Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 6s (2025)
https://www.jneonatalsurg.com

IoT-Enabled Smart Healthcare System for Heart Disease Prediction Using Deep Learning and Dimensionality Reduction

Preetha P*[1], Dr. A. Packialatha[2]

*[1]Research Scholar, Computer Science & Engineering, Vel's Institute of Science Technology and Advanced Studies, Tamilnadu, India.

[2]Associate Professor, Computer Science & Engineering, Vel's Institute of Science Technology and Advanced Studies, Tamilnadu, India.

Email ID: packialatha.se@velsuniv.ac.in

***Corresponding Author:**

Email ID: preetha.pandu@gmail.com

Cite this paper as: Preetha P, Dr. A. Packialatha, (2025) IoT-Enabled Smart Healthcare System for Heart Disease Prediction Using Deep Learning and Dimensionality Reduction. *Journal of Neonatal Surgery*, 14 (6s), 743-756.

ABSTRACT

Accurate predictive diagnostic systems have become essential because heart disease conditions are rising in frequency. Healthcare advancements led to broader patient data inclusion in medical databases that enhances heart disease diagnosis effectiveness. The current databases for ischemic heart disease struggle with four major problems involving both feature selection difficulties and insufficient sample sizes along with distribution imbalances and missing data points. The authors present a smart healthcare system powered by IoT with deep learning and dimensionality reduction methods to achieve accurate heart disease predictions. Through wearable IoT devices patients can give health data about their blood pressure and heart rate and oxygen saturation level and these readings are processed using deep learning algorithms based on Convolutional Neural Networks (CNNs). The predictive analysis utilizes principal component analysis alongside long short-term memory networks to accomplish efficient data simplification and performance improvement. The detection capabilities of the Attention Bidirectional Gated Recurrent Unit model (ABiGRU) improve with the help of a Parameter Optimization Approach (POA) to select its hyper parameters. The Kaggle dataset simulation revealed the system produced results of 96.80% accuracy on Dataset-I while attaining 94.80% accuracy on Dataset-II in addition to obtaining high precision values and recall scores and F1-scores. These predictive modeling outcomes prove that the system works effectively thus establishing itself as an accurate diagnostic tool for early heart disease prediction and improved patient healthcare treatment and demonstrating promising incorporation possibilities in IoT and deep learning based customized healthcare.

Keywords: *Internet of Things, Dimensionality Reduction, Heart Disease, Deep Learning, Long Short-Term Memory, Attention Bidirectional Gated Recurrent Unit, Principal Component Analysis*

1. INTRODUCTION

Health organizations worldwide must focus on building precise diagnostic systems because heart disease continues to rise across the globe as a vital public health issue. Healing mortality rates from heart diseases becomes possible through prompt detection accompanied by appropriate interventions because ischemic heart disease (IHD) stays among the principal causes of worldwide fatalities. Standard diagnostic procedures use human operators along with expert judgment but this leads to delayed results that exhibit human error. Modern technology advancements enable the creation of IoT-powered smart healthcare systems for providing immediate health monitoring services and detecting diseases early [1].

Modern people lead busy lives during their daily activities that result in increased stress levels and nervousness. Heart rate and blood pressure differ in every person with pulse rates between 60 to 100 beats per minute and blood pressure between 120/80 to 140/90. Heart Disease (HD) stands as the leading death-causing factor across the entire world. ‘‘Cardio’’ means ‘‘heart.’’ Heart disease exists as one of several heart diseases known as cardiologist disease. Heart disease remains the primary killer worldwide among all other diseases today. The cardiac conditions consist of Coronary artery disease, Myocardial infarction, Hypertrophic cardiomyopathy, Arrhythmia, Dilated cardiomyopathy, Heart failure, Mitral regurgitation, and congenital heart dis- ease. Based on advanced knowledge coupled with experience a prediction of heart disease stands as a complex medical challenge [2] [3]. The HD severity classification employs four different methods which include KNN, Naive Bayes, Decision Trees along with Genetic Algorithm. Treatment for the HD requires special caution due to its intricate nature. Failure to handle the situation with care could cause injury to the heart leading to possible death. Various metabolic disorders function as subjects for medical research while machine learning (ML) develops methods to analyze them. Classification of HD along with data analysis functions largely because of ML categorization methods [4].

Data mining functions as a method referred to as knowledge discovery from data (KDD) for computing the patterns enabled through knowledge found in extensive databases and information libraries and web platforms [5]. Evidence from the world health organization (WHO) report [6] indicates that 31% of global deaths occurred due to cardio- vascular disease (CVD) deaths reaching 17.9 million people in 2016. A combination of heart attack and stroke causes 85% of deaths in this population. The majority of deaths caused by CVD occur in developing nations along with countries from both low- and middle-income brackets. The 17 million precocious non-communicable disease fatalities from 2015 primarily took place in low and moderate-income nations where CVDs comprised 37% of cases. Detection of CVDs requires data mining algorithms according to research found in [7] and [8] while missing data represents a main obstacle to productivity in data analytics. When imputation of data is not correct it results in errors when predicting missing values [9].

2. RELATED WORKS

The research by M. A. Khan et al. [10] developed a supervised learning decision support system that analyzes wireless body sensor network healthcare data. The SSA system serves to select important features. The proposed M-DBN (modified deep belief network) technique gets evaluated against different approaches which include Artificial Neural Network, Conventional Neural network, and Deep Belief Network. The proposed method outperforms all other techniques currently in use. A better risk prediction system emerged from data mining techniques which T. Zheng et al. designed in their study [11]. The combination of CFS Subset Evaluation through the Best-First-Search method served to minimize data dimensions. Data originates from Cleveland Heart Disease Database together with PKU People’s Hospital’s Cardiology inpatient database. The SVM method requires two times longer execution than a random forest classifier implementation.

