Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Economics and Finance 4 (2012) 406 – 412

Procedia Economics and Finance

www.elsevier.com/locate/procedia

International Conference on Small and Medium Enterprises Development with a Theme "Innovation and Sustainability in SME Development" (ICSMED 2012)

Assessing Credit Risk: an Application of Data Mining in a Rural Bank

I Gusti Ngurah Narindra Mandala[a], Catharina Badra Nawangpalupi[a]*, Fransiscus Rian Praktikto[a]

[a]*Industrial Engineering Department, Parahyangan Catholic University, Ciumbuleuit 94, Bandung 14141, Indonesia*

Abstract

Credit risk assessment for secured loans is an important operation in banking systems to ensure the lenders pay the loans on schedule and to classify the bank as a well performing bank due to regulation. This paper aims to identify factors which are necessary for a rural bank (Bank Perkreditan Rakyat) to assess credit application. By aiming on the reduction of number of non-performing loans, current decision criteria on credit risk assessment are evaluated. Subsequently, a decision tree model is proposed by applying data mining methodology.

The credit risk assessment model is applied to PT BPR X in Bali that had 1082 lenders (11.99%) who had non-performing loans and were identified as bad loan cases. This made PT BPR X was categorized as a poorly performing bank.

Data mining is used to suggest a decision tree model for credit assessment as it can indicate whether the request of lenders can be classified as performing or non-performing loans risk. Using C 5.0 methodology, a new decision tree model is generated. This model suggests that new criteria in analyzing the loan application. The evaluation results show that if this model is applied, PT BPR X can reduce non-performing loans to less than 5% and the bank can be classified as a well performing bank.

***Keywords*: rural bank; data mining; non-performing loans; decision tree; credit risk assessment.**

* Corresponding author. Tel.: +62 22 2032655.
E-mail address: katrin@unpar.ac.id.

doi:10.1016/S2212-5671(12)00355-3

1. Introduction

Rural bank or Bank Perkreditan Rakyat is a bank that has a specific business activity in serving mutual loans and saving for rural areas and micro small medium enterprises (MSME). The term of BPR referring to rural bank is used in this paper. Undang-undang no 10 1998 defined rural bank, opposed to general banks, as a bank that has business activities without offering services in payment transactions which follows either conventional way or syariah principles. It means that a rural bank must only focuses on mutual loan and saving activities.

Bank Indonesia has a regulation that operational areas of a BPR are limited to a certain province (Bank Indonesia 2011). Bank Indonesia, on its website, informs that in April 2012 there are 1,667 conventional BPR which have in total more than 12 millions customers (or accounts) and about a quarter of these accounts are mutual loan account and the rests are saving accounts. In Bali Province alone, there are 137 active BPR.

As a rural bank, BPR is ideally owned by local entrepreneurs because BPR's main purpose is to develop local economy. It is important that a BPR maintains its operations as a healthy bank to ensure its sustainability to support local businesses and economy. A well performed and healthy BPR should have the following indicators (Bank Indonesia, 2011):

1. CAR (Capital Adequacy Ratio) > 8%
2. Productivity – Non Performing Loans (NPL) < 5%
3. ROA (Return on Asset) > 1,3%
4. Cash Ratio > 4%
5. LDR (Loan to Deposit Ratio) 80% - 95%

Non performing loan (NPL) indicator becomes a crucial indicator for a BPR. A BPR have no make sure that it gives loans to verified and accountable lenders so that the bank would not risk having high NPL rate. To ensure the bank gives loans to right persons, credit assessment is a critical decision making process.

This paper aims to evaluate the decision tree that a BPR had made and to identify what criteria should be considered first. A new decision tree model is recommended using data mining approach. A case study is used in a BPR in Bali (PT BPR X). This bank has served their customers with various products, such as saving, time deposit, western union service, and mutual loans. This bank has a NPL indicator of 11,99% that was way above the determined range. As a result, the study is focused to reduce this NPL indicator.

