International Journal on Engineering Technology and Infrastructure Development (InJET-InDev)
www.lec.edu.np/journal
Volume 2, Issue no. 2, Page 1-7
Received Date: 10/31/2025
Revised Date: 5/13/2026
Accepted Date: 5/24/2026
ISSN: 3021-999X (Print)
https://doi.org/10.3126/injet-indev.v2i2.95663

A Machine Learning Approach to Detect Depression from Student Mental Health Surveys

Roshan Shrestha[1], Rajad Shakya[2]

[1]*Department of Electronics and Computer Engineering, Tribhuvan University, Kathmandu, Nepal, shreeroshan60@gmail.com*
[2]*Department of Electronics and Computer Engineering, Tribhuvan University, Kathmandu, Nepal, shakyarajad1@gmail.com*

Abstract

The increasing instances of depression among students have become a significant global mental health concern, creating an urgent need for effective early detection methods. This study addresses the growing concern of depression among students by proposing a machine learning based system for predicting depression risk levels using structured questionnaire data. The dataset consisted of psychological and behavioral indicators, including interest loss, feelings of sadness, sleep issues, low energy, appetite issues, low self-worth, concentration difficulties, movement changes, self-harm thoughts, scholarship status, etc. After data collection, comprehensive preprocessing and feature selection techniques were applied to improve data quality, reduce redundancy, and enhance predictive performance. Random Forest and XGBoost classifiers were implemented to classify students into different levels of depression severity. Experimental results showed that both models performed effectively in identifying at-risk students, with XGBoost achieving the highest accuracy of 88.64% on unseen test data. Feature analysis revealed that sleep issues, feelings of sadness, low energy, and self-harm thoughts were among the most influential factors affecting prediction outcomes. The proposed system demonstrates the potential of machine learning techniques for early, scalable, and data-driven mental health assessment. It can support educational institutions, counselors, and healthcare professionals in identifying students who may require timely psychological intervention, thereby contributing to improved mental well-being and preventive mental healthcare in academic environments.

Keywords: Depression, Random Forest, XGBoost, Mental Health

1. Introduction

Depression is a common and serious mental health problem that affects how people feel, think, and handle daily life. It can lead to sadness, tiredness, trouble sleeping, and a lack of interest in activities. In students, depression can be especially harmful, as it can affect their academic performance, relationships, and overall well-being. Since students often face academic pressure, homesickness, financial worries, and social stress, they are more likely to experience mental health problems during their college or university years. Globally, studies indicate that around one in three university students experience depression symptoms. A meta-analysis of 64 studies including over 100,000 college students found depression of approximately 33.6%, and anxiety. 39.0%, with increasing rates after the COVID-19 pandemic Li et al. (2022). In low and middle income countries, the depressive symptoms among students sits at about 24.4% Akhtar et al. (2020). These numbers show how common depression is among students and why early detection is so important. To detect depression early, mental health professionals often use questionnaires. These tools ask people about their mood, sleep, energy, eating habits, and thoughts. For students, questionnaires can help to find risk of depression by looking at common signs like feeling down, having trouble concentrating, or thinking negatively about themselves. When used properly, these tools can guide early support and prevent serious mental health issues.

2. Related Works

Recent advancements in machine learning have shown great promise in predicting mental health issues among various populations. Rahman et al. (2020) conducted a systematic review on machine learning applications in mental health detection using Online Social Networks (OSNs). The study found that most researchers applied text analysis and machine learning techniques for early mental health prediction and highlighted OSNs as a cost-effective data source, while also identifying challenges related to data quality and expert valida-tion. Early studies explored basic classification models such as Logistic Regression and K-Nearest Neighbors

