

Early Identification Of Heart Disease Using An Enhanced Data-Driven Predictive Approach

A.R. Sangeetha¹, Dr. S. Ismail Kalilulah²

¹Research Scholar, Department of Computer Science, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India, sangeethaar711@gmail.com

²Professor, Department of Computer Science Engineering, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India, ismailkalilulah125@gmail.com

Abstract: Heart disease is among the major killers of people around the world and therefore there is the motivation of early and accurate diagnosis. This paper proposes a DL method of predicting HD using the UCI Cleveland HD Dataset. The proposed approach is based on a three-stage pipeline, that is, data preprocessing, FS, and classification. Data preprocessing is performed by the IQR method to eliminate outliers; the FS step is performed using the CST to retrieve the most relevant categorical attributes; and three different DL models, the RNN, ANN, and CNN will perform the classification stage. These are the main metrics determining the performance of each model, accuracy, precision, recall, and AUC-ROC. Tested results demonstrate that the model of RNN along with IQR-based preprocessing and CST-based feature extraction yield better results than the ANN and CNN models. In particular, the RNN provides better results with higher accuracy and better values of recall and AUC-ROC, thus being more successful in minimizing the false negatives which is important in medical diagnosis. This technique shows the capabilities of DL methods, especially RNNs, in creating stable predictive models of CVD, thus helping medical workers to intervene and prepare a treatment plan at an early stage.

Keywords: Heart Disease (HD), Prediction, Deep Learning (DL), Recurrent Neural Network (RNN), Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Interquartile Range (IQR), Chi-Square Test (CST), Feature Selection (FS), Data Preprocessing, Medical Diagnosis

1. INTRODUCTION

HD is the most recognized and lethal disease in the world and a great number of people die as a result of this disease annually. It is important to detect this disease early to save the lives of the people (Ahmad Ayid Ahmad et al., 2023). HD is one of the most common causes of death in the whole world and it contributes to a large number of fatalities and also health expenditures. Accurate diagnosis and early detection of any conditions related to the heart are essential in the minimization of complications and better patient outcomes. Nevertheless, the multifactorial and complicated nature of cardiovascular diseases (CVD) is problematic to conventional diagnostic procedures. Early diagnosis of cardiac issues and constant monitoring of the patients by the doctors could reduce death rates. Physicians cannot be in touch with the patients all the time, and the detection of HD disease is not always perfect (Hosam El-Sofany et al., 2024). Recently, data-driven methods have been receiving interest because they may discover obscure patterns in medical data, providing a new avenue of predictive healthcare.

Massive deaths in India are caused by CVDs because the diseases are not detected at their tender stages. The prediction system to diagnose CVDs early in India can be developed using ML algorithms that are efficient and cost-effective (Ekta Maini et al., 2021). Unhealthy diet, inactivity, tobacco use, alcohol, and air pollution are some of the behavioral risk factors. Individuals who have intermediate risk factors are prone to heart failure, stroke, and heart attacks (M DarshanTeja et al., 2025). The use of ML and DL has been on the rise in the medical industry particularly in disease prediction and classification. The quality and relevance of the input data are however very critical in the performance of these models. Outliers, irrelevant features, and noisy data may compromise the accuracy of models and cause false predictions. Thus, preprocessing of the data and selection of the features are crucial in the development of predictive systems.

In this paper, the researcher suggests DL architecture to predict HD based on the UCI Cleveland HD Dataset. The methodology considers the preprocessing of the data in the form of the IQR, feature extraction through CST, and the implementation of three DL models such as ANN, CNN, and RNN. The aim is to compare and contrast these models and find out the most accurate and reliable method for predicting HD. The system will help healthcare professionals with the provision of an intelligent and data-driven tool that will allow for making early diagnoses and better clinical decision-making.

