

ID3 Derived Fuzzy Rules for Predicting the Students Academic Performance

Anita Chaware¹, Dr. U.A. Lanjewar²

¹(P.G. Department of Computer Science, SNDT WU, Mumbai, India)

²(MCA, VMV college of Arts and Science, Nagpur, India)

Abstract: This paper presents a technique to use ID3 decision rules to produce fuzzy rules to get the optimized prediction of the students academic performance. In this paper, a the student administrative data for a class is used in order to classify the students final year marks in fuzzy logic prediction. This paper is using the machine learning approach to generate the rules so as to overcome the difficulties in a conventional approach like deriving fuzzy rules base from expert experience. This research provides us with: a way to produce meaningful and simple fuzzy rules; a method to fuzzify ID3-derived rules to deal with many inputs variables; and a de-fuzzification system to get the output in human understandable form. The ID3 tree is generated by the WEKA software and is utilized by the Fuzzy Inference System. A Fuzzy inference system was constructed to give the final crisp output. The ID3 was generated on 300 training data to get the better output. The output of our Fuzzy Student Performance Predictor was then tested on 50 test data to check for the accuracy.

Keywords: Student administrative data, Students performance, ID3 Decision tree, Fuzzy rules, Fuzzification, linguistic variables, Membership function, Defuzzification

I. Introduction

Many statistical, data mining and machine learning approaches have been developed for prediction of students Academic Performance by researchers from all over the world [1], [2]. Prediction of Student performance is an active area because it depends on various factors affecting the students performance like the demographic, personal information. Therefore, every researcher has different applied different method to get the optimized results with nearing to 100% accuracy. The higher education has to be improved if the research and skill among the students has to be developed in universities. Many education researchers and instructors have made extensive efforts in constructing effective models to predict student academic performance in a class [1]. This is to reduce the numbers of failures and dropouts of the school. Data mining is one of the major area which help us to predictions and classification. The classification, clustering, etc methods of data mining are capable of classifying the data in data warehouse having large amount of data. They are used to predict categorical class labels and classifies data based on training set, and hence can be used for classifying newly available data[3][4]. Thus it can be outlined as an inevitable part of data mining and has gained more popularity.

Instructional intervention is needed as students come from diverse background and to make the learning outcome better. To make instructional intervention, knowledge is acquired by modelling the previous database to find which students need it. One among many existing model is the predictive model to predict the results using factors affecting the results. Predictive models can be useful to the instructor to predict student academic performance and then take some proactive measures like designing an innovative and effective teaching and learning plan to help these academically at-risk students. Additionally, predictive models can reduce the dropout rate of students from relevant courses or programs.

Our objective was to explore the students' academic data of fresher students admitted in First year along with their First yr. marks with a view to classifying their performance using fuzzy logic technique. The paper in its next section 2 i.e. two discusses the scope and method used for this research. Section 3, we discussed briefly about ID3 decision tree and its generation by WEKA software. Section 4 explains how this decision tree can be converted into fuzzy rule. Section 5 we demonstrate the use of fuzzy logic concept to the student database and its how that fuzzy model of students performance predictor is used to predict the student final year marks. A 20 student database with its predicted values and actual values with their accuracy results are given in result section 6 Finally concluding the whole work and showing the future scope in section 7.

II. Scope And Method Used

This study is focused on developing and validating fuzzy mathematical model to predict student academic performance in the university. The predicted results are then validated with the actual results of students to find the accuracy of the model. The Academic performance of a student's is affected by number of factors like previous knowledge, interest, family background, motivation etc.[4]. In this paper we have considered the factors like the student's first year result and their class attendance, their HSC score Graduation

score, Parent income, Distance travel by them daily to college, to predict the final year result of a degree course. This model can be applied to any degree of any university and hence has a much greater scope in higher education. The uniqueness of this paper is that the student interest and motivation for the course can be found from their percentage of attendance in the classroom, it is one of the factors we have considered as input to our Linguistic Mamdani-type Fuzzy Rule-Based Systems – Fuzzy model for Student Performance predictor. Since the input to this fuzzy Inference system are more the number of rules will be more i.e if 6 input variables then $2^6=64$ rules will be possible. The usual Fuzzy system derive rulebase (i.e. set of fuzzy rules) from expert experience. In the system with 6 inputs it is difficult to get all the rules correct. Hence we have used the machine learning method called decision tree to generate the fuzzy rule base. The resulting tree is used to classify future samples i.e. Each branch of the decision tree generated by WEKA is formed as the rule of our fuzzy system. The whole methodology is thus, depicted in figure 1

