Matrix: Jurnal Manajemen Teknologi dan Informatika
Volume 13 Issue 2 Year 2023 Pages 52-67
DOI: http://dx.doi.org/10.31940/matrix.v13i2.52-67
URL: http://ojs2.pnb.ac.id/index.php/matrix

Determination of the best rule-based analysis results from the comparison of the Fp-Growth, Apriori, and TPQ-Apriori Algorithms for recommendation systems

***Moch Syahrir* [1]*, *Lalu Zazuli Azhar Mardedi* [2]**

[1,2] *Bumigora University, Indonesia*

*Corresponding Author: muhammadsyahriralfath@gmail.com

Abstract: The popular association rule algorithms are Apriori and fp-growth; both of these algorithms are very familiar among data mining researchers; however, there are some weaknesses found in the association rule algorithm, including long dataset scans in the process of finding the frequency of the item set, using large memory, and the resulting rules being sometimes less than optimal. In this study, the authors made a comparison of the fp-growth, Apriori, and TPQ-Apriori algorithms to analyze the rule results of the three algorithms. TPQ-Apriori is an algorithm developed from the Apriori algorithm. For experiments, the Apriori and fp-growth algorithms use RapidMiner and Weka tools, while the TPQ-apriori algorithm uses self-built application programs. The dataset used is the sales data for the Kopegtel NTB department store, which has been uploaded on the Kaggle site. As for the results of testing the base rules from the overall results of testing the rules with the good Kopegtel dataset for 100%, 50%, and 25% of the total volume of the dataset, a conclusion can be drawn that the larger the dataset to be processed, the results will be more optimal when using the fp-growth algorithm RapidMiner, but not optimal if the dataset to be processed is small. It is different from using the Apriori and Weka FP-growth algorithms, where the resulting rules are less than optimal if the dataset used is large and optimal if the dataset is small. Several rules do not appear in the fp-growth and Apriori Weka algorithms because the two algorithms do not have a tolerance value in Weka's tools for the support of the rules that will be displayed. Meanwhile, the TPQ-Apriori algorithm that has been developed is capable of producing optimal rules for both large datasets and small datasets.

Keywords: association rule, Fp-growth, apriori, TPQ-apriori, rapidminer

History Article: Submitted 11 April 2023 | Revised 13 May 2023 | Accepted 1 July 2023

How to Cite: M. Syahrir and L. Z. A. Mardedi, "Determination of the best rule-based analysis results from the comparison of the Fp-Growth, Apriori, and TPQ-Apriori Algorithms for recommendation systems", *Matrix: Jurnal Manajemen Teknologi Dan Informatika*, vol. 13, no. 2, pp. 52–67, 2023.

Introduction

Data mining is a scientific discipline that aims to extract knowledge and find patterns from large amounts of data by studying and developing algorithms [1], [2]. In its role, data mining consists of estimation, forecasting, classification, clustering, and association [3]. In using the five roles of data mining to extract knowledge and find patterns, they must be adjusted to the characteristics of the data, because the better the data that will be processed by the data mining algorithm, the better the results that will be obtained. The role of associations is widely implemented in business fields such as e-commerce, retail, and restaurants [4], [5], but it can also be applied to other fields such as software bug analysis and the analysis of biological and medical data [6], so we ensure that the benefits are quite widely used in various fields. Association, or what is commonly referred to in terms of data mining as the association rule," is one of the data mining techniques to search for and find a set of association rules between a combination of items [7]. Or, in another sense, it is an associative rule of the implications of a combination of relationships between an item [6], [8], and [9]. Commonly used and developed association rule algorithms are Apriori, Fp-Growth, Eclat, and Hash-Based [10], [11]. The resulting association rule algorithm, it is a rule that can be measured using support, confidence, lift ratio, leverage, conviction, and certainty factors. Support is the percentage combination of these items in the database; confidence is the strength of the relationship between items in the association rules; lift ratio is to test the value of the validity of the relationship between items; and leverage and conviction are to test how much influence and confidence there is between the antecedent and consequent.

The forerunner of the association algorithms is the Apriori algorithm, which was first presented in a seminar [12] and then repaired a year later [13]. In its implementation, Apriori can produce optimal rules, but the time used to scan datasets is very long because the approach used by Apriori in finding frequencies uses candidate generation, where all items must be traced to determine k-itemset candidates, where k-itemset means how many iterations will be generated from all traces and result in heavy use of memory. Repairs made by [14] with the fp- tree approach, which we now know as the fp-growth algorithm, are very good in terms of datasetscan time because it only does two scans of the dataset, but the rules produced by the fp-growth algorithm are not as optimal as Apriori, and also the memory usage is still large enough [15], which we now know as Frequent Pattern Growth (Fp-Growth). Then the last one that the author refers to is the TPQ-Apriori algorithm developed by [16]. This reference article has penetrated into one of t journals in Japan. The technique used in this research is to transform the horizontal format that has been used by the Apriori algorithm to a vertical format. This technique can reduce tuples so that the dataset dimensions are reduced, and the dataset will be processed into several partitions, which results in a faster dataset scan. The working process of the method proposed by the researcher [16].

To determine the state of the art of this research, the authors reviewed several papers relevant to the topics discussed. There are several journals that discuss the application and comparison of association rule algorithms. The first was in 2021, which was carried out by Aisyatul Maulidah and Fitra A. Bachtiar with the title "Application of the Association Rule Mining Method for the Ulsan Association for Aspects of Tourist Attractions" [7]. In this study, the aim was to find visitor recommendation patterns in Jatim Park 3, with 1067 Indonesian language review data. Comparison between Google reviews and the Apriori algorithm: the drawbacks are more to tourist recommendation reviews; there are no comparisons and details from the rule base for evaluating rule testing using support, confidence, and liftratio. The second was in 2022, which was carried out by Michael Henry et al. with the title "Implementation of an Apriori Algorithm for Music Genre" [17]. In this study, the aim of this research is that the pattern that is found can be a reference for music producers in terms of making or distributing their new music using the fp-growth algorithm with RapidMiner tools. Only using RapidMiner tools with the FP-algorithm growth without any comparison. Even though the dataset contains quite a lot of recorded data, there is no description of which part of the data was processed until it entered the FP-growth regression stage to evaluate rule testing using support and confidence.

The third was carried out by Rizky Wandri Anggi Hanafiah with the title "Analysis of Information Technology (IT) Goods Sales Patterns Using the FP-Growth Algorithm" [18]. In this study, the aim is to analyze marketing transaction patterns using the FP-growth algorithm with RapidMiner tools to look for correlation relationships to take policies, but the dataset used is only 70 transactions, and the lack of datasets used is very small. We find it difficult to prove that these results are maximized for the evaluation of rule testing using support and confidence.

Fourth in 2022, which was carried out by Komang Ardika Viantama and Painem with the title "Implementation of the Apriori Algorithm for Product Sales Analysis at Perjuangan Collection Stores" [19], in this study, the aim is to analyze sales transaction patterns at clothing retail stores using a web-based system using the Apriori algorithm. Disadvantages The application of Apriori algorithms based on the Algorithmmobile does not require analysis and comparison of algorithms. Only limited to making information systems for evaluating rule testing using support, confidence, and listrasio.

Lastly, in 2022, which was carried out by Zulham et al. with the title "Pattern Analysis of Drug Procurement System With FP-Growth Algorithm" [20], In this study, the aim was to analyze the correlation pattern of drug sales at the Medan Marela Health Center. Using Weka and Rapidman tools with the FP-growth algorithm Weaknesses in the dataset are not described in detail, even though they have been compared. preprocessing stages, and so forth, for the evaluation of rule testing using support and confidence.

Methodology

This section contains the stages of the research method. This stage is also used to explain the proposed solutions to research problems and to achieve the objectives of the research.

