Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

A Noble Methodology for Designing Expert System on Cloud Environment: A Case Study of Medical Diagnostic System

Mukti Priya[1], Pankaj Kr. Goswami[2], Kuntal Mukherjee[3]

[1]Research Scholar, Department of CSE, Sarala Birla University, Ranchi, Jharkhand, India.
[2]Associate Professor, Department of Eng./Comp. Sci. & Applied Sci., Sarala Birla University, Ranchi, Jharkhand, India.
[3]Assistant Professor, Department of CSE, Birla Institute of Technology, Ranchi, Jharkhand, India.
Email: muktipriya24@gmail.com

Cardiovascular disorders especially heart diseases become a major problem for global health and continue to place a heavy strain on people and healthcare systems. Early and precise diagnosis is the key to effective therapy and patient care for cardiac diseases. A wide variety of symptoms are found in many medical illnesses, and some disorders have symptoms similar to those of other diseases. Healthcare practitioners can find challenging to diagnose a patient's ailment due to its complexity correctly. This method's use of Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF) has shown promise in enhancing the identification of heart disorders utilizing medical data. A Heart Failure Prediction Dataset was utilized. In the pre-processing step, encode and normalize the data using encoder and Unit Vector Transformation methods. For feature extraction, lower the dimensionality of the transformed data by applying t-distributed Stochastic Neighbor (t-SNE) embedding. To use the Recursive Feature Elimination method to determine the most crucial attributes. Using Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF) for better results in a Python tool improves the diagnosis of heart disease. The proposed method outperformed existing methods using performance indicators such as accuracy, Precision, recall, sensitivity, specificity, and f-measures. This abstract highlight the promising future of this innovation and offers a glimpse into its potential to alter how we identify and treat medical diseases fundamentally.

Keywords: Medical diagnosis, Data Encode, Unit Vector Transformation, t-distributed Stochastic Neighbor embedding (t-SNE), Recursive Feature Elimination, Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF).

1. Introduction

Heart disease remains a major concern for the world's health, affecting individuals and healthcare systems worldwide. Efficiency in diagnosis and prompt care are two of the most important factors in lowering the morbidity and death rates that are caused by cardiac issues [1]. Despite this, medical professionals face a formidable obstacle when identifying cardiac illnesses. This is because cardiac disorders are characterized by highly complicated and overlapping symptomatology, in addition to their similarities to other medical problems. There has been a significant rise in the healthcare business due to the public's increased awareness of the significance of healthcare and the development of technology relevant to healthcare [2]. Despite this, delivering health care that is inexpensive and high-quality standards is challenging. In addition to the leading cause of mortality on a global scale, health problems related to the heart impact of those who are 65 years old or older [3]. The enormous rise, complexity, and high expenditures of chronic illnesses, which hurt society, are a weight on the worldwide population, both financially and physically. This burden is a result of the prevalence of chronic diseases. Modern technology has made it possible to develop complex Heart Disease Diagnosis Systems, completely changing how cardiovascular healthcare is provided [4]. This diagnostic device accurately evaluates heart health by combining artificial intelligence, machine learning algorithms, and cutting-edge medical imaging methods. These systems offer healthcare practitioners an in-depth understanding of a patient's cardiovascular profile by analyzing a wide range of patient data, such as medical history, genetic predispositions, and physiological factors, in real-time [5]. The system's ability to identify subtle patterns and abnormalities that can challenge conventional diagnostic techniques is improved by incorporating machine learning algorithms into the diagnostic process. These algorithms improve their predicting powers by absorbing large datasets and adjusting to new developments in diagnostic patterns and medical knowledge [6]. In addition to increasing accuracy, this dynamic learning process helps the system predict possible dangers, which allows proactive interventions and individualized treatment plans. A physician's duty is made more difficult because the diagnosis of cardiac disease depends on many variables to assist medical professionals in making judgments and reducing the number of mistakes that occur during diagnosis [7]. Classification systems make it possible for medical professionals to review medical data in a significant amount of detail. The implementation of these systems involves the creation of a model that is capable of classifying existing records by using data from samples [8]. To assist medical professionals in diagnosing individuals suffering from heart disease, several classification algorithms have been created and set to use as classifiers. Therefore, it is of the utmost importance to determine the underlying factors that account for cardiac problems [9]. Consequently, it is the utmost importance to discover the elements responsible for heart disorders and to build a system capable of diagnosing heart ailments.

1.1 Contribution of the study:

- Many medical conditions have a broad range of symptoms, and several diseases have symptoms similar to those of other disorders.
- Due to its intricacy, a patient's condition can be difficult for medical professionals to identify accurately.

- This method's use of Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF) has improved the diagnosing of cardiac problems using medical data.

