Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

Association Analysis of Printing and Photocopying Sales Data in Adzmi Art Shop Cirebon Uses the FP-Growth Algorithm

Suteja[1]*, Rudi Kurniawan[2], Yudhistira Arie Wijaya[3]

[1,2,3]*STMIK IKMI Cirebon*

Suteja045@gmail.com[1*], rudi226@gmail.com[2], yudhistira010471@gmail.com[3]

Abstract

In the digital era, transaction data analysis plays a crucial role in strategic decision-making, especially for SMEs such as Toko Adzmi Art in Cirebon Regency. This study aims to develop a sales data association model using the FP-Growth algorithm to identify product association patterns. Daily transaction data over a year were collected, processed through data cleaning, standardization, and transformation, and analyzed using RapidMiner software. Minimum support and confidence parameters were applied to evaluate the frequency and strength of product relationships. The results show that the combination of "Photocopy" and "Passport Photo" services has a confidence of 0.491 and a support of 0.061, with "Photocopy" as the most in-demand product (support 0.497). These findings open opportunities for bundling strategies and inventory optimization to enhance operational efficiency. This model provides an empirical foundation for SMEs to leverage data mining technology to improve competitiveness and customer satisfaction.

Keywords: *FP-Growth algorithm, data mining, product association, sales, printing, photocopying.*

1. Introduction

Rapid progress in the field of informatics has brought significant changes in various aspects of life, including the business and technology sectors. Data mining technology, especially in the trade sector, offers great opportunities for business actors to optimize transaction data management while supporting data-based strategic decision making[1]. One approach commonly used in transaction data analysis is the FP-Growth association algorithm, which is known to be effective in identifying association patterns from large data[2].In the context of the printing and photocopying business, such as at the Adzmi Art Shop, the application of this algorithm is very useful for analyzing customer behavior and identifying products with high levels of demand[3]. Such analytics not only help improve inventory management efficiency but also contribute to increased profitability through data-driven business strategies[4].

As the volume of transaction data generated by businesses increases, the challenges of managing and analyzing that data become increasingly complex. Transaction data that is not managed properly often becomes just a pile of information that does not provide added value for strategic decision making[5]. In the context of the Adzmi Art Store in Cirebon Regency, one of the challenges faced is a lack of understanding of customer purchasing patterns, especially regarding products that are often purchased simultaneously[6].
This kind of information is very important to help stores set priorities in inventory management and design effective marketing strategies[7]. Although various data analysis methods are available, small and medium enterprises (SMEs) often experience difficulties in adopting these technologies, mainly due to limited skilled human resources, financial constraints, and lack of adequate technological infrastructure[8].

These problems hamper the potential for businesses to utilize transaction data as a basis for developing more effective policies. Therefore, a data analysis approach is needed that is not only efficient but also easy to implement by SMEs[9]. One solution that can answer this need is the application of the FP-Growth algorithm. This algorithm is designed to efficiently identify association patterns or relationships between products in transaction data[10].

One of the methods used to support the implementation of the FP-Growth algorithm is Knowledge Discovery in Database (KDD). KDD is a systematic process that consists of several stages, namely data selection, data cleaning, data transformation, data analysis, and results evaluation. This method ensures that the data used is relevant, valid and ready for further processing. In this research, the KDD approach is used to identify relevant patterns in Adzmi Art Store transaction data, which in turn can support inventory management and data-based promotional strategies [11].

By utilizing the FP-Growth algorithm, Adzmi Art Store can recognize product combinations that are often purchased together by customers, so that it can support inventory management and more targeted promotional strategies[12]. This research aims to develop an association model that is applicable and relevant to the needs of the printing and photocopying business. This model is expected to make a real contribution in increasing operational efficiency and competitiveness of the Adzmi Art Store amidst increasingly tight business competition[13].

