GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

PREDICTION OF HEART DISEASE USING MACHINE LEARNING

A Project Report for Evolution 3 of Project - 2

Submitted by

ASHANSHU GAUTAM
(1613101185)

in partial fulfillment for the award of the degree

of

BACHELORE OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

Under the Supervision of

Mr. ANKIT KUMAR

Professor

APRIL/MAY - 2020

SCHOOL OF COMPUTING AND SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

Certified that this project report "PREDICTION OF HEART DISEASE USING MACHINE LEARNING" is the bonafide work of "ASHANSHU GAUTAM (1613101185)" who carried out the project work under my supervision.

SIGNATURE OF HEAD   **SIGNATURE OF SUPERVISION**

Dr. PRASHANT JOHRI   Mr. Ankit kumar

Professor & Dean,   **Professor**
School of Computing Science &   **School of computer science &**

Engineering   **Engineering**   TABLE OF

CONTENTS

Abstract.

Heart disease is one of the most critical human diseases in the world and affects human life very badly. In heart disease, the heart is unable to push the required amount of blood to other parts of the body. Accurate and on time diagnosis of heart disease is important for heart failure prevention and treatment. The diagnosis of heart disease through traditional medical history has been considered as not reliable in many aspects. To classify the healthy people and people with heart disease, noninvasive-based methods such as machine learning are reliable and efficient. In the proposed study, we developed a machine-learning-based diagnosis system for heart disease prediction by using heart disease dataset. We used seven popular machine learning algorithms, three feature selection algorithms, the cross-validation method, and seven classifiers performance evaluation metrics such as classification accuracy, specificity, sensitivity, Matthews' correlation coefficient, and execution time. The proposed system can easily identify and classify people with heart disease from healthy people. Additionally, receiver optimistic curves and area under the curves for each classifier was computed. We have discussed all of the classifiers, feature selection algorithms, preprocessing methods, validation method, and classifiers performance evaluation metrics used in this paper. The performance of the proposed system has been validated on full features and on a reduced set of features. The features reduction has an impact on classifiers performance in terms of accuracy and execution time of classifiers. The proposed machine-learning-based decision support system will assist the doctors to diagnosis heart patients efficiently.

Introduction.

Shrinkage and Selection Operator (LASSO), were also The heart disease (HD) has been considered as one of the complex and life deadliest human diseases in the world. In this disease, usually the heart is unable to push the required amount of blood to other parts of the body to fulfill the normal functionalities of the body, and due to this, ultimately the heart failure occurs [1]. The rate of heart disease in the United States is very high [2]. The symptoms of heart disease include shortness of breath, weakness of physical body, swollen feet, and fatigue with related signs, for example, elevated jugular venous pressure and peripheral edema caused by functional cardiac or noncardiac abnormalities [3]. The investigation techniques in early stages used to identify heart disease were complicated, and its resulting complexity is one of the major reasons that affect the standard of life [4]. The heart disease diagnosis and treatment are very complex, especially in the developing countries, due to the rare availability of diagnostic apparatus and shortage of physicians and others resources which affect proper prediction and treatment of heart patients [5]. The accurate and proper diagnosis of the heart disease risk in patients is necessary for reducing their associated risks of severe heart issues and improving security of heart [6]. The European Society of Cardiology (ESC) reported that 26 million adults worldwide were diagnosed with heart disease and 3.6 million were diagnosed every year. Approximately 50% of heart disease people suffering from HD die within initial 1-2 years, and concerned costs of heart disease management are approximately 3% of health-care financial budget [7].

classification and obtained a classification accuracy of 77%. The Cleveland dataset used [14] with global evolutionary approaches and achieved high prediction performance in accuracy. The study used feature selection methods for selection of features. Therefore, the classification performance of the approach depends on selected features. Gudadhe et al. [15] used multilayer perceptron (MLP) and support vector machine algorithms for heart disease classification. They proposed classification system and obtained accuracy of 80.41%. Kahramanli and Allahverdi [16] designed a heart disease classification system used a hybrid technique in which a neural network integrates a fuzzy neural network and artificial neural network. And the proposed classification system achieved a classification accuracy of 87.4%. Palaniappan and Awang [17] designed an expert medical diagnosing heart disease system and applied machine learning techniques such as Naive Bayes, decision tree, and ANN in the system. The Naive Bayes predictive model obtained performance accuracy 86.12%. The second best predictive model was ANN which obtained an accuracy and achieved a classification accuracy of 88.89%. Moreover, the proposed system could be easily deployed in healthcare information systems. Das et al. [19] proposed an ANN ensemble-based predictive model that diagnoses the heart disease and used statistical analysis system enterprise miner 5.2 with the classification system and achieved 89.01% accuracy, 80.09% sensitivity, and 95.91% specificity. Jabbar et al. [20] designed a diagnostic system for heart disease and used machine learning classifier multilayer perceptron ANN-driven back propagation learning algorithm and feature selection algorithm. The proposed system gives excellent performance in terms of accuracy. In order to diagnose heart

disease, an integrated decision support medical system based on ANN and Fuzzy AHP were designed by the authors in [12] which utilizes machine learning algorithm, artificial neural network, and Fuzzy analytical hierarchical processing. Their proposed classification system achieved a classification accuracy of 91.10%.

4

Material and Methods

2.1. Dataset. The "Cleveland heart disease dataset 2016" is used by various researchers [13] and can be accessed from online data mining repository of the University of California, Irvine. This dataset was used in this research study for designing machine-learning-based system for heart disease diagnosis. The Cleveland heart disease dataset has a sample size of 303 patients, 76 features, and some missing values. During the analysis, 6 samples were removed due to missing values in feature columns and leftover samples size is 297 with 13 more appropriate independent input features, and target output label was extracted and used for diagnosing the heart disease. The target output label has two classes in order to represent a heart patient or a normal subject. Thus, the extracted dataset is of 297*13 features matrix. The complete information and description of 297 instances of 13 features of the dataset is given in Table 1.

2.2. Methodology of the Proposed System. The proposed system has been developed with the aim to classify people with heart disease and healthy people. The performances of different machine learning predictive models for heart disease diagnosis on full and selected features were tested. Feature selection algorithms such as Relief, mRMR, and LASSO were used to select important features, and on these selected features, the performance of the classifiers was tested. The Cleveland heart disease dataset has been implemented in several studies [13] and is used in our study. The popular machine learning classifiers logistic regression, K-NN, ANN, SVM, DT, and NB were used in the system. The model's validation and performance evaluation metrics were computed. The methodology of the proposed system structured into five stages including (1) preprocessing of dataset, (2) feature selection, (3) cross validation method, (4) machine learning classifiers, and (5) classifiers' performance evaluation methods. Figure 1 shows the framework of the proposed system.

2.2.1. Data Preprocessing. The preprocessing of data is necessary for efficient representation of data and machine learning classifier which should be trained and tested in an effective manner. Preprocessing techniques such as removing of missing values, standard scalar, and Min Max Scalar have been applied to the dataset for effective use in the classifiers. The standard scalar ensures that every

feature has the mean 0 and variance 1, bringing all features to the same coefficient. Similarly, in Min Max Scalar shifts the data such that all features are between 0 and 1. The missing values feature row is just deleted from the dataset. All these data preprocessing techniques were used in this research.

2.2.2. Feature Selection Algorithms. Feature selection is necessary for the machine learning process because sometimes irrelevant features affect the classification performance of the machine learning classifier. Feature selection improves the classification accuracy and reduces the model execution time. For feature selection in our system, we used three well-known FS algorithms and these algorithms select important features.

In order to classify two classes 0 and 1, a hypothesis $h(\theta)$ $\theta^T \chi$ will be designed and threshold classifier output is $h\theta_{(\chi)}$

(1)   Relief Feature Selection Algorithm. Relief is a feature selection algorithm [21], which assigns weights to all the features in the dataset and these weights can be updated with passage of time. The important features to target have great weights value, and the remaining features have small weights. Relief uses the same techniques as in K-NN that determines the weights of features (see Algorithm 1) [22].

The pseudocode of Relief algorithm, the Relief algorithm iterated through m random training instances (R_k), was selected without replacement, where m is parameter. For each k, R_k is the "target" instance and the feature score vector W is updated [23].

 2). Minimal-Redundancy-Maximal-Relevance Feature Selection Algorithm. The mRMR chooses those features that are related to the target label. These selected features might be redundant variables which must be handled. The Heuristic search method is used in mRMR and selects optimum features that have maximum relevance and minimum redundancy. It checks one feature at a cycle and computes pairwise redundancy. The mRMR does not take care of the joint association of features [24]. The pseudocode mRMR algorithm is described in [25]. In this algorithm, main computation of mutual information (MI) between two features is computed. This function is calculated between each pair of features instead of many pairs of features; being irrelevant to the last result, mRMR is not suitable for large domain feature selection problems (see Algorithm 2).

(2)   Least Absolute Shrinkage and Selection Operator. Least absolute shrinkage and selection operator select features are based on updating the absolute value of features coefficient. Some coefficients values of features become zero, and these zero coefficients features are eliminated from features subset. The LASSO performs excellently with low coefficients feature values. The features having high values of coefficients will be included in selected feature subsets. In LASSO, some irrelevant features may be selected and include a subset of selected feature [26].

2.2.3. Machine Learning Classifiers. In order to classify the heart patients and healthy people, machine learning classification algorithms are used. Some popular classification algorithms and their theoretical background are discussed briefly in this paper.

