FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 1, January, 2026, pp 269 – 279
DOI: https://doi.org/10.33003/fjs-2026-1001-4543

MINING CLINICAL DATA FOR HYPERTENTION PREDICTION PERFORMANCE EVALUATION OF SIX SUPERVIVED LEARNING ALGORITHMS

Awuza Abdulrashid Egwa, Abdullahi Musa Bello, Sikirat Kehinde Aina and Saadatu Gimba

Department of Computer Science, Federal University Gashua, Yobe State, Nigeria.

*Corresponding authors' email: awuza@fugashua.edu.ng

ABSTRACT

Machine learning classifiers usage in medicine has been on a rise over the years due to increase in available clinical data. Hypertension, one of the cardiovascular heart diseases is a global health concern with several risk factors where early detection could prevent related heart diseases complications. With abundant clinical data piling within our health facilities here in Nigeria, little or non-research on hypertension prediction using local dataset has been done. This research work is aimed at developing models for predicting individuals with likelihood of hypertension even without necessary going through medical procedures. A sample of 294 clinical data from Specialist Hospital Gashua was used. Random Forest, Decision tree, Artificial neural network, Support vector machine, Logistic regression and Naïve bayes algorithms were used for classification. To minimised issues of over-fitting due to the number of available data, ADASYN balancing technique was used for data balancing. Using Root Means Square Error (RMSE), Accuracy, Recall, F1-score and ROC values as parameter matrix for evaluating the models, our results showed that machine learning can be utilised in predicting hypertension diseases for quick preventive measures. Furthermore, this research work further illustrate that Random Forest has proven to be the most efficient and best algorithms for hypertension prediction having performance supremacy over all others five algorithms with 92% prediction accuracy, 100% Precision, Recall values of 91%, 91% F1-score value, RMSE of 29% and AUC value of 97%. Thereby, stressing the significant of balancing dataset for enhanced model accuracy (and all other metrics).

Keywords: Hypertension, Imbalance data, Machine Learning, Prediction model, Supervised learning

INTRODUCTION

In the domain of artificial intelligence (AI) technology, the words Machine learning (ML) implies algorithms with advanced features and characteristics of supervised learning, unsupervised learning and re-enforcement for data related problem solving (Alhadeethy, Zeki & Shah, 2021). These algorithms are "trained" on specific datasets to generate mathematical models capable of generalizing to unseen data. ML has evolved into a valuable clinical tool, owing to advances in computational science that enable accurate modelling of complex diagnostic relationships (Sengupta et al., 2016). Compared to traditional reasoning based methods, ML models offer advantages in processing multidimensional clinical information with minimal bias (Mazzanti et al., 2018). The recent surge in medical data availability has further enhanced ML's potential in disease prediction and early diagnosis.

World Health Organization (WHO) revealed that hypertension remains the major global health issues, affecting over 1.1 billion persons, with the majority residing in low or middle-income countries (Haldar, 2013). It significantly enhances the possibility of heart attack (Fuchs & Whelton, 2020), heart failure (Kannan & Janardhanan, 2014), coronary artery disease, stroke, renal failure, and diabetes (Martinez-Rios et al., 2021). Approximately 45% death cause by cardiovascular issues and 51% death cause by stroke globally are attributed to elevated blood pressure (Haldar, 2013). Despite the availability of standard diagnostic implements such as sphygmomanometers, hypertension management remains costly and challenging for many individuals.

Medical studies emphasize that timely diagnosis, lifestyle modification, and adherence to treatment can substantially reduce hypertension's morbidity and mortality. However, because many patients are asymptomatic until advanced disease stages, hypertension is often described as a "silent killer" (Martinez-Rios, 2021). When present, symptoms may include headaches, dizziness, palpitations, chest pain, and nosebleeds (Ye et al., 2018). These challenges underscore the importance of predictive and monitoring systems, especially in developing countries like Nigeria, where proactive healthcare infrastructure remains limited.

Hypertension prediction using ML has been explored primarily through two methodological paradigms. The first treats blood pressure estimation as a regression task, using continuous signals such as photoplethysmography (PPG) and electrocardiography (ECG) to predict systolic and diastolic values (Martinez-Rios et al., 2021; Faul, 2019; Dev, Gaurav & Tiwari, 2018). The second approach formulates the problem as classification, estimating discrete hypertension categories based on clinical and demographic features (Islam et al., 2022). While regression models rely on raw physiological signals, classification models depend on structured clinical data.

A critical gap in existing studies is the scarcity of research utilizing locally generated datasets. Most prior works employ foreign data sources, limiting contextual relevance for Nigeria's population. This research aimed to develop models for hypertension prediction using locally sourced clinical and demographic datasets. It explores multiple supervised ML algorithms including Random forest (RF), Artificial neural network (ANN), Support vector machine (SVM), Logistic regression (LR), Naïve bayes (NB), and Decision tree (DT) just as in the work of (Chang et al., 2019; Nour & Polat, 2020; Tjahjadi et al., 2019; Alkaabi et al., 2020). Following Chang et al. (2019), Recursive Feature Elimination with Cross-Validation (RFECV) is employed to identify optimal predictors of hypertension outcomes.

To ensure robust and unbiased performance, this study applies the Adaptive Synthetic Sampling (ADASYN) technique (Chandramohan, 2022) to address class imbalance in the dataset. Model performances before and after balancing are

compared to assess interpretability and overall predictive reliability.

Related Literature

This review is organized in two parts: first, an investigation of classification based ML approaches applied for hypertension classification; and second, a discussion of dataset utilization, particularly those developed within Nigeria.

Classification algorithms have proven effective in predicting hypertension. López et al. (2018) employed LR on U.S. National Health Repository data, their model achieved 0.8977 AUC and 0.8023 accuracy. Similarly, Ye et al. (2018) proposed a risk estimation model using patients' one-year medical records, applying K-nearest neighbor (KNN) for imputing missing values and XGBoost for classification, resulting in AUC values of 0.870 and 0.917 for prospective and retrospective evaluations, respectively. Kwong et al. (2018) developed telemedicine-oriented models using multilayer perceptron and radial basis function neural networks, attaining accuracies of 94.28% and 91.06%, respectively.

Fitriyani et al. (2019) designed ensemble models integrating MLP, SVM, DT, and LR algorithms for predicting hypertension and type II diabetes using the Golinos dataset, achieving 75.78% accuracy for prehypertension and 85.73% for hypertension. Chang et al. (2019) constructed an outcome-based prediction model employing physical examination data and demonstrated that XGBoost, optimized with recursive feature elimination, achieved superior results (accuracy 94.3%, F1 = 0.875, AUC = 0.927). Leha et al. (2019) examined multiple ML algorithms for echocardiographic pulmonary hypertension diagnosis, with random forest regression achieving the best performance (AUC = 0.87, 95% CI 0.78–0.96).

In a large-scale cohort, Marin and Goga (2019) compared several classifiers RF, neural network, SVM, GBM, and KNN against LR across 56,762 Asian adults. LR achieved the highest AUC for CKD (0.905) and DM (0.768), while SVM and NN performed best for CVD (0.753) and HTN (0.780), though differences with LR were minimal. Nour and Polat (2020) applied PPG-BP datasets to C4.5, RF, LDA, and SVM classifiers, yielding accuracies between 92.7% and 99.5%.

