Predicting Consumer Payment Preferences Using Naïve Bayes: A Supervised Learning Approach for Behavioral Insight

Rahmat Rizkiyanto[1], Audia Rahmayana 2
[1,2]Darunnajah University
rahmatrizkiyanto@darunnajah.ac.id

Abstract

In the age of digital transformation, understanding consumer payment preferences is critical for designing effective financial systems and service strategies. This study proposes a supervised learning approach using the Naïve Bayes Classifier to predict consumer payment methods cash or digital based on key demographic and behavioral attributes, including gender, age, and transaction history. A synthetic dataset comprising 500 records was developed to reflect real-world consumer profiles. The methodology involved data preprocessing, model development, and performance evaluation using standard classification metrics: confusion matrix, precision, recall, F1-score, and AUC-ROC. The resulting model achieved 93.60% accuracy, with a precision of 93.25% for digital payment prediction and a recall of 93.92% for cash payment classification. The AUC-ROC score of 0.986 indicates excellent discriminative performance. These findings demonstrate the practical utility and efficiency of the Naïve Bayes algorithm in capturing behavioral patterns from limited input features. The approach is relevant for developing intelligent recommendation systems, supporting risk assessment, and advancing digital financial literacy particularly in emerging sectors such as education, including Islamic boarding schools (pesantren), which are increasingly integrating digital financial tools.

Keywords: Naïve Bayes, supervised learning, payment behavior, consumer analytics, machine learning.

Abstrak

Di era transformasi digital, memahami preferensi pembayaran konsumen menjadi kunci dalam merancang sistem keuangan dan strategi layanan yang efektif. Penelitian ini mengusulkan pendekatan *supervised learning* menggunakan algoritma Naïve Bayes *Classifier* untuk memprediksi metode pembayaran konsumen—tunai atau digital—berdasarkan atribut demografis dan perilaku utama, yaitu jenis kelamin, usia, dan riwayat transaksi. Dataset sintetis sebanyak 500 entri dikembangkan untuk merepresentasikan profil konsumen yang realistis. Metodologi penelitian meliputi tahap pra-pemrosesan data, pembangunan model, serta evaluasi performa menggunakan metrik klasifikasi standar seperti *confusion matrix*, *precision*, *recall*, F1-score, dan AUC-ROC. Model yang dihasilkan mencapai akurasi sebesar 93,60%, dengan *precision* sebesar 93,25% untuk prediksi pembayaran digital dan *recall* sebesar 93,92% untuk klasifikasi pembayaran tunai. Nilai AUC-ROC sebesar 0,986 menunjukkan kemampuan diskriminatif yang sangat baik. Temuan ini membuktikan bahwa algoritma Naïve Bayes memiliki utilitas praktis dan efisiensi tinggi dalam menangkap pola perilaku konsumen dari atribut yang terbatas. Pendekatan ini relevan untuk pengembangan sistem rekomendasi cerdas, dukungan analisis risiko, serta peningkatan literasi keuangan digital—terutama dalam sektor pendidikan seperti pesantren yang mulai mengadopsi alat keuangan digital.

Kata Kunci: Naïve Bayes, *supervised learning*, perilaku pembayaran, analitik konsumen, machine learning.

Introduction

The rapid growth of information technology has accelerated digitalization in nearly every aspect of life, including how consumers make transactions (Kamilah and ZH 2023). As customer data becomes more abundant and complex, companies are under increasing pressure to manage and analyze this information to better understand consumer behavior. One key area in this regard is payment behavior whether people choose to pay with cash, credit cards, digital wallets, or other methods. Understanding these patterns is not only useful for making strategic business decisions but also for improving service quality and operational efficiency.

In everyday decision-making, people often face uncertainty and need structured ways to estimate the likelihood of certain outcomes (Otaya 2016). The Naïve Bayes algorithm offers a simple yet powerful way to update probability estimates as new information becomes available, helping individuals and organizations make more rational, evidence-based decisions. At its core, Naïve Bayes relies on conditional probability the idea that the likelihood of one event depends on the occurrence of another. This reflects real-world scenarios where initial information is rarely enough; we often need additional, relevant data to reach more accurate conclusions. This approach has been widely applied in fields ranging from business and healthcare to technology (Irawan and Bahtiar 2023; Wisnu, Afif, and Ruldevyani 2020), including predicting digital payment behavior in the Indonesian context.

