

# Decision Supporting System of Granting Loans with Binary Logistic Regression

Samsul Pahmi <sup>a,1,\*</sup>, Ade Hudiana <sup>b,2</sup>, Lesri <sup>c,3</sup>, Santi Laswati Suryadi <sup>d,4</sup>, Lucas Cramer <sup>e,5</sup>, Bahadir Ozsut <sup>f,6</sup>

<sup>a</sup> Department Of Elementary Education Nusaputra University, Sukabumi, Indonesia

<sup>b,c,d</sup> Departement of Information System Nusaputra University, Sukabumi, Indonesia

<sup>e</sup> Departement of European Administrative Management Hochschule Harz of Applied Sciences, Halberstadt, Germany

<sup>f</sup> Departement of Business Administration Cukurova University, Adana, Turkey

<sup>1</sup> [samsul.pahmi@nusaputra.ac.id](mailto:samsul.pahmi@nusaputra.ac.id); <sup>2</sup> [ade.hudiana\\_SI15@nusaputra.ac.id](mailto:ade.hudiana_SI15@nusaputra.ac.id); <sup>3</sup> [lesri\\_SI15@nusaputra.ac.id](mailto:lesri_SI15@nusaputra.ac.id),

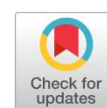
<sup>4</sup> [Santi.laswati\\_SI15@nusaputra.ac.id](mailto:Santi.laswati_SI15@nusaputra.ac.id), <sup>5</sup> [lucasramer@web.de](mailto:lucasramer@web.de), <sup>6</sup> [bahadirozsut@gmail.com](mailto:bahadirozsut@gmail.com)

\* Corresponding Author

Received 06 November 2021; revised 12 November 2021; acceted 20 November 2021

## ABSTRACT

This research aims to identify the existing problems on the Koperasi Karya Usaha Mandiri as skim credit for poor families by granting credit in groups. The issue raised, namely the influence of Decision Supporting System of Granting Loans against credit jam. This research is a type of quantitative research which uses 8 independent variables and 1 dependent variable. Method of data collection in this research is the observation, interviews, literature studies and documentation. This research method uses Binary Logistic Regression to analyze the determination decision granting loans to prospective members and cause bottlenecks in financing micro-credits in Koperasi Karya Usaha Mandiri Syariah.



## KEYWORDS

Binary Logistic  
Regression  
Granting Loans  
Decision System  
Supporting



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

## 1. Introduction

A cooperative loan is a form of the financial institution which plays an important role in helping to improve the welfare of society by way of lumping the fund and channel it in the form of credit or lending to the public. Credit is the ability to execute a purchase or hold a loan with payment will be made and suspended in an agreed period of time [1]. However, the current cooperative borrows much more highlights on the middle down of society to the lending target, or better known as Grameen Bank.

Credit is given to the poor is very important because it can help increase their incomes so they can get out of the problem of poverty by making use of such credit as a source of capital that is utilized in accordance with their condition [2], but in terms of this credit does not always run smoothly, there are times when issues arise.

Security issues are the main problems of credit to be aware of, because of the possibility that what can happen in the granting of credit is debt repayment or payment of congestion which is done by the customer. This can result poorly against the operational viability of cooperatives, therefore requiring a control system of good granting credit in order to ensure credit transmitted to the customer can run effectively and avoid all form of aberration [3].

The greater the credit is given to members, then troubled loans are also the possibility of getting bigger. One way to reduce bad debts, namely cooperative parties, must always use the correct credit analysis in the matter of granting credit to members. KUD Rahmat Widodo is one of KUD in Kebumen regency, Indonesia. KUD Rahmat Widodo had problems of bad credit that the percentage increased from the year 2011 to 2013 as follows: 21.70% in 2011, 25.49% in 2012 and 26.53% in 2013. The percentage is far from the limit maximum i.e. 5% according to the applicable provisions. This indicates that KUD Rahmat Widodo is having serious credit problems so that it will affect the performance of the cooperative [4]. Koperasi Karya Usaha Mandiri Syariah, is an institution which is

engaged in the field of Services, consultancy and development of micro financing that are specific for poor households in rural Indonesia by using the Grameen Bank approach.

