

Implementation Of Data Mining On Suzuki Motorcycle Sales In Gemilang Motor Prosperous With Apriori Algorithm Method

Anju Eliarsyam Lubis¹, Paska Marto Hasugian²)

^{1,2}Informatics Engineering Study Program, STMIK Pelita Nusantara, Jl. Iskandar Muda No.1
Medan, North Sumatra, Indonesia 20154

Email: anjulubis832@gmail.com¹, paskamarto86@gmail.com²

Abstract-The sale is part of the marketing that determine the survival of the company. With the sale, the company can achieve the goals or targets. To be a company that continues to grow in motorcycle sales, the company should be able to compete in increasing sales volume. Starting from the launch prodak the best in sophistication motorcycles, up to a very attractive price cuts the attention of consumers. Things like that already sangat often do, so the company can still compete, Motorcycles is a two-wheeled transportasi tool used more and more common people. From teenagers to old orag, not infrequently motorcycle including important sangat needs. If we do not have it feels very hard in activity quickly. Make sales without any restriction of sales data accumulate, until finally overwhelmed the company in terms of taking care of customer files. To find the most sales required Apriori Algorithm. Apriori algorithm, including the type of association rules on Data Mining. One stage of association that can produce an efficient algorithm is with high frequency pattern analysis. In an association can be determined by two benchmarks, namely: Support and Confidence. Support "penunang value" is the percentage of combinations of items in a database, and Confidence "value certainty" is strong correlation between the items in an association's rules. Apriori algorithm, including the type of association rules on Data Mining. One stage of association that can produce an efficient algorithm is with high frequency pattern analysis. In an association can be determined by two benchmarks, namely: Support and Confidence. Support "penunang value" is the percentage of combinations of items in a database, and Confidence "value certainty" is strong correlation between the items in an association's rules. Apriori algorithm, including the type of association rules on Data Mining. One stage of association that can produce an efficient algorithm is with high frequency pattern analysis. In an association can be determined by two benchmarks, namely: Support and Confidence. Support "penunang value" is the percentage of combinations of items in a database, and Confidence "value certainty" is strong correlation between the items in an association's rules.

Keywords : Sales, Data Mining, Apriori Algorithm.

1. Introduction

Data Mining is a branch of computer science that is very widely studied and used by computer engineers and programmers. Data mining is a concept to discover knowledge or information that is very valuable in the Database. Data Mining apply mathematics, statistical techniques, and artificial intelligence to decipher and identify a case study in the research. Data mining is widely used as research material in terms of finding a pattern or value in suatu large data base. Data Mining Applications memperoleh input data ie sales data, from the data obtained patterns and information for decision-making on Sejahtera Motor Gemilang in the future. Selection of the appropriate method or algorithm relies heavily on the purpose and process of Knowledge Discovery in Databases (KDD) as a whole. Data Mining Techniques for finding Associative rules or relationships between items called Association rule mining, One algorithm used to find the associative rule is Apriori Algorithm algorithm. Output of data mining can be used for decision-making process in the future. Sejahtera Motor Gemilang is an authorized dealer engaged in sales of Suzuki motorcycles and own service. In this era of technological advancement, a motorcycle is important to support the day-to-day activities. Although in motorcycle sales quite in demand, but often have problems to availability Unit or to place an order is often called pivot. Basically Suzuki Motor Gemilang yet to analyze specific data, for example in combining the booking unit in the same adjacent time, which would cause problems, so it does not know how the relationship between an item to another, and it will require precision and timing a lot. Apriori algorithm, help in the form of a combination candidate item then testing whether that combination meets the minimum parameters of support and confidence which is the minimum threshold value provided by the user. Apriori algorithm can be used in the sales process in Sejahtera Motor Gemilang, to give the relationship between the sales transaction data, in this case is the motorcycle that will be purchased will generate customer data.

2. Theory

Data mining is a process that employs one or more machine learning techniques (machine learning) to analyze and extract knowledge (knowledge) automatically. (Fajar Astuti Hermawati, 2017: 03). *Data mining* is a term that can digynakan for the decomposition of a knowledge discovery in sebuah Database. Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning in extracting and identifying sangat information useful, relevant knowledge from a variety of large databases. (Kusrini and Taufiq Emha Lutfi, 2009: 03). *Data mining* is a series of processes to explore additional value in the form of information that had been unknown to them manually from a database. Dihasilkan information obtained by extracting and recognizing the important or interesting patterns from data contained in the database. Data mining is a technique used to seek knowledge stored in large data bases, so often we called Knowledge Discovery in Databases (Retno Tri Vlandari, 2017: 01).

