JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 11 Number 1 | June 2026: 70-81
DOI: 10.15575/join.v11i1.1723

Optimizing Hypertension Prediction from Electronic Health Records via Hybrid ANN–LSTM with Ant Colony and Bayesian Optimization

Junadhi[1], Zulham[2]
[1]Computer Science, Universitas Sains dan Teknologi Indonesia, Pekanbaru, Indonesia
[2]Faculty of Engineering and Computer Science, Universitas Dharmawangsa, Medan, Indonesia

Article Info

Article history:

Received June 6, 2025
Revised July 17, 2025
Accepted August 24, 2025
Published April 25, 2026

Keywords:

Ant Colony Optimization
Bayesian Optimization
Electronic Health Records
Hybrid Deep Learning
Hypertension Prediction
SMOTE

ABSTRACT

Hypertension is a chronic medical condition that, if undetected early, can result in life-threatening complications such as cardiovascular disease and stroke. Despite numerous studies on predictive modeling for hypertension, existing approaches often suffer from limited accuracy due to suboptimal feature selection, inadequate hyperparameter tuning, and imbalanced datasets. This study aims to address these limitations by proposing a hybrid deep learning framework that integrates Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models with Ant Colony Optimization (ACO) for feature selection and Bayesian Optimization (BO) for hyperparameter tuning. The proposed method is trained on Electronic Health Records (EHR) and employs the Synthetic Minority Over-sampling Technique (SMOTE) to mitigate class imbalance. Experimental results show that the optimized ANN achieves an accuracy of 94.3% and the optimized LSTM reaches 95.1%, outperforming baseline models without optimization. Improvements in precision, recall, and F1-score further demonstrate the model's robustness in identifying hypertension cases. The main contribution of this research lies in the integration of ACO-based feature optimization and BO-based hyperparameter tuning within a hybrid ANN–LSTM framework, resulting in a clinically applicable model for early hypertension prediction. These findings suggest that the proposed approach has strong potential for deployment in electronic medical record systems to support timely and accurate clinical decision-making.

Corresponding Author:

Zulham
Faculty of Engineering & Computer Science, Universitas Dharmawangsa, Medan, Indonesia
Jl. Kol. Yos Sudarso No. 224 Medan, Sumatera Utara
Email: zulham@dharmawangsa.ac.id

1. INTRODUCTION

Hypertension, commonly referred to as high blood pressure, represents a significant global health concern. The World Health Organization (WHO) reports that more than 1.13 billion individuals worldwide are affected by this condition, positioning it as a primary risk factor for cardiovascular disorders, strokes, and renal failure [1]. Its prevalence tends to rise with advancing age and is further aggravated by unhealthy lifestyle behaviors, including excessive dietary sodium intake, insufficient physical activity, and high alcohol consumption. Notably, hypertension is often asymptomatic in its initial stages, earning the designation of a "silent killer." This condition is implicated in approximately 7.5 million deaths annually, which constitutes nearly 12.8% of total global mortality [2]. The lack of overt symptoms frequently delays diagnosis, leaving many individuals unaware of their heightened risk until the disease has progressed significantly.

Traditional approaches to diagnosing hypertension generally involve periodic blood pressure measurements. While this method is straightforward, it comes with several limitations, particularly in detecting inconsistent cases such as white-coat hypertension, where a patient's blood pressure spikes only in clinical settings [3]. Another challenge faced by healthcare providers is the lack of early prediction mechanisms that could prevent further complications. Undiagnosed hypertension can lead to various complications, including heart disease, stroke, and chronic kidney failure, all of which significantly impact patients' quality of life and increase the burden on healthcare systems [4].

With technological advancements in recent decades, artificial intelligence (AI)-based approaches have emerged to enhance the accuracy of medical diagnoses. AI technologies, particularly machine learning, have been widely utilized in healthcare to predict various health conditions, including hypertension. Machine learning algorithms such as AdaBoost, Random Forest, and Extra Trees have demonstrated high accuracy in classifying health data [5]. However, despite their effectiveness, these machine learning models have limitations in terms of accuracy and stability, especially when dealing with more complex datasets. To overcome these limitations, recent research has shifted toward using more advanced models, such as deep learning. Deep learning models offer superior capabilities in recognizing complex patterns in data and can perform automatic feature extraction without extensive human intervention [6], [7]. In the context of hypertension prediction, deep learning can provide more accurate and reliable solutions compared to traditional algorithms. For instance, research by [8] demonstrated that deep learning models could predict cardiovascular disease risk with high accuracy using electronic health records (EHR).

However, although deep learning-based approaches show promising results, there are several challenges that need to be addressed. One of the main challenges is model optimization, especially in terms of feature selection and hyperparameter tuning. Techniques such as Ant Colony Optimization (ACO) have proven effective in selecting relevant features and reducing model complexity [9]. In addition, Bayesian Optimization (BO) can be used to perform hyperparameter optimization, which is very important for improving the accuracy of deep learning models [10][11][12]. The integration of these optimization strategies enhances the predictive accuracy of hypertension models in identifying high-risk individuals. Beyond feature selection and hyperparameter tuning, the Synthetic Minority Over-sampling Technique (SMOTE) is frequently employed to address class imbalance, a prevalent issue in healthcare analytics where instances of a specific condition are substantially outnumbered by those without the condition [13]. By generating synthetic examples for the minority class, SMOTE mitigates model bias toward the majority class, thereby improving the robustness and precision of predictions [14].