The Grameen Bank that applied in KUMS is credit skimmed to poor families by granting of credit in groups (Rembug Center) and does not use the guarantee. Although the granting of credit in groups but in use and returns is the responsibility of each member of the construction [5]. Based on the background of the issue required a data mining technology may be the best choice in determining the decision lending cooperatives.

A method of data mining can be used to predict based on past data is classification. One method of classification i.e. C 4.5 can be used to predict nonperforming loans in case of submission of loan clients generating rules in the form of a decision tree. In this case, the researchers will use binary logistic regression analysis to figure out the best way to determine a decision granting a loan on the cooperative based on the variables that influence.[6]

## 2. The Literature Review

Credit or *cedere* in the greek language means a trust or *credo* which means I believe. The criteria used in general by financial institutions for credit granting there are five indicators better known as 5 C, which are character, capacity, capital, collateral and condition of economy [7]. One of the barriers to rural communities of lending is the ability to give guarantees as the terms of the lending bank in particular.

Cooperatives provide collateral requirements, however small and size are not the same as a bank. Because the cooperative does not require as much as a bank guarantee. So that it can facilitate cooperative members and fulfill loans. "Bad credit or problem loan is credit that has difficulty paying off due to factors or elements of the gap or due to conditions outside of the ability of the debtor.[8]

Many factors determine if the customer does not pay the installment credit smoothly. In one example is the economic situation that happened to that interfere with the business which was run. In a broader sense credit risk is the uncertainty or fluctuation of profit in credit activity [9]. Another factor is the lack of credit monitoring and supervision by the cooperative. They must ensure that the loans are utilized for the purpose for which they are defined. This can reduce the risk of bad debts in the cooperative [10].

Research Kaunang explained that the interest rate changes against a credit request show the regression coefficients of (-0.7983) means an increase in interest rates by 1% the credit request will be dropped by 79.83%. *ceteris paribus*. NPL negative and significant effect on the demand for credit. This is shown by the regression coefficient i.e. of NPL (0.6438). This means that each NPL decreased by 1% then the credit request will be up by 64.37%. *ceteris paribus* [11].

## 3. Method

### 3.1. Populations and Samples

The population in this study are all the borrowers that are in Koperasi Karya Usaha Mandiri Parungkuda, Indonesia. This research population is 190 people.

### 3.2. Data Collection

The collection of data required several stages that must be implemented to support the success of the research. These stages include interviews, observation, documentation and study of literature.

### 3.3. Research Variables

The variable in this study consists of the dependent variable is one that is the feasibility of granting the loan (Y) and 8 independent variables include gender (X1), education (X2), the amount of the loan (X3), earnings (X4), the number of families (X5), expenditure (X6), job type (X7) and the status of loaning (X8).

### 3.4. Data Analysis

In the process of this research consists of a number of factors that may affect the feasibility of granting the loan. As for the data analysis used in qualitative research, data consists of three stages of reduction, among other data, the presentation of data and data analysis.

Binary logistic regression analysis was used to describe the relationship between the response variable and binary data with the free variables in the form of scale intervals data and category. Binary logistic regression is actually the same as the multiple regression analysis, only this is bound variables variable dummy (0 and 1). Logistic regression does not require the normality assumption, although screening data outliers can still do.

Assumptions in binary logistic regression were as follows:

- a) Does not assume a linear relationship between the dependent and independent variables.
- b) The dependent variable must be the dichotomy (2 variables)
- c) Independent variable does not have to have the same diversity between groups of variables
- d) Category in the independent variable must be separated from each other or exclusive
- e) Sample that is required in the amount of the relatively large, minimum required up to 50 sample data for a predictor variables (free)

Unlike ordinary linear regression, binary logistic regression does not assume a relationship between the dependent and independent variable in linear regression.

These models are used in the binary logistic regression are :

$$\text{Log} (P / 1 - p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

Where  $p$  is the probability that  $Y = 1$ , and  $X_1, X_2, X_4$  are independent variables, and  $b$  is the coefficient of regression.

