International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 25(2), November, 2025, pp 33-46
DOI: https://doi.org/10.3329/ijss.v25i2.85733

Identifying the Determinants of Depression among Ever Married Women in Bangladesh: A Machine Learning Approach

Md. Salek Miah and Mohammad Ohid Ullah*

Department of Statistics, Shahjalal University of Science and Technology,
Sylhet-3114, Bangladesh
Emails: 2021134066@student.sust.edu; ohid-sta@sust.edu
*Correspondence should be addressed to Mohammad Ohid Ullah
(Email: ohid-sta@sust.edu)

[Received July 15, 2025; Accepted October 25, 2025]

Abstract

Mental health remains one of the most challenging public health concerns in the world, particularly in a developing country like Bangladesh. The objective of this study is to determine the key determinants of depression among ever-married women in Bangladesh. This study used 19,987 ever-married women aged 15-49 years in Bangladesh, from the 2022 Bangladesh Demographic and Health Survey (BDHS). The key features selection was selected by using Boruta algorithm. Total seven machine learning approaches, including Extreme Gradient Boosting (XGBoost), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbour's (KNN), and Neural Network (NN) were evaluated for predictive performance. All models were highly tuned with 5-fold cross-validation and tested on several complementary predictions including accuracy, precision, recall, F1 score, Matthews Correlation Coefficient (MCC), Cohen's Kappa, and Area Under the Receiver Operating Characteristic Curve (AUROC). The top-performing algorithm was observed XGBoost approach with the highest AUROC of 0.799, indicating strong discriminatory power between non-depressed and depressed women. Results from XGBoost approach revealed that anxiety, currently working, age group 45-49 years, internet user, menstruated last 6 weeks, pressure to pregnant, power of decision making etc. are most important determinants of the depression of married women in Bangladesh.

Keywords: Mental health, Depression, Boruta algorithm, Women, Machine learning.

AMS Classification: 62P25, 62P10, 68T09.

1. Introduction

Mental illness is a serious public health concern, especially in low- and middle-income countries (LMICs) like Bangladesh (WHO, 2022). In the last decade, mental illness has become more common and urgent, adding significantly to the worldwide burden of disease(Muzaffar et al., 2022). These figures identify a critical issue in public health that must be taken seriously in terms of research as well as policy action. Mental disorder is a leading and increasing global public health issue, contributing significantly to the overall disease burden, especially in low- and middle-income countries (LMICs) with a low level of mental health facilities (Xie et al., 2024). Social, economic, and cultural factors fuelling Asia include restricted access to effective care. South Asia is particularly faced with compelling mental health concerns due to the fast-paced socioeconomic changes and structural healthcare shortages (Mahmud et al., 2022). Bangladesh is an example, with increasing mental illness being sustained by limited infrastructure and stigma that necessitates imminent research as well as policy interventions to stem its increasing public health burden(Bangladesh Demographic and Health Survey 2022 Final Report, 2024). Previous studies have investigated mental health disorder by examining risk factors or association-based methods using logistic regression to investigate relationships between socio-demographic and environmental variables and mental health outcomes. Although such models are interpretable and offer initial insights, they fail in high-dimensional, nonlinear, and high-order interdependent data structures. Machine learning (ML), an extension of artificial intelligence, provides us with a rich set of tools for modelling this complexity(Rezapour & Elmshaeuser, 2022). Unlike conventional regression-based modelling, ML procedures are significantly better at uncovering nonlinear associations, interactions, and dependencies in large data sets without needing pre-specified assumptions. Ensemble algorithms such as Random Forest and Extreme Gradient Boosting (XGBoost) have been demonstrated to provide better predictive performance across diverse groups of people. Methods such as Recursive Feature Elimination (RFE) and SHAP (SHapley Additive exPlanations) values have been utilized to determine the top predictors in high-dimensional data(Xie et al., 2024; Xu et al., 2021). Additional methods including SMOTE (Synthetic Minority Oversampling Technique) have been effective in managing class imbalance, a pervasive problem in mental health data where the number of those affected is significantly lower than those not affected(Li et al., 2025a; Mitchell et al., 2023).Previous studies have compared different ML algorithms to examine the predictive performance and limitations of these algorithms in predicting mental health disorders(Huang et al., 2023b). Moreover, the aspects of digital exposures, social media, smartphone utilization, and access to digital technologies factors that play significant to mental health disorders remains unexplored in the Bangladeshi context(Huang et al., 2023a). Therefore, the mechanism of mental health disorder remains not clear. Therefore, the primary objective of the study is to identify the most important determinants of depression based on Patient Health Questionnaire -9 items (PHQ-9) in ever married women aged 15-49 in Bangladesh through machine learning feature selection methods. The secondary objective is to contrast the predictive performance of a variety of machine learning algorithms.

2. Theoretical framework and methodology of the study

2.1 Selection of the study area, sampling technique and data collection

This investigation used data from the Bangladesh Demographic and Health Survey (BDHS) 2022, a nationally representative, cross-sectional household survey employed under the authority of the National Institute of Population Research and Training (NIPORT) (https://www.niport.gov.bd), with technical assistance from ICF under the DHS program(https://dhsprogram.com). The BDHS applies a two-stage stratified cluster sampling technique, including all eight administrative divisions of Bangladesh include responses n=19,987 ever married women in Bangladesh. The Figure 1 represents the study methodology in details.

