International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

Ensemble Approach for Hypertension Risk Prediction Using Clinical and Demographic Features

Okebule Toyin, Oguntimilehin Abiodun, Abiola O.B

Department of Computing, Afe Babalola University, Nigeria. okebulet@abuad.edu.ng

Abstract: Hypertension, also known as high blood pressure, is a major risk factor for cardiovascular diseases and stroke, and it often progresses silently until severe complications arise. Early detection is therefore essential for timely management and prevention. Traditional screening methods, however, do not always integrate multiple risk factors for accurate and early identification. This study develops a hypertension prediction system using deep learning and ensemble machine-learning techniques based on a dataset containing demographic, clinical, and lifestyle features. A Multi-Layer Perceptron (MLP), Random Forest, and XGBoost were trained and evaluated, with the Random Forest achieving an accuracy of 87.13%, XGBoost 84.50%, and the MLP 76.28%. An ensemble of the three models achieved 94% accuracy, indicating improved stability and predictive capability. While the system performs well, limitations such as possible overfitting and population-specific bias are noted. The study contributes to AI-driven healthcare by demonstrating a practical approach for early hypertension risk prediction. Future work may involve expanding the dataset, incorporating additional clinical indicators, and improving model robustness across diverse populations.

Keywords: Hypertension Prediction, Deep Learning, Machine Learning, Ensemble Model, Healthcare Analytics.

1 INTRODUCTION

Hypertension is a widespread non-communicable disease affecting about 1.28 billion adults aged 30–79 years worldwide, with the majority residing in low- and middle-income countries [1]. Because hypertension often develops without obvious symptoms, many individuals remain undiagnosed until severe complications, such as stroke, heart attack, kidney failure, or vision impairment, have already occurred [2]. This silent progression makes hypertension both a clinical and public-health concern. Nearly half of affected individuals globally remain undiagnosed and untreated [3].

In Nigeria, cardiovascular diseases account for about 11% of all recorded deaths, and roughly 30% of adults are hypertensive [4]. Early detection is essential; however, traditional diagnostic methods rely heavily on sporadic blood-pressure checks, which may fail to identify early-stage hypertension, especially in younger or asymptomatic individuals [5]. Although affordable medications exist, systemic challenges, including low awareness, poor screening coverage, and weak follow-up practices, continue to hinder effective management. Globally, only 21% of hypertensive individuals achieve adequate blood-pressure control [6].

Hypertension is broadly classified into two types: primary (essential) and secondary. Primary hypertension develops gradually with no identifiable single cause. Still, it is influenced by several modifiable and non-modifiable factors such as age, obesity, sedentary lifestyle, excess sodium intake, alcohol consumption, and genetic predisposition [7]. Secondary hypertension has a more rapid onset and typically results from underlying medical conditions, including chronic kidney disease, endocrine disorders, vascular abnormalities, or substance abuse [8].

The burden of hypertension also has substantial social and economic implications. Healthcare systems in many low-resource settings face constraints such as limited diagnostic equipment, insufficiently trained personnel, low health literacy, and poor treatment adherence. These challenges underscore the need for innovative, scalable, and accessible methods for early detection. Artificial Intelligence (AI) and Deep Learning (DL) offer promising solutions by enabling the processing of large datasets and the identification of complex risk patterns. AI-based tools have been applied to non-invasive blood-pressure monitoring, risk stratification, and personalized care planning.

Recent studies demonstrate the feasibility of using smartphone-based photoplethysmography (PPG) signals with neural network models for blood pressure estimation [9], while transformer-based architectures such as TransfoRhythm have achieved strong predictive performance using single-channel PPG signals [10]. These AI-driven systems reduce reliance on traditional diagnostic equipment and enable remote or continuous monitoring, particularly valuable in underserved regions. Integration with wearable devices and electronic health records (EHRs) enhances real-time monitoring and early intervention [11]. Despite these advances, existing machine-learning models for hypertension prediction often rely on limited feature sets or single-model architectures, and few studies have evaluated ensemble techniques combining neural networks with tree-based learners.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

This creates a gap for systems that integrate diverse feature types and leverage complementary strengths of multiple algorithms. Additionally, many existing studies lack practical deployment, limiting their real-world applicability. To address this gap, ensemble learning is adopted in this study because it improves generalization by combining models with different learning behaviors and strengths. Tree-based models capture non-linear feature interactions, while neural networks identify deeper patterns; together, they provide a more balanced and robust prediction mechanism. Based on these motivations, the objectives of this study are to:

i) develop a hypertension prediction model using demographic, clinical, and lifestyle features.
ii) compare the performance of MLP, Random Forest, and XGBoost classifiers.
iii) design an ensemble model that combines the strengths of all three algorithms for improved accuracy.
iv) deploy the final model in a web-based application for practical use by healthcare professionals and individuals.

In line with these objectives, this study proposes an ensemble-based hypertension prediction system that supports early risk identification and enhances clinical decision-making.

