An Improved Ant Colony Optimization with Simulated Annealing based Fuzzy (IACOSAF) Classification System for Diagnosing Heart Disease in Patients

Dr.M. Nithya, Professor and Head, Dept. of CSE, VMKVEC, Salem, Tamilnadu.

S. Kavitha, Research Scholar, VMRF, Salem, Tamilnadu. E-mail:kavitha.murugesan@yahoo.com

Abstract--- The industry of healthcare is a broad domain with an ocean of data concerned to patients, along with enormous medical records growing on a day to day basis. In scientific terms, the healthcare industry is rich with information, but still lacks in-depth knowledge. Data Mining (DM) has enormous potential and plays an unavoidable role in this field. Today DM methodologies could be employed for the prediction of different kinds of disease. Nonetheless, data mining with its different analytical tools and schemes has a significant role in reducing the usage of difficult tests conducted on patients for detecting disease. The objective of the research is to apply and evaluate data mining strategy for the prediction of heart disease in a patient by extracting the risk levels from the dataset that employs essential parameters. The enhanced fuzzy classification based data mining classifier is used here on healthcare data Simulated Annealing (SA) and Adaptive Mutation Operators (AMO) have been integrated into ACO algorithm to resolve the low convergence rate and provides the capability of speeding up the convergence. The Improved Ant Colony Optimization with Simulated Annealing in Fuzzy (IACOSA) can work and focus more efficiently on the classification process to predict the major cause of health failure the heart disease.

Keywords--- ACO, AMO, Fuzzy, Optimization, Simulated Annealing, Classification, Health Care.

I. Introduction

A huge volume of data were stored in the database, data warehouses, or other kinds of data repositories, due to the rapid development in computerized data acquisition and storage modernization, Eg: World Wide Web (WWW). In the vast amount of raw data, valuable information lay hidden. It is a tedious process to extract the data from the overwhelming data without any powerful tool. To overcome this issue, a new way of approach called data mining emerged to extract the details from the huge amount of data, which works with the assistance of the ubiquitous modern computing device[9].

As the saying goes, “ Prevention is better than cure’’ predicting and preventing the heart diseases at an earlier stage has become a need of high priority because every year, the cause of death due to heart disease has increased rapidly due to changing lifestyle patterns coupled with lack of physical exercise.

Data mining approach is concerned here with figuring out the most useful ,non-trivial fact from the huge amount of data stored in the databases, data warehouses and database repositories [5].

Medical diagnosis is treated to be an important yet delicate job, which has to be performed with precision and efficiency. Its automation would be greatly advantageous. Clinical decisions are frequently taken on the basis of the intuition and experience of the doctor[17]. It is based on the knowledge rich data that stays concealed in the database, which results in unnecessary biases, errors and huge medical expenses that impact the quality of service given to patients. Data mining provides the means to create a knowledge-rich environment that can aid in considerably improving the quality of clinical decisions.

Fuzzy Logic [16] is dependent on intuitive reasoning and takes human subjectivity and imprecision into account. Dissimilar to conventional data mining methodologies like cluster or regression analysis, fuzzy logic facilitates the usage of classification through fuzzy propositions integrated by fuzzy logic operators.

Classification rules were created based on the datasets from these approaches. The created fuzzy rules will create the same rules again and again, to optimize this issue; ACO algorithm is brought-in. But, this algorithm has few demerits like trapping into local minimum and low convergence rate[4]. Simulated Annealing (SA) and Adaptive Mutation Operators have the capability and global convergence; and local search will increase the speed.

Received: 23 July 2018/Accepted: 21 Aug 2018

IACOSA algorithm is utilized to optimizing the generated rules and these rules were given to the fuzzy system, so this results in perfect decision[13]. The proposed IACOSAF algorithm efficiently identifies the heart disease and also it utilizes the accuracy as a metrics.

II. Background Knowledge

In the modern world, among various diseases that affect people, cardiovascular diseases are considered to the highest one [11]. The survey says, because of heart problems, more than 12 million deaths happens in an average, as a report given by the World Health Organization. It could be thus inferred that this disease increases the fatalities. It is a complicated process to rectify, because it has to be analyzed perfectly.

As per the National Heart, Lung, and Blood Institute, Heart disease is a common name given for an array of diseases, disorders and conditions affecting the heart and the blood vessels too. Heart disease can be deemed as the number one cause of death in women and men in the United States (US), and in addition, more than a million Americans suffer with myocardial infarctions. The symptoms of heart disease differ based on the particular kind of heart disease[11]. Chest pain is one common symptom of heart disease.

Han Kamber et al [10]referred that Data mining methodologies are utilized in several fields like Future Healthcare, Market Basket Analysis, Education, Manufacturing Engineering, Fraud Detection, Intrusion Detection, Lie Detection, Customer Segmentation, Financial Banking, Corporate Surveillance, Research Analysis, Criminal Investigation and Bio Informatics.

Yu et al [8] suggested one simple modified Ant Colony Optimization (ACO) algorithm for selecting the tumor-related marker genes, and Support Vector Machine (SVM) is employed as the classifier for evaluating the performance of the gene subset extracted.

Parthiban and Subramanian [6] demonstrated four algorithms consisting of Rule based, Decision Tree (DT), Naive Bayes (NB) and Artificial Neural Network (ANN) to the voluminous healthcare data. The Tanagra data mining tool is exploited for the process of data analysis.

Salama and Freitas [14] have investigated the usage of different classification quality measures for assessing the classifiers built by the ants. The goal of this examination is to find the way in which the usage of various evaluation measures impacts the quality of the output classifier with regard to predictive accuracy.

III. Issues in Ant-Colony Optimization (ACO) Algorithm

For resolving the combinatorial optimization issues, the Ant Colony Optimization (ACO) is utilized which is a meta-heuristic technique and it works on swarm intelligence[19]. ACO algorithm identifies the optimal path; its nature simulates that of an ant’s behavior, which searches for the food. The ant’s nature is marked by walking randomly, and in case it finds the food, it leaves a chemical substance on its path that it traverses which is termed as pheromone, through which the ants can recognize the food path.

In data mining concepts various researches has been done, by utilizing ACO algorithm. In this algorithm, solutions acquired are according to the pheromones deposited through the proceeding ants. The laid chemical (pheromones) dissolves over time. So, the solution depends on the way an identification of the food path from their home is determined.

