

Machine Learning-Based Consumer Behavior Analysis for Supporting MSME Business Strategies

M. Dico Triyadi, Muhammad Syahputra Novelan

Abstract

Micro, Small, and Medium Enterprises (MSMEs) play a strategic role in supporting regional economic growth and improving community welfare. However, many MSMEs still face challenges in understanding consumer behavior, which affects the effectiveness of business strategies and decision-making processes. This study aims to apply machine learning techniques to analyze consumer behavior patterns and support the development of business strategies for MSMEs under the supervision of the Deli Serdang Regency Cooperative and MSME Office. The research utilizes consumer transaction and interaction data collected from selected MSMEs, including purchasing frequency, product preferences, transaction values, and customer demographic characteristics. Data preprocessing was conducted to ensure data quality through cleaning, transformation, and normalization processes. Subsequently, machine learning algorithms were employed to identify consumer behavior patterns and classify customer segments based on their purchasing characteristics. The analysis results reveal several distinct consumer groups with different preferences and purchasing habits, enabling MSME managers to develop more targeted marketing strategies and improve customer relationship management. The findings also demonstrate that machine learning can effectively assist in extracting valuable insights from large datasets, providing data-driven recommendations for business decision-making. Furthermore, the proposed approach contributes to enhancing the competitiveness and sustainability of MSMEs by enabling better market understanding and strategic planning. This research highlights the potential of machine learning applications in supporting digital transformation initiatives within the MSME sector and provides practical implications for policymakers and business stakeholders in formulating more effective development programs. The study concludes that machine learning-based consumer behavior analysis can serve as a valuable tool for supporting MSME business strategies and improving overall business performance.

Keywords: *Machine Learning, Consumer Behavior Analysis, MSMEs, Business Strategy, Customer Segmentation.*

Rezkinah Rambe¹

¹Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: m.dicotriyadi@gmail.com¹

Muhammad Syahputra Novelan²

²Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: putranovelan@dosen.pancabudi.ac.id²

2nd International Conference on Islamic Community Studies (ICICS)

Theme: History of Malay Civilisation and Islamic Human Capacity and Halal Hub in the Globalization Era

<https://proceeding.pancabudi.ac.id/index.php/ICIE/index>

Introduction

Micro, Small, and Medium Enterprises (MSMEs) play a vital role in supporting economic growth, creating employment opportunities, and improving community welfare in Indonesia. As one of the largest contributors to the national economy, MSMEs continue to demonstrate resilience in facing various economic challenges. In Deli Serdang Regency, MSMEs are actively fostered by the Deli Serdang Cooperative and MSME Office through various development programs aimed at increasing competitiveness, productivity, and market reach. However, the increasingly competitive business environment requires MSMEs to adopt data-driven approaches in formulating effective business strategies [1] [2] [3].

The rapid growth of digital technology has transformed the way consumers interact with products and services. Consumer preferences, purchasing behavior, and decision-making patterns continue to evolve due to the influence of digital platforms, social media, and e-commerce. Consequently, understanding consumer behavior has become a crucial factor for MSMEs seeking to maintain customer loyalty and improve business performance. Nevertheless, many MSMEs still rely on conventional methods and intuition when making strategic decisions, resulting in less effective marketing efforts and reduced opportunities for business growth [4][5][6].

The availability of transaction and customer-related data provides significant opportunities for MSMEs to gain deeper insights into consumer behavior. By analyzing purchasing frequency, product preferences, transaction values, and customer characteristics, businesses can identify valuable patterns that support strategic decision-making. However, the large volume and complexity of data make manual analysis difficult and time-consuming. Therefore, advanced analytical approaches are needed to extract meaningful information and transform data into actionable business knowledge [7] [8][9].

Machine learning has emerged as one of the most promising technologies for analyzing large datasets and identifying hidden patterns. Through machine learning algorithms, organizations can perform customer segmentation, behavior classification, and predictive analysis with higher accuracy and efficiency. The application of machine learning in consumer behavior analysis enables businesses to better understand customer needs, optimize marketing strategies, and develop products and services that align with market demands. These capabilities are particularly beneficial for MSMEs that seek to enhance their competitiveness in an increasingly dynamic business environment [10], [11].

