Journal of Engineering and Applied Sciences 12 (23): 7358-7366, 2017
ISSN: 1816-949X
© Medwell Journals, 2017

Hybrid Decision Tree Fuzzy Rule Based Classifier for Heart Disease Prediction Using Chaotic Cuckoo Search Algorithm

Jagadeesh Gobal and Subhashini Narayan
School of Information Technology and Engineering,
Vellore Institute of Technology (VIT) University, Vellore, India

Abstract: Heart disease is the primary cause of death in all over the world and one of the primary diseases in developing countries. The disease identification and investigation needs a lot of prognosis and diagnosis with the high cost. To eliminate these burdens, we need to predict this disease in early stages with the help of machine learning techniques. The UCI repository heart data such as Cleveland, Hungarian dataset and Switzerland dataset for disease prediction. The proposed method consists of three stages: initially inconsistent, irrelevant and noisy data from the dataset using the fundamental fuzzy min. max. preprocessing, after that orthogonal locality preservation projection technique is applied to the processed data to reduce the dimension. At the second stage, chaotic cuckoo search algorithmetic approach is used for fitness evaluation, the combinations of cuckoo search algorithm, fuzzy and decision tree classifier can create a hybrid class. Information entropy algorithm will be sufficiently combined with cuckoo search algorithm achieves better classification accuracy. The result of classification accuracy of other existing algorithms is respectively 0.89, 0.93, 0.91. Consequently, the obtained results have shown very promising outcomes for the diagnosis of heart disease.

Key words: Preprocessing, cuckoo search, fuzzy, decision tree, classification, outcomes

INTRODUCTION

Heart disease is the main causes of morbidity and mortality in worldwide which leads to millions of deaths every year (Pu *et al*., 2012). The world health organization has predicted that 12 million deaths happen in worldwide every year because of the heart diseases. Most of the deaths in the United States and other urbanized countries happen because of heart diseases. A heart disease is a leading cause of death worldwide and the probable discovery at an earlier stage will avoid the seriousness. Most of the people experience chronic heart diseases. This disease is preferred to handle simple, ubiquitous and non-invasive sensors to find physiological information for early and primitive analysis (Huang *et al*., 2014). Contemporary clinical data attainment systems can incessantly check and storing measurements of vital patient signs such as Heart Rate (HR) and Blood Pressure (BP) and many days of hospitalization (Li-Wei *et al*., 2015). Physician's produces data with the prosperity of hidden information and it is not frequently used efficiently for forecasting (Masethe and Masethe, 2014). Atherosclerosis are a complete chronic disease characterized by the convention of prone-versus-one cative cells and lipids in the inside layer of the arteries (Rao and Kumar, 2013).

Literature review: Colman *et al*. (2011) have proposed an over disease-precise diagnostic choice technique which contains a one-versus-one attributes grade level and an attribute investigate level covered in the similar One-versus-one-rule Support Vector Machine (SVM) binary classifier. The projected technique varies from conventional methods in that it based on the choice of efficient characteristic subsets for distinctive a class from others by creating One-versus-one contrast. The Electrocardiograms (ECG) from the MIT-BIH arrhythmia database are used to estimate the projected special choice technique. The ECG characteristics accepted include inter-beat and intra-beat intervals, amplitude morphology, area morphology and morphological distance. Liu and Chen, 2015) have projected a comprehensive and experimental study of the new forecast of multi-label categorization. Total eight states of the art multi-label

Corresponding Author: Jagadeesh Gobal, School of Information Technology and Engineering,
Vellore Institute of Technology (VIT) University, Vellore, India

categorization techniques such as BR, CC, CLR, HOMER, Rakel, ECC, MLkNN and BRkNN are contrasted on 11 datasets. The investigational conclusion produces the major assistance of multi-label categorization.

Long *et al*. (2015) have projected a heart disease analysis system using jagged sets based quality fall and 23 intervals type 2 fuzzy logic system. It aims to hold with high-dimensional dataset confront and reservations. It exploits a 25 hybrid knowledge procedure including fuzzy c-mean clustering algorithm and limit alteration by chaos 26 firefly and inherited hybrid algorithms. This knowledge procedure is computationally costly, particularly 27 when in use with the high-dimensional dataset. The rough sets based quality fall using chaos 28 firefly algorithm is examined to discover optimal decline which so, decreases computational load 29 and improves the presentation of the system. The projected representation is helpful as a decision support system for 32 heart disease analysis. Weng *et al*. (2016) have projected, neural network ensembles models for medical analysis. The main reason of this study is to examine the presentation of different classifiers with entity classifiers concerned in a collection classifier and solo classifiers. Additionally, a variety of estimate criteria is used to inspect the presentation of these classifiers with real-life datasets. Saini *et al*. (2013) have proposed an algorithm which explained typical databases such as CSE and MIT-BIH Arrhythmia database. The digital band-pass filter was used to decrease artificial recognition originated by obstruction exists in ECG signal with QRS-recognition. The recognition rates of 99.89 and 99.81% were accomplished for CSE and MIT-BIH databases correspondingly.