1. **Research design**
   The research method used is an experimental earch method, with research stages including dataset collection, data pre-processing (pre-processing), methods used for comparison, experimentation and method testing, and evaluation.
2. **Data Collection**
   Collection or collection of data is the initial stage in which it is carried out, where from existing problems, related data will be collected in the form of datasets obtained from minimarket telkom cooperative minimarket data in the city of Mataram, NTB Indonesia, The dataset used that can be downloaded is the dataset obtained from the telecom employee cooperative minimarket in the city of Mataram-NTB. The dataset has been uploaded on the website www.kaggle.com for more details and can be downloaded at https://www.kaggle.com/datasets/syahrir12345678/datasetjualbelikopegtelntb.
3. **Initial Data Processing (Pre-processing)**
   Before entering the data algorithm model, it must be pre-processed to ensure the format that will enter the model is as expected. The better the data to be processed, the more optimal the results of the algorithmic process will be. In this study, the pre-processing process changes the format to the required form.
4. **Method Comparison**
   For comparison, we use three methods, namely the Apriori algorithm, the fp-growth algorithm, and the TPQ-apriori algorithm, while the tools used are Weka, Rapidminer, and self-designed application programs. There are three association rule algorithms that will be used in this study to analyze the rules formed by comparing the three algorithms.
5. **Apriori Algorithm**
   The apriori algorithm is one of the association rule algorithms with data collection techniques using an associative rule approach to determine the association relationship of an item combination. The importance of an associative rule can be determined by two parameters, namely support and confidence [6], [21]. Support (supporting value) is the percentage of that item combination e database. Confidence (and certainty value) is the strength of the relationship between items in the association rules. An association rule is said to be interesting if the support value is greater than the minimum support and the confidence value is greater than the minimum confidence. Meanwhile, to test the Three association rule algorithms willft ratio.
   a. Support Formulas:
   $$\text{Support}(A, B) = \frac{Jml\ TransaksiAdanB}{\sum Transaksi} x100\% \tag{1}$$
   b. Confidence Formula:
   $$Confidence(A, B) = \frac{JmlTransaksiAdanB}{\sum Transaksi\ A} x100\% \tag{2}$$
   c. LiftRatio Formula:
   $$Liftrasio = \frac{Support(A, B)}{SupportAxSupportB} \tag{3}$$
6. **Fp-Growth Algorithm**
   The fp-growth algorithm was developed from the Apriori algorithm; of course, the two are complementary. The fp-growth in the search process for item frequency is very good from the Apriori algorithm, but the resulting rules are not as good as the Apriori algorithm, and the memory usage is still quite large. The fp-growth algorithm is an algorithm from the association rules technique that can be used to determine the most frequently occurring data set (frequent itemset) in a data set by approaching the fp-tree concept [22].

7. TPQ-Apriori Algorithm

TPQ stands for Tid-list Vertical Partitioning Query and uses a vertical tid-list format technique with a query-based partitioning system. The algorithm is claimed to be able to process dataset scans in a frequency search for itemsets and is able to generate optimal rules [16]. The initial stage is carried out, namely pre-processing to adjust the format as needed with the dot SQL (.sql) extension. In the next stage, the dataset is partitioned, and each partition will be applied to a vertical tiling list approach. The goal with this approach is to trim the records. We can see the complete flow of the TPQ-Apriori algorithm in Figure 1.

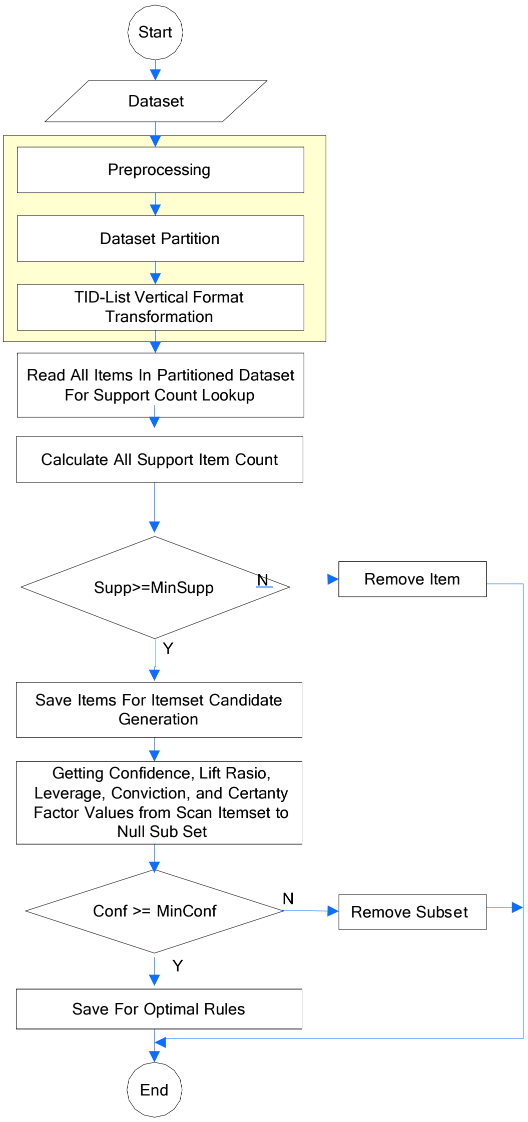


Figure 1. TPQ-Apriori flowchart

8. Experiment and Testing Methods

The experimental stages of the proposed method are:

a. Setting up datasets
b. Comparative analysis of the results of TPQ-Apriori and Fp-Growth in the rapidminner, Fp-Growth weka, and Apriori weka tools using the Telkom employee cooperative minimarket dataset in Mataram City-NTB

For rule testing, Fp-Growth will be compared in RapidMiner and Weka Tools, as well as Apriori in Weka Tools, while the dataset for rule testing is the Telkom employee cooperative minimarket dataset in Mataram (NTB). The dataset can be downloaded.

https://www.kaggle.com/datasets/syahrir12345678/datasetjualbelikopegtelntb. In this study, tests were carried out using the TPQ-Apriori tools, which were developed by themselves using Microsoft Visual Studio 2013 and MySQL Server 5.5 tools, which support parallelization.

9. Evaluation

The evaluation process will provide results regarding the value of support, confidence, lift ratio, conviction, leverage, and certainty factors. Support determines how often the rule is applied in the dataset [6], [21], [23], and [24]. Support is an indication of how often the item set appears in the dataset. Support can be formulated like Formula 5.

$$Support, s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \quad (5)$$

Where :
N = Transaction Totals
X = Antecedent
Y = Consequent

Confidence determinesof frequency of items in Y appearing in transactions that contain X. Confidence or belief in how often the rule or rules are proven to be true. Confidence can be formulated as in Formula 6.

$$confidence, c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{X} \quad (6)$$

To test whether a rule or a relationship between items is valid or not, a lift ratio is used. Lift ratio is a value that measures the magnitude of the relationship between the antecedent and consequent that is independent [6], [21], [23], and [24]. The lift ratio has a range from 0 to 1. Values close to 1 indicate that the antecedent and consequent have no dependence. Values far from 1 indicate that the antecedent provides information about the consequent. Or, with another understanding, if the lift is > 1, it lets us know to what extent two events are dependent on each other, which makes the rule potentially useful for predicting the consequent in a dataset. And if lift is 1, that lets us know that the items replace each other. This means that the presence of one item has a negative effect on the presence of another item, and vice versa.

$$lift(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X) * \sigma(Y)} \quad (7)$$

Where :
∞ = Infinity Or Null
σ = Support Count
U = Relate

Apart from the lift ratio, we can test it with another formula, namely conviction or the value of belief [6], [25], [23], [24]. Conviction is a value that measures the degree of implication of a rule. Conviction is very concerned about the direction of an association rule. Conviction indicates that conviction (X→Y) ≠ conviction (Y→X).

$$conv(X \rightarrow Y) = \frac{1 - \sigma(Y)}{1 - c(X \rightarrow Y)} \quad (8)$$

