Asian Journal of Information Technology 15 (3): 481-492, 2016
ISSN: 1682-3915

Heart Disease Prediction System Using Optimal Rough-Fuzzy Classifier Based on ABC

[1]T. Keerthika and [2]K. Premalatha
[1]Department of Information Technology,
Sir Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India
[2]Department of Computer Science and Engineering, Bannari Amman Institute of Technology,
Sathyamangalam, Tamil Nadu, India

Abstract: Worldwide heart disease forecast has been a major research over the past decade since the major reason of death is due to heart disease. Numerous researchers combined fuzzy technique with some other technique for proficient classification purpose in order to predict the heart disease, since the fuzzy is proficient only if proper fuzzy rules are specified in the rule base. At this point, we have introduced a rough-fuzzy classifier that shared rough set theory with the fuzzy set. Generally, there are three main steps taken part in the rough-fuzzy classifier such as: rule generation using rough set theory, rule optimization using Artificial Bee Colony (ABC) and prediction using fuzzy classifier. At first, the discernability matrix is framing by the given database. Reduct and core analysis is used to recognize the relevant attributes from the discernability matrix after that fuzzy rules are generated from the rough set theory. After that the set of rule is optimized. Then, with the assist of fuzzy rules and membership functions, the fuzzy system is intended so that the prediction can be carried out within the fuzzy system intended. Finally, the experimentation is carried out by means of the Cleveland, Hungarian and Switzerland datasets. From the results, we ensure that the proposed rough-fuzzy classifier outperformed the previous approach by achieving the accuracy of 87% in Hungarian and 80% in Switzerland datasets.

Key words: Heart disease prediction, Rough Set theory, artificial bee colony, fuzzy set, membership function, classifier

INTRODUCTION

Heart diseases are the number one cause of death globally: more people die annually from heart diseases than from any other cause. An estimated 17.3 million people died from heart diseases in 2008, representing 30% of all global deaths. Of these deaths, an estimated 7.3 million were due to coronary heart disease and 6.2 million were due to stroke (WHO, 2011). Recent research in the field of medicine has been able to identify risk factors that may contribute toward the development of heart disease but more research is needed to use this knowledge in reducing the occurrence of heart diseases. Diabetes, hypertension and high blood cholesterol have been established as the major risk factors of heart diseases. Life style risk factors which include eating habits, physical inactivity, smoking, alcohol intake, obesity are also associated with the major heart disease risk factors and heart disease (Mozaffarian *et al*., 2008; Poirier, 2008). There are studies showing that reducing these risk factors for heart disease can actually help in preventing heart diseases (Wood *et al*., 1998; Anderson *et al*., 1991). There are many studies and researches on the prevention of heart disease risk. Data from studies of population has helped in prediction of heart diseases based on blood pressure, smoking habit, cholesterol and blood pressure levels, Diabetes. Researchers have used these prediction algorithms in adapted form of simplified score sheets that allow patients to calculate the risk of heart diseases.

Diagnosis is a methodology for identification by process of elimination, of the nature of anything. Medicine, science, engineering, business, etc. are some of the areas that employ diagnosis. In medicine, diagnosis is "the recognition of a disease or stipulation by its apparent signs and symptoms" or "the analysis of the underlying physiological Biochemical cause's of a disease or condition" (Wolff, 2006). An important issue in medical diagnosis is the risk stratification which refers to the sorting of patients based on the severity of disease. This is vital owing to the reason that it can help in reducing the usage of beds, equipment and other medical resources.

Corresponding Author: T. Keerthika, Department of Information Technology, Sir Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India

Almost all the physicians are confronted during their formation by the task of learning to diagnose. Here, the problem of deducing certain diseases or formulating a treatment has to be solved by them on the basis of more or less specified observations and knowledge (Brause, 2001). In order to keep more of the relevant information constantly in mind the physicians are encouraged by continued training and recertification procedures. However, it is assured that most of what is known cannot be known by most individuals due to the fundamental limitations of human memory and recall coupled with the growth of knowledge. A good physician employs his knowledge experience and talent during the medical diagnosis procedure to diagnose a disease (Wolff, 2006). The diagnosis is then determined by taking the total available patients' status into account. The appropriate treatment is prescribed depending on the diagnosis and the entire process might be iterated. The diagnosis might be reconfigured, refined, or even rejected (Steimann and Adlassnig, 1998) in every iteration.

Literature review: For clinical decision support systems, artificial intelligence and data mining techniques presents a number of researches and literature. A handful of researches have been presented to support choice makers in the risk prediction of heart disease. A few of the considerable researches available in the literature are explained below.

Palaniappan and Awang (2008) have proposed a prototype heart disease prediction system has created using three data mining classification modelling approaches. The system extracted hidden information from a historical heart disease database. DMX query language and functions used to create and access the models. The models trained and validated against a test dataset. Lift Chart and Classification Matrix approaches used to identify the effectiveness of the models. All three models were able to extract patterns in response to the predictable state. The most efficient model to predict patients with heart disease appears to be Naive Bayes followed by neural network and decision tree.

Chen *et al.* (2001) have explained a heart disease prediction system that have assist medical professionals in analyzing a patient's heart disease mainly based on the medical data of the patient. There are three main steps. At first, 13 important medical features have been selected, i.e., age, sex, chest pain type, trestbps, cholesterol, fasting blood sugar, resting ecg, max heart rate, exercise induced angina, old peak, slope, number of vessels colored and thal. Secondly, for classifying heart disease, they have developed an artificial neural network algorithm based on these medical features. Nearly, 80% was obtained as the prediction accuracy.

