BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Liyanage D.R.Y

(IT18119718)

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

October 2021

BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Liyanage D.R.Y

(IT18119718)

The dissertation was submitted in partial fulfillment of the requirements for the B.Sc. Special Honors degree in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

October 2021

DECLARATION

I declare that this is our own work and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text. Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic, or another medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
Liyanage D.R.Y	IT18119718	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date

[Dr. Pradeepa Samarasinghe]

Signature of the supervisor: Date

[Ms. Vijani Piyawardana]

Signature of the supervisor: Date

[Ms. Shalindi Pandithakoralage]

ABSTRACT

Mental Health is the main component for the human well-being of childhood to senior years. But mental health disorders are preventing humans from living that intention of life. Depression is one of the most leading mental health disorders that can be common among people. The severe level of depression disordered people can go with harsh decisions like take their own life. Therefore, identify depressed people and influence them to go with medication is curial. However, the traditional method of identifying people is clinical psychologist meetings but the social standards on the mentally disordered person and go to a psychologist is extremally sheepish. Therefore people tend to hide the disorder and that makes critical side-effects on them. Moreover, the traditional methods having difficulties on monitor patients for two weeks of the time period. A modern and technological solution for this matter is to come up with a mobile application that users can identify depression analysis.

In this Paper, A early depression risk analysis based on a Phone usage system will be discussed. The system includes a phone usage analysis-based mobile application. The analysis is based on machine learning supervised learning algorithms and mobile applications on flutter language.

Keywords: Depression risk, Phone usage analysis, Machine learning, Flutter.

ACKNOWLEDGEMENT

Foremost I would like to express my sincere appreciation to our behavior analysis for depression detection research group supervisor Dr. Pradeepa Samarasinghe for guiding us to come up with our best ideas and complete the project successfully. Apart from our supervisor, I'd want to express my gratitude to Ms. Shalindi Pandithakoralage, our external - supervisor, for providing us with mental health assistance. Moreover, I want to thank our co-supervisor, Ms. Vijani Piyawardana, for her assistance with the research.

I appreciate the research project lecturer in charge, Dr. Janaka Wijekoon, and the CDAP team for assisting us with the research phases to complete the project in an effective manner.

Lastly, I want to thank my friend and family for their support throughout the research project process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EDRAP : Early Depression Risk Analysis on Mobile Phone Usage

KNN: K-Nearest Neighbour

SVM: Support Vector Machine

UI: User Interface

1. INTRODUCTION

1.1. Background & Literature Survey

1.1.1. Background

Healthy mental health is the balancing element that keeps human inner, exterior, and social well-being in a pleasant state. Mental health has an impact on not only how people think, but also how they feel about others, how they make decisions, and even how they communicate emotionally. From birth through old age, a healthy mind is critical to happiness. Mental health changes are a common state in every person's life at different times, but emotional down, sadness throughout a long period of time is referred to be a mental disorder. Mental illnesses are among the most dangerous factors. Anxiety disorders, depression, mood disorders, PTSD, and eating disorders, to name a few, are all common. According to the World Health Organization (WHO), depression is one of the primary causes of disability.

In society, depression has become a common mental health disorder. Worldwide more than 264 million people of all ages suffer from depression [1]. The disorder is most common among the 18 – 29 age group which is 21.0% [2]. Last long symptoms and severe stage depression are causes for suicide. With these conditions it is crucial to identify individuals that suffer from depression in the early stages, however, methods that identify depression has been designed more than 50 years ago which are face to face interviews, structured interviews conducted by clinical professionals But last decade researches have shown interest in understanding depression connections with technological growth because the technological impact on a daily basis is getting more absolute. Researchers have proved with few longitudinal studies' bidirectional associations between mobile phone addiction depression [3] [4]. Therefore, individuals can be recognized using an individual's phone usage activities since depression and phone addiction are bidirectional factors.

Phone usage addiction can depend on the frequency and time spent on mobile phones therefore Several studies have found relationships between the frequency or duration of mobile phone use and mobile phone addiction [5] [6]. 49.0 % which is more than

3.80 billion people active daily on social medial to keep relationships, access information, and for leisure there for social media directly shows impact our emotions. studies have found a positive relationship between social media addiction and depression, social media usage, and depression [7] [8]. When we consider phone usage late-night phone usage as another special area, we should consider researchers have found mantel health of teens that use late at night is at risk [9]. When we try to accomplish identify early depression risk by behavioral changes using phone usage, we must acknowledge social media use, late-night phone usage as important parameters.

1.1.2. Literature survey

Studies have proved that high phone usage among students and depression links together positively, some of them are studied by Kadir Demirci et al. "Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students" [10], In the study team, have proved that university students that have a high score of depression and anxiety can be monitored using mobile phone addiction. The research "Relationship between the Manner of Mobile Phone Use and Depression, Anxiety, and Stress in University Students" by Aleksandar Višnjić, Vladica Veličković et al. [11] aimed to identify the relationship of phone usage of university students and depression anxiety, stress. The team has shown that as a result of the study most university student population can be lead to mental disorders caused by intensive and modalities phone usage. Moreover the study by Sara Thomée et al. "Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults" the study identifies for both men and women, there were cross-sectional correlations between high versus low mobile phone use and stress, sleep disruptions, and depression symptoms. [12].

The School of Health and Social Development, Deakin University has researched to evaluate the relationship between the duration of time spent using the Internet for leisure, depressive symptoms among Australian adolescents they have shown depressive symptoms were most frequent among higher users of the Internet (3 or more hours per day) among females only Compared to adolescents reporting Internet use for two hours or less [13]. Addiction to social media can reflect suffering from depression disorder. The study by Amit Chowdhry shows that compare to people who use social media less frequently participants that use social media very frequently have 2.7 times the likelihood of depression and compared to people that spent less time, participants that spent the most total time on social media throughout the day had 1.7 times the risks of depression [14].

When it comes to app development for depression using phone usage data, we can find a study by Rafail-Evangelos Mastoras1, Dimitrios Iakovakis1, et al. track depression using Touchscreen typing pattern analysis data from the TypeOfMood app [15] . the team's aimed to develop a machine learning-based method to identify depression with the typing patterns of the individuals. to collect data from people the team has come up TypeOfMood mobile application. There is a study from Dartmouth College, USA conducted a study with the goal of the to see if there was a link between depression and the data acquired using an app and wearables. to tack depression using a mobile phone and wearable they took overuse phone usage from Student Life app as a factor to track depression [16].

In conclusion, most of the existing studies have proposed to prove the link between depression and phone usage and track depression symptoms according to phone usage data acquired using mobile apps. It will be an effective move if we can develop a mobile application that includes phone usage data obtained using phone activities log to behavior analysis for depressive disorder, may help people for self-awareness on the mental health state. In Sri Lanka studies have not been observed that concentrated on a similar approach of automated analysis on depressive disorder identification using unexpected behavioral shifts.

1.2. Research Gap

According to the Literature review, the majority of studies focused on establishing the relationships between depression and overuse of mobile phones usage. In addition, machine learning techniques will be used to forecast depression using data acquired from mobile applications that have been designed to collect individuals' phone usages. Table (1.1) shows the comparison of the research conducted in the area.

Table 1. 1: Comparison of the Conducted Researchers.

	[10] [11] [12]	[13]	[14]	[16]	Proposed System
Depression analysis based on phone usage pattern.	✗	✗	✓	✓	✓
Collect data without user contribution for the analysis.	✗	✗	✗	✓	✓
Explore the key factors that links depresion.	✓	✓	✓	✗	✓
Probability of early depression analysis based on phone usage pattern data.	✗	✗	✗	✗	✓

There are very few mobile applications that have been built to identify early depression risk using phone usage details, and the majority of them were created solely for the aim of collecting research data. The following analysis depicts the current state of mobile application design in the area.

The app "**PROSIT**" is an app that has been used to gather data for research conducted by Dalhousie University they have collected data through both apps and using self-report from individuals. When it comes to the features of the app it supports track data like exercise, sleep, call frequency, massages, users can self-report the details of the week. The app has only been used for the research cannot collect from the app store or the play store [17].

"App Usage" is an app developed to control the phone addiction of individuals. The app collects data related to phone usage which are the most used Apps, screen time.

Features of the app are Screen time and control mobile use, Control, limit screen time and digital habits, Digital history and Daily usage statistics, and Compare and limit screen time with previous data [18].

The app "**TypeOfMood app**" is developed by Rafail-Evangelos Mastoras et al to track depression using Touchscreen typing pattern analysis. In analysis app to collect data which are key pixel coordinates, timestamps of keypresses and releases, typing metadata (number of deletes, number of characters typed, typing session duration, deliberate long-press events, and the application where the user typed). The app is only used for research purposes can be found in the app store and the play store but is not installable [15].

The below table (Table 1.2) indicates the summary of an existing app that is related to depression analysis based on phone usage.

Table 1. 2: Comparison of Current Mobile Applications

	PROSIT	App Usage	TypeOfMood app	Proposed APP
Monitor phone usage characteristics related to depression	✓	✗	✓	✓
Collect data without user collaboration	✗	✓	✓	✓
Depression analysis based on Phone usage data	✗	✗	✗	✓
Extensive reports	✓	✗	✓	✓

2. RESEARCH PROBLEM

Depression has become a common mental health disorder all around the world and is most frequent among the younger generation. Academic problems, career achievements, social life, and habits that negatively impact health are some of the reasons for depression. If symptoms could not identify and hidden for long time life damages can occur therefore, identify depression in the early stage is crucial.

Sri Lanka has a huge negative thought toward mental health disorders individuals that suffer from depression commonly moved on hiding symptomatic behaviors and decline the medications with the fear of social scaling towards mental disorders. This problem can be led huge life damages. The methods that have using in clinics include self-assessment reports, interviews, surveys, and questioners, therefore, the accuracy is based on a human since the process is conducted at clinics and hospitals people are more likely to avoid visit psychiatrists. On the other hand, people need to identify behavioral changes of depression disorder before visit a psychiatrist because of the poor medical knowledge about depression most people don't have an idea they are suffering from depression or not.

Overuse of mobile phones has been linked to depression in studies, and it is also a behavioral shift in those who are depressed. There are apps that have been created to collect phone usage data for marketing purposes, as well as apps that have been created to limit individual phone usage by prohibiting the overuse of apps and activities.

To solve the aforementioned concerns, a self–awareness app based on mobile usage data of individuals employing machine learning analysis to predict early depression risk may be beneficial. Industrial shortcomings regarding depression disorder can be summaries as below,

- The individual difficulty is manually monitored over two weeks.
- A scarcity of automated tools that use phone usage data to identify early depression risk disorders.
- A scarcity of data-gathering systems that do not require the user's participation.
- Because of the detrimental impact on society, people tend to avoid taking drugs.

3. RESEARCH OBJECTIVES

The study's goal is to look at the early signs of depression that can lead to a person being depressed. The study's goal is to use a mobile application to detect activities that are likely to lead to depressive illness.

3.1. Main Objective

The main objective of the component is individual phone usage data analysis for early depression risk analysis identification.

3.2. Sub Objectives

The main objective is to divide them into sub-objectives based on the stages involved in achieving the main goal. Specific Objectives

1) Recognize the key factors of phone usage to identify depressive disorder changes.

There are several parameters that we can use to identify depression using phone usage, but we need to recognize the most important parameters therefore the accuracy of the system will be high.

2) Build a service to run in the mobile in order to log the key factor's data.

Build service to run on the background of the mobile phone and through the designed app and collect the parameter data that is necessary.

3) Based on data build a classifier to model, abnormal phone use based on mobile phone usage.

After the data preprocess model will be trained with the choose machine learning algorithm in order to understand the abnormal phone usage patterns of the individuals.

4) Predict the probability of abnormal phone use towards depression with the designed model.

Using the designed model Abnormal behavior towards depression will be predicted. The probability of the analysis will be provided in the designed model.

5) Integrate the classifier to the mobile app.

The prediction based on phone usage will be serialized and retrieve to the designed mobile application.

The additional, objectives related to the implementation of the mobile application can be explained as follows,

6) Implementing interfaces of the mobile application.

Proposed mobile application interface design and development will consider.

7) Mobile Application Development.

The backend development includes access phone usage data which need to analyze the early depression probability

3.3. Functional Requirements

- Examining a person's phone usage statistics to see whether they have any aberrant behavioral changes related to early depression risk.
- Predict the probability of having early depression symptoms based on analysis of phone usage data collected through the app.
- Two weeks early depression risk analysis must be provided to the user.
- When using the phone usage data, users must be given the relevant notifications.
- Generate the comprehensive report summary.

