The 2nd International Conference on Inovations
in Social Sciences Education and Engineering (ICoISSEE)
August 07th, 2021

Pattern Analysis of Health Equipment Procurement System Using the FP-Growth Algorithm

Zulham[1*], **Ibnu Rusydi**[2], **Ananda H Elyas**[3]

[1,2,3]*Software Engineering / Faculty of Engineering and Computer Science,Dharmawangsa University, Medan*
zulham@dharmawangsa.ac.id; ibnurusydi@dharmawangsa.ac.id;
nanda@dharmawangsa.ac.id
**Corresponding author: zulham@dharmawangsa.ac.id*

Abstract. *The development of today's technology world requires researchers to find efficient and effective procurement strategies. The strategy is carried out to improve accuracy in the procurement process, namely Data Mining. The technique used in this strategy is the FP-Growth Algorithm. FP-Gowth is a technique that can determine the set of data that most often appears (Frequent Itemset) in a data set. FP-Growth is a development of the Apriori technique which uses the concept of tree development in searching for Frequent Itemset. This study observes several variables, namely, Procurement Number, Period, Name of Medical Devices, Category, Number, Unit. The results of this study can later be used for institutional needs in the procurement process. Besides that, software is produced to support the performance of the institution so that results can be more effective.*

Keywords: *Data Mining, Association Rule, Frequent Itemset, FP-Growth*

1. INTRODUCTION

Current technological developments that are increasingly advanced in all fields require humans to find a strategy that can improve performance and quality in any case. In a government that is in direct contact with the community is demanded to be able to work more efficiently so that the results achieved can be directly felt by the community.

At present the Medan City Health Office, which is a distribution center for the needs of medical equipment needed by all Puskesmas (Community Health Centers) in the region. For this reason, the distribution and procurement of medical equipment greatly affect the community services so that the results achieved can be maximized in accordance with community expectations.

To maximize the procurement of medical devices so that there is no shortage of needs, an analysis needs to be carried out so that the problem does not occur.

Analysis of the procurement of medical devices uses the FP-Growth Algorithm, which has the ability to mine frequent Patterns (patterns that often appear in data sets) with efficiency. Based on the background above, a study was conducted under the title "Analysis of Patterns of Procurement of Medical Devices with Algorithms of FP-Growth (Case Study of Medan City Health Office)".

2. LITERATURE REVIEW

2.1 Knowledge Discovery in Database (KDD)

Fayyad and Dunham in (Azhari & Anshori, 2009) said Knowledge Discovery in Databases (KDD) to refer to the whole process of finding knowledge in large data sets.

2.2 Data Mining

According to Turban in (Gunadi and Sensuse, 2012) Data Mining is a term used to describe the discovery of knowledge in a database. Data Mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from various large databases.

According to Azhari & Anshori, (2009) Data Mining is the exploration and analysis of a number of data in order to find patterns and rules that are very important.

According to Siregar, (2014) Data Mining is a decision support process where we look for patterns of information in data.

2.3 Association Rule

According to Wandi, et al (2012) Association rule is one method that aims to find patterns that often appear among many transactions, where each transaction consists of several items, so this method will support the book lending recommendation system that will be borrowed by users through the discovery of patterns between items in transactions that occur in Bapersip

2.4 FP-Growth Algorithm

Gunadi (2012) said FP-Growth is one alternative algorithm that can be used to determine the set of data that most often appears (frequent itemset) in a data set.
According to Ruldeviani, et al (2008) the FP-Growth algorithm represents transactions using the FP-Tree data structure.
Melaini, (2015) FP-Growth Algorithm is the development of Apriori algorithm.
The characteristic of FP-Growth algorithm is the data structure used is a tree called FP-Tree. By using FP-Tree, the FP-Growth algorithm can directly extract frequent Itemset from FP-Tree.
Extracting frequent itemset using the FP-Growth algorithm will be done by generating a data tree structure or called FP-Tree. The FP-Growth method can be divided into 3 main stages:
1. The generation phase of the conditional pattern base,
2. The FP-Tree conditional generation stage, and
3. The frequent itemset search stage (Meilani, et al 2015).

