

The Application of Fuzzy-Rough Set Decision Tree for Credit Rating

Soetrisno^{a,*}, Grizelda Wahyuningtyas^a, Imam Mukhlash^a

^a Institut Teknologi Sepuluh Nopember, Faculty of Mathematics and Sciences, Department of Mathematics, Jl. Arief Rahman Hakim, Surabaya, Indonesia, 60111

* Corressponding author email address: soetrisno@matematika.its.ac.id

Abstract

Fuzzy decision tree is a data mining method which is a combination of fuzzy logic and decision tree. Integration of fuzzy logic concept in the decision tree intended to represent an uncertain condition and a very complex model. Construction of fuzzy decision tree using fuzzy rough techniques was done by looking under the value and significance levels for each factor to be analyzed. The problem discussed is to predict the potential success of a prospective customer credit through fuzzy decision tree by using the history data of existing credit customers. Parameters used are the amount of the credit, loan, mortgage interest (rate), customer turn over, and the long passage of the customer's business. From the simulation results, it is obtained a fuzzy decision tree model with an accuracy of 83%. With this application, a decision maker can determine the potential of prospective customers and prevent the occurrence of credit fail.

Keywords: Credit Scoring, Credit Risk, Fuzzy Rough Sets, Fuzzy Decision Tree, Data Mining

1. Introduction

The development of modern industrial technology has been growing rapidly create abundant provision of public goods. That way the people are easily attracted to consume goods with many choices available, according to the needs of every person. For middle and upper class society it does not have a significant effect, but for the people of the middle and lower income is only enough for daily needs, it affects their financial arrangements. Lower middle class people tend to buy goods, vehicles, and houses on credit. The Survey of Bank Indonesia (BI) showed that 74.7% of consumers use the facilities of mortgage to buy property (05/14/2008).

Many Banks provide credit facilities to its customers. Credit application can be submitted by the applicant to meet the requirements specified by the Bank. In lending activities common problem of bad loans or credit fail caused by the failure of partial repayment of loans to borrowers. This problem can actually be solved, one of them can be solved by identifying and predicting customers well before granting a loan by looking at the historical data of the loan.

The actual amounts of data can be a source of historical data that useful to find a pattern and new knowledge that can be useful in the future. In finding of new patterns or knowledge from these data requires a data mining methods that can solve the problem of "big data, the less information". Currently, data mining was developed in various business sectors such as marketing and banking (Han et al., 2011).

Credit rating (credit scoring) is the most common method used to evaluate the credit worthiness of credit applicants with regard attributes such as age, income, and marital status. The goal is to classify credit applicants into two classes based on the likelihood of their payment capabilities into a good applicant who tend to pay their credit and bad applicants are rejected because they failed to pay the financial obligations. Many methods have been used by Banks and financial institutions to increase the accuracy, and the most popular method is the method of credit scoring. In previous research method used was a soft computing and data mining method to perform credit scoring to help a credit analyst to find pattern in

the data set and predict the output using the techniques and computational tools (Lahsasna et al., 2010). One of the popular task in data mining to perform credit rating is decision tree method.

In identifying credit customers, previous studies often use decision tree methods as assessed quickly and effectively. In the study by (Yogi et al., 2009) the credit scoring models, are made of customer information such as income, type of mortgage, down payment, loan term, savings account, age, telephone and electricity bills, are processed using as many as 7 rules (rule), but the level of accuracy obtained is 79.57% . A study by (Wahyu, D. and Widyanto, M.R, 2011) concluded that the fuzzy decision tree method is better than the decision tree method that in this research method to be used is a fuzzy decision tree method.

Fuzzy decision tree is one of the methods for classifying data into several parts and can also be used to determine the order in which the desired corresponding data. In this paper, we developed data mining methods for credit scoring in the evaluation stage (application scoring). Application scoring can be used as a screening tool to separate the good and the bad applicants. Using of data mining methods is expected to help the credit evaluation process conducted by the Bank.

