Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

Association Rule to Increase Sales Using the Apriori Algorithm Method

Ermanto[1*], Abdul Halim Anshor[2], Asep Arwan Sulaeman[3], Sri Winarni[4]
[1,2,3,4]Informatics Engineering Study Program, Faculty of Engineering, Pelita Bangsa University, Indonesia
[1]ermanto@pelitabangsa.ac.id, [2]abdul.halim@pelitabangsa.ac.id, [3]asep.arwan@pelitabangsa.ac.id, [4]s.winar@gmail.com

open access

***Corresponding Author**

Article History:
Submitted: 27-06-2024
Accepted: 28-06-2024
Published: 08-07-2024

Keywords:
Data Mining, Association Rule, apriori algorithm, sales

ABSTRACT

The Apriori algorithm is a data mining technique used to find relationship patterns between items in a transaction dataset. In this context, the Apriori algorithm will be used to identify products that are often purchased simultaneously by customers. By understanding these purchasing patterns, companies can design more effective marketing strategies, such as strategic product placement, bundling package offers, and special promotions. This research involves several stages, starting from collecting sales transaction data, data preprocessing, applying the Apriori algorithm, to interpreting the results. The transaction data used is taken from the sales database of a retail store during a certain period. After the data is processed, the Apriori algorithm is applied to identify frequent itemsets and form association rules. The results of this research show that there are several significant purchasing patterns, such as a combination of product A and product B which are often purchased together. By applying data mining using the a priori algorithm method, you can find out which products sell the most. From the results of manual calculations it was found that consumers who bought RB 1060 would buy RB 1099 with 81% confidence, whereas using WEKA it was found that consumers who bought RB 1060 would buy RB 1099 with a confidence value of 82%.

INTRODUCTION

In order for a company's business to remain in demand by customers, the company needs to know the interests and needs of customers. This is a marketing strategy for a company so that the products being marketed remain in demand by customers (Wahyuni et al., 2021a). Knowing the interests and needs of customers is the key to success for a company so that the production it undertakes can be accepted (Ratih Wahyuningrum, 2018). To develop this business, it is necessary to predict consumer interest in the product. This can be done using data mining (Tualeka et al., 2021). These predictions aim to enable companies to plan marketing strategies, to manage their business so that customers remain interested (Wahyuni et al., 2021b). To predict the interest of these customers, you can use predictive techniques (Husdi & Dalai, 2023). Data mining is a technology for a company to make decisions (Saefudin & DN, 2019). The use of data mining to determine decisions has now been widely used. The application of data mining has been widely applied for the benefit of various fields (Priyanka Aggarwal, 2019).

Apart from making decisions, data mining can be used to dig into large databases (Putri Mai Sarah Tarigan et al., 2022). Association Rules are data mining technique rules for finding high frequency patterns between groups of items (Bimantoro & Wardhani, 2020). The process of finding these patterns can be done through analytical and statistical techniques (Kharis & Zili, 2022). This technique can be used to analyze customer shopping habits and to identify purchases made (Akbar et al., 2024). A classic example is shopping basket analysis, which aims to find products that are often purchased together.

With daily sales activities, PT NOK Indonesia is a manufacturing company that produces Oil Seals, O-Rings, Gaskets, Ultra Bushes, Chain-tensioners and Dampers which are automotive spare parts. Transaction activity will produce piles of data that get bigger and bigger, which can cause new problems. This activity requires storage media, so that transactions that require large storage media do not end up in a pile of trash (Erwin Stanley Harianja, 2020). As technology advances increasingly rapidly, the ability to collect, analyze and process data in databases also develops, so a way is needed to sort and select important data from databases. In the competitive business world, developers are required to find effective strategies that can increase sales of goods. Research at PT. NOK Indonesia is a quantitative research. whose specifications are systematic, planned and clearly structured from the beginning until the design is made. Research methods based on the philosophy of positivism are used to research certain populations or samples, collecting data using research instruments, quantitative/statistical data analysis, with the aim of testing predetermined hypotheses. The stages used use the CRISP-DM method, implemented by Weka.

LITERATURE REVIEW

Ultrabush is a motorbike spare part which functions to reduce vibrations caused by the engine, which can spread to the body due to contact between the body and the engine (Ariyani, 2021). Companies need analytical techniques to

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

find out consumer habits in purchasing products (Firdaus et al., 2021). This technique can use an a priori algorithm to test combinations in order to meet the minimum support and confidence parameters which are threshold values given by the user (Wijaya et al., 2024). The goal is to apply data mining techniques to increase product sales ultra bush (Sutrisno, 2020). The application of data mining using association rules with the Apriori algorithm method has proven effective in increasing sales of ultra bush products (Sibarani, 2020).

Application of Data Mining Using Association Rules to Support Marketing Strategies for Prospective New Students (Case Study of IBI Darmajaya)" is research by Nurjoko and Abdi Darmawan (Ratih Wahyuningrum, 2018). By using an a priori algorithm, the values displayed are support and confidence. For each pattern that is formed (Syaipul Amri, 2018). Association rules can help identify customer purchasing patterns and offer valuable insight into product combinations that are frequently purchased together. The main metrics used can measure the reliability of the resulting association rules. From trials on prospective student data and student data using the Association rule model, several rules were produced with a maximum confidence level of 81.1%. For information sources originating from newspaper media, the maximum trust level value is 77.8%.