2.4.1 FP-Tree

According to Melaini, (2015) FP-Tree is a data storage structure that is utilized. FP-Tree is built by mapping each transaction data into each particular path in the FP-Tree. Because in each mapped transaction, there may be transactions that have the same item, so the trajectories are possible to overwrite each other. The more transaction data that has the same item, the more effective the compression process with the FP-Tree data structure.
The FP-Tree is a tree with the following definitions:
1. FP-Tree is formed by a root labeled null, a set of sub-trees consisting of certain items, and a frequent header table.
2. Each node in the FP-Tree contains three important information, namely item labels, informing the type of item represented by the node, support count, representing the number of transaction paths that pass through the node, and connecting pointers that connect nodes with the same item label between nodes. -trajectory, marked with a dotted arrow.

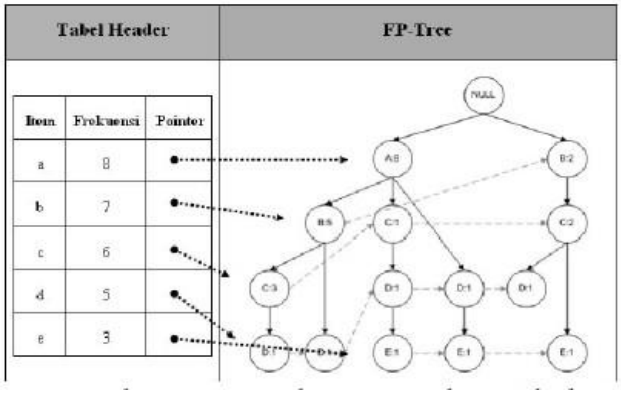

Figure 1. Example of FP-Tree and Header Table

3. RESEARCH METHODOLOGY

Before conducting research, a scientific method or appropriate and established rules are needed. It is intended that the results obtained from the study get maximum results. This research methodology contains the research framework to be carried out.

The research methodology will be used to identify the problems found, then make an analysis of these problems so that they will get the best solution to the problem or problem solving. In this study, it is necessary to solve problems in procuring medical devices which will later be distributed to Puskesmas in Medan City, so that the distribution of these medical devices can be more effective and on target than before.

3.1 Framework

In the research methodology there is a sequence of frameworks that must be followed, the sequence of this framework is a description of the steps that must be passed so that this research can run well and as expected. The framework to be followed can be seen in Figure 2.