According to (Shang and Qiang, 2008), the use of rough-fuzzy features selection mechanism allows the reduction for a low dimensionality features sets from sampels descriptions. The Rough-Set has recently emerged as another major mathematical approach for managing uncertainty that arises from inexact, noisy, or incomplete information. It is found to be particularly effective in the area of knowledge reduction (Petrosino and Salvi, 2006).

(Feng and Song, 2010) conducted a study using fuzzy rough sets to predict stock prices, get strong provisions of the securities markets and the relative economic data. The results showed that use of fuzzy rough sets and data mining can make more effective prediction outcomes. The application of rough fuzzy artificial neural network showed a great ability to generalize, to identify behavior patterns, and to allow the creation of an inference mechanism in high complex systems (Affonso et al., 2015). (Chen and Cheng, 2013) use hybrid

model based on rough set classifier for setting credit rating decision rule in banking industry. This research concludes that this model can reduce the number of attributes and optimize the number of factors in financial data.

The objective of this paper is to apply fuzzy rough set decision tree method to predict customer's credit rating of a micro credit of a bank and learn the advantages and disadvantages of this approach.

2. Basic Concept and Related Research

2.1. Credit Scoring

Credit is a payment mechanism that is very common in the community. A main function of credit is meets the needs of service to the community in order to facilitate trade, production and consumption of services even all of which are aimed at improving human welfare. Institution of credit providers and recipients should work together to achieve the desired goal. The receiving party has the purpose of obtaining credit financing sources easily and quickly. Party lender has the objective to benefit from the interest paid by the debtor. Besides the advantages that can be obtained, creditors should bear the risk of failed payment by the debtor. Creditor must evaluate each applicant's credit before acceptance or rejection decision is taken. Because the amount of the loan application, the creditor must be able to evaluate loan applicant with objective, accurate, and consistent. Evaluation can be aided by credit scoring. Isac defines credit scoring as a tool that involves the use of statistical models to evaluate all available information with the objective of making the credit decision. Credit ratings (credit scoring) is the most common technique used to evaluate the credit worthiness of credit applicants with all the attributes that they have. The benefit to be gained from the application of credit scoring is to increase the speed and consistency of the loan application process and enables the automation of the lending process (Lahsasna et al., 2010).

Credit scoring models built using a sample of past credit in large numbers. The sample is divided into two classes: good credit (payment is made on time) and the troubled credit (payment is made on time or cannot make a payment). Based on past patterns, a combination of borrower characteristics that distinguish good and bad borrowers produce value that represent risk estimation of each new borrower. Many of the techniques that can help in the development of credit scoring models are Support Vector Machine (SVM) and decision tree technique. Decision tree technique has become a popular technique because the resulting decision tree is easily interpreted and visualized (Zhai, 2011).

2.2. Data Mining

Data is a component that cannot be separated from daily operations in a company. Every day, a large amount of data generated by the company's operations. An organization can be flooded with various kinds of data. It is useless data collected and stored when the data did not use. The problem that arises is how to process the data so that it can display knowledges that are useful to management in making decisions. In the process of knowledge discovery can use data mining algorithms to discover useful patterns. Data mining is defined as a process that uses a variety of data analysis tools to discover patterns and relationships of data that can be used to

make precise predictions. Stages of the process of knowledge discovery is data cleansing, data integration, data selection, data transformation, data mining and pattern evaluation found.

Data mining is a process of finding patterns using statistical and mathematical techniques of vast numbers of records that can provide benefits for decision makers (Han et al., 2011). Data mining helps companies to obtain a pattern of data that is stored in the company database. The acquired knowledge will guide you in taking measures of business as the maintenance and improvement of the competitive level of the company's business.

