Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 26, Issues 5, pp. 2105–2117, DOI 10.12694/scpe.v26i5.4715

REVOLUTIONIZING CARDIAC PREDICTION BASED ON FOG-CLOUD-IOT INTEGRATED HEART DISEASE MODEL

AMIT KUMAR CHANDANAN,* MEENA RANI,† KIRAN SREE POKKULURI,‡ SUMAN SINGH,§ VAIBHAV JAISWAL,¶ POTU NARAYANA,‖ AND VANDANA ROY**

Abstract. In a time when technology is having a profound effect on medical applications, the rapid remote diagnosis of any cardiac disease has proven to be a formidable obstacle. These days, computers can swiftly process a large volume of patient medical records. Recent developments in the IoT and medical applications, such as the IoMT, have opened the possibility of data diffusion among numerous locations pertaining to patients. This study presents the IHDPM, an integrated model for the prediction of cardiac disease that takes into account dimensionality declining through PCA (principal component analysis), feature collection over sequential feature selection (SFS), and classifications through the random forest (RF) classifier. The proposed model outperforms over different unadventurous classification methods, including LR (logistic regression), NB (naive Bayes), SVM (support vector machine), KNN (K-nearest neighbors), DT (decision trees), and RF, according to experiments conducted using the CHDD (Cleveland Heart Disease Dataset) as of the UCI-ML source and the Python programming linguistic. Medical professionals may find the proposed model useful for making accurate diagnoses of cardiac patients. While DL approaches may produce more accurate prediction results, it would be supplementary operative to reduce the extents count before cluster generation to improve the results.

Key words: Internet of Medical Things; Support Vector Machine; K-nearest neighbors; random forest.

1. Introduction. Many people consider the Internet, which connected several networks and sparked the first digital revolution, to be the greatest innovation that has ever been. The second digital revolt, the IoT (Internet of Things), is happening right now, and it's crucial for long-distance communication. The proliferation of new technologies has led to an explosion in the number of IoT applications [1]. This will lead to an explosion in data production due to the proliferation of connected devices used for a wide range of tasks. One revolutionary network that provides a dispersed health care scheme that may treat any sickness in every home is the IoMT (Internet of Medical Things). Health providers may now keep tabs on their patients from afar thanks to Internet of Things (IoT) submissions in e-Healthcare schemes, and patients have easy access to these same services [2]. Globally, 17.9 million persons miss their survives each time to heart disease and cardiovascular illness, which is a major reason for concern rendering to the WHO (World Health Organization).

A rise in deaths caused by cardiovascular disease is related with an upsurge in mortality and the danger of numerous illnesses, rendering to the WHO. Figure 1.1 presents the structural design of the healthcare observing system. Risk factors for cardiovascular disease include unhealthy diet and obesity, insufficient physical activity, advancing age, substance abuse, and excessive alcohol consumption [3]. A number of conditions, including diabetes, high blood pressure, hypertension, and hyperglycemia, are considered to be risk factors for cardiovascular disease. This makes it difficult for medical doctors to arrive at a conclusive diagnosis of heart disease because the symptoms of the condition vary widely from one individual to the next and from one age

*Department of Computer Science and Engineering, Guru Ghasidas Vishwavidyalaya (A Central University), Bilaspur, (C G), India (chandanan.amit@ggu.ac.in).
†Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India (meena.rani@chitkara.edu.in).
‡Department of Computer Science and Engineering, Shri Vishnu Engineering College for Women, Bhimavaram, India (drkiransree@gmail.com)
§Rani Devi Durgavati Vishwavidyalaya, Jabalpur, MP, India (suman_singh2008@yahoo.com).
¶IES Institute of Pharmacy, IES University, Bhopal, Madhya Pradesh 462044, India (vaibhav.research@iesuniversity.ac.in).
‖Dept. of CSE, Stanley College of Engineering and Technology for Women, Hyderabad, India (potunarayana@gmail.com).
**DoEC, Gyan Ganga Institute of Technology and Sciences, Jabalpur, M.P., India (vandana.roy20@gmail.com).

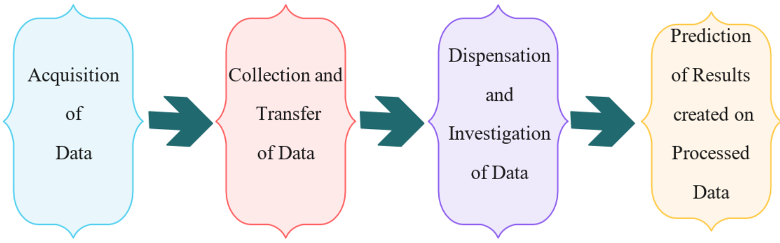

Fig. 1.1: Healthcare observing system structural design

group to another. Around two percent of adults and eight times more people aged 75 and older are affected by heart failure (HF), which is a prevalent and fatal ailment. This is similar to the situation in developed nations. There is a range of three to five percent of hospitalizations that are associated with heart failure. There is a possibility that heart disease may be the cause of severe mortality in the fifty to ten years to come [4][5]. The number of deaths is expected to surpass 24 million annually by the year 2030, according to the predictions. The authors of the study came to the conclusion that artificial intelligence classification methods, such as machine learning (ML) and data mining (DM), ought to be utilized in order to get accurate scientific outcomes in individuals who have been recognized or shown to have a sickness [6][7]. Machine learning, on the other hand, focuses on training and improving systems autonomously from previous experiences, as well as on creating computer programs that can learn from and apply these experiences. This is in contrast to data mining, which encompasses a variety of approaches for uncovering new information in vast datasets without taking into account any prior knowledge. In addition, the development of new computer technology helps in the process of diagnosis by providing feedback to medical professionals [8]. In the field of medicine, ML is utilized extensively for the purpose of diagnosis. The diagnosis process is simplified as a result, which in turn makes it simpler to forecast epidemics and evaluate genomic data.