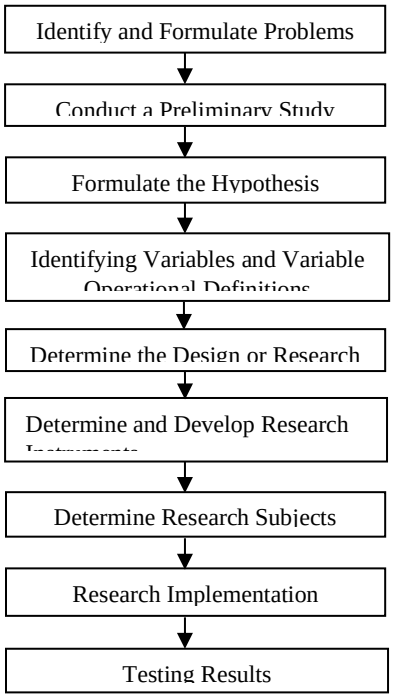

Figure 2. Framework

Based on Figure 2 above, the order of the framework is described as follows:
1. Identify and Formulate Problems.
As in the scientific method, scientific research must also depart from the existence of a problem to be solved. Before carrying out scientific research, it is necessary to identify the problem. The problem identification process is important so that the formulation of the problem becomes sharp and as a form of initial data that in scientific research it is necessary to solve problems through research. Problem identification is formulated according to the background of the problem, based on facts and data in the field. Problem identification is generally formulated in the form of declarative sentences, while the problem formulation is written in the form of interrogative sentences (in the form of questions).
2. Conduct a Preliminary Study.
In scientific research, it is necessary to conduct a preliminary study. Researchers can do this by tracing and understanding literature reviews for the theoretical foundation materials needed to formulate hypotheses and discuss research results later. A research is said to be good if it is based on a solid and relevant theoretical foundation. Many theories are in accordance with research, but they are not relevant. Therefore, it is necessary to make an effort to sort out the appropriate theory. In addition, a preliminary study conducted by researchers through literature review will be able to make research more focused on the problem under study so that it can facilitate the determination of what data will be needed later.
3. Formulate the Hypothesis.
Hypotheses need to be formulated in a scientific study, especially quantitative research. By stating the hypothesis, the scientific research conducted by the researcher will focus more on the issues raised. In addition to the formulation of the hypothesis, a researcher no longer needs to be bothered with data that he should not need, because the data he takes through research instruments is only data that is directly related to the hypothesis.
4. Identifying Variables and Variable Operational Definitions.
A variable in scientific research is a phenomenon that will or will not occur as a result of another phenomenon. Research variables really need to be determined so that the problems raised in a scientific research become clear and measurable.
5. Determine the Design or Research Design.
Research design is often referred to as research design. Research design is a procedure or applicative research steps that are useful as a guide in carrying out scientific research for the researcher concerned.
6. Determine and Develop Research Instruments.
What is a research instrument? Research instruments are tools used by researchers to collect the data they need. Various data collection tools and techniques can be selected according to the purpose and type of scientific research carried out. Each form and type of research instrument has its own advantages and disadvantages.
7. Determine Research Subjects.
People who are involved in scientific research and act as sources of data are called research subjects. Often the research subject is related to the research population and sample. If scientific research is carried out using a research sample in a research population, the researcher must be careful in determining it. This is because, research that uses samples as research subjects will conclude research results that are generally accepted for the entire population, even though the data taken is only a sample that is much smaller than the research population.
8. Research Implementation.
At this stage, the results of the research will be implemented so that the results can be

tested to find out whether the results are in accordance with what is expected.
9. Testing Results.
At this stage, the rule is re-tested or tested again using an existing data mining system. The tools used as system testing is Rapidminer.

4. RESULTS AND DISCUSSION

4.1 Data Analysis

The FP-Growth algorithm is an algorithm that can be used to analyze data structures using a characteristic called FP-Tree. FP-Tree is built by mapping each transaction data into each particular path in the FP-Tree. FP-growth algorithm can directly extract frequent itemset from by using FP-Tree. Data mining in the form of frequent itemset using the FP-Growth algorithm will be carried out by generating a tree data structure or called FP-Tree. The FP-Growth method can be divided into 3 main stages, namely:

1. The stage of generating the conditional pattern base,
2. FP-Tree conditional generation stage, and
3. The frequent itemset search stage.

These three stages are steps that will be taken to get frequent itemset. Gunadi and Senses (2014).

Input : FP-Tree Tree
Output : Rt Complete set of frequent . patterns
Method : FP-Growth (Tree, null)
Procedure : FP-Growth (Tree,)
{
01: if Tree contains single path P;
02: then for each combination (denoted) of the nodes in the do path path
03: generate pattern with support from nodes in path do ;
04: else for each a1 in the header of the do tree
}
05: generate pattern
06: wake up = a1 with support = a1 support
07: if Tree $\beta =$

4.2 Application of Data Mining

Data Mining can be done with many applications. Besides, the application is combined with several fields of science such as artificial intelligence, databases, statistics, mathematical modeling, image processing, and so on. So with this combination the application of data mining is becoming more widespread, one of which is in terms of procurement within government institutions.