Vafaie *et al*. (2014) have proposed an Atrial arrhythmias modification algorithm to differentiate between critical positions within the great atrial bundles resultant from unusually quick crucial movement in the atria may be imitated in a distorted P-Wave Morphology (PWM) in the ECG. The algorithm was realistic and could forecast the important atrial position with 85% correctness. Anooj (2012) have proposed a weighted fuzzy rule-based clinical decision support system be presented for the diagnosis of heart disease which uses production of subjective fuzzy rules. Fuzzy rule-based choice support system. The healthcare industry gathers enormous of heart disease data which is unfortunately to discover hidden information for effective decision. The significant progress made in the diagnosis and treatment of heart disease, further investigation is still needed (Vafaie *et al*., 2014).

MATERIALS AND METHODS

The objective of this study is to develop a prediction system for the classification for medical analysis of heart disease. Numerous methods were implemented to healthcare data sets for the prediction of future health care consumption such as expecting entity payments and disease risk for patients. From these works, it can be observed that feature selection methods can effectively increase the performance of single classifier algorithms in diagnosing heart disease (Pu *et al*., 2012). Noisy features and dependency relationships in the heart disease dataset can influence the diagnosis process. Typically, it is necessary to reduce the dimensions of the original feature set by a feature selection method that can remove the irrelevant and redundant features.

In machine learning environment, the dimension reduction is the process of reducing the number of random variables under conditional situations. The feature dimension reduction technique facilitate efficient data processing task and decrease the size of the data. It has also been empirically shown that, removing redundant features can result in significant performance improvement. There are several advantages to combining for OLPP approach as an integrated feature selection system for heart disease diagnosis. It can remove superfluous and redundant features more effectively. The OLPP algorithm can select relevant features for disease diagnosis, however, redundantfeaturesmay still exist in the selected relevant features. In such cases, the OLPP reduction algorithmcan remove remaining redundant features to offset this limitation of the ReliefF algorithm. The OLPP method helps to accelerate the reduction process and guide the search of the reducts and improve the quality of reducts. When unnecessary features are removed, moreimportant features can be extracted which will also improve the quality of reducts. It is obvious that the choice of an efficient feature selection method and an excellent classifier is extremely important for the heart disease diagnosis problem. Most of the common classifiers from the machine learning community have been used for heart disease diagnosis. It is now recognized hybrid iclassifier combination model that results in more precise decisions which reduces the variance of error estimation beyond the level that can be achieved by individual classifiers. In this study, we propose a hybrid classification system based on the chaotic based decision tree classifier approach for handling heart disease data set.

The system contains two subsystems: the OLPP feature selection subsystem and a chaotic decision tree classification system. First, the proposed method adopts

