

# IMPROVED FUZZY ASSOCIATION RULES FOR LEARNING ACHIEVEMENT MINING

Mimanshu Gupta<sup>1</sup>, Chandra Prakash Singh<sup>2</sup>, Rohit Miri<sup>3</sup>

Department of Computer Engineering  
<sup>1</sup>College of Science and Engineering  
<sup>2</sup>College of Science and Engineering  
<sup>3</sup>Dr.C V Raman University  
India

**Abstract**— The fuzzy logic and Data mining is very useful technology for finding the associate rules between the two variables or amongst variable. Our work is to find out the association of variables of diabetic's patient by proposed novel algorithm. We separated the data for male and female so that we will get the association rules separately for diabetic's patient. The result will alert the normal person that causes the diabetic's patient.

The proposed algorithm is divided in two parts. First part is calculated by fuzzy techniques and second part in mining association rule. The proposed algorithm is useful for prediction and association result. The proposed algorithm finds the fuzzy mining association rule in very easy way.

**Index Terms**— Fuzzy Logic, Data Mining and Triangular Membership Function.

## I. INTRODUCTION

The diabetic's is now a day very challenging diseases. Most of people have been suffering from this. If we find the main reason behind this or alert the people by finding the main attributes association of diabetic's patient. There is different parameter behind this. We will find the main two or three parameter that affects more or chance or diabetics is more. We have collected the diabetic's patient data of male and female separately from hospital. We will analyses the data separately by proposed novel algorithm. The proposed novel algorithm combines both the technology in more effective way. The triangular membership function is used for calculating the fuzzy sets. The proposed algorithm also reduces the data complexity. The prediction of data sets is also performed. The novel algorithm is divided into two parts. In the first part the range of linguistic variable is decided and fuzzy sets are calculated by triangular membership function and the second part the fuzzy based proposed novel algorithm is applied. The practical work is performed by environment.

### A. Association Rule Mining

Association rule mining is a technique for discovering unsuspected data dependencies and is one of the best known data mining techniques.

An association is a rule of the form:

If X then Y:

It is denoted as  $X \longrightarrow Y$

Example: If India wins in cricket, then sales of sweets go up.

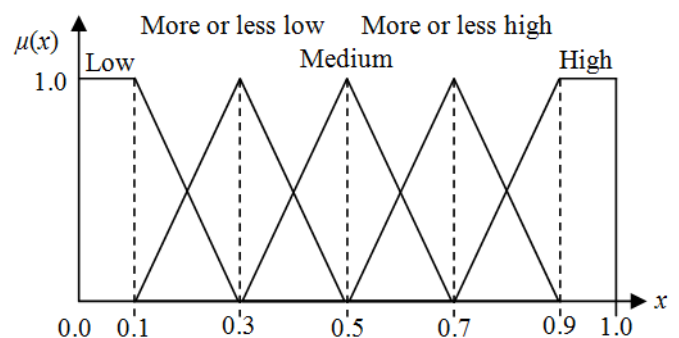
X is a India wins in cricket and Y is the prediction that sales of sweet go ups. So we say that we discover such a rule if we are able to conclusively says based on analyzing the data that whenever India wins in cricket the sales of sweet goes up. For any rule  $X \rightarrow Y$  or  $Y \rightarrow X$ , then X and y are called an "Interesting item set". Oder of the item sets x and Y is not an important.

### B. Fuzzy Logic

Fuzzy logic gives answer between 0 and 1. In other language we can say that fuzzy gives exact answer of any question. Traditionally logic represents using Boolean values i.e. 0 and 1 or yes and no.

### C. Fuzzy Set Operations

All the fuzzy sets use membership functions to represents fuzzy values. Membership Function are used to graphically represent fuzzy values. Triangular Membership Function is easy and used three Linguistic variables Low, Middle and High.



## II. REVIEW

In Previous Years, there is lot of research work has been done in this field.

