Analysis of the Use of the Naïve Bayes Method in Predicting Consumer Behavior Based on Historical Data: A Case Study of PT Syntax Corporation Indonesia

Agis Ahmad Rodiansjah

Universitas Neger Semarang

Corresponding email: agisahmad223@students.unnes.ac.id

Abstract

This study aims to analyze the effectiveness of the Naïve Bayes method in predicting consumer behavior in PT. Syntax Corporation Indonesia based on historical data that has been collected. In the digital era, the use of historical data to understand consumer behavior is an important step in a company's strategic decision-making. The Naïve Bayes method was chosen because it has efficient and accurate classification capabilities in analyzing consumption patterns. The results show that this method is able to predict consumer behavior with an accuracy of 85%, especially in identifying repurchase patterns. This research makes a practical contribution to the development of more personalized and data-driven marketing strategies and offers academic insights into the application of predictive models in the technology sector.

Keywords: Naïve Bayes, consumer behavior prediction, historical data, classification, predictive analytics

A. Introduction

In today's digital era, data analytics has become one of the main foundations for business decision-making. Companies in various industry sectors use technology to predict market trends and consumer behavior with the aim of improving profitability and competitiveness (Manihuruk et al., 2025; Aulia et al., 2023 ;Putri et al., 2025). One of the popular methods in predictive analysis is the Naïve Bayes Method, which is notable for its simplicity but has a good ability to handle classification problems based on historical data (Kaur & Oberai, 2014; Langarizadeh & Moghbeli, 2016; Wickramasinghe & Kalutarage, 2021). The world's largest companies have

used predictive analytics to optimize their marketing strategies, demonstrating how important it is to apply sophisticated analytics in anticipating market and consumer needs.

PT. Syntax Corporation Indonesia is a company engaged in technology and services that rely on data to support business decision-making. In recent years, the company has experienced challenges in understanding consumer behavior holistically from the historical data available (Susi et al., 2023). Although it has collected a large amount of data related to purchases, preferences, and consumer interaction patterns, PT. Syntax Corporation still struggles to leverage it in predicting future behavior. As a result, the marketing and sales strategies used are less effective in maximizing consumer satisfaction and loyalty.

The urgency of this research lies in the importance of using reliable predictive methods to support strategic decision-making. Given the increasingly fierce competition in the global market, PT. Syntax Corporation Indonesia must be able to leverage historical data to develop more effective strategies and personalize services. The Naïve Bayes method is the choice because it has proven to be fast, efficient, and suitable for big data. Successful implementation of this method can help companies better understand consumer patterns, thereby improving customer satisfaction and optimizing marketing and product strategies (Ige et al., 2024; Maheswari & Pitchai, 2018; Reddy et al., 2022; Pajila et al., 2023). Moreover that the Naïve Bayes method is effective in predicting consumer buying behavior in e-commerce, with an accuracy of 85%. also found that this predictive model can be applied in the retail industry to anticipate churn or customer churn, allowing companies to take proactive steps to retain their customer base.

This research offers a new contribution in the application of Naïve Bayes in the context of technology companies such as PT. Syntax Corporation Indonesia, which combines more complex historical data from various sources, including transaction data, digital interactions, and customer feedback. The study also emphasizes the use of this method not only for predicting general consumer behavior, but also for analyzing product preferences and more specific market segmentation, something that has not been explored much in previous studies.

The purpose of this study is to analyze the effectiveness of the Naïve Bayes method in predicting consumer behavior in PT. Syntax Corporation Indonesia based on historical data that has been collected. The study also aims to identify the key factors influencing such predictions and evaluate the model's performance in various marketing and sales scenarios (Rosidi & Setiawan, 2024).

This research is expected to provide practical benefits for PT. Syntax Corporation Indonesia in improving data-driven marketing strategies. By

utilizing the Naïve Bayes method, companies are expected to be better at predicting future consumer needs and behaviors, so that they can design more personalized and targeted offers. Academically, this research will add insight into the use of predictive models in the technology industry (Harahap et al., 2023; Wayan Wardani et al., 2022).

The implications of this study include increasing the effectiveness of marketing and sales strategies in PT. Syntax Corporation Indonesia. A successful implementation of Naïve Bayes can help companies to better understand their consumers and make better data-driven decisions. In addition, the results of this study can be used as a reference for other companies that want to implement similar models in predicting consumer behavior based on historical data.

