International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT22868

Pattern Recognition of Customer Spending Habits Using Apriori Algorithms in DataMining as an Inventory Strategy

Arief Jananto[1], Yohanes Suhari[2], Rara Sriartati Redjeki[3], Bambang Sudiyatno[4]

[1,2,3]Information Systems Department, Faculty of Information Technology and Industry, Stikubank University, Semarang, Indonesia

[*4]Department of Management, Fakulty of Economics and Business, Stikubank University, Semarang, Indonesia

Article Info

Publication Issue :
Volume 8, Issue 6
November-December-2022

Page Number : 104-115

Article History
Accepted: 06 Nov 2022
Published: 12 Nov 2022

ABSTRACT

The readiness of product inventory is very important, product shortages related to other products can make buyers disappointed and then cancel to buy products that were previously planned to be purchased at once. Sellers can experience a decrease in the number of sales to revenue. In this case, the seller needs to know the pattern of customer habits when making purchases by going through sales transaction data that has occurred. Association techniques can be used to analyze the pattern of interrelationships between items in transaction events. With the a priori algorithm as a popular association algorithm, the pattern of sales transaction data can be analyzed through the research stage. From the implementation of the algorithm with 1063 transaction data using 10% min-support and 75% min-confidence resulting in 4 association rules where 1) if you buy "kacer" and "love bird" you will buy "pentet" as much as 17% support, 2) if you buy "magpie" and "love bird" will also buy "pentet" at 16%, 3) if you buy "kacer" and "magpie" then you will buy "pentet" at 14%, 4) if you buy "anis" you will buy "pentet" of 11% with a confidence level of 76%, 81%, 84%, 77%, respectively. So, there are 5 main items that play a strong role in the rule that must be considered. Sellers can use the resulting item relationship patterns as consideration in managing inventory and structuring the items sold.

Keywords: Apriori, Inventory, Association Rules, Data Mining

I. INTRODUCTION

The readiness of product inventory to be marketed is very important. Sometimes there is a vacancy of a product related to another product that is desired by the buyer, which can make the buyer disappointed and then cancel to buy the product that was previously planned to be purchased at once. In fact, eventually move to another seller who is more complete. As a result, the seller will experience a decrease in the number of sales which of course has an impact on a decrease in revenue. The amount of availability of goods stored in the warehouse is very important because management can determine the minimum and maximum stock [1]. This means that there are still many sellers who do not or do not know the habits of customers and buyers when making purchases.

The availability of product items related to other product items that customers often buy simultaneously can be studied through sales data that has occurred. Sales transaction data is one of the things that can be used to make business decisions [2] but it can also be searched for or mined a very valuable knowledge. Constraints that arise in MSMEs are sometimes sales data cannot be studied directly, because apart from the large amount of data, but also because the data is still non-digital. The strategy for determining inventory stock, as conveyed by Kurniawati, et al [3], that information about the rules of association of strong linkages between products or itemset can be a recommendation in stocking goods and their placement. Where the rules of this association are obtained through processing a database of sales transactions that are already running from products in the form of mining.

Many methods and algorithms in data mining (data mining) can be applied, one of which is an association technique that seeks to analyze the pattern of linkages/connections between one item and another in a transaction event. An association algorithm can be used to determine the relationship of an item purchased with other items by looking at the behavior that is usually done by consumers [4]. Many problems can be solved by data mining techniques, such as the a priori algorithm used to determine the relationship between products through stored transaction data [5].

This study aims to implement the association method data mining technique using the a priori algorithm on sales transaction data in an MSME 'X' for 5 years and find association patterns between items in the sales transaction data. From the research, it is hoped that by knowing the pattern of relationships between items in sales transactions, in the future it is hoped that sellers can prepare product availability and placement better. Furthermore, with the availability of good products, buyers will be much easier and more comfortable in shopping so that it will increase purchases and more profits are received by sellers.

II. LITERATURE REVIEW

According to Khotimah, et al [6] in his research on the use of association rule mining to find the pattern of MSME owners with the fp-growth algorithm, it was stated that by using the attributes of gender and level of education, the results obtained were 13 rules with the lowest support value of 0.025 indicating that there were 4 groups. Where the first group has 1 rule if it is SMA, then female, the second group has 1 rule if it is female, then SMA, the third group has 5 rules and the fourth group has low support and confidence values with 6 rules.

Mining association rules as done by Azwar Anas [7] to choose a program of supporting activities for students at STIE-GK Muara Buliah. Research shows that the pattern of support activity programs that often appear simultaneously from student choices is computer and English programs with a support level of 22% and confidence 34%, especially in the combination of 2 items. In the combination of 3 items, students more often choose UKM, English and Computer programs with a lower support level of 18% but with a confidence level of up to 100%.

