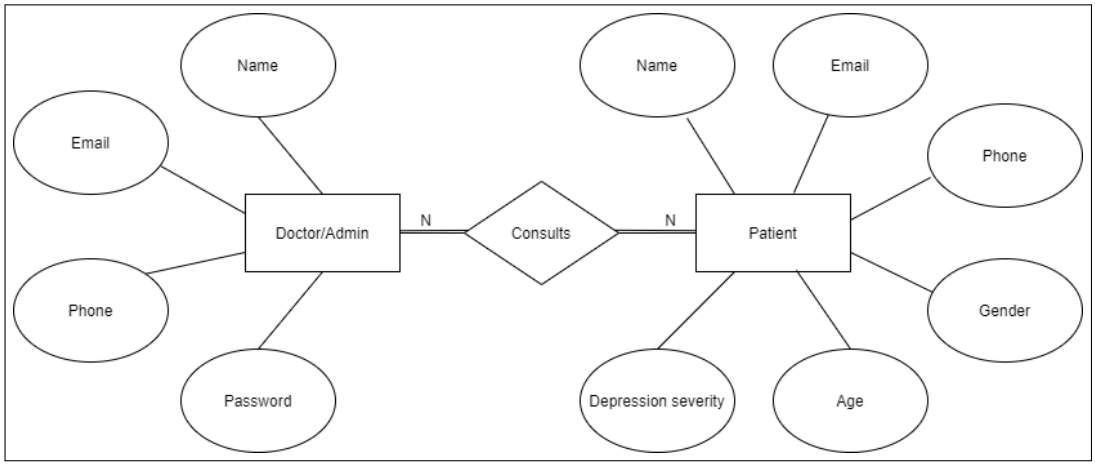


Figure 4.2: ER diagram

4.4 UML DIAGRAM

The UML diagram of the system is given below. The system illustrates that the actor needs to enter the valid login credentials to access the PHQ-9 questionnaire. After submitting the response to the PHQ-9 questionnaire the user would be redirected to facial and speech model, where the users facial and speech samples are collected. The features would be extracted and the model would classify the severity.

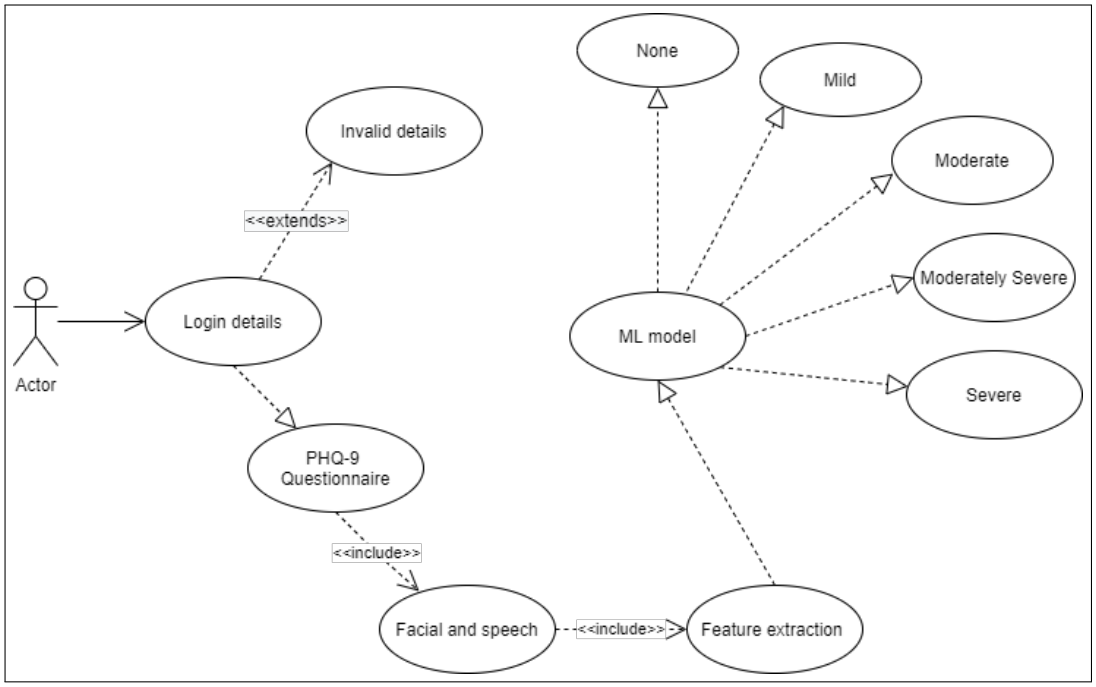


Figure 4.3: UML diagram

CHAPTER 5

PROJECT PLAN

5.1 PROJECT ESTIMATE

5.1.1 Reconciled Estimate

5.1.1.1 Cost Estimate

The model followed is the Constructive Cost Model(CO-COMO) for estimating the effort required in the completion of project. Like all estimation models, the CO-COMO model requires sizing information. This information can be specified in the form of

- Object Point
- Function Point(FP)
- Line of Source Code(KLOC)

For our project, we use the sizing information in the form of Lines Of Source Code.
Estimated lines of code, LOC = 1600
Equations:
The initial effort(Ei) in man-months is calculated using equation:
E = $ax(KLOC)^b$
Where, a = 3.0, b = 1.12, for a
semi-detached project
E = Efforts in person-hours
E = 4.5 PM
D = $ax(E)^b$
Where, a = 2.5,
b = 0.35, for a semi-detached project
D = Duration of Project in months
D = 5 Months

5.1.1.2 Time Estimates

C= D * Cp * hrs

Where, C = Cost of project,

10D = Duration in hours,

Cp = Cost incurred per person-hour

hrs = hours

Efforts E = 4.5 person-month

Total of 4.5 persons-months are required to complete the project successfully.

