JurnalMantik
Volume 4, Number 3, Nov 2020, pp. 1885-1899 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik

Classification of Feasibility of Credit for Candidated CS Finance Debtors Using Naïve Bayes Method

Nuraisana[1], Ellisa Purba[2]

[1,2]Teknik Informatika, STMIK Pelita Nusantara, Jl.Iskandar Muda No.1 Medan, Sumatera Utara, Indonesia

Email : nuraisana94@gmail.com , brpurbaellisamei12@gmail.com

ARTICLE INFO

Article history:
Received: 12/07/2020
Revised: 22/08/2020
Accepted: 30/09/2020

Keywords:
Credit; Debitor; Naïve Bayes; Data Mining

ABSTRACT

CS Finance is one of the central financing institutions in the two-wheeler finance industry. CS Finance, which was founded in 2010 under the name PT Central Santosa Finance. The problem that is often faced is when conducting administrative assessments to determine the right prospective debtor's eligibility. We need a system that can assist CS Finance in determining the feasibility of prospective debtors quickly and precisely. The method used in this research is Naïve Bayes. The data processed is data of prospective debtors. The variables used have been determined based on four attributes, namely character, capacity, capital, and conditions; testing is carried out using Rapidminer software, and the accuracy of the Naïve Bayes algorithm for predicting the feasibility of prospective debtors based on training data shows good performance, namely 80%. Hence, it is feasible for use. To make it easier for users to predict prospective borrowers' creditworthiness, a creditworthiness classification system for prospective debtors has been created in CS Finance using the web-based naïve Bayes method.

1. Introduction

Founded based on the belief to become the premier financial institution in the two-wheeler finance industry, CS Finance, founded in 2010 as PT Central Santosa Finance, is here to be a strong player in its field. Supported by BCA bank, the largest private bank in Indonesia, the company continues to develop its business by creating a competitive advantage to produce good results for consumers, ATPM, dealers and shareholders, and all CS Finance employees.

Purchasing a motorized vehicle with credit payments is currently quite attractive to the public due to the simple requirements that the community can fulfill. However, it will take a lot of time in manual services, and errors often occur in the administrative process.

A funding or capital agency can provide cash loans to customers by pledging securities. In the process of processing cash loans, several administrations must be fulfilled. The data will be managed by the admin or Credit Marketing Supervisor (CMS). In the loan application process, an artificial intelligence system is needed, which is expected to make it easier for funding agencies to prioritize customers who will be given loans.

Research like this has also been conducted, which discusses the provision of Yamaha motorcycle credit. The result of the system being built is to assist in providing recognition for Yamaha motorbikes. This system uses two assessment variables, namely installments and income variables, to obtain the proper installment value for the customer (LusiHerlinaSiagian, 2017).

There is previous research, namely the implementation of data mining to predict customers' creditworthiness on BMT BumiMizan Sejahtera Yogyakarta using the C4.5 algorithm. The variables that will be processed by data mining using the c.45 algorithm are the number of family dependents, the amount of financing proposed,the financing requirements, the time, the collateral, the total income, and the character; this system displays the creditworthiness of customers at BMT BumiMizan Sejahtera Yogyakarta with predicting the feasibility of Yes or no based on predetermined variables.

By making this system, it is hoped that it will facilitate credit analysts from CS Finance in conducting administrative assessments to determine the right prospective debtor's feasibility. The problem of recommendations for CS Finance for prospective debtors to take out motor vehicle loans can be resolved and become more effective and efficient. Based on the above background, thecreditworthinessclassificationof CS Finance is made using the Naive Bayes method.

2. Method

2.1 Data Mining

The term data mining has several equivalents, such as knowledge discovery or pattern recognition. The two terms have their accuracy. The term knowledge discovery or knowledge discovery is appropriate because data mining's primary purpose is to get the knowledge hidden in chunks of data. The term pattern recognition or pattern recognition is also relevantbecause the knowledge to be extracted is in the form of patterns that may also need to be removed from within the chunks of data at hand. If the term Data Mining is used in this paper, this is based more on the popularity of data mining activities. So, what exactly is data mining? There are many definitions of this term, and none have been standardized or agreed upon by all parties. However, this term has the essence (notion) as a scientific discipline whose primary purpose is to find my knowledge from the data or information we have. This activity is the primary concern of the Data Mining discipline. Several factors, including drive the continuing progress in the field of Data Mining

a) Rapid growth in the data set
b) Data storage in the data warehouse so that all companies have access to a reliable database.
c) There is an increase in data access through Web and internet navigation.
d) Pressure of business competition to increase market share in economic globalization.
e) Development of software technology for Data Mining (technology availability).
f) Great developments in compute capabilities and storage media capacity expansion. The terms data mining and knowledge discovery in databases (KDD) is often used interchangeably to describe the process of extracting hidden information in an extensive database. The two terms have different concepts but are related to one another. And one of the stages of the whole KDD process is Data Mining [8].

A. Data Mining Science

In its development, data mining has many different definitions. Some definitions of data mining in general:

According to Daryl Pregibon [10]. "Data mining is a mixture of statistics, artificial intelligence, and database research that is still developing."

So it can be concluded that data mining is a series of processes to explore added value in the form of information that has not been known manually from a database. The information generated is obtained by extracting and recognizing essential patterns from the data contained in the database.

