JITEEHA: Journal of Information Technology Application in Education, Economy, Health and Agriculture
Vol. 02, No. 01, February 2025, pp. 7~15, e-ISSN: 3090-3939
https://journal.iteeacademy.org/index.php/jiteeha/

Application of the Naive Bayes Data Mining Algorithm to Predict Used Motorcycle Purchase Decisions

Andreas Masdiyanto[1], Robert Baz Kiyosaki[2], Sudrajad Hakiki[3], Farrel Muhammad Raihan Akhdan[4], Tshering Peldon[5]
[1,2,3] Department of Informatics Engineering, University of Widyagama Malang, Borobudur Street No. 35, Malang City, East Java, Indonesia
[4] Department of Informatics Magister, Universitas Islam Indonesia, Yogyakarta, Indonesia
[5] Gaupel Lower Secondary School- Bhutan

Article Info

Article history:

Received January 01, 2025
Revised February 10, 2025
Accepted February 30, 2025

Keywords:

Naive Bayes
Purchasing Decision Prediction
Data Mining
Used Motorcycles
Gaussian Naive Bayes
Data Analysis

ABSTRACT

This study applies the Naive Bayes algorithm to predict the decision to purchase used motorcycles based on attributes such as model, year of manufacture, price, engine capacity, and transaction results. Utilizing the Gaussian Naive Bayes approach for continuous data, this research aims to develop a reliable predictive model and understand the most significant attributes influencing purchasing decisions. The test results show that the predictive model achieves an accuracy rate of 75%, indicating the effectiveness of the Naive Bayes algorithm in handling data classification. This study provides insights that can help industry players enhance their sales strategies based on accurate data analysis.

Corresponding Author:

Andreas Masdiyanto
Department of Informatics Engineering
Universitas Widya Gama Malang
Jl. Borobudur No. 35 Malang – Jawa Timur, Indonesia
Email: amasdiyanto13@gmail.com

1. INTRODUCTION

In today's digital era, the ease of making purchasing decisions has significantly increased due to the abundance of data availability[1]. This situation encourages many business players to shift their sales activities to online marketplaces, which are considered more efficient platforms for reaching a broader consumer base[2]. The used motorcycle industry, one of the significant sectors of commerce, has also experienced the impact of data technology. Consumers tend to consider various factors, such as price, model, year of manufacture, transmission type, and engine capacity (CC), before deciding to purchase a used motorcycle. However, the complexity of data and the variation in consumer preferences make it challenging for sellers and online platforms to effectively understand purchasing patterns. This is where data mining methods become relevant. Data mining is a field of study aimed at addressing challenges in extracting information from large-scale databases. It integrates various techniques such as statistics, machine learning, data visualization, pattern recognition, and database management[3]. One of the commonly used algorithms in data mining and the focus of this research is Naive Bayes. The Naive Bayes algorithm is a simple probabilistic-based classification method often used for document classification. It calculates probabilities by analyzing the frequency and combinations of values in a dataset, enabling classification based on the calculated probability values for each model[4], [5], [6]. Naive Bayes will be used as a straightforward yet effective probabilistic approach to predicting purchasing decisions based on the available attributes.

This research refers to several previous studies as references. Among them are: (1) Previous research using the Naive Bayes algorithm for data mining. One example is the study conducted by Kokom Komariah et al. in the manufacturing sector, where the Naive Bayes algorithm was used to create a stock prediction system that proved more accurate than manual methods, helping companies make decisions related to inventory management[7]. (2) Research in the education sector, such as the study by Detrinal Putra and Arief Wibowo, which used the Naive Bayes algorithm to predict students' interest in specific majors from 2016 to 2018 at SMA Yadika 5[8]. (3) Research in the SME sector by Erlin Elisa et al., which explored consumer purchasing patterns of Martista Ikhsan SME products in Batam City using the Naive Bayes algorithm. This helped business owners understand market trends, increase sales, expand market share, and innovate their products[9]. (4) Research in the social sector by Desi Laila Sari et al., which aimed to predict divorce potential in Central Aceh Regency. The researchers used data mining techniques to process divorce data and identify the most dominant factors affecting divorce, such as the age of the couple, length of marriage, number of children, and primary reasons for divorce. This research aimed to develop a web-based system that predicts divorce risk using historical data as a reference[10]. These diverse cases demonstrate the importance of data mining for organizations or companies in making decisions by extracting relevant information from existing data warehouses to draw meaningful conclusions[11].

