COMPUTING AND ALGORITHM INSIGHT
Volume 3, Issue 2, December 2025 | pp. 151-170
https://doi.org/10.63208/21018-120

OPEN ACCESS
Check for updates

ENHANCING HYPERTENSION PREDICTION ACCURACY THROUGH HYBRID MACHINE LEARNING FEATURE GENERATION AND LONG SHORT-TERM MEMORY NETWORKS

Tang Yifan[1] ✉, Sun Meiling[1], and Phan Quoc Bao[2]

School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China[1]; Department of Computer Science, Vietnam National University (VNU-HCMC), Ho Chi Minh City 70000, Vietnam[2]

ABSTRACT

The escalating prevalence of hypertension necessitates the development of sophisticated predictive mechanisms in healthcare. This study introduces a hybrid methodology synergizing machine learning and deep learning for feature augmentation and hypertension prognosis. We investigated five machine learning algorithms—Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), XGBoost (XGB), and Gradient Boosting (GB)—to generate predictive features, which were subsequently integrated with the original dataset to train a Long Short-Term Memory (LSTM) network. Comparative analysis reveals that the Gradient Boosting-based Feature Prediction combined with LSTM (GB-based FP + LSTM) yields superior performance, achieving a peak accuracy of 98.48% and an equivalent F1-score. Conversely, the LR-based integration demonstrated the lowest accuracy at 89.39%. These results suggest that the proposed GB-based FP + LSTM framework offers a highly reliable diagnostic instrument, providing clinicians and policymakers with an effective pathway for early and accurate hypertension detection.

Keywords: *Hypertension Prediction; Machine Learning; Deep Learning; Long Short-Term Memory (LSTM); Feature Generation.*

CORRESPONDING AUTHOR:
Tang Yifan
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China
contact: *yifan.tang@sjtu.edu.cn*

ARTICLE HISTORY
Received : July 21, 2025
Final Revised : December 1, 2025
Accepted : December 22, 2025
Published : December 30, 2025

© 2025 The Author(s) | Published by Integra Academic Press | This article is an open access article and distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license.

1. | INTRODUCTION

Hypertension, commonly referred to as high blood pressure, is a pervasive medical condition with significant consequences for human health [1], [2]. Its asymptomatic and gradual onset frequently masks its potential to inflict severe damage across multiple physiological systems [3]. Furthermore, persistent elevated blood pressure places a continuous strain on the heart, arteries, and other essential organs, thereby escalating the risk of critical complications such as myocardial infarction, stroke, renal failure, and cognitive impairment [4]. The deceptive nature of hypertension, which often remains undetected until it reaches advanced stages, emphasizes the urgent necessity for early detection and intervention [5]. Identifying hypertension at an incipient stage transcends the mere monitoring of numerical data; it is a vital preventative measure against life-threatening outcomes. specifically, early diagnosis empowers healthcare providers to deploy targeted interventions—ranging from lifestyle adjustments to pharmacological therapies—that can arrest disease progression and mitigate associated risks.

Given the substantial burden hypertension places on public health, the imperative to engineer effective detection methodologies is evident. Machine learning (ML), with its capacity to decipher intricate patterns within health data, has emerged as a promising frontier for transforming hypertension detection, potentially preserving lives through proactive, personalized healthcare strategies [6]. Conversely, while traditional blood pressure monitoring remains essential, it often faces limitations regarding continuous, real-time assessment, rendering it suboptimal for timely intervention.

The efficacy of machine learning in this domain stems from its ability to process massive and heterogeneous datasets, extracting subtle correlations and patterns that frequently escape conventional diagnostic methods [7], [8]. Unlike traditional techniques that depend on intermittent measurements, machine learning algorithms are capable of analyzing continuous streams of physiological data, providing a dynamic and comprehensive evaluation of blood pressure fluctuations [9], [10]. Additionally, these algorithms can synthesize data from diverse sources, including electronic health records, wearable technology, and lifestyle information, to form a holistic view of an individual's health status [11]. Crucially, ML models possess the ability to adapt and evolve as new data emerges, thereby refining their predictive accuracy. Algorithms such as Decision Trees (DT) and Logistic Regression (LR) are instrumental in detecting minute physiological deviations that precede overt hypertension, facilitating early intervention. Furthermore, machine learning aids in the construction of risk prediction models, allowing for the stratification of individuals based on susceptibility, which optimizes resource allocation and prioritizes preventative measures. Ultimately, machine learning facilitates a transition from reactive to proactive hypertension management [12], [13].

In summary, the integration of machine learning algorithms into healthcare technology signifies a paradigm shift in the detection of hypertension. By navigating complex datasets to analyze patterns and enable continuous monitoring, ML opens new pathways for early diagnosis and personalized management. These innovative solutions promise to revolutionize standard practices, offering a proactive approach that significantly improves patient outcomes. The convergence of technology and medical science not only deepens the understanding of hypertension but also propels the field toward a future where timely, data-driven interventions alleviate the burden of cardiovascular disease. This paper investigates the potential of machine

learning in hypertension detection, highlighting its critical role in modern healthcare. The collaboration between medical professionals and ML technologies is poised to redefine the standards of care for hypertension and beyond.

1.1 Contributions

This study offers the following primary contributions:

1) Feature Engineering via Machine Learning Classifiers: To expand the feature set, classifiers including Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), XGBoost (XGB), and Gradient Boosting (GB) are trained on the original features. The predictions generated by these classifiers are subsequently utilized as supplementary features, endowing the model with enhanced discriminatory capability. By integrating Long Short-Term Memory (LSTM) networks with dense layers and dropout regularization, the architecture effectively captures subtle dependencies, improving predictive performance for multi-class hypertension classification.
2) Enhanced Prediction through Hybrid ML and DL Approaches: This research aims to optimize hypertension prediction by synergizing machine learning-based feature engineering with deep learning-based prediction. The results indicate that the hybrid combination of GB and LSTM models achieved an F1-score of 98.48%. These findings offer valuable insights into the comparative performance of various models across multiple metrics.

1.2 Paper Organization

The remainder of this paper is structured as follows: Section II reviews related work concerning hypertension. Section III details the research methodology employed for hypertension detection. Section IV presents the experimental analysis, results, and discussion. Finally, Section V concludes the study and outlines directions for future research.

2. | RELATED WORK

A comprehensive review of existing literature is essential for contextualizing the current state of hypertension detection via machine learning. The intersection of healthcare and artificial intelligence has seen a surge in research activity, with a specific emphasis on applying machine learning to hypertension management [14], [15]. Significant contributions have utilized diverse datasets that include physiological measurements, electronic health records (EHRs), and lifestyle factors [16]. Furthermore, research has focused on developing and validating machine learning models designed to accurately predict blood pressure trends [17], [18]. Simultaneously, the integration of wearable technology with machine learning algorithms for continuous monitoring highlights the potential for real-time health assessment. For instance, smartphones and smartwatches equipped with photoplethysmography (PPG) capabilities are now being designed to measure blood pressure [19]. Additionally, recent investigations have been driven by advancements in feature engineering, model interpretability, and deep learning architectures [20]. While these studies collectively highlight the promise of machine learning in this domain, a critical evaluation of their methodologies, limitations, and generalizability remains necessary.

