



## OPTIMIZING RETAIL STRATEGY WITH APRIORI ALGORITHM FOR INFORMED DECISION-MAKING ON CUSTOMER PURCHASING PATTERNS

Yomei Hendra\*<sup>1</sup>, Putri Sakinah<sup>2</sup>, Muhammad Thoriq<sup>3</sup>, Aldo Eko Syaputra<sup>4</sup>

<sup>1234</sup>Universitas Adzka

Email : [yomeihendra@adzka.ac.id](mailto:yomeihendra@adzka.ac.id)

Vol. 17 No.4 2023

### Abstract

**Submit :**  
12/10/2023

**Accept :**  
14/12/2023

**Publish :**  
31/12/2023

In retail business, understanding customer purchase patterns is crucial for enhancing marketing strategies and product placement. This research focuses on analyzing product sales transactions at Minang Mart Lubuk Begalung branch, aiming to identify frequently purchased product combinations and understand the relationships between products through association rules. Data, subjected to preprocessing stages including data cleaning, outlier handling, and normalization, ensures consistent quality. Frequent itemset analysis using the Apriori algorithm with varying minimum support thresholds (0.003, 0.005, and 0.008) provides insights into customer purchase patterns. Association rules with a minimum confidence level of 0.5 and a minimum lift level of 1 yield significant findings, such as the combination of ADES NATURAL 24X600ML and CHEETOS JAGUNG BAKAR 40X40 GR. From these findings, Minang Mart can design more effective marketing strategies and enhance product placement in sales areas. Visualization of purchase patterns through graphs supports a more intuitive understanding of customer preferences. The research results are expected to contribute positively to optimizing sales strategies and strengthening the competitiveness of retail businesses in this era of intense competition.

**Keywords:** Association Rules, Purchase Patterns, Optimaizing Sales, Apriori Argorithm

<http://publikasi.ildikti10.id/index.php/jit>

DOI : <https://doi.org/10.22216/jit.v17i4>

PAGES : 1001-1011



## INTRODUCTION

In the digital era, retail companies operate amidst increasingly complex challenges and intense competition. Strategic decisions play a crucial role in ensuring their sustainability and success. Achieving competitive advantage requires a profound understanding of customer purchasing behavior, and one effective approach is leveraging the entire dataset of sales transactions owned by the company itself. Sales transaction data, considered a treasure trove for companies, holds immense potential to provide critical insights into customer preferences and shopping habits[1]. Patterns within this information can be extracted to gain knowledge[2]. By utilizing the Apriori algorithm as a method to uncover further customer purchasing patterns. This algorithm, which falls into the category of association rule growth algorithms, has proven effective in identifying patterns of relationships among items in transaction datasets[3]. By implementing the Apriori algorithm on sales transaction data, retail companies can unveil significant purchasing patterns and identify associative relationships between products frequently bought together[4]. One of the main advantages of the Apriori algorithm is its ability to filter out strong association rules from a set of transaction data. Consequently, this algorithm provides a profound understanding of cause-and-effect relationships between products, offering retail companies the opportunity to formulate more effective marketing strategies, enhance service

personalization, and optimize inventory management[5]. This study aims to delve into and explore the potential benefits that retail companies can gain through the implementation of the Apriori algorithm in the analysis of customer purchasing patterns. By understanding customer buying patterns, retail companies can make more informed strategic decisions, enhance customer experiences, and optimize their operations. The Apriori algorithm will be applied to the retail company's sales transaction data to identify significant purchasing patterns. The analysis of the results from applying this algorithm will provide valuable insights into customer shopping habits and product relevance, creating a strong knowledge base for strategic decision-making[6]. In the specific application of the Apriori algorithm in analyzing retail company transactions, several previous studies generally did not deeply explore the potential of this algorithm in comprehensively uncovering customer shopping patterns. Previous research tended to highlight an understanding of customer needs but lacked detailed exploration into how the Apriori algorithm could provide in-depth insights into customer shopping behavior, especially within dynamic retail environments. This study aims to fill this knowledge gap by focusing on a more in-depth implementation of the Apriori algorithm in retail transactional data, intending to provide more specific insights into customer purchase patterns and product success aspects within dynamic retail contexts.





