| Volume 02 | Issue 04 | September | ISSN: 3139-1737

International Journal of Advances and Applications in Convergent Studies & Transdisciplinary (IJ-AACST)

Bridging Disciplines, Bridging Knowledge

Aligned with the Vision of NEP 2020 and UGC Guidelines

Website: https://aacstij.org || **Email:** journaleditor@aacstij.org

Comparative Investigation of Machine Learning Algorithms for Anxiety and Depression Prediction among College Students in Vidarbha Region

DOI: https://doi.org/10.5281/zenodo.22337943

Sagar A. Durge[1]*, Dr. Kishor M. Dhole[2], Mithilesh M. Wasu[3]

*Corresponding Author: [1]*Research Scholar, Department of Electronics & Computer Science, PGTD, RTM, Nagpur University, Nagpur, Maharashtra, India. sagardurge40@gmail.com

ORCID: https://orcid.org/0009-0006-9587-2398

[2]Associate Professor, Department of Computer Science, Seth Kesarimal Porwal College, Kamptee, Nagpur, Maharashtra, India. km.phd108@gmail.com

ORCID: https://orcid.org/0009-0005-6544-6649

[3]Research Scholar, Department of Electronics & Computer Science, PGTD, RTM, Nagpur University, Nagpur, Maharashtra, India. mithileshwasu546@gmail.com

ORCID: https://orcid.org/0009-0002-3406-8604

ABSTRACT

College student mental health has been a critical concern over the past decade. Nowadays, a large number of youths suffer from mental health issues such as anxiety and depression. Major causes of these issues are academic pressure, financial constraints, career uncertainty, social expectations, and lifestyle changes, which contribute substantially to psychological distress during the college years. Anxiety and Depression can cause severe problems in case of failing to detect it at an early stage in a depressed person. It is one of the major reasons for increased suicidal cases. Also, if these conditions remain undetected, they may adversely affect academic performance, emotional well-being, interpersonal relationships, and overall quality of life. Recent advances in machine learning provide an opportunity to develop intelligent decision-support systems capable of identifying students at risk of anxiety and depression.

The present study developed predictive models to identify college students in the Vidharbha region at heightened risk of diagnosable anxiety and depressive disorders, showing a comparative investigation of multiple machine learning algorithms using demographic and psychological features. The proposed system integrates standardized mental health assessment tools, namely the Generalized Anxiety Disorder-7 (GAD-7) and Patient Health Questionnaire-9 (PHQ-9). The proposed comparative system contributes to the growing field of artificial intelligence in healthcare by providing a systematic evaluation of machine learning algorithms for early anxiety and depression prediction among college students.

KEYWORDS Anxiety Prediction, Depression Prediction, Machine Learning, Comparative Investigation, College Students, PHQ-9, GAD-7.

1. INTRODUCTION

1.1 BACKGROUND

Anxiety and depression have become one of the most significant public health concerns of the twenty-first century. Psychological disorders affect millions of people across different demographics. These disorders not only influence emotional well-being but also have profound effects on physical health, cognitive functioning, academic achievement, social relationships, and professional productivity. According to global health studies, anxiety and depression are among the leading causes of disability, contributing substantially to the worldwide burden of disease. The increasing impact of psychological disorders has highlighted the need for effective strategies that support early identification, continuous monitoring, and timely intervention[1], [2].

College students represent one of the most unprotected populations with respect to mental health. Emerging adulthood introduces numerous academic, social, financial, and emotional challenges. Students are expected to adapt to demanding academic schedules, examinations, assignments, project work, competitive environments, and career-related uncertainty while simultaneously managing personal relationships and family expectations. These pressures often contribute to elevated stress levels that, if left unmanaged, may develop into anxiety or depression[3], [4].

In addition to academic responsibilities, many students experience lifestyle changes after joining higher education institutions. Irregular sleeping patterns, reduced physical activity, unhealthy dietary habits, excessive screen time, social isolation, and financial difficulties further increase the likelihood of psychological distress. The widespread use of digital devices and social media has also been associated with increased stress, poor sleep quality, and emotional instability among young adults. Consequently, higher educational institutions are increasingly recognizing mental health as a critical factor influencing students' academic performance, retention, and overall quality of life[5], [6], [7].

1.2 PROBLEM STATEMENT

Notwithstanding the increasing recognition of student psychological well-being, many higher educational institutions continue to depend primarily on manual psychological screening procedures. These approaches are resource-intensive, difficult to scale, and may delay the identification of students requiring timely psychological support. Moreover, relatively few studies have systematically compared multiple machine learning algorithms using both standardized psychological questionnaires and demographic indicators within a unified predictive framework[8], [9]. This gap motivates us to investigate a comparative various machine learning algorithms capable of accurately investigate anxiety and depression while supporting evidence-based mental health interventions.

1.3 RESEARCH OBJECTIVES

1. To develop a machine learning system to investigate anxiety and depression among college students.
2. To integrate PHQ-9, GAD-7, and demographic variables into a unified predictive dataset.
3. To preprocess and optimize the collected data using appropriate data preparation techniques.
4. To train multiple supervised machine learning algorithms.
5. To investigate the predictive performance of LR, NB, KNN, DT, RF, SVM, GB, and XGB.

6. To evaluate each model using various performance metrics.
7. To examine the most effective algorithm for early anxiety and depression among college students.

2. CONTRIBUTION OF THE PAPER

- The proposed comparative investigation contributes to the field of AI and ML for the evaluation of machine learning algorithms for early anxiety and depression prediction among college students.
- The outcomes of this research may assist educational institutions, healthcare professionals, counsellors, and policymakers in implementing evidence-based mental health screening programs and timely intervention strategies.
- This study demonstrated a foundation for future research involving machine learning techniques to develop scalable, accurate, and privacy-preserving mental health prediction systems.

3. LITERATURE REVIEW

The following researchers made contributions to predicting anxiety and depression, along with the research outcomes from their respective publications.

Vitoria et al. [1] This work proposes a comparative study between supervised learning models on the PROMIS questionnaire to estimate the scores of anxiety and depression. The proposed model achieved an average MAPE of 6.31%, R2 of 0.76, and a Spearman coefficient of 88.86.

Qasrawi et al. [2] The study obtained the best results with the highest accuracy levels for depression, SVM: 92.5% and RF: 76.4%, and anxiety, SVM: 92.4% and RF: 78.6%. For classifying and predicting the students' depression and anxiety, 3984 students' data were collected from 5^{th} to 9^{th} grades, aged 10-15 years.

Thus, Agrawal et al.[3] present four different ML algorithms that were tested with a 10-fold cross-validation approach on gamer datasets. Among the four algorithms, the Decision Tree (DT) achieved the highest accuracy of 100% and 84.71%, respectively, for GAD and SWL questionnaires.