Building on this foundation, several studies have examined the integration of machine learning into clinical workflows. Developing predictive models that interface seamlessly with EHRs offers clinicians real-time insights into patient blood pressure trends [21], [22]. Moreover, enhancing the interpretability of these models addresses the challenge of

understanding the algorithmic decision-making process [23], [24]. Notably, the authors of [1] introduced ExHyptNet, a novel deep learning system utilizing wearable devices to detect hypertension via PPG signals. ExHyptNet processes these signals visually using EfficientNet, pre-trained for PPG analysis, while attention mechanisms and heatmaps improve interpretability for medical experts. This system outperforms existing methods in classification accuracy and elucidates the underlying reasoning through explainable modules. Consequently, ExHyptNet represents a significant advancement in non-invasive, continuous blood pressure monitoring, further emphasizing the importance of feature selection in Deep Neural Network (DNN) performance.

In another approach, the study in [25] proposed a "risk stratification" machine learning model to identify individuals susceptible to diabetes and hypertension, enabling efficient resource allocation. Using demographic data, medical history, and clinical test results from resource-constrained populations, the researchers evaluated several models, including Logistic Regression and Random Forest. Testing revealed that these trained models surpassed conventional stratification techniques, achieving an accuracy of 79.2%—a 13.5% improvement over prior methods—in limited-resource settings. The Random Forest model, in particular, demonstrated superior accuracy in predicting hypertension and diabetes risks.

Addressing the high prevalence of hypertension in South Asia, the research in [26] investigated the deployment of machine learning models for population-level prediction. By aggregating individual-level data from nationally representative surveys in Bangladesh, Nepal, and India, the study identified key determinants of hypertension. Various algorithms, including Decision Trees (DT), Random Forests (RF), Gradient Boosting Machines (GB), Extreme Gradient Boosting (XGB), Logistic Regression (LR), and Linear Discriminant Analysis (LDA), were trained and evaluated. The results indicated that machine learning models could predict hypertension with high accuracy, with LDA achieving approximately 90%, and strong performance observed in XGBoost, GBM, and LR. These advancements signal the field's maturation from proof-of-concept to practical implementations capable of reshaping hypertension detection and management.

Collectively, the reviewed studies underscore the potential of machine learning for hypertension detection and monitoring in clinical and real-world environments. The methodologies employed—ranging from EHR-integrated predictive models to deep learning for automated feature extraction—reflect a rapidly evolving field. The incorporation of wearable devices, feature selection, and ensemble learning further highlights the multifaceted nature of current research. However, as the domain progresses, addressing scalability, ethical implications, and interoperability within existing healthcare infrastructures is crucial. This section has explored recent studies, detailing their contributions and limitations to frame the comprehensive synthesis in subsequent sections. Table 1 provides a comparison of these previous techniques and their dataset characteristics.

Table 1. Comparison of Methods and Datasets

Models	Methods	Datasets
[1]	Deep Neural Network	Table 2 presents some of the important features used in the model training.
[25]	Random Forest Model	The variables dataset included are Age, Height, Weight, BMI, Heart rate, Random blood sugar, Waist circumference, Systolic blood pressure, Diastolic blood pressure, Urinations per night, Parental diabetes, Parental hypertension, Current smoker, Chest pain etc.
[26]	ML Models	The variables dataset included are blood pressure (BP), sociodemographic and economic factors, height, weight, hemoglobin, and random blood glucose

3. | RESEARCH METHODOLOGY

This investigation employs a structured systematic approach to hypertension prediction utilizing machine learning methodologies. The procedural framework and strategic approach implemented in this study are delineated in Figure 1. The proposed methodology leverages an extensive health dataset specifically curated for the non-invasive detection of cardiovascular diseases [27]. This dataset comprises 657 distinct records derived from 219 participants, with ages ranging from 20 to 89 years, covering prevalent conditions such as hypertension and diabetes, as quantified in Equation 1:

$$\begin{aligned} S &= 219 \\ R &= 657 \\ A_{min} &= 20 \text{ years old} \\ A_{max} &= 89 \text{ years old} \end{aligned} \tag{1}$$

Furthermore, data acquisition adhered to rigorous experimental protocols to ensure reliability. The study utilized this dataset to perform a thorough quality assessment of photoplethysmograph (PPG) signals, aiming to elucidate the intrinsic correlations between PPG waveforms and cardiovascular pathologies. The latent characteristic information embedded within these signals was utilized for the early, non-invasive screening of common cardiovascular diseases (CVDs), with a specific emphasis on hypertension.

Subsequently, a meticulous preprocessing phase was executed to optimize the dataset for analysis. Variables associated with specific health conditions—namely Diabetes, Cerebral Infarction, and Cerebrovascular Disease—were systematically eliminated, yielding a refined dataframe designated as dfnew. Categorical variables, including Sex (M/F) and Hypertension, underwent label encoding to convert them into numerical formats, thereby facilitating their seamless integration into the analytical workflow. Following preprocessing, a comprehensive Exploratory Data Analysis (EDA) was performed to uncover complex relationships within the data. A correlation matrix was computed to determine the strength and directionality of associations between variables. Visualized via a heatmap, this matrix was instrumental in guiding feature selection by providing insights into the interplay among various health parameters.

Moreover, feature selection was conducted to augment model efficiency by isolating the most pertinent variables. Specific attributes, including Heart Rate (b/m), Height (cm), subject ID, and Num., were excluded to generate the final refined dataset, denoted as dfnewfinal. This

rigorous curation ensures that the subsequent machine learning models are streamlined and optimized for capturing essential aspects of blood pressure prediction. Following a critical analysis of the target variable's distribution and the application of various feature analysis techniques, the dataset was reduced to the seven essential features detailed in Table 2.

Table 2. Feature Description

Features	Description
Sex	Stores Gender details of the patient that help in determining the type of pressure
Age	Stores patients age that helps in determining the severity and amount of disease
Weight	Stores weight details of patient, helps in the psychology of patient
Systolic Blood Pressure	It represents the maximum blood pressure on artery walls; it indicates cardiovascular health and is utilized in clinical procedures to detect various cardiovascular conditions that include hypertension
Diastolic Blood Pressure	It represents the minimum blood pressure on artery walls and indicates cardiovascular health and is utilized in clinical procedures to detect various cardiovascular conditions including hypertension
BMI	It stands for Body Mass Index, used to measure body fatness and is connected with the risk of several health conditions including cardiovascular issues, hypertension etc.

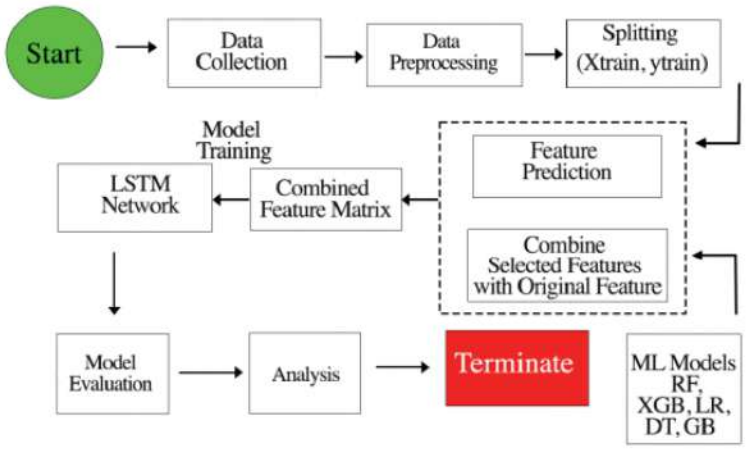


Figure 1. Proposed Methodology for Feature Prediction, Feature Addition, and Hypertension Prediction.

