Neural Comput & Applic
DOI 10.1007/s00521-016-2604-1

ORIGINAL ARTICLE

Performance evaluation of different machine learning techniques for prediction of heart disease

Ashok Kumar Dwivedi[1]

Received: 23 June 2016 / Accepted: 6 September 2016

Abstract Heart diseases are of notable public health disquiet worldwide. Heart patients are growing speedily owing to deficient health awareness and bad consumption lifestyles. Therefore, it is essential to have a framework that can effectually recognize the prevalence of heart disease in thousands of samples instantaneously. At this juncture, the potential of six machine learning techniques was evaluated for prediction of heart disease. The recital of these methods was assessed on eight diverse classification performance indices. In addition, these methods were assessed on receiver operative characteristic curve. The highest classification accuracy of 85 % was reported using logistic regression with sensitivity and specificity of 89 and 81 %, respectively.

Keywords Machine learning · Classification · Heart disease · Treatments · Artificial neural network · Support vector machine

1 Introduction

Conferring to the report of World Health Organization (WHO), heart diseases cause millions global deaths per year. Heart diseases cause deaths due to mental stress, work overload and many other sources. Healthcare data necessitate advance system for mining the pattern hidden within for effective decision making. State-of-the-art data mining approaches are applied to discern knowledge from clinical data for research in medical informatics, essentially in heart disease prophecy. Heart disease diagnosis is complicated nonetheless critical task that is essentially being accomplished precisely and proficiently. This task is frequently made on the understanding and acquaintance of doctor. This causes excessive time and cost. Consequently, an automatic medicinal system is designed that takes advantage of collected database and decision support system. This method can assist in establishing heart disease with less exertion.

Machine learning methods have drawn a high amount of thoughtfulness in the investigation community [7]. As described in numerous current studies, machine learning techniques have prospective in giving high accuracy in classification as linked to other processes for data classification. Accomplishing noticeable accuracy in prediction is vital as it can lead to an appropriate protection. Prediction accurateness may vary conditionally on different learning techniques. Consequently, it is critical to recognize contrivances proficient of producing high accuracy of prediction in heart diseases. Prediction accuracy accomplished in the commenced work is matched alongside the previous research work. The intent of this effort is to conceive a workflow of classification approaches based on machine learning methods for the operative diagnosis of heart disease. Machine learning classification is the most operational assessment creation methods for the real-world and scientific situations. The persistence is also to assess the behavior of different machine learning techniques for the classification of patients having presence and absence of heart disease. Furthermore, enactment of these methods has been estimated on different classification performance indices. In lieu of this, six machine learning techniques have been applied including artificial neural network

✉ Ashok Kumar Dwivedi
dwivedi.ashok@gmail.com

[1] Department of Bioinformatics, Computer Applications and Mathematics, Maulana Azad National Institute of Technology, Bhopal, MP 462003, India

(ANN), support vector machine (SVM), logistic regression, k-nearest neighbor (kNN), classification tree and Naïve Bayes. Moreover, the performance was compared using receiver operative characteristic (ROC) and calibration graph. Machine learning has been applied previously in biomedical field like in diabetic prediction [1–3], for association of heart disease and diabetes [4], in analysis of diabetes proteins [5], etc.

2 Materials and methods

StatLog heart disease dataset [6] available at UCI machine learning laboratory was used in this study [7]. This dataset consists of 270 samples with 150 samples without heart disease (absence) and 120 samples with heart disease (presence). For predicting heart diseases, 13 distinct parameters have been taken into account such as

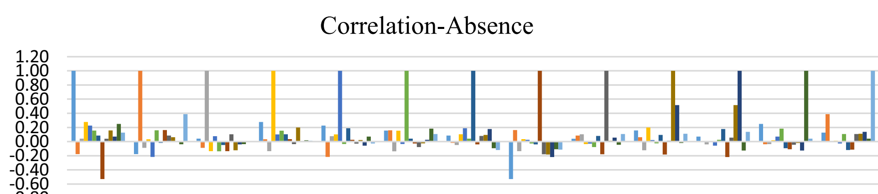
1. Age
2. Sex
3. Chest pain type (four values)
4. Resting blood pressure
5. Serum cholesterol in mg/dl
6. Fasting blood sugar > 120 mg/dl
7. Resting electrocardiographic results (values 0, 1 and 2)
8. Maximum heart rate achieved
9. Angina induced by Exercise
10. Peak Old = ST depression tempted by workout comparative to rest
11. Slant of the peak exercise ST segment
12. Numeral of major vessels (0–3) colored by fluoroscopy
13. Thal: 3 = normal; 6 = fixed defect; 7 = reversible defect