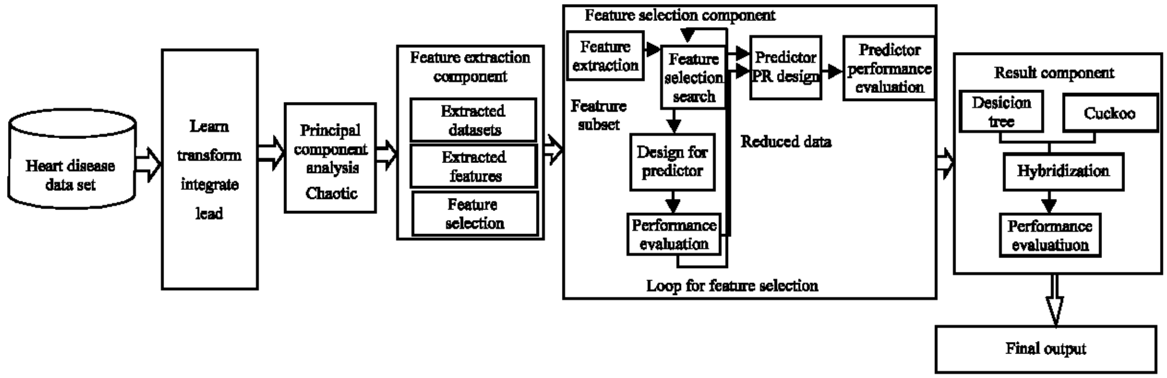

Fig. 1: Proposed heart disease prediction system

the mathematical min. max. model to obtain feature weights and select more relevant and important features with the help of orthogonal locality preservation projection s from heart disease datasets. Then, the feature estimation obtained from the first phase is used as the input for chaotic cuckoo search selection with classification decision tree subsystem. Finally, the performance can be obtained with help of chaotic cuckoo serch decision tree. To evaluate the performance of the proposed hybrid classification method, a confusion matrix, sensitivity, specificity and accuracy were used. The experimental results show that, the proposed method achieves very promising results using the k-fold validation method.

The main contributions of this study are summarized as follows. We propose a feature selection system eith the help of mathematicl min. max. model with orthogonal locality preservation projections to detect heart disease in an efficient and effective way. In the classification system, we propose an chaotic cuckoo search based ID3 as the base classifier. Hybrid learning can achieve better performance at the cost of computation than single classifiers. The experimental results show that, the ensemble classifier in this study is superior to three common classifiers. Compared with three classifiers and previous studies, the proposed diagnostic system achieved excellent classification results (Fig. 1).

Orthogonal locality preservation projection: This OLPP algorithm differs from principal component analysis and linear discriminant analysis as its manifold locality preservation. OLPP constructed an orthogonal basis function which produces the adjacency matrix for the class association between the data which is the main distinct from locality preserving projections. The practical consequences have confirmed that OLPP has got more discriminating and defensive authority while estimated with the locality preserving projections. The orthogonal expansion of LPP is known as the Orthogonal Locality Preserving Projections (OLPP).

Preprocessing: In the existent world data are normally imperfect, noisy and unpredictable. If imperfect means missing characteristic values, missing definite characteristic of attention or having only collective of data. Noisy means enclosing faults or outlines. Unpredictable means having inconsistency in codes or names. Mostly, we achieve a data cleaner, data integration, data transformation, data reduction, data discretization but in this study, we will reduce a dimension of the missing attributes.

Feature reduction using OLPP: The high amount of feature is a huge problem for classification. So, the feature dimension technique is implemented to decrease the fetters break without trailing the accuracy of classification. The OLPP algorithm varies from principal component analysis and linear discriminant analysis. Both the algorithms intend at aspect dimensionality reduction since the initial level of the algorithm is principal component analysis which assists in aspect fall. An adjacency chart is constructed by OLPP and the class association between the example sample points is best replicated by it. It is not simple to rebuild the data as Locality Preserving Projections (LPP) is non-orthogonal usually. By means of orthogonal locality preserving projection technique, this difficulty is a defeat which generates orthogonal basis function and can have extra locality preserving power than locality preserving projections. The sensible consequences have confirmed that OLPP has got more discriminating and defensive authority while estimated with the locality preserving projections. The orthogonal expansion of LPP is known as the Orthogonal Locality Preserving Projections (OLPP) (Fig. 2).

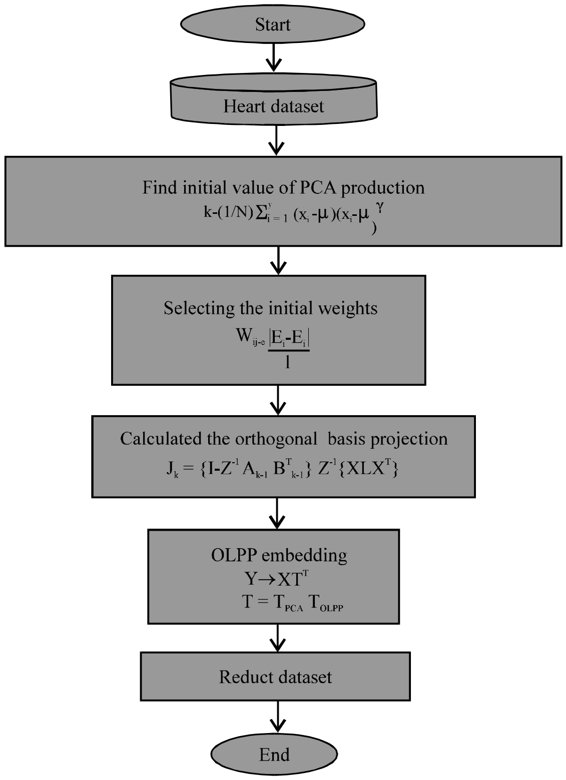

Fig. 2: OLPP algorithm for enhancement

RESULTS AND DISCUSSION

The procedures involved in OLPP: Principal components analysis is a technique that decreases data dimensionality by processing a covariance analysis between aspects. The PCA projection contains the subsequent steps: obtain a set of features from apreprocessed data. Let x be a matrix having the size of m×n. Compute the mean value:

$$\mu = \left(\frac{1}{N}\right)\sum_{i=1}^{N} x_i \tag{1}$$

Compute the covariance matrix:

$$K = \left(\frac{1}{N}\right)\sum_{i=1}^{N}(x_i\text{-}\mu)(x_i\text{-}\mu)^T \tag{2}$$

Compute the eigenvector e and eigenvalues λ of the covariance matrix. If, K is a square matrix, a non-zero vector e is an eigenvector of K if there is a scalar λ (eigen value) such that:

$$Ke = \lambda e \tag{3}$$

The eigen value and eigenvectors are prearranged and combined. The nth eigenvalue communicates to the n^{th} eigenvector. The conversion matrix of PCA is symbolized by T_{PCA}. By PCA projection, the extracted features are statistically uncorrelated and the grade of the novel data matrix is equivalent to the amount of attributes (dimensions).

Constructing the adjacency graph: Consider X = [E_1, E_2, ..., E_K] is a set of dataset. Consider G represents a graph with n nodes. The ith node communicates to the e-Mail E_i. A boundary is placed between nodes i and if E_i and E_j are "lock", i.e., E_i is between p adjacent neighbours of E_j or E_j is between p adjacent neighbours off E_i. If the group information is obtainable in any two nodes, we just place a border between that two nodes fitting into the similar group.

Choosing the weights: If the node i and j are connected, the weight W_{ij} is calculated using Eq. 1:

$$W_{ij} = e^{-\frac{\left\|E_i\text{-}E_j\right\|}{t}} \tag{4}$$

where, t (constant). If the node i and j are not connected means we put W_{ij} = 0. The weight matrix W of graph G models having the local structure of various data.

Computing the orthogonal basis functions: After discovering the weight matrix W, we compute the diagonal matrix M. A diagonal matrix M is denoted as whose entrance are column (or row) sums of W:

$$M_{ii} = \sum_j W_{ji} \tag{5}$$

After that, we calculate the Laplacian matrix L using diagonal matrix M and weight matrix W:

$$L = M\text{-}W \tag{6}$$

Let, O_1, O_2, ..., O_k be orthogonal basis vectors, we define:

$$A_{K-1} = [o_1, o_2, ..., o_{k-1}],\ B^T_{K-1} = A^T_{K-1} Z^{-1} A_{K-1} \tag{7}$$

where, Z^{-1} = XMX^T. The orthogonal basis vectors [O_1, O_2, ..., O_k] can be computed as follows: Compute O_1 as the eigenvector of $Z^{-1}XLX^T$ associated with the smallest eigenvalue. Compute O_2 as the eigenvector of associated with the smallest eigenvalue of J_k:

$$J_K = \{I\text{-}Z^{-1} A_{K-1} B_{K-1}^T\} Z^{-1} \{XLX^T\} \quad (8)$$

OLPP embedding: Let, T_{OLPP} = [O_1, O_2, O_3, ..., O_l] embedding is followed:

$$Y \rightarrow XT^T \quad (9)$$

$$T = T_{PCA}\, T_{OLPP} \quad (10)$$

Where:
T = The (transformation matrix
Y→ = One dimensional representation of X

This transformation matrix decreases the element of the attribute vectors of the heart disease prediction data. This element decreased characteristics, given to the categorization method.

Classification

Fuzzy: Fuzzy rule based classification is a technique of producing a diagram from a specified input to an output using fuzzy logic. The progression of fuzzification is calculated by implementing the subsequent Eq. 11 and 12:

$$ML = \min + \left(\frac{\max - \min}{3}\right) \quad (11)$$

$$XL = ML + \left(\frac{\max - \min}{3}\right) \quad (12)$$

Where:
ML = The minimum limit values of the featur M
XL = The maximum limit values of the feature M

By use Eq. 11 and 12 to compute the minimum and maximum limit values. And also and these regulations are supplied to produce the fuzzy values. These rules are given in to the cuckoo search.