2 Review Of Related Work

Kaushik et al. [12] analyzed trends in anxiety and depression symptoms across different demographic groups in the United States during the COVID-19 pandemic. Their dataset, collected between April 2020 and subsequent months, enabled the authors to examine the prevalence and temporal fluctuations of depressive symptoms. To conduct the study, seven deep learning models were implemented, including RNN architectures such as LSTM and GRU, which are well-suited for temporal data. Among the models, the LSTM achieved the highest training and validation accuracies of 86.39% and 82.67%, respectively, while the RNN slightly outperformed it in both metrics. GRU models performed similarly but with marginally lower validation accuracy. Their findings also showed that younger age groups reported higher levels of mental-health symptoms, and predictions indicated possible future periods of increased mental-health challenges.

A related study in [13] examined hypertension as a common cardiovascular condition that poses significant public-health concerns globally. The authors emphasized the importance of early prediction of hypertension risk to improve health outcomes. Using a dataset of 13 attributes, including gender, age, and smoking habits, the study applied several machine-learning models to identify features associated with hypertension. Random Forest achieved 87.26% accuracy and outperformed alternative approaches in precision, recall, and F1 Score, highlighting machine learning's ability to manage hypertension risk.

S. Lin et al. [14] explored hypertension as a major risk factor for various chronic diseases, especially in aging populations. To improve early detection, they proposed an online hypertension monitoring model (HRP-OG) that integrates reinforcement learning and generative feature replay using electronic health records (EHRs). The model transforms hypertension prediction into a sequential decision-making problem and incorporates sensor data for real-time monitoring. Using the MIMIC-III dataset, HRP-OG showed superior performance, particularly in nonstationary environments and limited-data conditions.

Another study in [15] focused on predicting intraoperative hypotension (IOH), a complication of general anesthesia. The authors proposed a multichannel deep learning framework combining CNNs and Transformers to analyze physiological waveforms, including ABP, ECG, PPG, and CO_2. Using retrospective data from 14,140 adult patients, their model achieved AUROC values of 0.943, 0.928, and 0.923 at 5, 10, and 15 minutes before the hypotensive event, respectively, demonstrating the usefulness of multimodal signal integration.

The study in [16] proposed a machine-learning framework for hypertension prediction using a range of models, including Random Forest, Logistic Regression, Support Vector Machine, AdaBoost, and XGBoost. Using over 53,000 adult records, AdaBoost achieved the highest baseline accuracy of 87.02%, while Random Forest achieved the highest AUC of 98.37%. For follow-up predictions, XGBoost and Random Forest performed strongly again. The study emphasized the importance of temporal data and recommended integrating additional clinical data to enhance predictive accuracy.

Abubakar et al. [17] developed a deep-learning-based hypertension model using transfer learning to reduce computational cost. By fine-tuning a pre-trained neural network on the PPG-BP dataset, the model achieved an accuracy of 81.34%. The authors suggested exploring ensemble and reinforcement-learning methods in future studies to improve real-time prediction capability. Islam et al. [18] conducted a large-scale study involving more than 818,000 participants from Bangladesh, Nepal, and India to identify risk factors for hypertension using machine-learning models. The XGBoost, GBM, Logistic Regression, and LDA models achieved high accuracy and recall values, while the Decision Tree model achieved a precision of 91%. Age and BMI emerged as the strongest predictors. The study demonstrated that machine learning can effectively support population-level screening.

In [19], machine learning and transfer learning models were developed to predict hypertension in children and adolescents across seven South American cities. Random Forest, XGBoost, and LightGBM achieved up to 90% accuracy for children, whereas transfer learning was required to improve performance for adolescents. The study identified dietary habits as major contributing factors.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

Martinez-Ríos et al. [20] proposed a machine-learning model for early detection of hypertension using photoplethysmography (PPG) signals and clinical variables such as age, BMI, and heart rate. Using the Wavelet Scattering Transform (WST) for feature extraction, the SVM classifier achieved an F1 Score of 76% and a recall of 90.47%. Although combining PPG and clinical features yielded slight gains, the study highlighted challenges, including overlapping feature distributions and physiological interference.

Unlike prior studies that focus on physiological signals, wearable-device data, pediatric cohorts, or region-specific datasets, this study contributes a feature-rich, structured dataset–based ensemble model that integrates demographic, clinical, and lifestyle variables into a unified predictive system. Furthermore, deploying the model as a fully functional web-based application distinguishes this work from earlier studies that remained purely experimental, demonstrating its practical applicability in real-world hypertension screening.

3 METHODOLOGY

3.1. Research Summary

This section presents the research methodology employed to achieve the study objectives. A mixed approach was adopted to gain sufficient insights into the data and to apply appropriate machine-learning techniques for hypertension prediction. The methodology includes dataset acquisition, preprocessing, model development, ensemble integration, and performance evaluation. The overall workflow is illustrated in Fig. 1.