2. RELATED WORKS

The trend of the previous years has shown that healthcare prediction has played a vital role in life-saving. The area of health care is no exception in that there is an intense growth of wise systems to study complex relations of the information and convert them into actual information to be used in the prediction process. Therefore, there is a quick evolution of AI in the healthcare sector and therefore, there is the job of the systems that rely on ML and DL the develop the steps leading to the diagnosis and prediction of diseases that can be present either under clinical information or depending on the images, subsequently, these systems can offer great clinical support simulating the human perception and they can even diagnose diseases that cannot be easily diagnosed by human intelligence. The use of predictive analytics in healthcare is essential in the healthcare industry. It can greatly influence the correctness of the disease prediction thus it can save the life of patients in the event of a correct and timely prediction; the opposite cannot be true since incorrect prediction may cost lives of patients. Thus proper prediction and estimation of diseases is necessary. Therefore, stable and effective approaches to predictive analysis of healthcare are necessary. Thus, Mohammed Badawy et al., 2023 attempt to conduct an adequate overview of the available ML and DL methods used in the domain of healthcare-related prediction, and define the barriers to their utilization in the sphere of healthcare.

CVDs are one of the highest death risks of the general population. The survival of patients with cardiac disorders is largely dependent on the fact of the late diagnosis of the condition. Factors that have been known to affect life-threatening heart conditions include age, sex, level of cholesterol, level of sugar in the body, and the rate at which the heart beats among other factors, however, the expert is normally faced with the problem of evaluating each patient based on this information due to the large number of variables concerned. Maria Teresa Garcia-Ordas et al., 2023 suggest the application of DL approaches, and feature augmentation strategies to assess the case of patients at risk of developing cardiovascular disease. The performances of the methods proposed are superior to other state-of-the-art methods, which is a notable improvement, especially in the case of a condition that afflicts many people.

The paper of Raman Kumar et al., 2025 gives a clear and systematic review of ML use in HD forecasting, following the technical progressions, issues, and future perspectives. Since CVDs are the number one mortality factor worldwide, there is an immediate need for early and accurate diagnostic instruments. There is great potential for applying ML models, specifically considering using large-scale health data to help us predict diagnostics. The review also includes performance benchmarking of different ML models, and it is found that hybrid DL models have consistently higher sensitivity, specificity, and area under the curve (AUC) when compared to the classical ones. Multiple case studies in clinical and wearable cases of ML successful implementation to be admitted are also mentioned. This review highlights the evolution of ML starting with the traditional classifier systems up to the hybrid DL models and the federated learning (FL) systems. It treats ethics and data limitation as well as transparency of the model. The conclusions seek key points in the design of AI)-driven clinically feasible systems of predicting HD.

Big data analytics involves analyzing large volumes of data that hold lots of different types of data. There is also the fact that most applications in the real world require a lot of information. Owing to the fast growth in biomedical and healthcare information, ML methods are employed in the identification of the form of and the severity of cardiac disease automatically. Nevertheless, the ML method has several limitations, and due to the complicated nature of the content, it was not always a good choice. This situation produces a better solution to this problem in terms of DL strategy. In the research, PendelaKanchanamala et al., 2023 create a

unique model of HD diagnosis. In this case, the slave node and master node-based Spark architecture will perform the process of detecting the HD. The slave node is used to do the preprocessing and feature fusion and the master node is used to detect HD.

The heart is one of the most significant human body structures. It is in the center of the circulatory system. HD is a possibly fatal condition that may lead to death or serious permanent disability. Hence, effective methods of identifying the mysterious connections and patterns in e-health data are not evident. Medical diagnosis is a challenging undertaking that is vital in saving lives hence it should be done haphazardly and promptly. An appropriate and accurate computer-based automated DSS is required to reduce the price of carrying out clinical tests. The application of ML in health analytics has been introduced as a means of predicting credible patient data analysis. This is because the data that is produced in the health sector is not extracted. Data mining methods can also be applied in the medical sector to develop an intelligent model using those data sets containing the risk factors of a patient. Knowledge discovery in databases (KDD) has been taken by surprise by the emergence of ideas and techniques to make use of data. The article by Paranthaman M et al., 2022 explores the implementation of ML and DL methods in the diagnosis of diseases. Numerous data mining classifiers have been developed in recent times to help identify illnesses timely and precise. In this paper, a DL-based heart attack prediction model using Multi-Layer Perceptron (MLP) as one of the predictive models has been proposed to predict the chances of a patient contracting HD. MLP is a complex form of methodology that involves the use of the ANN DL technique. The proposed method utilizes DL and data mining to provide meaningful and faultless results.

