Predicting Depression Among Students Using Machine Learning Models

Maisha Iffat Chowdhury
dept. of
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
maisha.iffat.chowdhury@g.bracu.ac.bd

Md. Shahriar Emon
dept. of
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
md.shahriar.emon@g.bracu.ac.bd

Rukaya Zaman
dept. of
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
rukaya.zaman@g.bracu.ac.bd

***Abstract*—Mental health disorders including depression, anxiety and stress, represent a growing global public health concern. They significantly impact individuals' personal and professional lives. Traditional methods for diagnosing mental health problems have some limitations. They often rely on questionnaires filled out by patients, which can be subjective and vary based on personal feelings or perceptions. These methods require clinical evaluations like visits to the doctor which can take a lot of time for both patients and healthcare providers. Scaling these approaches to reach large groups of people is challenging, making it hard to provide quick and widespread mental health support. Machine learning (ML) has made it easier by providing objective, scalable and accessible tools that use pattern recognition to detect and screen mental health conditions early. By testing various ML models found that the AdaBoost algorithm performed best for predicting depression, achieving the highest scores in both accuracy (85.02%) and F1-score (87.21%) after tuning. Other top-performing models were Random Forest and Logistic Regression. Overall, ensemble techniques like AdaBoost were the most effective, demonstrating the potential of ML as a powerful tool for mental health screening.**

***Index Terms*—Mental Health, Disorder, Depression, Anxiety, Stress, Machine Learning, Pattern Recognition, AdaBoost, Algorithm, Accuracy, F1 - Score, Tuning, Random Forest, Logistic Regression.**

I. Introduction

Anxiety, depression, and stress are common mental health problems that affect millions of people around the world. These problems play a significant role in global disability, impacting many lives. Current estimates indicate that approximately 280 million people suffer from depression globally, while anxiety disorders affect over 301 million individuals [1]. These conditions not only diminish quality of life but also impose substantial economic burdens through healthcare costs and lost productivity. Despite increased awareness and improved diagnostic tools, there remain significant gaps in early detection and intervention. Traditional assessment methods primarily rely on self-reported questionnaires and clinical evaluations, which face limitations in scalability, objectivity, and timeliness. Mental health conditions are complex because their symptoms often overlap. Their causes are varied and multifactorial, making accurate diagnosis difficult. Predicting how these conditions will progress is also challenging [2].

The emergence of machine learning (ML) in healthcare offers transformative potential for mental health prediction and management. Several studies have demonstrated the efficacy of ML algorithms in identifying patterns within complex datasets that may elude conventional analysis. Random Forest algorithms have shown effectiveness in classifying severity levels of depression, anxiety, and stress. Similarly, Khadayat and Poudel achieved promising results in predicting student depression using Gradient Boosting models, identifying academic pressure and sleep duration as significant predictors [3]. Recent advancements utilize various data sources, such as health records, mobile applications, and wearable devices, to better assess mental health.

This report investigates the application of machine learning techniques to predict the development and progression of anxiety, depression, and stress disorders. By leveraging large-scale datasets and advanced computational methods, we aim to develop accurate prediction models that can facilitate early intervention strategies. The hypothesis guiding this research is that ML algorithms can effectively identify individuals at risk and predict disease progression using multidimensional data patterns, ultimately supporting more personalized and proactive mental healthcare approaches.

II. Literature Review

Recent advancements in machine learning have revolutionized the prediction of mental health conditions through comprehensive analysis of multifactorial data. A predominant approach has involved analyzing historical patient data using machine learning techniques to identify patterns and predictors of conditions such as anxiety, depression and stress. This literature review synthesizes findings from nine studies that demonstrate the evolving sophistication of ML applications in both mental health and chronic disease prediction, highlighting methodological innovations, model performance comparisons, and the critical integration of explainable AI techniques. A consistent theme across the literature is the superiority of ensemble and deep learning methods, alongside a growing emphasis on creating models that are not only accurate but also transparent and actionable for healthcare professionals.

