ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 6, No.18 , November 2015
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

An Intelligent System based on Fuzzy Inference System to prophesy the brutality of Cardio Vascular Disease

Sivagowry S[1], Durairaj M[2]

[1] Department of Computer Science, Engineering and Technology, Bharathidasan University, Tiruchirapalli, Tamilnadu 620024, India
sivagowry87@gmail.com

[2] Department of Computer Science, Engineering and Technology, Bharathidasan University, Tiruchirapalli, Tamilnadu 620024, India
durairaj.bdu@gmail.com

Abstract

To unravel hidden relationships and diagnose diseases efficiently, Data Mining along with Soft Computing Techniques are used in several researches. Cardio Vascular Disease is a condition which leads to severe disability and death. Since the diagnosis involves vague symptoms and tedious procedures, diagnosis is usually time-consuming and erroneous. For the healthier analysis and treatment of heart disease based on brutality, an Intellectual, accurate and proficient investigative system is needed. For diagnosing heart disease with improved effectiveness, an Intelligent Fuzzy Inference System is needed. This paper illustrates how Fuzzy Inference System is used to envisage the severity of disease by constructing an effective Fuzzy Rule Base. It is also proved that a precision of 95.23% is obtained when Fuzzy System is used in severity prediction

Keywords: *Fuzzy Inference System, Fuzzy Rule Base, Severity prediction, Heart disease, Mamdani fuzzy system***.**

1. Introduction

The role of Data Mining in health care data is massive. The human decision making is optimal, but it is poor when the amount of data to be classified is huge. The enormous stress and overwork load resulted in poor / inaccurate decision making which may lead to disastrous consequences and cannot be allowed in medical field. The most exorbitant and harmful mistake is performing decision making process based on improper information acquired from medical data [1]. Institute of Medicine estimated that the effect of medical error accounts for about $17 to $29 billion, which is not declined since then. Medical history data, which comprise of number of tests and previous examination cycles, is essential to diagnose and devise future course of treatments on particular disease. It is conceivable to increase the benefit of Data mining [2], [3], [4] in health care by employing it as an intellectual symptomatic tool [5]. The researchers in the medical field have prospered in categorizing and prophesying the syndrome with the encouragement of Data mining techniques [6]. Association rules of Data Mining have been significantly used in health care data prediction [7], [8], [9, and 10]. The eventual goal of knowledge discovery is to identify factors which tend to improve the quality and effectiveness of health care system.

2. Heart Disease

The rise of health care cost is one of the universally confronting problems. The therapeutic term for Heart Disease / Heart Attack is Myocardial Infarction (MI) or Acute Myocardial Infarction (AMI). Heart attack emerges when there is indiscretion in the flow of blood and bruised heart muscles due to inadequate oxygen supply [11]. Jeopardy factors for Myocardial Infarction include smoking, high blood pressure, cholesterol, Diabetes, Family history, etc.,. Cardio Vascular Disease (CVD) clinical guidelines spotlight on the management of single risk factors [12, 13, and 14]. In majority of cases, it doesn't work since risk factors crop up in clusters, that is, the presence of single risk factor which indicates the presence of other risk factors too. It is apparent that the presence of multiple risk factors increases the sternness of CVD. World Health Organization (WHO) in the year 2008 testified that 30% of total global bereavements are due to Cardio Vascular Disease (CVD). WHO has announced India as global CVD capital [7]. Rapid urbanization and industrialization have led to the major cause of demise in India due to CVD [8]. It is estimated that numeral of CVD patients will increase from 300 million to 600 million by 2020 [15, 16]. By 2030, almost 25 million people will die from CVDs, mainly from heart disease and stroke [17], [18], [19]. These are projected to remain the CVD is the single leading cause of death. The CVD is also expected to be the leading cause of deaths in developing countries due to changes in lifestyle, work culture and food habits. Hence, more careful and efficient methods of diagnosing cardiac diseases and periodic examination are of high importance [11] [10] [20] to take preventive care.

