An AI-Powered Mental Health Monitoring Application: A Research on Mental Health Classification

by

Sayem Hossen
ID: CSE2101022074
Amit Hasan
ID: CSE2101022129
Monalisa Moon Joti
ID: CS2101022122
Supervised by
Sadia Tasnim Barsha

Submitted in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science and Engineering



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SONARGAON UNIVERSITY (SU)
January 2025

An AI-Powered Mental Health Monitoring Application: A Research on Mental Health Classification

by

Sayem Hossen
ID: CSE2101022074
Amit Hasan
ID: CSE2101022129
Monalisa Moon Joti
ID: CS2101022122

Supervised by
Sadia Tasnim Barsha

Submitted in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science and Engineering



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SONARGAON UNIVERSITY (SU)
January 2025

APPROVAL

The thesis and project titled "**An AI-Powered Mental Health Monitoring Application: A Research on Mental Health Classification**" submitted by Sayem Hossen (CSE2101022074), Amit Hasan (CSE2101022129), Monalisa Moon Joti (CS2101022122) to the Department of Computer Science and Engineering, Sonargaon University (SU), has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents.

Board of Examiners

Supervisor

Sadia Tasnim Barsha
Lecturer & Exam Coordinator,
Department of Computer Science & Technology
Sonargaon University (SU)

Examiner 1

(Examiner Name and Signature)
Department of Computer Science and Engineering
Sonargaon University (SU)

Examiner 2

(Examiner Name and Signature)
Department of Computer Science and Engineering
Sonargaon University (SU)

Examiner 3

(Examiner Name and Signature)
Department of Computer Science and Engineering
Sonargaon University (SU)

DECLARATION

We, hereby, declare that the work presented in this report is the outcome of the investigation performed by us under the supervision of Sadia Tasnim Barsha**,** Lecturer & Exam Coordinator**,** Department of Computer Science & Technology, Sonargaon University, Dhaka, Bangladesh. We reaffirm that no part of this thesis and project has been or is being submitted elsewhere for the award of any degree or diploma.

Countersigned

(Sadia Tasnim Barsha)
Supervisor

Signature

Sayem Hossen
ID: CSE2101022074

Amit Hasan
ID: CSE2101022129

Monalisa Moon Joti
ID: CS2101022122

ABSTRACT

The project discussed here deals with the conditions called Anxiety, Depression, Stress, and Suicidal tendencies, leading in a youth-oriented culture, under the broad area of Mental Health detection. A study is presented next that considers the labeled dataset comprising around 53,000 sentences describing seven states of minds, namely, Anxiety, Normal, Depression, Suicidal, Stress, Bipolar, and Personality Disorder. The different machine learning and deep learning models used include Roberta, GRU, LSTM, Logistic Regression, and Random Forest. Of these, the Roberta model produced a maximum accuracy of 81% and demonstrated an excellent performance in capturing contextual relationships within text. Further, this final model was then integrated with a chat-based application with the aim of real-time monitoring of states within user networks. This gives insight into good mental health and can, therefore, be helpful for intervention and support. The system provides a promising solution to address mental health concerns. This research contributes significantly to the field of mental health support, aiming to promote emotional well-being and foster healthier communities through AI-driven solutions.

ACKNOWLEDGMENT

At the very beginning, we would like to express my deepest gratitude to the Almighty Allah for giving us the ability and the strength to finish the task successfully within the schedule time.

We are auspicious that we had the kind association as well as supervision of **Sadia Tasnim Barsha**, Lecturer & Exam Coordinator, Department of Computer Science & Technology, Sonargaon University whose hearted and valuable support with best concern and direction acted as necessary recourse to carry out our project.

We are also thankful to all our teachers during our whole education, for exposing us to the beauty of learning.

Finally, our deepest gratitude and love to my parents for their support, encouragement, and endless love also deepest gratitude my juniors and senior for their support.

Table of Contents

Title **Page number**

APPROVAL ..... **ii**
DECLARATION ..... **iii**
ABSTRACT ..... **iv**
ACKNOWLEDGMENT ..... **v**
CHAPTER 1 ..... **1**
INTRODUCTION TO AI-POWERED MENTAL HEALTH MONITORING 1
1.1 Introduction **1**
1.2 Objectives **2**
1.3 Motivation **2**
1.4 Research Question **3**
1.5 Expected Outcome **3**
1.6 Layout of the Report **4**
CHAPTER 2 ..... **5**
BACKGROUND STUDY AND LITERATURE REVIEW 5
2.1 Introduction **5**
2.2 Literature Review **5**
2.3 Challenges **7**
2.3.1 Lack of Proper Bangla Dataset 7
2.3.2 Switching to English Dataset 7
2.3.3 Key Challenges: 7
CHAPTER 3 ..... **8**
RESEARCH METHODOLOGY 8
3.1 Introduction **8**
3.2 Study Toolkit **8**
3.2.1 Library: 8
3.2.2 Framework: 9
3.2.3 Computational Resources: 9
3.3 Methodology **9**
3.4 Dataset description **11**
3.5 Data Preprocessing **13**
3.6 Model Overview **15**

3.6.1 GRU (Gated Recurrent Unit) 15
3.6.2 LSTM (Long Short-Term Memory) 16
3.6.4 Logistic Regression 17
3.6.5 Random Forest 18
CHAPTER 4 19
EXPERIMENTAL RESULTS AND ANALYSIS 19
4.1 Introduction 19
4.2 Performance Measurement Tools 19
4.2.1 Confusion Matrix 19
4.2.2 Classification report 20
4.3 Results of Data Preprocessing 20
4.4 Model Performance 21
4.4.1 GRU 21
4.4.2 LSTM 23
4.4.3 Roberta 25
4.4.4 Logistic Regression 27
4.4.5 Random Forest 28
4.5 Model Comparison 30
CHAPTER 5 31
APPLICATION DEVELOPMENT PROCESS 31
5.1 Introduction 31
5.2 Entity-Relationship Diagram (ERD) 31
5.3 Analysis and Visualization 32
5.4 Technology Stack 38
5.5 Libraries and Tools 40
5.6 Interface 43
CHAPTER 6 46
CONCLUSION AND FUTURE WORKS 46
6.1 Conclusion 46
6.2 Limitation 46
6.3 Future Work 48
REFERENCE 49

LIST OF FIGURES

Figure Number	Figure Title	Page Number
Figure 3.3	Methodology	9
Figure 3.4.1	Emotion Categories Graph	11
Figure *3*.5.1	Visualize Average Length of Sentence	14
Figure 3.6.1	GRU Model	15
Figure 3.6.2	LSTM Model	15
Figure 3.6.3.1	Roberta Model	16
Figure 3.6.4.1	Logistic Regression Model	17
Figure 3.6.5.1	Random Forest Model	17
Figure 4.4.1.1	Training and Validation Loss	21
Figure 4.4.1.2	Confusion Matrix of GRU	22
Figure 4.4.2.1	Training and Validation Loss	23
Figure 4.4.2.2	Confusion Matrix of LSTM	24
Figure 4.4.3.1	Confusion Matrix of Roberta	25
Figure 4.4.4.1	Confusion Matrix of Logistic Regression	27
Figure 4.4.5.1	Confusion Matrix of Random Forest	28
Figure 4.5.1	Model Comparison	29
Figure 5.2.1	ER diagram	32
Figure 5.3.1.1	All Chat condition Distribution	33
Figure 5.3.1.2	All Chat Detailed Analysis	33
Figure 5.3.1.3	All chat Normal Trend	34
Figure 5.3.2.1	Self-Chat Condition Distribution	35
Figure 5.3.2.2	Self-Chat Details Analysis	36
Figure 5.3.2.3	Self-Chat Personally disorder Trend	36
Figure 5.3.3.1	Friends Chat Condition Distribution	37
Figure 5.3.3.2	Friends Chat Details Analysis	38
Figure 5.3.3.3	Friends Chat Normal Trend	39
Figure 5.6	Interface	48

LIST OF TABLES

Table Number	Table Title	Page No
Table 3.4.1	Dataset	10
Table 3.4.2	Dataset Category	11
Table 3.5.1	Clean Dataset	12
Table 3.5.2	Stop Words Removed Dataset	13
Table 4.2.1.1	Confusion Matrix	19
Table 4.3.1	Our Final Dataset	20
Table 4.4.1.1	Classification Report of GRU	21
Table 4.4.2.1	Classification Report of LSTM	23
Table 4.4.3.1	Classification Report of Roberta Model	25
Table 4.4.4.1	Classification Report of Logistic Regression	26
Table 4.4.5.1	Classification Report of Random Forest	28

LIST OF EQUATIONS

Equation Number	Equation Title	Page Number
Equation 1	Precision	19
Equation 2	Recall	19
Equation 3	F1-Score	19

LIST OF ABBREVIATIONS

1. **GRU** - Gated Recurrent Unit
2. **LSTM** - Long Short-Term Memory
3. **RoBERTa** - Robustly Optimized BERT Approach
4. **GPU** - Graphics Processing Unit
5. **SVM** - Support Vector Machine
6. **PMI** - Pointwise Mutual Information
7. **NLP** - Natural Language Processing
8. **CNN** - Convolutional Neural Network
9. **ERD** - Entity-Relationship Diagram
10. **TF-IDF** - Term Frequency-Inverse Document Frequency
11. **ML** Machine Learning
12. **DL** - Deep Learning

CHAPTER 1

INTRODUCTION TO AI-POWERED MENTAL HEALTH MONITORING

1.1 Introduction

In the last few years, the world has witnessed a generational shift in mental health. Due to increased exposure to social media, academic pressures, global uncertainties, and personal challenges, there is a sudden surge in emotional instabilities like anxiety, depression, and sadness among them. The global COVID-19 pandemic further exacerbated the issues and plunged youth into a mental health crisis. Studies have indicated that the younger generation is more vulnerable to emotional disorders that negatively affect their potential contribution to society.

Emotions are what define the behavioral pattern and decision-making process in every individual's life along with relationships. Emotional health among the young is very important because it builds up their coping mechanism regarding challenges in life and helps them maintain mental stability. Unresolved emotional problems, when related to anxiety, depression, and stress, escalate into dangerous mental conditions that might lower academic performance, deteriorate relationships, and even lead to suicidal tendencies in extreme cases. Therefore, it will be quite imperative to keep a keen eye on monitoring and understanding their emotions as a way to provide timely interventions for a much healthier society.

To handle such challenges, we propose an intelligent system that monitors the emotional states of youngsters. The proposed system can easily be integrated into chat applications, where a user can see the emotional state of his or her friends based on the text they are exchanging. For example, it will show the number of sentences exchanged by each friend that express depression, sadness, joy, and anxiety in the chat list. This innovative feature enables users to identify and react to their friends' mental health needs in real time.

Our system is based on a pre-built high-quality dataset from Hugging Face [nawax0x1/emotion detection], which contains more than 53,000 labeled sentences categorized into seven emotional states: Anxiety, Normal, Depression, Suicidal, Stress, Bipolar, and Personality Disorder. Several machine learning and deep learning models were experimented with in developing a robust emotion classification system, namely GRU, LSTM, Logistic Regression, and Random Forest. The score obtained varied between 71% and 71%. However, this model was outdone by the Roberta transformer, at 81% accuracy. It outperformed all the other models; hence, the Roberta model was considered to build our AI system.

Our research aims at making a potential contribution to society for the good of emotional awareness and mental health support. The developed AI-powered system will notify users if their friends show signs of depression, anxiety, or any other abnormal emotional state. In turn, the notification enables early detection and timely intervention that allows for creating a supportive community around the individual to help them avoid such emotional challenges before they develop.

Our proposed system has a number of key advantages. It allows the user to, in real time, monitor their emotions and hence identify any problems regarding mental health early enough. It facilitates social support by recognizing friends in emotional distress and providing meaningful interaction. It results in early detection and intervention before emotional problems escalate, thereby improving mental well-being. With easy integrations into chat applications, the system is very accessible, and the data-driven insights into emotional trends create valuable understanding both for users and professionals in view of better mental health practices.

1.2 Objectives

- **Real-time Emotion Monitoring:** It allows the users to trace the course of their emotional status in real time and, thus, motivates timely awareness about mental health problems.
- **Enhanced Social Support:** The system will identify friends in emotional distress and will encourage meaningful conversations and support.
- **Improved Mental Health:** The rationale behind early identification and intervention is that it will prevent the deterioration of emotional problems and contribute to better general mental health.
- **User-friendly Integration:** The system is designed in such a way that it can easily integrate with different chat applications out there, hence ensuring ease of use and accessibility.

1.3 Motivation

The motivation for this research is the increasing need to act upon the mental and emotional health issues of society. One can imagine the present scenario of society and social media, silently suffering from anxiety and depression among youth. Our society is getting worse day by day due to the issues which are being ignored and not resolved. Even in a close circle of friends, one friend may be depressed, but no one is aware of it.

This alarming situation outlines the importance of caring about one's close relations and society. If a person is suffering from depression or passing through emotional turmoil, on-time conversation can create a huge difference. Therefore, a system that can help one

understand the mental state of their friends is mandatory, so that they could present necessary support when needed. It is this drive of fostering an emotionally supportive society that forms the bedrock for our study and proposed AI-based emotion monitoring system.

1.4 Research Question

1. **Is the dataset and model able to perform real-world applications?**

   Yes, the dataset and the Roberta model were quite effective in classifying emotions with high accuracy; hence, it is very suitable for real-world applications. Indeed, the model identifies some important emotional states such as anxiety, depression, and stress that are highly critical in monitoring mental health conditions. Integrating this into chat applications gives users insights on trends in their emotions from which they may intervene if need be. The reliability of the model ensures its practicality in diverse scenarios, including mental health support, educational tools, and workplace wellness programs.

2. **How can people benefit from this research?**

   With the system, anyone will be able to recognize a friend or family member in emotional distress and indicate early signs of depression, anxiety, or any other mental health challenge for on-time support, which will thereafter return to the relationships, even stronger and empathetic. Such a system simply raises awareness of mental health itself and will help provide a forum of open dialogue about emotional challenges. This will enable people to seek and give emotional support to others, leading to community bonding.

3. **How can society benefit from this research?**
   By promoting early detection and intervention, the system contributes to improved mental health awareness and a more supportive community.

1.5 Expected Outcome

The strategy, therefore, has to be to develop and implement an AI-based solution embedded in chat applications to continuously monitor the mental health status of subscribers. It will scan the conversation and look out for signals associated with poor mental health conditions due to anxiety, depression, and stress. In case such abnormalities are spotted, the systems lead the user to an alert for the proper help in time. This solution is likely to promote better emotional awareness among friends and family through open communication and ways of

helping them deal with mental health issues. Eventually, the system will build a supportive community that cares about emotional health and pays attention to mental stability.

1.6 Layout of the Report

This report focuses on the construction process of one AI-powered emotional state monitoring system. First, a proper data set was taken for our study, of high importance since it ensures only reliable results later on. Secondly, cleaning, preprocessing of that data for which noise and missing value handling, preparation before modeling was going on. After that, we applied different ML and DL models such as GRU, LSTM, Logistic Regression, Random Forest, and Roberta for emotion classification. Hence, we will choose the final model that will be integrated into our software system. In fact, the whole process, starting from dataset preparation to model selection, took around three weeks, one week for each major task. This systematic approach thus made sure that the system was robust and reliable for real-world deployment.

