JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 4, No. 4. September 2022 DOI: https://doi.org/10.34288/jri.v4i4.194
Accredited rank 3 (SINTA 3), excerpts from the decision of the Minister of RISTEK-BRIN No. 200/M/KPT/2020

SALES LEVEL ANALYSIS USING THE ASSOCIATION METHOD WITH THE APRIORI ALGORITHM

Samuel[-1], Asrul Sani[-2*)], Agus Budiyantara[-3], Merliani Ivone[-4], Frieyadie[-5]

Teknik Informatika
Sekolah Tinggi Manajemen dan Ilmu Komputer Widuri
Jakarta, Indonesia
https://widuri.ac.id/
samuelsukmana100@gmail.com[1]; asrulsani@kampuswiduri.ac.id[2*)];
agusbudiyantara@kampuswiduri.ac.id[3]

Ilmu Kesejahteraan Sosial
STISIP Widuri
Jakarta, Indonesia
https://widuri.ac.id/
merlianiivone@gmail.com[4]

Sistem Informasi
Universitas Nusa Mandiri
Jakarta, Indonesia
https://nusamandiri.ac.id/
frieyadie@nusamandiri.ac.id[5]

(*) Corresponding Author

Abstract

The company does not yet know the pattern of consumer purchases because, so far, the sales transaction data has not been used correctly and does not have a unique method to determine consumer buying patterns. The problems on the company, this research was done to reprocess sales transaction data for 2018-2019 using data mining techniques with association methods and apriori algorithms. RapidMiner is a supporting application to find association rules derived from transaction data. Processed transaction data using the Knowledge Discovery in Database (KDD) approach. Thus, the company can determine consumer habits in buying goods from sales transaction data for 2018-2019. The results of this study are that in 2018, nine association rules were obtained, of which the best were CT G-246 ⇒ CT G-250 and CT G-250 ⇒ CT G-246. In 2019, nineteen association rules were received, of which the best were PN 0441, SK 0175 ⇒ SK 0530, and SK 0175, SK 0283, ⇒ SK 0530. From the best association rules, the goods in the Coat (imported), Pants, and Skirt categories are often bought together.

Keywords: Association Method, Apriori Algorithm, Association Rules, Knowledge Discovery in Database

Abstrak

Pihak perusahaan belum mengetahui pola pembelian konsumen karena selama ini data transaksi penjualan tidak dimanfaatkan dengan baik dan tidak mempunyai metode khusus untuk mengetahui pola pembelian komsumen. Untuk mengatasi permasalahan pada perusahaan, penelitian ini dibuat dengan tujuan untuk mengolah kembali data transaksi penjualan tahun 2018-2019 menggunakan teknik data mining dengan metode asosiasi dan algoritma apriori. RapidMiner merupakan aplikasi pendukung yang digunakan untuk menemukan aturan asosasi yang berasal dari data transaksi. Dalam pengolahan data transaksi menggunakan pedekatan Knowledge Discovery in Database (KDD). Dengan demikian, pihak perusahaan dapat mengetahui kebiasaan konsumen dalam membeli barang yang berasal dari data transaksi penjualan tahun 2018-2019. Hasil dari penelitian ini adalah pada tahun 2018, didapatkan sembilan aturan asosiasi yang dimana aturan asosiasi yang terbaik adalah CT G-246 ⇒ CT G-250 dan CT G-250 ⇒ CT G-246. Pada tahun 2019, didapatkan sembilan belas aturan asosiasi yang dimana aturan asosiasi terbaik adalah PN 0441, SK 0175 ⇒ SK 0530 dan

SK 0175, SK 0283 ⇒ SK 0530. Dari aturan asosiasi terbaik tersebut, barang-barang dengan kategori Coat (import), Pants, dan Skirt adalah kategori yang sering terbeli bersama.

Kata Kunci: Metode Asosiasi, Algoritma Apriori, Aturan Asosiasi; Knowledge Discovery in Database

INTRODUCTION

In the sales world, business people must develop strategies and advantages for their business. By carrying out sales activities daily, the data from these sales is increasing (Sikumbang, 2018). In addition to growing sales data, the stock of goods usually accumulates in warehouses because these goods are rarely purchased by consumers (Takdirillah, 2020). So far, this data is only used when calculating how many items are sold, how much sales turnover is per month, or looking at the stock of goods (Sani, 2016). Most companies, in terms of data collection, such as sales data from consumer data, only collect the data as archives and do not know what to do to process the data. The amount of available data not only confuses the company but also does not know what to do with the data, so it cannot produce the right decision for a company.