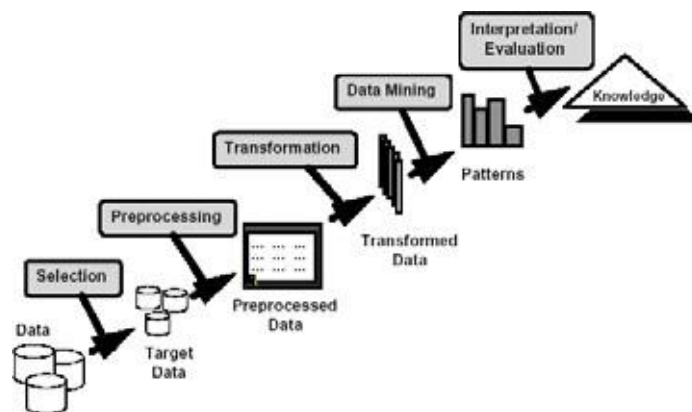


Figure 1. KDD process

3. Research methods

This stage is looking for a combination of items that meet the minimum requirements and the value of the support in the database. Support an item value is obtained by the following formula.

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

Meanwhile, the support of two items obtained from the formula 2 below:

$$Support(A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{\sum Transaksi \text{ Mengandung } A \text{ dan } B}{\sum Transaksi}$$

3.1 Analysis of High Frequency Pattern

Stages of the process of grouping data that can be performed on the a priori algorithm degan entering of data processing and then processed in the formation of itemset resulting combination selected. Priori algorithm stages in the process of data mining system is as follows:

a) Pattern Analysis of High Frequency

At this stage, where the search for a combination of items that meet a minimum requirement and value of support in the database. Support an item value is obtained by the following formula.

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

Table 1.
List Unit Suzuki

No.	Name of goods	Item code
1	Gsx-150R	GR
2	Gsx-150S	GS
3	All New Satria Fi-150	FU
4	Smash FI-110	SH
5	Address FI-110	AD
6	NEX II-110	NX

The above data is the data sample units to be used in this study. Accumulated sales of units of Suzuki in 20 transactions and is presented in the form of sprinkles diagram where each purchase on the mark degan numbers 1 and 0 is no purchase. The chart topping is as follows:

Table 2.
Sales Transaction Data

Transaction	Gsx-R	Gsx-S	Knight Fu	Smash	address	NEX-II
1	1	0	0	0	1	1
2	1	1	0	0	0	1
3	1	1	1	0	0	0
4	1	0	0	0	1	1
5	1	1	0	0	0	1
6	1	0	1	0	0	1
7	1	0	1	0	0	1
8	0	0	1	1	0	1
9	0	1	1	0	0	1
10	1	0	0	0	1	1
11	1	0	1	0	0	1
12	1	1	0	0	0	1
13	1	1	1	0	0	0
14	1	0	0	0	1	1
15	1	1	0	0	0	1
16	1	0	0	1	0	1
17	1	0	0	1	0	1
18	0	0	1	1	0	1
19	1	1	0	1	0	0
20	1	0	0	0	1	1

From table 2, then take as much as 20 transactions representing the overall data. In which the data will determine the value (support = 20% Confidence = 50%). Which meant that a consumer who buys GR and GS has a 50% chance to also buy NX. This rule is quite singnifikat because it represents 20% of the overall transaction records. The calculation is as follows:

$$\begin{aligned} \text{Support (GR)} &= \frac{17}{20} \times 100 = 85\% \\ \text{Support (GS)} &= \frac{8}{20} \times 100 = 40\% \\ \text{Support (FU)} &= \frac{8}{20} \times 100 = 40\% \\ \text{Support (SH)} &= \frac{5}{20} \times 100 = 25\% \\ \text{Support (AD)} &= \frac{5}{20} \times 100 = 25\% \\ \text{Support (NX)} &= \frac{17}{20} \times 100 = 85\% \end{aligned}$$

Table 4
List 1 item set

No.	Item name	amount	value Support
1	GR	17	85
2	GS	8	40
3	FU	8	40
4	SH	5	25
5	AD	5	25
6	NX	17	85

