**Predicting Depression, Anxiety, and Stress among University Students through Statistical and Machine Learning Algorithms**

Mohammad Kamal Hossain1, a, Arpita Haldar2, a, Md Sabbir Hossain3. a, b, Prosenjit Basak Arka3, a , Nazmuz Sayeed2,a

Mohammad Kamal Hossain

[kamalbsmrstu@gmail.com](mailto:kamalbsmrstu@gmail.com)

Arpita Haldar

[arpitahaldar229@gmail.com](mailto:arpitahaldar229@gmail.com)

Md Sabbir Hossain

[bmsabbirhossainsakib@gmail.com](mailto:bmsabbirhossainsakib@gmail.com)

Prosenjit Basak Arka

[prosenjitbasakarka@gmail.com](mailto:prosenjitbasakarka@gmail.com)

Nazmuz Sayeed

[nazmus.sayeed9445@gmail.com](mailto:nazmus.sayeed9445@gmail.com)

1. Associate Professor, Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh
2. Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh
3. Department of Statistics, Shahjalal University of Science and Technology, Sylhet, Bangladesh
4. Equal Contribution
5. Corresponding Author:

Md Sabbir Hossain,

Email: [bmsabbirhossainsakib@gmail.com](mailto:bmsabbirhossainsakib@gmail.com)

Phone: +8801311919804

**Abstract**

**Background**

Mental health disorders, including depression, anxiety, and stress, are significant global challenges affecting individuals' well-being and daily functioning. This study aims to evaluate the effectiveness of various statistical and machine learning (ML) algorithms in predicting these conditions among university students in Bangladesh to identify the best model for early detection in a resource-limited setting.

**Methods**

Data were collected from 384 students at Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh using stratified random sampling and self-administered questionnaires, including socio-demographic details and the BDASS-21 tool. Statistical and ML algorithms multinomial logistic regression (MLR), random forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), Gradient Boosting (GB), and K-Nearest Neighbors (KNN) were evaluated for predicting mental health severity, with performance metrics used to determine the most accurate model.

**Results**

The study found that most students experienced moderate levels of depression (30.7%), with high levels of anxiety, where 35.4% reported extremely severe anxiety. Stress levels were predominantly normal (37.5%), with some students experiencing severe stress (21.1%). Among ML models, SVM performed best in predicting depression, anxiety, and stress with high accuracy and balanced performance metrics, while NB showed the poorest performance across most measures. Key predictors of mental health issues included feelings of low self-worth, lack of motivation, and a sense of meaninglessness.

**Conclusion**

The SVM model proved superior in predicting mental health conditions among university students, outperforming other models in accuracy and consistency. KNN, RF, and GB also performed well, but SVM is recommended as the primary tool for mental health prediction, with complementary models enhancing overall prediction accuracy.

**Keywords:** Mental health, Depression, Anxiety, Stress, Machine Learning, University Student

**Introduction**

The World health Organization (WHO) highlights a significant global challenge, with approximately 450 million individuals grappling with mental health issues. Mental health disorders, specifically, rank as the fourth leading cause of the global disease burden and are projected to become the second leading cause by 2020 [1]. Psychological well-being is influenced by the levels of anxiety, depression, and stress.

Depression, a prevalent and severe mental illness, manifests through persistent feelings of sadness and a loss of interest, significantly impacting a person's ability to concentrate on daily tasks, work, and studies. Alarming statistics reveal that one in 15 adult’s experiences depression within a given year, underscoring its pervasive impact and its substantial role in over half of suicide attempts [2].

Anxiety is characterized by a distressing sense of inner turmoil, often accompanied by nervous behaviors such as pacing, physical symptoms, and excessive worrying [3]. Unlike fear, which is a response to an immediate or perceived threat, anxiety can also involve concerns about potential future threats. Factors such a family history of mental health issues, substance misuse, including caffeine, alcohol and poverty are contributors to anxiety [4-5].

Like anxiety and depression, stress is also a significant risk factor for health and well-being. Stress is described as a mental or physical reaction resulting from an individual's interaction with their environment, influenced by their cognitive evaluation of the stimuli [6]. Research has highlighted that stress is a complex phenomenon, varying greatly depending on individual temperament, experiences, situations, and environmental factors [7].

Globally, the prevalence of moderate to severe levels of depression stands at 60.8%, anxiety at 73%, and stress at 62.4% [8-9]. While often overlooked, mental health issues are increasingly posing a significant threat to low- and middle-income countries, including Bangladesh. In Bangladesh, nearly 7 million individuals are affected by depressive and anxiety disorders [10]. In 2012, an estimated 10,167 suicides occurred, and suicide attempts were reported by 4% of boys and 6% of girls among youths aged 13 to 17 [10]. The prevalence rates for depression, anxiety, and stress in Bangladesh have been reported to reach 54.3%, 64.8%, and 59.0%, respectively [11-12]. Therefore, researchers exhibit significant interest in the field of depression testing, employing various assessment methods such as DASS21, DASS42, PHQ-9, HADS, and the widely recognized The DASS21, short for Depression Anxiety Stress Scales, is a widely used self-report questionnaire comprising 21 items designed to assess the severity of depression, anxiety, and stress symptoms. Each domain contains seven items, with respondents rating the frequency or severity of experiences over the past week using a Likert scale. It serves as a standardized and reliable tool for evaluating emotional distress across these three domains, commonly employed in clinical and research settings for screening, monitoring, and aiding in diagnostic evaluation, treatment planning, and outcome assessment [13].

In recent years, researchers have employed machine learning (ML) algorithms to predict and classify anxiety, depression, and stress enhancing our understanding of these issues. A study conducted in India utilized five distinct classification techniques: Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). Among these methods, NB achieved the highest accuracy [14]. Another study applied various ML algorithms, including Logistic Regression, Catboost, NB, RF, and SVM for classification. The researchers discovered that Catboost achieved the highest accuracy and precision, with rates of 82.6% and 84.1%, respectively [15]. A study in the USA investigated predictors of depression and Post-Traumatic Stress Disorder (PTSD) among Twitter users, employing the Hidden Markov Model (HMM) to identify increased probabilities of PTSD. The analysis revealed that 31.4% of the dataset exhibited signs of depression, while 24% were affected by PTSD [16]. In a separate study, tweets from 135 participants on Amazon Mechanical Turk (MTurk) were analyzed using decision tree classification to assess suicide risk, achieving an impressive prediction accuracy of 92% [17]. Additionally, research conducted in China extracted real-time data from Twitter and used psychiatric stressors to label tweets associated with suicidal behavior. In this study, the Convolutional Neural Network (CNN) outperformed SVM and Extra Trees (ET), achieving a precision of 78% in identifying tweets with suicidal tendencies [18].

This study is focusing on university students, we aim to assess depression, anxiety, and stress diagnosis using multiple ML algorithms. In resource-constrained environments like Bangladesh, where adequate counseling resources are lacking, administrators struggle to evaluate students' mental health effectively. Thus, there's a pressing need for a computerized model capable of identifying students' mental health status. To determine the optimal model, we employed seven statistical and machine learning techniques including Multinomial Logistic Regression (MLR), RF, KNN, SVM, NB, and Gradient Boosting (GB), utilizing data from both male and female students at Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU), a public university in Bangladesh. The DASS-21 categorizes depression, anxiety, and stress into severity levels: normal, mild, moderate, and severe. Performance metrics such as Accuracy, Positive Predictive Value (PPV), Sensitivity, Negative Predictive Value (NPV), Specificity, Prevalence, Detection Rate (DR), Detection Prevalence (DP), and Balanced Accuracy (BA) were computed from confusion matrices to identify the best model for depression, anxiety and stress prediction.

