Original Article

A Comparative Study of Machine Learning Models for Hypertension Risk Assessment With Explainable Artificial Intelligence Integration

Saviour Inyang [1], Felix Njoku [2], Imeh Umoren [1], Ubong Etuk [3], Victoria Essien [4]

1. *Cybersecurity, Federal University of Technology, Ikot Abasi, NGA*
2. *Software Engineering, Federal University of Technology, Ikot Abasi, NGA*
3. *Cybersecurity, Akwa Ibom State University, Mkpat Enin, NGA*
4. *Computer Science, Akwa Ibom State University, Mkpat Enin, NGA*

Received: April 18, 2026 | Review began: May 05, 2026 | Review ended: May 31, 2026 | Published: June 09, 2026

Abstract

Hypertension is a major global health challenge and one of the leading causes of cardiovascular diseases, stroke, kidney failure, and premature death. Early identification and prediction of hypertension risk severity are essential for improving patient outcomes and supporting effective clinical decision-making. However, conventional diagnostic approaches often struggle to capture the complex relationships between clinical and lifestyle-related risk factors associated with hypertension. To address this challenge, this study proposes an explainable machine learning framework for identifying and predicting hypertension risk severity using patients' clinical and lifestyle data collected from the University of Uyo Teaching Hospital. The proposed framework incorporates data preprocessing techniques such as missing value handling, data normalization, and feature ranking using Principal Component Analysis to improve data quality and model performance. Several machine learning algorithms, including eXtreme Gradient Boosting (XGBoost), Random Forest, Artificial Neural Network, Support Vector Machine, and Decision Tree, were implemented and evaluated using stratified 10-fold cross-validation to reduce overfitting and ensure reliable model evaluation. Performance assessment was carried out using accuracy, precision, recall, F1-score, and area under the curve. The experimental results revealed that XGBoost and Random Forest achieved the highest accuracy of 98%, followed by Support Vector Machine with 95%, Artificial Neural Network with 94%, and Decision Tree with 89%. To enhance transparency and interpretability, Explainable Artificial Intelligence was integrated into the framework using Local Interpretable Model-Agnostic Explanations to identify the most influential features contributing to predictions. XGBoost emerged as the best-performing model due to its ability to effectively learn complex patterns within the dataset. The findings demonstrate that explainable machine learning models can significantly improve hypertension risk prediction while also promoting interpretability, trust, and informed decision-making for healthcare professionals and patients.

Categories: Explainable AI, Health Informatics, Machine Learning (ML)

Keywords: model, explainability, clinical features, machine learning, hypertension

Introduction

High blood pressure is a major contributor to cardiovascular diseases and early mortality around the world. Even after extensive research and clinical trials, the precise mechanisms behind hypertension are still not fully understood [1]. While there are numerous medications available, people with hypertension often find it challenging to keep their blood pressure levels in check [2]. According to the World Health Organization (WHO), it is defined as a systolic blood pressure ≥140 mmHg and/or a diastolic blood pressure ≥90 mmHg, based on the average of two or more readings taken at each of two or more visits after initial screening [3]. High blood pressure is a major health problem worldwide and is linked to many cases of heart disease, stroke, and early death. In 2015, an estimated 10.5 million deaths were linked to a systolic blood pressure of 110-115 mmHg or higher, and 7.8 million deaths were linked to a systolic blood pressure of 140 mmHg or more [4]. Hypertension poses a serious global health challenge by increasing the risk of cardiovascular diseases such as heart failure, coronary artery disease, and stroke, which are among the leading causes of premature death globally [5]. The widespread presence of this disease, combined with the critical impact it has on our health, makes hypertension a critical focus for public health solutions, interventions, and strategies [6]. The burden of hypertension not only affects the individuals suffering from the disease, but it also has economic impacts regarding healthcare costs, lowered productivity, and premature death [7].

Furthermore, Artificial Intelligence (AI) and Machine Learning (ML) are rapidly changing healthcare technology and strategies [8]. These technologies have offered promising solutions for tackling hypertension. AI models can analyze large patient datasets to identify individuals at high risk of developing hypertension and make recommendations on how they can prevent it [9]. These models have been used for several forms of predictive and preventive analyses; for example, logistic regression can determine whether someone suffers from hypertension or not. Other models have been created to identify pre-hypertensive patients. Golino et al. [10] were able to classify prehypertensive patients using data features that included Body Mass Index (BMI), Waist and Hip Circumference, as well as Waist-to-Hip Ratio. Another instance of the application of ML in hypertension is presented in [11], where diabetes and hypertension data were utilized to create an ensemble ML model to prevent and predict risk factors for both diabetes and hypertension. These ML models demonstrated high accuracy and efficiency by predicting and recognizing risk factors using complex data, with results greater than 90% [12,13]. Accurate risk assessment and evaluation of the severity of prehypertension and hypertension are important in making accurate and timely clinical decisions for patients [14]. Traditionally, these hypertension risk assessments were heavily based on manual processes, such as measuring blood pressure using medical devices and analyzing basic risk factors such as heredity, age, and weight. While these methods are helpful, they do not usually understand or incorporate the complexities of the human system [15]. In addition, manual assessments can be tedious, time-consuming, and subject to human error. Therefore, incorporating AI into hypertension prediction and risk assessment has its benefits. AI can provide deeper insights into the prognosis of hypertension in addition to the traditional methods of risk assessment [16].

Nevertheless, Explainable AI (XAI) builds on ML advancements by giving further interpretation to how ML model makes its predictions [17]. In healthcare, XAI is important because it will provide trust and promote a faster rate of adoption of AI technology [18], particularly for severe health cases like hypertension [19]. XAI tools like Shapley Additive explanations (SHAP) or Local Interpretable Model-Agnostic Explanations (LIME) can indicate the various contributions like age, weight, etc., to a model's prediction [20]. The interpretability on ML models not only supports personal care, it also complies with clinical guidelines and ethics. Several studies have implemented explainability into the field of hypertension. For example, Elshawi et al. [21] utilized various model-agnostic XAI tools to visualize how models predicted an individual's risk of developing hypertension using cardiorespiratory fitness data. El-Dahshan et al. [22] utilized XAI in their ExHyptNet model to explain how the model makes its predictions using PPG signals. XAI and ML have the capability to be viable solutions for tackling hypertension in a transparent, ethical, and advanced manner. This study aims to develop a robust and comprehensive framework for predicting hypertension risk severity using ML and XAI, with a focus on improving the accuracy and interoperability of ML models.

Materials And Methods

The method employed in this study is presented in procedural stages that are depicted in Figure *1*. The study methodology is sequentially explained in this section. The process starts with data collection and data preprocessing (which includes data cleaning, data normalization, and feature ranking), followed by modelling (which includes training and testing), performance evaluation, and finally model explainability using LIME. Figure *1* shows the framework for the classification of Hypertension Risk Severity.

