BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Chathuranaga W.W.P.K

(IT18119572)

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.

Chathuranaga W.W.P.K

(IT18119572)

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 acknowledgement 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 acknowledgement 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 other 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
Chathuranaga W.W.P.K	IT18119572	

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 most important aspect of life. Emotional, psychological, and social well-being is mostly based on good mental health. Mental illnesses, on the other hand, impede people from enjoying their lives to their full potential. Depression is one of the most frequent mental health problems that people suffer from. People with serious depression disorders may resort to drastic measures such as suicide. As a result, identifying depressed people and persuading them to take medicine is critical.

However, clinical psychologist sessions are the conventional technique of identifying people, but the social norms on a mentally ill person going to a psychologist are quite modest. As a result, people tend to hide their disorder, which has serious consequences for them. Furthermore, standard approaches have difficulty monitoring patients over two weeks.

The development of a mobile application that allows users to recognize depression analysis is a modern and technological answer to this problem. We hope to explain a method that analyzes early depression risk based on biometric data obtain invasively and non – invasively in this study. A mobile application based on phone usage analysis is included in the system. The research is based on supervised learning algorithms and mobile apps written in the flutter programming language.

Keywords: Depression risk, biometrics analysis, Machine learning, Flutter.

ACKNOWLEDGEMENT

I would like to express my gratitude to Dr. Pradeepa Samarasinghe, supervisor of our behavior analysis for the depression detection research group for giving us guidance to complete our undergraduate project successfully. Besides our supervisor, I want to thanks Ms. Shalindi Pandithakoralage our external - supervisor for proving us with guidance in the area of mental health aspects. I would like to thank our co-supervisor Ms. Vijani Piyawardana for giving us support on the research.

Moreover, I want to thank the lecturer in charge, Dr. Janaka Wijekoon, and the CDAP team for guiding us to complete the research in a timely manner.

Lastly, I am grateful to my friends and family for the support and encouragement provided while the research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EDRAHR: Early Depression Risk Analysis on Heart Rate

EDRASP: Early Depression Risk Analysis on Sleep Pattern

KNN: K-Nearest Neighbour

SVM: Support Vector Machine

UI: User Interface

1. INTRODUCTION

1.1. Background & Literature Survey

1.1.1. Background

Mental health keeps human emotional, psychological, and social well-being factors in comfort manner not only the way of thinking, how humans feel about people, human determinations, and even emotion exchange depends on Mental health. A healthy mind is an important factor for childhood to the elderly well-being. Every person has mental ups and downs in certain periods of life but when the mentality down becomes an ongoing part of life it is a mental illness we call a mental disorder. Mental disorders can be very among drastically points Anxiety disorders, Depression, mood disorders, post-traumatic stress disorder, and Eating disorders, etc. As information to World Health Organization (WHO), one of the leading causes of disability is Depression.

Depression has become the second leading cause of deaths age between 15 – 19 worldwide [1] therefore, Early-stage identification of depressive disorder is crucial. However, Identifying depression has been accomplished by a periodic assessment which can be structured or unstructured clinical interviews that use standard rating scales and advanced methods for severe individuals. The bright side is identification methods shifting from narrow methods which depend on scaling to broader anatomical and neurophysiological understanding of emotion, behavior, cognition, and their disorders [2]. We anticipate the development of a reliable mobile application that uses biometric data for early detection that will help improve the diagnosis and assessment of depression therefore, biometrical data from individuals obtained using wearables and non-wearables, can be crucial. Researchers have focused more on biometrical data such as heartbeat, sleep patterns, skin temperature, and mobility to study symptoms of depression ex: [3] [4].

When it comes to psychiatric illnesses sleep patterns of individuals take an important part in both a clinical symptom and an important curative target [5]. Therefore, sleep pattern data can be used as an important parameter to identify early depression changes moreover insomnia can be a sign of leading to depression [6]. Sleep disturbances are usually a risk factor for amplifying anxiety and depression [7]. Heart rate changes during sleep can be a biomarker of depression therefore heart rate is an important parameter to consider identifying early depression, moreover, increased and reduce HRV in both conscious and sleep is a link with depression [8]. When we seek to achieve identify early depression by behavioral changes using biometric data, we must examine heart rate and sleep pattern as important parameters.

1.1.2. Literature survey

One of the most frequent physical changes of a depressive patient is sleep, in the research Epidemiology of insomnia by Maurice M Ohayon have mentioned approximately 80% of individuals with a current major depressive episode have co-occurring sleep difficulties [9]. Besides sleep patterns, heart rate is another physical factor that we can recognize in depression patients therefore the research from the South China University of Technology, Guangzhou with Danni Kuang et al. proved Depression patients have lower HRV than healthy subjects. Therefore, HRV may be used to distinguish depression patients from healthy people by using Bayesian Networks [10].

A study from the Faculty of Science and Technology, Keio University, Kanagawa, Japan has developed a machine-learning algorithm to screen for depression and assess severity based on data from the wearable device [11]. In the study they have used measure step count, body movement, sleep time, heart rate, skin temperature, etc. during their daily activities. The wearable has been used as the data collection method for implementing the machine-learning algorithm. Moreover, the team has investigated the relationship between depression and each modality that they have used for the research.

Dartmouth College USA researched track depression using an app and wearable data they took heart rate through warble and some other parameters like sleep details, mobility through the app called The Student Life [12]. The study was conducted to identify the relationship with depression from collected data using app and wearables.

Furthermore, the team has used self-reported PHQ-8 and PHQ-4 depression scores for analysis purposes. As the discussion researches show that symptom features obtain from the phone, wearable sensors can predict whether or not a student of depressed within a week with 81.5% recall and 69.1% precision. as well as they have proved depressed persons have fewer travel movements and have more irregular mobility patterns.

There is a study from Media Lab, MIT, Cambridge, MA predicting the Hamilton Depression Rating Scale (HDRS) using data captured from E4 wearable wristbands and sensors in an Android phone [13]. The features for the study used are Physiological signals, Phone passive usage data, Interactive surveys, and Clinical measures. The E4 band was used to collect the data from users for the study. As the result of the study, the team has accomplished identifying the relationship with depressive symptoms, irregular sleep, less motion, fewer, incoming messages, less variability and etc.

A study by Isaac Moshe et al has aimed to explore the extent to which data from smartphones and wearable devices could predict symptoms of depression and anxiety. They have used wearable The Oura Ring which provides measurements related to sleep (total sleep time, sleep onset latency, wake after sleep onset, and time in bed) and heart rate variability (HRV). Through the study, research has proved significant positive associations between total sleep time and depression, time in bed, and depression [14].

Summing up, the existing studies have focused on show links between depression and biometrics, develop an algorithm to monitor depression using wearable data, and predict depression with wearable data. Most of the studies have acquired wearable data and but there can be people who refuse to use wearables. Therefore, the development of a mobile application that includes biometrical data obtained using invasive and non-invasive techniques to behavior analysis for depressive disorder, may help people for self-awareness on the mental health state. To the best of our knowledge, studies have not been found in Sri Lanka that focused on a similar approach of automated analysis on depressive disorder identification using unexpected biometric characteristic changes.

1.2. Research Gap

The Literature Servoy reveals that most of the studies focused on identifying the links between depression and features can be collected from wearables and mobile apps. As well as implement machine learning algorithms to predict depression from the data obtained with wearables. The Following Table (1.1) indicates the Comparison of current research papers.

Table 1. 1: Comparison of Conducted Researches.

	[11]	[12]	[13]	[14]	Proposed System
Depression analysis base on biometrical parameter changes.	✓	✗	✓	✓	✓
Screen and track depression using data that acquired using wearable	✓	✓	✓	✗	✓
Explore which data can be collected using a wearable to predict depression.	✗	✗	✗	✓	✓
Invasive method to acquire parameter data	✓	✓	✓	✓	✓
Noninvasive method as an alternative to acquire same parameter data.	✗	✗	✗	✗	✓
Probability of early depression analysis based on biometric data.	✗	✗	✗	✗	✓

A review has been conducted to identify the current mobile applications that focused on self-use depression analysis. There are a handful of mobile applications that have to develop considering depression disorder. We can them access through “Appstore” or “Play Store”.

The app “ Sleepio” is developed as Cognitive Behavioral Therapy for insomnia (CBTi). When the app collects data using an in-depth questionnaire from individuals then it builds a program for users to improve their sleep habits [15].

“StressScan” is a Stress level tester that collects heart rate through the phone camera then “Stress Scan” will analyze changes in the user’s heart rate interval and scientifically measure the level of your mental and physical stress on a scale of 1 to 100 [16].

" Welltory" app is also a Stress level tester but with more outputs like Measure stress, energy, and resilience with a smart. The result is evaluated using Track blood pressure and heart rate of the individual [17].

"Moodpath" is an assessment tool that aids in tracking individuals' physical and emotional well-being everyday mindfulness, sleep improvement, and stress management. The data is collected using self-reports every. As output app will retrieve a report weekly. Special features of the app are individuals can share reports with mental health professionals, self-compassion exercises are also available within the app [18]. Below table (Table 1.2) indicate a summary of existing app related to depression analysis based on invasive and noninvasive technique.

Table 1. 2: Comparison Of Current Mobile Applications

	Sleepio	**StressScan**	**Welltory**	**Moodpath**	**Proposed app**
Track most critical factors Biomatrix data	✗	✓	✓	✓	✓
Invasive method to collect Factor's data	✗	✗	✗	✗	✓
Non- invasive method to collect Factor's data	✗	✓	✓	✗	✓
Depression analysis based on Biomatrix data (HRV, Sleep pattern)	✗	✗	✗	✗	✓
Extensive report generation	✓	✓	✓	✓	✓

2. RESEARCH PROBLEM

Depression disorder has become a burden of disease all over the world with affecting more than 264 million people [1]. Depression is becoming the most common health problem among the younger generation, a considerable amount of the younger generation is affected by this mental health condition due to various reasons like self – esteems, social life, career achievements, and follow unhealthy practices for a long time. However, if we cannot identify depression in the early stages and inherit it for a long time it can cause liver damages.

When we come to Sri Lankan population, individuals endeavor to hide the symptom and avoiding medications with the negative social impact as well as most of the individual doesn't know they are suffering from depression because they don't have a solid idea about the early-stage behavioral changes that are signs of depression. The methods that have been used to identify depression are self-assessment reports, surveys, and questioners however these methods are conducting by clinical professionals therefore individuals tend to hide their disorder rather than visiting a psychiatrist.

There are biometric parameters that can be used to identify early-stage depression by behavioral changes therefore it will be effective if we can design self – awareness mobile application based on those biometric parameters. There are apps that have design considering biometrics data for depressed individuals but most of them have focused on therapy, which is giving daily plans for the effective day, sleep, etc. The limitation of those apps is they have collected data from users using surveys, questioners, or recordings.

In order to address the above issues, we can design a mobile application that is easily accessible, reliable, and accurate with the use of machine learning algorithm analysis of the biometric data collected using invasive and non-invasive methods.

The shortcoming in the industry regarding depressive disorder can be summarized below.

- Difficulties of two weeks manual monitoring individuals.
- The social impact that makes individuals avoid medication and understand symptoms.
- Lack of automated applications that focus on identify Early depression risk analysis disorder using biometric parameters.
- Lack of systems that acquire data without the user’s contribution.

3. RESEARCH OBJECTIVES

The objective of the research is to Early depression risk analysis signs that may lead to being a depressive person. The study proposes to focus on a mobile application to identify behaviors that could be highly anticipated to contribute to depressive disorder.

3.1. Main Objective

The main objective of the proposed research is to Acquire invasive and non-invasive data analysis for early depression risk analysis identification.

3.2. Specific Objectives

The Main object is dived into sub-objectives as following.

1) Recognize health parameters that can be collected using a wearable device to identify early depression symptoms.

There are several biometric data that we can use to identify depression, but we need to recognize the most suitable and accurate parameters that we can collect from the Invasive method.

2) Identify techniques as alternative methods to collect the same parameter data that collect from the wristwatch.

There should be an alternative method to collect the same parameter data that collect from wearable because there can be times warble won't work as well as individuals may not use the wearable.

3) Collect data using invasive (wearable) and non-invasive techniques and build methods to identify early depression symptoms

The model will be trained using a machine learning algorithm that is most accurate among several algorithms.

4) Predict the probability of abnormal physical changes towards depression with the designed models.

If the user can lead to depression or not will be identified from the prediction based on biometric data.

5) Integrate both models into the mobile app.

After Serialization of the prediction from model outputs will be retrieved to the designed mobile app.

There are additional objectives that include mobile application development is explained below,

6) Implementing interfaces of the mobile application.

Fronted development of the proposed mobile application will be complete in this stage.

7) Mobile Application development.

The mobile application development will be developed using flutter to collect the data from the user as well as to retrieve the analysis data from the user.

3.3. Functional Requirements

- Analysis biometrics parameter data to identify individuals with abnormal behavioral changes toward depression.
- Predict the probability of having early depression symptoms based on analysis of biometric data gathering invasively & non-invasively.
- Two weeks early depression risk analysis must be provided to the user.
- Users must be provided with necessary notifications when accessing the phone sensors.
- Generate the comprehensive report summary.

3.4. Non - Functional Requirements

- Security: The system must be aware that the wearable is connected without error to the application before collect data.
- Privacy: the system must be responsible to keep the confidentiality of the user data and analysis information.
- Reliability: Biometric data analysis must provide highly accurate probability results to the user about the depression
- Usability: The application should provide an analysis summary in a user-friendly and understandable manner, as well as the functions and navigations, must be simple and efficient.
- 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 check a daily summary of biometric user data.
- The user should be able to get two weeks' early depression risk analysis details.
- The user should be able to connect warble to the mobile application.
- The User should be able to review if the data collected from sensors properly.
- The user should be able to provide heart rate daily.

4. METHODOLOGY

4.1. Introduction

In this section, the Invasive non - invasive Early depression risk analysis application development process will be explained.

The development process consists of the following stages.

Requirement Gathering and Analysis: In the analysis, stage identifying the gaps and come up with a mobile application has been conducted as well as mobile application requirement organizing has been completed in this stage.

Mobile Application Design Thinking: As gathered information and analysis the "MindGuardian" mobile application design and features must be designed aiming at the target group of individuals. Therefore, design thinking and finalizing have been completed in this stage.

Analysis Model Implementation: The mobile application runs along with a depression risk analysis therefore, the analysis model implementations have been conducted with several sub-stages like data collection, pre-processing, implantation, model API implementation, and API deployment.

Mobile Accelerometer Sensor Data to Sleep Parameters: The section non – invasive developments consist of the development of the data accruing from user to mobile application non - invasively. This stage can be dived into sub-stages which are accelerometer convertors to sleep pattern implementation, API implementation of the converter, and API deployment.

Mobile App Development and Integration: The mobile application development stage consists of several sub-stages which are Invasively user data acquire implementation, Non – invasively data acquire implementation, model integration, UI implementation, and database integration has been conducted.

Testing: In this stage Model testing, Application front-end testing, Application back-end testing, and API testing has been completed.

