Predicting Depression Using Traditional Machine Learning Models

Ovindu Atukorala
Dept. of Computer Science & Engineering
University of Moratuwa
Colombo, Sri Lanka
ovindu.atukorala@gmail.com

Ilampoornan Raguparan
Dept. of Computer Science & Engineering
University of Moratuwa
Colombo, Sri Lanka
ilampooornan@gmail.com

Uthayasankar Thayasivam
Dept. of Computer Science & Engineering
University of Moratuwa
Colombo, Sri Lanka
rtuthaya@cse.mrt.ac.lk

***Keywords*—Depression, Traditional Machine Learning, Explainable AI, Mental Health Diagnostics, AI in Public Health**

I. Introduction

Depression is a common mental health disorder that significantly impacts emotional well-being and daily functioning. Early detection is vital but often limited by subjective and resource-intensive clinical assessments. The growing availability of structured health data presents an opportunity for machine learning (ML) to enable scalable, data-driven prediction tools. While deep learning (DL) models are powerful, their need for large datasets and lack of transparency restrict clinical adoption. This study focuses on traditional ML models, which strike a better balance between performance and interpretability, and explores how feature engineering and smart preprocessing can further improve their real-world utility.

II. Literature Review

Machine learning (ML), including deep learning (DL), is widely used for depression prediction. While DL models can detect complex patterns, they often need large datasets, high computational power, and lack interpretability [1], [2], limiting their clinical use where transparency and efficiency are essential [3].

Traditional ML models—such as logistic regression, SVMs, decision trees, and ensemble methods—have shown strong performance on structured data like surveys and medical records [4], [5]. Their interpretability and low computational needs suit real-world mental health screening, but many studies rely on minimal preprocessing and basic features, reducing predictive accuracy.

For instance, Battista et al. used decision trees in youth mental health surveillance but did not explore contextual behavioral features [5]. Similarly, logistic regression is commonly used with default variables and without robust data cleaning [4]. Our study addresses these gaps through domain-specific feature engineering—such as Sleep Score, Dietary Score, and City/Job prevalence scores—which embed real-world behavioral and demographic signals into the model [6].

Moreover, many prior works rely on naive imputation methods, which we address using XGBoost-based predictive imputation for better accuracy [7].

We also emphasize evaluation metrics suited to imbalanced data—using F1-score and AUC-ROC instead of accuracy alone [8]. Finally, unlike many studies, we develop a practical depression risk tool for clinicians, bridging the gap between ML research and real-world impact.

III. Materials and Methods

A. Dataset Overview

This study uses a dataset of 234,500 samples, split into 140,700 for training and 93,800 for testing. The target variable is binary (1 = Depressed, 0 = Not Depressed). The dataset includes 20 features spanning five categories: Demographics, Academic/Work, Mental Health, Lifestyle, and Risk Factors.

B. Data Preprocessing

1) Handling Missing Data: Missing values in City, Profession, Sleep Duration, and Dietary Habits were imputed using XGBoost-based predictive modeling to preserve data integrity.

2) Feature Engineering and Selection: To enhance predictive accuracy, several features were engineered. A Higher Education Score was introduced as a binary indicator for education above "Class 12." Dietary habits were transformed into a Dietary Score, weighted by depression prevalence. A City Score represented depression prevalence per city, while a Job Score represented depression prevalence per job. Finally, a Sleep Score was derived to assess the impact of sleep duration on depression risk.

Mutual Information (MI) was used for feature selection, retaining key variables while removing low-impact ones like CGPA.

3) Encoding Categorical Features: Categorical variables (Working Professional/Student, Suicidal Thoughts) were transformed using One-Hot Encoding. Rare categories (fewer than

https://doi.org/10.31705/ADScAI.2025.42

10 occurrences) in City, Profession, and Sleep Duration were removed to prevent overfitting.

C. Machine Learning Models

1) Models Used: Several traditional models were evaluated.

2) Hyperparameter Tuning: Grid Search and Randomized Search CV were used for optimization, with early stopping to prevent overfitting.

D. Evaluation Metrics

Performance was assessed using accuracy, precision, recall, F1-score (prioritized due to class imbalance), and AUC-ROC curve for classification evaluation.

E. Model Performance and Selection

Table I compares different machine learning models for depression prediction based on accuracy, precision, recall, F1-score, and AUC-ROC.

