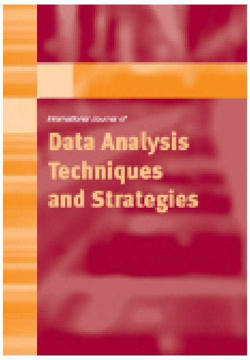


International Journal of Data Analysis Techniques and Strategies

ISSN online: 1755-8069 - ISSN print: 1755-8050
https://www.inderscience.com/ijdats

A user friendly anger and anxiety disorder prediction scheme using machine learning and a mobile application for mental healthcare

Mohammad Salah Uddin, Mahfuzulhoq Chowdhury

DOI: 10.1504/IJDATS.2024.10061933

Article History:

Received:	02 August 2023
Last revised:	02 December 2023
Accepted:	04 December 2023
Published online:	19 March 2024

A user friendly anger and anxiety disorder prediction scheme using machine learning and a mobile application for mental healthcare

Mohammad Salah Uddin and
Mahfuzulhoq Chowdhury*

Computer Science and Engineering Department,
Chittagong University of Engineering and Technology,
Chittagong – 4349, Bangladesh
Email: u1704123@student.cuet.ac.bd
Email: chowdhurymhoq@gmail.com
Email: mahfuzulhoq.cse05@gmail.com
*Corresponding author

Abstract: The growing prevalence of mental health disorders concerns has motivated the development of innovative technologies to support mental well-being. The previous literary works on mental healthcare did not investigate anger and anxiety disorder prediction by considering 23 features. There is a lack of mental healthcare assistance mobile applications in the literary works by considering anger and anxiety assessment and necessary emergency assistance features. To solve these issues, this paper initiates a machine learning model-based anger and anxiety prediction scheme by examining different machine learning algorithms. Our analytical results show that the logistic regression model shows better prediction results among all machine learning algorithms in terms of higher accuracy, precision, recall, and error rate. This paper presents a mental healthcare mobile application with anger and anxiety assessment, physical exercise suggestions, hospital search, doctor appointment booking, and emergency contact. The evaluation result shows the efficiency of the proposed scheme.

Keywords: anger and anxiety prediction; self-assessment tools; mental healthcare; mobile application; machine learning; logistic regression; support vector machine; SVM; K-nearest neighbours; KNN; decision tree; multi-class classification.

Reference to this paper should be made as follows: Uddin, M.S. and Chowdhury, M. (2024) 'A user friendly anger and anxiety disorder prediction scheme using machine learning and a mobile application for mental healthcare', *Int. J. Data Analysis Techniques and Strategies*, Vol. 16, No. 1, pp.47–81.

Biographical notes: Mohammad Salah Uddin received her BSc in Computer Science and Engineering from the Department of CSE, Chittagong University of Engineering and Technology in 2023. Her research interests include machine learning, deep learning, artificial intelligence, and mobile application development.

Mahfuzulhoq Chowdhury received his PhD in Telecommunications from the University of Quebec, INRS, Montreal, Canada. He is a faculty member at the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology since 2010. His research interests are related to machine learning, optimisation, mobile application development, game theory, resource allocation in future generation networks, cloud computing, IoT, age of information, and blockchain. He has published several research papers at highly cited IEEE journals, transactions, magazines, and conference proceedings as well as chapters of books.

1 Introduction

Anger and anxiety disorder can be termed as a mental illness, and it can seriously harm a person's social and personal life (Arif et al., 2020). Anger and anxiety need to be carefully assessed because they are important indicators of psychological distress and wellbeing. WHO (1993) states that worrying about various problems and experiencing excessive fear are symptoms of anxiety disorders. Among other symptoms, it may cause high headaches, high heart-bit racing, and high chest pain. Anger out of control can have detrimental long-term physical effects, including headaches, elevated blood pressure, and increased anxiety (Antonatos, 2023). Two emotions are inextricably linked in the rich and varied palette of human emotions. Anxiety is defined as the worry or fear you experience in response to a perceived threat. Anger is a threat response, but it is accompanied by a strong sense of annoyance. Researchers believe that these two emotions play an important role in our ability to detect and respond to danger (Healthline Authority, 2023). Anger and anxiety are closely related, despite the fact that this relationship is frequently disregarded. Anxiety is defined as worry and fear that are elevated above and beyond ordinary worries. Anger and anxiety can coexist for a variety of reasons. Anxiety is frequently triggered by overstimulation in a stressful situation coupled with a helpless feeling to control the stressor (Sagar, 2023). However, anger is often the result of frustration. Unspoken anxiety frequently turns into frustration, which eventually turns into anger. When anxiety gives rise to anger, it is typically a subconscious response to some underlying fear or concern that the person is experiencing. They use their anger as a means of regaining control over their anxious emotions in these circumstances (Sagar, 2023). Use the appropriate example of road rage. People who experience anxiety may become hostile toward other drivers as a result of traffic and congestion on the roads. This outward display of anger often conceals a deeper anxiety caused by internal pressures like impending obligations, mood swings, or work deadlines. The aggravation of traffic jams adds to their emotional turmoil.

Many psychological problems, such as stress, depression, and anxiety and anger management, are prevalent in today's society (Priya et al., 2020). The brain's reaction to stress over certain matters is anxiety. Anxiety can affect people at different times in their lives. Anxiety disorders are a class of mental illnesses that can lead to excessive and ongoing feelings of fear and anxiety. People may avoid their jobs, social events, interactions with others, schoolwork, family work, business work, and office work because of extreme anxiety (Sharma et al., 2021). People may experience frustration and helplessness in life. Appropriate treatment and monitoring of anger and anxiety disorders are required in order to prevent this problem and enable people to be productive.

A statistic indicates that mental health disorders affect approximately 25% of the world's population. Additionally, data indicates that approximately seven million Bangladeshis experience various mental health disorders, including anxiety, depression, and anger management (Arusha et al., 2020). People may experience stress, anger, and anxiety for a variety of reasons, including non-academic and academic issues, socioeconomic issues, environmental issues, cultural issues, and psychological issues. Their high levels of anxiety and rage can lead to unproductive social and personal lives. People in developing nations like Bangladesh typically do not give much thought to mental health issues like anxiety, depression, and anger.

Mental illness is unquestionably a health problem that influences a person's thinking, feelings, and social interactions. These issues have shown that mental illness, including depression, anxiety, and anger, has negative impacts on society at large and calls for the creation of novel preventative and intervention strategies. A critical first step in putting these strategies into practice is early mental health detection. Medical predictive analytics is expected to completely transform the healthcare sector, according to Miner et al. (2016). The patient's self-report, which necessitates the use of questionnaires meant to identify specific emotional or social patterns, is usually used to make the diagnosis of mental illness (Hamilton et al., 1967). Many people with mental illness or emotional disorders should be able to recover with the right care and treatment (American Psychiatric Association, 2013). The best way to deal with anger and anxiety, according to Sagar (2023), is to use an intelligent system for early detection and medical treatment (e.g., mobile application-based doctor consultation and machine learning-based mental disease prediction). In response to the rise in mental health issues like anger and anxiety as well as the need for effective medical care, researchers are looking into developing an intelligent system using machine learning that can be used to forecast future events and categorise mental health problems (Chung et al., 2022).

Therefore, the primary motivation of this work is to reduce the harmful effects of anxiety and anger via an intelligent system (Antonatos, 2023). To do so, a machine learning-based system for detecting anxiety and anger as well as a healthcare support system based on mobile applications are being developed. This paper's primary goal is to diagnose different mental illness levels (such as anxiety and anger disorders) by determining the emotional state of the subject through the use of machine learning algorithms, a standard psychological evaluation, and a mobile application-based mental healthcare assistance system.

In order to detect anxiety and anger in individuals by examining multiple features, the majority of the existing works (Priya et al., 2020; Sharma et al., 2021; Arusha et al., 2020; Sumathi et al., 2016; Ahmed et al., 2020; Hilbert et al., 2017) did not offer any machine learning-based classification model with high accuracy. Additionally, they did not offer any mobile applications that would have helped people with their anxiety and anger management. The current works lack an appropriate dataset for predicting anxiety and anger. Another gap in the works that are currently available is the incorporation of a machine learning-based system for predicting anger and anxiety into the mobile application. The majority of published works did not provide a suitable framework for analysing the relationship between anxiety and anger.

To help people with mental healthcare, an automated system is necessary that would use machine learning algorithms to diagnose anger and anxiety disorders. To prepare the dataset, a questionnaire is required that may include mental and physical conditions, current/past psychological health background, relationship status, and so on. For a more accurate prediction of mental disorders, different machine learning algorithms need to be examined in terms of accuracy, precision, recall, and error rate. This type of assistance system will assist people in identifying mental issues such as anger and anxiety that require counselling by analysing the prepared dataset. An Android-based mobile application is also required that will provide mental healthcare to those in need by assessing their anger and anxiety disorder level, emergency help, doctor suggestions, and hospital search, among others.

1.1 Contributions of this paper

By examining numerous machine learning classifiers and 23 features, this paper presents a user-friendly machine learning-based anger and anxiety disorder prediction scheme, addressing the drawbacks of earlier research. It should be mentioned that earlier studies did not examine a prediction scheme that examined both anger and anxiety disorders using multiple machine learning classifiers and 23 features. Unlike existing works, this work gives a multi-class classification approach (neither anger nor anxiety, anger only, anxiety only, both anger and anxiety) using machine learning by analysing the dataset. This paper also describes a multifeature-based mobile application that serves as a healthcare assistance system in addition to an anger and anxiety assessment.

The following list includes this paper's main contributions:

1 This paper constructs a robust prediction model capable of predicting whether an individual is experiencing mental health issues, such as anxiety or anger. This model leverages machine learning techniques to analyse relevant data and provide accurate predictions. The proposed scheme also selects suitable machine learning models for anger and anxiety prediction by assessing different machine learning classifiers such as logistic regression, random forest, support vector machine (SVM), K-nearest neighbours (KNN) and decision tree.

2 To train and evaluate the predictive model, this paper collects data from users and prepares a mental healthcare dataset with anger and anxiety issues. The data collection process involves gathering relevant information related to indicators and factors associated with anger and anxiety disorder.

3 This paper analyses the collected data and extracts valuable information regarding the presence and severity of anxiety and anger symptoms.

4 This paper develops a mental healthcare mobile application with an anger and anxiety disorder assessment feature, doctor search and appointment booking feature, hospital search feature, physical exercise and mental healthcare article reading feature, and emergency contact feature, among others.

5 This work also gives user evaluation results regarding the proposed mobile applications' necessity, user interface and design, and personalised recommendation query.

The existing works discussion will be given in Section 2. The proposed machine learning-based anger and anxiety prediction model will be discussed in Section 3. The proposed mental healthcare mobile application features will be discussed in Section 4. Section 5 holds the user-based evaluation results. Section 6 summarises the findings and future challenges of this work.

2 Related works

Anger and anxiety are two distinct emotional states that commonly coexist and interact within individuals. Anger can be defined as a strong feeling of displeasure or hostility, often accompanied by a desire to retaliate or express aggression. Anxiety, on the other hand, is characterised by feelings of fear, apprehension, and unease about future events or uncertain circumstances. Both anger and anxiety are normal and adaptive responses to various stimuli, but excessive or prolonged experiences of these emotions can lead to negative consequences for individuals. The relationship between anger and anxiety is complex and intertwined. It is not uncommon for individuals with anxiety to also experience anger, as anxiety can manifest as irritability, restlessness, or a sense of being on edge. Anxiety can heighten emotional reactivity, making individuals more prone to anger outbursts or aggressive behaviour as a way to cope with perceived threats or regain a sense of control. Similarly, anger can be a response to underlying feelings of anxiety or fear, as individuals may become frustrated or agitated when their expectations are not met or when they perceive a threat to their well-being. Furthermore, there is evidence suggesting that chronic or unmanaged anger can contribute to the development or exacerbation of anxiety disorders. Individuals who struggle with controlling their anger may experience heightened levels of stress and arousal, which can trigger or worsen symptoms of anxiety. The constant activation of the body's stress response system can lead to a vicious cycle, where anger fuels anxiety and vice versa. Study proves highly anxious individuals have been found to react to stress with increased irritability, hostility, anger attacks, indirect aggression, and direct aggression (Thompson et al., 2021).

Anger and anxiety are interconnected emotional states that can influence and exacerbate each other. Recognising and understanding the relationship between anger and anxiety is crucial for providing comprehensive mental health support and interventions to individuals experiencing these emotional challenges.

An anger rating scale and an anxiety rating scale are commonly used assessment instruments to evaluate an individual's level of anger and anxiety. These scales consist of a series of questions and inquiries that participants complete to provide insights into their emotional state. The primary goal of these scales is to assess the severity of anger and anxiety symptoms and identify individuals who may be at risk of developing anger issues or anxiety disorders. Psychologists or specialists utilise the Buss-Perry aggression questionnaire (BPAQ), which has gained recognition as a gold standard for measuring aggression. Psychologists and researchers widely utilise the BPAQ to assess various types of aggression (e.g., physical or verbal), anger issues, and hostility issues (Gerevich et al., 2007). By incorporating the BPAQ into the assessment process, psychologists can obtain comprehensive insights into an individual's aggressive tendencies, allowing for a more nuanced understanding of anger-related issues and providing valuable information for interventions and treatment planning. Additionally, the BPAQ's established reputation and extensive use in research make it a valuable tool for comparing aggression across different populations and studying the underlying mechanisms of aggressive behaviour.