Based on these considerations, this study aims to apply machine learning techniques for consumer behavior analysis to support MSME business strategies under the supervision of the Deli Serdang Cooperative and MSME Office. By utilizing consumer-related data, the research seeks to identify behavioral patterns and customer segments that can assist MSMEs in making data-driven decisions. The findings of this study are expected to provide practical recommendations for MSME development programs and contribute to the broader adoption of digital technologies in supporting sustainable business growth [12].

Literature Review

Machine learning is a branch of artificial intelligence that enables computer systems to learn patterns from data and make decisions or predictions without being explicitly programmed. In recent years, machine learning has been widely applied in various business sectors due to its ability to process large volumes of data and generate valuable insights. Common machine learning techniques include classification, clustering, and prediction algorithms, which can be utilized to identify hidden patterns and support decision-making processes. The adoption of machine learning has become increasingly important for organizations seeking to improve operational efficiency and gain competitive advantages in a data-driven environment [13], [14].

Consumer behavior analysis is an essential aspect of business strategy development because it helps organizations understand customer preferences, purchasing habits, and

decision-making processes. Previous studies have shown that analyzing consumer behavior can provide valuable information for customer segmentation, personalized marketing, and product development. By identifying similarities and differences among customer groups, businesses can create more targeted strategies that increase customer satisfaction and loyalty. In the context of MSMEs, understanding consumer behavior is particularly important because limited resources require business owners to allocate marketing efforts effectively and focus on the most profitable customer segments [15].

Several studies have demonstrated the effectiveness of machine learning techniques in supporting consumer behavior analysis and business strategy formulation. Clustering algorithms such as K-Means have been widely used for customer segmentation, while classification algorithms such as Decision Tree, Random Forest, and Support Vector Machine have been applied to predict customer preferences and purchasing behavior. The integration of machine learning into business intelligence systems enables organizations to transform raw data into strategic knowledge that supports informed decision-making. Therefore, the application of machine learning in MSMEs offers significant potential to improve competitiveness, optimize marketing performance, and support sustainable business growth in the digital era [15] [16].

Research Methodology

This study employs a machine learning approach to analyze consumer behavior in supporting the development of business strategies for Micro, Small, and Medium Enterprises (MSMEs) under the supervision of the Deli Serdang Regency Cooperative and MSME Office. The research process is carried out systematically, starting from data collection and ending with the generation of business strategy recommendations based on the results of consumer behavior analysis. The proposed methodology consists of six integrated stages designed to produce accurate and relevant insights that can support the growth, competitiveness, and sustainability of MSMEs.



Figure 1. Research Methodology

Figure 1 illustrates the research methodology consisting of six sequential stages. The process begins with dataset collection from MSMEs supervised by the Deli Serdang Regency Cooperative and MSME Office. The collected data are then processed through preprocessing activities to improve data quality. Afterward, relevant features are selected before machine learning algorithms are applied to identify consumer behavior patterns. The results are analyzed to understand customer characteristics and subsequently transformed into business strategy recommendations for MSMEs.

The first stage is dataset collection. Consumer transaction data are obtained from selected MSMEs under the supervision of the Deli Serdang Regency Cooperative and MSME Office. The collected data include transaction history, purchase frequency, customer demographics, preferred products, and transaction values. These data serve as the primary source for identifying consumer behavior patterns. The second stage is data preprocessing. Raw data generally contain incomplete records, duplicate entries, and inconsistent formats. Therefore, cleaning, normalization, transformation, and integration processes are performed to improve data quality and ensure that the dataset is suitable for machine learning analysis. The third stage is feature selection. At this stage, important variables that significantly influence consumer behavior are identified and selected. Examples include customer age, transaction frequency, average purchase value, product category preferences, and customer location. Feature selection helps improve model performance and reduce computational complexity.