The application of data mining can also be used to determine the layout or layout of the products being sold [4]. The layout of the goods sold, in this case clothing products with certain rules based on the analysis of daily sales transactions in the store, can increase the ease and speed of finding products and their combinations. The combination pattern obtained in this research is if someone buys KA0701 and KK-201 and SP-2001, the buyer will buy ST-651. The combination of items in this study reached 4 items to get the best-selling associative rule output of sales transactions.

The use of apriori algorithms in various studies to determine consumer behavior patterns as well as sales strategies and inventory control [8] [9] [10] [11]. In addition, data mining techniques can also be used as decision support in carrying out inventory management [12] and sales information systems using shopping cart analysis [13]. A research survey that applies parallel apriori algorithms that can be carried out on a big data platform [14]

Data Mining: According to Han & Kamber [15], briefly data mining can be interpreted as extracting or digging knowledge from large amounts of data. Data mining is an activity that includes the collection and use of historical data to find regularities, patterns, or relationships in large data sets. The output of this data mining can be used to improve future decision making. So, the term pattern recognition is now rarely used because it is part of the data mining [16]. The KDD (Knowledge Discovery in Databases) process consists of a series of transformation steps, from the preprocessing data process and the postprocessing data process from the extracted data. Input data can be stored in a variety of formats (flat file, spreadsheet, or relational table) and may reside in a centralized data repository or be distributed across multiple addresses. The purpose of the data preprocessing process is to convert the raw input data into a suitable format for further analysis. The steps taken include correcting dirty or duplicate data and selecting records and features that are relevant to the next data management process. Because of the many ways data can be collected and stored, the data processing process may be tiring and time-consuming in the overall knowledge discovery process [17]. The stages of KDD as a series of processes, data mining can be divided into several stages as follows:

1. Data, prepare the data to be processed by separating the data from the operational data.
2. Selection, selecting the required data through target data collection activities, determining attributes, determining samples and storing data or files
3. Pre-processing/Cleaning, from the data that has been selected then further cleaning is carried out which includes removing duplication, repairing inconsistent data, correcting other data errors.
4. Transformation, provides a format for the data to be mined according to the method and algorithm
5. Data Mining, as the main process in extracting and seeking useful knowledge and information.
6. Interpretation Evaluation, mining results will be displayed and interpreted in a form that is easily understood by users.
7. Knowledge, knowledge or information obtained is the main goal.

Technical Association Rules: Association method is a data mining technique to find a relationship between data [18], which is hidden in a data set [19] that shows conditions in a data set, where several attribute values will appear simultaneously. Association rule is one of the most numerous and important data mining algorithms in Data Mining [20].

Apriori Algorithm: The a priori algorithm is a standard algorithm presented by Agrawal and Srikant in 1994. This algorithm is intended to determine Frequent itemsets in Boolean association rules. The A-priori algorithm belongs to the Association Rules group in data mining [17]. Rules that show associations between a few fields/attributes are often known as affinity analysis or shopping basket analysis. Association analysis or association rule mining is a data mining technique to obtain the rules of a combination of items and is useful for finding interesting hidden relationships in large data sets [21] and this method can determine consumer buying behavior patterns [22]. There are two main benchmarks in this association technique, namely: support and confidence. Support (value of support / support) is the percentage of the combination of these items in the database, while confidence (value of confidence / certainty) is a measure of the strength of the relationship between items in the association rules.

a) High frequency pattern
At the initial stage, the minimum support requirement is determined which will be used as a limitation in the data frequency selection process.

$$\boldsymbol{Support}(\boldsymbol{A}) = \frac{\textbf{Jumlah Transaksi Berisi A}}{\textbf{Jumlah Seluruh Transaksi}} \quad \textbf{1)}$$

That is, the support value of an itemset A is the number of transactions containing itemset A divided by the total number of transactions in the database. The formula above is used for 1 itemset, while to find the support value for 2 itemset, the following formula can be given:

$$\boldsymbol{Support}(\boldsymbol{A},\boldsymbol{B}) = \boldsymbol{Support}(\boldsymbol{A} \cap \boldsymbol{B}) = \frac{\textbf{Jumlah Transaksi Berisi A dan B}}{\textbf{Jumlah Seluruh Transaksi}} \quad \textbf{2)}$$

b) From the association rules that state the strength of the itemset combination relationship in the next transaction, confidence is determined. To determine the association rules that are formed, a minimum of itemset must have two candidates A and B. In the rules formed, the associative law A -> B does not apply to B -> A. To determine the rule A->B used the formula:

$$\boldsymbol{Confidence}(\boldsymbol{A},\boldsymbol{B}) = \boldsymbol{Support}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\textbf{Jumlah Transaksi Berisi A dan B}}{\textbf{Jumlah Transaksi Berisi A}} \quad \textbf{3)}$$