The rest of the paper is divided into sections. The objective-based relevant studies are shown in section 2. The data collection and their proposed methods process are shown in section 3. The performance analysis and their discussion are conducted in sections 4 and 5. The paper concluded in section 6.

2. Related work:

Study [10] focused on many ANN-based automated decision support systems for cardiac illness diagnosis was suggested. The majority of these methods, however, feature pre-processing exclusive. Feature refinement and addressing the issues with the predictive model, specifically under fitting and overfitting, are the main topics of the study. Study [11] examined the clinical practice, the definition is based on integrated and iterative clinical evaluations that seek patients whose condition is deteriorating and who rely on certain treatments. Methods for risk classification, prognosis, and approaches to immediate and long- term care are discussed, which looks at the accepted definitions. Study [12] suggested efficiently anticipating, diagnosing, and treating various cardiac ailments; it gives a comprehensive overview of several machine learning algorithms and evaluates their performances. Article [13] suggested an innovative diagnostic method is presented in this publication. For feature selection, the proposed system uses a random forest model in conjunction with a random search algorithm (RSA) to diagnose heart failure. The proposed diagnostic system has fine-tuned using the grid search method. Study [14] suggested an Internet of Things (IoT) infrastructure with several data sources combined with computer- aided diagnostics, we can locate and monitor people who have heart failure is used to improve diagnosis in the planned healthcare system. Study [15] utilized the three top- performing algorithms of the Support vector machine (three kinds of SVM) in the research. These techniques were improved by data pre-treatment and normalization. Two evolutionary algorithms and particle swarm optimization with stratified 10-fold cross-validation were used to optimize classifier parameters and parallel select features. Study [16] examined the ideal machine learning model with rich heart disease healthcare data that can accomplish this aim. Machine learning-based heart disease prediction and diagnosis systems have been introduced lately. These systems cannot handle high-dimensional information because they lack a sophisticated framework for heart disease prediction using several data sources. Study [17] examined one approach that integrates genetic algorithms with a fuzzy logic model to aid medical professionals in the early detection of cardiac illness via prediction. Two modules comprise the model; one uses fuzzy rules for classification, while the other uses rough sets for selecting features related to heart disease. Article [18] suggested the pixels around it are faint; the method strongly influences determining variance in medical imaging. The machine learning algorithm's pooled area building uses clogged and plaque-filled blood arteries to bag narrowing veins and tissues.

2.1 Problem statement

Cardiovascular disorders, especially heart diseases, are a major global health concern. Accurate and early diagnosis is essential for effective therapy and patient care. Heart diseases

manifest a variety of symptoms, some of which overlap with symptoms of other medical conditions, making accurate diagnosis challenging. In heart problem diagnostics, our suggested CQR-DRF represents a significant step forward. Using dynamic characteristics and quantile regression improves the accuracy of diagnosing a wide range of symptoms, therefore eliminating the difficulties traditionally associated with overlapping presentations. In the battle against cardiac illnesses worldwide, this invention holds the potential for early and accurate identification, which will pave for more effective treatment and enhanced patient care.

3. Methods:

The present approach is developed to identify the categorization of diagnoses related to heart disease. This study aims to create a prediction mode for heart disease using the gathered data, CQR-DRF, and other beneficial processes. Predicting heart illness and providing helpful recommendations for a healthy heart are the primary goals of the model. Figure 1 shows the flow of the process to classify the heart disease diagnosis.

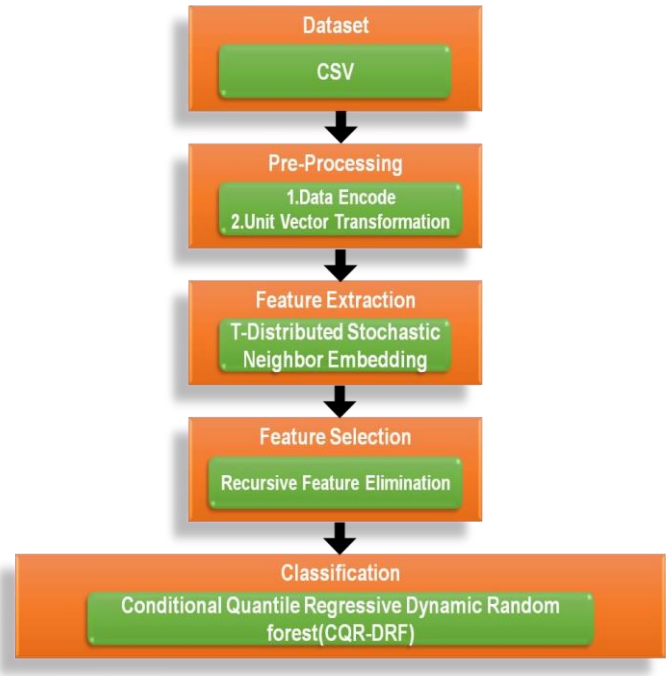

Figure1 Methods for Heart Disease Diagnosis System (source: Author)