The dataset used in this research includes information about motorcycle models, year of manufacture, price, transmission type, motorcycle type (moped, scooter), engine capacity, and transaction outcomes (purchase or no purchase). This dataset offers significant potential for analysis to identify key factors influencing purchasing decisions. The objective of this research is to develop a predictive model for purchasing decisions of used motorcycles using the Naive Bayes algorithm. The concept of utilizing, managing, and processing this dataset is similar to previous studies, which demonstrated that attributes such as price and customer interest provide valuable insights into purchasing patterns. For instance, Fitriana et al. utilized similar attributes to predict paint purchases with high accuracy, highlighting the importance of quality data for decision-making analysis[12]. Additionally, research by Arjun & Kiki used specific attributes such as brand and sales volume to determine optimal restocking patterns, indicating that multivariate attribute datasets can yield reliable and relevant predictions in a business context[13]. Thus, the attributes in this research dataset have the potential to provide significant predictions in understanding used motorcycle purchasing decisions. Furthermore, this study aims to identify the most significant attributes influencing purchasing decisions and to provide data-driven insights that can help industry players enhance the efficiency and effectiveness of their sales strategies.

1. RESEARCH METHODOLOGY

2. 1 Data Collcetion

The data used in this study was obtained from the Used Motorcycle Purchase dataset. The dataset contains the following attributes:

- Model (C1)
- Year of Production (C2)
- Price (C3)
- Engine Cubic Capacity (C4)
- Transaction (C5)

The collected dataset will be processed and analyzed using Data Mining methods with the Naive Bayes algorithm.

2. 2 Implementation of the Naive Bayes Algorithm and Gaussian Density Function

The collected and preprocessed data will be mined using the Naive Bayes algorithm in several steps, including calculating the probability of each attribute, determining the testing data, classifying the testing data, and obtaining the results of the testing data after classification. In processing the data, some attributes will have continuous values. Therefore, the Gaussian density function will be used for probability calculations, specifically for the Price (C3) attribute. The Gaussian density function, also known as the normal distribution function in data mining, is a probability function used to describe data distributions with a bell-shaped curve. This function is commonly used to calculate the probability or density at a specific point in continuous data distribution. In the Naive Bayes algorithm for continuous data, the Gaussian density function is used to calculate the probability of a feature based on the normal distribution of the data. This is referred to as Gaussian Naive Bayes.

2. 3 Model Evaluation

Accuracy calculations for the model or testing data are essential for metric evaluation. The results of this evaluation will demonstrate the effectiveness of the Naive Bayes algorithm in predicting purchasing decisions for used motorcycles.

2. RESULTS AND DISCUSSION

3. 1 Dataset Collection

The data used, which will also serve as the training data in this study, is obtained from a preprocessed dataset of Used Motorcycle Purchases. Below is the dataset:

Table 1. Used Motorbike Transactions (Training Data)

Motorbike	Attribute				Class
	Model (C1)	Year of Production (C2)	Price (C3)	Engine Cubic Capacity (C4)	Transaction (C5)
1	Supra X	2017	Rp8.000.000	125	Buy
2	Vario	2017	Rp14.200.000	125	Not Buy
3	Vario	2016	Rp11.400.000	125	Buy
4	Revo	2016	Rp7.000.000	110	Buy
5	Revo	2015	Rp7.000.000	110	Not Buy
6	Supra X	2015	Rp8.000.000	125	Buy
7	Vario	2016	Rp14.200.000	125	Not Buy
8	Genio	2019	Rp10.400.000	110	Buy
9	Vario	2016	Rp11.400.000	125	Not Buy
10	Vario	2017	Rp11.400.000	125	Buy
11	Genio	2019	Rp14.500.000	110	Not Buy

3. 2 Calculating the Ratio or Probability of Each Attribute and Class

Before calculating the likelihood, it is necessary to compute the ratio or probability of each attribute and class to use it in the Naive Bayes Classification calculations.

