https://dx.doi.org/10.4314/swj.v21i2.29

Full Length Research Article

PREDICTION OF DEPRESSION IN UNIVERSITY STUDENTS USING FIVE-FOLD CROSS-VALIDATION AND MACHINE LEARNING MODELS

Adedeji Oluwaseun Bukonla

Department of Computer and Information Sciences, Tai Solarin Federal University of Education, Ijagun, Ogun State, Nigeria

*Corresponding Author Email Address: adedejiob@tasued.edu.ng

ABSTRACT
Depression among university students has become a significant public health concern due to increasing academic pressure, financial instability, emotional stress, and social challenges. Early prediction of depression can help institutions implement timely interventions and improve student well-being. The study implements and compares four different classification algorithms: Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN). The analysis incorporates data augmentation techniques, handles class imbalance through SMOTE (Synthetic Minority Over-sampling Technique), and provides extensive model evaluation metrics, including standard train-test split evaluation and rigorous 5-fold cross-validation with aggregated confusion matrices. Data preprocessing involved handling missing values, normalization, label encoding, and feature selection. Model performance was evaluated using accuracy, precision, recall, F1-score, and Receiver Operating Characteristic Area Under Curve (ROC-AUC). Experimental results demonstrated that the Random Forest model achieves superior performance, with 91.5% accuracy and 0.904 cross-validation accuracy, significantly outperforming other architectures for predicting depression among students. The findings indicate that machine learning approaches can provide reliable tools for early depression risk identification in higher education institutions.

Keywords: Depression Prediction, University Students, Machine Learning, Random Forest, Support Vector Machine, Logistic Regression, Five-Fold Cross-Validation.

INTRODUCTION
Depression among students is emerging as a problem that seriously impairs their academic performance, personal life, and future career prospects. Depression is one of the most prevalent mental health disorders affecting university students worldwide (Ebert et. al., 2019; Niazi et. al., 2026). Students, particularly those in high-pressure academic environments, are more vulnerable to depression (Römhild et. al., 2025; Bhattacharjee et. al., 2025), resulting in poor academic performance, impaired social functioning, poor work performance, and an increased risk of self-harm or suicide (Mekonnen et. al., 2024; Mirza et. al., 2021). While some argue that university students are more likely to be advanced socio-economically, which is considered a protective element against depression (Moreno-Agostino, et. al., 2021; Mulaudzi, 2023) several factors make students more prone to depression, including low social support, family and relationship issues, financial constraints (adeem, et. al., 2023; Schelleman-Offermans, et. al., 2024) adjustment to a new environment, and lifestyle changes resulting in disturbed eating and sleep patterns (Zaidi, et. al., 2024; Afu, et. al., 2023).

Academic workload, examination pressure, financial challenges, social isolation, and uncertainty about future career opportunities contribute significantly to psychological distress among students (Saxena et al., 2025; Chowdhury and Das, 2025). The increasing prevalence of depression in higher education institutions has raised serious concerns among researchers, healthcare professionals, and policymakers (Nnadi et. al., 2026; Zeb et. al., 2025). Traditional approaches to diagnosing depression rely heavily on clinical interviews and self-report questionnaires, which may be time-consuming and insufficient for early detection in large student populations (Hassan et al., 2025; Li et al., 2025). Recent advances in artificial intelligence and machine learning have enabled automated prediction systems that identify mental health risks using behavioral and demographic data (Mumenin et al., 2025; Yan et al., 2025).

Machine learning algorithms have shown promising capabilities in healthcare prediction systems by learning hidden patterns from complex datasets. Techniques such as Logistic Regression, Support Vector Machines, Random Forests, and Artificial Neural Networks have been widely applied in medical diagnosis and psychological assessment (Arora et al., 2026; Alif et al., 2025).