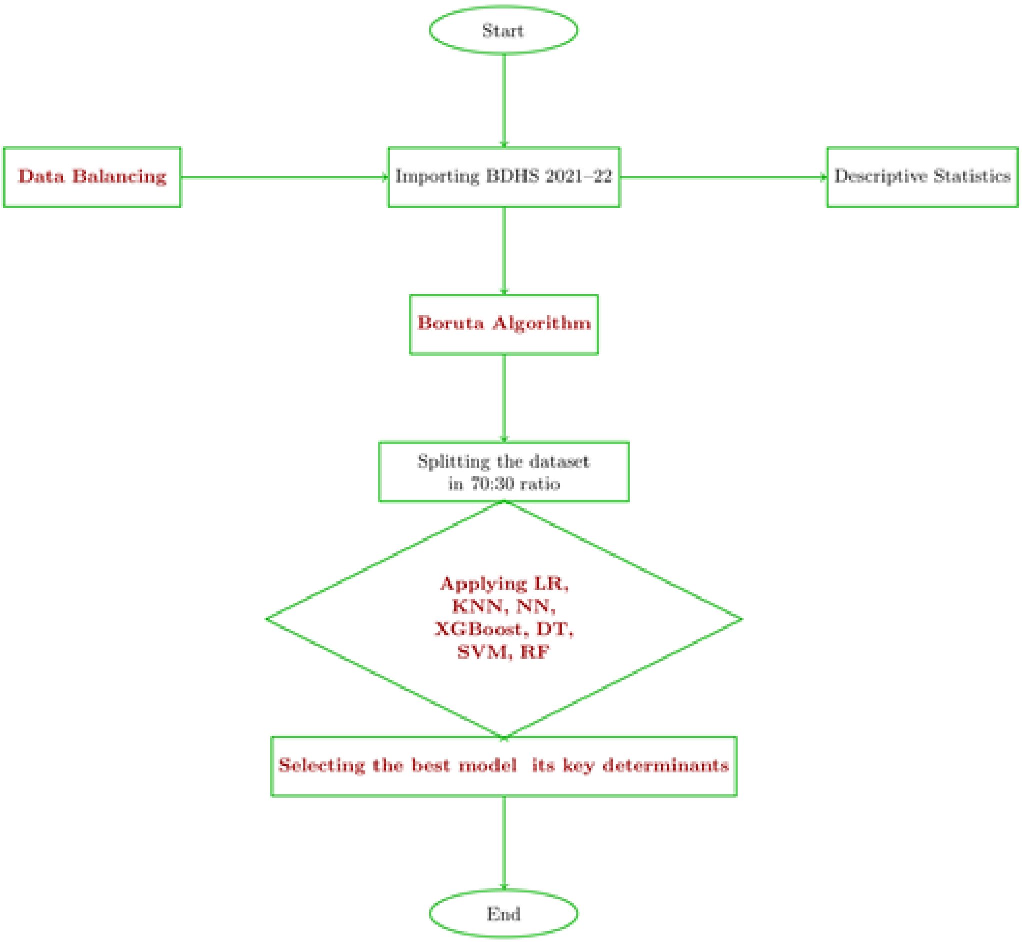

Figure 1: Flow Diagram of the Study Methodology

Target variables

The outcome variables of primary interest were created as dichotomous indicators using validated clinical cut-points. Depression was 10 or above on the PHQ-9 items, which represents yes symptoms of depression.

Predictor variables

Predictor variables were selected based on available evidence and theoretical frameworks and categorized as sociodemographic, behavioural, health-related, and social. Sociodemographic variables included, level of education (no education, primary, secondary, or above), and household wealth quintile (poorest to richest). Behavioural variables consisted of internet use, any use in the

12 previous months. Health-related variables comprised pregnancy at interview. Social factors were history of intimate partner violence (IPV), such as physical, emotional, or sexual IPV within the past 12 months. After applying Boruta algorithm for selecting feature 39 variables were selected from 35 variables. ICT ownership and ICT usage consists a composite variable including traditional technology such as ownership of mobile phone, TV, Radio, Refrigerator etc. Religion was dichotomized category as Muslim and Non- Muslim.

Missing Values and Data Quality

The data that was utilized in this research was properly screened to highlight non-existent or conflicting figures. Because the data were organized and purged in STATA 17 before analysis, there were no missing values related to any of the variables in the models. Thus, the analysis was carried out with all observations (N = 19,987), which led to full-case analysis and reliability of the model.

2.2 Theoretical framework

A number of machine learning algorithms were used to determine the determinants of depression that are important in ever-married women.

Decision Tree (DT): This is a hierarchical, rule-based classifier, which can support categorical as well as continuous variables. Even though interpretable, deep trees may overfit when they are too deep (Sarakshetrin et al., 2025).

Random Forest (RF): This is a collection of a large number of decision trees that enhances predictive accuracy and decreases over fitting especially with high dimensional data(Jiang et al., 2023).

Support Vector Machine (SVM): It builds an optimum hyperplane to separate classes which are competent in working with non-linear and high-dimensional data (Rahman et al., 2024).
Logistic Regression (LR): The LR is used to model binary outcome probabilities with coefficients that can be interpreted, which is why it is applicable to significant predictors (Samuel et al., 2024).

K-Nearest Neighbour's (KNN): The sample classification is determined by the majority group of the nearest neighbours; it is vulnerable to the scales of features and data preprocessing (Halder et al., 2024).

Extreme Gradient Boosting (XGBoost): A gradient-boosted tree-based algorithm that involves regularization, parallel computing; very accurate, and works well with structured data (Ostovar et al., 2025; Taskiran et al., 2025).

Artificial Neural Network (ANN): The complex nonlinear relations are represented in terms of interconnected layers but it needs to be tuned and consume a lot of computer power (Halder et al., 2024).