The ability and speed to process large amounts of data into meaningful information are needed by companies to formulate effective and efficient business strategies (Fitrina, Kustanto, & Vulandari, 2018). The large amount of data stored in the database must be converted into useful information using statistical, mathematical, artificial intelligence, or machine learning techniques. This process is known as data mining (Silvanie, 2020). Data mining is extracting information from large data sets to obtain valuable values in the form of essential knowledge that has not been known manually (Budiyasari et al., 2017). One of the methods found in data mining is the association method. The association method is used to find out consumer buying patterns in one transaction. This method aims to find a combination of items that appear simultaneously from the obtained association rules (Tana, Marisa, & Wijaya, 2018). Looking for association rules to detect concurrently purchased goods by consumers can use the apriori algorithm. The apriori algorithm is tasked with forming a pattern of combinations of items and high-frequency itemset and then testing whether the combination meets minimum support and minimum confidence requirements, which are the threshold values determined by the authors (Anggraeni, Hapsari Dita, Saputra, Ragisl, Noranita, 2013). In addition, to these two parameters, the lift parameter is also used to measure the importance of rules that have been formed based on the support and confidence values by providing information on whether product A was purchased together with product B (Wulandari & Ningsih, 2017). The supporting application used to find association rules is RapidMiner. The RapidMiner application is open-source software used as a solution in data mining, text mining, and predictive analysis (Wardani & Kristiana, 2020).

Several previous studies on data mining used the association method with the apriori algorithm to study the literature in search of similar references to obtain accurate sources. The following are related to previous studies, including the following:

In research conducted by (Wulandari & Ningsih 2017), the problem in this study is that the sales information system at the Pala Farma pharmacy has not been used to determine the proper decision-making in the inventory system. Hence, it is very influential on the level of sales. Applying the apriori algorithm can help find drug sales patterns based on the trend of selling drugs simultaneously. zIn addition, the test results with the apriori algorithm can also be used to arrange the drug layout closely to make it easier to find out the presence of the drug.

In research conducted by (Lestari & Hafiz, 2020), the problem in this study is that the Barbar Warehouse requires a method to analyze market share through sales patterns to determine consumer tendencies in buying goods. Issues in the Barbar Warehouse can be overcome by processing sales transaction data using data mining techniques with apriori algorithms to decide what products are sold and produce decisions in spending and sales.

Research conducted by (Putra, Raharjo, Sandi, Ridwan, & Prasetyo, 2019), the problem in this research is that the development of retail companies is currently overgrowing, so a unique strategy is needed to increase the company's turnover. The data mining process can be used to improve the company's turnover by using the apriori algorithm to generate association rules so that it can be used as a pattern of purchasing goods by consumers. In addition, the company can increase the company's turnover by looking at the results of processing sales transaction data using the apriori algorithm.

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v4i4.194
Accredited rank 3 (SINTA 3), excerpts from the decision of the Minister of RISTEK-BRIN No. 200/M/KPT/2020

The author researched companies engaged in garment retail. This company sells pants, skirts, minidresses, blazers, tunics, Etc. So far, the company still uses the difference between the goods produced and the goods sold in viewing the sales results of existing products. In addition, consumers make transactions to buy more than one type of goods.

Based on interviews conducted by the author with the HRD Manager of the company. In one year the goods sold are around three thousand pcs, meanwhile the sales target in one month is about two hundred pcs. Sales transaction recording activities using the Accurate application and Microsoft Excel. The company uses the application to take consumer suggestions, and consumers can see the types of goods that consumers like. The company does not use a unique method to determine consumer buying patterns. Hence, the company can find out the habits of consumers who buy more than one type of goods in one transaction by looking at the transaction records in the Accurate application and Microsoft Excel.

The use of data mining, and the association method with a priori algorithm, is a solution to provide a company picture of what goods are often purchased by consumers and the pattern of relationships between one interest and another as a benchmark for companies in producing goods. In addition, it provides information about the arrangement of items that should be placed close together so that consumers can easily find them. This way, sales transaction data can be helpful as more good information and not only stored as archives.

Based on the explanation above, this study aims to generate consumer buying patterns sourced from the company's sales transaction data in 2018-2019 using a supporting application, RapidMiner, to look for association rules. In addition, these results indicate what types of goods must be sold and come from any category.