Data mining or knowledge discovery database (KDD) aims to utilize data in data by processing it to produce new useful information [10]. The two terms have different concepts but are related to one another. And one of the stages in the whole KDD process is data mining '. The necessary KDD process can be explained as follows [10]:

1. **Data Selection**
   Selection (selection) of data from a set of operational data needs to be done before the KDD information mining phase begins. The selected data will be used for the data mining process, stored in one file, separate from the operational database
2. **Pre-processing / Cleaning**
   Before the data mining process can be carried out, it is necessary to carry out a cleaning process on the data focused on KDD. The cleaning process includes removing duplicate data, checking for inconsistent data, and correcting errors in data, such as printing errors (typography). An enrichment process is also carried out, namely the process of "enriching" existing data with data or other information that is relevant and necessary for KDD, such as external data or information
3. **Transformation**
   Coding is transforming the data that has been selected, so that the data is suitable for the process Data mining. The coding process in KDD is a creative process and is very dependent on the type or pattern of information to be searched for in the database
4. **Data mining**
   Data mining involves looking for patterns or interesting information in selected data using specific techniques or methods. Knowledge discovery in a database (KDD) is defined as the extraction of potential, implicit, and new information from various data sets. The process of knowledge discovery involves the results of the data mining process (the process of extracting a pattern of trends in data), then converting the results accurately into information that is easy to understand

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

5. Interpretation / Evaluation

Information patterns resulting from the data mining process need to be displayed in a form easily understood by parties who may have an interest. This stage is part of the Knowledge Discovery in Database (KDD) process called interpretation. This stage includes examining whether the patterns or information found contradict previously existing facts or hypotheses.

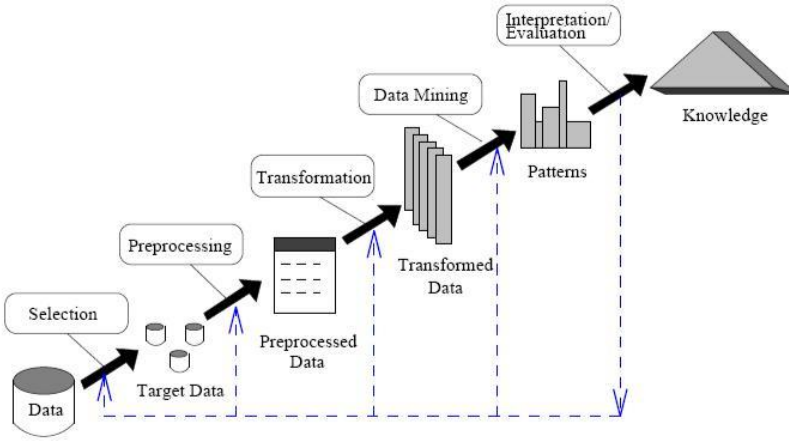
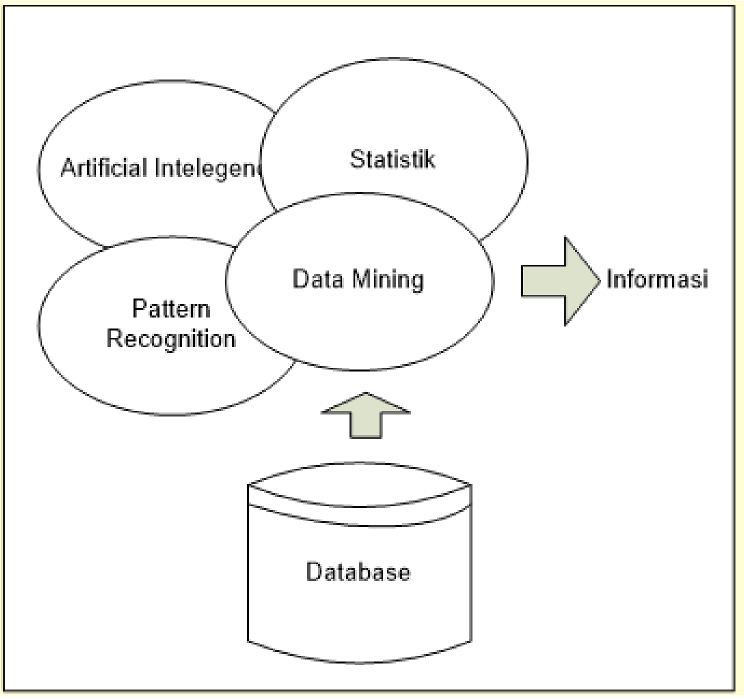


Fig 1. Stages of the data mining process

B. Relationship between Data mining and Other Sciences

Data mining is not an entirely new field. One of the difficulties in defining data mining is that data mining inherits many aspects and techniques from established scientific fields. Figure 2 shows that data mining has long roots in science, such as artificial intelligence, statistics, databases, and Pattern recognition.



Source: (Prasetyo 2014)

Fig 2 The Root Science of Data mining

2.2. Naïve Bayes

Naïve Bayes is a classification method in data mining using probability and statistical techniques, as described by a British scientist named Thomas Bayes. The Naïve Bayes algorithm is a class decision technique, using conditional mathematical probability calculations, that is, the decision value is correct, based on object information [7].

The word Naïve, which seems derogatory, comes from the assumption of independence of the effect of the value of an attribute from the probability in a given class on the number of other features. The use of the Bayes theorem in the Naïve Bayes algorithm is by combining prior probability and conditional probability in a formula that can be used to calculate the probability of each possible classification. The independence model provides the best solution. The effectiveness of the Naïve Bayes method is also seen in examples, and further empirical comparisons, with the same results, are found in Domingos and Pazzani.

Bayes decision theory is a fundamental statistical approach to pattern recognition. Bayesian theory is the possibility of future events that can be calculated by determining previous experiences' frequency. In terms of classification, the Bayes algorithm must have problems that can be seen statistically [6].

