



Application of Data Mining for Sales Strategy at Ria Busana Using the A priori Algorithm

Nurayni Sinabang¹, Paska Marto Hasugian²

Informatics Engineering Study Program, STMIK Pelita Nusantara, Medan, Indonesia.

Article Info

Article history:

Received Okt 20, 2020

Revised Nov 24, 2020

Accepted Des 31, 2020

Keywords:

Data Mining,
Apriori Algorithm,
Sales Strategy.

ABSTRACT

Ria Busana supermarket is a national-scale fashion retailer spread across several cities in Indonesia. Where this company every day must meet consumer needs and are required to be able to take the right decisions in determining sales strategies. Problems that often occur are the accumulation of transaction data, product placement that is not optimal and the absence of an application used to determine sales patterns at Ria Busana. Data mining techniques are widely used to solve many problems, one of which is the a priori algorithm method used to obtain information about associations between products by utilizing a transaction database. Algorithm a priori is a logic of calculation which combines a product with another product through product history data. Based on the results of the trial with the a priori algorithm, it is hoped that it can help Ria Busana in determining the placement of goods according to products that consumers often buy so that from determining the right product for promotion which is done right on target.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Nurayni Sinabang,
Informatics Engineering Study Program,
STMIK Pelita Nusantara, Medan, Indonesia,
Jl. Iskandar Muda No.1, Merdeka, Kec. Medan Baru, Kota Medan, Sumatera Utara 20154.
Email: nuraynisinabang@gmail.com

1. INTRODUCTION

Ria Busana Supermarkets is a national fashion retailer spread across several cities in Indonesia. Companies that provide clothing products according to the trend, quality and affordable prices. Where this company every day must meet consumer needs and are required to be able to make the right decisions in determining sales strategies. This supermarket also always tries to provide the best service and a pleasant shopping atmosphere. However, due to the intense competition with other supermarkets, the right strategy is needed so that this fashion retailer can survive.

Based on the background of the problem above, the identification of the problem in this study is the accumulation of unused sales transaction data, the absence of an effective product sales strategy, suboptimal product placement, and no application to determine sales patterns. Based on the background that has been mentioned above, the authors describe the scope of the problem, namely the data that will be processed is the product sales data for May 2020 at Ria Busana, for the minimum support used is 10% and 50% confidence, the applications used in the design Web based with MySQL database. The formulation of the problem on research are How can analyze product sales transaction data at Ria Clothing? how do you apply the a priori method to increase product sales at Ria Busana? How can designing application data mining to determine the strategy of selling at Ria Clothing? The purpose of this research is for analyze product sales transaction data at Ria Clothing, apply A priori Algorithm to increase sales of products in the Ria Clothing, designing application of

data mining to determine patterns in the Ria Clothing sales. Based on the above background, the benefits of this research are able implement data mining with algorithm A priori strategy of selling, No Ria Clothing, can become a reference for students who will conduct research to develop further with different problems, give benefits Ria Busana as input to improve a more effective sales strategy, helping management in designing a sales strategy application at Ria Busana.

Data mining is a series of processes to extract additional value in the form of information that has not been known manually from a database. The resulting information is obtained by extracting and recognizing patterns that are important or interesting from the data contained in the database. Data mining is mainly used to find knowledge contained in large databases, so it is often called knowledge Discovery in Database (KDD) (Retno Tri Vlandari, 2017: 01). A priori algorithm is a basic algorithm proposed by Agrawal and Srikan in 1994 to find frequent itemset s in Boolean association rules. The main ideas of the a priori algorithm are: first, looking for frequent itemset (the set of items that meet the minimum support) from the transaction database, second - eliminating itemset with low frequency based on a predetermined minimum level of support. Next, build association rules from itemset that meet the minimum confidence value in the database (Dewi Listriani, Anif Hanifa Setianingrum, Fenty Eka MA, 2016). Sales are one of the marketing steps of a company, so that the company can get profits as a result of which the company's operational activities can continue to run. Sales are the gathering of a buyer and a seller with the aim of exchanging goods and services based on appropriate considerations. (Anthoy, Andeka Rocky Tanaamah, Agustinus Fritz Wijaya. 2017,138).