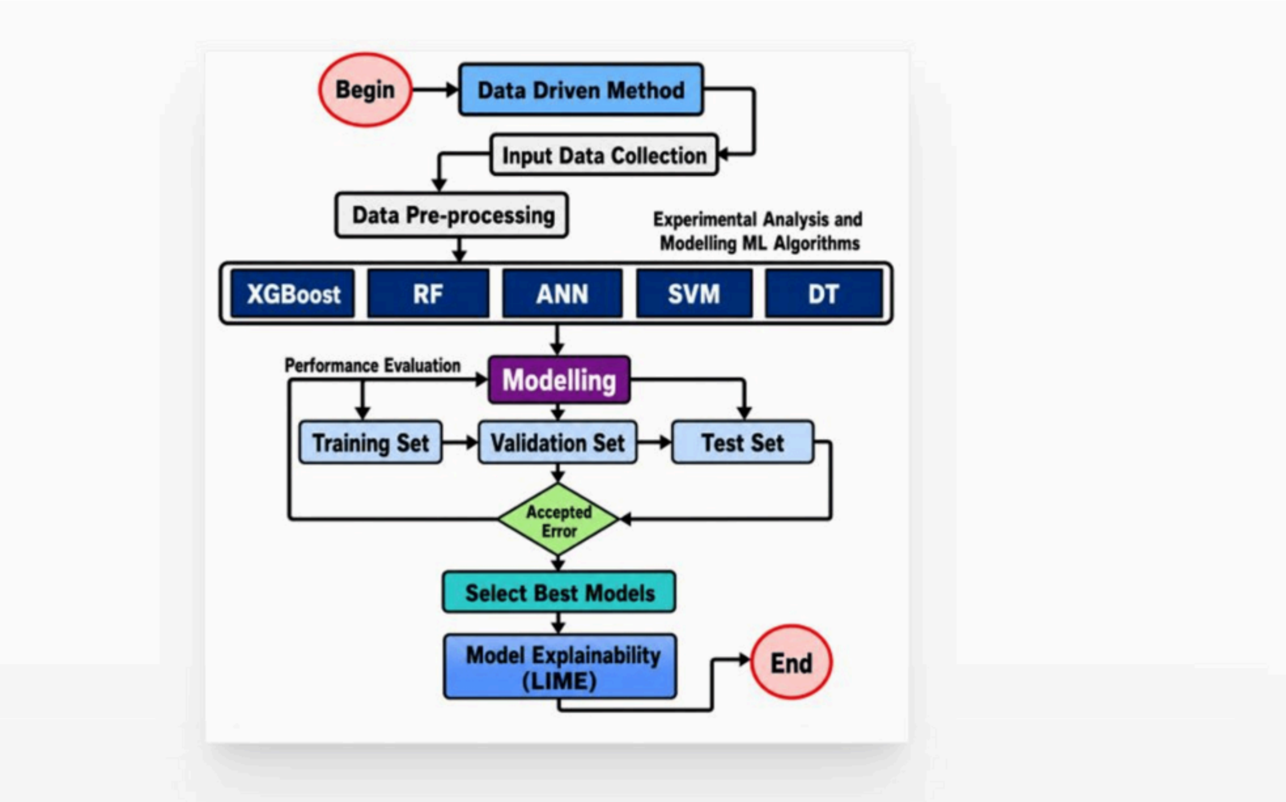

FIGURE 1: Framework for XAI-Based Model for Classification and Prediction of Hypertension Risk Severity

ANN, Artificial Neural Network; DT, Decision Tree; LIME, Local Interpretable Model-Agnostic Explanations; RF, Random Forest; SVM, Support Vector Machine; XAI, Explainable Artificial Intelligence; XGBoost, eXtreme Gradient Boosting

Data collection

The data were collected from the cardiology department of the University of Uyo Teaching Hospital and can be accessed in [23]. The patients constituted a cross-section of those who were diagnosed with hypertension and had been receiving treatment at the University of Uyo Teaching Hospital for a minimum of one year. Also, a chief medical expert in the cardiology unit, Dr. Umoh, was present during the data-gathering process to assist with issues and guide the process of data collection. The data provide an all-inclusive view of hypertension risk severity by identifying key features and factors necessary for in-depth analysis and decision-making. The description of the data is presented in Table *1*.

Features	Datatype	Description
Age	Integer	Age represents the patient's age in years, which influences hypertension risk.
Systolic Pressure (SP)	Integer	Systolic pressure measures the pressure in arteries during heartbeats, indicating hypertension severity.
Diastolic Pressure (DP)	Integer	Diastolic pressure measures arterial pressure between heartbeats, essential for diagnosing hypertension.
Family History (FH)	Boolean	Familia indicates genetic predisposition to hypertension based on family medical history.
Sodium	Boolean	Sodium tracks dietary sodium intake, a key factor affecting blood pressure levels.
Cholesterol	Number	Cholesterol measures blood cholesterol levels, influencing cardiovascular and hypertension risks.
Fasting Blood Sugar (FBS)	Number	Fasting Blood Sugar monitors glucose levels after fasting, often linked with hypertension.
Diabetic	Boolean	Diabetic identifies if the patient has diabetes or not, a comorbidity increasing hypertension risk.
Alcohol	Boolean	Alcohol assesses alcohol consumption, which can contribute to blood pressure variations.
Body Mass Index (BMI)	Number	Body Mass Index evaluates weight-to-height ratio, a significant factor in hypertension risk.
Risk Severity (RS)	Factor	Risk Severity determines the hypertension severity level (high or low) for individual patients.

TABLE 1: Datasets Description

Exploratory data analysis

In exploratory data analysis, it encompasses the discovery of patterns, anomalies, and the testing of hypotheses through the visual and statistical representation of data points from the dataset that will be used for ML techniques. It is important to understand the data before applying machine learning algorithms to it. This can be achieved through careful observation, where useful information regarding the correlation between variables and any inconsistencies in the data can be obtained. Figure *2* is therefore an illustration of the structure of the dataset.