(1) Logistic Regression. A logistic regression is a classification algorithm [27–29]. For binary classification problem, in order to predict the value of predictive variable y when $y \in [0, 1]$, 0 is negative class and 1 is positive class. It also uses multiclassification to predict the value of y when $y \in [0, 1, 2, 3]$.

at 0.5. If the value of hypothesis $h\theta_{(x)} \geq 0.5$, it will predict y 1 which mean that the person has heart disease

6

Table:

S. no.	Feature name	Feature code	Description	Domain of values (minmax)
1	Age	AGE	Age in years	$30 < age_1 < 77$
2	Sex	SEX		0
3	Type of chest pain	CPT	4123nonanginal pain Female a typical angina a typical angina asymptomatic Male	1 2
4	Resting blood pressure	RBP	Fasting blood sugar$_{false)}$>120 mg/dl true;	94-200
5	Fasting blood sugar >120 mg/dl	SCH	0 mm Hg admitted at the hospital	120–564
6	Serum cholesterol	FBS	In mg/dl	1
7	Resting electrocardiographic results	RES	0 normal	0 0
8	Maximum heart rate achieved	MHR	12 having ST-Thypertrophy—	1 2 71–202
9	Exercise-induced angina	EIA	1_0 yes$_{no}$	0 1
10	Old peakexercise relative to rest ST depression induced by	OPK	—	0–6.2

11	PES	Slope of the peak exercise ST segment	1 $_2$up sloping $_{flat}$		1
			3 down sloping		2
					3
12	VCA	Number of major vessels (0–3) colored by fluoroscopy	—		0
					1

———

Figure 1: A hybrid intelligent system framework
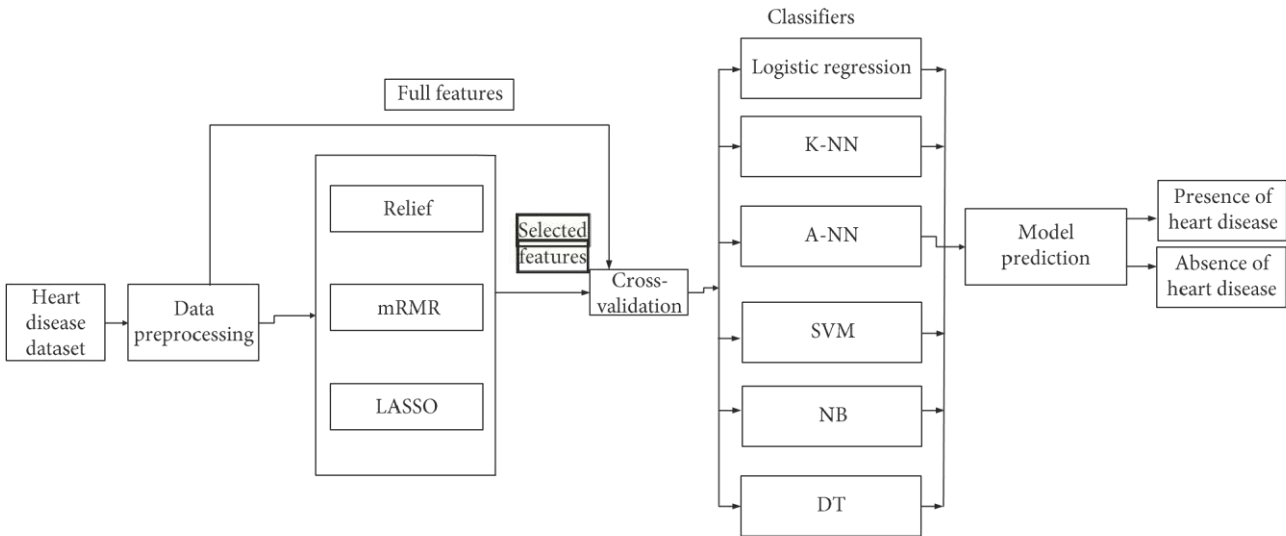
predicting heart disease. and if value of $\leq h\theta(x\leq)< 0.5$, then predict y 0 which shows $h\theta$-(zx) $g\theta\tau X,$ $-z$ (1) that the person is healthy.

condition 0Hence, the prediction of logistic regression under the $h\theta(x)$ 1 is done. Logistic regression sigmoid where written as follows:Similarly, the logistic regression cost function can be $g(z)$ $1/(1 + x)$ and $h\theta(x)$ $1/(1 + x)$. function can be written as follows:

RELIEF Algorithm

    Require: for each training instance set S, a vector of feature values and the class value n ← number of training instances a ← number of features

Parameter: m ← number of random training instances out of n used to update W

    Initialize all feature weights $W[A]$: 0.0

      For k: 1 to m do

Randomly select a "target" instance R_k

    Find a nearest hit "H" and nearest miss (instances)

        For A: 1 to a do

        $W[A]$: $W[A] - \text{diff}(A, R_k, H)/m + \text{diff}(A, R_k, M)/m$ **End for**

    End for

Return the weight vector W of feature scores that compute the quality of features

2) Support Vector Machine. The SVM is a machine learning classification algorithm which has been mostly used for classification problems [30–32]. SVM used a maximum margin strategy that transformed into solving a complex quadratic programming problem. Due to the high performance of SVM in classification, various applications widely applied it [4, 33].

In a binary classification problem, the instances are separated with a hyperplane $w^T x + b$ 0, where w and d are dimensional coefficient vectors, which are normal to the hyperplane of the surface, b is offset value from the origin, and x is data set values. The SVM gets results of w and b. w can be solved by introducing Lagrangian multipliers in the linear case. The data points on borders are called support vectors. The solution of w can be written as w $\sum^n_{l1} \alpha_i y_i x_i$, where n is the number of support vectors and y_i are target labels to x. The value of w and b are calculated, and the linear discriminant function can be written as follows:

$$g(x)\ \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i x_i^T x + b \right).\qquad (3)$$

The nonlinear scenario, for kernel trick and decision function, can be written as follows:

$$g(x)\ \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K\left(x_i, x\right) + b \right).\qquad (4)$$

The positive semidefinite functions obey Mercer's condition as kernel functions [32].

(3)   Naive Bayes. The NB is a classification supervised learning algorithm. It is based on conditional probability theorem to determine the class of a new feature vector. The NB uses the training dataset to find out the conditional probability value of vectors for a given class. After computing the probability conditional value of each vector, the new vectors class is computed based on its conditionality probability. NB is used for text-concerned problem classification [34].

(4)   Artificial Neural Network. The artificial neural network is a supervised machine learning algorithm [35] and is a mathematical model that integrates neurons that pass messages. The ANN has three components including inputs, outputs, and transfer functions. The input units take extraordinary values and weights, which are modified during the training process of the network. The output of the artificial neural network is calculated for the known class; the weight is recomputed using the error margin between the output of predicted and actual class. ANN is designed by the integration of neurons. This different combination of neurons from different structures is just like multilayer perception [36].

(5)   Decision Tree Classifier. A decision tree is a supervised machine learning algorithm [35, 37]. A decision tree shape is just a tree where every node is a leaf node or decision node. The techniques of the decision tree are simple and easily understandable for how to take the decision. A decision tree contained internal and external nodes linked with each other. The internal nodes are the decision-making part that makes a decision and the child node to visit the next nodes. The leaf node on the other hand has no child nodes and is associated with a label.

(6)   K-Nearest Neighbor. K-NN is a supervised learning classification algorithm. K-NN algorithm [35] predicts the class label of a new input; K-NN utilizes the similarity of new input to its inputs samples in the training set. If the new input is same the samples in the training set. The K-NN classification performance is not good. Let (x, y) be the training observations and the learning function $h: X \longrightarrow Y$, so that given an observation x, $h(x)$ can determine y value.

Validation Method of Classifiers. We used k-fold cross-validation (CV) method and four performance evaluation metrics in this research paper. The details are given in following subsections:

(1) K-Fold Cross-Validation. In k-fold cross-validation, the data set is divided into k equal size of parts, in which $k − 1$ groups are used to train the classifiers and remaining part is used for checking outperformance in each step. The process of validation is repeated k times. The classifier performance is computed based on k results. For CV, different values of k are selected. In our experiment, we used k 10 because its performance is good. In the 10-fold CV process, 90% data were used for training and 10% data were used for testing purpose. The process was repeated 10 times for each fold of process, and all instances in the training and test groups were randomly divided over the whole dataset prior to selection training and testing new sets for the new cycle. Lastly, at the end of the 10fold process, averages of all performance metrics are computed.

Performance Evaluation Metrics. In order to check the performance of the classifiers, various performance evaluation metrics were used in this research. We used confusion matrix, every observation in the testing set is predicted in exactly one box. It is 2×2 matrix because there are 2 repose classes. Moreover, it gives two types of correct prediction of the classifier and two types of classifier of incorrect prediction. Table 2 shows the confusion matrix.

Table 2: Confusion matrix.

From confusion matrix, we compute the following:

TP: predicted output as true positive (TP), we concluded that the HD subject is correctly classified and subjects have heart disease.

TN: predicted output as true negative (TN), we concluded that a healthy subject is correctly classified and the subject is healthy.

FP: predicted output as false positive (FP), we concluded that a healthy subject is incorrectly classified that they do have heart disease (a type 1 error).

FN: predicted output as false negative (FN), we concluded that a heart disease is incorrectly classified that the subject does not have heart disease as the subject is healthy (a type 2 error).

1 shows that positive case means diseased, and 0 shows that a negative case means healthy.

Classification accuracy: accuracy shows the overall performance of the classification system as follows:

$$\text{classification accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%.$$

Classification error: it is the overall incorrect classification of the classification model which is calculated as follows:

$$\text{classification error} = \frac{FP + FN}{TP + TN + FP + FN} \times 100\%. \quad (6)$$

Sensitivity: it is the ratio of the recently classified heart patients to the total number of heart patients. The sensitivity of the classifier for detecting positive instances is known as "true positive rate." In other words, we can say that sensitivity (true positive fraction) confirms that if a diagnostic test is positive and the subject has the disease. It can be written as follows:

$$\text{Sensitivity(Sn)/recall/true positive rate} = \frac{TP}{TP + FN} \times 100\%.$$

7). *Specificity*: a diagnostic test is negative and the person is healthy and is mathematically written as follows:

$$\text{specificity(Sp)} = \frac{TN}{TN + FP} \times 100\%. \quad (8)$$

Precision: the equation of precision is given as follows:

$$\text{precision} = \frac{TP}{TP + FP} \times 100\%. \quad (9)$$

MCC: it represents the prediction ability of a classifier with values between [−1, +1].

If the value of the MCC classifier is +1, this means the classifier predictions are ideal. −1 indicates that classifiers produce completely wrong predictions. The MCC value near to 0 means that the classifier generates random predictions. The mathematical equation of MCC is as follows:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \times 100\%.$$

(1) ROC and AUC. The receiver optimistic curves analyze the prediction capability of the machine learning classifiers used for classification. ROC analysis is a graphical-based representation which compares the "true positive rate" and "false positive rate" in the classification results of machine learning algorithm. AUC characterizes the ROC of a classifier. The larger the value of AUC .

Experimental Results and Discussion.

This section of the paper involved the discussion on the classification models and outcomes from different perspectives. First, we checked the performance of different machine learning algorithms such as logistic regression, k-nearest Neighbor, artificial neural network, support vector machine, Naive Bayes, and decision tree on Cleveland heart disease dataset on full features. In the second, we used feature selection algorithm Relief, mRMR, and LASSO for important features selection. In third classifiers, performances were checked on selected features. Also, the k-fold cross-validation method was used. In order to check the performance of classifiers performance evaluation metrics were applied. All features were normalized and standardized before applying to classifiers. All computations were performed in Python on an Intel(R) Core™ i5 -2400CPU @3.10GHz PC.

