+IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Multi Parametric Approach Using Fuzzification on Heart Disease Analysis

Upasana Juneja[*1], Deepti Dhingra[2]
[*1,2] Department of CSE, PIET, Samalkha, Panipat, Kurukshetra University, India
upasana_jdkps@yahoo.com

Abstract

The aim of this study is to define a fuzzy based prediction system that will accept the patient basic information as well as the symptoms as input and identify the chances of heart disease. Heart Disease Prediction System is the system that helps to predict the heart disease mainly cardiovascular. The importance of heart disease prediction system can be visualized from the fact that heart disease is one of the diseases that causes highest mortality rate. The present system helps in diagnosis of heart disease effectively and reduces the overall cost. Various techniques like Artificial Neural Network, Naïve Bayes etc can be used to implement this system .Data Mining operations is the improvement over the statistical methods that enables a user to perform the future analysis. A Fuzzy based soft computing approach is been implemented on multiple parameters to predict the heart disease. In this paper, a fuzzy database management system is introduced to diagnose the severity of the heart disease of a patient by using existing data in the common database systems. Crisp values are transformed into fuzzy values through the fuzzification and Transformation of fuzzy set into crisp values is called Defuzzification.

**Keywords**:Fuzzy System, Heart Disease, Operator, Dataset, Data Mining, and Prediction based.

Introduction

Medical history data comprises of a number of tests essential to diagnose a particular disease. It is possible for the healthcare industries to gain advantage of data mining by employing the same as an intelligent diagnostic tool .Data mining extract the patterns in the process of knowledge discovery in the database. As the dataset has grown in size and complexity, new emerging field of data mining provides new techniques and methods which help to analyse And understand large bodies of data. Data mining involves some common methods that are association rule learning, Clustering, classification, regression, summarization and sequential pattern matching. Classification is one of the important techniques of data mining. Classification is the processing of finding a set of models (or functions) which describe and distinguish data classes or concepts.

A medical diagnostic system cannot be implemented without an expert personal. In this present we have defined a medical application to identify the person heart disease or the chances of heart disease. For this heart disease analysis we have used one of the major soft computing techniques called fuzzy logic. The fuzzy logic is one of the intelligent schemes to deal with uncertainty. The fuzzy logic takes the intelligent decision based on the

characteristics of current population set. The fuzzy rule is implemented on the physical characteristics of a person.

The influence of data mining on the quality of Health Care cannot be understated. All Health Care organizations retain detailed and comprehensive records of patient data. Trends and patterns identified in these records can positively impact the quality of Health Care. The huge amounts of patient data make identification of these trends a difficult task. The data mining applications built for this purpose can make this very simple and produce efficient results.
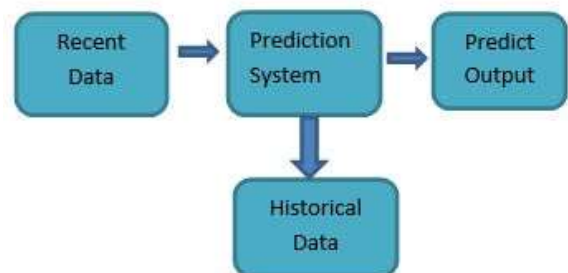
In this present work the data mining is been studied under the health care system. Health care system is one of the crucial applications of data mining where the accurate results are expected after the appropriate

processing and even then the obtained results cannot be used directly without the expert concern. In clinical area, the data mining is used to perform the prediction, classification, clustering etc.

Clustering: When we have a large dataset, then instead of processing the dataset individually, the dataset is subdivided to the smaller units called clusters clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). There are number of clustering approaches such as C-Means, K-means clustering approach etc[12].

Decision Tree: Decision tree, as the name suggest is the tree based approach in which the decision are represented by the parent nodes and the associated events are represented by the child nodes. This kind of algorithm is used basically to perform the data classification

Linear Regression: It is the statistical approach that work as the filtration as well as the analytical approach to perform the prediction of the data values. The regression is basically the analysis of an attribute with respect to one or more attributes of the same dataset. The regression model is then used to predict the result of an unknown dependent variable, given the values of the independent variables.

