



Analysis of E-Commerce Consumer Purchasing Patterns Using the Naïve Bayes Algorithm

Khairunnisa*, Muhammad Iqbal

Master of Information Technology, Faculty of Postgraduate Studies, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: ^{1,*}nisak030720@gmail.com, ²muhammadiqbal@dosen.pancabudi.ac.id

(* : nisak030720@gmail.com)

Submitted: 22/01/2026; Accepted: 20/02/2026; Published: 30/03/2026

Abstract-The rapid growth of e-commerce has generated a large volume of transactional data; however, its utilization is often limited to sales reports. In fact, transaction data holds significant strategic potential to be analyzed in order to understand customer purchasing patterns. This study aims to analyze consumer purchasing behavior in e-commerce by applying the Naïve Bayes algorithm as a probabilistic classification method. The dataset used in this research consists of 30 transaction records, including attributes such as purchase frequency, transaction value, product type, payment method, and promotion, with the purchasing decision (buy or not buy) as the class attribute. The research stages include calculating prior probabilities and conditional probabilities for each attribute with respect to the purchasing decision classes. The results indicate that the majority of transactions belong to the buy class with a probability of 0.63, while the not-buy class has a probability of 0.37. Purchase frequency and transaction value are identified as the most influential factors, where medium to high purchase frequency and medium to large transaction values show a strong tendency toward buying decisions. Furthermore, electronic products and non-cash payment methods, particularly e-wallets and credit cards, exhibit high probabilities associated with purchase decisions. Promotions also contribute positively to encouraging purchases, although they are not the sole determining factor. Overall, this study demonstrates that the Naïve Bayes algorithm is effective in identifying customer purchasing patterns and can support strategic decision-making in e-commerce businesses.

Keywords: E-commerce; Consumer Purchasing Patterns; Naïve Bayes; Transaction Data; Classification.

1. INTRODUCTION

The development of information and communication technology has brought significant changes to commercial activities, particularly through the emergence of e-commerce platforms. E-commerce enables buying and selling processes to be carried out online with a wide market reach and easy access for consumers. Along with the increasing number of transactions, e-commerce platforms generate large volumes of data that reflect consumer behavior and preferences. This transaction data includes important information such as purchase frequency, transaction value, preferred product types, payment methods used, and the influence of promotions on purchasing decisions. However, in practice, this data is still often utilized merely as operational reports and sales archives, and thus has not been fully optimized to generate strategic information that is valuable for business decision-making[1].

Understanding consumer purchasing patterns is a key factor for the success of e-commerce businesses. By comprehending how consumers behave when making purchases, companies can design more effective marketing strategies, increase conversion rates, and maintain customer loyalty[2],[4]. Consumer purchasing patterns in e-commerce are not formed simply; rather, they are influenced by various interrelated factors. These factors include consumers' shopping habits, purchasing power, product characteristics, ease of payment methods, and the presence of promotional programs. The complexity of interactions among these factors makes conventional data analysis approaches insufficient for capturing consumer behavior comprehensively[5].

In this context, the use of data mining and machine learning techniques becomes a relevant solution for processing and analyzing e-commerce transaction data. One algorithm widely used in consumer behavior analysis is the Naïve Bayes algorithm. Naïve Bayes is a probabilistic-based classification method that calculates the likelihood of a class based on the combination of attributes possessed by the data[6],[8]. This algorithm operates based on the assumption that each attribute is independent of the others. In other words, the presence or value of one attribute does not influence or depend on another attribute. This simplification allows the model to perform probability calculations efficiently while still producing reliable classification results in many cases[9]. Although this assumption is simple, Naïve Bayes is known for its high accuracy and computational efficiency, especially on categorical data such as e-commerce transaction data[10].

Several previous studies serve as references for this research, including the study conducted by Andi Diah Kuswanto et al. in 2024 on shopping trend analysis using the Naïve Bayes algorithm. The study utilized consumer transaction data and applied a quantitative research approach. The dataset was divided into 80% training data and 20% testing data to evaluate the model performance. The results showed that the Naïve Bayes algorithm performed effectively in identifying and analyzing consumer behavior patterns. In addition, the study demonstrated that Naïve Bayes provides efficient computation and reliable classification accuracy. These findings support the use of the Naïve Bayes method for analyzing purchasing behavior in e-commerce contexts[11]. The study conducted by Pingki Ans Saputra et al. in 2025 examined the dominant factors influencing Generation Z's purchase intentions on Shopee using the Naïve Bayes algorithm. This research utilized survey data collected from 100 respondents. Several variables were analyzed, including price, product reviews, delivery speed, payment methods, and social influence. The dataset was processed to identify patterns and determine the most influential attributes affecting purchase decisions. The results revealed that price and delivery speed were the most dominant factors influencing Generation Z consumers. Furthermore, the model achieved an



accuracy level of 73.75%, indicating that the Naïve Bayes algorithm performed adequately in predicting purchase intentions[12].

