ISSN : 2231-0711

Anu Ragavi.V et al | International Journal of Computer Science Engineering and Technology(IJCSET) | Dec 2014 | Vol 4, Issue 12,388-391

Fuzzy Inferenceing for Coronary Artery Disease Screening

Anu Ragavi.V [1] And Devi Priya.R[2]

[1,2] *Department of Information Technology,Kongu Engineering College,Perundura,India.*

Abstract: **An expert system is a computer software that attempts to mimic the reasoning of a human specialist. Expert systems typically consist of three parts: a knowledge base , an inference engine and interface that allows the user to express the problem in a human language. The traditional medical expert system is simple and widely used system but manual screening of Coronary Artery Disease (CAD) patients is not an easy task because the number of medical experts in comparison with number of patients is less. No clinical risk factors based expert system has been reported to screen the patients at an early stage. Also, the severity of the disease is not predicted from the system. Another con behind medical expert system is that higher level tests like ECG and angiogram are needed to identify the disease where poor people finds difficult to pay and undergo the tests. In the proposed approach, fuzzy expert system is used in place of medical expert system. Fuzzy inferenceing involves Mamdani approach. This approach involves cost effective screening which also identifies the severity of CAD. No higher level tests like ECG and angiogram are needed to identify the disease. The system is useful when the expert strength is less. The proposed fuzzy inferenceing for Coronary Artery Disease screening using clinical parameters identifies the risk percentage efficiently.**

Keywords: **Expert System, Coronary Artery Disease (CAD), Risk Factors, Fuzzy Inferenceing.**

I. INTRODUCTION

Coronary artery disease (CAD) also known as atherosclerotic heart disease is the most common type of heart disease caused by building up of plaque along the inner walls of the arteries of the heart. The inside of the arteries develop plaques of different sizes. The plaque deposits are soft on the inside with a hard fibrous “cap” covering the outside surface. If the hard surface cracks or tears, the soft fatty substance inside gets exposed. This narrows the lumen of arteries and reduces the oxygen-rich blood flow to the heart eventually leading to heart attack. Coronary Artery Disease turns out to be the leading cause of death for both men and women and it accounts for nearly 600,000 deaths every year. The most common symptom of coronary artery disease is **angina** (also called angina pectoris). Angina is otherwise referred to as chest pain. It is also depicted as chest discomfort, heaviness, pressure, aching, burning, numbness, fullness, or squeezing. Angina is felt in the chest but also in the left shoulder, arms, neck. Other symptoms for coronary artery disease includes shortness of breath, irregular heartbeats, skipped beats or a “flip-flop” feeling in chest, faster heartbeat, dizziness, nausea, extreme weakness, sweating. CAD can lead to other serious problems such as heart attack, stroke or death.

Papaloukas et al [4] proposed a rule-based expert system that rapidly and reliably detects ECG changes suggestive of ischemia in long duration ECGs. The performance of ECG is not affected by the noise and thus clarifies the type of each detected episode. This new system will increase the diagnostic and prognostic accuracy but the ECG tests seem costly for the poor people.

Gamberger et al [6] proposed an approach for active mining of patient records aimed at discovering patient groups at high risk for Coronary Heart Disease (CHD). The approach proposes active expert involvement in the knowledge discovery process: data gathering, cleaning and transformation, subgroup discovery, statistical characterization of induced subgroups with their interpretation, and the evaluation of results.

Chi-Ming et al [3] proposed an expert system by using data mining technique called Bayesian model for detecting Coronary Artery Disease (CAD). In addition, this study provides an evaluation of CAD detection before a cardiac angiography but the severity of CAD cannot be calculated.

Hannan et al [10] proposed a system where the collected 300 patient’s information is coded and normalized into 13 different excel sub-sheets. All patients’ data are trained by using Support Vector Machine (SVM) and Feed Forwards Back Propagation (FFBP). The analysis model by using SVM and FFBP of Artificial Neural Network (ANN) gives less appropriate result for medical prescription for heart disease patient. However, techniques that can improve the speed and performance should be used.