2. Research Method

This research method uses an experimental method with a Knowledge Discovery In Database (KDD) approach process, which is a quantitative research method. The experimental method aims to identify relationships between variables by manipulating the available data. In this research, the experimental method was applied to analyze purchasing patterns for printing and photocopying products at the Adzmi Art Shop, Cirebon Regency. This pattern will be analyzed using the FP-Growth algorithm with a focus on support, confidence and lift ratio values to find product associations that are most often purchased together.

2.1. Data Analisys Techniques

This research uses data analysis techniques to extract information from Adzmi Art Store transaction data to identify product purchasing patterns. The Knowledge Discovery in Database (KDD) approach is applied because it is capable of processing complex and large data, through systematic stages starting from data selection to evaluation of results. The FP-Growth algorithm is used to discover product association patterns, generating insights that are useful for inventory management and marketing strategies for adzmi art stores.

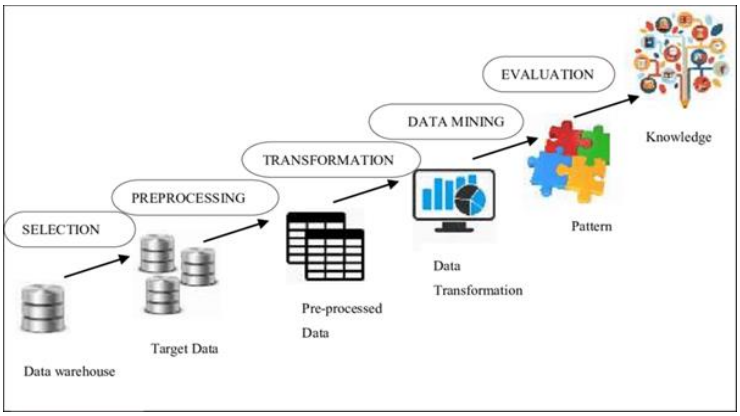

Fig. 1: Knowledge Discovery In Database[11]

A. Selection
The initial stage of data analysis is selecting data that is relevant to the research objective, namely identifying product purchasing patterns at the Adzmi Art Store. Data was taken from the store's digital transaction recording system, covering 2,084 transactions with 10 attributes, during the period September 2023 to October 2024. This process was carried out carefully to ensure the data used was valid and supported analysis of purchasing patterns for products that are often purchased together.

B. Preprocessing
At this stage, the data collected is processed through a cleaning process to ensure its quality and suitability, including removing duplicate data, eliminating incomplete data, and standardizing the format so that the data is ready to be analyzed with software. These steps are taken to ensure the accuracy and consistency of the data, so that the analysis results can support the identification of optimal purchasing patterns.

C. Transformation
The transformation stage aims to change the raw data into a format that is suitable for the FP-Growth algorithm. This process includes converting transaction data into itemset table format, grouping products purchased together in each transaction for association analysis, as well as adjusting the data structure to be compatible with the analysis software. This transformation ensures the data is ready to be used to efficiently identify purchasing patterns.

D. Data Mining
The main stage in the KDD process is the application of the FP-Growth algorithm to identify relationship patterns between products. This process begins with determining minimum parameters, namely determining support and confidence values for significant patterns. Next, the FP-Growth algorithm is applied to find product combinations that are frequently purchased together. The result is the generation of association patterns, which show relationships between products based on support, lift ratio and confidence values.

E. Evaluation
The results of the FP-Growth analysis were evaluated to ensure the patterns found were in accordance with the research objectives. Based on the evaluation of the support, confidence and lift ratio values, it is recommended to offer bundling with discounts on Photocopy and Photocopy services which are often used together, as well as discounts for Regular Color Print services if purchased together with Photocopy. In addition, stores can develop

loyalty programs by awarding points for use of Photocopy services, which can be exchanged for discounts or additional services.