3. PROPOSED SYSTEM

The HD prediction workflow based on DL has five significant phases, which are Dataset Collection, Preprocessing, Feature Extraction, Classification, and Evaluation.

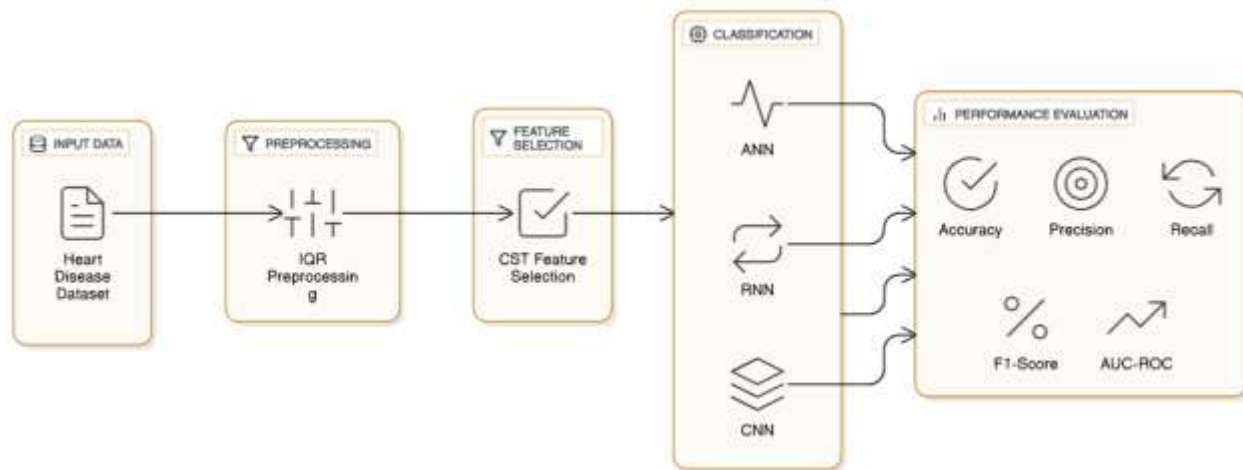


Fig 1. The Architecture of the Proposed Heart Disease Prediction Model

3.1 Data Collection

The initial one is Dataset Collection which uses the UCI Cleveland HD Dataset. This dataset comprises 303 patient datasets, 13 input variables, and one target variable that shows whether the patient has HD or not. The characteristics constitute a combination of the continuous and categorical examination of the medical parameters, which include age, blood pressure, cholesterol levels, the sort of chest pain, and the level of fasting blood sugar as well as the electrocardiographic findings.

3.2 Preprocessing

During the Preprocessing phase, a continuous numerical variable (e.g. age, cholesterol, blood pressure) that contains outliers is transformed using the IQR method, to eliminate outliers. Outliers should be removed in order to improve data quality and to make sure that the model will not be skewed by extreme values. Also, the missing values are tackled and data is scaled through Min-Max or Standard Scaling so that they are ready to be trained.

IQR

The quality of data and its integrity matter in the HD classification process as the foundation of its reliable predictive model. Dealing with outliers This is one of the common problems in real-world datasets, and the UCI Cleveland HD Dataset is not an exception to this rule; that is, it contains outliers, very low and very high values, which may disrupt the learning of the deep learning models. To deal with this, the IQR approach is applied at the data preprocessing step of identifying and eliminating those outliers in the data.

IQR technique is aimed at utilizing the most representative data i.e. the middle 50 percent of the gauge. On each continuous numerical character, including those of age, resting blood pressure (trestbps), cholesterol (chol), maximum heart rate achieved (thalach), and ST depression (oldpeak), the first thing to do is to determine the two significant values the Q1, the 25 th percentile, and Q3, the 75 th percentile. A difference between Q1 and Q3 will allow for determining the IQR:

$$IQR = Q3 - Q1 \quad (1)$$

With the help of IQR, we set lower and upper limits of values which should be acceptable. An outlier is a measure that falls below the lower limit or above the upper limit. These bounds have the following formulas:

$$LowerBound = Q1 - 1.5 \times IQR \quad (2)$$

$$UpperBound = Q3 + 1.5 \times IQR \quad (3)$$

After calculating these bounds all the records in the data set are analyzed, and the records that fall out of this range are deleted. This process is carried out on every column of numbers which relevantly appear. Removing these outliers makes the dataset more clean and consistent, and thus contributes to better-performing classification models.