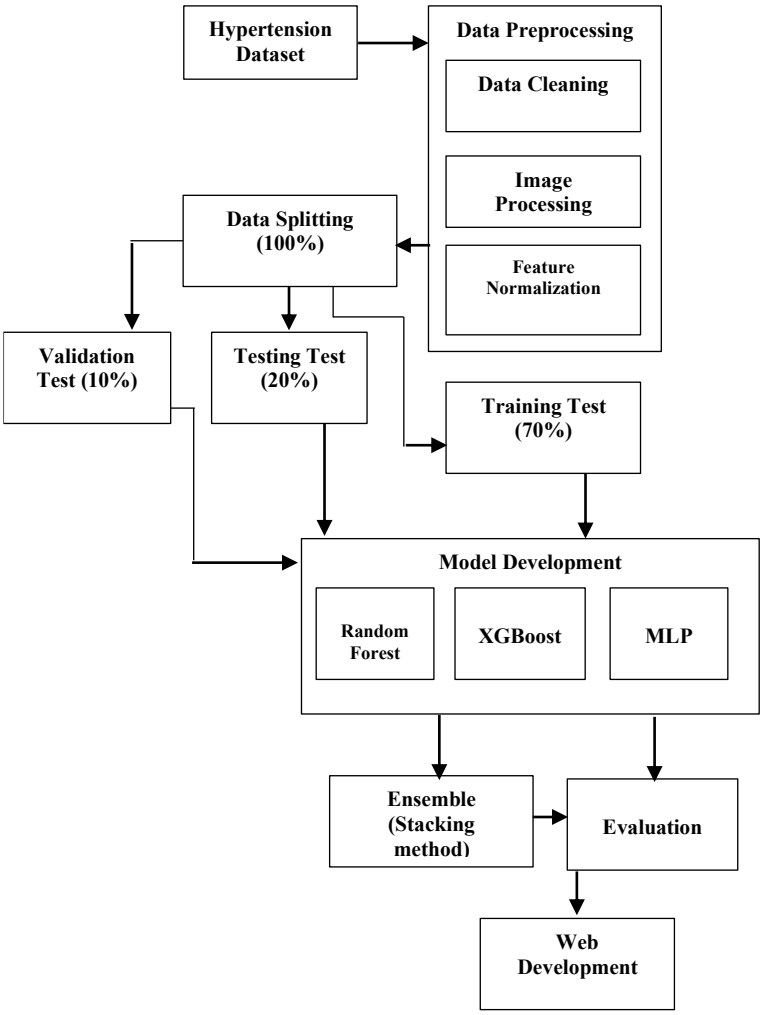

Fig. 1. Block Diagram of the proposed System

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

The dataset used in this study was sourced from Kaggle and comprises structured/tabular hypertension-risk features, including demographic, clinical, and lifestyle attributes such as age, BMI, blood pressure readings, family history, and behavioral indicators.

3.2. Data Preprocessing

To ensure the dataset's integrity and reliability, comprehensive preprocessing was performed before developing the predictive models.

Data Cleaning

Structured health records were cleaned to remove duplicate entries, correct inconsistencies, and handle missing values. Python libraries such as *Pandas* and *NumPy* were used to ensure completeness and accuracy across all variables.

Encoding

Categorical or non-numeric attributes, such as gender, lifestyle habits, and medical history, were converted to numerical representations suitable for machine learning algorithms. One-hot encoding was used for nominal features, while ordinal features were label-encoded. The target variable (hypertension status) was encoded to distinguish hypertensive from non-hypertensive cases.

Feature Normalization

Feature scaling was applied to standardize the feature ranges. Numerical attributes were normalized using *scikit-learn*'s standard scaler to improve model stability and convergence during training.

Loss Function

Because hypertension prediction is a binary classification problem, binary cross-entropy was used as the loss function to measure the discrepancy between the predicted probabilities and the actual class labels.

Data Splitting

To support effective training and evaluation, the dataset was partitioned into three subsets:

- 70% for training
- 20% for testing
- 10% for validation

This splitting ensures fair performance assessment on unseen data.

3.3. Model Development

Three models were developed and compared in this study: Multi-Layer Perceptron, Random Forest, and XGBoost. Each model provides unique strengths, and their combination enhances predictive performance.

Multilayer Perceptron (MLP)

The MLP is a feed-forward neural network consisting of an input layer, one or more hidden layers, and an output layer. Forward propagation is defined as:

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)} \quad (1)$$

$$a^{(l)} = \sigma(z^{(l)}) \quad (2)$$

where $W^{(l)}$and $b^{(l)}$ represent weights and biases, respectively, and σ is the activation function. The cross-entropy loss function is given by:

$$L = -\frac{1}{n}\sum_{c=1}^{C}\sum_{c=1}^{C} y_{ic}\log(\hat{y}_{ic}) \quad (3)$$

Weights are updated using gradient descent:

$$W^{(l)} \leftarrow W^{(l)} - \eta\frac{\partial L}{\partial W^{(l)}} \quad (4)$$

Although tree-based models often outperform neural networks on tabular data, the MLP was included because it can capture deeper non-linear interactions among features and contributes complementary learning behavior to the ensemble.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

Random Forest

Random Forest constructs multiple decision trees using different subsets of the training data and aggregates their predictions. For classification, the final prediction is obtained through majority voting:

$$\hat{y} = \text{mode}(T_1(x), T_2(x), \ldots, T_n(x)) \quad (5)$$

Random Forest helps reduce overfitting, improves variance control, and performs well with mixed feature types.