Not only that, this research's distinctive approach involves the utilization of various support value variations in analyzing purchasing patterns. This strategy aims to unveil patterns that might be overlooked when using only one support value. Through these variations, it is expected to gain deeper insights into the relationship between products in customer transactions, as well as discover more consistent and significant associations among frequently co-purchased items. These efforts strengthen the uniqueness and added value of this study in comprehensively exploring customer purchasing patterns. This research holds significant strategic implications. By contributing new insights into the analysis of customer purchasing patterns using the Apriori algorithm, it is expected to assist retail companies in optimizing their strategic decisions. The findings from this study could positively impact operational efficiency, enhance customer satisfaction, and improve competitiveness in a rapidly evolving market. In the context of the increasingly dynamic retail business, this research aims to make a substantial contribution to industry knowledge and practices. By employing the Apriori algorithm, it is anticipated that retail companies can gain deeper insights into customer purchasing patterns, paving the way for strategic improvements and a competitive edge in the ever-evolving market.

## RESEARCH METHODS

This research aims to identify and analyze customer purchasing patterns to provide in-depth insights for the retail company. By utilizing the Apriori algorithm, this study seeks to uncover inter-product relationships frequently purchased together by customers[7], opening opportunities to enhance marketing strategies, improve customer experiences, and optimize inventory management[8].



Figure 1. Research Stages

A critical step at the beginning of the research is to thoroughly analyze the challenges faced by the retail industry in the context of customer purchasing patterns. Highlighting the existing challenges, including changes in consumer





behavior, market trend fluctuations, and intensifying competition, is key to understanding the complexity of the business landscape.

Through this analysis, we identified several aspects that became the focus of the research, including but not limited to:

- How can unexpected changes in consumer preferences influence customer purchasing patterns, and why is recognizing these elements at the core of efforts to understand market dynamics?
- Facing rapid changes in the business environment, what are the main challenges encountered by retail companies in optimizing their marketing strategies, especially in attracting customer attention and stimulating additional purchases?
- Why is efficiency and effectiveness in inventory management considered a crucial element in maintaining the smooth operation of retail companies, and how can errors in demand forecasting or inventory management significantly impact profits?
- In the context of improving customer experience, how can retail companies understand and implement personalized services and intelligent product recommendations to maintain and attract customer interest?

After completing the problem analysis stage, the next step is the literature review phase related to the title. A comprehensive literature review is conducted to understand the concept of customer purchase patterns and gain a solid understanding of the application of

the Apriori algorithm in the context of the retail industry. Literature sources include specialized books, recent research articles, and other relevant literature. This stage involves a thorough examination of scholarly journals, books, and other related literature to gather insights into existing theories, methodologies, and findings related to customer purchase patterns and the Apriori algorithm. The goal is to build a strong theoretical foundation and identify gaps in current knowledge that the research can address. The insights gained from the literature review will inform the conceptual framework and methodology of the research, ensuring that it contributes to the existing body of knowledge in a meaningful way.

## RESULT

This research utilizes the Apriori algorithm, a classic method in data mining, to process data and generate recommendations for retailers[1]. The algorithm is chosen because it not only identifies frequently sold products but also provides recommendations for related products, enhancing the potential for customer purchases[4].

There are three stages in the fundamental analysis of associations:

- 1) High-Frequency Pattern Analysis  
This process searches for item combinations that meet the minimum support score in the database. The support score of an item is calculated using the formula[10]:





$$Support(A) = \frac{\sum \text{containing transaction A}}{\text{total transactions}} \times 100\% \quad (1).$$

while the support score of two items is calculated by the formula:

$$Support(A, B) = \frac{\sum \text{containing transaction A and B}}{\text{total transaction}} \times 100\% \quad (2).$$

## 2) Association Rule Generation

After obtaining high-frequency patterns, the next step is to find association rules. This process involves determining rules that meet the minimum confidence level. The confidence score for the rule  $A \rightarrow B$  is calculated using the formula[11]:

$$Confidence(A \Rightarrow B) = \frac{\sum \text{containing transaction A and B}}{\sum \text{containing transaction A}} \times 100\% \quad (3).$$

## 3) Itemset Relationships

Independence occurs between itemsets A and B if  $P(A \cup B) = P(A)P(B)$ ; conversely, itemsets A and B are correlated as events. The lift between the occurrence of A and B is measured by the formula[12]:

$$Lift(A, B) = \frac{P(A \cup B)}{P(A)P(B)} \quad (4).$$

Association rules are then derived from the set of items with support and confidence as the threshold.