2. The Existing Work Done. Fog computing is quite useful for health care submissions that operate in real-time. Typically, sensors count are applied to gather health information from patients. This information is then directed to a entry device, which can be a mobile phone, tablet, laptop, or desktop computer with specific application software installed [9]. The transmission can be done manually or wirelessly (via Bluetooth). After then, it's possible to send the data to a central workstation or PC for analysis. When it's not there, a network of inexpensive computers called Fog nodes—typically Raspberry Pi and Arduino—step in [10].

E-Healthcare systems have benefited from fog computing since they require quick and reliable results. Electronic health records can be more quickly updated using this method. Examining delicate statistics at the entry devices can improve privacy associated to cloud manner by reducing data propagation to the network [11]. By storing and processing data locally, the need to connect to data centers is removed. It reduces the strain on the network. A scalable design makes it possible to grow. Consequently, e-Healthcare systems are impacted by Fog Computing. From 2012 forward, publications began discussing fog computing and its impact on the Internet of Things (IoT), and the lists have kept growing without ever reaching saturation. Infrastructure that is enabled by fog is covered in half of these articles [12][13]. For the most part, these gadgets came from healthcare and intelligent environments. Effective and value-added therapy can be providing to responsible affected role through the healthcare system's facilitation of regular patient-consultant communication. When it comes to healthcare and the Internet of Things, fog computing is king [14].

The autonomous rehabilitation system ADM was built around ontologies, according to the researchers. This rehabilitation method has alleviated a lot of concerns for the elderly [15]. It is imperative that healthcare facilities utilizing Fog take security precautions seriously. The global health awareness system was investigated by the authors using a multitier Cloud system. It would make data exchange easier and help find outbreaks sooner. The health awareness system's data helps in the battle against illness. Important messages and data pertaining to world health must be sent via the WBAN (Wireless Body Area Network). To handle and distribute data, the suggested design makes use of MapReduce [16]. Using Fog Computing to attach the operator to the

Table 2.1: Comparison of the various autonomous rehabilitation, global health awareness, and health monitoring systems

Study/System	**Description**	**Technologies/ Methods Used**	**Benefits**	**Challenges**	**Evaluation Metrics**
Autonomous Rehabilitation System (ADM) [23]	Built around ontologies to aid elderly rehabilitation.	Ontologies	Alleviates elderly concerns	Security precautions necessary in healthcare facilities using Fog	N/A
Global Health Awareness System [24]	Investigated by means of a multitier Cloud scheme for easier data exchange and outbreak detection.	Cloud system, WBAN, MapReduce, Fog Computing, CASB	Facilitates data exchange and early outbreak detection, robust data security	Ensuring secure and stable data transmission	N/A
Real-Time IoT Monitoring [25]	Tested IoT devices for real-time vitals tracking during exercise.	IoT devices, Bayesian Belief Networks (BBN)	Monitors health during exercise, assesses severity of health issues	Significant risks if not performed correctly, IoT security issues	N/A
Low-Cost Health Monitoring System [26]	Developed using a Fog coat and energy-effective sensor bulges.	Fog Computing, nRF sensor nodes	Low power consumption, quick decision-making	Inherent IoT security risks, unpredictability	N/A
Enhanced Healthcare IoT and Cloud Services [27]	Aimed at reducing network latency using fuzzy and reinforcement learning.	Fuzzy and reinforcement learning, IoT batch workloads	Reduces delays, real-time decision making	Limited-service capacity, high-capacity network utilization	N/A
ML Algorithms on CHDD Dataset [28]	Employed ML algorithms to evaluate heart disease prediction.	LR, SVM-linear, NNs (RBNF, GRNN, SLP)	Precise results with LR and SVM-linear	Complexity of more composite models	Accuracy, F1 score, ROC curves
EML Technique on CHDD Dataset [29]	Applied EML technique to heart disease prediction.	Extreme Learning Machine (EML)	Suitable for big data, reduced learning time	Limited to CHDD dataset	Accuracy (80%)
DT vs. NB Classifier on CHDD Dataset [30]	Compared performance of DT and NB classifiers on heart disease prediction.	Decision Tree (DT), Naive Bayes (NB)	DT performed better than NB	Limited comparison scope	Accuracy

Cloud, the authors were able to build a stable Fog system. A three-layer system reduces communication costs. The author's own design, the Cloud access security broker (CASB), uses many security enforcement mechanisms to further protect Health Fog data [17][18][19].

While players were warming up for a game, Bhatia et al. tested how well Internet of Things (IoT) devices could track their vitals in real time. Exercising regularly has several health benefits, including lowering stress levels, but it also has significant dangers if not performed correctly. This study employs Bayesian Belief Networks (BBN) to assess the severity of health issues. Here is a tabular comparison of the various autonomous rehabilitation, global health awareness, and health monitoring systems [20][21].

This table 2.1 provides a concise comparison of the different approaches, highlighting their technologies, benefits, challenges, and evaluation metrics.

After developing an advanced ELM technique with an adaptive boosting algorithm, researchers tested it on four different datasets: CHDD, HHDD, LHDD, and SHDD. They used accurateness as the calculating degree and found that the advanced EML models improved upon previous research by achieving accuracies of 80.14%, 89.12%, 77.78%, and 96.72% for CHDD, HHDD, LHDD, and SHDD, respectively.