Bayesian decision theory is a fundamental statistical approach to pattern recognition *(pattern recognition*). The Bayesian theory is the possibility of future events that can be calculated by determining previous experiences' frequency. The use of the Bayes algorithm in terms of classification must have problems that can be seen statistically [6]

Classification refers to the task of predicting the class label for a point that is not labeled. The Bayes classification (full) uses the Bayes theorem to predict the class to maximize theprobability *posterior*. Its main task is to estimate the combined probability of each class's density function, which is modeled through a multivariate normal distribution. Classification Naive Bayes assumes that the attributes are independent but still very strong for many applications [9]

The training Dataset consists of n points xi in the space dimension d, and let yi indicates a class for each point, with Yi ∈ {c1, c2, ..., ck}. The Bayes classification directly uses the Bayes theorem to predict the class for the new test example, x. estimate the posterior probability P (ci | x) for each ci class, and choose the greatest class probability. The predicted class for x is given as the formula [9]:

$$y = \frac{\text{argmax } \{P\ (ci|x)\}}{ci} \text{ ………………… (1)}$$

Bayes allows us to invert theprobabilities *posterior* in terms of the previous Likelihood and probability, as follows [9]:

$$P\ (ci|x) = \frac{P\ (x|ci).P\ (ci)}{P\ (x)} \text{…… …………..(2)}$$

Where P (x | ci) is the probability, defined as the probability of observing x assuming that the true class is ci, P (ci) is the prior probability of class ci, and P (x) is the probability of observing x from one of the k classes, is given as follows [9]:

$$P\ (x) = \sum_{j=1}^{k} P\ (x|cj).P\ (cj) \text{ …………… (3)}$$

Estimating *thePriorProbability*

To classify values, we have to calculate the likelihoods and probabilities in advance directly from the training dataset D. Let show the part of the point in D labeled with class ci [9]:

$$Di = \{xj \in D\ |xj\ \ yj = ci\} \text{ ……… .(4)}$$

Let the dataset size D is given as | *D* | = *n*, and let the class-specific size of each subset be given as | | =. The prior probability for a class can be estimated as follows [9]:

$$P(c\ i) = \frac{ni}{n} \text{ ……………………… .. (5)}$$

Estimating thevalue *Likelihood*

To estimate the probability P (x | ci), we must estimate the combined probability x across all dimensions. d, that is, we have to estimate [9]

$$Px = (x1, x2, ..., xd) \mid ci \text{ ……………………(6)}$$

3. Results and Discussion

Overview Stages Discussion and Analysis describes the process of analyzing a problem and an overview of the application of methods or algorithms to solve problems. To support data analysis in search of knowledge(knowledge), it will calculate manual data in CS Financefrom the file's *customer survey*. In this case, Rapid Miner software will be usedwhere this application will be very helpful in calculating and analyzing data. Based on the framework that has been discussed from the description of Chapter 3, the stages of analysis and results in following the steps and sequences are as follows:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

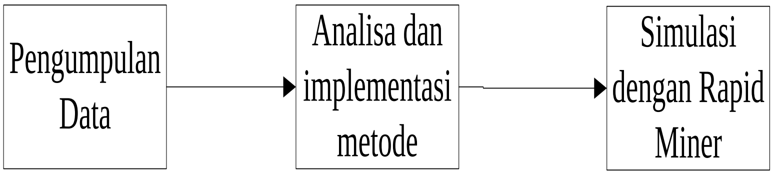


Fig 3. Chart of Research Methods

Based on the research methodology in the research methodology, the work stages consist of defining problems, problem analysis, determination of research objectives, data collection, analysis and implementation of methods, decision-making, and system design. The data that has been collected will then be summarized as material for a motorcycle creditworthiness analysis. After the prospective debtor's feasibility analysis is carried out, the identification of the prospective debtor will be generated.

3.1 Knowledge Base

In this study, the determination of a prospective debtor's eligibility involves several interrelated criteria between one measure and another. This criterion is assumed to be the input variable, which is a character, capacity, capital, collateral, the condition of the output results from the assessment.

Tabel 1

Survey Feasibility Analysis

No	Criteria	Information	Value of Sub-criteria
1	Character	AO / RO does not enter OD2	25
		cooperative and not challenging to confirm	25
		Information on survey and checks for logical/suitable environment	25
		Well known to the surrounding community	25
2	capacity	Net income> 2 installments	60
		Spouses work/have income	40
3	Capital	has assets that have been paid off / own	60
		the number of employees / employee> 3	20
		businesses own property	20
4	The condition	age> 25 years / married	10
		years debtor 21-55 years	10
		dependents <3	10
		age guarantor 21 - 55 years	10
		lives in the current place of residence> 2 years	10
		house is not located in a low-risk area	15
		self-employed/professional with a length of business> 2 years	10
		has a business license/practice	10
		contract business premises> 1 year	5
		strategic business locations	10

Ket.

- AO: Aditional Order(Old creditisstillrunning/existing, thenapplyfor new credit again)
- RO: Repeat Order(The previous motorbike loan has beencompleted/paidoff,thenappliesfor a motorbike loanagain)
- OD: Over Due(latein paying over 30 days / 2 months)

And the following is the transformation table for the prospective debtor's assessment

Tabel 2

Survey Feasibility Analysis

No	Criteria	Value	Information
1	Character	0-25	Low
		50	Medium
		75-100	High
2	Capacity	0-40	Low
		60-100	High

No	Criteria	Value	Information
3	Capital	0-20	Low
		40	Medium
		60-100	High
4	Conditions	0-45	Low
		50-70	Medium
		75-100	High

3.2 Debtor Data

The steps that must be taken to determine a prospective debtor are to use the rule table data usingexisting data (data-driven) alongwithregulatory data based on debtor feasibility survey data, along with data on debtors who already have a decision from CS finance,

Tabel 3
Conversion Table Ratings Borrower

No. Register Debtor	Character	Capacity	Capital	Conditions	Decree
204	High	High	High	Medium	Worth
438	Medium	High	High	Moderate	Net Worth
355	Low	High	Medium	High	Unfit
303	High	Low	Low	Medium	Worth
492	High	High	High	High	Worth
253	High-	High	low	Medium	Worth
298	High	High	High	Medium	Worth
303	Low	Low	High	Medium	Unfit
355	Medium	High	High	High	Worth
194	Low	High	High	Low	Unfit
456	High	High	Low	Low	Worth
267	Medium	Low	High	Medium	Unfit
150	High	High	Low	High	Worth
223	Medium	Average	Average	Average	Worth

3.3 Debtor data identification using the Naïve Bayes algorithm.

In identifying data in this study, severalsteps aretakentoobtain data that can be processed using the Naïve Bayes method. Someoftheactionstaken in data identification are the collection of data or information collected.