Naive Bayes or Bayes' Rule: is the basis for many machine-learning and data mining methods. The rule (algorithm) is used to create models with predictive capabilities. It provides new ways of exploring and understanding data.

Literature Survey
Asha Raj Kumar performed a work," Diagnosis Of Heart Disease Using Data mining Algorithm". The data classification is based on supervised machine learning algorithms which result in accuracy of time taken to build the algorithm. The classification of data is done by Tanagra tool and the data is evaluated using 10-fold cross validation and the results are compared.

Abdel Badeeh M. Salem performed a work," Intelligent Technologies for Medical eLearning". The new forms of artificial intelligence- based education software have been created in the last few years that allow the computer to cat as an intelligent tutor. Such intelligent tutoring systems (ITSs) can adjust its tutorial to the knowledge, experience, strengths, and weaknesses of student; it may also be carry a natural language dialogue.

Himigiri Danapana worked on," Effective Data Mining Association Rules for Heart Disease Prediction System". This paper gives a survey of current techniques of knowledge discovery in databases using data mining techniques that are in use in today's medical research particularly in Heart Disease Prediction. Various experiments have been conducted by the author under the decision tree based data mining operation to obtain the effective results.

Yi Mao defined a data mining technology so that information will be derived regarding the disease status. The patient has designed to identify the signs of clinical deterioration and to provide the early identification of serious clinical events. Presented system is about to provide reliable early alarms for patients at the general hospital wards (GHWs).

Muhammad Hariz Muhamad Adnan worked on," Hybrid Approaches Using Decision Tree, Naïve Bayes, Means and Euclidean Distances for Childhood Obesity Prediction". This paper studied childhood obesity prediction using different data mining techniques and proposed hybrid approaches to improve the accuracy of the prediction.

Method
Data
We consider a set of m diseases D, and define a collective set of n features F relevant to these diseases. Usually we have n>>m. Let:
$$D = \{d_1, d_2, d_3, \ldots, d_m\}$$
$$F = \{f_1, f_2, f_3, \ldots, f_n\}$$
To specify the symptoms of a patient, he would be checked against all features in the set F and a value would be assigned to each feature. The values are selected from the set:
{Very Low, Low, Moderate, High, Very High}.

Parameters
1) *Blood Pressure*: It is the major factor that can predict the chances of heart disease. Generally a heart patient has the higher blood pressure. According to the fuzzy rule set we have divided the patient blood pressure in four major categories called "Very High", "High", "Normal" and "Low". If the blood pressure is greater than 182, it is considered very high. If the blood pressure is between 154 and 182 it is considered high blood pressure. If it is between 125 and 154 then it is considered medium otherwise it is considered "Low*".*
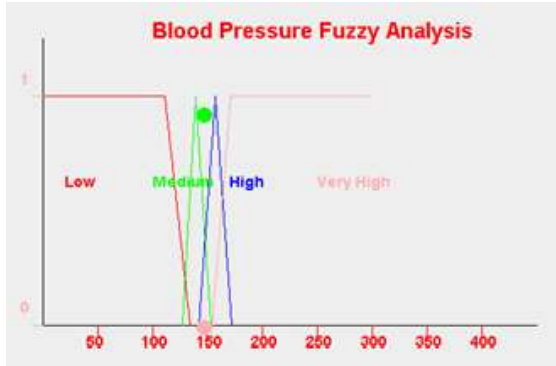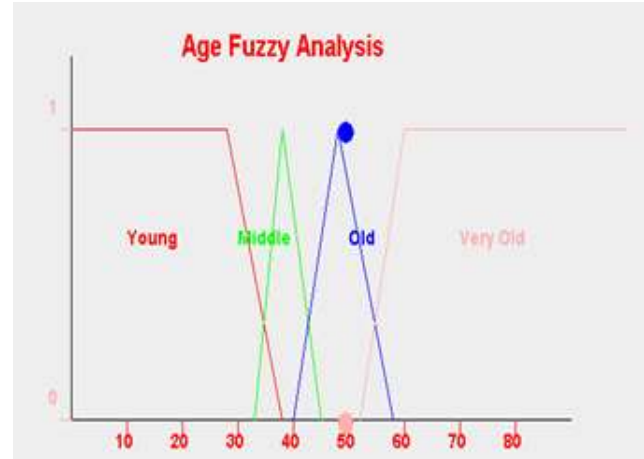
Blood Pressure Fuzzy Analysis