Machine learning (artificial intelligence) offers a promising approach to identifying individuals at risk of suicidal ideation or attempts, which is crucial for preventing suicide (Weng et. al., 2020). However, a significant limitation of many current models is their dependence on clinical data or psychological factors that are not easily accessible. These models require specific data points, such as the Patient Health Questionnaire (PHQ) depression score, which are typically only available through medical or psychological evaluations. The reliance on clinical data limits the models' practical use for the public, who may not have access to such detailed assessments. For example, if someone is worried about a loved one's mental health, they likely will not have the necessary clinical mental health assessment scores to use these predictive models effectively.

This issue highlights a critical gap in the utility of these models for everyday situations and underscores the challenge of understanding at-risk behaviors in a domestic setting. The limitation points to a critical need to develop models that can operate on more commonly observed data points.

Our study aims to identify and leverage social and behavioral predictors of depression among Nigerian university students. Note that the objective is not to introduce a new predictive model to replace clinical assessment, but to assess whether social and

https://dx.doi.org/10.4314/swj.v21i2.29

behavioral predictors observable by friends, family, or individuals themselves can help predict depression.
This study aims to develop a predictive framework for identifying depression among university students using five-fold cross-validation. The major contributions of this study include:

1. Development of a machine learning-based depression prediction model.
2. Comparative analysis of LR, SVM, RF, and ANN algorithms.
3. Use of five-fold cross-validation to improve generalization performance.
4. Identification of key factors contributing to depression among students.
5. Provision of a scalable framework suitable for institutional mental health monitoring.
6. Early warning decision support system by predicting identification before clinical diagnosis, enables proactive intervention, and supports institutional mental health management

Depression among students is a serious global mental health concern, affecting academic performance, emotional well-being, and long-term development. While traditional diagnostic tools like self-reported questionnaires and clinical interviews are useful, they often suffer from subjectivity, recall bias, and limited scalability (Kotnala et. al., 2025). Several studies have explored the application of machine learning in mental health prediction.
A study by Al-Ghani (2025) reported that machine learning algorithms could accurately identify depressive symptoms using behavioral datasets. Similarly, Ademoyegun et. al. (2025 demonstrated that Random Forest classifiers achieved superior performance in predicting student anxiety and depression compared to traditional statistical models.
Research conducted in Nigeria has also shown increasing concern regarding student mental health. University students face socioeconomic instability, unemployment concerns, and academic competition, all of which increase vulnerability to depression (Ademogun, 2025).
Despite the growing body of literature, many existing studies suffer from limitations such as:

i. Small sample sizes,
ii. Lack of cross-validation,
iii. Limited feature engineering,
iv. Inadequate comparison among multiple algorithms.

This study addresses these gaps by employing five-fold cross-validation and comparing multiple machine learning approaches using standardized evaluation metrics.

Critical comparison of Related works

Criterion	Ademoyegun et al. (2025)	Nadeem et al. (2023)	Li et al. (2026)
Focus Area	Academic Performance Depression	Eating disorders and depression	AI-based depression prediction
Study Design	Cross Sectional	Cross sectional	Deep Learning Predictive Study
Geographic Context	Nigeria	Pakistans	International datasets
Depression Assessment	DASS- 21	DASS- 21	Student Depression datasets
Machine learning used	No	No	Yes (GLNet)
Explainability Technique	No	No	SHAP
Largest Predictor	Psychological Stress	Anxiety	Academic Pressure and Financial Stress
Sample size Strength	Moderate (510)	Small (79)	Very large (27,898)
Main Limitation	No Predictive modelling	Small sample size	Lack of localized validation

The reviewed studies collectively demonstrate that a complex interaction of academic, psychological, behavioral, and socioeconomic factors influences student depression. Ademoyegun et al. established the mediating role of psychological stress and sleep duration in the relationship between academic performance and depression among Nigerian students. Nadeem et al. highlighted the significant associations between eating disorders, anxiety, stress, and depression among healthcare students in Pakistan. Li et al. advanced the field by applying explainable deep learning techniques, demonstrating that academic pressure, financial stress, suicidal ideation, and poor sleep habits are among the strongest predictors of depression.
Despite these contributions, a notable gap remains in the development and validation of machine learning models using African university datasets that simultaneously integrate academic, psychological, behavioral, and socioeconomic variables. This gap provides a strong justification for this study on predicting depression among university students using machine learning.

