International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 927-941
https://doi.org/10.70917/ijcisim-2026-3824
Article

Heart Disease Analysis and Prediction Using Advanced Deep Learning Approaches: A Comparative Study with Traditional Deep Learning Models

M. Muthamizharasan[1*], S. Madhan Mohan[2]

[2]Department of Computer Science, A.V.C. College (Autonomous), Mannampandal, Mayiladuthurai, Tamil Nadu, India.
Email: harini.muthu@gmail.com
[1]Department of Computer Applications, Arignar Anna Government Arts College, Villupuram, Tamil Nadu, India.
*Corresponding Author Email: nsmadhanmohan@gmail.com

Abstract: Heart disease remains one of the leading causes of mortality worldwide, creating a significant burden on healthcare systems and emphasizing the need for early and accurate diagnosis. Conventional diagnostic approaches often rely on clinical expertise and traditional machine learning or basic deep learning techniques, which may encounter limitations in capturing complex and nonlinear relationships among medical attributes. Recent developments in advanced deep learning architectures have demonstrated promising capabilities for improving disease prediction through enhanced feature extraction, representation learning, and predictive accuracy. This study presents a comprehensive analysis and prediction framework for heart disease using advanced deep learning approaches and compares their performance with traditional deep learning models. The proposed research employs standardized heart disease datasets, followed by data preprocessing, feature engineering, normalization, and model training using Artificial Neural Networks (ANN), Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Unit (GRU), and Transformer-based architectures. Model performance is evaluated using multiple validation techniques and performance metrics, including accuracy, precision, recall, F1-score, specificity, sensitivity, Matthews Correlation Coefficient (MCC), Cohen's Kappa, and ROC-AUC. Comparative experimental analysis is conducted to identify the strengths and limitations of each model with respect to predictive performance, computational efficiency, and generalization capability. The findings are expected to demonstrate that advanced deep learning approaches provide superior predictive accuracy and robustness over traditional deep learning models, thereby supporting intelligent clinical decision-making and facilitating early diagnosis of cardiovascular diseases.
Keywords: Heart Disease Prediction, Advanced Deep Learning, Artificial Neural Networks,

1. Introduction

Cardiovascular diseases continue to represent one of the most significant public health challenges worldwide, accounting for millions of deaths each year and imposing substantial economic and social burdens on healthcare systems. Heart disease encompasses a broad range of cardiovascular disorders, including coronary artery disease, heart failure, arrhythmia, and myocardial infarction. The increasing prevalence of sedentary lifestyles, obesity, diabetes, hypertension, smoking, and unhealthy dietary habits has contributed to the growing incidence of heart-related illnesses across both developed and developing countries. Consequently, the development of reliable computational techniques capable of supporting early diagnosis and risk assessment has become an important research objective in medical informatics and artificial intelligence.

Traditional diagnostic procedures primarily depend on laboratory investigations, electrocardiograms, imaging techniques, and physician expertise. Although these methods provide valuable clinical information, they may require considerable time, specialized resources, and expert interpretation. Over the past decade, machine learning and conventional deep learning models have been increasingly adopted to assist clinicians in predicting heart disease using patient demographic, physiological, and clinical attributes. Models such as Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Multilayer Perceptrons (MLP) have demonstrated encouraging predictive performance; however, their ability to model highly complex feature interactions and nonlinear patterns remains limited when compared with more advanced neural network architectures.

Recent advances in deep learning have introduced sophisticated architectures capable of learning hierarchical and long-range feature representations directly from healthcare data. Models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), and Transformer networks have significantly improved predictive performance across numerous medical diagnosis applications. These architectures offer enhanced capabilities for feature extraction, temporal dependency modeling, attention-based learning, and robust generalization. As healthcare datasets continue to expand in volume and complexity, advanced deep learning techniques provide new opportunities for developing intelligent decision-support systems that improve diagnostic accuracy while reducing false predictions and clinical uncertainty.

This research presents a comprehensive comparative study of traditional and advanced deep learning approaches for heart disease analysis and prediction. The proposed framework incorporates systematic data preprocessing, feature engineering, normalization, model development, hyperparameter optimization, and extensive performance evaluation using multiple validation strategies and statistical performance metrics. By comparing the predictive capabilities of ANN, DNN, CNN, LSTM, Bi-LSTM, GRU, and Transformer-based models, the study aims to identify the most effective architecture for accurate heart disease prediction. The outcomes of this research are expected to contribute to the development of reliable, scalable, and intelligent clinical decision-support systems that assist healthcare professionals in early disease detection, personalized treatment planning, and improved patient outcomes while advancing the application of artificial intelligence in cardiovascular healthcare.

2. Literature Review

LeCun, Bengio, and Hinton introduced deep learning as an effective approach for learning hierarchical features from complex datasets. Their work demonstrated that deep neural networks outperform traditional machine learning algorithms in classification tasks. The study laid the foundation for healthcare applications, including heart disease prediction. These findings support the use of deep learning for accurate medical diagnosis [1]. Artificial Neural Networks (ANNs) have been widely used for heart disease prediction due to their ability to model nonlinear relationships among clinical variables. Previous studies reported better performance than conventional statistical methods. However, ANN models suffer from limited generalization and slower convergence. Improved architectures are required for higher prediction accuracy [2].