The paper is organized as follows: Section III gives a magnitude about the existing literature work regarding

Fuzzy logic in Heart disease prediction. Section IV describes the data set collected for experimentation. Section V describes how the Fuzzy Inference System evaluates the severity of heart disease based on Fuzzy Rule base. Section VI discusses the results. Section VII concludes the paper.

3. Review of literature based on Fuzzy Logic in Heart Disease Analysis

A new Particle Swarm Optimization based Fuzzy expert system was proposed which involves four stages [21]. Nearest hot deck imputation is used to remove the missing data. A Fuzzy Expert System is generated based on set of rules. So the system can used to provide interpretations for decisions. The Fuzzy Expert System which can be developed by using the set of rules and tuned Member Functions improves the accuracy. Mamdani fuzzy inference system is used for Fuzzification of crisp sets. The Centre of Gravity (COG) method is employed for defuzzification process. 93.27% of accuracy was obtained by using the Fuzzy Expert System.

Existing system was studied and a disease prediction system which is based on fuzzy is proposed. The accuracy provided by this system is more than 90% [22]. A neuro fuzzy network was designed to identify and classify the coronary artery disease by using MATLAB [23]. Sugeno-based fuzzy expert system is used. The sensitivity and specificity obtained are 1 and 0.88 respectively. This shows that the network have an acceptable degree of accuracy.

A data mining technique which is based on rough set theory and fuzzy logic is proposed. It has two phases which includes clustering and classification. Clustering is based on rough set theory and fuzzy are used for classifying the clusters obtained by using the rough set theory. Complexity in generating the rules based on Fuzzy logic is reduced since rough set theory is used prior to fuzzy logic. MATLAB is used for implementation [24]. Sensitivity, specificity and accuracy are the parameters which were used to measure the performance of the proposed system. From the result, the data set of Switzerland database provides better results than other. The proposed method can also used to deal with uncertainty problem.

Advanced fuzzy resolution mechanism [25] was conducted for diagnosing the heart disease. MATLAB Fuzzy Logic Tool box is used to generate the rules. It has five layers, and each layer has its own node. Accuracy is used as a metric to compare the working of proposed algorithm with existing methods. Accuracy obtained is 94.11% which is higher when compared with other methods.

Two systems were developed to diagnosis heart disease based on MATLAB [26]. Multi Layer Perceptron (MLP) network is the base for developing the first system. Second system was developed based on Adaptive Neuro Fuzzy Inference System (ANFIS). 80% of data was used for training and 20% is used for testing. The accuracy obtained while training by using ANFIS was 100% whereas by using MLP it was 90.74%. But, while testing, MLP outperforms ANFIS in accuracy

4. Data set description

The Data set used for experimentation is taken from Data mining repository of the University of California, Irvine (UCI). Data of Cleveland Data set, Hungary Data set, Switzerland Data set, Long beach and Statlog Data set are collected. Cleveland, Hungary, Switzerland and Va Long Beach data set contain 76 attributes. Among all the 76 attributes, 14 attributes are taken for experimentation, since all the published experiments refer to using subset of 14 attributes. Researchers in the medical domain mostly use Cleveland data set and Statlog data set for testing purpose. This is because all the other data set has more number of missing values than Cleveland data set [27]. The Table 1 describes the data attributes used in this work.

Among the 13 attributes, only 5 attributes are chosen by using the knowledge of Intelligent Hybrid Quick Reduct Particle Swarm Optimization Algorithm [28] which does not affect the accuracy in classification and prediction. The chosen five attributes are cp, exang, slope, ca and thal

5. Fuzzy Logic in Severity Prediction

The following is the algorithm which is used to construct the rule base Fuzzy Inference System for severity prediction.

