

Apriori Method for sales optimization of rafin's snack product as UMKM

Alvin Bryan Saputra^{*1}, Erza Sofian²

Sekolah Tinggi Ilmu Manajemen dan Ilmu Komputer ESQ

Abstract

Received: 19 December 2022

Revised: 29 December 2022

Accepted: 16 January 2023

This research is a quantitative research using a descriptive method that adopts CRISP-DM as a research stage. The purpose of this research is to find out what snack products sell the most and determine marketing strategies from the association rules obtained to increase sales. The data source used is primary data, namely sales transaction data for the period January 2021 to December 2021 obtained from the owner of Rafin's Snack house. The results of this study show that the apriori algorithm can be used to find association rules and determine the most sold snack products and can determine sales strategies, with a minimum Support and minimum Confidence determined, namely Support 30% and Confidence 50%. The association rules found, it is known that the most sold snack products are Salted Egg and Hot Bomb, and the right marketing strategy is to bundle products. The a priori algorithm is influential in knowing what snack products sell the most and can determine marketing strategies, from the association rules obtained, it is known that if consumers buy Salted Egg, 67% (certainty that consumers buy items) will buy Hot Bomb, and if consumers buy Hot Bomb, 75% will buy Salted Egg, and marketing strategies by bundling Salted Egg and Hot Bomb products with products that are less attractive to consumers, namely, shredded catfish, coffee, and banana sale.

Keywords: Data Mining, Algoritma Apriori, Support, Confidence

(*) Corresponding Author:

a.bryan.s@students.esqbs.ac.id, erza.sofian@esqbs.ac.id

How to Cite: Saputra, A., & Sofian, E. (2023). Apriori Method for sales optimization of rafin's snack product as UMKM. *International Journal of Education, Information Technology, and Others*, 6(1), 113-122. <https://doi.org/10.5281/zenodo.7663022>

INTRODUCTION

The development of the business world in the current era of globalization is increasingly rapid and is accompanied by progress in the field of technology-based information systems. This also brings a major change in the level of competition between business people, thus requiring business people to find a strategy that can increase sales and marketing of the products being sold (Purnia & Warnilah, 2017).

UMKM are trading businesses that are managed by individuals or business entities and comply with business criteria in a small scope or also in a micro scope (Wijoyo, 2021). Rafin's Snack House is an MSME assisted by the Pringsewu Regency Koperindag Service that produces food made from processed catfish skin, cassava, potatoes, taro, bananas, peanuts, taro, coffee, sweet potatoes, and also shredded catfish. Founded in 2017 which is located at Jalan Jendral Sudirman Pringsewu Selatan, Pringsewu District, Pringsewu Regency, Lampung. Until now, Rafin's snack has been exporting its products abroad, such as to Egypt, Japan, and Singapore.

Based on the Explanation of the owner of Rafin's snack, sales transaction data for Rafin's snacks have only been used as archives for UMKM Rafin's Snack

and only become historical data for UMKM, so this transaction data is useless and useless data.

According to Buulolo (2020), in order for sales transaction data to be useful and useful, a data processing algorithm is needed, one of the data processing algorithms is the a priori algorithm. The a priori algorithm is a data collection algorithm with associative rules (Association Rule) to find associative relationships for a combination of items (Sibarani, 2020). The Association Rule is carried out through a mechanism for calculating the support and confidence of a relationship item. In this study there is a relationship between the items to be analyzed, thus the application of the Apriori Algorithm can be used to find out the analysis of sales transaction patterns for Rafin's snack products, also to help find which snack products sell the most.

Researchers used the Apriori algorithm to search for the most sold snack products and help determine the marketing strategy at Rafin's Snack House SMEs from sales transaction data. After obtaining the most sold snack products, the researcher hopes to be able to help Rafin's Snack House SMEs to make it easier to determine marketing strategies to increase sales. Based on previous research, it states that the association rules produced by the a priori algorithm have a higher level of accuracy compared to those produced by the FP-Growth algorithm (Takdirillah, 2020).