The paper presents an imperialist competitive algorithm based meta-heuristic solution for selecting significant heart disease features [12]. The methodology stands out from competing optimization methods since it provides FS genetics with another optimal outcome. The K-nearest neighbor method functions as the approach for conducting the categorization process. Application of the proposed algorithm enhanced the FS method efficiency according to evaluation outcome measures. This research obtained data from two sources: the UCI ML repository’s HD data and the Tehran Shahid Rajaei hospital. C. A. Subasini et al. [13] suggested a feature optimization approach to find optimal variable size n-gram characteristics that serve supervised ML operations termed ‘‘discrete weights based n-gram feature selection.’’ Predicting disease severity through medical records requires consideration of pre-processing methods in addition to selecting proper attributes followed by evaluating attribute FS and deciding on the classification approach.

A framework using Internet of Medical Things technology analyzes HD data through MSSO and ANFIS systems to achieve higher prediction accuracy. Through the Levy flight method the proposed MSSO-ANFIS brings an enhancement in search performance capabilities. Maximum fitness values were attained by the Levy-based crow search algorithm for FS throughout all its iterations. The proposed MSSO-ANFIS strategy achieves higher F1-score and precision rates together with accuracy and recall while maintaining the lowest classification error when compared to present methodologies [14].

The SSA represents a revolutionary optimization model introduced by S. Joteppa et al. [15]. Carnivorous mammals use gliding behavior as an effective method to cover extensive ranges. The research demonstrates thorough evidence that SSA achieves better performance in terms of convergence rate and optimization accuracy when compared against known optimizers. The proposed technique gets verified through its implementation in real-time HD research as a test of its usable features. The last outcome confirms SSA provides highly accurate solutions alongside a fast convergence rate when compared to alternative existing optimizers.

3. MATERIALS AND METHODS

3.1 Dataset

The data collection consists of a network of objects where each instance has its own report that shows its represented data and attributes represent individual dataset attributes. The analysis draws its information from 'Cleveland, Switzerland, Long Beach, and Hungary datasets alongside data obtained from UCI repository [16] and Kaggle [17]. The data collection consists of 76 features but 14 features prove particularly useful in diagnosing heart disease. The predictive class attribute functions as the final attribute in most examination cases. A table has been provided to show the specifications of the data set attributes Table 1.

Random Forest Algorithm

Random Forest is an ensemble learning algorithm that creates multiple decision trees during training and outputs the mode (for classification) or average (for regression) of the predictions from individual trees. It shows in the Figure 1.

Collect a dataset that includes various features related to heart disease. For example, the **Cleveland Heart Disease Dataset** from the UCI repository includes features like:

- Age
- Sex
- Chest pain type
- Blood pressure
- Cholesterol levels
- ECG results
- Maximum heart rate
- Exercise-induced angina
- ST depression induced by exercise
- Slope of peak exercise ST segment
- Number of major vessels colored by fluoroscopy
- Thalassemia (Thal) status (fixed vs. reversible)

Split the dataset into **training** and **testing** sets. A common practice is using an **80-20 split**, where 80% of the data is used for training the model, and 20% is reserved for testing.

- ➢ n_estimators: The number of decision trees in the forest.
- ➢ max_depth: The maximum depth of each tree.
- ➢ min_samples_split: The minimum number of samples required to split an internal node.
- ➢ max_features: The number of features to consider when looking for the best split.

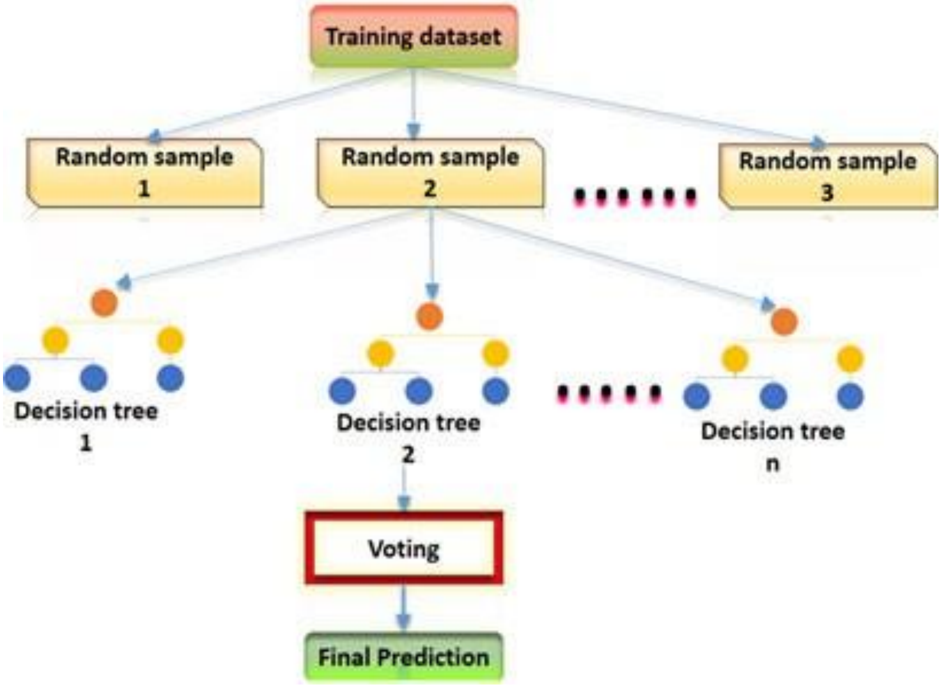

Figure 1: Random Forest Algorithm

Once the model is trained and evaluated, it can be used to predict heart disease status (e.g., **0 = No Heart Disease, 1 = Heart Disease**) for new, unseen data.

