TSABIT
Journal of Computer Science
https://journal.ilmubersama.com/index.php/tsabit

Click here and write your Article Category

Implementation of the Naïve Bayes Algorithm in Credit Card Approval Classification at Bank Central Asia (BCA)

Surya Lesmana [1], Firahmi Rizky [2]

[1,2] Department of Information System, Faculty of Computer Science and Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan, 20238, North Sumatra, Indonesia

ARTICLE INFORMATION

Received: February 00, 00
Revised: March 00, 00
Available Online: April 00, 00

KEYWORDS

Naïve Bayes Algorithm, BCA Bank Credit Card Approval, Use Training Set

CORRESPONDENCE

Phone: +62822 8415 5036
E-mail: steelsurya70@gmail.com

A B S T R A C T

Credit card approval is a financial facility that allows a person or business entity to borrow money to purchase products and repay it within a mutually agreed time. Researchers used a credit card approval dataset (Clean Data) of 200 data with 6 atributes. In conducting the analysis, researchers used the WEKA tools. In this study the author wants to classify the data to predict whether or not it is appropriate to grant a credit card. The method used is the Naïve Bayes Algorithm method. In this study, the data was split by 60:40, 70:30, and 80:20. The author also obtained different accurasy result, namely in the Weka data Testing tool 20% using the use Training Set obtained the highest accuracy of 87.68% compared to the lowest accuracy in the 30% testing data using the use Training Set, which is 80.60% The Naïve Bayes Algorithm can be said to be one of the effective Algorithm both in the terms of calculations and final results where the test can be used as a basis for credit card approval.

INTRODUCTION

Granting a Bank Central Asia (BCA) credit card to a customer is difficult and carries a very high risk. Applications require clear data analysis, a clear record of problems with other banks, and documents that convince the bank to grant the card, such as employment information, monthly income, residence, age, and many other factors. The bank's analysis department must consider many factors. Consequently, many customers fail, and it's crucial to identify patterns in the information that can be used to predict the outcome [1,2].

The process of extracting this information is called data mining. Data mining is the process of analyzing data from different perspectives and summarizing it into important information. Data mining activities also include collecting and using historical data to find regularities, patterns, or relationships in the data [3]. Meanwhile, the process of predicting using classification methods. Classification is a very common task in everyday life, essentially involving dividing objects so that each is assigned to one of a number of complementary and exclusive categories known as classes. The classification method (classifier) is a systematic approach to building a classification model from a set of input data.

Each method uses a learning algorithm to obtain a model that best reflects the relationship between the set of attributes and class labels in the input data. With the development of information technology, particularly in the field of data mining, which has become widely used in information systems, several data mining classification methods are used for prediction, such as the Naïve Bayes Classifier, Decision Tree, Neural Network, K-Nearest Neighbor, and others. According to PP [4,5], the advantage of using Naïve Bayes is that this method requires only a small amount of training data to determine

https://doi.org/10.12345/xxxxx

the parameter estimates needed in the classification process. Naïve Bayes often performs much better in most complex real-world situations than expected.

The Naïve Bayes algorithm is a classification algorithm using probability and statistics methods invented by British scientist Thomas Bayes. It predicts future opportunities based on past experiences. The Naïve Bayes algorithm uses a simple probabilistic method that calculates a set of probabilities by summing the frequencies and combinations of values from a given dataset. Naïve Bayes classification is a simple classification algorithm whose performance is comparable to Decision Tree and Neural Network classification.

Based on the author's observations, to help Bank Central Asia (BCA) credit analysts increase their caution in making credit granting decisions, an application is needed to predict whether or not a credit applicant is worthy when analyzing credit based on data and information from previous credits.

METHOD

Naïve Bayes Algorithm

The Naïve Bayes algorithm is one of the algorithms found in classification techniques. Naïve Bayes is a classification method using probability and statics, proposed by British scientist Thomas Bayes. It predicts future probabilities based on past experience, hence the term Bayes' Theorem. This theorem is combined with Naïve Bayes, which assumes that the conditions between attributes are independent. Naïve Bayes classification assumes that the presence or absence of a particular characteristic of one class is unrelated to the characteristics of other classes.