4.2.1 Variable Selection

The sample data tested in this study were 1 period of data on the procurement of medical equipment at the Medan City Health Office, the format is as follows:

1. Procurement Number
2. Period
3. Name of Medical Device
4. Categories
5. Quantity
6. Unit

As for what is taken as a variable in the association, namely the Procurement Number, Name of Medical Device, Category, Amount, Unit and these four variables will later be used as the basis for forming the FP-Growth algorithm which is divided into three main steps, but first the formation of a tree using FP-Tree algorithm. Frequent Pattern Tree (FP-Tree) is a compressed representation of data entry.

Table 1. Frequent Items

Items	*Frequent*	*Support*	
A	26	26/39	66,6%
B	25	25/39	64,1%
C	10	10/39	25,6%
D	13	13/39	33,3%
E	7	7/39	17,9%
F	10	10/39	25,6%
G	12	12/39	30,7%
H	17	17/39	43,5%
I	16	16/39	41%
J	15	15/39	38,4%
K	14	14/39	35,8%
L	8	8/39	20,5%
M	11	11/39	28,2%
N	13	13/39	33,3%
O	21	21/39	53,8%

To determine the minimum frequency itemset, the researcher set 35%. Therefore the minimum support in this study is 35%, so the frequency that appears most often is taken above the minimum itemset frequency value of 35%, namely: items A, B, O, H, I, J, and K. determining the minimum support is not used to build FP-Tree. These items will have an effect and will be entered into the FP-tree, the rest of the seven variables can be discarded because they have no significant effect.

Furthermore, data filtering is carried out on the sample data used. Filtering is done by removing itemsets other than the selected itemset.

Then the next step is to form an FP-Tree The image below provides an illustration of the formation of an FP-tree after reading the TID.

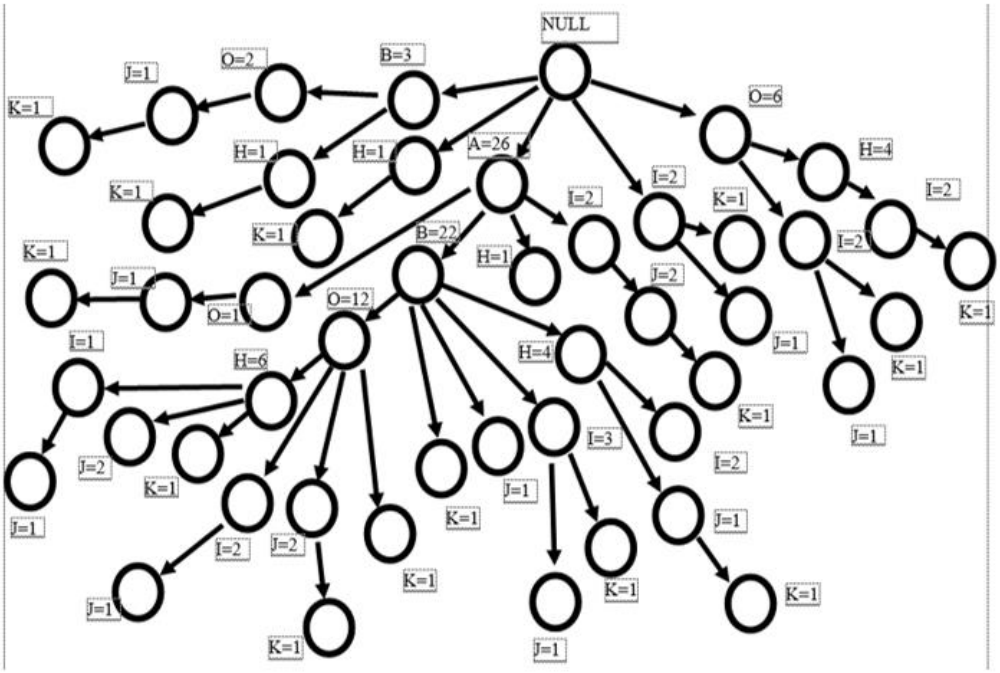

Figure 3. Result of FP-tree Formation After TID 39

Another parameter is confidence, which is a measure of how much validity an Association Rules are. An Association Rules with confidence equal to or greater than the minimum confidence can be said to be a valid association rule. From the results of the minimum support above the chosen one, the next step is to determine the

confidence value. The minimum confidence value in this study is set at 50%.