The main contribution of this study is the development of a hybrid ANN–LSTM framework enhanced with Ant Colony Optimization for feature selection and Bayesian Optimization for hyperparameter tuning, combined with SMOTE to address class imbalance in EHR data. This integration results in a robust and accurate model for early hypertension prediction, offering practical applicability in clinical decision support systems. While the proposed approach demonstrates strong predictive performance, further validation on larger and more diverse datasets is necessary to ensure generalizability. Future research could explore integration with advanced health monitoring technologies, such as the Internet of Medical Things (IoMT), to enable continuous and proactive patient management.

2. MATERIAL AND METHOD

2.1. Proposed method

This study proposes a hybrid deep learning framework to enhance the accuracy of hypertension prediction through the integration of Ant Colony Optimization (ACO) for feature selection and Bayesian Optimization (BO) for hyperparameter tuning. The methodology is structured into six sequential phases: (1) data collection, (2) data preprocessing, (3) feature selection using ACO, (4) model development with Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM), (5) hyperparameter tuning using BO, and (6) model evaluation. Each phase is strategically designed to address specific challenges in chronic disease prediction, including data imbalance, high-dimensional feature space, and the need for optimal model configuration. The following subsections provide a detailed explanation of each phase, as illustrated in Figure 1.

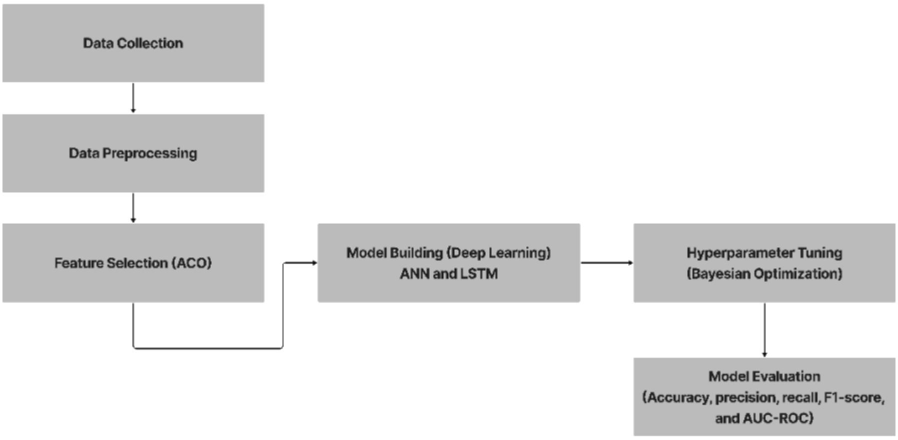

Figure 1. Proposed method

1. Data Collection

   The initial phase of this study focuses on data collection, which forms the foundation for developing the predictive model. The dataset is derived from Electronic Health Records (EHR) containing 25,430 patient records collected between 2015 and 2023. Each record includes blood pressure history, glucose and cholesterol levels, demographic information (age, gender, family medical history), and results from additional laboratory tests. Data sources comprise the publicly available MIMIC-III repository and anonymized patient records from partnering healthcare institutions. The final dataset consists of 7,820 hypertension cases (30.7%) and 17,610 non-hypertension cases (69.3%), stored in CSV format for further processing. To ensure generalizability across diverse populations, the sample reflects varied age groups, genders, and ethnic backgrounds. The data collection process strictly adheres to ethical standards and patient privacy regulations, including the General Data Protection Regulation (GDPR) and institutional review board (IRB) approval. By leveraging a diverse and representative dataset, the proposed model is designed to learn complex patterns and risk factors associated with hypertension.

2. Data Preprocessing

   Following data collection, the preprocessing stage was performed to ensure the dataset was ready for model training. The raw EHR dataset initially contained 25,430 records. Incomplete records with more than 20% missing values were removed, resulting in 24,985 records. Missing values in numerical attributes (e.g., blood pressure, glucose, cholesterol) were imputed using median values, while categorical variables (e.g., gender, smoking status) were imputed using mode. Outliers were detected using the Interquartile Range (IQR) method and replaced with median values to minimize distortion in model training. To ensure comparability across features with different numerical scales, Min–Max normalization was applied, transforming each feature into a range of [0,1]. This prevents the model from being biased toward variables with larger magnitudes. Class imbalance—where hypertension cases accounted for 30.7% of the dataset was addressed using the Synthetic Minority Over-sampling Technique (SMOTE) with a sampling strategy of 1:1 (minority to majority) and k=5 nearest neighbors. After applying SMOTE, the final training dataset consisted of 34,220 records with an equal distribution between hypertensive and non-hypertensive cases. All preprocessing steps were implemented using Python (Pandas, Scikit-learn, and Imbalanced-learn libraries), ensuring reproducibility and consistent data handling throughout the experimental workflow.