Forthe number of antecedent and consequence items sold simultaneously in a dataset is more than we would expect using Leverage. A value of 0 indicates antecedent and consequent independent. Leverage has a value range from -0.25 to 0.25. Can be formulated as in formula 9.

$$lev(X \rightarrow Y) = \sigma(X \cup Y) - (\sigma(X) * \sigma(Y)) \quad (9)$$

In this study, additional evaluation is needed to make a decision about assessing the relationship or correlation rule that is formed using the certainty factor method. The certainty factor method, according to David McAllister "is a method for proving whether a fact is certain or uncertain in form, a metric that is usually used in expert systems".

$$CF(X \rightarrow Y) = \frac{c(X \rightarrow Y) - \sigma(Y)}{1 - \sigma(Y)} \tag{10}$$

Results and Discussions

Experiments to overcome the problems of scanning old datasets, the rules that are formed, and the use of memory and processor, which is still quite large, by integrating the TID-List Vertical approach and data partitioning Where data partitioning is used to partition datasets so that the dataset volume can be partitioned to a smaller size than the original dataset. And for each dataset that has been partitioned, the initial table format partition results will be transformed to a vertical form, and with these two approaches, the frequency search process in the developed Apriori algorithm becomes faster. The runtime testing process is repeated 10 times to ensure consistent dataset scan times.

The following is a comparison of the results of testing the Rule Apriori Tools Weka, Fp-Growth Tools Rapidminer, and TPQ-Apriori. The results of testing the three algorithms in different tools have been obtained. To simplify the comparison, the results are put together in tabular form. As for the tabular form for the section:

1. TPQ-Apriori algorithm on TPQ-Apriori tools, with rapidminer's Fp-Growth Algorithm for 100% of Kopegtel datasets. Can be seen in Table 1.
2. TPQ-Apriori Algorithm on TPQ-Apriori tools, with Weka's Fp-Growth Algorithm for 100% Kopegtel dataset. Can be seen in Table 2.
3. TPQ-Apriori algorithm on TPQ-Apriori tools, with Weka's Apriori Algorithm for 100% of Kopegtel datasets. Can be seen in Table 3.
4. The TPQ-Apriori algorithm on the TPQ-Apriori tools, with the rapidminer Fp-Growth Algorithm for 50% of the Kopegtel dataset. Can be seen in Table 4.
5. TPQ-Apriori algorithm on TPQ-Apriori tools, with Weka's Fp-Growth Algorithm for 50% of the Kopegtel dataset. Can be seen in Table 5.
6. TPQ-Apriori algorithm on TPQ-Apriori tools, with Weka's Apriori algorithm for 50% of the Kopegtel dataset. Can be seen in Table 6.
7. TPQ-Apriori algorithm on TPQ-Apriori tools, with rapidminer's Fp-Growth Algorithm for 25% of the Kopegtel dataset. Can be seen in Table 7.
8. TPQ-Apriori algorithm on TPQ-Apriori tools, with Weka's Fp-Growth Algorithm for 25% of the Kopegtel dataset. Can be seen in Table 8.
9. The TPQ-Apriori algorithm on the TPQ-Apriori tools, with the Weka Apriori Algorithm for 25% of the Kopegtel dataset. Can be seen in Table 9.

Table 1. Results of the TPQ-Apriori algorithm rule with fp-growth rapidminer for 100% of the Kopegtel dataset

No	Antecedent	Consequent	Item set	Freq	Supp	TPQ-Apriori				Fp-Growth Rapidminer			
						Conf	Lift	Conv	Leve	CF	Conf	Lift	Conv
1	MIE 89, SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	4	48	11.62%	1.00	3.06	∞	0.08	1.00	1.00	3.06	∞
2	KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	3	67	16.22%	0.99	3.01	67.31	0.11	0.99	0.99	3.01	45.77
3	MIE 89, KUDA MAS MAKARONI	KUDA MAS STIK	3	64	15.50%	0.98	3.01	33.66	0.10	0.97	0.98	3.01	43.75
4	SOSIS SONICE SAPI, KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	4	49	11.86%	0.98	3.00	33.66	0.08	0.97	0.98	3.00	33.66
5	KUDA MAS STIK, KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	4	43	10.41%	0.98	3.13	34.38	0.07	0.97	0.98	3.13	30.26

No	Antecedent	Consequent	Item set	Freq	Supp	TPQ-Apriori				Fp-Growth Rapidminer			
						Conf	Lift	Conv	Leve	CF	Con f	Lift	Conv
6	MIE 89, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97	0.98	2.99	29.62
7	KUDA MAS MAKARONI, KELAPA MUDA,JAGOAN NEON	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97	0.98	2.99	29.62
8	KUDA MAS MAKARONI	KUDA MAS STIK	2	122	29.54%	0.95	2.89	13.46	0.19	0.93	0.95	2.89	12.40
9	KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	3	69	16.71%	0.95	2.89	13.46	0.11	0.93	0.95	2.89	12.28
10	SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	3	80	19.37%	0.95	2.91	13.46	0.13	0.93	0.95	2.91	14.14
11	KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	3	69	16.71%	0.95	3.03	13.75	0.11	0.93	0.95	3.03	12.55
12	KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	3	67	16.22%	0.94	3.02	11.46	0.11	0.91	0.94	3.02	12.21
13	SOSIS SONICE SAPI, KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	4	49	11.86%	0.94	3.02	11.46	0.08	0.91	0.94	3.02	11.92
14	SOSIS SONICE SAPI, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	46	11.14%	0.94	2.87	11.22	0.07	0.91	0.94	2.87	10.99
15	MIE 89, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	43	10.41%	0.93	2.99	9.82	0.07	0.90	0.93	2.99	10.54
16	SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	3	80	19.37%	0.92	2.94	8.60	0.13	0.88	0.92	2.94	8.55
17	SOSIS SONICE SAPI, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	46	11.14%	0.92	2.95	8.60	0.07	0.88	0.92	2.95	8.60
18	KUDA MAS STIK	KUDA MAS MAKARONI	2	122	29.54%	0.90	2.89	6.88	0.19	0.85	0.90	2.89	7.14
19	MIE 89, SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	4	48	11.62%	0.89	2.85	6.25	0.08	0.84	0.89	2.85	6.19
20	MIE 89, KUDA MAS STIK	KUDA MAS MAKARONI	3	64	15.50%	0.88	2.81	5.73	0.10	0.83	0.88	2.81	5.58
21	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK	3	44	10.65%	0.88	2.69	5.61	0.07	0.82	0.88	2.69	5.61
22	KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	3	44	10.65%	0.88	2.82	5.73	0.07	0.83	0.88	2.82	5.73
23	MIE 89, JAGOAN NEON	KUDA MAS STIK	3	43	10.41%	0.88	2.68	5.61	0.07	0.82	0.88	2.68	5.50
24	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK, KUDA MAS MAKARONI	4	43	10.41%	0.86	2.91	5.03	0.07	0.80	0.86	2.91	5.03
25	MIE 89, TEPUNG WHITE BEAR	GULA PTPN	3	84	20.34%	0.85	1.98	3.81	0.10	0.74	0.85	1.98	3.77
26	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS STIK	3	52	12.59%	0.84	2.57	4.21	0.08	0.76	0.84	2.57	4.17
27	MIE 89, GULA PTPN	TEPUNG WHITE BEAR	3	84	20.34%	0.82	1.99	3.26	0.10	0.69	0.82	1.99	3.32
28	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS MAKARONI	3	50	12.11%	0.81	2.58	3.62	0.07	0.72	0.81	2.58	3.55
29	TEPUNG WHITE BEAR	GULA PTPN	2	137	33.17%	0.80	1.87	2.86	0.15	0.65	0.80	1.87	2.87

From the analysis results in Table 1, all evaluation values between the Fp-Growth rapidminer algorithm and the algorithm proposed in the TPQ-Apriori tools are all the same starting from confidence, liftration, and leverage, but there are some differences in the conviction values from the test, but the values are close. The rules generated by the two algorithms are 29 rules; if we refer to the threshold that we set, namely the minimum support of 0.1 and the minimum confidence of 0.8, the resulting rule is very optimal. In the TPQ-Apriori tools developed in this study, there is an additional formula for evaluating rules other than lift ratio, leverage, and conviction, namely the certainty factor. The certainty factor is one of the methods used to make it easier for us to understand the results of evaluating the possibility of determining whether or not a rule applies. This makes it easier for us to assess that the antecedent and consequent relationships in the rule have useful links and information. The calculation of conviction refers to the standard conviction formula to ensure that the conviction value in TPQ-Apriori is correct.

