A Hybrid AI-Based Model for Depression Detection in Bangla Social Media Posts Using the BSMDD Dataset

By

Pijush Chandro Roy
ID: 183-31-260

This Thesis Report Presented in Partial Fulfilment of the Requirements of the Degree of Masters of Science in Electronics and Telecommunication Engineering

Supervised by

Md. Taslim Arefin
Associate Professor and Head
Department of ICE
Faculty of Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

January, 2025

APPROVAL

This research titled **"A Hybrid AI-Based Model For Depression Detection In Bangla Social Media Posts Using The BSMDD Dataset "** submitted by **Pijush Chandro Roy** to the department of Information and Communication Engineering (ICE), Daffodil International University, has been accepted as satisfactory for the partial fulfilments of the requirements for M.Sc. in Electronics and Telecommunication Engineering (ETE) and approved as to its style and contents. The presentation was held on 25th January 2025.

Board of Examiners

Md. Taslim Arefin — **Chairman**
Associate Professor and Head
Department of ICE
Faculty of Engineering
Daffodil International University

Professor Dr. A. K. M. Fazlul Haque — **Internal Examiner**
Professor
Department of ICE
Faculty of Engineering
Daffodil International University

Engr. Md. Zahirul Islam — **Internal Examiner**
Assistant Professor
Department of ICE
Faculty of Engineering
Daffodil International University

Md. Ashraful Alam Bhuiyan — **External Examiner**
Chief Technology Officer (CTO)
Bank Alfalah, Bangladesh

DECLARATION

This certifies that the thesis titled **"A Hybrid AI-Based Model for Depression Detection in Bangla Social Media Posts Using the BSMDD Dataset"** is the result of our study in partial fulfillment of the requirements for a master's degree in the Department of Electronics and Telecommunication Engineering under the supervision of (Mr. Md. Taslim Arefin), Research Lecturer, Department of Electronics and Telecommunication Engineering.

Supervised By

Md. Arefin Taslim
Associate Professor and Head
Department of ETE
Faculty of Engineering
Daffodil International University

Submitted By

Pijush Chandro Roy
ID: 183-31-260
Department of ETE
Faculty of Engineering
Daffodil International University

DEDICATION

This Thesis is dedicated to my parents and all of my respected teachers.

ACKNOWLEDGEMENT

I am deeply grateful to the multitude of individuals who have played pivotal roles in the success of this thesis. This academic journey has been incredibly enriching, not only in terms of professional knowledge but also on a personal level. I attribute my achievements to the unwavering guidance and support I have received throughout this endeavour.

I am particularly thankful to the divine for granting me the strength and resilience to complete this thesis successfully. It is through this divine guidance that I have been able to reach this significant milestone. Furthermore, I owe a debt of gratitude to my esteemed thesis supervisor, Md. Arefin Taslim Associate professor and Head Department of Electronic and Telecommunication Engineering at the University of Daffodil International University (DIU). Their mentorship, encouragement, and unwavering support have been instrumental in shaping the course of my research. Their thought-provoking inquiries have prompted me to explore concepts from various perspectives, ultimately enhancing the quality of my work. I am immensely grateful for their guidance and support throughout this transformative journey.

I also extend my thanks to the dedicated teaching staff of the department, whose constant encouragement, support, and guidance have contributed significantly to my academic growth. Your collective wisdom and commitment to education have been indispensable to my development.

Abstract

Depression is a significant healthcare problem worldwide, impacting countless people of all ages annually. This paper presents a novel approach for depression identification by using complex datasets and employing various machine learning models such as LightGBM, XGBoost, Naïve Bayes, and Random Forest. The study emphasizes how it is possible to change the face of mental health diagnosis with those models relying on the data. In the future, the analysis should aim at improving model interpretability, reducing bias in the algorithms, using multi-source data as well ethical issues such as rights of privacy and consent. These innovations are designed to help improve the diagnostic capabilities, promote inclusiveness in the outcomes, and support the means of real-time monitoring their health of the individual. Among the models tested, LightGBM proved to be the best as it achieved an accuracy of 83.67% and an F1 score of 84.06%. XGBoost had a higher accuracy rate while Naive Bayes had a higher recall rate. Random Forest also proved effective, performing well in all areas metrics which show their different benefits for dealing with issues in mental health care.

Chapter One: Introduction

Depression, which is one of the most common mood disorders, negatively affects the mental state of millions of people across the globe, manifested by persistent depressed mood, disinterest, and the inability to carry out any activities. The term depression comes from the Latin word “deprimere” [1], which means to press down, and it has been written about as far back as ancient scriptures from Hippocrates, and remains one of the biggest medical challenges. In a report compiled by The World Health Organization (WHO), it was revealed that more than 264 million people [2] equivalent to 3.4 percent of the world's population are depressed with the population aged 15-29 years being the most mostly affected. The increasing observed trend has also exacerbated the occurrence of high suicide rates, where 800,000 people [3] take their own lives every year making suicide the fourth main cause of death among the youth. The impact of depression is most emphasized in poorer low- and middle-income countries as society and the economy strengthen the difficulties associated with mental health.

In the modern context, where the world is more interconnected than ever before, people increasingly turn to social media for various issues rather than using them as a coping mechanism, which has become a new reality with social networking platforms as nine out of ten teenage girls aged 13-19 use Instagram as of 2021 showing the definition shift. Over the years, it has been observed that users spend more than 2 hours on user generated content websites such as Facebook, Twitter, and Pinterest. Looking at this high level of interaction there have been questions with regard to how social media impacts a person’s mental well-being. Research shows that excessive use of these platforms can exert anxiety, stress and depressive symptoms due to fear of missing out, inadequacy and social comparison [4], [5], [6], [7]. However, there are also instances where depression can be detected from social media [8], [9], [10], which can also help lift depressive feelings.

Social media content has grabbed people’s attention because many people use it to communicate about depression. For instance, people with depression can be active users of social media, posting photographs and text whilst continuously switching their behavior patterns. They can post content with negative tones, use more of negative words, and even use social media as an outlet when they can’t sleep. Identifying such social media patterns can help protect predisposed individuals by using chatbots, and other means that can positively manipulate them. While most research in this area focuses on English-language data, there is a growing need to extend these efforts to low-resource languages like Bangla, which is spoken by over 228 million people worldwide, primarily in

Bangladesh. Bengali, as a language of both culture and emotional depth, also poses a complexity and opportunity to detect depression. Bangladesh is a country that is in the course of developing and in fact a large proportion of social media active users create and share content in their vernacular. There is an urgent case to investigate Bangla content on social media as it provides an important perspective on the mental health situation in a country where depression is often not treated or spoken about. This necessitates building strong NLP systems specifically for Bangla which can deal with the phonological and syntactic diversity featured in the language.

Recent machine learning (ML) as well as deep learning (DL) models have been very much effective in traces of depressive thoughts in social media [11], [12], [13]. Naive Bayes, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) are some examples of classical ML models that have been used with textual features like Term Frequency-Inverse Document Frequency (TF-IDF) and linguistic markers. These approaches have been effective in depression language classification such as the high frequency of first-person pronouns, possessive pronouns, and negative words. By means of example, it is DL, particularly LSTM networks and their derivatives that are performing best as a problem is the sequential nature and contextual nature of the problem being aided by the text data.

In the context of Bangla depression detection, the researchers have come up with novel techniques to address the issues related to the low availability of linguistic resources. Mumu et el. [14] used a hybrid LSTM-Convolutional Neural Network (CNN) model that was intended to target the analysis of Bangladeshi social media posts. While the CNN layers were doing the feature extraction, the LSTM layers were responsible for sequential dependencies resulting into a good classification. Likewise, Uddinet el. in 2019, [15] constructed a five-layered LSTM model while others worked with GRUs and BiLSTMs to improve results. These trends suggest the feasibility of Bangla specific depression detection systems, especially when coupled with user behavior and emotional tone features. There are several issues related to the Bangladesh social media landscape including language barriers, low availability of datasets, and the unavailability of substantive content. There are several areas that need innovation that include the inconsistent use of Bangla and English language. The emergence of English transliterated Bangla also makes content creation robust. LSTM-based models perform faster but become ineffective as the sentence length increases. As Bangla social media primarily consists of long sentences, attention mechanisms would make a meaningful impact on the content creation process.

Clinical studies for detecting accurate measures of depression [16] have started to integrate the analysis of user-generated images alongside user behaviors. Relevant literature indicates that the combination of the two gives a better picture for depression analysis. Productive measures such as these may further diversify the understanding of how depression manifests itself in Bangladesh and the ways media services can be created to support this. There are serious implications associated with the development of Bangla depression detection systems. These systems may help to reduce the prevailing stigma around mental health in Bangladesh by detecting risk populations and enabling timely interventions. Social media networks can play a preventative measure by introducing AI-driven tools that can track and assess user activity along with recommending and presenting content tailored to users' needs. This appears as a response to international efforts towards use of technology in mental health promotion which is particularly important in the fast digitalizing world.

The ability to determine whether a content in Bangla posted on social media is written in the context of depression is a unique convergence of technology, language and mental health. This project has the potential to enhance the understanding of depression in the context of Bangladesh and its management by building on existing studies and language specific issues related to the Bangla language. The convergence of available AI tools and NLP for Bengali language has possibility to revolutionize the provision of care in the field of mental health resulting in a safe and friendly digital environment for everyone around. In this study we try to garner a machine learning model that will precisely detect depression and thus help in the research od mental health and NLP.

1.1 Background of Research

The World Health Organization describes depression as a mental disorder that is characterized by loss of interest, sadness, or fatigue [17]. It is one of the main contributors of disability in the world along with being many social and economic problems. Everyone regardless of age [15], gender [18] or social background [19] suffers from depression but people in low and middle income countries have it the worst because of lack of resources, poverty and stigma surrounding mental health. Additionally, many regions like Bangladesh remain in isolation when it comes to awareness and tackling mental health including depression which leads to it being underdiagnosed and undertreated.

AI and machine learning hasn't always been the case but in the recent years experts have been looking for AI or NLP to help tackle the ever-increasing demand for mental health services. Along with this, social media sites have also gradually become an important aspect of communication and an essential source of data for understanding emotions and mental states. With these platforms, users have the ability to post, comment and send messages allowing them to express themselves. Such expressions often reveal subtle indicators of psychological distress [20], making social media a valuable resource for depression detection.

Across a number of countries, it has been shown that machine learning (ML) and deep learning (DL) approaches are effective in extracting information from social media for gauging mental health status. These methods use language patterns, users' behavioral features, as well as complex data sets to recognize depressive signs. For instance, it has been shown that depressed individuals tend to use first-person pronouns more, have a higher inclination for negative visual images, as well as post interactions at unusual times, e.g. during the circadian trough. Such behaviors and language patterns support the concept of AI being adopted to provide tools to diagnose depression.

There is a trend in the literature that suggests that trawling for depression amongst the English has been done extensively by applying both traditional ML algorithms and deep learning models. Traditional machine learning approaches such as SVM and Random Forest are feature-based algorithms that use tf-idf and diversity in lexis as features. These algorithms are good, but they require lots of effort in coming up with parameters, and it may take a long time and be very specific to a given area. However, Deep Learning models do also include Long Short Term Memory (LSTM) networks, Gated Recurrent Units, (GRU), and Transformer based architectures that are becoming popular because of their ability of performinge high-level pattern recognition directly from data.