The study conducted by Bintang Zulfikar Ramadhan et al. in 2022 analyzed sentiment in e-commerce application reviews using the Naïve Bayes algorithm. This research collected 500 user reviews for each e-commerce application and applied the Knowledge Discovery in Databases (KDD) method as the analytical framework. Several data split scenarios were tested, including 80:20, 70:30, and 60:40, to evaluate model performance. The experimental results showed that the best performance was achieved using the 80:20 data split scenario, particularly for the Shopee application. In this configuration, the Naïve Bayes algorithm obtained an accuracy of 92%, indicating its effectiveness in classifying user sentiment and analyzing opinions in e-commerce reviews[13]. The study conducted by Anastasya Nadia Puspitasari, Yulian Findawati, and Yunianita Rahmawati in 2024 examined sentiment in e-commerce user tweets using the Naïve Bayes classification method. The research focused on analyzing opinions expressed on social media and categorized sentiments into positive and negative classes. Data preprocessing steps were applied to clean and prepare tweet data before classification. The performance of the model was evaluated using accuracy, precision, and recall metrics. The experimental results showed that the Naïve Bayes method achieved strong performance, with an average accuracy of 92.00%, precision of 90.35%, and recall of 100%, indicating its effectiveness in sentiment classification tasks[14].

Based on several previous studies, the Naïve Bayes algorithm has proven to be effective in analyzing e-commerce data, including purchasing patterns, consumer purchase intentions, and user sentiment analysis. The results of these studies demonstrate a relatively high level of accuracy and the ability to identify key factors influencing consumer behavior. Therefore, these studies provide a strong foundation for adopting the research title Analysis of Consumer Purchasing Patterns in E-Commerce Using the Naïve Bayes Algorithm, as this method is relevant, reliable, and suitable for understanding consumer purchasing patterns in a more structured manner.

Nevertheless, there remains a need to further examine the application of the Naïve Bayes algorithm in analyzing e-commerce consumer purchasing patterns by considering various relevant transaction attributes. Each e-commerce platform has distinct consumer characteristics and transaction patterns, necessitating analysis tailored to the context of the data used. This study focuses on analyzing e-commerce consumer purchasing patterns by utilizing purchase frequency, transaction value, product type, payment method, and promotions as predictor variables, with the purchase decision as the class variable.

Based on this background, the aim of this study is to analyze e-commerce consumer purchasing patterns using the Naïve Bayes algorithm and to identify the factors that most significantly influence purchase decisions. The results of this research are expected to provide valuable insights for e-commerce managers in developing more targeted marketing strategies, optimizing payment methods, and determining which product types should be prioritized. Furthermore, this study is anticipated to serve as an academic reference for the development of research related to the application of machine learning algorithms in analyzing consumer behavior in the e-commerce sector.

2. RESEARCH METODOLOGY

2.1 Consumer Purchase Pattern Analysis

Consumer purchase pattern analysis is an analytical approach aimed at examining and understanding consumer behavior based on transaction data generated in e-commerce activities[15]. This approach is conducted by processing historical transaction data to identify trends, regularities, and relationships among variables that influence consumer purchasing decisions. Through this analysis, meaningful patterns can be discovered, enabling understanding of customer behavior, improving prediction accuracy, and supporting strategic decision making for marketing, sales, and business development activities. [16]. In this study, the analysis focuses on transaction attributes such as purchase frequency, transaction value, product type, payment method, and promotional availability. Through the data analysis process, researchers can obtain a more comprehensive understanding of consumer behavior, including factors that tend to encourage or inhibit purchases. The results of this purchase pattern analysis can then be utilized as a basis for strategic decision-making, particularly in marketing planning, product management, and determining more targeted promotional strategies. Thus, consumer purchase pattern analysis plays an important role in supporting the development of data-driven e-commerce business strategies[17].

2.2 Naïve Bayes Algorithm

The Naïve Bayes algorithm is one of the classification algorithms that belongs to the probabilistic approach in the field of machine learning. This method applies Bayes' theorem to calculate the probability of a data instance belonging to a particular class based on its attributes. It assumes that each attribute contributes independently to the final classification result. Due to its simplicity, the algorithm is easy to implement and requires relatively small training data. Despite the independence assumption, Naïve Bayes often produces competitive performance. Therefore, it is widely used in various applications such as text classification, sentiment analysis, spam detection, and consumer behavior prediction tasks [18]. This algorithm calculates the probability of a class based on the attributes possessed by the data. It assumes that each attribute is independent of one another. This independence assumption simplifies computation, allowing efficient



processing while still producing reliable classification results for various problems, including consumer behavior analysis and prediction tasks [19]. Although this independence assumption is simple, the Naïve Bayes algorithm has proven to provide good performance in various classification cases, especially for data with categorical attributes[20]. In this study, the Naïve Bayes algorithm is used to analyze and classify e-commerce consumer purchasing decisions into two classes, namely Buy and Not Buy. The classification process is carried out by calculating class probabilities based on attributes such as purchase frequency, transaction value, product type, payment method, and promotion.[8]. The class with the highest probability value is selected as the classification result. The basic formula used in the Naïve Bayes algorithm is as follows:

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)} \tag{1}$$

With the following explanation:

$P(H|X)$ = Probability of hypothesis H given data X

$P(X|H)$ = Probability of data X given hypothesis H

$P(H)$ = Prior probability of hypothesis H

$P(X)$ = Probability of data X

2.3 Tahapan Penelitian

To ensure that the research is conducted systematically and in a structured manner, this study is carried out through several interconnected stages. Each stage is carefully designed to support the achievement of the research objectives and to produce valid and reliable conclusions. The research process begins with identifying the problem and determining the study scope. This is followed by collecting relevant data from appropriate sources. The collected data then undergoes preprocessing to improve quality and consistency. Next, the selected analytical method is applied to analyze the data. The results obtained are evaluated and interpreted to generate meaningful insights. The overall flow is illustrated in the research framework figure.

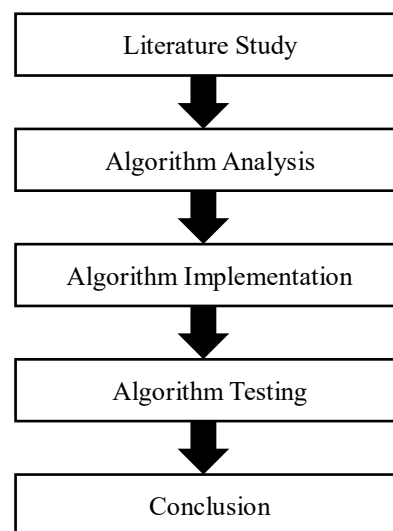


Figure 1. Research Framework

The following is a description of the research framework in Figure 1, which explains the research flow and stages systematically. This framework shows the research process starting from the literature study to obtain theoretical foundations, followed by data collection and preprocessing of e-commerce transaction data. Next, the Naïve Bayes algorithm is applied to analyze purchase patterns, followed by testing of classification results, and ending with drawing conclusions and providing research recommendations.

a. Literature Study

The initial stage of the research is conducted through a literature study to obtain an in-depth theoretical understanding related to e-commerce concepts, consumer behavior, data mining, and the Naïve Bayes algorithm. The reference sources used include textbooks, scientific articles, and research journals relevant to the research topic. This literature study aims to build a theoretical foundation and identify appropriate approaches for analyzing consumer purchase patterns.

b. Algorithm Analysis

At this stage, the researcher collects e-commerce transaction data used as the research object. The obtained data then undergoes selection and cleaning (data preprocessing) to ensure data quality. This process includes checking data completeness, standardizing attribute formats, and grouping attribute values into appropriate categories. The prepared data is then used as a dataset in the analysis process using the Naïve Bayes algorithm.



c. Algorithm Implementation

The analysis stage is conducted by calculating class probabilities and the probability of each attribute against the class using the Naïve Bayes algorithm. This calculation includes the prior probability of purchasing decisions and the conditional probability of each transaction attribute. The results of these probability calculations are used to identify consumer purchase patterns and determine the factors that most influence purchasing decisions.

d. Algorithm Testing

Setelah algoritma diterapkan, tahap selanjutnya adalah melakukan pengujian terhadap hasil klasifikasi. Pengujian dilakukan dengan membandingkan hasil prediksi algoritma dengan data aktual untuk melihat kesesuaian pola yang terbentuk. Evaluasi hasil ini bertujuan untuk menilai kemampuan algoritma Naïve Bayes dalam menganalisis pola pembelian konsumen e-commerce serta mengidentifikasi kekuatan dan keterbatasan metode yang digunakan.

e. Conclusion

The final stage of the research is conclusions based on the analysis and testing results that have been conducted. The conclusions are structured to answer the research objectives and provide an overview of e-commerce consumer purchase patterns obtained through the application of the Naïve Bayes algorithm. In addition, this stage also presents recommendations for further research development and the application of research results in the e-commerce business context.

3. RESULT AND DISCUSSION

3.1 Consumer Purchase Pattern Data Analysis

The analysis of consumer purchase pattern data in this study was conducted using an e-commerce transaction dataset presented in Table 1. The dataset consists of 30 transaction records, each representing consumer behavior in making or not making purchases on an e-commerce platform. Each transaction record contains six main attributes: Purchase Frequency, Transaction Value, Product Type, Payment Method, Promotion, and Purchase Decision. These attributes were selected because they are considered relevant in describing the characteristics and tendencies of consumer behavior in the purchasing decision-making process.