Similarly, the GAD-7 scale is an adequate measure to detect generalised anxiety symptoms (Monteiro et al., 2017). It consists of questions that measure various manifestations of anxiety, such as excessive worry, restlessness, difficulty concentrating, and sleep disturbances. Psychologists and healthcare professionals employ the GAD-7 scale to determine the degree of anxiety experienced by individuals and to assess whether further evaluation or diagnosis of an anxiety disorder is necessary. These rating scales, including the BPAQ for anger assessment and the GAD-7 for anxiety assessment, play vital roles in identifying, quantifying, and understanding the presence of anger and anxiety in individuals. They provide valuable insights that help guide interventions, treatment planning, and support for individuals experiencing anger and anxiety-related difficulties.

In Kapoor et al. (2022), the authors explore various machine-learning models for predicting anxiety disorders based on survey questionnaires. However, it has some limitations. Notably, it fails to analyse the connection between anger and anxiety, a crucial aspect of mental health. Additionally, the absence of an integrated user interface restricts the models' practicality for individuals seeking mental health support. Furthermore, the study does not include the development of an application to aid users' mental well-being. Addressing these limitations would improve the models' comprehensiveness and accessibility for real-world use, offering more holistic support to individuals facing mental health challenges. In Walsh et al. (2018), the authors investigated the relationship between anger and anxiety in young individuals. They discovered that elevated levels of anger were associated with a higher likelihood of being diagnosed with generalised anxiety disorder (GAD). However, the research has several limitations. Firstly, the small sample size of only 40 participants might impact the study's generalisability. Secondly, the analysis solely focused on children, limiting its applicability to other age groups. Additionally, the absence of a dedicated model for predicting anger and anxiety may restrict the accuracy of predictions.

In Thompson et al. (2021), the authors show that anxiety sensitivity holds promise as a treatment target for anger symptoms. However, the study has certain limitations. Firstly, it lacks classification for detecting anxiety and anger in individuals. Secondly, the absence of machine learning utilisation for prediction hinders the model's potential accuracy. Moreover, the research fails to provide mental support or suggestions for individuals in need, potentially limiting its practical impact. Lastly, no app or user interface has been developed to support those dealing with mental health disorders. Addressing these limitations could enhance the study's applicability and efficacy, offering better support and insights for individuals experiencing anxiety and anger-related issues. In Hilbert et al. (2017), the authors developed an SVM-based ML model that used multi-modal bio-behavioural data from GAD. To classify the mental disorder subjects from healthy subjects, they also used clinical questionnaire data along with the MRI data. The MRI data provides good accuracy but they did not get good accuracy with questionnaires data. The limitation of this paper is that their clinical questionnaires were score-based. Also, they did not develop an app or user interface.

In Ahmed et al. (2020), the authors described an ML-based approach to detect depression and anxiety. They applied several algorithms such as linear regression, KNN or K-nearest neighbour, SVM or support vector machines, LDA, and CNN or convolutional neural network. The limitation of this paper is that they did not implement any improvement monitoring system after therapy. In Priya et al. (2020), the authors used image processing to better understand the patients' psychological states (e.g., anxiety, depression, and post-traumatic disorder). The limitation of this paper is that they did not consider the patient's feelings and background history.

In Deeba et al. (2004), the authors developed an anxiety scale for Bangladeshi people based on an online survey from a different perspective of Bangladeshi people. The limitation of this paper is that they did not apply any model to predict anxiety. In Sumathi et al. (2016), the authors compared eight machine learning methods for categorising the dataset according to various mental health issues. The limitation of this paper is that their dataset was not large enough to get better accuracy. Their work is limited only to children rather than adults. In Masri et al. (2012), the authors used AI or artificial intelligence methods in the diagnosis of mental health. Their reasoning techniques are rule-based reasoning, fuzzy logic, and fuzzy-genetic algorithm. The limitation of this paper is that their approach was rule-based reasoning, which is pretty straightforward. Their prototype is only for a small area in Malaysia. They did not present any immediate help service for life-threatening situations.

In Burkhardt et al. (2006), the authors developed a concept of an emotion-aware VoiceXMLbased portal to predict anger situations from customer care call centre data. This can be useful measurement and emotion-aware dialogue strategies. The limitation of this paper, the voice data can be in low bandwidth and may cause an accuracy fall. Also, they did not apply any counselling-based anger prediction model.

By examining the tweets, profile features, social media data, and hybrid knowledge of the authors, the work in Gupta et al. (2023) developed a bi-directional long short-term memory, or bi-LSTM, based depressed social user detection system. Several machine learning algorithms (KNN, SVM, and random forest) were examined in Siddique et al. (2022) to predict depression in Bangladeshi university students. Both the partitional CNN model and the LSTM model were employed in the work in Ismail et al. (2022) to predict the verbal offence or insult pattern by monitoring social media comments. A model for the detection of chronic kidney disease was created by Khamparia et al. (2020) by combining principal component analysis (PCA) and SVM techniques.

Differing from the existing works, this paper develops a user-friendly anger and anxiety disorder prediction scheme using machine learning techniques. This paper also developed a mobile application for anger and anxiety disorder assessment-based mental healthcare assistant facilities.

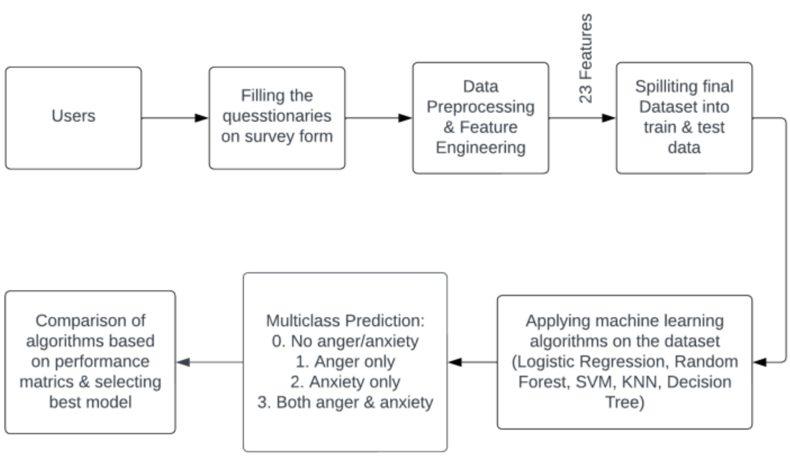
3 Proposed model

This section provides an overview of our model's design for predicting anxiety traits and anger issues. It details the steps of data collection and cleaning, ensuring dataset reliability.

3.1 System design

The system design process, as depicted in Figure 1, encompasses the overall structure and flow of our model. It outlines the steps from data collection to model training, highlighting the key components and their interconnections. The procedure for our prediction model begins with administering questionnaires to individuals, which cover various aspects such as family income support, psychological health background, academic performance, political involvement, relationship problems, drug addiction, and experiences of ragging/bullying. The collected data is then organised in an Excel sheet for further analysis. Next, a data cleaning process is conducted to eliminate incomplete or erroneous entries, ensuring data quality and reliability. This step is crucial to prepare the dataset for analysis. To train the model, we employ a range of machine-learning algorithms, including logistic regression, KNN, SVM, decision tree, and random forest. Each algorithm has its strengths in different scenarios, some performing better with smaller datasets while others with larger datasets. By utilising multiple algorithms, we aim to maximise the accuracy of our model in identifying individuals who may be experiencing anger and anxiety disorders.

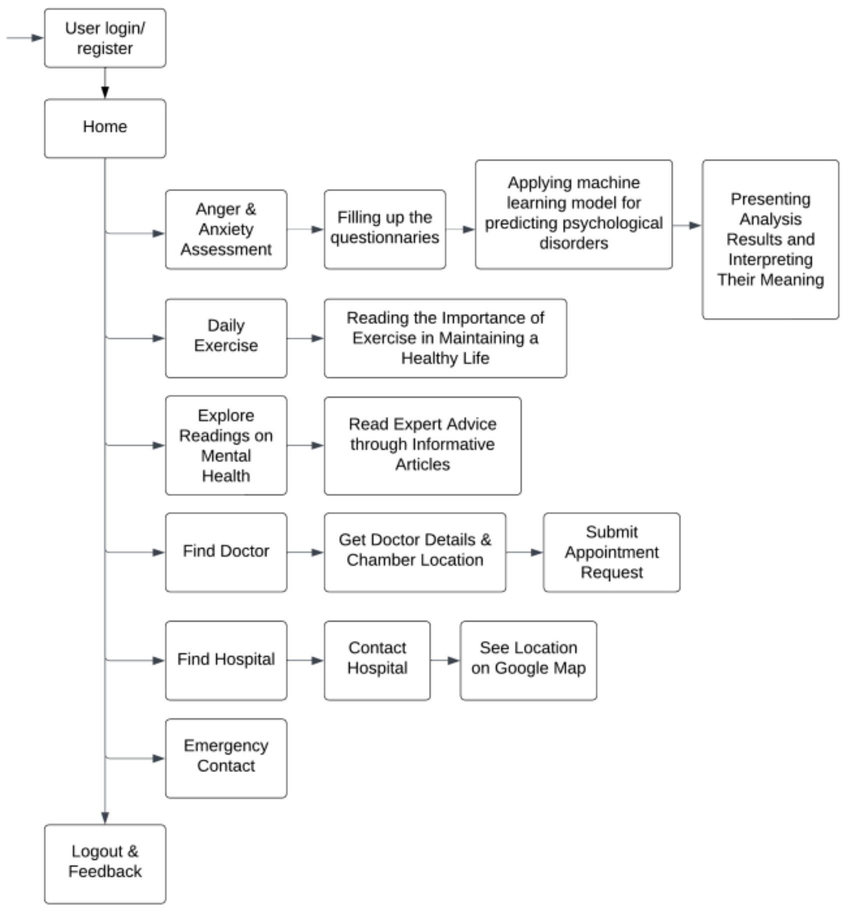
Figure 1 System design for anger and anxiety analysis



With a dataset of only 1,001 survey responses, fewer than 30 features, and a requirement for results that are easily interpreted, traditional machine learning models make more sense for this application than deep learning. Deep learning is probably not going to work well on a dataset with fewer than 500,000 items and fewer than 30 features. It is very difficult for the deep learning algorithms to learn the features with a small dataset. Additionally, there is a good chance that the algorithm will overfit. Conventional machine learning models are a more realistic option in this particular situation because they require less data and provide more transparency in their decision-making process. For this reason, instead of looking into deep learning models, we looked into conventional machine learning models in this case. Therefore, using deep learning models to predict anger and anxiety is outside the scope of this work and will be considered as a future research issue.

Figure 2 illustrates the complete data flow diagram of the Android-based mobile application for anger and anxiety assistance. Upon logging in, users will be presented with a questionnaire to assess the underlying causes of their anxiety and anger. The questionnaire responses will be analysed to identify psychological problems such as experiences of bullying, addiction issues, or recent personal difficulties. Users will have the option to share their concerns with a psychologist who is an existing user of the application. Users will also have the ability to request appointments with available psychologists through the application. When a psychologist logs into the application, they will be able to view user information, along with the specific psychological problems and any personal messages shared by the users. Psychologists can then approve appointments and schedule sessions with individual users, who will receive notifications regarding their scheduled appointments.

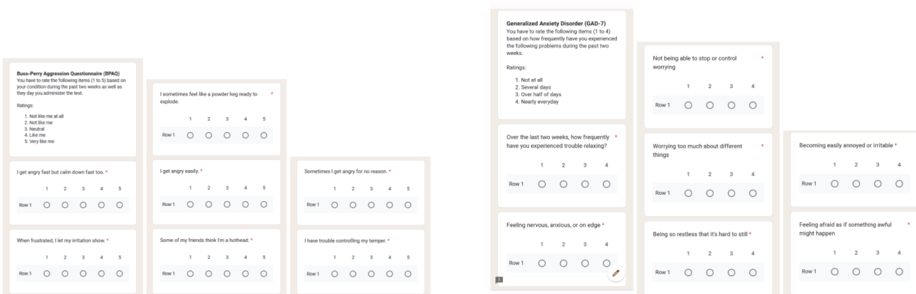
Figure 2 Block diagram of Android application



3.2 Data collection and dataset preparation

In this section, we embark on the analysis of the collected data to achieve two primary objectives. Firstly, the data is utilised to conduct a comprehensive analysis using various machine learning algorithms. Two distinct datasets are considered for this analysis. The first dataset, named S-9, comprises nine features, offering insights into students' personal lives. The second dataset, named S-14, includes data from the Buss-Perry aggression questionnaires (BPAQ) scale (Gerevich et al., 2007) and generalised anxiety disorder (GAD-7) scale (Monteiro et al., 2017), providing information on aggression and anxiety symptoms. Figures 3(a) and 3(b) displayed the BPAQ and GAD-7 scaling methods, respectively.

Figure 3 (a) Seven items of Buss-Perry aggression questionnaires (BPAQ) (b) Seven items of generalised anxiety disorder (GAD-7)



(a) (b)

To ensure the dataset's accuracy and relevance, extensive research was conducted to identify the most significant factors associated with anger and anxiety. These factors were carefully integrated into the questionnaire, forming the foundation of our initial dataset discussed below.

