International Journal of Informatics, Information System and Computer Engineering

Association Analysis with Apriori Algorithm for Electronic Sales Decision Support System

R. Fenny Syafariani

Mathematics Department, Faculty of Ocean Technology Engineering and Informatics, Universiti Malaysia Terengganu, Malaysia

Corresponding Email: r.fenny.syafariani@email.unikom.ac.id

ABSTRACTS

The purpose of this study was to determine the level of dependence of various items in order to dig up information on what items are dependent on other items. The method used in this research is descriptive analysis with a qualitative approach through a priori algorithm. The results show that the association analysis of the 26 transactions taken is 76.47%. A consumer who buys a laptop electronic device has the possibility to also buy an electronic mouse.

ARTICLE INFO

Article History:
Submitted/Received 05 Jan 2022
First Revised 18 Feb 2022
Accepted 25 Mar 2022
First Available Online 01 Apr 2022
Publication Date 01 Jun 2022

Keywords:
Technology,
Information System,
Apriori,
Algorithm,
Decision Support System,
Electronic sales

1. INTRODUCTION

Data mining is a data processing method to find patterns from the data obtained (Ordila et al., 2020). There are many methods in data mining. One method that is often used is the association method or association rule, more precisely using the Apriori Algorithm. The data generated from the sales process or transaction data is processed by association rules to find out information related to product purchases made by buyers (Riszky et al., 2019).

There are various kinds of electronic goods that are sold such as Laptops, Printers, Mouse and so on. Sales transactions continue to grow every day and cause huge data storage (Purnia et

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

al., 2017). Most sales transaction data is only used as an archive without being used properly. However, this data set contains very useful information. With the application of association analysis or association rule mining in this discussion, it is hoped that association rules can be found between a combination of items. So that obtained a knowledge of the application of the concept of association mining analysis through the search for support and confidence.

The previous research discussed "Application of Data Mining for Analysis of Consumer Purchase Patterns with the Fpgrowth Algorithm on Motorcycle Spare Part Sales Transaction Data" (Fajrin et al., 2018). This research is tested in order to influence consumer buying patterns, because each consumer's buying pattern is different. This needs to be analyzed further so that it can produce useful information, as well as maximize the benefits that can be obtained. Then the next research that has been done previously is discussing "Data Mining Analysis for Clustering Covid-19 Cases in Lampung Province with the K-Means Algorithm" (Nabila et al., 2021). This study was to analyze data on Covid-19 cases in order to find out the grouping of the Covid-19 case problems in Lampung Province. The grouping of data on Covid-19 cases in Lampung Province was carried out using the Clustering method with the K-Means algorithm. The results of DBI validation using manual calculations and using the help of RapidMiner tools have differences, in this case manual calculations have better results than using RapidMiner tools, but the results of both calculations are both close to 0 which means the evaluated clusters produce good clusters.

In the previous research conducted using the fpgrowth algorithm analysis method, and subsequent research using the k-means method in conducting the analysis. This research is "Association Analysis with Apriori Algorithm for Decision Support System for Selling Electronic Goods" (Riszky et al., 2019). Data mining and a priori algorithms are very useful to find out the reality of the frequency of sales of electronic goods that are most in demand by consumers. so that it can be used as very useful information in making decisions to prepare stocks of what types of electronic goods are needed in the future.

2. METHOD

In this study using descriptive analysis method with a qualitative approach (Rahmawati et al., 2018). While in data processing using data mining techniques. The algorithm approach used is the a priori algorithm. The process of forming a combination of itemsets pattern and making rules starts from data analysis. The data used is data on sales of electronic goods, then followed by the formation of a combination of itemsets pattern and from an interesting combination of itemsets, association rules are formed. Then the data is made in tabular data format (Tana et al., 2018; Syahril et al., 2020; Simbolon, 2019). In relation to the application used in the test, it is an application that uses one of the Microsoft Excel databases with data in tabular data, then the sales transaction data (electronic goods data out), is converted into binary form (Triansyah et al., 2018). After that the formation of a combination of two elements with a minimum value of the frequency of occurrence = 15 and a minimum value of

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

confidence = 75%. To calculate support and confidence, the following formula is used:

$$\text{Support} = \frac{\Sigma items\ purchased\ at\ once}{\Sigma total\ transaction} \times 100\%$$

$$\text{Confidence} = \frac{\Sigma items\ purchased\ at\ once}{\Sigma transactions\ in\ the\ antecedent\ section} \times 100\ \%$$

3. RESULTS AND DISCUSSION

3.1. Preprocessing

The dataset used as a test sample in this study uses 26 transaction data. In the data there are several items of electronic goods sold, namely printers, laptops, chargers, and mice. The following is a Table 1 of transaction data that is used as a sample.