RESEARCH METHODS

Types of Research

This study used qualitative research methods. Qualitative research aims to find answers to an event or question by applying systematic scientific procedures. It is carried out using various methods such as interviews, observations, and documents (Sugiyono, 2015).

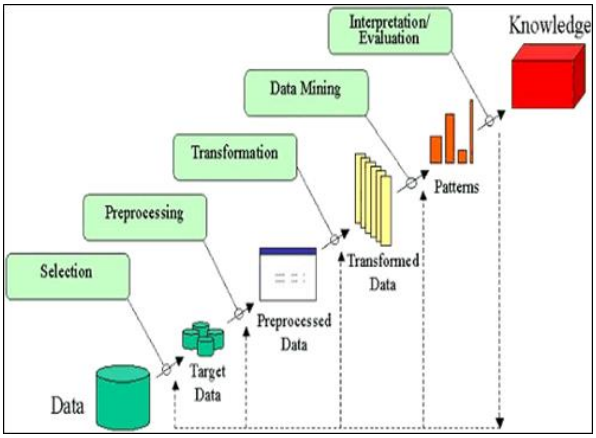
Time and Place of Research

The author researched retail garment companies to get sales transaction data from the company. The time to carry out research in that place is one month.

Research Target/Subject

Targets/subjects in this study will produce consumer buying patterns on the company using transaction history data that occurred in 2018-2019 and applies data mining techniques, namely the association method with the apriori algorithm, to find the association rules.

Knowledge Discovery in Database (KDD)



Source: (Amrin Amrin, 2017)

Figure 1. Process on KDD

Based on Figure 1, data mining, also known as Knowledge Discovery in Database or KDD, is an activity that includes collecting and using historical data to find regularities, patterns, or relationships in large data sets. The process in KDD are:

1. Data selection

At this stage, the selection of data for use based on the purpose is stored in a separate file from operational data. In the operational data, there are attributes of date, purchase ID, item code, and the number of items sold, but only two points are used, namely, the date and object code, because the information contained in it already represents the explanation needed for a research indicator. Furthermore, the data that has been selected will be saved into a separate excel file.

2. Prepossessing

At this stage, cleaning the data is the point of KDD's attention. This data cleaning includes removing duplicate data, checking inconsistent data, and correcting typographical errors. From the author's checking, one item code was miswritten (typography), namely CT G--248. The correct writing was CT G-248.

3. Transformation

After the data has passed the selection and refinement stage, it will be converted to tabulation format. So the tabulated data contains binary numbers 0 and 1. 0 means damaging / no goods are purchased while one means positive/there are goods purchased.

4. Data mining

At this stage, the process of searching for new patterns or valuable information by combining specific methods and algorithms. So after the tabulated data has been created, it will then be processed using the RapidMiner application to find the association rules from using the association method with the apriori algorithm.

5. Evaluation

At this stage, evaluate the results of the data mining process against the association rules by looking at the highest support, confidence, and lift parameters.

Data, Instruments, and Data Collection.

This study's primary data comes from interviews, and the company's sales transaction data is given directly by HRD Manager. The secondary data comes from journals and books related to this research. To support this research, the author uses several data collection methods. In observation and literature study, the author makes observations and goes directly to the field to get important information about the sales system at the company. At this stage, interviews were conducted with the HRD Manager and employees involved in product sales transactions to get the data needed in the research.

Data Analysis Technique

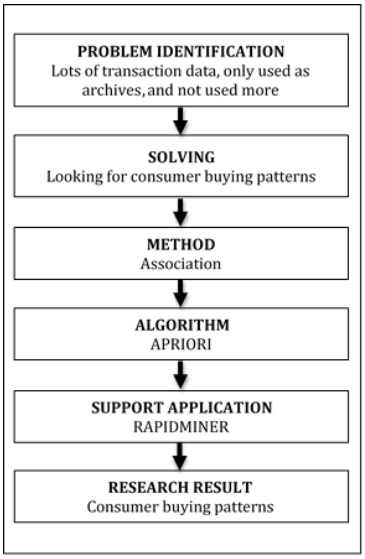


Figure 2. Research Stages

The explanation in Figure 2 above is:

1. Problem identification

The research stage begins with determining the research problem. The company has a lot of sales transaction data, but it is only used as an archive without being reused. Whereas the sales transaction data can be reprocessed to obtain essential information.

2. Solving

To overcome the problems faced by the company, sales transaction data can be reused to find consumer buying patterns so that relationships between items are found.