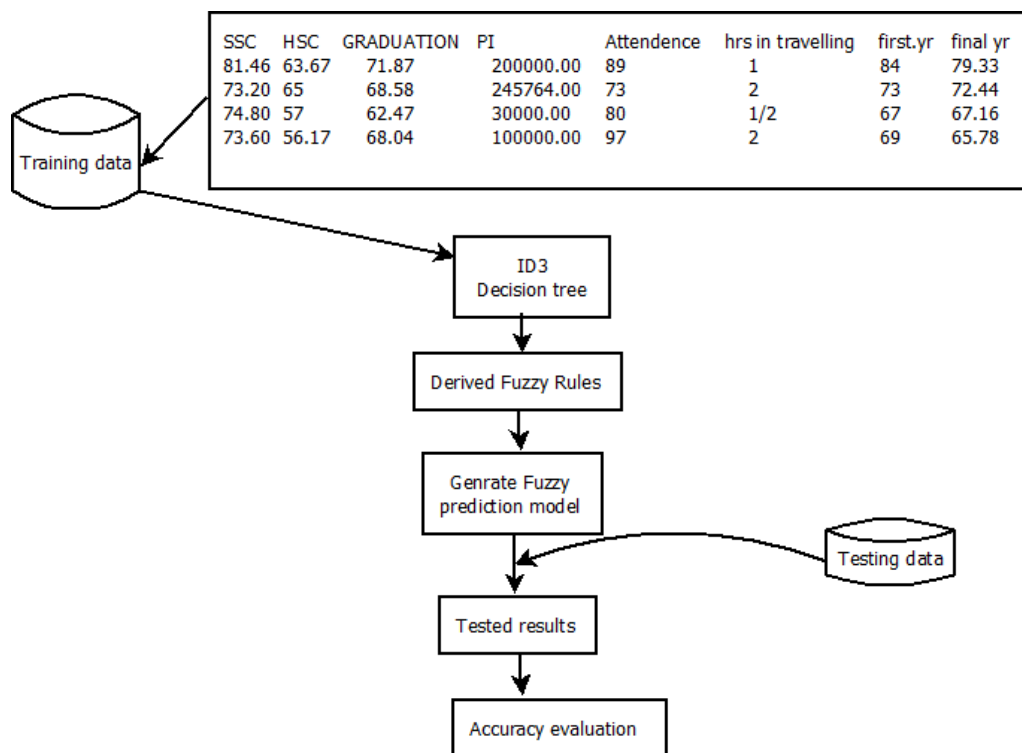


Fig 1: Flow of our Fuzzy Model of Student Performance Predictor.

III. ID3-Decision Tree Using Weka

The tool used to generate the ID3 (Iterative Dichotomiser 3) decision tree is WEKA. The full form of WEKA: Waikato Environment for Knowledge Learning. WEKA is a computer program that was developed by the student of the University of Waikato in New Zealand for the purpose of identifying information from raw data gathered from agricultural domains [6]. Data preprocessing, classification, clustering, association, regression, visualization and feature selection these standard data mining tasks are supported by WEKA. It is also well-suited for developing new machine learning schemes. It is an open source application which is freely available. The input dataset or file have the format ARFF format. The WEKA Explorer will use these automatically if it does not recognize a given file as an ARFF file. In case of having data in excel, the file can be converted into .csv is also recognized by WEKA. Classify tab in WEKA Explorer is used for the classification purpose. A large different number of classifiers are used in WEKA such as bayes, function, tree etc. The steps to apply classification techniques on data set and get result in WEKA:

Step 1: Decision tree works on class-labeled training tuples hence the numerical data need to convert each attribute of the dataset into groups/class-labels

Step 2: Take the input dataset in ARFF / CSV format.