Analyzing and implementing data mining using a priori algorithms is research that aims to increase food sales. This research was carried out in three stages, namely: data collection, data preprocessing and finally application of the Apriori algorithm (Mulya et al., 2019). Meanwhile, Hutasuhut in his research used an iterative approach to increase sales (Hutasuhut et al., 2019). Research to increase sales carried out by Kamil Erwansyah, et al (Erwansyah et al., 2021) is research that processes sales data. The goal is to find out products that can be recommended to consumers. Using the association method, so that sales in the store can increase. Using sales data in 3 months, and choosing 2 products. By using the association method with an a priori algorithm, a product shopping recommendation pattern was obtained at the Avis Mobile store.

METHOD

Type of Research

The type of research used in this research is quantitative research. Quantitative research methods are a type of research whose specifications are systematic, planned and clearly structured from the start until the creation of the research design. Quantitative research methods, as stated by Sugiyono (2014: 13), are: "Research methods based on the philosophy of positivism, used to research certain populations or samples, data collection using research instruments, quantitative/statistical data analysis, with the aim of testing hypothesis that has been established "

a. Data Type

The type of data used in this research is quantitative data. Quantitative data is a type of data that can be measured or calculated directly, in the form of information or explanations expressed in numbers or in the form of numbers (Sugiyono, 2014; 15). In this research, the quantitative data required are: Type of item and number of transactions

b. Data source

The data source used in this research is secondary data from Ultra Bush Sales Data in PDF form for 4 months (May-September 2018). According to Sugiyono, (2014:131) secondary data is a source of research data obtained by researchers indirectly through intermediary media (obtained and recorded by other parties). Secondary data is in the form of evidence, notes or historical reports compiled in published archives. The data used in this research was obtained from weekly reports published by the company which are available on the PT NOK Indonesia production wall.

c. Method of collecting data

The methods used to collect data and information required in preparing this thesis include:

1) Interview Method
   Ask questions that support the problem. Interviews were conducted with direct questions and answers to the parties at PT. Indonesian NOK regarding the required data. This interview was conducted with the Leader and permanent employees in the ultrabush production section of PT. NOK Indonesia.
2) Observation Method
   Observations are carried out by observing the system to create a model of the processes in the system that are currently running.
3) Literature Study Method
   It is a method of collecting data by studying literature and all other literature that is deemed necessary and supportive.

Cross-Industry Standard Process for Data Mining (CRISP-DM) Research Stages

The stages that the author carried out in this research used the CRISP-DM method with the following steps:

a. Business Understanding Phase

The aim of this research is to look for relationships between goods which are often purchased by customers simultaneously, to make it easier to manage stock of goods. In the initial stage, researchers looked for a dataset by knowing the weekly schedule data on the Bushing Production Line from March 2018 to September 2018.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

b. Data Understanding Phase

The data source used in this research is secondary data obtained from weekly data curing schedules from March 2018 to September 2018.

Table 1. Example of Sales Transaction Table Dataset

	A	B	C	D	E
1	**No.**	**Transaction date**	**Items**	**Customers**	
2	1	3/27/2018	RB 1095	SUZUKI	
3		3/27/2018	RB 1104	MITSUBA	
4		3/27/2018	RB 1060	YAMAHA	
5		3/27/2018	RB 1087	MITSUBA	
6		3/27/2018	RB 1099	YAMAHA	
7	2	3/28/2018	RB 1060	YAMAHA	
8		3/28/2018	RB 1087	MITSUBA	
9		3/28/2018	RB 1099	YAMAHA	
10	3	3/29/2018	RB 1087	MITSUBA	
11		3/29/2018	RB 1099	YAMAHA	
12		3/29/2018	RB 1104	SGS	
13		3/29/2018	RB 1060	YAMAHA	
14	4	04/01/2018	RB 1104	SGS	
15		04/01/2018	RB 1099	YAMAHA	
16		04/01/2018	RB 1111	MITSUBA	
17		04/01/2018	RB 1060	YAMAHA	
18	5	04/02/2018	RB 1104	MITSUBA	
19		04/02/2018	RB 1099	YAMAHA	
20		04/02/2018	RB 1060	YAMAHA	
21		04/02/2018	RB 1087	MITSUBA	
22		04/02/2018	RB 1095	SUZUKI	
23	6	04/03/2018	RB 1099	YAMAHA	
24		04/03/2018	RB 1111	MITSUBA	
25		04/03/2018	RB 1113	SGS	

Table 2. Explanation of variables

Variable	Information
No	Serial number
Transaction date	Date created
Items	Name of item purchased
Customers	Consumers who buy

c. Data Processing Phase

Of the existing transaction data, not all of it is processed. On In this research, the data used is only transactions that are more than 1 type of item, not how many items are purchased because What is sought is the relationship between goods.

d. Selection of processed data

The data processed is transactions involving more than 1 type of goods and used are the number and name of the item. Transaction table on (figure 1) there are variables "No", "Production Date", "Item", "Customer" but not all of these variables are used, only "No" and "Item" are needed.

e. Transaction Data Tabular Format

Tabular data format is a data format in the form of 1 and 0 or data format in binary form. In connection with the application used in testing, it is an application that uses a Microsoft Excel database with data in tabular data form, so the sales transaction data is converted into binary form. Process The conversion is that the slip number of the data to be tested is made horizontally downwards, while all types of items will be a vertical attribute, so that it forms like a table, based on real sales transaction data, the meeting point between the item name and number will be binary 1, while those that are not meeting points will be binary 0. Example of the results of the sales transaction data conversion process to data format in tabular form the data is like the following table:

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

Table 3. Example of Tabulator Format

NO	RB 1099	RB 1060	RB 1095	RB 1104	RB 1087	RB 1067	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
1	1	1	1	1	1	1	1	0	1	0	0	1
2	1	1	1	0	1	0	0	1	0	1	0	0
3	1	1	0	1	1	1	0	0	1	1	0	0
4	1	1	0	1	0	1	1	0	0	1	1	0
5	1	1	1	1	1	1	0	0	0	1	0	0
6	1	1	0	0	0	0	1	0	1	0	0	0
7	1	1	0	1	1	1	0	1	0	0	1	1
8	1	0	0	1	1	0	1	0	0	0	0	0
9	0	1	0	1	1	0	0	1	1	0	1	0
10	1	0	1	1	1	0	1	0	0	0	0	0
11	0	1	0	1	0	1	0	1	0	1	0	0
12	1	0	0	1	0	0	1	0	0	0	0	0
13	1	1	0	0	0	1	0	1	1	0	0	0
14	1	1	1	1	0	1	1	0	0	0	0	0
15	0	0	0	1	0	1	1	0	0	1	1	1
16	1	0	1	0	0	1	1	0	0	0	0	0
17	1	1	0	0	0	1	0	0	0	0	0	0
18	1	1	0	0	0	1	0	1	0	0	0	0
19	1	1	1	1	0	1	0	0	0	1	0	0

f. Modeling Phase

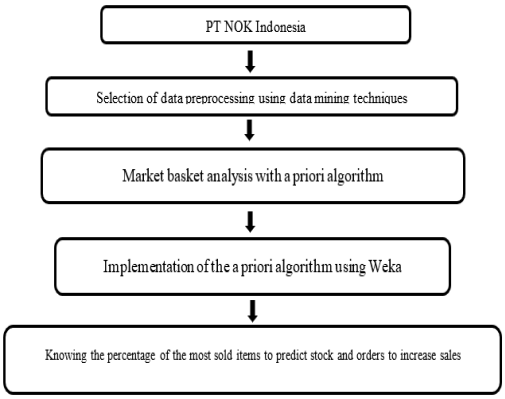


Figure 1. Modeling Phase Image

Information :

a) Ultrabush sales transaction data that will be used in testing
b) Datasets
c) Market basis analysis:
   - Merger
   - Pruning with minimum support and minimum confidence
d) Implementation using Weka
e) Find out the sales results of the most sold products for into stock and order.

g. Evaluation Phase

The evaluation phase is carried out to obtain quality and effectiveness before distribution. Has the model met the initial objectives and has solved the problems in this research, as well as making decisions regarding the use of the results from data mining.

h. Deployment Phase

After carrying out the evaluation, the next phase is the distribution phase, namely the results of this research can be used as recommendations by the PT NOK Indonesia data manager in increasing sales to manage stock of goods and positioning of goods by paying attention to the goods that are most often purchased simultaneously.

RESULT

Preparing Data

In this research, researchers used sales transaction data ultrabush from PT. Nok Indonesia. This data amounts to 117 data from 413 sales transaction data, which contains production date, item and customer. The following is a picture of PT Nok Indonesia's ultrabush sales transaction data.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

Table 4. Sales Transaction Data
(Source: PT. Nok Indonesia Sales Transaction Data)

M17

No.	Transaction date	Items	Customers
1	3/27/2018	RB 1095	SUZUKI
	3/27/2018	RB 1104	MITSUBA
	3/27/2018	RB 1060	YAMAHA
	3/27/2018	RB 1087	MITSUBA
	3/27/2018	RB 1099	YAMAHA
2	3/28/2018	RB 1060	YAMAHA
	3/28/2018	RB 1087	MITSUBA
	3/28/2018	RB 1099	YAMAHA
3	3/29/2018	RB 1087	MITSUBA
	3/29/2018	RB 1099	YAMAHA
	3/29/2018	RB 1104	SGS
	3/29/2018	RB 1060	YAMAHA
4	04/01/2018	RB 1104	SGS
	04/01/2018	RB 1099	YAMAHA
	04/01/2018	RB 1111	MITSUBA
	04/01/2018	RB 1060	YAMAHA
5	04/02/2018	RB 1104	MITSUBA

This data will be analyzed using an a priori algorithm with the aim of finding out how to utilize data mining to produce information on sales of ultrabush products at PT. NOK Indonesia.

Data Preprocessing

From the transaction data in table 4, data preprocessing will be carried out using data mining techniques. Preprocessing can be seen in table 5.

Table 5. Image of Transaction Data Preprocessing

O8

TRANSACTION	RB 1099	RB 1060	RB 1095	RB 1104	RB 1067	RB 1087	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
1	Y	Y	Y	Y	N	Y	N	N	N	N	N	N
2	Y	Y	Y	N	N	N	N	N	N	N	N	N
3	Y	Y	N	Y	N	Y	N	N	N	N	N	N
4	Y	Y	N	Y	N	N	Y	N	N	N	N	N
5	Y	Y	Y	Y	N	Y	N	N	N	N	N	N
6	Y	Y	N	N	N	N	Y	N	Y	N	N	N
7	Y	Y	N	N	N	Y	N	N	N	N	N	N
8	Y	N	N	N	N	Y	Y	N	N	N	N	N
9	N	Y	N	N	N	Y	N	N	N	N	N	N
10	Y	N	Y	Y	N	Y	Y	N	N	N	N	N
11	N	Y	N	Y	N	N	N	N	N	N	N	N
12	Y	N	N	Y	N	N	Y	N	N	N	N	N
13	Y	Y	N	N	N	N	N	Y	Y	N	N	N
14	Y	Y	Y	Y	N	N	Y	N	N	N	N	N
15	N	N	N	Y	Y	N	N	N	N	N	N	N
16	Y	N	N	N	N	N	Y	N	N	N	N	N
17	Y	Y	N	N	Y	N	N	N	N	N	N	N
18	Y	Y	N	N	Y	N	N	Y	N	N	N	N

Data preprocessing will be used for manual calculations with using an a priori algorithm. From the data above, the next stage is to form a tabular format, which can be seen in table 6.

