JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER

p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

Application of Apriori Algorithm in Data Mining to Find Consumer Purchasing Patterns in Supermarkets

Purwa Hasan Putra[1,*], Desilia Selvida[2], Muhammad Syahputra Novelan[3]

[1] Jurusan Teknik Komputer dan Informatika, Teknologi Rekayasa Perangkat Lunak, Politeknik Negeri Medan, Medan, Indonesia
[2] Fakultas Ilmu Komputer dan Teknologi Informasi, Ilmu Komputer, Universitas Sumatera Utara, Medan, Indonesia
[3] Pascasarjana, Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia
Email: [1]pputra@polmed.ac.id , [2]desilia.selvida@usu.ac.id, [3]putranovelan@dosen.pancabudi.ac.id
(*Email Corresponding Author: pputra@polmed.ac.id)
Received: June 19, 2025 | Revision: June 20, 2025 | Accepted: June 22, 2025

Abstract

The development of information technology has encouraged the use of transaction data in the retail world to gain deeper business insights. One method used in data mining is the Apriori algorithm, which is able to identify consumer purchasing patterns through association analysis between products. This study aims to apply the Apriori algorithm in finding product combination patterns that are often purchased together by consumers in supermarkets. The data used are sales transactions that have gone through a preprocessing process, including product category classification and transformation into a basket format. The results of the analysis show that products such as biscuits, detergents, and household appliances have the highest support values individually, while product combinations such as (milk, drinks, soap & shampoo, cosmetics) also appear consistently in transactions. The application of the Apriori algorithm with a certain minimum support threshold is able to produce frequent itemsets that represent consumer shopping habits. These findings can be used to develop promotional strategies, product arrangement, and category-based recommendation systems. Thus, this study proves that the Apriori algorithm can be used effectively in the context of data mining to support business decision making in the retail sector, especially supermarkets.

Keywords : Data Mining, Apriori, Market Basket Analysis, Association, Buying Patterns

1. INTRODUCTION

In the ever-growing digital era, data has become a very valuable asset for various sectors, including the retail sector such as supermarkets [1]. Every transaction a consumer makes stores important information that can be analyzed to uncover consumer patterns and behavior. The process of extracting useful information from large amounts of data is known as data mining [2].

One of the data mining approaches that is widely used in the retail context is market basket analysis, which aims to find associative relationships between products that are purchased together. In this context, the Apriori algorithm is one of the most popular algorithms for finding associations or itemset patterns that often appear in transaction data [3]. One method that can be used to extract valuable information from transaction data is data mining, especially the Association Rule Mining technique. This technique is able to identify relationships or patterns of association between products in consumer purchase transactions. One of the most popular algorithms in this technique is the Apriori algorithm. The Apriori algorithm is used to find frequent itemsets (groups of items that are often purchased together) and generate association rules that represent consumer purchase patterns [4][5][6].

Apriori works on the principle that a subset of itemsets that occur frequently should also occur frequently. Using minimum support and confidence thresholds, this algorithm allows supermarket management to determine consumer purchasing tendencies, such as "if a customer buys bread, then they are likely to also buy butter."[7].

The application of the Apriori algorithm in transaction analysis in supermarkets not only provides insight into consumer behavior, but can also be used to develop marketing strategies, product layout, bundling promotions, and stock management. Thus, the integration of data mining technology and retail management practices can improve operational efficiency as well as customer satisfaction [8] [9].

Study [10] This study successfully applied data mining techniques with the Apriori algorithm to produce association rules in the form of rules. One of the product combinations, namely Aqua and ABC Milk Coffee, showed the highest support value of 50% and confidence of 60%, and a lift ratio of 2.50, which indicates a strong positive correlation between the two products. The resulting rules are expected to help Surya Supermarket owners in developing more optimal business strategies, such as managing stock, arranging products on shelves, and providing promotional recommendations, which in turn can encourage increased sales.