Ephzibah and Sundarapandian [5] proposed a system with the help of evolutionary techniques like genetic algorithms and neural networks. This expert system will help the doctors to arrive at a conclusion about the presence or absence of heart disease in patients. This is an enhanced system that precisely classifies the presence of the heart disease. Results show that the severity of the disease cannot be obtained correctly.

From the literature, it is clear that the CAD risk factors may not be always known in advance. The existing medical expert systems classify presence or absence of the disease whereas the severity of the disease (CAD) is not predicted. The system is not efficient with more population is more with less number of experts. Also higher level tests like ECG and angiogram are needed to identify the disease.

II. Proposed Scheme

In the existing system, such as MYCIN, INTERNIST which are multiple disease expert systems that studies on 18-lead electro cardiogram (ECG) analysis , 12-lead electro cardiogram (ECG) analysis, wavelet transformed ECG analysis, Exercise stress testing analysis, Myocardial Infarction analysis, heart valve disease analysis have been done extensively for detection of cardiovascular diseases . Thus the problems that were faced in the existing systems are, need for higher level tests like ECG and Angiogram, severity could not be predicted and manual screening is not possible. In proposed system, fuzzy inferenceing by using easily available clinical parameters is used to identify CAD risk without using the higher level tests like ECG and Angiogram, to find the severity of CAD, to have an early detection by using clinical parameters.

Fuzzy logic is a multi-valued logic. Fuzzy logic includes 0 and 1 as extreme cases of truth but also includes the various states of truth in between 0 and 1. This work includes inferenceing by Mamdani Approach.

A. CAD Risk Factors

The risk factors identified are mainly based on Indian population. They are age, smoking habit, Body Mass Index(BMI), Systolic Blood Pressure(SBP), Diastolic Blood pressure(DBP), Blood Sugar(BS), Total Cholesterol(TC), Low Density Lipoprotein (LDL), Tri-GLyceride (TGL), High Density Lipoprotein (HDL)

1) *Age*: Age is a non-modifiable prime risk factor. As age grows, CAD risk increases significantly.

2) *Smoking*: Smokers are almost twice as likely to have a heart attack compared with people who have not addicted to smoking. By controlling and stopping smoking servers, risk associated with CAD can be reduced and benefits to large extend. Toxins present in the blood contribute to the development of plaque in the coronary arteries due to smoking thus resulting in obstruction of blood flow causing heart attack.

3) *Obesity*: The distribution of body fat is a determinant of cardiovascular risk. Major changes in lifestyle patterns have occurred leading to a trend towards decreasing physical activity resulting in increase in obesity. It is quantified by Body Mass Index (BMI) where it is defined as

$$\text{BMI} = \frac{\text{Weight}}{\text{Height}^2}$$

4) *Hypertention*: Hypertension puts strain on the heart, leading to heart disease and coronary artery disease. To evaluate risk, both Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) have been considered.

5) *Diabetes*: Diabetes is a long term condition in which there is too much glucose in blood because human body can't use insulin properly or can't make enough insulin. People with diabetes are also at risk for high blood pressure, unhealthy cholesterol levels, blood clotting problems etc..

6) *Lipid profile*: The lipid profile increases the risk of CAD. There are different constituents of blood cholesterol along with Total Cholesterol (TC) contributing to risk of CAD. High quantity of Low Density Lipoprotein (LDL) increases the risk of CAD whereas high quantity of TC level and ratio of TC to High Density Lipoprotein (HDL) cholesterol increase the risk of CAD. Lipid profile has been quantified by TC, Tri-Glyceride (TGL), LDL and HDL.

B. Sample Data Collection

Total of 500 patients' data were collected from Sudha hospital, Erode, India. The participating patients' minimum age was restricted to 30 years and the subjects with prior history of myocardial infarction (MI), arrhythmia and any type of coronary heart diseases were excluded from our investigation.