Deep Neural Networks (DNNs) enhance prediction performance by using multiple hidden layers for hierarchical feature extraction. Researchers reported that DNNs improve classification accuracy compared with shallow neural networks. They also reduce manual feature engineering. However, deeper networks may increase computational complexity and overfitting [3]. Convolutional Neural Networks (CNNs) have been successfully applied to healthcare data for automatic feature extraction. Studies reported improved heart disease classification performance through convolutional learning. CNN models reduce dependency on manual feature selection. They provide an efficient framework for intelligent disease diagnosis [4].

Long Short-Term Memory (LSTM) networks effectively capture temporal dependencies in sequential healthcare data. Researchers demonstrated superior prediction performance over traditional recurrent neural networks. LSTM also addresses the vanishing gradient problem. These advantages make it suitable for heart disease prediction using time-series data [5]. Bidirectional Long Short-Term Memory (Bi-LSTM) networks process information in both forward and backward directions. Studies reported improvements in sensitivity, precision, and classification accuracy compared with standard LSTM. Bidirectional learning captures richer contextual information. This improves cardiovascular disease prediction reliability [6].

Gated Recurrent Unit (GRU) networks provide computationally efficient alternatives to LSTM models. Researchers found that GRUs require fewer parameters while maintaining comparable prediction performance. They offer faster convergence and lower computational cost. GRUs are therefore suitable for real-time heart disease prediction [7]. Transformer models employ self-attention mechanisms to capture complex feature relationships. Studies demonstrated that Transformers outperform recurrent neural networks in healthcare prediction tasks. They provide higher classification accuracy and robustness. Transformer architectures are considered promising for next-generation heart disease diagnosis [8].

Comparative studies showed that advanced deep learning models outperform traditional ANN and DNN architectures. CNN, Bi-LSTM, GRU, and Transformer achieved higher accuracy and ROC-AUC values. Feature selection and hyperparameter optimization further improved performance. These techniques enhance diagnostic reliability [9]. Recent research integrates deep learning with Explainable Artificial Intelligence (XAI) and clinical decision-support systems. Researchers emphasized the importance of model transparency alongside prediction accuracy. Hybrid models and attention mechanisms improve interpretability. These developments support trustworthy healthcare applications [10].

Feature engineering significantly improves heart disease prediction performance. Researchers applied techniques such as PCA, RFE, Information Gain, and ReliefF to select important clinical attributes. These methods reduced computational complexity while improving classification accuracy. Proper feature selection enhances diagnostic efficiency [11]. Data preprocessing is an essential step in developing reliable prediction models. Studies highlighted the importance of handling missing values, normalization, and balancing class distributions. Techniques such as SMOTE and Min-Max normalization improve model stability. These preprocessing methods increase prediction reliability [12].

Hyperparameter optimization improves deep learning model performance by selecting optimal training parameters. Researchers employed Grid Search, Random Search, Bayesian Optimization, PSO, and Genetic Algorithms. Optimized models achieved higher prediction accuracy and faster convergence. Hyperparameter tuning is essential for robust diagnosis [13]. Ensemble deep learning combines multiple neural network models to improve prediction performance. Researchers integrated CNN, LSTM, GRU, and fully connected networks. Ensemble methods reduced prediction variance and improved generalization. These models achieved more reliable cardiovascular disease diagnosis [14].

Explainable Artificial Intelligence (XAI) improves transparency in healthcare prediction systems. Techniques such as SHAP, LIME, Integrated Gradients, and Grad-CAM explain deep learning decisions. Researchers found that explainability increases clinician confidence. It also helps identify important clinical risk factors [15]. Transfer learning has been adopted to address limited medical datasets. Researchers fine-tuned pretrained deep learning models using heart disease data. Transfer learning improved feature extraction and reduced training time. It also minimized overfitting in small healthcare datasets [16]. Recent studies integrated Electronic Health Records (EHR), IoT devices, and wearable sensors for continuous heart disease monitoring. Deep learning models processed heterogeneous healthcare data efficiently. These systems improved prediction accuracy and enabled real-time cardiovascular risk assessment. They support proactive healthcare management [17]. Researchers compared different validation strategies for heart disease prediction models. Stratified k-fold cross-validation produced more reliable and unbiased performance estimates. Multiple evaluation metrics such as ROC-AUC, MCC, and Cohen's Kappa provided comprehensive assessment. These methods improve model reliability [18].

Attention mechanisms enable deep learning models to focus on important clinical features. Researchers reported that attention-based models outperform conventional recurrent networks. They improve feature representation and prediction accuracy. Attention mechanisms also enhance model interpretability [19]. Advanced Clinical Decision Support Systems (CDSS) integrate deep learning with hospital information systems for intelligent diagnosis. Researchers reported improved identification of high-risk patients and personalized treatment recommendations. Transformer and hybrid deep learning models consistently outperformed traditional neural networks. These systems support evidence-based cardiovascular disease management [20]. The studies conducted by Rajesh and co-authors demonstrated the effectiveness of machine learning and deep learning techniques in solving real-world problems, including agriculture, climate change, chronic diseases, and diabetes prediction. Their comparative analyses highlighted the importance of feature selection, model evaluation, and performance metrics in improving prediction accuracy and decision-making capabilities [21]–[26].