These models capture interdependencies in sequences and contextual information available in texts, therefore this would enable them to understand the semantics of depressive language better.

The research on mental health using social media in English is already moving on a larger scale, while similar initiatives in low-resourced languages such as Bangla are still under development. Despite being spoken by 228 million people, is the seventh most spoken language in the world yet it is still not represented in the academic research regarding Natural Language Processing. This language barrier creates several hurdles to mental health research in Bangladesh, where such content is more likely to be available on social media in abundance. A lot of Bangladeshi users post content in Bangla language using both Bangla and English script or a transliteration of both which adds a layer of complexity to text analysis. Therefore, while developing depression detection systems for banlgla language these social cultural and linguistics factors need to be taken into a count.

The last few years have seen several research attempts for Bangla depression detection aiming for various ML and DL techniques. To classify Bangla texts, features of the languages were used employing classical methods, such as Naive Bayes and SVM. For example, researchers have employed a lexicon of depression words and n-grams to study Bangla language social media postings. However, these approaches usually have difficulty in dealing with the informal and mixed-script aspects of Bangla social media text.

But the development of DL models brings new possibilities of detecting depression in Bangla speaking people. Long Short Term Memory (LSTM) networks which are known to be capable of dealing with long range dependencies in sequential data have been used for Bangla text analysis with great success. Researchers also combined LSTM with Convolutional Neural Networks (CNNs) for better performance. These models incorporate CNN layers for the extraction of spatial features and LSTM layers for the modeling of sequential information, hence, enhancing the analysis of the texts. Models for detecting Bangla depression have also been improved with the introduction of attention mechanisms that help to concentrate the models on the most relevant parts of the input sequence.

Even though there have been notable improvements, the scope of Bangla-specific NLP tools and resources remains grim. There is a lack of available annotated datasets and the available corpora are generally insufficient in diversity for robust model training. Also, preprocessing and feature extraction is problematic due to the informal structure of Bangla social media language which is code-switched and transliterated. In order to fit them into a Bangla context, complicated text normalization systems need to be constructed and pre-trained transformer-based models for Bangla are needed. The models such as BERT [21] and its multilingual variants have helped in improving the understanding of the context behind the Bangla language.

Multimodal data is another special focus in depression detection research. While a lot of research looks into text data alone, a number of other sources like images, videos and behavioral data provide more context with respect to mental health. For example, mood related features can be detected through visual image analysis employing CNN [22], while behavioral analysis can assist in detecting abnormal posting habits related to emotional issues of the users. For Bangla depression detection, the use of multi-modality could lead to improvement in operational effectiveness and robustness of AI powered models.

The outcomes of effective Bangla depression detection are significant. The use of AI and NLP technologies can help build systems that help screen and flag vulnerable individuals for help. Such systems can assist specialists in mental healthcare with basic non-invasive diagnoses and investigations thereby offloading some pressure from the healthcare systems. Furthermore, AI-powered systems can also help to change the stigma around depression by facilitating conversations about it and encouraging education through personalized tips and resources.

Nonetheless, the ethical implications of the use of depression detection systems also needs attention. Data security, consent, and bias in the algorithms are important issues when dealing with sensitive areas like mental health. These systems will only be effective and widely used if they are easy to understand, fair to everyone, and safe to use. In the case of Bangladesh, cultural and social issues hole up these developments so that they are relevant and sensitive to the stakeholders.

In brief, the existing research on Bangla depression detection lies at the intersection of mental health, technology and linguistics. While advancements made globally in depression detection are of great significance, the special features associated with Bengali language require a different focus. Addressing these challenges and utilizing the most advanced AI and NLP methods makes it possible to develop solutions that can potentially improve the lives of millions of speakers of Bengali language.

1.2 Motivation of the Work:

The increasing rate of depression along with its adverse effects on individuals and on society at large should be a cause of concern. This is even more worrying in Bangladesh as there is poor consciousness on mental health, poor funding and social issues. There are many people where conventional means of locating depression and fighting it are not viable, hence the search for new efficient and more universal solutions. The developing state of the art technologies especially in artificial intelligence and natural language processing offer a great chance to address this problem by providing automatic and prompt assessment of depression through text.

Bangla, one of the most popular languages in the world, is also an essential part of the global Online community. Today, social media is the arena where people express their private feelings, while, at the same time, disclosing their mental state in an implicit manner. The lack of artificial intelligence tools that could measure an individual's mental state to be more effective is the result of the fact that the language of Bangladesh is highly featured, yet underrepresented in NLP studies. This eases the need to consider depression diagnosis based on a specific Bangladeshi view, which incorporates the linguistic characteristics of the language and the culture of the population.

Moreover, the effect of transliterated code-switching in the informal Bangla social media text also has an impact that the current available global models do not effectively address. In order to solve any of these problems, it is essential to create new strategies that will blend different areas of linguistics with the application of modern machine learning and deep learning systems.

This study aims to provide Bangla depression detection, which is one of the emerging trends in research aimed at using technology for the humanitarian effort. The creation of such tools would help not only practitioners in spotting people more prone to stress, but also recognize the culture of discrimination by encouraging normal conversations on mental related issues. This motivation stems from a commitment to advancing inclusive and accessible mental healthcare solutions, ensuring no linguistic or cultural group is left behind.

1.3 Problem Statement:

Depression is a health issue that must be handled under public expenditure as it leads to disability and reduces productivity. While textually, depression can be detected using automated systems as a result of the recent advances in NLP (Natural Language Processing), such research has focused on languages such as English, Chinese, and Spanish which have a powerful influence worldwide. Unfortunately, languages such as Bangla remain vastly underrepresented in this field resulting in a huge unexplored area in terms of development and implementation of depression detection systems for the Bangla speakers.

Another feature of NLP based systems is that the Bangla language is written in a complex script, has a great deal of morphology and in many cases, English is used as a code switch during informal communication which makes it even more challenging. The use of Romanized Bangla and the English language, multilingual, informal in nature, social media content hurts any text processing and analysis. Existing models have primarily been developed using English and other languages which are resource dense and therefore are unsuited to Bangla and the sociocultural dimensions that Bangla entails.

Furthermore, the overall ability to train and assess machine learning models is poorly in step with the depression annotations because of the unavailability of public annotated data sets for Bangla language. This is made worse by the social perceptions of mental health in Bangladesh since the people may be hesitant to take part in research studies or show their emotions resulting in data being less and erratic.

Although there is a great demand, research on depression detection in the Bangla language is still a broad area largely unexplored and for the most part the existing techniques in this field rather have been largely implemented but never evaluated. The situation leaves room for research which means that the Bangla speakers would have more options to combat and address the mental health issues that they face.

In order to fill this gap, it will be necessary to build strong models which are also able to understand the culture of Bangla people in relation to their depression. This type of work would also be pertinent in the field of mental health as well as providing robust means for primay intervention for the relevant communities in order to enhance the mental health status of Bangla –speaking people.

1.4 Proposed Solution

The aim of this research is to build a cost-effective system capable of efficiently detecting depression in Bangla text with the help of classical ML algorithms. Deep learning algorithms, on the other hand, are expensive and need considerable resources and vast volumes of data. Still, some classical algorithms still offer great performance with slightly fewer resources. The focus of the research is to implement and assess the efficiency of four popular algorithms namely Random Forest, Naive Bayes, XGBoost, and LightGBM in the context of the Bangla language.

Out of all the voting schemes, Random Forest is the most widely used because it is an ensemble method that can be obtain reliable results even when the data is quite noisy and disorganized, as is often the case for mental health-related data. In addition, the Naive Bayes algorithm is a probabilistic classifier that is very effective in various text classification tasks and is able to make fast predictions which can be useful in case of wider application. XGboost and LightGBM are also efficient since they extensive use of gradient boosting. This allows the methods to effectively capture complex patterns in the data while still using low computational resources.

The proposed solution will utilize the various preprocessed Bangla text datasets to overcome any language issues. Adequate representation for the datasets will be achieved by the feature extraction techniques which will include TF – IDF and word embeddings. Since these classical models may be assessed in terms of accuracy, precision, recall and F1 score this study intends to establish the best suitable algorithm for the purpose of Bangla language depression detection and thus improve the affordability and viability of mental health approaches.

1.5 Research Objective:

The research objectives of this thesis are as follows.

- Determine the efficacy of Random Forest, Naive Bayes, XGBoost, LightGBM models when detecting depression in Bangla, using accuracy, precision, recall, and F1 score as evaluation criteria.

- Find the best feature extraction methods for Bangla text datasets such as TF-IDF and Word embeddings and implement them.

- Compare the resource consumption and computational costs of the models in order to assess their applicability in low resource situations.

- Seek to enhance the existing knowledge of Bangla mental health detection by providing evidence of feasibility of classical machine learning techniques on Bahasa- low resource language.

1.6 Thesis Outline:

The subsequent sections provide a detailed outline of the entire study. Hereafter is a summary of each chapter.

Chapter 1 provides an overview of depression as a significant global public health challenge. It highlights the widespread impact of depression, affecting millions worldwide, with a higher prevalence among young adults aged 15–29 years. The chapter discusses the social and economic factors that exacerbate depression, especially in low- and middle-income countries, and the alarming association between depression and suicide rates. Recognizing the limitations of traditional diagnostic methods, this chapter introduces the study's approach: leveraging machine learning models to predict depressive tendencies, thereby enabling early detection and intervention.

Chapter 2 presents a comprehensive literature review on depression, analyzing prior research on its causes, symptoms, and global prevalence. The review delves into biological, psychological, and socio-economic factors that contribute to the onset of depression. It further explores the role of social media as both a risk factor for and a potential tool in mental health diagnostics. Various machine learning models employed for detecting depressive tendencies are examined, with an emphasis on their strengths and limitations. The review identifies gaps in the application of advanced machine learning techniques, particularly in utilizing user-generated data for real-time depression prediction, and underscores the need for accurate, scalable, and ethical solutions.

Chapter 3 details the methodology for developing machine learning models to predict depression. The study focuses on four machine learning algorithms—LightGBM, XGBoost, Random Forest and Naïve Bayes which explains and shows their suitability for this task. Data collection involved curating datasets derived from social media and other sources, with preprocessing steps like cleaning, imputation, and feature extraction to ensure data quality. LightGBM was selected for its efficiency with large datasets, while XGBoost was chosen for its ability to handle complex data relationships. Evaluation metrics such as accuracy, precision, recall, F1 score, and ROC AUC were used to assess model performance.

Chapter 4 presents the results of the machine learning models and discusses their implications. The chapter includes a comparative analysis of the models, highlighting their respective strengths and limitations. Additionally, it discusses the broader implications of machine learning in mental health, including ethical considerations and the potential for integrating multimodal data sources.

Chapter 5 synthesizes the findings of the study, emphasizing the critical role of machine learning in advancing mental health diagnostics. The chapter highlights the need for context-specific model selection based on application requirements, such as prioritizing recall in clinical settings or precision in public health initiatives. It also identifies areas for future research, including improving model interpretability, addressing biases in training data, and ensuring the ethical use of machine learning in sensitive applications. By leveraging advanced algorithms like LightGBM and XGBoost, this study demonstrates the potential of AI-driven approaches to revolutionize mental health care and reduce the global burden of depression.

