Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Sales Association Analysis at the Donatkoe Factory Store Which is Upgraded using the Fp-Growth Algorithm

Dwina Aurelia Agustina[1*], Nining Rahaningsih[2], Raditya Danar Dana[3], Cep Lukman Rohmat[4]

[1,2,3,4] STMIK IKMI Cirebon
dwinaaurelia410@gmail.com[1*]

Abstract

In the retail industry, especially food and beverage, understanding customer buying patterns is crucial for effective stock management and marketing strategies. Donatkoe Factory stores faced challenges in identifying items that were frequently purchased at the same time, which often led to operational inefficiencies and lowered profitability. Association analysis is needed to uncover purchasing patterns to support data-driven decision-making. This study uses the FP-Growth algorithm to analyze transaction data at Donatkoe Factory stores. The parameters used are support, confidence, and elevator to evaluate the strength of the relationship between items. Transaction datasets are analyzed to find combinations of products that are frequently purchased together. The results of the analysis showed several product combinations with strong associations, such as Donuts with Pizza (confidence 0.814; elevator 1.031) and Donuts with Fruit Salad (confidence 0.821; elevator 1,039). The combination with the highest confidence was Donuts with Pizza, Fruit Salad, and Buko Pandan (confidence 0.842; elevator 1,066). These findings indicate that the FP-Growth algorithm is effective in identifying relationships between items, so it can support marketing strategies such as adjacent product placements, bundling, or special promotions. The results of this study also provide insights for Donatkoe Factory stores to improve operational efficiency and customer satisfaction through data-driven decisions.

Keywords: *FP-Growth Algorithm, Purchasing Patterns, Association Rules, Donatkoe Factory Store, Data Mining.*

1. Introduction

Rapid developments in the field of Informatics have had a significant impact on various aspects of life, including technology, business, and education. In today's digital era, the use of information technology is very important to improve operational efficiency and effectiveness in various sectors. One of the important applications of information technology is in data analysis, which allows for better decision-making based on accurate and relevant information. In a business context, sales data analysis can provide valuable insights into consumer behavior and market trends, which in turn can improve marketing and sales strategies. Therefore, this study focuses on the application of the FP-Growth algorithm to improve the analysis of sales associations in Donatkoe Factory stores, which is expected to make a significant contribution to data-driven business management.

Although many previous studies have used the Apriori algorithm for the analysis of purchasing patterns, there is still a gap in the application of the FP-Growth algorithm, especially in the context of the food and beverage industry. Previous research has shown that many companies still struggle to manage their inventory and understand consumer buying patterns effectively[1]. This can lead to a shortage or excess of stock, which in turn negatively impacts the company's profitability. In addition, with increasing competition in the market, it is important for companies to have a data-driven marketing strategy. Therefore, this study aims to fill the gap by applying the FP-Growth algorithm to analyze sales data at Donatkoe Factory Stores and provide more precise recommendations for stock management and marketing strategies. In the previous study, there were various studies that applied the FP-Growth algorithm in sales data analysis [2]. shows that the FP-Growth algorithm can be used to analyze sales data in Toko Berkah, which helps management in stock management and promotion, although this study does not discuss in depth the influence of the results of the analysis on the broader marketing strategy. Furthermore, it applies the FP-Growth algorithm to grocery sales data, emphasizing the importance of identifying purchasing patterns for inventory management, but does not explore how these patterns can be integrated into more strategic business decisions [3]. [4] also examined the application of the FP-Growth algorithm in drug sales, showing that this analysis can improve the efficiency of marketing strategies and provide valuable insights to support marketing strategy decisions, although it does not discuss the link between purchasing patterns and consumer behavior in more depth. These findings show that although the FP-Growth algorithm has been applied in a variety of contexts, there is still a gap in understanding how the results of the analysis can directly influence strategic and operational decisions at the managerial level, which is the main focus of this study.

The main goal of this study is to improve the analysis of sales associations in Donatkoe Factory Stores by applying the FP-Growth algorithm. This study aims to identify significant purchasing patterns, which can assist management in decision-making related to stock of goods and marketing strategies. Thus, this research is expected to fill the knowledge gap in the application of the FP-Growth algorithm in the retail sector, especially in the context of food sales.