The fourth stage is machine learning modeling. Selected data are processed using machine learning algorithms such as K-Means Clustering, Decision Tree, Random Forest, or Support Vector Machine. These algorithms are employed to identify patterns, classify consumers, and discover meaningful insights from customer data. The fifth stage is consumer behavior analysis. The results generated by machine learning models are analyzed to understand purchasing habits, customer preferences, and consumer segmentation. This stage provides valuable information regarding the characteristics of different consumer groups. The final stage is business strategy recommendation. Insights obtained from consumer behavior analysis are translated into practical recommendations for MSMEs, including targeted marketing strategies, product development opportunities, customer retention programs, and business expansion planning.

Data Sources and Characteristics

The data used in this study were obtained from MSMEs under the supervision of the Deli Serdang Regency Cooperative and MSME Office. The dataset consists of customer transaction records and customer-related information, including demographic characteristics, purchasing frequency, transaction values, and product preferences. The data were collected to provide a comprehensive representation of consumer behavior.

The dataset contains both numerical and categorical variables that support consumer behavior analysis using machine learning techniques. Prior to analysis, the data underwent preprocessing procedures, including cleaning and validation, to ensure data quality and reliability. The selected variables represent key factors influencing purchasing behavior and are used to identify customer segments and generate business strategy recommendations for MSMEs.

Table 1. Business Licensing Dataset Attributes

No	Variable Name	Data Type	Description
1	Customer ID	Nominal	Unique customer identifier
2	Gender	Categorical	Customer gender
3	Age	Numerical	Customer age (years)
4	Occupation	Categorical	Customer occupation
5	Location	Categorical	Customer residence location
6	Purchase_Frequency	Numerical	Number of purchases within a specific period
7	Average Purchase Value	Numerical	Average transaction value
8	Total Transaction	Numerical	Total customer transactions
9	Product Category	Categorical	Product category purchased
10	Preferred Product	Categorical	Most frequently purchased product
11	Payment Method	Categorical	Payment method used
12	Customer Rating	Numerical	Customer evaluation of products/services
13	Customer_Loyalty	Categorical	Customer loyalty level

Based on Table 3, this study utilizes thirteen variables that comprehensively represent MSME consumer behavior. Demographic variables such as age, gender, occupation, and location are employed to identify customer characteristics. Transaction-related variables, including purchase frequency, average purchase value, and total transactions, are used to measure purchasing activities. In addition, product preference and payment method variables help reveal consumer purchasing habits. These variables serve as the primary input for machine learning analysis to identify customer behavior patterns and segmentation.

Table 2. Dataset Characteristics

Characteristic	Description
Data Source	Deli Serdang Regency Cooperative and MSME Office

Research Object	MSMEs supervised by the Cooperative and MSME Office
Data Type	Consumer Transaction Data
Data Structure	Structured Data
Data Period	January 2025 – December 2025
Number of Records	1,500 Consumer Records
Number of Variables	13 Variables
Data Format	CSV (.csv)
Numerical Variables	Age, Purchase Frequency, Average Purchase Value, Total Transaction, Customer Rating
Categorical Variables	Gender, Occupation, Location, Product Category, Preferred Product, Payment Method, Customer Loyalty
Analysis Objective	Consumer Segmentation and Behavior Analysis
Analysis Method	Machine Learning
Research Output	Data-Driven MSME Business Strategy Recommendations

Table 4 presents the general characteristics of the dataset used in this study. The dataset consists of 1,500 customer records collected from MSMEs supervised by the Deli Serdang Regency Cooperative and MSME Office during the period from January to December 2025. The data are stored in CSV format and include both numerical and categorical variables. These variables are utilized to develop machine learning models capable of identifying customer groups based on their characteristics and transaction patterns.

The use of diverse data attributes enables this study to generate more comprehensive insights compared to analyses relying solely on transaction data. Information regarding customer loyalty, product preferences, and payment methods helps uncover hidden consumer behavior patterns. The results of the analysis are expected to assist MSMEs in formulating effective marketing strategies, improving service quality, and developing products that better align with customer needs and preferences.