Table 1. Transaction Data Table

Transaction	Item
E1	Printer,Laptop
E2	Laptop, Printer
E3	Charger, Printer
E4	Printer, mouse, Printer
E5	Charger, Printer, Printer
E6	Laptop, Printer
E7	Printer, Laptop, Charger
E8	Printer, Laptop, Printer
E9	Printer, Laptop
E10	Printer, Printer
E11	Printer, Laptop, Printer
E12	Laptop, Printer
E13	Charger, Printer, Printer
E14	Printer, Laptop
E15	Printer, Laptop, Charger
E16	Printer, Charger
E17	Charger, Printer
E18	Charger, Printer, Printer
E19	Printer, Laptop, Printer

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 1 (Continue). Transaction Data Table

Transaction	Item
E20	Charger, Printer, Printer
E21	Laptop, Printer
E22	Printer, Laptop, Charger
E23	Printer, Laptop
E24	Printer, mouse, Printer
E25	Laptop, Printer
E26	Printer, mouse, Printer

3.2. Transaction Data Tabular Format

The application used in the test is a Microsoft Excel database so that the data must be converted into binary form (Sianturi et al., 2018). The conversion process is that the slip number of the data to be tested is made horizontally downwards, while all types of items will become vertical attributes, so that they form like a table, based on real sales data (electronic goods data out) the meeting point between the name of the electronic type and the number slip will become binary 1, while those that do not become meeting points will become binary 0. The results of the conversion process of sales transaction data to data format in tabular data form are as shown in Table 2.

Table 2. Data in the form of tabular data

Transaction	Printer	Mouse	Laptop	Charger
E1	1	0	1	0
E2	0	1	1	0
E3	1	0	0	1
E4	1	1	1	0
E5	1	1	0	1
E6	0	1	1	0
E7	0	1	1	1
E8	1	1	1	0
E9	1	0	1	0

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 2 (Continue). Data in the form of tabular data

Transaction	Printer	Mouse	Laptop	Charger
E10	1	1	0	0
E11	1	1	1	0
E12	0	1	1	0
E13	1	1	0	1
E14	1	0	1	0
E15	0	1	1	1
E16	0	1	0	1
E17	1	0	0	1
E18	1	1	0	1
E19	1	1	0	0
E20	1	1	0	1
E21	0	1	1	0
E22	0	1	1	1
E23	1	0	1	0
E24	1	1	1	0
E25	0	1	1	0
E26	1	1	1	0
	17	20	17	10

In the tabular data table above, the number of occurrences (electronic items that come out) of each item is: Printer = 17, Mouse = 20, Laptop = 17, and Charger = 10.

3.3. Formation of Two Elements Combination Pattern

With a minimum value of the frequency of occurrence Φ= 15. In the form table tabular data, there is one electronic item that does not meet the provisions of the frequency limit, namely Charger = 10, so in the formation of the pattern of these two elements we make a combination of pairs of 3 electronic items, namely Printer-Mouse, Laptop Printer, Mouse-Laptop. Tables 3-5 show the elements combination pattern.

