JURNAL INFOTEL
Informatics - Telecommunication - Electronics
Website: https://ejournal.ittelkom-pwt.ac.id/index.php/infotel
ISSN: 2085-3688; e-ISSN: 2460-0997

Implementation of association rule using apriori algorithm and frequent pattern growth for inventory control

Imam Riadi[1], Herman[2], Fitriah[3,*], Suprihatin[4], Alwas Muis[5], Muhajir Yunus[6]
[1,4]Department of Information System, Universitas Ahmad Dahlan
[2,3,5,6]Master Program of Informatics, Universitas Ahmad Dahlan
[1,2,3,4,5,6]Jl. Ahmad Yani, Tamanan Banguntapan, Bantul 55191, Indonesia
[*]Corresponding email: 2207048004@webmail.uad.ac.id

Received 8 June 2023, Revised 13 November 2023, Accepted 21 November 2023

Abstract — Inventory control is needed to maintain smooth operations and improve product efficiency. The problem of procurement of goods that are not by needs is a problem that often occurs in various companies. This problem has the potential to affect business processes that hinder increased sales. Based on this, a strategy is needed in managing inventory to increase customer satisfaction. This study aims to find combinations of goods to be analyzed, resulting in an Association Rule that can be used by store owners to manage the inventory of goods. The algorithms used in this study are Apriori and Frequent Pattern Growth (FP-Growth). Apriori algorithm methods and FP-Growth are two techniques used in data mining to find combinations of goods resulting in the Association Rule. This research begins with literature study and data acquisition as the first step in data pre-processing. Next, data analysis is performed to identify patterns and relationships between items in the dataset. At the evaluation stage, the lift ratio method is used to determine the validity of association rules. This is an important step to ensure that the rules found have relevance and are being analyzed so that they can be used for better decision-making. The results of the implementation of a priori and FP-Growth algorithms for 2-item analysis show that they both produce the same number of association rules, *i. e.* 16 rules. There is a similarity in the rules with the highest lift ratio, which is 14. In addition, some other rules with a lift ratio of 5 are also found in both algorithms. Next, the calculation of the 3-item set lift value, both show the same lift ratio value. This indicates that they are effective in extracting association rules from the same dataset, for both 2-itemset and 3-itemset.

Keywords – apriori, association rule, data mining, frequent pattern growth, lift ratio

I. Introduction

Data is very important in today's era and can support decisions on various aspects. Making decisions based on facts from data in the field is a smart move and can improve the quality of decisions. One example is in the sales process, especially in retail stores that provide various products [1], [2]. Every day, retail stores make many sales transactions, so the sales data generated is also very much. However, the data should not only be used as a record without added value but must be used for business purposes. Transaction data storage will continue to grow over time, so data must be optimized to improve efficiency [3], [4].

Goods control is an important aspect that needs to be managed properly so that a business can run smoothly and efficiently. The development of information technology has begun to be utilized by competition between rulers, so developers need to find the right strategy to meet needs and ensure customer satisfaction. One of the strategic innovations in business is to look for connections or linkages between different sets of products so that they can be packaged and sold simultaneously [5]–[7]. This innovation provides a solution to sales and inventory problems because product mismatches increase sales. However, finding relationships between goods is a difficult process due to large product problems, especially when merchants have thousands of products [8]–[11].

For inventory management, data mining algorithms can be used to find patterns in inventory data, one of which is by applying the association rule technique and the method that can be applied in this study is market basket analysis [12]–[14]. Market basket analysis is

Fig. 1. Research flow.

one method in data mining [15], [16]. Data mining is a complex data analysis process that involves statistical, mathematical, and artificial intelligence techniques to explore and discover hidden patterns in large data sets [17]–[19]. The goal of data mining is to get valuable and useful information from data to make business decisions. Data mining can be applied to different types of data such as customer data, inventory data, transaction data, marketing data, and other data [20]–[22].

The data mining process includes several steps such as data preprocessing, selection of appropriate data mining methods, method execution, evaluation of results, and interpretation of results. Some commonly used data mining techniques include classification, regression, clustering, association, and ranking models. Each of these techniques has a different purpose and can be applied to different types of data [23], [24]. In business, data mining can help optimize inventory management, improve product efficiency, improve customer experience, and make more accurate business decisions [25], [26]. Data mining can also help identify hidden trends and patterns in data, allowing companies to respond more effectively to changing market or customer needs [27]–[32].

Two data mining algorithms that are often used in inventory analysis are the apriori algorithm and FP-Growth. Both of these algorithms can be used in the context of inventory control to find patterns of association and high frequency between inventory items [33]–[35]. Apriori algorithm is an algorithm in data mining that is used to find patterns that often appear (frequent itemsets) in a collection of transaction data. This algorithm works by dividing the dataset into smaller subsets and then looking for items that often appear in each subset. After that, the found itemsets will be combined to form larger itemsets and then recalculated in frequency. This process is done iteratively until no more frequently occurring itemsets can be found. Apriori algorithms are useful in data analysis and marketing, where we can find frequently emerging consumer buying patterns and create more effective marketing strategies [36].

The FP-Growth algorithm uses a top-down approach to generate a tree data structure called FP-Tree, which represents all transactions in a more concise form. Then, the FP-Growth algorithm will extract frequent itemsets from the FP-Tree using a recursive approach, which is a more efficient alternative to apriori algorithms, especially when the dataset is very large and complex. These two algorithms have the same goal of finding patterns in datasets, but apriori and FP-Growth algorithms have different approaches and techniques for extracting relevant information from data. Natural inventory control [37]–[40].

Some previous research has been related to the utilization of association rule algorithms. For example, research conducted by Ardiantoro and Sunarmi [41] (2019) aims to analyze the game patterns of badminton players which is one of the popular sports in Indonesia. In this study, data was collected and processed by dividing the playing field into several play areas. In addition, Apridonal *et al.* [42] (2019) conducted research to develop an application that can identify sales patterns using the Association Rule method and apriori algorithm. Hu *et al.* [43] also used frequency pattern growth algorithms in their study to analyze the stability of public transport trips using association rule mining. The next research was conducted by Elisa and Azwanti [44] (2019) who used the association rule method with the FP-Growth algorithm to analyze the frequency of purchasing 3 kg LPG gas.