Diversity of the population will be lost through the premature convergence and hence the algorithm gets bound with local optima. Hence, it is required to manage the diversity and to make the tradeoff among the diversification and intensification through joining two or more algorithms, in order to give high-quality solutions and also to increase the execution time [3] while considering the large-scale instances. The performance of these studies will desperately get reduced, in ACO algorithm.

To hybridize elitist ant system[12] with SA, mutation, and local search, no research has been worked-out. So, in our work, Travelling Salesman Problem (TSP), a new hybrid elitist ant system[7] with SA, mutation operator, and local search procedure with shortest time is brought-in to find most probable features to predict the disease. To get away from local optima[2], SA assists ACO. Alternatively, defining the basic solution of SA is a complex task. So, ACO helps in creating SA basic solution. The performance of algorithm, expansion of diversity population and restricting of immature convergence, will be improved by bringing in mutation operation to ACO algorithm. Either SA or mutation, depends on the diversity level of the population, once after enforcing any one, elitist ant system proceed through the local search procedure to increase the speed of the convergence.

Received: 23 July 2018/Accepted: 21 Aug 2018

Nature Inspired Meta-Heuristics

Stochastic algorithms are classified into two varieties: heuristic and meta-heuristic, but their variation is less. Loosely speaking heuristic refers to 'to find' or 'to discover by trial and error'. Quality solutions to a tough optimization issue are detected in a moderate timing, but optimal solution isn't assured. These algorithms works fine occasionally, but not regularly. To proceed from local search to the global scale, randomization helps. Hence, nearly entire meta-heuristic algorithm aims to be flexible for global optimization. Practically, an acceptable solution for a difficult problem is produced by Heuristics, which works in a trial and error method[20]. The tediousness makes it impractical to obtain the possible solution or combination, when the target is to identify the suitable solution in a reasonable time.

Intensification and diversification are the two major components of meta-heuristic algorithms. Diversification creates diverse solutions, to explore the search space on the global scale, whereas intensification concentrates on the search in a local region researching the details which gives a perfect solution in these areas.

In many ways the meta-heuristic algorithms can be classified, population-based and trajectory-based is one among them. For instance: genetic algorithms work is s based on the former one, which utilizes a set of strings and in Particle Swarm Optimization (PSO) multiple agents or particles were utilized[1]. Alternatively, SA utilizes the single agent or solution, which proceeds in the design space or search space in a piecewise style. But perfect solution is usually accepted, while a not-so-good move is accepted with a specific probability. In the search space, the steps and moves draw a route, with a non-zero probability that this route can attain the global optimum.

IV. General Procedure of Simulated Annealing (SA)

SA is a trajectory-based optimization technique and generally it is an iterative enhancement strategy with the standards that assume the greater configuration many times. To resolve the combinatorial optimization issue, the 1st try is to enforce SA, which was in 80s of the last century [15]. A survey of SA, its establishment in way of theory and SA was aspired by physical annealing process of solids, where it is heated in the beginning and then cooled slowly to attain its state of energy.

In a thermodynamic system, metropolis acceptance criterion was that the initial state which is selected at energy G and temperature T. The constant value T, the initial configuration of the system is bothered to generate a new configuration and the modification in energy ΔG is computed. The new configuration is approved unconditionally if $\Delta G$is negative because it is confirmed that if ΔG is positive with a probability provided by the Boltzmann factor shown in (4.1) to prevent caught in the local optima:

$$exp^{-\Delta Cost/Temperature} \quad - \quad 1.1$$

The process gets iterated till it attains the good sampling statistics for the present temperature i.e feature selected and it will be reduced and then the process again gets iterated till a frozen state is attained (free energy state) at $T = 0.$ The analogy among the states of a physical system and optimization issue is provided as follows:

1. The present configuration of the thermodynamic system is same as the present solution of the optimization issue;
2. The thermodynamic system energy is same to the objective function of optimization issue; and
3. Ground status of the thermodynamic system is same to the global minimum of the optimization issue.

Algorithm 1.1: Simulated Annealing (SA)

```
Simulated Annealing
Begin
Solution        =   Initial Solution
Cost            =   Evaluate(Solution)
Temperature     =   Initial Temperature
While (Temperature > Final Temperature) do
New Solution    =   Mutate(Solution)
New Cost        =   Evaluate(New Solution)
ΔCost           =   New Cost – Cost
if (ΔCost ≤ 0) OR (e−ΔCost/Temperature > Rand) then
Cost            =   New Cost
Solution        =   New Solution
End
Temperature     =   Codling rate × Temperature
End
Return the best solution
End
```


Received: 23 July 2018/Accepted: 21 Aug 2018

V. Proposed Methodology

Initially, the input medical dataset is given to the proposed work. Based on these dataset, the rule for fuzzy rules generation for classification is done. From the generated fuzzy rules process, observed that the same rules are generated repeatedly.

To overcome the problem, an Improved Ant Colony Optimization with SA (IACOSA) Algorithm is introduced to generate optimized rules. These rules are given into the fuzzy to perform the fuzzification process; finally the rules generated should undergo the defuzzification process, which helps for the decision making about the diseases as illustrated in figure 1.1.