R (R Language): According to Drs. I. M. Tirta, Dip.Sc, M.Sc., Ph.D. in his book "R Statistics Program Guide" (electronic version) states that R is an open-source software package that can be obtained through free downloads on the website http://www.r-project.org/ as well as through http://cran.rproject.org/. Historically, R was included in the S language family, where the two main programs were written in the S language. S-Plus was developed commercially with R built through an open-source concept. Both at that time were distinguished by the side of the interface where the S-Plus was displayed with complete menus and a GUI interface. While R is displayed in CLI (Command Line Interface) mode as the main display. With this open-source model, R is growing quite fast with lots of package donations from contributors. R's internal ability to analyze statistical data has resulted in R being categorized as a data processing package (statistical package). Furthermore, many libraries or libraries were developed from contributors for certain analyzes so that R has almost parallel capabilities with S-Plus. However, in terms of ease of use, it is still not equal.http://tirtamade.blog.unej.ac.id/wp-content/uploads/2014/10/RcomManSld.pdf

III. METHODS AND MATERIAL

The research method used in this study is descriptive analysis with a quantitative approach, concentrating on analyzing the data obtained, to get a clear picture of a condition from the data through presenting, collecting, and analyzing data and using new information obtained to answer the problem being researched. The stages of the research can be seen in Figure 1.

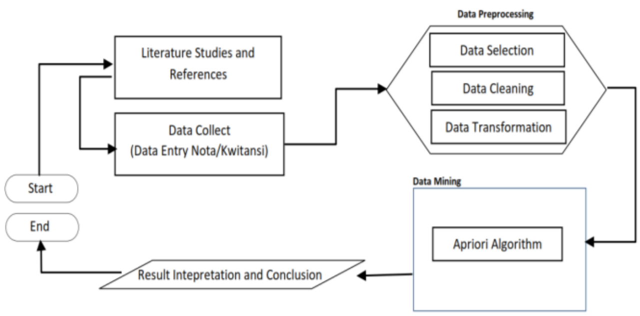

Figure 1. Stages of research carried out

Data Analysis Method: While the analysis of the data in the application of data mining uses the Knowledge Discovery in Databases (KDD) stage. The use of association techniques with apriori algorithms is assisted by the R language in RStudio and Microsoft Excel. Data collection was carried out by digitizing data from receipt form to computerized form using Microsoft Excel and obtained data of 1063 transaction numbers with a total data of 2717 records. The data input into Excel includes transaction number, transaction date, buyer's name/description, item name, purchase amount. The following is an example of sales data as shown in table 1.

TABLE 1
EXAMPL OF SALES DATA

No	Original Item Name	New Item Name
1	love bird	A
2	poksai	B
3	pentet	C
4	murai	D
5	anis	E
6	kacer	F
7	pleci	G
8	kenari	H
9	perkutut	I
10	bekisar	J
11	merpati	K
12	ternak love bird	L

Cleaning, from the data that has been obtained then further cleaning is carried out which includes the removal of several duplicate data, correction of inconsistent data. Correction of data errors can be in the form of removing leading and trailing spaces from item name entries, deleting empty data, and changing all data entries to lowercase. Giving format to the data to be mined according to the method and algorithm is a transformation process. From the data that has gone through preprocessing, in the name of the item there are many variations of the name of the item which is the same item. Therefore, the name of the goods is simplified to the name of the raw or original goods of the type. Examples of goods name magpie, magpie 1, magpie 2, ordinary magpie, blue magpie, light blue magpie, dark blue magpie, thin shirt magpie, red magpie, magpie no. From the process of standardizing the name of the goods, 12 kinds of goods were finally obtained, which then further facilitated the data processing in association pattern mining, the name of the goods was further transformed into 1 capital letter, so that it became as shown in table 2.

TABLE 2
LIST OF ITEM NAMES

No Not e	Date	Buyer	Name of Goods	Amoun t
1	7-Jan-14	agung	love bird	5
1	7-Jan-14	agung	poksai	5
1	7-Jan-14	agung	pentet	5
1	7-Jan-14	agung	murai	10
2	15-Jan-14	non-subscribe2	love bird	5
2	15-Jan-14	non-subscribe2	anis	5
2	15-Jan-14	non-subscribe2	poksai	5
2	15-Jan-14	non-subscribe2	kacer	5
2	15-Jan-14	non-subscribe2	pentet	5
2	15-Jan-14	non-subscribe2	murai	5
2	15-Jan-14	non-subscribe2	pentet	5

As the main process in extracting and seeking useful knowledge and information, the algorithm implementation process is a priori. The a priori algorithm procedure can be presented in the flowchart in Figure 2.