Deep learning techniques have also shown promise in hypertension prediction. López et al. (2020) enhanced their earlier LR model using a multilayer perceptron, improving AUC and specificity. Tjahjadi et al. (2020) applied bidirectional LSTM networks with time-frequency feature extraction, achieving accuracy, sensitivity, and specificity values of 97.33%, 100%, and 94.87%, respectively. Their subsequent work using KNN achieved perfect F1-scores for normotensive and prehypertensive groups and 90.8% for hypertensive patients.

Alkaabi et al. (2020) analyzed 987 Qatari records using RF, LR, and DT algorithms, all performing comparably, with RF attaining the highest AUC (86.9%). Zhang et al. (2020) developed neural networks trained on retinal fundus images, achieving AUCs of 0.880 (hyperglycaemia), 0.766 (hypertension), and 0.703 (dyslipidaemia), confirming the potential of image-based hypertension prediction.

Islam et al. (2022) applied ML to data from Bangladesh, Nepal, and India, identifying age and BMI as key hypertension determinants. XGBoost, GBM, LR, and LDA models achieved accuracy up to 90%, with XGBoost performing best overall. Martinez-Rios et al. (2022) combined PPG features derived from wavelet scattering transforms with clinical data via early and late fusion. Although SVM achieved 71.42% accuracy and a 76% F1-score, the fusion approaches did not outperform separate feature analyses.

Finally, Fang et al. (2023) integrated KNN and LightGBM to predict five-year hypertension risk, obtaining 86% accuracy and 92% recall, highlighting the utility of hybrid ML models in early hypertension detection and prevention.

MATERIALS AND METHODS

Review of Dataset

Various machine learning algorithms have been applied in the development of hypertension prediction models, as highlighted in previous studies. This section reviews the most commonly utilized datasets and their characteristics to identify frequently used variables.

Among publicly accessible resources, the Medical Information Mart for Intensive Care (MIMIC) datasets is the most extensively employed. These datasets provide comprehensive clinical data suitable for predictive modelling but require formal access requests and ethical review prior to use (Johnson et al., 2016). Another widely used dataset is the PPG-BP dataset, available through the Figshare repository (Liang et al., 2018). Although the PPG-BP dataset is well-structured and requires minimal pre-processing, it presents limitations related to class imbalance and small sample size, containing only 219 instances. This restricts its suitability for training deep learning models, which typically demand larger datasets to prevent over-fitting.

Design Flow Diagram

The design flow of the propose model shows sequence of activities performed by the model. It shows the sample data as input to the pre-processing stage which involves data cleaning, data transformation and then features selection to select most important features for the model, then the learning and testing stages followed by model evaluation as shown in Figure 1.

Data Collection

The dataset used in this study was obtained from Specialist Hospital Gashua, Yobe State, Nigeria. It comprises both demo-graphic and clinical attributes of individual relevant to hypertension risk assessment. Sixteen variables were recorded, including Gender, Age, Marital Status, Family History of Hypertension, Blood Pressure Medication, Employment Status, Smoking Status, Smoking Rate per Day, Diabetes Status, Total Cholesterol Level, Systolic Blood Pressure, Heart Rate, Diastolic Blood Pressure, Body Mass Index (BMI), Glucose Level, and Risk Classification. These variables collectively represent key predictors of hypertension and associated cardiovascular conditions.

Dataset Description

Data were collected manually from the hospital's archived clinical records and subsequently digitized for analysis. All variables were encoded numerically to facilitate machine learning model training and evaluation. The categorical (discrete) variables include Gender, Marital Status, Family History of Hypertension, Employment Status, Smoking Status, Blood Pressure Medication, Diabetes Status, and Risk Classification. The continuous variables comprise Age, Smoking Rate per Day, Total Cholesterol Level, Systolic Blood Pressure, Diastolic Blood Pressure, BMI, Glucose Level, and Heart Rate.

This structured encoding enabled uniform data representation and ensured compatibility with supervised machine learning algorithms for hypertension prediction.