3.4 Probability of Occurrence

The probability of occurrence ispredictingtheeventofanoutcome based on the grouping of classes.

A. Character Probability The

following is the Character Probability table

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Tabel 4
Character Probability

Character	Numberofevents		Probability	
	Eligible	Ineligible	Eligible	Ineligible
Low	0	3	0/8	3/6
Medium	2	2	2/8	2/6
High	6	1	6/8	1/6
Total	8	6	8/8	6/6

The value of 0 in the low row of the number of events is feasible to be taken from table 5.4 where the criteria for low-value characters get a decent decision as much as 0. The valueof 3 in thecolumn'sshortrowthenumberof incidents is not feasible to be taken from table 5.4. The criteriaforlow-value characters get a decision not possible as much as 32, like wise for the short row and column height feasible and not feasible. Andthevalueof 0/8 in the low and viable row probability column is taken by the amount of 0 is taken from the appearance of a lowcostfor a feasible decision on the character criteria in table 5.4. The valueof 8 is taken from the possible Decision that appears in table 5.4,whichis as many as 8, andhowtogetthe medium and high row infeasible probability column

B. CapacityProbabilities

The character following the probability table:

Tabel 5
ProbabilityCapacity

Capacity	Number of Events		Probability	
	Eligible	Ineligible	Eligible	Ineligible
Low	0	3	0/8	3/6
Average	1	0	1/8	0/6
High	7	3	7/8	3/6
Total	8	6	8/8	6/6

The value of 0 in the low row of the number of events is feasible to be takenfromtable 5.4,where the criteria for low-value capacity get a decent decision as much as 0. The valueof 3 in the column's short row, the numberof incidentsis not feasible, is taken from table 5. The criteriafor the low-value capacity to get a decisionis not possible as much as 3, so it is for the row height column is feasible and not feasible.

And the value of 0/8 in the feasible and low row probability column is taken by the amount of 0 is taken from the appearance of a lowcostfor a possibledecisiononthe capacity criterion in table 5.4. The value eight is taken from the feasible Decision that appears in table 5.4,whichis as many as 8, and how to get high row unworthy probability column value.

C. Probability Kapital

Below is a table of probabilities Kapital

Tabel 6
KapitalProbability

Kapital	Number of Events		Probability	
	Eligible	Ineligible	Eligible	Ineligible
low	3	1	3/8	1/6
medium	1	1	1/8	1/6
high	4	4	4/8	4/6
Total	8	3	8/8	6/6

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

The value of 3 in theshortrowofthe number of events is feasible to be takenfromtable 5.4,wherethecriteria for low-value capital get a proper decision as much as 3. The valueof 1 in thecolumn'sshortrow, thenumberofincidentsis not feasible,istakenfromtable 5.4. The criteriaforlow-value capital get a decision not possible as much as 1, likewise for medium rows and column height feasible and improper.

And the value of 3/8 in the feasible probability column and the short row is taken by the amount three is taken from the appearance of a lowcostfor a possibledecisiononthe capacity criterion in table 5.4. The value eight is taken from the feasible Decision that appears in table 5.4,whichis as many as 8, and how to get the medium and high row infeasible probability column.

D. ProbabilityConditions

Probability Table The following condition

Tabel 7

Table Condition Probability

Condition	Number of Events		Probability	
	Eligible	Ineligible	Eligible	Ineligible
Low	1	1	1/8	1/6
Medium	4	4	4/8	4/6
High	3	1	3/8	1/6
Total	8	6	8/8	6/6 The

Value 1 in the low row column number of incidents is feasible to be taken from table 5.4, where the criteria for low-value conditions get a proper decision as much as 1. The value 1 in the low row-column the number of incidents is not feasible to be taken from table 5.4, wherein criteria with low-value conditions get an ineffective decision as much as one and moderate rows and column height possible and not feasible.

And the value of 1/8 in the feasible probability column and the short row is taken by the amountone is taken from the appearance of a low cost for a possible decision on the condition criteria in table 5.4. The value eight is taken from the feasible Decision that appears in table 5.4, which is as much as 8, and how to get the medium and high row infeasible probability column.

E. Decision Probability

Below is the recap distribution table of theprobability **Decision**

Tabel 8

Decision Probability Decision

Decision	Probability		
Eligible	8	8/14	0.571
Ineligible	6	6/14	0.429

3.5 Calculating the value *Likelihoodstep*

The nextis to calculate thevalue *Likelihood*; the *Likelihood* is the word equation of the probability used as a parameter, an example of a case with potential debtor data as follows:

A. Characters: Low
B. Capacity: High
C. Capital: Low
D. Condition: High

From the data above, thevalue can be calculated, the *likelihood* probability of appearance of each of the criteria, such as:

Likelihood

Eligible = $\frac{0}{8}$x $\frac{7}{8}$x $\frac{3}{8}$x $\frac{3}{8}$ x $\frac{8}{14}$