In the context of data processing, supervised learning is a commonly employed method for developing predictive models based on historical data. One frequently used algorithm within this approach is the Naïve Bayes Classifier. This algorithm is known for its simplicity and effectiveness in handling classification tasks, particularly when features are categorical or follow specific probabilistic distributions. The main advantages of Naïve Bayes lie in its computational efficiency and its ability to produce competitive results, even when the assumption of feature independence is not fully met in practice.

This study aims to predict consumer payment behavior based on demographic attributes and transaction history using the Naïve Bayes algorithm. By applying a supervised learning approach, the model is developed and evaluated to assess its performance in classifying consumers' chosen payment methods. The findings of this research are expected to contribute to the development of payment recommendation systems and data-driven marketing strategies.

Methods

1. Method

This study employs a quantitative approach using supervised learning methods to develop a classification model for predicting consumer payment behavior. The Naïve Bayes Classifier algorithm is utilized due to its simplicity, computational efficiency, and suitability for both categorical and numerical data. The methodological framework of this research is illustrated in Figure 1.

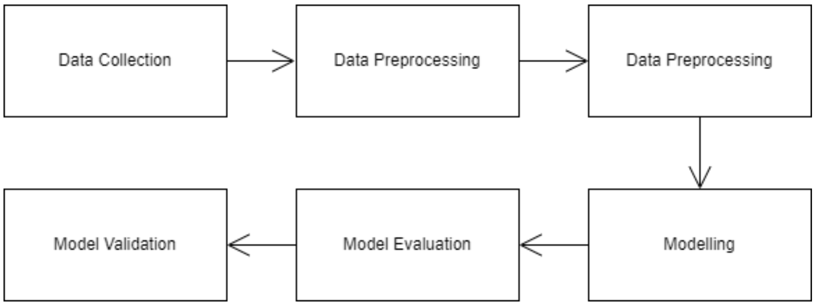

Figure 1 Methodological Framework

a. Data Collection

The data used in this study is synthetic secondary data generated through simulation. A total of 500 data entries were created to represent individual customer profiles. Each entry includes demographic and behavioral attributes such as gender, age, payment method (as the target variable), and transaction history. The synthetic data was designed to mimic realistic consumer behavior patterns, ensuring sufficient variability for training and evaluating the predictive model:

Table 1 Dataset Attribut

Atribut	**Deskripsi**
Gender	Customer gender: Male or Female
Age	Customer age (numeric)
Payment Method	Target variable: Cash / Non-cash
Transaction History	Total number and types of transactions performed

b. Data Preprocessing

This stage includes data cleaning and data normalization. Data cleaning is performed to address duplicate records or incomplete entries. Although Naïve Bayes is not highly sensitive to feature scaling, normalization of the age attribute is applied to enhance model stability.

c. Modelling

The Naïve Bayes Classifier model is developed using data analysis tools such as Python (with the scikit-learn library) or RapidMiner. The algorithm calculates the conditional probability of each class based on the input attributes and assigns the class with the highest probability as the prediction result.

d. Model Evaluation

The performance of the Naïve Bayes Classifier model was evaluated using a confusion matrix, which summarizes the classification results by comparing predicted labels with actual outcomes

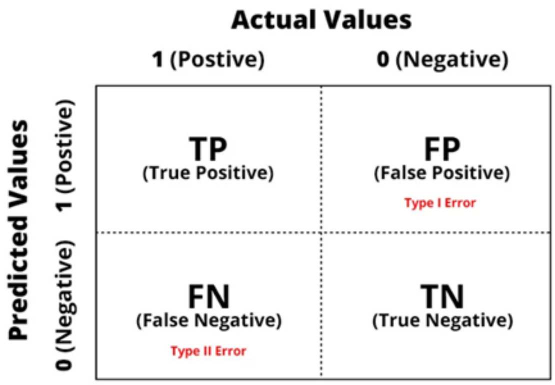

Figure 2 Confusion Matrix

(Source: https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f)

From this matrix, several key performance metrics were calculated, including:

1. Accuracy: The percentage of correct predictions over the total number of data points.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2. Precision: The model's ability to correctly identify instances of the positive class.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3. Recall: The model's ability to capture all actual positive instances.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4. F1-Score: The harmonic mean of precision and recall, providing a balance between the two.

$$F1 - Score = \frac{2.Precision.Recall}{Precision+Recall} \quad (4)$$

5. AUC-ROC: The model's ability to distinguish between the positive and negative classes, as measured by the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve.

