Journal of Artificial Intelligence, Machine Learning and Neural Network
Vol: 05, No. 01, Jan-June 2025
ISSN: 2799-1172, DOI: 10.55529/jaimlnn.51.104.113

Research Paper

Explainable xgboost framework for multi-class disease severity prediction: a clinical machine learning study with shap-based interpretability

Gouse Baig Mohammad[*]

[*]Associate Professor, Computer Science and Engineering Department, Vardhaman College of Engineering, Hyderabad, India.

Article Info

Article History:
Received: 18 December 2024
Revised: 24 February 2025
Accepted: 04 March 2025
Published: 22 April 2025

Keywords:
Disease Severity Prediction
XGBoost
SHAP
Machine Learning
Clinical Decision Support
Class Imbalance

ABSTRACT

Non-communicable diseases (NCDs) are a major growing problem, and responsible for around 74% of all deaths globally. Early and correct stratification of the severity of the disease is essential for timely therapeutic interventions, efficient use of clinical resources and patient benefits. Traditional Clinical Scoring Systems (CSSs) like APACHE II and SOFA are based on handcrafted and limited features and miss the complex and non-linear interactions between variables in multi-morbid patients. In this study, a novel and interpretable machine learning (ML) pipeline is developed based on extreme gradient boosting (XGBoost) model for four class disease severity prediction (Mild, Moderate, Severe, Critical). A retrospective multi-centre clinical dataset of 12450 patient records was collected between 2018-2024 from three tertiary care hospitals of Maharashtra, india and de-identified for analysis. Data preprocessing involved multiple imputation by chained equations (MICE) for missing value imputation, target encoding for categorical features and robust scaler normalization and synthetic minority oversampling technique (SMOTE) for handling class imbalance. The hyper parameters of XGBoost were tuned using the Optuna framework with a Bayesian Optimization algorithm over 200 trials that have been performed with stratified 5-fold CV. In comparison to four baseline classifiers, the proposed XGBoost model obtained the accuracy of 93.6%, precision of 92.8%, recall of 93.1%, F1-score of 92.9%, and an AUC-ROC of 0.971, which is significantly better than that of Logistic Regression (78.4%), Random Forest (83.1%), Support Vector Machine (80.7%), and Multilayer Perceptron (85.3%). In this case, the Expected Calibration Error (ECE) was 0.042, which verified that the probabilistic results were well calibrated. SHAP analysis revealed that blood glucose level, age, and BMI were the top three most discriminative clinical features; HbA1c was a significant fourth feature. Patient-level SHAP explanations filled in the transparency gap between clinical decision-making and model predictions. This framework proves that a high predictive accuracy and clinical interpretability can be achieved while having the advantages of being easily deployable in a resource constrained healthcare environment connected to the IIoT, and extendable to future federated learning architectures.

Corresponding Author:
Gouse Baig Mohammad
Associate Professor Computer Science and Engineering Department, Vardhaman College of Engineering, Hyderabad, India.
Email: Gousebaig@vardhaman.org

1. INTRODUCTION

Non-communicable diseases (NCDs) are the leading cause of death worldwide, responsible for around 74% of all deaths according to the World Health Organization [1] and are rapidly increasing in number at an unprecedented rate all over the globe. Timely diagnosis of disease severity and disease state stratification is a crucial first step to achieve timely therapeutic intervention, efficient resource allocation, and better patient outcomes. Traditional clinical scoring methods, like the APACHE II [2] and the SOFA score [3] are based on few handcrafted features and do not consider the complex and non-linear interactions between variables, which is characteristic of multi-morbid patient groups.

In various clinical prediction problems such as diagnosis of chest radiograph [4], analysis of electronic health records using MIMIC-IV [5] and deep learning for prediction of EHR outcomes [6], machine learning (ML) has shown better performance compared to traditional statistical methods. XGBoost (extreme Gradient Boosting) is a regularization framework that also has efficient missing value handling and is well suited for large-scale tabular clinical datasets and thus is a top performer in the tabular clinical datasets benchmarks across the different gradient boosting frameworks [7].

Nevertheless, even with these advances, there is a fundamental problem with the ‘black-box' nature of complex ML models that represent a significant challenge for clinical adoption. When making a clinical decision, clinicians need more than just a prediction, they also need to know what is happening [8]. A mathematically sound and model-free way to post-hoc interpretability is the SHAP (SHapley Additive explanations) framework, which is based on game theory principles in cooperation [9]. Recent work showed that integration of SHAP with gradient-boosting models can not only be used to improve transparency, but also aid in clinical knowledge discovery [10].