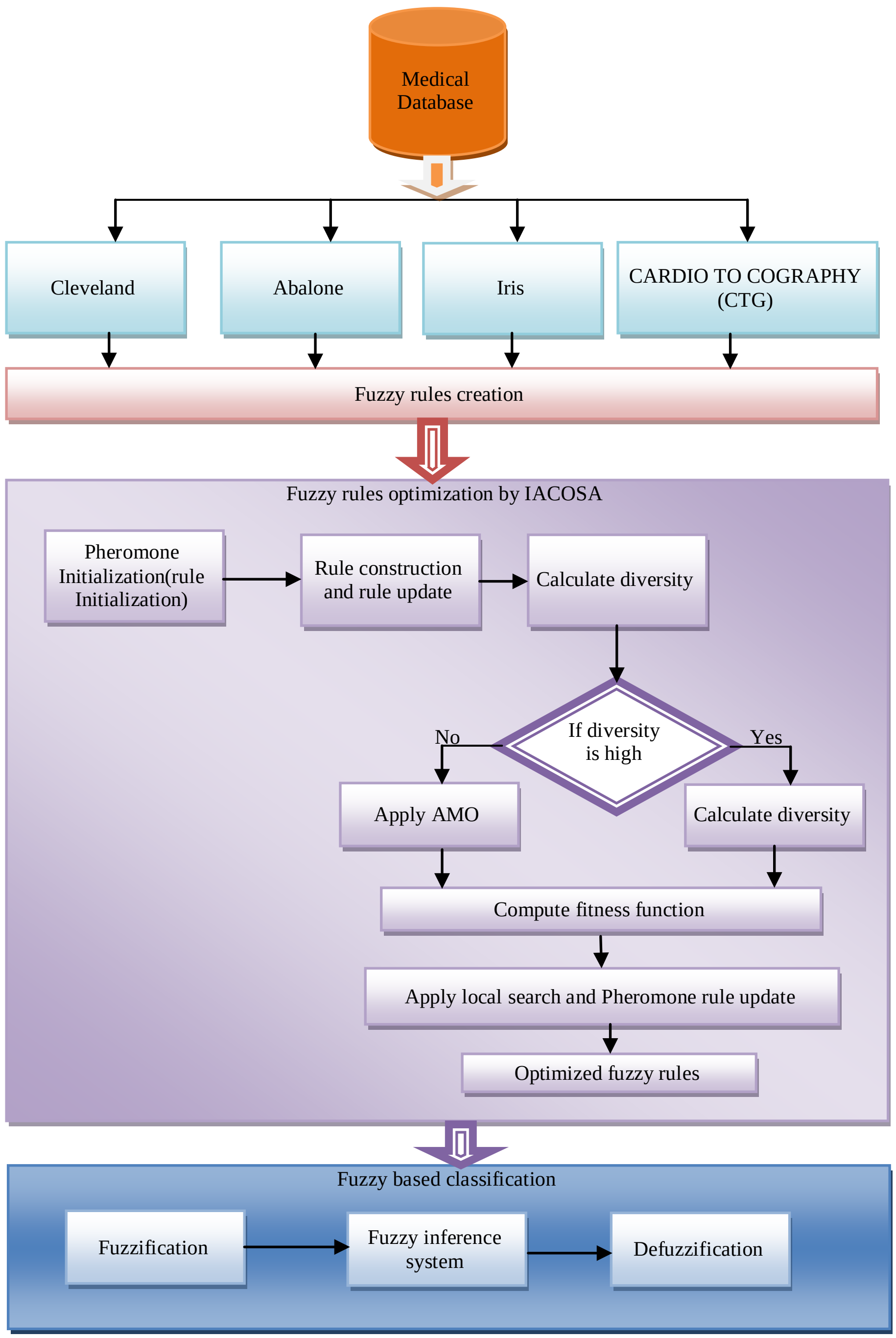

Figure 1.1: Block Diagram for Heart Disease Prediction

Received: 23 July 2018/Accepted: 21 Aug 2018

Improved Ant Colony Optimization Simulated Annealing (IACOSA) Algorithm

Fuzzy rule creates rules. Here, Simulated Annealing (SA) elitist ant system with Adaptive Mutation Operator (AMO)[16] and local search for resolving the rule optimization is brought-in. Provided with n number of rules, initially, the elitist ant system produces population with m ants and every ant select a city as its own source of path to begin. The pheromone intensity level among any two cities is initialized with small positive constant s_0.

The iterations counter which is set to zero initially, will record the repetitions implemented in the proposed algorithm. For each Interval (Int) while deploying the algorithm, where i is a predefined number, the elitist ant system will call either SA or Adaptive Mutation Operators (AMO), according to the diversity of the elitist ant system, and also to enhance the performance of the algorithm. If the diversity is higher than 0.5, intensification is required, that can be accomplished by the simulated annealing for ratio of solutions pool. If in case, the diversity is less than 0.5; the algorithm will loss its diversity and remains in the local minima. So, the algorithm demands to raise the diversity by enforcing the AMO with a predefined likelihood. By computing the Euclidean Distance (ED), the diversity among the fitness of the ants in the population will be measured.

$$ED = \frac{\overline{ds} - ds_{min}}{ds_{max} - ds_{min}} \qquad 1.2$$

where 'd' is average distance among the fitness of the best ant and the fitness of the rest of the ants in the population. $ds_{\min}$ and $ds_{\max}$ are the distances of the worst ant fitness and the second best ant fitness from the fitness of best ant correspondingly (Herrera and Lozano 1996). ED ranges between 0 and 1. If ED is low, most ants in the population focus around the best ant and so the convergence will be accomplished. If it is high, most of the ants are not biased near the existing best ant.

Therefore, ED gives a description for the population variation and the lack of similarity between ants. Using the subsequent equation (1.3) the quality of a rule is calculated.

$$Q = \left[\frac{TruePositive}{TruePos + FalseNeg}\right] * \left[\frac{TrueNegative}{FalsePos + FalseNeg}\right] \qquad 1.3$$

where True Positive(TP) is the number of cases wrapped by the rule and containing the similar class as that forecasted by the rule, False Positive(FP) is the number of cases wrapped by the rule and containing a dissimilar class from that forecasted by the rule, False Negative(FN) is the number of cases that are not wrapped by the rule, while containing the class predicted by the rule, True Negative(TN) is the number of cases that are not wrapped by the rule which have a dissimilar class from the class forecasted by the rule. The following are the steps of the proposed algorithm.

Step 1: After the initial stage mentioned above, each ant in the population builds its tour by applying the transition rule followed by the local pheromone update rule:

1. **Transition Rule:** The i^{th} ant decides the next city j to be visited according to (1.4).The term selection is based on the probability as specified by the equation (1.4)

$$p_{ij(t)} = \frac{\tau_{ij(t)}.\eta_{ij}}{\sum_{i`}^{a}\sum_{j}^{bi}\tau_{ij}(t).\eta_{ij}, \forall i \in I} \qquad 1.4$$

   where η_{ij} is a problem-dependent heuristic value for $term_{ij}$, τ_{ij} is the amount of pheromone presently accessible (at time t) on the connection among attribute i and value I is the set of characteristics that are not so far applied by the ant in the domain of attribute I.

2. **Local Pheromone Update Rule:** After all ants complete their tours, the local update rule of the pheromone trails is applied for each route according to the following equation (4.5).

$$\tau_{ij}(t+1) = \tau_{ij}(t) + \tau_{ij}(t) * Q \ \forall term_{ij} \in \qquad 1.5$$

Thus, the local update rule is decreasing the pheromone trails by a constant factor (pheromone evaporation).