Thus the researcher decided to use the a priori algorithm, so the results are used for the development of increased sales and marketing of Rafin's snack products

BACKGROUND

Apriori Algorithm

The Apriori algorithm is the most well-known algorithm for finding high-frequency patterns (Purnia & Warnilah, 2017). The quality of the data collected is an important factor in determining decisions when implementing the a priori algorithm on data sets (John & Shaiba, 2019). The a priori algorithm is suitable when there are several item relationships to be analyzed. In research conducted by (Silva et al., 2019) an a priori algorithm was used to classify and find associations between products for each customer group.

The apriori algorithm is divided into several stages which are called narratives or passes (Sutradana, 2017). The stages are:

- a. Formation of candidate itemset. The k-itemset candidate is formed from the (k-1) itemset combination obtained from the previous iteration. The method of the a priori algorithm is pruning k-itemset candidates whose subset contains k-1 items not included in a high-frequency pattern with k-1 length.
- b. Calculation of support from each k-itemset candidate. Support for each k-itemset candidate is obtained by scanning the database to calculate the number of transactions that contain all items in the k-itemset candidate. This is also a feature of the a priori algorithm where it is necessary to calculate by means of the entire database the longest k-itemset.
- c. Set a high frequency pattern. High-frequency patterns that contain k items or k-itemsets are determined from k-itemset candidates whose support is greater than the minimum support.
- d. If no new high-frequency pattern is found, then the whole process is stopped.

Association Rules

Association Rules is a data mining technique for determining associative rules or combination patterns of an item. One of the stages of association analysis that produces an efficient algorithm is frequent pattern mining. The importance of an association can be determined by two benchmarks, namely: support and confidence. Support (support value) is the percentage of item combinations in the database, confidence (certainty value) is the strength of the relationship between items in the association rules (Ristianingrum, 2017).

1. High frequency pattern analysis

Looking for an itemset combination that meets the minimum requirements of the support

value of an item. Value support for an itemset:

$$Support(A) = \frac{Jumlah\ Transaksi\ mengandung\ A}{Total\ Transaksi} \times 100\%$$

2. Formation of itemset combinations

the support value of the 2 itemset is obtained by using the formula:

$$Support(A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{Jumlah\ Transaksi\ mengandung\ A\ dan\ B}{Transaksi} \times 100\%$$

3. Formation of association rules

Confidence is the strength of the relationship between items in associative rules.