Duration of project D = 5 months The approximate duration of project is 5 months.

5.1.2 Project Resources

The framework used for front-end is Vue.js and the language used for back-end is python. The platforms used are VScode, Jupyter Notebook and MongoDB. The datasets used are RAVDESS, TESS, SFC and LFW for predicting the output with higher accuracy.

5.2 RISK MANAGEMENT

This section discusses Project risks and the approach to managing them.

5.2.1 Risk Identification

1. Skill Level: The skill level could be less
2. Image Data Quality: The images may not be of good quality
3. Audio Data Quality: There audio may contain noise
4. Low Performance Speed: Due to large comparison of dataset, the system processing could get slow

5.2.2 Risk Analysis

The risks for the Project can be analyzed within the constraints of time and quality

5.2.3 Overview of Risk Mitigation, Monitoring, Management

Following are the details for each risk.

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Skill Level	Low	High	High	High
2	Image Data Quality	Low	Medium	High	High
3	Audio Data Quality	Medium	Medium	High	High
4	Low Performance Speed	Medium	Medium	Medium	Medium

Table 5.1: Risk Table

Probability	Value	Description
High	Probability of occurrence is	$> 75\%$
Medium	Probability of occurrence is	$26 - 75\%$
Low	Probability of occurrence is	$< 25\%$

Impact	Value	Description
Very high	$> 10\%$	Schedule impact or Unacceptable quality
High	$5 - 10\%$	Schedule impact or Some parts of the project have low quality
Medium	$< 5\%$	Schedule impact or Barely noticeable degradation in quality Low Impact on schedule or Quality can be incorporated

Risk ID	1
Risk Description	Skill Level
Category	Work Distribution.
Source	Project Plan Document.
Probability	Low
Impact	High
Response	Mitigate
Strategy	Planning and assigning the job accordingly
Risk Status	Identified

Risk ID	2
Risk Description	Image Data Quality
Category	Dataset
Source	Input
Probability	Low
Impact	High
Response	Mitigate
Strategy	Use good quality of data for training the images
Risk Status	Identified

Risk ID	3
Risk Description	Audio Data Quality
Category	Dataset
Source	Input
Probability	Medium
Impact	High
Response	Mitigate
Strategy	Collecting multiple audio samples
Risk Status	Identified

Risk ID	4
Risk Description	Low Performance Speed
Category	Design
Source	Input.
Probability	Medium
Impact	High
Response	Mitigate
Strategy	Parallel processing could be used
Risk Status	Identified

5.3 PROJECT SCHEDULE

5.3.1 Project task set

Major Tasks in the Project stages are:

- Task 1: Project Planning
- Task 2: Data cleaning and Preprocessing
- Task 3: Selecting and Training the Model
- Task 4: Implementation and Application Development
- Task 5: Deploying the Application
- Task 6: Testing and Documentation

5.3.2 Task network

Project tasks and their dependencies are noted in this diagrammatic form.

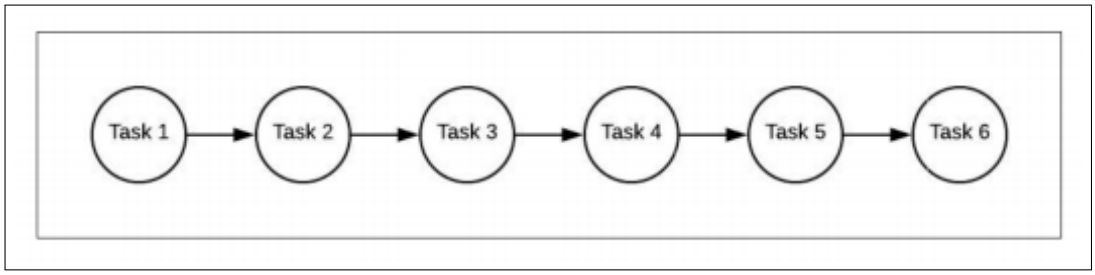


Figure 5.1: Task network graph

5.3.3 Timeline Chart

This project is basically divided in 6 different task. Each one followed by the other one.

- 1^{st} Aug-1^{st} Oct: During this time period literature survey and project planning is done. Choosing the right platform and languages to be used. Planning the model to be used and its working. Everything related to this is discussed in this time period

- 2^{nd}Oct - 10^{th}Nov : Data collection and preprocessing will be done in this time duration. Preprocessing the data is very important for the model to execute properly. The data might contain some blank values or different data formats so converting and normalizing the data according to the requirements of the model is very important

- 11^{th} Nov- 15^{th} Dec: The model is trained using the preprocessed data. The model is trained and tested for various inputs to get the highest accuracy

- 16^{th} Dec-15^{th} Feb: Web application development phase will be initiated and the graphical user interface will be designed in such a way that it would be user friendly and interactive

- 16^{th} Feb-15^{th} Apr: This application would be tested by different individuals so that it could help us fix the bugs and resolve the problems

- 16^{st} Apr - 10^{th} May: Documentation

5.4 TEAM ORGANIZATION

5.4.1 Team structure

The team consists of 4 members, where the tasks related to the project are distributed equally. The team members and the assigned responsibilities are:

- Harsimran Singh Dhillon : Model design and algorithms implemented for the model

- Himanshu Garud : Model design and algorithms implemented for the model

- Ojas Ahire : Dataset collection and testing

- Aniruddha Kulkarni : Documentation and testing

5.4.2 Management reporting and communication

A dedicated time is allotted for communication between team members. In this time duration team members communicate by using google meet an online platform for

meetings. Doubts and queries were solved during this meet and progress were noted in log book.