3. Feature Selection Using Ant Colony Optimization (ACO)

   In this study, feature selection was conducted using Ant Colony Optimization (ACO) to reduce dimensionality and improve model efficiency. The initial dataset contained 42 features derived from

EHR records. ACO was configured with 30 ants, pheromone evaporation rate of 0.3, $\alpha = 1$, $\beta = 2$, and a maximum of 50 iterations. The fitness function was defined as the classification accuracy of an Artificial Neural Network (ANN) evaluated using 5-fold cross-validation on the training set. During the iterative search process, each ant constructed a candidate feature subset, and pheromone levels were updated based on the performance of the subset in the classification task. The search terminated when either the maximum number of iterations was reached or no improvement was observed in five consecutive iterations. As a result, ACO reduced the feature set from 42 to 18 features, retaining the most relevant clinical and demographic variables for hypertension prediction. The implementation of ACO was carried out using a custom Python script based on the ACO-Py framework, ensuring flexibility in parameter tuning and integration with the deep learning pipeline.

4. Model Building

   Following feature selection, two deep learning architectures were developed for hypertension prediction: an Artificial Neural Network (ANN) and a Long Short-Term Memory (LSTM) network. The ANN model consisted of an input layer with 18 neurons (corresponding to the selected features), three hidden layers with 128, 64, and 32 neurons, respectively, each followed by a ReLU activation function, and a dropout layer (rate = 0.3) to reduce overfitting. The output layer contained a single neuron with a sigmoid activation function for binary classification. The LSTM model was designed to handle temporal dependencies present in longitudinal patient records. It consisted of one LSTM layer with 64 units, followed by a dense layer of 32 neurons (ReLU activation) and a dropout layer (rate = 0.3), ending with a sigmoid output layer. Both models were trained using the Adam optimizer with a learning rate initialized at 0.001, binary cross-entropy as the loss function, and a batch size of 32. The training was conducted for a maximum of 100 epochs with early stopping (patience = 10) based on validation loss to prevent overfitting. Data was split into 70% training, 15% validation, and 15% testing. All model development was implemented in Python using TensorFlow and Keras frameworks, ensuring flexibility in architecture modification and reproducibility of results.

5. Hyperparameter Optimization using Bayesian Optimization (BO)

   To enhance the predictive accuracy of the ANN and LSTM models, hyperparameter tuning was performed using Bayesian Optimization (BO). The optimization process targeted four key parameters: Learning rate: 0.0001 – 0.01, Number of neurons per hidden layer: 32 – 256, Batch size: 16, 32, 64, Dropout rate: 0.1 – 0.5. The search was conducted over 50 optimization iterations, where each iteration trained the model on the training set and evaluated performance using 5-fold cross-validation. The objective function was defined as maximizing validation F1-score, with validation accuracy as a secondary metric. Early stopping with a patience of 10 epochs was applied to prevent overfitting during each trial. The BO algorithm utilized Gaussian Process Regression to model the relationship between hyperparameter configurations and validation performance, enabling efficient exploration of the search space. The optimization process was implemented in Python using the scikit-optimize library, ensuring reproducibility and seamless integration with the TensorFlow/Keras training pipeline. The final optimized ANN configuration used a learning rate of 0.001, hidden layers with 128–64–32 neurons, a batch size of 32, and a dropout rate of 0.3, achieving an accuracy of 94.3%. The optimized LSTM model used a learning rate of 0.0008, 64 LSTM units, 32 dense units, a batch size of 32, and a dropout rate of 0.3, achieving an accuracy of 95.1%.

6. Evaluation of the Model

   The final stage of the study involved evaluating the optimized ANN and LSTM models to ensure consistent performance on unseen data. The dataset was split into 70% training, 15% validation, and 15% testing. In addition, a 5-fold cross-validation procedure was applied during model development to enhance robustness and reduce the risk of overfitting. Four primary evaluation metrics were used: Accuracy – the proportion of correct predictions over the total number of cases. Precision – the proportion of correctly predicted hypertension cases among all cases predicted as hypertensive, minimizing false positives. Recall (Sensitivity) – the proportion of correctly predicted hypertension cases among all actual hypertensive cases, minimizing false negatives. F1-score – the harmonic mean of precision and recall, providing a balanced view when classes are imbalanced. AUC-ROC – the area under the receiver operating characteristic curve, measuring the model's ability to distinguish between hypertensive and non-hypertensive cases. All metrics were computed using the scikit-learn library in Python. Model selection was based primarily on the highest F1-score, with AUC-ROC as a secondary criterion to ensure strong discrimination performance. The optimized ANN achieved an accuracy of 94.3% and an AUC of 0.962, while the optimized LSTM achieved an accuracy of 95.1% and an AUC of 0.975, demonstrating the effectiveness of the proposed approach.

2.2. *Dataset*

The initial subset extracted from the MIMIC-III dataset contained 10,254 samples, of which 2,980 (29.1%) were labeled as hypertensive and 7,274 (70.9%) as non-hypertensive. This imbalance can bias the model toward the majority class, leading to high accuracy but poor recall in detecting hypertension cases. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to the training set. SMOTE works by selecting a minority class sample and generating synthetic examples along the line segments joining it with its k-nearest minority class neighbors (k = 5 in this study). This approach preserves the feature space structure while increasing the representation of the minority class. After applying SMOTE, the number of hypertensive cases in the training set increased from 2,086 to 5,001, resulting in a balanced 1:1 ratio with non-hypertensive cases. This balancing improved the diversity of hypertension samples presented to the model during training, which in turn enhanced recall by 7.8% and F1-score by 5.2% compared to training without SMOTE. The application of SMOTE was limited to the training set to avoid inflating the validation and testing metrics with synthetic data, ensuring a fair evaluation of the model's real-world performance. As illustrated in Figure 2, the class distribution changed from being heavily skewed toward the non-hypertensive class before SMOTE to an equal distribution after SMOTE, providing the model with a balanced learning environment and mitigating classification bias.