3. Result and Discussion

This research uses transaction data from the Adzmi Art Store in Cirebon Regency as a source and applies the FP-Growth algorithm to analyze sales patterns. As a guide, this research follows the Knowledge Discovery in Databases (KDD) stages to ensure a systematic and structured process. Through this stage, the collected data is processed to produce insights that are useful for inventory management and marketing strategies. This is the result of each step taken:

3.1. Selection

The first stage, namely Selection, is an important initial step to ensure that the data used is accurate and in accordance with the analysis objectives. The data used in this research comes from daily transactions at the Adzmi Art Shop in Cirebon Regency. This data contains information about the types of products sold and the frequency of sales in a certain time period. This information is very important for analyzing purchasing patterns and identifying products that are often purchased together. The dataset used is an Excel file entered into the retrieve operator in Rapidminer with a total of 2,084 transaction data records, consisting of 10 default attributes.

Table 1: Result of the retieve operator

No.	Atribut	Type Data	Keterangan
1.	No. Transaksi	Integer	**Nomer transaksi**
2.	Tgl. Transaksi	Nominal	**Tanggal transaksi**
3.	Nama Pelanggan	Nominal	**Nama pelanggan yang membeli**
4.	Metode Bayar	Nominal	**Metode pembayaran yang digunakan**
5.	Nama Barang	Nominal	**Nama barang**
6.	Harga Modal	Nominal	**Harga modal**
7.	Type Harga Jual	Nominal	**Harga satuan**
8.	Harga Jual	Integer	**Harga jual**
9.	Qty	Integer	**Jumlah produk yang dibeli**
10.	Jumlah	Integer	**Jumlah keseluruhan produk yang dibeli**

Next, the step that needs to be taken is to apply the set role operator to assign an ID to the dataset. This process aims to determine the columns that will be used as unique identifiers for each transaction, which is very important to ensure accuracy in subsequent analysis. parameter settings in the Set Role operator, as summarized in Table 2. This table explains the configuration of the column selected as ID, which functions to provide unique identification for each transaction. Determining this ID is very important to ensure that the data can be analyzed accurately in the next analysis stage.

Table 2: Parameter Setrole

No	Parameter	Isi Parameter
1	attribute name	**No.Transaksi**
2	target role	**Id**

Based on the results of applying the Set Role operator, information regarding ID settings in the dataset is reflected in Table 3. which is presented below. This table provides an explanation of the settings made to the ID column, which supports the smoothness of the subsequent data analysis process.

Table 3: Results from the setrole operator

No.	Hasil Set role	Keterangan hasil
	Record	**891**
	Special Attribute	**1**
	Reguler Atribute	**376**
	Attribute :	
1.	No.Transaksi	**Integer**
2.	count(Nama Barang)_A4 1 RIM	**Integer**
3.	count(Nama Barang)_AMPLOP 90 APS	**Integer**
4.	count(Nama Barang)_AMPLOP 90 APS BESAR	**Integer**
5.	count(Nama Barang)_AMPLOP POLOS 1 PAK 104	**Integer**
6.	count(Nama Barang)_AP 230	**Integer**
7.	count(Nama Barang)_Alat Printer	**Integer**
8.	count(Nama Barang)_Asturo Osaka	**Integer**
9.	count(Nama Barang)_BANNER	**Integer**
…..	……...	**……**
376.	count(Nama Barang)_photopaper	**Integer**

3.2. Preprocessing

The data cleaning process is carried out at the preprocessing stage to handle missing data or inconsistent values. Before this stage is carried out, a preliminary analysis is carried out to verify whether there are any attributes in the dataset that have missing values or show inconsistencies. Based on statistical analysis of the dataset, as seen in table 4, it was identified that several attributes did not contain missing values. To ensure the consistency of the dataset, each record is checked in detail, and the analysis results show that the dataset finally has consistent values.