CHAPTER 6

PROJECT IMPLEMENTATION

6.1 OVERVIEW OF PROJECT MODULES

- Questionnaire

  This module consists of rule based learning model. In this module the user needs to answer a set of questions, these questions are defined by professionals and are based on users past 14 days of experience. The rule based learning model then classifies the severity of depression in an individual. This model have a specificity rate of 88 percent and sensitivity rate of 89 percent.

- Facial

  In this module the system collects the facial features of the user. Required preprocessing steps are applied to remove noise and disturbance from the collected samples. These samples are then given to the applied DeepFace algorithm which classifies the facial features and gives the dominant facial feature as output out of all the samples collected.

- Acoustic

  This is the last module of the system in which the user needs to read a paragraph and the audio samples of the user will be collected by the system. The speech of the user is captured using pyaudio and the important features are extracted using MLP classifier to classify the user speech. The classified speech output is then mapped with the facial emotion classification and the questionnaire classification to get the final output.

6.2 TOOLS AND TECHNOLOGY USED

- Tools: VScode, Jupyter Notebook and MongoDB compass
- Technology: Vue.js, javascript, python and mongodb

6.3 ALGORITHM DETAILS

6.3.1 MLP Classifier

- MLP classifier is used for classifying the speech emotion classification
- The dataset used for classification are RAVDESS and TESS
- The speech samples are collected dynamically using pyaudio
- The features of the speech samples were extracted using librosa library
- The functions defined for extraction of speech features are : MFCC, MEL, CHROMA
- The dataset is loaded and is split into 75 and 25 percent for training and testing the model respectively
- MLP Classifier is initialized and trained
- Collected speech samples are tested for classification using this trained model

6.3.2 DeepFace Architecture

- DeepFace is used for classifying the facial features of the user
- DeepFace in default configuration compares faces with VGG-Face model
- Cosine similarity metric is used for finding out the similarity or differences
- OpenCV is used to capture the camera frames at run time
- Haarcascade frontal face is used preprocessing the collected sample data
- DeepFace analyses the test sample data and classifies the output

CHAPTER 7

SOFTWARE TESTING

7.1 TYPES OF TESTING

- Unit Testing : The Facial model and the Speech model are run and tested individually using pytest. Pytest is the best unit testing framework. PHQ-9 Questionnaire model is tested using Jest.
- User Interface Testing : Jest is used to test the user interface. Jest is a JavaScript test framework that is focused on simplicity. The login page, sign up page, and other links of the website are tested

7.2 TEST CASES AND TEST RESULTS

- PHQ-9 Questionnaire Module :
  - Test Case 1: Output Changes dynamically after changing the input
  - Test Case 2: All the options and their associated values of outputs were compared
  - Test Case 3: Responses getting submitted
- Facial Feature Module :
  - Test case 1: No output if there is no human face detected (No human in frame)
  - Test case 2: No output if the user is look in different direction (when face is more tilted in different direction)
  - Test case 3: Improper lightning conditions (too much light or low light affects the classification results)
- Audio :
  - Test case 1: Classified emotion as the output of the speech
  - Test case 2: No output if more than one person's voice or lot of noise is detected
  - Test case 3: No output if the mic is being accessed somewhere else

- User details
  - Test case 1: Prompts invalid details(if user enters invalid phone number digits and password validation)
  - Test case 2: Prompts successful registration and redirects to login page(After registration)
  - Test case 3: Prompts invalid password(when user enters a wrong password in login page)
  - Test case 4: Successful login and redirects to home page(when user enters correct credentials)

CHAPTER 8

RESULTS

8.1 OUTCOMES

- PHQ-9 questionnaire classifies the severity of depression with a sensitivity of 89 percent
- DeepFace classifies emotion of the person with an accuracy of 93 percent
- MLP Classifier classifies emotion using acoustic features with an accuracy of 92.75 percent