Table 1. Consumer Purchase Dataset

No	Frekuensi Beli	Nilai Transaksi	Jenis Produk	Metode Pembayaran	Promo	Keputusan Pembelian
1	High	Large	Electronics	Credit Card	No	Buy
2	Medium	Moderate	Accessories	E-Wallet	Yes	Buy
3	Low	Small	Fashion	COD	Yes	Not Buy
4	Medium	Moderate	Electronics	Transfer	No	Buy
5	High	Large	Electronics	Credit Card	Yes	Buy
6	Low	Small	Daily Needs	COD	No	Not Buy
7	Medium	Moderate	Accessories	E-Wallet	No	Buy
8	High	Large	Electronics	E-Wallet	Yes	Buy
9	Low	Small	Fashion	Transfer	Yes	Not Buy
10	Medium	Moderate	Fashion	E-Wallet	Yes	Buy
11	High	Large	Electronics	Credit Card	No	Buy
12	Low	Small	Accessories	COD	No	Not Buy
13	Medium	Moderate	Daily Needs	Transfer	Yes	Buy
14	High	Large	Electronics	E-Wallet	No	Buy
15	Low	Small	Fashion	COD	Yes	Not Buy
16	Medium	Moderate	Accessories	Transfer	No	Buy
17	High	Large	Electronics	Credit Card	Yes	Buy
18	Low	Small	Daily Needs	COD	No	Not Buy
19	Medium	Moderate	Electronics	E-Wallet	Yes	Buy
20	High	Large	Electronics	Transfer	No	Buy
21	Low	Small	Fashion	COD	Yes	Not Buy
22	Medium	Moderate	Accessories	E-Wallet	Yes	Buy
23	High	Large	Electronics	Credit Card	No	Buy
24	Low	Small	Daily Needs	Transfer	No	Not Buy
25	Medium	Moderate	Fashion	E-Wallet	Yes	Buy
26	High	Large	Electronics	Credit Card	Yes	Buy
27	Low	Small	Accessories	COD	No	Not Buy
28	Medium	Moderate	Electronics	Transfer	Yes	Buy
29	High	Large	Electronics	E-Wallet	No	Buy
30	Low	Small	Fashion	COD	Yes	Not Buy



3.2 Naïve Bayes Algorithm Implementation

E-commerce transaction data that has been stored is generally only utilized as sales archives and financial reports, and has not been maximized to generate strategic information. In fact, this data has great potential to be analyzed to identify customer purchasing patterns. Therefore, in this study, the Naïve Bayes algorithm is used to analyze and classify product purchasing decisions on an e-commerce platform.

The Naïve Bayes algorithm is a probabilistic-based classification method that works by calculating the probability of a class based on a combination of attributes possessed by a data instance. This method applies Bayes' Theorem and assumes that each attribute is independent of one another. Although this assumption is relatively simple and does not always match real-world conditions, Naïve Bayes has proven to be effective and efficient in various classification cases. Its advantages lie in fast computation, relatively small training data requirements, and the ability to produce reasonably good accuracy. Therefore, this algorithm is widely used in consumer behavior analysis, including predicting purchasing decisions.

1. Class Probability (Prior Probability)

The initial step in implementing the Naïve Bayes algorithm is determining the class probability (prior probability). Class probability represents the initial likelihood that a transaction belongs to a particular decision class before considering other attributes. In this study, the classes used are Buy and Not Buy, which describe the final consumer decision in e-commerce transactions. Class probability is calculated by comparing the number of data points in each class to the total number of transaction data used as the research dataset. This probability serves as the basis for further calculations in determining posterior probabilities in the Naïve Bayes algorithm. Mathematically, class probability is calculated using the following formula:

$$P(A) = \frac{n(A)}{n(S)} \tag{2}$$

Where:

P(A) = probability of class A,

n(A) = number of data in class A,

n(S) = total number of data in the dataset.

Based on observations of the e-commerce transaction dataset used in this study, the following data distribution was obtained:

Number of data with Buy decision: 19 transactions,

Number of data with Not Buy decision: 11 transactions,

Total number of transactions: 30 data.

Using the class probability formula, the initial probability for each purchase decision is calculated as follows:

$$P(Buy) = \frac{19}{30} = 0,63$$

$$P(Not Buy) = \frac{11}{30} = 0,37$$

The calculation results indicate that the majority of transaction data in the dataset leads to the Buy decision, with a probability of 0.63. Meanwhile, the Not Buy decision has a probability of 0.37, showing a lower occurrence in the dataset. This difference suggests that consumers are more inclined to complete purchases rather than cancel or abandon transactions. The higher probability of the Buy class also reflects stronger purchasing behavior within the observed data. These probabilities serve as an initial representation of consumer tendencies before considering other attributes. Therefore, understanding this distribution is important for classification. A summary of these class probabilities is presented in Table 2.

