ADVANCING HEART DISEASE PROGNOSIS WITH A HYBRID MACHINE AND DEEP LEARNING APPROACH

Ibrahim Abunadi [1], Lakshmana Kumar Ramasamy [2]

[1] Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[2] Computer and Information Science, Higher Colleges of Technology, Ras Al Khaimah, United Arab Emirates

ABSTRACT

Heart disease is a leading cause of death globally, claiming approximately 17 million lives each year. Often, these deaths are due to heart failure, a condition where the heart cannot supply enough blood to meet the body's needs. To improve diagnosis and treatment, healthcare professionals increasingly rely on electronic medical records. These records are invaluable for detecting subtle patterns in symptoms and test results that might otherwise go unnoticed. In the realm of medical data analysis, data mining techniques have shown promise in predicting the outcomes of cardiovascular diseases. However, major challenges can occur—overfitting and managing large dimensions of data—can hinder their effectiveness. To address these issues, this paper proposes a novel method that simplifies the data through feature selection, making this model not only more efficient but also easier to understand. Specifically, we introduce a new framework that combines advanced feature selection algorithms (sequential forward and backward, or CSFB) with a blend of traditional machine learning and cutting-edge deep learning techniques. Utilizing algorithms like Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and a deep learning classifier (Dl4jMlpClassifier), this method refines the data to improve predictions of heart disease outcomes. This work findings confirm that this integrated approach -CSFB feature selection combined with the CMD (Combined Machine and Deep learning) algorithm effectively identifies crucial data features and reliably predicts patient survival rates. This advancement holds significant potential for enhancing heart disease diagnostics and patient care strategies.

KEYWORDS

Heart disease prediction, Cardiovascular disease (CVD) prognosis, Deep learning, Feature selection, Overfitting prevention

1. INTRODUCTION

Coronary heart disease (CHD) is considered one of the most multifaceted life-threatening human diseases. Usually, in such conditions, the heart cannot carry enough blood to the whole body to carry out its standard functions [1]. Thus, it finally causes heart failure (HF). In the United States, the incidence of heart disease is high, where one death occurs every half a minute [2]. Manifestations of coronary artery disease include physical deformities, swelling, and fatigue of the legs [3].

Furthermore, coronary artery disease results are so significant, specifically in developing countries, that the lack of integrated diagnostic machines, physicians, and other assets can impact

David C. Wyld et al. (Eds): SIGI, CSTY, AI, NMOCT, BIOS, AIMLNET, MaVaS, BINLP – 2025
pp. 01-21, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.151901

the therapy of heart disease patients and hurt their expectations [4]. Therefore, a precise and rational examination of patients' CHD threat is essential to decrease their risk of acute heart disease and enhance cardiovascular security. The European Cardiovascular Society reports show that 26 million older peoples have coronary artery disease each year [5]. Half of the people with CHD may die within 1–2 years, worrying about their treatment costs. Therefore, to reduce the number of deaths of this type, heart disease should be stopped early. Therefore, healthcare should give more significance to heart disease forecasting. Data mining methods enable diagnosing CHD at an initial stage by utilizing the patient's health record [6]. To diagnose heart disease, many learning use ML techniques.

Machine learning is considered a branch of artificial intelligence (AI) [7]. The purpose of ML, in general, is to understand and discover the organization and format of data and to make predictions. Although ML is a domain in computer technology, it is different from conventional calculation techniques. On the one hand, conventional computer algorithms are a set of explicitly planned algorithms used by computers to solve a problem. On the other hand, ML techniques allow computers to train input data and perform arithmetic for results within a given range [8]. Technology users can now use ML to create a database of sample logs and perform decision-making processes using input data. For example, facial recognition technology allows social media sites to tag users and distribute pictures of friends [9]. The technology of optical character recognition (OCR) alters text images to movable shapes [10]. ML-based recommendation engine recommends what movie to watch, subsequently using the preferences of the user [11]. ML-based self-driving cars will be available to customers soon [12].

However, good preprocessing is essential in using an ML technique. The big challenge is choosing suitable strategies to minimize the use of unnecessary features in the dataset. In general, healthcare records consist of large data with several features. Therefore, the application of ML for this large data with several features impacts the techniques' accuracy [13]. To get the best result in prediction, you need to find the appropriate feature selection techniques. The main goal is to identify the key features and classification techniques for developing an accurate forecast model. Feature selection is the automatic selection of the most sophisticated features for ML techniques through intelligent analysis of available information and it does not require the use of expert awareness [14]. Feature selection is the process of reducing the number of input variables when developing a predictive model. It is desirable to reduce the number of input variables to both reduce the computational cost of modelling and, in some cases, to improve the performance of the model.

The most important reasons for utilizing feature selection are as follows [15]: Utilizing feature selection, a consumer can eliminate unnecessary features that will not disturb or alter the result of their technique. For example, if a user attempts to forecast the cost of a home in Spain, utilizing features that comprise the climate circumstances in China, then these features would possibly not be beneficial. These types of inappropriate features will reduce the performance of an ML technique by introducing noise. Fewer features usually mean faster training models: for parameter models such as linear or logistic regression, there are fewer weights to calculate, and for non-parameter models such as Random Forest of decision trees, there are fewer features to evaluate in each category.

When models are put into production, fewer features mean less work for the application-making team that uses the model. Users can reduce the integration time for the application by using the feature selection algorithm. While a user maintains the vital features, removing the ones that their feature selection techniques direct him to eliminate also makes their process easier. A technique that has 25 features is significantly more effective than the one that has 200 features.

It is much easier to debug a model with fewer features of abnormal behavior than a multi-feature model. Furthermore, deep learning (DL) is a kind of AI that emulates specific skill types of persons [16]. DL is a significant component of data science that comprises predictive modelling. It is advantageous to big data wranglers working on gathering, examining, and understanding big data; DL makes this procedure faster and simpler. DL is very popular for its neural network, for example, recurrent neural networks (RNN) [17]. When other ML techniques use the predictive analysis method to recognize patterns automatically, DL is designed to model neurons in the human brain.

DL techniques use the functional structure of the human brain. To understand DL, we need to understand how the complex nervous system in the human body works. The nervous system is made up of neurons. These neurons are capable of absorbing the data that is transmitted to our bodies. In addition, neurons can recognize data over time. Artificial neural networks (ANN) further use this "learning" principle. An ML technique is utilized to parse information, learn from inputted information, and create learned decisions using what was discovered. Fundamentally, DL is being used to form layers to build an ANN [18]. However, it can, on its own, train and create smart results. Table 1 compares DL versus ML through their characteristics [19].

Table 1. Comparison of Deep Learning and Machine Learning

Characteristics	**Deep Learning**	**Machine Learning**
Time to Training	Typically takes more time due to model complexity and larger datasets, but can be reduced with techniques like transfer learning.	Often takes less time but can vary depending on the model and data size.
Accuracy	Generally provides high accuracy on large datasets but can overfit on small datasets.	Can perform well on small datasets and may outperform deep learning in certain cases with proper feature engineering.
Size of the Data	Performs better with large datasets and may struggle with small data	Can train effectively on small datasets with proper feature engineering.
Hyperparameter Tuning	Highly flexible and offers various tuning options but can be computationally expensive.	Limited tuning capabilities but simpler to optimize in practice.
Hardware Dependency	Often requires GPUs for efficient training, especially for large models.	Can train on CPUs effectively and may not require specialized hardware.

DL provides vast accuracy, but ML provides less accuracy. Furthermore, DL requires vast data, but ML can train more on small data [20]. In this way, we can understand that one technique solves the defect of another technique. Therefore, this paper proposes the combination of sequential forward and backward (CSFB) feature selection algorithm with ML and DL (CMD) algorithms. Figure 1 shows the system Architecture of the proposed CSFB.