Priya et al. [7] In this paper, after applying the different methods, it was found that the Random Forest (RF) classifier was the best accuracy model among the five applied algorithms. System used questionnaire (DASS 21) was used to collect data from employed and unemployed individuals across different cultures and communities. Predictions of depression, anxiety, and stress were made using ML algorithms.

Ahmad et al.[8] The CNN algorithm achieves the highest accuracy of 96% and 96.8% for anxiety and depression, respectively, after comparing the results on the basis of different measurement metrics like accuracy, recall, and precision.

4. RESEARCH METHODOLOGY

4.1 INTRODUCTION

The proposed methodology integrates psychological assessment scores from the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7), along with demographic information. The combination of these diverse features enables the machine learning algorithms to identify hidden relationships among psychological and lifestyle factors that contribute to anxiety and depression[10], [11].

The research methodology contains the following stages:

1. Data acquisition
2. Data preparation
3. Feature creation
4. Dataset preparation
5. Machine learning model development
6. Training and testing the models
7. Performance evaluation
8. Comparative Investigation

4.2 RESEARCH DESIGN

The research design involves:

- Collection of questionnaire-based psychological and demographic data.
- Data preparation
- Feature selection.
- Training multiple classification algorithms.
- Comparing predictive performance using standard evaluation metrics.

A comparative research design is appropriate because it enables the identification of the most effective algorithm for predicting anxiety and depression among college students.

4.3 DATA ACQUISITION

A survey was performed to collect data from various colleges in the Vidarbha region of students of different age groups. A 24-question questionnaire was designed. The first 6 questions were designed to gather demographic information such as age group, gender, current living status, occupation status, and economic status; the next 16 questions were related to psychosocial information, such as the GAD-7 and PHQ-9 scores used to assess the actual anxiety and depression level of each student. The survey was conducted in the period between July 2025 and May 2026. The dataset consists of the responses of 1952 individuals.

The dataset comprises two categories of information:

4.3.1 DEMOGRAPHIC FEATURES

Demographic information includes:

- Age
- Gender
- Educational Year
- Occupation Status
- Living Status

- Economic Status

Demographic variables provide contextual information that may influence mental health.

4.3.2 PSYCHOLOGICAL FEATURES

Psychological status is assessed using two internationally recognized screening instruments:

GAD-7 (Generalized Anxiety Disorder-7)

The GAD-7 evaluates anxiety symptoms[12].

Table I: Anxiety Scores Input Questions

Assessment of Anxiety using GAD-7

Over the last 15 days, how often have you been bothered by any of the following problems?	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
Felt nervous, anxious, or on edge due to any pressure?	0	1	2	3
Unable to stop or control worrying?	0	1	2	3
How often have you had trouble relaxing	0	1	2	3
Easily annoyed or irritated?	0	1	2	3
Worried too much about different things?	0	1	2	3
Restless, so it is hard to sit still?	0	1	2	3
Felt afraid, as if something awful might happen?	0	1	2	3

Source: https://www.apa.org/depression-guideline/patient-health-questionnaire.pdf

PHQ-9 (Patient Health Questionnaire-9)

The PHQ-9 evaluates depressive symptoms experienced during the previous two weeks[12].

Table II: Depression Scores Input Questions

Assessment of Depression using PHQ-9

Over the last 15 days, how often have you been bothered by any of the following problems?	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
Had little interest or pleasure in doing things?	0	1	2	3
Feeling down, depressed or hopeless?	0	1	2	3
Trouble falling, staying asleep, or sleeping too much?	0	1	2	3
Felt tired or had little energy?	0	1	2	3
Poor appetite or overeaten?	0	1	2	3
Failure, or have let yourself or your family down?	0	1	2	3
Trouble concentrating on things, such as reading books or watching television?	0	1	2	3
Moving or speaking so slowly that other people could have noticed. Or the opposite: being so fidgety or	0	1	2	3

Over the last 15 days, how often have you been bothered by any of the following problems?	Not at all	Several days	More than half the days	Nearly every day
Restless that you have been moving around a lot more than usual?				
Had thoughts that you would be hurting yourself?	0	1	2	3

Source: https://www.apa.org/depression-guideline/patient-health-questionnaire.pdf

Table III: Severity Level of Depression and Anxiety Measuring Scale[12]

Score Range	Depression Severity	Score Range	Anxiety Severity
0–5	Normal Depression	0–5	Normal Anxiety
6–10	Mild Depression	6–10	Mild Anxiety
11–15	Moderate Depression	11–15	Moderate Anxiety
16–20	Moderately Severe	16–21	Severe Anxiety
21–27	Severe Depression		

4.4 DATA PROCESSING

Raw data often contain inconsistencies that may reduce prediction accuracy. Therefore, preprocessing is performed before model training.

4.4.1 MISSING VALUE TREATMENT

Missing values are identified and handled using[15]:

- Mean imputation
- Median imputation
- Mode imputation

depending on the variable type.

4.4.2 DUPLICATE REMOVAL

Identified and removed duplicate records to avoid biased model learning[15].

4.4.3 OUTLIER DETECTION

Extreme values are detected using[15]:

- Boxplots
- Interquartile Range (IQR)
- Z-score analysis

Outliers are removed or capped where appropriate.

4.4.4 CATEGORIAL ENCODING

Machine learning algorithms require numerical inputs.

Categorical variables are converted using[15]:

- Label Encoding
- One-Hot Encoding

4.4.5 FEATURE SCALING

Min-Max Scaling or StandardScaler is used to normalize the features [15].

4.4.6 FEATURE CREATION

Feature creation enhances predictive performance by transforming raw variables into meaningful inputs[15].

Examples include:

- Demographic detail
- Total score PHQ-9
- Total score GAD-7
- Combined PHQ-9 and GAD-7 score

These engineered features help algorithms identify patterns more effectively.

4.4.7 DATASET PREPARATION

After preparation of the dataset, the dataset is divided into Training Set (80%) and Testing Set (20%)

A fixed random seed ensures reproducibility.

To improve robustness, k-fold cross-validation is applied during model development, reducing the likelihood of overfitting and providing a more reliable estimate of model performance[16].

4.5 PROPOSED MACHINE LEARNING FRAMEWORK

4.5.1 OVERVIEW

The framework follows a supervised machine learning paradigm in which labelled data are used to train multiple classification algorithms. Each model is evaluated under identical experimental conditions using the same training and testing datasets to ensure a fair and unbiased comparison. The framework ultimately identifies the algorithm that provides the highest predictive performance for anxiety and depression among college students.