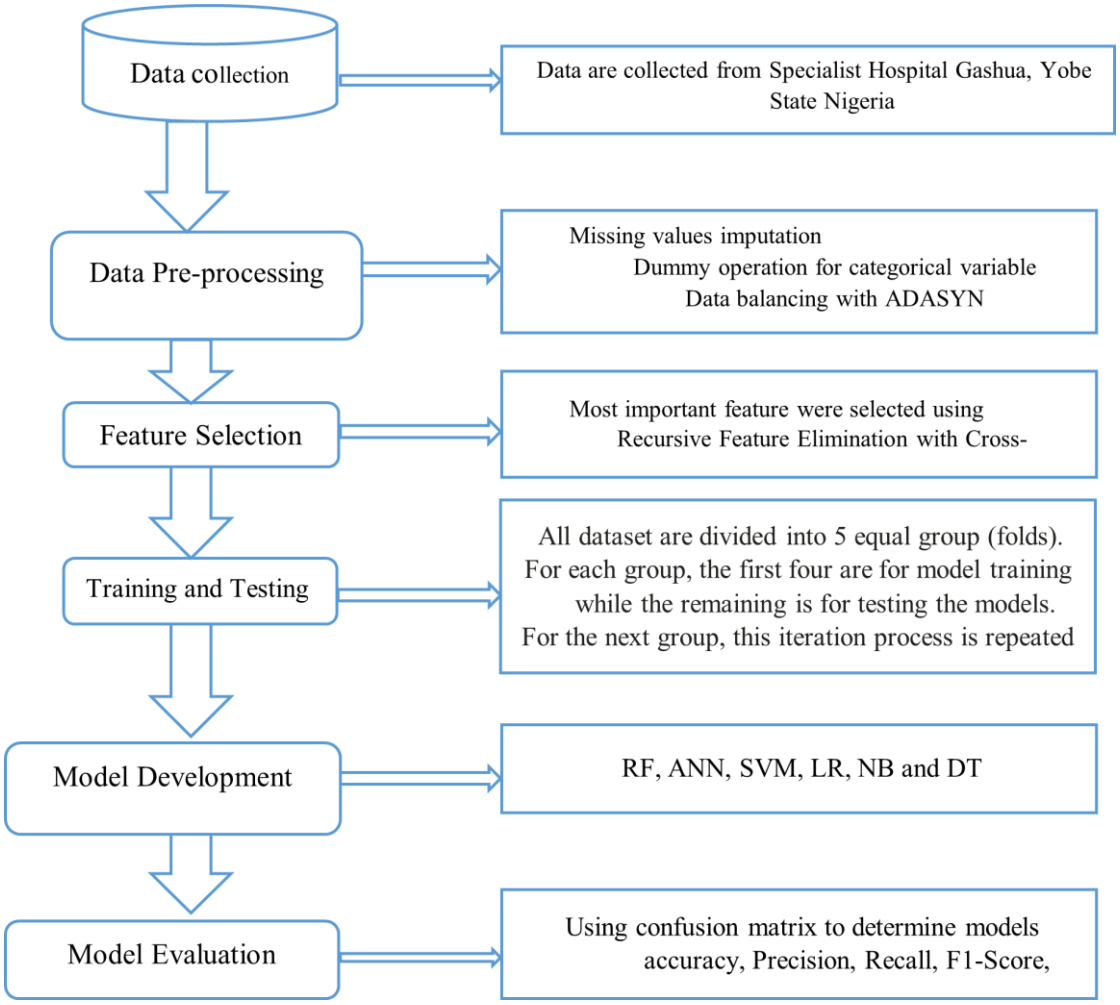


Figure 1: Proposed System Design Workflow

Table 1: Dataset Description

S/N	Variable	Description	Type
1	Gender	Male = 1, Female = 0	Categorical (Numeric)
2	Age	Age of the patient	Continuous (Numeric)
3	Married	Marital Status: 1=Married, 0= Not Married	Categorical (Numeric)
4	FHHP	Family history of Hypertension: 1=Yes, 0=No	Categorical (Numeric)
5	Smoker_ST	Smoking Status: 1=Yes, 0=No	Categorical (Numeric)
6	CigPerDay	Number of cigarette smoke per day	Continuous (Numeric)
7	BpMeds	Blood pressure medication: 1= Yes, 0= No	Categorical (Numeric)
8	Diabetes	Diabetes status: 1= Yes, 0= NO	Categorical (Numeric)
9	Emp_Status	Employment status: 1= Employed, 0= Not Employed	Categorical (Numeric)
10	Totcho	Total cholesterol level	Continuous (Numeric)
11	SysBp	Systolic blood pressure	Continuous (Numeric)
12	DiaBp	Diastolic blood pressure	Continuous (Numeric)
13	BMI	Body Mass Index	Continuous (Numeric)
14	HeartRate	Heart rate	Continuous (Numeric)
15	Glucose	Glucose level	Continuous (Numeric)
16	Risk	Hypertension Risk status: 1=Yes, 0=No	Target (Numeric)

Data Pre-processing

Data pre-processing is an important step in preparing datasets for ML applications. It ensures data quality, consistency, and suitability for model training. Since the dataset in this study was collected from a single clinical source (Specialist Hospital Gashua) no major inconsistencies such as mismatched data types or mixed categorical entries (e.g., "male" vs. "man") were observed. The dataset was examined for missing values and potential outliers. In addition, a correlation analysis was conducted to determine the significance and interdependence of features relative to the target variable (*Risk*).

Feature Selection

Feature selection involves reducing the number of attributes in a dataset to retain only those that significantly influence the

predictive outcome. This process enhances model efficiency, interpretability, and accuracy by removing redundant or irrelevant features. According to Chandrashekar and Sahin (2014), feature selection methods are generally categorized into filter**,** wrapper**,** and embedded techniques.

For this work, Recursive Feature Elimination with Cross-Validation (RFECV) was used to identify and retain the most relevant features for hypertension prediction (Chandramohan, 2022). This approach iteratively eliminates less significant variables and evaluates model performance at each step, ensuring optimal feature subset selection for robust model training.

Adopted Algorithms

This research implemented six supervised ML algorithms, consisting of two tree based models; RF and DT and four non tree based models; ANN, SVM, NB and LR.

The DT algorithm employs a hierarchical tree structure where internal nodes represent decision conditions based on dataset features, and leaf nodes indicate classification outcomes. Each node evaluates a condition and splits the data according to the feature's information gain, ultimately enhancing classification accuracy (Zhou et al., 2021).

The RF algorithm, an ensemble of multiple decision trees, combines the bagging technique with random subspace selection. It aggregates results from several independent estimators to improve generalization and reduce over-fitting (Breiman, 2001).

The ANN mimics the behaviour of biological neurons, enabling nonlinear and complex pattern learning. It consists of interconnected layer; input, hidden, and output where neurons transmit weighted signals. In this study, a three layer ANN architecture was adopted, comprising an input layer that receives predictors, a hidden layer for feature processing, and an output layer for classification (Zhang et al., 2020).

The SVM constructs an optimal hyper-plane that separates data points into distinct classes with maximal margin. It performs effectively on linearly separable data and can be extended to higher dimensional feature spaces using kernel transformations (Zhou, 2016).

NB algorithm is a probabilistic classifier based on Bayes' theorem**,** which computes posterior probabilities of class membership and assigns observations to the class with the highest likelihood (Dugan et al., 2015).

Lastly, LR**,** a generalized linear model, estimates the relationship between dependent and independent variables through a logistic function. It is widely used in medical prediction studies for its interpretability and robustness in binary classification (Bagley et al., 2001).

Adaptive Synthetic Sampling (ADASYN)

A common challenge in ML modelling is class imbalance**,** where one category significantly outnumbers another (Brennan, 2018). Models trained on imbalanced datasets tend to be biased toward the majority class, resulting in poor predictive performance for minority classes. Previous studies have addressed this issue using Synthetic Minority Over-sampling Technique (SMOTE) (Fitriyani et al., 2019). However, SMOTE does not account for varying data densities and may generate overlapping or noisy samples near class boundaries.

To overcome these limitations, this study applied the Adaptive Synthetic Sampling (ADASYN) algorithm, an advanced variant of SMOTE. ADASYN automatically determines the number of synthetic samples required for each minority instance based on data density distribution. It generates more synthetic samples near complex or sparsely represented regions, enabling the model to focus on harder-to-learn examples. Randomized perturbations are added to the generated points to enhance sample realism and reduce over-fitting.

Consequently, the dataset balanced using ADASYN provided a more representative data distribution, improving model robustness, interpretability, and classification performance compared to traditional resampling techniques.

RESULTS AND DISCUSSION

Implementation Results and Findings

To evaluate model performance and generalizability, a 5-fold cross-validation technique was employed during the training and testing phases. This method assesses how well models perform on unseen data, following the approaches of Alkaabi et al. (2020) and Fang et al. (2021). In this approach, the entire dataset was randomly partitioned into five equal subsets (folds)**.** Four folds were used for model training in each iteration, while the remaining fold served as the testing set. The process was repeated five times, with each fold used exactly once as the test set. The performance metrics obtained from all iterations were then averaged to yield a more reliable estimate of the model's predictive accuracy.

The adoption of cross-validation helped prevent over-fitting**,** ensuring that the models generalized well to new, unseen data rather than merely memorizing the training patterns. Furthermore, this approach reduced evaluation bias and facilitated effective hyper-parameter optimization**,** leading to improved model robustness and reliability.

Confusion Matrix Evaluation

In evaluating hypertension prediction models, it is essential to visualize the classification outcomes using a confusion matrix. A confusion matrix is a square matrix that compares the actual classes (displayed on the vertical axis) with the predicted classes (displayed on the horizontal axis). In this study, instances predicted as having a likelihood of hypertension are treated as positive cases, while those without likelihood are negative cases.

The TP, FP, TN, and FN values derived from the confusion matrix are subsequently used to compute key evaluation metrics such as accuracy, precision, recall, and F1-score, RMSE and AUC value as illustrated in Figure 2 and Figure 3.