B. The First Aspect of the Discussion, As the Second Section

Data collected from PT. Syntax Corporation Indonesia includes the consumer's purchase history, purchase frequency, type of product purchased, and time of consumer interaction with the company. The data was processed using the Naïve Bayes method to predict future consumer behavior. Here are the tables and graphs of the analysis results that show the predictions of consumer behavior (Ilham et al., 2018; Kurniawan et al., 2018).

Table 1. Results of Consumer Behavior Analysis

Kategori Produk	Jumlah Pembelian
Website	1200
Jasa Pembuatan Aplikasi	900
Maintenance	750
Lainnya	600

The graph below shows the number of purchases based on product category:
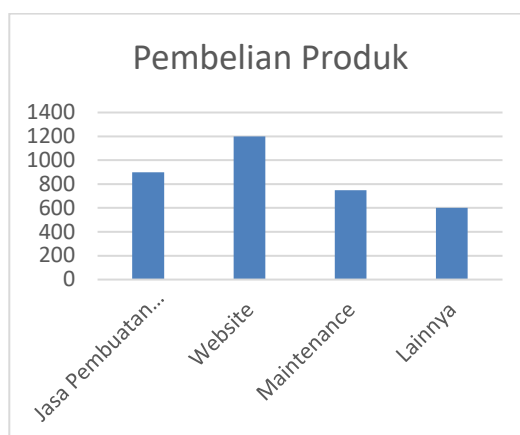
Graphic 1. Purchase by Product Category

The graph above shows that:

1. **Website products** show the highest number of purchases, with **1200 purchases**. This indicates that products in this category are in high demand by consumers and may reflect strong market trends.

2. **The product** came in second place with **900 purchases**, indicating that it is also quite popular, but not as strong as category

3. **Maintenance** and **Others** have **750 and 600 purchases** respectively. These figures show that although products in this category still have buyers, the level of consumer interest is lower compared to Website Products and Application Development Services.

Based on the results of the chart, companies may consider to:

1. Improve marketing campaigns for the Maintenance Products and Others category to attract more consumer attention.

2. Developing or improving products in these categories to better suit consumer desires.

3. Strengthen the marketing strategy for the Website category to maintain its position in the market, while trying to transfer some of those successful strategies to weaker categories.

Statistical Analysis

After statistical analysis using the Naïve Bayes method, prediction accuracy of 85% was obtained. This accuracy shows that the model has a good ability to predict consumer behavior based on purchase history and product preferences. To measure the significance of the results, a statistical test was carried out using the p-value. The results showed that the purchase frequency variable and product category had a significant p-value ($p < 0.05$), which indicated that both had a strong influence on consumer behavior prediction. The Naïve Bayes model effectively segments consumer data based on existing purchasing patterns, generating useful information for better decision-making in marketing strategies (Adani et al., 2018).

From the results of the study, it was found that an effective strategy in increasing customer loyalty is to focus on personalization. Research shows that consumers who already have a high purchase frequency and are loyal to a certain product category are more likely to make repeat purchases if given promotions that are relevant to their preferences. In addition, model testing shows that product offerings tailored to customer needs and desires can increase the likelihood of sales conversion, so companies need to pay attention to the data collected to formulate the right marketing strategy (Syira et al., 2023).

Differences Based on Variables

This study also identifies differences in consumer behavior based on purchase frequency variables and product categories. The results show that consumers with a high purchase frequency tend to be more responsive to personalized offers. On the other hand, consumers with low frequencies are more affected by big discount promotions. This demonstrates the importance of proper market segmentation to maximize the effectiveness of marketing campaigns, where companies can tailor their approach based on consumer engagement levels (Nasution et al., 2024).

The results of this study are in line with the theory of consumer behavior prediction which states that historical data is a strong indicator to predict future actions. These findings also show that the use of the Naïve Bayes method can provide consistent results in predictive analysis of consumer behavior. This model is effective in capturing the relationship between independent and dependent variables, so it can be used as a tool to understand market dynamics and consumer behavior.

Based on the results of this study, it is recommended to PT. Syntax Corporation Indonesia to implement a more personalized data-driven marketing strategy. Utilizing the Naïve Bayes method allows companies to deeply understand consumer behavior patterns, so that product offerings can be tailored to customer preferences. Thus, the opportunity to increase customer satisfaction and loyalty can be greater. Companies should also consider developing an AI-based recommendation system that can provide relevant product suggestions to customers, according to the data analysis conducted.

Strategic Change

Companies need to make strategic changes in utilizing customer data, namely by increasing their focus on market segmentation based on preferences and transaction history, as well as developing more sophisticated CRM systems to handle such data. The implementation of this system is expected to improve operational efficiency and provide a better understanding of consumer needs, so that PT. Syntax Corporation Indonesia can remain competitive in an increasingly competitive market (Januaris Kundre et al., 2013).