Table 6. Tabular Format Image

O7

NO	RB 1099	RB 1060	RB 1095	RB 1104	RB 1087	RB 1067	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
1	1	1	1	1	1	1	1	0	1	0	0	1
2	1	1	1	0	1	0	0	1	0	1	0	0
3	1	1	0	1	1	1	0	0	1	1	0	0
4	1	1	0	1	0	1	1	0	0	1	1	0
5	1	1	1	1	1	1	0	0	0	1	0	0
6	1	1	0	0	0	0	1	0	1	0	0	0
7	1	1	0	1	1	1	0	1	0	0	1	1
8	1	0	0	1	1	0	1	0	0	0	0	0
9	0	1	0	1	1	0	0	1	1	0	1	0
10	1	0	1	1	1	0	1	0	0	0	0	0
11	0	1	0	1	0	1	0	1	0	1	0	0
12	1	0	0	1	0	0	1	0	0	0	0	0
13	1	1	0	0	0	1	0	1	1	0	0	0
14	1	1	1	1	0	1	1	0	0	0	0	0
15	0	0	0	1	0	1	1	0	0	1	1	1
16	1	0	1	0	0	1	1	0	0	0	0	0
17	1	1	0	0	0	1	0	0	0	0	0	0
18	1	1	0	0	0	1	0	1	0	0	0	0
19	1	1	1	1	0	1	0	0	0	1	0	0

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

Calculation

In this research, we will use an a priori algorithm with 117 data.

Table 7. Sales Transaction Table

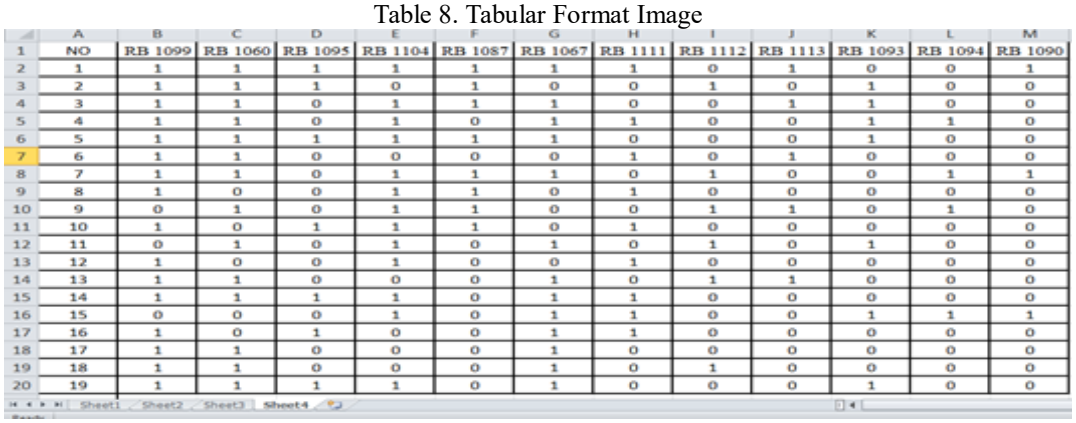
No	Transaction
1	RB 1095, RB 1104, RB 1060, RB 1087, RB 1099, RB 1111, RB 1113, RB 1090, RB 1067
2	RB 1060, RB 1095, RB 1099, RB 1087, RB 1112, RB 1093
3	RB 1087, RB 1099, RB 1104, RB 1060, RB 1113, RB 1093, RB 1067
4	RB 1104, RB 1099, RB 1111, RB 1060, RB 1093, RB 1094, RB 1067
5	RB 1104, RB 1099, RB 1060, RB 1087, RB 1095, RB 1067, RB 1093
6	RB 1099, RB 1111, RB 1113, RB 1060
7	RB 1060, RB 1087,RB 1099, RB 1104, RB 1112, RB 1090, RB 1094, RB 1067
8	RB 1087, RB 1099, RB 1111, RB 1104
9	RB 1060, RB 1087, RB 1104, RB 1112, RB 1113, RB 1094
10	RB 1099, RB 1104, RB 1087, RB 1111, RB 1095
11	RB 1060, RB 1104, RB 1067, RB 1112, RB 1093
12	RB 1104, RB 1099, RB 1111
13	RB 1099, RB 1112, RB 1113, RB 1060, RB 1067
14	RB1060, RB 1095, RB 1099, RB 1111, RB 1104, RB 1067
15	RB 1104, RB 1067, RB 1111, RB 1093, RB 1094, RB 1090
16	RB 1099, RB 1111, RB 1095, RB 1067
17	RB 1060, RB 1099, RB 1067
18	RB 1067, RB 1099, RB 1112, RB 1060
19	RB 1060, RB 1099, RB 1095, RB 1093, RB 1104, RB 1067
20	RB 1093, RB 1099, RB 1060, RB 1104
21	RB 1060, RB 1087, RB 1095, RB 1104, RB 1099, RB 1067
22	RB 1099, RB 1111, RB 1095, RB 1067, RB 1113, RB 1090, RB 1094
23	RB 1099, RB 1111, RB 1112, RB 1095, RB 1067, RB 1087
24	RB 1111, RB 1099, RB 1112, RB 1090, RB 1104
25	RB 1111, RB 1112, RB 1099, RB 1104
26	RB 1111, RB 1099, RB 1104, RB 1060, RB 1067
27	RB 1087, RB 1099, RB 1060, RB 1095, RB 1104, RB 1112, RB 1067
28	RB 1099, RB 1060, RB 1067, RB 1113, RB 1093, RB 1094, RB 1090
29	RB 1067, RB 1060, RB 1087, RB 1104
30	RB 1104, RB 1087, RB 1111, RB 1112, RB 1099
31	RB 1099, RB 1094, RB 1087, RB 1111, RB 1104
32	RB 1099, RB 1112, RB 1111, RB 1060, RB 1104, RB 1067
33	RB 1094, RB 1104, Rb 1087, RB 1099, Rb 1060, RB 1113, RB 1090, RB 1067
34	RB 1099, RB 1111, RB 1112, RB 1095, Rb 1067
35	RB 1099, RB 1111, RB 1113, RB 1112
36	RB 1099, RB 1112, RB 1060, RB 1094
37	RB 1067, RB 1095, RB 1099, RB 1104, RB 1060, RB 1087
38	RB 1099, RB 1111, RB 1060, RB 1095, RB 1112
39	RB 1099, RB 1111, RB 1113, RB 1067
40	RB 1067, RB 1095, RB 1104, RB 1099, RB 1090
41	RB 1095, RB 1104, RB 1099, RB 1087, RB 1060, RB 1067