CHAPTER 2

BACKGROUND STUDY AND LITERATURE REVIEW

2.1 Introduction

Emotion recognition has become an important part of modern technology, since people are increasingly expressing their feelings through text on social media, messaging apps, and other platforms. Some of the emotions that shape our communication and interaction with others. These emotions, detected by computer systems, can improve mental health tools, enhance online experiences, and create better human-computer interactions.

In the last years, much attention has been focused on advanced machine learning and deep learning techniques in emotion analysis from text data. Deep learning models, like Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and transformers such as Roberta, show more accurate results by learning deeper patterns in language.

The paper reviews some of the works related to emotion detection, ranging from the early approaches that use sentiment analysis models to modern methods employing deep learning techniques. It then highlights various applications related to health, education, and e-commerce in this work. These techniques will, therefore, be of help in the development of better applications related to monitoring emotion apps and mental health tools. This review will outline strengths and shortcomings of some of the works in this domain, which will pave the way for more effective solutions with regard to emotion classification.

2.2 Literature Review

G. Pattun et al. developed a novel method for fine-grained emotion classification from user-generated text by leveraging the power of GPT embeddings. Their study emphasized the semantic richness of GPT embeddings, which increases the accuracy of machine learning models on the task of emotion detection. They presented how GPT-based models can be leveraged to enhance the understanding of sentiment analysis, leading toward advanced emotion classification approaches [1]. R. Padate et al. proposed an unsupervised machine learning approach to classify WhatsApp messages into emotions. The research, by using NLP with a PMI-based algorithm, did classify messages into six emotions with a precision of 92.75%. It is very effective in handling informal languages and slang; hence, it can be developed for real-world applications in text-based emotion detectors and sentiment analyses. [2]. R. R. Malagi et al. proposed the detection of emotions from textual data by employing supervised machine learning models. They considered Decision Trees, Naive Bayes, SVM, Logistic Regression, k-Nearest Neighbors, and Random Forest in this work. The six basic emotions considered in this work are joy, fear, anger, love, surprise, and

sadness. Pre-processing techniques such as stemming, tokenization, and stop-word removal have improved the results of this model and solved some of the challenges regarding emotion analysis. [4]. R. K. Kaliyar et al. have proposed a deep learning-based framework for emotion intensity and classes extraction from tweets. It utilized several models, such as CNN and LSTM, to tackle various tasks on emotion intensity regression, ordinal classification, and multi-label classification with ordinal class dependencies, while gaining 77.64% in ordinal classification. This study placed strong emphasis on fine-grained intensity scoring and showed how deep learning-based approaches may become effective in the task of detecting emotions from tweets [5].

G. S. M. Mursalin et al. have proposed a deep learning approach for emotion recognition in Bengali-English code-mixed text, addressing the challenges of low-resource language processing. They created a corpus of 10,221 sentences and evaluated various word embedding techniques, including Word2Vec and FastText. Their BiLSTM model with Word2Vec achieved the highest accuracy of 76.1%, outperforming other methods in classifying six emotions: disgust, sadness, joy, anger, fear, and surprise [6]. N. V. Vinay et al. explored emotion detection using deep learning, emphasizing its ability to capture detailed representations of emotions and facial expressions. Their study highlighted the superiority of deep learning over traditional machine learning in emotion recognition tasks. Applications range from mental health diagnostics to enhancing virtual assistants. By leveraging large datasets and advanced algorithms, this research underscored the growing potential for innovation in emotion detection across diverse domains [7]. W. Wang and W. Ding proposed an algorithm improvement for SVM classification, optimized with a Genetic Algorithm, to classify emotions in E-learning systems. The proposed algorithm proved very promising and efficient in improving a system's performance by classifying learning emotions effectively [8]. A. Pandiaraj et al. conducted a comparative study of various deep learning and machine learning techniques to carry out sentiment analysis on 10,000 tweets. They came up with a hybrid model of word embedding algorithm and proposed the website for categorizing the tweets into three classes: positive, negative, and neutral [9]. H. A. Sayyed et al. proposed a deep learning framework in which LSTM and RNN classify the textual data for emotions, which include six categories of human feelings: joy, sadness, anger, fear, love, and surprise. Their research improves the accuracy of emotion classification and also focuses on comparative analysis in sentiment models [10].

2.3 Challenges

Apart from these, in this particular study, we encountered quite a few issues that considerably hampered our progress and results in research. Emotion detection has automatically been a serious and challenging task, and focusing on the development of an emotion monitoring system brought on several key obstacles:

2.3.1 Lack of Proper Bangla Dataset

One of the major challenges we faced was the unavailability of a proper Bangla emotion dataset. Despite searching on the internet and testing multiple Bangla datasets, none of them gave good results; the accuracy ranged from 50-60%. This is because it is tough to model emotions in a low-resource language like Bangla, where nuances within the language further complicate classification.

2.3.2 Switching to English Dataset

After trying many experiments and failing with Bangla datasets, we decided that an English dataset would serve the purpose better. This decision was made to take full advantage of available high-quality datasets and to enhance the performance of our emotion classification models.

2.3.3 Key Challenges:

- **Finding a Proper Dataset**: Finding a proper dataset that would match the emotion detection task was time-consuming because it needed really high-quality and labeled data.
- **Data Cleaning**: For this, the datasets had to be preprocessed by removing the noise, correcting errors, and matching consistency.
- **Identifying Stop Words for Emotions**: The model had to be trained to filter out the stop words related to emotions so that it could understand the emotional content of the text.
- **Balancing the Dataset**: Where one or more emotions had too much representation in the dataset, this skewed the dataset, leading to potential bias in the model's predictions.
- **Model Fitting**: GRU, LSTM, and Roberta training models require further tuning of the parameters for desired accuracy.
- **Integrating the Model into the App**: The deployment of the trained model in the app, which was then made to work for real-time emotion classification, is one more technical challenge.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The section describes our structured approach to the development of an AI emotion monitoring system. The methodology followed in this study was well-planned to ensure that the proposed solution was appropriate, reliable, and valid. We initiated the selection of a dataset to train and test our models. The dataset from Hugging Face contained more than 53,000 labeled sentences in seven emotional categories—a strong backbone for our study.

Cleaning and preprocessing of data were done to eliminate inconsistencies and better prepare it for modeling. This was further followed by several machine learning and deep learning techniques, namely, GRU, LSTM, Logistic Regression, Random Forest, and Roberta, in classifying emotions. After iterative experimentation, the best model selected was Roberta with an accuracy of 81%. The final step was to embed the best model into a chat-based application so that real-time monitoring of emotion would be possible. This structured methodology ensures consistency and an effective way to identify states of emotion.

3.2 Study Toolkit

Tools and materials used in this research are very diverse for the effective analysis of data, model training, and evaluation. The Pandas and NumPy packages were employed to manipulate the data and compute the numerical ones; Matplotlib did the plotting, showing insight from data and performance into the models under study. On the other hand, TensorFlow and PyTorch deep learning frameworks were relevant in the implementation and fine-tuning of Roberta, GRU, LSTM, and CNN. The computation was performed on a Kaggle P100 GPU with 16 GB memory and was found to be very well-suited for training all the deep learning models. Next, Dart and Flutter were leveraged to design the application used for integrating an AI-powered emotion monitoring system by way of an intuitive chat interface.

3.2.1 Library:

- Pandas: For data manipulation and analysis.
- NumPy: For numerical computations.
- Matplotlib: For data visualization.

3.2.2 Framework:

PyTorch: For implementing and fine-tuning models such as Roberta, GRU, and LSTM.

3.2.3 Computational Resources:

- Kaggle P100 GPU: 16 GB memory for training complex models and managing high computational demands.

3.3.4 Application Development Tools:

- Dart and Flutter: For developing chat-based application and integrating the AI system seamlessly.

3.3 Methodology

The methodology for this research is structured around a systematic approach to analyzing Mental Health datasets and classifying them into meaningful categories. This dataset contains a mix of structured and unstructured data, requiring preprocessing to make it usable for machine learning models.

The next step is data preprocessing, which involves cleaning the data to remove unnecessary elements such as noise, irregularities, or irrelevant content. This step ensures that the data is standardized and ready for further analysis. Following preprocessing, Stop Word Removal is performed to eliminate common words (such as conjunctions, prepositions, and articles) that do not contribute significantly to the sentiment or meaning of the comments.

After the removal of the stop word, the preprocessed data undergoes tokenization, whereby text is split either into individual words, phrases, or tokens. These steps are quintessential in transforming the text so that machine learning algorithms may start analyzing the said text efficiently and effectively.

Then, the processed and tokenized data is fed to the Training Classifier. At this step in a sequence, a number of deep learning models—among which the major are Roberta, GRU, LSTM, Logistic Classifier, and Random Forest—are trained on this dataset to grasp some knowledge on pattern improvement from the data and the relationship that exists between or among them. These models, in turn, fine-tune to optimize performances related to categorically classifying mental health into 'Anxiety', 'Normal', 'Depression', 'Suicidal', 'Stress', 'Bipolar', and 'Personality disorder'.

The models then head for evaluation, where the performance of each is gauged by measures such as accuracy, precision, recall, and F1-score. This helps in realizing how effective each model is to offer room for improvements on weak ones.

These models are then used for prediction by being applied to new, unseen mental health data for their correct classification. The workflow is complete in its approach toward the objectives of the study and thus contributes valuably to NLP.

We further developed a chat application that integrated our best-performing model to monitor and visualize users' emotional well-being. Not only does the chat application provide a platform for communication, but it also offers real-time feedback on the state of mental health among friends through graphical representations and categorized insights. Users can view the breakdown of their friends' current emotional states, how many are anxious, depressed, or joyful. The feature encourages awareness and, perhaps, proactive conversations within circles that make life supportive. The application is designed in a user-friendly way, making it very easy to use, while its strong backend, powered by deep learning models, ensures the correct detection and analysis of emotions.

3.4 Dataset description

The data used for this work was obtained from the Hugging Face [nawax0x1/emotion detection] repository. Labels represent seven classes of emotional state (Anxiety, Normal, Depression, Suicidal, Stress, Bipolar, Personality Disorder), totaling over 53 thousand labeled sentences in the corpus. The labels correspond to different emotional states, from actual mental health conditions to just general feelings.

This data is divided into structured and unstructured data and needs to be pre-processed to ensure it is structured for training machine learning models. In the graph attached, you can see how these categories are distributed and how balanced the data for the emotional states is. Normal and Depression are the two largest categories, and Personality Disorder is the smallest. It is a comprehensive dataset that provides a strong basis to train and evaluate emotion classification models.

Statement	**Status**
"Oh my gosh."	Anxiety
"Trouble sleeping, confused mind, restless heart."	Anxiety
"All wrong, back off dear, forward doubt. Stay."	Anxiety
"I've shifted my focus to something else but I..."	Anxiety
"I'm restless and restless, it's been a month now."	Suicidal
"Nobody takes me seriously. I've (24M) dealt with this."	Depression
"Selfishness. 'I don't feel very good, it's like...'"	Anxiety
"Is there any way to sleep better? I can't sleep properly."	Normal
"Public speaking tips? Hi, all. I have to give a speech soon."	Anxiety
"I have really bad door anxiety! It's not about the door, it's something else."	Personality Disorder

Table 3.4.1: Dataset

Emotion Category	Number of Sentences
Normal	16,351
Depression	15,404
Suicidal	10,653
Anxiety	3,888
Bipolar	2,877
Stress	2,669
Personality Disorder	1,201

Table 3.4.2: Dataset category

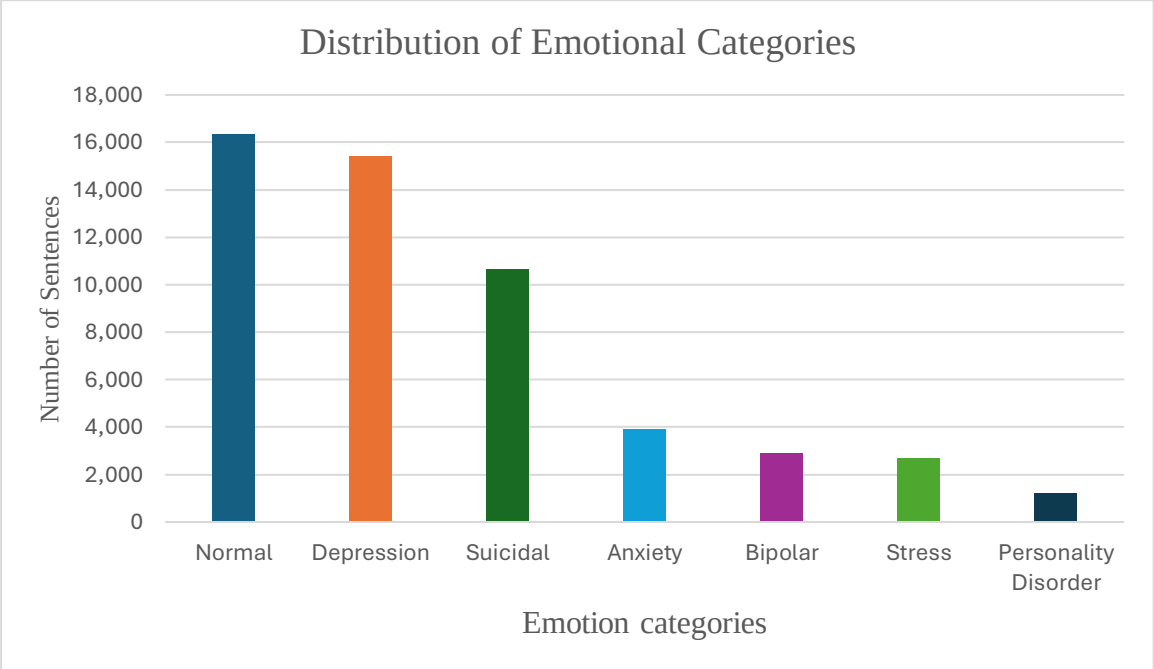


Figure 3.4.1: Emotion Categories Graph

3.5 Data Preprocessing

In the cleaning process of the raw dataset, English punctuation marks, stop words, and other unnecessary elements were removed to prepare the data for machine learning models. These stop words, including words such as “and,” “the,” and “is,” were removed to retain only the most relevant words for emotion classification. Additionally, punctuation marks and unrelated elements were eliminated to standardize the dataset and minimize noise. This critical cleaning step ensured the data was in a proper format, allowing machine learning models to effectively identify patterns and relationships for accurate emotion classification