Based on previous research that has used FP-Growth and apriori algorithms for various purposes, such as analyzing game patterns, travel stability, developing children's toy applications, purchasing fertilizers, transportation trip stability, and purchasing 3 kg LPG gas, the FP-Growth algorithm is more often used because it is considered superior to apriori [45]. Therefore, this study will also use the FP-Growth algorithm to find association rules to help store owners manage their product inventory. Although using the same algorithm, the main difference lies in the use of different data, namely clothing sales data from the city of Bengkulu. The purpose of this study is to find combinations of goods so that store owners know the right placement of goods and analyze itemsets to obtain association rules using apriori and FP-Growth algorithms.

II. Research Method

This research stage begins with a literature review and then data acquisition is carried out for processing. Next, analyze by implementing the apriori algorithm and FP-Growth in goods control. The research flow can be seen in Fig. 1.

A. Literature Study

The study of literature is a relatively simple research method, but it is very important in this process of inquiry. This method involves collecting data from various literature sources, such as journals, books, papers, theses, theses, and related research articles. Data collection is done by reading, recording, and processing information from these sources. The process of data analysis in literature studies begins with evaluating the most relevant and current scientific work with this research. The selection of literature starts from the latest and passes on to the older ones. Any previous research relevant to this topic will be summarized to evaluate the alignment of the issues raised with the issues studied in this study. An important and relevant part of the research problem is also considered. By referring to the relevant literature, this research can design a system that can be used for inventory control based on the association rule method with apriori algorithm and FP-Growth.

B. Data Acquisition

Data acquisition, the process of collecting data from various sources, holds significant benefits in various contexts. Firstly, data acquisition provides essential information needed for making informed decisions, be it in business, research, or scientific fields. Furthermore, the data gathered can be used for monitoring the performance of systems, devices, or processes, as well as for analysis and pattern discovery that can help identify trends, weaknesses, opportunities, and threats. Data acquisition also plays a crucial role in scientific research, where the obtained data is used to support or refute hypotheses, identify trends, and draw scientific conclusions. In business, data acquired through acquisition can be used to optimize processes, reduce costs, and enhance quality. It also supports better decision-making, reducing uncertainty in the decision-making process. Additionally, data acquisition fuels the development of intelligent systems such as artificial intelligence (AI) and machine learning, which are used for classification, prediction, and pattern recognition. The collected data is also employed for security purposes, detecting anomalies, and developing products that better match customer needs. Data acquisition also has a role in compliance and audit, used to track and verify adherence to applicable regulations and standards. Thus, data acquisition is a key step in optimizing various aspects of life and business.

The clothing data acquisition process begins with the request submission stage which aims to obtain a complete copy of sales data from the cash register system. This data includes all clothing sales transactions recorded in the period from February 1 to March 31, 2023. During that period, this data will include detailed information about each transaction, including details about products sold, number of items recorded, and total transactions. This data will be the basis for a more in-depth analysis, allowing the owner of the figure to understand sales trends, manage inventory, and evaluate sales performance. The following item names in the GalleryNCA store are shown in Table 1.

Table 1. The Name of the Item in the GalleryNCA Store

Initials	Item name	Initials	Item name
K	Kaos	BBH	Baby hanger
RI	Rok include	KKB	Kaos kaki baby
KL	Kaos lemon	KJP	Kemeja jeans pdk
LS	Lejing s. M. L	CA	Celana amrico
SC	Stelan clobie	SHC	Stelan carnaval
SCP	Stelan clonie p	LX	Lejing xl
CAK	Cd agree kids	LB1	Lejing baby 123
DK	Daster kalong	ST0	Bh2telan timy 0
CKN	Celana katun	KS	Kemeja salur
CK	Celana kerja	HY	Hayami
RB	Rok baloteli	KD	Kaos destino
BKN	Blus katun	STS	Stelan timy s.l
BF	Blus filimili	SCO	Shot codigo
HL	Handuk lancer	BH	Bh kawat
PS	Pp sielie/meong	CVM	Cd vaya m.l
KN	Kaos noren	BI	Blus impor
STX	Stelan timy xl	DKS	Dres kaos
LK	Lejing korea	TI	Tunik impor
CKJ	Celana kerja J	GD7	Gamis de 7
JV	Jamper valvet	CKR	Cp karakter
B	Blus	TK	Tunik katun
BK	Blus korea	SVS	Stelan velvet s.m.
RKN	Rok katun	SS	Stelan sibulat
TP	Tunik platinum	J	Jumper
EJ	Egoist jumbo	CT	Cp tessa
BS	Baby shop	DHK	Daster hap katun
SI	Stelan impor	H	Hoki
SM	Stelan mld	GI	Gaun impor
SMI	Stelan milady		

C. Preprocessing

Before data can be processed with apriori and FP-Growth algorithms, a crucial first step is to preprocess the data. Pre-processing includes two main stages, namely cleaning and data transformation. The cleaning phase includes deleting invalid data, grouping missing values, or handling outlier data to ensure good data quality. In addition, the data transformation stage includes changing the data format, normalization, or encoding categorical data to numerical form so that it can be used by the algorithm. By careful preprocessing, the data will be more readily used by apriori and FP-Growth algorithms, resulting in more accurate and informative analysis results. Thus, preprocessing data becomes an important step in the process of extracting association patterns and analyzing itemsets.

D. Analysis

This study conducts process analysis using data mining methodology by applying apriori and FP-Growth algorithms to determine the frequent itemset that will be used as a basis in formulating the best rules or recommendations in purchasing goods.