Fuzzy Logic Algorithmic steps:
Input:
cp, exang, slope, ca, thal
Output:
Heart Disease Severity- Low, Mild, Severe
Begin
Input crisp value
Set Mamdani fuzzy model, with fuzzy if-then rules.
Assign fuzzy member for the input of each variables.
Output variable is predicted by using the input.
Generate rules

Calculate Membership function by using Guassian Membership function given in equation (1)

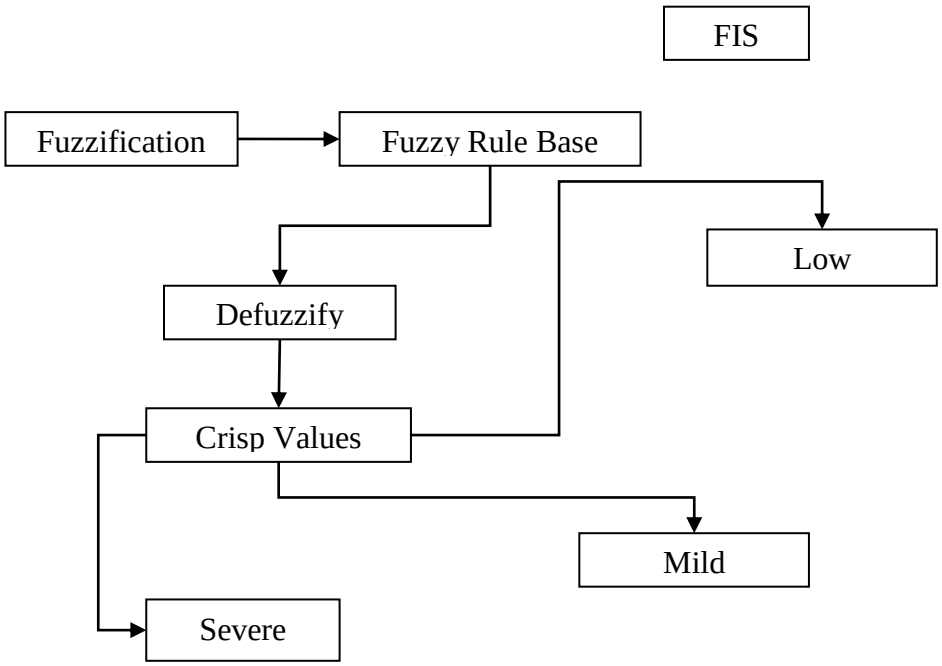

Figure 1 : Fuzzy Inference System for Severity Prediction

$$Guassian\ Membership\ Function = \mu(x,a,b) = e^{\frac{-(x-b)^2}{2a^2}} \quad .....(1)$$

Fuzzy rule base Construction.
Defuzzification of the fuzzy value to get the final crisp output. The Centroid method is used for defuzzification.

$$Centroid = ZCOA = \frac{\int_Z \gamma_A(z) z dz}{\int_z \gamma_A(z) dz} \quad .....(2)$$

End

The Figure 1 describes how the Fuzzy Inference System is designed for predicting the severity of Cardio Vascular Disease.

The five inputs cp, slope, ca, exang, thal and the output severity is converted into fuzzy variables as follows:

The input chest pain (cp) has four types namely typical, atypical, non-angina and asymptomatic. It is assigned the fuzzy values as 1,2.3 and 4. The figure 2 shows the membership function plot of the input variable cp

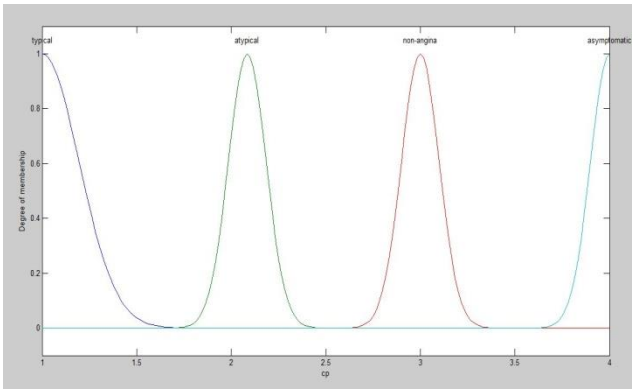

Figure 2: Membership function plot for the input variable cp

The next input variable exercise induced angina (exang) has two fields yes and no. It has 1 for yes and 0 for no. The field slope has three fields as shown in Table 2. The Table 2 shows the fuzzy value assigned for the crisp input for the input field slope