(KNN) for mental health prediction. Vaishnavi et al. (2022) used Logistic Regression, KNN, Decision Trees, Random Forest, and Stacking techniques to classify individuals based on mental health conditions. Their results showed that stacking achieved the highest accuracy of 81.75%, outperforming simpler models. Sahu and Debbarma (2022) conducted a comparative study on students using several machine learning algorithms including Logistic Regression, Decision Trees, Random Forest and KNN. Further improvement in model performance was observed by Jain et al. (2024), who applied machine learning models to a Kaggle mental health dataset. Among various algorithms tested, the *Decision Tree classifier* provided the highest accuracy of 82%, with good precision and AUC scores. Some studies explored mental health prediction in more specialized groups.Sau and Bhakta (2017) focused on elderly patients, achieving an impressive 89% accuracy using the Random Forest model. When tested on a separate set of patients, the model achieved 91% accuracy, confirming strong generalizability with a false positive rate of just 10%. In terms of predicting mental health service utilization, Sharma et al. (2021) tested various algorithms and found that *Support Vector Machine (SVM)* outperformed others with an accuracy of 82.5%. Advanced ensemble methods have also been explored. Adeniji et al. (2022) proposed a hybrid Random Forest–Artificial Neural Network (ANN) model using a Bagging Ensemble approach. Their system predicted the likelihood of developing a mental disorder with a weighted average accuracy, precision, and recall of 84.5%, with precision alone reaching 82.5%. More recently, Mumenin et al. (2024) proposed a depression screening framework for university students using the GHQ-12 questionnaire combined with machine learning techniques. Using data collected from 804 students across universities in Bangladesh, the study evaluated 16 ML models and found that the Extremely Randomized Tree (ET) classifier achieved the best performance with an accuracy of 90.26%. The findings highlighted the effectiveness of ensemble-based approaches for early depression detection and emphasized the influence of socio-demographic and career-related factors on student mental health.

3. Methodology

A. System Architecture

B. Data Collection

This dataset was collected from university students as part of a mental health survey focused on depression, anxiety, and stress levels. The data specifically used here is focused on depression and is based on responses to PHQ-9 (Patient Health Questionnaire-9)-style questions. It originates from the MHP Dataset (Mental Health Prediction), available through open academic repositories and research publications.

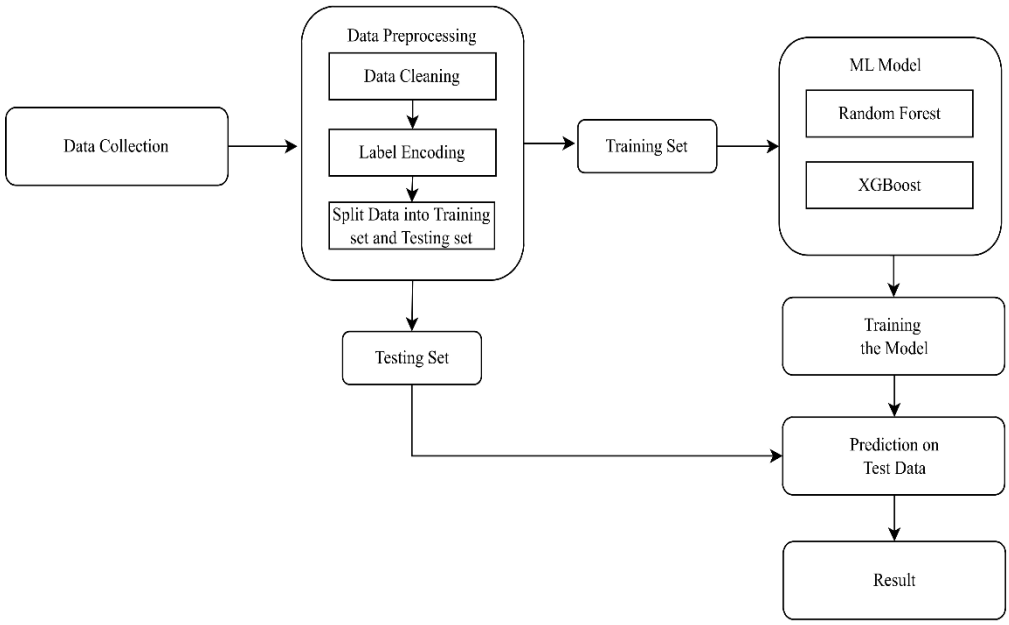

Figure 1: System Block Diagram

C. Dataset Description

The dataset is in csv format that contains self-reported psychological and demographic information of university students to assess depression levels. The dataset is designed for mental health prediction using

machine learning.