4.2. System Flow.

The smartphone application "MindGuardian" Invasive and non-invasive procedures the early depression risk analysis approach begins with the user installing the app and connecting the MI band 5 for invasive data collection or non-invasive data collection utilizing the accelerometer and phone flasher. For the user's convenience, both invasive and non-invasive approaches have been upgraded. Users can contribute biometric data using either an invasive or non-invasive technique. The raw biometric data is cleaned and processed as needed for analysis before being sent to the Firebase database. The analysis is being carried out using the database data that has been accessed. The health information and analytical summary are displayed to the user as the final step. In detail, procedures are explained in section (5.1).

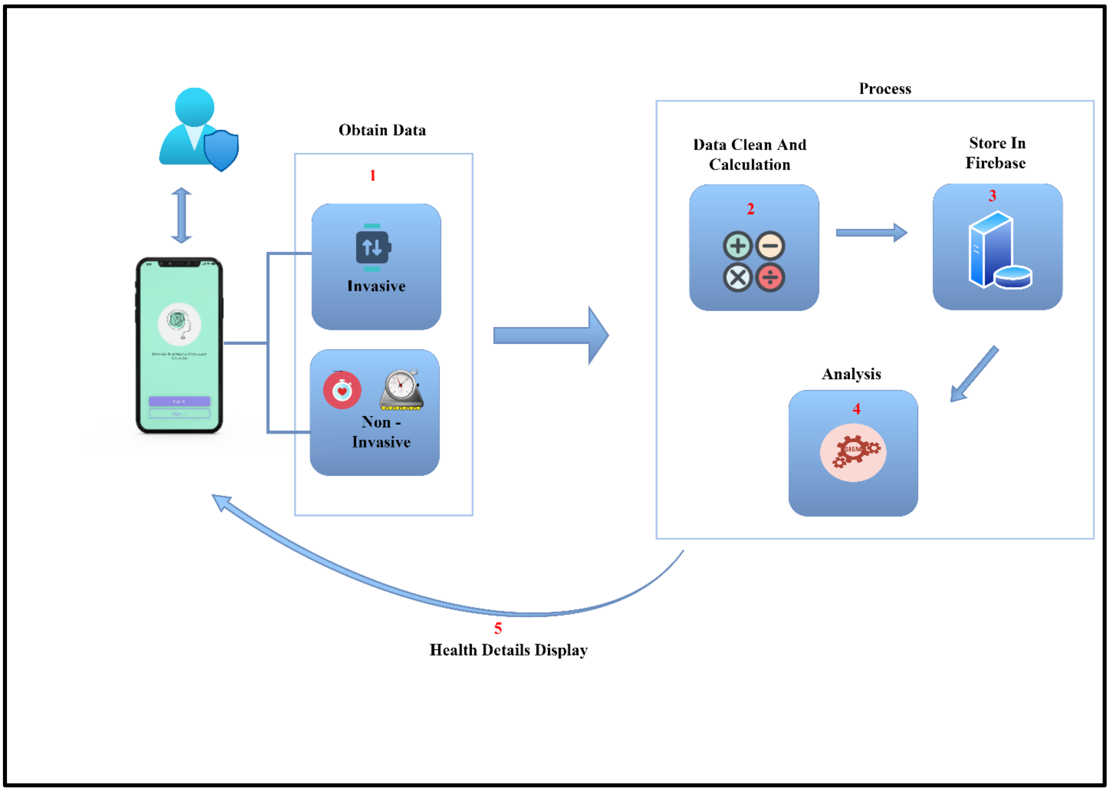


Figure 4. 1: Invasive and Non-Invasive System process.

4.3. Mobile Application Design Thinking

Requirement Gathering and Analysis

Initially, the requirements gathering has been conducted with Clinical Psychologist moreover, the current researchers within the area have been analyzed to identify the gaps of the researchers as section 1.2, TABLE (1.1).

The current mobile application features and biometric parameter analysis methods have been studied to identify the downsides of the current apps for improving them within our "MindGuardian" application. The current app analysis has been discussed in section 1.2 TABLE (1.2).

Design Thinking.

The application design and features have been organized according to the necessities and cognitive level of the target research group. The application design phase has been conducted and confirmed by Clinical psychologists in order to provide better performance to the users. Figure (4.2) despite the application flow with UI designs. Following aspects have been followed conducting mobile application designing.

The simplicity of the features: The target group of the application users is mentally complex-minded individuals therefore when designing the features of the application, they have to be simple as possible as well as features have to not make the individuals depressed or anxious. Moreover, the login steps, devices connecting, heart rate measuring, and certain using guidelines have been provided in a simple manner.

Health summary exhibit in a relaxed manner: As mentioned in the previous point the users of the "MindGuardian" is special with that when we display the mental health summary we had to focus on not use straightforward verbal statements to display summary since the mental health summary also can be a depressing point for the individuals. Therefore, the design consists of very mild language.

Clean and user-friendly design: We have to be specially focused on the application screen designing. Manly the colors, elements even text. Because the users must be not making users stressed or depressed by using the application. Therefore, the Mobile application user interface design must be consisted of clean elements, simple text styles, and must have used relaxing and pleasing colors.

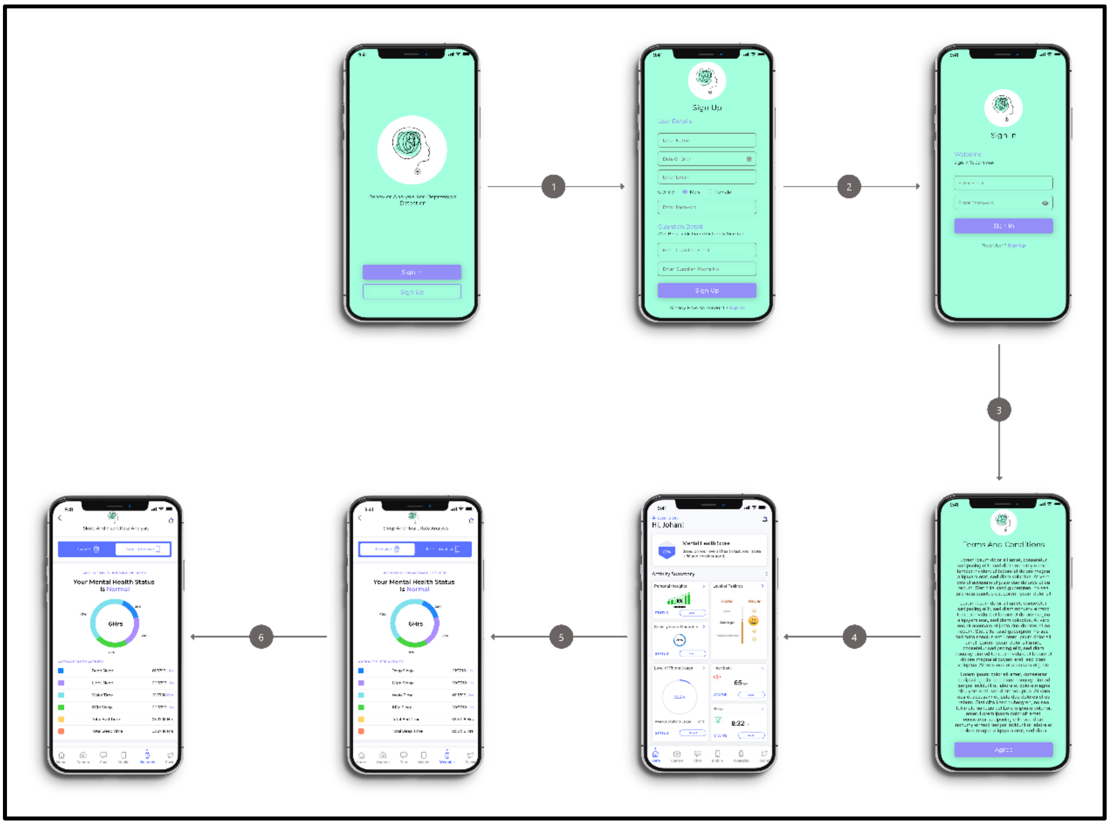


Figure 4. 2: User Flow of Invasive and Non – Invasive Analysis

As shown in figure (4.2) users can log in or register to the “Mind Guardian” mobile application. In the registration, basic user pieces of information will be asked and after submission, it will lead to a user agreement in which the user provides consent to access the data through the mobile application. in the same stage, the privacy policy agreement will be provided by the application. user can next connect the MI band with the “Mind Guardian” for invasive biometric data access. Furthermore, users can get a full summary of the activities from the home screen and an invasive noninvasive early depression risk analysis summary can be check by using the required navigations. In the Invasive and non – invasive tabs, users can view the summary of sleep details, heart rate details, and analysis of early depression risk and 14 days monitoring data. Users can provide heart rate daily by clicking the button in non – invasive tab.

4.4. Commercialization Aspects of the Product

When analyzing early depression risk, biomimetic parameter changes must be taken into account. Early identification of depression can reduce the risk of life-threatening consequences, but the social impact might lead to people refusing to take drugs.

Because the application is intended to tell the responsible guardian about the individual's mental health state, self-awareness applications can address individuals' worries and assist them to acquire appropriate and efficient medications. The proposed mobile application's economic value is listed below.

Benefits to the public

- An expanded review of behavioral changes without user contribution.
- Daily summary of sleep pattern and heart rate data.
- Automatic biweekly monitoring without user contribution.
- Easly install and can be used without prescription
- No privacy abusing.
- User can check their mental status without concerning social fear.
- Clinical phycologists can start immediate treatment soon as a two-week analysis.

Product Releasing.

- The MindGuardian mobile application can be downloaded from “play store” and installed on the mobile application.

4.5. Tools and Technologies

In this section, the tools and technologies used for model implementation and mobile app development will be discussed.

Let's look at the tools and technologies used for model implementation.

Table 4. 1: Tools and Technologies Used to Model Implementation.

Tools & Technology	Use
Environment→ Anaconda 3	Anaconda is an environment that is specified for python and R language management and deployment. Therefore, the anaconda is used as the environment for machine learning model development.
IDE → Jupyter Notebook	Jupyter Notebook is an open-source data science software that provides space to perfume scientific programs. This has been used as the IDE for the model implementation process.
Library → Scikit-Learn	Scikit-Learn is a free python library that provided a very of classification, regression, and clustering models. This library has been used for python classification model implementation.
Python Pickel	Python Pickel use or serialization and deserializations of the models. Therefore, this module has been used to obtain a model pickle.

API – Framework →Python fast API	Python fast API is a framework used for restful API development. Therefore, the framework has been used for model API development.
API Implementation environment → Spyder	Spyder is an open-source cross-platform environment. Spyder has been used for model API implantation.
API Deployment → Heroku	Heroku is a cloud platform that supports several languages. The model APIs are deployed in Heroku.

Let's look at the tools and technologies used for mobile application development implementation.

Table 4. 2: Tools and Technologies Used to Mobile Application Implementation.

Tools & Technology	Use
IDE→Android Studio	Android Studio is the official development platform for google android. Therefore, mobile development has been conducted on android studio.
Language →Flutter & Dart	Flutter is an open-source UI development language and dart is a client development language. These languages have been used for mobile application development.
Database → Firebase Real-Time	Firebase is a mobile and web application database service by google. The data storing in the application is completing by Firebase in Real-Time.

5. IMPLEMENTATION & TESTING.

5.1. Implementation.

5.1.1. Model Implementation

The "MindGuardian" Mobile Applications' main feature is biometric parameter analysis for early depression risk identification. Therefore, the development process consists of a machine learning analysis stage.

In this section, the machine learning procedure will be discussed with data collection, data pre-processing, model implementation, Model API implementation, and API deployment. Figure (5.1) indicates the process stages of the Model implementation.

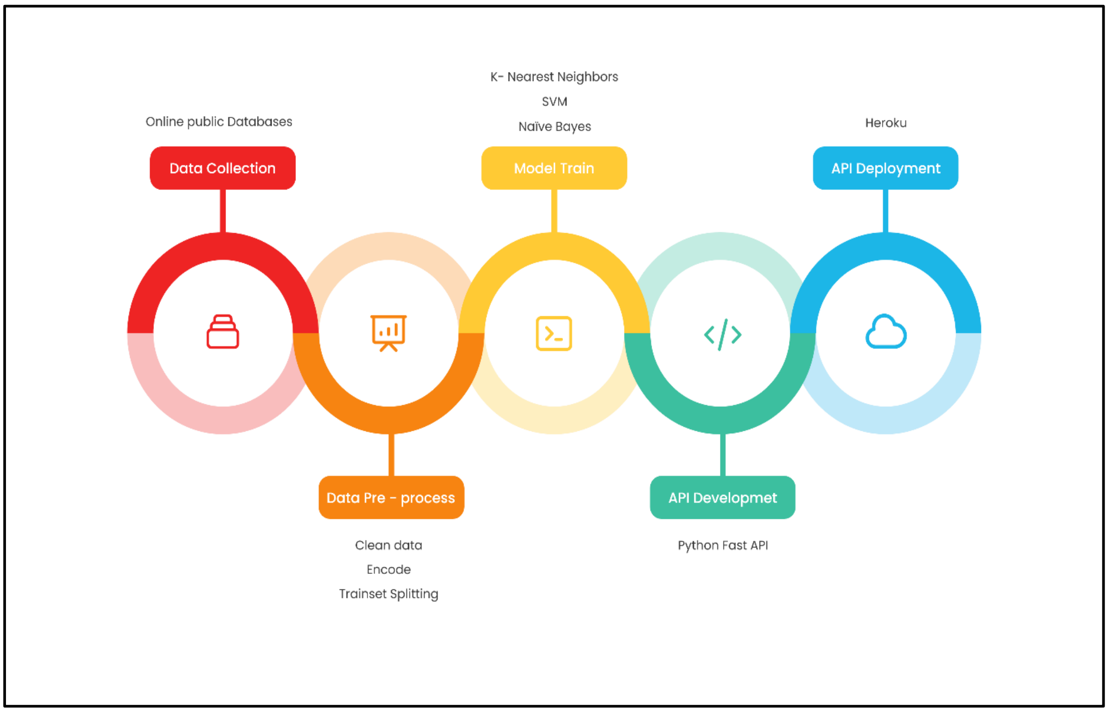


Figure 5. 1: Invasive and Non – Invasive Model Analysis Process.

Data Collection

As the initial step for early depression analysis with sleep patterns and heart rate model implementation data sets needed to be collected. Therefore, the best possible way of finding out a large number of datasets is public online datasets that are available for researchers. One of the common public datasets is researchers used is Kaggle but when it comes to data that goes with mental health is much confidential therefore, finding data from Kaggle was a fail attempt.

As the next attempt goes through the available dataset from researchers In their papers. In this approach, we have collected a dataset for heart rate and sleep patterns that are labeled as depressed and healthy individuals. Therefore, both dataset heart rate and sleep parameters have been collected from publicly available data.

The heart rate data was collected from the Health assessment of French university students, and risk factors associated with mental health disorders research by Harvard Medical School and University of Nice Sophia Antipolis students [19]. The total number of individual data that was in the dataset was 4184. Out of that amount they have labeled 483 individuals as depressed students. This data set has been used for the EDRAHR (Early Depression Risk Analysis Heart Rate).