Table 2. Class Probability of Purchase Decision

Purchase Decision	Number of Data	Probability
Buy	19	0,63
Not Buy	11	0,37

These class probabilities are then used as the main component in calculating posterior probabilities in the Naïve Bayes algorithm. The prior probabilities are combined with the conditional probabilities of each transaction attribute, such as purchase frequency, transaction value, product type, payment method, and promotion. Through this combination, the algorithm determines the likelihood of each purchase decision class for a given data instance. This process helps identify patterns and relationships within the dataset. Consequently, determining class probabilities becomes a crucial foundation for further analysis. It supports accurate classification results and plays an essential role in analyzing consumer purchase patterns based on e-commerce transaction data.

2. Probability of Purchase Frequency Variable

After determining class probability (prior probability), the next step in the Naïve Bayes algorithm is calculating the conditional probability of each attribute value for each purchase decision class. In this section, the analyzed attribute is Purchase Frequency, which describes how often consumers conduct transactions on the e-commerce platform. The Purchase Frequency attribute in the dataset is classified into three categories: Low, Medium, and High. Conditional



probability is calculated by determining the proportion of each purchase frequency category within each decision class, namely Buy and Not Buy. This probability indicates the tendency of an attribute value to appear when a certain class occurs. Mathematically, conditional probability is calculated using the formula:

$$P(F = f_i|K) = \frac{\text{Number of data with } F=f_i \text{ in class } K}{\text{Total number of data in class } K} \tag{3}$$

Based on the grouping of transaction data, the distribution of purchase frequency against purchase decisions is presented in Table 3. The table shows the number of transactions within each purchase frequency category, which are then associated with the corresponding purchase decision classes, namely Buy and Not Buy. This distribution helps illustrate how often consumers with different purchase frequencies tend to complete or cancel transactions. The information obtained from this table is important for identifying consumer purchasing behavior patterns. Furthermore, the distribution serves as a fundamental basis for calculating conditional probabilities in the Naïve Bayes algorithm, which are later used to determine classification results and predict consumer purchase decisions.

Table 3. Probability of Purchase Frequency Variable

Purchase Frequency	Buy	not Buy	P(F Buy)	P(F Not_Beli)
Low	0	11	0,00	1,00
Medium	10	0	0,53	0,00
High	9	0	0,47	0,00

The probability values in the table are obtained from the following calculations:

Probability of Purchase Frequency = Medium for class Buy:

$$P(F = \text{Medium}|Buy) = \frac{10}{19} = 0.53$$

Probability of Purchase Frequency = High for class Buy:

$$P(F = \text{High}|Buy) = \frac{9}{19} = 0.47$$

Probability of Purchase Frequency = Low for class Not Buy:

$$P(F = \text{Low}|Not Buy) = \frac{11}{11} = 1$$

Meanwhile, no data was found with low purchase frequency in the Buy class, and no data with medium or high purchase frequency in the Not Buy class, resulting in conditional probabilities of zero for those combinations. This distribution indicates a clear separation between classes based on purchase frequency. Consumers with low purchase frequency consistently fall into the Not Buy category, suggesting limited engagement with the platform. In contrast, consumers with medium and high purchase frequency are strongly associated with the Buy class. This pattern demonstrates that purchase frequency significantly influences purchasing decisions and can serve as a reliable indicator for predicting consumer behavior in e-commerce transactions.

3. Probability of Transaction Value Variable

The next stage in implementing the Naïve Bayes algorithm is calculating the conditional probability for the Transaction Value attribute with respect to each purchase decision class. The Transaction Value attribute represents the nominal amount of transactions made by consumers within a certain period and is classified into three categories: Small, Medium, and Large. Conditional probability is calculated by comparing the number of occurrences of each transaction value category in each purchase decision class (Buy and Not Buy) to the total data in that class. This probability reflects the likelihood of a certain transaction value appearing when a particular purchase decision occurs. Mathematically, the conditional probability for the Transaction Value attribute is formulated as follows:

$$P(N = n_i|K) = \frac{\text{Number of data with } N = n_i \text{ in class } K}{\text{Total number of data in class } K} \tag{4}$$

Based on the processing of the e-commerce transaction data, the distribution of transaction values against purchase decisions is presented in Table 4. The table illustrates the relationship between transaction value magnitude and consumer tendency to make or not make purchases. From the distribution, consumers with medium to high transaction values tend to make purchases more frequently than those with low transaction values. This pattern indicates that transaction value is an important factor influencing consumer purchasing decisions.