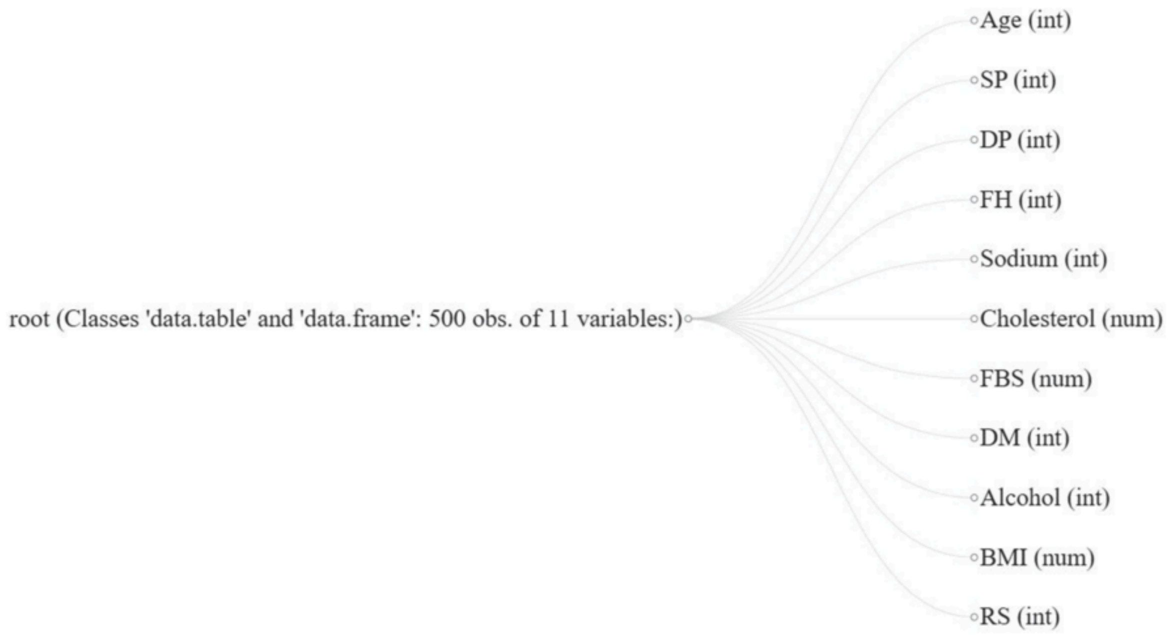

FIGURE 2: Data Structure

BMI, Body Mass Index; DP, Diastolic Pressure; DM, Diabetes Mellitus; FBS, Fasting Blood Sugar; FH, Family History; RS, Risk Severity; SP, Systolic Pressure

Furthermore, a correlation analysis to measure the relationship between the input variables was conducted. The correlation matrix indicates that most input variables are weakly to moderately associated, suggesting limited multicollinearity among predictors. The strongest positive relationship is observed between systolic pressure (SP) and diastolic pressure (DP) ($r = 0.75$), while age shows moderate positive correlations with several clinical variables, including cholesterol, fasting blood sugar (FBS), and diabetes status. In contrast, alcohol consumption demonstrates negative correlations with most variables, and BMI shows only weak associations across the matrix. Overall, these results suggest that the predictors capture related information but not highly redundant information, which supports their use in the subsequent analysis. Figure *3* depicts the correlation matrix for the input features in the dataset.

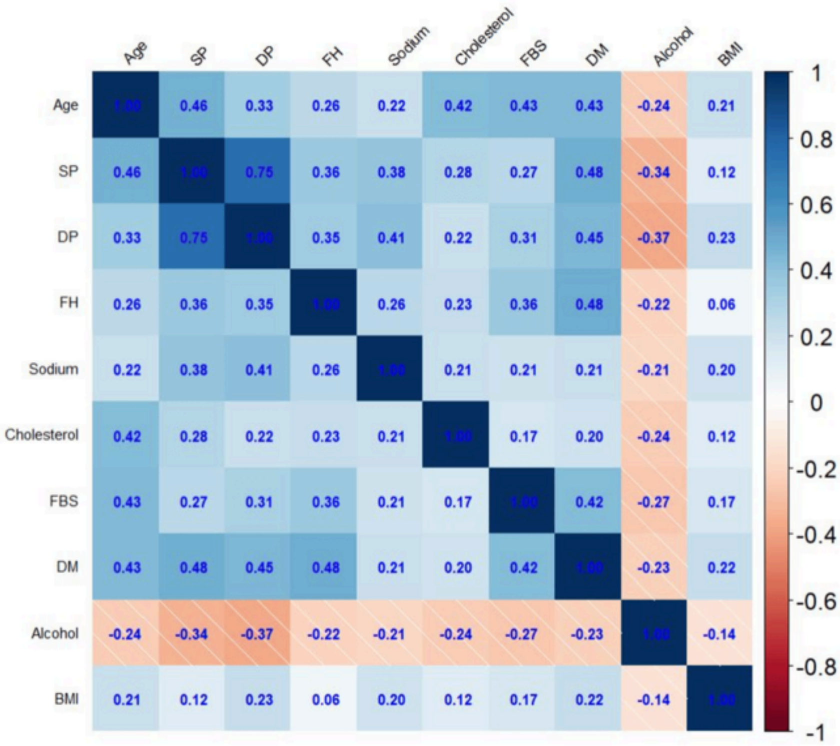

FIGURE 3: Correlation Matrix for Input Variables

BMI, Body Mass Index; DP, Diastolic Pressure; DM, Diabetes Mellitus; FBS, Fasting Blood Sugar; FH, Family History; RS, Risk Severity; SP, Systolic Pressure

Data preprocessing

Data preprocessing in this study involves data cleaning, including handling missing values, data normalization using the min-max method, and data splitting. The original data cannot be directly supplied to the ML algorithms because Non-Available (NA) values were present in some of the features in the dataset. For example, the Age and DP features contained NA values. To handle this, the mean of each column was taken and used to replace the NA values in the respective columns. Data normalization was necessary due to the large variation in the distribution of values across the features. The importance of this strategy is that it preserves all data associations, meaning that it does not lead to predisposition. Although the min-max method was applied, the scaling process remains within the value interval for prediction, but the conforming dispersal of the related variables remains exclusive to the existing value range [24]. Furthermore, this study ranks individual features in the dataset based on the Principal Component Analysis (PCA) method. To rank the features, the dataset, which involves multiple dimensions, was considered. Also, by computing the mean and the dimensions of the dataset, the covariance matrix was computed in Equation (1).

$$CO(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y}) \quad (1)$$

The computation of eigenvectors and their corresponding eigenvalues was carried out, which necessitated the sorting of the eigenvectors in ascending order. To transform the dataset into a new subspace, an eigenvector matrix was used, where each feature could be ranked as a principal component. Figure *4* shows the various principal components in the dataset and ranks the components based on their importance.