3.4. Non - Functional Requirements

- Privacy: The system must be in charge of maintaining the privacy of user data and analysis information.
- Security: Before collecting data, the system must be aware that the user has provided consent to access the phone usage data.
- Usability: The application should provide an analytical summary that is user-friendly and clear, as well as easy and efficient operations and navigations.

- Reliability: The phone usage analysis must present the customer with highly accurate probability results on depression.
- Availability: The application must be accessible for users 24/ 7.
- Localization: Must consider about Sri Lankan population's social norms, income, and technology accessibility, Knowledge.

3.5. User Requirements

- Users should be able to see a daily breakdown of their phone usage.
- The user should be able to obtain the results of a two-week early depression risk analysis.
- The user should be allowed to consent to the collection of mobile usage data.
- The user should be able to check whether the data has been obtained correctly.

4. METHODOLOGY

4.1. Introduction

This section explained the process stages of the phone utilization early depression risk analysis application Implementations. Moreover, this section consists of a System diagram and technologies that are used for model development and mobile application development.

The process can be dived into the following stages,

Information Gathering and Analysis: In this stage, necessary prior requirements, current applications, and application gaps have been identified to come up with a convent mobile application.

Mobile Application Designing: In the stage of mobile application designing identified requirements have been organized and identify the features of the "MindGuardian" mobile application according to the target research group and finalized the Mobile application designs.

Model Implementation: The Application consists of an early depression risk analysis. Therefore, Dataset Collection, Preprocessing, model implementation, model API Implementation, and API deployment have been conducted in this stage.

App categorizing method Implementation: The mobile Application needed to categorize the user's used mobile application therefore, in this stage API implementation for app categorization and API deployments has been conducted.

Mobile Application Development and Integration: In this stage, Mobile application development has been completed along with several substages which are Acquire user data implementation, Frontend development, API integration, and Database Integration.

Testing: In the testing stage Model Testing, API Testing, Application front–end Testing, and Mobile Application back-end Testing have been completed.

4.2. System Diagram and Flow.

Obtaining user phone usage data is the first step in the "MindGuardian" mobile application phone usage for the early depression risk analysis process. Data about phone usage is collected without the user's consent. The collected phone usage data is being cleaned and processed to ensure that the data formats are correct for analysis. The data is processed and saved in the Firebase database for later use. The firebase access data is being analyzed for mental health analysis and prediction. Finally, the viewer sees the created health summary and analysis data. In the next sections, we'll go through the basics of each step.

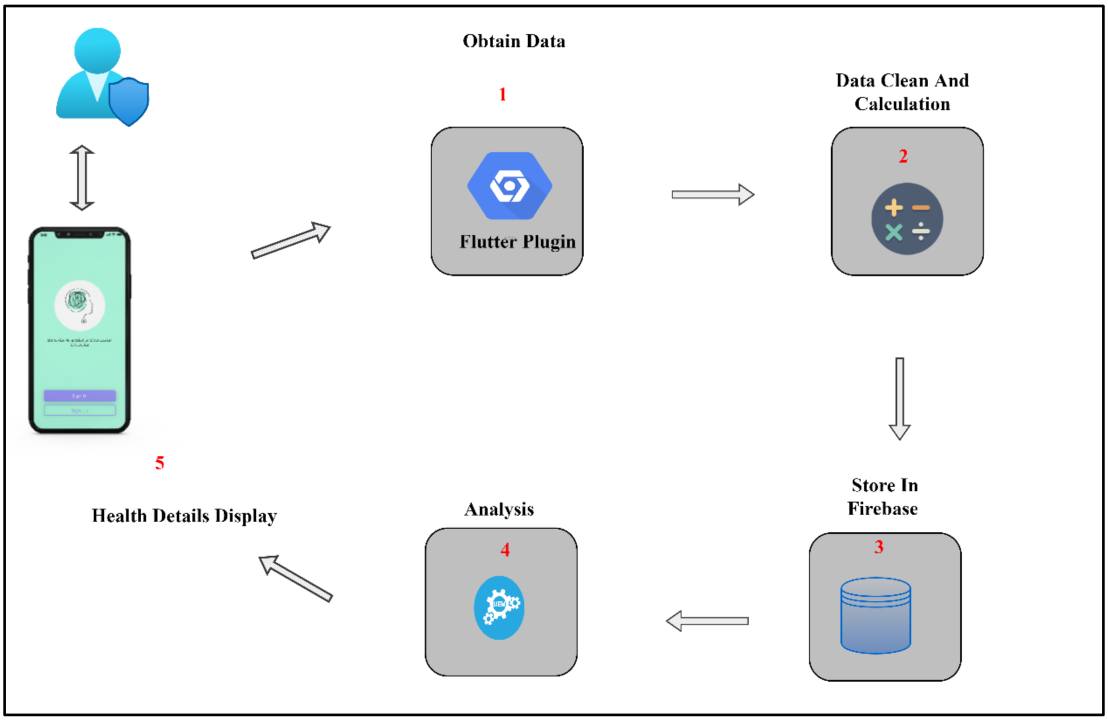


Figure 4. 1: Phone Usage Analysis System Flow.

4.3. Mobile Application Designing

Requirement gathering and analysis.

Requirements gathering has been conducted in the early stage of the research. The information about the research area has been studied with Clinical Psychologist. Furthermore, the current application within the depression identification area has been studied to identify gaps. The research paper gap has been discussed in section (1.2 TABLE).

On other hand, we have looked into a current mobile application that is in use to exam what are the features and how the application comes up with the analysis. Identifying the drawbacks of the current application can make improvements to our "MindGuardian" application. The application gap has been discussed in section (1.2 TABLE).

Design Thinking

The Mobile application features and design need to be arranged according to the identified necessities in the information-gathering phase. Moreover, the research target group cognitive level, mentality level needed to be considered when designing the application. Mobile application designs have been conducted and confirmed by Clinical psychologists in order to fulfill understanding of the target group usability. . Figure (4.2) shows the mobile app UI designs and flow.

Below aspects have been considered when mobile application design thinking,

Simple features: The target users group of the mobile application is a mentally unstable, ill group because of that when coming up with the features and stapes of the mobile application they need to be super simple in this approach users will not be stressed or anxious by using the "MindGuardian". Therefore, log in to the application, guidelines giving permission to certain connectives designed in a simple manner.

Easy and simple health summary displaying: "MindGuardian" mobile application users are mentally special conditioned individuals, In this situation providing mental health summaries with straightforward language is not suitable because it can be a point to make their health conditions worse. Therefore, we have focused on providing the summary in light and simple language in a more relaxing manner.

User-friendly and relaxing design: As previously pointed out designing the UI according to the target group is a bit of a challenging task because the "MindGuardian" application screens needed to be clean and simple to make users not depressed and stress. Therefore, simple text, mild colors, and less complex elements have been used for UI designing.

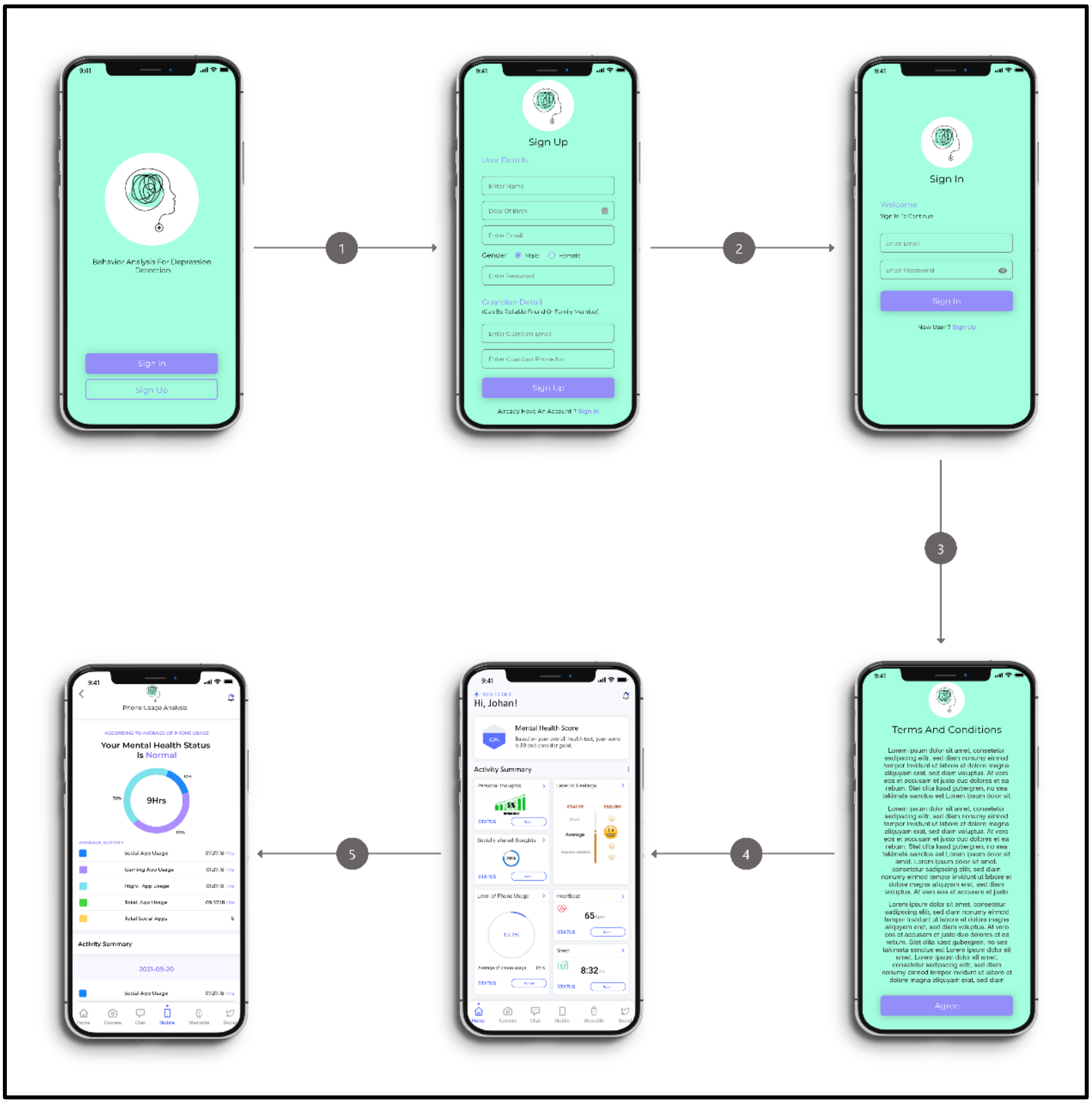


Figure 4. 2: Phone Usage User Flow.

The flow of the “MindGuardian” mobile application early depression risk analysis on phone usage start with a user login to the system or register to the system with basic information. The application leads to the user agreement screen which collects users' consent for access to phone usage data and provides privacy policy agreement to the user. After that user can view the home screen with all component health summary details. Users can access the phone usage health summary and information from navigating from the footer. The Phone usage summary screen provides mental health analysis on average 14 days data and graphical user’s phone usage utilization analyses are provided.

4.4. Commercialization Aspects of the Product

Individuals must be monitored for abnormal behavioral changes in order to detect early depression symptoms and obtain effective therapy, as well as a high chance of saving lives that could otherwise be lost to suicide. However, negative societal attitudes concerning depression cause many to hide their symptoms and avoid taking drugs.

Not only does the suggested system can be monitored by a guardian or a professional about his or her mental health level, but it also addresses particular difficulties and helps them comprehend their mental health level. The following is a description of the proposed mobile application's economic worth.

The general public benefits

- An in-depth examination of behavioral changes without user input.
- A daily overview of phone usage information.
- Automatic biweekly monitoring without the need for human input.
- Simple to set up and use without a prescription
- There will be no invasion of privacy.
- The user can assess their mental health without the worry of societal repercussions.
- As soon as a two-week analysis is completed, clinical phycologists can begin treatment.

Release of a product.

- The mobile application “MindGuardian” is available for users to download and install from the "Google Play" store.

4.5. Tools and Technology.

Tools technologies that have been used for mobile application development and model implementation are discussed in this section.

Tools and technologies that have been used for model implementation are explained in Table (4.1).

Table 4. 1: Model Implemented Tools and Technologies.