According to the study conducted by Owens et al. (2012), higher levels of anxiety and depression were found to be associated with lower academic performance. This suggests that an individual's academic condition can have a significant impact on their mental health, particularly in terms of anxiety (Savage et al., 2017). According to Bögels et al. (2006), the intergenerational transmission of anxiety involves a variety of family factors. These factors contribute to the likelihood of anxiety being passed down from one generation to another. The study emphasises that family dynamics, parenting styles, and the modelling of anxious behaviours all play significant roles in shaping the transmission of anxiety within families (van Santvoort et al., 2015). Numerous studies indicate a strong connection between drug addiction or alcohol misuse and anxiety and anger. In Goodwin et al. (2013), the authors found a relationship between drug addiction and anxiety disorders in a study of 5,788 adults in the USA. In Nichols et al. (2008), the authors reported that anger was significantly associated with increased smoking, drinking, and marijuana use among 2961 middle school students in New York City. The experience of sadness resulting from the death or loss of a loved one is an important feature for predicting anxiety and anger. In Bonanno et al. (2004), the authors provided insights into the psychological impact of bereavement and highlight how sadness can significantly impact mental health. Additionally, in Rosner et al. (2011), the authors explored the complexities of grief and its treatment, shedding light on the emotional turmoil individuals may experience following a loss. The impact of financial problems on an individual's anxiety and anger is significant. A study carried out by the University of Southampton and the Solent NHS Trust, focusing on first-year British undergraduate students, established a clear connection between financial difficulties, concerns about debt, and mental health issues like anger, anxiety, depression, and alcohol use (Richardson et al., 2017).

Figure 4 Prepared dataset (see online version for colours)

	F	G	H	I	J	K	L	M	N
1	Are you hap	Do you have	Are you add	Sadness fro	Do you have	Violence in f	Do you ever	If yes does i	How many h
2	Yes	No	Yes	Yes	Yes	Yes	Yes	Rarely	1 to 3 hours
3	No	No	No	No	Yes	No	Yes	Rarely	3 to 6 hours
4	Yes	No	Yes	No	No	No	Yes		3 to 6 hours
5	No	No	No	No	No	No	Yes		Greater than
6	No	No	No	No	Yes	Never	Yes	Rarely	1 to 3 hours
7	Yes	No	No	No	Yes	Never	Yes	Rarely	3 to 6 hours
8	No	No	No	No	No	Never	No		3 to 6 hours
9	No	No	No	No	No	Often	Yes	Rarely	3 to 6 hours
10	Yes	No	No	Yes	No	Never	No	Never	1 to 3 hours
11	Yes	No	No	No	Yes	Never	No		Less than 1
12	Yes	No	No	Yes	No	Never	Yes		Greater than
13	No	No	No	No	No	Rarely	No		1 to 3 hours
14	No	No	No	No	No	Rarely	No		1 to 3 hours
15	No	No	No	No	No	Rarely	No		3 to 6 hours
16	No	No	No	Yes	Yes	Rarely	Yes	Often	3 to 6 hours
17	Yes	No	No	No	Yes	Never	Yes	Rarely	1 to 3 hours
18	Yes	No	Yes	No	No	Rarely	Yes	Often	1 to 3 hours
19	No	No	Yes	Yes	Yes	Rarely	No		Greater than
20	No	No	No	No	Yes	Never	Yes	Often	Greater than
21	No	Yes	No	No	No	Rarely	Yes	Rarely	3 to 6 hours
22	No	No	No	Yes	No	Rarely	Yes	Rarely	Greater than
23	No	No	No	Yes	Yes	Never	Yes	Rarely	Greater than
24	Yes	Yes	No	Yes	No	Never	Yes	Rarely	3 to 6 hours
25	Yes	No	No	No	Yes	Often	No		1 to 3 hours
26	No	No	No	Yes	No	Never	Yes	Rarely	3 to 6 hours
27	No	Yes	Yes	No	No	Never	No		1 to 3 hours
28	No	Yes	No	Yes	Yes	Often	Yes	Most of the t	1 to 3 hours
29	No	Yes	No	Yes	Yes	Often	Yes	Often	3 to 6 hours
30	Yes	No	No	No	No	Never	No		1 to 3 hours
31	No	No	No	No	No	Rarely	No	Rarely	3 to 6 hours
32	Yes	No	No	No	No	Never	Yes	Rarely	3 to 6 hours
33	No	No	No	No	No	Never	No	Never	1 to 3 hours

	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	I get angry fas	When frustrat	I sometimes f	I get angry ea	Some of my f	Sometimes I	I have trouble	Over the la	Feeling ne	Not being a	Worrying t	Being so r	Becoming	Feeling afr
2	4	5	4	3	4	1	2	2	1	1	3	4	4	3
3	2	2	2	2	3	1	3	1	3	3	1	1	1	1
4	2	1	1	2	2	4	2	3	1	1	1	1	1	1
5	3	1	2	2	2	2	2	1	1	1	1	1	1	1
6	4	3	3	3	4	1	2	2	2	1	4	3	2	3
7	3	2	4	2	3	1	3	2	2	4	4	3	4	2
8	2	1	3	2	2	4	2	4	4	4	4	3	2	2
9	4	5	3	5	3	3	4	2	2	4	2	1	3	3
10	4	4	1	2	2	2	4	1	2	2	2	1	1	2
11	3	3	3	3	3	3	3	3	3	3	3	3	3	3
12	3	2	2	2	2	2	2	3	2	2	2	3	2	3
13	2	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	4	3	2	3	2	2	4	3	3	3	1	2	2	2
16	4	1	3	3	1	3	1	1	2	2	2	3	2	2
17	4	2	1	4	3	4	3	4	2	3	3	3	2	4
18	5	4	1	4	2	4	3	2	3	2	2	2	2	1
19	1	3	3	1	2	5	5	2	2	2	2	2	2	2
20	3	2	1	2	3	2	3	4	3	4	4	2	2	4
21	4	4	2	4	2	2	4	4	3	4	3	3	4	2
22	2	3	5	3	1	3	5	1	4	3	3	2	1	2
23	5	4	4	4	4	2	3	4	4	4	4	4	4	4
24	3	5	1	1	1	5	4	2	2	4	4	4	1	4
25	4	4	1	3	1	1	1	3	1	3	3	1	1	3
26	4	3	3	2	2	2	2	2	3	1	2	2	3	1
27	4	4	4	3	3	2	3	1	2	2	1	2	1	1
28	5	5	5	4	5	5	4	3	3	4	4	3	3	3
29	4	2	3	4	4	1	2	3	3	1	2	2	2	1
30	2	2	1	1	1	1	1	1	2	2	2	2	1	1
31	2	1	2	1	2	2	1	1	1	2	2	2	1	2
32	3	2	3	2	4	2	2	2	2	2	2	1	2	1
33	3	1	1	2	2	3	2	2	3	2	1	1	2	1

Table 1 Dataset labelling

BPAQ score	*GAD-7 score*	*Emotional state*	*Label*
<17	<10	No anger/anxiety	0
≥17	<10	Anger only	1
<17	≥10	Anxiety only	2
≥17	≥10	Both anger and anxiety	3

Violence in the family is a contributing factor in the prediction of anxiety and anger due to its well-documented impact on individual mental health. Research has consistently shown that exposure to violence within the family can have detrimental effects on an individual's psychological well-being. Sadof et al. (1976) provide an insightful perspective on this issue. Additionally, the study conducted by McCloskey et al. (1995) further emphasise the relationship between family violence and mental health, specifically anxiety. Being a victim of bullying can have significant effects on an individual's mental well-being, contributing to increased levels of anxiety and anger. In Copeland et al. (2013), the authors conducted a study on adult psychiatric outcomes of bullying, highlighting the long-lasting impact of childhood and adolescent victimisation. In Hawker et al. (2000), the authors reviewed two decades of research on peer victimisation and psychosocial maladjustment. Including the feature of hours spent on social media in the study of anxiety and anger prediction is crucial due to the evidence provided by Keles et al. (2020) in their systematic review. Figure 4 illustrates the entirety of our dataset. We have collected our dataset from 1,001 participants via a survey form. The 23 questionnaire of our dataset is shown in Figure 11. We have validated our dataset by three experts. One professor expert from Chittagong Engineering University (CUET), one professor from Dhaka University, and one doctor expert from Chittagong medical college. The prediction model is based on a scaling system. For the BPAQ scores, ranging from 1 to 5, a score of 1 indicates 'very unlike me', while a score of 5 signifies 'very like me'. The BPAQ scores are summed, and scores above 17 are categorised as anger, while scores below 17 are considered as non-anger. Similarly, in the GAD-7 scale, the scores measure the severity of generalised anxiety disorder symptoms and range from 0 to 3. A score of 0 implies 'not at all', while a score of 3 denotes ‘nearly every day’. Scores greater than 10 indicate anxiety, while scores less than 10 indicate non-anxiety. Table 1 shows the details of the labelling.

3.2.1 Data cleaning and data preprocessing

Finally, after gathering data from 1,001 students, we decided to begin making predictions based on the information we had gathered. However, before applying the prediction algorithms to the data, we cleaned it to eliminate as many outliers as possible to improve accuracy. As a result, we produced different data versions. Our survey had three parts. The first part contained students’ information like family income support, present/past psychological health background, academic result, political involvement, relationship problem, drug addiction, victim to ragging/bullying, etc. The second part contained BPAQ scaling and the third part contained GAD-7 scaling. Since our prediction model will be based on this scaling system, both the scaling were very important for us. After collecting the dataset through the consent form, a data cleaning process was conducted to ensure the quality and integrity of the data. Any missing values in the dataset were identified and addressed. Missing values could arise from respondents choosing not to answer certain questions or technical issues during data collection. Mean imputation techniques were employed to fill in the missing values based on the characteristics of the other variables. Then, the BPAQ and GAD-7 scores were scaled to numerical values to facilitate the machine learning process. The BPAQ scores originally range from 1 to 5, with 1 indicating ‘very unlike me’ and 5 indicating ‘Very like me’. These scores were mapped directly to the 1 to 5 scale. The GAD-7 scores, which measure the severity of generalised anxiety disorder symptoms, ranging from 0 to 3, with 0 indicating ‘not at all’ and 3 indicating ‘Nearly every day’. Based on the scaled BPAQ and GAD-7 scores, a new column called ‘emotional state’ was created to categorise the participants into four distinct groups: ‘neither anger nor anxiety’, ‘only anger’, ‘only anxiety’, and ‘both anger and anxiety’. The ‘emotional state’ column was derived by analysing the scaled BPAQ and GAD-7 scores are as follows. If the scaled BPAQ score and the scaled GAD-7 score were both less than the threshold, the participant was labelled as having ‘neither anger nor anxiety’. If the scaled BPAQ score was greater than the threshold and then scaled GAD-7 score was less than the threshold, the participant was labelled as ‘only anger’. If the scaled BPAQ score was less than the threshold and the scaled GAD-7 score was greater than the threshold, the participant was labelled as having ‘only anxiety’. If both the scaled BPAQ score and the scaled GAD-7 score were greater than the threshold, the participant was labelled as ‘both anger and anxiety’. By creating the ‘emotional state’ column, the dataset was enriched with a categorical variable that encapsulated the participants’ emotional states, allowing for further analysis and prediction of anxiety and anger levels using machine learning algorithms. This dataset has 9 features and we named it S-9. The other dataset contained the standard questions from BPAQ and GAD-7 which has 14 different personality trait questions and we named it S-14. Then we used our five prediction algorithms to see what might unfold. We ran all the algorithms based on the final label and our features on two separate datasets to see how accurate the predictions were.

3.2.2 Data visualisation

To understand the significance of the data, Figure 5 shows the data visualisation diagram based on different questionnaires.

Figure 5 (a) Data comparison for academic condition (b) Data comparison for family history of mental illness (c) Data comparison for drug addiction (d) Data comparison for sadness (e) Data comparison for financial problem (f) Data comparison for family violence history (g) Data comparison for bullying (h) Data comparison for social media addiction (see online version for colours)

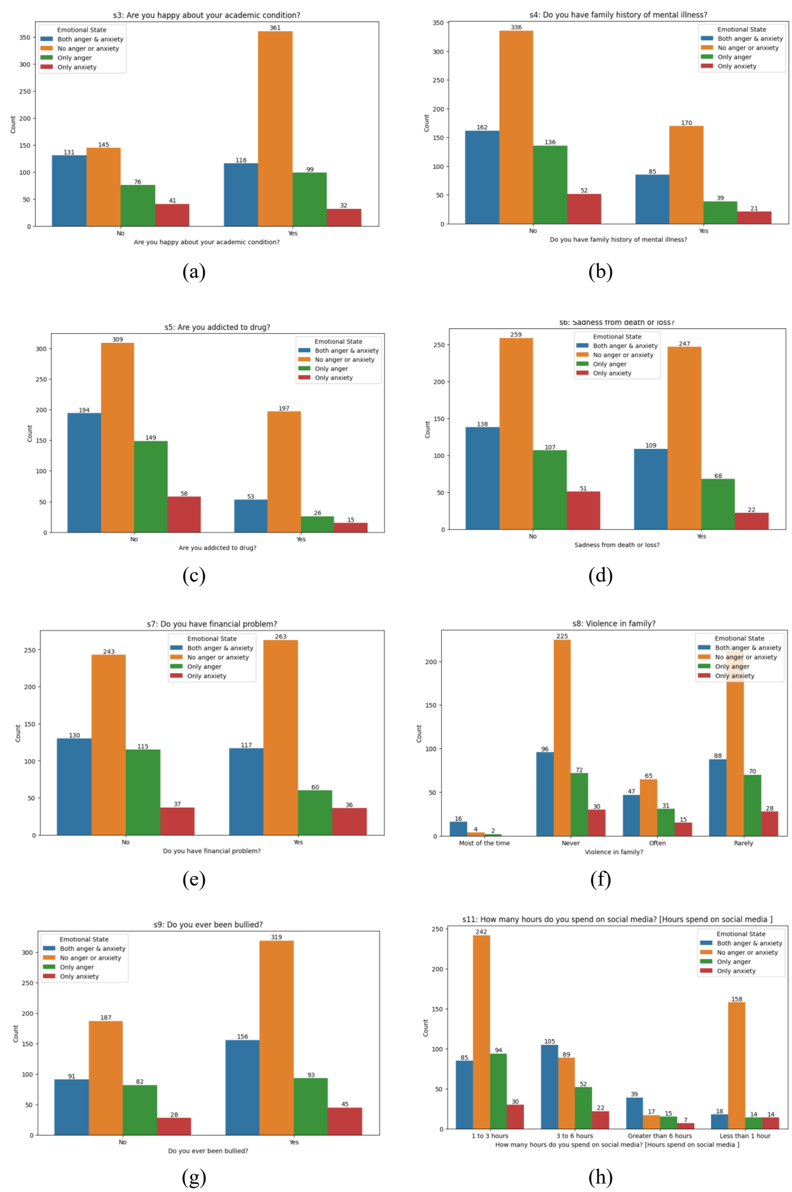


Figure 5(a) shows that a majority of students reported being happy with their academic condition. Among 1,001 students, 608 expressed happiness. Within the group of happy students, 361 were free from anger/anxiety, 116 experienced both anger and anxiety, 99 had only anger, and 32 had only anxiety. Among the 393 students were not happy with their academic condition, 131 experienced both anger and anxiety, 145 were free from anger/anxiety, 76 had only anger, and 41 had only anxiety. Figure 5(b) shows that a majority of students reported no family history of mental illness. Among 1,001 students, 315 have a family history of mental illness. Within the group with a family history of mental illness, 170 were free from anger/anxiety, 85 experienced both anger and anxiety, 39 had only anger, and 21 had only anxiety. Among the 686 students who had no family history of mental illness, 162 experienced both anger and anxiety, 336 were free from anger/anxiety, 136 had only anger, and 52 had only anxiety.