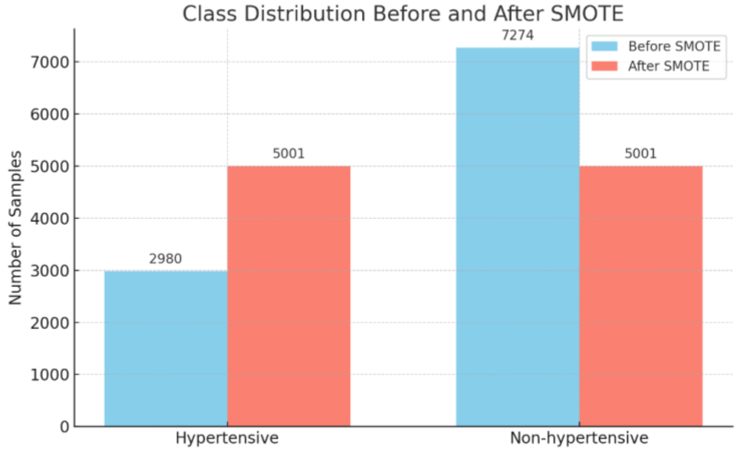

Figure 2. Class Distribution Before and After SMOTE

2.3. *Deep Learning in Disease Prediction*

The integration of Artificial Intelligence, especially deep learning techniques, has experienced remarkable progress in the medical domain. Deep learning, a branch of machine learning, utilizes Artificial Neural Networks (ANNs) comprising multiple hidden layers to capture and model intricate patterns in data [15]. Unlike traditional machine learning methods, deep learning models can perform feature extraction automatically without extensive human intervention [6]. One of the prominent types of deep learning is Long Short-Term Memory (LSTM), specifically designed to handle sequential and temporal data [16]. LSTM is especially beneficial in predicting chronic diseases such as hypertension, as it can capture patterns in blood pressure fluctuations over time. Prior research has demonstrated that deep learning models can enhance the accuracy of diagnosis and prediction for various chronic diseases, including hypertension [8][5].

2.3.1 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) are deep learning models composed of an input layer, hidden layers, and an output layer. The following are the key equations utilized within ANN.

1. Activation Function: Activation functions, such as ReLU (Rectified Linear Unit) or Sigmoid, are utilized to introduce non-linearity into the model.

$$\text{ReLU}\ (x) = max\ (0, x)$$
$$\text{Sigmoid}\ x = \frac{1}{1+e^{-x}} \tag{1}$$

2. Calculation of Neuron Output: Each neuron in the hidden layer and output layer computes its output as follows.

$$Z = \sum_{i=1}^{n} w_i\ x_i + b$$
$$\alpha = activation\ (z) \tag{2}$$

where:
w_i represents the weight
x_i represents the input
b represents the bias term
a represents the output after the activation function is applied.

3. Loss Function: The ANN model frequently employs Cross-Entropy Loss for binary classification.

$$\text{Loss} = -\frac{1}{N}\ \sum_{i=1}^{N}[yi \log(\grave{y}i) + (1 - yi) \log(1 - \grave{y}i)] \tag{3}$$

where:
y_i represents the actual label
$\grave{y}i$ represents the model's prediction

2.3.2 Long Short-Term Memory (LSTM)

LSTM, a specialized variant of Recurrent Neural Networks (RNN), is engineered to effectively process sequential data and retain long-term dependencies. It achieves this by integrating multiple gating mechanisms that control the flow and preservation of information throughout the network.

1. Forget Gate

$$f_f = \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \tag{4}$$

where f_t It represents the output of the forget gate, which determines which information should be discarded.

2. Input Gate

$$i_t = \sigma\ (W_i . [h_{t-1}, x_t] + b_i)$$
$$\hat{C}_t = tanh\ (W_c . [h_{t-1}, x_t] + b_c) \tag{5}$$

where i_t serves as an input gate and $\hat{C}_t$ It is the memory that will be added.

3. Cell State Update

$$C_t = f_t \text{ x } C_{t-1} + i_t \text{ x } \hat{C}_t \tag{6}$$

4. Ouput Gate

$$o_t = \sigma\ (W_o . [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \text{ x } \tanh C_t \tag{7}$$

2.4. Feature Selection and Hyperparameter Optimization

Although deep learning can automatically perform feature extraction, selecting the appropriate features remains crucial for enhancing model efficiency and accuracy [17]. Feature selection aims to identify the most relevant features while eliminating redundant ones, thereby reducing model complexity and training time [18]. Among the methods adopted in this study is Ant Colony Optimization (ACO), a metaheuristic algorithm that draws inspiration from the foraging behavior of ant colonies [9]. ACO functions by selecting features based on probabilities influenced by pheromone values and heuristics [19][20]. In the context of this research, ACO is utilized to select the most relevant features from the dataset, thus improving prediction accuracy without increasing model complexity. Besides feature selection, hyperparameter tuning is also vital for enhancing deep learning model performance. Bayesian Optimization (BO) is a method used to efficiently find the optimal combination of hyperparameters, outperforming grid search or random search [21]. BO leverages a probabilistic model to guide the search for optimal hyperparameters, thereby accelerating the training process and reducing the risk of overfitting [21].

