An Ensemble based Extreme Learning Machine for Cardiovascular Disease Prediction

[1]R. Subha, [2]K. Anandakumar, [3] A. Bharathi

[1]Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, Tamil Nadu, India.
[2]Assistant Professor, Department Computer Science Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu, India.

[3]Professor, Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu, India.

Abstract

Today, numerous doctors manage healthcare data utilizing medicinal services data framework it contains huge measure of information, used to extract hidden information for prediction of cardiovascular disease. Heart disease is the main reason for death in the world over the last decade. The main objective of this research is to develop a Heart Disease Prediction System to predict the presence of heart disease using machine learning algorithms. Extreme Learning Machine (ELM) is a new class of Single-Hidden Layer Feed Forward Neural Network (SLFN), which is simple in theory and fast in implementation and it reported that it suffers from over fitting. These conventional classification methods are reported to have several limitations in their diagnostic ability. To rectify these problems in ELM, this work proposed an Ensemble ELM. Consequently, proposal of constructing an ensemble of several predictors on the training set using various sets of random parameters where the parameters of each predictor are updated according to particular criterions, and then making decisions for testing samples through majority voting with the ensemble. The experimental results show that proposed method outperforms well and provides better classification accuracy to predict the presence of cardiovascular disease.

Keywords: Cardiovascular Disease, Diagnosis, Prediction, Extreme Learning Machine and Weighted Extreme Learning Machine.

INTRODUCTION

Heart disease is the main source of death for the two men and women in the world over recent years, with the greater part of the death happening in men [13]. One in every four people is afflicted with and dies of heart disease, and in the United States, over 610,000 afflicted Americans lose their lives annually. It is known that heart is essential organ in human body part if that organ gets affected then it also affects the other vital parts of the body. It is just a pump, which pumps blood through the body. In the event that circulation of blood in body is wasteful the organs like brain endure and if heart quits working inside and out, death happens within minutes. Life is totally subject to effective working of the heart. The term Heart disease alludes to ailment of heart and blood vessel system inside it. The factors have been shown that increases the risk of Heart disease such as Family history, Smoking, Poor diet, High blood pressure, High blood cholesterol, Obesity, physical inactivity and Hyper tension. Factors like these are utilized to examine the Heart disease [11]. Much of the time, diagnosis is for the most part in view of patient's present test outcomes and doctor's experience. Accordingly, the diagnosis is a complex task that requires much experience and high expertise.

The heart circulatory system is composed of the heart and blood vessels, including arteries, veins, and capillaries. The term 'cardiovascular ailment' that speaks to a category of heart disease includes a wide assortment of conditions that furious the heart and the blood vessels and the route in which blood is pumped and flowed in the body. Cardio vascular disease affects the heart circulatory system and damages the system also damages the valves resulting in heart attack or heart failure [8]. To keep away from such a circumstance a clinical expert system is produced to expert system the heart disease and to diminish the level heart failure and death. Therefore, it is very important for people predict the heart disease by an automated method of classification techniques.

Classification techniques of Machine Learning Algorithms [4] play a significant role in Prediction. Machine learning algorithm [17] can altogether help in tackling the medicinal services issues by creating classifier frameworks that can help doctors in expecting sicknesses in beginning times. In any case, extracting information from huge data it might be heterogeneous, unorganized and high dimensional and may contain noise and outliers. Most appropriate Extreme Learning Machine has been chosen and validating their performances in terms of accuracy and precision.

Problem Statement

In present world there are many scientific technologies which help doctors in taking clinical decisions but they might not be accurate. Heart disease prediction system can help therapeutic experts in predicting condition of heart, in light of the clinical data of patients fed into the system. Doctors may in some cases neglect to take accurate decisions while diagnosing the heart disease of a patient, consequently heart disease prediction systems which utilize machine learning algorithms aid such cases to get accurate outcomes. There are many tools available which use prediction algorithms but they have some flaws. The

vast majority of the devices can't deal with enormous information and most are not incorporated, not sent on cloud and subsequently not open on the web. There are many hospitals and healthcare industries which collect huge amounts of patient data which becomes difficult to handle with currently existing systems. The objective of this work is to predict more accurately the presence of heart disease with reduced number of attributes and deploying the algorithms to overcome the existing limitations.

The rest of the paper is organized as follows. Section II deals the related works about the cardio vascular system and their prediction issues and the existing methods. Section III deals with the Extreme Learning Machine and proposed method Ensemble Extreme Learning. Section IV, gives the result and performance analysis. Finally, the overall proposed method concludes in section V

RELATED WORK

Heart disease is the main source of death in the world over recent years. Researchers have been utilizing a few data mining systems in the diagnosis of heart disease. Support vector machine are a cutting-edge method in the field of machine learning and have been effectively utilized as a part of various fields of use. Parthiban *et.al* (2012) [4] uses classification algorithm like Naive Bayesian and Support vector machine used for prediction utilizing attributes from diabetic's diagnosis to discover whether diabetic patient is experiencing heart disease with showing levels. From the experimental results obtained, it can be seen that the classifier displays a high classification accuracy i.e. 94.60% generally speaking. Subsequently this SVM model can be prescribed for the classification of the diabetic dataset.

Patel *et.al* (2015) [5] presents new model that improves the Decision Tree accuracy in distinguishing heart disease patients. It thinks about various algorithms of Decision Tree classification looking for better execution in heart disease analysis utilizing WEKA. The algorithms which are tried by J48 algorithm, Logistic model tree algorithm and Random Forest algorithm. The objective of this investigation is to extract hidden patterns by applying data mining strategies, which are imperative to heart diseases and to anticipate the nearness of heart disease in patients. The current datasets of heart disease patients from Cleveland database of UCI vault is utilized to test and justify the performance of decision tree algorithms.