The method that will be used in this study is a data mining approach with the FP-Growth algorithm. The research process began with the collection of sales transaction data from the Donatkoe Factory Store during the period of October 2024. Once the data is collected, the next step is to preprocess to ensure the quality of the data to be analyzed. Then, the FP-Growth algorithm will be applied to identify frequent buying patterns. The results of this analysis will be evaluated and used to design more effective marketing strategies. This approach is in line with previous research that showed the effectiveness of the FP-Growth algorithm in association analysis [5].
If the objectives of this research are achieved, the results obtained will provide new insights into consumer purchasing patterns at the Donatkoe Factory Store. These findings can be used by practitioners to design more effective marketing strategies, such as strategic product placement and targeted promotions. In addition, the results of this study can also be a reference for other researchers who are interested in the field of sales data analysis and the application of the FP-Growth algorithm. Thus, this research not only contributes to the development of science in the field of Informatics, but also has a positive impact on the development of technology in data analysis [6].

2. Theoretical Foundations

2.1. FP-Growth

The FP-Growth algorithm is a development of the Apriori algorithm. The FP-Growth algorithm is one of the alternative algorithms that can be used to determine the most frequently occurring datasets in a dataset [7].

2.2. Knowledge Discovery Database

Knowledge Discovery in Databases (KDD) is a sub-field of data mining. The Knowledge Discovery in Database (KDD) process is a step in data analysis that involves several steps/stages. Knowledge discovery in database (KDD) is one of the most important procedures and projects that must be carried out to collect data and identify historical data in order to obtain a pattern contained in the relationships between data in a large database [8].

3. Contents

3.1. Selection

The initial stage is data selection, which is the stage where analysis is carried out on relevant data. The data used comes from cake sales transactions at the Donatkoe Factory Shop.
Uses the Excel Read operator, which serves to read Excel files as input. This component has two connection ports: the "file" port as the input, which is used to receive the Excel file, and the "out" port as the output, which provides the data that has been read from the Excel file. The results of the data reading process using the read excel operator can be seen in the table below.
shows transaction data in the form of a binary matrix, where each column represents an available item (such as Donuts, Banana Cake, Pizza, etc.), and each row represents a specific transaction. The number "1" in the cell indicates that the item was purchased in the transaction, while the number "0" indicates that the item was not purchased. For example, in the first transaction (Transaction ID 1), Donuts, Pizza, Cromboloni, Buko Pandan, and Chiffon Koe.

Table 1: Sales Transaction Data

Transaction ID	Donat	Banana cake	Pizza	Cromboloni	Milk Bun	Buko Pandan	Fruit Salad	Brownies Block	Koe bread	Chiffon koe
1	1	0	1	1	0	1	0	0	0	1
2	1	1	0	0	0	0	1	0	1	0
3	1	0	1	1	1	1	1	0	0	1
4	1	1	1	1	0	1	1	1	1	1
5	1	0	1	1	1	0	0	0	0	0
6	0	1	0	0	0	1	1	1	1	1
7	1	0	1	1	1	1	0	1	1	1
8	1	1	1	0	0	0	1	0	1	0
9	1	1	0	1	0	1	1	1	0	1
10	1	0	1	1	1	1	1	0	1	0
11	0	1	0	0	0	0	1	1	0	0
12	1	1	1	1	1	1	0	0	1	1
13	1	0	0	0	0	0	1	1	1	1
14	1	1	1	1	0	1	1	0	0	0
15	1	1	1	1	1	0	0	0	1	1
16	0	0	0	1	0	1	1	1	0	0
17	1	1	1	0	0	1	0	0	1	1
18	1	1	1	0	1	1	1	1	1	1
19	1	0	0	1	1	0	0	0	1	0
20	1	1	1	0	0	1	1	0	0	1
21	0	0	1	1	1	0	0	1	1	1
22	1	1	0	0	0	1	1	0	0	1
23	1	0	1	1	1	1	1	1	1	0
24	1	1	0	0	0	1	0	0	0	1

25	0	1	1	1	1	0	1	1	1	0
26	1	0	1	0	0	1	1	0	0	0
27	0	1	0	1	1	0	0	1	1	1
28	1	1	1	1	1	1	1	0	1	0
29	1	0	1	0	0	1	0	0	1	1
30	1	0	0	1	1	1	1	1	0	1
31	1	1	1	1	1	0	0	0	1	1

3.2. Preprocessing

The preprocessing process began by entering a dataset consisting of 404 transaction data collected from September to October 2024. At this stage, the Filter Example operator will be applied to filter the data on cake sales transaction data from the Donatkoe Factory Store based on the no missing attributes condition, that is, data that does not have missing values will be deleted from the dataset. Using the Examples Filter parameter in the RapidMiner software. This operator is used in the preprocessing stage to filter transaction data based on attributes that do not have empty values, such as deleting empty or incomplete data[9].