3. Method

The method used in this study is the association method which aims to find the link between goods purchased at one time from the company's sales transaction data in 2018-2019.

4. Algorithm

Support is a parameter that proves the level of dominance of an item/itemset from the total transactions. Confidence is a parameter that describes the relationship between two items conditionally (Oktaviani, TM Napitupul, Sarkawi, & Yulianti, 2019). The apriori algorithm finds each rule in the association method that has met the minimum support and confidence requirements.

The following steps are carried out in the calculation with the apriori algorithm (Qoniah & Priandika, 2020):

a. Analysis of the highest frequency pattern

This stage is looking combination of items that meet the minimum requirements of the support value of an object. The formula is:

$$Supp\ (A) = \frac{\sum trans\ contain\ A}{total\ trans} \quad \text{..... (1)}$$

The formula to find the support value of two items is:

$$Supp\ (A \cap B)\ = \frac{\sum trans\ contain\ A\ and\ B}{total\ trans} \quad \text{..... (2)}$$

b. Establishment of association rules

After all the highest frequency patterns are found, then the association rules that meet the minimum confidence requirements will be searched with the formula:

$$Conf\ =\ P(B|A) = \frac{\sum trans\ contain\ A\ and\ B}{total\ trans\ contain\ A} \quad \text{..... (3)}$$

In addition to the support and confidence parameters, the lift parameter is also needed to measure how important the rule formed from the support and confidence parameters is. The lift

parameter shows whether or not the transaction process is valid and provides information on whether product A was indeed purchased at the same time as product B (Sani, 2016).

$$Lift = \frac{supp\ (A \cap B)}{supp\ A\ x\ supp\ B} \quad \text{..... (4)}$$

The value of lift > 1 shows that these rules have a strong relationship. Meanwhile, if the lift value is < 1, the rule has not related, or the associative strength is low.

5. Support application

The author uses RapidMiner as a supporting application to analyze the company's sales transaction data in 2018-2019 to get the rules from the sales transaction data.

6. Research result

This research produces consumer buying patterns so that from these results, it can be seen what types of goods need increased sales and come from any category. In addition, from the design of purchases, the company or management can also find the excellent positioning of goods, so they are easy to find.

RESULTS AND DISCUSSION

Product sales transaction data will continue to grow daily and cause massive data storage. Most product sales transaction data are only used as archives without being appropriately used. Hence, companies need a method or technique to turn any data into helpful information for decision-making using the association method with apriori algorithms to determine consumer buying patterns.

Name of Item Codes

Table 1. List of Item Codes and Item Names

Code	Name
BR	Blazer
BL / BL G	Blouse
CT / CT G	Coat
DR / DR G	Dress
JS	Jumpsuit
MD	Mini Dress
MM	Muslim
PN / PN G	Pants
SC	Scarf
SK / SK G	Skirt
ST	Suit
TN	Tunic

Table 1 shows the item code with the letter G as an imported item. The number code behind the letter is the type of fabric used. For example, CT G-250 (CT = Coat, G = imported, 250 = type of fabric used).

Input Data

The test is carried out by inputting data, namely the date and item codes from 2018-2019.

	A	B	C	D	E	F	G	H	I	J	K
1	Januari			Februari			Maret			April	
2	08/01/2018	DR 1274		02/02/2018	PN 0773		09/03/2018	DR 1151		04/04/2018	BR 1235
3	10/01/2018	DR 1152		06/02/2018	BL 1279			DR 1194			DR 1249
4	12/01/2018	PN 0677			DR 1265			PN 1106		05/04/2018	DR 1231
5		PN 0838		07/02/2018	DR 1231			TN 1278			DR 1273
6		PN 1239			SK G-198		12/03/2018	BL 1279			DR 1274
7	13/01/2018	SC 1272		08/02/2018	BL 0673			DR 0864			DR 1286
8	16/01/2018	BL 0673			BL 1233			DR 1290			DR 1290
9		BL 1039			BL 1251			DR 1293			SK 1237
10		BL 1232			BR 1119			DR 1297			TN 1294
11		BL 1233			TN 1217			PN 1053		06/04/2018	DR 1194
12		BL 1267		09/02/2018	DR 1188			SK 0283			DR 1314
13		BL 1269		10/02/2018	PN 1238			SK 0755			DR 1320
14		BL 1279		12/02/2018	SK 0755		13/03/2018	DR 0798		07/04/2018	DR G-094
15		BL 1280		13/02/2018	DR 0864			DR 1131		09/04/2018	PN 0844
16		BR 0774			DR 1020		14/03/2018	BR 1234		10/04/2018	BL 0673
17		BR 1119			DR 1225			BR 1235			BL 1177
18		BR 1234			DR 1250			CT G-193			BR 1234
19		BR 1235			DR 1290			DR 1249			DR 1259
20		BR 1302			DR 1291			DR 1256			DR G-242
21		BR 1303		15/02/2018	DR G-176			DR 1296			PN 1106
22		DR 0961			PN 0441			DR G-237			PN 1238