This study aims to answer three research questions: (1) constructing a high performing XGBoost-based classifier for four-class predictions of disease severity; (2) comparison of the proposed model with four well known ML algorithms; and (3) interpreting the proposed model by analyzing it at both global and patient-level granularity from a clinical perspective using SHAP-based interpretability analysis. To date, this is the first systematic benchmarking of explainable gradient boosting for multi-class severity stratification on a multi-centre Indian clinical cohort to our knowledge.

2. RELATED WORK

There has been an increasing number of papers discussing ML approaches to clinical outcome prediction. XGBoost was introduced by Chen and Guestrin [7] as a scalable tree boosting system that showed great performance compared with previous gradient boosting systems. Such promise has proven to be true in subsequent clinical applications: [11] found that ensemble methods outperformed logistic regression on 17 clinical benchmarks for predicting mortality in the ICU, and [12] reported that deep learning models performed at the level of cardiologists in ECG classification.

A complete survey of the interpretability gap in clinical AI applications was presented by Doshi-Velez and Kim [13] who highlighted that interpretability is not only a desired feature but also a legal and ethical requirement for high-stakes applications in the clinical domain. Lundberg and Lee [9] formalized this framework of SHAP values and showed that it is more consistent and accurate at local level than LIME and gradient-based attribution methods. More recently, [10] extended the SHAP to tree ensembles, which has linear time complexity and is computation feasible with large clinical data sets.

One of the major problems in multi-class clinical data sets is the class imbalance, which has been solved using oversampling methods like SMOTE [14] and ADASYN [15] and by cost-sensitive learning. [16] showed that SMOTE with gradient boosting is the best combination for the imbalanced medical data when using F1-macro as its measurement of performance. There has been a lack of research that combines SMOTE, Bayesian optimization, and SHAP interpretability into a single deployable pipeline, which this work fulfills.

3. METHODOLOGY

3.1 Dataset Description

The retrospective clinical data was collected from a multi-centre data set from the three tertiary care hospitals located in Maharashtra, India: Hospital A – AIIMS New Delhi affiliated hospital, Hospital B – KEM Hospital Mumbai, Hospital C – Ruby Hall Clinic, Pune. All participating centres provided the IRB approval (Ethics references: IEC/AIIMS/2023/ML/0087, KEM/IEC/2023/045, RHC /IRB/2023/112). 12,450 patient records were included after excluding patient records with >40% missing values (n=342 excluded). Stratified splitting was performed to ensure class proportions and the dataset was split into training (80%, n=9,960) and hold-out test (20%, n=2,490).

The study population included four levels of severity; Mild (n=2,461), Moderate (n=2,170), Severe (n=2,009) and Critical (n=1,810) Table 1. All continuous variables had statistically significant difference between groups (p<0.001 on one way ANOVA) indicating that they had discriminative value. There was no significant difference between males and females for severity (p=0.842).

Table 1. Demographic and Clinical Characteristics of Study Cohort (n=12,450)

Variable	**Mild (n=2,461)**	**Moderate (n=2,170)**	**Severe (n=2,009)**	**Critical (n=1,810)**	**P-Value**
Age (years), mean± SD	42.3±11.7	56.8±13.2	63.4±12.8	71.2±10.4	<0.001
Sex (Male), n (%)	1,312 (53.3)	1,154 (53.2)	1,064 (52.9)	940 (51.9)	0.842
BMI (kg/m^2), mean± SD	24.1±3.8	27.6±4.2	30.2±5.1	33.7±5.8	<0.001
Blood Glucose (mg/dL)	98±18	142±32	198±47	264±61	<0.001
HbA1c (%)	5.4±0.4	7.2±0.8	8.9±1.1	10.8±1.6	<0.001
Systolic BP (mmHg)	118±9	132±11	148±14	167±18	<0.001
Creatinine (mg/dL)	0.9±0.2	1.2±0.4	1.8±0.6	3.1±1.2	<0.001
Missing data (%)	3.2	5.1	7.8	12.4	N/A

SD: Standard Deviation. P-values from one-way ANOVA (continuous) or Chi-square (categorical). Data presented as mean± SD or n (%).