2.3. Decision Tree

Decision tree is an approach that is very popular and practical in machine learning to solve classification problems. The concept of the decision tree is basically converts the data into a hierarchy and decision tree rules which on further developments can be simplified by eliminating branches or rules are not necessary. While accompanying attributes declare a parameter data created as a criterion in the formation of the tree. One attribute is an attribute that declares the completion of data per data item referred to classification. The attribute has the value that are called the instance (Wahyu and Widyanto, 2011). Suppose instance attributes have income in the form of low, medium, and high.

Decision tree consists of nodes that form a tree rooted, which means that the tree is directed by a node called the root which does not have the input. All other nodes have exactly one input. A node with no output is called internal nodes. And all in addition to the nodes are called leaf or often known as a decision node. For each leaf, a decision has been made and applied to all observations in the leaf. The Leaf is the end result of grouping data from decision tree classification divides into groups.

2.4. Fuzzy Information System

A fuzzy information system (*FIS*) by Wang et al. in 2001 refers to the four tuple information systems, ie $(U, A \cup C, V, f)$ where $U = \{x_1, x_2, \dots, x_n\}$ is the set of a finite object is not empty and for each and for each x_i represented as $x_i = \{a_{i1}, a_{i2}, \dots, a_{im}\}$. Where A is a finite set of fuzzy conditional attributes. At *FIS*, noted $A = \{a_1, a_2, \dots, a_n\}$ where $A_i (1 \leq i \leq n)$ represents a conditional attribute consisting of a set of fuzzy linguistic term $FLT_i\{A_{i1}, A_{i2}, \dots, A_{in}\} (1 \leq i \leq n)$. C symbolizes a fuzzy decision attribute with a set of $FLT_C\{C_1, C_2, \dots, C_m\}$. Each *FLT* are assumed to be a fuzzy set in U .

2.5. Fuzzy Set Theory

Here are given two definitions that support the theory of fuzzy.

Definition 2.1 Fuzzy Set. Given a crisp universe of the X , fuzzy set A (more precisely, a fuzzy subset of X) is given by the membership function $\mu_A(x): X \rightarrow [0,1]$, and the value $\mu_A(x)$ is interpreted as the degree of member x in fuzzy set A . The group of all fuzzy subsets of X is denoted as $F(X)$.

Definition 2.2 Fuzzy Numbers. A fuzzy number A is an unnormalized convexity ($\sup_x \mu_A(x) = 1$) exceeds the fuzzy set of real numbers with a continuous membership function that has more than one mean value $x_0 \in \mathbb{R}, \mu_A(x_0) = 1$.

Subsets provide a basic introduction to the theory of fuzzy sets. Fuzzy set theory allows an object to have a partial member of more than one set. This is done through the introduction of a function known as a member function (membership function), where the mapping of the complete set of an object X in a set that is known as a member space. The definition of a fuzzy set is

If X is a collection of objects generally denoted by x and then a fuzzy set A in X is an ordered set:

$$A = \{(x, \mu_A(x)) | x \in X\} \quad (1)$$

$\mu_A(x)$ is called the membership function of x in A which maps X to the member space M .

2.6. Fuzzy Rough Set

Fuzzy rough sets developed by *D. Dubois* integrate the concept of vagueness (equity) and something that cannot be distinguished well. Fuzzy and fuzzy rough set have used in data classification, for example to meteorological data in (Winda and Mukhlash, 2014).

Definition 2.3 Fuzzy Equivalence Classes. Suppose U is the universe, R is a fuzzy equivalence relation of U . Fuzzy equivalence class, $[x]_R$ defined by

$$\mu_{[x]_R} = \mu_{R(x,y)}$$

Definition 2.4 Fuzzy Approximation. Suppose U is the universe, X and P are two fuzzy sets in U , U/P is a fuzzy partition fuzzy of U . For a given $x \in U$, fuzzy approximation P -lower dan fuzzy approximation P -upper of X defined by (Jensen and Shen, 2005).

$$\mu_{P-x}(x) = (\mu_F(x), \inf_{y \in U} \max \{1 - \mu_F(y), \mu_F(x)\}) \quad (2)$$

$$\mu_{\bar{P}x}(x) = \sup_{F \in \frac{U}{P}} \min(\mu_F(x), \sup_{y \in U} \min\{1 - \mu_F(y), \mu_F(x)\}) \quad (3)$$

Tuple $(PX, \bar{P}X)$ is called fuzzy rough set.