TABLE I: Performance Comparison and Model Selection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	78.2	72.4	68.9	70.6	0.81
Naïve Bayes	76.1	70.8	65.5	68.0	0.79
Decision Tree	83.5	76.3	74.1	75.2	0.85
Random Forest	85.0	80.1	75.8	77.9	0.88
XGBoost (Final Model)	87.2	83.4	79.5	81.4	0.91

Given its superior performance, ability to handle class imbalance, and interpretability, XGBoost was selected as the final model for depression prediction.

IV. Results and Discussion

A. Results Summary

The high performance of XGBoost (AUC-ROC = 0.91) demonstrates its strong capability in distinguishing between depressed and non-depressed individuals. This suggests that traditional ML models, when properly engineered, can be effective tools for mental health diagnostics. XGBoost's high precision (83.4%) and recall (79.5%) indicate its balanced performance, ensuring both reliable predictions and low false negatives, making it a practical choice for healthcare and educational settings where early detection is critical.

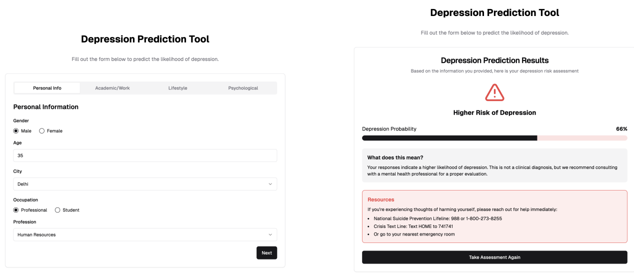
B. Depression Risk Among Students and Younger Individuals

The analysis revealed a higher depression prevalence among younger individuals, particularly students, aligning with studies linking academic pressure, financial instability, and future uncertainty to mental health risks. Also, individuals without higher education exhibited increased depression risk, likely due to economic insecurity and limited job opportunities.

V. Conclusion

This study confirms that traditional machine learning models, particularly XGBoost, offer an interpretable and effective approach to depression prediction. Key predictors include suicidal thoughts, family history, sleep duration, academic pressure, and financial stress. The findings also highlight the vulnerability of students and young professionals, emphasizing the need for targeted mental health interventions.

A major advantage of traditional machine learning over deep learning is its explainability, making it suitable for clinical applications. To bridge the gap between research and practice, this study introduces a user-friendly depression risk assessment tool designed for clinicians. The interactive user interface provides probabilistic risk scores, supporting evidence-based decision-making and enhancing early diagnosis while maintaining transparency.



(a) Dashboard interface

(b) Risk prediction results

Fig. 1: Screenshots of the depression risk assessment tool.

VI. Future Directions

Future research should integrate real-time behavioral data (e.g., wearables, smartphone activity) to improve accuracy. Clinical validation with psychologists is crucial for ethical AI implementation. Expanding scalable ML-based screening tools will enhance early detection in healthcare and education, while UI improvements will aid clinical adoption.

Future versions will show not only the depression risk but also the features influencing each prediction, helping professionals make more informed, targeted decisions. By bridging machine learning and clinical expertise, this research promotes data-driven, scalable, and proactive mental health care solutions.

References

[1] F. Fan, J. Xiong, M. Li, and G. Wang, "On Interpretability of Artificial Neural Networks: A Survey," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 6, pp. 741–760, 2020.

[2] A. B. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: A scoping review of methods and applications," *Psychological Medicine*, vol. 49, no. 9, pp. 1426–1448, 2019.

[3] A. Holzinger et al., "What do we need to build explainable AI systems for the medical domain?," *arXiv preprint arXiv:1712.09923*, 2017.

[4] D. B. Dwyer, P. Falkai, and N. Koutsouleris, "Machine learning approaches for clinical psychology and psychiatry," *Annual Review of Clinical Psychology*, vol. 14, pp. 91–118, 2018.

[5] K. Battista et al., "Examining the use of decision trees in population health surveillance research: An application to youth mental health data in the COMPASS study," *Health Promotion and Chronic Disease Prevention in Canada*, vol. 43, no. 2, pp. 73–86, 2023.

[6] B. Mihaljevic and A. Slišković, "Context-aware features for mental health prediction in university students," *International Journal of Environmental Research and Public Health*, vol. 19, no. 9, p. 5303, 2022.

[7] J. M. Jerez et al., "Missing data imputation using statistical and machine learning methods in a real breast cancer problem," *Artificial Intelligence in Medicine*, vol. 50, no. 2, pp. 105–115, 2010.

[8] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLOS ONE*, vol. 10, no. 3, e0118432, 2015.