Cardio vascular disease influences the heart circulatory system and damages the system valves bringing about heart attack or heart failure. To stay away from such a circumstance a clinical expert system is produced to distinguish CVD ahead of time and to decrease the level heart failure and death. Kumar *et.al* (2014) built up a programmed framework for the classification of ICU patients utilizing ANN techniques for decision-making is executed. The basic leadership was performed utilizing highlights extricated from ECGs. This expert system implements the neural network to diagnose the heart diseases. The back propagation algorithm is utilized to prepare the neural network for diagnosing the cardio vascular illness and to take

preparatory activities. The result of training process will be the error level related with the original data. In view of the error level a choice will be taken that a patient has this specific level of risk related with him. The proposed approach exhibited a superior performance in terms of classification accuracy and also simple to implement and use, as it only requires the ECG signal to determine the patients' states.

ELM and its enhanced variant are just in view of the empirical risk minimization principle, which may experience the ill effects of over fitting. Mao *et.al* (2014) [6] consolidated the basic risk minimization standard into the (weighted) ELM. The M-WELM can be summed up to cost sensitive learning and can likewise manage information with imbalanced class circulation as the WELM. Then again, it's over fitting risk can be diminished by considering both the observational and structural risks simultaneously. From the experimental results, M-WELM algorithm demonstrates the best execution against the other revealed ELM calculation and SVR algorithm, especially when using less training samples.

Prerana *et.al* (2015) designed an algorithm for accurate prediction of heart disease risk level. PAC algorithm is built utilizing existing machine learning algorithms. Popular machine learning algorithms to decide the heart disease risk level and to help the specialists effectively anticipate the same. Hadoop single node cluster is utilized to process Big Data. Map Reduce code is actualized for the composed algorithms. At long last the comparisons between the algorithms is done which encourages the users to figure out which algorithm demonstrates the highest accuracy. The interface is easy to use and the application is all inclusive open on cloud. Depending on the increasing requirement multi nodes can be added to the cluster to decrease the execution time and process more data.

A binary classification technique, Probabilistic Extreme Learning Machine (called P-ELM) is proposed by Zhao *et.al* (2011) to improve the reliability of the classification of an unknown object. The P-ELM algorithm may restrain vulnerability of the extreme learning machine prediction in the different trials of simulation because of the introduction of input weights and bias, which would damage the reliability of the classification for the new object. ELM is coordinated with thickness techniques and Bayesian decision theory so as to consider the uncertainty of the predictions in ELM. Huang et al (2012) [2] shows that both LS-SVM and PSVM can be simplified further and a unified learning framework of LS-SVM, PSVM, and other regularization algorithms referred to Extreme Learning Machine (ELM) can be built.

Zhang *et.al* [3] proposed a fast and efficient classification technique as ELM algorithm. The ELM randomly identify the all hidden node parameters and then analytically make a decision on the output weights. It has good simplification process and it can be executed naturally. In ELM the nonlinear activation functions are used such as sigmoid, sine, hard limit, radial basis functions and complex activation functions.

METHODOLOGY

Classification is one of the most important decision-making techniques for selecting data. In this paper, the main aim of

research is to build Intelligent Heart Disease Prediction System [8] to predict the data as presence of heart disease for improving the classification accuracy. An ensemble-based ELM (EN-ELM) algorithm is introduced where ensemble learning and cross-validation is fixed into the training phase, so as to alleviate the overtraining problem and enhance the predictive stability. EN-ELM is robust and efficient for classification. The block diagram is shown in Figure 1. This section comprises of existing Extreme Learning Machine (ELM) and proposed Ensemble Extreme Learning Machine (EELM) method used for prediction of cardio vascular system by various classifiers are briefly explained. They are briefly explained below.
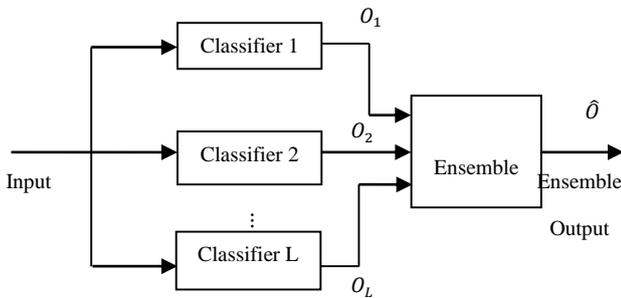

Figure 1. Ensemble ELM Structure

A. Extreme Learning Machine

Extreme Learning Machine (ELM) mainly applied for Single Hidden Layer Feed forward Neural Networks (SLFNs) it is the process of randomly selecting the input weights and systematically determines the output weights of SLFNs. This algorithm tends to the best generalization performance at extremely fast learning speed [14].

ELM has several significant features which are differ from traditional learning algorithms applied for feed forward neural networks. The learning speed of ELM could be completed in seconds or less than seconds for many traditional applications. In traditional algorithm there exists a virtual speed barrier in which the algorithms cannot process and it is not unusual way to take long time for train a feed-forward network using classic learning algorithms for uncomplicated applications [6]. The ELM has better simplification performance compared with gradient based learning algorithms such as back propagation. The gradient based learning algorithms and some other learning algorithms may face many issues such as local minima, improper learning rate and over fitting, etc. The methods are implemented to overcome the above issues such as weight decay and stopping methods.

