National University of Computer and Emerging Sciences

Suicide Ideation

Maryam Mukhtar …… L17-4080
Manha Zamir…...…… L17-4096
Husnain Akhtar…… L16-4366
Adeel Tahir……...……L16-4224

Supervisor: Dr. Muhammad Usman Awais

B.S. Computer Science
Final Year Project
April 2021

Anti-Plagiarism Declaration

This is to declare that the above publication produced under the:

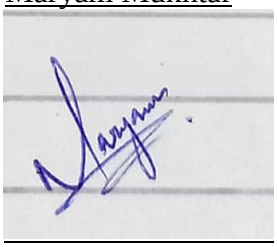
Title: Suicide Ideation

is the sole contribution of the author(s) and no part hereof has been reproduced on **as it is** basis (cut and paste) which can be considered as Plagiarism. All referenced parts have been used to argue the idea and have been cited properly. I/We will be responsible and liable for any consequence if violation of this declaration is determined.

Date: 30-12-2020

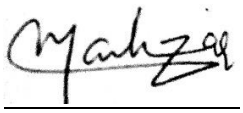
Student 1

Name: Maryam Mukhtar

Signature: 

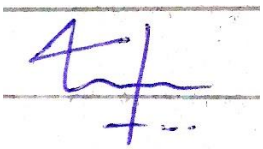
Student 2

Name: Manha Zamir

Signature: 

Student 3

Name: Adeel Tahir

Signature: 

Student 4

Name: Husnain Khan

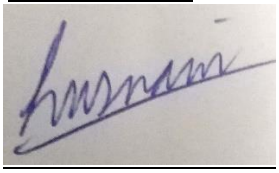
Signature: 

Table of Contents

List of Tables

List of Figures

Abstract

Suicide is a critical issue in our modern society. Early detection of the negative emotions that may lead to suicide can have a considerate impact on a person. Such diagnosis and prevention of suicide attempts should be addressed to save people's life. Our android application, Suicide Ideation, aims for early detection of an individual's inclination towards suicide. Such mental state is not a mental illness but a potential consequence of various mental disorders involving depression, anxiety and stress. These disorders could act as warning signs and can contribute to produce valuable information, used by our system, to find the level of severity of a person's depression, anxiety and stress. For this purpose, the system takes into account the use of DASS42 scale to check the severity levels of each of the negative emotional state. Moreover, in order to predict the if an individual is at suicidal risk our not, we use an NLP Model combined by TF-IDF Vectorizer and Multinomial Naïve Bayes. This model is trained on the dataset collected from an online platform Reddit, and is combination of two subreddits r/depression and r/suicidewatch, both corresponding to depression and suicide, respectively. It contains posts and comments that refer to suicidal texts and thoughts of people. Apart from this, users can interact with the system that evaluates these risk factors through speech-to-speech conversation using Urdu NLP webservices. It provides the results and keeps a history of these records as well to observe the change in their thought processes through these records. Hence, our application could help detect individuals at suicidal risks along with early diagnosis of mental disorders including depression, anxiety and stress.

Chapter 1: Introduction

Suicide defines as the act of deliberately causing one's own death. Every year, nearly 800,000 people die due to suicide, which is one person every 40 seconds. Suicide is more common among young people, and according to a 2017 WHO survey, 1.4% of all deaths were from suicide [1]. A significant problem in our community has been the growing suicide rate over the past few years.

Suicide ideation is a pre requisite of suicide attempts. Not all suicide ideators tend to actually go through the attempt. Nevertheless, both of them do not differ in their acquired capabilities for suicide. The suicidal thoughts still validate enough risk for early diagnosis and prevention. It can be preventable which all starts by knowing what to look for and what to do. Some of the major reasons behind suicide ideation is anxiety, depression and stress. Considering this, designing a digital android-based solution could really help these people through diagnosis.

The project includes an android application that identifies whether a person is suffering from any suicidal tendencies or not. Moreover, it calculates the severity levels of depression, anxiety and stress for an individual as well. Initially, a conversation between a Chabot and the user would take place by asking a couple of questions related to suicide and each of the above three states. The conversation for Depression, Anxiety and Stress will be based on DASS dataset as for suicide ideation text will be predicted using a trained model. The means of communication between the user and the Bot will be through an interactive voice response system in Urdu using CLE NLP webservices [2]. After the conversation has ended, the system will calculate the results depicting the rates of depression, anxiety, stress or suicidal risk, depending on the type of test chosen.
Finally, based on user's responses, our machine learning model will predict whether the user is at suicidal risk or not. User can also view their history for attempted tests and a therapist recommendation list for consultation.
The first chapter comprises of the introduction followed by goals and scope of our project. The second chapter includes literature review. Our project’s requirement and design are discussed in the third chapter including functional and nonfunctional requirements, system architecture, database design, use cases and GUIS etc. The fourth chapter discusses prototype implementation and the last chapter includes conclusion.

1.1 Goals and Objectives

The primary goals and objectives of this project includes

- Providing a human-type communication bot which will analyze user’s depression, anxiety, stress and suicide risk.
- Providing on time prevention to users through our diagnosis of the aspects related to emotional disturbance leading to suicidal tendencies
- A measure for each of the three states by an evaluative conversation between user and our AI bot
- Analyzing and predicting the concerned issues using Machine Learning algorithm
- Maintaining user’s previous records of test to observe change in results with the passage of time

1.2 Scope of the Project

The scope of our project is limited as:

- Predicting user's severity level for depression, anxiety and stress scales using machine learning
- Predicting suicide ideation using an NLP trained model
- Providing interactive communication between user and AI model by integrating Urdu CLE NLP web-services
- NLP involves:
    - English text to Urdu text
    - Urdu text to Urdu speech
    - Urdu speech to Urdu text
    - Urdu text to English text
- Checking history for the previously attempted tests
- Therapist recommendation for consultation
- No provision of treatment in the application domain

Chapter 2: Literature Survey / Related Work

The use of artificial intelligence and machine learning methods to predict the suicidal tendencies as compared to the clinical techniques has taken over a rise over the past few years. With the advancement in technology, several methods have been applied to identify the risks involving for an individual carrying a mental state of such sort. This involves researches for data related to medical as well as social suicide prediction. Our concern is the social suicide prediction, in which researchers have applied machine learning methods on data available on social media platforms representing suicidal tendencies and risks [3]. Apart from it, our work also includes finding the severity levels of depression, anxiety and stress. Some of the related work is as follows:

2.1 Assessment of Anxiety, Depression and Stress using Machine Learning Models

In [3] eight different machine learning algorithms have been applied on two different databases related to depression, anxiety and stress to identify the severity levels for each category. The main target was to find which of the eight applied methods result in greater accuracy in identifying the severity level i.e., Extremely severe, severe, moderate, mild and normal. Each level of category corresponds to a class that is, five classes for five levels. Hence, the study corresponds to a multiclass classification.

2.1.1 Dataset

The datasets used for this study are DASS42 and DASS21. A total of 39776 user instances are included in DASS42 collected online. The data collection for DASS21 was also online using google forms. As the name suggests, it includes 42 and 21 questions respectively. For each question, the responses are scaled from 1 to 4:
1 – Did not apply to me at all
2 – Applied to some degree, or some of the time
3 – Applied to a considerable degree, or a good part of the time
4- Applied to me very much, or most of the time

The total score was calculated for each response to label the dataset based on the severity levels.

2.1.2 Method

As discussed before, this study uses multiclass classification to predict the five levels for depression, anxiety and stress. The eight algorithms belong to the family of probabilistic, neural network, nearest neighbor and tree-based classification. Hybrid techniques were also applied to visualize the efficiency of the training model compared to single classifiers.

2.1.3 Results

For DASS42, neural network worked the best in showing the accurate results as compared to the other groups. Among the neural network, Radial Basis Function Network (RBFN) showed the best results (97.48 for Anxiety, 96.02 for Depression and 96.12 for Stress). After this, the hybrid methods showed better results but were considered time consuming.
For DASS21, some differences were recorded due to a smaller number of instances for this dataset. According to the research, random forest classifier showed 100% accuracy for anxiety, MLP and RBFN showed 96.55 accuracy for stress and depression, respectively.

2.2 Exploratory Analysis of Social Media Prior to a Suicide Attempt

In this research [4] focuses on providing the exploratory analysis for the patterns in the language revolved around suicide attempt. It highlights the emotions indulged in the attempt by the users and how such users can be identified using machine learning. Data from online social media platform 'Twitter' has been analyzed to monitor these alarming behaviors and emotions by quantifying the language. This helped in determining patterns of the users who have survived one or more suicidal attempts. The patterns were majorly directed in distinguishing users who attempt suicide, contrary to a neurotypical person. Also, the difference between a user before and after the attempt.

2.2.1 Dataset

The dataset used contains the public statements of the users who addressed their suicidal attempts on twitter. It was observed that their sample data did not contain people apart from youth. Only public data was observed containing 554 direct statements from people who have tried to take their own life. Female users were recorded more than male, most likely aged 15-29.

2.2.2 Analysis

Some analysis was found prior to the prediction through exploratory data analysis. A comparison was made on the type of user that tried to attempt suicide. A user matched to their neurotype controls and a user prior to and after taking their life. Following observations were drawn in [4]:

- Neurologically typical people while giving their statements on Twitter tend to use emoticons more than a user before an attempt
- Users talk more about suicide after trying to take their life than before the attempt
- Users tend to use more self-focused language (similar to the findings in case of depression) before attempt

2.2.3 Method

For this research, they diverge from most NLP since it discards the stop words from the data, which in this case is important to keep track of due their significance in verifying the time of their experience and carrying psychological meaning.
For character language models, character n gram language models followed by logistic regression via scikit learn were applied. 10-fold cross validation was applied while calculating the measures for the model. To observe the emotional states, a novel corpus with automatically induced emotion labels were collected which resulted in guilt being a prominent emotion for people attempting suicide.

2.2.4 Results

The results suggested that the users counted as suicide attempters differ from the rest of the population. About 70% of the results correctly identified the suicide attempters. As far as the emotional states are concerned, tweets from the people who took their own life were more emotional than those who didn't. Users exhibiting emotions like anger and sadness in the tweets before the attempt tend to match neurologically typical people more after the attempt. Moreover, the timeline for their emotions before the attempt shows sadness for a time, which later is mixed with anger as they come close to the time suicide was attempted. Moreover, suicide attempters used less emoticons in their tweets comparatively. Hence, indicating that

even though these users carry more emotions than neurotypicals, they still use less emojis to express their feelings through the tweets.

2.3 Supervised Learning for Suicidal Ideation Detection in Online User Content

This experimental study focuses on suicide ideation based on the online content provided by users on Reddit and Twitter. It involves feature extraction for six different sets along with a comparative analysis of six supervised machine learning algorithms to discern suicidal tendencies in the content. It delivers the knowledge of suicidal thoughts from a data analytic perspective [5]. Instead of using simple benchmarks for the identification of suicide ideation, this research uses statistical, topic, linguistic, syntactic and word embedding features [5] for detection. Different classifiers are used to compare the results including traditional classifiers and neural networks.

2.3.1 Datasets

The datasets have been collected from the online social platform Reddit, with a subreddit called SuicideWatch (SW). The content involved a title and a body. The data instances falling under this subreddit were categorized as “Suicide texts”. Apart from it, some non-suicide texts were also garnered from other general subreddits. Reddit datasets are divided into two categories. One contains the combination of the subreddit SuicideWatch and some other subreddit. The second contains a combination of six subreddits, one of which being SW.
The other type of dataset is from Twitter which was collected from a keyword filtering technique. It comprises of 10,288 tweets with 594 about suicide ideation.

2.3.2 Method

Following are the methods involved for this research.

2.3.2.1 Feature Extraction

Six sets of features were extracted for our model. Statistical features were carried out to provide the number of words, tokens, sentences in the body. Parts of speech (POS) were extracted a feature for syntactic attribute of the texts. Some linguistic statistical information was extracted by LIWC that covered pronouns, emotions, tense, work, money, swear words, female and male preferences for suicide and non-suicide texts. TF-IDF was employed as a word frequency feature to calculate the number of times a word appears in texts. For word-embedding, word2vec was used to devise the semantic representation.

2.3.2.2 Classification

Six classification models were used including SVM, Logistic Regression, Random forest, LSTM, GBDT and XGBoost. A comparison between these classifiers was made to detect suicide ideation with greater accuracy.

2.3.3 Result

According to the research, XGBoost resulted in best performance (96.56%) as compared to the other supervised learning models on the first type of reddit dataset. LSTM showed the worst performance since it does not require feature extraction. The rest showed better accuracy than that of LSTM. In the second type of reddit dataset, XGBoost resulted in better performance in two of the five categories. LSTM and Random Forest showed more accurate results in the rest. In case of Twitter data, Random Forest outperforms all the other machine learning models.

Hence, suicide ideation can be detected by using the classification models proposed for best accuracy.

From the above related work, we can see that machine learning techniques can be applied under this domain to demonstrate suicidal ideation and its dependencies. Many researchers are focused with such findings related to social media data as it covers data globally, redefining the constraints that might be applied due to limited area of research. Hence, ML prediction for suicidal tendencies show greater significance in influencing related outcomes.

Chapter 3: Requirements and Design

All the functional and nonfunctional requirements are discussed in this chapter along with software and hardware requirements of our project. System architecture, database design, all the user modules and use cases describing the flow of application are also being discussed. It also includes the Graphical User Interface of the application provided below.

3.1 Functional Requirements

Following are the functional requirements of our system.

3.1.1 Functional Requirements for General Users

- The system allows the user to access the application if registered
- The system allows the user to change password/retrieve old password
- The system allows the user to take a Depression, Anxiety and Stress test
- The system allows the user to respond to the conversation through speech in Urdu
- The system allows the user to check their results for the attempted test
- The system allows the user to view profile
- The system allows the user to view history for their previous test records
- The system allows the user to contact a therapist from the provided list

3.1.2 Functional Requirements for System

- The system generates credentials and authenticates user.
- The system updates password
- The system asks the user some questions regarding the test type they chose through text and speech
- The system calculates the score for each test taken
- The system displays the scores for the test taken
- The system displays the severity levels of depression, anxiety and stress
- The system keeps a record of the tests on the dates attempted
- The system refers to therapist recommendation

3.2 Non-Functional Requirements

Following are the nonfunctional requirements of our system.

3.2.1 Sustainability

- The system is accessible from play store and Android device meeting the specified hardware and software requirements

3.2.2 Performance

- The system should not take more than 3 seconds to load the initial screen
- The system should predict the final result within 5 sec

3.2.3 Usability

- System will have an intuitive and user-friendly design, making it easy to use
- Users would be able to interact/response the system within 2 seconds of viewing the interface

3.2.4 Reliability

The system will be available 24/7 to the users.

3.3 Hardware and Software Requirements

Following are the hardware and software requirements of our project

3.3.1 Hardware Requirements

For development, the hardware requirements are the following:

- Desktop with 4GB RAM and 64 bit operating system
- High speed online servers
- Ethernet connection or wireless adapted (Wifi)

For usage, the hardware requirements are the following:

- Internet speed of at least 2Mbps required.

3.3.2 Software Requirements

For development, the software requirements are the following:

- Python: 3.6.8
- Pycharm/Jupyter notebook
- Android Studio
- Firebase DB server
- Python libraries (NumPy (>= 1.13.3 SciPy (>= 0.19.1), Scikit-learn, Pandas)
- CLE NLP web services

3.4 System Architecture

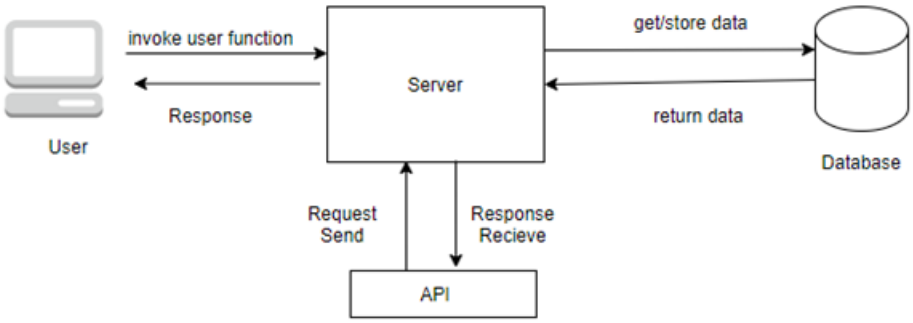


Figure 1: High Level System Architecture

Diagram for high level system architecture

As shown in the high-level system architecture, we will be implementing client-server architecture where user can invoke a particular function to the server for it to perform. The server contains two main modules (explained in the Low-level architecture), that interacts with the Database and NLP API. The database contains the required data needed to corresponding to the execution of the requests made by user. The API would send and receive

requests from server to perform its tasks. After performing the requested function, the server shall return the response to the user showing the status of their request.