From table 4, it will look for $\geq 20\%$ minimum support of 20 sales transactions. From these results we will obtain data that meet the minimum support are:

Table 5
Data Meets Value Support

No.	Item code	Support Minimum 20%
1	GR	85%
2	GR	40%
3	FU	40%
4	SH	25%
5	AD	25%
6	NX	85%

From table 5, the next will be the formation of C2 or said degan formation degan 2 item minimum amount of support = 20%, and Confidence = 50%, completed degan the following formula:

$$\text{Support (A, B)} = \frac{\sum \text{Transaksi Mengandung A dan B}}{\sum \text{Transaksi}}$$

Here is a sample calculation of the formation of 2-itemsets:

$$\text{Support (GR-GS)} = \frac{11}{20} \times 100 = 55\%$$

$$\text{Support (GR-FU)} = \frac{7}{20} \times 100 = 35\%$$

$$\text{Support (GR-SH)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (GR-AD)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (GR-NX)} = \frac{20}{20} \times 100 = 100\%$$

$$\text{Support (GS-FU)} = \frac{1}{20} \times 100 = 5\%$$

$$\text{Support (GS-SH)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (GS-AD)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (GS-NX)} = \frac{6}{20} \times 100 = 30\%$$

$$\text{Support (FU-SH)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (FU-AD)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (FU-NX)} = \frac{8}{20} \times 100 = 40\%$$

$$\text{Support (SH-AD)} = \frac{0}{20} \times 100 = 0\%$$

$$\text{Support (SH-NX)} = \frac{1}{20} \times 100 = 5\%$$

$$\text{Support (AD-NX)} = \frac{1}{20} \times 100 = 5\%$$

Table 6
List of 2-itemsets

No.	Item code	amount	Support%
1	GR-GS	11	55%
2	GR-FU	7	35%
3	GR-SH	0	0%
4	GR-AD	0	0%
5	GR-NX	20	100%
6	GS-FU	1	5%
7	GS-SH	0	0%
8	GS-AD	0	0%
9	GS-NX	6	30%
10	FU-SH	0	0%
11	FU-AD	0	0%
12	FU-NX	8	40%
13	SH-AD	0	0%
14	SH-NX	1	5%
15	AD-NX	1	5%

From Table 6, the next will seek the support of $\geq 20\%$ minimum of 20 makadiperoleh sales transaction data that meets the minimum value of support:

Table 7
Category 2-itemsets

No.	Combination 2 Item	amount	value Support
1	GR-GS	11	55%
2	GR-NX	20	100%
3	GS-NX	6	30%
4	FU-NX	8	40%

Furthermore, to find association rules then do steps confidence value calculation of each item listed in Table 7, based on the following formula:

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

Here is a sample calculation of the value formation of Confidence:

$$\text{Confidence } P(GR \rightarrow GS) = 55\% \frac{11}{20} \times 100\% =$$

$$\text{Confidence } P(GR \rightarrow NX) = 100\% \frac{20}{20} \times 100\% = 100$$

$$\text{Confidence } P(GS \rightarrow NX) = 30\% \frac{6}{20} \times 100\% =$$

$$\text{Confidence } P(FU \rightarrow NX) = 40\% \frac{8}{20} \times 100\% =$$

Table 8.
2-itemsets With Confidence Value

No.	Combination 2 Item	amount	confidence
1	GR-GS	11	55%
2	GR-NX	22	100%
3	GS-NX	6	30%
4	FU-NX	8	40%

Furthermore, let's say the minimum value $\geq 50\%$ confidence, then aturanasosiasi formed are as follows:

Table 9
Association rules (Rule Asosiasi)

No.	Combination 2 Item	amount	Support	confidence
1	GR-GS	11	55%	55%
2	GR-NX	22	100%	100%

From the stages are carried out, the items that meet the support and confidence $\geq 50\%$, so that by the rules of the association which is formed from table 9, is. It can be concluded as follows:

- If you buy products Suzuki GSX-150R, it will buy products Suzuki GSX-150S. With support at 55% and 64% Confidence.
- If you buy products Suzuki GSX-150R, it will buy Suzuki Nex-II products. With the support of 110% and 129% Confidence.