2.7. Fuzzy Decision Tree

Decision tree technique has been shown to independently interpret problems and can solve the problem on a large scale, but this technique is also known as the classification technique with a high level of instability with respect to interference in training data and way of presenting the data with a high variance. Fuzzy theory can increase endurance when performing classification of new cases in a decision tree. Fuzzy logic led to improvements in this aspect because the elasticity of fuzzy sets. The method aim shave been studied in detail and crisp been compared with alternative method and the results showed much improvement from accuracy level of prediction results, shown with much reduced variance models. In addition, fuzzy logic is also more stable at a level better than parameters interpreted. Fuzzy decision tree based on fuzzy rough techniques is a new criterion based on a meeting between fuzzy conditional attributes with attribute fuzzy decision to select attributes that will be expanded. Fuzzy conditional attributes will be selected as the most important attribute to be expanded. For a given FIS, every fuzzy conditional attribute has a different contribution to the fuzzy decision attribute, relationship

interests of a fuzzy attribute to other fuzzy attributes defined by Zhai, 2011.

Definition 2.5 Fuzzy Attribute Significance

$$\mu_{POS_P Q} = \sup_{x \in \frac{U}{Q}} \mu_{P_x}(x) \quad (4)$$

In the literature (Jensen and Shen, 2005), the above equation represents the degree of membership of an object x in the fuzzy positive region.

Definition 2.6 Significance Fuzzy with linkages to other attributes. Let P and Q are two fuzzy attributes in a given FIS, the relationship of importance (significance) of P with connection on Q defined by:

$$\tau_P(Q) = \frac{\sum_{x \in U} \mu_{POS_P(Q)}(x)}{|U|} \quad (5)$$

Where $|U|$ is the number of the universe U .

As in (Zhai, 2011), the algorithm to generate fuzzy rough decision tree model as below:

Step 1: Prepare FIS from the value of membership function of each attribute.

Step 2: Choose the attribute that will be expanded.

Step 2.1: For every conditional fuzzy attribute A_i and every fuzzy linguistic term A_{iki} ($1 \leq i \leq n$), significance from A_i that corresponded to fuzzy decision attribute was calculated using Eq. (5).

Step 2.2 : Choose the highest value of i_o using equation:

$$i_o = \text{Argmax}_{1 \leq i \leq n} \{ \tau_{A_i}(C) \}$$

A_{i_o} is an expanded attribute.

Step 3: If the goal condition is not satisfied, the partition U , and choose attribute recursively that is expanded until a fuzzy decision tree successfully generated.

Step 4: Extract the classification of fuzzy rules from fuzzy decision tree T .

3. Design and Implementation ASI System

3.1. Design of Credit Scoring Applications

In the design of the system in this study, we carried out the design data and design process.

3.1.1. Design Data

In accordance with the issues raised in this study, there are five conditional attribute data to be used, i.e ceiling, period, rate, turnover, and company age. As for the decision attribute, there are two decisions, namely succeeded and failed.

3.1.2. Design Process

In the design process of the development of credit scoring application involves java programming and calling the data in the database. General application design is shown in Fig. 1.

3.2. Implementation and Application

The following describes the implementation of the processes required program as described in the design process into the algorithm. The process of the process is as follows:

1. The process of formation of Fuzzy Information System.

2. The process of calculating the value of the relationship of importance (significance) between attributes
3. The prediction process customer credit.

