Implementation of Naive Bayes Method for Granting Fisherman Business Credit

[1]I Gede Totok Suryawan, [2*]I Kadek Surya Arimbawa, [3]I Gede Iwan Sudipa
[1,2*,3]Fakultas Teknologi dan Informatika , Program Studi Teknik Informatika, Institut Bisnis dan Teknologi Indonesia
Email: totok.suryawan@gmail.com[1], suryaarimbawa@gmail.com[2*],iwansudipa@instiki.ac.id[3]

Keywords

Decision Support System, Fisherman Business Credit, Naive Bayes, Village Credit Institution

Abstract. The Lembaga Perkreditan Desa (LPD) is a village financial institution engaged in the savings and loan industry, provides credit as one of its services. Midway through 2021, the Jimbaran Traditional Village LPD issued a new credit product, the Fisherman Business Credit (KUN), to assist the Jimbaran villagers who are experiencing economic hardships due to Covid-19. The rapid increase in credit applications at the Jimbaran Traditional Village LPD, particularly for fisherman business loans, necessitates a more comprehensive analysis of incoming fisherman business loan application data to prevent a repeat of the poor credit decisions from the previous year. On the other hand, the community, especially those whose credit applications are denied, mandates transparency in the selection process for credit assessment. Current credit evaluation procedures are rife with subjectivity, necessitating software that can provide transparency by generating scores from each existing credit application using scientific methods. In this study, a credit granting decision support system was developed that evaluates each application for a business credit line from a fisherman at LPD Desa Adat Jimbaran using the nave bayes method. Using 340 data on prospective credit recipients including loan amount, collateral value, income, expenses, time period, other obligations, and credit history, it is determined that more prospective creditors are eligible than ineligible, with 321 declared eligible and 19 declared ineligible. The average accuracy result was 94.31%, with the first experiment yielding the highest accuracy at 95.30% and the third experiment yielding the lowest accuracy at 94.67%.

1. INTRODUCTION

Lembaga Perkreditan Desa is a village financial institution engaged in the business of saving and borrowing, LPD not only plays its function as a financial institution that serves the financial transactions of the village community but has become a solution to the limited access to funds for rural communities who are in fact a community group with limited economic capacity. LPD Desa Adat Jimbaran is a financial institution owned by a traditional village domiciled in the area of Desa Adat Jimbaran. The profits of LPD Desa Adat Jimbaran in 2020 experienced a drastic decline due to the Covid-19 pandemic sweeping the world. Even so, LPD Desa Adat Jimbaran is still trying to be present to provide the principle of benefits while encouraging the economic revival of the traditional village krama by launching a new product, namely the Fisherman Business Credit (KUN) in June 2021.

With the rapid development of credit applications at the Jimbaran Traditional Village LPD, especially fishing business credit, this requires the LPD to be more thorough and observant in analyzing incoming fisherman business credit application data. This is done to minimize bad loans like in 2020 which resulted in the turnover of funds and profits being stagnant. On the other hand, the community demands transparency in the selection process for assessing credit applications, especially people whose credit applications are rejected. The current credit assessment process is full of subjectivity and requires software that can solve this problem. One system that can help policy makers decide on credit recipients is a decision support system (SPK).[1].

Decision support systems can also expand the capabilities of decision makers, save time needed to solve problems, produce solutions more quickly, are able to provide various alternatives in decision making, strengthen decision maker confidence and provide a competitive advantage for the organization as a whole.[1]–[4]. Therefore, in this study a decision support system has been developed to assist the LPD of Jimbaran Traditional Village in making credit granting decisions using the Naive Bayes method. The Naive Bayes method itself is a method that can be used in terms of decision making to get better results on a classification problem[5]–[7]. The Naive Bayes method is also considered to have good potential in classifying documents compared to other classification methods in terms of computational accuracy and efficiency.[8], [9]. The performance of the Naïve Bayes method in the developed system is measured using the Confusion Matrix method. This method can produce performance ratings in the form of numbers so that it makes it easier to analyze algorithm performance[10].