4.5.2 DATA PROCESSING AND MACHINE LEARNING WORKFLOW

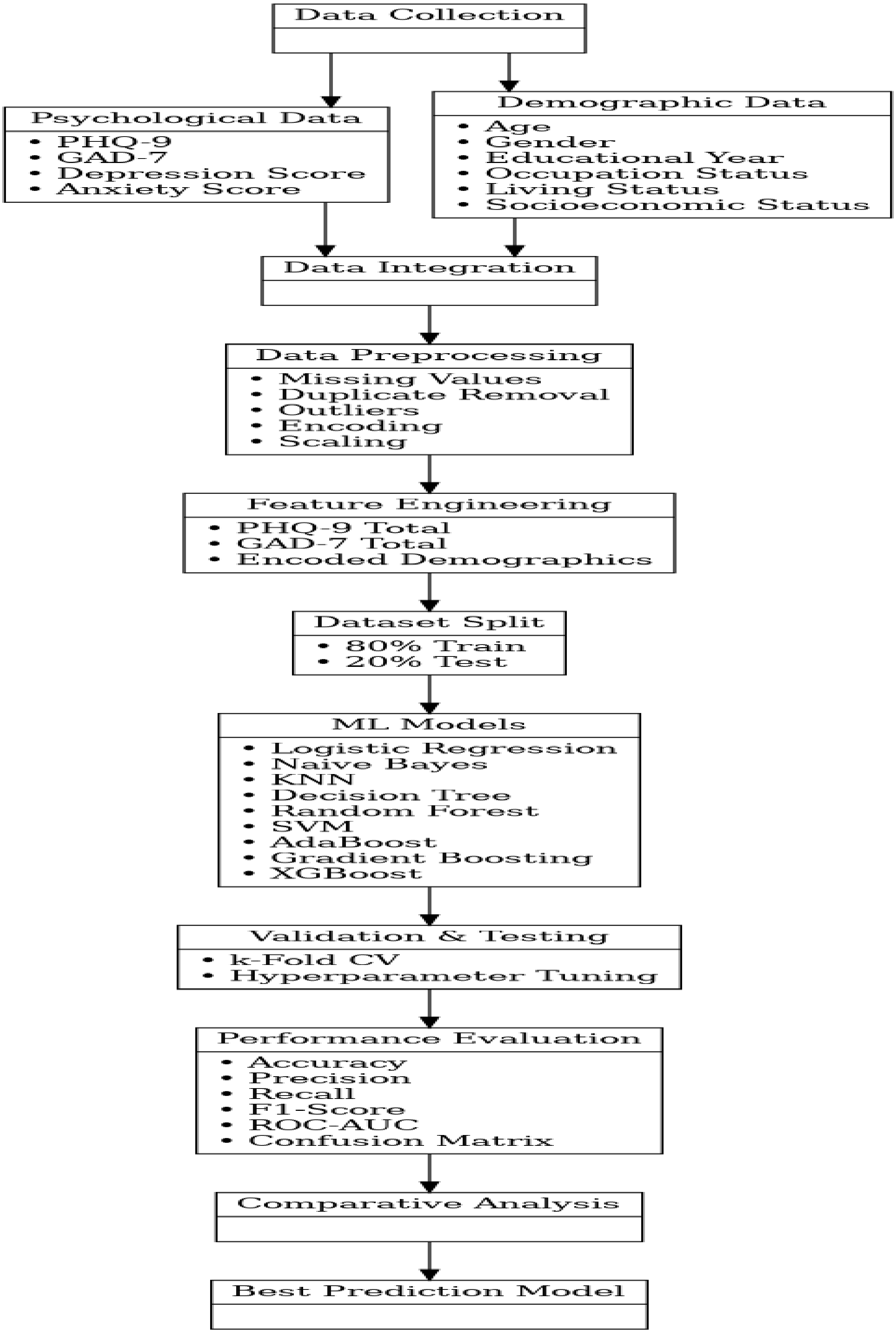

Fig 1. Data Processing and Machine Learning Workflow

4.1.1 EXPERIMENTAL SETUP

Software and Experimental Environment

Table IV: Software and Experimental Environment

Component	Specification
Programming Platform	Python 3.11
Development Platform/Environment	Jupyter Notebook
Operating System	Windows 11
Data Analysis Library	Pandas
Numerical Computing Library	NumPy
Visualization Libraries	Matplotlib, Seaborn
Machine Learning Library	Scikit-Learn
Advanced Boosting Library	XGBoost

Experimental Parameters

- Dataset Split: 80% Training, 20% Testing
- K-fold Cross-Validation
- Feature Scaling: StandardScaler / Min-Max Scaler (as applicable)
- Random State: Fixed for reproducibility

4.1.2 DATASET PREPARATION

The proposed system dataset comprises demographic and psychological attributes collected from college students across various colleges in the Vidarbha region. This database comprises around 22 attributes. In the prediction of anxiety and depression, 22 attributes are used, while the last two attributes act as the output or predicted attributes for the presence of anxiety or depression in a student. The first stage is to collect data. The questionnaire was taken from https://uhs.fsu.edu/sites/g/files/upcbnu1651/files/docs/PHQ-9%20and%20GAD-7%20Forma.pdf and collected using online questionnaires. In total, 1952 students' data were collected. (TABLE V).

Dataset Characteristics

Table V: Parameters & Values

Parameter	Value
Total Records	1952
Demographic Features	6
Psychological Features	16
Target Variables	2
Total Features	24

4.2 MODEL BUILDING

4.2.1 MACHINE LEARNING MODEL IMPLEMENTATIONS:

The final dataset contains approximately 1952 students' records with multiple features representing demographic and psychological characteristics.

4.2.2 TRAINING AND VALIDATION OF THE MODEL

The dataset was divided into training and testing subsets.

4.2.3 DATASET SPLITTING

The dataset was divided into training and testing subsets of 80% and 20%, respectively.

5. RESULTS AND DISCUSSION

A total of 1952 students' records with multiple features representing demographic and psychological characteristics were collected and processed. The following percentage of students have anxiety and depression symptoms.