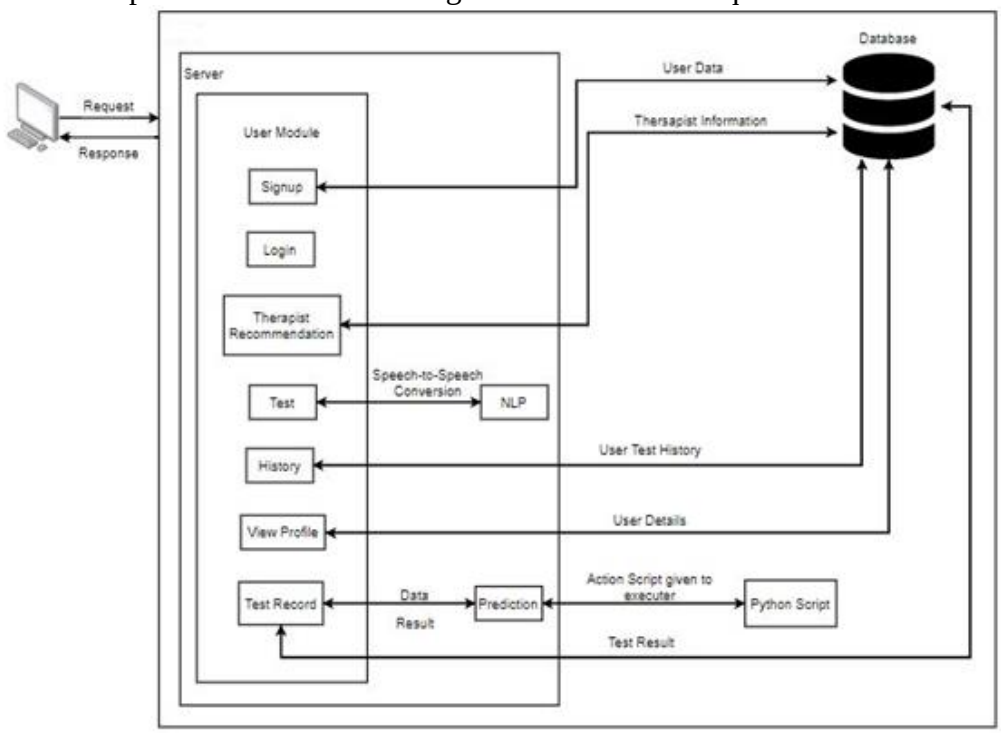


Figure 2: Low Level System Architecture

Diagram for low level system architecture

The low-level architecture contains the following aspects:

3.4.1 User Module

The user module that contains the requests the user will send to the server to perform said operations. The user needs to authenticate their account through the database before having access to other services of the system.

3.4.2 Database

Our database is a firebase database that contains the following components:

- User record
- History
- Test record
- Therapists

This data can be accessed through the server upon the functions that carry these requests.

3.4.3 NLP

In order to take a test, the server will access the NLP API for speech-to-speech conversation.

3.4.4 User Module

While the conversation takes place, the user's response shall be predicted in the 'Prediction' that will execute these actions based on the received data. Once all the predictions are completed, the results generated are saved in the database to be viewed by the user upon the related operation selection.

3.4.5 Subsystem Architecture

Details of all the subcomponents along with their diagrams are shown below.

3.4.5.1 Signup Component

The signup component contains is available to users before accessing the application. This component requires the validation sub-component to show that the user is authenticated. These results are stored in the Storage component later.

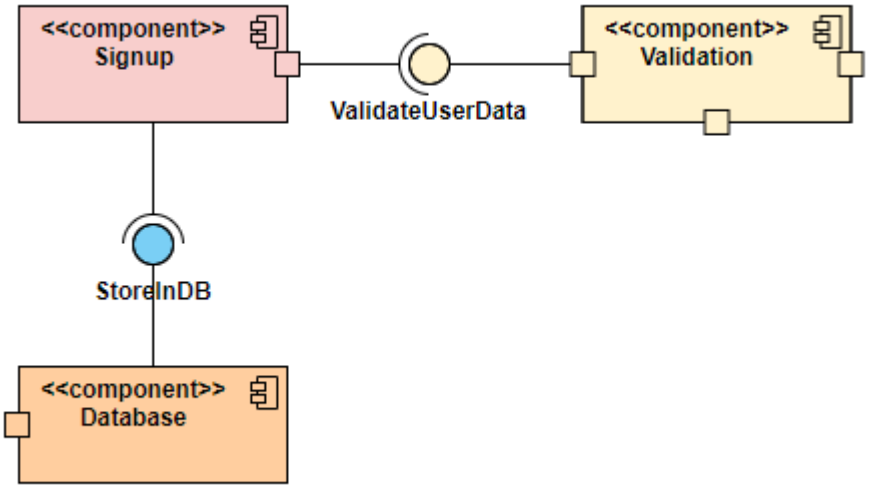


Figure 3: Signup Component
Component Diagram of Signup

3.4.5.2 Therapist Recommendation Component

This component can be used by the user to access the list of therapists stored in the database.

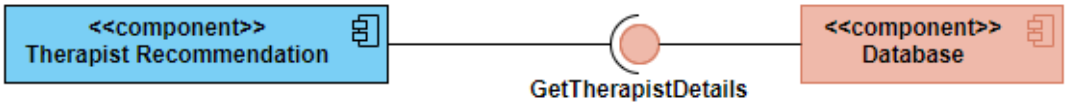


Figure 4: Therapist Recommendation Component
Component Diagram of Therapist

3.4.5.3 Test Component

The test component is used to access the three types of test available to users. These include Depression, Anxiety and Stress Test. It also contains Suicide Ideation test based on the three types. Here, the test record contains the results for each test after prediction.

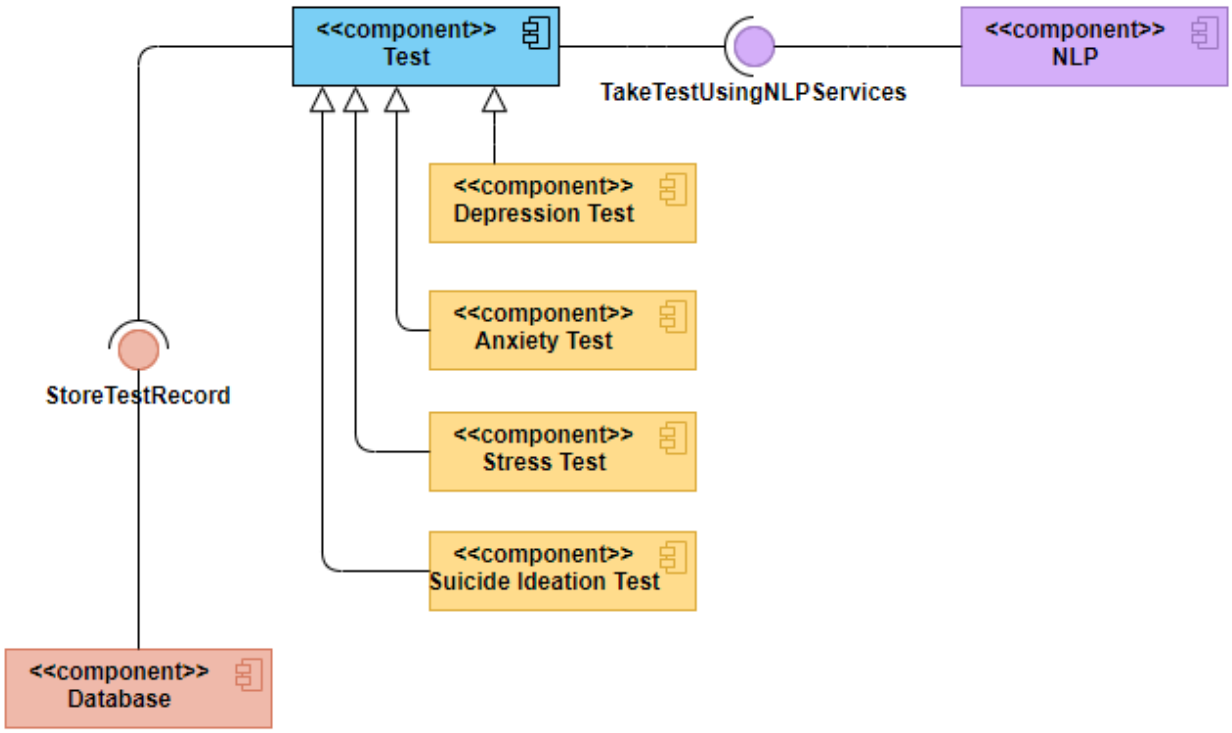


Figure 5: Test Component
Component Diagram of Test

3.4.5.4 History Component

This component contains the results of the attempted tests of any type by the user. It accesses the database which itself has the stored results from Test Record component.

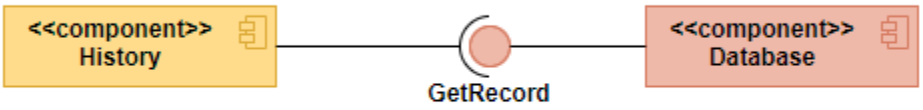


Figure 6: History Component
Component Diagram of History

3.4.5.5 View Result Component

In this component, user can view the scores for Depression, Anxiety, Stress and Suicide Ideation scores. These scores are attained through the test component.

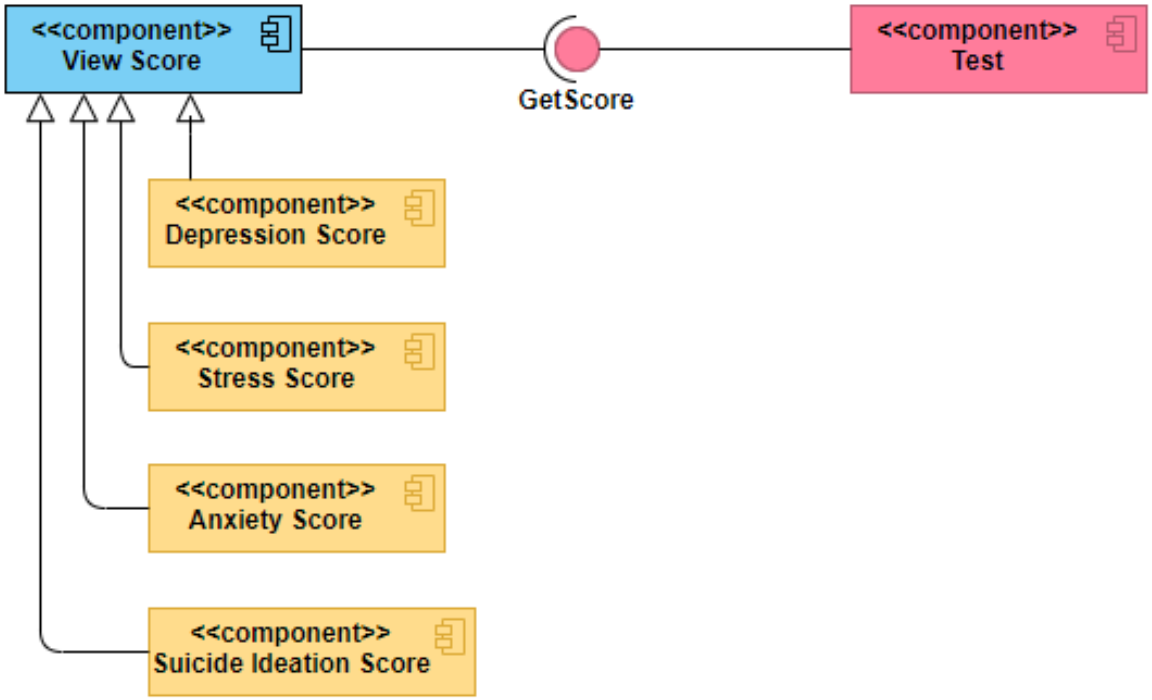


Figure 7: View Result Component
Component Diagram of View Result

3.5 Architectural Strategies

Following architectural strategies are considered for our approach.

3.5.1 Android Studio, Scikit-Learn, Firebase

We are using Android Studio because it makes mobile app development easy because of its open-source platform. Android Studio also has better User Interface and is also considered a stable IDE. The frontend will be created through XML, while backend will be done in Java. Also, we are using Scikit-Learn - Multinomial Naïve Bayes to train our model. Scikit-learn is a python module that provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python [7]. It can be used with Python. Firebase database is a real-time database that stores data in a JSON object and provide real-time data to all connected clients.

3.5.2 Future for enhancing software

Currently, for the scope of our final year project we have only created an android app however in future we plan on creating a fully functional web application that not only predicts user suicide tendencies but also recommend health practitioners based on user's location.

3.5.3 User interface paradigm

For the user interface, the eight golden rules for interface designing would be utilized which are part of the standards of human computer interaction laid by Ben Schneiderman. This is essential for implementation of a good UI design. Overall interface will be user friendly so that novice users will not have to face any difficulties to use our system.

3.5.4 Error detection and recovery

If a user carries out any restricted action, in that case a proper error handling will be carried out by giving him appropriate response and facilitating the user to correct his mistakes and perform the action again.

3.5.5 Memory management policies

Our system will not store any user data offline. Users should have an internet connection to use our system and all the data will be fetched from database placed on cloud at runtime.

3.6 Use Cases

All the use cases and their tables are mentioned below.

3.6.1 Login

Name	Login
Actors	User
Summary	The user shall provide their email and password on the login form and after successful verification, redirect the user to the home page.
Pre-Conditions	The user must be in the database records. The user must not already be logged in.
Post-Conditions	The user's session is successfully established and shall be redirected to the home page.
Special Requirements	None

Basic Flow			
Actor Action		**System Response**	
1	The user opens the login page.	2	The login page is displayed asking for email and password.
3	The user enters valid email and password.	4	The system verifies the email and password, establishes a session for the user and redirects the user to the home page.
Alternative Flow			
3	The user enters invalid email or password.	4-A	The system responds with an error message: *Incorrect email or password entered.*

3.6.2 Signup

Name	Signup		
Actors	User		
Summary	The user shall provide their required credentials on the signup form and after successful verification, redirect the user to the login page.		
Pre-Conditions	User must not have an account already registered with the email.		
Post-Conditions	The user's session is successfully established and shall be redirected to the home page.		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user opens the signup page.	2	The signup page is displayed asking for a full name, email and password.
3	The user enters a valid name	4	The system verifies it
5	The user enters a valid email	6	System verifies it
7	The user selects Gender (Male, Female or Other)	8	Gender Selected
9	Enter Age	10	System verifies
11	The user enters a valid username	12	System verifies
13	The user enters a valid password	14	System verifies and establishes a session for the user redirects to the home page
Alternative Flow			
3	The user enters an invalid name	4-A	The system responds with an error message: *Name can only comprise characters*
5	The user enters an email already in use	6-A	The system responds with an error message: *Email already in use*
5	The user enters an invalid email	6-B	The system responds with an error: *Invalid email address*
11	The user enters an invalid username	12-A	The system responds with an error: *Username already exists*
13	The user enters an invalid password	14-A	The system responds with an error: *Password must have at least one uppercase letter(s)*

3.6.3 Forget Password

Name	Forget Password		
Actors	User		
Summary	The user can reset his password if he has forgotten his account password		
Pre-Conditions	The user must be in the database records. The user must not already be logged in.		
Post-Conditions	User's account password will be updated		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the "Forget Password" option	2	System prompts the user to enter email id.
3	User enters email id	4	System sends verification pin to registered email id
5	User enters the PIN sent by system	6	System prompts the user to enter new password
7	User enters new password	8	System updates the user password
Alternative Flow			
3	The user enters an invalid email	4-A	The system responds with an error: *Invalid email address*
7	User enters invalid password	8-A	System responds with an error: *Password must contain at least one uppercase letter(s)*

3.6.4 View Profile

Name	View Profile		
Actors	User		
Summary	The user can view their profile details		
Pre-Conditions	The user must be logged in		
Post-Conditions	The user can change password or redirect back to home page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the "View Profile"	2	The system displays user details including Name, Username, Email, Age, Gender and Password alongwith an option to 'Change Password'
No Alternative Flow			

3.6.5 Change Password

Name	Change Password		
Actors	User		
Summary	The user can reset his password if he has forgotten his account password		
Pre-Conditions	The user must be logged in and have selected view profile.		
Post-Conditions	User's account password will be updated		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the "Change Password" option	2	Reset password page is displayed asking for new password
3	User enters new password	4	System verifies password and update is made
Alternative Flow			
3	The user enters an invalid password	4-A	The system responds with an error: *Password must have at least one uppercase letter(s)*