The Sleep parameter data has been obtained from the research by Stockholm University, the University of Sao Paulo, and the University of Surrey [20]. The dataset has been a collection of light exposure data and sleep pattern data of individuals. From the dataset, we have obtained 71 depressive individuals' sleep data for the EDRASP (Early Depression Risk Analysis on Sleep Patterns).

Data Pre-Process

Data pre-processing is an important stage that needs to conduct prior to model train in order to prepare the dataset as a necessary condition to feed to the model. This phase includes clean data, encode data, and trainset splitting. Therefore, the Gathered data must be pre-processed before being fed into the model. This section will be an explanation of how the pre-processing phase has been conducted before EDRAHR and EDRASP.

Frist both datasets heart rate data and sleep pattern data have been gone through the data cleaning step which including add relevant headings, clean missing values. Clean missing values can be done in two ways, deleting the rows that have missing values, replacing the missing values with mean, median, or mode. clean the missing values of the dataset have been conducted by deleting the rows that have missing values. Figure (5.2) indicates before and after cleaning data of the heart rate data and sleep pattern data.

Sleep Data — Before

Out[3]:

	Gender	Age	SleepQualitycat	Sleepinesscat	Bedtimef	Uptimef	Totaltimef	Deepsleepf	REMSf	LightSleepf	Depression
0	1	33.000000	0	0	20.333334	30.0	9.666666	79.0	135.0	355.0	0
1	1	64.000000	0	0	20.333334	29.0	8.666666	90.0	112.0	297.0	0
2	1	39.000000	0	0	21.000000	29.0	8.000000	42.0	159.0	242.0	1
3	1	36.000000	0	0	20.500000	31.0	10.500000	51.0	228.0	327.0	1
4	1	22.000000	0	0	19.500000	29.0	9.500000	90.0	150.0	321.0	0
...	...	...	...	...	...	...	...	...	...	...	...
531	1	39.685146	0	0	24.000000	32.0	8.000000	61.0	138.0	247.0	0
532	1	32.142368	0	0	24.000000	33.5	9.500000	78.0	129.0	342.0	0
533	1	25.111567	0	0	22.000000	30.0	8.000000	47.0	206.0	201.0	1
534	0	22.691307	0	0	23.000000	34.0	11.000000	92.0	160.0	385.0	0
535	0	41.388092	0	0	23.500000	32.0	8.500000	68.0	127.0	271.0	0

536 rows × 11 columns

Sleep Data — After

Out[4]:

	Gender	Age	SleepQualitycat	Sleepinesscat	Bedtimef	Uptimef	Totaltimef	Deepsleepf	REMSf	LightSleepf	Depression
0	1	33.000000	0	0	20.333334	30.0	9.666666	79.0	135.0	355.0	0
1	1	64.000000	0	0	20.333334	29.0	8.666666	90.0	112.0	297.0	0
2	1	39.000000	0	0	21.000000	29.0	8.000000	42.0	159.0	242.0	1
3	1	36.000000	0	0	20.500000	31.0	10.500000	51.0	228.0	327.0	1
4	1	22.000000	0	0	19.500000	29.0	9.500000	90.0	150.0	321.0	0
...	...	...	...	...	...	...	...	...	...	...	...
531	1	39.685146	0	0	24.000000	32.0	8.000000	61.0	138.0	247.0	0
532	1	32.142368	0	0	24.000000	33.5	9.500000	78.0	129.0	342.0	0
533	1	25.111567	0	0	22.000000	30.0	8.000000	47.0	206.0	201.0	1
534	0	22.691307	0	0	23.000000	34.0	11.000000	92.0	160.0	385.0	0
535	0	41.388092	0	0	23.500000	32.0	8.500000	68.0	127.0	271.0	0

530 rows × 11 columns

Heart rate — Before

Out[8]:

	Age (4 levels)	Gender	Heart rate (bpm)	Abnormal heart rate	Depressive symptoms
0	20	0	75.0	0	0
1	19	0	65.0	0	0
2	20	0	65.0	0	1
3	19	0	70.0	0	0
4	18	0	NaN	NaN	0
...	...	...	...	...	...
4179	19	1	68.0	0	0
4180	18	1	80.0	0	0
4181	19	0	70.0	0	0
4182	20	1	81.0	0	0
4183	18	0	80.0	0	1

4184 rows × 5 columns

Heart rate — After

Out[9]:

	Age (4 levels)	Gender	Heart rate (bpm)	Abnormal heart rate	Depressive symptoms
0	20	0	75.0	0	0
1	19	0	65.0	0	0
2	20	0	65.0	0	1
3	19	0	70.0	0	0
8	18	1	72.0	0	1
...	...	...	...	...	...
4179	19	1	68.0	0	0
4180	18	1	80.0	0	0
4181	19	0	70.0	0	0
4182	20	1	81.0	0	0
4183	18	0	80.0	0	1

3798 rows × 5 columns

Figure 5. 2: Comparison of Before and After Cleaning Data.

Data encoding has been conducted after the Data cleaning step, in this step, datasets have been converted to float for necessary model training steps. As the last step of the pre-processing phase heart rate data and sleep pattern data sets went through the trainset splitting step which is dived the datasets for training and testing. The splitting was conducted with the common norm of the industry, the 8: 2 method which is 80% for training and 20% for testing. Note that the above mention pre-processing steps have been performed in Jupiter notebook with the use of "scikit-learn" libraries. Figure (5.3) exhibits the encoding and train test split-step of heart rate data, sleep pattern data.

```
In [16]: data = preprocessing.StandardScaler().fit(data).transform(data.astype(float))
         data[0:1]

Out[16]: array([[ 1.31158263, -0.85319077,  0.11971494, -0.28590749]])
```

Train Test Split

```
In [6]: from sklearn.model_selection import train_test_split

        train_data,test_data,train_target,test_target=train_test_split(data,target,test_size=0.2)
```

Figure 5. 3: Data Encoding and Trainset splitting.

Model Train

In this section, the model training process will be explained,

The biometric data analysis contains EDRAHR and EDRASP Both analyses have been conducted with supervised learning algorithms. Supervised learning algorithms have been used according to the labeled dataset that we have obtained. EDRAHR and EDRASP have been conducted with three supervised learning algorithms which are K– Nearest Neighbors (KNN), Support vector machine (SVM), and Naïve Bayes.

Prior to fit the dataset to models, as common steps import necessary libraries, import the dataset to the notebook, Standardize the Variables and define x with features, define y with the label which is “Depressive symptoms” have conducted.

Let's looking at the model trained of EDRAHR, the heart rate data set has been trained with the KNN model with the identified best K value. When finding the best K value, the value is marked as 5. The SVM algorithm with heart rate dataset trained with a kernel with degree 1. Next, the same dataset has trained with Naïve Bayes to GaussianNB classifier. Figure (5.4) exhibits the Heart Rate models.

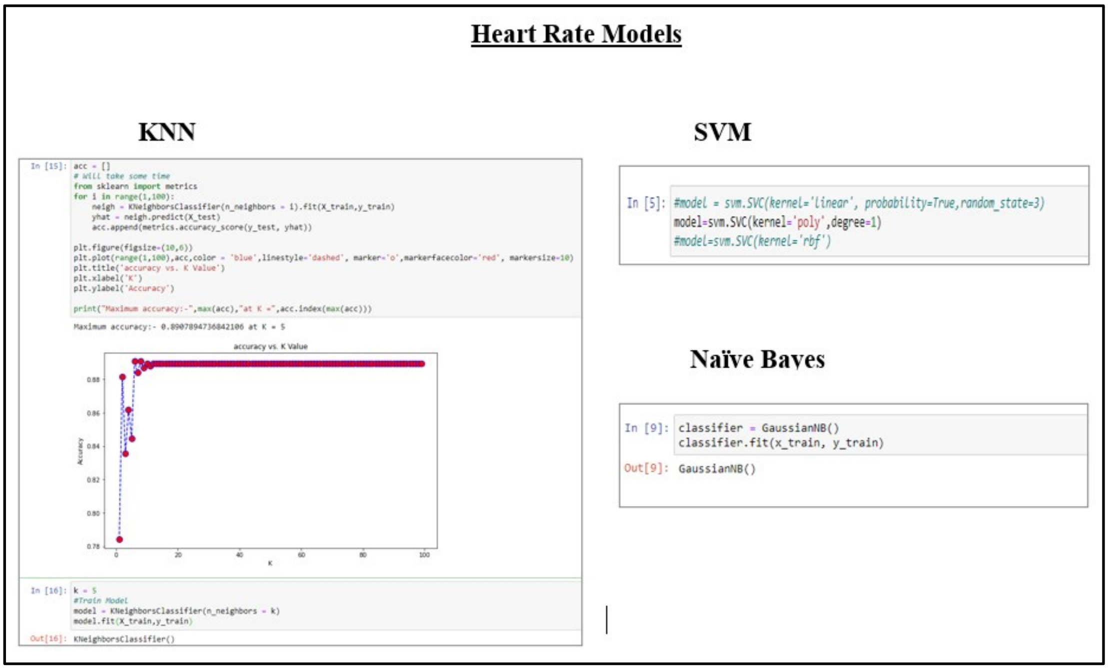


Figure 5. 4: Heart Rate Analysis Algorithms.

EDRASP has trained with the KNN algorithm for the best k value which is 5. After identifying the best k value, the KNeighbors Classifier fitted with the training dataset. Sleep pattern dataset also trained with SVM algorithm with kernel poly, degree 1. Moreover, the Naïve Bayes algorithm with GaussianNB trained for the same sleep pattern dataset. Figure (5.5) exhibits the sleep pattern models.

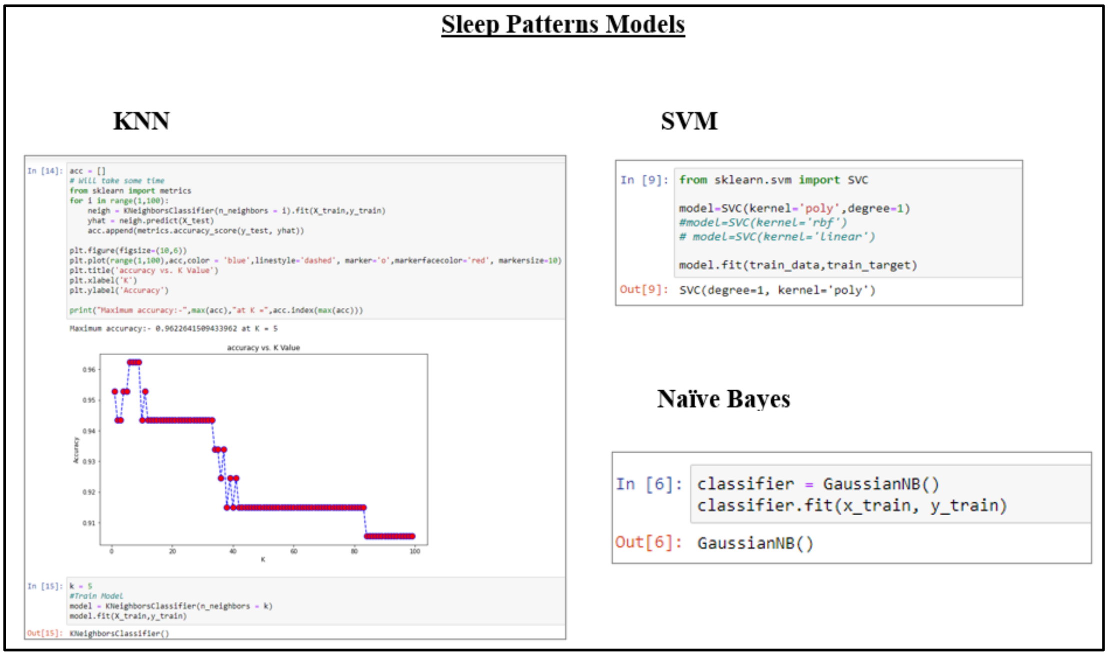


Figure 5. 5: Sleep Pattern Analysis Algorithms.

After all, models have been trained in the confusion matrix, ROC curves plotted, and dump pickle files for the purposes of further discussions of the models. The libraries that are used for these steps are “sklearn.metrics” and “pickle”.

All the above models have tiered on Jupiter notebook with the use of “scikit-learn” libraries.

API Implementation and Deployment.

API implementation has been conducted for model integration to the mobile application. Therefore, in this section API implementation process will be discussed.

Frist the best accuracy and fitted model will be chosen for integration to the mobile application. Selecting the best mode will be discussed in future sections (6). From EDRAHR models KNN has used for API and from EDRASP models Naïve Bayes have been chosen. The pickle file of the chosen model was used to write the API. In the API, parameters that need to be passed are defined to get a response. Python fast API has been used for development and Spyder is used as IDE.

After the Development of the APIs they have deployed in Heroku, APIs are deployed for the purpose of access from an external source. the APIs can be tested using postman or Swagger. The testing has been explained in section (5.2.3). Figure (5.6) shows the EDRASP and EDRAHR model APIs deployment.

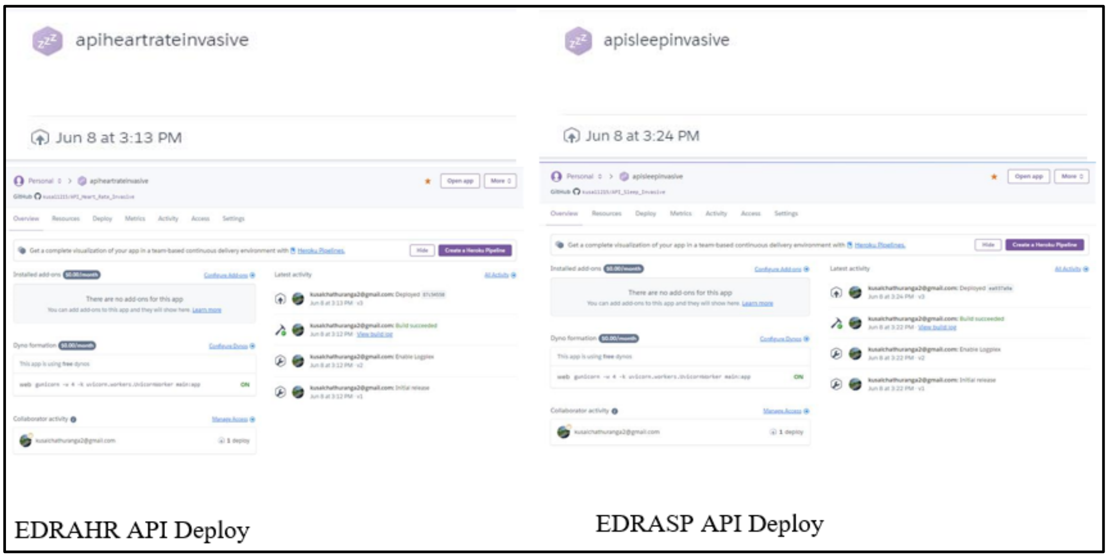


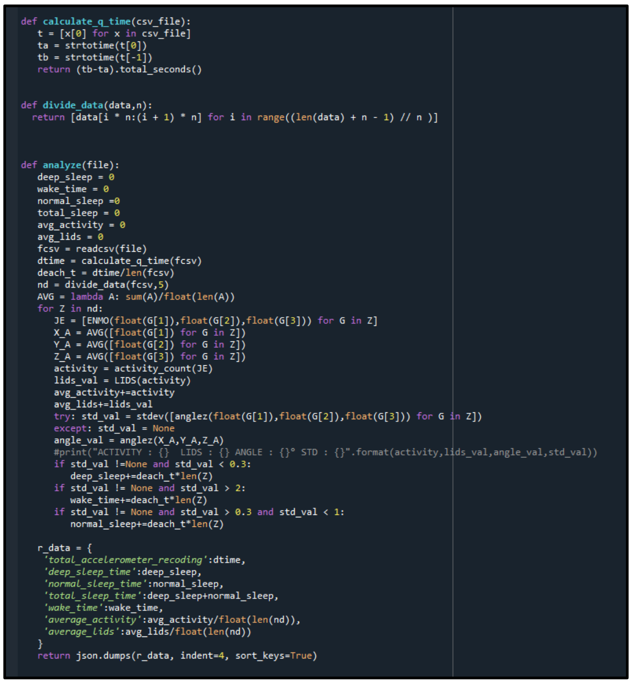
Figure 5. 6: Heart Rate and Sleep Model API Deployment.