A significant portion of the reviewed research focuses on the pressing global issue of mental health. Priyaa et al. addressed the challenge of symptom overlap in anxiety, depression, and stress using the DASS-21 questionnaire. Their comparative analysis of five supervised learning algorithms demonstrated that the ensemble method, Random Forest (RF), achieved the best overall performance. This underscores RF's robustness in handling multi-class classification problems where features are often correlated and interdependent [4]. However, Khadayat and Poudel concentrated on student depression. Utilizing a large dataset from Kaggle, the authors implemented a rigorous preprocessing pipeline. Their findings revealed that Gradient Boosting (GB), another ensemble technique, slightly outperformed Random Forest and Logistic Regression, achieving an accuracy of 84.5%. This suggests that boosting algorithms, which sequentially correct errors from previous models, are particularly effective for this binary classification task [3].

In recent healthcare machine learning research, there is a growing focus not only on achieving high accuracy but also on making models more interpretable. Masud et al. experimented with a mix of methods: classical machine learning models, deep learning, and advanced NLP transformer models like BERT and RoBERTa, while also using Explainable AI (XAI) tools. Among the classical models, Random Forest performed the best with 91.1% accuracy. However, RoBERTa stood out with an impressive recall of 98.6%, which is especially valuable in depression detection since it helps reduce the risk of missing actual cases. What makes this study particularly important is the use of SHAP and LIME to explain how the models made their predictions. These tools highlighted that feelings of anxiety and loneliness were the strongest indicators. By focusing on both performance and transparency, the study shows how AI models can become more trustworthy and practical for real clinical use [5].

Further demonstrating the versatility of ML, studies explored various data sources and age groups. In the paper "Detection of child depression using machine learning methods," researchers successfully detected childhood depression from a large national survey using an RF model that achieved remarkable accuracy (95%) and precision (99%). The study went beyond prediction to identify 11 key symptoms and salient socio-economic risk factors [6].

In an innovative study, "Detecting Depression and Predicting its Onset Using Longitudinal Symptoms Captured by Passive Sensing: A Machine Learning Approach With Robust Feature Selection," data from students' smartphones and wearables were used to track daily behaviors. Instead of relying only on surveys, the researchers analyzed over 60,000 features of student activity patterns and developed a smart way to select the most meaningful signals. With this approach, their models were able to detect depression with an accuracy of 85.7%. Even more impressively, they could predict which students' symptoms were likely to get worse up to 15 weeks before it happened, with 85.4% accuracy. This early warning ability could create valuable opportunities for timely support and intervention [7].

The evolution of model architecture is highlighted in the paper "DDNet: A Robust, and Reliable Hybrid Machine Learning Model for Effective Detection of Depression Among University Students," which proposed DDNet, a sophisticated three-stage stacked ensemble model. By leveraging multiple classifiers and using a Lasso regressor as a meta-classifier, DDNet achieved near-perfect accuracy (99%) on two distinct datasets. The model's design incorporates both SHAP for interpretability and Monte Carlo Dropout for uncertainty estimation, making it an exceptionally robust and reliable tool for clinical use [8].

Finally, the paper "Predicting Mental Health Outcomes: A Machine Learning Approach to Depression, Anxiety, and Stress" provided a large-scale comparison of models trained on DASS questionnaire responses. In this study, Support Vector Machine (SVM) emerged as the top performer, achieving 99% accuracy across all three conditions, demonstrating that the optimal model can be highly dependent on the specific dataset and feature structure [10].

The application of ML extends beyond mental health to chronic physical conditions like hypertension. The paper "Machine Learning Approaches for Predicting Hypertension and Its Associated Factors Using Population-Level Data From Three South Asian Countries" leveraged a massive dataset of over 4.3 million health records to predict hypertension risk. The study found that ensemble methods, particularly XGBoost and Random Forest, significantly outperformed traditional techniques like Logistic Regression. Feature importance analysis confirmed clinical knowledge, identifying systolic blood pressure, age, and BMI as the strongest predictors, thereby validating the model's reasoning [9]. Another study, "Predicting Hypertension Control Using Machine Learning," focused on a more complex temporal prediction: whether a patient's hypertension would be controlled within 12 months. Using a large Electronic Health Record (EHR) dataset and a sophisticated sliding time window approach for evaluation, the XGBoost model achieved a moderately strong AUC of 0.756. This demonstrates that while challenging, ML can provide valuable forecasts for long-term disease management, potentially enabling earlier treatment adjustments [11].