**Methods**

**Data Source and Study Design**

The present cross-sectional study collected primary data from students of BSMRSTU across different academic years and departments using a stratified random sampling technique. A multi-indicator survey was designed to explore various topics related to depression, anxiety, and stress among BSMRSTU students. A sample size of 384 participants was gathered for this purpose. Participation was voluntary, and participants were assured of the confidentiality of their information and individual identity. After obtaining consent, the data collection process commenced. Under the close supervision of the research student and principal investigator, five well-trained graduate students collected the primary data. A training session on data collection procedures was conducted before the face-to-face interviews. Data was collected via a self-administered questionnaire comprising two sections: the first section focused on socio-demographic, socioeconomic, and behavioral traits, while the second section utilized the BDASS-21 tool. The detailed questionnaires are available in the supplementary file’ssection [**Questionnaires**].

**Features**

The study gathered a comprehensive range of socio-demographic information, including age, gender, year of study, faculty of study, CGPA in honors, religion, academic performance, accommodation type (hall, mess/home), permanent residence (urban or rural), family living systems, socioeconomic status (upper, middle, and lower class), relationship status (single, in a relationship, or married), and parents' educational and occupational backgrounds. BSMSTU comprises three Institutes and eight Faculties, including 34 Departments spanning Engineering, Science, Biological Science, Social Science, Humanities, Business Studies, Agriculture, and Law. Data collection included participants from all faculties but excluded institutes as they do not admit undergraduate students. Participants from Agriculture and Law were combined due to their small sample size for statistical analysis. Behavioral factors were also included in the study. Participants were asked whether they smoked cigarettes ("Yes" or "No"), and whether they exercised for at least 20 minutes per day, including activities such as walking, playing sports, games, cycling, or swimming. Following Mamun and Griffiths (2019), participants categorized their daily average sleep time as normal (6-7 hours), short (<6 hours), or long (>7 hours). Additionally, they were questioned about their Internet usage and the number of hours spent studying each week.

**Outcome Variable**

The primary outcome variables in this study were depression, anxiety, and stress, assessed using the 21-item Bangla Depression Anxiety Stress Scale (BDASS-21). The BDASS-21, a short version of the original 42-item scale developed by Lovibond and Lovibond (1995), was a self-report instrument designed to measure the severity of these three mental health conditions [21]. The Bangla version of DASS-21, validated by Alim et al. (2017a, b), Mamun et al. (2019), was utilized [12,19-20]. The scale comprised 21 items divided into three subscales: depression (seven items), anxiety (seven items), and stress (seven items), each rated on a four-point Likert scale ranging from never (0) to always (3). The severity of depression, anxiety, and stress was categorized using the following rating ranges: for depression, normal (0–9), mild (10–13), moderate (14–20), severe (21–27), and extremely severe (28 and above); for anxiety, normal (0–7), mild (8–9), moderate (10–14), severe (15–19), and extremely severe (20 and above); and for stress, normal (0–14), mild (15–18), moderate (19–25), severe (26–33), and extremely severe (34 and above).

**Statistical Techniques**

**Multinomial Logistic Regression**

MLR is almost the same as general LR. The difference is: In Multinomial Logistic Regression, there’s more than two categories of dependent variables. This algorithm is used as a classification method and a generalized form of logistic regression to solve multiclass problems.

**Machine Learning Techniques**

The machine learning algorithms were applied using the R programming language with RStudio version 3.5. These algorithms predicted the proportion of students experiencing depression, anxiety, and stress based on the severity of their symptoms. The study employed several supervised ML algorithms, including MLR, RF, NB, SVM, GB, and KNN, to the collected data for early detection of depression, anxiety, and stress. The goal was to identify the algorithm that most accurately predicted the severity of depression, anxiety, and stress. Here is a summary of each ML algorithm.

**Random Forest**

The RF algorithm features hyperparameters that include the number of trees and the maximum depth of each tree, which influence the evaluation of model interactions and decision-making criteria. It builds multiple decision trees and aggregates their results to achieve more precise and dependable predictions [22].

**Naïve Bayes**

In ML, NB is known as the most straightforward algorithm which uses basically Bayes theorem and simply creates a probabilistic model. But the performance might be slow, and the result also might be less efficient if the size of the data was large [23].

**Support Vector Machine**

SVM is a supervised classification algorithm founded on statistical learning theory. It aims to determine the best decision function by defining a hyperplane that effectively separates two classes. SVM employs four commonly used kernel functions: linear, polynomial, sigmoid, and radial basis function (RBF). In this study, the RBF kernel was utilized [24].

**K-Nearest Neighbors**

KNN, a non-parametric algorithm, is widely used for both classification and regression tasks. Being an instance-based learning method, KNN classifies new instances by comparing them to the most similar examples in the training set, with similarity assessed through distance metrics. The key parameter in KNN is the K value, which denotes the number of nearest neighbors considered. When K is set to 1, the new instance is assigned to the class of its closest neighbor [25].

**Gradient Boosting**

The GB classifier combines multiple weak models in a sequential manner to build new models iteratively. Each subsequent model aims to reduce the loss function. GB employs the gradient descent technique to evaluate the loss function. To prevent overfitting, it is essential to stop boosting at the right time, using specific stopping criteria. These criteria might include setting a maximum number of models or defining a threshold for predictive accuracy [26].

**Performance Measures**

Performance metrics are essential for evaluating the effectiveness of predictive models, particularly in classification tasks. Performance metrics such as Accuracy, PPV, Sensitivity, NPV, Specificity, Prevalence, DR, DP, and BA were calculated to evaluate the ML algorithms.

The performance measurements are calculated using the following equations (1 – 9) from the confusion matrix in **table 1**.

**[Insert Table 1 Here]**

**Table 1:** Confusion Matrices from different ML Algorithms for Depression, Anxiety, and Stress

**Results**

**Descriptive statistics of depression, anxiety and stress among university students**

Among the 384 university students surveyed, the majority exhibited moderate levels of depression (30.7%), followed by normal (24.7%), mild (18.2%), extremely severe (15.9%), and severe (10.4%) levels. Anxiety levels were notably high among the students, with 35.4% experiencing extremely severe anxiety. This was followed by moderate anxiety (26.8%), severe anxiety (15.4%), normal anxiety (15.1%), and mild anxiety (7.3%). In terms of stress, a significant portion of the students reported normal stress levels (37.5%). This was followed by mild (18.5%), severe (21.1%), moderate (17.4%), and extremely severe (5.5%) stress levels [**Table 2**].

**[Insert Table 2 Here]**

**Table 2:** The Distribution of Depression, Anxiety, and Stress among University Students

**Evaluation of different ML model performance for predicting depression among university students**

**Table 3** evaluates the performance of various ML models for predicting depression using different performance measures, including Accuracy, PPV, Sensitivity, NPV, Specificity, Prevalence, DR, DP, and BA. Starting with Accuracy, the SVM model leads with a value of 0.829, followed by KNN at 0.763, indicating their robustness in overall correct predictions. RF and MLR show reasonable performance with accuracies of 0.7105 and 0.6974, respectively, while NB and GB lag behind with accuracies of 0.579 and 0.684.

In terms of PPV, the RF achieves perfect scores (1) in most of its categories, showcasing its reliability in predicting true positive cases. SVM and GB also score 1 in certain categories, reflecting their strength in predicting positive cases correctly. MLR and KNN also show strong PPV values, particularly MLR with a high of 0.936 in category 1, indicating reliable performance. On the other hand, Naïve Bayes performs worst in predicting true positives cases.

