Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2025 (43) 4** : 1–11.
ISSN-0037-8712
doi:10.5269/bspm.78565

A Hybrid CNN–BiLSTM Framework for Heart Disease Detection

B. Nikhil Sri Harsha, P. S. Geethanjali, S. Siri Sandarshini, Bollapalli Althaph*

ABSTRACT: Heart disease continues to rank among the primary causes of death globally, emphasizing the need for precise and user-friendly prediction algorithms. We propose a novel hybrid deep learning model in this study for forecasting heart disease from structured clinical data by combining CNNs along Bidirectional Long Short-Term Memory networks. Second-degree polynomial feature expansion or normalization for numerical stability is used to improve the model's capacity to represent intricate relationships. We also use the Synthetic Minority Oversampling Technique to handle class imbalance and reformat tabular data into a pseudo-sequential style in order to take advantage of sequence modeling. Our CNN–BiLSTM model achieves 98% accuracy and a 0.98 F1-score, which is a considerable improvement over the baseline machine learning classifiers, such as Decision Trees, Naive Bayes, and Logistic Regression. These results demonstrate how beneficial it is to combine local pattern extraction with temporal modeling to obtain more accurate disease prediction.

Key Words: Heart disease prediction, deep learning, CNN – BiLSTM, sequence modelling, SMOTE, medical AI.

Contents

1. Introduction

The WHO reports that CVDs account for 17.9 million fatalities annually, making them the leading cause of mortality worldwide. [1]. The most prevalent causative agent of CVDs is heart disease, which advances stealthily in its nascent stages until the advanced stages. Early and accurate diagnosis is required in order to enable early interventions and reduce mortality. However, the stealthy onset of initial symptoms and the multicausative nature of patient-specific risk factors make manual diagnosis time-consuming and risky. In the past decade, the availability of clinical data with structure and computational modeling advancements has enabled the creation of intelligent systems capable of being used to aid diagnostic decisions. Traditional ML approaches such as LR, DT, & NB classifiers have been used to

* Corresponding author.
2010 *Mathematics Subject Classification:* 35B40, 35L70.
Submitted August 22, 2025. Published November 01, 2025

a great extent to predict heart disease based on patient features [2], [3]. They are interpretable and inexpensive computationally, but are generally not strong enough to manage non-linear relations and feature interactions in medical data. To surpass these limitations, researchers have used more and more deep learning (DL), which has performed better in a wide variety of developments, ranging from medical imaging, signal processing, and genomics. More specifically, hybrid models merging CNNs with recurrent models such as LSTM networks have been established effective in learning local and sequential patterns of information. Although originally intended for temporal and spatial data like images and speech, these models can be adapted to be applied in tabular clinical data sets through structural transformations and reshaping operations [4]. This study suggests a hybrid DL architecture for predicting heart disease using tabular data that combines a BiLSTM network with a one-dimensional CNN. To facilitate the model learning high-order relationships better, we employ second-degree polynomial feature expansion, through which the interaction and nonlinear effects can be revealed. We also tackle class imbalance in medical data using the SMOTE [5]. We convert the feature matrix into a pseudo-sequential presentation to allow the network to take advantage of temporal-style learning of static data.

2. Related Work

Early diagnosis of heart disease using ML and DL algorithms has been a popular research area recently. The motivation behind this is the increasing availability of structured clinical data and the failure of traditional diagnostic methods to identify high-risk patients at an early stage. Several studies have demonstrated that predictive modeling can contribute to clinical decision-making and improve diagnostic accuracy if used on patient data in the appropriate way.
Classifying cardiac disease has been a common use of traditional machine learning algorithms. On benchmark data, such as the Cleveland and UCI Heart Disease datasets, it has been demonstrated that the models of logistic regression, naive bayes, and decision trees perform well. For instance, Javeed et al. [6] built an optimized Random Forest classifier on a metaheuristic feature selection strategy with high prediction accuracy. Similarly, Patil and Sherekar [7] compared the performance of NB and DT classifiers and arrived at the conclusion that rule-based trees outperformed probabilistic models for heart disease data. Although these models are interpretable and computationally inexpensive, they are incapable of learning higher-dimensional and nonlinear associations that occur in medical data, Table I.