In the feature extraction phase, the authors utilized the autoregressive (AR) Burg method. In the classification phase, they tested five different classifiers such as C4.5 DT, KNN, ANN, SVM, and RF using ECG signals from the BIDMC-CHF database, which stands for Beth Israel Deaconess Medical Centre Congestive Heart Failure. The authors asserted that the RF classifier achieved a classification accuracy of 100%. The authors also considered calculating events such as accuracy, sensitivity, specificity, FM, and the ROC curve. The subjects included 15 individuals.

After taking the CHDD dataset's evaluative measure accuracy into account, the authors developed a decision provision scheme that syndicates ANN with Fuzzy_AHP (fuzzy analytic hierarchy process) to improve upon conventional ANN methods. Their findings suggest that the suggested approach could attain a typical estimate accuracy of 91.10%.

3. The motivation for the proposed work. According to a variety of study and online sources, there have been numerous studies on Fog Computing concepts that are based on smart homes, smart cities, etc. One possible reason for conducting new research is that it appears to be relatively low in e-Healthcare systems. According to projections, the numeral of sensor nodes in usage round the world will reach 50 billion in 2025, and 75 billion in 2030. Data generated by these equipments will be enormous and will require constant processing time. With the increasing demand for Internet of Things (IoT) solutions, fog computation has appeared as a viable alternative to cloud computing. Global citizens have grown increasingly worried about the increasing mortality toll from long-term health conditions in recent years. Additionally, it is a social issue to deliver medical benefits to these patients in real-time. Having the capability to remotely and in real-time diagnose patients is crucial.

4. Research Objective.

- Develop an integrated model (IHDPM) for predicting cardiac disease using PCA for dimensionality lessening, consecutive feature assortment, and random forest (RF) classification.
- Evaluate the presentation of the projected prototypical against different conventional classification techniques using the Cleveland Heart Disease Dataset (CHDD) from the UCI-ML repository.
- Demonstrate that the projected model provides accurate diagnoses and could be a valuable tool for medical professionals.

5. The Proposed Work. The number of individuals with a reduced heart rate is growing annually, making heart disease the most dangerous health problem. As a result, more and more patient data is becoming publicly available, which is great news for practitioners because it allows them to draw useful insights using DM approaches and then apply those insights to diagnosis and prediction using ML and classification techniques. The suggested model is based on the idea of using algorithms in a step-by-step manner to achieve better prediction. The primary process of the prediction system is illustrated in Figure 5.1 by the subsequent stages.

Here is a detailed breakdown of how this proposed model operates:

1. The suggested model begins with gathering relevant data. Our team retrieved the CHDD dataset from the UCI-ML cloud storage.
2. In this stage, missing value imputation is applied to pre-process the data.
3. In the third step, the PCA technique is used to reduce dimensionality.
4. The SFS algorithm is used for feature selection in this stage.
5. Split the Featured Dataset in Half: One Half for Training and the Other Half for Testing.
6. The sixth step is to apply the RF classifier algorithm to the classification task.
7. Various evaluation metrics obtained by the suggested model are computed in this step.
8. Involves using the proposed model to forecast the occurrence of heart disease.

5.1. Data Set Information. The UCI-ML repository's CHDD was utilized in this research work. The processed data set, which includes 303 cases with just 14 attributes, has been considered by the greatest number

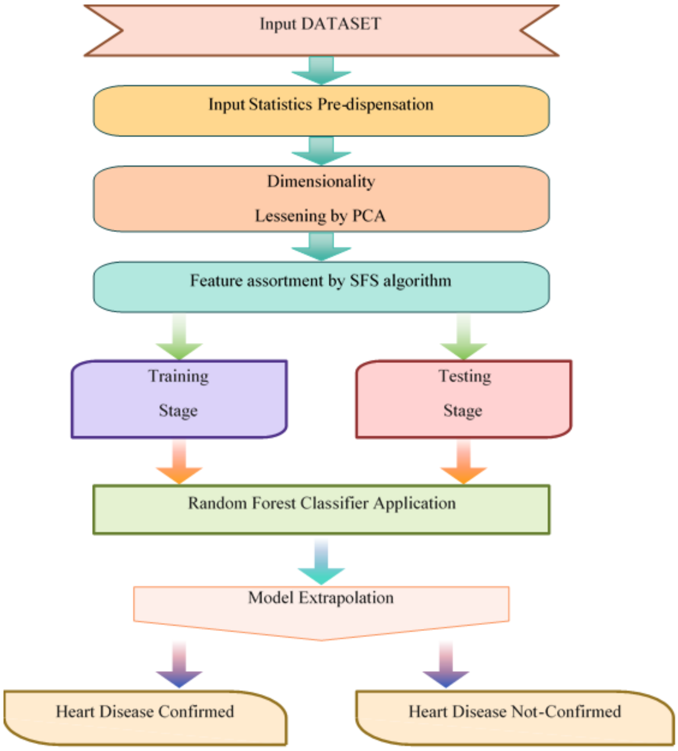

Fig. 5.1: The Proposed system Model

of writers, while this dataset, which has 76 attributes, is far more extensive. Age, sex (female or male), type of chest discomfort, and a brief explanation for each of the thirteen qualities we've chosen are Measurement of blood pressure in millimeters of mercury taken upon admission to the hospital, Test results for electrocardiogram (ECG) and cholesterol (in milligrams per deciliter) Fasting blood sugar levels (FBS) When Thalach reached his maximum heart rate, Gaelic for movement in relation to inactivity, A number of main arteries, the slope of the ST segment, the exercise-encouraged angina measure, and the numeral of main vessels (Ca) Check the thalassemia level and the cardiac disease classification.

