 p-ISSN 2827-8585
e-ISSN 2827-8062

International Journal of Multidisciplinary Research and Literature
IJOMRAL
Vol. 1, No. 5, September 2022 pp. 481-600
Journal Page is available at http://ijomral.esc-id.org/index.php/home

APPLICATION OF DATA MINING BY USING A PRIORI ALGORITHM TO IMPROVE CUSTOMER PURCHASING DECISIONS AT MIKAMART BLITAR STORE

Nurul Farida[1*], Mukh Taofik Chulkamdi[2], Zunita Wulansari[3]
[1,2,3] Blitar Islamic University, Indonesia
Corresponden Author: *nurifarida27@gmail.com*[1]

Abstract

Product arrangement, known as product display, is the layout or way of structuring products, especially products companies apply to attract consumers. Although one of the scopes of sales in the market is the delivery of information through galleries, storefronts require specialization and the ability to arrange products or displays following the standards and specifications of the company and the product itself. The resulting output directly affects the level of sales and information to be achieved in a store, especially in modern retail stores with a self-service format (self-service) such as minimarkets.

To implement a display system, Data Mining is applied as a determinant of the layout of the merchandise. According to (Kusrini & Lutfi 2009), data mining is used to extract knowledge from databases. One of the many methods in Data Mining is the Apriori Algorithm. The a priori algorithm is a type of association rule in data mining. Association rule or association rule mining is used to find associative rules between item combinations to be used as wisdom to make decisions and provide a stock of products that consumers prefer.

Kata kunci: Data Mining, Apriori Algorithm, Display Product

INTRODUCTION

The importance of a business, a job that is very easy to achieve with management, a business development defined by many things such as in new product development, sales concept and marketing concept define a company concept. Therefore, the sales concept and marketing concept activities cannot separate from promotional activities (sales promotion) and display activities. Display activities (product arrangement) are activities of a company to display merchandise both inside and outside the room to influence potential consumers directly or indirectly on the goods to be sold.

To implement the display system, Data Mining is applied as a determinant of the layout of the merchandise. Understanding Data Mining, according to (Kusrini & Lutfi, 2009), is a term used to extract knowledge from databases. One of the many methods in Data Mining is the Apriori Algorithm. The a priori algorithm is a type of association rule in data mining. Association rule or association rule mining is used to find associative rules between item combinations. An example of associative rules that we can apply in everyday life is the application in supermarkets. The owner of the supermarket can find out how likely a customer is to buy instant noodles and sausages. With the information or knowledge obtained, if instant noodles and sausages are known that they are often purchased together, the supermarket owner can make placement the two products close together so that buyers will easily take the two products. The supermarket owner will get higher profits after implementing the association rules.

Based on the description above, the researcher has a solution to solve the problems that exist in the stall, namely, by applying data mining to determine product arrangement using an a priori algorithm.

The results can be used as a guide for companies to determine the display of goods and promote goods that are not selling well so that they can participate or sell quickly. In addition, customer satisfaction and company profits can maximize, and the risk of losing unsold products will reduce.

According to research (Nursikuwagus & Hartono, 2016), implementing the a priori algorithm to find the most sold products can help develop association rules. This association rule is obtained based on the Itemset selection on each Transaction. Thus the results obtained can be used to assist decision-makers. Another result is that this implementation can help stock up on products liked by many buyers and increase product inventory.

METHOD

This research was conducted at the supermarket "Mika Mart" Maliran village, Blitar Regency, approximately 7 km from the city of Blitar with the name of the owner, Mr. Miwanto Suharsono. The steps taken by the researcher in this study were problem analysis, where data were collected and analyzed to provide an overview of the problem at the research site. Applying an a priori algorithm to determine best-selling products can help develop association rules. This association rule is derived from selecting a set of entities for each Transaction. Therefore, the results obtained can be used to help decision-makers. Another result is that this implementation can help many customers find their desired products and improve product offerings. Association analysis is known to apply to shopping cart content analysis in supermarkets, and association analysis is often referred to as market basket analysis.

