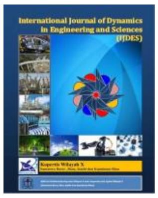
INTERNATIONAL JOURNAL of DYNAMICS in ENGINEERING and SCIENCES (IJDES)
LLDIKTI WILAYAH X
Vol. 8, No. 2, Oct 2023
e-ISSN: 2502-0692



OPTIMIZING IT BUSINESS MANAGEMENT USING THE FP-GROWTH ALGORITHM

Ritna Wahyuni[1]
[1] Institut Teknologi Sawit Indonesia

doi.10.22216/ jod.v8i2.2449
*Correspondence should be addressed to ritnawahyuni@itsi.ac.id

Article Information

Submitted : 12 Mei 2023

Accepted : 20 Sept 2023

Published : 01 Oct 2023

Abstract

Expert Computer Store is a store that sells various computer equipment and computer equipment. The problem faced by shop owners is that the bookkeeping of inventory and sales is still done manually, and there is no complete sales data analysis related to the accumulation of goods due to lack of sales and empty goods. Because the business process is carried out manually, the shop owner has difficulty in analyzing what items sell quickly and what items are bought together. This causes store owners to be unable to meet market demand and has an impact on decreasing buyer interest and sales levels. This study uses 23 transaction data of computer equipment sales products as samples. The algorithm used to help stores find customer buying patterns is the FP-Growth Algorithm, so that it can help business owners in managing the number of orders in order to meet market demand. The result of this research is a decision in the procurement of goods at the Expert Computer Store. which can be seen from the frequency pattern of purchasing building materials made by customers.

Keywords: Data mining, FP-Growth, Association, it Bussines Management

1. Introduction

The rapid development of digital technology is characterized by increased access to information. This development is always accompanied by the emergence of highly sophisticated technology and advanced innovation. The benefits of computers are very broad, ranging from the legal field of education, and the economic field, especially trade [1].

The need for quality information is a major factor in designing the use of information technology. The presence of business management information technology can reduce the error rate when revenue and followed by increased business output.

Data mining is a process that includes statistical, mathematical, artificial intelligence, and machine learning techniques that can identify and generate hidden information from data sets. And data mining is a method of using special techniques to find hidden information and patterns in data[2].

Inventory is a pile of goods that are deliberately stored as assets or stockpiles given the scarcity of these goods. As inventory management management is something that needs special attention from the business [3].

It management actors in a company can manage stock by providing the most demanded products, because providing stock inventory needs is a strategy in achieving maximum sales targets. [4]

In previous research conducted by Satlea Suhada, et al with the title Application of the fp-growth algorithm to determine consumer purchasing patterns. Explaining that the fp-growth algorithm is suitable for determining consumer purchasing patterns in sales transaction data. The resulting association provisions can be used as support in decision making by company management such as helping with marketing strategies, ensuring product layout arrangements, and setting discounts for a mixture of certain items that are often purchased at the same time by consumers[5].

The problem faced by the owner of the Expert Computer shop is that the bookkeeping of inventory and sales is still done manually, and there is no complete sales data analysis related to the accumulation of goods due to lack of sales and empty goods, because the business process is manual, it is difficult for shop owners to analyze which products are in demand and which items are bought together. This has an impact on reduced buyer interest and sales because store owners cannot meet market demand.

The Frequent Pattern-Growth algorithm method is an alternative that can already be used to determine the most frequently occurring frequent itemset in a set of data. The Frequent Pattern-Growth algorithm uses the concept of tree construction, also known as Frequent Pattern-Tree when searching for frequent itemsets, rather than getting candidate results like the Apriori algorithm. This concept, rather than the Frequent Pattern-Growth algorithms, maximizes speed. [6]

One of the most frequently occurring alternative Frequent Pattern-Growth algorithms that determines the set of data that can be used for (frequent itemset) in a data set. [7]

2. METHODS

The stages of research are a sequence of processes or steps that will be taken in completing this research. The stages of this research are as follows:

1. Field Research

This method is done by coming directly to the Expert Computer Store to study existing problems and collect all data related to the research. This research was carried out by making observations or observations and conducting interviews.

2. Library Research

This research was conducted to research and collect information and sources related to the research. Information and sources are obtained from books, the internet, research reports, scientific journals, theses, and literature.