Table 4. Probability of Transaction Value Variable

Transaction Value	Buy	Not Buy	P(N Buy)	P(N Noy Buy)
Small	0	11	0,00	1,00
Medium	10	0	0,53	0,00
large	9	0	0,47	0,00

The probability values above are obtained from the following calculations:

Probability of Transaction Value = Medium for class Buy:



$$P(N = \text{Medium}|\text{Buy}) = \frac{10}{19} = 0.53$$

Probability of Transaction Value = Large for class Buy:

$$P(N = \text{Large}|\text{Buy}) = \frac{9}{19} = 0.47$$

Probability of Transaction Value = Small for class Not Buy:

$$P(N = \text{Small}|\text{Not Buy}) = \frac{11}{11} = 1$$

Based on these results, no transaction data with small value was found in the Buy class, and no medium or large transaction values were observed in the Not Buy class, resulting in conditional probabilities of zero for those combinations. This distribution clearly separates the influence of transaction value on purchase decisions. The analysis indicates that transaction value is a highly influential factor in consumer purchasing behavior. Consumers with small transaction values tend not to proceed with purchases, which may reflect lower interest, budget limitations, or exploratory browsing behavior. Conversely, medium to large transaction values are strongly associated with Buy decisions. This suggests that higher spending levels reflect stronger commitment, greater product interest, and increased consumer confidence in completing transactions on the e-commerce platform.

4. Probability of Product Type Variable

The next step in classification using the Naïve Bayes algorithm is calculating conditional probabilities for the Product Type attribute for each purchase decision class. This attribute represents product categories offered on the e-commerce platform and is classified into four groups: Electronics, Accessories, Fashion, and Daily Needs. The calculation aims to determine how likely a product type is associated with Buy or Not Buy decisions based on historical transaction data.

Mathematically, the conditional probability for Product Type is formulated as follows:

$$P(J = j_i|K) = \frac{\text{Number of data with } J = j_i \text{ in class } K}{\text{Total Number of data in class } K} \tag{5}$$

Based on the dataset processing, the distribution of product types against purchase decisions is shown in Table 5. The table illustrates how each product category relates to Buy and Not Buy decisions. This information provides an overview of consumer preferences, identifies categories with higher purchase likelihood, and helps understand purchasing tendencies across different product types in e-commerce transactions.

Table 5. Probability of Product Type Variable

Product Type	Buy	Not Buy	P(J Buy)	P(J Not Buy)
Electronics	10	0	0,53	0,00
Accessories	4	2	0,21	0,18
Fashion	3	7	0,16	0,64
Daily Needs	2	2	0,10	0,18

The probability values in the table above are obtained from the following calculations:

Probability of Product Type = Electronics for the Buy class:

$$P(J = \text{Elektronics}|\text{Buy}) = \frac{10}{19} = 0.53$$

Probability of Product Type = Fashion for the Not Buy class:

$$P(N = \text{Fashion}|\text{Not Buy}) = \frac{7}{11} = 0.64$$

The analysis results show that electronic products have the highest probability in relation to purchase decisions, indicating that e-commerce consumers tend to be more interested in purchasing product categories with high utility value and long-term usage characteristics. This may be due to the high demand for technology products and consumers' trust in e-commerce platforms to provide electronic items. Conversely, fashion products show the highest probability in the not-buy class, indicating that consumers tend to be more cautious when purchasing fashion products online. Factors such as uncertainty in size, material, color, and the suitability of the product with personal preferences may influence consumers' decisions not to proceed with the purchase. Meanwhile, the accessories and daily necessities categories show a relatively balanced distribution between buy and not-buy decisions. These findings indicate that purchase decisions in these two categories are strongly influenced by other factors, such as price, promotions, and payment methods, which will be analyzed in the following variables. Thus, the probability calculation results for the product type variable illustrate that product category is an important factor in shaping e-commerce consumer purchasing patterns. This information can be utilized as a basis for developing product grouping strategies, determining promotional priorities, and optimizing product offerings on e-commerce platforms.

5. Probability of Payment Method Variable

At this stage, conditional probabilities for the Payment Method attribute are calculated for each purchase decision class (Buy and Not Buy). This variable represents how consumers complete transactions and is classified into four categories: E-Wallet, Credit Card, Bank Transfer, and Cash on Delivery (COD). Mathematically:

$$P(M = m_i|K) = \frac{\text{Number of data with } M = m_i \text{ in class } K}{\text{Total number of data in class } K} \tag{6}$$





Based on the results of processing the e-commerce transaction data, the distribution of payment methods with respect to purchase decisions is presented in Table 6 below. This table shows a comparison between various payment methods, such as cash and non-cash, and consumer decisions to buy or not buy. Through this data presentation, it can be identified which payment methods are most frequently used and their influence on consumer purchasing decision tendencies.