This method, also known as the Naive Bayes Classifier, applies a supervised technique for future object classification by assigning class labels to instances/records with conditional probabilities. Conditional probability is a measure of the likelihood of an event occurring based on other events that have been (assumed, presumed, stated, or proven) on the classification of training data that has been labeled with a class. For example, a fraudulent transaction has been marked as transactional data. The Naive Bayes classifier demonstrates high accuracy and speed when applied to large databases. This method is often used in solving problems in the field of machine learning because this method is known to have a high level of accuracy with simple calculations.

The primary classification technique employed in this study is the Naïve Bayes (NB) algorithm, a probabilistic supervised machine learning method based on Bayes' Theorem. Naïve Bayes has become one of the most widely used classification algorithms because of its computational simplicity, fast training process, and ability to produce competitive classification performance even when working with relatively small datasets. In financial decision support systems, Naïve Bayes has been extensively applied to credit scoring, loan approval prediction, fraud detection, customer segmentation, and risk assessment due to its capability of estimating posterior probabilities efficiently.

In the context of credit card approval at Bank Central Asia (BCA), the Naïve Bayes classifier estimates the probability that an applicant belongs to either the Approved or Rejected class based on the applicant's financial profile and demographic characteristics. Rather than relying solely on fixed decision rules, the classifier learns statistical relationships from historical credit card application data. Consequently, every prediction is made by calculating the probability of each possible class and assigning the applicant to the class with the highest posterior probability.

One of the major assumptions underlying the Naïve Bayes algorithm is the conditional independence assumption, which states that every predictor variable contributes independently to the target class once the class label is known. Although this assumption is rarely fully satisfied in real-world banking datasets, numerous empirical studies have demonstrated that Naïve Bayes remains highly effective because it simplifies probability estimation while maintaining satisfactory predictive accuracy. The independence assumption significantly reduces computational complexity, allowing the algorithm to perform efficiently on high-dimensional datasets with numerous applicant attributes.

Dataset Review

The data collection method used in this study was to collect data from a Bank Central Asia credit card customer dataset located at Jalan Prof. H. M. Yamin No. 245 A-B, Sei Kera Hilir II, Medan Perjuangan District, Medan City, North Sumatra 20111, to be used as testing material in this study. Data collection was conducted to obtain the results and information

needed to achieve the research objectives. The author took a dataset from the Bank Central Asia credit card customer dataset at Jalan Prof. H. M. Yamin No. 245 A-B, Sei Kera Hilir II, Medan Perjuangan District, Medan City, North Sumatra 20111 and was given the title "Credit Card Approval" which amounted to 200 data with 6 attributes, namely, "Age, Income, Employment Status, Home Ownership, Credit Card History, Results" as labels (classes). In the study, using 2/3 for training data and 1/3 for testing data. In general, the proportion of training data and testing data is 60:40, 70:30, or 80:20, even 90:10 if the dataset is relatively large. Then the author will divide the data for training data as much as 60%, 70% and 80%, then testing data as much as 20%, 30% and 40%.

Classification

Classification is defined as a form of data analysis that extracts a model that will be used to predict class labels. Classes in classification are the most unique attributes of a dataset and the independent variables in statistics. H. Susana (2022) states: "Classification is a data mining framework that incorporates all concepts extracted from a training dataset into classes distinct from pre-existing data classes." L. Faisal and M.F. DYH Tanjung (2022) define: "Classification methods are the process of finding models or functions that explain or differentiate concepts or data classes, with the aim of predicting the class of an object whose label is unknown." Therefore, it can be concluded that classification is the grouping of classes based on existing data classes to predict the class of an unknown object.

RESULTS AND DISCUSSION

The results of the application of the Naive Bayes algorithm in credit card classification at Bank Central Asia (BCA) were obtained using web-based software that starts from logging in as in Figure 1.