Study [11] This paper proposes the addition of category parameters in the Apriori algorithm to limit the formation of similar or repetitive relationship patterns. The results obtained show that increasing the minimum support and minimum confidence values significantly affects the number and quality of association rules generated. In testing with a minimum support value of 0.06 and confidence of 0.6 compared to a minimum support of 0.1 and confidence of 0.8, the number of recommendations decreased by 80%. However, there was an increase in the average support value of 35% and confidence of 32%, indicating an increase in the quality of the association. The addition of category parameters has also been shown to have an impact on the number of recommendations generated. One example is a combination of 4 itemsets from the

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

categories Lamps, Cables, Fittings, and Sockets, which produces one recommendation with an average support of 0.056 and confidence of 0.529.

Based on the research description above, it can be concluded that most previous studies have applied the Apriori algorithm to identify purchasing patterns and determine product bundles in various sectors, such as supermarkets and wedding services. The main objectives of these studies are to increase sales, manage inventory, and provide product bundle recommendations based on the relationship between items that are frequently purchased together. However, this study has a significant difference because it specifically focuses on electronic stores. In addition, although the Apriori algorithm has been proven to be able to identify relationships between products in transaction data, the results obtained have not been fully utilized directly to design ideal product bundles [12]. This is due to the possibility of grouping products with the same type. Therefore, this study aims to overcome this problem by integrating the Apriori algorithm into a web-based system equipped with additional parameters in the form of categories. With this parameter, it is expected that the results of the Apriori algorithm will be more relevant and able to minimize the emergence of product packages that have the same type [13].

Therefore, this study aims to apply the Apriori algorithm in exploring consumer purchasing patterns in supermarkets, in order to provide recommendations that can support more targeted business decision making.

2. RESEARCH METHODOLOGY

This study uses a quantitative approach with an applied experimental method. The main focus of this study is the application of the Apriori algorithm in data mining to find consumer purchasing association patterns, as well as testing the effectiveness of additional parameters in the form of product categories to improve the quality of product package recommendations at figure 1:

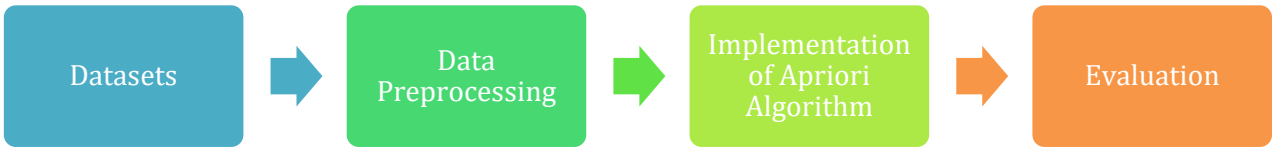

Figure 1. Research Methodology

2.1 Datasets

The data used in this study is sales transaction data obtained from the self-service cashier system in digital format (CSV/Excel). Each transaction contains information in the form of a transaction ID, purchase date, and a list of items purchased by the consumer [11].

2.2 Data Preprocessing

It is the process of changing raw data or commonly known as raw data collected from various sources into cleaner information that can be used for further processing. Before the analysis process is carried out, the raw data will be cleaned and prepared with the following steps: Removing duplicate or irrelevant data. Normalizing product names to be uniform. Grouping products into certain categories (eg: food, drinks, electronics, etc.). Changing transaction data into a basket format (shopping cart) according to the needs of the Apriori algorithm [14].

2.3 Implementation of Apriori Algorithm

At this stage, the Apriori algorithm is applied to find associations between products with the following steps: Determining the minimum support and minimum confidence values. Identifying frequent itemsets based on support values. Forming association rules based on confidence values and calculating the lift ratio to measure the strength of the association. Adding category parameters to limit the formation of associations between products of the same type [15].