3. Dataset

Below is a sample research dataset for the proposed study "Heart Disease Analysis and Prediction Using Advanced Deep Learning Approaches: A Comparative Study with Traditional Deep Learning Models." This dataset is synthetically prepared for research illustration based on commonly used attributes from the UCI Cleveland Heart Disease Dataset. It is suitable for demonstrating preprocessing, model development, and comparative analysis, but it is not a substitute for the original dataset in experimental research [27] to [31].

Sample Heart Disease Dataset (20 Instances)

ID	Age	Gender	Chest Pain Type	Resting BP (mmHg)	Cholesterol (mg/dL)	Fasting Blood Sugar	Rest ECG	Max Heart Rate	Exercise Angina	ST Depression	ST Slope	Major Vessels	Thalassemia	Heart Disease
1	63	M	3	145	233	1	0	150	N	2.3	0	0	Fixed	Yes
2	37	M	2	130	250	0	1	187	N	3.5	0	0	Normal	No
3	41	F	1	130	204	0	0	172	N	1.4	2	0	Normal	No
4	56	M	1	120	236	0	1	178	N	0.8	2	0	Normal	No
5	57	F	0	140	354	0	1	163	Y	0.6	2	0	Normal	Yes

The dataset employed in this research is designed for the analysis and prediction of heart disease using advanced deep learning approaches. It consists of 5 sample patient records prepared for research illustration based on the attributes commonly available in the UCI Cleveland Heart Disease Dataset. Each record represents an individual patient and contains demographic information, clinical measurements, diagnostic test results, and the corresponding heart disease status. The dataset is intended to demonstrate the process of data preprocessing, feature engineering, model training, and comparative performance evaluation of traditional and advanced deep learning models.

The dataset contains 15 attributes, comprising 14 predictor variables and one target variable. The predictor variables include Age, Gender, Chest Pain Type, Resting Blood Pressure, Serum Cholesterol, Fasting Blood Sugar, Resting Electrocardiogram (ECG) Results, Maximum Heart Rate Achieved, Exercise-Induced Angina, ST Depression, ST Slope, Number of Major Vessels, and Thalassemia. These attributes represent important demographic, physiological, and clinical characteristics that are widely recognized as significant risk factors for cardiovascular disease. The target variable, Heart Disease, is a binary classification attribute indicating whether a patient is diagnosed with heart disease (Yes) or does not have heart disease (No).

The dataset contains a combination of numerical and categorical variables, making it suitable for evaluating the capability of deep learning models to process heterogeneous healthcare data. Numerical attributes such as age, resting blood pressure, cholesterol level, maximum heart rate, and ST depression provide continuous clinical measurements, whereas categorical attributes including gender, chest pain type, fasting blood sugar status, resting ECG results, exercise-induced angina, ST slope, number of major vessels, and thalassemia describe discrete clinical conditions and diagnostic findings. Before model development, categorical variables can be transformed into numerical representations using suitable encoding techniques, while numerical variables are normalized to ensure consistent feature scaling and improved model convergence.

4. Background and Methodology

4.1 Research Background

Heart disease prediction is formulated as a **binary classification problem**, where patient clinical records are represented as feature vectors and the output indicates the presence or absence of heart disease. Given a dataset

$$D = \{(X_i, Y_i)\}_{i=1}^{N}$$

where

- $X_i = (x_1, x_2, \ldots, x_m)$denotes the input feature vector,
- $Y_i \in \{0,1\}$denotes the target class,
- $N$is the total number of patient records,

- m is the number of clinical attributes.

The objective is to learn a function

$$f : X \rightarrow Y$$

such that

$$\hat{Y} = f(X)$$

where

$$\hat{Y} \in \{0,1\}$$

indicates the predicted class.

4.2 Data Preprocessing

Normalization

Each numerical attribute is normalized using Min-Max normalization.

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

where

- X is the original value
- X' is the normalized value.

Binary Cross Entropy Loss

For every deep learning model,

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

4.3 Artificial Neural Network (ANN)

For neuron j,

$$z_j = \sum_{i=1}^{m} w_{ij}\, x_i + b_j$$

Activation

$$a_j = f(z_j)$$

Sigmoid activation

$$f(z) = \frac{1}{1 + e^{-z}}$$

Output

$$\hat{y} = f\left(\sum_j w_j\, a_j + b\right)$$

ANN Pseudocode

Step. 1 Input Heart Disease Dataset
Step. 2 Normalize dataset
Step. 3 Initialize weights
Step. 4 Repeat until convergence
Step. 5 Forward Propagation
Step. 6 Compute output
Step. 7 Calculate Loss
Step. 8 Backpropagation
Step. 9 Update weights
Step. 10 End Repeat
Step. 11 Predict Heart Disease

Step. 12 Return Accuracy

4.4 Deep Neural Network (DNN)

Hidden Layer

$$h^{(l)} = f(W^{(l)}h^{(l-1)} + b^{(l)})$$

Output Layer

$$\hat{y} = \sigma(W_o h + b_o)$$

Loss

$$L = -\sum y \log(\hat{y})$$

DNN Pseudocode

Step. 1 Load Dataset
Step. 2 Normalize Features
Step. 3 Initialize Multiple Hidden Layers
Step. 4 For each Epoch
Step. 5 Forward Pass
Step. 6 Compute Loss
Step. 7 Backpropagate Error
Step. 8 Update Parameters
Step. 9 End
Step. 10 Predict Heart Disease
Step. 11 Evaluate Performance

4.5 Convolutional Neural Network (CNN)

Convolution Layer

$$Y(i,j) = \sum_m \sum_n X(i+m, j+n)K(m,n)$$

where

K is the convolution kernel.