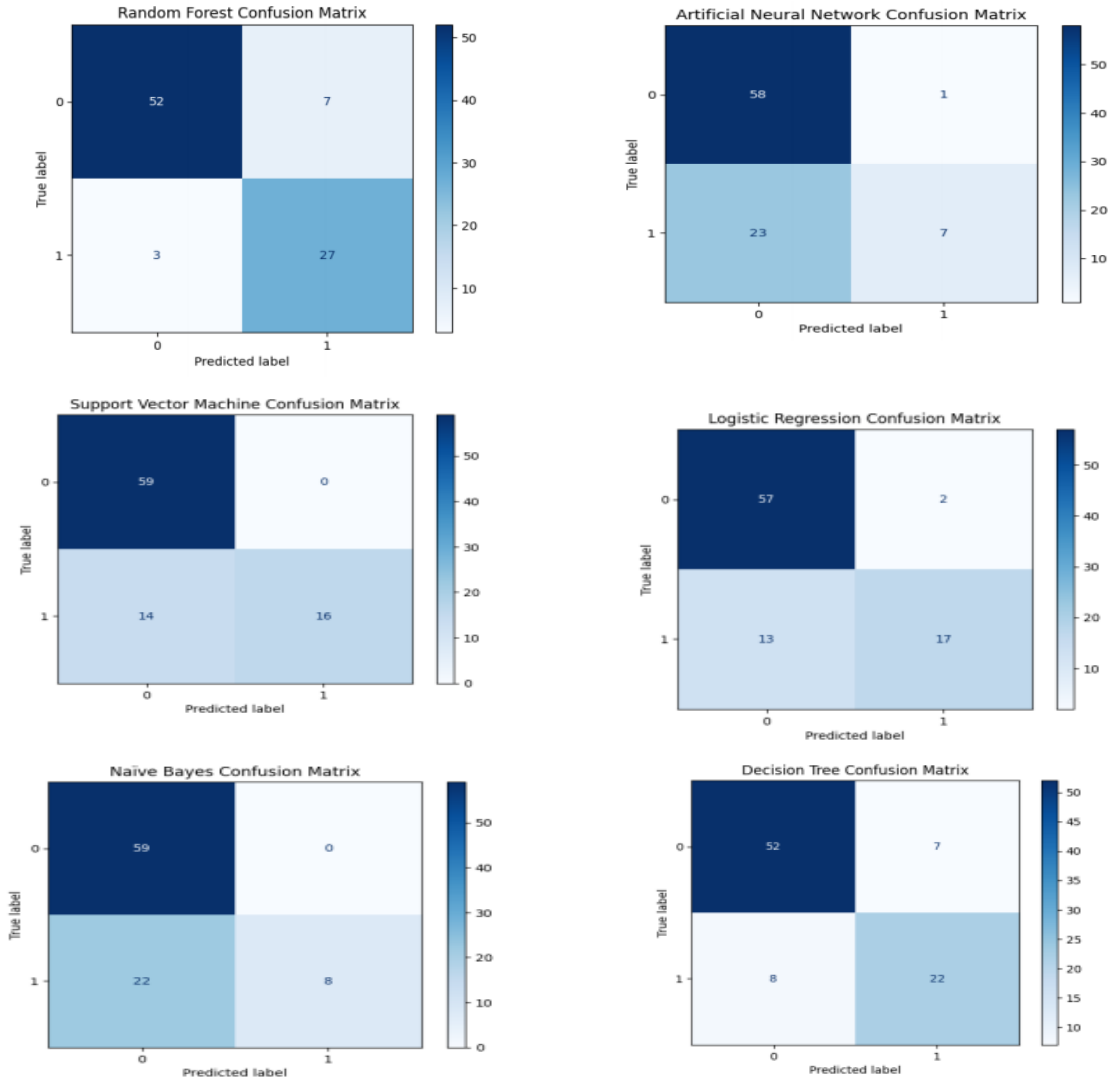
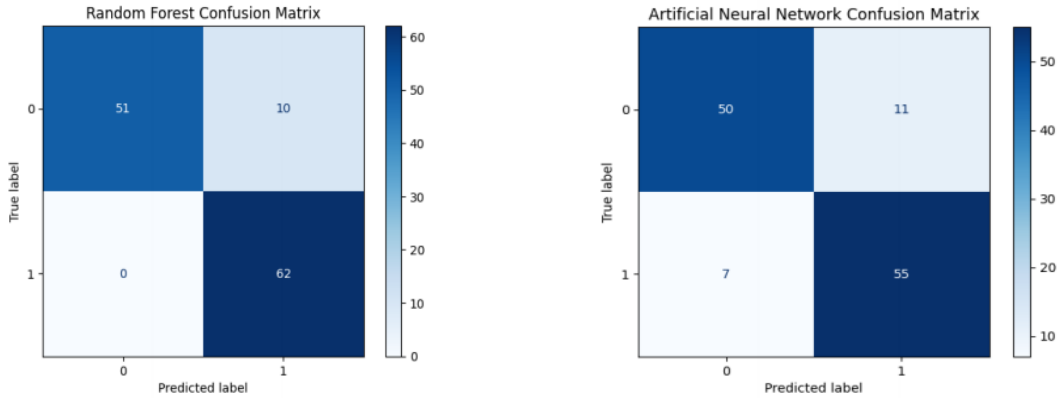


Figure 2: Models Confusion Matrix (Before Data Balancing)



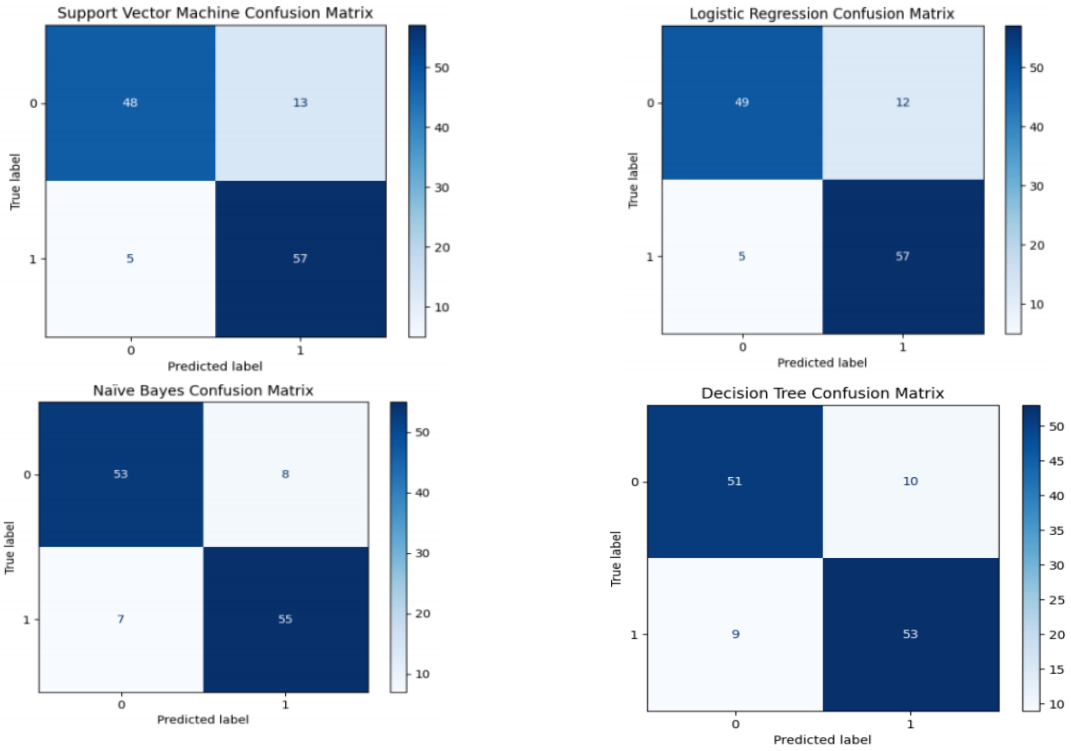


Figure 3: Models Confusion Matrix (After Data Balancing**)**

The results of the models performances evaluation are presented in Table 2 and Table 3 with same as a bar chart plot in Figure 4 and Figure 5. Other parameter matrix used models performance evaluation is the ROC curve in obtaining the Area Under the curve (AUC). The AUC varied within 0 and 1 and is presented here in Figure 4 and Figure 5.