Table 2. Results of the tpq-Apriori algorithm rule with fp-growth weka for 100% of the Kopegtel dataset

No	Antecedent	Consequent	Item set	Freq	Supp	TPQ-Apriori					Fp-Growth Weka		
						Conf	Lift	Conv	Leve	CF	Conf	Lift	Conv
1	MIE 89, SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	4	48	11.62%	1.00	3.06	∞	0.08	1.00			
2	KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	3	67	16.22%	0.99	3.01	67.31	0.11	0.99	0.99	3.01	45.77
3	MIE 89, KUDA MAS MAKARONI	KUDA MAS STIK	3	64	15.50%	0.98	3.01	33.66	0.10	0.97			
4	SOSIS SONICE SAPI, KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	4	49	11.86%	0.98	3.00	33.66	0.08	0.97	0.98	3.00	33.66
5	KUDA MAS STIK, KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	4	43	10.41%	0.98	3.13	34.38	0.07	0.97	0.98	3.13	30.26
6	MIE 89, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97			
7	KUDA MAS MAKARONI, KELAPA MUDA,JAGOAN NEON	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97	0.98	2.99	29.62
8	KUDA MAS MAKARONI	KUDA MAS STIK	2	122	29.54%	0.95	2.89	13.46	0.19	0.93	0.95	2.89	12.40
9	KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	3	69	16.71%	0.95	2.89	13.46	0.11	0.93	0.95	2.89	12.28
10	SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	3	80	19.37%	0.95	2.91	13.46	0.13	0.93	0.95	2.91	14.14
11	KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	3	69	16.71%	0.95	3.03	13.75	0.11	0.93	0.95	3.03	12.55
12	KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	3	67	16.22%	0.94	3.02	11.46	0.11	0.91	0.94	3.02	12.21
13	SOSIS SONICE SAPI, KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	4	49	11.86%	0.94	3.02	11.46	0.08	0.91	0.94	3.02	11.92
14	SOSIS SONICE SAPI, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	46	11.14%	0.94	2.87	11.22	0.07	0.91	0.94	2.87	10.99
15	MIE 89, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	43	10.41%	0.93	2.99	9.82	0.07	0.90			
16	SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	3	80	19.37%	0.92	2.94	8.60	0.13	0.88	0.92	2.94	8.55
17	SOSIS SONICE SAPI, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	46	11.14%	0.92	2.95	8.60	0.07	0.88	0.92	2.95	8.60
18	KUDA MAS STIK	KUDA MAS MAKARONI	2	122	29.54%	0.90	2.89	6.88	0.19	0.85	0.90	2.89	7.14
19	MIE 89, SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	4	48	11.62%	0.89	2.85	6.25	0.08	0.84			
20	MIE 89, KUDA MAS STIK	KUDA MAS MAKARONI	3	64	15.50%	0.88	2.81	5.73	0.10	0.83			
21	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK	3	44	10.65%	0.88	2.69	5.61	0.07	0.82	0.88	2.69	5.61
22	KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	3	44	10.65%	0.88	2.82	5.73	0.07	0.83	0.88	2.82	5.73
23	MIE 89, JAGOAN NEON	KUDA MAS STIK	3	43	10.41%	0.88	2.68	5.61	0.07	0.82			
24	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK, KUDA MAS MAKARONI	4	43	10.41%	0.86	2.91	5.03	0.07	0.80	0.86	2.91	5.03
25	MIE 89, TEPUNG WHITE BEAR	GULA PTPN	3	84	20.34%	0.85	1.98	3.81	0.10	0.74			
26	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS STIK	3	52	12.59%	0.84	2.57	4.21	0.08	0.76	0.84	2.57	4.17
27	MIE 89, GULA PTPN	TEPUNG WHITE BEAR	3	84	20.34%	0.82	1.99	3.26	0.10	0.69			

No	Antecedent	Consequent	Item set	Freq	Supp	TPQ-Apriori					Fp-Growth Weka		
						Conf	Lift	Conv	Leve	CF	Conf	Lift	Conv
28	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS MAKARONI	3	50	12.11%	0.81	2.58	3.62	0.07	0.72	0.81	2.58	3.55
29	TEPUNG WHITE BEAR	GULA PTPN	2	137	33.17%	0.80	1.87	2.86	0.15	0.65	0.80	1.87	2.87

From the test results in Table 2, there are 20 rules generated by Fp-Growth Weka. If we refer to the threshold set, namely the minimum support 0.1 and minimum confidence 0.8, the rules produced by the Fp-Growth algorithm in Weka tools are not optimal because there are 9 rules not found. And if we analyze the results from Table 6, what causes rule 9 not to be found is that of the 9 rules, all the rules that are not found are rules that have items or attributes with the name "MIE 89." The item or attribute "MIE 89" in the Kopegtel dataset is the item that has the highest frequency in the first iteration, namely 207 occurrences. In its implementation, when referring to the Fp-Growth algorithm developed by Jiawei Han, the item with the highest support count or with the highest frequency in the first iteration will be the initial node, and this initial node is not included in the path formed in the Fp-tree. For information, the column marked in orange is a rule that was not found.

Table 3. Results of the tpq-Apriori algorithm with Apriori weka for 100% of the Kopegtel dataset

No	Antecedent	Consequent	Itemset	Freq	Supp	TPQ-Apriori					Apriori Weka	
						Conf	Lift	Conv	Leve	CF	Supp	Conf
1	MIE 89, SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	4	48	11.62%	1.00	3.06	Infinity	0.08	1.00		
2	KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	3	67	16.22%	0.99	3.01	67.31	0.11	0.99	16.22%	0.99
3	MIE 89, KUDA MAS MAKARONI	KUDA MAS STIK	3	64	15.50%	0.98	3.01	33.66	0.10	0.97	15.50%	0.91
4	SOSIS SONICE SAPI, KUDA MAS MAKARONI, JAGOAN NEON	KUDA MAS STIK	4	49	11.86%	0.98	3.00	33.66	0.08	0.97	11.86%	0.98
5	KUDA MAS STIK, KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	4	43	10.41%	0.98	3.13	34.38	0.07	0.97	10.41%	0.98
6	MIE 89, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97		
7	KUDA MAS MAKARONI, KELAPA MUDA,JAGOAN NEON	KUDA MAS STIK	4	43	10.41%	0.98	2.99	33.66	0.07	0.97	10.41%	0.98
8	KUDA MAS MAKARONI	KUDA MAS STIK	2	122	29.54%	0.95	2.89	13.46	0.19	0.93	29.54%	0.95
9	KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	3	69	16.71%	0.95	2.89	13.46	0.11	0.93	16.71%	0.95
10	SOSIS SONICE SAPI, KUDA MAS MAKARONI	KUDA MAS STIK	3	80	19.37%	0.95	2.91	13.46	0.13	0.93	19.37%	0.95
11	KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	3	69	16.71%	0.95	3.03	13.75	0.11	0.93	16.71%	0.95
12	KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	3	67	16.22%	0.94	3.02	11.46	0.11	0.91	16.22%	0.94
13	SOSIS SONICE SAPI, KUDA MAS STIK, JAGOAN NEON	KUDA MAS MAKARONI	4	49	11.86%	0.94	3.02	11.46	0.08	0.91	11.86%	0.94
14	SOSIS SONICE SAPI, KUDA MAS MAKARONI, KELAPA MUDA	KUDA MAS STIK	4	46	11.14%	0.94	2.87	11.22	0.07	0.91	11.14%	0.94
15	MIE 89, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	43	10.41%	0.93	2.99	9.82	0.07	0.90		
16	SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	3	80	19.37%	0.92	2.94	8.60	0.13	0.88	19.37%	0.92
17	SOSIS SONICE SAPI, KUDA MAS STIK, KELAPA MUDA	KUDA MAS MAKARONI	4	46	11.14%	0.92	2.95	8.60	0.07	0.88	11.14%	0.92
18	KUDA MAS STIK	KUDA MAS MAKARONI	2	122	29.54%	0.90	2.89	6.88	0.19	0.85	29.54%	0.90
19	MIE 89, SOSIS SONICE SAPI, KUDA MAS STIK	KUDA MAS MAKARONI	4	48	11.62%	0.89	2.85	6.25	0.08	0.84		
20	MIE 89, KUDA MAS STIK	KUDA MAS MAKARONI	3	64	15.50%	0.88	2.81	5.73	0.10	0.83	15.50%	0.94