Step 3: Apply the classifier algorithm on the whole data set.

Step 4: Note the accuracy given by it and time required for execution.

The data required for the decision tree algorithm is :

1. HSC {U=0-35,P=35-45,S=45-59, F= 60-70, D=70-100 }
2. Graduation {U=0-35,P=35-45,S=45-59, F= 60-70, D=70-100 }
3. First yr. {U=0-40,P=40-49,S=50-59, F= 60-69, D=70-100 }
4. Final yr. {U=0-40,P=40-49,S=50-59, F= 60-69, D=70-100 }
5. PI in lakhs { vl=0-1, l=1-2, m=2-3, h=3-4, vh= 5 and above }
6. Attendance { vl=0-49, l=50-64,m=65-74,h=75-100 }
7. Distance travel { 0-29=vl, 30-59=l, 60-89=m, 90-120=h, 121 above = vh }

Table 1: The classes of students data with labels

The snapshot of converted file looks like the given below in Table 2.

HSC	GRADUATION	FY	Attendance	Final Yr.	Parent income	Distance travelled
D	D	D	POOR	D	MEDIUM	more
D	D	D	good	D	MEDIUM	more
F	S	D	good	F	LOW	normal
F	F	D	avg	D	LOW	more
S	F	D	good	F	LOW	more
D	D	D	POOR	D	LOW	more
F	S	D	avg	F	LOW	normal
D	D	D	avg	D	MEDIUM	more
D	S	D	POOR	D	LOW	more
D	F	D	avg	F	LOW	less
F	D	D	good	D	LOW	less
D	D	D	avg	D	MEDIUM	less
F	F	D	POOR	F	MEDIUM	more

Table 2: The class to learn is: final year marks of student

The first few branches of WEKA generated resultant tree is given below(due to space constraints):

Test mode:10-fold cross-validation === Classifier model (full training set) === ID3

```

FY = D
| PI = M
| | HSC = F
| | | Attend = H
| | | | DT = L: D
| | | | DT = M: D
| | | Attend = M: D
| | | Attend = L: D
| | HSC = S: F
| | HSC = P: F
| | HSC = D: D
| PI = VL: D
| PI = L: D
| PI = VH
| | HSC = F: F
| | HSC = S: D
    
```

IV. Tree to Fuzzy rules

Each branch of the ID3 tree represent a rule which can be further utilized as predictor of the the decision to make. We can put the rules in the Fuzzy linguistic format i.e. IFTHEN

The fuzzy rules are written as:

If <fuzzy proposition> then <fuzzy proposition>

The fuzzy proposition can be atomic(i.e. HSC is High) or compound (i.e. HSC is High and PI is low).

The following is a subset of the ID3 derived fuzzy rulebase we got from the WEKA generated tree:

- IF FIRST.YR =D & PI=M & HSC=F & ATTEND = H & DT = L THEN FINAL.YR = D
- IF FIRST.YR =D & PI=M & HSC=F & ATTEND = H & DT = M THEN FINAL.YR = D
- IF FIRST.YR =D & PI=M & HSC=S & ATTEND = M THEN FINAL.YR = D
- IF FIRST.YR =D & PI=M & HSC=S & ATTEND = L THEN FINAL.YR = D
- IF FIRST.YR =D & PI=M & HSC=S = THEN FINAL.YR = F
- IF FIRST.YR =D & PI=M & HSC=P THEN FINAL.YR = D