Several previous studies that implemented the Naïve Bayes method included those conducted by[11]which discusses the implementation of the Naive Bayes method in classifying beneficiaries of the Family Hope Program (PKH) in Minggiran Kediri Village. The test results show that the developed system produces an accuracy of 93.33%.Other research conducted by[12] who built a system for determining the eligibility of granting credit to Adira Finance with test results yielding an accuracy value of 85%. Research by[13] develop a decision support system for determining employee status at PT. Emsonic Indonesia uses the Naive Bayes method with an accuracy rate of 94%.The Naïve Bayes method used by[14]for the acceptance of new employees at PT. Sasmito.While Siregar implements the Naïve Bayes method in a decision support system for determining employee monthly incentives at Edene Sayangku Café & Bakery[15].

Some of the research that has been done shows that the Naïve Bayes method can be implemented as a decision support system with very good performance with an accuracy rate above 85%. Making a decision support system for granting fishing business credit using the Naive Bayes method for the Jimbaran Traditional Village LPD in this study is expected to help speed up the performance of LPD employees, especially those in the credit department, so that they can more quickly calculate the feasibility predictions for granting fishing business credit. With this decision support system, it is hoped that it can provide transparency to the public by providing decision results based on data and the use of scientific methods that are believed to have excellent performance with a high degree of accuracy.

2. METHODS

Research Stages

The research stages are an overview regarding the flow of research carried out in carrying out this research from beginning to end. The first stage in this research is to identify the problems that occur and determine research objectives. After the research objectives have been set, proceed with data collection and conducting literature studies related to the research theme.

Data collection was carried out using several methods including interviews, observation, documentation and library research. Interviews were conducted with the Head of the Jimbaran Traditional Village LPD to explore the process of granting fishing business loans, the problems that occurred and the expected solutions. The documentation obtained is in the form of credit application data, credit recipient data, and records of credit payments that have been made by LPD customers.

The next stage is to apply the data that has been collected into the naïve Bayes method starting from setting the criteria, class probabilities, probabilities of each class to the final probabilities of each class in the form of decisions produced by the naïve Bayes method. The decision-making process using the Naïve Bayes method is then implemented into a system starting from designing the system, coding, to evaluating and testing the system. System testing was carried out using the black box testing method, while the performance of the implemented Naïve Bayes method was tested using the confusion matrix method.

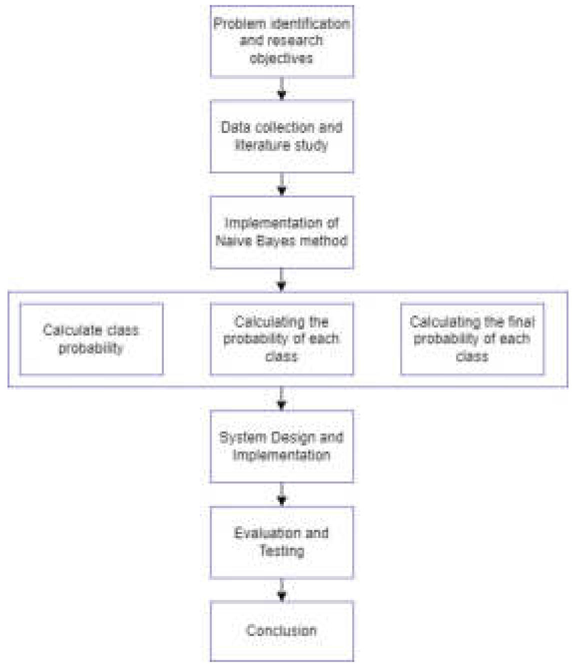

Figure 1. Research Stages

The Naïve Bayes method

This decision support system supports a variety of classification strategies, including the Naive Bayes method. The bayes theorem, also known as the naive bayes method, can foresee future opportunities based on past experience. In the classification procedure, the naive bayes method [16] employs probability and statistical methods. This method has great classification potential in terms of accuracy and data computation. Naive Bayes is extensively utilized in classification techniques and sentiment analysis, particularly with Twitter data, which employs multiple methods including Unigram Naive Bayes, Multinomial Naive Bayes, and Maximum Entropy Classification. The primary characteristic of Naive Bayes classification is obtaining a robust hypothesis for each condition or event. Here is the Naive Bayes theorem's equation [17], [18]:

$$P(H|X) = \frac{P(X|H).\ P(H)}{P(X)} \quad (1)$$

Where:
X : Data with unknown class
H : Hypothesis data is a specific class
P(H|X) : The probability of the hypothesis H under condition X
P(H) : Probability hypothesis H
P(X|H) : The probability of X based on the condition of the hypothesis
P(X) : Probability hypothesisX

