

CHAPTER III

RESEARCH METHODOLOGY

3.1. Research Instrument

In this study, the author used observation method to collect and get the data that will be used and performed in this study. Direct observation is conducted to the financing company which is PT. Multindo Auto Finance to ask and request the data. During the observation, the company agrees and gives a confirmation that the data can be exported and used in this study. As described before, the data used to be observed in this study was taken from PT. Multindo Auto Finance Semarang. The data was taken on April, 2015 directly from PT. Multindo Auto Finance Semarang.

3.2. Data Collection Method

The data used in this study is a data provided by the company which is PT. Multindo Auto Finance Semarang. There are two kinds of data, which are customer data and customer transaction data. Both of the data provided by the company are the data that will be used in this study.

Besides that, this study also used another data resources to help the author in arranging this final year project as described in Chapter II. Those data resources are some of literature review that has been used, which are:

1. Journals about credit cases and credit activities
2. Data Mining Books (Text Book and E-Book)

3.3. Data Analysis Technique

Cross-Industry Standard Process for Data Mining (CRISP-DM) that has been developed provides a general data mining process as a problem solving strategy for this study. In CRIPS-DM, a data mining project is divided into

six phase. This study follows the steps of CRISP-DM as a standard data mining process in business environment.

3.3.1. Business Understanding Phase

PT. Multindo Auto Finance is a one among many finance companies which is worked as a consumer finance agency. PT. Multindo Auto Finance, based on its business license, is carries on business in the area of consumer financing. Currently, PT. Multindo Auto Finance has two divisions that have been divided based on the type of the vehicles. Those are Motorcycle Division and Cars Division. In this study, the author focus on the cars division.

Until now, PT. Multindo Auto Finance is giving fast, appropriate, and flexible finance solution for people to own cars. Finance solution offered by PT. Multindo Auto Finance to the customer is form as a credit loan. Credit activity in PT. Multindo Auto Finance is similar to any credit activity conducted by bank. It may also have a credit problems e.g. loss credit.

In PT. Multindo Auto Finance, credit can be classified into three types which are regular credit, problem account, and loss credit. Regular credit is a credit with on time repayment or within 0 – 29 days late. Problem account is a credit with a problem which is late in repayment within 30 – 59 days. Loss credit is a credit which has late in the repayment more than 60 days (60 – UP).

As a big finance company that has 62 branches office in Sumatra, Jawa, and Bali, PT. Multindo Auto Finance has many customers with different profiles. Moreover, with the demand of credit applicants to apply the loan, PT. Multindo Auto Finance need to be concern on the credit risk that may appears on giving the loan. Thus, it is needed to classify the customer credit data before credit loan is accepted to minimize the risks in credit repayment.

3.3.2. Data Understanding Phase

Data used in this study was taken from PT. Multindo Auto Finance Semarang. The data given by the company is already in excel format. There are two data given, which are customer data and customer transaction data.

The customer data given by PT. Multindo Auto Finance is the data about customer profile that has been registered as a credit customer in this company. This is the main data used in this study. This customer data consists of 10264 data records with 12 attributes. While, for the customer transaction data is the data of customer credit repayment report in period of 2014. This data is an additional data used as a selector in data preparation phase. This data consists of 62747 records data with 7 attributes.

All of data that have been retrieved from PT. Multindo Auto Finance will be further processing in the data preparation phase. The sample data of customer data and customer transaction data used in this study can be seen in the table 3.1 and table 3.2.