Figure 5(c) shows that a majority of students reported they are not addicted to drugs. Among 1,001 students, 291 were addicted to drugs. Within the group of addicted, 53 experienced both anger and anxiety, 197 were free from anger/anxiety, 26 had only anger, and 15 had only anxiety. Among the 710 students who had not been addicted to any drugs, 309 were free from anger/anxiety, 194 experienced both anger and anxiety, 149 had only anger, and 58 had only anxiety. Figure 5(d) shows that 555 students reported no sadness from losing someone. Among 1,001 students, 446 had sadness from death or loss. Within the group sadness, 109 experienced both anger and anxiety, 247 were free from anger/anxiety, 68 had only anger, and 22 had only anxiety. Among the 555 students who had no sadness from death or loss, 259 were free from anger/anxiety, 138 experienced both anger and anxiety, 107 had only anger, and 51 had only anxiety.

Figure 5(e) shows that 525 students reported no financial problems and 476 had financial problems. Within the group of financial problems, 117 experienced both anger and anxiety, 263 were free from anger/anxiety, 60 had only anger, and 36 had only anxiety. Among the 525 students who had no financial problems, 243 were free from anger/anxiety, 130 experienced both anger and anxiety, 115 had only anger, and 37 had only anxiety. Figure 5(f) offers insights into predicting anger and anxiety about the feature of violence in the family. The comparison reveals several key observations. In instances categorised as 'most of the time', a coexistence of anger and anxiety is reported when there is a history of violence in the family. Conversely, in cases where there is no violence, neither anger nor anxiety are present. When examining the category of 'never', it is evident that the absence of violence corresponds to a lower occurrence of both anger and anxiety. However, in instances of violence, both emotions are frequently reported. The categories of 'often' and 'rarely' show mixed distributions, with varying levels of anger and anxiety reported in the presence and absence of violence. The data comparison underscores the need for further analysis, as the prediction of anger and anxiety involves additional factors beyond the feature of violence in the family.

Figure 5(g) shows that being ever been bullied is indeed an important feature for predicting anger and anxiety. Among the total of 1,001 surveyed students, 156 students who reported having been bullied showed a higher incidence of both anger and anxiety. This indicates a strong association between bullying experiences and the coexistence of anger and anxiety. On the other hand, among the students who have never been bullied (reported as 'no'), 91 students experience both anger and anxiety, suggesting that factors other than bullying may also contribute to these emotions. It is worth noting that being bullied is associated with a higher likelihood of experiencing either only anger (82 students) or only anxiety (28 students). In Figure 5(h) the comparison highlights notable differences in the prevalence of anger and anxiety based on social media usage. Students who spend 1 to 3 hours on social media exhibit a higher occurrence of both anger and anxiety (85 instances). Conversely, those who spend less than 1 hour show a lower incidence of anger and anxiety (18 instances). Additionally, spending greater than 6 hours on social media is associated with a moderate occurrence of anger and anxiety (39 instances). The data suggest a potential relationship between increased social media usage and heightened levels of anger and anxiety.

3.2.3 Evaluation of framework

Our study aims to determine anxiety and anger in individuals. We utilised multiclass classification algorithms and visualised our data using histograms. We have selected the suitable ML model for prediction by evaluating the accuracy, precision, recall, specificity, error rate, and f-measure using five algorithms. Higher accuracy and f-measure indicate better performance, ensuring the correct identification of anxious and angry individuals. The comparison of algorithm performance helped determine the most suitable one. Reducing false negative and false positive rates improved the model's effectiveness. Mean absolute percentage error (MAPE) and root mean squared error (RMSE) was calculated for model comparison. We used two datasets, F-21 and F-15, and employed Python libraries such as pandas and scikit-learn for data analysis and classification. The dataset was divided into training and test data, with a 75% and 25% split respectively. SVM, KNN, logistic regression, random forest, and decision tree algorithms were employed for model prediction. We have compared our suitable ML approach (logistic regression) with existing works on SVM-based prediction (Hilbert et al., 2017) and KNN-based prediction (Ahmed et al., 2020). It can be noted that we have implemented both methods based on our experimental settings.

3.2.4 Comparison of different machine learning algorithms

Figure 6(a) shows the confusion matrix for logistic regression. We can see that, in the first row, there were 98 true positives (TP), 0 true negatives, 2 false positives (FP), and 0 false negatives (FN). In the second row, there were 0 FPs, 39 TNs, 0 TPs, and 0 FNs. In the third row, there were 2 FPs, 0 FNs, 11 TPs, and 1 TN. In the fourth row, there were 0 FPs, 0 FNs, 0 TPs, and 48 TNs. The logistic regression model achieved an accuracy of 97%, indicating that it correctly classified 97% of the samples. The precision score of 95% suggests a high proportion of true positive predictions relative to the total number of positive predictions. The recall score of 94% indicates that the model successfully identified 94% of the actual positive instances. The error rate was 2%, reflecting a low rate of misclassification by the model. The F-measure, which combines precision and recall, resulted in a value of 94%. The MAPE value of 2% suggests a low percentage of absolute errors in the model's predictions.

Figure 6(b) highlights the confusion matrix for the SVM classifier. We can see that, in the first row, there were 100 true positives (TP), 0 true negatives, 0 false positives (FP), and 0 false negatives (FN). In the second row, there were 3 FPs, 36 TNs, 0 TPs, and 0 FNs. In the third row, there were 4 FPs, 0 FNs, 9 TPs, and 1 TN. In the fourth row, there were 0 FPs, 1 FN, 0 TP, and 47 TNs. The SVM model achieved an accuracy of 95%, indicating that it correctly classified 95% of the samples. The precision score of 94% suggests a high proportion of true positive predictions relative to the total number of positive predictions. The recall score of 87% indicates that the model successfully identified 87% of the actual positive instances. The error rate was 5%, reflecting a moderate rate of misclassification by the model. The F-measure, which combines precision and recall, resulted in a value of 90%. The MAPE value of 9% suggests a moderate percentage of absolute errors in the model's predictions.

Figure 6 (a) Confusion matrix for logistic regression for 23 features (b) Confusion matrix for SVM for 23 features (Hilbert et al., 2017) (c) Confusion matrix for KNN for 23 features (Ahmed et al., 2020) (d) Confusion matrix for decision tree for 23 features (e) Confusion matrix for random forest for 23 features (see online version for colours)

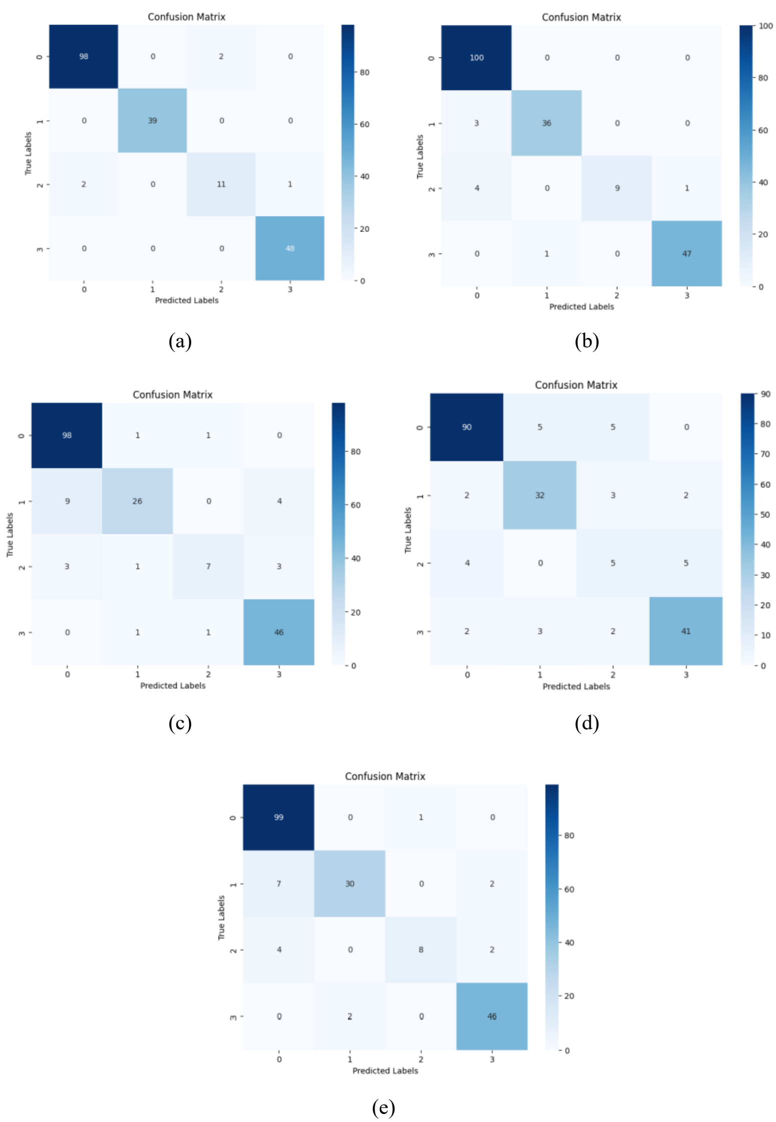


Figure 6(c) depicts the confusion matrix for the KNN classifier. We can see that, in the first row, there were 98 true positives (TP), 1 true negative, 1 false positive (FP), and 0 false negative (FN). In the second row, there were 9 FPs, 26 TNs, 0 TPs, and 4 FNs. In the third row, there were 3 FPs, 1 FN, 7 TP, and 3 TNs. In the fourth row, there were 0 FP, 1 FN, 1 TP, and 46 TNs.

The KNN algorithm exhibited its prowess in classification, achieving an accuracy of 89% and demonstrating an impressive precision of 86%. It excelled incorrectly labelling a significant proportion of the samples, with a relatively low rate of false positives. While its recall score of 80% indicated the potential for improvement in capturing all positive instances, the algorithm struck a harmonious balance with an F-measure of 83%.

Figure 6(d) gives the confusion matrix for the decision tree classifier. We can see that, in the first row, there were 90 true positives (TP), 5 true negatives, 5 false positives (FP), and 0 false negatives (FN). In the second row, there were 2 FPs, 32 TNs, 3 TPs, and 2 FNs. In the third row, there were 4 FPs, 0 FNs, 5 TPs, and 5 TNs. In the fourth row, there were 2 FPs, 3 FNs, 2 TPs, and 41 TNs. The decision tree model demonstrated its capabilities with an accuracy of 87% but showcased relatively lower precision at 75% and recall at 74%. The error rate stood at 13%, while the F-measure reached 75% and the MAPE was 20%.

Figure 6(e) provides the confusion matrix for the random forest classifier. We can see that, in the first row, there were 99 true positives (TP), 0 true negative, 1 false positive (FP), and 0 false negative (FN). In the second row, there were 7 FPs, 30 TNs, 0 TP, and 2 FNs. In the third row, there were 4 FPs, 0 FN, 8 TPs, and 2 TNs. In the fourth row, there were 0 FP, 2 FNs, 0 TP, and 46 TNs.

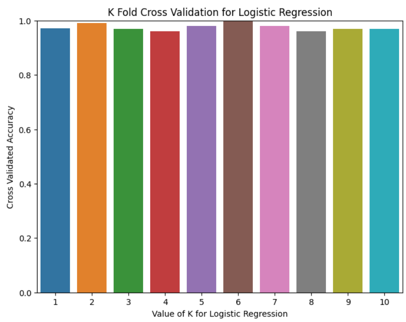
The random forest model yielded impressive results, achieving an accuracy of 91%. With a precision score of 91% and a recall score of 82%, the model demonstrated strong performance in correctly identifying positive instances while minimising false positives. The error rate was reduced to 8%, indicating a lower rate of misclassification compared to the decision tree model. The F-measure of 86% highlighted the model's balanced performance between precision and recall. The MAPE value of 17% indicated a relatively low percentage of absolute errors.