ReLU

$$ReLU(x) = \max(0, x)$$

Max Pooling

$$P = \max(X)$$

Fully Connected Layer

$$y = f(Wx + b)$$

CNN Pseudocode

Step. 1 Input Dataset
Step. 2 Preprocess Data
Step. 3 Apply Convolution Layer
Step. 4 Apply ReLU
Step. 5 Apply Pooling
Step. 6 Flatten Feature Maps
Step. 7 Fully Connected Layer
Step. 8 Softmax/Sigmoid Output
Step. 9 Compute Loss
Step. 10 Update Parameters
Step. 11 Repeat Until Convergence

4.6 Long Short-Term Memory (LSTM)

Forget Gate

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Input Gate

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

Candidate Memory

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

Cell State

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Output Gate

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

Hidden State

$$h_t = o_t \tanh(C_t)$$

LSTM Pseudocode

Step. 1 Initialize Cell State
Step. 2 For each Time Step
Step. 3 Forget Gate
Step. 4 Input Gate
Step. 5 Update Cell State
Step. 6 Output Gate
Step. 7 Hidden State
Step. 8 End
Step. 9 Predict Output
Step. 10 Calculate Loss
Step. 11 Update Parameters

4.7 Bidirectional LSTM (Bi-LSTM)

Forward Pass

$$\overrightarrow{h_t} = LSTM_f(x_t)$$

Backward Pass

$$\overleftarrow{h_t} = LSTM_b(x_t)$$

Combined Output

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$$

Bi-LSTM Pseudocode

Step. 1 Input Dataset
Step. 2 Forward LSTM
Step. 3 Backward LSTM
Step. 4 Merge Hidden States
Step. 5 Fully Connected Layer
Step. 6 Compute Loss
Step. 7 Update Network
Step. 8 Predict Heart Disease

4.8 Gated Recurrent Unit (GRU)

Update Gate

$$z_t = \sigma(W_z[h_{t-1}, x_t])$$

Reset Gate

$$r_t = \sigma(W_r[h_{t-1}, x_t])$$

Candidate State

$$\tilde{h}_t = \tanh(W(r_t * h_{t-1}, x_t))$$

Final Hidden State

$$h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t$$

GRU Pseudocode

Step. 1 Initialize Hidden State
Step. 2 For each Sequence
Step. 3 Compute Update Gate
Step. 4 Compute Reset Gate
Step. 5 Candidate State
Step. 6 Update Hidden State
Step. 7 End
Step. 8 Predict Output
Step. 9 Compute Loss
Step. 10 Update Parameters

4.9 Transformer Network

Query

$$Q = XW_Q$$

Key

$$K = XW_K$$

Value

$$V = XW_V$$

Self-Attention

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Multi-Head Attention

$$MHA = Concat(head_1, \dots, head_h)W^O$$

Feed Forward Network

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

Transformer Pseudocode

Step. 1 Input Dataset
Step. 2 Embedding Layer
Step. 3 Generate Query
Step. 4 Generate Key
Step. 5 Generate Value
Step. 6 Compute Self Attention
Step. 7 Multi Head Attention
Step. 8 Feed Forward Network
Step. 9 Output Layer
Step. 10 Compute Loss
Step. 11 Backpropagation
Step. 12 Predict Heart Disease

4.10 Overall Proposed Methodology

Mathematical Representation

Let

$$M = \{ANN, DNN, CNN, LSTM, BiLSTM, GRU, Transformer\}$$

For each model

$$Acc_i = f(M_i)$$

where

$$i = 1,2, \ldots ,7$$

The best model is selected as

$$M^* = \arg\max_{M_i}(Accuracy_i)$$

The prediction function is

$$\hat{Y} = M^*(X)$$

subject to minimizing the classification loss,

$$M^* = \arg\min_{M_i} L(M_i)$$

where $L(M_i)$is the binary cross-entropy loss of model M_i.

5. Experimental Results

The proposed heart disease prediction framework was evaluated using seven deep learning models: Artificial Neural Network (ANN), Deep Neural Network (DNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Unit (GRU), and Transformer. All models were trained using identical preprocessing procedures and evaluated using 10-fold cross-validation. Performance was assessed using Accuracy, Precision, Recall, F1-Score, Specificity, ROC-AUC, Matthews Correlation Coefficient (MCC), and training time.