XGBoost

XGBoost builds decision trees sequentially, with each tree correcting errors from previous trees. The final prediction is:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \quad (6)$$

XGBoost includes regularization, second-order optimization, and efficient handling of complex, non-linear relationships.

Ensemble Learning

To leverage the complementary strengths of Random Forest, XGBoost, and MLP, a soft-voting ensemble method was used. The final prediction is:

$$\hat{y}_{final} = \arg\max (w_1 P_{RF} + w_2 P_{XGB} + w_3 P_{MLP}) \quad (7)$$

where P_{RF}, P_{XGB}, and P_{MLP} are model probabilities and w_1, w_2, w_3 are performance-based weights.

3.4. Hyperparameter Tuning

To optimize model performance, hyperparameters were adjusted using a combination of default settings and empirical tuning.

- Random Forest: number of trees (100), maximum depth (None), criterion ("gini").
- XGBoost: learning rate (0.1), maximum depth (6), number of estimators (100), subsampling (0.8).
- MLP: hidden-layer sizes (e.g., 64–32), activation function ("relu"), and learning rate (0.001).

While exhaustive tuning was not performed, these settings were chosen in accordance with standard practice and preliminary experimentation.

3.5. Model Evaluation

The performance of the models was assessed using accuracy, precision, recall, and F1-score.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$F1 = 2 \times \frac{Precision \cdot Recall}{Precision+Recall} \quad (11)$$

These metrics offer a balanced evaluation of predictive performance, especially for health-related classification tasks.

4 Implementation, Results, and Discussion

4.1. Model Evaluation

The models were implemented in Python using Google Colab. Deep-learning and machine-learning algorithms were evaluated using the *Scikit-learn* library with standard performance metrics, including accuracy, precision, recall, and F1-score. The overall objective was to assess the effectiveness of the three individual models and the ensemble model in predicting the risk of hypertension. Fig. 2 illustrates the relationship between age distribution and hypertension status.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

As expected, older individuals tend to show a higher likelihood of hypertension compared to younger groups. Although this trend is observable, the degree of variation is somewhat limited due to the dataset's characteristics. The dataset also shows a *mild imbalance* between hypertensive and non-hypertensive classes, but not enough to bias the models significantly.

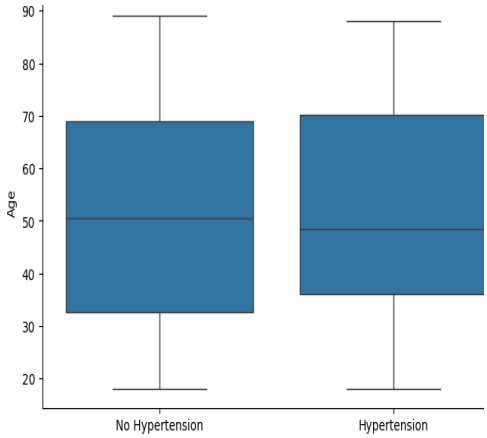

Fig. 2. Relationship between age distribution and hypertension status.

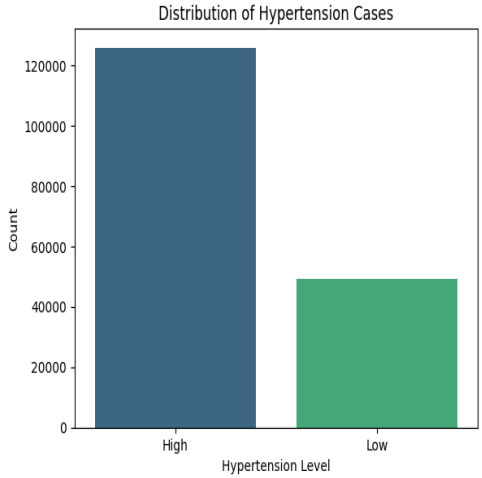

Fig. 3. Distribution of hypertension status to examine class balance.

Fig. 3 shows the distribution of hypertension status, indicating how the dataset is divided between high- and low-risk classes. This visualization helps confirm that the dataset is not severely skewed.

Table 1. Precision, Recall, F1-Score, and Accuracy of the Deep Learning Models

Model	Accuracy	Precision	Recall	F1-Score
Ensemble	94%	0.91	0.92	0.90
Random Forest	87%	0.83	0.87	0.84
XGBoost	84%	0.79	0.78	0.73
Neural Network (MLP)	76%	0.71	0.72	0.70

The Random Forest model performed strongly, particularly in identifying hypertensive individuals, with a recall of 0.87. This makes it well-suited for medical screening tasks where minimizing false negatives is critical. Its ability to capture non-linear feature interactions contributed to its accuracy of 87%. The XGBoost classifier achieved an accuracy of 84%, confirming the effectiveness of gradient boosting in handling complex and structured health data. Although its F1-score was slightly lower, it remained competitive and contributed significantly to the ensemble. The Neural Network (MLP) achieved an accuracy of 76%, which is lower than that of the tree-based models. Even so, the neural network captured distinct feature relationships that tree models may not fully capture, making it a valuable component of the ensemble.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