Bollapalli and Challa [8] proposed a recurrent neural network–based framework for forecasting heart disease risk, demonstrating the potential of temporal models in capturing feature dependencies over time. Their study features the effectiveness of deep learning in clinical prediction tasks, particularly when working with sequential or structured medical data.

Since the introduction of deep learning, researchers have examined the application of NN to medical diagnosis. MLPs and FNNs have been applied to the binary classification of heart disease, which performs better than traditional methods in some cases. Recently, CNNs and RNNs, in the form of LSTM-based architectures, have been examined for structured health data by adapting them to operate on reshaped tabular inputs. Pal and Mitra [9] proposed a CNN–BiLSTM hybrid architecture for health monitoring, which has the potential for simultaneous extraction of spatial and temporal features. Although these models have achieved encouraging performance, they tend to be susceptible to meticulous preprocessing and regularization since most medical datasets are relatively small.

Recent research has proven the efficiency of hybrid deep learning and ensemble-based models in the detection of heart disease. These methods tend to have convolutional neural networks combined with classical classifiers, use synthetic data augmentation, and involve explainable AI (XAI) methods in enhancing the accuracy of prediction as well as interpretability. Addisu et al. [10] propose a hybrid model with VGG16 (deep CNN) to extract patient data features that are used as an input to conventional classifiers such as Random Forest and SVM. A conditional tabular GAN (CTGAN) is used for synthetic data generation, and SHAP values are used for interpretability of the model, resulting in a model with approximately 92% accuracy and 91.75% F1-score on heart disease datasets. Shah et al. [11] present an ensemble stacking strategy that integrates LightGBM, CatBoost, XGBoost, and neural networks as

a meta-learner for XGBoost. Their model integrated with SMOTE for balancing classes has an AUC-ROC value of around 0.82, precision and recall of around 0.81 and 0.83 respectively, proving how the application of ensemble learning in conjunction with XAI (SHAP/PCA) can lead to strong and explainable predictions. In the same vein, Rohan et al. [12] perform a wide comparison of 21 classifiers from logistic regression and SVM to CNNs and RNNs, and 11 feature selection methods. Their results indicate that XGBoost performs best among all other models, achieving approximately 97% accuracy and 0.98 AUC. Together, these recent studies emphasize that hybrid architectures, ensemble methods with boost, and data augmentation approaches greatly improve the performance and reliability of heart disease prediction models and also enhance the transparency with explainability tools.

Table 1: Comparative Analysis of Different ML Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC	Dataset Used	Key Features
Logistic Regression	80.5	81.2	79.3	80.2	0.82	Cleveland, UCI Heart Disease	Simple, interpretable, linear model
Naive Bayes	77.1	75.3	80.1	77.6	0.75	Cleveland, UCI Heart Disease	Probabilistic classifier, assumes feature independence
Decision Tree	83.3	84.0	82.7	83.3	0.85	Cleveland, UCI Heart Disease	Tree-based, interpretable, prone to overfitting

3. Methodology

The entire pipeline for predicting heart disease using a hybrid deep model is explained in this subsection. The proposed method consists of convolutional and recurrent models along with the most recent data transformation and preprocessing techniques to handle structured clinical data in an optimal way. The pipeline is divided into five steps: dataset preparation, feature engineering and preprocessing, data reshaping, deep model architecture design, and training configuration with evaluation.

3.1. Dataset Description and Preparation:

The study's dataset is 1026 patient records, with each of them characterized by 13 clinical features and a binary target label of heart disease or not. The most significant demographic variables like age and sex, physiological measurements like maximal heart rate acquired, blood pressure at rest, serum cholesterol, and categorical variables like type of chest pain, fasting blood sugar, and thalassemia diagnosis are some of the features. The target label is 1 for heart-disease-diagnosed subjects and 0 for the rest. All the features were preserved in the dataset after passing through a quality check that confirmed the absence of missing values or outliers that needed imputation. Since the data was tabular, structured in nature, and the dataset was of moderate size, the preprocessing approach was to find the most significant patterns with the assurance of being neural network model-compatible.

