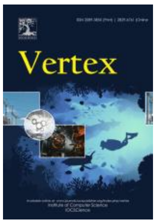

ISSN 2089-385X (Print) | 2829-6761 (Online), Volume 15, No. 2 (2026), pp. 25-33

Vertex

Published by: Institute of Computer Science (IOCS)

Analyzing Consumer Purchasing Behavior in Electrical Supply Stores Using Association Rules

Yunas Akbar[1], Tiwuk Wahyuli Prihandayani[2], Novianti Madhona Faizah[3], Luky Fabrianto[4], Ryan Rakryan[5]

[1,3,5]Computer Science Department, Faculty of Engineering, Universitas Tama Jagakarsa, Jakarta, Indonesia

[2]Information Systems Department, Faculty of Computer Science, Mercu Buana University, Jakarta, Indonesia

[4]Digital Business Department, Faculty of Economic & Business, Universitas Nusa Mandiri, Depok, Indonesia

Abstract

This study aims to improve inventory management and to analyze consumer purchasing behavior through data exploration and the application of association rule mining. The dataset used in this research consists of sales transaction records of electrical products collected over a one-year period. Due to the wide variety of items sold, product categorization is conducted to support more effective analysis and interpretation of purchasing patterns. The method applied in this study is association rule mining using the Apriori algorithm. This method is employed to discover relationships and co-occurrence patterns among items in transaction data. The minimum thresholds used in this study are support ≥ 10% and confidence ≥ 30%, ensuring that only significant and reliable association rules are generated. The results of the analysis reveal several important patterns, with the strongest rule identified as: “Lakban, Switch, and Socket → Cable,” which has a confidence value of 46%. This indicates that customers who purchase Lakban, switches, and sockets have a 46% likelihood of also purchasing cables. The findings provide insights into customer purchasing behavior that can be utilized to optimize inventory control, improve product arrangement, and develop effective cross-selling strategies. Furthermore, this study demonstrates that the application of association rule mining can support data-driven decision-making, enhance operational efficiency, and contribute to increased sales performance and customer satisfaction.

Article Info

Article history:

Received : May 19,2026
Revised : May 25, 2026
Accepted : Jun 05, 2026

Keywords:

Apriori;
Association Rules;
Confidence;
EDA;
Support;

Corresponding Author:

Novianti Madhona Faizah,
Computer Science Department, Faculty of Engineering,
Universitas Tama Jagakarsa,
Jl. TB Simatupang No.152, Jakarta Selatan 12530, Indonesia.
novianti@jagakarsa.ac.id

Introduction

In today’s highly competitive business environment, companies are required to develop effective sales strategies and reliable inventory management systems(Mantik et al., 2023). However, many small and medium enterprises (UMKM) still face challenges in utilizing their transaction data optimally(Angraeni et al., 2025). Although sales data are continuously recorded, they are often stored without further analysis, resulting in missed opportunities to extract valuable insights for improving business performance. This condition highlights the need for a data-driven approach to support decision-making processes, particularly in understanding customer purchasing behavior and optimizing stock management(Bandyopadhyay et al., 2021).

To address this problem, this study proposes the application of data mining techniques, specifically Association Rule Mining using the Apriori algorithm, to analyze sales transaction data(Kumar et al., 2023). In addition, Exploratory Data Analysis (EDA) is employed as an initial step to understand data characteristics, detect patterns, and identify anomalies(Widiastuti et al., 2025). The study focuses on optimizing support and confidence thresholds to generate meaningful and relevant association rules that can support business strategies such as product placement and cross-selling.

The main objective of this research is to identify purchasing patterns and relationships among products based on transaction data. Furthermore, this study aims to uncover trends, detect outliers, and provide insights that can enhance inventory management and improve customer shopping experience(Bunda, 2020).

From a theoretical perspective, data mining is a multidisciplinary field that integrates statistics, machine learning, and artificial intelligence to extract useful knowledge from large datasets. Association Rule Mining, commonly applied in market basket analysis, is widely used to identify relationships between items and support decision-making in retail environments(Fabrianto, 2022).