To leverage the strengths of all three models, a soft-voting ensemble strategy was used. This approach combines the variance reduction of Random Forest, the gradient-boosting optimization of XGBoost, and the pattern-recognition capabilities of MLP. As a result, the ensemble achieved 94% accuracy, outperforming the individual models. This demonstrates that combining complementary learning patterns improves overall prediction stability and reduces misclassification. Although repeated trials were not extensively performed, the results were stable across the available runs. Future work will include multiple experimental repetitions to measure statistical robustness. While the performance metrics are strong, reliance on a single dataset limits generalizability. External validation using datasets from different regions or clinical settings is recommended to further strengthen the model's robustness.

4.2. Web Development

The study integrated the best-performing model, the ensemble classifier with an accuracy of 94%, into a functional web-based application hosted on Hugging Face Spaces:

https://huggingface.co/spaces/chris-ai/HypertensionPrediction

Fig. 4 illustrates the user interface of the hypertension risk predictor. The application is designed to be intuitive and interactive, allowing users, including healthcare professionals, to enter patient information via sliders, dropdown menus, and numeric fields. This design supports ease of use and enhances accessibility for individuals without technical expertise. The interface shown in Fig. 5 demonstrates how users can adjust various clinical and lifestyle parameters to generate a predicted risk score. This interaction simulates a simple decision-support process and helps users understand how different variables contribute to hypertension risk.

Once the patient data is submitted, the application generates an output window as shown in Fig. 6, followed by a clear risk classification, High or Low Risk and a corresponding probability score (Fig. 7). For example, one prediction for a Nigerian patient resulted in a Low-Risk classification with a probability of 38.68%, demonstrating how the system provides straightforward, interpretable results. Beyond simple interaction, the web app demonstrates real-world applicability by translating the model into a practical tool that can be incorporated into clinical workflows, community health-screening programs, or patient self-assessment. It also supports telemedicine efforts, particularly in low-resource settings such as Nigeria, where access to routine medical screening may be limited. By providing risk assessment through a browser interface, the application enables early detection, improves awareness, and may help reduce delays in seeking medical attention. This deployment underscores the potential of machine-learning systems not only as research tools but as scalable digital health solutions that can assist healthcare providers and at-risk populations.

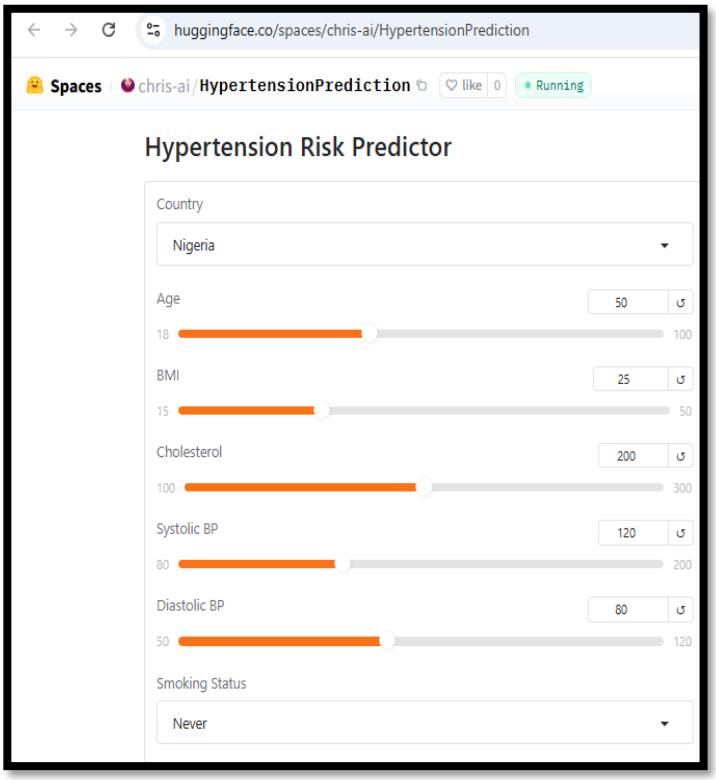

Fig 4. Hypertension Risk Predictor interface.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