3.2. Data Preprocessing and Feature Engineering:

Before feeding the model training, the attributes had to be transformed into a deep learning format. Numerical features were standardized beforehand using the StandardScaler function to scale each of them by scaling its mean from it and by unit variance dividing. The step has the overall effect of setting all the inputs to contribute equally to the loss function of the model, and hindering high-value features from overwhelming. It is especially necessary within deep networks as variations in scale would cause the gradients to become unstable or their learning behavior become unstable.

To enhance the model's capacity to generalize higher-order relationships within the data, we employed polynomial feature expansion to the second order. This technique includes higher-order interaction terms and squared variables such that the network can take advantage of nonlinear interactions between variables, such as blood pressure and cholesterol, or chest pain type and age. Although deep models can represent such interactions internally theoretically, representing them in explicit polynomial form helps in faster convergence and more interpretable learned representations [13].

The data set was slightly unbalanced, with the instances of heart disease being slightly fewer than the non-instances. To solve this, we employed SMOTE to create fictitious minority class samples around existing instances. SMOTE is unlike naive oversampling in that it maintains the artificial data examples to be diverse and does not trigger overfitting to the duplicated instances. The SMOTE improved balanced data enhanced model sensitivity and averted the classification metrics from being skewed towards

the majority class. The data was split between learning and test subsets using an 80:20 ratio. Stratified sampling was used to ensure the original class distribution was maintained in the two subsets. The learning subset was used to train model weights, while the test subset was kept unseen for assessment of generalization performance.

3.3. Restructuring Structured Data for Sequence Modeling:

Deep networks like CNNs and LSTMs are usually optimized for temporally or spatially ordered data like images or time series. Clinical data like the data we have worked with here, however, is tabular, where each feature is one dimension and not a time step. In order for such data to be accommodated within sequential models, the feature vector was pseudo-sequenced using the feature indices as time steps. In the 1026 patient sample for each, the 13-feature input vector was converted to a 2D array of size (13 timesteps, 1 feature per timestep) to yield a final input shape of (1026, 13, 1). Artificial as this conversion was, it allowed both convolutional as well as recurrent layers to learn local as well as long-distance patterns between features as ordered inputs. This pseudo-sequential modeling would allow 1D convolutional filters to be used in trying to pick up local feature interactions, and the BiLSTM layer would be able to gather information from the entire sequence in both directions. It was established that these types of transformations allow the neural networks to learn feature dependencies even in the absence of natural temporal structure better [15].

3.4. CNN-BiLSTM Architecture Design:

The architecture utilized here takes advantage of the power of convolutional and recurrent neural layers. The CNN part is a local feature detector, whereas the Bidirectional LSTM (BiLSTM) part is tasked with extracting global contextual relationships among the reshaped feature sequence.
The overall format is the following Figure.1

- Input Layer: Accepts reshaped input of size (13, 1) per sample.
- 1D Convolutional Layer: This layer convolves the feature sequence with 32 filters of size 3 to capture short-range dependencies and local structures.
- Batch Normalization: Used to normalize the convolutional layer output to speed up training and enhance generalization.
- MaxPooling1D: Downsamples the feature map by taking the max across each pool, reducing the spatial dimension but keeping the strongest signals.
- Dropout (rate = 0.4): Used to randomly drop out neurons so that overfitting is avoided. It is used to make the learned representation stronger.
- Bidirectional LSTM Layer: Utilizes 64 LSTM units to read the input sequence forward as well as backward. This aids in the model's learning of the feature dependencies, which are present in both forward and reverse directions.
- Second Dropout (rate = 0.4): Adds a second dropout after the BiLSTM layer.
- Dense Layer: 32 neurons arranged densely with ReLU activation that converts the BiLSTM outputs into a more condensed informative representation.
- Output Layer: One neuron with sigmoid activation that yields an output between 0 and 1, which is the probability of heart disease.