Data sample comprised of total of 270 samples with 13 features for each samples. Samples with the nonappearance of heart disease were treated as positive class, and samples with presence of heart disease were treated as negative samples for analysis purpose. The correlation between thirteen parameters of "absence" and "presence" samples indicates the high correlation between the parameters of these two classes of samples as shown in Figs. 1 and 2. Figure 1 clearly indicates that age parameter is positively correlated with all other parameters except "sex" and "max heart rate" in samples without heart disease (absence) Fig. 1. Similarly, positively correlation between other parameters in "absence" class can be observed in Fig. 1. Similarly, higher bars as positive side and lower number of small bars at negative sides in correlation plots of parameters indicate that these parameters are highly correlated with each other in "presence" class as well Fig. 2.

3 Description of the algorithms

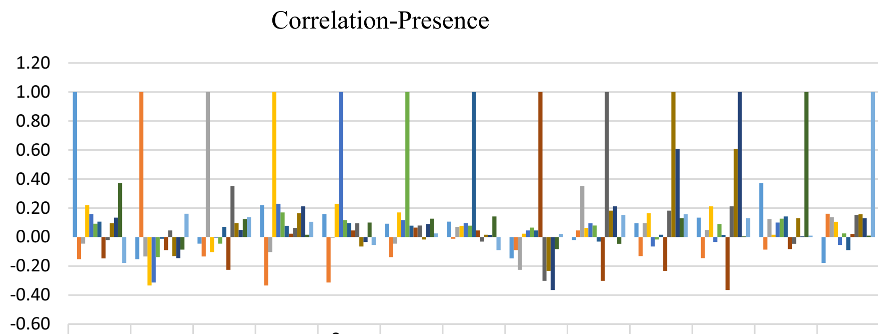
3.1 Artificial neural network

Artificial neural network (ANN) models endeavor to impressionist the organization and functions of organic neural networks. Basic construction of every ANN is



	Age	Sex	Chest-Pain-Type	RBP	Serum Cholestrol	BloodSugar	ECG	MaxHeartRate	Angina	OldPeak	SlopeST	Vessels	TNFR
Age	1.00	-0.18	0.04	0.28	0.23	0.16	0.08	-0.53	0.04	0.16	0.07	0.25	0.13
Sex	-0.18	1.00	-0.09	0.03	-0.22	0.16	-0.02	0.16	0.08	0.06	0.00	-0.04	0.39
Chest-Pain-Type	0.04	-0.09	1.00	-0.14	0.08	-0.14	-0.05	-0.14	0.10	-0.12	-0.04	-0.04	0.00
RBP	0.28	0.03	-0.14	1.00	0.10	0.15	0.10	0.04	-0.04	0.20	0.00	0.02	0.01
Serum Cholestrol	0.23	-0.22	0.08	0.10	1.00	-0.03	0.19	0.03	-0.03	0.02	-0.06	0.07	-0.03
BloodSugar	0.16	0.16	-0.14	0.15	-0.03	1.00	0.04	-0.03	-0.08	-0.03	0.03	0.18	0.11
ECG	0.08	-0.02	-0.05	0.10	0.19	0.04	1.00	-0.04	0.08	0.09	0.18	-0.09	-0.12
MaxHeartRate	-0.53	0.16	-0.14	0.04	0.03	-0.03	-0.04	1.00	-0.18	-0.18	-0.22	-0.11	-0.11
Angina	0.04	0.08	0.10	-0.04	-0.03	-0.08	0.08	-0.18	1.00	0.01	0.06	-0.05	0.11
OldPeak	0.16	0.06	-0.12	0.20	0.02	-0.03	0.09	-0.18	0.01	1.00	0.52	-0.02	0.11
SlopeST	0.07	0.00	-0.04	0.00	-0.06	0.03	0.18	-0.22	0.06	0.52	1.00	-0.13	0.14
Vessels	0.25	-0.04	-0.04	0.02	0.07	0.18	-0.09	-0.11	-0.05	-0.02	-0.13	1.00	0.04
TNFR	0.13	0.39	0.00	0.01	-0.03	0.11	-0.12	-0.11	0.11	0.11	0.14	0.04	1.00