- .
- .
- IF FIRST.YR =D & PI=M THEN FINAL.YR = D

V. Linguistic Mamdani-type Fuzzy Rule-Based Systems

Based on the FRBS developed by Zadeh in 1965 [7] Mamdani Fuzzy Rule-Based Systems (FRBS) constitute an extension of classical Rule-Based Systems, because they deal with fuzzy rules instead of classical logic rules. Mamdani FRBSs are composed of four main components: a Knowledge Base (KB), an Inference System and the Fuzzification and Defuzzification Interfaces. The KB —composed of Data Base (DB) . and Rule Base (RB)— stores the available knowledge about the problem in the form of linguistic “IF-THEN” rules. The Inference System puts into effect the inference process on the system inputs making use of the information stored in the KB. The Fuzzification Interface establishes a mapping between crisp values in the input domain universe U of the system outputs and fuzzy sets defined in the same universe of discourse. On the other hand, the Defuzzification Interface develops the opposite operation by defining a mapping between fuzzy sets defined in the output domain and crisp values defined in the same universe[8].

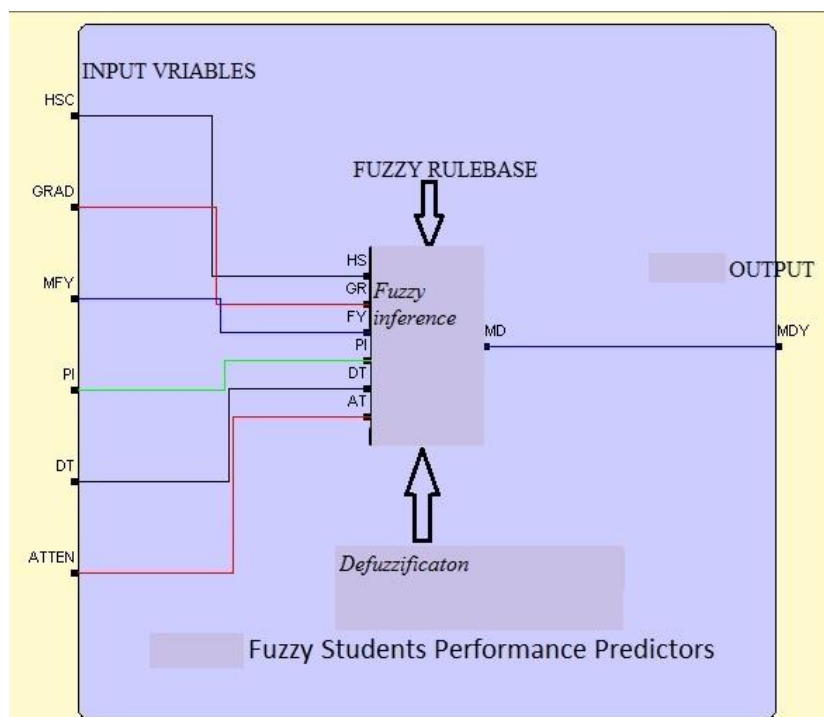


Fig 2.depicts the simple layout of the Fuzzy Model of Students Performance Predictors

5.1 Fuzzification

The purpose of fuzzification is to map the inputs from a set of given inputs to values from 0 to 1 using a set of input membership functions. The input called as antecedents given to this fuzzy model was the first year marks and the Attendance. To compute the fuzzification of the inputs we need to compute the Degree of membership function ($\mu(x)$) of each antecedents (inputs), using the formulas given below. The inputs to the fuzzy system are the linguistic variables are "FY" , "PI" , "HSC" , "Attend" , "DT" and "DY". The linguistic values for "FY" are MU, MP, MS, MF, MD representing under fail marks, pass class, second class, first class and Distinction . The membership function for FY in the universe of discourse, $U=\{MU,MP,MS,MF,MD\}$, are defined below:

- MU [0, 50]
- MP [40,60]
- MS [50,70]
- MF [60,80]
- MD [70,100]

$$\mu_{MU}(x) = \begin{cases} 0 & \text{if } x < 0 \text{ or } x \geq 40 \\ (x - 0) / (50 - 0) & \text{if } 0 < x < 50 \\ 1 & \text{if } x = 50 - 0/2 \end{cases} \quad 5.1$$

$$\mu_{MP}(x) = \begin{cases} 0 & \text{if } x < 40 \text{ or } x \geq 60 \\ (x - 40) / (60 - 40) & \text{if } 40 < x < 60 \\ 1 & \text{if } x = 60 - 40/2 \end{cases} \quad 5.2$$