2. RESEARCH METHOD

The research framework is a collection of concepts arranged systematically in solving problems so that the research objectives are good.

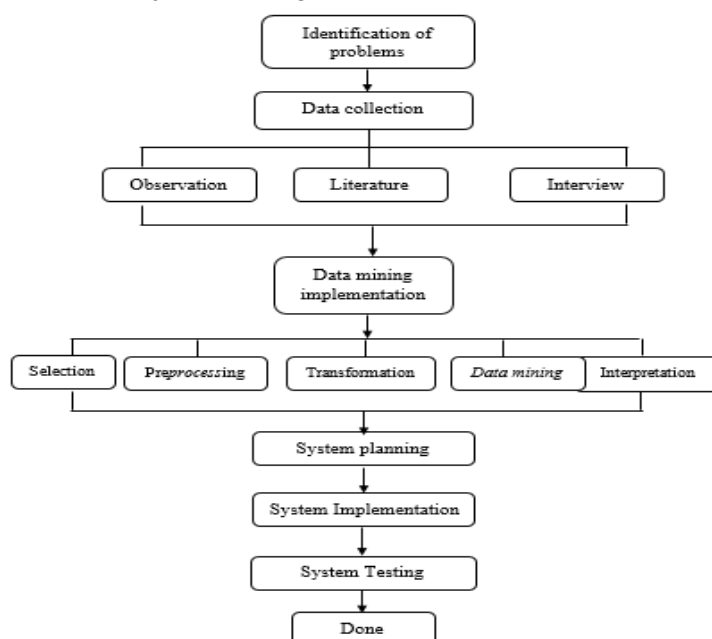


Figure 1. Research framework

Identification of Problems, these problems include the accumulation of unutilized sales transaction data, and the absence of an application to process the data. Data collection form literature is the process of finding references that match the problem, Observation is the activity of a process or object by observing directly to obtain clear information, Interview Is an oral communication process that aims to gather certain information in a structured way and carried out by two or more people. Data mining implementation that is Selection Aims to transform raw data into a format suitable for analysis. Preprocessing Aims to ensure that the results of the data mining process that are integrated into decision support systems are really valid results. Transformation is process or what is called the coding process is the process of transforming selected data so that the data becomes

useful and can be traced. Data mining is the process of finding suitable patterns for selected data using certain techniques or methods. Interpretation the pattern generated from the data mining process with a priori algorithm needs to be displayed in a form that is easily understood by the party who will use the system created, namely Ria Busana. System planning the system created is designed using UML. Web Based System Implementation, and last System Testing At this stage, activities aimed at measuring and assessing the results of the research that have been carried out are carried out.

3. RESULTS AND DISCUSSION

Analysis with the A priori Algorithm

In this section, we will describe the process of forming a product sales analysis at Ria Busana starting from discretion, determining *Support*, *Confidence*, to association rules. In this study, the minimum value of *support* used was 10% and the minimum *confidence* value was 50%.

Table 1. Test Data

NO	DATE	PRODUCT CODE	PRODUCT NAME
1	01-05-2020	2440	3 PANK
2	01-05-2020	0007	ADIYAT
3	01-05-2020	2270	ADONIS
10	02-05-2020	2482	AG
66	12-05-2020	0010	AKBAR
67	12-05-2020	0014	ALPHA

Itemset Formation 1

The process of forming a *support* for 1 *itemset* by utilizing the *Support* formula, where the minimum amount of *support* value specified is 10%. With a description of the calculation as follows:

$$\text{Support}(A) = \frac{\sum \text{transactions containing } A}{\sum \text{transactions}} \times 100\%$$

$$\text{Support}(3 \text{ PANK}) = \frac{10}{12} \times 100\% = 83.33\%$$

Table 2. Itemset 1

PRODUCT NAME	NUMBER OF TRANSACTIONS	SUPPORT (%)
3 PANK	10	83.33
ADIYAT	7	58.33
ADEEVHA	2	16.67
AGREE	10	83.33
Etc ..		