C. Fuzzy Inferenceing By Mamdani Approach

The basic structure of a fuzzy inference system consists of the following conceptual components

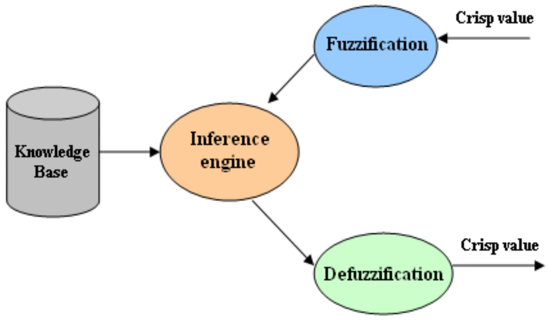

Fig. 1 Fuzzy Inference System

1) *Fuzzification:* In the traditional crisp set approach an element either fully belongs to a set or it does not belongs to the set. Fuzzification is the process of changing a crisp value into a fuzzy value. This is achieved with the different types of fuzzifiers (membership functions).The membership function is used to associate a grade to each linguistic term. In the fuzzy system, the CAD risk factors are considered as the input fuzzy variables. Each fuzzy variable has fuzzy values calculated from the membership function. Among the membership functions available, trapezoidal membership function is used because of its generality and capacity to contain more fuzzy information. The equation of trapezoidal membership function is,

$$\mu_A(x) = \begin{cases} 0, & (x < a)\ and\ (x > d) \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \frac{d-x}{d-c} & c \le x \le d \end{cases}$$

where $\mu_A(x)$ describes the membership function of variable *x* which is defined over the range of [a,d] with $a < b \le c < d$ The values of a, b, c and d for different values of a linguistic variable have been determined based on doctor's experience.

The interval [b, c] is called core of the trapezoidal function, where [a,d] is called support of the trapezoidal function and [a,b] and [c,d] are called boundaries of trapezoidal functions.

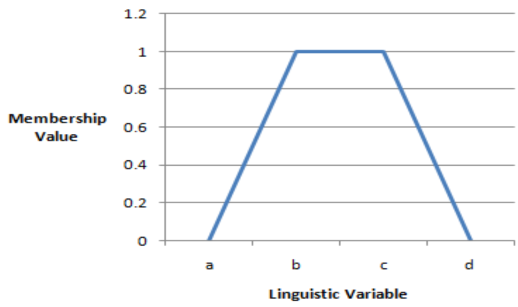

Fig. 2 Trapezoidal Membership Function

The output variable of fuzzy system is ''Risk'' which is defined as fuzzy variable and the fuzzy values are low, medium and high risk.

2) *Knowledge base*: A rule base/knowledge base, which contains a selection of fuzzy rules encoded in computer understandble form. A **fuzzy rule** is defined as a conditional statement in the form:

IF *x is A*
THEN *y is B*

where x and y are linguistic variables; A and B are linguistic values determined by fuzzy sets on the universe of discourse X and Y, respectively.

The knowledge base is dynamic in nature that is new rules may be added or invalid rules may be deleted. Also, the patients may not know the values for all his/her risk factors. To defeat the above problems, in the present expert system rules are grouped as modules. The grouping criteria of the rules is based on the antecedent part in the production rules. The present system has 10 modules since ten risk factors are considered. Module one has the rules containing only one risk factor and it can be any one of the available ten risk factors. Module two contains the rules that involve two risk factors. Similarly, module ten has the rules containing all the risk factors. In Table I some sample fuzzy rules have been shown.