MATERIALS AND METHODS

Research Design

This study adopted an experimental machine learning research design for predicting depression among university students.

Data Characteristics

The dataset comprises student records with the following attributes:

a) **Demographic features**: Age, Gender, Marital status

https://dx.doi.org/10.4314/swj.v21i2.29

b) **Academic features**: Course of study, Year of study, CGPA (Cumulative Grade Point Average)
c) **Clinical indicators**: Depression status, Panic attack history, Treatment-seeking behavior
d) **Target variable**: Anxiety diagnosis (binary: Yes/No)

Following preprocessing and augmentation, the dataset was expanded from approximately 100 records to 1,000 records while preserving the original class distribution.

Data Preprocessing Methodology

The preprocessing pipeline implements several critical transformations:

Column Standardization: Column names undergo systematic cleaning to ensure consistency and compatibility with Python's syntax requirements. Spaces are replaced with underscores, and special characters are removed or standardized. This standardization is crucial for programmatic access to features.

Missing Value Handling: The dataset undergoes rigorous quality control through:

i. Detection and removal of records with missing age values
ii. Conversion of age to numeric format with coercion of invalid entries
iii. Validation of categorical responses against expected values

Categorical Encoding: Machine learning algorithms require numerical input, so categorical variables must be encoded. Label encoding converts each unique category to a distinct integer value. This approach preserves the categorical nature of variables while making them computationally accessible.

Dataset Description

The dataset consists of student-related psychological, academic, and demographic attributes.

Features Used

Variable	Description
Age	Student age
Gender	Male/Female
Academic Pressure	Level of academic stress
CGPA	Academic performance
Financial Stress	Economic difficulties
Sleep Quality	Sleeping pattern quality
Social Support	Family and peer support
Social Isolation	Degree of loneliness
Substance Use	Drug/alcohol involvement
Anxiety History	Previous anxiety symptoms
Depression	Target variable (0 = No, 1 = Yes)

Data Preprocessing

Data preprocessing steps included:

1) Removal of duplicate records,
2) Missing value imputation,
3) Label encoding of categorical variables,
4) Feature normalization using Min-Max scaling,
5) Train-test splitting,
6) Feature selection.

Five-Fold Cross-Validation

To improve model reliability and reduce overfitting, five-fold cross-validation was implemented.

$k=5$

The dataset was partitioned into five equal subsets. During each iteration, four subsets were used for training and one for testing. This process was repeated five times, and the average performance metrics were computed.

Machine Learning Models

Logistic Regression (LR)

Logistic Regression is a supervised classification algorithm suitable for binary outcomes.

$$P(Y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\cdots+\beta_nX_n)}}$$

LR estimates the probability of depression occurrence based on predictor variables.

Support Vector Machine (SVM)

Support Vector Machine constructs an optimal hyperplane for classification.

$$f(x)=w^Tx+b$$

SVM performs well in high-dimensional datasets and handles nonlinear boundaries using kernel functions.

Random Forest (RF)

Random Forest is an ensemble learning algorithm based on multiple decision trees.

$$RF(x)=\frac{1}{N}\sum_{i=1}^{N}T_i(x)$$

RF reduces overfitting and improves predictive accuracy through bootstrap aggregation.

Artificial Neural Network (ANN)

ANN mimics biological neural systems through interconnected neurons.

$$y=f\left(\sum_{i=1}^{n}w_ix_i+b\right)$$

The ANN model employed:

a) Input layer,
b) Hidden layers,
c) Output layer,
d) Backpropagation learning.