1) Stages of implementation of apriori algorithm

The first step in forming association rules is to determine the minimum value of support to get high frequency. The higher the support value, the more often the itemset appears, and this can be useful information in association analysis, in which case the researcher sets the minimum support value at 0.2 % with (1).

$$Support(A) = \frac{\sum Contains\ A}{\sum Total\ Transactions} \times 100\% \quad (1)$$

The formula is used to measure the extent to which the itemset or element (A) appears in a dataset transaction. Data that has met the support set of frequently occurring items is merged, and the set of non-conforming items will be eliminated, and the set that meets will be used in iterations. The second step is to search for 2-itemset. The value of the 2-itemset is calculated using (2).

$$Support\ (A \rightarrow B) = \frac{\sum Contains\ A\ and\ B}{\sum Total\ Transaction} \times 100\% \quad (2)$$

The value containing A and B is the value of the two items of goods in the transaction, dividing the value containing A and B by the total amount of all transactions. The third step 3-itemset support value is obtained using (3).

$$Support\ (A,B,C) = \\ = \frac{\sum Contains\ A,\ B,\ and\ C}{\sum Total\ Transaction} \times 100\% \quad (3)$$

Confidence is one of the important aspects of association analysis, which is used to measure the extent to which an association or association rule applies. In this context, we will understand how strong the relationship between two itemsets, namely A and B or A, B and C. Confidence, measured in percentages, gives an idea of how often itemset B appears in transactions that also contain itemset A or how often Item C appears in transactions that also contain items A and B, in this study the minimum value of confidence has been set at (0.8 %) To determine the value of confidence, it is obtained using (4).

$$Confidence = P(A \mid B) = \\ = \frac{Number\ of\ Transactions\ A\ and\ B}{Number\ of\ Transactions\ A} \times 100\% \quad (4)$$

2) *Stages of FP-Growth algorithm implementation*

The search for associations using the FP-Growth method is a development of apriori algorithms. In the FP-Growth algorithm, confidence and support values are applied similarly to apriori, but the difference lies in the more efficient approach used by FP-Growth. FP-Growth reduces complexity by involving only one iteration through a dataset of transactions, in contrast to apriori which requires multiple iterations. FP-Growth implements three main stages:

- Conditional pattern base generation phase: At this stage, FP-Growth identifies patterns that appear repeatedly in the transaction dataset.
- Conditional FP-Tree development phase: FP-Growth builds a data structure known as a conditional FP-Tree. This structure describes the relationship between itemsets that appear together in a transaction.
- Frequent itemset search phase by utilizing TID (Transaction Identity): In this stage, FP-Growth uses the conditional FP-Tree that has been built to efficiently generate frequent itemset. This includes the calculation of support and confidence values.

E. Evaluation

Association analysis is the validity of a rule determined using the lift ratio method. Lift ratio is a common indicator used to determine whether an association rule is considered valid or not. An association rule is considered valid when the lift ratio value in the rule exceeds 1, and the higher the lift ratio value, the stronger the rule. In apriori and FP-Growth algorithms, lift ratio is used to evaluate the strength of the association rule found to calculate lift ratio can use (5).

$$Lift\ Ratio = \frac{Confidence(A,B)}{Benchmark\ Confidence(A,B)} \quad (5)$$

For benchmark confidence use (6).

$$Benchmark\ Confidence = \frac{N_c}{N} \quad (6)$$

III. Result

This section discusses research data, pre-processing, apriori implementation, and FP-Growth implementation.

A. Research Data

Data collection in this study was carried out through clothing sales transactions recorded in the form of shopping receipts. Every time a customer makes a purchase of clothes in the store, the information recorded in the shopping receipt includes details such as the type of clothes purchased, the number of items, the transaction number. The data used was 140 transactions and 57 items of goods which were sales samples at the GalerryNca store in Bengkulu City. Transaction dataset in tabular form as in Table 2.

Table 2 is a dataset of goods transactions, the name of the goods has been changed to initials. To find out the full name of each item, refer to Table 1.

B. Pre-processing

The data processing process is carried out with the aim of finding patterns of similarity of purchased goods based on daily purchase report data. Sales report data every day at first is just ordinary information then after processing it will be very useful information for future business progress and improvement. Therefore,

Table 2. Transaction Dataset

Transaction No	Name of Goods			
	Item 1	Item 2	Item 3	Item 4
210105.001	K	-	-	-
210105.002	RI	K	-	-
210105.003	KL	-	-	-
210105.004	LS	SC	SCP	-
210105.005	CAK	-	-	-
···	···	···	···	···
210205.009	GI	ST	KJK	CA
210208.010	K	BS	-	-
210208.011	BKN	BF	-	-
210208.012	HL	RKN	-	-
210208.013	PS	CK	TP	-
210208.014	KN	PS	RB	-
210208.015	EJ	K	SI	-

Table 3. Product Transaction Tabulation Results

No.	Transaction No	Good Items			
		K	*RI*	···	G
1.	210.105.001	1	0	···	0
2.	210.105.002	1	1	···	0
3.	210.105.003	0	0	···	0
4.	210.105.004	0	0	···	0
5.	210.105.005	0	0	···	0
6.	210.105.006	0	0	···	0
···	···	···	···	···	···
140.	210.208.015	0	0	···	0

researchers will test the results of using apriori and FP-Growth to determine the relationship between buying patterns.

In the sales data table, the product sold has been processed into a tabular form with binary contents, a value of 1 indicates that the product has been sold, while a value of 0 indicates that the product has not been sold. The data view can be seen in Table 3. From a tabular format it is easier to know how many items were purchased in each transaction to look for high-frequency patterns by creating a 1-itemset.

C. Apriori Implementation

1) High frequency

The stages of finding a high frequency with a predefined minimum support (0.2 %) using (1), the number of transaction items (A) divided by the total number of transactions and multiplied (100 %), items that do not meet the minimum support will be removed and items that meet the minimum support can be seen in Table 4.

2) Minimun support 2-itemset

Table 4 displays items that have met the minimum support for 1-itemset then a calculation will be carried out with 2-itemset where items are purchased simultaneously in one transaction as many as 2 items (A and B) with the same minimum support provisions (0.2 %). The calculation results using (2) are shown in Table 5. After getting the result of the 2-itemset then proceed to look for a combination of 3-itemset.

3) Minimum Support 3-itemset

Results Transaction data with 2-itemset shown in Table 4 that has met the minimum support value will be continued with the calculation of 3-itemset, by calculating transactions of items A, B, and C with (3) display the results of the 3-itemset calculation in Table 6. From the results of the calculation of the minimum Support 3-itemset, then the determination of association rules will be carried out from 2-itemset and 3-itemset data that meet the minimum value of support and minimum confidence.