Sensitivity, a crucial measure for identifying true positive cases, is highest in SVM, KNN and RF models, which achieve perfect sensitivity value (1) in one category, ensuring all actual positive cases are identified. MLR also performs well with high sensitivity values (above 0.609), ensuring a good detection rate of true positives.

For NPV, which measures the ability to identify true negatives, SVM, KNN, RF, and NB score 1 in one category each, highlighting their efficiency in ruling out non-depressive cases. MLR and GB follow closely with lowest values of 0.830, ensuring reliable identification of negatives.

Specificity, indicating the true negative rate, sees perfect scores (1) in RF and NB in several categories, ensuring accurate identification of non-depressive cases. SVM and GB models also show high specificity values (1), suggesting strong performance in this metric. MLR and KNN also perform well with the lowest value of 0.811.

Analyzing DR, SVM, Random Forest, and MLR display consistent values, suggesting stable model behavior across different categories. GB and KNN also exhibit balanced DR, indicating consistent model performance. On the other hand, NB performs worst in DR.

DP values indicate the proportion of predicted positive cases, with SVM, GB, and MLR models showing balanced performance. Finally, BA, which combines sensitivity and specificity, is highest in SVM at 0.971, followed by KNN (0.956) and RF (0.921), showcasing overall strong model performance. On the other hand, NB performs worst in terms of BA.

[**Table 3**].

**[Insert Table 3 Here]**

**Table 3:** Evaluation of performance measurements from various ML algorithms for Depression

**Evaluation of different ML model performance for predicting Anxiety among university students**

**Table 4** evaluates the performance of various ML models for predicting anxiety using different performance measures. Starting with Accuracy, KNN leads with a value of 0.811, demonstrating its robustness in making correct predictions. The SVM model follows closely with an accuracy of 0.797, indicating strong overall performance. RF and GB show reasonable accuracy with values of 0.716 in both cases, while MLR and NB have lower accuracies, 0.689 and 0.541, respectively, with NB showing particularly poor performance.

In terms of PPV, KNN achieves high values with a maximum of 0.931, reflecting its reliability in predicting true positives. SVM and GB also perform well with high PPV values of 0.926 and 0.92, respectively. RF and MLR also display strong PPV values, though not consistently perfect. In contrast, the NB model shows the least PPV performance in most of the categories.

Sensitivity, crucial for identifying true positive cases, is highest for KNN, which achieves perfect sensitivity (1) in several categories, ensuring it identifies all actual positive cases. SVM also demonstrates high sensitivity values of 0.926. MLR, RF and GB exhibit moderate sensitivity, with RF showing particularly strong performance in detecting anxiety. On the other hand, NB shows the worst sensitivity value of 0 in several categories, reflecting its poor performance in predicting positive cases.

For NPV, which measures the ability to identify true negatives, KNN consistently scores high value (1) in multiple categories, indicating their efficiency in ruling out non-anxiety cases. Like KNN, the rest of the model follows with strong NPV with lowest value of 0.851.

Specificity, reflecting the true negative rate, NB consistently scores high value (1) in multiple categories. KNN models, with perfect scores (1) also performed well. The rest of the models also perform well with the lowest specificity values of 0.851, suggesting the effectiveness in identifying non-anxiety cases.

In DR, MLR and SVM exhibit balanced values, indicating stable performance. RF, KNN and GB also display consistent detection rates, suggesting reliable model behavior across different categories. In contrast, NB performs worst in terms of DR. Like DR, DP values are balanced across GB, and SVM. Finally, BA, which combines sensitivity and specificity, is highest in KNN at 0.976, followed by SVM at 0.942, showcasing overall strong model performance [**Table 4**].

**[Insert Table 4 Here]**

**Table 4:** Evaluation of performance measurements from various ML algorithms for Anxiety

**Evaluation of different ML model performance for predicting stress among university students**

**Table 5** evaluates the performance of various ML models for predicting stress using different performance measures. Starting with accuracy, SVM leads with a value of 0.84, indicating strong overall performance in making correct predictions. KNN follows closely with an accuracy of 0.811, showing robust performance. RF and GB exhibit solid performance with accuracies of 0.773 and 0.8, respectively. MLR and NB have lower accuracies, 0.573 and 0.373, respectively, with NB showing particularly poor performance.

In terms of PPV, GB achieves high values with a maximum of 1.0, reflecting its reliability in predicting true positives. SVM and KNN also perform well, with high PPV values of 0.964 and 0.931, respectively. RF, MLR and GB display strong PPV values. The NB model performs worst in terms of predicting true positives.

Sensitivity, which measures the ability to identify true positive cases, is highest for SVM, RF, KNN and GB achieving perfect sensitivity (1) in several categories. MLR also demonstrates high sensitivity, notably 0.821. Like other metrics, Naïve Bayes performs worst in terms of sensitivity.

For NPV, which measures the ability to identify true negatives, SVM, KNN, GB, and RF consistently score high with perfect value of 1, indicating their efficiency in ruling out non-stress cases. MLR and NB show more variable NPV performance, particularly NB shows worst performance.

Specificity, reflecting the true negative rate, is high across NB, KNN, and GB models, showing perfect scores (1) in multiple categories. The rest of the models also demonstrate strong specificity with the lowest value of 0.839, suggesting effective identification of non-stress cases.

In case of detection rate, SVM exhibits balanced values, indicating stable performance. RF and GB also display consistent DR, suggesting reliable model behavior across different categories. Like NPV and PPV, NB produces the worst performance in terms of DR.

Like DR, DP is balanced across SVM, GB, MLR and RF. Finally, BA, which combines sensitivity and specificity, is highest in RF and GB at 0.992, followed closely by SVM at 0.986, showcasing overall strong model performance [**Table 5**].

**[Insert Table 5 Here]**

**Table 5:** Evaluation of performance measurements from various ML algorithms for Stress

Many variables are used to predict the depression, anxiety and stress severity of a person. Every of them is not equally important at all. Among the variables, some are found to be most responsible for depression, anxiety and stress. These variables are shown in **Figure 1**.

**[Insert Figure 1 Here]**

**Figure 1:** Important variables for depression

The feelings of having nothing to look forward to, of low self-worth, difficulty in finding motivation, absence of positive emotions, lack of enthusiasm, feeling downhearted, and experiencing a sense of meaninglessness in life are significant predictors for depression, anxiety and stress when using the DASS-21 assessment tool **[Figure 1]**.

**Discussion**

University students face a heightened susceptibility to depression and other mental health challenges, leading to potential detrimental effects on both their well-being and academic performance [27-29]. This alarming scenario has prompted an increase in public concern for the mental health of university students, with depression, anxiety and stress emerging as a prominent focus in public health discussions. In response to this growing concern, this study research endeavors to employ various statistical and ML models for the classification and prediction of depression, anxiety and stress among university students in Bangladesh. The application of a multi-class classification technique was employed to assess five distinct severity levels, utilizing the DASS-21 scale.

The findings of this study revealed that the prevalence of moderate to extremely severe levels was 57% for depression, 77.6% for anxiety, and 44% for stress. A study conducted at a university in Bangladesh discovered that the prevalence rates for moderate to extremely severe levels of depression, anxiety, and stress were 52.2%, 58.1%, and 24.9%, respectively [30]. Another study found that the prevalence rates for moderate to extremely severe levels of depression and anxiety were 69.5% and 61%, respectively [31].