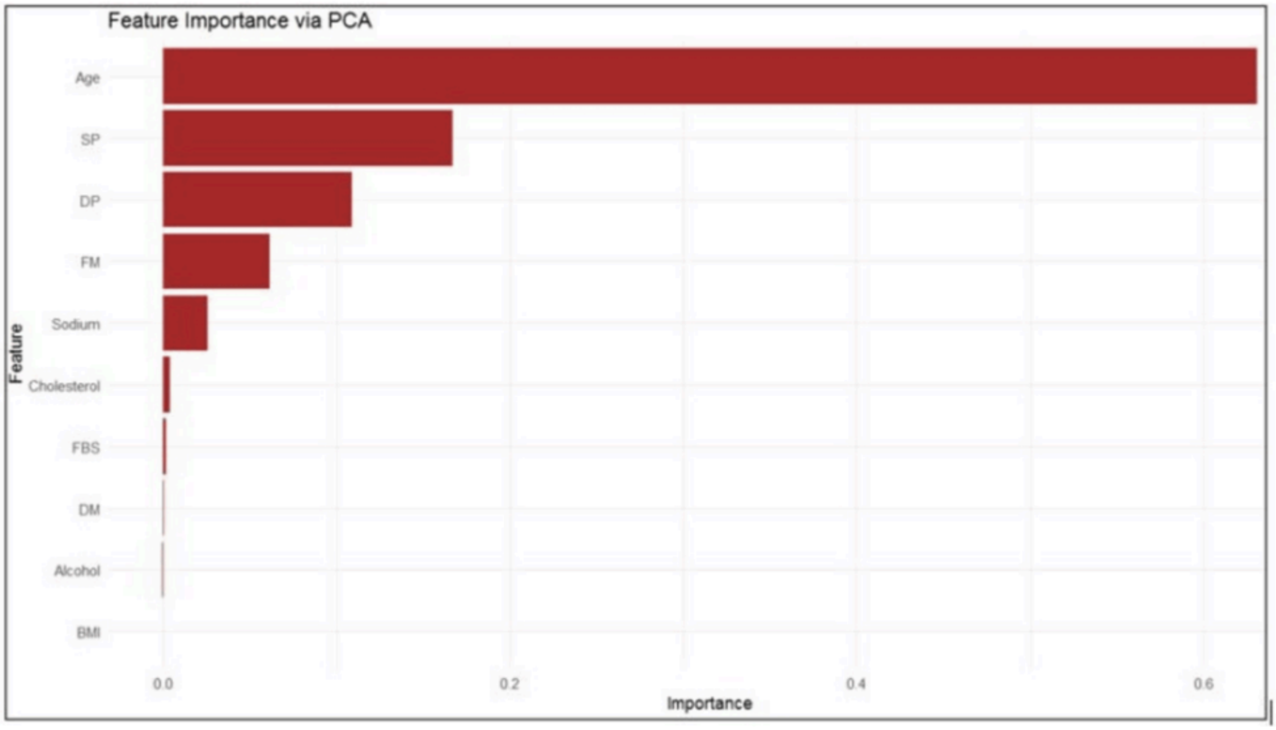

FIGURE 4: Feature Importance

BMI, Body Mass Index; DP, Diastolic Pressure; DM, Diabetes Mellitus; FBS, Fasting Blood Sugar; FH, Family History; RS, Risk Severity; SP, Systolic Pressure

Data splitting

At this phase, the original dataset was divided into two units: a test set (30%) and a training set (70%).

Modelling

We proposed a comparison of five ML models for the hypertension severity dataset. These ML models include eXtreme Gradient Boosting (XGBoost), Random Forest, Artificial Neural Network (ANN), Support Vector Machine (SVM), and Decision Tree, respectively. Consequently, the dataset was subjected to Stratified K-Fold cross-validation (CV) with k = 10 for the training of each model.

Performance evaluations

Here, the efficiency of the models will be assessed for hypertension risk severity. After the learning process of the model is completed, the ML model is evaluated based on its ability to correctly identify recall, F1-score, Cohen's kappa statistic, precision, and area under the curve (AUC). The mathematical formulations applied in this study for the performance metrics are presented in Equations 2-5, as shown in [25]. Where TP represents True Positive, TN represents True Negative, FN represents False Negative, and FP represents False Positive, respectively, as presented in Equations (2-5).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1-Score} = \frac{2TP}{2TP + FP + FN} \tag{5}$$

XAI

XAI models have become critical in disease prediction, specifically in hypertension management. This results in the need for justifying why ML model decisions must be transparent for numerous reasons. Lack of transparency makes the process through which AI models reach a certain output difficult to understand, hence the problem of making strategic actions based on their results [26]. Frequently used methods for model explainability include LIME and SHAP. LIME is one of the most widely used methods in the interpretation of the predictions of any classifier [27], while SHAP is based on cooperative game theory and gives equal weight to all features by assigning a single value. It calculates the contribution of a feature for a specific prediction by considering all feature possibilities [28]. This study will employ LIME for the interpretation of all ML classifiers utilized in this study based on the selected features relevant to the overall model classification.

System requirements

The successful implementation operation of the ML model for hypertension risk prediction requires specific system resources to ensure efficient computation, scalability, and robustness. These requirements are categorized into hardware and software components.

Hardware Requirements

The hardware infrastructure provides the necessary computational support for implementing the ML model, performing training, testing, and evaluation processes. The minimum hardware specifications include:

a. Intel Core i5 or higher, 2.5 GHz processor

b. 8 GB RAM

c. 500 GB Hard Disk Drive

d. LCD

Software Requirements

The software environment is essential for designing, simulating, and validating the ML models. The following software components are required:

a. 64-bit Windows 10 or higher operating system

b. Rstudio Editor

c. R (with libraries such as caret, tidyverse, explainers for preprocessing and training of the models to XAI for the hypertension risk data)

Results And Discussion

Experimental setup

As a result of carrying out cross-validation training, a stratified 10-fold cross-validation was utilized in this study to train the dataset. This technique is one strategy that is employed to evaluate ML models while ensuring that each fold of the dataset preserves the percentage of samples for each class. After the evaluation of the ML models, the two best-performing models were subjected to XAI in order to enhance transparency regarding the features that contributed to the success of the models' prediction process.

Results

Figure 5 shows the cross-validation plots for the five models used in this study. XGBoost (#boosting iteration) presents a line graph showing performance metrics across various boosting iterations. The performance of the XGBoost model depends on the tuning of max_tree_depth, colsample_bytree, subsample, and eta. It is observed that higher accuracy is achieved with larger tree depths; for example, depth = 3 performs better than depth = 1. Also, higher values of subsample (1.0) generally perform better than 0.5 or 0.75. Optimal results occur with a balanced colsample_bytree value, where values like 0.8 perform consistently well. In Random Forest (# randomly selected predictors), cross-validation shows a plot of accuracy against the number of randomly selected predictors. The highest accuracy is observed when the number of predictors is moderate (e.g., 6). It is observed that either too few or too many predictors (e.g., 2 or 10) result in lower performance, highlighting the importance of fine-tuning the number of predictors for optimal model accuracy. In the ANN (#hidden units), cross-validation illustrates accuracy in relation to hidden units and weight decay. It is observed that a weight decay of three hidden layers of 3 had increased in accuracy at specific configurations of the neurons. For the SVM (cost), model accuracy after cross-validation varies based on its tuning parameter which is the cost function. It is observed that an increase in the cost parameter results in an increase in accuracy, implying that the SVM model's performance improves with higher cost values. The Decision Tree (complexity parameter) model is evaluated in relation to the complexity parameter during cross-validation. It is observed that as the complexity parameter increases, accuracy decreases, therefore lower values of the complexity parameter provide better model prediction capability.