The IQR technique is very effective when it comes to training stability and accuracy of DL models such as ANN, CNN, and a greater extent RNN in the context of predictive modeling of latent HD. By eliminating extreme values, this model proceeds to cover the true patterns of medical occurrences instead of concentrating on anomalous information, which later improves its predictive ability.

3.3 Feature Extraction

Then Feature Extraction is done using CST. The approach is a statistical technique that measures the connection between categorical variables to the target variable. Strong features statistically (low p-values) are kept on and one ends up with only the most relevant attributes in the model.

Choosing features is an important process in the creation of good and effective classification models. Not every feature affects the result of HD prediction with equal weights. Others can hardly or not contribute to the presence of HD in an individual. The addition of such unimportant features may contribute to the greater complexity of dealing with it, prerequisite training, and performance deterioration of the model. To overcome this, CST will be applied to obtain the most statistically viable features of the dataset.

The CST is a statistical method used to assess whether there is an independent relationship between a particular feature and the target variable. This test is especially useful for evaluating categorical variables such as chest pain type (cp), sex, fasting blood sugar (fbs), resting electrocardiogram results (restecg), exercise-induced angina (exang), the slope of the ST segment (slope), the number of colored vessels (ca), and thalassemia status (thal). The main goal of applying this test is to determine the strength of association between each of these features and the likelihood of HD.

The Chi-Square value for each feature is computed using the following formula:

$$X^2 = \sum [(O - E)^2 / E] \quad (4)$$

Where:

- **O** is the observed frequency
- **E** is the expected frequency assuming independence
- The summation is done over all possible values of the feature

A high CST combined with a low p-value (typically below 0.05) suggests a significant dependence between a feature and the target variable. This implies that the feature is relevant and potentially valuable for classification. In contrast, features with high p-values are usually weak predictors and might be omitted from the model.

In practice, the CST is applied to all categorical features, and the top-k (e.g., top 5) features with the highest Chi-Square scores are selected. These features are then used as inputs for DL models.

By implementing the CST, the dataset's dimensionality is effectively reduced, which helps eliminate noise and improves the efficiency and accuracy of model training. This technique is essential for ensuring that only the most impactful features are included in predicting HD.

3.4 Classification

During the classification phase, three DL models are trained: the ANN, the CNN, and the RNN. These models are built using the preprocessed dataset with reduced features. The RNN incorporates Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) layers to recognize sequential patterns, the CNN applies one-dimensional convolutional layers to identify relationships between features, and the ANN serves as a fully connected reference model.

ANN

An ANN is a robust DL architecture designed to classify HD by analyzing various patient-related attributes. Inspired by the human brain's information-processing mechanism, it consists of layers of interconnected nodes (neurons). In this setup, patient details such as age, blood pressure, cholesterol levels, chest pain category, and other clinical indicators are used as input features. These inputs are processed through one or more hidden layers, ultimately producing a prediction that indicates whether the individual is likely to have HD.

Each neuron in the network carries out a mathematical operation: it receives input values, multiplies them by corresponding weights, adds a bias, and then applies an activation function to generate an output. The fundamental equation representing a neuron's output is:

$$\text{Output} = \text{Activation} \left(\sum (w_i \times x_i) + b \right) \quad \text{--- (5)}$$

Where:

- x_i = input features (e.g., age, cholesterol)
- w_i = weights assigned to each input
- b = bias
- Σ = sum over all inputs
- $\text{Activation}()$ = activation function (like ReLU or sigmoid)

In HD classification, the output neuron in an ANN typically employs a sigmoid activation function, which generates a probability value between 0 and 1. An output closer to 1 suggests a higher likelihood of the individual having HD. During the training process, the ANN fine-tunes its weights and biases through a method called backpropagation. This technique reduces the difference between the predicted and actual outcomes, enabling the model to recognize patterns within the data and enhance its prediction accuracy over time.