Performance Evaluation Metrics

The following evaluation metrics were used:

Accuracy

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

Precision

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

Recall

$$Recall=\frac{TP}{TP+FN}$$

F1-Score

$$F1=2\times\frac{Precision\times Recall}{Precision+Recall}$$

ROC-AUC
ROC-AUC measures the classifier's ability to distinguish between depressed and non-depressed students.

EXPERIMENTAL RESULTS

Model Performance Comparison

Table 1 summarizes the model evaluation results from applying SMOTE to Machine Learning algorithms in a 5-fold cross-validation. Four Machine Learning classifiers were simulated. The models' performance was evaluated using a variety of criteria, including accuracy, precision, recall, F1 score, and advanced metrics like Matthews Correlation Coefficient (MCC) and Cohen's Kappa. These evaluation criteria gave a detailed assessment of the models' performance.
The Logistic Regression model without SMOTE consistently achieved an accuracy of 69%, while with SMOTE it was 68.5%, underscoring its reliability. The 66.9% precision with SMOTE and 66.3% precision without SMOTE, respectively, shows that the model correctly predicts depression 65% and 66% of the time. The recall (68.3% with SMOTE and 67.6% without SMOTE) indicates that 31.5% and 32.4% of actual depression cases are missed, respectively. The F1-Scores of 67% with SMOTE and 66.5% without SMOTE balance precision and recall, indicating fair model performance. The MCCs of 0.3516 with SMOTE and 0.3380 without SMOTE indicate that the model has moderate performance in predicting depression.

Excellent results are shown by the Random Forest (RF), achieving 91.5% accuracy with SMOTE and 91.0% without SMOTE, and a very high precision of 90% with SMOTE and 90.0% without SMOTE, suggesting a low proportion of false positives. The recall (91.8%) for both SMOTE-with and SMOTE-without is higher than that of Logistic Regression, capturing more actual anxiety cases. In contrast, the F1-Score (90.8%) for both SMOTE-with and SMOTE-without indicates a balanced model with high precision and recall. The Matthew's Correlation Coefficient (MCC) 0.8176 with SMOTE and 0.8176 without SMOTE indicates a strong positive relationship and strong prediction.

The Support Vector Machine shows very good results: 84.5% accuracy with SMOTE and 83% without SMOTE, and high precision of 82% with SMOTE and 81% without SMOTE, suggesting a low proportion of false positives. The recall (82.5%) with SMOTE and (84.3%) without SMOTE are improved recalls over Logistic Regression, capturing more actual anxiety cases. In contrast, the F1-Scores (832%) with SMOTE and (81.8%) without SMOTE indicate a balanced model with both high precision and recall. The Matthew's Correlation Coefficient (MCC) 0.6685 with SMOTE and 0.6421 without SMOTE indicates a moderate positive relationship and moderate prediction.

Lastly, the Artificial Neural Network (ANN) model yielded an accuracy of 91.5% with SMOTE and 88.5% without SMOTE, surpassing the Support Vector Machine in terms of accuracy, with a high precision of 89.9% with SMOTE and 90.4% without SMOTE, respectively, which suggests a low proportion of false positives. The recall (92.1%) with SMOTE and (84.2%) without SMOTE are improved recalls over Logistic Regression, capturing more actual depression cases. In contrast, the F1-Scores (90.8%) with SMOTE and (86.2%) without SMOTE indicate a balanced model with both high precision and recall. The Matthew's Correlation Coefficient (MCC) 0.8204 with SMOTE and 0.7425 without SMOTE indicates a strong positive relationship and strong prediction.

Support Vector Machine, on the other hand, achieves 83% accuracy without SMOTE and 84% with SMOTE; lastly, the ANN achieves 88.5% accuracy without SMOTE and 89.5% with SMOTE. However, a more in-depth look at the precision and recall values suggests they may not fully capture the subtle nuances of university students' daily behaviors. The RF model achieved the highest accuracy of 91.5% across iterations, while the LR and SVM models achieved 89.5% accuracy with SMOTE, closely matching the RF model's performance. This demonstrates that both models have equivalent accuracy, allowing greater flexibility in choosing the best-suited model for specific use cases.