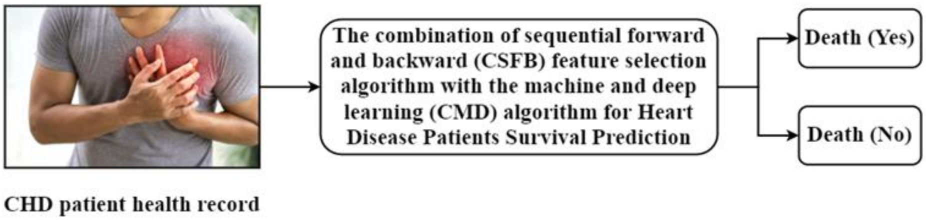

Figure 1: CSFB System Architecture.

In the existing studies, the researchers have used the sequential forward or sequential backward feature selection technique. However, in this work, we have combined both types of sequential feature selection. As a result, it selects standard features from both types and provides the best and optimal features compared to existing techniques. Moreover, it uses CMD techniques. Therefore, compared with individual learning techniques, the combination provides the best prediction results. Furthermore, many researchers have focused on predicting whether a person will have heart disease or not. However, this paper focuses on predicting whether or how long a patient who has heart disease will survive or not.

The remaining sections are explained as follows. The existing works related to heart disease prognosis utilizing the required features and numerous data mining techniques are discussed in Section II. Section III discusses the problem definition. Section IV addresses the description of the dataset taken for this paper. Section V discusses the CSFB feature selection algorithm with the CMD algorithm to predict heart disease patient survival. Section VI presents the simulation results of the proposed work. Lastly, Section VII provides the conclusion.

2. Materials and Methods

This section reviews related works on cardiovascular disease forecasting based on various blends of feature selection and classification techniques.

2.1. Feature Selection with Ml Techniques for Heart Disease Data

Haq et al. [21] utilized three techniques for feature selection: absolute shrinkage and selection operator (LASSO), Maximum Relevance, Minimum Redundancy (mRMR), and Relief and used seven techniques for ML, for example, SVM, logistic regression (LR), ANN, KNN, RF, DT, and NB for classification. In many cases, the LR reached a maximum accuracy of 89% for heart disease prediction.

Mohan et al. [22] provided a Hybrid Random Forest with a Linear Model (HRFLM) for heart disease forecasting. The HRFLM method combined the characteristics of the linear method and RF. As a result, HRFLM was verified for entirely precise heart disease forecasting. Furthermore, this research used different feature combinations and numerous classification techniques. Here, the RF technique attained 88.7% more accuracy compared to other methods. Pawlovsky [23] implemented an ensemble distance for the KNN algorithm to forecast cardiovascular disease, and this method attained an accuracy of 85%. Two configurations were developed within the ensemble. The first used three distances (Euclid, Manhattan, and Mahalanobis) and the second used five distances (Canberra and Mahalanobis, Sorensen, Chebyshev, Manhattan, and Euclid). He also added the weighted average accuracy that each distance provides while utilizing the KNN technique.

Vijayashree et al. [24] presented a fitness function (FF) for particle swarm optimization (PSO) with the assistance of SVM. The goal of the FF is to decrease the feature count and raise the accuracy significantly. They compared the effectiveness of a presented PSO–SVM with numerous previous feature selections, for example, information gain, Chi-square, CFS, Relief, PSO algorithm, and filtered attributes. Moreover, the SVM algorithm is compared with multiple algorithms, for example, Multilayer Perceptron (MLP), RF, and Naïve Bayes.

Magesh et al. [25] presented a heart disease forecasting technique, namely cluster-based DT Learning (CDTL). This technique primarily contains five vital stages. First, the original set was separated by the target label distribution. From the higher sharing models, other possible class combinations have been developed. Third, for each class-set mixture, important features are identified by entropy. Fourth, an entropy-based partition has been developed with significant key features. Finally, in these entropy clusters, RF effectiveness is noteworthy in the entire features of predicting heart disease. Using their CDTL technique, the RF algorithm attained 89.30% enhanced forecast accuracy from 76.70% accuracy (with no CDTL). Gárate-Escamila et al. [26] presented a dimensionality reduction technique that also identified features of cardiovascular disease by using a ChiSqSelector (CHI) approach. The dataset utilized for their work was attained from the UCI repository. The dataset had 74 features with a label that they verified through six ML classification algorithms. Chi-square, Principal Component Analysis (PCA) and RF provided 98.7 % accuracy.

Aggrawal et al. [27] provided a sequential feature selection algorithm for discovering the mortalities of heart disease patients. Numerous ML techniques (RF, GBC, LDA, KNN, SVM, and DT) were utilized. Numerous performance metrics, especially accuracy, were computed to confirm the outcomes attained through the sequential forward selection (SFS) technique. The effects demonstrated that for RF Classifier_FS, the SFS technique attained 86.67% accuracy. Gokulnath et al. [28] presented an optimization function in terms of SVM. This optimization function is utilized in the genetic algorithm (GA) technique for choosing the considerable features and the likelihood of obtaining cardiovascular disease. The outcomes of the SVM-GA are compared with the previous techniques, such as filtered subset, CFS, Relief, and Chi-square, consistency, subset, information gain, GA, gain ratio, filtered attribute, and one attribute-based algorithm. The SVM-GA provided 88.34% accuracy compared to others. This method was established in the matrix LABoratory (MATLAB) surroundings by a dataset gathered from the cardiovascular disease dataset.

Bashir et al. [29] conducted the forecasting of heart disease in the medical domain by utilizing data science. However, the accuracy of the forecast is still needed to be enhanced. Therefore, their study deliberates on feature selection algorithms where numerous cardiovascular disease data are utilized for the research on how to increase the accuracy of feature selection techniques. By utilizing the rapid miner tool, RF, Naïve Bayes, LR, SVM, LR, and DT the techniques are being used as feature selection algorithms, and the development of the same is demonstrated by the outcomes in terms of the attained accuracy. Ahmed et al. [30] provided an actual time scheme for forecasting cardiovascular disease from clinical information streams explaining the present health condition of a patient. The principal aim of their work was to discover the most favorable ML technique that attains tremendous accuracy for the prediction of cardiovascular disease. Two kinds of feature selection techniques, relief and univariate feature selection, were utilized to choose significant features from the dataset. The authors compared four ML techniques: DT, SVM, RF, and LR classifier by the chosen features. The authors applied cross-validation with ML to boost accuracy. An essential qualification of this work is the ability to efficiently manage Twitter data streams containing patient data. This is done by incorporating Apache Kafka with Apache Spark. Their outcomes demonstrated that the RF algorithm attained an immense accuracy of 94.9%.

Khourdifi et al. [31] used the fast correlation-based feature selection (FCBF) technique to filter unnecessary features and to enhance the accuracy of cardiovascular disease prediction. After that, they performed classification using various classification techniques, for example, KNN, SVM, Naïve Bayes, RF, and MLP, enhanced through PSO integrated with ant colony optimization (ACO) algorithm. The integrated approach is used in the cardiovascular disease dataset; the outcomes showed the effectiveness and strength of the integrated approach in classifying different kinds of records for cardiovascular disease prediction. Thus, this work used various ML techniques and compared the outcomes using multiple evaluation metrics. Le et al. [32] provided an automated heart disease forecasting technique using feature selection with the DM method that used the presented signs and medical data allocated to the patient. In their approach, heart disease features were weighted and reorganized using their weights and rank distributed through the infinite latent feature selection (ILFS) approach. A soft margin SVM was utilized to categorize the subgroups of chosen features in various cardiovascular disease classes. They concluded that the efficiency of their process for heart disease forecast creation was quite high and exact; the effectiveness of their approach was the most satisfactory and had achieved 90.65% of accuracy and 0.96 area under the curve (AUC) for differentiating "absence" of heart disease with "no absence" of heart disease.