In real applications, the number of hidden N nodes will always be less than the number of training samples N and the training error cannot be made exactly zero but can be a nonzero training error ϵ. The hidden node parameters a_i and (input weights and biases or centers and impact factors) of ELM need not be tuned during training and may simply assigned with random values according to continuous sampling distribution. If the number of neurons in the hidden layer is equal to the number of samples, then H is square and invertible. Otherwise, the system of equations needs to be solved by numerical methods, concretely by solving

$$\|H(w_1, \ldots, w_M, b_1, \ldots, b_M)\hat{\beta} - T\| = min_\beta \|\beta - T\| \text{ -------(1)}$$

The result that minimizes the norm of this least squares equation is

$$\hat{\beta} = H^\dagger T \text{ ------------------------- (2)}$$

Where $H^\dagger$ was the Moore-Penrose generalized inverse of matrix H.

The three important properties are

- Minimum training error.
- Smallest norm of weights and best generalization performance.
- The minimum norm least-square solution of Hβ = T is unique, $\hat{\beta} = H^\dagger T$

The ELM learning algorithm looks much simpler and it gives accurate result when compared to other algorithms. Extreme Learning Machines provides better solutions and possesses a unique features to deal with issues such as regression, uncertainty and (multi-class) classification tasks [3].

B. Weighted Extreme Learning Machine (WELM)

Weighted extreme learning machine for imbalance learning, which defined a N x diagonal matrix **W** associated with every training sample x_i. Usually if training data xi comes from a minority class (assumed to be positive class), the associated weight Wii will be set relatively larger than other [6]. To maximize the marginal distance and to minimize the weighted cumulative error with respect to each sample, an optimization problem mathematically are written as

Minimize $\|H\beta - T\|^2$,$\|\beta\|$ ------------------------------ (3)

Where $T = [t_{1,\ldots\ldots} t_N]$

Minimize $L_{P_{ELM}} = \frac{1}{2} \|\beta\|^2 + \frac{1}{2} CW \sum_{i=1}^{N}\|\xi_i\|^2$ --------- (4)

Subject to h (xi) β = $t_i^T - \xi_i^T$ i=1,…N

Where h(xi) is the feature mapping vector in the hidden layer with respect to xi, β represents the output weight vector connecting the hidden layer and output layer, and C is the regularization parameter to represent the trade-off between the minimization of training errors and the maximization of the marginal distance. ξi, the training error of sample xi, is caused by the difference of the desired output ti and the actual output h (xi).
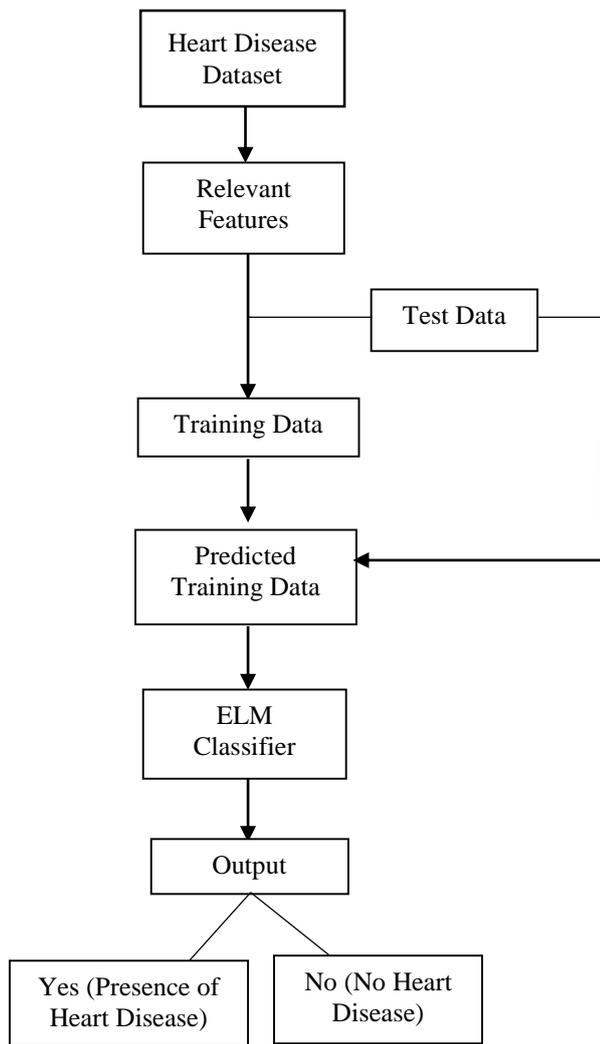
Figure 2. Flowchart for the traditional Classifiers

C. *Ensemble Classification*

Classification can be defined as the process to approximate I/O mapping from the given observation to the optimal solution [6]. Generally, classification tasks consist of two parts: feature selection and classification. Feature selection is a transformation process of observations to obtain the best pathway to get to the optimal solution. Therefore, considering multiple features encourages obtaining various candidate solutions, so that estimate more accurate solution to the optimal than any other local optima.

When it have multiple features available, it is important to know which of features should be used. Theoretically, as many features may concern, it may be more effective for the classifier to solve the problems. But features that have overlapped feature spaces may cause the redundancy of irrelevant information and result in the counter effect such as over fitting. Therefore, it is more important to explore and utilize independent features to train classifiers, rather than increase the number of features use. Correlation between feature sets can be induced from the distribution of feature numbers, or using mathematical analysis using statistics.

Meanwhile, there are many algorithms for the classification from machine learning approach, but none of them is perfect. However, it is always difficult to decide what to use and how to set up its parameters. According to the environments the classifier is embedded, some algorithm works well and others not. It is because, depending on the algorithms, features and parameters used the classifier searches in different solution space. These sets of classifiers produce their own outputs, and enable the ensemble classifier to explore more wide solution space.