- In July 2013[1], [J.Preethi] She has done comparison of efficiency by using two technologies. First technology apply mining association rule by using genetic algorithm. The second technology used is fuzzy mining association rule. They compared both the technology and found that the accuracy or efficiency of fuzzy logic is more as compare to genetic algorithm.
- In April 2008[2], [Jesús Alcalá-Fdez, Rafael Alcalá, María José Gacto, Francisco Herrera] proposed to define the range or domain of linguistic variable of fuzzy logic by using genetic algorithm. It is based on the 2-tuples linguistic representation model i.e. age and weight allowing us to adjust the context associated to the linguistic term membership functions.
- In Aug 2013[3], [Nandita Rane, Madhuri Rao] they used fuzzy mining association rule to early detection of diabetics. They used large data sets from medical center and perform data mining association rule on this data sets. The method not only can find direct factors but also find indirect factors that cause type 2 diabetes mellitus which may help health doctors to explore their data and understand the discovered rules better.
- In 2012[4], [Ziauddin, Shahid Kammal, Khaiuz Zaman Khan, Muhammad Ijaz Khan] Authors try to explore the application of fuzzy mining association rule. They explained the usefulness of important of association rule from last 15 years. Much research has been done on mining association rule. On the field of latest research it plays vital role for prediction of data set, association rule among the attributes. Uncertainty condition has been solved by using fuzzy association rule.
- In 2011[5], [Jr-Shian Chen, Hung-Lieh Chou, Ching-Hsue Cheng, Jen-Ya Wang] They defined the domain of fuzzy linguistic variable by using cumulative probability distribution approach (CPDA) by using mean and standard deviation.
- In 2008 [6], [E. Ramaraj, K. Ramesh Kumar, N. Venkatesan] Proposed three mining association rule i.e. AprioriTID, TRAApriori, HEA Algorithm. They have reduces the time complexity of algorithm in efficient way. Among three the TRApriori is very fast data mining association algorithm.
- In 2004[7], [Tzung- Pei Hong, Chan-ShengKuo, Shyue-LiaWang] In this paper, the authors are concentrated on reduced computational time by using fuzzy mining association rule.

## III. PROPOSED WORK

We have proposed the algorithm to eliminate the large data sets into smaller data sets that is very helpful to reduce the complexity between the candidate item sets.

The Male Dataset used throughout this paper shown in **Table 1**. The Male Dataset contains 15 instances, which is characterized by the following attributes – (I) Age (denoted Age), (II) Height (denoted Height), (III) Weight (denoted Weight), (IV) High Blood Pressure (denoted H.B.P.), (V) Cholesterol (denoted CL).

All these Attributes are numerical values. The proposed model is introduced in detail as follows.

S. No.	Sex	Age	Height	Weight	H.B.P	CL
1	M	65	155	65	150	198
2	M	64	154	62	140	165
3	M	69	152	64	140	158
4	M	70	153	61	120	148
5	M	68	160	65	120	198
6	M	65	148	62	125	197
7	M	62	159	52	130	196
8	M	61	160	51	120	135
9	M	64	160	58	140	136
10	M	63	162	64	150	206
11	M	72	164	59	145	204
12	M	71	168	53	140	245
13	M	58	140	55	145	182
14	M	55	150	48	150	172
15	M	59	152	68	140	218

**TABLE1.** The Male Data Set

**STEP:1** Partition the attributes by using linguistic variable, finding the mean of the statistical of male and female data separately.

S.N.	Attributes	Low	Middle	High
1	Age	(32,32,43)	(37.6,49.06,60.42)	(55,76,76)
2	Height	(160,160,163)	(161,166,172)	(169,180,180)
3	Weight	(58,58,66)	(62.3,69,75.6)	(72,87,87)
4	H.B.P.	(120,120,132)	(127.7,137.3,146.9)	(142,150,150)
5	Cl	(125,125,170)	(156,185.8,215.5)	(200,250,250)

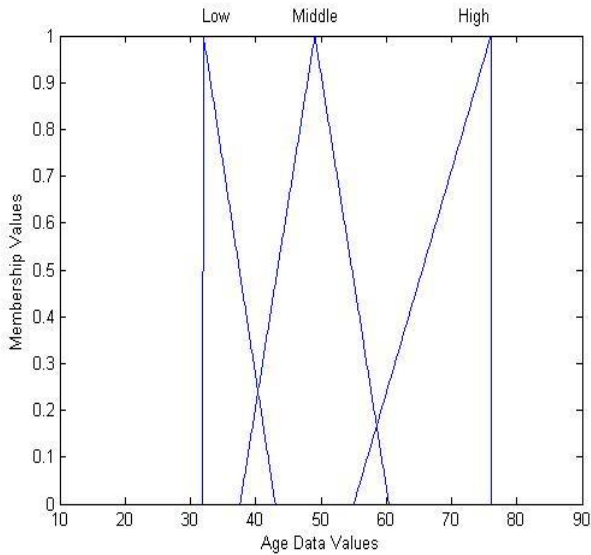
**TABLE2.** Range of Linguistic Variable

In this Step, We define Standard deviation, Mean, Minimum and Maximum values.