Usman et al. [33] presented two associated cuckoo encouraged algorithms, namely, the cuckoo optimization algorithm (COA) and cuckoo search algorithm (CSA), for feature selection on a cardiovascular disease dataset. Both algorithms utilized the common filter technique while subgroup creation. They concluded that CSA worked superiorly to COA in both a few features and predictive accuracy in all datasets. Finally, CSA performed better on all datasets compared to existing approaches. Kohli et al. [34] used various classification techniques. Each had its benefit on three individual disease datasets (heart disease, diabetes, and breast cancer). The feature selection for all datasets was passed through the backward technique using the p-value analysis. The outcomes of the technique make the design of the appliance of ML stronger in the untimely discovery of diseases.

Latha et al. [35] examined an approach, namely the ensemble approach, which was employed for enhancing the accuracy of feeble plans through merging several classifiers. A diagnostic approach was used to decide how the ensemble approach could enhance the forecasting accuracy of heart disease. The outcomes of the work specify that ensemble approaches, namely, boosting and bagging, are efficient in enhancing the predictive accuracy of feeble classification algorithms and express satisfying effectiveness in diagnosing heart disease risk. The assistance of ensemble classification attained an utmost boost of 7% accuracy for feeble classification algorithms. The performance of the procedure was also improved with the feature selection execution, and also the outcomes demonstrated noteworthy enhancement in the forecasting accuracy.

2.2. Feature Selection with DL Techniques for Heart Disease Data

Uyar et al. [36] construct a cardiovascular disease forecasting method using GA with recurrent fuzzy neural networks (RFNN). The UCI Cleveland cardiovascular disease dataset was used. The consequences demonstrated that 91.78% accuracy was attained. Yazid et al. [37] discussed a forecasting method using an ANN through a parameter-tuning structure. Cleveland and Statlog datasets were used to assess the efficiency of the method. The outcomes demonstrated that this method could create colossal accuracy (90% for Statlog and 90.9% for Cleveland).

Ali et al. [38] proposed a combined DL with feature fusion technique for cardiovascular disease forecasting. Initially, the feature synthesis technique combined the taken-out features from sensed information with electronic health data to create precious medical records. Next, the info gain method removed unnecessary features, and it also chose the significant ones that reduced the calculation load and boosted the efficiency of the method. Additionally, the conditional probability approach calculated the particular feature weight of all classes, which also enhances the system's performance. Lastly, the combined DL technique was trained for cardiovascular disease forecasting. This technique assessed cardiovascular disease information also compared to conventional classification techniques. This technique attained 98.5% accuracy which was superior to previous approaches. They concluded that their approach was very efficient in forecasting cardiovascular disease compared with other existing approaches.

Khan et al. [39] presented an internet of things (IoT) structure for assessing cardiovascular disease very precisely by using the Modified Deep CNN (MDCNN). The smartwatch and a device for monitoring the heart, which was fixed to the patient, examined the blood pressure (BP). The MDCNN was used to categorize the obtained sensed information as abnormal or usual. The system performance was assessed by comparing the MDCNN with prior LR and DL. They concluded that the cardiovascular disease forecasting based on MDCNN performs superior to existing techniques. The MDCNN attained an accuracy of 98.2% which is considered an improvement to previous classification techniques. Ali et al. [40] concentrated on both, that is, features' refinement and removal of underfitting and overfitting issues through the forecasting technique. By evading the approach to these issues, it could demonstrate high-quality effectiveness in training datasets with the testing datasets. Unsuitable configuration of the network with unnecessary features frequently results in fitting the training dataset. To remove unrelated features, the authors recommended operating the Chi-square statistical technique when the finest configured deep neural network (DNN) is explored by utilizing a comprehensive exploration plan. The potency of the mixed process called Chi-square Dnn evaluates its performance by comparing it through the usual ANN and DNN techniques. This technique attained a 93.33% forecasting accuracy. The attained outcomes were hopeful compared with the formerly described techniques. Study findings suggest that physicians can use diagnostic methods to accurately predict heart disease.

Darmawahyuni et al. [41] presented an imitation that could analyze CHD with improved efficiency by using conventional analytic techniques. A few researchers implemented the technique using CNN techniques; however, the outcomes were not highly effective. Based on the regular neural network (NN), the DNN was proposed as a model for this work. Known as NN, it is a supervised learning technique that specializes in categorization. Using the DNN technique, binary classification was performed to analyze CHD absent or CHD present. This technique attained 96% accuracy, 99% sensitivity, and 92% specificity.

Kong et al. [42] presented the graph-embedded deep feedforward network (GEDFN) for combining exterior relational data with features in the DNN structure. The technique was capable of attaining sparse links among network layers to avert overfitting. The resulting superior classification accuracy and easy-to-explain feature selection outcomes recommend that this technique is a helpful addition to present classification techniques in feature selection processes. Chen et al. [43] implemented a CNN technique to discover cardiac arrhythmia (CA) using a vast 12-lead electrocardiogram (ECG) dataset. Their technique attained an F-measure of 0.84. An additional study identified which simultaneous CA was sufficiently forecasted for 476 patients by numerous kinds of CA analysis in the dataset. Utilizing merely single-lead information provided effectiveness that was merely slightly inferior to utilizing the complete 12-lead information for guiding a VR, with V1 being very famous. The authors broadly regarded these outcomes in the circumstance of their contract with and significance to medical observation.

Krishnan et al. [44] presented DNN structures, for example, a feedforward neural network (FNN) and one-dimensional CNN (1D-CNN) for phonocardiogram (PCG) signal classification. The authors intended to mechanize the feature selection procedure employed in the study of the PCG signal. The unique PCG signal was reduced to 500 Hz. After that, the unique PCG signals are separated into lesser time sections of 6 s epoch. Savitzky–Golay filter was utilized to repress the high-frequency noise in the signal through data point softening. The implemented dataset was subsequently presented as an input for the DNN structures. The FNN technique presented a higher accuracy of 0.86. Dutta et al. [45] presented an effective neural network with a convolution layer for categorizing considerably class-imbalanced medical datasets. While most of the ML models used in this class of data are subject to class inequality even after adjusting for class-specific weights, their two-layer CNN reveals the pliability of inequalities with reasonable consistency in class-specific effectiveness. Provided extremely imbalanced data, it is frequently tricky to concurrently attain superior class 1 accuracy with large class 0 accuracy, because the size of the testing dataset rises. The authors used a two-stage technique: initially, the authors used feature weight evaluation based on LASSO, followed by a greater number of voting-based discoveries of significant features. After that, the key features are uniformly modified by utilizing an entirely bonded layer, which is an important step before sending the layer output to the next transformation stage. The authors further presented a training schedule per epoch, like an imitated annealing procedure to improve the accuracy of classification. Despite a superior class inequity of the NHANES dataset, the analysis verified that their CNN structure had 77% power of classification for properly categorizing the appearance of CHD with 81.8% for properly classifying the nonappearance of CHD cases in a test dataset, that is, 85.70% of the whole dataset.