Hannan *et al.* (2010) presented the applicability of RBF has investigated for prediction of medical prescription of heart disease. Totally, 300 patients data records were collected from Sahara Hospital, Aurangabad. The information contained of all the parameters required to decide whether the patient was suffering from heart disease or not. The experimental data converted into binary form. The collected information coded, normalized and entered into 13 different excels sub-sheets. Patient's data trained by using radial basis Function and the outcome was compared with the original medicines provided by the doctor. Around 75 samples were tested with this outcome of the Radial Basis Function. It was found that the result of the testing data by using RBF was satisfactory.

Tavares *et al.* (2013) have proposed a Non-invasive data was used to develop a model able to predict if a child is or not cardiac prone. Four machine learning methods and three techniques to solve the imbalance problem were investigated. For the chosen dataset and two UCI datasets, preliminary experiments have shown that the method for dealing with imbalancing datasets recently introduced that did not perform well. So, the given method was discarded. The method proposed in the given study (Weighted SVM+MLP) outperformed the other methods in 3 out of 4 performance metrics. In this context, the damage of classifying a healthy person as cardiac was much smaller than classify a cardiac person as healthy. It caused a delay on the treatment, thereby increasing the severity of the disease.

Amin *et al.* (2013) presented a data mining techniques and approaches applied in patient medical dataset has resulted in innovations, standards and decision support system that have significant success in improving the health of patients and the overall quality of medical services. But, they have needed systems which could predict heart diseases in early stages. In the given study, a new hybrid model of neural networks and genetic algorithm to optimize the connection weights of ANN so as to improve the performance of the artificial neural network. The system used identified important risk factors for the prediction of heart disease and it does not require costly medical tests. With using hybrid data mining techniques they have designed more accurate clinical decision support systems for diagnosis of diseases. They have build an intelligent system which could predict the disease using risk factors hence saving cost and time to undergo medical tests and check-ups and ensuring that the patient monitored his health on his own and plan preventive measures and treatment at the early stages of the diseases.

Jabbar *et al.* (2013) have given a research study in which they have presented a lazy data mining approach for heart disease classification. They have applied information centric attribute measure PCA to generate class association rules. The given class association rules used to predict the occurrence of heart disease. The system was designed for Andhra Pradesh population. Andhra Pradesh was in risk of more death due to heart disease. Heart disease handled successfully if more research encouraged developing prediction system in the given area. It has predicted that CVD the most important cause of mortality in India by the year 2015 and AP was in risk of CVD. Hence, a decision support system should be proposed to predict the risk score of a patient which helped in taking precautionary steps like balanced diet and medication which in turn increase life time of a patient.

Rajeswari *et al.* (2011) have proposed Clinical Decision Support System (CDSS) using artificial intelligence techniques for heart disease risk classification. Das *et al.* (2009) have introduced a methodology which used SAS base software 9.1.3 for diagnosing of the heart disease. A neural network ensemble method was in the centre of the proposed system. The methods are of applying machine learning or artificial intelligence methods to aid the cardiologists. Being motivated by the works, we have intended to contribute on heart disease diagnosis system.

Soni *et al.* (2011) have presented an intelligent and effective heart attack prediction system using weighted associative classifier. Different weights have been assigned to the attributes after consulting with expert doctor. Experimental results revealed that WAC was an efficient approach for the extraction of significant patterns from the heart disease dataset. These patterns were stored in rule base in the form of Prediction rules. A little modification had been incorporated in the database and instead of considering 5 class label (4 for four types of heart disease and 1 for no heart disease) they have considered only 2 class labels 1 for "Heart Disease" and another for "No Heart Disease" as the data set was having less number of records for different types of heart disease. The maximum accuracy (81.51%) has been achieved using support value 25% and confidence to be 80%.

From the above literature, the data analysis methods used in most of the heart disease prediction methods cannot supply apparent and direct reason for the decisions prepared to examine the risk factors for cardiovascular diseases as they are depend on neural networks. For this reason, a technique based on simply obtained features talented of calculating the risk level of computer-aided diagnosis and providing elucidation for the decisions made would be of enormous clinical value. Hence, the soft computing technique in meticulous the fuzzy logic technique could be used for assessing the risk level of heart patients in mounting the clinical decision support system of heart disease diagnosis.

Problem definition: Predictive data mining is becoming an essential instrument for researchers and clinical practitioners in medicine. The goal of predictive data mining in clinical medicine is to derive models that can use patient specific information to predict the outcome of interest and to thereby support clinical decision-making. Predictive data mining methods may be applied to the construction of decision models for procedures such as prognosis, diagnosis and treatment planning which once evaluated and verified may be embedded within clinical information systems. The real-life data mining applications are attractive since they provide data miners with varied set of problems, time and again. Working on heart disease patients databases is one kind of a real-life application. The detection of a disease from several factors or symptoms is a multi-layered problem and might lead to false assumptions frequently associated with erratic effects.