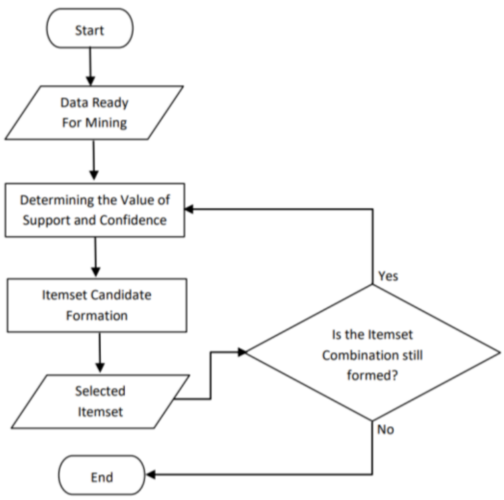

Figure 2. Apriori algorithm procedure

The presentation of rules or association rules resulting from the mining process with association techniques is interpreted into a model that is easily understood by users. Association rules are displayed in the form if then (IF THEN)

Implementation of the Apriori Algorithm Manually: In this study, the implementation of a data mining algorithm that uses Microsoft Excel and R language tools in the RStudio application is determined by a minimum support (minsupp) of 10% and a minimum confidence (minconf) of 75%. The implementation was carried out in 2 experimental models, where the first experiment was implemented manually with the help of Microsoft Excel and RStudio using the first 100 transaction data and the second experiment for all transaction data using RStudio. In the manual implementation process, the first 100 transaction data are converted into tabular data format (data tabulation) to calculate the frequency of items that meet the requirements. The presentation of data in tabular format is shown in table 3., to calculate the number of items in each transaction and the number of transactions for each item. Where the value of one indicates a purchase transaction and 0 there is no purchase. For example, in transaction number 3, the items purchased are {C, D, E}.

TABLE 3
EXAMPLE OF TABULAR DATA PRESENTATION

No Note	A	B	C	D	E	F	G	H	I
1	1	1	1	1	0	0	0	0	0
2	1	1	1	1	1	1	0	0	0
3	0	0	1	1	1	0	0	0	0
4	0	0	1	1	1	0	0	0	0
5	0	0	1	1	1	0	0	0	0
6	0	0	1	1	1	0	0	0	0
7	0	0	1	1	1	0	0	0	0

Formation of n-Itemset: The process of forming the n-itemset starts from selecting items that meet the predetermined minsupp=10% calculated using formula 1) whose results can be seen in table 4. Of the 9 items, there are 6 items that meet the minsupp.

TABLE 4
DATA PROCESSING RESULTS ITEMSET

No	Name of Goods	Number of Transactions	Support
1	A	37	37%
2	B	21	21%
3	C	80	80%
4	D	61	61%
5	E	50	50%
6	F	27	27%

Furthermore, after a combination of 1 item is formed and the item that meets the minimum supply is selected, it is continued by compiling a combination of 2 items until the most n-itemset.

Combination of 2 itemset up to the most n-itemset: For the formation of 2-itemset and so on until the most n-itemset combinations that meet the minsupp are obtained, they are presented in the tables below. The 2-itemset combination is shown in Table 5.

TABLE 5:
2-ITEM SET COMBINATION

No	2-Itemset	Amount	Support
1	AB	13	13%
2	AC	29	29%
3	AD	21	21%
x	xx	xx	xxx
x	xx	xx	xxx
x	xx	xx	xxx
14	DF	14	14%
15	EF	12	12%

Combination of 3 itemset: In the process of forming a combination of 3 itemset which is also called C3, based on the combination of 2 itemsets that meet the minsupp, 34 combinations are formed 3-itemset. From the process of forming a combination of 3 itemsets that meet the minsuoo results obtained 21 combinations as shown in Table 6.

TABLE 6
3-ITEM SET COMBINATION

No	3-Itemset	Amount	Support
1	ABC	12	12%
2	ABD	11	11%
3	ABE	10	10%
xx	xxx	xx	xxx
xx	xxx	xx	xxx
20	DFC	14	14%
21	CEF	10	10%

Next from C3 is continued for the 4-itemset or C4 combination. Of the 6 items combined with 4-itemset, 15 combinations are obtained, and 4 itemset combinations are obtained that meet the minimum supply as shown in Table 7.

