Fuzzy Decision Tree to Predict Student Success in Their Studies

Heri Bambang Santoso[a]

[a]Departement of Computer Science, Universitas Pakuan, Bogor,16143 Indonesia

Corresponding author e-mail address: heri.bambang@unpak.ac.id

Abstract

The number of students graduating on time is one of the important aspects in the assessment of accreditation of a university. But the problem is still a lot of students who exceed the target time of graduation. Therefore, the prediction of graduation on time can serve as an early warning for the university management to prepare strategies related to the prevention of cases of drop out. The purpose of this research is to build a model using fuzzy decision tree to form the classification rules are used to predict the success of a student's study using fuzzy inference system. Results of this study was generated model of the number of classification rules are 28 rules when the value θr is 98% and θn is 3%, with the level of accuracy is 95.85%. Accuracy of Fuzzy ID3 algorithm is higher than ID3 algorithms in predicting the timely graduation of students.

Keywords: fuzzy decision tree, fuzzy inference system, fuzzy ID3, ID3

1. Introduction

The quality of a university than can be seen by the average length of its graduates get a job, can also be seen by the average length of study students. The study program is obliged to monitor the progress of students study. Prediction graduate on time can serve as an early warning to the performance of students study. Furthermore, the overall prediction can be used as a reference in evaluating the educational process, curriculum, and other matters relating to education. To make a prediction can be done in various ways, one of which can be done by using a data mining techniques. Kwik Kian Gie School of Business has a dataset at the Academic Information System that has not been fully utilized. It is unfortunate if the dataset is so large is not used for mining the information contained therein.

Classification is one of the methods in data mining to determine the class of a record label in the data (Han and Kamber, 2006). Classification techniques are the focus of this research is a decision tree. In the decision tree method, if the attributes used continuous type, then it should be done discretization to divide the range of values on the attributes using the point of intersection, where the cut points used will determine the value of a domain with clear boundaries that may occur misclassification. Determination of the cutoff point is crucial. The concept of fuzzy logic is an alternative to state something that cannot be defined precisely. In the fuzzy set, the role of the degree of membership as a determinant of the presence of elements in a set is very important.

Some research in the field of university academic has been done by using decision tree classification techniques. Some of research on the classification and evaluation of student performance using decision tree algorithm C4.5 and Naive Bayesian (Vasani and Gawali, 2014). Applying algorithms ID3 and C4.5 to predict the performance of students in the first semester (Adhatrao et al., 2013). Applying decision tree algorithm ID3, C4.5 and CART to predict the performance of students majoring in engineering (Yadav and Pal, 2012). Predict the performance of students Purvachal University in India using decision tree algorithm ID3, C4.5, CART (Yadav et al., 2012). Fuzzy inference systems to predict student learning achievement based on the national exam, a test of academic potential, and learning motivation (Mustafidah and Aryanto, 2012).

Some research related to the application of the technique of fuzzy decision tree is the research on fuzzy decision tree for software effort estimation (Idri and Elyassami, 2011). Fuzzy decision tree for evaluation of employee performance (Li et al., 2012). Fuzzy ID3 Algorithm for Effective Prediction of Bankruptcy based on the Qualitative factors (Martin et al., 2012). Fuzzy decision tree using the data iris (Yun et al., 2014).

Based on the research that has been done using decision tree classification techniques, this research will be done graduation data classification using decision tree to predict the timely graduation of students by using fuzzy inference system (Mustafidah and Aryanto, 2012). However, in studies using fuzzy inference system still uses its own set of rules established or not established based on the data so that it can potentially cause the resulting level of accuracy is not the maximum. So in this research will build a model student graduation data classification using fuzzy decision tree method. Results of the rules of fuzzy decision tree classification is then used in the fuzzy inference system as a classification rule base that will be used to predict the timely graduation of students. Results from this research can be used by the university management, to provide treatment and warning against students who are expected to graduate is not timely.

2. Materials and Methods

2.1. Materials

Data obtained from the graduation data of Kwik Kian Gie School of Business in 2015, majoring in accounting and management from 2008 to 2010. The data is a combination of several tables, which comes from the dataset student profiles, student transcripts, and the presence of students. The data used is the data that is already clean or does not contain the missing value, relevant and not redundant. The result of the merger of the three datasets, generating 410 records and 6 attributes relevant to the graduation on time. These attributes include GPA (Grade Point Average) 1st semester, 2nd semester GPA, values of pancasila, the value of citizenship, attendance, and the average value of two high school classes. Attribute values of pancasila and citizenship values will be taken the average of both values and serve as a new attribute with the name attribute behavior. As for the attributes of attendance will represent the discipline of a student then this attribute will be named as attributes discipline. So that the final result obtained five attributes among 1st semester GPA, 2nd semester GPA, Discipline, Behavior, and Raport. As for the target class is categorized into two categories: passed on time ($\leq$ 48 months of the study period) and category not passed on time (study period> 48 months).