Combination of 2 *itemset*

The process of forming a *support* for 2 *itemset* by utilizing the *Support* formula, where the minimum amount of *support* value specified is 10%. The data used for the combination of 2 *itemset* is the data in table 2 above, with a description of the calculations as follows:

$$\text{Support}(A, B) = \frac{\sum \text{Transactions Contain } A \text{ and } B}{\sum \text{Transactions}} \times 100\%$$

$$\text{Support}(3 \text{ PANK, ADIYAT}) = \frac{7}{12} \times 100\% = 58.33\%$$

Table 3. *itemset* 2

PRODUCT NAME	NUMBER OF TRANSACTIONS	SUPPORT (%)
3 PANK, ADIYAT	7	58.33
3 PANK, ADEEVHA	2	16.67
3 PANK, ADONIS	4	33.33
Etc ..		

The same calculation is done to find the support value for itemset 3 to itemset 7

7 itemset combination

The process of forming a *support* for 7 itemset by utilizing the *Support* formula, where the minimum amount of *support* value specified is 10%.

$$\text{Support (3 PANK, ADIYAT, AGREE, AGREE KIDS, ALENA, AG, ALFA)} = \frac{2}{12} \times 100\% = 16,67\%$$

Table 4. Itemset 7

PRODUCT NAME	NUMBER OF TRANSACTION	SUPPORT (%)
3 PANK, ADIYAT, AGREE, AGREE KIDS, ALENA, AG, ALFA	2	16.67%

Establishment of Association Rules

After all high frequency patterns are found, then look for the association rule that meets the minimum requirements for *confidence* by calculating the associative *confidence* rule $A \rightarrow B$.

Minimum *confidence* = 50%

The *confidence* value of the rule $A \rightarrow B$ is obtained by the formula:

$$\text{Confidence (A)} = \frac{\sum \text{transactions contain A and B}}{\sum \text{transactions containing A}} \times 100\%$$

$$\text{Confidence (3 PANK, AG)} = \frac{9}{10} \times 100\% = 90\%$$

The same calculation to find the *confidence* value is carried out for all itemset 2 so that the results are as in Table 5 below:

Table 5. Result of Confidence 2 Itemset Calculation

PRODUCT NAME	NUMBER OF TRANSACTIONS	SUPPORT (%)	CONFIDENCE (%)
3 PANK, AG	9	75	90%
3 PANK, AGREE	8	66.67	80%
3 PANK, AGREE KIDS	6	50	60%
ADIYAT, AG	6	50	85.71%
ADEEVHA, AG	2	16.67	100%
Etc ..			

The same calculation is carried out to find the confidence value of itemset 3 to itemset 7

Table 6. Results of the 7 Itemset Confidence Calculation

PRODUCT NAME	NUMBER OF TRANSACTIONS	SUPPORT (%)	CONFIDENCE (%)
3 PANK, ADIYAT, AGREE, AGREE KIDS, AG, ALFA, ALENA	2	16.67	50%

The minimum *confidence* specified is 50%, the result is an association rule based on predetermined parameters, namely a minimum *support* of 10% and a minimum *confidence* of 50%, several rules or *rules* are formed, namely:

1. If consumers buy 3 PANK products, they will buy AG with a *confidence* value of 90%.
2. If consumers buy 3PANK, ADIYAT products, they will buy AGREE with a *confidence* value of 50%.
3. If consumers buy 3 PANK, ADIYAT, AG products, they will buy ADEEVHA with 100% *confidence* value.
4. If consumers buy 3 PANK, AGREE, AGREE KIDS, AKBAR products, they will buy ALENA with a *confidence* value of 75%.
5. If consumers buy 3 PANK, ADIYAT, AGREE, AGREE KIDS, AG products, they will buy ALFA with a *confidence* value of 50%.

6. If consumers buy 3 PANK, ADIYAT, AGREE, AGREE KIDS, AG, ALFA products, they will buy ALENA with a *confidence* value of 50%
7. If consumers buy 3 PANK, ADIYAT, AGREE, AGREE KIDS, ALENA, AG products, they will buy ALFA with a *confidence* value of 50%.

After the calculation, the system implementation is carried out which includes:

1. Program *Login* Page

To display the login Enter the registered username and password. The Enter button functions to agree and enter the main menu. This page appears for the first time when the system starts.

