http://infor.seaninstitute.org/index.php/infokum/index
JURNAL INFOKUM, Volume 9, No. 2,Juni 2021 ISSN : 2302-9706

IMPLEMENTATION OF APRIORI ALGORITHM IN PREDICTING CAR PARTS

Syahman Putra Pratama Nst[1], Ibnu Rasyid Muthe[2], Rahma Mutiah[3]
[1,2,3]Information Management, Labuhan Batu University, Rantauprapat, Indonesia
syahmanp.3ips@gmail.com[1], ibnurasyidmunthe@gmail.com[2]

Article Info
Received 25 March 2021
Revised 25 April 2021
Accepted 6 June 2021

Abstract

In the procurement of spare parts for the car production process, information or knowledge related to auto parts is needed so that the company's production processes and results are effective and efficient. The purpose of this study is to apply an a priori algorithm for prediction of auto parts that often appear. The research subject used is the BMW car production data of PT Gaya Motor. The research model used is market basket analysis. The stages of the research carried out include: (1) Data Collection; (2) Training Data; (3) Formation of Association Rule, (4) Lift Ratio Test, and (5) Drawing Conclusion. The research results obtained are the most widely produced BMW car types in 2018 are the BMW 320 and BMW 7 Series. So the company can use these results to determine strategies related to the procurement of spare parts for that type of car. Based on the Lift Ratio Test that has been carried out, there are two very strong and valid rules to be used in the prediction of BMW auto parts, namely the BMW 320 and BMW 7 SERIES.

Keywords: Apriori, Data Mining, Prediction, Spare Parts.

1. Introduction

Along with the times that require high mobility, the need for transportation for daily activities is also increasing. This provides opportunities for automotive companies. For a company, this is an opportunity to dominate the market in the transportation sector, one of which is cars. So we need the right way to increase the production of quality cars. In the procurement of spare parts for the production process, related information or knowledge is needed so that the purchase and inventory of spare parts is in accordance with production information. One solution for company owners is to use data mining to gain knowledge regarding related spare parts for decision making in car production in the future. According to Larose, Data mining itself is a process of extracting information to find important patterns (pattern recognition) in piles of data in the database so that it becomes knowledge (knowledge discovery). The functions in data mining include description functions, estimation functions, prediction functions, classification functions, clustering functions and association functions [1].

One of the data mining functions that are very useful in production activities is prediction. Prediction function is a process to find patterns from data by using several variables to predict other variables whose type or value is unknown [2]. Through predictions, making various decisions has a clear basis and is proven to be profitable in various economic activities such as the use of electricity in housing to industrial scale [3] [4]. Therefore, the car industry also requires data mining to maximize decision making.

One algorithm that is often used is a priori. Based on previous studies, the use of the a priori algorithm was chosen because this algorithm is effective and efficient in determining the association pattern of an item. In addition, this algorithm is also easy to implement and does not take long in the process. The type of research used in this research is experimental research. This research is used to apply data mining based on associations using a priori algorithm, to facilitate companies or analysts in knowing the company's production patterns. There are several studies related to the a priori algorithm and its implementation in various fields. The first research was conducted by Ristianingrum and Sulastri (2017) with the research title "Implementation of Data Mining Using the Apriori Algorithm". The a priori algorithm in this study helps in determining which spare parts products are purchased based on

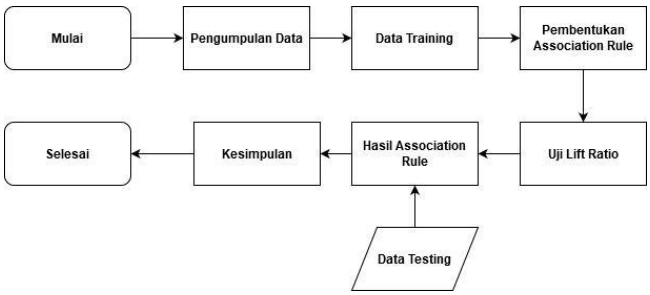
consumer transaction patterns that can be identified from the information and services that appear [7]. The a priori algorithm for determining spare parts was also carried out by Saputra et al (2016). In this study, the existing sales data is processed using the a priori algorithm, which produces any spare parts products purchased by consumers, especially outer tires and engine oil. Delta Motor can increase the stock of engine oil and tire products [8]. Research using a priori algorithms that become other references include: by Prabawanti (2018) [9], Sholik and Salam (2018) [10] and also conducted by Aditya et al (2016) [11]. Because of the effectiveness and efficiency of this algorithm, the application also needs to be applied to PT.