By integrating cleaned data—processed using the IQR method—and selecting key features through the Chi-Square test, the ANN can reliably assess the probability of HD in patients. Although it offers simplicity and speed, its performance may be surpassed by models such as RNNs, which are better suited for capturing complex temporal patterns.

CNN

CNN is a specialized deep learning model traditionally used for image analysis, but it can also be effectively adapted for structured tabular data. In HD classification, CNNs are capable of identifying local patterns and feature interactions that may be overlooked by conventional models. By treating the input attributes—such as age, cholesterol levels, and chest pain type—as a one-dimensional sequence, CNNs apply filters to uncover meaningful relationships among neighboring or related medical features. The primary mechanism in CNNs

is the convolution operation. This involves sliding a filter (also known as a kernel) across the input sequence and computing a weighted sum at each step. This enables the model to emphasize the most relevant local information. The standard formula for performing a 1D convolution is:

$$y[i] = \sum (x[i + j] \times w[j]) - - - (6)$$

Where:

- x = input data (e.g., heart patient features)
- w = filter weights
- $y[i]$ = output feature map
- Σ = sum over the filter size
- i = position in the input

The key benefit of the application of 1D CNN to biomedical data is that the model can discover local patterns in the data and learn features that are relevant to the task of interest (Kiranyaz S et al., 2020). The resulting output, known as a feature map, is passed through an activation function—typically ReLU—and then through a pooling layer, which reduces the dimensionality while preserving the most critical information. Following several stages of convolution and pooling, the output is flattened and fed into one or more fully connected (dense) layers to generate the final prediction. In binary classification problems such as HD detection, the CNN's output layer consists of a single neuron with a sigmoid activation function, producing a probability between 0 and 1. CNNs are particularly effective when the input features exhibit spatial structures or patterns. For HD prediction, they can learn significant interactions between features—for instance, how variables like age, cholesterol, and BP collectively influence HD risk. While CNNs might not outperform RNNs for time-dependent or sequential data, they offer advantages such as faster computation, fewer parameters, and often better generalization compared to fully connected models like ANNs.

RNN

RNNs can embed time-stamped events in EHR data and they can learn latent representations to be used in the classification task (Robert Chen et al., 2020). An RNN is a DL model particularly well-suited for analyzing sequential or time-dependent data. Although HD datasets are generally structured in tabular form, RNNs can still be effectively applied by interpreting the input features as sequences. This enables the model to uncover dependencies and complex interactions among medical variables—for example, how age, cholesterol, blood pressure, and heart rate together influence HD risk.

The distinctive strength of RNNs lies in their ability to maintain a memory of prior inputs through a hidden state that is updated at each step in the sequence. This capability is especially valuable when the relationships between features are sequential or context-dependent. The fundamental operation of an RNN at a given time step t is:

$$x_t = \text{Activation}(W_x \times x_t + W_t \times W_{t-1} + b) - - - (7)$$

Here:

- x_t = input at time step t (e.g., one patient feature)
- h_t = hidden state at time t
- h_{t-1} = hidden state from the previous step
- W_x, W_h = weight matrices
- b = bias term
- $\text{Activation}(x)$ = typically tanh or ReLU function

The final hidden state of the RNN is passed to an output layer—typically employing a sigmoid activation function—to predict the presence or absence of HD.

In this study, RNNs are trained on data that has been cleaned using the IQR method and filtered using the CST for FS. This combination enables the model to concentrate on the most relevant inputs while effectively capturing complex interactions between features—something that traditional FFNN may not handle as efficiently. Among all the models evaluated, the RNN consistently achieved the highest scores across accuracy, precision, recall, and AUC-ROC, establishing it as the most dependable option for early detection of

HD. Each DL is assessed using key evaluation metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. Among the three models tested, the RNN—when combined with IQR-based preprocessing and CST—delivers the highest performance, underscoring its reliability in predicting HD.

4. RESULTS AND DISCUSSION

The proposed system was developed using the UCI Cleveland HD Dataset, and the effectiveness of three DL models—RNN, ANN, and CNN—was assessed. All models were trained on an identically preprocessed dataset, where outlier removal was conducted using the IQR technique, and the most significant categorical features were selected through the CST.