Association analysis is one of the data mining techniques with various other data mining techniques. Specifically, this is one of the steps that attracted the attention of researchers to create an effective algorithm with high-frequency pattern analysis (frequency pattern mining). The importance of an associative rule can be determined by two parameters: support and confidence. Support (support value) is the proportion of the combination of these items in the database, while confidence (value) is the strong correlation between items in the association rules.

Association rules are expressed in the form:
{bread, butter} -> {milk} {support = 40%, confidence = 50%}

The rule means that “50%” of transactions in the database that process bread and butter items also process milk items. then "40%" of transactions recorded in the database process those three items." Therefore, it can be concluded: "A consumer who buys bread and butter has a 50% probability of buying milk. This rule is significant because it represents 40% of the transaction records". Association analysis is a process to determine all association rules that meet the minimum requirements for support (minimum support) and minimum requirements for confidence (minimum confidence).

1. High-frequency pattern analysis

   This stage is looking for a combination of items that meet the minimum requirements of database support value. The following formula obtains the support value of an item.

$$\text{Support(A)} = \frac{\text{Transaction amount contains A}}{\text{Total Transaction.}}$$

The element set is the set of an itemset in 1, and the element set is the set of elements containing k elements. For example, {tea, sugar} is 2 itemset, and {tea, sugar, bread} is 3 itemset. The set of frequency elements indicates the set of elements whose occurrence frequency exceeds the specified minimum (ϕ). for example, = 2, the set of all elements whose frequency can be greater than or equal to 2 times is called frequency. The set that occurs frequently is denoted by F., While the two-point support value is obtained from Equation (2) below.

$$\text{Support (A, B)} = P(A \cap B)$$
$$\text{Support (A, B)} = \frac{\Sigma \text{ Transaction amount contains A and B}}{\Sigma \text{ Transaction}}$$

2. Formation of Association Rules

If you have found a high-frequency pattern, then find the association rules that meet the minimum requirements for confidence (certainty) by calculating the certainty of the associative rules A->B. The certainty value of the rule A->B is obtained from the following formula.

$$\text{Confidence} = P(B \mid A) = \frac{\Sigma \text{ The number of transactions contains A and B}}{\Sigma \text{ transactions}} \times 100\%$$

FINDINGS AND DISCUSSIONS

In this system, data collected from interviews and the research director at the research site will be used as a reference to know the problems faced. Sample data is obtained from the research location in table 1.

Table 1 Sales Data

No	Date	Note	Goods	Amount	Process	Discount	Total Price
1	01/05/2022	MKPJ22030170008	Fresco ground coffee 7gr	2	5.000	0	10.000
2	01/05/2022	MKPJ22030170011	Cusson powder 350 gr pink	3	20.000	0	60.000
3	01/07/2021	MKPJ22030170017	MP-116. big mechanical pencil	3	3.000	0	9.000
4	01/07/2021	MKPJ22030170008	R/GG Signature	4	19.000	0	76.000
5	01/07/2021	MKPJ22030170004	A4 vision drawing book	3	5.000	0	15.000
6	01/05/2022	MKPJ22030170013	C-TIK 6*10 plastic	3	3.000	0	9.000
7	01/05/2022	MKPJ22030170002	Fresco ground coffee 7gr	1	5.000	0	5.000
8	01/05/2022	MKPJ22030170004	butterfly ruler	1	6.000	0	6.000

9	01/05/2022	MKPJ22030170019	A4 vision drawing book	3	5.000	0	15.000
10	01/05/2022	MKPJ22030170003	Omela klg thick sweet 370 ml	6	10.000	0	60.000
11	01/05/2022	MKPJ22030170019	MP-116. big mechanical pencil	3	3.000	0	9.000
12	01/05/2022	MKPJ22030170019	butterfly ruler	1	6.000	0	6.000

The data in the table is normalized by grouping data on items with notes with the same code into one transaction data.