The STATA 17 was used to perform all the data management and R Studio 4.5.1 was used to execute and test the machine learning algorithms.

Performance Evaluation Metrics

Accuracy, precision (recall), F1-score, Cohen, and Cohen (Kappa) were used to measure model performance. Accuracy: This measures the percentage of instances that are correctly classified; it is however deceptive in imbalance data sets.

Accuracy: What percentage of all the instances is correctly identified. A high degree of accuracy can mask poor results on the minority group (Ostovar et al., 2025).

Sensitivity (Recall): The value of the fraction of true positives that is identified, which is important when it is expensive to miss the positive cases (Taskiran et al., 2025).

Precision: The ratio of predicted positives to the actual ones, which is the reliability of the model in the targeted interventions (Li et al., 2025b).

F1-Score: The harmonic average of recall and precision, which provides a more balanced measure of data with unbalanced data sets(Zhu et al., 2025).

Cohen Kappa: Measures the agreement between the predicted and observed classifications, including the correction of chance agreement, which is a stronger measure of assessment compared to accuracy in skewed data (Hakkal & Lahcen, 2024).

AUROC: (Assuming it providing in your original metrics) Accessibility of the model to differentiate between classes at all threshold levels (Wang et al., 2021).

Matthews Correlation Coefficient (MCC):
Reliable for imbalanced datasets, since it takes into account both the positive and negative classes equally (Islam et al., 2024).

3. Results and Discussion

3.1 Features Selection

Based on previous study the study considers the following predictors gives Tables 1(supplementary). The figure 2 highlights the important feature selection using Boruta algorithm. The variable of anxiety was by far the strongest predictor, with a mean importance value considerably greater than all the other variables. This finding reinforces the paramount position of psychological distress in influencing the outcome and singles out anxiety as a prime target for future modelling and interventions strategies. In addition to anxiety, several other variables exhibited strong and consistent significance. These included age group, decision-making control, employment visit history, household breakdown, and employment-derived purchasing power. An even larger set of confirmed variables comprised fields such as digital access (e.g., ICT ownership, internet usage), financial standing (wealth, employment funds), and household composition (children category, living situation), reflecting the multidimensionality of determinants to the outcome. Conversely, four variables now abstaining, religion_bin, (religion category) IPV experience, and IPV justification were dropped, indicating minimal or no extra predictive ability. Two additional characteristics number of unions and menstruation in the last 6 weeks were highlighted as tentative, suggesting unclear relevance that warrants potential investigation in sensitivity analyses or other model specifications. The output of the Boruta algorithm offers a data-driven, explicit basis for model specification. The inclusion of 29 confirmed predictors from 35 predictors optimizes the interpretability and generalisability of subsequent analyses while allowing the elimination of non-informative variables to reduce model complexity [**Error! Reference source not found.**].

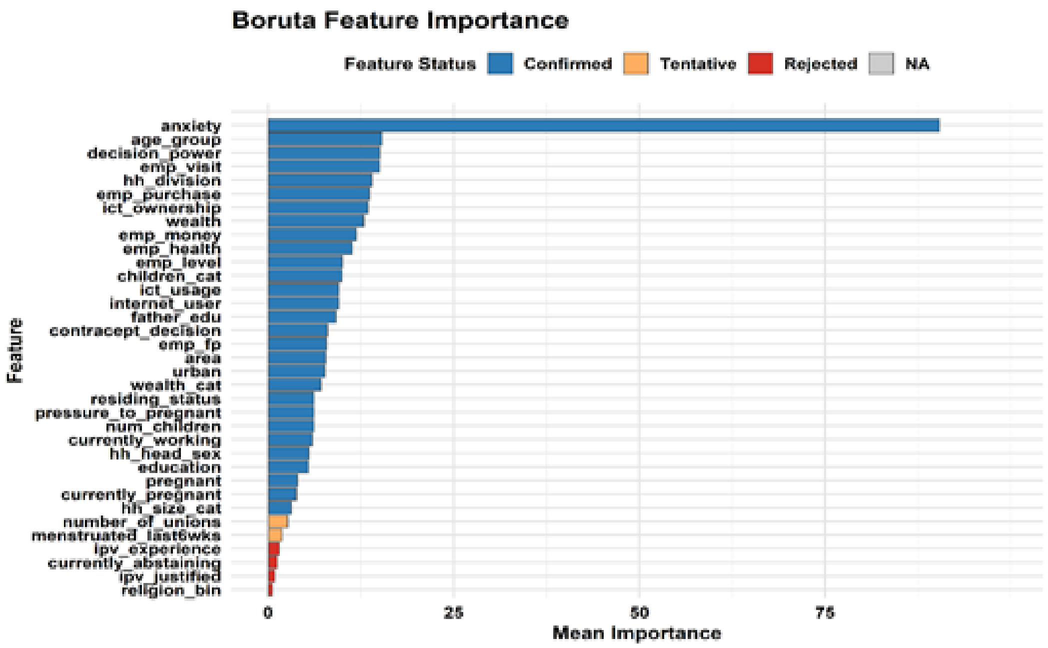

Figure 2: Boruta Feature Importance

3.2 Performance Metrics

Table 1: Performance Metrics of Machine Learning Models

Metric	Random Forest	Decision Tree	KNN	Logistic Regression	SVM	XGBoost	Neural Network
Accuracy	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Precision	0.58	0.55	0.06	0.55	0.61	0.56	0.05
Recall	0.34	0.39	0.70	0.33	0.28	0.34	0.65
F1	0.43	0.45	0.12	0.41	0.39	0.43	0.10
MCC	0.42	0.44	0.08	0.40	0.40	0.42	0.04
Kappa	0.41	0.43	0.03	0.39	0.37	0.40	0.01
AUROC	0.76	0.72	0.60	0.77	0.74	0.79	0.76

The lack of intimate partner violence and behavioural abstaining variables shows that these variables less potential. This subset of validate features yields high stake predictive modelling.