3.6.6 Logout

Name	Logout		
Actors	User		
Summary	The user can log out of their account		
Pre-Conditions	The user must already be logged in.		
Post-Conditions	The user shall be redirected to login page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user logouts	2	Login page is redirected
No Alternative Flow			

3.6.7 Depression Test

Name	Depression Test		
Actors	User		
Summary	The user shall select the ‘Depression Test’ type		
Pre-Conditions	The user must be logged in		
Post-Conditions	The user will be redirected to the conversation page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the ‘Depression’ test type	2	The system redirects to the conversation page
No Alternative Flow			

3.6.8 View Depression Score

Name	View Depression Score		
Actors	User		
Summary	The user can check their depression rate based on all of the attempted tests		
Pre-Conditions	The user must attempt Depression Test		
Post-Conditions	The user can redirect to depression test score display page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the “View depression score”	2	The system displays depression rate based on set of previously asked questions.
No Alternative Flow			

3.6.9 Anxiety Test

Name	Anxiety Test		
Actors	User		
Summary	The user shall select the ‘Anxiety Test’ type		
Pre-Conditions	The user must be logged in		
Post-Conditions	The user will be redirected to the conversation page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the ‘Anxiety’ test type	2	The system redirects to the conversation page
No Alternative Flow			

3.6.10 View Anxiety Score

Name	View Anxiety Score		
Actors	User		
Summary	The user can check their anxiety rate based on all of the attempted tests		
Pre-Conditions	The user must attempt Anxiety Test		
Post-Conditions	The user can redirect to anxiety test score display page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the “View anxiety score”	2	The system displays anxiety rate based on set of previously asked questions.
No Alternative Flow			

3.6.11 Stress Test

Name	Stress Test		
Actors	User		
Summary	The user shall select the ‘Stress Test’ type		
Pre-Conditions	The user must be logged in		
Post-Conditions	The user will be redirected to the conversation page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the ‘Stress test type	2	The system redirects to the conversation page
No Alternative Flow			

3.6.12 View Stress Score

Name	View Stress Score		
Actors	User		
Summary	The user can check their stress rate based on all of the attempted tests		
Pre-Conditions	The user must attempt Stress Test		
Post-Conditions	The user can redirect back to stress test score display page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the “View stress score”	2	The system displays stress rate based on set of previously asked questions.
No Alternative Flow			

3.6.13 View Suicide Ideation Rate

Name	View Suicide Ideation Rate		
Actors	User		
Summary	The user can check their suicide ideation rate based on all of the attempted tests		
Pre-Conditions	The user must attempt Depression Test, Anxiety Test and Stress Test		
Post-Conditions	The user can redirect back to home page		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects the ‘Stress test type	2	The system redirects to the conversation page
No Alternative Flow			

3.6.14 View History

Name	View History		
Actors	User		
Summary	The user shall view all their records for the attempted tests		
Pre-Conditions	The user must be logged in		
Post-Conditions	The scores for the attempted tests will be displayed or ‘No history found’ message shall be displayed if no tests have been attempted previously		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects ‘View History’	2	Redirects to ‘History’ page
No Alternative Flow			

3.6.15 Contact Therapist

Name	Contact therapist		
Actors	User		
Summary	The user can contact a therapist from the system’s recommendations based on results		
Pre-Conditions	The user must view their results		
Post-Conditions	The system redirects to a page to show the therapist list for contact		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects ‘Contact Therapist”	2	Redirects to a new page to view the list for recommendations
No Alternative Flow			

3.6.16 Select Therapist

Name	Select therapist		
Actors	User		
Summary	The user can select a therapist of their choice from the list for details		
Pre-Conditions	The user must select 'Contact therapist'		
Post-Conditions	The system can redirect to home page or logout		
Special Requirements	None		
Basic Flow			
Actor Action		**System Response**	
1	The user selects preferred therapist from list	2	Shows the details of the selected therapist
No Alternative Flow			

3.7 GUI

This section includes the graphical user interface of the application.

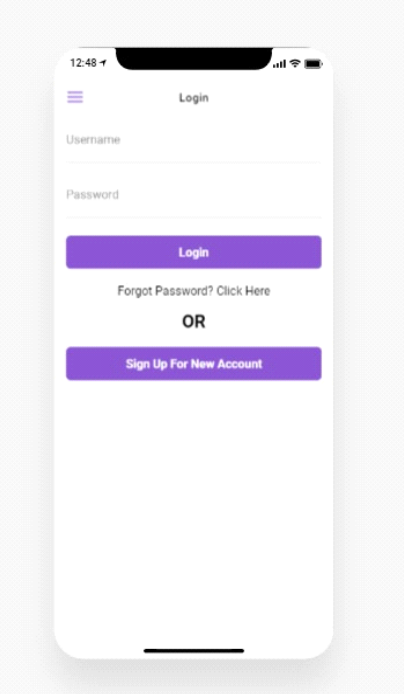


Figure 8: Login
This GUI shows login page for user

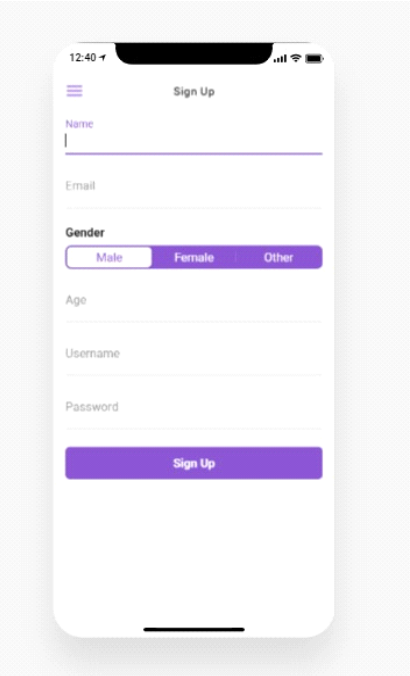


Figure 9: Signup
This GUI shows Signup page for user

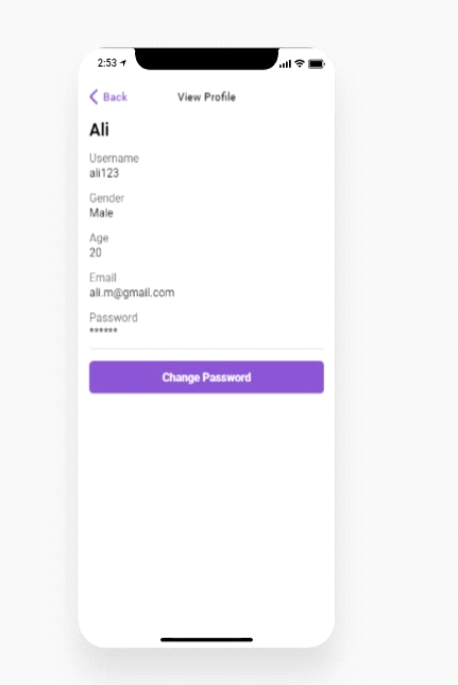


Figure 10: View Profile
This GUI shows user's profile details

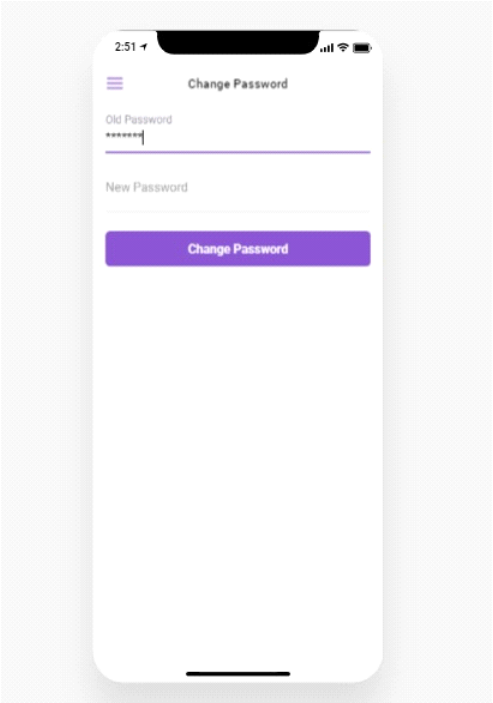


Figure 11: Change Password
This GUI shows the requirements to change password for user

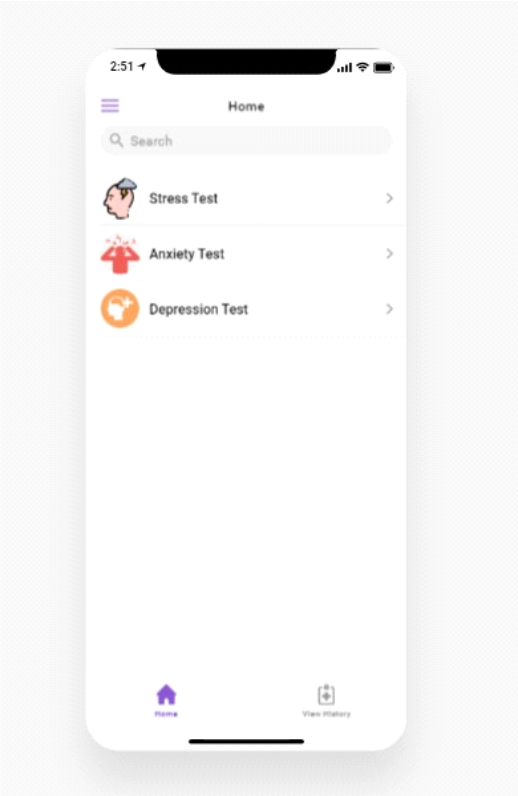


Figure 12: Home page (Screen 1)
This GUI shows home page (Screen 1)

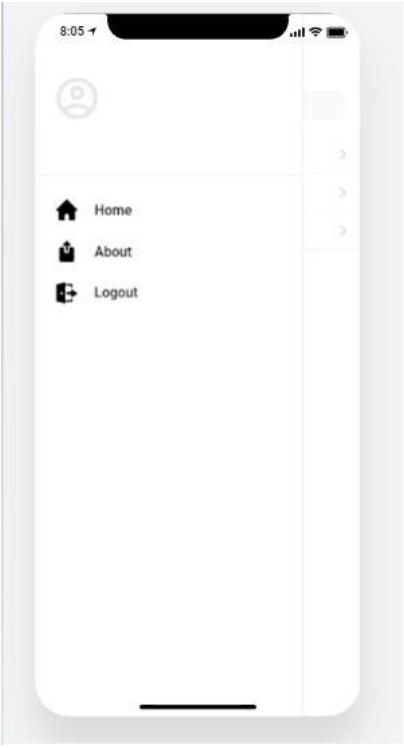


Figure 13: Home page (Screen 2)
This GUI shows home page (Screen 2)

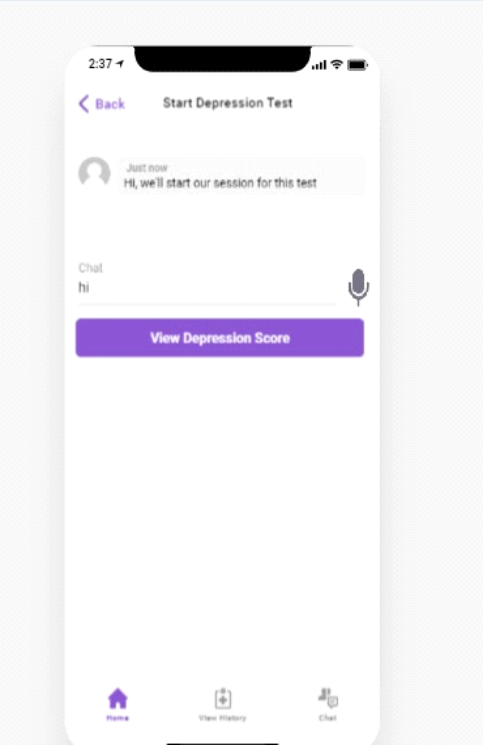


Figure 14: Chat for Depression Test
This GUI shows the conversation chat for Depression Test

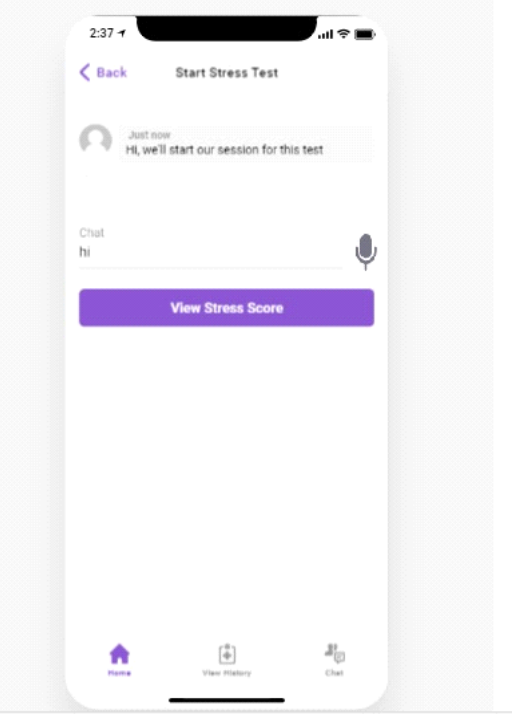


Figure 15: Chat for Stress Test
This GUI shows the conversation chat for Stress Test

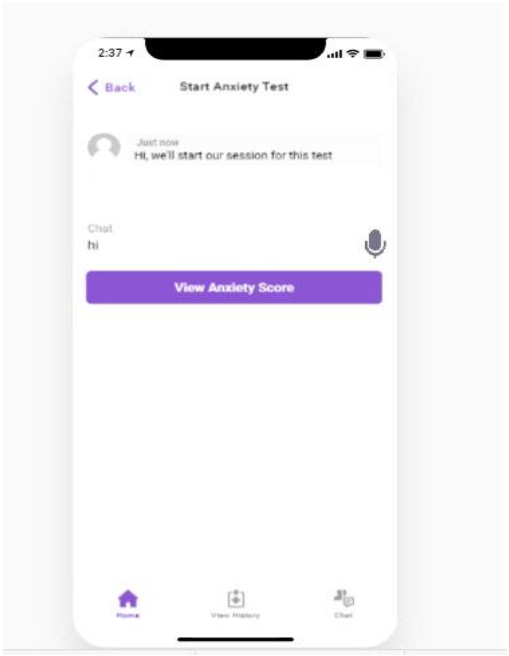


Figure 16: Chat for Anxiety Test
This GUI shows the conversation chat for Anxiety Test

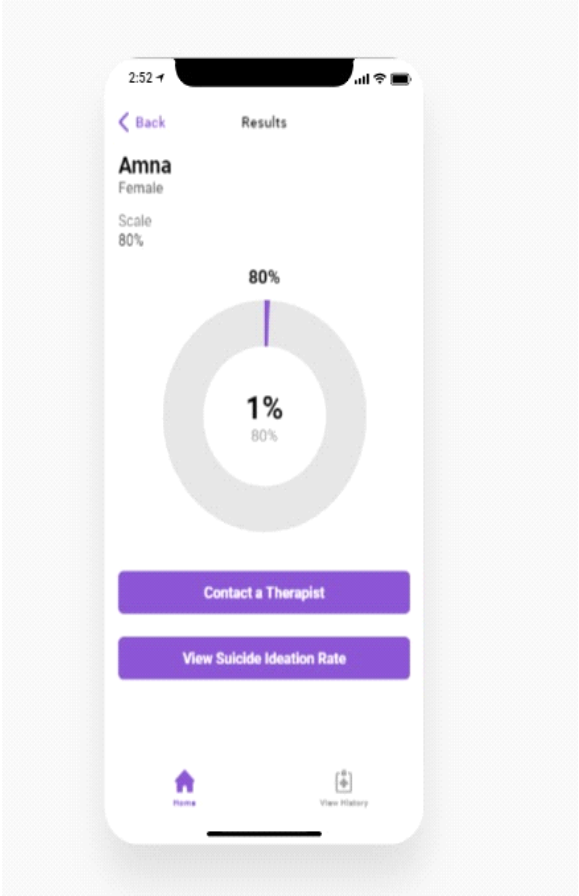


Figure 17: View Depression/Anxiety/Stress Score
This GUI shows the results for the selected test