3.1 Dataset

With an estimated 17.9 million fatalities annually, the greatest cause of mortality globally is cardiovascular diseases (CVDs), which account for 31% of all fatalities. Four out of five cardiovascular disease deaths occur as a result of heart attacks and strokes, with one-third of these occurrences involving people younger than 70 years old. This dataset includes 11 variables that can be used to predict the possibility of a heart illness. Heart failure is a frequent occurrence by CVDs. The dataset was collected from the

Kaggle(Source:https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction)

3.2 Data pre-processing using Data Encode and Unit Vector Transformation

Data encoding and unit vector transformation are two methods used in data pre-processing techniques to modify the initial information for better representation and analysis.

3.2.1 Data Encode

Data encoding is a pre-processing method that changes the data's format or structure to prepare data for analysis, storage, or transmission. Data compression for storage space reduction, encoding text into numerical representations (e.g., using word embeddings or one- hot encoding), and changing categorical variables into numerical values are examples of encoding methods.

3.2.1.1 One-hot encoding

One instance of an encoding technique is known as one-hot encoding. The initial feature vector is extended to a multidimensional matrix in one-hot encoding. The dimension of the matrix is equal to the number of states included inside this feature, and each size represents a different state. When this processing is performed, it results in the assertion of one dimension of the feature matrix for a certain condition, which is typically '1'. All other sizes of the state are considered to be zero.

3.2.2Unit Vector Transformation

This method uses the Pythagorean Theorem (${u_w}^2 + {u_z}^2 = u^2$)to find the vector's magnitude or hypotenuse(equation (1)).

$$\left|\left|\vec{u}\right|\right| = \sqrt{u_w * u_w + u_z * u_z} \qquad (1)$$

Normalize the vector by scaling it down to 1 by dividing each component to the vector's magnitude. A vector containing 10 elements would have a value of 1 when split by 10. Dividing all other members by the same number of 10 is necessary to scale down to vector size 1. This approach can alter the vector's length while keeping its direction unchanged. A new variable w' with a range of [0, 1] can be created while executing unit vector transformations in equation (2).

$$w' = \frac{w}{||w||} \qquad (2)$$

3.3 Feature extraction using T-Distribute Stochastic Neighbour Embedding (T-SNE)

Data visualization and dimensionality reduction are two areas where the t-distributed stochastic neighbour embedding approach and principal component analysis (PCA) are comparable algorithms. Apart from the approach, t-SNE differs primarily in its capacity to describe nonlinear connections in the data and in its ability to translate the original data's high-dimensional structure into a lower-dimensional space while preserving it. By calculating the distance between each point in the initial dataset, the t-SNE method generates pair wise similarities using a Gaussian kernel. Afterward, the process creates probability distributions across sets of points, where the similarity is proportional to the pairwise similarity. A lower-dimensional space maps the resultant chosen items to a comparable probability distribution.

The technique aims to discover a lower-dimensional representation while preserving the original data structure by minimizing the difference between the two distributions chosen in equation (3).

$$o_{ji} = \frac{o_{i|j} + o_{j|i}}{2m} \quad (3)$$

In the equations, $oi|j$ and $oj|i$ show the inverse conditional probability of a point j given point i and the whole count of data elements, denoted as n. Notably, for t-SNE, oii and ojj are both zero, and oji is equivalent to oji. The computation of pair-wise similarities within the original space is performed using equation (4).

$$r_{ji} = \frac{\left(1+\left||z_j - z_j|\right|^2\right)^{-1}}{\sum_{l \neq 1} \left(1+||z_l - z_k||^2\right)^{-1}} r (4)$$

It is used to generate potential map $z = z_1, z_2, \ldots, z_n$in the space with fewer dimensions. At first, these candidates are chosen randomly, usually from a small-variance, origin-cantered Gaussian distribution. By minimizing the Kullback-Leibler divergence, which is given by, t-SNE finds the optimum mapping relations in equation (5)

$$LK(O||R) = \sum_{ji} \; o_{ij} \; log \; log \left(\frac{o_{ji}}{r_{ji}}\right) (5)$$

3.4Feature selection using Recursive Feature Elimination (RFE)