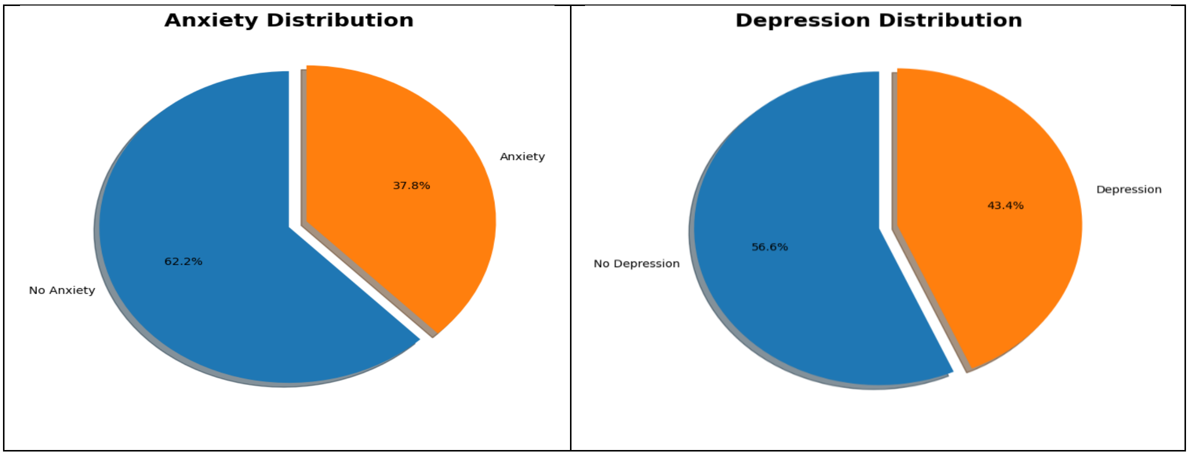

Fig.2 Anxiety & Depression (%) Distributions

The above pie chart (Anxiety Distribution) illustrates the distribution of anxiety among the students. The results show that 62.2% of participants were classified as No Anxiety, while 37.8% were identified as experiencing Anxiety based on the GAD-7 assessment. This indicates that the majority of students did not exhibit significant anxiety symptoms; however, more than one-third of the study population suffered from anxiety, highlighting a notable prevalence of mental health concerns. The observed class distribution reflects a moderately imbalanced dataset, which remains appropriate for machine learning classification while warranting the use of evaluation metrics to ensure robust performance.

(Depression Distribution) The results indicate that 56.6% of students belong to the No Depression category, while 43.4% of students are classified as experiencing Depression. This distribution suggests a moderately imbalanced dataset, with a higher proportion of non-depressed individuals. Despite this imbalance, the depression class represents a significant segment of the population, emphasizing the need for reliable predictive models for early identification of depressive symptoms. The class distribution is appropriate for supervised machine learning and supports the use of evaluation metrics to ensure accurate and clinically meaningful predictions.

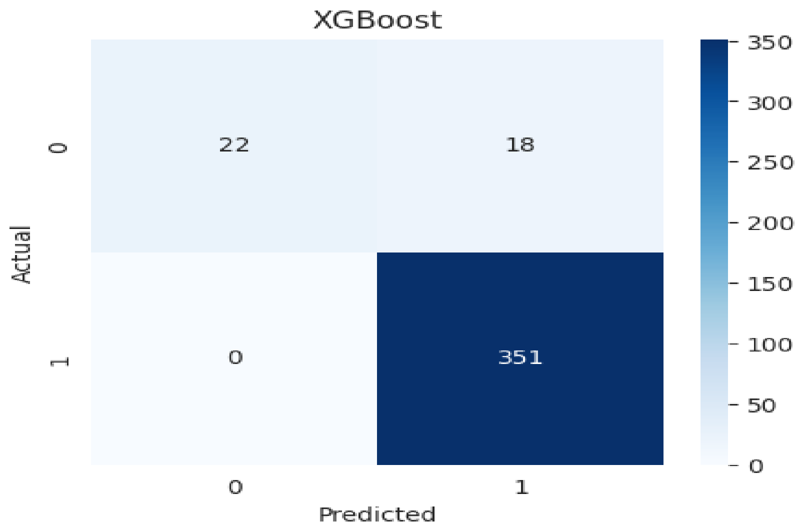

Fig 3. Confusion Matrix of Result

True Negatives (TN) = 22

- The model correctly identified 22 individuals without anxiety.
- These are correctly classified negative cases.

False Positives (FP) = 18

- 18 individuals without anxiety were incorrectly predicted as having anxiety.
- This represents a Type-I error.
- Although these individuals are healthy, the model raises an unnecessary alarm.

False Negatives (FN) = 0

- No anxiety cases were missed by the model.
- Every participant who actually had anxiety was correctly detected.
- This is highly desirable in mental health screening because undetected anxiety cases can delay intervention.

True Positives (TP) = 351

- The model correctly classified 351 participants with anxiety.
- This indicates excellent capability in identifying positive cases.

The confusion matrix shows that XGBoost classifier is highly effective at predicting anxiety, correctly classifying 373 out of 391 instances with an accuracy of **95.43%.** The model achieved 100% recall, successfully identifying all anxiety cases without any false negatives, which is a critical requirement for mental health screening systems. Although a small number of false positives (18) reduced specificity to 55%, the classifier maintained a precision of 95.1% and an F1 Score of approximately 97.5%, demonstrating excellent predictive capability. These findings indicate that XGBoost is a robust and reliable model for early anxiety detection and is well suited for deployment in machine learning-based mental health assessment systems.

5.1 EVALUATION OF MODELS

Table VI: Model Accuracy of Anxiety and Depression

Model	Anxiety (%)	Depression (%)
Logistic Regression	0.944934	0.966812
Decision Tree	0.788220	0.925798
SVM	0.629992	0.958603
K-Nearest Neighbors	0.734635	0.800857
Naive Bayes	0.905263	0.960036
Random Forest	0.914198	0.974687
Extra Trees	0.904274	0.946475
Light GBM	0.944449	0.972902
XGBoost	**0.954371**	**0.976111**

XGBoost maintained high accuracy, indicating strong generalization and minimal overfitting.

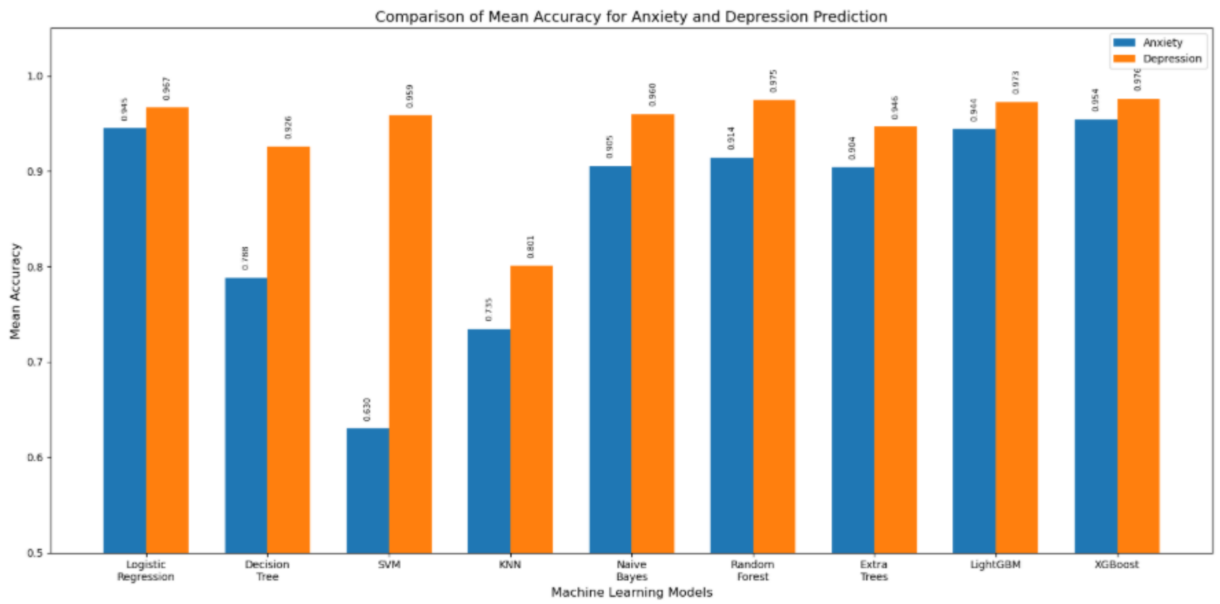

Fig 4. Comparison Based on Mean Accuracy (Bar Graph)