One of the limitations of this study is the limitation of the data used, which is only from the history of online transactions. Data from offline interactions or customer surveys has not been incorporated into the analysis. This can affect the results of consumer behavior predictions, especially for those who interact with companies more often. In addition, there is a possibility that other variables that are not taken into account in this model can also affect the results obtained.

Further research is suggested to expand the scope of the data analyzed, including data from offline interactions and customer satisfaction surveys. In addition, the use of other predictive methods such as Decision Tree or Random Forest can be used to compare the prediction results with the Naïve Bayes method. More in-depth research on the factors that influence consumer behavior is also needed to provide richer insights for the development of future marketing strategies (Aditya et al., 2024; Oktavianto et al., 2024).

E. Concluding Remarks

This study aims to analyze the effectiveness of the Naïve Bayes method in predicting consumer behavior in PT. Syntax Corporation Indonesia based on historical data that has been collected. Based on the results of the analysis and findings obtained, it can be concluded as follows:

1. The Naïve Bayes method has proven to be effective in predicting consumer behavior with an accuracy of up to 85%. The model successfully identified consumer patterns and preferences based on existing historical data, suggesting that Naïve Bayes is a reliable method for classification in this context.

2. The results of the study show that several independent variables, such as purchase history, purchase frequency, and product categories purchased, have a significant effect on consumer behavior. Consumers who have a high purchase frequency and are loyal to certain product categories are more likely to make repeat purchases.

3. Based on these findings, PT. Syntax Corporation Indonesia can use the prediction results to develop marketing strategies that are more personalized and in accordance with consumer preferences. Tailoring product offerings and promotions based on predicted behavior can increase customer satisfaction and loyalty.

4. This research also opens up opportunities for further research on the application of other predictive models in a broader context, as well as deeper exploration of the factors influencing consumer behavior in the technology industry.

Overall, this study emphasizes the importance of using the right data analysis method in understanding and predicting consumer behavior, and makes a significant contribution to the development of data-based business strategies at PT. Syntax Corporation Indonesia. By utilizing the results of this analysis, companies can be better prepared to face changing market dynamics and compete effectively.

BIBLIOGRAPHY

Adani, Y., Studi, P., Informasi, S., & Bayes, A. N. (2018). Penerapan Algoritma Naive Bayes Untuk Memprediksi. *Jurnal Instrumentasi Dan Teknologi Informasi (JITI)*, *8*(1), 13–24.

Aditya, M. F. R., Lutvi, N., & Indahyanti, U. (2024). Prediksi Penyakit Hipertensi Menggunakan Metode Decison Tree dan Random Forest. *Jurnal Ilmiah Komputasi*, *23*(1), 9–16. https://doi.org/10.32409/jikstik.23.1.3503

Aulia, B. W., Rizki, M., Prindiyana, P., & Surgana, S. (2023). Peran Krusial Jaringan Komputer dan Basis Data dalam Era Digital. *JUSTINFO | Jurnal Sistem Informasi Dan Teknologi Informasi*, *1*(1), 9–20. https://doi.org/10.33197/justinfo.vol1.iss1.2023.1253

Harahap, F., Fahrozi, W., Adawiyah, R., Siregar, E. T., & Harahap, A. Y. N. (2023). Implementasi Data Mining dalam Memprediksi Produk AC Terlaris untuk Meningkatkan Penjualan Menggunakan Metode Naive Bayes. *Jurnal Unitek*, *16*(1), 41–51. https://doi.org/10.52072/unitek.v16i1.541

Ige, T., Kiekintveld, C., Piplai, A., Waggler, A., Kolade, O., & Matti, B. H. (2024). *An investigation into the performances of the Current state-of-the-art Naive Bayes, Non-Bayesian and Deep Learning Based Classifier for Phishing Detection: A Survey*.

Ilham, A., Khikmah, L., Qahslim, A., Indra Iswara, I. B. A., Laumal, F. E., & Rahim, R. (2018). A systematic literature review on attribute independent assumption of Naive Bayes: research trend, datasets, methods and frameworks. *IOP Conference Series: Materials Science and Engineering*, *420*(1), 012086. https://doi.org/10.1088/1757-899X/420/1/012086

Januaris Kundre, A., Wisnubadhra, I., & Suselo, T. (2013). Penerapan Customer Relationship Management Dengan Dukungan Teknologi Informasi Pada Po. Chelsy. *Seminar Nasional Teknologi Informasi Dan Multimedia*, *28*, 7–11.