Number of classes = 11

- **C1**
- Supra X: Buy = 2 / 6 = 0,33; Not buy = 0 / 5 = 0
- Vario: Buy = 2 / 6 = 0,33; Not Buy = 3 / 5 = 0,6
- Revo: Buy = 1 / 6 = 0,16; Not Buy = 1 / 5 = 0,16
- Genio: Buy = 1 / 6 = 0,16; Not Buy = 1 / 5 = 0,16

Table 2. Probability of C1

Model	Probability	
	Buy	Not Buy
Supra X	2 / 6	0 / 5
Vario	2 / 6	3 / 5
Revo	1 / 6	1 / 5
Genio	1 / 6	1 / 5

- **C2**
- 2015: Buy = 1 / 6 = 0,33; Not Buy = 1 / 5 = 0,2
- 2016: Buy = 2 / 6 = 0,33; Not Buy = 2 / 5 = 0,4
- 2017: Buy = 1 / 6 = 0,16; Not Buy = 1 / 5 = 0,2
- 2019: Buy = 1 / 6 = 0,16; Not Buy = 1 / 5 = 0,2

Table 3. Probability of C2

Year of Production	Probability	
	Buy	**Not Buy**
2015	1 / 6	1 / 5
2016	2 / 6	2 / 5
2017	2 / 6	1 / 5
2019	1 / 6	1 / 5

- **C3**

For attribute C3, the data is continuous. Therefore, the probability of the data will be calculated using the Gaussian Density Function or Gaussian Naive Bayes. The rows of data in the training dataset, specifically for attribute C3, will be sorted based on the values of the class attribute (Buy or Not Buy). Then, the data will be used to calculate the mean (μ) and standard deviation (σ). Below is the Probability Table for C3.

Table 4. Probability of C3

n	Buy	Not Buy
1	Rp8.000.000	Rp14.200.000
2	Rp11.400.000	Rp7.000.000
3	Rp7.000.000	Rp14.200.000
4	Rp8.000.000	Rp11.400.000
5	Rp10.400.000	Rp14.500.000
6	Rp11.400.000	-
Average	Rp9.366.667	Rp12.260.000
Deviation Standard	Rp1.932.529	Rp3.199.687

- **C4**

- 110: Buy = 2 / 6 = 0,33; Not Buy = 1 / 5 = 0,4
- 125: Buy = 4 / 6 = 0,6; Not Buy = 3 / 5 = 0,6

Table 5. Probability of C4

Engine Cubic Capacity	Probability	
	Buy	**Not Buy**
110	2 / 6	2 / 5
125	4 / 6	3 / 5

- **C5**

Beli = 6 / 11 = 0,54; Tidak Beli = 5 / 11 = 0,45

3. 3 Testing Data

To make predictions on a target model or relevant data, new data is required to evaluate how well the model can make predictions on unknown data. The researcher will use 8 rows of data as the Test Data. Below is the Test Data:

Table 6. Data *Test*

Motorbike	**Attribute**				**Class**
	Model (C1)	**Year of Production (C2)**	**Price (C3)**	***Engine Cubic Capacity* (C4)**	**Transaction (C5)**
1	Supra X	2017	Rp8.550.000	125	?
2	Vario	2017	Rp12.000.000	125	?
3	Genio	2019	Rp11.400.000	110	?
4	Revo	2017	Rp9.000.000	110	?
5	Supra X	2015	Rp7.750.000	125	?
6	Revo	2015	Rp8.000.000	110	?
7	Vario	2016	Rp13.500.000	125	?
8	Genio	2019	Rp10.400.000	110	?

The prediction target that will be predicted is in the class of a data model. In this case, what will be predicted is Transaction (C5).

3. 4 Gaussian Density Classification (C3)

In the Data Model for this case, it can be observed that there is continuous data in the Price attribute (C3). This data will first be classified independently using the Gaussian Density function before being included in the Naive Bayes Algorithm testing calculations. The Gaussian Density function is defined by the following formula:

$$f(X) = \frac{1}{\sqrt{2 * \pi * \sigma}} * e^{-\frac{(X-\mu)^2}{2 * \sigma^2}}$$

Explanation:
- f(x): The probability density function of the normal distribution for the value x.
- π: Pi (≈ 3.14159).
- e: Exponential constant (≈ 2.7183).
- x: The attribute value for which the probability is to be calculated.
- μ: The mean of the attribute for a specific class, in this case, the C3 attribute.
- σ: The standard deviation of the attribute for a specific class, in this case, the C3 attribute.