Table 1: Heart Disease Attributes

Attribute	**Description**	**Domain of Value**
Age	Age of the patient in years	Integer (0-120)
Sex	Gender of the patient	Male, Female
Chest Pain Type	Type of chest pain experienced by the patient	1: Typical Angina, 2: Atypical Angina, 3: Non-anginal pain, 4: Asymptomatic
Resting Blood Pressure	Resting blood pressure (in mmHg)	Integer (0-200)
Serum Cholesterol	Serum cholesterol level (in mg/dL)	Integer (0-600)
Fasting Blood Sugar	Blood sugar level after fasting (greater than 120 mg/dL indicates high blood sugar)	0: False, 1: True
Resting Electrocardiographic Results	Resting electrocardiogram results (measuring heart electrical activity)	0: Normal, 1: ST-T wave abnormality, 2: Left ventricular hypertrophy
Max Heart Rate	Maximum heart rate achieved during exercise	Integer (60-220)
Exercise Induced Angina	Indicates whether the patient experienced angina during exercise	0: No, 1: Yes
Oldpeak	ST depression induced by exercise relative to rest (indicates severity of heart disease)	Float (-2.6 to 6.2)
Slope of the Peak Exercise ST Segment	Slope of the ST segment during exercise (used to determine the severity)	1: Upsloping, 2: Flat, 3: Downsloping
Number of Major Vessels	Number of major vessels colored by fluoroscopy (related to heart blockage)	Integer (0-3)
Thalassemia	Thalassemia, a blood disorder related to hemoglobin	3: Normal, 6: Fixed defect, 7: Reversible defect
Target	Whether the patient has heart disease or not (the class attribute)	0: No, 1: Yes

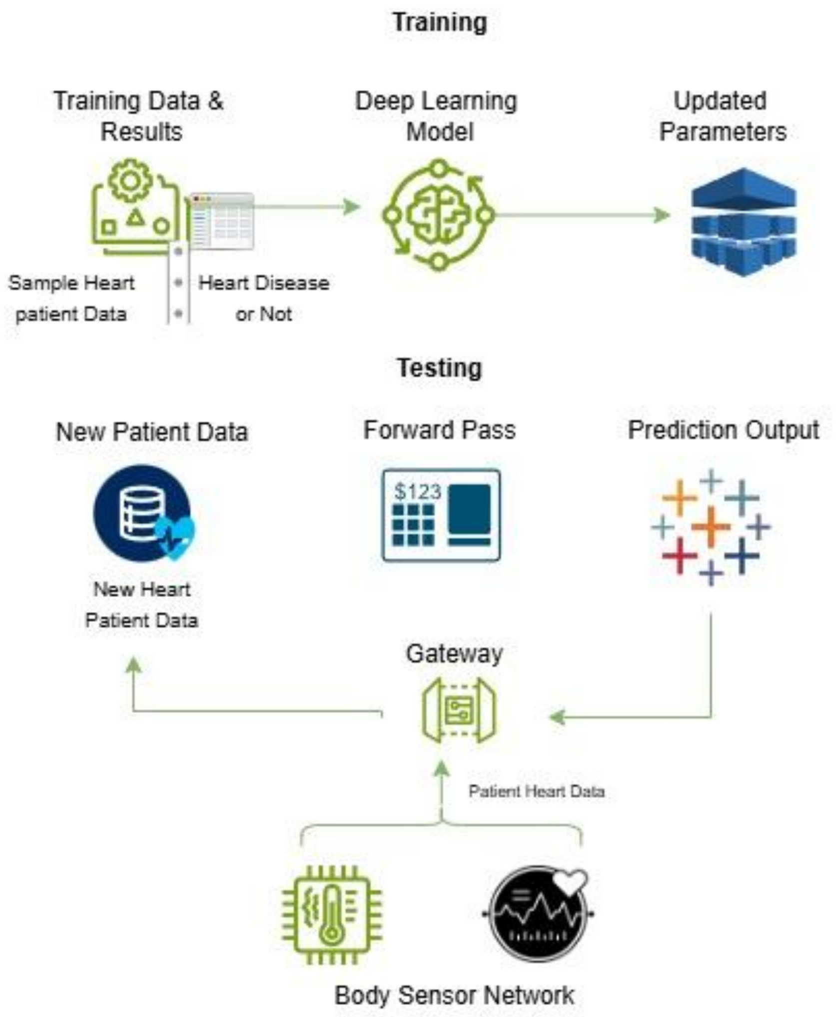

Figure 2: Training and Testing Datasets

3.1 Significance of the Attribute

The UCI HD dataset contains all properties which might affect HD prediction while the importance ratings act as hyper parameters [18]. Table 2 shows the importance of the property, description, domain value, and hyper parameter. Figure 2 shows the Training and Testing Datasets.

Table 2: Significance of the Attribute

Attribute	**Description**	**Domain of Value**	**Significance of the Attribute (Hyper parameter)**
Age	Age of the patient in years	Integer (0-120)	Significant for risk stratification; higher age is often correlated with greater risk.
Sex	Gender of the patient	Male, Female	Significant for gender-specific risks; gender can impact heart disease prevalence.
Chest Pain Type	Type of chest pain experienced by the patient	1: Typical Angina, 2: Atypical Angina, 3: Non-anginal pain, 4: Asymptomatic	High impact in determining the likelihood of heart disease, especially typical and atypical pain.