3. Laboratory Research

This research is related to the hardware and software used in the research. Conducted to test the accuracy of the system by practicing the results of the analysis and trying programming on a computer.

4. Analysis

Analysis of problem identification, research conducts several analyzes to produce the expected system, in the analysis process there are 3 stages, namely the stages of data analysis, process analysis, and system analysis.

5. Design

In this design, researchers use the Unified Modeling Language (UML), to design the system. As well as designing the interface of the system being built.

3. RESULT

Discussion of the results of research and testing obtained is presented in the form of theoretical descriptions, both qualitatively and quantitatively.

The data source used in this research is unprocessed data in the form of transaction data. The recorded data has product name and product code attributes.

Table 1 Transaction Sample Data

No	Product Name	Product Code
1	Acer Aspire 3050 4/256 GB	KD0001
2	Gamepad Double Black	KD0002
3	Acer Aspire A514 I5 GEN 11 4/512 GB	KD0003
4	Flashdisk Sandisk 16 GB	KD0004
5	Asus A416J I7 GEN 10	KD0005
6	Gamepad Single Black	KD0006
7	Asus M1403QA R5-5600 8/512 GB	KD0007
8	Keyboard + Mouse Rexus KM8	KD0008
9	HP 14S-FQ2002AU R5-5652 8/512 GB	KD0009
10	Flashdisk Sandisk 128 GB	KD0010
11	HP 14s-DQ0508TU CELL 4/256 GB	KD0011
12	HDD External 1TB	KD0012
13	Lenovo IP 1 R3-5300 8/512 GB	KD0013
14	Flashdisk Sandisk Fit USB CZ430 32 GB 3.0	KD0014
15	Lenovo V14 I3 GEN 11 4/256 GB	KD0015
16	Converter HDMI-VGA	KD0016
17	Mouse Wireless Logitech 186	KD0017

Table 2 grouped transaction data

No	Products
1	KD0005, KD0014, KD0012
2	KD0011, KD0006, KD0012
3	KD0007, KD0010, KD0014
4	KD0013, KD0014
5	KD0001, KD0016, KD0014
6	KD0003, KD0014, KD0002, KD0016
7	KD0005, KD0014, KD0016
8	KD0009, KD0014, KD0002
9	KD0011, KD0012, KD0017, KD0004
10	KD0015, KD0012, KD0017, KD0006
11	KD0003, KD0008, KD0017
12	KD0015, KD0008, KD0006, KD0002
13	KD0007, KD0010, KD0012, KD0017
14	KD0011, KD0010, KD0014, KD0016
15	KD0005, KD0006, KD0008
16	KD0013, KD0017, KD0012, KD0004
17	KD0011, KD0012, KD0014, KD0006
18	KD0011, KD0002, KD0004, KD0014
19	KD0015, KD0012, KD0017, KD0010
20	KD0001, KD0002, KD0008
21	KD0009, KD0017, KD0012, KD0006
22	KD0009, KD0012, KD0002, KD0017
23	KD0013, KD0017, KD0008, KD0012

Table 3 Frequency of Occurrence of Each Item/Product

No	Product Code	Product Name	Frequency
1	KD0001	Acer Aspire 3050 4/256 GB	2
2	KD0002	Gamepad Double Black	6
3	KD0003	Acer Aspire A514 I5 GEN 11 4/512 GB	2
4	KD0004	Flashdisk Sandisk 16 GB	3
5	KD0005	Asus A416J I7 GEN 10	3
6	KD0006	Gamepad Single Black	6
7	KD0007	Asus M1403QA R5-5600 8/512 GB	2
8	KD0008	Keyboard + Mouse Rexus KM8	5
9	KD0009	HP 14S-FQ2002AU R5-5652 8/512 GB	3
10	KD0010	Flashdisk Sandisk 128 GB	4
11	KD0011	HP 14s-DQ0508TU CELL 4/256 GB	5
12	KD0012	HDD External 1TB	11
13	KD0013	Lenovo IP 1 R3-5300 8/512 GB	3
14	KD0014	Flashdisk Sandisk Fit USB CZ430 32 GB 3.0	10
15	KD0015	Lenovo V14 I3 GEN 11 4/256 GB	3
16	KD0016	Converter HDMI-VGA	4
17	KD0017	Mouse Wireless Logitech 186	9

After calculating the frequency of occurrence of each item, it is known that items that are above the minimum support value = 17% where above the minimum support the product will be influential and will be included in the FP-tree, the rest can be discarded because they do not meet the minimum support = 17%.