5.1.2. Accelerometer Data to Sleep Data convert Method Implementation.

On non – invasive user data obtaining method sleep parameter data collecting the use of the accelerometer. In this section, the method of accelerometer data conversion to sleep parameter data will be explained.

At night the mobile application has developed to acquire accelerometer data which comes in the formant of X, Y, Z, and time stamp. These data need to be converted to sleep data (total sleep, Deep Sleep, Light Sleep hours) for the EDRASP. Therefore, the best possible process for accomplishing this task was to implement a python program. As our requirement of obtaining sleep parameter data from accelerometer data, a python program has been developed and implement an API to access the program within the "MindGuardian" mobile Application. The python program is a modification of the Package Called "SleepPy" [21]. Following Figure (5.7) indicates the implementation of the program.

```python
from math import *
import csv
from datetime import datetime
from  statistics import stdev
import json

'''
ENMO : The Eucliden Norm Minus One
its to correlate with magnitude of acceleration and
human energy

'''
def ENMO(x,y,z):
  return max(0,sqrt((x*x)+(y*y)+(z*z))-1)



def activity_count(ENMO_VAL):
  fx = lambda x : max(0,x-0.002)
  return sum([fx(x) for x in ENMO_VAL])


'''
LIDS : Locomoter inactivity During Sleep
involves non linear conversion of locomoter activity
'''
def LIDS(activity_c):
  return 100.0/(activity_c+1)



'''
Angle related to Z axis
'''
def anglez(x,y,z):
  return atan(z/sqrt((x*x)+(y*y)))*(180/pi)



def readcsv(file):
  data = []
  cr = csv.reader(open(file,'r'),delimiter=',')
  for x in cr:
    if len(x) == 4:
      try: data.append(x)
      except: pass
  return data


def strtotime(text):
  return datetime.strptime(text,"%Y-%m-%d %H:%M:%S.%f")


def calculate_q_time(csv_file):
   t = [x[0] for x in csv_file]
   ta = strtotime(t[0])
   tb = strtotime(t[-1])
```



```python
def calculate_q_time(csv_file):
   t = [x[0] for x in csv_file]
   ta = strtotime(t[0])
   tb = strtotime(t[-1])
   return (tb-ta).total_seconds()


def divide_data(data,n):
  return [data[i * n:(i + 1) * n] for i in range((len(data) + n - 1) // n )]



def analyze(file):
   deep_sleep = 0
   wake_time = 0
   normal_sleep =0
   total_sleep = 0
   avg_activity = 0
   avg_lids = 0
   fcsv = readcsv(file)
   dtime = calculate_q_time(fcsv)
   deach_t = dtime/len(fcsv)
   nd = divide_data(fcsv,5)
   AVG = lambda A: sum(A)/float(len(A))
   for Z in nd:
      JE = [ENMO(float(G[1]),float(G[2]),float(G[3])) for G in Z]
      X_A = AVG([float(G[1]) for G in Z])
      Y_A = AVG([float(G[2]) for G in Z])
      Z_A = AVG([float(G[3]) for G in Z])
      activity = activity_count(JE)
      lids_val = LIDS(activity)
      avg_activity+=activity
      avg_lids+=lids_val
      try: std_val = stdev([anglez(float(G[1]),float(G[2]),float(G[3])) for G in Z])
      except: std_val = None
      angle_val = anglez(X_A,Y_A,Z_A)
      #print("ACTIVITY : {}  LIDS : {} ANGLE : {}° STD : {}".format(activity,lids_val,angle_val,std_val))
      if std_val !=None and std_val < 0.3:
         deep_sleep+=deach_t*len(Z)
      if std_val != None and std_val > 2:
         wake_time+=deach_t*len(Z)
      if std_val != None and std_val > 0.3 and std_val < 1:
         normal_sleep+=deach_t*len(Z)

   r_data = {
    'total_accelerometer_recoding':dtime,
    'deep_sleep_time':deep_sleep,
    'normal_sleep_time':normal_sleep,
    'total_sleep_time':deep_sleep+normal_sleep,
    'wake_time':wake_time,
    'average_activity':avg_activity/float(len(nd)),
    'average_lids':avg_lids/float(len(nd))
   }
   return json.dumps(r_data, indent=4, sort_keys=True)
```

Figure 5. 7: Accelerometer data to Sleep Data Method Implementation

The developed program needs to be accessed from a mobile application therefore, API has been developed. The API is implemented using python fast API and deployed in Heroku. An accelerometer data CSV file should be pass to the API to obtain sleep parameter data. Following Figure (5.8) indicated the API implementation and Figure (5.9) indicates the Deployment.

```python
@app.route('/analyze', methods = ['POST'])
def success():
    if request.method == 'POST':
      f = request.files['file']
      if f.filename.split('.')[-1] in ['CSV','csv']:
        f.save(f.filename)
        return html_r.format(json2html.convert(json = analyze(f.filename)))
      else:
        return 'invalid file format ! '


@app.route('/upload', methods = ['POST'])
def successapi():
    if request.method == 'POST':
      f = request.files['file']
      if f.filename.split('.')[-1] in ['CSV','csv']:
        f.save(f.filename)
        return analyze(f.filename)
      else:
        return "{'error':'invalid file format !'}"

'''
if __name__ == '__main__':
    app.run(debug = True,port=3004)
'''
```

Figure 5. 8: Accelerometer data to Sleep Data API Implementation.

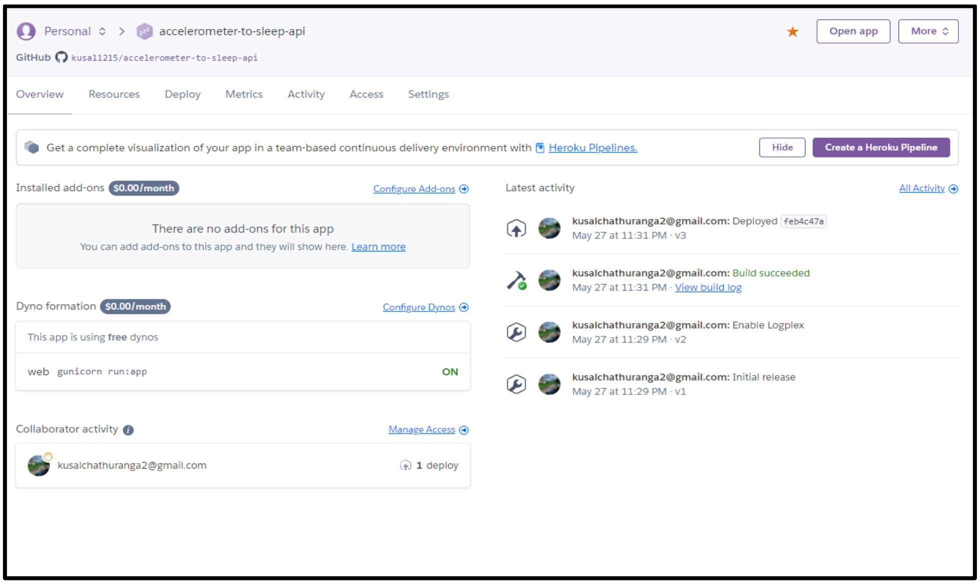


Figure 5. 9: Accelerometer data to Sleep Data API Deployment.

5.1.3. Mobile Application Development

In this section invasive user data obtain methods development, non - invasively user data obtained method development, Model Integration, UI design, and database integration will be explained.

Mobile development has been developed section-wise. First, let's looking at the invasive user data obtain development.

Invasively user data obtain methods development

The method selected to collect biometrics data from the user is via a digital watch. The data that need for the EDRASP are total sleep, deep sleep, light sleep, Rapid Eye moment sleep (REM) hours. For EDRAHR heart rate. To acquire these data, we have selected MI band 5. The process integrates the MI band 5 with a mobile application include Google Fit Authorization and Flutter FIT KIT Package integration. Google Fit Authorization process consists of Create a project in the Google cloud platform, Install the google fit kit library to the created project and Create a Client id for android in the project []. The next step was fit kit package integration and backend developments. Even the fit kit package is a full packet to integrate developing an app but yet while development issues have occurred. Figure (5.10) shows the steps of the Google Fit Authorization process and figure (5.11) shows fit kit package integration and MI Band.

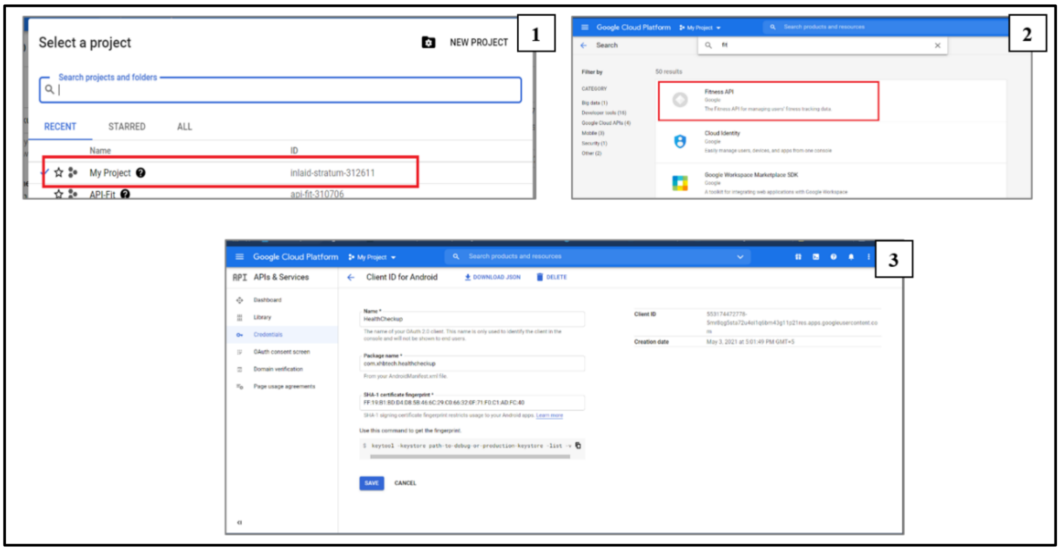


Figure 5. 10: Google Fit Authorization process

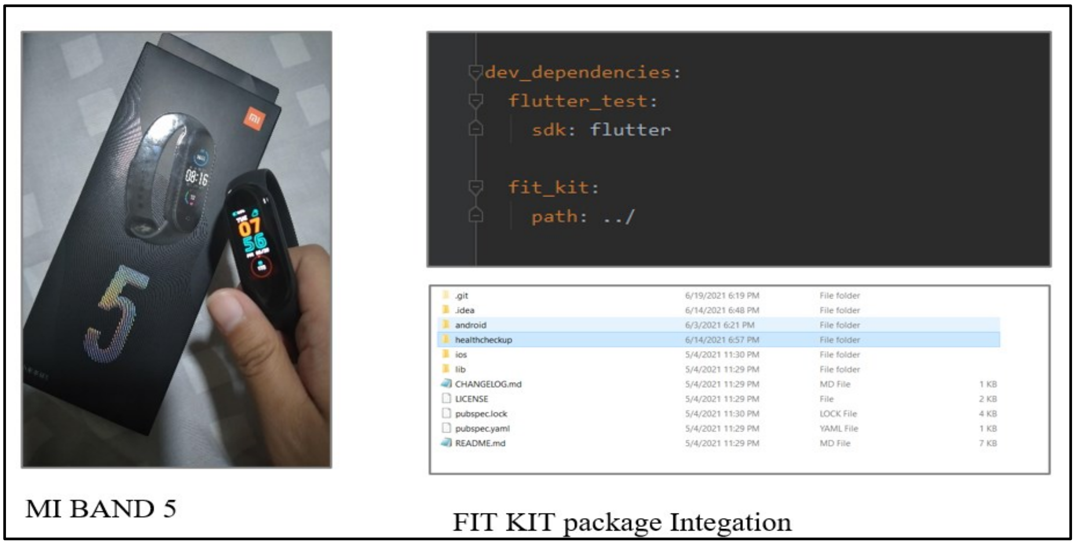


Figure 5. 11: fit kit package integration and MI Band

After integration of the package, the biometric data can be obtained from the user, but yet there is a final step and need to be performed which is Clean accessed Data and Calculation. The Data retrieving from the package is not readable, therefore, proceed with some steps to make them Readable. The Unreadable data have been categorized to store in firebase and Unusable data will be Ignored. The categories are Heart rate, Deep sleep, REM sleep, Light Sleep, Total Sleep. The mobile application development has been performed in Android studio using flutter and, dart languages. The before and after states of data retrieving are indicating in Figure (5.12).