**STEP:2** Find the membership function of both the male and female data.

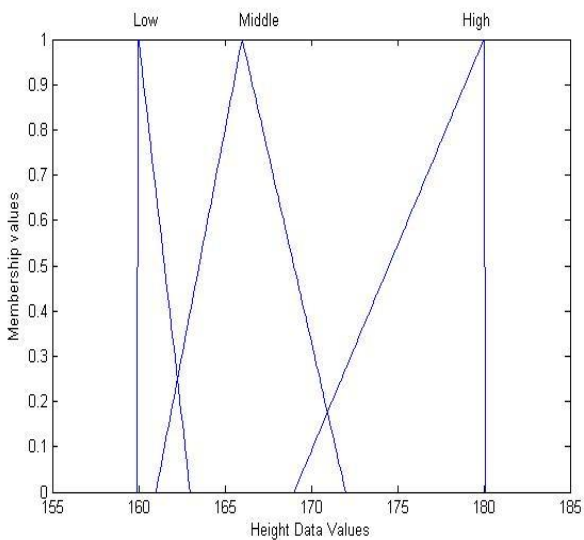
In this Step, We draw Membership Function through Standard Deviation, Mean, Minimum and Maximum Value.

❖ **Membership function for the Age Data Set-**



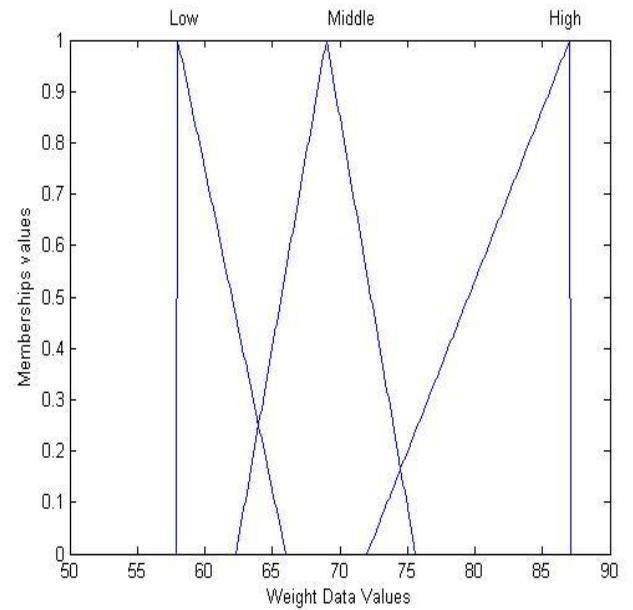
**Fig.1-** Triangular membership function for Age

❖ **Membership function for the Height Data Set-**



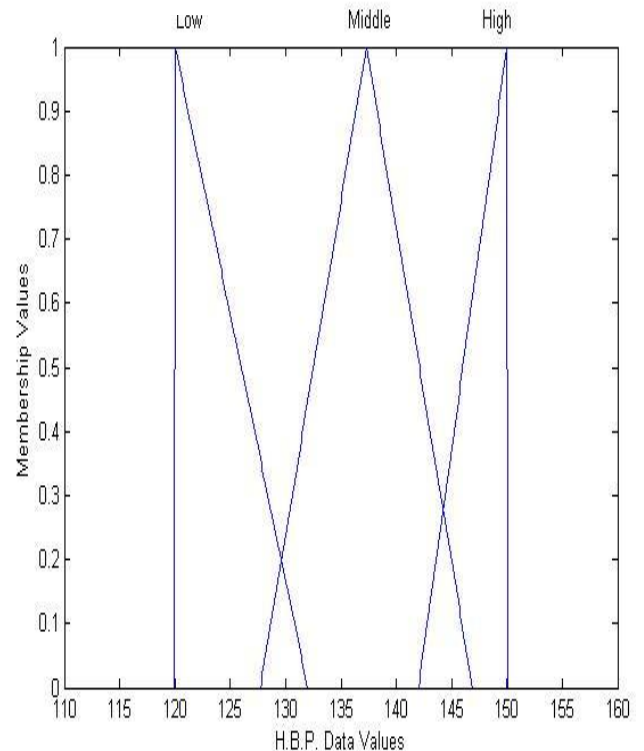
**Fig.2-** Triangular membership function for Height

❖ **Membership function for the Weight Data Set-**



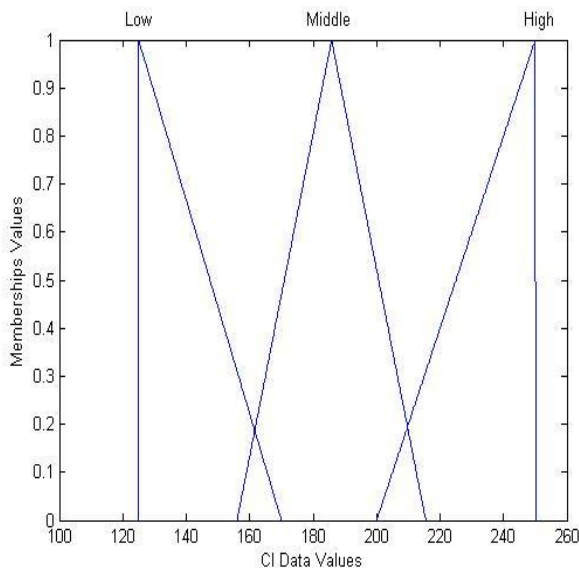
**Fig.3-** Triangular membership function for Weight