Tools & Technology	Use
Environment→ Anaconda 3	Anaconda is a development and deployment environment for the Python and R programming languages. As a result, the anaconda is utilized as a development environment for machine learning models.
IDE → Jupyter Notebook	Jupyter Notebook is an open-source data science application that allows you to run scientific programs in a virtual environment. For the model implementation process, this was utilized as the IDE.
Library → Scikit-Learn	Scikit-Learn is a free Python package that comes with a wide range of classification, regression, and clustering

	models. This package was used to create Python classification models.
Python Pickel	Python Pickel makes advantage of model serialization and deserialization. As a result, this module was used to create a model pickle.
Python fast API (API – Framework)	Python fast API is a framework for creating RESTful APIs. As a result, the framework was utilized to create model APIs.
Spyder (API Implementation environment)	Spyder is a cross-platform open-source application. Model API insertion has been done with Spyder.
Heroku (API Deployment)	Heroku is a cloud platform that can be used in a variety of languages. Heroku is where the model APIs are hosted.

The tools and technologies used for mobile application development are explained in Table (4.2).

Table 4. 2: Model Implemented Tools and Technologies.

Tools & Technology	Use
Android Studio (IDE)	The official development platform for Google Android is Android Studio. As a result, Android Studio was used to construct the mobile app.
Flutter & Dart (Language)	Dart is a client programming language, while Flutter is an open-source user interface development language. Mobile application development has been done with these languages.
Firebase Real-Time (Database)	Google's Firebase is a database solution for mobile and online applications. Firebase Real-Time completes the data storage in the application.
Lavarl (API develpment)	Laravel is a PHP web framework that allows you to create web apps. Laravel use for the implementation of the app categorizing API implementation.

5. IMPLEMENTATION & TESTING

5.1. Implementation

5.1.1. Model Implementation

Mobile application "MindGuardian" consists of a phone usage analysis toward depression risk. Therefore the development process has a model implementation phase.

The stage model implementation procedure is contained Obtain data, pre-processing, Train Model, API development, and Deploy APIs. In this section, all the above components will be discussed. Figure (4.3) exhibits the procedure of the Model implementation.

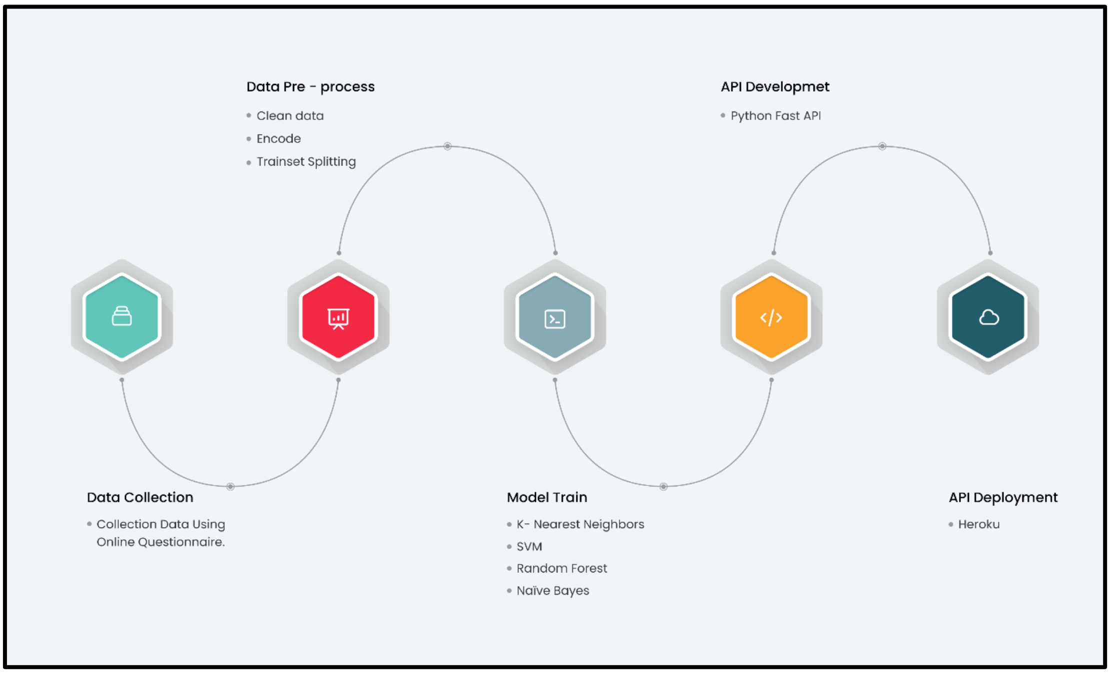


Figure 5. 1: Phone Usage Analysis Steps.

Obtain Datasets.

Model implementation procedure initially starts with collect data for phone usage analysis towards early depression risk. The most common and convented method of collecting data for reaches is the use of publically available online datasets. Therefore, we have tied to fined phone usage data of depressed individuals from the most common online data collection Kaggle. Since the depressed individuals, data is extremely confidential finding data from public databases or find data from current research papers was a failed effort.

With the above issue faced we decided to collect data for Early Depression Risk Analysis on Phone Usage (EDRAP) using a google foram questionnaire. The questionnaire has been designed under the guidance of a research Clinical Psychologist with ensuring the privacy of the user's data confidentiality. Moreover, the questionnaire consists of the questions of users' phone usage habits and world standard Patient Health Questionnaire-9 (PHQ-9) questions [19]. The PHQ-9 questions have been used for identifying the level of depression of the individuals. Figure (5.2) indicate the scoring of the PHQ-9 scale.

» Next Steps | Evidence | Creator Insights

MANAGEMENT

PHQ-9 Management Summary

Score	Depression severity	Comments
0-4	Minimal or none	Monitor; may not require treatment
5-9	Mild	Use clinical judgment (symptom duration, functional impairment) to determine necessity of treatment
10-14	Moderate	
15-19	Moderately severe	Warrants active treatment with psychotherapy, medications, or combination
20-27	Severe	

CRITICAL ACTIONS

- Perform suicide risk assessment in patients who respond positively to item 9 "Thoughts that you would be better off dead or of hurting yourself in some way."
- Rule out bipolar disorder, normal bereavement, and medical disorders causing depression.

Figure 5. 2: PHQ-9 Scale Scoring.

From the questionnaire, we were able to collect 351 responses. The responses have been the age range of 16 – 30 individuals, 203 female, and 148 male. After collecting the data, the PHQ-9 scores have been calculated. After the calculation, we select point 9 as the cutoff point for label depression. Therefore, we identified 100 individuals who shows the possibility of risk towards early depression. The scoring and labeling were conducted by a research Clinical Psychologist. The below figure (5.4) shows the foram response.

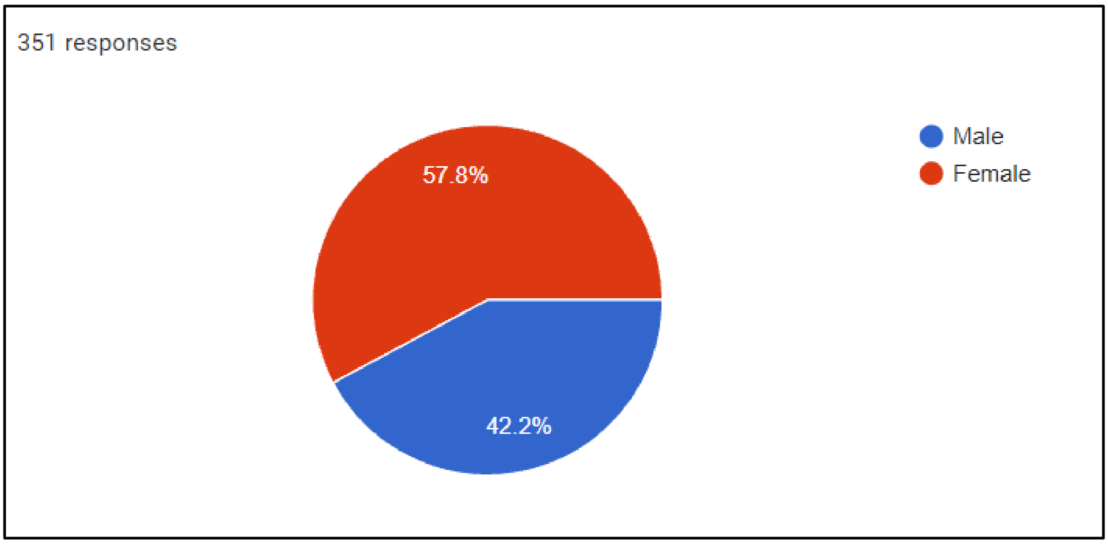


Figure 5. 3: Foram Response.

Data Pre-processing

Data pre-processing is a crucial step when it comes to model implementation. In this step, obtained data will be shep to the necessary way of feeding into the model. This step consists of data cleaning, data encoding, and trainset splitting. Therefore, the collected data scored data has to been gone through pre-processing step prior to model training. The pre-processing steps of EDRAP will be discussed in this section.

As the first step of the data pre-processing data set needed to be cleaned, The ways of cleaning are adding convenient headings for the data columns and clean null values.

We can clean null values in 2 different ways, remove null value raws or replace the null value cells with mean, mode, or median. Along with these two methods, the phone usage dataset has been cleaned null values with remove null value raws. Next, data encoding has been performed to the data set. Data encoding does is convert all the values into float values for purpose of model training performance. The last step of the data pre-processing stage is to train test split the dataset for model train and testing. The dataset has been split into train and test with the most usual industrial norm which is the 8: 2 method, 80% train, and 20% test. Figure (5.5) shows the Encoding and Trainset splitting of the phone usage dataset.

All above preprocessing steps have been performed in Jupiter notebook with the use of “scikit-learn” libraries.

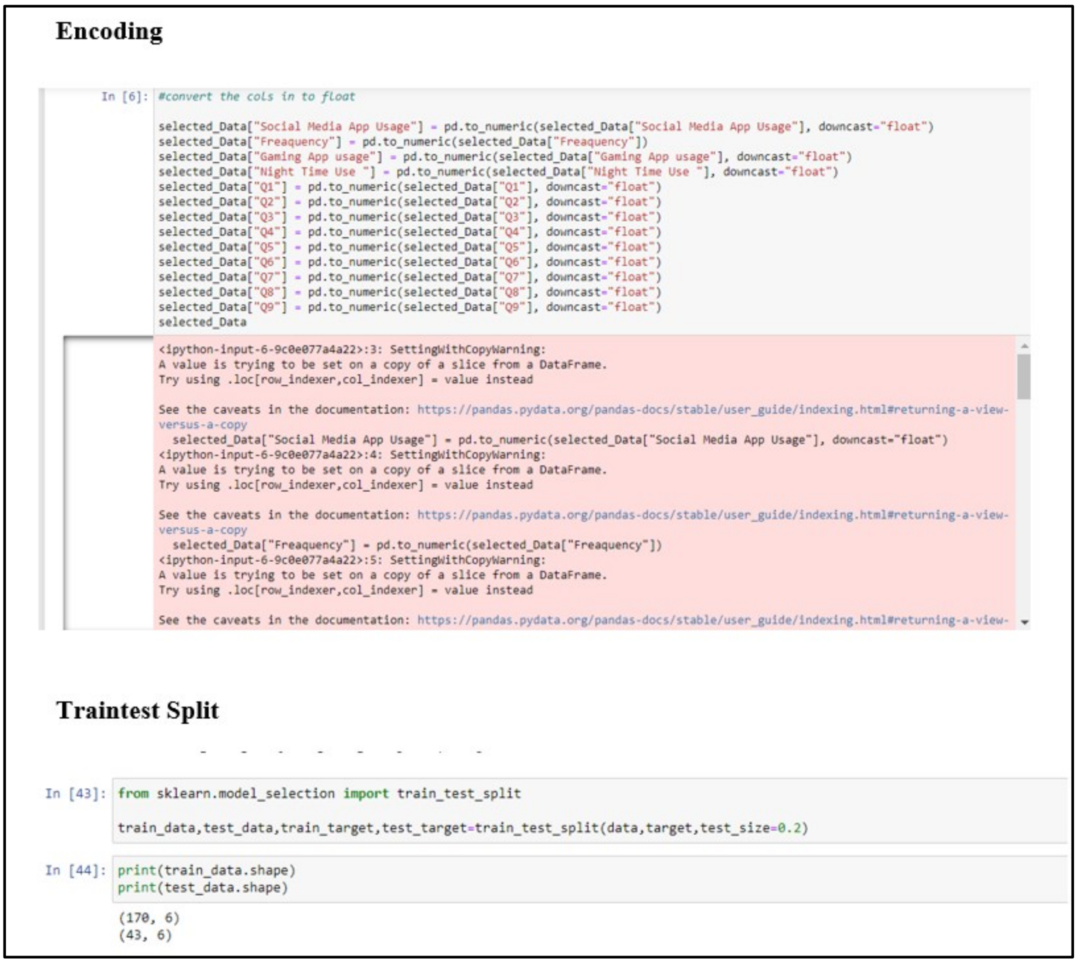


Figure 5. 4: Data Encoding and Trainset splitting

Model Implementation.