The data used for this research encompasses product purchase transactions at Minang Mart's Lubuk Begalung branch from July 13, 2023, to

July 20, 2023. The information contained in this data includes invoice numbers, product names, and the quantity of products purchased in each transaction, as illustrated in Table 1.

No	Faktur	Produk
1	MTR3818SL00116 82	Kg Shortcake Coklat 225 Gr
2	MTR3818SL00116 82	Indomilk Uht Coklat 190 MI
3	MTR3818SL00116 82	Lifebuoy Shp Anti Hairfall 170 MI
4	MTR3818SL00116 82	Sampoerna Mild 16
5	MTR3818SL00116 83	Paddle Pop Rainbow Pwr 48X60MI
...	....	....
1212	MTR3818SL00123 40	Aqua Mineral 600MI
1213	MTR3818SL00123 40	Zwitsal Bb Bath Nat Hair N Body 200MI
1214	MTR3818SL00123 40	Zwitsal Bb Bath Nat Hair N Body 200MI
1215	MTR3818SL00123 40	Formula Ripple Med
1216	MTR3818SL00123 40	Lifebuoy Bw Lemon Fresh Btl 100 MI

Table 1. Transaction Data

Before entering the analysis stage, the data undergoes several preprocessing steps to ensure accuracy and consistency. In this study, the preprocessing focus is only on:

- **Data Cleaning**  
This step aims to address missing values or missing values within the dataset. This can be done by removing rows or columns that have missing values, filling in missing values with the mean or median value, or using data imputation techniques.
- **Identification and Handling of Outliers**  
Outliers are values that significantly differ from the majority of the data. This step is important to address outliers that







might affect your analysis. You can use statistical methods like the Interquartile Range (IQR) method or other techniques to detect and handle outliers, such as removing them or transforming them into more representative values.

- Data Normalization

Data normalization helps to ensure consistency in the scale of numeric variables. This allows for better comparison between variables that have different scales. Normalization methods, such as Min-Max Scaling or Z-score Normalization, can be used to transform the scale of numeric variables into a specific range or a normal distribution.

In this section, we elaborate on the results of the frequent itemsets analysis found in customer purchase data using the Apriori algorithm, implemented with the Python programming language. Experiments were conducted with variations in the minimum support threshold (0.003, 0.005, and 0.008) to understand how the results vary. The minimum support threshold is the minimum percentage of the total transactions required for an itemset to be considered frequent. We aim to understand to what extent this threshold affects the results of the analysis.

With a minimum support threshold value of 0.003, 271 combinations of frequent itemsets were successfully identified. This threshold allowed us to identify a significant number of frequent itemsets, encompassing various combinations of

products in customer purchase transactions.

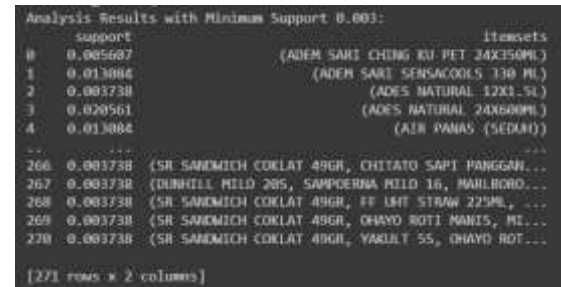


Figure 2. Analysis Results Minimum Support: 0.003

Increasing the minimum support threshold to 0.005 narrowed down the list of frequent itemsets. Nevertheless, it still managed to detect significant itemsets, albeit covering more common products. A total of 137 itemset combinations were identified at this threshold.