2.4 Evaluation

The evaluation was conducted to measure the success and effectiveness of the implementation of the Apriori algorithm in finding consumer purchasing patterns based on transaction data in supermarkets. This evaluation aims to assess the quality of the association rules formed and their impact on the implementation of business strategies, such as product recommendations, inventory management, and goods arrangement [16].

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

3. RESULTS AND DISCUSSION

This study uses the Apriori algorithm to analyze consumer transaction data to find frequent purchasing patterns in supermarkets. After the data preprocessing process in the form of data cleaning, product grouping by category, and data transformation into a basket format, the Apriori algorithm is applied with certain minimum support and minimum confidence values. The application results show that several product categories have a high frequency of occurrence, such as:

3.1 Datasets

This dataset is a sales transaction data in a supermarket that records consumer purchase details based on invoices or shopping receipts. Each row represents one product item purchased in one transaction.

Table 1. Research Methodology Product Purchase Dataset

	InvoiceNo	InvoiceDate	Branch_Splr	Branchname_Splr	Barcodeid	Product_Category
0	536365	2020-12-01 08:26:00	13	Medan	8992753282401	SUSU
1	536365	2020-12-01 08:26:00	13	Medan	8999909192034	MINUMAN
2	536365	2020-12-01 08:26:00	13	Medan	8999909007147	SABUN & SAMPHOO
3	536365	2020-12-01 08:26:00	13	Medan	8999909172234	KOSMETIK
4	536365	2020-12-01 08:26:00	19	Medan	8999909005860	SNACK

*) data source 2025

3.2 Data Preprocesing

Data preprocessing is an important step before applying data mining algorithms such as Apriori. The goal is to clean, tidy, and prepare raw data so that it can be analyzed properly.

Table 2. Product Purchase Dataset

	Product_category InvoiceNo	Electrical Tools	Household appliance	Rice	Biscuits	Detergent
0	536365	0.0	0.0	0.0	0.0	0.0
1	536365	0.0	0.0	0.0	0.0	0.0
2	536365	0.0	0.0	0.0	0.0	0.0
3	536365	0.0	0.0	0.0	0.0	0.0
4	536365	0.0	0.0	0.0	0.0	0.0

*) data source 2025

3.3 Implementation of Apriori Algorithm

The implementation of the Apriori algorithm is carried out through several main stages, starting from the preparation of transaction data to the formation of association rules. The implementation process of the Apriori algorithm is carried out in stages, starting from data preparation, applying algorithm logic, to evaluating the results. Each step plays an important role in ensuring that the patterns found truly reflect consumer shopping behavior and can be utilized in making decisions for self-service businesses.

Table 3. Results of the Apriori Implementation Process

	support	itemsets
0	0.326438	(Alat Rumah Tangga)
1	0.348726	(Biskuit)

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

2	0.172516	(Bumbu)
3	0.166359	(Cokelat)
4	0.273488	(Detergen)
...	...	...
250	0.102204	(Snack, Minuman, Sabun & Samphoo, Kosmetik)
251	0.117473	(Susu, Minuman, Sabun & Samphoo, Kosmetik)
252	0.102081	(Parfum, Sabun & Samphoo, Kosmetik, Obatan)
253	0.106760	(Parfum, Susu, Sabun & Samphoo, Kosmetik)
254	0.102327	(Parfum, Susu, Minuman, Sabun & Samphoo)

*) data source 2025

The table displays the results of the implementation of the Apriori algorithm in the form of a frequent itemsets table, which is a collection of products that are often purchased together by consumers based on transaction data in supermarkets.
support Shows how often a combination of products (itemsets) appears in all transactions. The higher the value, the more often the items are purchased together.
itemsets A combination of products or product categories that often appear together in one transaction (frequent itemsets).
Row 0:
Itemset: (Household Appliances)
Support: 0.326 → means that 32.6% of all transactions contain the product category "Household Appliances".
Row 1:
Itemset: (Biscuits)
Support: 0.348 → this product appears in 34.8% of transactions, meaning it is very commonly purchased.
Row 250:
Itemset: (Snacks, Beverages, Soap & Shampoo, Cosmetics)
Support: 0.102 → this combination of four categories was purchased together in 10.2% of all transactions. While more complex, this is a valuable cross-category consumer shopping pattern for bundling promotions or shelf organization.