Among the algorithms compared in Table 2, the maximum accuracy of 97% was achieved by the logistic regression model, while the minimum accuracy of 87% was obtained by the decision tree model. Based on this analysis, it can be concluded that the logistic regression model performed the best in terms of accuracy among the evaluated algorithms. With a higher accuracy rate, it demonstrated superior classification capabilities, correctly identifying a larger proportion of samples compared to the other models.

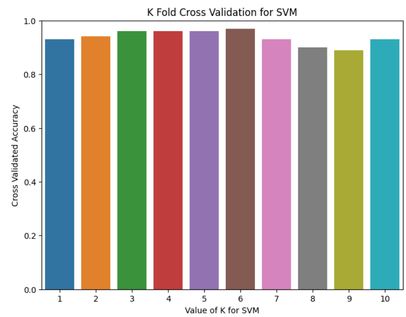
Table 2 Comparison among algorithms for 23 features (without cross validation)

Classifier	*Accuracy (%)*	*Precision (%)*	*Recall (%)*	*Error rate (%)*	*F-measure (%)*	*MAPE (%)*
Logistic regression	97%	95%	94%	2%	94%	2%
SVM (Hilbert et al., 2017)	95%	94%	87%	5%	90%	9%
KNN (Ahmed et al., 2020)	89%	86%	80%	10%	83%	16%
Decision tree	87%	75%	74%	13%	75%	20%
Random forest	91%	91%	82%	8%	86%	17%

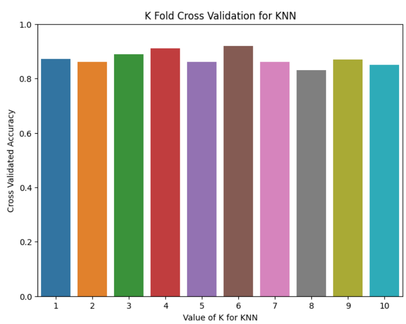
Figure 7 (a) Cross-validated accuracy for logistic regression (b) Cross-validated accuracy for SVM (Hilbert et al., 2017) (c) Cross-validated accuracy for KNN (Ahmed et al., 2020) (d) Cross-validated accuracy for decision tree (e) Cross-validated accuracy for random forest (see online version for colours)



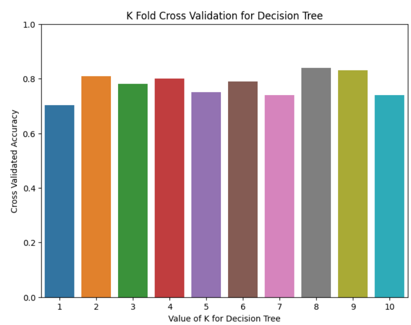
(a)



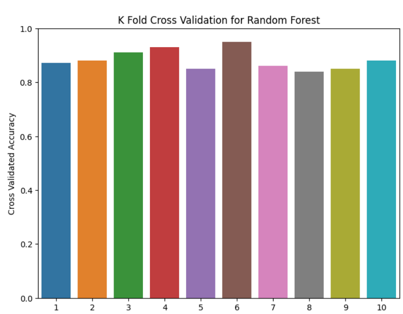
(b)



(c)



(d)



(e)

Figure 7 highlights the different models accuracy with K-fold cross-validation (k = 1 to 10). If K = 5, then we are dividing the provided dataset into five folds and performing a train and test on each fold. One fold is used for testing purposes during each run, with the remaining folds being used for training purposes and subsequent iterations (Analytics Vidhya, 2022). Figure 7(a) shows that the cross-validated accuracy of the logistic regression model is consistently high, with a range of 96.0% to 99.5% over 10 folds.

The highest accuracy of 99.5% is obtained when K = 6. This consistency demonstrates the model's strong generalisation abilities and suggests how resilient it is to various data subsets. The lack of extreme outliers in the accuracy values demonstrates the stability and dependability of the model in producing precise predictions under a variety of cross-validation conditions. Figure 7(b) indicates that SVM achieved cross-validated accuracy scores between 89.0% and 97.0% over ten folds. They achieved the second highest cross-validated accuracy of 97% when K = 6. The model continues to maintain a high degree of accuracy, demonstrating its capacity to manage a variety of data subsets and produce accurate predictions. Figure 7(c) demonstrates that the cross-validated accuracy of KNN's performance varies, ranging from 83.0% to 92.0%. The achieved fourth highest validated accuracy of 92%, when k = 6. The model exhibits flexibility in response to various data attributes, albeit with some sensitivity to particular folds. In spite of this fluctuation, KNN generally maintains an accuracy level that is moderate.

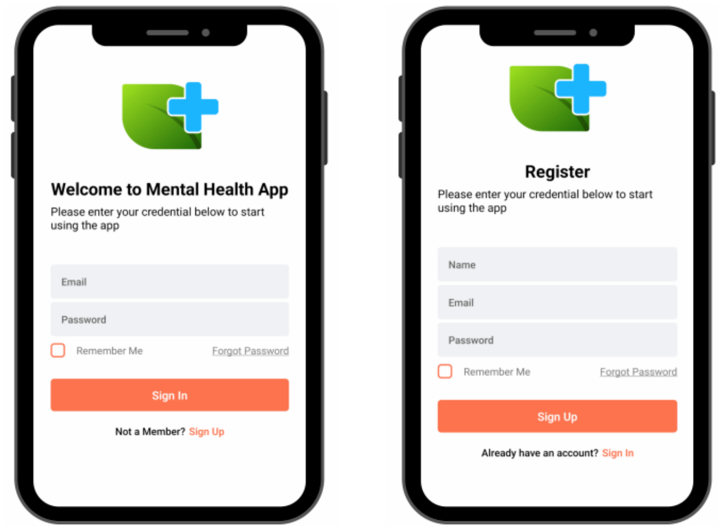
Figure 7(d) informs us that the accuracy of the decision tree model varies tenfold, from 70.2% to 84.0%. Even though the accuracy of certain folds is lower than others, the model performs admirably, identifying underlying patterns in the data and making accurate predictions. The achieved fifth highest validated accuracy of 84%, when k = 8. Figure 7(e) visualises that the random forest consistently performs well, achieving validated accuracy scores between 85.0% and 95.0%. The achieved the third highest validated accuracy of 95%, when k = 6. The ensemble model's robustness is evident, maintaining high accuracy across diverse folds. This highlights its effectiveness in handling complex relationships within the dataset and making accurate predictions. From the aforementioned accuracy results with and without cross-validation, it can be said that for our dataset, the logistic regression model is superior to other ML models. Therefore, the logistic regression model can be used for anger and anxiety prediction purposes.

4 Mobile application development

This section will discuss the features of our proposed mobile application. The mental health app's user-centric mobile interface aims to promote mental well-being and provide valuable resources. Users are welcomed with a clean login page [see Figure 8(a)] to access the app's features, while new users can create an account through the register feature in Figure 8(b). The login page [see Figure 8(a)] offers a straightforward design, featuring essential fields for email and password entry. A 'forgot password?' link facilitates easy password recovery. The register page presents a comprehensive form, requesting users' names, email addresses, and secure passwords. Upon completing the form, users can tap 'register' to gain access to the app's full features.

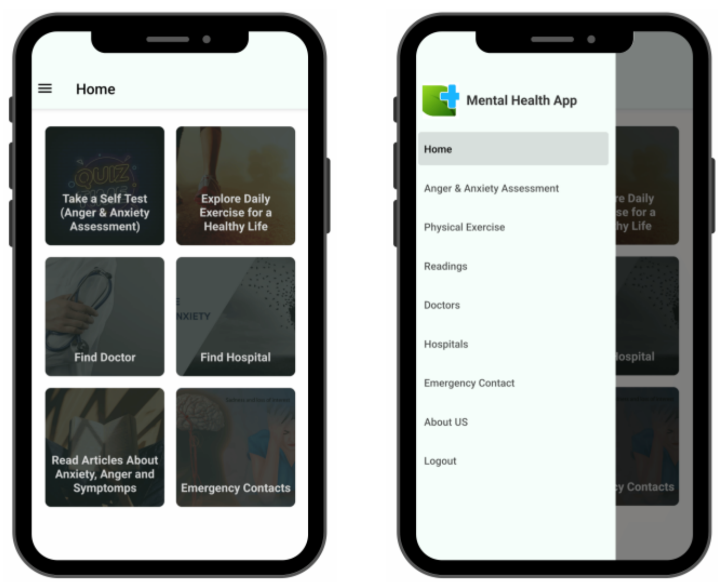
After login, the user can access the home page in Figure 9(b). The navigation menu is shown in Figure 9(a). The home interface serves as the central hub and the starting point for users after they log in to our mental health app. Designed with user convenience in mind, the home interface in Figure 9(b) offers easy navigation to various sections and features within the app. Users can navigate through the navigation bar as well [see Figure 9(a)]. Upon logging in, users are welcomed to the home interface, where they can seamlessly access different sections, including the anger and anxiety assessment feature, physical exercise suggestion page, doctors search, emergency contact, article reading, and hospital search page.

Figure 8 (a) Login and (b) register process (see online version for colours)



(a) (b)

Figure 9 (a) Home and (b) navigation interface (see online version for colours)



(a) (b)

The anger and anxiety assessment feature of our mental healthcare app employs advanced machine learning techniques to provide personalised insights to users in Figure 10. Utilising Python and hosting the model on AWS servers, we offer a seamless and efficient experience. The assessment user interface presents users with 23 carefully crafted questions, each requiring them to select one option that best represents their feelings. Upon submission of the answer regarding asked questions, our powerful machine learning classifier model processes the 23 features and predicts one of four classes: neither anger nor anxiety, anger only, anxiety only, or both anger and anxiety. The anger and anxiety prediction outcome is shown in Figure 11. The predictive power of our model is driven by a sophisticated algorithm that analyses user responses to assess their mental state accurately. This helps users gain valuable self-awareness and understand their emotional well-being.

Figure 10 Anger and anxiety assessment (see online version for colours)

Figure 11 Anger and anxiety assessment prediction outcome (see online version for colours)

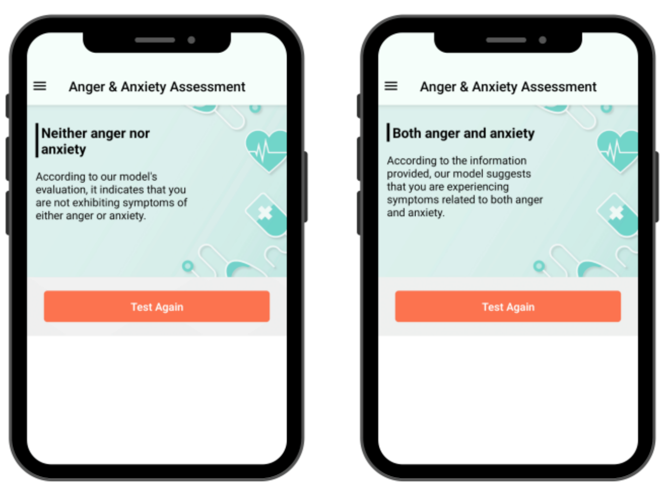
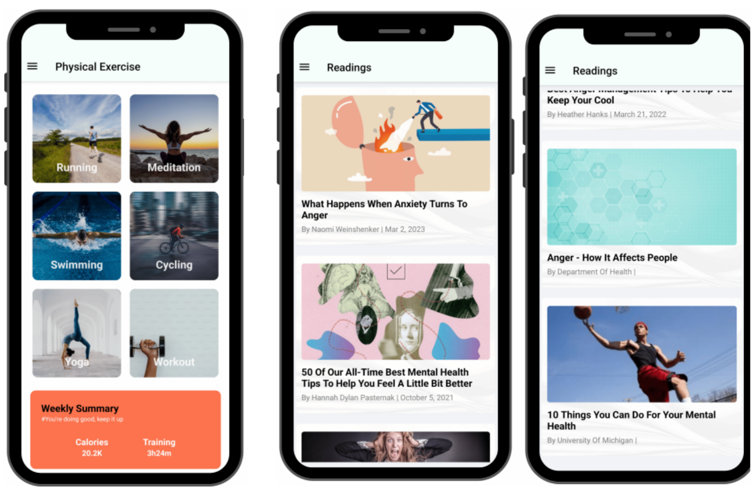


Table 3 Applying different test cases to evaluate the prediction model

Feature	*Test-1*	*Test-2*	*Test-3*	*Test-4*
Q1: Are you happy about your academic condition?	Yes	No	No	No
Q2: Do you have family history of mental illness?	No	Yes	No	No
Q3: Are you addicted to drugs?	No	Yes	Yes	No
Q4: Do you have sadness from death or loss?	No	Yes	No	No
Q5: Do you have financial problems?	No	Yes	Yes	Yes
Q6: Violence in family?	Never	Most of the time	Rarely	Often
Q7: Hare you ever boon bullied?	No	Yes	Yes	Yes
Q8: Does bullying still affect you?	Never	Most of the time	Rarely	Often
Q9: How many hours do you spend on social media?	Less than 1 hour	Greater than 6 hours	1 to 3 hours	1 to 3 hours
Q10: I get angry fast but calm down fast too	not like me at all	Very like me	Like me	Like me
Q11: When frustrated, I let my irritation show	Not like me at all	Very like me	Like me	Not like me at all
Q12: I sometimes feel like a powder keg ready to explode	Not like me at all	Very like me	Not like me	Not like me at all
Q13: I get angry easily.	Not like me at all	Very like me	Neutral	Not like me at all
Q14: Some of my friends think I'm a hothead	Not like me at all	Very like me	Not like me	Not like me
Q15: Sometimes I get angry for no reason	Not like me at all	Very like me	Not like me	Not like me
Q16: I have trouble controlling my temper	Not like me at all	Very like me	Not like me	Not like me
Q17: Over the last two weeks, how frequently have you experienced trouble relaxing?	not at all	Nearly everyday	Several days	Over half of the days
Q18: Feeling nervous, anxious, or on edge	Not at all	Nearly everyday	Several days	Nearly everyday
Q19: Not being able to stop or control worrying	Not at all	Nearly everyday	Several days	Nearly everyday
Q20: Worrying too much about different things	Not at all	Nearly everyday	Not at all	Nearly everyday
Q21: Being so restless that it's hard to still	Not at all	Nearly everyday	Not at all	Not at all
Q22: Becoming easily annoyed or irritable	Not at all	Nearly everyday	Not at all	Over half of the days
Q23: Feeling afraid as if something awful might happen	Not at all	Nearly everyday	Not at all	Nearly everyday
Model prediction	Neither anger nor anxiety	Both anger and anxiety	Anger only	Anxiety only

Table 3 presents the outcomes of different test cases conducted to assess the performance of our model. Each row in the table corresponds to a specific feature, and the columns represent different test cases. The values within the table cells represent the answers provided for each feature in each test case. The final row of the table represents the model's predictions for each test case, classifying them into four categories: neither anger nor anxiety, both anger and anxiety, anger only, or anxiety only. Figure 12(a) shows the physical exercise feature of our app. The physical exercise section of our mental health app is dedicated to promoting overall well-being by providing users with a diverse range of exercise options to reduce mental health distress and maintain a healthy lifestyle. Within this section, users have access to a variety of exercises, including running, meditation, swimming, cycling, yoga, and workouts. Each exercise option is accompanied by informative articles that explore the importance of the exercise in the context of mental health. The article reading interface is shown in Figure 12(b). These articles highlight the numerous benefits of incorporating the specific exercise into one's routine, shedding light on its positive impact on mental clarity, emotional regulation, and stress reduction.

