

Analysis of Data Mining Associations on Drug Sales at Pharmacies with APRIORI Techniques

Salam Irianto Nadeak¹, Yusmar Ali²

^{1,2} Politeknik Negeri Media Kreatif, Indonesia

¹salamirianto77@gmail.com, ²yusalim.msi@gmail.com

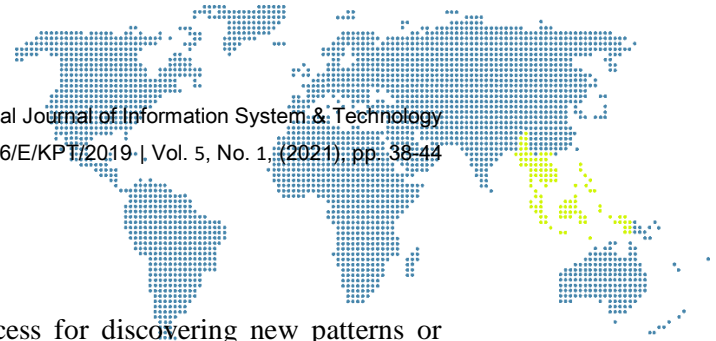
Abstract

The purpose of the research is to utilize artificial intelligence techniques in analyzing drug sales. Sources of data used are observations and interviews with shop owners. The method used as a solution is the Association method with the Apriori technique. By using the RapidMiner software, the test results are obtained using a minimum support of 25% and a minimum of 60% confidence as many as 5 rules with a predetermined itemset. The results of the study are expected to provide information and make it easier for related parties to find combinations of selling items. The results of this analysis can be used by pharmacies for marketing strategies and product promotion.

Keywords: Data Mining, Drug Sales, Association, Apriori Algorithm

1. Introduction

Sales transaction data shall be available to a pharmacy or shops selling their goods. Due to the large number of transactions occurring daily, daily transaction information will increase data. A pharmacy must have important information to boost sales by processing transaction data, which have collected useful information, in order to increase sales. This helpful data can be used for a corporate policy, namely by using the information in a corporate strategy. To develop your company, a pharmacy certainly needs a business strategy. Data mining [1], [2] is an activity that seeks out interesting patterns in large amounts of data and stores them in a data warehouse or other form of information storage. Data mining patterns must be simple to understand and practical [3]. The Apriori algorithm [4–6] is a frequent itemset search algorithm that employs association rule technology. One of the many pharmacies in the area is "United." This pharmacy is on Jalan Sangnawaluh Rasaan, not far from residential areas. "United" pharmacies always strive to provide high-quality services and products, but they still fall short of competing with other pharmacies in the area. As a result, "ABC" pharmacies must be able to comprehend consumer needs. One method is to keep various drugs available in the pharmacy warehouse. Data mining can be used for analytical technology, namely analyzing consumer needs, to determine which drugs are purchased by consumers. The Apriori algorithm can help shape the possible combinations of candidate products and then test whether those combinations meet the user's minimum support and trust parameters. Every day, there are more and more buying and selling activities for services to consumers at the "United" Pharmacy, resulting in a larger data pile. Although the "ABC" pharmacy's service activities and transactions have not encountered any issues thus far, the large amount of data collected at one time can become an impediment to improving services, making it difficult for pharmacies to analyze the types of products that are most in demand and least desired by consumers.



2. Research Methodology

2.1 Data Mining

Data mining is an iterative and interactive process for discovering new patterns or models in large databases that are completely effective, useful, and easy to understand. Data mining patterns must be simple to understand, novel, and useful. Data mining is the process of discovering interesting patterns or information in selected data by employing specific techniques or methods. In data mining, the relationship sought can be one between two or more in one dimension [7].

2.2. Association

The Association method is to display the confidence or relationship between items. The Association method includes two stages, namely finding the most frequent combination of an itemset and defining the conditions and results (for the conditional association method). In determining an association method, there is an interestness measure obtained from the results of data processing with certain calculation data [8].