Fig. 1 Linear correlation absence



	Age	Sex	Chest-Pain-Type	RBP	Serum Cholestrol	Blood Sugar	ECG	MaxHeartRate	Angina	OldPeak	Slope ST	Vessels	TNFR
Age	1.00	-0.15	-0.05	0.22	0.16	0.09	0.11	-0.15	-0.02	0.09	0.13	0.37	-0.18
Sex	-0.15	1.00	-0.13	-0.33	-0.31	-0.14	-0.01	-0.09	0.04	-0.13	-0.15	-0.09	0.16
Chest-Pain-Type	-0.05	-0.13	1.00	-0.10	-0.01	-0.05	0.07	-0.23	0.35	0.10	0.05	0.12	0.14
RBP	0.22	-0.33	-0.10	1.00	0.23	0.17	0.08	0.02	0.06	0.16	0.21	0.02	0.10
Serum Cholestrol	0.16	-0.31	-0.01	0.23	1.00	0.12	0.10	0.05	0.09	-0.07	-0.03	0.10	-0.05
BloodSugar	0.09	-0.14	-0.05	0.17	0.12	1.00	0.08	0.06	0.08	-0.02	0.09	0.13	0.02
ECG	0.11	-0.01	0.07	0.08	0.10	0.08	1.00	0.05	-0.03	0.02	0.01	0.14	-0.09
MaxHeartRate	-0.15	-0.09	-0.23	0.02	0.05	0.06	0.05	1.00	-0.30	-0.23	-0.37	-0.08	0.02
Angina	-0.02	0.04	0.35	0.06	0.09	0.08	-0.03	-0.30	1.00	0.18	0.21	-0.05	0.15
OldPeak	0.09	-0.13	0.10	0.16	-0.07	-0.02	0.02	-0.23	0.18	1.00	0.61	0.13	0.16
SlopeST	0.13	-0.15	0.05	0.21	-0.03	0.09	0.01	-0.37	0.21	0.61	1.00	0.00	0.13
Vessels	0.37	-0.09	0.12	0.02	0.10	0.13	0.14	-0.08	-0.05	0.13	0.00	1.00	0.01
TNFR	-0.18	0.16	0.14	0.10	-0.05	0.02	-0.09	0.02	0.15	0.16	0.13	0.01	1.00

Fig. 2 Linear correlation presence

artificial neuron, to be exact, a simple mathematical function. These models have three sets of rules: summation, multiplication and activation. At the entry of artificial neuron, the inputs are weighted and it means that every input value is multiplied with diverse weight. In the inner of artificial neuron is a sum function that sums all weighted bias and inputs. At the production of artificial neuron, the summation of earlier weighted inputs and bias is passed through an activation function (transfer function).

The yield of a characteristic ANN with K constituents is given by Eq. (1) [8, 9]:

$$y(x) = \sum_{i=1}^{k} w_i y_i(x) \tag{1}$$

where y_i is the output of net i and w_i is the weight linked with the net.

ANN may have different architectures; in this exertion, multilayer perceptron model (MLP) was applied [10, 11]. MLP is organized in layers like a multistage directed graph. Every node at every layer acquires an input from the linked node of preceding layer, and then it calculates value of a function and delivers input to the associated node in the following layer. The layers are labeled as "input layer", "hidden layers" and "output layers". The intermediary layers that do not have straight linking with input and output are named as hidden layers. The initiation of hidden layers and output layers is calculated by a function, which is a weighted summation of the inputs they receive, and then passes this through an activation function. Neural network has extensive applications in various fields like forecasting, commercial modeling, economics, medicinal applications, etc. [12–14]. These methods also have been useful in the field of bioinformatics [15–17]. Further solicitations of ANN models have been discussed in [18] and in references within.

3.2 Support vector machine method for classification

V. Vapnik first introduced support vector machine (SVM) in his work of theory of statistical learning [19, 20]. SVM is a technique of supervised machine learning used for classification and regression. The purpose of these approaches is to resolve problems unswervingly devoid of solving any intermediary problem. SVM integrates the competence to overcome the problem of over-fitting by expending the concept of structural risk minimization. Here, SVM was used for binary classification having two categories Absence and Presence of heart disease, for $y_i = +1, -1$ correspondingly. Technique can be directly extended for multiclass classification by constructing multiple two-class classifiers [19]. The SVM classifiers examine for the optimum extrication hyperplane which in between from the two classes [21]. This optimum unscrambling hyperactive plane has many satisfactory statistical characteristics. SVC is defined first for the linearly distinguishable case. Competences of SVC can be additionally extended using kernel trick of building nonlinear decision planes. Finally, difficulties of noisy data can be controlled by familiarizing slack variables.

3.3 Naïve Bayes classifier

One of the best operative classifiers is Bayesian network [22–26]. These networks are made up of network like structures with accompanying conditional probabilities. The BN construction is directed acyclic graph wherein nodes resemble to province variables and edges among nodes resemble to dependences between variables. Henceforth, Naïve Bayes classifiers are techniques of finding the suitable classification for a dataset wherever definite fundamental conventions are met.