Rather of relying on a single classifier, the clinical sector makes heavy use of classification algorithms to sort data into many categories based on specific criteria. Patients' risk of cardiovascular disease has been projected using a diversity of machine learning classification methods.

5.2. PCA (Principal Component Analysis) based Reduction. Taking fewer variables into account is called dimensionality reduction. It can be employed to reduce data while preserving structure or to extract hidden features from unstructured information. You can choose from a number of different dimensionality lessening methods. Principle component analysis (PCA) was investigated as a means of dimensionality reduction in this study. PCA is a method for reducing dataset dimensionality, making data more interpretable, and reducing data loss all at once. PCA is a feature extraction method that precisely integrates our input variables while removing the less important ones. Large datasets are becoming more common, but they are also becoming more difficult to interpret.

Step 1: Data Standardization

1. Center the data:

$$X_{\text{centered}} = X - \bar{x} \tag{5.1}$$

2. Standardize the dataset:

$$X_{\text{std}} = \frac{X_{\text{centered}}}{s} \tag{5.2}$$

Where, s is the standard deviation of X.

Step 2: Compute Covariance Matrix

1. Covariance matrix of the standardized data:

$$C = \frac{TX_{\text{std}}}{X_{\text{std}}} \tag{5.3}$$

Step 3: Eigenvalues and Eigenvectors

1. Solve for the eigenvalues:

$$(C - \lambda I) = 0 \tag{5.4}$$

2. Compute the eigenvectors:

$$(C - \lambda_i I)v_i = 0 \tag{5.5}$$

3. Normalize the eigenvectors:

$$\|v_i\| = 1 \tag{5.6}$$

Step 4: Sort Eigenvalues and Eigenvectors

1. List eigenvalues in descending order:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \tag{5.7}$$

2. Arrange corresponding eigenvectors:

$$V = [v_1, v_2, \ldots, v_p] \tag{5.8}$$

3. Form the matrix of sorted eigenvectors:

$$W = [v_1, v_2, \ldots, v_k] \tag{5.9}$$

Step 5: Select Principal Components

1. Choose the top k eigenvectors:

$$W_k = [v_1, v_2, \ldots, v_k] \tag{5.10}$$

Step 6: Scheme Data onto Principal Components

1. Transform the original data:

$$Z = X_{\text{std}} W_k \tag{5.11}$$

2. Reconstruct the data:

$$X_{\text{reconstructed}} = Z W_k^T + \bar{x} \tag{5.12}$$

Step 7: Compute Explained Variance

1. Total variance:

$$\text{Total Variance} = \lambda_1 + \lambda_2 + \ldots + \lambda_p \tag{5.13}$$

2. Explained variance by k components:

$$\text{Explained Variance} = \lambda_1 + \lambda_2 + \ldots + \lambda_k \tag{5.14}$$

3. Explained Variance Ratio:

$$\text{Explained Variance Ratio} = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_k}{\lambda_1 + \lambda_2 + \ldots + \lambda_p} \tag{5.15}$$

Step 8: Evaluate Performance

1. Calculate reconstruction error:

$$\text{Reconstruction Error} = \|X - X_{\text{reconstructed}}\|_2 \tag{5.16}$$

2. Compute the Frobenius norm:

$$\|X - X_{\text{reconstructed}}\|_F \tag{5.17}$$

Step 9: Scree Plot Creation

1. Plot eigenvalues:

$$y = \{\lambda_1, \lambda_2, \ldots, \lambda_p\} \tag{5.18}$$

2. Set x-axis as component number:

$$x = \{1, 2, \ldots, p\} \tag{5.19}$$

3. Scree plot: Scree Plot: (x, y)

Step 10: Determine Optimal Components

1. Apply Kaiser Criterion: $\lambda_i \geq 1$

Step 11: Validate with Cross-Validation

1. Split data into training and validation sets: $\{\ X_{\text{train}}, X_{\text{val}}\ \}$
2. Fit PCA on training data:

$$W_{\text{train}} = \text{PCA}(X_{\text{train}}) \tag{5.20}$$

3. Validate on test data:

$$X_{\text{val\_transformed}} = X_{\text{val}} W_{\text{train}} \tag{5.21}$$

Step 12: Apply PCA to New Data

1. Standardize new data:

$$X_{\text{new\_std}} = sX_{\text{new}} - \bar{x} \tag{5.22}$$

2. Project onto principal components:

$$Z_{\text{new}} = X_{\text{new\_std}} W_k \tag{5.23}$$

3. Transform back if needed:

$$X_{\text{new\_reconstructed}} = Z W_k^T + \bar{x} \tag{5.24}$$

5.3. SFS (Sequential Feature Selection). The primary goal of implementing FS approaches is to improve the classifiers' accuracy. An assortment of elements from greedy search make up the SFS algorithms. These are essential for reducing a d-dimensional feature intergalactic to a k-dimensional feature subspace, where k is less than or equal to d. These notions are based on the idea of automatically selecting a subset of attributes that are applicable to the task at hand. As an added bonus, FS has a dual purpose: first, to increase compute efficiency by decreasing generalization error; second, to eliminate irrelevant characteristics or noise. When feature selection is integrated, a wrapper concept like SFS becomes even more useful. When it comes to SFS, you can choose between four different types: forward, backward, forward floating, and backward floating.