Statement	Statement cleaned
Oh my gosh	Oh my gosh
Trouble sleeping, confused mind, restless heart...	Trouble sleeping confused mind restless heart...
All wrong, back off dear, forward doubt. Stay...	All wrong back off dear forward doubt Stay in...
I've shifted my focus to something else but i'm...	I’ve shifted my focus to something else but i’m...
I'm restless and restless, it's been a month now...	I’m restless and restless its been a month now...
Every break, you must be nervous, like something...	Every break you must be nervous like something...
I feel scared, anxious, what can I do? And maybe...	I feel scared anxious what can I do and maybe...
Have you ever felt nervous but didn't know why?	Have you ever felt nervous but didn’t know why
I haven't slept well for 2 days; it's like i'm...	I haven’t slept well for 2 days it’s like i’m...
I'm really worried, I want to cry.	Im really worried I want to cry

Table 1.5.1: Clean Dataset

Statement	Statement cleaned	Statement_without_stopwords
Oh my gosh	Oh my gosh	Oh gosh
Trouble sleeping, confused mind, restless heart...	Trouble sleeping confused mind restless heart...	Trouble sleeping confused mind restless heart...
All wrong, back off dear, forward doubt. Stay...	All wrong back off dear forward doubt Stay in...	Wrong back dear forward doubt Stay restless...
I've shifted my focus to something else but i'm...	I’ve shifted my focus to something else but i’m...	I’ve shifted focus something else I’m still worried
I'm restless and restless, it's been a month now...	I’m restless and restless its been a month now...	I’m restless restless month boy mean
Every break, you must be nervous, like something...	Every break you must be nervous like something...	Every break must nervous like something wrong...
I feel scared, anxious, what can I do? And maybe...	I feel scared anxious what can I do And maybe...	Feel scared anxious may family us protected
Have you ever felt nervous but didn't know why?	Have you ever felt nervous but didn’t know why	Ever felt nervous didn’t know
I haven't slept well for 2 days; it's like i'm...	I haven’t slept well for 2 days its like i’m...	Haven’t slept well 2 days like I’m restless huh
I'm really worried, I want to cry.	I’m really worried I want to cry	I’m really worried want cry

Table 3.5.2: Stop words remove dataset

Understanding the structure of your dataset, including average text lengths for each category, is crucial for assessing its impact on model performance. Uneven text lengths can introduce bias, especially in models like GRU and LSTM that rely heavily on sequential data. Categories with much longer or shorter texts may hinder the model's ability to learn and generalize effectively, resulting in lower accuracy. Analyzing text length helps us fine-tune hyperparameters, such as sequence length and batch size, to handle these variations effectively. This ensures balanced learning across all categories, enhancing the model's robustness and overall accuracy.

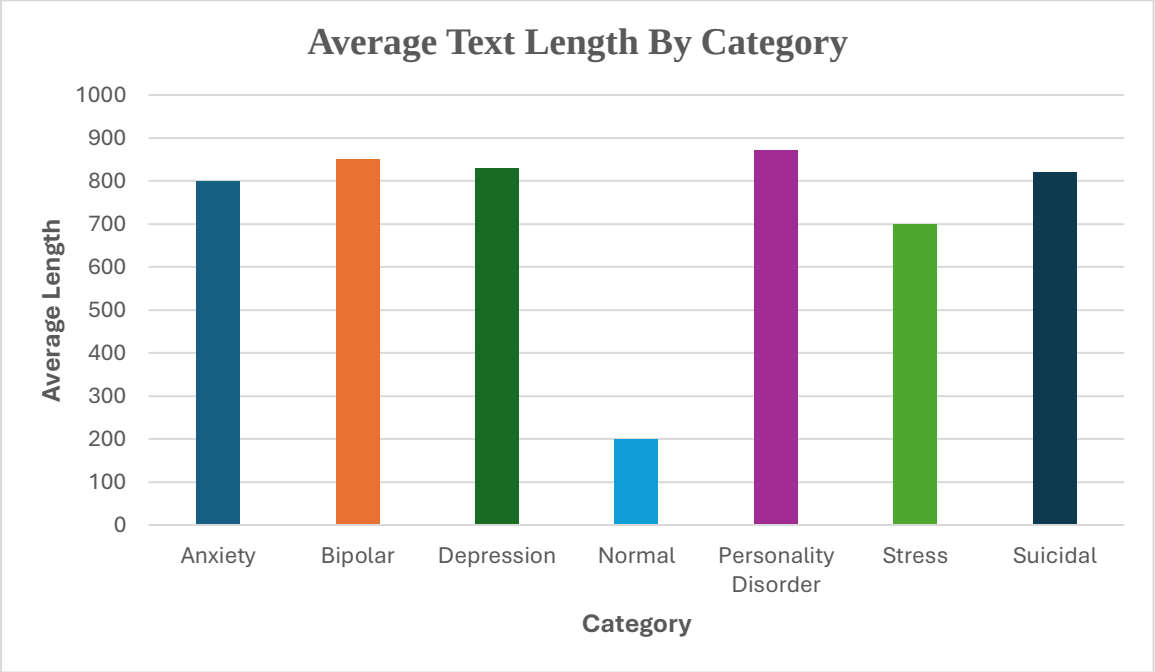


Figure 3.5.1: Visualize average length of sentence

3.6 Model Overview

In this study, five machine learning and deep learning models were used for emotion classification. Each model has distinct architecture and advantages for analyzing text data. Below are the explanations of the five models:

3.6.1 GRU (Gated Recurrent Unit)

The Gated Recurrent Unit (GRU) is a type of recurrent neural network (RNN) designed to process sequential data efficiently. By using gating mechanisms to control the flow of information, GRU effectively captures long-term dependencies while being less computationally demanding than Long Short-Term Memory (LSTM) networks. It features two key gates: the update gate and the reset gate, which help retain relevant information and discard unnecessary details in the sequence.

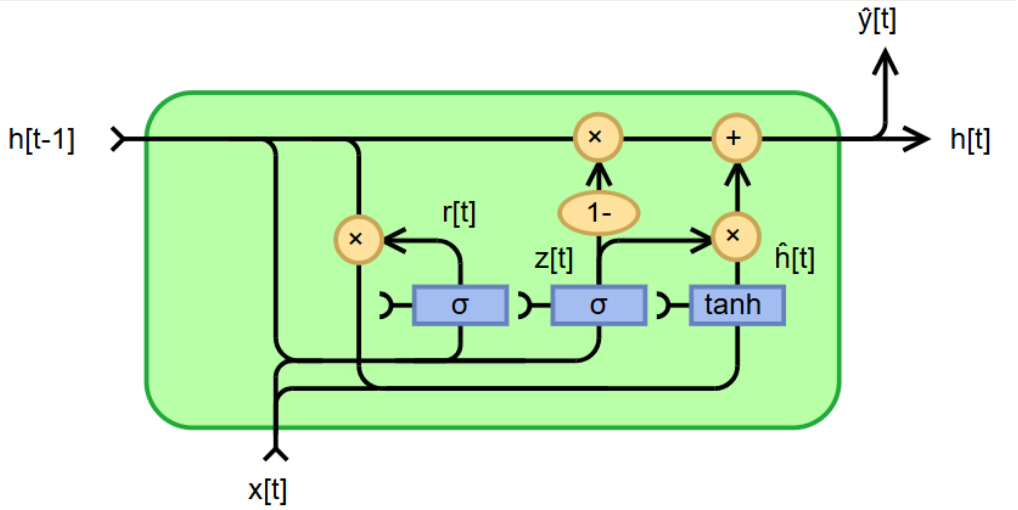


Figure 3.6.1: GRU model

3.6.2 LSTM (Long Short-Term Memory)

Long Short-Term Memory (LSTM) is an advanced RNN architecture developed to address the vanishing gradient problem in traditional RNNs. It uses three gates—input, forget, and output gates—to manage the flow of information through the network. This allows LSTM to effectively capture both short-term and long-term dependencies in text, making it highly effective for tasks like sentiment and emotion analysis.

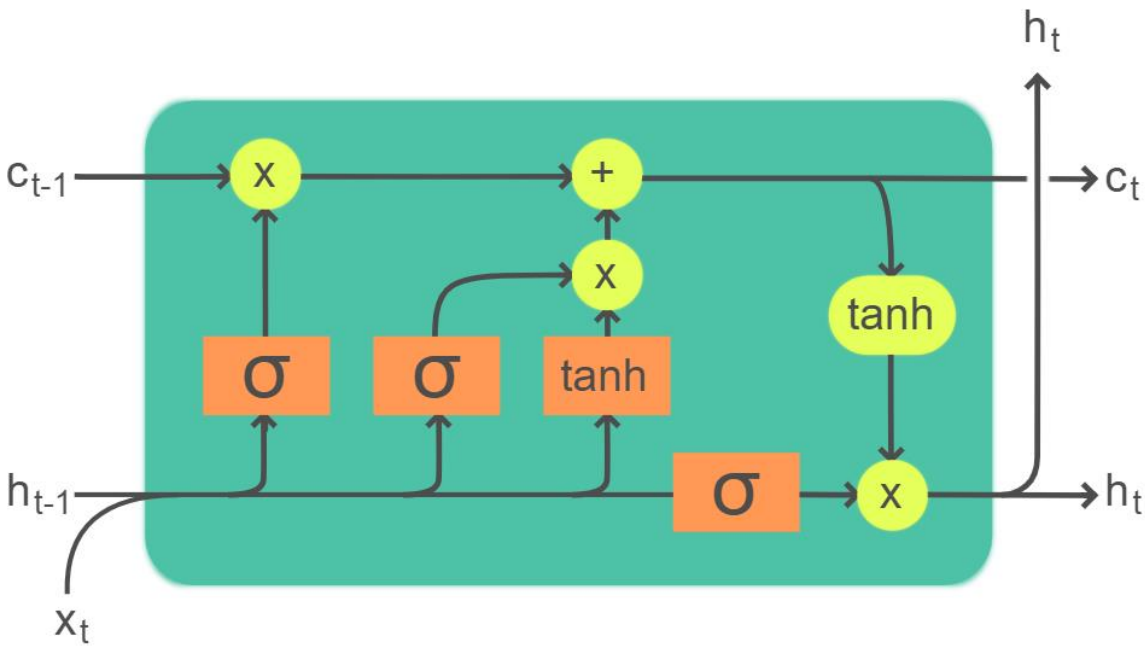


Figure 3.6.2: LSTM model

`3.6.3 Roberta (Robustly Optimized BERT Pretraining Approach)

Roberta is a transformer-based model developed to fine-tune BERT's training process and boost its performance. It leverages self-attention mechanisms to capture intricate dependencies in text and is pretrained on a large dataset. Roberta's deep understanding of context makes it particularly effective for natural language processing tasks, such as emotion classification.

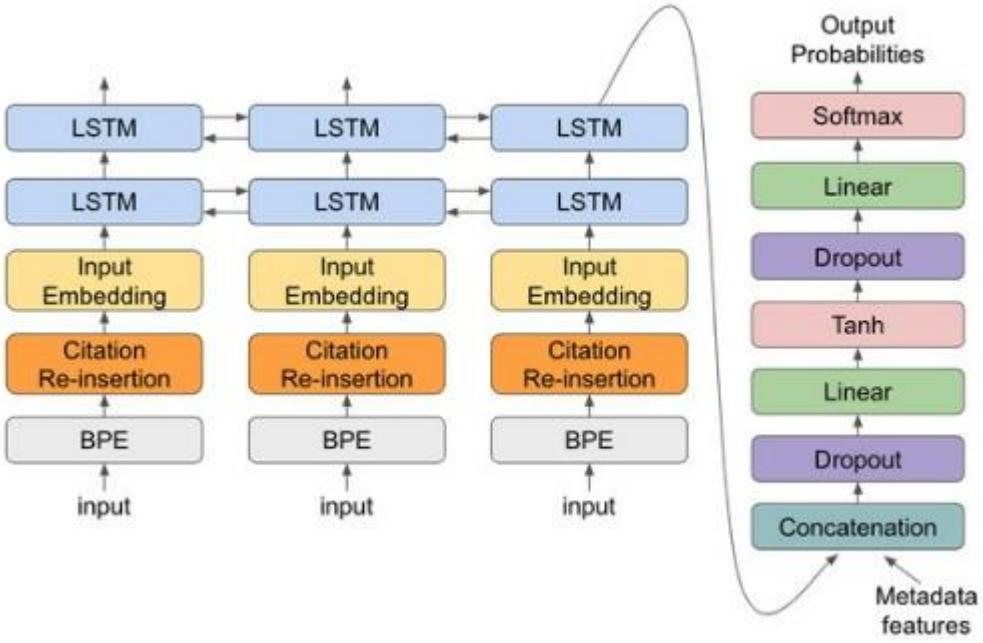


Figure 3.6.3.1: Roberta Model

3.6.4 Logistic Regression

Logistic Regression is a straightforward and easy-to-understand machine learning algorithm designed for binary or multi-class classification. It uses a sigmoidal function to transform input features into probabilities, making it particularly effective for data with linear relationships. While simple, it proves to be a strong foundation for emotion classification tasks, especially when paired with features like TF-IDF or word embeddings.

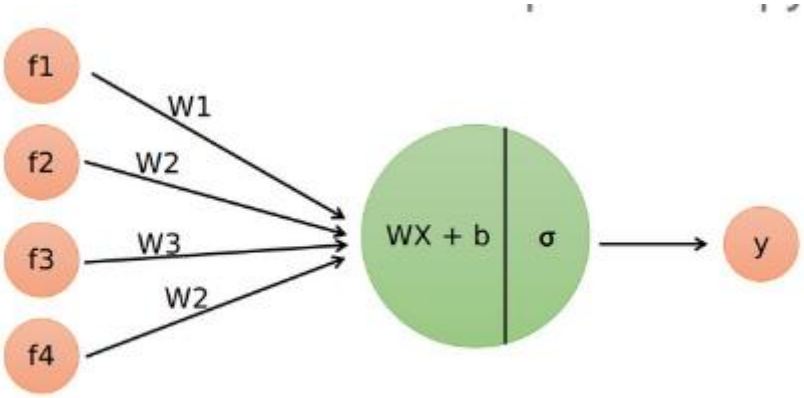


Figure 3.6.4.1: Logistic Regression model

3.6.5 Random Forest

Random Forest is an ensemble learning algorithm that creates multiple decision trees during training and makes decisions based on the majority vote for classification tasks. It is resilient to overfitting and works effectively with both categorical and continuous data. By aggregating predictions from multiple trees, Random Forest delivers high accuracy, making it an excellent choice for text classification tasks.