3. RESULTS AND DISCUSSION

Data Sets

In this study, customer data was used for LPD Jimbaran Badung Bali Traditional Village for the period June 2021 to July 2022 with a total data of 340 customers. This customer data has the attributes of name, loan amount, collateral value, income, time period, expenses, other debts, credit history, and creditworthiness status. After the transformation, the customer's raw data is written using a range of

http://ejournal.seaninstitute.or.id/index.php/InfoSains
Jurnal Info Sains : Informatikan dan Sains, Volume 13, No 01, 2023
E-ISSN.2797-7889, P-ISSN.2089-3329

amounts according to the real conditions of the data, and then given the status of a customer who is eligible and not eligible to receive credit. The dataset used in this study can be seen in Table 1 below.

Table 1. Dataset of alternative credit recipient candidates

No	Name	Loan Amount	The value of collateral	Income	Time period	Expenditure	Other Debt	Credit History	Status
1	Sudra I Wayan	1,000,000 to 10,000,000	>10,000,000 to 20,000,000	>5,000,000	24 Months	>2,500,000 to 5,000.00	1,000,000 to 10,000,000	Never and smoothly	Worthy
2	Sumerta I Made	>10,000,000 to 20,000,000	>20,000,000	>3,500,000 to 5,000,000	36 Months	>2,500,000 to 5,000,000	>10,000,000 to 20,000,000	Never and less smoothly	Not feasible
3	Sumartayasa I Nyoman	>20,000,000	>20,000,000	>5,000,000	36 Months	>2,500,000 to 5,000,000	0	Never and smoothly	Worthy
4	Sekar I Wayan	>10,000,000 to 20,000,000	>10,000,000 to 20,000,000	>3,500,000 to 5,000,000	12 months	1,000,000 to 2,500,000	1,000,000 to 10,000,000	Never and smoothly	Worthy
5	Soma I Made	>10,000,000 to 20,000,000	>10,000,000 to 20,000,000	>5,000,000	36 Months	1,000,000 to 2,500,000	1,000,000 to 10,000,000	Never	Worthy
6	Run I Wayan	>10,000,000 to 20,000,000	>10,000,000 to 20,000,000	>5,000,000	36 Months	1,000,000 to 2,500,000	1,000,000 to 10,000,000	Never and smoothly	Worthy
7	Sarta I Wayan	>20,000,000	>20,000,000	>5,000,000	36 Months	>5,000,000	>10,000,000 to 20,000,000	Never and less smoothly	Not feasible
8	Undah I Nyoman	1,000,000 to 10,000,000	>10,000,000 to 20,000,000	>3,500,000 to 5,000,000	24 Months	1,000,000 to 2,500,000	0	Never	Worthy
9	Merta I Wayan	>10,000,000 to 20,000,000	>10,000,000 to 20,000,000	>3,500,000 to 5,000,000	36 Months	1,000,000 to 2,500,000	0	Never and smoothly	Worthy
10	One I Wayan	>20,000,000	>20,000,000	>5,000,000	36 Months	1,000,000 to 2,500,000	0	Never	Worthy
…	…	…	…	…	…	…	…	…	…
340	Redu I Nyoman	>10,000,000 to 20,000,000	>10,000,000 to 20,000,000	>3,500,000 to 5,000,000	36 Months	1,000,000 to 2,500,000	0	Never and smoothly	Worthy

Naïve Bayes Calculation Method

Before testing the developed system, at this stage a manual calculation is carried out to obtain the probability value of each class. In applying for fishing business loans to the Jimbaran Traditional

Village LPD, the calculation results yielded 2 classes, namely the "Eligible" and "Not Eligible" classes. The method of calculation is to find the number of feasible and infeasible data from the total dataset, then divide it by the total dataset.