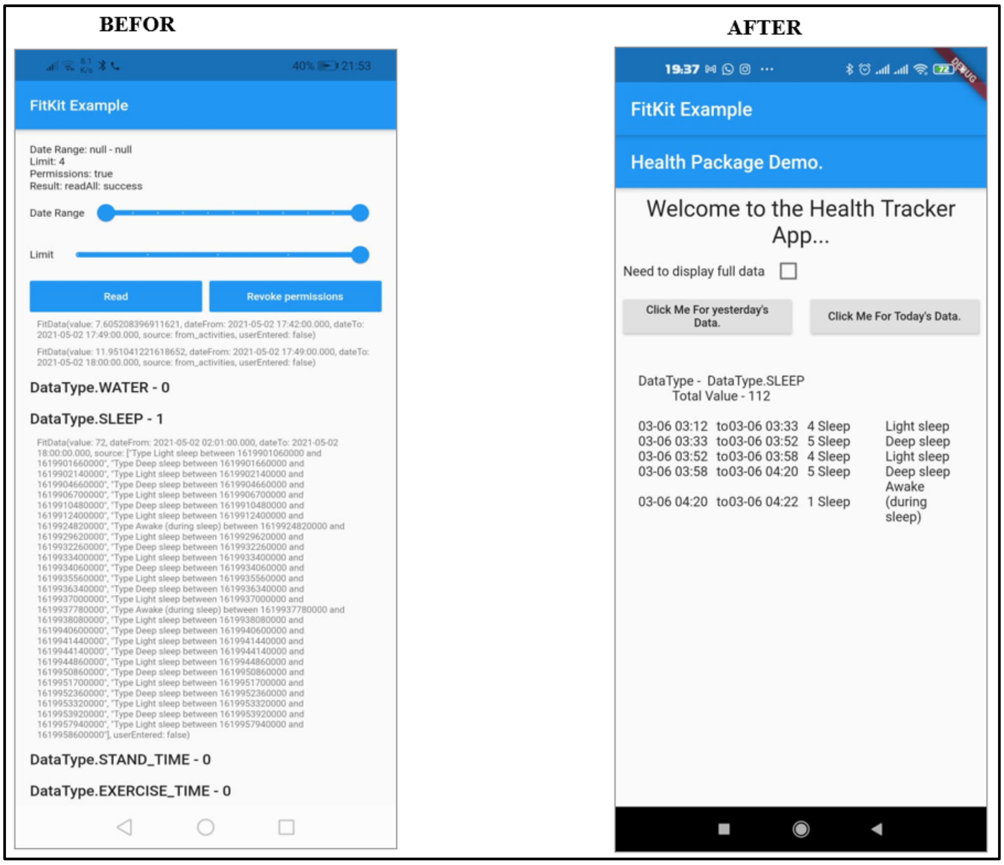


Figure 5. 12: Before and After States of Data Retrieving

Non - Invasively user data obtain methods development

User's biometric data have been collected using both invasively and non-invasively. The invasive method is used to obtain data was a digital watch (explained in section (5.1.2)) and non-invasively obtained sleep data and heart rate proceed as two developments.

Non - invasively collection of sleep data: this part contains the collect mobile accelerometer sensor data by a mobile application integrated plugin and data are converting into sleep parameters using Implemented Python Program. The method of accelerometer data to sleep parameter data has been discussed in section (5.1.2). Therefore, in this section, how the whole process works will be explained.

The method of collected accelerometer data is by a flutter plugin designed to collect raw sensor data from a mobile application. Then the collected accelerometer data have put into a CSV file in the mobile application process. As the last step, the CSV file has passed to the "Mobile Accelerometer Sensor Data to Sleep Parameters API" to get sleep parameter data as a response. The API provides total sleep, Deep Sleep, Light Sleep hours of the user. There is a certain way that user should keep their mobile phone with them in their sleep to mobile application to collect reliable Accelerometer sensor data. That is to keep the app running in the background, not device locked, and placing the phone is important. Figure (5.13) shows how mobile phones place on the bed at night. Figure (5.14) indicates the flutter plugin integration to collect Accelerometer sensor data.

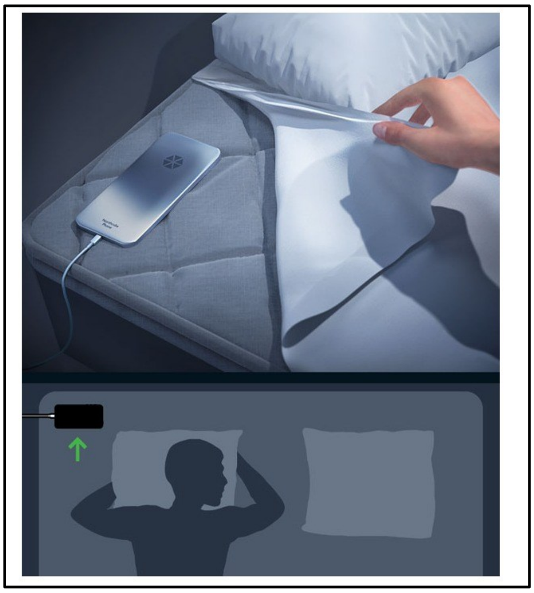


Figure 5. 13: How Mobile Phones Place on the Bed

```yaml
# pubspec.yaml
name: rawsensors
description: A flutter plugin that lets you access raw data from sensors, the goal of this plugin is to expose every sensors from the And
version: 0.0.3
homepage: https://github.com/TrAyZeN/rawsensors

environment:
  sdk: ">=2.1.0 <3.0.0"
  flutter: ^1.10.0
```

Figure 5. 14: Flutter Plugin Integration

Non - invasively collection of heart rate data: The heart rate collected from users non-invasively done by flutter dependency. The dependency has been integrated into the mobile application with the necessary resources made the method of the user to provide heart rate for the analysis. There is a technique that users need to follow in order to provide data to the mobile application which is User Need to cover Both camera and Flash With a finger. The method indicates in figure (5.16) and figure (5.15) indicates the integration of the dependency.

```
environment:
  sdk: ">=2.1.0 <3.0.0"

dependencies:
  flutter:
    sdk: flutter
  cupertino_icons: ^0.1.2
  charts_flutter: ^0.9.0
  wakelock: ^0.1.4+1
  camera: 0.7.0+4

dev_dependencies:
  flutter_test:
    sdk: flutter
  flutter_launcher_icons: ^0.7.3
```

Figure 5. 15: Camera Integration Of The Dependency

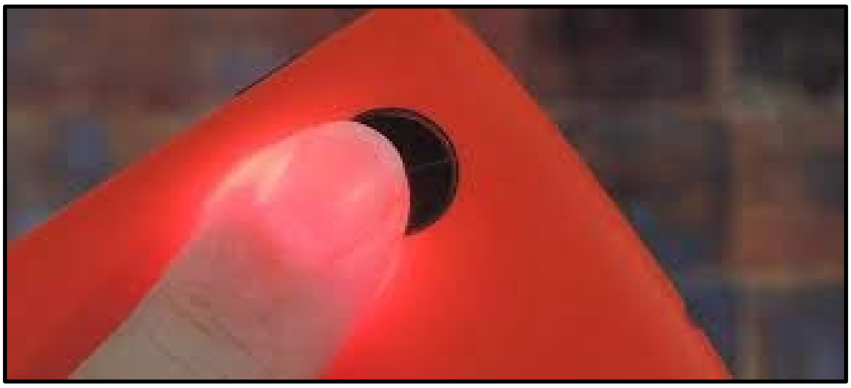


Figure 5. 16: How Finger Should Place to Camera and Flasher.

There is a calculation of getting the heart rate because, at one time the sensor provide several heart rates therefore, the calculation used was to get the average heart rate of the several heart rates. Figure (5.17) shows how the Heart rate can be collected from the mobile app.

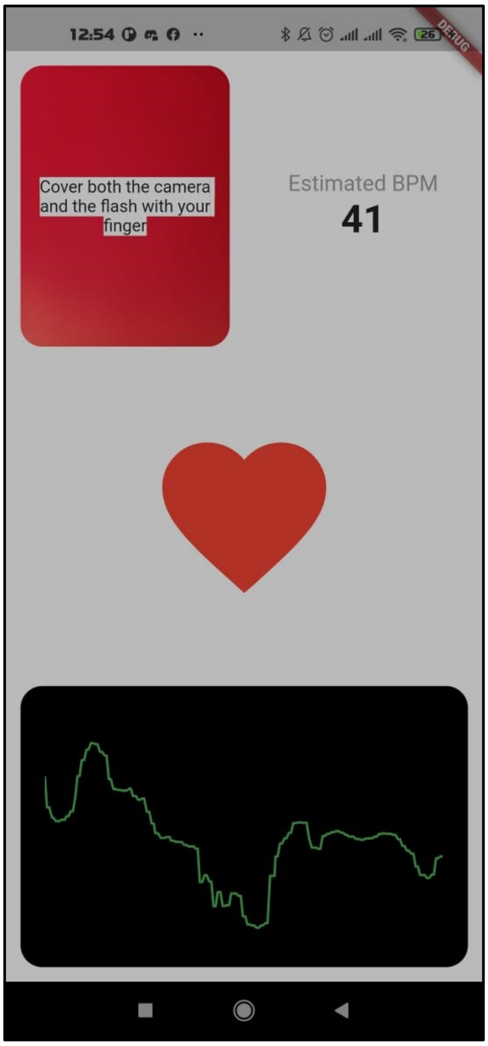


Figure 5. 17: The Heart Collecting Method.

Store Data and Database Integration.

After Obtaining the data from either method invasively or non - invasively the data will be calculated and organized then stored in the database for analysis purposes. The calculation code segments are indicated in Figure (5.18).

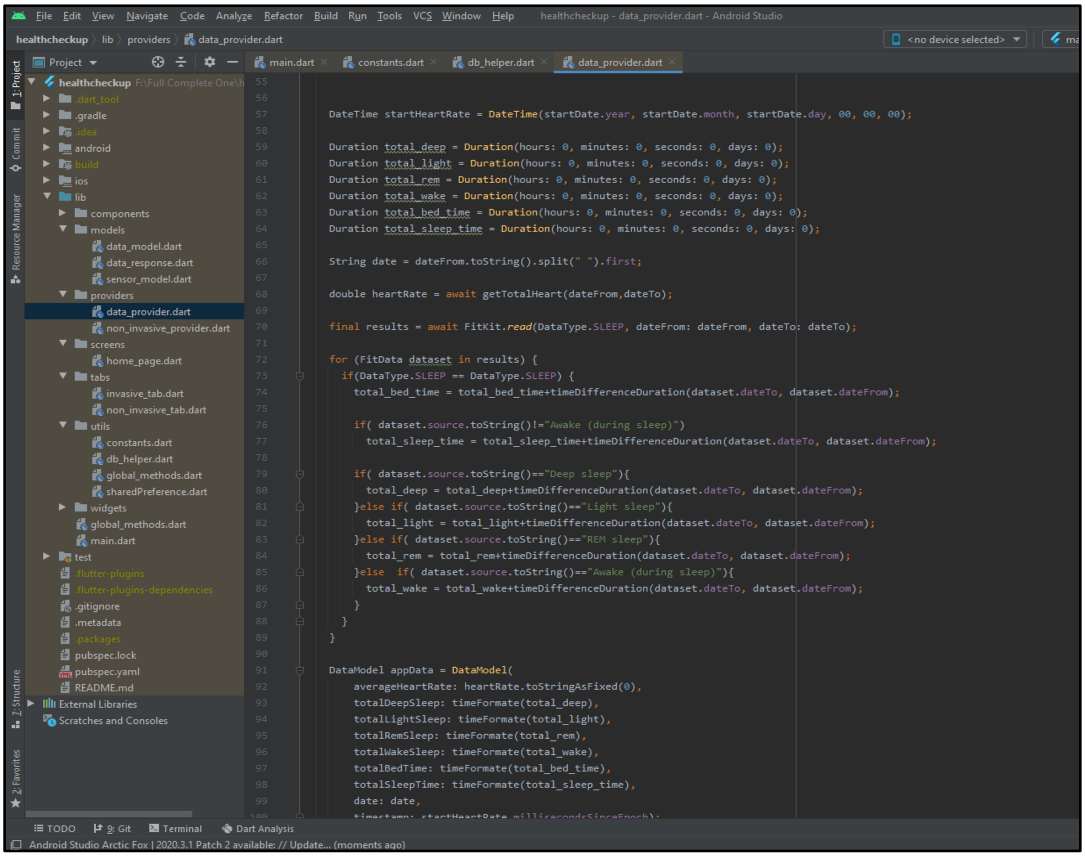


Figure 5. 18: Calculation Code Segments

The used database for the mobile application was firebase real-time database and there were some steps to follow in order to connect firebase to the application. The steps have been followed according to the Installation & Setup on Android by firebase [22].

The analysis has gone through 14 days therefore, some rules have been added to the firebase. The rule is to eliminate the 15th-day data after 14 days have been completed. Figure (5.19) indicates the rules that have been added to firebase. Figure (5.21) database integration to the application and Figure (5.20) stored data example in DB.

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

Figure 5. 20: Firebase Rules

Figure 5. 19: Stored Data In Firebase.

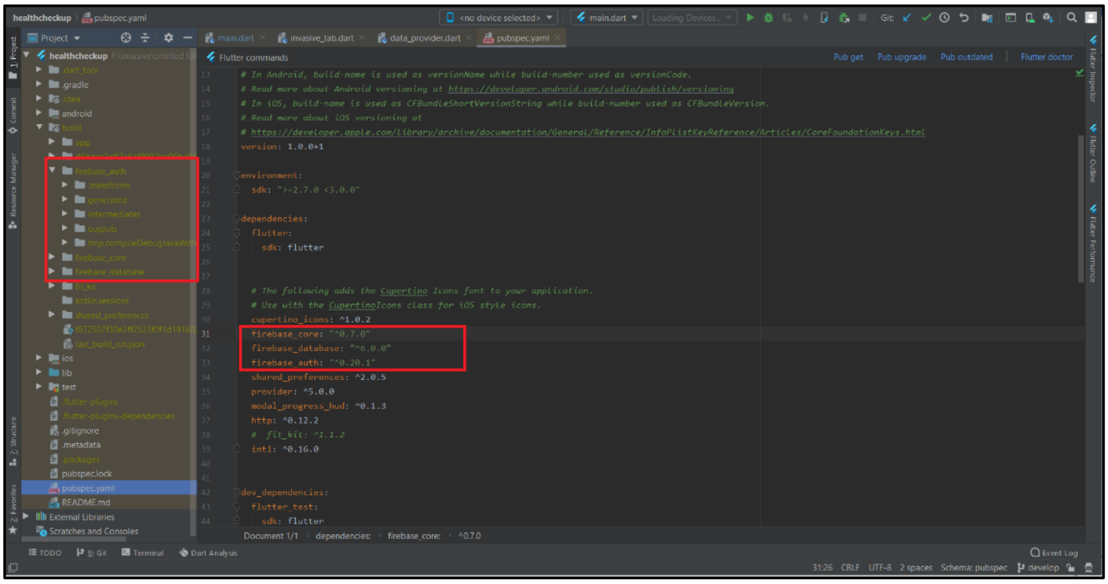


Figure 5. 21: Firebase Integration.

Model and Accelerometer API integration to the Mobile Application.

After data have been stored in the database, by accessing the data predictions have been conducted. Moreover, the accelerometer data convert to sleep data API have integrated to the mobile application. APIs defined at global methods for using them in respectively "getPredictionHeart Rate", "getsleepPrediction", and "getsleepdatafromscv". Figure (5.22) exhibits API initiation and Figure (5.23) shows API integrations.