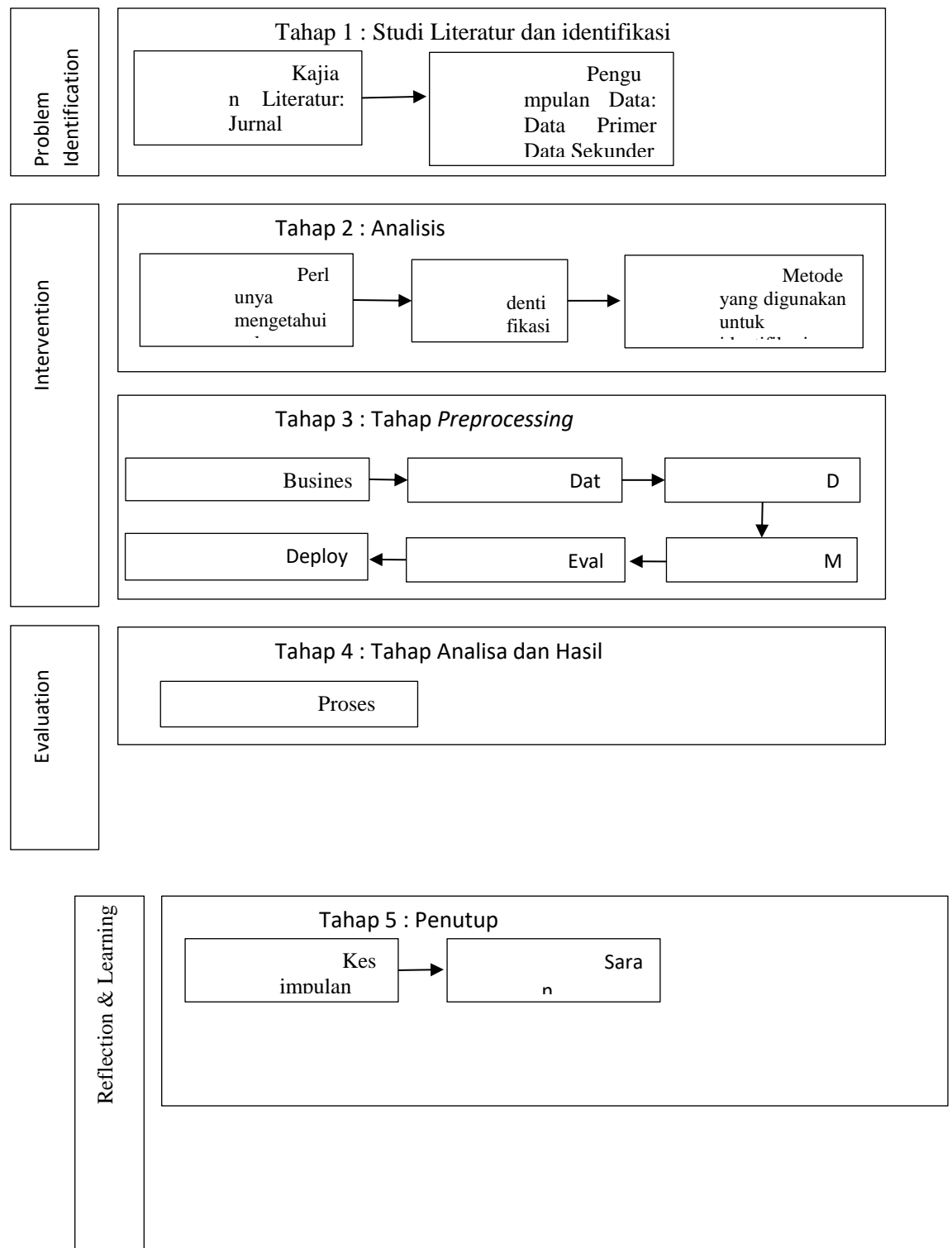
$$Confidence (A \Rightarrow B) = \frac{Jumlah\ transaksi\ mengandung\ (A\ dan\ B)}{jumlah\ transaksi\ mengandung\ (A)} \times 100\%$$

RESEARCH METHODELOGY

The research method used in this study is a descriptive analysis method and the approach used is a quantitative approach in which this research emphasizes its analysis on numeric data or numbers, the aim is to get a clear picture of a situation based on the data that has been obtained by presenting, collecting and analyze data to become new information that is used to analyze the problem being studied.

Research stage

CRISP-DM provides a standardized process for data mining that can be used into common problem-solving strategies in businesses or research units. This study adopted CRISP-DM as a research stage



Pic. 1 Research Gate

The analysis stage and the results carried out produce information in the form of relationships between two or more itemsets from existing sales transaction data so that association rules can be found through the a priori algorithm

Data Understanding

The data used is data from consumer purchase transactions that are stored in a database for a period of 12 months of operation at Rafin's Snack House from January 2021 to December 2021. The data that will be accumulated is obtained from monthly sales taken from the top three or the most sold monthly reports. can be seen in table 1