Figure 2. *Login* page

2. Transaction Data Page Display

The transaction data page is a page that displays the number, transaction, data, date and account. The following is a display of the transaction data page which can be seen in Figure 4.

No	Transactions	Data	Date	Action
1	T000001	3 PANK	2020-05-01	
2	T000001	ADIYAT	2020-05-01	
3	T000001	ADONIS	2020-05-01	
4	T000001	AGREE	2020-05-01	
5	T000001	AGREE KIDS	2020-05-01	

Figure 3. Weather Data Transactions

3. A priori Process page view

In this display, the a priori calculation process is carried out by entering the start date and end date of the transaction, then entering the minimum support and confidence, then clicking the calculate button.

1. Determination of Itemset Combinations

- a. Itemset Priori Process 1

After clicking the calculate button, itemset 1 to itemset 7 will appear with the number of transactions and support values.

C1 (Kandidat 1-itemset)			
No	Item1	Qty	Support
1	ag	11	91.67%
2	agree	10	83.33%
3	3 pank	10	83.33%
4	agree kids	8	66.67%
5	adiyat	7	58.33%
6	akbar	5	41.67%
7	alena	4	33.33%
8	alfa	4	33.33%
9	adonis	4	33.33%
10	adeevha	2	16.67%
11	agape 88	1	8.33%
12	algim	1	8.33%

Figure 4. The itemset priori process 1

b. Itemset Priori Process 7

C7 (Kandidat 7-itemset)									
No	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Qty	Support
1	3 pank	adiyat	agree	agree kids	alena	ag	alfa	2	16.67%
2	3 pank	agree	agree kids	akbar	alena	ag	alfa	1	8.33%
3	adiyat	agree	agree kids	akbar	alena	ag	alfa	1	8.33%
4	3 pank	adiyat	agree kids	akbar	alena	ag	alfa	1	8.33%
5	3 pank	adiyat	agree	akbar	alena	ag	alfa	1	8.33%
6	3 pank	adiyat	agree	agree kids	akbar	ag	alfa	1	8.33%
7	3 pank	adiyat	agree	agree kids	akbar	alena	ag	1	8.33%
8	3 pank	adiyat	agree	agree kids	akbar	alena	alfa	1	8.33%

Figure 5. Itemset Priori Process 7

2. Establishment of Association Rules

a. Itemset Association 2

After clicking the calculate button, the association or relationship of each itemset will appear starting from itemset 2 to itemset 7.

Asosiasi (2-itemset)						
No	Rule	Support	Confidence		Sup. * Conf.	Lift Ratio
1	If Consumers Buy 3 pank Then Buy alena	33.33%	4/4	100%	33.33	2
2	If Consumers Buy agree Then Buy alena	33.33%	4/4	100%	33.33	2
3	If Consumers Buy agree Then Buy agree kids	66.67%	8/8	100%	66.67	2
4	If Consumers Buy agree Then Buy akbar	41.67%	5/5	100%	41.67	2
5	If Consumers Buy adiyat Then Buy adeevha	16.67%	2/2	100%	16.67	2
6	If Consumers Buy agree kids Then Buy alena	33.33%	4/4	100%	33.33	2
7	If Consumers Buy ag Then Buy alfa	33.33%	4/4	100%	33.33	2
8	If Consumers Buy 3 pank Then Buy adeevha	16.67%	2/2	100%	16.67	2
9	If Consumers Buy 3 pank Then Buy adiyat	58.33%	7/7	100%	58.33	2
10	If Consumers Buy 3 pank Then Buy adonis	33.33%	4/4	100%	33.33	2
11	If Consumers Buy agree Then Buy alfa	33.33%	4/4	100%	33.33	2

Figure 6. Itemset Association 2

b. Itemset Association 7

Asosiasi (7-itemset)						
No	Rule	Support	Confidence	Sup. * Conf.	Lift Ratio	
1	If Consumers Buy 3 pank, adiyat, agree, agree kids, ag, alfa Then Buy alena	16.67%	4/2	50%	8.33	1
2	If Consumers Buy 3 pank, adiyat, agree, agree kids, alena, ag Then Buy alfa	16.67%	4/2	50%	8.33	1

Figure 7. Itemset Association 7

4. CONCLUSION

In the research on the application of data mining for sales strategies at Ria Clothing using the Apriori algorithm, several conclusions were obtained, namely can be solved by data mining techniques using a priori algorithm with association rules to determine products that consumers often

buy simultaneously. Ria Busana can provide satisfaction to consumers because the products they often buy are combined in one shelf so that consumers can save more time in buying products that they often buy simultaneously so that consumers feel more comfortable shopping.