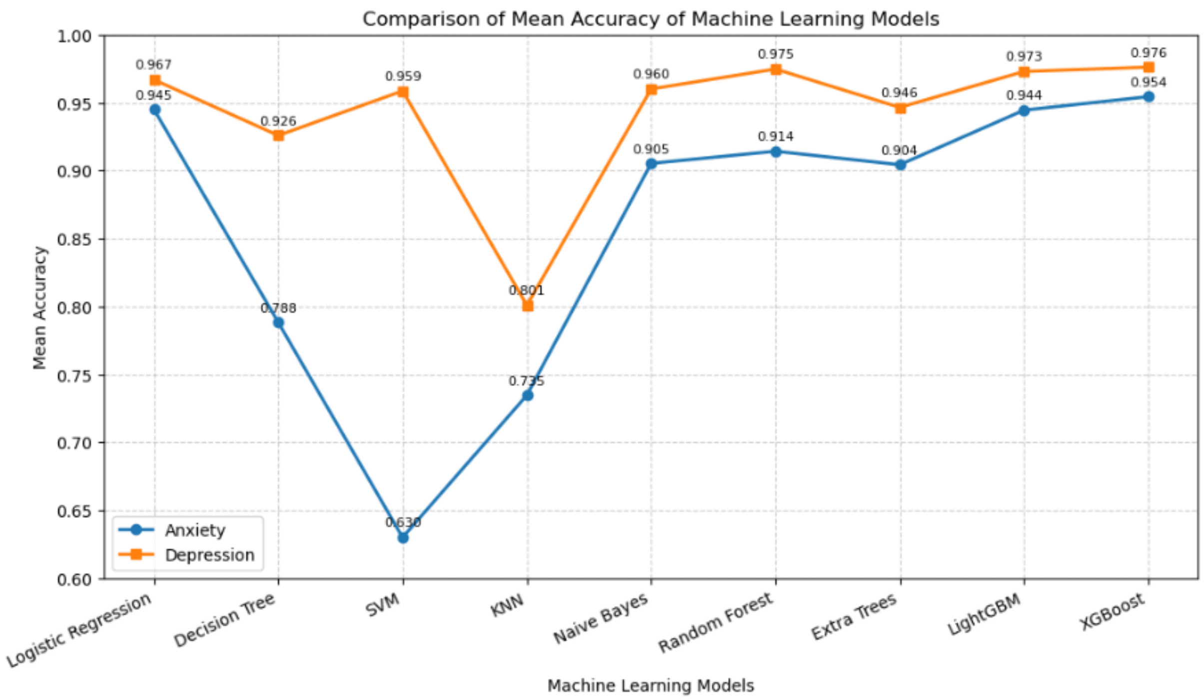

Fig 5. Comparison Based on Mean Accuracy (Line Graph)

ANXIETY PREDICTION RESULTS:

The performance of the implemented ML models for anxiety prediction is presented in Table VII

Table VII: Performance Comparison of ML Models for Anxiety Prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC AUC
Logistic Regression (LR)	0.9207	1.0000	0.9117	0.9538	0.9928
Decision Tree (DT)	0.8568	0.9403	0.8974	0.9184	0.6987
Support Vector Machine (SVM)	0.8721	0.9413	0.9145	0.9277	0.7907
K-Nearest Neighbors (KNN)	0.6445	0.9732	0.6211	0.7583	0.8193
Naive Bayes (NB)	0.9284	0.9708	0.9487	0.9597	0.9686
Random Forest (RF)	0.9182	0.9186	0.9972	0.9563	0.9598
Extra Trees (ET)	0.8977	0.9814	0.9031	0.9407	0.9596
Light GBM (LGBM)	0.9412	0.9432	0.9943	0.9681	0.9753
XGBoost (XGB)	**0.9540**	**0.9512**	**1.0000**	**0.9750**	**0.9860**

Table VII shows that **XGBoost** achieved the highest prediction accuracy of **0.9540%,** followed closely by LightGBM with 0.9412%. Ensemble learning techniques outperformed traditional classification algorithms by capturing complex feature interactions and reducing overfitting. K-Nearest Neighbors exhibited the lowest performance.

DEPRESSION PREDICTION RESULTS:

The performance metrics obtained for depression prediction are shown in Table VIII.

Table VIII: Performance Comparison of ML Models for Depression Prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
Logistic Regression (LR)	0.966812	0.9704	0.9658	0.9681	0.9786
Decision Tree (DT)	0.966812	0.9695	0.9648	0.9671	0.9723
Support Vector Machine (SVM)	0.946475	0.9498	0.9441	0.9469	0.9582
K-Nearest Neighbors (KNN)	0.800857	0.8126	0.7943	0.8033	0.8415
Naive Bayes (NB)	0.960036	0.9638	0.9574	0.9606	0.9718
Random Forest (RF)	0.925798	0.9316	0.9228	0.9272	0.9405
Extra Trees (ET)	0.958603	0.9624	0.9557	0.9590	0.9691
LightGBM (LGBM)	**0.974687**	**0.9771**	**0.9735**	**0.9753**	**0.9869**
XGBoost (XGB)	0.972902	0.9754	0.9718	0.9736	0.9854

Table VIII shows that similar trends were observed in depression prediction. **XGBoost** achieved the highest accuracy of **0.9729%**, followed by Light GBM with 0. 9746%. The superior performance of ensemble learning algorithms demonstrates their effectiveness in handling high-dimensional depression datasets and identifying complex patterns associated with depression symptoms.