No	Antecedent	Consequent	Itemset	Freq	Supp	TPQ-Apriori					Apriori Weka	
						Conf	Lift	Conv	Leve	CF	Supp	Conf
21	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK	3	44	10.65%	0.88	2.69	5.61	0.07	0.82	10.65%	0.88
22	KELAPA MUDA, JAGOAN NEON	KUDA MAS MAKARONI	3	44	10.65%	0.88	2.82	5.73	0.07	0.83	10.65%	0.88
23	MIE 89, JAGOAN NEON	KUDA MAS STIK	3	43	10.41%	0.88	2.68	5.61	0.07	0.82		
24	KELAPA MUDA, JAGOAN NEON	KUDA MAS STIK, KUDA MAS MAKARONI	4	43	10.41%	0.86	2.91	5.03	0.07	0.80	10.41%	0.86
25	MIE 89, TEPUNG WHITE BEAR	GULA PTPN	3	84	20.34%	0.85	1.98	3.81	0.10	0.74		
26	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS STIK	3	52	12.59%	0.84	2.57	4.21	0.08	0.76	12.59%	0.84
27	MIE 89, GULA PTPN	TEPUNG WHITE BEAR	3	84	20.34%	0.82	1.99	3.26	0.10	0.69		
28	SOSIS SONICE SAPI, JAGOAN NEON	KUDA MAS MAKARONI	3	50	12.11%	0.81	2.58	3.62	0.07	0.72	12.11%	0.81
29	TEPUNG WHITE BEAR	GULA PTPN	2	137	33.17%	0.80	1.87	2.86	0.15	0.65	33.17%	0.80

From the results of testing the rules in Table 3, the rules produced by the Apriori Weka algorithm are 22. Referring to the specified threshold, namely a minimum support of 0.1 and a minimum confidence of 0.80, the resulting rules are not optimal because there are 7 rules that are not found, although they are slightly better than Fp-Growth Weka because there are 2 additional rules to be found. However, this rule has a different trust value from that produced by TPQ-Apriori; see Table 7 in rows 3 and 20, which are colored orange. Now we prove the correct confidence value. In Table 3, row 3, pay attention to the rule "MIE 89, KUDA MAS MACARONI => KUDA MAS STIK." The total support count or frequency is 64, while Transactions for the Kopegtel dataset of 100% volume is 413 transactions, so the result is 0. We multiply 1549 by 100% to get 15.49%, while to calculate the confidence, we divide 64 by the number of support or frequency of the antecedents "MIE 89, KUDA MAS MAKARONI", The support for the antecedent count "MIE 89, KUDA MAS MAKARONI" in the Kopegtel dataset is 65, so the result is 0.98. Thus, the value of trust in the TPQ-Apriori is correct, while the result of the Apriori Weka is 0.91,which is not quite right when referring to the trust formula. This value is also strengthened by the results of the Fp-Growth Rapid Miner. See Table 1, row 3, where the confidence value is 0.98. For information, the orange color is a rule that was not found, and the green color is a miscalculation of the confidence value. The number of support antecedents "MIE 89, KUDA MASMAKARONI" in the Kopegtel dataset is 65, so the result is 0.98, thus the trust value in TPQ-Aprioriis correct, while the Apriori Weka result is 0.91, which is not quite right when referring to the formula trust.

This value is also strengthened by the results of the Fp-Growth Rapid Miner. See Table 1, row 3, where, the confidence value is 0.98. For information, the orange color is a rule that was not found,d and the green color is a miscalculation of the confidence value. The number of support antecedents "MIE 89, KUDA MAS MAKARONI" in the Kopegtel dataset is 65, so the result is 0.98, thus the trust value in TPQ-Apriori is correct, while the Apriori Weka result is 0.91, which is not quite right when referring to the formula trust. This value is also strengthened by the results of the Fp-Growth Rapid Miner. See Table 1, row 3, where, the confidence value is 0.98. For information, the orange color is a rule that was not found, and the green color is a miscalculationof the confidence value. Because the implementation of the formula used is support 64 divided by 314 and confidence 64 divided by 65.

Table 4. Results of the tpq-Apriori algorithm rule with fp-growth rapidminer for 50% of the Kopegtel dataset

No	Antecdent	Consequent	Freq	TPQ-Apriori						Fp-Growth Rapidminer			
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift	Conv
1	AMARTA COKLAT KACANG	AMARTA CHORY	29	0.10	0.97	8.26	29.43	0.09	0.97				
2	KOPYOR MANGGA	KOPYOR MELON	40	0.13	0.91	5.66	9.33	0.11	0.89	0.13	0.91	5.66	9.23

No	Antecdent	Consequent	Freq	TPQ-Apriori						Fp-Growth Rapidminer			
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift	Conv
3	KOPYOR MELON	KOPYOR MANGGA	40	0.13	0.83	5.66	5.02	0.11	0.80	0.13	0.83	5.66	5.12
4	AMARTA CHORY	AMARTA COKLAT KACANG	29	0.10	0.83	8.26	5.29	0.09	0.81				
5	KUDA MAS MAKARONI	KUDA MAS STIK	37	0.12	0.82	4.24	4.48	0.09	0.78	0.12	0.82	4.24	4.53
6	FRENTA STROBERI	FRENTA ANGGUR	30	0.10	0.81	6.06	4.56	0.08	0.78	0.10	0.81	6.06	4.58
7	TEPUNG BERAS ROSE BRAND	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.12	0.90	0.00	0.00
8	TEPUNG WHITE BEAR	TELER JUMBO	8	0.03	0.00	0.00	0.00	0.00	0.00	0.23	0.90	0.00	0.00
9	GULA PTPN	TELER JUMBO	9	0.03	0.00	0.00	0.00	0.00	0.00	0.24	0.89	0.00	0.00
10	GULA PTPN, TEPUNG WHITE BEAR	TELER JUMBO	6	0.02	0.00	0.00	0.00	0.00	0.00	0.14	0.87	0.00	0.00
11	MIE 89, GULA PTPN	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.11	0.87	0.00	0.00
12	MIE 89	TELER JUMBO	14	0.05	0.00	0.00	0.00	0.00	0.00	0.31	0.87	0.00	0.00
13	TEPUNG KETAN ROSE BRAND	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.11	0.86	0.00	0.00
14	KUDA MAS MAKARONI	TELER JUMBO	8	0.03	0.00	0.00	0.00	0.00	0.00	0.12	0.82	0.00	0.00
15	FRENTA STROBERI	TELER JUMBO	7	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.81	0.00	0.00
16	SOSIS SONICE SAPI	TELER JUMBO	17	0.06	0.00	0.00	0.00	0.00	0.00	0.23	0.80	0.00	0.00