4. Results and Discussion

From these results it obtained the discussion of a display program as beriku. Implementation of this sale transaction analysis applications using Visual Studio 2010 with the programming language used is Visual Basic .Net and Crystal Report.

a) Form Item

Form item serves to display data items of goods sold on PT. Sejahtera Motor Gemilang. In this form, the admin can add items by selecting the Add button, change the items by selecting the Change button and delete the data by selecting the Remove button.

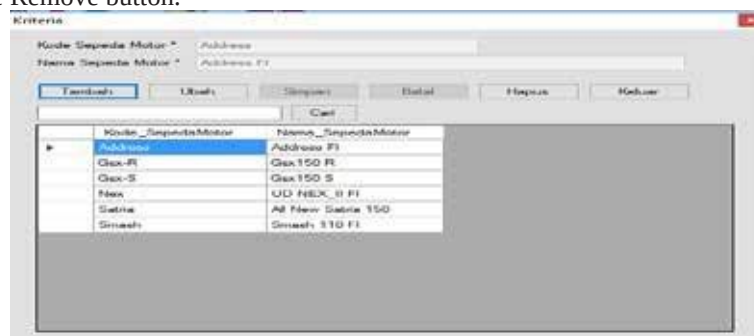


Figure 2. Form Item

b) Transaction Form

Transaction form display data on the sales transactionPT. Sejahtera Motor Gemilang. In this form, the admin can add transaction by choosing the Add button, change the items by selecting the Change button and delete the data by selecting the Remove button.

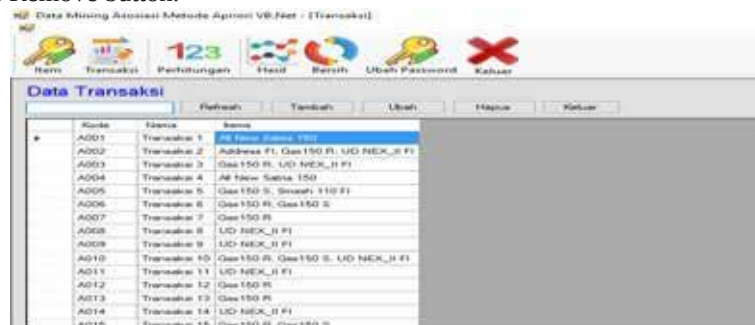


Figure 3. Transaction Form

c) Add Transaction Form

After admin selecting the plus button on the form of the transaction, the system will display the form added transactions. The transaction code will be generated automatically by the system. Admin and then enter the name of a transaction and then select the items that go into the sales transaction.

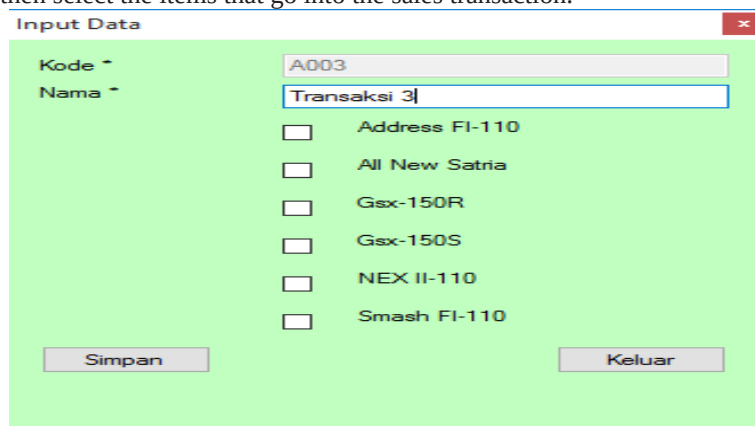


Figure 4. Add Transaction Form

d) Form Analysis

Form analysis is a form of analysis of sales transactions in PT. Sejahtera Motor Gemilang. To perform the analysis, the admin will include the value of Min Min Support and Confidence. After that, select the calculate button to perform the analysis with Apriori algorithm. To save the results of this analysis, it can choose the save button and print.