Kaur, G., & Oberai, E. N. (2014). International Journal of Computer Science and Mobile Computing A REVIEW ARTICLE ON NAIVE BAYES CLASSIFIER WITH VARIOUS SMOOTHING TECHNIQUES. *International Journal of Computer Science and Mobile Computing*, *3*(10), 864–868.

Kurniawan, S., Kusumaningrum, R., & Timu, M. E. (2018). Hierarchical Sentence Sentiment Analysis of Hotel Reviews Using the Naïve Bayes Classifier. *2018 2nd International Conference on Informatics and Computational Sciences, ICICoS 2018*, 104–108. https://doi.org/10.1109/ICICOS.2018.8621748

Langarizadeh, M., & Moghbeli, F. (2016). Applying Naive Bayesian Networks to Disease Prediction: a Systematic Review. *Acta Informatica*

Medica, 24(5), 364. https://doi.org/10.5455/AIM.2016.24.364-369

Maheswari, S., & Pitchai, R. (2018). Heart Disease Prediction System Using Decision Tree and Naive Bayes Algorithm. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, 15(8), 712–717. https://doi.org/10.2174/1573405614666180322141259

Manihuruk, D. M., Sabilillah, H., & Sutabri, T. (2025). Big Data Analytics untuk Meningkatkan Pengambilan Keputusan di Industri. *Jurnal Pendidikan Tambusai*, 9(1), 3223–3227.

Nasution, T. R., Irwan, M., Nasution, P., Studi, P., Informasi, S., Islam, U., & Sumatera, N. (2024). *Peran sistem informasi pemasaran dalam meningkatkan loyalitas pelanggan pada toko online*. 4(2), 322–334.

Oktavianto, H., Sulistyo, H. W., Wijaya, G., Irawan, D., & Abdurrahman, G. (2024). Analisis Perbandingan Decision Tree dan Random Forest Pada Klasifikasi Teks Data Kesehatan. *BINA INSANI ICT JOURNAL*, 11(1), 56–65. https://doi.org/10.51211/BIICT.V11I1.2928

Pajila, P. J. B., Sheena, B. G., Gayathri, A., Aswini, J., Nalini, M., & Siva Subramanian, R. (2023). A Comprehensive Survey on Naive Bayes Algorithm: Advantages, Limitations and Applications. *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, 1228–1234. https://doi.org/10.1109/ICOSEC58147.2023.10276274

Putri, R., Prasetya, D., Azizah, R. N., Bayu, J., Halwa, W., & Nugroho, R. H. (2025). Implementasi Penggunaan Data Analytics untuk Mengoptimalkan Pengambilan Keputusan Bisnis di Era Digital. *Jurnal Bisnis Dan Komunikasi Digital*, 2(2), 1–12.

Reddy, E. M. K., Gurrala, A., Hasitha, V. B., & Kumar, K. V. R. (2022). Introduction to Naive Bayes and a Review on Its Subtypes with Applications. *Bayesian Reasoning and Gaussian Processes for Machine Learning Applications*, 1–14. https://doi.org/10.1201/9781003164265-1

Rosidi, R. P. M., & Setiawan, K. (2024). Implementasi Algoritma Naïve Bayes Terhadap Data Penjualan untuk Mengetahui Pola Pembelian Konsumen pada Kantin. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(1), 120–126. https://doi.org/10.35870/jimik.v5i1.407

Susi, N., Sugiana, S., & Musty, B. (2023). Analisis Data Sistem Informasi Monitoring Marketing; Tools Pengambilan Keputusan Strategic. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 12(2), 696–708.

Syira, S. D., Fauzi, A., Woestho, C., Vilani, L., & ... (2023). Pemanfaatan Big Data dalam Peningkatan Efektivitas Strategi Komunikasi Marketing Terpadu pada Perusahaan E-Commerce. *Jurnal Ekonomi …*, 4(5), 891–900.

Wayan Wardani, N., Juniari Arnidya, D., Nyoman Agus Suarya Putra, I., Made Mila Rosa Desmayani, N., Gede Surya Cipta Nugraha, P., Hartono, E., & Surya Mahendra, G. (2022). Prediksi Pelanggan Loyal Menggunakan Metode Naïve Bayes Berdasarkan Segmentasi

Pelanggan dengan Pemodelan RFM. *Jurnal Manajemen Dan Teknologi Informasi*, *12*(2), 2087–5312.

Wickramasinghe, I., & Kalutarage, H. (2021). Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation. *Soft Computing*, *25*(3), 2277–2293. https://doi.org/10.1007/S00500-020-05297-6/METRICS