The 2nd International Conference on Multidisciplinary Engineering (2nd ICOMDEN 2024)

The 2nd International Conference on Multidisciplinary Engineering: 00023 (2024) https://doi.org/10.29103/icomden.v2.xxxx
Volume 2, 2024
eISSN: xxxx-xxxx
Research Original Article

Penerapan Metode Naïve Bayes Dalam Mengoptimalkan Kinerja Pemasaran Pada PT. Semen indonesia

Mahesa Anabrang Reglisalo[1] **Dahlan Abdullah**[2] **Yessy Afrillia**[3]
[1]Malikussaleh University, Informatic Engineering,Lhokseumawe, Indonesia,mahesa.190170099@mhs.unimal.ac.id
[2]Malikussaleh University, Informatic Engineering,Lhokseumawe, Indonesia, dahlan@unimal.ac.id
[3]Malikussaleh University, Informatic Engineering,Lhokseumawe, Indonesia, yesy.afrillia@unimal.ac.id
Corresponding Author: mahesa.190170099@mhs.unimal.ac.id

Abstract
This study examines the application of the Naïve Bayes method to improve marketing performance at PT. Semen Indonesia. In an increasingly competitive business environment, effective data management is crucial for strategic decision-making. Currently, PT. Semen Indonesia utilizes the SAP system to manage sales and financial data, but it lacks an automated system to analyze marketing performance. This research aims to develop a Naïve Bayes-based classification system to monitor marketing performance, considering attributes such as profit, market share, sales volume, and customer satisfaction. The Naïve Bayes method was chosen for its accuracy in handling large-scale data and its ability to provide fast and efficient predictions. Marketing performance data is processed using this method to categorize marketing performance as "good" or "poor." The analysis results show that the developed system achieves a classification accuracy of 43.75% for the "good" category and 56.25% for the "poor" category. This system assists management in designing more effective marketing strategies by leveraging historical data to predict trends and market needs.
Keywords: Naïve Bayes, marketing performance, PT. Semen Indonesia, data analysis, classification system, profit, market share

Introduction
In an increasingly competitive era, companies need to employ data-driven strategies to maintain a competitive advantage. Daily data collection has become crucial for modern businesses to obtain up-to-date information, enabling decisions that meet elements of effectiveness, efficiency, and risk mitigation [1]. PT. Semen Indonesia utilizes SAP technology for data management in areas like sales, HR, and finance. However, SAP has not yet been applied for in-depth analysis of sales performance summaries, making it unclear whether sales trends are increasing or decreasing. Thus, a classification system is required to support the analysis process.

The use of strong marketing knowledge has been proven to enhance marketing capabilities. These capabilities are built through a learning process in which knowledge is applied repeatedly to solve marketing issues [2]. PT. Semen Indonesia needs to improve its targeted marketing strategies to achieve optimal marketing performance, both in terms of profit and market share. Naïve Bayes was chosen for its ability to process numerical data, making it easier and more accurate to analyze sales patterns based on attributes like product type, time, and sales volume [3].

Previous research by Ningsih, Windarto, and Chairul Fadlan indicated that the Naïve Bayes Classifier method is effective in optimizing product sales with an accuracy of over 65% on a sales dataset with 120 records. Based on these results, this study focuses on developing a Naïve Bayes-based classification system that is expected to assist PT. Semen Indonesia in managing and optimizing marketing performance efficiently and accurately [4].

Literature Review
A. Data Mining

Data mining is an automated process for extracting useful information from large databases through statistical, mathematical, artificial intelligence, and machine learning methods[5]. The purpose of data mining is to uncover hidden patterns in data that can provide valuable insights for companies[6]. Factors driving the development of data mining include rapid data growth, data access via the internet, and advancements in data mining technology[7].

B. Classification in Data Mining

Classification is a method in data mining used to group objects based on specific characteristics. Algorithms such as

Naïve Bayes, k-nearest neighbor, and decision trees are commonly used techniques for classification[8]. Classification is essential because it helps predict the category or class of an object based on available data.