Figure 3. Example of Transaction Data in 2018

Furthermore, the sales transaction data will be made into tabulated data which is analyzed using the association method with the apriori algorithm. Tabulated data is saved in excel format. Figure 3 is an example of sales transaction data in 2018.

	B	C	D	E	F	G	H	I
1	BL 0624	BL 0673	BL 0818	BL 0920	BL 0958	BL 0976	BL 1039	BL 1157
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	1	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0
20	0	1	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0

Figure 4. Example of Data Tabulation

Figure 4 shows that the tabulated data contains binary numbers 0 (negative) and 1 (positive). Then the tabulation data is imported using the RapidMiner application. The RapidMiner application will read all the attribute lists according to the table structure in the tabulated data that has been created, as shown in Figure 5.

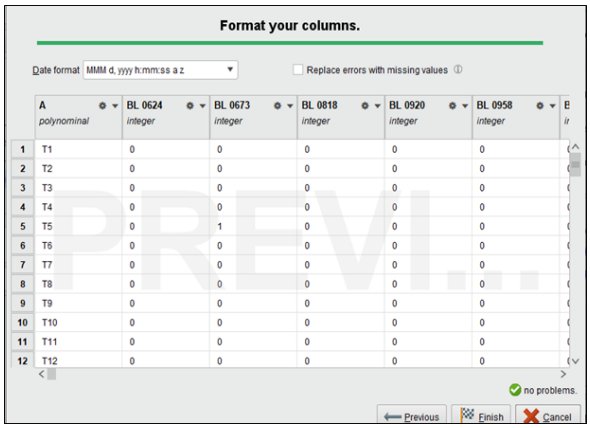


Figure 5. Tabulated Data Attribute List Display in RapidMiner

Minimum Limit of Support and Confidence Parameters

The minimum support and confidence limits are determined through observation from several experiments on the database used. The minimum support limit used is 0.015, and the minimum confidence limit used is 0.4. The determination of these values is adjusted to produce good association rules.

Association Rule Results

The number of association rules produced provides many regulations in the sales transaction data, thus providing many possibilities for decision-making. Not all association rules as a reference, so only association rules that have the highest confidence and lift values are used.

1. Rules in 2018

No.	Premises	Conclusion	Support	Confidence	Lift
1	PN 0814	PN 1008	0.019	0.400	7.100
2	PN 0161	PN 0441	0.028	0.500	8.875
3	PN 0441	PN 0161	0.028	0.500	8.875
4	BR 1234	BL 1232	0.019	0.571	24.343
5	DR 1288	DR 1189	0.019	0.571	20.286
6	DR 1189	DR 1288	0.019	0.667	20.286
7	BL 1232	BR 1234	0.019	0.800	24.343
8	CT G-246	CT G-250	0.019	1	53.250
9	CT G-250	CT G-246	0.019	1	53.250

Figure 6. Association Rules 2018 in RapidMiner

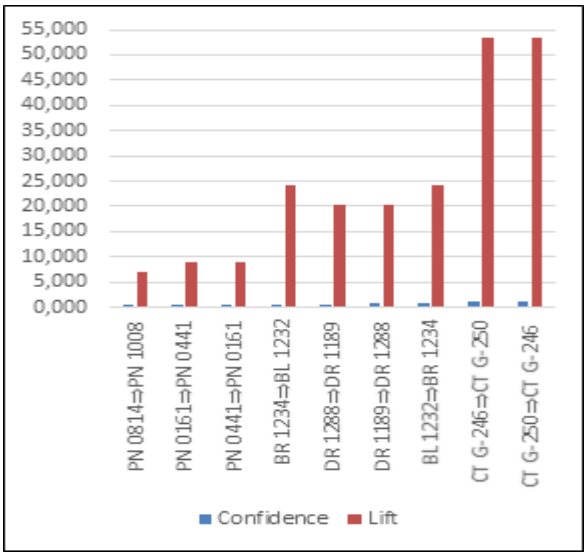


Figure 7. Association Rules Graph in 2018

The regulations obtained achieved the highest results by rules number 8 and 9 because they got the highest confidence results with a magnitude of 100% and produced the highest lift with a volume of 53,250. In Figure 6 and Figure 7, the rules produced in 2018 were nine. The explanation of these rules is:

- Someone who buys a CT G-246 (Coat) has a 100% chance of purchasing a CT G-250 (Coat). This possibility can be valid because it gets a lift of 53,250.
- Someone who buys a CT G-250 (Coat) has a 100% chance of purchasing a CT G-246 (Coat). This possibility can be valid because it gets a lift of 53,250.