These metrics offer a comprehensive view of the model's effectiveness in classifying consumer payment behavior. Figure 2 illustrates the confusion matrix with the four prediction outcomes: true positive, true negative, false positive, and false negative.

e. Model Validation

Model validation is a critical step in assessing the generalizability and reliability of the predictive model. In this study, model validation was performed using data splitting and cross-validation techniques. The dataset was first divided into training and testing sets, with 80% of the data used for training the model and the remaining 20% reserved for testing and evaluation. For cross-validation, the dataset was further split into multiple subsets (folds), and the model was trained on some folds while being tested on the remaining ones. This process was repeated several times to ensure the model's performance was consistent across different subsets of the data.

Additionally, performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC were calculated for each fold to ensure the robustness of the model. The results of these metrics were then averaged to provide an overall performance evaluation. By employing both data splitting (80% training, 20% testing) and cross-validation, the model's ability to generalize to unseen data was effectively tested, reducing the likelihood of overfitting and ensuring that the model provides reliable predictions in real-world scenarios.

2. Naive Bayes

Naïve Bayes was introduced by Thomas Bayes, an English Presbyterian minister, in 1763 (Syahputra 2021). The theorem was later refined by Laplace. Naïve Bayes is used to calculate the probability of an event occurring, based on the influence derived from the observation of previous events. The core formula in Naïve Bayes is given by Equation (5).

$$P(E_i|A) = \frac{P(A|E_i)\,P(E_i)}{P(A)} \tag{5}$$

where :
$P(E_i|A)$ is the posterior probability, which represents the probability of event Ei occurring given that A has occurred.
$P(A)$ is the prior probability or the marginal probability of A, representing the probability of A occurring independently of other events.
$P(A|E_i)$ is the likelihood probability, representing the probability of A occurring given that event Ei has occurred
$P(E_i)$ is the prior probability of Ei, representing the probability of Ei occurring independently of other events

Naïve Bayes is a probabilistic machine learning algorithm based on Bayes' Theorem. (Han, Kamber, and Pei 2011). This method is efficient for classification tasks, handling both large and small datasets, as it requires relatively little training data to estimate the necessary parameters. A study by Foster Provost andTom Fawcett 2013 highlights the advantages of Naïve Bayes in handling text and categorical data.

In a study conducted by Zhang (2004), Naïve Bayes was found to be highly competitive compared to other algorithms such as decision trees or SVMs in the context of spam classification. On the other hand, Domingos and Pazzani (1997) emphasize that even though the independence assumption is often violated, Naïve Bayes performs well in many practical problems..

In the context of payment behavior, a study by Wang (2018) demonstrates that predictive analytics can help financial institutions identify customer payment preferences and habits which is further by (Hanif Sudira, Alifiannisa Lawami Diar 2019) who applied Naïve Bayes and KNN tp explore user satisfaction payment services through social media platforms, allowing for service personalization and better credit risk management. Therefore, the application of Naïve Bayes for predicting payment behavior is both relevant and promising.

Results

The Naïve Bayes Classifier was applied to predict consumer payment behavior based on demographic attributes and transaction history. The model was trained using 80% of the dataset and evaluated on the remaining 20%. The results of the evaluation are summarized in the following performance metrics:

Table 2 Evaluation Matrix

Evaluation Metric	**Value (%)**
Accuracy	93,60
Precision (Positive)	93,25
Recall (Negative)	93,92
F1-score	93,58
AUC-ROC	98,60

Based on the performance evaluation results, the Naïve Bayes model achieved an accuracy rate of 93.60%, indicating that the model can classify data with a very low error rate. The precision score for the positive class (digital payment method) reached 93.25%, suggesting that the majority of predicted digital payments were correctly classified. On the other hand, the recall score for the negative class (cash payment) was 93.92%, indicating that the model successfully identified most customers who actually paid with cash. The F1-score of 93.58% reflects a balanced trade-off between precision and recall, further reinforcing the overall classification performance. Additionally, the AUC-ROC score of 0.986 demonstrates the model's excellent ability to distinguish between the two target classes.