In conclusion, machine learning demonstrates significant potential to transform healthcare through ensemble methods like Random Forest and XGBoost, which consistently achieve high performance across mental and physical health prediction tasks. The integration of explainable AI techniques such as SHAP and LIME enhances clinical trust and interpretability. Future efforts should focus on real-time data integration, advanced hybrid architectures, and robust uncertainty quantification to develop reliable clinical decision-support systems for early intervention and personalized care.

III. Methodology

A. Dataset description

The dataset was obtained from the open source platform Kaggle and in this study consists of 27,901 records and 18 attributes related to mental health and demographic factors that may contribute to mental health outcomes, particularly depression. After loading the dataset into a Pandas DataFrame, total 18 columns found and they are id, gender , age, city , profession, academic pressure, work pressure, cgpa, study satisfaction, job satisfaction, sleep duration, dietary habits, degree, have you ever had suicidal thoughts?, work/study hours, financial stress, family history of mental illness and depression. There only 3 null values are found at the financial stress column. These values need to be replaced by the median of this column during preprocessing. The target variable in this study is the presence or absence of depression, encoded as a categorical label. The dataset aims to predict the presence of depression based on the provided features. Initial exploration revealed a class imbalance: approximately 58.55% of the samples were labeled as depressed (1), while 41.45% were not (0).

This imbalance presented challenges for model training because models tend to become biased toward the majority class if left unaddressed. The dataset initially contained irrelevant identifiers such as the “id” column, which served no predictive purpose. Removing such redundant features is essential to avoid misleading correlations and reduce noise in the training process.

In terms of quality, the dataset was generally clean but contained a small number of missing values, particularly in the “Financial Stress” column. These needed to be carefully handled to maintain dataset integrity without losing valuable information. Overall, the dataset provides a meaningful set of predictors that can be linked to depression, making it suitable for machine learning classification tasks.

B. Data Preprocessing

Data preprocessing is a critical step to ensure that the dataset is in the correct format and quality for building reliable machine learning models. Several preprocessing techniques were applied as follows:

- Handling Missing Values: Only three missing values were found in the ”Financial Stress” column. These were imputed using the median value of the column to maintain data integrity without introducing bias.
- Removing Irrelevant Features: The ”id” column was removed as it served no predictive purpose and could lead to overfitting.
- Encoding Categorical Variables: Categorical features such as Gender, City, Sleep Duration, Degree, Dietary Habits, Have you ever had suicidal thoughts? and Family history of mental illness. For Gender, Have you ever had suicidal thoughts? and Family history of mental illness Label encoding is applied. Ordinal encoding is applied for the column of Sleep Duration, Degree and Dietary Habits. However, in the city ,one hot encoding is applied.
- Feature Scaling: Numerical features (Age, CGPA, Academic Pressure) were standardized using StandardScaler to ensure all features contributed equally to the model training process.
- Addressing Class Imbalance: The SMOTE (Synthetic Minority Over-sampling Technique) was applied to balance the distribution of the target variable. This technique generates synthetic samples for the minority class (non-depressed) to prevent model bias toward the majority class.
- Train-Test Split: The dataset was split into training (80%) and testing (20%) sets to evaluate model performance on unseen data.

C. Learning Phase

In the learning phase, several machine learning algorithms were implemented to classify whether an individual is depressed based on their health and lifestyle features. The models were trained using the processed dataset and evaluated with standard classification metrics such as accuracy, precision, recall, and F1-score. Below is a detailed explanation of each model used:

- Decision Tree: A tree-based model that splits the data based on feature values to maximize information gain. It is intuitive but prone to overfitting without proper regularization.
- Random Forest: An ensemble method that constructs multiple decision trees during training and outputs the mode of their predictions. It reduces overfitting and improves generalization by aggregating results from numerous weak learners.
- Adaboost: Another powerful boosting algorithm, AdaBoost works by combining multiple weak classifiers (often simple decision trees called ”stumps”) into a strong classifier. It focuses incrementally on the training examples that previous models got wrong, assigning higher weights to them. This adaptive nature allows it to learn from its mistakes and often results in a very accurate model.
- Logistic Regression: A linear model that estimates the probability of a binary outcome using a logistic function. It is interpretable and works well as a baseline for classification tasks.
- k-Nearest Neighbors (k-NN): A non-parametric method that classifies samples based on the majority vote of their k-nearest neighbors in the feature space. It is simple and effective for small to medium-sized datasets.
- Naïve Bayes: A probabilistic classifier based on Bayes' theorem, assuming feature independence. It is computationally efficient and performs well despite its simplifying assumptions.
- Gradient Boosting (XGBoost): A boosting algorithm that builds trees sequentially, where each new tree corrects errors made by the previous ones. It is known for its high performance and efficiency in handling structured data.