Table 5. Results of the Tpq-Apriori Algorithm rule with Fp-Growth Weka for 50% of the Kopegtel dataset

No	Antecdent	Consequent	Freq	TPQ-Apriori						Fp-GrowthWeka			
				Supp	Conf	Lift	Conv	Leve	CF	Conf	Lift	Conv	Leve
1	AMARTA COKLAT KACANG	AMARTA CHORY	29	0.10	0.97	8.26	29.43	0.09	0.97				
2	KOPYOR MANGGA	KOPYOR MELON	40	0.13	0.91	5.66	9.33	0.11	0.89	0.91	5.66	7.39	0.11
3	KOPYOR MELON	KOPYOR MANGGA	40	0.13	0.83	5.66	5.02	0.11	0.80	0.83	5.66	4.55	0.11
4	AMARTA CHORY	AMARTA COKLAT KACANG	29	0.10	0.83	8.26	5.29	0.09	0.81				
5	KUDA MAS MAKARONI	KUDA MAS STIK	37	0.12	0.82	4.24	4.48	0.09	0.78	0.82	4.24	4.03	0.09
6	FRENTA STROBERI	FRENTA ANGGUR	30	0.10	0.81	6.06	4.56	0.08	0.78	0.81	6.06	4.01	0.08

Table 6. Results of the Tpq-Apriori Algorithm with Apriori Weka for 50% of the Kopegtel dataset

No	Antecdent	Consequent	Freq	TPQ-Apriori						Apriori Weka		
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift
1	AMARTA COKLAT KACANG	AMARTA CHORY	29	0.10	0.97	8.26	29.43	0.09	0.97			
2	KOPYOR MANGGA	KOPYOR MELON	40	0.13	0.91	5.66	9.33	0.11	0.89	0.13	0.91	5.66
3	KOPYOR MELON	KOPYOR MANGGA	40	0.13	0.83	5.66	5.02	0.11	0.80	0.13	0.83	5.66
4	AMARTA CHORY	AMARTA COKLAT KACANG	29	0.10	0.83	8.26	5.29	0.09	0.81			
5	KUDA MAS MAKARONI	KUDA MAS STIK	37	0.12	0.82	4.24	4.48	0.09	0.78	0.12	0.82	4.24
6	FRENTA STROBERI	FRENTA ANGGUR	30	0.10	0.81	6.06	4.56	0.08	0.78	0.10	0.81	6.06

The test results in Table 4 are a comparison of the rules generated by the Fp-Growth rapidminer and TPQ-Apriori algorithms; the records marked in orange in rows 1 and 4 are rules that were not found by the Fp-Growth rapidminer algorithm. While the records that are colored grayscale are the rules generated by the Fp-Growth rapidminer algorithm, these rules are below the specified minimum support. The data contains 10 rules, but only 2 are displayed, so there are not too many in the displayed table. The rule should not meet the requirements to display. For example, we are calculating support for record 7, the support count from the rule "TEPUNG BERAS ROSE BRAND => TELER JUMBO" is 4, so $Support = \frac{4}{299} = 0.013$ multiplied by 100% the result is 1% even though the minimum support set is 0.1 or 10% of ΣTransactions, 299 is ΣTransactions of the kopegtel dataset for 50% of the volume.

The results of the rule test in Table 8 are a comparison of the rules produced by TPQ-Apriori and Fp-Growth Weka where there are 2 rules not found in Fp-Growth Weka. This is very reasonable because the frequency of the 2 rules that are not found is the rule "AMARTA CHOCOLATE BEANS => AMARTA CHORY" and vice versa is 29 using the formula $Support = \frac{29}{299} =$ 0.097 in TPQ-Apriori this value is rounded up to 0.10 or 10% of ΣTransactions. So that this rule appears on TPQ-Apriori. Meanwhile, Fp-Growth Weka has no tolerance for the specified minimum support value.

Table 7. Results of the tpq-Apriori algorithm rule with fp-growth rapidminer for 25% of the Kopegtel dataset

No	Antecedent	Consequent	Freq	TPQ-Apriori						Fp-Growth Rapidminer			
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift	Conv
1	FRENTA ANGGUR, FRENTA LEMON	FRENTA STROBERI	15	0.07	1.00	8.36	∞	0.06	1.00	0.07	1.00	8.36	∞
2	FRENTA STROBERI, FRENTA JERUK	FRENTA ANGGUR	14	0.07	1.00	8.36	∞	0.06	1.00				
3	KOPYOR MANGGA	KOPYOR MELON	19	0.09	0.90	7.56	8.80	0.08	0.89	0.09	0.90	7.56	9.24
4	FRENTA ANGGUR,FRENTA JERUK	FRENTA STROBERI	14	0.07	0.88	7.32	7.34	0.06	0.86				
5	FRENTA STROBERI, FRENTA LEMON	FRENTA ANGGUR	15	0.07	0.88	7.38	7.34	0.06	0.86	0.07	0.88	7.38	7.48
6	FRENTA STROBERI	FRENTA ANGGUR	20	0.10	0.80	6.69	4.40	0.08	0.77	0.10	0.80	6.69	4.40
7	FRENTA ANGGUR	FRENTA STROBERI	20	0.10	0.80	6.69	4.40	0.08	0.77	0.10	0.80	6.69	4.40
8	TEPUNG WHITE BEAR	TELER JUMBO	3	0.01	0.00	0.00	0.00	0.00	0.00	0.15	0.91	1.06	1.63
9	FRENTA COLA	TELER JUMBO	2	0.01	0.00	0.00	0.00	0.00	0.00	0.10	0.91	1.06	1.58
10	GULA PTPN	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.21	0.90	1.05	1.38
11	TEPUNG BERAS ROSE BRAND	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.16	0.89	1.04	1.36
12	FRENTA LEMON, FRENTA COLA	TELER JUMBO	2	0.01	0.00	0.00	0.00	0.00	0.00	0.07	0.88	1.03	1.22
13	GULA PTPN, TEPUNG WHITE BEAR	TELER JUMBO	3	0.01	0.00	0.00	0.00	0.00	0.00	0.10	0.87	1.02	1.10
14	TEPUNG KETAN ROSE BRAND	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.15	0.86	1.01	1.06
15	MIE 89	TELER JUMBO	8	0.04	0.00	0.00	0.00	0.00	0.00	0.21	0.84	0.98	0.92
16	KUDA MAS MAKARONI	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.84	0.98	0.90
17	FRENTA STROBERI	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.84	0.98	0.90
18	TEPUNG BERAS ROSE BRAND, TEPUNG KETAN ROSE BRAND	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.84	0.98	0.90
19	FRENTA LEMON	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.83	0.97	0.86
20	ICE COCO	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.11	0.82	0.96	0.80
21	KUDA MAS RUJAK	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.09	0.82	0.96	0.79
22	KUDA MAS STIK	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.80	0.93	0.72
23	FRENTA ANGGUR	TELER JUMBO	5	0.02	0.00	0.00	0.00	0.00	0.00	0.10	0.80	0.93	0.72

No	Antecedent	Consequent	Freq	TPQ-Apriori						Fp-Growth Rapidminer			
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift	Conv
24	TELER JUMBO, FRENTA ANGGUR	FRENTA STROBERI	4	0.02	0.00	0.00	0.00	0.00	0.00	0.08	0.80	6.69	4.40
25	FRENTA STROBERI, FRENTA ANGGUR	TELER JUMBO	4	0.02	0.00	0.00	0.00	0.00	0.00	0.08	0.80	0.93	0.72

Table 8. Results of the tpq-Apriori algorithm rule with fp-growth weka for 25% of the Koegtel dataset