Zhang et al. [46] proposed a DNN-based automated sleep level score using the polysomnography (PSG) dataset. To capture high-sequence records, spectrograms were created from electromyography, electrophotography, and electroencephalography records and transmitted to the neurological network. A whole collection of 580 PSGs not incorporated in the training data was utilized to evaluate technique accuracy and bias by Cohen's kappa (κ), per-level accuracy, and weighted F1 score. The optimal NN model is made up of spectrograms in the input layer, fed to the CNN layer and long short-term memory layer (LSTM) to attain a weighted F1 score of 0.87. Hannun et al. [47] implemented a DNN to categorize 12 rhythm classes utilizing 91,232 1-lead ECGs from 53,549 patients who utilized a 1-lead ambulatory ECG tracking tool. While authenticated in opposition to autonomous test data interpreted through the agreement group of team-certified working cardiologists, the DNN attained the Receiver Operating Characteristic (ROC) curve of 0.95. The average F1 score was 0.84. These findings showed that an end-to-end DL technique could categorize various unique arrhythmias from 1-lead ECG with tremendous analytical effectiveness like that of heart specialists. If verified in a medical environment, this technique can decrease the ratio of misdiagnosing automated ECG analysis. It could also enhance the effectiveness of specialist ECG reading by precisely checking or prioritizing very urgent circumstances.

Cai et al. [48] presented a DL technique with tremendous accuracy for atrial fibrillation (AF) discovery. The technique built a one-dimensional deep densely linked neural network (DDNN) to discover AF an ECG waveform with a length of 10 s. From the testing data, the DDNN attained tremendous effectiveness with an accuracy of 99.35%. Its massive practical nature showed that the network contained a vast possibility for medical computer assistance, analysis of AF, or upcoming viewing of AF on the wearable device. Kumar et al.[52] presented a Survey on blockchain for the industrial Internet of Things. Khan et al.[53] worked on an Efficient, Ensemble-Based Classification Framework for Big Medical Data. Pooranam, N. et al. [54] proposed a decision support mechanism to improve a secured system for clinical processes using Blockchain Technique. Reddy G. Thippa et al.[55] developed the hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis. The same authors in [56] worked on the early detection of diabetic retinopathy using a PCA-firefly-based D deep learning model. Gadekallu, Thippa Reddy et al.[57] suggested that Deep neural networks predict diabetic retinopathy where the efficiency was proved well.

Kim et al. [49] presented a neural network (NN) technique for forecasting CHD threat by Feature Correlation Analysis (FCA), namely, NN-FCA. It consisted of two steps. Initially, the feature selection step that creates features agreeing with the significance of forecasting CHD hazards was graded. Feature correlation analysis level, which studies the relationship among feature relationships with the data of all NN forecast results. The ROC curve of this technique was 0.749. The NN-FCA FCA originated to be superior to FRS in terms of CHD hazard forecast. Additionally, this technique provided a superior ROC curve and precise forecasting of CHD hazards.

Baviskar et al. [50] presented a feature selection and classification technique for the forecasting of cardiovascular disease. For feature selection, the improved GA and PSO have been executed. The RNN and LSTM had been executed for classification. At last, the authors concluded that LSTM, when merged with PSO, demonstrated 93.5% accuracy than other existing techniques. Thus, this technique could be used in the clinical domain for precise cardiovascular disease forecasting. Furthermore, Table 2 demonstrates the outline of the related work. In [58], four well-known Machine Learning (ML) algorithms—Decision Tree Induction, Support Vector Machine (SVM), Naive Bayes Classifier, and Random Forest Classifier—are investigated using two well-known dimensionality reduction techniques, Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA), using publicly available Cardiotocography (CTG) dataset from University of California and Irvine Machine Learning Repository. The experiment's findings demonstrate that PCA performs better than LDA in all the measures.

3. PROBLEM DEFINITION

Analyzing multifaceted data and mining large amounts of data is time-consuming and difficult. Therefore, data reduction methods, for example, data cube integration, attribute subgroup selection, dimension reduction, numerical reduction, data uniqueness, and concept hierarchical creations, are used in this study. When the user uses any of the data reduction methods, they should minimize using a dataset that is devoid of missing valuable data. One more challenge with classification is the accuracy of the classification algorithm. The accuracy relies not only on the classifier but also on the feature selection algorithm. Selecting unsuitable features can lead to aggravation of the classification problem. The feature selection algorithm has an important role that increasing the effectiveness of the classification. Although various types of feature selection algorithms are obtainable, the finest algorithm must be selected to enhance the classification accuracy to choose the suitable features, and the feature selection method must spend less space and time for better performance. The biggest requirement of the feature selection method is to be able to sustain a multiclass high-dimensional dataset. By utilizing a limited number of features, the existence of the disease can be detected more accurately. Previously, real-world data could be utilized to diagnose a lot of diseases early on. When real-world datasets are utilized, the research focuses on preprocessing data; the finest feature selection method is identified and used to filter out inappropriate and unwanted features, and the accuracy of classification is also estimated.

4. DATASET DESCRIPTION

The "Heart Failure Clinical Record Dataset 2020" is utilized in this paper and could be attained from the UCI ML repository [15]. It has the health records of 299 HF patients. The patients consisted of 105 women and 194 men, all under 40–95 years of age. All 299 patients had left ventricular systolic failure. There were 13 features in the dataset that report medical, physical, and lifestyle data, as shown in Table 3. Some features are binary: anemia, hypertension, diabetes, gender, and smoking. For example, if the hematocrit levels are less than 36%, the hospital physician assumes that a patient has anemia.

Table 2. Dataset Feature Description

Feature	Description	Measurement	Range
Age	The age of the patient	Years	40 to 95
Anaemia	Decreased haemoglobin	Boolean	0, 1
High BP	If a patient has a high BP	Boolean	0, 1
Creatinine phosphokinase (CPK)	The amount of CPK in the blood	mcg/L	23 to 7861
Diabetes-mellitus	If the patient has Diabetes-mellitus	Boolean	0, 1
Ejection fraction	How many % well left or right ventricle pump blood with each heartbeat	Percentage	14 to 80
Sex	Female or male	Binary	0, 1
Platelets	Platelets in the blood	kiloplatelets/mL	25.01 to 850.0
Serum-creatinine	The amount of Serum-creatinine in the blood	mg/dL	0.50 to 9.40
Serum-sodium	The amount of Serum-sodium in the blood	mEq/L	114 to 148
Smokes	If the patient has smoking habits	Boolean	0, 1
Time	Follow up Time	Days	4 to 285
(Target) death event	If the patient dies in the follow-up time	Boolean	0, 1

5. METHODOLOGY

Coronary artery disease outcomes are very significant, especially in developing countries, where a lack of integrated diagnostic machines, physicians, and other assets can affect the treatment of heart disease patients and affect their prospects. Therefore, an accurate and rational assessment of CHD risk in patients is necessary to reduce their risk of acute heart disease and improve cardiovascular safety. Half of the people with CHD may die within 1-2 years of worrying about the costs of their treatment. Therefore, to reduce treatment costs, this work proposes the CSFB feature selection algorithm and the CMD algorithm for forecasting cardiovascular disease patient survival. The heart disease patient's survival prediction flow diagram that is based on the combination of CSFB and CMD algorithms is shown in Figure 2. CSFB algorithm is used to select optimal features and then the CMD algorithm is used to predict CHD patient survival using selected optimal features.