Cuckoo search: Cuckoo search algorithm is a meta heuristic algorithm which was presented by the reproductive activities of the cuckoos and reduces to carry out. They contain many nests in cuckoo search. Every egg point to a resolution and an egg of cuckoo specify a fresh explanation cuckoo search is inspired by smart incubation behavior of a type of birds called cuckoos in nature for the single-object case, incuckoo search, the number of eggs, cuckoos and nests are equal and each one represents an available solution. cuckoo search is well capable of finding the best solutions by continuously using potentially better solution to replace a not-so-good cuckoo in the population. Conceptually, extremely simple, cuckoo search is very easy to implement the new and improved explanation is replacing the majority useful explanation in the nest. The following representative system is selected by the cuckoo search algorithm every egg in a nest symbolizes an explanation and a cuckoo egg symbolizes a novel explanation. The intent is to utilize the novel and probably improved egg to restore a not-so-good egg of cuckoo in the nests. Though, this is the fundamental situation which means one cuckoo for each nest but the amount of the method can be amplified by integrating the possessions that every nest can have more than one egg which corresponds to a set of explanations. The procedure of clustering is indicated:

- The only one egg at a time is laid by cuckoo. Cuckoo dumps its egg in a randomly chosen nest
- The number of available host nests is fixed and nests with high quality of eggs will carry over to the next generations
- In case of a host bird discovered the cuckoo egg it can throw the egg away or abandon the nest and build a completely new nest (Algorithm 1)

Algorithm 1; Cuckoo search via. Levy flights:

```
Begin Step 1:
        Initialization. Set the generation counter G = 1
        initialize the population P of n host nests randomly
        set the discovery rate pa
 Step 2:
       While G<MaxGeneration do Sort the population as per their fitness
       Get a cuckoo randomly (say, i) and replace its solution by performing
       Levy flights
       Evaluate it's fitness F_i. Choose a nest among n (say, j) randomly
        if (F_i<F_j) replace j by the new solution
        end if A fraction (pa) of the worse nests is abandoned and new ones are
       built. Keep the best solutions/nests
       Sort the population and find the current best
       Pass the current best to the next generation. G = G+1
 Step 3: end while
 End
```

Initialization phase: The population (m_i where i = 1, 2, ..., n) of host nest is commenced randomly.

Generating new cuckoo phase: The Levy flights is used a cuckoo to selected at random and it generates new explanations. After that, the created cuckoo is evaluated by the aim utility for finding the worth of the explanations.

Fitness evaluation phase: Assess the fitness function based on the equation and after that choose the best one:

$$\text{Fitness} = \text{Maximum popularity} \quad (13)$$

Updation phase: Modify the primary explanation by Levy flights in which cosine transform is engaged. The main.

parameters of CS are the step size α and discovery rate pa that characterizes the variations of the global best step size and their values have a great influence on the convergent speed and how CS performs. The basic CS is well of finding the best solutions but the solutions are still swinging slightly around the optima. The step size used in CS remains unchanged. The improved CS method with a chaotic varying step size α may be better than the basic one which may also accelerate its convergence. By normalizing all chaotic maps, their variations are always in [0, 2]. After normalization, chaotic maps are able to tune step size α and this improved CS method is referred as the chaotic CS elitism strategy is introduced into CCS to protect better strategy.

Zakharov chaotic map:

$$\int f(\vec{x}) = \sum_{i=1}^{N}(x_i)^2 + \sum_{i=1}^{N}(0.5ix_i)^2 + \sum_{i=1}^{N}(0.5ix_i)^4 \quad (14)$$

Reject worst nest phase: The bad nests are removed in this level, depend on their option values and new ones are constructed. Afterward, depend on their suitability task the finest explanations are marked. After that, the finest explanations are predictable and marked as optimal solutions.

Stopping criterion phase: Until, the maximum iteration accomplishes this process is replicated. The optimized consequence will be inspected for to determine of software excellence. The particular procedure is apparently established in flow chart. It's exposed in below.

Decision tree: Decision tree form is quick reliable, effortless to preserve and correct in the preparation course area ID3 Algorithm is choosing which quality to examination at every node in the tree. We would like to choose the quality that is mainly helpful for classifies examples what is a high-quality quantitative quantify of the bad of a quality, we will describe an arithmetical property called information increase that calculates how well a known quality divides the preparation instances based on their target classification. ID3 utilize this information increase calculates to choose between the candidate at every phase while increasing the tree.