8.2 SCREEN SHOTS

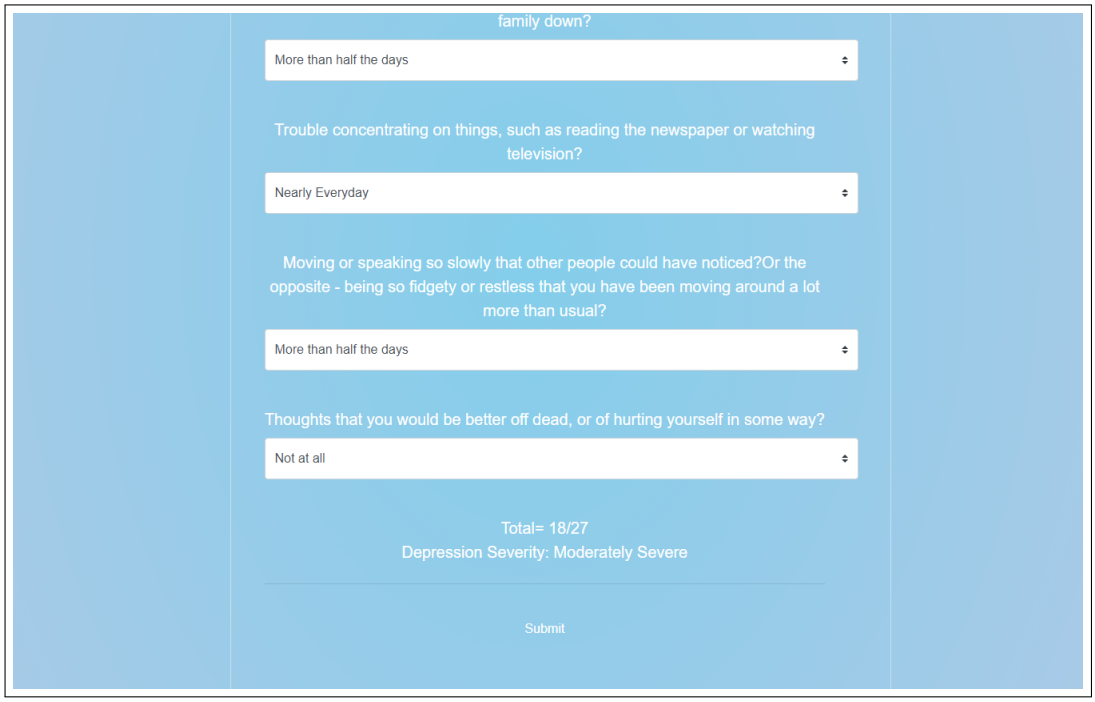


Figure 8.1: PHQ-9 questionnaire

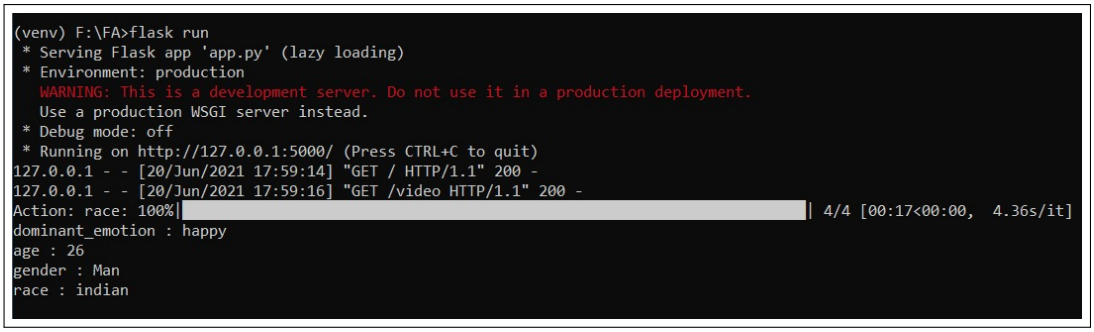


Figure 8.2: Facial Features Classification

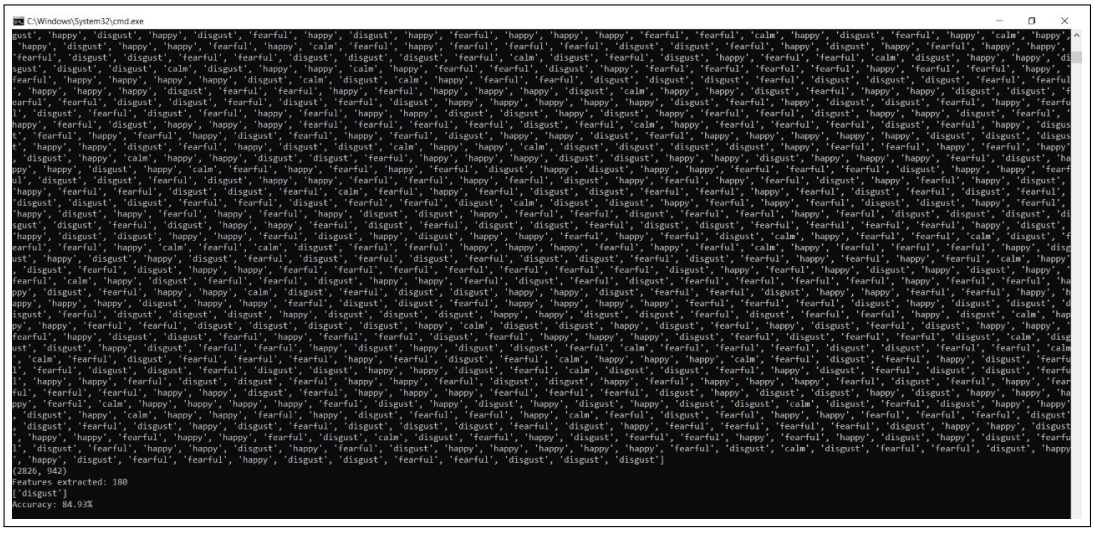


Figure 8.3: Speech Features Classification

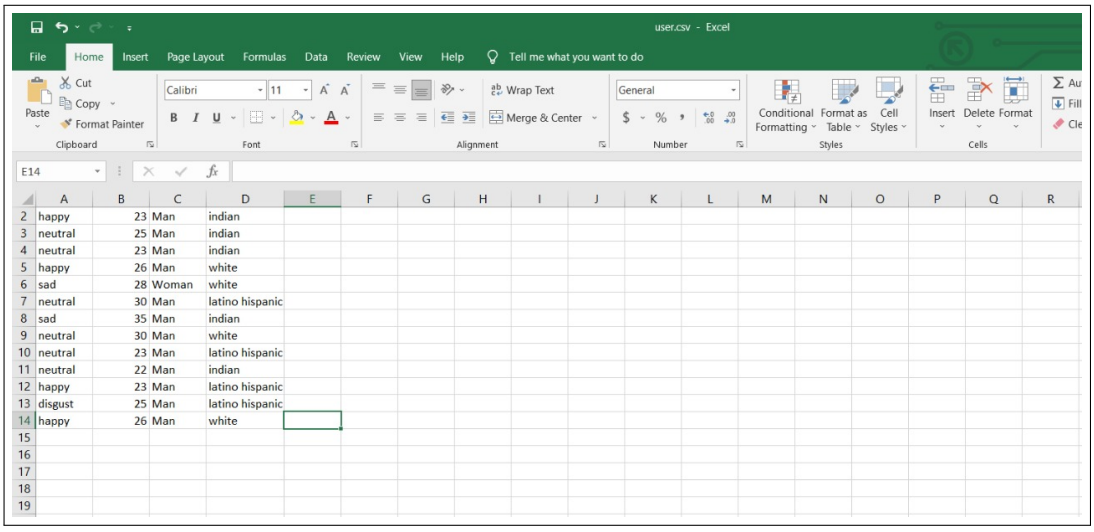


Figure 8.4: Output stored in CSV file

CHAPTER 9

CONCLUSIONS

9.1 CONCLUSION

Depression is a major problem faced by many people in today's world. More than 264 million people from all ages suffer from depression. Unlike sentiments it is not visible in external appearance, so many of the people don't even know that they are suffering from depression. One of the major cause of suicides occurring is depression, so detecting depression in people is very important. This system will detect depression with higher accuracy, so people could do the treatment as soon as possible without more affecting their health. Depression reduces productivity so many companies could use this project to keep a track on their employees health condition and could help them getting the treatment on time.

This system is divided into 2 parts, one is the PHQ-9 questionnaire and the other one is audio and visual samples collected while reading a paragraph. These 2 parts are very important to attain higher accuracy in determining the output. Output with less accuracy could create a false alarm so for higher accuracy these 2 steps should be done properly. With a good internet connectivity and good quality of camera and microphone, this application would be able to run efficiently with less lag possible.

9.2 FUTURE WORK

- Building an android application of the system
- Improving the efficiency and reliability of the system
- Making the system more stable with each update possible