Table 4. Minimum Support 1-itemset

Transaction	Items	Support (%)
LK	0.179	17.86
CK	0.121	12.14
STX	0.114	11.43
RI	0.086	8.57
BF	0.071	7.14
CAK	0.071	7.14
HL	0.071	7.14
BK	0.064	6.43
K	0.057	5.71
B	0.043	4.29
PS	0.043	4.29
JV	0.036	3.57
KN	0.036	3.57
EJ	0.029	2.86
LS	0.029	2.86
RB	0.028	2.86
BKN	0.021	2.14
CKJ	0.021	2.14
DK	0.021	2.14
LB1	0.021	2.14
LX	0.021	2.14
RKN	0.021	2.14
SCP	0.021	2.14
SI	0.021	2.14
SMI	0.021	2.14
ST0	0.021	2.14
TP	0.021	2.14

Table 5. Minimum Support 2-itemset

Transaction Items		Support	(%)
LK	STX	0.1	10
LK	RI	0.071	7.1
LK	CAK	0.057	5.7
STX	RI	0.036	3.6
STX	CAK	0.021	2.1
BF	BK	0.021	2.1
BF	BKN	0.021	2.1
CK	RB	0.021	2.1

4) Apriori association rules

The association rule is obtained by calculating using (4) based on the results of the 2-itemset value that satisfies the support and confidence values, the results can be seen in Table 7.

A total of 16 2-itemset association rules are obtained from the minimum support and minimum confidence values that have been determined. Furthermore, the results of the association rule with the 3-itemset are shown in Table 8. Based on the results of the analysis, 24 associations rules 3-itemset were obtained.

Table 6. Minimum Support 3-itemset

Transaction Items			Support	(%)
LK	STX	RI	0.036	3.60 %
LK	STX	CAK	0.021	2.10 %

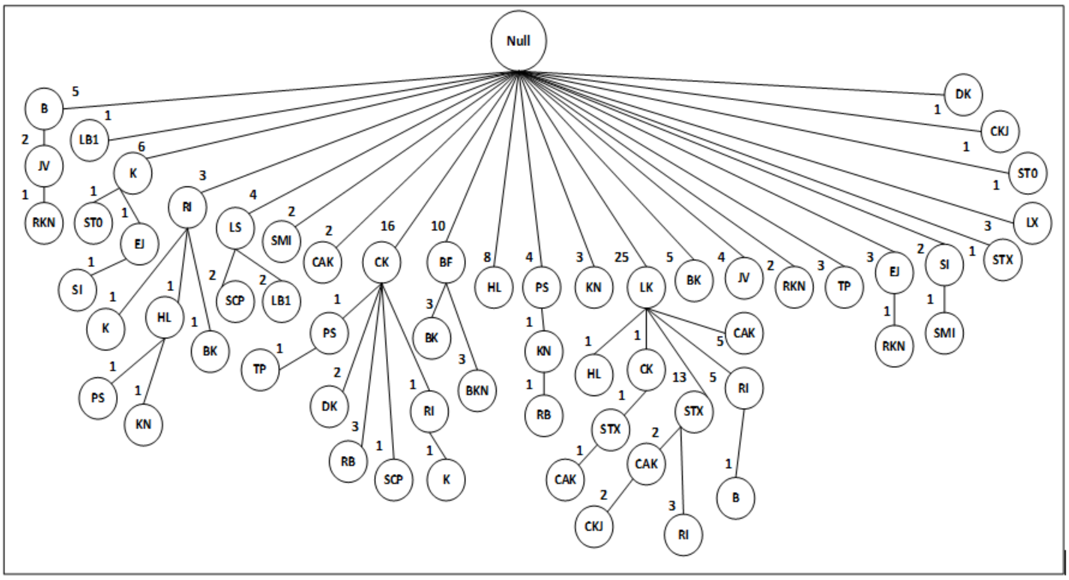

Fig. 2. FP-Tree formation results.

Table 7. Association Rules Apriori 2-itemset

No.	Rules	Support (%)	Confidence (%)
1	If buying LK → STX	10	56
2	If buying STX → LK	10	88
3	If buying LK → RI	7	40
4	If buying RI → LK	7	83
5	If buying LK → CAK	6	32
6	If buying CAK → LK	6	80
7	If buying STX → RI	4	31
8	If buying RI → STX	4	42
9	If buying STX → CAK	2	19
10	If buying CAK → STX	2	30
11	If buying BF → BK	2	33
12	If buying BK → BF	2	33
13	If buying BF → BKN	2	30
14	If buying BKN → BF	2	30
15	If buying CK → RB	2	18
16	If buying RB → CK	2	75

Table 8. Association Rules Apriori 3-itemset

No.	Rules	Supp-ort (%)	Confiden-ce (%)
1	If buying LK → STX ^ RI	2	20
2	If buying LK → RI ^ STX	2	20
3	If buying LK ^ STX → RI	2	20
4	If buying LK ^ RI → STX	2	20
5	If buying STX → LK ^ RI	2	31
6	If buying STX → RI ^ LK	2	31
7	If buying STX ^ LK → RI	2	31
8	If buying STX ^ RI → LK	2	31
9	If buying RI → LK ^ STX	2	42
10	If buying RI → STX ^ LK	2	42
11	If buying RI ^ STX → LK	2	42
12	If buying RI ^ LK → STX	2	42
13	If buying LK → STX ^ CAK	4	12
14	If buying LK → CAK ^ STX	4	12
15	If buying LK ^ CAK → STX	4	12
16	If buying LK ^ STX → CAK	4	12
17	If buying STX → LK ^ CAK	4	19
18	If buying STX → CAK ^ LK	4	19
19	If buying STX ^ CAK → LK	4	19
20	If buying STX ^ LK → CAK	4	19
21	If buying CAK → LK ^ STX	4	30
22	If buying CAK → STX ^ LK	4	30
23	If buying CAK ^ LK → STX	4	30
24	If buying CAK ^ STX → LK	4	30

D. FP-Growth Implementation

1) Creation of FP-Tree trees

From data that has met the minimum value of high-frequency support. Then the items are sorted by priority as seen from the largest frequency value, sorted according to priority as shown in Table 9. Next, the data that has been sorted by priority will be formed into an FP-Tree tree. It starts from Null and will be followed by an itemset in the transaction until all transactions form a tree as shown in Fig. 2.