The flow chart depicted in Fig.1 describes the process of preprocessing the data so that it becomes *FIS* and fuzzy decision tree construction process based on fuzzy rough technique.

This flow chart describes the process of preprocessing the data so that it becomes *FIS* and fuzzy decision tree construction process based on fuzzy rough technique. After the model of fuzzy decision tree model is obtained, then the model will be tested to determine the level of accuracy.

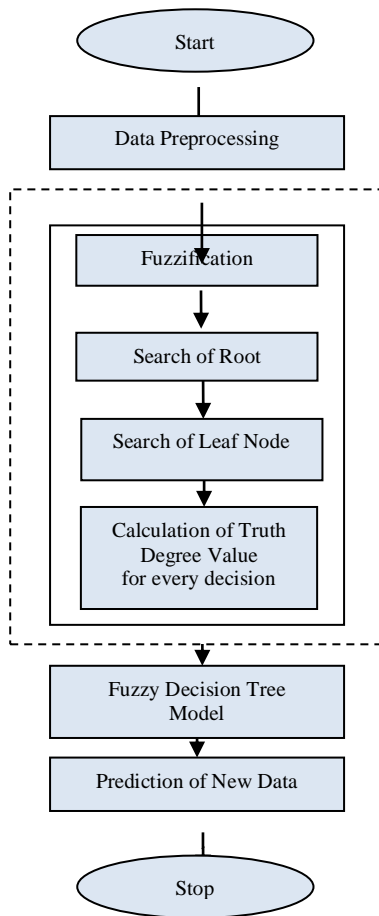


Fig. 1. The design of general application

4. Testing and Results

4.1. Environmental testing applications

The device used in the test system consists of multiple hardware and software. The hardware used is a computer with Intel (R) Core (TM) i7-2630 CPU @ 2:00 GHz QM 2:00 GHz RAM 8 GB, and 750 GB Hard Disk. While the software used is the Operating System Windows 7 Ultimate 64-bit software and Net-Beans IDE 7.0.1 for system development and open source software MySQL 5.5 as database management system.

4.2. Input data of application testing

Data input of the application form that includes the number of data the customer, maximum credit, period, interest rate, turnover, and company age. The data used as inputs 1000 of data the customer. Views from some of the data input is shown in Fig. 2.

NO	PLAFOND	J.K WAKTU	RATE	OMZET	LAMA USA	KOLEKTI
148	10000000	12	12.0	2000000	12	0
147	7000000	24	12.0	2000000	12	0
146	20000000	24	12.0	2000000	12	0
145	20000000	24	12.0	2000000	12	0
144	10000000	24	12.0	2000000	8	0
143	13000000	24	12.0	1800000	7	0
142	10000000	18	12.0	1800000	6	0
141	10000000	24	12.0	1750000	36	0
140	10000000	24	12.0	1750000	24	0
139	10000000	24	12.0	1750000	18	0
138	8000000	24	12.0	1750000	12	0
137	10000000	24	12.0	1750000	10	0
136	20000000	24	12.0	1750000	10	0
135	15000000	24	12.0	1750000	8	0
134	7000000	24	12.0	1750000	6	0
133	5000000	12	12.0	1500000	36	0

Fig. 2. The display of input data of application testing.

4.3. Steps of the application testing

The following shows a test of the system already built. The test is done by displaying the login page, the home page, the page appearance of the data, the data page that has been converted into the *FIS*, the root of the search page, search page leaf node, page fuzzy decision tree models, accuracy test page, and the page prediction prospective customers.

1. The process of change in the initial data in the form of *FIS*

In this process sought membership function of the initial data using trapezoidal fuzzy equations.

Alpha is used to obtain the desired degree of truth by the user. Large degree of truth is 1 - alpha.

2. Test the root search page

On this page will look for the value of the greatest significance between fuzzy conditional attributes with fuzzy decision attribute using Eq. (5).