The [Figure 2**:** Boruta Feature Importance

3.2 Performance Metrics

Table 1] revealed the comparison of model performance of seven machine learning algorithms Random Forest, Decision Tree, k-Nearest Neighbour's (KNN), Logistic Regression, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Neural Network on different performance metrics determined the strongest classifier of the given prediction task. Among them, XGBoost gave the best result on the overall metric with the highest Area Under the Receiver Operating Characteristic Curve (AUROC) at 0.79, demonstrating outstanding discriminative ability among classes. Though uniform high accuracy ranged from 0.95 to 0.95 across models, this was not sufficient on its own considering the underlaying class imbalance. Precision and recall metrics revealed that there were considerable trade-offs; XGBoost maintained a better balance between precision at 0.56 and recall at 0.34 compared to other models. Its average recall means that it omitted some positive examples, while KNN and Neural Network model had greater recall but at the expense of their considerably low precision, resulting in a flood of false positives. The F1 score also reinforced XGBoost's excellently balanced performance with a comparable value of 0.43. Strong measures against class imbalance, such as the Matthews Correlation Coefficient (MCC) and Cohen's Kappa, were strongest in Decision Tree and Random Forest models, although XGBoost was still quite better. In general, the performances Favor XGBoost as the best model given its high discriminative power and well-balanced metric profile. Other higher-precision models but lower-recall models, like KNN and Neural Networks, are less preferable because they contain high false positive rates.

The **[**Error! Reference source not found.**]** represents precision and recall metrics further elucidate model strengths and weaknesses. XGBoost and Random Forest models offer a well-balanced trade-off, with precision values of approximately 0.57 and recall values near 0.35, leading to competitive F1 scores around 0.43. These metrics suggest that these models efficiently identify positive cases while controlling for false positives. In contrast, the Neural Network and KNN models achieve high recall (0.65 and 0.70), indicating better sensitivity, but at the expense of very low precision (0.05 and 0.06), resulting in false positives. This trade-off may limit their practical utility where precision is critical. Support Vector Machine and Logistic Regression demonstrate moderate precision but lower recall, implying they may underperform in identifying all positive instances.