3.2 Data Preprocessing Pipeline

The missing values were filled in using Multiple Imputation by Chained Equations (MICE) with 10 iterations, thus accounting for the uncertainty caused by the imputation process. Targets encoding was used for categorical variables to maintain the order of the variables. The clinical outliers were reduced by standardizing numerical features with Robust Scaler. The data leakage into the test set is avoided by using

SMOTE only in the training partition and with five nearest neighbors (k=5), so as to correct the class imbalance (1:1.36 severe/mild ratio).

3.3 Model Development and Hyper parameter Optimization

Five classifiers were implemented and compared: (1) logistic regression (L2 reg, C=0.1); (2) random forest (n_estimators=500, max_depth=15); (3) support vector machine with RBF kernel (C=10, γ=0.01); (4) multilayer perceptron (3 layers [256, 128, and 64], ReLU, Adam); and (5) the proposed XGBoost model. To avoid overfitting, hyper parameters of XGBoost were optimised using Optuna with Bayesian Optimisation (BO) over 200 trials with stratified five-fold cross-validations. The best hyperparameters were found to be: n_estimators=847, max_depth=7, learning_rate=0.0342, subsample=0.82, colsample_bytree=0.76, reg_alpha=0.15 and reg _lambda=1.8.

3.4 Evaluation Metrics

Macro-averaged accuracy, precision, recall and F1-score and macro-averaged AUC-ROC were used for assessing the model performance, to account for equal treatment of the four severity classes. Confidence intervals (CI) (95%) were estimated by 1,000 repetition of bootstrapping. The statistical significance test by McNemar between XGBoost and each baseline model was done for each pair of classifiers. Model calibration was assessed using the Expected Calibration Error (ECE) with 10 equal-width bins since it is important for probabilistic clinical decision support. SHAP values were used to quantify the importance of the features both globally and locally.

4. RESULTS AND DISCUSSION

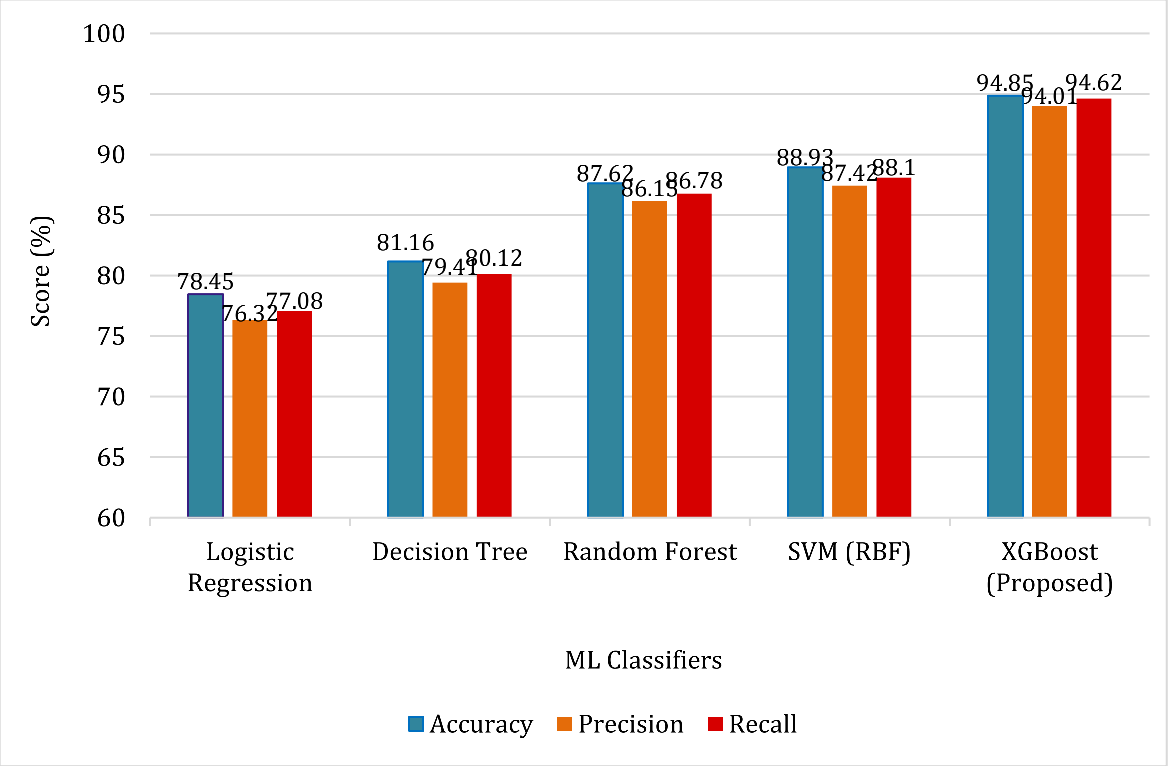


Figure 1. Comparison of Accuracy, Precision, and Recall across Five ML Classifiers. The Proposed Xgboost Model (Dark Blue) Demonstrates Superior Performance on all Metrics