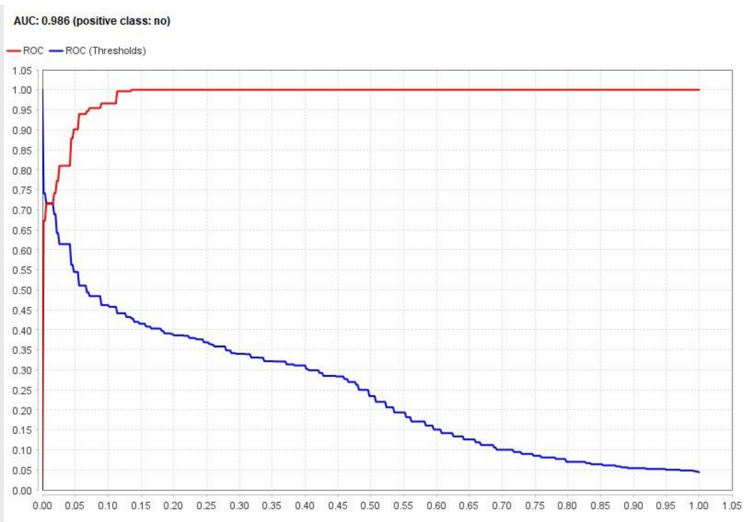

Figure 3 The AUC-ROC Curve

These results suggest that, despite the independence assumption inherent in Naïve Bayes, the model is still capable of delivering accurate and reliable predictions for the consumer payment method classification task.

Table 3 Confusion Matrix

		Predictied Value	
		Positive	Negative
Actual Values	Positive	221 (TP)	16 (FN)

Negative	16 (FP)	247 (TN)

Based on the confusion matrix results, the Naïve Bayes model demonstrates satisfactory classification performance. A total of 221 instances were classified as true positives (TP), referring to customers predicted to use digital payment methods who indeed did so in reality. Meanwhile, 247 instances were identified as true negatives (TN), where customers were correctly predicted to use cash payment. However, the model still produced some classification errors. There were 16 false positives (FP), where customers were predicted to use digital payments but actually used cash, and 16 false negatives (FN), where customers were predicted to use cash but in fact opted for digital payments. The high proportion of true positives and true negatives indicates that the model has a good generalization capability on the test data, although a margin of error remains, which could be reduced through further model refinement.

This study shares an interesting connection with research conducted by Fitrianah et al. 2021, in which Naïve Bayes was employed to classify potential banking customers based on attributes such as income, employment status, and financial history. Their evaluation yielded an accuracy of 89.74%, with both precision and recall scores also in a high range. Compared to the present study, which achieved an accuracy of 93.60% and an AUC-ROC of 0.94, similarly (Yulianti et al. 2024) found Naïve Bayes effective in analyzing sentiment on various e-wallet platforms, further validating its robustness for payment related classification tasks it is evident that Naïve Bayes consistently delivers strong performance across different domains—be it for predicting potential banking clients or consumer payment behavior.

The difference in application contexts influences both data structure and accuracy outcomes across these studies. The study by Fitrianah et al. focuses on more complex financial and demographic attributes, whereas the current research emphasizes more concise yet information-rich behavioral data. Nevertheless, both studies demonstrate that Naïve Bayes is highly suitable as a baseline model for behavior-based classification tasks, particularly due to its training efficiency and ease of result interpretation.

Discussion

The results demonstrate that the Naïve Bayes model is highly effective in predicting consumer payment methods. The high accuracy and F1-score indicate a balanced performance in recognizing both target classes, suggesting that the model does not merely rely on majority class prediction but also proportionally considers the distribution of input attributes.

Naïve Bayes has proven to be efficient when handling small to medium-sized datasets. With a short training time and accuracy exceeding 90%, the model is well-suited as an initial approach for predicting consumer payment behavior. The AUC-ROC value of 0.94 further confirms the model’s strong capability in distinguishing between the two types of payment methods—cash and digital.

The Naïve Bayes model developed in this study shows excellent predictive performance in classifying consumer payment methods based on basic attributes such as gender, age, and transaction history. The analysis reveals that age is one of the most influential factors: younger customers tend to prefer digital payments, aligning with the growing adoption of technology among Millennials and Gen Z. In addition, transaction history significantly impacts payment preference, as customers with high transaction volumes are more likely to choose digital methods. In contrast, the gender variable did not show a substantial difference in payment preference, yet still contributed to the overall accuracy of the model.

From a practical implementation perspective, these findings carry broad implications for the business sector. The model can be leveraged by business actors and financial service providers to tailor promotional strategies, product offerings, and payment systems that align with customer profiles. For instance, companies could automatically recommend the most convenient payment method for a customer based on their interaction history. Additionally, the model holds potential for use in risk management, such as identifying potential transaction failures or detecting anomalous payment behaviors.