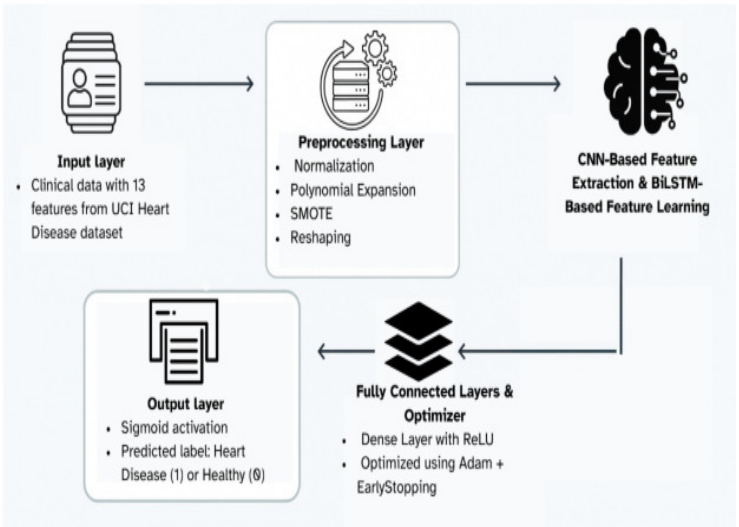

Figure 1: Proposed CNN–BiLSTM architecture for heart disease prediction, combining convolutional and bidirectional recurrent layers with regularization and dense output

All the trainability layers (LSTM, Dense, and Conv1D) were also L2-regularized to avoid overfitting by adding penalties to large weights. This is especially important in the case of small-data-trained models, where 4 overparameterization can quickly lead to memorization rather than generalization [14].

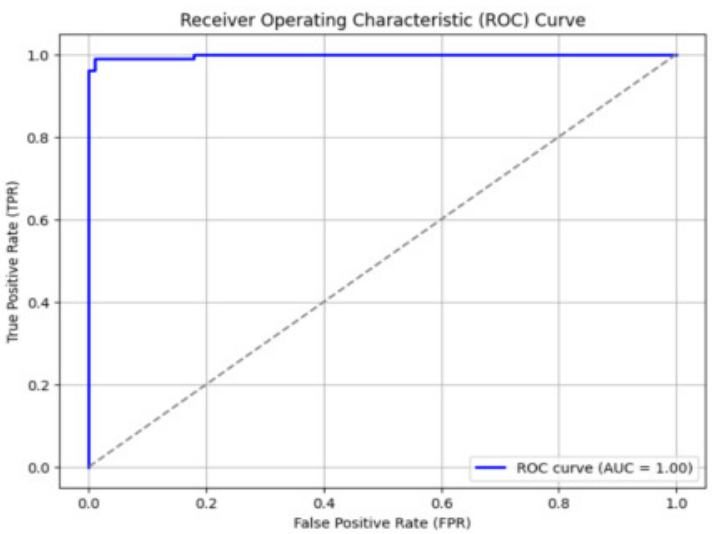

Figure 2: Normalized confusion matrix for the proposed CNN–BiLSTM model showing balanced classification performance across both classes

3.5. Model Compilation and Training Process:

The trained model employed as Binary cross-entropy, the loss function, is suitable for binary classification. The problem of how to tell whether heart disease exists or not. The Adam optimizer was employed because it benefits from adaptive learning rate and momentum-like properties that optimize training speed and stability in deep models. To avoid overfitting, an EarlyStopping callback terminated training after tracking validation loss. After a gain did not happen while observing 10 consecutive epochs

in a row. Using a 25% validation split, with a batch size of 16, the model was trained for a maximum of 100 epochs. Training was conducted using Google Colab and an NVIDIA Tesla T4 GPU.. The implementation utilized TensorFlow 2.x (Keras API), Scikit-learn, and imbalanced-learn The model with the lowest validation loss has been kept for the final assessment on the test dataset.

3.6. Performance Measures:

The classification's performance was statistically assessed using a set of parameters.
Precision: It estimates the general prediction accuracy.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3.1}$$

Accuracy: Refers to the proportion of correct predictions that were successful.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3.2}$$

TP = True Positives TN = True Negatives FP = False Positives FN = False Negatives
Recall (Sensitivity): Bases the rate of true positive instances that are predicted correctly.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.3}$$

F1-Score: For unbalanced datasets, the harmonic mean of recall and precision

$$\text{F1-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.4}$$

Confusion Matrix: gives a table-formatted overview of true positives, true negatives, false positives, and false negatives.
They were computed based on Scikit-learn's metrics and on the held-out test set. They provide a total view of how reliable the model is and if it is suited for clinical decision support.