Table 6. Probability of Payment Method Variable

Payment Method	Buy	Not Buy	P(M Buy)	P(M Not Buy)
E-Wallet	7	0	0,37	0,00
Credit Card	6	0	0,32	0,00
Bank Transfer	6	2	0,32	0,18
COD	0	9	0,00	0,82

The probability values in the table above are obtained from the following calculations:

Probability of Payment Method = E-Wallet for the Buy class:

$$P(M = E - wallet|Buy) = \frac{7}{19} = 0.37$$

Probability of Payment Method = COD for the Not Buy class:

$$P(M = COD|Not Buy) = \frac{9}{11} = 0.82$$

The analysis results indicate that payment methods have a significant influence on consumer purchase decisions. The Cash on Delivery (COD) method has a very high probability in the Not Buy class, indicating that consumers who choose the COD method tend not to complete the purchase process. This condition may be caused by factors such as stock uncertainty, order cancellations, or low purchase commitment from consumers who select payment upon delivery. In contrast, the E-Wallet and Credit Card payment methods show high probabilities in the Buy class. This indicates that consumers who use non-cash and digital-based payment methods have a greater tendency to complete purchase transactions. Factors such as convenience, speed, security, and incentives such as cashback or digital payment promotions may serve as the main drivers in purchase decisions. Meanwhile, the Bank Transfer method shows a relatively balanced distribution between buy and not-buy decisions, indicating that this method is still used by consumers with varying levels of purchase commitment.

6. Probability of Promotion Variable

The Promotion variable in this study is used to represent the presence of special offers, such as discounts, cashback, or price reductions, provided by the e-commerce platform to consumers during transactions. The promotion variable is classified into two categories, namely Yes (promotion available) and No (no promotion available). The analysis of this variable aims to determine the extent to which the presence of promotions influences consumer purchase decisions. The calculation is performed by determining the conditional probability of each value of the Promotion variable with respect to each purchase decision class, namely Buy and Not Buy. These probabilities are calculated based on the proportion of the occurrence of the Promotion attribute in each decision class. Mathematically, the conditional probability for the Promotion variable is formulated as follows:

$$P(P = p_i|K) = \frac{\text{Number of data with } P = p_i \text{ in class } K}{\text{Total number of data in class } K} \tag{7}$$

Based on the results of processing e-commerce transaction data, the distribution of the Promotion variable with respect to purchase decisions is presented in Table 7 below. The table illustrates the relationship between the presence of promotions and consumers' decisions to make a purchase or not. From the presented data, it can be seen that transactions accompanied by promotions tend to have a higher number of buy decisions compared to transactions without promotions. This indicates that promotions play a role in attracting consumer interest, although they are not the sole determining factor. This analysis helps to more clearly understand the influence of promotional strategies on consumer purchasing behavior.

Tabel 7. Probability of Promotion Variable

Promotion	Buy	Not Buy	P(P Buy)	P(P Not Buy)
Yes	10	6	0,53	0,55
No	9	5	0,47	0,45

The probability values in the table above are obtained from the following calculations:

Probability of Promotion = Yes for the Buy class:

$$P(P = Yes|Buy) = \frac{10}{19} = 0.53$$

Probability of Promotion = Yes for the Not Buy class:

$$P(Yes|Not Buy) = \frac{6}{11} = 0.55$$



The analysis results indicate that the presence of promotions has a positive influence on purchase decisions, as reflected by the relatively high probability value of Promotion = Yes in the Buy class. This suggests that promotions remain one of the factors capable of attracting consumers' interest in making purchases on e-commerce platforms. However, the calculation results also show that the probability value of Promotion = Yes in the Not_Buy class is still quite large. This finding indicates that the presence of promotions does not always guarantee a purchase, since consumer decisions are also influenced by other factors, such as previous purchase frequency, transaction value, product type, and the payment method used. Conversely, conditions without promotions still show the occurrence of purchase decisions, although with a lower probability compared to conditions with promotions. This suggests that some consumers continue to make purchases based on their needs, product loyalty, or trust in the e-commerce platform, without fully depending on promotional offers.

4. CONCLUSION

Based on the implementation results of the Naïve Bayes algorithm on e-commerce transaction data, it can be concluded that this method can be effectively used to analyze and classify consumer purchase decisions. The class probability calculation results show that most transaction data fall into the Buy class, indicating that consumers in the dataset tend to make purchases more frequently than cancel transactions. This finding reflects a relatively active consumer behavior pattern in shopping on e-commerce platforms. The conditional probability analysis for each variable shows that purchase frequency and transaction value are the most influential factors in purchase decisions. Consumers with low purchase frequency and small transaction values are more often classified into the Not_Buy class, whereas consumers with moderate to high purchase frequency and medium to large transaction values have a strong tendency to make purchases. This indicates that loyalty level and spending amount play important roles in shaping consumer decisions. The product type variable also has an influence, where electronic products have the highest probability for the Buy decision, while fashion products are more frequently associated with the Not_Buy decision. In addition, digital-based payment methods such as e-wallets and credit cards are proven to better support purchase completion compared to the COD method. Promotional factors also contribute to increasing purchase interest, although they are not the sole determinant. Overall, this study shows that the Naïve Bayes algorithm is capable of providing useful insights in understanding e-commerce consumer purchasing behavior patterns.