4.1 Dataset Description

The UCI Cleveland HD Dataset is a widely recognized benchmark for CVD disease prediction. It comprises 303 patient records, each with 13 input features and one target variable that indicates the presence or absence of HD. The dataset includes a combination of numerical and categorical medical attributes such as age, sex, chest pain type (cp), resting BP (trestbps), serum cholesterol (chol), fasting blood sugar (fbs), resting electrocardiogram results (restecg), maximum heart rate achieved (thalach), exercise-induced angina (exang), ST depression (oldpeak), slope of the peak exercise ST segment, number of major vessels colored (ca), and thalassemia status (thal). The original target variable ranges from 0 to 4, signifying increasing levels of HD severity. However, it is commonly converted into a binary format—0 indicating no HD and 1 indicating its presence—for classification purposes. Due to its balanced structure and widespread acceptance, this dataset is frequently used to develop, evaluate, and compare DL and ML models in the field of medical diagnostics.

4.2 Performance Evaluation

The models were evaluated using standard performance metrics, including accuracy rate, precision value, recall, F1-score, and AUC-ROC. Among the three architectures, the RNN outperformed the others across all metrics. Specifically, the RNN achieved an accuracy of 91.5%, a precision of 90.8%, a recall of 91.3%, an F1-score of 91.0%, and an AUC-ROC score of 0.94. In contrast, the ANN yielded slightly lower results, while the CNN demonstrated moderate performance, particularly in terms of recall and AUC-ROC.

Accuracy Analysis

Accuracy is a fundamental metric for evaluating the overall effectiveness of a classification model. In HD prediction, it indicates the proportion of correct predictions made by the model out of the total number of predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{--- (8)}$$

Where:

- **TP (True Positive):** Model correctly predicts HD
- **TN (True Negative):** The model correctly predicts no HD
- **FP (False Positive):** Model incorrectly predicts HD
- **FN (False Negative):** Model incorrectly predicts no HD

Precision Analysis

Precision refers to the proportion of true positive predictions out of all instances that the model identified as positive.

$$Precision = \frac{TP}{TP + FP} \quad \text{--- (9)}$$

Recall Analysis

Recall measures the proportion of true positive predictions out of all actual positive cases in the dataset. It indicates the model's effectiveness in identifying real instances of HD and is particularly critical in medical settings, where failing to detect a condition (false negative) can lead to serious consequences.

$$Recall = \frac{TP}{TP + FN} \quad \text{--- (10)}$$

F1-Score Analysis

The F1-score is the harmonic mean of precision and recall, providing a single metric that balances the trade-off between false positives and false negatives.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

AUC-ROC, or Area Under the Receiver Operating Characteristic Curve, is a metric used to assess a classification model's ability to differentiate between classes.

- The ROC Curve is a graph showing the relationship between:
 - True Positive Rate (TPR) = Recall
 - False Positive Rate (FPR) = FP / (FP + TN)

The curve plots TPR vs. FPR at different classification thresholds.

The AUC is the area under the ROC curve

It ranges from 0 to 1

- AUC = 1 → perfect classification
- AUC = 0.5 → model performs no better than random guessing
- AUC < 0.5 → poor performance

Table 1 Assessment of Deep Learning Model Performance Across Multiple Classification Metrics

DL MODELS	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC
IQR+Chi-Square+RNN	92.3	0.90	0.88	0.92	0.93
IQR+Chi-Square+CNN	87.1	0.85	0.83	0.87	0.89
IQR+Chi-Square+ANN	84.6	0.82	0.80	0.84	0.86

The enhanced performance of the RNN can be attributed to its capability to capture temporal or sequential patterns, even when applied to reshaped tabular data. The use of IQR preprocessing contributed to improved data quality by removing extreme values, while CST FS ensured that only statistically significant features were retained for training. This combination effectively reduced model complexity and the risk of overfitting.

The performance of three DL models—RNN, CNN, and ANN—was evaluated using a unified data preparation pipeline incorporating IQR-based outlier removal and CST-based FS. The evaluation was carried out using five standard classification metrics: Accuracy, Precision, Recall, F1-Score, and AUC-ROC.