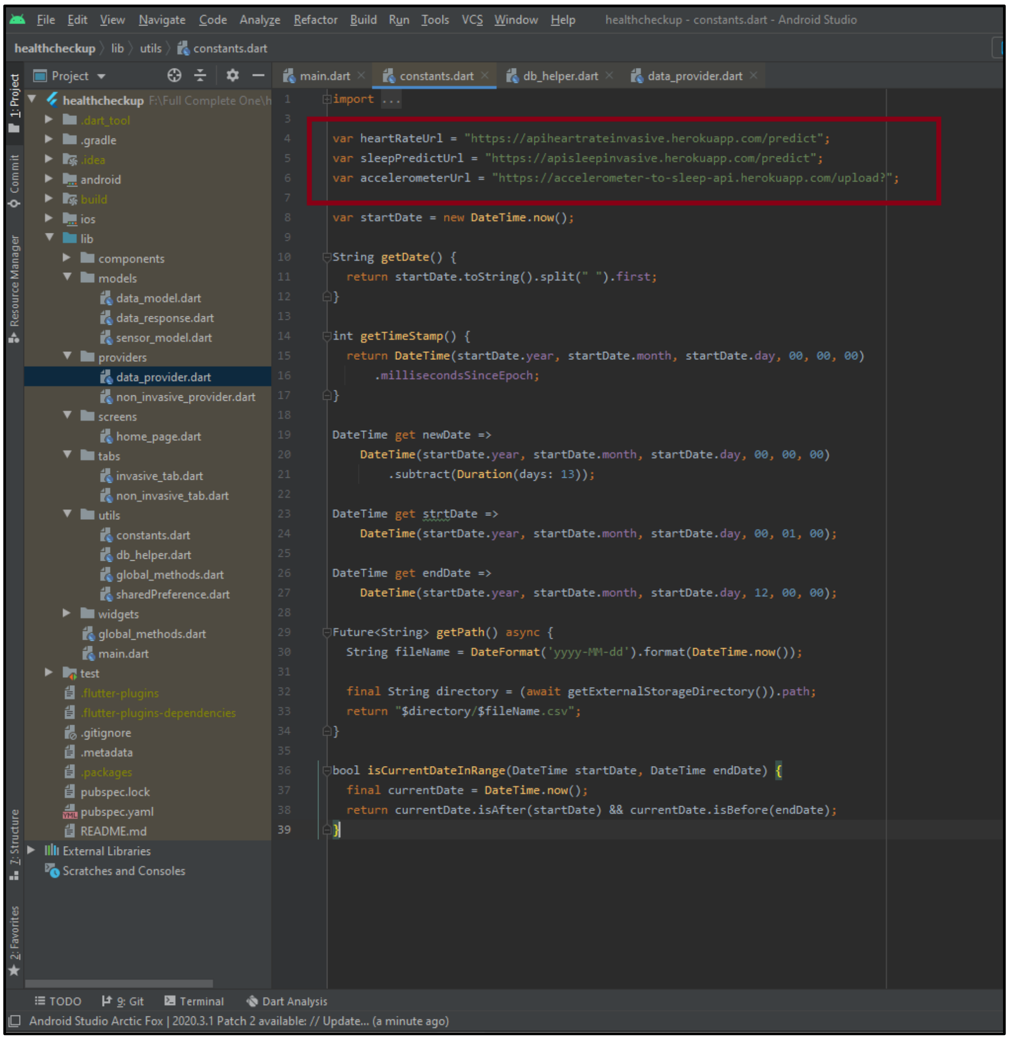


Figure 5. 22: API Initiation

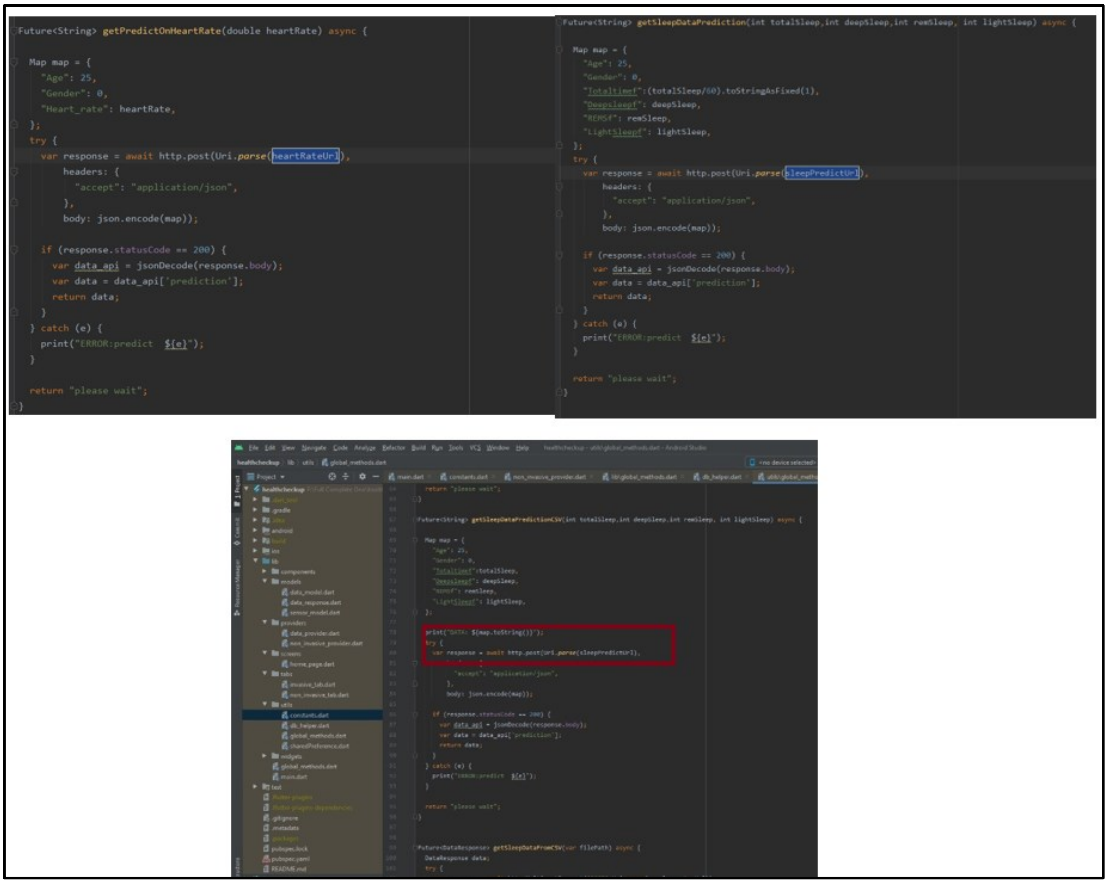


Figure 5. 23: API Integrations

After obtaining the predictions from model APIs the data will be pass to the frontend. Moreover, Accelerometer to sleep data API provides sleep parameters those data will be used for predictions and necessary calculations. The end of the calculations and organizing the data passing to the frontend have been completed.

5.2. Testing

The testing phase was the final stage in all of the implementations. In this section EDRASP and EDRAHR model API testing, Acceleometr data to sleep data API testing, mobile development front end testing, mobile development back end testing will be explained.

5.2.1. EDRSP and EDRHR Model Testing

Models have been examined to identify the best models for EDRAHR and EDRASP. After that using the choose models which are the KNN model for EDRAHR and Naïve Bayes model for EDRASP. The discussion of selecting the best model is discussed in section (6.2).

Data from two people from each test dataset have been selected for model evaluation. Test the High Risk of Depression (HRD) and the Low Risk of Depression (LRD) of the EDRAHR, EDRASP Test cases are following,

Table 5. 1: Model Test Cases

Test case	Scenario	Tested data	Expected Output	Actual Output	Satus
T001	Verify whether the EDRASP model Naïve Bayes identifies a High Risk of Depression person.	Gender: Male Age: 25 Total Sleep: 8.0 Deep Sleep: 42 REM Sleep: 159 Light Sleep: 242	1	HRD (1)	Pass
T002	Verify whether the EDRASP Naïve Bayes model identifies a High not Risk of Depression person.	Gender: Male Age: 25 Total Sleep: 6.5 Deep Sleep: 78 REM Sleep: 90	0	LRD (0)	Pass

		Light Sleep: 201			
T003	Verify whether the EDRAHR KNN model identifies a High Risk of Depression person.	Gender: Female Age: 18 Heart rate: 60	1	HRD (1)	Pass
T004	Verify whether the EDRAHR KNN model identifies a High not Risk of Depression person.	Gender: Female Age: 19 Heart rate: 80	0	LRD (0)	Pass

According to the EDRASP and EDRAHR model testing table the person with Total Sleep of 8h, Deep Sleep of 42min, REM Sleep with 159min And Light Sleep with 242h tend to have HRD, moreover, the person with a Heart rate of 60mbp shows HRD.

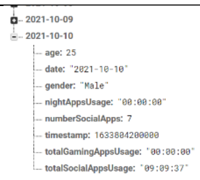
5.2.2. Mobile Application Testing

Mobile Application testing has been conducted to ensure the process and functionalities work as design as well as in the required manner. Mobile application testing has been conducted using test cases. Table (5.2) indicate test cases

Table 5. 2: Mobile Application Test Cases.

Test Case	**Test Scenario**	**Test Step**	**Test Data**	**Expected result.**	**Status**
T005	Verify whether the MI band 5	Connect watch the MI fit app.	Google account email password.	Display Sleep data and heart rate data that have	**Pass**

	connected with the mobile Application.	Connect MI fit with google fit account. Login to the Mind guardian app with the google fit connected email.		been collected.	
T006	Verify whether heart rate can be obtained from phone camera and flasher	Place the finger on the mobile camera and the flasher which in the way both cover.	----------	Display collected heart rate on the screen.	**pass**
T007	Verify whether accelerometer data can be obtained from the application.	Place the mobile phone beside the bed while the mobile application run background.	-----------	Check CSV file store in the mobile application Internal storage/ Android/ data/ com.xhbtech.healthcheckup folder.	**pass**

T008	Verify whether the invasively acquired heart rate display in the application.	Check MI band connected with Google fit.	Day:2021 – 10 - 05 Heart rate: 75bpm	Display heart rate on the invasive tab.	**Pass**
T009	Verify whether the invasively acquired sleep pattern data display in the application.	Check MI band connected with Google fit.	Day:2021 – 10 - 05 Deepsleep:00:46:00 LightSleep: 03:14:00 REMsleep: 00:14:00 Total sleep: 04:14:00	Display sleep pattern on the invasive tab.	**Pass**
T010	Check whether the collected data pass to the firebase.	Log into the firebase account using email and password.		Show date, timestamp. Total bedtime, total deep sleep, total REMsleep, total sleep time, total	**Pass**

				wake time, and heart on firebase.	
T011	Verify whether the mobile application plugins are integrated properly.	Add the plugins to the pubspec.yaml. Run “flutter pub get” on the console.	firebase_core: "^0.7.0" camera: 0.7.0+4 csv: ^4.1.0 sensors: ^2.0.3	Display 0 errors on the console.	**Pass**

5.2.3. API Testing

The EDRAHR and EDRASP analysis model API as well as the Accelerometer data to sleep data API has been created and released. The APIs are required to be tested as the final step in the implementation process.

The EDRASP analysis Naïve Bayes model API was tested using Postman. The data in the request body was submitted in JSON format, and the response was also received in JSON format. The Request's URL, Body, and Response are as follows:

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

Body: JSON

{
"Gender": 0,
"Age": 20,
"Totaltimef": 9,
"Deepsleepf": 20,
"REMSf": 60,

"LightSleepf": 220

}

Response:

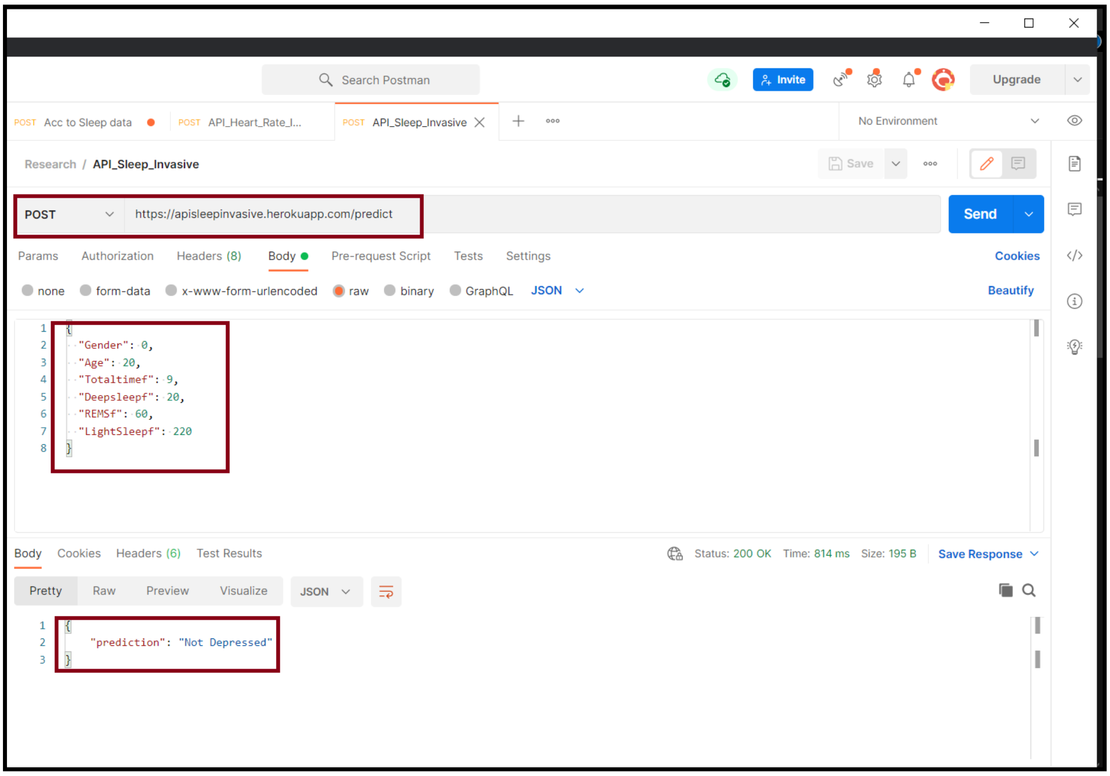


Figure 5. 24: Sleep API Testing.