No	Antecedent	Consequent	Freq	TPQ-Apriori							Fp-Growth Weka		
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift	Conv
1	FRENTA ANGGUR, FRENTA LEMON	FRENTA STROBERI	15	0.07	1.00	8.36	∞	0.06	1.00	0.07	1.00	8.36	13.21
2	FRENTA STROBERI, FRENTA JERUK	FRENTA ANGGUR	14	0.07	1.00	8.36	∞	0.06	1.00				
3	KOPYOR MANGGA	KOPYOR MELON	19	0.09	0.90	7.56	8.80	0.08	0.89	0.09	0.90	7.56	6.16
4	FRENTA ANGGUR,FRENTA JERUK	FRENTA STROBERI	14	0.07	0.88	7.32	7.34	0.06	0.86				
5	FRENTA STROBERI, FRENTA LEMON	FRENTA ANGGUR	15	0.07	0.88	7.38	7.34	0.06	0.86	0.07	0.88	7.38	4.99
6	FRENTA STROBERI	FRENTA ANGGUR	20	0.10	0.80	6.69	4.40	0.08	0.77				
7	FRENTA ANGGUR	FRENTA STROBERI	20	0.10	0.80	6.69	4.40	0.08	0.77				

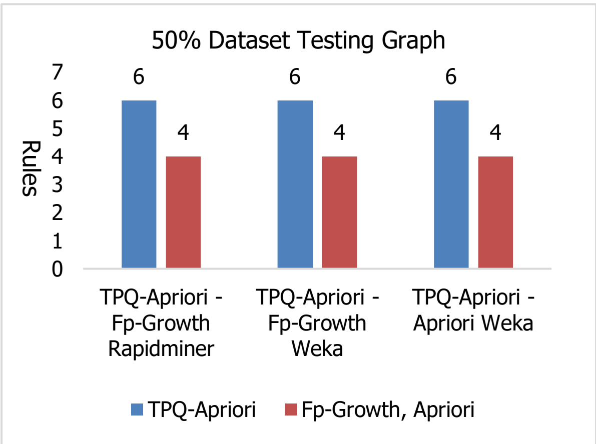
Table 9. Results of the tpq-Apriori algorithm with Apriori weka for 25% of the Kopegtel dataset

No	Antecedent	Consequent	Freq	TPQ-Apriori						Apriori Weka		
				Supp	Conf	Lift	Conv	Leve	CF	Supp	Conf	Lift
1	FRENTA ANGGUR, FRENTA LEMON	FRENTA STROBERI	15	0.07	1.00	8.36	∞	0.06	1.00	0.07	1.00	8.36
2	FRENTA STROBERI, FRENTA JERUK	FRENTA ANGGUR	14	0.07	1.00	8.36	∞	0.06	1.00			
3	KOPYOR MANGGA	KOPYOR MELON	19	0.09	0.90	7.56	8.80	0.08	0.89	0.09	0.90	7.56
4	FRENTA ANGGUR,FRENTA JERUK	FRENTA STROBERI	14	0.07	0.88	7.32	7.34	0.06	0.86			
5	FRENTA STROBERI, FRENTA LEMON	FRENTA ANGGUR	15	0.07	0.88	7.38	7.34	0.06	0.86	0.07	0.88	7.38
6	FRENTA STROBERI	FRENTA ANGGUR	20	0.10	0.80	6.69	4.40	0.08	0.77	0.10	0.80	7.69
7	FRENTA ANGGUR	FRENTA STROBERI	20	0.10	0.80	6.69	4.40	0.08	0.77	0.10	0.80	7.69

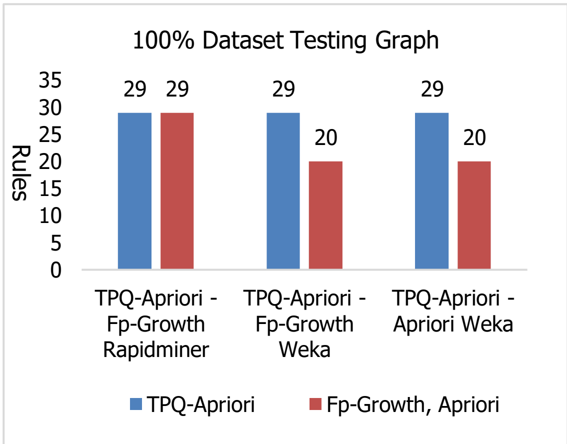
The results of testing the rules in Tables 8 and 9 are a comparison of the rules produced by TPQ-Apriori with Apriori Weka and Fp-Growth Weka, where there are 2 rules not found in Apriori Weka. It's the same with Fp-Growth, which has no tolerance for support values. From the results of Table 7, the results of a comparison of the rules produced by TPQ-Apriori with Fp-Growth rapidminer at TPQ-Apriori yielded the 7 best rules with a minimum threshold support of 0.07 and a minimum confidence of 0.8. A minimum support of 0.07 was chosen because it uses a minimum support of 0.1 on Fp-Growth weka, and Apriori weka do not produce the best rules. Meanwhile, Fp-Growth RapidMiner produced 23 rules, but only 5 of them matched the specified threshold. The rest is the wrong rule, starting from the 8th record to the 25th record, because the support is below the specified minimum support. Table 7 shows the results of testing the rules from TPQ-Apriori with Fp-Growth Weka, where the rules from TPQ-Apriori are 7 rules while the rules from Fp-Growth Weka are only 3 rules. With this result, TPQ-Apriori is more optimal.

Table 9 shows the results of testing the rules from TPQ-Apriori with Apriori Weka, where the rules from TPQ-Apriori are 7 rules and the rules from Apriori Weka are 5 rules. TPQ-Apriori is more optimal, but the rules from Apriori Weka are better than those from Fp-Growth Weka for 25% of the Koegtel dataset volume. From the overall results of testing the rules with good Koegtel datasets for 100%, 50%, and 25% of the total dataset volume, a conclusion can be drawn that the larger the dataset to be processed, the more optimal the results will be when using the Fp-Growth Rapidminer algorithm, but not optimal if the dataset to be processed is a small dataset. It's different from using the Apriori and Fp-Growth Weka algorithms. Where the resulting rule is less than optimal if the dataset used is large and optimal if the dataset is small. Some rules are not displayed in the algorithms in the Fp-Growth Weka and Apriori Weka algorithms because the two algorithms do not have a tolerance value for the support of the rules that will be displayed. Meanwhile, the method proposed and developed in the form of a tool called TPQ-Apriori is capable of producing optimal and consistent rules for both large and small datasets.

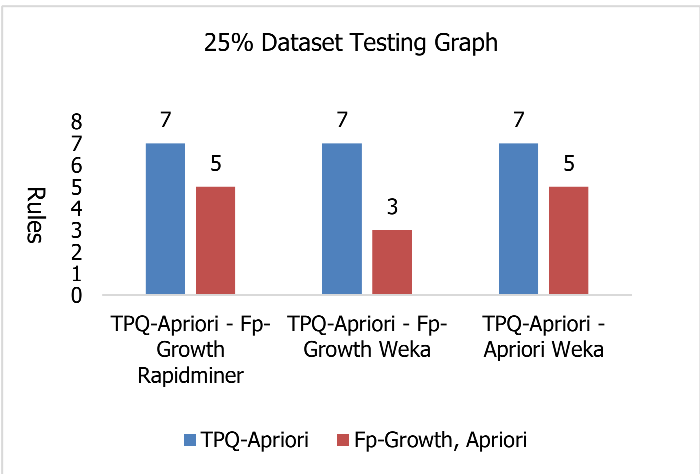
From the results of the analysis of all comparisons of evaluation values for support, confidence, lift ratio, and leverage values, there is no difference, while for conviction there are differences, but the values are close to the same, but they are also the same. But what is unique here is that the conviction value between those produced by the Fp-Growth algorithm in Weka tools is different from the results of the Fp-Growth algorithm in Rapidminer tools. And also, there are rules that are not correctly found or displayed by the Fp-Growth algorithm with RapidMiner tools. Even though this rule has a support value that is very far from fulfilling the minimum support limit that has been set, Thus, it can also be concluded that the rules generated by Fp- Growth RapidMiner are inconsistent. Meanwhile, the rules generated by Fp-Growth Weka and Apriori Weka are slightly more consistent, although sometimes there are a few rules that cannotbe displayed. Overall, they are quite consistent. And finally, the method proposed with the TPQ-Apriori tools is able to produce optimal and consistent rules. Fp-Growth in RapidMiner tools, as well as Fp-Growth and Apriori in Weka tools. To make iteasier, we can see it as shown in Figure 2.