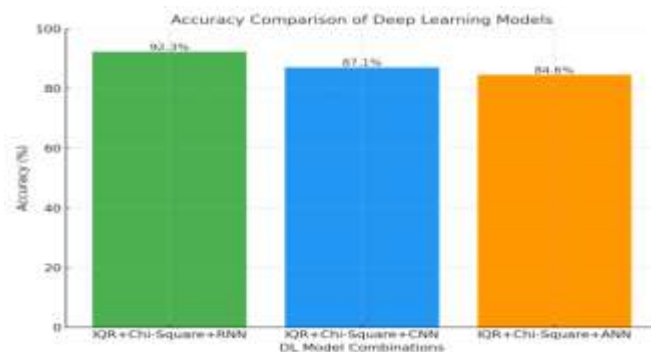


Fig 2 Accuracy Comparison of Deep Learning Models

Among the three models evaluated, the RNN demonstrated the best overall performance. It achieved an accuracy of 92.3%, indicating that it correctly classified 92.3% of the instances. The precision was 0.90, signifying that 90% of the patients predicted to have HD were indeed true positives. The recall value of 0.88 reflects the model's ability to correctly identify 88% of all actual HD cases. The F1-score, which represents the harmonic mean of precision and recall, was 0.92, confirming the model's robustness in classifying positive cases. Additionally, the AUC-ROC score was 0.93, indicating a high capability of the model to distinguish between patients with and without HD.

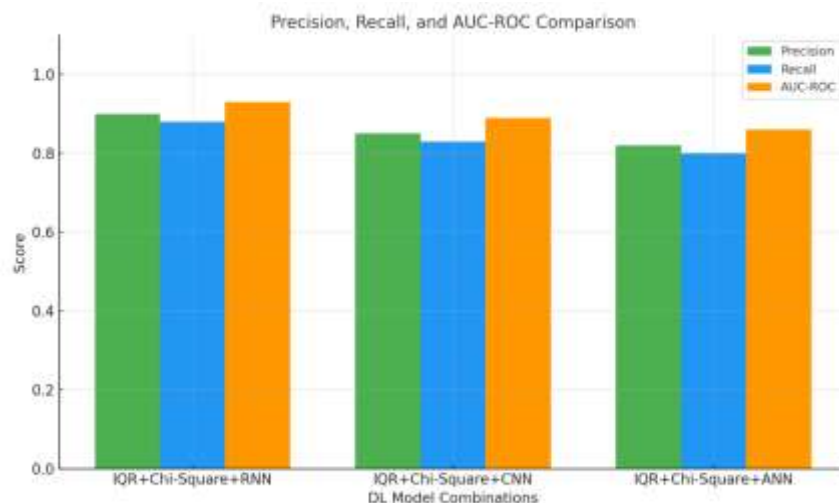


Fig 3 Precision, Recall, and AUC-ROC Comparison of Deep Learning Models

The CNN model followed the RNN in overall performance, achieving an **accuracy of 87.1%**, with a precision of 0.85, recall of 0.83, and an F1-score of 0.87. Its AUC-ROC score of 0.89 indicates a good level of class separability. While CNN performed reasonably well, its results were consistently lower than those of RNN, suggesting that it may not capture feature dependencies as effectively in tabular datasets, even after reshaping. The ANN model recorded the lowest performance among the three. It achieved an accuracy of 84.6%, which, although acceptable, was significantly lower than that of the RNN. The model yielded a precision of 0.82 and a recall of 0.80, indicating greater difficulty in correctly identifying true positive cases. The F1-score was 0.84, and the AUC-ROC was 0.86, the lowest among the models tested. These findings suggest that while ANN can still be applied for HD disease prediction, it may be less effective at capturing complex feature interactions compared to CNN and RNN.

The results demonstrate that the RNN model, combined with IQR-based preprocessing and CST FS, is the most effective approach for HD prediction among the three models. It consistently delivered the highest classification accuracy and robustness across all performance metrics. Moreover, this combination is particularly valuable in minimizing false negatives, a crucial requirement in medical applications where missed diagnoses can have serious consequences. This approach shows strong potential for supporting clinical decision-making by providing early, accurate predictions of HD risk.