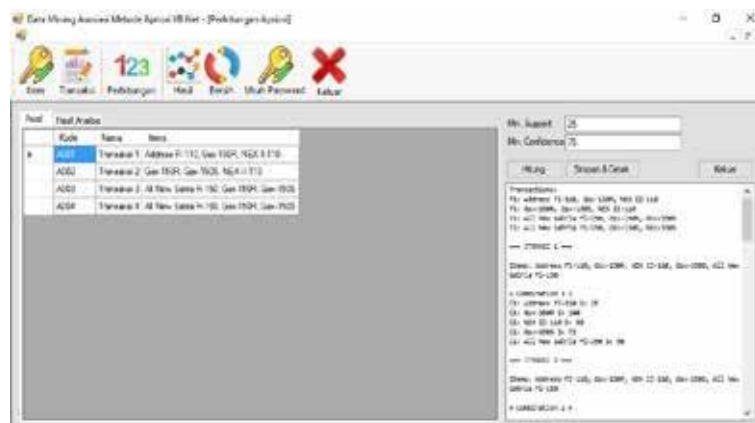


Figure 5. Analysis Form Sales Transactions

e) Form Result Analysis

After analyzing, the system will display the results of the analysis on the Analysis Results tab.



Figure 5. Form Result Analysis

f) Reports Analysis Results

The analysis results are displayed using Crystal Report to then be printed as a report on the analysis.

LAPORAN HASIL ASOSIASI				
Asosiasi	Support	Confidence	Supp*Conf	Lift Ratio
Jika membeli Gsx-R maka akan membeli Nex	100.00	50.00	25.00	1.00
Jika membeli Gsx-S maka akan membeli Satria 150	50.00	100.00	25.00	1.00
Jika membeli Gsx-R maka akan membeli Address	50.00	50.00	25.00	1.00
Jika membeli Gsx-R maka akan membeli Gsx-S	100.00	50.00	25.00	1.00
Jika membeli Nex maka akan membeli Gsx-R	100.00	50.00	25.00	1.00
Jika membeli Satria maka akan membeli Nex	50.00	50.00	100.00	1.00
Jika membeli Address maka akan membeli Nex	50.00	50.00	25.00	1.00
Jika membeli Smash maka akan membeli Satria	50.00	100.00	25.00	1.00

Figure 6. Reports Analysis Results

5. Conclusion

From the results of research by the author, the conclusions drawn and suggestions as follows:

- The existing data on Sejahtera Motor Gemilang can be decomposed properly, without any buildup again.
- Apriori algorithm is able to determine the Association Rule on sales of Suzuki motorcycles.
- Sejahtera Motor Gemilang greatly assisted by the Motorcycle Sales Application.
- It is expected that in future studies, decomposition motorcycle sales data can be done online.
- Apriori algorithm can find Association Rule on Suzuki's motorcycle sales, by calculating the value of Support and Confidence.
- Visual Basic 2010 application with MySQL storage is very helpful in creating perogram motorcycle sales.

6. Reference

- PM Hasugian, "Testing Algorithm Apriori By Application Weka In," J. Mantik Penusa, 2017.
- TR Vulandari, "Stages of Data Mining," in Data Mining, Theory and Applications Rapiirminer, 2017.
- SF Rodiyansyah, "Apriori Algorithm to Shopping Cart Analysis on Sales Transaction Data," *Infotech*, 2015.
- J. Ipmawati, Kusri, and E. Taufiq Lutfi, "Comparison of Text Mining Classification Techniques On Sentiment Analysis," *Indones. J. Netw. Secur.*, 2017.
- A. -, Marisa F., and D. Purnomo, "Application of Apriori Algorithm Against Data Warehouse Store Sales in BM," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, 2016.
- I. Kurnawan, Marisa F., and D. Purnomo, "Implementation of Data Mining With Apriori Algorithm For," *J. Teknol. and Manaj. Inform.*, 2018.
- I. Mubarak, Marisa F., and D. Purnomo, "apriori algorithm implementation for determining the satisfaction of students of service online university SIM Widayagama Malang," *J. Teknol. and Manaj. Inform.*, 2017.
- A. Nursikuwagus and T. Hartono, "Apriori algorithm implementation for web-based sales analysis by," *Symmetrical J. Tech. Mechanical, Electrical and Sciences Komput.*, 2016.
- SZ Ninuk Wiliani, "cashier ticket design applications watch the ball bareng X X cashier IN A location with visual basic 2010 and MYSQL," etc. Nusant. PGRI Kediri, 2017.
- A. Anthony, AR Tanaamah, and AF Wijaya, "Analysis and Design of Information Systems Warehouse Sales Based on Stock-Based Client Server (Case Study Grocery Store 'Restu you')," *J. Teknol. Inf. and Science Komput.*, 2017.