C. Naïve Bayes Algorithm

Naïve Bayes is a simple yet effective probability-based classification algorithm. This algorithm predicts probabilities based on past experience, known as Bayes' Theorem[9] . This method is fast, simple, and suitable for large data sets as it requires minimal training data to accurately estimate parameters[10]. The main formula of Bayes' Theorem enables probability calculation by assuming independence among variables[3].

D. Marketing Performance

Marketing performance is a measure of the success of a company's marketing activities, including sales volume, customer count, profit, and sales growth [2]. Market-oriented companies generally have better marketing performance, reflected in increased sales and market share[11]. Indicators of marketing performance include sales growth, customer growth, and profit growth.

E. Previous Research

Previous studies on Naïve Bayes demonstrate the effectiveness of this method in various classification applications. For example, Handoko & Neneng used Naïve Bayes to diagnose pregnancy-related diseases with 77% accuracy[12], while Wijaya & Dwiasnati, implemented it in drug sales with an accuracy level of 88% [13]. Other research, such as Fadlan et al. [4], successfully classified eligibility for Rastra rice recipients, and Nurdiawan & Salim, applied this method to optimize marketing strategy with an optimization level of over 65%.

Materials & Methods

the methodology used in this research, conducted at PT. Semen Indonesia in July 2023 to facilitate data collection for system development. Key steps included collecting data directly from the company, conducting a literature review on Naïve Bayes, analyzing the collected data, and designing a MySQL database to store marketing performance data. The system was developed using PHP and MySQL, with an interface enabling users to classify data based on attributes like month, year, market share, customer satisfaction, sales volume, production, and profit. This Naïve Bayes-based system classified marketing performance as "good" or "poor," providing a structured approach for optimizing the company's marketing strategy.

A. System Application Design Scheme

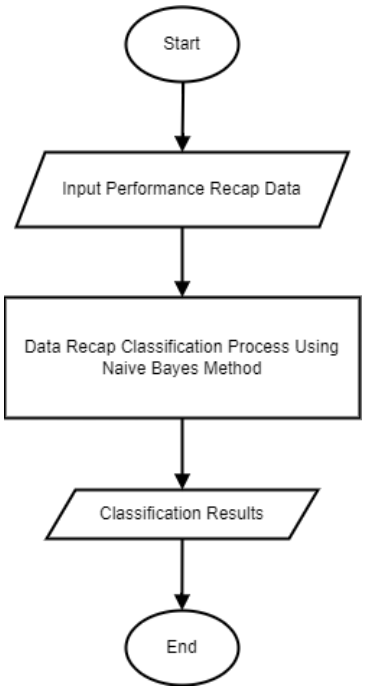

Figure 1. System Aplication Scheme

Here is the translation of the system flow process:

1. The system process starts.
2. The user inputs marketing performance data, with attributes such as profit, market share, sales volume, and customer satisfaction.
3. Next, the user clicks on the classification process, and the Naive Bayes algorithm calculation is carried out by the system.
4. After that, the classification results are displayed, showing whether the performance is optimal or not, depending on the data provided.

Results and Discussion

This study addresses the challenge of optimizing marketing performance by employing the Naïve Bayes algorithm at PT. Semen Indonesia. By systematically categorizing and analyzing historical data, this research demonstrates the viability of Naïve Bayes as a tool for predictive marketing classification, identifying key indicators associated with high and low performance. The results of this research are organized into several key components: data attribute analysis, Naïve Bayes implementation, manual calculation illustrations, calculation results, comprehensive analysis of findings, and an overarching conclusion.