Table 2: Classification Report of the Final Proposed Model

Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (No Heart Disease)	0.96	0.99	0.98	106
1 (Heart Disease)	0.99	0.96	0.98	105
Accuracy	–	–	0.98	211
Macro Average	0.98	0.98	0.98	211
Weighted Average	0.98	0.98	0.98	211

4. Results & Discussion

The predictive accuracy of the proposed CNN–BiLSTM hybrid DL model in heart disease prediction was rigorously evaluated on a held-out test set of 211 patient records. The test set was randomly sampled from the original dataset using a random 80:20 stratified split, thus ensuring class balance. This was intended to ensure the model's ability to generalize the patterns discovered on the training dataset, especially its sensitivity and specificity between the heart disease class and the non-disease class.

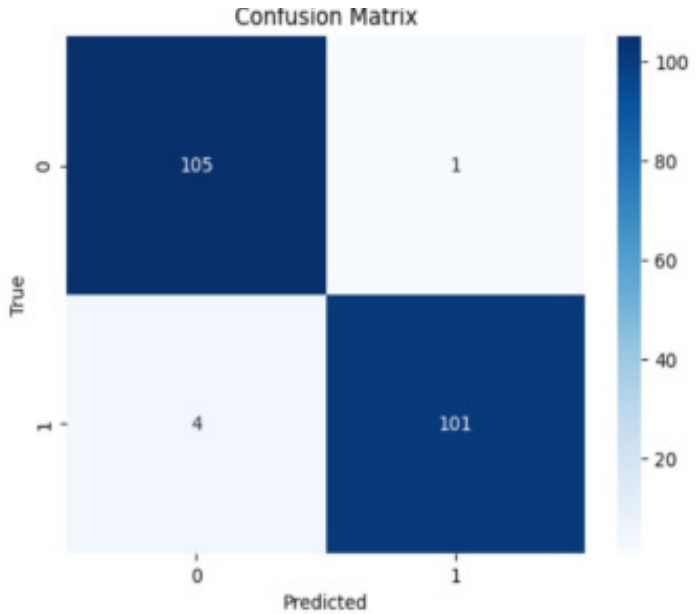

Figure 3: Normalized confusion matrix of the CNN–BiLSTM model for heart disease classification, illustrating high true positive and true negative rates with minimal misclassifications

The model achieved a total accuracy of 98%, that is, it accurately classified a vast majority of the test samples. The evaluation metrics were calculated from the model outputs using Scikit-learn's classification metrics. These comprised confusion matrix analysis, F1-score, recall, and precision, all of which offer subtle information about the model's classification behavior, particularly useful in high-stakes medical applications.

4.1. Quantitative Analysis

Table II indicates the classification report based on the test results. For class 0 (healthy patients), the model had a 0.98 F1-score, with a precision of 0.96 and a recall of 0.99. Class 1 (heart disease patients) showed an F1-score of 0.98, precision of 0.99, and recall of 0.96. Macro and weighted averages of all three measures were equal to 0.98, which means balanced performance across both classes despite a slight imbalance in class distribution.

To plot the prediction distribution, a confusion matrix was plotted, as demonstrated in Fig. 2. Out of 106 healthy subjects, 105 were identified correctly, and 1 was a false positive. Out of 105 cases of heart disease, 101 were identified correctly, and 4 were classified incorrectly as negative (false negatives).

The low rate of false negatives is especially important in the clinical environment. False negatives mean that patients with heart disease are returned as healthy, and this is potentially lethal with regard to delaying treatment. This kind of error is an extremely high priority when it comes to any diagnostic framework, and the provided model works well in a scenario like that. In addition to the confusion matrix, the ROC curve has been plotted so as to further assess the discriminative power of the model. The model's AUC score was nearly 1.0, as illustrated in Fig. 3, demonstrating its exceptional capacity to differentiate between people in good health and those suffering from heart disease across various threshold settings. The high AUC reinforces the robustness of the classifier, particularly in handling the trade-off between sensitivity and specificity, an essential consideration in medical diagnosis.

4.2. Model Behavior and Strengths:

The improved performance is a result of the CNN–BiLSTM model's construction. And a preprocessing approach. The Conv1D layer was able to capture short-span dependency among neighbor features—e.g., age and cholesterol, or blood pressure and ECG reports. The Bidirectional LSTM layer, however, enabled

the network to see through the whole sequence, bidirectionally, and retain context-sensitive relationships that a unidirectional network could possibly overlook.