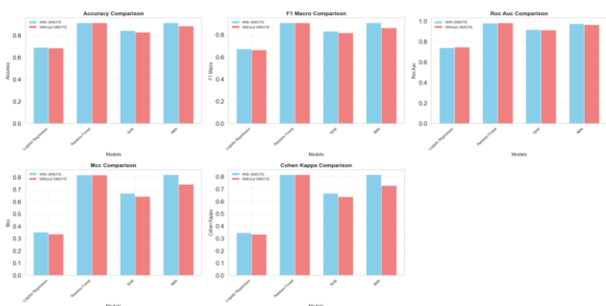
Table 1: Experimental results of the four models over a 5-fold cross-validation

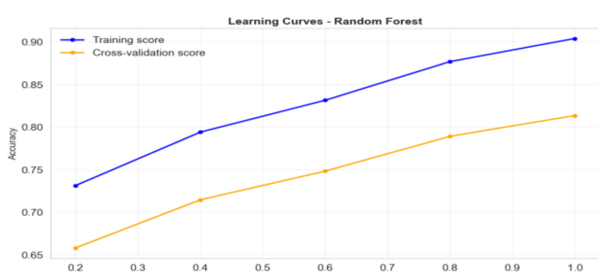
Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	MCC	CV	Stability
Logistic Regression								
SMOTE	69%	0.6687	0.6832	0.6711	0.7394	0.3516	0.6510	±0.0571
NO SMOTE	68.5%	0.6625	0.6758	0.6649	0.7439	0.3380		
Random Forest								
SMOTE	91.5%	0.9000	0.9178	0.9075	0.9794	0.8176	0.9040	±0.0194
NO SMOTE	91.5%	0.9000	0.9178	0.9075	0.9818	0.8176		
Support Vector Machine								
SMOTE	84.5%	0.8254	0.8434	0.8323	0.9132	0.6685	0.79940	±0.0857
NO SMOTE	83%	0.8104	0.8320	0.8177	0.9115	0.6421		
Artificial Neural Network								
SMOTE	91.5%	0.8993	0.9213	0.9080	0.9732	0.8204	0.7900	±0.1316
NO SMOTE	88.5%	0.9035	0.8416	0.8624	0.9080	0.7425		

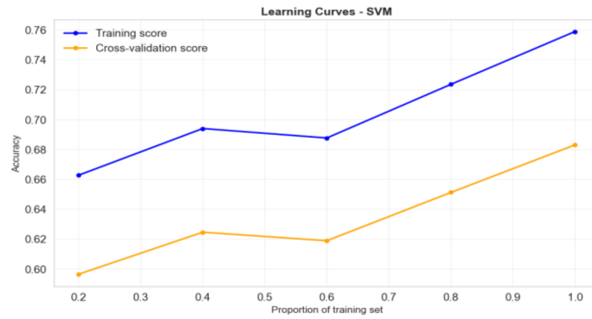
The Random Forest model outperformed other models across all evaluation metrics.

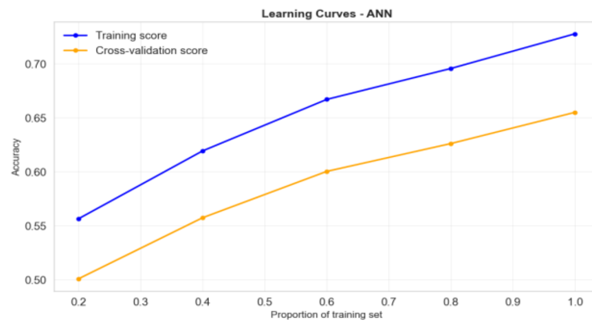
Science World Journal Vol. 21(No 2) 2026
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
https://dx.doi.org/10.4314/swj.v21i2.29