Table 1: Description of the data attributes

No	*Name*	*Description*
1	Age	Age in Years
2	Sex	1=male, 0=female
3	cp	Chest pain type(1 = typical angina, 2 =atypical angina, 3 = non-anginal pain, 4 = asymptomatic)
4	Trestbps	Resting blood sugar(in mm Hg on admission to hospital)
5	chol	Serum cholesterol in mg/dl
6	fbs	Fasting blood sugar>120 mg/dl(1= true, 0=false)
7	Restecg	Resting electrocardiographic results(0 = normal, 1 = having ST-T wave abnormality, 2 = left ventricular hypertrophy)
8	Thalach	Maximum heart rate
9	Exang	Exercise induced angina
10	Oldpeak	ST depression induced by exercise relative to rest
11	Slope	Slope of the peak exercise ST segment (1=upsloping, 2=flat, 3= downsloping)
12	ca	Number of major vessels colored by fluoroscopy
13	thal	3= normal, 6=fixed defect, 7= reversible defect
14	Num	Class(0=healthy, 1=have heart disease)

Table 2: Fuzzy value for the slope

Upsloping	1
flat	2
downsloping	3

The number of major vessels colored by fluoroscopy is indicated by the input variable ca The Gaussian membership function is used to plot the membership value for all the fields. The final input variable is thal which is the size and location of injured muscle after the heart attack. The Fuzzy value for each of the value is given in Table 3. The value 3,6 and 7 indicates normal, fixed defect and reversible defect respectively.

Table 3: Fuzzy value for thal

normal	3
fixed defect	6
reversible defect	7

The figure 3 shows the Fuzzy Inference System designed for the extraction of fuzzy rule in predicting the severity of Heart Disease. The Figure 9 shows the description of the FIS designed using Fuzzy logic for severity prediction. After designing the FIS , the rule base is constructed by using the if-then rules. Number of rules framed for the system is 63. The Mamdani system is used for designing the FIS. Because the system has widespread application than the Sugeno model and it is more suitable system for the human input. So, the Mamdani system is used for designing the FIS for predicting the severity of the Heart disease.

Name = **Pref**
Type = **Mamdani**
NumInputs = **5**
InLabels = **cp**
exang
slope
ca
thal
NumOutputs = **1**
OutLabels = **Severity**
NumRules = **63**
AndMethod = **min**
OrMethod = **max**
ImpMethod = **min**
AggMethod = **max**
DefuzzMethod = **centroid**

Figure 4: Description of FIS for prediction

Fuzzy rule base is constructed using the reduced attributes. Rule viewer shows the strength of each rule and the surface viewer plot the output with respect to every input. The figure 5 shows the rule viewer of the prediction system designed.

6. Results and Discussion

The defuzzified value obtained by using centroid method is compared with original output and manual output calculation. The Table 4 shows the comparison of each value obtained. In each case, it was observed that there is no much difference in the output observed which supports the use of Fuzzy Logic in severity diagnosis. The accuracy in severity prediction has been increased. It was up to 95.86% when using Fuzzy LogicWhile using the Fuzzy logic the whole prediction level has been increased. The figure 6 shows the increase in accuracy when Fuzzy Logic is employed. The whole FIS performs well in predicting the severity level and also there is a tremendous amount of increase in accuracy after using the FIS.