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 3. Two Elements Combination Pattern (Printer, Mouse)

Transaction	Printer	Mouse	f
E1	1	0	S
E2	0	1	S
E3	1	0	S
E4	1	1	P
E5	1	1	P
E6	0	1	S
E7	0	1	S
E8	1	1	P
E9	1	0	S
E10	1	1	P
E11	1	1	P
E12	0	1	S
E13	1	1	P
E14	1	0	S
E15	0	1	S
E16	0	1	S
E17	1	0	S
E18	1	1	P
E19	1	1	P
E20	1	1	P
E21	0	1	S
E22	0	1	S
E23	1	0	S
E24	1	1	P
E25	0	1	S

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 3 (Continue). Two Elements Combination Pattern (Printer, Mouse)

Transaction	Printer	Mouse	f
E26	1	1	P
Total (P)			11

Table 4. Two Elements Combination Pattern (Printer, Laptop)

Transaction	Printer	Laptop	f
E1	1	1	P
E2	0	1	S
E3	1	0	S
E4	1	1	P
E5	1	0	S
E6	0	1	S
E7	0	1	S
E8	1	1	P
E9	1	1	P
E10	1	0	S
E11	1	1	P
E12	0	1	S
E13	1	0	S
E14	1	1	P
E15	0	1	S
E16	0	0	S
E17	1	0	S
E18	1	0	S
E19	1	0	S
E20	1	0	S
E21	0	1	S
E22	0	1	S
E23	1	1	P
E24	1	1	P
E25	0	1	S
E26	1	1	P
Total (P)			9

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 5. Two Elements Combination Pattern (Printer, Laptop)

Transaction	Mouse	Laptop	f
E1	0	1	S
E2	1	1	P
E3	0	0	S
E4	1	1	P
E5	1	0	S
E6	1	1	P
E7	1	1	P
E8	1	1	P
E9	0	1	S
E10	1	0	S
E11	1	1	P
E12	1	1	P
E13	1	0	S
E14	0	1	S
E15	1	1	P
E16	1	0	S
E17	0	0	S
E18	1	0	S
E19	1	0	S
E20	1	0	S
E21	1	1	P
E22	1	1	P
E23	0	1	S
E24	1	1	P
E25	1	1	P
E26	1	1	P
Total (P)			13

From the tables of the 2 elements above, P means that the items are sold together, while S means that there are no items that are sold together or there is no transaction. Σ represents the number of Frequency items set. So that in the pattern of these two elements, the support value is obtained, namely:

- Printer – Mouse = 11
- Printer – Laptop = 9
- Mouse – Laptop = 13

3.4. Formation of Three Elements Combination Pattern

The combination of the 2 elements in the table above, we can combine into 3 elements. For the set formed on these 3 elements are Laptop, Printer, Mouse (see Table 3).

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Table 6. Three Elements Combination Pattern (Printer, Mouse, Laptop)

Transaction	Printer	Mouse	Laptop	f
E1	1	0	1	S
E2	0	1	1	S
E3	1	0	0	S
E4	1	1	1	P
E5	1	1	0	S
E6	0	1	1	S
E7	0	1	1	S
E8	1	1	1	P
E9	1	0	1	S
E10	1	1	0	S
E11	1	1	1	P
E12	0	1	1	S
E13	1	1	0	S
E14	1	0	1	S
E15	0	1	1	S
E16	0	1	0	S
E17	1	0	0	S
E18	1	1	0	S
E19	1	1	0	S
E20	1	1	0	S
E21	0	1	1	S
E22	0	1	1	S
E23	1	0	1	S
E24	1	1	1	P
E25	0	1	1	S
E26	1	1	1	P
Total (P)				5

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

It can be seen from the pattern table of the 3 elements above, the items that were sold simultaneously were Laptop – Printer - Mouse = 5

So, the support value in the 3-element pattern table is 5.

3.4. Association Rules

Calculating the support and confidence values of each frequent itemset so that candidate association rules appear (Lestari, 2017). To calculate support and confidence, the following formula is used:

$$\text{Support} = \frac{\Sigma items\ purchased\ at\ once}{\Sigma total\ transaction} \times 100\%$$

So, that the results are obtained as in the Table 7.

$$\text{Confidence} = \frac{\Sigma items\ purchased\ at\ once}{\Sigma the\ number\ of\ transactions\ in\ the\ antecedent\ section} \times 100\ \%$$

From the table above, the support and confidence have been determined. then select the association rules that meet the minimum confidence of 75%, so that the association rules are obtained, which are as follows:

From the results of the analysis that has been carried out, there is 1 product association rule that meets the minimum confidence limit, namely Laptop - Mouse. Then the results obtained are "76,47% of transactions that contain Laptop electronics also contain Mouse electronics. And 50% of all transactions that contain these two items". With Apriori Algorithm analysis can be applied to assist marketing strategies in a company or institutions. Data mining and a priori algorithms are very useful to find out the relationship between the frequency of sales of electronic goods that are most in demand by customers, so that they can be used as very valuable information in making decisions to prepare stocks of what types of electronic goods are needed in the future.