Recently, lot of researchers combined fuzzy technique with other technique like neural network, genetic algorithm and decision tree for efficient classification purpose. The reason behind combining the fuzzy with other techniques is that the inability of fuzzy classifier in better prediction without providing the proper fuzzy rules. So, neural network, decision tree and genetic algorithms are used to generate the fuzzy rules and the rules generated from those techniques are then given to the fuzzy rule base for classification process. Here, we have used rough set theory to generate the fuzzy rules and in the fuzzy classifier, the rules generated from the rough set theory are used for prediction with the same step defined. Generally, there are two main steps taken part in the rough-fuzzy classifier such as:

- Rule generation using rough set theory
- Prediction using fuzzy classifier

MATERIALS AND METHODS

Proposed method: Over the past decade, heart disease prediction has been a considerable research as the main cause of death worldwide is due to heart disease (Partu and Khamiss, 2009). In order to forecast the heart disease, numerous researchers combined fuzzy technique with some other method for efficient classification purpose as the fuzzy is efficient only if proper fuzzy rules are specified in the rule base. Now, we have brought in an optimal rough-fuzzy classifier that combined rough set theory with the fuzzy set. Generally, there are three main steps taken part in the rough-fuzzy classifier such as:

- Rule generation using rough set theory
- Rule optimization using Artificial Bee Colony (ABC) optimization
- Prediction using fuzzy classifier

In Fig. 1, the block diagram of the suggested method is displayed. From the Fig. 1, it demonstrates that overall function of the executed method. We regard as patient data base as an input in the method. Data base enclose data like high blood pressure, cholesterol, stress, fats, smoking, diabetes, drug abuse, etc. In the initial step, patient data base is separate as training and testing data. Now, training data is supplied to rule generation. The rule for the classifier is produced by the rule generation next the output is given to the rule optimization. Now, for choosing the best set of rule, the optimization is employed then it sends to the forecast step. Currently, the best set of rule is classified then the classified output is sending to the decision step now the decision is ready if the patient has heart disease or not.

Rule generation: In order to classify the data, rule generation is employed to produce the set of rules. For the generation purpose, Rough Set theory is implemented in the proposed method.

Rough Set theory: Rough Set theory assuring as a dominant theory that dealing with imperfect data. It is a mathematical device to deal with imprecision and insecurity. Using the Rough Set theory, the process of rule generation is made. One of the most important benefits of Rough Set theory is that it does not require any preliminary or further information about data. Let regard as ‘D’ is the data base of heart disease patient containing of ‘X’ number of record and ‘Y’ number of attributes. The detailed process of generating the rules using rough set theory is explained using the below steps.

Step 1: Information/decision system (Table 1 and 2).

Information system: A training data set is signified as a table, where each row signifies a new case or patient (or object) and each column signifies a variable or examination (an attribute) that can be measured for each object. By a user the attribute value is given. This table is known as an information system:

$$B = (X, Y);\ X \in B \tag{1}$$

Where:

X = A non-empty finite set of object

Y = A non-empty finite set of attributes

Example: A very simple information system is shown in Table 1. There are four cases or objects and three attributes like (fats, smoking and stress).

Decision system: In a supervised learning setting, a distinguished feature indicates the outcome (z) i.e., where

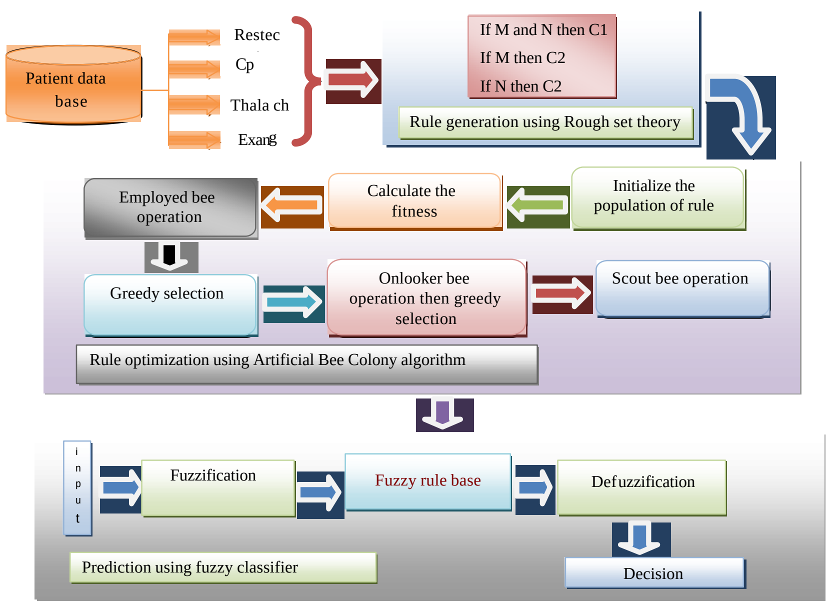

Fig. 1: Block diagram of proposed method

Table 1: An example of information system

Case	Y1	Y2	Y3
X1	12	6	0.20
X2	8	6	0.20
X3	12	6	0.20
X4	12	2	0.12

Table 2: An example of decision system

Case	Y1	Y2	Y3	Decision
X1	12	6	0.20	C1
X2	8	6	0.20	C2
X3	12	6	0.20	C1
X4	12	2	0.12	C2

C1 means decision is Yes and C2 means decision is No

the outcome or classification is recognized which is called the decision attribute. The further features (Y) are referred to as condition attributes. Tables sticking to these conditions are called decision tables. The decision system is any information system of the form that is given below:

$$B = (X, Y \cup \{z\});\ z \notin B \tag{2}$$

A small example decision table can be found in Table 2.

Step 2; Indiscernibility (similarity): Idea of indiscernibility is basic to Rough Set theory. Generally, two objects in a decision table are indiscernible in which one can never be among them on the basis of a given set of attributes. The decision table can be unreasonably large in part because it is redundant in at least two ways; the similar objects may be signified in numerous times or some of attributes may be superfluous. This difficulty is made cleared beneath.