The expected results of this study include the discovery of that can be utilized to optimize product arrangement, improve inventory control, and enhance sales performance. Ultimately, this research is expected to contribute to more efficient business operations, better customer satisfaction, and data-driven strategic planning significant (Nurhidayanti et al., 2022).

From a theoretical perspective, data mining is a computational approach used to explore and analyze large datasets in order to discover meaningful patterns, relationships, and knowledge. It integrates techniques from statistics, machine learning, and artificial intelligence to support data-driven decision-making. Among various data mining techniques, association rule mining is widely applied to identify relationships between items within transactional data, commonly known as market basket analysis(Fujo et al., 2022).

The process of Knowledge Discovery in Databases (KDD) provides a systematic framework that includes data selection, preprocessing, data mining, pattern evaluation, and knowledge presentation(Shu & Ye, 2023). In this context, preprocessing plays a crucial role in ensuring data quality through steps such as data cleaning, transformation, and feature selection. Furthermore, Exploratory Data Analysis (EDA) is employed as an initial stage to understand data characteristics, identify trends, and detect anomalies that may influence the analysis(Nosiel et al., 2021).

This study utilizes the Apriori algorithm, a well-established method in association rule mining, to generate frequent itemsets and discover relationships between products based on support and confidence measures. Through iterative candidate generation and evaluation, Apriori is capable of identifying significant association rules that reflect consumer purchasing patterns. The integration of EDA and product-category transformation with Apriori-based market basket analysis for electrical retail transaction data.

Method

This study employs a descriptive quantitative approach to analyze transactional data and uncover purchasing patterns(Al-Sayyed et al., 2025). The descriptive method is used to systematically organize, summarize, and interpret historical data in order to support data-driven decision-making (Profitabilitas Pada Perusahaan Perbankan Yang Terdaftar Di Bursa Efek Indonesia Bayu Wulandari et al., 2022). The research focuses on applying exploratory data analysis and association rule mining to identify relationships among products based on transaction records, as illustrated in Figure 1.

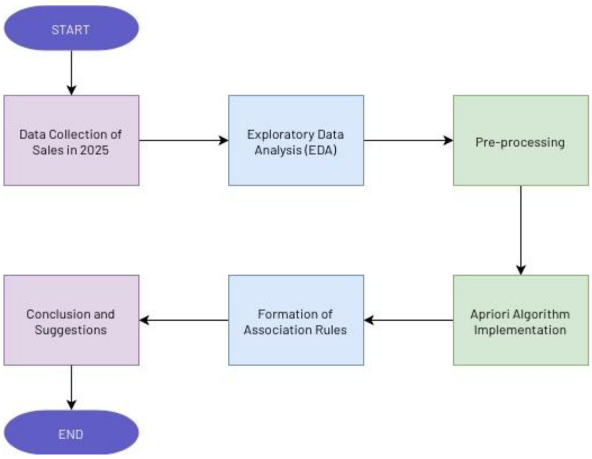

Figure 1. Research Method

Datasets

The dataset used in this study consists of sales transaction data collected over a one-year period (January–December 2025) from a retail business in the electrical goods sector. The dataset includes 1,039 transactions, 42 product types, and 1,814 items sold with varying quantities. The data were obtained from the store's computerized sales records and then prepared for further analysis. Tabel 1 shows all attributes in dataset.

Tabel 1. Attributes

Attribute	Description
id_penjualan	Unique code for each transaction
Tgl	Transaction date
nama_barang	Name of the item sold
qty	Quantity of items sold
harga	Unit price of the item
tot_bayar	Total amount to be paid

Pre-processing

Prior to analysis, a preprocessing stage was conducted to ensure data quality and suitability. This stage includes (1) attribute addition, where products are grouped into relevant categories (e.g., cables, switches, sockets, lighting) to improve interpretability (Fadhli et al., 2018)(Wahyudi et al., 2019), and (2) data transformation, where transaction data are converted into a binary format (1 = purchased, 0 = not purchased) to support market basket analysis (*Data Mining Menggunakan Android, Weka, Dan SPSS - Indah Werdiningsih, S.Si., M.Kom., Barry Nuqoba, S.Si., M.Kom., Muhammadun, S.Si., M.Si. - Google Books*, n.d.)(Fabrianto, 2022).