TABLE 7
4-ITEM SET COMBINATION

No	4-Itemset	Amount	Support
1	ABCD	10	10%
2	ABCE	10	10%
3	ACDE	14	14%
4	BCDE	12	12%

After getting the 4-Itemset combination, then proceed to the 5-Itemset combination. From the day of calculation of transaction data that meets the 5-itemset, only 9% is obtained for the support value, this means that the 5-itemset combination does not meet the minsupp=10% limit value. Thus, it means that the maximum combination that meets the 10% support in the first 100 transaction data is the 4-itemset combination.

Henceforth, from each combination of itemset from 2 to 4 that meets the minsupp, the confidence value can be calculated as shown in table 8. From the calculation 21 combinations that meet the minsupp and mincof.

TABLE 8
COMBINATION OF ITEMS MEETS
MISUPP 10% AND MINCONF 75%

No	2-Itemset	Amount	Support	Confidece
1	AC	29	29%	78%
2	BC	17	17%	81%
3	BD	16	16%	76%
4	DC	57	57%	93%
xx	xxx	xx	xxx	xxxx
xx	xxx	xx	xxx	xxxx
20	ABCE	10	10%	83%
21	BCDE	12	12%	86%

Then if it is translated into the actual name of the goods then the association rules above are as shown in table 9.

TABLE 9

ASSOCIATION RULES WITH ITEM NAMES

No	Association Rules	Amount	Support	Confidece
1	If buy love birds --> buy pentet	29	29%	78%
2	If buy poksai --> buy pentet	17	17%	81%
3	If buy poksai --> buy murai	16	16%	76%
4	If buy murai --> buy pentet	57	57%	93%
xx	xxxxxxxxxxxxxxxxxxxxxxxx	xx	xxx	xxxx
xx	xxxxxxxxxxxxxxxxxxxxxxxx	xx	xxx	xxxx
20	If buy love bird and poksai and pentet --> buy anis	10	10%	83%
21	If buy poksai and pentet and murai --> buy anis	12	12%	86%

Apriori Algorithm Implementation Using Rstudio: The implementation process with RStudio starts with installing and loading the required packages with the commands Install.packages(“packagename”) and library(packagename). Where each package has a function readxl# to read/import datasets stored in Microsoft excel format, arules# forming frequent itemset in data mining association, arulesViz# visualization of rules and frequent itemset in data mining association, grid# make a plot at a certain location in the plotting area, matrix# create large vectors more effectively.

Data reading and converting into transactional data

1) *sales<read_excel("D:/Script/Apriori100.xlsx",col_names=FALSE) #* read and enter the file 'apriori100.xlsx' into RStudio.
2) *transactiondata100 <- ddply (sales, c("NoNota"), function(df1)paste(df1$Name of Goods,collapse = ","))#* converted into transactional data.
3) *write.csv (transactiondata,"D:/Script/transactiondata100NB.csv", quote = FALSE, row.names = FALSE)#* save transactional data to drive
4) *storetransaction<- read. transactions ("D:/Script/transaction data100NB.csv ", format = 'basket', sep=',') #* read transactional data from drive

Establishment of rules

5) *data_rules=apriori (shop transactions, parameters) = list (support=0.10, confidence=0.75, minlen=2))*
6) *is. redundant(data_rules)*
7) *inspect (data_rules [is. redundant(data_rules)])*
8) *inspect (data_rules [! is. redundant(data_rules)])*

IV. RESULTS AND DISCUSSION

RCode Execution Results: The results of the source code execution (rule forming) are obtained as follows:

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen target ext

0.75 0.1 1 none FALSE TRUE 5 0.1 210 rules TRUE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 10

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[111 item(s), 101 transaction(s)] done [0.00s].

sorting and recoding items ... [6 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 3 4 done [0.00s].

writing ... [25 rule(s)] done [0.00s].

creating S4 object ... done [0.00s].

Figure 3. Results of the apriori function execution process

In Figure 3. it can be explained, from the transaction data that the source code execution has been carried out based on the minimum support and confidence parameters, 25 rules were obtained. Furthermore, the process was also carried out to eliminate duplicate/redundant rules, so that in the end a total of 20 rules were obtained.

lhs rhs support confidence coverage lift count
[1] {D} => {C} 0.5643564 0.9344262 0.6039604 1.17 9713 57 [2] {E} => {C} 0.4455446 0.9000000 0.495 0495 1.136250 45 [3] {D,E} => {C} 0.3465347 1.0000 000 0.3465347 1.262500 35
……………………………………………………………
…………………..
[19] {B,C,D} => {E} 0.1188119 0.8571429 0.1386139 1. 731429 12
[20] {A,B} => {D} 0.1089109 0.8461538 0.1287129 1. 401009 11

Figure 4. Results of the non-redudant apriori rule

From Figure 4 with a minimum support of 0.10 and a confidence of 0.75, 20 rules are generated which are sorted by the largest support.