A. Counting the number of classes or labels

P(Decent) = 322/340 = 0.947
P(Not Eligible) = 18/340 = 0.053

B. Counting the amount of data for each class

Loan Amount:
P(1,000,000 to 10,000,000) Eligible = 120/322 = 0.373
P(1,000,000 to 10,000,000) Not Eligible = 1/18 = 0.036
P(10,000,001 to 20,000,000) Eligible = 133/322 = 0.413
P(10,000,001 to 20,000,000) Not Eligible = 9/18 = 0.500
P(>20,000,000) Decent = 69/322 = 0.214
P(>20,000,000) Not Eligible = 8/18 = 0.444

Guarantee:
P(5,000,000 to 10,000,000) Eligible = 54/322 = 1.689
P(5,000,000 to 10,000,000) Not Eligible = 2/18 = 0.111
P(10,000,001 to 20,000,000) Eligible = 201/322 = 0.624
P(10,000,001 to 20,000,000) Not Eligible = 13/18 = 0.722
P(>20,000,000) Decent = 67/322 = 0.208
P(>20,000,000) Not Eligible = 3/18 = 0.167

Income:
P(1,000,000 to 3,500,000) Eligible = 102/322 = 0.317
P(1,000,000 to 3,500,000) Not Eligible = 11/18 = 0.611
P(3,500,001 to 5,000,000) Eligible = 163/322 = 0.506
P(3,500,001 to 5,000,000) Not Eligible = 4/18 = 0.222
P(>5,000,000) Decent = 57/322 = 0.177
P(>5,000,000) Not Eligible = 3/18 = 0.167

Expenditure:
P(1,000,000 to 2,500,000) Eligible = 26/322 = 0.081
P(1,000,000 to 2,500,000) Not Eligible = 8/18 = 0.444
P(2,500,001 to 5,000,000) Eligible = 59/322 = 0.183
P(2,500,001 to 5,000,000) Not Eligible = 7/18 = 0.389
P(>5,000,000) Decent = 1/322 = 0.003
P(>5,000,000) Not Eligible = 3/18 = 0.167

Credit History:
P(Ever & Current) Eligible = 9/322 = 0.028
P(Ever & Current) Not Eligible = 0/18 = 0
P(Ever & Substandard) Eligible = 40/322 = 0.124
P(Ever & Substandard) Not Eligible = 2/18 = 0.111
P(Never Crashed) Decent = 273/322 = 0.848
P(Never Crashed) Not Eligible = 16/18 = 0.889
P(Never) Eligible = 273/322 = 0.848
P(Never) Not Eligible = 16/18 = 0.889

Other Payables:
P(None) Eligible = 260/322 = 0.807
P(None) Not Eligible = 3/18 = 0.167
P(1,000,000 to 10,000,000) Eligible = 54/322 = 0.168
P(1,000,000 to 10,000,000) Not Eligible = 7/18 = 0.389
P(10,000,001 to 20,000,000) Eligible = 7/322 = 0.022
P(10,000,001 to 20,000,000) Not Eligible = 7/18 = 0.389
P(>20,000,000) Decent = 1/322 = 0.003
P(>20,000,000) Not Eligible = 1/18 = 0.056

Time period:
P(12 Months) Eligible = 9/322 = 0.028
P(12 Months) Not Eligible = 0/18 = 0
P(24 Months) Eligible = 40/322 = 0.124
P(24 Months) Not Eligible = 2/18 = 0.111
P(36 Months) Eligible = 273/322 = 0.848
P(36 Months) Not Eligible = 16/18 = 0.889

C. Calculating the Final Probability of Each Class

Furthermore, in calculating the final probability in each class using training data and converting it into a predetermined value with sub criteria. Then from each sub-criteria and class probability values are multiplied, and from the two results that have been determined in each class and compare the highest value. If the feasible class has the highest value, then the result is feasible and vice versa if the class is not feasible has the highest value, then the result is not feasible.

Design and Implementation

A. Database Design

The naïve Bayes method in this study was implemented into a system that has an interface and can be accessed by several users such as the head of the credit admin section, the head of the credit analysis section, the head of the credit department, and the head of the LPD with different access rights. To facilitate access, a structured database design was created as shown in Figure 2 below.