3.1. Result of Selected Features by Relief Feature Selection Algorithm. Relief [38], FS algorithm, selects important features on the basis of features weight. The most important 6 features were selected by Relief are given in Table 3. The rank of features on which the features are selected is shown in Figure 2. According to the results, the most important features for the diagnosis of heart disease are THA and EIA. We performed experiments on different numbers of selected features but the performances of classifiers on 6 features were very good that we only reported the performance of classifiers on 6 features in our simulation results. Additionally, only six important feature information and descriptions are tabulated in the paper. Table 3 shows the important selected features.
Figure 2 shows the ranking of important features by Relief.

3.2. Result of Selected Features by mRMR Feature Selection Algorithm. The selected important 6 features by mRMR FS based on mutual information are represented in Table 4. Also, Figure 3 shows the important features rank. In scores graph, chest pain is an important feature for heart disease prediction. We performed experiments on different numbers of selected

Table 3: Features selected by Relief algorithm and their ranking.

Order	Feature	Feature name	Feature code	Scores
1	13	Thallium scan	THA	0.247
2	9	Exercise-induced angina	EIA	0.227
3	3	Type of chest pain	CPT	0.217
4		Slope of the peak exercise ST segment		
			PES	0.131
	11			
5	12	Number of major vessels (0–3) colored by fluoroscopy	VCA	0.128

features but the performances of classifiers on 6 features were very good. Therefore, we only reported the performance of classifiers on 6 features in our simulation results. Table 4 shows important selected features by mRMR FS algorithm.

Figure 3 shows the important features selected by mRMR.

3.3. Result of Selected Features by LASSO Features Selection Algorithm. The LASSO selects highly related features to target as true and the remainder as false. The LASSO ranks the important features. In Table 5, the six important features are listed because the classifiers performances were excellent on these features. Table 5 shows the important selected features.

Figure 4 shows the important features selected by LASSO FS algorithm.

The important features score are presented in Figure 4 with features scores. These three tables show the important features for the diagnosis of heart disease. Moreover, FBS has a low score in important features scores so it means that FBS features have no influence on the prediction of heart disease, and additionally, three feature selection algorithms have not been selected for heart disease diagnosis which has been shown in Figures 2–4, respectively.

3.4. Results of K-Fold Cross-Validation for Classifiers Performance on Full Features (n 13). In this experiment, the full features of the dataset were checked on seven machine learning classifiers with 10-fold cross-validation methods. In 10-fold CV, 90% was used for training the classifiers and only 10% was tested. Finally, the average metrics of 10-fold methods were computed. Moreover, different parameters values were passed through classifiers. Table 6 describes the 10-fold cross-validation results of seven classifiers with full features.

In Table 6, the logistic regression showing good performance that has 84% classification accuracy, 85% specificity, 83% sensitivity, 89% MCC, and 84% AUC. The specificity value of logistic regression was 85% showing the probability that a diagnostic test was negative, and the person does not have the heart disease. Moreover, 83% sensitivity shows the probability that the diagnostic test positive and MCC was 89%.

For the K-NN classifier, we performed experiments with different values of k 1, 3, 5, 9, and 13. However, at k 9, the
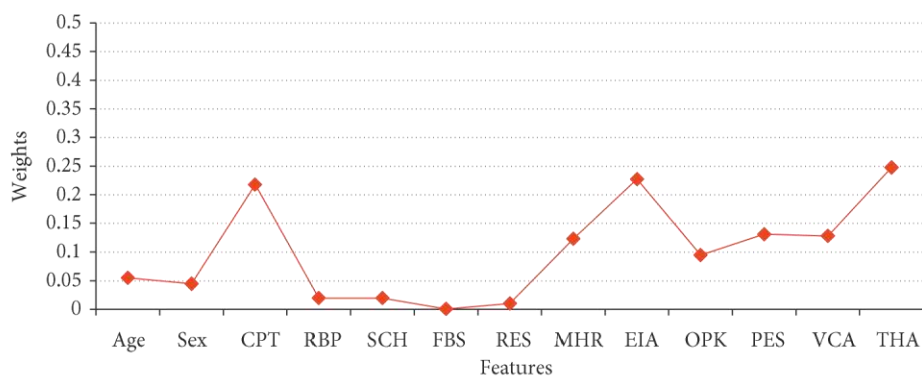

Figure 2: Important features selected by Relief FS.

Table 4: Features selected by mRMR algorithm and their ranking.

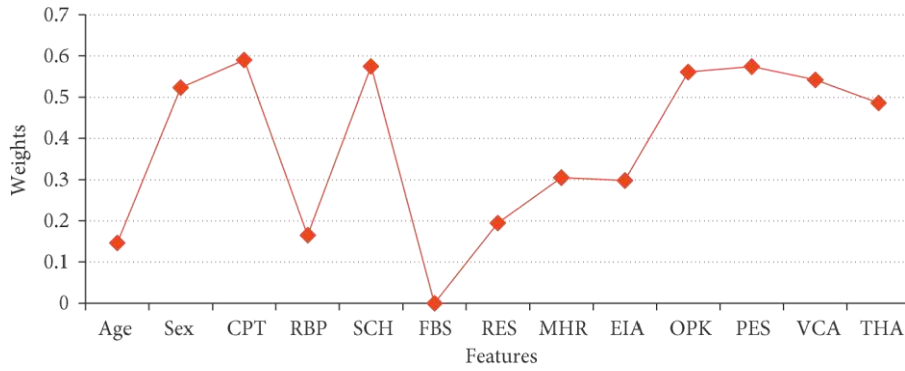| Order	Feature	Feature name	Feature code	Score
1	3	Type chest pain	CPT	0.590
2	5	Serum cholesterol	SCH	0.575
3	11	SlopofST	PES	0.574
4	12	Fluoroscopy	VCA	0.542
5	2	Sex	SEX	0.523
6	13	Thallium scan	THA	0.486

Figure 3: Important features selected by mRMR FS algorithm.

Important features selected by LASSO FS algorithm.

Order	Feature	Feature name	Feature code	Score
1	2	Sex	SEX	0.15
2	12	Fluoroscopy	VCA	0.14
3	9	EIAgina	EIA	0.13
4	3	Type chest pain	CPT	0.10
5	11	SlopofST	PES	0.08
6	13	Thallium scan	THA	0.08

performance of K-NN was excellent as shown in Figure 5. The artificial neural network was trained on different number of inputs and hidden neurons, and then it produced output. After this, with 13 inputs, 16 hidden neurons units, and the last layer having 2 units, it gives output. The ANN classifier achieved 73% accuracy, 74% specificity, and 73% sensitivity. The SVM kernel RBF at C 100 and g 0.0001 has 88% specificity, 78% sensitivity, and 86% accuracy. Similarly, SVM using linear kernel has the best specificity 78%, sensitivity 75%, and accuracy 75%. The NB was the second best classifier that has specificity 87%, sensitivity 78%, and accuracy 84%. The decision tree has specificity 76%, sensitivity 68%, and accuracy 74%. The decision tree has 74% accuracy, 76% specificity, and 68% sensitivity. The random forest classifier with classification accuracy 83%, specificity 70%, and sensitivity 94% is given. Figure 5 shows the classification performance of K-NN with different values of k.

Figure 6 shows the performance of classifiers with 10fold CV on full features.

As shown in Figure 6, the performance of SVM outperformed to the other five classifiers in term of accuracy, sensitivity, and specificity. The predictive accuracy of SVM (RBF) was 86%, sensitivity 78%, and specificity 88%. The second important classifier was NB which has specificity 87%, sensitivity 78%, and classification accuracy 83%. The worst performance was observed for ANN out of five classifiers in terms of accuracy, sensitivity, and specificity
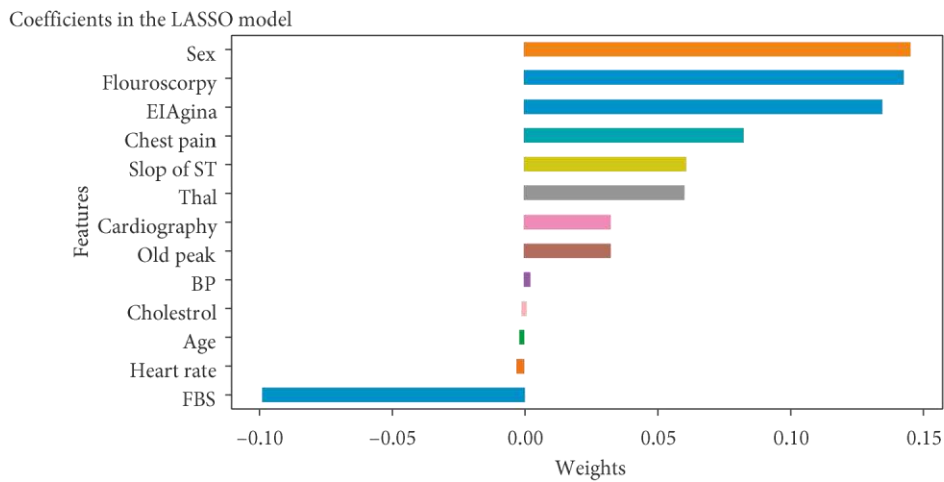
Figure 4: Important features scores and ranks selected by LASSO.

Table 6: 10-fold CV classification performance evaluation of different classifiers on Cleveland heart disease dataset on full features.

Predictive model	Classifiers performance evaluation metrics					
	Accuracy (%)	Specificity (%)	Sensitivity (%)	MCC	AUC (%)	Processing time (s)
Logistic regression (C 10)	84	85	83	89	84	19.213
K-nearest neighbor (K-NN, K 9)	76	74	73	76	73	29.400
Artificial neural network (13, 16, 2)	74	73	74	50	69	21.600
SVM (kernel RBF, C 100, g 0.0001)	86	88	78	85	86	15.234
SVM (kernel linear)	75	78	75	78	74	18.239
Naive Bayes	83	87	78	80	84	34.101
Decision tree	74	76	68	75	76	21.911
Random forest (100)	83	70	94	82	83	15.121
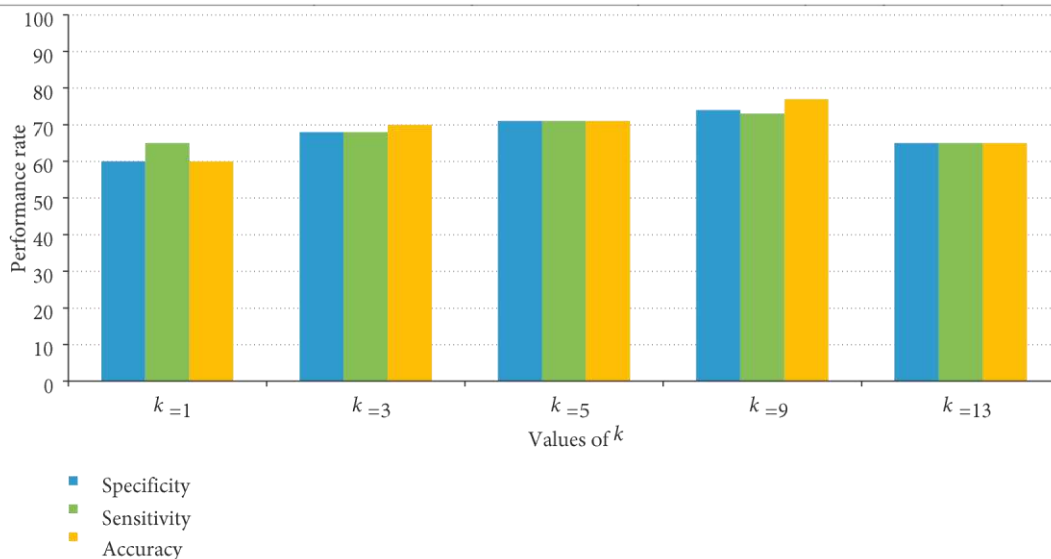
Figure 5: Performance of K-NN on different values of k.

which were 73%, 73%, and 74%, respectively. Figure 7 shows the classifiers processing time in seconds with 10-fold CV.