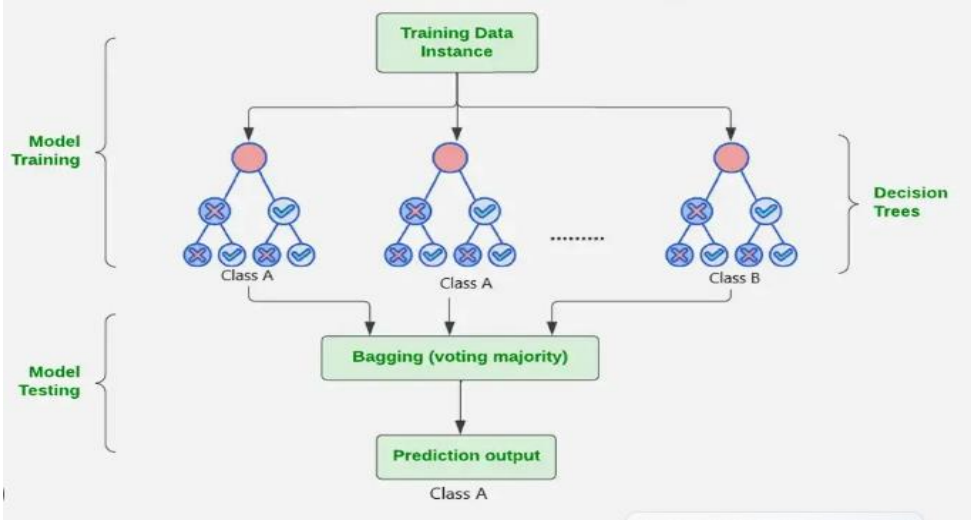


Figure 3.6.5.1: Random Forest model

CHAPTER 4

EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Introduction

This section presents the evaluation of our proposed models, including Roberta, GRU, LSTM, Logistic Regression, and Random Forest. To measure the performance of these models, we relied on key metrics such as the classification report, accuracy, F1-score, recall, and support. Confusion matrices were also created to provide a clearer view of the models' performance across various categories. Despite working with an imbalanced dataset and encountering several challenges during the study, we made a concerted effort to achieve high accuracy and reliable results. By examining the strengths and weaknesses of each model, we identified the most effective approach for emotion classification. This evaluation not only compares the models' performances but also sheds light on the impact of data preprocessing, class balancing, and model fine-tuning on the final results. The findings here set the stage for further refinement and practical application of the proposed system.

4.2 Performance Measurement Tools

4.2.1 Confusion Matrix

A confusion matrix is a highly effective tool for evaluating the performance of a classification model by comparing actual labels with predicted ones. It offers detailed insights into the model's true and false predictions. The matrix is composed of four key elements: True Positives (TP), representing correct predictions for a specific class, and True Negatives (TN), where the model accurately identifies instances that do not belong to a class. False Positives (FP) occur when the model incorrectly predicts a class, while False Negatives (FN) signify instances where the model fails to predict a class correctly.

By analyzing these components, it becomes easier to pinpoint areas where the model may struggle, such as handling class imbalances or managing overlapping features across categories. Targeting reductions in FP and FN rates enhances both precision and recall, leading to more dependable predictions. Ultimately, the confusion matrix serves as a valuable resource for understanding a model's strengths and weaknesses across various emotion categories.

	Predicted class	
True Class	True Positive	False Positive
	False Negative	True Negative

Table 4.2.1.1`: Confusion Matrix

4.2.2 Classification report

A classification report provides a detailed breakdown of a model's performance across each class using four key metrics:

- Precision: Precision is the ratio of correctly predicted positive instances to the total predicted positives. It indicates the model's ability to avoid false positives (FP). Higher precision means fewer misclassifications of a given class.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)+False Potives (FP)}}$$ *Equation 1: Precision*

- Recall (Sensitivity): Recall measures the ratio of correctly predicted positive instances to all actual positives. It reflects the model's ability to capture true positives (TP) and avoid false negatives (FN).

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP)+False\ Negatives\ (FN)}$$ *Equation 2: Recall*

- F1-Score: The F1-score is the harmonic mean of precision and recall, providing a balanced measure of a model's accuracy, especially when the dataset is imbalanced.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall}$$ *Equation 3: F1-Score*

- Support: Support represents the number of actual occurrences of each class in the dataset. It is a helpful metric for understanding class imbalance and the amount of data available for each category.

4.3 Results of Data Preprocessing

Data preprocessing was a crucial step in improving the quality of the dataset. We processed the raw data by removing punctuation, stop words, and other irrelevant elements. Missing values were handled by replacing them with empty strings, and tokenization was applied

to transform text into numerical sequences. These efforts resulted in a clean, balanced dataset, significantly enhancing the model's accuracy and performance.

Statement_without_stopwords	Status
find coming depression never truly goes away g	Depression
say called doctor told stopping 10mg Escitalopram	Depression
f***Ing stupid want kill death anxiety means w	Suicidal
think take year two	Normal
Affronting worlds edit I’m native directly trans	Personality disorder
Hi all Let introduce 22y old male student deign	Depression
anyone else constantly think fact ‘BP makes r	Bipolar
get ticket	Normal
live children overcome try keep alive two beauty	Suicidal
don't ability cope anymore I'm trying lot things	Stress

Table 4.3.1: Our final dataset

4.4 Model Performance

4.4.1 GRU

The GRU (Gated Recurrent Unit) model effectively demonstrated its ability to handle the sequential nature of text data for emotion classification. It was trained on a balanced dataset over 10 epochs, with class weights applied to address class imbalance. Key metrics, such as precision, recall, and F1-score, were analyzed for each category. According to the classification report, the model achieved an overall accuracy of 70%. Categories like "Normal" and "Depression" performed well, achieving high precision and recall due to a sufficient number of samples. However, lower scores in categories such as "Stress" and "Personality Disorder" underscored the challenges posed by data imbalance.

The confusion matrix revealed that the model excelled at predicting "Normal" and "Depression" but struggled with misclassifications in closely related mental health categories like "Stress" and "Anxiety." Overall, the GRU model serves as a strong foundation, with potential for improved classification performance through further fine-tuning and data augmentation.

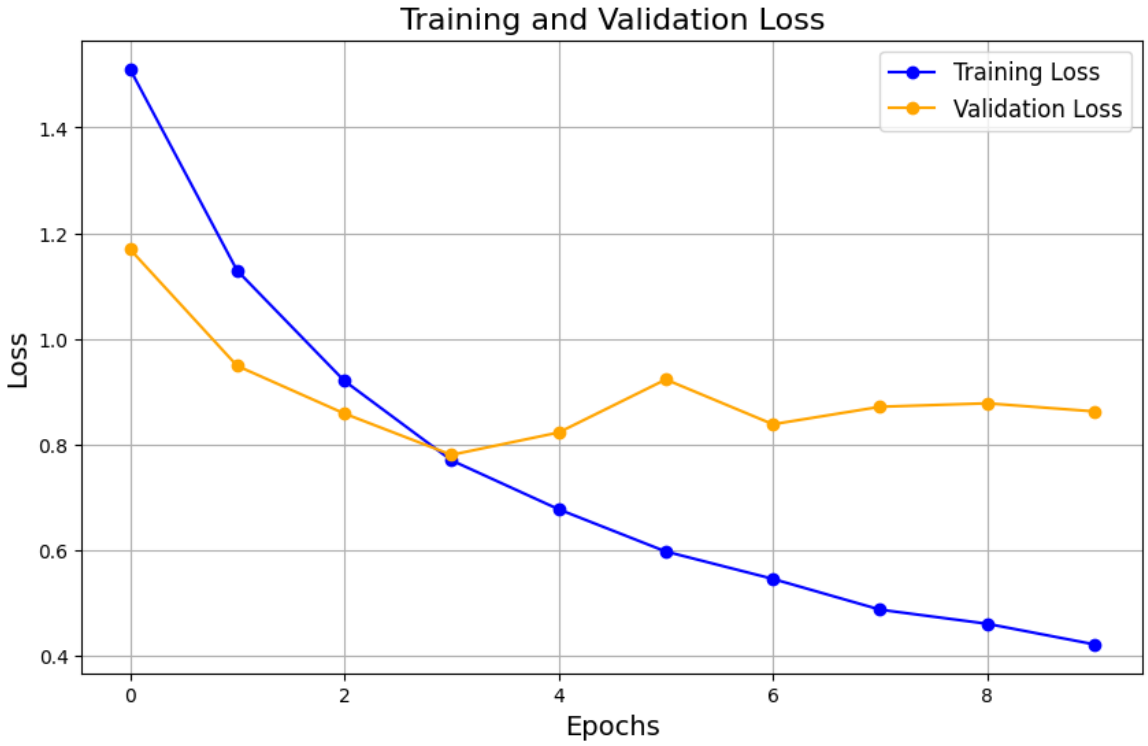


Figure 4.4.1.1: Training and Validation Loss

	Precision	Recall	F1-score	Support
Anxiety	0.62	0.68	0.65	779
Bipolar	0.55	0.59	0.57	580
Depression	0.74	0.56	0.64	3100
Normal	0.94	0.87	0.9	3327
Personality disorder	0.41	0.56	0.47	248
Stress	0.41	0.5	0.45	557
Suicidal	0.56	0.75	0.64	2018
accuracy			0.70	-10609
macro avg	0.6	0.64	0.62	10609
weighted avg	0.72	0.70	0.70	10609

Table 4.4.1.1: Classification report of GRU

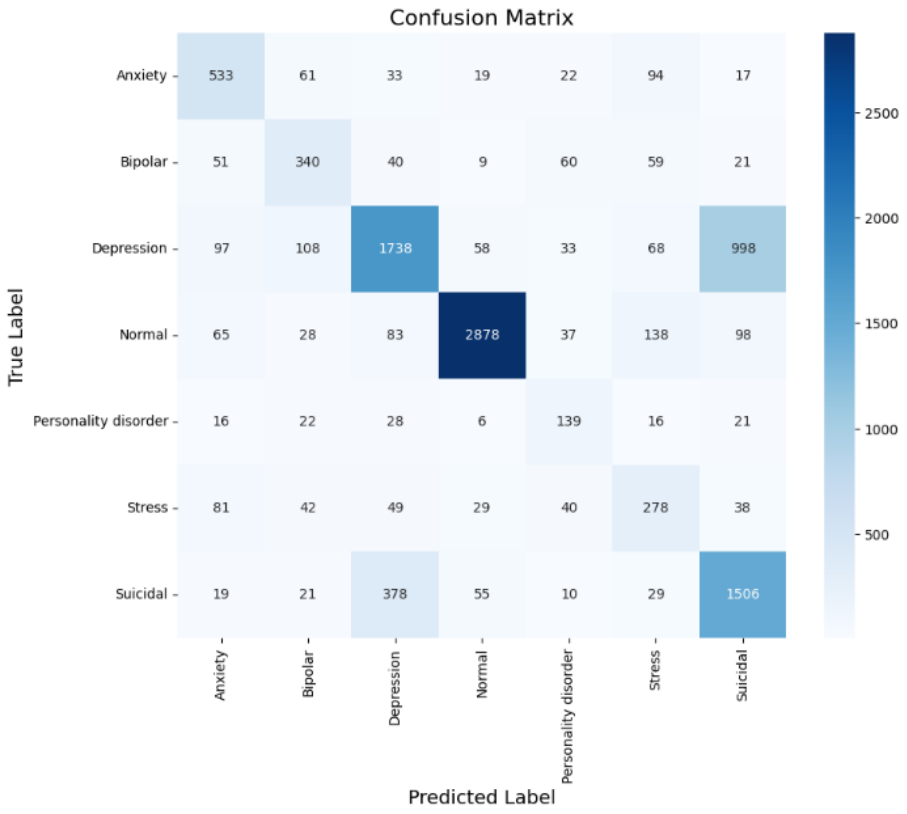


Figure 4.4.1.2: Confusion Matrix of GRU

4.4.2 LSTM

The LSTM model, for the sequential nature of the text data, worked pretty efficiently, with quite good general accuracy of 70%. This model could capture long-range dependencies in the textual sequences pretty effectively. In regard to the metrics of performance presented in the report for classification, great results were performed for such classes as "Normal" and "Depression" because there were more samples in those categories, so better precision, recall, and F1-scores were provided for them.

The results for categories like "Stress" and "Personality Disorder" in turn came out to be relatively lower, which means there is great room for improvement, especially with minority class imbalance issues. The confusion matrix suggested that while it provided an accurate performance for most labels, both for "Normal" and "Suicidal," misclassifications of some instances existed, especially among the classes having overlapping emotional states, such as anxiety/stress.

Given these challenges, the LSTM model still could classify emotions with reasonable accuracy and therefore stands out as a very promising candidate for fine-tuning and application in real-world Mental health monitoring systems.

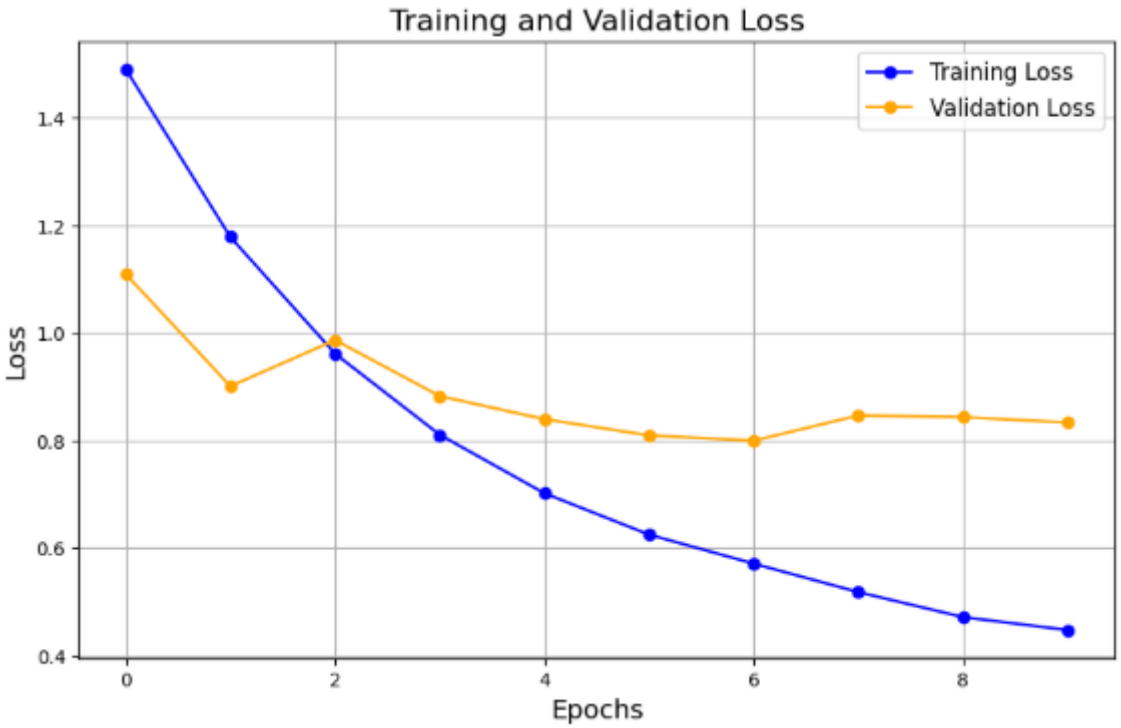


Figure 4.4.2.1: Training and Validation Loss

	Precision	*Recall*	*F1-score*	*Support*
Anxiety	0.7	0.64	0.67	779
Bipolar	0.54	0.62	0.58	580
Depression	0.75	0.51	0.61	3100
Normal	0.92	0.9	0.91	3327
Personality disorder	0.41	0.61	0.49	248
Stress	0.45	0.54	0.49	557
Suicidal	0.56	0.79	0.66	2018
accuracy			0.71	10609
macro avg	0.62	0.66	0.63	10609
weighted avg	0.73	0.70	0.71	10609