Categorization of Goods Types

To simplify the large number of types of goods, categorization is necessary to facilitate the association rule method in obtaining models, and the resulting models have good confidence values. Categorization follows the item placement pattern proposed by (St Sujana, n.d.). Product categorized was performed based on functional similarity and practical retail classification to ensure that the resulting association rules reflect purchasing behavior at the product-category level rather than at the individual product level.

Association Rules

The main method applied in this research is association rule mining using the Apriori algorithm, which is widely used to identify frequent itemsets and generate association rules from large transactional datasets (Wijaya, 2017). The procedure is conducted chronologically as follows: (A) generating candidate itemsets, (B) calculating support values, as shown in formula (1), (C) pruning itemsets below the minimum support threshold, as shown in formula (2), (D) generating higher-order itemsets iteratively, and (E) forming association rules based on frequent itemsets, as shown in formula (3)(Homepage et al., 2025). This approach aligns with the Knowledge Discovery in Databases (KDD) framework, which emphasizes systematic pattern extraction and evaluation (Nurhidayanti et al., 2022).

Support for an item,

$$Support\ (A) = \frac{Number\ of\ transaction\ that\ consist\ \ A}{Total\ transaction} \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

Support for two items,

$$Support\ (A \cup B) = \frac{Number\ of\ transaction\ that\ consist\ \ A\ and\ B}{Total\ transaction} \ldots\ldots\ldots\ldots\ldots \textit{(2)}$$

Confidence value from rule A=>B,

$$Confidence = P\ A\ \mid\ B = \frac{Number\ of\ transaction\ that\ consist\ \ A\ and\ B}{Number\ of\ transaction\ that\ consist\ \ A} \ldots\ldots\ldots\ldots \textit{(3)}$$

The evaluation of the generated rules is based on standard metrics, including support and confidence, which measure the frequency and reliability of the association rules(Patel et al., 2022). In this study, minimum thresholds for support and confidence are applied to ensure that only meaningful and actionable rules are selected for analysis and decision-making.

To ensure the robustness and reproducibility of the experimental process, this study defines a structured experimental setup. The dataset is first divided into transactional representations without applying data splitting, as the Apriori algorithm operates on unsupervised pattern discovery rather than predictive modeling(Fabrianto et al., 2021). The preprocessing stage includes data cleaning to remove inconsistencies and duplicates, followed by categorical grouping of products and binary transformation of transaction data. All experiments are conducted using the WEKA data mining tool, which provides a reliable implementation of the Apriori algorithm and facilitates parameter tuning(Attwal et al., 2020).

The experimental configuration involves setting minimum support and confidence thresholds to filter significant association rule(Fadhli et al., 2018)s. In this study, the thresholds are empirically determined at support ≥ 10% and confidence ≥ 30%, based on preliminary observations to balance between rule significance and rule quantity. Several iterations are performed by adjusting these thresholds to analyze their impact on the number and quality of generated rules. The evaluation focuses on identifying the most meaningful rules that reflect actual purchasing behavior, avoiding overly general or trivial associations(Padma & Mishra, 2022).

Furthermore, the quality of the generated rules is assessed using multiple evaluation metrics, including support, confidence, lift, and leverage. Lift is used to measure the strength of association beyond random chance, while leverage evaluates the difference between observed and expected co-occurrence. The results are then interpreted in the context of business applications, such as product placement optimization and cross-selling strategies. This experimental setup ensures that the findings are not only statistically valid but also practically relevant for decision-making in retail environments(Ahn & Kim, 2023).

Results

This study analyzes sales transaction data collected over a one-year period, consisting of 1,039 transactions, 42 product types, and 1,814 items sold. The dataset includes attributes such as transaction ID, date, product name, quantity, unit price, and total payment, Tabel 2 sample of transactions. These data were processed through several stages, including Exploratory Data Analysis (EDA), preprocessing, and association rule mining using the Apriori algorithm(Sahoo et al., 2019).