The formation of association rules consists of two steps:

a) Mode analysis at high frequencies

At this point, the database is searched for a combination of items that meet the minimum support value requirements. Support for the rule "X => Y" is the likelihood of an attribute or set of attributes X and Y appearing in a transaction at the same time.

$$\text{Support } (X \Rightarrow Y) = P(X \cap Y) \tag{1}$$

X => Y = items that appear in the same order P (XY) = probability of X and Y transactions divided by total number of transactions.

b) Formation of the association's rules

After locating all high-frequency patterns, compute the confidence of the if X then Y rules in order to identify association rules that meet the minimum confidence requirements..

$$\text{Confidence } (X \Rightarrow Y) = P(Y | X) \tag{2}$$

X => Y = items that appear in the same order P (Y | X) = the probability of the number of transactions that contain X and Y divided by the number of transactions that contain X.

2.3. Apriori Algorithm

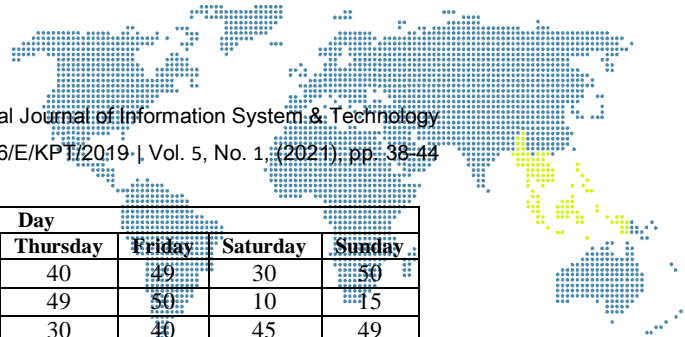
The Apriori algorithm was proposed in 1994 by Agrawal and Srikant to determine frequent itemsets for boolean association rules. Apriori algorithm makes use of prior knowledge about itemsets with frequent occurrences, also known as itemset frequency frekuensi [9].

3. Result and Discussion

The information used in this study is from drug sales in "ABC" pharmacies. In this study, 8 data samples were collected, and the collected data will be calculated by looking for support and confidence values.

Table 1. Data used

No	Itemset	Day						
		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	Antangin Jrg Cair	60	56	55	50	60	15	22
2	Ambeven Kapsul	55	10	52	56	55	15	21



No	Itemset	Day						
		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
3	Insto Tetes Mata	50	55	53	40	49	30	50
4	Albothyl 5 ml	40	40	45	49	50	10	15
5	Alvita Tablet	45	15	30	30	40	45	49
6	Andalan	25	26	10	15	10	25	20
7	Andalan Laktasi	15	16	19	25	27	14	11
8	Alpara Tablet	20	10	15	11	25	20	10

The manual calculation of the Apriori Algorithm is as follows:

a) Transaction Patterns in Drug Sales at “ABC” Pharmacies

Based on drug sales data from “United” Pharmacies. The transaction patterns are obtained by analyzing the eight types of drugs that occur the most frequently on a daily basis, as shown in table 2 below.

Table 2. Drug Sales Transaction Pattern

Day	Itemset
Monday	Antangin Jrg Cair, Ambeven Kapsul, Insto Tetes Mata
Tuesday	Antangin Jrg Cair, Insto Tetes Mata, Albothyl 5 ml
Wednesday	Antangi Jrg Cair, Insto Tetes Mata, Ambeven Kapsul
Thursday	Ambeven Kapsul, Antangin Jrg Cair, Albothyl 5 ml
Friday	Antangin Jrg Cair, Ambeven Kapsul, Albothyl 5 ml
Saturday	Alvita Tablet, Insto Tetes Mata, Andalan
Sunday	Insto Tetes Mata, Alvita Tablet, Antangin Jrg Cair

b) Tabular Format Creation

Tabular format of daily transaction data, when formed, will look like table 3 below:

Table 3. Tabular Format of Transaction Data

Itemset	Day						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Antangin Jrg Cair	1	1	1	1	1	0	1
Ambeven Kapsul	1	0	1	1	1	0	0
Insto Tetes Mata	1	1	1	0	0	1	1
Albothyl 5 ml	0	1	0	1	0	0	0
Alvita Tablet	0	0	0	0	0	1	1
Andalan	0	0	0	0	0	1	0

c) High Frequency Pattern Analysis

1. Formation of 1 Itemset

The process of forming C1 or called 1 itemset with a minimum amount of support = 15%. With the following formula:

$$\text{Support (A)} = \frac{\sum \text{Transaksi yang mengandung A}}{\sum \text{Transaksi}} * 100\% \quad (3)$$

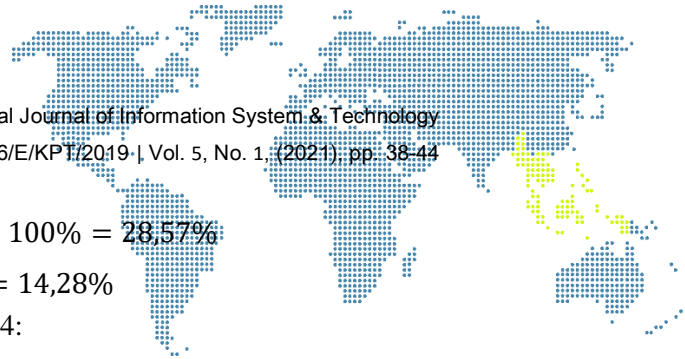
The following is the calculation of the formation of 1 itemset:

$$\text{Support Antangin Jrg Cair} = \frac{\sum \text{Transaksi Antangin Jrg Cair}}{\sum 7} = \frac{6}{7} * 100\% = 85,71\%$$

$$\text{Support Ambeven Kapsul} = \frac{\sum \text{Transaksi Ambeven Kapsul}}{\sum 7} = \frac{4}{7} * 100\% = 57,14\%$$

$$\text{Support Insto Tetes Mata} = \frac{\sum \text{Transaksi Insto Tetes Mata}}{\sum 7} = \frac{5}{7} * 100\% = 71,42\%$$

$$\text{Support Albothyl 5 ml} = \frac{\sum \text{Transaksi Albothyl Cair}}{\sum 7} = \frac{2}{7} * 100\% = 28,57\%$$



$$\text{Support Alvita Tablet} = \frac{\Sigma \text{Transaksi Alvita Tablet}}{\Sigma 7} = \frac{2}{7} * 100\% = 28,57\%$$

$$\text{Support Andalan} = \frac{\Sigma \text{Transaksi Andalan}}{\Sigma 7} = \frac{1}{7} * 100\% = 14,28\%$$

Based on the description above, it can be made table 4:

Table 4. Support of each Itemset

Itemset	amount	Support
Antangin Jrg Cair / A	6	85,67%
Ambeven Kabsul / B	4	57,11%
Aito Tetes Mata / C	5	71,42%
Albothyl 5 ml / D	2	28,57%
Alvita Tablet / E	2	28,57%
Andalan / F	1	14,28%

2. Combination of 2 Itemset

The process of forming C2 or called 2 itemset with a minimum amount of support = 20%. With the following formula:

$$\text{Support}(A,B) = \frac{\Sigma \text{Transaksi yang mengandung } A,B}{\Sigma \text{Transaksi}} * 100\% \quad (4)$$

The following is the calculation of the formation of 2 itemsets:

$$\text{Support}(A,B) = \frac{\Sigma \text{Transaksi A dan B}}{\Sigma 7} = \frac{4}{7} * 100\% = 57,14\%$$

$$\text{Support}(A,C) = \frac{\Sigma \text{Transaksi A dan C}}{\Sigma 7} = \frac{4}{7} * 100\% = 57,14\%$$

$$\text{Support}(A,D) = \frac{\Sigma \text{Transaksi A dan D}}{\Sigma 7} = \frac{2}{7} * 100\% = 28,57\%$$

$$\text{Support}(A,E) = \frac{\Sigma \text{Transaksi A dan E}}{\Sigma 7} = \frac{1}{7} * 100\% = 14,28\%$$

$$\text{Support}(A,F) = \frac{\Sigma \text{Transaksi A dan F}}{\Sigma 7} = \frac{0}{7} * 100\% = 0\%$$