Table 4.4.2.1: Classification report of LSTM

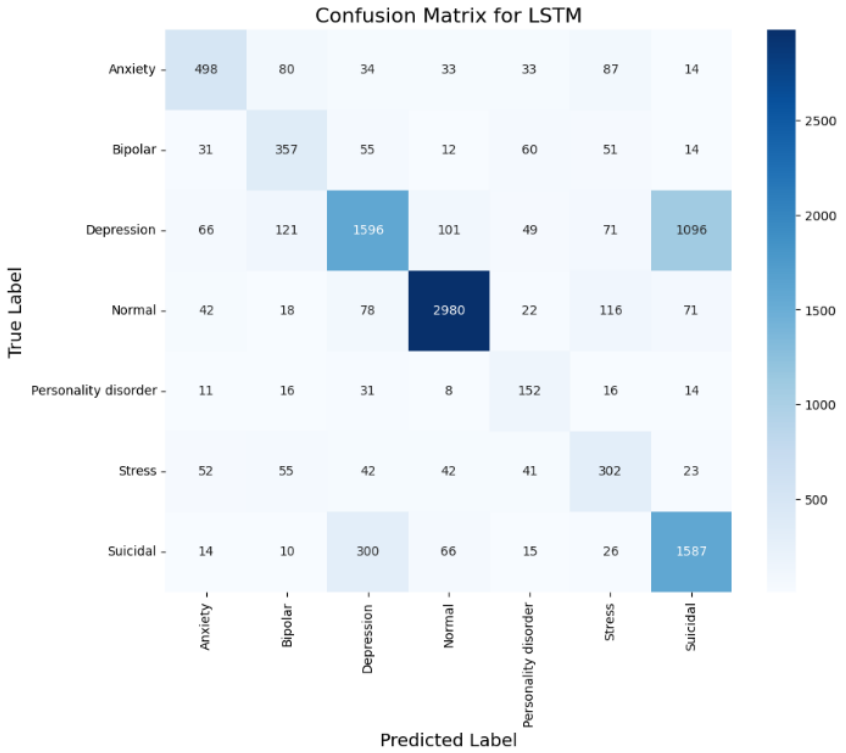


Figure 4.4.2.2: Confusion Matrix of LSTM

4.4.3 Roberta

The best performance among all models tested was given by the Roberta model, reaching an overall accuracy of 81%. The classification report is well segregated in better precision, recall, and F1-scores for classes like "Anxiety," "Bipolar," "Depression," and "Normal." Concretely, "Normal" and "Anxiety" states are very well detected, with F1-scores equal to 0.95 and 0.87, respectively.

Although the overall performance of the model is impressive, there are some shortcomings with regard to classifying minority classes such as "Personality Disorder" and "Stress," reflected by their relatively low F1-scores. The confusion matrix below indicates that, for most of the classes, correct classification was made but sometimes it went wrong—most of the times between the classes "Stress" and "Anxiety," and "Suicidal" and "Depression."

	Precision	***Recall***	***F1-score***	***Support***
Anxiety	0.85	0.9	0.87	779
Bipolar	0.75	0.86	0.8	580
Depression	0.81	0.68	0.74	3100
Normal	0.96	0.95	0.95	3327
Personality disorder	0.62	0.69	0.65	248
Stress	0.79	0.61	0.69	557
Suicidal	0.65	0.81	0.72	2018
accuracy			0.81	10609
macro avg	0.77	0.78	0.77	10609
weighted avg	0.82	0.81	0.81	10609

Table 4.4.3.1: Classification report of Roberta Model

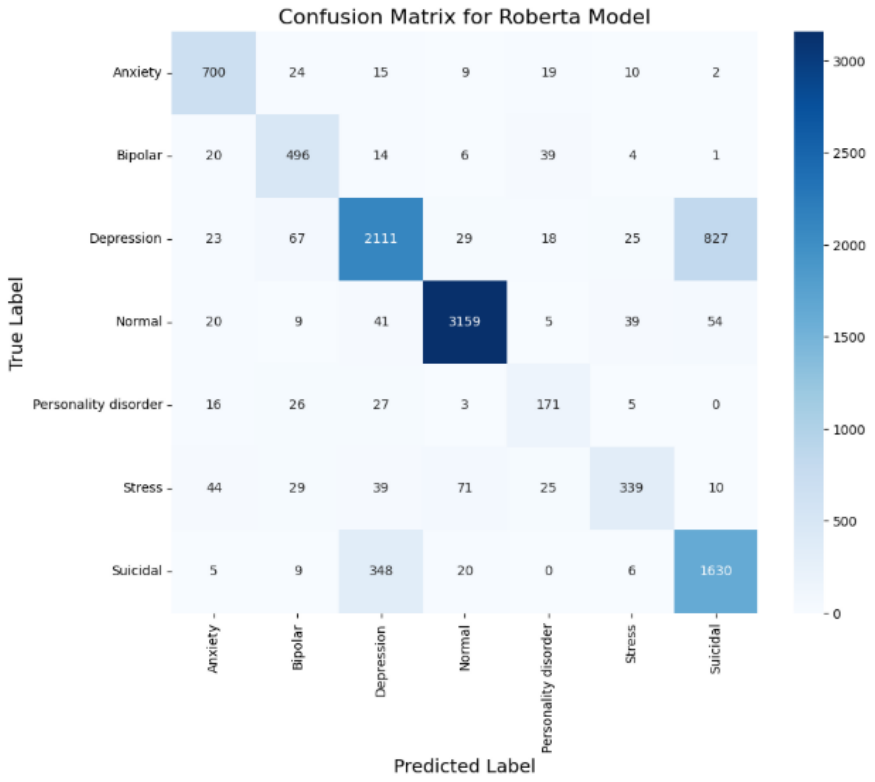


Figure 4.4.3.1: Confusion Matrix of Roberta

4.4.4 Logistic Regression

The Logistic Regression model performed well in the classification task of emotional states, with an overall accuracy of 77%. From the classification report, it can be seen that the model performed very well for categories like "Normal" and "Anxiety," since both precision and recall had high values owing to the large number of samples in the classes.

By contrast, results for the model were poor concerning minority classes: for instance, in the "Personality Disorder" and "Stress," metrics reflect very poor recall and low F1 scores. The confusion matrix therefore shows that regarding such major emotional category discrimination—"Depression" or "Suicidal"—the results by this particular model turned out to be remarkably brilliant, while there is misclassification in quite identical kinds of emotional states. As usual, misclassification occurred between them, particularly classes close together, like "Bipolar" and "Anxiety."

Overall, the Logistic Regression model presents a very solid baseline with a good balance between precision and recall for this task. The model should serve well for any emotion classification task. The model will be further improved by solving class imbalance and fine-tuning feature engineering.

	Precision	*Recall*	*F1-score*	*Support*
Anxiety	0.82	0.77	0.8	779
Bipolar	0.88	0.67	0.76	580
Depression	0.72	0.74	0.73	3100
Normal	0.86	0.96	0.91	3327
Personality disorder	0.64	0.52	0.58	248
Stress	0.73	0.49	0.58	557
Suicidal	0.68	0.66	0.67	2018
accuracy			0.77	10609
macro avg	0.76	0.69	0.72	10609
weighted avg	0.77	0.77	0.77	10609

Table 4.4.4.1: Classification report of Logistic Regression

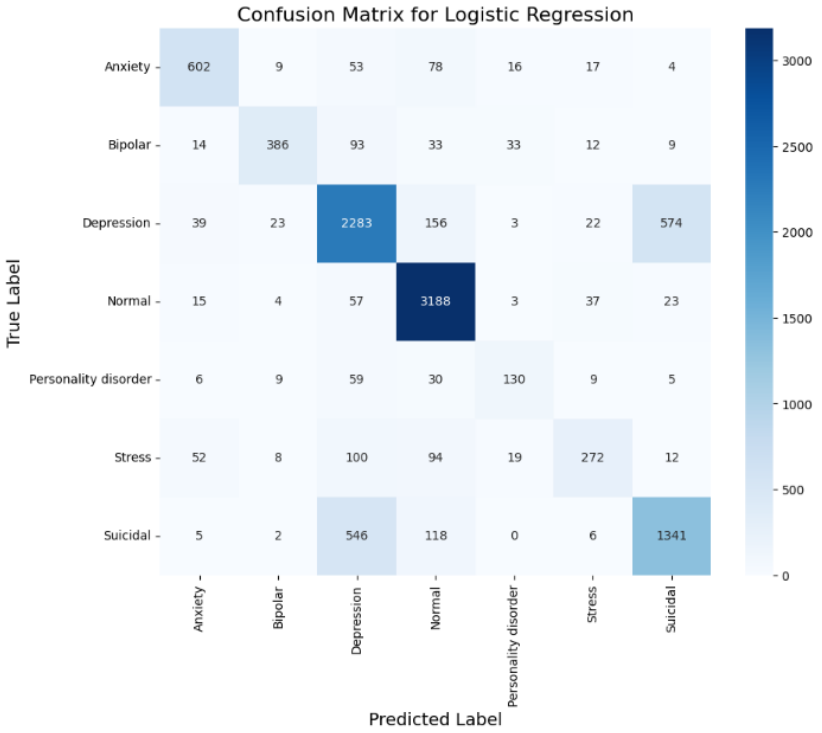


Figure 4.4.4.1: Confusion Matrix of Logistic Regression

4.4.5 Random Forest

This represents the performance of the Random Forest model on emotion classification, achieving an overall accuracy of 71%. According to the classification report, the model predicted reasonably well for the categories "Normal" and "Depression," whereas its performance in the "Stress" and "Anxiety" categories was not as strong, particularly under minority class conditions.

The confusion matrix showed that the model performed well in terms of classifying instances from dominant categories, while misclassifications in overlapping emotions, such as "Bipolar" and "Anxiety," were more frequent. While it could handle large datasets with strong predictive capabilities, further performance gains could be realized by balancing the dataset and fine-tuning hyperparameters.

Overall, Random Forest provides a strong foundation for emotion classification, but it needs further refinement to achieve better performance on underrepresented emotional states.

	Precision	*Recall*	*F1-score*	*Support*
Anxiety	0.87	0.53	0.66	779
Bipolar	0.96	0.46	0.62	580
Depression	0.58	0.8	0.67	3100
Normal	0.83	0.95	0.89	3327
Personality disorder	0.65	0.41	0.5	248
Stress	0.91	0.24	0.38	557
Suicidal	0.69	0.49	0.57	2018
Accuracy			0.71	10609
Macro avg	0.78	0.56	0.61	10609
Weighted avg	0.74	0.71	0.7	10609

Table 4.4.5.1: Classification report of Random Forest

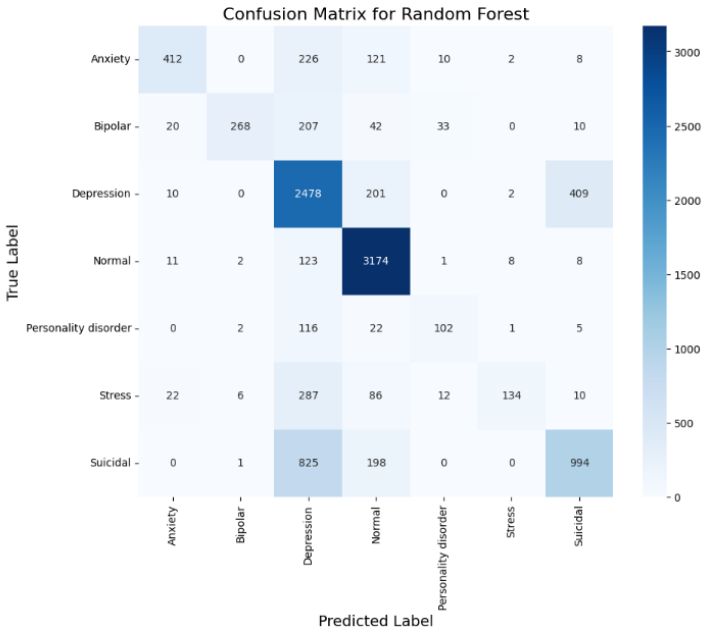


Figure 4.4.5.1: Confusion Matrix of Random Forest

4.5 Model Comparison

It gives the performance of the Random Forest model on emotion classification, with an overall accuracy of 71%. From the classification report, the model performed relatively well in classes "Normal" and "Depression," but performance in the "Stress" and "Anxiety" classes was not that good, especially in minority class conditions.

From the confusion matrix, it was very evident that the model performed well with classifying instances from dominant categories, while misclassifications in overlapping emotions, like "Bipolar" and "Anxiety," were higher. While it could deal with big datasets with good predictive capabilities, this performance could be further improved by balancing the dataset and hyperparameter tuning.

Overall, Random Forest provides a good baseline for emotion classification, but it needs further refinement to achieve better performance on underrepresented emotional states.

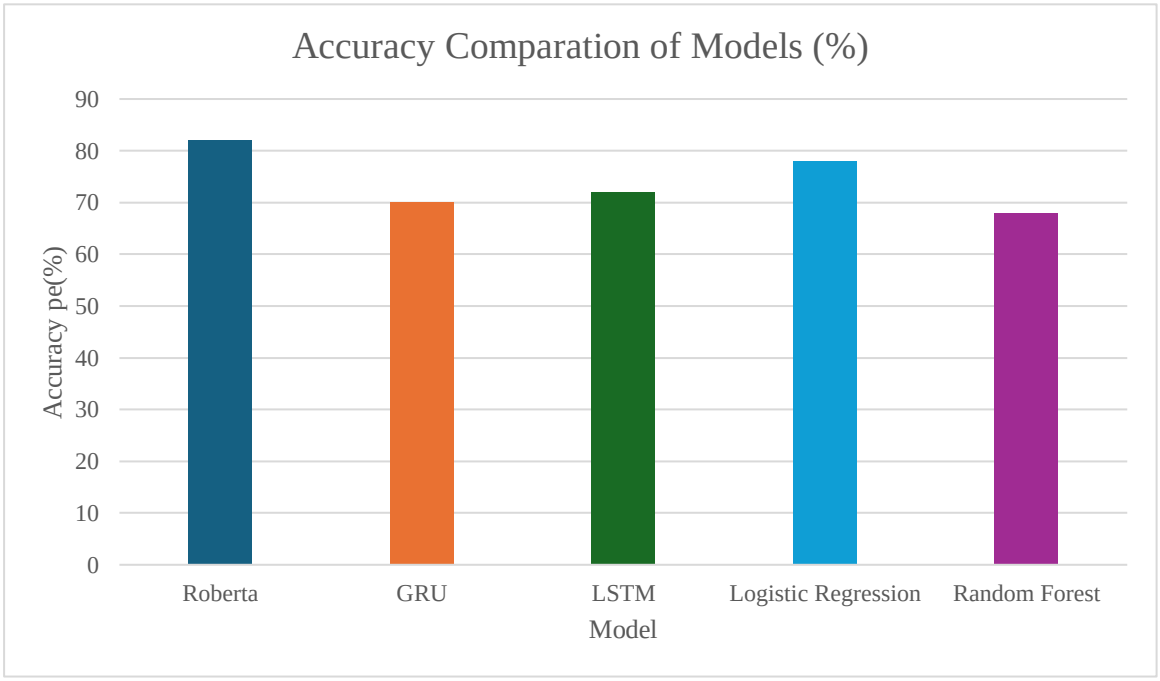


Figure 4.5.1: Model Comparison

CHAPTER 5

APPLICATION DEVELOPMENT PROCESS

5.1 Introduction

While access to modern technologies and an increasingly connected digital space has increased, anxiety, depression, and stress have become significant concerns, especially among today's youth. Social media and instant messaging platforms, although offering avenues for expression, often exacerbate these challenges. This application seeks to harness the power of machine learning and data analysis to monitor mental health through chat-based communication, providing insights that benefit individuals and society as a whole.