Tabel 2. Transaction sample

id_penjualan	tgl	nama_barang	qty	harga	tot_bayar
202501016600	1/4/2025	LED 15 Watt Philips	3	55000	165000
202501016609	1/4/2025	Kabel 2 x 1.5 (m)	10	12000	120000
202501016612	1/4/2025	Stopkontak 1	4	15000	60000
202501016611	1/2/2025	LED 9 Watt Philips	4	40000	160000
202501016618	1/3/2025	Kabel 2 x 1.5 (m)	10	12000	120000

The EDA process was conducted to understand the structure and characteristics of the dataset. Descriptive statistical analysis was used to summarize the data, while incomplete or inconsistent records were identified for further cleaning. The results of EDA indicate that the dataset is suitable for association analysis after preprocessing, Tabel 3 shows statistic summary and Tabel 4. one year sales recap.

Tabel 3. Statistic summary

	qty	harga	tot_bayar
Mean	4	29199	69748
Standard Error	0	654	1479
Median	2	15000	55000
Mode	1	12000	60000
Standard Deviation	4	27846	63003
Range	19	135000	413000
Minimum	1	5000	7000
Maximum	20	140000	420000
Sum	6380	52967000	126523000
Count	1814	1814	1814

Tabel 4. Sales recapitulation

Month	Total (Rp)
January	12,999,000
February	11,233,000
March	9,391,000
April	10,205,000
May	10,087,000
June	11,113,000
July	13,711,000
August	8,643,000
September	12,125,000

October	9,055,000
November	9,431,000
December	8,530,000

Categorization of Items Types

To support the implementation of the Apriori algorithm, the *nama_barang* (item name) attribute is simplified by grouping products into their respective categories. This transformation is performed to improve computational efficiency and effectiveness during the analysis process. There are 42 items sold, it is necessary to simplify or categorize them to get good analysis results, so that the 42 types of items sold become 14 categories, as shown in Tabel 5 and figure 2 illustrated distribution of all categories.

Tabel 5. Categorization

Categories	Item name		
Screwdriver	Obeng setting	Obeng Godwin	
Cable	Kabel 2 x 1.5 (m)	Kabel 3 x 1.5 (m)	Kabel Audio
Electrical Tape	Selotip Nito		
Light bulb	LED 9 Watt Philips	LED 15 Watt Philips	LED 5 Watt Philips
Power plug	T Broco	Steker Broco	
Switch and Socket	Stop Kontak 2 outbouw	Stopkontak 1	
Pliers	Tang potong	Tang multi	

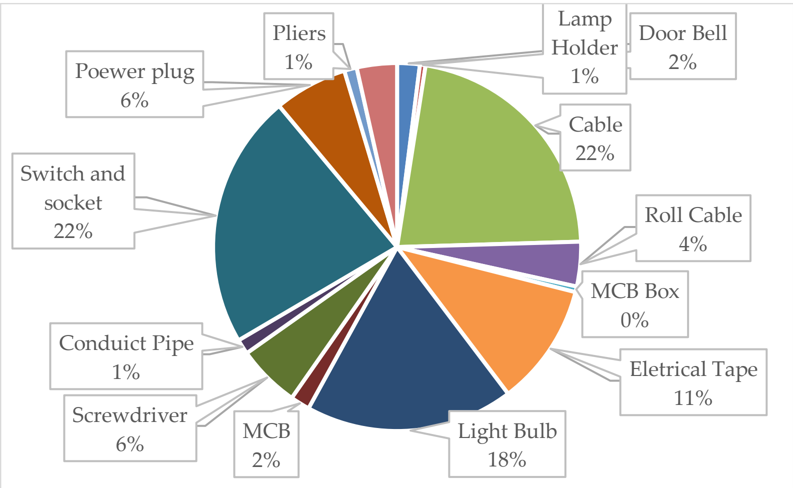

Figure 2. Items distribution after categorization

Attribute Transformations

In the preprocessing stage, data cleaning and duplicate removal were performed to improve data quality. Furthermore, product grouping was applied to classify items into relevant categories. The dataset was then transformed into a binary format, where each product category is represented as "1" (purchased) or "0" (not purchased), Tabel 6 and 7 shows the before and after transformations. These transformations enable efficient implementation of the Apriori algorithm in identifying item associations.