We have applied this idea to a classification framework as shown in Figure 3. If there are 'k' features and 'n' classifiers, there are $k \times n$ feature-classifier combinations. There are $k \times n C_m$
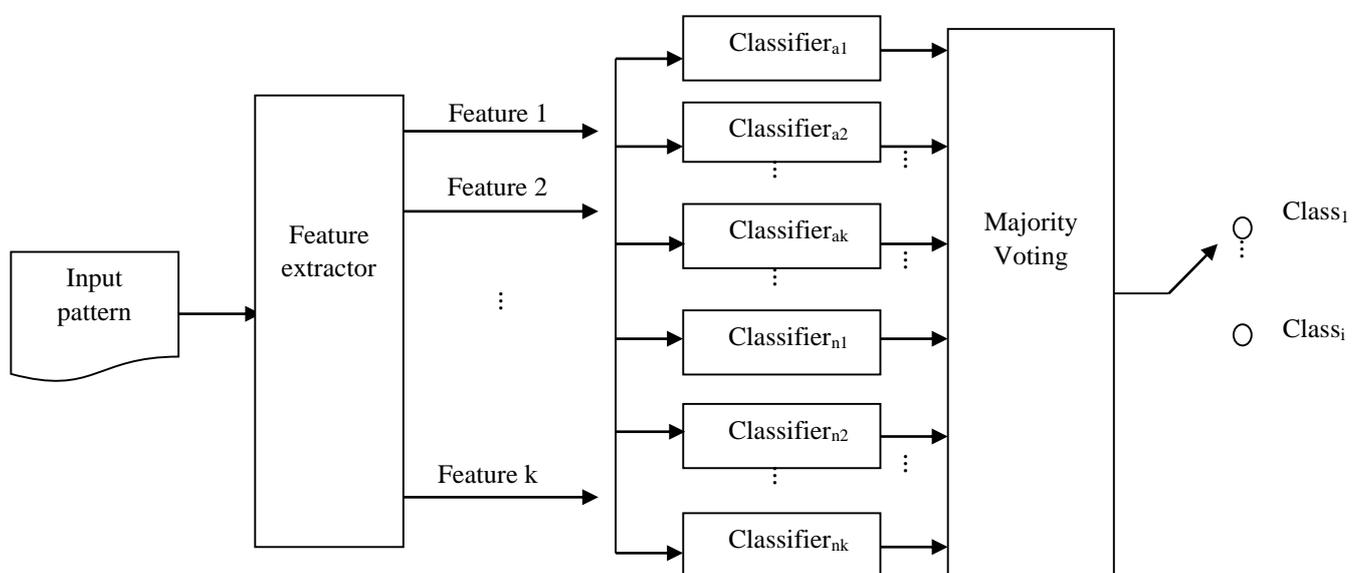

Figure 3. Flowchart for the Ensemble Classifiers

possible ensemble classifiers when 'm' feature-classifier combinations are selected for ensemble classifier. Then classifiers are trained using the features selected, finally a majority voting is accompanied to combine the outputs of these classifiers. After classifiers with some features are trained independently produce their own outputs, final answer will be judged by a combining module, where the majority voting method is adopted.

The traditional extreme learning machines are based on the empirical risk minimization principle and the training error minimization principle, whose drawback is that it is likely to suffer from over fitting, which reduces the generalization capability consequently. According to the statistical theory, the actual risks include the empirical and structural risks, and a model with good generalization performance should be able to balance empirical and structural risks to obtain the best compromise. The Ensemble principle into the ELM algorithm and proposed Ensemble ELM model are employed.

The steps of the proposed EELM algorithm can be summarized as follows.

Algorithm: EN-ELM

Inputs

Sequence of N examples $L = \{(x_j, t_j) | x_j \epsilon R^n, t_j \epsilon R^m, j = 1,2, \ldots, N\}$

Number of iterations K, number of hidden nodes $\tilde{N}$, cross validation fold number R.

Initialization

Randomly generate w_i and b_i where i=1,2,……,$\tilde{N}$.

Partition training set into R subsets so that R groups of data are obtained. In each groups (R-1)N⁄R samples are used for training and the remaining N⁄R samples for validation.

Assign $\hat{w}_i$ and $\hat{b}_i$ with the values of w_i and b_i, respectively. Then calculate mean values of accuracies as well as $\|\beta\|$ for validation sets and store the values into $\widehat{CA}$ and $\|\hat{\beta}\|$.

For k = 1, 2, …….., K

1) Randomly generate w_i^k and b_i^k where $i = 1,2, \ldots \ldots, \widetilde{N}$
2) Calculate mean of classification accuracy CA_r^k on R subsets where r = 1,2, …….R and store the values into CA^k and the norm of β^k, where $CA^k = (1/R) \sum_{r=1}^{R} CA_r^k. \|\beta^k\| = (1/R) \sum_{r=1}^{R} \|\beta_r^k\|$
3) If CA^k<CA or $\|\beta^k\| > \|\hat{\beta}\|$

   Update parameters as $w_i^k = \hat{w}_i, b_i^k = \hat{b}_i.$.

   Else

   Set $\hat{w}_i = w_i^k, \hat{b}_i = b_i^k$.

   End if

End for

Hypothesis Ensemble Construction

For r = 1,2,… R

Given a testing instance (x, **t**), evaluate the hypothesis ensemble $\{\hat{h}_1^r, \hat{h}_2^r, \hat{h}_K^r\}$ on X,

For k=1,2,…,K

a) Obtain updated hidden node weights and biases w_i^k and b_i^k and call ELM algorithm to predict target for X using partial training data L^*, i.e., (R-1) subsets.

b) Set $v_{k,c}^r$ to 1 if $\hat{h}_K^r$ is predicted as class c, otherwise, its value is zero, where c=1,2,…,C.