Attribute	**Description**	**Domain of Value**	**Significance of the Attribute (Hyper parameter)**
Resting Blood Pressure	Resting blood pressure (in mmHg)	Integer (0-200)	Hyperparameter that affects prediction; high resting BP is a critical indicator of heart disease.
Serum Cholesterol	Serum cholesterol level (in mg/dL)	Integer (0-600)	Important for assessing lipid-related risks; high cholesterol often indicates cardiovascular risk.
Fasting Blood Sugar	Blood sugar level after fasting (greater than 120 mg/dL indicates high blood sugar)	0: False, 1: True	Hyperparameter used to check for diabetes; high blood sugar is a significant risk factor for heart disease.
Resting Electrocardiographic Results	Resting electrocardiogram results (measuring heart electrical activity)	0: Normal, 1: ST-T wave abnormality, 2: Left ventricular hypertrophy	Highly significant for diagnosing heart conditions; abnormalities are strong indicators of heart disease.
Max Heart Rate	Maximum heart rate achieved during exercise	Integer (60-220)	Hyperparameter that reflects physical fitness; lower max heart rate may indicate a risk factor.
Exercise Induced Angina	Indicates whether the patient experienced angina during exercise	0: No, 1: Yes	Significant for assessing the intensity of heart disease and exercise tolerance.
Oldpeak	ST depression induced by exercise relative to rest (indicates severity of heart disease)	Float (-2.6 to 6.2)	Strongly linked with heart disease severity; higher values are critical in diagnosing the condition.
Slope of the Peak Exercise ST Segment	Slope of the ST segment during exercise (used to determine the severity)	1: Upsloping, 2: Flat, 3: Downsloping	Hyperparameter that correlates with the severity of heart issues during exertion.
Number of Major Vessels	Number of major vessels colored by fluoroscopy (related to heart blockage)	Integer (0-3)	Crucial for detecting blockage; a higher number of blocked vessels indicates increased severity.
Thalassemia	Thalassemia, a blood disorder related to hemoglobin	3: Normal, 6: Fixed defect, 7: Reversible defect	Important for differentiating between heart disease and other conditions; helps assess overall health.
Target	Whether the patient has heart disease or not (the class attribute)	0: No, 1: Yes	Target attribute for classification; helps the model in final decision-making to predict heart disease.

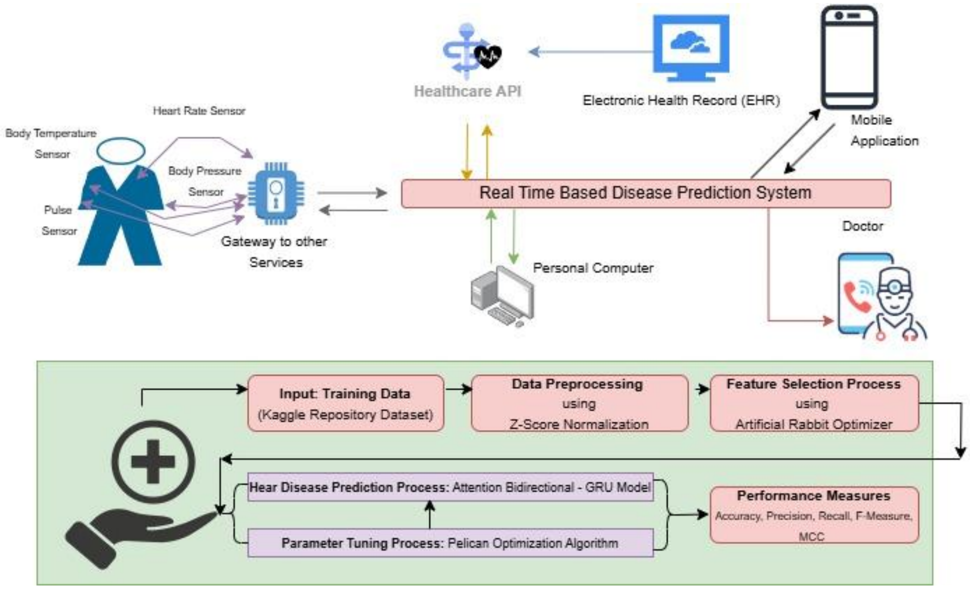

Figure 3: Overall Architecture of the Heart Disease Prediction Model

4. DIMENSIONALITY REDUCTION TECHNIQUES

Feature selection represents the model development process that decreases the number of variables used as input. Reduction of total input variables leads to cheaper modeling expenses as well as enhanced model performance. Statistical-based FS methods examine the correlation between input features and target attributes before finally selecting those variables whose relationship to the target variable is most robust. Statistical measurement techniques for both kinds of variables in the dataset determine their results yet these processes execute efficiently and achieve success [19], [20]. An ML practitioner faces challenges when performing filter-based analysis because they need to choose suitable statistical measures for their datasets.

Three categories of methods exist in supervised learning which include the wrapper and filter and intrinsic approaches. Statistical measures help determine variable correlations during filter-based FS until the most important features become visible for selection. Developing statistical measures for FS requires special attention to understand both the input variable data type and response variable data type during the development phase [21].

$$H\left(P_D\right) = -\sum_{i=1}^{l} P_{D_i} \log_2 P_{D_i} \tag{1}$$

$$IH\left(P_D\right) = \sum_{i=1}^{l} \beta_k^{\gamma} . H^s\left(P_D\right) + \left(\frac{\sum_{i=1}^{l} \beta_k^{\gamma} - 1}{\left(1-\gamma\right)} \right) \tag{2}$$

$$H^s\left(P_D\right) = -\sum_{i \in P_{D_i}} \mu\left(P_{D_i}\right) . \log_2 \rho\left(P_{D_i}\right) \tag{3}$$

- Unsupervised: The analysis must exclude the use of the target variable.

Example: Correlation

- Supervised: Use the target variable.
- Wrapper: Seek feature subsets that demonstrate high performance values. Example: Recursive Feature Elimination.
- Filter: Feature selection needs to be based on features that demonstrate meaning to the objective.