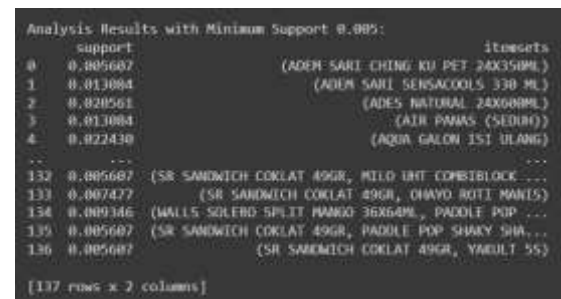
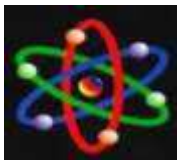


Figure 3. Analysis Results Minimum Support: 0.005

At the threshold level of 0.008, only a few highly significant frequent itemsets were identified, encompassing combinations of very common products. A total of 57 itemset combinations were successfully identified at this threshold level.





```

Hasil Analisis dengan Minimum Support 0.008:
Support                               Itemsets
0 0.013884 (ADEM SARI SENSACDOOLS 338 ML)
1 0.020561 (ADES NATURAL 24X600ML)
2 0.013884 (AIR PANAS (SEDUH))
3 0.022430 (AQUA GALON ISI ULANG)
4 0.037383 (AQUA MINERAL 1500ML)
  
```

Figure 4. Analysis Results Minimum Support: 0.008

These results indicate that the minimum support threshold level of 0.005 is an effective choice for further analysis. This level strikes a balance between sensitivity to significant purchasing patterns and reducing the number of itemsets for more efficient analysis. As a result, the threshold level of 0.005 will be used in the next stages of this research.

In this stage, association rule analysis using the Apriori algorithm is conducted to detect patterns in customer purchases and relationships between products frequently bought together. This analysis utilizes a minimum support threshold level of 0.005, resulting in 26 association rules with significant support, confidence, and lift values, implying potential implications for marketing strategies. Here are the results of the association rule analysis:

antecedent	support	consequent	support	support	confidence	LIFT
0	0.020561	0.011215	0.005047	0.272727	24.31012	
1	0.011215	0.020561	0.005047	0.500000	24.31012	
2	0.010222	0.013084	0.007477	0.444444	15.900254	
3	0.013084	0.010222	0.007477	0.576923	15.900254	
4	0.030928	0.033703	0.005047	0.005396	1.480311	
5	0.037283	0.100025	0.005047	0.150000	1.000311	
6	0.034053	0.022430	0.005047	0.175000	10.710706	
7	0.024310	0.014951	0.005047	0.250000	16.710706	
8	0.041121	0.013084	0.007477	0.103018	12.000204	
9	0.013084	0.041121	0.007477	0.576923	12.000204	
10	0.041121	0.013084	0.005047	0.103018	10.020076	
11	0.013084	0.041121	0.005047	0.020721	10.020076	
12	0.010222	0.010002	0.005047	0.133333	17.011113	
13	0.010002	0.010222	0.005047	0.100000	17.011113	
14	0.010025	0.014951	0.003346	0.002593	6.132238	
15	0.014951	0.010025	0.003346	0.025000	6.132238	
16	0.000346	0.041121	0.005047	0.000000	24.500000	
17	0.041121	0.000346	0.005047	0.100364	14.500000	
18	0.041121	0.011215	0.007477	0.101018	4.021400	
19	0.011215	0.041121	0.007477	0.100101	4.021400	
20	0.014951	0.000346	0.003346	0.257183	22.003250	
21	0.000346	0.014951	0.003346	0.025000	22.003250	
22	0.041121	0.010222	0.005047	0.100364	8.100004	
23	0.010222	0.041121	0.005047	0.133333	8.100004	
24	0.014951	0.041121	0.005047	0.175000	0.133333	
25	0.041121	0.014951	0.005047	0.100364	0.133333	

Figure 5. Association Rule Results with Minimum Support: 0.005

The analysis results reveal some interesting association rules, especially those with high confidence. For example, the first rule indicates that customers purchasing "WALLS SOLERO SPLIT MANGO 36X64ML" are likely to also buy "PADDLE POP RAINBOW PWR 48X60ML" with a confidence level of 62.5%. The high lift value (23.88) indicates a strong and significant relationship between these two products.

This analysis provides a deep understanding of customer preferences and purchasing patterns at Minang Mart, laying the groundwork for strategic decision-making related to product placement, promotions, and other marketing strategies.

Based on the analysis of Minang Mart transaction data, this research provides a deeper understanding of customer purchasing patterns and the relationships between products. Association rules reveal correlations between products, allowing for a better understanding of customer preferences. Some rules highlight strong relationships between the purchase of specific products. Here are the top 5 Association Rules from this study:

- 1) Rule 1: If a customer buys ADES NATURAL 24X600ML, they are likely to buy CHEETOS JAGUNG BAKAR 40X40 GR (Confidence: 50%, Lift: 24.32).
- 2) Rule 2: If a customer buys CHEETOS JAGUNG BAKAR 40X40 GR, they are likely to buy ADES NATURAL 24X600ML (Confidence: 27.27%, Lift: 24.32).