❖ **Membership function for the H.B.P. Data Set-**



**Fig.4-** Triangular membership function for H.B.P

❖ **Membership function for the CL Data Set-**



**Fig.5-** Triangular membership function for CI

**STEP:3** Fuzzifies the male and female data separately. In this step, we find fuzzy values through membership function(0 to max 1 value).

**STEP:4** Find out the minimum support value and minimum confidence support value of male data. In Our example 4 is minimum support value and 0.7 is confidence threshold value.

**STEP:5** Find out the candidate item set C1 for the male data separately.

**STEP:6** Find out the five largest values for the data separately.

**STEP:7** If only single attributes comes in an output then take another two maximum values. It means if our example consist maximum values with the same attributes then take another attributes but our example has no this type condition occurs.

**STEP:8** Find out the L1 item sets.

**STEP:9** Find out the C2 item sets.

**STEP:10** Repeat the steps 7 and 8 until the  $L_{i+1}$  large item sets is null.

**STEP:11** Find out the fuzzy association rule.

**IV. CONCLUSION**

The proposed algorithm is fast and finds out the important attributes that causes the diabetics. The proposed algorithm is more beneficial for normal person for alerting the diseases or for prediction of diabetic's diseases. The data is analysis periodically individually for both male and female data sets. The Association among diabetic's attribute are calculates. The Prediction is performed from diabetic's data sets for altering the people.

**V. FUTURE SCOPE**

For future work, we can predict the result of association rule without applying the Data Mining association algorithm based on the fuzzy classification technique. There will be possibility to improve or upgrade this algorithm. For Future work we can apply mixed approach to find the association rule.

**VI. ACKNOWLEDGMENT**

I would like to express my sincere thanks to my guide **Mr. Chandra Prakash Singh** for his helpful support and valuable advice during our research work.

**REFERENCES**

- [1] J. Preethi "Temporal Outlier Detection using Fuzzy logic and Evolutionary Computation" Proceedings of International Conference on Optical Imaging Sensor and Security, Coimbatore, Tamil Nadu, India, July 2-3, 2013.
- [2] Jesús Alcalá-Fdez, Rafael Alcalá, María José Gacto, Francisco Herrera "Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms(Science Direct)2008.
- [3] Nandita Rane, Madhuri Rao "Association Rule Mining on Type2 Diabetes using FP-growth association rule International Journal Of Engineering And Computer Science, August 2013.
- [4] Ziauddin, Shahid Kammal Khaiuz Zaman Khan, Muhamm "Research on Association Rule Mining" Advances in Computational Mathematics and its Applications (ACMA) ,2012.
- [5] Jr-Shian Chen, Hung-Lieh Chou, Ching-Hsue Cheng, Jen-Ya Wang Wang CPDA Based Fuzzy association rules for learning achievement mining, 2009 international conference on machine learning and computing IPCSIT vol3 (2011) IACSIT PRESS SINGAPUR.
- [6] E. Ramaraj, K . Rameshkumar, N. Venkatesan" A better Performed transaction Reduction algorithm for mining frequent item set from large voluminous database" Volume 5, Issue 1, Feb 2008.
- [7] Tzung-PeiHonga, Chan-Sheng Kuob, Shyue-Liang Wangc "A fuzzy AprioriTid mining algorithm with reduced computational time" Volume 5, Issue 1, December 2004.
- [8] K Sathesh Kumar and M Hemlath, "An innovative potential on rule optimization using fuzzy artificial bee colony" Research journal of applied science , engineering and Technology.volume2627-2633,ISSN:2040-7459,E-

issn:2040-7467, April 05 2014.

- [9] Neelukhare, Neeruadlakha, K R pardasani “An algorithm for mining multidentional fuzzy association rule” international journal of computer science and information security vol 5, no-1,2009.
- [10] Thomas sudkomp “Refinement of temporal constraints in fuzzy association ” volume 48(2008) 14 June 2007.
- [11] Tzung-Pei Honga,Chan-Sheng Kuob,Shyue-Liang Wangc “A fuzzy AprioriTid mining algorithm with reduced computational time” Volume 5, Issue 1, December 2004, Pages 1–1.