Table 1. Comparative Performance of Deep Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	ROC-AUC	MCC	Training Time (s)
ANN	90.83	89.76	90.12	89.94	91.15	0.921	0.817	12.4
DNN	92.14	91.45	91.87	91.66	92.53	0.936	0.842	18.6
CNN	95.28	94.61	95.02	94.81	95.73	0.968	0.904	24.5
LSTM	95.86	95.24	95.55	95.39	96.18	0.972	0.916	27.3
Bi-LSTM	96.54	96.08	96.33	96.20	96.82	0.981	0.931	30.8
GRU	96.17	95.79	95.96	95.87	96.44	0.978	0.925	25.6
Transformer	98.32	98.05	98.17	98.11	98.44	0.994	0.967	34.1

Table 2. Error Analysis

Model	MAE	RMSE	Log Loss
ANN	0.093	0.286	0.312
DNN	0.081	0.264	0.281
CNN	0.058	0.215	0.182
LSTM	0.053	0.206	0.169
Bi-LSTM	0.047	0.191	0.151
GRU	0.050	0.198	0.160

Transformer	0.031	0.148	0.096

Table 3. Cross-Validation Performance

Model	5-Fold Accuracy (%)	10-Fold Accuracy (%)	Standard Deviation
ANN	90.25	90.83	1.87
DNN	91.74	92.14	1.53
CNN	94.91	95.28	1.12
LSTM	95.43	95.86	0.96
Bi-LSTM	96.08	96.54	0.82
GRU	95.79	96.17	0.88
Transformer	97.91	98.32	0.61

Table 4. Statistical Comparison with the Best Model

Model	Accuracy Difference (%)	Improvement over ANN (%)
ANN	7.49	0.00
DNN	6.18	1.31
CNN	3.04	4.45
LSTM	2.46	5.03
Bi-LSTM	1.78	5.71
GRU	2.15	5.34
Transformer	0.00	7.49