It addresses the issue of concealed emotional distress by analyzing and visualizing mental health states based on chat data. This enables early detection and timely intervention by identifying users experiencing negative emotional states, fostering a supportive digital environment. Real-time insights from the application aim to reduce mental health stigma, improve emotional well-being, and encourage young people to seek help and support one another. The following sections will outline the detailed development process of this application. These analysis tools empower users to better understand their own emotions and those of their friends, fostering a supportive and self-aware community. By doing so, the application promotes mental health and emotional resilience.

5.2 Entity-Relationship Diagram (ERD)

The development of this application relies on a carefully designed database structure. The ERD represents the relationships between core entities that enable the application's functionality.

- **Users**: Includes personal details such as name, email, and university, serving as the primary entity for system users.
- **Chats**: Tracks group and individual chat sessions, including participants, last messages, and unread counts.
- **Messages**: Stores chat content with timestamps and delivery status.
- **Analysis Reports**: Contains summaries of emotional analysis, including statistical breakdowns and category summaries.

The ERD ensures a clear data flow and seamless integration for real-time emotion detection and analysis.

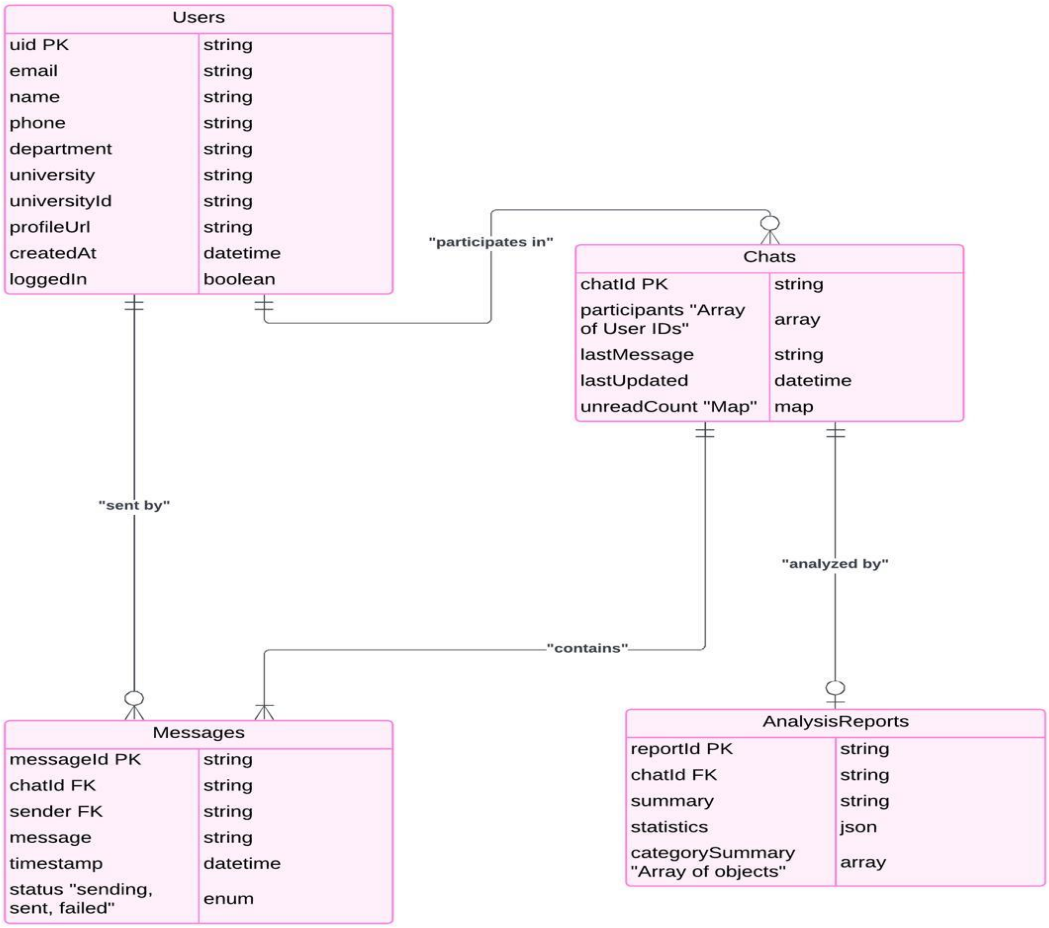


Figure 5.2.1: ER diagram

5.3 Analysis and Visualization

To offer meaningful insights, the application provides various analysis and visualization tools to monitor mental health based on chat data. These visualizations help users better understand their emotional states, identify potential risks, and provide support when needed. This section is divided into three key mental health overviews:

5.3.1 All Chat Mental Health Overview

This feature provides a comprehensive summary of emotional trends across all chat sessions. It includes the following visualizations:

1. **Condition Distribution (Pie Chart)**: Displays the proportion of each emotional category (e.g., Anxiety, Depression, Stress) across all chats.

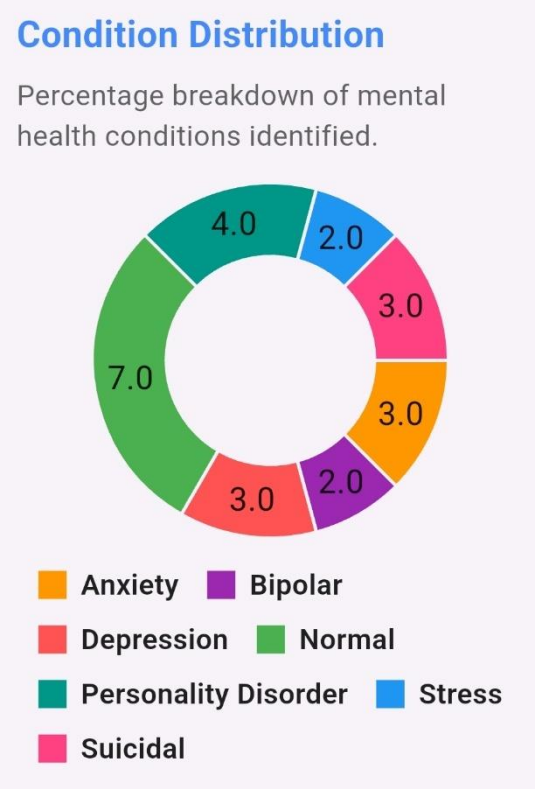


Figure 5.3.1.1: all chat condition distribution

2. **Detailed Analysis Percentage**: Highlights the percentage breakdown of emotional states, offering detailed insights into the data.

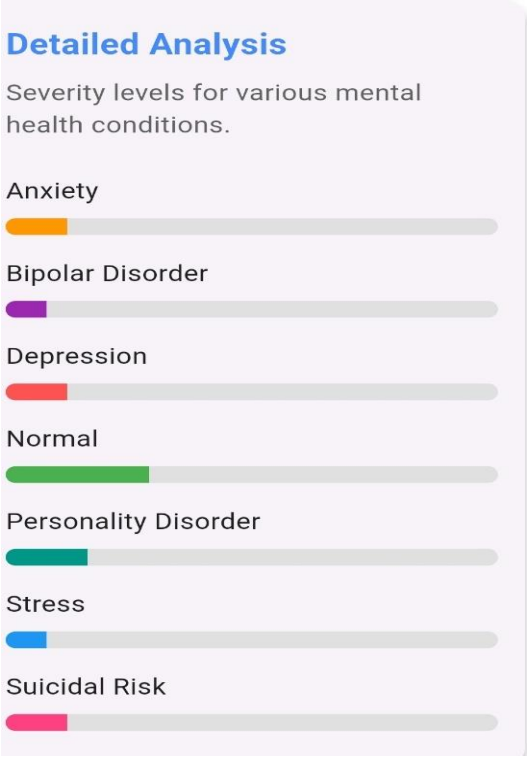


Figure 5.3.1.2: All chat Detailed Analysis

3. **7 Days Trend (Line Graph)**: Tracks emotional trends over the past week for all seven categories, showing fluctuations in emotional states.

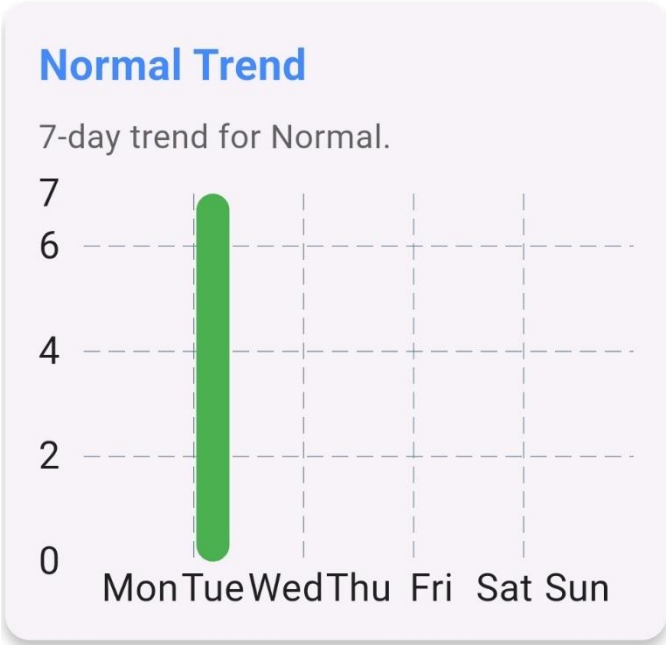


Figure 5.3.1.3: All chat normal trend

5.3.2 Self Chat Mental Health Overview

This section focuses on an individual user's mental health patterns based on their own chat data. The following visualizations are provided:

1. **Condition Distribution (Pie Chart)**: Displays the emotional state distribution of the user's own chats.

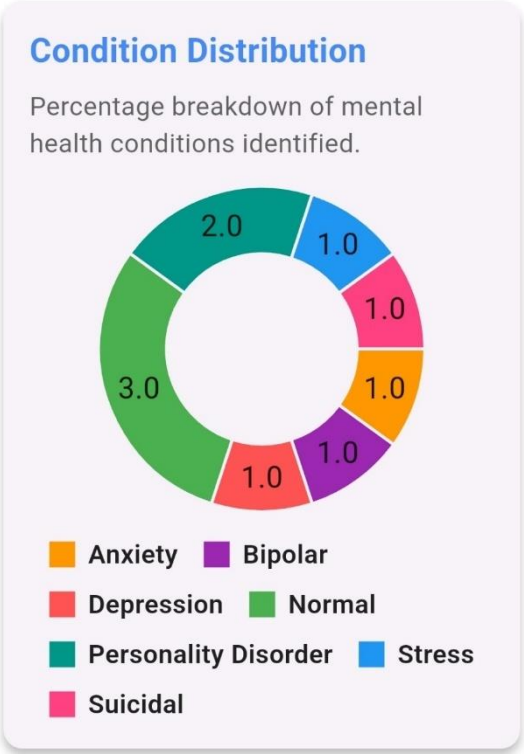


Figure 5.3.2.1: Self Chat Condition Distribution

2. **Detailed Analysis Percentage**: Provides an in-depth breakdown of the emotional categories detected in the user's chats.

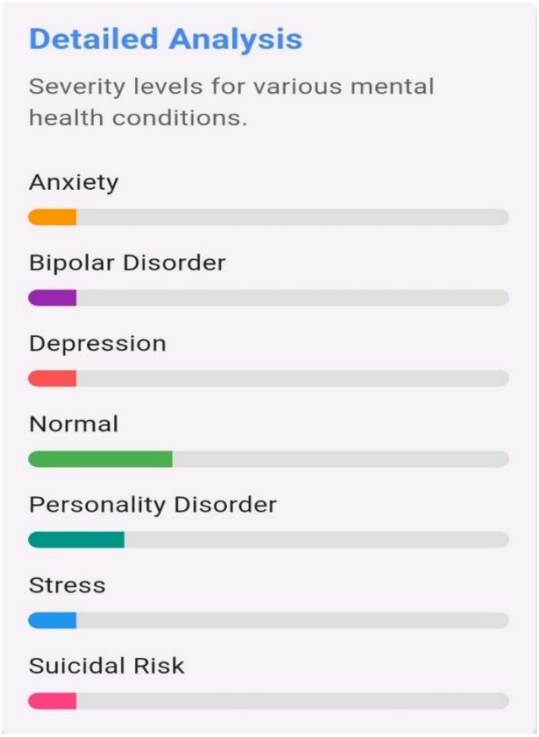


Figure 5.3.2.2: Self Chat Details Analysis

3. **7 Days Trend (Line Graph)**: Monitors changes in the user's emotional states over the last seven days, helping users track their mental well-being.

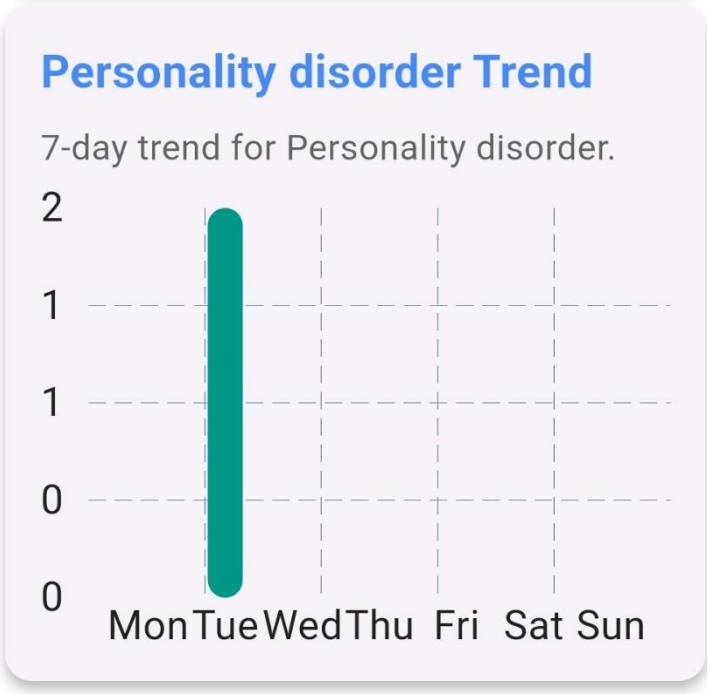


Figure 5.3.2.3: Self Chat Personally disorder Trend

5.3.3 Friends Chat Mental Health Overview

This feature offers insights into the mental health states of friends within chat groups. The visualizations include:

1. **Condition Distribution (Pie Chart)**: Shows the emotional distribution among friends' chats.