Table 2 Normalized Sales Data

No	Date	Note	Goods	Amount	Process	Discount	Total Price
1	01/05/2022	MKPJ22030170002	Fresco ground coffee 7gr	1	5.000	0	5.000
2	01/05/2022	MKPJ22030170003	Omela klg thick sweet 370 ml	6	10.000	0	60.000
3	01/05/2022	MKPJ22030170004	butterfly ruler	1	6.000	0	6.000
4	01/05/2022	MKPJ22030170004	A4 vision drawing book	3	5.000	0	15.000
5	01/05/2022	MKPJ22030170008	R/GG Signature	4	19.000	0	76.000
6	01/05/2022	MKPJ22030170008	Fresco ground coffee 7gr	2	5.000	0	10.000
7	01/05/2022	MKPJ22030170011	Cusson powder 350 gr pink	3	20.000	0	60.000
8	01/05/2022	MKPJ22030170013	C-TIK 6*10 plastic	3	3.000	0	9.000
9	01/05/2022	MKPJ22030170017	MP-116. big mechanical pencil	3	3.000	0	9.000
10	01/05/2022	MKPJ22030170019	butterfly ruler	1	15,500	0	15,500
11	01/05/2022	MKPJ22030170019	A4 vision drawing book	3	5.000	0	15.000
12	01/05/2022	MKPJ22030170019	MP-116. big mechanical pencil	3	3.000	0	9.000

The application calculations will occur manually if the min support is 2, the min confidence is 2, and the number of transactions is 20. The formula to calculate the support value is:

$$\text{Support(A)} = \frac{\text{Transaction amount contains A}}{\text{Total Transaction}}.$$

Support (A, B) = P (A ∩ B)

$$\text{Support (A, B)} = \frac{\Sigma \text{ The number of transactions contains A and B}}{\Sigma \text{ Transaksi}}$$

Table 3 Itemset 1

No	Item 1	Item 2	Amount	Support
1	Fresco ground coffee 7gr	2	10	Getaway
2	Omela klg thick sweet 370 ml	1	5	Did not pass
3	butterfly ruler	2	10	Getaway
4	A4 vision drawing book	2	10	Getaway
5	R/GG Signature	1	5	Did not pass
6	Cusson powder 350 gr pink	1	5	Did not pass
7	C-TIK 6*10 plastic	1	5	Did not pass
8	MP-116. big mechanical pencil	2	10	Getaway

Suppose the support value owned is more than the value specified by the user. Items that pass will be used as data for the 2nd itemset formation process.

Table 4 Selected Itemset 1

No	Items	Amount	Support
1	Fresco ground coffee 7gr	2	10
2	butterfly ruler	2	10
3	A4 vision drawing book	2	10
4	MP-116. big mechanical pencil	2	10

Itemset 1 that passes can be used to form itemset two by combining item 1 with different item 2. Then a scanning process for transaction data on itemset 2 with normalized sales data is carried out.

Table 5 Itemset 2

No	Item 1	Item 2	Amount	Support	
1	Fresco ground coffee 7gr	Butterfly ruler	1	5	Did not pass
2	Fresco ground coffee 7gr	A4 vision drawing book	1	5	Did not pass
3	Fresco ground coffee 7gr	MP-116. big mechanical pencil	1	5	Did not pass
4	butterfly ruler	A4 vision drawing book	2	10	Getaway
5	butterfly ruler	MP-116. big mechanical pencil	2	10	Getaway
6	A4 vision drawing book	MP-116. big mechanical pencil	2	10	Getaway
7	A4 vision drawing book	Butterfly ruler	2	10	Getaway

Table 6 Itemset 2 that passed

No	Item 1	Item 2	Amount	Support

1	butterfly ruler	A4 vision drawing book	2	10
2	butterfly ruler	MP-116. big mechanical pencil	2	10
3	A4 vision drawing book	MP-116. big mechanical pencil	2	10

Itemset 2 that passes will be processed for the formation of itemset 3, namely the process of combining items where in itemset 3, the normalized sales data is checked. The number and value of support are calculated.

Table 7 Itemset 3

No	Item 1	Item 2	Item 3	Amount	Support	
1	Butterfly Ruler	A4 vision drawing book	MP-116. big mechanical pencil	2	10	Getaway

Table 8. Itemset 3 that passed

No	Item 1	Item 2	Item 3	Amount	Support
1	Butterfly Ruler	A4 vision drawing book	MP-116. big mechanical pencil	2	10

After being in the itemset 3 stage, the items in the sales data cannot use to form the next itemset because the number of past transactions has reached the maximum limit. Then the process that will be carried out next is the calculation of the confidence value.