Example: Statistical Methods

- Intrinsic: The training process of certain algorithms lets them pick feature subsets on their own.

Example: DT

Metaheuristic algorithms need implementation due to multiple key elements as shown in this list [35].

- The nature of optimization problems leads to certain instances where one cannot establish a method for locating the complete solution.
- The decision-maker faces less difficulty understanding metaheuristic algorithms from their standpoint.
- Using metaheuristic algorithms to solve problems includes two functions: learning and determining precise answers.
- To create mathematical definitions researchers commonly ignore complex areas that occur in actual problems. When using incorrect data for parameter determination one could encounter significantly greater errors than the partial solution offered by metaheuristic methods.

4.1 Attention Bidirectional Gated Recurrent Unit model

Attention Bidirectional Gated Recurrent Unit (ABiGRU) represents a potent combination which merges Bidirectional Gated Recurrent Unit (BiGRU) with an attention mechanism. Sequential data processing and relevant feature extraction through attention-dependent sequencepart analysis is the goal of this model design. ABiGRU technology serves multiple domains for time-series prediction and medical diagnostics applications alongside natural language processing applications particularly in heart disease diagnosis [22]. Figure 3 shows the Overall Architecture of the Heart Disease Prediction Model.

$$InG(F) = IH(P_D) - H(F) \tag{4}$$

$$S = \frac{\sum_i \left(P_{D_i} - \mu\right)^3}{n\sigma^3} \tag{5}$$

$$K = \frac{\sum_i \left(P_{D_i} - \mu\right)^4}{n\sigma^4} \tag{6}$$

- Bidirectional Processing: The ability to consider both past and future contexts improves performance in tasks where both directions of the data are essential for prediction.
- Attention Mechanism: By focusing on the most relevant features in the sequence, the attention mechanism improves the model's accuracy and interpretability.
- Memory Efficiency: The GRU structure allows the model to handle long-term dependencies efficiently without the need for more complex architectures like LSTMs.
- Enhanced Performance: ABiGRU is particularly effective in tasks that require both contextual understanding and a focus on key data points, such as time-series analysis and sequential decision-making.

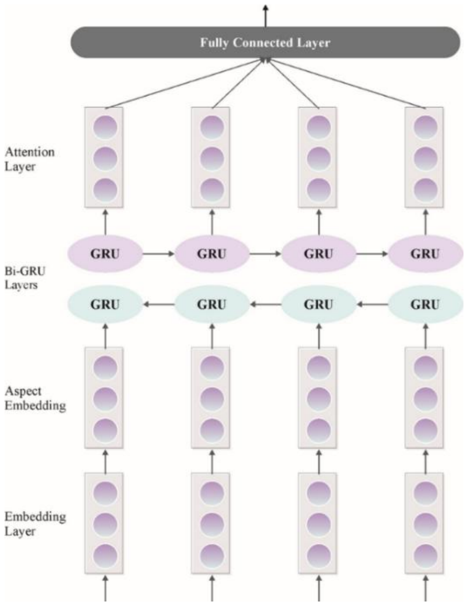

Figure 4: The Architecture of ABiGRU Model

4.2 Convolutional Neural Networks (CNNs):

Deep learning algorithms under the Convolutional Neural Networks (CNNs) specialize in processing orderly data structures including images and both images and videos and time-series data. The ability of CNNs to conduct spatial pattern extraction through convolution operations makes them efficient tools for processing data while becoming top tools for object detection sequence prediction and image classification [23]. CNNs find specific application in healthcare assessments for predicting heart diseases by analyzing ECG signals and medical images through time-series data analysis.

$$I_t = \sigma\left(w_{FSI}.FS_t + w_{HI}.H_{t-1} + B_I\right) \tag{7}$$

$$F_t = \sigma\left(w_{FSF}.FS_t + w_{HF}.H_{t-1} + B_F\right) \tag{8}$$

$$C_t = F_t.C_{t-1} + I_t.\tanh\left(w_{FSC}.FS_t + w_{HC}.H_{t-1} + B_C\right) \tag{9}$$

$$O_t = \sigma\left(w_{FSO}.FS_t + w_{HO}.H_{t-1} + B_O\right) \tag{10}$$

Despite their primary application in images the concept of CNN networks functions effectively with medical data especially time-series data which includes wearables and ECG readings along with medical imaging data such as X-rays and MRIs. CNNs extract automatic hierarchical data features which include:

- For ECG signals: CNNs can automatically identify abnormal patterns in the electrical activity of the heart, which are indicators of potential heart diseases.
- For heart rate or blood pressure monitoring: CNNs can identify trends and sudden fluctuations that are associated with cardiovascular issues.

The convolutional layers analyze input raw data to extract vital features after which predictions regarding heart disease are processed through subsequent layers.

- Input Data: Time-series data from ECG sensors or heart rate monitors.
- Feature Extraction: The CNN extracts key patterns from the raw data.
- Classification: The fully connected layers classify the data, determining whether the input data indicates a healthy heart or potential heart disease.
- Output: The system outputs a prediction, such as the likelihood of ischemic heart disease, which can then be used to guide medical interventions.