Table 1: Dataset Features and Their Data Types

Feature Name	Data Type
Age	object
Gender	object
University	object
Deaprtment	object
Academic Year	object
CGPA	object
Scholarship	object
Interest Loss	int64
Feeling Down	int64
Sleep Issues	int64
Low Energy	int64
Appetite Issues	int64
Low Self Worth	int64
Concentration Issues	int64
Movement Change	int64
Self-Harm Thoughts	int64
Depression Score	int64
Depression Level	object

- Age, Gender, University, Department, Year, CGPA describe the academic and personal background of the student.
- Scholarship: Indicates whether the student receives financial aid or not. This can relate to financial stress, which is often linked to mental health.
- Depression Indicators (PHQ-9 inspired questions): There are 9 standardized questions derived from the PHQ-9 (Patient Health Questionnaire-9), which is a widely used tool in mental health screening. These questions measure symptoms like loss of interest, hopelessness, sleep disturbance, low energy, appetite changes, feelings of failure, poor concentration, retardation and suicidal ideation.Depression Level: A categorical label (Minimal, Mild, Moderate, moderately severe, Severe) derived from the score, suitable for classification tasks.

D. Data Preprocessing

In the data pre-processing phase, several steps are involved to ensure that the data set is ready for analysis.

1. Conversion of object types
   Age which is in the format of ((18-22), (23-26) and CGPA are replaced by the value in their range, making it more compatible with algorithms that require numerical input.

2. Missing Values
   No missing values are detected in the dataset.

3. Encoding categorical variable
   Label Encoding is a technique that is used to convert binary categorical variables into numerical format. For example, for the binary feature Scholarship, Label Encoding converts categories such as Yes and No into 1 and 0, respectively, allowing machine learning models to process the data. Also, Response from the PHQ-9 is encoded with label encoder.

4. Feature Selection
   Feature selection involves identifying the most relevant features for the model while discarding less important ones. In this case, features derived from demographic data and PHQ-9 responses were considered. Based on this analysis, features such as University and Department were dropped as they

were found to be less important for predicting depression risk. Furthermore, the Depression Score was excluded, as it is highly correlated with the level of depression, making it redundant for prediction purposes.

5. Label transformation
   Depression scores, which represent the sum of nine mental health indicators from the PHQ-9 questionnaire, were transformed into categorical labels to help classification. The scores were mapped as follows: *No depression* = 0, *Minimal* = 1, *Mild* = 2, *Moderate* = 3, *Moderately Severe* = 4, and *Severe* = 5. This transformation enables the model to predict different levels of depression based on the numerical score, simplifying the classification task.
6. Train-Test Split

   The dataset was divided into two parts: the Training set (80%) and the Testing set (20%). To ensure that the class distribution is consistent across both sets, stratified sampling was used.

E. Performance Measures

All performance measures used in the analysis like confusion matrix, accuracy, precision, recall and f1-score are explained in this section.

Accuracy.

Accuracy is a fundamental metric that is used to evaluate the performance of classification models. It measures the proportion of correct predictions made by the model out of all predictions, which is calculated as in equation 1.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad \text{(Equation 1)}$$

Where,

TP=True Positive TN=True Negative FP=False Positive FN=False Negative **Precision**

Precision measures the accuracy of positive predictions made by a model. It is defined as the ratio of true positives to the sum of true positives and false positives. Mathematically,

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad \text{(Equation 2)}$$

Recall

Recall is a metric that measures the ability of a model to correctly identify all relevant instances (true positives) within a dataset as in equation 3.

$$Recall = \frac{TP}{TP + FN} \qquad \text{(Equation 3)}$$

Specificity

Specificity (also called the true negative rate) measures how well a model identifies negative cases. It is given by equation 4.

$$Specificity = \frac{TN}{TN + FP} \qquad \text{(Equation 4)}$$

F1 Score

The F1 score is a metric that is used to evaluate the performance of machine learning models, particularly in classification tasks. It provides a balanced measure of a model's precision and recall. It is given by equation 5.

$$F1\ Score = \frac{2 * (Precision * Recall)}{Precision + Recall} \qquad \text{(Equation 5)}$$

F. Random Forest

Random Forest is an ensemble method that is used for classification and regression. It builds a large number of decision trees while training the model and predicts the mode of the predictions (classification) or mean (regression) of the individual trees. The random part of Random Forest is that it allows randomness into the tree building process.