(a)



(b)



(c)

Figure 2. Comparison Chart

We can see the data referred to by the graph; the data depicted in blue are the results from TPQ-Apriori, while the orange ones are the results from Fp-Growth and Apriori from both RapidMiner and Weka tools. While the division of the dataset used is 100%, 50%, and 25%.

Conclusions

The Apriori algorithm is one part of the association rule algorithm with an associative approach. The Apriori algorithm is easy to implement and has the advantage of being able to produce optimal rule combinations. However, the weakness is the very long dataset scan time. Long dataset scans are caused by the process of searching for the frequency of items or support counts in the dataset. Items in the association rule algorithm are attributes or features that will be searched for their frequency or support count. So the bigger the volume, the more items. The length of time the dataset is scanned will have an impact on the amount of memory and processor usage. The priority in this study is to compare three algorithms, namely Apriori, fp-growth, and TPQ-Apriori, with the same dataset to see and ensure the formation of optimal and consistent rules. The test was carried out using the NTB Telkom employee cooperative dataset, which can be downloaded on the Kaggle site.

As for the results of testing the basic rules of the overall results of testing the rules with the Kopegtel dataset both for 100%, 50%, and 25% of the total dataset volume, it can be concluded that the larger the dataset to be processed, the results will be more optimal if using the RapidMiner fp-growth algorithm, but not optimal if the dataset to be processed is a small dataset. In contrast to using the FP-growth Apriori and Weka algorithms, where the resulting rules are less optimal if the dataset used is large and optimal if the dataset is small. The methods used for comparison are the traditional Apriori algorithms fp-growth and TPQ-Apriori, Traditional Apriori, FP-growth, and TPQ-Apriori. For rule testing, the results of the TPQ-Apriori rule were compared to the Fp-Growth algorithm in the Rapidminer and Weka tools and the Apriori algorithm in the Weka tools. From the test results, TPQ-Apriori is able to be more optimal and consistent.

References

[1] C. C. Aggarwal, Data Mining The Textbook, USA: Springer, 2015.
[2] I. H. Witten and E. Frank, M. A. Hall, Data Mining Practical Mechine Learning (Third Edition). Boston: Morgan Koufmann, 2011.
[3] D. T. Larose, Data Mining Methods and Models. John Wiley & Sons, 2006.
[4] R. A. Purba, "Application design to help predict market demand using the waterfall method," Matrix J. Manaj. Teknol. dan Inform., vol. 11, no. 3, pp. 140–149, 2021.
[5] I. M. Fauzan, A. Ihsan, F. A. Mahmudi, A. R. Mubarak, and C. Apriono, "Business transformation from connectivity to digital: case study 'PT. MNO Indonesia,'" MATRIX J. Manaj. Teknol. dan Inform., vol. 12, no. 2, pp. 79–90, 2022, doi: 10.31940/matrix.v12i2.79-90.
[6] C. C. Aggarwal and J. Han, Frequent Pattern Mining. Switzerland: Springer Cham, 2014.
[7] A. Maulidah and F. A. Bachtiar, "Penerapan Metode Association Rule Mining untuk Asosiasi Ulasan Terhadap Aspek Tempat Wisata Jawa Timur Park 3," J. Teknol. Inf. dan Ilmu Komput., vol. 8, no. 5, p. 1029, 2021, doi: 10.25126/jtiik.2021854417.
[8] K. S. Kumar and R. M. Chezian, "A survey on association rule mining using apriori algorithm," International Journal of Computer Applications, vol. 45, no. 5, pp. 47-50, May 2012. doi: 10.1109/ACCT.2015.69.
[9] M. Syahrir and F. Fatimatuzzahra, "Association Rule Integrasi Pendekatan Metode Custom Hashing dan Data Partitioning untuk Mempercepat Proses Pencarian Frekuensi Item-set pada Algoritma Apriori," MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput., vol. 20, no. 1, pp. 149–158, 2020, doi: 10.30812/matrik.v20i1.833.
[10] R. Fitria, W. Nengsih, and D. H. Qudsi, "Implementasi algoritma fp-growth dalam penentuan pola hubungan kecelakaan lalu lintas", J. Sistem Inf. (J. Inf. Sys.), vol. 13, no. 2, pp. 118-124, Oct. 2017.
[11] X. Wu and V. Kumar, The Top Ten Algorithms in Data Mining. New York: Chapman and Hall/CRC (1st edition), April 9, 2009.

[12] R. Agrawal, T. Imieliński, and A. Swami, "Mining Association Rules Between Sets of Items in Large Databases," ACM SIGMOD Rec., vol. 22, no. 2, pp. 207–216, 1993, doi: 10.1145/170036.170072.
[13] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules," in VLDB '94:Proceedings of the 20th International Conference on Very Large Data Bases, 1994, pp. 487–499.
[14] J. Han, M. Kamber, and J. Pei, Introduction. 2012.
[15] E. Darrab, "Vertical Pattern Mining Algorithm for Multiple Support Thresholds," Procedia Comput. Sci., vol. 112, pp. 417–426, 2017, doi: 10.1016/j.procs.2017.08.051.
[16] M. Syahrir and A. Anggrawan, "Improvement of Apriori Algorithm Performance Using the TID-List Vertical Approach and Data Partitioning," Int. J. Intell. Eng. Syst., vol. 16, no. 2, pp. 176–191, 2023, doi: 10.22266/ijies2023.0430.15.
[17] M. Henry, W. Chandra, and A. Zahra, "Implementation of Apriori Algorithm for Music Genre Recommendation," J. Online Inform., vol. 7, no. 1, p. 110, 2022, doi: 10.15575/join.v7i1.819.
[18] R. Wandri and A. Hanafiah, "Analysis of Information Technology (IT) Goods Sales Patterns Using the FP-Growth Algorithm," IT J. Res. Dev., vol. 6, no. 2, pp. 130–141, 2022, doi: 10.25299/itjrd.2022.8155.
[19] I. K. A. Viantama dan P. Painem, "Implementasi Algoritma Apriori Untuk Analisis Penjualan Produk Pada Toko Perjuangan Collection", SENAFTI, vol. 1, no. 1, hlm. 438–444, Sep 2022.
[20] Z. Zulham, E. E. Putri, and B. S. Hasugian, "Pattern Analysis of Drug Procurement System With FP-Growth Algorithm," J. Online Inform., vol. 7, no. 1, p. 70, 2022, doi: 10.15575/join.v7i1.841.
[21] M. Mandeep, "Efficient Ordering Policy for Imperfect Quality Items Using Association Rule Mining," Encycl. Inf. Sci. Technol. Third Ed., no. August 2014, pp. 773–786, 2014, doi: 10.4018/978-1-4666-5888-2.ch074.
[22] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," ACM SIGMOD Record, vol. 29, no. 2, pp. 1–12, 2000.
[23] C. Zhang and S. Zhang, Association Rules Mining. Berlin, Heidelberg: Springer, 2002.
[24] P. F. Viger, J. C. Wei Lin, R. Nkambou, B. Vo, and V. S. Tseng, High-Utility Pattern Mining Theory, Algorithms and Applications. Switzerland: Springer, 2019.
[25] S. Mita, Y. Yamazoe, T. Kamataki, and R. Kato, "Metabolic activation of a tryptophan pyrolysis product, 3-amino-1-methyl-5H-pyrido [4,3-b] indole (Trp-P-2) by isolated rat liver nuclei", Cancer Letters, vol. 14, no. 3, pp.261-266. 1981.