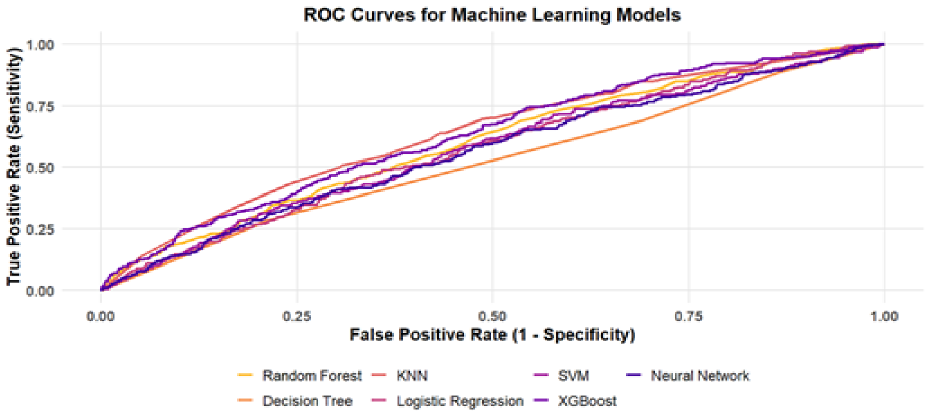

Figure 3: Machine Learning algorithm Evaluation

4.1. Evaluation

This study provides a comprehensive evaluation of seven machine learning algorithms including Random Forest, Decision Tree, K-Nearest Neighbour's (KNN), Logistic Regression, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Neural Network applied to a highly imbalanced binary classification task. The results indicate that XGBoost had the best performance, as reflected in its best Area Under the Receiver Operating Characteristic Curve (AUROC) of 0.79, confirming its excellent discriminative capability. While overall accuracy across all models was always high (0.95), such measures are insufficient in imbalanced datasets and are misleading in the absence of other sensitivity and specificity measures(Islam et al., 2024). Precision and recall must be traded off in predictive modelling(Islam et al., 2024), particularly in medical and epidemiological settings where false positive and false negative costs are highly divergent. XGBoost supported a reasonable trade-off with precision of 0.56 and recall of 0.34, showing its ability to detect positive instances without creating an intolerably high false alarm load. In comparison, algorithms such as KNN and Neural Networks, with recall (0.701 and 0.650 respectively) greater than XGBoost, had extremely poor precision (0.06 and 0.05 respectively), producing an intolerably high false-positive load(Huang et al., 2023b; Willinger et al., 2009). These effects lower clinical utility and can lead to resource wastage and patient depression. Logistic Regression and SVM were of moderate accuracy and low sensitivity, highlighting the inevitable trade-offs with model selection. The MCC and Cohen's Kappa statistics, which are superior performance measurement under class. imbalance, also validated the advantage of tree-based ensemble methods. Random Forest and Decision Tree classifiers yielded strong MCC and Kappa scores, but XGBoost performed respectably, attesting to its strength and reliability (Hua et al., 2025; Jiang et al., 2023). Confusion matrix evaluations confirmed these findings, showing that XGBoost was effective in minimizing false positives while maintaining an acceptable true positive detection rate, albeit possibly optimizable sensitivity. The ROC curve plot revealed XGBoost's superior ability to maintain high true positive rates at low false positive rates, an attribute of supreme relevance to clinical and public health contexts where false positive implications can be considerable (Tharwat, 2018; Zhenya & Zhang, 2021). The characteristic renders it more suitable for application in risk stratification, disease screening, and other precision medicine applications(Chicco & Jurman, 2020; Poudel et al., 2022). Taken collectively, these findings are in line with the current literature positioning XGBoost as an advanced classifier of complex, imbalanced data sets, given its gradient boosting design with the ability to proficiently handle data heterogeneity and nonlinearities. However, the modest recall observed means that further methodological enhancements are required. The use of advanced class-balancing methods such as Synthetic Minority Oversampling Technique (SMOTE), cost-sensitive learning, or ensemble techniques may be able to enhance sensitivity at a cost to precision. Our study finding [Table 2] represents that top determinants in line with (Boughorbel et al., 2017; Luque et al., 2019) . Limitations to this work include reliance on a single dataset, with the possibility of affecting generalizability, and exclusion of deep learning models more sophisticated than basic neural networks. Hybrid modelling architectures and external verification across different populations should be explored in future research to provide robustness validation. In summary, XGBoost seems to be the optimal algorithm for the task under conditions of class imbalance, achieving the best balance between specificity and sensitivity under conditions of class imbalance. Model selection with regard to context-dependent misclassification costs remains of the utmost importance. Combining class imbalance correction techniques offers an intriguing path towards further optimizing predictive accuracy and clinical utility. The study finds the top most 20 determinants are anxiety, currently working, age group 45-49 years, internet users, menstruated last 6 weeks, household wealth, women decision making power, education, ICT ownership, division, IPV, household size, pressure to pregnant etc.

Table 2: Top 20 Feature of Best Model

Feature	Our Model Top Rank	Prior Study Support	Description
Anxiety (yes)	Yes	Yes	Anxiety was the strongest predictor, consistent with global evidence of high comorbidity and predictive power of anxiety symptoms for depression (Halder et al., 2024; Taskiran et al., 2025).
Currently working (yes)	Yes	Yes	Young adults are more likely to follow mental health practices (Cuenca, 2022).
Age groups (20–24, 25–29, 30–34, 35–39, 45–49)	Yes	Yes	Risk rises with age, particularly 45–49 years; middle-aged individuals are more susceptible due to cumulative stressors and hormonal changes; internet exposure interacts with mental health (Li et al., 2025b).
Internet users (yes)	Yes	Yes	Higher internet use among young adults is linked to increased anxiety and depression, possibly due to social comparison or greater mental health awareness (Mujahid et al., 2024).
Menstruated last 6 weeks (yes)	Yes	Yes	Reproductive health markers, such as menstrual status, are associated with depression risk(Kaya et al., 2025) .
Pressure to become pregnant (yes)	Yes	Yes	Pressure regarding pregnancy contributes to mood disturbance and depression (Dunne et al., 2024).
Household size (4+ members)	Yes	Yes	Residential crowding is significantly associated with depression .(Jiang et al., 2023)
Intimate Partner Violence (yes)	Yes	Yes	Experiencing IPV doubles the risk of depression compared to non-exposed women (Jiang et al., 2023).
Household division (Rangpur, Khulna)	Yes	Yes	Geographic disparities influence depression prevalence (Naznin et al., 2025).
ICT ownership	Yes	Yes	Access to ICT promotes awareness and may reduce depressive risk(Sarakshetrin et al., 2025) .
Education (Secondary)	Yes	Yes	Secondary education acts as a protective factor against depression (Poudel et al., 2022).
Decision-making power	Yes	Yes	Women's decision-making power reflects a protective factor against depression (Sergeev et al., 2024).
Household wealth (Rich)	Yes	Yes	Wealthier households indicate lower depression risk (Fazal et al., 2024).

The difference in prevalence between women aged 45-49 and 60-64 years can be attributed to the fact that family and social changes later in life might have contributed to the higher prevalence rate of depression in women aged 45- 49 years. In the Bangladeshi society, there are numerous cases where the elder women suffer emotional trauma as their adult sons start living independently with their families, and become socially isolated and neglected. In addition, women can lose control or power over family decisions as children advance and acquire or run family wealth. Together with augmented health anxieties and even care giving needs, these changes can significantly elevate the psychological stress levels thus further endangering the group of people with depression tendencies at this age (Bangladesh Demographic and Health Survey 2022 Final Report, 2024; WHO, 2022).

4.2. Limitations

Household income was not specified as an explanatory variable (since the DHS 2022 does not present data of household income), but the variables pertaining to the household income were also represented in the dataset: household wealth and empowerment of women when making household financial decisions. The empowerment variable is the degree of the economic autonomy and portrays the socioeconomic status within the household. Due to the conceptual relationship between the wealth and empowerment indicators, the empowerment variable was retained to prevent the occurrence of multicollinearity but reflect the economic participation and decision-making power of women that are strongly associated with the outcomes of mental health. Besides, due to outcome variable imbalance's categories the overall performances not optimal.