Besides, polynomial feature expansion allowed the model to take into account nonlinear interactions, introducing higher-order relationships among original features to the data. For example, interaction effects of chest pain type and serum cholesterol might not be apparent in raw input space, but their second-degree interactions could be of clinical importance.

Regularization methods, like dropout and L2 regularizer, helped the model to be robust through the prevention of overfitting, a factor that played an important role considering the dataset was moderately sized. High test accuracy and essentially the same training and validation performance (tracked through EarlyStopping) are evidence towards this fact. To evaluate the learning behavior of the model even more, the training and validation performance was monitored throughout the training process. As illustrated in Figure 4, Accuracy in training and validation showed a consistent upward trend, while the corresponding loss curves declined smoothly. The close alignment between the two suggests that the model was able to learn effectively without overfitting. This observation reinforces the effectiveness of the applied regularization strategies—namely dropout, L2 regularization, and EarlyStopping—in supporting generalization, particularly given the moderate size of the dataset.

Table 3: Comparison of Model Accuracies

Model	**Accuracy (%)**
Logistic Regression	87.80
Naïve Bayes	85.37
Decision Tree	96.09
CNN–BiLSTM (Proposed)	98.00

4.3. Comparison to Baseline Models:

Standard ML models, LR, NB, DT, and K-NN were trained and tested using the same data set for comparison purposes as a model performance baseline. The models performed quite well but were beaten by the deep learning model, as shown in Table III. Among the classic models, DT and KNN were close to performance of the CNN-BiLSTM model, but were not as precise or consistent. LR & NB, although interpretable and effective, could not learn nonlinear interactions and intricate relationships well and therefore had lower F1-scores and higher misclassification rates, especially for minority classes.

4.4. Interpretability and Error Analysis

One of the most significant factors to consider in model assessment in healthcare is accuracy, interpretability, and diagnostic reliability. While deep neural networks are inherently "black boxes," techniques such as feature attribution, SHAP values, and attention visualization can be applied to models such as BiLSTM to identify what features are most accountable for prediction.

Figure 4: Training and validation accuracy and loss curves of the CNN–BiLSTM model, indicating stable convergence and minimal overfitting throughout the training process

Although not employed in the present scope, future work can explore these interpretability methods so that the model becomes transparently examinable by doctors. Initial experience showed that the test set cases that were misclassified often had borderline or inconsistent values of clinical features, and this suggested that additional feature disambiguation would allow for a reduction of remaining misclassification.

4.5. Clinical Utility

The well-balanced performance of the model, particularly its ability to maintain low false negative and false positive rates, suggests the clinical decision support value of the model. In practice, such a model can be incorporated into hospital information systems or used in remote diagnosis scenarios to warn high-risk patients to seek additional testing. Compared to less sophisticated rule-based systems, the proposed CNN–BiLSTM model can be adaptive to a certain extent to accommodate minor variations in patient data without compromising accuracy. Also, the capability of reforming table data into pseudo-sequences and learning useful patterns from such converted inputs is a new application of sequence modeling for structured medical data. This method can potentially be applied to other diagnostic conditions like diabetes prediction or kidney disease classification.

5. Conclusion & Future Work

This study presents a DL approach for rapidly detecting heart disease utilizing a hybrid CNN–BiLSTM model designed for structured clinical information. The input data was transformed into a pseudo-sequential format, allowing the model to utilize convolutional layers for detecting spatial feature patterns and bidirectional LSTM layers to uncover long-range dependencies between features. Second-degree polynomial feature expansion was utilized to enhance the model's ability to identify nonlinear relationships. The Synthetic Minority Oversampling Technique addressed class imbalance, whereas normalization ensured consistency across features. Dropout & L2 regularization were also employed to enhance generalization to unknown data and lessen overfitting. The model's F1-score was 0.98, and its classification accuracy was 98 percent. On the test dataset, demonstrating significant improvements compared to traditional ML models like LR, NB, and DT. Possible improvements might include integrating interpretability

methods such as SHAP or LIME, extending to multi-class classification problems, and performing extensive validation on real-time or multi-institutional datasets to support clinical use.