6. Result Analysis and Discussion. Finding the confusion matrix, a matrix of real class to expected class, is the primary goal of performance evaluation in prediction. This matrix can be used by many evaluative measures to work on. Predictions can be made using a diversity of evaluation metrics, as well as Accuracy, Precision, F-Measure, MCC, False Positive Rate, and False Negative rate.

After completing the necessary pre-processing on the data, the dimensionality reduction technique PCA is applied to lessen dimensions count. Following this, the statistics is processed for feature selection using SFS. To make sense of the distance metric, it will probably minimize the number of dimensions first. Afterwards, the data were divided into two halves, training and testing, at a ratio of 7:3. When that, various evaluation metrics are computed when the suggested model is put into action. Table 6.1 shows the results of comparisons with six different traditional categorization methods using the aforementioned evaluation metrics. If we compare the results with the other traditional ML methods, we can see that the suggested system IHDPM attained the greatest accuracy at 83.42% (Figure 5.1). Figures 6.1, 6.2, and 6.3 indicate that the suggested solution outperforms the alternatives when it comes to F-Measure computation, Precision, and Recall, respectively. Therefore, it's safe to say that clinical doctors are using this recommended approach to make the best decisions when it comes to predicting cardiac problems.

Manuscripts for FoMT take into account both performance and network parameters for potential evaluations, which provide evidence that Fog Computing concepts outperform Cloud Computing concepts. Using a consistent set of performance metrics, performance evaluations aim to establish a real-to-expected class confusion matrix.

Accuracy. When compared to the total number of cases that were investigated, the percentage of true outcomes, which includes both true positives and true negatives.

$$\text{Accuracy} = \frac{TN + TP}{\text{Total Data Sample}} \times 100 \tag{6.1}$$

where

- TP = True Positives,
- TN = True Negatives,
- FP = False Positives,
- FN = False Negatives.

Sensitivity (Recall or True Positive Rate, TPR). The percentage of genuine positives that are appropriately identified as being positive tests.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \tag{6.2}$$

Precision (Positive Predictive Value, PPV). This refers to the percentage of positive outcomes that are actually positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6.3}$$

F-measure (F1 Score). The harmonic mean of an individual's recall and precision.

$$\text{F-measure} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 \tag{6.4}$$

True Positive Rate (TPR). The percentage of genuine positives that are appropriately identified as being positive tests.

$$\text{TPR} = \frac{TP}{TP + FN} \times 100 \tag{6.5}$$

False Positive Rate (FPR). The percentage of false positives that are mistakenly identified as positives.

$$\text{FPR} = \frac{FP}{FP + TN} \times 100 \tag{6.6}$$

Table 6.1: Comparative Study of Different Classifiers with the Proposed Algorithm

Evaluation Parameters	**Support Vector Machine**	**K-Nearest Neighbour**	**Random Forest**	**Proposed Algorithm**
Accuracy (%)	79.06	76.52	81.29	83.56
Precision	0.824	0.782	0.821	0.834
Sensitivity	0.824	0.832	0.887	0.912
F-Measure	0.827	0.794	0.862	0.871
TPR (%)	81.14	85.12	88.75	91.37
TNR (%)	73.72	71.86	74.28	75.83
FPR (%)	26.32	39.84	27.34	28.18
FNR (%)	18.67	16.27	12.52	10.24

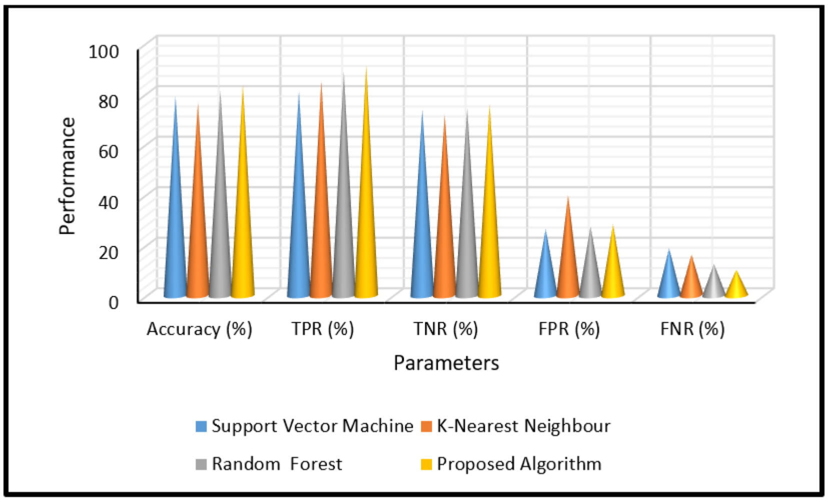

Fig. 6.1: The comparison of the proposed methods based on different evaluation parameters

True Negative Rate (TNR or Specificity). The fraction of actual negatives that are appropriately detected within the data set.

$$\text{TNR} = \frac{TN}{TN + FP} \times 100 \tag{6.7}$$

False Negative Rate (FNR). The percentage of true positives that are mistakenly recognized as negatives in the data.

$$\text{FNR} = \frac{FN}{TP + FN} \times 100 \tag{6.8}$$

In cardiac disease prediction, accuracy evaluates the overall correctness of the model, while sensitivity (TPR) assesses its ability to identify patients with the disease. Precision measures the reliability of positive diagnoses. The F-measure balances precision and recall, provided that a single performance metric. TPR and TNR gauge true positive and true negative rates, respectively, while FPR and FNR help understand the model's error rates in misclassifying patients. The table 6.1 summarized the performance of these parameters when evaluated for different algorithms.