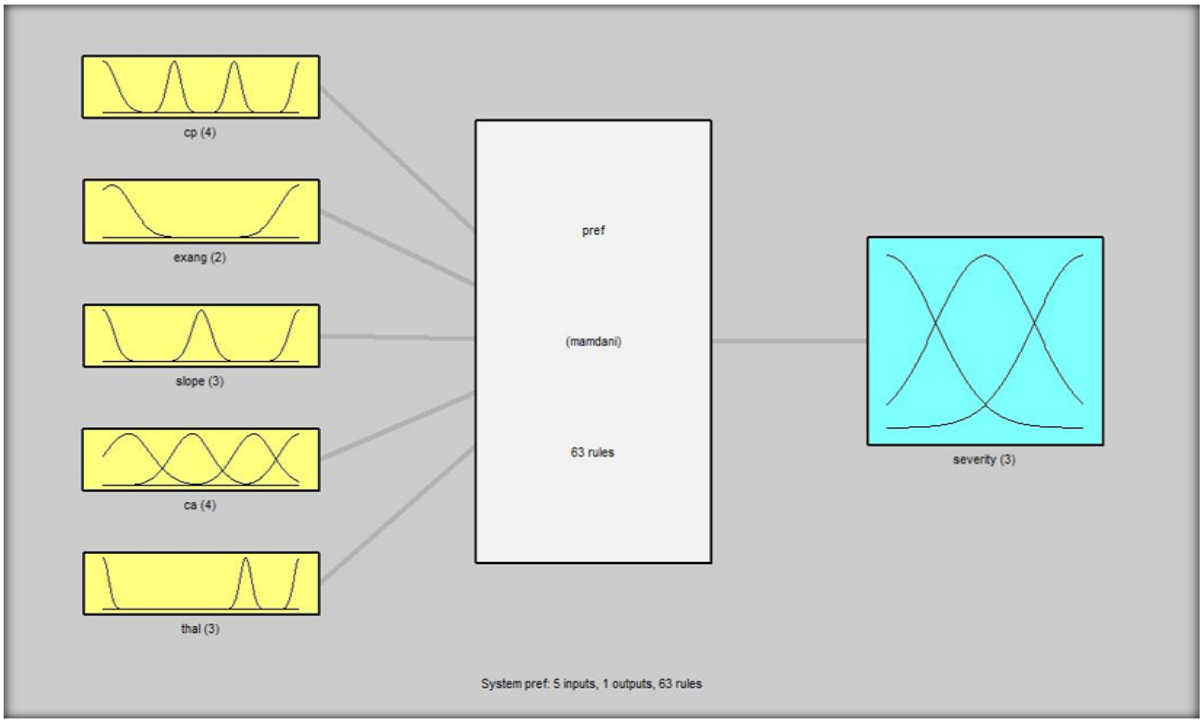

Figure 3: FIS designed for severity prediction

Figure 5: Rule viewer of the FIS in severity prediction

Table 4: Comparing of the output value

Original output	Manual calculation	System output
1	1.32	1.4
1	1.54	2
1	1.4	1.4
1	1.38	1.4
1	1.25	1.4
1	1.13	1.4
1	1.4	1.4
1	1.0	1.4
1	1.0	1.4
1	1.2	2
2	2.11	2
2	2.13	2
2	2.42	2
2	2.0	2
2	2.0	2
2	2.42	1.9
2	2.0	2.1
2	2.31	2
2	2.01	2
2	2.0	2
3	3.11	2.5
3	3.18	2.5
3	2.29	2.6
3	2.68	2.6
3	3.11	2.78
3	3.22	2.6
3	2.55	2.6
3	3.0	2.98
3	3.0	2.5
3	3.0	2.6

The Table 4 compares the original output with the result obtained from the system and mathematical calculation. It is observed that there is no much variation in the outputs when compared which justifies the use of the proposed FIS with Rule base for predicting the severity of the disease.