2.2. Methods

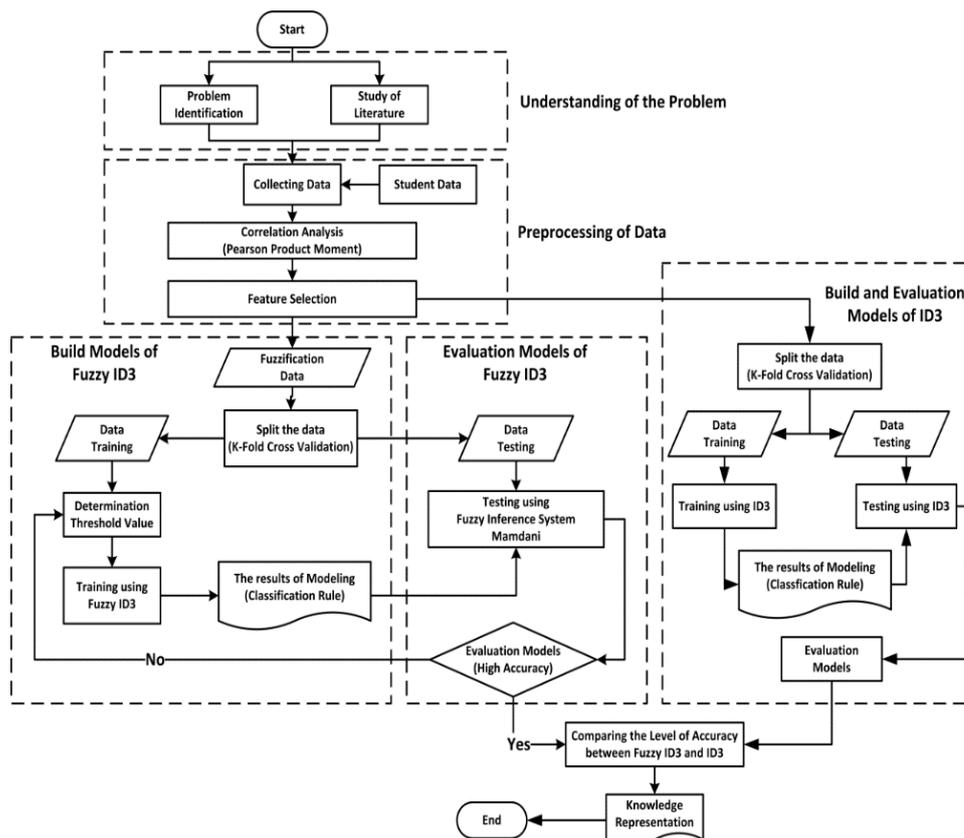Framework in this study can be described in a flow chart as can be seen in Figure 1.

Figure 1: Research Framework

2.2.1. Machine Learning

2.2.1.1. Fuzzy ID3 Algorithm

Fuzzy ID3 algorithm is an efficient algorithm to create a fuzzy decision tree. The algorithm of fuzzy ID3 is as follows (Liang, 2005; Mujiarto et al., 2019):

a. Create a *Root* node that has a set of fuzzy data with membership value 1.
b. If a node t with a fuzzy set of data D satisfies the following conditions, then it is a leaf node and assigned by the class name.
   - The proportion of class C_k is greater than or equal to θ_r,
   $$\frac{|D^{Ci}|}{D} \geq \theta_n \tag{1}$$
   - The number of a data set is less than θ_n
   - There are no attributes for more classifications
c. If a node D does no satisfy the above conditions, then it is not a leaf-node. And an new sub-node is generated as follow:
   - For A_i's ($i=1,\dots,L$) calculate the information gain, and select the test attribute A_{max} that maximizes them.
   - Devide D into fuzzy subset $D_1,\dots,D_m$ according to A_{max}, where the membership value of the data in D_j is the product of the membership value in D and the value of $F_{max,j}$ of the value of A_{max} in D.
   - Generate new node $t_i,\dots,t_m$ for fuzzy subsets $D_1,\dots,D_m$ and label the fuzzy sets $F_{max,j}$ to edges that connect between the nodes t_j and t.
   - Replace D by D_j ($j=1,2,\dots,m$) and repeat from 2 recursively.