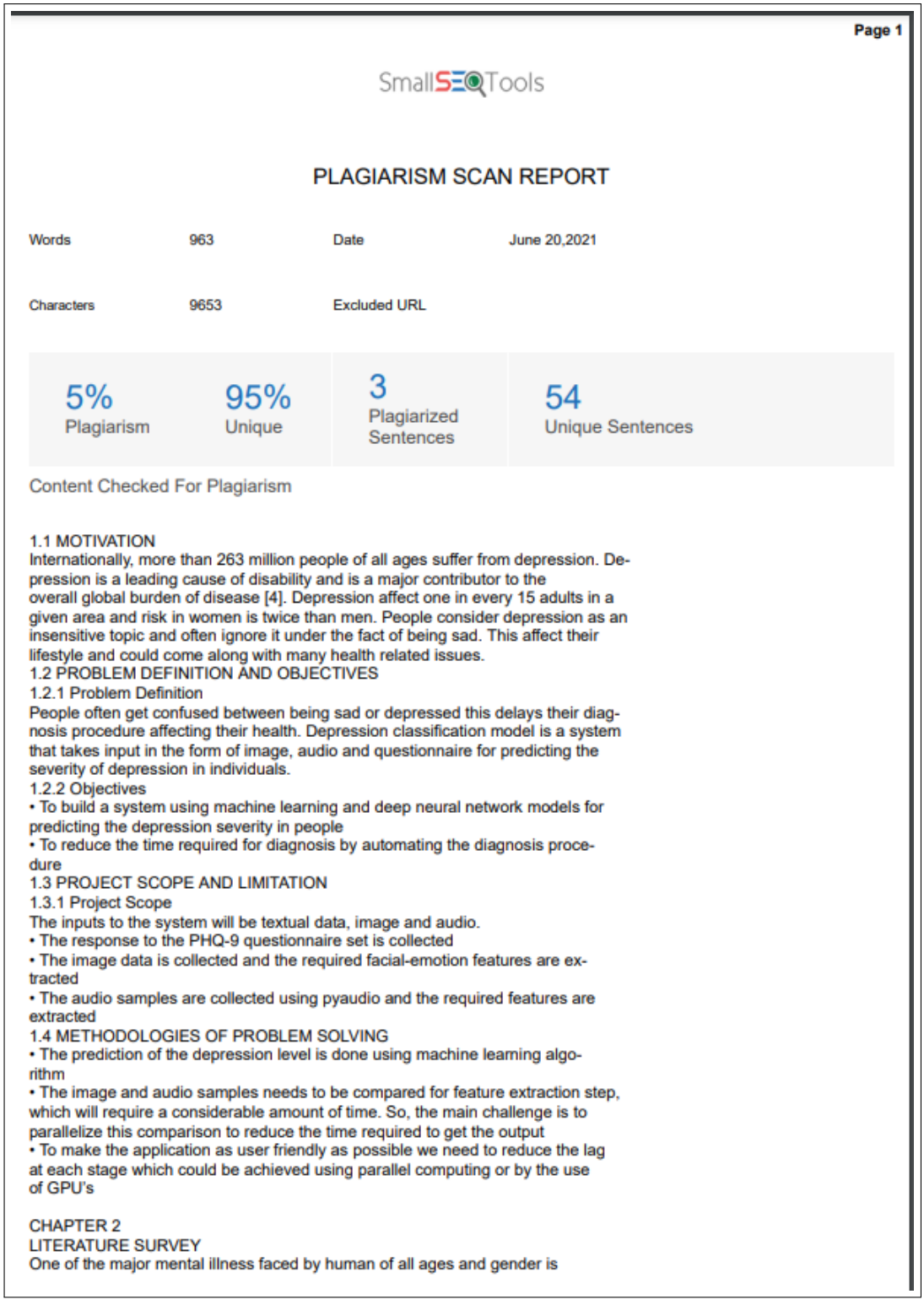
9.3 APPLICATIONS

- In companies, to keep a track on the employees health so that performance of the employee is maintained
- In schools and colleges to check whether the students are suffering from depression or not, if yes then the Institution could help the students to get over from depression

- People being anywhere could diagnose themselves at their convenience and later could refer to a professional

ANNEXURE A

PLAGIARISM REPORT



Page 1

SmallSEOTools

PLAGIARISM SCAN REPORT

Words	963	Date	June 20,2021
Characters	9653	Excluded URL	

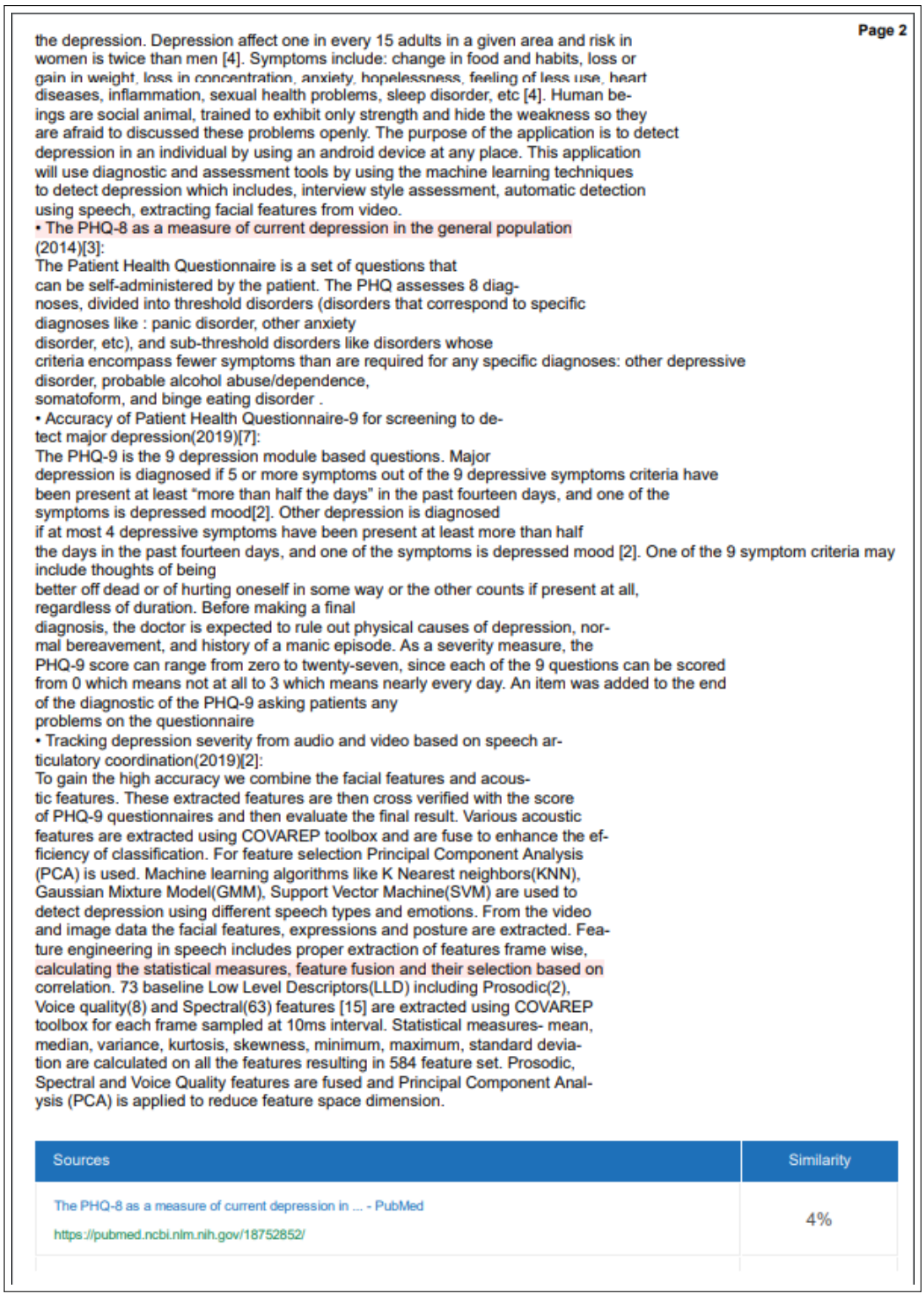
5% Plagiarism	95% Unique	3 Plagiarized Sentences	54 Unique Sentences

