http://ejournal.seaninstitute.or.id/index.php/InfoSains
Jurnal Info Sains : Informatika dan Sains, Volume 12, No 02, 2022
E-ISSN.2797-7889, P-ISSN.2089-3329

Analysis Of Consumer Purchasing Patterns At Giant Atk Store Using The Apriori Method

Reza Alamsyah[1], Indra Sidabutar[2]
[1]STMIK Methodist, Informatics Engineering, Binjai, Indonesia, [2*]STMIK Methodist, Information Systems, Binjai, Indonesia
Emial: indra3753@gmail.com

Keywords

Purchasing Patterns, Data Mining, A priori

Abstract. Data on purchases by consumers in storage will continue to grow. So the sales application at the Atk Gyant shop is less effective. Meanwhile, in the sales transaction data, there is valuable information such as knowing a trend or pattern carried out by consumers. Data Mining has several characteristics and functions. The characteristic of data mining which aims to find tendencies in combination patterns is to use the Apriori method. The Apriori method can be applied to analyze purchasing patterns by Atk Gyant store consumers with the results of 10 patterns, namely the highest support value for 1 item is 40%, 2 items is 26% and 3 items is 19%.

1. INTRODUCTION

In the era of globalization, in trade, consumer purchasing patterns need to be considered to facilitate the business carried out. Sometimes consumers have a little difficulty in finding the product they want to buy [1]. This is caused by less effective product preparation by the shop owner. This has a big impact on economic principles in achieving profits [2]. Purchasing pattern trends also often result in consumer disappointment [3]. The layout of goods is a very determining variable in the business world because it greatly influences consumer shopping behavior [4], [5]

Toko Atk Giant is a business that operates in the field of trading office stationery. Every day, office stationery at the Atk Giant store will be sold to consumers. So that the data purchased by consumers in storage will increase in size [6]. This can result in less effective sales applications that have been established. Meanwhile, in the sales transaction data, there is valuable information if carried out in-depth analysis [7], [8].

The benefit that can be gained from carrying out in-depth analysis on a sales database is by knowing a trend or pattern carried out by consumers so that based on this pattern it can be used as a guideline in arranging the layout of the goods being arranged. This is a common business strategy[9]. Analyzing sales data can be done using data mining techniques.

Data Mining is a process of the concept of extracting information in a database or what is often called Knowledge Discovery in Database (KDD)[10],[11]. However, to carry out the data mining stage, you must first clean attribute data that is not used in the analysis process or is often referred to as the data cleaning stage [12], [13]. Data Mining has several characteristics and functions. The characteristic of data mining which aims to find tendencies in combination patterns is by using the Apriori method [14]. Apriori is very efficient in forming combination patterns based on predetermined support and confidence values [15]. The sales data that will be analyzed in this research is sales transaction data for office stationery at the Giant ATK store. This data is sales data for a period of 100 days, namely from January 1 2022 – April 10 2022. And the minimum support value determined is 18% and minimum confidence 40%.

2. METHOD

Data mining is one of the Knowledge Discovery in Database (KDD) series. KDD relates to integration techniques and scientific discovery, interpretation and visualization of patterns in a number of data sets. The KDD stages are as follows:

1. Data Selection
   Existing data will be selected for data and attributes that will be used for the next process. Data selection from a set of operational data is carried out before the information mining stage. The selected data will be used for data processing mining, and stored in a separate file from operational data

2. Data Cleaning
   The Cleaning stage is the selected data that will undergo preliminary processing and data cleaning. Before the data mining process is carried out, cleaning needs to be carried out with the aim of removing duplicate data, eliminating inconsistent data
3. Data Transformation
   It is a transformation process of data that has been selected to be converted into a form suitable for processing.
4. Data Mining
   Data mining is the process of looking for interesting information or patterns in data that has been selected using certain techniques or methods. The choice of method or algorithm is very important because it greatly influences the goals and process of KDD overall.
5. Interpretation
   The resulting information pattern needs to be displayed in a form that is easily understood by interested parties

To find association rules from a data set, what you have to do is look for Frequent itemsets first. Frequent itemset is a collection of items that often appear together. Whether an association is important or not can be determined using two benchmarks, namely the support value and the confidence value. Support is the supporting value or percentage of a combination of an item in the database, while confidence is the certainty value, namely the strength of the relationship between items in the association [16]

Support is the percentage of a certain combination of items. Support value can be determined using the following formula[17]:

$$Support\ (A) = \frac{\sum \text{Transaksi Mengandung A}}{\text{Total Transaksi}}$$

Association rules are used to meet the minimum confidence requirements. The Confidence value from the A->B rule can be obtained using the following formula [17]:

$$Confidence\ P\ (B|A) = \frac{\sum \text{Transaksi Mengandung A dan B}}{\sum \text{Transaksi Mengandung A}}$$