4.3 Parameter Optimization Approach (POA)

The Parameter Optimization Approach (POA) exists as a methodology that improves machine learning and deep learning model performance through selecting the best hyperparameter configuration for each model. The performance of deep learning models extensively depends on the correct selection of hyperparameter values which includes learning rate together with batch size and number of layers and additional model-specific parameters.

$$Fu_{sc} = \sum_{i}^{m}\left(sc_{NN} + sc_{DCNN} + sc_{LSTM}\right) \tag{11}$$

$$sc_i = \frac{sc_i - \min sc_i}{\max sc_i - \min sc_i} \tag{12}$$

$$IFu_{sc} = \left(\sum\left(x_0 * sc_{NN}\right) + \sum\left(x_1 * sc_{DCNN}\right) + \sum\left(x_2 * sc_{LSTM}\right)\right) \tag{13}$$

- Learning Rate: Controls how much the model updates with respect to the loss function.
- Batch Size: Number of samples used in one update of the model during training.
- Number of Layers and Neurons: In deep learning models, the depth and width of the neural network architecture can significantly impact model performance.
- Activation Functions: The choice of activation function (e.g., ReLU, Tanh, Sigmoid) affects the model's ability to capture complex relationships in the data.
- Regularization Parameters: Techniques like L1/L2 regularization or Dropout rates to prevent overfitting.

- Optimizer: The optimization algorithm (e.g., Adam, SGD, RMSprop) influences how the model's weights are updated during training.
- Epochs: The number of times the entire training dataset is passed through the model.

```
from sklearn.model_selection import GridSearchCV
param_grid = {'learning_rate': [0.001, 0.01, 0.1],
'batch_size': [16, 32, 64],
'epochs': [10, 20, 30]}
grid_search = GridSearchCV(estimator=Model(), param_grid=param_grid, cv=3)
grid_search.fit(X_train, y_train)
```

- Learning rate: In heart disease prediction, optimizing the learning rate is critical for achieving convergence without overshooting the optimal weights.
- Epochs and batch size: These affect how quickly and effectively the model learns, and tuning them ensures faster and more accurate convergence.

4.4 Confusion Matrix

A **confusion matrix** is a useful tool for evaluating the performance of a classification model, such as one used for heart disease prediction. It helps visualize how well the model is performing by comparing the predicted labels with the actual labels.

True Positive (TP): The number of patients who actually have heart disease and are predicted to have heart disease.

True Negative (TN): The number of patients who do not have heart disease and are predicted to not have heart disease.

False Positive (FP): The number of patients who do not have heart disease but are predicted to have heart disease (Type I error).

False Negative (FN): The number of patients who have heart disease but are predicted not to have heart disease (Type II error).

$$Accuracy = \frac{TP+TN}{TP+T\ N+FP+FN} \tag{14}$$

$$Precision = \frac{TP}{TP+FP} \tag{15}$$

$$Recall = \frac{TP}{TP+FN} \tag{16}$$

$$F1 - Score = 2 * \frac{Precision*Recall}{Precisio+*Recall} \tag{17}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{18}$$

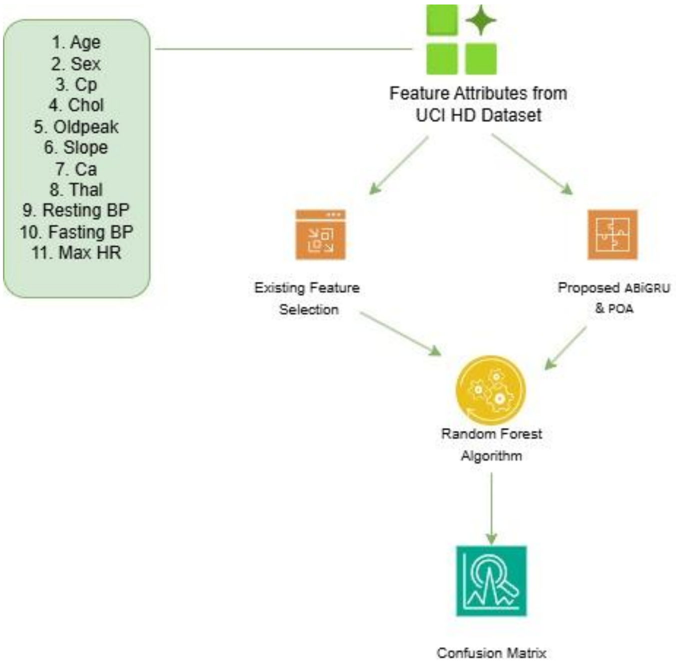

Figure 5: Performance Comparison of the Proposed ABiGRU and POA Models with the Existing Model

5. RESULTS AND DISCUSSION

The primary phase of Experiment workflow uses UCI HD data set which awaits pre-processing. Among the preprocessed HD data set the 13 features include 'Age, Sex, Cp, Trestbps, Chol, Fbs, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, and Thal.' This experiment employed the RF algorithm for the purpose of increasing classification precision. The confusion matrix serves as the model's performance evaluation method. As depicted in Figure 3 the experiment workflow shows the procedures. At this stage relevant attributes get selected for use in the classification process of Ischemic Heart Diseases. The 13 attributes in UCI HD datasets establish critical clinical information which needs to be strongly emphasized. Medical records serve an essential function for both determining HD severity and making essential diagnoses of the condition. Figure 6 shows the Performance Comparison of the Proposed ABiGRU and POA Models with the Existing Model.

Table 3: Comparison of the Efficacy of Different Algorithms

Techniques	Accuracy	Precision	Recall	F1-Measure
MLP-NN	74.22	73.71	78.89	78.41
SVM	75.52	75.31	79.45	78.87
RF	74.22	81.12	79.42	80.55
NB	78.13	80.6	87.08	80.21
Deep CNN	95.03	93.33	93.42	83.61
SE	93.5	93.82	93.87	93.36
Fused ML	93.77	95.4	92.12	93.85
DNN	76.76	78.83	72.49	89.74
DBN	80.15	81	83	80.25
LR	79.09	76	80.15	77.21
Proposed ABiGRU & POA	97.24	97.34	96.51	96.92