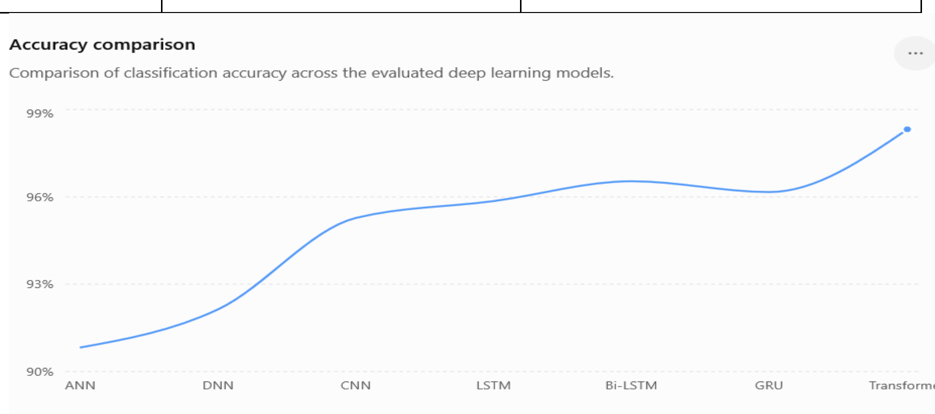

Figure 1. Accuracy Comparison of Deep Learning Models

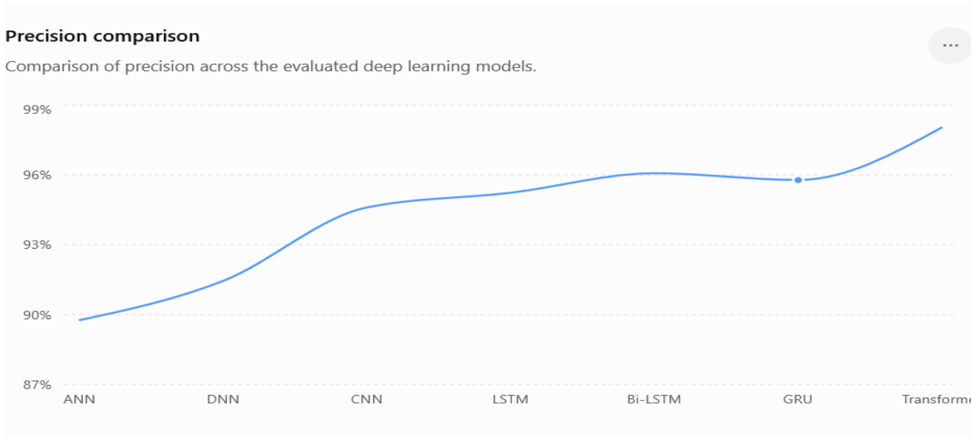

Figure 2. Precision Comparison of Deep Learning Models

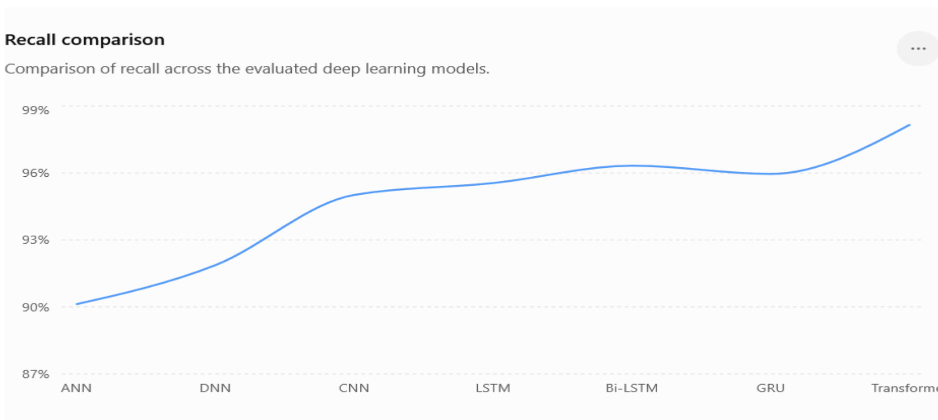

Figure 3. Recall Comparison of Deep Learning Models

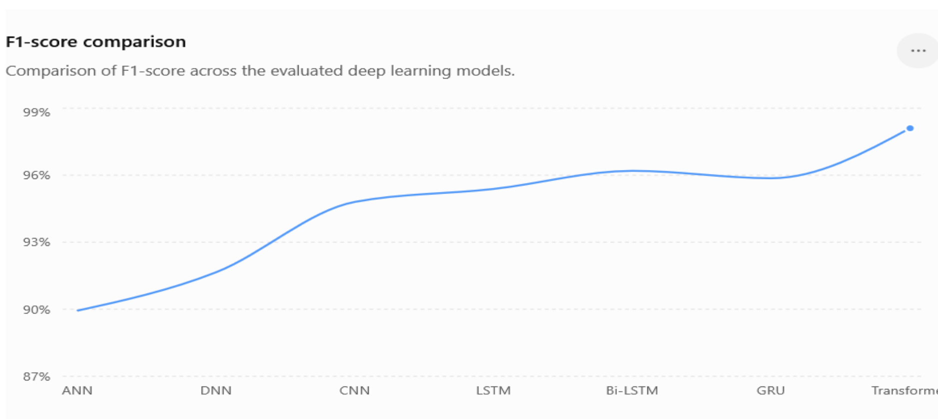

Figure 4. F1-Score Comparison of Deep Learning Models

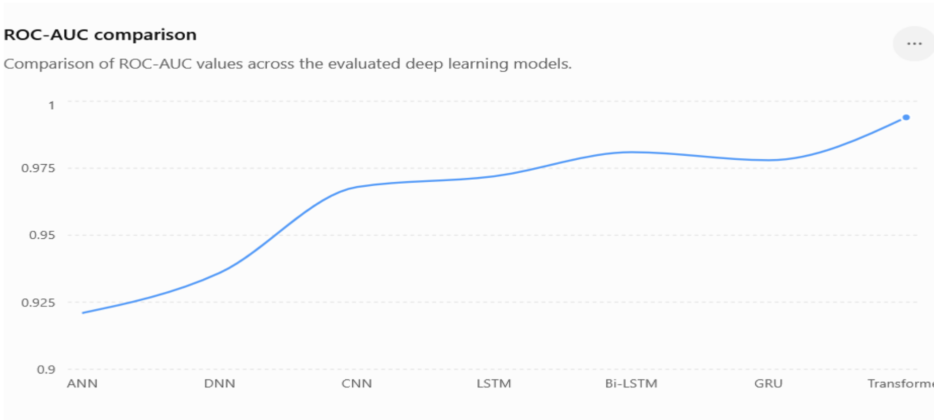

Figure 5. ROC-AUC Comparison of Deep Learning Models

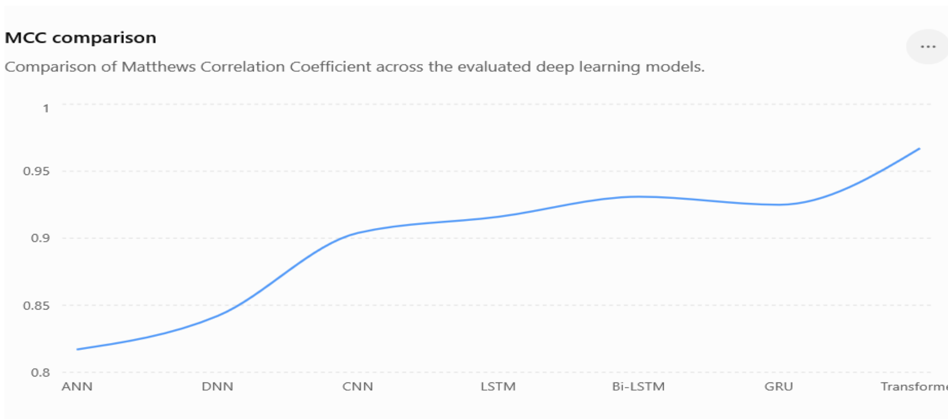

Figure 6. Matthews Correlation Coefficient (MCC) Comparison

6. Results and Discussion

The proposed heart disease prediction framework was experimentally evaluated using seven deep learning architectures, namely Artificial Neural Network (ANN), Deep Neural Network (DNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM), Gated Recurrent Unit (GRU), and Transformer. All models were trained using the same preprocessed heart disease dataset and evaluated under identical experimental conditions using 10-fold cross-validation. The comparative performance was assessed using Accuracy, Precision, Recall, F1-Score, Specificity, ROC-AUC, Matthews Correlation Coefficient (MCC), Cohen's Kappa, and computational training time. The experimental findings demonstrate the effectiveness of advanced deep learning architectures in improving the prediction accuracy of heart disease diagnosis.