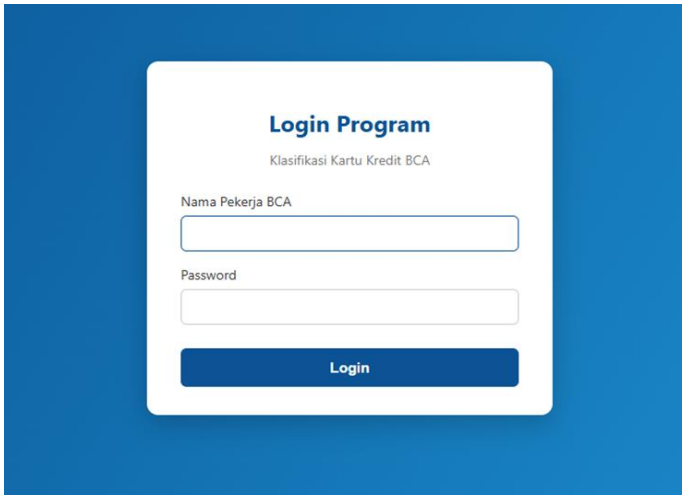

Figure 1. User Login

Figure 1. is the login form that serves as access to the dashboard or main menu. Users must correctly enter their username and password to access the dashboard, as shown in Figure 2.

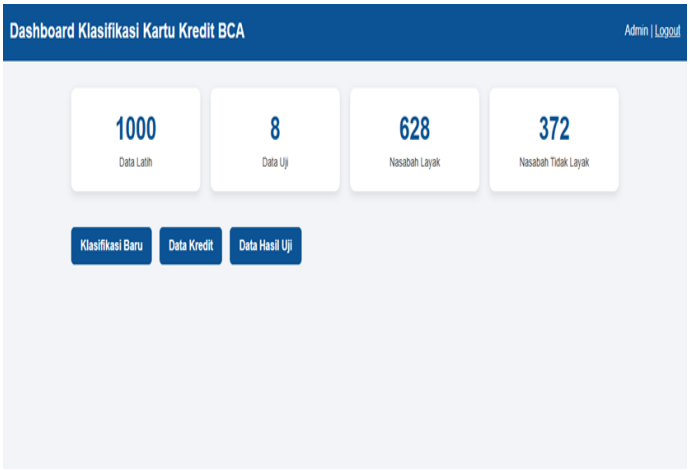

Figure 2. BCA Credit Card Dashboard

Figure 2 is a dashboard that displays several data processing menus, including new classification, credit data, and test results. Clicking on credit data will display the following screen: Figure 3.

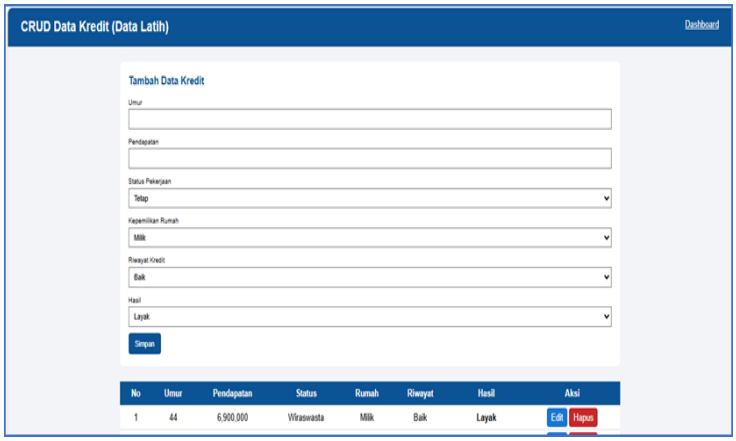

Figure 3. Credit Data

Figure 3 shows a credit data form that allows users to process credit data by saving, editing, and deleting it. If the user enters data in the text box and clicks save, the data is saved in the database. If the user clicks edit, the data editing form appears. If the user clicks delete, the data is deleted according to the selection. If the user clicks new classification, the screen shown in Figure 4 appears.