2.2.1.2. Fuzzy Entropy and Information Gain

Information gain is a statistical value that is used to select attributes that will expand the tree and produce a new node on ID3 algorithm. An entropy is used to define the value of information gain. Entropy is defined as follows:

$$H_s(s) = \sum_i^N -p_i * log_2(P_i) \tag{2}$$

With P_i is the ratio of class C_i on the set of examples $S = \{X_1, X_2,\dots, X_k\}$.

$$P_i = \frac{\sum_{j=1}^k x_j \in C_i}{S} \tag{3}$$

There are two special cases that occur in the classification boolean, the first is if all the members of the set S has the same type, then the value of entropy is 0 (zero). This means it does not happen uncertainties classification.

$$H_s(S) = -1 * log_2(1) - 0 * log_2(0) = 0 \tag{4}$$

Second, if the number of positive examples is equal to the number of negative examples, the entropy value is 1 (one), it indicates the maximum classification uncertainties.

$$H_s(S) = -0.5 * log_2(0.5) - 0.5 * log_2(0.5) = 1 \tag{5}$$

To expand the attributes, which are based on data from the set of instances, must first be defined standard measure information gain. Information gain is used as a measure of selection of attributes, which is the result of a reduction in entropy of the set of sample after dividing the sample size set by the number of attributes. Information gain for attributes *A* is defined as follows:

$$G(S, A) = H(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} H(S_v) \tag{6}$$

With a weight $W_i = \frac{|S_v|}{|S|}$ is the ratio of the data on a set of *v* attributes example. In the fuzzy data set, there is an adjustment formula to calculate the entropy values for attributes and information gain due to the data expression fuzzy. Here is the equation to find the value of fuzzy entropy of all the data:

$$H_f(s) = H_s(S) = \sum_{i}^{N} -P_i * log_2(P_i) \tag{7}$$

To determine the fuzzy entropy and information gain of an attribute *A* on fuzzy ID3 algorithm (FID3) used the following equation:

$$H_f(S, A) = -\sum_{i-1}^{c} \frac{\sum_{j}^{N} \mu_{ij}}{S} log_2 \frac{\sum_{j}^{N} \mu_{ij}}{S} \tag{8}$$

$$G_f(S) = H_f(S) - \sum_{v \subseteq A}^{N} \frac{|S_v|}{|S|} * H_f(S_v, A) \tag{9}$$

With μ_{ij} is the value of membership of a pattern to-*j* for the i-th class. $H_f(S)$ shows the entropy of the set S of training data on the node. $|S_v|$ is the size of the subset $S_v \subseteq$ S of training data X_j with attributes *v*. $|S|$ indicates the size of the set *S* (Liang, 2005).

2.2.1.3. Fuzziness Control Threshold and Leaf Decision Threshold

Determine fuzziness control threshold (θr) and leaf decision threshold (θn) to get the best model and with high accuracy (> 80%) for predicting student graduation on time. For example, if the threshold value used for θr θn 80% and 20%, but still produces a low model accuracy or less than 80% then conducted another experiment combination threshold to achieve higher accuracy.

If the learning process of a fuzzy decision tree is stopped until all the data sample on each leaf-node becomes a member of a class it will produce low accuracy. Therefore, to improve accuracy, the process of learning to be terminated early or tree pruning in general (Liang, 2005).

As a result, two thresholds are defined (Umano *et al.*, 1994).

- *Fuzziness control threshold (FCT) / θr*
  If the proportion of the data set of class Ck is greater than or equal to the threshold value θr, the expansion of tree stopped.
- *Leaf decision threshold (LDT) / θn*
  If the number of members of the data set at a node is smaller than the threshold θn, then the tree expansion is stopped.

2.2.1.4. ID3 Algorithm

ID3 algorithm is a decision-making algorithm most widely used because it is easy to use and effectiveness (Liang, 2005). ID3 algorithm developed by Ross Quinlan (Quinlan, 1986). ID3 algorithm will do the greedy search for all possible decision tree. To determine the root node and other attributes, ID3 calculating entropy values to get the information gain. The more diverse set of sample data, the greater the entropy value.

Entropy values were in the range of 0 to 1. The formula calculating the entropy values are as follows:

$$H(S) = \sum_{i}^{n} -P_i log_2(P_i)$$
(10)

Where H(S) is the entropy value of the data sample S. N is the number of classes on the attributes, while P_i is the number of samples for class i or the ratio of class. After searching the value of entropy, then the next can be calculated the value of information gain. Information gain is a measure of the effectiveness of an attribute or parameter to classify data.

To calculate it can be used the following formula:

$$G(S, A) = H(S) - \sum_{v \in Value(A)} \left| \frac{S_v}{S} \right| H(S_v)$$
(11)

Where *G(S,A)* is the information gain A of the sample *S. H(S)* is the value of the entropy *S. Values(A)* is the set of possible values for the attribute A. $|S_v|$ is the number of samples to the value *v.* $|S|$ is the total number of samples. $H(S_v)$ is the entropy for samples that have value *v.*

3. Results and Discussion

3.1. Attribute Correlation Analysis

Based on correlation test was done to all the attributes using the Pearson product moment correlation results obtained index value of each attribute is summarized in Table 1. As seen in Table 1, all attributes have a negative correlation to graduate on time. 2nd semester GPA attribute has the highest correlation with coefficient -0.74 (74%). Negative correlation has a meaning that the higher the value of an attribute predictor of the long study period covered will be the smaller or faster. These attributes are then used in this study.