There are so many types of cars in circulation today. One type of car with a high population and well-known in Indonesia is the BMW. The automotive manufacturer BMW is the parent company of Rolls-Royce Motor [5]. The object of this research is the BMW car production data at PT Gaya Motor. PT. Motor Style [6]. In making decisions regarding the prediction of spare parts for production, PT Gaya Motor has not implemented data mining, especially the a priori algorithm. the use of a priori algorithm is very much needed for decision making in buying what spare parts will be needed and the amount according to the type of car produced.

A priori algorithm is one of the association techniques in data mining. The a priori algorithm can assist in providing information on recommendations for stocking goods and facilitating the arrangement and placement of spare parts for the car production process at PT Gaya Motor. The formulation of the problem in this research is how to apply the a priori algorithm for prediction of BMW auto parts at PT Gaya Motor.

2. Research methods

The research method used is the Association model or Market Basket Analysis. The type of data used in this study is secondary data obtained from the company's official website. This research was conducted in several stages, the following are the steps in conducting the research. The steps or stages taken to obtain association rules on car production data at PT Gaya Motor Jakarta Utara are: (1) Data Collection; (2) Training Data; (3) Formation of Association Rule, (4) Lift Ratio Test, and (5) Drawing Conclusion. More complete research stages are presented in the following flowchart.



Picture 1. Research Stages

A. Data collection

PT. Gaya Motor is precisely located at Jalan Gaya Motor Raya No. 3 Sunter II, Tanjung Priok, North Jakarta. At first PT. Gaya Motor is located on Jalan Sulawesi Ujung No. 22 Tanjung Priok (at the port of Tanjung Priok). The data used in this study is secondary data through documentation, namely the Car Production Data of PT. Gaya Motor in January-December 2018 [6] [12].

B. Training Data

At this stage, the selection of appropriate data and the creation of data tables are carried out so that it is easy to apply the a priori algorithm. The data used are selected and analyzed. While the data transformation is done by changing the form of production data into tables. After the data is ready to be processed, the next step is the application of the a priori algorithm.

C. Formation of Association Rules

The application of the a priori algorithm goes through two processes and benchmarks, namely support and confidence. Support (supporting value) is the percentage of the combination of these items in the database, while confidence (certainty value) is the strength of the relationship between items in the association rule [13]. In the formation of the production pattern by the a priori algorithm, there are two stages. The first stage is to find the frequent itemset set of items that meet the minimum support value. Then the second stage is to form an association pattern from the frequent itemset that has been obtained using the confidence value [14]. After the two patterns are carried out, the results are analyzed and described to find research conclusions.

D. Test Lift Ratio

Lift Ratio test is a method used to measure the strength of associations that have been formed. Lift Ratio can be used as a determinant of the strength and validity of the association pattern. The following is the formula for Lift Ratio [13].

$$Lift\ Ratio = \frac{Confidence\ (A,B)}{Benchmark\ Confidence\ (A,B)} \quad (1)$$

Benchmark Confidence is calculated by the following formula.

$$Benchmark\ Confidence = \frac{Nc}{N}$$

Information :
Nc = Number of transactions with consequent items
N = Number of database transactions

E. Drawing Conclusion

In the formation of association patterns by the a priori algorithm, there are two stages, namely looking for frequent itemsets to form association patterns. After the two patterns are carried out, the results are analyzed and described to find research conclusions.