= 0 x 0.874 x 0.375 x 0.375 x 0.5714

= 0

likelihood

Ineligible = $\frac{3}{6}$x $\frac{3}{6}$x $\frac{1}{6}$ x $\frac{1}{6}$ x $\frac{6}{14}$

$= 0.500x0,500 \text{ x } 0.1667 \text{ x } 0.1667 \text{ x } 0.4286$

$= 0.0030$

3.6 Calculating Prior Probabilities

Calculating the probability value is done by calculating the value by normalizing the *Likelihood* so that the amount obtained is 0.01875value is *Likelihood* feasible and 0.0000 *likelihood is* not feasible, how to calculate this probability value is as follows:

Probability Eligible $= \frac{0}{0+0,0208}$

$= \frac{0}{0,0208} = 0$

$= 0\%$

Probability InEligible $= \frac{0,0208}{0+0,0208}$

$= \frac{0,0208}{0+0,0208} = 1$

$= 100\%$

From the final result, the feasible probability value is 0% and the unfeasible probability is 100%, so the high probability will be superior, it can be concluded that the prospective debtor is eligible to become a debtor in CS. Finance

3.7 Simulation with Rapid Miner

Here is a calculation with a rapid miner to ensure manual counting and the program that will be made in accordance with naïve Bayes calculations and testing the accuracy of the naïve Bayes algorithm on the eligibility data of prospective debtors.

A. Creating Training data

Here is a design for creating naïve Bayes training data on a rapid miner according to manual calculations

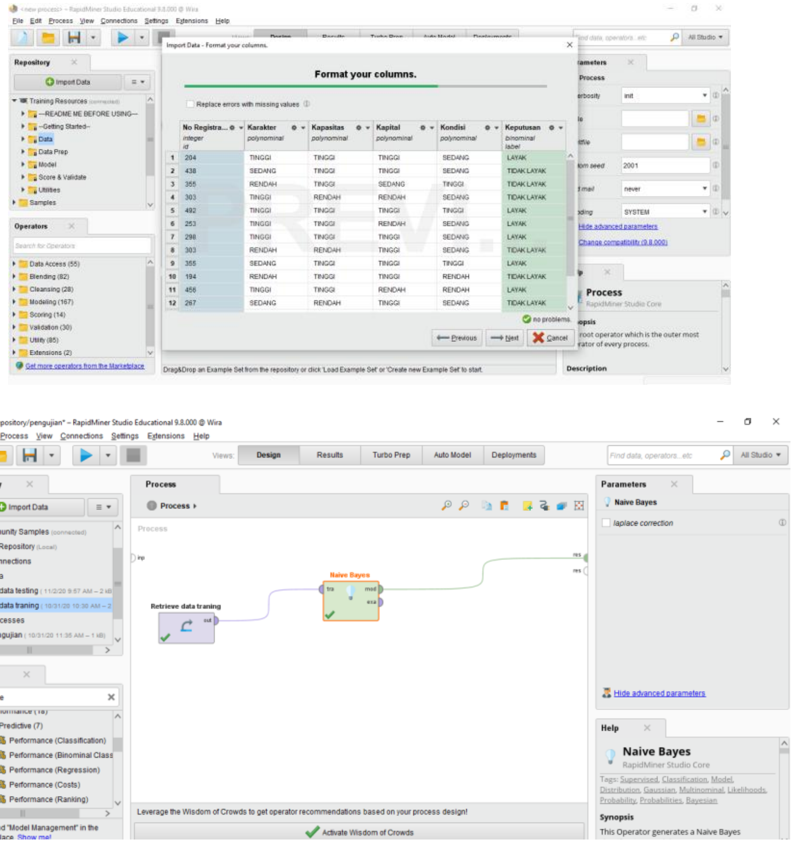


Fig 4. Designs for creating naïve Bayes training data on a rapid miner

Here are the results of the training data.

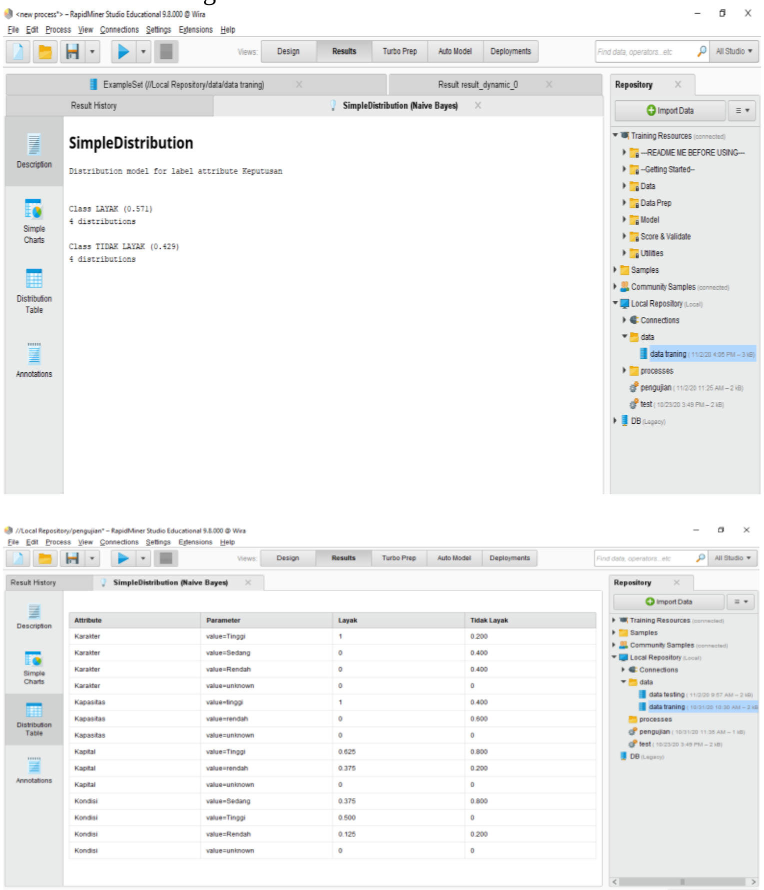
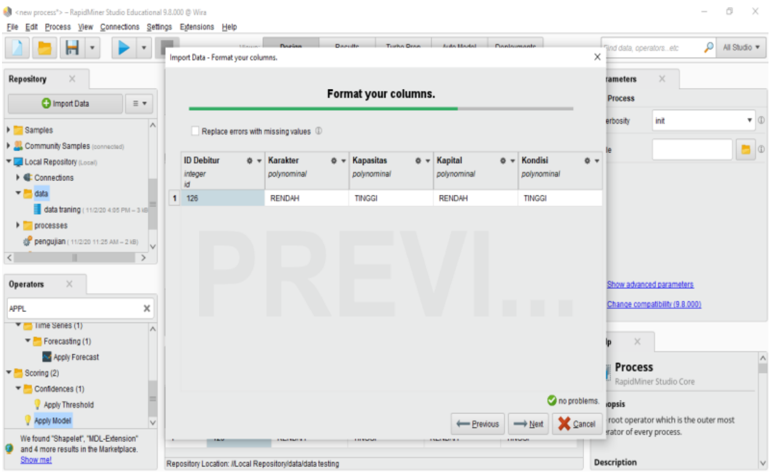