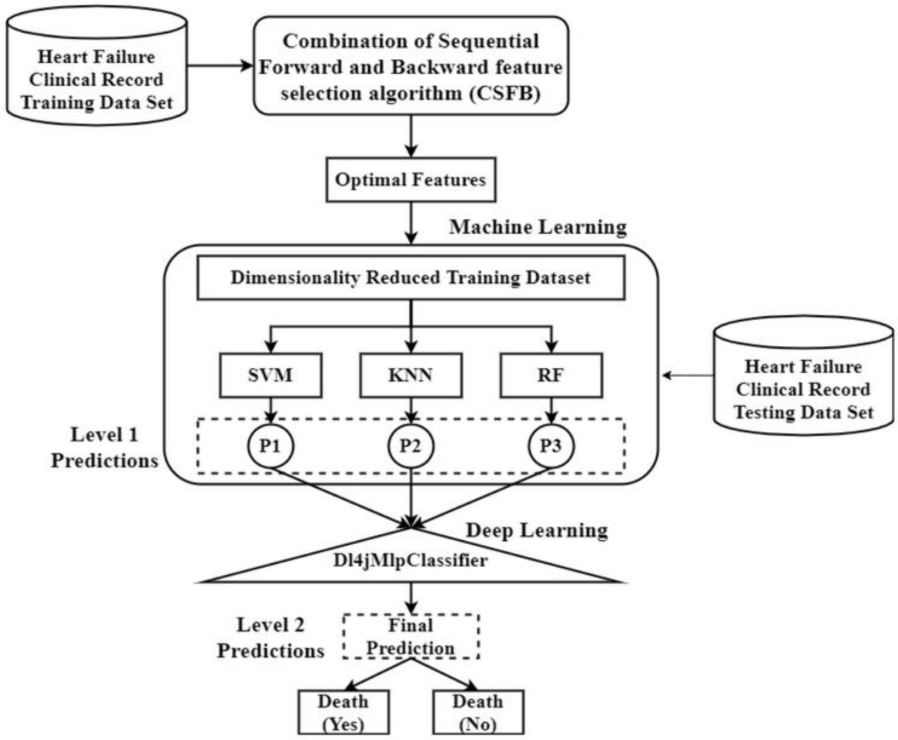

Figure 2: Proposed heart disease patient's survival prediction flow diagram.

5.1. Combination of Sequential Forward and Backward Feature Selection Algorithm

Feature selection is the identification of a subgroup of unique features, which can achieve higher quality classification results. Initially, decreasing the number of features will decrease overfitting. Instead, a better understanding of the features, their interaction, and response variables can be achieved. Presume that the whole set of features is F, here ∣F∣ ◎d is the features count in F. Feature selection is done from the subset of best features S ⊆ F where ∣S∣ ◎m reliable by a particular situation. m is the future count in S. It should be noted that there are feasible 2∣F∣ ◎ 1 subsets, creating a comprehensive analysis of the finest subset that is unfeasible because of the NP-hardness (nondeterministic polynomial-time hardness). Total realistic feature selection methods use particular heuristics to lead the exploratory procedure.

Sequential feature selection is the most familiar technique of feature selection. It has two mechanisms:

- A criterion is also known as an objective function, which tries to reduce the entire possible feature subgroups. General criteria are misclassification ratio (for classification techniques) and mean squared error (for regression techniques).
- A sequential exploration technique that attaches or eliminates features to and from the subset, respectively, of the candidate when assessing criteria. A complete evaluation of the criterion in the entire 2n subgroups of the n feature dataset is generally difficult; sequential searches go merely one way, constantly increasing or consistently decreasing the set of the candidates.

The technique contains two variations, namely 1) sequential forward selection (SFS) and 2) sequential backward selection (SBS). In SFS, additional features will continue to be attached to the empty candidate set until the criteria are reduced. The steps to perform SFS:

- Initially, the finest single feature is chosen (that is, utilizing a few criterion functions).
- After that, pairs of features are created by utilizing one of the residual features, and these finest and second finest pairs are chosen.
- Subsequently, triplets of features are created by utilizing one of the residual features, and these two finest features and the finest triplet are chosen.
- This process maintains until a predefined feature count is chosen.

SFS works finest when the optimal subset is small.

1. Sequential Backward Selection (SBS)

In SBS, features will be continuously eliminated from the entire candidate set until the elimination of extra features increases the criteria. Steps to perform SBS:

- Initially, the criterion function is calculated for the entire n features.
- After that, all features are removed, one at a time, and the criterion function is calculated for entire subsets with n◎ 1 characteristic. In addition, the unnecessary feature is also removed.
- Subsequently, all features among the residual n◎ 1 are removed, one at a time, and the unnecessary feature is removed to create a subset which has n ◎ 2 features.
- This process maintains a predefined feature count left.
- SBS works well when the optimum subset is big.

2. Combination of Sequential Forward and Backward Feature Selection

CSFB applies SFS and SBS at the same time:

- SFS is worked from the empty set.
- SBS is worked from the set.
- To assure that SFS along with SBS joins for a similar result:
- Features already chosen through SFS are not deleted through SBS.
- Features previously deleted through SBS are not added through SFS.

The following Algorithm 1 shows the proposed CSFB algorithm.

Algorithm 1: Combination of Sequential Forward and Backward feature selection

Input	**:**	Training Dataset (X)
Output	**:**	Dimensionality Reduced Training Dataset with Selected Features (DR)
Step 1	**:**	Start SFS with $Y_F = \{\emptyset\}$
Step 2	**:**	Start SBS with $Y_B = P$
Step 3	**:**	Choose the finest feature $p^+ = \arg\max J(Y_{Fq} + p)$ $p \notin Y_{Fq}$ $p \in Y_{Bq}$ $Y_{Fq+1} = Y_{Fq} + p^+$ J(YFq+1) = TrainingModule(YFq+1, X)
Step 4	**:**	Eliminate the bad feature

		$p^- = \arg\max J(Y_{Bq} - p)$ $p \in Y_{Bq}$ $p \notin Y_{Fq+1}$ $Y_{Bq+1} = Y_{Bq} - p^-$; $q = q + 1$ $J(YBq+1) = TrainingModule(YBq+1, X)$
Step 5	**:**	Go to Step 3
Step 6	**:**	SF = Extract standard features from Y_F and Y_B
Step 7	**:**	DR = Reduce the dimensionality of data based on selected features SF

5.2. Combination of ML And DL Algorithm (CMD)

Classification is the arrangement of things or organisms into orderly collections by using their similarities. Classification is further called taxonomy. An easy way to appreciate an instance of an email is by classifying emails as "spam" or "not spam." Conventionally, the ML method has been utilized for classification. This technique used mathematical concepts, for example, statistics, linear programming, neural networks, and DTs. However, recently, the DL method has gained fame because of its preeminence in accuracy when trained by enormous data. However, both ML and DL have both advantages and disadvantages. DL provides high accuracy, but ML gives lesser accuracy. Moreover, DL requires extensive data, but ML can train on more minor data.

Furthermore, DL takes more time for training, but ML takes less time. Therefore, the DL technique solves the disadvantage of the ML technique and similarly, the ML technique solves the disadvantage of the DL technique. Therefore, this paper proposes the CMD algorithm. SVM, KNN, and RF algorithms are used for ML, and the Dl4jMlpClassifier algorithm is used for DL in the CMD algorithm. In the CMD algorithm, the Dl4jMlpClassifier algorithm is used as the stacking classifier approach. A stacking classifier is an ensemble technique where the output from multiple classifiers is sent as an input to a metaclassifier for the job of the final classification. The stacking classifier technique could be an extremely competent method to execute a multiclassification issue. The individual classification techniques, commonly known as base learning, could be merged by building the metaclassifier for the result forecast task. It could be completed by stacking the results jointly from each classification algorithm and passing it as an input to the metaclassifier. In the CMD algorithm, SVM, KNN, and RF algorithms are used as base classifiers, and the Dl4jMlpClassifier classifier is utilized as metaclassifiers. The Weka software is utilized for the execution of SVM, KNN, RF, and Dl4jMlpClassifier algorithms with the CMD algorithm. The following Algorithm 2 shows the proposed CMD algorithm.