Chapter Two: Literature review

This chapter includes a detailed scrutiny and specifics of particular matters alongside other strategies which form part of different models. This part seeks to present an adequate description of already completed academic and empirical work that bear relevance to the creation of text within an automated system for detecting of depression. The literature review is undertaken about the research work done in the field to help outline the research context, gaps in existing knowledge and highlight how this work contributes to the advancement of knowledge in predicting preteen depression and its influencers. This chapter positions prior knowledge and literature on the subject at hand to support the research question and endorse the outcomes and recommendations of the study.

2.1 Recent works

In a landmark investigation that was conducted in 2009, Costafreda et al. [23] investigated the possibilities of structural neuroanatomy in diagnosing and predicting the outcome of treatment for patients suffering from major depressive disorder (MDD). The research focused on analysis of smoothed gray matter voxel-based intensity values as revealed by T1-weighted MRI from 37 patients in acute depression episodes and 37 healthy patients. Voxel-based morphometry and support vector machines (SVM) with ANOVA as a filter method were employed by the researchers to structurally assess the neuroanatomy and its diagnostic accuracy. The results showed an overall accuracy of 67.6%, sensitivity of 64.9% and specificity of 70.3%. These results demonstrate low diagnostic accuracy which implies that volumetric brain structures are not robust enough to differentiate depressed patients from healthy individuals. However, more importantly, the study aimed to determine whether the structural information could help in predicting the outcome of the intervention. Of the thirty fluoxetine or cognitive behavioral therapy (CBT) patients, how structural brain patterns were assessed, accurately predicted clinical remission in 88.9% of the patients before any treatment, thus also proving to be sensitive and specific in predicting patients who will respond to the antidepressant medication. This underlines the usefulness of neuroanatomical markers in making treatment choices and tailoring the treatment approach in MDD. At least in the context of these findings, the role of structural neuroimaging in predicting clinical response is reiterated. However, with their over reliance on structural markers, the neuroimaging techniques managed only to achieve a relatively low diagnostic accuracy which means that more work needs to be done in order to make the structural markers more useful and reliable by integrating them with functional imaging or genetic analysis or behavioral data, for example. Thus, Costafreda et al. come up with the first steps towards instruments of a neurobiological nature for the control of depression, indicating the potential of structural neuroanatomy to guide the formulation of treatment strategies while at the same time stressing its limitations as a single tool for diagnosis. The study opens up the research to further development of multi modal strategies for the purpose of improving the diagnostic and prognostic importance of neuroimaging in mental health disorders.

In 2008, Fu et al. [24] attempted to investigate the neural correlates of depression by employing whole-brain pattern analysis of fMRI scans during an implicit sad facial affect recognition task. Their research showed that neural patterns of those who were in an acute depressive episode can be distinguished from healthy controls with an impressive accuracy of 86%, sensitivity of 84%, and specificity of 89% using a linear support vector machine (SVM) and leave-one-out cross-validation. The study revealed the existence of a distributed network in the occipital, temporal and limbic regions that contributed significantly to classification. These results also demonstrated the clinical relevance of certain neuroimaging biomarkers for diagnosing depression and assessing the effectiveness of its treatment. The predictive power of the linear models to classify treatment responders versus non-responders was preliminarily established, but these findings were statistically insignificant due to small number of subjects. Fu et al. argued the preferential use of whole-brain approaches over region-of-interest analyses to minimize bias and improve reproducibility. On the other hand, they mentioned that some shortcomings such as focusing on a single dentous medication agent (fluoxetine), the event- related fMRI design, and limited number of participants. Therefore, multi-centre strudies with moderate sample size and complementery measures such as resting state imaging and linking clinical and imaging information are needed. This work brings to fruition the neurobiological tool approach on diagnosis of psychiatric disorders.

Depression is a global concern, and the indications are that this condition will only get worse over the coming years. Over the years, several researches have been undertaken with a view to strengthening the possibility of early interventions through early detection. To find Bangladeshi pxrn abusive comments, in 2022, Sultana et al. [25] took 5000 comment s from social networking sites Facebook, Twitter, and You Tube and tested on six machine learning algorithms, Random Forest, K-Nearest Neighbor, Support Vector Machine, Multinomial Naive Bayes, Logistic Regression, Gradient Boosting to detect abusive comments. Out of this, SVM shows the highest accuracy of 85.7 percent in performing the other models. In the same year Hoque et al. [26] tudied four levels of depression namely, nondepression, mild depression, moderate depression, and severe depression using social media post in Bengali language. The authors explain that with the use of XLM-RoBERTa, a pre-trained multilingual model, this model produced better results with increased F1, while scoring 61.11%, which raised the accuracy to 60.89% compared to the prior models/methods. Similarly, Zim et al. [27] took advantage of NLP techniques in the analysis of the Bengali social media data. The study aimed to solve single and multiple-label classification problems as in the case of binary classification that sought to classify all sentences into either positive or negative sentiments. Employing a Skip-gram model they were able to obtain accuracy rates of 77% for binary classification and 69% for multi-label classification. Vasha et al. [28] examined the depressive state using a sociocultural perspective, including social media postings, texting and commenting on depression. They applied Term Frequency-Inverse Document Frequency (TF-IDF) for vectorization and used several classifiers, including SVM, LR, DT, RF, MNB and KNN. The best accuracy rate observed was 75.15% with the SVM classifier. On the contrary, Chanda et al. [29] aimed to analyse the linguistic content of tweets which focused on mental health issues during COVID19, as well as flu virus, lockdown and social distance. It consisted of 1,500 English tweets posted to the social media platform and specifically targeting tweets on EmoCT (EmotionCOVID19-Tweet) dataset. The study employed two models to address the issue of overwhelming sadness and anxiety in these tweets, and SVM yielded the best accuracy of 71%. In the same year, Rahman et al. [30] performed some experiments seeking to Fig. 1: Workflow diagram of our proposed method. analyze Bangla sentiment and emotions from text analysis using multiple models in addition to deep learning models. They applied a Bidirectional Gated Recurrent Unit (Bi-GRU), which obtained the maximum performance of 69% accuracy in their tests. They explained that deep learning architectures have great potential for emotion recognition in Bangla, and so they should be explored further. Monjoor et al. in their research aimed to understand sentiment, which is the evaluation of people's perception, attitude and judgment towards a multitude of people or objects. This was accomplished in their work [31] where they classified about 17,000 pieces of texts as neutral, positive and negative. The paper presented a satisfactory assessment of the performance of Deep Learning-based Recurrent Neural Network (RNN) on the original and translated (Bengali and English) texts, reporting an accuracy rate of 82.9% on the Bengali texts and 84.5% on the English translated versions. Also in 2021, Alvi et al [32] studied Bangla text sentiment analysis through a Gated Recurrent Neural Network (GRNN) model. To strengthen their results, they employed a 10-fold cross- validation approach three times. Their model achieved the highest accuracy of 78.41% and a lowest accuracy rate of minimum 76.34%. Moreover, Ahmed et al. [33] introduced an LSTM- based technique with a feedback layer for multilabel sentiment classification of Bangla text integrating the word2vec word embedding approach. Using a small, stratified dataset of Bangla sentences, the researchers proposed the LSTM model and achieved an impressive accuracy of 94% in classifying sentiment polarity, outperforming RNN and SVM.

Table 1: Comparison Table of the related literature.

Author	Patient Sample	Implemented model	Results
[23]	2009	Support vector machines	Accuracy: 67.6% Sensitivity: 64.9% Specificity: 70.3%
[24]	2008	Support vector machines (linear kernel)	Accuracy: 86% Sensitivity: 84% Specificity: 89%
[25]	2023	SVM	Acc: 75.15%
[26]	2022	XLM-RoBERTa	Acc: 60.89%
[27]	2022	Skip-gram model	Acc: 77%
[28]	2022	SVM	Acc: 85.7%

[29]	2022	SVM	Acc: 71%
[30]	2021	BiGRU	Acc: 69%
[31]	2021	RNN	Acc: 82.9%
[32]	2024	GRNN	Acc: 78.41%

The compared literatures have been given in table 1.

2.2 Research Gap

The literature on detecting depression and sentiment analysis in text highlights significant advancements, but several gaps remain that warrant further exploration. Many studies have focused on deep learning models like XLM-RoBERTa, Bi-GRU, and LSTM, which have demonstrated promising results but require substantial computational resources and large datasets, making them less accessible for resource-constrained settings. Conversely, while classical machine learning models like SVM and Random Forest have shown potential, their performance is often limited by feature representation techniques, such as TF-IDF, that may not fully capture the semantic richness of text. Moreover, much of the existing research is skewed towards English datasets, with relatively few studies addressing low-resource languages like Bangla. Even within the Bangla language domain, studies often emphasize sentiment classification or basic depression detection, rather than tackling nuanced distinctions such as different levels of depression severity. Additionally, while approaches utilizing social media data have been explored, these often rely on small, non-diverse datasets, limiting the generalizability of the findings.

Integration of multimodal data sources is another gap, in that, most of the studies limit themselves to the use of text data and neglect to incorporate user demographics and behavioral patterns which could also boost predictive capability. Last, however, the application of similar diagnostic approaches to text systems is still an unexplored area, hence the gap in comprehensive and clinically relevant context detection frameworks, even though Fu et al. and Costafreda et al. showed optimal results for using neuroimaging data in predicting treatment response.

Chapter Three: Methodology

Our study will now move to the center stage starting with this chapter and discuss ML models as well as an in-depth methodology used to predict depression in social media text. We aim at building a powerful predictive model that makes use of state-of-the-art machine learning techniques applied on social media for the purpose of identifying depression. The chapter will cover process of data collection, preprocessing steps, feature selection, model development and evaluation metrics.

3.1 Dataset description

BSMDD dataset published by Madeley Data on 14th October 2024 is a rich resource of Bengali social media and its analysis provides a huge scope in the field of mental health. The current dataset contains 21,910 samples being the result of an ongoing collaboration with the Ahsanullah University of Science and Technology, which includes 10,961 instances that were tagged as Depressed and 10,949 instances that were tagged as Non-Depressed. This data BSMDD has been translated, and labeled by qualified, and mental health trained native Bengali speakers, therefore the data is valid that was obtained from various social media platforms. Due to the usability of this dataset for the general population, it is beneficial for research in the field of language and depression, because this dataset consists of relevant data according to the Bengali language and language used in the social media posts.

The words that have been used in training are visible in the word-cloud in figure 1.