Figure 1: Confusion matrices for the four models with five (5) fold cross-validation with SMOTE









DISCUSSION

The superior performance of Random Forest can be attributed to its ensemble structure, which effectively handles nonlinear relationships and reduces overfitting. The ANN model also achieved strong performance by learning complex hidden patterns in the dataset.SVM demonstrated robust classification capabilities, particularly in handling multidimensional data. Logistic Regression achieved the lowest accuracy because of its limited ability to model complex nonlinear relationships. Feature importance analysis revealed that: Academic pressure, Sleep quality, Financial stress, Social isolation, and Anxiety history were the strongest predictors of depression among university students. These findings align with existing mental health studies indicating that psychosocial and academic stressors significantly influence depression prevalence among students.

Conclusion

This study presented a machine learning framework for predicting depression among university students using five-fold cross-validation. Four machine learning models—LR, SVM, RF, and ANN—were evaluated. The Random Forest algorithm achieved the best overall performance. This comprehensive analysis demonstrates that machine learning techniques, particularly Random Forest, can effectively predict student anxiety using demographic and academic data. The Random Forest model achieved 91.5% accuracy with excellent cross-validation stability (90.4% ± 1.94%), significantly outperforming Logistic Regression, SVM, and ANN architectures. The application of SMOTE provided meaningful improvements, particularly for ANN, which showed a 4.56% increase in F1-Score. The 5-fold cross-validation with aggregated confusion matrices establishes a rigorous methodological framework for mental health prediction research. These findings contribute to the growing body of research on computational approaches to mental health screening in educational settings, offering practical tools for early intervention

https://dx.doi.org/10.4314/swj.v21i2.29

and support. Future work should focus on external validation with diverse student populations and integration of additional behavioral and academic features. The results demonstrate that machine learning techniques can effectively support early identification of depression risk in higher education environments. Early detection systems can assist universities in implementing targeted interventions and counseling support for vulnerable students.

Recommendations

The following recommendations are proposed:

1. Universities should integrate AI-driven mental health screening systems.
2. Counseling units should monitor students with high academic stress.
3. Depression awareness programs should be strengthened.
4. Institutions should provide financial and emotional support services.
5. Future systems should incorporate real-time monitoring using mobile applications.

Future Work

Future studies may:

- Use larger real-world datasets,
- Apply deep learning techniques,
- Integrate wearable sensor data,
- Develop explainable AI models using SHAP and LIME,
- Explore multimodal mental health prediction systems.

REFERENCES

Adeem, T, Gul, B, Chhapra, R, Mahr, F, Pirani, S., & Asad, N (2023). Eating disorders in medical and nursing students of a private university: Relationship with depression, anxiety, and stress. Nursing20, 30–38.Google Scholar

Ademoyegun AB, Ibitoye A, Afolabi J, Awotidebe T, Mbada C. (2025). Academic performance and depression among university students: The serial mediating role of sleep duration and psychological stress. J CONTEMP STUD EPIDEMIOL PUBLIC HEALTH. 2025;6(1):ep25004.https://doi.org/10.29333/jconseph/17215

Afu, OM, Oguche, TE, Sammani, ZU, and Baba, G (2023). Relationship between insecurity, depression, and students' academic achievement in Nigeria: Implications for guidance. International Journal of Education and National Development 1 (3), 54–74.Google Schola

Al-Azzam, M, Abuhammad, S, Abdalrahim, A., and Hamdan-Mansour, AM (2021). Predictors of depression and anxiety among senior high school students during the COVID-19 pandemic: The context of home quarantine and online education. The Journal of School Nursing 37 (4), 241–248. https://doi.org/10.1177/1059840520988548.Google Scholar