1. Support Vector Machine (SVM)

Algorithm 2: Combination of Machine learning and Deep learning (CMD)		
Input	**:**	Training Dataset (TrD), Testing Dataset (TeD)
Output	**:**	Heart Disease Patient Survival Prediction (SP) for TeD
Step 1	**:**	CR_{svm} ← Classify TrD based on SVM
Step 2	**:**	CR_{knn} ← Classify TrD based on KNN
Step 3	**:**	CR_{rf} ← Classify TrD based on RF
Step 4	**:**	P1 ← Predict TeD based on SVM with CR_{svm}
Step 5	**:**	P2 ← Predict TeD based on KNN with CR_{knn}
Step 6	**:**	P3 ← Predict TeD based on RF with CR_{rf}
Step 7	**:**	SP ← Predict TeD based on Dl4jMlpClassifier with P1, P2 and P3

The SVM technique is a supervised ML technique that uses arithmetical theory. SVM fundamentally works like the linear partition between two information points to recognize two classes in the multidimensional surroundings. SVM utilizes an extensive collection of nonlinear features, which is task autonomous. They contain an intelligent method for avoiding overfitting. They have a unique scheme for using numerous characteristics without the need for almost precise calculation which seems essential. The primary aim of this technique is to increase the edge between the classes and reduce the distance between the hyperplane points. Furthermore, SVM splits the data into two collections of vectors beneath the n-dimensional vector space. The SVM technique creates a hyperplane environment, with each component varying in a separated linear line. The hyperplane idea is provided for the information partition using the most extensive distance study to discover the classes. To decrease the error ratio, the most extensive edge classification algorithm is explained.

2. K-Nearest Neighbor

KNN is a consistent learning sequence computation. KNN computes class names that could forecast additional data; if the novel data is identical, the examples in the training dataset are similar. During the training stage, the KNN algorithm will record the data points. In the testing stage, the distance from the query data point to the training stage points is computed to categorize all points in the testing data. Different distances could be computed, however, the very famous is the Euclidean distance (ED). The KNN working could be explained based on the below steps:

- Pick the k number of neighbors.
- Compute the ED of picked k number of neighbors.
- Obtain the KNNs as stated by the computed ED.
- Among these k neighbors, calculate the number of instances in all groups.
- Allocate novel instances for that group in which the number of neighbors is high.

3. Random Forest (RF)

RF constructs a lot of distinctive selection trees during training. The summary of the anticipations of entire trees decides the last forecast, the technique of organization, or the standard anticipation of repetition. For choosing significant attributes, it is typically for entire trees in the "RF" stage. The full significance of elements in each tree is decided and also separated through the total number of trees.

$$\mathrm{RFfi}_i = \frac{\sum_{j \in \text{all trees}} \mathrm{normfi}_{ij}}{T}$$

Where i in RFfi_i is computed from the total number of trees in the RF algorithm, normfi_{ij} is the normalized attribute significance for i in tree j, and T is the tree count.

4. Dl4jMlpClassifier

Dl4jMlpClassifier is one of the DL classification algorithms that allow the formation of spontaneously deep feedforward neural networks, for example, CNN. Dl4jMlpClassifier is a core technique of WekaDeeplearning4j, constructed in a Weka package; therefore, it creates Deeplearning4j methods that are obtainable in the entire Weka environment. The Dl4jMlpClassifier could be together utilized for regression and classification by selecting suitable loss functions. In DL, a CNN is a class of DNN. They are further known as shift-invariant ANNs, namely, SIANN, using the disseminated weight structural design of the convolution kernels, which slide alongside inputs. The attributes present conversion equivalent responses called attribute maps. CNN's are regular versions of MLP. MLPs typically denote an entirely linked network, that is, all neurons at one layer are linked to entire neurons at the subsequent layer. The "full link" of these networks builds them prone to overfitting data. The usual method of regularization, or avoiding overfitting, is penalizing parameters through trimming connectivity. CNNs have various techniques for regularization; they benefit from the hierarchical pattern of information and assemble designs of rising difficulty, utilizing lesser and easier ways embossed on their filter.

6. Results

The proposed model has been built to filter patients who have survived the cardiovascular disease. For feature selection, the CSFB algorithm is utilized to choose significant features; it also exhibits classifiers for these chosen features. To predict heart disease patients' survival, the CMD algorithm is used. This algorithm has three well-known ML classifiers: SVM, KNN, and RF, and one DL classifier, namely, Dl4jMlpClassifier, to process the model and calculate evaluation measurements, for example, confusion matrix, accuracy and precision, recall and f-measure.

A confusion matrix is a table-like structure that contains real values and forecasted values that are referred to as true positive and true negative. It is explained in four fractions: the first is a true positive (TP), in which a value is forecasted as "yes," and it was actually "yes." The next one is false positive (FP), in which the value is predicted as "yes," but it is "no." The third one is a false negative (FN), in which a value was predicted as "no," but it was actually "yes." The last one is true negative (TN), in which the value was predicted as "no," but it was actually "no."

Accuracy

Accuracy is the fraction of the whole quantity of perfect predictions and the whole quantity of forecasts. Accuracy denotes the % of each correctly forecasted data point shown in (2).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Table 5 shows the accuracy comparison of SVM, KNN, RF, and CMD algorithms.

Table 5. Comparison of SVM, KNN, RF, and CMD Algorithms based on Accuracy

Algorithm	Accuracy (in %)
SVM	69.8
KNN	74.0
RF	75.0
CMD	99.3

Figure 3 shows the accuracy comparison for the heart disease dataset. Compared with other classification algorithms, the CMD algorithm provides the highest accuracy. It is because the CMD algorithm is used in the CSFB algorithm for optimal feature selection. This feature selection algorithm minimizes the dimensionality of the dataset. This procedure assists to boost the effectiveness of the CMD algorithm. Therefore, the CMD algorithm provides the highest accuracy.

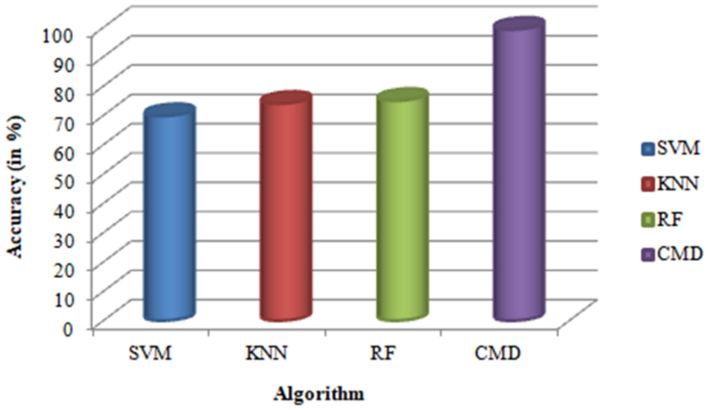

Figure 3: Accuracy comparison of SVM, KNN, RF, and CMD algorithm set

Precision

Precision is the ratio between the true positive and the entire positive. Precision is known as a measure of exactness or quality or positive predictive value, which is shown in (3).

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Table 6 demonstrates the comparison of the SVM, KNN, RF, and CMD algorithms based on precision.

Table 6. Precision Comparison of SVM, KNN, RF, and CMD Algorithms

Algorithm	Precision
SVM	0.714
KNN	0.608
RF	0.799
CMD	0.978

Figure 4 shows the comparison of precision for the heart disease dataset. Compared with other classification algorithms, the precision of the proposed CMD algorithm is large. Since the CMD algorithm is used, both ML and DL benefit. Therefore, it provides the highest precision.