Numerous attributes can hold greater significance for the analysis. There is a correlation between the existence of redundant parts and the effectiveness of the classification algorithm. Consequently, the forecast can get partial as a consequence. Recursive Feature Elimination, known as RFE, is a technique used to identify the characteristics most responsible for the predicted predictions. The dimensionality of the dataset is reduced because of this, and the effective features created from the restricted subsets of features are preserved. The deletion of less significant characteristics is accomplished via a method based on recursive feature elimination (RFE) and carried out in a step-by-step manner. The use of RFE is advantageous in the process of improving RF models for breach detection. Choosing features from a training dataset is made easier using this feature selection approach, which is efficient. It is possible to generate an input feature set without compromising classification accuracy by removing characteristics that are not relevant to the job. It is an iterative strategy that is used until a certain amount of features are discovered, and the term recursive refers to this approach. Prior to generating a feature ranking (from high to low), it first determines the significance of each feature by analyzing its contribution to the classification process using an RF classifier. The model is re-trained with the amended features, and the classification performance is achieved with the new feature set. After this, the traits considered to be the least importance are eliminated. This is an iterative process until the feature set isdevoid of any elements.

3.5Classification method for heart disease diagnosis system using Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF)

The system employs advanced techniques such as Conditional Quantile Regressive Dynamic Random Forest (CQR-DRF) to diagnose heart disease. To improve classification accuracy,

this method blends DRF with CQR. The system uses quantile regression for detailed risk assessment and dynamic modelling for adaptability, guaranteeing accurate and reliable heart disease classification.

3.5.1 Dynamic Random Forest (DRF)

An assembly of random trees is produced using a classification method known as the Dynamic random forest (DRF) algorithm. The algorithm for creating classification and regression trees (CART) is the decision tree approach that represents the norm. A CART tree comprises several nodes, including central, branching, terminal, and connecting edges. The most crucial processes are selecting the most appropriate variables to serve as nodes and determining the most proper points of division between those vertices. This guarantees that the progeny nodes will be more precise than their parents. The CART technique uses the Gini index to determine the degree of impureness associated with each node. Provided that it has a node s as well as the anticipated probability of each class, an expression for the Gini index at node$s$ is $p(c|t)(t = 1, \dots T)$ below in equation (6-8)

$$G(t) = \sum_{s_{1+c_2}} p(c_1|t)p(t) = 1 - \sum_{c=1}^{c} p^2(c.t) \quad (6)$$

Let s serve as the node's point of branching t, resulting in the partitioning of the node into two halves, with a percentage assigned to each, p_R , tis equal to the number of samples that match to t_R, and a proportion, p_L , is set to t_L , i.e., 1 .This means that the decrease in the impurity Gini score is becausep_R $p_L$as described here

$$\Delta G(s,t) = G(t) - p_R G(t_R) - p_L G(t_L) \quad (7)$$

The ideal variation t^* and the ideal dividing point i^* are determined to have the effect of reducing the Gini impurity the most.

$$s^*, j^* = arg\ argmax s, j \Delta G(s,t) \quad (8)$$

The CART algorithm makes repeated calls to the above function to produce a tree. Bagging and random subspace theories are combined in an ensemble model called an arbitrary decision forest derived from the CART method. To be more explicit, the CART method educates a collection of decision trees using each non-leaf node containing a bootstrap sample and a set of randomly selected variables. Every one of the forest's trees has the potential to develop its maximum height to an endless level as long as all of the leaf nodes remain unadulterated. DRF algorithm is shown in Algorithm 1.

Algorithm 1: Dynamic Random Forest (DRF)

1: $procedure RANDOM FOREST$

2: for 1 $to Tdo$

3: $Drawn points D, with\ replacement from D$

4: $Build a full decision/regression tree on D$

BUT: $each\ split only considers k features, picked uniformly at random$

$new features for every split$

5: $Prunetreetominimizeout - of - bagerror$

6: $endfor$

7: $AverageallTtrees$

8: $endprocedure$

3.5.2 Conditional Quantile Regression (CQR)

The Conditional Quantile regression (CQR) statistical technique is based on the conventional conditional mean regression. Estimating the conditional median of the target variable can be accomplished by reducing the mean error to the greatest extent possible. Similarly, when it comes to CQR, minimizing the CQR error might lead to the computation of the conditional quantiles.Assuming$Z$ is the variable of random response $w_l(l = 1,2, \dots, n)$ and a collection of arbitrary explanatory variables. The cumulative distribution function with conditions$E(z)$, can be expressed as in equation (9)

$E(z) = O(Z \leq z|W_1 = w_1, \dots, w_l = w_l)$(9)

Regarding a cumulative distribution function, the α –quantile r_α can be defined as in equation (10)

$r_\alpha = \{z \,|E(z) \geq \alpha\}$ (10)

In the traditional QR, □-quantile r_α creates an equation for linear regression by substituting in equation (11-12)

$r_\alpha(w_1, \dots, w_l) = \beta^S w$ (11)

Where,

$\beta = [\beta_1, \beta_1, \dots, \beta_n]^S, w = [w_1, w_2, \dots w_k]^S$(12)