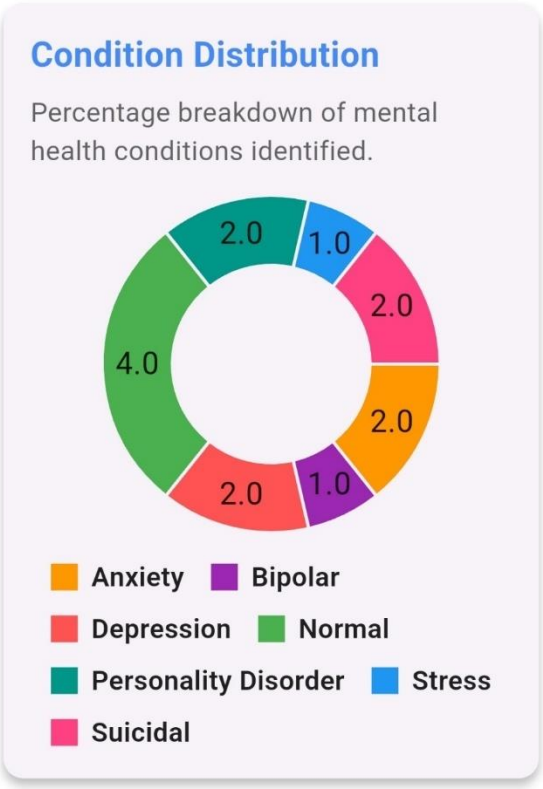


Figure 5.3.3.1: Friends Chat Condition Distribution

2. **Detailed Analysis Percentage**: Provides a percentage breakdown of each emotional state within the friend group.

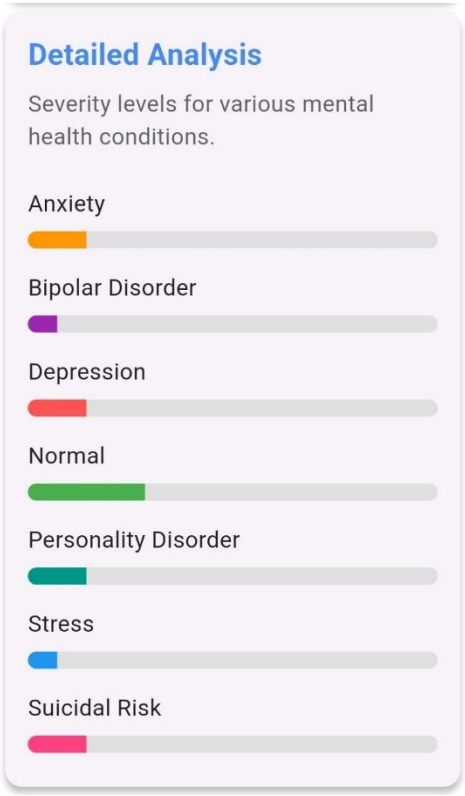


Figure 5.3.3.2: Friends Chat Details Analysis

3. **7 Days Trend (Line Graph)**: Tracks emotional trends within the friend group over a seven-day period, highlighting changes in collective emotional health.

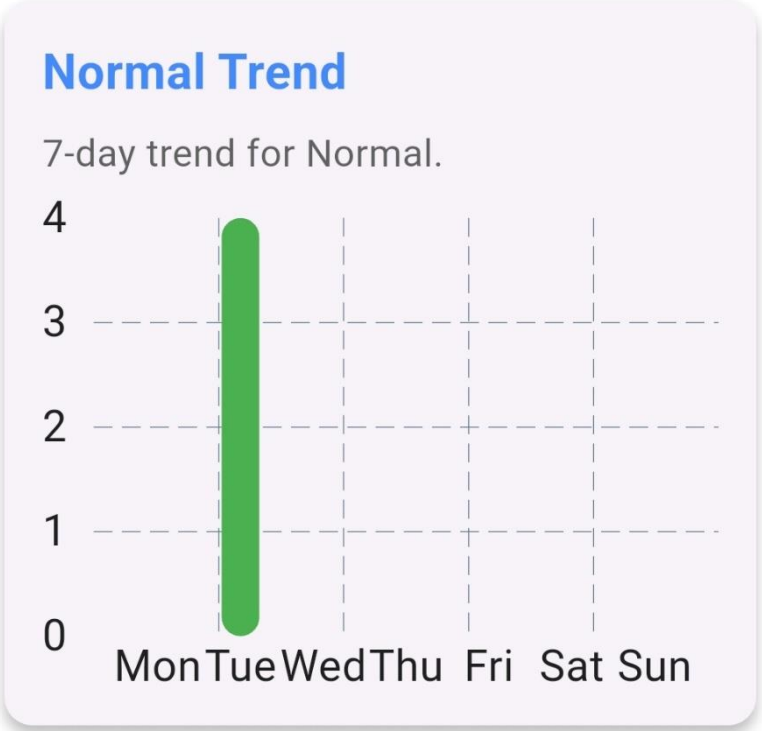


Figure 5.3.3.3: Friends Chat Normal Trend

5.4 Technology Stack

The development and implementation of this AI-powered mental health monitoring application relied on a carefully curated technology stack. This stack was selected to ensure scalability, efficiency, and robustness while addressing the functional and technical requirements of the system. Each component played a vital role in ensuring that the application met the needs of its users effectively.

5.4.1 Framework: Fintler Framework

The **Fintler Framework** served as the foundation for the seamless integration of machine learning models into the user interface. Its primary purpose was to bridge the gap between complex backend functionalities and the frontend interface.

Key features of the Fintler Framework include:

- **Real-Time Data Visualization**: Facilitates the real-time analysis and graphical visualization of emotional trends derived from chat data.

- **Modularity**: The framework's modular structure allowed for efficient implementation of additional features such as analytics and data logging.
- **Cross-Platform Support**: Ensured compatibility with mobile and web platforms, providing a unified user experience across devices.

5.4.2 Programming Language: Dart

The application was developed using **Dart**, a client-optimized programming language designed for developing fast and responsive user interfaces. Dart's compatibility with the **Flutter** framework was a significant factor in its selection.

Why Dart was chosen:

- **High Performance**: Dart's ahead-of-time (AOT) compilation ensured that the application was optimized for performance, offering a smooth user experience.
- **Rich Ecosystem**: Provided extensive libraries and tools that simplified the development process.
- **Cross-Platform Development**: Allowed developers to write a single codebase that could run seamlessly on both Android and iOS platforms, ensuring consistency in functionality and appearance.
- **Stateful Hot Reload**: Enhanced productivity during development by allowing changes to be instantly visible without restarting the entire application.

5.4.3 Backend API: FastAPI

To handle the backend operations of the application, **FastAPI** was employed. It is a modern web framework for building APIs with Python that is known for its speed and efficiency.

The role of FastAPI in this application:

- **Machine Learning Model Integration**: FastAPI served as the gateway for integrating the predictive model (e.g., Roberta), enabling emotion classification from chat inputs.
- **High Performance**: Its asynchronous nature ensured that multiple requests could be handled concurrently, reducing latency and enhancing the user experience.
- **Ease of Development**: FastAPI's automatic generation of interactive API documentation using OpenAPI specifications simplified testing and debugging.
- **Scalability**: FastAPI's lightweight design ensured that the backend could scale to handle increasing user demands and concurrent predictions.

5.4.4 Database: Firebase Firestore

The application's data storage needs were addressed using **Firebase Firestore**, a real-time NoSQL database. Firebase was chosen for its ability to synchronize data across users and devices instantaneously, ensuring smooth and efficient interactions.

Key benefits of Firebase Firestore:

- **Real-Time Synchronization**: Updates made by users were reflected instantly across all devices, maintaining consistency and accuracy in chat data.
- **Scalability**: Its horizontally scalable architecture enabled the application to handle large volumes of user data without compromising performance.
- **Secure Data Storage**: Firebase provides built-in security rules to restrict access to sensitive user information, ensuring compliance with privacy standards.
- **Offline Capabilities**: Firestore allowed data to be accessed and updated even in offline mode, with changes synchronized once the device reconnected to the internet.
- **Seamless Integration with Other Firebase Services**: By leveraging services like Firebase Authentication, the application offered a secure and efficient way to manage user access.

5.5 Libraries and Tools

The libraries and tools employed in the application's development were critical to achieving its functionality, aesthetics, and user experience. Each library was selected for its specific utility and seamless integration with the chosen framework and programming language.

5.5.1 Libraries

1. **cupertino icons: ^1.0.8**
    a. **Purpose**: Provides a comprehensive set of high-quality icons that follow Apple's Cupertino design guidelines.
    b. **Role in the Application**:
        i. Ensures a visually consistent and professional look for iOS users.
        ii. Enhances the overall aesthetics of the application by offering elegant iconography.
        iii. Examples of Use: Navigation icons, buttons, and status indicators.

2. **firebase_core**

a. **Purpose**: Serves as the foundational library for integrating Firebase services into Flutter applications.
b. **Role in the Application**:
    i. Acts as the entry point for initializing Firebase within the application.
    ii. Ensures a seamless connection between the app and Firebase backend services such as Firestore and Firebase Authentication.
    iii. Examples of Use: Loading Firebase configuration and enabling other Firebase libraries to function.

3. **firebase_auth**
    a. **Purpose**: Facilitates secure user authentication using Firebase's authentication services.
    b. **Role in the Application**:
        i. Provides tools for implementing various authentication methods, such as email/password, Google, or social media logins.
        ii. Ensures the security and privacy of user accounts with Firebase's built-in security protocols.
        iii. Examples of Use: Login screens, user registration, and session management.

4. **cloud_firestore**
    a. **Purpose**: Enables interaction with the Firebase Firestore NoSQL database for data storage and retrieval.
    b. **Role in the Application**:
        i. Supports real-time synchronization of user data, such as chat messages and emotional analysis results.
        ii. Provides efficient querying and filtering capabilities for managing large datasets.
        iii. Examples of Use: Storing and retrieving chat data, emotional analysis summaries, and user details.

5. **fl_chart**
    a. **Purpose**: A charting library for creating beautiful and interactive data visualizations in Flutter applications.
    b. **Role in the Application**:
        i. Used to generate charts that visually represent emotional trends and user analytics.
        ii. Supports a variety of chart types, including line graphs, bar charts, and pie charts, for displaying data insights effectively.

iii. Examples of Use: Displaying emotional trends over time, category-wise breakdowns of emotions.

6. **flutter_svg**
    a. **Purpose**: A library for rendering scalable vector graphics (SVG) images in Flutter applications.
    b. **Role in the Application**:
        i. Ensures high-quality images that scale seamlessly across devices with varying screen sizes and resolutions.
        ii. Provides flexibility for using vector assets such as icons, logos, and illustrations.
        iii. Examples of Use: Adding high-resolution graphics to dashboards and analysis screens.

7. **flutter_launcher_icons: ^0.14.2**
    a. **Purpose**: Simplifies the process of generating and managing app launcher icons for multiple platforms.
    b. **Role in the Application**:
        i. Automates the creation of platform-specific launcher icons, ensuring they meet the required dimensions and design guidelines.
        ii. Adds a professional touch to the application by supporting custom branding through unique launcher icons.
        iii. Examples of Use: Setting the app's launcher icon for Android and iOS.

8. **intl**
    a. **Purpose**: Provides tools for internationalization (i18n) and localization (l10n) in Flutter applications.
    b. **Role in the Application**:
        i. Allows the application to adapt to different languages, date formats, and number formats based on user preferences or geographic location.
        ii. Ensures a personalized experience for users from diverse backgrounds.
        iii. Examples of Use: Formatting dates and times, displaying messages in multiple languages.

5.5.2 Summary

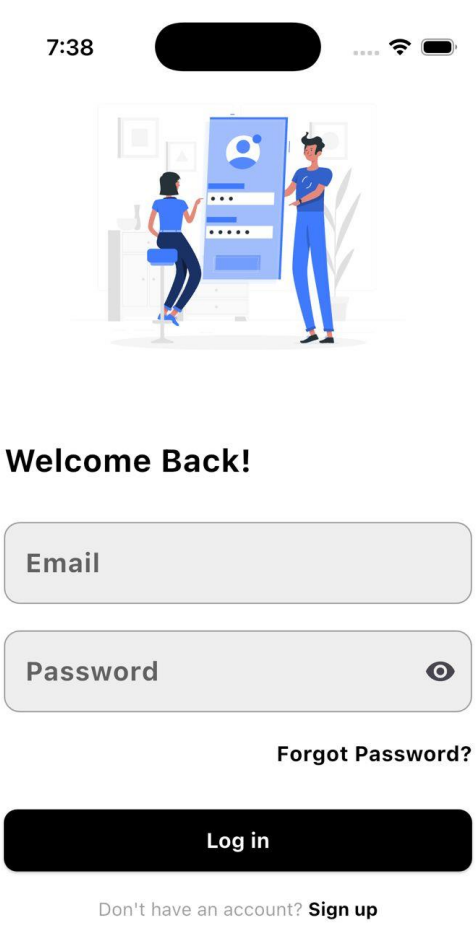
These libraries collectively enhanced the application's functionality and user experience. They ensured:

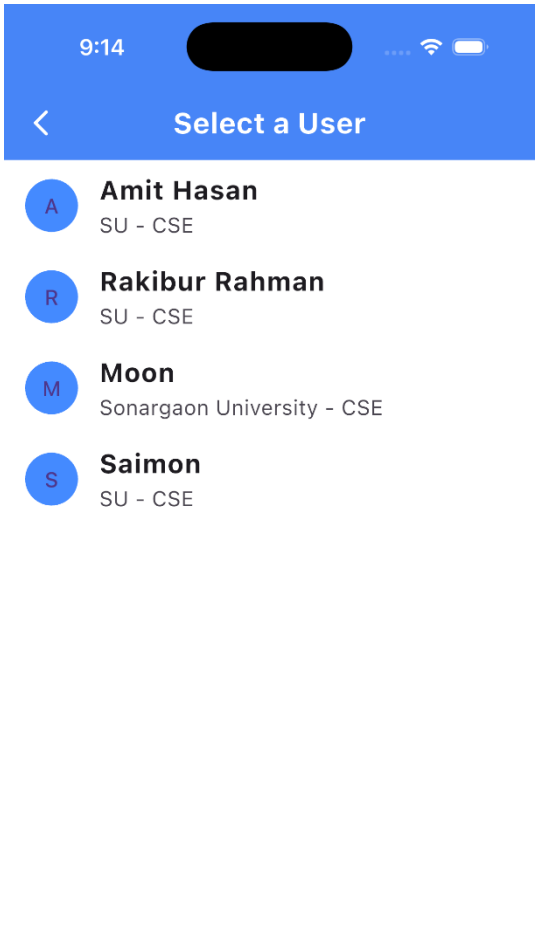
- Smooth and secure data interactions (Firebase libraries).
- Visually appealing and intuitive designs (cupertino_icons, flutter_svg, fl_chart).
- Seamless user management (firebase_auth).
- A global-ready application with internationalization features (intl).