End for

End for

The EELM is able to be generalized to cost sensitive learning and can also deal with data with imbalanced class distribution as the EELM. On the other hand, it's over fitting risk can be reduced by considering both the empirical and structural risks simultaneously. The obtained results are used to predict the presence of heart diseases. The performance of proposed method is evaluated in terms of certain parameters to improve the classification accuracy.

EXPERIMENTAL RESULTS

The performance of the proposed method are evaluated to predict the presence of heart disease. The heart disease data sets, which were used in this research, were obtained from the Heart Disease Databases available in the UCI Machine Learning Repository [13]. These databases contain data information on heart disease clinical instances, contributed by the Cleveland Clinic Foundation (CCF), [12] consists of 303 records. In the experiments, the samples are randomly divided into two sample groups rare 70% for training and the remaining 30% for test. The process is repeated randomly train-test procedure and to calculate the classification accuracy and prediction error of every algorithm. In which the heart disease dataset are split into two non-overlapping sets as training set and a testing set. In which training set consists of 80%, 50%, or 30% of the heart disease datasets while the testing set consists of the remaining 20%, 50%, or 70% of the heart disease dataset. The regression models are trained on the training set and the results are then tested on the testing set.

The performance of proposed method are measured under confusion matrix. The confusion matrix is obtained to calculate the accuracy of classification. A confusion matrix shows how many instances have been assigned to each class. In this experiment the two classes are considered as Yes (heart disease) and No (no heart diseases) as 2x2 confusion matrix. Table 1 shows the confusion matrix.

Table 1. Confusion Matrix

	Presence of Heart Disease	Absence of Heart Disease
Predicted Heart Disease	TP	FN
Predicted No Heart Disease	FP	TN

- TP (True Positive) - It denotes the number of records classified as true while they were actually true.
- FN (False Negative) - It denotes the number of records classified as false while they were actually true.
- FP (False Positive) - It denotes the number of records classified as true while they were actually false.
- TN (True Negative) - It denotes the number of records classified as false while they were actually false.

$$Accuracy(A) \; = \; \frac{(TP + TN)}{(TP \; + \; TN \; + \; FP \; + \; FN)}$$

$$Precision \; = \; \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

The proposed Ensemble Extreme Learning Machine are compared with Modified Weighted Extreme Learning Machine (M-WELM), Weighted Extreme Learning Machine (W-ELM) [7], Modified Extreme Learning Machine (M-ELM) [6], Extreme Learning Machine (ELM) [3], Support Vector Machine (SVM) [5], Bayesian Network [8] and K-NN [16]. Table 2 shows the performance comparison of various techniques on heart disease datasets for 80% samples are used for training in terms of accuracy, precision, recall and execution time. These performance metrics are used to evaluate the algorithms.

Table 2. Performance Comparison of EELM (80% sample for training and 20% for testing)

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	Execution Time (Seconds)	Error Rate
EELM	**99.5**	**97.5**	**70**	**0.31**	**0.01**
M-WELM	98	95.6	71	0.32	0.02
W-ELM	96.3	93.2	73.6	0.39	0.04
M-ELM	94.2	91.5	75.5	0.40	0.06
ELM	92.3	89.5	78.5	0.45	0.08
SVM	91.6	86.3	80.3	0.48	0.08
Bayesian Network	86.5	84.2	82.3	0.52	0.13
KNN	82.1	81.6	86.3	0.51	0.18

Table 2 shows the performance comparison of proposed EELM for 80% sample for training and 20% sample for testing. It is clear that the proposed EELM achieves high classification accuracy of 98% and better results to predict the presence of heart disease.

Table 3. Performance Comparison of EELM (50% sample for training and 50% for testing)

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	Execution Time (Seconds)	Error Rate
EELM	**96.7**	**93.2**	**74**	**0.37**	**0.03**
M-WELM	95.2	92.5	75	0.38	0.05
W-ELM	92.3	90.2	78.6	0.42	0.08
M-ELM	86.2	87.6	79.5	0.45	0.14
ELM	84.3	85.3	81.5	0.49	0.16
SVM	81.6	82.6	83.5	0.51	0.19
Bayesian Network	75.6	79.6	85.6	0.56	0.24
KNN	72.6	75.8	89.5	0.58	0.27

Table 3 shows the performance comparison of proposed EELM for 50% sample for training and 50% sample for testing. It is clear that the proposed EELM achieves high classification accuracy and better results to predict the presence of heart disease.

Table 4 shows the performance comparison of proposed EELM for 30% sample for training and 70% sample for testing. It is clear that the proposed EELM achieves high classification accuracy and better results to predict the presence of heart disease.

