

Analysis of Apriori Algorithm in Mining Drug Sales Data at Ridos Hospital

E S Ompusunggu, S Efendi, P H Putra

Abstrac: This study will apply the Apriori Algorithm in the field of health and the determination of drug purchase patterns. Data mining processing to find out what drugs are purchased by consumers, can be done with analytical techniques from consumer buying habits. The results of this study are: In this study the authors apply a priori algorithm using transaction data 20 on the 8th month transaction, resulting in 80% support value and confident 50% value. The result of applying the a priori method with a minimum support of 50% using 20 transaction data is if you buy abbotic medicine 500 mg, and buy a 14 terumo abocath, then buy astemizole. The application of a priori algorithm is based on the calculation of support and confidence values. in the process of calculating the support and confidence values it will be more difficult, if the data you want to process is large. The a priori method used is quite effective in providing the final combination of drugs that are often purchased by consumers. The level of accuracy of testing using a priori method is 96%.

I. INTRODUCTION

This research will apply the Apriori Algorithm in the field of health and determination of drug purchase patterns. Data mining processing to find out what drugs are purchased by consumers, can be done with analytical techniques from consumer buying habits. Detection of drugs that are often purchased simultaneously is done using the Association Rule, in which data items are taken from a relational database. The process uses the Apriori Algorithm, which functions to form candidate combination items, then tested whether the combination meets the minimum support parameters and minimum confidence which is the threshold value given by the user.

Data analysis is done by creating an application that helps the Ridos General Hospital Pharmacy obtain knowledge in the form of drug sales transaction patterns within a particular month period.

II. RELATED RESEARCH

J Du, et al, in his research proposed DC_Apriori which was enhanced to analyze the shortcomings of a priori algorithms. With algorithms can optimize the set connection of items, and can reduce the number of operating connections [1].

Revised Manuscript Received on December 22, 2018

E S Ompusunggu, Student, Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

S Efendi, Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

P H Putra Student, Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

K R Thakre, et al, in implementing a system with the aim of extracting patterns from most datasets, for the purpose of knowledge achieved. A priori algorithms utilize less memory, less time calculation and less CPU usage when mining datasets [2].

Ali Akbar, et al, in this study proved that from execution time and iteration perspective is far superior, a priori algorithm can be used to reduce costs [3].

III. PROPOSED METHOD

A. Data Mining

Data mining is the process of looking for a pattern or information using certain techniques or methods. Data mining uses statistical techniques, mathematics, artificial intelligence, and machine learning, to extract and identify useful information and related knowledge of large data or databases [4].

B. Algorithm Apriori

Apriori algorithm is a highly popular pattern-finding algorithm in data mining technique. This algorithm is aimed at finding an itemset combination which has a certain value of frequency according to the desired criteria or filter. This algorithm is proposed by R. Agrawal and R. Srikant. The result of apriori algorithm can be used to help the management make decisions [3].

Apriori algorithm uses a simple data structure, and the execution process is clear, but when the dimension of affairs is large and the minimum support is small, it will reduced the execution efficiency greatly [5].

IV. RESULTS AND DISCUSSION

In analyzing the needs by using a priori algorithm. High frequency patterns can be known about the types of drugs most often purchased by consumers.

The following is the completion of the C3 formation process or referred to as 3 itemset with a minimum number of support = 50% with the following formula:

$$\text{Support (A,B)} = \frac{\sum \text{transactions contain A, B, and C}}{\text{transactions}} \dots (1)$$

$$\text{Confidence (A,B)} = \frac{\sum \text{transactions contain A, B, dan C}}{\text{transactions}} \dots (2)$$