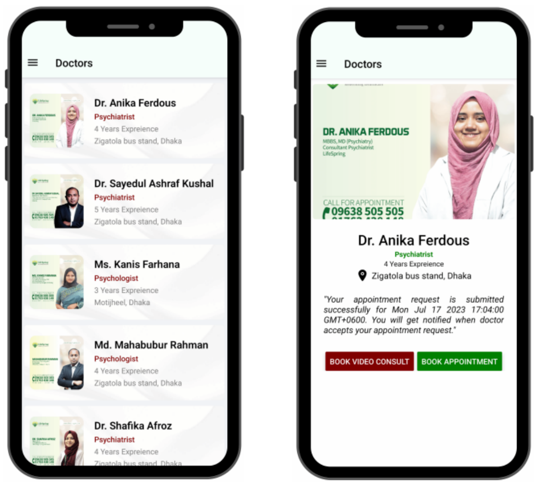
Figure 12 (a) Physical exercise interface (b) Article readings interface (see online version for colours)



(a) (b)

The find doctors interface [see Figures 13(a) and 13(b)] in our mental health app provides users with a convenient way to access a comprehensive list of licensed psychologists and psychiatrists. It serves as a valuable resource for individuals seeking professional support and counselling to address their mental health concerns. Users can browse through the profiles of different doctors, which include their credentials, areas of expertise, and patient reviews. Upon selecting a specific doctor, users are directed to the doctor's details interface, where they can initiate the process of booking an appointment [see Figure 13(a)]. Through this interface, users can request a preferred date and time for the appointment, along with any additional notes or specific requirements they may have. Once the appointment request is submitted, the doctor will be notified and can review the details provided by the user. The doctor has the option to accept or deny the appointment request based on their availability. In the case of acceptance, the user will receive a confirmation, and the appointment will be scheduled [see Figure 13(b)]. This interface aims to streamline the process of finding and booking appointments with mental health professionals, ensuring that users can easily access the support they need to improve their mental well-being.

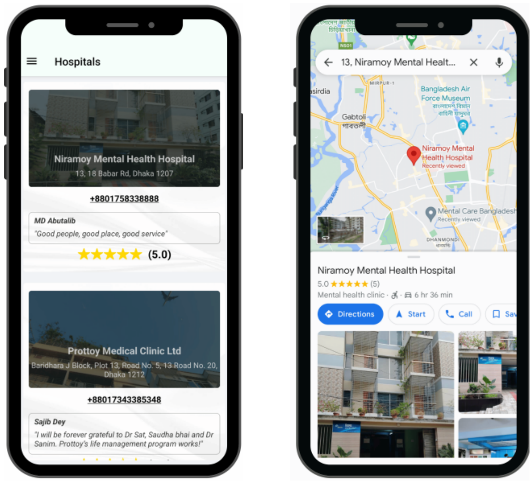
Figure 13 (a) Find doctors and (b) appointment booking interface (see online version for colours)



(a) (b)

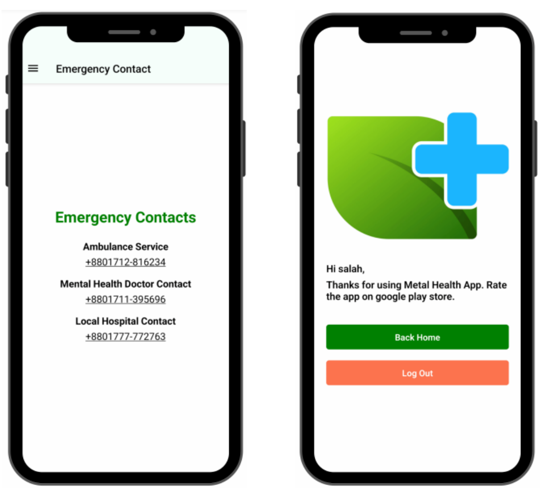
The find hospital interface of our app is shown in Figure 14. This interface serves as a valuable tool for users to locate and access nearby hospitals that offer mental health services. It provides a user-friendly platform where individuals can find essential information about various hospitals, enabling them to make informed decisions about their healthcare needs. The interface presents a list of hospitals, each displaying crucial details such as the hospital name, address, contact number, reviews, and ratings [see Figure 14(a)]. This information allows users to assess the quality and suitability of the hospital based on the experiences of other patients. Upon selecting a specific hospital from the list, users have the option to view the hospital's location on Google Maps [see Figure 14(b)]. This feature simplifies navigation and helps users easily locate the hospital, ensuring a smooth and hassle-free visit. The find hospital interface aims to empower users by providing them with comprehensive information about mental health care facilities in their area. By offering hospital reviews, ratings, and location services, the app ensures that users can access the most suitable and reliable healthcare resources to address their mental health needs effectively. The emergency contact interface in our mental health app is shown in Figure 15(a). In times of urgent need or crisis, this emergency contact interface serves as a quick and accessible resource to contact essential emergency services. The interface displays three primary contact numbers such as ambulance service, doctor contact number, and hospital contact number. The logout interface of our mental health app is depicted in Figure 15(b). This logout interface offers a simple and user-friendly a way for users to conclude their session and provide feedback. Upon choosing to log out, users are presented with two options: 'back home' and 'logout'. Alongside these buttons, a brief message invites users to share their valuable feedback by rating the app on the Google Play Store.

Figure 14 Find hospitals interface (see online version for colours)



(a) (b)

Figure 15 (a) Emergency contact and (b) logout and feedback interface (see online version for colours)



(a) (b)

5 Evaluation results

We have collected feedback from thirty-two users via a survey regarding the proposed app feature and its usefulness. This feedback helped us to examine the functionality of our mobile application. The users responded to our survey and evaluated the app on several issues. We chose the following questions to respond to on a scale of 1–5 to assess how effectively the app performs:

1 rate how user-friendly and visually appealing the app's interface is (user interface and design)

2 rate the necessity of the app's mental health articles and resources in terms of their usefulness and relevance (helpful resources)

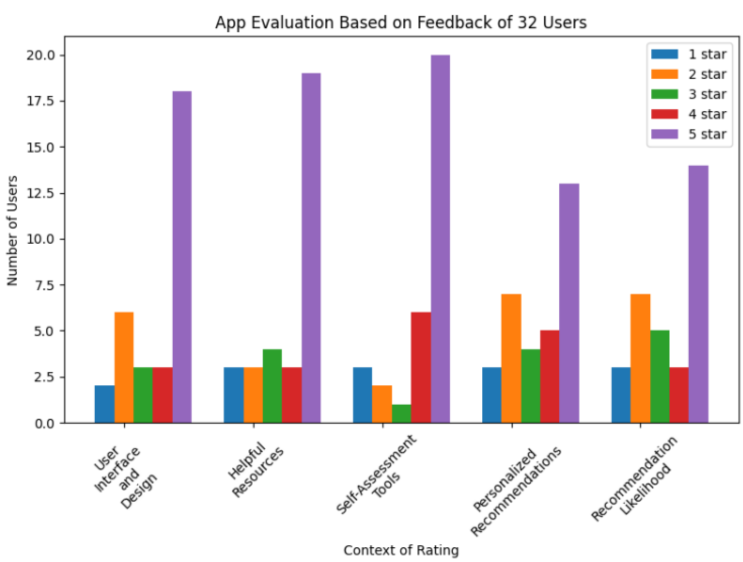
3 evaluate the effectiveness of the app's self-assessment tools in understanding your mental well-being (self-assessment tools)

4 rate how satisfied you are with the app's ability to provide personalised recommendations for improving mental health (personalised recommendation)

5 indicates how likely you are to recommend this app to others seeking mental health support (recommendation likelihood).

From Figure 16, we can see that the majority of the users selected a 5-star rating (suitable and useful). Whereas, the number of users with low ratings (1 star) is very low. According to Figure 16 user based on review analysis, it can be seen that the majority of users liked our application features and hinted at the usefulness of the proposed application.

Figure 16 Mobile application evaluation result (see online version for colours)



5.1 Comprehensive overview of the proposed system procedure, materials, and analysis

In this subsection, we will deliver a comprehensive overview of the proposed machine learning-based anger and anxiety disorder prediction procedure.

Figure 17 Initial dataset upload procedures (see online version for colours)

```
[ ] from google.colab import drive
    drive.mount('/content/drive')
[ ] !pip install --upgrade scikit-learn
    import sklearn
    sklearn.__version__
[ ] import pandas as pd
    from sklearn.preprocessing import LabelEncoder
[ ] df = pd.read_excel('/content/drive/MyDrive/mental health assessment model/Mental Health Assessment Dataset.xlsx')
[ ] df.shape
    (1001, 31)
[ ] df.head()
```

Academic Year [Academic Year]	Department [Department]	Are you happy about your academic condition?	Do you have family history of mental illness?	Are you addicted to drug?	Sadness from death or loss?	Do you have financial problem?	...	Over the last two weeks, how frequently have you experienced trouble relaxing? [Row 1]	Feeling nervous, anxious, or on edge [Row 1]	Not being able to stop or control worrying [Row 1]	Worrying too much about different things [Row 1]	Being so restless that it's hard to still [Row 1]	Becoming easily annoyed or irritable [Row 1]	Feeling afraid as if something awful might happen [Row 1]	Do you think you have anger problem?	Do you think you have anxiety?	Mobile number (if you like to win 50 tk mobile recharge)
4th year	Engineering	Yes	No	Yes	Yes	Yes	...	2	1	1	3	4	4	3	Yes	Yes	NaN

Figure 18 Cleaning data procedures (see online version for colours)

```
[ ] df.drop(df.columns[[0, 1, 30]], axis=1, inplace=True)
    for i, col in enumerate(df.columns):
      print(i, col)

    0 Age [Age]
    1 Academic Year [Academic Year]
    2 Department [Department]
    3 Are you happy about your academic condition?
    4 Do you have family history of mental illness?
    5 Are you addicted to drug?
    6 Sadness from death or loss?
    7 Do you have financial problem?
    8 Violence in family?
    9 Do you ever been bullied?
    10 If yes does it still effect you?(ANSWER ONLY IF YOU ARE BULLIED) [Row 1]
    11 How many hours do you spend on social media? [Hours spend on social media ]
    12 I get angry fast but calm down fast too. [Row 1]
    13 When frustrated, I let my irritation show. [Row 1]
    14 I sometimes feel like a powder keg ready to explode. [Row 1]
    15 I get angry easily. [Row 1]
    16 Some of my friends think I'm a hothead. [Row 1]
    17 Sometimes I get angry for no reason. [Row 1]
    18 I have trouble controlling my temper. [Row 1]
    19 Over the last two weeks, how frequently have you experienced trouble relaxing? [Row 1]
    20 Feeling nervous, anxious, or on edge [Row 1]
    21 Not being able to stop or control worrying [Row 1]
    22 Worrying too much about different things [Row 1]
    23 Being so restless that it's hard to still [Row 1]
    24 Becoming easily annoyed or irritable [Row 1]
    25 Feeling afraid as if something awful might happen [Row 1]
    26 Do you think you have anger problem?
    27 Do you think you have anxiety?
```

