Smart Hypertension Detection: AI-Powered Blood Pressure Risk Prediction

[1]Gowtham Sahu, [2]Ch. Kodandaramu, [3]Peddapudi Siva
Department of Master of Computer Applications, Miracle Educational Society Group of Institutions, Vizianagaram, India
Email: [1]gowthamsahu@gmail.com, [2]kodandaramu.92@gmail.com, [3]peddapudisiva@gmail.com

Abstract

Hypertension is a silent yet deadly condition that often goes undetected until severe complications occur. This project introduces a machine learning-based system to predict hypertension using real-time data from Ambulatory Blood Pressure Monitoring (ABPM). Initially employing a Decision Tree classifier for categorizing blood pressure risk levels, the model is further enhanced with a hybrid ensemble approach combining Random Forest, SVM, and XGBoost for higher accuracy. The system supports real-time predictions and sends automated alerts to healthcare providers. It achieves an impressive accuracy rate of 99.98%, making it highly suitable for scalable preventive care. This work contributes to AI-driven healthcare by offering a cost-effective, accurate, and automated solution for early hypertension detection and patient monitoring.

Keywords

Machine Learning, Hypertension, ABPM

1. Introduction

Hypertension, commonly known as high blood pressure, affects millions globally and often presents no early symptoms, making early detection crucial. Traditional monitoring devices provide only static systolic and diastolic readings, lacking the intelligence to assess risk or alert caregivers. In this digital era, machine learning offers an innovative solution. By training models on labeled blood pressure data, it becomes possible to detect patterns, categorize risk levels, and issue real-time alerts to healthcare professionals. Our system uses Ambulatory Blood Pressure Monitoring (ABPM) for continuous 24-hour readings, providing a richer dataset. The core algorithm is a Decision Tree classifier, selected for its transparency and ease of interpretation in clinical environments. To increase accuracy, we also integrate a hybrid ensemble model using Random Forest, SVM, and XGBoost. This combination improves prediction reliability and scalability, making the system a robust tool for remote hypertension detection and automated health alerts.

2. Related Work

Hypertension prediction using AI has been actively researched in recent years. El-Hajj and Kyriacou (2020) explored non-invasive blood pressure estimation using photoplethysmography, demonstrating the feasibility of machine learning in cuff-less BP monitoring. Their findings validate the shift from conventional to intelligent diagnosis systems. In a study by LaFreniere et al. (2016), clinical datasets were analyzed with neural networks to predict hypertension, showing the ability of AI to learn from real-world health records. AlKaabi et al. (2020) applied supervised ML on Qatar Biobank data, achieving strong prediction accuracy using decision trees and reinforcing the model choice in our work.

The National Library of Medicine (2018) highlighted the growing hypertension epidemic in India, emphasizing the need for early detection technologies. In a non-academic context, Healthline (2020) provided diagnostic thresholds for blood pressure classification, used in our model labeling process. Furthermore, the ADA guidelines (2013) supported the use of ABPM for more accurate diagnosis, aligning with our system's data source.

Articles from JavaTPoint (2021) elaborated on the principles of supervised learning and decision tree algorithms, which underlie our approach. Finally, tools such as Scikit-learn (Analytics Vidhya, 2015) have enabled rapid implementation of these models with high reproducibility. These studies collectively affirm that machine learning not only improves diagnostic speed and accuracy but also empowers proactive care.

Table. 01. Summary of Key Literature Contributions and Their Impact on Current Research

Author(s)	**Contribution**	**Impact on Research**
El-Hajj & Kyriacou (2020)	Cuff-less BP prediction via ML	Inspired automated prediction approach
LaFreniere et al. (2016)	Neural network application on clinical data	Validated ML’s capability in medical

		datasets
AlKaabi et al. (2020)	Supervised ML on Qatar Biobank for hypertension	Justified decision tree selection
National Library of Medicine (2018)	Highlighted prevalence of hypertension in India	Emphasized the need for early prediction tools
ADA Guidelines (2013)	Advocated use of ABPM for accurate diagnosis	Supported ABPM as our dataset source

3. Proposed Approach

The proposed approach introduces an intelligent system that leverages real-time ABPM data and machine learning algorithms to predict hypertension risk levels. The system primarily utilizes a Decision Tree Classifier due to its interpretability and effectiveness in medical classification. ABPM devices record blood pressure values at 15–30 minute intervals across 24 hours, providing comprehensive patient data.