In the present investigation, various factors were identified as significant predictors of depression, anxiety and stress among university students. The feelings of having nothing to look forward to, of low self-worth, difficulty in finding motivation, absence of positive emotions, lack of enthusiasm, feeling downhearted, and experiencing a sense of meaninglessness in life are significant predictors for depression. These factors have been identified as significant in accordance with findings from various previous studies [32-33].

Based on the evaluation of various ML models for predicting depression, anxiety, and stress, this study found SVM model consistently demonstrates superior performance across all three conditions. For depression, SVM leads in accuracy (0.8289), sensitivity, NPV, specificity, and balanced accuracy. For anxiety, KNN slightly outperforms SVM in accuracy and balanced accuracy, but SVM still shows strong performance in PPV and sensitivity. In predicting stress, SVM again leads with the highest accuracy and balanced accuracy and exhibits strong PPV and sensitivity. Overall, while KNN performs exceptionally well for anxiety, SVM proves to be the most robust and consistent model across depression, anxiety, and stress, making it the best overall choice for predicting mental health conditions among university students. Other models, Random Forest (RF), and Gradient Boosting also demonstrate commendable performance in predicting depression, anxiety, and stress. However, the MLR model, while showing reasonable performance, often lags in comparison to SVM and KNN in terms of accuracy and balanced accuracy. NB, on the other hand, consistently performs the worst across all conditions, with particularly low accuracy scores, poor sensitivity, and low balanced accuracy, indicating its limited reliability in predicting mental health conditions among university students. Similar findings were reported in a study conducted in India, where RF, DT, SVM, Ada, CatBoost, and Extreme Gradient Boosting (EGB) were utilized to classify the levels of depression, anxiety, and stress into normal, mild, moderate, severe, and extremely severe categories. In their experiments on the dataset, the SVM achieved superior performance for predicting depression, anxiety, and stress, respectively [34]. In a similar study utilizing six supervised ML techniques (KNN, RF, LR, DT, SVM, and NB) on a Kaggle dataset, the models were compared for accuracy and performance in predicting depression. Through 10-fold cross-validation, the study found that SVM outperformed other methods, achieving an accuracy of 83.32% [14]. In a comparable study utilizing GB classifier, RF, LR, and SVM for depression prediction, both SVM and LR proved effective. However, LR demonstrated slightly higher accuracy and quicker execution time compared to SVM [35].

**Recommendation**

Based on the findings of this study, it is recommended that universities implement mental health screening and intervention programs utilizing advanced machine learning models, particularly the SVM, which has demonstrated superior performance in predicting depression, anxiety, and stress. Given the high prevalence rates of these conditions among students, integrating SVM into mental health initiatives can enhance early detection and management. Additionally, focusing on key predictors such as feelings of having nothing to look forward to, low self-worth, difficulty in finding motivation, absence of positive emotions, lack of enthusiasm, feeling downhearted, and a sense of meaninglessness is crucial. Addressing these factors through targeted interventions can significantly improve the effectiveness of mental health support. While SVM proves to be the most robust model, other models like KNN, RF, and GB also show commendable performance and could serve as valuable complementary tools. By incorporating these models and focusing on the identified predictors, universities can develop a comprehensive mental health strategy that ensures timely support and effective intervention for students.

**Conclusion**

The study highlights the superior performance of the SVM model in predicting depression, anxiety, and stress among university students, outperforming other models in accuracy, sensitivity, and balanced accuracy. While KNN excels in predicting anxiety and models like RF and GB show commendable results, SVM remains the most robust and consistent across all conditions. NB consistently underperforms, demonstrating limited effectiveness in accurately predicting mental health conditions. These findings recommend prioritizing SVM for mental health prediction applications, while also considering complementary models to enhance overall prediction accuracy and reliability.

**Declarations**

**Ethics approval and consent to participate**

This study has received ethical approval from the Institutional Review Board (IRB) of BSMRSTU. This study utilized data obtained from self-reported questionnaires administered to university students. Informed consent was obtained from all participants prior to data collection, in accordance with guidelines provided by the BSMRSTU University Research Cell.

**Consent for publication**

Not Applicable**.**

**Availability of data and materials**

The data is available on request from the corresponding author.

**Competing interests**

The authors declare no conflict of interest.

**Funding**

BSMRSTU Research Cell, received funding for this research work from the People's Republic of Bangladesh (project code: BSMRSTU-RC-24-025).

**Authors' contributions**

MKH, PBA, MSH, AH, NS cleaned, processed the dataset, and prepared the draft of the manuscript. All the authors reviewed the manuscript several times. All the authors have equal contribution to conduct the study.

**Acknowledgements**

We express gratitude to the BSMRSTU Research Cell for providing funding for data collection purposes for this study. We extend our thanks to all the respondents who participated in the study.