3. Results and Discussion

The data obtained is processed into a dataset that is ready for use. The data used are BMW car production data with several types during 2018. The following is a list of BMW car types to be analyzed, which can be seen in Table 1.

Table 1. BMW Car Type

No	*Type*	*Type*
1.	BMW 320	Car
2.	BMW 328/330	Car
3.	BMW 520	Car
4.	BMW 528	Car
5.	BMW 7 SERIES	Car

The first process is to determine the itemset table for each transaction. Car items are represented by the letters one to five. One for the BMW 320 type, two for the BMW 328 type and so on as shown in table 1. Transaction data is made monthly from January-December and is denoted by the number 1 for the month of January, the number 2 for the month of February and so on. The following are the itemsets that have been created, which can be seen in Table 2.

Table 2. Data Item Set

No	*Itemset*
1	1,5,4
2	5,1,2
3	4,5,3
4	4,2,1
5	1,4,5

6	1.5,3
7	1,2,5
8	1,2,3
9	3,4,2
10	1,4,3
11	4,1,5
12	3,2,4

A. Application of the Apriori Algorithm

The application of the a priori algorithm is carried out in the following stages.

1) Formation of 1 Itemset: the minimum specified is 30% (Golden Rule). The following is the formation of 1 itemset using the formula below [15].

$$\frac{\sum transaksi\ mengandungA}{\sum transaksi} \boldsymbol{x}\ \mathbf{100\%} \quad (2)$$

Table 3 is the result of forming 1 itemset.

Table 3. Formation of 1 Itemset

No	*Type*	*Support*
1	1	75%
2	2	50%
3	3	50%
4	4	66.67%
5	5	58.3%

2) Formation of 2 Itemsets: The process of forming 2 itemsets with a minimum support of 30% is done using the following formula [15].

$$\frac{\sum transaksi\ mengandungAdanB}{\sum transaksi} \boldsymbol{x}\ \mathbf{100\%} \quad (3)$$

The results of the formation of 2 itemsets are shown in Table 4.

Table 4. Formation of 2 Itemset

No	Type	Support
1.	1, 2	**33.33%**
2.	1, 3	25%
3.	1, 4	**41.67%**
4.	1.5	50%
5.	2, 5	16.67%
6.	2, 4	25%
7.	2, 3	20%
8.	3, 4	**33.33%**
9.	3, 5	16.67%
10.	5, 4	**33.33%**

From the formation of 2 itemsets with a minimum support of 30% in Table 5, several combinations that meet the minimum support are obtained, namely:

Table 5. Combinations that meet the minimum support

No	*Type*	*Support*
1.	1, 2	33.33%
2.	1, 4	41.67%
3.	3, 4	33.33%
4.	5.4	33.33%

3) Formation of 3 Itemsets: The process of forming 3 itemsets with a minimum support of 30% is carried out using the following formula [15].

$$\frac{\sum transaksi\ mengandung A,B dan C}{\sum transaksi} x\ 100\% \quad (4)$$

Table 6 is the result of the formation of 3 itemset

Table 6. Formation of 3 Itemset

No	*Type*	*Support*
1.	1,5,3	8.33%
2.	1,5,2	16.67%
3.	1,4,5	16.67%
4.	2,1,3	8.33%
5.	2,3,4	16.67%
6.	2,3,5	0%
7.	3,4,5	8.33%

Because in the 3 itemsets above there is no support value above 30%, the association rules will be formed using 2 itemsets.

4) FormationAssociation Rules: The next step is to find the confidence value of the 2 itemset in Table 7. At this stage, the minimum confidence applied is 60%. The formula used is as follows [15].

$$\frac{\sum transaksi\ mengandung A\ dan\ B}{\sum transaksi\ mengandung\ A} x\ 100\% \quad (5)$$

Table 7. Formation of Association Rules

Rules	Confidence
If it produces a BMW 320, it also produces a BMW 328/330	44.44%
If it produces a BMW 328/330, it also produces a BMW 320	**66.67%**
If it produces a BMW 320, it also produces a BMW 528	45%
If you produce a BMW 528, you will also produce a BMW 320	40%
If it produces BMW 320, it also produces BMW 7 SERIES	**66.67%**
If it produces the BMW 7 SERIES, it also produces the BMW 320	**85.714 %**
If it produces a BMW 520, it also produces a BMW 528	**66.67%**
If it produces a BMW 528, it also produces a BMW 520	50%

Discussion

Based on the final association rules. The most widely produced BMW car types are the BMW 320 and the BMW 7 Series and the two types of cars are closely related. The rules of this association can be used as a reference for making strategies and managing companies in increasing production effectiveness.