Figure 1: The dataset Word-Cloud

The BSMDD dataset has been given in figure 2

1	text	label
2	মানসিক শারীরিকভাবে অসুস্থ ক্লান্ত পুরো জীবন শারীরিক স্বাস্থ্য সমস্যার সাথে লড়াই করেছি মানসিক স্বাস্থ্যের সমস্যাগুলির সাথে লড়াই করছি উভয়	1
3	দয়া সাথে থাকুন অত্যন্ত দীর্ঘ আপনাকে পড়তে উত্সাহিত তথ্য সহায়ক পারো।এবং এজন্ত এছাড়াও নোট প্রচুর প্রযুক্তিগত আলোচনা চলেছে দয়া হতাশ	1
4	জানতাম সাথে ভুল লোক খারাপ জীবন কাটিয়েছে সম্পূর্ণরূপে কার্যকরী কতটা করুণ হারানোর অজুহাত নিজেকে বলতাম বাবামায়ের দোষ ধ্বংস চি	1
5	অনেটিভ ইংরেজি স্পিকারের অনুসরণ বিরক্তিকর অপ্রতিরোধ্যভাবে দীর্ঘ চিৎকার রাশিয়ার যুবকের জীবন সম্পর্কে দ্রুত উঁকি আগ্রহী স্বাগতম পোস্ট	1
6	অনেটিভ ইংরেজি স্পিকারের অনুসরণ বিরক্তিকর অপ্রতিরোধ্যভাবে দীর্ঘ চিৎকার রাশিয়ার যুবকের জীবন সম্পর্কে দ্রুত উঁকি আগ্রহী স্বাগতম পোস্ট	1
7	প্রথমে আপনাকে প্রশ্ন উপস্থাপন যাচ্ছি পারবেন দিয়েছি আপনাকে সরবরাহ চাই গুরুতর বিষণ্নতা ভালোবাসো জীবনে ব্যক্তি সম্পর্কে চিন্তা আপনা	1
8	রেফারেন্সের বয়সী পুরুষ থাকার সাইক ওয়ার্ড বের হয়েছি একেবারে মানসিক অ্যালকোহল বাঁকানোর স্ন্যাপচ্যাটে জীবনের সুরক্ষার মাতাল অঙ্গভঙ্গি	1
9	দৈর্ঘ্যের ক্ষমাপ্রার্থী বোঝানোর বের আলিঙ্গন বলার প্রয়োজন সবকিছু বন্ধুদের মিস বুধবার রাতে অপরের জায়গায় কাটাতে অপরের মুখ হাসতে পান	1
10	উপকারী সাব দয়া আপলোড পোস্ট সেরা জায়গা মোবাইলে আছি ত্রুটির কিছুটা দুঃখিত পারো বর্তমানে চিন্তায় আটকে আছি কোথাও পোস্ট না। বান্ধ	1
11	দীর্ঘ পোস্ট জীবনের সারসংক্ষেপ টাইম লাইনটি একটু ঘোলাটে শুধুমাত্র কারণে কষ্টে আছি এটিকে প্রকাশ চাই টাইপো বিরাম চিহ্ন ব্যাকরণগত ত্রুটি	1
12	সাতাই বুক পেতে চাই বন্ধু পরিবারের সাথে প্রতিক্রিয়া দেখাতে চাই অবস্থা কারো খারাপ মধ্যবিত্ত এসেছি বাবামা আছেন সত্যিই যত্ন নেন বন্ধু আছেন	1
13	কাঁদতে কাঁদতে বোতল ওয়াইন পান সময় লিখছি শিরোনামটি অবিলম্বে পাথরের নীচে আঘাত করেছি ভেবেছিলাম নীচে নামতে পারব বিষণ্নতা সম্প	1
14	সম্পর্কের উদ্বেগের কারণে এসেছি মেয়ের সম্পর্কে প্রচুর উদ্বেগ থাকার বিষয়ে পোস্ট করেছি চিনতাম সম্ভ্রান্ত মেয়ে গ্রাউন্ডেড মিষ্টি অদ্ভুত কারণে রস	1
15	শহরতলির টেনেসিতে বড় হয়েছি সীমান্তরেখার মৌলবাদী খ্রিস্টান সম্প্রদায়ের শৈশব অত্যন্ত আরামদায়ক সুখী আশ্রয়ে বাবামাকে ভালোবাসি সত্যিক	1
16	রবিবার রাতে গলায় ফাঁস আত্মহত্যার শেষ বন্ধ রেখেছিলাম জেগে উঠে হিংস্রভাবে হাঁপাতে হাঁপাতে দৌড়ে বাথরুমে ছুঁড়ে ফেলেছিলাম বাথরুমের মে	1
17	হ্যাঁ এমনকি দ্বিতীয় তৃতীয়বার পোস্ট করেছি সাতা আশা করছিলাম হতাশার বিন্দুতে পৌঁছতে পারব যতবার পোস্ট মানসিক অবস্থা শেষের খারাপ আ	1
18	পড়ার সময় নেয় অগ্রিম ধন্যবাদ বর্তমানে গভীর বিষণ্নতায় রয়েছি বাড়িতে জীবন অবিশ্বাস্যভাবে হতাশ অতীতে বিবেচনা হালকা হতাশাজনক অব	1
19	হাই জানি কোথায় কীভাবে পরিস্থিতি বর্ণনা মস্তিষ্ক কিছুটা কুয়াশাচ্ছন্ন পুরো পাঠ্যটি জগাখিচুড়ি আশা সাথে সহ্য অবশেষে ভয়ের মুখোমুখি হয়েছি উ	1
20	চাই জীবনের শেষ কিশোর বয়সে নিছক সম্পূর্ণ নরকীয় পারিবারিক অন্যান্য অপব্যবহারের ক্রমাগত কলেজের সংগ্রাম আপত্তিজনক রোমান্টিক সম্	1
21	সপ্তাহের বয়সী চলেছি। ওয়াশিংটন রাজ্যের শ্বেতাঙ্গ ছেলে প্রচুর ভিডিও গেম খেলে বড় হয়েছি স্কুলে ভাল করিনি পড়াশোনায় সময় ব্যয় করিনি মা বাচ্	1

Figure 2: BSMDD dataset

3.2 Data Preprocessing

This section will deal with how the dataset was processed for further analysis and model training. The steps include data cleaning, imputing, encoding and scaling.

3.2.1 Data Cleaning:

The clean_data function performs a comprehensive data cleaning process focused on text preprocessing. It begins by removing all numeric characters from the text, followed by eliminating alphabetic characters to retain only relevant content. Punctuation marks, including special characters such as '।' (commonly used in certain languages), are stripped from the text using translation tables. Emojis are replaced with an empty string to ensure only textual data remains. To refine the text further, common stop words are removed, enhancing the semantic relevance of the remaining content. Extra whitespace is standardized by replacing multiple spaces with a single space and trimming any leading or trailing spaces. The text is then transformed using a stemming process, reducing words to their root forms, which aids in normalizing variations of the same word. Finally, the cleaned text is tokenized into individual words for further analysis or processing. This method ensures the data is thoroughly prepared for subsequent text analytics or machine learning tasks.

3.2.2 Text Vectorization:

The second part, related to K-Fold cross-validation, is implemented in the following way. The data splitting step is implemented in the train_with_K_fold function where preprocessing steps are performed. The data in the column 'text' undergoes transformation into vectors via a TF-IDF Vectorizer type transformation, which suits the documents weight based on frequency of the term in the corpus. The vectorizer also takes care of the dimensionality problem by compressing the feature space to 5,000 of the most salient terms. This indeed yields the X, a numerical array suitable for rather the training of machine learning models. At the same time the column 'label' is transformed into y which would be used as the classification target variable.

3.2.3 K-Fold Validation:

In order to avoid data overfitting and to assess the model qualitatively the function goes through K-Fold Cross-Validation. The dataset is partitioned into k subsets, default is five, and each partition is used alternately for testing and training purposes, after shuffling the data, some random_state parameter e.g random_state=42 is used to ensure reproducibility of the test result. This by all means reduces the non-response bias and also provides an overall view of how well the model performs with the different partitions of the data.

3.3 Model Development

The models that have been used and prepared have been discussed in this section. Mainly, LGBM and GRU have been used for prediction.

3.3.1 LightGBM (LGBM)

LightGBM [34] is a gradient boosting framework that was developed by Microsoft with ability to process huge amount of data and is a fast and efficient implementation. This algorithm uses histogram based method to speed up training compared to standard gradient boosting method [34]. Considering its side features, it is quite efficient for tasks with large amounts of data and intricate models. In LightGBM, the builds of decision trees are done in iterations such that each tree rectifies the wrong predictions of its predecessors thus reducing the loss function. It offers many of the sophisticated capabilities such as the ability to directly handle categorical variables, tree growth by leaves instead of by levels and efficiently managing missing values.

This Tree wise grown of LGBM is shown in figure 3.

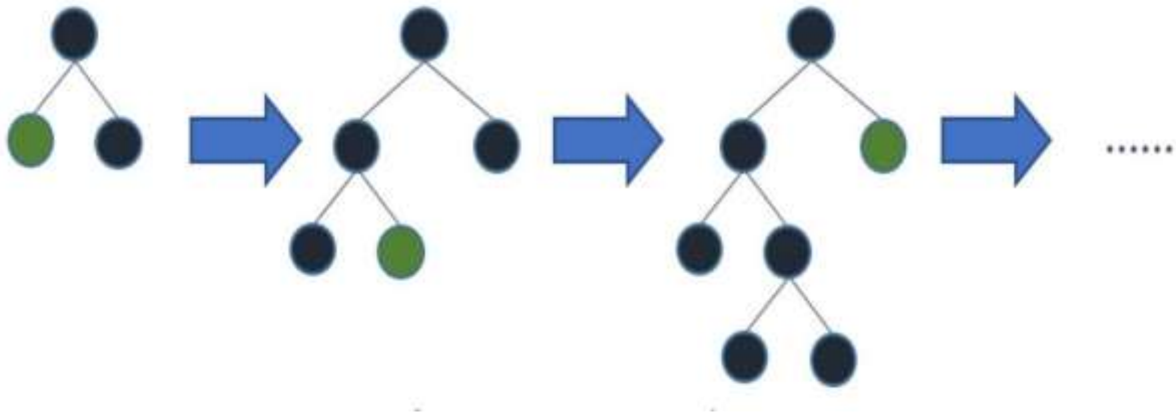

Fig 3: Light Gradient Boosting model (LGBM) and its tree wise growth

3.3.2 Random Forest:

Random Forest [35] is one of the most effective ensemble learning algorithms that is used in tasks of both classification and regression by aggregating the predictions of a number of decision trees in order to make more accurate predictions. Includes bagging to increase the robustness of the trees as each tree is fit on a random sample of the data set with replacement. For classification, the prediction is done using the majority of votes, while for regression, the average value of all the predictions is done. High-dimensional data and multicollinearity are quite severe problems, but they are not a concern with Random Forest. It randomly selects a subset of features at each split which reduces overfitting and helps with generalization. The model also computes feature importance scores which tell how useful was each feature in the decision- making process. It is relatively difficult to develop an ensemble of models as it is computation heavy, though this is compounded by its parallelizable nature, making it appropriate for big data. Its accuracy, robustness, ability to work with imbalanced and missing data among other factors makes Random Forest appealing as it requires very little parameter adjustments and operates in an effective manner in various tasks including but not limited to medical imaging and natural language tasks.