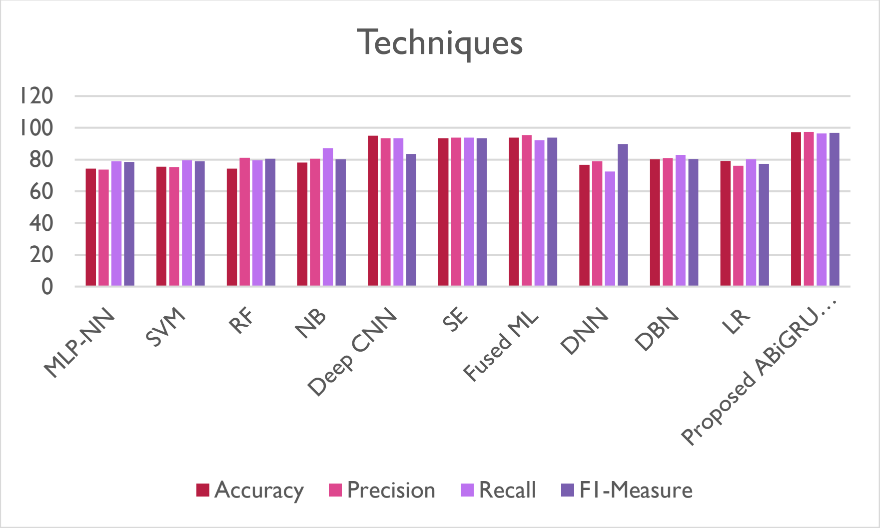

Figure 6: Comparison of the Efficacy of Different Algorithms

Table 3 shows the comparison findings. The implemented ABiGRU & POA model demonstrates superior performance compared to other tested methods resulting in 97.24 percent maximum accuracy. This model functioned as a classification system which enables real-time diagnosis purposes.

Table 4: Comparative Analysis of the Existing Method with the Proposed Method

Classifiers	Accuracy	Sensitivity	Specificity
Bagging SMO	89.1	90.67	85.21
Navie Bayes	62.21	50	95.55
SMO	89.09	90.67	85.21
Neural Network	86.11	90.67	74.86
ICA-KNN	90.02	90.67	74.86
Proposed ABiGRU & POA	97.24	97.66	97.1

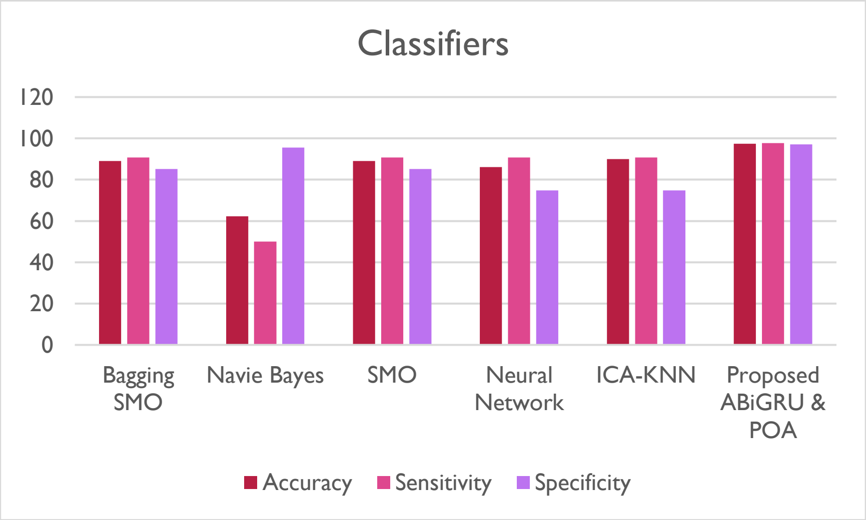

Figure 7: Comparative Analysis of the Existing Method with the Proposed Method

Table 4 shows the Comparative Analysis of the Existing Method with the Proposed Method. The implemented ABiGRU & POA model demonstrates superior performance compared to other tested methods resulting in 97.66 percent maximum Sensitivity. This model functioned as a classification system which enables real-time diagnosis purposes. Figure 7 depict the suggested model, ABiGRU and POA, Random forest, and Extra tree model bar charts, respectively.

Experimental testing of various feature selection methods established the superior performance of the proposed model because it processes high-order data and unlimited attributes and achieves best convergence and balanced exploration-exploitation results. Based on studies [16] the model achieved 93% results before optimization yet obtained 97.24% accuracy after applying the proposed optimization. Four variables continue to exist: ‘‘Cp’’, ‘‘restecg’’, ‘‘oldpeak’’, ‘‘Ca’’, ‘‘thal’’ in the final analysis process. The discarded features are:’’Age’’. The input features include ‘Cp’ and ‘restecg’ and ‘oldpeak’ and ‘Ca’ and ‘thal’.

6. CONCLUSION

A new verification of the Ischemic Heart Disease Squirrel Search Optimization feature selection algorithm takes place using the UCI heart disease dataset. Through the proposed IHDSSO model medical practitioners can select the most vital features that serve as strong indicators for classifying heart disease. The examination proves that heart disease prediction requires primary features consisting of 'Cp', 'restecg', 'oldpeak', 'Ca' and 'thal'. Both the proposed IHDSSO model applied with random forest classifier achieved an accuracy rate of 98.38% for accurate prediction of heart diseases. PDHH-SSO as a proposed model offers substantial support to medical diagnostic solutions for detecting Ischemic heart disease. The existing study requires improvements regarding its convergence accuracy and speed of convergence.

REFERENCES

[1] S. Mohan, C. Thirumalai, and G. Srivastava, ‘‘Effective heart disease prediction using hybrid machine learning techniques,’’ IEEE Access, vol. 7, pp. 81542–81554, Jan. 2019.

[2] R. Soundharya, D. Cenitta, and R. V. Arjunan, ‘‘Information concealment and redemption through data anonymization technique,’’ *J. Adv. Res. Dyn. Control Syst.*, vol. 10, no. 7, pp. 22–26, 2018.