The steps in the use of binary logistic regression analysis are as follows :

- a) A Test of the Model Significance

To know the influence of independent variables against dependent variables together (overall) in the model, can use Likelihood Ratio Test, whose hypotheses are as follows:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$  (There is no influence between independent variables simultaneously against the dependent variables)

$H_1 : \text{at least one } \beta_j \neq 0$  (no influence at least one independent variable against dependent variable)

For  $j = 1, 2, \dots, p$

The test statistic used is :

$$G^2 = -2 \ln \frac{L_0}{L_p}$$

With :

$L_0$ : Maximum Likelihood of Reduced Model or a model consisting of a constant course

$L_p$ : Maximum Likelihood of the Full Model or with all the independent variables.

$G^2$  statistics follows the chi-squared distribution with non-degrees of  $p$  so the hypothesis is rejected if the  $p$ -value  $< \alpha$ , meaning independent variable  $X$  simultaneously affects dependent variable  $Y$ .

- b) Test the Fit of the Model

Test the fit of the model is used to evaluate the suitability of the model with data, the value of the observations obtained the same or close to that expected in the model. As for the model used must meet the Goodness of Fit (GoF). A model categorized meets the GoF if there is conformity between

the data entered in the model with observed data. In logistic regression, a method used to test the feasibility of a model can be measured by the value of the *chi-square* test *with Hosmer and Lemeshow*. This test can be done by looking at the value of goodness of fit test that measured the value of chi-square on a significant level of 5%.

The decision of the acceptance of the hypothesis is based on the following considerations.

$H_0$  : models are hypothesized to fit with the data

$H_1$  : models are hypothesized to not fit with the data

c) A partial Test and the establishment of Model (Prediction Parameters)

In general the purpose of the analysis is to find a suitable model between the model with data. Prediction Parameters testing (coefficient  $\beta$ ) can be done partially through Test Wald with whose hypotheses are as follows :

$H_0$  :  $\beta_j = 0$  (Independent variable does not have a significant influence towards dependent variable)

For  $j = 1, 2, \dots, p$

With test statistics as follows:

$$W = \left[ \frac{\beta_j}{Se(\beta_j)} \right]^2$$

The hypothesis will be rejected if the p-value  $< \alpha$  that means independent variable  $X_j$  affects dependent variable  $Y$ .

d) Odds Ratio ( $\theta$ )

Odds Ratio is a good indicator in showing your chances of success and failure based on the specific conditions of the data. This has a very important role in logistic regression because it is used as a descriptive statistics. Odds Ratio defined as the ratio of the odds for  $X_j = 1$  against  $X_j = 0$ . Odds ratio stated the risk propensity of the influence of observation with  $X_j = 1$  is how to fold when compared with observation  $X_j = 0$ . For continuous-scale independent variables then the interpretation of coefficients  $\beta_j$  of the regression models logistic is an increased  $c$  unit of independent variable will cause the risk of occurrence of  $Y = 1$ , if the  $\exp(c \cdot \beta_j)$  times greater.

$$\theta = \frac{[\pi(1)/[1-\pi(1)]]}{[\pi(0)/[1-\pi(0)]]}$$

3.5. Framework

In figure 1 below, it appears that the problem that occurs is a poor credit system that causes wrong decisions on the part of management. Meanwhile, the variables studied in the form of gender, education, number of families, work and so on and some data analysis tools are used to produce decisions in performance predictions and facilities from management.

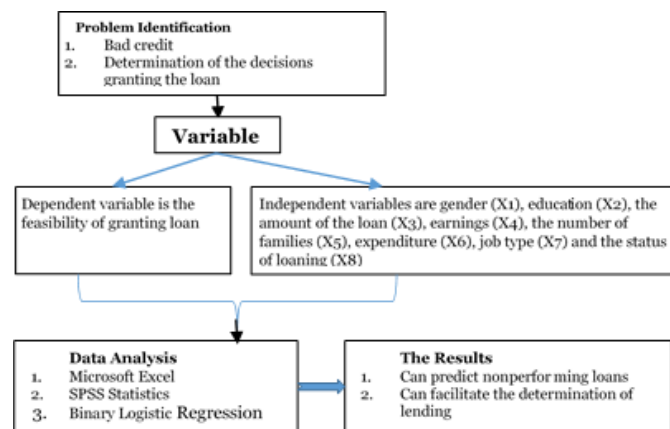


Fig. 1. Research flow

Based on figure 1. Research Flow describes the identification of problems with bad credit and the determination of lending decisions. The variables used are gender, education, loan amount, income, family number, expenses, type of employment and loan status. With the identification of problems and types of variables that have been determined then data analysis is carried out using Microsoft Excel, SPSS Statistics and Binary Logistic Regression that can predict problematic credit and can facilitate credit determination.