Let B = (X, Y) be an information system, next with any $R \subseteq X$ there is related an equivalence relation $IND_B(R)$:

$$IND_B(R) = \{(a, a) \in X \mid \forall x \in R, x(a) = x(a)\} \tag{3}$$

where, $IND_B(R)$ is called the R-indiscernibility relation. If (a, a')$\in IND_B(R)$. The equivalence class of the R-indiscernibility relation are indicated $(a)_R$. Initial step of Indiscernibility is to discover the equivalence class the second step is to discover the Discernability matrix (D_{ij}) and the next step is to discover the relative discernability function. The equivalence class is signifying in Table 3, here the equivalence class is denoted as 'E'.

The generation of discernability matrix explained below using sample data provided in Table 4. The discernability matrix developed by comparing the attributes of equivalence class. Here, every class is compared to other class in terms of different attributes.

Table 3: Equivalence class

Equivalence class	Y1	Y2	Y3	Decision
E1(X1 and X3)	12	6	0.20	C1
E2	8	6	0.20	C2
E3	12	2	0.12	C2

Table 4: Discernability matrix

Equivalence class	E1	E2	E3
E1	-	Y1	Y2 and Y3
E2	Y1	-	Y1-3
E3	Y2 and Y3	Y1-3	-

The axes come below equivalence classes where the cells have the condition attributes which distinguish it among those classes. For instance, the Y1 attributes is only thing that distinguishes among equivalence classes E1 and E2 whereas all three condition attributes are dissimilar among equivalence classes E2 and E3.

To find the comparative discernability function is the final step of indiscernibility. To calculate the discernability function we are using the discernability matrix which provide the minimum set of features required to distinguish a given class from the others. To assess the relative discernability function, just take each row in turn with concatenate the cells with logical ANDs and concatenate the attributes within each cell with logical ORs. The relative discernability function of our example is given below:

- F(E1) = (Y1) AND (Y2 OR Y3)
- F(E2) = (Y1) AND (Y1 OR Y2 OR Y3)
- F(E3) = (Y2 OR Y3) AND (Y1 OR Y2 OR Y3)

Minimum sets of attributes are necessary in order to distinguish each class. Consider the relative discernability function for E1: to distinguish between E1 and E2 we would like to consider the Y1 attributes; to distinguish between E1 and E3 we could employ either Y2 or Y3. We go for the following step in order to get the minimum features.

Step 3; reduct and core analysis: Reduct and core study is a significant step of rough set theory categorize the significance of the attributes so that the relevant attributes can be able to recognize without compromising the accuracy of classification:

$$\text{Core} = \bigcap \text{reduct}$$

Reduct is computed by taking the relative discernability function and eliminating unnecessary attributes. For instance, in F(E2) we only require Y1 attributes to satisfy the entire function. In F(E3) we require either Y2 or Y3 to satisfy the function. Reduct for our example:

- R(E1) = (Y1 AND Y2) OR (Y1 AND Y3)
- R(E2) = Y1
- R(E3) = Y2 OR Y3

Step 4; rules: The rules are easily constructed once the reduct has been calculated by overlying the reduct over the originating decision table and reading off the value. A decision rule is described to be a report on the form "if M then N" where the condition M is a set of elementary conditions linked by "AND" and the decision N is a set of possible outcomes linked by "OR". For example:

- If Y1 = 12 and Y2 = 6 or Y1 = 12 and Y3 = 0.2 then decision C1
- If Y1 = 12 then decision C2
- If Y2 = 6 or Y3 = 0.2 then decision C2

If a rule contained "or" it was split into two rules to keep the things simple.

- If Y1 = 12 and Y2 = 6 then decision C1
- If Y1 = 12 and Y3 = 0.2 then decision C1
- If Y1 = 12 then decision C2
- If Y2 = 6 then decision C2
- If Y3 = 0.2 then decision C2

Rough Set theory creates the set of rules by this way. If the set of rule is created next the rule is enduring the optimization process.

Rule optimization: Rule optimization is employed to optimize the set of rules and choose the best rule among them. There is an obligation to find the best rule for that Artificial Bee Colony (ABC) algorithm is suggested.

Artificial Bee Colony algorithm: Artificial Bee Colony (ABC) is inspired by the intelligent behavior of honey bees. The position of food source signifies a possible solution in ABC algorithm to the optimization crisis and the nectar amount of a food source corresponds to the quality (or fitness) of the related solution. During population, the number of the employed bees or the onlooker bees is equivalent to the number of solutions. Employed bees, onlooker bees and scout bees are the three main components present in ABC algorithm.

Employed bees: Within the hive, employed bees are pooled with the food sources and they put back the information to the onlookers concerning the nectar quality of the food sources they are using.

Onlooker bees: According to the information presented by the employed bees within the hive, employed bees movement have been noticed by the onlooker to make a choice of one food source to employ.

Scout bees: The employed bees whose food source is discarded turn out to be scout and looking for novel food source randomly.

ABC algorithm, the rules are considered as a population. The fitness value is computed in order to discover the best rule. The fitness value is a function of which chromosome is tested for its suitability to the problem in hand. Using the subsequent formula (Eq. 4) fitness value is computed:

$$F_i(S_{ij}, D^a_{D_{ij}}) = \begin{cases} a = a+1 \text{ if } S_{ij} = \; = D^a_{D_{ij}} \\ a \text{ otherwise} \end{cases} \tag{4}$$

Where:

S_{ij} = The solution
D_{dij} = The discretized format
a = The number of rule replication

The set of rule is diminished with the assist of ABC optimization algorithm, in order to reduce the computational complexity. In Fig. 2, the functional procedure of ABC algorithm is illustrated as below. The main step of ABC optimization algorithm is explained below:

- Initialize the population of rules
- Evaluate the population
- Cycle = 1
- Repeat
- Produce new solution for the employed bee by using the following Eq. 5:

$$\begin{aligned} &S_{ij} = M_{ij} + \gamma_{ij}(M_{ij} - M_{kj}); \\ &j \in \{1, 2, \ldots.. D\} \text{ and } i, k \in \{1, 2, \ldots\ldots N\}, k \neq i \end{aligned} \tag{5}$$

Where:

M_{ij} = The jth parameter of the ith employed bee
S_{ij} = A new solution for Mij in the jth dimension
M_{kj} = The neighbor bee of Mij in employed bee population
γ = A number randomly chosen in the range of (-1,1)
D = The dimension of the problem
N = The number of employed bee

- Use greedy selection process for the employed bees
- Calculate the probability values p_i for the rules X_{ij} using the fitness of the solution:

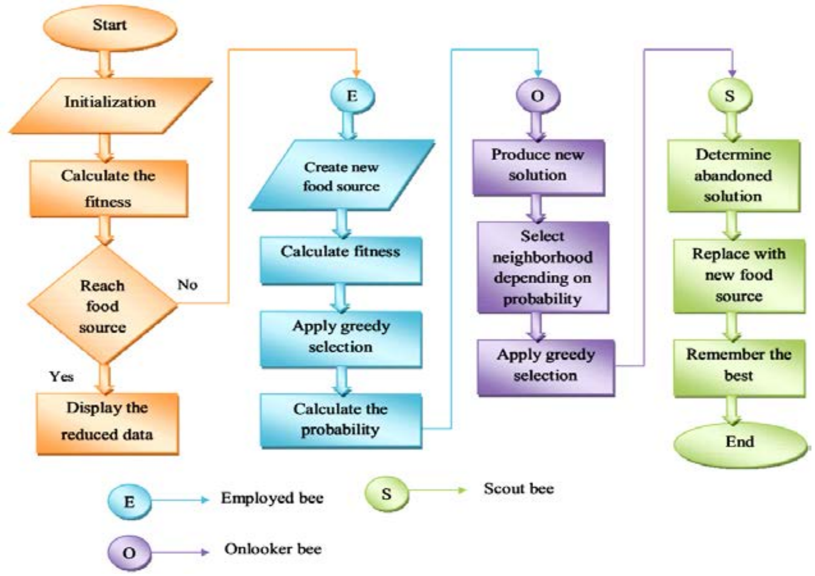

Fig. 2: Flowchart for ABC algorithm

$$P_i = \frac{\text{fitness}_i}{\sum_{n=1}^{SN} \text{fitness}_n} \qquad (6)$$

where, fit_i is a fitness value of ith employed bee. For each onlooker bee, produce a new solution S_{ij} by Eq. 7:

$$S_{ij} = M_{ij} + \gamma_{ij} (M_{ij} - M_{kj}) \qquad (7)$$

In the neighborhood of the solution selected depending on the P_i and evaluate it. Apply selection process between S_i and M_i based on the greedy method for onlooker bee. If scout production is completed, determine the abandoned solutions by using limit parameters for the scout if it exits replace it with a new randomly produced solution by:

$$M_i^j = M_{min}^j + \text{rand}[0,1](M_{max}^j - M_{min}^j)$$
$$i = \{1,2,......N\}; \quad j = \{1,2,.......D\} \qquad (8)$$

Where:

M_i^j = A parameter to be optimized the ith employed bee on the dimension j of the D-dimensional solution space

N = A number of employed bee

M_{max}^j and M_{min}^j = Upper and lower boundaries for M_i^j

- Memorize the best solution achieved so far
- Cycle = Cycle+1
- Until cycle reaches the maximum iteration

The best set of rule is forecasted from the above step and it is preceded for the further process.

Prediction: Using the fuzzy system, Prediction of heart disease is performed. The detailed procedure of fuzzy system is made cleared in the beneath segment.

Fuzzy system: This study explained the fuzzy system proposed for the prediction method. Figure 3 illustrates the detailed procedure involved in the proposed fuzzy system. The most important ideas behind a fuzzy system use the concept of linguistic variables to make decisions based on fuzzy rules and thus get a better response compared to a system by means of crisp values.

Designing of fuzzy system: Proposing of fuzzy system has three significant steps:

- Fuzzification
- Fuzzy inference engine
- Defuzzification

Fuzzification modifies the crisp input to a linguistic variable with the membership function collected in the fuzzy knowledge base. Fuzzy inference engine with the assist of If-Then type fuzzy rules, alters the fuzzy input into the fuzzy output. Defuzzification alters the fuzzy