5. Conclusions and policy implications

This study shows that among all the machine learning models, XGBoost performs better than the others in handling binary imbalanced classification problems and has the best AUROC along with the best precision-recall balance. While having overall high accuracy but uniformly across models, the evaluation places importance on the disadvantage of applying accuracy to imbalanced data and highlights the sensitivity-specificity trade-offs. Tree-based ensemble methods, particularly XGBoost, reflected improved performance to suppress false positive rates at respectable true positive rates, critical for clinical and epidemiological applications. While models like KNN and Neural Networks offered higher recall, their poor precision made them less desirable with high false-positive rates. The results establish XGBoost's robustness against data heterogeneity and class imbalance, although there are future improvement's., class-balancing techniques and blend models to enhance recall and generalizability. This study findings reflects urgent need for screening mental health interventions integrated to both digital and reproductive health. Lastly, model choice should be guided by the specific misclassification cost within the target application. Our investigation finds top most 5 determinants are anxiety, currently working, age group 45-49 years, internet users, menstruated last 6 weeks etc. Developing community-based programs for both mental and reproductive health with online activity, we can reduce the depression. Future work includes enhancing recall, exploring hybrid, or blended modelling, and extending surveillance to additional sociodemographic environments so that predictive models remain actionable and generalizable.

Authors' contributions: Md. Salek Miah (MSM) design the study, analysing the data and written the manuscript initially. Afterwards, Mohammad Ohid Ullah (MOU) finalize the manuscript.

Acknowledgement: Bangladesh Demographic and Health Survey 2022(BDHS) for providing free cost of data. The author appreciates the reviewer's constructive dialogue, which has led to significant improvements in the paper.

References

[1] Rutowski, T., Lu Y, Shriberg, E., Chlebek, P., Harati, A., and Oliveira, R. Depression and Anxiety
Prediction Using Deep Language Models and Transfer Learning.
[2] Muzaffar, R., Koly, K. N., Choudhury, S., Biswas, MAAJ, Kader, S. B, Abdullah, R., et al. (2022). Generalized anxiety disorder among Bangladeshi university students during COVID-19 pandemic: gender specific findings from a cross-sectional study. Discover Mental Health. 2022 Dec 1;2(1).
[3] Xie, H., Tan, Z., and Lv, X. (2024). Application of Artificial Intelligence in Financial Risk Management in a Company. In: ACM International Conference Proceeding Series. Association for Computing Machinery; 2024. p. 349–54.
[4] Duan, Z., Wang, Y., Ta,o Y., Bower, J. L., Yu, R., Wang, S., et al. (2019). Relationship between trait neuroticism and suicidal ideation among postpartum women in China: Testing a mediation model. J Affect Disord. 2019 Sep 1; 256:532–5.
[5] Moon, Y., and Woo, H. (2024). Key risk factors of generalized anxiety disorder in adolescents: machine learning study. Front Public Health. 2024;12.
[6] Mahmud, S., Mohsin, M., Muyeed, A., Nazneen, S., Sayed, Md. A., Murshed, N., et al. (2022). Machine Learning Approaches for Predicting Suicidal Behaviors among University Students in Bangladesh during the COVID-19 Pandemic [Internet]. 2022. Available from: https://www.researchsquare.com/article/rs-2069873/v1
[7] Chowdhury, A. H., Rad, D., and Rahman, M. S. (2024). Predicting anxiety, depression, and insomnia among Bangladeshi university students using tree-based machine learning models. Health Sci Rep. 2024 Apr 1;7(4).
[8] Nayan, M., Uddin, M., Hossain, M., Alam, M., Zinnia, M., Haq, I., et al. (2022). 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 Journal of Social Health and Behavior. 2022 Apr 1;5(2):75–84.
[9] Park, J. H., Shin, Y. Bin, Jung, D., Hur, J. W., Pack, S. P., Lee, H. J., et al. (2024). Machine learning prediction of anxiety symptoms in social anxiety disorder: utilizing multimodal data from virtual reality sessions. Front Psychiatry. 2024;15.
[10] Rezapour, M., and Elmshaeuser, S. K. (2022). Artificial intelligence-based analytics for impacts of COVID-19 and online learning on college students' mental health. PLoS One. 2022 Nov 1;17(11 November).
[11] Mitchell, B., Chivers, G., and Drife, J. (2023). Weight Monitoring and Antipsychotics: Are We Compliant With NICE Guidelines? BJPsych Open. 2023 Jul;9(S1):S9–S9.
[12] Sakib, N., Islam, Md. K., and Faruk, T. (2025). A Machine Learning Approach for Multi-Level Anxiety Screening among University-Going Students using Wireless EEG Signals. In Institute of Electrical and Electronics Engineers (IEEE); 2025. p. 21–5.
[13] Rezapour, M., and Elmshaeuser, S. K. (2022). Artificial intelligence-based analytics for impacts of COVID-19 and online learning on college students' mental health. PLoS One. 2022 Nov 1;17(11 November).
[14] Mahmud, S., Mohsin, M., Muyeed, A., Nazneen, S., Sayed, Md. A., Murshed, N., et al. (2022). Machine Learning Approaches for Predicting Suicidal Behaviors among University Students in Bangladesh during the COVID-19 Pandemic [Internet]. 2022. Available from: https://www.researchsquare.com/article/rs-2069873/v1
[15] Vu, T., Dawadi, R., Yamamoto, M., Tay, J. T., Watanabe, N., Kuriya, Y., et al. (2025). Prediction of depressive disorder using machine learning approaches: findings from the NHANES. BMC Med Inform Decis Mak. 2025 Dec 1;25(1).
[16] Saki,b N., Islam, M. K., and Faruk, T. (2024). A Machine Learning Approach for Multi-Level Anxiety Screening among University-Going Students using Wireless EEG Signals. In:

Proceedings - 2024 7th Asia Conference on Cognitive Engineering and Intelligent lnteraction, CEII 2024. Institute of Electrical and Electronics Engineers Inc.; 2024. p. 21–5.
[17] Ballı, M., Dogan, A. E, Senol, S. H., and Eser, H. Y. (2025). Machine learning based identification of suicidal ideation using non-suicidal predictors in a university mental health clinic. Sci Rep. 2025 Dec 1;15(1).
[18] Xu, X., Peng, H., Sun, L., Bhuiyan, MZA., Liu, L., and He, L. (2021). FedMood: Federated Learning on Mobile Health Data for Mood Detection. 2021 May 20; Available from: http://arxiv.org/abs/2102.09342
[19] Mitchell, B., Chivers, G., and Drife, J. (2023). Weight Monitoring and Antipsychotics: Are We Compliant With NICE Guidelines? BJPsych Open. 2023 Jul;9(S1):S9–S9.
[20] Li, Y., Yang, Y., Song, P., Duan, L., and Ren, R. (2025). An improved SMOTE algorithm for enhanced imbalanced data classification by expanding sample generation space. Sci Rep. 2025 Dec 1;15(1).
[21] Park, J. H., Shin, Y. Bin, Jung, D., Hur, J. W., Pack, S. P., Lee H. J,, et al. (2024). Machine learning prediction of anxiety symptoms in social anxiety disorder: utilizing multimodal data from virtual reality sessions. Front Psychiatry. 2024;15.
[22] Huang, H. H., Lu, C. J., Jhou, M. J., Yang, W. J., Chang, H. C., Chen, M. S., et al. (2023). Using a Decision Tree Algorithm Predictive Model for Sperm Count Assessment and Risk Factors in Health Screening Population. Risk Manag Healthc Policy. 2023;16:2469–78.
[23] Sarakshetrin, A., Sujimongkol, C., Rongmuang, D., Chantra, R., and Nimwatanakul, S. (2025). Support Vector Machine-Based Prediction Model for Healthcare Workforce Transition Success Under Decentralization. F1000Res. 2025 Jan 9;14:49.
[24] Jiang, Z., Chen, J., Feng, L., Jin, M., Liu, S., Wang, L., et al. (2023). Associations between maternal occupational exposures and pregnancy outcomes among Chinese nurses: a nationwide study. Reprod Health. 2023 Dec 1;20(1).
[25] Rahman, R., Talukder, A., Das, S., Saha, J., and Sarma, H. (2024). Understanding and predicting pregnancy termination in Bangladesh: A comprehensive analysis using a hybrid machine learning approach. Medicine (United States). 2024 Jun 28;103(26):e38709.
[26] Samuel, O., Zewotir, T., and North, D. (2024). Application of machine learning methods for predicting under-five mortality: analysis of Nigerian demographic health survey 2018 dataset. BMC Med Inform Decis Mak. 2024 Dec 1;24(1).
[27] Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., and Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. J Big Data. 2024 Dec 1;11(1).
[28] Taskiran, S. F., Turkoglu, B., Kaya, E., and Asuroglu, T. (2025). A comprehensive evaluation of oversampling techniques for enhancing text classification performance. Sci Rep. 2025 Dec 1;15(1).
[29] Ostovar, A., Davari, D. D., and Dzikuć, M. (2025). Determinants of Design with Multilayer Perceptron Neural Networks: A Comparison with Logistic Regression. Sustainability (Switzerland). 2025 Mar 1;17(6).
[30] UNICEF Pakistan ©Nutrition Wing of MoNHSR&C Designed By Human Design Studios.
[31] Li, Y., Yang, Y., Song, P., Duan, L., and Ren, R. (2025). An improved SMOTE algorithm for enhanced imbalanced data classification by expanding sample generation space. Sci Rep. 2025 Dec 1;15(1).
[32] Zhu, S., Huang, Z., Chen, X., Jiang, W., Zhou, Y., Zheng, B., et al. (2025). Construction and evaluation of machine learning-based prediction model for live birth following fresh embryo transfer in IVF/ICSI patients with polycystic ovary syndrome. Journal of Ovarian Research. 2025 Dec 1;18(1).
[33] Hakka,l S., and Lahcen, A. A. (2024). XGBoost To Enhance Learner Performance Prediction. Computers and Education: Artificial Intelligence. 2024 Dec 1;7.

[34] Wang, H., Li, J., Liu, H., Guo, F., Xue, T., Guan, T., et al. (2021). Association of maternal exposure to ambient particulate pollution with incident spontaneous pregnancy loss. Ecotoxicol Environ Saf. 2021 Nov 1;224.

[35] Islam, M. M., Shoukot, Jahan. Kibria, N. M., Kumar, S., Roy, D. C., and Karim. M. R. (2024). Prediction of undernutrition and identification of its influencing predictors among under-five children in Bangladesh using explainable machine learning algorithms. PLoS One. 2024 Dec 1;19(12).

[36] Fernández. A, García, S, Herrera. F., and Chawla, N. V. (2018). SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. Vol. 61, Journal of Artificial Intelligence Research. 2018.

[37] Powers DMW, Ailab. Evaluation: from precision, recall and f-measure to ROC, informedness, markedness & correlation.

[38] Willinger, M., Ko ,C. W., and Reddy, U. M. (2009). Racial disparities in stillbirth risk across gestation in the United States. Am J Obstet Gynecol. 2009;201(5):469.e1-469.e8.

[39] Huang, H. H., Lu, C. J., Jhou, M. J., Yang, W. J., Chang, H. C., Chen, M. S., et al. (2023). Using a Decision Tree Algorithm Predictive Model for Sperm Count Assessment and Risk Factors in Health Screening Population. Risk Manag Healthc Policy. 2023;16:2469–78.