The EDRAHR analysis KNN model API was tested using Postman. The data in the request body was submitted in JSON format, and the response was also received in JSON format. The Request's URL, Body, and Response are as follows:

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

Body: JSON

{

"Age": 20,

"Gender": 0,

"Heart_rate": 65

}

Response:

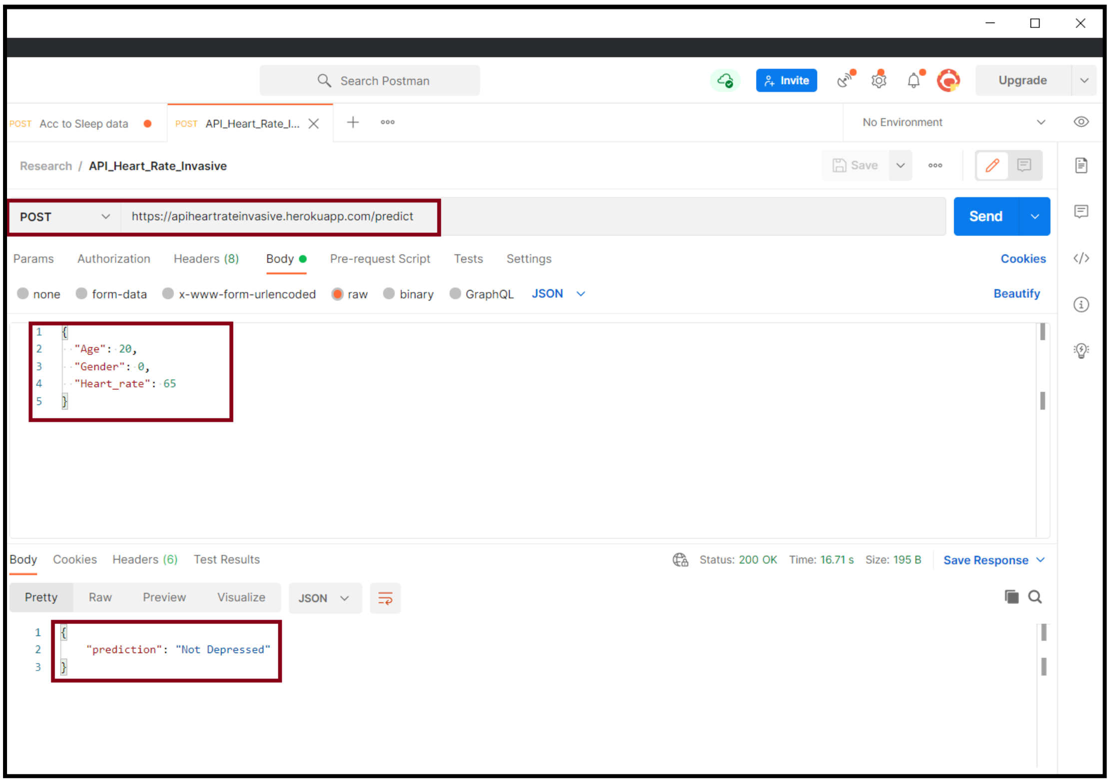


Figure 5. 25: Heart Rate API Testing

The Accelerometer data to sleep data conversion API has been tested using postman and the body contains the CSV file of nighttime mobile accelerometer data. The response was received in JSON format. The Request's URL, Body, and Response are as follows:

GET request: https://accelerometer-to-sleep-api.herokuapp.com/upload

Response:

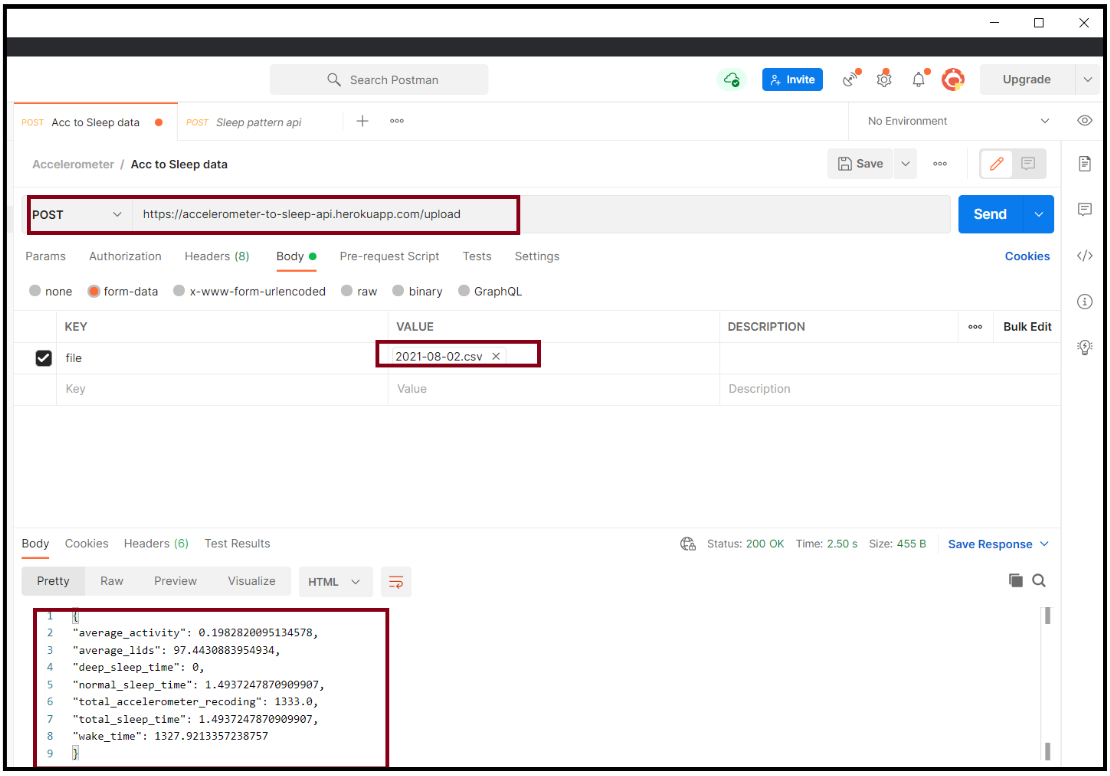


Figure 5. 26: Accelerometer Data To Sleep Data API Testing.

6. RESULTS & DISCUSSION

6.1. Results

The preferable models for EDRAHR and EDRASP analysis have been chosen according to the results of the algorithms that have attained the results of the models well the invasive and non-invasive data obtaining methods results have been explained in this section.

EDRAHR and EDRASP Model Results.

As described in Section 2.2.3, the EDRAHR was carried out using three algorithms. The maximum accuracy of the KNN algorithm trainset was 89%, while the minimum k was designated as k = 5 with a 10% misclassification rate. Get an AUC of 84%, a sensitivity of 50%, and a specificity of 94 % for the test data model. Meanwhile, with kernel poly at degree = 1, the Support Vector Machine (SVM) algorithm achieves 86% accuracy. The Naive Bayes algorithm produced a test accuracy of 86%, as well as an AUC of 86%, a sensitivity of 50%, and a specificity of 93%. The following Figure (6.1) indicates the Confusion Matrix of the KNN model.

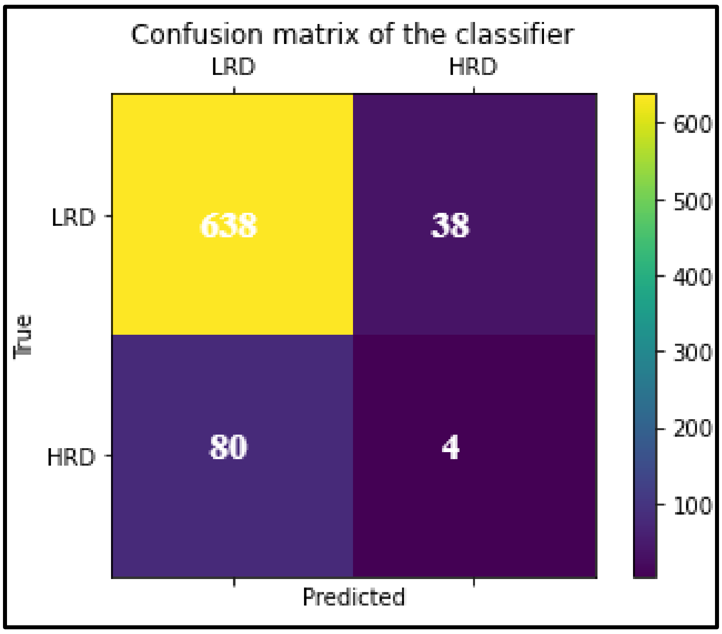


Figure 6. 1: Confusion Matrix of the KNN Model.

The EDRASP AUC has a KNN of 95%, a model sensitivity of 60%, and a specificity of 99%. The test dataset had a 95% accuracy, whereas the training dataset had a 96% accuracy at k = 5 with a 3.7% misclassification rate. When we looked at the SVM for the same sleep pattern dataset, we found that the model had 95% test accuracy, a 5% misclassification rate, an AUC of 85%, and model sensitivity and specificity of 65 and 100%, respectively. Furthermore, the Naive Bayes algorithm achieved 92% test accuracy, with an 80% sensitivity and 96% specificity. Note that the dataset split for all the models was the ratio of 80 training and 20 test sets. The following Figure (6.2) exhibits the ROC curves of the Naive Bayes model.

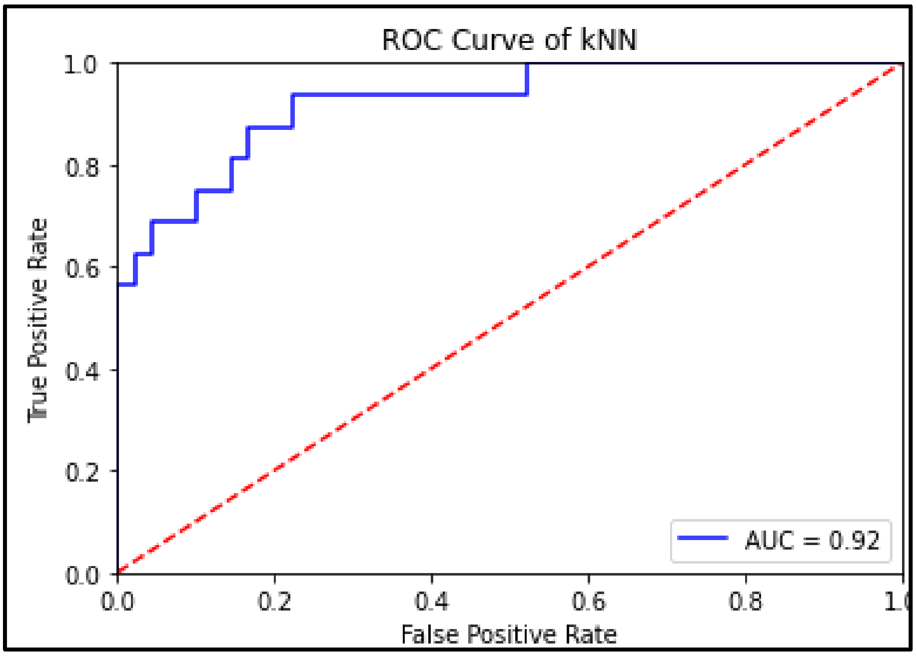


Figure 6. 2: ROC Curves of the Naive Bayes Model

Invasive and Non-invasive data obtaining method Results.

The accelerometer data is collected at night, and as a result of the data collection, a CSV file is saved in the phone's storage for later usage. The examples of collected data are shown in Figure (5.29).

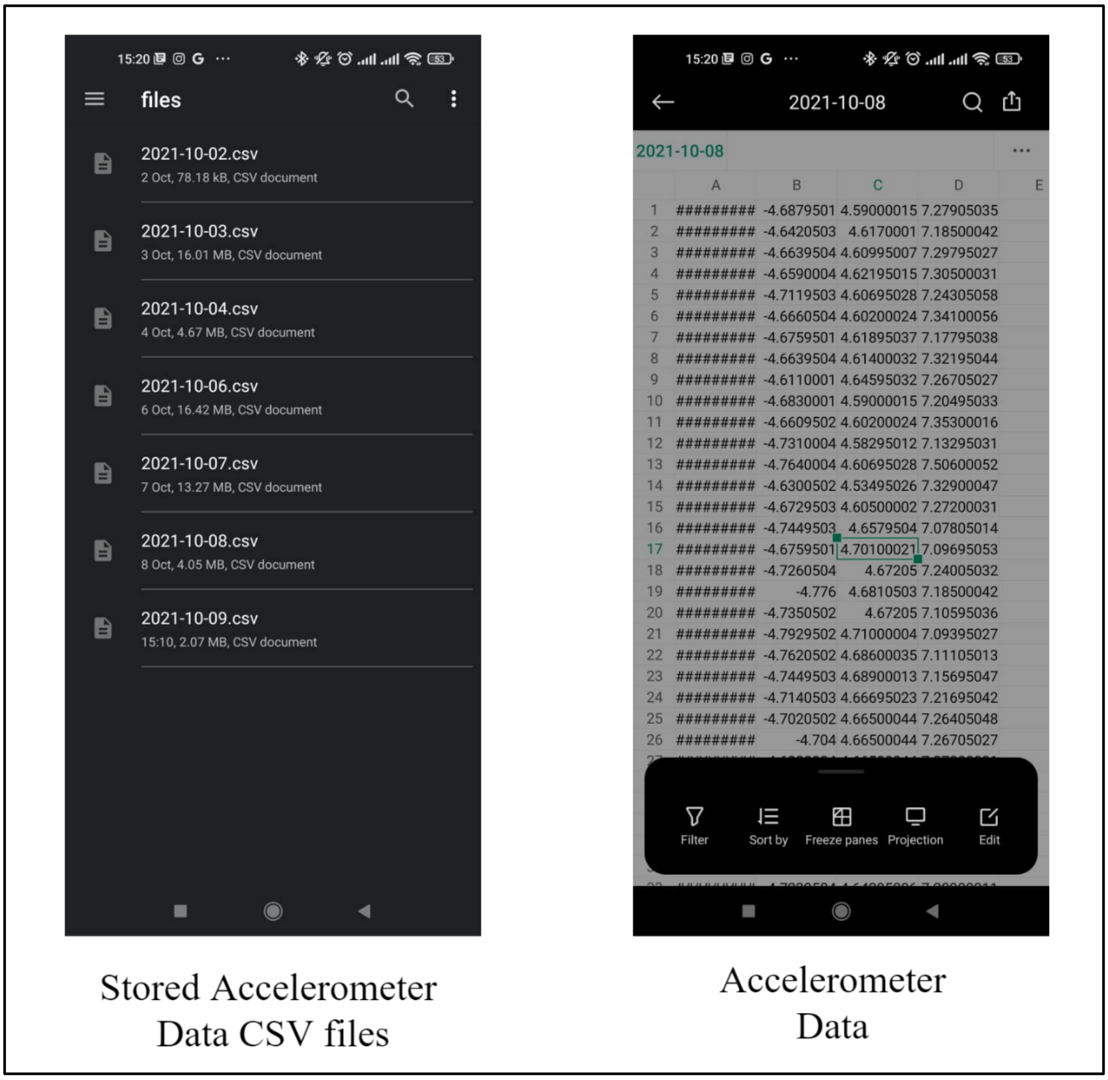


Figure 6. 3: Accelerometer Collected Data.

6.2. Discussion.

The outcomes of the EDRAHR and EDRASP trained algorithms have been examined in order to determine each algorithm's performance. The identification of algorithms led to the development of a preferred model for EDRAHR and EDRASP analyses. The following table () indicates the comparative performance evaluation of the above-explained algorithms.

Table 6. 1: Model Algorithm Performance Comparison.

	Algorithm	**Accuracy**	**PRECISION**	**RECALL**	**F1-SCORE**
EDRAHR	KNN	0.84	0.80	0.84	0.82
	SVM	0.86	0.76	0.87	0.82
	Naive Bayes	0.86	0.43	0.50	0.46
EDRASP	KNN	0.95	0.95	0.95	0.95
	SVM	0.96	0.96	0.96	0.96
	Naive Bayes	0.92	0.88	0.80	0.84

According to the Classification report findings, the SVM and Naive Bayes algorithms achieved significantly higher accuracy than the KNN, although their sensitivity percentages are significantly lower. When considering the total results of the Classification report, the KNN model has an 84 % test accuracy and 0.84 sensitivity, implying that it can correctly identify an individual with the disease in 84 % of cases. As a result, the KNN algorithm is the best option for EDRAHR.

As a result of the observations of EDRASP analysis algorithms, the KNN outperforms SVM and Naive Bayes with model accuracy. However, when all algorithms are compared in terms of sensitivity, Naive Bayes outperforms them all. This means the Naive Bayes algorithm has the most possibility of identifying the individual with depression risk as positive. As a result, the Naive Bayes model is the better option for EDRASP.