Throughout the scope of this investigation, the primary objective is to construct a system that is capable of performing autonomous, real-time self-diagnosis by utilizing a combination of fog, cloud, and internet of things

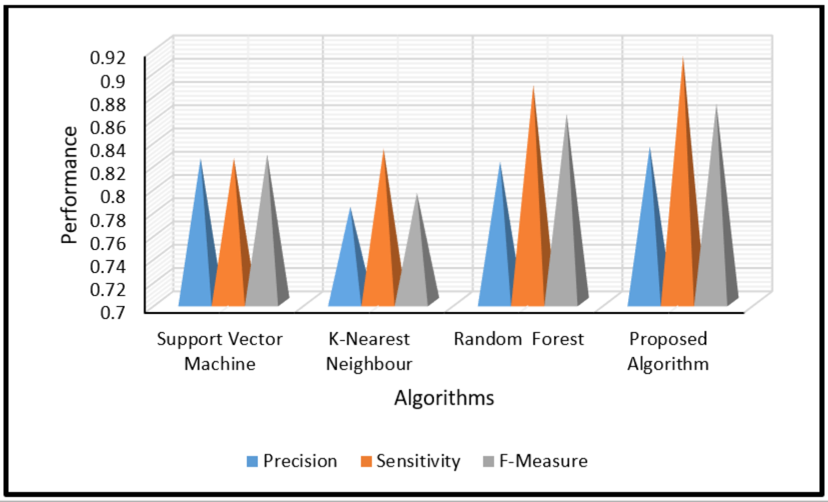

Fig. 6.2: The comparison of the proposed methods based on different evaluation parameters

Table 6.2: Assessment of the Projected Technique with Some State-of-the-Art Models

Methods	**Accuracy (%)**	**Precision (%)**	**Sensitivity (%)**	**Specificity (%)**	**F-Measure (%)**
ELM with Boosting [14]	81.32	82.18	74.00	87.34	76.24
DT, NN, RS, SVM, and NB [16]	88.91	89.20	88.49	82.70	86.90
SVM, MLP, LR, and DT [17]	92.54	92.57	92.46	91.08	92.48
ML and CART [20]	93.25	96.18	91.27	91.27	93.48
Proposed Work	94.21	96.82	91.24	92.62	93.94

architectures. Through the utilization of principal component analysis (PCA) to reduce the dimensionality of the data, support vector machines (SFS) to choose features, and a random forest (RF) classifier, an integrated model has been developed for the purpose of predicting heart disease. For the purpose of carrying out the experiments, the CHDD and the programming language Python are utilized. It turns out that the model that was proposed is light years ahead of the other categorization strategies that are more conventional that are now in use. An analysis of the results reveals that the proposed system IHDPM, which is depicted in Figure 6.1, achieved the highest possible accuracy of 83.56% in comparison to the other conventional machine learning approaches. Additionally, when compared to other ways, the system that was recommended performs better than those other approaches in terms of precision, recall, and F-measure calculation, as shown in figures 6.2.

Further the projected technique has evaluated with the state-of-the-art research work on the same field and compared based on different evaluation parameters tabulated in table 6.2.

From figure 6.3, it can be concluded that the proposed method outperforms than existing method by showing improved accuracy of 94.21%, improved precision of 96.82%.

In figure 6.4, the other parameters, which include sensitivity, specificity, and F-measure, are displayed. These parameters have respective performance levels of 91.24%, 92.62%, and 93.94%. The increased presentation that was achieved with the projected method is demonstrated by the performance evaluation of the methods, which demonstrates that the approach that was projected was successful. It is feasible that the method that has been developed will assist cardiologists in improving the accuracy of their diagnoses. Following the appearance of the trained model that was proposed, the question of whether or not it is capable of assisting patients with remote self-diagnosis through the use of cellphones has surfaced.

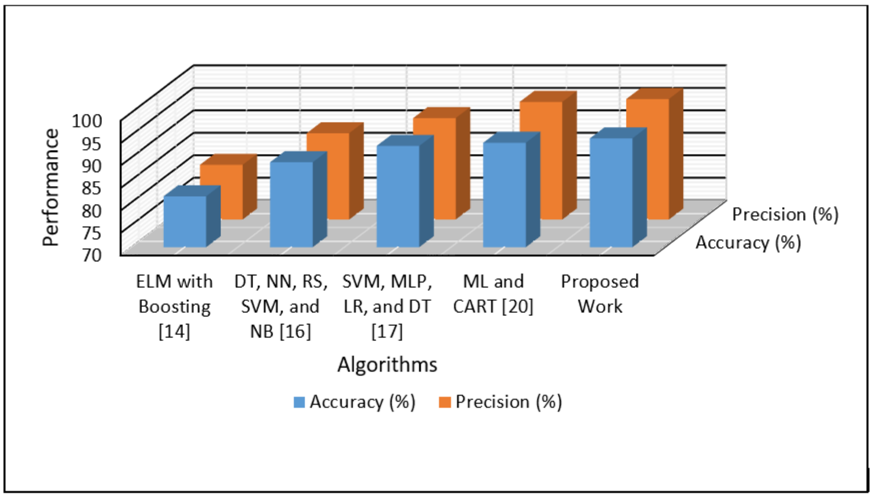

Fig. 6.3: The comparison of the proposed methods based on different evaluation parameters