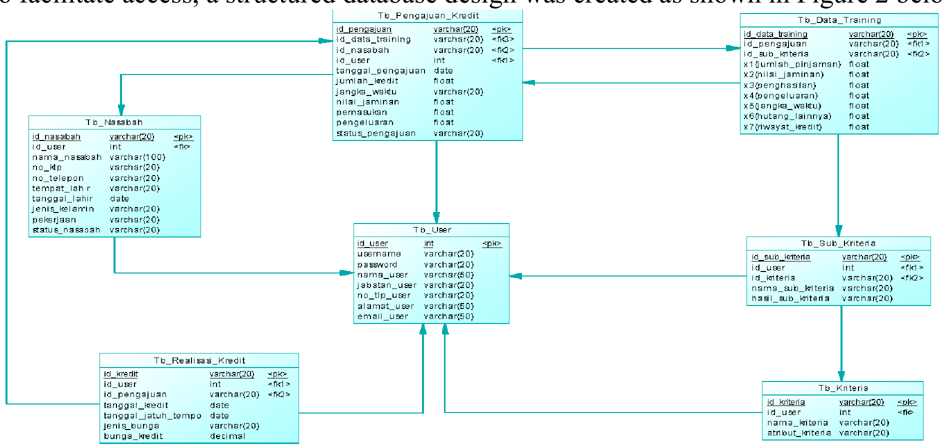

Figure 2. Structure of the Fishermen Business Credit Granting System

As seen in Figure 2, the database structure is implemented with seven tables, namely the user table, customer table, credit application table, criteria table, sub-criteria table, training data table, and credit realization table. The user table is used to store user data that can use the system, the customer table is used to store LPD customer data, and the credit realization table is used to store credit recipient data. The criteria and sub-criteria table is used to store attribute data for the Naive Bayes method, the credit application table is used to store data on prospective credit recipients and this table will relate to the training data table to store learning data carried out by the Naïve Bayes method.

B. System Implementation

The system developed to implement the Naïve Bayes method in the process of receiving LPD fishing business loans at the Jimbaran Traditional Village facilitates users in managing user data, criteria data, sub-criteria data, and customer data. The resulting system can manage credit application transactions, process credit by implementing the Naïve Bayes method, credit realization to produce credit application and realization reports. The results of the eligibility assessment process for prospective credit recipients can be seen in Figure 3 below.

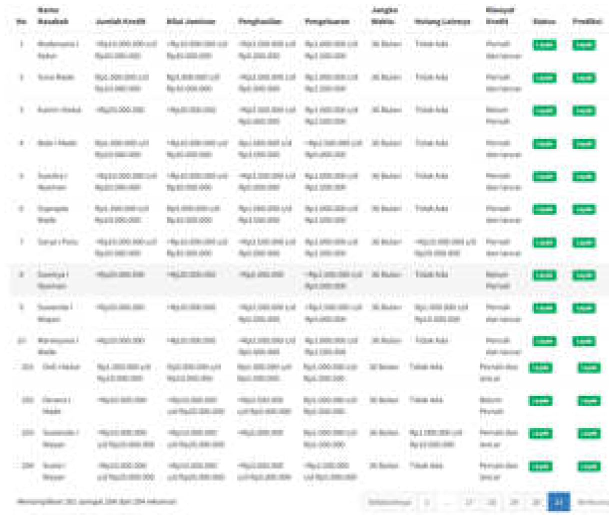

Figure 3. Prospective Credit Recipient Eligibility Results

Figure 3 shows that the credit recipient's eligibility assessment uses seven criteria, namely the amount of credit, collateral value, income, expenses, time period, other debts, and credit history. Using these criteria, each prospective customer is assessed using the Naive Bayes method, resulting in two decisions, namely prospective customers who are eligible and who are not eligible to receive credit. Details of the probability value of each decision can be seen in Figure 4 below.