To get the quantiles, have to estimate the β. Consequently, the quantile regression error function is represented as equation (13-14)

$F\{K_\alpha(z, r_\alpha \mid W_1 = w_1, \dots, W_l = w_l)\}$ (13)

whereK_α is a test function, as,

$K_\alpha(z, r_\alpha) = \{\alpha(z - r_\alpha)\ z - r_\alpha > 0, (\alpha - 1)(z - r_\alpha)\ z - r_\alpha \leq 0$ (14)

The next step is to construct confidence intervals using the CQR findings. As an example, the intervals for 90% predictions can be represented by equation (15)

$OJ(w) = [r_{0.005}(w), r_{0.95}(w)]$ (15)

4. Results

A 2.33 GHz CPU, 4 GB of RAM, and Windows 8.1 were used throughout the experiment. Python was utilized throughout the testing process. The data from CQR-DRF approaches based on heart disease prediction was provided in this research. The ability of a model to anticipate the outcomes of a classification strategy for the diagnosis of heart disease is

measured by statistics such as accuracy, specificity, sensitivity, Precision, f1-score, and recall.Support Vector Machine (SVM) [19], Linear Regression (LR) [19], and Bidirectional Long Short-Term Memory (Bi-LSTM) [20] all fall short of the performance of our suggested technique, CQR-DRF.

4.1 Accuracy

The term accuracy is used in the context of heart disease prediction refer to the degree to which the predictive model is accurate in identifying people as having or not having heart disease. The capacity of the model to produce accurate predictions in both positive and negative scenarios that the metric of a measure. Both Table 1 and Figure 2 illustrate the accuracy of the heart disease classification.

Table1. Comparison of Accuracy (Source: Author)

Methods	Accuracy (%)
SVM [19]	85
LR [19]	88
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

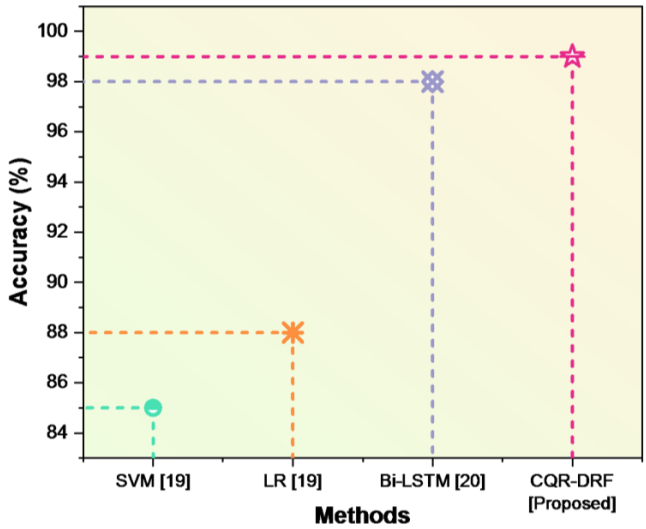

Figure2. Accuracy (Source: Author)

The CQR-DRF approach that was presented exhibits a better accuracy of 99%, exceeding SVM [19] with (85%), LR [19] with (88%), and even surpassing the accuracy reached by Bi-LSTM [20] with (98%).

4.2 Precision

Precision measures the accuracy of a predictive model's favorable heart disease predictions. It is the ratio of real positive predictions to true and false positives. High Precision means the model makes reliable positive predictions with few false positives. False positives in heart disease prediction can cause, testing, and treatments for healthy people. Table 2 and Figure 3 show the precision of heart disease classification.

Table 2. Comparison of Precision (Source: Author)

Methods	Precision (%)
SVM [19]	81
LR [19]	87
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

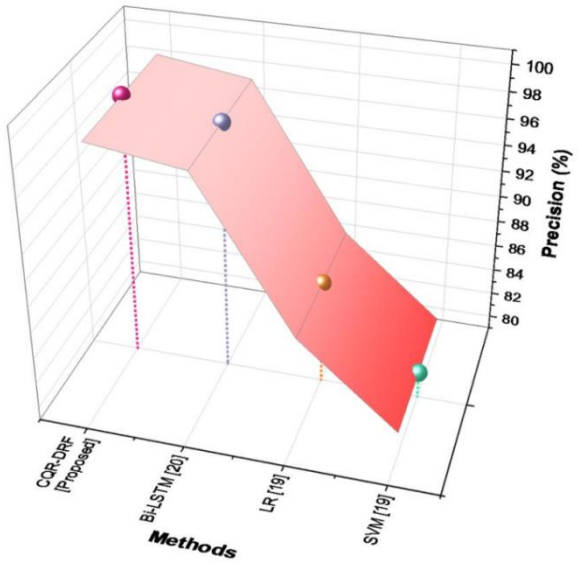

Figure3. Precision (Source: Author)

Based on the provided data, the CQR-DRF method achieved the highest precision rate of 99%, outperforming SVM (81%), LR (87%), and Bi-LSTM (98%) in that specific precision rate.