Table 4. Performance Comparison of EELM (30% sample for training and 70% for testing)

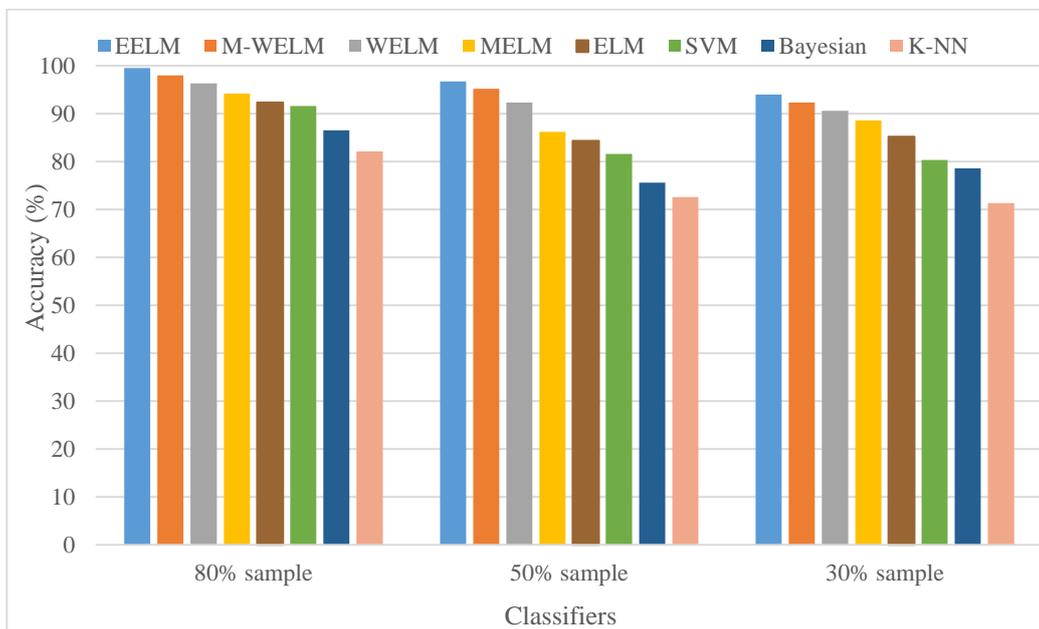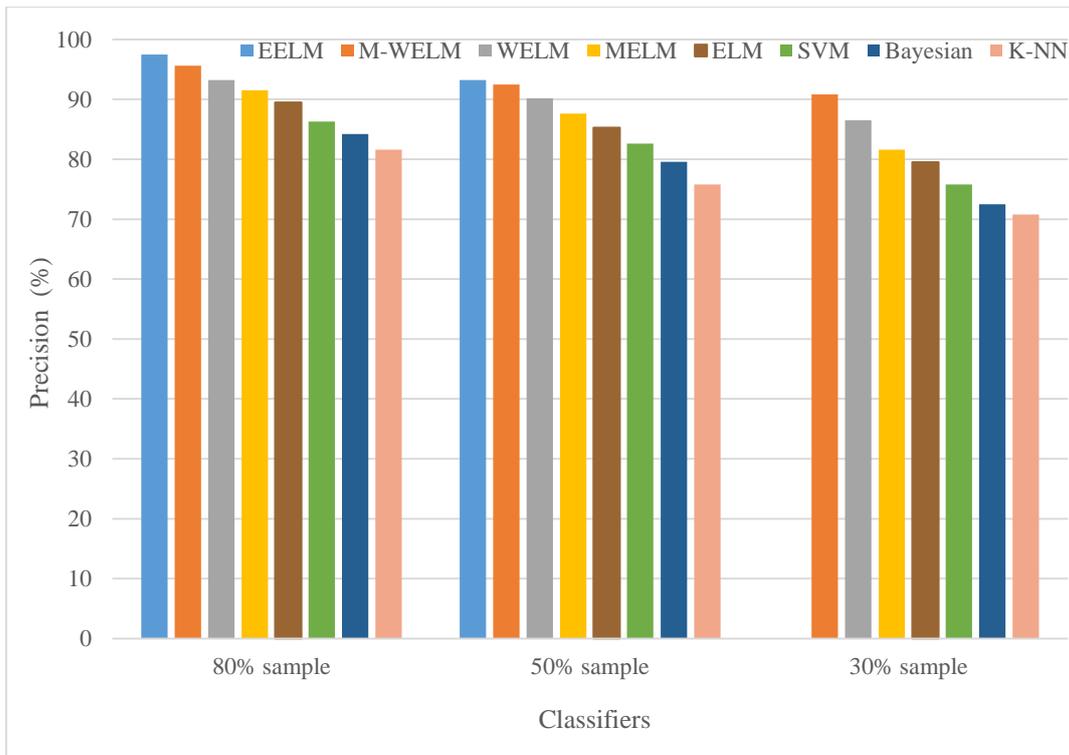| Classifiers	Accuracy (%)	Precision (%)	Recall (%)	Execution Time (Seconds)	Error Rate
EELM	**94**	**92.3**	**78.7**	**0.40**	**0.06**
M-WELM	92.3	90.8	79.8	0.42	0.08
W-ELM	90.6	86.5	82.5	0.45	0.09
M-ELM	88.6	81.6	84.6	0.48	0.11
ELM	85.2	79.5	86.5	0.51	0.15
SVM	80.3	75.8	88.5	0.55	0.2
Bayesian Network	78.6	72.5	89.3	0.59	0.21
KNN	71.3	70.8	91.6	0.62	0.29