2. Rules in 2019

No.	Premises	Conclusion	Support	Confidence	Lift
1	PN 0814	PN 0838	0.017	0.400	7.633
2	PN 0161	PN 0773	0.022	0.417	7.951
3	PN 0773	PN 0161	0.022	0.417	7.951
4	SK 0530	PN 0441	0.017	0.444	8.481
5	SK 0283	SK 0175	0.017	0.444	9.253
6	SK 0530	PN 0441, SK 0175	0.017	0.444	25.444
7	SK 0283	SK 0175, SK 0530	0.017	0.444	16.963
8	SK 0530	SK 0175, SK 0283	0.017	0.444	25.444
9	SK 0175	SK 0530	0.026	0.545	13.879
10	SK 0283	SK 0530	0.022	0.556	14.136

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v4i4.194
Accredited rank 3 (SINTA 3), excerpts from the decision of the Minister of RISTEK-BRIN No. 200/M/KPT/2020

No.	Premises	Conclusion	Support	Confidence	Lift
11	SK 0530	SK 0283	0.022	0.556	14.136
12	PN 1120	PN 0838	0.017	0.571	10.905
13	SK 0530	SK 0175	0.026	0.667	13.879
14	SK 0175, SK 0530	PN 0441	0.017	0.667	12.722
15	SK 0175, SK 0530	SK 0283	0.017	0.667	16.963
16	SK 0283, SK 0530	SK 0175	0.017	0.800	16.655
17	PN 0441, SK 0175	SK 0530	0.017	1	25.444
18	PN 0441, SK 0530	SK 0175	0.017	1	20.818
19	SK 0175, SK 0283	SK 0530	0.017	1	25.444

Figure 8. Association Rules 2019 in RapiMiner

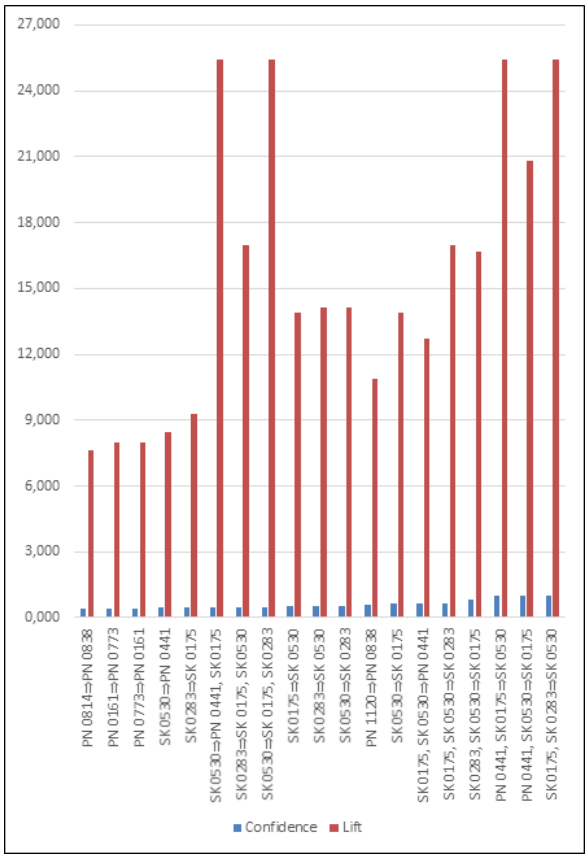


Figure 9. Association Rules Graph in 2019

From Figure 8 and Figure 9, the rules produced in 2019 were nineteen rules. From the regulations obtained, the highest results were achieved by rules number 17 and 19 because they got the highest confidence results with a magnitude of 100% and produced the highest lift with a volume of 25,444. The explanation of these rules is:

- Someone who buys PN 0441 (Pants) and SK 0175 (Skirt) has a 100% chance of purchasing SK 0530 (Skirt). This possibility can be valid because it gets a lift of 25,444.
- Someone who buys SK 0175 (Skirt) and SK 0283 (Skirt) has a 100% chance of also purchasing SK 0530 (Skirt). This possibility can be valid because it gets a lift of 25,444.