The procedure of model training for phone usage dataset is explained in the following section,

The EDRAP has been conducted with several algorithms. The algorithms have been chosen according to the dataset. since the phone usage dataset is a labeled dataset all the training has been conducted with supervised learning algorithms. Random Forest, K– Nearest Neighbors (KNN), Support vector machine (SVM), and Naïve Bayes are the algorithms that have been used for training.

Several common steps have been conducted before fit the pre-processed datasets to the models. relevant libraries are imported to the notebook, the dataset is imported to the notebook, Variables were Standardized and x defined with features, y defined with the label "Depression".

The Phone usage dataset trained for random forest classifier performed under "RandomForestClassifier" with estimators of 20 and random state 0. KNN algorithm best k have calculated for train the phone usage data set with KNN. The best k marked as 14. The next algorithm to train the dataset was SVM, the algorithm trained with kernel poly, degree 2. Lastly, the same dataset has trained to Naïve Bayes for GaussianNB. Figure (5.5) shows model Training.

As the last step of the model trained the confusion matrix, Roc curves and bump pickle file step performed. In order to perform these steps, relevant libraries (sklearn.metrics, pickle)have been used. The model implementation phase has been performed with Jupiter notebook using "scikit-learn" libraries.

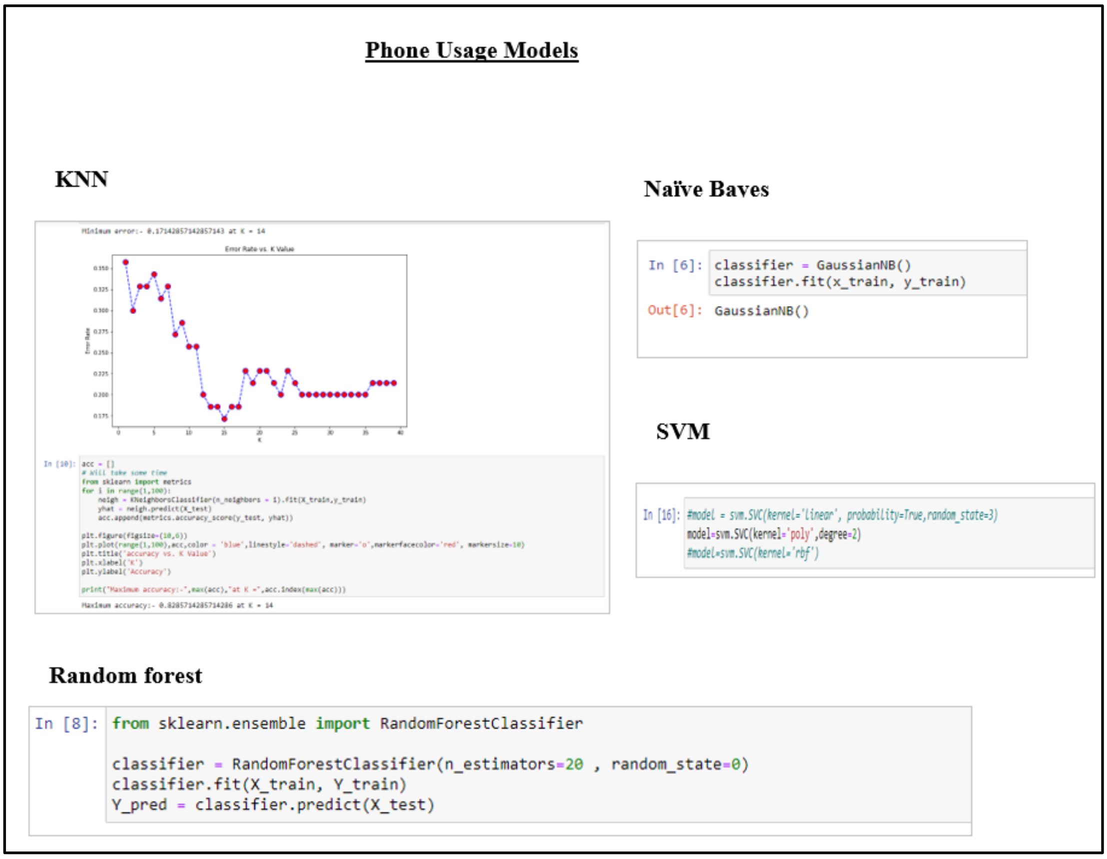


Figure 5. 5: Phone Usage Analysis Algorithms.

API Development and Deployment.

APIs have been developed for EDRAP model integration to the mobile Application “MindGuardian”. The API development and deployment process will be explained in the following section.

After the model implementation, the best accuracy model that can use for EDRAP is chosen. API development has cooperated with the pickle file of the best model. The model selection process will be discussed in section (6). KNN model pickle file has been selected for the development of the EDRAP API. Using the Pickel file API has been written. To get the response from the API the parameres that data need to be passed have been defined in the API. API development is done with python fast API and used IDE Spyder. Figure (5.6) exhibits the API development of EDRAP.

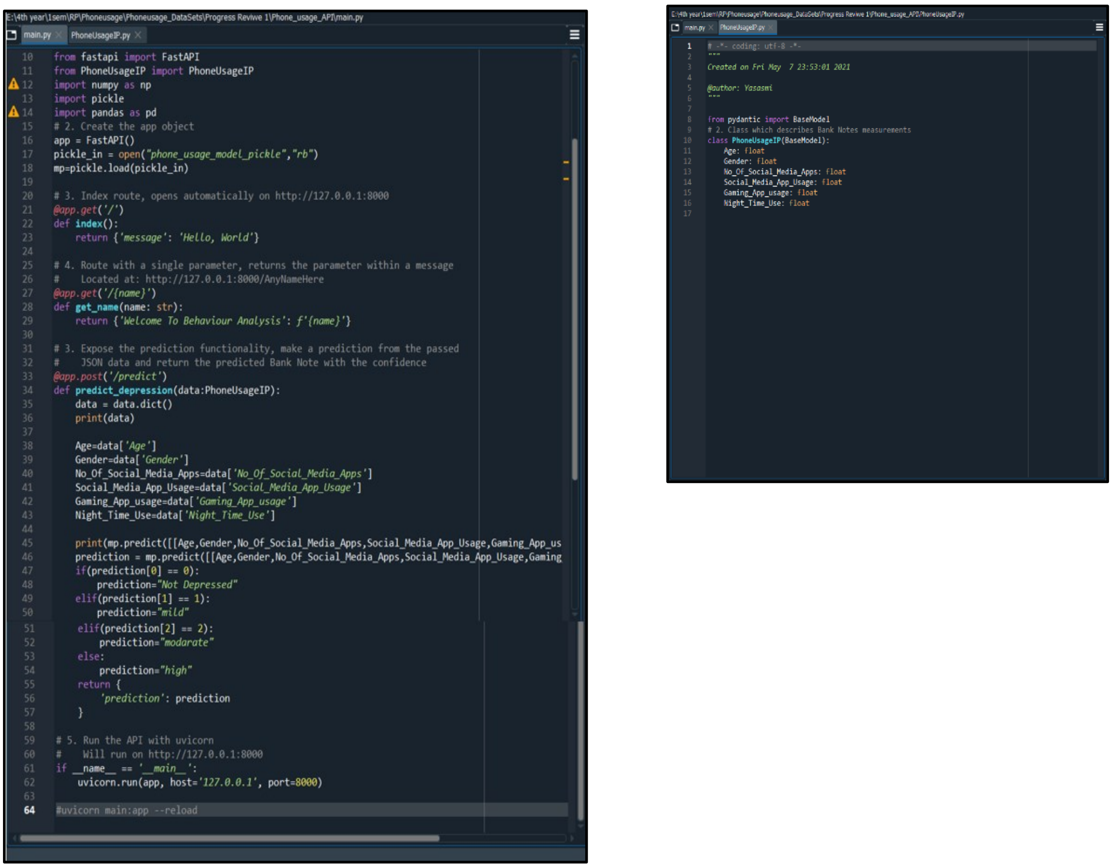


Figure 5. 6: API Development of EDRAP

In order to access the developed API from the external source, it needed to be deployed. Therefore, the Developed API is deployed in Heroku. Testing of the API has been discussed in the future section (5.2.3). Figure (5.7) shows the deployment of the API.

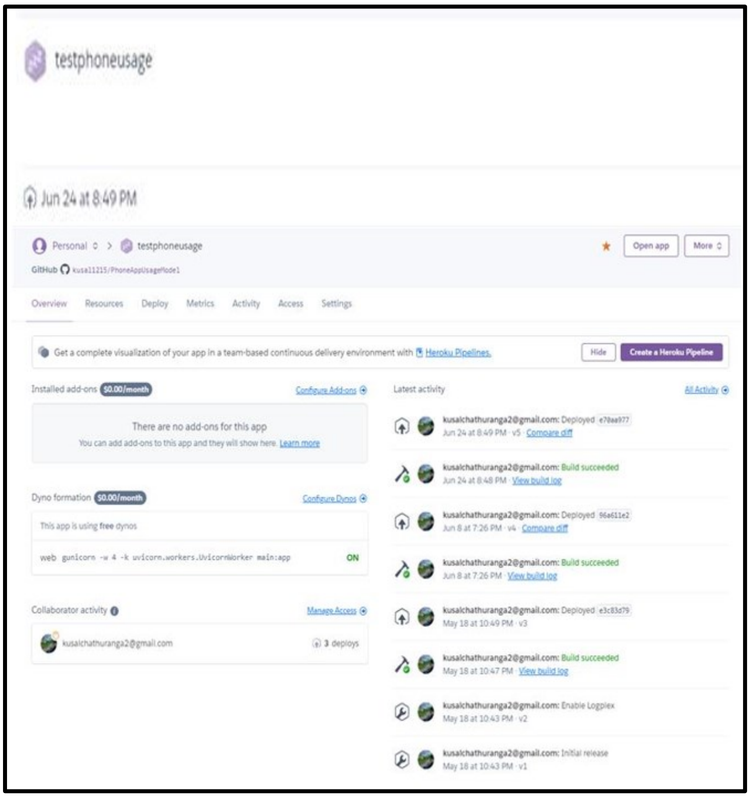


Figure 5. 7: EDRAP Model API Deployment

5.1.2. App Categorizing Method Implementation

In the process of mobile application development of EDRAP mobile app that user used in day to day lief needed to be categorized into social media apps, gaming apps, in order to work with the analysis. Therefore, a method to categorize the apps was crucial. In this section, the process of implementing a method to categorizes will be discussed.

Before categorizing the apps into main two categories which social media and gaming apps, needed to obtain the type of apps grouped in the google play store. The best way to access the type of mobile app is by implementing an API. Therefore, API has been implemented to access the type of mobile application.

The method of the obtain type is by access the Google Play store Browser HTML selector and get the type by passing the mobile application package name. Figure (5.8) shows the method of accessing the selector path.