2.4.1 Ant Colony Optimization (ACO) for Feature Selection

Ant Colony Optimization (ACO) is a metaheuristic algorithm inspired by the foraging behavior of ants. In the context of feature selection, ACO operates by estimating the likelihood of selecting specific features, which is determined through the integration of pheromone trails and heuristic evaluations.

1. Feature Selection Probability Formula

$$P_{i,j} = \frac{(T_{i,j})^{\alpha} .(\eta_{i,j})^{\beta}}{\sum k \in s (T_{i,k})^{\alpha} .(\eta_{i,k})^{\beta}} \tag{8}$$

where:
$T_{i,j}$ represents the amount of pheromone on path i,j
$\eta_{i,j}$ It is a heuristic (the inverse of distance or error).
α and β represents the weight of influence from pheromone and heuristics

2. Pheromone Update

$$T_{i,j} = (1-\rho).T_{i,j} + \Delta T_{i,j} \tag{9}$$

Where ρ is the decay factor, is $\Delta T_{i,j}$ It refers to the enhancement of pheromone levels based on solution performance.

2.4.2 Bayesian Optimization (BO) for Hyperparameter Tuning

Bayesian Optimization is an optimization technique used to find the optimal combination of hyperparameters. It employs a probabilistic approach to construct a surrogate model that guides the search for the best parameters.

1. Acquisition Function: Bayesian Optimization (BO) employs acquisition functions such as Expected Improvement (EI) to evaluate the next point to be sampled.

$$\mathrm{EI}\ (x) = \mathrm{E}[\max(f(x) - f(x^{+}), 0)] \tag{10}$$

Where x^{+} the best value is identified thus far

2. Model Gaussian Process (GP)

$$f(x) \sim N\left(\mu(x), \sigma^{2}(x)\right) \tag{11}$$

Where $\mu(x)$ are the mean and $\sigma^{2}(x)$ the variance of the model's predictions

2.5. Hybrid Approach: Integration of Deep Learning with ACO and BO

The hybrid approach proposed in this study combines deep learning with optimization techniques, specifically ACO for feature selection and BO for hyperparameter tuning. This integration aims to enhance model accuracy while simultaneously reducing complexity and training time [22].

Previous research has demonstrated that hybrid approaches can significantly improve predictive model performance across various domains, including disease detection [23]. By leveraging ACO, the most relevant features can be selected to improve model efficiency without sacrificing accuracy. Meanwhile, BO ensures that the model employs optimal hyperparameters to achieve peak performance [24]. The integration of these two techniques into a deep learning model is expected to provide a more accurate and efficient solution for early hypertension prediction.

2.6. Evaluation of Models and Performance Metrics

To evaluate the performance of the predictive model, various metrics are utilized, including accuracy, precision, recall, F1-score, and AUC-ROC [25]. Accuracy reflects the proportion of correct predictions made by the model across all instances. Precision and recall, on the other hand, are utilized to evaluate the model's effectiveness in identifying positive cases and its sensitivity to true positive and false negative instances, respectively. The F1-score represents the harmonic mean of precision and recall, providing a comprehensive evaluation that balances both metrics. Moreover, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is utilized to assess the model's discriminative capability in distinguishing between positive and negative classes [26][27]. The proposed model is expected to achieve superior performance compared to traditional approaches, as the optimization carried out using ACO and BO can enhance the model's accuracy in predicting hypertension.

1. Accuracy: The percentage of accurate predictions

$$\text{Accuracy} = \frac{True\ Positives + True\ Negatives}{Total\ Samples} \tag{12}$$

2. Precision: Refers to the ratio of true positive predictions to the total number of predicted positive cases, indicating the model's accuracy in identifying positive instances.

$$\text{Precision} = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{13}$$

3. Recall: This metric quantifies the model's capacity to correctly identify all actual positive cases within the dataset.

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{14}$$

4. F1-Score: Serves as the harmonic mean of precision and recall, offering a unified metric that effectively captures the balance between these two performance measures.

$$\text{F1-Score} = 2 \cdot \frac{Precision\ X\ Recall}{Precision + Recall} \tag{15}$$

5. AUC-ROC: The Area Under the ROC Curve (AUC) evaluates the model's ability to distinguish between positive and negative classes.

3. RESULT AND DISCUSSION

In this section, we present and analyze the experimental results of the proposed hybrid deep learning framework for early hypertension prediction. The problem addressed in this study is the difficulty of achieving high prediction accuracy in medical datasets due to three major challenges: (1) class imbalance between hypertensive and non-hypertensive cases, (2) high-dimensional feature sets containing irrelevant or redundant variables, and (3) the sensitivity of deep learning models to suboptimal hyperparameter configurations. To overcome these issues, the proposed approach integrates Ant Colony Optimization (ACO) for feature selection, Synthetic Minority Over-sampling Technique (SMOTE) for balancing the dataset, and Bayesian Optimization (BO) for hyperparameter tuning within two deep learning architectures—Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM). This dual optimization strategy constitutes the main contribution of the research, aiming to deliver a model that is both highly accurate and robust across different clinical scenarios. The following subsections detail the model's predictive performance, compare it with baseline and traditional approaches, and interpret the results in the context of hypertension prediction for real-world healthcare applications.