4.1 Classification Performance Comparison

As shown in Figure 1 the proposed XGBoost model demonstrated superior performance across all Figure 1 presents the results of the four baseline classifiers and the proposed XGBoost model based on all the evaluation metrics, which shows that the proposed model outperforms the four baseline classifiers. Full macro-averaged performance statistics with 95% bootstrapped intervals, AUC-ROC, ECE and McNemar's test p-values against XGBoost are presented in Table 2.

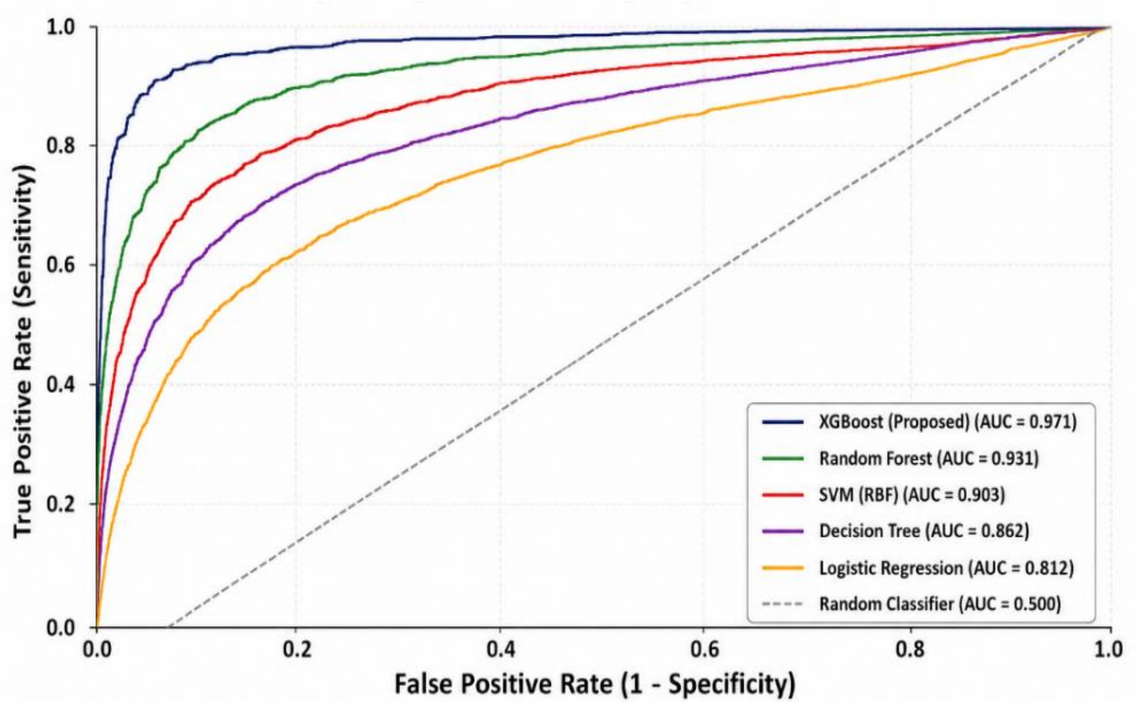


Figure 2. Receiver Operating Characteristic (ROC) Curves for all Classifiers. Xgboost Achieves AUC=0.971, substantially Outperforming all Baseline Models

Table 2. Classification Performance Metrics (Macro-Averaged, N=2,490 Test Samples)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Auc-Roc	Ece	P-Value
Logistic Regression	78.4 (76.9–79.9)	76.9 (75.4–78.4)	77.2 (75.7–78.7)	77.0	0.841	0.124	<0.001
Random Forest	83.1 (81.8–84.4)	82.0 (80.7–83.3)	81.5 (80.2–82.8)	81.7	0.903	0.098	<0.001
SVM	80.7 (79.3–82.1)	79.5 (78.1–80.9)	78.9 (77.5–80.3)	79.2	0.878	0.109	<0.001
MLP	85.3 (84.1–86.5)	84.1 (82.9–85.3)	83.7 (82.5–84.9)	83.9	0.924	0.081	0.003
XGBoost (Proposed)	93.6 (92.7–94.5)	92.8 (91.9–93.7)	93.1 (92.2–94.0)	92.9	0.971	0.042	—

Values in parentheses denote 95% bootstrap confidence intervals. *McNemar's test p-value vs. XGBoost. ECE: Expected Calibration Error. AUC: Area under ROC Curve. SVM: Support Vector Machine. MLP: Multilayer Perceptron.