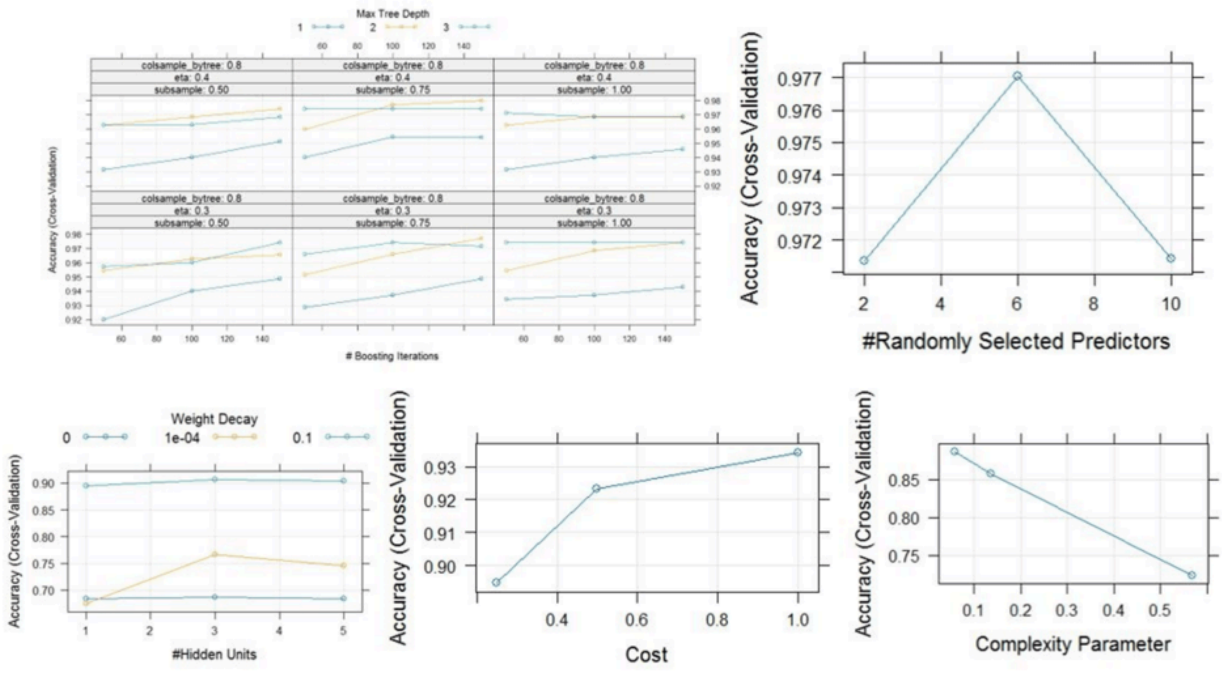

FIGURE 5: Cross Validation Plots for the XGBoost, RF, ANN, SVM and DT

ANN, Artificial Neural Network; DT, Decision Tree; RF, Random Forest; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting

Furthermore, the confusion matrices for the five ML models evaluated in this study are presented in Figure 6. Figure 6 shows the actual and predicted labels for each of the ML model using the 30% test set for which the actual values were known.

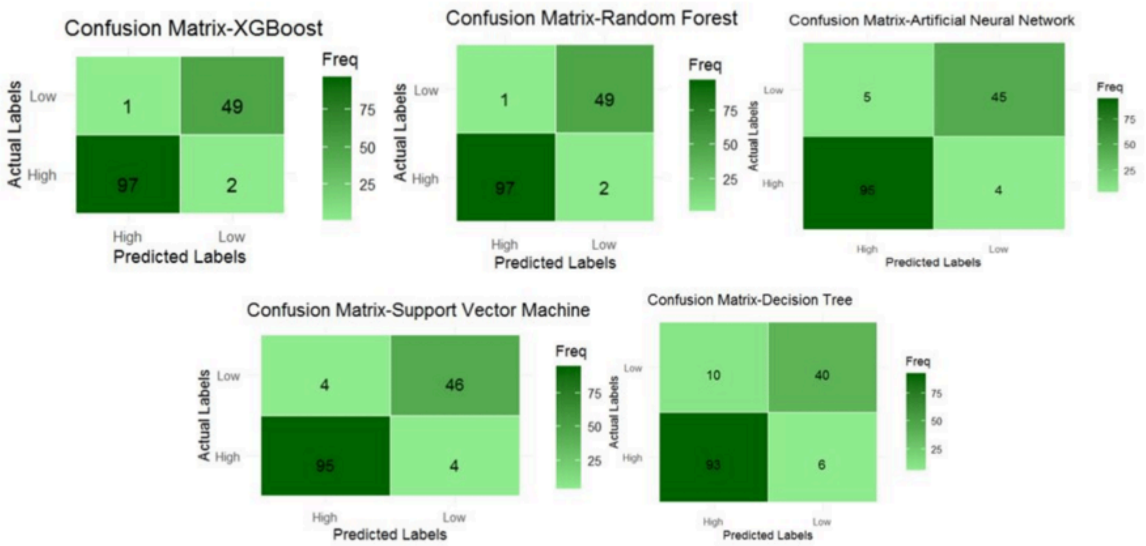

FIGURE 6: Confusion Matrix Plots for the XGBoost, Random Forest, Artificial Neural Network, Support Vector Machine, and Decision Tree

Again, the ROC curve alongside its AUC was computed for each ML model in this study; the results of the ROC-AUC are shown in Figure 7. The ROC curve was employed in order to determine the degree to which the models were able to differentiate between high- and low-risk severity of hypertension. If the ROC value is closer to 1, it indicates that the model's measure of separability is good. On the other hand, if the value is close to 0, it indicates that the model has the worst possible measure of separability.

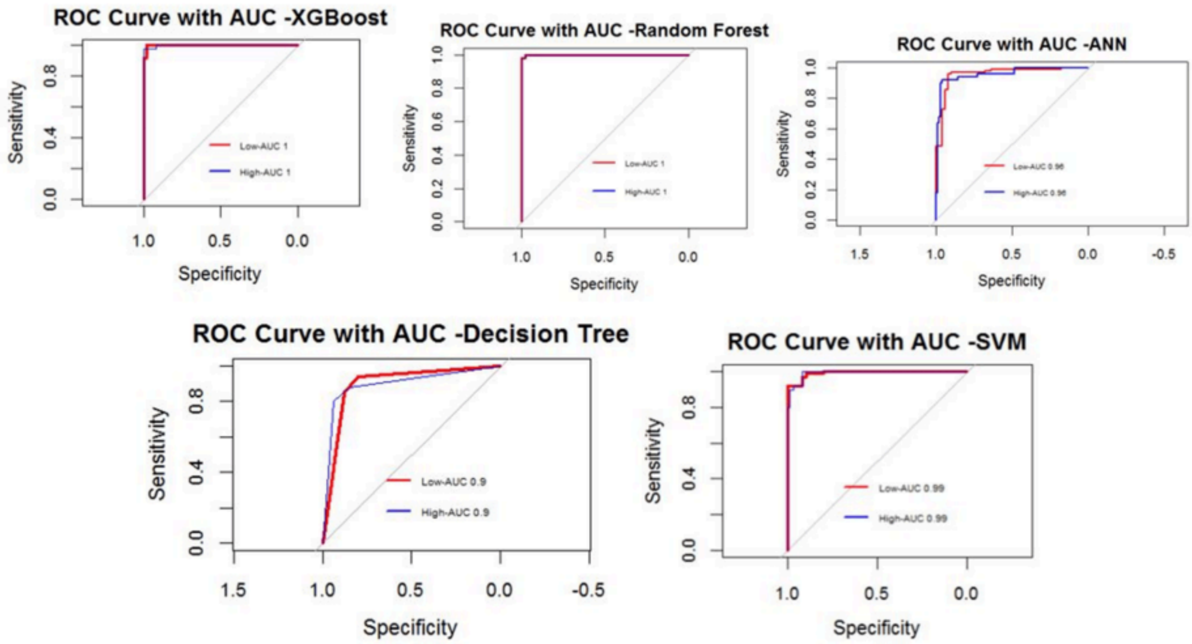

FIGURE 7: ROC Curve with AUC Plots for the XGBoost, Random Forest, ANN, SVM and Decision Tree

ANN, Artificial Neural Network; AUC, Area Under the Curve; ROC, Receiver Operating Characteristic; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting

The results of the ML classifiers presented in Table 2 indicate that both XGBoost and Random Forest outperform the other models in terms of all evaluated metrics. They achieve the highest Accuracy (0.98), Recall (0.98), Precision (0.99), and F1-score (0.99). Moreover, their AUC (1.00) suggests perfect discrimination ability, and their kappa (0.96) reflects excellent agreement between predicted and actual values. The ANN and SVM models perform slightly lower but still maintain competitive metrics. ANN achieves an Accuracy of 0.94, with Recall (0.96), Precision (0.95), F1-score (0.96), and a strong AUC (0.96). Similarly, SVM performs marginally better than ANN in terms of Precision (0.96) and AUC (0.99), but its kappa (0.88) is slightly higher than ANN's 0.86, indicating better reliability. The Decision Tree model demonstrates the weakest performance, with an Accuracy of 0.89 and the lowest scores across all metrics. Its AUC (0.90), F1-score (0.92), and kappa (0.75) reflect moderate performance, which may be attributed to its relatively simple structure compared to ensemble and advanced models.