ISSN 1943-023X

Received: 23 July 2018/Accepted: 21 Aug 2018

Step 2: Calculate the route cost of each ant. After that, apply the global pheromone update rule in which the amount of pheromone is added to the best route which has the lowest cost. This rule is defined in (1.6).

$$\Delta\tau_{ij}^{gb}(t) = \begin{cases} \dfrac{e}{L^{gb}(t)} & if \quad edge(i,j) \in T^{gb} \\ 0, & otherwise \end{cases} \tag{1.6}$$

where T^{gb} is the best route, $L^{gb}(t)$ is the distance of the best route, and *e* is a positive integer. This means that the edges belonging to the global-best tour get an additional amount of pheromone each time the pheromone is updated.

Step 3: Calculate the diversity of the population. If the diversity is high, the algorithm needs intensification by applying SA. Otherwise, the algorithm needs to regain the diversity by applying the Adaptive Mutation Operators (AMO).

Simulated Annealing (SA): If in case the diversity value is higher than 0.5, then the elitist ant system will utilize the current approach to improve the chosen ant by enciphering its tour which is attained by the ant system into SA. Here, the tour of the chosen ant is enciphered as the initial state of SA.

Initialize the system temperature to T_0. Two bit in the ant *A* is chosen for interchange and to create the new mutate ant A' and its energy (cost) is computed appropriately. If the energy of ant A' is good than ant *A* or a random number created among 0 and 1 is less than the Boltzmann factor as determined by (4.1), then the new ant A' is authorized. Else, ant *A* stays without modification.

It is required to minimize the temperature by annealing schedule or by factor $0.0 < \alpha < 1.0$. Repeatedly, it will be iterated till T_0 attains a predefined low temperature. If tour path of ant A' is good than ant *A*, then the improved ant A' is added in the ant's population. Else, the ant *A* remains without modifications and returns back into the ant's population.

Adaptive Mutation Operators (AMO): Three AMO is brought-in to manage the diversity of the suggested IACOSA algorithm, for exploring the new research area. The elitist ant system will utilize the mutation operation to improve the chosen ant, if the diversity value is less than 0.5.

Three mutation operators are: the Cauchy mutation operator, the Gaussian mutation operator, and the Levy mutation operator. All three has similar initial selection ratio of 1/3. Every operator is enforced based on the selection ratio n and it's offspring fitness is computed. Continuously, the appropriate mutation operator is selected randomly and manages entire mutation behaviour in the entire ACO. $\text{Prog}_i(t)$ of operator i at generation t is determined, to describe the updating equation for the adaptive mutation operator in ACO.

$$prog_i(t) = \sum_{j=1}^{M_i} f\left(p_j^i(t)\right) - \min\left[f\left(p_j^i(t)\right), f\left(c_j^i(t)\right)\right] \tag{1.7}$$

where $p_j^i(t)$ and $c_j^i(t)$ indicates a parent and its child created by mutation operator *i* at generation correspondingly, and M_i is the number of ants which choose the select mutation operator *i* to mutate. The reward value $\text{reward}_i(t)$ of operator *i* at generation *t* is determined as:

$$Reward_i(t) = \exp\left(\frac{prog_i(t)}{\sum_{j=1}^{N} prog_j(t)}.\alpha + \frac{s_i}{M_i}(1-\alpha)\right) + c_i p_i(t) - 1 \tag{1.8}$$

where s_i is the ants count whose children have good fitness than themselves after being mutated by mutation operator *i*, $p_i(t) \in Q$ is the selection ratio of mutation operator *i* at generation *t*, *α* is a random weight between

Received: 23 July 2018/Accepted: 21 Aug 2018

$(0,1)$, N is mutation operator's count and c_i is a penalty factor for mutation operator i, which is determined as :

$$c_i = \begin{cases} 0.9, & if \ \ s_i = 0 \ \ and \ \ p_i(t) = \max_{j=1\,to N}\left(p_j(t)\right) \\ 1, & Otherwise \end{cases} \quad 1.9$$

The selection ratio of the present best operator will be minimized, if the earlier best operator doesn't offer at the present generation. The selection ratio of the mutation operator 'i' is updated with the above equation which depends on the following equation (1.10)

$$p_i(t+1) = \frac{reward_i(t)}{\sum_{j=1}^{N} reward_j(t)}(1 - N * \gamma) + \gamma \quad 1.10$$

where γ is the minimum selection ratio for every mutation operator that is set as 0.01 for entire experiment. The factors considered in selection ratio update equation are

- the progress value,
- ratio of successful mutations,
- previous selection ratio and
- the minimum selection ratio.

The next significant parameter for the adaptive mutation operator is the frequency of updating the selection ratios of mutation operators, which is the selection ratio of every mutation operator which can be updated at a fixed frequency, e.g., every U_f generation, rather than each generation.

Initial values are allotted for all the mutation ratios of these operators. Eg: 0.1.Every mutation operator [18] is enforced by its mutation ratio. The progress value is computed after the mutation operation by utilizing the following equation:

$$progress_i(t) = \sum_{j=1}^{M_i} \max\left[f\left(p_j^i(t)\right), f\left(c_j^i(t)\right) - f\left(p_j^i(t)\right)\right] \quad 1.11$$

where progress $_i(t)$ is the progress value of operator i at generation t, f is the fitness of an individual, $p_j^i(t)$ and $c_j^i(t)$ are the parent and its offspring created by mutation operator i at generation(t)and M_i denotes the individual's count to choose the mutation operator i to mutate. The mutation ratio of operator i is updated based on their average progress value at generation (t), based on the following equation (1.12)

$$p_i(t+1) = \frac{progress_i(t)}{\sum_{j=1}^{N} progress_j(t)}\left(P_{mutation} - N * \delta\right) + \delta \quad 1.12$$

where $p_i(t)$ is the mutation ratio of mutation operator i at generation t, N is the mutation operator's count, $\delta = 0.01$ is the minimum mutation ratio for every mutation operator and $P_{mutation}$ denotes the initial mutation probability.

In this process, an ant is selected randomly with a $P_{mutation}$ probability. If a random number between 0 and 1 is less than the predefined mutation rate, then the algorithm selects an ant to mutate. Two bits in the ant A are selected randomly for exchange to generate a new mutate ant A' and calculate its tour.