Table 3.1 Sample of Customer Data

NopinAll	RealisasiDate	MerkName	CategoryName	occupation	Grossincome	Tanggungan	LoanType	AngsuranReal	Tenor	total pinjaman	OS PINjaman PER 31 DES 14
0012002101-010	22/02/2014	SUZUKI	LOW PICKUP	Wiraswasta	9,000,000	0	ADDDB	1,450,600	36	52,221,600.00	37,715,600
0012002217-009	05/09/2014	TOYOTA	LOW MPV	Wiraswasta	16,000,000	3	ADDM	3,835,756	36	138,087,216.00	122,744,192
0012050079-003	25/06/2014	SERTIFIKAT	LAINNYA	Karyawan Swasta	4,700,000	2	ADDM	1,089,574	36	39,224,664.00	31,597,646
0012050131-002	25/03/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	16,800,000	5	ADDDB	5,097,011	48	244,656,528.00	198,783,429
0012050213-005	11/12/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	24,000,000	2	ADDM	4,889,901	24	117,357,624.00	112,467,723
0012050320-006	27/10/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	14,400,000	3	ADDDB	3,687,207	48	176,985,936.00	169,611,522
0012050333-006	07/03/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	48,000,000	2	ADDDB	18,937,298	12	227,247,576.00	56,811,894
0012050486-003	23/09/2014	DAIHATSU	PICKUP	Wiraswasta	22,500,000	3	ADDM	1,730,300	24	41,527,200.00	34,606,000
0012050550-003	07/04/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	12,750,000	2	ADDM	5,192,251	36	186,921,036.00	140,190,777
0012050575-004	18/06/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	6,000,000	3	ADDM	3,943,612	36	141,970,032.00	114,364,748
0012050594-004	23/12/2014	HYUNDAI	SEDAN	Pendidikan	4,000,000	2	ADDDB	1,761,114	36	63,400,104.00	63,400,104
0012050596-002	18/01/2014	MITSUBISHI	LIGHT TRUCK	Wiraswasta	18,000,000	3	ADDDB	5,998,401	48	287,923,248.00	227,939,238
0012050643-003	13/09/2014	TOYOTA	LOW MPV	Wiraswasta	25,000,000	3	ADDM	2,990,751	36	107,667,036.00	95,704,032

Explanation:

NopinAll	Customer credit ID
RealisasiDate	Date for the company to realize the credit loan
MerkName	The vendor of the vehicle manufacturer
CategoryName	Type of the vehicle
Occupation	Customer daily work
Grossincome	Customer income in a month
Tanggungsan	Number of people that become the responsibility of the customer
LoanType	Type of loan given by the company
AngsuranReal	Amount of interest that customer need to repay every month to the company
Tenor	Credit repayment time
Total pinjaman	Total amount that customer need to repay including the credit interest
OS PINjaman PER 31 DES 14	Rest of total amount that customer need to repay to the company per December 31 th , 2014

Table 3.2 Sample of Customer Transaction Data

ID	Apldate	RealisasiDate	HARI TGK	BucketDueJT	AngsuranReal	LastPaidDate
0012002101-010	28/02/2014	22/02/2014	0	CURRENT	1,450,600	
0012002101-010	31/03/2014	22/02/2014	0	CURRENT	1,450,600	20/03/2014
0012002101-010	30/04/2014	22/02/2014	0	CURRENT	1,450,600	22/04/2014
0012002101-010	31/05/2014	22/02/2014	9	008-029 hari	1,450,600	22/04/2014
0012002101-010	30/06/2014	22/02/2014	0	CURRENT	1,450,600	21/06/2014
0012002101-010	26/07/2014	22/02/2014	0	CURRENT	1,450,600	22/07/2014
0012002101-010	31/08/2014	22/02/2014	0	CURRENT	1,450,600	19/08/2014
0012002101-010	30/09/2014	22/02/2014	0	CURRENT	1,450,600	20/09/2014
0012002101-010	31/10/2014	22/02/2014	0	CURRENT	1,450,600	29/10/2014
0012002101-010	30/11/2014	22/02/2014	8	008-029 hari	1,450,600	29/10/2014
0012002101-010	31/12/2014	22/02/2014	0	CURRENT	1,450,600	27/12/2014
0012002217-009	30/09/2014	05/09/2014	0	CURRENT	3,835,756	05/09/2014
0012002217-009	31/10/2014	05/09/2014	0	CURRENT	3,835,756	09/10/2014
0012002217-009	30/11/2014	05/09/2014	0	CURRENT	3,835,756	06/11/2014
0012002217-009	31/12/2014	05/09/2014	0	CURRENT	3,835,756	06/12/2014
0012050079-003	30/06/2014	25/06/2014	0	CURRENT	1,089,574	25/06/2014
0012050079-003	26/07/2014	25/06/2014	0	CURRENT	1,089,574	22/07/2014
0012050079-003	31/08/2014	25/06/2014	0	CURRENT	1,089,574	26/08/2014
0012050079-003	30/09/2014	25/06/2014	0	CURRENT	1,089,574	25/09/2014
0012050079-003	31/10/2014	25/06/2014	0	CURRENT	1,089,574	24/10/2014
0012050079-003	30/11/2014	25/06/2014	0	CURRENT	1,089,574	27/11/2014
0012050079-003	31/12/2014	25/06/2014	0	CURRENT	1,089,574	22/12/2014
0012050131-002	31/03/2014	25/03/2014	0	CURRENT	5,097,011	
0012050131-002	30/04/2014	25/03/2014	0	CURRENT	5,097,011	26/04/2014
0012050131-002	31/05/2014	25/03/2014	0	CURRENT	5,097,011	23/05/2014
0012050131-002	30/06/2014	25/03/2014	0	CURRENT	5,097,011	25/06/2014
0012050131-002	26/07/2014	25/03/2014	0	CURRENT	5,097,011	25/07/2014
0012050131-002	31/08/2014	25/03/2014	0	CURRENT	5,097,011	26/08/2014
0012050131-002	30/09/2014	25/03/2014	0	CURRENT	5,097,011	29/09/2014
0012050131-002	31/10/2014	25/03/2014	0	CURRENT	5,097,011	29/10/2014
0012050131-002	30/11/2014	25/03/2014	0	CURRENT	5,097,011	26/11/2014
0012050131-002	31/12/2014	25/03/2014	0	CURRENT	5,097,011	26/12/2014
0012050213-005	31/12/2014	11/12/2014	0	CURRENT	4,889,901	11/12/2014
0012050320-006	31/10/2014	27/10/2014	0	CURRENT	3,687,207	
0012050320-006	30/11/2014	27/10/2014	0	CURRENT	3,687,207	26/11/2014
0012050320-006	31/12/2014	27/10/2014	0	CURRENT	3,687,207	27/12/2014