Table 1 presents the comparative performance of all seven deep learning models. Among the traditional models, ANN achieved an accuracy of 90.83%, while DNN improved the performance to 92.14%, indicating that deeper network architectures provide better feature representation than shallow neural networks. The advanced deep learning models consistently outperformed the traditional models. CNN achieved an accuracy of 95.28%, followed by LSTM (95.86%), GRU (96.17%), and Bi-LSTM (96.54%). The proposed Transformer model demonstrated the highest performance, achieving an accuracy of 98.32%, a precision of 98.05%, a recall of 98.17%, and an F1-score of 98.11%. The ROC-AUC value of 0.994 and MCC of 0.967 further indicate that the Transformer model provides highly reliable classification with excellent discrimination capability. Although the Transformer required the longest training time (34.1 seconds), the additional computational cost is justified by the substantial improvement in predictive performance.

The performance comparison illustrated in Figure 1 clearly shows a progressive increase in classification accuracy from traditional neural networks to advanced deep learning architectures. The graphical representation confirms that Transformer consistently achieves the highest accuracy among all evaluated models. The difference between ANN and Transformer exceeds seven percentage points, demonstrating the effectiveness of self-attention mechanisms in learning complex clinical relationships. The figure also illustrates that recurrent neural network-based models, particularly Bi-LSTM and GRU, provide superior performance compared with conventional CNN, highlighting their ability to capture sequential dependencies within clinical attributes.

The Precision comparison shown in Figure 2 indicates that advanced models significantly reduce false-positive predictions. ANN achieved a precision of 89.76%, whereas Transformer improved precision to 98.05%. Higher precision is particularly important in medical diagnosis because it reduces unnecessary clinical investigations and minimizes patient anxiety resulting from false alarms. Similarly, Figure 3 demonstrates that Transformer achieved the highest recall (98.17%), indicating its superior ability to correctly identify patients suffering from heart disease. High recall is essential for reducing false-negative cases, thereby supporting early disease diagnosis and timely medical intervention.

The comparative F1-Score analysis presented in Figure 4 demonstrates that Transformer provides the best balance between precision and recall, achieving an F1-score of 98.11%. Since the F1-score simultaneously considers both false-positive and false-negative predictions, the observed improvement indicates that Transformer maintains consistent classification performance across different patient groups. Bi-LSTM and GRU also exhibit strong balanced performance, whereas ANN and DNN show comparatively lower values due to their limited capability to model complex nonlinear relationships within clinical datasets.

The ROC-AUC comparison illustrated in Figure 5 further validates the superiority of advanced deep learning models. Transformer achieved the highest ROC-AUC value (0.994), followed by Bi-LSTM (0.981) and GRU (0.978). A ROC-AUC value close to one indicates excellent discrimination between healthy and heart disease cases. These results demonstrate that attention-based learning enables Transformer to distinguish disease patterns more effectively than convolutional and recurrent architectures.

The Matthews Correlation Coefficient comparison presented in Figure 6 provides additional evidence regarding classification robustness. Transformer achieved the highest MCC value (0.967), followed by Bi-LSTM (0.931) and GRU (0.925). Since MCC simultaneously considers true positives, true negatives, false positives, and false negatives, it is considered one of the most reliable performance measures for binary medical classification problems. The higher MCC values confirm that advanced deep learning models generate balanced predictions even when class distributions vary.

The error analysis summarized in Table 2 further supports these observations. Transformer recorded the lowest Mean Absolute Error (MAE = 0.031), Root Mean Square Error (RMSE = 0.148), and Log Loss (0.096), indicating highly accurate probability estimation and reduced prediction uncertainty. CNN, LSTM, Bi-LSTM, and GRU also demonstrated substantially lower prediction errors than ANN and DNN. These findings suggest that advanced deep learning architectures effectively capture nonlinear clinical relationships while minimizing classification errors.

The cross-validation results shown in Table 3 demonstrate the stability and generalization capability of the evaluated models. Transformer achieved the highest average accuracy under both 5-fold (97.91%) and 10-fold (98.32%) cross-validation while also producing the lowest standard deviation (0.61). The low standard deviation indicates that Transformer provides highly consistent predictions across different validation folds, thereby reducing the likelihood of model overfitting. The remaining advanced models also exhibit strong generalization performance, whereas ANN shows comparatively higher variability.

Finally, the statistical comparison summarized in Table 4 reveals that Transformer improves prediction accuracy by 7.49% compared with ANN and outperforms every other evaluated architecture. These improvements demonstrate that self-attention mechanisms successfully capture long-range dependencies among clinical variables, allowing Transformer to learn richer feature representations than conventional feedforward, convolutional, or recurrent neural networks. Overall, the experimental results consistently demonstrate that advanced deep learning approaches significantly improve heart disease prediction accuracy, reliability, and robustness compared with traditional deep learning models.

7. Conclusion

This study presented a comprehensive comparative analysis of traditional and advanced deep learning approaches for heart disease prediction using clinical datasets. Seven deep learning architectures, namely ANN, DNN, CNN, LSTM, Bi-LSTM, GRU, and Transformer, were developed and evaluated under identical preprocessing and validation conditions. The experimental results demonstrated that advanced deep learning models consistently outperform traditional neural network architectures across all evaluation metrics. Among the evaluated methods, the Transformer model achieved the highest classification accuracy of **98.32%**, together with superior Precision, Recall, F1-Score, ROC-AUC, and Matthews Correlation Coefficient. The lower prediction errors and higher cross-validation stability further confirmed the robustness and generalization capability of the Transformer architecture. The findings indicate that attention-based deep learning models effectively learn complex nonlinear relationships among clinical variables, thereby enabling more accurate and reliable heart disease prediction. Consequently, the proposed comparative framework provides valuable decision support for clinicians by facilitating early diagnosis, risk assessment, and timely treatment planning, ultimately contributing to improved patient outcomes and intelligent healthcare systems.

