International Journal of Advanced Technology in Engineering and Science
Vol. No. 10, Issue No. 02, February 2022
www.ijates.com
ijates
ISSN 2348 - 7550

AN EXTRACTING RULE SET FOR CLASSIFYING MEDICAL DATA MINING FOR ACCURATE AND PREDICTION OF SYNDROME

Mrs.Dukitha.M [1], Dr.A.Banumathi [2]

Assistant Professor [1], Assistant Professor [2],

[1]Department of MCA, Er.Perumal Manimekalai College of Engineering, Hosur

[2]Department of Computer Science, Government Arts College, Karur

[1]dukitham@gmail.com, [2]Banukarthikeyan7811@gmail.com

ABSTRACT:

The Medical industry research is significant to the society and human values. This reduces the guess works and improves decision making with the pattern generated from the history of patient information. The complication of medical domain and manual analysis and extracting precise and accurate pattern of decision is a life-challenging task. Diagnosing with the maximum knowledge enriches the quick and accurate decisions. Though human decision making is best, it is poor when there are huge amounts of data to be classified. In many cases the symptoms are more prone to many of the diseases. Predicting the level of disease is a most significant task for finding the accurate treatment and early prevention of human life. Fuzzy Logic (FL) consists of probabilistic logic, which infers the rule set for classifying the medical data with fractional values of accuracy. This classification supports for accurate prediction and controlling the disease syndrome.

Keywords :Data Mining, Fuzzy Rule Set, Medical Data Classification and Control of Disease , Fractional Computation of accuracy.

1. INTRODUCTION

In day-to-day life medical advancement creates medical data a incredible growth. The complexity of medical domain makes the practitioner difficult in decision making. As the diseases are more diverse, the Medical practitioner suffers from the problem of exact classification of disease based on the symptoms notified on the patient. Historical symptoms analysis gives constructive effects in early prediction of disease and fast decision making for new treatment. In such discoveries pattern recognition evolves with the clustering and classification among different patterns assist in knowledge detection. All expert diagnosis may subject to error, the incompleteness of information, and infuses by improbability. The correction of this inaccuracy requires strong assumptions which are difficult in practice. All the defined methods are efficient with the less voluminous of data and needs training

data for prediction. Whereas Fuzzy logic (FL) is used improving the accuracy of classification and prediction of diseases with high voluminous of data and requires no training needs.

2. FUZZY LOGIC

FL consists of valued logic in which the truth value of variables may be any real number between 0 and 1or many-valued logic. Rather than fixed and exact reasoning, it provides approximate reasoning. It provides an organized, perceptive and mathematic way of managing uncertainty. FL can have multi values instead of fixed values (Asma Hashmi & Muhammad Saleem 2015). As a replacement for the term "Completely true" or "Completely false, in FL they are allocated with a value which corresponds to their degree of accuracy. In fuzzy systems, the values are represented in a range from 0 to 1. 0. 0 stands for complete falseness and 1.0 stands for absolute accuracy.FL allows the machine to respond ‘to a certain degree'.

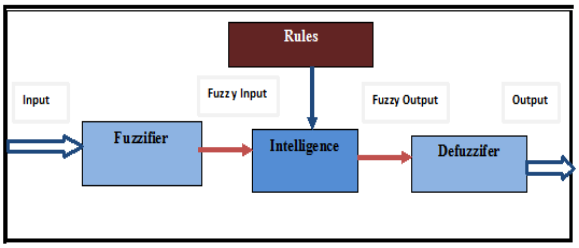

Fig. 1. Fuzzy Logic

3. DEDUCTION CLASSIFIER

The computational intelligence with logical information inference can be achieved by a projected Fuzzy Inference Classifier (FIC). It is a model of mapping from the known input to predict an output using FL and it is used to detain both complex stochastic and deterministic uncertainties. This model generates fuzzy classification rules using support and count framework.