The minimum Support specified is 20%, so the combination of 2 itemsets that do not meet the minimum Support will be removed, as shown in Table 5 below:

Table 5. Minimum Support 2 itemset 20%

Itemset	amount	Support
Antangin Jrg Cair, Ambeven Kabsul	4	57,14%
Antangin Jrg Cair, Aito Tetes Mata	4	57,14%
Antangin Jrg Cair, Albothyl 5 ml	2	28,57%
Ambeven Kabsul, Aito Tetes Mata	2	28,57%
Aito Tetes Mata, Alvita Tablet	2	28,57%

3. Combination of 3 Itemset

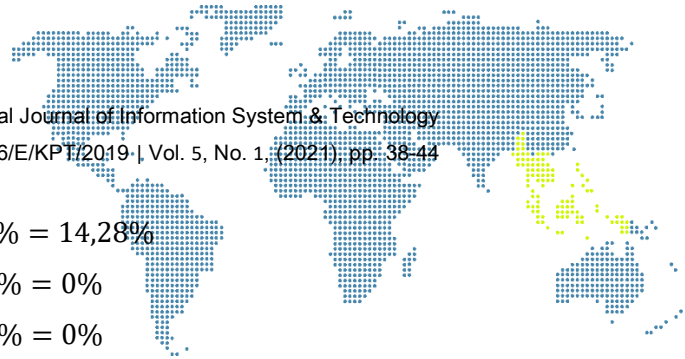
The process of forming C3 or called 3 itemset with a minimum amount of 20%. With the following formula:

$$\text{Support}(A,B \text{ dan } C) = \frac{\Sigma \text{Transaksi yang mengandung } A,B \text{ dan } C}{\Sigma \text{Transaksi}} * 100\% \quad (5)$$

The following is the calculation of the formation of 3 itemset:

$$\text{Support}(A,B \text{ dan } C) = \frac{\Sigma \text{Transaksi } A,B \text{ dan } C}{\Sigma 7} = \frac{2}{7} * 100\% = 28,57\%$$

$$\text{Support}(A,B \text{ dan } F) = \frac{\Sigma \text{Transaksi } A,B \text{ dan } F}{\Sigma 7} = \frac{0}{7} * 100\% = 0\%$$



$$Support(A,C \text{ dan } D) = \frac{\Sigma \text{Transaksi } A,C \text{ dan } D}{\Sigma 7} = \frac{1}{7} * 100\% = 14,28\%$$

$$Support(A,C \text{ dan } E) = \frac{\Sigma \text{Transaksi } A,C \text{ dan } E}{\Sigma 7} = \frac{0}{7} * 100\% = 0\%$$

$$Support(B,D \text{ dan } F) = \frac{\Sigma \text{Transaksi } B,D \text{ dan } F}{\Sigma 7} = \frac{0}{7} * 100\% = 0\%$$

Based on the description above, it can be made table 6:

Table 6. Combination of 3 itemset

Itemset	Jumlah	Support
Antangin Jrg Cair, Ambeven Kapsul, Insto Tetes Mata	2	28,57%
Antangin Jrg Cair, Insito Tetes Mata, Albothyl 5 ml	1	14,28%
Antangin Jrg Cair, Ambeven Kapsul, Andalan	1	14,28%
Aito Tetes Mata, Alvita Tablet, Andalan	1	14,28%
Antangin Jrg Cair, Aito Tetes Mata, Alvita Tablet	1	14,28%

The minimum support specified is 25%, then 3 combinations that meet the requirements will form an association rule.