1. Data Attributes and Target Classification

The dataset used in this study includes monthly records from January 2020 to December 2023. Each data point represents a snapshot of marketing metrics, which collectively influence PT. Semen Indonesia's market positioning and customer satisfaction. The attributes analyzed in this study are as follows:

a. Market Share: Categorized into five ranges: 1-20%, 21-40%, 41-60%, 61-80%, and 81-100%. Each range provides insights into customer engagement and market penetration.
   · Customer Satisfaction: Categorical, with binary values of "Satisfied" and "Not Satisfied." This attribute is integral for understanding consumer response to the company's offerings.
b. Sales Volume: Segmented into ranges: <30 million, 30-40 million, 40-50 million, 50-60 million, and >60 million tons. This metric captures overall product demand in various markets.
c. Production Volume: Like sales volume, this attribute reflects the operational output segmented across multiple production ranges.
d. Profit: Expressed in Indonesian Rupiah (IDR), profit values provide a financial metric tied directly to market success.
e. Performance Classification: The target variable classifies marketing performance as either "Good" or "Poor," establishing the basis for the Naïve Bayes classification.

This categorical and continuous data provides a foundation for probabilistic classification, allowing the Naïve Bayes algorithm to evaluate the likelihood of each class (Good or Poor) based on observed attribute patterns.

2. Naïve Bayes Implementation Process

The Naïve Bayes method was chosen due to its efficacy in probabilistic classification, particularly with datasets that assume independence among attributes. The implementation involved multiple steps:

a. Data Preprocessing: Data preparation involved categorizing continuous variables into discrete ranges to meet Naïve Bayes requirements. For instance, sales and production volumes were grouped into predefined ranges, facilitating probabilistic calculation.
b. Calculation of Prior Probabilities: Using historical data, prior probabilities were computed for each target class. This step established a foundational likelihood for both "Good" and "Poor" outcomes:
   P(Good)=21/48=0.4375
   P(Poor)=27/48=0.5625

These values indicate that, historically, "Poor" performance has been slightly more prevalent than "Good" performance within the dataset

3. Conditional Probability Calculations

For each attribute, conditional probabilities were calculated within each class. These conditional probabilities enable the model to estimate the likelihood of observing each attribute value given a particular classification (Good or Poor).

4. Model Testing and Validation

After calculating the probabilities, the model was tested on the dataset, validating its classification accuracy and ability to predict new entries effectively.

An example of the Naïve Bayes process was undertaken to illustrate the classification mechanics. Here's a breakdown of the calculation:

Prior Probability Calculation

For "Good" performance: $P(\ Good\ ) = \frac{21}{48} = 0.4375$

For "Poor" performance: $P(\ Poor\ ) = \frac{27}{48} = 0.5625$

Step 2: Conditional Probability for Market Share (e.g., Range 1-20%)

Probability within "Good" classification: $P(\ MarketShare = 1-20\% \mid Good\ ) = \frac{3}{21} = 0.1429$

Probability within "Poor" classification: $P(\ MarketShare = 1-20\% \mid Poor\ ) = \frac{2}{27} = 0.0741$

Step 3: Other Attribute Conditional Probabilities (e.g., Customer Satisfaction)

Probability of "Satisfied" within "Good": $P(\ Satisfied \mid Good\ ) = \frac{14}{21} = 0.6667$

Probability of "Not Satisfied" within "Poor": $P(\ Not\ Satisfied \mid Poor\ ) = \frac{10}{27} = 0.3704$

The above calculations are repeated for all attributes and values to build a complete conditional probability table. This step-by-step calculation process helps to illustrate how the algorithm assesses each attribute's impact on classification.

A. Results of Naïve Bayes Classification

The calculation results provided meaningful insights into PT. Semen Indonesia's marketing performance. Some highlights from the calculated probabilities are as follows:

a. Market Share: Higher market share ranges, particularly 61-100%, were associated with greater likelihoods of "Good" performance outcomes. For example, the range 81-100% has probabilities of 0.2857 for "Good" and 0.2963 for "Poor," reflecting its critical role in determining positive outcomes.
b. Customer Satisfaction: This attribute showed a strong correlation with "Good" performance. Satisfied customers were more commonly found in the "Good" category, with a probability of 0.6667. Conversely, dissatisfaction among customers was more prevalent in the "Poor" class, with a probability of 0.3704.
c. Sales and Production Volumes: While less influential than customer satisfaction and market share, sales and production volumes also contributed to the classification process. For instance, a sales volume of 30-40 million tons had a probability of 0.0476 for "Good" and 0.2593 for "Poor."