Step 4: Apply the local search procedure for further enhancement.

Step 5: Again, apply the global pheromone update rule according to equation (4.6).

Step 6: If the termination condition is satisfied, then return the best route with its length. Otherwise, go to step 1. Figure 1.2 shows the flowchart of the proposed algorithm.

Perform all these steps in the ACO optimization. These rules are given into the fuzzy to perform the fuzzification process; finally the rules generated should undergo the defuzzification process, which helps for the decision making about the disease.

Received: 23 July 2018/Accepted: 21 Aug 2018

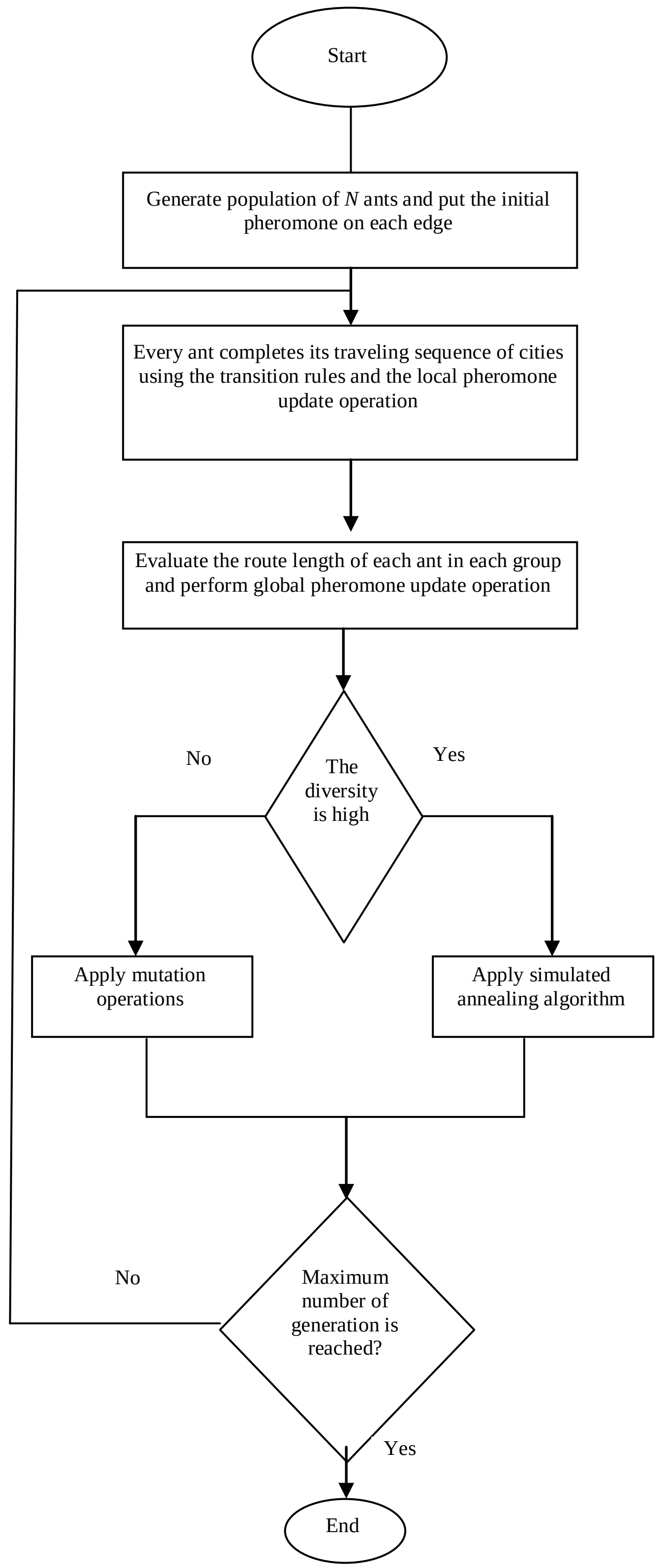

Figure 1.2: The Flowchart of the Proposed IACOSA Algorithm

In this fuzzy based classification method IACOSAF the heart disease is predicted comparatively better with more accuracy.

ISSN 1943-023X

Received: 23 July 2018/Accepted: 21 Aug 2018

Fuzzy Logic based Classification

Fuzzy based classification is a method of generating a mapping from a given input to an output using fuzzy logic. Then, the mapping gives a basis, from which decisions can be generated. Membership Functions, Logical Operations, and If-Then Rules are used in the Fuzzy Logic based Process. The Stages of Fuzzy are,

1. Fuzzification,
2. Fuzzy Inference Engine and
3. Defuzzification.

Fuzzification: Adapts the crisp input to a linguistic variable with the membership function accumulated in the fuzzy knowledge base.

Fuzzy inference engine: With the help of If-Then type fuzzy rules, alters the fuzzy input into the fuzzy output.

Defuzzification: Defuzzification changes the fuzzy output of the inference engine to crisp using membership function equivalent to those exploited by the fuzzifier. Crisp rules are fuzzified inference system through the triangular membership function in our effort. Fuzzification is required as a degree of membership function is defined for each member of set. The fuzzy system predicts the results more specifically with the optimized membership function.

VI. Results and Discussion

The experimental result of fuzzy based classifier is discussed below. The proposed system is implemented using MATLAB 2014 and the experimentation is performed with i5 processor of 3GB RAM.

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 choose for the evaluation purpose are TP, TN, FP and FN, Specificity, Sensitivity, Accuracy, F measure.

a)Performance Analysis

The performance of the proposed heart diseases prediction methods of human beings are evaluated by the six metrics Precision, Recall, F-Measure, Sensitivity, Specificity and Accuracy. The results of proposed work help to analyze the efficacy of the prediction process. The subsequent table 4.1 tabulates the results. Here, only the results of the four datasets are considered.

Table 1.1: Results of the Proposed IACOSAF based Prediction System with Datasets

Dataset	Performance metrics (IACOSAF) (%)				
	Sensitivity	Specificity	Precision	F-measure	Accuracy
CTG	88	80	84.6154	81.553	88
Abalone	85.7143	74.5455	81.0811	83.33	85.714
Iris	88	84	84.6154	86.274	88
CHD	93.3333	94.7619	97.2705	95.261	94.60

From table 1.1, the evaluation metrics are analyzed for the four different numbers of datasets; it observes the efficiency of proposed IACOSAF classification system. The results of the measures are Sensitivity, Specificity, Precision, F-Measure and Accuracy is graphically represented in figure 1.3.

