Data Mining of Data Processing Items using Apriori Algorithm in Selly Sport & Electronic Shop in Perbaungan

Muhammad Iqbal Syahputra Barus

Informatics Engineering Study Program, STMIK Pelita Nusantara, Medan

Article Info	ABSTRACT
	Data processing for the process of calculation or transformation of input data into information that is easy to understand. In addition, data processing is a process consisting of data storage and data handling activities.Data Mining is one of the fastest growing fields due do the huge need for added value from large-scale databases that are accumulating more and more as information grows. The general defenition of Data it self is not known manually from a data set. By showing the correlation of previously unknown data, the store owner can make the decision to progress the The Selly Sport & Electronic Perbaungan.Data Mining is used many places and fielda of application can also vary, data mining learn what are the main factors inthe accuracy of the target purchase of a product by consumers. Business intelligence is the process of converting data into information. Apriori Algorithm is one of the data mining algorithms in the formation of association of rule mining.Algorithm mining is the process of extracting information from a database, followed by doing frequent item/ itemset in formation of association rule mining in order to get the minimum value of support and minimum confidence value. *This is an open access article under the CC BY-SA license.*

Corresponding Author:

Muhammad Iqbal Syahputra Barus,
Informatics Engineering Study Program,
STMIK Pelita Nusantara, Medan,
Jl. St. Iskandar Muda No. 1 Medan.
Email: muhammadiqbalsyahputrabarus18@yahoo.com

1. INTRODUCTION

Data Mining is used in various places and fields and its application can be applied in various fields, in data mining things are studied which are the main factors in the accuracy of purchasing a product by consumers. Business intelligence is the process of converting data into information. From the collection of information available, the pattern will be taken into knowledge. Techniques in data mining are, classification, clustering, association rule, regression, forecasting, sequence analysis, and deviation analysis. Data processing is processing of data or a combination of various types of processing of data to make the data useful in accordance with the desired results can be used immediately.

Selly Sport & Electronic Perbaungan shop sells various items including electronic items, sporting items, children's toys, musical instruments and others. Selly Sport & Electronic Stores Perbaungan in terms of data processing, namely processing inventory data, processing transaction sales transaction data still using records using books to find out the number of sales items available, so that shop owners will experience difficulties in determining sales strategies, having difficulty knowing the goods many of which sell well in a certain period of time by knowing the support and confidence , because the owner must open his notebook to find out what items have been sold, have

not been sold, and of course this will require a long time to check the accounting records and this is one of the weaknesses of manually recording item data.

In addition, the owner also has difficulty in providing goods that are desired by the buyer because of the position of the goods that have not been grouped. Looking at the conditions that exist in Selly Sport & Electronic Stores Perbaungan in solving existing problems, it takes a computerized application.

2. RESEARCH METHOD

The stages in this research are described in the form of a diagram in order to understand each stage carried out. The stages of this research can be seen in Figure 1 below:
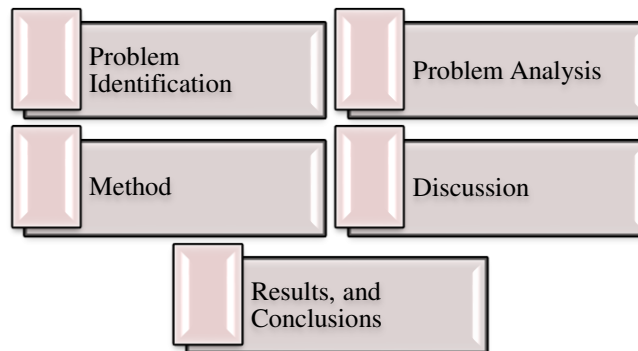

Figure 1. Diagram of Methods and Research Stages

2.1. Basic theory
A. Artificial Intelligence

Data mining is the process of getting useful information from a large database. Data mining can be interpreted as extracting new information taken from large data that helps in decision making. The term data mining is also called knowledge discovery [1].

Data mining is a process that employs one or more learning techniques to analyze and extract knowledge automatically. Knowledge Discovery in Databases (KDD) is the application of scientific methods to data mining [2].

The purpose of data mining techniques is to try to find benefits from a set of data. Viewed from the disciplines used, data mining is a multidisciplinary science that involves various disciplines such as databases, artificial intelligence, information science, high performance computing, visualization, machine learning, statistics, artificial neural networks, mathematical modeling, information retrieval and information extraction and recognition of a pattern. Currently data mining has also developed into one of the various concepts of other disciplines, such as web mining and text mining [3].

B. Apriori Algorithm

This algorithm controls the development of candidate itemsets from the results of frequent itemsets with support-based pruning to eliminate unattractive itemset y by setting minsup [4]. Apriori algorithms are also defined as a process for finding all a priori rules that qualify for support and minimum requirements for confidence [5].