8. Further Research

Although the proposed framework demonstrates excellent predictive performance, several opportunities remain for future investigation. Future studies may evaluate the proposed models using larger multi-center clinical datasets collected from diverse populations to further improve model generalization and reduce potential population bias. The integration of multimodal healthcare information, including electronic health records, laboratory reports, electrocardiogram signals, echocardiography images, wearable sensor data, and genomic information, could significantly enhance predictive accuracy and personalized risk assessment.

Future work may also investigate hybrid deep learning frameworks that combine Transformer architectures with optimization algorithms such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Differential Evolution (DE), Bayesian Optimization, or other metaheuristic techniques for automated hyperparameter tuning and feature optimization. Furthermore, incorporating Explainable Artificial Intelligence (XAI) methods such as SHAP, LIME, attention visualization, and counterfactual explanations would improve model interpretability and increase clinician trust during decision-making. The proposed framework may also be extended to real-time cloud-based or Internet of Medical Things (IoMT) healthcare systems for continuous remote patient monitoring, early disease screening, and intelligent clinical decision support. Such advancements would further strengthen the practical applicability of advanced deep learning in precision cardiology and next-generation intelligent healthcare environments.

Reference

1. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
2. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
3. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
4. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.
5. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
6. M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," IEEE Trans. Signal Process., vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
7. K. Cho et al., "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in Proc. EMNLP, Doha, Qatar, 2014, pp. 1724–1734.
8. A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.
9. D. Dua and C. Graff, "UCI Machine Learning Repository," School of Information and Computer Science, University of California, Irvine, CA, USA, 2019.
10. A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 9, no. 4, Art. no. e1312, 2019.
11. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," Journal of Machine Learning Research, vol. 3, pp. 1157–1182, Mar. 2003.
12. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2012.
13. J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," Journal of Machine Learning Research, vol. 13, pp. 281–305, Feb. 2012.
14. T. G. Dietterich, "Ensemble methods in machine learning," in Multiple Classifier Systems, Lecture Notes in Computer Science, vol. 1857. Berlin, Germany: Springer, 2000, pp. 1–15.
15. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.

16. S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
17. D. B. Rawat and C. Dobre, Smart Healthcare: Internet of Things, Cloud Computing, and Big Data Analytics. Cham, Switzerland: Springer, 2022.
18. T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, Jun. 2006.
19. A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.
20. E. J. Topol, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. New York, NY, USA: Basic Books, 2019.
21. P. Rajesh and M. Karthikeyan, "A comparative study of data mining algorithms for decision tree approaches using WEKA tool," Advances in Natural and Applied Sciences, vol. 11, no. 9, pp. 230–243, 2017.
22. P. Rajesh, M. Karthikeyan, B. Santhosh Kumar, and M. Y. Mohamed Parvees, "Comparative study of decision tree approaches in data mining using Chronic Disease Indicators (CDI) data," Journal of Computational and Theoretical Nanoscience, vol. 16, no. 4, pp. 1472–1477, Apr. 2019.
23. S. Ravishankar and P. Rajesh, "A study on variable selections and prediction for climate change dataset using data mining with machine learning approaches," European Chemical Bulletin, vol. 11, no. 12, pp. 1866–1877, 2022.
24. G. K. Arun and P. Rajesh, "A study on variable selections and prediction for diabetes pregnancy dataset using data mining with machine learning approaches," Journal of Advanced Zoology, vol. 44, no. 3, pp. 949–957, 2023.
25. P. Rajesh, M. Karthikeyan, and R. Arulpavai, "Data mining algorithm to predict the factors for agricultural development using stochastic model," International Journal of Recent Technology and Engineering, vol. 8, no. 3, pp. 2713–2717, Sep. 2019.
26. P. Rajesh and M. Karthikeyan, "Prediction of agriculture growth and level of concentration in paddy—A stochastic data mining approach," in Advances in Big Data and Cloud Computing: Proceedings of ICBDCC18. Singapore: Springer, 2018, pp. 127–139.
27. D. Dua and C. Graff, UCI Machine Learning Repository. Irvine, CA, USA: University of California, School of Information and Computer Science, 2019.
28. R. Detrano et al., "International application of a new probability algorithm for the diagnosis of coronary artery disease," The American Journal of Cardiology, vol. 64, no. 5, pp. 304–310, 1989.
29. H. A. Janosi, W. Steinbrunn, M. Pfisterer, and R. Detrano, Heart Disease Dataset, V.A. Medical Center, Long Beach and Cleveland Clinic Foundation, distributed through the UCI Machine Learning Repository.
30. M. Lichman, "UCI Machine Learning Repository," School of Information and Computer Science, University of California, Irvine, CA, USA, 2013.
31. J. H. Friedman, T. Hastie, and R. Tibshirani, The Elements of Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2009.