Table 1: Attribute correlation test using Pearson Product Moment

	Correlations					
	1st Semester GPA	2nd Semester GPA	Behavioral	Discipline	Raport	Graduate on time
1st Semester GPA	1	0.901**	0.776**	0.320**	0.716**	-0.664**
2nd Semester GPA	0.901**	1	0.711**	0.404**	0.792**	**-0.740****
Behavioral	0.776**	0.711**	1	0.258**	0.594**	-0.583**
Discipline	0.320**	0.404**	0.258**	1	0.354**	-0.687**
Raport	0.716**	0.792**	0.594**	0.354**	1	-0.645**
Graduate on time	-0.664**	-0.740**	-0.583**	-0.687**	-0.645**	1

**. Correlation is significant at the 0.01 level (2-tailed).

3.2. Fuzzyfication Data

In this research used data mining techniques are fuzzy decision tree, therefore the data used must be represented in the form of fuzzy. The process begins with a membership function of each attribute is used.

3.2.1. GPA attributes

Attribute 1st Semester GPA and 2nd Semester GPA is divided into 4 groups or linguistic terms, namely low ($x <=$ 2.75), medium ($2.65 <= x <= 3$), high ($2.9 <= x <= 3.6$), and very high ($3.5 <= x <= 4$). Fuzzy set for each linguistic term use curves with a trapezoidal shape as shown in Figure 2.

$$\mu_{Low}(x) = \begin{cases} 1 & ; x \leq 2.65 \\ \dfrac{2.75 - x}{0.1} & ; 2.65 \leq x\ 2.75 \\ 0 & ; x \geq 2.75 \end{cases}$$

$$\mu_{VeryHigh}(x) = \begin{cases} 0 & ; x \leq 3.50 \\ \dfrac{x - 3.50}{0.1} & ; 3.50 \leq x \leq 3.60 \\ 1 & ; x \geq 3.60 \end{cases}$$

$$\mu_{Medium}(x) = \begin{cases} 0 & ; x \leq 2.65 \\ \dfrac{x - 2.65}{0.1} & ; 2.65 \leq x \leq 2.75 \\ 1 & ; 2.75 \leq x \leq 2.90 \\ \dfrac{3.00 - x}{0.1} & ; 2.90 \leq x \leq 3.00 \\ 0 & ; x \geq 3.00 \end{cases}$$

$$\mu_{High}(x) = \begin{cases} 0 & ; x \leq 2.90 \\ \dfrac{x - 2.90}{0.1} & ; 2.90 \leq x \leq 3.00 \\ 1 & ; 3.00 \leq x \leq 3.50 \\ \dfrac{3.60 - x}{0.1} & ; 3.50 \leq x \leq 3.60 \\ 0 & ; x \geq 3.60 \end{cases}$$
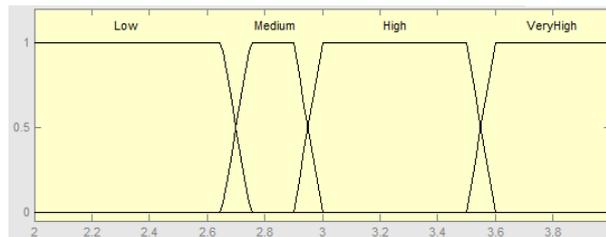

Figure 2: Fuzzy set of GPA attributes

3.2.2. Raport attributes

Attribute raport are divided into 4 groups or linguistic terms, namely low ($x <= 7$), medium ($6.75 <= x <= 7.5$), high ($7.25 <= x <= 8.25$), and very high ($8 <= x <= 10$). Fuzzy set for each linguistic term use curves with a trapezoidal shape as shown in Figure 3.