Fig 5 is the result of naïve Bayes training data on rapid miners

B. Creating Testing data on Rapid Miner

The following is a design for creating naïve Bayes Testing data on Rapid Miner according to manual calculations



Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

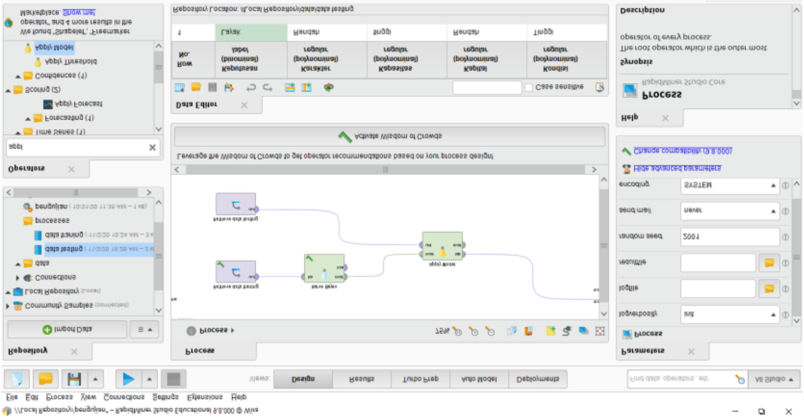


Fig 6. The design of testing the data naïve Bayes on rapidminner

Here is the result of a data testing

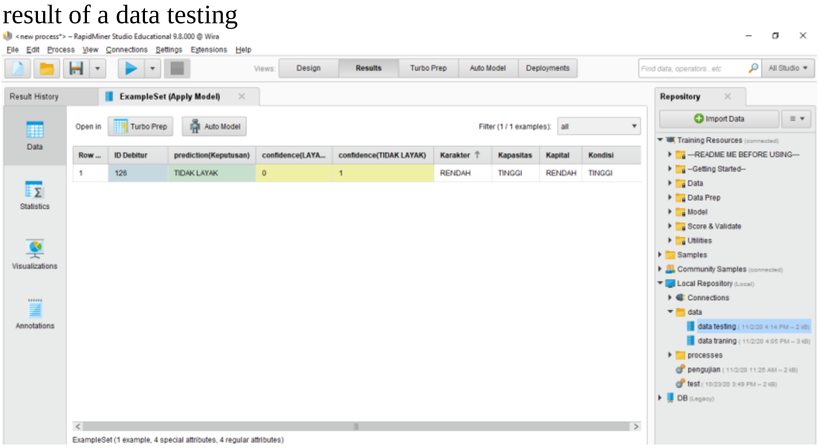


Fig 7. The results predicted naïve Bayes in accordance with the manual calculation

Of the above picture of the results of prediction (Decision) declared debtor is declared NOT WORTH with confidence DECENT 0 and confidence NOT WORTH 1 results this is according to manual calculation. Furthermore, 5 data testing will be made. To find out the accuracy of this naïve Bayes method for training data

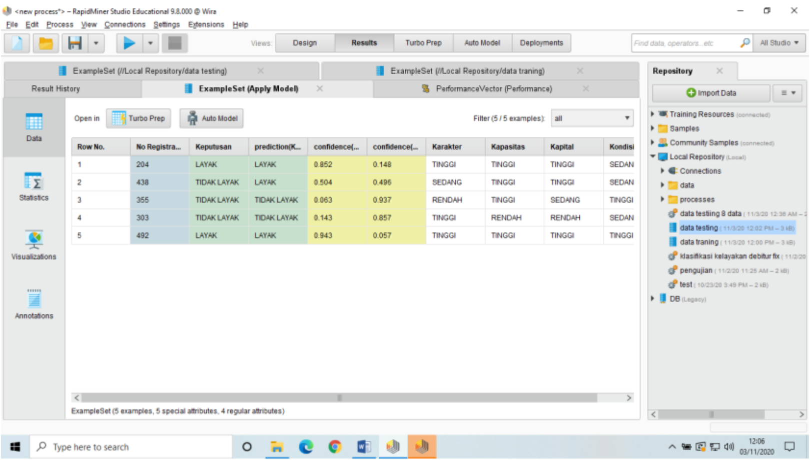


Fig 8. Prediction results with 5 data testing.