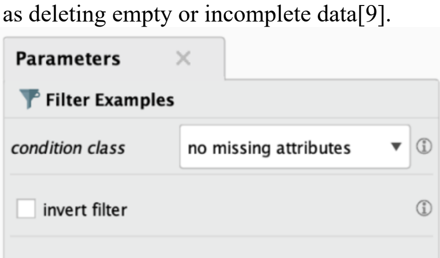

Fig. 1: Parameter Operator Filter Examples.

In this section, users can configure operator behavior to filter data as needed. There is a condition class option set to "no missing attributes" which means that only data without empty values will be selected. In addition, there is an invert filter option that, if enabled, will reverse the filter condition (selecting data with an empty value instead of a complete one). This setting ensures that only valid data is passed to the next process in the analysis. In the next stage of the preprocessing process, the Remove Duplicates step is performed to ensure that each transaction in the dataset is unique. This stage aims to remove duplicate transactions, i.e. transactions with the exact same value in each attribute. It uses the Remove Duplicates operator parameter, which serves to identify and remove duplicate values from the Donatkoe Factory Sales Transaction Data dataset. This operator ensures that only one copy of the repetitive data is stored, improving the quality and reliability of the data for subsequent analysis.

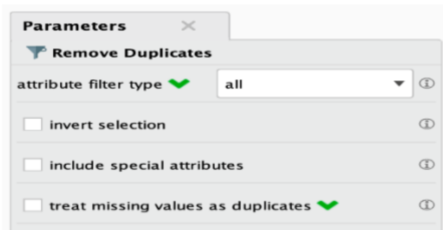

Fig. 2: Operator Parameter Remove Duplicate.

shows the Remove Duplicates operator parameter, the selection of the attribute filter type option with the value all is done to ensure that all attributes in the dataset are thoroughly checked to detect duplicates. This way, each attribute will be comprehensively compared, so that any entries identified as duplicates can be accurately removed. This approach was chosen to prevent missing potential duplication due to the limitation of the inspected attributes, resulting in a more valid and reliable dataset for further analysis.

3.3. Transformation

The next stage is the transformation with Numerical to Binominals and Remap Binominals operators. The Numerical to Binominals operator functions to convert numerical data into binomial data (two categories), while the Remap Binominals operator functions to remap the binomial categories to suit the needs of further analysis or processing.

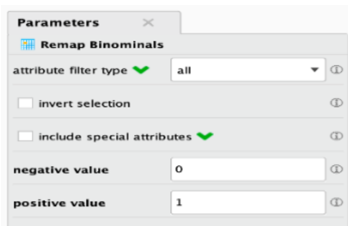

Fig. 3: Parameter Remap Binominals.

The results of data transformation using Numerical to Binominals and Remap Binominals operators. Each column represents a type of food, such as Donuts, Banana Cake, Pizza, Crombolloni, Milk Bun, Buko Pandan, Fruit Salad, Brownies, Roti Koe, and Chiffon Koe. Each row represents a transaction, where a value of true indicates that the food item was purchased, while false indicates that the item was not purchased. This transformation helps in converting numerical data into a binominal form, making it easier to analyze purchasing patterns or relationships between products in transactions.

Donat	Banana cake	Pizza	Cromboloni	Milk Bun	Buko Pand...	Salad Buah	Brownies S...	Roti koe	Chiffon koe
true	false	true	true	false	true	false	false	false	true
true	true	false	false	false	false	true	false	true	false
true	false	true	true	true	true	true	false	false	true
true	true	true	true	false	true	true	true	true	true
true	false	true	true	true	false	false	false	false	false
false	true	false	false	false	true	true	true	true	true
true	false	true	true	true	true	false	true	true	true
true	true	true	false	false	false	true	false	true	false
true	true	false	true	false	true	true	true	false	true
true	false	true	true	true	true	true	false	true	false
false	true	false	false	false	false	true	true	false	false

Fig. 4: Results of the transformation of Donatkoe Factory sales transactions.