In Figure 7, the processing time of each classifier in which SVM processing time was 15.234 seconds which is computationally very fast as compared with other classifiers is shown. Figure 8 shows the AUC values of different classifiers with k-fold CV.

AUC for both training and testing of SVM was 86% and 85%, respectively, which shows that SVM covered 86% and 85% area which was greater as compared with other
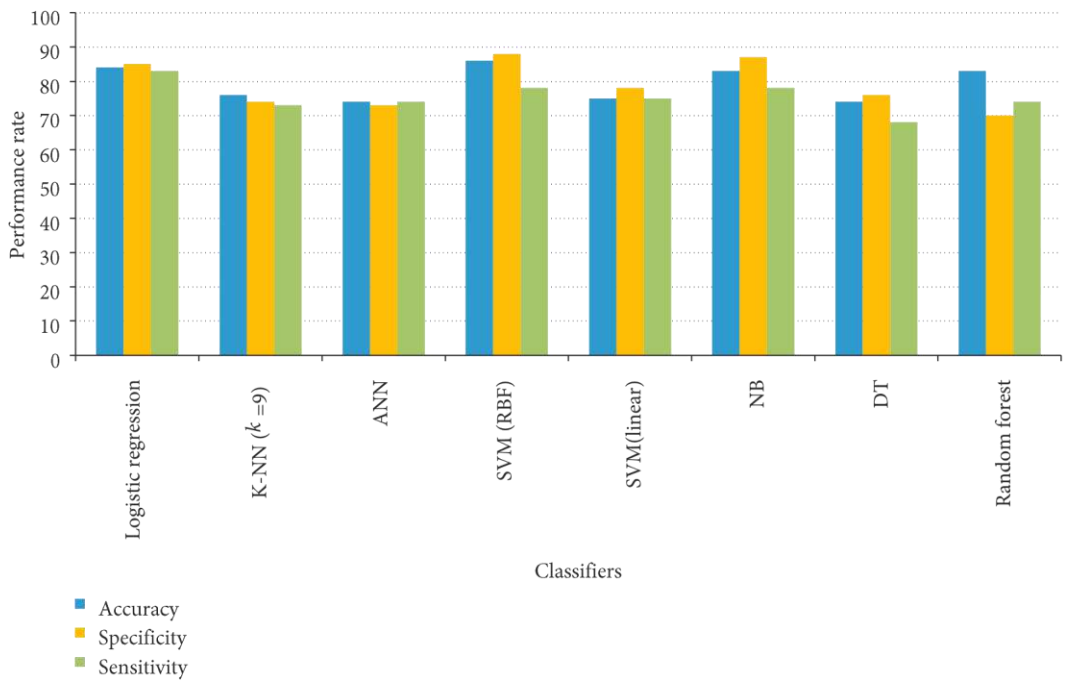
16

Figure 6: Performance of different classifiers with 10-fold CV on full features.

Figure 7: Classifiers processing time in seconds.

classifiers. The larger value of AUC shows the more effective performance of classifiers. The AUC of classifiers is shown in Figure 8.

3.5. Results of K-Fold Cross-Validation (k 10) Classifier Performance on Selected Features (n 6) by Relief FS Algorithm. In this experiment, selected features by Relief FS algorithm were checked on seven machine learning classifiers with 10-fold cross-validation methods. In 10-fold CV, 90% was used for training the classifiers and only 10% was tested. Finally, the average metrics of 10-fold methods were computed. Moreover, different parameters values were passed through classifiers. Initially, we trained and tested the classifiers with the most important 3 features; second time, we fed 4 features, then 6 important features, similarly fed 8, 10 important features; and finally, we used 12 important features. The performances of classifiers were pretty good on 6 important features. Hence 7 tables for 10-fold crossvalidation were formed but we only described the performance of classifiers on 6 important features in Table 7. And for better demonstration of results, some graphs have been created for classification accuracy, specificity, sensitivity, MCC, and processing time. These performance metrics were computed automatically.

According to Table 7, the logistic regression at hyperparameters *C* 100 showed a very good performance, and 89% accuracy, 98% specificity, and 77% sensitivity were
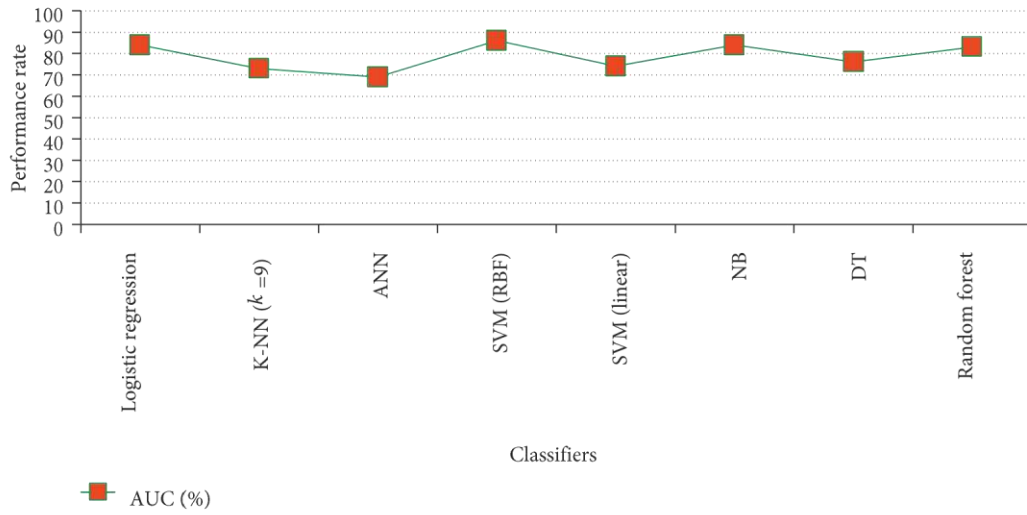
Figure 8: The AUC of different classifiers.

Table 7: 10-fold CV Classification performance of different classifiers on selected features by Relief FS algorithm when $n = 6$.

Predictive model	Classifiers performance evaluation metrics						
	Turning parameters	Accuracy (%)	Specificity (%)	Sensitivity (%)	MCC	AUC (%)	Processing time (s)
Logistic regression	C 1	88	98	76	88	87	16.213
	C 10	87	98	76	88	87	16.200
	C 100	89	98	77	89	88	16.111
	C 0.001	74	98	47	72	73	16.233
K-nearest neighbor	K 1	80	73	78	80	80	24.400
	K 3	75	80	72	76	76	24.500
	K 7	74	78	71	75	75	24.600
	K 9	73	78	70	75	73	24.611
	K 13	70	69	71	70	71	21.777
Artificial neural network	16	77	2		100	50	69	21.600
	20	54	96	5		50	68	22.101
SVM (kernel RBF)	C 100, g 0.0001 C 1,	87	95	78	86	87	14.134
	g 0.01	79	82	81	79	80	14.139
	C 10, g 0.001	75	84	68	76	77	14.255
SVM (kernel linear)	C 10, g 0.0001 C	78	95	55	78	74	18.139
	100, g 0.0001	80	97	60	79	79	18.222
Naive Bayes	—	85	87	78	80	84	34.101
Decision tree	100	74	85	66	75	76	20.911
	500	73	84	65	74	74	20.899
Random forest	100	83	93	70	82	83	15.121
	50	85	94	74	82	84	14.330
	25	82	94	70	82	82	14.199

obtained along with 89% MCC. The AUC value of logistic regression is 88%, and the processing time is 16.111 seconds. The performance at C 0.001 logistic regression obtained an accuracy of 74%, 98% specificity, and 47% sensitivity along with 72% MCC. Moreover, the AUC value of logistic regression was 73%, and the processing time was 16.233 seconds.

For K-NN, we fed different values of K 1, 3, 7, 9, and 13 but at k 1 the K-NN shows good performance with 88% accuracy at computation time 24.400 seconds. However, at k 13, the K-NN performance was not good. The artificial neural networks were formed as multilayer perceptron (MLP), and in MLP, a different number of hidden neurons were used. At 16 hidden neurons, the MLP gives good results. ANN obtained 77% accuracy at 16 hidden neurons, and at 20 hidden neurons, poor performance was observed.

The performance of SVM (RBF) at C 100 and g 0.0001 was good as compared to other values of C and g as shown in Table 7. SVM (kernel RBF) obtained accuracy 87%, specificity 95%, sensitivity 78%, MCC 86%, and AUC 87%. The computational time was 14.134 seconds. SVM (kernel linear) at C 100 and g 0.0001 obtained accuracy 80%, specificity 98%, and sensitivity 60% with computational time 18.222 seconds. The NB obtained classification accuracy of 85%, specificity 87%, and sensitivity 78% with processing time 34.101 seconds. We applied 100 and 500 trees for ensemble classifiers. The ensemble with 100 gives 74% accuracy, 85% specificity, 66% sensitivity, and 75% MCC. The computational time was 20.911 seconds. The performance of ensemble at 500 was little poor and obtained 73% accuracy, 84% specificity, and 65% sensitivity, and processing time was 20.889 seconds. For random forest, 100, 50, and 25 iterations were applied. At 100, accuracy 83%, specificity 93%, sensitivity 70%, and MCC 82% were obtained. The AUC value at 100 was 83%. The processing time was 15.121 seconds. The random forest at 50 has very pretty good performance and obtained classification accuracy of 85%, specificity of 94%, sensitivity of 74%, and MCC of 82%, and AUC was 84%. Table 7 show the 10-fold CV classifiers performance on selected features by Relief FS algorithm.

Figure 9 shows the performance of classifiers on 6 important selected features by Relief FS with 10-fold CV.