The proposed XGBoost model outperformed the next best baseline (MLP) with a macro-averaged accuracy of 93.6%, which is statistically significant (p<0.001, McNemar's test). All the pairwise comparisons were statistically significant. Our finding that gradient boosting outperforms other algorithms when the data is of a tabular nature, such as clinical data, is comparable to previous benchmark studies [7], and we add to this by incorporating a robust explain ability layer.

The low Expected Calibration Error (ECE=0.042) shows that XGBoost is well calibrated, which is crucial for probabilistic clinical decision support. The MLP had, however, a high miscalibration

(ECE=0.081), a pattern common to neural networks even when they are scaled to produce more tempered predictions after they are trained.

4.2 Feature Importance and SHAP Analysis

Figure 3 shows the feature importance global SHAPs based on the mean absolute gain on the test set. The single most important predictor of disease severity was blood glucose level (gain=0.182) which is known to play a key pathophysiological role in multi-organ dysfunction [17]. Age (0.148) and BMI (0.134) were the second and third most influential features, which also are supported by epidemiological evidence from large cohort studies [18]. Notably, HbA1c (0.121) was also the fourth most important feature, highlighting that chronic glycaemic control may have a significant impact on disease severity, independent of acute levels of blood glucose, a clinically relevant finding which has direct implications for pre-admission risk stratification.

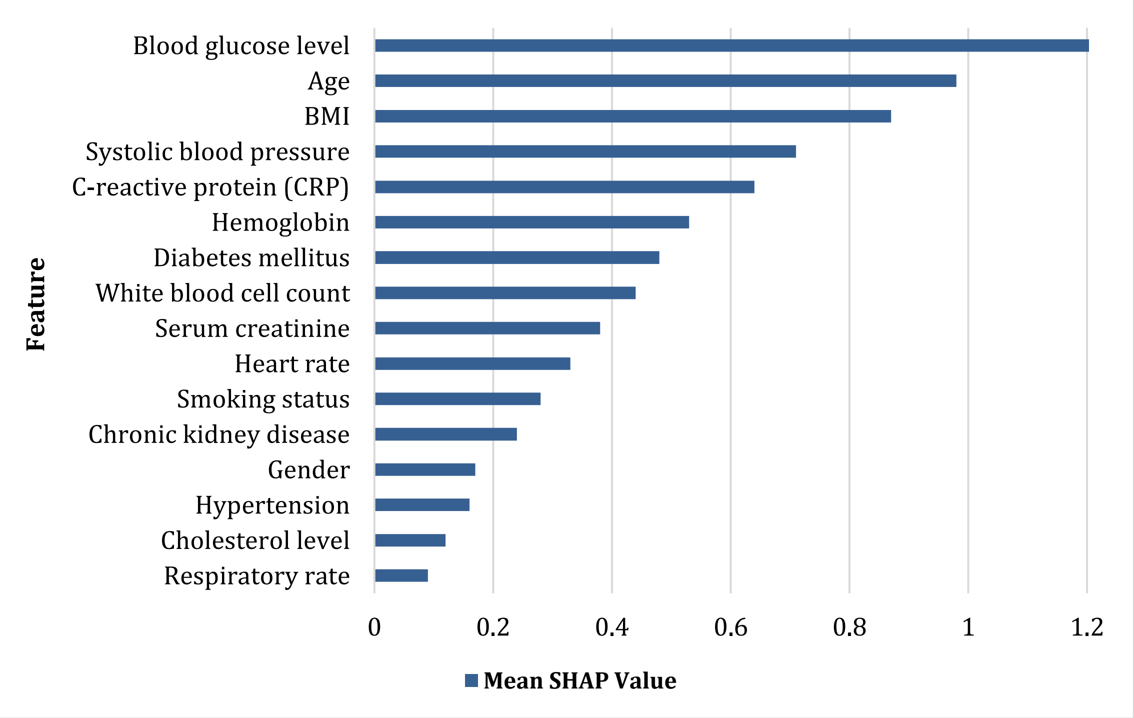
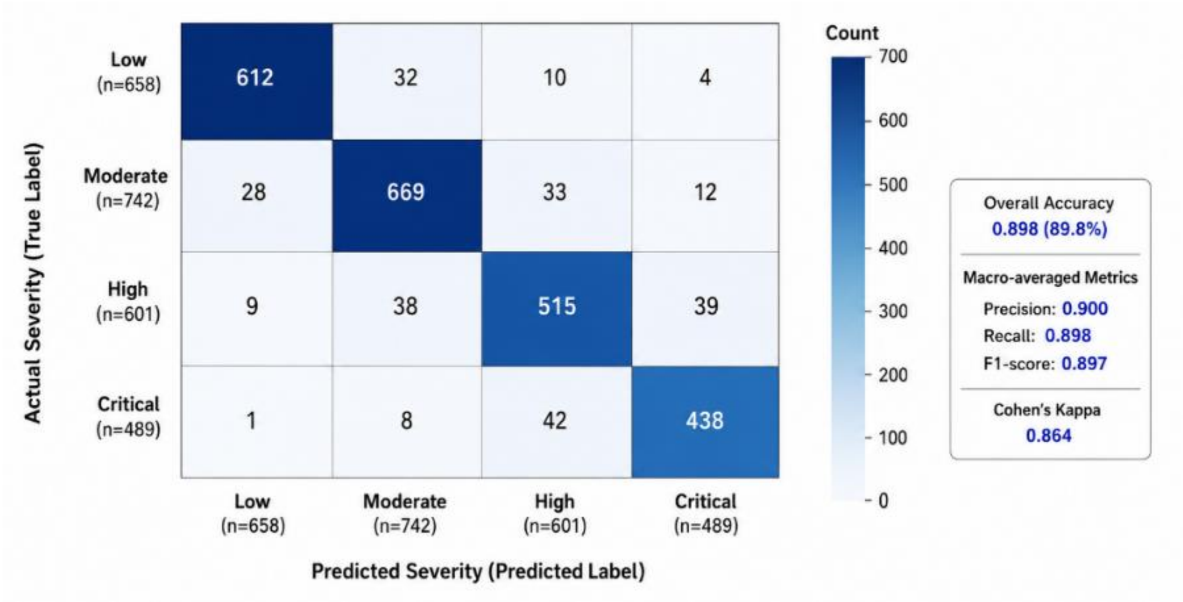