Analysis of Rule Results in 2018-2019

Analyzing the results is to discover more about the critical information generated from the rules.

1. In 2018, PN 0814 was bought with PN 1008, but in 2019, PN 0814 was purchased with PN 0838.
2. In 2018, PN 0161 was bought with PN 0441, but in 2019, PN 0161 was purchased with PN 0773.
3. In 2018, PN 0441 was bought with PN 0161, but in 2019, PN 0441 was purchased with SK 0175 and 0530.
4. In 2018, the pants category was dominated. Meanwhile, in 2019, it was dominated by the pants and skirts category.
5. It can happen because of the possible influence of clothing styles that continue to grow, and the average person follows the latest clothing models. In 2018, the Coat (imported) category became the best association rule. But in 2019, the type of pants and skirts became the best association rules.
6. In 2018 and 2019, the highest support was achieved by PN 0161 and PN 0441 at 0.028, while the lowest support was achieved by PN 0814, SK 0283, SK 0530, PN 1120, PN 0838, PN 0441, and SK 0175 at 0.017. It means that if the items get a higher support value, they have a significant level of dominance in the transaction data.
7. In 2018 and 2019, the highest confidence was achieved by CT G-246, CT G-250, PN 0441, SK 0175, SK 0530, and SK 0283 of 1,000, while the lowest confidence was achieved by PN 0814, PN 1008, and PN 0838 of 0.400. They have a strong relationship if the items get a higher confidence value.
8. In 2018 and 2019, the highest lift was achieved by CT G-246 and CT G-250 of 53,250, while the lowest was achieved by PN 0814 and PN 1008 of 7,100. It means that if the items get a higher lift value, these items have the truth that they can be bought together.
9. In 2019, PN 0441, SK 0530, ⇒ SK 0175 also received the highest confidence with 1,000, but the lift produced was slightly lower at 20.818. It has little effect on the strength of the association of the items.
10. In 2018, PN 0161 ⇒ PN 0441 and PN 0441 ⇒ PN 0161 received the highest support, but the

strength of the relationship between these items can be said to be low because they produce confidence of 0.500 while the highest confidence is 1,000.

11. In 2019, SK 0175 ⇒ SK 0530 and SK 0530 ⇒ SK 0175 received the highest support, but the strength of the relationship between these items can be said to be low because they produce confidence with a large 0.026 while the highest confidence is 1,000.

Recommendation for The Company

The results of this study are expected to provide input and an overview of the company's management to take advantage of data mining technology, especially the association method. The company can make decisions about producing goods from the resulting association rules, especially fast-selling items.

The arrangement of goods in groups based on the results of the association rules must be appropriate because the better the association rules produced, the more goods are often bought together. Companies can place items that have the best association rules close together, making it easier for consumers to find the items they want. Another thing that can be done is the creation of a sales brochure that places items with the best association rules on one page so that consumers can easily see these items.

CONCLUSIONS AND SUGGESTIONS

Conclusion

Observations made at the company found that the attributes of sales transaction data can be used to perform data mining analysis using the association method with the apriori algorithm. Results obtained from the RapidMiner application in 2018, the Coat (import) category became the best association rule, and in 2019, the Pants and Skirt category became the best association rule. The association rules resulting from using the RapidMiner application between 2018 and 2019 are different because consumer buying patterns are constantly changing and influenced by clothing styles.

Other important information that the company can know is that the arrangement of goods based on the highest support can be placed at the beginning because consumers most often purchase these goods, and the arrangement of goods based on the highest confidence can be placed side by side because of the possibility that consumers will buy these goods together.

Suggestion

An accurate association rule can be made by analyzing it every three or four months of the year because consumer buying patterns are constantly changing. Sales transaction data used is not only in 2018 and 2019 but can be from 2020 and 2021 to get more critical information from the resulting association rules.

The use of other methods with suitable algorithms can be combined for further research to obtain more information. For example, combining the classification method with the naïve Bayes algorithm to determine the categories of goods that sell and do not sell. In addition, a data mining application can be developed so that it is easy to use for data processing.