Distinguishing itself from datasets that prioritize secondary factors such as diabetes, uric acid levels, or smoking history, this dataset isolates the fundamental determinants of hypertension: blood pressure readings. It specifically incorporates Systolic and Diastolic Blood Pressure alongside Body Mass Index (BMI)—a critical metric for assessing body fat and cholesterol-induced cardiovascular risks—as the primary diagnostic indicators. To mitigate potential biases arising from data disparity, the study rigorously examined the distribution ratios within the hypertension variable. This quantification process highlighted the prevalence of various hypertensive categories, which were further elucidated through visual representation in a count plot, offering a holistic view of the case distribution.

Subsequently, the dataset underwent a systematic partition into independent features (X) and the dependent target variable (y). This stratification is crucial for model training and validation, facilitating effective generalization to novel data. To ensure dataset integrity and analytical completeness, missing values were addressed through zero-imputation. These foundational procedures—encompassing preprocessing, exploratory data analysis (EDA), feature selection, and data partitioning—establish a robust framework for the research.

Consequently, the refined dataset is optimally prepared for the development and evaluation of machine learning algorithms aimed at blood pressure prediction.

Figure 2 delineates the correlation coefficients (CC) among the dataset features, quantifying the magnitude and directionality of linear associations. The matrix reveals the interrelationships between variables, where positive and negative values signify direct and inverse correlations, respectively. Notably, the analysis yields a coefficient of -0.16 between age and hypertension, suggesting a weak negative correlation wherein advanced age is marginally associated with a reduced risk profile, though the link is tenuous. Conversely, the correlation between weight and blood pressure is approximately 0.2, indicating a weak positive association where increased weight corresponds slightly to elevated blood pressure levels.

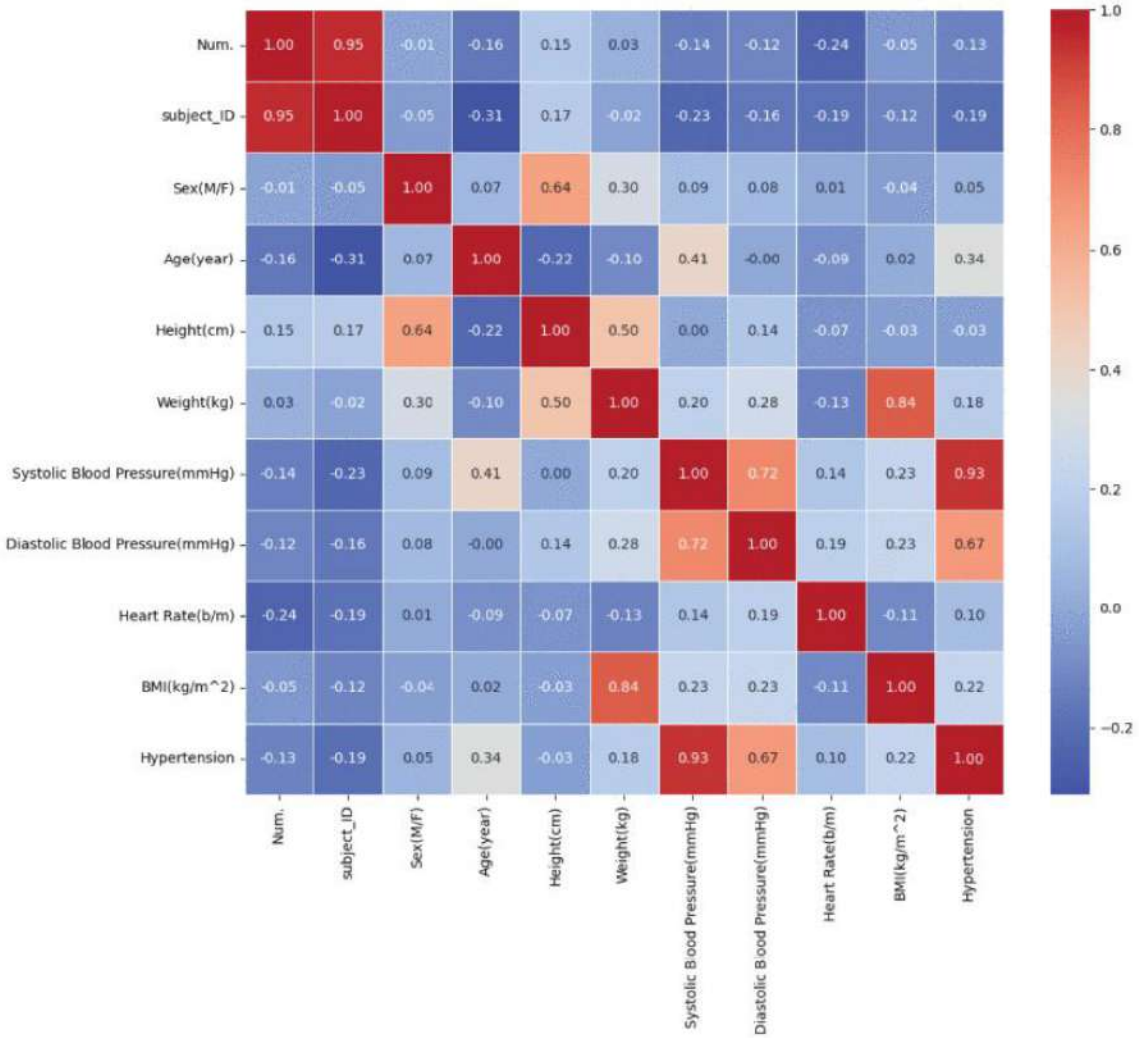


Figure 2. Matrix showing the correlation among variables in the employed dataset.

3.1 ML Models With DL Model

This research investigated a suite of machine learning classifiers—Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), XGBoost (XGB), and Gradient Boosting (GB)—to execute Feature Prediction (FP). To enhance predictive capability, the feature sets were flattened, and these classifiers were trained on the training partition (X_{train}, y_{train}). The resulting predictions were subsequently integrated with the primary dataset as supplementary features. Following this integration, the sequences were padded to standardized lengths. The

resulting combined and padded sequences, denoted as X_{train_padded} and X_{test_padded}, were reshaped to ensure compatibility with the Long Short-Term Memory (LSTM) network architecture.

Subsequently, a sequential neural network was defined for multi-class classification, comprising an LSTM layer succeeded by densely connected layers. The architecture was configured with an input shape corresponding to the padded sequences and incorporated dropout regularization to mitigate the risk of overfitting. The output layer utilized the softmax activation function to facilitate multi-class classification tasks. The model was compiled utilizing the categorical cross-entropy loss function and the Adam optimizer. Training was conducted using the combined feature set over a span of 50 epochs with a batch size of 32, employing callbacks for early stopping and model checkpoints to preserve the optimal model based on validation accuracy. The best-performing model was subsequently loaded for evaluation on the test set ($X_{test_padded}, y_{test_categorical}$).