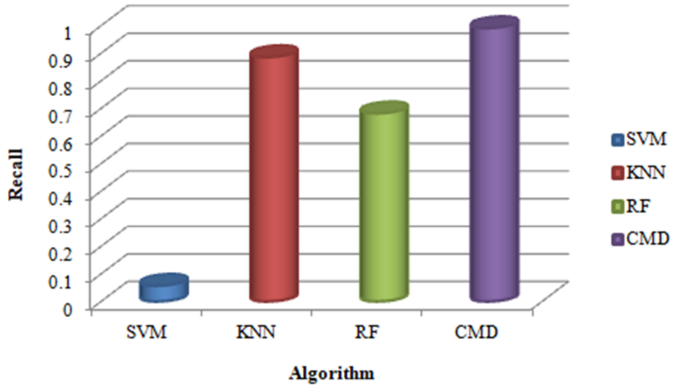

Figure 4: Precision comparison of SVM, KNN, RF, and CMD algorithms set.

Recall

The recall is a measure of accurate identification of true positives. The recall is also known as sensitivity, which is shown in (4).

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

Table 7 shows the recall comparison of SVM, KNN, RF, and CMD algorithms.

Table 7. Recall Comparison of SVM, KNN, RF, and CMD Algorithms

Algorithm	Recall
SVM	0.057
KNN	0.886
RF	0.682
CMD	0.991

Figure 5 shows the comparison of recall for the heart disease dataset. Compared with SVM, KNN, and RF, the recall of the CMD algorithm is high.

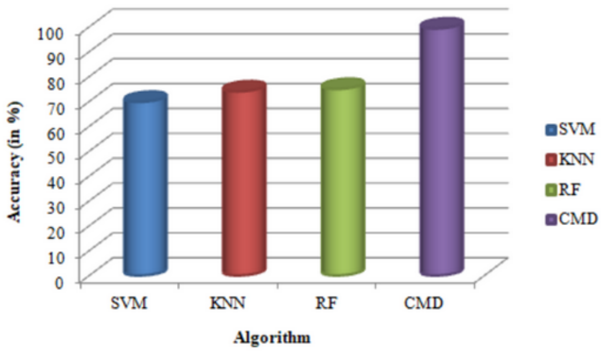

Figure 5: Comparison of recall for SVM, KNN, RF, and CMD algorithms.

F-Measure

The harmonic mean of precision and recall is called F-measure, which is shown in (5).

$$F - Measure = 2\,\frac{Precision.Recall}{Precision + Recall} \quad (5)$$

Furthermore, Table 8 demonstrates the F-measure comparison of SVM, KNN, RF, and CMD algorithms.

Table 7. F-Measure Comparison of SVM, KNN, RF, and CMD Algorithms

Algorithm	F-Measure
SVM	0.105
KNN	0.699
RF	0.889
CMD	0.989

Figure 6 shows the comparison of the F-measure for the heart disease dataset. Compared with other classification algorithms, the F-measure of the proposed CMD algorithm is high.

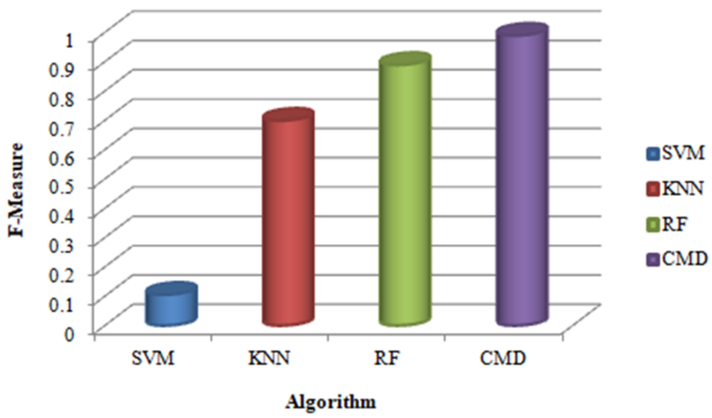

Figure 6: F-measure comparison of SVM, KNN, RF, and CMD algorithms set.

7. Conclusions

Health around the world is deteriorating gradually because of heart disease. A significant reason for death associated with the heart is that it is not detected early. Therefore, the clinical profession should give more importance to the prognosis of heart disease patient survival. This paper initiated a forecasting method by utilizing the CSFB feature selection algorithm with the CMD algorithm. The CSFB algorithm combines both types of sequential feature selection algorithms. As a result, it selected essential features from both types and provided the best and optimal features. Furthermore, SVM, KNN, and RF algorithms have been used for ML and Dl4jMlpClassifier algorithm has been used for DL in the CMD algorithm. In the CMD algorithm, the Dl4jMlpClassifier algorithm was used in the stacking classifier approach. The experimental results showed that, compared with existing classification algorithms, the CMD algorithm provides high accuracy, precision, recall, and f-measure as 99.3%, 0.978, 0.991, and 0.989, correspondingly. It concluded that the CMD algorithm predicts heart disease patients' survival efficiently. However, the CMD algorithm combines both ML and DL, which takes up more time and space. In the future, an enhanced CMD algorithm (ECMD) algorithm will be required to deal with this problem.

REFERENCES

[1] Gonsalves, A. H., Thabtah, F., Mohammad, R. M. A., & Singh, G. (2019, July). Prediction of coronary heart disease using machine learning: An experimental analysis. In Proceedings of the 2019 3rd International Conference on Deep Learning Technologies (pp. 51-56).

[2] Centers for Disease Control and Prevention. 2021. Heart Disease Facts | cdc.gov. [online] Available at: <https://www.cdc.gov/heartdisease/facts.htm> [Accessed 20 November 2021].

[3] Aggrawal, R. I. T. U., & Pal, S. A. U. R. A. B. H. (2021). Multi-Machine Learning Binary Classification, Feature Selection and Comparison Technique for Predicting Death Events Related to Heart Disease. International Journal of Pharmaceutical Research, 13(1).

[4] Moss, A. J., Williams, M. C., Newby, D. E., & Nicol, E. D. (2017). The updated NICE guidelines: cardiac CT as the first-line test for coronary artery disease. Current cardiovascular imaging reports, 10(5), 15.

[5] Timmis, A., Townsend, N., Gale, C. P., Torbica, A., Lettino, M., Petersen, S. E., ... & Vardas, P. (2020). European Society of Cardiology: cardiovascular disease statistics 2019. European heart journal, 41(1), 12-85.

[6] Nalluri, S., Saraswathi, R. V., Ramasubbareddy, S., Govinda, K., & Swetha, E. (2020). Chronic heart disease prediction using data mining techniques. In Data engineering and communication technology (pp. 903-912). Springer, Singapore.

[7] Seetharam, K., Shrestha, S., & Sengupta, P. P. (2019). Artificial intelligence in cardiovascular medicine. Current treatment options in cardiovascular medicine, 21(5), 1-14.

[8] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 7, 81542-81554.

[9] Tanikawa, C., & Chonho, L. (2021). Machine Learning for Facial Recognition in Orthodontics. In Machine Learning in Dentistry (pp. 55-65). Springer, Cham.

[10] Sharma, R., Kaushik, B., & Gondhi, N. (2020, March). Character recognition using machine learning and deep learning a survey. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 341-345). IEEE.

[11] Roy, S., Sharma, M., & Singh, S. K. (2019, October). Movie Recommendation System Using Semi-Supervised Learning. In 2019 Global Conference for Advancement in Technology (GCAT) (pp. 1-5). IEEE.

[12] Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. Social studies of science, 48(1), 25-56.