Figure 3. Accuracy Comparison of Various Classifiers
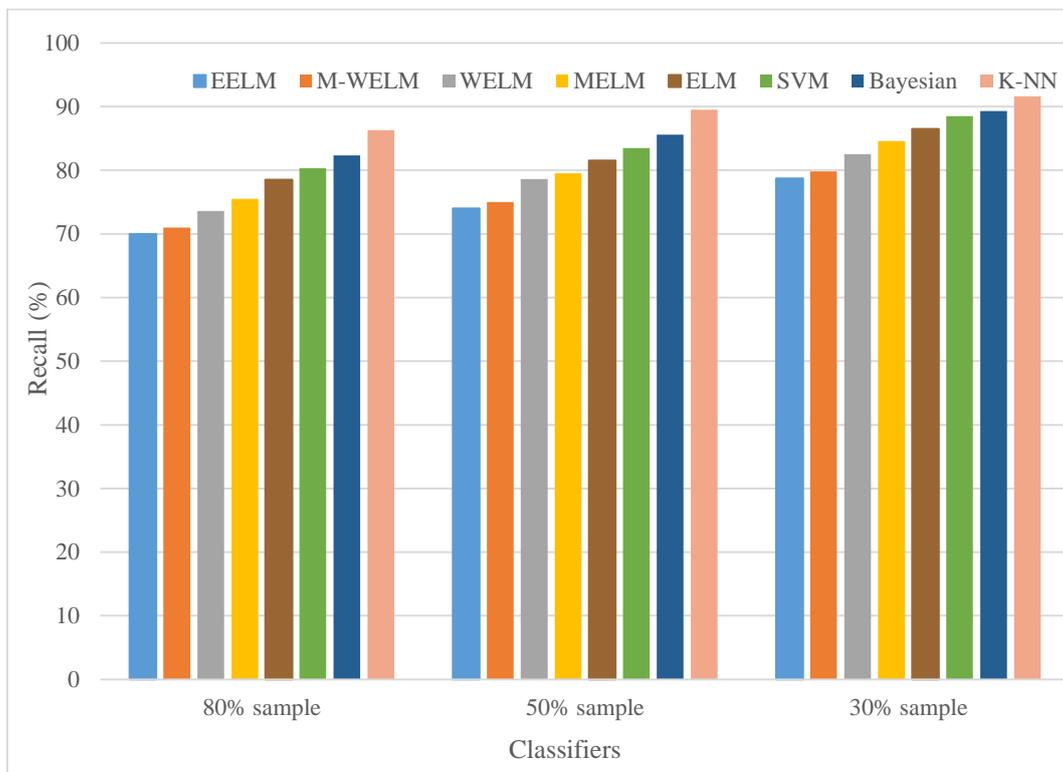
Figure 4. Precision Comparison of Various Classifiers

Figure 5. Recall Comparison of Various Classifiers
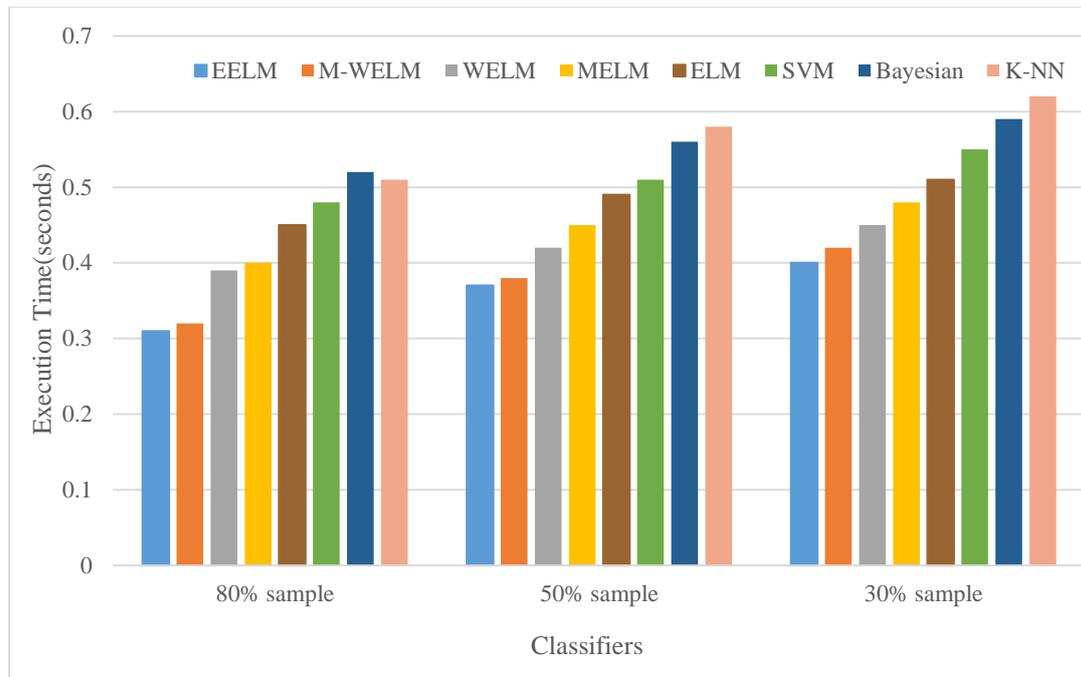
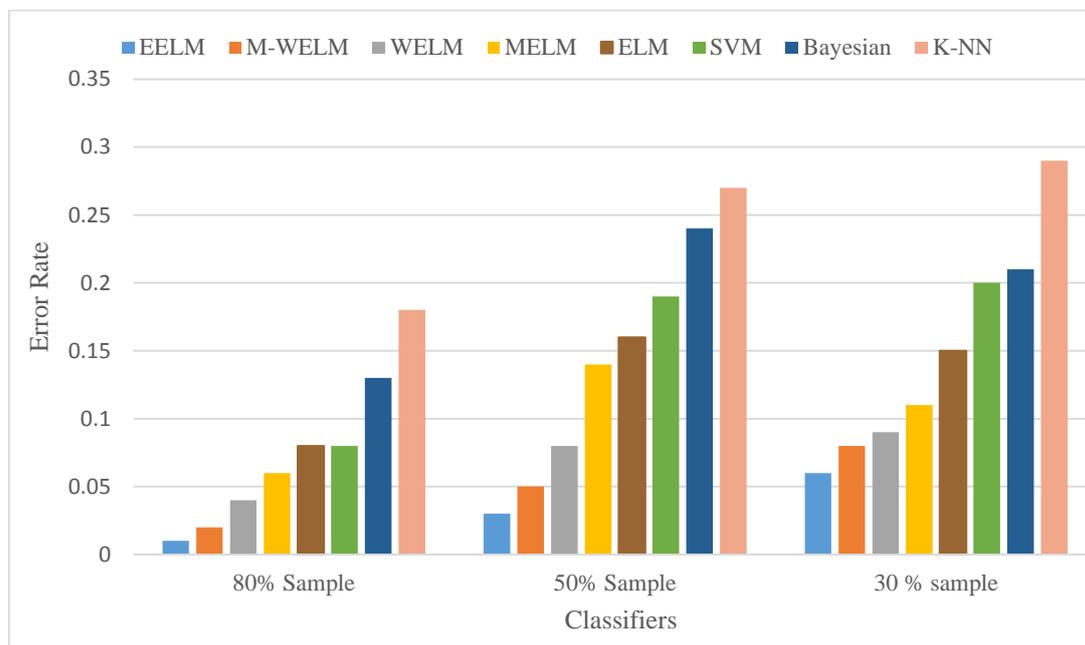**Figure 6.** Execution Time Comparison of Various Classifiers