Table 9. Item Transactions by Priority

No.	Transaction Code	Items
1	210105.001	K
2	210105.002	RI, K
3	210105.004	LS, SCP
4	210105.005	CAK
5	210105.006	CK, DK
6	210105.007	CK, RB
7	210105.008	BF, BK
8	210105.009	HL
9	210105.010	PS
10	210105.011	KN
...	...	...
122	210208.013	CK, PS, TP
123	210208.014	PS, KN, RB
124	210208.015	K, EJ, SI

2) Generating conditional pattern base

At this stage, the FP-Growth algorithm will break down the FP-Tree results in Fig. 2 based on each suffix and produce as mentioned in Table 9, and the item names will be used as code for the next step sorted from small to largest.

3) Generating conditional FP-Tree

To find the conditional FP-Tree, the step is to sum the number of existing supports, and the larger number of supports will replace them with the conditional FP-Tree. The support count is the number of occurrences of an itemset in all transactions in the dataset. Itemsets with a larger number of supports are considered more important because they appear more frequently in the

Table 10. Conditional Pattern Base

Item	Conditional Pattern Base
CKJ	{{LK, STX, CAK :2} {:1}}
BKN	{{BF:3}}
SCP	{{CK:1}, {LS:2}}
SI	{{K, EJ:1}, {:2}}
SMI	{{SI:1}, {:2}}
LX	{{:3}}
LB1	{{LS:2}, {:1}}
ST0	{{K:2}, {:1}}
RKN	{{B, JV:1}, {:2}}
TP	{{:3}}
DK	{{CK:2}, {:1}}
EJ	{{K:1}, {:3}}
LS	{{:4}}
RB	{{PS, KN:1}, {CK:3}}
JV	{{B, JV:2}, {:4}}
KN	{{PS:1}, {RI:1}, {:3}}
PS	{{CK:1}, {RI, HL:1}, {:4}}
B	{{LK, RI:1}, {:5}}
K	{{CK, RI:1}, {RI:1}, {:6}}
BK	{{RI:1}, {BF:3}, {:5}}
CAK	{{:2}, {LK, CK, STX:1}, {LK, STX:1}, {LK:7}}
BF	{{:10}}
HL	{{:8}, {RI:1}, {LK:1}}
RI	{{LK, STX:5}, {LK :7}, {CK:1}, {:3}}
STX	{{LK: 14}, {LK, CK :1}, {:1}}
CK	{{LK:1}, {:16}}
LK	{{:25}}

dataset. The results of the conditional FP-Tree are shown in Table 11. The result of the Conditional FP-Tree is determined by the calculation of items that meet the minimum support requirements in the conditional pattern base.

Table 11. Conditional FP-Tree

Item	Conditional Pattern Base	Conditional FP-Tree
BKN	{BF:3}	BF:3
RB	{{PS, KN:1}, {CK:3}}	CK:3
BK	{RI:1}, {BF:3}, {:5}}	BF:3
CAK	{{:2}, {LK, CK, STX:1}, {LK, STX:1}, {LK:7}}	LK:7
RI	{{LK, STX:5}, {LK :7}, {CK:1}, {:3}	LK: 8, STX: 3
STX	{{LK: 14}, {LK, CK :1}, {:1}	LK:14

4) Frequent itemset stages

The next step is to form a Frequent Itemset by combining sets and subsets of the Conditional FP-Tree with items. The results of the frequent itemset stage can be seen in Table 12, which is a collection of frequently occurring items that meet the minimum support value of Conditional FP-Tree. The result of frequent items will be found as the minimum value of support using the set of 2 items and the set of 3 items. To find the support value for a set of 2 items using (2), and for a set of 3 items using (3), the results of finding the support value for the 2-Itemset are shown in Table 13.

After a combination of 2-itemset. Furthermore, the combined analysis results of the 2-itemset will proceed to find the relationship with the 3-itemset set before calculating the association rules to be used in decision making by retail owners or related parties, then the results of the 3-itemset calculation are shown in Table 14.

Table 12. Frequent Itemset

Item	Conditional Pattern Base	Conditional FP-Tree	Frequent Itemset
BKN	{BF:3}	BF:3	{BF-BKN: 3}
RB	{{PS, KN:1}, {CK:3}}	CK:3	{CK-RB:3}
BK	{{RI:1}, {BF:3}, {:5}}	BF:3	{BF-BK:3}
CAK	{{:2}, {LK, CK, STX:1}, {LK, STX:2}, {LK:5}}	LK:5, STX:3	{LK-CAK:8, STX-CAK:3} {LK, STX, CAK:3}
RI	{{LK, STX:5}, {LK:7}, {CK:1}, {:3}	LK-STX:3, LK: 7	{LK, STX, RI:5} {LK, RI :10}
STX	{{LK: 14}, {LK, CK:1}, {:2}}	LK:14	{LK-STX:14}

Table 13. Support 2-itemset

Items	Transaction	Total Transactions	Support (%)
BF-BKN	3	140	2
CK-RB	3	140	2
BF-BK	3	140	2
LK-CAK	8	140	6
STX-CAK	3	140	2
LK-RI	7	140	5
STX-RI	3	140	2
LK-STX	14	140	10

The FP-Growth method can be used in determining the choice of product item pairs. The algorithm focuses on providing a choice of product item pairs based on the results of sales transactions, generating association rules.

5) FP-Growth association rules

From the results of itemsets that often appear, the association rules will be formed using (4) so as to produce trust values that can be used in decision making by store owners, the results of association rules with 2-itemsets are shown in Table 15.