#### 4. Results and Discussion

##### 4.1. Description of Data

##### 4.1.1. Case Processing Summary

**Table 1.** Case Processing Summary

		N	Percent
Unweighted Cases <sup>a</sup>	Included in Analysis	190	100.0
	Missing Cases	0	.0
	Total	190	100.0
Unselected Cases		0	.0
Total		190	100.0

The table above i.e. Case Processing Summary is a summary of the data samples taken, namely 190 samples. The output of the Case Processing Summary explains that there is no sample lost/missing.

##### 4.1.2 Encoding Regresi Logistic

**Table 2.** Dependent Variable Encoding

Original Value	Internal Value
non-current credit payments	0
good credit payment	1

The table above illustrates the results of the process of data input that is used on the dependent variable, i.e., the category of "Not Smooth" code 0 and category "Smoothly" with code 1.

##### 4.1.3 Iteration History

**Table 3.** Iteration History<sup>a,b,c</sup>

Iteration		-2 Log likelihood	Coefficients
			Constant
Step 0	1	168.268	1.368
	2	165.761	1.647
	3	165.742	1.674
	4	165.742	1.674

The Iteration History table above the values -2 Log Likelihood i.e. 165,742. The value of the Degree of Freedom (df) = N – 1 = 1 = 190-189 and the value of the Chi-Square (X<sup>2</sup>) tables on df 189 at 0.05

significance level = 165.742 A value of  $-2 \text{ Log Likelihood } (165,742) > X^2 \text{ table } (165.742)$  so to reject  $H_0$ , then shows that the model before entering the independent variable was NOT a FIT with the data.

#### 4.1.4 Classification Table

**Table 4.** Classification Table<sup>a,b</sup>

Observed		Predicted			
Step 0	Y	No	Y		Percentage Correct
		Yes	No	Yes	
	Overall Percentage		0	30	100.0
			0	160	84.2

In classifying observations known percentage of model is  $160/190 = 84.2\%$ . With the number of samples as much as 190 people, there were 160 people are precise in his classification by the logistic regression model.

#### 4.1.5 Pseudo R Square

**Table 5.** Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	.000 <sup>a</sup>	0.582	1

From the table above, to explain the dependent variable by looking at the ability of the independent variable, then use Cox and Snell R Square & Nagelkerke R Square. These values are called also with Pseudo R-Square or if on linear regression (OLS) better known as the R-Square. Nagelkerke R Square value of 1 and Cox & Snell R Square 0.582, which suggests that the ability of the independent variable in explaining the dependent variable is 1 or 100% and there are  $100\% - 100\% = 0\%$  other factors outside the model that describes the dependent variable.

#### 4.1.6 Prediction Parameters

In the table Variable in the equation above is known to all independent variables the P value test wald (Sig)  $> 0.05$ , which means that each variable has a significant partial influence towards Y in the model.  $X_2, X_3, X_4, X_5, X_6, X_7, X_8$ , has a value of Sig Wald  $> 0.05$  so it does not reject  $H_0$  or the means of education, loan amount, income, number of families, spending, type of work and description of efforts do not provide significant partial influence against the occurrence of credit.

#### 4.1.7 Odds Ratio

Exp (B) value shows the magnitude of influence or so-called Odds Ratio. Variable education has an odds ratio of 1,345, then it can be said that people with higher education then prone credit smoothly into 1,187 times. Then it can be said that the trend experienced credit smoothly on each variable depends on the magnitude of the values of the Odds Ratio.