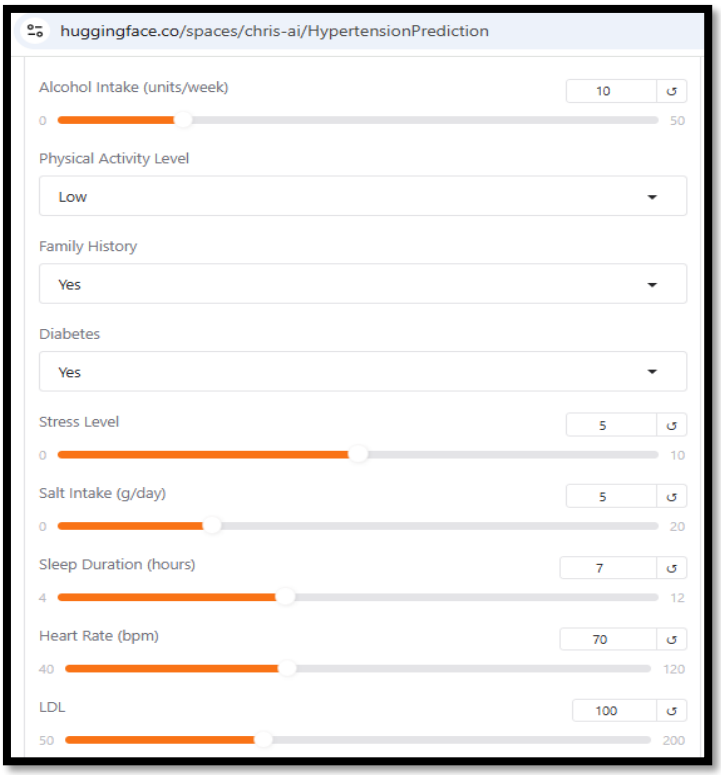

Fig 5. An interactive interface showing user inputs.

Fig. 5 shows the platform's interactive features, which allow users to adjust personal health parameters to generate a risk prediction. Fig. 6 shows the output screen following data submission, and Fig. 7 illustrates the classification result and risk percentage.

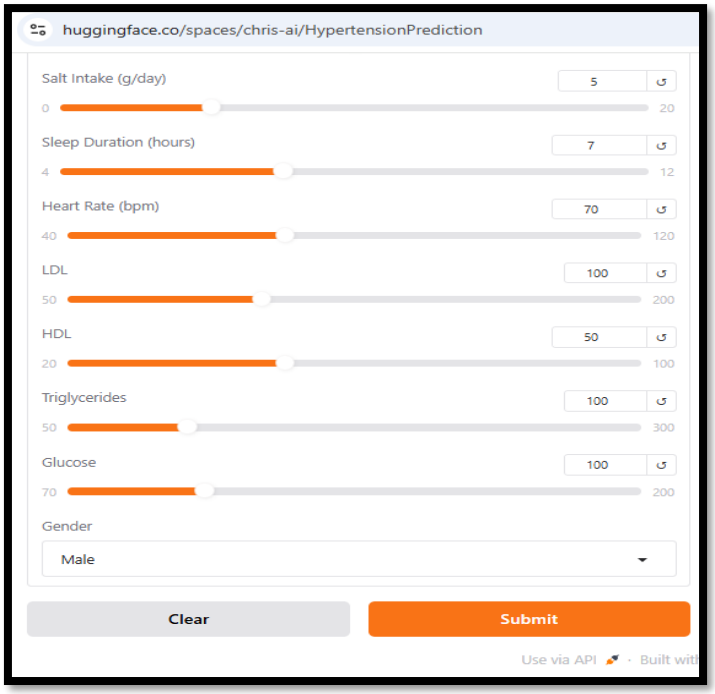

Fig. 6. Output Screen After Submission of Patient Data

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

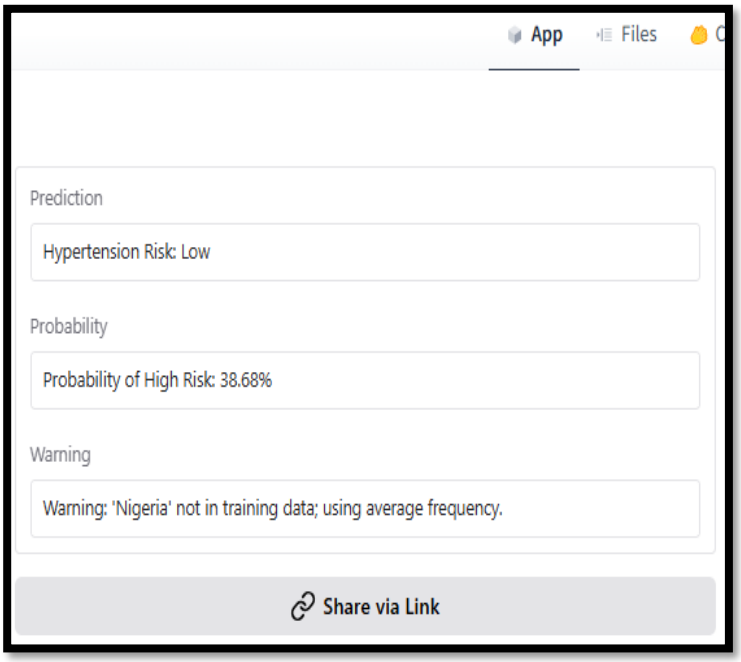

Fig. 7. Classification (High/Low Risk) with percentage score.

For example, the result shown in Fig. 7 indicates that a Nigerian patient's input data resulted in a Low Risk classification with a score of 38.68%, demonstrating how the application provides straightforward and interpretable output for practical use.