Figure 19 Column mapping procedure (see online version for colours)

```
col_map = {}
original_df = df
columns = df.columns
df.columns = [f's{i}' for i in range(len(df.columns))]
for i, (col, prev) in enumerate(zip(df.columns, columns)):
  col_map[col] = prev
for key in col_map.keys():
  print(key, col_map[key])
```

```
s0 Age [Age]
s1 Academic Year [Academic Year]
s2 Department [Department]
s3 Are you happy about your academic condition?
s4 Do you have family history of mental illness?
s5 Are you addicted to drug?
s6 Sadness from death or loss?
s7 Do you have financial problem?
s8 Violence in family?
s9 Do you ever been bullied?
s10 If yes does it still effect you?(ANSWER ONLY IF YOU ARE BULLIED) [Row 1]
s11 How many hours do you spend on social media? [Hours spend on social media ]
s12 I get angry fast but calm down fast too. [Row 1]
s13 When frustrated, I let my irritation show. [Row 1]
s14 I sometimes feel like a powder keg ready to explode. [Row 1]
s15 I get angry easily. [Row 1]
s16 Some of my friends think I'm a hothead. [Row 1]
s17 Sometimes I get angry for no reason. [Row 1]
s18 I have trouble controlling my temper. [Row 1]
s19 Over the last two weeks, how frequently have you experienced trouble relaxing? [Row 1]
s20 Feeling nervous, anxious, or on edge [Row 1]
s21 Not being able to stop or control worrying [Row 1]
s22 Worrying too much about different things [Row 1]
s23 Being so restless that it's hard to still [Row 1]
s24 Becoming easily annoyed or irritable [Row 1]
s25 Feeling afraid as if something awful might happen [Row 1]
s26 Do you think you have anger problem?
s27 Do you think you have anxiety?
```

Figure 20 Labelling anger and anxiety dataset (see online version for colours)

Anger Rating The Buss-Perry Aggression Questionnaire (BPAQ) consists of 29 self-administered items rated on a 5-point Likert Anger scale
(items 15-21) subscale. The score is the sum or the ratings (1 to 5) for its items. **Range 7 - 35, Midpoint: 21**

```
[ ] for i in range(len(df)):
      df.loc[i, 'anger_score'] = df.iloc[i, 12:19].sum()
```

Anxiety Scoring

Anxiety Score - Range: 0 - 21, Scores of 5, 10, and 15 are taken as the cut-off points for mild, moderate and severe anxiety

```
[ ] # Converting from 1-4 to 0-3 score
    for i in range(len(df)):
      for j in range(19, 26):
        df.iloc[i, j] = df.iloc[i, j]-1
```

```
[ ] for i in range(len(df)):
      df.loc[i, 'anxiety_score'] = df.iloc[i, 19:26].sum()
```

Labeling anxiety & anger

```
[ ] def labeling(row):
      # 3 - Both anger & anxiety
      # 2 - Only anxiety
      # 1 - Only anger
      # 0 - Not anger & anxiety

      anger_threshold = 17
      anxiety_threshold = 10

      if row['anger_score'] > anger_threshold and row['anxiety_score'] > anxiety_thres
        return 3
      elif row['anxiety_score'] > anxiety_threshold:
        return 2
      elif row['anger_score'] > anger_threshold:
        return 1
      else:
        return 0

    df['y'] = df.apply(labeling, axis=1)
```

```
[ ] # Create a dictionary to map values in column 'y' to emotional states
    emotional_state_mapping = {
        3: 'Both anger & anxiety',
        2: 'Only anxiety',
        1: 'Only anger',
        0: 'No anger or anxiety'
    }

    # Add the 'Emotional State' column based on the mapping
    df['Emotional State'] = df['y'].map(emotional_state_mapping)
```

```
[ ] main_df = df.copy()
    df.head()
```

	s0	s1	s2	s3	s4	s5	s6	s7	s8	s9	...	s22	s23	s24	s25	s26	s27	anger_score	anxiety_score	y	Emotional State
0	21-23	4th year	Engineering	Yes	No	Yes	Yes	Yes	Yes	Yes	...	2	3	3	2	Yes	Yes	23.0	11.0	3	Both anger & anxiety
1	24-26	4th year	Engineering	No	No	No	No	Yes	No	Yes	...	0	0	0	0	No	No	15.0	4.0	0	No anger or anxiety
2	18-20	4th year	Engineering	Yes	No	Yes	No	No	No	Yes	...	0	0	0	0	No	No	14.0	2.0	0	No anger or anxiety
3	21-23	4th year	Engineering	No	No	No	No	No	No	Yes	...	0	0	0	0	No	Yes	14.0	0.0	0	No anger or anxiety
4	21-23	4th year	Engineering	No	No	No	No	Yes	Never	Yes	...	3	2	1	2	Yes	Yes	20.0	10.0	1	Only anger

5 rows × 32 columns

Figure 21 Data visualisation procedures (see online version for colours)

```
[ ] import matplotlib.pyplot as plt
    import seaborn as sns

    def data_visual(col):
        # Group the DataFrame by col and 'Emotional State' and count the occurrences
        grouped_df = df.groupby([col, 'Emotional State']).size().unstack()

        # Reset the index to convert the columns into categorical variables
        grouped_df = grouped_df.reset_index()

        # Melt the DataFrame to convert it into a long format
        melted_df = pd.melt(grouped_df, id_vars=col, value_vars=grouped_df.columns[1:], var_name='Emotional State')

        # Plot bar plot using seaborn
        plt.figure(figsize=(10, 6))
        ax = sns.barplot(x=col, y='value', hue='Emotional State', data=melted_df)

        # Set labels and title
        plt.xlabel(col_map[col])
        plt.ylabel('Count')
        plt.title('{}: {}'.format(col, col_map[col]))

        # Add total count numbers on top of each bar
        for p in ax.patches:
            ax.annotate(format(p.get_height(), '.0f'), (p.get_x() + p.get_width() / 2, p.get_height()), ha='center', va='bottom')

        # Show the plot
        plt.show()

    data_visual('s3')
```

Figure 22 Label encoding procedures (see online version for colours)

Label Encoding

```
[ ] # Drop unnecessary columns
    df = df.drop(columns=['anger_score',  'anxiety_score'])
    df = df.drop(columns = ['s0', 's1', 's2'])
```

```
[ ] le = LabelEncoder()
    for col in df.columns[:-1]:
      df[col] = le.fit_transform(df[col].astype(str))

      # Print the label encoding mapping
      print("{} Mapping:".format(col))
      for label, encoded_label in zip(le.classes_, le.transform(le.classes_)):
          print(f"{label} -> {encoded_label}")

    s3 Mapping:
    No -> 0
    Yes -> 1
    s4 Mapping:
    No -> 0
```

```
[ ] main_df = df.copy()
    main_df.head()
```

	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	...	s20	s21	s22	s23	s24	s25	s26	s27	y	Emotional State
0	1	0	1	1	1	3	1	3	0	3	...	0	0	2	3	3	2	1	1	3	Both anger & anxiety
1	0	0	0	0	1	1	1	3	1	1	...	2	2	0	0	0	0	0	0	0	No anger or anxiety
2	1	0	1	0	0	1	1	3	1	1	...	0	0	0	0	0	0	0	0	0	No anger or anxiety
3	0	0	0	0	0	1	1	3	2	2	...	0	0	0	0	0	0	0	1	0	No anger or anxiety
4	0	0	0	0	1	1	1	3	0	3	...	1	0	3	2	1	2	1	1	1	Only anger

5 rows × 27 columns

Drop questionnaries columns

```
[ ] df = df.drop(df.iloc[:, 9:23], axis=1)
```

```
[ ] df.head()
```

Figure 23 Feature engineering procedures (see online version for colours)

We have two dataset. F-9 and F-14, Combinning both dataset features we get 23 features. Based on 23 features our new models.

```
main_df.head()
```

	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	...	s20	s21	s22	s23	s24	s25	s26	s27	y	Emotional State
0	1	0	1	1	1	3	1	3	0	3	...	0	0	2	3	3	2	1	1	3	Both anger & anxiety
1	0	0	0	0	1	1	1	3	1	1	...	2	2	0	0	0	0	0	0	0	No anger or anxiety
2	1	0	1	0	0	1	1	3	1	1	...	0	0	0	0	0	0	0	0	0	No anger or anxiety
3	0	0	0	0	0	1	1	3	2	2	...	0	0	0	0	0	0	0	1	0	No anger or anxiety
4	0	0	0	0	1	1	1	3	0	3	...	1	0	3	2	1	2	1	1	1	Only anger

5 rows × 27 columns

```
[ ] from sklearn.model_selection import train_test_split

    df = main_df.copy()
    df = df.drop(['s26', 's27'], axis=1)
```

```
[ ] # Correlation matrix

    corr_matrix = df.corr().round(2)

    plt.figure(figsize=(14, 10))
    plot = sns.heatmap(corr_matrix, annot = True)
```

Figure 24 Model implementation procedures (see online version for colours)

Model Implementation

1. KNN

```
[ ] from sklearn.neighbors import KNeighborsClassifier
    from sklearn.metrics import accuracy_score
    import matplotlib.pyplot as plt
    import seaborn as sns
```

```
knn = KNeighborsClassifier()

# X_train = X_train.drop(["s4", "s7"], axis=1)
knn.fit(X_train, y_train)

print('Test score:', knn.score(X_test, y_test))
```

2. SVM

```
[ ] from sklearn.svm import SVC
    from sklearn.multiclass import OneVsOneClassifier
```

```
[ ] svm_model = OneVsOneClassifier(SVC())

    svm_model.fit(X_train, y_train)

    # print('Train score:', svm_model.score(X_train, y_train))
    print('Test score:', svm_model.score(X_test, y_test))

    Test score: 0.5870646766169154
```

3. Logistic Regression

```
from sklearn.linear_model import LogisticRegression

from sklearn.linear_model import LogisticRegression
log_model=LogisticRegression(max_iter=3000)
log_model.fit(X_train,y_train)

print('Score', log_model.score(X_test, y_test))
```

4. Decision Tree

```
[ ] from sklearn import tree

    tree_model = tree.DecisionTreeClassifier()
    tree_model = tree_model.fit(X_train, y_train)

    # print('Train score:', tree_model.score(X_train, y_train))
    print('Test score:', tree_model.score(X_test, y_test))
```

Figure 25 Evaluating model (see online version for colours)

```
[ ] df.to_csv('procced.csv')
```

```
[ ] X_train, X_test, y_train, y_test = train_test_split(df.iloc[:, :-2], df['y'], test_size=0.2,random_state=42)
    X_train
```

Function for evaluating models

```
[ ] import numpy as np
    import matplotlib.pyplot as plt
    import seaborn as sns
    from sklearn.metrics import confusion_matrix, accuracy_score, precision_score, recall_score

    def plot_confusion_matrix(y_true, y_pred, labels):
        cm = confusion_matrix(y_true, y_pred, labels=labels)
        plt.figure(figsize=(8, 6))
        sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=labels, yticklabels=labels)
        plt.xlabel('Predicted Labels')
        plt.ylabel('True Labels')
        plt.title('Confusion Matrix')
        plt.show()

    def evaluate_model(model, X_test, y_test, labels):
        # Model predictions
        y_pred = model.predict(X_test)

        # Accuracy
        accuracy = accuracy_score(y_test, y_pred)
        print("Accuracy:", accuracy)
```

```
        # Recall
        recall = recall_score(y_test, y_pred, average='macro')
        print("Recall:", recall)

        # # Confusion matrix
        # cm = confusion_matrix(y_test, y_pred)
        # tn, fp, fn, tp = cm.ravel()

        # # Specificity
        # specificity = tn / (tn + fp)
        # print("Specificity:", specificity)

        # Error Rate
        error_rate = 1 - accuracy
        print("Error Rate:", error_rate)

        # F-Measure
        f_measure = 2 * (precision * recall) / (precision + recall)
        print("F-Measure:", f_measure)

        # MAPE (Mean Absolute Percentage Error)
        def mean_absolute_percentage_error(y_true, y_pred):
          y_true, y_pred = np.array(y_true), np.array(y_pred)

          # Check for division by zero or NaN values
          mask = y_true != 0
          y_true = y_true[mask]
          y_pred = y_pred[mask]
```

```
[ ] def apply_model():
      #Logistic Regression
      log_model=LogisticRegression(max_iter=3000)
      log_model.fit(X_train,y_train)

      print('Logistic Regression:')
      evaluate_model(log_model, X_test, y_test, labels=[0, 1, 2, 3])
      joblib.dump(log_model, 'classifier.joblib')

      #SVM
      svm_model = OneVsOneClassifier(SVC())
      svm_model.fit(X_train, y_train)

      print('SVM:')
      evaluate_model(svm_model, X_test, y_test, labels=[0, 1, 2, 3])


      #KNN
      knn = KNeighborsClassifier()
      knn.fit(X_train, y_train)

      print('KNN:')
      # evaluate_model(knn, X_test, y_test, labels=[0, 1, 2, 3])
```

```
      #Decision Tree
      tree_model = tree.DecisionTreeClassifier()
      tree_model = tree_model.fit(X_train, y_train)

      print('Decision Tree:')
      evaluate_model(tree_model, X_test, y_test, labels=[0, 1, 2, 3])


      #Random Forest
      random_forest = RandomForestClassifier()
      random_forest.fit(X_train, y_train)

      print('Random Forest:')
      evaluate_model(random_forest, X_test, y_test, labels=[0, 1, 2, 3])
```

```
[ ] apply_model()

    Logistic Regression:
    Accuracy: 0.9751243781094527
    Precision: 0.9514364207221351
    Recall: 0.9414285714285714
    Error Rate: 0.02487562189054726
    F-Measure: 0.9464060396064825
    MAPE: 0.024752475247524754
```

The first step of our work is dataset collection and dataset preparation. The reader can refer to Section 3.2 for the dataset preparation procedure. Next, we have uploaded the dataset using the Google Colab platform (see Figure 17 for source code details).