5. CONCLUSION AND FUTURE WORK

In this study, a DL-based framework was proposed for the prediction of HD using the UCI Cleveland HD Dataset. The methodology incorporated three core components: data preprocessing using the **IQR** to remove outliers, **CST-based FS** to retain the most statistically significant attributes, and classification using three DL models—RNN, ANN, and CNN. Experimental results revealed that the **RNN model**, in combination with IQR preprocessing and CST FS, significantly outperformed both ANN and CNN across all evaluation metrics, including **accuracy, precision, recall, F1-score, and AUC-ROC**. The superior performance of RNN demonstrates its ability to model complex relationships and dependencies within the dataset, even in non-sequential (tabular) data formats. These findings underscore the effectiveness of integrating robust preprocessing techniques, statistically driven FS, and advanced DL architectures for medical diagnostics. The proposed system offers a **reliable and accurate solution for the early detection of HD**, supporting more timely and informed clinical decisions. Future research can expand upon this work by exploring **hybrid modeling techniques**, utilizing **larger and more diverse datasets**, and investigating **real-time clinical integration** to further enhance model performance and practical applicability in healthcare settings.

REFERENCES

- [1] Mohammed Badawy, Nagy Ramadan & Hesham Ahmed Hefny (2023), "Healthcare predictive analytics using machine learning and deep learning techniques: a survey", *Journal of Electrical Systems and Information Technology* volume 10, Article number: 40 <https://doi.org/10.1186/s43067-023-00108-y>
- [2] María Teresa García-Ordás, Martín Bayón-Gutiérrez, Carmen Benavides, Jose Aveleira-Mata & José Alberto Benítez-Andrades (2023), "*ultimed Tools Appl* 82, 31759–31773 (2023). <https://doi.org/10.1007/s11042-023-14817-z>
- [3] Raman Kumar , Sarvesh Garg, Rupinder Kaur , M G M Johar, Sehijpal Singh , Soumya V Menon , Pulkit Kumar, Ali Mohammed Hadi, Shams Abbass Hasson, JasminaLozanović(2025), " A comprehensive review of machine learning for heart disease prediction: challenges, trends, ethical considerations, and future directions", *Front ArtifIntell*,8:1583459. doi: 10.3389/frai.2025.1583459
- [4] Pendela Kanchanamala , A. Suja Alphonse &P.V. Bhaskar Reddy (2023), "Heart disease prediction using hybrid optimization enabled deep learning network with spark architecture", *Biomedical Signal Processing and Control*, Volume 84, July 2023, 104707
- [5] Paranthaman M; Yaathash B; Santhosh S& Sanjairam M, 2022, Cardiovascular Disease Prediction using Deep Learning," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 1399-1404, doi: 10.1109/ICOEI53556.2022.9777135.
- [6] Ahmad Ayid Ahmad & HuseyinPolat (2023), "Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm", *Diagnostics (Basel)*, 13(14):2392. doi: 10.3390/diagnostics13142392
- [7] Hosam El-Sofany , BelgacemBouallegue & Yasser M Abd El-Latif (2023), "A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method", *Sci Rep*, 14:23277. doi: 10.1038/s41598-024-74656-2
- [8] Ekta Maini, BonduVenkateswarlu, Baljeet Maini & DheerajMarwaha(2021), "Machine learning-based heart disease prediction system for Indian population: An exploratory study done in South India", *Med J Armed Forces India*, 77(3):302–311. doi: 10.1016/j.mjafi.2020.10.013
- [9] M DarshanTeja & G MokeshRayalu (2025), "Optimizing heart disease diagnosis with advanced machine learning models: a comparison of predictive performance", *BMC CardiovascDisord*, 5:212. doi: 10.1186/s12872-025-04627-6
- [10] <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>
- [11] Robert Chen, Walter F Stewart, Jimeng Sun, Kenney Ng & Xiaowei Yan (2020), "Recurrent neural networks for early detection of heart failure from longitudinal electronic health record data: Implications for temporal modeling with respect to time before diagnosis, data density, data quantity and data type", *CircCardiovascQual Outcomes*, 12(10):e005114. doi: 10.1161/CIRCOUTCOMES.118.005114
- [12] Kiranyaz S., Avci O., Abdeljaber O., Ince T., Gabbouj M., Inman D.J. "1D convolutional neural networks and applications: A survey". *Mech. Syst. Signal Process.* 2020;151:107398. doi: 10.1016/j.ymssp.2020.107398.