The strategic selection of these specific techniques—including loss functions, gradient descent optimization, activation functions, dropout, and cross-entropy—was predicated on their potential to minimize prediction errors, optimize model parameters, and introduce the necessary non-linearity to capture complex data patterns. Collectively, these components contribute to the reliability, scalability, and interpretability of the framework, which are critical requisites for healthcare applications. Finally, the model's efficacy was rigorously assessed by comparing predictions (y_{pred}) against true labels (y_{test_true}). Comprehensive performance metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and ROC curves, were calculated to gauge the effectiveness of the developed machine learning model in predicting hypertension.

3.1.1 RF-Based FP and LSTM

Random Forest (RF) constitutes a robust ensemble learning methodology that synthesizes multiple decision trees to augment predictive accuracy and structural stability. Operationally, the algorithm generates a multitude of decision trees derived from varied dataset subsets, determining the final prediction via an aggregation mechanism such as majority voting or averaging. A distinct advantage of RF is its inherent resistance to overfitting, notably within high-dimensional data environments. Moreover, the algorithm yields feature importance scores, which are instrumental for deciphering intricate data dependencies and optimizing feature selection for predictive tasks. The mathematical formulation for this ensemble aggregation is depicted in Equation 2:

$$\hat{y} = RF(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \tag{2}$$

Equation 2 demonstrates the process by which the random forest model consolidates the individual outputs of constituent decision trees to formulate a unified prediction.

3.1.2 GB-Based FP and LSTM

Gradient Boosting (GB) constitutes a machine learning methodology that iteratively assembles a cohort of learners, typically decision trees, in a sequential manner. The primary objective is to minimize a specified loss function by introducing models designed to rectify errors committed by preceding models. The algorithm continuously refines these models based on the residual differences—or gradients—in the forecasts, a process that underpins the designation "gradient" boosting. This approach demonstrates significant efficacy in both

regression and classification assignments, which frequently necessitate robust predictive capabilities.

$$\hat{y} = GB(x) = \sum_{k=1}^{K} f_k(x) \quad (3)$$

As delineated in Equation 3, gradient boosting synthesizes the output of multiple weak learners to generate a final prediction.

3.1.3 LR-Based FP and LSTM

Linear Regression (LR) is a learning technique utilized to model the relationship between a dependent variable, y, and one or more independent variables, x. Predicated on the assumption of a linear correlation between these variables, the method seeks to establish an optimal linear model by minimizing the sum of squared residuals. Due to its inherent simplicity and high interpretability, LR is extensively adopted for forecasting outcomes and deriving inferential conclusions.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \quad (4)$$

Linear Regression calculates the coefficients (β) that minimize the variance between observed y values and those predicted by the linear model. This process results in a formula for forecasting y using x values, as presented in Equation 4.

3.1.4 XGB-Based FP and LSTM

XGBoost (Extreme Gradient Boosting) represents a state-of-the-art advancement of the gradient boosting technique, optimized for superior computational speed and performance. Rather than solely minimizing the loss function gradient, this procedure generates an ensemble of weak learners (predominantly decision trees) at each iteration using optimized objectives. Characterized by a two-stage boosting strategy, the capability to regularize training via L1/L2 norms, and support for parallel processing, the XGBoost method is highly efficient and effective across both regression and classification tasks. Furthermore, its scalability and integration with libraries such as Sklearn have secured its widespread adoption in machine learning competitions.

$$\hat{y} = XGB(x) = \sum_{k=1}^{K} f_k(x) \quad (5)$$

As shown in Equation 5, the algorithm utilizes the outputs of numerous weak learners to predict a final output for the input X by successfully adding each weak learner to minimize the overall loss function gradient.

3.1.5 DT-Based FP and LSTM

The Decision Tree (DT) stands as a prominent supervised learning algorithm, extensively applied in both classification and regression contexts. It operates by recursively partitioning the data space into regions defined by feature values, with the objective of reducing impurity (in classification) or variance (in regression) at every node. Along each branch, the terminal element represents a prediction or outcome. Decision trees are characterized by their clear, logical structure and ability to process heterogeneous data types, including numeric and categorical data. Consequently, they require careful fitting to manage variability within complex datasets.

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} w_i x_i + b\right) \quad (6)$$

The decision tree output $f(x)$, illustrated in Equation 6, is primarily utilized for classification, where signs are applied within the equation to designate classes as positive (1) or negative (0).

3.2 Employed Techniques

3.2.1 Loss Function

The cross-entropy loss function was selected due to its superior efficacy in the discriminator's role of distinguishing between the two classes. This metric quantifies the divergence between the actual label probability distribution and the distribution deduced by the model, making it highly pertinent for the binary classification task of hypertension detection. By imposing a substantial penalty on incorrect classifications, this function facilitates precise model fine-tuning, thereby enhancing both overall accuracy and the rate of convergence. The loss function is mathematically defined in Equation 7:

$$\text{Categorical Cross-Entropy Loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log(p_{ij}) \tag{7}$$

3.2.2 Optimization

Gradient descent was utilized as the fundamental optimization algorithm, as it is essential for minimizing the loss function in neural network training. Specifically, the Adam optimizer was chosen over alternatives such as AdaGrad and RMSProp. Adam is advantageous because it adapts the learning rate for each parameter individually, making it particularly suitable for the current dataset, which may exhibit noisy or sparse gradients. The decision to employ Adam was further justified by its proven ability to handle non-stationary objectives and accelerate model convergence. The optimization parameters for Adam are represented in Equation 8:

$$\theta_{t+1} = \theta_t - \frac{\alpha\widehat{m_t}}{\sqrt{\widehat{v_t}}+\epsilon} \tag{8}$$

3.2.3 Activation Function

The Rectified Linear Unit (ReLU) was implemented as the activation function for neurons within the hidden layers of the LSTM network, as shown in Equation 9. ReLU is favored for its computational simplicity and its effectiveness in mitigating the vanishing gradient problem, a common challenge in deep network architectures. By introducing non-linearity, ReLU enables the model to learn complex, high-level patterns, thereby augmenting its capacity to capture intricate data characteristics associated with hypertension.

$$\text{ReLU}(x) = \max(0, x) \tag{9}$$

3.2.4 Regularization

To mitigate the risk of overfitting, dropout regularization was integrated into the model architecture. This technique operates on the principle of randomly deactivating a subset of neurons during the training phase, compelling the deep neural network to learn robust features that facilitate better generalization to novel datasets. Dropout was selected due to its flexibility and demonstrated success in enhancing the generalization capabilities of deep learning models, particularly in scenarios involving small to moderately sized datasets. The dropout mechanism is defined in Equation 10:

$$\text{Dropout}(x) = \begin{cases} \frac{x}{1-p}, & \text{with probability } p \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

3.3 Software Tools and Libraries

The experimental framework of this study was established utilizing a suite of open-source software tools and libraries. Python was employed as the core programming language to facilitate model development and data analysis. For efficient array manipulation and complex numerical computations, the project utilized the robust library NumPy [28]. The

implementation of machine learning classifiers, such as Random Forest, along with the calculation of performance metrics including accuracy, precision, recall, and F1-score, was conducted using the scikit-learn library [29]. Furthermore, the development and training of the LSTM network were facilitated by TensorFlow [30] and Keras [31], two sophisticated deep learning frameworks developed by Google. For data visualization purposes, including the generation of various graphs, confusion matrices, and ROC curves, the libraries Matplotlib [32] and Seaborn [33] were employed. These specific tools were selected based on their operational stability, ease of integration, and extensive community support.