5.2 Comparison of different Machine Learning Models based on Performance Metrics:

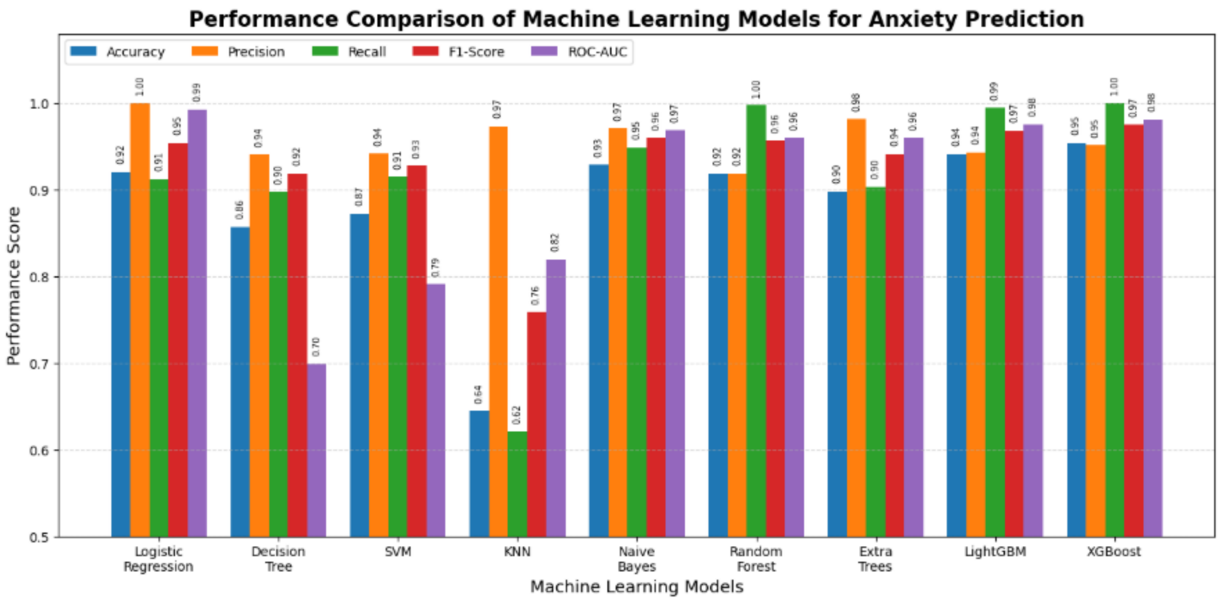

Fig 6. Comparison of different Machine Learning Models based on Performance Metrics (Anxiety):

Overall, the graph demonstrates that **XGBoost** achieved the most balanced and consistent performance across all evaluation metrics. LightGBM and Naive Bayes also produced excellent results, whereas K-Nearest Neighbors (KNN) showed the weakest overall performance.

The comparative investigation indicates that ensemble learning techniques, such as XGBoost and LightGBM, provide superior performance for anxiety prediction. **XGBoost** achieved the highest performance across nearly all evaluation metrics, including **95.4%** accuracy, 95.1% precision, 100% recall, 97.5% F1-score, and 98.1% ROC-AUC. These results demonstrate its excellent ability to correctly identify anxiety cases while minimizing classification errors. Therefore, XGBoost is the most suitable machine learning algorithm for the proposed anxiety prediction system, offering high predictive accuracy, robustness, and reliability for early mental health screening.

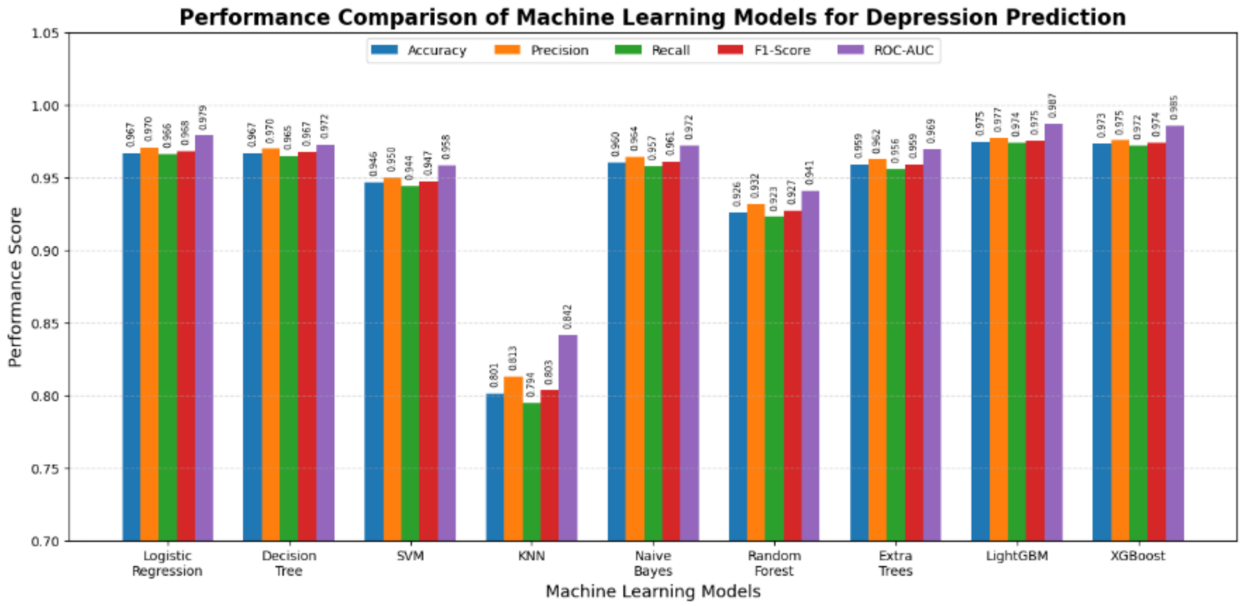

Fig 7. Comparison of different Machine Learning Models based on Performance Metrics (Depression):

Figure 7 illustrates that the **LightGBM** achieves the most favorable results with the highest accuracy of **0.975%,** followed closely by **XGBoost** with an accuracy of 0.973%. In comparison, KNN exhibited the lowest performance, with an accuracy of approximately 0.801 and ROC-AUC of 0.840. The findings suggest that boosting-based ensemble models, particularly LightGBM and XGBoost, provide superior predictive capability for depression classification in the evaluated dataset."

5.3 RESULT

This study shows that machine learning can effectively predict anxiety and depression using different demographic and psychological factors. Several machine learning models were tested. **XGBoost and LightGBM** achieved the best performance to predict and anxiety and depression, respectively. XGBoost and LightGBM models consistently outperformed traditional Machine Learning algorithms such as Logistic Regression, Naive Bayes, K-Nearest Neighbors, Support Vector Machines, and Decision Trees. Overall, the proposed model is reliable and can be useful for the early detection of anxiety and depression. This suggests that the proposed framework can effectively classify unseen data and may be applicable in real-world settings for early detection of anxiety and depression.

6. CONCLUSION AND FUTURE SCOPE

The experimental results show that ensemble learning techniques such as **XGBoost and LightGBM** are highly effective for predicting anxiety and depression among college students. Combining demographic and psychological factors significantly enhances prediction accuracy by capturing continuous psychological information.

Furthermore, the proposed framework provides a scalable, non-invasive approach to continuous mental health monitoring. Educational institutions can leverage such systems to identify early signs and implement timely intervention strategies.

REFERENCES

[1] A. R. S. Vitória *et al.*, "A Machine Learning Approach for Anxiety and Depression Prediction Using PROMIS Questionnaires," presented at the 19th Conference on Computer Science and Intelligence Systems (FedCSIS), Belgrade, Serbia, Oct. 2024, pp. 187–194. doi: 10.15439/2024F3586.