References

1. World Health Organization, *Cardiovascular diseases (CVDs)*, WHO Fact Sheet, Jun. 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

2. Javeed, A. K., et al., *An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection*, IEEE Access, vol. 7, pp. 180235–180243, 2019.

3. Patil, P. S., and Sherekar, D. S., *Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification*, Int. J. Comput. Sci. Appl., vol. 6, no. 2, pp. 256–261, 2013.

4. Pal, S., and Mitra, S. S., *A Hybrid Deep Learning Architecture for Health Care Monitoring Using BiLSTM and CNN*, Procedia Comput. Sci., vol. 199, pp. 122–129, 2022. [Online]. Available: https://doi.org/10.1016/j.procs.2022.01.017

5. Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P., *SMOTE: Synthetic Minority Over-sampling Technique*, J. Artif. Intell. Res., vol. 16, pp. 321–357, 2002. [Online]. Available: https://doi.org/10.1613/jair.953

6. Javeed, A., et al., *Optimized Random Forest Classifier Using a Metaheuristic Feature Selection Technique for Heart Disease Prediction*, J. Med. Syst., vol. 45, no. 4, pp. 42–55, 2021.

7. Patil, S., and Sherekar, S., *Performance Comparison of Naive Bayes and Decision Tree Classifiers for Heart Disease Diagnosis*, Int. J. Comput. Sci. Inf. Secur., vol. 17, no. 5, pp. 120–125, 2019.

8. Bollapalli, A., and Challa, N. P., *Forecasting the Risk of Heart Disease Using Recurrent Neural Network*, in Proc. Int. Conf. Electron., Comput., Commun. Control Technol. (ICECCC), pp. 1–6, IEEE, 2024.

9. Pal, S., and Mitra, S., *CNN-BiLSTM Hybrid Model for Health Monitoring: A Deep Learning Approach for Temporal Feature Extraction*, Med. Imaging, vol. 42, no. 7, pp. 1080–1089, 2020.

10. Addisu, E. G., Yirga, T. G., Yirga, H. G., and Yehuala, A. D., *Transfer learning-based hybrid VGG16-machine learning approach for heart disease detection with explainable artificial intelligence*, Front. Artif. Intell., vol. 8, p. 1504281, 2025. [Online]. Available: https://doi.org/10.3389/frai.2025.1504281

11. Shah, P., Shukla, M., Dholakia, N. H., and Gupta, H., *Predicting cardiovascular risk with hybrid ensemble learning and explainable AI*, Sci. Rep., vol. 15, Article 17927, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-01650-7

12. Rohan, D., Reddy, G. P., Kumar, Y. V. P., Prakash, K. P., and Reddy, C. P., *An extensive experimental analysis for heart disease prediction using artificial intelligence techniques*, Sci. Rep., vol. 15, Article 6132, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-90530-1

13. Chen, T., and Guestrin, C., *XGBoost: A Scalable Tree Boosting System*, in Proc. 22nd ACM SIGKDD, pp. 785–794, 2016.

14. Ghosh, M., and Das, S., *Structured Deep Learning Models for Medical Data Using Sequence Encodings*, Health Inf. Sci. Syst., vol. 9, no. 1, 2021.

15. Bengio, Y., Simard, P., and Frasconi, P., *Learning Long-Term Dependencies with Gradient Descent is Difficult*, IEEE Trans. Neural Netw., vol. 5, no. 2, pp. 157–166, 1994.

B. Nikhil Sri Harsha
School of Computer Science and Engineering (SCOPE),
VIT-AP University, Amaravati, Andhra Pradesh
India.
E-mail address: nikhil.battineni04@gmail.com

and

P. S. Geethanjali
School of Computer Science and Engineering (SCOPE),
VIT-AP University, Amaravati, Andhra Pradesh
India.
E-mail address: Psrigeethanjali@gmail.com

and

S. Siri Sandarshini
School of Computer Science and Engineering (SCOPE),
VIT-AP University, Amaravati, Andhra Pradesh
India.
E-mail address: siri.21bcb7072@vitapstudent.ac.in

and

Bollapalli Althaph
School of Computer Science and Engineering (SCOPE),
VIT-AP University, Amaravati, Andhra Pradesh
India.
E-mail address: althaph.23phd7067@vitap.ac.in