The use of lift ratio to evaluate the strength of an association rule that has been formed. Lift Ratio is a comparison between the confidence of a rule with the benchmark confidence value that has been obtained. While Benchmark Confidence is a comparison between the number of all consequent items to the total number of transactions that occur [13]. The lift ratio calculation that has been done can be seen in Table 8.

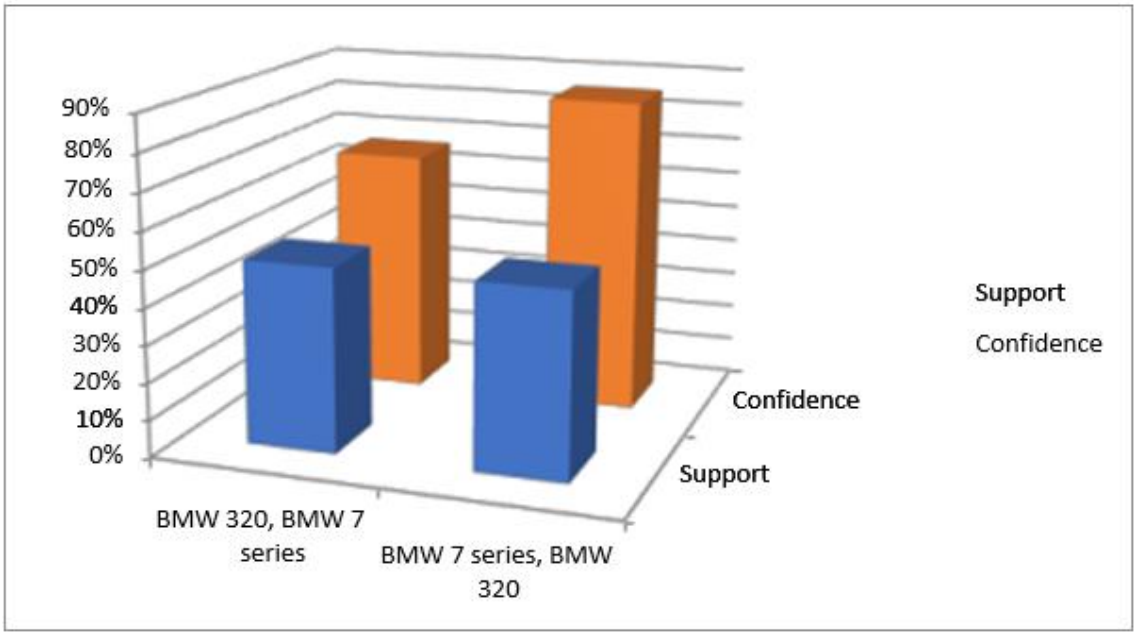


Figure 2. Final Association Rules Graph

Table 8. Lift Ratio Test

If	Then	Confidence	Support	Nc	Bc	Lift Ratio
BMW 320	BMW 328/330	44.44%	83.33%	6	6/12=0.5	0.44/05=0.88
BMW 328/330	BMW 320	66.67%	50%	10	10/12=0.83	0.67/0.83=0.807
BMW 320	BMW 528	45%	83.33%	8	8/12=0.67	0.45/0.670=0.67
BMW 528	BMW 320	40%	66.67%	10	10/12=0.83	0.4/0.83=0.482
BMW 320	**BMW SERIES 7**	**66.67%**	**83.33%**	**7**	**7/12=0.58**	**0.67/0.58=1.155**
BMW 7 SERIES	**BMW 320**	**85.71%**	**58.33%**	**10**	**10/12=0.83**	**0.85/0.83=1.024**
BMW 520	BMW 528	66.67%	50%	8	0.67	0.67/0.67=1
BMW 528	BMW 520	50%	66.67%	6	6/12=0.5	0.5/0.5=1