3.4. Data Mining

At this stage, we can identify or find support and confidence values to determine sales patterns at Donatkoe Factory Stores by using the FP-Growt algorithm [10]. The implementation of the FP-Growth algorithm to find clothing sales patterns at the Donatkoe Factory Store can be seen in the following steps:

Fig. 5: Operator Algorithm FP-Growth.

FP-Growth operator used to analyze item frequency patterns in transaction data. In figure 4.12, there is a shopping cart icon that represents the transaction data. The texts "exa" and "fre" may refer to some process or data that is fed into and out of the FP-Growth algorithm, perhaps related to the example or frequency parameter in the context of data processing. In this study, the support metric will be set or set as much as 0.7 or 70%.

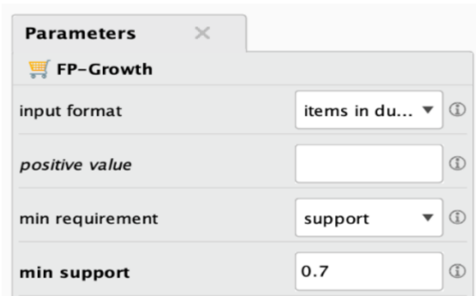

Fig. 6: Parameter Operator FP-Growth.

parameter configuration on the FP-Growth operator used in this study. In this configuration, there are some important parameters that can be adjusted such as the mean support set at 0.7 or 70%, which indicates that an item or combination of items must appear in at least 70% of the total transaction to be considered significant. This setting ensures that the FP-Growth algorithm will only find patterns that have a high frequency of occurrence in the dataset used.

To complement the frequency pattern analysis with the FP-Growth algorithm, the Create Association Rules operator is also used after the pattern search process is complete. The Create Association Rules operator can be seen in the image below:

Fig. 7: Operator Association Rules.

The Create Association Rules operator functions to generate associative rules from the frequency patterns found by FP-Growth. In this study, the confidence parameter in the Create Association Rules operator is set to a value of 0.8 or 80%. This setting ensures that the resulting rules have a high level of confidence, i.e. at least 80% of transactions that contain antecedent items also contain consequent items. Thus, the combination of the FP-Growth algorithm and the Create Association Rules operator allows the study to identify patterns and rules with a high level of relevance as well as the strength of significant relationships in the analyzed transaction data.

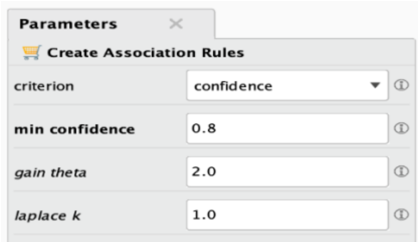

Fig. 8: Parameter Operator Create Association Rules.

The Create Association Rules operator used in this study to generate associative rules from the patterns found. In this configuration, several important parameters have been set, including a criterion set to "confidence," which means that the rules will be filtered based on confidence level. Furthermore, the mean confidence is set at 0.8 or 80%, so the rule will only be generated if it has a confidence level of at least 80%, ensuring that the resulting rule has high relevance. The theta gain parameter is set at a value of 2.0, which may be used to control the weight or relevance of the rules found based on additional settings in the algorithm. Meanwhile, the laplace k is set at a value of 1.0, which is likely to be used to smooth out calculations or reduce reliance on scarce data. With this setup, the Create Association Rules operator is optimized to generate associative rules that have a high level of confidence and significant relevance in pattern analysis on transaction data.

3.5. Evaluation/Interpretation

In this evaluation and interpretation stage, the analysis will focus on the pattern of product purchases in Donatkoe Factory stores to answer two main objectives. First, an analysis of purchasing patterns was carried out using the FP-Growth algorithm to identify items that are often purchased together. Based on the resulting Association Rules, researchers can identify items that are frequently purchased at the same time (frequent itemsets) by looking at the combination of items that appear in the association rules. Each association rule describes the relationship between two sets of items, where if one set of items is purchased, then the other items tend to be purchased together. In this case, we focus on the combination of items that appear as the premise of the rule and look at the confidence level to measure the strength of the association. The results of the association can be seen in the image below.