2. Literature Review

Credit assessment

Credit assessment is a study to identify a feasibility of a loan application. It is performed to assess if a potential lender has business activities that are feasible, marketable, profitable and id the laon can be paid on time (Rivai 2006). Usually credit assessment is done by banking account officer which can be a part of an assessment committee. This assessment is done to analyse all factors involved in credit application, such as business financial performance and credit rating of the lenders. Rivai (2006) define 6C's analysis for credit assessment, which are:

1. Character. Character is the lender credit rating, which assess the willingness to pay and the ability to pay based on defined agreement. Character is the most important factor in credit assessment.
2. Capital. Capital is the amount of available initial fund of the lender. The higher the fund, the more the lender is considered to be able to repay the loan.
3. Capacity. Capacity is the ability of the lender to pay the annuity. It is evaluated by the income and the expected income during the period of the loan payment. The capacity is needed to assess the ability to pay from the lender.
4. Collateral. Collateral is lender's properties which should be given to the bank as a warrsnty for the loan. The collateral should have value similar to the loan applied by the lenders.

5. Condition of Economy. Condition of economy, which is political, social, cultural and economical situation that may cause the business sustainability of the lender.
6. Constraint. Constraint is all limitation and barriers that may make the business cannot continue the activities, such as regulation and resource scarcity.

Data mining

According to Gartner Group in Larose, data mining (2005) is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data sotred in repositories, using pattern recognition technologies as well as statistical and mathematical techniques.

There are six phases in standard procedure in performing data mining. Based on Cross – Industry Standard Process for Data Mining (CRISP – DM). Those are (Larose 2005),

1. Business understanding phase, which consist of the determination of business objectives, data mining problem definition, and data mining strategy preparation.
2. Data understanding phase, which consist of the collection of initial data, data exploration, and data evaluation.
3. Data preparation phase, which consists of data selection, data transformation, and data cleaning.
4. Modeling phase which consist of data modeling which main task is to model the data to represent the situation.
5. Evaluation phase, which evaluate the result of the model and to evaluate if the result is in line with the business objectives.
6. Deployment phase is an implementation phase.

Figure 1 shows the phases of the data mining standard process. Decision tree is a method in data mining which has association rules to process massive data. Decision tree identifies a collection of variables and associated rules that influence the decision making and analyze how the impact of these variables (Olshon & She 2007). Decision tree is a classification method that creates decision nodes connected with branches which consist of root node and leaf node where the root node is the top node of the decision tree. All attributes in decision node are tested and evaluared to show if they can generate another note, otherwise terminating node is generated (Larose 2005).

C5.0 algorithm is the newest version of decision tree methodwhich is the development version of ID3 (Iterative Dichotomiser 3) (Berry & Linoff 1997). This algorithm prunes decision tree model by identifying the error rate in each node and assumes the error rate at the predecessor node as the worst error. The errorr rates are then compared and the decision tree model chosen is the model that has the least error rate (Berry & Linoff 1997). C5.0 algorithm has a decision tree with unlimited nodes. Compared to other akgorithm in decision tree model (such as CART algorithm and heuristic models), this algorithm can provide more branches which resulted in the quicker process in generating nodes. Figure 2 shows the differences process of generating nodes between C.50 and CART algorithms.

3. Data Mining for Credit Assessment

As has been mentioned earlier, PT BPR X has a quite alarming performance wheren the NPL is much higher than 5% (11,99%). This is the result evaluation using 10 years of lenders data at PT BPR X. Given that this NPL rate is high, this research is carried out to evaluate the credit assessment criteria and to siggest better model to lower the NPL rate.