**References**

1. Organization WH. Investing in mental health [Internet]. World Health Organization; 2003 [cited 2024 Mar 31]. Available from: <https://iris.who.int/handle/10665/42823>
2. American psychiatric association. Annual-Report-2013 [Internet]. [cited 2024 Mar 31]. Available from: https://www.psychiatry.org/, 2013.
3. Seligman ME, Walker EF, Rosenhan DL. Abnormal psychology. (No Title). 2001.
4. Craske MG, Stein MB, Eley TC, et al. Anxiety disorders. *Nat Rev Dis Prim*. 2017; **3**(1):17024.
5. Kendler KS. Major depression and generalised anxiety disorder. *Focus*. 2004; **2**(3): 416-425.
6. Lazarus RS, Folkman S. *Stress, Appraisal, and Coping*. Springer Publishing; 1984: 1-60.
7. Vijaya VS, Karunakaran P. A study on stress among intermediate students. *Innov Thoughts Int Res J*. 2013; **1**: 21-25.
8. Faruk MdO, Ching U, Chowdhury KUA. Mental health and well-being of indigenous people during the COVID-19 pandemic in Bangladesh. Heliyon. 2021 Jul;7(7):e07582.
9. Bayram N, Bilgel N. The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. Social psychiatry and psychiatric epidemiology. 2008 Aug;43:667-72.
10. World Health Organization (2020b) Mental Health: Current mental health situation in Bangladesh. Available at: <http://www.searo.who.int/bangladesh/mental-health/en/>
11. Alim SA, Kibria SM, Uddin MZ, Nessa M, Wahab MA. Translation of DASS 21 into Bangla and validation among medical students. Bangladesh Journal of Psychiatry. 2014;28(2):67-70.
12. Al Mamun MA, Griffiths MD. The association between Facebook addiction and depression: A pilot survey study among Bangladeshi students. Psychiatry research. 2019 Jan 1;271:628-33.
13. Sahlan F, Hamidi F, Misrat MZ, Adli MH, Wani S, Gulzar Y. Prediction of Mental Health Among University Students. Int J Perceptive Cogn Comput. 2021 Jul 16;7(1):85–91.
14. Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. Procedia Computer Science. 2020 Jan 1;167:1258-67.
15. Sau A, Bhakta I. Screening of anxiety and depression among seafarers using machine learning technology. Informatics in Medicine Unlocked. 2019 Jan 1;16:100228.
16. Reece AG, Reagan AJ, Lix KL, Dodds PS, Danforth CM, Langer EJ. Forecasting the onset and course of mental illness with Twitter data. Scientific reports. 2017 Oct 11;7(1):13006.
17. Braithwaite SR, Giraud-Carrier C, West J, Barnes MD, Hanson CL. Validating machine learning algorithms for Twitter data against established measures of suicidality. JMIR mental health. 2016 May 16;3(2):e4822.
18. Du J, Zhang Y, Luo J, Jia Y, Wei Q, Tao C, Xu H. Extracting psychiatric stressors for suicide from social media using deep learning. BMC medical informatics and decision making. 2018 Jul;18:77-87.
19. Alim, S.M. A. H.M., Kibria, S. M. E., Islam,M. J., Uddin, M. Z., Nessa,M., Wahab, M. A., & Islam, M. M. (2017a). Translation of DASS 21 into Bangla and validation among medical students. Bangladesh Journal of Psychiatry, 28(2), 67–70.
20. Alim, S. M. A. H. M., Rabbani, M. G., Karim, E., Mullick, M. S. I., Al Mamun, A., & Khan, M. Z. R. (2017b). Assessment of depression, anxiety and stress among first year MBBS students of a public medical college, Bangladesh. Bangladesh Journal of Psychiatry, 29(1), 23–29.
21. Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 1995;33(3):335-343.
22. Breiman L. Random forests. Machine learning. 2001 Oct;45:5-32.
23. Rish I. An empirical study of the naive Bayes classifier. InIJCAI 2001 workshop on empirical methods in artificial intelligence 2001 Aug 4 (Vol. 3, No. 22, pp. 41-46).
24. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural processing letters. 1999 Jun;9:293-300.
25. Ali N, Neagu D, Trundle P. Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets. SN Applied Sciences. 2019 Dec;1:1-5.
26. Rahman S, Irfan M, Raza M, Moyeezullah Ghori K, Yaqoob S, Awais M. Performance analysis of boosting classifiers in recognizing activities of daily living. International journal of environmental research and public health. 2020 Feb;17(3):1082.
27. Manthorpe NS Jill. Responding to students’ mental health needs: Impermeable systems and diverse users. J Ment Health. 2001 Jan 1;10(1):41–52.
28. Changes in counseling center client problems across 13 years. [Internet]. [cited 2024 Mar 31]. Available from: <https://psycnet.apa.org/record/2003-04125-010>
29. Eisenberg D, Gollust SE, Golberstein E, Hefner JL. Prevalence and Correlates of Depression, Anxiety, and Suicidality Among University Students. Am J Orthopsychiatry. 2007;77(4):534–42.
30. Mamun, M.A., Hossain, M.S. & Griffiths, M.D. Mental Health Problems and Associated Predictors Among Bangladeshi Students. *Int J Ment Health Addiction* **20**, 657–671 (2022). <https://doi.org/10.1007/s11469-019-00144-8>
31. Islam, S., Akter, R., Sikder, T. *et al.* Prevalence and Factors Associated with Depression and Anxiety Among First-Year University Students in Bangladesh: A Cross-Sectional Study. *Int J Ment Health Addiction* **20**, 1289–1302 (2022). <https://doi.org/10.1007/s11469-020-00242-y>
32. Kumar P, Garg S, Garg A. Assessment of anxiety, depression and stress using machine learning models. Procedia Computer Science. 2020 Jan 1;171:1989-98.
33. Budiyanto S, Sihombing HC, IM FR. Depression and anxiety detection through the Closed-Loop method using DASS-21. TELKOMNIKA (Telecommunication Computing Electronics and Control). 2019 Aug 1;17(4):2087-97.
34. Singh S, Gupta H, Singh P, Agrawal AP. Comparative analysis of machine learning models to predict depression, anxiety and stress. In2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART) 2022 Dec 16 (pp. 1199-1203). IEEE.
35. Obagbuwa IC, Danster S, Chibaya OC. Supervised machine learning models for depression sentiment analysis. Frontiers in Artificial Intelligence. 2023 Jul 19;6:1230649.

**List of Tables and Figures**

**Table 1:** Confusion Matrices from different ML Algorithms for Depression, Anxiety, and Stress.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Depression** | | | | | | **Anxiety** | | | | | | **Stress** | | | | | |
|
| **Multinomial Logistic Regression** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 15 | 0 | 1 | 0 | 0 | 1 | 5 | 1 | 1 | 0 | 0 | 1 | 23 | 3 | 1 | 1 | 0 |
| 2 | 3 | 10 | 4 | 0 | 0 | 2 | 0 | 3 | 2 | 1 | 0 | 2 | 4 | 6 | 2 | 1 | 0 |
| 3 | 1 | 4 | 14 | 2 | 2 | 3 | 6 | 1 | 12 | 0 | 1 | 3 | 1 | 2 | 7 | 5 | 2 |
| 4 | 0 | 0 | 2 | 6 | 2 | 4 | 0 | 0 | 4 | 8 | 3 | 4 | 0 | 1 | 3 | 5 | 0 |
| 5 | 0 | 0 | 2 | 0 | 8 | 5 | 0 | 0 | 1 | 2 | 23 | 5 | 0 | 2 | 0 | 4 | 2 |
|  | | | | | | |  | | | | | |  | | | | | |
| **Random Forest** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 19 | 6 | 3 | 0 | 0 | 1 | 6 | 1 | 2 | 0 | 0 | 1 | 26 | 6 | 0 | 0 | 0 |
| 2 | 0 | 4 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 0 | 2 | 2 | 5 | 1 | 0 | 0 |
| 3 | 0 | 4 | 20 | 6 | 3 | 3 | 2 | 1 | 16 | 2 | 2 | 3 | 0 | 1 | 9 | 1 | 0 |
| 4 | 0 | 0 | 0 | 2 | 0 | 4 | 0 | 0 | 0 | 6 | 2 | 4 | 0 | 2 | 3 | 14 | 0 |
| 5 | 0 | 0 | 0 | 0 | 9 | 5 | 1 | 1 | 2 | 3 | 23 | 5 | 0 | 0 | 0 | 1 | 4 |
|  | | | | | | |  | | | | | |  | | | | | |
| **SVM** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 16 | 1 | 1 | 0 | 0 | 1 | 9 | 2 | 0 | 0 | 0 | 1 | 27 | 1 | 0 | 0 | 0 |
| 2 | 3 | 12 | 2 | 0 | 0 | 2 | 1 | 2 | 3 | 1 | 0 | 2 | 1 | 12 | 4 | 0 | 0 |
| 3 | 0 | 1 | 17 | 0 | 1 | 3 | 1 | 1 | 15 | 0 | 1 | 3 | 0 | 1 | 8 | 2 | 0 |
| 4 | 0 | 0 | 3 | 8 | 1 | 4 | 0 | 0 | 2 | 8 | 1 | 4 | 0 | 0 | 1 | 12 | 0 |
| 5 | 0 | 0 | 0 | 0 | 10 | 5 | 0 | 0 | 0 | 2 | 25 | 5 | 0 | 0 | 0 | 2 | 4 |
|  | | | | | | |  | | | | | |  | | | | | |
| **KNN** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 19 | 5 | 0 | 0 | 0 | 1 | 11 | 2 | 1 | 0 | 0 | 1 | 28 | 10 | 1 | 0 | 0 |
| 2 | 0 | 5 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 4 | 3 | 0 | 0 |
| 3 | 0 | 4 | 22 | 6 | 0 | 3 | 0 | 3 | 18 | 5 | 0 | 3 | 0 | 0 | 8 | 1 | 0 |
| 4 | 0 | 0 | 0 | 1 | 1 | 4 | 0 | 0 | 1 | 4 | 0 | 4 | 0 | 0 | 1 | 15 | 4 |
| 5 | 0 | 0 | 0 | 1 | 11 | 5 | 0 | 0 | 0 | 2 | 27 | 5 | 0 | 0 | 0 | 0 | 0 |
|  | | | | | | |  | | | | | |  | | | | | |
| **Naïve Bayes** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 19 | 8 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 28 | 14 | 13 | 16 | 4 |
| 2 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 6 | 21 | 8 | 8 | 3 | 7 | 3 | 15 | 3 | 2 | 3 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 4 | 5 | 4 | 2 | 5 | 8 | 25 | 5 | 0 | 0 | 0 | 0 | 0 |
|  | | | | | | |  | | | | | |  | | | | | |
| **Gradient Boosting** |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |  | 1 | 2 | 3 | 4 | 5 |
| 1 | 17 | 3 | 1 | 0 | 0 | 1 | 7 | 1 | 1 | 0 | 0 | 1 | 28 | 6 | 0 | 1 | 0 |
| 2 | 0 | 6 | 3 | 0 | 0 | 2 | 1 | 2 | 3 | 0 | 0 | 2 | 0 | 8 | 5 | 0 | 0 |
| 3 | 2 | 5 | 15 | 3 | 1 | 3 | 3 | 1 | 14 | 2 | 1 | 3 | 0 | 0 | 6 | 0 | 0 |
| 4 | 0 | 0 | 4 | 5 | 2 | 4 | 0 | 1 | 2 | 7 | 3 | 4 | 0 | 0 | 2 | 14 | 0 |
| 5 | 0 | 0 | 0 | 0 | 9 | 5 | 0 | 0 | 0 | 2 | 23 | 5 | 0 | 0 | 0 | 1 | 4 |