When comparing invasive versus non-invasive biometric data, invasive data is more reliable and accurate because it is collected using a digital watch that the user wears all the time. The heart rate data obtained from non – invasive and invasively shows a 5 to 6 Mbps difference since when collecting non – invasive data it obtain an average of 5 heart rates. The sleep data collected by the MI band 5 is extremely reliable, and there are always discrepancies between non-invasively and invasively collected sleep data. The reason for this is that when collecting accelerometer data using sensors, the mobile phone can move its position to where we want it to be, and the mobile application can stop running in the background after a while these reasons preventing data collection from the accelerometer sensor. Therefore, even the non – invasive method integrated for user convince we always influence users to use invasive method for more accurate analysis.

The difficulty to find depressed individuals was the challenging part of the project. Moreover, due to the pandemic situation finding a real dataset was a difficult task therefore, models' did not have tested on real depressed individuals this can be marked as the limitation of the project.

7. CONCLUSION

The aim of the project component was to come up with invasive and non-invasive collected biometric data analysis for early depression risk identification of the individuals. The machine learning supervised learning algorithms were used to create analytical models. The sleep pattern analysis (EDRSP) was carried out with 92% accuracy using the Naive Bayes method, while the heart rate analysis (EDRAHR) was carried out with 84% accuracy using the KNN model. The models have been extending to integrate to the mobile application for making an analysis prediction with invasively and non – invasively obtained data from the individuals.

REFERENCE

[1] WHO, "Depression," World Health Organization, [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression. [Accessed 16 February 2021]

[2] S. N. Haber and S. L. Rauch, "Neurocircuitry: A Window into the Networks Underlying Neuropsychiatric Disease," Neuropsychopharmacology, vol. 35, no. 1, pp. 1-3, 2009.

[3] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh and H. Ohsaki, "Recognizing Depression from Twitter Activity," Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 2015.

[4] 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.

[5] M. J. Kalucy, R. Grunstein, T. Lambert and N. Glozier, "Obstructive sleep apnoea and schizophrenia – A research agenda," Sleep Medicine Reviews, vol. 17, no. 5, pp. 357-365, 2013.

[6] P. L. Franzen and . D. J. Buysse, "Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications," Remission in Depression, vol. 10, no. 4, pp. 473-481, 2008.

[7] D. Neckelmann, A. Mykletun and A. A. Dahl, "Chronic Insomnia as a Risk Factor for Developing Anxiety and Depression," Sleep, vol. 30, no. 7, pp. 873-880, 2007.

[8] M. W. Agelink, C. Boz, H. Ullrich and J. Andrich, "Relationship between major depression and heart rate variability.," Psychiatry Research, vol. 113, no. 2, pp. 139-149, 2002.

[9] M. M. Ohayon, "Epidemiology of insomnia: what we know and what we still need to learn," Sleep Medicine Reviews, vol. 6, no. 2, pp. 97-111, 2002.

[10] D. Kuang, "Depression recognition according to heart rate variability using Bayesian Networks," Journal of Psychiatric Research, vol. 95, pp. 282-287, 2017.

[11] Y. Tazawa, K.-c. Liang, M. Yoshimura, M. Kitazawa, Y. Kaise and A. Takamiya, "Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning," Heliyon, vol. 6, no. 2, 2020.

[12] 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.

[13] A. Ghandeharioun†, S. Fedor†, L. Sangermano, D. Ionescu, J. Alpert, C. Dale, D. Sontag and R. Picard, "Objective Assessment of Depressive Symptoms with," 2017.

[14] I. Moshe, Y. Terhorst, K. O. Asare, L. B. Sander, D. Ferreira, H. Baumeister, C. M. David and L. Pulkki-Råback , "Predicting Symptoms of Depression," January 2021.

[15] nhs.uk, "Sleepio app - NHS," Play.google.com, [Online]. Available: https://www.nhs.uk/apps-library/sleepio/. [Accessed 14 Frbruary 2021]

[16] G. Play, "Stressscan App," Play.google.com, [Online]. Available: https://play.google.com/store/apps/details?id=com.dumsco.stressscan. [Accessed 14 February 2021].

[17] G. Play, "Welltory App," Play.google.com, [Online]. Available: https://play.google.com/store/apps/details?id=com.welltory.client.android [Accessed 14 February 2021].

[18] G. Play, "Moodpath App," Play.google.com, [Online]. Available: https://play.google.com/store/apps/details?id=de.moodpath.android&hl=en&gl=US. [Accessed 15 February 2021].

[19] Tran, Antoine, Laurie, Geghre, Nicolas, Darmon, David, Rampal and Marion, "Health assessment of French university students and risk factors associated with mental health disorders," vol. 12, no. 11, 2018.

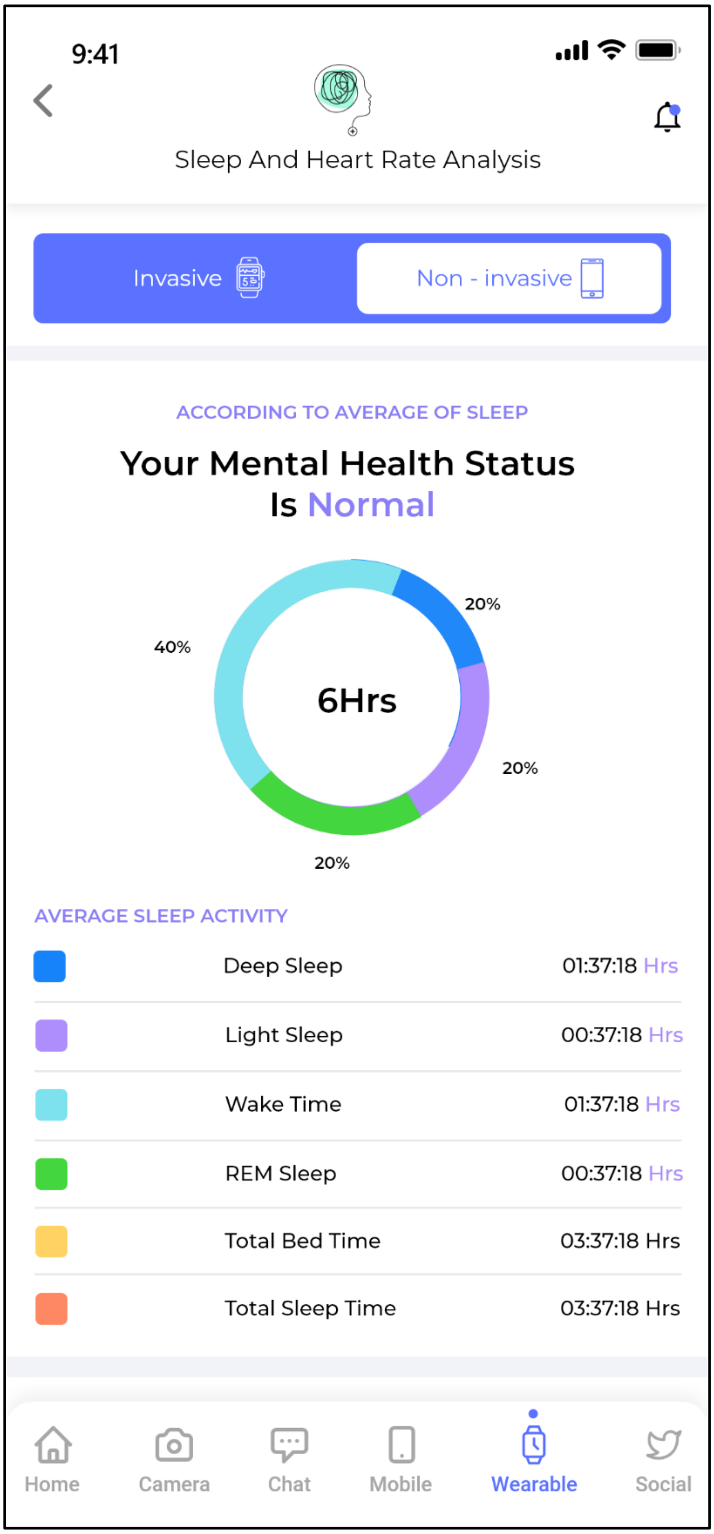
[20] M. E. Cristina, V. S. S. D. J., M. C. Roberta, L. and A. , "Natural Light Exposure, Sleep and Depression among Day Workers and Shiftworkers at Arctic and Equatorial Latitudes," vol. 10, no. 4, p. e0122078, 2015.

[21] Y. Christakis, N. Mahadevan and S. Patel, "SleepPy: A python package for sleep analysis from accelerometer data," [Online]. Available: https://joss.theoj.org/papers/10.21105/joss.01663. [Accessed 22 March 2021].

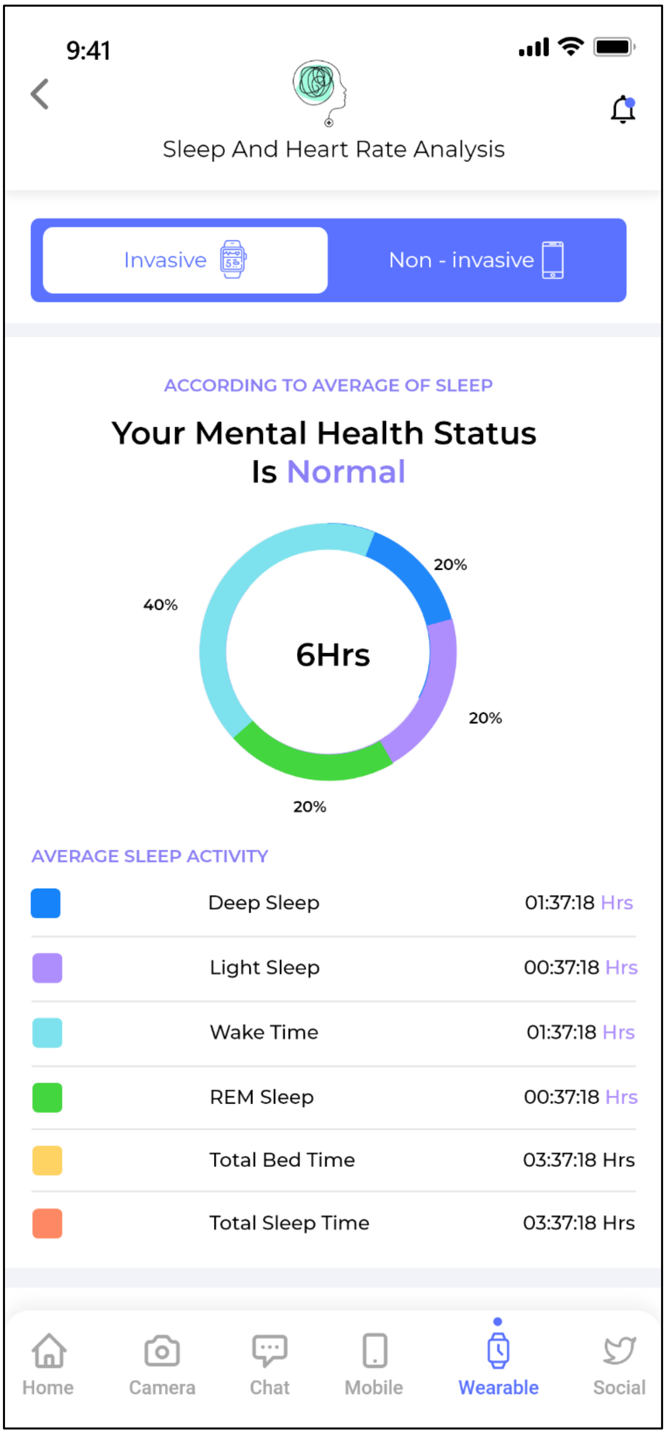
[22] Firebase, "Installation & Setup on Android | Firebase Documentation," Firebase, [Online]. Available: https://firebase.google.com/docs/database/android/start. [Accessed 21 June 2021].

Appendix A: User Interface.

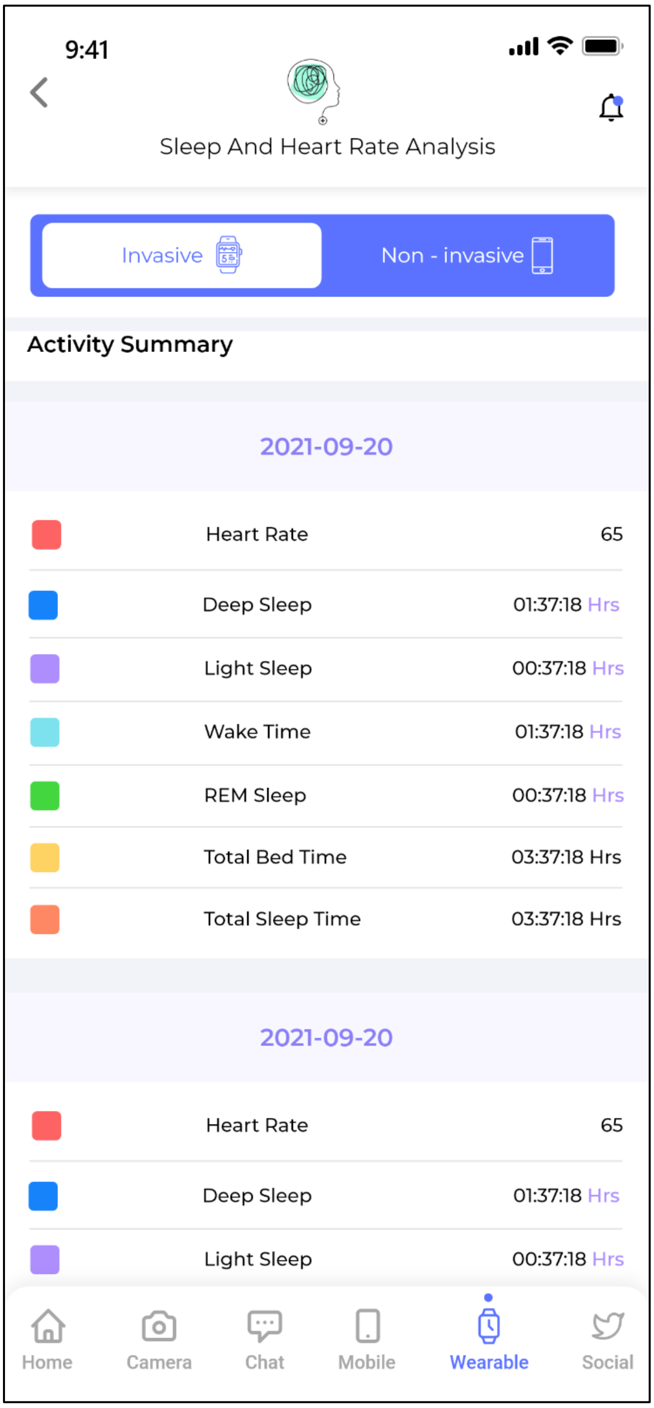
Non – Invasive sleep pattern Summary



Invasive sleep pattern Summary



Daily Biometric data Summary



Weekly Heart rate Summary