REFERENCES

- [1] Al Irsyadi, FY (2014). Implementation of Data Warehouse and Data Mining for Determining Strategic Plans for Batik Sales (Case Study of Mahkota Laweyan Batik). *Communion* , VI (1), 42–58. <https://publikasiilmiah.ums.ac.id/handle/11617/4491>
- [2] Annur, H. (2019). Data Mining Application Determining Car Variation Sales Strategy Using K-Means Clustering Method. *Upgris Journal of Informatics* , 5 (1). <https://doi.org/10.26877/jiu.v5i1.3091>
- [3] Dharmawan, KD, & Sari, WS (2016). Website Development Using Iconix Process Method For Computer Sales Strategy In Cv. Citra Mandiri Semarang. (*Journal of Information Systems*) , 1 (2), 193–201. <https://publikasi.dinus.ac.id/index.php/joins/article/view/1309>
- [4] Haryanto, D., Oslan, Y., & Dwiyan, D. (2011). Shopping Cart Analysis Implementation with Association Rules Using Apriori Algorithm in Motorcycle Parts Sales. *Journal of Buana Informatics* , 2 (2), 81–94. <https://doi.org/10.24002/jbi.v2i2.311>
- [5] Kusrini, Lutfi taufiq Emha, (2009), *Data Mining Algorithm*, Andi Publisher, Yogyakarta, Research and Development, Division, (2004), *PHP and MySQL Program Applications*, Andi & Madcoms Publisher, Yogyakarta. West Java Regional Art Inventory, Cultural Park management center. <http://jagoanana.wordpress.com/category/document-classification/>, (15 November 2010). http://fpmipa.upi.edu/staff/yudi/stop_words_list.txt, (15 November 2010).
- [6] Latifah, VN, Furqon, MT, & Santoso, N. (2018). Implementation of a Modified-Apriori Algorithm to Determine Sales Patterns as a Strategy for Placing Goods and Promos. *Journal of Information Technology and Computer Science Development (J-PTIHK) Universitas Brawijaya* , 2 (10), 2829–2834.
- [7] Listriani, D., Setyaningrum, AH, & MA, FE (2016). Application of the Association Method Using the Apriori Algorithm in the Application of Consumer Shopping Patterns (Case Study of Gramedia Bintaro Bookstore). *Journal of Informatics Engineering Vol 9 No. 2*, Jakarta State Islamic University , 9 (2), 120–127. <http://journal.uinjkt.ac.id/index.php/ti/article/view/5602/3619>
- [8] Putri, TU, Izman, MH, & Dian, S. (2014). Application of Data Mining to Determine Sales Strategy at Gramedia Bookstore Using Clustering Method. *Bina Darma University* , I (Sales strategy), 1–10.
- [9] Sari, P., & Sinaga, B. (2018). 9. Data Mining Applications Using Apriori Algorithms For The Largest Sales Of Products At Cv. Sakura Photo. *E-Jurnal.Pelitanusantara.Ac.Id* , 22 (1), 34–38. <http://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/349>
- [10] Tama, B. (2010). Determining Sales Strategy Using Association Rules in the Context of CRM. *Generic Journal* , 5 (1), 79435.
- [11] AS Rosa, and M. Saladin. 2014. *Structural and Object-Oriented Software Engineering*. Bandung: Informatics.
- [12] Hidayatullah, Priyanto, and Jauhari Khairul Kawistara. 2017. *WEB programming*. Bandung. Bandung Informatics.
- [13] Vulandari Tri Retno 2017. *Data Mining Theory and Rapidminer Application*. Surakarta. Penernit Gava Media