[1] {D} => {C} 0.5643564 0.9344262 0.6039604 1.17 9713 57

From the biggest rule, if you buy D, you will buy C. That is, if someone buys magpie, they will buy pentet with support 0.5643564 and confidence 0.9344262 with a lift of 1.179713 with a total of 57 transactions from the total number of transactions, namely 100 transactions. Here if the data is in the form of the original item name.

>inspect(data_rules[!is.redundant(data_rules)])
lhs rhs support confidence coverage lift count
[1] {murai} => {pentet} 0.5643564 0.9344262 0.60396 04 1.179713 57
[2] {anis} => {pentet} 0.4455446 0.9000000 0.4950495 1.136250 45
[3] {anis,murai} => {pentet} 0.3465347 1.000000 0 0.3465347 1.262500 35
……………………………………………………………
…………………… *[19] {murai,pentet,poksai} => {anis} 0.1188119 0.8571429 0.1386139 1.731429 12*
[20] {love bird,poksai} => {murai} 0.1089109 0.8461 538 0.1287129 1.401009 11

While all sales data used in this study were 1063 transactions, the results of association rules or 4 rules were obtained.

Apriori
Parameter specification:
confidence minval smax arem aval originalSupport maxtime su
0.75 0.1 1 none FALSE TRUE 5 0.1 2 10 rules TRU

Algorithmic control:
filter tree heap memopt load sort verbose
0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 10
set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[1077 item(s), 1064 transaction(s)] done [0.00
sorting and recoding items ... [6 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 done [0.00s].
writing ... [4 rule(s)] done [0.00s].
creating S4 object ... done [0.00s].

lhs rhs support confidence coverage lift count
[1] {kacer,murai} => {pentet} 0.1381579 0.8352273 0.1654135 1
[2] {love bird,murai} => {pentet} 0.1607143 0.8066038 0.199248
[3] {anis} => {pentet} 0.1137218 0.7707006 0.1475564 1.346511
[4] {kacer,love bird} => {pentet} 0.1701128 0.7605042 0.223684

The rule results obtained when presented in the form of a diagram using the plot function are as shown in Figure 5.

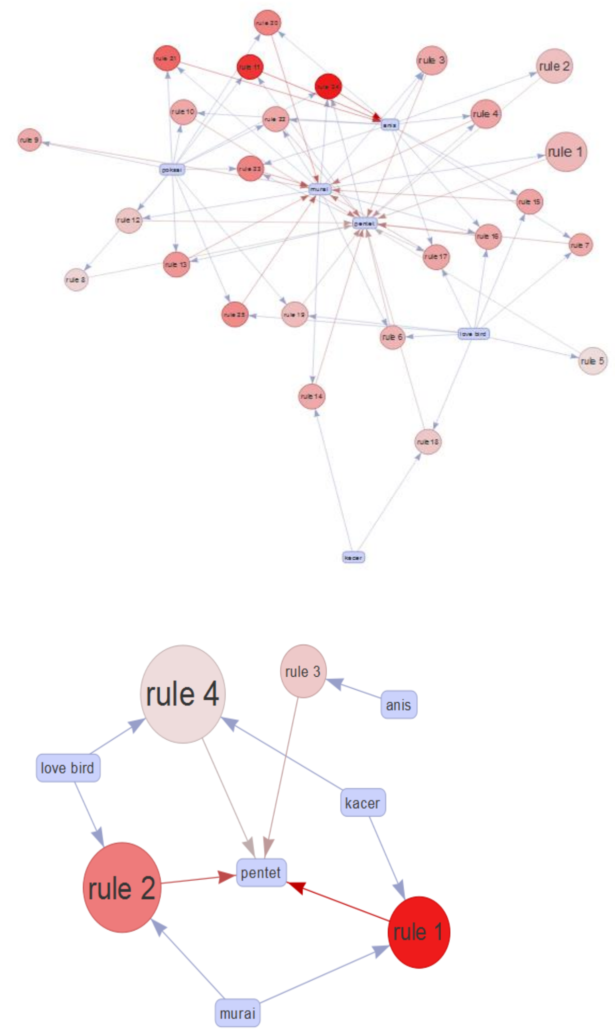

Figure 5. Plot of the rules obtained from the first 100 transactions and all transactions

Sales data association results analysis: The strength level of the association rule can be calculated by multiplying the support value by its confidence. So that the calculation results can be seen in table 10. While the calculation of accuracy is based on the amount of support from each rule obtained.