In decision tree knowledge, ID3 (Iterative Dichotomiser 3) is an algorithm proposed by Ross Quinlan employed to produce a decision tree from the dataset. ID3 is classically used in the machine learning and natural language dispensation fields. The decision tree method contains build a tree to form the categorization procedure. Once a tree is constructing, it is functional to every tuple in the database and consequences in categorization for that tuple. The subsequent problems are resolved by most decision tree algorithms:

- Choosing splitting attributes
- Ordering of splitting attributes
- Number of splits to take
- Balance of tree structure and pruning
- Stopping criteria

The ID3 Algorithm is a categorization algorithm depend on information entropy, its fundamental design is that all instances are drawn to dissimilar classed depending on dissimilar values of the state quality set; its center is to establish the finest categorization quality form state quality sets. The algorithm decides information increase as quality collection criteria as a rule the quality that has the uppermost information increase is chosen as the dividing quality of existing node in order to create information entropy that, the separated subsets require minimum. According to the dissimilar values of the quality, branches can be recognized and the procedure is recursively described on every branch to generate other nodes and branches until all the examples in a branch fit in to the similar group. To choose the dividing qualities, the idea of entropy and information increase are utilized.

Entropy: Given probabilities $p_1, p_2, ..., p_s$ where, $\Sigma\ p_i = 1$, entropy is defined as:

$$H(p_1, p_2, ..., p_s) = \sum -(p_i \log p_i) \quad (15)$$

Entropy discovers the quantity of array in a known database state. A value of H = 0 recognize a completely categorized set. The sup error the entropy which means the better in the possible to develop the categorization method.

Information gain: ID3 decide the divide quality with the maximum increase in information where increase is defined as dissimilarity between how much information is desirable after the divide. This is considered by formative the dissimilarity between the entropies of the unique dataset and the subjective amount of the entropies from each of the subdivided datasets. The formula used for this reason is:

$$G(D, S) = H(D) - \sum P(D_i)H(D_i) \quad (16)$$

Dataset description: The proposed fuzzy based classifier is experimenting with the three datasets namely Cleveland, Hungarian, Switzerland. These datasets are taken from the UCI machine learning repository.

Cleveland data: This data set was compiled by Robert Detrano and PhD holder at V.A Medical Center, Long Beach and Cleveland Clinic Foundation the main task is to detect the presence of heart disease in the patients. This database encloses 76 characteristics but the researchers handle only 14 out of them. All the papers are related to the use of subset indicate the existence 14 of the 76 features that are present in the processed in the Cleveland data set. The field includes an integer constant that can take any value from 0-4. It is the integer value of 0 no presence of illness and 1-4 are disease existence. The "final" field indicates the presence of coronary artery disease in the patients. Six of the examples have missing values; class distributions are 54% heart disease absent and 46% heart disease present.

Hungarian data: Owing to a vast percentage of missing values three of the characteristics have been rejected, however, the format of the data is precisely the similar as that of the Cleveland data. About 34 examples of the database were rejected on account of missing values and 261 examples were present. Class distributions are 62.5% heart disease not present and 37.5% heart disease present.

Switzerland data: More number of missing values is in Switzerland data. It encloses 123 data instances and 14 features. Class distributions are 6.5% heart disease not present and 93.5% heart disease present.

Evaluation metrics: An evaluation metric is used to evaluate the effectiveness of the proposed system. It consists of a set of measures that follow a common underlying evaluation methodology some of the metrics that, we have choose for our evaluation purpose are specificity, sensitivity, accuracy.

Sensitivity: The measure of the sensitivity is the proportion of actual positives which are accurately recognized. It relates to the capacity of test to recognize positive results:

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

Where:
TP = True Positive
FN = False Negative

Table 1: Truth table for experimental outcome

	Cleveland		Hungarian		Switzerland	
-	0	1	0	1	0	1
0	162	2	180	8	5	3
1	136	3	94	12	115	0

Table 2: Confusion matrix for all datasets

	Condition as determined by the standard truth	
Experimental outcome	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Table 3: Comparison of accuracy values in proposed vs. existing methods of Switzerland Cleveland, Hungarian datasets

Accuracy	Fuzzy with CS and DT	Fuzzy with CS	Fuzzy	SVM
Switzerland	0.890	0.884	0.844	0.835
Cleveland	0.933	0.874	0.844	0.785
Hungarian	0.911	0.846	0.834	0.814

Specificity: The measure of the specificity is the extent of negatives which are properly recognized. It relates to the capacity of test to recognize negative results:

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

Where:
TN = True Negative
FP = False Positive

Accuracy: Accuracy of the proposed method is the ratio of the total number of TP and TN to the total number of data (Table 1):

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

Confusion matrix: The confusion matrix for the heart disease prediction system obtained by combining the evidence of overall recognition performance is shown in Table 2 for all datasets. a confusion matrix for all datasets such as Cleveland, Hungarian and Switzerland datasets are present in Table 2.