G. XGBoost

XGBoost (Extreme Gradient Boosting) is a highly effective and efficient machine learning algorithm that utilizes the gradient boosting framework. XGBoost is often used as a classification algorithm as well as regression task. XGBoost builds an ensemble of decision trees sequentially, with each tree correcting the mistakes of the previous trees. The regularization techniques introduced in XGBoost optimize the loss function and maximize the predictive accuracy, while also reducing the chances of overfitting.

4. Results and Discussion

This study applied two machine learning model: Random Forest and XGBoost to classify students into six categories of depression severity based on questionnaire data. The effectiveness of each model was evaluated using classification reports, test accuracy, and cross-validation accuracy. Random Forest achieved a test accuracy of 87%, with the highest individual class performance, while XGBoost slightly outperformed Random Forest, with a test accuracy of 88.64%.

The findings clearly shows that XGBoost is the best model which achieved the highest test accuracy (88.64%) and cross-validation mean accuracy (88.55%), outperforming Random Forest across most evaluation metrics. XGBoost showed superior generalization and maintained consistent performance across all classes, making it particularly effective for this multi-class classification task.

Table 2: Classification Report: Random Forest vs XGBoost

Class	**Random Forest**			**XGBoost**		
	Precision	Recall	F1-score	Precision	Recall	F1-score
No Depression (0)	0.90	1.00	0.95	1.00	1.00	1.00
Minimal Depression (1)	0.94	0.79	0.86	0.83	0.79	0.81
Mild Depressionn(2)	0.91	0.87	0.89	0.90	0.87	0.88
Moderate Depressio (3)	0.78	0.85	0.81	0.84	0.87	0.85
Moderately Severe (4)	0.85	0.83	0.84	0.86	0.88	0.87
Severe Depression (5)	0.95	0.95	0.95	0.95	0.93	0.94
Accuracy		0.87			0.89	
Macro Avg	0.89	0.88	0.88	0.90	0.89	0.89
Weighted Avg	0.88	0.87	0.87	0.89	0.89	0.89