Tabel 6. Real quantity

id_penjualan	Cable	MCB Box	Electrical Tape	Light bulb	MCB	Screwdriver	Switch and Socket
202401016596	5	0	0	3	0	0	1
202401016597	1	0	2	2	0	0	2
202401016598	3	0	2	4	2	1	1

202401016599	0	1	0	3	0	1	3
202401016600	0	0	2	1	0	0	1

Tabel 7. Data transformation

id_penjualan	Cable	MCB Box	Electrical Tape	Light bulb	MCB	Screwdriver	Switch and Socket
202401016596	1	0	0	1	0	0	1
202401016597	1	0	1	1	0	0	1
202401016598	1	0	1	1	1	1	1
202401016599	0	1	0	1	0	1	1
202401016600	0	0	1	1	0	0	1

The Apriori algorithm was applied using a minimum support threshold of ≥ 10% and a minimum confidence threshold of ≥ 30%. Based on the support analysis, four product categories met the minimum support criteria: electrical cables, switches and sockets, light bulbs, and electrical tape, as shown at Tabel 8. These frequent itemsets were then used to generate association rules.

Tabel 8. Categories supports

Categories	Support
Cable	22.54%
Switch and Socket	21.65%
Light bulb	17.20%
Electrical Tape	10.91%
Power plug	6.64%
Screwdriver	5.63%
Cable Roll	4.03%
Test pen	3.74%
Door Bell	2.14%
MCB	1.84%
Conduit Pipe	1.42%
Pliers	1.19%
Lamp holder	0.53%
MCB Box	0.53%

The results show several significant association patterns. For two-item combinations, the rule "Electrical Cable → Light Bulb" achieved the highest confidence value of 41%, indicating a strong likelihood that customers purchasing cables will also purchase light bulbs, as shown in Tabel 9. Other notable rules include "Light Bulb → Switch and Socket" (36%) and "Cable → Electrical Tape" (35%). These findings demonstrate meaningful relationships between commonly co-purchased electrical products.

Tabel 9. Two items rules

2 Items rules			No of transaction	Confidence
Cable	=»	Electrical Tape	65	35%
Cable	=»	Light bulb	75	41%
Cable	=»	Switch and Socket	120	33%
Electrical Tape	=»	Light bulb	29	34%
Electrical Tape	=»	Switch and Socket	50	27%
Light bulb	=»	Switch and Socket	85	36%

For higher-order combinations, the strongest rule identified is "Electrical Tape, Switch and Socket → Cable" with a confidence value of 46%. This indicates that customers who purchase these three items have a high probability of also purchasing cables. Another significant rule is "Cable and Light Bulb → Switch and Socket" with a confidence of 41%, suggesting a consistent purchasing pattern among complementary products. Tabel 10 shows the 10 best rules on this research.

Tabel 10. Best Rules

No	Items Combination			No of transaction		Third items	No of transaction	Confidence
1	Electrical Tape	,	Switch and Socket	50	=»	Cable	23	46%
2	Cable	,	Light bulb	75	=»	Switch and Socket	31	41%
3	Light bulb	,	Switch and Socket	85	=»	Cable	31	36%
4	Cable	,	Electrical Tape	65	=»	Switch and Socket	23	35%
5			Electrical Tape	184	=»	Cable	65	35%
6	Electrical Tape	,	Light bulb	29	=»	Cable	10	34%
7			Switch and Socket	365	=»	Cable	120	33%
8			Cable	380	=»	Switch and Socket	120	32%
9			Light bulb	290	=»	Switch and Socket	85	29%
10			Electrical Tape	184	=»	Switch and Socket	50	27%

Overall, the results demonstrate that association rule mining using the Apriori algorithm can effectively identify purchasing patterns and relationships among products. These findings can be utilized to optimize product placement, improve cross-selling strategies, and enhance inventory management. The study confirms that data-driven analysis provides valuable insights for improving business performance and customer experience.

From a broader perspective, this study supports the transition from conventional retail management to smart retail analytics, where historical transaction data are systematically analyzed to generate actionable insights. Such capabilities are aligned with Industry 4.0 principles that emphasize data-driven decision-making, operational efficiency, and intelligent business processes.

Conclusion

This study shows that the APriori algorithm effectively identifies consumer purchasing patterns, with the strongest rule "Electrical Tape and Switch → Cable" achieving a confidence of 46%. The results indicate that cables and switches are key products with strong associations to other items, providing important insights into customer buying behavior.