The lift ratio test results can be said to be strong and valid if the value is more than 1.00 [13]. Based on the lift ratio test table that has been carried out, there are two rules that have a lift ratio value of more than 1 (lift ratio>1). This shows that the two rules are valid to be used to predict BMW car parts by type

4. Conclusion

The purpose of this study is to apply an a priori algorithm for prediction of auto parts that often appear. Based on the research that has been done, it can be concluded that the test results are as follows: a) The most widely produced BMW car types in 2018 are the BMW 320 and BMW 7 Series. So that the company can use these results to determine strategies related to the procurement of spare parts for that type of car, b) Based on the Lift Ratio Test that has been carried out, there are two very strong and valid rules to be used in the prediction of BMW auto parts, namely the BMW 320 and BMW 7 SERIES.

Reference

[1] EP Gunawan and R. Wardoyo, "An Expert System Using Certainty Factor for Determining Insomnia Acupoint," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 12, no. 2, p. 119, 2018, doi:10.22146/ijccs.26328.

[2] V. Mayora, B. Ginting, and FA Sianturi, "LABORATORY USING SIMPLE ADDITIVE WEIGHTING (SAW) METHOD IN GRANMED HOSPITAL," *J. Inform. Nusant lamp.*, vol. 4, no. 2, pp. 1–7, 2019, [Online]. Available: http://e-jurnal.pelitanusantara.ac.id/index.php/JIPN/article/view/606.

[3] B. Sasangka and A. Witanti, "Expert System for Diagnosing Acute Respiratory Infections in Children Using Bayes' Theorem," *JMAI (Jurnal Multimed. Artif. Intell.*, 2019, doi:10.26486/jmai.v3i2.83.

[4] JT Informatics, T. Informatics, SI Industri, T. Informatika, and PC Riau, "Expert System for Diagnosing Kidney Disease Using Methods," *Sis. Experts in Diagnosing Kidney Disease Using the Method. Hill Climbing*, 2012.

[5] T. Hidayat, HN Nasution, SWR Nasution, and R. Fauzi, "EXPERT SYSTEMS TO DIAGNOSE LUPUS DISEASE USING CERTAINTY FACTOR METHOD," *J. Educ. Dev.*, 2019, doi:10.37081/ed.v7i3.1201.

[6] B. Sinaga, PM Hasugian, and AM Manurung, "Expert System Diagnosing Smartphone Damage," vol. 3, no. 1, pp. 333–339, 2018.

[7] W. Purba, D. Ujung, T. Wahyuni, L. Sihaloho, and J. Damanik, "DESIGNING ONLINE TICKET BOOKING INFORMATION SYSTEMS AT KMP . IHAN BATAK BASED," vol. 3, no. 2, pp. 65–75, 2020.

[8] RA Ramadhani, "Expert System for Hepatitis Diagnosis Using the K-NN Method for Primary Health Services," no. November, 2017.

[9] AJ Aziz Amrullah and E. Ekojono, "EXPERT SYSTEM FOR DIAGNOSIS OF LUNG DISEASE WITH FORWARD CHAINING METHOD," *J. Inform. Polynema*, 2017, doi:10.33795/jip.v2i1.48.

[10] Minarni and A. Fadhillah, "Expert System in Detecting Rice Plant Diseases," *J. Dyn.*, vol. 2, no. 1, pp. 11–15, 2017, doi:10.22216/JoD.2017.V2.11-15.

[11] D. Kiray and FA Sianturi, "Diagnose Expert System Computer Malfunction Certainty Factor Method," *J. Comput. Networks, Archit. High Performance. Comput.*, 2020, doi:10,47709/cnapc.v2i1.358.