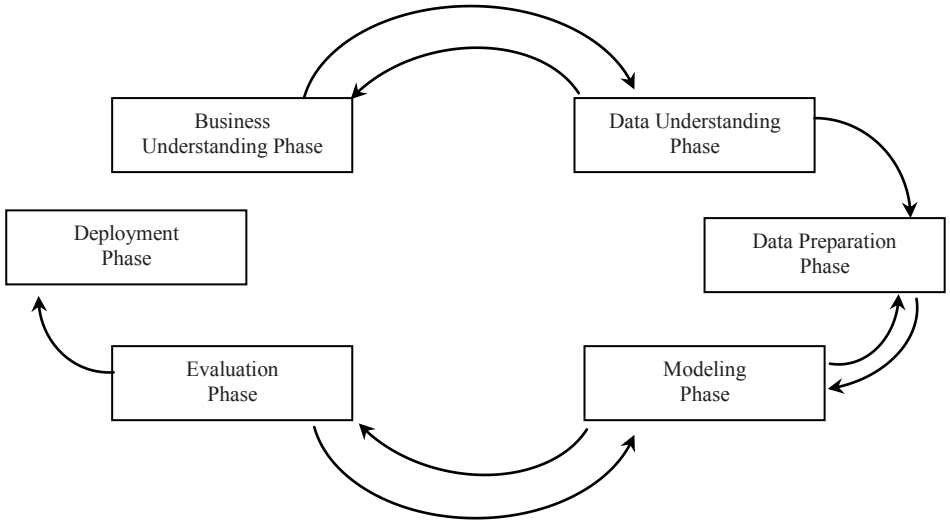

Fig. 1 CRISP – DM (Source: Larose 2005)

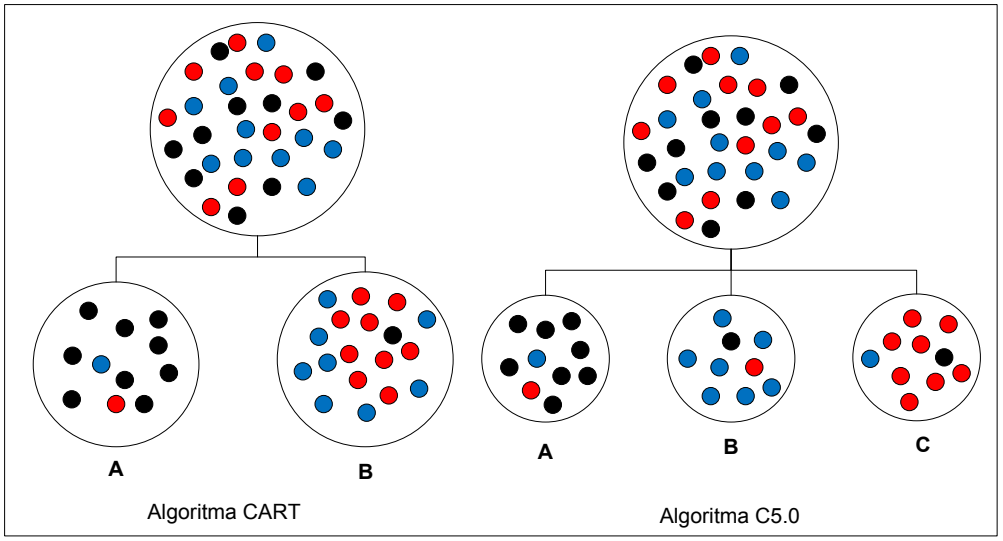

Fig. 2. CART and C5 algorithms

Credit assessment process at PT BPR has followed the procedure of submitting application, data verification, credit assessment process and approval or disapproval decision. The data used by PT BPR X for assessment are as follows:

1. Gender
2. Age
3. Credit amount
4. Monthly income
5. Expenses each month
6. Current payment per month

7. Saving (income – expenses – payment)
8. Collateral types
9. Collateral values
10. Loan period
11. Type of business activities
12. Source of funding
13. Previous credit status/rating