These results reflect the effectiveness of Naïve Bayes in segmenting marketing data and indicate the algorithm's ability to predict performance outcomes based on specific marketing attributes.

B. Analysis of Classification Results

The analysis of the results revealed several key findings regarding attribute impact and potential marketing strategies:

a. Key Drivers of Marketing Success: The probabilities indicate that market share and customer satisfaction are primary drivers of marketing performance. Higher market shares, combined with positive customer satisfaction, were consistently associated with the "Good" class. Therefore, focusing on these two areas could yield more favorable outcomes for PT. Semen Indonesia.
b. Influence of Sales and Production Metrics: While secondary to customer satisfaction, sales and production volumes also played a notable role in performance classification. For example, production volumes in the range of 50-60 million tons were more likely associated with "Poor" outcomes, suggesting the need for optimized production efficiency.
c. Strategic Implications: The findings highlight areas of strategic focus for PT. Semen Indonesia, particularly improving market share and customer satisfaction. The data suggests that maintaining high satisfaction rates and expanding market share can substantially enhance the likelihood of achieving "Good" performance outcomes.

C. Probability Tables and Visual Representations

Table 1. Probability table

Attribute	Value	Probability (Good)	Probability (Poor)
Market Share	$1-20\%$	0.1429	0.0741
Market Share	$81-100\%$	0.2857	0.2963
Customer Satisfaction	Satisfied	0.6667	0.6296
Customer Satisfaction	Not Satisfied	0.3333	0.3704
Sales Volume	$30-40$ million	0.0476	0.2593
Production Volume	$50-60$ million	0.1429	0.2593

Interpretation: This table provides a breakdown of the likelihood for each attribute range within the "Good" and "Poor" classes. For instance, market shares between 1-20% show limited association with "Good" outcomes, whereas market shares above 60% are more indicative of positive results.

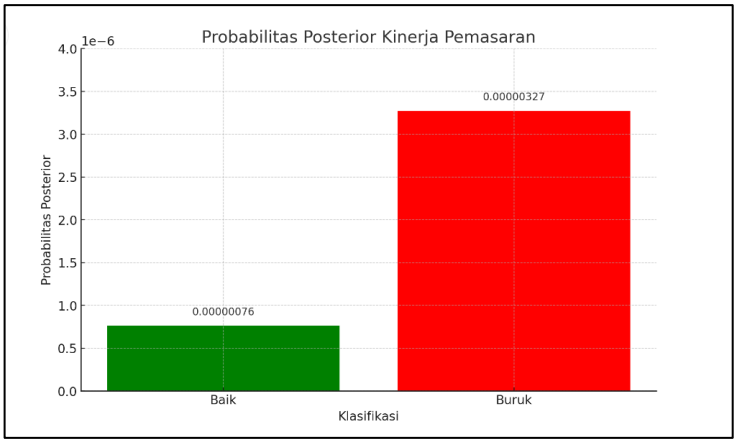

Figure 2. Probalitas

A bar chart visualizes the probability of "Good" and "Poor" outcomes across various market share ranges, emphasizing that higher market shares align with improved outcomes.

D. Conclusion of Research Findings

The implementation of Naïve Bayes has demonstrated its potential as a powerful tool for PT. Semen Indonesia's marketing optimization. The classification results reveal clear patterns that associate higher market share and customer satisfaction with improved marketing performance. provide a predictive framework for marketing performance, assisting PT. Semen Indonesia in focusing on areas of strategic importance to achieve consistent success. This approach highlights the viability of data-driven strategies to support decision-making in competitive markets.