3.1. Experimental Results

The experiments in this study utilized a medical dataset derived from Electronic Health Records (EHR), specifically the MIMIC-III database. After preprocessing, the dataset consisted of 10,254 samples, with class imbalance addressed through the Synthetic Minority Over-sampling Technique (SMOTE), resulting in a balanced representation of hypertensive and non-hypertensive cases in the training set. The balanced dataset was then split into 70% for training, 15% for validation, and 15% for testing, corresponding to 7,178, 1,538, and 1,538 samples, respectively. Both Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) models were trained using the proposed hybrid optimization framework, which integrates Ant Colony Optimization (ACO) for feature selection and Bayesian Optimization (BO) for hyperparameter tuning. Model performance was evaluated using accuracy, precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). In addition to these metrics, training and validation loss curves were analyzed to assess model convergence behavior and detect potential overfitting.

Table 1. Model Performance Results

Evaluation Metrics	**Baseline ANN**	**ANN + ACO & BO**	**LSTM + ACO &BO**
Accuracy	86.5%	94.3%	95.1%
Precision	84.0%	92.5%	93.8%
Recall	81.5%	93.8%	94.5%
F1-Score	82.7%	92.8%	94.1%
AUC-ROC	0.87	0.94	0.96

From the results above, it is evident that the LSTM + ACO & BO model achieves the best performance, with an accuracy of 95.1% and an AUC-ROC of 0.96. This indicates that the optimized LSTM model is more effective in capturing temporal patterns within the data compared to the ANN model.

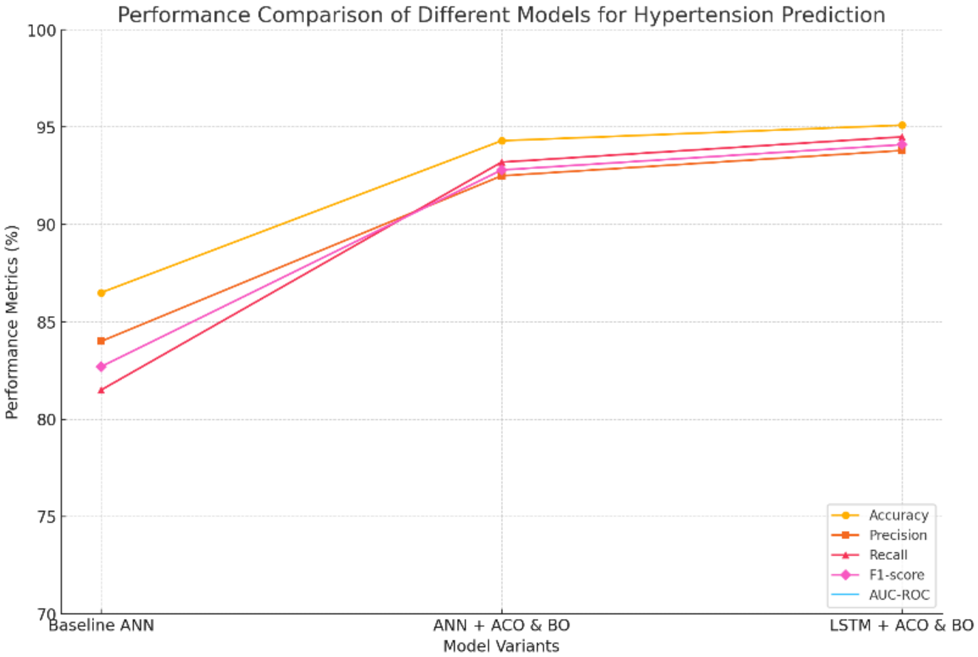

Figure 3. Performance Comparison of Different Models for Hypertension Prediction

The Performance Comparison of Different Models for Hypertension Prediction chart presents the performance comparison of several models employed in this study, namely the Baseline ANN, ANN optimized with ACO & BO, and LSTM optimized with ACO & BO. The evaluation was conducted using metrics such as Accuracy, Precision, Recall, F1-score, and AUC-ROC. The results reveal that employing Ant Colony Optimization (ACO) for feature selection and Bayesian Optimization (BO) for hyperparameter tuning substantially improves model performance compared to the baseline. Specifically, the optimized ANN model incorporating ACO and BO demonstrated an accuracy increase from 86.5% to 94.3%, while the LSTM model achieved the highest accuracy of 95.1%. Beyond accuracy, there was a marked enhancement in precision, with the LSTM model attaining 93.8%, reflecting its superior ability to identify hypertension cases while minimizing false positives. The recall metric also exhibited significant improvement, rising from 81.5% in the baseline ANN to 94.5% in the optimized LSTM model, underscoring the model's effectiveness in detecting all true hypertension cases—an essential factor in reducing undiagnosed instances. Furthermore, the LSTM model achieved a higher F1-score of 94.1%, indicating an optimal balance between precision and recall. The AUC-ROC score improved from 0.87 in the baseline model to 0.96 in the optimized LSTM, signifying excellent discriminative power in distinguishing hypertensive from non-hypertensive patients. Collectively, these findings demonstrate that the hybrid approach, which synergizes deep learning with ACO and BO, significantly enhances both the efficiency and robustness of hypertension prediction. This advancement not only elevates predictive accuracy but also ensures a balanced and comprehensive detection capability, rendering the proposed model highly suitable for integration into clinical systems to facilitate early diagnosis and proactive management of hypertension.