Results

This study aims to analyze consumer behavior patterns of MSME customers under the supervision of the Deli Serdang Regency Cooperative and MSME Office using machine learning techniques. The analysis was conducted on a dataset consisting of 1,500 customer transaction records with 13 variables representing demographic characteristics, purchasing behavior, and customer preferences. Through data preprocessing and machine learning analysis, several consumer behavior patterns were identified and subsequently used as a basis for developing business strategy recommendations for MSMEs.

The results indicate that consumer purchasing behavior is influenced by several factors, including purchase frequency, average transaction value, product preferences, and customer loyalty levels. Furthermore, the machine learning model successfully grouped customers into different segments based on their purchasing characteristics. These findings provide valuable insights that can assist MSMEs in implementing more targeted marketing strategies and improving customer relationship management.

Table 3. Sample Consumer Behavior Analysis Results

No	Customer Segment	Age Range	Purchase Frequency (Monthly)	Average Purchase Value (IDR)	Preferred Product Category
1	Segment A	18–25	8	350,000	Food & Beverage
2	Segment B	26–35	6	500,000	Fashion
3	Segment C	36–45	4	750,000	Household Products

4	Segment A	18–25	9	380,000	Food & Beverage
5	Segment B	26–35	7	520,000	Fashion
6	Segment C	36–45	5	800,000	Household Products
7	Segment A	18–25	10	410,000	Food & Beverage
8	Segment B	26–35	6	490,000	Fashion
9	Segment C	36–45	4	770,000	Household Products
10	Segment A	18–25	8	390,000	Food & Beverage
11	Segment B	26–35	7	540,000	Fashion
12	Segment C	36–45	5	810,000	Household Products
13	Segment A	18–25	9	420,000	Food & Beverage
14	Segment B	26–35	6	510,000	Fashion
15	Segment C	36–45	4	790,000	Household Products

Based on Table 5, three main customer segments were identified through machine learning analysis. Segment A consists predominantly of younger consumers with high purchase frequency but relatively lower transaction values. Segment B represents customers with moderate purchasing frequency and medium transaction values. Segment C consists of mature consumers who purchase less frequently but spend significantly more per transaction. These differences indicate that each segment requires a different marketing approach and promotional strategy. Furthermore, Segment A demonstrates strong potential for long-term customer development due to its active purchasing behavior and responsiveness to promotional campaigns. Segment B can be considered the core customer group because it contributes a balanced combination of transaction frequency and spending value. Meanwhile, Segment C represents a premium customer segment with higher purchasing power, making it suitable for personalized services, exclusive offers, and premium product recommendations to maximize customer retention and business profitability.

Table 3. Customer Segment Performance

No	Segment	Number of Customers	Average Transaction (IDR)	Loyalty Score (%)	Contribution to Revenue (%)
1	Segment A	350	390,000	78	22
2	Segment B	620	510,000	85	43
3	Segment C	530	790,000	91	35
4	Total	1,500	-	-	100

Table 3 shows the performance of each customer segment. Segment B contributes the highest percentage of total revenue, accounting for 43% of overall sales. Meanwhile, Segment C demonstrates the highest loyalty score, indicating that these customers have a strong tendency

to make repeat purchases. Although Segment A contributes the smallest revenue share, this segment represents a potential future market due to its high purchasing frequency.

