Jurnal ICT : Information and Communication Technologies, 16 (2) (2025) 145-156

Published by: Marq & Cha Institute

Jurnal ICT : Information and Communication Technologies

Journal homepage: www.ejournal.marqchainstitute.or.id/index.php/JICT

Data Mining for Drug Inventory Using Web-Based FP-Growth Method

Della Puspita[1], Wirhan Fahrozi[2]

[1,2] Sistem Informasi, Universitas Potensi Utama, Medan, Indonesia

Article Info

Article history

Received : Oct 06, 2025
Revised : Oct 22, 2025
Accepted : Oct 25, 2025

Keywords:

Association Rules;
Data Mining;
FP-Growth Algorithm;
Inventory Management;
Pharmacy System.

Abstract

Pharmacy inventory management plays a critical role in ensuring product availability and preventing financial losses due to overstock or stock shortages. However, many pharmacies, including Global Medcare Pharmacy in Medan, still experience challenges in maintaining optimal inventory levels due to the absence of data-driven management systems. This study aims to develop and implement a website-based data mining application using the FP-Growth algorithm to identify frequent itemsets and uncover association patterns within pharmaceutical sales transactions. The FP-Growth algorithm was applied to 300 transaction records to generate frequent item combinations with a minimum support threshold of 3%. The results reveal strong associations among specific drugs, such as Amlodipine 10mg, Azithromycin 500mg, and Cetirizine 10mg, with confidence levels reaching up to 100%. These findings demonstrate that FP-Growth effectively identifies purchasing patterns that can guide pharmacies in forecasting demand, managing stock levels, and designing promotional bundles. The practical implication of this research is that integrating FP-Growth into pharmacy information systems can enhance decision-making accuracy, improve service quality, and increase operational efficiency. Nevertheless, the study is limited to a single-site dataset and static analysis; future research should employ larger datasets and hybrid predictive approaches for real-time implementation across multiple pharmacy networks.

Corresponding Author:

Della Puspita
Sistem Informasi,
Universitas Potensi Utama,
Jl.K.L Yos Sudarso KM 6.5 Tj.Mulia, Medan, 20241, Indonesia
Email : dellapuspita8262@gmail.com

1. Introduction

in today's increasingly competitive pharmaceutical industry, efficient inventory management plays a vital role in ensuring profitability and sustainability. Pharmacies that are unable to anticipate demand patterns often experience either excess inventory or frequent stock-outs, both of which can negatively affect customer satisfaction and financial performance (I Made Satrya Ramayu, 2022; Tarwoto et al., 2023). Predicting the market share of best-selling medications enables managers to make data-driven ordering decisions, thereby minimizing financial losses caused by unsold drugs (Widisa Adi Kumara et al., 2024). In Medan, Global Medcare Pharmacy faces significant challenges in maintaining adequate drug stock levels, often resulting in consumer dissatisfaction and loss of market share. Similarly, Tiara Farma Pharmacy lacks an integrated inventory-control system, making it difficult to accurately monitor drug supply levels and demand fluctuations (Pertiwi & Tyoso, 2024). A website-based data-mining system presents an accessible and scalable solution for such pharmacies to manage inventory in real time and enhance operational efficiency (Wijaya et al., 2024).

From a theoretical standpoint, inventory management draws upon the principles of the Economic Order Quantity (EOQ) model, safety-stock optimization, and service-level theory (Hashad, Wah, & Alnoor, 2021). These models emphasize balancing ordering costs, holding costs, and stock-out risks to minimize total inventory costs while maintaining service quality. When combined with data analytics, these frameworks enable dynamic forecasting and reorder optimization (Hartomo, Yulianto, & Suharjo, 2020). In parallel, the Knowledge Discovery in Databases (KDD) framework provides the theoretical basis for extracting hidden patterns and associations from large datasets (Liu, 2014). Association rule mining (ARM) is one of the most widely used techniques under KDD, facilitating the discovery of relationships among frequently co-purchased items, which can significantly improve inventory decision-making (Dwiputra et al., 2023).

The FP-Growth algorithm, proposed by Han et al. (2004), is a frequent-pattern mining algorithm designed to overcome the limitations of the Apriori algorithm, particularly its computational inefficiency in generating candidate itemsets. Unlike Apriori, FP-Growth constructs a compact data structure called the FP-Tree to efficiently mine frequent itemsets without multiple database scans (Aqliyah et al., 2024). Studies have demonstrated that FP-Growth outperforms Apriori in terms of execution time and memory utilization for large-scale retail data (Amsury et al., 2023; Almahsa et al., 2023). Moreover, FP-Growth has been successfully applied in real-world data-mining applications for sales forecasting, product bundling, and inventory optimization (Padillah, Tambunan, & Nasution, 2022; Aguststewan & Handoko, 2022). This makes FP-Growth particularly suitable for pharmaceutical data, which often contain repetitive and categorical transactional patterns (Nur Aulia et al., 2024).