Table 2: Models Evaluation Performance (Before Data Balancing)

S/N	Algorithms	Accuracy	Precision	Recall	F1-Score	RMSE	AUC
1	RF	0.89	0.95	0.88	0.91	0.34	0.95
2	ANN	0.73	0.72	0.78	0.83	0.52	0.84
3	SVM	0.84	0.81	0.98	0.89	0.40	0.95
4	LR	0.85	0.81	0.97	0.88	0.41	0.94
5	NB	0.75	0.73	0.71	0.84	0.50	0.95
6	DT	0.83	0.87	0 .88	0.87	0.41	0.81

Table 3: Models Performance Evaluation (After Data Balancing)

S/N	Algorithms	Accuracy	Precision	Recall	F1-Score	RMSE	AUC
1	RF	0.92	1.0	0.91	0.91	0.29	0.97
2	ANN	0.85	0.88	0.82	0.85	0.38	0.94
3	SVM	0.85	0.91	0.79	0.84	0.38	0.93
4	LR	0.86	0.91	0.80	0.85	0.37	0.94
5	NB	0.88	0.87	0.87	0.88	0.35	0.96
6	DT	0.85	0.85	0.84	0.84	0.39	0.85

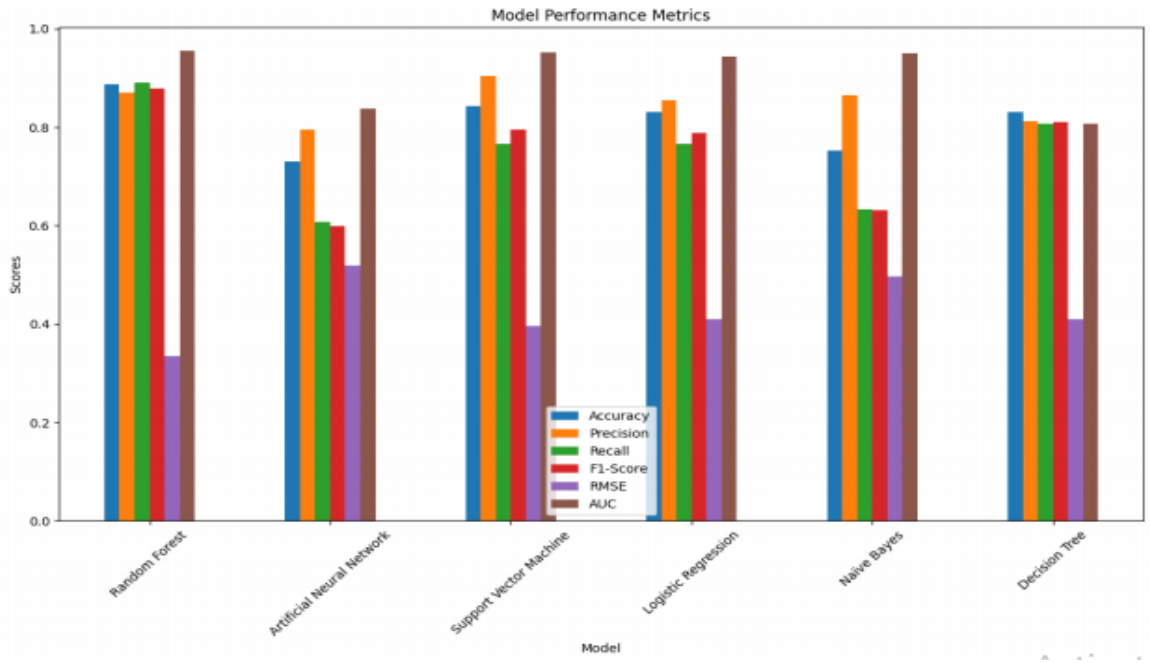


Figure 4: Bar Chart for Models Performances (Before Data Balancing)

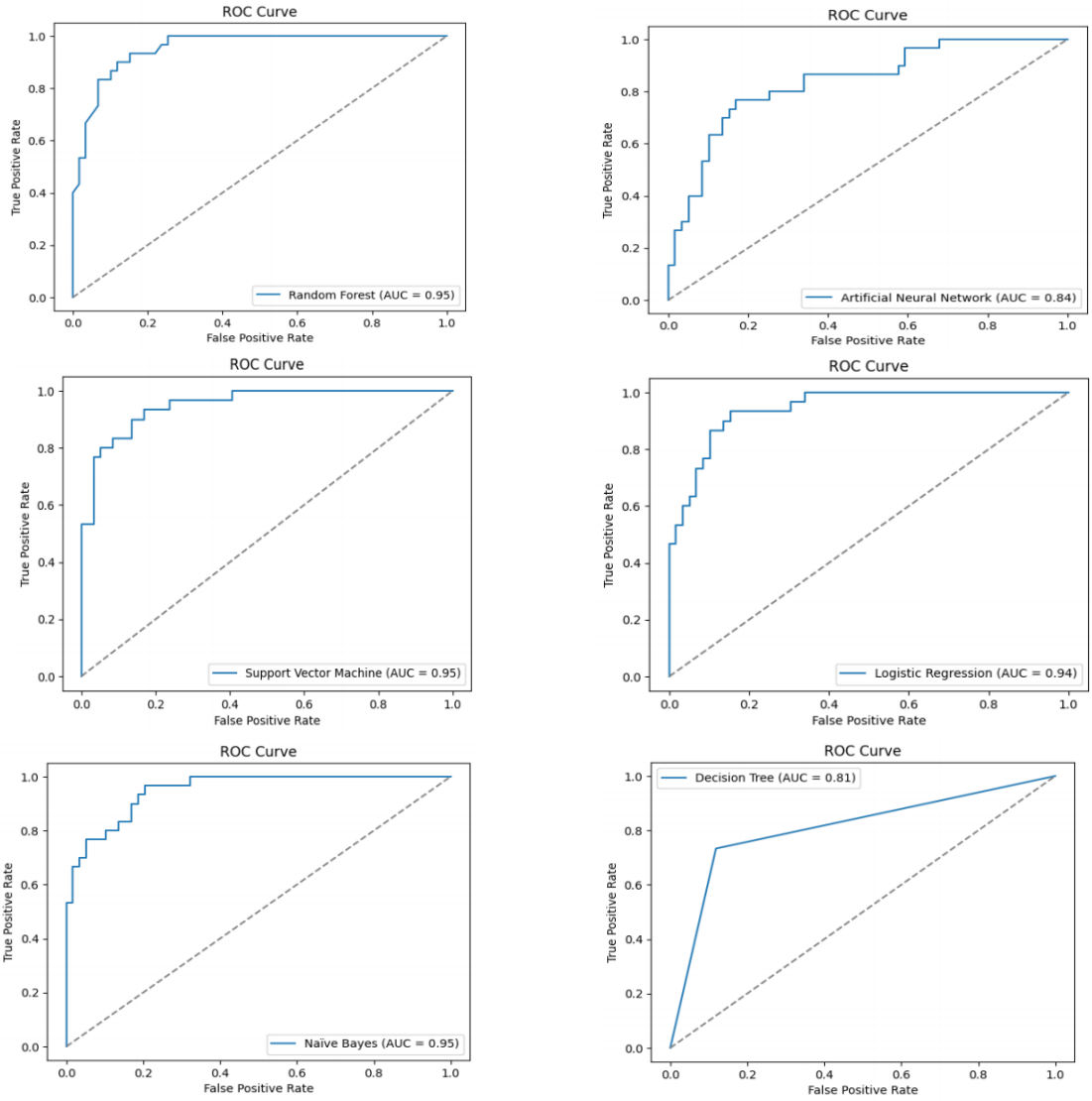


Figure 5: ROC Plot for the Models (Before Data Balancing)