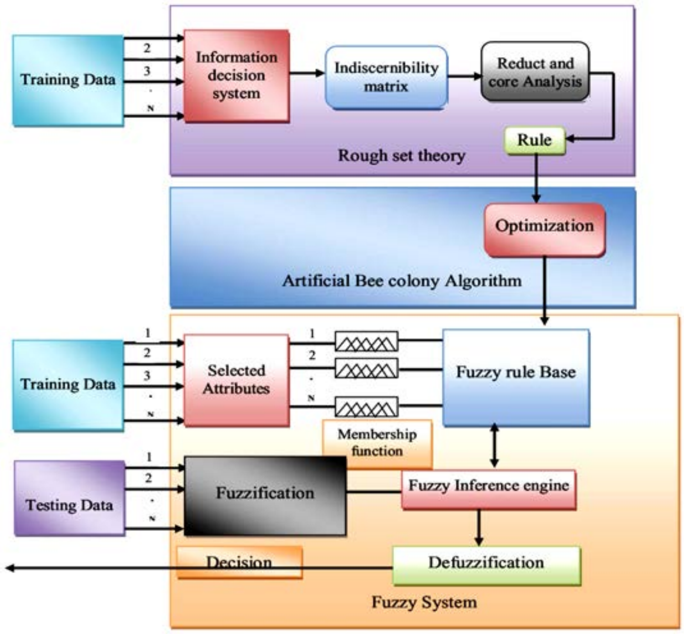

Fig. 3: The detailed procedure involved in the proposed rough-fuzzy classifier

output of the inference engine to crisp by means of membership function equivalent to those utilized by the fuzzfier. Crisp rules are fuzzified inference system through the triangular membership function in our work. Fuzzification is necessary as a degree of membership function is specified for each member of set. The fuzzy system forecasts the results more precisely with the optimized membership function. The detailed procedure involved in the proposed rough-fuzzy system is shown in Fig. 3.

When we are planning the fuzzy system, the two main significant steps to be noticed are definition of fuzzy membership function and fuzzy rule base. Fuzzy Membership function: The membership function is proposed by selecting the proper membership function. Now, we have selected the triangular membership function to amend over the information into the fuzzified value. The triangular membership function contains three vertices l, m and n of in a fuzzy set Q (l: lower limit and n: upper limit where membership degree is zero, m: the centre where membership degree is one). The formula used to compute the membership values is depicted as Eq. 9:

$$f(h)=\begin{cases} 0 & \text{if } h \le l \\ \dfrac{h-l}{m-l} & \text{if } l \le h \le m \\ \dfrac{n-h}{n-m} & \text{if } m \le h \le n \\ 0 & \text{if } h \ge n \end{cases} \tag{9}$$

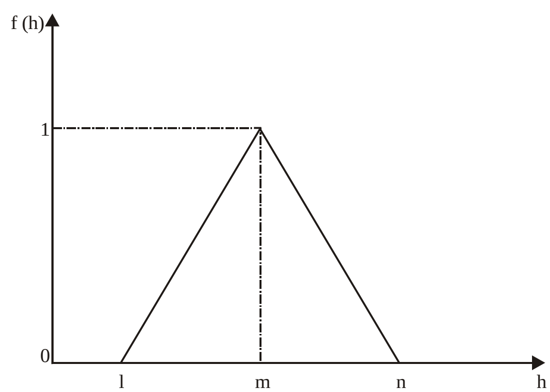

Fig. 4: Triangular membership function

For a single fuzzy set, Fig. 4 shows a triangular membership function. Now, we can see that at l and n the value is zero and it attains constantly to a maximum of value one at the centre point m between the l and n. The triangular membership function demonstrated in beneath.

Rule base: We previously created the fuzzy rule set by means of rough set theory that is specified in the fuzzy rule base. The rule base encloses a set of fuzzy rule in the form of, If Y1 = 12 and Y2 = 6 then decision C1, If Y2 = 6 then decision C2.

Using the fuzzy system the prediction of heart disease is performed. The testing data with decreased attribute is specified to the fuzzy logic system where the test data is changed to the fuzzified value based on the fuzzy membership function. Next, based on the membership function, in the rule base system, the fuzzified input is coordinate with the fuzzy rules defined. After that in the defuzzification, the outcome is specified, now the fuzzified value is transformed to crisp value then the calculation is prepared.

Decision: The decision is produced based on the prediction whether the test data belongs to the heart disease or not.

RESULTS AND DISCUSSION

The experimental result of Rough fuzzy classifier is conversed below. Using MATLAB 2014, the proposed system is executed and the experimentation is carried out with i5 processor of 3GB RAM.

Dataset description: The proposed rough fuzzy classifier is tested with the three dataset namely Cleveland, Hungarian and Switzerland. These three datasets are got from the UCI machine learning repository.

Cleveland data: This data base encloses 76 characteristics on the other hand, all allocated tests refer to exploiting a subset of 14 of them. Particularly, ML researchers employ only the Cleveland database still today. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the Cleveland database have contemplated on simply attempting to differentiate presence (values 1-4) from absence (value 0). The names and social security numbers of the patients were lately eliminated from the database, substituted with dummy values. Due to missing values, six of the examples have been discarded. Class distributions are 54% heart disease absent, 46% heart disease present.

Hungarian data: Owing to an enormous percentage of missing values three of the characteristics have been rejected however the format of the data is exactly the similar as that of the Cleveland data. Due to missing values, thirty-four of the examples have been discarded and 261 examples were there. Class distributions are around 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: In order to assess the efficiency of the proposed system an evaluation metric is employed. General evaluation methodology that contains a set of measures that pursue a some of the metrics that we have chosen for our evaluation purpose are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN), Specificity, Sensitivity, Accuracy, F measure.

Sensitivity: To evaluate of the sensitivity is the percentage of real positives which are precisely known. It relates to the capacity of test to identify positive results:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP+FN})} \tag{10}$$

Where:
TP = True Positive
FN = False Negative

Specificity: To evaluate of the specificity is the level of negatives which are suitably known. It relates to the capacity of test to spot the negative results:

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN+FP})} \tag{11}$$

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:

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{(\text{TN} + \text{TP} + \text{FN} + \text{FP})} \tag{12}$$

F-measure: The F-measure can be employed as a single measure of performance of the test. It is the harmonic mean of precision and recall.

$$\text{F} = 2\times\left(\frac{\text{Precision}\times\text{recall}}{\text{Precision}+\text{recall}}\right) \tag{13}$$

Precision or positive predictive value is the ratio of True Positive to the number of True and False positive value:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{14}$$