$$\mu_{MS}(x) = \begin{cases} 0 & \text{if } x < 50 \text{ or } x \geq 70 \\ (x - 50) / (70 - 50) & \text{if } 50 < x < 70 \\ 1 & \text{if } x = 70 - 50/2 \end{cases} \quad 5.3$$

$$\mu_{MF}(x) = \begin{cases} 0 & \text{if } x \leq 60 \text{ or } x \geq 80 \\ (x - 60) / (80 - 60) & \text{if } 60 < x < 80 \\ 1 & \text{if } x = 80 - 60/2 \end{cases} \quad 5.4$$

$$\mu_{MD}(x) = \begin{cases} 0 & \text{if } x \leq 70 \text{ or } x \geq 100 \\ (x - 70) / (100 - 70) & \text{if } 70 < x < 100 \\ 1 & \text{if } x = 100 - 70/2 \end{cases} \quad 5.5$$

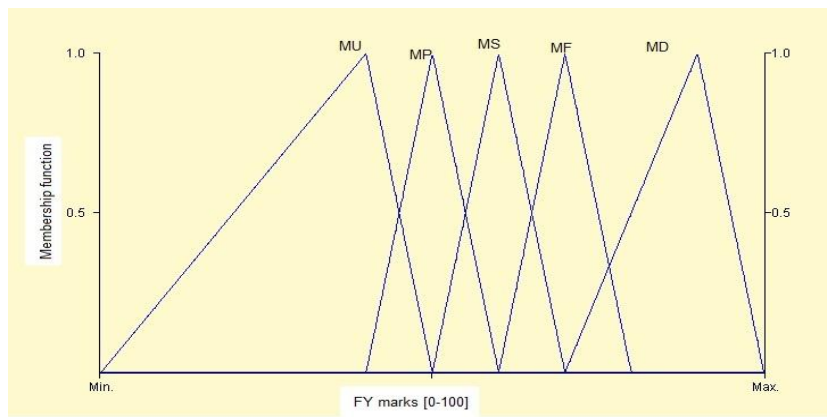


Fig. 3

Hence the Membership functions of the FY

first.yr	MU	MP	MS	MF	MD
84	0	0	0	0	0.933333
73	0	0	0	0.7	0.2
67	0	0	0.3	0.7	0
69	0	0	0.1	0.9	0
45	0.2	0.5	0	0	0
72	0	0	0	0.8	0.133333
58	0	0.2	0.8	0	0
65	0	0	0.5	0.5	0
80	0	0	0	0	0.666667
63	0	0	0.7	0.3	0

Table1: different membership functions value obtained for FY input variable

All the other input variables along with their membership function are depicted in Fig 3, Fig 4.

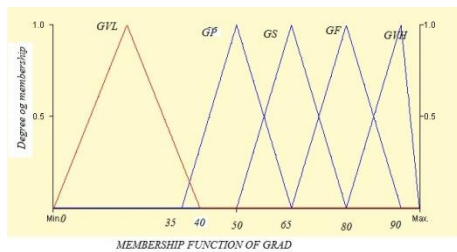


Fig 3: membership functions of Grad input variable

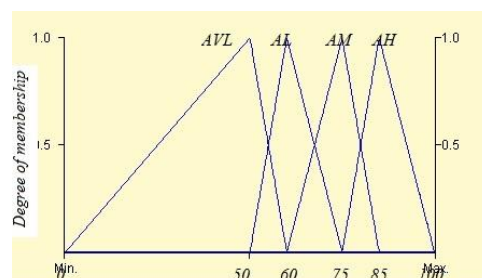


Fig 4: membership functions of Attend input variable

5.2 Rule Implication

To apply the Mamdani implication rule for in our example , the following rule applies

1. The or connective is replaced with the max (union) operator.
2. The Maximum of the all membership functions is evaluated for the antecedent part of the fuzzy rules.
3. The Mamdani Implication operator (i.e., min operator) is applied between the resulting antecedent membership function and the consequent membership function.