3. RESULTS AND DISCUSSION

Results

Based on calculations using the Apriori method on the data used as test data, the results of this research are as follows:

Association Rule According to Min. Support And Min. Confidence

Item Type Initialization	Transaction A	Transaction A->B	Support (%)	Confidence (%)
PNG => KRT	37	26	26%	70%
PNG => MK	37	24	24%	65%
PNG => TS	37	26	26%	70%
PNG => FP	37	20	20%	54%
KRT => MK	40	21	21%	53%
KRT => TS	40	21	21%	53%
MK => FP	36	25	25%	69%
MK => TS	36	23	23%	64%
PNG => KRT, MK	37	19	19%	51%
PNG => MK, TS	37	20	20%	54%

1. PNG => KRT
   It means : If consumers buy graphite pencils then consumerswill buy writing paper.
2. PNG => MK
   It means : If consumers buy graphite pencils then consumersgoing to buy a paper folder.
3. PNG => TS
   It means : If consumers buy graphite pencils then consumerswill buy stamp ink.

4. PNG => FP
   It means : If consumers buy graphite pencils then consumerswill buy stamp ink.
5. KRT => MK
   It means : If consumers buyWriting Paper then consumers will buy Paper Folders
6. KRT => TS
   It means : If consumers buyWriting Paper then consumers will buy Stamp Ink
7. MK => FP
   It means : If a consumer buys a paper folder then the consumergoing to buy a plastic Folder.
8. MK => TS
   It means : If a consumer buys a paper folder then the consumerwill buy stamp ink
9. PNG => KRT, MK
   It means : If consumers buy graphite pencils then conSumen will buy writing paper and paper folders
10. PNG => MK, TS
    It means : If consumers buy graphite pencils then conSumen will buy writing paper and paper folders

Discussion

Based on the data obtained at the Gyant ATK Shop, the next stage carried out was data selection. After carrying out the data selection stages, the resulting data is as follows:

Table 1. Selection Data

No	Name of goods	Initialization	Number of Transactions
1	Graphite pencil	PNG	37
2	Colored pencils	PNW	17
3	Mechanical pencil	PNM	9
4	Gel pen	PLG	11
5	Water ink pen	PTA	10
6	Rollerball pen	PLR	3
7	Ballpoint pen	PLB	3
8	Permanent marker	SP	4
9	Whiteboard markers	SW	5
10	Highlighter marker	SH	5
11	Eraser	PHP	5
12	Ink eraser	IPM	4
13	Writing paper	KRT	40
14	Printer paper	KRP	3
15	Photocopy paper	KRF	5
16	Envelope paper	KRA	7
17	Lined notebook	BCB	7
18	Boxed notebook	BCK	13
19	Spiral notebook	BCS)	3
20	Staples	SLS	10
21	Stapler	SLRs	3
22	paper map	MK	36
23	Plastic folder	F.P	37
24	Straight ruler	PLU	36
25	Curved ruler	PLE	6
26	Geometry ruler	PGM	3
27	Masking tape	ST	7
28	Cutter	CT	2
29	Scissors	GT	8
30	Type X	TX	4

No	Name of goods	Initialization	Number of Transactions
31	Glue stick	L.S	2
32	Liquid glue	LC	3
33	Date stamp	STG	2
34	Address stamp	SAL	5
35	Ink stamp	T.S	36
36	Paper sticker	SK	13
37	Address labels	LA	2
38	Labels of various types	LBJ	2
39	String of raffia	TR	13
40	Paper binder	PK	35
41	Paper clip	K.K	13
42	Printer ink	T.P	7
43	Calculator	KL	12
44	Phone book	BT	35
45	Calendar	KLD	37
46	World map	PD	2
47	Flow diagram	DA	2
48	Art product	PS	2
49	Art materials	BS	1
50	Whiteboard	PPT	1

The next thing to do is clean the data which is known as data cleaning. Data cleaning at this stage aims to eliminate sales data at Giant Atk Stores that meet the minimum support value, as in the following 5 cases:

1. Support Pensil Frafit=$\frac{37}{100}$*100%=37%
2. Support Kertas Tulis=$\frac{40}{100}$*100%=40%
3. Support Map Kertas=$\frac{36}{100}$*100%=36%
4. Support Folder Plastik=$\frac{37}{100}$*100%=37%
5. Support Penggaris Lurus=$\frac{36}{100}$*100%=36%
6. Support Tinta Stempel=$\frac{36}{100}$*100%=36%
7. Support Pengikat Kertas=$\frac{35}{100}$*100%=35%
8. Support Buku Telepon=$\frac{35}{100}$*100%=35%

The results of data cleaning as a whole, data that meets minimum support are as follows:

Table 2. Data Cleaning

No	Name of goods	Initialization	Number of Transactions	Support (%)
1	Graphite pencil	PNG	37	37%
2	Writing paper	KRT	40	40%
3	paper map	MK	36	36%
4	Plastic folder	F.P	37	37%
5	Straight ruler	PLU	36	36%
6	Ink stamp	T.S	36	36%
7	Paper binder	PK	35	35%
8	Phone book	BT	35	35%

The sales data for Atk Giant stores that meet the minimum support value as in table 2 are as follows:

Table 3. Product Sales Transactions that Meet Min Support

No	Date	Sales Transactions							
		PNG	KRT	MK	F.P	PLU	T.S	PK	BT
1	01-Jan-22	1	1	1	1	1	0	0	0
2	02-Jan-22	0	0	1	1	0	0	1	0
3	03-Jan-22	0	1	1	1	0	1	0	1
4	04-Jan-22	0	0	1	0	0	1	0	1
5	05-Jan-22	0	0	1	1	0	1	1	1
6	06-Jan-22	1	0	1	1	0	1	0	0
7	07-Jan-22	1	1	1	1	0	1	1	1
8	08-Jan-22	1	0	1	1	1	1	1	1
9	09-Jan-22	1	1	1	1	0	1	1	1
10	10-Jan-22	1	1	1	1	0	1	0	1
11	11-Jan-22	1	1	1	1	1	0	1	1
12	12-Jan-22	1	1	1	0	0	1	1	0
13	13-Jan-22	1	1	1	1	1	0	1	1
14	14-Jan-22	0	1	1	1	0	1	1	1
15	15-Jan-22	1	1	1	0	0	1	0	0
16	16-Jan-22	1	1	0	1	0	1	1	1
17	17-Jan-22	1	1	0	1	1	1	0	1
18	18-Jan-22	1	1	1	1	1	1	1	0
19	19-Jan-22	1	1	1	1	0	1	0	0
20	20-Jan-22	1	1	1	1	0	1	1	0
21	21-Jan-22	1	1	1	0	1	0	0	1
22	22-Jan-22	1	1	1	1	1	1	1	1
23	23-Jan-22	0	1	0	1	0	1	1	1
24	24-Jan-22	1	1	0	0	1	1	0	1
25	25-Jan-22	1	1	0	0	0	1	0	1
				etc					

Based on the data in table 3, the next stage of data mining is carried out, namely by applying the a priori method iteration, namely finding the support value for a combination of items that meets the minimum support value.

Table 4. Combination of 2 Items

Item Type Initialization	Transaction A->B	Support (%)
PNG, KRT	26	26%
PNG, MK	24	24%
PNG, FP	20	20%
PNG, PLU	13	13%
PNG, T.S	26	26%
PNG, PK	12	12%
PNG, BT	15	15%
KRT, MK	21	21%
KRT, FP	18	18%
KRT, PLU	15	15%
KRT, TS	21	21%
KRT, PK	14	14%
KRT, BT	17	17%
MK, FP	25	25%
MK, PLU	16	16%
MK, TS	23	23%
MK, PK	15	15%

MK, BT	16	16%
FP, PLU	17	17%
FP, TS	19	19%
FP, PK	13	13%
FP, BT	14	14%
PLU, TS	8	8%
PLU, PK	7	7%
PLU, BT	10	10%
TS, PK	12	12%
TS, BT	16	16%
PK, BT	10	10%

Combinations of sales data itemsets at Giant Atk Stores that meet the minimum support value, as in the following 5 cases:

1. Support Pensil Grafit→Kertas Tulis=26/100*100%=26%
2. Support Pensil Grafit→Map Kertas=24/100*100%=24%
3. Support Pensil Grafit→Tinta Stempel=26/100*100%=26%
4. Support Map Kertas→Folder Plastik=25/100*100%=25%
5. Support Map Kertas→Tinta Stempel=23/100*100%=23%

The results of searching for the 2nd iteration of support values that meet the minimum support are as follows:

Table 5. Combination of 2 Items that Meet Min Support

Item Type Initialization	Transaction A->B	Support (%)
PNG, KRT	26	26%
PNG, MK	24	24%
PNG, T.S	26	26%
PNG, FP	20	20%
KRT, MK	21	21%
KRT, TS	21	21%
MK, FP	25	25%
MK, TS	23	23%
FP, TS	19	19%

The next step is to carry out the 3rd iteration, namely finding the min support value for 3 item combinations. The results are as follows:

Table 6. Combination of 3 items

Item Type Initialization	Transaction A->B	Support (%)
PNG, KRT, MK	19	19%
PNG, KRT, TS	17	17%
PNG, KRT, FP	15	15%
PNG, MK, TS	20	20%
PNG, MK, FP	16	16%
MK, FP, TS	14	14%
MK, FP, PNG	15	15%
MK, TS, KRT	15	15%
MK, TS, FP	15	15%