$$\mu_{Low}(x) = \begin{cases} 1 & ; x \leq 6.75 \\ \dfrac{7.00 - x}{0.25} & ; 6.75 \leq x\ 7.00 \\ 0 & ; x \geq 7.00 \end{cases}$$

$$\mu_{Medium}(x) = \begin{cases} 0 & ; x \leq 6.75 \\ \dfrac{x - 6.75}{0.25} & ; 6.75 \leq x\ 7.00 \\ 1 & ; 7.00 \leq x \leq 7.25 \\ \dfrac{7.50 - x}{0.25} & ; 7.25 \leq x \leq 7.50 \\ 0 & ; x \geq 7.50 \end{cases}$$

$$\mu_{High}(x) = \begin{cases} 0 & ; x \leq 7.25 \\ \dfrac{x - 7.25}{0.25} & ; 7.25 \leq x \leq 7.50 \\ 1 & ; 7.50 \leq x \leq 8.00 \\ \dfrac{8.25 - x}{0.25} & ; 8.00 \leq x \leq 8.25 \\ 0 & ; x \geq 8.25 \end{cases}$$

$$\mu_{VeryHight}(x) = \begin{cases} 0 & ; x \leq 8.00 \\ \dfrac{x - 8.00}{0.25} & ; 8.00 \leq x \leq 8.25 \\ 1 & ; x \geq 8.25 \end{cases}$$
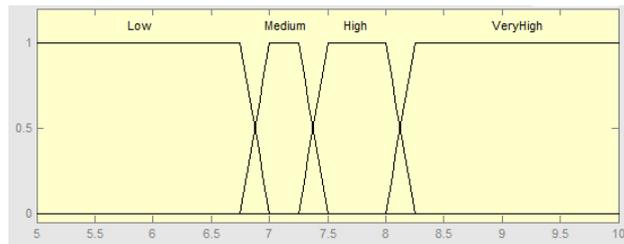
Figure 3: Fuzzy set of raport attributes

3.2.3. Behavioral attributes

Behavior attributes were divided into 4 groups or linguistic terms, ie less ($x <= 65$), medium ($63 <= x <= 70$), good ($68 <= x <= 80$), and excellent ($78 <= x <= 100$). Fuzzy set for each linguistic term use curves with a trapezoidal shape as shown in Figure 4.

$$\mu_{Less}(x) = \begin{cases} 1 & ; \ x \leq 63 \\ \dfrac{65 - x}{2} & ; \ 63 \leq x \ 65 \\ 0 & ; \ x \geq 65 \end{cases}$$

$$\mu_{Medium}(x) = \begin{cases} 0 & ; \ x \leq 63 \\ \dfrac{x - 63}{2} & ; \ 63 \leq x \leq 65 \\ 1 & ; \ 65 \leq x \leq 68 \\ \dfrac{70 - x}{2} & ; \ 68 \leq x \leq 70 \\ 0 & ; \ x \geq 70 \end{cases}$$

$$\mu_{Good}(x) = \begin{cases} 0 & ; \ x \leq 68 \\ \dfrac{x - 68}{2} & ; \ 68 \leq x \leq 70 \\ 1 & ; \ 70 \leq x \leq 78 \\ \dfrac{80 - x}{2} & ; \ 78 \leq x \leq 80 \\ 0 & ; \ x \geq 80 \end{cases}$$

$$\mu_{Excellent}(x) = \begin{cases} 0 & ; \ x \leq 78 \\ \dfrac{x - 78}{2} & ; \ 78 \leq x \leq 80 \\ 1 & ; \ x \geq 80 \end{cases}$$
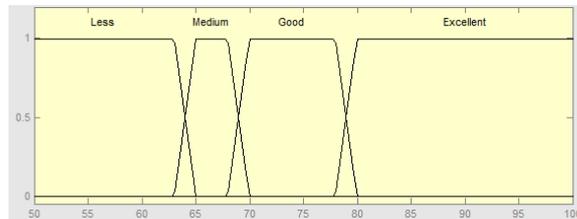

Figure 4: Fuzzy set of behavioral attributes

3.2.4. Discipline attributes

Discipline attributes are divided into 3 groups or linguistic terms, ie less ($x <= 180$ present), medium (175 present $<= x <= 190$ present), and good (185 present $<= x <= 196$ present). Fuzzy set for each linguistic term use curves with a trapezoidal shape as shown in Figure 5.

$$\mu_{Less}(x) = \begin{cases} 1 & ; \ x \leq 175 \\ \dfrac{180 - x}{5} & ; \ 175 \leq x \ 180 \\ 0 & ; \ x \geq 180 \end{cases}$$

$$\mu_{Medium}(x) = \begin{cases} 0 & ; \ x \leq 175 \\ \dfrac{x - 175}{5} & ; \ 175 \leq x \leq 180 \\ 1 & ; \ 180 \leq x \leq 185 \\ \dfrac{190 - x}{5} & ; \ 185 \leq x \leq 190 \\ 0 & ; \ x \geq 190 \end{cases}$$

$$\mu_{Good}(x) = \begin{cases} 0 & ; \ x \leq 185 \\ \dfrac{x - 185}{5} & ; \ 185 \leq x \leq 190 \\ 1 & ; \ x \geq 190 \end{cases}$$
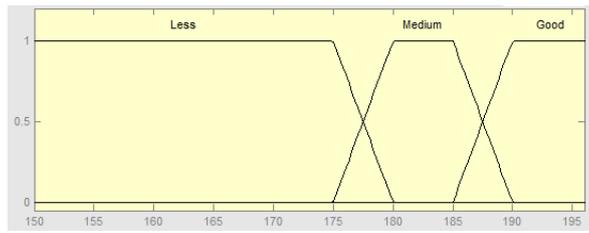
Figure 5: Fuzzy set of discipline attributes