4.4. Results of the application testing

The results of the application testing based on test data is

$$\tau_{max. credit}(C) = \frac{\sum_{x \in U} \mu_{POS_{max. credit}(C)}(x)}{|U|} = 0,8480$$

$$\tau_{time period}(C) = \frac{\sum_{x \in U} \mu_{POS_{time period}(C)}(x)}{|U|} = 0,7638$$

$$\tau_{rate}(C) = \frac{\sum_{x \in U} \mu_{POS_{rate}(C)}(x)}{|U|} = 0,7427$$

$$\tau_{omzet}(C) = \frac{\sum_{x \in U} \mu_{POS_{omzet}(C)}(x)}{|U|} = 0,8284$$

$$\tau_{company age}(C) = \frac{\sum_{x \in U} \mu_{POS_{company age}(C)}(x)}{|U|} = 0,8181$$

And conditional attributes are selected as the root is a maximum credit, because the value of the maximum credit is the greatest significance.

4.4.1. Determining the leaf node

At this stage, this algorithm look for the value of significance among other attributes with branches. The results of the calculation of significance for the low branches are:

$$\begin{aligned} \tau_{time\ period}(maximum\ credit\ is\ low) &= 0.5097 \\ \tau_{rate}(maximum\ credit\ is\ low) &= 0.4749 \\ \tau_{omzet}(maximum\ credit\ is\ low) &= 0.6193 \\ \tau_{company\ age}(maximum\ credit\ is\ low) &= 0.5290 \end{aligned}$$

Because of the highest value is omzet attribute, then the omzet chosen as the leaf node. The results of calculations for the standard branch are:

$$\begin{aligned} \tau_{time\ period}(maximum\ credit\ is\ standard) &= 0.1373 \\ \tau_{rate}(maximum\ credit\ is\ standard) &= 0.2746 \\ \tau_{omzet}(maximum\ credit\ is\ standard) &= 0.4120 \\ \tau_{company\ age}(maximum\ credit\ is\ standard) &= 0.5493 \end{aligned}$$

The highest value is in the high attributes of company age, but great value attribute significance high operating in a high branch is greater. Likewise with turnover attributes. So as the third highest value, attribute Rate elected as a leaf node. The results of calculations for the high branches are:

$$\begin{aligned} \tau_{time\ period}(high\ maximum\ credit) &= 0.1966 \\ \tau_{rate}(high\ maximum\ credit) &= 0.3933 \\ \tau_{omzet}(high\ maximum\ credit) &= 0.5900 \\ \tau_{company\ age}(high\ maximum\ credit) &= 0.7866 \end{aligned}$$

Because the highest value on the high attributes of company age, the high company age was elected as the leaf nodes.

4.4.2. Test stage of tree models results

On this stage, it is shown that fuzzy decision tree models generated with degrees of truth, assuming that $\beta = 0.75$. Generated model of fuzzy decision tree is shown in Fig. 3.



Fig. 3. Fuzzy decision tree model.

And the rules of inference that used by the system as below:

- Rule 1.** IF max. credit IS Low AND omzet IS Low THEN Fail
- Rule 2.** IF max. credit IS Low AND omzet IS Standard THEN Success
- Rule 3.** IF max. credit IS Low AND omzet IS High THEN Success
- Rule 4.** IF max. credit IS Standard AND rate IS Low THEN Success
- Rule 5.** IF max. credit IS Standard AND rate IS High THEN Success
- Rule 6.** IF max. credit IS High AND company age IS Low THEN Fail
- Rule 7.** IF max. credit IS High AND company age IS High THEN Success

4.2.3. Test the level of accuracy of the model

On this page, it is shown the comparison of prediction using fuzzy decision tree models that have been established with the reality of data.

In this system, the input data is data the customer and a predicted outcome is a decision that is successful if the customer deserve to be given a loan or fail if the customer is not feasible given the loan.