ML Models	Accuracy	Recall	Precision	F1-score	AUC	Kappa
XGBoost	0.98	0.98	0.99	0.99	1.00	0.96
RF	0.98	0.98	0.99	0.98	1.00	0.96
ANN	0.94	0.96	0.95	0.96	0.96	0.86
SVM	0.95	0.96	0.96	0.96	0.99	0.88
DT	0.89	0.94	0.90	0.92	0.90	0.75

TABLE 2: Test Set Results on ML Models

ANN, Artificial Neural Network; AUC, Area Under the Curve; DT, Decision Tree; ML, Machine Learning; RF, Random Forest; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting

Furthermore, in this study, out of the five models selected for severity risk prediction, XGBoost and Random Forest algorithms performed the best in terms of all performance metrics. The provided visualizations in Figure *8* depict the results of the LIME analysis for both Random Forest and XGBoost models. Each figure demonstrates feature importance and explanation summaries for the respective models, providing insights into their predictive behavior and how individual features contribute to specific predictions.

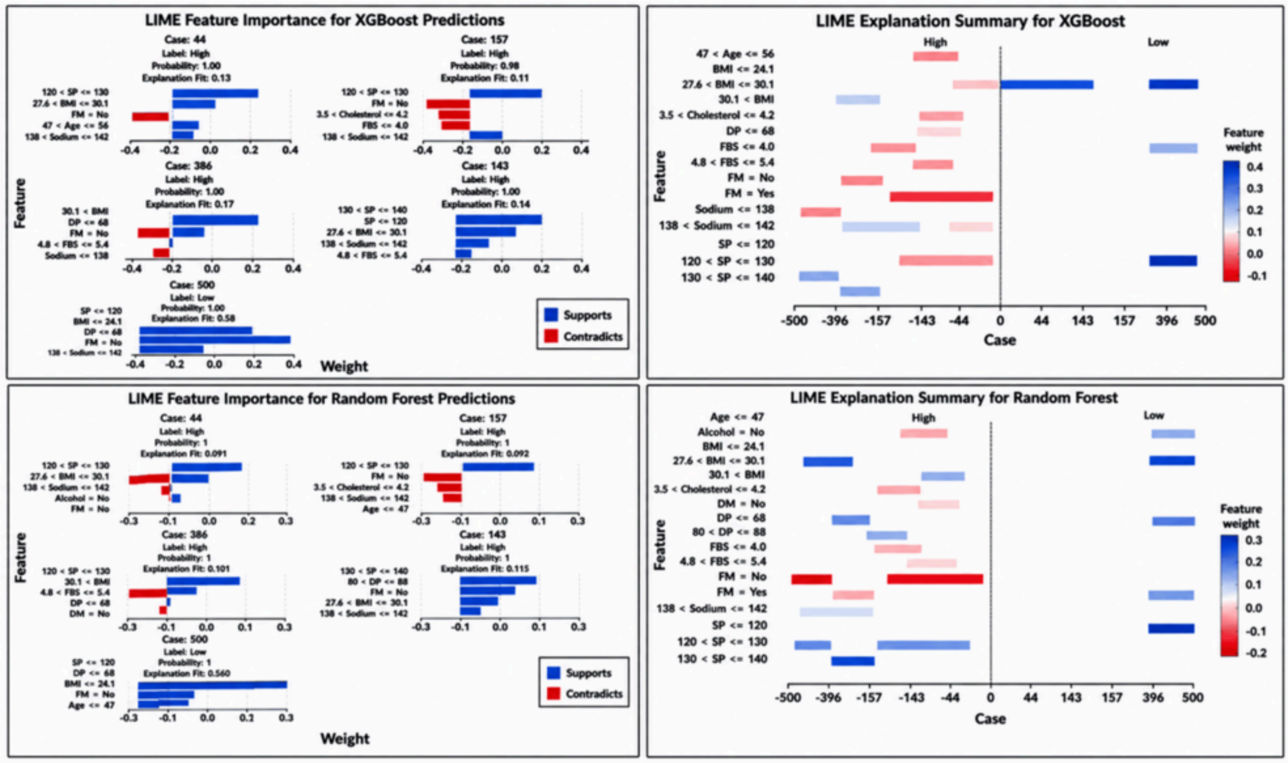

FIGURE 8: Lime Feature Importance and Explanations for XGBoost and Random Forest

LIME, Local Interpretable Model-Agnostic Explanations; XGBoost, eXtreme Gradient Boosting

To further evaluate the accuracy of the models and the XAI integration in the models performance, benchmarking this study with recently existed research in domain of hypertension prediction or cardiovascular disease prediction was carried out, which is presented in Table 3.

How to cite this article:
Inyang S, Njoku F, Umoren I, et al. (June 09, 2026) A Comparative Study of Machine Learning Models for Hypertension Risk Assessment With Explainable Artificial Intelligence Integration. Cureus J Comput Sci 3 : es44389-026-00111-0. DOI https://doi.org/10.7759/s44389-026-00111-0

Reference	Year	Integration of XAI		Model	Accuracy of Models Used				
					XGBoost	SVM	ANN	RF	DT
		SHAP	LIME						
[29]	2025	✓		GB, CatB, LGBM, RF, SVM, ANN, XGBoost		0.756	0.768	0.80	
[30]	2024	Nil	Nil	SVM , DT, GLM		0.90			0.83
[31]	2024	Nil	Nil	LR, SVM, RF, ANN, CRNet		94.6	0.971	0.963	
[32]	2024	✓		RF, XGBoost, LGBM, CatB, EnAutoML	0.98			0.97	
[33]	2024	✓		LGBM, LR, RF, SVM, NB, KNN		0.748		0.755	
[34]	2025	Nil	Nil	XGBoost, RF, SVM, DT	0.90	0.87		0.89	0.82
[35]	2025	✓		XGBoost, RF, LR, NB, KNN		0.897		0.88	
[36]	2025	✓		SVM, RF, KNN, DT, LR, NB, VotEn, StackEn		0.876		0.894	0.845
[37]	2025	✓		GLMnet, OFS, CART, OCT, RF,	0.734	0.568		0.833	