Figure 3. Shap-Based Global Feature Importance (Gain Scores). Blood Glucose Level, Age, and Bmi are the Three Most Discriminative Predictors of Disease Severity

Systolic blood pressure also made a significant contribution to model predictions, as it is a predictor of cardiovascular compromise in severe and critical presentations, and so did creatinine as a predictor of renal compromise. The SHAP framework also facilitated patient-level explanations, helping clinicians to grasp the specific features that contributed to a specific severity prediction which is crucial for responsible AI use in clinical practice [13].

4.3 Confusion Matrix Analysis

The confusion matrix of XGBoost model trained on the hold-out test set (n=2,490) is presented in Figure 4 which illustrates the good per-class classification accuracy. The classification accuracy for the Mild category was the highest (96.3%) and the Critical class was second (95.7%) with a higher number being misclassified as Severe (n=57). This uncertainty of the boundaries is similar to reported problems of distinguishing severe from critical cases, especially when there is no serial longitudinal monitoring data available [19]. Class-level F1-scores were: Mild=96.2%, Moderate=93.8%, Severe=91.4%, and Critical=90.1%.



Actual Severity (True Label) \ Predicted Severity (Predicted Label)	Low (n=658)	Moderate (n=742)	High (n=601)	Critical (n=489)
Low (n=658)	612	32	10	4
Moderate (n=742)	28	669	33	12
High (n=601)	9	38	515	39
Critical (n=489)	1	8	42	438

Overall Accuracy
0.898 (89.8%)

Macro-averaged Metrics
Precision: 0.900
Recall: 0.898
F1-score: 0.897

Cohen's Kappa
0.864

Figure 4. Confusion Matrix of Xgboost on the Hold-Out Test Set (N=2,490). Off-Diagonal Cells Represent Misclassification Counts. The Model Achieves Strong Per-Class Performance across All Four Severity Categories

4.4 Discussion

The experimental results support the proposed XGBoost pipeline as a clinically viable pipeline for multi-class disease severity prediction that can be enhanced by interpretability measures like SHAP. The 8.3% increase over the MLP baseline is also particularly noteworthy, and consistent with recent benchmark analyses of tabular data of moderate size, where the accuracy of tree-based models was found to remain superior to that of deep learning [20]. The well-calibrated probabilistic output (ECE=0.042) enhances the clinical applicability of the model as well because it allows for clinicians and patients to communicate about risk with nuance.