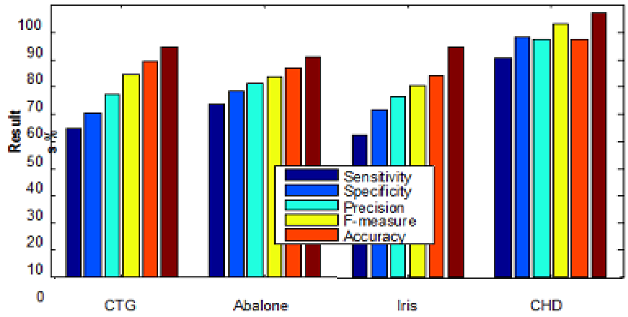

Figure 1.3: Performance Measures such as Precision, Recall, F-Measure, Specificity and Sensitivity, Accuracy (IACOSAF Classifier)

Received: 23 July 2018/Accepted: 21 Aug 2018

The sensitivity results of four dataset samples are 88%, 85.71%, 88%, and 93.33% and the specificity results for the all the dataset samples are 80.00%, 74.545%, 84.00% and 94.761%. The precision values for the four datasets are 84.6154%, 81.081%, 84.6154%, and 94.7619%. The F-measure values for the four datasets are 81.553%, 83.33%, 86.27%, and 95.2612%. At finally the accuracy values for the four datasets are 88%, 85.714%, 88%, and 94.60%.

b) Comparative Analysis

The literature review works are compared in this section with the proposed work to show that the proposed work is better than the state-of-art works. Established that the proposed work helps to attain very good accuracy in heart disease prediction of medical dataset using Fuzzy classifier. And also establishes the prediction accuracy outcome by comparing with other classifiers. The comparison outcomes are presented in the following table 1.2.

Table 1.2: Comparison of Sensitivity Values in Proposed IACOSAF Classifier Vs. Existing Classifiers

Datasets	**Sensitivity (%)**			
	Fuzzy	**ACOF**	**IACOF**	**IACOSAF**
CTG	60	76	84.00	88
Abalone	57.1429	74.2857	80.00	85.7143
Iris	48	76	80.00	88
CHD	69.0476	82.1429	87.62	93.3333

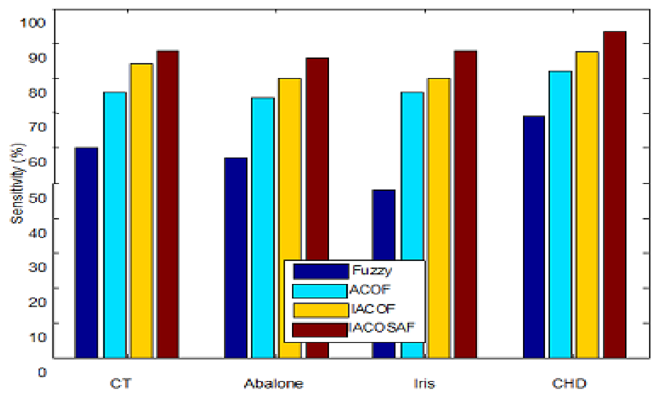

Figure 1.4: Sensitivity Comparison between Proposed IACOSAF Vs. Existing Fuzzy Classifiers

Figure 1.4 explains the sensitivity outcomes of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers. The sensitivity results of IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 88%, 84%, 76%, and 60% respectively for CTG samples, it concludes that the proposed IACOSAF system is 4%, 12% and 28% higher when compared with IACOF, ACOF and fuzzy classifiers respectively.

For Abalone dataset, the sensitivity results of the IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 85.7143%, 80.00%, 74.28%, and 57.1429% respectively, here the proposed IACOSAF classifier is 5.7143%, 11.4286% and 28.5714% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

For Iris dataset, sensitivity results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 88%, 80%, 76%, and 48% respectively, here the proposed IACOSAF classifier is 8%,12% and 40% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

For CHD dataset, sensitivity results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 93.33%, 87.62%, 82.1429%, and 69.048% respectively, here the proposed IACOSAF classifier is 5.7133%, 11.1904%, and 24.2857% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively (Refer table 1.2).

Table 1.3: Comparison of Specificity Values in Proposed IACOSAF Classifier Vs. Existing Classifiers

Datasets	**Specificity(%)**			
	Fuzzy	**ACOF**	**IACOF**	**IACOSAF**
CTG	37.5000	67.5000	72.5	80
Abalone	58.1818	66.3636	69.09	74.5455
Iris	56.0000	68	72	84
CHD	66.6667	88.0952	90	94.7619

ISSN 1943-023X
Received: 23 July 2018/Accepted: 21 Aug 2018

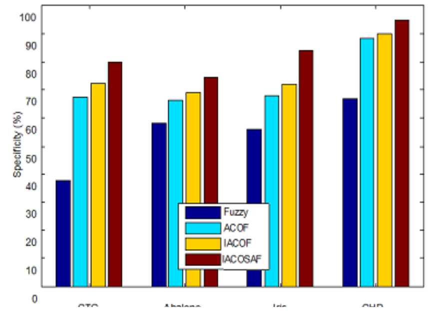

Figure 1.5: Specificity Comparison between Proposed IACOSAF Vs. Existing Fuzzy Classifiers

Figure 1.5 explains the specificity outcomes of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers.

The specificity results of IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 80%, 72.5%, 67.5000%, and 37.5000% respectively for CTG samples, it concludes that the proposed IACOSAF system is 7.5%, 12.5% and 42.5% higher when compared with IACOF, ACOF and fuzzy classifiers respectively.

For Abalone dataset, the specificity results of the IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 74.546%, 69.09%, 66.364%, and 58.182% respectively, here the proposed IACOSAF classifier is 5.455%, 8.18% and 16.3637% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively. Similarly for Iris dataset and CHD dataset proposed IACOSAF higher when compared with IACOF, ACOF and Fuzzy classifiers respectively (Refer table 1.3).