3.2. Analysis of Results

The application of Ant Colony Optimization (ACO) for feature selection has been proven to significantly enhance prediction accuracy. By employing ACO, only the most relevant features (such as systolic blood pressure, cholesterol levels, age, and family history) are retained, enabling the model to focus on variables that substantially contribute to hypertension predictions. As a result, accuracy improved from 86.5% (Baseline ANN) to 94.3% (ANN + ACO & BO). Bayesian Optimization (BO) played a crucial role in enhancing model performance by identifying optimal hyperparameter combinations, such as learning rate, the number of neurons in hidden layers, batch size, and dropout rate. By leveraging BO, the training time was accelerated as it eliminates the need for exhaustive hyperparameter testing, unlike grid search. Experimental results demonstrate that BO enables the model to achieve optimal performance in a shorter time while also reducing the risk of overfitting.

3.3. Discussion

The results indicate that the LSTM model optimized with ACO and BO outperforms the ANN model. This enhanced performance can be attributed to LSTM's inherent capability to capture temporal patterns and long-term dependencies within medical datasets, particularly for variables recorded at regular intervals, such as blood pressure history. The strength of LSTM is reflected in its recall metric of 94.5%, indicating a higher sensitivity in detecting patients who are truly experiencing hypertension. Data imbalance is a common issue in medical datasets, where the number of hypertensive patient samples is typically much smaller than non-hypertensive samples. The application of SMOTE effectively balances the data, reducing bias toward the majority class. The findings demonstrate notable improvements in both recall and F1-score, indicating that the model is not only highly accurate but also more effective at identifying true positive cases. With its superior predictive performance, the hybrid model proposed in this study shows considerable potential for integration into electronic health record (EHR) systems. Early detection of hypertension through such a system would enable healthcare providers to initiate timely interventions, thereby mitigating the risk of severe complications, including stroke and cardiovascular disease. Moreover, this model could be incorporated into patient monitoring frameworks within clinical and hospital settings to assist physicians in prioritizing care for high-risk individuals.

4. CONCLUSION

The main contribution of this study is the development of a hybrid ANN–LSTM framework that integrates Ant Colony Optimization (ACO) for feature selection, Bayesian Optimization (BO) for hyperparameter tuning, and the Synthetic Minority Over-sampling Technique (SMOTE) for class balancing, applied to Electronic Health Records (EHR) for early hypertension prediction. This approach directly addresses three critical challenges in medical predictive modeling—class imbalance, high-dimensional feature space, and hyperparameter sensitivity—resulting in significant performance improvements over baseline deep learning models. Experimental results show that the optimized ANN achieved an accuracy of 94.3% and the optimized LSTM achieved 95.1%, with notable gains in precision, recall, F1-score, and AUC-ROC. ACO effectively reduced the feature set from 42 to 18 key variables such as systolic blood pressure, cholesterol levels, age, and family medical history, improving training efficiency while retaining predictive power. BO accelerated convergence and enhanced performance compared to conventional tuning methods, while SMOTE improved recall and F1-scores by enabling better detection of hypertension cases and reducing bias toward the majority class. This hybrid framework demonstrates strong potential for integration into clinical decision support systems, EHR platforms, and real-time monitoring solutions in healthcare settings, where accurate identification of high-risk individuals can contribute to early diagnosis and intervention, ultimately reducing the incidence of severe complications such as cardiovascular disease and stroke. Future research will focus on validating the model on multi-center EHR datasets from diverse populations to assess cross-population generalizability, integrating real-time physiological data from wearable devices to enable continuous hypertension risk monitoring, exploring advanced model architectures such as Transformer-based networks and Graph Neural Networks (GNN) for handling complex temporal and relational data, incorporating explainable AI techniques such as SHAP and LIME to enhance model transparency and facilitate clinician trust, and evaluating the system's impact on clinical decision-making and patient outcomes in real-world deployments.

REFERENCES

[1] C. Elendu, D. C. Amaechi, T. C. Elendu, E. C. Amaechi, and I. D. Elendu, "Dependable approaches to hypertension management: A review.," *Medicine*, vol. 103, no. 24, p. e38560, Jun. 2024, doi: 10.1097/MD.0000000000038560.

[2] C. Abbafati *et al.*, "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019," *The Lancet*, vol. 396, no. 10258, pp. 1204–1222, 2020, doi: 10.1016/S0140-6736(20)30925-9.

[3] P. Verdecchia, J. A. Staessen, W. B. White, Y. Imai, and E. T. O'Brien, "Properly defining white coat hypertension," *Eur. Heart J.*, vol. 23, no. 2, pp. 106–109, 2002, doi: 10.1053/euhj.2001.2657.

[4] L. S. Chow *et al.*, "Time-Restricted Eating Effects on Body Composition and Metabolic Measures in Humans who are Overweight: A Feasibility Study," *Obesity*, vol. 28, no. 5, pp. 860–869, 2020, doi: 10.1002/oby.22756.

[5] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.