5 Limitations, Novelty, and Practical Implications

5.1. Limitations of the Study

Although the proposed hypertension prediction framework achieved strong performance, several limitations should be considered. First, the dataset used in this study was obtained from a single online source and may not fully represent the broader population. As a result, the findings may be influenced by sampling bias. Since the model was trained on a dataset not regionally representative, predictions for African demographic groups (as shown in the web demo) may not fully reflect demographic-specific risk, and external validation is needed to address this limitation. Second, the study relied mainly on demographic, clinical, and lifestyle attributes. Important factors such as genetic markers, medication adherence history, stress indicators, and longitudinal trends were not included. Incorporating such data could enhance the model's ability to capture more detailed patterns of hypertension. Third, the model was trained on cross-sectional data, limiting its ability to analyze changes in blood pressure or risk levels over time. Longitudinal datasets may improve predictive depth and early-warning capabilities. Although the ensemble model performed well, interpretability remains limited due to the black-box nature of neural networks and boosted trees. Clinicians may require more transparent explanations of how predictions are generated, especially in critical healthcare decision-making.

5.2 Novelty and Contribution of the Study

This study contributes to research on hypertension prediction in several ways. The proposed approach integrates three different learning algorithms, MLP, Random Forest, and XGBoost—into a unified ensemble model. This combination leverages complementary strengths, allowing the system to capture diverse patterns within the dataset more effectively than any single model. Unlike many prior studies that focus on physiological signals, wearable data, or region-specific clinical datasets, this research uses a feature-rich, structured dataset comprising demographic, lifestyle, and clinical indicators. Additionally, deploying the final model as a web-based application demonstrates practical usability and extends the work beyond purely theoretical analysis. The ensemble model achieved 94% accuracy, demonstrating its potential as a reliable screening tool for early hypertension detection.

5.3 Practical Implications

The findings of this study have significant practical implications. The ensemble model, with its high predictive accuracy, can support early identification of individuals at risk of hypertension. This may help clinicians prioritize high-risk cases and enable timely intervention. Deploying the system as a web application further extends its usability.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

It can be integrated into telemedicine platforms, community health screening programs, and remote patient monitoring initiatives. This is particularly beneficial in low-resource settings, where access to routine health assessments may be limited. By providing an accessible and user-friendly interface, the system encourages awareness, preventive care, and informed decision-making. The platform can also support researchers and public health stakeholders by providing an additional digital tool for monitoring high-risk groups, planning intervention strategies, and improving outreach efforts.

6 CONCLUSIONS

This study demonstrates the effectiveness of a deep learning approach for predicting hypertension using a comprehensive dataset comprising demographic, clinical, and lifestyle attributes. The ensemble model—combining a Multi-Layer Perceptron, Random Forest, and XGBoost classifiers—achieved an accuracy of 94%, underscoring its strong potential for early detection and risk assessment. By integrating the strengths of diverse learning algorithms, the ensemble method was able to capture complex and nonlinear patterns within the data that traditional models may overlook, thereby improving its ability to distinguish between hypertensive and non-hypertensive cases accurately. The integration of the model into a web-based application provides an intuitive and interactive platform for healthcare professionals to assess hypertension risk in real time. This practical deployment highlights the usefulness of artificial intelligence as a supportive tool in early diagnosis, screening, and preventive management of hypertension—a major global cause of morbidity and mortality. To further enhance predictive performance, future research should expand the dataset to include individuals from diverse geographic, demographic, and clinical backgrounds, thereby improving model generalizability and minimizing bias. Additionally, incorporating real-time biometric measurements from wearable devices may enable continuous monitoring and support earlier identification of hypertension risk, ultimately strengthening the role of AI-driven tools in modern healthcare systems.

ACKNOWLEDGMENT

The authors are grateful to the Founder and the Management of Afe Babalola University, Ado-Ekiti, Nigeria, for providing a conducive research environment for this work.

FUNDING INFORMATION

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

ETHICS STATEMENT

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

STATEMENT OF CONFLICT OF INTERESTS

The authors declare no conflicts of interest related to this study.

LICENSING

This work is licensed under a Creative Commons Attribution 4.0 International License.

REFERENCES

[1] E. Ahmed, O. Maphasha, and S. Okeke, "Knowledge, attitudes and practices on hypertension among patients in a district hospital," *South African Family Practice*, vol. 67, no. 1, pp. e1–e8, May 2025, doi: 10.4102/safp.v67i1.6094.

[2] B. Zhou, P. Perel, G. A. Mensah, and M. Ezzati, "Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension," *Nature Reviews Cardiology*, vol. 18, no. 11, pp. 785–802, May 2021, doi: 10.1038/s41569-021-00559-8.

[3] A. A. Okesina *et al.*, "Prevalence of undiagnosed hypertension and associated factors in Ndera sector, Gasabo district of Rwanda: a cross-sectional study," *BMC Public Health*, vol. 24, no. 1, p. 2495, Sep. 2024, doi: 10.1186/s12889-024-19999-1.