Table 7. Association Rules Candidate List

If antecedent, then consequent	Support	Confidence
Printer, Mouse	11/26*100 %= 42,30%	11/17*100 %= 64,70%
Mouse, Printer	11/26*100 %= 42,30%	11/20*100 %= 55%
Printer, Laptop	9/26*100%=34,61%	9/17*100%=52,94%
Laptop, Printer	9/26*100%=34,61%	9/17*100%=34,61%
Mouse, Laptop	13/26*100%=50%	13/20*100%=65%
Laptop, Mouse	13/26*100%=50%	13/17*100%=76,47%

Table 8. Association Rules List

If antecedent, then consequent	Support	Confidence
Laptop, Mouse	13/26*100%=50%	13/17*100%=76,47%

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

4. CONCLUSION

A priori algorithm is used in conducting association analysis to determine the level of dependence of various items to explore information on what items have dependence on other items based on 26 transaction records that are sampled. The author performs an association analysis calculation from the samples taken so that the result is that 76.47% of a consumer who buys laptop electronics has the possibility to also buy mouse electronics. And 50% of all transactions that contain these two items.

REFERENCES

Fajrin, A. A., & Maulana, A. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, *5*(01), 1-10.

Lestari, N. (2017). Penerapan Data Mining Algoritma Apriori Dalam Sistem Informasi Penjualan. *Jurnal Edik Informatika Penelitian Bidang Komputer Sains dan Pendidikan Informatika*, *3*(2), 103-114.

Nabila, Z., Isnain, A. R., Permata, P., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi*, *2*(2), 100-108.

Ordila, R., Wahyuni, R., Irawan, Y., & Sari, M. Y. (2020). Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma Clustering (Studi Kasus: Poli Klinik Pt. Inecda). *Jurnal Ilmu Komputer*, *9*(2), 148-153.

Purnia, D. S., & Warnilah, A. I. (2017). Implementasi Data Mining Pada Penjualan Kacamata Menggunakan Algoritma Apriori. *IJCIT (Indonesian Journal on Computer and Information Technology)*, *2*(2).

Rahmawati, F., & Merlina, N. (2018). Metode Data Mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, *6*(1), 9-20.

Riszky, A. R., & Sadikin, M. (2019). Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan. *Jurnal Teknologi dan Sistem Komputer*, *7*(3), 103-108.

Riszky, A. R., & Sadikin, M. (2019). Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan. *Jurnal Teknologi dan Sistem Komputer*, *7*(3), 103-108.

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584

Sianturi, F. A. (2018). Penerapan Algoritma Apriori Untuk Penentuan Tingkat Pesanan. *Jurnal Mantik Penusa*, *2*(1).

Simbolon, P. H. (2019). Implementasi Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori (Studi Kasus: Srikandi Cash Credit Elektronic dan Furniture). *JURIKOM (Jurnal Riset Komputer)*, *6*(4), 401-406.

Syahril, M., Erwansyah, K., & Yetri, M. (2020). Penerapan Data Mining untuk menentukan pola penjualan peralatan sekolah pada brand wigglo dengan menggunakan algoritma apriori. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, *3*(1), 118-136.

Tana, M. P., Marisa, F., & Wijaya, I. D. (2018). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Pada Toko Oase Menggunakan Algoritma Apriori. *JIMP-Jurnal Informatika Merdeka Pasuruan*, *3*(2).

Triyansyah, D., & Fitrianah, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *InComTech: Jurnal Telekomunikasi dan Komputer*, *8*(3), 163-182.

DOI: https://doi.org/10.34010/injiiscom.v3i1.8089
p-ISSN 2810-0670 e-ISSN 2775-5584