Suppose the FY =84 . The or connective is replaced with the union operator. Using the fuzzification union operator of Equation 7.6 yields

$$\begin{aligned} \mu_{FY}(84) &= \mu_{MU}(84) \cup \mu_{MP}(84) \cup \mu_{MS}(84) \cup \mu_{MF}(84) \cup \mu_{MD}(84) \\ &= \text{MAX} [\mu_{MU}(84), \mu_{MP}(84), \mu_{MS}(84), \mu_{MF}(84), \mu_{MD}(84)] = \text{MAX} [0,0,0,0,0.9] \\ &= 0.9 \end{aligned}$$

So the Degree of membership for FY = 84 =0.9

Below is the table of all linguistic variables along with their Degree of membership.

HSC	$\mu_{HSC}(x)$	GRAD	$\mu_{GRAD}(x)$	PI	$\mu_{PI}(x)$	Attend	$\mu_{ATTEND}(x)$	DT	$\mu_{DT}(x)$	FY	$\mu_{FY}(x)$
64	0.9	72	0.5	200000	0.5	89	0.9	1	1	84	0.9
65	1	69	0.7	245764	0.8	73	0.2	2	1	73	0.7
57	0.5	62	0.8	30000	0.6	80	0.4	0.5	1	67	0.7
56	0.6	68	0.7	100000	0.5	97	0.2	2	1	69	0.9
46	0.7	73	0.6	96000	0.5	73	1	2	1	45	0.5
70	0.7	70	0.7	100000	0.5	75	0.8	1	1	72	0.8
42	0.5	66	0.9	50000	1	81	0.5	1	1	58	0.8
50	1	56	0.6	150000	1	87	1	1	1	65	0.5
71	0.6	62	0.8	100000	0.5	87	1	1	1	80	0.7
47	0.8	60	0.7	50000	1	87	1	0.5	1	63	0.7
54	0.7	67	0.9	100000	0.5	87	1	0.5	1	60	1
62	0.8	59	0.6	200000	0.8	60	0.8	0.5	1	57	0.7
68	0.8	66	0.9	200000	0.8	80	0.4	2	1	70	1
64	0.9	69	0.7	150000	1	80	0.4	2	1	62	0.8
47	0.8	58	0.5	75000	0.5	87	1	1	1	57	0.7
60	0.7	60	0.7	200000	0.8	67	0.6	0.15	0.3	66	0.6
70	0.7	57	0.6	100000	0.5	93	0.6	0.15	0.3	66	0.6
62	0.8	78	0.9	250000	0.8	93	0.6	0.15	0.3	71	0.9
52	0.8	71	0.6	90000	0.4	67	0.6	0.15	0.3	64	0.6

Table 2: The degree of membership for all input variables or rule ascendant

5.3 Defuzzification

The implication result obtained for each rule should be aggregated and defuzzified to obtain a single crisp values. The centroid defuzzification technique is used here as shown by:

$$x^* = \frac{\int \mu_{output}(x) \cdot x \, dx}{\int \mu_{output}(x) \, dx} \tag{5.6}$$

where x^* is de-fuzzified output, μ_{DY} is the aggregated resultant membership function of the all output fuzzy sets, and x is the universe of discourse.

Fuzzification gives the membership function which is in range between 0-1 and does not give a clear cut idea to regarding the problems output or is not understandable to human hence the defuzzification. The output has to be numeric or crisp. It is desired that the output, to come up with a single crisp value or numeric values from the FIS. i.e. to predict the result one should come up with a crisp number to tell what was the final year score for the student . This crisp number is obtained in a process known as defuzzification. The method used in this paper is centre of gravity for defuzzifying.

Below is the resultant of the output of rule fired for the first student in the list or database.