4.3 Recall

In classification tasks like heart disease prediction, recall (sensitivity or true positive rate) is utilized to assess a predictive model and the ratio of real positive predictions to positive incidents. Missing a positive case (a cardiac patient) is more important than misclassifying a negative one. High recall indicates the model catches most positives and lowers false negatives. The Recall of heart disease classification is shown in Table 3 and Figure 4.

Table 3. Comparison of Recall (Source: Author)

Methods	Recall (%)
SVM [19]	86
LR [19]	94
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

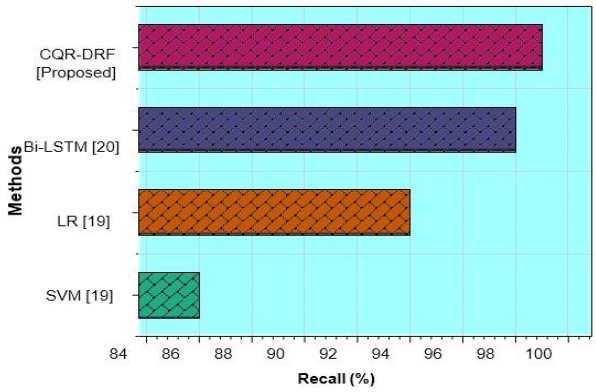

Figure 4. Recall (Source: Author)

A recall rate of 99% was demonstrated using the suggested CQR-DRF strategy, in accordance with the methodologies and recall percentages provided. When it came to performance, our technique beat SVM (86%), LR (94%), and even Bi-LSTM (98%).

4.4 Specificity

Specificity in heart disease prediction means a model can reliably identify non-patients. If the model detects people without heart disease, it avoids false positives. In heart disease prediction, a high specificity means the model can identify those who are not at risk, reducing the need for unneeded interventions or treatments. Table 4 and Figure 5 show the specificity of heart disease classification.

Table 4. Comparison of Specificity (Source: Author)

Methods	Specificity (%)
SVM [19]	75
LR [19]	81
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

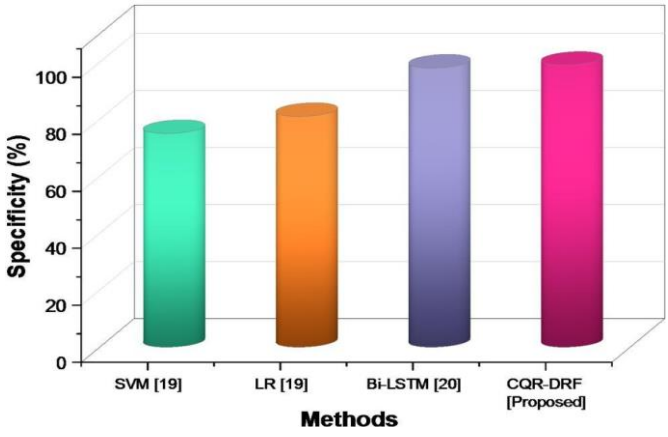

Figure5. Specificity (Source: Author)

In percentage terms, the CQR-DRF technique outperformed SVM (75%), LR (81%), and even Bi-LSTM (98%) with a better specificity of 99% during the comparative analysis.

4.5 Sensitivity

Heart disease prognosis depends on sensitivity, True Positive Rate, or Recall. It shows how many real heart disease patients a prediction model or diagnostic test identified. Sensitivity assesses a model's ability to identify heart disease patients among all individuals having it. Heart disease prediction requires a high-sensitivity model to identify the most positive instances and minimize false negatives. Table 5 and Figure 6 illustrate the sensitivity of heart disease classification.

Table 5. Comparison of Sensitivity (Source: Author)

Methods	Sensitivity (%)
SVM [19]	85
LR [19]	94
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