3.4 Logistic regression classifier

Logistic regression is a discriminative classification technique [27] that works on real-valued input vector. The measurement of input vector, to be categorized, is known as features or predictors. Logistic regression can also be applied with multiclass classification. In logistic regression, probability P of a dichotomous event can be deliberated as rising from Bernoulli trial and can be associated with examining event [27–30].

3.5 *k*-nearest neighbor

The k-nearest neighbor (kNN) is an instance-based learning technique that does not frame a comprehensive theoretical model commencing the training examples. As an alternative, it uses the somewhat simpler conception and the instance nearby in the input space is probably to fit in the similar class [31].

Symptomatically, a kNN categorizes an example to the class, which seems most persistently among its k nearby neighbors. k is a constraint for fine-tuning the classification work.

3.6 Classification trees

Classification tree is a method that characterizes the data in the form of n-ary hierarchy through every node, and branch has a definite accompanying outcome, probability and weights. Root of the tree is designated normally by computing the entropy or by its converse, the information gained.

3.7 Classification performance measurement

Six supervised machine learning techniques were applied for the classification of heart disease samples. Tenfold cross validation was for the evaluation of classification performance. The classification models were evaluated on the eight quality measures [32]. These quality measures for the analysis of classification were investigated. Samples with absence of heart disease were considered as positive class, and samples with presence of heart disease were considered as negative class. Here,

True positive (TP)—number of samples with absence of heart disease predicted as absence of heart disease.

False positive (FP)—number of samples with presence of heart disease predicted as absence of heart disease.

True negative (TN)—number of samples with presence of heart disease predicted as presence of heart disease.

False negative (FN)—number of samples actually have absence of heart disease predicted as presence of heart disease.

The quality measures are defined as below:

Classification accuracy The ratio of occurrences that are appropriately categorized by the classification learner. Means ratio of suitably predicted samples to total number of examples.

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

Sensitivity, recall or true positive rate The ratio of perceived positive instance with the total positive instances, e.g., the ratio of total predicted absence with total absence.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$

Specificity or true negative rate Specificity is the proportion of perceived negative samples with all negative samples, means ratio of predicted presence with total samples with presence of heart disease.

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

Precision or positive predictive value Precision is the ratio of true positive (absence classified as absence) with all instances classified as positive (total samples classified as absence).

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$

Negative predictive value Ratio of correctly predicted negative (presence of heart disease) to the total predicted negative (total samples classified as presence of heart disease).

$$\text{NPV} = \text{TN}/(\text{TN} + \text{FN})$$

False positive rate

$$\text{FP rate} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

Rate of misclassification Proportion of incorrectly classified samples to the total number of samples is known as rate of misclassification. There are two types of incorrect classifications. If presence of heart disease was classified as absence of heart disease (E1), it is defined as type I error. If absence of heart disease was classified as presence of heart disease (E2), it is demarcated as type II error.

$$\text{Rate of misclassification} = (\text{E1} + \text{E2})/\text{total number of samples.}$$

3.8 F1 measure

F measure is the harmonic mean among precision and recall. For the best performance, its requisite is one, and for foulest performance, it is zero.

$$F_1 = \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}.$$

4 Results and discussion

The performance of six machine learning methods was assessed for the prediction of heart disease using 13 parameters discussed in methods and materials section. Total 270 samples with 150 with absenteeism of heart disease and 120 samples with incidence of heart disease were taken into account. Data samples were partitioned in tenfold, every fold was applied in testing, and remaining folds were used as training during cross validation [33].

Table 1 Confusion matrix for tenfold cross validation using logistic regression

Logistic regression			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	133 (85.3 %)	17 (14.9 %)	150
Presence	23 (14.7 %)	97 (85.1 %)	120
Total predicted	156	114	270

Learning parameters—training error cost (C): 1.0, regularization-type: L1 (absolute weights), normalization: yes

Confusion matrix of prediction results is tabulated in Table 1 through Table 6 for logistic regression, support vector machine (SVM), artificial neural network (ANN), *k*-nearest neighbor (kNN), classification tree and Naïve Bayes, respectively.

Figure 3 shows predictions of these machine learning models. It is evident from the results that logistic regression predicts highest number of true positives (absence of heart disease predicted as absence) (Table 1; Fig. 3) and SVM predicts highest number of true negatives (presence of heart disease predicted as presence of heart disease) (Table 2; Fig. 3).

ANN confusion matrix (Table 3) indicates that it has second highest true positives (Fig. 3).

Table 4 shows the confusion matrix of Naïve Bayes classifier, which specifies that this classifier gives third highest number of true positives and true negatives (Fig. 3).

Tables 5 and 6 display the confusion matrix of classification tree and kNN, respectively, which designate that these two are worst performer in the sense of lowest TN.