Al-Garni, A.M., Shati, A.A., Almonawar, N.A. (2025). Prevalence of depression, anxiety, and stress among students enrolled at King Khalid University: a cross-sectional study. *BMC Public Health***25**, 354 (2025). https://doi.org/10.1186/s12889-025-21277-7

Alif, M. A. U., Ahmed, M., Ali M. Y., & Shamsi, R. (2025). A Comparative Analysis of Machine Learning Algorithms for Predicting Depression in Bangladeshi University Students. *IEEE 4th International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON*), Dhaka, Bangladesh. 760–765, doi: 10.1109/RAAICON69033.2025.11502398.

Alshahrani, AM, Al-Shahrani, MS, Miskeen, E, Alharthi, MH, Alamri, MMS, Alqahtani, MA, & Ibrahim, ME (2024). Prevalence of depressive symptoms and its correlates among male medical students at the University of Bisha, Saudi Arabia. Healthcare12(6), 640. MDPI.Google Scholar

Arora, K., Joshi, K., Manjunadha, K.S. (2026). A Four-Country Comparative Review on Stress, Anxiety, and Depression. In: Rathore, V.S., Polkowski, Z., Dey, N. (eds) Expert Applications and Security. NICE-TEAS 2025. Lecture Notes in Networks and Systems, vol 1823. Springer, Cham. https://doi.org/10.1007/978-3-032-20114-0_16

Bhattacharjee, S., Hossain, M.F., Akhy, S. (2025). Predicting depression among public university students in Bangladesh using machine learning models. *Discov Ment Health* *5*, 191 (2025). https://doi.org/10.1007/s44192-025-00332-0

Chowdhury, S., & Das, D. K. (2025). Navigation of a Research Study on Student Depression Incorporating Various Machine Learning Approaches. 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Tirupur, India, 2025, pp. 1854-1859, doi: 10.1109/ICIMIA67127.2025.11200976.

Ebert D. D, Buntrock C, & Mortier P. (2019). Prediction of major depressive disorder onset in college students. *Depress Anxiety*. 2019;36:294–304. https://doi.org/10.1002/da.22867

Hassan, A. O., Jamal, I. M., ShuabDaud Ahmed, &AbdifatahUgas Abdullahi. (2025). Predicting Student Depression using Machine Learning: A Comparative Analysis of Machine Learning Algorithms for Early Depression Detection in Students.*AITU SCIENTIFIC RESEARCH JOURNAL*, *4*(1), 28–35. https://doi.org/10.63094/AITUSRJ.25.4.1.4

Heckler, W. F., de Carvalho, J. V., & Barbosa, J. L. (2022). Machine learning for suicidal ideation identification: A systematic literature review. Computers in Human Behavior. 128:107095. https://doi.org/10.1016/j.chb.2021.107095

Kotnala, A., Arora, N. & Srivastava, S. (2025). Detecting Student Depression Using Non-Clinical Measures with Explainable Predictive Modeling. *International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN*), Greater Noida, India. 648–652, doi: 10.1109/EmergIN67762.2025.11450836.

Li X, Hu Y, Chen H, Wang X, & Gong F (2026). Explaining factors influencing students' depression with a deep learning approach. Front. Psychol. 16:1684274. doi: 10.3389/fpsyg.2025.1684274

Li Y, Sun B, Wu X, & Li Y (2025). An explainable analysis of depression status and influencing factors among nursing students. Front. Psychiatry 16:1696139. doi: 10.3389/fpsyt.2025.1696139

.Mekonnen, CK, Abate, HK, Beko, ZW, Zegeye, AF, and Azagew, AW (2024). Prevalence of depression among medical students in Africa: Systematic review and meta-analysis. PLoS One19(12), e0312281. https://doi.org/10.1371/journal.pone.0312281.Google Scholar