Empirical research has widely adopted FP-Growth in inventory-related domains. For instance, Zarkhasy and Satria (2024) applied FP-Growth to analyze item correlations in retail warehouses, while Guntoro and Hutabarat (2023) used it to optimize spare-parts inventory management. In the healthcare sector, Agushinta and Putri (2019) utilized FP-Growth to analyze drug purchase patterns and identify frequently co-prescribed medications. Similarly, Nugroho et al. (2024) employed FP-Growth in pharmacy sales data to support procurement and marketing strategies. These studies indicate that the FP-Growth algorithm provides actionable insights that enhance forecasting accuracy and inventory planning, thereby improving operational efficiency (Wijaya et al., 2024; Tarwoto et al., 2023). Nonetheless, most prior implementations are limited to general retail or manufacturing settings, leaving a gap in its application to pharmacy-specific inventory systems in Indonesia (Pertiwi & Tyoso, 2024).

This study aims to fill that research gap by developing a website-based data-mining application that applies the FP-Growth algorithm for pharmacy inventory management. The proposed system will analyze transaction data to discover frequent itemsets of co-purchased medications, providing evidence-based recommendations for restocking and procurement decisions (Nugroho et al., 2024). By integrating FP-Growth with pharmacy inventory data, this study offers both theoretical and practical contributions—combining inventory management principles with data-driven decision support to minimize stock-outs, reduce overstock, and improve resource allocation (Dwiputra et al., 2023; Widisa Adi Kumara et al., 2024). The novelty of this research lies in its domain-specific adaptation of FP-Growth for pharmacy operations in Medan, offering a scalable and implementable framework for enhancing pharmaceutical supply-chain management (Zarkhasy & Satria, 2024).

2. Research Methodolgy

This research uses a quantitative approach. Quantitative research is a type of research that uses numerical or numerical data to analyze phenomena and answer research questions. This research typically aims to test hypotheses, measure relationships between variables, or identify patterns or trends. Data can be collected through surveys, experiments, observations, or existing statistical data. Quantitative research often uses statistical analysis techniques to process and draw conclusions from the results.

Quantitative research aims to gain an understanding of how something is constructed, how it is built, or how it works. Quantitative research is generally driven by a hypothesis, which is formulated and rigorously tested, with the goal of proving it incorrect. The quantitative aspect emphasizes that measurement is fundamental because it provides the link between observation and the formalization of models, theories, and hypotheses. The outcome of quantitative research is the development of models, theories, and hypotheses related to natural phenomena. This research aims to obtain information about drug inventory management, so the research was conducted following the steps shown in Figure 1.

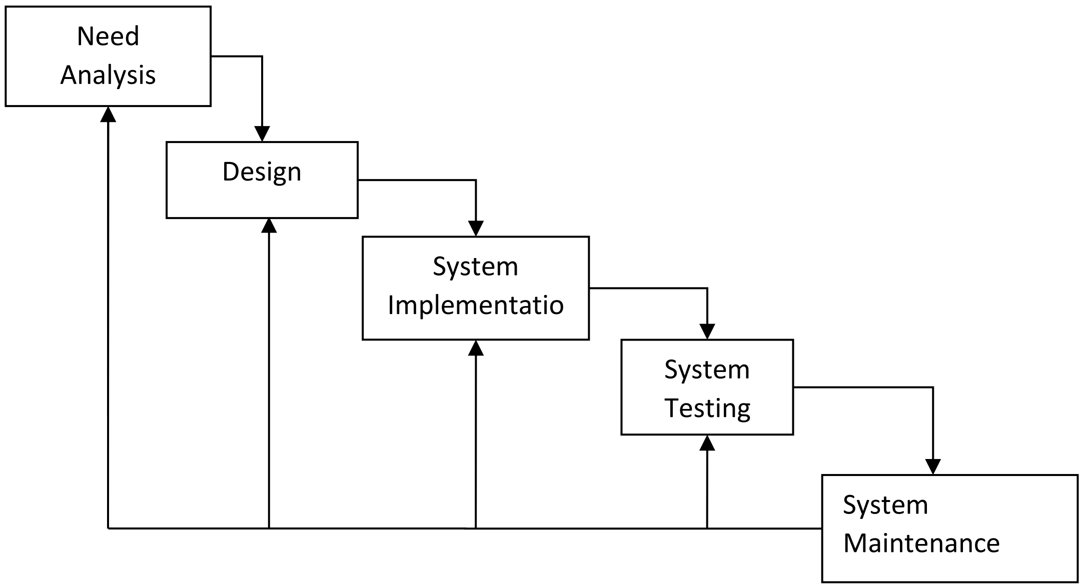

Figure 1 Research Method

The stages in Figure I.1 are explained as follows:

a. Data Collection

Data collection is not simply about collecting existing data; it must also be able to describe the data and contribute to knowledge. Based on its source, data is divided into two categories:

- Primary data, which is data collected directly from the data source. This data collection requires more time and costs than secondary data. Examples of primary data sources include questionnaires, observations, interviews, and experiments conducted directly by researchers.
- Secondary data, which is data obtained from researchers/other parties, even if the data has previously been used differently. Secondary data can be obtained relatively quickly and at a low cost. The secondary data source is Global Medcare Pharmacy. This study uses primary data. The data collected from Tiara Farma Pharmacy is medication inventory.

b. Initial data processing

The collected data is processed to reduce irrelevant data or data with missing attributes. Processing also involves converting redundant (excessive) values or values with excessive diversity into smaller groups to facilitate model development.

c. Proposed model/method

This section describes the flow of the proposed model/method and explains how it works. This model/method is schematically illustrated and accompanied by calculation formulas. The proposed model/method will be developed from the processed data, and the results of the model processing will be compared with the existing model.

d. Experimentation and Model Testing

This section describes the experiments conducted to develop the model and explains how to test the resulting model.

e. Evaluation and Validation of Results

This section describes the prediction results using a data mining algorithm. Validation is used to ensure that the results will be consistent when run independently. Performance is measured by comparing the algorithm's accuracy values.

3. Results and Discussion

Association Rule is a method that aims to find patterns that frequently appear in many transactions. Association Rule Mining is a procedure for finding relationships between items in a dataset. It begins by finding frequent itemsets, which are combinations that frequently appear in an itemset and must satisfy minsup. Association Rules have the form LHS (Left Hand Shake) -> RHS (Right Hand Shake) with the interpretation that if every item in the LHS is purchased, it is likely that the items in the RHS are also purchased. Two important measurements for a rule are Support and Confidence. Support (s) is defined as the percentage of records to the total records of the entire database. Support can be written in the following formula:

$$Support\ (\mathrm{X}) = \frac{\text{The number of x}}{Total\ number\ of\ transactions\ in\ the\ database\ (\mathrm{D})} \tag{1}$$

Meanwhile, the support value of the 2 items is written in the following formula:

$$\text{Support (X,Y)} = \text{P (X} \cap \text{Y)}$$
$$\text{Support (X,Y)} = \frac{\sum \text{The number of X and Y}}{\sum \text{Transaction}} \tag{2}$$

So, for example, if the support for an item is 0.1%, it means that only 0.1 percent of transactions contain this item. Meanwhile, confidence (c) is defined as the percentage of the number of transactions containing X ∪ Y to the total number of records containing X. Confidence can be written in the following formula:

$$\text{Confidance (X | Y)} = \frac{\text{Support (XY)}}{\text{Support (X)}} \tag{3}$$

Confidence is a measure of the strength of an association rule. If we assume the confidence of X->Y is 80%, this means that 80 percent of transactions containing X will also contain Y. Association rule mining can be viewed as a two-step process:

1. Find all frequent itemsets, understanding that each itemset will appear as often as at least a predetermined minimum support (min-sup).
2. Generate strong association rules from frequent itemsets, understanding that these rules must meet both minimum support and minimum confidence.

Case study:
The FP-Growth method is divided into three main stages:
1. Conditional pattern base generation.
2. Conditional FP-Tree generation.
3. Frequent itemset search.

Based on the previous research stages, we obtained drug sales transaction data for Global Medcare Pharmacy in July 2025, as shown in Table 1 below:

Table 1. Drug Sales Transaction Data for Global Medcare Pharmacy

Date	No	Transaction Id	Products Sold
01 Jul 2025	1	TRX001	Paracetamol 500mg, Amoxicillin 500mg, Cetirizine 10mg, Azithromycin 500mg
	2	TRX002	Ibuprofen 400mg, Azithromycin 500mg, Captopril 25mg, Loratadine 10mg
	3	TRX003	Azithromycin 500mg, Cetirizine 10mg, Antasida, Amlodipine 10mg, Amoxicillin 500mg, Loratadine 10mg
	4	TRX004	Simvastatin 20mg, Vitamin C 1000mg, Omeprazole 20mg, Ranitidine 150mg, Ibuprofen 400mg, Azithromycin 500mg
	5	TRX005	Cetirizine 10mg, Omeprazole 20mg, Loratadine 10mg, Ibuprofen 400mg
	6	TRX006	Loratadine 10mg, Metformin 500mg, Cetirizine 10mg, Amoxicillin 500mg
	7	TRX007	Captopril 25mg, Vitamin C 1000mg, Salbutamol Inhaler, Ibuprofen 400mg, Cetirizine 10mg
	8	TRX008	Azithromycin 500mg, Captopril 25mg, Omeprazole 20mg, Loratadine 10mg, Vitamin C 1000mg
	9	TRX009	Ibuprofen 400mg, Antasida, Azithromycin 500mg, Metformin 500mg
02 Jul 2025	10	TRX010	Ranitidine 150mg, Omeprazole 20mg, Metformin 500mg, Vitamin C 1000mg, Paracetamol 500mg, Simvastatin 20mg
	11	TRX011	Azithromycin 500mg, Simvastatin 20mg, Ranitidine 150mg, Cetirizine 10mg, Amlodipine 10mg
	12	TRX012	Azithromycin 500mg, Salbutamol Inhaler, Amlodipine 10mg, Simvastatin 20mg
	13	TRX013	Antasida, Cetirizine 10mg, Simvastatin 20mg, Azithromycin 500mg, Amlodipine 10mg
	14	TRX014	Azithromycin 500mg, Captopril 25mg, Vitamin C 1000mg, Cetirizine 10mg, Metformin 500mg, Omeprazole 20mg
	15	TRX015	Azithromycin 500mg, Captopril 25mg, Simvastatin 20mg, Ranitidine 150mg, Loratadine 10mg
	16	TRX016	Loratadine 10mg, Vitamin C 1000mg, Salbutamol Inhaler, Captopril 25mg, Amoxicillin 500mg
	17	TRX017	Antasida, Vitamin C 1000mg, Omeprazole 20mg, Paracetamol 500mg, Captopril 25mg
	18	TRX018	Captopril 25mg, Ibuprofen 400mg, Vitamin C 1000mg, Loratadine 10mg
	19	TRX019	Paracetamol 500mg, Antasida, Amlodipine 10mg, Metformin 500mg, Amoxicillin 500mg
03 Jul 2025	20	TRX020	Metformin 500mg, Antasida, Captopril 25mg, Paracetamol 500mg
	21	TRX021	Antasida, Vitamin C 1000mg, Omeprazole 20mg, Salbutamol Inhaler, Amlodipine 10mg
	22	TRX022	Metformin 500mg, Azithromycin 500mg, Ranitidine 150mg, Captopril 25mg, Vitamin C 1000mg, Simvastatin 20mg
	23	TRX023	Simvastatin 20mg, Amlodipine 10mg, Vitamin C 1000mg, Loratadine 10mg
	24	TRX024	Loratadine 10mg, Amoxicillin 500mg, Salbutamol Inhaler, Cetirizine 10mg, Vitamin C 1000mg, Amlodipine 10mg
	25	TRX025	Simvastatin 20mg, Salbutamol Inhaler, Antasida, Ranitidine 150mg, Vitamin C 1000mg
	298	TRX298	Ibuprofen 400mg, Ranitidine 150mg, Loratadine 10mg, Amoxicillin 500mg, Amlodipine 10mg
	299	TRX299	Antasida, Ibuprofen 400mg, Loratadine 10mg, Amlodipine 10mg
	300	TRX300	Captopril 25mg, Azithromycin 500mg, Amoxicillin 500mg, Metformin 500mg

1. Determining FP-Growth

To determine FP-Growth, the following steps are required:

(a) FP-Tree Formation

The FP-tree is constructed by representing each transaction data point in a specific path. The initial step in FP-tree formation is to establish a root, often referred to as null. Before forming the FP-tree, it is necessary to identify the frequent item set. Afterward, the data is sorted by frequency of occurrence and selected based on a specified support of 3%. The frequent item set search is shown in Table 2.

$$Support\ (\mathrm{X}) = \frac{\text{The number of x}}{Total\ number\ of\ transactions\ in\ the\ database\ (\mathrm{D})}$$

a. $Support$ (Paracetamol 500mg) = $\frac{104}{300}$ * 100% = 35%

b. $Support$ (Amoxicillin 500mg) = $\frac{93}{300}$ * 100% = 32%

c. $Support$ (Cetirizine 10mg) = $\frac{105}{300}$ * 100% = 35%

d. $Support$ (Azithromycin 500mg) = $\frac{111}{300}$ * 100% = 37%

e. $Support$ (Ibuprofen 400mg) = $\frac{95}{300}$ * 100% = 32%

f. $Support$ (Captopril 25mg) = $\frac{91}{300}$ * 100% = 30%

g. $Support$ (Loratadine 10mg) = $\frac{99}{300}$ * 100% = 33%

h. $Support$ (Antasida) = $\frac{98}{300}$ * 100% = 33%

i. $Support$ (Amlodipine 10mg) = $\frac{111}{300}$ * 100% = 37%

j. $Support$ (Simvastatin 20mg) = $\frac{99}{300}$ * 100% = 33%

k. $Support$ (Vitamin C 1000mg) = $\frac{111}{300}$ * 100% = 37%

l. $Support$ (Omeprazole 20mg) = $\frac{102}{300}$ * 100% = 34%

m. $Support$ (Ranitidine 150mg) = $\frac{96}{300}$ * 100% = 32%

n. $Support$ (Metformin 500mg) = $\frac{92}{300}$ * 100% = 31%

o. $Support$ (Salbutamol Inhaler) = $\frac{91}{300}$ * 100% = 30%

From the calculation above, the results of the item set frequency can be seen as follows:

Table 2 Frequent item set search

No	Product	Sold Frequency	Support
1	Paracetamol 500mg	104 / 300	35 %
2	Amoxicillin 500mg	93 / 300	31 %
3	Cetirizine 10mg	105 / 300	35 %
4	Azithromycin 500mg	111 / 300	37 %
5	Ibuprofen 400mg	95 / 300	32 %
6	Captopril 25mg	91 / 300	30 %
7	Loratadine 10mg	99 / 300	33 %
8	Antasida	98 / 300	33 %
9	Amlodipine 10mg	111 / 300	37 %
10	Simvastatin 20mg	99 / 300	33 %
11	Vitamin C 1000mg	111 / 300	37 %
12	Omeprazole 20mg	102 / 300	34 %
13	Ranitidine 150mg	96 / 300	32 %
14	Metformin 500mg	92 / 300	31 %
15	Salbutamol Inhaler	91 / 300	30 %

To find frequent itemsets and association rules using the FP-Growth algorithm, there are several steps as follows:

1) First, calculate the minimum support.

The minimum support is 3%, so the minimum support count is (3% * 300) = 9, which means the minimum support count is 9. Since the minimum support count is 9, there are items that must be removed. For more details, see Table 3. All products meet the minimum support count and none are removed.

Table 3. Frequency of all items for all transactions

No	Product	Sold Frequency	Support
1	Paracetamol 500mg	104 / 300	35 %
2	Amoxicillin 500mg	93 / 300	31 %
3	Cetirizine 10mg	105 / 300	35 %
4	Azithromycin 500mg	111 / 300	37 %
5	Ibuprofen 400mg	95 / 300	32 %
6	Captopril 25mg	91 / 300	30 %
7	Loratadine 10mg	99 / 300	33 %
8	Antasida	98 / 300	33 %
9	Amlodipine 10mg	111 / 300	37 %
10	Simvastatin 20mg	99 / 300	33 %
11	Vitamin C 1000mg	111 / 300	37 %
12	Omeprazole 20mg	102 / 300	34 %
13	Ranitidine 150mg	96 / 300	32 %
14	Metformin 500mg	92 / 300	31 %
15	Salbutamol Inhaler	91 / 300	30 %

1) Prioritize items with the highest frequency
After eliminating product items below the minimum support count, the next step is to prioritize items with the highest frequency, as shown in Table III.5. Table III.5 shows that Amlodipine 10mg, Azithromycin 500mg, and Vitamin C 1000mg have the highest frequency, at 111, and so on.

Table 4. Order of product items based on the highest frequency

No	Product	Sold Frequency	Priority
1	Amlodipine 10mg	111	1
2	Azithromycin 500mg	111	2
3	Vitamin C 1000mg	111	3
4	Cetirizine 10mg	105	4
5	Paracetamol 500mg	104	5
6	Omeprazole 20mg	102	6
7	Loratadine 10mg	99	7
8	Simvastatin 20mg	99	8
9	Antasida	98	9
10	Ranitidine 150mg	96	10
11	Ibuprofen 400mg	95	11
12	Amoxicillin 500mg	93	12
13	Metformin 500mg	92	13
14	Captopril 25mg	91	14
15	Salbutamol Inhaler	91	15

2) Order items by priority
After the highest and lowest frequencies are determined, the items are reordered by priority. At the beginning of the process, the sales transaction data totaled 111. The resulting order (ordering) is shown in Table III.6.

Table 5. Result of product item ordering by priority

No.	ID Transaction	Item
1	TRX38	["Amlodipine 10mg","Antasida","Ibuprofen 400mg","Captopril 25mg"]
2	TRX225	["Amlodipine 10mg","Azithromycin 500mg","Antasida","Ranitidine 150mg"]
3	TRX36	["Amlodipine 10mg","Azithromycin 500mg","Cetirizine 10mg","Captopril 25mg"]
4	TRX186	["Amlodipine 10mg","Azithromycin 500mg","Cetirizine 10mg","Paracetamol 500mg"]
5	TRX47	["Amlodipine 10mg","Azithromycin 500mg","Cetirizine 10mg","Simvastatin 20mg"]
6	TRX247	["Amlodipine 10mg","Azithromycin 500mg","Metformin 500mg","Captopril 25mg"]
7	TRX37	["Amlodipine 10mg","Azithromycin 500mg","Paracetamol 500mg","Amoxicillin 500mg"]
8	TRX72	["Amlodipine 10mg","Azithromycin 500mg","Ranitidine 150mg","Salbutamol Inhaler"]
9	TRX39	["Amlodipine 10mg","Azithromycin 500mg","Simvastatin 20mg","Ranitidine 150mg"]
10	TRX12	["Amlodipine 10mg","Azithromycin 500mg","Simvastatin 20mg","Salbutamol Inhaler"]
296	TRX126	["Simvastatin 20mg","Antasida","Ranitidine 150mg","Metformin 500mg"]
297	TRX116	["Simvastatin 20mg","Ibuprofen 400mg","Metformin 500mg","Salbutamol Inhaler"]
298	TRX235	["Antasida","Ranitidine 150mg","Ibuprofen 400mg","Metformin 500mg"]
299	TRX231	["Antasida","Ranitidine 150mg","Ibuprofen 400mg","Amoxicillin 500mg","Metformin 500mg","Salbutamol Inhaler"]
300	TRX115	["Ranitidine 150mg","Amoxicillin 500mg","Captopril 25mg","Salbutamol Inhaler"]