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

42	RB 1099, RB 1112, RB 1111, RB 1113, RB 1060, RB 1067
43	RB 1113, RB 1111, RB 1099, RB 1060, RB 1095, RB 1104, RB 1093
44	RB 1099, RB 1111, RB 1113, RB 1060, RB 1067
45	RB 1111, RB 1067, RB 1090, RB 1095, RB 1104
46	RB 1067, RB 1099, RB 1112, RB 1111, RB 1095, RB 1087, RB 1104
47	RB 1111, RB 1113, RB 1099, RB 1060, RB 1104, RB 1067
48	RB 1060, RB 1099, RB 1087, RB 1095
49	Rb 1095, RB 1067, RB 1104, RB 1099, RB 1111, RB 1090
50	RB 1060, RB 1099, RB 1087, RB 1095, RB 1111
51	RB 1095, RB 1067, RB 1104
52	RB 1060, RB 1099, RB 1087, RB 1090
53	RB 1060, RB 1095, RB 1090, RB 1067
54	RB 1095, RB 1067, RB 1060, RB 1113, RB 1093, RB 1094
55	RB 1090, RB 1060, RB 1099, RB 1104, RB 1067
56	RB 1099, RB 1112, RB 1111, RB 1090
57	RB 1087, RB 1090, RB 1095, RB 1093, RB 1067
58	RB 1060, RB 1099, RB 1093, RB 1094, RB 1090, RB 1067
59	Rb 1104, RB 1095, RB 1112
60	RB 1099, RB 1090, RB 1111, RB 1113
61	RB 1113, RB 1112, RB 1099, RB 1060, RB 1067, RB 1087, RB 1104
62	RB 1099, RB 1111, RB 1095, RB 1104, RB 1090
63	RB 1104, RB 1099, RB 1095, RB 1093
64	RB 1093, RB 1104, RB 1060, RB 1099, RB 1113, RB 1094, RB 1067
65	RB 1090, RB 1060, RB 1104, RB 1099, RB 1067
66	RB 1094, RB 1113, RB 1111
67	RB 1090, RB 1067, RB 1060
68	RB 1094, RB 1099, RB 1111, RB 1093
69	RB 1087, RB 1090, RB1094
…..	
117	RB 1093, RB 1095, RB 1090, RB 1099, Rb 1111, RB 1087

Table 8. Tabular Format Image



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	NO	RB 1099	RB 1060	RB 1095	RB 1104	RB 1087	RB 1067	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
2	1	1	1	1	1	1	1	1	0	1	0	0	1
3	2	1	1	1	0	1	0	0	1	0	1	0	0
4	3	1	1	0	1	1	1	0	0	1	1	0	0
5	4	1	1	0	1	0	1	1	0	0	1	1	0
6	5	1	1	1	1	1	1	0	0	0	1	0	0
7	6	1	1	0	0	0	0	1	0	1	0	0	0
8	7	1	1	0	1	1	1	0	1	0	0	1	1
9	8	1	0	0	1	1	0	1	0	0	0	0	0
10	9	0	1	0	1	1	0	0	1	1	0	1	0
11	10	1	0	1	1	1	0	1	0	0	0	0	0
12	11	0	1	0	1	0	1	0	1	0	1	0	0
13	12	1	0	0	1	0	0	1	0	0	0	0	0
14	13	1	1	0	0	0	1	0	1	1	0	0	0
15	14	1	1	1	1	0	1	1	0	0	0	0	0
16	15	0	0	0	1	0	1	1	0	0	1	1	1
17	16	1	0	1	0	0	1	1	0	0	0	0	0
18	17	1	1	0	0	0	1	0	0	0	0	0	0
19	18	1	1	0	0	0	1	0	1	0	0	0	0
20	19	1	1	1	1	0	1	0	0	0	1	0	0

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

DISCUSSION

Testing the Apriori Algorithm using WEKA

Testing the analysis results is very important to ensure whether the analysis results are in accordance with the results of data processing carried out manually in CHAPTER III. The correctness can be tested using one of the WEKA data mining software applications. The tool used for data processing in this research is WEKA version 3.8 and make sure it is installed on the computer.

a. All variables used are saved in a Microsoft Excel document file with the file name "data.xslx" as seen in table 9.