Explanation:

ID	Customer Credit ID
Apldate	Date of the transaction recorded
RealisasiDate	Date for the company to realize the credit loan
HARITGK	Amount of late day for the customer to repay the credit
BucketDueJT	Time of late classification
AngsuranReal	Amount of interest that customer need to repay every month to the company
LastPaidDate	Date of the last time to repay the credit each month

3.3.3. Data Preparation Phase

In this study, not all of the data will be used in the process. In this phase, data preparation is conducted to prepare the data before it is ready to be processed. The first step conducted is data selection process. In this process, the real data that will be used for this study is selected. The main data that in this study is customer data. But, as explained before, not all of the data in customer data will be used in this study. To decide which data that will be used or not in customer data, it is based on the customer transaction data.

Customer transaction data show the transaction of each customer in credit repayment process within 2014. Each customer has a different repayment record. There are customers that already complete with 12 transactions, but there are also customers with only 1 or 2 transaction. Based on that case, the data selection process is conducted. Customer with credit repayment transactions between 9 – 12 transactions is selected. While, customer with credit repayment less than 9 is not selected. Unless, the customer who has shown the indicator of credit problem.

From this selection process, 3253 records of customer data are selected as training set. This training set is the real data set used in this study. Moreover, in the training set, attribute status is added as a label in this data. The context of attribute status is based on the credit categorization from the company. After the selection process, the next step conducted is labeling process. Labeling process is a process in choosing the attribute of the data used as a predictor variable and target variable.

Table 3.3 Detail Attributes in Training Set

Attribute	Utilization Detail		
NopinAll	×	No	-
Status	√	Yes	Target variable
RealisasiDate	×	No	-
MerkName	√	Yes	Predictor variable
CategoryName	√	Yes	Predictor variable
Occupation	√	Yes	Predictor variable
Grossincome	√	Yes	Predictor variable
Tanggungan	√	Yes	Predictor variable
LoanType	√	Yes	Predictor variable
AngsuranReal	√	Yes	Predictor variable
Tenor	√	Yes	Predictor variable
Total pinjaman	×	No	-
OS PINjaman PER 31 DES 14	×	No	-

The table above consists of attributes in training set used in this study. Each attributes has its own integrity and has been elected. The attributes with yes indicator (√) show that those attributes is elected to be used in this study. While, the attributes with no indicator (×) show that those attributes will not be used in this study. The sample of final data used in this study can be seen in the table 3.4.