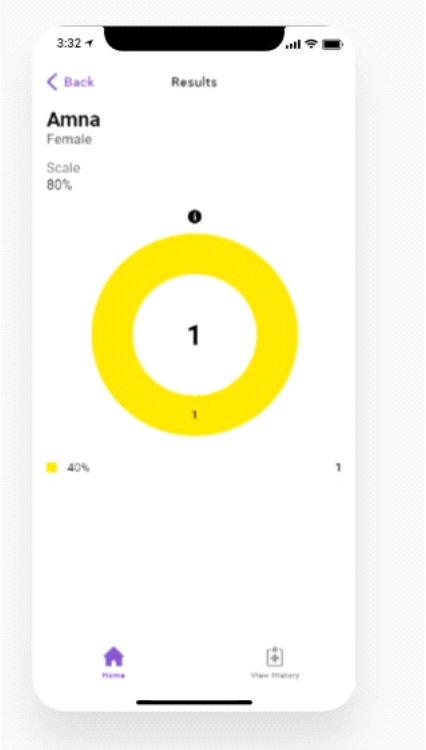


Figure 18: View Suicide Ideation Rate
This GUI shows the 'Suicide Ideation rate' results if all tests have been attempted

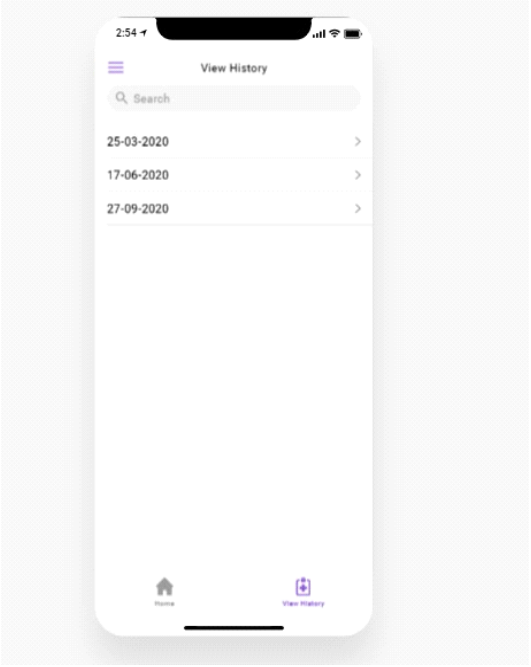


Figure 19: View History
This GUI shows History page for user

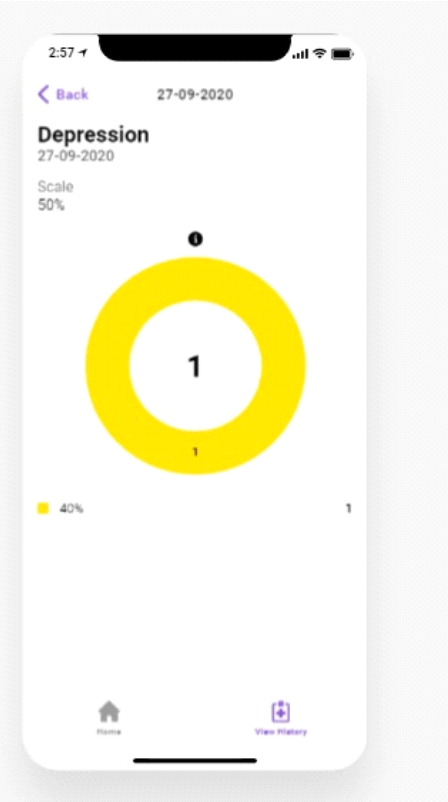


Figure 20: View Depression History
This GUI shows history for depression test attempted on the selected date

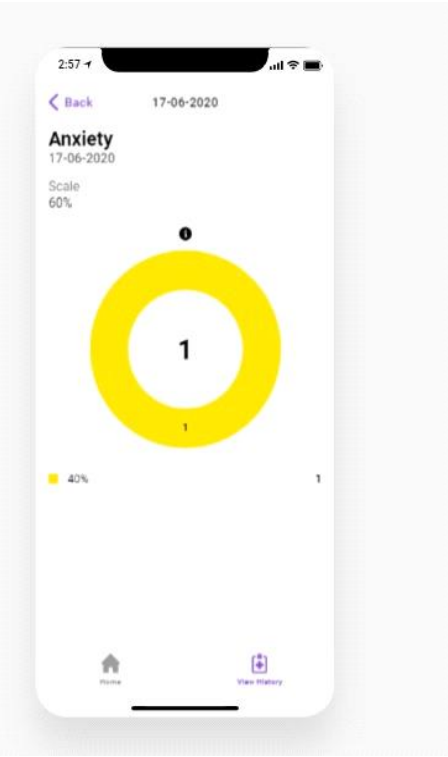


Figure 21: View Anxiety History
This GUI shows history for anxiety test attempted on the selected date

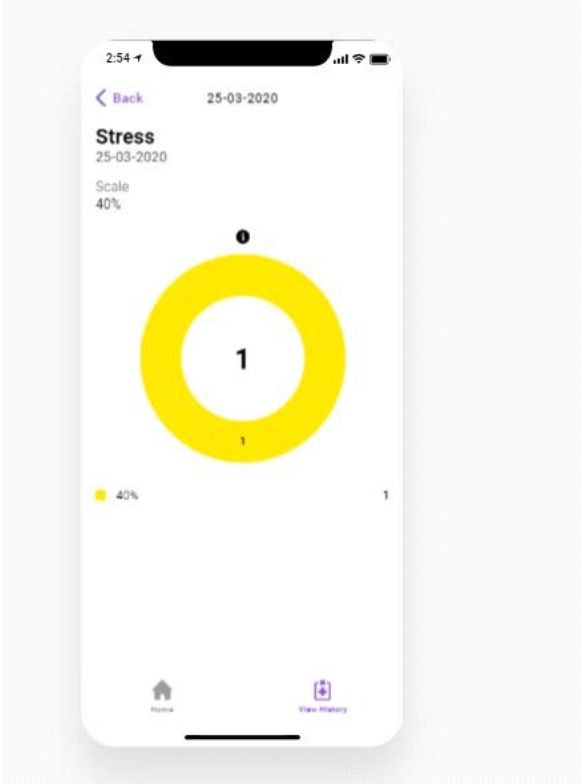


Figure 22: View Stress History
This GUI shows history for stress test attempted on the selected date

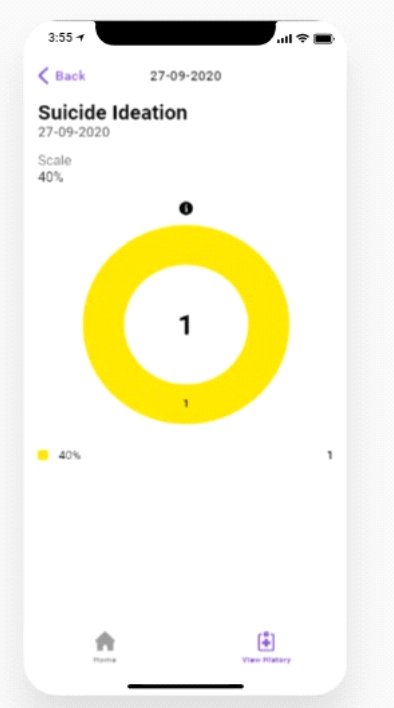


Figure 23: View Suicide Ideation Rate History
This GUI shows suicide ideation rate history

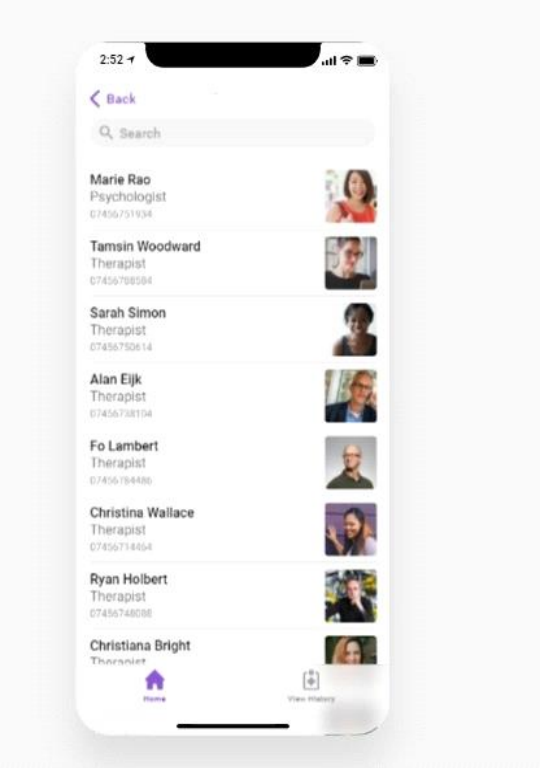


Figure 24: Contact Therapist
This GUI shows list of therapist recommendations

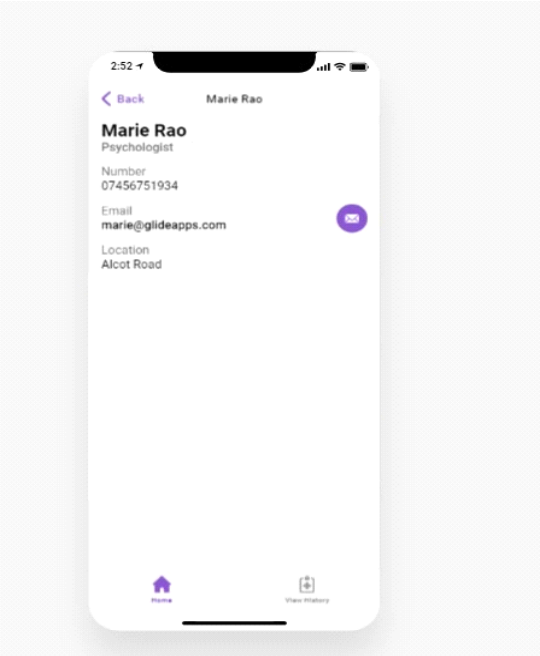


Figure 25: Therapist Details
This GUI shows details for therapist

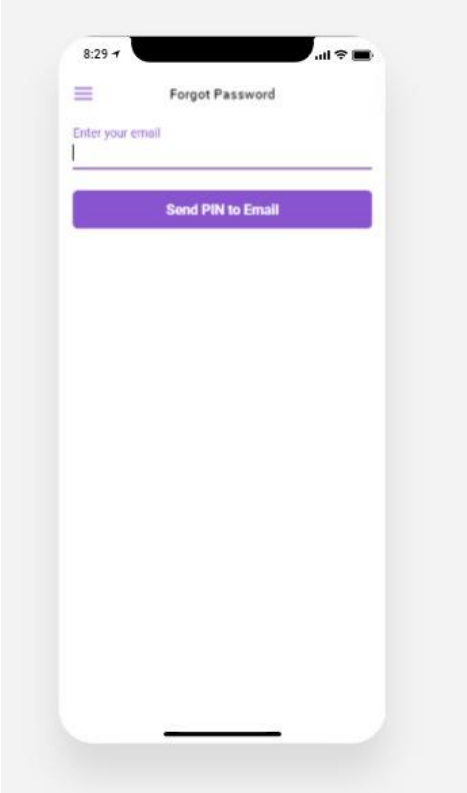


Figure 26: Forgot Password (Screen 1)
This GUI shows Forgot Password (Screen 1)

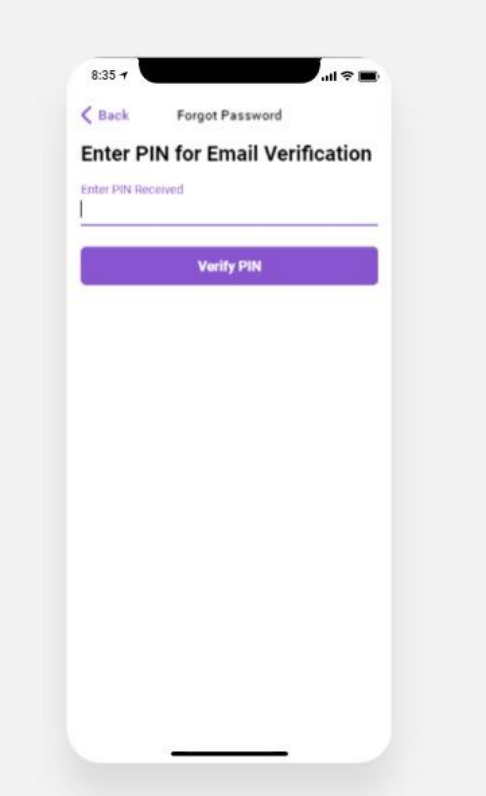


Figure 27: Forgot Password (Screen 2)
This GUI shows Forgot Password (Screen 2)

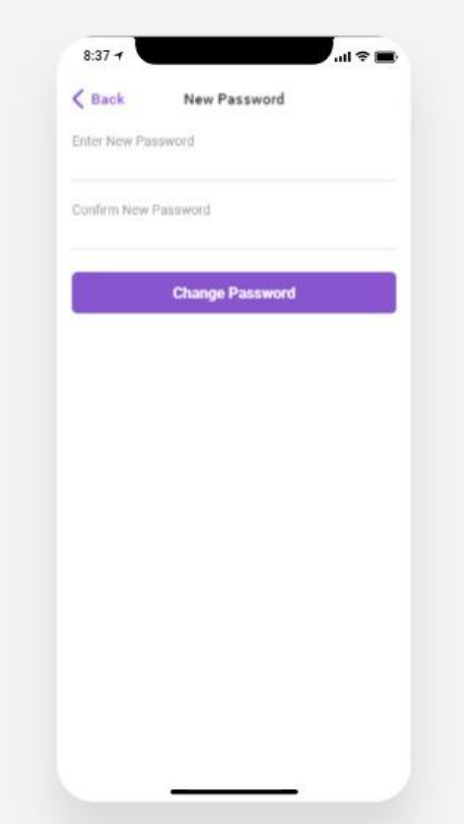


Figure 28: Forgot Password (Screen 3)
This GUI shows Forgot Password (Screen 3)

3.8 Database Design

This section includes the Entity Relationship diagram and data dictionary.

3.8.1 ER Diagram

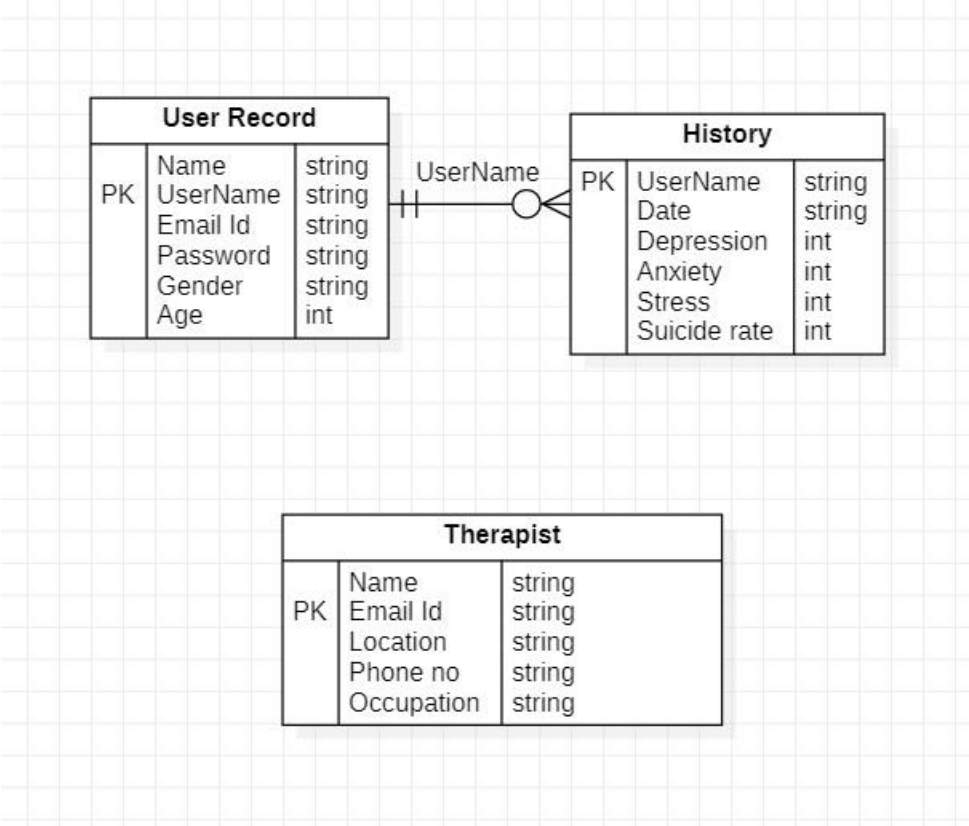


Figure 29: ER Diagram
This is the ER Diagram for our database

3.8.2 Data Dictionary

This sub section includes all the data dictionary tables required according to our database design.

3.8.2.1 User Record

Table 1: User Record Data Dictionary
This is the data dictionary of user record.