This stage looks for item combinations that meet the minimum requirements of the support value in the database. The support value of an item is obtained by using the following formula : [6].

$$Support\ (A) = \frac{\sum transaction\ of\ A}{total\ transaction}\ x\ 100\%$$

To find a support value of two item are obtained using the formula :

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

$$Support\ (A, B) = \frac{\sum transaction\ of\ A\ and\ B}{\sum transaction}\ x\ 100\%$$

All the high frequency patterns were found, then a rule was found that fulfilled the minimum requirements for confidence by calculating the confidence in the associativeA☐B rules obtained by the following formula : [7].

$$Confidance = \frac{\sum transaction\ of\ A\ and\ B}{\sum transaction\ of\ A} x\ 100\ \%$$

To determine the association rules to be chosen, they must be sorted by Support x Confidence. Rules are taken as many as n rules that have the greatest results.

3. RESULTS AND DISCUSSION

The Apriori algorithm is one of the algorithms in data mining to look for frequent items / itemset on transactional databases. The Apriori algorithm is the basic algorithm proposed by Agrawal & Srikant in 1994 to determine freeguent itemsets for association rules for data mining.

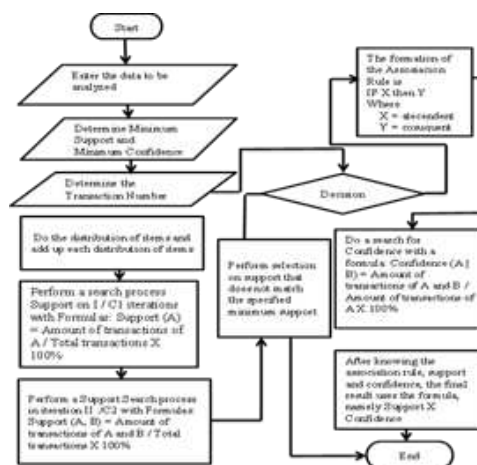To find out the way that is made with a priori algorithm can be seen in the flowchart below :

Figure 2. Flowchart about Algorithms Apriori

Data analysis is the process carried out in a study to solve problems related to a case that wants to be studied, in order to get a solution in solving problems. In research, things that are considered include, raw data, methods used in completing a case and the results of the analysis carried out. Sales transaction data for one item is carried out by a cleaning process to find out the number of items sold and know the value of support. The results of transaction data are one item to find out which items are sold and know the value of support.

Table 1. Data Analysis Transaction For One Item Support (Item / C1)

No.	Item name	Amount (∑)	Output
1	**Aquarium Tools**	2	0,3 %
2	**Basket Ball**	3	0,5 %
3	**Hannoch Lamp 5 Watt**	4	0,7 %
4	**Fishing Rod**	3	0,5 %
5	**Car Fan**	3	0,5 %
6	**Tube Lamp**	9	0,7 %
7	**Vacuum Cleaner**	4	0,7 %
8	**Guitar String**	18	3,4 %
9	**Head Flashlight**	5	0,9 %

The results of the cleaning process for these items can be seen in the table below :

Table 2. Data Analysis Results Transaction one item support (item / C1) that has been cleared.

No.	Item name	Amount (∑)	Output
1	**Basket Ball**	3	0,5 %
2	**Hannoch Lamp 5 Watt**	4	0,7 %
3	**Fishing Rod**	3	0,5 %
4	**Car Fan**	3	0,5 %
5	**Tube Lamp**	9	0,7 %

6	Vacuum Cleaner	4	0,7 %
7	Guitar String	18	3,4 %
8	Head Flashlight	5	0,9 %

The results of one item transaction data are then paired with other items so that it becomes two items to find out which items are sold and know the value of the support. The following is a table of goods data (items) that are used with total transactions = 523 to find out two items (items / itemset) with Threshold (Threshold) Min Support two items = 0.5%.

Table 3. Transaction Data Analysis Results two items support (item / C2).

No.	Item name	Amount (∑)	Output
1	Aquarium Tools, Traning Pants	2	0,3 %
2	Basket Ball, Fishing Rod	3	0,5 %
3	Car Fan, Tube Lamp	4	0,7 %
4	Hannoch Lamp 5 Watt, Head Flashlight	4	0,7 %
5	Vacuum Cleaner, Guitar String	4	0,7 %

The results of the cleaning process for these items can be seen in the table below:

Table 4. Data Analysis Results Transactions of two items support (item / C2) were cleared

No.	Item name	Amount (∑)	Output
1	Aquarium Tools, Traning Pants	2	0,3 %
2	Basket Ball, Fishing Rod	3	0,5 %
3	Car Fan, Tube Lamp	4	0,7 %
4	Hannoch Lamp 5 Watt, Head Flashlight	4	0,7 %
5	Vacuum Cleaner, Guitar String	4	0,7 %

Based on calculations made on the pair of item can be known the amount of confidence value. The following is an item data table that is used to find out items with the Min Threshold Confindace one item = 0.6 %.