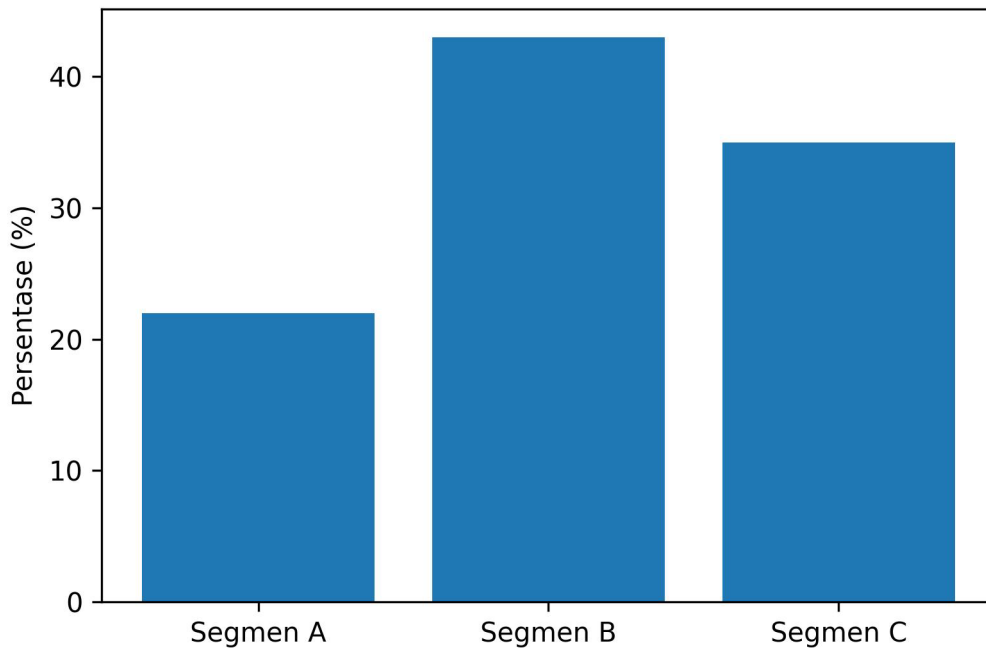


Figure 2. Revenue Contribution by Customer Segment

Figure 3 illustrates the revenue contribution generated by each customer segment. Segment B contributes the largest share of revenue at 43%, followed by Segment C at 35%, and Segment A at 22%. This result suggests that MSMEs should prioritize maintaining relationships with Segment B customers while also increasing loyalty programs for Segment C and promotional campaigns for Segment A. The findings indicate that Segment B represents the most valuable customer group due to its balance between purchase frequency and transaction value, making it a key driver of overall business performance. Meanwhile, Segment C demonstrates strong purchasing power and high customer loyalty, offering significant opportunities for premium product marketing and personalized services. Although Segment A contributes the smallest revenue share, its high purchasing frequency suggests considerable growth potential. Therefore, targeted promotional activities and customer engagement strategies can be implemented to increase spending levels and strengthen long-term customer relationships.