Content Checked For Plagiarism

1.1 MOTIVATION
Internationally, more than 263 million people of all ages suffer from depression. Depression is a leading cause of disability and is a major contributor to the overall global burden of disease [4]. Depression affect one in every 15 adults in a given area and risk in women is twice than men. People consider depression as an insensitive topic and often ignore it under the fact of being sad. This affect their lifestyle and could come along with many health related issues.
1.2 PROBLEM DEFINITION AND OBJECTIVES
1.2.1 Problem Definition
People often get confused between being sad or depressed this delays their diagnosis procedure affecting their health. Depression classification model is a system that takes input in the form of image, audio and questionnaire for predicting the severity of depression in individuals.
1.2.2 Objectives
• To build a system using machine learning and deep neural network models for predicting the depression severity in people
• To reduce the time required for diagnosis by automating the diagnosis procedure
1.3 PROJECT SCOPE AND LIMITATION
1.3.1 Project Scope
The inputs to the system will be textual data, image and audio.
• The response to the PHQ-9 questionnaire set is collected
• The image data is collected and the required facial-emotion features are extracted
• The audio samples are collected using pyaudio and the required features are extracted
1.4 METHODOLOGIES OF PROBLEM SOLVING
• The prediction of the depression level is done using machine learning algorithm
• The image and audio samples needs to be compared for feature extraction step, which will require a considerable amount of time. So, the main challenge is to parallelize this comparison to reduce the time required to get the output
• To make the application as user friendly as possible we need to reduce the lag at each stage which could be achieved using parallel computing or by the use of GPU's

CHAPTER 2
LITERATURE SURVEY
One of the major mental illness faced by human of all ages and gender is

Figure A.1: Plagiarism Report-1



Page 2

the depression. Depression affect one in every 15 adults in a given area and risk in women is twice than men [4]. Symptoms include: change in food and habits, loss or gain in weight, loss in concentration, anxiety, hopelessness, feeling of less use, heart diseases, inflammation, sexual health problems, sleep disorder, etc [4]. Human beings are social animal, trained to exhibit only strength and hide the weakness so they are afraid to discussed these problems openly. The purpose of the application is to detect depression in an individual by using an android device at any place. This application will use diagnostic and assessment tools by using the machine learning techniques to detect depression which includes, interview style assessment, automatic detection using speech, extracting facial features from video.

• The PHQ-8 as a measure of current depression in the general population (2014)[3]:
The Patient Health Questionnaire is a set of questions that can be self-administered by the patient. The PHQ assesses 8 diagnoses, divided into threshold disorders (disorders that correspond to specific diagnoses like : panic disorder, other anxiety disorder, etc), and sub-threshold disorders like disorders whose criteria encompass fewer symptoms than are required for any specific diagnoses: other depressive disorder, probable alcohol abuse/dependence, somatoform, and binge eating disorder .

• Accuracy of Patient Health Questionnaire-9 for screening to detect major depression(2019)[7]:
The PHQ-9 is the 9 depression module based questions. Major depression is diagnosed if 5 or more symptoms out of the 9 depressive symptoms criteria have been present at least "more than half the days" in the past fourteen days, and one of the symptoms is depressed mood[2]. Other depression is diagnosed if at most 4 depressive symptoms have been present at least more than half the days in the past fourteen days, and one of the symptoms is depressed mood [2]. One of the 9 symptom criteria may include thoughts of being better off dead or of hurting oneself in some way or the other counts if present at all, regardless of duration. Before making a final diagnosis, the doctor is expected to rule out physical causes of depression, normal bereavement, and history of a manic episode. As a severity measure, the PHQ-9 score can range from zero to twenty-seven, since each of the 9 questions can be scored from 0 which means not at all to 3 which means nearly every day. An item was added to the end of the diagnostic of the PHQ-9 asking patients any problems on the questionnaire

• Tracking depression severity from audio and video based on speech articulatory coordination(2019)[2]:
To gain the high accuracy we combine the facial features and acoustic features. These extracted features are then cross verified with the score of PHQ-9 questionnaires and then evaluate the final result. Various acoustic features are extracted using COVAREP toolbox and are fuse to enhance the efficiency of classification. For feature selection Principal Component Analysis (PCA) is used. Machine learning algorithms like K Nearest neighbors(KNN), Gaussian Mixture Model(GMM), Support Vector Machine(SVM) are used to detect depression using different speech types and emotions. From the video and image data the facial features, expressions and posture are extracted. Feature engineering in speech includes proper extraction of features frame wise, calculating the statistical measures, feature fusion and their selection based on correlation. 73 baseline Low Level Descriptors(LLD) including Prosodic(2), Voice quality(8) and Spectral(63) features [15] are extracted using COVAREP toolbox for each frame sampled at 10ms interval. Statistical measures- mean, median, variance, kurtosis, skewness, minimum, maximum, standard deviation are calculated on all the features resulting in 584 feature set. Prosodic, Spectral and Voice Quality features are fused and Principal Component Analysis (PCA) is applied to reduce feature space dimension.