From the data above, it can be seen that the results of the Header Table record all items that pass the minimum support, and point to the first node of the item in the FP-Tree as in the table below:

No	Drug Name	Sold Frequency	Prioritas
1	Amlodipine 10mg	111	1
2	Azithromycin 500mg	111	2
3	Vitamin C 1000mg	111	3
4	Cetirizine 10mg	105	4
5	Paracetamol 500mg	104	5

No	Drug Name	Sold Frequency	Prioritas
6	Omeprazole 20mg	102	6
7	Loratadine 10mg	99	7
8	Simvastatin 20mg	99	8
9	Antasida	98	9
10	Ranitidine 150mg	96	10
11	Ibuprofen 400mg	95	11
12	Amoxicillin 500mg	93	12
13	Metformin 500mg	92	13
14	Captopril 25mg	91	14
15	Salbutamol Inhaler	91	15

The conditional pattern base is a subdatabase containing prefix paths and suffix patterns. The conditional pattern base is generated using the previously constructed FP-Tree. To find frequent itemsets, the paths ending with the smallest to the largest support count are first determined. The following table shows the formation of the conditional pattern base:

Table 6 Conditional Pattern Base Generation

No	Suffix	Subset
1	Azithromycin 500mg	[Amlodipine 10mg] => 21 [Vitamin C 1000mg] => 17 [Amlodipine 10mg, Vitamin C 1000mg] => 1
2	Cetirizine 10mg	[Azithromycin 500mg] => 19 [Amlodipine 10mg, Azithromycin 500mg] => 6 [Vitamin C 1000mg] => 16 [Azithromycin 500mg, Amlodipine 10mg] => 7 [Azithromycin 500mg, Vitamin C 1000mg] => 6 [Amlodipine 10mg, Vitamin C 1000mg] => 5 [Vitamin C 1000mg, Azithromycin 500mg] => 4 [Amlodipine 10mg] => 16 [Azithromycin 500mg, Vitamin C 1000mg, Amlodipine 10mg] => 1
3	Paracetamol 500mg	[Azithromycin 500mg, Cetirizine 10mg] => 4 [Vitamin C 1000mg] => 11 [Amlodipine 10mg] => 12 [Amlodipine 10mg, Azithromycin 500mg] => 5 [Cetirizine 10mg] => 11 [Amlodipine 10mg, Cetirizine 10mg] => 7 [Azithromycin 500mg, Amlodipine 10mg] => 3 [Amlodipine 10mg, Azithromycin 500mg, Cetirizine 10mg] => 4 [Azithromycin 500mg] => 12 [Amlodipine 10mg, Vitamin C 1000mg] => 4 [Vitamin C 1000mg, Azithromycin 500mg] => 2 [Azithromycin 500mg, Vitamin C 1000mg, Cetirizine 10mg] => 2 [Vitamin C 1000mg, Cetirizine 10mg] => 2 [Azithromycin 500mg, Vitamin C 1000mg] => 1 [Azithromycin 500mg, Vitamin C 1000mg, Amlodipine 10mg] => 1 [Vitamin C 1000mg, Amlodipine 10mg] => 2

In determining the Association rule, it is as follows:

1. Number of transactions for Metformin 500mg,Salbutamol Inhaler => Vitamin C 1000mg = 12 transaction
2. Number of transactions for Antasida,Salbutamol Inhaler => Paracetamol 500mg =11 transaction

3. Number of transactions for Antasida,Salbutamol Inhaler => Cetirizine 10mg = 11 transaction
4. Number of transactions for Amoxicillin 500mg,Salbutamol Inhaler => Loratadine 10mg = 12 transaction
5. Number of transactions for Loratadine 10mg,Salbutamol Inhaler => Amoxicillin 500mg = 12 transaction
6. Number of transactions for Amlodipine 10mg,Salbutamol Inhaler => Vitamin C 1000mg = 11 transactiondan seterusnya

   Meanwhile, the transactions obtained from the Conditional Pattern base that pass the minimum support criteria are as follows:

   Therefore, the association formation process is as follows:

   Confidance (Metformin 500mg,Salbutamol Inhaler => Vitamin C 1000mg) $= \frac{12}{27} * 100\%$ = 44.44%

   Confidance Antasida,Salbutamol Inhaler => Paracetamol 500mg $= \frac{11}{27} * 100\%$ = 40.74%

   Confidance Antasida,Salbutamol Inhaler => Cetirizine 10mg $= \frac{11}{27} * 100\%$ = 40.74%