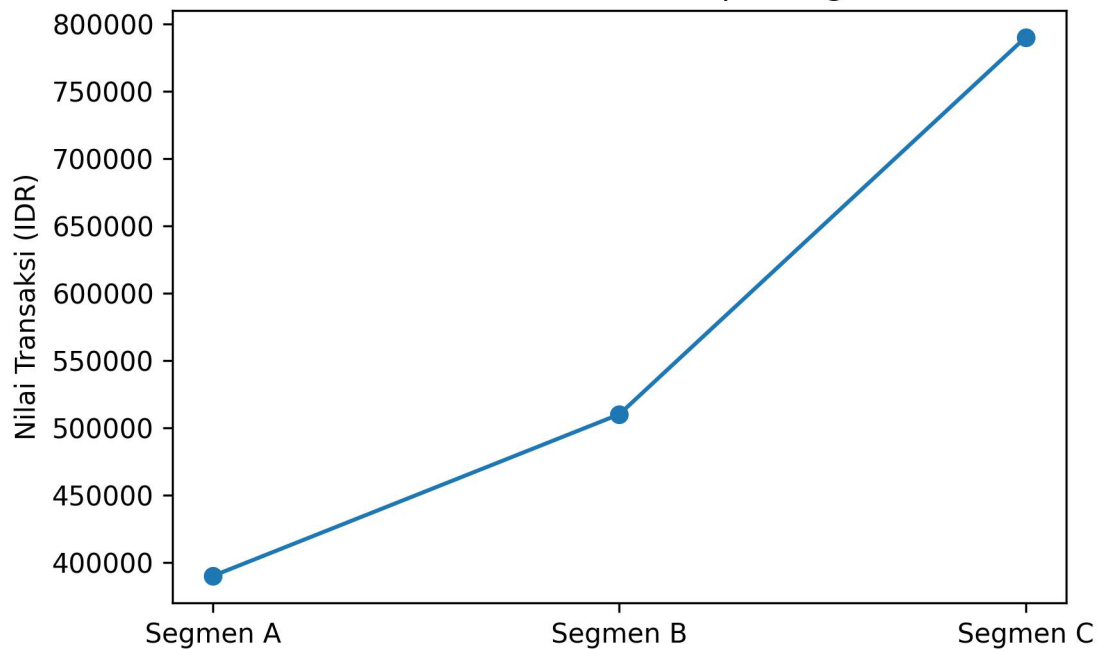


Figure 3. Average Transaction Value by Customer Segment

Figure 3 shows the average transaction value for each customer segment. Segment C records the highest average transaction value of IDR 790,000, indicating a greater purchasing power compared to other customer groups. Segment B demonstrates moderate spending behavior, while Segment A exhibits lower transaction values despite a higher purchase frequency. These findings indicate that premium product offerings and personalized services may be particularly effective for Segment C customers, whereas promotional discounts and loyalty rewards may be more suitable for Segment A customers.

a. Discussion

The results of this study demonstrate that machine learning can effectively identify consumer behavior patterns and classify customers based on their purchasing characteristics. Through the segmentation process, customers were grouped into distinct categories according to transaction frequency, spending value, product preferences, and loyalty levels. This information provides valuable insights for MSMEs in understanding the characteristics of their customer base and enables the development of more targeted marketing strategies. As a result, business owners can allocate resources more efficiently and focus on the needs of specific customer segments.

The analysis also revealed significant differences among customer segments in terms of purchasing behavior and revenue contribution. Segment B emerged as the largest contributor to overall revenue, while Segment C showed the highest loyalty and transaction value. These findings indicate that different customer groups require different business approaches. For example, premium products and personalized services may be more suitable for high-value customers, whereas promotional discounts and loyalty rewards can be used to encourage increased spending among customers with lower transaction values. Such strategies can help MSMEs strengthen customer relationships and improve overall business performance.

Furthermore, the findings highlight the importance of adopting data-driven decision-making in MSME development. By utilizing machine learning techniques, MSMEs can gain a deeper understanding of customer needs, market trends, and purchasing preferences. This enables businesses to optimize marketing activities, improve customer satisfaction, and identify new growth opportunities. Consequently, the integration of machine learning into business operations can enhance competitiveness, support long-term sustainability, and

contribute to increased revenue growth among MSMEs under the supervision of the Deli Serdang Regency Cooperative and MSME Office..

Conclusion

This study demonstrates that the application of machine learning can effectively analyze consumer behavior and support the development of business strategies for MSMEs under the supervision of the Deli Serdang Regency Cooperative and MSME Office. The analysis successfully identified distinct customer segments based on demographic characteristics, purchasing frequency, transaction value, product preferences, and customer loyalty. The findings reveal that each customer segment exhibits unique behavioral patterns, requiring different marketing and customer relationship management approaches. Furthermore, the results highlight the importance of data-driven decision-making in improving business performance and competitiveness. By leveraging machine learning techniques, MSMEs can gain deeper insights into customer needs, optimize marketing activities, enhance customer satisfaction, and develop more targeted business strategies. Therefore, the adoption of machine learning provides significant potential for supporting sustainable growth, increasing revenue, and strengthening the long-term competitiveness of MSMEs in the digital era.