Figure 7. Error Rate Comparison of Various Classifiers

Figure 3-7 shows the performance comparison of Proposed EELM with the traditional classifiers such as M-WELM, W-ELM, M-ELM, ELM, SVM, Bayesian Network and K-NN classifier in terms of accuracy, precision, recall, Execution time and Error Rate. From the graph it is clear that the proposed EELM shows better classification accuracy and robust for 80% sample for training and 20% testing when compared to other existing techniques. The highest training dataset 80% achieves good classification accuracy and it easily predict the presence of heart disease.

CONCLUSION

Heart Disease Prediction System is developed by employing Ensemble Extreme Learning Machine technique. This proposed method extracts hidden knowledge from a heart disease database. This is the most effective model to predict patients with heart disease of reasonable accuracy. The proposed method reduces the over fitting issues during learning process. From results it has been seen that Ensemble Extreme Learning Machine provides accurate results and excellent predictive performance as compare to other techniques.

Therefore, proposed EELM technique for diagnosis the cardiovascular disease rendering better prediction and robust strategy.

REFERENCES

[1] Olatunji, S.O., Selamat, A. and Raheem, A.A.A., 2010, May. Modeling permeability prediction using extreme learning machines. In Mathematical/Analytical Modelling and Computer Simulation (AMS), 2010 Fourth Asia International Conference on (pp. 29-33). IEEE.

[2] Huang, G.B., Zhu, Q.Y. and Siew, C.K., 2004, July. Extreme learning machine: a new learning scheme of feedforward neural networks. In Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on (Vol. 2, pp. 985-990). IEEE.

[3] R. Zhang, Y. Lan, G.-B. Huang, and Z.-B. Xu, "Universal approximation of extreme learning machine with adaptive growth of hidden nodes," IEEE Trans. Neural Netw. Learn. Syst.,, vol. 23, no. 2, pp. 365–371, 2012

[4] Parthiban, G. and Srivatsa, S.K., 2012. Applying machine learning methods in diagnosing heart disease for diabetic patients. International Journal of Applied Information Systems (IJAIS), 3, pp.2249-0868.

[5] Patel, J., TejalUpadhyay, D. and Patel, S., 2015. Heart Disease Prediction Using Machine learning and Data Mining Technique. Heart Disease, 7(1).

[6] Cho, S. B., and Ryu. J. Classifying gene expression data of cancer using classifier ensemble with mutually exclusive features. Proceedings of the IEEE 90, no. 11 (2002): 1744-1753.

[7] Pattekari, S.A. and Parveen, A., 2012. Prediction system for heart disease using Naïve Bayes. International Journal of Advanced Computer and Mathematical Sciences, 3(3), pp.290-294.

[8] Nahar, J., Imam, T., Tickle, K.S. and Chen, Y.P.P., 2013. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. Expert Systems with Applications, 40(1), pp.96-104.

[9] Kumari, V.A. and Chitra, R., 2013. Classification of diabetes disease using support vector machine. International Journal of Engineering Research and Applications, 3(2), pp.1797-1801.

[10] Anbarasi, M., Anupriya, E. and Iyengar, N.C.S.N., 2010. Enhanced prediction of heart disease with feature subset selection using genetic algorithm. International Journal of Engineering Science and Technology, 2(10), pp.5370-5376.

[11] Giri, D., Acharya, U.R., Martis, R.J., Sree, S.V., Lim, T.C., Ahamed, T. and Suri, J.S., 2013. Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. Knowledge-Based Systems, 37, pp.274-282.

[12] Cleveland database: http://archive.ics.uci.edu/ml/datasets/Heart+Disease

[13] Dangare, C.S. and Apte, S.S., 2012. Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), pp.44-48.

[14] Chen, H.L., Wang, G., Ma, C., Cai, Z.N., Liu, W.B. and Wang, S.J., 2016. An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson′s disease. Neurocomputing, 184, pp.131-144.

[15] Zhao, L.J., Diao, X.K., Yuan, D.C. and Tang, W., 2011. Enhanced classification based on probabilistic extreme learning machine in wastewater treatment process. Procedia Engineering, 15, pp.5563-5567.

[16] Shouman, M., Turner, T. and Stocker, R., 2012. Applying k-nearest neighbour in diagnosing heart disease patients. International Journal of Information and Education Technology, 2(3), p.220.

[17] Obenshain, M.K., 2004. Application of data mining techniques to healthcare data. Infection Control & Hospital Epidemiology, 25(8), pp.690-695.