Table 4: Result from the Replace missing values

NO	ATRIBUT	TIPE DATA	MISSING VALUE
1	NO. Transaksi	Integer	**0**
2	AP 230	Integer	**0**
3	BANNER 200*100	Integer	**0**
4	BINGKAI 3R	Integer	**0**
5	BUKU PRESTASI	Integer	**0**
6	COPY A3	Integer	**0**
7	ETONA 10	Integer	**0**
…	……	Integer	**0**
376	Photopaper	Integer	**0**

3.3. Transformation

The data transformation stage is carried out to convert raw data into a format that is more suitable for the FP-Growth algorithm. Because the data type in the dataset is numeric, while the FP-Growth association algorithm requires data in binominal format, a transformation step is needed to convert numeric data to binominal. This conversion process is carried out using the Numerical to Binominal operator, which converts numeric values into two categories. This process can be seen more clearly in the table below.

Table 5: Result from the operator Numerical to binominal

No	Name	type
1	No. Transaksi	**Integer**
2	AP 230	**Binominal**
3	BANNER 200*100	**Binominal**
4	BINGKAI 3R	**Binominal**
5	BINGKAI 12R	**Binominal**
6	BUKU PRESTASI	**Binominal**
7	COPY A3	**Binominal**
8	ETONA 10	**Binominal**
9	FOTOCOPY	**Binominal**
……		
100	UNDANGAN CUSTOM	**Binominal**

3.4. Data Mining

The data mining stage at the Adzmi Art Store, Cirebon Regency uses the FP-Growth algorithm to identify patterns of products that are often purchased together. This algorithm utilizes the FP-Growth operator and the FP-Tree data structure to calculate item combinations efficiently, especially on small to medium scale datasets. The results of the analysis produce association rules that support data-based stock management and marketing strategies. The FP-Growth operator configuration used can be seen in the following table, which describes the analysis steps.

Table 6: Parameters in the FP-growth operator

No.	Parameter	Isi
1.	input format	**items in dummy coded columns**
		Setiap item dalam kumpulan data transaksi direpresentasikan dalam kolom tersendiri, dengan nama item digunakan sebagai nama kolom. Nilai binominal True/False digunakan untuk menunjukkan keberadaan item untuk setiap transaksi, di mana True menunjukkan bahwa item tersebut ada dalam keranjang, dan False menunjukkan sebaliknya.
2.	min requirement	**Support**
	Parameter ini menyediakan dua metode berbeda untuk menentukan cutoff, menghilangkan kumpulan item yang jarang terjadi.	**Persentase kombinasi item dalam database**

To obtain association rules, you must use an operator called Create Association Rules. This operator plays a role in exploring relationships between items in the dataset, which allows identification of association patterns after the FP-Growth stage. The following is a display of the Create Association Rules operator used in this analysis process.

Table 7: Parameters in the Association Rules operator

No.	Parameter	Isi
1.	criterion	**confidence**
2.	Min confidence	**0.1**

Analysis with the FP-Growth algorithm identified eight items that are frequently purchased and two product combinations that often appear together in transactions. These findings provide insights for promotional strategies, product grouping and stock management. Stores can use this data to optimize stock and design promotions that suit customer shopping patterns, improving operational efficiency and customer satisfaction. Details of the Frequent Item Set results, including support, confidence, and lift scores, are presented in Table 8.

Table 8: The result of the FP-growth operator

frequentItemSets(Fp-Growth)				
No.	Size	Support	**Item**	**Item 2**
1.	1	0.497	**Fotocopy**	
2.	1	0.162	**Print warna biasa**	
3.	1	0.123	**Pas photo**	
4.	1	0.110	**Print warna 50%**	
5.	1	0.091	**Print Bw <10**	
6.	1	0.066	**Laminating**	
7.	2	0.061	**Fotocopy**	**Print warna biasa**
8.	2	0.061	**Fotocopy**	**Pas photo**

Association Rule analysis shows the relationship between two products that are often purchased together. This pattern can be used to increase sales and design more effective marketing strategies, such as product bundling or targeted promotions. This has the potential to increase profits and customer satisfaction. Table 9 presents product combinations with support, confidence and lift values, providing information for optimizing business strategies.