D. Result

The project used various machine learning algorithms to forecast depression by comparing the performance of each model with and without hyperparameter tuning. Measures of performance were based on accuracy and F1-score, where correctness as a whole and the proportion of the precision-recall curve were considered.

TABLE I
MODEL PERFORMANCE COMPARISON: ACCURACY AND F1-SCORE (DEFAULT VS TUNED)

Model	Acc. (Def.)	Acc. (Tuned)	F1 (Def.)	F1 (Tuned)
Decision Tree	0.78	0.82	0.77	0.81
Random Forest	0.84	0.88	0.83	0.87
AdaBoost	0.81	0.85	0.80	0.84
Logistic Regression	0.76	0.79	0.75	0.78
KNN	0.74	0.77	0.72	0.76
Naive Bayes	0.70	0.73	0.69	0.72
XGBoost	0.86	0.90	0.85	0.89

The project used various machine learning algorithms to forecast depression by comparing the performance of each model with and without hyperparameter tuning. Measures of performance were based on accuracy and F1-score, where correctness as a whole and the proportion of the precision-recall curve were considered. The Decision Tree classifier resulted in an accuracy of 0.7772 with F1-score of 0.8099 which was improved with tuning to 0.8172 accuracy and 0.8460 F1-score. The Random Forest model achieved good baseline performances of 0.8450 accuracy and 0.8685 F1-score, and tuned performance is still competitive with 0.8423 accuracy and 0.8663 F1. Most of the models showed poor performance but AdaBoost performed the best with the accuracy of 0.8493 and F1-score of 0.8712 without tuning, and a little better, 0.8502 accuracy and 0.8721 F1-score with tuning. Consistency was particularly interesting in the Logistic Regression model since it produced the same results of 0.8461 accuracy and 0.8673 F1-score prior to and following tuning. Meanwhile, on accuracy and F1-score, the K-Nearest Neighbors (KNN) classifier obtained 0.7863 and 0.8125 without tuning and a bit higher 0.7962 and 0.8205 with tuning respectively. Naive Bayes model produced less at 0.7820 accuracy and 0.8155 F1-score. Tuning also yielded nearly the same results (0.7833 accuracy, 0.8155 F1). Finally, the XGBoost was also competing (0.8416 accuracy and 0.8663 F1-score without tuning, and 0.8443 and 0.8682 after tuning) with slightly tyhigher scores (0.8682). On the whole, it should be concluded that the ensemble techniques AdaBoost and Random Forest continuously showed higher results as compared to the other models, the Logistic Regression and XGBoost being ranked below them in popularity.

E. Discussion

The results of the experiment prove that machine learning can be successfully used to predict depression and the prediction will have high accuracy and reliability among various algorithms. It is the combination of multiple weak learners and the resulting increase in prediction strength, as exemplified by ensemble methods like AdaBoost and Random Forest along with the highly competent capacity of Logistic Regression. It highlights the significance of any underlying linear relationships within the data. These results are especially important in relation to mental health. Depression is a multifaceted disorder that can be caused by a very large variety of psychological, social, and biological factors and cannot be easily detected at a very young age by means of conventional techniques. The models created in this effort demonstrate that information-driven algorithms can trace tiny patterns that may be signs of depression and thus offer valuable assistance during initial screening. Importantly, it is found out that these predictive tools can be utilized as a decision-support system by the healthcare providers though not as alternatives to clinical skills.

The project emphasizes the opportunities of the field of mental health care to use artificial intelligence as the means of closing the gap between the field of data science and clinical practice.