Based on Table 15 FP-Growth algorithm obtains

Table 14. Support 3-itemset

Item	Transaction	Total Transactions	Support (%)
LK-STX-CAK	3	140	2
LK-STX-RI	5	140	4

Table 15. Association Rules FP-Growth 2-itemset

No.	Rules	Support (%)	Confidence (%)
1	If buying BF → BKN	0.02	0.3
2	If buying BKN → BF	0.02	1
3	If buying CK → RB	0.02	0.18
4	If buying RB → CK	0.02	0.75
5	If buying BF → BK	0.02	0.3
6	If buying BK → BF	0.02	0.33
7	If buying LK → CAK	0.06	0.32
8	If buying CAK → LK	0.06	0.8
9	If buying STX → CAK	0.02	0.19
10	If buying CAK → STX	0.02	0.30
11	If buying → RI	0.05	0.28
12	If buying RI → LK	0.05	0.58
13	If buying STX → RI	0.02	0.19
14	If buying RI → STX	0.02	0.25
15	If buying LK → STX	0.1	0.56
16	If buying → LK	0.1	0.88

16 association rules for 2-itemset, then the calculation of 3-itemset association rules is carried out using (3) and (4) results are shown in Table 16.

Table 16. Association Rules FP-Growth 3-itemset

No.	Rules	Support (%)	Confidence (%)
1	If buying LK → STXˆCAK	2	12
2	If buying LK → CAKˆSTX	2	12
3	If buying LKˆCAK → STX	2	12
4	If buying LKˆSTX → CAK	2	12
5	If buying STX → LKˆCAK	2	19
6	If buying STX → CAKˆLK	2	19
7	If buying STXˆCAK → LK	2	19
8	If buying STXˆLK → CAK	2	19
9	If buying CAK → LKˆSTX	2	30
10	If buying CAK → STXˆLK	2	30
11	If buying CAKˆLK → STX	2	30
12	If buying CAKˆSTX → LK	2	30
13	If buying LK → STXˆRI	4	12
14	If buying LK → RIˆSTX	4	20
15	If buying LKˆSTX → RI	4	20
16	If buying LKˆRI → STX	4	20
17	If buying STX → LKˆRI	4	20
18	If buying STX → RIˆLK	4	31
19	If buying STXˆLK → RI	4	31
20	If buying STXˆRI → LK	4	31
21	If buying RI → LKˆSTX	4	31
22	If buying RI → STXˆLK	4	42
23	If buying RIˆSTX → LK	4	42
24	If buying RIˆLK → STX	4	42

After calculation, 24 association rules were found for the 3-itemset of the dataset used. Furthermore, as an evaluation of the two algorithms, a lift ratio calculation will be carried out to determine the extent to which the rules are relevant to the existing data. The lift ratio is a useful metric to measure the extent to which an association is stronger than what is expected. Thus, an evaluation with a lift ratio will help determine whether the association rules found are really meaningful and can be used in further decision-making.

IV. Discussion

The calculations that have been carried out using apriori and FP-Growth algorithms aim to identify the rules of significant associations. Furthermore, for the evaluation of the association rules that have been found, the calculation of the lift ratio value will be carried out using (5). The use of lift ratio as an evaluation metric in association analysis is based on fundamental scientific considerations, namely, to measure the accuracy and significance of relationships between items in association rules. The results of calculating the lift ratio for the 2-itemset are shown in Table 17.

Results from apriori and FP-Growth algorithms for 2-itemsets show differences and similarities in the association rules found. Both algorithms produce the same number of association rules, which is 16 rules. There is a similarity in the rule with the highest lift ratio, where both find the rule ”If you buy BF, then buy BKN” with a lift ratio of 14. In addition, some other rules with a lift ratio of 5 are also found in both algorithms, such as ”If Buying STX then buying LK” and ”If Buying RI then buying LK”. However, differences also exist between the two. In the apriori algorithm, there is only one rule with the highest lift ratio, while in the FP-Growth algorithm there are two rules with the highest lift ratio. However, in the apriori algorithm, the lowest lift ratio is 3, while in the FP-Growth algorithm, there are two rules that get the lowest lift with a value of 2. Furthermore, the calculation of lift values for the 3-itemset hasip is shown in Table 18.

Both algorithms get the same lift ratio value, this shows that both are equally effective in extracting association rules from the same dataset. In this case, the analysis does show that the results of both algorithms are fully consistent, and there is no difference in the rule-associations generated by apriori and FP-Growth.

V. Conclusion

Based on the research that has been carried out, it can be concluded by applying a priori algorithm and FP-Growth has succeeded in finding a combination of 2-itemset and 3-itemsets. The results of the 2-itemset analysis found 16 association rules. Meanwhile, 3-itemsets get 24 association rules. This can be able to help store owners place goods according to the criteria of the item.

References

[1] E. Wanti, L. P. Maharrani, R. H. A. Prasetya, N. W. Tripustikasari, and G. N. Ikhtiagung, “Optimation economic order quantity method for a support system reorder point stock,” *Int. J. Electr. Comput. Eng.*, vol. 10, no. 5, pp. 4992-5000, 2020.

[2] K. Nissa and M. T. Siregar, “Analisis pengendalian persediaan bahan baku kain kemeja poloshirt menggunakan metode economic order quantity (EOQ) di PT bina busana internusa,” *Int. J. Soc. Sci. Bus.*, vol. 1, no. 4, pp. 271, 2017.

[3] 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.

[4] Fitriah, I. Riadi, and Herman, “Analisis data mining sistem inventory menggunakan algoritma apriori,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 3, no. 1, pp. 118–129, 2023.

[5] S. E. S. Careza Rizky and Y. Sudarso, “Analisis perbandingan metode EOQ dan metode POQ dengan metode min-max dalam pengendalian persediaan bahan baku pada PT sidomuncul pupuk nusantara,” *Admisi dan Bisnis*, vol. 17 No 1, no. ISSN 1411 – 4321, pp. 11–22, 2016.

[6] S. A. Rachmawati, L. Syafirullah, and M. N. Faiz, “Perancangan sistem pengendalian persediaan barang menggunakan metode EOQ dan ROP berbasis web,” *Semin. Nas. Terap. Ris. Inov. Ke-6*, vol. 6, no. 1, pp. 778–786, 2020.

[7] E. Fatma and D. S. Pulungan, “Analisis pengendalian persediaan menggunakan metode probabilistik dengan kebijakan backorder dan lost sales,” *J. Tek. Ind.*, vol. 19, no. 1, pp. 38, 2018.

[8] S. Bagui, K. Devulapalli, and J. Coffey, “A heuristic approach for load balancing the FP-growth algorithm on MapReduce,” *Array*, vol. 7, pp. 100035, 2020.