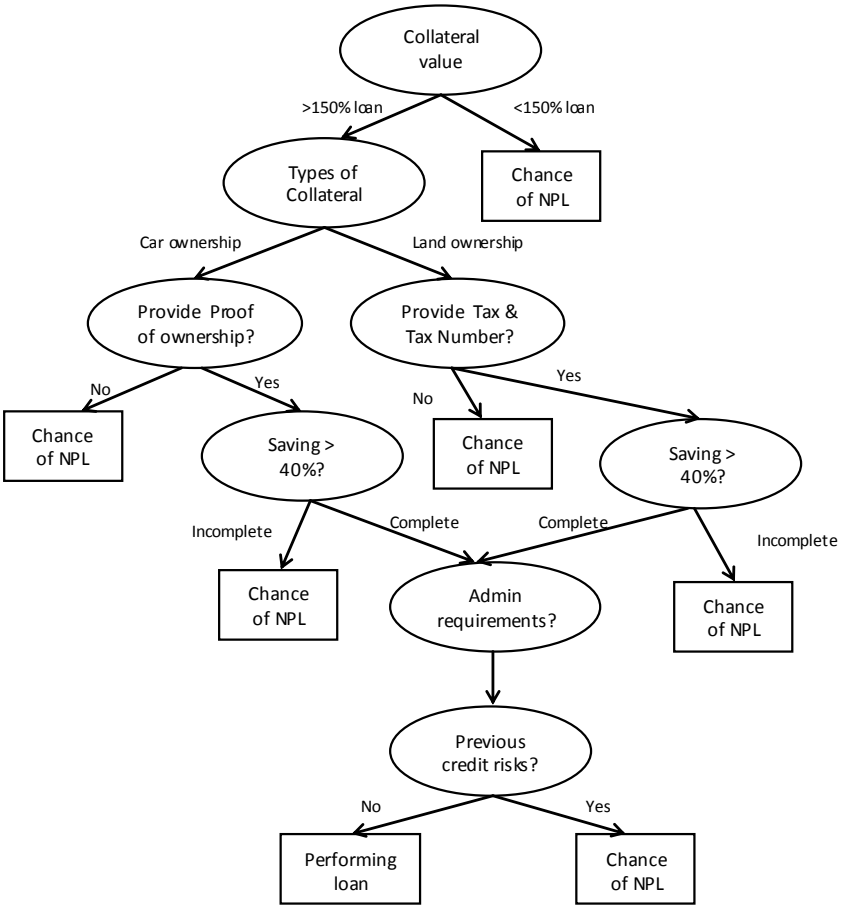

Figure 3: Credit assessment criteria

The current model of credit assessment is illustrated in Figure 3. This credit assessment criteria has collateral values (syarat agunan) as a root node, and the terminating node is previous credit rating (kritis credit sebelumnya). The other nodes are only types of collateral, completeness of required documents and saving. The current credit assessment model actually did not use all the attributes. Thus, credit assessment model is revised using data mining process. Algorithm C5.0 is used for creating new credit assessment model.

For modeling phase of the data mining process, there are 5 discrete/ non-continuous variables, which are gender, types of collateral, type of business activities, source of funding, credit status and the use of loan. There

are 8 continuous variables, which are age, monthly income, expenses each month, current payment per month, saving, collateral values, and loan period. Based on modeling phase, the decision tree is generated. This model can be seen in Figure 4.