Table 9. Initial Transaction Data

NO	RB 1099	RB 1060	RB 1095	RB 1104	RB 1087	RB 1067	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
1	1	1	1	1	1	1	1	0	1	0	0	1
2	1	1	1	0	1	0	0	1	0	1	0	0
3	1	1	0	1	1	1	0	0	1	1	0	0
4	1	1	0	1	0	1	1	0	0	1	1	0
5	1	1	1	1	1	1	0	0	0	1	0	0
6	1	1	0	0	0	0	1	0	1	0	0	0
7	1	1	0	1	1	1	0	1	0	0	1	1
8	1	0	0	1	1	0	1	0	0	0	0	0
9	0	1	0	1	1	0	0	1	1	0	1	0
10	1	0	1	1	1	0	1	0	0	0	0	0
11	0	1	0	1	0	1	0	1	0	1	0	0
12	1	0	0	1	0	0	1	0	0	0	0	0
13	1	1	0	0	0	1	0	1	1	0	0	0
14	1	1	1	1	0	1	1	0	0	0	0	0
15	0	0	0	1	0	1	1	0	0	1	1	1
16	1	0	1	0	0	1	1	0	0	0	0	0
17	1	1	0	0	0	1	0	0	0	0	0	0
18	1	1	0	0	0	1	0	1	0	0	0	0

The values in the form of numbers are transformed into letters which have a certain range of values.

Table 10. Letter Format Transaction Data

TRANSACTION	RB 1099	RB 1060	RB 1095	RB 1104	RB 1067	RB 1087	RB 1111	RB 1112	RB 1113	RB 1093	RB 1094	RB 1090
1	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y
2	Y	Y	Y	N	N	Y	N	Y	N	Y	N	N
3	Y	Y	N	Y	Y	Y	N	N	Y	Y	N	N
4	Y	Y	N	Y	Y	N	Y	N	N	Y	Y	N
5	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	N
6	Y	Y	N	N	N	N	Y	N	Y	N	N	N
7	Y	Y	N	Y	Y	Y	N	Y	N	N	Y	Y
8	Y	N	N	Y	N	Y	Y	N	N	N	N	N
9	N	Y	N	Y	N	Y	N	Y	Y	N	Y	N
10	Y	N	Y	Y	N	Y	Y	N	N	N	N	N
11	N	Y	N	Y	Y	N	N	Y	N	Y	N	N
12	Y	N	N	Y	N	N	Y	N	N	N	N	N
13	Y	Y	N	N	Y	N	N	Y	Y	N	N	N
14	Y	Y	Y	Y	Y	N	Y	N	N	N	N	N
15	N	N	N	Y	Y	N	Y	N	N	Y	Y	Y
16	Y	N	Y	N	Y	N	Y	N	N	N	N	N
17	Y	Y	N	N	Y	N	N	N	N	N	N	N
18	Y	Y	N	N	Y	N	N	Y	N	N	N	N
19	Y	Y	Y	Y	Y	N	N	N	N	Y	N	N
20	Y	Y	N	Y	N	N	N	N	N	Y	N	N
21	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N
22	Y	N	Y	N	Y	N	Y	N	Y	N	Y	Y
23	Y	N	Y	N	Y	Y	Y	Y	N	N	N	N
24	Y	N	N	Y	N	N	Y	Y	N	N	N	Y

Then the "Data" file will be saved as a CSV type file (*.CSV). Furthermore, if the file is opened from Microsoft Word, Notepad, or another text editor program, it will appear that it has changed to a compressed format. Run the WEKA tool until the GUI Chooser appears as in figure 2.

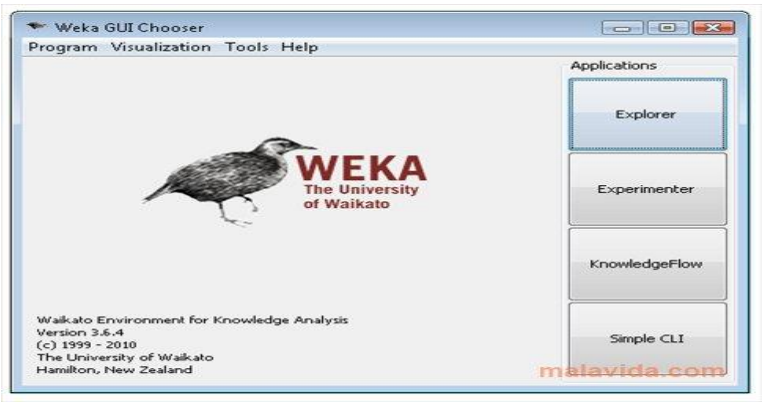


Figure 2. Tools WEKA

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

After running the WEKA tool, it will display the calculation results of the a priori algorithm, bringing up the association rules found from the process. As in figure 3.