Comparative analysis: The proposed work helps to attain very good accuracy for the heart disease database using improved cuckoo search classifier. The fuzzy with cuckoo search, fuzzy and SVM algorithm are utilized for the comparison of work. The comparison outcomes are presented in Table 3 (Fig. 3-6).

The improved good accuracy outcomes of heart disease classification are presented by our proposed work. In comparison with the existing fuzzy algorithm gives very less accuracy values for the evaluation measures. The performance of the proposed knowledge

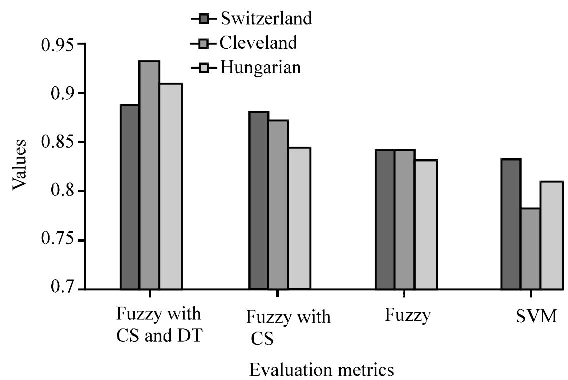

Fig. 3: Graph for compared accuracy result in evaluation metrics; Accuracy prediction of heart disease data

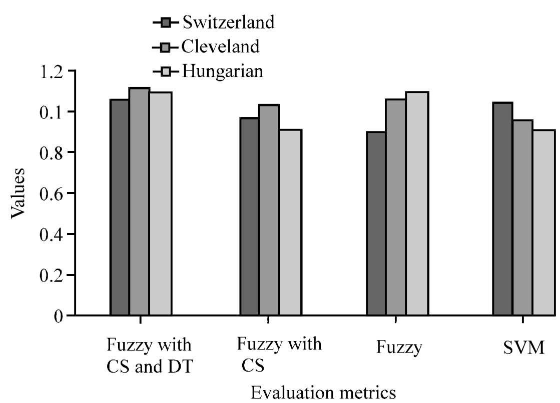

Fig. 4: Graph for compared sensitivity result in evaluation metrics; Sensitivity prediction of heart disease data

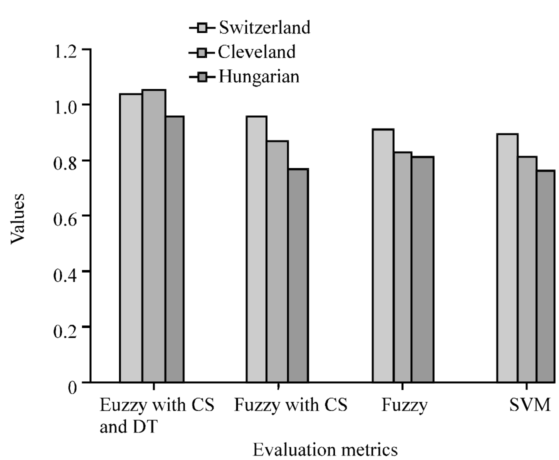

Fig. 5: Graph for compared specificity result in evaluation metrics; Specificity prediction of heart disease dataset

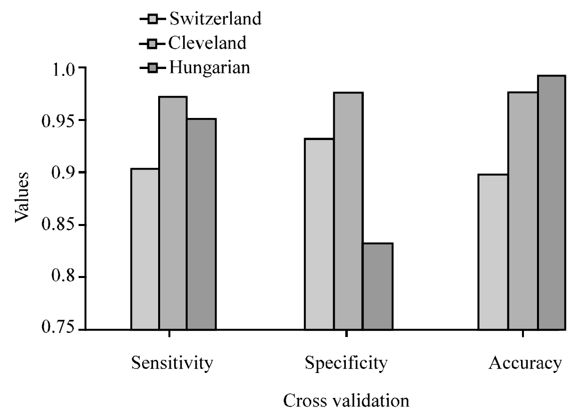

Fig. 6: Graph for cross validation

Table 4: Comparison of sensitivity values in proposed vs. existing method of Switzerland Cleveland, Hungarian datasets