As shown in Figure 9, the classification accuracy of logistic regression at C 100 was 89% at 6 important features at 10-fold cross-validation with respect to other classifiers.

The SVM (kernel RBF at C 100, g 0.0001) was the second best classifier and obtained 87% accuracy, the SVM (kernel linear at c 100, g 0.0001) obtained 80% accuracy. The accuracy of K-NN at k 1 was 80%. The ANN obtained classification accuracy of 77% at 16 hidden neurons. The NB accuracy was pretty good, 85%. DT accuracy at 100 was 74%. The random forest accuracy is 85%. So from Figure 9, logistic regressions at 6 important features give better results as compared to other classifiers. The specificity of logistic regression is 98% which is high among others classifiers; SVM (RBF) specificity is 95%; and SVM linear specificity values is 97%. Moreover, the lowest specificity of ANN was 2%. K-NN at k 1 has specificity of 73%. DT and random forest have 85% and 94% specificity, respectively. The sensitivity of ANN was 100%; logistic regression has 77%, K-NN sensitivity was 78%. The poor sensitivity of SVM (linear) was 55%. Figure 10 shows the ACU values of classifiers of 6 important features selected by Relief FS with 10-fold CV.

The ROC AUC values of classifiers at 6 important features are also shown in Figures 10. The AUC values of logistic regression and SVM (RBF) are 88% and 87%, respectively, which are large as compared to other classifiers. DT and K-NN have poor AUC values 76% and 69%, respectively. Figure 11 shows the processing time of classifiers at six important features selected by Relief with 10-fold CV.

The processing time of classifiers on six important features by Relief at suitable classifiers parameters is shown in Figure 11. The logistic regression processing time was 16.111 seconds. SVM (RBF) has processing time of 14.134 seconds, and random forest processing time was 14.333 seconds. The processing time of these three classifiers was lower and K-NN, DT, and NB processing time were 24.400 seconds, 20.911 seconds, and 34.101 seconds, respectively. Figure 12 shows MCC of classifiers at six important features selected by Relief with 10-fold CV.

The MCC of different classifiers on six important features was excellent as shown in Figure 12. According to Figure 12, logistic regression and SVM (RBF) had high MCC values while ANN and DT were lowest MCC values on six important features by Relief with 10-fold cross-validation. Table 7 shows 10-fold CV of classifiers with selected features by Relief.

3.6. Results with K-Fold Cross-Validation of Classifiers Performance on Selected Features(n = 6)by mRMRFS Algorithm. In this experiment, the selected features by mRMR FS algorithm were checked on seven machine learning classifiers with 10-fold cross-validation methods. In 10-fold CV, 90% was used for training the classifiers and only 10% was tested. Finally, the average metrics of 10 folds were computed. Moreover, different parameters values were passed through classifiers. Firstly, we trained and tested the classifiers with the important 3 features; second time, we fed 4 features, then 6 important features, similarly fed 8, 10 important features; and finally, used 12 important features. The performance of classifiers was good enough on 6 important features. Hence, 8 tables for 10-fold cross-validation were formed, but in this paper, we only described the performance of classifiers on 6 important features in Table 8 because the overall performance of classifiers at 6 important features was good as compared to the performance on experiments on 3, 4, 8, 10, and 12 important features. For better demonstration of the results, some graphs have been created for classification accuracy, specificity, sensitivity, MCC, processing time, and ROC AUC. All these performance metrics were computed automatically. Table 8 shows the 10-fold CV classification performance of different classifiers on selected features by mRMR FS algorithm.

From Table 8, the logistic regression at hyperparameters C 100 was a very good performance, and 78% accuracy, 88% specificity, and 67% sensitivity were obtained along with 78% MCC. The AUC value of logistic regression was 79%, and processing time was 2.159 seconds, while other values of C performance were not good. For K-NN, we fed different values of K 1, 3, and 7 but at k 7, K-NN shows good performance with 62% accuracy and computation time was 10.144 seconds. However, at k 3, the K-NN performance is not good. The artificial neural networks were formed as MLP, and in MLP, a different number of hidden neurons were used. At 16 hidden neurons, the MLP gives good results. ANN obtained 63% accuracy at 16 hidden neurons, and at 20 hidden neurons, poor performance was observed and 47% accuracy was obtained.

The performance of SVM (RBF) at C 100 and g 0.0001 was good as compared to other values of C and g as shown in Table 8. SVM (kernel RBF) obtained accuracy 77%, specificity 88%, sensitivity 65%, MCC 76%, and AUC 77%. The computational time was 60.589 seconds. SVM (kernel linear) at C 100 and g 0.0001 obtained accuracy 70%, specificity 100%, sensitivity 35%, and MCC 71% with computational time 10.179 seconds. The NB obtained classification accuracy 84%, specificity 90%, sensitivity 77%, and MCC 83% with processing time 1.596 seconds. We applied 100 and 50 trees for ensemble classifiers. The ensemble with 100 gives 57% accuracy, 55% specificity, 60% sensitivity, and 58% MCC. The computational time was 1.902 seconds. The performance of ensemble at 50 was good and obtained 60% accuracy, 54% specificity, and 67% sensitivity, and processing time was 1.831 seconds.
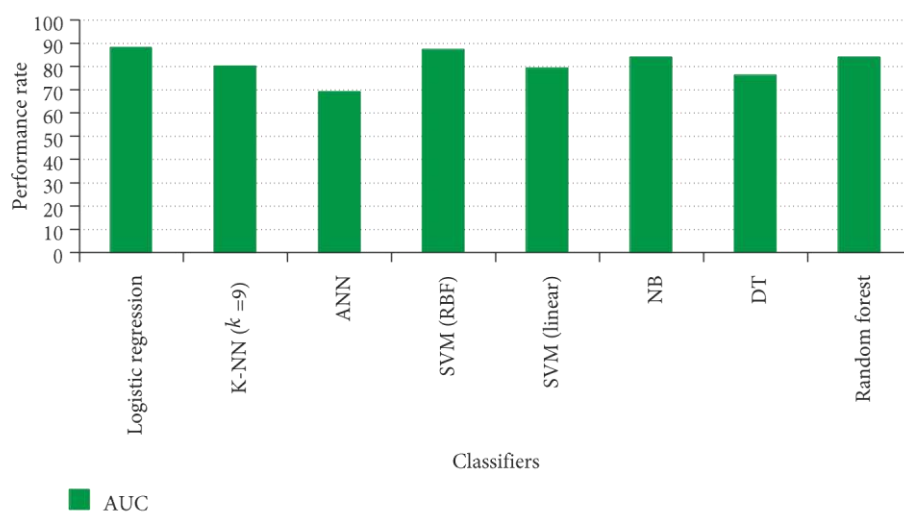
For
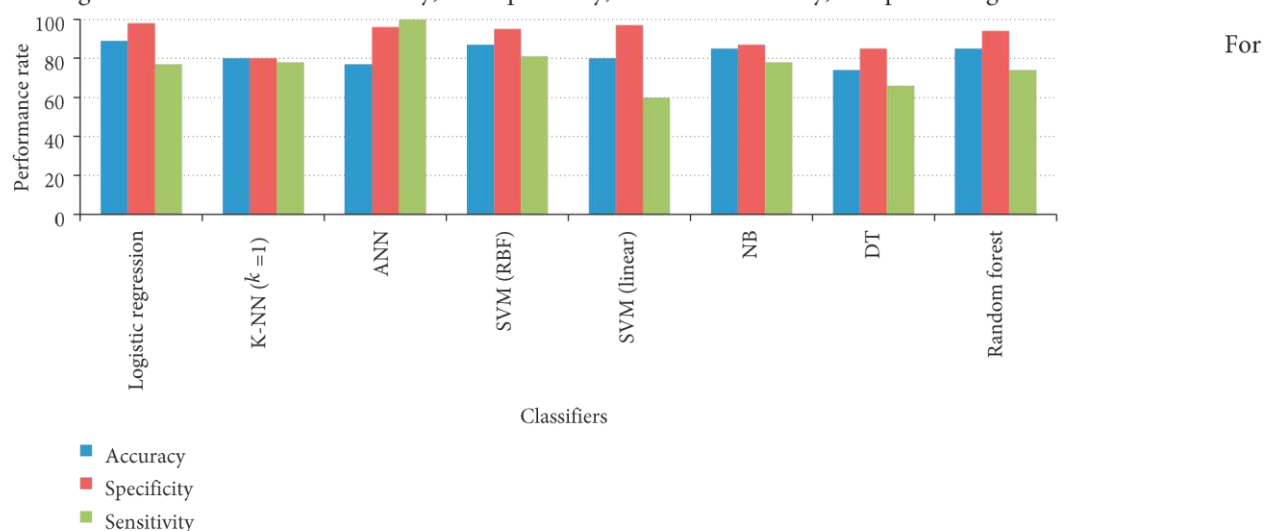
Figure 9: Performance of classifiers on six important selected features by Relief FS with 10-fold CV.

random forest, 100 and 50 iterations were applied. At 100, accuracy 66%, specificity 69%, sensitivity 62%, and MCC 66% were obtained. The AUC value at 100 was 65%. The processing time was 1.100 seconds. The random forest at 50 shows pretty good performance and classification accuracy 67%, specificity 70%, sensitivity 62%, and MCC 66% were obtained, and AUC was 68%. The computational time was 2.220 seconds. Figure 13 shows the performance of classifiers on six important features selected by mRMR FS algorithm with 10-fold CV.

