Implementation Of The Naive Bayes Algorithm For Cooperative Loan Prediction At Xyz Savings And Loan Cooperative

Rini Amalya Br Hutagalung [1], Mhd. Zulfansyuri Siambaton [2], Tasliyah Haramaini [3]

[1]Universitas Islam Sumatera Utara, riniamalya03@gmail.com.

[1]Universitas Islam Sumatera Utara, zulfansyuri@ftuisu.ac.id. [1]Universitas Islam Sumatera Utara, tasliyah@ftuisu.ac.id

ARTICLE INFO

ABSTRACT

This study aims to implement the Naive Bayes algorithm to predict the payment feasibility of prospective loan applicants at XYZ Cooperative. The data includes applicant profiles such as age, job status, income, and willingness to participate in preparatory training held by the cooperative. Naive Bayes was selected due to its high efficiency in classification tasks. The results show that the algorithm effectively classifies applicants into eligible and non-eligible categories with satisfactory accuracy, aiding decision-making processes in the cooperative.

Naive Bayes, Loan Prediction, Cooperative, Classification, Data Mining.

INTRODUCTION

Credit is the provision of funds or comparable claims, which are made based on a loan agreement between a bank or financial institution and another party, the borrower is required to repay his debt within an agreed time period. (Borman & Wati, 2020)

Credit analysis is crucial, as one of the contributing factors to bad debt is a lack of due diligence on the part of cooperatives in granting loans. This problem can be addressed by accurately identifying and predicting customers before granting loans, utilizing historical loan data. (Rifqo & Wijaya, 2017)

Naive Bayes is a simple probability classification method that can calculate all possibilities by combining a number of combinations and the frequency of a value from a database. An algorithm utilizes Bayes' theorem to estimate all independent and mutually exclusive attributes that a value can assign to a variable class. (Rachman & Handayani, 2021)

The Naive Bayes method has the advantages of fast calculation, a simple algorithm, and high accuracy. The Naive Bayes method only requires a small amount of training data to determine the parameter estimates needed in the classification process. The Naive Bayes Classifier algorithm is easier to use because it has a short calculation process. (Putro et al., 2020)

According to research conducted by (Atina et al., 2024) entitled "Application of the Naive Bayes Algorithm in a Cooperative Loan Eligibility Determination System," the study concluded that the Naive Bayes method is an appropriate method for determining eligibility. The Pangarso Budi Utomo Roso Manunggal Jati Community Organization Cooperative built a web-based cooperative loan eligibility determination system using the PHP programming language.

Based on research conducted by (Hanifatun & Zahrotun (2025) on "The Application of Data Mining in Determining Customer Creditworthiness," Farras and his colleagues concluded that the results of this research have produced an

application that can predict creditworthiness for BUMDes Gedong Gincu customers using the Naive Bayes approach. Test results using 186 data sets using the confusion matrix method obtained an accuracy of 67%. However, based on the SUS test by users, the result was 82.25. This indicates that this application is good and suitable for use by the Gedong Gincu Village-Owned Enterprise (BUMDes) because it has a SUS score above 80.

The research conducted by (Efendi et al., 2022), concerning the classification of installment payment performance using the Naive Bayes algorithm (Case Study: Customer Data of the Bina Bersama Sharia Financing Savings and Loans Cooperative), concluded that this customer data classification system is used to provide information regarding the payment status of prospective customers, whether they are current, substandard, or non-current in paying credit installments, using the Naive Bayes algorithm.

The role of cooperatives is crucial and strategic for the national economy, especially in facing competition in the current global market era, which also presents a new opportunity to foster development and achieve a just and prosperous society. (Lidwina Mulinbota Moron et al., 2023)

Customers can be likened to kings who must be served, but this does not mean that everything must be left entirely to them. Efforts to meet customer needs must be carried out while still considering the profitability aspect. Customer satisfaction is a crucial factor in maintaining sustainable customer relationships. (Reza Nurul Ichsan, Ahmad Karim, 2021)
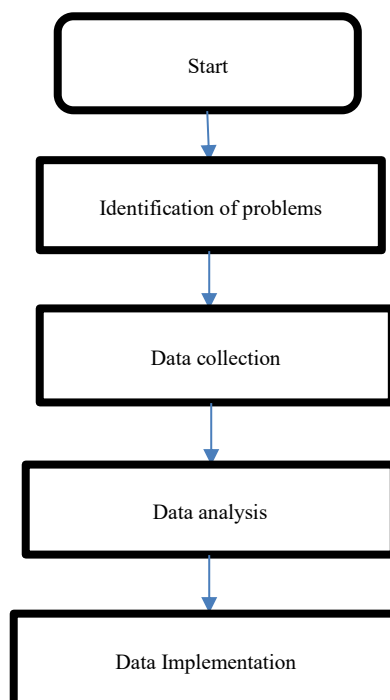
When customers are satisfied, they will remain loyal to the product they purchased. Therefore, banks must prioritize service quality to maintain customer satisfaction. Service quality has a significant influence on customer satisfaction, contributing to increased customer trust, facilitating the formation of customer loyalty, and providing benefits for the bank itself. (Satria & Diah Astarini, 2023)
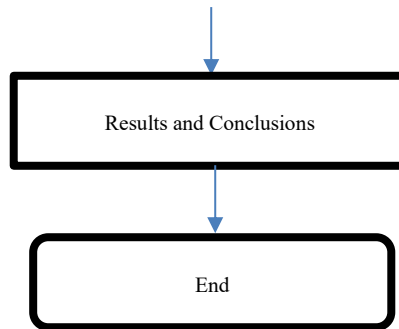
Customer loyalty is essential for banks that want to maintain their presence and be recognized in their business. Acquiring new customers is difficult, but retaining existing customers is even more challenging. The positive impact of retaining existing customers is the addition of new customers due to testimonials from existing customers. (Ahmadi, 2023)