2) Heart Beat: Heart beat is another major factor that can increase the chances of heart disease in a patient. Higher the Heart Beat value more chances of heart disease. According to the study we have divided the Heart Beat level in a patient in three major categories called "High", "Normal" and "Low". If the Heart Beat Level is greater than 184, it is considered high. If Heart Beat level is between 140 and 184 it is considered Average Heart Beat otherwise it is considered "Low".
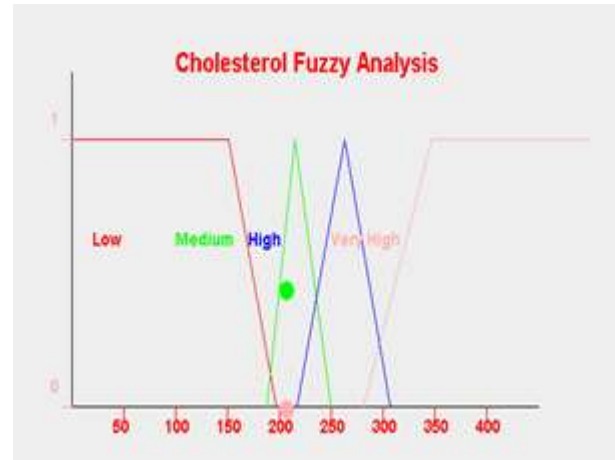

Heart Beat Fuzzy Analysis

3) Age: Age is another major factor that predicts the chances of heart disease in a patient. According to the study we have divided the Age of a patient in four major categories called "Young", "Middle", and "Old" and "Very Old". If the Age is greater than 52, it is considered Very Old. If the age is between 40 and 52, it is considered Old, between 33 and 45 is middle otherwise it is considered Young.
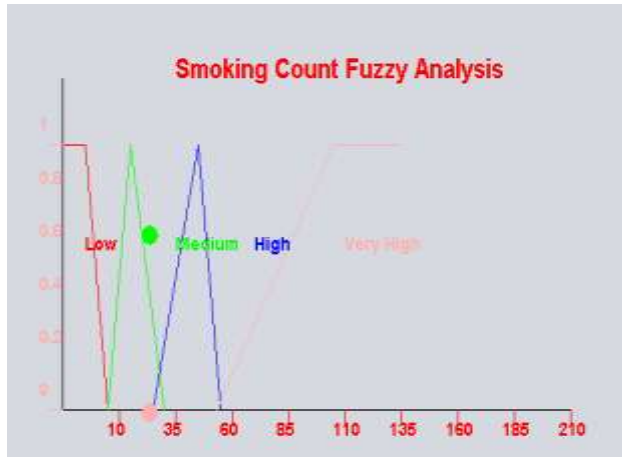

AgeFuzzy Analysis

4) Cholesterol: Cholesterol is another major factor that can increase the chances of heart disease in a patient. Higher the Cholesterol value more chances of heart disease. According to the study we have divided the cholesterol level in a patient in four major categories called "Very High", "High", "Normal" and "Low". If the cholesterol level is greater than 240, it is considered very high. If cholesterol level is between 217 and 240 it is considered high blood pressure. If it is between 180 and 200 then it is considered medium otherwise it is considered "Low".

Cholesterol Level Fuzzy Analysis

5) Smoking: Smoking is another major factor that can increase the chances of heart disease in a patient. Higher the value of smoke count more chances of heart disease.

Smoking Level Fuzzy Analysis

6) Sugar Level: Sugar Level is another major factor that can increase the chances of heart disease in a patient. Higher the level of sugar more chances of heart disease. According to the study we have divided the sugar level in a patient in four major categories called "Very High", "High", "Normal" and "Low".