Fields	Data Types	Examples
Name	String	Maryam
Username	String	maryam25
Email Id	String	maryammukhtar@gmail.com
Password	String	Maryam123
Gender	String	F
Age	Int	21

3.8.2.2 History

Table 2: History Data Dictionary
This is the data dictionary of History

Fields	Data Types	Examples
Username	String	maryam25
Date	String	28-10-2020
Stress	Int	45
Depression	Int	50
Anxiety	Int	60
Suicide Rate	Int	60

3.8.2.3 Therapist

Table 3: Therapist Data Dictionary
This is the data dictionary of Therapists

Fields	Data Types	Examples
Email Id	String	maryammukhtar@gmail.com
Name	String	Maryam
Location	String	Doctors Hospital
Phone Number	String	0333-1234567
Occupation	String	Psychiatrist

3.9 System Requirements

Following are the hardware and software requirements for the users and developers.

3.9.1 Hardware Requirements

For development, the hardware requirements are the following:

- Desktop with 4GB RAM and 64 bit operating system
- High speed online servers
- Ethernet connection or wireless adapted (Wifi)

For usage, the hardware requirements are the following:

- Internet speed of at least 2Mbps required.

3.9.2 Software Requirements

For development, the software requirements are the following:

- Python: 3.6.8
- Pycharm
- Android Studio

- Firebase DB server
- Python libraries (NumPy (>= 1.13.3 SciPy (>= 0.19.1), Scikit-learn)
- CLE NLP web services

3.10 Design Considerations

Following are the issues that need to be addressed or resolved before attempting to devise a complete design solution:

3.10.1 Assumptions and Dependencies

Following are the assumptions or dependencies regarding the usage of the software:

- The application works for android devices
- The internet facility is available to the person using this application
- Internet speed of at least 3Mbps
- User is familiar with the basic knowledge of using a mobile application
- The end user can converse with the AI bot in Urdu language
- Therapist recommendation list could be extended or reduced

3.10.2 General Constraints

Following are the global limitations or constraints that have a significant impact on the design of the system's software:

- A consistent internet connection is required to use the application
- The user must keep their account’s password safe to keep their records secure
- The user must complete a test (conversation) in order to view their results for the selected test type
- The system would fail to recognize irrelevant answers or any other input language except Urdu
- Application should work on android version 10.0 and above
- Data will not be distributed to any third party and all data shall be stored in encrypted form
- The performance of the system depends on the android version, internet connection speed, etc.
- HTTPS protocol will be used by the system for communication over the internet.
- Users shall be able to use the functionalities of the application that have valid registration and credentials for login

3.11 Development Methods

Our project, Suicide Ideation, follows the agile development method. Agile is the most productive and effective choice. We are using the scrum model more specifically. The main reason to follow this method is to deliver working software frequently with a preference to the shorter timescale and provide early and continuous delivery. The complete project is divided into various sprints which consists of functionality-based tasks. The team members shall

collaborate for a discussion for every sprint. Through an iterative process, we can update for each phase during the development process.
Following sequence would be followed for every functionality:

1. Requirement gathering
2. Product backlog defining and sprint planning
3. Design integration
4. Development
5. Test and release

This helps in dividing the project in small modules and makes development of the product more efficient.

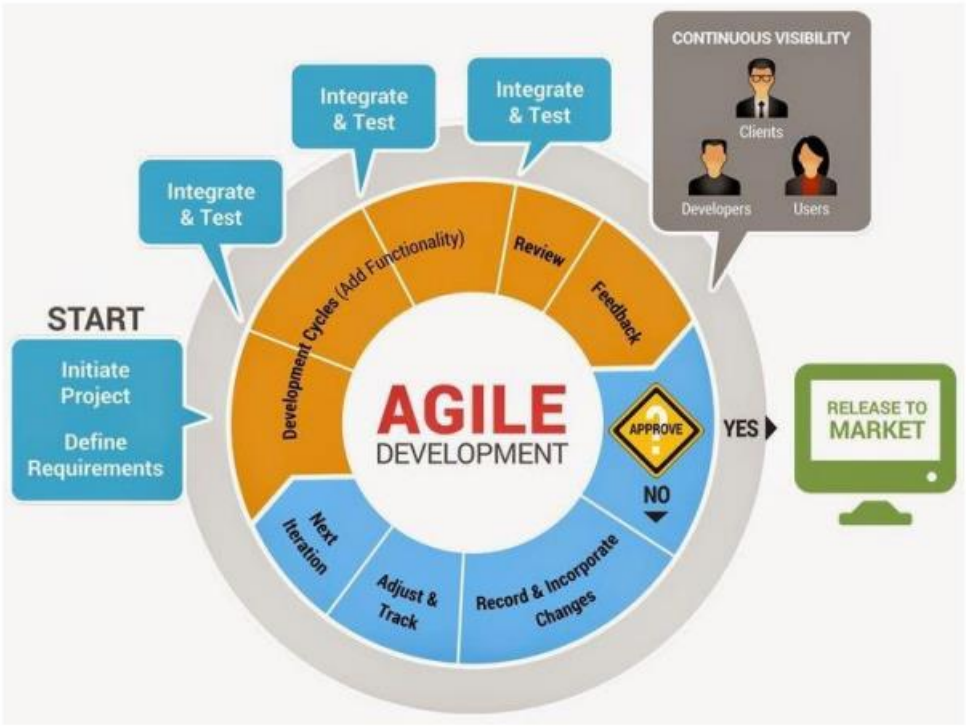


Figure 30: Agile Software Development in [6]

The project development is an iterative process where next sprint is started upon successful completion of previous sprint.

3.12 Class diagram

Following is the class diagram of our project.

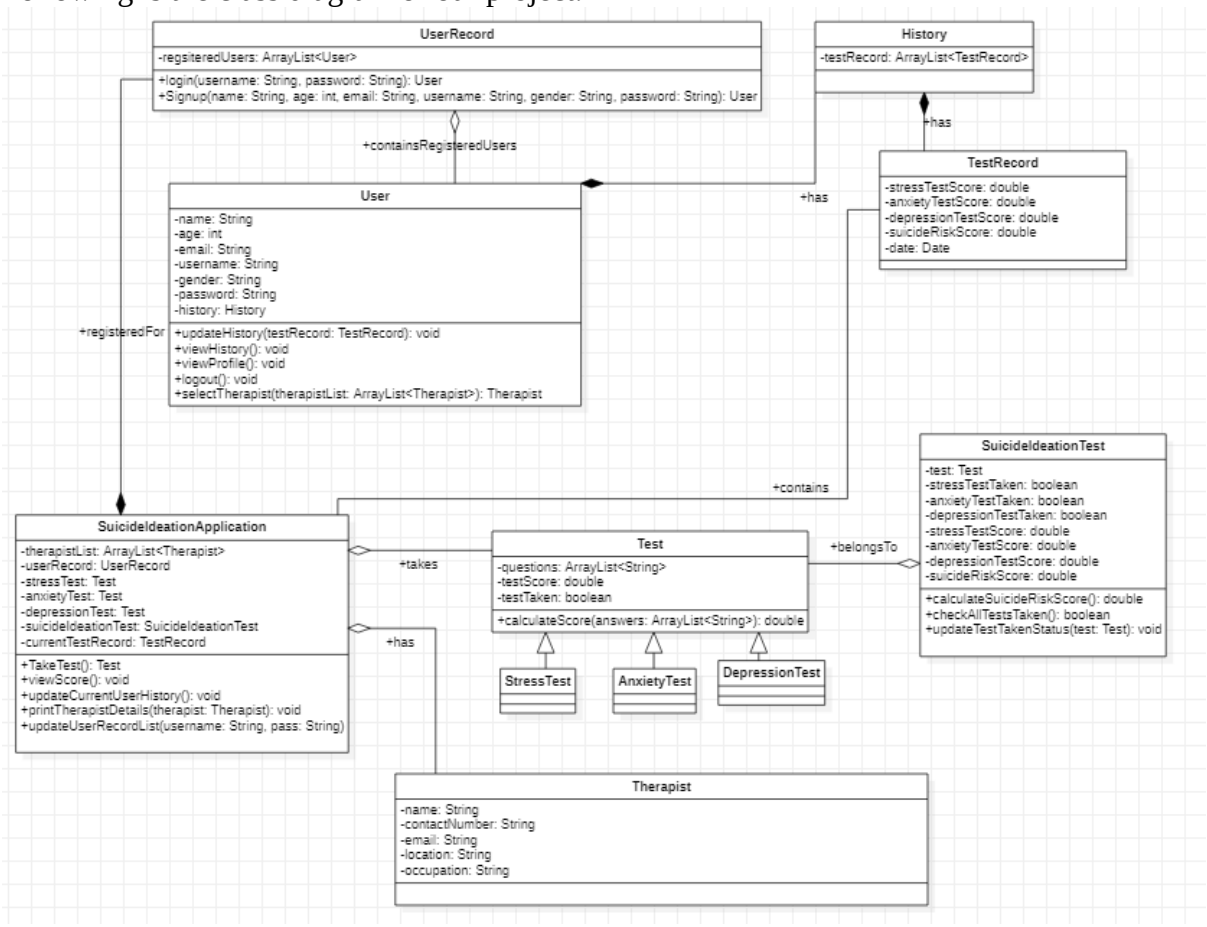


Figure 31: Design Class Diagram

This figure shows Design Class Diagram

3.13 Sequence diagram

Sequence diagrams of all the use cases are shown below.