In binary classification, recall is called sensitivity:

$$\text{Recall} = \text{Sensitivity} \tag{15}$$

Table 5: Condition and terms of TP, TN, FT and FN

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

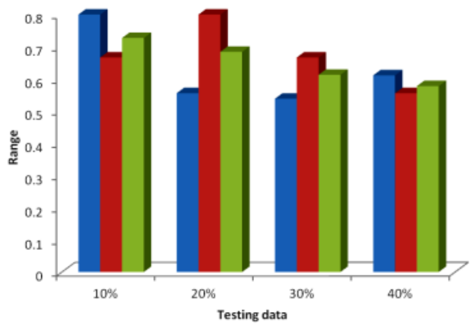

Fig. 5: Chart for sensitivity, specificity, accuracy in Cleveland dataset

The test data result can be positive means predicting that the person has the heart disease or negative means predicting that the person does not have the heart disease:

- True Positive = unhealthy people correctly identified as unhealthy
- True Negative = healthy people correctly identified as healthy
- False Positive = unhealthy people incorrectly identified as unhealthy
- False Negative = healthy people incorrectly identified as healthy (Table 5)

Evaluation metrics for Cleveland data: Figure 5 shows the evaluation metrics for Cleveland dataset. In proposed method, we calculate the evaluation metrics with 90% of attributes as training data and 10% of remaining as testing data and up to 60% of attributes as training data and 40% of remaining as testing data. When the training data is increase the accuracy of the proposed method tends to be improved.

Evaluation metrics for Hungarian data: Figure 6 shows the evaluation metrics for Hungarian dataset.

Evaluation metrics for Switzerland data: Evaluation metrics for Switzerland dataset is shown in Fig. 7.

Sample dataset: This study explained the sample dataset of Cleveland, Hungarian and Switzerland; only some of them are demonstrated in Table 6. The exciting information is that the attributes selected from all three datasets are precisely similar.

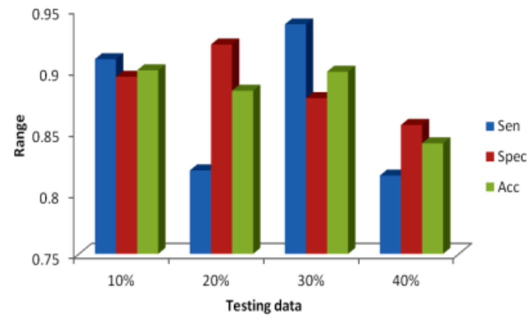

Fig. 6: Chart for sensitivity, specificity, accuracy in Hungarian dataset

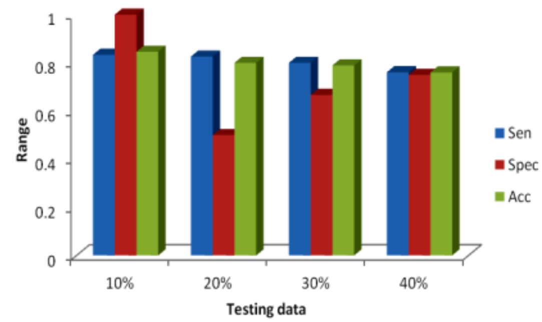

Fig. 7: Chart for sensitivity, specificity, accuracy in Switzerland dataset

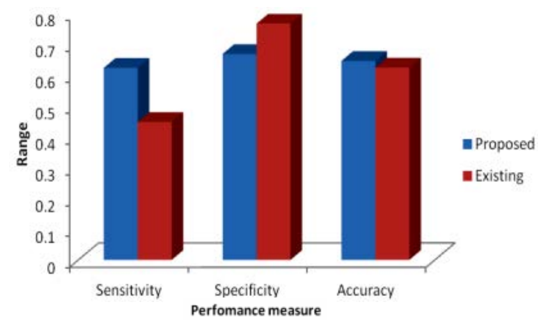

Fig. 8: Chart for cleveland data set is compared to the existing method

Performance analysis comparison of proposed methodology with the existing methods: The presentation of our proposed rough fuzzy classifier work is examined based on the evaluation measures sensitivity, specificity, accuracy and F-measure. The values are tabulated in the subsequent table.

There are three dataset values are engaged in this and each of which is tabulated. In performance measure of Cleveland data set is compared to the existing method (Anooj, 2012) (Fig. 8).

Table 6: Sample dataset of Cleveland, Hungarian and Switzerland

Cleveland selected attribute	Hungarian attribute index	Switzerland selected attribute	Attribute index	Selected attribute	Attribute index
Restecg (resting electrocardiographic results)	6	Restecg	6	Restecg	6
Thalach (aximum heart rate achieved)	8	Thalach	8	Thalach	8
Cp (chest pain)	12	Cp	12	Cp	12
Chol (serum cholesterol)	9	Chol	9	Chol	9
Num (diagnosis of heart disease)	10	Num	10	Num	10

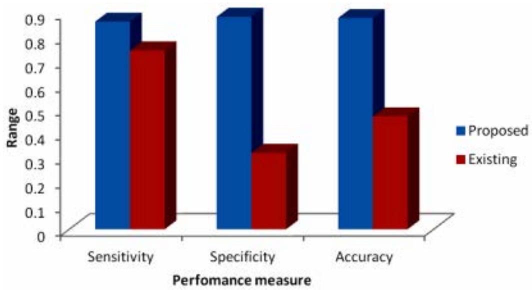

Fig. 9: Chart for Hungarian data set is compared to the existing method

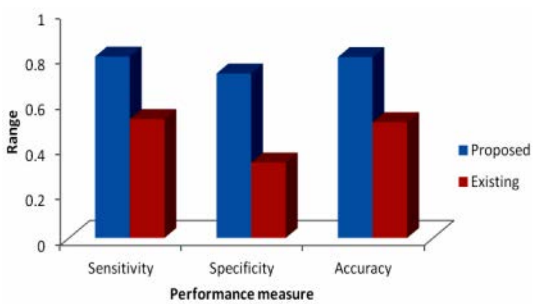

Fig. 10: Chart for Switzerland data set is compared to the existing method

Performance measure of Hungarian data set is compared to the existing method (Fig. 9). Performance measure of Switzerland data set is compared to the existing method (Fig. 10).