As shown in Figure 13, the classification accuracy of logistic regression at C 100 is 78% for 6 features of 10-fold cross-validation. The SVM (kernel RBF at C 100 and g 0.0001) obtained 77% accuracy; the SVM (kernel linear at C 100 and g 0.0001) obtained 70% accuracy. The accuracy of K-NN at k 7 was 62%. The ANN obtained classification accuracy 63% at 16 hidden neurons. The NB accuracy was 84% which is as compared with other classifiers. DT accuracy at 100 was 57% while on 50 it was 60%. The random

forest accuracy is 67%. Figure 13 shows that NB classification accuracy at 6 features give better results as compared with other classifiers. The specificity and sensitivity of logistic regression was 88% and 66% at C 100, respectively. SVM (RBF) at C 100 and g 0.0001 specificity and sensitivity were 88% and 65%, respectively. SVM linear specificity was 100% and sensitivity was 35%. Moreover, the specificity of ANN was 67% and sensitivity was 58% at 16 hidden neurons. K-NN at k 7 specificity was 73% and sensitivity was 61%. DT at 50 has specificity and sensitivity 54% and 67%, respectively. Random forest at 50 iterations has 70% and 62% specificity and sensitivity, respectively. Lastly, the best classifiers in terms of accuracy was NB and has accuracy 84%, in terms of specificity, SVM linear at C 100 and g 0.0001 was good and obtained 100% and sensitivity of ANN was 98% as compared with other
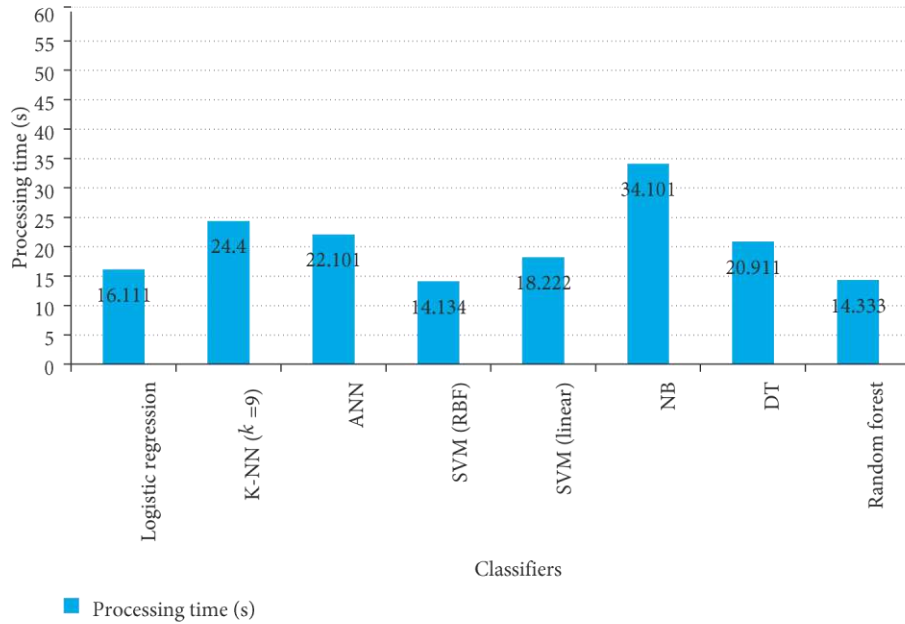
Figure 11: Processing time of classifiers on six important features selected by Relief with 10-fold CV.

Figure 12: MCC of classifiers at six important features selected by Relief with 10-fold CV.

classifiers at 6 important features selected by mRMR FS. Figure 14 shows the AUC of Classifiers on six important features selected by mRMR FS algorithm with 10-fold CV.

The ROC AUC values of classifiers at 6 features are shown in Figures 14. The AUC value of logistic regression, SVM (RBF), and NB were 79%, 77%, and 84%, respectively, which were large as compared with other classifiers. DT, K-NN, and ANN had poor AUC values of 61%, 65%, and 66%, respectively. The ROC AUC of Naive Bayes was 84% on selected features with k folds cross-validation as compared with other classifiers. Figure 15 shows the processing time of classifiers on selected feature s by mRMR with 10-fold CV.

The computational time of classifiers on the six important features by mRMR FS algorithm with suitable classifiers parameters is shown in Figure 15. The logistic regression processing time was 2.159 seconds. SVM (RBF) has processing 60.589 seconds, and random forest processing time was 2.222 seconds. DTprocessing time was 1.831 seconds, and NB time

was 1.596 seconds. The processing time of SVM (RBF) was large as compared to other classifiers. The lowest processing time of NB was 1.596 seconds as compared to other classifiers. Figure 16 shows MCC of classifiers at selected features by mRMR FS algorithm with 10-fold CV.

The MCC of different classifiers at 6 features was excellent as shown in Figure 16. According to the graph, the logistic regression MCC value was 78%. The K-NN MCC at k 7 was 62 which is same aANN. SVM (RBF) MCC was 76%, and SVM (Linear) MCC was 68%. The NB, DT, and random forest MCC were 83%, 60%, and 66%, respectively. The high value of MCC shows better performance of classifiers. Therefore, the performance of NB was good, and it's MCC was 83% at selected features by mRMR feature selection algorithm. Logistic regression and SVM (RBF) performances were also good on reduced features.

Table 8: 10-fold CV classification performance of different classifiers on selected features by mRMR FS algorithm when n 6.

| Predictive model		Classifiers performance evaluation metrics						
	Turning parameters	Accuracy (%)	Specificity (%)	Sensitivity (%)	MCC		AUC (%)	Processing time (s)
Logistic regression	C 1 C	74	82	66	74		74	2.313
	10	75	82	67	74		75	2.352
	C 100	78	88	67	78		79	2.159
K-nearest neighbor	K 1	57	57	58	57		63	1.784
	K 3	56	56	55	56		55	1.742
	K 7	62	62	61	62		65	10.144
Artificial neural network	16	63	67	58	62		66	30.802
	20	47	4	98	51		50	23.483
SVM (kernel RBF)	C 100, g 0.0001 C	77	88	65	76		77	60.589
	10, g 0.001	66	71	60	65		67	59.132
SVM (kernel linear)	C 10, g 0.0001 C	58	23	70	60		59	12.567
	100, g 0.0001	70	100	35	68		71	10.179
Naive Bayes	—	84	90	77	83		84	1.596
Decision tree	100	57	55	60	58		57	1.902
	50	60	54	67	60		61	1.831
Random forest	100	66	69	62	66		65	1.121
	50	67	70	62	66		68	2.220
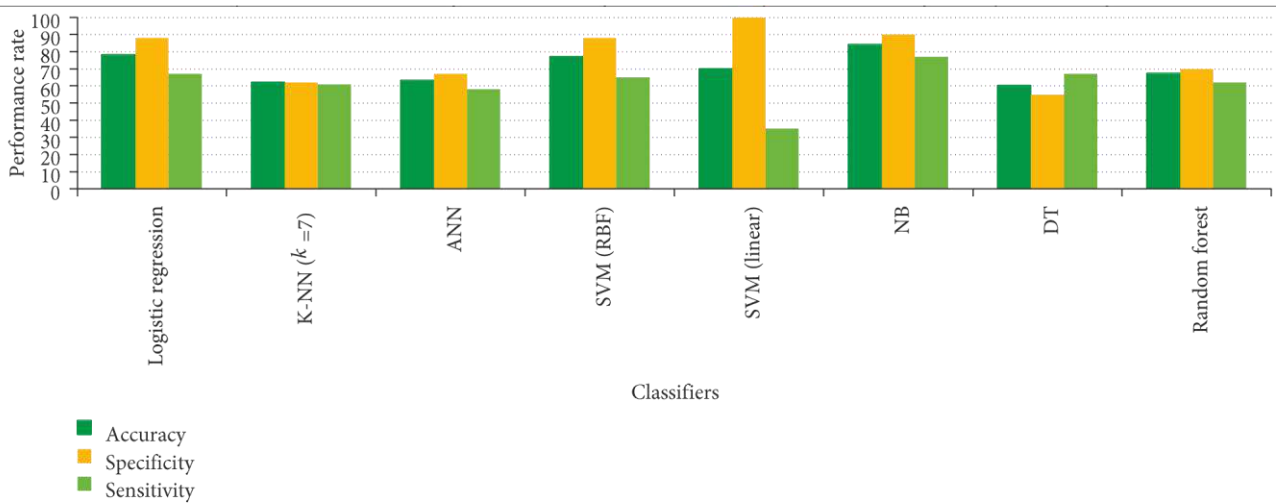
Figure 13: Performance of classifiers on six important features selected by mRMR FS algorithm with 10-fold CV.

3.7. Results of K-Fold Cross-Validation (k 10) Classifiers Performance on Selected Features (n 6) by LASSO FS Algorithm. In this section, the selected features by LASSO feature selection algorithm were checked on seven machine learning classifiers with 10-fold cross-validation method. In 10-fold CV, 90% was used for training the classifiers and 10% was used for testing. Moreover, different parameters values were passed through classifiers. Firstly, we used 3 features; second time, we fed 4 features and then 6 features, similarly 8, 10 important features; and finally, we used 12

important features. The performances of classifiers were good on 6 features. Hence, 8 tables for 10-fold cross-validation were formed but we only described the performance of classifiers on 6 important features in Table 9, because the overall performance of classifiers at 6 important features was good as compared with the performance of 3, 4, 8, 10, and 12 important features. For better demonstration of results, some graphs have been created. Additionally, performance evaluation metrics were computed automatically. Table 9 shows 10fold CV classification performance of different classifiers on selected features by LASSO FS algorithm.

According to Table 9, logistic regression at hyperparameters C 10 obtained 87% accuracy, 96% specificity, and 76% sensitivity along with 87% MCC. The AUC of logistic regression was 88%, and the processing time was 0.008 seconds, while other values of C performance were not as good as compared to C 10. We used different values of k 1, 3, 5, and 7 for K-NN but at k 1, K-NN shows good performance with 85% accuracy, 94% specificity, 74% sensitivity, and 84% MCC, and computation time was 0.0002 seconds. However, at k 7, the K-NN performances were
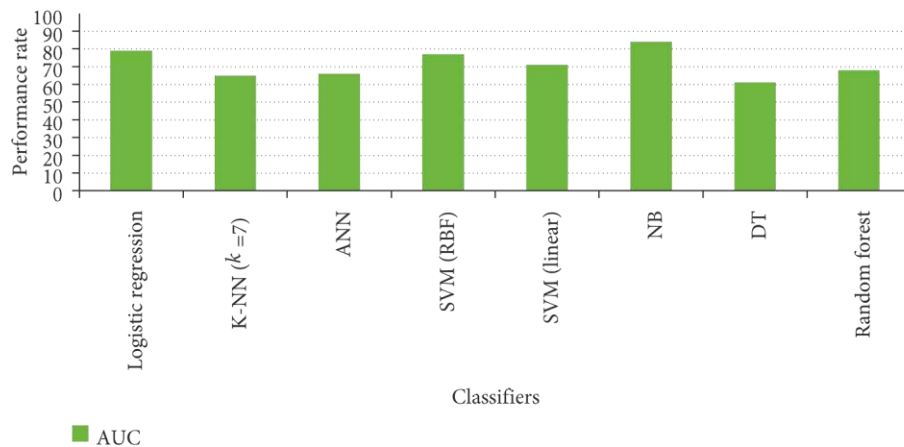