The data above is the prediction result with 5 data; there is one prediction that is not suitable, namely on row 2. The Decision gets an incorrect value feasible, but the prediction gets a decent amount. Here are the results of the accuracy from testing 5 data testing.

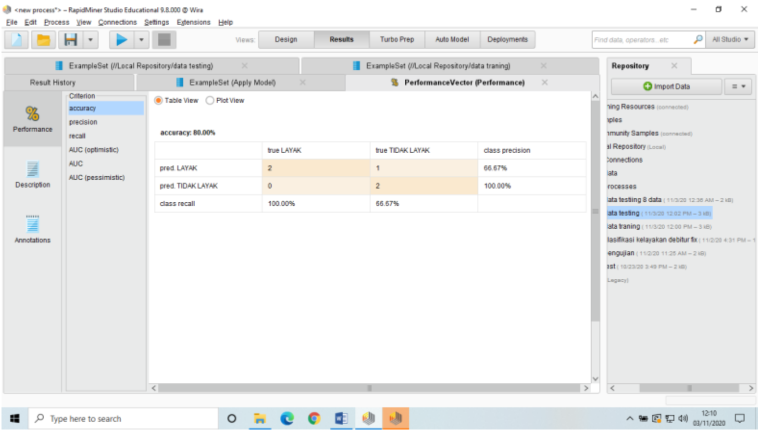


Fig 9. Performance accuracy five testing dataaccuracy

As in the picture above, thevalue in testing fiveof this data is 80%.
Next will be tested with the number of testing data 10.

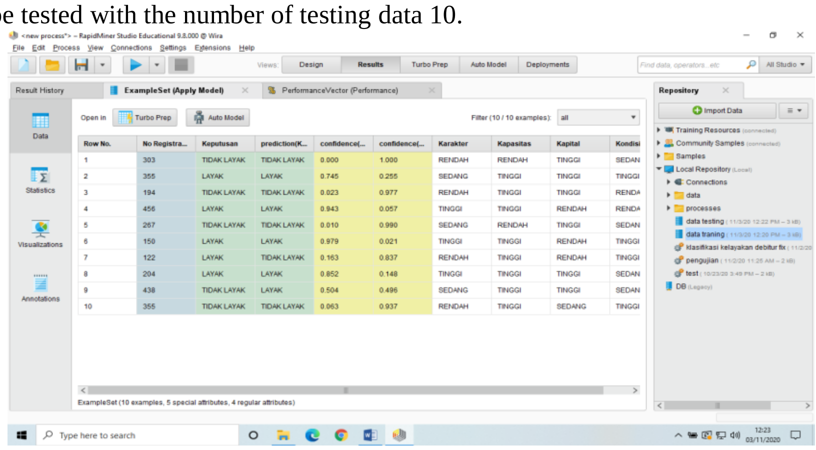


Fig 10. Prediction results with ten testing data

Following the image above, there are two prediction errors, namely at 7 and 9; after that, the ten testing data's accuracy will be seen.

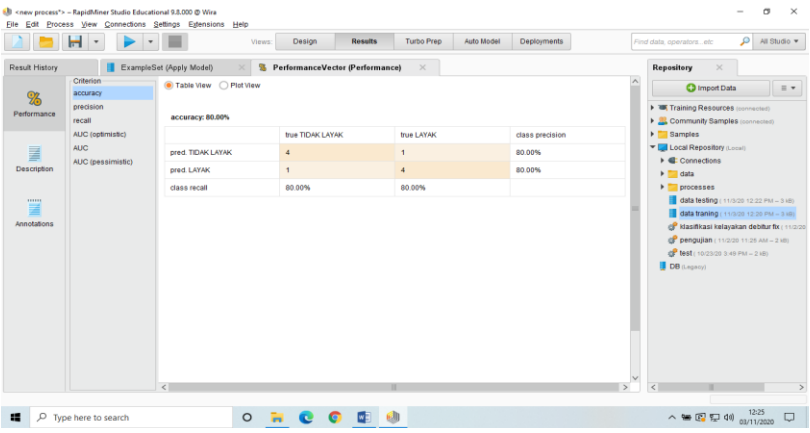


Fig 11. is theaccuracy of the testing data with 10 data

In accordance with the picture above that the naïve algorithm for predicting potential debtor eligibility reaches 80 % likewise for 5 data also yields the same thing a is 80% so that this algorithm is feasible to use because the accuracy rate is above or equal to 80%.

3.8 System Implementation

To make it easier for users to use the eligibility classification system for prospective debtors in CS Finance, it was built using the PHP programming language and WEB-based. Here is the implementation

A. Login Form

To be able to enter the main page, you must first pass the login form, the user is required to fill in the username and password correctly, here is the display of the login form

Fig 12. Display the login form

B. the Main Page

After successfully logging in, the user will then be directed to the main page/dashboard of the system; on this page, there is a choice of training data or testing data; here is the main page's display.

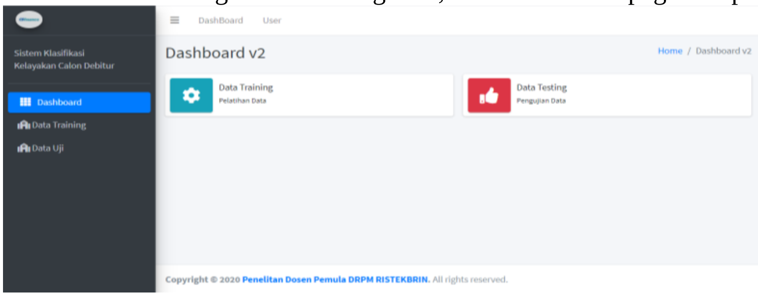


Fig 13. Main page

C. Input Data Training Form

Next is the Input training data form. The training data input form functions to create and enter training data as a reference for predicting prospective debtors' feasibility using the naïve Bayes algorithm.