Table 2. Selected Minimum Support Itemset with Confidence value

No	*Itemset*	Qty	*Support*	*Confidence*
1	A,B	22	56 %	(22/26) x 100% = 84 %
2	B,A	22	56%	(22/25) x 100% = 88 %
3	A,O	13	33 %	(13/26) x 100% = 50 %
4	B,O	14	35 %	(14/25) x 100% = 56 %
5	J,A	12	30%	(12/15) x 100% = 80 %
6	A,J	12	30%	(12/26) x 100% = 46 %
7	O,A	13	33 %	(13/21) x 100% = 61 %
8	O,B	14	35 %	(14/21) x 100% = 66 %
9	A,B,O	12	30 %	(12/22) x 100% = 54 %
10	A,O,B	12	30%	(12/13) x 100% = 92 %
11	B,O,A	12	30%	(12/14) x 100% = 85 %
12	B-O,A	12	30%	(12/25) x 100% = 48 %
13	O-A,B	12	30%	(12/21) x 100% = 57 %
14	B-A,O	12	30%	(12/25) x 100% = 48 %

Table 1. Selected Minimum Support Itemset with Confidence value

Furthermore, with a minimum confidence value of 50%, the association rules formed are as follows:

Table 3. Association Rule

No	*Itemset*	Qty	*Support*	*Confidence*
1	A,B	22	56 %	84 %
2	B,A	22	56%	88 %
3	A,O	13	33 %	50 %
4	B,O	14	35 %	56 %
5	J,A	12	30%	80 %
6	O,A	13	33 %	61 %
7	O,B	14	35 %	66 %
8	A,B,O	12	30 %	54 %
9	A,O,B	12	30%	92 %
10	B,O,A	12	30%	85 %
11	O-A,B	12	30%	57 %

From the steps that have been carried out above, the items that meet the largest Support x Confidence and minimum confidence = 50% in Table 4.12, so that based on the association rules formed, the following conclusions can be drawn:

Rule 1: If you buy Adult Stethoscope (A), then you will buy Children's Stethoscope (B) with 56% support and 84% confidence.

Rule 2: If you buy a Children's Stethoscope (B), then you will buy an Adult Stethoscope (A) with 56% support and 88% confidence.

Rule 3: If you buy Adult Stethoscope Medical Devices (A), then you will buy X-Ray Double (O) Viewer Medical Devices with 33% support and 50% confidence.

Rule 4: If you buy a Children's Stethoscope (B), then you will buy an X-Ray Double (O) Viewer Medical Device with 35% support and 56% confidence.

Rule 5: If you buy a Magnifying Glass Medical Device (J), you will buy an Adult Stethoscope Medical Device (A) with 30% support and 80% confidence.

Rule 6: If you buy X-Ray Double (O) Viewer Medical Devices, you will buy Adult Stethoscope Medical Devices (A) with 33% support and 61% confidence.

Rule 7 : If you buy Medical Devices Viewer X-Ray Double (O), then you will buy Medical Devices Stethoscope for Children (B) with 35% support and 66% confidence.

Rule 8: If you buy Adult Stethoscope (A) and Children's Stethoscope (B), it will be followed by buying X-Ray Double (O) Viewer Medical Devices with 30% support and 54% confidence.

Rule 9: If you buy Adult Stethoscope Medical Devices (A) and X-Ray Double (O) Viewer Medical Devices, you will be followed by buying Children's Stethoscope Medical Devices (B) with 30% support and 92% confidence.

Rule 10: If you buy Children's Stethoscope (B) and Viewer X-Ray Double (O) Medical Device, you will buy Adult Stethoscope Medical Device (A) with 30% support and 85% confidence.

Rule 11: If you buy X-Ray Double (O) Viewer Medical Devices, you will buy Adult Stethoscope Medical Devices (A) and Children's Stethoscope Medical Devices (B) with 30% support and 57% confidence.