Figure 14: Area under the curve (AUC) of classifiers on six important features selected by mRMR FS algorithm with 10-fold CV.
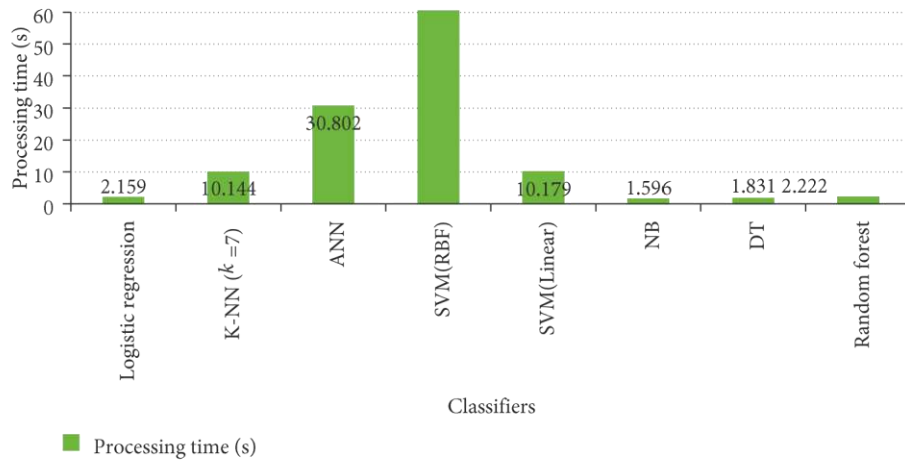
Figure 15: Processing time (s) of classifiers on selected feature *s* by mRMR with *k*-fold CV.
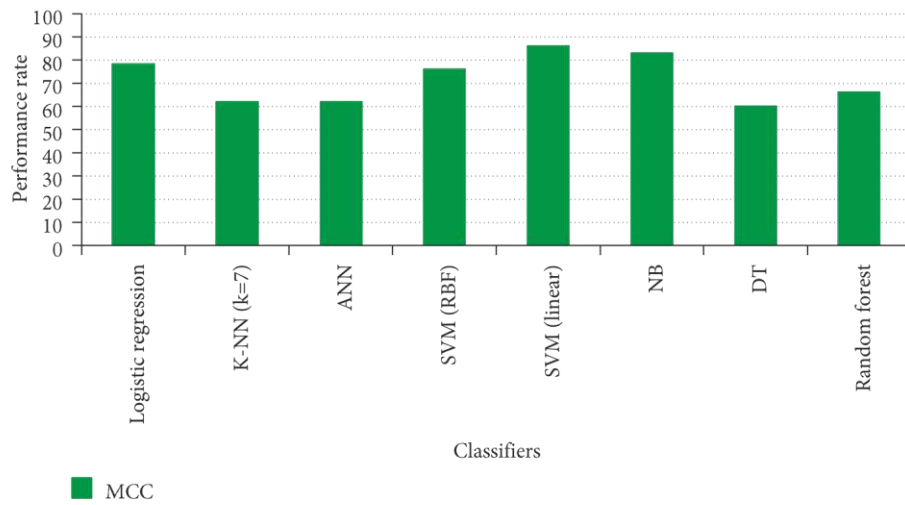

Figure 16: MCC of classifiers at selected features by mRMR FS algorithm with 10-fold CV.

Table 9: 10-fold CV classification performance of different classifiers on selected features by LASSO FS algorithm when *n* = 6

Predictive model	Turning parameters	Classifiers performance evaluation metrics					
		Accuracy (%)	Specificity (%)	Sensitivity (%)	MCC	AUC (%)	Processing time (s)
Logistic regression	*C* 1 *C*	85	94	74	84	86	0.012
	10	87	97	76	87	88	0.019
	C 0.1	83	90	75	84	84	0.069
K-nearest neighbor	*K* 1	85	94	74	84	85	0.024
	K 3	84	94	72	85	83	0.016
	K 7	81	88	73	84	80	1.799
Artificial neural network	16	86	94	77	85	85	7.650
	20	82	94	70	82	81	7.362

		40	71	88	38	69	69	7.400
SVM (kernel RBF)	C 10, g 0.0001	85	94	74	85	84	0.019	
	C 100, g 0.001	88	96	75	88	89	0.009	
SVM (kernel linear)	C 10, g 0.0001 C	84	96	74	85	85	0.023	
	100, g 0.0001	82	96	75	84	84	0.005	
Naive Bayes	—	83	88	78	82	82	6.591	
Decision tree	100	84	92	73	83	84	2.606	
	50	83	90	70	83	83	2.774	
Random forest	100	83	92	72	82	83	0.017	

The performance of SVM (RBF) at C 100, g 0.0001 was good as compared to other values of C and g as shown in Table 7. SVM (kernel RBF) obtained accuracy 88%, specificity 96%, sensitivity 75%, MCC 85%, and AUC 84%. The computational time was 0.002 seconds. SVM (kernel linear) at C 10 and g 0.0001 obtained accuracy 84%, specificity 96%, sensitivity 74%, and MCC 85% with computational time 0.003 seconds. The NB obtained classification accuracy 83%, specificity 88%, sensitivity 77%, and MCC 82% with processing time 6.591 seconds. We applied 100 and 50 trees for ensemble classifiers. The ensemble with 100 gives 84% accuracy, 92% specificity, 73% sensitivity, and 84% MCC. The computational time was 2.606 seconds. The performance of ensemble at 50 was also good and obtained 83% accuracy, 90% specificity, 70% sensitivity, 83% MCC, and processing time was 12.774 seconds. For random forest, 100 and 50 iterations were applied. At 100, accuracy 66%, specificity 69%, sensitivity 62%, and MCC 66% were obtained. The AUC value at 100 was 65%. The processing time was 1.100 seconds. The random forest at 50 has pretty good performance and obtained classification accuracy 83%, specificity 92%, sensitivity 72%, and MCC 82% and AUC was 83%. The computational time was 0.017 seconds. Figure 17 shows performance of classifiers on six features selected by LASSO FS algorithm with 10-fold CV.

The performance of classifiers is shown in Figure 17. According to Figure 17, in terms of classification, accuracy of SVM (RBF) at C 100 and g 0.0001 was 88% on selected features which was good as compared to other classifiers. Logistic regression accuracy was 87%, and ANN accuracy was 86%. These three classifiers on selected features by LASSO give a good performance. Additionally, in terms of specificity, logistic regression obtained 97%, and SVM (RBF) at C 100, g 0.0001 was good and obtained 96% and sensitivity of ANN was 77% and Naive Bayes 78% as compared to other classifiers at 6 important features CV.

selected by LASSO FS algorithm. Figure 18 shows AUC on six important features selected by LASSO FS algorithm with 10fold

The ROC AUC graph of classifiers at 6 important features is shown in Figure 18. The AUC values of logistic regression and SVM (RBF) were 88% and 89%, respectively, which were large as compared to other classifiers. The AUC values K-NN, ANN, DT, and NB were 85%, 85, 84%, and 82%, respectively. Figure 19 shows processing time of classifiers on six important features selected by LASSO FS algorithm with 10-fold CV.

The computational time of classifiers on 6 important selected features LASSO FS algorithm with suitable classifiers parameters is shown in Figure 19. The logistic regression processing time was 0.008 seconds. SVM (RBF) has a processing time of 0.009 seconds, and random forest processing time was 0.017 seconds. DT processing time was 2.606 seconds, and NB time was 6.591 seconds. The processing time of ANN time was 7.650 seconds large as

compared to other classifiers. The lowest processing time of K-NN at k 1 was 0.002 seconds as compared to other classifiers. Figure 20 shows MCC of classifiers on six important features selected by LASSO FS algorithm with 10-fold CV.

The MCC of different classifiers on six important features was good enough as shown in Figure 20. According to the graph, the logistic regression MCC value was 87%. The K-NN MCC at k 1 was 85% which is same as A-NN. SVM (RBF) MCC was 88%, and SVM (linear) MCC was 85%. The NB, DT, and random forest MCC were 82%, 83%, and 82%, respectively. The high value of MCC shows better performance of classifiers. Therefore, SVM (RBF) MCC was 88%, and it is a good predictive model for heart disease prediction. According to the results of three feature selection algorithms, the performance of best classifiers with their evaluation metrics has been shown in Table 10 using 10-fold cross-validation.
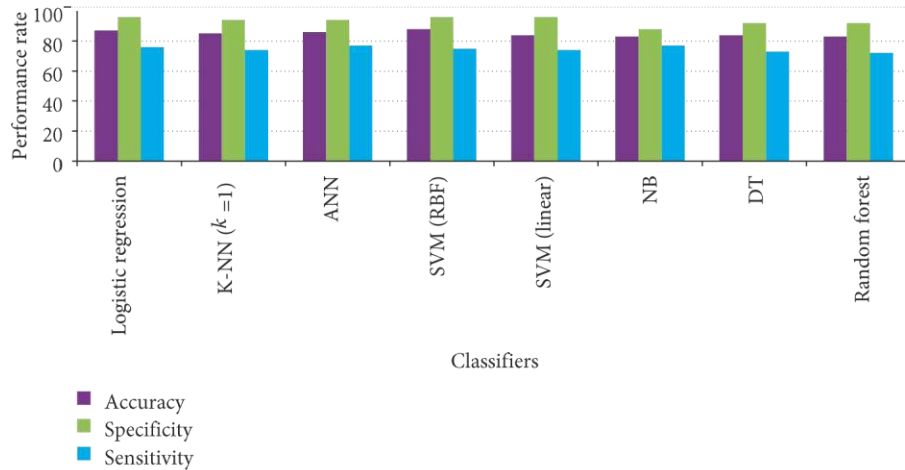
Figure 17: Performance of classifiers on six important features selected by LASSO FS algorithm with 10-fold CV.
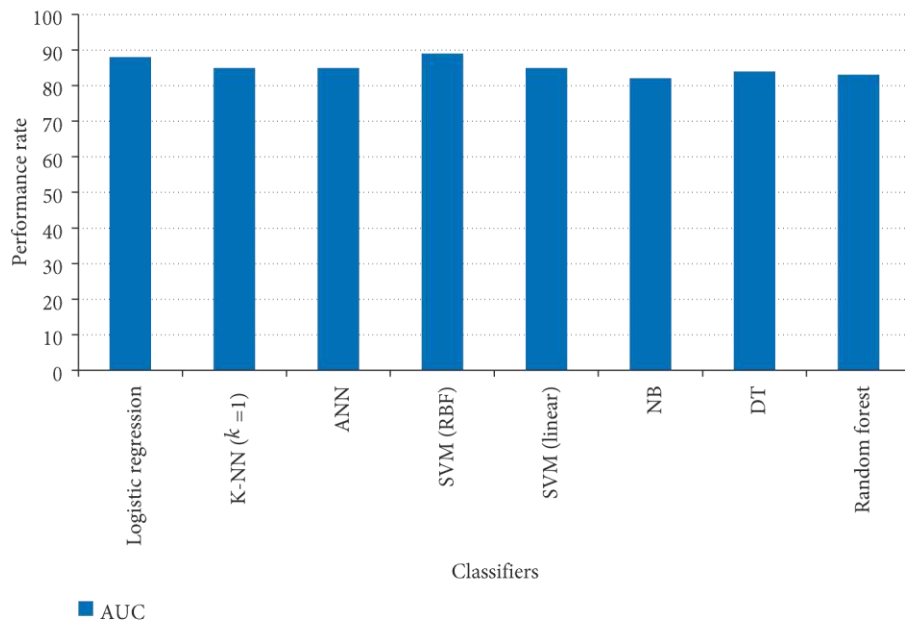

Figure 18: AUC on 6 important features selected by LASSO FS algorithm with 10-fold CV.