[13] Ramalingam, V. V., Dandapath, A., & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. International Journal of Engineering & Technology, 7(2.8), 684-687.

[14] Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. Neurocomputing, 300, 70-79.

[15] De Silva, K., Jönsson, D., & Demmer, R. T. (2020). A combined strategy of feature selection and machine learning to identify predictors of prediabetes. Journal of the American Medical Informatics Association, 27(3), 396-406.

[16] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. Nature medicine, 25(1), 24-29.

[17] Mishra, D., Naik, B., Sahoo, R. M., & Nayak, J. (2020, February). Deep recurrent neural network (Deep-RNN) for classification of non-linear data. In Springer (pp. 207-215).

[18] Ghasemi, F., Mehridehnavi, A., Perez-Garrido, A., & Perez-Sanchez, H. (2018). Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks. Drug Discov. Today, 23(10), 1784-1790.

[19] Sewak, M., Sahay, S. K., & Rathore, H. (2018, June). Comparison of deep learning and the classical machine learning algorithm for malware detection. In 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (pp. 293-296). IEEE.

[20] Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2020). A survey of deep learning and its applications: a new paradigm to machine learning. Archives of Computational Methods in Engineering, 27(4), 1071-1092.

[21] Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mobile Information Systems, 2018.

[22] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542-81554.

[23] Pawlovsky, A. P. (2018, January). An ensemble-based on distances for a kNN method for heart disease diagnosis. In 2018 International conference on electronics, information, and communication (ICEIC) (pp. 1-4). IEEE.

[24] Vijayashree, J., & Sultana, H. P. (2018). A machine learning framework for feature selection in heart disease classification using improved particle swarm optimization with support vector machine classifier. Programming and Computer Software, 44(6), 388-397.

[25] Magesh, G., & Swarnalatha, P. (2021). Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction. Evolutionary Intelligence, 14(2), 583-593.

[26] Gárate-Escamila, A. K., El Hassani, A. H., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. Informatics in Medicine Unlocked, 19, 100330.

[27] Aggrawal, R., & Pal, S. (2020). Sequential feature selection and machine learning algorithm-based patient's death events prediction and diagnosis in heart disease. SN Computer Science, 1(6), 1-16.

[28] Gokulnath, C. B., & Shantharajah, S. P. (2019). An optimized feature selection based on a genetic approach and support vector machine for heart disease. Cluster Computing, 22(6), 14777-14787.

[29] Bashir, S., Khan, Z. S., Khan, F. H., Anjum, A., & Bashir, K. (2019, January). Improving heart disease prediction using feature selection approaches. In 2019 16th international bhurban conference on applied sciences and technology (IBCAST) (pp. 619-623). IEEE.

[30] Ahmed, H., Younis, E. M., Hendawi, A., & Ali, A. A. (2020). Heart disease identification from patients' social posts, machine learning solution on Spark. Future Generation Computer Systems, 111, 714-722.

[31] Khourdifi, Y., & Bahaj, M. (2019). Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. International Journal of Intelligent Engineering and Systems, 12(1), 242-252.

[32] Le, H. M., Tran, T. D., & Van Tran, L. A. N. G. (2018). Automatic heart disease prediction using feature selection and data mining techniques. Journal of Computer Science and Cybernetics, 34(1), 33-48.

[33] Usman, A. M., Yusof, U. K., & Naim, S. (2018). Cuckoo inspired algorithms for feature selection in heart disease prediction. International Journal of Advances in Intelligent Informatics, 4(2), 95-106.

[34] Kohli, P. S., & Arora, S. (2018, December). Application of machine learning in disease prediction. In 2018 4th International conference on computing communication and automation (ICCCA) (pp. 1-4). IEEE.

[35] Latha, C. B. C., & Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Informatics in Medicine Unlocked, 16, 100203.

[36] Uyar, K., & İlhan, A. (2017). Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks. Procedia computer science, 120, 588-593.

[37] Yazid, M. H. A., Satria, M. H., Talib, S., & Azman, N. (2018, October). Artificial neural network parameter tuning framework for heart disease classification. In 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI) (pp. 674-679). IEEE.

[38] Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on deep ensemble learning and feature fusion. Information Fusion, 63, 208-222.

[39] Khan, M. A. (2020). An IoT framework for heart disease prediction based on an MDCNN classifier. IEEE Access, 8, 34717-34727.

[40] Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An automated diagnostic system for heart disease prediction based on x2 statistical model and optimally configured deep neural network. IEEE Access, 7, 34938-34945.

[41] Darmawahyuni, A., Nurmaini, S., & Firdaus, F. (2019). Coronary heart disease interpretation based on deep neural network. Computer Engineering and Applications Journal, 8(1), 1-12.

[42] Kong, Y., & Yu, T. (2018). A graph-embedded deep feed-forward network for disease outcome classification and feature selection using gene expression data. Bioinformatics, 34(21), 3727-3737.

[43] Chen, T. M., Huang, C. H., Shih, E. S., Hu, Y. F., & Hwang, M. J. (2020). Predicting cardiac arrhythmias by a Challenge-best deep learning neural network model. science, 23(3), 100886.

[44] Krishnan, P. T., Balasubramanian, P., & Umapathy, S. (2020). Automated heart sound classification system from unsegmented PCG using deep neural network. Physical and Engineering Sciences in Medicine, 1-11.

[45] Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. Expert Systems with Applications, 159, 113408.

[46] Zhang, L., Fabbri, D., Upender, R., & Kent, D. (2019). Automated sleep stage scoring of the sleep heart health study using deep neural networks. Sleep, 42(11), zsz159.

[47] Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine, 25(1), 65-69.

[48] Cai, W., Chen, Y., Guo, J., Han, B., Shi, Y., Ji, L., ... & Luo, J. (2020). Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network. Computers in biology and medicine, 116, 103378.

[49] Kim, J. K., & Kang, S. (2017). Neural network-based coronary heart disease risk prediction using feature correlation analysis. Journal of healthcare engineering, 2017.

[50] Baviskar, V., Verma, M., & Chatterjee, P. (2020, December). A Model for Heart Disease Prediction Using Feature Selection with Deep Learning. In International Advanced Computing Conference (pp. 151-168). Springer, Singapore.

[51] Heart Failure Clinical Record Dataset 2020 https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+records

[52] Kumar, R. L., Khan, F., Kadry, S., & Rho, S. (2021). A Survey on blockchain for industrial Internet of Things. Alexandria Engineering Journal.

[53] Khan, F., Siva Prasad, B. V. V., Syed, S. A., Ashraf, I., & Ramasamy, L. K. (2021). An Efficient, Ensemble-Based Classification Framework for Big Medical Data. Big Data.

[54] Pooranam, N., Ignisha Rajathi, G., Lakshmana Kumar, R., & Vignesh, T. (2021). Decision Support Mechanism to Improve a Secured System for Clinical Process Using Blockchain Technique. In Internet of Things, Artificial Intelligence and Blockchain Technology (pp. 241-258). Springer, Cham.

[55] Reddy, G. T., Reddy, M., Lakshmanna, K., Rajput, D. S., Kaluri, R., & Srivastava, G. (2020). Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis. Evolutionary Intelligence, 13(2), 185-196.

[56] Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., Ra, I. H., & Alazab, M. (2020). Early detection of diabetic retinopathy using PCA-firefly based deep learning model. Electronics, 9(2), 274.

[57] Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., & Srivastava, G. (2020). Deep neural networks to predict diabetic retinopathy. Journal of Ambient Intelligence and Humanized Computing, 1-14.

[58] Reddy, G. T., Reddy, M. P. K., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). Analysis of dimensionality reduction techniques on big data. IEEE Access, 8, 54776-54788.