Table 1. Data Understanding

NO	NAMA PRODUK	JAN/pcs	FEB/pcs	MAR/pcs	APR/pcs	MEL/pcs	JUN/pcs	JUL/pcs	AGU/pcs	SEP/pcs	OKT/pcs	NOV/pcs	Des/pcs
1	SALTED EGG	653	83	913	448	331	2339	340	3032	1927	281	1591	2559
2	HOT BOMB	990	198	1426	758	218	1795	0	2995	543	1473	1577	2776
3	POTATO	569	169	218	228	299	771	0	243	818	119	513	907
4	BANANA	234	718	2054	500	96	0	200	1115	686	1778	436	2032
5	PEANUTS	53	40	0	0	39	0	0	80	0	0	1031	0
6	TARO	279	618	192	905	40	310	0	42	750	0	940	878
7	CASSAVA	491	440	496	647	816	130	324	716	860	0	0	1484
8	UBI UNGU	34	70	4114	0	0	0	0	0	0	0	154	0
9	KOPI	0	0	0	0	0	0	0	0	0	0	0	0
10	PISANG SALE	0	0	0	0	125	0	62	0	34	0	0	0
11	ABON PATIN	0	0	0	0	0	242	0	0	303	0	0	0
	JUMLAH	3303	2336	9413	3486	1964	5587	926	8223	5921	3651	6242	10636

Data Preparation

Based on sales transactions for Rumah Rafin's snacks, transactions are accumulated. The accumulation of snack product sales transactions is obtained from monthly sales taken from the top 3 monthly reports, can be seen in table 4.2

Table . 1 Sales Transaction Pattern

No	Itemset
1	Hot Bomb, Salted egg, Potato
2	Banana, Taro, Cassava
3	Ubi Ungi, Banana, Hot Bomb
4	Taro, Hot Bomb, Cassava
5	Cassava, Salted Egg, Potato
6	Salted Egg, Hot Bomb, Potato
7	Salted Egg, Cassava, Banana
8	Salted Egg, Hot Bomb, Banana
9	Salted Egg, Cassava, Potato
10	Banana, Hot Bomb, Salted Egg
11	Salted Egg, Hot Bomb, Peanuts
12	Hot Bomb, Salted egg, Banana

the results of the snack sales transaction pattern, a tabular format of monthly transaction data on snack sales is made which will make it easier to find out how many items are sold in each transaction, see table 3

Table 2 Matriks

Bulan	Salted Egg	Hot Bomb	Potato	Banana	Peanut	Taro	Cassava	Ubi Ungu
1	1	1	1	0	0	0	0	0
2	0	0	0	1	0	1	1	0
3	0	1	0	1	0	0	0	1
4	0	1	0	0	0	1	1	0
5	1	0	1	0	0	0	1	0
6	1	1	1	0	0	0	0	0
7	1	0	0	1	0	0	1	0
8	1	1	0	1	0	0	0	0
9	1	0	1	0	0	0	1	0
10	1	1	0	1	0	0	0	0
11	1	1	0	0	1	0	0	0
12	1	1	0	1	0	0	0	0

Model (Modeling)

Formation of 1 Itemsets

The following is a calculation for the formation of 1 itemset based on the tabular table of transaction data above. The goal is to find out the support value of each item in an existing transaction. The minimum support value determined by the researcher is the minimum support value of 30%. The process of forming 1 itemset uses the equation :

$$Support(A) = \frac{Jumlah\ Transaksi\ mengandung\ A}{Total\ Transaksi} \times 100\%$$

Table . 3 Formating of 1 Itemset

No	Itemset	Amount of Itemset	Support (%)
1	Salted Egg	9	75%
2	Hot Bomb	8	67%
3	Potato	4	33%
4	Banana	5	42%
5	Peanuts	1	8%
6	Taro	2	17%
7	Cassava	3	25%
8	Ubi Ungu	1	8%

After forming the support value of each item, the analysis can determine the minimum support value to be used. The specified minimum support is 30%, so those who do not meet the minimum support will be eliminated items that meet the minimum support value are Saltede Egg, Hot Bomb, Potato, and Banana

Formation of 2 Itemsets

Furthermore, the calculation for the formation of 2 itemsets, Salted Egg, Hot Bomb, Potato, and Banana with a minimum amount of support of 30%, can be seen in the table below. The process of forming 2 itemsets uses the equation:

$$Support(A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{Jumlah\ Transaksi\ mengandung\ A\ dan\ B}{Transaksi} \times 100\%$$