Sources	Similarity
The PHQ-8 as a measure of current depression in ... - PubMed https://pubmed.ncbi.nlm.nih.gov/18752852/	4%

Figure A.2: Plagiarism Report-2

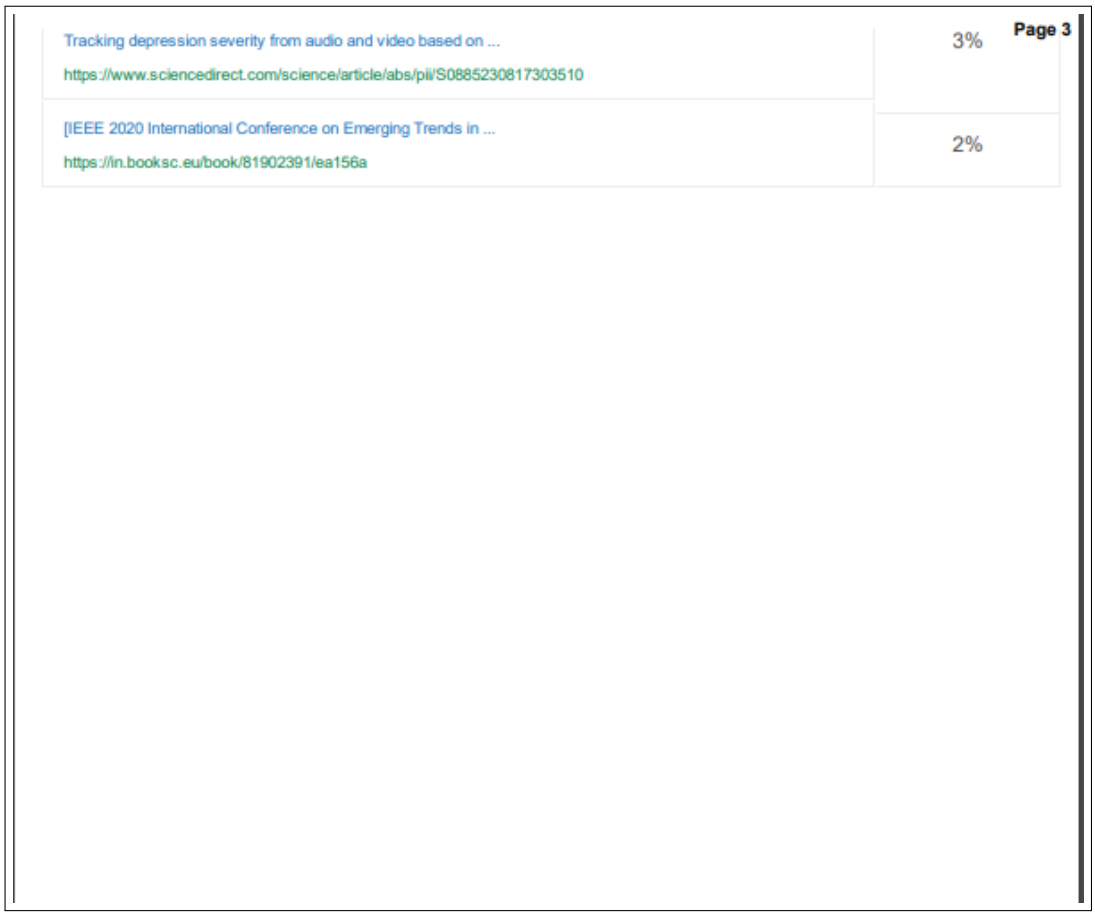


Figure A.3: Plagiarism Report-3

ANNEXURE B

REFERENCES

[1] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. F. Quatieri *"A review of depression and suicide risk assessment using speech analysis,"* Speech Communication, vol. 71, pp. 10–49, 2015

[2] J. R. Williamson, D. Young, A. A. Nierenberg, J. Niemi, B. S. Helfer, and T. F. Quatieri, *"Tracking depression severity from audio and video based on speech articulatory coordination,"* Computer Speech & Language, vol. 55, pp. 40–56, 2019.

[3] K. Kroenke, T. W. Strine, R. L. Spitzer, J. B. Williams, J. T. Berry, and A. H. Mokdad, *"The phq-8 as a measure of current depression in the general population,"* Journal of affective disorders, vol. 114, no. 1-3, pp. 163–173, 2014.

[4] https://www.who.int/news-room/fact-sheets/detail/depression

[5] https://www.kaggle.com/ashishbansal23/emotion-recognition

[6] G. Degottex, J. Kane, T. Drugman, T. Raitio, and S. Scherer, *"Co-varep—a collaborative voice analysis repository for speech technolo- gies,"* 2014 ieee international conference on acoustics, speech and signal processing (icassp). IEEE, 2014, pp.960–964.

[7] Brooke Levis, doctoral student1, Andrea Benedetti, associate professor2, Brett D Thombs *"Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression"* Published 09 April 2019