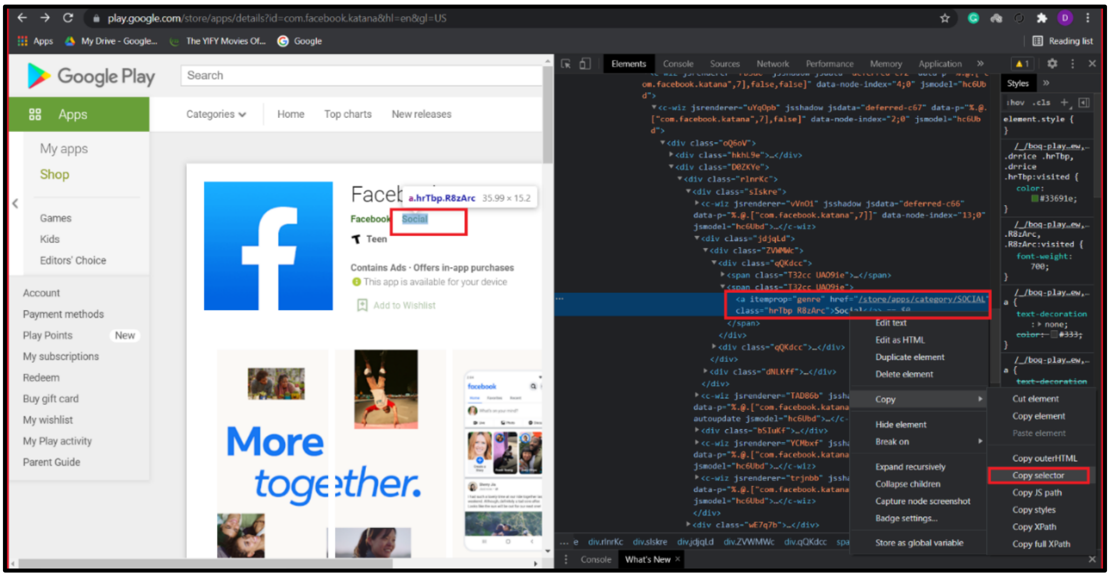


Figure 5. 8: Method of Accessing the Selector Path

To this process, API has been implemented. The types of apps grouped by google are communication, social, entertainment for social apps and arcade, action, adventure, puzzle, etc for mobile gaming apps. The collected app types have been categorized into main two categories in the mobile implementation. For implementation Laravel framework with Library crawler has been used on IDE PhpStorm.This section is discussed in section (5.2.3). The API has The following figure () indicated the API development of access the app type.

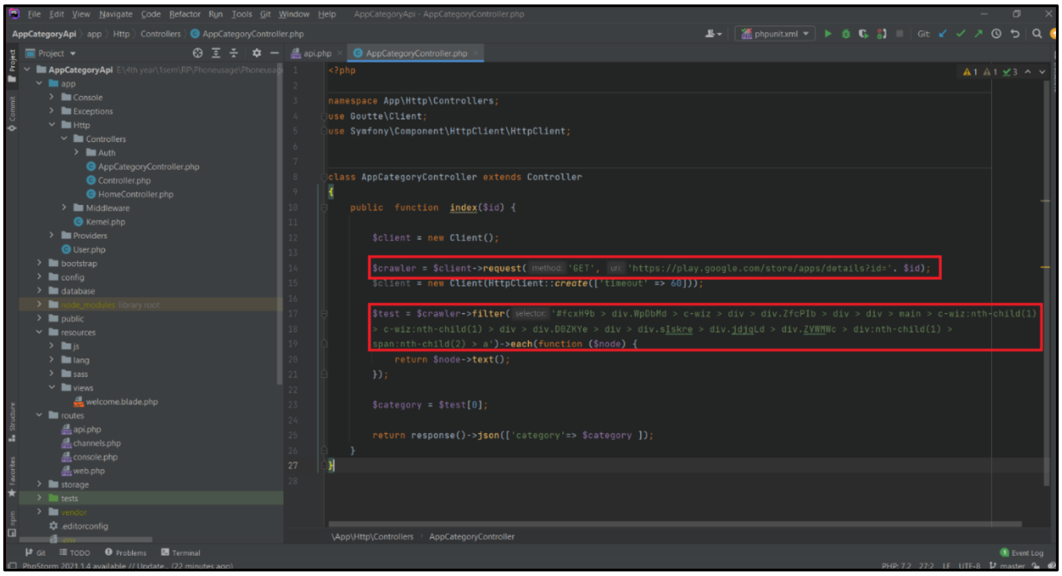


Figure 5. 9: API Development of Access the App Type

The developed API has been deployed in Heroku for access from the mobile application. Testing of the API has been discussed in section (5.2.3). Figure (5.10) shows the deployment.

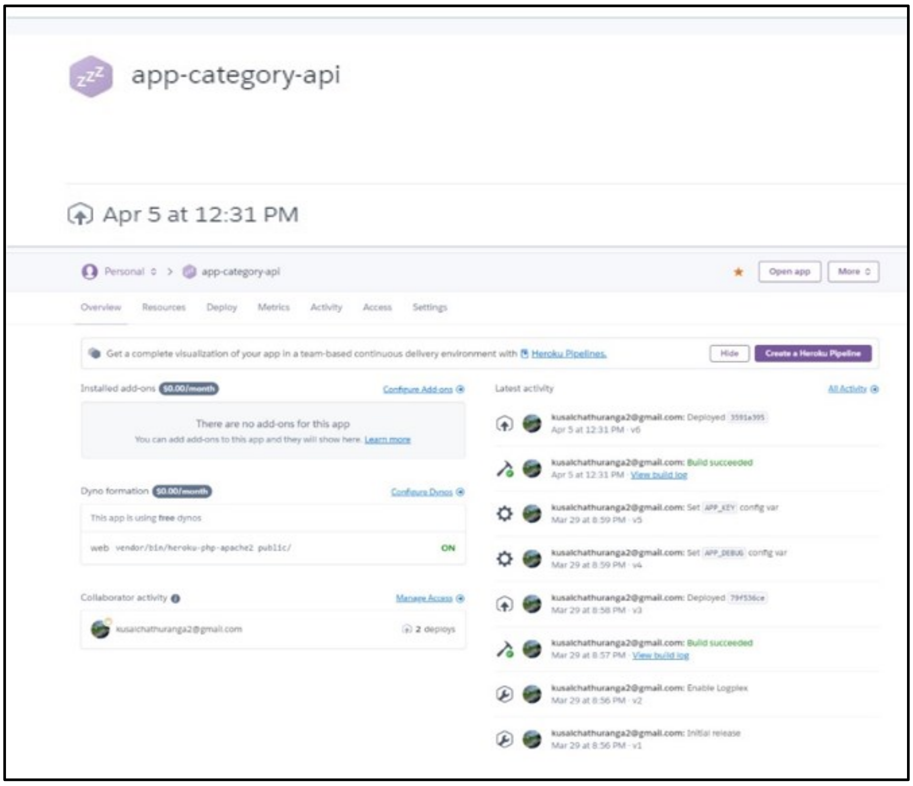


Figure 5. 10: App Type Classifying API Deployment

5.1.3. Mobile Application Development and Integration

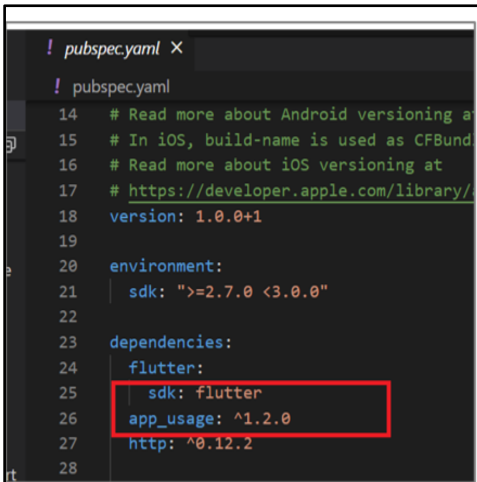
User phone usage data obtain development, model integration, calculations database integration developments will be explained in this section.

Mobile development has been conducted step by step. Let’s see how each step's development has been completed.

Obtain user phone usage

The first phase of the mobile development was to identify a method to acquire mobile utilization details for EDRAP. For the analysis of user’s total number of social apps, Total hours of social app usage, Total hours of gaming app usage, and Total hours of nighttime app usage need. Therefore, obtain these data without users’ contribution was the intention. The method used to obtain the above-identified data is used flutter “app - usage” package. The package has integrated to the developing mobile application and obtain data. Before instigate, the package the min SDK version of the mobile application has changed. Because the package can not apply if the min version was less than 21. Therefore the SDK version was changed accordingly. Figure (5.11) exhibits the package integration and Min SDK version update.

After the integration, the information of the apps has been obtained. But yet there



App _ Usage package integration.

Min sdk vesion.

Figure 5. 11: Package Integration and Min SDK Version Update

was a long way to get the data as the requirements to the EDRAP. The data have gone through several stages of cleaning and pre-processing to come as, the Total number of social apps, Total hours of social app usage, Total hours of gaming app usage, and Total hours of nighttime app usage.

Frist the raw obtained app details needed to be categorized into main two groups which are social and gaming apps. The API developed for App Categorizing is used in this stage the API has integrated into the mobile application and then used in "getCategroyWithPackageName" method to passed the mobile application package name to get the type of apps. The mobile application package name was data that we have acquired from the package "app_usage". Then the apps have been Categorized into main groups according to the types. The systems apps have been ignored in the categorization. Figure (5.12) shows the App Categorizing API integration and Figure (5.13) shows the application separation to main categories.