Table . 4 formation of 2 Itemsets

2 Itemset	Amount of Itemset	Support (%)
Salted Egg, Hot Bomb	6	50%
Salted Egg, Potato	4	33%
Salted Egg, Banana	4	33%
Hotbomb, Potato	2	17%
Hotbomb, Banana	4	33%
Potato, Banana	0	0%

Formation of 3 Itemsets

Then the calculation for the formation of 3 itemsets, Salted Egg, Hot Bomb, Potato, and Banana with a minimum amount of support of 30%, can be seen in the table above. The process of forming 3 itemsets uses the equation:

$$Support(A, B, dan C) = P(A \cap B \cap C)$$

$$= \frac{Jumlah\ Transaksi\ mengandung\ A,\ B\ dan\ C}{Transaksi} \times 100\%$$

Table . 5 Formation of 3 Itemsets

3 Itemset	Amount of Itemset	Support (%)
Salted Egg, Hot Bomb, Potato	2	17%
Salted Egg, Hot Bomb, Banana	3	25%

Because there is no combination of 3 Itemsets that meets the minimum support, 2 Itemsets will be used to further determine the minimum confidence.

Formation of Association Rules

After all the high-frequency patterns have been found, association rules that meet the minimum requirements for confidence by calculating the confidence of the associative rules $A \rightarrow B$ have a minimum confidence of 50%, the confidence value of the rules $A \rightarrow B$ is obtained from the following equation:

$$Support(A, B) = \frac{Jumlah\ Transaksi\ mengandung\ A\ dan\ B}{Transaksi} \times 100\%$$

$$Confidence = P(B|A) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{jumlah\ transaksi\ mengandung\ (A)} \times 100\%$$

Table . 6 Formation of Association of Rules

Itemset	Support	Confidence
Salted Egg, Hotbomb	50%	67%
Hot Bomb, Salted Egg	50%	75%
Hot Bomb, Potato	17%	25%
Potato, Hot Bomb	17%	50%
Potato, Banana	0%	0%
Banana, Potato	0%	0%

Evaluation

Based on the minimum support and minimum confidence that has been determined, it can be seen in the table below:

Table . 7 Evaluation

Rules	Support	Confidence
If buy Salted Egg then will buy Hot Bomb	50%	67%
If buy Hot Bomb then will buy Salted Egg	50%	75%

From the two rules above, it can be concluded that the most sold snack products at Rafin's snack houses, namely Salted Egg and Hot Bomb.

Deployment

The apriori algorithm was successfully implemented to calculate and find the association rules according to the two parameters in the association rules, namely the minimum support value and the minimum confidence value. The results of the association rules are taken from sales transaction data and calculated using Ms. Excel. In this study, the researchers determined a minimum support value of 30% and a minimum confidence value of 50%. From processing transaction data Ms. Excel, the results of the association rules are obtained as follows:

Table . 9 Deployment

Itemset	Support	Confidence
Salted Egg, Hotbomb	50%	67%
Hot Bomb, Salted Egg	50%	75%

From the resulting association rules, there are 2 rules that meet the specified minimum support and minimum confidence values

1. Support 50% means that 50% of all transactions analyzed show that Salted Egg and Hot Bomb sold the most while Confidence of 67% indicates the level of trust or it can be said that if a consumer buy Salted Egg then there is a 67% chance that he will buy Hot Bomb.
2. Support 50% means that 50% of all transactions analyzed show that Hot Bomb and Salted Egg are the most sold while Confidence of 75% indicates the level of trust or it can be said that if a consumer buy Hot Bomb then there is a 75% chance that he will buy Salted Egg.