Based on the 3rd search iteration and continued with the presentation of data in table 6, the combination of 3 items that meet min support is as follows:

PNG, KRT, MK with a support value of 19%

PNG, MK, TS with a support value of 20%

Next, the 4th iteration is carried out, namely a combination of 4 items based on a combination of 3 items that meet the min value. support. The combination results obtained are: PNG, KRT, MK, TS and have 14 transactions so the support value is 14%. Therefore, none of the combination patterns

of 4 items are fulfilled and then the iteration is stopped. The next step is to find the confidence value for each combination of items using the formula explained in the methodology above.

4. CONCLUSIONS

Based on the results of the analysis, the conclusion obtained is that the Apriori method can be applied in analyzing purchasing patterns by consumers by recording the min value. support and min value. confidence is determined by the leader or even the owner. At the Atk Gyant Store there are 10 rules which are then said to be consumer purchasing patterns with the highest support value for 1 item being 40%, 2 items being 26% and 3 items being 19%. And for a combination of 4 items not found.

REFERENCE

[1] Bervan and Evan. 2008. Retail Management. Jakarta: Erlangga Publishers
[2] Indrianto, Nur., and Supomo. 2009.Business Research Methodology for Accounting and Management, First Edition. BPFE. Yogyakarta
[3] Sulistiorini, beautiful, (2017). The Influence of Store Atmosphere on Consumer Purchasing Decisions at Eramart Tenggiri Samarinda. Business Administration E-Journal. 5(2):11-520
[4] Masibbuk, I.,Silcyljeova Moniharapon, & Imelda Wj Ogi. (2019). The Influence Of Product Completeness, Service Quality And Layout On Purchasing Decisions (Case Study At Golden Supermarket In Manado City). EMBA Journal, 7(1), pp. 411-420
[5] Umam, K. and MansurM. (2017), The Influence of Store Location, Store Layout and Store Atmosphere on Consumer Satisfaction in Shopping at the An-nur Kita Cooperative PP An-nur 3 Murah Banyu Bululawang Malang. .E–Management Research Journal, 6(1), pp. 29-44
[6] Finn Lee S, Juan Santana, Data Mining: Forecasting Company Business, Jakarta: PT Elex Media Komputindo, 2010
[7] ME Saputri, "The Effect of Consumer Behavior Toward the Online Purchase of Fashion Products of Zalora Indonesia," Sosiotechnology, vol. 15, no. 2, pp. 291–297, 2016
[8] Salam, A., & Sholik, M. (2018). Implementation of the Apriori Algorithm to Search for Associations of Goods Sold on E-commerceOrderMas.Techno. com,17(2), 158-170
[9] Rodiyansyah, SF (2016). Apriori Algorithm for Shopping Cart Analysis on Sales Transaction Data, Infotech Journal, 36–39
[10] Herman, H. (2017). Analysis of Selection Patterns for Information Systems Science Concentration at STMIK TIME. TIMES Journal, 6(2), 37-43
[11] AWO Gama, IKGD Putra and IPA Bayupati, "Implementation of the Apriori Algorithm to Find Frequent Itemsets in Shopping Carts," Electrical Technology Scientific Magazine, vol. 15, no. 2, 2016
[12] Dennis, et al. Learn Data Mining with RapidMiner. Jakarta: Dennis Aprilla C, 2013
[13] Junianto, E.,& Rizal Rachman, "Application of Apriori and FP-Tree Data Mining Methods in Sales of Educational Media (Case Study: Oisha Smartkids)", Indonesian Journal on Computer and Information Technology (IJCIT), 5(2), pp. 117-125
[14] Batubara, Dinda Nabila, and Agus Perdana Windarto. 2019. "Data Mining Classification Analysis of the Level of Satisfaction of Visitors to the Pematang Siantar Animal Park Using Algorithms." COMIC (National Conference on Information and Computer Technology) 3(1):588–92
[15] Nursikuwagus, A., & Hartono, T. (2016). Implementation of a priori algorithms for web-based sales analysis. Symmetric: Journal of Mechanical Engineering, Electrical and Computer Science, 7(2), 701-706
[16] Santoso, H., Hariyadi, IP, & Prayitno, P. (2016). Data Mining Analyzes Product Purchasing Patterns Using the Apriori Algorithm Method. SEMNASTEKNOMEDIA ONLINE, 4(1), 3-7-19
[17] Auliadaya, SA,et al. (2019), "Analysis Of Product Purchase Patterns At Cimahiruko Store Using Algorithm Methods", Journal of Information Management (JAMIKA), 9(1), pp. 58-69