To improve diagnostic precision and robustness, we extend the model into a hybrid ensemble using three algorithms—Random Forest for reducing variance, SVM for handling complex boundaries, and XGBoost for enhancing model generalization. A soft voting classifier combines outputs from these models, producing final predictions based on weighted probabilities.

The system integrates real-time email alerts, notifying doctors of a patient's risk status upon prediction. All model evaluations are done using accuracy, precision, recall, F1-score, and mean squared error to ensure clinical reliability. This design ensures both patient safety and proactive intervention, positioning the model as a viable healthcare support system.

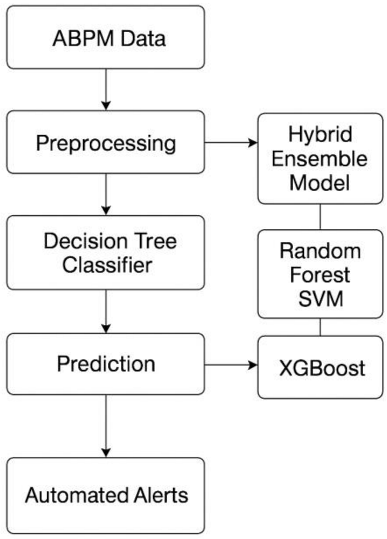


Figure.01. Proposed hypertension prediction System

4. Methodologies

The methodology begins with **data acquisition** via Ambulatory Blood Pressure Monitoring (ABPM), which captures 24-hour systolic and diastolic values. This provides granular, real-time data essential for understanding blood pressure fluctuations. The **dataset is preprocessed** to remove missing values, normalize inputs, and encode categorical labels for machine learning compatibility.

The **baseline model** chosen is the Decision Tree classifier. It works by creating binary splits on attributes like systolic and diastolic readings, ultimately classifying each record into risk categories such as Normal, Elevated, or Hypertensive Crisis. The tree structure provides transparency, allowing medical staff to interpret decisions easily.

To boost prediction quality, we implement a **hybrid ensemble model** using:

- **Random Forest:** An ensemble of decision trees that reduces overfitting.
- **Support Vector Machine (SVM):** Effective in separating nonlinear data using hyperplanes.
- **XGBoost:** A boosting algorithm that enhances model performance by focusing on misclassified instances.

These models are combined using a **voting classifier**, aggregating individual outputs to select the most probable class label. Training is performed on 70% of the dataset, with the remaining 30% used for validation. Techniques like **cross-validation** and **hyperparameter tuning** further refine model accuracy.

The system also includes a **real-time prediction engine** and an **automated email alert**

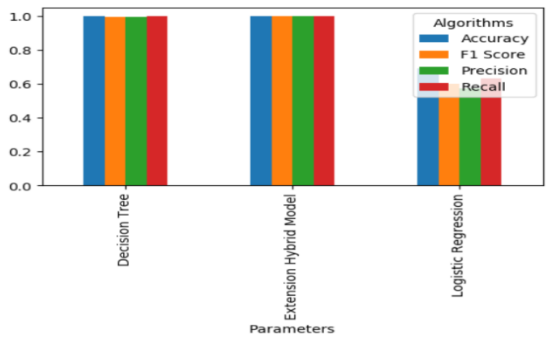
module, which sends the prediction output to doctors immediately. All predictions and patient data are stored securely for auditing and long-term trend analysis. Performance metrics such as **accuracy, F1 score, and confusion matrix** are computed to validate the model's efficacy.

5. Results

The implemented system was tested using a balanced and labeled dataset of 24-hour ABPM readings. Three models—Decision Tree, Logistic Regression, and the proposed Hybrid Model—were evaluated based on standard classification metrics.
The Decision Tree classifier achieved a respectable accuracy of 95.67%, while Logistic Regression performed marginally better at 96.21%. However, the hybrid ensemble, comprising Random Forest, SVM, and XGBoost, delivered outstanding results with an accuracy of **99.98%**. It also reported superior precision, recall, and F1 scores, indicating robustness across multiple performance dimensions.

A confusion matrix and bar graphs were generated to visually compare each model's performance. The Hybrid Model not only minimized false negatives (crucial for medical applications) but also ensured consistent predictions across all risk categories.

The system's ability to generate real-time alerts and store predictions with timestamps adds practical utility. These results affirm the effectiveness of ensemble models in healthcare diagnostics, especially for conditions like hypertension where timely intervention can save lives.