Conclusions

This study concludes that the Naïve Bayes algorithm is an effective tool for classifying and predicting marketing performance at PT. Semen Indonesia. By analyzing historical data on key attributes such as market share, customer satisfaction, sales volume, and production levels, the algorithm reveals that higher market share and positive customer satisfaction are the most significant indicators of "Good" marketing performance. Sales and production volumes also contribute to classification outcomes, with certain ranges correlating more with "Poor" performance, highlighting areas for potential operational improvements.

The results suggest practical strategies for PT. Semen Indonesia, emphasizing the importance of expanding market share and improving customer satisfaction to increase the likelihood of favorable performance. The success of the Naïve Bayes approach in this context underscores its applicability in predictive analytics for marketing, providing a scalable and data-driven framework for strategic decision-making. This study lays the groundwork for future research to further enhance performance predictions, potentially by incorporating additional attributes or exploring more advanced machine learning methods.

References

[1] D. Dahri, F. Agus, and D. M. Khairina, "Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas Mulawarman," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 11, no. 2, p. 29, 2016, doi: 10.30872/jim.v11i2.211.

[2] A. Utaminingsih, "Pengaruh orientasi pasar, inovasi, dan kreativitas strategi pemasaran terhadap kinerja pemasaran pada UKM kerajinan rotan di Desa Teluk Wetan, Welahan, Jepara," *Media Ekon. Dan Manaj.*, vol. 31, no. 2, pp. 77–87, 2016.

[3] *et al.*, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 1, pp. 177–188, 2023, doi: 10.29408/jit.v6i1.7529.

[4] C. Fadlan, S. Ningsih, and A. P. Windarto, "Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra," *J. Tek. Inform. Musirawas*, vol. 3, no. 1, p. 1, 2018, doi: 10.32767/jutim.v3i1.286.

[5] R. Nurul Arifin, "Implementasi Algoritma Frequent Pattern Growth (FP-GROWTH) Menentukan Asosiasi Antar Produk (Study KAsus Nadia Mart)," *Dok. Karya Ilm.*, pp. 0–1, 2015.

[6] B. N. , T. R. , Syamsul Bahri, "Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak)," *Coding J. Komput. dan Apl.*, vol. 4, no. 3, 2016, doi: 10.26418/coding.v4i3.17034.

[7] Styawati, Nurkholis Andi, and Anjumi Krisma Nur, "Analisis Pola Transaksi Pelanggan Menggunakan Algoritme Apriori," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 619–626, 2021.

[8] C. Darujati *et al.*, "Pemanfaatan Teknik Supervised Untuk," *Text*, vol. 16, no. 1, pp. 1–8, 2012.

[9] M. Furqan, M. K. Sastia H Wibowo, S. Kom, and E. Dedy Abdullah, ST. M, "Sistem Persediaan Obat Pada Puskesmas Menggunakan Metode Naive Bayes (Studi Kasus Puskesmas Talang Tinggi Seluma) Skripsi. Universits Muhammadiyah Bengkulu," *Univ. Muhamadiyah Bengkulu*, vol. 18, no. 60, pp. 1–8, 2016.

[10] E. Darnila, M. Maryana, and M. Azmi, "Aplikasi Klasifikasi Status Gizi Balita Menggunakan Metode Naïve Bayes Berbasis Android," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 5, no. 2, pp. 135–141, 2021, doi: 10.46880/jmika.vol5no2.pp135-141.

[11] A. Lutfi and Firmansyah, "Peningkatan Kinerja Pemasaran Umkm Berdasarkan Orientasi Kewirausahaan, Pengetahuan Pemasaran Dan Keadaan Lingkungan," *J. Bisnis, Manajemen, Dan Inform.*, vol. 18, no. II, pp. 86–102, 2021, doi: 10.26487/jbmi.v18i2.13977.

[12] M. Ridho Handoko, "Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 50–58, 2021, [Online]. Available: http://jim.teknokrat.ac.id/index.php/JTSI

[13] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.