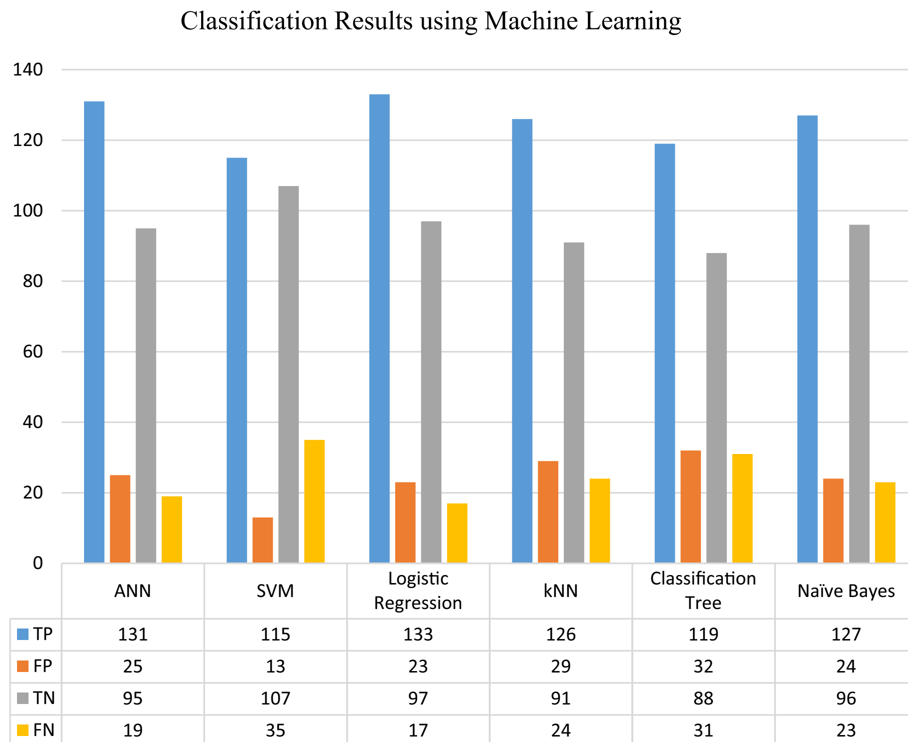
Table 7 elucidates different classification recital measurements, and Fig. 4 plots four classification performance measurements specifically classification accuracy, specificity, sensitivity and precision.

Figure 4 specifies that logistic regression outperformed over all other machine learning methods with maximum classification accuracy of 85 % while second highest classification accuracy is achieved by ANN (84 %). Additionally, these two methods have shown utmost sensitivity of 89 and 87 % correspondingly. SVM achieves highest specificity of 89 % which indicates that this classifier is most suitable for identification of patients with heart disease (presence class). Moreover, SVM also has the highest precision of 90 %.

Figure 5 associates negative predictive value, rate of misclassification and F1 measures of these machine learning methods. This graph evidently demonstrates that kNN has highest negative predictive value of 85 % whereas it also beats all other techniques on F1 measure. SVM shows lowest FP rate of 0.11, and kNN has lowest rate of misclassification (15 %).

4.1 Performance evaluation using ROC

Another measure for classification is ROC (receiver operating curve) [34], which is a curvature connived on FP ratio (*x*-axis) and TP ratio (*y*-axis). ROC is unbiased of both classes and valuable when the number cases of both classes



	ANN	SVM	Logistic Regression	kNN	Classification Tree	Naïve Bayes
TP	131	115	133	126	119	127
FP	25	13	23	29	32	24
TN	95	107	97	91	88	96
FN	19	35	17	24	31	23

Fig. 3 Classification results of machine learning techniques, true positive (TP)—number of samples with absence of heart disease predicted as absence of heart disease. False positive (FP)—number of samples with presence of heart disease predicted as absence of heart disease. True negative (TN)—number of samples with presence of heart disease predicted as presence of heart disease. False negative (FN)—number of samples actually have absence of heart disease predicted as presence of heart disease

Table 2 Confusion matrix for tenfold cross validation using support vector machine (SVM)

SVM			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	115 (89.8 %)	35 (24.6 %)	150
Presence	13 (14.7 %)	107 (85.1 %)	120
Total predicted	128	142	270

Learning parameters—kernel: RBF, $e-^{0.0050*(x-y).\ (x-y)}$, cost (C): 0.1, numeric precision: 0.0005, estimation of class probabilities: yes, normalization of data: yes

Table 3 Confusion matrix for tenfold cross validation using artificial neural network (MLP BP)

ANN			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	131 (84.0 %)	19 (16.7 %)	150
Presence	25 (16.0 %)	95 (83.3 %)	120
Total predicted	156	114	270