4. | EXPERIMENTATION ANALYSIS, RESULTS AND DISCUSSIONS

This section details the experimental findings, offering a comprehensive analysis of the machine learning models' efficacy in detecting hypertension. Executed within a Google Colab environment, the experiments yielded results detailed in Table 3, with the associated computational time recorded in Table 4. These outcomes highlight the effectiveness of various model combinations, shedding light on their respective accuracy, precision, recall, and F1-scores. The experimental protocol utilized a dataset split of 70% for training and 30% for testing. These results establish a critical basis for understanding the synergistic relationship between base classifiers and the LSTM neural network, providing significant insights into their collective potential for enhancing hypertension prediction in clinical settings. Consequently, Table 3 delineates the experimental results of the trained models, showcasing the performance metrics for the diverse combinations of base classifiers integrated with the LSTM neural network.

Table 3. Experimental Results (Performance Metrics)

Model Combination	Accuracy	Precision	Recall	F1-score
RF based FP + LSTM	96.97	97.03	96.97	96.93
GB based FP + LSTM	98.48	98.55	98.48	98.48
LR based FP + LSTM	89.39	90.41	89.39	89.30
XGB based FP + LSTM	95.45	95.96	95.45	95.44
DT based FP + LSTM	96.97	97.20	96.97	96.97

Table 4. Computational Time

Model Combination	Time(seconds)
RF based FP + LSTM	10.29
GB based FP + LSTM	11.36
LR based FP + LSTM	14.16
XGB based FP + LSTM	9.76
DT based FP + LSTM	13.68

Table 5 details the structural configuration of the deep LSTM model, specifying the layer type, output dimensionality, and parameter count for each constituent component. The architecture initiates with an LSTM layer that produces an output shape of (1, 8) and encompasses 416 parameters. This is immediately followed by a dense layer yielding a (64,) output shape with 576 parameters. Subsequent dense layers are engineered with output dimensions of (64,) and (32,), requiring 4,160 and 2,080 parameters, respectively. To counteract overfitting, a dropout layer is incorporated to randomly deactivate specific neurons during the training process. The framework terminates with a dense output layer characterized

by a (4,) output shape—corresponding to the four target classes—and 132 parameters. This design exemplifies a strategic synthesis of LSTM and dense layers, customized for the specific classification task. Furthermore, Table 6 outlines the specific parameter settings utilized for the classifiers within the predictive modeling framework.

Table 5. LSTM Architecture

Layer Type	Output Shape	Parameters
LSTM	(1, 8)	416
Dense	(64,)	576
Dense	(64,)	4160
Dropout	(64,)	-
Dense	(32,)	2080
Dense (Output)	(4,)	132

Table 6. Classifier Parameters

Classifier	Parameters
LSTM	- Units: 64 - Return Sequences: True - Activation: Relu - Loss: Categorical Cross Entropy - Optimizer: Adam
GB	- Loss: log - Learning Rate: 0.1 - Criterion: Friedman Mse - N Estimators: 100
DT	- Criterion: Gini - Max Depth: None - Min Samples Split: 2
LR	- LogisticRegression (all default)
XGBoost	- Max Depth: 3 - Learning Rate: 0.1
RF	- Number of Trees: 100 - Max Depth: None

The initial implementation of Random Forest (RF)-based Feature Prediction (FP) integrated with LSTM yielded favorable results, attaining an accuracy of 96.97%. This outcome substantiates the efficacy of synthesizing tabular data processing with ensemble learning methodologies. Subsequently, the Gradient Boosting (GB)-based FP and LSTM hybrid emerged as the superior model, achieving a peak accuracy of 98.48%. This enhancement underscores the significant potency of coupling boosting algorithms with LSTM architectures for hypertension prognosis. Conversely, the Logistic Regression (LR)-based framework exhibited suboptimal performance, with accuracy limited to 89.39%. This discrepancy suggests inherent constraints in applying linear relationship modeling to complex hypertension datasets, a limitation further reflected in the model's compromised precision and recall metrics regarding the identification of hypertensive instances.

Meanwhile, the XGBoost (XGB)-based configuration demonstrated robust performance with 95.45% accuracy, highlighting the algorithm's adaptability in capturing intricate data patterns to aid prediction. The Decision Tree (DT)-based approach yielded results analogous to the RF model, affirming the reliability of decision tree structures in tabular data contexts. Collectively, these findings indicate that ensemble methods—particularly GB and RF—

function synergistically with LSTM to provide robust detection capabilities. While LR faces challenges with health data complexities, necessitating careful model selection, advanced boosting techniques like XGB and consistent estimators like DT prove to be effective tools for significantly improving predictive accuracy in hypertension detection.

Figure 3 illustrates the confusion matrices derived from the experiments. The analysis of precision and recall, grounded in the results presented in Table 3, provides critical insights into the classification proficiency of the machine learning models, particularly regarding their aptitude for minimizing false positives. The mathematical formulations for precision, recall, and F1-score are defined in Equations 11, 12, and 13, respectively:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{12}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

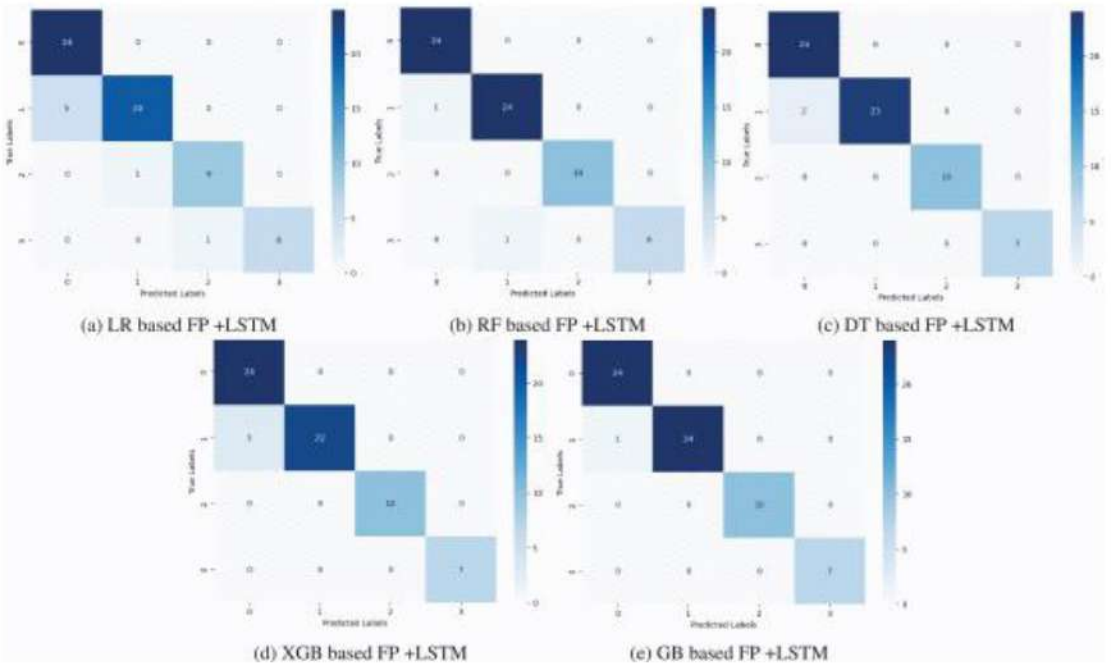


Figure 3. Confusion Matrix for Evaluating Model Performance.