3.3.3 XGBoost:

Indeed, XGBoost [36] is an extremely efficient machine learning algorithm that is based on an advanced version of gradient boosting. It incorporates several innovations, such as second-order Taylor expansions for precise updates and L1 and L2 regularization methods to counteract overfitting, XGBoost consists of building a whole sequence of decision trees by minimizing a loss function. So it is also important on the other hand to mention that XGBoost develops a recursive algorithm that aims to speed up the computation of the above-stated trees by parallel processing. These characteristics make it highly scalable and easier to use. The model's features make it exclusively suitable for various purposes and tasks, including the ability to customize objective functions and evaluation metrics. This algorithm is probably one of the best and robust in terms of accuracy and efficiency, as it regularly wins the competitions in machine learning or in tons of cases in real life applications. It is worth mentioning that the model is complex but relatively easy to work with due to the numerous available documentation materials and possibilities to support implementations. Thus it is very popular and can be used in predictive modeling, customer satisfaction, and anomaly detection tasks.

3.3.4 Naïve Bayes:

Considering the probabilistic approach of machine learning, which computes classes based on a given set of features, the Naïve Bayes [37] algorithm is perhaps one of the simplest ones. As its terminology describes, the model assumes that all features are independent, that is, the presence of one feature does not affect other features, a rather unrealistic but computationally easy and easy-to-implement assumption. Naive Bayes is powerful enough to perform text classification tasks such as spam detection and sentiment classification despite this "naive" assumption. As a rule, most independent features pertaining to a single class determine the output, though people generally employ it with small datasets and consider it rather robust in the presence of outliers. All classes of problems which can be approximated with the distribution featuring a class have the highest probabilities given priors and likelihood of features. However, given the inherent criterion of a feature being a high probability one, the model suffers on features exhibiting high correlation. Naive bayes is interpretable, requires less training time and is known to be suitable for classification applications which are simple and don't require time long computations.

3.4 Evaluation Criteria

During the evaluation of the model, 20 percent went for testing, from the dataset was divided into two groups of 80% and 20%, to make sure that overfitting could avoided. In this section, we discuss the evaluation metrics that help us definitively access the models' performances.

Accuracy: The level of correctness or wrongness of the model's predictions is sometimes referred to as the accuracy measure. It is an indefinite but general assessment of the performance of the model even though it is not very meaningful for imbalanced datasets because it cannot tell the nature of the errors being made.

$$Accuracy = \frac{\sum TP+TN}{\sum TP+FN+FP+TN}$$

F1 Score: A very important metric which relates to both precision and recall is referred to as the F1 measure. The accuracy of the model is dependent on the distribution of the classes within the majority of the datasets provided. It is described mathematically as the average of precision and recall giving importance` to both positive and negative errors. A situation when the value of the F1 measure is adequate indicates that the model has the capability of reducing errors while identifying the positive cases correctly

$$F1 = \frac{TP}{TP+\frac{1}{2}(FP+FN)}$$

Precision: Precision focuses on only one aspect of performance: the accuracy of the model's positive predictions. That is, the number of true positive results from the positive predictions. A significantly high precision value means that a model has successfully reduced unnecessary positive predictions and this is helpful in situations where the consequences of making unnecessary positive predictions are very high.

$$Precision = \frac{TP}{TP+\text{TP}}$$

Recall: Recall measures how well the model identifies every single actual positive example being a ratio of true positives to total number of actual positives. A high recall rate is important when one needs to classify all instances of positives even when there are many false alarms involved.

Chapter Four: Result and analysis

4.1 LGBM prediction:

The evaluation matrices for the LGBM predictions have been given in table no 2.

Table 1: Evaluation of Predictions by LGBM

Evaluation Metric	**Values**
Accuracy	0.836
Precision	0.834
Recall	0.840
F1 Score	0.834

The confusion matrix for LGBM has been shown in figure 4.

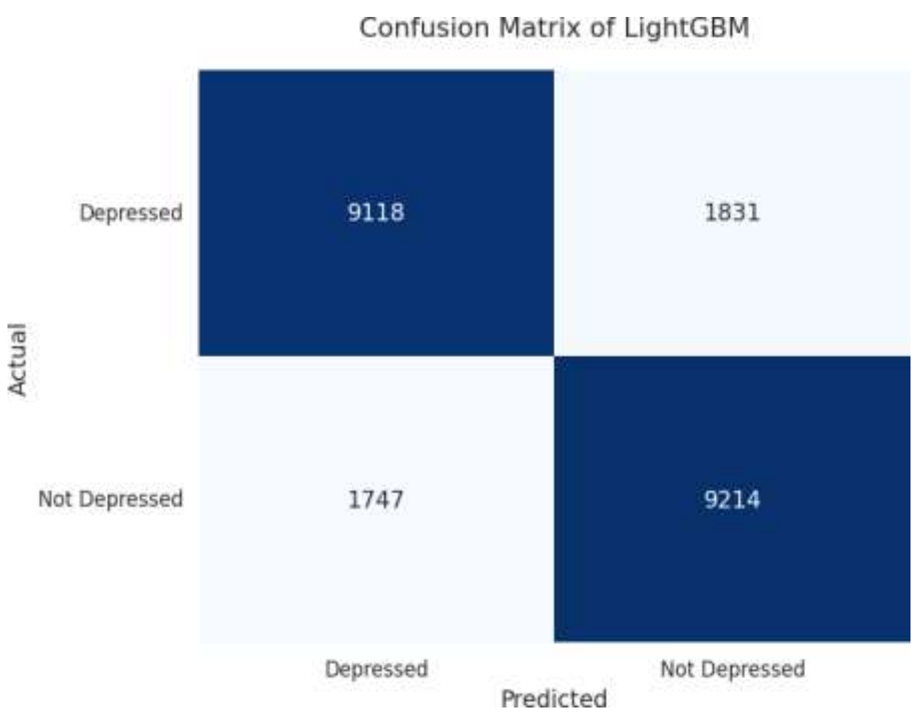

Fig 4: Confusion model through LGBM

4.2 XGBoost predictions:

The evaluation matrices for the XGBoost predictions have been given in table no 3.

Table 2: Evaluation of Predictions by XGBoost

Evaluation Metric(Average)	**Values**
Accuracy	0.827
Precision	0.829
Recall	0.824
F1 Score	0.826

The confusion matrix for XGBoost has been shown in figure 5.

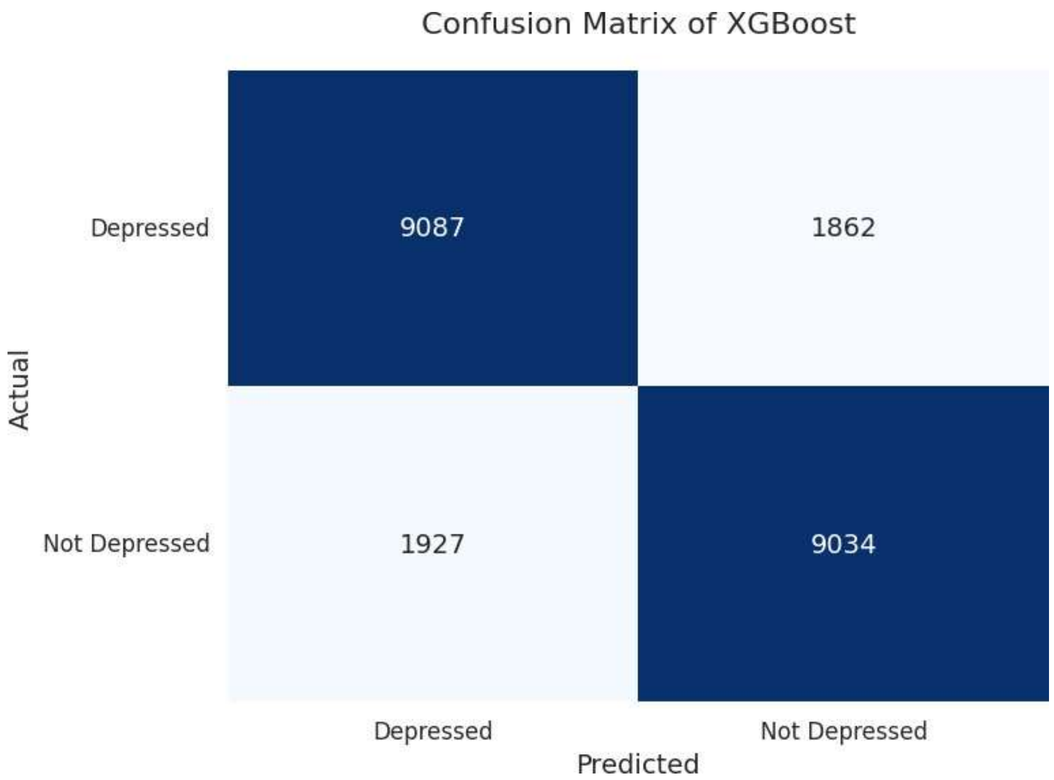

Fig 5: Confusion model through XGBoost

4.3 Naiye Bayes predictions:

The evaluation matrices for the XGBoost predictions have been given in table no 3.
Table 3: Evaluation of Predictions by Naïve Bayes

Evaluation Metric(Average)	**Values**
Accuracy	0.820
Precision	0.796
Recall	0.862
F1 Score	0.827

The confusion matrix for Naïve Bayes has been shown in figure 6.

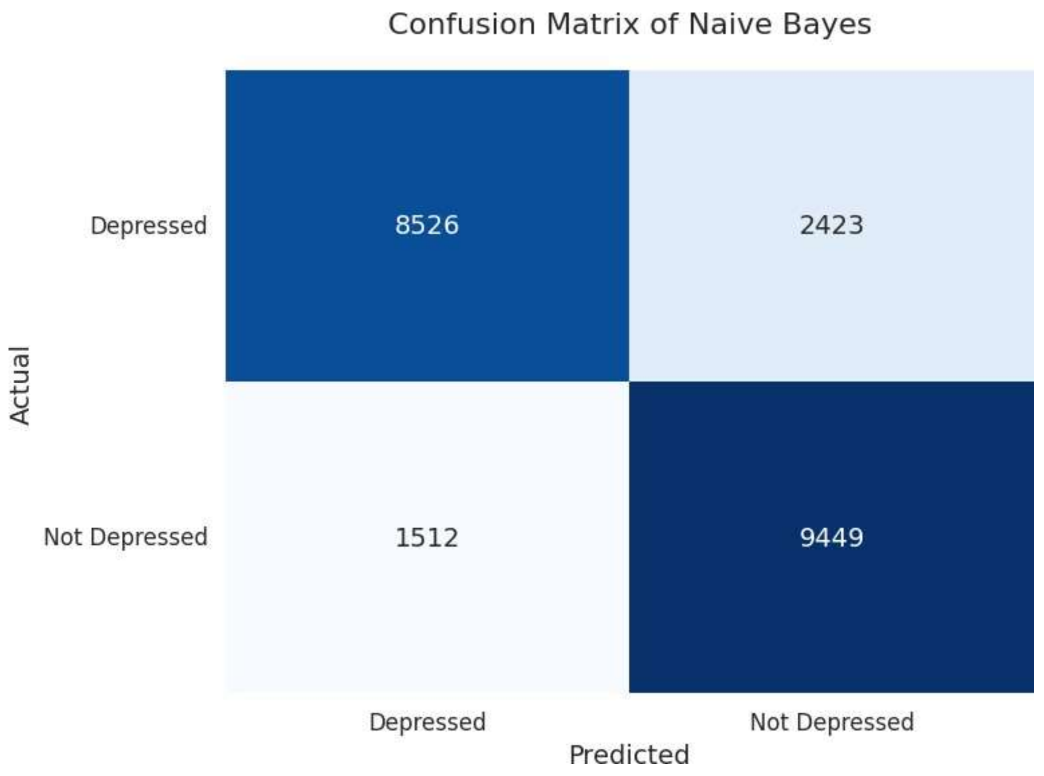

Fig 6: Confusion model through Naïve Bayes

4.4 Random Forest predictions:

The evaluation matrices for the Random Forest predictions have been given in table no 4.
Table 4: Evaluation of Predictions by Random Forest

Evaluation Metric(Average)	**Values**
Accuracy	0.830
Precision	0.819
Recall	0.849
F1 Score	0.833

The confusion matrix for Random Forest has been shown in figure 7.