[40] Hua, Y., Stead, T. S., George, A., and Ganti, L. (2025). Clinical Risk Prediction with Logistic Regression: Best Practices, Validation Techniques, and Applications in Medical Research. Academic Medicine & Surgery. 2025 Mar 7;

[41] Leeflang, MMG., and Kraaijpoel, N. (2018). Systematic reviews of diagnostic test accuracy in CMI. Vol. 24, Clinical Microbiology and Infection. Elsevier B.V.; 2018. p. 1115–6.

[42] Zhenya, Q., and Zhang, Z. (2021). A hybrid cost-sensitive ensemble for heart disease prediction. BMC Med Inform Decis Mak. 2021 Dec 1;21(1).

[43] Tharwat, A. (2018). Classification assessment methods. Applied Computing and Informatics. 2018;17(1):168–92.

[44] Chicco, D., and Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020 Jan 2;21(1).

[45] Poudel, S., Razee, H., Dobbins, T., and Akombi-Inyang, B. (2022). Adolescent Pregnancy in South Asia: A Systematic Review of Observational Studies. Vol. 19, International Journal of Environmental Research and Public Health. MDPI; 2022.

[46] Luque, A., Carrasco, A., Martín, A., and de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. Pattern Recognit. 2019 Jul 1;91:216–31.

[47] Boughorbel, S., Jarray, F., and El-Anbari, M. (2017). Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PLoS One. 2017 Jun 1;12(6).

[48] Mujahid, M., Kına EROL, Rustam, F., Villar, M. G., Alvarado, E. S., De La Torre Diez, I., et al. (2024). Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering. J Big Data. 2024 Dec 1;11(1).

[49] Kaya, Y., Aydın, Y., Kaya, C., Tahta, T., and Çelik, Ö. (2025). A Novel Machine Learning Model for Predicting Natural Conception Using Non-Laboratory-Based Data. Reproductive Sciences. 2025 Aug 1;

[50] Dunne, J., Foo, D., Dachew, B. A., Duko, B., Gebremedhin, A. T., Nyadanu, S. D., et al. (2024). Diabetic and hypertensive disorders following early pregnancy loss: a systematic review and meta-analysis. Vol. 71, eClinicalMedicine. Elsevier Ltd; 2024.

[51] Naznin, S., Uddin, M. J, and Kabir, A. (2025). Identifying determinants of under-5 mortality in Bangladesh: A machine learning approach with BDHS 2022 data. PLoS One. 2025 Jun 1;20(6 June).

[52] Sergeev, S., Diakova, I., and Nadirashvili, L. (2024). Neural networks pipeline for quality management in IVF laboratory. Journal of IVF-Worldwide. 2024 Oct 23;2(4).

[53] Fazal, S., Shahid, Y., Nadhman, A., Amjad, N., Naeem, M., Khan, S., et al. (2024). Maternal Malnutrition and Its Associated Adverse Pregnancy Outcomes and Childbirth in Dadu District, Sindh, Pakistan. Pakistan Journal of Public Health. 2024 Aug 5;14(Special.NI):127–33.
[54] WHO (2024). World mental health report: Transforming mental health for all. 2022.
[55] Bangladesh Demographic and Health Survey 2022 Final Report [Internet]. 2024. Available from: www.niport.gov.bd

Appendix: Descriptive Statistics for All variables

Variable Name	Categories	Description
Household Division	Barishal, Chattogram, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, Sylhet	Geographic region where the household is located in Bangladesh. Dhaka has the highest representation (15.12%).
Area	Rural, Urban	Residential classification. 65.12% of households are rural.
Household Head Sex	Female, Male	Gender of household head. Male-headed households dominate (87.87%).
Pressure to Pregnant	No, Yes	Whether the respondent faces pressure to become pregnant. Rare (3.09% affirmative).
Currently Working	No, Yes	Employment status. 69.95% of respondents are not working.
Father's Education	No education, Primary incomplete, Primary complete, Secondary incomplete, Secondary complete or higher	Education level of the respondent's father. 34.88% have incomplete secondary education.
Internet User	No, Yes	Internet usage. 71.11% do not use the internet.
Depression	No, Yes	Self-reported depression. 4.82% report experiencing depression.
Anxiety	No, Yes	Self-reported anxiety. 4.11% report experiencing anxiety.
ICT Ownership	0, 1, 2, 3, 4	Number of ICT devices owned (e.g., phones, computers). 31.67% own 2 devices.
ICT Usage	0, 1, 2, 3	Frequency of ICT use. 49.94% do not use ICT.
Age Group	15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49	Respondent's age. Largest group: 25–29 years (18.12%).
Education	None/Primary, Secondary, Higher	Respondent's education level. 46.39% have secondary education.
Number of Children	0, 1, 2	Children per household. 58.97% have 1 child.
Pregnant	No, Yes	Current pregnancy status. 6.17% are pregnant.
Decision Power	0–5	Empowerment score (0 = lowest, 5 = highest). 49.21% score 5.
Wealth Quintile	1 (Poorest) to 5 (Richest)	Household wealth distribution. Richest quintile: 22%.
IPV Experience	No, Yes	Intimate partner violence experience. 13.09% report IPV.
Residing Status	Not with husband/partner, With husband/partner	Living arrangement. 84.80% live with their partner.
Currently Abstaining	No, Yes	Abstinence from sexual activity. Rare (3.77% affirmative).
Menstruated Last 6 Weeks	No, Yes	Menstrual activity. 69.64% menstruated recently.
Employment in Decisions	Health, Purchase, Money, Visit, FP	Participation in household decisions. Highest in FP decisions (89.81%).