Table 1.4: Comparison of Precision Values in in Proposed IACOSAF Classifier Vs. Existing Classifiers

Datasets	**Precision (%)**			
	Fuzzy	**ACOF**	**IACOF**	**IACOSAF**
CTG	54.5455	74.5098	79.2453	84.6154
Abalone	63.4921	73.7589	76.7123	81.0811
Iris	52.1739	70.3704	74.074	84.6154
CHD	80.5556	93.2432	87.610	97.2705

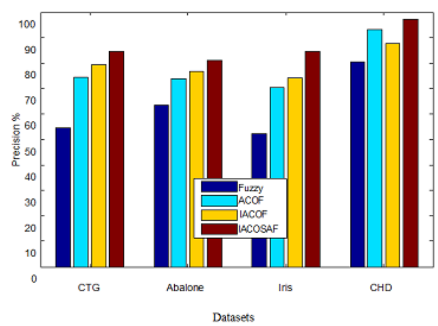

Figure 1.6: Precision Comparison of Proposed IACOSAF Vs. Existing Fuzzy Classifiers

Figure 1.6 explains the precision outcomes of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers.

The precision results of IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 84.615%, 79.245%, 74.51%, and 54.546% respectively for CTG samples, it concludes that the proposed system is 5.3%, 10.10% and 30.0699% higher when compared with IACOF, ACOF and fuzzy classifiers respectively.

For Abalone dataset, the precision results of the IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 81.081%, 76.712%, 73.759%, and 63.492% respectively, here the proposed IACOSAF classifier is 4.3688%, 7.322% and 17.589% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

For Iris dataset , precision results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 84.615%, 74.074%, 70.37%, and 52.174% respectively, here the proposed IACOSAF classifier is 10.5414%,14.245% and 32.4415% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

ISSN 1943-023X

Received: 23 July 2018/Accepted: 21 Aug 2018

For CHD dataset, precision results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 97.271%, 87.61%, 93.243%, and 66.667% respectively, here the proposed IACOSAF classifier is 9.6605%, 4.0273%, and 16.7149% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively (refer table 4.4).

Table 1.5: Comparison of F-Measure Values in Proposed IACOSAF Classifier Vs. Existing Classifiers

Datasets	**F-Measure (%)**			
	Fuzzy	**ACOF**	**IACOF**	**IACOSAF**
CTG	57.1429	75.2475	81.5534	81.553
Abalone	60.1504	74.0214	78.3217	83.33
Iris	50.00	73.07	76.92	86.2745
CHD	74.3590	87.3418	90.9765	95.2612

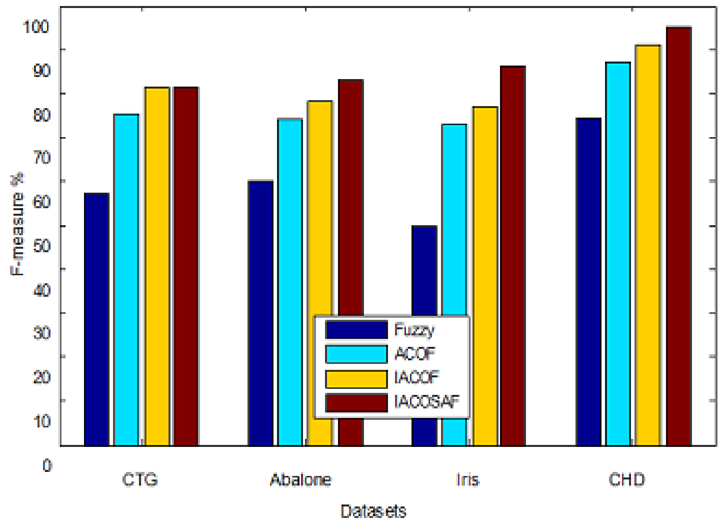

Figure 1.7: F-Measure Comparison of Proposed IACOSAF Classifier Vs. Existing Fuzzy Classifiers

Figure 1.7 explains the F-measure outcomes of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers. The F-measure results of IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 81.553%, 81.553%, 75.248%, and 57.413% respectively for CTG samples, it concludes that the proposed system is 0%, 6.3055% and 24.4101% higher when compared with IACOF, ACOF and fuzzy classifiers respectively.

For Abalone dataset, the F-measure results of the IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 83.33%, 78.322%, 74.021%, and 60.15% respectively, here the proposed IACOSAF classifier is 5.0083%, 9.3086% and 23.1796% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

For Iris dataset, F-measure results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 86.275%, 76.92%, 73.07%, and 50% respectively, here the proposed IACOSAF classifier is 9.3545%,13.2045% and 36.2745% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively.

For CHD dataset, F-measure results of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 95.261%, 90.977%, 87.342%, and 74.359% respectively, here the proposed IACOSAF classifier is 4.2847%, 7.9194%, and 20.9022% higher when compared with IACOF, ACOF and Fuzzy classifiers respectively (Refer table 1.5).

Table 1.6: Comparison of Accuracy Values in Proposed IACOSAF Classifier Vs. Existing Classifiers

Datasets	**Accuracy(%)**			
	Fuzzy	**ACOF**	**IACOF**	**IACOSAF**
CTG	50	72.2222	78.889	84.44
Abalone	57.6000	70.8000	76.71	80.80
Iris	52.00	72.00	76.00	86
CHD	68.2540	84.1270	88.412	93.8095

Figure 1.8 explains the accuracy outcomes of the proposed IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers.

ISSN 1943-023X

Received: 23 July 2018/Accepted: 21 Aug 2018

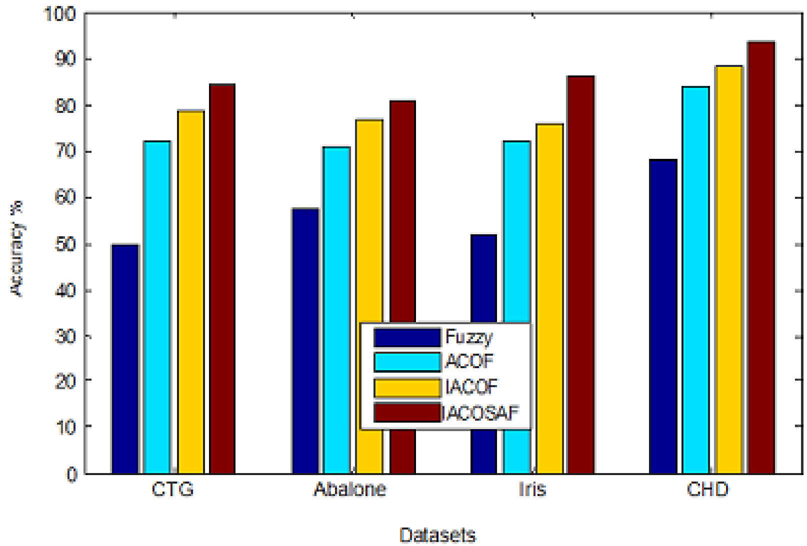

Figure 1.8: Accuracy Comparison of Proposed IACOSAF Vs. Existing Fuzzy Classifiers

The accuracy results of IACOSAF classifier, existing IACOF, ACOF and fuzzy classifiers are 84.44%, 78.889%, 72.22%, and 50.00% respectively for CTG samples, it concludes that the proposed IACOSAF system is 5.551%, 12.2178% and 34.44% higher when compared with IACOF, ACOF and fuzzy classifiers respectively.