Table 10 shows that logistic regression accuracy was the best (89%) on selected features by Relief FS algorithm as compared to mRMR and LASSO feature selection algorithms with 10-fold cross-validation. Hence, in terms of accuracy, Relief FS algorithm is the best for important feature selection and logistic regression is the suitable classifier for classification of heart disease and healthy subjects. Specificity of classifiers as shown in Table 10 indicates that specificity of SVM is the best on mRMR FS algorithm as compared to the specificity of Relief and LASSO feature selection algorithms. The mRMR FS algorithm selected import features for correct classification of healthy people. Additionally, AUC values of SVM (RBF) with LASSO FS give best results with respect to other classifiers and feature selection algorithms.

The sensitivity of the classifier ANN (MLP) with 16 hidden neurons is the best (100%) on the selected features by Relief FS algorithm and correctly classified the people with heart disease and normal people. The sensitivity of the classifier Naive

Bayes on a selected feature by LASSO FS algorithm has the worst results. In the case of MCC, Relief selects most suitable features with classifier logistic regression and achieved best MCC as compared to the MCC values of mRMR and LASSO FS algorithm. The AUC of classifier SVM (RBF) with C 100 and g 0.001 on 6 selected features selected by LASSO FS algorithm gives the best results. The other feature selection algorithms (Relief and mRMR) in case the AUC are the worst FS algorithms. The computation time of different classifiers with six selected features by Relief, mRMR, and LASSO FS algorithms is given in Table 10. The computation time of LASSO features selection is low as compared to Relief and mRMR FS algorithms. uracy of Naive Bayes was 84% and SVM has
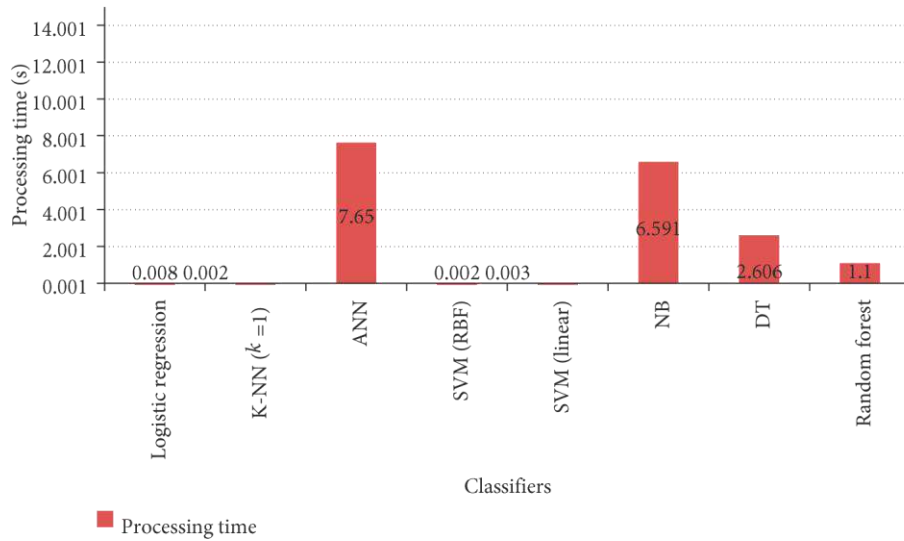
Figure 19: Processing time of classifiers on six important features selected by LASSO FS algorithm with 10-fold CV.
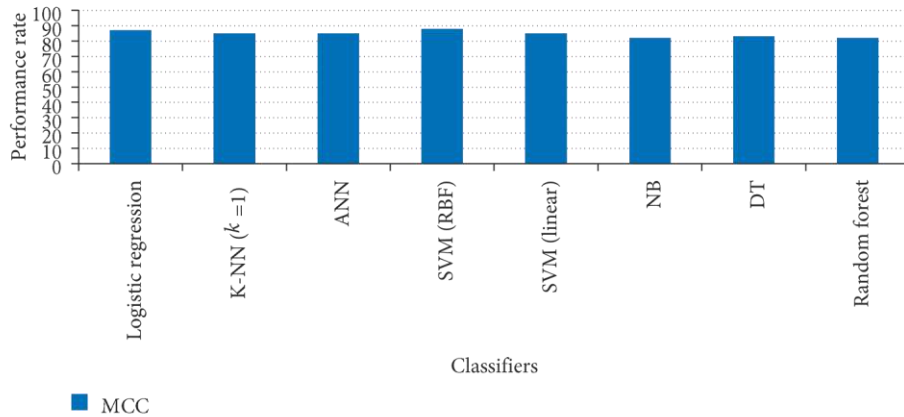

■ MCC

Figure 20: MCC of classifiers on six important features selected by LASSO FS algorithm with 10-fold CV.

Table 10: Excellent performance metrics results and best classifiers with feature selection algorithms for n = 6 with 10-fold CV.

	Best performance evolution metrics and best classifiers				
FSS	The Best accuracy (%) and classifier	The Best sensitity (%) and classifier	The best MCC and classifiers	The Best AUC and classifiers	The Best process time(s) & classifier
Relief	89 logistic regression with C=100	98 logistic regression with C=100	100ANN	89 logistic regression with C=100	14.134 SVM

Table 11 shows that the classification accuracy of logistic regression increased from 84% to 89% on reduced features. Similarly, SVM (RBF) accuracy increased from 86% to 88% with reduced features. Hence, the feature selection algorithms select important features which increased the performance of the classifiers and reduced the execution time as well. The designing of a diagnosis system for heart disease prediction using FS with classifiers will effectively improve performance.

Conclusions.

.

In this research study, a hybrid intelligent machine-learningbased predictive system was proposed for the diagnosis of heart disease. The system was tested on Cleveland heart disease dataset. Seven well-known classifiers such as logistic regression, K-NN, ANN, SVM, NB, DT, and random forest were used with three feature selection algorithms Relief, mRMR, and LASSO used to select the important features. The K-fold cross-validation method was used in the system for validation. In order to check the performance of classifiers, different evaluation metrics were also adopted. The feature selection algorithms select important features that improve the performance of classifiers in terms of classification accuracy, specificity, and sensitivity, MCC and reduced the computation time of algorithms. The classifiers logistic regression with 10-fold cross-validation showed best accuracy 89% when selected by FS algorithm Relief. Due to the good performance of logistic regression with Relief, it is a better predictive system in terms of accuracy.

In terms of specificity, SVM (linear) with a feature selection, algorithm mRMR performance was the best as compared to the specificity of logistic regression with FS algorithms Relief and LASSO as shown in Table 10. The SVM (linear) with mRMR-based system will correctly classify the health people. The best sensitivity was 100% of classifier ANN (MLP) with 16 hidden neurons on selected features by Relief. The classifier Naive Bayes with LASSO FS algorithm has the worst sensitivity. The ANN with Relief correctly classified the heart disease people. The classier logistic regression MCC was 89% on selected features by Relief FS algorithm as shown in Table 10. The execution time of SVM with LASSO FS algorithm is the best as compared to other features algorithms and classifiers. Feature selection algorithms should be used before classification to improve the classification accuracy of classifiers as shown in Table 11. Hence, through FS algorithms, we can reduce the computation time and improve the classification accuracy of classifiers.

FS algorithms select important features that are related to discriminate HD from healthy people. According to FS algorithms, the most important and suitable features are Thallium scan, type chest pain, and exercise-induced angina; the results of all the three FS algorithms show that the feature fasting blood sugar is not suitable for classification of heart disease and healthy people. The performance of classifiers with Relief FS algorithm important features selection is excellent as compared to mRMR and LASSO.

The novelty of this research work is developing a diagnosis system for HD. The system used three FS algorithms, seven classifiers, one cross-validation method, and performance evaluation metrics for HD diagnosis. The system was tested on Cleveland heart disease dataset to classify HD and healthy subjects. Designing a decision support system through machine-learning-based method will be more suitable for diagnosis of heart disease. Additionally, some irrelevant features reduced the performance of the diagnosis system and increased the computation time. So another innovative dimension of this study was the usage of feature selection algorithms to choose best features that improve the classification accuracy as well as reduce the execution time of the diagnosis system. In the future, we will perform more experiments to increase the performance of these predictive classifiers for heart disease diagnosis by using others feature selection algorithms and optimization techniques.

REFERENCES.

[1] V. Manikantan and S. Latha, "Predicting the analysis of heart disease symptoms using medicinal data mining methods", International Journal of Advanced Computer Theory and Engineering, vol. 2, pp.46-51, 2013.

[2] Sellappan Palaniappan and Rafiah Awang, "Intelligent heart disease prediction system using data mining techniques", International Journal of Computer Science and Network Security, vol.8, no.8, pp. 343350,2008.

[3] K.Srinivas, Dr.G.Ragavendra and Dr. A. Govardhan," A Survey on prediction of heart morbidity using data mining techniques",International Journal of Data Mining & Knowledge Management Process (IJDKP) vol.1, no.3, pp.14-34, May 2011.

[4] G.Subbalakshmi, K.Ramesh and N.Chinna Rao," Decision support in heart disease prediction system using Naïve Bayes", ISSN: 0976-5166, vol. 2, no. 2.pp.170-176, 2011.

[5] T.John Peter , K. Somasundaram, "An Empirical Study on Prediction of Heart Disease using classification data mining technique" IEEE- International Conference On Advances In Engineering, Science And Management (ICAESM - 2012) March 30, 31, 2012.

[6] Shamsher Bahadur Patel, Pramod Kumar Yadav and Dr. D. P.Shukla, "Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques",IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS),Volume 4, Issue 2 (Jul. - Aug. 2013.

[7] Niti Guru, Anil Dahiya, Navin Rajpal, "Decision Support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review, Vol. 8, No. I (January - June 2007.

[8] M.A.Nishara Banu, B.Gomathy, "Disease Forecasting System Using Data Mining Methods," International Conference on Intelligent Computing Applications,2014

[9] Feixiang Huang, Shengyong Wang, and Chien- Chung Chan, "Predicting Disease By Using Data Mining Based on Healthcare Information System", IEEE International Conference on Granular Computing,2012.

[10] Chotirat "Ann" Ratanamahatana and Dimitrios Gunopulos,"Scaling up the Naive Bayesian Classifier:Using Decision Trees for Feature