This strategic combination of libraries contributed significantly to achieving the application’s goals and delivering a high-quality user experience.

5.6 Interface

The interface of the AI-powered mental health monitoring application is designed to provide a user-friendly and intuitive experience. It seamlessly integrates advanced emotion detection features into a visually appealing and responsive layout. The application offers real-time insights into users' emotional states through interactive graphs, pie charts, and trend visualizations. Key sections include personalized dashboards for self-monitoring and detailed overviews of friends' emotional well-being, ensuring accessibility and clarity. Screenshots below highlight the app's sleek design and functional elements, emphasizing simplicity and ease of navigation for users of all technical backgrounds. This interface aims to foster emotional awareness and proactive support.


7:38
Welcome Back!
Email
Password
Forgot Password?
Log in
Don't have an account? Sign up


9:14
Select a User
A
Amit Hasan
SU - CSE
R
Rakibur Rahman
SU - CSE
M
Moon
Sonargaon University - CSE
S
Saimon
SU - CSE

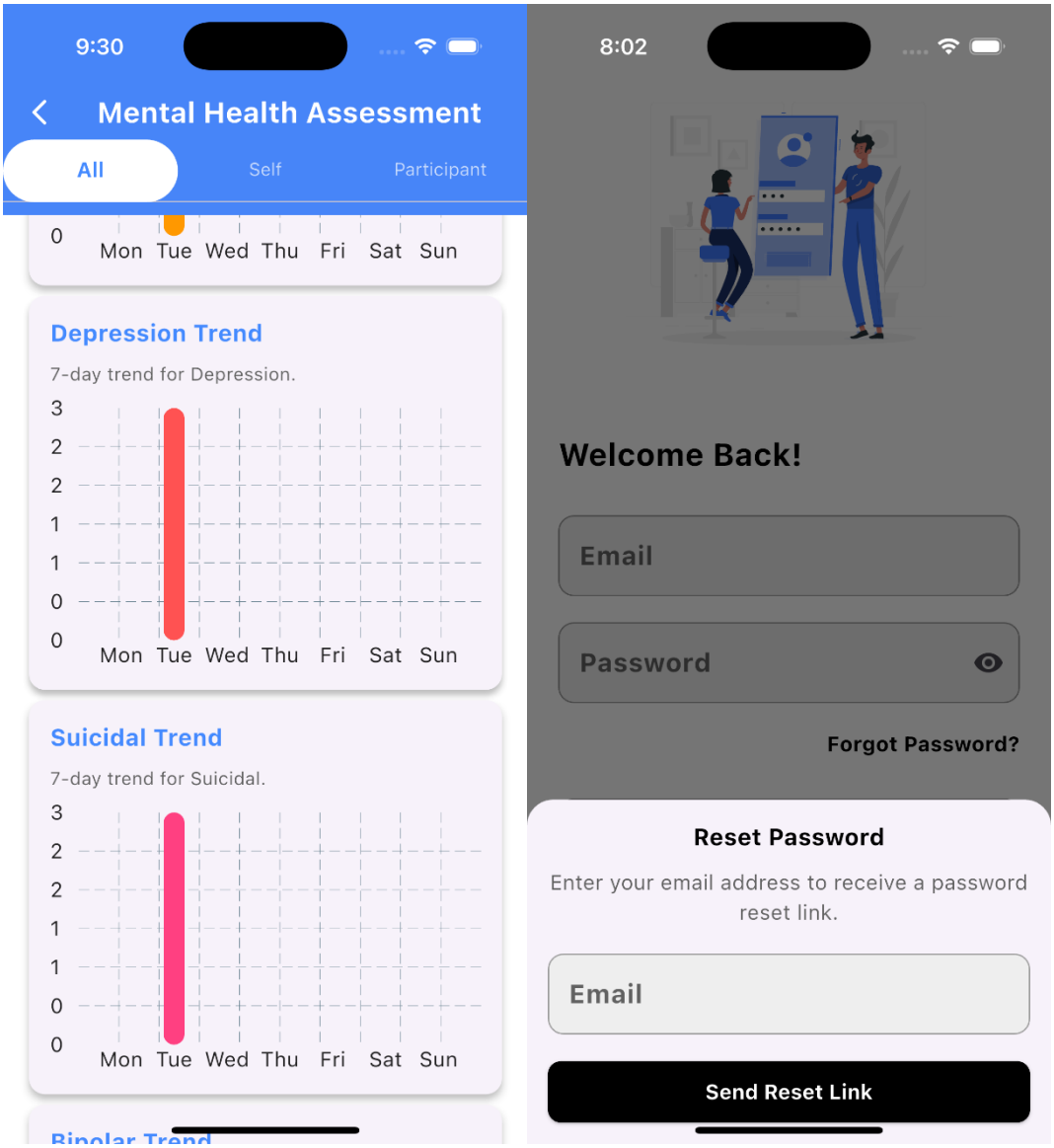


Fig 5.6 : Interface

CHAPTER 6

CONCLUSION AND FUTURE WORKS

6.1 Conclusion

Our research points to a new direction in addressing emotional challenges among the younger generation. By leveraging the latest machine learning models and embedding them into a user-friendly application, we aim to create an environment that supports emotional well-being. This system not only enhances social interactions but also serves as a tool for early detection and intervention, contributing to a healthier and more emotionally aware society.

This study addresses the critical problem of emotion detection from textual data, with a focus on the mental health of the younger generation. Using a dataset of over 53,000 sentences categorized into seven emotional states—Anxiety, Normal, Depression, Suicidal, Stress, Bipolar, and Personality Disorder—we evaluated various machine learning and deep learning models, including Roberta, GRU, LSTM, Logistic Regression, and Random Forest. Among these, Roberta demonstrated the highest accuracy of 81%, showcasing its effectiveness in understanding contextual information.

The processed and refined data enabled reliable emotion classification, which has been integrated into a user-friendly chat application. This application provides real-time insights into users' mental health, facilitating timely support and intervention. The proposed system has the potential to make a significant impact by promoting mental well-being, fostering awareness, and enabling early detection of emotional distress.

6.2 Limitation

While this research successfully demonstrates the development and implementation of an AI-powered mental health monitoring application, several limitations were identified that should be addressed in future studies:

1. Dataset Limitations

- **Language Bias**: The dataset used for training and testing the models primarily consisted of English text, limiting the applicability of the system to non-English-speaking users. The lack of a robust Bangla dataset also restricted the system's ability to cater to regional users effectively.
- **Class Imbalance**: Some emotional categories, such as "Stress" and "Personality Disorder," were underrepresented in the dataset, which may have affected the model’s performance in accurately classifying these emotions.
- **Contextual Constraints**: The dataset did not account for multimodal data (e.g., audio, images, or video), which could provide richer context for detecting emotions.

2. Model Performance

- **Challenges with Minority Classes**: Despite achieving high accuracy for majority classes like "Normal" and "Depression," the models struggled with minority classes, such as "Stress" and "Personality Disorder," leading to lower precision and recall scores.
- **Dependency on Text Quality**: The models' performance heavily depends on the quality and clarity of the input text. Informal language, slang, or ambiguous phrasing in real-world chats may reduce prediction accuracy.

3. Limited Real-Time Capabilities

- The system currently performs predictions on pre-processed chat data, which may introduce a slight delay in real-time scenarios. Enhancing real-time processing capabilities would improve the system's usability.
-

4. Generalization Issues

- The models were trained and evaluated on a specific dataset and may not generalize well to unseen data from other domains, such as professional communications or academic texts.
- Emotional context may vary significantly across cultures and user demographics, which the current system does not address.

5. Lack of Multimodal Integration

- The study focused solely on textual data for emotion detection, excluding other valuable modalities such as speech, facial expressions, or physiological signals that could enhance the accuracy and depth of emotion recognition.

6. User Privacy and Ethical Concerns

- The application stores and processes sensitive user data, raising potential privacy concerns. Although Firebase provides strong security measures, the study did not address additional ethical considerations, such as informed consent and user control over their data.
- Lack of explicit mechanisms to prevent misuse of emotional insights (e.g., manipulation or discrimination) by third parties.

7. Scope of Application

- The current system is designed for individual and group chat-based applications and may not be easily extended to other platforms, such as social media or workplace communication tools, without significant modifications.

6.3 Future Work

While the proposed system achieved promising results, there are several areas for improvement and expansion. Future work could include the following:

1. **Expanding the Dataset**: Incorporating multilingual datasets, especially in Bangla, to make the system more inclusive and applicable to a wider audience.
2. **Improving Model Performance**: Exploring advanced transformer-based models such as BERT or fine-tuning Roberta further to achieve higher accuracy and better handling of minority classes like "Stress" and "Personality Disorder."
3. **Real-Time Monitoring**: Enhancing the application to support real-time emotion monitoring and continuous data collection to adapt to users' evolving emotional states.
4. **Cross-Platform Integration**: Expanding the application to support multiple platforms, such as web, mobile, and desktop, using frameworks like Flutter for seamless integration.
5. **Incorporating Additional Features**: Adding speech-to-text emotion detection and facial emotion recognition to broaden the scope of emotion monitoring.
6. **Mental Health Support Features**: Integrating mental health resources, including automated recommendations, relaxation techniques, and professional contact options, to provide comprehensive support to users in distress.

REFERENCE

[1] G. Pattun and P. Kumar, "Emotion Classification using Generative Pre-trained Embedding and Machine Learning," *2023 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT)*, San Salvador, El Salvador, 2023, pp. 1-6, doi: 10.1109/ICMLANT59547.2023.10372980.

[2] R. Padate, A. Gupta, P. Chakrabarti and A. Sharma, "Emotion Recognition from WhatsApp Text Messages Using Unsupervised Machine Learning," *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Kirtipur, Nepal, 2024, pp. 1871-1876, doi: 10.1109/I-SMAC61858.2024.10714817.

[3] R. Padate, A. Gupta, P. Chakrabarti and A. Sharma, "Emotion Recognition from WhatsApp Text Messages Using Unsupervised Machine Learning," *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Kirtipur, Nepal, 2024, pp. 1871-1876, doi: 10.1109/I-SMAC61858.2024.10714817.

[4] R. R. Malagi, Y. R, S. P. T. K, A. Kodipalli, T. Rao and R. B. R, "Emotion Detection from Textual Data Using Supervised Machine Learning Models," *2023 4th International Conference for Emerging Technology (INCET)*, Belgaum, India, 2023, pp. 1-5, doi: 10.1109/INCET57972.2023.10170212.

[5] R. K. Kaliyar, K. Ram, A. Sharma, S. Tiwari, N. Ahuja and M. Agrawal, "Affects in Tweets with Real Time Emotions using Deep Learning Techniques: A Novel Approach," *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 2020, pp. 17-21, doi: 10.1109/Confluence47617.2020.9057913.

[6] G. S. M. Mursalin, S. D. Bappon and M. I. Khan, "A Deep Learning Approach for Recognizing Textual Emotion from Bengali-English Code-Mixed Data," *2022 25th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, 2022, pp. 330-335, doi: 10.1109/ICCIT57492.2022.10054872.

[7] N. V. Vinay, C. S. Kumar, J. Asipalli, M. Mouli and G. Sindhu Madhuri, "Detection of AI Empathy Using Deep Learning," *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, Bangalore, India, 2023, pp. 1-5, doi: 10.1109/CSET58993.2023.10346939.

[8] W. Wang and W. Ding, "Research of improved SVM model based on GA in E-learning emotion classification," *2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems*, Hangzhou, China, 2012, pp. 919-923, doi: 10.1109/CCIS.2012.6664310.

[9] A. Pandiaraj, P. Nagaraj, B. V. Durga, A. Tejasri, M. T. Teja and M. Navi Venkata Pavan Kumar, "A Comparative Study on Emotion AI using Machine

Learning and Deep Learning Models," *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2023, pp. 900-904, doi: 10.1109/ICACCS57279.2023.10113020.

[10] H. A. Sayyed, S. Rushikesh Sugave, S. Paygude and B. N Jazdale, "Study and Analysis of Emotion Classification on Textual Data," *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatre, India, 2021, pp. 1128-1132, doi: 10.1109/ICCES51350.2021.9489204.

[11] S. Kataru, K. King and L. Fernando, "Machine Learning-Based Early Detection and Intervention for Mental Health Issues in Children," *2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)*, Osaka, Japan, 2024, pp. 2001-2007, doi: 10.1109/COMPSAC61105.2024.00320

[12] Y. Ren, Q. Min and A. Jain, "Analysis of College Students' Mental Health Based on Machine Learning Algorithm," *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*, Raichur, India, 2024, pp. 1-4, doi: 10.1109/ICICACS60521.2024.10498768.

[13] H. K. Marrapu, B. Maram and P. Reddi, "New Analytic Framework of Public Mental Health Prediction Using Data Science," *2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, Villupuram, India, 2022, pp. 1-6, doi: 10.1109/ICSTSN53084.2022.9761324.

[14] A. Mittal, L. Dumka and L. Mohan, "A Comprehensive Review on the Use of Artificial Intelligence in Mental Health Care," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1-5, doi: 10.1109/ICCCNT56998.2023.10308255.

[15] Y. Sun, "Mental Health Classification and Diagnosis System Based on Random Forest Algorithm," *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2024, pp. 1-6, doi: 10.1109/ICDSIS61070.2024.10594527.

[16] S. Jage, S. Chaudhari, M. Jatte, A. Mhatre and V. Mane, "Predicting Mental Health Illness using Machine Learning," *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIANCON58793.2023.10270445.

[17] V. Laijawala, A. Aachaliya, H. Jatta and V. Pinjarkar, "Classification Algorithms based Mental Health Prediction using Data Mining," *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2020, pp. 1174-1178, doi: 10.1109/ICCES48766.2020.9137856.

[18] S. Verma, C. Sharma, G. Aggarwal and P. Upadhya, "Artificial Intelligence-Based Approach for Classification and Prediction of Mental Health," *2024 14th International Conference on Cloud Computing, Data Science & Engineering*

(Confluence), Noida, India, 2024, pp. 708-713, doi: 10.1109/Confluence60223.2024.10463203.

[19] J. Gao and G. Shi, "Mental Health Evaluation of College Students Based on Similar Trajectory Clustering Algorithm," *2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Beihai, China, 2021, pp. 828-831, doi: 10.1109/ICMTMA52658.2021.00189.

[20] N. Pritam, K. S. Gill, M. Kumar, R. Rawat and D. Banerjee, "Classification of Student Mental Health Analysis using Logistic Regression and other classification techniques through Machine Learning Methods," *2024 3rd International Conference for Innovation in Technology (INOCON)*, Bangalore, India, 2024, pp. 1-5, doi: 10.1109/INOCON60754.2024.10512216.

[21] R. A. Rahman, F. Haziqah Mohamad Zaini, M. S. Nizam Mohd Danuri and A. Amin, "The Sentiment Analysis on Mental Health Awareness by Non-Governmental Organisation's Twitter," *2022 International Visualization, Informatics and Technology Conference (IVIT)*, Kuala Lumpur, Malaysia, 2022, pp. 185-190, doi: 10.1109/IVIT55443.2022.10033345.