Table I Sample fuzzy rules

RULE NO	FUZZY RULES
R1	If age is young and smoking is low smoker then risk=low
R2	If age is old and smoking is high smoker and SBP is highly increased then risk=high
R3	If age is medium and smoking is medium smoker and SBP is highly increased and DBP is medium increased and blood sugar is medium increased then risk=medium
R4	If age is young and smoking is medium smoker and SBP is slightly increased and DBP is medium increased and blood sugar is medium increased and LDL is medium increased then risk=medium
R5	If age is young and BMI is slightly increased and SBP is slightly increased and DBP is slightly increased and FBS is slightly increased and TC is slightly increased and LDL is slightly increased and TGL is slightly increased and HDL is low and chest pain significance is low then risk = low

3) *Inference engine*: It defines the membership functions used in the fuzzy rules and also incorporates the ability of reasoning for deciding which rules to be selected and fired for risk prediction. The fuzzy inference engine in the present system selects fuzzy rules and uses Mamdami inferencing to produce the fuzzy risk output. For each rule, Mamdami rule is applied as below.

To evaluate the disjunction of the rule antecedents, OR fuzzy operation is used.

$$\mu_{A\cup B}(x) = \max\,[\mu_A(x), \mu_B(x)]$$

Similarly, to evaluate the conjunction of the rule antecedents, AND fuzzy operation is used.

$$\mu_{A\cap B}(x) = \min\,[\mu_A(x), \mu_B(x)]$$

4) *Defuzzification*: **Defuzzification** is the process of converting fuzzy values in to crisp value, given fuzzy sets and corresponding membership degrees. The input for the defuzzification process is the aggregate output fuzzy set and the output is a single number. Centroid method is esteemed to defuzzify the fuzzy risk output to crisp risk percentage given by,

$$x^* = \frac{\int \mu_i(x)\, x\, dx}{\int \mu_i(x)\, dx}$$

where ∫ is the algebraic integration.

III. Performance Analysis

The output is expressed in three different fuzzy sets namely low, medium and high risk. The low risk patients are considered as normal patients and both medium and high risk patients are considered as abnormal (CAD) patients. The performance of fuzzy expert system is analyzed by taking three parameters into account. From the obtained results, it is inferred that the prediction of severity of CAD is improved by using fuzzy inferenceing. The following parameters are taken into account for comparison:

1. Specificity
2. Sensitivity
3. Accuracy

The Specificity is calculated by using

$$Specificity = \frac{TN}{TN+FP}$$

The Sensitivity is calculated by using

$$Sensitivity = \frac{TP}{TP+FN}$$

The Accuracy is calculated by using

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

- True positive (TP): It denotes the number of abnormal patients correctly classified by the model.
- True negative (TN): It denotes the number of normal patients correctly classified by the model.
- False positive (FP): It denotes the number of healthy patients wrongly classified as abnormal patients by the model.
- False negative (FN): It denotes the number of abnormal patients wrongly classified as normal patients by the model.

TABLE II
COMPARISON OF PERFORMANCE MEASURES

S.NO	PARAMETERS	NO.OF PATIENTS	PERFORMANCE EVALUATION	
			WITHOUT FUZZY (%)	WITH FUZZY (%)
1	Specificity	100	75.56	78.80
		300	78.8	80.05
		500	83.33	84.48
2	Sensitivity	100	85.00	88.88
		300	89.50	90.2
		500	95.85	95.31
3	Accuracy	100	72.66	73.70
		300	78.23	81.35
		500	84.20	90.08