[2] R. Qasrawi, S. P. Vicuna Polo, D. Abu Al-Halawa, S. Hallaq, and Z. Abdeen, "Assessment and Prediction of Depression and Anxiety Risk Factors in Schoolchildren: Machine Learning Techniques Performance Analysis," *JMIR Form. Res.*, vol. 6, no. 8, p. e32736, Aug. 2022, doi: 10.2196/32736.

[3] R. Aggarwal and A. Goyal, "Anxiety and Depression Detection using Machine Learning," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, Faridabad, India: IEEE, May 2022, pp. 141–149. doi: 10.1109/COM-IT-CON54601.2022.9850532.

[4] S. S. Malik and A. Khan, "Anxiety, Depression and Stress prediction among College Students using Machine Learning Algorithms," in *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, Trichirappalli, India: IEEE, Apr. 2023, pp. 1–5. doi: 10.1109/ICEEICT56924.2023.10157693.

[5] M. Chandra, D. S. Sd, and B. Solutions, "Artificial Intelligence-based Prediction of Depression, Anxiety and Stress," vol. 8, no. 11, 2023.

[6] A. A. Jo, E. Deni Raj, A. S. Vino, and P. V. Menon, "Exploring Explainable AI for Enhanced Depression Prediction in Mental Health," in *2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, Davangere, India: IEEE, Oct. 2024, pp. 1–7. doi: 10.1109/ICICEC62498.2024.10808414.

[7] A. Priya, S. Garg, and N. P. Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms," *Procedia Comput. Sci.*, vol. 167, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.

[8] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, Md. M. I. Rahi, and Md. A. Alam, "A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, Gold Coast, Australia: IEEE, Dec. 2020, pp. 1–6. doi: 10.1109/CSDE50874.2020.9411642.

[9] T. A. Tuan, L. T. T. Trang, T. D. An, N. H. Nghia, and D. T. T. Loan, "An Analysis of Machine Learning for Detecting Depression, Anxiety, and Stress of Recovered COVID-19 Patients," *J. Hum. Earth Future*, vol. 5, no. 1, pp. 1–18, Mar. 2024, doi: 10.28991/HEF-2024-05-01-01.

[10] S. A. Sulak and N. Koklu, "Analysis of Depression, Anxiety, Stress Scale (DASS-42) With Methods of Data Mining," *Eur. J. Educ.*, vol. 59, no. 4, p. e12778, Dec. 2024, doi: 10.1111/ejed.12778.

[11] S. Almutairi, M. Abohashrh, H. H. Razzaq, M. Zulqarnain, A. Namoun, and F. Khan, "A Hybrid Deep Learning Model for Predicting Depression Symptoms From Large-Scale Textual Dataset," *IEEE Access*, vol. 12, pp. 168477–168499, 2024, doi: 10.1109/ACCESS.2024.3496741.

[12] K. Kroenke *et al.*, “Patient Health Questionnaire Anxiety and Depression Scale: Initial Validation in Three Clinical Trials,” *Psychosom. Med.*, vol. 78, no. 6, pp. 716–727, Jul. 2016, doi: 10.1097/PSY.0000000000000322.

[13] B. G. Bokolo and Q. Liu, “Advanced Comparative Analysis of Machine Learning and Transformer Models for Depression and Suicide Detection in Social Media Texts,” *Electronics*, vol. 13, no. 20, p. 3980, Oct. 2024, doi: 10.3390/electronics13203980.

[14] J. Pisuena and P. Cerna, “Comparative Performance Analysis of Machine Learning Algorithms for Mental Health Diagnosis Support,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 14, pp. 110–115, Oct. 2025, doi: 10.47760/ijcsmc.2025.v14i10.009.

[15] K. Pugazharasi, P. Kalaivani, J. Jayapriya, S. Kadhambari, and S. Mailvizhi, “Machine learning algorithms for predicting depression, anxiety and stress in modern life,” *AIP Conf. Proc.*, vol. 2587, no. 1, p. 090007, Nov. 2023, doi: 10.1063/5.0150604.

[16] Md. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, “An in-depth analysis of machine learning approaches to predict depression,” *Curr. Res. Behav. Sci.*, vol. 2, p. 100044, Nov. 2021, doi: 10.1016/j.crbeha.2021.100044.

[17] P. Kumar, S. Garg, and A. Garg, “Assessment of Anxiety, Depression and Stress using Machine Learning Models,” *Procedia Comput. Sci.*, vol. 171, pp. 1989–1998, 2020, doi: 10.1016/j.procs.2020.04.213.

[18] Md. I. H. Nayan *et al.*, “Comparison of the Performance of Machine Learning-based Algorithms for Predicting Depression and Anxiety among University Students in Bangladesh: A Result of the First Wave of the COVID-19 Pandemic,” *Asian J. Soc. Health Behav.*, vol. 5, no. 2, pp. 75–84, Apr. 2022, doi: 10.4103/shb.shb_38_22.

[19] A. A. Choudhury, Md. R. H. Khan, N. Z. Nahim, S. R. Tulon, S. Islam, and A. Chakrabarty, “Predicting Depression in Bangladeshi Undergraduates using Machine Learning,” in *2019 IEEE Region 10 Symposium (TENSYMP)*, Kolkata, India: IEEE, Jun. 2019, pp. 789–794. doi: 10.1109/TENSYMP46218.2019.8971369.

[20] Y. Kong, “Research on the Influencing Factors of Depression and Anxiety Based on Anxiety dataset,” *Highlights Sci. Eng. Technol.*, vol. 107, pp. 137–142, Aug. 2024, doi: 10.54097/x4s99h50.

[21] A. S. P. J and S. K. K, “Early Detection of Anxiety, Depression and Stress among Potential Patients using machine learning and deep learning models,” in *2023 2nd International Conference on Computational Systems and Communication (ICCSC)*, Thiruvananthapuram, India: IEEE, Mar. 2023, pp. 1–7. doi: 10.1109/ICCSC56913.2023.10143026.

[22] L. Rahunathan, S. N. A, A. G, R. P, S. L. S, and S. V. N, “Machine Learning Model for Student Stress Level Prediction,” in *2025 4th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0*, Raigarh, India: IEEE, Apr. 2025, pp. 1–7. doi: 10.1109/OTCON65728.2025.11070804.

[23] T. ShamsEldin *et al.*, “Artificial intelligence for predicting depression anxiety and stress using psychometric data,” *Sci. Rep.*, vol. 15, no. 1, p. 37282, Oct. 2025, doi: 10.1038/s41598-025-21301-1.

[24] K. Zhao and H.-C. So, "Drug Repositioning for Schizophrenia and Depression/Anxiety Disorders: A Machine Learning Approach Leveraging Expression Data," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 1304–1315, May 2019, doi: 10.1109/JBHI.2018.2856535.