- 3) Rule 3: If a customer buys POP MIE AYAM 75 GR, they are likely to buy AIR PANAS (SEDUH) (Confidence: 57.14%, Lift: 33.97).
- 4) Rule 4: If a customer buys AIR PANAS (SEDUH), they are likely to buy POP MIE AYAM 75 GR (Confidence: 44.44%, Lift: 33.97).
- 5) Rule 5: If a customer buys SAMPOERNA MILD 16, they are likely to buy AQUA MINERAL 1500ML (Confidence: 5.56%, Lift: 1.49).

In an effort to enrich the understanding of customer purchasing patterns, we present visualization graphs that reflect the findings of the analysis. The use of the Apriori algorithm with a minimum support threshold of 0,005 resulted in 26 significant association rules. Here is a graphical representation that illustrates some information:

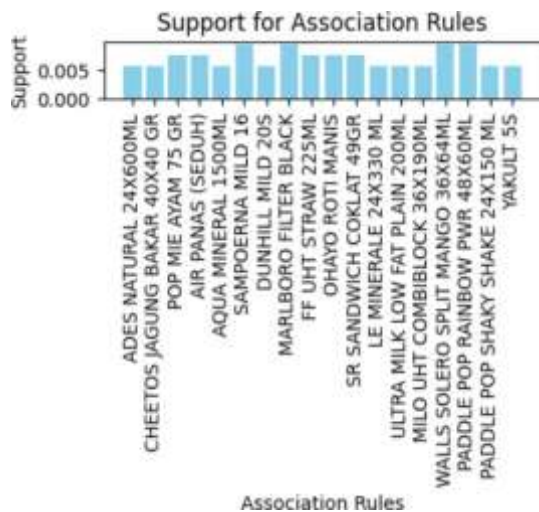


Figure 6. Support Graph of Association Rule Results

Association rules with high Support values indicate combinations of products that are popular and commonly purchased together. High data points on the graph can be the focus of marketing and sales strategies as they reflect significant purchasing patterns.



Figure 7. Confidence Graph of Association Rule Results

Association rules with high Confidence indicate a strong tendency for simultaneous purchases. Focusing on association rules with high Confidence can assist in designing product placement or joint promotions.





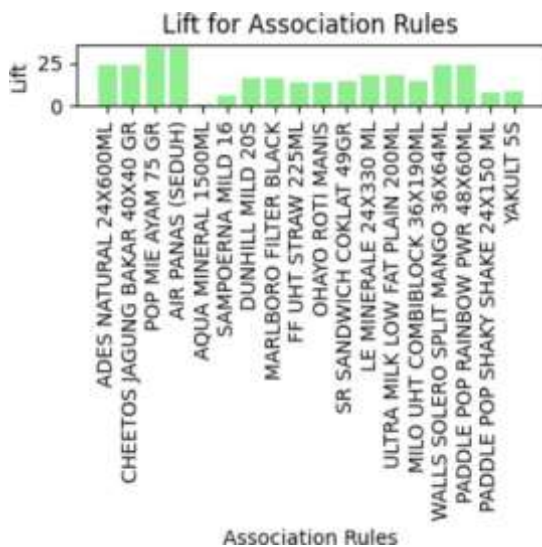


Figure 8. Lift Graph of Association Rule Results

Association rules with high Lift indicate that the purchase of one product increases the likelihood of purchasing another. This can provide valuable insights for optimizing product placement in the store or creating sales bundles. This graph provides a deeper understanding of customer purchasing behavior and can guide strategic decisions to enhance product sales and marketing. However, it's important to note the limitations of these findings. As preliminary guidance, these results offer opportunities for Minang Mart to improve cross-product sales and adjust store layouts. Yet, for a more comprehensive understanding of customer purchasing patterns within a dynamic retail context, further research might be necessary. In conclusion, these findings indicate potential for Minang Mart to enhance their marketing strategies by comprehensively understanding customer purchasing patterns.