Table 9: The result of the Association rules

AssociationRules				
No.	**Premises**	**Conclusion**	**Support**	**Confidence**
1.	Print warna biasa	Fotocopy	0.061	0.375
2.	Pas photo	Fotocopy	0.061	0.491

The following diagram provides a visual representation of the analyzed data, helping to facilitate understanding of the patterns and relationships found

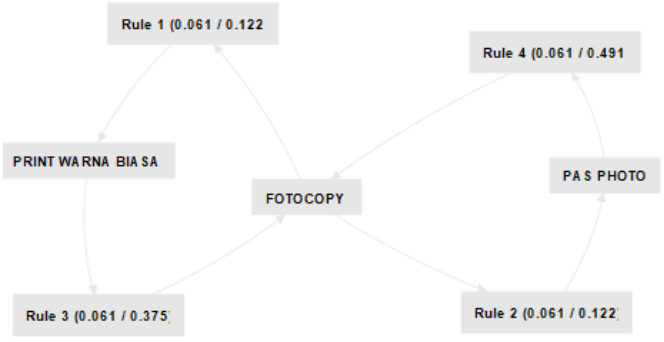

Fig. 2: visualization of association rules

3.5. Evaluation

The results of experimental activities evaluated on the dataset produced by frequent itemsets show that the FOTOCOPY product has the highest support of 0.497, which means this product is the most frequently purchased. Apart from that, the combination of FOTOCOPY and PAS FOTO products has support of 0.061 and confidence of 0.491, indicating that these two products are often purchased together. The resulting association rules provide strategic insights, such as opportunities to offer product bundling and stock management priorities. The results of this analysis are then translated into strategic recommendations to improve operational efficiency and customer satisfaction at the Adzmi Art Store.

Association Rules:

[PRINT WARNA BIASA] → [FOTOCOPY] (support: 0.061) (confidence: 0.375)
[PAS PHOTO] → [FOTOCOPY] (support: 0.061) (confidence : 0.491).

4. Conclusions and Suggestions

4.1. Conclusion

The results of this research reveal a number of important findings related to transaction data analysis at the Adzmi Art Shop, Cirebon Regency using the FP-Growth algorithm:

1. The research succeeded in identifying significant patterns of association between products. For example, the Pas Foto service has a close relationship with Fotocopy, with confidence of 0.491 and support of 0.061, indicating that these two services are often purchased together.
2. Photocopy services occupy the position as the product with the highest sales frequency, with a support value reaching 0.497. This makes this service a superior product that is most popular with customers compared to other products such as regular color printing and passport photos.
3. Through the FP-Growth algorithm, support and confidence values can be used to describe product relationships, thereby providing strategic insight in stock management and designing marketing strategies.
4. The FP-Growth algorithm has proven to be an effective tool for analyzing purchasing patterns on small to medium scale transaction data, making it relevant for small and medium enterprises (SMEs) such as Toko Adzmi Art.

4.2. Suggestions

Based on the research results and conclusions that have been obtained, the following are several recommendations and suggestions that can be proposed:

1. Optimization of Stock Management and Promotions It is recommended that the Adzmi Art Store prioritize stock management of products with high demand, such as photocopies, to ensure their availability. Apart from that, promotions in the form of bundling between Photocopy services and Photocopy services or Regular Color Print can be implemented to increase transaction volume.
2. Increasing Further Research To get a more in-depth picture, future research could consider including external factors such as seasonal influences or price trends. Additionally, other algorithms such as Apriori can be used to compare the results and effectiveness of the FP-Growth algorithm.
3. Wider Utilization of Data Mining Technology In addition to analyzing purchasing patterns, data mining technology can be used to understand customer preferences in more depth and support personalization-based marketing strategies.

References

[1] B. S. Pranata and D. P. Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service)," *Bull. Inf. Technol.*, vol. 1, no. 2, pp. 83–91, 2020.