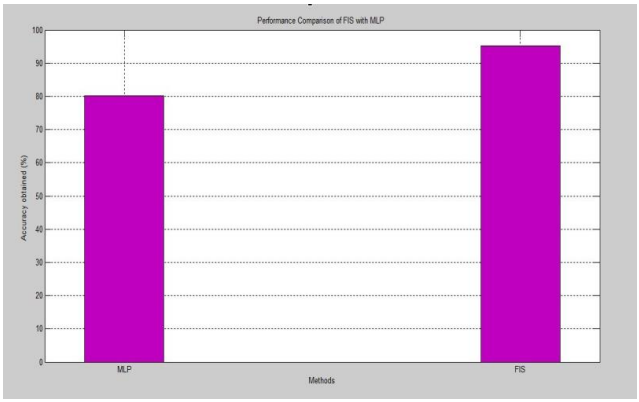

Figure 6 : Performance of FIS over MLP

The Figure 6 compares the performance of MLP over the proposed FIS with 63 rules framed. It is evident that the prediction accuracy has been increased from 80 to 90% which supports the use of FIS in predicting disease severity. The Figure 7 plot the training error with regard to number of epoch. It is observed that the training error has been reduced at the end of 50th epoch.

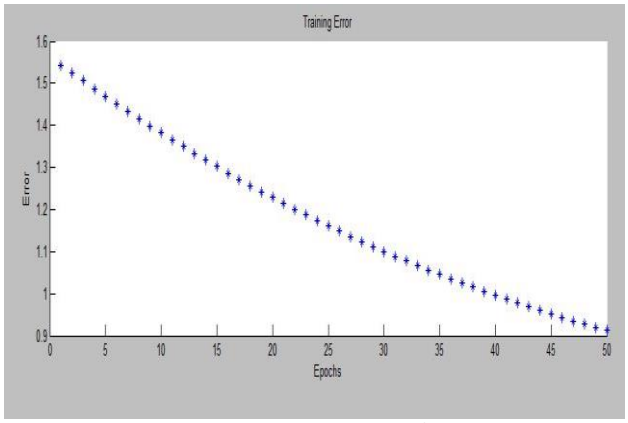

Figure 7: Training error Vs no. of epochs

7. Conclusion

The most important and difficult task in medicine is Medical diagnosis. The problem here is detecting a disease from several factors or symptoms, since it may lead to false assumptions with unpredictable results. The results obtained indicate that the proposed approach can be used to induce fuzzy rules from data by providing good balance between accuracy and readsability. Primary prevention is recommended for promoting healthy life style and habits through increased awareness and consciousness, to prevent development of any risk factors and a system to predict possibility of heart disease risk for prevention. The FIS designed from optimal data set has produced an accuracy of 95.23%. Hence it can be

suggested as a valid tool for medical practitioners in their preliminary stage.