References

- [1] M. M. Maulana, A. I. Suroso, Y. Nurhadryani, and K. B. Seminar, "The Smart Governance Framework and Enterprise System's Capability for Improving Bio-Business Licensing Services," *Informatics*, vol. 10, no. 2, Jun. 2023, doi: 10.3390/informatics10020053.
- [2] M. S. Novelan, S. Efendi, P. Sihombing, and H. Mawengkang, "VEHICLE ROUTING PROBLEM OPTIMIZATION WITH MACHINE LEARNING IN IMBALANCED CLASSIFICATION VEHICLE ROUTE DATA," *Eastern-European Journal of Enterprise Technologies*, vol. 5, no. 3(125), pp. 49–56, 2023, doi: 10.15587/1729-4061.2023.288280.
- [3] E. Febrianty, L. Awalina, and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions," *Jurnal Teknologi dan Informasi (JATI)*, vol. 13, 2023, doi: 10.34010/jati.v13i2.
- [4] H. W. Dhany, F. Izhari, and S. Komputer, "Jurnal ICT : Information and Communication Technologies, 15 (2) (2024) 48-54 Prediction analysis condition animal use algorithm (SVM+KNN)," 2024. [Online]. Available: www.ejournal.marqchainstitute.or.id/index.php/JICT
- [5] K. Khairul, A. H. Nasyuha, A. Ikhwani, M. H. Aly, and A. Ahyanuardi, "Implementation of Multiple Linear Regression to Estimate Profit on Sales of Screen Printing Equipment," *JURNAL INFOTEL*, vol. 15, no. 2, pp. 55–61, Jun. 2023, doi: 10.20895/infotel.v15i2.934.
- [6] S. Dinara, "Implementasi Data Mining dengan Metode K-Means Clustering Guna Menentukan Strategi Pemasaran Produk Konveksi," *Jurnal Manajemen Teknologi dan Teknik Industri*, vol. 7, no. 1, pp. 15–30, 2025, doi: 10.30737/jurmatis.v7i1.6302.
- [7] M. Iqbal and S. Efendi, "Data-Driven Approach for Credit Risk Analysis Using C4.5 Algorithm," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 14, no. 1, pp. 11–20, May 2023, doi: 10.21512/comtech.v14i1.8243.
- [8] Z. Sitorus, D. Prayogi, M. A. Rizko, A. G. Suteja, and M. R. Harahap, "Implementation of the Insertion Sort Algorithm to Sort Positive Integers in Ascending Order Using Flowgorithm," *Journal of Information Technology, computer science and Electrical Engineering (JITCSE)*, vol. 1, no. 3, pp. 323–328, 2024, doi: 10.30596/jitcse.
- [9] A. Putera Utama Siahaan, N. Azizah Harahap, R. Yuni Simanullang, and P. Wanny, "Analysis of Inpatient Data Using Cluster Analysis on Simulation Dataset," *Bulletin of*

- Information Technology (BIT)*, vol. 6, no. 1, pp. 33–39, 2025, doi: 10.47065/bit.v5i2.1830.
- [10] W. Cetera, W. Gogołek, A. Żołnierski, and D. Jaruga, “Potential for the use of large unstructured data resources by public innovation support institutions,” *J. Big Data*, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00610-6.
 - [11] M. Overton, S. Larson, L. J. Carlson, and S. Kleinschmit, “Public data primacy: the changing landscape of public service delivery as big data gets bigger,” *Global Public Policy and Governance*, vol. 2, no. 4, pp. 381–399, Dec. 2022, doi: 10.1007/s43508-022-00052-z.
 - [12] W. Jin, “Challenges and Innovative Countermeasures Faced by Public Administration in the Context of Big Data and Internet of Things,” *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/8949365.
 - [13] J. Ilmiah, E. & Erudisi, A. Zahara, M. I. Kabullah, and R. E. Putera, “JIEE: Effectiveness of the OSSRBA (Online Single Submission Risk Based Approach) System in Business Licensing Services in Payakumbuh City DPMPTSP,” vol. 3, pp. 22–40, 2023.
 - [14] R. Farta Wijaya, F. Kurniawan, R. R. Putrai, A. Alvin, F. Sains, and D. Teknologi, “OPTIMIZING THE MANAGEMENT OF VILLAGE ACTIVITIES THROUGH INFORMATION SYSTEMS FOR TRANSPARENCY AND ACCOUNTABILITY IN PERTUMBUKAN VILLAGE, WAMPU DISTRICT,” 2024.
 - [15] L. Cao, “On Machine Learning and Public Administration,” *Frontiers in Management Science*, vol. 1, no. 1, Aug. 2022, doi: 10.56397/fms.2022.08.01.
 - [16] S. Kasus, : Hoyweapstore, D. Triyansyah¹, and D. Fitrianah², “Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing”, doi: 10.22441/incomtech.v8i2.4174.