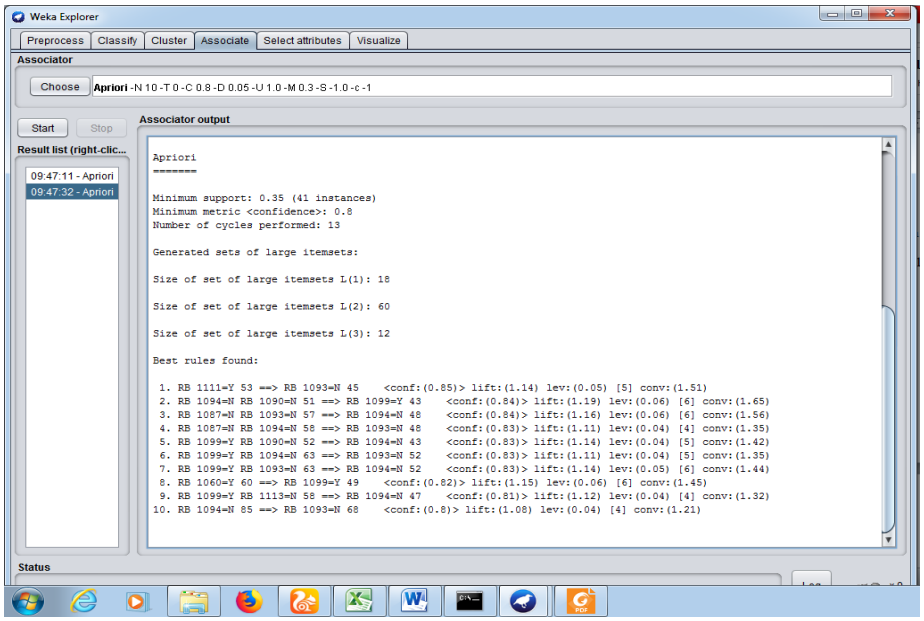


Figure 3. Results of a priori calculations with WEKA
(Source: Research, 2018)

Test Results

The results of a priori testing with Weka with the percentage of items sold to increase sales at PT NOK Indonesia are as follows:

1. If you buy RB 1111 then you will not buy RB 1093 with 85% confidence (there are RB 1111 transactions but no RB 1093 transactions)
2. If you don't buy RB 1094 and don't buy RB 1090 then you won't buy RB 1099 with 84% confidence (there are RB 1094 and RB 1090 transactions but no RB 1099 transactions)
3. If you don't buy RB 1087 and don't buy RB 1093 then you won't buy RB 1094 with 84% confidence (no RB 1087, RB 1093 and RB 1094 transactions)
4. If you don't buy RB 1087 and don't buy RB 1094 then you won't buy RB 1093 with 83% confidence (no RB 1087, RB 1094 and RB 1093 transactions)
5. If you buy RB 1099 and don't buy RB 1090 then you won't buy RB 1094 with 83% confidence (there are RB 1087 transactions but no RB 1090 and RB 1094 transactions)
6. If you buy RB 1099 and don't buy RB 1094 then you won't buy RB 1093 with 83% confidence (there are RB 1099 transactions but there are no RB 1094 and RB 1093 transactions)
7. If you buy RB 1099 and don't buy RB 1093 then you won't buy RB 1094 with 83% confidence (there are RB 1099 transactions but there are no RB 1093 and RB 1094 transactions)
8. If you buy RB 1060 you will buy Rb 1099 with 82% confidence (there are RB 1060 and RB 1099 transactions)
9. If you buy RB 1099 and don't buy RB 1113 then you won't buy RB 1094 with 81% confidence (there are RB 1099 transactions but no RB 1113 and RB 1094 transactions)
10. If you don't buy RB 1094 and don't buy RB 1093 with 80% confidence (no RB 1094 and RB 1093 transactions)
11. So based on the results of manual calculations using the basic methodology of association analysis with a minimum support value of 30% and a minimum confidence value of 80%, the resulting trend for ultra bush item types is:

Table 11. Calculation Results

No	Calculation Techniques	Results
1	Manual calculation	If you buy RB 1060 you will buy RB 1099 with 81% confidence
2	Calculations using WEKA	Jika membeli RB 1060 maka akan membeli RB 1099 dengan *confidence 82%*

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

With a minimum support value of 30% and a minimum confidence value of 80%, no strong association rules were found for the Ultra Bush product in this transaction data example. However, there are several itemsets that are often purchased together with Ultra Bush, namely Shampoo, Conditioner, and Soap, even though the resulting association rules do not meet the specified confidence threshold. This suggests that even if co-purchasing patterns exist, the strength of the association may not be strong enough for a particular marketing strategy without adjusting the analysis parameters.

CONCLUSION

From the analysis and discussion described in the research above, the researcher can conclude that Data Mining uses the Association Rule to increase sales of ultra bush by applying data mining using the a priori algorithm method, that is, with this method it is possible to identify the most sold product items. From the results of manual calculations it was found that consumers who bought RB 1060 would buy RB 1099 with 81% confidence, whereas using WEKA it was found that consumers who bought RB 1060 would buy RB 1099 with a confidence value of 82%. So with the conclusion of the calculation results, it can be hoped that PT NOK Indonesia can increase sales by marketing products in other types by examining what the advantages of the most sold products are compared to other products and can increase stock supplies.

REFERENCES

Akbar, R., Sukmawati, U. S., & Katsirin, K. (2024). Analisis Data Penelitian Kuantitatif. *Jurnal Pelita Nusantara*, *1*(3), 430–448. https://doi.org/10.59996/jurnalpelitanusantara.v1i3.350

Ariyani, M. (2021). Mekanisme Pengelolaan Persediaan Sparepart Sepeda Motor Honda Pada PT. Bintang Motor Jaya, Tbk Cabang Cirebon. *Exchall: Economic Challenge*, *2*(1), 71–105. https://doi.org/10.47685/exchall.v2i1.101

Bimantoro, T., & Wardhani, A. K. (2020). IMPLEMENTASI ALGORITMA PARTITIONING AROUND MEDOIDS DALAM PENGELOMPOKAN RESTORAN. *Indonesian Journal of Technology, Informatics and Science (IJTIS)*, *2*(1), 33–36. https://doi.org/10.24176/ijtis.v2i1.5651

Erwansyah, K., Andika, B., & Gunawan, R. (2021). Implementasi Data Mining Menggunakan Asosiasi Dengan Algoritma Apriori Untuk Mendapatkan Pola Rekomendasi Belanja Produk Pada Toko Avis Mobile. *J-SISKO TECH (Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD)*, *4*(1), 148. https://doi.org/10.53513/jsk.v4i1.2628

Erwin Stanley Harianja. (2020). Pengembangan Media Penyimpanan dalam Sistem Berkas (Studi Kasus Mahasiswa STMIK Eresha). *Jurnal Ilmu Komputer*, *3*(2), 5–9.