Figure 4 illustrates the Receiver Operating Characteristic (ROC) curves, which serve as a critical metric for evaluating the performance of classification models. These curves delineate the trade-off between True Positive and False Positive rates across a spectrum of decision thresholds, effectively quantifying the model's ability to discriminate between positive and negative classes, where a larger Area Under the Curve (AUC) signifies superior discriminatory power.

In terms of precision, the hybrid model combining Gradient Boosting (GB) and LSTM demonstrates exceptional performance (98.55%), indicating a minimal error rate regarding the misclassification of non-hypertensive instances. Conversely, the integration of Logistic Regression (LR) and LSTM yields a lower precision (90.41%), suggesting a comparatively higher incidence of false positives. Regarding recall—defined as the ratio of true positives to the total actual positives—the metric assesses the sensitivity of the models in detecting the

condition. The GB and LSTM combination excels with a recall of 98.48%, underscoring its efficacy in identifying the vast majority of hypertension cases. In contrast, the LR and LSTM combination lags with a recall of 89.39%, revealing limitations in comprehensively capturing the full extent of hypertensive instances.

Collectively, these precision and recall metrics provide nuanced insights into the operational strengths and weaknesses of each predictive framework. In summary, the comparative analysis of these model combinations elucidates the distinct capabilities and limitations of various machine-learning approaches for hypertension detection. These findings have broader implications for the selection of appropriate models tailored to specific health datasets and highlight the potential of ensemble methods to significantly enhance predictive accuracy.

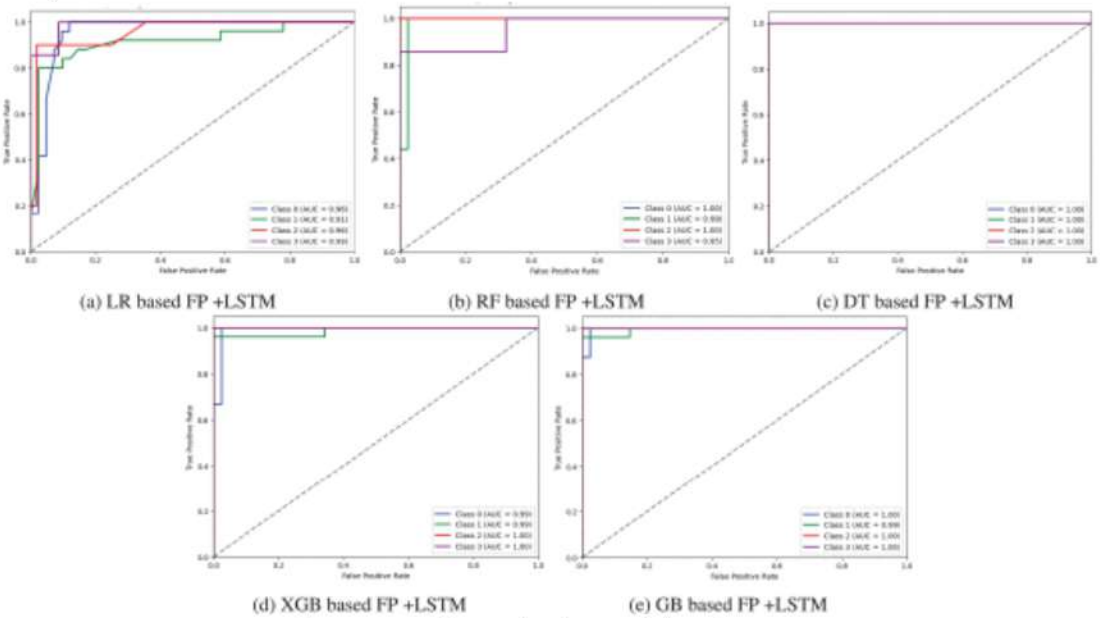
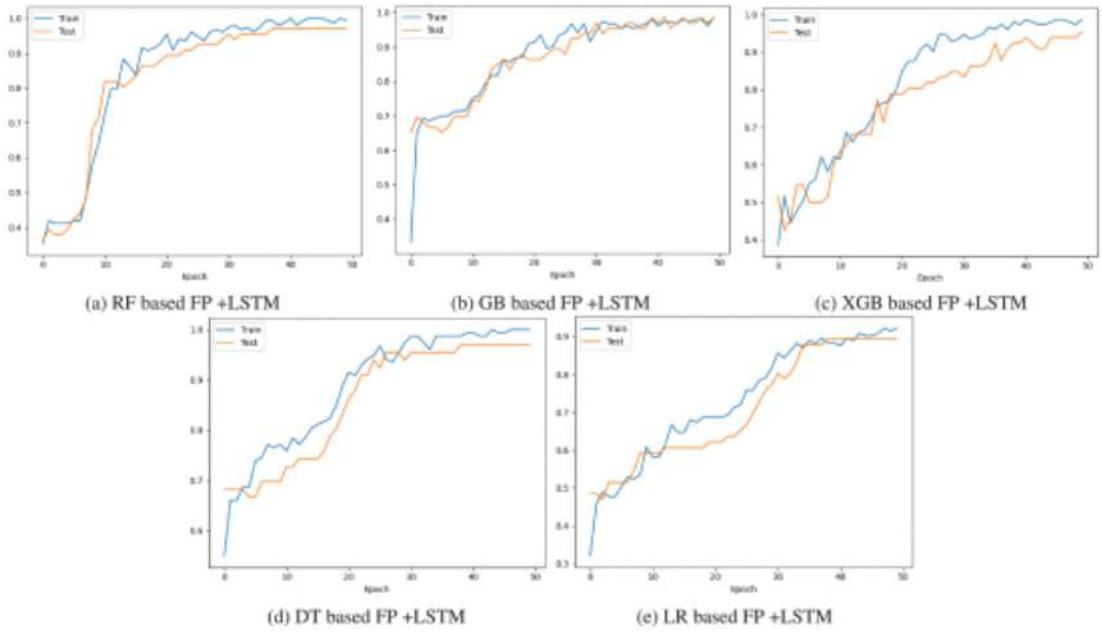


Figure 4. ROC for Evaluating Model Performance.

The trained models underwent concurrent evaluation using validation data to monitor performance trajectories throughout the training phase. This assessment utilized Model Accuracy and Model Loss curves, as illustrated in Figure 5 and Figure 6, respectively. The training curves demonstrated rapid rates of improvement, offering valuable insights into the learning dynamics of the models. By scrutinizing the accuracy curves, it was possible to validate the efficacy of the training process and identify any potential anomalies or inconsistencies in performance. Specifically, while the LR-based FP + LSTM model depicted in Figure 5e yielded lower overall results, it did not exhibit any structural irregularities.