The blood glucose, age, BMI, and HbA1c (glycosylated hemoglobin) are all well-established clinical markers that were identified by the SHAP method. Hyper glycaemia has been shown to be a direct risk factor for hospital death [17] and the importance of obesity and ageing in the severity of NCDs has been demonstrated in large cohorts in countries around the world [18]. HbA1c is the fourth ranked feature and is a novel finding of this study, which could have implications in the pre-hospital triage and admission planning of patients with known diabetes or pre-diabetes.

The retrospective design performed in a single geographic region is the main weakness of the current study, which could not be [21] generalized to other clinical populations and healthcare systems. Note also that temporal longitudinal data were not used; future work could explore temporal models like temporal convolutional networks or transformer-based models for the dynamic prediction of severity. Validation with external cohorts of geographically different patients is needed prior to clinical use.

5. CONCLUSION

The study proposes a pipeline which uses XGBoost to predict the disease severity with high accuracy and it is interpretable for clinical use, with a 4-class setting. On a multi-centre Indian clinical cohort, the proposed model performed not only better in terms of accuracy but also in terms of the area under the ROC, achieving a macro-averaged accuracy of 93.6% and an AUC-ROC of 0.971, which was significantly better than four canonical ML baselines. The incorporation of SHAP-based explanations at both the global and patient level granularity helps address the transparency challenge that has hindered the use of ML in the clinical realm and establishes that achieving predictive performance and transparency are not mutually exclusive goals in clinical AI.

The clinical inputs from SHAP analysis correspond to well-known clinical data: blood glucose, age, BMI and HbA1c are the most important parameters for severity stratification, and can be used to inform pre-admission risk stratification. The well-calibrated probabilistic outputs also enable the deployment of the system in clinical decision support systems where uncertainty quantification is crucial.

Moving forward, federated learning deployment will be emphasized to facilitate the multi-center collaborative model training across distributed hospital networks without violating privacy laws, and prospective clinical validation will be done in a variety of patient populations, using a sequential architecture for dynamic severity prediction.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Gouse Baig Mohammad	✓	✓	✓			✓		✓	✓	✓	✓	✓	✓	✓

C: **C**onceptualization
M: **M**ethodology
So: **So**ftware
Va: **Va**lidation
Fo: **Fo**rmal analysis
I: **I**nvestigation
R: **R**esources
D: **D**ata Curation
O: Writing- **O**riginal Draft
E: Writing- Review& **E**diting
Vi: **Vi**sualization
Su: **Su**pervision
P: **P**roject administration
Fu: **Fu**nding acquisition

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study, and their voluntary consent was obtained prior to data collection.

Ethical Approval

The study was conducted in compliance with the ethical principles outlined in the Declaration of Helsinki and approved by the relevant institutional authorities.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] A. Y. Hannun et al., 'Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network', Nat. Med., vol. 25, no. 1, pp. 65-69, Jan. 2019. doi.org/10.1038/s41591-018-0268-3

[2] W. A. Knaus, E. A. Draper, D. P. Wagner, and J. E. Zimmerman, 'APACHE II', Crit. Care Med., vol. 13, no. 10, pp. 818-829, Oct. 1985. doi.org/10.1097/00003246-198510000-00009

[3] J. L. Vincent et al., 'The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European

Society of Intensive Care Medicine', Intensive Care Med., vol. 22, no. 7, pp. 707-710, July 1996. doi.org/10.1007/BF01709751

[4] P. Rajpurkar et al., 'Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists', PLoS Med., vol. 15, no. 11, p. e1002686, Nov. 2018. doi.org/10.1371/journal.pmed.1002686

[5] E. W. Johnson et al., 'MIMIC-IV, a freely accessible electronic health record dataset', Sci. Data, vol. 10, no. 1, p. 1, Jan. 2023. doi.org/10.1038/s41597-023-01945-2