The results of the primary data from the XYZ Savings and Loan Cooperative will be tested manually, and the remainder will be processed through a system designed as training data. To calculate the training data, the Naive Bayes classifier method is used, which involves several stages, each of which involves calculations using a predetermined formula. (Huda et al., 2023)

Furthermore, prediction can also be defined as a method of estimating or predicting quantitatively or qualitatively what will happen in the future based on relevant past data. This prediction is used to make systematic and pragmatic predictions based on relevant past data. (M. Afriansyah et al., 2024)

RESEARCH METHODS

Start

↓

Identification of problems

↓

Data collection

↓

Data analysis

↓

Data Implementation

Results and Conclusions

End

1. Problem Identification

The goal of this step is to clearly understand the problem and be able to address it effectively. This allows for a work plan and the necessary elements of this research to be determined.

2. Data Collection

The dataset used in this research was collected through observational methods, by observing and studying problems in the field related to the research object, namely information regarding prospective customer data.

3. Data Analysis

This stage is the data selection stage, aimed at obtaining clean data ready for use in the research. Selection was carried out by considering data consistency, missing values, and redundant data. Next, the data was processed using the PHP programming language with Performance, which serves as a tool to validate and ensure data reliability to ensure information accuracy.

4. Data Implementation

In the testing phase, each new data item is analyzed by calculating the joint probability for each class using the Naive Bayes formula. The class with the highest probability

value is selected as the classification result for the data. Thus, this algorithm simplifies the decision-making process in classification through an efficient and effective probabilistic approach.

5. Results and Conclusions

After implementing the PHP programming language, the next step is to review the results of the XYZ Savings and Loans application. Conclusions are then drawn based on the results obtained.

RESULTS AND DISCUSSION

The results of the probability calculations for the Eligible class show that the majority of data comes from females (81.5%), while only 18.5% come from males. For the Ineligible class, a similar distribution pattern is observed: 81.5% come from females and 18.5% from males.

Despite the smaller number of males, the probability calculations indicate that the proportion of ineligible males is slightly higher than the proportion of eligible females.

Table 1 Job Probability

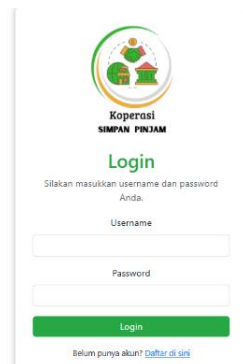| Job Position	Eligible	Not Eligible
First Expert	727	183
First	24	12
Skilled	39	3
Beginner Implementer	0	1
Implementer	0	3

Junior	0	8

The majority of potential customers categorized as Ineligible came from the First Expert position, at 87%. This position also constituted the largest overall group in the data, so this dominance is quite reasonable. The First Position also contributed a significant proportion, at 0.57% of the total Ineligible customers.

Nevertheless, the presence of data with these positions remains important because it can influence the system's prediction results, especially if the position attribute is used as a primary parameter in the classification process.

The position with the highest probability of being in the Eligible class was the First Expert with 87%, followed by the First position with 87%.

Within the Ineligible class, the largest distributions were still held by the First Expert (87%) and First (0.57%). Other positions, such as Skilled (0.14%), Junior (0.38%), and Beginner Implementer (0.04%), had smaller distributions but still influenced the classification model.

Figure 4.2 Login display on the prospective customer prediction system

By pressing the Login button, users will be able to enter the XYZ Savings and Loans Cooperative's customer prediction application system. This system is used to help the cooperative identify and predict potential customers to join and use the savings and loan services provided by the XYZ Savings and Loans Cooperative.

Figure 4 3 Registration Page.

This image displays a web page with the title "Register" at the top. This page is a registration form for new users who wish to join and log into the savings and loan cooperative system to predict potential customers at XYZ.

The purpose of this page is to facilitate the registration process for prospective customers who wish to use the savings and loan services provided by XYZ Savings and Loan Cooperative. Through this registration form, prospective customers can register and gain access to the cooperative's customer prediction system.