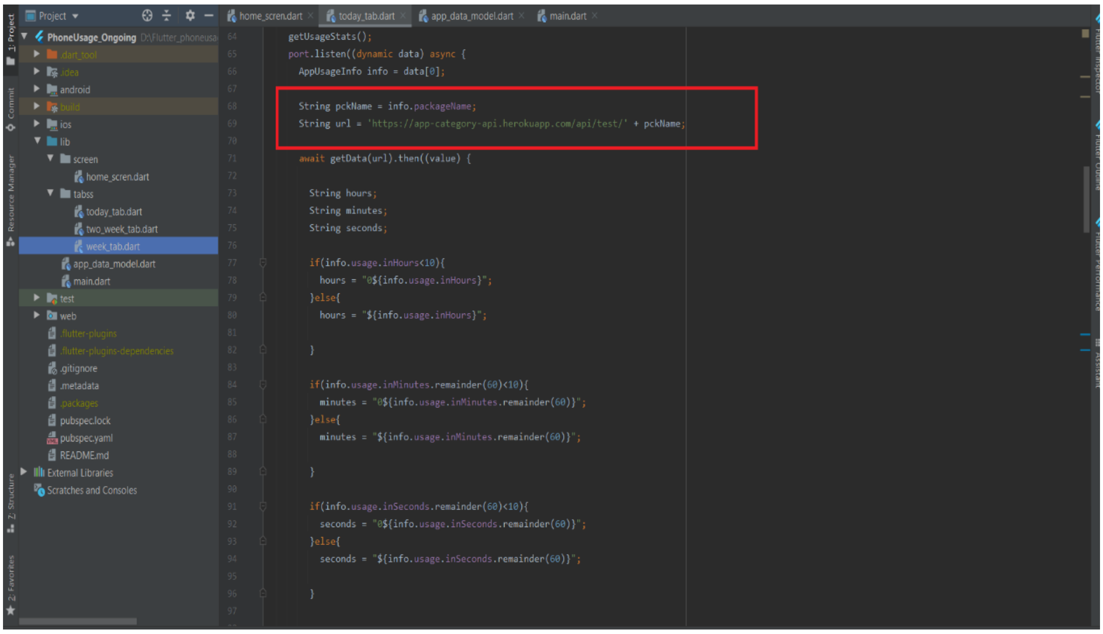


Figure 5. 12: App Categorizing API Integration

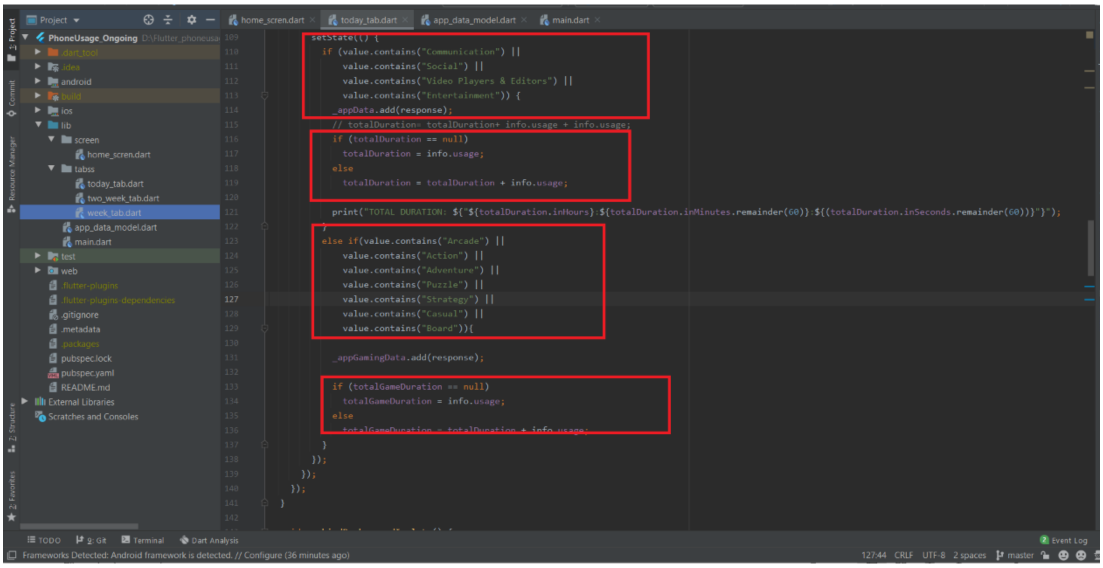
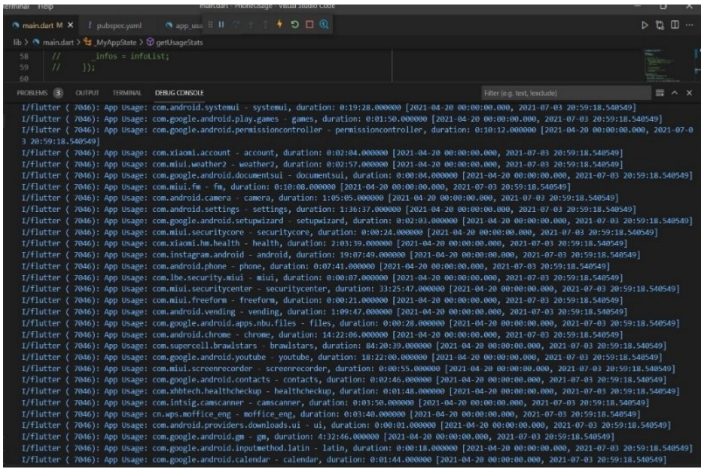
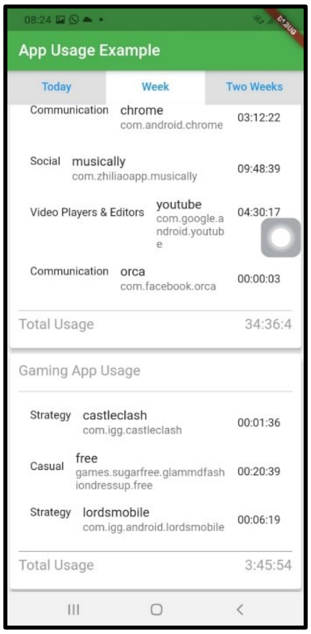


Figure 5. 13: Application Separation to Main Categories Implementation.

After the categorizing the other organizing parts have been conducted which are calculate apps nighttime usage, calculation the total apps that have been used, and calculating the total hours of app usages. The before and after organizing the data shows in Figure (5.14).



Before organize.



After organize.

Figure 5. 14: Comparison Of Organizing the Data

Database Integration and Store data.

In this step, the obtained and cleaned data have stored in the database. Before that, an appropriate database must be integrated into the application. The database that is used for data storing is the firebase realtime database. There were some steps to follow in order to integrate the firebase into the mobile application. The steps have been followed accordingly as in the Installation & Setup on Android by firebase [20]. Furthermore, some rules have been added to the firebase for the process of getting remove the 15th-day data after 14 days have been completed. This rule has been added because the EDRAP is for 14 days of monitoring. Below Figures respectively indicate the firebase rules, stored data, and firebase integration.

```
"rules": {
  "UsersData": {
    "$uid": {
        ".indexOn": ["timestamp"],
    },
  },
  "GoogleFitData": {
    "$uid": {
        ".indexOn": ["timestamp"],
    },
  },
  "Non_Invasive": {
    "$uid": {
        ".indexOn": ["timeStamp"],
    },
  },
 ".read": "auth.uid != null",
  ".write": "auth.uid != null"
}
```

Figure 5. 15: Firebase Rules.

Figure 5. 16: Store Data In Firebase.

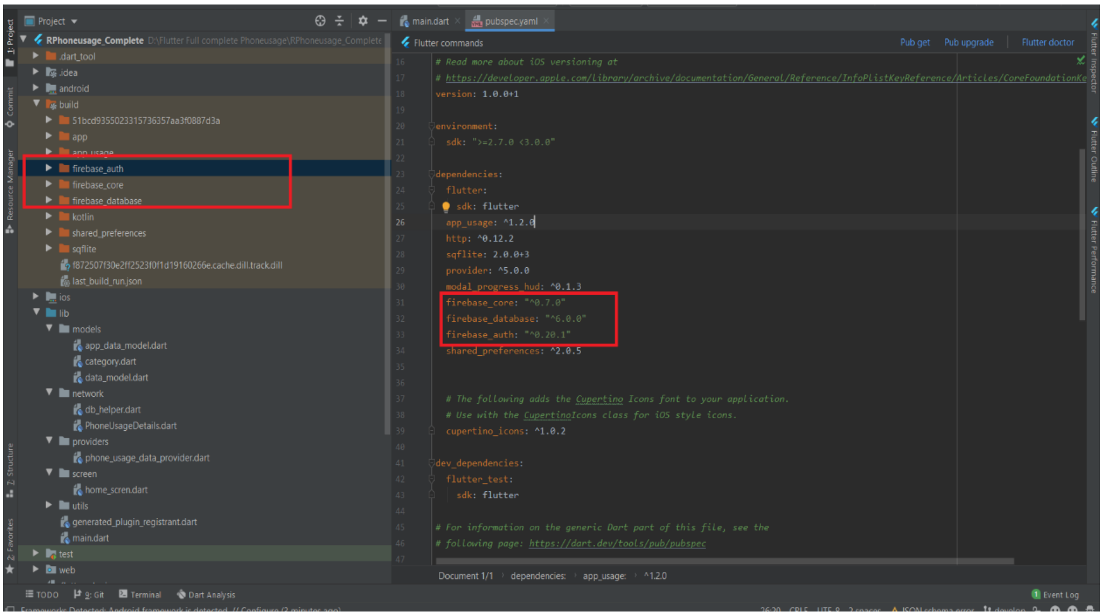


Figure 5. 17: Firebase Integration

Model Integration to the mobile Application.

After storing the data for further analysis. A calculation process has been conducted after the data have access from the database. Calculate total hours of nighttime usage, calculate the total apps used and etc have consisted in the step. Next, the data passed to the prediction stage to get the analysis output but for getting predictions, the developed model APIs needed to integrate into the mobile application. Frist the APIs have define globally and used in necessary methods which are “getSocialApp”, “getGamingAppDuration” and “getPredictData”. Below Figures respectively indicate the API initialization, API use in relevant methods.

As the last step, the organized data and prediction output have passed to the frontend. In order to pass the data to the frontend, the front-end developments need to be complete as design in the requirement analysis phase.

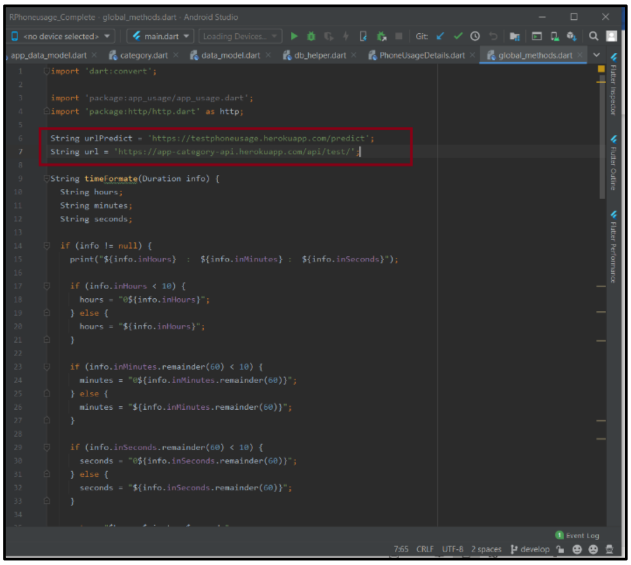


Figure 5. 18: API Initialization

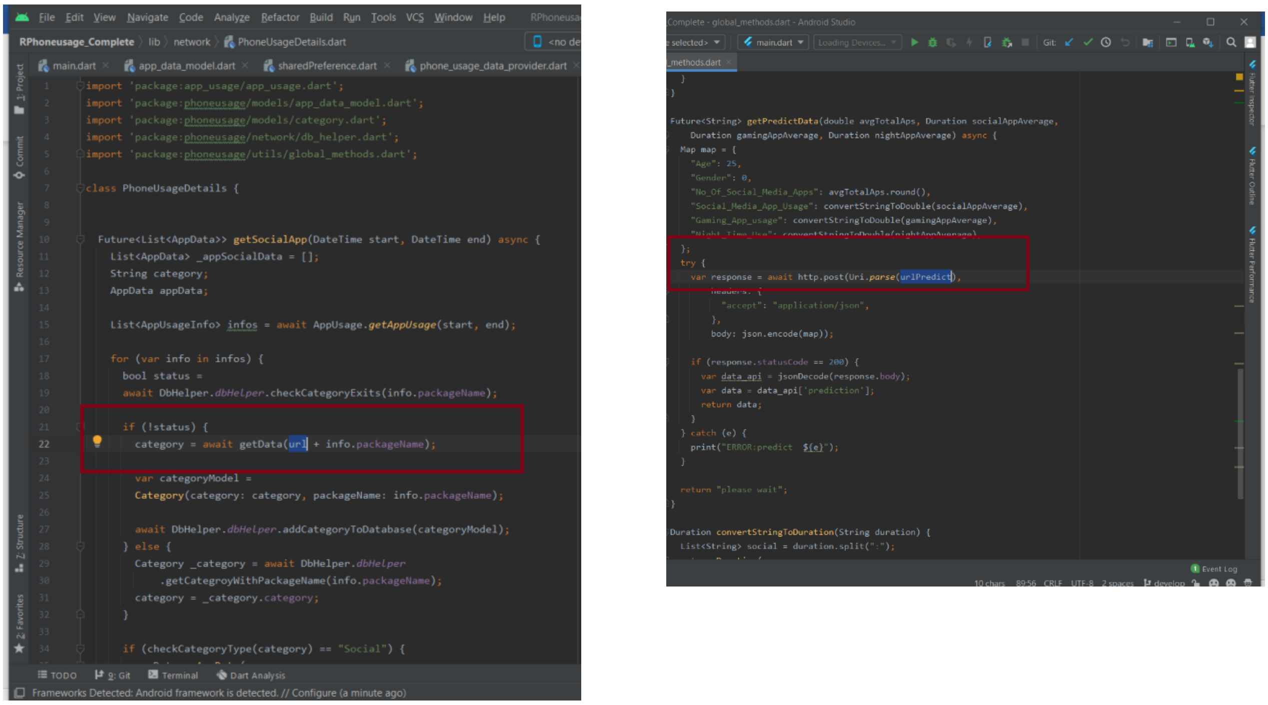


Figure 5. 19: API Using

5.2. Testing

The last step of all the implementations ended with the testing phase. In this phase, all the implementation testings have been conducted. In this section, model testing, API testing, mobile front-end testing, and mobile backend testing will be explained.

5.2.1. EDRAP Model Testing

The models were reviewed and discussed after they were deployed for EDRAP to determine the optimal accuracy and performance. The data from the dataset has been used to evaluate the chosen model KNN for EDRAP. Section () explains the analysis and discussion.

The data of two people were chosen to test the High Risk of Depression (HRD) and the Low Risk of Depression (LRD) of the EDRAP model. The test cases are shown below in Table (5.1).

Table 5. 1: Model Test Cases.

Test Case	**Test data**	**Expected value**	**Actual Value**	**Status**
T001	No Social Media Apps: 9 Social Media App usage:6 Gaming App Usage: 7 Night Usage: 12	1	HRD (1)	Pass
T002	No Social Media Apps: 5 Social Media App usage:4 Gaming App Usage: 3 Night Usage: 3	0	LRD (0)	Pass

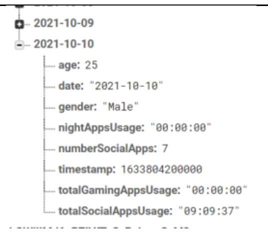
According to the EDRAP model testing table, High usage of social media applications, gaming apps, and nighttime phone app usage of 9 hours, 6 hours, and 12 hours, respectively, indicates a high risk of depression symptoms. On the other hand, someone with 4 hours of social media usage, 3 hours of gaming app usage, and 3 hours of night usage has a low risk of depression.