				XGBoost, NB, SVM					
[38]	2025	Nil	Nil	LR, XGBoost, RF, DT, CatB, ANN	0.633		0.62	0.578	0.60
Proposed Model	2026		✓	XGBoost, RF, ANN, SVM, DT	0.98	0.95	0.94	0.98	0.89

TABLE 3: Benchmarking Machine Learning Models for Hypertension Risk Assessment with XAI Integration

ANN, Artificial Neural Network; CatB, Categorical Boosting; CART, Classification and Regression Trees; CRNet, CardioRiskNet; DT, Decision Tree; GB, Gradient Boosting; GLM, Generalized Linear Model; KNN, K-Nearest Neighbors; LGBM, Light Gradient Boosting Machine; LIME, Local Interpretable Model-agnostic Explanations; LR, Logistic Regression; OCT, Optimal Classification Trees; OFS, Online Feature Selection; NB, Naive Bayes; RF, Random Forest; SHAP, SHapley Additive exPlanations; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting; XAI, Explainable Artificial Intelligence

Discussions

XGBoost and RF are the most effective classifiers, demonstrating exceptional performance across all metrics, followed by ANN and SVM, which exhibit slightly lower but still strong performance. The Decision Tree, while functional, lags significantly behind, suggesting limited applicability in high-stakes predictive tasks compared to the other models. For XAI, the Random Forest model illustrates the relative contribution of features to the prediction outcomes. The RF LIME feature importance plot in Figure *6* highlights that features such as SP, FBS, and BMI played significant roles in shaping the model's decisions. Positive contributions (in blue) and negative contributions (in red) are shown for each prediction case, suggesting which features supported or contradicted a given classification. The explanation fit values also indicate how well the local explanations align with the overall prediction probability, reinforcing the model's reliance on these features. The RF LIME explanation summary (in Figure *6*) aggregates feature importance across cases, showing that age, alcohol consumption, BMI, and SP consistently emerge as significant contributors. The weight of these features varies across cases, with some features having contrasting roles (e.g., BMI may support a "high" label in one instance but contradict it in another). This variability underscores the complexity of the Random Forest model's decision boundaries and its ability to adapt to varying feature interactions. For the XGBoost model, Figure *6* provides similar insights but reveals distinct feature importance patterns, which shows that XGBoost also relies heavily on SP, FBS, and BMI.

However, the distribution of supports and contradictions across cases differs slightly, with XGBoost demonstrating more defined contributions for certain features. For example, SP appears to have a consistently high influence on predictions, often dominating the explanation fit scores. The LIME explanation summary for XGBoost (in Figure *6*) shows a broader reliance on features like cholesterol, sodium levels, and DP compared to Random Forest. This suggests that XGBoost may capture additional subtle interactions in the data, leading to a more diverse feature utilization. The presence of contrasting contributions for features such as BMI and cholesterol reflects the nuanced decision-making of XGBoost, which is known for its ability to capture non-linear relationships. Comparing the two models, both Random Forest and XGBoost emphasize key clinical features such as SP, BMI, and FBS. However, XGBoost seems to leverage a wider range of features and exhibits stronger local explanations for certain cases. This could indicate that XGBoost, with its gradient-boosting approach, captures complex patterns more effectively than the Random Forest's ensemble of decision trees. These differences highlight the importance of model selection and interpretation when dealing with medical data, where understanding feature contributions can provide critical insights into patient outcomes.

The proposed solution in this study involves developing an integrative ML approach to predict hypertension severity risk using XGBoost, Random Forest, ANN, SVM, and Decision Tree models. The study incorporates XAI (LIME) to provide interpretability and insights into model predictions. The results achieved demonstrate that XGBoost and Random Forest outperform other models with an accuracy of 98% and an AUC of 1.00, indicating excellent predictive capability. ANN and SVM also performed well, while the Decision Tree model showed the weakest performance. The LIME analysis further highlighted the key features influencing predictions, such as SP, FBS, and BMI. What is unique in this study is the application of multiple ML models for severity risk classification, and leveraging XAI (i.e., LIME) to improve transparency in hypertension prediction, which is not extensively covered in previous studies.

Conclusions

This study developed a comprehensive framework for the identification and prediction of hypertension risk severity levels using machine learning models alongside their explainability techniques. Among the different machine learning models evaluated, XGBoost and Random Forest achieved very high predictive accuracy (98%), outperforming all other models. The best-performing model was XGBoost, which efficiently exploited a wide range of features, capturing complex patterns to yield more accurate predictions with stronger local explanations compared to the other models. LIME explanations of the models provided rich insights into their decision-making through key clinical features such as SP, FBS, and BMI, with additional contributions from cholesterol and DP for XGBoost. The insights from the explainability analysis are more interpretable and reliable, thereby showing good applicability in a clinical setup. These results indicate the potential of explainable machine learning models in improving hypertension management by facilitating better clinical decision-making and personalized treatment strategies. Further studies should be conducted to validate the framework with larger and more diverse datasets, integrating the insights into real-world clinical workflows for further improvement in the management and treatment of hypertension.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Saviour Inyang, Felix Njoku, Imeh Umoren, Ubong Etuk, Victoria Essien

Acquisition, analysis, or interpretation of data: Saviour Inyang, Felix Njoku, Imeh Umoren, Ubong Etuk, Victoria Essien

Drafting of the manuscript: Saviour Inyang, Felix Njoku, Imeh Umoren, Ubong Etuk, Victoria Essien

Critical review of the manuscript for important intellectual content: Saviour Inyang, Felix Njoku, Imeh Umoren, Ubong Etuk, Victoria Essien

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors would like to acknowledge the School of Computing and Information Technology, Federal University of Technology Ikot Abasi, for providing the smooth environment to conduct this research. The data and study materials are available upon request to the corresponding author.