Each C3 attribute in each row of the Test Data will be calculated and will refer to all variable names or values in the class from the Test Data.

1. X = 8550000

$$f(8550000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(8550000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000019$$

$$f(8550000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(8550000 - 12260000)^2}{2 * 3199687^2}} = 0{,}00000006$$

2. X = 12000000

$$f(12000000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(12000000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000008$$

$$f(12000000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(12000000 - 12260000)^2}{2 * 3199687^2}} = 0{,}000000124$$

3. X = 11400000

$$f(11400000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(11400000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000012$$

$$f(11400000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(11400000 - 12260000)^2}{2 * 3199687^2}} = 0{,}000000120$$

4. X = 9000000

$$f(9000000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(9000000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000020$$

$$f(9000000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(9000000 - 12260000)^2}{2 * 3199687^2}} = 0{,}00000007$$

5. X = 7750000

$$f(7750000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(7750000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000015$$

$$f(7750000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(7750000 - 12260000)^2}{2 * 3199687^2}} = 0{,}000000046$$

6. X = 8000000

$$f(8000000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(8000000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000016$$

$$f(8000000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(8000000 - 12260000)^2}{2 * 3199687^2}} = 0{,}000000051$$

7. X = 13500000

$$f(13500000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(13500000 - 9366667)^2}{2 * 1932529^2}} = 0{,}00000002$$

$$f(13500000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(13500000 - 12260000)^2}{2 * 3199687^2}} = 0{,}00000012$$

8. X = 10400000

$$f(10400000 \mid Buy) = \frac{1}{\sqrt{2 * \pi * 1932529}} * e^{-\frac{(10400000 - 9366667)^2}{2 * 1932529^2}} = 0{,}0000002$$

$$f(10400000 \mid Not\ Buy) = \frac{1}{\sqrt{2 * \pi * 3199687}} * e^{-\frac{(10400000 - 12260000)^2}{2 * 3199687^2}} = 0{,}00000011$$

Following are the calculation results in table form.

Table 7. Gaussian Density Classification (C3)

Motorbike	Buy	Not Buy
1	0,00000019	0,00000006
2	0,00000008	0,000000124
3	0,00000012	0,000000120
4	0,00000020	0,00000007
5	0,00000015	0,000000046
6	0,00000016	0,000000051
7	0,00000002	0,00000012
8	0,0000002	0,00000011

3. 5 Testing the Naive Bayes algorithm

After calculating the classification for the C3 attribute using the Gaussian Density function, all the attributes whose probabilities have been calculated will be computed again to determine the Naive Bayes likelihood before calculating the probability classification result. Here is the formula for testing the Naive Bayes Algorithm:

$$P(X) = \frac{P(X|C) * P(C)}{P(X)}$$

Explanation:
- P(C|X): The probability of the class (buy/not buy) given the attribute X.
- P(X|C): The probability of attribute X appearing in a specific class.
- P(C): The prior probability of each class.
- P(X): The total probability of attribute X.

The following are the results of testing the Naive Bayes Algorithm:

Table 8. *Naive Bayes Likelyhood*

Buy	Not Buy
0,0000000076	0
0,0000000033	0,0000000041
0,0000000006	0,00000000087
0,000000002	0,000000001
0,0000000029	0
0,00000000081	0,00000000037
0,00000000085	0,0000000076
0,0000000009	0,00000000077

3. 6 Classification of Probability Values and Normalization of Prediction Decisions

Once the Naive Bayes likelihood calculation is complete, the next step is to calculate the Classification Probability Value for each row of the Test Data that has been evaluated in the Naive Bayes likelihood.