CONCLUSION

We have introduced a new classifier namely, rough-fuzzy classifier by blending rough set theory with the fuzzy set. Here, rule generation was done using rough set theory and the optimization was done by ABC and also the prediction was carried out by fuzzy classifier. Two important steps utilized in rule generation process are reduct and core analysis and the discernability matrix formation. Then, the fuzzy system is designed with the help of fuzzy rules given by rough set theory and membership functions. Subsequently, the presence of heart disease is identified by inputting the data to the fuzzy system. Finally, the experimentation is carried out using the Cleveland, Hungarian and Switzerland datasets and the performance was analyzed with sensitivity, specificity, accuracy and F-measure. From the results, we ensured that the proposed rough-fuzzy classifier outperformed the previous approach by achieving the accuracy of 87% in Hungarian datasets.

REFERENCES

Amin, S.U., K. Agarwal and R. Beg, 2013. Genetic neural network based data mining in prediction of heart disease using risk factors. Proceedings of the 2013 IEEE Conference on Information and Communication Technologies (ICT), April 11-12, 2013, IEEE, JeJu Island, South Korea, pp: 1227-1231.

Anderson, K.M., P.M. Odell, P.W. Wilson and W.B. Kannel, 1991. Cardiovascular disease risk profiles. Am. Heart J., 121: 293-298.

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.

Brause, R.W., 2001. Medical Analysis and Diagnosis by Neural Networks. In: Medical Data Analysis. Crespo, J., V. Maojo and F. Martin (Eds.). Springer Berlin Heidelberg, Berlin, Germany, pp: 1-13.

Chen, A.H., S.Y. Huang, P.S. Hong, C.H. Cheng and E.J. Lin, 2011. HDPS: Heart disease prediction system. Proceedings of the Conferences on Computing in Cardiology, September 18-21, 2011, IEEE, Hangzhou, China, ISBN: 978-1-4577-0612-7, pp: 557-560.

Das, R., I. Turkoglu and A. Sengur, 2009. Effective diagnosis of heart disease through neural networks ensembles. Expert Syst. Appl., 36: 7675-7680.

Hannan, S.A., A.V. Mane, R.R. Manza and R.J. Ramteke, 2010. Prediction of heart disease medical prescription using radial basis function. Proceedings of the 2010 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), December 28-29, 2010, IEEE, Coimbatore, India, ISBN: 978-1-4244-5965-0, pp: 1-6.

Jabbar, M.A., B.L. Deekshatulu and P. Chandra, 2013. Heart disease prediction using lazy associative classification. Proceedings of the 2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), March 22-23, 2013, IEEE, Kottayam, India, ISBN: 978-1-4673-5089-1, pp: 40-46.

Mozaffarian, D., P.W. Wilson and W.B. Kannel, 2008. Beyond established and novel risk factors lifestyle risk factors for cardiovascular disease. Circ., 117: 3031-3038.

Palaniappan, S. and R. Awang, 2008. Intelligent heart disease prediction system using data mining techniques. Proceedings of the International Conference on Computer Systems and Applications, March 31-April 4, 2008, Doha, pp: 108-115.

Partu, F.A. and N.N. Khamiss, 2009. Heart diseases diagnosis expert system based on multichannel adaptive resonance theory (MART). Asian J. Inform. Technol., 8: 37-46.

Poirier, P., 2008. Healthy lifestyle even if you are doing everything right, extra weight carries an excess risk of acute coronary events. Circ., 117: 3057-3059.

Rajeswari, K., V. Vaithiyanathan and P. Amirtharaj, 2011. A novel risk level classification of ischemic heart disease using artificial neural network technique -An Indian case study. Int. J. Mach. Learn. Comput., 1: 231-235.

Soni, J. U. Ansari, D. Sharma and S. Soni, 2011. Intelligent and effective heart disease prediction system using weighted associative classifiers. Int. J. Comput. Sci. Eng., 3: 2385-2392.

Steimann, F. and K.P. Adlassnig, 1998. Fuzzy Medical Diagnosis. In: Handbook of Fuzzy Computation, Ruspini, E.H., P.P. Bonissone and W. Pedrycz (Eds.). IOP Publishing, Oxford.

Tavares, T.R., A.L. Oliveira, G.G. Cabral, S.S. Mattos and R. Grigorio, 2013. Preprocessing unbalanced data using weighted support vector machines for prediction of heart disease in children. Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), August 4-9, 2013, IEEE, Dallas, TX., pp: 1-8.

WHO., 2011. Global Atlas on Cardiovascular Disease Prevention and Control. World Health Organization, Geneva, Switzerland.

Wolff, J.G., 2006. Medical diagnosis as pattern recognition in a framework of information compression by multiple alignment, unification and search. Decis. Support Syst., 42: 608-625.

Wood, D., G.D. Backer, O. Faergeman, I. Graham and G. Mancia *et al*., 1998. Prevention of coronary heart disease in clinical practice. J. Hypertens., 16: 1407-1414.