3.2.5. Attribute of Target Class

Attributes graduate on time will be referred to as Class, represented by two linguistic variables are "graduate on time" and "not graduating on time". Both linguistic terms are defined as follows: a) graduate on time = 1 (≤ 48 months of the study period), b) not graduating on time = 2 (study period> 48 months).

3.3. Results of Formation Model of Fuzzy ID3

Data from the transformation that has divided using 10-fold cross validation, given the same treatment by Fuzzy ID3 algorithm to perform the training process. The training process is done 240 times. For each training set, the training process is done 24 times, by changing the value θr 6 times, namely 75%, 80%, 85%, 90%, 95%, and 98%, and for each of the same value θr, given the value θn different, namely 3%, 5%, 8%, and 10%. The average number of rules generated by the training process and the total execution time can be seen in Table 2 and Table 3.

Table 2: Average number of classification rules of Fuzzy ID3

FCT (θr)	LDT (θn)			
	3%	**5%**	**8%**	**10%**
75%	12	9	7	7
80%	16	12	9	8
85%	17	12	9	9
90%	18	12	10	9
95%	22	16	13	12
98%	25	19	14	13

Table 3: Average total time of formation of the model of Fuzzy ID3 (in seconds)

FCT (θr)	LDT (θn)			
	3%	**5%**	**8%**	**10%**
75%	0.458	0.376	0.341	0.323
80%	0.514	0.412	0.368	0.327
85%	0.523	0.408	0.366	0.343
90%	0.649	0.494	0.414	0.387
95%	0.923	0.741	0.504	0.439
98%	0.922	0.742	0.501	0.455

The values of θr and θn are used in Table 2 are selected based on the results of the experiment, because with these values, the number of rules generated significant changes. In Table 2 and Table 3 shows that the most significant improvement occurred when the value of θr is increased from 90% to 95%. This condition is caused due to expansion during the first training set is done, many sub-node that the proportion of one class has reached a value above 90%, so that the sub-node does not need to be expanded further. In Table 2 shows that the higher the value of θr it will cause the number of rules generated is also rising, this is because before a node is dominated by a class and the proportion of the class above or equal to the value of θr then the tree will continue to be expanded. The opposite occurs if the value of θn the higher the resulting rules tends to decrease.
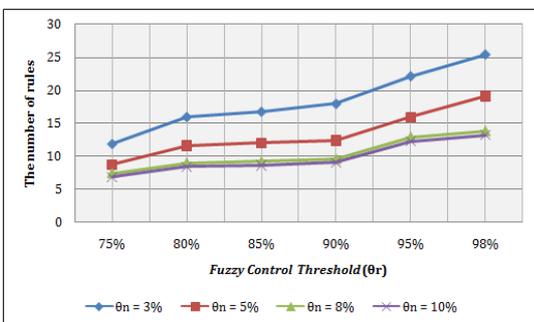

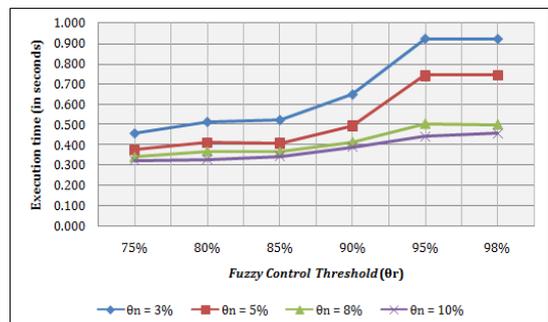**Figure 6:** Comparison of average number of rules of Fuzzy ID3

Figure 7: Comparison of average total time of formation model of Fuzzy ID3

In Figure 7 shows that the higher value of θr so that the longer the execution time to form a decision tree. It is inversely linearly with the number of rules that are formed as shown in Figure 6, that the more the number of rules that forms the execution time to form a decision tree is also increasing.

3.4. Testing Model of Fuzzy ID3

To measure the accuracy of the model generated in the training phase, the process of testing carried out 240 times of testing. Testing process is done by inserting a rule which is derived from the training process into a FIS Mamdani to determine the class of each record in the testing set. For one training process done once the testing process.