**Table 2:** The Distribution of Depression, Anxiety, and Stress among University Students

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Level of depression** | | **Level of anxiety** | | **Level of stress** | |
| **Frequency** | **Percent** | **Frequency** | **Percent** | **Frequency** | **Percent** |
| **Normal** | 95 | 24.7 | 58 | 15.1 | 144 | 37.5 |
| **Mild** | 70 | 18.2 | 28 | 7.3 | 71 | 18.5 |
| **Moderate** | 118 | 30.7 | 103 | 26.8 | 67 | 17.4 |
| **Severe** | 40 | 10.4 | 59 | 15.4 | 81 | 21.1 |
| **Extremely severe** | 61 | 15.9 | 136 | 35.4 | 21 | 5.5 |
| **Total** | 384 | 100 | 384 | 100 | 384 | 100 |

**Table 3:** Evaluation of performance measurements from various ML algorithms for Depression

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Accuracy** | **PPV** | | **Sensitivity** | | **NPV** | | **Specificity** | | **Prevalence** | | **Detection Rate** | | **Detection Prevalence** | | **Balanced Accuracy** | |
|  |
| **Multinomial Logistic Regression** | 0.697 | 1 | 0.938 | 1 | 0.790 | 1 | 0.933 | 1 | 0.983 | 1 | 0.250 | 1 | 0.197 | 1 | 0.211 | 1 | 0.886 |  |
| 2 | 0.588 | 2 | 0.714 | 2 | 0.932 | 2 | 0.887 | 2 | 0.184 | 2 | 0.132 | 2 | 0.224 | 2 | 0.801 |  |
| 3 | 0.609 | 3 | 0.609 | 3 | 0.830 | 3 | 0.830 | 3 | 0.303 | 3 | 0.184 | 3 | 0.303 | 3 | 0.719 |  |
| 4 | 0.600 | 4 | 0.750 | 4 | 0.970 | 4 | 0.941 | 4 | 0.105 | 4 | 0.079 | 4 | 0.132 | 4 | 0.846 |  |
| 5 | 0.800 | 5 | 0.667 | 5 | 0.939 | 5 | 0.969 | 5 | 0.158 | 5 | 0.105 | 5 | 0.132 | 5 | 0.818 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Random Forest** | 0.711 | 1 | 0.679 | 1 | 1.000 | 1 | 1.000 | 1 | 0.842 | 1 | 0.250 | 1 | 0.250 | 1 | 0.368 | 1 | 0.921 |  |
| 2 | 1.000 | 2 | 0.286 | 2 | 0.861 | 2 | 1.000 | 2 | 0.184 | 2 | 0.053 | 2 | 0.053 | 2 | 0.643 |  |
| 3 | 0.755 | 3 | 0.870 | 3 | 0.930 | 3 | 0.755 | 3 | 0.303 | 3 | 0.263 | 3 | 0.434 | 3 | 0.812 |  |
| 4 | 1.000 | 4 | 0.250 | 4 | 0.919 | 4 | 1.000 | 4 | 0.105 | 4 | 0.026 | 4 | 0.026 | 4 | 0.625 |  |
| 5 | 1.000 | 5 | 0.750 | 5 | 0.955 | 5 | 1.000 | 5 | 0.158 | 5 | 0.118 | 5 | 0.118 | 5 | 0.875 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **SVM** | 0.829 | 1 | 0.889 | 1 | 0.842 | 1 | 0.948 | 1 | 0.964 | 1 | 0.250 | 1 | 0.211 | 1 | 0.237 | 1 | 0.904 |  |
| 2 | 0.706 | 2 | 0.857 | 2 | 0.966 | 2 | 0.919 | 2 | 0.184 | 2 | 0.158 | 2 | 0.224 | 2 | 0.888 |  |
| 3 | 0.895 | 3 | 0.739 | 3 | 0.895 | 3 | 0.962 | 3 | 0.303 | 3 | 0.224 | 3 | 0.250 | 3 | 0.851 |  |
| 4 | 0.667 | 4 | 1.000 | 4 | 1.000 | 4 | 0.941 | 4 | 0.105 | 4 | 0.105 | 4 | 0.158 | 4 | 0.971 |  |
| 5 | 1.000 | 5 | 0.833 | 5 | 0.970 | 5 | 1.000 | 5 | 0.158 | 5 | 0.132 | 5 | 0.132 | 5 | 0.917 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **KNN** | 0.763 | 1 | 0.792 | 1 | 1.000 | 1 | 1.000 | 1 | 0.912 | 1 | 0.250 | 1 | 0.250 | 1 | 0.316 | 1 | 0.956 |  |
| 2 | 0.833 | 2 | 0.357 | 2 | 0.871 | 2 | 0.984 | 2 | 0.184 | 2 | 0.066 | 2 | 0.079 | 2 | 0.671 |  |
| 3 | 0.688 | 3 | 0.957 | 3 | 0.977 | 3 | 0.811 | 3 | 0.303 | 3 | 0.290 | 3 | 0.421 | 3 | 0.884 |  |
| 4 | 0.500 | 4 | 0.125 | 4 | 0.905 | 4 | 0.985 | 4 | 0.105 | 4 | 0.013 | 4 | 0.026 | 4 | 0.555 |  |
| 5 | 0.917 | 5 | 0.917 | 5 | 0.984 | 5 | 0.984 | 5 | 0.158 | 5 | 0.145 | 5 | 0.158 | 5 | 0.951 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Naïve Bayes** | 0.579 | 1 | 0.655 | 1 | 1.000 | 1 | 1.000 | 1 | 0.825 | 1 | 0.250 | 1 | 0.250 | 1 | 0.382 | 1 | 0.912 |  |
| 2 | NAN | 2 | 0.000 | 2 | 0.816 | 2 | 1.000 | 2 | 0.184 | 2 | 0.000 | 2 | 0.000 | 2 | 0.500 |  |
| 3 | 0.488 | 3 | 0.913 | 3 | 0.939 | 3 | 0.585 | 3 | 0.303 | 3 | 0.276 | 3 | 0.566 | 3 | 0.749 |  |
| 4 | NAN | 4 | 0.000 | 4 | 0.895 | 4 | 1.000 | 4 | 0.105 | 4 | 0.000 | 4 | 0.000 | 4 | 0.500 |  |
| 5 | 1.000 | 5 | 0.333 | 5 | 0.889 | 5 | 1.000 | 5 | 0.158 | 5 | 0.053 | 5 | 0.053 | 5 | 0.667 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Gradient Boosting** | 0.684 | 1 | 0.810 | 1 | 0.895 | 1 | 0.964 | 1 | 0.930 | 1 | 0.250 | 1 | 0.224 | 1 | 0.276 | 1 | 0.912 |  |
| 2 | 0.667 | 2 | 0.429 | 2 | 0.881 | 2 | 0.952 | 2 | 0.184 | 2 | 0.079 | 2 | 0.118 | 2 | 0.690 |  |
| 3 | 0.577 | 3 | 0.652 | 3 | 0.840 | 3 | 0.793 | 3 | 0.303 | 3 | 0.197 | 3 | 0.342 | 3 | 0.722 |  |
| 4 | 0.455 | 4 | 0.625 | 4 | 0.954 | 4 | 0.912 | 4 | 0.105 | 4 | 0.066 | 4 | 0.145 | 4 | 0.768 |  |
| 5 | 1.000 | 5 | 0.750 | 5 | 0.955 | 5 | 1.000 | 5 | 0.158 | 5 | 0.118 | 5 | 0.118 | 5 | 0.875 |  |