   Confidance Amoxicillin 500mg,Salbutamol Inhaler => Loratadine 10mg omg $= \frac{12}{26} * 100\%$ = 46.15%

   Confidance Loratadine 10mg,Salbutamol Inhaler => Amoxicillin 500mg $= \frac{12}{29} * 100\%$ = 41.38%

Table 7. Confidence Results

No	Rule	Frequency of Sales A,B,C / AB	AB Sold Frequency	Confidence
1	Metformin 500mg,Salbutamol Inhaler => Vitamin C 1000mg	12	27	44.44 %
2	Antasida,Salbutamol Inhaler => Paracetamol 500mg	11	27	40.74 %
3	Antasida,Salbutamol Inhaler => Cetirizine 10mg	11	27	40.74 %
4	Amoxicillin 500mg,Salbutamol Inhaler => Loratadine 10mg	12	26	46.15 %
5	Loratadine 10mg,Salbutamol Inhaler => Amoxicillin 500mg	12	29	41.38 %
6	Amlodipine 10mg,Salbutamol Inhaler => Vitamin C 1000mg	11	27	40.74 %
7	Amlodipine 10mg,Vitamin C 1000mg => Salbutamol Inhaler	11	19	57.89 %
8	Amlodipine 10mg,Salbutamol Inhaler => Omeprazole 20mg	12	27	44.44 %
9	Ibuprofen 400mg,Salbutamol Inhaler => Vitamin C 1000mg	10	23	43.48 %
10	Amlodipine 10mg,Metformin 500mg => Vitamin C 1000mg	10	21	47.62 %
70	Amoxicillin 500mg,Paracetamol 500mg => Amlodipine 10mg	18	35	51.43 %
71	Amlodipine 10mg,Azithromycin 500mg => Paracetamol 500mg	13	14	92.86 %
72	Azithromycin 500mg,Vitamin C 1000mg => Cetirizine 10mg	11	16	68.75 %

73	Amlodipine 10mg,Azithromycin 500mg => Cetirizine 10mg	14	14	100 %
74	Amlodipine 10mg,Cetirizine 10mg => Azithromycin 500mg	14	35	40 %
75	Cetirizine 10mg => Azithromycin 500mg	43	105	40.95 %

From the calculations above, the association with the highest percentage is:

1. If (Azithromycin 500mg, Vitamin C 1000mg, then Metformin 500mg) will be provided = 106%.
2. If (Amlodipine 10mg, Azithromycin 500mg, then Cetirizine 10mg) will be provided = 100%.

Discussion

The results of this study demonstrate the effectiveness of the FP-Growth algorithm in identifying frequent itemsets and strong association rules from pharmaceutical sales transactions at Global Medcare Pharmacy. By applying a minimum support threshold of 3%, a total of 15 frequent items were identified, with *Azithromycin 500mg*, *Amlodipine 10mg*, and *Vitamin C 1000mg* emerging as the most frequently sold products, each appearing in 37% of transactions. The subsequent construction of the FP-Tree enabled efficient pattern discovery without generating candidate itemsets, thus optimizing computational performance compared to the Apriori algorithm. The generated conditional pattern bases revealed strong correlations among several drug combinations frequently purchased together. For instance, the association rule *Amlodipine 10mg, Azithromycin 500mg → Cetirizine 10mg* achieved a confidence level of 100%, indicating that every transaction containing *Amlodipine* and *Azithromycin* also included *Cetirizine*. Similarly, the rule *Amlodipine 10mg, Azithromycin 500mg → Paracetamol 500mg* recorded a high confidence of 92.86%, suggesting a strong purchasing relationship among these drugs. The analysis also found other meaningful associations, such as *Azithromycin 500mg, Vitamin C 1000mg → Cetirizine 10mg* (68.75%) and *Amlodipine 10mg, Vitamin C 1000mg → Salbutamol Inhaler* (57.89%). These relationships provide valuable insights into consumer purchasing behavior and can assist pharmacies in optimizing inventory management and marketing strategies by predicting complementary product demand. Moreover, the FP-Growth algorithm's ability to process 300 transactions efficiently underscores its scalability for larger datasets. Overall, the findings validate FP-Growth as a robust and efficient approach for market basket analysis in pharmaceutical retail, capable of uncovering hidden purchasing patterns that can improve decision-making in stock replenishment, promotional bundling, and sales forecasting.

4. Conclusion

The findings of this study confirm that the FP-Growth algorithm is a highly effective data mining technique for identifying frequent itemsets and generating strong association rules from pharmaceutical sales transactions, enabling more accurate forecasting and data-driven decision-making in inventory management. By uncovering significant relationships among frequently co-purchased drugs—such as Amlodipine 10mg, Azithromycin 500mg, and Cetirizine 10mg—the algorithm provides valuable insights for optimizing stock control, minimizing shortages, and supporting targeted promotional strategies. These results underscore the practical implication that integrating FP-Growth into pharmacy information systems can enhance operational efficiency and responsiveness to consumer demand. However, the research is limited by its use of a single pharmacy dataset and the absence of real-time transaction integration, which may constrain the generalizability of the findings. Future research should address these limitations by employing larger, multi-site datasets, integrating dynamic inventory variables, and exploring hybrid approaches that combine FP-Growth with predictive models such as machine learning-based demand forecasting. Such advancements would not only refine algorithmic performance but also strengthen the strategic application of data mining in the broader healthcare supply chain ecosystem.