CONCLUSION

Based on the results of the research, the conclusions that were obtained were that the researchers succeeded in analyzing data on sales of rafin's snack house sales in the period January 2021 - December 2021. The Apriori algorithm can be implemented to determine the most sold products and assist in determining marketing strategies. Association rules resulting from the application of the Apriori Algorithm with a minimum support value of 30% (the strong combination of these items in the database) and a minimum confidence value of 50% (the strong relationship between items in the association rules) produce two association rules. The two association rules are as follows:

1. If consumers buy Salted Egg then 67% (consumer certainty in buying items) will buy Hot Bomb
2. If consumers buy Hot Bomb then 75% (consumer certainty in purchasing items) will buy Salted Egg.

From the two association rules obtained, it can be seen which snack products sell the most, namely Rafin's Salted Egg and Hot Bomb snacks, so that Rafin's Snack can make a marketing strategy to increase sales by bundling Salted Egg and Hot Bomb products with snack products that are not much in demand by consumers, namely shredded catfish, coffee, and banana sale in accordance with rafins' snack sales data for the period January 2021 to December 2021. Based on the association rules that have been produced in this study, it can be developed as a business strategy by UMKM Rafin's snack house. The right strategy that is in accordance with the rules of the association is obtained, namely product bundling. Bundling Salted Egg and Hot Bomb products with snack products that were not selling, namely shredded catfish, coffee and banana sale according to Rafins' snack sales data for the period January 2021 to December 2021.

The limitation of this research is that the initial data obtained still uses monthly summary data, it would be better if the initial data uses data per sales transaction so that more and better knowledge is obtained.

REFERENCES

- Adha, N., Sianturi, L. T., & Siagian, E. R. (2017). Implementasi Data Mining Penjualan Sabun Dengan Menggunakan Metode Apriori (Studi Kasus : PT. Unilever). *Majalah Ilmiah INTI*, 12(2), 219–223.
- Anggraini, Y., Pasha, D., & Damayanti Setiawan, A. (2020). Sistem Informasi Penjualan Sepeda Berbasis Web Menggunakan Framework Codeigniter (Studi Kasus : Orbit Station). *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 1(2), 64–70.
- Edastama, P., Bist, A. S., & Prambudi, A. (2021). Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm. *International Journal of Cyber and IT Service Management*, 1(2), 159–172. <https://doi.org/10.34306/ijcitsm.v1i2.46>
- Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1120–1126. <https://doi.org/10.29207/resti.v5i6.3588>

- John, M., & Shaiba, H. (2019). Apriori-Based Algorithm for Dubai Road Accident Analysis. *Procedia Computer Science*, 163, 218–227. <https://doi.org/10.1016/j.procs.2019.12.103>
- Kurnia, Y., Isharianto, Y., Giap, Y. C., Hermawan, A., & Riki. (2019). Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm. *Journal of Physics: Conference Series*, 1175(1). <https://doi.org/10.1088/1742-6596/1175/1/012047>
- Patimah, S., & Hasugian, P. S. (2020). Implementation Of Data Mining In Determining Sales Pattern Of Snack Products Using Apriori Algorithm (Case Study : PT Siantar Top Tbk). *Journal of Intelligent Decision Support System (IDSS)*, 3(4), 1–9.
- Purnia, dini silvi, & Warnilah, ai ilah. (2017). Implementasi Data Mining Pada Penjualan kacamata Dengan Menggunakan Algoritma Apriori. *Indonesian Journal on Computer and Information Technology*, 2(2), 31–39.
- Sibarani, A. J. P. (2020). Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(2), 262–276. <https://doi.org/10.35957/jatisi.v7i2.195>
- Sikumbang, E. D. (2018). Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori. *Jurnal Teknik Komputer AMIK BSI (JTK)*, Vol 4, No.(September), 1–4.
- Silva, J., Varela, N., López, L. A. B., & Millán, R. H. R. (2019). Association rules extraction for customer segmentation in the SMES sector using the apriori algorithm. *Procedia Computer Science*, 151(2018), 1207–1212. <https://doi.org/10.1016/j.procs.2019.04.173>
- Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic : Jurnal Pendidikan Informatika*, 4(1), 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>