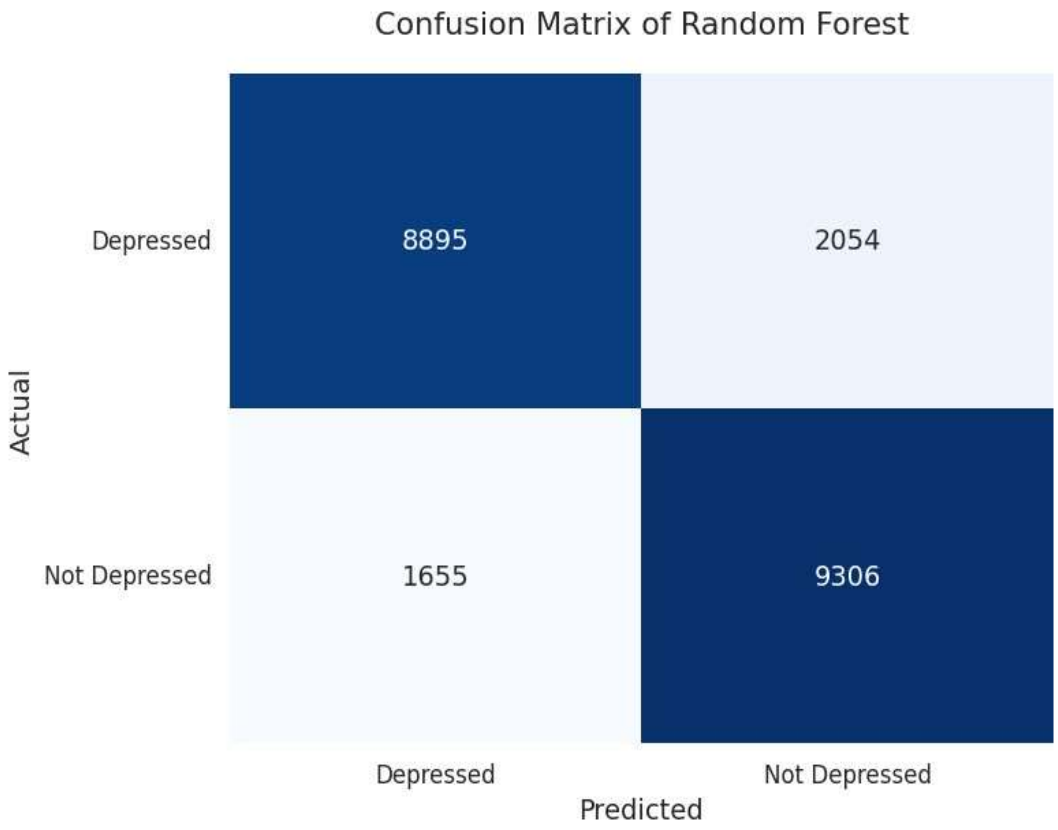

Fig 7: Confusion model through Random Forest

4.5 Combined Results

The combined table 4 shows the evaluation metric values of the implemented models for clear understanding.

Table 5: Evaluation of Combined Results

Model	Accuracy	Precision	Recall	F1
Naive Bayes	0.820402	0.796022	0.86211	0.827676
XGBoost	0.827065	0.829149	0.824199	0.826650
Random Forest	0.830717	0.819209	0.849031	0.833823
LightGBM	0.836696	0.834261	0.840618	0.837409

In figure 8, we see the combined accuracy.

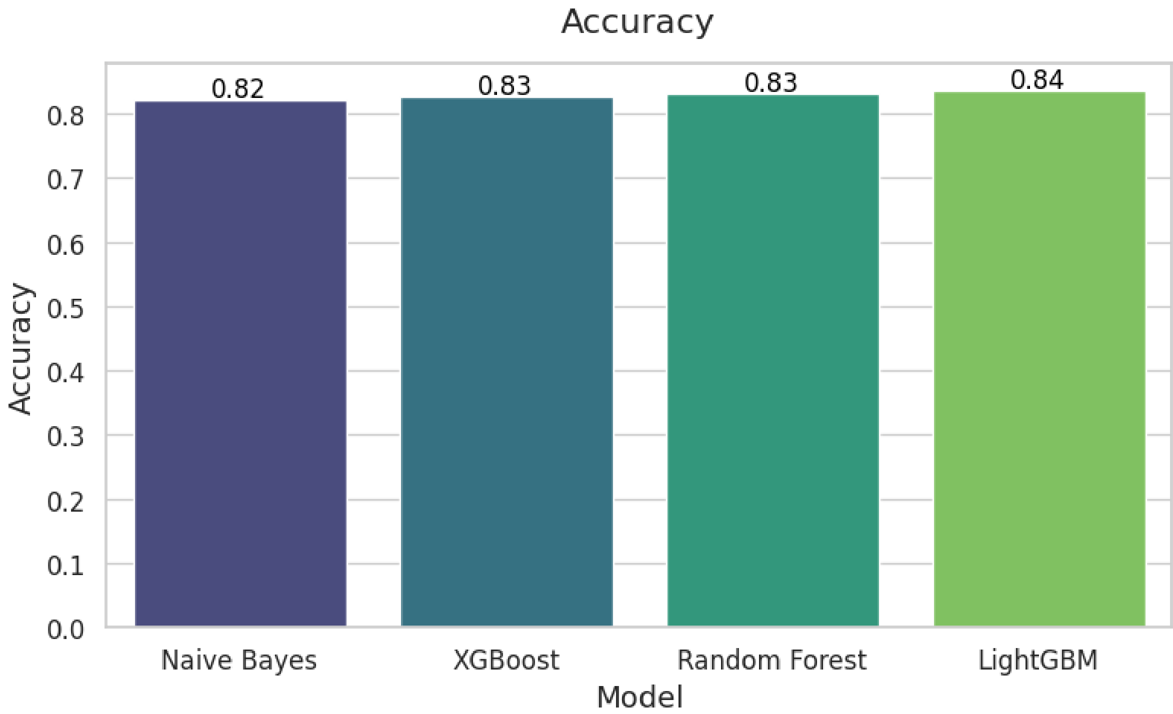

Figure 8: Combined Accuracy

The measurement of the accuracy is to assess the number of correctly classified observations over the total number of the observations. It gives a high level overview of the performance of the model as a whole. In the assessment made, LightGBM was able to use data with the highest accuracy of 83.67%, this was the best performance of the model in classifying the instances. However, Naive Bayes, registered the worst accuracy in this criteria using only 82.04%.

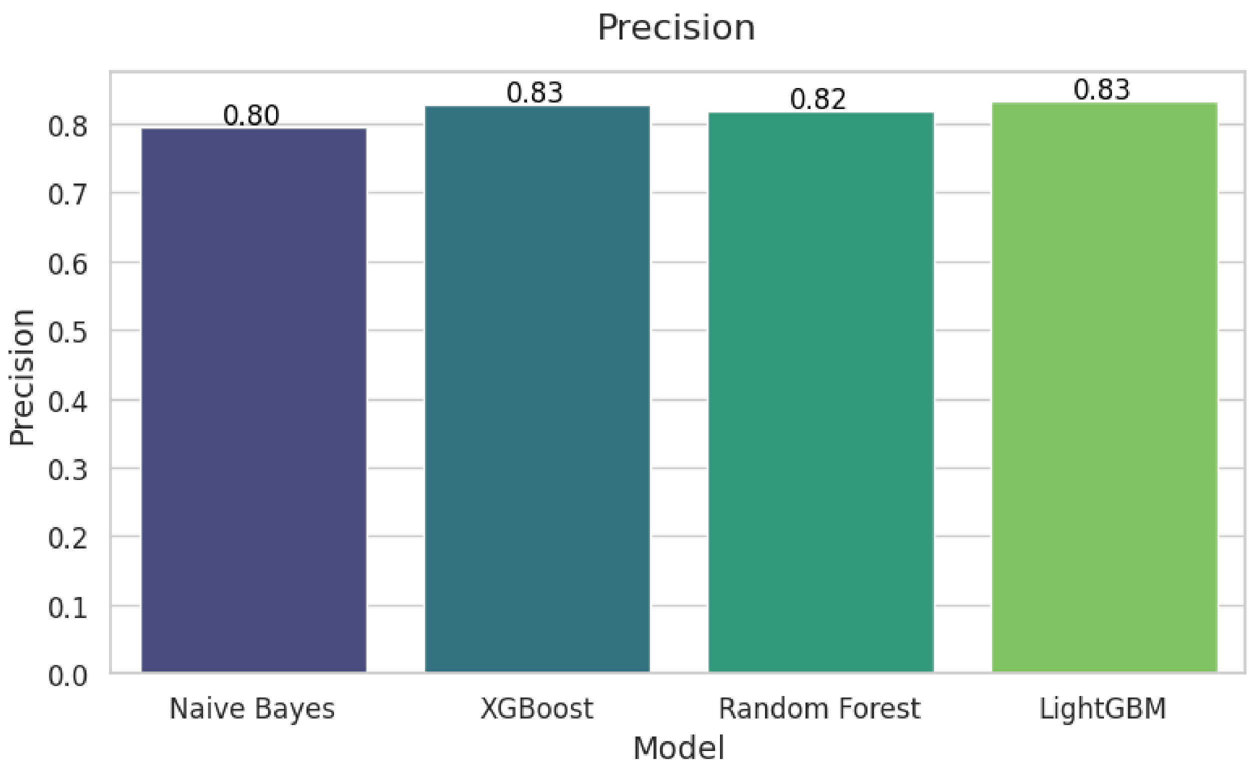

Figure 9: Combined Precision

The precision of a model is expressed as a ratio between the actual true positive prediction and the total number of positive predictions. It reflects the ability of the model to get false positives. With the highest precision of 83.91% XGBoost appears to be the best predictor of the positive cases among the models. However, Naive Bayes is recorded as the most unaccurate model in this area with a precision of 79.60%. The combined precision in seen in figure 9.