Naive Bayes Calculation Results

The final result of the system to predict prospective customers at the XYZ savings and loan cooperative is a system that applies the Naïve Bayes Algorithm to calculate the number of probabilities for the eligible class and the Ineligible class, where this class is very necessary for the XYZ Savings and Loan Cooperative in predicting prospective customers who are eligible to receive loans from the Cooperative.

Of the 1000 datasets uploaded to the XYZ savings and loan cooperative, there are 790 people who have and fall into the eligible criteria, while for the Ineligible criteria, there are 210 people who fall into the ineligible criteria, where prospective customers who meet the eligible criteria have a minimum age of 21 years and have an income of 490000.

Figure 4.8 Naive Bayes Calculation Results

CONCLUSION

1. The implementation of the Naive Bayes algorithm in determining the eligibility of prospective customers at a savings and loan cooperative was successful. Using a dataset of 1,000 prospective customers, the Naive Bayes algorithm was able to classify the data into two categories: "Eligible" and "Not Eligible," based on attributes such as age, gender, and salary.

2. The eligibility criteria established in this study were age between 21 and 45 years and a minimum salary of IDR 4,900,000. Based on these criteria, the classification results showed that 790 data (79.1%) fell into the "Eligible" category and 210 data (20.9%) fell into the "Not Eligible" category.

3. The calculation results showed P(Eligible) = 0.791 and P(Not Eligible) = 0.210. This indicates that the majority of prospective customers met the eligibility criteria.

REFERENCES

Borman, R. I., & Wati, M. (2020). Application of Data Maining in Classifying Data of Members of Sejahtera Cooperative of Bandarlampung Using Naïve Bayes Algorithm. Scientific Journal of the Faculty of Computer Science, 09(01), 25–34.

Rachman, R., & Handayani, R. N. (2021). Naive Bayes Algorithm Classification in Predicting the Smoothness of MSME Terrace Rental Payments. Informatics Journal, 8(2), 111–122. https://doi.org/10.31294/ji.v8i2.10494

Rifqo, M. H., & Wijaya, A. (2017). Implementation of the Naive Bayes Algorithm in Determining Credit Granting. Pseudocode, 4(2), 120–128. https://doi.org/10.33369/pseudocode.4.2.120-128

Putro, H. F., Vulandari, R. T., & Saptomo, W. L. Y. (2020). Application of the Naive Bayes Method for Customer Classification. Journal of Information and Communication Technology (TIKomSiN), 8(2). https://doi.org/10.30646/tikomsin.v8i2.500

Atina, V., Permatasari, H., Srirahayu, A., Duta, U., & Surakarta, B. (2024). APPLICATION OF THE NAIVE BAYES ALGORITHM IN A DETERMINATION SYSTEM. November, 461–471.

Hanifatun, F., & Zahrotun, L. (2025). Jurnal Informatika: Jurnal Pengembangan IT: Application of Data Mining in Assigning Customer Credit Eligibility at Gedong Gincu Village-Owned Enterprises Using the Naive Bayes Method. 10(1), 226–236. https://doi.org/10.30591/jpit.v9ix.xxx

Efendi, D. M., Mintoro, S., . S., Lubis, S. H., & Lestari, S. (2022). CLASSIFICATION OF INSTALLMENT PAYMENT PERFORMANCE USING THE NAIVE BAYES ALGORITHM (Case Study: Customer Data of the Bina Bersama Sharia Financing

Savings and Loans Cooperative). Journal of Information and Computers, 10(1), 57–61. https://doi.org/10.35959/jik.v10i1.305

Lidwina Mulinbota Moron, Henrikus Herdi, & Yoseph Darius Purnama Rangga. (2023). The Influence of Work Culture on Employee Performance at the Ikamala Savings and Loans Cooperative. Jurnal Kompetitif, 12(1), 1–14. https://doi.org/10.52333/kompetitif.v12i1.56

Reza Nurul Ichsan, Ahmad Karim. (2021). Service Quality on Customer Satisfaction at PT. Jasa Raharja Medan. Journal of Social and Humanities Education Research, 6(1), 54–57. https://doi.org/10.32696/jp2sh.v6i1.729

Satria, F., & Diah Astarini. (2023). The Effect of Service Quality and Customer Satisfaction on Customer Loyalty Mediated by Customer Satisfaction. Trisakti Journal of Economics, 3(1), 1911–1924. https://doi.org/10.25105/jet.v3i1.16124

Ahmadi, M. A. (2023). Strategic Management in Increasing Bank Customer Loyalty. Nitro Journal of Financial Banking Management, 7(1), 15–25. https://doi.org/10.56858/jmpkn.v7i1.167

Huda, M., Fawaid, A., & Slamet. (2023). Implementation of Behaviorist Learning Theory in the Learning Process. August, 1(4), 64–72. https://doi.org/10.51903/pendekar.v1i4.291

M. Afriansyah, Joni Saputra, Ardhana, V. Y. P., & Yuan Sa'adati. (2024). An Efficient Naive Bayes Algorithm for Classifying Raja Bananas Based on Color Features. Journal of Information Systems Management and Digital Business, 1(2), 236–248. https://doi.org/10.59407/jismdb.v1i2.438