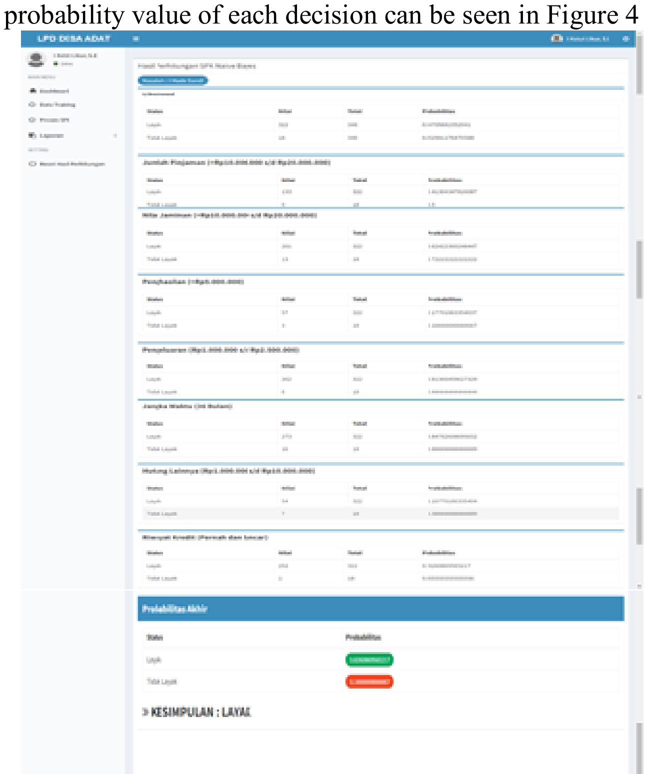

Figure 4. Probability Values

Testing

The final stage of this research is to test the system that has been developed. System functional testing was carried out using the blackbox testing method and the performance of the naïve Bayes method was measured using the confusion matrix. Based on the results of functional testing of the 12 system features tested, all have been running well according to their respective functions. The performance of the naïve Bayes method was measured by three experiments using 204 data, 238 data, 272 and 340 customer data. The following Table 2 is the documentation of the results of the tests that have been carried out.

Table 2. Test Results

Test	accuracy	Precision	recall
204 Data	95.30%	99.00%	95.74%
238 Data	94.82%	99.37%	95.36%
272 Data	94.67%	99.35%	95.22%
340 Data	94.71%	99.38%	95.24%
Average	**94.31%**	**99.28%**	**95.46%**

As shown in Table 2, the test was carried out four times by measuring accuracy, precision and recall. The highest accuracy results were obtained in the first test with an accuracy value of 95.30% while the lowest results were obtained in the third test with a value of 94.67%, and an average accuracy result of 94.31%. Details of accuracy, precision, and recall of the tests carried out can be seen in the figure below.

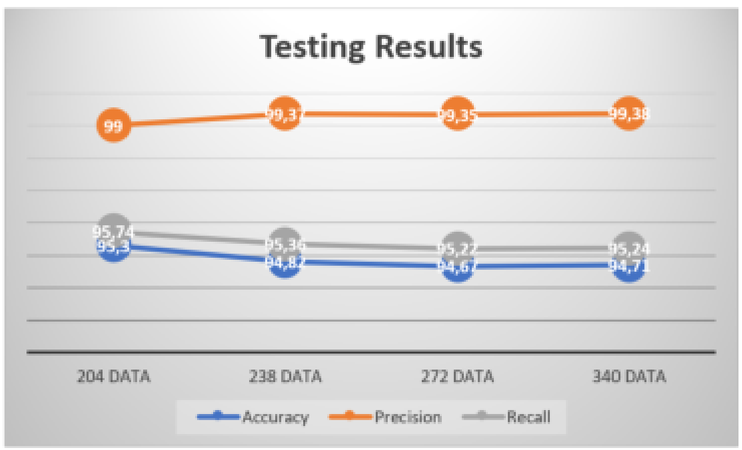

Figure 5. Test Results

4. CONCLUSION

Based on the results of the tests, all of the system features developed in this study functioned properly. Using 340 data on prospective fishermen business credit recipients of LPD Desa Adat Jimbaran, which includes loan amount, collateral value, income, expenses, time period, other debts, and credit history, it was determined that prospective creditors who are deemed feasible outnumber those who are deemed unfit, with 321 deemed feasible and 19 deemed unfit. The first experiment yielded the highest accuracy with a result of 95.30%, while the third experiment yielded the lowest accuracy with a result of 94.67%; the average accuracy result was 94.31%.

REFERENCES

[1] I. G. T. Suryawan and W. G. M. Negara, "Rancang Bangun Sistem Pendukung Keputusan Penerima Beasiswa Dengan Metode Profile Matching," *J. Teknol. Inf. dan Komput.*, vol. 7, no. 3, 2021.

[2] G. S. Mahendra *et al.*, *SISTEM PENDUKUNG KEPUTUSAN (Teori dan Penerapannya dalam berbagai Metode)*. PT. Sonpedia Publishing Indonesia, 2023.