3.5. Performance Evaluation of Fuzzy ID3

Performance evaluation of Fuzzy ID3 algorithm can be determined by calculating the average accuracy of the entire process of testing at 10 different sets of testing. Evaluation of the performance of Fuzzy ID3 on the value θr and θn different, can be seen in Table 4 and Figure 4.

Table 4. Performance evaluation of *Fuzzy* ID3

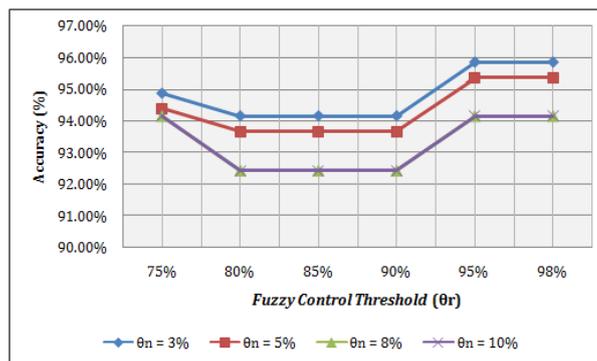| FCT (θr)	LDT (θn)			
	3%	5%	8%	10%
75%	94.88%	94.39%	94.15%	94.15%
80%	94.15%	93.66%	92.44%	92.44%
85%	94.15%	93.66%	92.44%	92.44%
90%	94.15%	93.66%	92.44%	92.44%
95%	95.85%	95.37%	94.15%	94.15%
98%	95.85%	95.37%	94.15%	94.15%

Figure 8: Comparison of average level accuracy of Fuzzy ID3

From Table 4 and Figure 8 it can be seen that the performance of Fuzzy ID3 increased if the value θr greater or smaller value θn. By looking at the average value of accuracy of testing results, it appears that the level of accuracy is decreased when the value θr increased from 75% to 80% and accuracy increased significantly when increased from 90% to 95%. The average level of accuracy at the time of the value θn is 8% and 10% have the same accuracy rate changes. The highest accuracy value produced during θr value is 95% and 98% by value θn is 3% which is the average value of accuracy was 95.85%.

3.6. Performance Comparison of Fuzzy ID3 and ID3

At this stage, the analysis of testing of fuzzy decision tree algorithm using fuzzy ID3 algorithm and decision tree using ID3 algorithm. Data used in the ID3 algorithm is data graduation same with fuzzy ID3 algorithm that 410 data records. Training method used is the same that is using 10-fold cross validation. Before the training, these data should be transformed into several categories. The number of categories and range of data on these attributes equated with the transformation of fuzzy ID3. Detail data transformation can be seen in Table 5.

Table 5: Transformation of data attributes for ID3 algorithm

Attributes	Data Type	Range Value of Data
1st Semester GPA	Nominal	Low = 2.00 – 2.65
		Medium = 2.66 – 3.00
		High = 3.01 – 3.50
		Very High = 3.51 – 4.00
2nd Semester GPA	Nominal	Low = 2.00 – 2.65
		Medium = 2.66 – 3.00
		High = 3.01 – 3.50
		Very High = 3.51 – 4.00
Behavioral	Nominal	Less = 56 - 63
		Medium = 64 - 70
		Good = 71 - 80
		Excellent = 81 – 100
Discipline	Nominal	Less = 154 - 175
		Medium = 176 - 190
		Good = 191 – 196
Raport	Nominal	Low = 6.00 - 6.75
		Medium = 6.76 – 7.5
		High = 7.6 – 8.25
		Very High = 8.26 – 10.00

This is done to determine whether using a fuzzy decision tree in the form of a decision tree models would be better if compared with the decision tree method. Table 6 below are the results of performance comparison between fuzzy ID3 and ID3.

Table 6: Performance comparison between Fuzzy ID3 and ID3

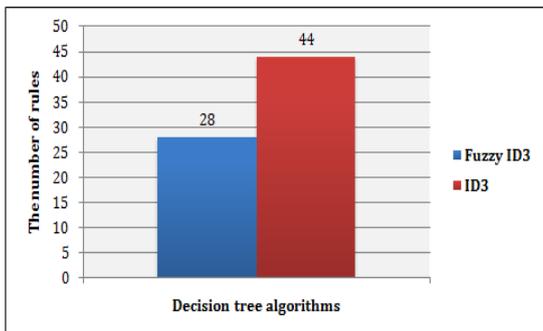| Criteria	Algorithms	
	Fuzzy ID3	ID3
Number of rules	28	44
Accuracy	95.85%	93.41%
Error	4.15%	6.59%