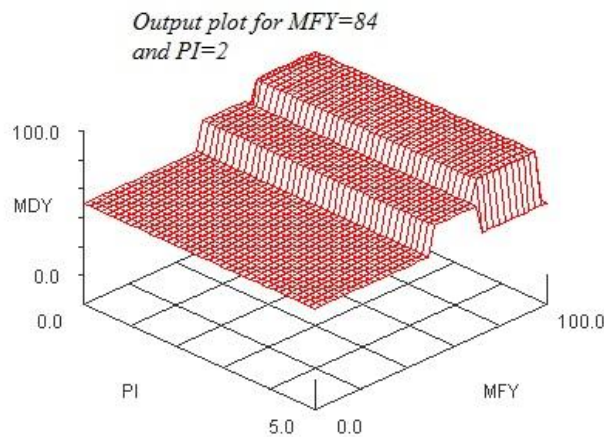


Fig 6: Resultant surface of the fuzzy model of Students Performance Predictor with inference rules as in last rule given in the snapshot of rule list.i.e IF FIRST.YR =D & PI=M THEN FINAL.YR = D

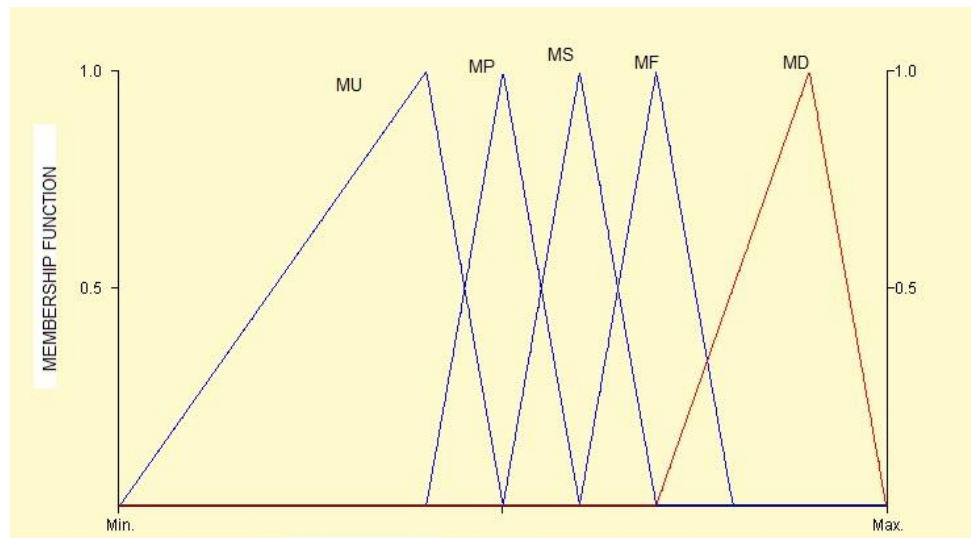


Fig 7.Aggregation of the result obtained for Rule FY=D and PI=M →DY=D

VI. Results

The primary experiment was done on the set of 55 students database . for few of the the results were very near to actual but for some they were at a difference of 10-15. The table given below depict the snapshot of few results taken.

HSC	GRAD	PI	Atten	Dist. Trav	First. Yr	Predicted results	Actual results	Error
64	72	200000	89	1	84	78.72	79.33	-0.61
65	69	245764	73	2	73	83.98	72.44	11.54
57	62	30000	80	1	67	81.7	67.16	14.54
56	68	100000	97	2	69	69.99	65.78	4.21
46	73	96000	73	2	45	57.82	58.84	-1.02
70	70	100000	75	1	72	86.37	70.53	15.84
42	66	50000	81	1	58	67.58	63.02	4.56
50	56	150000	87	1	65	69.995	66.58	3.42
71	62	100000	87	1	80	86.1	72.80	13.30
47	60	50000	87	1	63	77.345	63.02	14.33
54	67	100000	87	1	60	69.99	62.13	7.86
62	59	200000	60	1	57	64.54	63.91	0.63
68	66	200000	80	2	70	84.89	65.56	19.33
64	69	150000	80	2	62	75.58	68.80	6.78
47	58	75000	87	1	57	69.99	68.93	1.06
60	60	200000	67	0	66	74.6	62.18	12.42
70	57	100000	93	0	66	80.52	69.51	11.01
62	78	250000	93	0	71	72.36	71.33	1.03
52	71	90000	67	0	64	74.49	68.49	6.00
69	57	15000	100	0	68	69.99	66.40	3.59