TABLE 10

CALCULATION OF THE STRENGTH OF APRIORI ASSOCIATION RULES

lhs	rhs	support	confidence	Strength
[1] {kacer,murai} =>	{pentet}	0.1381579	0.835227 3	0.115393
[2] {love bird,murai} =>	{pentet}	0.1607143	0.8066038	0.129633
[3] {anis} =>	{pentet}	0.1137218	0.7707006	0.087645
[4] {kacer,love bird} =>	{pentet}	0.1701128	0.7605042	0.129371
Apriori Algorithm Accuracy		0.5827068		
The Power of the Apriori Algorithm				0.115511

So, in general, if the four rules are arranged into an association rule, then:

No	Association Rules	support	confidence	Strength
1	If buy kacer and love bird, then will buy pentet	17 %	76%	13%
2	If buy murai dan love bird, then will buy pentet	16 %	81 %	13%
3	If buy kacer dan murai, then will buy pentet	14 %	84 %	12%
4	If buy anis, then will buy pentet	11 %	77 %	9%

From the association rules of the algorithm results, there are 5 items involved in it, namely kacer, magpie, pentet, love bird, anis. Meanwhile, if based on the number of sales of each item in the dataset used kacer (397), magpie (385), pentet (609), love bird (666), anis (157).

With this result, it is hoped that it can be a consideration for MSME owners "X" that the inventory of each item of goods can be adjusted to the level of linkage between items of goods and the amount of support in the rules or association rules that are formed.

V. CONCLUSION

Based on the results of research activities carried out, conclusions can be drawn as follows:

By using a minimum support of 0.10 and a minimum confidence of 0.75, the obtained rules or association rules using the a priori algorithm for the first 100 transactions are 20 rules and 4 rules using the R language in RStudio for all data used in the study.

The pattern of the relationship between the items in the sales data which is the pattern of customer shopping habits, it is found that if someone buys kacer and love bird, they will buy pentet as much as 17% support from 1063 transactions with a confidence level of 76%, then if they buy magpie and love bird then I will also buy 16% pentet with 81% confidence level, if I buy kacer and magpie, I will buy 14% pentet with 84% confidence level and finally if I buy anis, I will buy 11% pentet with 77% confidence level. So, there are 5 main items that play a role in this rule, namely love bird (666), pentet (609), magpie (385), kacer (397) and anis (157). Thus, it is expected that the seller can use the resulting item relationship pattern as consideration in managing inventory and structuring the goods sold.

VI. REFERENCES

[1]. M. S. I. R. S. H. M. S. Adnyana, "Market Basket Analysis For Procurement Of Food Stock Using Apriori Algorithm And Economic Order Quantity," International Journal of Engineering and Emerging Technology, p. 149, 2020.

[2]. R. Maman Novian, "Analysis of the Application of Customer Purchase Mining Data on Paint Sales Using Apriori Algorithm (Case Study:PT Indowarna Cemerlang Indonesia)," Bit-Tech, vol. 2, no. 3, p. 131, 2020.

[3]. A. E. K. B. D. Laela Kurniawati, "Implementasi Algoritma Apriori untuk Menentukan Persediaan Spare Part Compressor," CESS(Jurnal Of Computer Engineering System and Acience) Vol. 4 No. 1 Januari 2019, pp. 6-10, 2019.

[4]. L. S. Z. L. W. L. N. C. R. H. Li Zhou, "Study on a storage location strategy based on clustering and association algorithms," Soft Computing , vol. 24, no. 8, p. 5499–5516, 2020.

[5]. P. M. H. Nurayni Sinabang, "Application of Data Mining for Sales Strategy at Ria Busana Using the A priori Algorithm," Login : Jurnal Teknologi Komputer, vol. 14, no. 2, pp. 121-127, 2020.

[6]. D. E. S. A. L. A. S. Tutik Khotimah, "ASSOCIATION RULE MINING UNTUK MENEMUKAN POLA PEMILIK UMKM," Prosiding SNATIF, pp. 517-522, 2018.

[7]. A. Anas, "Penggalian Kaidah asosiasi dalam Memilih Program Kegiatan Pendukung Mahasiswa STIE-GK Muara Bulian," MEDIASISFO Vol. 11, No.1, April 2017, pp. 709-722, 2017.

[8]. C. M. A. A. F. D. J. Ahmad Heru Mujianto, "Consumer Customs Analysis Using the Association Rule and Apriori Algorithm for Determining Sales Strategies in Retail Central," in ICENIS, Semarang, 2019.