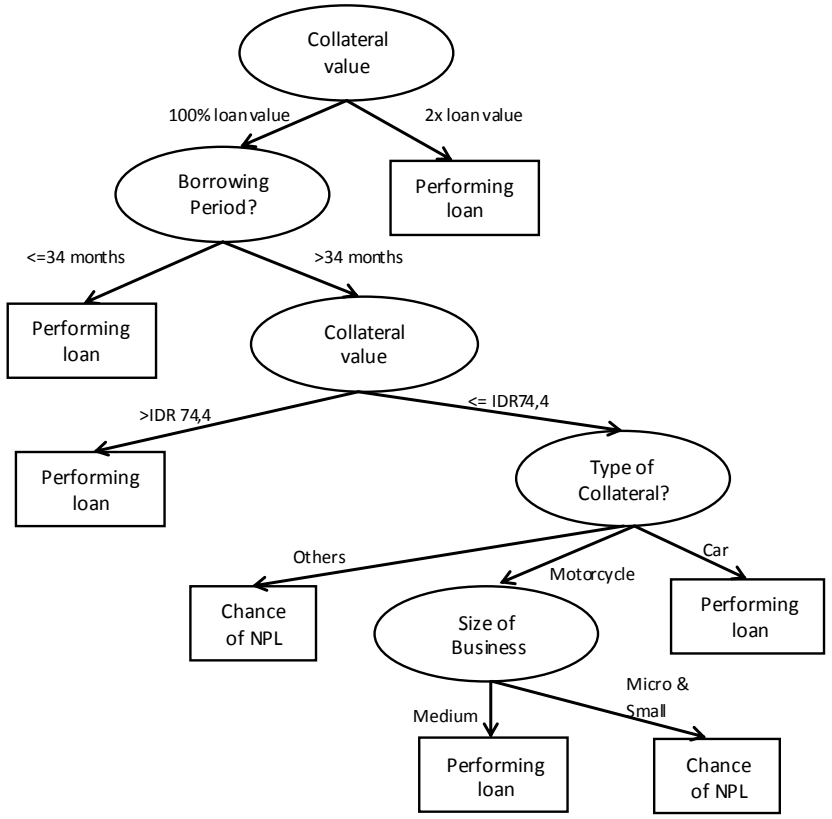

Figure 4: Proposed credit assessment criteria

This model show that collateral values is the root node and the followed leaf node are loan period and collateral values. The terminating node is type of collateral and type of business activities. After replicating this model using the data, it is evaluated that NPL rate become 3%, which is lower than the required rate and much lower than the current rate.

4. Discussions

Based on the model built, it is confirmed that collateral values is the most important criterion in credit assessment, shown by the collateral values is the root node. The proposed model has credit period as the first leaf node followed by the collateral values.

Although there are many other variables in credit assessment criteria, the model shows that not many of those variables are relevant for credit assessment criteria.

This model has been evaluated using the 84% data from 1028 data as the evaluation data. In terms of the values of loan, model has also improved the value of non-performing loans. Table 1 shows the comparison of the predicted performing loans based on current decision making and the proposed model from data mining

methods. The model has shown a good result, and PT BPR X is suggested to apply this model for their credit assessment process.

Table 1. Result compartison between current and proposed models

	Current (in thousand rupiahs)	**%**	**Proposed model (in thousand rupiahs)**	**%**	***Test* (in thousand rupiahs)**	**%**
Performing loan	36.996.600	88,01	36.531.700	94,14	10.519.900	96,75
Non performing loan	5.038.800	11,99	2.275.600	5,86	353.700	3,25
Total	42.035.400	100,00	38.807.300	100,00	10.873.600	100,00

References

Bank Indonesia, 2011. Model Bisnis Bank Perkreditan Rakyat: Referensi Mengelola BPR (Business Model of Bank Perkreditan Rakyat: A Reference of BPR Management), Direktorat Kredit, BPR dan UMKM Bank Indonesia.

Berry, Michael J.A. dan Linoff, Gordon S. 2004, *Data Mining Techniques, 1st ed.,* John Wiley & Sons Inc., Indianapolis, Indiana.

Berry, Michael J.A. dan Linoff, Gordon S. 1997, *Data Mining Techniques, 2nd ed.,* John Wiley & Sons Inc., Indianapolis, Indiana.

Larose, D.T. 2005, *Discovering Knowledge in Data,* John Wiley & Sons Inc., Hoboken, New Jersey.

Rivai, Veithzal dan Permata, Andria V. 2007, *Credit Management Handbook*: Teori, Konsep, Prosedur, dan Aplikasi Panduan Praktis Mahasiswa, Bankir, dan Nasabah (Credit Manegement Handbook: Theory, Concept, Procedures and Application - Practical Guide for Students Banks and Customers), PT RajaGrafindo Persada, Jakarta, Indonesia.

Suki, Ryoichi P. 2010, Definisi Bank (Bank Definition), [Online], Diakses dari: http://www.scribd.com/doc/11320386/Definisi-Bank [2011, 10, Februari].