The formula to calculate the Confidence value:

$$\text{Confidence} = P(B \mid A) = \frac{\Sigma \text{ Transaction amount contains A and B}}{\Sigma \text{ Transaction contains A}}$$

Table 9 Confidence of Itemset 3

No	X => Y	Support XUY	Support X	Confidence	Information
1	Butterfly ruler, vision book A4 = > giant mechanical pencil MP-116	10	10	100	Getaway
2	A4 vision book, mechanical pencil prominent MP-116 = > Butterfly ruler	10	10	100	Getaway
3	Butterfly ruler, butterfly ruler => A4 vision drawing book	10	10	100	Getaway
4	Butterfly ruler = > Big MP-116 mechanical pencil, vision book A4	10	10	100	Getaway
5	Vision A4 drawing book => Butterfly ruler, big mechanical pencil MP-116	10	10	100	Getaway
6	Big MP-116 mechanical pencil => Vision A4 drawing book, butterfly ruler	10	10	100	Getaway

Table 10 Confidence in itemset 2

No	X => Y	Support XUY	Support X	Confidence	Information
1	Butterfly ruler => A4 vision drawing book	10	10	100	Getaway
2	A4 vision book => Butterfly ruler	10	10	100	Getaway
3	Butterfly ruler => Big MP-116 mechanical pencil	10	10	100	Getaway
4	Big MP-116 mechanical pencil => Butterfly ruler	10	10	100	Getaway
5	Vision A4 drawing book => Big MP-116 mechanical pencil	10	10	100	Getaway
6	Giant MP-116 mechanical pencil => A4 vision drawing book	10	10	100	Getaway

In tables 9 and 10, Table X => Y represents item => item. The way to read it is if item X, then item Y.

Then Support X U Y is the support value of item X in column X => Y, Support X is the support value of Y in column X=>Y, and then from the two values of Support X U Y and Support X, the Confidence value can calculate.

Table 11 Association Rules formed

No	X => Y	Confidence
1	Butterfly ruler, A4 vision drawing book => giant mechanical pencil MP-116	100
2	Vision A4 drawing book, mechanical pencil prominent MP-116 => Butterfly ruler	100
3	Giant MP-116 mechanical pencil, butterfly ruler => A4 vision drawing book	100
4	Butterfly ruler => Big MP-116 mechanical pencil, vision book A4	100
5	Vision A4 drawing book => Butterfly ruler, big mechanical pencil MP-116	100
6	Big MP-116 mechanical pencil => A4 vision drawing book, butterfly ruler	100
7	Butterfly ruler => A4 vision drawing book	100
8	A4 vision book => Butterfly ruler	100
9	Butterfly ruler => Big MP-116 mechanical pencil	100
10	Big MP-116 mechanical pencil => Butterfly ruler	100
11	A4 vision drawing book => giant mechanical pencil MP-116	100
12	Giant MP-116 mechanical pencil => A4 vision drawing book	100

After all, processes are carried out, and it can conclude that the Structuring Solution consists of the association rules in table 11, which have been normalized to table 12.

Table 12 Setup Solutions

1	If the customer buys the BUTTERFLY LINER, THE VISION A4 PICTURE BOOK, the customer will also buy the BIG MECHANIC PENCIL MP-116.
	Solution: Place Nearby