CONCLUSION

From the results of the discussion, the FP-Growth algorithm can determine the relationship between 1 item of medical device and other medical devices (Rule Association). This is very helpful for institutions to be able to work more effectively and efficiently.

In addition, the results obtained from this study can be used as samples for other institutions so that in the future each purchase can be adjusted to the results in the FP-Growth algorithm.

REFERENCES

Gunadi & Sensuse (2012), Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Metode Algoritma Apriori dan Frequent Pattern Growth (FP-Growth). Vol .4, No.1.

Azhari & Anshori (2009), Pendekatan Aturan Asosiasi Untuk Analisis Pergerakan Saham, ISSN: 1979-2328.

Meilani & Azinar (2015), Penentuan Pola Yang Sering Muncul Untuk Penerima Kartu Jaminan Kesehatan Masyarakat (Jamkesmas) Menggunakan Metode FP-Growth, ISSN: 2089-1121.

Hermawati (2009), Dalam Buku “Data Mining”.

Triyanto (2014), Association Rule Mining Untuk Penentuan Rekomendasi Promosi Produk, Vol 5 No 2.

Siregar (2014), Implementasi Data Mining Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori, Vol VII, Nomor :1.

Ruldeviyani & Fahrian (2008), Implementasi Algoritma-algoritma Association Rules Sebagai Bagian Dari Pengembangan Data Mining Algoritms Collection, KSN&I08-043

A. Mehay, K. Singh, and N. Sharma, “Analyze Market Basket Data using FP-growth and Apriori Algorithm,” Int. J. Recent Innov. Trends Comput. Commun., vol. 1, no. 9, pp. 693–696, 2013.

N. Padhy, P. Mishra, and R. Panigrahi, “The Survey of Data Mining Applications And Feature Scope Keywords Data mining task, Data mining life cycle , Visualization of the data mining model , Data mining Methods, Data mining applications,” Int. J. Comput. Sci. Eng. Inf. Technol., vol. 2, no. 3, pp. 43–58, 2012, doi: 10.5121/ijcseit.2012.2303.

S. R. Siregar, “IMPLEMENTASI DATA MINING PADA PENJUALAN TIKET PESAWAT MENGGUNAKAN ALGORITMA APRIORI (Studi Kasus : Jumbo Travel Medan),” vol. 152, pp. 152–156, 2014.

W. Aprianti, J. Permadi, and Oktaviyani, “Penerapan Algoritma Apriori untuk Transaksi Penjualan Obat pada Apotek Azka,” Semin. Nas. Mat. dan Apl., no. February, pp. 436–442, 2017.

Azhari and Anshori, “Pendekatan aturan asosiasi untuk analisis pergerakan saham,” Knowl. Creat. Diffus. Util., vol. 2009, no. semnasIF, pp. 183–189, 2009.

P. Priti and S. Jikitsha, “Improved Apriori Algorithm using Bottom – up Approach,” vol. 4, no. 1, pp. 153–156, 2015.

Zulham and A. I. Ulya, “Analisis Pola sistem pengadaan alat kesehatan denagn algoritma fp-

growth," J. Manaj. Inform. dan Tek. Komput., vol. 3, no. April, pp. 40–45, 2018.
V. Mangla, C. Sarda, and T. Nadu, "Improving the efficiency of Apriori Algorithm in Data Mining," vol. 3, no. 3, pp. 393–396, 2013.
A. Saad and A. Alghamdi, "Efficient Implementation of FP Growth Algorithm-Data Mining on Medical Data," vol. 11, no. 12, pp. 7–16, 2011.
A. Singh, J. Agarwal, and A. Rana, "Performance Measure of Similis and FP-Growth Algorithm," Int. J. Comput. Appl., vol. 62, no. 6, pp. 25–31, 2013, doi: 10.5120/10085-4712.
R. K. Soni, P. N. Gupta, and P. A. Sinhal, "An FP-Growth Approach to Mining Association Rules," Int. J. Comput. Sci. Mob. Comput., vol. 2, no. February, pp. 1–5, 2013.