References

Adawiyah, Q., Defit, S., & Sumijan. (2024). Penerapan Algoritma K-Means Clustering untuk Mengelompokkan Rekomendasi Metode Kontrasepsi Berbasis Machine Learning di Puskesmas. Jurnal KomtekInfo, 300–305. https://doi.org/10.33365/jkt.v5i3.1204

Afisyah, A., Winata, H., & Syahputra, Y. H. (2021). Implementasi Data Mining dengan Metode FP-Growth untuk Strategi Promosi pada Toko Cool Kids Plaza Medan Fair. Jurnal Cyber Tech. http://ojs.trigunadharma.ac.id/index.php/jct/article/view/4055

Ayu, A., Perdana Windarto, A., & Suhendro, D. (2021). Implementasi Data Mining dengan Metode FP-Growth terhadap Data Penjualan Barang sebagai Strategi Penjualan pada CV. A & A Copier. RESOLUSI: Rekayasa Teknik Informatika dan Informasi, 2(2), 67–75. https://djournals.com/resolusi

Boy, A. F., Yakub, S., Ishak, I., & Azmi, Z. (2022). Implementasi Data Mining pada Pengaturan Distribusi Barang dengan Menggunakan Algoritma FP-Growth. Journal of Science and Social Research, 5(2), 431. https://doi.org/10.54314/jssr.v5i2.947

Filki, Y. (2022). Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa. Jurnal Informatika Ekonomi Bisnis, 4, 166–171. https://doi.org/10.37034/infeb.v4i4.166

Fitra Andikos, A. (2019). Perancangan Aplikasi Multimedia Interaktif sebagai Media Pembelajaran Pengenalan Hewan pada TK Islam Bakti 113 Koto Salak. Indonesia Jurnal Sakinah: Jurnal Pendidikan dan Sosial Islam, 1(1), 34–49.

Han, J., Pei, J., & Kamber, M. (2022). Data Mining: Concepts and Techniques (4th ed.). Morgan Kaufmann.

Hermiati, R., Asnawati, A., & Kanedi, I. (2021). Pembuatan E-Commerce pada Raja Komputer menggunakan Bahasa Pemrograman PHP dan Database MySQL. Jurnal Media Infotama, 17(1), 54–66. https://doi.org/10.37676/jmi.v17i1.1317

Muhamad. (2022). Sistem Informasi Penjualan Online Berbasis Mobile pada Supermarket Kasimura. Jurnal Processor, 17(1), 34–45. https://doi.org/10.33998/processor.2022.17.1.1190

Muhammad Aldi Zarkashy, Ismarmiaty, I., & Ria Rismayati. (2024). Analisa Pola Pembelian Produk dengan Menggunakan Metode FP-Growth pada Toko Heron. Journal of Information Systems Management and Digital Business, 1(2), 107–115. https://doi.org/10.59407/jismdb.v1i2.312

Nanda, A. P., Pramono, D. E. H., & Hartati, S. (2020). Menentukan Tingkat Kepuasan Mahasiswa terhadap Pelayanan Akademik Menggunakan Metode Algoritma K-Means. Jurnal Sistem Informasi dan Telematika, 11(1), 23–28.

Nu, S. M. P., Hikam, M., & Metode, D. (2024). Penerapan Algoritma Data Mining dalam Sistem Prediksi Penjualan. Jurnal Teknologi dan Sistem Informasi, 6(2), 343–349.

Padillah, A., Satria Tambunan, H., Nasution, R. A., & Artikel, G. (2022). Penerapan Algoritma FP-Growth dalam Menganalisa Pola dan Tata Letak Bahan Makanan. JOMLAI: Journal of Machine Learning and Artificial Intelligence, 1(4), 2828–9099. https://doi.org/10.55123/jomlai.v1i4.1673

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2018). Introduction to Data Mining (2nd ed.). Pearson.

Wahyusari, R. (2023). Penerapan Algoritma FP-Growth untuk Menemukan Pola Peminjaman Alat pada Workshop Teknik Mesin. LOGIC: Jurnal Ilmu Komputer dan Pendidikan, 1(3), 406–411. https://journal.mediapublikasi.id/index.php/logic/article/view/2745

Zyen, A. K., Wibowo, G. W. N., Widiastuti, N. A., & Wahono, B. B. (2023). Klasterisasi Penerima Bantuan Fasilitas Sekolah di MI Datuk Singaraja Menggunakan Metode K-Means. Jurnal Teknik Informatika, 2(1), 25–38. https://journal.unisnu.ac.id/JTINFO/article/view/482