3.13.1 Login

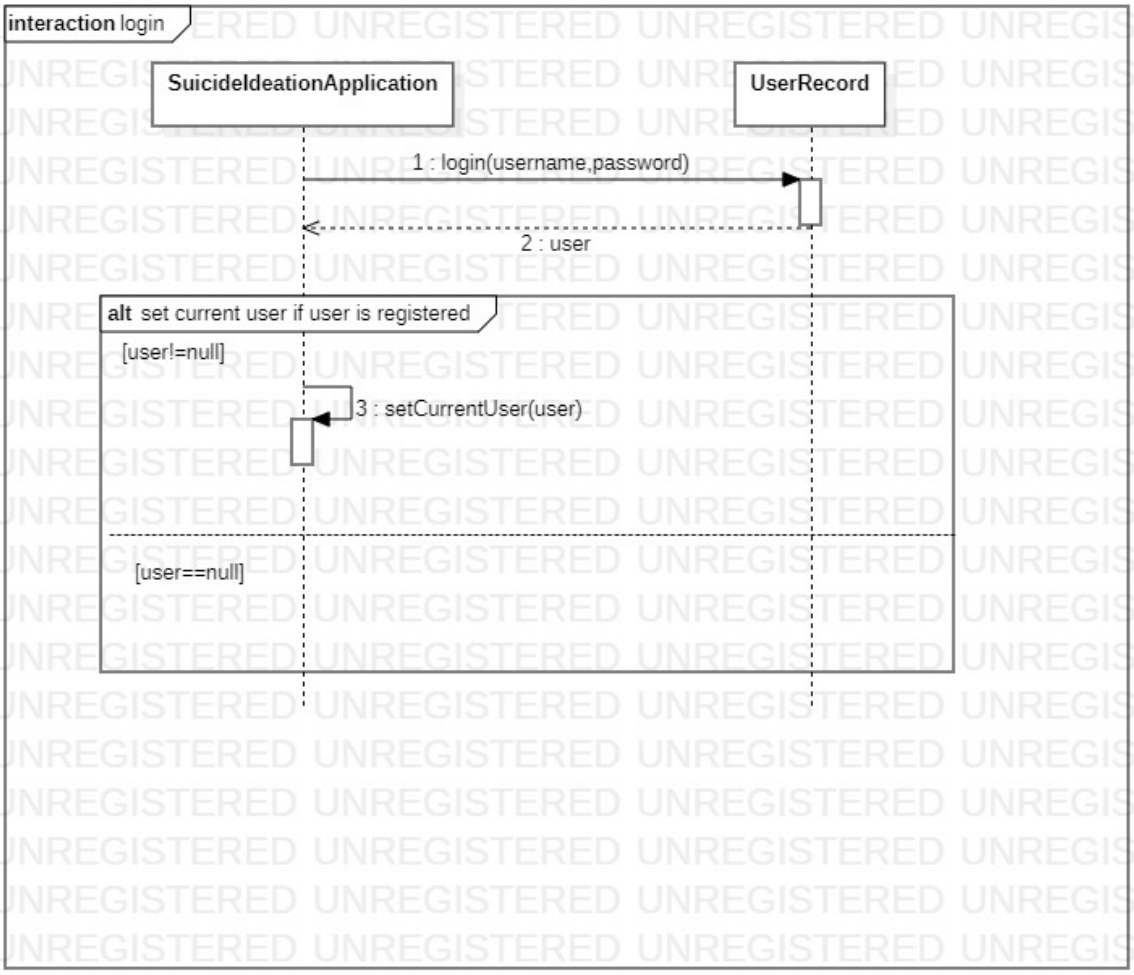


Figure 32: Login

This figure shows Sequence Diagram for Login

3.13.2 Signup

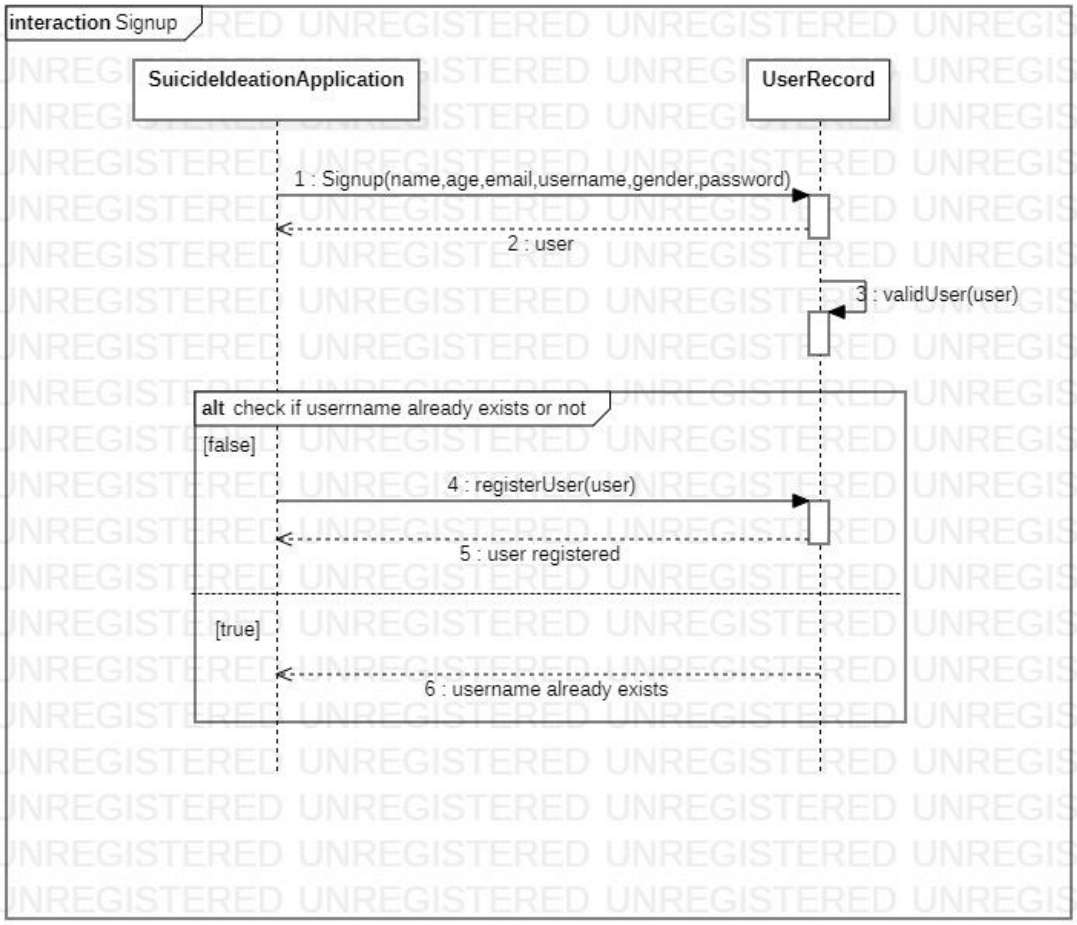


Figure 33: Signup
This figure shows Sequence Diagram for Signup

3.13.3 Logout

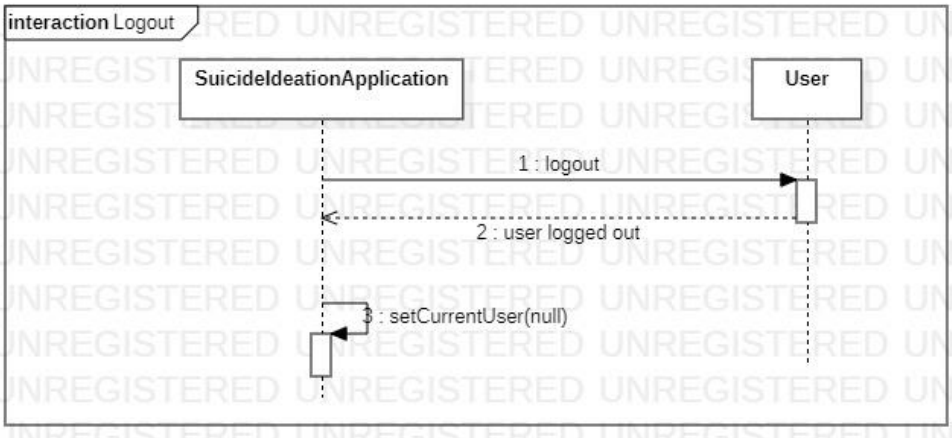


Figure 34: Logout
This figure shows Sequence Diagram for Logout

3.13.4 Change Password

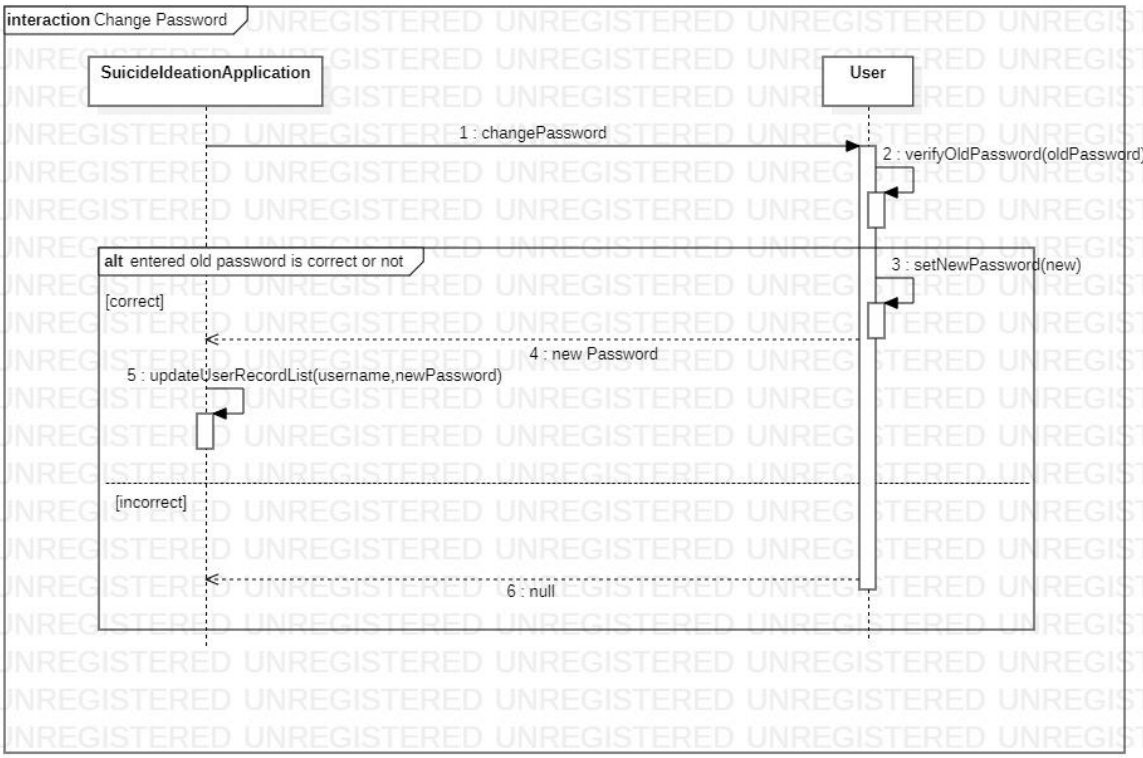


Figure 35: Change Password
This figure shows Sequence Diagram for Change Password

3.13.5 Forget Password

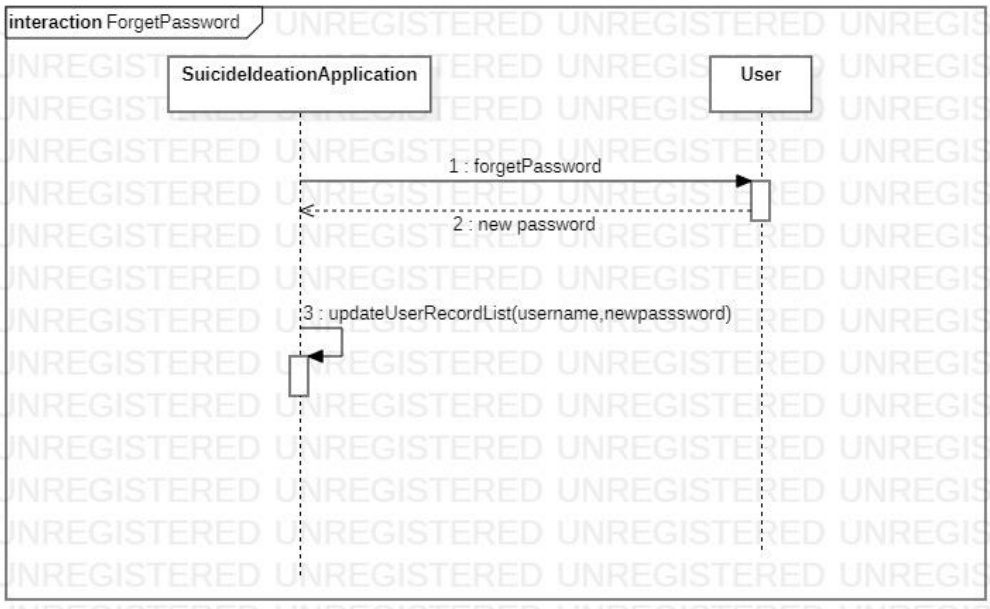


Figure 36: Forget Password
This figure shows Sequence Diagram for Forget Password

3.13.6 View Profile

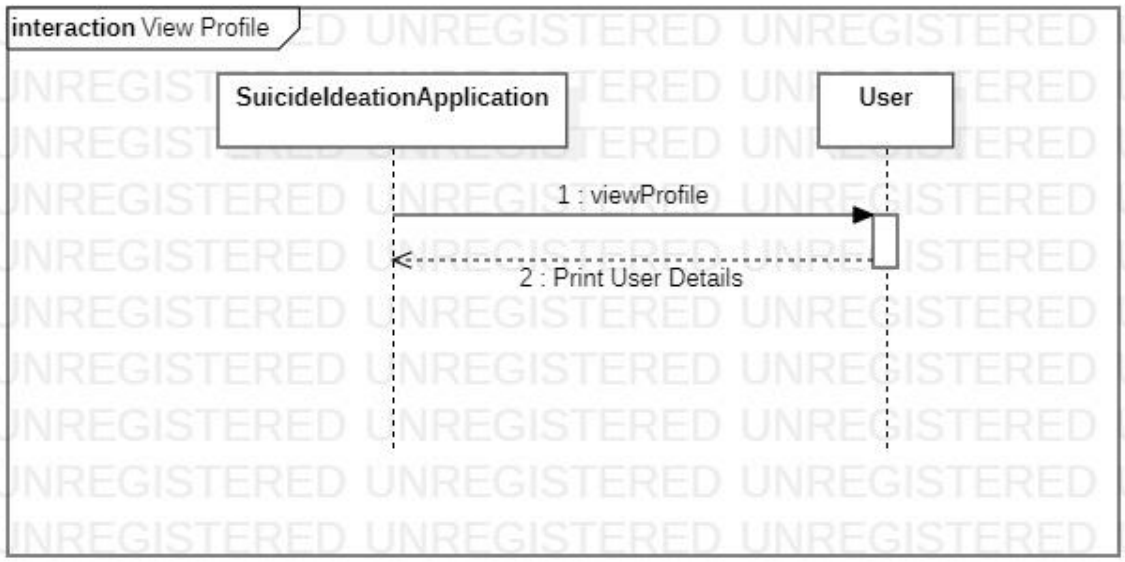


Figure 37: View Profile
This figure shows Sequence Diagram for View Profile

3.13.7 View History

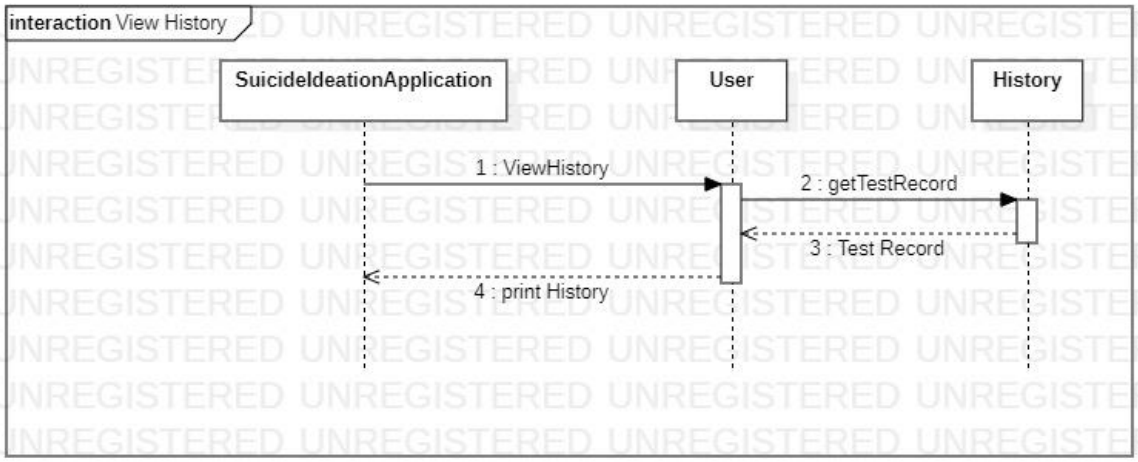


Figure 38: View History
This figure shows Sequence Diagram for View History

3.13.8 Take Test

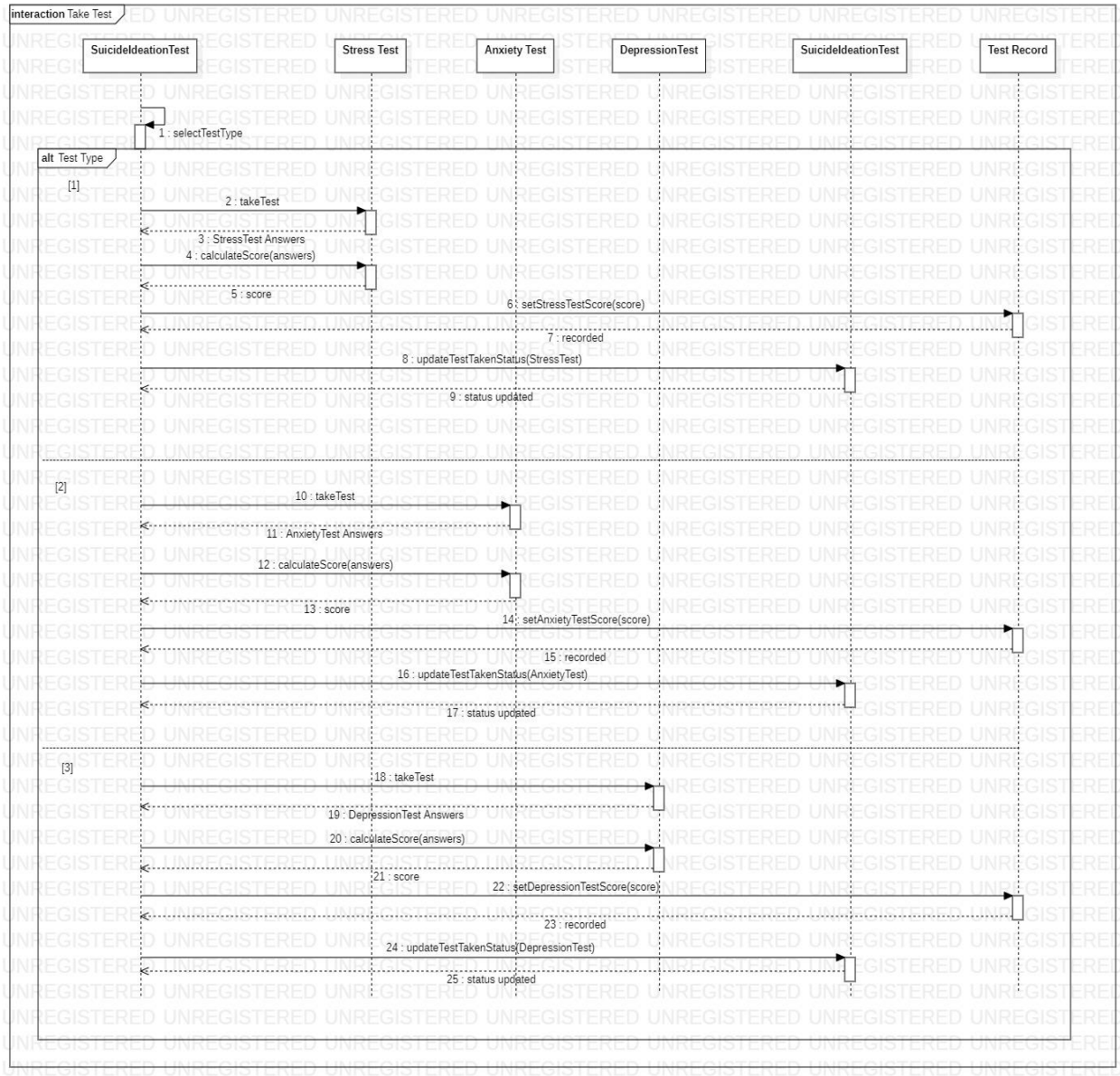


Figure 39: Take Test
This figure shows Sequence Diagram for Take Test

3.13.9 View Anxiety Score

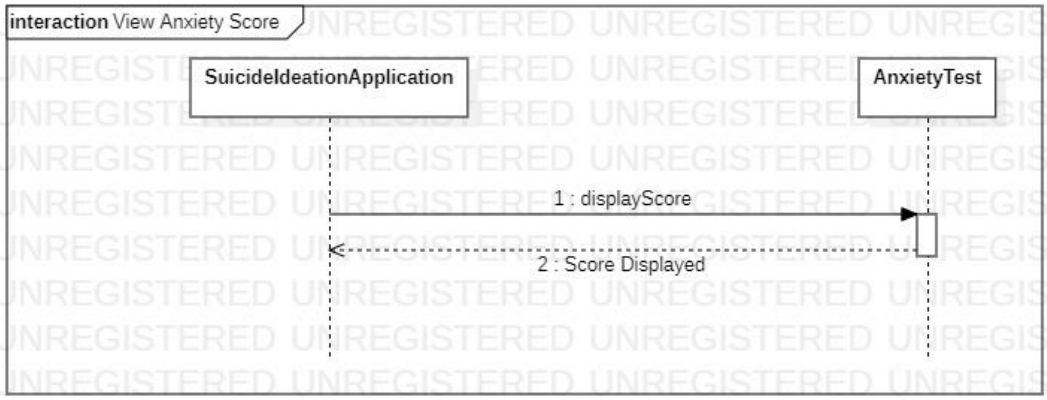


Figure 40: View Anxiety Score
This figure shows Sequence Diagram for View Anxiety Score

3.13.10 View Depression Score

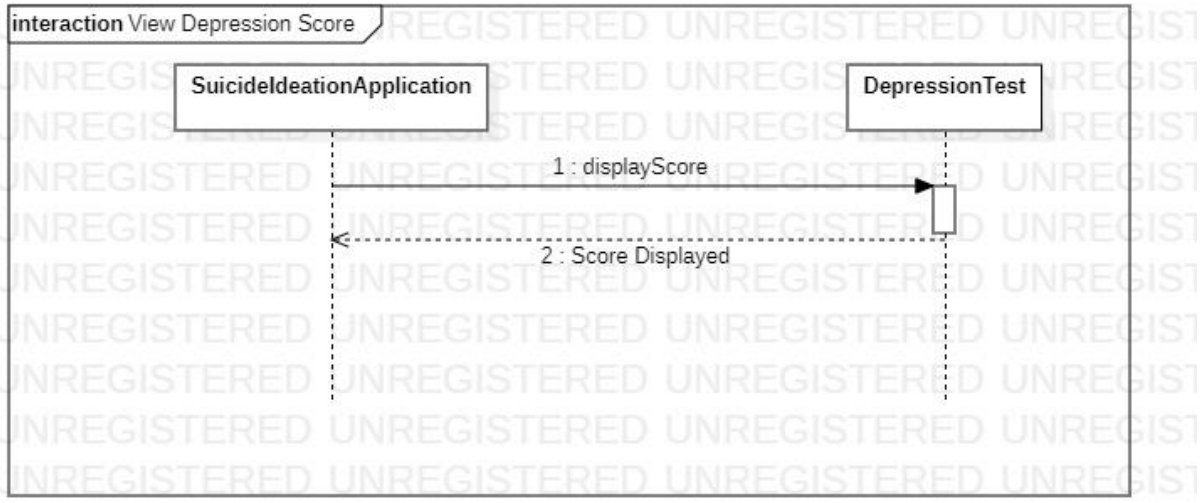


Figure 41: View Depression Score
This figure shows Sequence Diagram for View Depression Score