1. Probabilitas Beli = 0,0000000076 / (0,0000000076 + 0) = 1
   Probabilitas Tidak Beli = 0 / (0,0000000076 + 0) = 0
2. Probabilitas Beli = 0,0000000033 / (0,0000000033 + 0,0000000041) = 0,45
   Probabilitas Tidak Beli = 0,0000000041 / (0,0000000033 + 0,0000000041) = 0,55
3. Probabilitas Beli = 0,0000000006 / (0,0000000006 + 0,00000000087) = 0,41
   Probabilitas Tidak Beli = 0,00000000087 / (0,0000000006 + 0,00000000087) = 0,59
4. Probabilitas Beli = 0,000000002 / (0,000000002 + 0,000000001) = 0,79
   Probabilitas Tidak Beli = 0,000000001 / (0,000000002 + 0,000000001) = 0,21

And so on, until the end of the rows of Test Data that have been evaluated in the Naive Bayes likelihood. Below are the results of the Probability Value Classification calculations presented in a table.

Table 9. Classification of Probability Values

Buy	Not Buy
1	0
0,45	0,55
0,41	0,59
0,79	0,21
1	0
0,68	0,32
0,10	0,90

0,54	0,46

After that, the prediction decision result for Buy or Not Buy on used motorcycle purchases can be determined by normalizing the data based on the Probability Value Classification results with the following condition: If Buy < Not Buy, then the normalization result is Not Buy; otherwise, the normalization result will be Buy.
Thus, the Test Data provided earlier will yield the following Test Data results.

Table 10. Testing Data Results

Motorbike	**Attribute**				**Class**
	Model (C1)	**Year of Production (C2)**	**Proce (C3)**	***Engine Cubic Capacity* (C4)**	**Transaction (C5)**
1	Supra X	2017	Rp8.550.000	125	Beli
2	Vario	2017	Rp12.000.000	125	Tidak Beli
3	Genio	2019	Rp11.400.000	110	Tidak Beli
4	Revo	2017	Rp9.000.000	110	Beli
5	Supra X	2015	Rp7.750.000	125	Beli
6	Revo	2015	Rp8.000.000	110	Beli
7	Vario	2016	Rp13.500.000	125	Tidak Beli
8	Genio	2019	Rp10.400.000	110	Beli

In this section, the test class data that has been previously tested and obtained results will be compared with the Training Data class to calculate the accuracy of the prediction. After comparing and determining the number of correct and incorrect predictions, the prediction accuracy will be calculated using the following formula:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Data\ in\ the\ Data\ Test}$$

The following are the results of the accuracy calculation.

Table 11. Accuracy Data

Motorbike	**Training Data Class Value**	**Data Test Class Value**	**Correct Prediction**	**Incorrect Prediction**
1	Buy	Buy	1	0
2	Not buy	Not Buy	1	0
3	Buy	Not Buy	0	1
4	Buy	Buy	1	0
5	Not Buy	Buy	0	1
6	Buy	Buy	1	0
7	Not Buy	Not Buy	1	0
8	Buy	Buy	1	0
		Total	6	2
		Accuracy or Inaccuracy (%)	75	25

Based on the table above, it is known that the percentage for Correctly Classified Instances is 75%. Meanwhile, the percentage for Incorrectly Classified Instances is 25%. Out of the 8 used motorcycle purchase data, 6 instances were correctly classified, and 12 instances of used motorcycle purchases were incorrectly classified.

3. CONCLUSION

This research successfully applied the Naive Bayes algorithm to predict used motorcycle purchase decisions based on a dataset that includes attributes such as model, production year, price, engine capacity, and transaction. The algorithm uses a probabilistic approach to calculate the likelihood of a transaction falling into the "Buy" or "Don't Buy" category. By integrating the Gaussian Naive Bayes method for continuous attributes such as price, this research was able to improve accuracy in processing complex data.

The test results show that the prediction model achieved an accuracy rate of 75%, with most of the class predictions being correctly classified. This accuracy rate reflects the effectiveness of the Naive Bayes algorithm in handling data classification problems. Furthermore, further analysis indicates that attributes such as price, model, and engine capacity significantly influence the purchase decision.

ACKNOWLEGDEMENTS

In the planning, research process, and finalization of this study and the preparation of this journal, our research team faced various challenges. Therefore, we extend our gratitude to God Almighty for His guidance and blessings throughout this research.

We also express our deepest appreciation and gratitude to Mrs. Fitri Marisa, S.Kom., M.Pd., Ph.D., for her guidance, support, and dedication during the research and preparation of this journal. Her commitment and insights have been a valuable inspiration to our team.

Additionally, we would like to thank our fellow researchers and all related parties who have made significant contributions to data collection, analysis, and the preparation of this journal. The collaboration and support provided have greatly assisted us in achieving satisfactory research results.