5.2.2. Mobile Application Testing

Testing of the mobile application was carried out to guarantee that the process and functionalities perform as intended and in the desired manner. Using test cases, mobile application testing was conducted. Table (5.2) indicates the mobile application test cases.

Table 5. 2: Mobile Application Test Cases.

Test Case	**Test Scenario**	**Test Step**	**Test Data**	**Expected result.**	**Status**
T003	Verify whether the user can give consent to access the phone usage data.	Install the app to mobile phone. On the consent button that we can give consent.	------	If consent is provided correctly the application flow leads to the home screen.	**Pass**
T004	Verify whether the collected phone usage data display to the user	Go to the phone usage icon from navigation.	Date: 2021 – 10 – 04 Social app usage 10:35:19 Gaming app usage: 00:00:00 Night app usage: 9:15:22 Total social app: 8	Display phone usage data with date and digrams.	**Pass**

T005	Check whether the collected data pass to the firebase.	Log into the firebase account using email and password.		Show date, number of social apps, total gaming app usage, total social app usage on firebase db.	**pass**
T006	Verify whether the mobile application plugins are integrated properly.	Add the plugins to the pubspec.yaml. Run “flutter pub get” on the console.	app_usage: ^1.2.0 firebase_database: "^6.0.0"	Display 0 errors on the console.	**Pass**

5.2.3. API Testing

EDRAP analysis model API and App categorizing API have been developed and deployed. As the last step of implementation, the APIs needed to be Tested.

Postman was used to testing the KNN model API. The request body data was sent in JSON format, and the response was obtained in JSON format as well. The URL, Body, and Response of the Request are as follows,

Post request: https://testphoneusage.herokuapp.com/predict

Body: JSON

{
  "Age": 24,
  "Gender": 1,
  "No_Of_Social_Media_Apps": 8,
  "Social_Media_App_Usage": 6,
  "Gaming_App_usage": 12,
  "Night_Time_Use": 10
}

Response:

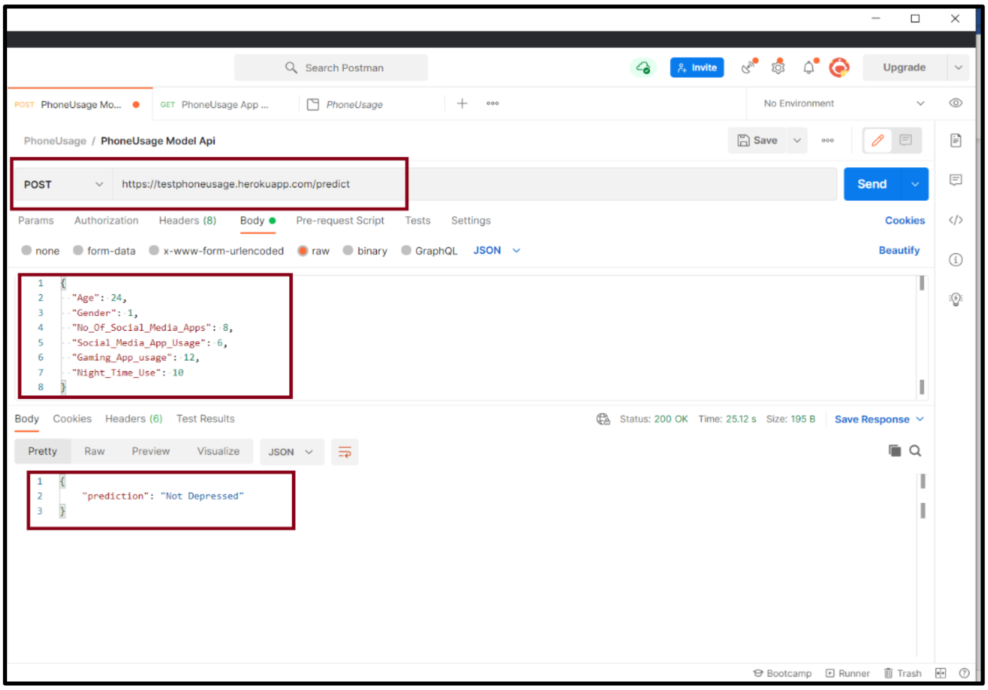


Figure 5. 20: EDRAP KNN API Testing

The App categorizing API has been Tested using Postman and the GET request is passed along with the app package name. to get the Type of the app. The Request URL, Body, and Response indicates as following,

GET request: https://app-category-api.herokuapp.com/api/test/com.facebook.mlite

Response:

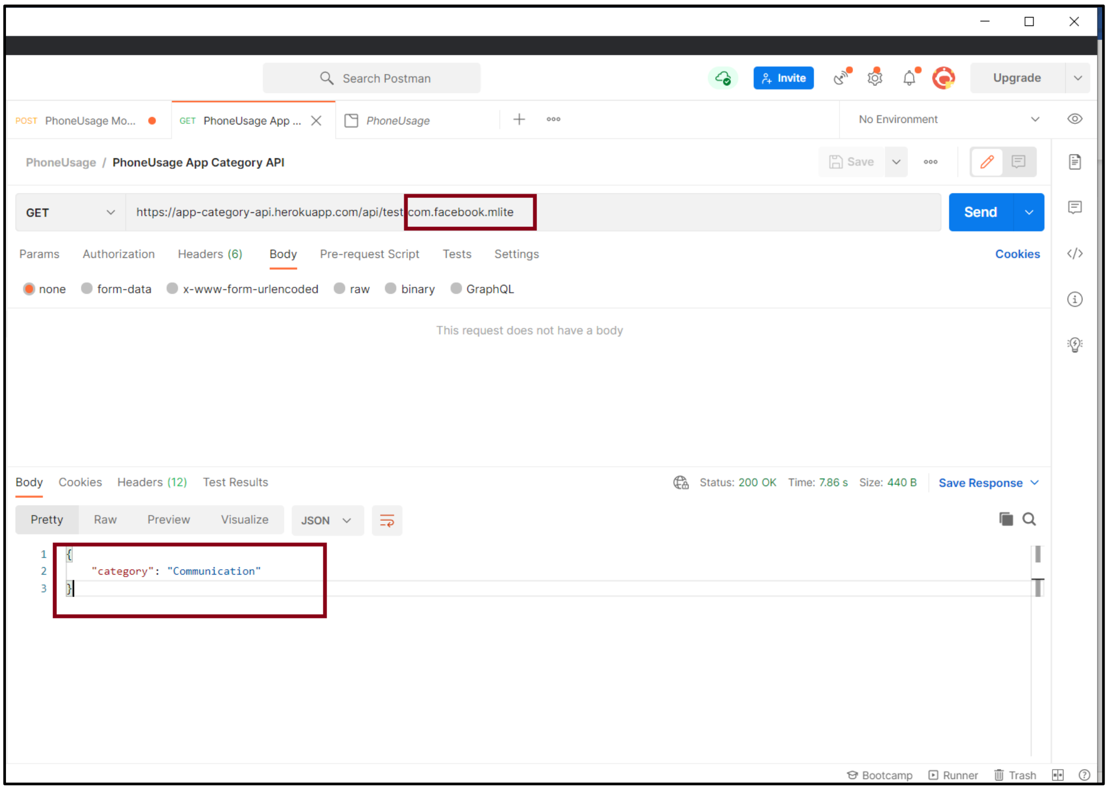


Figure 5. 21: App categorizing API Testing

6. RESULTS & DISCUSSION

6.1. Results

The preferred models for EDRAP analysis have been chosen based on the findings of the algorithms, as well as the app categorizing technique results, which have been presented in this section.

Model Results

At k = 14, the KNN algorithm achieved 81% test accuracy, whereas the train set maximum accuracy was 82%. The AUC of the model is 81%, the sensitivity is 31%, and the specificity is 87%, and the model has a 17% misclassification rate. The random forest model has a 62% overall accuracy, while sensitivity and specificity are 37% and 73%, respectively, with a 38% misclassification rate. Even though the SVM achieved 72% test accuracy, 63% AUC, and 27% misclassification rate, the model did not reach a decent sensitivity rate. The Naive-Bayes algorithm, on the other hand, achieved 76% AUC, 65% sensitivity, and 88% specificity. The Confusion Matrix of the KNN model is shown in Figure (6.1).

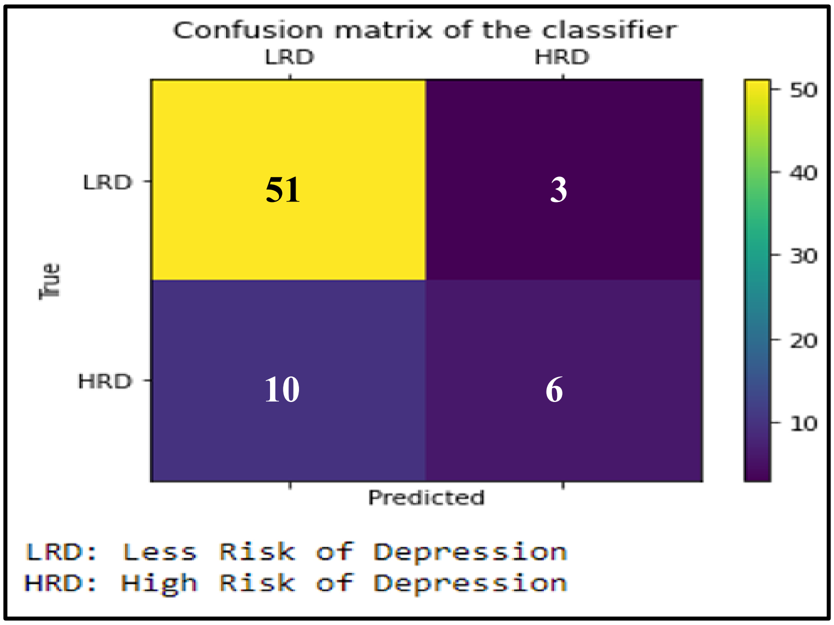
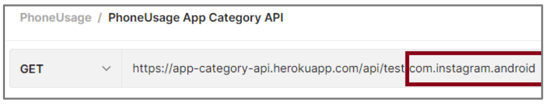
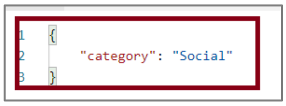
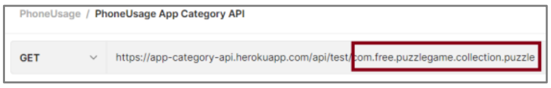
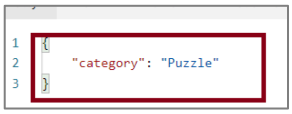
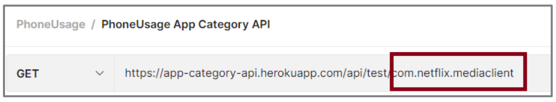
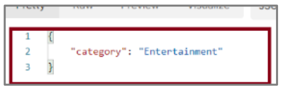
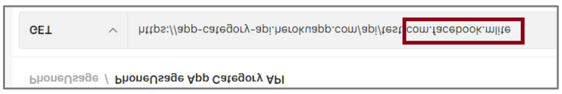
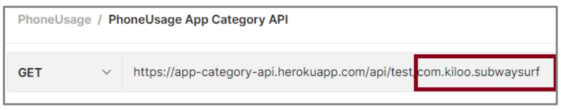


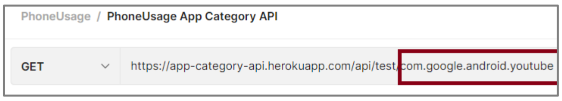
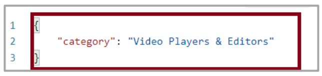
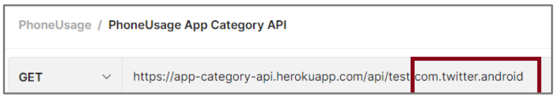
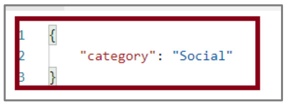
Figure 6. 1: Confusion Matrix of the KNN model

App Categorizing Method Results.

As mentioned in section (5.1.2), the app categorizing technique was created from the ground up. The method's output is the creation of an API, which allows us to pass the package name of a mobile app and obtain the app type. The following are some common apps that can see in mobile phones and type that obtain from API,

Table 6. 1: App Categorizing Method Performance.