Hidden layer neurons: 10, regularization factor: 1.5, max iterations: 2000

Table 4 Confusion matrix for tenfold cross validation using *k*-nearest neighbors (kNN)

kNN			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	126 (81.3 %)	24 (20.9 %)	150
Presence	29 (18.7 %)	91 (79.1 %)	120
Total predicted	155	115	270

Learning parameters—metrics: euclidean, continuous attributes: normalized, unknown values ignored: no, number of neighbors: 5, weighting: by distances

Table 5 Confusion matrix for tenfold cross validation using classification trees

Classification trees			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	119 (78.8 %)	31 (26.1 %)	150
Presence	32 (21.2 %)	88 (73.9 %)	120
Total predicted	151	119	270

Learning parameters—attribute selection: information gain, binarization: no binarization, pruning: two instances in leaves, recursively merge leaves with same majority class: yes, pruning with m-estimate: $m = 2$

Table 6 Confusion matrix for tenfold cross validation using Naïve Bayes

Naïve Bayes			
	Predicted class		
	Absence	Presence	Actual total
Actual class			
Absence	127 (84.1 %)	23 (19.1 %)	150
Presence	24 (15.9 %)	96 (80.7 %)	120
Total predicted	151	119	270

Learning parameters—probability estimation: relative frequency, LOESS window size: 0.5, number of points in LOESS: 100, adjust classification threshold: no

fluctuate through training. Area under ROC must be near to 1 for best classifier. Figure 6 designates that SVM beats all other techniques in prediction of absence of heart disease whereas logistic regression and ANN beat other methods in the prediction of presence of heart disease (Fig. 7).

5 Conclusion

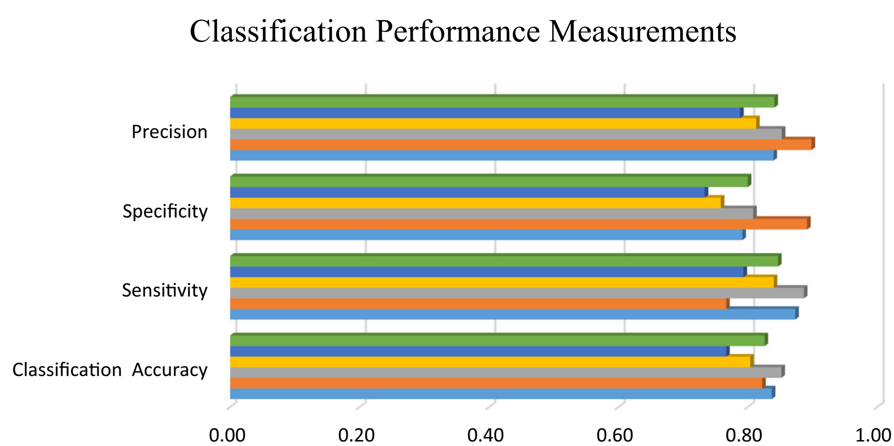
Prediction of heart disease may save the life of human beings and can have significant impact on its treatment. This exertion provides a workflow founded on machine learning techniques for the prediction of presence or absence of heart disease. This examination has affianced

Table 7 Classification performance measure indices for tenfold cross validation using machine learning techniques

	CA	Sens.	Spec.	Pre.	NPV	FPR	RMC	F1
ANN	0.84	0.87	0.79	0.84	0.83	0.21	0.16	0.86
SVM	0.82	0.77	0.89	0.90	0.75	0.11	0.18	0.83
Logistic regression	0.85	0.89	0.81	0.85	0.85	0.19	0.15	0.87
kNN	0.80	0.84	0.76	0.81	0.79	0.24	0.20	0.83
Classification tree	0.77	0.79	0.73	0.79	0.74	0.27	0.23	0.79
Naive Bayes	0.83	0.85	0.80	0.84	0.81	0.20	0.17	0.84

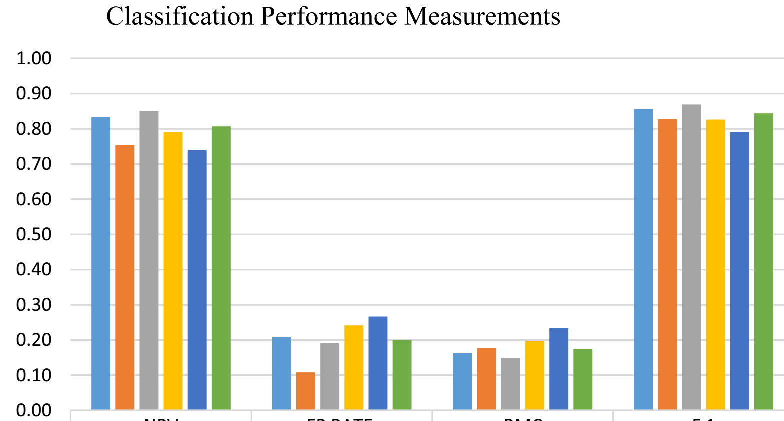
Classification accuracy (*CA*), sensitivity (*Sens.*), specificity (*Spec.*), precision (*Pre.*), negative predictive value (*NPV*), false positive rate (*FPR*), rate of misclassification (*RMC*), F1 measure (*F1*)