[9]. A. F. T. Y. P. Sunardi, "Market Basket Analysis to Identify Stock Handling Patterns and Item Arrangement Patterns Using Apriori Algorithms," Jurnal Ilmu Komputer dan Informatika, vol. 6, no. 1, pp. 33-41, 2020.

[10]. A. S. ,. A. B. M. I. F. Samuel, "SALES LEVEL ANALYSIS USING THE ASSOCIATION METHOD WITH THE APRIORI ALGORITHM," JURNAL RISET INFORMATIKA, vol. 4, no. 4, pp. 331-340, 2022.

[11]. S. S. M. A. S. L. R. ,. H. T. S. B. S. Panjaitan, "Implementation of Apriori Algorithm for Analysis of Consumer Purchase Patterns," in The International Conference on Computer Science

and Applied Mathematic, Parapat, Indonesia, 2019.

[12]. B. H. N. Vivek Ware, "Decision Support System for Inventory Management using Data Mining Techniques," International Journal of Engineering and Advanced Technology (IJEAT), vol. 3, no. 6, pp. 164-168, 2014.

[13]. G. S. B. D. H. S. R. D. Alexander Setiawan, "Data Mining Applications for Sales Information System Using Market Basket Analysis on Stationery Company," in 2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT), Denpasar, Indonesia, 2018.

[14]. E. S. A. A. A. B. A.L.SAYETH SAABITH, "PARALLEL IMPLEMENTATION OF APRIORI ALGORITHMS ON THE HADOOP-MAPREDUCE PLATFORM- AN EVALUATION OF LITERATURE," Journal of Theoretical and Applied Information Technology , vol. 85, no. 3, pp. 321-351, 2018.

[15]. M. K. Jiawei Han, Data Mining: Concepts and Techniques, Second Edition, San Francisco: Morgan Kaufmann, 2006.

[16]. B. Santoso, Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis, Yogyakarta: Graha Ilmu, 2009.

[17]. E. T. L. Kusrini, Algoritma Data Mining, Yogyakarta: Andi Publiser, 2009.

[18]. A. S. Hena Lisnawati, "Data Mining with Associated Methods to Predict Consumer Purchasing Patterns," I.J. Modern Education and Computer Science, vol. 5, pp. 16-28, 2020.

[19]. A. D. S. H. S. S. J. W. Michael S Packianather, "Data mining techniques applied to a manufacturing SME," in 10th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '16, 2017.

[20]. D. M. P. Dr. M. Dhanabhakyam, "A Survey on Data Mining Algorithm for Market Basket Analysis," Global Journal of Computer Science and Technology, vol. 11, no. 11, 2011.

[21]. M. Kang, "Market Basket Analysis: Identify the changingtrens of market data," Procedia Computer Science, pp. 78 - 85, 2016.

[22]. E. Irfiani, "Application of Apriori Algorithms to Determine Associations in Outdoor Sports Equipment Stores," SINKRON : Journal Publications & Informatics Engineering Research, vol. 3, no. 2, pp. 218-222, 2019.

[23]. D. A. Nurdin, "Penerapan Data Mining untuk Menganalisis Penjualan Barang dengan Menggunakan Metode Arpiori Pada Supermarket Sejahtera LhokSeumawe," Techsi Vol. 6 No.1, April 2015, pp. 133-155, 2015.

[24]. H. F. A. S. Sanjani, "Implementasi Data Mining Penjualan Produk Pakaian Dengan Algoritma Apriori," IJAI(Indonesian Journal of Applied Informatics) Vol. 4 No. 1 Tahun 2019, pp. 23-29, 2019.

[25]. R. N. Arifin, "Implementasi Algoritma Frequent Pattern Growth(FP-Growth) Menentukan Asosiasi antar produk(Studi Kasus : Nadiamart)," Jurnal Teknik ITS, pp. 68-76, 2015.

[26]. C. D. L. Daniel T. Larose, Discovering Knowledge in Data: An Introduction to Data Mining, 2nd Edition, Jhon Wiley & Sons Inc, 2005.

[27]. E. Buulolo, Data Mining Untuk Perguruan Tinggi, Yogyakarta: Deepublish, 2020.

Cite this article as :

Arief Jananto, Yohanes Suhari, Rara Sriartati Redjeki, Bambang Sudiyatno, "Pattern Recognition of Customer Spending Habits Using Apriori Algorithms in DataMining as an Inventory Strategy", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 6, pp. 104-115, November-December 2022. Available at doi : https://doi.org/10.32628/CSEIT22868

Journal URL : https://ijsrcseit.com/CSEIT22868