3.13.11 View Stress Score

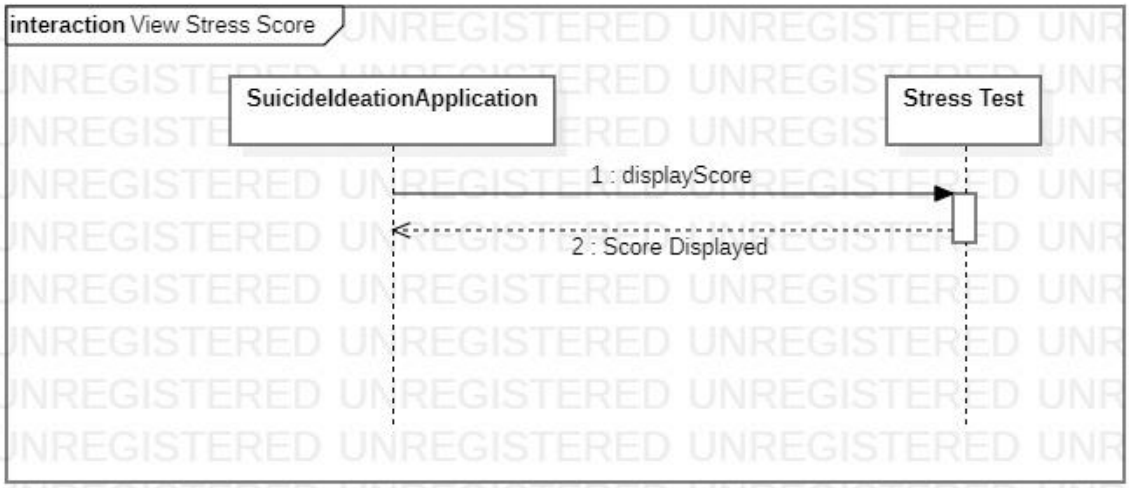


Figure 42: View Stress Score
This figure shows Sequence Diagram for View Stress Score

3.13.12 View Suicide Ideation Score

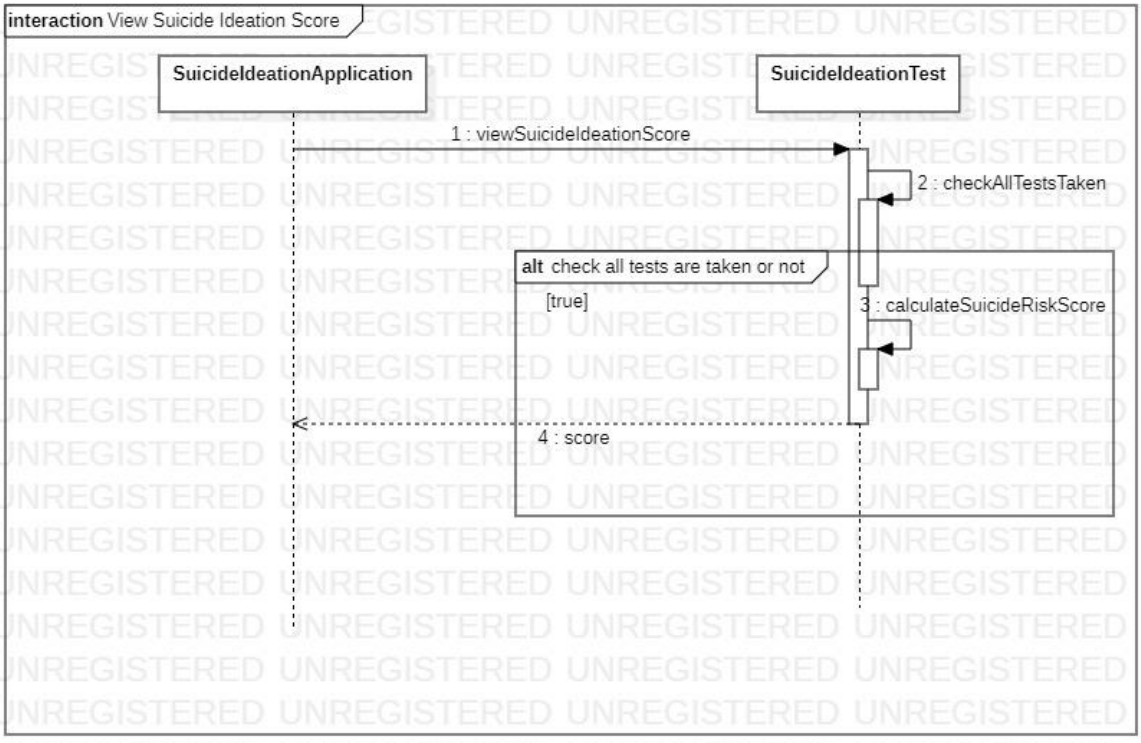


Figure 43: View Suicide Ideation Score
This figure shows Sequence Diagram for View Suicide Ideation Score

3.13.13 Select Therapist

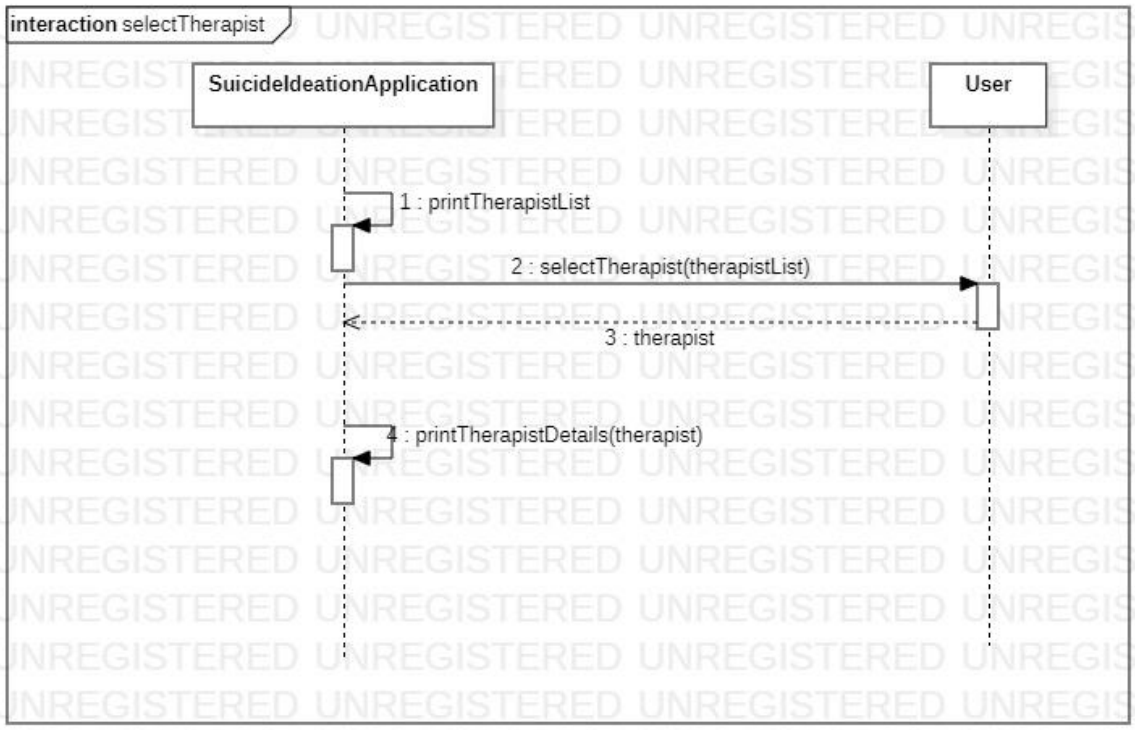


Figure 44: Select Therapist
This figure shows Sequence Diagram

3.14 Policies and Tactics

Following are the design policies and tactics that influences the implementation of our project.

3.14.1 Libraries and Database

For our application, we require libraries including numPy, pandas and keras for classification of the dataset in Python. Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python [7]. Multiclass classification shall be implemented using Keras sequential model with our machine learning algorithm (RBF). Moreover, firebase database is used which is a real-time database that stores data in a JSON object and provides real-time data to all connected clients.

3.14.2 Algorithm

Our application uses a combination of two machine learning models i.e., TF-IDF Vectorizer Multinomial Naïve Bayes. This hybrid outperformed other models and worked best as a training model and also for finding the best true positives and higher AUC as compared to other models used for prediction. [3].

3.14.3 Interface

Interfaces are developed, following the standards of Human Computer Interaction and the golden rules for interfaces, especially the following:

- Strive for consistency
- Prevent errors

- Keep users in control
- Reduce short-term memory load

These design rules would help us develop an application that keeps the user engaged and maintains user-flow.

3.14.4 NLP API

Natural Language Processing API is used for the speech-to-speech conversation in Urdu. By far, we shall integrate these web-services provided by the Center for Language Engineering (CLE) by Al-Khwarizmi Institute of Computer Science (KICS), UET [2].

3.14.5 Testing

First, we need to address the test cases that correspond to the aspects of the system. Then, unit testing will be used to substantiate the whole system.

3.14.6 Performance

A slowed down or crash could lead to user disinterest to continue using the application. It could also lead in user losing their data being processed. To avoid this, we need to consider the memory and network management, error condition and exception handling to avoid hindrance in the performance.

3.14.7 Ensuring Requirements Traceability

Validation and verification of the product will be given importance throughout. Forward traceability will be performed to ensure that the project is moving in the right direction. Backward traceability will be performed to ensure that you are not expanding the scope of the project by adding features or functionality that was not a part of the original requirements.

3.14.8 Use of product

The application provides a signup component, upon which the user can validate and have access to use the services. Later, the user can start a conversation by selecting the test types, start a conversation, check results, view previous history record and contact a therapist.

To conclude, this chapter includes all the information required to understand the system's specifications and architectural design. The main functionalities of the framework are also set out in this chapter and how it works using various variables and techniques. All the main techniques and software development techniques are explained and this chapter gives all the specifications and design detail.

Chapter 4: Implementation and Test Cases

This chapter includes the information about all the work done in the project till now. The system proposed here is an android based application and training the model using Machine Learning Techniques.

4.1 Implementation

4.1.1 Machine Learning

This chapter describes the Machine Learning implementation [8] for detecting suicide ideation in individuals in detail.

4.1.1.1 Data Collection

For this project, we are using different datasets. One is a labeled dataset, DataforModel, combined from two subreddits including r/SuicideWatch and r/Depression. It contains the text from the posts related to depression and suicide, and a label (is_suicide) that depicts true = 1 for posts related to suicide and false = 0 for posts related to depression.
The other dataset is attained using Reddit API to scrape posts from the subreddit r/SuicideWatch. This dataset can be combined and used later in order to increase our training data and accuracy.

4.1.1.2 Data Preprocessing

While dealing with NLP tasks, raw data should be preprocessed first in order to use it to train our model for better results. This includes removing noise, tokenizing the words and perform exploratory data analysis (EDA) to derive meaningful conclusions from the data.
This step has already been performed on the data that has been used.

4.1.1.2.1 Model Implementation

Using natural language processing to identify suicidal posts, following are the steps performed to implement and evaluate our model.

4.1.1.2.2 Loading Dataset

After loading our dataset, DataforModel, a baseline accuracy was established. It results in 0.5166, meaning that the predicted accuracy must be outperformed by our model than this score.

4.1.1.3 Selecting the appropriate column to draw features from

To select the best column to pick our features from, fit the model using scikit-learn-feature extraction text - CountVectorizer for tokenization and scikit-learn – naivebayes, Multinomial Naïve Bayes.
The confusion matrix represents the following parameters:
True Positives (TP) - Predict that an entry is from the SuicideWatch subreddit and it results correctly.
True Negatives (TN) - Predict that an entry is from the depression subreddit and it results correctly.
False Positives (FP) - Predict that an entry is from the SuicideWatch subreddit and it results incorrect.

False Negatives (FN) - Predict that an entry is from the depression subreddit but the entry is actually from SuicideWatch. It means that the results show negative on a person who is inclined to take their life.
After applying the above, it gives the best column to draw our features from. This column contains the cleaned version of the posts and generalizes well i.e., with training score of 0.95 gives a test score of 0.67. Moreover, it also results in better AUC and recall.

4.1.1.4 Finding the best model

Next we compare the results for combinations of models with three different vectorizers for prediction. These combinations are:

- Count Vectorizer – Multinomial Naïve Bayes
- Count Vectorizer – Standard Scalar – K Neighbor Classifier
- Count Vectorizer – Standard Scalar – Logistic Regression
- TF-IDF Vectorizer – Multinomial Naïve Bayes
- TF-IDF Vectorizer – Standard Scalar – K Neighbor Classifier
- TF-IDF Vectorizer – Standard Scalar – Logistic Regression
- Hashing Vectorizer – Multinomial Naïve Bayes
- Hashing Vectorizer – Standard Scalar – K Neighbor Classifier
- Hashing Vectorizer – Standard Scalar – Logistic Regression

These hybrid models were analyzed and compared in terms of AUC, training and test rate accuracy, recall, precision, best score and F1 score.

4.1.2 Android App Implementation

Following section discusses the android app implementation done so far. This includes both the front end and backend.

4.1.2.1 User Authentication

We have used Firebase Authentication for the purpose of sign up and login activities of users in our android application. Users can sign up with their credentials and on completion a verification email is sent to the users, through which they can verify their accounts. After this, the users can successfully login to their accounts through the login screen.

4.1.2.2 Forget Password

If a user forgets the account password in that case the account's password can be changed by selecting the forget password option from the login screen, after which a pin is sent to the user's email. By entering this pin, users can change/update their account password.

4.1.2.3 View Profile

By selecting the view profile option users can view their profile information. The data that is displayed on the screen is fetched through the online firebase storage.

4.1.2.4 Change Password

The user's account password can be changed by selecting the change password option from the view profile screen.

4.1.2.5 User Therapist List

Once logged in, the users can view a list of therapists and their details for consultation and checkups. Therapist's details are stored on Firebase real time cloud database and can be updated through the firebase console.

4.1.2.6 Chatbot

Based on user selection, chatbot screen for the respective test appears where user can interact with the bot through audio conversation. This part uses our NLP API from the server that shall be integrated later. This API, provided by CLE NLP Web services, uses text-to-speech (TTS) and speech-to-text(STT). For that, it uses Apache Http Client and Apache Http Core libraries to build the connection to the web service. It returns a JSON Response containing a wave file or a text message based on the type of input provided in TTS and STT, respectively. In case of SST, the input in wave file format is first converted into a base64 string that requires Apache Common Codec package. Also, we established an SSL connection since the website belongs to 'https'. Chatbot asks questions which are stored in online database.

4.2 Test case Design and description

Following are the test cases for the functionalities of our project and their description. Table 3 shows the test case ids with respect to their titles.