Fig 14. Input Data Training Form

D. Data Training Form

Furthermore, the data that has been inputted will enter the training data, which can be seen as in the following

Fig 15. List of Training Data

E. Testing Data Input Form

After completing the training data, start entering the testing data to see predictions on the testing data according to the data in the training data; here is a display of the testing data input form

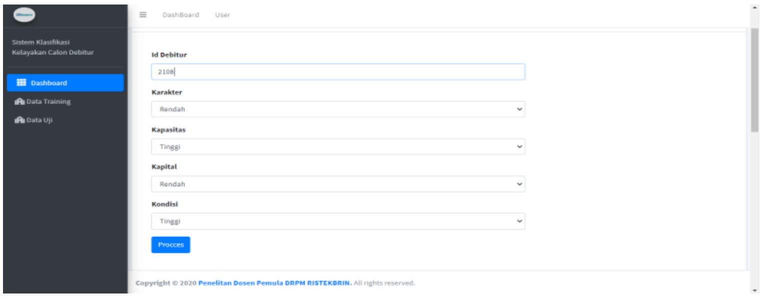


Fig 16. Testing data input form

F. Data Testing Results Form

After all testing data input forms are filled according to the testing data in manual calculations, then by pressing the process button, the prediction results will appear from the prospective debtor's classification system on CS Finance with naïve Bayes algorithms.

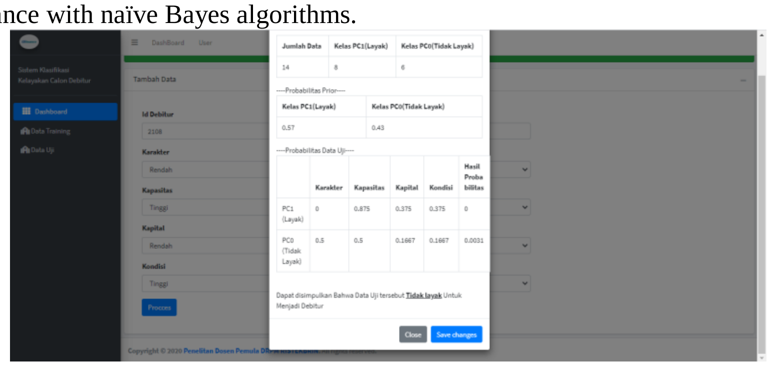


Fig 17. Prediction results from testing data.

There is information on the number of preliminary priority data in the testing result form and the probability of testing data according to manual calculations. The conclusion is that the prospective debtor is not eligible to become a debtor.

4. Conclusion

Following are the conclusions of this study:

a) The naïve Bayes algorithm has been implemented with four attributes to obtain the debtor's eligibility class's prediction class.

b) The application implemented can help the marketing survey section at CS Finance predict prospective borrowers' feasibility.

c) The system has been tested and obtained a good performance accuracy to predict prospective debtors' feasibility on CS Finance with an accuracy value of 80%.

5. References

Journal Article

[1]. Ardun dan Windha M.P, (2015) *"Impelementasi data mining Untuk Prediksi kelayakan kredit Nasabah Pada BMT Bumi Mizan Sejahtera Yogyakarta menggunakan metode C.45"* Naskah Publikasi STMIK AMIKOM Yogyakarta

[2]. Arga, Lestari dan Sutanta, (2017) *"Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritme C4.5 (Studi Kasus: Jurusan Teknik Informatika, Institut Sains & Teknologi Akprind Yogyakarta)"* Jurnal SCRIPT Vol. 5 No. 1 Desember 2017 ISSN: 2338-6304

[3]. Deny C.M dan Achmad W.K, (2015) *"klasifikasi data debitur untuk menentukan Kelayakan kredit dengan menggunakan metode Naive bayes"*Skripsi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang

[4]. Ekawijana, A., &Heryono, H. (2016). *"Composite Naive Bayes Classification and semantic method to enhance sentiment accuracy score "*International Conference on Cyber and IT Service Management, 2016 4th, ISBN Elektronik: 978-1-4673-8443-8, DOI:10.1109/citsm.2016.7577591

[5]. J. Soares, A.J. Santoso, Suyoto (2017) *"The ApplicationOf Data mining For Old Student To Prediction Studies Using Naive BayesAndAdaboostMethod"*International Journalof Advances in Electronics and Computer Science, Volume-4, Issue-11, ISSN: 2393-2835

[6]. Nofriansyah. D, Erwansyah. K, Ramadhan. M, (2016) *"Penerapan Data mining denganAlgoritma Naive Bayes ClasifieruntukMengetahuiMinatBeliPelangganterhadapKartu Internet XL (StudiKasus di CV. Sumber Utama Telekomunikasi)"*Jurnal SAINTIKOM Vol.15, No. 2, Mei 2016, ISSN : 1978-6603

[7]. Putri. A. N (2017) *"Penerapan Naive Bayesian Untuk Perankingan Kegiatan Di Fakultas Tik Universitas Semarang"* Jurnal SIMETRIS, Vol 8 No 2 November 2017 ISSN: 2252-4983.

[8]. Perwira, Y. (2018). *Penerapan Data Mining Dalam Menentukan Calon Credit Marketing Oficer (CMO) Baru Di PT. WOM Finance Dengan Menggunakan Metode Asosiasi* , *2*, 151-160. http://www.e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/721

[9]. Zaki, M. J., &Wagner, Meira, J. (2014). *Data mining and analysis: Fundamental concepts and algorithms*. Cambridge University Press.

[10]. Prasetyo, Eko. (2014). Data Mining - Mengolah Data Menjadi Informasi menggunakan Matlab. Yogyakarta: CV. Andi Offset.