Similarly for Abalone Dataset, Iris and CHD dataset proposed IACOSAF classifier is higher when compared with IACOF, ACOF and Fuzzy classifiers respectively (Refer table 1.6).

VII. Conclusion and Future Work

An improved fuzzy prediction system with three phases-fuzzy rule generation, Rule Optimization and Classification have been implemented with Rules generated based on the dataset then these rules are optimized by using IACOSA. Then the optimized rules are given in to the fuzzy so named as IACOSAF prediction system. The Adaptive Mutation Operation (AMO) to Ant Colony Optimization (ACO) algorithm had enhanced the algorithm performance, expanded the diversity of population, and inhibit the premature convergence.

The performance measures of sensitivity, specificity, accuracy, Recall, precision, and f-measures, were evaluated for the proposed method. The efficiency of the classification is very high by presenting very good accuracy outcomes and also the prediction of heart disease gives very accurate outcomes. From the outcomes, hybrid IACOSAF classifier had reduced the consumed number of iterations for convergence that outperforms the other classifiers by providing very good accuracy.

In heart disease identification, every individual has particular values for Blood Pressure (BP), cholesterol and Pulse Rate (PR). Forecasting the risk level of every person dependent upon gender, age, cholesterol, Blood pressure, pulse rate turns out to be a main task. Erstwhile in case the dataset samples turn out to be vast, forecasting the risk levels can become more problematic.

In order to resolve these issues feature selection (FS) algorithms or attribute reduction algorithms are concentrated in future. The FS algorithm chooses most significant features from the dataset samples so as to decrease the calculation time, and raises prediction accurateness. Hence this is reserved as a possible scope for the forthcoming work for present research.

Acknowledgement

The author gratefully acknowledge the support and facilities provided by Dept. of CSE,VVIT and also extend their thanks to the Head of the Dept. of CSE Prof. Dr. M. Nithya Research Supervisor, VMKVEC-Salem, for the immense support and guidance during the course of work.

Received: 23 July 2018/Accepted: 21 Aug 2018

References

[1] Herrera, F. and Lozano, M. Adaptation of genetic algorithm parameters based on fuzzy logic controllers. *Genetic Algorithms and Soft Computing* **8** (1996) 95-125.

[2] Dorigo, M., Maniezzo, V. and Colorni, A. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **26** (1) (1996) 29-41.

[3] Yang, X.S. *Nature-inspired metaheuristic algorithms*. Luniver press, 2004.

[4] Chiang, Y.M, Chiang, H.M. and Lin, S.Y. The application of ant colony optimization for gene selection in microarray-based cancer classification. *International Conference on Machine Learning and Cybernetics,* 2008, 4001-4006.

[5] Olson, D.L. and Delen, D. Advanced data mining techniques. *Springer Science & Business Media*, 2008.

[6] Parthiban, L. and Subramanian, R. Intelligent heart disease prediction system using CANFIS and genetic algorithm. *International Journal of Biological, Biomedical and Medical Sciences* **3** (3) (2008) 1-4.

[7] Zhou, Y. Runtime analysis of an ant colony optimization algorithm for TSP instances. *IEEE Transactions on Evolutionary Computation* **13** (5) (2009) 1083-1092.

[8] Yu, H., Gu, G., Liu, H., Shen, J. and Zhao, J. A modified ant colony optimization algorithm for tumor marker gene selection. *Genomics, proteomics & bioinformatics* **7** (4) (2009) 200-208.

[9] Data, G.O. *Data Mining: Concepts and Techniques*. San Francisco, 2010.

[10] Han, J, Pei, J. and Kamber, M. *Data mining: concepts and techniques*. Elsevier, 2011.

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

[12] Kabir, M.M, Shahjahan, M. and Murase, K. A new hybrid ant colony optimization algorithm for feature selection. *Expert Systems with Applications* **39** (3) (2012) 3747-3763.

[13] Anooj, P.K. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University-Computer and Information Sciences* **24** (1) (2012) 27-40.

[14] Salama, K.M. and Freitas, A.A. Investigating the impact of various classification quality measures in the predictive accuracy of ABC-Miner. *IEEE Congress on Evolutionary Computation,* 2013, 2321-2328.

[15] Aarts, E, Korst, J. and Michiels, W. Simulated annealing. *Search methodologies*, 2014, 265-285.

[16] Anusuya, V. and Kavitha, R. A New Hybrid Ant based genetic algorithm–fuzzy shortest path problem. *International Journal of Applied Mathematics & Statistical Sciences (IJAMSS)* **4** (2) (2015) 7-14.

[17] Dewan, A. and Sharma, M. Prediction of heart disease using a hybrid technique in data mining classification. *2nd International Conference on Computing for Sustainable Global Development,* 2015, 704-706.

[18] Korejo, I. and Yang, S. Comparative Study of Adaptive Mutation Operators for Genetic Algorithms. *Journal of Emerging Technology and Advanced Engineering* **2** (2) (2012) 1-10.

[19] Kavitha, S. and Nithya, M. An Efficient Ant Colony Optimization and Data Mining Technique (ACODMT) based on Fuzzy for Heart Disease Prediction. *International Journal of Printing, Packaging & Allied Sciences* **4** (6) (2016).

[20] Kavitha, S. and Nithya, M. An Efficient Cardiovascular Disease Prediction Through Pheromone Based ACO with Hybrid Fuzzy Logic. *International Journal of Scientific Research* (2016).

ISSN 1943-023X

Received: 23 July 2018/Accepted: 21 Aug 2018