REFERENCES

- [1] I. T. Laiya and D. Fahriani, "Analisis Penggunaan E-Commerce terhadap Peningkatan Pendapatan UMKM Pengrajin Tas," vol. 8, no. 2, pp. 185–193, 2025.
- [2] N. Sutisna and Sutrisna, "Implementasi Sistem Informasi dalam Mendukung Perilaku Pembelian Terhadap Keputusan Pembelian E-Commerce," vol. 2, no. 1, pp. 20–30, 2023.
- [3] P. Salsabila, E. Wahyudin, G. Dwilestari, K. Kaslani, and F. Subhiyanto, "Penerapan Algoritma Fp-Growth Untuk Mengetahui Pola Pembelian Konsumen Di Warung Makan Dede," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 1, pp. 1221–1128, 2024, doi: 10.36040/jati.v8i1.8964.
- [4] V. C. Nisa and F. N. Khasanah, "Algoritma Apriori Dalam Identifikasi Pola Pembelian Konsumen Pada Produk Minuman," vol. 8, no. 2, pp. 156–164, 2023.
- [5] L. Sasabone, E. Sudarmanto, and S. Adiwijaya, "Pengaruh E-commerce dan Kemudahan Transaksi Terhadap Perubahan Pola Konsumsi Dalam Era Digital Di Indonesia," vol. 1, no. 01, pp. 32–42, 2023.
- [6] N. S. Fauzia and R. D. Dana, "Implementasi Algoritma Naive bayes dalam Klasifikasi Status Kesejahteraan Masyarakat Desa Gunungsari," 2023.
- [7] Rayuwati, H. Gemasih, and I. Nizar, "IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMREDIKSI TINGKAT," 2022.
- [8] M. Z. Haq *et al.*, "Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial," vol. 13, pp. 2020–2025, 2024.
- [9] E. Laia and M. Yamin, "Penerapan Algoritma Naïve Bayes dalam Menganalisis Sentimen pada Review Pengguna E-Commerce," vol. 4, no. 1, pp. 305–316, 2023, doi: 10.30865/klik.v4i1.1186.
- [10] M. Wahyudi and M. S. Novelan, "PERBANDINGAN KINERJA K-NN DAN NAIVE BAYES DALAM TATA KELOLA TEKNOLOGI INFORMASI UNTUK MENGANALISA TINGKAT KEPUASAN PENGGUNA APLIKASI SHOPEE," *J. Sci. Soc. Res.*, vol. 8, no. 3, pp. 3424–3431, 2025.
- [11] A. D. Kuswanto, S. I. Puro, J. Hariyan, R. Rafliansyah, M. R. Aziz, and P. V. Rajagukguk, "Analisa Data Shopping Trends Menggunakan Algoritma Klasifikasi Dengan Metode Naive Bayes," vol. 2, no. 3, 2024.
- [12] P. A. Saputra, S. S. Irawan, and R. Prianto, "Analisis Faktor Dominan Minat Beli Generasi Z di Shopee Menggunakan Algoritma Naïve Bayes," 2025.
- [13] B. Z. Ramadhan, I. Riza, and I. Maulana, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Dengan Menggunakan Algoritma Naïve Bayes," vol. 6, no. 2, pp. 220–225, 2022.
- [14] A. N. Puspitasari, Y. Findawati, and Y. Rahmawati, "ANALISIS SENTIMEN TWEET PENGGUNA E-COMMERCE DENGAN MENGGUNAKAN METODE KLASIFIKASI NAÏVE BAYES," vol. 9, no. 3, pp. 1123–1132, 2024.
- [15] M. S. Mauludin, A. D. Saputra, A. Z. Sari, and I. Munawaroh, "Analisis Perilaku Konsumen Dalam Transaksi Di e-Commerce," vol. 1, no. 1, 2022.
- [16] A. Zhalifunnas, A. Fauzi, M. Z. Augestri, and R. Rahman, "Analisis Peluang Bisnis Melalui Pemahaman Perilaku Konsumen dengan Pendekatan Matematika Ekonomi," vol. 2, no. 1, pp. 29–35, 2024.



- [17] G. Oktavia, “ANALISIS POLA PEMBELIAN KONSUMEN BERDASARKAN DATA HISTORIS DENGAN METODE ASOSIASI DI PT EXCELLINDO PERKASA,” *Comput. Sci. Ind. Eng.*, vol. 10, no. 3, pp. 88–98, 2024.
- [18] T. F. Mustafa and H. Alfianti, “KLASIFIKASI BERITA PALSU BERBAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE BAYES BERBASIS WEB,” pp. 657–663, 2025.
- [19] F. A. Indriyani, A. Fauzi, and S. Faisal, “Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine Tiktok application sentiment analysis using naïve bayes algorithm and support vector machine,” vol. 10, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.
- [20] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” vol. 8, no. April, pp. 714–725, 2024, doi: 10.30865/mib.v8i2.7458.