**Table 4:** Evaluation of performance measurements from various ML algorithms for Anxiety

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **PPV** | | **Sensitivity** | | **NPV** | | **Specificity** | | **Prevalence** | | **Detection Rate** | | **Detection Prevalence** | | **Balanced Accuracy** | |
|  |
| **Multinomial Logistic Regression** | 0.689 | 1 | 0.714 | 1 | 0.455 | 1 | 0.911 | 1 | 0.968 | 1 | 0.149 | 1 | 0.068 | 1 | 0.095 | 1 | 0.711 |  |
| 2 | 0.500 | 2 | 0.600 | 2 | 0.971 | 2 | 0.957 | 2 | 0.068 | 2 | 0.041 | 2 | 0.081 | 2 | 0.778 |  |
| 3 | 0.600 | 3 | 0.600 | 3 | 0.852 | 3 | 0.852 | 3 | 0.270 | 3 | 0.162 | 3 | 0.270 | 3 | 0.726 |  |
| 4 | 0.533 | 4 | 0.727 | 4 | 0.949 | 4 | 0.949 | 4 | 0.149 | 4 | 0.108 | 4 | 0.203 | 4 | 0.808 |  |
| 5 | 0.885 | 5 | 0.852 | 5 | 0.917 | 5 | 0.917 | 5 | 0.365 | 5 | 0.311 | 5 | 0.351 | 5 | 0.894 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Random Forest** | 0.716 | 1 | 0.667 | 1 | 0.546 | 1 | 0.923 | 1 | 0.952 | 1 | 0.149 | 1 | 0.081 | 1 | 0.122 | 1 | 0.749 |  |
| 2 | 0.500 | 2 | 0.400 | 2 | 0.957 | 2 | 0.971 | 2 | 0.068 | 2 | 0.027 | 2 | 0.054 | 2 | 0.686 |  |
| 3 | 0.696 | 3 | 0.800 | 3 | 0.922 | 3 | 0.870 | 3 | 0.270 | 3 | 0.216 | 3 | 0.311 | 3 | 0.835 |  |
| 4 | 0.750 | 4 | 0.546 | 4 | 0.924 | 4 | 0.968 | 4 | 0.149 | 4 | 0.081 | 4 | 0.108 | 4 | 0.757 |  |
| 5 | 0.767 | 5 | 0.852 | 5 | 0.909 | 5 | 0.851 | 5 | 0.365 | 5 | 0.311 | 5 | 0.405 | 5 | 0.852 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **SVM** | 0.797 | 1 | 0.818 | 1 | 0.818 | 1 | 0.968 | 1 | 0.968 | 1 | 0.149 | 1 | 0.122 | 1 | 0.149 | 1 | 0.893 |  |
| 2 | 0.286 | 2 | 0.400 | 2 | 0.955 | 2 | 0.928 | 2 | 0.068 | 2 | 0.027 | 2 | 0.095 | 2 | 0.664 |  |
| 3 | 0.833 | 3 | 0.750 | 3 | 0.911 | 3 | 0.944 | 3 | 0.270 | 3 | 0.203 | 3 | 0.243 | 3 | 0.847 |  |
| 4 | 0.727 | 4 | 0.727 | 4 | 0.952 | 4 | 0.952 | 4 | 0.149 | 4 | 0.108 | 4 | 0.149 | 4 | 0.840 |  |
| 5 | 0.926 | 5 | 0.926 | 5 | 0.957 | 5 | 0.957 | 5 | 0.365 | 5 | 0.338 | 5 | 0.365 | 5 | 0.942 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **KNN** | 0.811 | 1 | 0.786 | 1 | 1.000 | 1 | 1.000 | 1 | 0.952 | 1 | 0.149 | 1 | 0.149 | 1 | 0.189 | 1 | 0.976 |  |
| 2 | NAN | 2 | 0.000 | 2 | 0.932 | 2 | 1.000 | 2 | 0.068 | 2 | 0.000 | 2 | 0.000 | 2 | 0.500 |  |
| 3 | 0.692 | 3 | 0.900 | 3 | 0.958 | 3 | 0.852 | 3 | 0.270 | 3 | 0.243 | 3 | 0.351 | 3 | 0.876 |  |
| 4 | 0.800 | 4 | 0.364 | 4 | 0.896 | 4 | 0.984 | 4 | 0.149 | 4 | 0.054 | 4 | 0.068 | 4 | 0.674 |  |
| 5 | 0.931 | 5 | 1.000 | 5 | 1.000 | 5 | 0.957 | 5 | 0.365 | 5 | 0.365 | 5 | 0.392 | 5 | 0.979 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Naïve Bayes** | 0.541 | 1 | NAN | 1 | 0.000 | 1 | 0.851 | 1 | 1.000 | 1 | 0.149 | 1 | 0.000 | 1 | 0.000 | 1 | 0.500 |  |
| 2 | NAN | 2 | 0.000 | 2 | 0.932 | 2 | 1.000 | 2 | 0.068 | 2 | 0.000 | 2 | 0.000 | 2 | 0.500 |  |
| 3 | 0.500 | 3 | 0.750 | 3 | 0.887 | 3 | 0.722 | 3 | 0.270 | 3 | 0.203 | 3 | 0.405 | 3 | 0.736 |  |
| 4 | NAN | 4 | 0.000 | 4 | 0.851 | 4 | 1.000 | 4 | 0.149 | 4 | 0.000 | 4 | 0.000 | 4 | 0.500 |  |
| 5 | 0.568 | 5 | 0.926 | 5 | 0.933 | 5 | 0.596 | 5 | 0.365 | 5 | 0.338 | 5 | 0.595 | 5 | 0.761 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Gradient Boosting** | 0.716 | 1 | 0.778 | 1 | 0.636 | 1 | 0.939 | 1 | 0.968 | 1 | 0.149 | 1 | 0.096 | 1 | 0.122 | 1 | 0.802 |  |
| 2 | 0.333 | 2 | 0.400 | 2 | 0.956 | 2 | 0.942 | 2 | 0.068 | 2 | 0.027 | 2 | 0.081 | 2 | 0.671 |  |
| 3 | 0.667 | 3 | 0.700 | 3 | 0.887 | 3 | 0.870 | 3 | 0.270 | 3 | 0.189 | 3 | 0.284 | 3 | 0.785 |  |
| 4 | 0.539 | 4 | 0.636 | 4 | 0.934 | 4 | 0.905 | 4 | 0.149 | 4 | 0.095 | 4 | 0.176 | 4 | 0.771 |  |
| 5 | 0.920 | 5 | 0.852 | 5 | 0.918 | 5 | 0.957 | 5 | 0.365 | 5 | 0.311 | 5 | 0.338 | 5 | 0.905 |  |