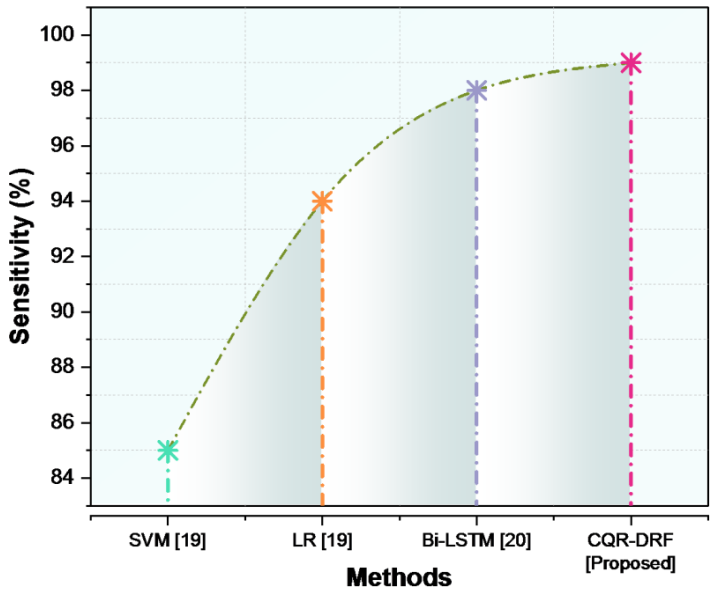

Figure6. Sensitivity (Source: Author)

The proposed CQR-DRF strategy outperforms the SVM (85%), LR (94%), and Bi-LSTM (98%) approaches in terms of performance, displaying an outstanding sensitivity of 99%.

4.6 F1-Score

The F1-score is employed in binary classification issues like heart disease prediction. The F1 score balances accuracy and recall. It scores 0 to 1, with 1 is the best. A model with a higher F1 score better recognizes positive instances and avoids false positives. A high F1-score indicates a model's ability to identify heart disease patients while preventing false positives and negatives. Heart disease classification F1-scores are shown in Figure 7 and Table 6.

Table 6. Comparison of F1-Score (Source: Author)

Methods	F1-Score (%)
SVM [19]	82
LR [19]	90
Bi-LSTM [20]	98
CQR-DRF [Proposed]	99

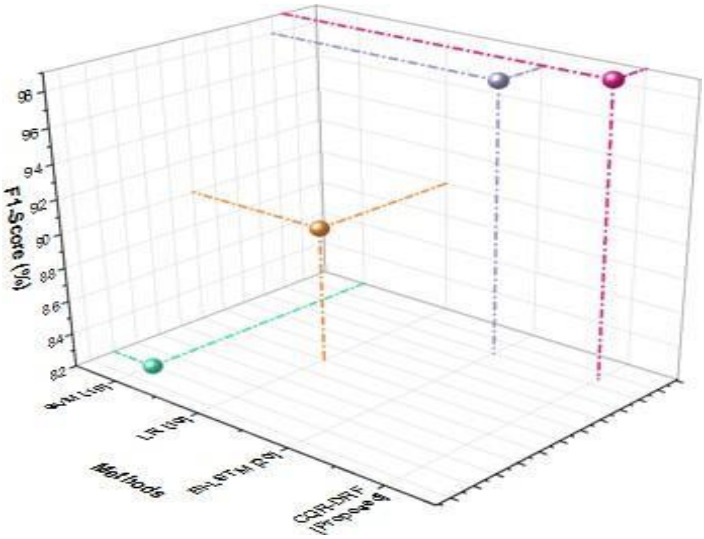

Figure7. F1-Score (Source: Author)

LR achieves an F1-Score of 90%, Bi-LSTM has a high accuracy of 98%, SVM achieves 82%, and the recommended CQR-DRF technique outperforms all of these methods with an outstanding F1-Score of 99%.

5. Discussion

The accuracy of diagnostics for heart disease has been enhanced by applying several machine learning models, which is an important part of healthcare. There are limitations to the current models that need to be considered. These include SVM [19], LR [19], and Bi-LSTM [20]. Although SVM and LR are used, they need help to understand the complex patterns linked to heart disease since they have trouble collecting complex nonlinear correlations in the data. Despite its proficiency with sequential data, the temporal correlations in the physiological measures related to heart health could be difficult for Bi-LSTM to understand. The inability to understand the reasoning behind a certain diagnosis is another issue with these models due to their lack of explanation. Our proposed CQR-DRF model was set onward. However, there are obstacles to implementing innovative models, such as they are interpretable and working with current healthcare systems. Therefore, even if CQR-DRF's appearance indicates declaration, it is imperative to examine practical implementation issues and continuously refine CQR-DRF to ensure its efficacy in real-world scenarios, including the detection of heart disease.

6. Conclusion

The utilization of CQR-DRF in the classification approach for the detection of heart disease exhibits encouraging outcomes, as stated in the conclusion. This innovative method enhances prediction accuracy by providing a robust framework for the rapid and dependable detection of heart problems. Thishelps more effective healthcare decision-making, ultimately benefiting patients via better outcomes. The proposed system outperforms prior models that are considered state-of-the-art in the field of heart disease prediction with accuracy, Precision, recall, sensitivity, specificity, and F1-score of 99%, 99%, 99%, 99%, 99%, and 99%, respectively. The incorporation of edge computing in future, which allows for urgent analysis to be performed at the edge levels, can help address the inherent limits of the Cloud, such as increased latency and bandwidth usage, and the exponential growth of IoT data into this investigation. With fog/edge computing, healthcare organizations can improve their efficiency illness prediction, faster reaction times, and more adaptable medical decision- making, and the entire quality of service will be enhanced.