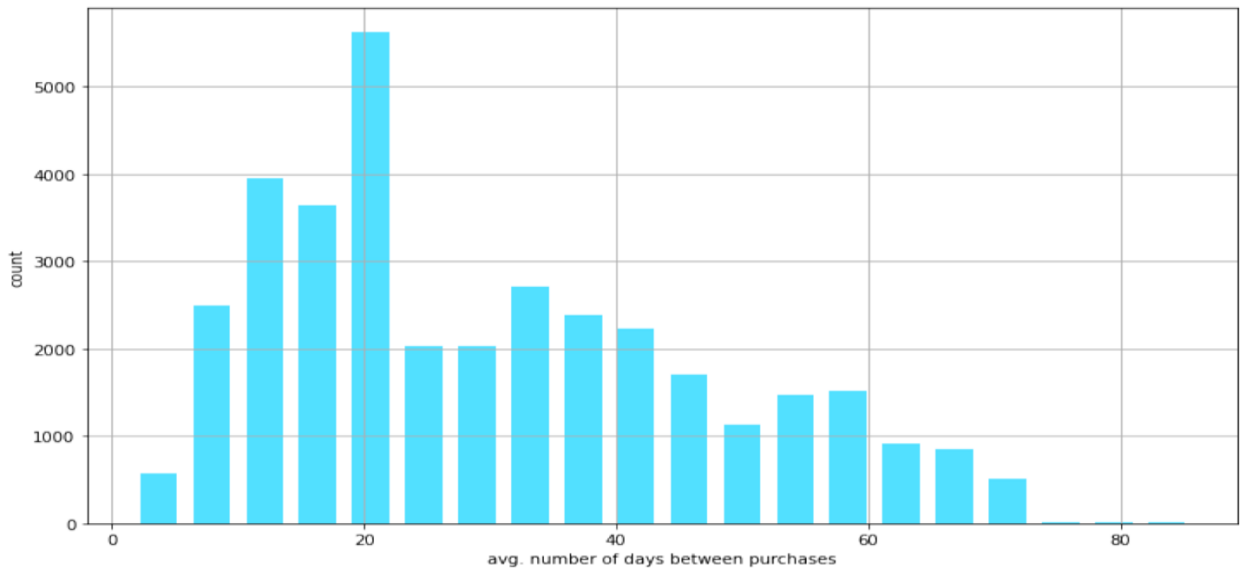

Figure 2. Apriori Implementation Results

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

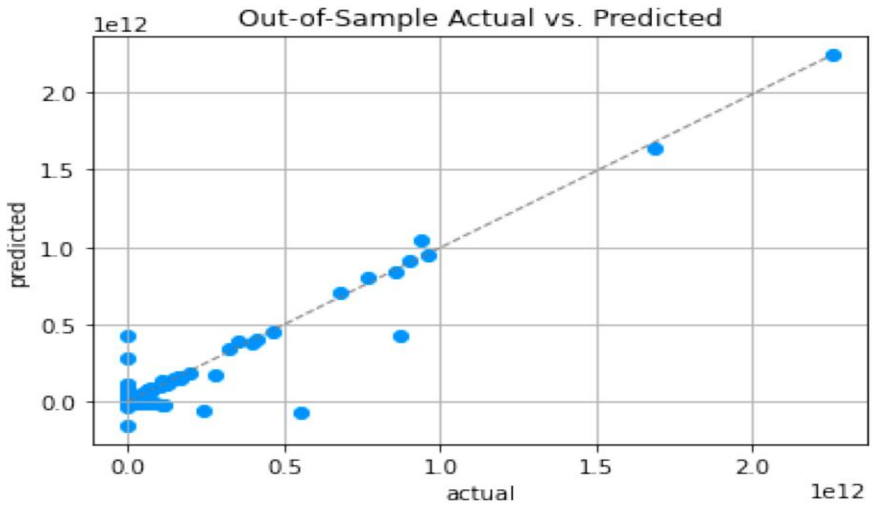

Figure 3. Implementation Results

The figure shows a frequency histogram of the average number of days between purchases by consumers. This is a very important result of consumer buying behavior analysis in the context of customer retention and shopping habits analysis. X-axis (horizontal): shows the average number of days between one purchase and the next. Y-axis (vertical): shows the number of consumers who have this pattern (frequency / number of data). Peak around 20 days, The highest column is around 20 days, meaning that the majority of customers shop again every 20 days on average. Consumer Distribution Tends to Decrease After 30-40 Days, After 30 days, the number of consumers drops significantly. This shows that consumers are less likely to shop again after that time span, which could mean an increased risk of churn or loss of customers. Positively Skewed Distribution, Most of the data is concentrated in the 5–40 days range, with a tail extending to the right (up to 80+ days). This means that there is a small portion of consumers who rarely shop again, perhaps seasonal or non-loyal buyers. Most customers have a repeat shopping pattern within ±20 days. This can be used to compile recommendations for promotion times, marketing email frequencies, and stock scheduling.

4. CONCLUSION

Based on the results of the implementation of the Apriori algorithm on consumer transaction data in supermarkets, a number of frequent itemsets were obtained that show patterns of purchasing products that often appear together. The support value of each product combination provides an overview of how often the product or combination of products is purchased by consumers in the overall transaction. From these results, the following conclusions can be drawn: Popular Individual Products Products with the highest support values are found in the following categories: Biscuits (34.87%) Household Appliances (32.64%) Detergent (27.34%) This shows that these products are among the most frequently purchased individually by consumers. Consistent Product Combinations Purchased Together Several product category combinations with support above 10% indicate cross-category shopping patterns, such as: (Milk, Drinks, Soap & Shampoo, Cosmetics) – 11.74% (Perfume, Soap & Shampoo, Cosmetics, Medicine) – 10.21% (Perfume, Milk, Drinks, Soap & Shampoo) – 10.23% This indicates a tendency for simultaneous purchases for daily necessities and personal care products. Promotion Strategy Development Opportunities These patterns can be utilized for: Arranging bundling promotions based on items that are frequently purchased together. More strategic shelf arrangement to make it easier for consumers to find related products. Developing a product recommendation system on a web-based or mobile supermarket platform. Overall, the Apriori algorithm has proven effective in identifying consumer purchasing patterns. The results of this analysis provide insights that can be used to support decision-making in more targeted and data-based stock management, sales, and marketing.

REFERENCES

[1] D. Puspita and R. Syahri, “Implementasi Algoritma Apriori Untuk Menentukan Strategi,” vol. 8, no. 2, pp. 371–377, 2023, doi: 10.47065/bit.v5i2.156.

[2] V. S. Rifania, S. Saniman, and A. Azlan, “Penerapan Algoritma Apriori Dalam Mencari Pola Pembelian Konsumen,” *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 2, no. 2, p. 201, 2023, doi: 10.53513/jursi.v2i2.5750.

[3] Nurdin and D. Astika, “Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Pada Supermarket Sejahtera Lhokseumawe,” vol. 6, no. 1, pp. 134–155, 2015, doi: 10.29103/TECHSI.V7I1.184.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 4 No 1 June 2025
Page: 221-226
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v4i1.392

[4] B. H. Situmorang, A. Isra, D. Paragya, and D. A. A. Adhieputra, "Apriori Algorithm Application for Consumer Purchase Patterns Analysis," *Komputasi J. Ilm. Ilmu Komput. dan Mat.*, vol. 21, no. 1, pp. 15–20, 2024, doi: 10.33751/komputasi.v21i1.9260.