## CONCLUSION

In this study, the findings from the association analysis provide a strong insight into customer purchasing patterns at Minang Mart. With a deeper understanding of the relationships between products, Minang Mart has the potential to enhance their marketing strategies. Better arrangements in product placements, such as CHEETOS JAGUNG BAKAR 40X40 GR and ADES NATURAL 24X600ML, as well as tailored promotions for product combinations like POP MIE AYAM 75 GR and HOT WATER (SEDUH), could be effective strategic steps.

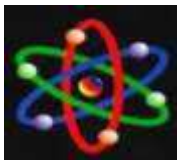
## ACKNOWLEDGMENT

I want to express my deepest gratitude to Adzka University for their invaluable financial support throughout this research. Their funding has been a crucial catalyst in conducting this study, and I truly appreciate the trust and opportunity they have given me. I hope that the results of this research will contribute meaningfully to the world of science and society. Thank you.

## BIBLIOGRAPHY

- [1] R. Alfianzah, R. I. Handayani, and M. Murniyati, "Implementation of Apriori Algorithm Data Mining for Increase Sales," *Sinkron*, vol. 5, no. 1, p. 17, 2020, doi: 10.33395/sinkron.v5i1.10587.
- [2] Lismardiana and D. E. B. N. Herman Mawengkang, "Pengembangan Algoritma Apriori





- Untuk Pengambilan Keputusan,” *Keputusan Dev. Apriori Algorithm Decis.*, vol. 4, no. 2, pp. 110–121, 2019, [Online]. Available: <https://jurnal.kominfo.go.id/index.php/jtik/article/view/823>.
- [3] R. Husna, R. Lestari, and Y. Hendra, “Inventory model of goods availability with apriori algorithm,” *J. Phys. Conf. Ser.*, vol. 1317, no. 1, 2019, doi: 10.1088/1742-6596/1317/1/012019.
- [4] Ismai, “Penerapan Algoritma Apriori Pada Penjualan Kopi Arabica,” *Simmastek Uisu*, no. Dm, pp. 118–124, 2021.
- [5] E. Elisa, “Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 2, pp. 472–478, 2018, doi: 10.29207/resti.v2i2.280.
- [6] N. N. Merliani, N. I. Khoerida, N. T. Widiawati, L. A. Triana, and P. Subarkah, “Penerapan Algoritma Apriori Pada Transaksi Penjualan Untuk Rekomendasi Menu Makanan Dan Minuman,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 8, no. 1, pp. 9–16, 2022, doi: 10.25077/teknosi.v8i1.2022.9-16.
- [7] P. A. Permatasari, Linawati, and L. Jasa, “Analysis of ShoppingCart in Retail Companies Using Apriori Algorithm Method and Model Profset,” *Int. J. Eng. Emerg. Technol.*, vol. 5, no. 2, pp. 52–60, 2020.
- [8] Y. Fang, R. Wang, M. Guo, and Y. Hou, “Product bundling for online supermarkets by frequent itemset mining and optimization approach,” *Procedia Comput. Sci.*, vol. 207, pp. 4434–4441, 2022, doi: 10.1016/j.procs.2022.09.507.
- [9] E. Alma, E. Utami, and F. Wahyu Wibowo, “Implementasi Algoritma Apriori untuk Rekomendasi Produk pada Toko Online Implementation of Apriori Algorithms for Product Recommendations at Online Stores,” *Citec J.*, vol. 7, no. 1, 2020.
- [10] U. N. Kumalasari, K. Gede, D. Putra, I. Putu, and A. Dharmadi, “Implementasi Algoritma Apriori untuk Menemukan Pola Pembelian Konsumen pada Perusahaan Retail,” *JITTER-Jurnal Ilm. Teknol. dan Komput.*, vol. 1, no. 2, 2020.
- [11] A. Silvanie, “Pencarian Frequent Itemset Dengan Algoritma Apriori Dan Python,” *J. Nas. Inform.*, vol. 1, No. 2, no. 2, pp. 103–113, 2020.
- [12] H. Kusumo, E. Sedyono, and M. Marwata, “Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi,” *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 49, 2019, doi: 10.21580/wjit.2019.1.1.4000.
- [13] M. Krishnamoorthy and R. Karthikeyan, “Pattern mining algorithms for data streams using itemset,” *Meas. Sensors*, vol. 24, no.





June, p. 100421, 2022, doi:  
10.1016/j.measen.2022.100421.