[3] *World Health Organization*. [Online]. Available: https://www.who. int/news-room/fact-sheets/detail/cardiovascular-diseas%es-(cvds)

[4] S. C. Yogaamrutha, D. Cenitta, and R. V. Arjunan, ‘‘Forecast of coronary heart disease using data mining classification technique,’’ *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 4, pp. 25–36, 2019.

[5] D. Cenitta, R. V. Arjunan, and K. Prema, ‘‘Missing data imputation using machine learning algorithm for supervised learning,’’ in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2021, pp. 1–5.

[6] D. Cenitta, R. V. Arjunan, and K. Prema, ‘‘Cataloguing of coronary heart malady using machine learning algorithms,’’ in *Proc. 4th Int. Conf. Electr., Comput. Commun. Technol. (ICECCT)*, Sep. 2021, pp. 1–6.

[7] T. S. L. V. Ayyarao, N. S. S. Ramakrishna, R. M. Elavarasan, N. Polumahanthi, M. Rambabu, G. Saini, B. Khan, and B. Alatas, ‘‘War strategy optimization algorithm: A new effective Metaheuristic algorithm for global optimization,’’ *IEEE Access*, vol. 10, pp. 25073–25105, 2022.

[8] J. Nourmohammadi-Khiarak, M.-R. Feizi-Derakhshi, K. Behrouzi, S. Mazaheri, Y. Zamani-Harghalani, and R. M. Tayebi, ‘‘New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection,’’ Health Technol., vol. 10, no. 3, pp. 667–678, May 2020.

[9] F. A. M. Al-Yarimi, N. M. A. Munassar, M. H. M. Bamashmos, and M. Y. S. Ali, ‘‘Feature optimization by discrete weights for heart disease prediction using supervised learning,’’ Soft Comput., vol. 25, no. 3, pp. 1821–1831, Feb. 2021.

[10] M. A. Khan and F. Algarni, ‘‘A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO- ANFIS,’’ IEEE Access, vol. 8, pp. 122259–122269, 2020.

[11] T. Zheng and W. Luo, ‘‘An improved squirrel search algorithm for optimization,’’ Complexity, vol. 2019, pp. 1–31, Jul. 2019.

[12] M. Kader, ‘‘Comparative study of five metaheuristic algorithms for team formation problem,’’ in Human-Centered Technology for a Better Tomorrow. Cham, Switzerland: Springer, 2022, pp. 133–143.

[13] C. A. Subasini, S. P. Karuppiah, A. Sheeba, and S. Padmakala, ‘‘Devel- oping an attack detection framework for wireless sensor network-based healthcare applications using hybrid convolutional neural network,’’ Trans. Emerg. Telecommun. Technol., vol. 32, no. 11, Nov. 2021, Art. no. e4336.

[14] M. S. Sanaj and P. M. Joe Prathap, ‘‘Nature inspired chaotic squirrel search algorithm (CSSA) for multi objective task scheduling in an IAAS cloud computing atmosphere,’’ Eng. Sci. Technol., Int. J., vol. 23, no. 4, pp. 891–902, Aug. 2020.

[15] S. Joteppa, S. K. Balraj, N. Cheruku, T. R. Singasani, V. Gundu, and A. Koithyar, ''Designing a smart IoT environment by predicting chronic kidney disease using kernel based xception deep learning model,'' Revue d'Intelligence Artificielle, vol. 38, no. 1, pp. 303–312, Feb. 2024.

[16] K. Venkatrao and S. Kareemulla, ''HDLNET: A hybrid deep learning network model with intelligent IoT for detection and classification of chronic kidney disease,'' IEEE Access, vol. 11, pp. 99638–99652, 2023.

[17] E. A. Refaee and S. Shamsudheen, ''A computing system that integrates deep learning and the Internet of Things for effective disease diagnosis in smart health care systems,'' J. Supercomput., vol. 78, no. 7, pp. 9285–9306, May 2022.

[18] K. Venkatrao and K. Shaik, ''CAD-CKD: A computer aided diagnosis system for chronic kidney disease using automated BiGSqENet in the Internet of Things platform,'' Evolving Syst., vol. 15, pp. 1487–1502, Mar. 2024.

[19] A. Ashraf, Z. Qingjie, W. H. K. Bangyal, and M. Iqbal, ''Analysis of brain imaging data for the detection of early age autism spectrum disorder using transfer learning approaches for Internet of Things,'' IEEE Trans. Consum. Electron., vol. 70, no. 1, pp. 4478–4489, Feb. 2024.

[20] A. Sundas, S. Badotra, G. S. Shahi, A. Verma, S. Bharany, A. O. Ibrahim, A. W. Abulfaraj, and F. Binzagr, ''Smart patient monitoring and recommendation (SPMR) using cloud analytics and deep learning,'' IEEE Access, vol. 12, pp. 54238–54255, 2024.

[21] Y. Cao, F. Yang, Q. Tang, and X. Lu, ''An attention enhanced bidirectional LSTM for early forest fire smoke recognition,'' IEEE Access, vol. 7, pp. 154732–154742, 2019.

[22] Y. Cao, X. Su, Q. Tang, S. You, X. Lu, and C. Xu, ''Searching for better spatio-temporal alignment in few-shot action recognition,'' in Proc. Adv. Neural Inf. Process. Syst., vol. 35, 2022, pp. 21429–21441.

[23] Y. Cao, Q. Tang, F. Yang, X. Su, S. You, X. Lu, and C. Xu, ''Re-mine, learn and reason: Exploring the cross-modal semantic correlations for language- guided HOI detection,'' in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2023, pp. 23492–23503. Surgery New2