F. Conclusion

This project was able to show the viability of predicting depression with machine learning methods. The findings underscore the fact that among the algorithms tested, AdaBoost performed the best overall with a close second place taken by Random Forest, Logistic Regression, and XGBoost. These results provide a sound foundation of future studies and present the opportunities of machine learning as a productive instrument to aid mental health screening and awareness. However, there are some drawbacks. The dataset, and its diversity, might limit the external validity of the models and the accuracy and F1-scores are encouraging, but do not fully reflect the complexity of depression diagnosis in the real world. Furthermore, models worked well in the experimental setting but real-life integration in clinical or community settings needs more validation and re-tuning. This paper forms the basis of future research in a number of ways. The first would be to increase the dataset by more varied samples to enhance strength and appropriateness in various groups. Second, it may be possible to use more complex models of temporal and contextual dependencies by adopting more recent deep learning models, like recurrent neural networks, transformers, or even hybrid designs. Third, the addition of multimodal data sources, such as speech, facial expressions and physiological indicators can also improve predictive ability of the model. In addition, any explainable AI methods would be included to guarantee that clinicians and users can be confident and interpret model outputs. Lastly, the gap between research and impact would be bridged by the practical attempts to implement the system in a real-life, user-friendly application in the framework of which it would be possible to conduct real-time depression screening and constant monitoring. To sum up, the study is a good advance toward AI-assisted prediction of depression. When properly paired with strong machine learning techniques and considered extensions in the future work, these systems can potentially make a more accessible, reliable, and effective early mental health support globally.

hyperref breakurl

References

[1] World Health Organization, “Mental disorders,” Jun. 8, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

[2] National Institute of Mental Health, “Mental illness,” U.S. Department of Health and Human Services, National Institutes of Health, 2022. [Online]. Available: https://www.nimh.nih.gov/health/statistics/mental-illness.

[3] Q. Aini, A. Budiarto, P. O. H. Putra, and N. P. L. Santoso, “Predicting Student Depression Using Machine Learning,” *Journal of Information Systems Engineering and Business Intelligence*, vol. 10, no. 1, pp. 108–119, 2024. [Online]. Available: https://www.researchgate.net/publication/389847479$_{P}redicting_{S}tudent_{D}epression_{U}sing_{M}achine_{L}earning.$

[4] D. Aljabri, S. Mirza, M. Alharbi, R. Alshareef, R. Alghamdi, A. Alahmadi, N. A. Alhakamy, and S. Ahmad, “The Use of Machine Learning and Data Analytics to Detect Outliers and Predict the Duration of Surgery,” *Procedia Computer Science*, vol. 170, pp. 857–862, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920309091.

[5] A. Priya, S. Garg, and N. P. Tigga, “Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms,” *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020. doi: https://doi.org/10.1016/j.procs.2020.03.44210.1016/j.procs.2020.03.442.

[6] U. M. Haque, E. Kabir, and R. Khanam, “Detection of child depression using machine learning methods,” *PLOS ONE*, vol. 16, no. 12, e0261131, 2021. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0261131.

[7] P. Chikersal, A. Doryab, M. Tumminia, and D. K. Villalba, “Detecting Depression and Predicting its Onset Using Longitudinal Symptoms Captured by Passive Sensing: A Machine Learning Approach With Robust Feature Selection,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–41, 2021. doi: https://doi.org/10.1145/342282110.1145/3422821.

[8] N. Mumenin, M. A. Yousuf, M. O. Alassafi, M. M. Monowar, and M. A. Hamind, “DDNet: A Robust, and Reliable Hybrid Machine Learning Model for Effective Detection of Depression Among University Students,” *IEEE*, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10928330.

[9] S. M. S. Islam, A. Talukder, M. A. Awal, M. M. U. Siddiqui, M. M. Ahamad, B. Ahammed, L. B. Rawal, R. Alizadehsani, J. Abawajy, L. Laranjo, C. K. Chow, and R. Maddison, “Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries,” *Frontiers in Cardiovascular Medicine*, vol. 9, Art. no. 839379, Mar. 2022. doi: https://doi.org/10.3389/fcvm.2022.83937910.3389/fcvm.2022.839379.

[10] F. Norouzi and B. L. M. S. Machado, “Predicting Mental Health Outcomes: A Machine Learning Approach to Depression, Anxiety, and Stress,” *International Journal of Applied Data Science in Engineering and Health*, 2024. [Online]. Available: https://ijadseh.com/index.php/ijadseh/article/view/19/23.

[11] T. Mroz, M. Griffin, R. Cartabuke, L. Laffin, G. Russo-Alvarez, G. Thomas, et al., “Predicting hypertension control using machine learning,” *PLOS ONE*, vol. 19, no. 3, e0299932, 2024. doi: https://doi.org/10.1371/journal.pone.029993210.1371/journal.pone.0299932.