References

1. Taler SJ: Initial treatment of hypertension. New England Journal of Medicine. 2018, 378:636-644. 10.1056/nejmcp1613481
2. Coffman TM: Under pressure: the search for the essential mechanisms of hypertension. Nature Medicine. 2011, 17:1402-1409. 10.1038/nm.2541
3. Beaney T, Wang W, Schlaich MP, et al.: Global blood pressure screening during the COVID-19 pandemic: results from the May Measurement Month 2021 campaign. Journal of Hypertension. 2023, 41:1446-1455. 10.1097/hjh.0000000000003488
4. Kannel WB: Elevated systolic blood pressure as a cardiovascular risk factor. The American Journal of Cardiology. 2000, 85:251-255. 10.1016/s0002-9149(99)00635-9
5. Wandile PM: Hypertension and comorbidities: A silent threat to global health. Hypertension and Comorbidities. 2024, 1:1-7. 10.46439/Hypertension.1.001
6. Zhou B, Perel P, Mensah GA, Ezzati M: Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. Nature Reviews Cardiology. 2021, 18:785-802. 10.1038/s41569-021-00559-8
7. Berek PA, Irawati D, Hamid AYS: Hypertension: A global health crisis. Annals of Clinical Hypertension. 2021, 5:008-011. 10.29328/journal.ach.1001027
8. Kumar K, Kumar P, Deb D, Unguresan M-L, Muresan V: Artificial intelligence and machine learning based intervention in medical infrastructure: a review and future trends. Healthcare. 2023, 11:207. 10.3390/healthcare11020207
9. López-Martínez F, Schwarcz A, Núñez-Valdez ER, García-Díaz V: Machine learning classification analysis for a hypertensive population as a function of several risk factors. Expert Systems with Applications. 2018, 110:206-215. 10.1016/j.eswa.2018.06.006
10. Golino HF, de Brito Amaral LS, Duarte SFP, Gomes CMA, de Jesus Soares T, dos Reis LA, Santos J: Predicting increased blood pressure using machine learning. Journal of Obesity. 2014, 2014:637635. 10.1155/2014/637635
11. Fitriyani NL, Syafrudin M, Alfian G, Rhee J: Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. IEEE Access. 2019, 7:144777-144789. 10.1109/access.2019.2945129
12. Tjahjadi H, Ramli K, Murfi H: Noninvasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis. IEEE Access. 2020, 8:20735-20748. 10.1109/access.2020.2968967
13. Martinez-Ríos E, Montesinos L, Alfaro-Ponce M, Pecchia L: A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data. Biomedical Signal Processing and Control. 2021, 68:102813. 10.1016/j.bspc.2021.102813
14. Wang MC, Lloyd-Jones DM: Cardiovascular risk assessment in hypertensive patients. American Journal of Hypertension. 2021, 34:569-577. 10.1093/ajh/hpab021
15. Poznyak AV, Sadykhov NK, Kartuesov AG, Borisov EE, Melnichenko AA, Grechko AV, Orekhov AN: Hypertension as a risk factor for atherosclerosis: Cardiovascular risk assessment. Frontiers in Cardiovascular Medicine. 2022, 9:959285. 10.3389/fcvm.2022.959285
16. Chaikijurajai T, Laffin LJ, Tang WHW: Artificial intelligence and hypertension: recent advances and future outlook. American Journal of Hypertension. 2020, 33:967-974. 10.1093/ajh/hpaa102
17. Shin D: The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies. 2021, 146:102551. 10.1016/j.ijhcs.2020.102551
18. Tiwari R: Explainable AI (XAI) and its applications in building trust and understanding in AI decision making. International Journal of Scientific Research in Engineering and Management. 2023, 7:10.55041/ijsrem17592
19. Aman Z, Qidwai MA: Society, healthcare, and artificial intelligence: navigating the intersections of innovation and ethical considerations. Intersection of Human Rights and AI in Healthcare. Qidwai M (ed): IGI Global Scientific Publishing, Hershey, PA;

2025. 479-496. 10.4018/979-8-3693-7051-3.ch020

20. Vimbi V, Shaffi N, Mahmud M: Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. Brain Informatics. 2024, 11:10. 10.1186/s40708-024-00222-1

21. Elshawi R, Al-Mallah MH, Sakr S: On the interpretability of machine learning-based model for predicting hypertension. BMC Medical Informatics and Decision Making. 2019, 19:146. 10.1186/s12911-019-0874-0

22. El-Dahshan ESA, Bassiouni MM, Khare SK, Tan RS, Acharya UR: ExHyptNet: An explainable diagnosis of hypertension using EfficientNet with PPG signals. Expert Systems with Applications. 2024, 239:122388. 10.1016/j.eswa.2023.122388

23. Hypertension Risk Severity Data Set. Kaggle. (2026). https://www.kaggle.com/datasets/saviourjoshuainyang/hypertension-risk-severity-data-set.

24. Umoren I, Abe O, Ansa G, Inyang S, Umoh I: A new index for intelligent classification of early syndromic of cardiovascular (CVD) diseases based on electrocardiogram (ECG). European Journal of Computer Science and Information Technology. 2023, 11:1-21. 10.37745/ejcsit.2013/vol11n4121

25. Ajagbe SA, Awotunde JB, Florez H: Intrusion detection: a comparison study of machine learning models using unbalanced dataset. SN Computer Science. 2024, 5:1028. 10.1007/s42979-024-03369-0

26. Doshi-Velez F, Kim B: Towards a rigorous science of interpretable machine learning. arXiv. 2017, 10.48550/arXiv.1702.08608

27. Ribeiro MT, Singh S, Guestrin C: 'Why should i trust you?' Explaining the predictions of any classifier. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, 1135-1144. 10.1145/2939672.2939778

28. Lundberg S, Lee SI: A unified approach to interpreting model predictions. arXiv. 2017, 10.48550/arXiv.1705.07874

29. Shishehbori F, Awan Z: Enhancing cardiovascular disease risk prediction with machine learning models. arXiv. 2024, 10.48550/arXiv.2401.17328

30. Umoh I, Essien V, Inyang S: Optimizing hypertension risk classification through machine learning. International Journal of Computer Applications. 2024, 186:21-29. 10.5120/ijca2024923500

31. Talaat FM, Elnaggar AR, Shaban WM, Shehata M, Elhosseini M: CardioRiskNet: a hybrid AI-based model for explainable risk prediction and prognosis in cardiovascular disease. Bioengineering. 2024, 11:822. 10.3390/bioengineering11080822

32. Shah STH, Shah SAH, Khan II, et al.: Data-driven classification and explainable-AI in the field of lung imaging. Frontiers in Big Data. 2024, 7:1393758. 10.3389/fdata.2024.1393758

33. Sekaran K, Zayed H: Identification of novel hypertension biomarkers using explainable AI and metabolomics. Metabolomics. 2024, 20:124. 10.1007/s11306-024-02182-3

34. Napa KK, Govindarajan R, Sathya S, Murugan JS, Vijayammal BKP: Comparative analysis of explainable machine learning models for cardiovascular risk stratification using clinical data and shapley additive explanations. Intelligence-Based Medicine. 2025, 12:100286. 10.1016/J.IBMED.2025.100286

35. Okpono J, Asedegbega J, Sanyaolu T: Enhancing the predictive accuracy of machine learning models for project management challenges: a comparative study of advanced techniques for handling imbalanced data and improving model interpretability. TIJER - International Research Journal. 2024, 11:a107-a128.

36. Ganie SM, Pramanik PKD, Zhao Z: Ensemble learning with explainable AI for improved heart disease prediction based on multiple datasets. Scientific Reports. 2025, 15:13912. 10.1038/s41598-025-97547-6

37. Yang Y, Liao CY, Keyvanshokooh E, Shao H, Weber MB, Pasquel FJ, Garcia GGP: A responsible framework for assessing, selecting, and explaining machine learning models in cardiovascular disease outcomes among people with type 2 diabetes: methodology and validation study. JMIR Medical Informatics. 2025, 13:e66200.

38. Al Younis SM, Ghosh SK, Raja H, Alskafi FA, Yousefi S, Khandoker AH: Prediction of heart failure risk factors from retinal optical imaging via explainable machine learning. Frontiers in Medicine. 2025, 12:1551557. 10.3389/fmed.2025.1551557

How to cite this article:
Inyang S, Njoku F, Umoren I, et al. (June 09, 2026) A Comparative Study of Machine Learning Models for Hypertension Risk Assessment With Explainable Artificial Intelligence Integration. Cureus J Comput Sci 3 : es44389-026-00111-0. DOI https://doi.org/10.7759/s44389-026-00111-0