Based on these findings, it is recommended to optimize product placement by grouping frequently co-purchased items and to apply product bundling strategies to increase sales. Cables can be prioritized as a main product due to their strong associations. Future research may integrate additional data mining techniques and include customer-related data to obtain deeper insights.

Future research should investigate the integration of Apriori with other data mining techniques, such as clustering and classification, as well as customer profiling attributes. Such integration may provide deeper insights into customer segmentation, purchasing preferences, and personalized marketing strategies, thereby enhancing the effectiveness of smart retail analytics and decision-making.

References

Ahn, B., & Kim, B. (2023). A Decision-Making Model for Selecting Product Suppliers in Crop Protection Retail Sector. *Administrative Sciences 2023, Vol. 13, Page 97*, *13*(4), 97. https://doi.org/10.3390/ADMSCI13040097

Al-Sayyed, H. F., AbuZeinah, Z. F., & Al-Awwad, N. J. (2025). Food purchasing pattern is correlated to gender, educational level, and income: Results from the food purchasing habits survey. *PLOS ONE*, *20*(6), e0322781. https://doi.org/10.1371/JOURNAL.PONE.0322781

Angraeni, D. F., Rahaningsih, N., Dana, R. D., & Rohmat, C. L. (2025). PEMANFAATAN ALGORITMA K-MEANS DALAM ANALISIS DATA PENJUALAN TOKO BUYUNG UPIK JS DI LAZADA. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(2), 2830–7062. https://doi.org/10.23960/JITET.V13I2.6438

Attwal, K., Mathematical, A. D.-A. and A. in, & 2020, undefined. (2020). Exploring data mining tool-Weka and using Weka to build and evaluate predictive models. *Mililink.Com*, *19*(6), 451–469. https://www.mililink.com/upload/article/131953194aams_vol_196_april_2020_a3_p451-

469_kanwal_preet_singh_attwal.pdf

Bandyopadhyay, S., Thakur, S. S., & Mandal, J. K. (2021). Product recommendation for e-commerce business by applying principal component analysis (PCA) and K-means clustering: benefit for the society. *Innovations in Systems and Software Engineering*, *17*(1), 45–52. https://doi.org/10.1007/S11334-020-00372-5/METRICS

Bunda, Y. P. (2020). Algoritma FP-Growth Untuk Menganalisa Pola Pembelian Oleh-Oleh (Studi Kasus Di Pusat Oleh-Oleh Ummi Aufa Hakim). *Riau Journal of Computer Science*, *06*(01), 34–44.

Data Mining Menggunakan Android, Weka, dan SPSS - Indah Werdiningsih, S.Si., M.Kom., Barry Nuqoba, S.Si., M.Kom., Muhammadun, S.Si., M.Si. - Google Books. (n.d.)

Fabrianto, L. (2022). Shopping Behavior At Certain Times And Weeks Using Data Mining Methods Analysis Of Consumer. *Mantik*, *6*(2), 1946.

Fabrianto, L., Faizah, N. M., Prasetyo, J. H., Prakoso, B. S., & Wiharso, G. (2021). Aturan Asosiasi Antar Item Terjual pada Data Penjualan Minimarket Milik Komunitas di Hari Besar Tertentu Menggunakan Algoritma Apriori. *JURNAL AKUNTANSI, EKONOMI Dan MANAJEMEN BISNIS*, *9*(2), 208–215. https://doi.org/10.30871/JAEMB.V9I2.3621

Fadhli, M., Ketua Dewan Editor Zulfan, Mk., Editor Pelaksana Munawir, M., Baihaqi, M., Sekretaris Yeni Yanti, M., Mitra Bestari Ir Yuwaldi Away, M., Gani, T. A., & Melinda, Me. (2018). Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, *1*(2). https://doi.org/10.32672/JNKTI.V1I2.771

Fujo, S. W., Subramanian, S., & Khder, M. A. (2022). *Customer Churn Prediction in Telecommunication Industry Using Deep Learning*. https://doi.org/10.18576/isl/110120