To the 200 testing data, it was obtained that 166 of data is successfully predicted correctly by the system (look at Table 1), and the average level of accuracy that is obtained based on the testing data is 83%.

Table 1

The results of accuracy test.

No.	Number of Training Data	Number of Test Data	Level of Accuracy
1	500	200	83,0%
2	600	200	86,0%
3	700	200	82,0%
4	800	200	84,5%

5. Conclusions

In this paper, we proposed an decision tree algorithm based on fuzzy rough set for credit rating prediction in a micro credit bank in Indonesia. The main steps of this approach are data preprocessing, fuzzification process, finding root and leaf nodes, calculation of truth degree value for every decision, fuzzy decision tree modeling, generate rules, and make prediction of new data with decision rules formed.

Based on the analysis of the results of system testing conclude that this system has been successfully classify success or fail of new borrower with the degree of truth. This method is able to produce a rule that has an accuracy rate of 83% for testing data. The advantage of this method is that this method produce relatively high accuration of prediction. However, the limited of testing data causes the capability of this method to predict future condition needs further research.

References

- Affonso, C., Sassi, R.J., and Barreiros, R.M. (2015). Biological Image Classification using Rough-Fuzzy Artificial Neural Network. *Expert Systems with Applications*. vol. 000, pp. 1-7.
- Aprianti, W. and Mukhlash, I. (2014), The Application of Rough Set and Fuzzy Rough Set based Algorithm to Classify incomplete Meteorological Data, *Proceedings of International Conference on Data and Software Engineering (ICoDSE)*, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7062674>
- Chen, Y. S., & Cheng, C. H. (2013). Hybrid models based on rough set classifiers for setting credit rating decision rules in the global banking industry. *Knowledge-Based Systems*, 39, 224-239.
- Han, J. and Kamber, M. (2011). *Data Mining Concepts and Techniques*. Morgan Kaufman: USA.
- Hullermeier, E. and Vanderlooy, S. (2010). Why Fuzzy Decision Trees are Good Rankers. *Fuzzy Systems paper in IEEE*.
- Lahsasna, A., Ainon, R.N. and Wah, T.Y. (2010). Credit Scoring Models Using Soft Computing Methods: A Survey. *The International Arab Journal of Information Technology*. 7(2).
- Olaru, C. and Wehenkel, L. (2003). A Complete Fuzzy Decision Tree Technique. *Journal of Fuzzy Sets and Systems*.
- Petrosino, S. and Salvi, G. (2006). Rough Fuzzy Set based Scale Space Transforms and their use in Image Analysis. *International Journal of Approximate Reasoning*. vol. 41, pp. 212-228.
- Shang, C. and Qiang, S. (2008). Aiding zction. In *Proceedings of the IEEE international conference on fuzzy systems*. pp. 976-982.
- Smith III and James, F. (2010). *Fuzzy Logic Resource Manager: Evolving Fuzzy Decision Tree Structure that Adapts in Real-Time*. Naval Research Laboratory. Code 5741. Washington, D.C.
- Wahyu, D. and Widyanto, M.R. (2011). Studi Perbandingan Metode Pohon Keputusan dan Pohon Keputusan Fuzzy pada Klasifikasi Penutup Lahan. *Jurnal Ilmiah KURSOR*. 6(1).
- Xiao-Feng, H., and Song-Song, L. (2010). Research on Predicting Stock Price by using Fuzzy Rough Set. In *Management Science and Engineering (ICMSE), International Conference on IEEE*, (pp. 1124-1130).
- Yogi Y.W., F. Rianti, P., Dan Viviane A.S. (2009). Evaluasi Pemohon Kredit Mobil Di PT "X" Dengan Menggunakan Teknik Data Mining Decision Tree. *Symposium Nasional RAPI VIII*.
- Zhai, J. H. (2011). Fuzzy Decision Tree based on Fuzzy-Rough Technique. *Soft Computing*, 15(6), 1087-1096.