Sugar Level Fuzzy Analysis

Fuzzy With Operators

A) Union (logical OR)- The membership of an element in the union of two fuzzy sets is the larger of the memberships in these sets.

(A OR B)=max ((A), (B))
e.g., (tall OR small) = max ((tall), (small))

B) Intersection (logical AND) - The membership of an element in the intersection of two fuzzy sets is the small of the memberships in these sets.

(A AND B) = min ((A), (B))
e.g., (tall AND small) = min ((tall), (small))

User can select the input values for different attributes as well as the criticality of these values is defined.

Fuzzy Without Operators

Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based.

Fuzzy logic includes 0 and 1 as extreme cases of truth but also includes the various states of truth in between, example; a cup of coffee is not exact "Hot" or "Cold" but ".48 of coldness".

In 'Fuzzy without Operators' two queries are considered

a) Fuzzy Query
Select * from patient info where Age is 'Nominal Young', Blood Pressure is 'Nominal Low', Cholesterol is 'Nominal Low', Heartbeat is 'Nominal Low', Smoking is' Nominal Low', Sugar Level is 'Nominal Low'.

b) SQL Query
Select * from patient info where agey>=0 and agey<=.3 and blow>=0 and blow<=.3 and clow>=0 and clow<=.3 and hlow>=0 and hlow<=.3 and smokel>=0 and smokel<=.3 and sugarl>=0 and sugarl<=.3

The fuzzy query is here performed on different attributes including age, blood pressure, cholesterol level, smoking, sugar level and heartbeat. The input is also taken in the form of accepted fuzzy values and the criticality level of these values. Based

on this fuzzy query, the defuzzification is here performed and relatively records are accessed from the database and presented in the form of table.

Fuzzy Without Operators

Age Membership Function

Here 1ˢᵗ figure is showing two queries and below is showing the age data retrieved on the basis of query.

E. Prediction Value

In the end Prediction Value is found which gives the criticality of disease in patient. A final decision is also taken for each patient.

The decision can be "Very Low", "Low", "Medium", "High", "Very High".
The numerical value is the mean of the numerical values corresponding to the degree of membership at which the membership function is scaled.

Decision And Prediction Value

Conclusion

The work is under the intelligent system that can be adapted by a doctor. In this work we have taken a parameter based fuzzification that will perform the analysis based on some parameters. The present work is the analysis on the patient symptom information based on which a pre-level decision is taken to identify the chances of a heart disease. Specifically, the whole application software finds the frequent illnesses with medication. This research provides important facts like correlations between medical issues related to disease finally.

Future Directions

In this thesis, supervised learning techniques have been used. this thesis can be extended further by using various other supervised learning techniques like density based clustering, Naive Bayes, Artificial Neural Network based classification and frequent pattern matching based association. It is also possible to extend the proposed methodology on other type of disease and analysis can be carried out.

References

[1] Milan Kumari," Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction", IJCST ISSN: 2229-4333 (Print)/ISSN:0976-8 491
[2] Jyoti Soni," Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", International Journal of Computer Applications (0975 – 8887)

[3] *Mr. Dhiraj Pandey," Prediction system to support medical information system using data mining approach", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622*

[4] *Jyoti Soni," Intelligent and Effective Heart Disease Prediction System using Weighted Associative Classifiers", International Journal of Computer Applications (0975 – 8887) Volume 17– No.8, March 2011*

[5] *Dr. D. Raghu," Probability based Heart Disease Prediction using Data Mining Techniques", IJCST ISSN : 0976-8491 (Online) | ISSN: 2229-4333(Print)*

[6] *Shantakumar B.Patil," Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network", European Journal of Scientific Research ISSN: 0975-3397 Vol. 3 No. 6 June 2011 2385*

[7] *T Srinivasan," Knowledge Discovery in Clinical Databases with Neural Network Evidence Combination".*

[8] *Sellappan Palaniappan," Intelligent Heart Disease Prediction System Using Data Mining Techniques", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.8, August 2008*

[9] *Cleveland database: http://archive. ics. uci. edu/ml/datasets/Heart+Disease*

[10] *T Srinivasan," Knowledge Discovery in Clinical Databases with Neural Network Evidence Combination".*