Table 3.4 Sample of Final Data

Status	MerkName	CategoryName	occupation	grossincome	Tanggung	LoanType	AngsuranReal	Tenor
Lancar	SUZUKI	LOW PICKUP	Wiraswasta	9,000,000	0	ADDB	1450,600	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	16,800,000	5	ADDB	5,097,011	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	48,000,000	2	ADDB	18,937,298	12
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	12,750,000	2	ADDM	5,192,251	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	18,000,000	3	ADDB	5,998,401	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	10,000,000	2	ADDB	3,871,112	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	14,700,000	3	ADDM	5,881,867	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	16,000,000	2	ADDM	5,386,384	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	9,000,000	3	ADDM	4,791,488	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	12,000,000	4	ADDM	4,499,951	36
Lancar	TOYOTA	LIGHT TRUCK	Wiraswasta	12,500,000	2	ADDB	4,335,801	48
Macet	MITSUBISHI	LIGHT TRUCK	Wiraswasta	19,800,000	2	ADDM	5,881,867	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	10,500,000	1	ADDM	4,022,484	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	25,000,000	2	ADDB	6,813,156	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	15,200,000	2	ADDM	3,898,000	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	9,000,000	2	ADDM	4,791,488	36
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	12,000,000	3	ADDM	5,114,667	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	11,250,000	3	ADDM	5,881,867	48
Lancar	MITSUBISHI	LIGHT TRUCK	Wiraswasta	16,000,000	2	ADDB	5,509,401	48
Lancar	TOYOTA	MINIBUS	Pemerintahan	4,968,900	2	ADDM	2,275,724	36

3.3.4. Modeling Phase

The model proposed to be used in this study is NBC model. This model will be applied in RapidMiner application. Thus, accuracy checking in this study performed by using RapidMiner Ver.5.3.013 framework.

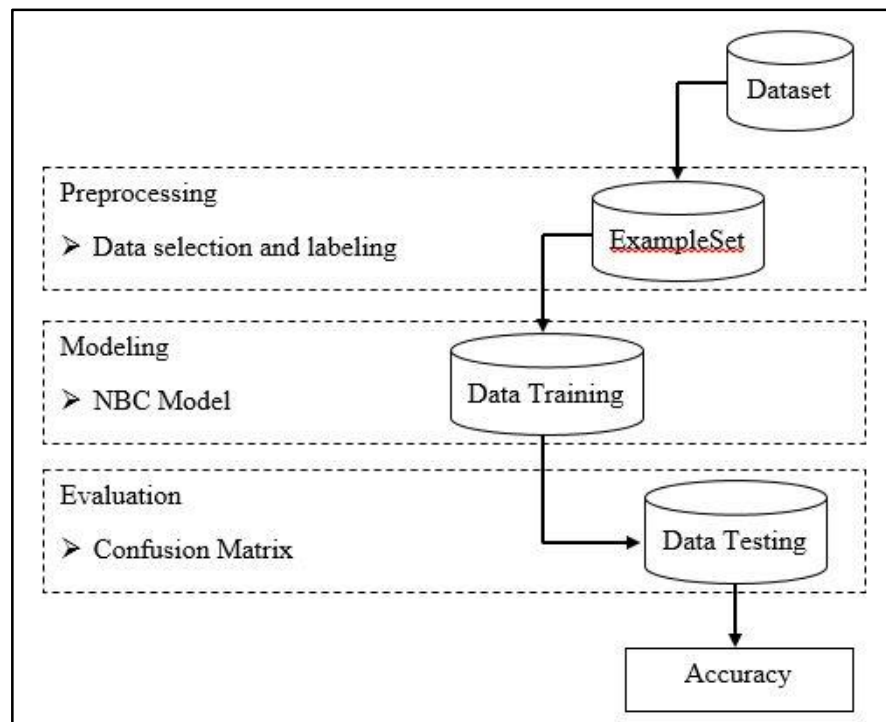


Figure 3.1 Proposed Model for The Study

3.3.5. Evaluation Phase

In this phase, validation and accuracy measuring from the result model are performed. Several techniques for validation and calculation of the accuracy are available in RapidMiner Ver.5.3.013. In this study, validation process is performed using split validation with relative split type and stratified sampling type. Moreover, for calculating the accuracy of the result, confusion matrix is performed.

Table 3.5 Sample of Confusion Matrix

Classification	True Lancar	True Macet	True Bermasalah
Pred. Lancar	2016	67	52
Pred. Macet	108	7	3
Pred. Bermasalah	21	3	0

Based on the table, the accuracy level can be calculate by:

$$Accuracy = \left(\frac{2016 + 7 + 0}{2016 + 67 + 52 + 108 + 7 + 3 + 21 + 3} \right) \times 100\%$$

$$Accuracy = \left(\frac{2023}{2277} \right) \times 100\% = 88.84\%$$

3.3.6. Deployment Phase

The result of this study is an analysis that can be proposed as a Decision Support System (DSS) for the company which is PT. Multindo Auto Finance Semarang. This result can be used by the company to classified credit customers based on credit status in the credit repayment. Moreover, this result can also be used as a predictor by predicting the credit applicant status in the future. Besides that, the result of this study can also be used as a reference in further research.