Table 17. Ratio 2-itemset Algorithm Apriori and FP-growth

No.	Algorithm			
	Apriori		FP-Growth	
	Rules	Lift Ratio	Rules	Lift Ratio
1	If buying LK → STX	5	If buying BF → BKN	14
2	If buying STX → LK	5	If buying BKN → BF	14
3	If buying LK → RI	5	If buying CK→ RB	6
4	If buying RI → LK	5	If buying RB → CK	6
5	If buying LK → CAK	4	If buying BF → BK	5
6	If buying CAK → LK	4	If buying BK → BF	5
7	If buying STX → RI	4	If buying LK → CAK	4
8	If buying RI → STX	4	If buying CAK → LK	4
9	If buying STX → CAK	3	If buying STX → CAK	3
10	If buying CAK → STX	3	If buying CAK → STX	3
11	If buying BF → BK	5	If buying LK → RI	3
12	If buying BK → BF	5	If buying RI → LK	3
13	If buying BF → BKN	14	If buying STX → RI	2
14	If buying BKN → BF	4	If buying RI → STX	2
15	If buying CK → RB	6	If buying LK → STX	5
16	If buying RB → CK	6	If buying STX → LK	5

Table 18. Ratio 2-itemset Algorithm Apriori and FP-growth

	Apriori		FP-Growth	
	Rules	Lift Ratio	Rules	Lift Ratio
1	If buying LK → STXˆCAK	6	If buying LK → STXˆCAK	6
2	If buying LK → CAKˆSTX	6	If buying LK → CAKˆSTX	6
3	If buying LKˆCAK → STX	2	If buying LKˆCAK → STX	2
4	If buying LKˆSTX → CAK	1	If buying LKˆSTX → CAK	1
5	If buying STX → LKˆCAK	2	If buying STX → LKˆCAK	2
6	If buying STX → CAKˆLK	3	If buying STX → CAKˆLK	3
7	If buying STXˆCAK → LK	9	If buying STXˆCAK → LK	9
8	If buying STXˆLK → CAK	2	If buying STXˆLK → CAK	2
9	If buying CAK → LKˆSTX	3	If buying CAK → LKˆSTX	3
10	If buying CAK → STXˆLK	5	If buying CAK → STXˆLK	5
11	If buying CAKˆLK → STX	5	If buying CAKˆLK → STX	5
12	If buying CAKˆSTX → LK	15	If buying CAKˆSTX → LK	15
13	If buying LK → STXˆRI	10	If buying LK → STXˆRI	10
14	If buying LK → RIˆSTX	10	If buying LK → RIˆSTX	10
15	If buying LKˆSTX → RI	2	If buying LKˆSTX → RI	2
16	If buying LKˆRI → STX	4	If buying LKˆRI → STX	4
17	If buying STX → LKˆRI	6	If buying STX → LKˆRI	6
18	If buying STX → RIˆLK	6	If buying STX → RIˆLK	6
19	If buying STXˆLK → RI	3	If buying STXˆLK → RI	3
20	If buying STXˆRI → LK	16	If buying STXˆRI → LK	16
21	If buying RI → LKˆSTX	4	If buying RI → LKˆSTX	4
22	If buying RI → STXˆLK	4	If buying RI → STXˆLK	4
23	If buying RIˆSTX → LK	21	If buying RIˆSTX → LK	21
24	If buying RIˆLK → STX	8	If buying RIˆLK → STX	8

[9] L. Linwei, W. Yiping, H. Yepiao, L. Bo, M. Fasheng, and D. Ziqiang, "Optimized apriori algorithm for deformation response analysis of landslide hazards," *Comput. Geosci.*, vol. 170, 2023.

[10] L. Zheng, "Research on e-commerce potential client mining applied to apriori association rule algorithm," in *Proc. - 2020 Int. Conf. Intell. Transp. Big Data Smart City, ICITBS 2020*, pp. 667–670, 2020.

[11] W. Thurachon and W. Kreesuradej, "Incremental association rule mining with a fast incremental updating frequent pattern growth algorithm," *IEEE Access*, vol. 9, pp. 55726–55741, 2021.

[12] M. Yasir, M. A. Habib, M. Ashraf, S. Sarwar, M. U. Chaudhry, M. Shahwani, M. Ahmad, CH. M. N. Faisal, "D-GENE: Deferring the GENEration of power sets for discovering frequent itemsets in sparse big data," *IEEE Access*, vol. 8, pp. 27375–27392, 2020.

[13] J. R. D. Arcos and A. A. Hernandez, "Efficient apriori algorithm using enhanced transaction reduction approach," in *2019 IEEE 13th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, 03-04 October 2019, Bali, Indonesia, pp. 97–101.

[14] M. Fauzy, K. R. Saleh W, and I. Asror, "Penerapan metode association rule menggunakan algoritma apriori pada simulasi prediksi hujan wilayah kota Bandung," *J. Ilm. Teknol. Inf. Terap.*, vol. II, no. 2, pp. 221–227, 2016.

[15] G. G. Prabawa, I. G. M. Darmawiguna, and I. M. A. Wirawan, "Pengembangan sistem pendukung keputusan pengendalian persediaan barang menggunakan metode economic order quantity (EOQ) dan min-max berbasis web (Studi kasus: Apotek sahabat qita)," *J. Nas. Pendidik. Tek. Inform.*, vol. 7, no. 2, pp. 107, 2019.

[16] S. Fauziah and Ratnawati, "Penerapan metode FIFO pada sistem informasi persediaan barang," *J. Tek. Komput.*, vol. 4, no. 1, pp. 98–108, 2018.

[17] E. Callens, "Financial instruments entail liabilities: Ether, bitcoin, and litecoin do not," *Comput. Law Secur. Rev.*, vol. 40, Apr 2021.

[18] P. Giudici and I. Abu-Hashish, "What determines bitcoin

exchange prices? A network VAR approach," *Financ. Res. Lett.*, vol. 28, pp. 309–318, Mar 2019.

[19] D. G. Anghel, "A reality check on trading rule performance in the cryptocurrency market: Machine learning vs. technical analysis," *Financ. Res. Lett.*, vol. 39, Mar 2021.