Our next step is data preprocessing. In this step, we removed unnecessary columns from the dataset. This step was crucial to enhancing the simplicity and efficiency of our subsequent analyses, ensuring that only relevant and meaningful data was retained (see Figure 18 for the source code details). Following the data cleaning process, we streamlined the dataset further by renaming the remaining columns from S0 to S27. This systematic column mapping was implemented to facilitate easy access to specific features, fostering a more intuitive and organised dataset structure (see Figure 19 for column mapping step source code details). Our next step is labelling the dataset. In order to categorise emotional states, a new column named 'emotional state' was introduced based on predefined thresholds for anger and anxiety. This labelling process involved assigning specific emotional states to each data point, providing a foundation for subsequent analyses and model training (see Figure 20 for dataset labelling details, along with Table 1). Next, we performed data visualisation. Categorical columns within the dataset were visually explored using count plots, with the additional dimension of color representing the 'emotional state'. These visualisations served to illustrate the relationships between different features and the final outcome of emotional states, offering valuable insights into potential patterns or correlations (see Figure 21 for details on the source code and Figure 5 for some data visualisation results). Next, to prepare the dataset for machine learning algorithms, we employed the LabelEncoder from the sklearn.preprocessing module (see Figure 22 for source code details). This process converted categorical data into numerical values, ensuring compatibility with various machine learning models. Our next step is feature engineering. A correlation matrix was constructed to analyse the relationships between different features in the dataset. Subsequently, two sub-datasets, namely S-9 and S-14, were created. S-9 focused on user personal attributes, while S-14 included GAD and BPAQ symptoms. This segregation facilitated a more targeted exploration of specific feature sets (see Figure 23 for source code details). Our next step is model implementation. Several machine learning algorithms, including KNN, SVM, logistic regression, decision trees, and random forests, were implemented on the dataset. The models were trained using two different sets of features (9 features in S-9 and 23 features in S-9 and S-14). Notably, the inclusion of GAD and BPAQ symptoms in S-14 resulted in higher accuracy, indicating the relevance of these additional features (see Figure 24 for source code details). Our next step is model evaluation (see Figure 25 for details). Performance metrics such as accuracy, precision, recall, error rate, and f-measure were computed and compared across different models. After thorough evaluation, logistic regression emerged as the best-performing model and was chosen for deployment in subsequent stages. The selected logistic regression model was saved using the joblib.dump method. This step was crucial for preserving the trained model's parameters, allowing for seamless integration into our mobile application. To assess the generalisation ability of the models, 10-fold cross-validation was performed on the five implemented models. This involved splitting the dataset into 10 subsets, training the models on 9 subsets, and validating on the remaining subset. The process was repeated, ensuring robustness and reliability in predicting emotional states with unseen data.

Next, we will provide a comprehensive overview of the proposed mobile application development. The initial phase of our mobile application development involved planning to ensure the usability of our prediction model for end-users. We opted for React Native to build the frontend, providing a versatile solution that supports Android, iOS, and web versions. Additionally, Firebase was chosen as the backend for its rapid development capabilities, streamlining the overall development process. Before moving into development, we designed the user interface (UI) using Figma. This step allowed us to create a prototype of the app, visualising the layout and flow and ensuring a user-friendly and intuitive design. The frontend of the mobile app was implemented using React Native. Various user interfaces were developed to enhance user interaction, including login and register (see Figure 8), home (see Figure 9), anger and anxiety assessment (see Figures 10 and 11), anger and anxiety assessment prediction outcome (see Figure 11), physical exercise interface [see Figure 12(a)], article readings interface [see Figure 12(b)], find doctors and appointment booking interface (see Figure 13), find hospitals interface (see Figure 14), emergency contact and logout (see Figure 15) and feedback interface. Each interface was carefully created to provide a seamless and comprehensive user experience. The backend of our mobile application was developed using Firebase. The Firebase Realtime Database was utilised to store essential data such as user information, user readings, hospital and doctor information, and other relevant app data. This ensured efficient data management and retrieval for various functionalities within the app. To seamlessly integrate our prediction model into the mobile app, a Python FastAPI backend was implemented. The trained model, saved using joblib, was deployed as a FastAPI API endpoint. The Firebase backend was connected to the FastAPI backend using a REST API, enabling the app to perform predictions and receive outcomes. The FastAPI backend was hosted on an Amazon AWS Lightsail server, ensuring reliable communication between the mobile application and the prediction model. By combining these elements – React Native for frontend development, Firebase for the backend, and FastAPI for model integration – we created a robust and user-friendly mobile application capable of assessing and predicting emotional states, providing relevant information, and facilitating essential functionalities such as finding doctors or hospitals and appointment booking.

6 Conclusions

This paper provides a user-friendly anger and anxiety disorder prediction model by examining five machine learning models (i.e., SVM, KNN, decision tree, random forest, logistic regression) and utilising 23 key features. This paper also develops a mental healthcare dataset based on a user survey-based questionnaire approach and the collected dataset is validated by the healthcare professionals. The multi-class classification approach (neither anger nor anxiety, anger only, anxiety only, both anger and anxiety) using machine learning yields valuable insights into mental health assessment. The results demonstrated that the logistic regression model is best suited for the anger and anxiety disorder prediction scheme with 97% accuracy. This work also offers a user-centric mental healthcare mobile application with personalised recommendations and resources. The integration of an anger and anxiety assessment feature, reading articles, physical exercise suggestions, doctor search, hospital search feature, and access to mental health professionals enhances the proposed mobile applications' ability to promote mental well-being. The future work for the mental health app involves improving data collection and increasing the number of participants to enhance accuracy. Emphasis will be placed on enhancing the machine learning models, incorporating deep learning methods, and optimising feature selection. Moreover, other mental health disorders will be considered to expand the app's scope and accommodate a broader audience. Personalised recommendation features will be integrated to offer custom assistance, and the app will be adapted to serve both doctors and hospitals, allowing users to access a wide network of mental health professionals. Video call consultation facilities will be introduced to facilitate real-time interactions with experts. Additionally, exercise tracking with smartwatch implementation will empower users to monitor their physical activities about mental well-being. Finally, the app will be made accessible to users from diverse demographics, ensuring a wider reach and impact. Through these efforts, an even more effective and inclusive mental health app will be developed for all users.

References

Ahmed, A. et al. (2020) 'A machine learning approach to detect depression and anxiety using supervised learning', *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, Gold Coast, Australia, pp.1–6.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*, American Psychiatric Association, Reston, VA, USA.

Analytics Vidhya (2020) *K-Fold Cross Validation Technique and its Essentials* [online] https://www.analyticsvidhya.com/blog/2022/02/k-fold-cross-validation-technique-and-its-essentials/ (accessed November 2023).

Antonatos, L. (2023) *Anxiety and Anger: Connections, Treatments, and How to Cope* [online] https://www.choosingtherapy.com/anxiety-and-anger/ (accessed November 2023).

Arif, M. et al. (2020) 'Classification of anxiety disorders using machine learning methods: a literature review', *Insights of Biomedical Research*, Vol. 4, No. 1, pp.95–110.

Arusha, A.R. et al. (2020) 'Prevalence of stress, anxiety and depression due to examination in Bangladeshi youths: a pilot study', *Children and Youth Services Review*, Vol. 116, pp.1–7.

Bonanno, G.A. et al. (2004) 'Loss, trauma, and human resilience: have we underestimated the human capacity to thrive after extremely aversive events?', *The American Psychologist*, Vol. 59, No. 1, pp.20–28.

Burkhardt, F. et al. (2006) 'Detecting anger in automated voice portal dialogs', *INTERSPEECH, ICSLP*, pp.1053–1056.

Bögels, S.M. et al. (2006) 'Family issues in child anxiety: attachment, family functioning, parental rearing and beliefs', *Clinical Psychology Review*, Vol. 26, No. 7, pp.834–856.

Chung, J. et al. (2022) 'Mental health prediction using machine learning: taxonomy, applications, and challenges', *Applied Computational Intelligence and Soft Computing*, Vol. 2022, Article ID 9970363, pp.1–19.

Copeland, W.E. et al. (2013) 'Adult psychiatric outcomes of bullying and being bullied by peers in childhood and adolescence', *JAMA Psychiatry*, Vol. 70, No. 4, pp.419–426.

Deeba, F. et al. (2004) 'Development of an anxiety scale for Bangladeshi population', *Bangladesh Psychological Studies*, Vol. 14, pp.39–54.

Gerevich, J. et al. (2007) 'The generalizability of the Buss-Perry aggression questionnaire', *International Journal of Methods in Psychiatric Research*, Vol. 16, No. 3, pp.124–136.

Goodwin, R.D. et al. (2013) 'Anxiety disorders and drug dependence: evidence on sequence and specificity among adults', *Psychiatry and Clinical Neurosciences*, Vol. 67, No. 3, pp.1–12.

Gupta, G.K. et al. (2023) 'Depression detection using semantic representation based semi-supervised deep learning', *International Journal of Data Analysis Techniques and Strategies*, Vol. 15, No. 3, pp.217–237.

Hamilton, M. et al. (1967) 'Development of a rating scale for primary depressive illness', *British Journal of Social and Clinical Psychology*, Vol. 6, No. 4, pp.278–296.

Hawker, D.S. et al. (2000) 'Twenty years' research on peer victimization and psychosocial maladjustment: a meta-analytic review of crosssectional studies', *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, Vol. 41, No. 4, pp.441–455.

Healthline Authority (2023) *Is There a Connection between Anxiety and Anger* [online] https://www.healthline.com/health/anxiety/anxiety-and-anger (Accessed July 2023).

Hilbert, K. et al. (2017) 'Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: a multimodal machine learning study', *Brain and Behavior*, Vol. 7, No. 3, pp.1–11.

Ismail, M.M.B. et al. (2022) 'Insult detection using a partitional CNN-LSTM model', *International Journal of Data Analysis Techniques and Strategies*, Vol. 14, No. 4, pp.336–349.

Kapoor, A. et al. (2022) ''Prediction of anxiety disorders using machine learning techniques', *2022 IEEE Bombay Section Signature Conference (IBSSC)*, India, pp.1–6.

Keles, B. et al. (2020) 'A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents', *International Journal of Adolescence and Youth*, Vol. 25, No. 1, pp.79–93.

Khamparia, A. et al. (2020) 'A novel integrated principal component analysis and support vector machines-based diagnostic system for detection of chronic kidney disease', *International Journal of Data Analysis Techniques and Strategies*, Vol. 12, No. 2, pp.99–113.

Masri, R.Y. et al. (2012) 'Employing artificial intelligence techniques in mental health diagnostic expert system', *2012 International Conference on Computer and Information Science (ICCIS)*, Kuala Lumpur, Malaysia, pp.495–499.

McCloskey, L.A. et al. (1995) 'The effects of systemic family violence on children's mental health', *Child Development*, Vol. 66, No. 5, pp.1239–1261.

G. Miner et al. (2016) *Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost-Effectiveness for Healthcare Administration and Delivery Including Medical Research*, 1st ed., Academic Press, Inc., Cambridge, MA, USA.

Monteiro, S.O.M. et al. (2017) '(Re)examining the factorial structure of the generalized anxiety disorder-7 in a college students sample', *European Psychiatry*, Vol. 41, pp.1–10.

Nichols, T.R. et al. (2008) 'Examining anger as a predictor of drug use among multiethnic middle school students', *The Journal of School Health*, Vol. 78, No. 9, pp.480–486.

Owens, M. et al. (2012) 'Anxiety and depression in academic performance: an exploration of the mediating factors of worry and working memory', *School Psychology International*, Vol. 33, No. 4, pp.433–449.

Pavlidis, I. et al. (2001) 'Thermal image analysis for anxiety detection', *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, Vol. 2, pp.315–318.

Priya, A. et al. (2020) 'Predicting anxiety, depression and stress in modern life using machine learning algorithms', *Procedia Computer Science*, Vol. 167, pp.1258–1267.

Richardson, T. et al. (2017) 'A longitudinal study of financial difficulties and mental health in a national sample of British undergraduate students', *Community Mental Health Journal*, Vol. 53, No. 3, pp.344–352.

Rosner, R. et al. (2011) 'Treatment of complicated grief', *European Journal of Psychotraumatology*, Vol. 2.

Sadof, R.L. et al. (1976) 'Violence in families: an overview', *The Bulletin of the American Academy of Psychiatry and the Law*, Vol. 4, No. 4, pp.292–296.

Sagar, K. (2023) *Is There a Connection between Anxiety and Anger?* [online] https://www.cadabamshospitals.com/anxiety-and-anger-connection-and-treatment/ (accessed November 2023).

Savage, J. et al. (2017) 'The effect of academic achievement on aggression and violent behavior: a meta-analysis', *Aggression and Violent Behavior*, Vol. 37, pp.91–101.

Sharma, A. et al. (2021) 'Understanding importance of clinical biomarkers for diagnosis of anxiety disorders using machine learning models', *PloS One*, Vol. 16, No. 5, pp.1–5.

Siddique, A.B. et al. (2022) 'A machine learning-based approach to predict university students' depression pattern and mental healthcare assistance scheme using Android application', *International Journal of Data Analysis Techniques and Strategies*, Vol. 14, No. 2, pp.122–139.

Sumathi, M. et al. (2016) 'Prediction of mental health problems among children using machine learning techniques', *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 1, pp.1–5.

Thompson, J.S. et al. (2021) 'The role of anxiety sensitivity in anger symptomatology: results from a randomized controlled trial', *Journal of Anxiety Disorders*, Vol. 83, pp.1–10.

van Santvoort, F. et al. (2015) 'The impact of various parental mental disorders on children's diagnoses: a systematic review', *Clinical Child and Family Psychology Review*, Vol. 18, No. 4, pp.281–299.

World Health Organization (WHO) (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research* [online] https://apps.who.int/iris/handle/10665/37108 (accessed July 2023).

Walsh, L.M. et al. (2018) 'The relationship between anger and anxiety symptoms in youth with anxiety disorders', *Journal of Child and Adolescent Counseling*, Vol. 4, No. 2, pp.117–133.