REFERENCES

Amrin Amrin. (2017). Data Mining Dengan Algoritma Apriori untuk Penentuan Aturan Asosiasi Pola Pembelian Pupuk. *Paradigma*, *XIX*(1), 74–79. https://doi.org/https://doi.org/10.31294/p.v19i1.1836

Anggraeni, Hapsari Dita, Saputra, Ragisl, Noranita, B. (2013). Aplikasi Data Mining Analisis Data Transaksi Penjualan Obat Menggunakan Algoritma Apriori. *Mining of Massive Datasets*, *2*(January 2013), 5–20. Retrieved from https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part

Budiyasari, V. N., Studi, P., Informatika, T., Teknik, F., Nusantara, U., & Kediri, P. (2017). Implementasi Data Mining Pada Penjualan kacamata Dengan Menggunakan Algoritma Apriori. *Indonesian Journal on Computer and Information Technology*, *2*(2), 31–39.

Fitrina, N., Kustanto, K., & Vulandari, R. T. (2018). Penerapan Algoritma Apriori Pada Sistem Rekomendasi Barang Di Minimarket Batox. *Jurnal Teknologi Informasi Dan Komunikasi (TIKomSiN)*, *6*(2), 21–27. https://doi.org/10.30646/tikomsin.v6i2.376

Lestari, A. F., & Hafiz, M. (2020). Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse. *INOVTEK Polbeng - Seri Informatika*, *5*(1), 96. https://doi.org/10.35314/isi.v5i1.1317

Oktaviani, A., TM Napitupul, G., Sarkawi, D., & Yulianti, I. (2019). Penerapan Data Mining Terhadap Penjualan Pipa Pada Cv. Gaskindo Sentosa Menggunakan Metode Algoritma Apriori. *Jurnal Riset Informatika*, *1*(4), 167–

172. https://doi.org/10.34288/jri.v1i4.96

Putra, J. L., Raharjo, M., Sandi, T. A. A., Ridwan, R., & Prasetyo, R. (2019). Implementasi Algoritma Apriori Terhadap Data Penjualan Pada Perusahaan Retail. *Jurnal Pilar Nusa Mandiri*, *15*(1), 85–90. https://doi.org/10.33480/pilar.v15i1.113

Qoniah, I., & Priandika, A. T. (2020). Analisis Market Basket Untuk Menentukan Asossiasi Rule Dengan Algoritma Apriori (Studi Kasus: Tb. Menara). *Jurnal Teknologi Dan Sistem Informasi*, *1*(2), 26–33.

Sani, A. (2016). Analisa Penjualan Retail dengan Metode Association Rule untuk Pengambilan Keputusan Strategis Perusahaan: Studi Kasus PT. *XYZ. Infotech*, (September). Retrieved from https://www.researchgate.net/profile/Asrul-Sani/publication/327680554_Analisa_Penjualan_Retail_Dengan_Metode_Association_Rule_Untuk_Pengambilan_Keputusan_Strategis_Perusahaan_Studi_Kasus_Pt_Xyz/links/5b9e8660299bf13e60373b02/ANALISA-PENJUALAN-RETAIL-DENGA

Sikumbang, E. D. (2018). Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori. *Jurnal Teknik Komputer AMIK BSI (JTK)*, *Vol 4*, *No.*(September), 1–4.

Silvanie, A. (2020). Pencarian Frequent Itemset dengan Algoritma Apriori dan Python. Studi kasus: Data Transaksi Penjualan Eceran Online di UK. *Jurnal Nasional Informatikan*, *1*(2), 103–113.

Sugiyono. (2015). *Metode Penelitian Kuantitatif, Kualitatif dan R & D*. Alfabeta, Bandung.

Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic : Jurnal Pendidikan Informatika*, *4*(1), 37–46. https://doi.org/10.29408/edumatic.v4i1.2081

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. *J I M P - Jurnal Informatika Merdeka Pasuruan*, *3*(2), 17–22. https://doi.org/10.37438/jimp.v3i2.167

Wardani, F. A. K., & Kristiana, T. (2020). Implementasi Data Mining Penjualan Produk Kosmetik Pada PT. Natural Nusantara Menggunakan Algoritma Apriori. *Paradigma - Jurnal Komputer Dan Informatika*, *22*(1), 85–90. https://doi.org/10.31294/p.v22i1.6520

Wulandari, D. A. N., & Ningsih, L. (2017). Data Mining Market Basket Analysis Menggunakan Algoritma Apriori Untuk Menentukan Persediaan Obat. *Konferensi Nasional Imu Sosial & Teknologi (KNiST)*, *1*(1), 227–235.