Sensitivity	Fuzzy with CS and DT	Fuzzy with CS	Fuzzy	SVM
Switzerland	0.905	0.831	0.825	0.90
Cleveland	0.977	0.934	0.969	0.85
Hungarian	0.917	0.804	0.970	0.84

Table 5: Comparison of specificity values in proposed vs. existing method of Switzerland Cleveland, Hungarian datasets

Specificity	Fuzzy with CS and DT	Fuzzy with CS	Fuzzy	SVM
Switzerland	0.961	0.886	0.85	0.83
Cleveland	0.978	0.805	0.77	0.75
Hungarian	0.886	0.716	0.76	0.71

Table 6: Results of the proposed improved cuckoo search classification

Cross validation	Sensitivity	Specificity	Accuracy
Switzerland	0.891	0.9170	0.886
Cleveland	0.953	0.9578	0.958
Hungarian	0.935	0.8250	0.973

extraction in heart disease prediction methods evaluated by the three metrics sensitivity, specificity and accuracy. The results of proposed work help to analyze the efficiency of the prediction process. The subsequent Table 4-6 tabulates the results.

CONCLUSION

A hybrid decision tree with cuckoo search based classification method was proposed in this study. The approach starts with preprocessing technique which remove noisy, irrelevant, inconsistent data from the database. Then, feature dimension reduction method will be applied to reduce the feature's space using Orthogonal Locality Preserving Projection (OLPP) algorithm. In next classification phase fuzzy decision tree classifier effectively combined with Chaotic Cuckoo Search (CCS) algorithm which produce a fusioin classifier and leads for accurate classification. UCI repository will be utilized and the performance of the proposed system gives higher prediction rates compare with other existing techniques. Performance can be measure by means of sensitivity, specificity and accuracy. The efficiency of the classification is very high by presenting very good accuracy outcomes and also the classification of heart disease prediction is gives very accurate outcomes. From the outcomes, we have showed that the improved cuckoo search classifier utilized in our proposed work outperforms the other existing classifiers by facilitated very good accuracy. Thus, we can observe that our proposed work is better than other existing works for the heart disease prediction.

REFERENCES

Anooj, P.K., 2012. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. J. King Saud Univ. Comput. Inf. Sci., 24: 27-40.

Colman, M.A., O.V. Aslanidi, J. Stott, A.V. Holden and H. Zhang, 2011. Correlation between P-wave morphology and origin of atrial focal tachycardia-insights from realistic models of the human atria and torso. IEEE. Trans. Biomed. Eng., 58: 2952-2955.

Huang, S.C., P.H. Hung, C.H. Hong and H.M. Wang, 2014. A new image blood pressure sensor based on PPG, RRT, BPTT and harmonic balancing. IEEE. Sens. J., 14: 3685-3692.

Li-Wei, H.L., R.P. Adams, L. Mayaud, G.B. Moody and A. Malhotra *et al*., 2015. A physiological time series dynamics-based approach to patient monitoring and outcome prediction. IEEE. J. Biomed. Health Inf., 19: 1068-1076.

Liu, S.M. and J.H. Chen, 2015. An empirical study of empty prediction of multi-label classification. Expert Syst. Appl., 42: 5567-5579.

Long, N.C., P. Meesad and H. Unger, 2015. A highly accurate firefly based algorithm for heart disease prediction. Expert Syst. Appl., 42: 8221-8231.

Masethe, H.D. and M.A. Masethe, 2014. Prediction of heart disease using classification algorithms. Proceedings of the World Congress on Engineering and Computer Science, Volume 2, October 22-24, 2014, San Francisco, USA., pp: 1-4.

Pu, L.N., Z. Zhao and Y.T. Zhang, 2012. Investigation on cardiovascular risk prediction using genetic information. IEEE. Trans. Inf. Technol. Biomed., 16: 795-808.

Rao, V.S.H. and M.N. Kumar, 2013. Novel approaches for predicting risk factors of atherosclerosis. IEEE. J. Biomed. Health Inf., 17: 183-189.

Saini, I., D. Singh and A. Khosla, 2013. QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. J. Adv. Res., 4: 331-344.

Vafaie, M.H., M. Ataei and H.R. Koofigar, 2014. Heart diseases prediction based on ECG signals classification using a genetic-fuzzy system and dynamical model of ECG signals. Biomed. Signal Process. Control, 14: 291-296.

Weng, C.H., T.C.K. Huang and R.P. Han, 2016. Disease prediction with different types of neural network classifiers. Telematics Inf., 33: 277-292.