References

[1] www.nytimes.com/2009/07/29/opinim/29hall.html.
[2] K.Rajeswari, "Prediction of Risk Score for Heart Disease in India using Machine Intelligence", *IPCSIT*, Vol 4, 2011
[3] Latha Parthiban and R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", *International Journal of Biological and Life Sciences*, Vol 3(3), pp157-160,2007.
[4] Liangxiao. J, Harry.Z, Zhihua.C and Jiang.S "One Dependency Augmented Naïve Bayes", *ADMA*, pp 186-194, 2005
[5] Huan Liu and Hiroshi Motoda, Rudy Setiono and Zheng Zhao. "Feature Selection: An Everlasting Frontier in Data Mining", *JMLR: The 4th Workshop on Feature Selection and Data Mining*, 2010.
[6] Rafiah Awang and Palaniappan. S "Web based Heart Disease Decision Support System using Data Mining Classification Modeling techniques" , *Proceedings of iiWAS*, pp 177-187, 2007
[7] Carlos Ordonez, Edward Omincenski and Levien de Braal "Mining Constraint Association Rules to Predict Heart Disease", *Proceeding of 2001, IEEE International Conference of Data Mining, IEEE Computer Society*, ISBN-0-7695-1119-8, 2001, pp: 433-440
[8] Deepika. N, "Association Rule for Classification of Heart Attack patients", *IJAEST*, Vol 11(2), pp 253-257, 2011.
[9] Durairaj.M, and Meena.K" A Hybrid Prediction System using Rough Sets and Artificial Neural Network", *International Journal of Innovative Technology and Creative Engineering*, Vol 1(7), July 2011.
[10] Durairaj.M, Sivagowry.S. "A Survey on Particle Swarm Optimization and Rough Set Theory in Feature Selection for Heart Disease Prediction." *International Journal of Computer Science and Mobile Computing (IJCSMC)* Vol 4 (3) (2015): 87-92.
[11] Sudha.A, Gayathri.p and Jaishankar. N "Utilization of Data Mining Approaches for prediction of life Threatening Disease Survivability", *IJAC* (0975-8887), Vol 14(17), March 2012.
[12] Rafiah Awang and Palaniappan. S "Intelligent Heart Disease Prediction System Using Data Mining techniques", *IJCSNS*, Vol 8(8), pp 343-350, Aug 2008
[13] Sivagowry.S, Dr. Durairaj. M. (2014). PSO-An Intellectual Technique for Feature Reduction on Heart Malady Anticipation Data. *International journal of Advanced Research in Computer Science and Software Engineering*, *4*(9), 610-621.
[14] Sivagowry.S, Dr. Durairaj.M. "A Pragmatic Approach of Preprocessing the Data Set for Heart Disease Prediction." *International Journal of Innovative Research in Computer and Communication Engineering* 2.11 (2014): 6457-6465.
[15] World Health Organization. Strategic priorities of the WHO Cardiovascular Disease programme. Available online at URL: http://www.who.int/whr/200. Last accessed February 2006.
[16] Chen A.H., "HDPS: Heart Disease Prediction System", *Computing in Cardiology*, ISSN 0276-6574, pp 557-560, 2011.
[17] en.wikipedia.org/wiki/myocardial_infarction
[18] Nidhi Bhatia and Kiran Jyothi, "A Novel Approach for heart disease diagnosis using Data Mining and Fuzzy logic", *IJCA*, Vol 54(17), pp 16-21, September 2012.
[19] Sivagowry, S., M. Durairaj, and A. Persia. "An empirical study on applying data mining techniques for the analysis and prediction of heart disease." *Information Communication and Embedded Systems (ICICES), 2013 International Conference on*. IEEE, 2013.
[20] Durairaj.M, Sivagowry.S. "Feature Diminution by Using Particle Swarm Optimization for Envisaging the Heart Syndrome." *International Journal of Information Technology and Computer Science(IJITCS)* (7).2 (2015): 35-43.
[21] Muthukarupan S and Erib M.J, " A hybrid PSO based fuzzy expert system for the diagnosis of Coronary heart system", *Expert systems with Application*, Vol 39, pp 11657-11665, Elsevier, 2012
[22] Neera Pathania and Rithika, "FIS Rule based Heart Disease Prediction System to predict the risk level of Heart Disease", *IJCEM*, Vol 7, October 2014.
[23] Saeed Ayat, Asieh Khosravanian, "Identification and classification of Coronary Artery disease patients using Neuro-Fuzzy Inference system", *Journal of Mathematics and Computer Science*, Vol 13, pp 16-141, 2014.
[24] Saravana kumar, Tholkapia arasu, "Rough Set Theory and Fuzzy Logic based warehousing of Heterogeneous clinical database".
[25] AVS Kumar, " Generating rules for Advanced Fuzzy Resolution mechanism to diagnosis Heart Disease" , *International Journal of Computer Application*, Vol 77(11), pp 6-12, September 2012.
[26] Mohd. A.M. Abushariah, Assal A.M Alquadah, Omar Y.Adwan, Rana M.M. Yousef, " Automatic Heart Disease Diagnosis system based on Artificial Neural Network and ANFIS approach", *Journal of Software Engineering and Application*, pp 1055-1064, 2014
[27] UCI machine learning repository: http://archive.ics.uci.edu/ml/datasets/Heart+Disease: Last visited 8th August, 2014.
[28] Durairaj.M, Sivagowry.S, " An Intelligent Hybrid Quick Reduct Particle Swarm Optimization Algorithm for feature Reduction in Cardiac Disease Prediction", *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)*, Volume 12(2), March-May 2015**,** pp.163-173