[20] B. Let, K. Sobanski, W. Swider, and K. Wlosik, "What drives the popularity of stablecoins? Measuring the frequency dynamics of connectedness between volatile and stable cryptocurrencies," *Technol. Forecast. Soc. Change*, vol. 189, 2023.

[21] S. Zhang and G. Mani, "Popular cryptoassets (bitcoin, ethereum, and dogecoin), gold, and their relationships: Volatility and correlation modeling," *Data Sci. Manag.*, vol. 4, pp. 30–39, Des 2021.

[22] L. J. Liebi, "Is there a value premium in cryptoasset markets?," *Econ. Model.*, vol. 109, Apr 2022.

[23] A. S. Hoong Lee, L. S. Yap, H. N. Chua, Y. C. Low, and M. A. Ismail, "A data mining approach to analyse crash injury severity level," *J. Eng. Sci. Technol.*, vol. 16, pp. 1–14, 2021.

[24] S. Wang, J. Cao, and P. S. Yu, "Deep learning for spatio-temporal data mining: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 1–21, 2019.

[25] A. De Vries, "Cryptocurrencies on the road to sustainability: Ethereum paving the way for Bitcoin," *Patterns*, vol. 4, no. 1. Cell Press, 13 Januari 2023.

[26] I. Chalkiadakis, A. Zaremba, G. W. Peters, and M. J. Chantler, "On-chain analytics for sentiment-driven statistical causality in cryptocurrencies," *Blockchain Res. Appl.*, vol. 3, no. 2, 2022.

[27] S. Syahriani, "Penerapan data mining untuk menentukan pola penjualan sepatu menggunakan metode algoritma apriori," *Bina Insa. Ict J.*, vol. 9, no. 1, pp. 43, 2022.

[28] F. Rahmawati and N. Merlina, "Metode data mining terhadap data penjualan sparepart mesin fotocopy menggunakan algoritma apriori," *PIKSEL Penelit. Ilmu Komput. Sist. Embed. Log.*, vol. 6, no. 1, pp. 9–20, 2018.

[29] Y. A. Ünvan, "Market basket analysis with association rules," *Commun. Stat. - Theory Methods*, vol. 50, no. 7, pp. 1615–1628, 2021.

[30] P. H. Agapito, G. Milano, M. Guzzi and M. Cannataro, "Mining association rules from disease ontology," in *Proc. - 2019 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2019*, pp. 2239–2243, 2019.

[31] G. Zhang, C. Liu, and M. Tao, "Data mining technology based on association rules algorithm," *Int. J. Mechatronics Appl. Mech.*, vol. 2019, no. 5, pp. 106–112, 2019.

[32] Z. Chen and B. Ye, "Association rule analysis of petrochemical fire accidents based on the apriori algorithm," in *2019 9th Int. Conf. Fire Sci. Fire Prot. Eng. ICFSFPE 2019*, pp. 0–4, 2019.

[33] E. Elisa, "Market basket analyis pada mini market ayu dengan algoritma apriori," *Jurnal Resti*, vol. 1, no. 2, 2018.

[34] A. W. O. Gama, I. K. G. D. Putra, and I. P. A. Bayupati, "Implementasi algoritma apriori untuk menemukan frequent itemset dalam keranjang belanja," *Maj. Ilm. Teknol. Elektro*, vol. 15, no. 2, pp. 21–26, 2016.

[35] A. Salam, J. Zeniarja, W. Wicaksono, and L. Kharisma, "Pencarian pola asosiasi untuk penataan barang dengan menggunakan perbandingan algoritma apriori dan Fp-Growth (study kasus distro epo store Pemalang)," *Dinamik*, vol. 23, no. 2, pp. 57–65, 2019.

[36] M. D. Febrianto and A. Supriyanto, "Implementasi algoritma apriori untuk menentukan pola pembelian produk," *Jurikom*, vol. 9, no. 6, pp. 2010–2020, 2022.

[37] F. AR. Mado, "Analisis persediaan bahan baku produk usaha sale pisang industri rumah tangga 'Sofie' di kota Palu," *e-J. Agrotekbis*, vol. 4, no. 2, pp. 204–209, 2016.

[38] W. N. Setyo and S. Wardhana, "Implementasi data mining pada penjualan produk di CV cahaya etya menggunakan algoritma fp-growth," *Petir*, vol. 12, no. 1, pp. 54–63, 2019,

[39] Z. Abidin, A. K. Amartya, and A. Nurdin, "Penerapan algoritma apriori pada penjualan suku cadang kendaraan roda dua (studi kasus: Toko prima motor sidomulyo)," *J. Teknoinfo*, vol. 16, no. 2, pp. 225, 2022.

[40] E. Irfiani, "Application of apriori algorithms to determine associations in outdoor sports equipment stores," *SinkrOn*, vol. 3, no. 2, pp. 218, 2019.

[41] L. Ardiantoro and N. Sunarmi, "Badminton player scouting analysis using frequent pattern growth (FP-growth) algorithm," *J. Phys. Conf. Ser.*, vol. 1456, no. 1, 2020.

[42] W. Apridonal M, Y. Choiriah, and A. Akmal, "Penerapan data mining menggunakan metode association rule dengan algoritma apriori untuk analisa pola penjualan barang," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 5, no. 2, pp. 193–198, 2019.

[43] S. Hu, Q. Liang, H. Qian, J. Weng, W. Zhou, and P. Lin, "Frequent-pattern growth algorithm based association rule mining method of public transport travel stability," *Int. J. Sustain. Transp.*, vol. 15, no. 11, pp. 879–892, 2021.

[44] E. Elisa and N. Azwanti, "Algoritma fp-growth untuk menganalisa frekuensi pembelian gas elpiji 3 kg," *INTENSIF J. Ilm. Penelit. dan Penerapan Teknol. Sist. Inf.*, vol. 3, no. 1, pp. 69, 2019.

[45] R. M. Anggraeni, "Perbandingan algoritma apriori dan algoritma fp-growth untuk perekomendasi pada transaksi peminjaman buku di perpustakaan Universitas Dian Nuswantoro," *Tek. Inform.*, pp. 1–6, 2014.