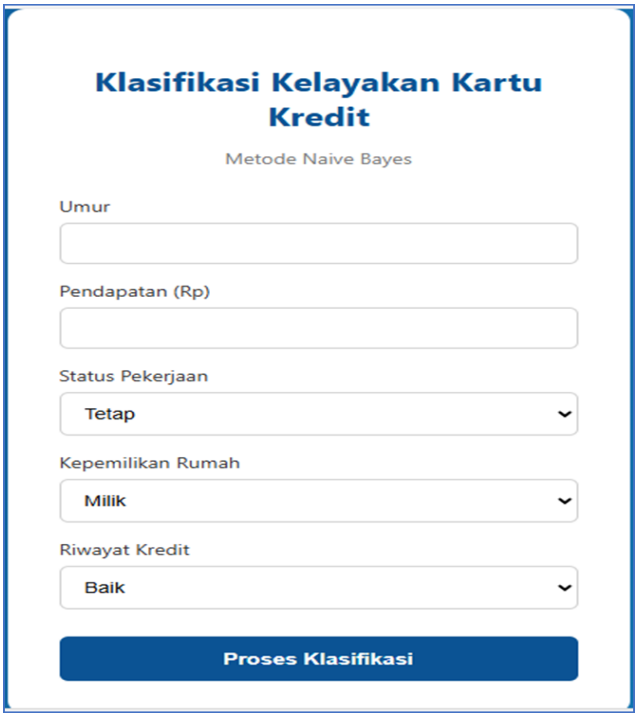

Figure 4. New Classification

Figure 4 represents a new classification. If the user enters all the data and then clicks Classify, it will be saved in the database table, and the new classification results will appear, as shown in Figure 5. Based on the calculation results, the posterior probability value for the Feasible class is greater than that for the Infeasible class. Therefore, the tested data is classified into the Feasible class. The summary of the classification data can be seen in Figure 4.5. This form shows all the test results obtained during the user's several tests.

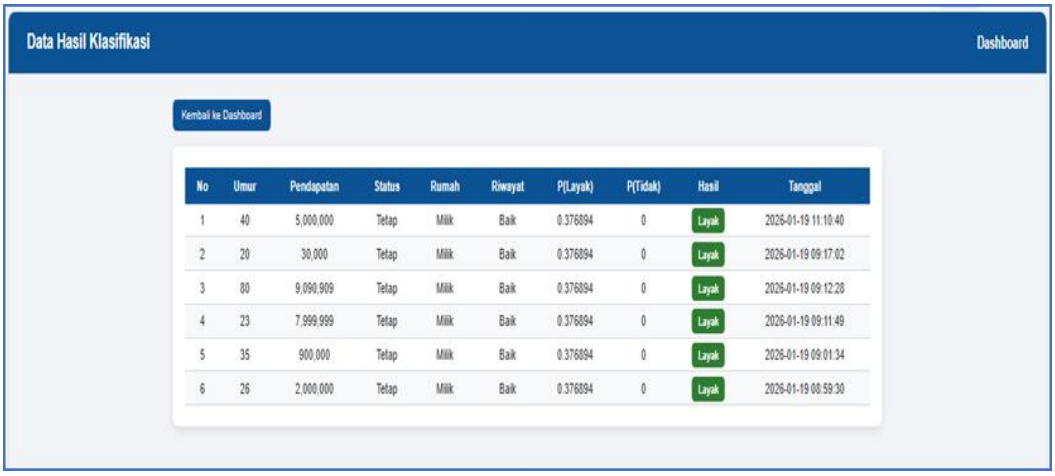

Figure 5. Naive Classification Results

CONCLUSION

Based on the research results and discussion regarding the application of the Naive Bayes method in the feasibility classification process, the following conclusions can be drawn: The Naive Bayes method is capable of conducting a systematic, probability-based feasibility classification process. This is demonstrated through the calculation of prior probability, conditional probability, and posterior probability, which results in objective classification decisions. The results of the prior probability calculations indicate that the Feasible class has a higher initial probability than the Infeasible class. This reflects the distribution of the training data, which is dominated by the Feasible class, thus influencing the final classification results. The obtained conditional probabilities indicate that the attributes Employment Status, Home Ownership, and Credit History have a significant influence on class determination. These attributes have a high probability value in the Feasible class and a zero value in the Infeasible class, making them strong indicators in the classification process. The posterior probability calculations indicate that the Feasible class has a significantly higher probability value than the Infeasible class. Based on these results, the tested data is classified into the Feasible class. This proves that the Naive Bayes method can provide clear and measurable classification decisions. Overall, the application of the Naive Bayes method in this study can help the feasibility decision-making process more accurately and efficiently compared to conventional methods, because it is based on historical data analysis and structured mathematical calculations.