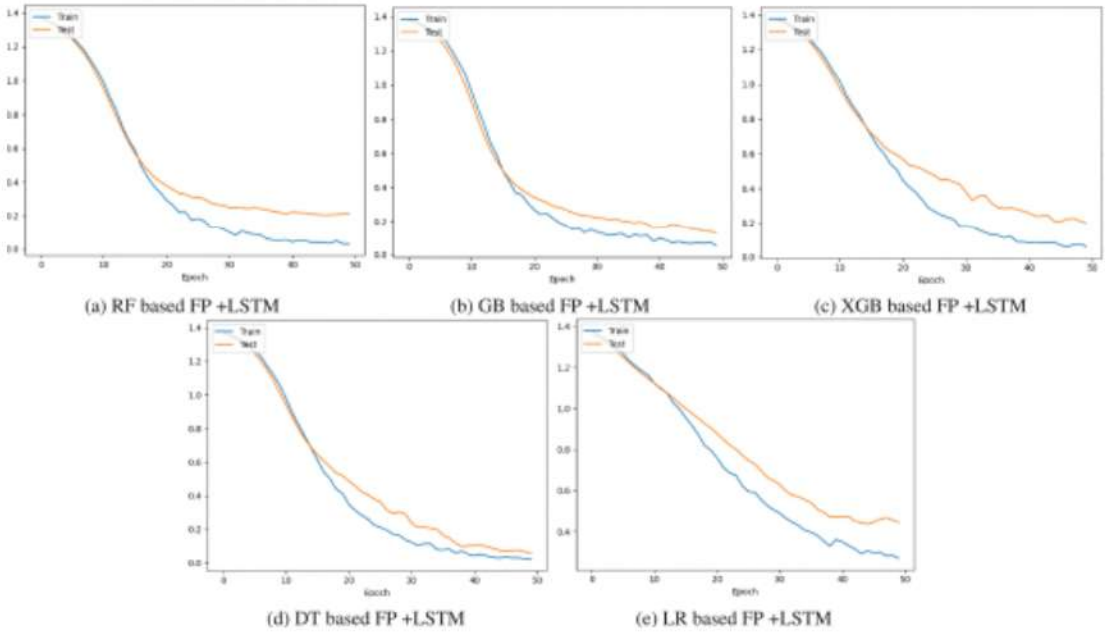
Conversely, the loss curves reflect the models' capacity to minimize errors during training. A consistently descending trajectory indicates convergence toward an optimal solution, whereas distinct variations or peaks may signal issues such as overfitting or instability. In this study, the majority of models displayed stable descending curves devoid of irregularities; however, minor deviations were observed in the LR-based FP + LSTM (Figure 6e) and the XGB-based FP + LSTM (Figure 6c). Ultimately, the analysis of these loss curves was instrumental in assessing the overall training progress and the models' ability to generalize to unseen data.



(a) RF based FP +LSTM (b) GB based FP +LSTM (c) XGB based FP +LSTM

(d) DT based FP +LSTM (e) LR based FP +LSTM

Figure 5. Model Accuracy Curves.



(a) RF based FP +LSTM (b) GB based FP +LSTM (c) XGB based FP +LSTM

(d) DT based FP +LSTM (e) LR based FP +LSTM

Figure 5. Model Loss Curves.

4.1 Discussion

The proposed model demonstrated robust performance on the dataset, with high accuracy rates observed across all hybrid configurations, including Random Forest (RF), Gradient Boosting (GB), Logistic Regression (LR), XGBoost (XGB), and Decision Tree (DT) integrated with the LSTM architecture. Specifically, the GB-based integration yielded superior results, attaining an accuracy of 98.48%, while the RF and DT variants both achieved 96.97%. Meanwhile, the XGB and LR combinations produced accuracy outcomes of 95.45% and 89.39%, respectively. The efficacy of this approach is attributed to the novel feature production technique, which enriched the input data, thereby enabling the LSTM model to interpret

complex patterns and data dynamics more efficiently. This innovation significantly optimized the LSTM's capability to handle intricate data structures across diverse domains.

In contrast, prior techniques exhibited limitations, primarily characterized by lower accuracy and reduced efficiency in learning complex patterns. Alternative models, such as Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), are prone to overfitting, particularly when applied to small or irregular datasets. Furthermore, these architectures remain vulnerable to adversarial attacks, where minor input variations can precipitate inaccurate predictions. Notably, the GB-based FP + LSTM model achieved a performance of 98.48%, surpassing not only the other comparative models but also the baseline approach referenced in [1] by a margin of 2.29%. Table 7 provides a comparative overview of the baseline method against the proposed approach.

Table 7. Performance Comparison With [1]

Reference	Accuracy
[1]	96.20%
GB based FP +LSTM	98.48%

Beyond the primary findings, several alternative models also demonstrated performance superior to the baseline approach. Specifically, both the RF-HDLM and DT-HDLM architectures achieved a commendable accuracy of 96.97%, further validating the efficacy of the proposed methodologies.

5. | CONCLUSION AND FUTURE WORK

This study establishes the theoretical underpinnings of a novel classification system designed for the early diagnosis of hypertension, synergizing machine learning-based feature engineering with deep learning architectures. In the predictive engineering phase, algorithms such as Gradient Boosting, Random Forest, Logistic Regression, Decision Trees, and XGBoost were utilized; however, the integration of Long Short-Term Memory (LSTM) networks with these models was pivotal in optimizing predictive precision. Notably, the hybrid configuration employing Gradient Boosting-based feature predictions in conjunction with the LSTM model achieved a paramount accuracy of 98.48%. The demonstrated reliability and high accuracy of this proposed model hold significant value for healthcare practitioners and policymakers. Facilitating the prompt diagnosis of hypertension is critical for averting severe sequelae, including myocardial infarction, stroke, and renal failure. Unlike standard diagnostic tools that rely on sporadic and less granular measurements, this model's capacity to analyze and predict trends from extensive, raw datasets represents a sophisticated advancement in diagnostic capability. Consequently, the implications of this model extend to transforming clinical management through early detection. The integration of machine learning and deep learning into clinical practice offers a pathway toward proactive prevention and intervention for cardiovascular conditions, potentially reducing both healthcare expenditures and mortality rates.

This research yields substantive findings regarding the utility of feature engineering combined with state-of-the-art deep learning models for disease identification. While the synergy between Gradient Boosting and LSTM networks demonstrates high sensitivity to analyzed dependencies, it remains a subject for further investigation to enhance model robustness. The modularity of this methodology suggests its potential applicability to other

chronic conditions. Future research avenues may involve incorporating additional covariates, such as genetic profiles and dietary habits, to further refine predictive accuracy. Additionally, the integration of real-time data acquisition via wearable technology and Electronic Health Records (EHRs) could facilitate continuous patient monitoring and the generation of automated alerts for both patients and medical personnel. It is also imperative that future studies address potential data biases to ensure the model's generalizability across diverse demographics. Finally, strict adherence to data privacy standards and legal compliance remains a critical consideration to protect patient information and maintain trust.