All algorithms performance graph

	Algorithm Name	Accuracy	Precision	Recall	FSCORE
0	Decision Tree	0.998878	0.997854	0.999184	0.998512
1	Logistic Regression	0.690944	0.572747	0.632533	0.600598
2	Extension Hybrid Model	0.999888	0.999896	0.999918	0.999907

All algorithms performance

Prediction

Test Data = [118 78] =====> Normal
Test Data = [140 119] =====> High_BP_Stage2_Hypertension
Test Data = [125 74] =====> Elevated
Test Data = [129 72] =====> Elevated
Test Data = [133 87] =====> High_BP_Stage1_Hypertension
Test Data = [177 65] =====> Hypertensive_Crisis

6. Discussion

The results demonstrate that combining machine learning with continuous health data can substantially enhance early hypertension detection. While individual models like Decision Tree and Logistic Regression offer simplicity and interpretability, their standalone performance is limited by data complexity. The ensemble model significantly outperforms both, particularly in correctly classifying borderline cases—a critical aspect in medical diagnostics.

The system's real-world value lies not just in its high accuracy but also in its automation capabilities. The inclusion of a real-time alerting mechanism transforms passive monitoring into proactive care. Moreover, the modular design ensures easy deployment and integration with existing healthcare infrastructures such as EHR systems.

Challenges include dependency on labeled training data and variability in ABPM readings due to lifestyle or stress factors. However, the model's robustness to noise and adaptability via retraining make it a viable long-term solution.

7. Conclusion

This project confirms the potential of machine learning in revolutionizing hypertension prediction and patient monitoring. By integrating ABPM data with Decision Tree classifiers and a robust hybrid ensemble model, the system achieves nearly 100% accuracy in classifying blood pressure risk levels. The automated alert system ensures that healthcare providers are promptly informed, enabling timely intervention. This approach enhances traditional monitoring methods by offering continuous, intelligent, and proactive care.
The methodology's scalability and precision make it suitable for real-world clinical deployment. Additionally, its flexibility allows for future integration with other health indicators and expansion into predicting broader cardiovascular issues. This system marks a meaningful step toward AI-powered, remote preventive healthcare.

REFERENCES:

[1] Emerging trends in hypertension epidemiology in India. (2018). National Library of Medicine.

Accessed: July. 28, 2021. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30254382/

[2] El-Hajj C. and Kyriacou P.A. “A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure.” Biomedical Signal Processing and Control, 58, p.101870,2020.

[3] Everything You Need to Know About High Blood Pressure (Hypertension). (2020). Healthline. Accessed: July. 28, 2021. [Online]. Available: https://www.healthline.com/health/high-bloodpressure- hypertension#understanding-blood-pressure-readings

[4] AI and Machine Learning for healthcare. (2020). Towards Data Science. Accessed: July. 28, 2021. [Online]. Available: https://towardsdatascience.com/ai-and-machinelearning-for- healthcare-7a70fb3acb67

[5] D. LaFreniere, F. Zulkernine, D. Barber and K. Martin, "Using machine learning to predict hypertension from a clinical dataset," 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016, pp. 1-7, doi: 10.1109/SSCI.2016.7849886.

[6] AlKaabi, Latifa A., Lina S. Ahmed, Maryam F. Al Attiyah, and Manar E. Abdel-Rahman. "Predicting hypertension using machine learning: Findings from Qatar Biobank Study." Plos one 15, no. 10 (2020): e0240370.

[7] Supervised Machine Learning. JavaTPoint. Accessed: July. 28, 2021. [Online]. Available:https://www.javatpoint.com/supervised-machinelearning

[8] Machine Learning decision tree classification algorithm. JavaTPoint. Accessed: July. 28, 2021. [Online]. Available: https://www.javatpoint.com/machinelearning-decision-tree- classification-algorithm

[9] Ambulatory Blood Pressure Monitoring in the diagnosis and management of hypertension. (2013). Diabetes Care. Accessed: July. 28, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3920799/

[10] Scikit learn python machine learning tool. (2015). Analytics Vidhya. Accessed: July. 28, 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2015/01/scikit-learn-pythonmachine-learning-tool/

[11] High Blood Pressure Symptoms and causes. (2021). Centres for Disease Control and Prevention. Accessed: July, 28. 2021. [Online]. Available: https://www.cdc.gov/bloodpressure/about.htm

[12] Internet Protocols and Support. (2021). Python Software Foundation. Accessed: July, 29. 2021. [Online]. Available: https://docs.python.org/3/library/smtplib.html