Fig. 4 Classification performance measurements of machine learning algorithm



	Classification Accuracy	Sensitivity	Specificity	Precision
Naïve Bayes	0.83	0.85	0.80	0.84
Classification Tree	0.77	0.79	0.73	0.79
kNN	0.80	0.84	0.76	0.81
Logistic Regression	0.85	0.89	0.81	0.85
SVM	0.82	0.77	0.89	0.90
ANN	0.84	0.87	0.79	0.84

Naïve Bayes Classification Tree kNN Logistic Regression SVM ANN



	NPV	FP RATE	RMC	F 1
ANN	0.83	0.21	0.16	0.86
SVM	0.75	0.11	0.18	0.83
kNN	0.85	0.19	0.15	0.87
Naïve Bayes	0.79	0.24	0.20	0.83
Logistic Regression	0.74	0.27	0.23	0.79
Classification Tree	0.81	0.20	0.17	0.84

Fig. 5 Classification performance measurements of machine learning algorithm

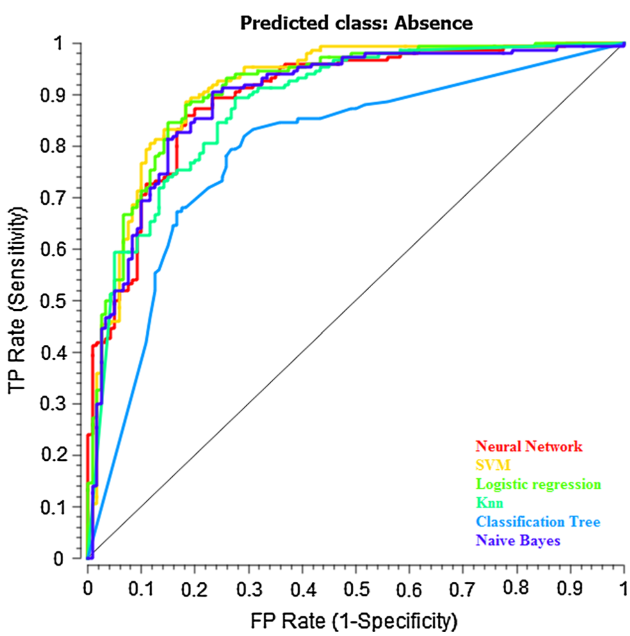


Fig. 6 ROC for six classification techniques tested for absence

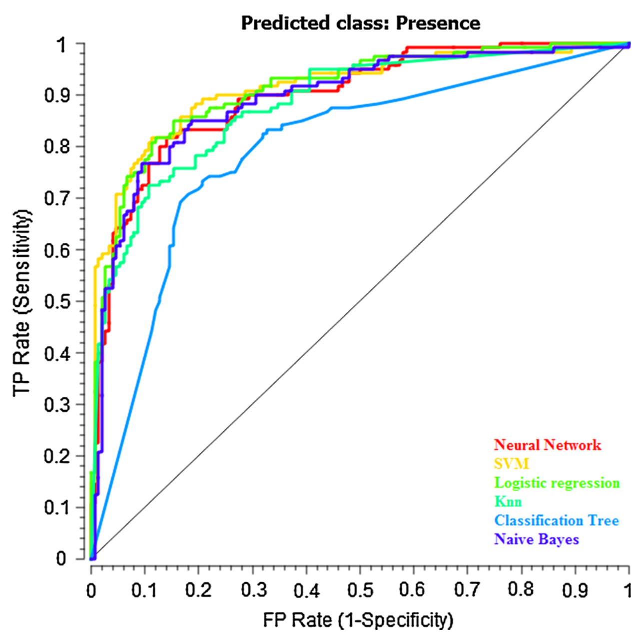


Fig. 7 ROC for five classification techniques tested for positive samples

six classical machine learning methods for prophecy of heart disease based on numerous meticulous parameters. These methods were validated using tenfold cross validation and evaluated in terms of various performance measurements. Furthermore, this exertion can be extended for prediction of heart disease by gathering the measurable from various clinics that can provide the more exhaustive model. The recitation can be used for operational heart disease predictions.