Figure 9: Comparison of the number of rules of Fuzzy ID3 and ID3
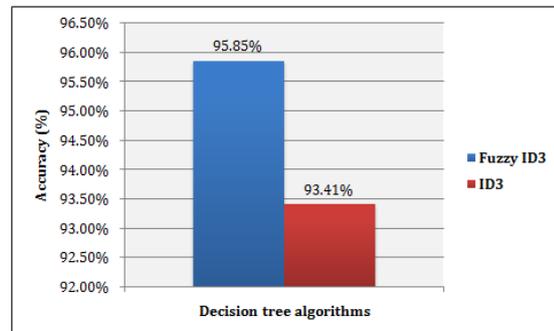

Figure 10: Comparison of the accuracy between Fuzzy ID3 and ID3

In Figure 9 and Figure 10 can be seen that the number of rules generated by the fuzzy ID3 algorithm less when compared with the ID3 algorithm. But fuzzy ID3 algorithm has a higher accuracy of the ID3 algorithm. From the results of this experiment, it can be concluded that the model generated by an algorithm that uses fuzzy approach is better when compared to the decision tree algorithm which does not use fuzzy approach in terms of classification and prediction study period students.

4. Conclussion

This research has successfully developed a model based on the data graduation by using fuzzy ID3 algorithm to establish the rules of classification that is used to predict the timely graduation of students using FIS Mamdani. The best number of classification rules are generated an average of 25 rules when the value of θr is 98% and θn is 3%,

with the level of accuracy is 95.85%. The higher the value of θr and the lower the value of θn then the accuracy will be higher. Based on the classification rules are established, the most decisive factor one can graduate on time is 2nd Semester GPA. The model produced by Fuzzy ID3 algorithm has better accuracy rate than the ID3 algorithm. In addition, this study also produce a simple application that can be used to create a classifier based on fuzzy ID3 algorithm using Matlab. In the next study the formation process decision tree can use the genetic algorithm to optimize fuzzy decision tree (FDT) in order to obtain genetically optimized fuzzy decision tree (G-FDT) in predicting timely graduation of students in an effort to improve the accuracy of the model is obtained.

Acknowledgments

References

Adhatrao, K., Gavkar, A., Dhawan, A., Jha, R., & Honrao, V. (2013). Predicting students' performance using ID3 and C4. 5 classification algorithms. *International Journal of Data Mining & Knowledge Management Process, 3*(5), 39-52.

Elyassami, S., & Idri, A. (2011). Applying fuzzy ID3 decision tree for software effort estimation. *International Journal of Computer Science Issues*, *8*(4), 131-138.

Han, J and Kamber, M. (2006). *Data Mining Concepts and Techniques*. San Fransisco: Morgan Kaufmann Publisher.

Liang, G., van den Berg, J., and Kaymak, U. (2005). *A comparative study of three Decision Tree algorithms: ID3, Fuzzy ID3 and Probabilistic Fuzzy ID3.* Netherlands: Erasmus University.

Li, Y., Jiang, D., and Li, F. (2012). The application of generating fuzzy ID3 algorithm in performance evaluation. *Procedia Engineering*, *29*, 229-234.

Martin, A., Balaji, S., & Venkatesan, V. P. (2012). Effective prediction of bankruptcy based on the qualitative factors using FID3 algorithm. *International Journal of Computer Applications*, *43*(21), 28-32.

Mujiarto, Komaro, M., Mohamed, M. A., Rahayu, D. S., Sanjaya, W. S. M., Mamat, M., & Sambas, A. (2019). Subiyanto,"Colored object detection using 5 Dof robot arm based adaptive Neuro-Fuzzy method,". *Indonesian Journal of Electrical Engineering and Computer Science*, *13*(1), 293-299.

Mustafidah, H., and Aryanto, D. (2012). Fuzzy Inference Systems to Predict Student Learning Achievement Based on the National Exam, a Test of Academic Potential, and Learning Motivation, *JUITA*, *2*(1), 1-7.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*(1), 81-106.

Umanol, M., Okamoto, H., Hatono, I., Tamura, H. I. R. O. Y. U. K. I., Kawachi, F., Umedzu, S., and Kinoshita, J. (1994, June). Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. In *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference* (pp. 2113-2118). IEEE.

Vasani, V. P., & Gawali, R. D. (2014). Classification and performance evaluation using data mining algorithms. *International Journal of Innovative Research in Science, Engineering and Technology*, *3*(3), 10453-10458.

Yadav, S. K., Bharadwaj, B., and Pal, S. (2012). Data mining applications: A comparative study for predicting student's performance, *International Journal of Innovative Technology & Creative Engineering*, *1*(12), 13-19.

Yadav, S. K., and Pal, S. (2012). Data mining: A prediction for performance improvement of engineering students using classification, *World of Computer Science and Information Technology Journal, 2*(2), 51-56.

Yun, J, Seo, J. W and Yoon, T. (2014). Fuzzy Decision Tree. *International Journal of Fuzzy Logic System*, *4*(3), 7-11.