REFERENCES

Journal Article from the Internet
[1] NA Rahman (2024) Waiting Period of Telkom University Student Alumni Using K-Nearest Neighbor and Naïve Bayes
[2] PP Wulan dan H Basri (2024) Analisis Sentimen Terhadap Layanan Nasabah Bank Menggunakan Klasifikasi Naïve Bayes. Vol 3 No 2
[3] H Aliyah, S Masyithoh (2024) Peran Teknologi Digital Dalam Meningkatkan Kualitas Pembelajaran Di Sekolah. Vol 1 No 2
[4] H Hartatik, AY Rukhmana, E Efitra, IR Mukhlis (2023) TREN TECHNOPRENEURSHIP : Strategi dan Inovasi Pengembangan Bisnis Kekinian Dengan Teknologi Digital [21-25]
[5] FJ Perwitosari (2023) Penerapan Data Mining Untuk Penilaian Kinerja Di PT Riksa Djaya Menggunakan Metode Naïve Bayes Classification. Vol 5 No 1 Sains dan Teknologi

[6] Zulherry, A., Riadi, I., & Umar, R. (2026). Anomaly Detection in Cloud Device-Based Information Technology Infrastructure Using Isolation Forest Algorithm. Journal Of Informatics And Telecommunication Engineering 9 (2)
[7] Sari, I.P., & Zulherry, A. (2025). Development of A Smart Monitoring System for IoT–Based Tide Observation. Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal, 6 (2), 18-23
[8] Zulherry, A., Gunawan, M., & Sari, I.P. (2025). Development of an Android-Based Smart Health Monitoring Device for Heartbeat Detection. Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal, 6 (2), 43-47
[9] J Jefri, Z Fatah (2025) Klasifikasi Data Mining Untuk Memprediksi Kelulusan Mahasiswa Menggunaka Metode Naïve Bayes. Vol 2 No 1
[10] H Susan (2022) Penerapan Model Klasifikasi Metode Naïve Bayes Penggunaan Akses Internet. Vol 4 No 2
[11] Ichsan, A., Zulherry, A., Lubis, T.A., & Shahnaz, B.A.Z. (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. IJATCoS: Indonesian Journal of Applied Technology, Computer and Science 2 (1)
[12] Amada, E.P., & Zulherry, A. (2025). Klasterisasi Minat Dan Bakat Siswa Menggunakan Metode X-Means Berbasis Web: Studi Kasus SMA Negeri 1 Hamparan Perak. sudo Jurnal Teknik Informatika 4 (4), 276-283.
[13] Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal 2 (2), 156-163
[14] DYH Tanjung (2022) Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Data Pengisian ATM. Vol 7 No 1 Universitas Potensi Utama.
[15] SA Amono (2023) Evaluasi Model Decision Tree Pada Keputusan Kelayakan Kredit. Vol 2 No 2
[16] M Clementine (2022) Prediksi Nasabah Bank Menggunakan Klasifikasi ID3. Vol 17 No 1
[17] Basri, M., & Zulherry, A. (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyahan. Ar-Rasyid: Jurnal Pendidikan Agama Islam, 5(1), 65-73.
[18] Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. 2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 486-491
[19] A Syakur, RP Putra (2024) Analisasi Metode Naïve Bayes Classifier Untuk Prediksi Persetujuan Kredit Vol 13 No 1
[20] N Suwimayanti, R Hadi (2024) Implementasi Metode Normalisasi dan Seleksi Fitur Dalam Optimasi Algoritma KNN Untuk Klasifikasi Data Bank. Vol 12 No 2
[21] RO Merdiyanti, K Kusrini (2023) Analisis Sentimen Pengguna Aplikasi Bank Syariah Indonesia Dengan Menggunakan Algoritma Support Vector Machine (SVM). Vol 4 No 1
[22] IM Panggabean (2023) Analisis Prediksi Kelayakan Nasabah Kredit Menggunakan Algoritma Rndom Forest Menggunakan PEGA dan WEKA. Vol 5 No 2