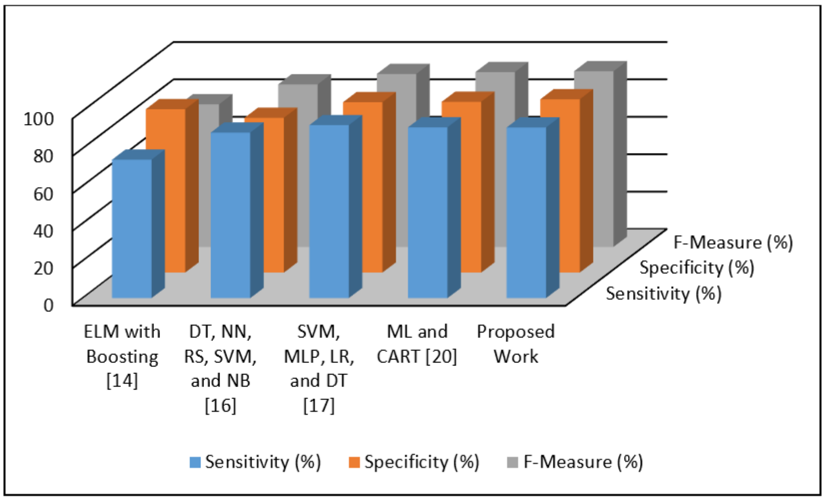

Fig. 6.4: The Sensitivity, Specificity and F-Measure Comparison of the proposed method

7. Conclusion and Future Scope. An efficient strategy for the prediction of cardiac illness is presented in this study article. By removing irrelevant and unsuitable topographies from the dataset and focusing on the utmost relevant ones for the cataloguing task, this suggested model aims to enhance performance and obtain more accurate predictions of cardiac illnesses utilizing classifiers. The suggested system achieves dimensionality reduction with PCA, FS with SFS, and classification with RF. The outcomes are next evaluated in comparison to the alternative classification methods, specifically SVM, KNN, and RF. By doing trials using the CHDD dataset, which is obtained from the UCI-ML repository, and the Python programming language, it has been determined that the suggested model outperforms more traditional approaches.

The suggested system will aid medical professionals in making accurate diagnoses and predictions about heart patients, and it has the potential to pave the way for additional research into prediction using various datasets, which could lead to important new information about heart disorders. While DL approaches may produce more accurate prediction results, it would be more operative to reduce the dimensions count before cluster generation to improve the results. The next step in our research should be to determine if our trained model can be used for remote cardiac patients to do self-diagnostics in real-time.

REFERENCES

[1] M. A. Lakhan, S. Mohammed, S. Kozlov, and J. J. P. C. Rodrigues, "Mobile-fog-cloud assisted deep reinforcement learning and blockchain-enable IoMT system for healthcare workflows," *Transactions on Emerging Telecommunications Technologies*, Sep. 2021, doi: https://doi.org/10.1002/ett.4363.

[2] C. Mouradian, D. Naboulsi, S. Yangui, R. H. Glitho, M. J. Morrow and P. A. Polakos, "A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 416-464, First quarter 2018, doi: 10.1109/COMST.2017.2771153.

[3] S. Svorobej et al., "Simulating Fog and Edge Computing Scenarios: An Overview and Research Challenges," *Future Internet*, vol. 11, no. 3, p. 55, Feb. 2019, doi: 10.3390/fi11030055.

[4] K.-L. Lai and J. I. Z. Chen, "Development of Smart Cities with Fog Computing and Internet of Things," *Journal of Ubiquitous Computing and Communication Technologies*, vol. 3, no. 1, pp. 52–60, May 2021, doi: https://doi.org/10.36548/jucct.2021.1.006.

[5] Cisco Knowledge Networking, "Cisco Global Cloud Index: Forecast and Methodology, 2015-2020," white paper, Cisco Public, San Jose, 2016.

[6] A. Lakhan, M. A. Mohammed, J. Nedoma, R. Martinek, P. Tiwari, and N. Kumar, "DRLBTS: deep reinforcement learning-aware blockchain-based healthcare system," *Scientific Reports*, vol. 13, no. 1, p. 4124, Mar. 2023, doi: https://doi.org/10.1038/s41598-023-29170-2.

[7] S. Tanwar et al., "Human Arthritis Analysis in Fog Computing Environment Using Bayesian Network Classifier and Thread Protocol," *IEEE Consumer Electronics Magazine*, vol. 9, no. 1, pp. 88-94, 1 Jan. 2020, doi: 10.1109/MCE.2019.2941456.

[8] L. Greco, G. Percannella, P. Ritrovato, F. Tortorella, and M. Vento, "Trends in IoT based solutions for health care: Moving AI to the edge," *Pattern Recognition Letters*, vol. 135, pp. 346–353, Jul. 2020, doi: https://doi.org/10.1016/j.patrec.2020.05.016.

[9] A. A. Alli and M. M. Alam, "The fog cloud of things: A survey on concepts, architecture, standards, tools, and applications," *Internet of Things*, vol. 9, p.100177, Mar. 2020, doi: https://doi.org/10.1016/j.iot.2020.100177.

[10] R. M. Abdelmoneem, A. Benslimane, and E. Shaaban, "Mobility-Aware Task Scheduling in Cloud-Fog IoT-Based Healthcare Architectures," *Computer Networks*, p. 107348, Jun. 2020, doi: https://doi.org/10.1016/j.comnet.2020.107348.

[11] S. Shukla, V. Roy and A. Prakash, "Wavelet Based Empirical Approach to Mitigate the Effect of Motion Artifacts from EEG Signal," *2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT)*, 2020, pp. 323-326, doi: 10.1109/CSNT48778.2020.9115761.