Table1. Tabel Item Support

Transactions Data	Support
Abbotic 500 Mg	$(14/20) \times 100\% = 70.00\%$
Abocath 14 Terumo	$(16/20) \times 100\% = 80.00\%$
Adalat Oros 30 Mg	$(3/20) \times 100\% = 15.00\%$
Aesculap No.23	$(2/20) \times 100\% = 10.00\%$
Alinamin F 25 Mg / 10 MI	$(1/20) \times 100\% = 5.00\%$
Alprazolam 1 Mg	$(1/20) \times 100\% = 5.00\%$
Ambroxol 30 Mg	$(3/20) \times 100\% = 15.00\%$
Atemizole	$(11/20) \times 100\% = 55.00\%$
Betametason	$(2/20) \times 100\% = 10.00\%$
Bromhexin	$(5/20) \times 100\% = 25.00\%$
Desoksimetason	$(2/20) \times 100\% = 10.00\%$
Diponium	$(1/20) \times 100\% = 5.00\%$
Glafenin	$(1/20) \times 100\% = 5.00\%$
Karbosistein	$(1/20) \times 100\% = 5.00\%$

Alinamin F 25 Mg / 10 MI, Diponium	100 = 5 %
---------------------------------------	-----------

The combination of 3 and 4 itemset can meet a minimum of 50% support then the next is the formation of an association. After all the high frequency patterns have been found, then the rules of the association that satisfy for confidence are sought by calculating the associative confidence rules A - B.

Based on Table 2. the drugs most often purchased by consumers are Abbotic 500 Mg, Abocath 14 Terumo, Astemizole so that the drug is the best-selling drug bought by the consumer with the medicine that is most often purchased by consumers, the pharmacy can strategize in determining the purchase of drugs to maintain availability drugs needed by consumers and can also arrange the layout of drugs based on a combination of drug itemset formed.

V. CONCLUSION

In this study the author applies a priori algorithm using transaction data 20 on the 8th month transaction, resulting in 80% support value and confident 50% value. The result of applying the a priori method with a minimum support of 50% using 20 transaction data is if you buy abbotic medicine 500 mg, and buy a 14 terumo abocath, then buy astemizole. The application of a priori algorithm is based on the calculation of support and confidence values. in the process of calculating the support and confidence values it will be more difficult, if the data you want to process is large. The a priori method used is quite effective in providing the final combination of drugs that are often purchased by consumers. The level of accuracy of testing using a priori method is 96%.

Tabel2. Associatonrules table

Nama Item	Jumlah	Confidant
Abbotic 500 Mg, Abocath 14 Terumo, Alprazolam 1 Mg	1	$(1/20) \times 100 = 5\%$
Abbotic 500 Mg, Abocath 14 Terumo, Astemizole	10	$(10/20) \times 100 = 50\%$
Abbotic 500 Mg, Abocath 14 Terumo, Astemizole, Betametason	3	$(3/20) \times 100 = 15\%$
Abbotic 500 Mg, Aesculap No.23, Ambroxol 30 Mg, Betametason	1	$(1/20) \times 100 = 5\%$
Abbotic 500 Mg, Bromhexin, Desoksimetason	2	$(2/20) \times 100 = 10\%$
Abocath 14 Terumo, Abocath 14 Terumo, Aalat Oros 30 Mg, Karbosisten	1	$(1/20) \times 100 = 5\%$
Abocath 14 Terumo, Adalat Oros 30 g, Ambroxol	1	$(1/20) \times 100 = 5\%$
Abocath 14 Terumo, Astemizole, Betametason	4	$(4/20) \times 100 = 20\%$
Abocath 14 Terumo, Bromhexin, Glafenin	1	$(1/20) \times 100 = 5\%$
Adalat Oros 30 Mg, Ambroxol 30 Mg, Bromhexin, Bromhexin	1	$(1/20) \times 100 = 5\%$
Aesculap No. 23,	1	$(1/20) \times$

REFERENCES

1. Jiaoling Du, et al.: *Research and improvement of apriori algorithm*, Sixth international conference on information science and technology, 2016.
2. Kajal R. Thakre, et al.: *Implementation on an approach for mining of datasets using apriori hybrid algorithm*, App. international conference on trends in electronics and informatics, 2017.
3. Ali akbar, et al.: *Shortest-path problem solving in the installation of data/ internet network using apriori algorithm*.. 2017.
4. Yuli Mardi : *Data mining : klasifikasi menggunakan algoritma c4.5*.. Jurnal edik informatika, 2017.
5. Shaosong Yang, et al : *The parallel improved apriori algorithm research based on spark*.. International conference on frontier of computer science and technology, 2015.