Firdaus, A. A., Iksan, N., Sadiah, D. N., Sagita, L., & Setiawan, D. (2021). Penerapan Algoritma Apriori untuk Prediksi Kebutuhan Suku Cadang Mobil. *Jurnal Sistem Dan Teknologi Informasi (Justin)*, *9*(1), 13. https://doi.org/10.26418/justin.v9i1.41151

Husdi, H., & Dalai, H. (2023). Penerapan Metode Regresi Linear Untuk Prediksi Jumlah Bahan Baku Produksi Selai Bilfagi. *Jurnal Informatika*, *10*(2), 129–135. https://doi.org/10.31294/inf.v10i2.14129

Hutasuhut, M., Octavina, D., & Halim, J. (2019). Penerapan Data Mining dalam Menganalisa Pola Kelayakan Siswa Pada Kelas Unggulan Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) pada SMP N. 2 Rantau Selatan. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, *18*(2), 154. https://doi.org/10.53513/jis.v18i2.154

Kharis, S. A. A., & Zili, A. H. A. (2022). Learning Analytics dan Educational Data Mining pada Data Pendidikan. *JURNAL RISET PEMBELAJARAN MATEMATIKA SEKOLAH*, *6*(1), 12–20. https://doi.org/10.21009/jrpms.061.02

Mulya, M. F., Rismawati, N., & Alifi, R. R. (2019). ANALISIS DAN IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI UNTUK MENINGKATKAN PENJUALAN PADA KANTIN UNIVERSITAS TANRI ABENG. *Faktor Exacta*, *12*(3), 210. https://doi.org/10.30998/faktorexacta.v12i3.4541

Priyanka Aggarwal. (2019). Impact of Corporate Governance on Corporate Financial Performance. *IOSR Journal of Business and Management*, *13*(3), 1–5.

Putri Mai Sarah Tarigan, Jaya Tata Hardinata, Hendry Qurniawan, Muhammad Safii, & Riki Winanjaya. (2022). IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI DALAM MENENTUKAN PERSEDIAAN BARANG (STUDI KASUS : TOKO SINAR HARAHAP). *Jurnal Sistem Informasi, Teknologi Informasi Dan Komputer*, *12*(2), 51–61.

Ratih Wahyuningrum. (2018). ANALISIS STRATEGI E-MARKETING UNTUK MENINGKATKAN MINAT BELI ONLINE. *Jurnal Manajemen Bisnis*, *21*(3), 275–290.

Saefudin, S., & DN, S. (2019). PENERAPAN DATA MINING DENGAN METODE ALGORITMA APRIORI UNTUK MENENTUKAN POLA PEMBELIAN IKAN. *JSiI (Jurnal Sistem Informasi)*, *6*(2), 36. https://doi.org/10.30656/jsii.v6i2.1587

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.4185

Sibarani, A. J. P. (2020). Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *7*(2), 262–276. https://doi.org/10.35957/jatisi.v7i2.195

Sutrisno, S. (2020). PENERAPAN ALGORITMA APRIORI UNTUK MENCARI POLA PENJUALAN PRODUK DANA PADA PT BANK RAKYAT INDONESIA (PERSERO) TBK KANCA JAKARTA PASAR MINGGU. *Jurnal Sistem Informasi Dan Informatika (Simika)*, *3*(1), 12–26. https://doi.org/10.47080/simika.v3i1.834

Syaipul Amri. (2018). Pengaruh Kepercayaan Diri (Self Confidence) Berbasis Ekstrakurikuler Pramuka Terhadap Prestasi Belajar Matematika Siswa Sma Negeri 6 Kota Bengkulu . *Jurnal Pendidikan Matematika Raflesia*, *3*(2), 78–85.

Tualeka, S., Alameka, F., & Wanti Wulan Sari, N. (2021). IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI PENJUALAN DAN PENEMPATAN STOK BARANG PADA CV PASTI JAYA HOUSEWARE DENGAN MENGGUNAKAN ALGORITMA APRIORI. *SEMINASTIKA*, *3*(1), 115–123. https://doi.org/10.47002/seminastika.v3i1.258

Wahyuni, S., Sulistianingsih, I., Hermansyah, Hariyanto, E., & Cindi Veronika Lumbanbatu, O. (2021a). Data Mining Prediksi Minat Customer Penjualan Handphone Dengan Algoritma Apriori. *JURNAL UNITEK*, *14*(2), 10–19. https://doi.org/10.52072/unitek.v14i2.243

Wahyuni, S., Sulistianingsih, I., Hermansyah, Hariyanto, E., & Cindi Veronika Lumbanbatu, O. (2021b). Data Mining Prediksi Minat Customer Penjualan Handphone Dengan Algoritma Apriori. *JURNAL UNITEK*, *14*(2), 10–19. https://doi.org/10.52072/unitek.v14i2.243

Wijaya, A., Faqih, A., Solihudin, D., Rohmat, C. L., & Eka Permana, S. (2024). PENERAPAN ASSOCIATION RULES MENGGUNAKAN ALGORITMA APRIORI UNTUK IDENTIFIKASI POLA PEMBELIAN. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(6), 3871–3878. https://doi.org/10.36040/jati.v7i6.8270