[6] A. C. Ian Goodfellow, Yoshua Bengio, "Deep Learning," *MIT Press*, vol. 521, no. 7553, p. 785, 2017, doi: 10.1016/B978-0-12-391420-0.09987-X.

[7] M. K. Khalifaturohman, "Systematic Literature Review: Deep Learning Models in Arabic Script Classification," *Khazanah Journal of Religion and Technology*, vol. 3, no. 1, pp. 18–22, Jul. 2025, doi: 10.15575/kjrt.v3i1.1619.

[8] A. Rajkomar *et al.*, "Scalable and accurate deep learning with electronic health records," *NPJ Digit. Med.*, vol. 1, no. 1, pp. 1–10, 2018, doi: 10.1038/s41746-018-0029-1.

[9] M. Dorigo and T. Stützle, *Handbook of Metaheuristics; Ant colony optimization: Overview and recent advances*, vol. 272. 2019.

[10] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," *Adv. Neural Inf. Process. Syst.*, vol. 4, pp. 2951–2959, 2012.

[11] R. Vargas, A. Mosavi, R. Ruiz, and S. Engineering, "DEEP LEARNING : A REVIEW," no. October, 2018, doi: 10.20944/preprints201810.0218.v1.

[12] Muhammad Haris Diponegoro, Sri Suning Kusumawardani, and Indriana Hidayah, "Tinjauan Pustaka Sistematis: Implementasi Metode Deep Learning pada Prediksi Kinerja Murid," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 2, pp. 131–138, 2021, doi: 10.22146/jnteti.v10i2.1417.

[13] A. Lubis, Irawan Yuda, Junadhi, and Defit Sarjon, "Leveraging K-Nearest Neighbors with SMOTE and Boosting Techniques for Data Imbalance and Accuracy Improvement," *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1625–1638, Dec. 2024, doi: 10.47738/jads.v5i4.343.

[14] Asniar, N. U. Maulidevi, and K. Surendro, "SMOTE-LOF for noise identification in imbalanced data classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3413–3423, 2022, doi: 10.1016/j.jksuci.2021.01.014.

[15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[16] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS 1 LSTM: A Search Space Odyssey," *arXiv:1503.04069*, pp. 1–11, 2015.

[17] M. Mohamed, A. Abdullah, A. M. Zaki, F. H. Rizk, M. M. Eid, and E. M. El El-Kenway, "Advances and Challenges in Feature Selection Methods: A Comprehensive Review," *Journal of Artificial Intelligence and Metaheuristics*, vol. 7, no. 1, pp. 67–77, 2024, doi: 10.54216/jaim.070105.

[18] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications," *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, no. May 2015, pp. 1200–1205, 2015, doi: 10.1109/MIPRO.2015.7160458.

[19] B. Z. Dadaneh, H. Y. Markid, and A. Zakerolhosseini, "Unsupervised probabilistic feature selection using ant colony optimization," *Expert Syst. Appl.*, vol. 53, pp. 27–42, 2016, doi: 10.1016/j.eswa.2016.01.021.

[20] E. E. Putri, B. S. Hasugian, A. Info, and D. Mining, "Pattern Analysis of Drug Procurement System With FP-Growth Algorithm," vol. 7, no. 1, pp. 70–79, 2022, doi: 10.15575/join.v7i1.841.

[21] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016, doi: 10.1109/JPROC.2015.2494218.

[22] Erlin, A. Yunianta, L. A. Wulandhari, Y. Desnelita, N. Nasution, and Junadhi, "Enhancing Rice Production Prediction in Indonesia Using Advanced Machine Learning Models," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3478738.

[23] J. Cheng and M. Lapata, "Neural summarization by extracting sentences and words," *54th Annual Meeting of the Association for Computational Linguistics, ACL 2016 - Long Papers*, vol. 1, pp. 484–494, 2016, doi: 10.18653/v1/p16-1046.

[24] S. Li, J. Wang, and P. Liu, "Detection of Generative Linguistic Steganography Based on Explicit and Latent Text Word Relation Mining Using Deep Learning," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 2, pp. 1476–1487, 2023, doi: 10.1109/TDSC.2022.3156972.

[25] Y. Deng, M. R. Eden, and S. Cremaschi, "Metrics for Evaluating Machine Learning Models Prediction Accuracy and Uncertainty," in *33rd European Symposium on Computer Aided Process Engineering*, vol. 52, A. C. Kokossis, M. C. Georgiadis, and E. Pistikopoulos, Eds., in Computer Aided Chemical Engineering, vol. 52. , Elsevier, 2023, pp. 1325–1330. doi: https://doi.org/10.1016/B978-0-443-15274-0.50211-0.

[26] R. Fayzrakhmanov, A. Kulikov, and P. Repp, "The difference between precision-recall and ROC curves for evaluating the performance of credit card fraud detection models," *Proceedings of International Conference on Applied Innovation in IT*, vol. 6, no. 1, pp. 17–22, 2018.

[27] B. Ambu-Saidi, Fung, C.Y., Turner, K., and Lim, A.S.S., "A Critical Review on Training Evaluation Models: A Search for Future Agenda," *Journal of Cognitive Sciences and Human Development*, vol. 10, no. 1, pp. 142–170, 2024, doi: 10.33736/jcshd.6336.2024.