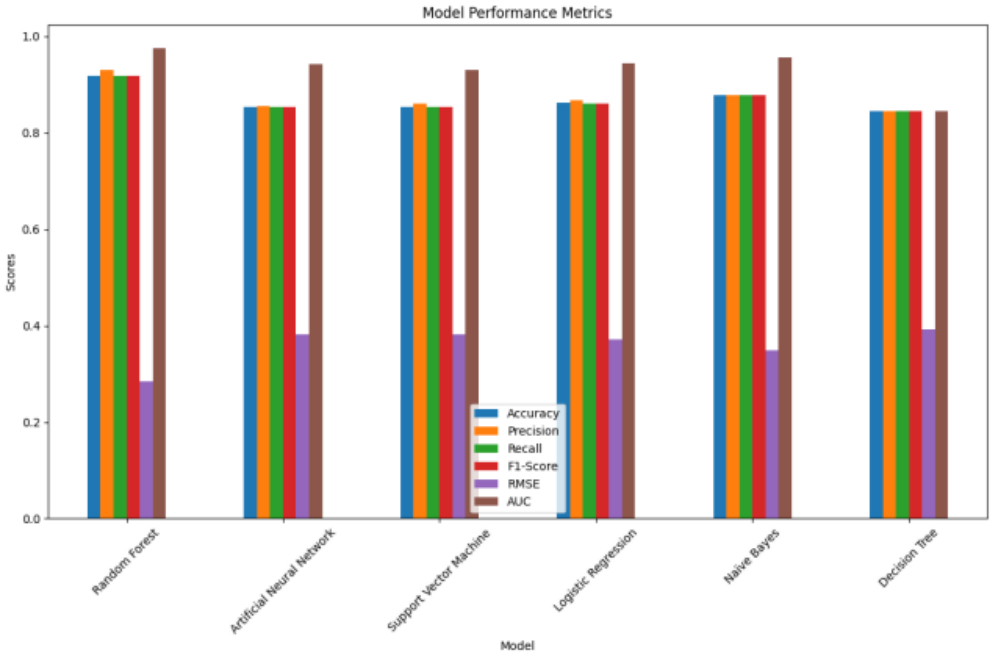


Figure 6: Bar Chart for Models Performances (After Data balancing)

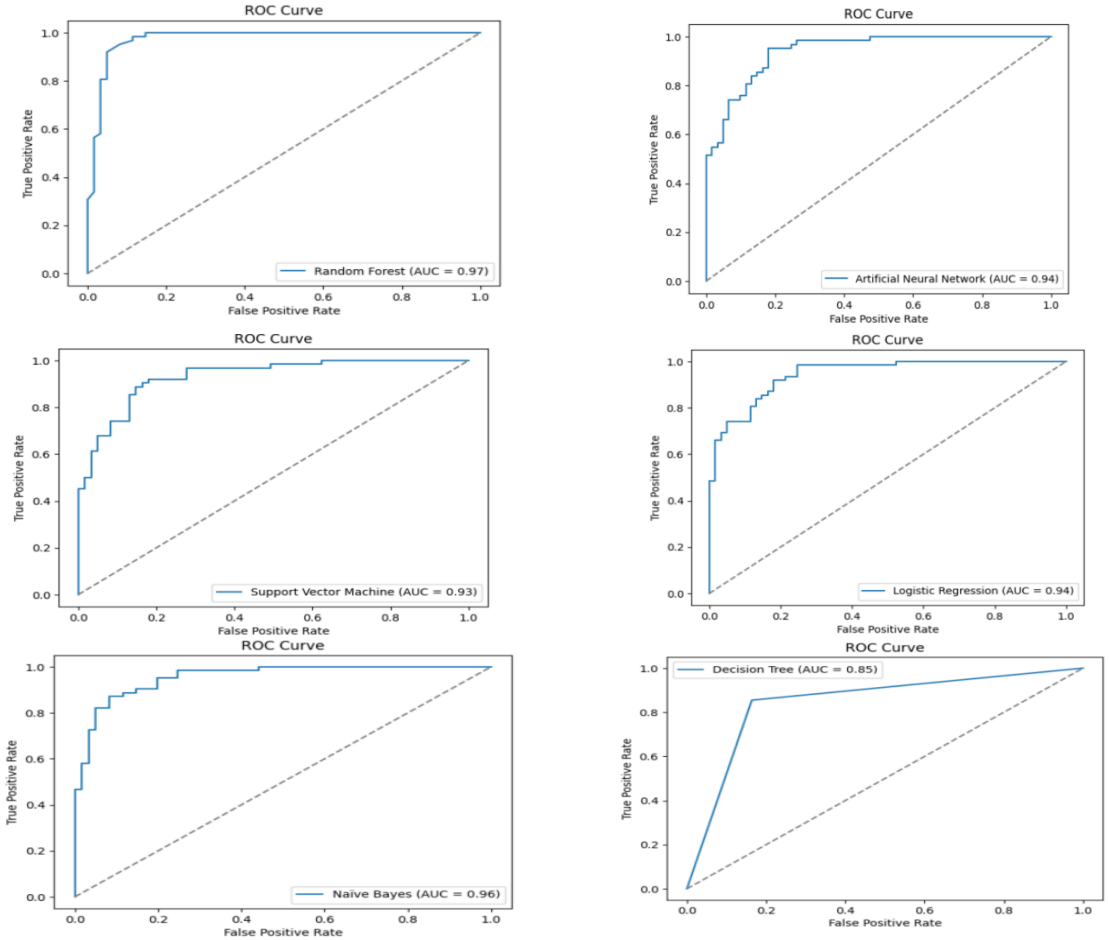


Figure 7: ROC Plot for the Models After Data Balancing

Discussion

Table 2 presents the performance of all models prior to data balancing. The RF model achieved the highest of 89% accuracy, 95% precision, 88% recall, 91% F1-score, 34% RMSE, and 95% AUC. The ANN yielded an accuracy of 73%, precision of 72%, recall of 78%, F1-score of 83%, RMSE of 52%, and AUC of 84%. The Support Vector Machine (SVM) attained an accuracy of 84%, precision of

81%, recall of 98%, F1-score of 89%, RMSE of 40%, and AUC of 95%. Similarly, the LR model produced an accuracy of 85%, precision of 81%, recall of 97%, F1-score of 88%, RMSE of 41%, and AUC of 94%. The NB model achieved 75% accuracy, 73% precision, 71% recall, 84% F1-score, 50% RMSE, and 95% AUC. Finally, the DT model recorded 83% accuracy, 87% precision, 88% recall, 87% F1-score, 41% RMSE, and 81% AUC. These results are illustrated in the bar chart (Figure 6) and ROC curve (Figure 7).

Building machine learning models with relatively small datasets can often result in over-fitting, under-fitting, or limited generalizability (Ramampiandra et al., 2023). Given the modest dataset size in this study, and unlike many existing works that neglect data balancing (Nour & Polat, 2020), the Adaptive Synthetic Sampling (ADASYN) technique was employed to mitigate class imbalance and enhance model robustness.