AssociationRules

```
Association Rules
[Pizza, Milk Bun] --> [Donat] (confidence: 0.800)
[Buko Pandan, Cromboloni] --> [Donat] (confidence: 0.800)
[Donat, Salad Buah, Milk Bun] --> [Pizza] (confidence: 0.800)
[Salad Buah] --> [Donat] (confidence: 0.803)
[Buko Pandan, Roti koe] --> [Donat] (confidence: 0.805)
[Pizza, Roti koe] --> [Donat] (confidence: 0.809)
[Pizza, Salad Buah, Roti koe] --> [Donat] (confidence: 0.810)
[Buko Pandan] --> [Donat] (confidence: 0.813)
[Pizza] --> [Donat] (confidence: 0.814)
[Pizza, Salad Buah, Milk Bun] --> [Donat] (confidence: 0.814)
[Pizza, Banana cake] --> [Donat] (confidence: 0.816)
[Pizza, Salad Buah] --> [Donat] (confidence: 0.821)
[Salad Buah, Buko Pandan] --> [Donat] (confidence: 0.832)
[Pizza, Buko Pandan] --> [Donat] (confidence: 0.833)
[Pizza, Salad Buah, Buko Pandan] --> [Donat] (confidence: 0.842)
```

Fig. 9: Association Rules Result

The following are the items that are often purchased at the same time based on the Association Rules shown in the figure:
Pizza, Milk Bun --> Donat (confidence: 0.800)
The combination of Pizza and Milk Bun is often purchased along with Donuts.
Buko Pandan, Cromboloni --> Donuts (confidence: 0.800)
Buko Pandan and Cromboloni are often purchased along with Donuts.
Donat, Salad Buah, Milk Bun --> Pizza (confidence: 0.800)
Combinations of Donuts, Fruit Salads, and Milk Buns are often purchased along with Pizza.
Salad Buah --> Donat (confidence: 0.803)
Fruit Salad is often purchased along with Donuts.
Buko During, Roti koe --> Donat (confidence: 0.805)
The combination of Buko Pandan and Roti koe is often purchased along with Donuts.
Pizza, Roti koe --> Donat (confidence: 0.809)
Pizza and Roti koe are often purchased along with Donuts.
Pizza, Salad Buah, Roti koe --> Donat (confidence: 0.810)
Combinations of Pizza, Fruit Salad, and Bread koe are often purchased along with Donuts.
Buko Pandan --> Donat (confidence: 0.813)
Buko Pandan is often purchased along with donuts.
Pizza --> Donat (confidence: 0.814)
Pizza is often purchased along with Donuts.
Pizza, Salad Buah, Milk Bun --> Donat (confidence: 0.814)
Pizza Combinations, Fruit Salads, and Milk Buns are often purchased along with Donuts.
Pizza, Banana cake --> Donat (confidence: 0.816)
Pizza and Banana cake are often purchased along with Donuts.
Pizza, Salad Buah --> Donat (confidence: 0.821)
Pizza and Fruit Salad are often purchased along with Donuts.
Fruit Salad, Buko Pandan --> Donuts (confidence: 0.832)

The combination of Fruit Salad and Buko Pandan is often purchased along with Donuts.
Pizza, Buko Pandan --> Donat (confidence: 0.833)
The combination of Pizza and Buko Pandan is often purchased along with Donuts.
Pizza, Fruit Salad, Buko Pandan --> Donuts (confidence: 0.842)
Pizza Combinations, Fruit Salads, and Buko Pandan are often purchased along with Donuts.
Items that are often purchased together.
Using high confidence, we can conclude that combinations of items such as Pizza, Fruit Salad, Buko Pandan, and Donuts are often bought together. These rules can be used to understand customer buying patterns and design more effective marketing strategies, such as placing relevant products together.

4. Conclusion

Based on research on product purchase patterns at Donatkoe Factory, the FP-Growth algorithm has proven to be effective in identifying customer purchase patterns. The results of the analysis show a strong relationship between Donuts as a primary product and various other products, providing important insights to support more precise marketing strategies. Here are two key takeaways from the study:
This study uses the FP-Growth algorithm to analyze transaction data at Donatkoe Factory. The results of the analysis showed a significant purchase pattern, especially the relationship between Donuts as the main product and various other products, such as Pizza, Buko Pandan, Fruit Salad, and Roti Koe. The FP-Growth algorithm successfully identifies product combinations that are often purchased along with donuts. For example, the combination of Donuts, Pizza, and Roti Koe has a confidence of 0.809, while the combination of Donuts, Pizza, Fruit Salad, and Buko Pandan has a confidence of 0.842. These findings provide important insights to improve marketing strategies and product management.