[4] A. E. Ojo, D. B. Ojji, D. E. Grobbee, M. D. Huffman, and S. a E. Peters, "The burden of cardiovascular disease attributable to hypertension in Nigeria: A modelling study using Summary-Level data," *Global Heart*, vol. 19, no. 1, p. 50, Jan. 2024, doi: 10.5334/gh.1332.

[5] D. Adeloye *et al.*, "Prevalence, awareness, treatment, and control of hypertension in Nigeria in 1995 and 2020: A systematic analysis of current evidence," *Journal of Clinical Hypertension*, vol. 23, no. 5, pp. 963–977, Feb. 2021, doi: 10.1111/jch.14220.

[6] "46% of adults with hypertension worldwide unaware they have it: WHO." https://www.aa.com.tr/en/world/46-of-adults-with-hypertension-worldwide-unaware-they-have-it-who/2899954

[7] P. K. Whelton *et al.*, "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines," *Hypertension*, vol. 71, no. 6, pp. 1269–1324, Nov. 2017, doi: 10.1161/hyp.0000000000000066.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4 | Issue 4 | pp. 01-11 | Oct-Dec 2025.
www.ijeresm.com | eISSN – 2583-4894
Received: 06 Sep 2025 | Revised: 24 Nov 2025 | Accepted: 01 Dec 2025 | Published: 07 Dec 2025

[8] S. Hegde, I. Ahmed, and N. R. Aeddula, "Secondary hypertension," *StatPearls - NCBI Bookshelf*, Jul. 30, 2023. https://www.ncbi.nlm.nih.gov/books/NBK544305/

[9] J. Esmaelpoor, M. H. Moradi, and A. Kadkhodamohammadi, "A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals," *Computers in Biology and Medicine*, vol. 120, p. 103719, Apr. 2020, doi: 10.1016/j.compbiomed.2020.103719.

[10] A. Arjomand, A. Boudesh, F. Bayatmakou, K. B. Kent, and A. Mohammadi, "TransfoRhythm: a transformer architecture conductive to blood pressure estimation via solo PPG signal capturing," *arXiv.org*, Apr. 15, 2024. https://arxiv.org/abs/2404.15352

[11] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24–29, Dec. 2018, doi: 10.1038/s41591-018-0316-z.

[12] P. Kaushik, K. Bansal and Y. Kumar, "Deep Learning-Based Prediction of Anxiety and Depression Symptoms using Demographic and Temporal Data," 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bangalore, India, 2025, pp. 1-6, doi: 10.1109/IITCEE64140.2025.10915280.

[13] S. Kaur, K. Bansal, and Y. Kumar, "Machine Learning-Based Prediction System for Risk Assessment of Hypertension using symptoms investigations," *International Journal of Experimental Research and Review*, vol. 46, pp. 139–149, Dec. 2024, doi: 10.52756/ijerr.2024.v46.011.

[14] S. Lin, H. Yan, S. Zhou, Z. Qiao, and J. Chen, "HRP-OG: Online Learning with Generative Feature Replay for Hypertension Risk Prediction in a Nonstationary Environment," *Sensors*, vol. 24, no. 15, p. 5033, Aug. 2024, doi: 10.3390/s24155033.

[15] K. Yang, M. Ren, J. Xu and X. Zeng, "Dynamic Prediction of Intraoperative Hypotension Based on Hemodynamic Monitoring Data With a Transformer-Based Deep Learning Model," *in 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Lisbon, Portugal, 2024, pp. 5166-5173, doi: 10.1109/BIBM62325.2024.10821970.

[16] K. Bora *et al.*, "A machine learning approach to predict hypertension using cross-sectional & two years follow up data from a health & demographic cohort of Assam, North East India," *The Indian Journal of Medical Research*, vol. 161, no. 4, pp. 394–405, May 2025, doi: 10.25259/ijmr_881_2024.

[17] Abubakar. B. Bada, A. B. Garko, and D. Gabi, "Hypertension Prediction Using Deep Learning With Transfer Learning Techniques," *FUDMA Journal of Sciences*, vol. 8, no. 6, pp. 257–263, Dec. 2024, doi: 10.33003/fjs-2024-0806-2855.

[18] S. M. S. Islam *et al.*, "Machine learning approaches for predicting hypertension and its associated factors using Population-Level data from three South Asian countries," *Frontiers in Cardiovascular Medicine*, vol. 9, p. 839379, Mar. 2022, doi: 10.3389/fcvm.2022.839379.

[19] K. Araujo-Moura, L. Souza, T. A. De Oliveira, M. S. Rocha, A. C. F. De Moraes, and A. C. Filho, "Prediction of hypertension in the pediatric population using machine learning and transfer learning: a multicentric analysis of the SAYCARE study," *International Journal of Public Health*, vol. 70, p. 1607944, Mar. 2025, doi: 10.3389/ijph.2025.1607944.

[20] E. Martinez-Ríos, L. Montesinos, and M. Alfaro-Ponce, "A machine learning approach for hypertension detection based on photoplethysmography and clinical data," *Computers in Biology and Medicine*, vol. 145, p. 105479, Apr. 2022, doi: 10.1016/j.compbiomed.2022.105479.