Following the application of ADASYN and re-evaluation using the confusion matrix (Figure 3), the updated results are presented in Table 3. Consistent with the findings of Fitriyani et al. (2019), all classifiers demonstrated notable improvement across all performance metrics.

Post-balancing, the RF model achieved the highest overall performance, with accuracy of 92% (5%), precision of 100% (+5%), recall of 91% (+1%), F1-score of 91% (+7%), RMSE of 29% (−5%), and AUC of 97% (+2%). The ANN model improved to 85% accuracy (+12%), 88% precision (+16%), 82% recall (+4%), 85% F1-score (+2%), 38% RMSE (−14%), and 94% AUC (+10%). The SVM model recorded 85% accuracy (+1%), 91% precision (+10%), 79% recall, 84% F1-score, 38% RMSE (−2%), and 94% AUC.

Similarly, the LR model yielded 86% accuracy (+5%), 91% precision (+10%), 80% recall, 85% F1-score, 37% RMSE (−4%), and 94% AUC. The NB model improved to 88% accuracy (+13%), 87% precision (+14%), 87% recall, 88% F1-score (+4%), 35% RMSE (−15%), and 96% AUC (+1%). Finally, the DT model recorded 85% accuracy (+2%), 85% precision, 84% recall, 84% F1-score, 39% RMSE (−2%), and 85% AUC (+4%).

These improvements collectively demonstrate that Random Forest remains the most effective algorithm for hypertension prediction, outperforming all others across key metrics. The results are consistent with those reported by (Nour & Polat, 2020), (Mahdi & Goga, 2025) and Islam et al. (2022). The enhanced performance is visualized in Figure 5 (bar chart) and Figure 6 (ROC curve).

Feature Ranking

To determine the relative influence of each variable on model performance, feature importance ranking was conducted. Using the Random Forest algorithm identified as the most effective classifier in this study, the importance of individual features contributing to hypertension prediction was evaluated. The ranking results, as displayed in Figure 8, highlight the most influential clinical and demographic predictors of hypertension within the dataset.

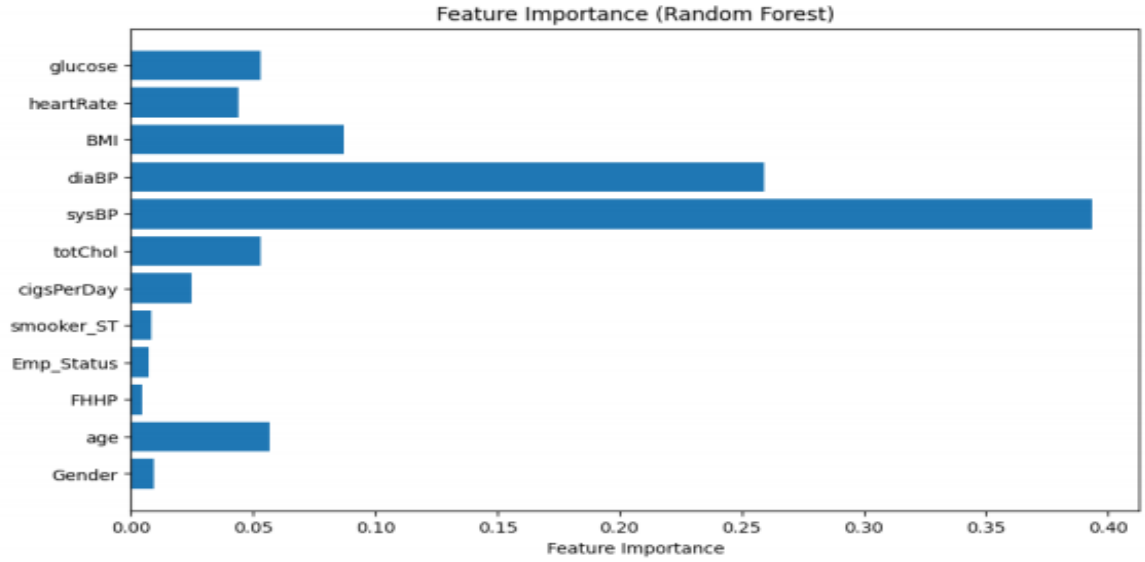


Figure 8: Features Ranking

Feature Importance Analysis

From the feature importance plot, Systolic Blood Pressure (SysBP) emerged as the most influential predictor of hypertension. SysBP represents the arterial pressure during cardiac contraction, with normal values typically below 140 mmHg**.** This was closely followed by Diastolic Blood Pressure (DiaBP)**,** which measures arterial pressure during cardiac relaxation. Both variables are physiologically central to hypertension risk assessment, hence their strong predictive contributions.

Next in importance was the Body Mass Index (BMI)**,** indicating that higher body mass is strongly associated with elevated hypertension risk. Glucose level**,** total cholesterol (TotChol)**,** and age followed as moderate contributors, reinforcing the established link between metabolic and cardiovascular health factors.

Conversely, features such as family history of hypertension (FHHP)**,** employment status**,** smoking status**,** and gender were ranked as the least significant in influencing the predictive performance of the models. This suggests that, within the studied population, clinical indicators play a more dominant role than socio-demographic factors in predicting hypertension risk.

CONCLUSION

This study developed and evaluated several ML models for the prediction of hypertension using a locally generated dataset from Specialist Hospital Gashua, Yobe State, Nigeria. Six classification algorithms RF, DT, ANN, SVM, LR, and

NB were implemented and assessed using Accuracy, Precision, Recall, F1-score, Root Mean Square Error (RMSE), and Area under the Curve (AUC) as performance metrics.

The results demonstrated that Random Forest consistently outperformed all other classifiers, achieving 92% accuracy, 100% precision, 91% recall, 91% F1-score, RMSE of 29%, and AUC of 97% after applying the Adaptive Synthetic Sampling (ADASYN) technique for data balancing. All models showed improved performance post-balancing, confirming the significant role of handling class imbalance in enhancing model generalization and predictive accuracy. Our findings affirm that ML techniques, particularly ensemble methods such as Random Forest, can serve as effective tools for early hypertension detection and risk assessment, reducing dependence on invasive diagnostic procedures.

RECOMMENDATIONS

i. Adoption in Clinical Decision Support Systems: Healthcare institutions should consider integrating ML-based predictive systems, such as Random Forest models, into diagnostic workflows for early screening and preventive care.
ii. Expansion of Local Datasets: Future research should focus on collecting larger and more diverse datasets across multiple healthcare centres in Nigeria to improve model robustness and external validity.
iii. Incorporation of Lifestyle and Genetic Factors: Including additional features such as diet, stress levels, and genetic predisposition may further enhance prediction accuracy.
iv. Model Interpretability: Further exploration of explainable AI (XAI) techniques is recommended to ensure that clinicians can interpret ML predictions with confidence.
v. Real-time Implementation: Developing mobile or web-based platforms that utilize these trained models could aid in remote health monitoring and community-level hypertension prevention.

Overall, this research underscores the potential of machine learning as a transformative approach to predictive healthcare, particularly in resource-constrained settings like Nigeria, where early detection can significantly reduce hypertension-related morbidity and mortality.

REFERENCES

Alhadeethy, N. F. A., Zeki, A. M., & Shah, A. (2021). Deep Learning Model For Predicting And Detecting Overlapping Symptoms Of Cardiovascular Diseases In Hospitals Of UAE. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(14), 5212-5224.