**Table 5:** Evaluation of performance measurements from various ML algorithms for Stress

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **PPV** | | **Sensitivity** | | **NPV** | | **Specificity** | | **Prevalence** | | **Detection Rate** | | **Detection Prevalence** | | **Balanced Accuracy** | |
|  |
| **Multinomial Logistic Regression** | 0.573 | 1 | 0.821 | 1 | 0.821 | 1 | 0.894 | 1 | 0.894 | 1 | 0.373 | 1 | 0.307 | 1 | 0.373 | 1 | 0.858 |  |
| 2 | 0.462 | 2 | 0.429 | 2 | 0.871 | 2 | 0.885 | 2 | 0.187 | 2 | 0.080 | 2 | 0.173 | 2 | 0.657 |  |
| 3 | 0.412 | 3 | 0.539 | 3 | 0.897 | 3 | 0.839 | 3 | 0.173 | 3 | 0.093 | 3 | 0.227 | 3 | 0.689 |  |
| 4 | 0.556 | 4 | 0.313 | 4 | 0.833 | 4 | 0.932 | 4 | 0.213 | 4 | 0.067 | 4 | 0.120 | 4 | 0.622 |  |
| 5 | 0.250 | 5 | 0.500 | 5 | 0.970 | 5 | 0.916 | 5 | 0.053 | 5 | 0.027 | 5 | 0.107 | 5 | 0.708 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Random Forest** | 0.773 | 1 | 0.813 | 1 | 0.929 | 1 | 0.954 | 1 | 0.872 | 1 | 0.373 | 1 | 0.347 | 1 | 0.427 | 1 | 0.901 |  |
| 2 | 0.625 | 2 | 0.357 | 2 | 0.866 | 2 | 0.951 | 2 | 0.187 | 2 | 0.067 | 2 | 0.107 | 2 | 0.653 |  |
| 3 | 0.818 | 3 | 0.692 | 3 | 0.938 | 3 | 0.968 | 3 | 0.173 | 3 | 0.120 | 3 | 0.147 | 3 | 0.830 |  |
| 4 | 0.737 | 4 | 0.875 | 4 | 0.964 | 4 | 0.915 | 4 | 0.213 | 4 | 0.187 | 4 | 0.253 | 4 | 0.895 |  |
| 5 | 0.800 | 5 | 1.000 | 5 | 1.000 | 5 | 0.986 | 5 | 0.053 | 5 | 0.053 | 5 | 0.067 | 5 | 0.992 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **SVM** | 0.840 | 1 | 0.964 | 1 | 0.964 | 1 | 0.979 | 1 | 0.979 | 1 | 0.373 | 1 | 0.360 | 1 | 0.373 | 1 | 0.972 |  |
| 2 | 0.706 | 2 | 0.857 | 2 | 0.966 | 2 | 0.918 | 2 | 0.187 | 2 | 0.160 | 2 | 0.227 | 2 | 0.888 |  |
| 3 | 0.727 | 3 | 0.615 | 3 | 0.922 | 3 | 0.952 | 3 | 0.173 | 3 | 0.107 | 3 | 0.147 | 3 | 0.784 |  |
| 4 | 0.923 | 4 | 0.750 | 4 | 0.936 | 4 | 0.983 | 4 | 0.213 | 4 | 0.160 | 4 | 0.173 | 4 | 0.867 |  |
| 5 | 0.667 | 5 | 1.000 | 5 | 1.000 | 5 | 0.972 | 5 | 0.053 | 5 | 0.053 | 5 | 0.080 | 5 | 0.986 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **KNN** | 0.811 | 1 | 0.786 | 1 | 1.000 | 1 | 1.000 | 1 | 0.952 | 1 | 0.149 | 1 | 0.149 | 1 | 0.189 | 1 | 0.976 |  |
| 2 | NAN | 2 | 0.000 | 2 | 0.932 | 2 | 1.000 | 2 | 0.068 | 2 | 0.000 | 2 | 0.000 | 2 | 0.500 |  |
| 3 | 0.692 | 3 | 0.900 | 3 | 0.958 | 3 | 0.852 | 3 | 0.270 | 3 | 0.243 | 3 | 0.351 | 3 | 0.876 |  |
| 4 | 0.800 | 4 | 0.364 | 4 | 0.896 | 4 | 0.984 | 4 | 0.149 | 4 | 0.054 | 4 | 0.068 | 4 | 0.674 |  |
| 5 | 0.931 | 5 | 1.000 | 5 | 1.000 | 5 | 0.957 | 5 | 0.365 | 5 | 0.365 | 5 | 0.392 | 5 | 0.979 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Naïve Bayes** | 0.373 | 1 | 0.373 | 1 | 1.000 | 1 | NAN | 1 | 0.000 | 1 | 0.373 | 1 | 1.000 | 1 | 1.000 | 1 | 0.500 |  |
| 2 | NAN | 2 | 0.000 | 2 | 0.813 | 2 | 1.000 | 2 | 0.187 | 2 | 0.000 | 2 | 0.000 | 2 | 0.500 |  |
| 3 | NAN | 3 | 0.000 | 3 | 0.827 | 3 | 1.000 | 3 | 0.173 | 3 | 0.000 | 3 | 0.000 | 3 | 0.500 |  |
| 4 | NAN | 4 | 0.000 | 4 | 0.787 | 4 | 1.000 | 4 | 0.213 | 4 | 0.000 | 4 | 0.000 | 4 | 0.500 |  |
| 5 | NAN | 5 | 0.000 | 5 | 0.947 | 5 | 1.000 | 5 | 0.053 | 5 | 0.000 | 5 | 0.000 | 5 | .5.8 |  |
|  | | | | | | | | | | | | | | | | | |  |
| **Gradient Boosting** | 0.800 | 1 | 0.800 | 1 | 1.000 | 1 | 1.000 | 1 | 0.851 | 1 | 0.373 | 1 | 0.373 | 1 | 0.467 | 1 | 0.926 |  |
| 2 | 0.615 | 2 | 0.571 | 2 | 0.903 | 2 | 0.918 | 2 | 0.187 | 2 | 0.107 | 2 | 0.173 | 2 | 0.745 |  |
| 3 | 1.000 | 3 | 0.462 | 3 | 0.899 | 3 | 1.000 | 3 | 0.173 | 3 | 0.080 | 3 | 0.080 | 3 | 0.731 |  |
| 4 | 0.875 | 4 | 0.875 | 4 | 0.966 | 4 | 0.966 | 4 | 0.213 | 4 | 0.187 | 4 | 0.213 | 4 | 0.921 |  |
| 5 | 0.800 | 5 | 1.000 | 5 | 1.000 | 5 | 0.986 | 5 | 0.053 | 5 | 0.053 | 5 | 0.067 | 5 | 0.992 |  |

**Figure 1:** Important variables for depression