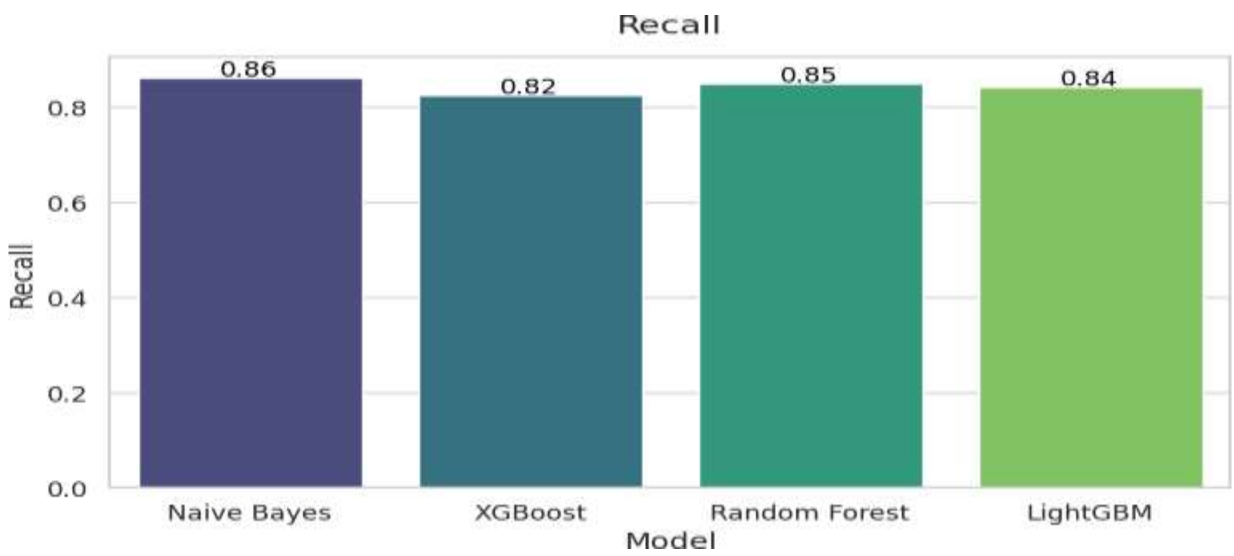

Figure 10: Combined Recall

Recall is also referred to as sensitivity and it assesses the number of actual positive cases that were identified by the model. Its role is to highlight the performance of the model in cases of false negatives. On massacring the other two models, Naive Bayes had the greatest recall at 86.21%, this means that there were more true positives correctly predicted. XGBoost scored the worst in recall out of the three with a score of 82.42%, indicating a better chance of missing actual positive cases. The combined recall in seen in figure 10.

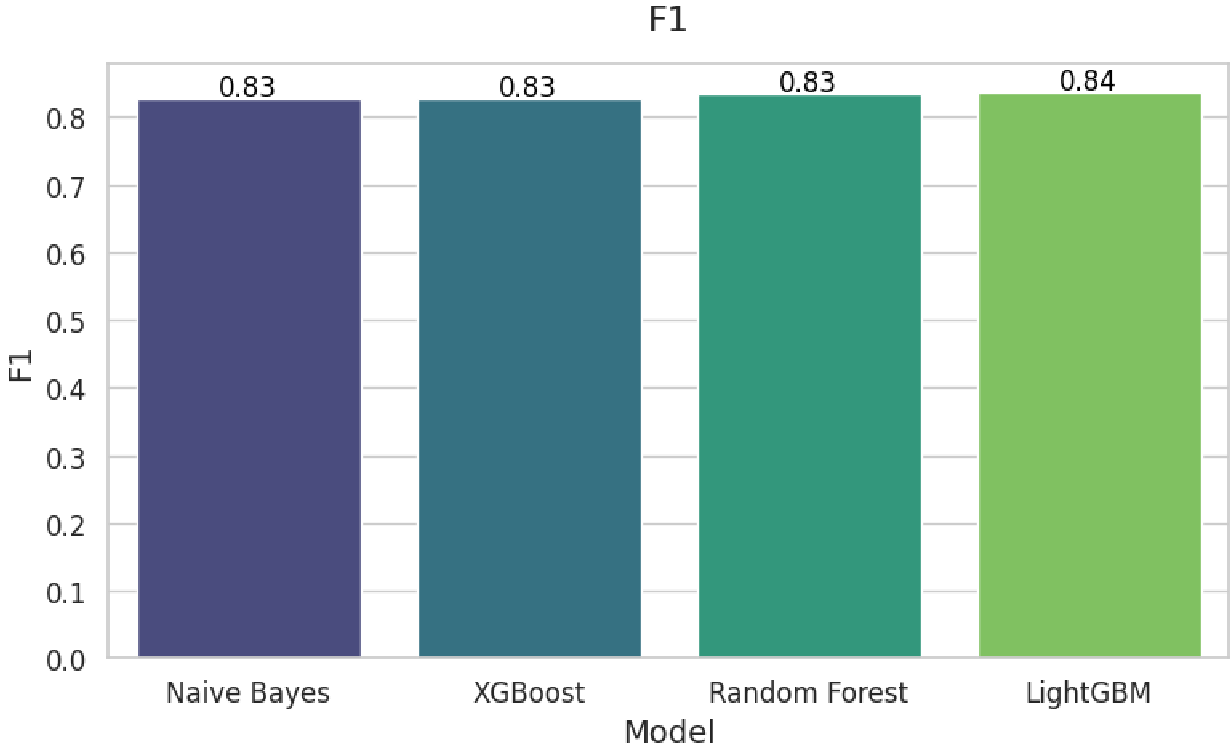

Figure 11: Combined F1-score

F1 score is the measure of effectiveness which considers both the precision and recalla mot useful when the cost of ignoring something is similarly level. False Positives and False Negatives. Light GBM's prcison and recall levels are the most balanced with an F1 score of 84.06%. It has the best F1 score. On the other hand, a less than average F1 score makes Naive Bayes a relatively unbalanced model since it has a high recall but low precision that is not minimal. The combined f1-score in seen in figure 11.

4.5 Discussion

Mental health disorder come in many forms, however Depression has been consistent in its deleterious development across the globe. To make this unequivocal clear Depression affects over 264 million people making it one of the most acute challenges we face today. The chronic feeling of sadness and low in cognitions and activities does not only have psychological effects but alters the socioeconomics as well. The strongest affect of this disorder has been noted amongst youth aged 15-29, which doesn't get easier due to perceptional factors and low to none availability of mental health resources in developing countries. The rising rates of suicide, where close to 800,000 people die of this every year are another factor deserving attention due to the search for solutions. As social media platforms emerge, so do the complexities attached to it. seeing how certain oversaturation of social media promotes anxiety, depression, and stress through the medium of social comparison and even fear of not fitting in. However, these platforms contain important information that can be pivotal in the early signs of depression. The use of machine learning algorithms, which are best suited to discover patterns within large and complicated datasets, seems to provide an effective solution to the improvement of mental health diagnosis and its intervention approaches. This research applied four algorithms namely LightGBM, XGBoost, Naive Bayes and Random Forest in depression analysis and made a comparative study of their performance based on accuracy, precision, recall, and F1 score.

On the other hand, the results underline the advantages and disadvantages of each algorithm, and in this case, it was LightGBM that proved the best model in depression prediction. LightGBM had the highest accuracy of 83.67% and the F1 score of 84.06 %, this made the model to perform better than all the others overall taking into consideration the precision and recall. It is also likely that the combination of these features and the ability to manage enormous amounts of data assisted in its success. XGBoost is well known to be the best implementation of gradient boosting with a precision of 83.91%, which shows that it is able to effectively reduce the number of false positives. That is especially useful in the case of mental illness where a patient is tagged as depressed mistakenly, which can cause embarrassment or anxiety for that person. On the other hand, Naive Bayes performed well in recall with an 86.21 percent ratio, which suggests it has the ability to capture many true positives. This characteristic is desirable in cases where one of the most important requirements is that no patient suffering from depression is left unidentified even with the risk of high false positives. Random Forest's performance was uniform across all metrics with an accuracy of 83.07% and F1 score of 83.33% making it an excellent option for detection tasks related to depression. These observations emphasize the need of aligning the choice of algorithms with the aims of some particular mental health application. For example, the case that requires high precision would likely prefer XGBoost, while the opposite would rely on Naive Bayes. Given

LightGBM's impressive results, it appears to be a good choice in the setting of real-world applications requiring precise calibration. This study not only validates the utility of machine learning in mental health but also opens avenues for future research to enhance model interpretability, address ethical considerations, and integrate multimodal data sources to improve diagnostic accuracy further.

Chapter Five: Conclusion and Future work

5.1 Conclusion

Depression is one of the most serious challenges faced by society today. It has such a far reaching influence on the individual and the society at large, that one is left with no choice but to come out with new approaches in ways to both detect it and treat it. It is true that many traditional means of addressing mental health issues will remain important, but the application of machine learning and other digital instruments in this area makes it possible to affect changes in a revolutionary manner. The work examined the active use of four algorithms in the detection of depression among patients in complex datasets; LightGBM, XGBoost, Naive Bayes and Random Forest were the algorithms under evaluation with emphasis on their ability to mitigate this problem.

It was found that LightGBM was the best among the models with an outstanding accuracy and F1 score. The capability of LightGBM to effectively work with large, complicated datasets without sacrificing computational performance makes this model one of the best candidates for the real life applications in the screening for depression and other mental disorders. Because there is a good compromise between precision and recall with these procedures, there minimization of both false positive rate and false negative rate which are important in accurate classification in mental health since inaccuracy can lead to adverse consequences. XGBoost has good precision which makes it an excellent fit for instances where there is need to cut down the number of false positives for example when carrying out large scale testing and diagnosisIn contrast, Naive Bayes performed remarkably well in recall which is beneficial if the goal is to capture as many cases of depression as possible including at the expense of some false positives. Random Forest showed a strong and steady performance in most metrics, thus it can serve as a dependable and stable moderate performer across various mental health interventions.

This comparison of applications of these algorithms is very important for integration of machine learning in the area of mental health. It appears that the decision to use a certain algorithm should be governed by the intended use of the application. For instance, in the clinical context where it is of utmost importance not to overlook any case of depression, then it is preferred to use models with high recall such as Naive Bayes. On the other hand, in the public health programs that aim to avoid the stigma associated with overdiagnosis, XGBoost is the recommended model due to its high precision.

Apart from the efficiency metrics that this research covers, there are wider implications of the use of ICTs with the 'business' of mental health. The possibility to access large volumes of information from the web sources, especially social networking sites, opens new avenues for effective preventive health care. Nevertheless, there are also potential threats to manage including privacy concerns, bias of algorithms, and assurance of transparency in ethical practices. Indeed, ethical challenges are more valid while dealing with sensitive areas such as health in which an indiscretion or abuse of information may have life changing effects on a person. There is a need for future studies to concentrate on creating models that are properly explained, and accurately placed within the context of various types of data, while also trying to bridge the gap in the digital age in order to allow the fair distribution of such technology.

As a final point, machine, may be, has great potential in improving detection of depression such that the depression diagnosis maybe more accurate, done in time and will be more easily deployed over a wider population. When the process of choosing algorithms is suited to applications' objectives and ethical and technical issues are resolved, the potential of the technology to contribute greatly to alleviating the depression burden should be tremendous. The results of this research not only show the efficiency of the machine learning models such as LightGBM but there is also an opportunity to use AI-based tools in order to make the approach to mental health more active and data centered. Provided that more innovative tools are developed and more stakeholders come on board, the use of machine learning in mental health care would be quite effective in increasing access and saving lives.