d) Rules for Association Formation

After identifying all high-frequency patterns, the association rules that meet the minimum Confidence requirements are found by calculating the Confidence of the associative rules AB with a minimum confidence value of 60%. The following formula yields the AB rule's confidence value:

$$Confidance = P(A/B) = \frac{\Sigma \text{Transaksi yang mengandung } A}{\Sigma \text{Transaksi } A} * 100\% \quad (4)$$

The following is a Confidence calculation:

$$Confidance = P(A/B) = \frac{\Sigma \text{Transaksi } A \text{ dan } B}{\Sigma 6} = \frac{4}{6} * 100\% = 66,66\%$$

$$Confidance = P(A/C) = \frac{\Sigma \text{Transaksi } A \text{ dan } C}{\Sigma 6} = \frac{4}{6} * 100\% = 66,66\%$$

$$Confidance = P(A/D) = \frac{\Sigma \text{Transaksi } A \text{ dan } D}{\Sigma 6} = \frac{2}{6} * 100\% = 33,33\%$$

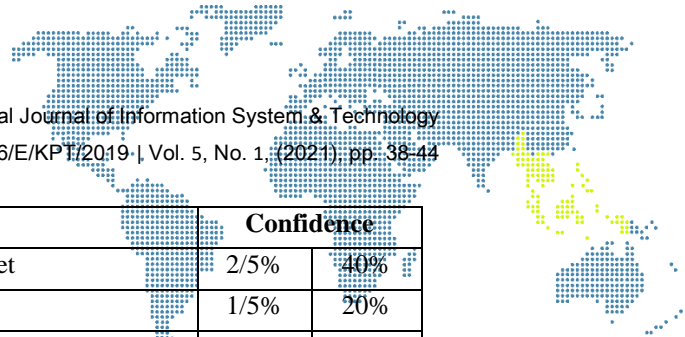
$$Confidance = P(A/E) = \frac{\Sigma \text{Transaksi } A \text{ dan } E}{\Sigma 6} = \frac{1}{6} * 100\% = 16,66\%$$

$$Confidance = P(A/F) = \frac{\Sigma \text{Transaksi } A \text{ dan } F}{\Sigma 6} = \frac{0}{6} * 100\% = 0\%$$

From the combination of 2 itemsets that have been determined, it can be seen the magnitude of the support value and the confidence value of the candidate association as shown in the following table:

Table 7. Candidate Association Rules

Rules	Confidence	
If you buyAntangin Jrg Cair Then Will Buy Ambeven Kapsul	4/6 %	66,66%
If you buyAntangin Jrg Cair Then Will Buy Aito Tetes mata	4/6 %	66,66%
If you buyAntangin Jrg Cair Then Will Buy Albothyl 5 ml	2/6 %	33,33%
If you buyAntangin Jrg Cair Then Will Buy Alvita Tablet	1/6%	16,66%
If you buyAntangin Jrg Cair Then Will Buy Andalan	0/7%	0,00%
If you buyAmbeven Kapsul Then Will Buy Aito Tetes mata	2/4%	50%
If you buyAmbeven Kapsul Then Will Buy Albothyl 5 ml	1/4%	25%
If you buyAmbeven Kapsul Then Will Buy Alvita Tablet	0/4	0
If you buyAmbeven Kapsul Then Will Buy Andalan	0/4%	0
If you buyAito Tetes Mata Then Will Buy Albothyl 5 ml	1/5%	20%



Rules	Confidence	
	If you buyAito Tetes Mata Then Will Buy Alvita Tablet	2/5%
If you buyAito Tetes Mata Then Will Buy Andalan	1/5%	20%
If you buyAlbothyl 5 ml Then Will Buy Alvita Tablet	0/2%	0
If you buyAlbothyl 5 ml Then Will Buy Andalan	0/2%	0%
If you buyAlvita Tablet Then Will Buy Andalan	1/1%	100%

e) Final Association Rules

Based on the candidate association rules in table 7, those who meet a minimum of 25% support and a minimum of 60% confidence can be seen in the table below:

Table 8. Final Association Rules

Rules	Support	Confidence
If you buyAntangin Jrg Cair Then Will Buy Ambeven Kabsul	57.14%	66,66%
If you buyAntangin Jrg Cair Then Will Buy Aito Tetes mata	57,14%	66,66%
If you buyAntangin Jrg Cair Then Will Buy Albothyl 5 ml	28,57%	33,33%
If you buyAmbeven Kabsul Then Will Buy Aito Tetes mata	28,57%	50,00%
If you buyAito Tetes Mata Then Will Buy Alvita Tablet	28,57%	40,00%

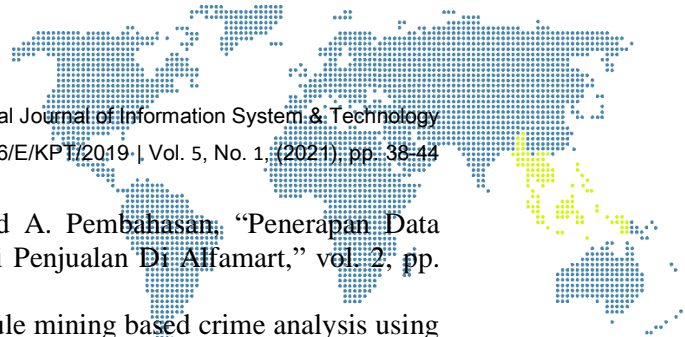
4. Conclusion

From the test results using a minimum support of 25% and a minimum confidence of 60%, it produces a possible rule.

- a) If you buyAntangin Jrg Cair, Then Will Buy Ambeyen Kapsul generate value support 57,14% and confidence 66,66%.
- b) If you buyAntangin Jrg Cair, Then Will Buy Insto Tetes Mata generate value support 57,14% and confidence 66,66%.
- c) If you buyAntangin Jrg Cair, Then Will Buy Albothyl 5 ml generate value support 28,57% and confidence 33,33%.
- d) If you buyAmbeyen Kapsul, Then Will Buy Insto Tetes Mata generate value support 28,57% and confidence 50,00%.
- e) If you buyInsto Tetes Mata, Then Will Buy Alvita Tablet generate value support 28,57% and confidence 40,00%.

References

- [1] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, "Bagian 1: Kombinasi Metode Klastering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia)," *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, p. 855, 2020, doi: 10.30865/mib.v4i3.2312.
- [2] A. P. Windarto, J. Na, and A. Wanto, "Bagian 2 : Model Arsitektur Neural Network dengan Kombinasi K- Medoids dan Backpropagation pada kasus Pandemi COVID-19 di Indonesia," vol. 4, pp. 1175–1180, 2020, doi: 10.30865/mib.v4i4.2505.
- [3] S. Kasus, P. Pt, G. Gunadi, and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) :," vol. 4, no. 1, 2012.
- [4] R. D. Jayapana and Y. Rahayu, "Analisis pola pembelian konsumen dengan algoritma apriori pada apotek rahayu jepara," *UG Jurnal*, vol. Vol. 8, pp. 1–6, 2015.



- [5] N. Agustiani, D. Suhendro, K. Apriori, and A. Pembahasan, “Penerapan Data Mining Metode Apriori Dalam Implementasi Penjualan Di Alfamart,” vol. 2, pp. 300–304, 2020.
- [6] H. Jantan and A. Z. M. Jamil, “Association rule mining based crime analysis using apriori algorithm,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 1.5 Special Issue, pp. 18–24, 2019, doi: 10.30534/ijatcse/2019/0581.52019.
- [7] P. Obat and M. Algoritma, “Data mining,” vol. 2, no. 2, pp. 1–7, 2013.
- [8] Putri Septiani Azura, M Safii, and Fitri Rizki, “Analisa Penjualan Produk Oli Dengan Metode Data Mining,” vol. 2, no. 1, pp. 8–15.
- [9] Dewi Listriani, Anif Hanifa Setyaningrum, and Fenty Eka M.A, “Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro),” vol. 9, no. 2, pp. 120–127, 2016.