VII. Conclusion

The choice of fuzzy implication rule is very important while designing a fuzzy predictor system. A technique to produce fuzzy rules based on the ID3 approach and to get optimized results with greater accuracy was the aim of this paper. We applied the ID3 Derived rules on the students database available with the instructors to predict their final years marks. The presented technique provides us with: a Simple way to produce meaningful fuzzy rules; a method to fuzzification of ID3-derived rules to deal with fuzzy data; A method for defuzzification so as to get the output in the human understandable form. ID3 being the simple decision tree generate huge rulebase which has to read into our fuzzy inference system . Many of the rules are generating repetitive information which can be discarded. Hence as per the [14] the pruned decision tree can be used. The rules generated by the ID3 can be optimized by pruning of the tree like C4.8, or WEKA J48 can be used and hence we may get better accuracy.

References

Journal Papers:

- [1]. B.K. Bhardwaj and S. Pal, Data Mining: A prediction for performance improvement using classification, International Journal of Computer Science and Information Security, Vol. 9(4), 2011.
- [2]. V. Juana-Maria and F. Manuel, How does one assess the accuracy of academic success predictors? ROC analysis applied to university entrance factors, International Journal of Mathematical Education in Science and Technology, Vol. 39 (3), 2008, pp 325–340.
- [3]. Emerson, T and Taylor, B., "Comparing Student Achievement Across Experimental and Lecture-Oriented Sections of a Principles of Microeconomics Course," Southern Economic Journal 703:672-693 (2004).
- [4]. Cindy P. Veenstra, "A Strategy for Improving Freshman College Retention" , Journal for Quality and Participation, ASQ, January 2009.
- [5]. P.H. Zhu, "Learning factors for technical school students," Taipei college of business journal, 12, pp.81-94, 2007.
- [6]. Waikato ML Group. User Manual Weka: The Waikato Environment for Knowledge Analysis. Department of Computer Science,
- [7]. University of Waikato (New Zealand), June 1997
- [8]. K.A. Rasmani and Q. Shen, Data-driven fuzzy rule generation and its application for student academic performance evaluation; Springer, Applied Intelligence, 25, 2006, 305–319.
- [9]. F. Herrera and E. Herrera-Viedma, Linguistics decision analysis: Steps for solving decision problem under linguistics information, Fuzzy sets and system management, 115, 2000, 67-82.
- [10]. H. Hellendoorn and C. Thomas, (1993) "Defuzzification in fuzzy controllers," J.Intell. Fuzzy Syst., Vol. 1, pp. 109–123
- [11]. Agrawal, R., Imielinski, T., Swami, A., "Database Mining: A Performance Perspective", IEEE Transactions on Knowledge and Data Engineering, pp. 914-925, December 1993

Books:

- [12]. George J. Klir and Tina Folger, Fuzzy sets uncertainty and information, Eastern Economy edition PHI, 2003, PP 254 – 260
- [13]. S.Rajasekaran and G.A. vijayalakshmi pai , "Neural networks, fuzzy Logic and Genetic Algorithms-Synthesis and Application" , PHI, 210-216
- [14]. Tan P.-N., Steinbach M. and Kumar V., Introduction to Data Mining, Addison Wesley, 2006.
- [15]. T. J. Ross, Fuzzy Logic with Engineering Applications, MacGraw-Hill, Inc, 1995.

Proceedings Papers:

- [16]. L. T. Koczy: Fuzzy If .. Then Rule Models and Their Transformation Into One Another, IEEE Trans. Systems, Man, and Cybernetics – Part A, vol. 26, pp. 621 – 637, 1996.

Websites :

- [17]. <http://www.cs.waikato.ac.nz/ml/weka/>
- [18]. H.J. Walberg, "Improving educational productivity," Philadelphia: Laboratory for School Success. Retrieved April 2005, from <http://www.temple.edu/lss/pdf/publications/pubs2003-I.pdf>, 2003