AlKaabi, L. A., Ahmed, L. S., Al Attiyah, M. F., & Abdel-Rahman, M. E. (2020). Predicting hypertension using machine learning: Findings from Qatar Biobank Study. *Plos one*, *15*(10), e0240370.

Al-Zaben, A., Fora, M., & Obaidat, A. (2018, March). Detection of premature ventricular beats from arterial blood pressure signal. In *2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME)* (pp. 17-19). IEEE.

Bagley SC, White H, Golomb B.A. (2001) Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *J Clin Epidemiol.* 54:979–85. https://doi.org/10.1016/s0895-4356(01)00372-9

Breiman L. 2001) Random forests, machine learning 45. *J Clin Microbiol.* (2:199–228. doi: 10.1023/A:1010933404324

Chandramohan, R. M. (2022). *Stroke Detection and Prediction Using Deep Learning Techniques and Machine Learning Algorithms* (Doctoral dissertation, Dublin, National College of Ireland).

Chang, W., Liu, Y., Xiao, Y., Yuan, X., Xu, X., Zhang, S., & Zhou, S. (2019). A machine-learning-based prediction method for hypertension outcomes based on medical data. *Diagnostics*, *9*(4), 178.

Chang, W., Liu, Y., Xiao, Y., Yuan, X., Xu, X., Zhang, S., & Zhou, S. (2019). A machine-learning-based prediction method for hypertension outcomes based on medical data. *Diagnostics*, *9*(4), 178.

Chandrashekar, G. & Sahin, F. A (2014) Survey on feature selection methods. *Comput. Electr. Eng.* https://doi.org/10.1016/j.compeleceng.2013.11.024

Dey, J., Gaurav, A., & Tiwari, V. N. (2018, July). InstaBP: cuff-less blood pressure monitoring on smartphone using single PPG sensor. In *2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 5002-5005). IEEE.

Fang, M., Chen, Y., Xue, R., Wang, H., Chakraborty, N., Su, T., & Dai, Y. (2023). A hybrid machine learning approach for hypertension risk prediction. *Neural Computing and Applications*, *35*(20), 14487-14497.

Faul, A. C. (2019). *A concise introduction to machine learning*. CRC Press.

Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2019). Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *Ieee Access*, *7*, 144777-144789.

Fuchs, F. D., & Whelton, P. K. (2020). High blood pressure and cardiovascular disease. *Hypertension*, *75*(2), 285-292.

Haldar, R. N. (2013). Global brief on hypertension: silent killer, global public health crisis. *Indian J Phys Med Rehabil*, *24*(1), 2.

Islam, S. M. S., Talukder, A., Awal, M. A., Siddiqui, M. M. U., Ahamad, M. M., Ahammed, B., ... & Maddison, R. (2022). Machine learning approaches for predicting hypertension and its associated factors using population-level data from three south asian countries. *Frontiers in Cardiovascular Medicine*, *9*, 839379.

Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L. W. H., Feng, M., Ghassemi, M., ... & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific data*, *3*(1), 1-9.

Kwong, E. W. Y., Wu, H., & Pang, G. K. H. (2018). A prediction model of blood pressure for telemedicine. *Health informatics journal*, *24*(3), 227-244.

LaFreniere, D., Zulkernine, F., Barber, D., & Martin, K. (2016, December). Using machine learning to predict

hypertension from a clinical dataset. In *2016 IEEE symposium series on computational intelligence (SSCI)* (pp. 1-7). IEEE.

Leha, A., Hellenkamp, K., Unsöld, B., Mushemi-Blake, S., Shah, A. M., Hasenfuß, G., & Seidler, T. (2019). A machine learning approach for the prediction of pulmonary hypertension. *PloS one*, *14*(10), e0224453.

Liang, Y., Chen, Z., Liu, G., & Elgendi, M. (2018). A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China. *Scientific data*, *5*(1), 1-7.

Liang, Y., Chen, Z., Ward, R., & Elgendi, M. (2018). Photoplethysmography and deep learning: enhancing hypertension risk stratification. *Biosensors*, *8*(4), 101.

López-Martínez, F., Núñez-Valdez, E. R., Crespo, R. G., & García-Díaz, V. (2020). An artificial neural network approach for predicting hypertension using NHANES data. *Scientific Reports*, *10*(1), 10620.

López-Martínez, F., Núñez-Valdez, E. R., Crespo, R. G., & García-Díaz, V. (2020). An artificial neural network approach for predicting hypertension using NHANES data. *Scientific Reports*, *10*(1), 10620.

Lopez-Martinez, F., Schwarcz, A., Núñez-Valdez, E. R., & Garcia-Diaz, V. (2018). Machine learning classification analysis for a hypertensive population as a function of several risk factors. *Expert Systems with Applications*, *110*, 206-215. Mahadi, N., & Borodo, S. M. (2025). Enhancing Sentiment Analysis For Hausa Language With Improved Hausa Text Stemmer (Hts) And Machine Learning Models. *Fudma Journal Of Sciences*, *9*(12), 629-638.

Marin, I., & Goga, N. (2019, September). Hypertension detection based on machine learning. In *Proceedings of the 6th Conference on the Engineering of Computer Based Systems* (pp. 1-4).

Martinez-Ríos, E., Montesinos, L., & Alfaro-Ponce, M. (2022). A machine learning approach for hypertension detection based on photoplethysmography and clinical data. *Computers in Biology and Medicine*, *145*, 105479.

Martinez-Ríos, E., Montesinos, L., Alfaro-Ponce, M., & Pecchia, L. (2021). A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data. *Biomedical Signal Processing and Control*, *68*, 102813.

Mazzanti, M., Shirka, E., Gjergo, H., & Hasimi, E. (2018). Imaging, health record, and artificial intelligence: hype or hope?. *Current cardiology reports*, *20*, 1-9.

Nour, M., & Polat, K. (2020). Automatic classification of hypertension types based on personal features by machine learning algorithms. *Mathematical Problems in Engineering*, *2020*, 1-13.

Ramampiandra, E. C., Scheidegger, A., Wydler, J. & Scuwiirth, N. (2023) A comparison of Machine Learning and Statistical species Distribution Models. "Quantifying over-fitting supports model interpretation" Ecological Modelling 481 (2023) 110353.

Tjahjadi, H., & Ramli, K. (2020). Noninvasive blood pressure classification based on photoplethysmography using k-nearest neighbors algorithm: a feasibility study. *Information*, *11*(2), 93.

Tjahjadi, H., Ramli, K., & Murfi, H. (2020). Noninvasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis. *IEEE Access*, *8*, 20735-20748.

Ye, C., Fu, T., Hao, S., Zhang, Y., Wang, O., Jin, B., ... & Ling, X. (2018). Prediction of incident hypertension within the next year: prospective study using statewide electronic health records and machine learning. *Journal of medical Internet research*, *20*(1), e22.

Zhang, L., Yuan, M., An, Z., Zhao, X., Wu, H., Li, H. & Wu, W. (2020). Prediction of hypertension, hyperglycemia and dyslipidemia from retinal fundus photographs via deep learning: A cross-sectional study of chronic diseases in central China. *PloS one*, *15*(5), e0233166

Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA, et al. (2021) Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population representative studies with 104 million participants. *Lancet.* 398: 957–80. https://doi.org/10.1016/S0140-6736(21)01330-1