App Name	Package Name	Type
Instagram	com.instagram.android 	Social 
Super Brain Plus (gaming App)	com.free.puzzlegame.collection.puzzle 	Puzzle 
Netflix	com.netflix.mediaclient 	Entertainment 
Facebook	com.facebook.mlite 	Communication 
Subway Surf (Game)	com.kiloo.subwaysurf 	Arcade 

Youtube	com.google.android.youtube 	Video Player & Editor 
Twitter	com.twitter.android 	Social 

6.2. Discussion

To determine each algorithm's performance, the results of the EDRAP trained algorithms were analyzed. The construction of a suitable model for EDRAP analysis resulted from the identification of methods. The comparative performance evaluation of the above-explained algorithms is shown in table (6.1).

Table 6. 2: Model Algorithm Performance Comparison.

	Algorithm	**Accuracy**	**PRECISION**	**RECALL**	**F1-SCORE**
EDRAP	KNN	81%	0.72	0.74	0.73
	SVM	72%	0.52	0.72	0.61
	Naive Bayes	76%	0.69	0.65	0.66
	Random Forest	62%	0.64	0.63	0.63

When comparing the results of the algorithms, KNN and SVM have the almost same different highest sensitivity which is 74% and 72%, but SVM's accuracy was significantly lower than KNN's. The specificities of the SVM, Random Forest, and Naive Bayes algorithms are scale in a similar range, and the KNN algorithm has a greater proportion of specificity factors. As we analyze the findings, the KNN has the highest accuracy, with a 74% chance of correctly identifying a person at risk of depression. As a result, the KNN model is the best to use for future forecasting. Note that all of the models' datasets were split into 80 training and 20 test sets.

The app classifying API performed as expected and delivered all of the categories required for categorizing users in a mobile app. However, there was a restriction of the API in that it took a long time to acquire app type when the API was integrated into the mobile application. Backend techniques were employed to overcome the problem at that time. Aside from that flaw, the app classifying method's results are dependable and appropriate for usage in the app.

Finding data for phone usage analysis from internet databases was an impossible task therefore analysis has continued with the collected data from a foram. Moreover, Due to the pandemic situation, and obtaining data from individuals was also challenging. Limit in place. The EDRAP model has not been tested on a real-life dataset of depressed people. The reason for this is that due to the Covide – 19 pandemic, locating a truly depressed person dataset was a difficult task.

7. CONCLUSION

The project complement's main goal was to come up with an analysis that can identify individuals with early depression risk according to their phone usage patterns. The analysis was performed with supervised learning classification algorithms to come up with the most preferable model for analysis. The KNN model has achieved the most suitable performance with accuracy, F1 – score, recall, and AUC. In order to make the analysis convenient and identifying the concerns in the industry, the model has been extended to integrate with a mobile application for a two-week monitoring process.

REFERENCE

[1] WHO, World Health Organization, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression. [Accessed 18 February 2021].

[2] M. A. Villarroel and E. P. Terlizzi, "Products - Data Briefs - Number 379 - September 2020," Cdc.gov, 2021. [Online]. Available: https://www.cdc.gov/nchs/products/databriefs/db379.htm. [Accessed 6 January 2021].

[3] S. Jun, "The reciprocal longitudinal relationships between mobile phone addiction and depressive symptoms among Korean adolescents," Computers in Human Behavior, vol. 58, no. 0, pp. 179-186, 2016.

[4] R. J. J. M. van den Eijnden, G.-J. Meerkerk, A. A. Vermulst, R. Spijkerman and R. C. M. E. Engels, "Online communication, compulsive internet use, and psychosocial well-being among adolescents: A longitudinal study.," Developmental Psychology, vol. 44, no. 3, pp. 655-665, 2008.

[5] S.-S. Cha and B.-K. Seo, "Smartphone use and smartphone addiction in middle school students in Korea: Prevalence, social networking service, and game use," Health Psychology Open, vol. 5, no. 1, 2018.

[6] y. J. Jin and . D. H. Kim, "A study on the addicted use of mobile phone among the high school students.," vol. 27, no. 1, pp. 140 - 153, 2005.

[7] P. So-Young, Y. Sonam, C.-S. Shim and H. Jang, "Long-Term Symptoms of Mobile Phone Use on Mobile Phone Addiction and Depression Among Korean Adolescents," International Journal of Environmental Research and Public Health, vol. 16, no. 19, p. 3584, 2019.

[8] L. y. Lin, J. E. Sidani, A. Shensa, A. Radovic, E. Miller, J. B. Colditz, B. L. Hoffman, L. M. Giles and B. A. Primack, "association between social media use and depression among u.s. young adults," Depression and Anxiety, vol. 33, no. 4, pp. 323-331, 2016.

[9] "Late-night screen time puts teens' sleep and mental health at risk," Cbsnews.com, 2021. [Online]. Available: https://www.cbsnews.com/news/late-night-cell-phone-screen-time-teens-mental-health-sleep/. [Accessed 17 February 2021].

[10] K. Demirci, M. Akgönül and A. Akpinar, "Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students," Journal of Behavioral Addictions, vol. 4, no. 2, pp. 85-92, 2015.

[11] Višnjić, V. Veličković, D. Sokolović, M. Stanković, K. Mijatović, M. Stojanović, Z. Milošević and O. Radulović, "Relationship between the Manner of Mobile Phone Use and Depression, Anxiety, and Stress in University Students," International Journal of Environmental Research and Public Health, vol. 15, no. 4, p. 697, 2015.

[12] S. Thomée, A. Härenstam and M. Hagberg, "Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults - a prospective cohort study," 2020.

[13] E. Hoare, K. Milton, C. Foster and S. Allender, "Depression, psychological distress and Internet use among community-based Australian adolescents: a cross-sectional study," BMC Public Health, vol. 17, no. 1, 2017.

[14] A. Chowdhry , "Research Links Heavy Facebook And Social Media Usage To Depression," 2017.

[15] R.-E. Mastoras, D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. Kassie, T. Alsaadi, A. Khandoker and L. J. Hadjileontiadis, "Touchscreen typing pattern analysis for remote detection of the depressive tendency," Scientific Reports, no. 1, 2019.

[16] R. Wang, W. Wang, A. DaSilva, J. F. Huckins, W. M. Kelley, T. F. Heatherton and A. T. Campbell, "Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing," Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 2, no. 1, pp. 1-26, 2018.

[17] "New app tracks your mental health based on your phone usage," bodyandsoulau, 2021. [Online]. Available: https://www.bodyandsoul.com.au/mind-body/wellbeing/theres-now-an-app-that-tracks-your-mental-health-based-on-how-you-use-your-phone/news-story/4b837e48191ecc524cea49154355093d. [Accessed 15 February 2021].

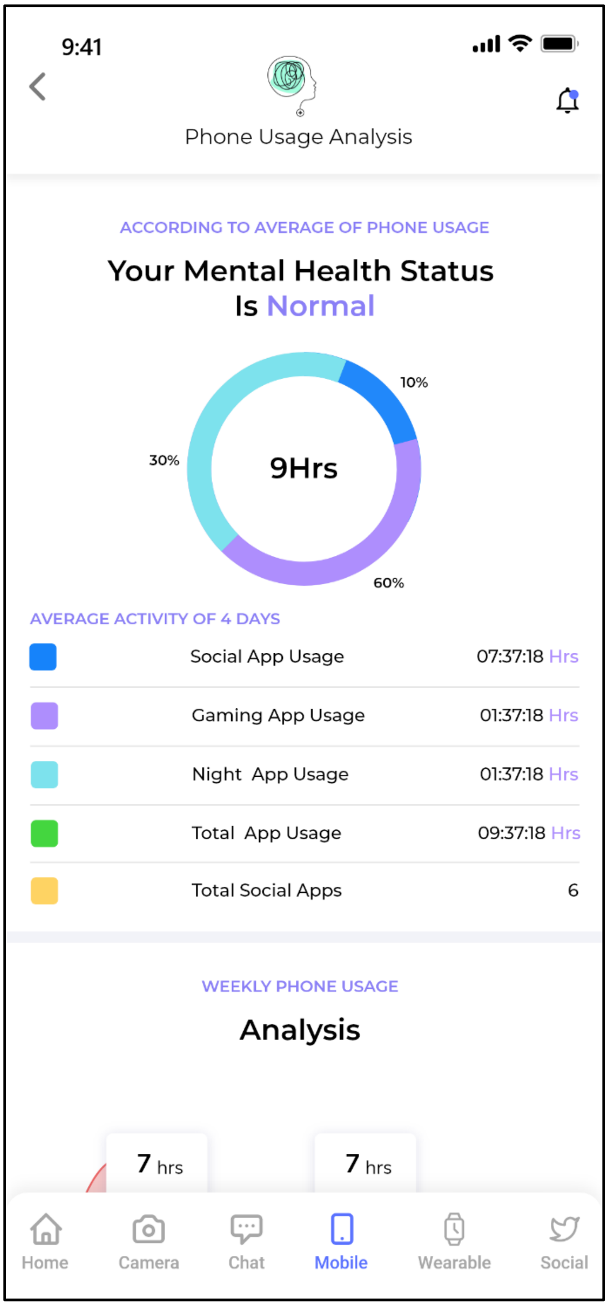
[18] G. Play, "App Usage," Play.google.com, [Online]. Available: https://play.google.com/store/apps/details?hl=en&id=com.a0soft.gphone.uninstaller. [Accessed 15 February 2021].

[19] D. K. Kroenke, "PHQ-9 (Patient Health Questionnaire-9)," Mdcalc, [Online]. Available: https://www.mdcalc.com/phq-9-patient-health-questionnaire-9.

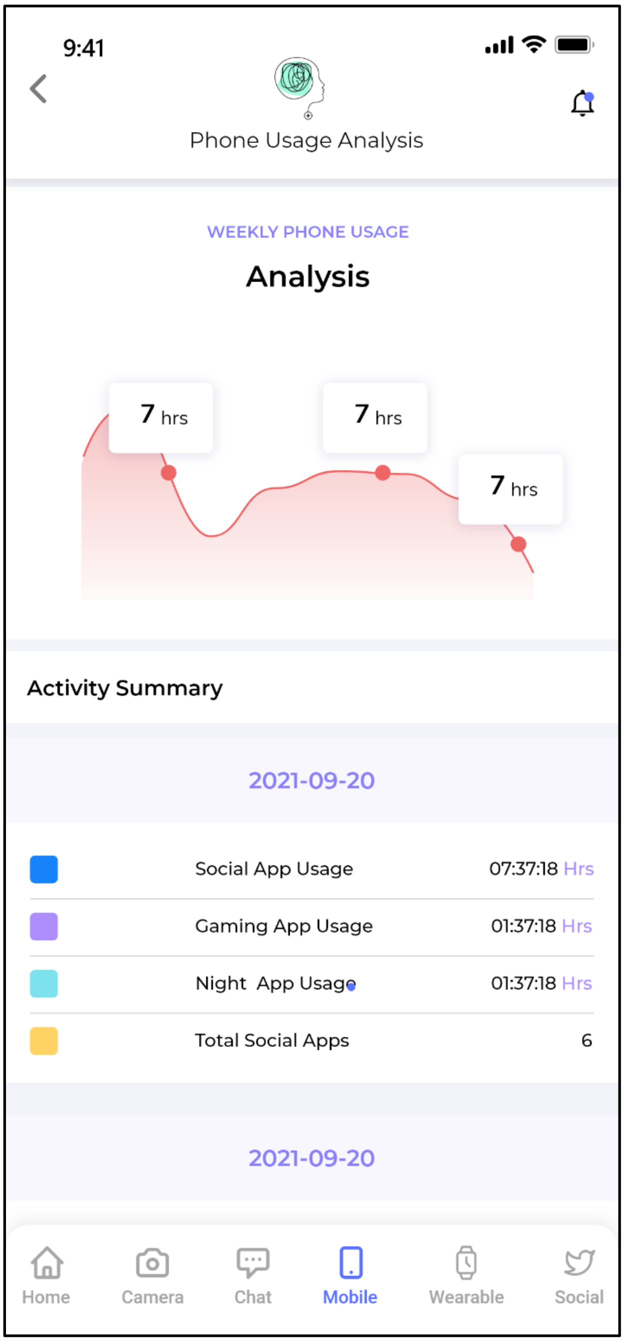
[20] Firebase, "Installation & Setup on Android | Firebase Documentation," Firebase, [Online]. Available: https://firebase.google.com/docs/database/android/start. [Accessed 21 June 2021].

Appendix A: UI Wireframes

Average Mental Health Summary



Total Phone Usage Graph.



Daily Phone usage details.