Acknowledgments I am greatly thankful to Department of Biotechnology, New Delhi, for providing Bioinformatics Infrastructure Facility of DBT at Maulana Azad National Institute of Technology, Bhopal.

Compliance with ethical standards

Conflict of interest None.

References

1. Farran B, Channanath AM, Behbehani K, Thanaraj TA (2013) Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait—a cohort study. BMJ open 3(5):e002457
2. Heydari M, Teimouri M, Heshmati Z, Alavinia SM (2015) Comparison of various classification algorithms in the diagnosis of type 2 diabetes in Iran. Int J Diabetes Dev Ctries 36(2):1–7
3. Bansal A, Agarwal R, Sharma R (2015) Determining diabetes using iris recognition system. Int J Diabetes Dev Ctries 35(4):432–438
4. Kalaiselvi C, Nasira GM (2015) Classification and prediction of heart disease from diabetes patients using hybrid particle swarm optimization and library support vector machine algorithm. Int J Comput Algorithm 4(1):2278–2397
5. Bhramaramba R, Allam AR, Kumar VV, Sridhar G (2011) Application of data mining techniques on diabetes related proteins. Int J Diabetes Dev Ctries 31(1):22–25
6. King RD (1992) Statlog databases. Department of Statistics and Modelling Science, University of Strathclyde, Glasgow
7. Bache K, Lichman M (2013) UCI machine learning repository [http://archive.ics.uci.edu/ml]. University of California, School of Information and Computer Science. Irvine, CA
8. García-Pedrajas N, Hervás-Martínez C, Ortiz-Boyer D (2005) Cooperative coevolution of artificial neural network ensembles for pattern classification. Evolut Comput IEEE Trans 9(3):271–302
9. Yao X, Liu Y (1998) Making use of population information in evolutionary artificial neural networks. Syst Man Cybern Part B: Cybern IEEE Trans 28(3):417–425
10. Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Cambridge
11. Haykin S (2010) Neural networks: a comprehensive foundation, 1994. Mc Millan, New Jersey
12. Bahrammirzaee A (2010) A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Comput Appl 19(8):1165–1195
13. Hoptroff RG (1993) The principles and practice of time series forecasting and business modelling using neural nets. Neural Comput Appl 1(1):59–66
14. Azar AT (2013) Fast neural network learning algorithms for medical applications. Neural Comput Appl 23(3–4):1019–1034
15. Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M, Haussler D (2000) Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc Natl Acad Sci 97(1):262–267
16. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D (2000) Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 16(10):906–914
17. Ranawana R, Palade V (2005) A neural network based multi-classifier system for gene identification in DNA sequences. Neural Comput Appl 14(2):122–131
18. Yasdi R (2000) A literature survey on applications of neural networks for human-computer interaction. Neural Comput Appl 9(4):245–258
19. Vapnik VN, Vapnik V (1998) Statistical learning theory, vol 2. Wiley, New York
20. Vapnik V (2000) The nature of statistical learning theory. Springer, Berlin
21. Burges CJ (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2(2):121–167
22. Shafer G, Pearl J (1990) Readings in uncertain reasoning. Morgan Kaufmann Publishers Inc., Burlington
23. Heckerman D, Geiger D, Chickering DM (1995) Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn 20(3):197–243
24. Jensen FV (1996) An introduction to Bayesian networks, vol 210. UCL press, London
25. Peral J (1988) Probabilistic reasoning in intelligent systems. Morgan Kaufmann, San Mateo, Cali fornia 12:241–288
26. Castillo E (1997) Expert systems and probabilistic network models. Springer, Berlin
27. Hosmer DW Jr, Lemeshow S (2004) Applied logistic regression. Wiley, New York
28. Schumacher M, Roßner R, Vach W (1996) Neural networks and logistic regression: part I. Comput Stat Data Anal 21(6):661–682
29. Vach W, Roßner R, Schumacher M (1996) Neural networks and logistic regression: part II. Comput Stat Data Anal 21(6):683–701
30. Hajmeer M, Basheer I (2003) Comparison of logistic regression and neural network-based classifiers for bacterial growth. Food Microbiol 20(1):43–55
31. Aha DW (1997) Lazy learning. Kluwer academic publishers, Berlin
32. Kanmani S, Uthariaraj VR, Sankaranarayanan V, Thambidurai P (2007) Object-oriented software fault prediction using neural networks. Inf Softw Technol 49(5):483–492
33. Geisser S (1993) Predictive inference, vol 55. CRC Press, Boca Raton
34. Metz CE (1978) Basic principles of ROC analysis. In: Freeman LM (ed) Seminars in nuclear medicine. vol 4. Elsevier, pp 283–298