Mirza, AA, Milaat, WA, Ramadan, IK, Baig, M, Elmorsy, SA, Beyari, GM, Halawani, MA, Azab, RA, Zahrani, MT, and Khayat, NK (2021). Depression, anxiety, and stress among medical and non-medical students in Saudi Arabia: An epidemiological comparative cross-sectional study. Neuroscience Journal26(2), 141–151. https://doi.org/10.17712/nsj.2021.2.20200131.Google Scholar

Moazzam M, Asif MF, Farooq SMY. (2026). Prevalence and epidemiology of depression symptoms among Pakistani students: a systematic review and meta-analysis (2000–2025). *Cambridge Prisms: Global Mental Health*. 2026;13:e6. doi:10.1017/gmh.2025.10125

Moreno-Agostino, D, Wu, Y., Daskalopoulou, C, Hasan, M., Huisman, M., & Prina, M (2021). Global trends in the prevalence and incidence of depression: A systematic review and meta-analysis. Journal of Affective Disorders281, 235–243.Google Scholar

Mulaudzi, IC (2023). Challenges faced by first-year university students: Navigating the transition to higher education. Journal of Education and Human Development 12 (2), 79–87.Google Scholar

Mumenin, N., Abu Yousuf, M., Alassafi, M. O., Mostafa Monowar, M., & Abdul Hamid, M. (2025). DDNet: A Robust and Reliable Hybrid Machine Learning Model for Effective Detection of Depression Among University Students. *IEEE Access*, 13: 49334–49353. doi: 10.1109/ACCESS.2025.3552041Niazi Y,

Nnadi, L. C., Isiwu, C. P., Ding, D., Muepu D. M. and Watanobe, Y. (2026). Multi-Level Explainable AI for Predicting Student Depression Risk: Global, Subgroup, and Individual Insights. *IEEE Access*, 14. 6271-6286. doi: 10.1109/ACCESS.2026.3652631.

Saxena, V., Agarwal, R., Stuti, A., Anand, T., Pandey, T., & Verma, S. (2025). Student Depression Prediction: A Multimodal Machine Learning Approach. *5th International Conference on Advancement in Electronics & Communication Engineering (AECE),* Ghaziabad. 949-954, doi: 10.1109/AECE67531.2025.11386544.

Schelleman-Offermans, K., Dito, B. B., Gudeta, K. H., Fourie, E., Kebede, S. W., Mazzucato, V., & Jonas, K. J. (2024). Socio-economic inequities in mental health problems and wellbeing among women working in the apparel and floriculture sectors: Testing the mediating role of psychological capital, social support, and tangible assets. *BMC Public Health*24(1), 1157.Google Scholar

Weng, J. C., Lin, T. Y., Tsai, Y. H., Cheok, M. T., Chang, Y. P., & Chen, V. H. (2020). An autoencoder and machine learning model to predict suicidal ideation with brain structural imaging. Journal of Clinical Medicine. 2020; 9(3):658. Epub 20200229. https://doi.org/10.3390/jcm9030658 PMID: 32121362; PubMed Central PMCID: PMC7141277.

Yan, N., Zhang, C., & Wang, Y. (2021). Chinese burden of depressive disorders from 1990 to 2021 and projections for 2030: analysis of data from the Global Burden of Disease Study. *BMC Psychol* **13**, 30. https://doi.org/10.1186/s40359-025-02349-0

Zaidi, T., Zafar, M, Ilyas, A, Khan, M, Ghani, R, Naz, R., & Mubbashir, A. (2023). Association of burnout and depression symptoms and their prevalence among medical students in Karachi, Pakistan. *Russian Open Medical Journal*12(1), 104. https://doi.org/10.15275/rusomj.2023.0104.Google Scholar

Zeb, I., Khan, A.; Su, Yuhao, PhD. (2025). Spectrum of mental health among students: A scientometric analysis of 20 years of research on depression and anxiety. *Medicine* 104(49): p. e46751, December 05, 2025. | DOI: 10.1097/MD.0000000000046751