Table 4: Component ID and Name Mapping

This table shows component id and name references to test cases

Component ID	Component Name
1	Validation Component
2	Login Component
3	Test Component
4	NLP Component
5	Therapist Recommendation Component
6	Profile Management Component
7	Score Component
8	History Component
9	Logout Component

Table 5: Test Case ID and Name Mapping

This table shows names of different test cases and id

Test Case ID	Test Case Name
1	Login

2	Login (Alternative - 4A)
3	Signup Account
4	Signup Account (Alternative – 4A)
5	Signup Account (Alternative – 4B)
6	Forget Password
7	Forget Password Alternative – 4A)
8	View Profile
9	Logout
10	Contact Therapist
11	Select Therapist
12	Change Password
13	Change Password (Alternative – 4A)
14	Depression/Anxiety/Stress/Suicide Ideation Test
15	Chat
16	View History
17	View Results

4.2.1 Login

Validation Component			
Component ID - 1			
Test Case ID:	*4.2.1*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Login*
Revision History:	*None*		
Objective	*User should be able to login successfully.*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Login Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User must be registered in the system.*		
Step No.	**Execution description**	**Procedure result**	

1	*User opens the login page.*	*Login page is displayed asking for email and password.*
2	*Enter valid credentials and press login.*	*System verifies the credentials, establish a session for the user and redirects the user to the dashboard.*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.2 Login (Alternative-4A)

Validation Component			
Component Id - 1			
Test Case ID:	*4.2.2*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Login*
Revision History:	*None*		
Objective	*To test login functionality on invalid credentials*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Login Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should be registered in system.*		
Step No.	**Execution description**	**Procedure result**	
1	*User opens the login page.*	*Login page is displayed asking for email and password.*	
2	*User enters invalid credentials*	*System prompts the error message: Incorrect email or password.*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.3 Signup Account

Signup Component			
Component Id – 2			
Test Case ID:	*4.2.3*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Signup*
Revision History:	*None*		
Objective	*User should be able to sign up to a new account*		

Product/Ver/Module:	*Suicide Ideation/ v1/ Signup Module*	
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*	
Assumptions:	*User's device should be connected to Internet.*	
Pre-Requisite:	-	
Step No.	**Execution description**	**Procedure result**
1	*User opens the signup page.*	*Signup is displayed asking for user information.*
2	*Enter valid information and press signup*	*Verification email is sent and user is successfully registered.*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.4 Signup Account (Alternative-4A)

Signup Component			
Component Id – 2			
Test Case ID:	*4.2.4*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Signup*
Revision History:	*None*		
Objective	*To test sign up functionality if invalid credentials are sent*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Signup Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	-		
Step No.	**Execution description**	**Procedure result**	
1	*User opens the signup page.*	*Signup is displayed asking for user information.*	
2	*Enter Invalid input and press signup*	*System prompts a message:* *User not registered successfully*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.5 Signup Account (Alternative-4B)

Signup Component			
Component Id - 2			
Test Case ID:	*4.2.5*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Signup*
Revision History:	*None*		
Objective	*To test sign up functionality if empty credentials are sent*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Signup Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	-		

Step No.	**Execution description**	**Procedure result**
1	*User opens the signup page.*	*Signup is displayed asking for user information.*
2	*Leave the credentials empty and select sign up*	*System prompts a message:* *Please enter the required credentials*

Comments:

The system works as expected

☒*Passed* ☐*Failed* ☐*Not Executed*

4.2.6 Forget Password

Validation Component			
Component Id - 1			
Test Case ID:	*4.2.6*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Forget Password*
Revision History:	*None*		
Objective	*User should be able to reset password successfully.*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Forget Password Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should be registered in database*		

Step No.	**Execution description**	**Procedure result**
1	*User clicks on forget password.*	*Reset password screen is displayed.*

2	*User enter email and click on reset password button.*	*Reset password link is sent to the user email.*
3	*User opens the reset password link*	*New password text field appears.*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.7 Forget Password (Alternative-4A)

Validation Component			
Component Id – 1			
Test Case ID:	*4.2.7*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Forget Password*
Revision History:	*None*		
Objective	*User should not be able to reset password successfully*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Forget password module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should be registered in database*		
Step No.	**Execution description**	**Procedure result**	
1	*User clicks on forget password.*	*Reset password screen is displayed.*	
2	*User enter invalid email and click on reset password button.*	*Invalid email message is displayed*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.8 View Profile

Profile Management Component			
Component Id - 6			
Test Case ID:	*4.2.8*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*View Profile*
Revision History:	*None*		
Objective	*User should be able to view his profile*		
Product/Ver/Module:	*Suicide Ideation/ v1/ View Profile module*		
Environment:	*Software: Any Modern Android Device*		

	Hardware: Android Phone	
Assumptions:	*User's device should be connected to Internet*	
Pre-Requisite:	*User should be logged in to the app*	
Step No.	**Execution description**	**Procedure result**
1	*User clicks on View profile*	*User's profile is displayed showing user's details and the option to change to password*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.9 Logout

Logout Component			
Component Id - 9			
Test Case ID:	*4.2.9*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Logout*
Revision History:	*None*		
Objective	*User should be able to logout from the account*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Logout Module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should be logged in to the app.*		
Step No.	**Execution description**	**Procedure result**	
1	*User clicks on logout.*	*User logged out successfully.*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.10 Contact Therapist

Therapist Recommendation Component			
Component ID - 5			
Test Case ID:	*4.2.10*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Contact Therapist*
Revision History:	*None*		
Objective	*User should be able to view list of available therapists*		

Product/Ver/Module:	*Suicide Ideation/ v1/ Contact Therapist module*	
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*	
Assumptions:	*User's device should be connected to Internet.*	
Pre-Requisite:	*User should be logged in to the app.*	
Step No.	**Execution description**	**Procedure result**
1	*User selects Contact Therapist*	*Therapist List is displayed*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.11 Select Therapist

Therapist Recommendation Component			
Component ID - 5			
Test Case ID:	*4.2.11*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*06/07/2021*	**Use Case Reference(s):**	*Select Therapist*
Revision History:	*None*		
Objective	*User should be able to select Therapist from therapists list*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Select Therapist module.*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should have selected Contact Therapist*		
Step No.	**Execution description**	**Procedure result**	
1	*User clicks on Therapist button*	*Selected therapist information is displayed*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.12 Change Password

Validation Component			
Component ID - 1			
Test Case ID:	*4.2.12*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Change Password*
Revision History:	*None*		
Objective	*User should be able to change password successfully*		

Product/Ver/Module:	*Suicide Ideation/ v1/ Change password module*	
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*	
Assumptions:	*User's device should be connected to Internet.*	
Pre-Requisite:	*User should have selected View Profile..*	
Step No.	**Execution description**	**Procedure result**
1	*User clicks on Change password.*	*Change password screen is displayed.*
2	*User enter email and click on change password button.*	*Change password link is sent to the user email.*
3	*User opens the change password link.*	*New password text field appears.*
Comments: *The system works as expected*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.13 Change Password (Alternative 4A)

Validation Component			
Component ID - 1			
Test Case ID:	*4.2.13*	**QA Test Engineer:**	*Husnain*
Test case Version:	*1*	**Reviewed By:**	*Adeel*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Change Password*
Revision History:	*None*		
Objective	*To test the change password functionality if it receives an empty field*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Change password module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should have selected View Profile..*		
Step No.	**Execution description**	**Procedure result**	
1	*User clicks on Change password.*	*Change password screen is displayed.*	
2	*User clicks on change password button with an empty email field*	*System prompts an empty field error*	
Comments: *The system works as expected*			
☒*Passed* ☐*Failed* ☐*Not Executed*			

4.2.14 Depression/Anxiety/Stress/Suicide Ideation Test

Test Component			
Component Id - 3			
Test Case ID:	*4.2.14*	**QA Test Engineer:**	*Manha*
Test case Version:	*1*	**Reviewed By:**	*Maryam*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Depression Test /Anxiety Test /Stress Test /Suicide Ideation Test*
Revision History:	*None*		
Objective	*User can start a conversation related to Depression, Anxiety, Stress or Suicide Ideation test*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Test module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User's device should be connected to Internet.*		
Pre-Requisite:	*User should be logged in to the system*		

Step No.	**Execution description**	**Procedure result**
1	*User selects Depression, Anxiety, Stress or Suicide Ideation Test*	*User is redirected to Chat page*

Comments:

The system works as expected

☒ *Passed* ☐ *Failed* ☐ *Not Executed*

4.2.15 Chat

NLP Component			
Component Id - 4			
Test Case ID:	*4.2.15*	**QA Test Engineer:**	*Manha*
Test case Version:	*1*	**Reviewed By:**	*Maryam*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*Depression Test, Anxiety Test, Stress Test, Suicide Ideation Test*
Revision History:	*None*		
Objective	*User can converse with the Chatbot related to the chosen test type*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Chat module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User should be connected to Internet*		
Pre-Requisite:	*User should select a test type from home page*		

Step No.	**Execution description**	**Procedure result**
1	*User selects a voice icon to record a response in Urdu*	*System displays users message in Urdu text*

		System continues conversation in Urdu voice command and text
Comments: *This module uses NLP API for speech to text and text to speech conversion. It successfully converts user voice command to Urdu text using Android Intent that is further converted to English text for backend processing. On the other hand, the text to be used by the system is stored in English text, which is translated to Urdu text and further converted to Urdu speech using CLE NLP Text to Speech API. System displays the text and uses voice command to deliver the message.*		
☒*Passed* ☐*Failed* ☐*Not Executed*		

4.2.16 View History

History Component			
Component ID - 8			
Test Case ID:	*4.2.16*	**QA Test Engineer:**	*Manha*
Test case Version:	*1*	**Reviewed By:**	*Maryam*
Test Date:	*29/05/2021*	**Use Case Reference(s):**	*View History*
Revision History:	*None*		
Objective	*User can view their history for previously attempted tests*		
Product/Ver/Module:	*Suicide Ideation/ v1/ History module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User should have attempted any test previously*		
Pre-Requisite:	*User should be logged in to the system*		
Step No.	**Execution description**	**Procedure result**	
1	*User selects View History*	*User is redirected to the History page*	
2	*User selects a test from the list of previously attempted test records*	*User can view the results for the selected test*	
Comments:			
☐*Passed* ☐*Failed* ☒*Not Executed*			

4.2.17 View Results

Score Component			
Component Id - 7			
Test Case ID:	*4.2.17*	**QA Test Engineer:**	*Manha*
Test case Version:	*1*	**Reviewed By:**	*Maryam*

Test Date:	*29/05/2021*	**Use Case Reference(s):**	*View Depression Score, View Stress Score, View Anxiety Score, View Suicide Ideation Score*
Revision History:	*None*		
Objective	*User can view their results for the attempted test*		
Product/Ver/Module:	*Suicide Ideation/ v1/ Results module*		
Environment:	*Software: Any Modern Android Device* *Hardware: Android Phone*		
Assumptions:	*User should be connected to the internet*		
Pre-Requisite:	*User must have attempted a test completely*		

Step No.	**Execution description**	**Procedure result**
1	*User selects View Results*	*User is redirected to the Results page, showing their test results*

Comments:

☐*Passed* ☐*Failed* ☒*Not Executed*

4.3 Test Case Metrics

Following are the test metrics and their respective results after the testing process.

Metric	**Value**
Number of Test cases	17
Number of Test cases passed	15
Number of Test cases failed	2
Test case defect density	11.76%
Test case effectiveness	88.2%
Traceability matrix	Traceability matrix is provided separately as excel file

Chapter 5: Experimental Results and Analysis

5.1 Machine Learning

This section comprises of the results that were concluded after comparing the performances of the defined models.

5.1.1 Choosing the best model

Among the above models, the model that outperforms all in terms of highest recall score i.e., will predict the suicidal risk potential better is the combination of two models **TF-IDF Vectorizer and Multinomial Naïve Bayes.** TF-IDF assigns scores to the words in our feature through tokenization and Multinomial Naïve Bayes makes the prediction of a word falling in a certain category. With only a variation of 0.01 from train to test data scores, it generalizes quite well and is quite far from overfitting.

On the other hand, Hashing Vector + Multinomial NB also outperforms other hybrid models in terms of highest AUC. Even though this model had a slightly better AUC, it was not selected in contrast to TF-IDF + Multinomial NB due to the better accuracy for true positives i.e., prediction of suicidal risk. Logistic Regression typically works well when some of the variables are correlated and hence did not provide considerable scores.

5.1.2 Evaluating the model

The parameters for the two models, Multinomial Naïve Bayes with TF-IDF Vectorizer and with Hashing Vectorizer, that performed well were fine-tuned twice to improve their prediction. The overall result chart can be seen in the table below.

Table 6: Model Accuracy Results

This table shows accuracy results of various hybrid models

Model	AUC	Precision	Recall	Best score	Confusion Matrix	Train accuracy	Test accuracy	Baseline Accuracy	Specificity	F1 Score
cvec+ multi_nb	0.72	0.67	0.67	0.65	{'TP':160, 'FP':71, 'TN': 159,'FN':85}	0.68	0.67	0.52	0.69	0.67
cvec + ss + knn	0.60	0.58	0.58	0.60	{'TP':144, 'FP':100,'TN': 130,'FN':101}	0.72	0.58	0.52	0.57	0.58
cvec + ss + logreg	0.73	0.69	0.69	0.65	{'TP':173, 'FP': 75, 'TN': 155, 'FN': 72}	0.69	0.69	0.52	0.67	0.69
tvec + multi_nb	0.73	0.68	0.68	0.65	{'TP':169, 'FP': 77, 'TN': 153, 'FN': 76}	0.68	0.68	0.52	0.67	0.68
tvec + ss + knn	0.56	0.54	0.54	0.60	{'TP':146, 'FP':118,'TN': 112, 'FN': 99}	0.74	0.54	0.52	0.49	0.54
tvec + ss + logreg	0.73	0.67	0.67	0.65	{'TP':160, 'FP': 72, 'TN': 158, 'FN': 85}	0.68	0.67	0.52	0.69	0.67
hvec + multi_nb	0.77	0.69	0.68	0.72	{'TP':148, 'FP': 54, 'TN': 176, 'FN': 97}	0.89	0.68	0.52	0.77	0.68

hvec + ss + knn	0.51	0.75	0.52	0.52	{'TP':245, 'FP':229,'TN': 1, 'FN': 0}	0.52	0.52	0.52	0.00	0.36
hvec + ss + logreg	0.65	0.62	0.62	0.63	{'TP':155, 'FP': 90, 'TN': 140, 'FN': 90}	1.00	0.62	0.52	0.61	0.62
hvec + multi_nb(tuning)	0.75	0.68	0.68	0.69	{'TP': 170, 'FP': 75, 'TN': 155, 'FN': 75}	0.82	0.68	0.52	0.67	0.68
tvec + multi_nb(tuning)	0.75	0.69	0.69	0.68	{'TP':172, 'FP': 76, 'TN': 154, 'FN': 73}	0.71	0.69	0.52	0.67	0.69
hvec + multi_nb (tuning_2)	0.76	0.68	0.68	0.69	{'TP':167, 'FP': 74, 'TN': 156, 'FN': 78}	0.84	0.68	0.52	0.68	0.68
tvec + multi_nb (tuning_2)	0.75	0.70	0.70	0.68	{'TP':173, 'FP': 71, 'TN': 159, 'FN': 72}	0.71	0.70	0.52	0.69	0.70

Hence, the final model to be deployed to our system is the combination of TF-IDF and Multinomial Naïve Bayes (tuning 2). It optimizes well on the dataset with an AUC of 0.75 and test accuracy of 0.70. With increased number of training data instances, these scores could improve.

Chapter 6: Conclusion

Suicide along with many other mental health issues has taken over the youth and public. Suicide rates are increasing day by day and people are yet not ready to open-up about their mental health issues. To overcome this problem of discomfiture, we aim to develop a system to detect and calculate an individual's depression, anxiety, stress and suicidal rates using machine learning and artificial intelligence. Our system will be consisting of an AI bot that will communicate with the user using Urdu NLP web services provided by Al-Khwarizmi Institute of Computer Science, UET to make it more user friendly. This API comprises of text-to-speech and speech-to-text conversion.

For suicide prediction, we selected the dataset that has been collected from two different subreddits on Reddit. These subreddits provide a platform to talk about their depression (r/depression) and suicidal thoughts (r/suicidewatch) via posts.

Experiment [8] shows that among different NLP models that were applied to detect the accuracy of suicide ideation, a combination of TF-IDF and Multinomial Naïve Bayes outstood all other classifiers by giving the best performance as compared to other supervised learning models. On the other hand, we will be using DASS42 datasets to calculate depression, anxiety and stress scores.

Currently, we are working to integrate the Urdu NLP API with our system to provide a speech to speech platform for the users. From future perspective, our system can be further enhanced by adding therapist recommendation based on the user's location and providing some mental and physical exercises to help with their depression, anxiety, stress and suicidal thoughts.

References

[1] "Suicide: one person dies every 40 seconds", *https://www.who.int/news/item/09-09-2019-suicide-one-person-dies-every-40-seconds*, Sept 09, 2019.

[2] “CLE Webservices”, *https://tech.cle.org.pk/*, 2018

[3] "Assessment of Anxiety, Depression and Stress using Machine Learning Models”, *https://www.researchgate.net/publication/341906771_Assessment_of_Anxiety_Depression_and_Stress_using_Machine_Learning_Models*, January 2020.

[4] “Exploratory Analysis of Social Media Prior to a Suicide Attempt”, *http://www.info-suicide.be/wp-content/uploads/2017/01/Exploratory-Analysis-of-Social-Media-Prior-to-a-Suicide-Attempt.compressed.pdf*, June 16, 2016.

[5] “Supervised Learning for Suicidal Ideation Detection in Online User Content”, *https://www.hindawi.com/journals/complexity/2018/6157249/*, Sept 09, 2018.

[6] M. Chao, “Introduction to Agile Software Development”, medium.com, December 2018”, *https://medium.com/@mcccccc/introduction-to-agile-software-development-fb31b20d0d4e*, April 30, 2020.

[7] “A Gentle Introduction to Scikit-Learn: A Python Machine Learning Library”, *https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/* , July 24, 2019.

[8] “Goodbye world: using Natural Language Processing to identify suicidal posts”, *https://github.com/hesamuel/goodbye_world* , March 14, 2021.