5. Suggestions

Based on the results of the research that has been carried out, some suggestions that can be given for Donatkoe Factory are as follows:
Given the buying patterns found between products such as Pizza, Fruit Salad, Buko Pandan, and Donuts, it is recommended to place these products in close proximity to each other in the store. A more strategic arrangement can make it easier for customers to find the products they are looking for, as well as increase the chances of selling products that are often purchased together. To increase sales, stores can consider offering product bundling promotions based on the associations found, such as bundling Pizza with Donuts, or Buko Pandan with Fruit Salad.

References

[1] A. Rizky Mangunsong, V. Sihombing, and I. Rasyid Munthe, "Development of a Product Recommendation System Based on Purchase Patterns with a Priori Algorithm Approach," *J. Computation. and Sist. Inf.*Vol. 7 No. 1 pp. 82–86, 2024, doi:10.55338/jicomsi.v7i1.2718.

[2] S. Herdyansyah, E. H. Hermaliani, L. Kurniawati, and S. R. Sri Rahayu, "Analysis of the Association Rule Method Using the FP-Growth Algorithm on Sales Data (Case Study of Toko Berkah)," *J. Khatulistiwa Inform.*, vol. 8, no. 2, pp. 127–133, 2020, doi: 10.31294/jki.v8i2.9277.

[3] Y. Syach Putra, R. Kurniawan, and Y. Arie Wijaya, "Application of Data Mining Using the Fp-Growth Algorithm on Basic Food Sales Data," *JATI (Journal of Mhs. Tek. Inform.*Vol. 8 No. 1 pp. 561–567, 2024, doi:10.36040/caste.V8I1.8391.

[4] K. Kharomiyah, N. Rahaningsih, and R. D. Dana, "Analysis of Drug Sales Linkage through the Application of FP-Growth Algorithm to Optimize Marketing Strategies," *J. SAINTIKOM (Journal of Science Manaj. Inform. and Computers)*, Vol. 23 No. 1 p. 57, 2024, doi: 10.53513/j.v23i1.9495.

[5] F. Achmad, O. Nurdiawan, and Y. Arie Wijaya, "Analysis of Consumer Purchase Transaction Patterns in Health Retail Stores Using the FP-Growth Algorithm," *JATI (Journal of Mhs. Tek. Inform.*, Vol. 7 No. 1 pp. 168–175, 2023, doi:10.36040/caste.v7I1.6210.

[6] A. R. Wibowo and A. Jananto, "Implementation of FP-Growth Algorithm Association Method Data Mining in Retail Companies," *Inspir. J. Technol. Inf. and Commune.*, vol. 10, no. 2, p. 200, 2020, doi:10.35585/inspir.v10i2.2585.

[7] P. Palupiningsih and B. Prayitno, "Implementation of the Fp-Growth Algorithm for Determining MSME Product Recommendations based on Purchase Frequency," *J. Teknoinfo*, vol. 17, pp. 493–501, 2023.

[8] E. Nurarofah, R. Herdiana, and N. Dienwati Nuris, "Application of Associations Using the Fp-Growth Algorithm on Sales Transaction Patterns in Bakery Stores," *JATI (Journal of Mhs. Tek. Inform.*, Vol. 7 No. 1 pp. 353–359, 2023, doi:10.36040/caste.v7I1.6299.

[9] N. Salsabila, N. Sulistiyowati, and T. Nur Padilah, "Search for Drug Use Patterns Using FP-Growth Algorithm," *J. Appl. Informatics Comput.*, vol. 6, pp. 120–128, Aug. 2022, doi: 10.30871/jaic.v6i2.4187.

[10] I. Iddrus, D. Wulansari, and J. Dharmawan, "Application of Data Mining in Finding Association Patterns of Data Tracer Study Using the Fp-Growth Algorithm (Case Study of Upt. Career Service Center of Wiraraja University)," *J. Tek.*, vol. 16, no. 1, pp. 1–6, 2024, doi:10.30736/jt.v16i1.1173.