[2] R. Wahyusari, "Penerapan Algortima FP-Growth Untuk Menemukan Pola Peminjaman Alat Pada Workshop Teknik Mesin," *Log. J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 3, pp. 406–411, 2023, [Online]. Available: https://journal.mediapublikasi.id/index.php/logic/article/view/2745%0Ahttps://journal.mediapublikasi.id/index.php/logic/article/download/2745/1211

[3] R. Aditiya, S. Defit, and G. W. Nurcahyo, "Prediksi Tingkat Ketersediaan Stock Sembako Menggunakan Algoritma FP-Growth dalam Meningkatkan Penjualan," *J. Inform. Ekon. Bisnis*, vol. 2, pp. 67–73, 2020, doi: 10.37034/infeb.v2i3.44.

[4] A. Padillah, H. Satria Tambunan, R. A. Nasution, and G. Artikel, "Penerapan Algoritma FP-Growth dalam Menganalisa Pola dan Tata Letak Bahan Makanan," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 4, pp. 2828–9099, 2022, doi: 10.55123/jomlai.v1i4.1673.

[5] V. Suryani, S. Defit, and Y. Yunus, "Analisa Data Profil Pelanggan Menggunakan Algoritma FP-Growth," *J. Inf. Teknol.*, vol. 2, pp. 7–15, 2020, doi: 10.37034/jidt.v2i1.8.

[6] F. Achmad, O. Nurdiawan, and Y. Arie Wijaya, "Analisa Pola Transaksi Pembelian Konsumen Pada Toko Ritel Kesehatan Menggunakan Algoritma Fp-Growth," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 168–175, 2023, doi: 10.36040/jati.v7i1.6210.

[7] F. Firmansyah and A. Yulianto, "Market Basket Analysis for Books Sales Promotion using FP Growth Algorithm, Case Study : Gramedia Matraman Jakarta," *J. Informatics Telecommun. Eng.*, vol. 4, no. 2, pp. 383–392, 2021, doi: 10.31289/jite.v4i2.4539.

[8] F. Mahardika, N. Alfiah, and R. Bagus Bambang Sumantri, "Penerapan Metode FP Tree dan Frequent Pattern Growth pada Penerimaan Mahasiswa Baru STMIK," *Blend Sains J. Tek.*, vol. 1, no. 3, pp. 226–234, 2023, doi: 10.56211/blendsains.v1i3.176.

[9] F. Z. Ghassani, Asep Jamaludin, and Agung Susilo Yuda Irawan, "Market Basket Analysis Using the Fp-Growth Algorithm To Determine Cross-Selling," *J. Inform. Polinema*, vol. 7, no. 4, pp. 49–54, 2022, doi: 10.33795/jip.v7i4.508.

[10] M. Yudho Ardianto and S. Adinugroho, "Penentuan Tata Letak Produk menggunakan Algoritma FP-Growth pada Toko ATK," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 9, pp. 3826–3832, 2021.

[11] & H. Z. Diaz Aji Sasongko, "Implementasi Pengolahan Data Dengan Menggunakan Algoritma Fp-Growth Untuk Menentukan Frequent Item Set Pada Penyediaan Oli Motor (Studi Kasus: Bengkel Cahaya Motor)," *J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 3, pp. 643–654, 2023, [Online]. Available: https://journal.mediapublikasi.id/index.php/logic

[12] P. Sari, L. I. Kesuma, A. F. Oklilas, and M. A. Buchari, "Simulasi Algoritma Apriori dan FP-Growth Dalam Menentukan Rekomendasi Kodefikasi Barang Pada Transaksi Persediaan," *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 1135–1152, 2024, doi: 10.33022/ijcs.v13i1.3632.

[13] D. A. Istiqomah, Yuli Astuti, and Siti Nurjanah, "Implementasi Algoritma Fp-Growth Dan Apriori Untuk Persediaan Produk," *J. Inform. Polinema*, vol. 8, no. 2, pp. 37–42, 2022, doi: 10.33795/jip.v8i2.845.