5.2 Areas of Future Research

Future research in the area of application of artificial intelligence in analyzing signals of depression would be directed towards enhancing the effectiveness, equity and ethical aspects of use of these models. One key concern is understanding and trusting the model outputs— increasing model interpretability and explainability. On the other side, advanced algorithms such as LightGBM or XGBoost are powerful, but their highly nonlinear nature causes difficulties in comprehending the rule for making predictions. In this regard, future research should incorporate Explainable AI approaches that can shed some light on how decisions are made and the rationale behind them. Simultaneously, multicultural incorporation should be investigated alongside analysis of voice indicators, facial features or even physiological markers as addition or override to written or numerically coded data. This way, depression would be approached in a more holistic and accurate representation since a wider range of an individual's actions and emotions would be included.

Moreover, they considered algorithmic bias that is addressed could promote equity in the predictions made. Where there are data restraints, concerns of bias are often present resulting in lower prediction outcomes in some demographics than others resulting in inequity. Therefore, the next stage should be about moving towards fair algorithms which are robust to diversity of population and do not bias certain populations during the training process. In addition, the approaches able to provide the real-time monitoring and the interventions need to be developed. Following that machine learning models can be implemented within mobile or social media applications so that that the individual's mental well-being can be constantly managed through ongoing comprehensive relevant interventions. The catch is that this comes at the risk of privacy and informed consent issues. As a result, ethical concerns need to be placed for all future studies and researchers need to work on how data will be secured so that he/she sensitive health data will not be abused and their privacy respected. Furthermore, participants' engagement should also be encouraged in such a way that they are able to have a say on how their data is utilized and what the implications of using such technologies would be. In conclusion, these issues need to be resolved so that the applicability of machine learning in the field of mental health is maximized, not only enhancement of diagnosis but also allowing application in a humane, fair and user oriented manner.

References

[1] R. I. Shader, "How Patients Describe Their Depressions: D Words," *J. Clin. Psychopharmacol.*, vol. 41, no. 4, p. 364, Aug. 2021, doi: 10.1097/JCP.0000000000001438.

[2] S. Amiri, "Unemployment associated with major depression disorder and depressive symptoms: a systematic review and meta-analysis," *Int. J. Occup. Saf. Ergon.*, vol. 28, no. 4, pp. 2080–2092, Oct. 2022, doi: 10.1080/10803548.2021.1954793.

[3] M. Nordentoft and A. Erlangsen, "Suicide—turning the tide," *Science*, vol. 365, no. 6455, pp. 725–725, Aug. 2019, doi: 10.1126/science.aaz1568.

[4] E. B. O'Day and R. G. Heimberg, "Social media use, social anxiety, and loneliness: A systematic review," *Comput. Hum. Behav. Rep.*, vol. 3, p. 100070, Jan. 2021, doi: 10.1016/j.chbr.2021.100070.

[5] B. Keles, N. McCrae, and A. Grealish, "A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents," *Int. J. Adolesc. Youth*, vol. 25, no. 1, pp. 79–93, Dec. 2020, doi: 10.1080/02673843.2019.1590851.

[6] A. Vannucci, K. M. Flannery, and C. M. Ohannessian, "Social media use and anxiety in emerging adults," *J. Affect. Disord.*, vol. 207, pp. 163–166, Jan. 2017, doi: 10.1016/j.jad.2016.08.040.

[7] A. Shensa, J. E. Sidani, M. A. Dew, C. G. Escobar-Viera, and B. A. Primack, "Social Media Use and Depression and Anxiety Symptoms: A Cluster Analysis." Accessed: Jan. 04, 2025. [Online]. Available: https://www.ingentaconnect.com/content/png/ajhb/2018/00000042/00000002/art00011

[8] N. V. Babu and E. G. M. Kanaga, "Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review," *SN Comput. Sci.*, vol. 3, no. 1, p. 74, 2022, doi: 10.1007/s42979-021-00958-1.

[9] A. Malhotra and R. Jindal, "Deep learning techniques for suicide and depression detection from online social media: A scoping review," *Appl. Soft Comput.*, vol. 130, p. 109713, Nov. 2022, doi: 10.1016/j.asoc.2022.109713.

[10] D. William and D. Suhartono, "Text-based Depression Detection on Social Media Posts: A Systematic Literature Review," *Procedia Comput. Sci.*, vol. 179, pp. 582–589, Jan. 2021, doi: 10.1016/j.procs.2021.01.043.

[11] M. L. Joshi and N. Kanoongo, "Depression detection using emotional artificial intelligence and machine learning: A closer review," *Mater. Today Proc.*, vol. 58, pp. 217–226, Jan. 2022, doi: 10.1016/j.matpr.2022.01.467.

[12] D. Liu, X. L. Feng, F. Ahmed, M. Shahid, and J. Guo, “Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review,” *JMIR Ment. Health*, vol. 9, no. 3, p. e27244, Mar. 2022, doi: 10.2196/27244.

[13] S. Bhadra and C. J. Kumar, “An insight into diagnosis of depression using machine learning techniques: a systematic review,” *Curr. Med. Res. Opin.*, vol. 38, no. 5, pp. 749–771, May 2022, doi: 10.1080/03007995.2022.2038487.

[14] T. F. Mumu, I. J. Munni, and A. K. Das, “Depressed People Detection from Bangla Social Media Status using LSTM and CNN Approach,” *J. Eng. Adv.*, vol. 2, no. 01, pp. 41–47, Mar. 2021, doi: 10.38032/jea.2021.01.006.

[15] A. H. Uddin, D. Bapery, and A. Arif, “Depression Analysis from Social Media Data in Bangla Language Applying Deep Recurrent Neural Networks,” 2019. doi: 10.13140/RG.2.2.26377.08805/1.

[16] A. R. Brunoni, M. Lopes, and F. Fregni, “A systematic review and meta-analysis of clinical studies on major depression and BDNF levels: implications for the role of neuroplasticity in depression,” *Int. J. Neuropsychopharmacol.*, vol. 11, no. 8, pp. 1169–1180, Dec. 2008, doi: 10.1017/S1461145708009309.

[17] R. C. Kessler, J. M. Haro, S. G. Heeringa, B.-E. Pennell, and T. B. Üstün, “The World Health Organization World Mental Health Survey Initiative,” *Epidemiol. Psychiatr. Sci.*, vol. 15, no. 3, pp. 161–166, Sep. 2006, doi: 10.1017/S1121189X00004395.

[18] J. S. Girgus and K. Yang, “Gender and depression,” *Curr. Opin. Psychol.*, vol. 4, pp. 53– 60, Aug. 2015, doi: 10.1016/j.copsyc.2015.01.019.

[19] D. M. Fergusson and L. J. Woodward, “Mental health, educational, and social role outcomes of adolescents with depression,” *Arch. Gen. Psychiatry*, vol. 59, no. 3, pp. 225–231, Mar. 2002, doi: 10.1001/archpsyc.59.3.225.

[20] K. N. Hampton, “Social Media and Change in Psychological Distress Over Time: The Role of Social Causation,” *J. Comput.-Mediat. Commun.*, vol. 24, no. 5, pp. 205–222, Sep. 2019, doi: 10.1093/jcmc/zmz010.

S. Senn, M. Tlachac, R. Flores, and E. Rundensteiner, “Ensembles of BERT for Depression Classification,” in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2022, pp. 4691–4694. doi: 10.1109/EMBC48229.2022.9871120.

[21] R. A. Sukamto, Munir, and S. Handoko, “Learners mood detection using Convolutional Neural Network (CNN),” in *2017 3rd International Conference on Science in Information Technology (ICSITech)*, Oct. 2017, pp. 18–22. doi: 10.1109/ICSITech.2017.8257079.

[22] S. G. Costafreda, C. Chu, J. Ashburner, and C. H. Y. Fu, “Prognostic and Diagnostic

Potential of the Structural Neuroanatomy of Depression," *PLOS ONE*, vol. 4, no. 7, p. e6353, Jul. 2009, doi: 10.1371/journal.pone.0006353.

[23] C. H. Y. Fu *et al.*, "Pattern Classification of Sad Facial Processing: Toward the Development of Neurobiological Markers in Depression," *Biol. Psychiatry*, vol. 63, no. 7, pp. 656–662, Apr. 2008, doi: 10.1016/j.biopsych.2007.08.020.

[24] S. Sultana, M. Redoy, J. Nahian, A. K. M. Masum, and S. Abujar, *Detection of Abusive Bengali Comments for Mixed Social Media Data Using Machine Learning*. 2022. doi: 10.21203/rs.3.rs-2379359/v1.

[25] M. Hoque and U. Salma, "Detecting Level of Depression from Social Media Posts for the Low-resource Bengali Language," *J. Eng. Adv.*, vol. 4, pp. 49–56, Jun. 2023, doi: 10.38032/jea.2023.02.003.

[26] S. Zim *et al.*, "Exploring Word2vec Embedding for Sentiment Analysis of Bangla Raw and Romanized Text," 2023, pp. 677–691. doi: 10.1007/978-981-19-6634-7_48.

[27] Z. Vasha, B. Sharma, I. Esha, J. Nahian, and J. Polin, "Depression detection in social media comments data using machine learning algorithms," *Bull. Electr. Eng. Inform.*, vol. 12, pp. 987–996, Apr. 2023, doi: 10.11591/eei.v12i2.4182.

[28] K. Chanda, S. Roy, H. Mondal, and R. Bose, "To Judge Depression and Mental Illness on Social Media Using Twitter," *Univers. J. Public Health*, vol. 10, pp. 116–129, Mar. 2022, doi: 10.13189/ujph.2022.100113.

[29] R. Rahman, S. Hasan, and F. Rubel, "Identifying Sentiment and Recognizing Emotion from Social Media Data in Bangla Language," Dec. 2022, pp. 36–39. doi: 10.1109/ICECE57408.2022.10089138.

[30] S. Monjoor *et al.*, "A Comparative study for sentiment analysis of raw and translated text," Jan. 2021, pp. 220–231. doi: 10.1049/icp.2021.2678.

[31] S. Monjoor *et al.*, "A Comparative study for sentiment analysis of raw and translated text," Jan. 2021, pp. 220–231. doi: 10.1049/icp.2021.2678.

[32] N. Alvi, K. Talukder, and A. H. Uddin, "Sentiment Analysis of Bangla text using Gated Recurrent Neural Network," Feb. 2021.

[33] A. Ahmed and M. Yousuf, "Sentiment Analysis on Bangla Text Using Long Short-Term Memory (LSTM) Recurrent Neural Network," 2021, pp. 181–192. doi: 10.1007/978-981-33- 4673-4_16.

[34] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. Accessed: Apr. 05, 2024.

[Online]. Available:
https://proceedings.neurips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

[35] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, “Random Forests and Decision Trees,” *Int. J. Comput. Sci. IssuesIJCSI*, vol. 9, Sep. 2012.

[36] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[37] G. Webb, “Naïve Bayes,” 2016, pp. 1–2. doi: 10.1007/978-1-4899-7502-7_581-1.