REFERENCES

1. S. T. U. Raju, S. A. Dipto, M. I. Hossain, M. A. S. Chowdhury, F. Haque, A. T. Nashrah, A. Nishan, A. Ahmad, M. Z. Chowdhury, and M. Hashem, "A novel technique for continuous blood pressure estimation from optimal feature set of PPG signal using deep learning approach," Tech. Rep., 2023.
2. A. F. Constant, E. V. Geladari, and C. V. Geladari, "The economic burden of hypertension," in *Hypertension and Cardiovascular Disease*. Cham, Switzerland: Springer, 2016, pp. 351–359.
3. E.-S.-A. El-Dahshan, M. M. Bassiouni, S. K. Khare, R.-S. Tan, and U. R. Acharya, "ExHyptNet: An explainable diagnosis of hypertension using EfficientNet with PPG signals," *Expert Syst. Appl.*, vol. 239, Apr. 2024, Art. no. 122388.
4. J. R. Ivy and M. A. Bailey, "Pressure natriuresis and the renal control of arterial blood pressure," *J. Physiol.*, vol. 592, no. 18, pp. 3955–3967, Sep. 2014.
5. L. P. Aiello, M. T. Cahill, and J. S. Wong, "Systemic considerations in the management of diabetic retinopathy," *Amer. J. Ophthalmol.*, vol. 132, no. 5, pp. 760–776, Nov. 2001.
6. G. F. S. Silva, T. P. Fagundes, B. C. Teixeira, and A. D. P. C. Filho, "Machine learning for hypertension prediction: A systematic review," *Current Hypertension Rep.*, vol. 24, no. 11, pp. 523–533, Nov. 2022.
7. D. LaFreniere, F. Zulkernine, D. Barber, and K. Martin, "Using machine learning to predict hypertension from a clinical dataset," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2016, pp. 1–7.
8. S. Ali, A. Hashmi, A. Hamza, U. Hayat, and H. Younis, "Dynamic and static handwriting assessment in Parkinson's disease: A synergistic approach with C-Bi-GRU and VGG19," *J. Comput. Theories Appl.*, vol. 1, no. 2, pp. 151–162, Dec. 2023.
9. J. R. Sutton, R. Mahajan, O. Akbilgic, and R. Kamaleswaran, "PhysOnline: An open source machine learning pipeline for real-time analysis of streaming physiological waveform," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 1, pp. 59–65, Jan. 2019.
10. B. Rush, L. A. Celi, and D. J. Stone, "Applying machine learning to continuously monitored physiological data," *J. Clin. Monit. Comput.*, vol. 33, no. 5, pp. 887–893, Oct. 2019.
11. K. Kario, "Management of hypertension in the digital era: Small wearable monitoring devices for remote blood pressure monitoring," *Hypertension*, vol. 76, no. 3, pp. 640–650, Sep. 2020.

12. M. Z. I. Chowdhury, I. Naeem, H. Quan, A. A. Leung, K. C. Sikdar, M. O'Beirne, and T. C. Turin, "Prediction of hypertension using traditional regression and machine learning models: A systematic review and meta-analysis," *PLoS ONE*, vol. 17, no. 4, Apr. 2022, Art. no. e0266334.
13. A. Cai, Y. Zhu, S. A. Clarkson, and Y. Feng, "The use of machine learning for the care of hypertension and heart failure," *JACC, Asia*, vol. 1, no. 2, pp. 162–172, Sep. 2021.
14. C. D. Toit, "Survey and evaluation of hypertension machine learning research," *J. Amer. Heart Assoc.*, vol. 12, no. 9, 2023, Art. no. e027896.
15. A. Shrivastava, M. Chakkaravarthy, and M. A. Shah, "A new machine learning method for predicting systolic and diastolic blood pressure using clinical characteristics," *Healthcare Analytics*, vol. 4, Dec. 2023, Art. no. 100219.
16. M. Asadullah, M. M. Hossain, S. Rahaman, M. S. Amin, M. S. A. Sumy, M. Y. A. Parh, and M. A. Hossain, "Evaluation of machine learning techniques for hypertension risk prediction based on medical data in Bangladesh," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 31, no. 3, p. 1794, Sep. 2023.
17. A. Leha, K. Hellenkamp, B. Unsöld, S. Mushemi-Blake, A. M. Shah, G. Hasenfuß, and T. Seidler, "A machine learning approach for the prediction of pulmonary hypertension," *PLoS ONE*, vol. 14, no. 10, Oct. 2019, Art. no. e0224453.
18. E. Kogan, E.-M. Didden, E. Lee, A. Nnewihe, D. Stamatiadis, S. Mataraso, D. Quinn, D. Rosenberg, C. Chehoud, and C. Bridges, "A machine learning approach to identifying patients with pulmonary hypertension using real-world electronic health records," *Int. J. Cardiol.*, vol. 374, pp. 95–99, Mar. 2023.
19. K. B. Kim and H. J. Baek, "Photoplethysmography in wearable devices: A comprehensive review of technological advances, current challenges, and future directions," *Electronics*, vol. 12, no. 13, p. 2923, Jul. 2023.
20. E. S. Muckley, J. E. Saal, B. Meredig, C. S. Roper, and J. H. Martin, "Interpretable models for extrapolation in scientific machine learning," *Digit. Discovery*, vol. 2, no. 5, pp. 1425–1435, 2023.
21. W. K. Cornwell, T. Tran, L. Cerbin, G. Coe, A. Muralidhar, K. Hunter, N. Altman, A. V. Ambardekar, C. Tompkins, M. Zipse, M. Schulte, K. O'Gean, M. Ostertag, J. Hoffman, J. D. Pal, J. S. Lawley, B. D. Levine, E. Wolfel, W. M. Kohrt, and P. Buttrick, "New insights into resting and exertional right ventricular performance in the healthy heart through real-time pressure-volume analysis," *J. Physiol.*, vol. 598, no. 13, pp. 2575–2587, Jul. 2020.
22. R. Miranda, C. Alves, A. Abelha, and J. Machado, "Data platforms for real-time insights in healthcare: Systematic review," *Proc. Comput. Sci.*, vol. 220, pp. 826–831, Jan. 2023.
23. A. Bibal and B. Frénay, "Interpretability of machine learning models and representations: An introduction," in *Proc. 24th Eur. Symp. Artif. Neural Netw., Comput. Intell. Mach. Learn.*, 2016, pp. 77–82.
24. P. J. G. Lisboa, S. Saralajew, A. Vellido, R. Fernández-Domenech, and T. Villmann, "The coming of age of interpretable and explainable machine learning models," *Neurocomputing*, vol. 535, pp. 25–39, May 2023.
25. J. J. Boutilier, T. C. Y. Chan, M. Ranjan, and S. Deo, "Risk stratification for early detection of diabetes and hypertension in resource-limited settings: Machine learning analysis," *J. Med. Internet Res.*, vol. 23, no. 1, Jan. 2021, Art. no. e20123.
26. S. M. S. Islam, A. Talukder, M. A. Awal, M. M. U. Siddiqui, M. M. Ahamad, B. Ahammed, L. B. Rawal, R. Alizadehsani, J. Abawajy, L. Laranjo, C. K. Chow, and R.

Maddison, "Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries," *Frontiers Cardiovascular Med.*, vol. 9, Mar. 2022, Art. no. 839379.

27. M. Elgendi, V. Galli, C. Ahmadizadeh, and C. Menon, "Dataset of psychological scales and physiological signals collected for anxiety assessment using a portable device," *Data*, vol. 7, no. 9, p. 132, Sep. 2022.
28. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, and R. Kern, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.
29. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.
30. M. Abadi, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," 2016, arXiv:1603.04467.
31. F. Chollet, "Keras," Tech. Rep., 2015. [Online]. Available: https://github.com/keras-team/keras
32. J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007.
33. M. Waskom, "Seaborn: Statistical data visualization," *J. Open Source Softw.*, vol. 6, no. 60, p. 3021, Apr. 2021.