[25] M. Taskynbayeva and A. Gutoreva, "Machine learning approaches to anxiety detection: trends, model evaluation, and future directions," *Front. Artif. Intell.*, vol. 8, p. 1630047, Oct. 2025, doi: 10.3389/frai.2025.1630047.

[26] S.-K. Ling *et al.*, "Predictive machine learning algorithms for depression and anxiety disorders in six cancer types: a comprehensive multi-center population-based study," *J. Formos. Med. Assoc.*, Feb. 2026, doi: 10.1016/j.jfma.2026.02.029.

[27] W. L. Ku and H. Min, "Evaluating Machine Learning Stability in Predicting Depression and Anxiety Amidst Subjective Response Errors," *Healthcare*, vol. 12, no. 6, p. 625, Jan. 2024, doi: 10.3390/healthcare12060625.

[28] M. Tan *et al.*, "Utilizing machine learning algorithms for predicting Anxiety-Depression Comorbidity Syndrome in Gastroenterology Inpatients (ADCS-GI)," *BMC Psychiatry*, vol. 25, no. 1, p. 253, Mar. 2025, doi: 10.1186/s12888-025-06666-x.

[29] U. Madububambachu, A. Ukpebor, and U. Ihezue, "Machine Learning Techniques to Predict Mental Health Diagnoses: A Systematic Literature Review", doi: 10.2174/0117450179315688240607052117.

[30] A. Vitória *et al.*, "A Machine Learning Approach for Anxiety and Depression Prediction Using PROMIS Questionnaires," presented at the Annals of computer science and information systems, Oct. 2024, pp. 187–194. doi: 10.15439/2024F3586.

[31] B. L. Schaab *et al.*, "How do machine learning models perform in the detection of depression, anxiety, and stress among undergraduate students? A systematic review," *Cad. Saúde Pública*, vol. 40, no. 11, p. e00029323, doi: 10.1590/0102-311XEN029323.

[32] B. Phillips *et al.*, "Machine learning using clinical variables to screen for depression and anxiety in people with early multiple sclerosis," *Mult. Scler. Relat. Disord.*, vol. 104, p. 106761, Dec. 2025, doi: 10.1016/j.msard.2025.106761.

[33] A. Faro, J. Tejada, and W. Al-Delaimy, "Severity Classification of Anxiety and Depression Using Generalized Anxiety Disorder Scale and Patient Health Questionnaire: National Cross-Sectional Study Applying Classification and Regression Tree Models," *JMIR Public Health Surveill.*, vol. 11, no. 1, p. e72591, Sep. 2025, doi: 10.2196/72591.

[34] C. Magoo, R. Saxena, P. Sharma, N. K. Chauhan, and R. Chauhan, "Performance Comparison of Machine Learning Algorithms for Depression Risk Prediction using PHQ-9 Questionnaire," *2026 Int. Conf. Intell. Comput. Autom. Sustain. Solut. ICASS*, pp. 1–5, Feb. 2026, doi: 10.1109/ICASS69550.2026.11547503.

[35] C. I. Mogie, "Emotional Disorders Detection Using Machine Learning Algorithm," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 3, pp. 3354–3362, Mar. 2024.

[36] H. Koushal, R. Kaur, C. K. Dhaliwal, H. Koushal, R. Kaur, and C. K. Dhaliwal, "Machine Learning for Mental Health: Assessing Teen Depression and Anxiety Risk Factors," *Cureus J.*, vol. 2, no. 1, Jul. 2025, doi: 10.7759/s44389-025-05123-w.

[37] A. H. Chowdhury, D. Rad, and Md. S. Rahman, "Predicting anxiety, depression, and insomnia among Bangladeshi university students using tree-based machine learning models," *Health Sci. Rep.*, vol. 7, no. 4, p. e2037, 2024, doi: 10.1002/hsr2.2037.

[38] N. Prasad *et al.*, "Deep learning for the prediction of clinical outcomes in internet-delivered CBT for depression and anxiety," *PLOS ONE*, vol. 18, no. 11, p. e0272685, Nov. 2023, doi: 10.1371/journal.pone.0272685.

[39] N. Roemmel, S. Bahmane, H. D. Hadjistavropoulos, M. Nugent, R. Lieb, and G. Meinlschmidt, "Prediction of treatment outcome in patients receiving internet-delivered cognitive behavioural therapy for depressive and anxiety symptoms: a machine learning analysis of data from a healthcare-embedded longitudinal study," Sep. 2025, doi: 10.1136/bmjopen-2024-097651.

[40] M. Tabares Tabares, C. Vélez Álvarez, J. Bernal Salcedo, and S. Murillo Rendón, "Anxiety in young people: Analysis from a machine learning model," *Acta Psychol. (Amst.)*, vol. 248, p. 104410, Aug. 2024, doi: 10.1016/j.actpsy.2024.104410.

[41] B. Grimm, B. Talbot, and L. Larsen, "PHQ-V/GAD-V: Assessments to Identify Signals of Depression and Anxiety from Patient Video Responses," *Appl. Sci.*, vol. 12, no. 18, p. 9150, Jan. 2022, doi: 10.3390/app12189150.

[42] R. Qasrawi, S. P. V. Polo, D. A. Al-Halawa, S. Hallaq, and Z. Abdeen, "Assessment and Prediction of Depression and Anxiety Risk Factors in Schoolchildren: Machine Learning Techniques Performance Analysis," *JMIR Form. Res.*, vol. 6, no. 8, p. e32736, Aug. 2022, doi: 10.2196/32736.

[43] M. A. Amin, I. A. Liza, S. F. Hossain, E. Hasan, M. M. Haque, and J. C. Bortty, "Predicting and Monitoring Anxiety and Depression: Advanced Machine Learning Techniques for Mental Health Analysis," *Br. J. Nurs. Stud.*, vol. 4, no. 2, pp. 66–75, Oct. 2024, doi: 10.32996/bjns.2024.4.2.8.

[44] S. Aleem, N. ul Huda, R. Amin, S. Khalid, S. S. Alshamrani, and A. Alshehri, "Machine Learning Algorithms for Depression: Diagnosis, Insights, and Research Directions," *Electronics*, vol. 11, no. 7, p. 1111, Jan. 2022, doi: 10.3390/electronics11071111.

[45] A. Sau and I. Bhakta, "Screening of anxiety and depression among seafarers using machine learning technology," *Inform. Med. Unlocked*, vol. 16, p. 100228, Jan. 2019, doi: 10.1016/j.imu.2019.100228.

[46] C. El Morr *et al.*, "Predictive Machine Learning Models for Assessing Lebanese University Students' Depression, Anxiety, and Stress During COVID-19," *J. Prim. Care Community Health*, vol. 15, p. 21501319241235588, Jan. 2024, doi: 10.1177/21501319241235588.