[3] I. G. I. Sudipa *et al.*, *PENERAPAN DECISION SUPPORT SYSTEM (DSS) DALAM BERBAGAI BIDANG (Revolusi Industri 4.0 Menuju Era Society 5.0)*. PT. Sonpedia Publishing Indonesia,

2023.
[4] Y. Boari, R. Megavitry, P. J. Pattiasina, H. T. Ramdani, and H. Munandar, "The Analysis Of Effectiveness Of Mobile Learning Media Usage In Train Students' Critical Thinking Skills," *Mudir J. Manaj. Pendidik.*, vol. 5, no. 1, pp. 172–177, 2023.
[5] D. Alita, I. Sari, A. R. Isnain, and S. Styawati, "Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa," *J. Data Min. Dan Sist. Inf.*, vol. 2, no. 1, pp. 17–23, 2021.
[6] I. G. I. Sudipa, I. M. D. P. Asana, K. J. Atmaja, P. P. Santika, and D. Setiawan, "Analisis Data Kepuasan Pengguna Layanan E-Wallet Gopay Menggunakan Metode Naïve Bayes Classifier Algorithm," *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 3, pp. 726–735, 2023.
[7] B. Kwintiana *et al.*, *DATA SCIENCE FOR BUSINESS: Pengantar & Penerapan Berbagai Sektor*. PT. Sonpedia Publishing Indonesia, 2023.
[8] D. Berrar, "Bayes' theorem and naive Bayes classifier," *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.*, vol. 403, p. 412, 2018.
[9] A. Firmansyah, F. Ramadhani, and E. Fauzan, "Sistem Pendukung Keputusan Penentuan Status Karyawan Menggunakan Metode Naïve Bayes," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 2, no. 02, pp. 85–94, 2020.
[10] M. R. Fanani, "Algoritma Naive Bayes Berbasis Forward Selection Untuk Prediksi Bimbingan Konseling Siswa," *J. Disprotek*, vol. 11, no. 1, pp. 13–22, 2020.
[11] F. K. Pratama, D. W. Widodo, and N. Shofia, "Implementasi Metode Naïve Bayes dalam Mengklasifikasi Penerima Program Keluarga Harapan (PKH) Desa Minggiran Kediri," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 2021, vol. 5, no. 3, pp. 23–28.
[12] D. Dewi and F. Satria, "Algoritma Naive Bayes Untuk Menentukan Kelayakan Pemberian Kredit Pada Adira," *PROCIDING KMSI*, vol. 6, no. 1, pp. 8–13, 2018.
[13] A. Firmansyah and F. Ramdhani, "Sistem Pendukung Keputusan Penentuan Status Karyawan Menggunakan Metode Naive Bayes (Studi Kasus: PT. Emsonic Indonesia)," *J. SIGMA*, vol. 11, no. 1, pp. 1–8, 2020.
[14] T. D. Pengestuti, F. T. Anggraeny, and E. P. Mandyartha, "Rancang Bangun Sistem Pendukung Keputusan Penerimaan Karyawan Baru Menggunakan Naive Bayes Classifer (Studi Kasus PT. Sasmito)," *J. Inform. dan Sist. Inf*, vol. 1, no. 3, pp. 1072–1080, 2020.
[15] V. M. M. Siregar, "Sistem Pendukung Keputusan Penentuan Insentif Bulanan Pegawai Dengan Menggunakan Metode Naïve Bayes," *Sist. J. Sist. Inf.*, vol. 7, no. 2, pp. 78–86, 2018.
[16] R. Mehra, M. K. Bedi, G. Singh, R. Arora, T. Bala, and S. Saxena, "Sentimental analysis using fuzzy and naive bayes," in *Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017*, 2018, vol. 2018-Janua, no. Iccmc, pp. 945–950. doi: 10.1109/ICCMC.2017.8282607.
[17] G. Triyono and D. Ginting, "Comparative Analysis Performance of Naïve Bayes and K-NN Using Confusion Matrix and AUC To Predict Insurance Fraud," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 4, pp. 2293–2300, 2022.
[18] A. B. Downey, *Think Bayes*. " O'Reilly Media, Inc.," 2021.