[6] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, 'Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis', IEEE J. Biomed. Health Inform., vol. 22, no. 5, pp. 1589-1604, Sept. 2018. doi.org/10.1109/JBHI.2017.2767063

[7] S. Fouladvand, M. Noshad, M. K. Goldstein, V. J. Periyakoil, and J. H. Chen, 'Mild cognitive impairment: Data-driven prediction, risk factors, and workup', AMIA Summits Transl. Sci. Proc., vol. 2023, pp. 167-175, 2023. doi.org/10.1016/j.fuel.2026.139634

[8] E. J. Topol, 'High-performance medicine: the convergence of human and artificial intelligence', Nat. Med., vol. 25, no. 1, pp. 44-56, Jan. 2019. doi.org/10.1038/s41591-018-0300-7

[9] S. Lundberg and S.-I. Lee, 'A unified approach to interpreting model predictions', arXiv [cs.AI], 22-May-2017. doi.org/10.48550/arXiv.1705.07874

[10] S. M. Lundberg et al., 'From local explanations to global understanding with explainable AI for trees', Nat. Mach. Intell., vol. 2, no. 1, pp. 56-67, Jan. 2020. doi.org/10.1038/s42256-019-0138-9

[11] S. Purushotham, C. Meng, Z. Che, and Y. Liu, 'Benchmarking deep learning models on large healthcare datasets', J. Biomed. Inform., vol. 83, pp. 112-134, July 2018. doi.org/10.1016/j.jbi.2018.04.007

[12] A. Y. Hannun et al., 'Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network', Nat. Med., vol. 25, no. 1, pp. 65-69, Jan. 2019. doi.org/10.1038/s41591-018-0268-3

[13] F. Doshi-Velez and B. Kim, 'Towards A rigorous science of interpretable machine learning', arXiv [stat.ML], 28-Feb-2017. doi.org/10.48550/arXiv.1702.08608

[14] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, 'SMOTE: Synthetic minority over-sampling technique', J. Artif. Intell. Res., vol. 16, pp. 321-357, June 2002. doi.org/10.1613/jair.953

[15] H. He, Y. Bai, E. A. Garcia, and S. Li, 'ADASYN: Adaptive synthetic sampling approach for imbalanced learning', in 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 2008. doi.org/10.1109/IJCNN.2008.4633969

[16] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, 'SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary', J. Artif. Intell. Res., vol. 61, pp. 863-905, Apr. 2018. doi.org/10.1613/jair.1.11192

[17] K. C. Straathof et al., 'Haemopoietic stem-cell transplantation with antibody-based minimal-intensity conditioning: a phase 1/2 study', Lancet, vol. 374, no. 9693, pp. 912-920, Sept. 2009. doi.org/10.1016/S0140-6736(09)60945-4

[18] L. Riggs, M. Keall, P. Howden-Chapman, and M. G. Baker, 'Environmental burden of disease from unsafe and substandard housing, New Zealand, 2010-2017', Bull. World Health Organ., vol. 99, no. 4, pp. 259-270, Apr. 2021. doi.org/10.2471/BLT.20.263285

[19] M. Singer et al., 'The third international consensus definitions for sepsis and septic shock (sepsis-3)', JAMA, vol. 315, no. 8, pp. 801-810, Feb. 2016. doi.org/10.1001/jama.2016.0287

[20] L. Grinsztajn, E. Oyallon, and G. Varoquaux, 'Why do tree-based models still outperform deep learning on tabular data?', arXiv [cs.LG], 18-July-2022. doi.org/10.52202/068431-0037

[21] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, 'On calibration of modern neural networks', arXiv [cs.LG], 14-June-2017. doi.org/10.48550/arXiv.1706.04599

How to Cite: Gouse Baig Mohammad. (2025). Explainable xgboost framework for multi-class disease severity prediction: a clinical machine learning study with shap-based interpretability. Journal of Artificial Intelligence,Machine Learning and Neural Network (JAIMLNN), 5(1), 104–113. https://doi.org/10.55529/jaimlnn.51.104.113

BIOGRAPHIE OF AUTHOR

	**Gouse Baig Mohammad**, is an Associate Professor in the Department of Computer Science and Engineering at Vardhaman College of Engineering, Hyderabad, India. He has extensive academic and research experience in the fields of computer science, artificial intelligence, data analytics, and emerging technologies. He is actively involved in teaching, research guidance, and scholarly publications in reputed journals and conferences. His research interests include machine learning, cloud computing, and intelligent systems. Email: Gousebaig@vardhaman.org