It makes use of

- Fuzzy set
- Computing membership of the fuzzy input variables
- Fuzzy inference and
- Defuzzification

3.1 Fuzzy Set

A fuzzy set (FS) is any site that permits its members to have different grades of membership degree. FS is defined by two values, the FS member x denoted as FS (x) and the membership function ((x)). In the fuzzy theory, an FS A in X is defined as

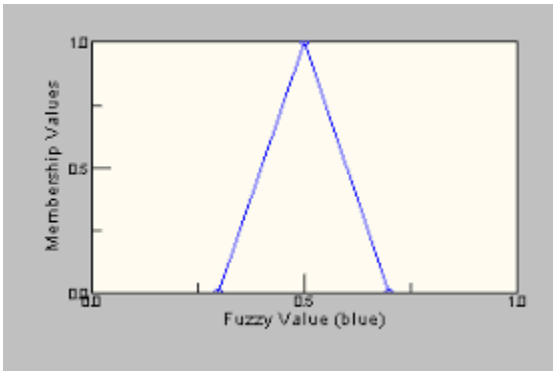

Fig. 1. Fuzzy set

- ***3.2 Membership Function***

Membership function (MF) defines a fuzzy set of mapping crisp values from its data set attributes to the set of associated degree of membership values in the interval [0-1]. It is also known as fuzzy input. A label is a descriptive name used to identify an MF. The number of label corresponds to the number of areas that the boundary ought to be divided, such that every label describes a region of behavior. The type of MF used to compute degree is based on the base set

- ***3.3 Fuzzy Inference***

The inference is the process of deriving new knowledge from the existing knowledge from the database. Fuzzy reasoning adopts probabilistic inference procedure which uses the input values, fuzzy sets (FS), available fuzzy if – then rules and known fact to conclude probabilistic FS and probabilistic fuzzy rules (FR). Consuming these probabilistic elements, a distinctive defuzzification is applied to infer the output of reasoning means. Fuzzy
Fuzzy inference is performed in three steps

- Identify the input, output attributes, FS and fuzzy numbers
- Fuzzification of the input variables,
- Rule evaluation

3.3.1.Identify the input, output attributes, fuzzy set and fuzzy numbers

Input variables and their value ranges are identified to test the proposed fuzzy evaluation method and output variable is Disease Closure
(DC)
FS are
, , ,
And
$Dc\{verylow, low, moderate, high, very\ hig$

The probable types of membership values are called the representation of Fuzzy set attributes are listed in Table 1.

Table 1.Representation of Fuzzy input variables, fuzzy membership variables

Fuzzy Input Variables	Fuzzy
i_1	1
	2
	3
i_2	1
	2
	3
3	1
	2
	3

3.3.2 Fuzzification

Fuzzification may be defined as the process of transforming a crusty set to a fuzzy set or a fuzzy set to fuzzier set and are accompanied by associating degree of membership ranging from 0 to 1. Where x is fuzzy membership variables can be applied for low, moderate and high. denotes minimum limit (L) values of the feature, and Indicates maximum limit (values of the symptom . By using the membership function, the minimum and maximum limit values for other input symptoms are also computed.

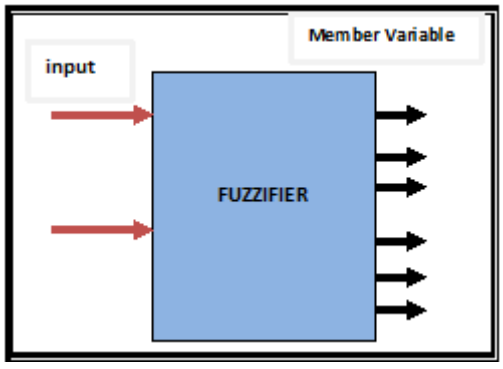

Fig. 4. ***Fuzzification***

3.3.3Fuzzy Evaluation

Fuzzy Evaluation endures the following steps

- Fuzzy Rule Generation
- Computing probability for Fuzzy Rules
- Aggregating the output of each rule

*3.1**: Fuzzy Rule Generation***

FR discovery is one of the major procedures of the fuzzy classification. The major part of FR construction is to construct membership functions and to determine a set of suitable fuzzy rules. The membership function and the FR set for this work were developed with the assistance of medical experts' knowledge base. The rules are constructed with the fuzzy values generated from the fuzzification process. The number of the FR is based on the number of the attribute data pairs. This rule base may contain the incompatible and redundant rules. In order to determine this, the support of a rule is examined by estimating the number of data that presents the same rule in each type of class. Then the FR in the rule base is prioritized according to their supports.

