

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