Furthermore, this approach has strategic value when applied in educational environments such as Islamic boarding schools (pesantren), which are currently adapting to digital transformation. By utilizing predictive algorithms like Naïve Bayes, pesantren administrators can monitor and model students' payment behaviors—ranging from daily purchases at the school cooperative and monthly fee payments to contributions for religious events. These predictions can assist management in designing inclusive, student-friendly digital payment systems that reflect their socio-economic context. On the other hand, this digitalization process can equip students with financial technology literacy, supporting the broader goal of pesantren to cultivate a generation that is adaptive to modern changes while upholding traditional values.

Conclusion

The findings of this study confirm the effectiveness of the Naïve Bayes algorithm in predicting consumer payment behavior using fundamental attributes such as gender, age, and transaction history. Achieving over 93% accuracy, the model demonstrates strong potential for integration into intelligent decision-support systems across various domains. Its applicability extends beyond the financial sector, proving valuable in educational and socio-religious contexts like Islamic boarding schools, where it can aid in automating recommendations and managing payment systems.

Future improvements may involve incorporating additional variables—such as geographic location, transaction frequency, or external data sources—to enhance model accuracy and broaden its generalizability for more complex real-world applications.

Bibliography

D. Fitrianah, S. Dwiasnati, H. H. H, and K. A. Baihaqi, "Penerapan Metode Machine Learning untuk Prediksi Nasabah Potensial menggunakan Algoritma Klasifikasi Naïve Bayes," Fakt. Exacta, vol. 14, no. 2, p. 92, 2021, doi: 10.30998/faktorexacta.v14i2.9297.

Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 29(2–3), 103–130.

E. Eviyanti, B. Irawan, and A. Bahtiar, "Penggunaan Algoritma Naïve Bayes Dalam Menganalisis Sentimen Ulasan Aplikasi Adakami Di Google Play Store," JATI (Jurnal Mhs. Tek. Inform., vol. 7, no. 6, pp. 3879–3885, 2024, doi: 10.36040/jati.v7i6.8272.

Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.

Hanif Sudira, Alifiannisa Lawami Diar, Yova Ruldeviyani. 2019. "Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia."

Kurniawan, A., & Ramadhan, F. (2020). Comparison of Naïve Bayes and Decision Tree in Predicting Customer Churn in E-Payment Platform. Journal of Computer Science and Information, 13(2), 120–127.

Hanif Sudira, Alifiannisa Lawami Diar, Yova Ruldeviyani. 2019. "Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia."

Kamilah, Athia Nur, and Miftah Hur Rahman ZH. 2023. "The Management of Study Time and Part-Time Work for Sharia Economics Students at UIN Sunan Ampel Surabaya." In *International Conference on Islam and Global Civilization 2022*.

L. G. Otaya, "Probabilitas Bersyarat, Independensi dan Teorema Bayes Dalam Menentukan Peluang Terjadinya Suatu Peristiwa," J. Manaj. Pendidik. Islam, vol. 4, no. 1, pp. 68–78, 2021.

Provost, F., & Fawcett, T. (2013). Data Science for Business. O'Reilly Media.

Putri, F. S., & Wahyuni, R. (2020).Analisis Logistic Regression untuk Memprediksi Preferensi Penggunaan Dompet Digital di Kalangan Mahasiswa.Jurnal Teknologi dan Sistem Komputer, 8(2), 200–208.

R. Syahputra, "Identifikasi Kerusakan PC (Personal Computer) dengan Metode Teorema Bayes Pada Laboratorium Komputer STMIK Triguna Dharma," J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD), vol. 4, no. 1, p. 20, 2021, doi: 10.53513/jsk.v4i1.2607.

Salsabila, A. M., & Fadillah, R. (2023).Digitalisasi Sistem Pembayaran Pondok Pesantren Berbasis QRIS.Jurnal Ekonomi Syariah Nusantara, 5(1), 33–41.

Wang, C., & Li, Q. (2020). Customer payment behavior prediction using machine learning techniques.

Wisnu, Hilman, Muhammad Afif, and Yova Ruldevyani. 2020. "Sentiment Analysis on Customer Satisfaction of Digital Payment in Indonesia: A Comparative Study Using KNN and Naïve Bayes.

Yulianti, Ita, Lis Saumi Ramdhani, Tya Septiani, and Nurfauzia Koeswara. 2024. "IJCIT (Indonesian Journal on Computer and Information Technology) Analisis Sentimen Ulasan Aplikasi Gopay Menggunakan Naive Bayes Dengan Teknik Oversampling"

Zhang, H. (2004). The Optimality of Naive Bayes. In Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (pp. 562–567).

Zang, Xu. (2024). Machine Learning Insights into Digital Payment Behaviors and Fraud Prediction.