References

1. Sleeman KE, De Brito M, Etkind S, et al. The escalating global burden of serious health-related suffering: projections to 2060 by world regions, age groups, and health conditions. The Lancet Global Health, 2019; 7(7),883-892, 10.1016/S2214-109X(19)30172-X
2. Wang Q, Su M, Zhang M et al. Integrating digital technologies and public health to fight Covid-19 pandemic: key technologies, applications, challenges and outlook of digital healthcare. International Journal of Environmental Research and Public Health, 2021; 18(11),6053,10.3390/ijerph18116053
3. Usak M, Kubiatko M, Shabbir MS, et al. Health care service delivery based on the Internet of things: A systematic and comprehensive study. International Journal of Communication Systems, 2020; 33(2), 4179, 10.1002/dac.4179.
4. Tuli S, Basumatary N, Gill SS, Kahani et al. An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments. Future Generation Computer Systems, 2020;104,187-200, 10.1016/j.future.2019.10.043
5. Zhou J, Du M, Chang S et al. Artificial intelligence in echocardiography: detection, functional evaluation, and disease diagnosis. Cardiovascular ultrasound, 2021;19(1), 1-11,
6. Gill AY, Saeed A, Rasool S, et al. Revolutionizing Healthcare: How Machine Learning is Transforming Patient Diagnoses-a Comprehensive Review of AI's Impact on Medical Diagnosis. Journal of World Science, 2023; 2(10), 1638-1652, 10.58344/jws.v2i10.449.
7. Lysaght T, Lim HY, Xafis V et al. AI-assisted decision- making in healthcare: the application of an ethics framework for big data in health and research. Asian Bioethics Review, 2019; 11,299-314,.
8. Sulaiman MA. Evaluating data mining classification methods performance in internet of things applications. Journal of Soft Computing and Data Mining, 2020; 1(2),11-25.
9. Alzaharani ME and Hannan SA. Diagnosis and Medical Prescription of Heart Disease Using FFBP, SVM and RBF. 2019; 5(1),6-15.
10. Ali L,Rahman A, Khan A et al. An automated diagnostic system for heart disease prediction based on ${\chi^{2}} $ statistical model and optimally configured deep neural network. Ieee Access, 2019; 7,34938-34945,10.1109/ACCESS.2019.2904800.
11. Truby LK and Rogers JG. Advanced heart failure: epidemiology, diagnosis, and therapeutic

approaches. Heart failure, 2020; 8 (7) 523-536.10.1016/j.jchf.2020.01.014
12. Riyaz L, Butt MA, Zaman Met al. Heart disease prediction using machine learning techniques: a quantitative review. In International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Springer Singapore, 2022; 3, 81-94. 10.1007/978-981-16-3071-2_8
13. Javeed A, Zhou S, Yongjian L, et al. An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection. IEEE access, 2019; 7180235- 180243.10.1109/ACCESS.2019.2952107
14. Abdel-Basset M, Gamal A, Manogaran G et al. A novel group decision-making model based on neutrosophic sets for heart disease diagnosis. Multimedia Tools and Applications, 2020; 79, 9977-10002.10.1007/s11042-019-07742-7
15. Abdar M, Książek W, Acharya UR, et al. A new machine learning technique for an accurate diagnosis of coronary artery disease. Computer methods and programs in biomedicine, 2019; 179, 104992.10.1016/j.cmpb.2019.104992
16. Ali F, El-Sappagh S, Islam SRet al. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 2020; 63208-222.10.1016/j.inffus.2020.06.008
17. Reddy GT, Reddy MPK, Lakshmanna K, et al. Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis. Evolutionary Intelligence, 2020; 13,185-196.10.1007/s12065-019-00327-1
18. Chen DJIZ and Hengjinda P. Early prediction of coronary artery disease (CAD) by machine learning method-a comparative study. Journal of Artificial Intelligence and Capsule Networks, 2021; 3(1) 17-33.10.36548/jaicn.2021.1.002
19. Mansur Huang NS, Ibrahim Z and Mat Diah N. Machine learning techniques for early heart failure prediction. Malaysian Journal of Computing (MJoC), 2021; 6(2) 872-884.10.1109/MAJICC56935.2022.9994093
20. Nancy AA, Ravindran D, Raj Vincent PDet al. Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. Electronics, 2022; 11 (15) 2292. 10.3390/electronics11152292