Homepage, J., Aria Hibnastiar, N., Setiawan, A. F., & Susanto, E. H. (2025). Penerapan Algoritma Apriori dalam Menentukan Rekomendasi Paket Produk. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *5*(1), 321–331. https://doi.org/10.57152/MALCOM.V5I1.1782

Kumar, B., Roy, S., Sinha, A., Iwendi, C., & Strážovská, Ľ. (2023). E-Commerce Website Usability Analysis Using the Association Rule Mining and Machine Learning Algorithm. *Mathematics*, *11*(1). https://doi.org/10.3390/math11010025

Mantik, J., Fabrianto, L., Prasetyo, J. H., Faizah, N. M., & Solichatun, S. (2023). Inventory management for essential oil UMKM: enhancing business performance with data mining. *Jurnal Mantik*, *7*(2), 702–711. https://doi.org/10.35335/MANTIK.V7I2.3909

Nosiel, N., Sriyanto, S., & Maylani, F. (2021). Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Pada UMKM Gerabah. *Prosiding Seminar Nasional Darmajaya*, *1*, 72–86.

Nurhidayanti, D., Kurniawati, I., & Artikel, S. (2022). Implementasi Algoritma Apriori Dalam Menemukan Association Rules Pada Persediaan Sparepart Motor. *Innovation in Research of Informatics (INNOVATICS)*, *4*(2), 62–67. https://jurnal.unsil.ac.id/index.php/innovatics/article/view/5300

Padma, A. P., & Mishra, A. K. (2022). Forecasting on Stock Market Time Series Data Using Data Mining Techniques. *Dogo Rangsang Research Journal* , *9*(1), 351–358.

Patel, P., Sivaiah, B., & Patel, R. (2022). Relevance of Frequent Pattern (FP)-Growth-Based Association Rules on Liver Diseases. *Lecture Notes in Networks and Systems*, *431*, 665–676. https://doi.org/10.1007/978-981-19-0901-6_58/COVER

Profitabilitas Pada Perusahaan Perbankan Yang Terdaftar Di Bursa Efek Indonesia Bayu Wulandari, T., Veronica, V., & Prima Indonesia, U. (2022). Pengaruh Dana Pihak Ketiga, Risiko Kredit, Loan to Deposit Ratio Dan Struktur Modal Terhadap Profitabilitas Pada Perusahaan Perbankan Yang Terdaftar Di Bursa Efek Indonesia. *Management Studies and Entrepreneurship Journal (MSEJ)*, *3*(2), 325–335. https://doi.org/10.37385/MSEJ.V3I2.414

Sahoo, K., Samal, A. K., Pramanik, J., & Pani, S. K. (2019). Exploratory data analysis using python. *International Journal of Innovative Technology and Exploring Engineering*, *8*(12), 4727–4735. https://doi.org/10.35940/ijitee.L3591.1081219

Shu, X., & Ye, Y. (2023). Knowledge Discovery: Methods from data mining and machine learning. *Social Science Research*, *110*, 102817. https://doi.org/10.1016/J.SSRESEARCH.2022.102817

St Sujana, A. (n.d.). *PARADIGMA BARU DALAM MANAJEMEN RETAIL MODERN*. Retrieved July 24, 2022, from www.grahailmu.com

Wahyudi, E. N., Utomo, A. P., & Mariana, N. (2019). Pengelompokan Jenis Usaha Umkm Kota Semarang Dalam Rangka Proses Pembinaan Dan Pendampingan Untuk Pengembangan Usaha Dengan Teknik

Data Mining. *Dinamik*, *24*(1), 13–20. https://doi.org/10.35315/dinamik.v24i1.7840

Widiastuti, E., Kusanti, J., & Sulistyowati, H. (2025). Enhancing Recurrent Neural Network Efficacy in Online Sales Predictions with Exploratory Data Analysis. *International Journal of Advanced Computer Science & Applications*, *16*(2), 282. https://doi.org/10.14569/IJACSA.2025.0160230

Wijaya, K. N. (2017). Analisa Pola Frekuensi Keranjang Belanja Dengan Algoritma Apriori (Studi kasus: Minimarket Adi). *Annual Research Seminar (ARS)*, *3*(1), 9–12. http://seminar.ilkom.unsri.ac.id/index.php/ars/article/view/1707