2	If the customer buys the VISION A4 IMAGE BOOK, THE BIG MECHANICAL PENCIL MP-116, then the customer will also buy BUTTERFLY RULER.
	Solution: Place Nearby
3	If the customer buys MECHANICAL PENCIL BIG MP-116, BUTTERFLY LINER, then the customer will also buy VISION DRAWING BOOK A4.
	Solution: Place Nearby
4	If the customer buys the BUTTERFLY RULER, then the customer will also buy the BIG MECHANICAL PENCIL MP-116, VISION A4 GRAPHIC BOOK.
	Solution: Place Nearby
5	If the customer buys the VISION A4 PICTURE BOOK, then the customer will also buy BUTTERFLY RULER, BIG MECHANICAL PENCIL MP-116.
	Solution: Place Nearby
6	If the customer buys MECHANICAL PENCIL BIG MP-116, then the customer will also buy DRAWING BOOK VISION A4, BUTTERFLY RULER.
	Solution: Place Nearby
7	If the customer buys the BUTTERFLY RULER, then the customer will also buy the VISION A4 GRAPHIC BOOK.
	Solution: Place Nearby
8	If the customer buys the VISION A4 PICTURE BOOK, then the customer will also buy the BUTTERFLY RULER.
	Solution: Place Nearby
9	If the customer buys BUTTERFLY RULER, then the customer will also buy MECHANICAL PENCIL BIG MP-116.
	Solution: Place Nearby
10	If the customer buys PENCIL MEKANIK BIG MP-116, then the customer will also buy BUTTERFLY RULER.
	Solution: Place Nearby
11	If the customer buys the VISION A4 GAMBLE BOOK, then the customer will also buy MECHANICAL PENCIL BIG MP-116.
	Solution: Place Nearby
12	If the customer buys MECHANICAL PENCIL BIG MP-116, then the customer will also buy VISION DRAWING BOOK A4.
	Solution: Place Nearby

CONCLUSION

Based on the research design and discussion, it can infer that the following is the conclusion:

1. Application of Data Mining to determine product structuring uses a priori algorithm to:
   a. Obtain sales data from the research site. The sales data obtained is the original sales data from the research site.
   b. Normalize sales data according to system requirements. That is by ignoring the contents of the column other than the date and item column—the same note groups items.
   c. Implement a priori algorithm into the application.
   d. Enter normalized sales data into the application, then start the analysis process.
   e. The analysis results are obtained by the emergence of calculations and structuring solutions.

2. Displaying Product Rank as a value-added application in the following way:
   a. The results obtained are product rankings with ordering the highest number of transactions to the smallest number of transactions.
   b. The top product ranking is the product that has the greatest number of transactions. The lower it is, the fewer transactions that occur on the product. Users can determine the exact and appropriate layout and price by knowing the number of transactions for each product.

REFERENCES

Aprizal, Hasriani, & Ningsih. (2016). *Implementasi Data Mining Untuk Penentuan Posisi*. Techno.COM, 8.

Hariguna, Hasanah, & Susanti. (2018). Sales Transaction Data Analysis Using Apriori Algorithm to Determine the Layout of The Goods. *International Journal of Informatics and Information Systems*, 8.

Kusrini, & lutfi. (2009). *Algoritma Data MIning*. Yogyakarta: Andi.

Listriani, S. E. (2016). Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen. *Jurnal Teknik Informatika*, 8.

Maharani, Hasibuan, Silalahi, Nasution, Mesran, Suginam, et al. (2017). Implementasi Data Mining untuk Pengaturan Layout Minimarket dengan Menerapkan Association Rule. *Jurnal Riset Komputer (JURIKOM)*, 6.

Marsono. (2019). Penerapan Data Mining Pengaturan Pola Tata Letak Barang Pada Berkah Swalayan Untuk Strategi Penjualan Menggunakan Algoritma Apriori. *Informatika Dan Teknologi Jaringan*, 6.

Nursikuwagus, & Hartono. (2016). *Implementasi Algoritma Apriori untuk Analisis Penjualan Dengan Berbasis Web.* Simetris, 6.

Purba, & Ginting. (2018). Implementasi Algoritma Apriori Untuk Mencari Relasipada Transaksi Pembelian Alat-Alat Kesehatan (Studi Kasus: RS. ESTOMIHI). *Majalah Ilmiah INTI*, 7.

Tarigan. (2017). Penerapan Algoritma Apriori untuk Menentukan Barang Terlaris (Studi Kasus: PT. Indomarco Prismatama Medan). *Majalah Ilmiah INTI*, 2.

Winata, A. D. (2017). Analisis Dan Prediksi Penjualan Produk Terlaris Distro "Root Shoes" Dengan Aplikasiandroid. *Electronic Theses and Dissertations*, 18.

Wulandari, & Rahayu. (2014). Pemanfaatan Algoritma Apriori Untuk Perancangan Ulang Tata Letak Barang Di Toko Busana. *SNATI*, 6.