Table 5. Transaction Data Analysis Results Sales item confidence

No.	Rule	Output
1	If you buy an Aquarium Tools, then buy a Training Pants	20 %
2	If you buy a Basket Ball, then buy a Fishing Rod	100 %
3	If you buy a Car Fan, then buy a Tube Lamp	100 %
4	If you buy a Hannoch Lamp 5 Watt, then buy a Head Flashlight	100 %
5	If you buy a Vacuum Cleaner, then buy a Guitar String	100 %

Based on the pair of items adjusted to the rule. The results of calculating the pair of items can be seen in the table below :

Table 6. Results of Analysis of Sales Transaction Confidence items the cleaning.

No.	Rule	Output
1	If you buy a Basket Ball, then buy a Fishing Rod	100 %
2	If you buy a Car Fan, then buy a Tube Lamp	100 %
3	If you buy a Hannoch Lamp 5 Watt, then buy a Head Flashlight	100 %
4	If you buy a Vacuum Cleaner, then buy a Guitar String	100 %

Based on the calculation of the support multiplied by confidence, the association rule can be known.

Table 7. Item Sales Transaction Data Analysis Results, Support x Cleansing confidence.

No.	Rule	Output
1	If you buy a Basket Ball, then buy a Fishing Rod	0,005
2	If you buy a Car Fan, then buy a Tube Lamp	0,007

| 3 | **If you buy a Hannoch Lamp 5 Watt, then buy a Head Flashlight** | **0,007** |
| 4 | **If you buy a Vacuum Cleaner, then buy a Guitar String** | **0,007** |

Based on the results of calculation of support and confidence, it can be made a graph of the final results of the analysis of item sales transaction data.
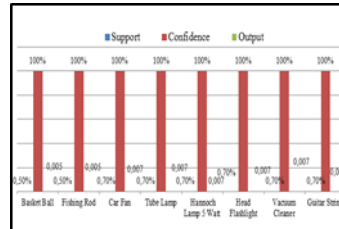

Figure 3. Graph of Final Results Data Analysis Sales Transaction item

4. CONCLUSION

After the discussion in the previous chapters, then in this chapter the researcher will present conclusions and suggestions that can be taken from the descriptions in the previous chapters. The items used in this study consisted of 103 items. The Data Mining process is carried out with cleaning and determines the minimum 0.5% support only up to two itemset. After determining confidence with minimum confidence 60% From the support and confidence, the four Rule results are obtained, name: If you buy Basketball then buy a Fishing Rod, with a result of 0.005, If you buy Car Fan then buy TL Lights, with a result of 0.007, If you buy Hannoch five Watt then buy a Head Flashlight, with a result of 0.007, If you buy Vacuum Cleaner, buy Guitar String ,with a result of 0.007. The owner can find out the visitor's interest in buying an item, so that the owner can determine the number of items that must be stocked and the items that are not stocked (based on the conclusion of number four). To adjust the location of items in the display rack can be determined the position of the item (based on the conclusion of number 4).

REFERENCES

[1] E. Prasetyo, *Data mining : konsep dan aplikasi menggunakan MATLAB*, 1st ed. Yoyakarta: CV ANDI Offset, 2012.

[2] F. A. Hermawati, *Data Mining*, I. Yoyakarta: CV ANDI Offset, 2013.

[3] R. YOTENKA and P. . Prof. Drs. Subanar, "ASSOCIATION RULE MINING DENGAN PENERAPAN ALGORITMA APRIORI DALAM PENENTUAN POLA HUBUNGAN ANTAR RAWI HADIS," *Repos. UGM*, 2015, [Online]. Available: http://etd.repository.ugm.ac.id/penelitian/detail/80186.

[4] N. Wandi, R. A. Hendrawan, and A. Mukhlason, "Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalian Association Rule Menggunakan Algoritma Apriori," *J. Tek. ITS*, vol. 1, pp. 1–5, 2012.

[5] M. A. Syaifullah, "Implementasi Data\Mining Algoritma Apriori Pada Sistem Penjualan," 2010, [Online]. Available: http://vokasi.uho.ac.id/statistika/assets/download/151212070630algortima apriori jurnal.pdf.

[6] Kusrini and E. T. Luthfi, *Algoritma Data Mining*, I. Yoyakarta: CV ANDI Offset, 2009.