#### 4.1.8 The Logistic Regression Equation

Based on the values in the calculation of B above, then the model equation is formed are as follows:

Probability or Predicted =

$$\frac{(Exp(-49.600 + (0.297x1) + 0 + 0 + (0.448x1) + 0 + (0.212x1) + (0.315x1)))}{1 + (Exp(-49.600 + (0.297x1) + 0 + 0 + (0.448x1) + 0 + (0.212x1) + (0.315x1)))}$$

**Table 6.** Variables in the Equation

	B	S.E.	Wald	Df	Sig.	Exp(B)	95% C.I.for EXP(B)	95% C.I.for EXP(B)	
Step 1*	X2	.297	7876.578	.000	1	1.000	1.345	Lower	Lower
	X3	.000	.002	.000	1	.992	1.000	.000	.
	X4	.000	.002	.000	1	1.000	1.000	.997	1.003
	X5	.448	2331.393	.000	1	1.000	1.566	.997	1.003
	X6	.000	.002	.000	1	.999	1.000	.000	.
	X7	.212	5813.988	.000	1	1.000	1.236	.995	1.005
	X8	.315	2126.774	.000	1	1.000	1.371	.000	.
	Constant	-49.600	17877.717	.000	1	.998	.000		

The variable table in the equation shows the variable value (X2) of education, (X3) the number of loans, (X4) income, (X5) the number of families, (X6) expenditures, (X7) types of work and loan status (X8). With a constant value of -49,600.

#### 4. Conclusion

After doing a logistic regression analysis involving 8 variables consisting of gender, education, loan amount, income, number of families, expenses, type of work, and loan status, based on the results and discussion of the research, it can be concluded that, based on the results and discussion of the research it can be concluded that the factors that affect the loan status in the Mandiri Syariah Karya Usaha Cooperative are education (X2), loan amount (X3), income (X4), number of family (X5), expenditure (X6), type of work (X7) and loan status (X8). The model equation formed is  $\ln P/1-P = -49.600+0.297+0+0+0.448+0+0.212+0.315$ .

#### References

- [1] F. Alanshari and D. Marlius, "Procedures for Providing Mortgage Loans at PT Bank Tabungan Negara (Persero) Tbk Bukittinggi Sub-Branch," *Akad. finance. and Banking. Padang*, no. 2014, pp. 1–11, 2018.
- [2] W. Cahyawan and S. Viciawati Machdum, "Institutionalist Approach to Microfinance and Its Role in Poverty Reduction: Problems and Solutions," *Sosio Inf.*, vol. 5, no. 1, pp. 64–80, 2019, doi:10.33007/inf.v5i1.1685.
- [3] N. Royan, "Chapter 1: Introduction," *Health Profile. kab*, vol. 41, no. 2005, pp. 1–9, 2015.
- [4] Y. Nawangsari and IS Putra, "Analysis of the Internal Control System for Credit Provision in Reducing Bad Loans in the Kharisma Mitra Karya Savings and Loan Cooperative," *J. Kompil. Eco Science.*, vol. 8, no. 1, pp. 67–80, 2016, [Online]. Available: <http://journal.stieken.ac.id/index.php/kompilek/article/view/278>.
- [5] M. Fasha, "By : UNIVERSITAS ISLAM RIAU," 2021.
- [6] N. Using, M. Classification, P. Cooperative, S. Borrow, and M. Raya, "obtained a Bachelor Degree in Computers Compiled by : Iqbal Taufiq Ahmad Nur," 2018.
- [7] T. Mardiana, "Application of the TOPSIS Algorithm in a Lending Decision Support System in Employee Cooperatives," *J. Khatulistiwa Inform.*, vol. 5, no. Vol 5, No 2 (2017): Period December 2017, pp. 128–138, 2017.
- [8] E. Fitriani, "Analysis of Non-Performing Loans at PT. BPR Supra Bandung," *J. Financial Science. and Banking.*, vol. 3, no. 1, pp. 1689–1699, 2013.

- [9] A. Nursyahriana, M. Hadjat, and I. Tricahyadinata, "Analysis of Factors Causing Bad Credit," *Forum Ekon.*, vol. 19, no. 1, p. 1, 2017, doi:10.29264/jfor.v19i1.2109.
- [10] NLASH Sari, AAS Indrawati, and S. Putrawan, "The Settlement of Bad Loans at the Savings and Loans Cooperative Karya Artha Sedana and Ksp. Wirartha Utama in South Denpasar City," *Kertha Semaya J. Huk Science.*, vol. 7, no. 8, p. 1, 2019, doi:10.24843/km.2019.v07.i08.p05.
- [11] G. Kaunang, "The level of interest rates on loans and bad loans affects the demand for MSME loans in Indonesia," *Emba*, vol. 1, no. 3, pp. 920–930, 2013.