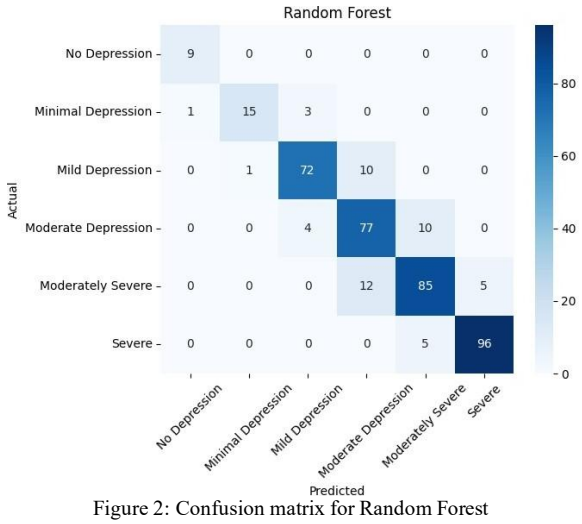

Figure 2: Confusion matrix for Random Forest

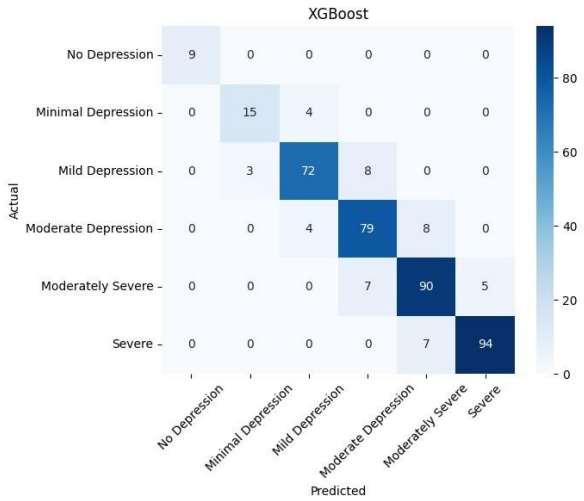

Figure 3: Confusion matrix for XGBoost

Table 3: Best Hyperparameters for XGBoost Model

Hyperparameter	Value
n_estimators	200
max_depth	4
learning_rate	0.1
subsample	0.8
colsample_bytree	0.7
reg_lambda	1
eval_metric	mlogloss

5. Conclusion

This study shows that machine learning models, particularly Random Forest and XGBoost, are effective tools for early detection of depression severity among students based on questionnaire data. By using impor-tant features such as sleep habits, stress levels, and previous mental health issues, the models perform well in classifying students, with XGBoost reaching up to 88.64% accuracy on new data. The model helps group students into clear risk levels, making it useful for early health support in schools or health centers.

Future Work

I aim to explore advanced feature engineering techniques to improve model predictive accuracy. Additionally, integrating real-time data pipelines and deploying the system for live mental health monitoring and prediction are promising directions. This would enable timely interventions and support systems, making the solution more practical and impactful in real-world applications.

Acknowledgement

I would like to thank the Department of Electronics and Computer Engineering, Thapathali Campus for giving us the opportunity, resources, and support during this study. I would also like to thank our lecturer, Er. Rajad Shakya, for his support, guidance, and constructive feedback throughout to help with the successful completion of this research work.

References

Adeniji, O., Adeyemi, S., and Ajagbe, S. (2022). An improved bagging ensemble in predicting mental disorder using hybridized random forest-artificial neural network model. *Informatica*, 46(4).

Akhtar, P., Ma, L., Waqas, A., Naveed, S., Li, Y., Rahman, A., and Wang, Y. (2020). Prevalence of depression among university students in low- and middle-income countries (lmics): a systematic review and meta-analysis. *Journal of Affective Disorders*, 274:911–919. Epub 2020 May 24.

Jain, V., Kumari, R., Bansal, P., and Dev, A. (2024). Mental health predictive analysis using machine-learning techniques. In *International Conference on Smart Computing and Communication*, pages 103–115. Springer.

Li, W., Zhao, Z., Chen, D., Peng, Y., and Lu, Z. (2022). Prevalence and associated factors of depression and anxiety symptoms among college students: a systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry*, 63(11):1222–1230. Epub 2022 Mar 16.

Mumenin, N., Kabir Hossain, A., and Hossain, M. (2024). Screening depression among university students utilizing ghq-12 and machine learning. *Heliyon*, 10.

Rahman, R. A., Omar, K., Mohd Noah, S. A., Danuri, M. S. N. M., and Al-Garadi, M. A. (2020). Application of machine learning methods in mental health detection: A systematic review. *IEEE Access*, 8:183952–183964.

Sahu, S. and Debbarma, T. (2022). Mental health prediction among students using machine learning techniques. In *International Conference on Frontiers of Intelligent Computing: Theory and Applications*, pages 529–541. Springer.

Sau, A. and Bhakta, I. (2017). Predicting anxiety and depression in elderly patients using machine learning technology. *Healthcare Technology Letters*, 4(6):238–243.

Sharma, M., Mahapatra, S., Shankar, A., and Wang, X. (2021). Predicting the utilization of mental health treatment with various machine learning algorithms. *Mental Health*, 1(2):3.

Vaishnavi, K., Kamath, U., Rao, B., and Reddy, N. (2022). Predicting mental health illness using machine learning algorithms. In *Journal of Physics: Conference Series*, volume 2161, page 012021. IOP Publishing.