3.2: Computing Probability Inference for Fuzzy Rules

General form of fuzzy rule is

= if is ,.....,and/or is , then y is ,k, i=1,2,...N.Where is number of rules, ... are input variables, y is output variables; The FIS is accomplished with the use of the FR and the examination is done with the help of data sets.

For each generated rule, compute the firing strength This firing strength measures the strength of the areas belonging to the fuzzy region covered by the rule.

Step 3.3: Aggregating the output of each rule

Fuzzy rule support is computed as the sum of the weight of all the rules whose class label is true, divided by the sum of rules. And is given by

where, i=1..n ,N- number of records in the rule base.

Combine the fuzzified consequent probable rule sets, by combining the rule strength and the output membership function with maximum grade compatibility

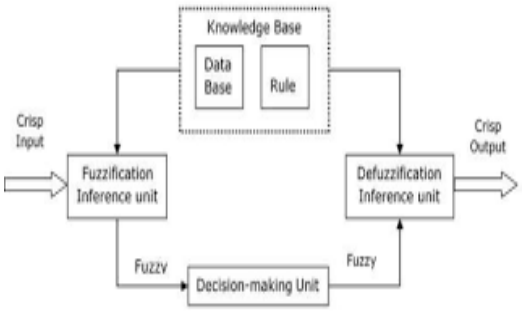

Fig. 5. Probabilistic Fuzzy Inference Model

3.4 Defuzzification

After the aggregation process, there is a fuzzy set for each output variable. To convert the fuzzy output into crisp output, probabilistic defuzzification is required in this method. Probabilistic centroid of set defuzzification function is used for defuzzifying fuzzy output. The weighted average of a data points is used to compute centroid.

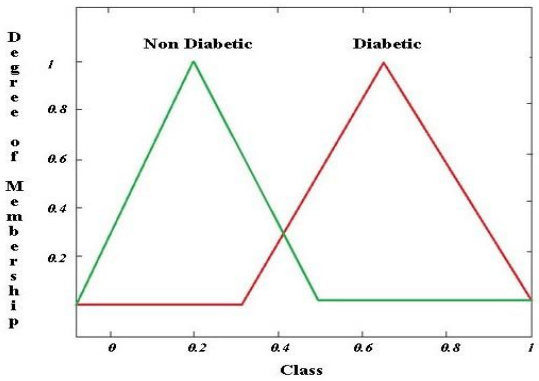

Fig 6. Classification of Output Fuzzy Set

Probability Computation of a Disease

In this section result analysis has been carried out to analyze the possibilities of disease infection, the experimental study generates three classes to represent the high, low and medium possibility of disease infection. The attributes are analyzed on the basis of their fuzzy values. If a patient has 50% of symptoms of diabetes, then he/she may prove to be diabetes with medium probability. Similarly, if a patient has 40% of symptoms, then he/she may prone with low probability.

Table 4. shows the probability and its range description.

Table 2. Description of the Probability of the Disease Diagnosis

Probability	Range Description
0 -.40	Low
0.40- 0.60	Medium
0.60 -1.0	High

4. CONCLUSIONS

Thus the pattern recognition progress with the clustering and classification among different patterns assist in knowledge detection .FL is comparable to person decision making and the ability to produce accurate decisions even

with incomplete and noisy data. It also produces the Boolean outputs which helps to predict ad control the disease syndrome earlier with accuracy

REFERENCES

[1]. .Ahmad, P, Qamar, S and Rizvi, SQA.(2015). Techniques of Data Mining In Healthcare: A Review. *International Journal of Computer Applications, volume (120)*

[3]. Bai, L, Liang, J, Dang, C and Cao, F.(2013).A novel fuzzy clustering algorithm with between-cluster information for categorical data. Fuzzy Sets and Systems, volume(215)

[5]. Dash, T, Nayak, SK and Behera, H. (2015).Hybrid Gravitational Search and Particle Swarm Based Fuzzy MLP for Medical Data Classification. *Computational Intelligence in Data Mining, Springer, volume*(1)

[6]. Haraty, RA, Dimishkieh, M and Masud, M.(2015).An enhanced k-means clustering algorithm for pattern discovery in healthcare data.*International Journal of Distributed Sensor Networks, Hindawi,*

[7]. Jamsutkar, K, Patil, V and Meshram, BB. (2013).Query Processing Strategies in Distributed Database. *Journal of Engineering Computers & Applied Sciences, volume(2)*