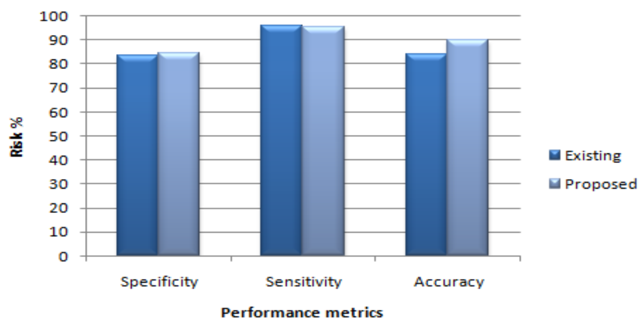

Fig.3 Comparison graph

IV. CONCLUSION

The fuzzy expert system uses easily available clinical parameters like age, smoking, height, weight, SBP, DBP, chest pain details and some low cost laboratory tests- TC, LDL, TGL and HDL. No higher level medical tests like ECG, angiogram, etc. are needed for predicting severity level of the patients by this methodology at least in its early stage. Moreover the system detects the degree of severity of subjects in terms of low (normal), medium and high risk prediction (abnormal). The system is easy to handle and can be applied as an adjunct tool to the cardiologists for automated mass screening of apparently healthy CAD patients in the developing countries where population is a major concern. Fuzzy logic is based on rule based approach and the rules have been developed by the knowledge engineers while interviewing a group of cardiologists. The rules are not developed by applying learning algorithms on the dataset so there is no data driven nature in the rule base. These rules have been expressed in the form of production rules (IF–THEN) and linguistic variables have been included for easy accessibility. The application of fuzzy logic approach has made the expert system capable of dealing with the ambiguity present in decision making in the medical domain. The present screening system is a single disease expert system. In future, the number of rules that has been framed can be optimized by using suitable optimization algorithm to find a best rule.

REFERENCES

[1] A.A.Ajlan (2007), '*Medical Expert Systems- HDP and PUFF*', International Journal of Engineering Trends and Technology, Vol. 04, No. 8, pp. 29-63.

[2] P.K.Anooj (2012), '*Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules*', International Journal of Computer and Information Sciences, Vol. 24, No. 6, pp. 27-40.

[3] C.M.Chu, W.C.Chien, C.H. Lai, Hans Bernd Bludau, Huei-Jane Tschai, Lu Pai, Shih-Ming Hsieh, Nian Fong Chu, Angus Klar, Reinhold Haux and Thomas Wetter (2009), '*A Bayesian Expert System for Clinical Detecting Coronary Artery Disease*', Journal of Medical Sciences, Vol. 29, No. 4, pp. 187-194.

[4] C.Papaloukas, I.Dimitrios Fotiadis, Aristidis Likas, Christos S. Stroumbis and K. Lampros Michalis (2002), '*Use of a Novel Rule-based Expert System in the Detection of Changes in the ST Segment and the T Wave in Long Duration ECGs*', Journal of Electrocardiology, Vol. 35, No. 1, pp. 160-197.

[5] E.P.Ephzibah and V. Sundarapandian (2012), '*A Neuro Fuzzy Expert System For Heart Disease Diagnosis*', An International Journal of Computer Science and Engineering, Vol. 2, No.1, pp. 160-187.

[6] D. Gamberger, N. Lavrac and G. Krstacic (2003) , '*Active subgroup mining: a case study in coronary heart disease risk group detection*' , Artificial Intelligence in Medicine, Vol. 28, No. 1, pp. 27–57.

[7] S.J.Gath and R.V.Kulkarni (2012), '*Expert System for Diagnosis of the Myocardial Infarction*', International Journal of Computer Science and Information Technologies, Vol. 3, No. 6, pp. 5315-5321.

[8] D.Ge, L.Sun, J.Zhou and Y.Shao (2009), '*Discrimination of the myocardial infarction stages by subjective feature extraction*', Computer Methods and Programs in Biomedicine, Vol. 95, No. 3, pp. 270–279.

[9] V.Mahesh, A.Kandaswamy and R.Venkatesan (2009), '*Diagnosis of Heart Disease Using Data Mining Algorithm*', International Journal of Recent Trends in Engineering, Vol. 2, No. 3, pp. 360-390.

[10] S.A.Hannan, V. D.Bhagile, R. R.Manza and R. J.Ramteke (2010), '*Diagnosis and Medical Prescription of Heart Disease Using Support Vector Machine and Feed forward Back propagation Technique*', Vol. 2, No. 6, pp. 2150- 2159.