[12] D. Fernández-Cerero, A. Jakóbik, A. Fernández-Montes, and J. Kołodziej, "GAME-SCORE: Game-based energy-aware cloud scheduler and simulator for computational clouds," *Simulation Modelling Practice and Theory*, vol. 93, pp. 3–20, May 2019, doi: https://doi.org/10.1016/j.simpat.2018.09.001.

[13] J. Hasenburg, M. Grambow, E. Grünewald, S. Huk and D. Bermbach, "MockFog: Emulating Fog Computing Infrastructure in the Cloud," *2019 IEEE International Conference on Fog Computing (ICFC)*, Prague, Czech Republic, 2019, pp. 144- 152, doi: 10.1109/ICFC.2019.00026.

[14] K. H. Miao, J. Miao, and G. Miao, "Diagnosing Coronary Heart Disease Using Ensemble Machine Learning," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, 2016.

[15] A. Alqhatani, S. B. Khan, "IoT-Driven Hybrid Deep Collaborative Transformer with Federated Learning for Personalized E-Commerce Recommendations: An Optimized Approach," *Scalable Computing: Practice and Experience*, vol. 25, no. 5, 2024.

[16] A. A. Khan, K. K. Almuzaini, V. D. J. Macedo, S. Ojo, V. K. Minchula, and V. Roy, "MaReSPS for energy efficient spectral precoding technique in large scale MIMO-OFDM," *Physical Communication*, vol. 58, 2023, doi: https://doi.org/10.1016/j.phycom.2023.102057.

[17] A. Esfahani and M. Ghazanfari, "Cardiovascular disease detection using a new ensemble classifier," *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, Tehran, Iran, 2017, pp. 1011-1014, doi: 10.1109/KBEI.2017.8324946.

[18] M. Ganesan and N. Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models," *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, India, 2019, pp. 1-5, doi: 10.1109/ICSCAN.2019.8878850.

[19] S. Shaik, J. Hall, C. Johnson, Q. Wang, R. Sharp, and S. Baskiyar, "PFogSim: A simulator for evaluation of mobile and hierarchical fog computing," *Sustainable Computing: Informatics and Systems*, vol. 35, p. 100736, Sep. 2022, doi: https://doi.org/10.1016/j.suscom.2022.100736.

[20] Y. Li, W. Qu, and Z. Zhang, "Intelligent Algorithm Operation and Data Management of Electromechanical Engineering Power Communication Network based on the Internet of Things," *Scalable Computing: Practice and Experience*, vol. 25, no. 5, 2024.

[21] O. U. Akgul, W. Mao, B. Cho, and Y. Xiao, "VFogSim: A Data-driven Platform for Simulating Vehicular Fog Computing Environment," 2022.

[22] V. Roy, P. K. Shukla, A. K. Gupta, V. Goel, P. K. Shukla, and S. Shukla, "Taxonomy on EEG Artifacts Removal Methods, Issues, and Healthcare Applications," *Journal of Organizational and End User Computing (JOEUC)*, vol. 33, no. 1,

pp. 19-46, 2021, doi: 10.4018/JOEUC.2021010102.

[23] O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, ”An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction,” *Expert Systems with Applications*, vol. 68, pp. 163–172, Feb. 2017, doi: https://doi.org/10.1016/j.eswa.2016.10.020.

[24] L. Ali et al., ”An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure,” in IEEE Access, vol. 7, pp. 54007-54014, 2019, doi: 10.1109/ACCESS.2019.2909969.

[25] R. Zhou, “Research on Intelligent Building Integrated Cabling System Based on Internet of Things and Deep Learning”, Scalable Computing: Practice and Experience, Vol. 25 No. 5, 2024.

[26] S. Stalin, V. Roy, P. K. Shukla, A. Zaguia, M. M. Khan, P. K. Shukla, A. Jain, ”A Machine Learning-Based Big EEG Data Artifact Detection and Wavelet-Based Removal: An Empirical Approach”, Mathematical Problems in Engineering, vol. 2021, Article ID 2942808, 11 pages, 2021. https://doi.org/10.1155/2021/2942808.

[27] D. Mienye, Y. Sun, and Z. Wang, “An improved ensemble learning approach for the prediction of heart disease risk,” Informatics in Medicine Unlocked, vol. 20, p. 100402, 2020, doi: https://doi.org/10.1016/j.imu.2020.100402.

[28] S. Gupta, “Classification of Heart Disease Hungarian Data Using Entropy, Knnga Based Classifier and Optimizer,” International Journal of Engineering & Technology, vol. 7, no. 4.5, p. 292, Sep. 2018, doi: https://doi.org/10.14419/ijet.v7i4.5.20092.

[29] M. Jan, A. A. Awan, M. S. Khalid, and S. Nisar, “Ensemble approach for developing a smart heart disease prediction system using classification algorithms,” Research Reports in Clinical Cardiology, vol. Volume 9, pp. 33–45, Dec. 2018, doi: https://doi.org/10.2147/rrcc.s172035.

[30] W. Liu, “The Trajectory Data Mining Model for College Students in Compus Life and Academic Management”, Scalable Computing: Practice and Experience, Vol. 25, Issues 5, pp. 3225–3240, DOI 10.12694/scpe.v25i5.2349.

Edited by: Manish Gupta
Special issue on: Recent Advancements in Machine Intelligence and Smart Systems
Received: Aug 25, 2024
Accepted: Oct 14, 2024