[5] F. P. Fey, "Application of the Apriori Algorithm to Purchase Patterns," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 553–561, 2023, doi: 10.33022/ijcs.v12i2.3105.

[6] E. Omol, D. Onyango, L. Mburu, and P. Abuonji, "Apriori Algorithm and Market Basket Analysis to Uncover Consumer Buying Patterns : Case of a Kenyan Supermarket," vol. 5, no. 2, pp. 51–63, 2024, doi: 10.36805/bit-cs.v5i2.6082.

[7] N. A. Hibnastiar, A. F. Setiawan, and E. H. Susanto, "Application of Apriori Algorithm for Product Bundle Recommendations Penerapan Algoritma Apriori dalam Menentukan Rekomendasi Paket Produk," vol. 5, no. January, pp. 321–331, 2025.

[8] N. K. Ulfadiyah, P. I Ketut Gede Darma, and D. I Putu Arya, "Implementasi Algoritma Apriori untuk Menemukan Pola Pembelian Konsumen pada Perusahaan Retail," *JITTER - J. Ilm. Teknol. dan Komput.*, vol. 1, no. 2, 2020, [Online]. Available: https://ocs.unud.ac.id/index.php/jitter/article/view/69743.

[9] M. T. Anwar, H. D. Purnomo, M. Novita, and C. H. Primasari, "Implementasi Metode Asosiasi Apriori Untuk Mengetahui Pola Beli Konsumen Dan Rekomendasi Penempatan Produk Pada Swalayan Xyz," *Dinamik*, vol. 25, no. 1, pp. 29–38, 2020, doi: 10.35315/dinamik.v25i1.7747.

[10] S. D. Putri and S. Sitohang, "Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori," *Comput. Sci. Ind. Eng.*, vol. 9, no. 7, pp. 160–167, 2023, doi: 10.33884/comasiejournal.v9i7.7889.

[11] A. I. Zalukhu, D. Sartika, and S. Wahyuni, "Penerapan Algoritma Apriori untuk Optimasi Strategi Penjualan Berdasarkan Analisis Pola Pembelian di Torsa Cafe," vol. 5, no. 4, pp. 295–304, 2024, doi: 10.47065/bit.v5i2.1715.

[12] A. Hermawan, B. Wicaksono, and ..., "Implementasi Algoritma Apriori pada Market Basket Analysis terhadap Data Penjualan Produk Supermarket," *... Alam, Kebumian dan ...*, vol. 2, no. 5, 2024, [Online]. Available: https://journal.arimsi.or.id/index.php/Algoritma/article/view/137%0Ahttps://journal.arimsi.or.id/index.php/Algoritma/article/download/137/228.

[13] W. D. Ramadana, N. Satyahadewi, and H. Perdana, "Penerapan Market Basket Analysis Pada Pola Pembelian Barang Oleh Konsumen Menggunakan Metode Algoritma Apriori," *Bul. Ilm. Math. Stat. dan Ter.*, vol. 11, no. 3, pp. 431–438, 2022.

[14] P. H. Putra, B. Purba, and Y. A. Dalimunthe, "Random forest and decision tree algorithms for car price prediction," vol. 1, no. 2, pp. 81–89, 2023.

[15] N. A. Pradipta and R. D. H. Untari N, "Implementasi Algoritma Apriori Untuk Analisis Pola Pembelian Produk Donat Bolong," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, p. 268, 2024, doi: 10.35889/jutisi.v13i1.1778.

[16] T. Tukino, "Analisis Pola Pembelian Konsumen Pada PT Indoritel Makmur Internasional Tbk Menggunakan Metode Algoritma Apriori," *J. Desain Dan Anal. Teknol.*, vol. 2, no. 1, pp. 75–83, 2023, doi: 10.58520/jddat.v2i1.23.