REFERENCES

[1] S. Rahmasari, "Strategi Adaptasi Bisnis di Era Digital: Menavigasi Perubahan dan Meningkatkan Keberhasilan Organisasi," 2023.

[2] N. Purwati, "ANALISA PENGARUH IKLAN TANPA LABEL HARGA PADA MEDIA SOSIAL MENGGUNAKAN ALGORITMA NAIVE BAYES," *Jurnal Informatika, Manajemen dan Teknologi*, vol. 24 Nomor 1, pp. 45–50, Jun. 2022, doi: 10.23969/infomatek.v24i1.4622.

[3] R. Zein and Suparni, "IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI MINAT BELI MAKANAN SECARA ONLINE PADA GENERASI MILENIAL," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, pp. 12727–12734, Dec. 2024.

[4] E. Manalu, F. A. Sianturi, and M. R. Manalu, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTRIES," *Manajemen Dan Informatika Pelita Nusantara*, vol. 2, pp. 16–21, 2017.

[5] A. Bsi Bogor, "Pemilihan Fitur Pada Analisis Sentimen Review Travel Online Menggunakan Algoritma Naïve Bayes Dalam Penerapan Mutual Information Dan Particle Swarm Optimization (PSO) Lisda Widiastuti," *IJCIT (Indonesian Journal on Computer and Information Technology*, vol. 3, no. 1, 2018, [Online]. Available: www.sitejabber.com,

[6] I. M. Praman, I. W. Sudiarsa, and P. Nugraha, "Penerapan Algoritma Naive Bayes Untuk Prediksi Penjualan Produk Terlaris Pada CV Akusara Jaya Abadi," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, No. 4, pp. 518–533, Dec. 2023.

[7] Kokom Komariyah, Rahaditya Dasuki, Dias Bayu Saputra, Saeful Anwar, and Gifthera Dwilestari, "Klasifikasi Stok Barang Menggunakan Algoritma Naïve Bayes Pada Pt.Dharma Electrindo Manufacturing," *KOPERTIP : Jurnal Ilmiah Manajemen Informatika dan Komputer*, vol. 4, no. 2, pp. 35–41, Jun. 2020, doi: 10.32485/kopertip.v4i2.117.

[8] D. Putra and A. Wibowo, "Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes," *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, vol. 2, pp. 84–92, 2020.

[9] E. Elisa, Y. Mardiansyah, and R. Fauzi, "POLA PEMBELIAN KONSUMEN TERHADAP PRODUCT UMKM MARTISTA IKHSAN DENGAN ALGORITMA NAIVE BAYES," *Jurnal Manajemen Informatika & Sistem Informasi*, vol. 5, Nomor 2, pp. 159–167, Jun. 2022, doi: 10.36595/misi.v5i2.

[10] D. Laila Sari, M. Saputra, and H. Gemasih, "PENERAPAN DATA MINING DALAM PROSES PREDIKSI PERCERAIAN MENGGUNAKAN ALGORITMA NAIVE BAYES DI KABUPATEN ACEH TENGAH," *Jurnal Teknik Elektro dan Informatika*, vol. 4 No 1, pp. 23–35, Jan. 2022, doi: 10.14716/ijtech.v0i0.0000.

[11] Susanto, E. Dewi Sri Mulyani, and I. Ratnasari Nurhasanah, "Penerapan Data Mining Classification Untuk Prediksi Perilaku Pola Pembelian Terhadap Waktu Transaksi Menggunakan Metode Naïve Bayes," *Konferensi Nasional Sistem & Informatika 2015*, pp. 313–318, Oct. 2015.

[12] A. Fricco and K. Setiawan, "Implementasi Algoritma Naive Bayes Menggunakan Particle Swarm Optimization Untuk Prediksi Penjualan Produk Air Mineral," *Jurnal Sains dan Teknologi*, vol. 5, no. 1, pp. 193–197, 2023, doi: 10.55338/saintek.v5i1.1430.

[13] F. Harahap, N. E. Saragih, E. T. Siregar, and H. Sariangsah, "Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat," *Jurnal Ilmiah Informatika*, vol. 9 No. 01, pp. 19–23, May 2021.