Table 4 Transactions customized with Frequent List

No	Products
1	KD0012, KD0014,
2	KD0012, KD0006, KD0011
3	KD0014, KD0010
4	KD0014,
5	KD0014, KD0016
6	KD0014, KD0002, KD0016
7	KD0014, KD0016,
8	KD0014, KD0002,
9	KD0012, KD0017, KD0011
10	KD0012, KD0017, KD0006
11	KD0017, KD0008
12	KD0002, KD0006, KD0008
13	KD0012, KD0017, KD0010
14	KD0014, KD0011, KD0010, KD0016
15	KD0006, KD0008
16	KD0012, KD0017
17	KD0012, KD0014, KD0006, KD0011
18	KD0014, KD0002, KD0011
19	KD0012, KD0017, KD0010
20	KD0002, KD0008
21	KD0012, KD0017, KD0006
22	KD0012, KD0017, KD0002
23	KD0012, KD0017, KD0008

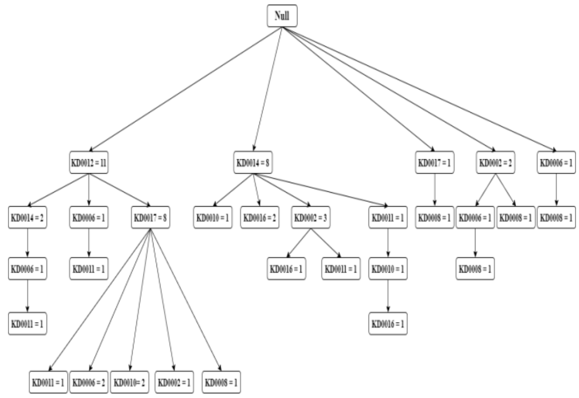


Figure 1. FP-Tree formation results after TID 3 reading

The picture is an explanation of the formation of the Fp-Tree after the reading is obtained after doing TID 23, which contains Null-KD0012=11, KD0017=8, KD0008=1.

To find Frequent itemset from table 4.6, it is necessary to first determine the path that ends with the smallest support count, namely

KD0016 followed by KD0010, KD0008, KD0011, KD0002, KD0006, KD0017, KD0014, and ended with KD0012.

After checking the frequent itemset for several suffixes, the results are summarized in the table below.

Table 5 Frequent Itemset (Suffix) results

No	*Suffix*	*Frequent Itemset*
1	KD0016	{KD0016}{KD0016,KD0014}{KD0016,KD0011}{KD0016, KD0010}{KD0016, KD0002}{KD0016, KD0002, KD0014}{KD0016, KD0010, KD0014}{KD0016, KD0011, KD0014}{KD0016, KD0010, KD0011,

No	*Suffix*	*Frequent Itemset*
		KD0014}
2	KD0010	{KD0010}{KD0010,KD0012} {KD0010,KD0014}{KD0010, KD0017}{KD0010, KD0011} {KD0010,KD0017, KD0012} {KD0010, KD0011, KD0014}
3	KD0008	{KD0008}{KD0008, KD0012} {KD0008, KD0017} {KD0008, KD006}{KD0008,KD0002} {KD0008,KD0017,KD0012} {KD0008, KD0002, KD0006}
4	KD0011	{KD0011} {KD0011, KD0012} {KD0011, KD0014} {KD0011, KD0006} {KD0011, KD0017} {KD0011, KD0002} {KD0011, KD0006} {KD0011, KD0006, KD0012} {KD0011, KD0014, KD0012} {KD0011, KD0002, KD0014} {KD0011, KD0006, KD0014} {KD0011, KD0006, KD0014, KD0012}
5	KD0006	{KD0006} {KD0006, KD0012} {KD0006, KD0014} {KD0006, KD0017} {KD0006, KD0014, KD0012} {KD0006, KD0017, KD0012}
6	KD0002	{KD0002} {KD0002, KD0006} {KD0002, KD0014} {KD0002, KD0012} {KD0002, KD0017} {KD0002, KD0017, KD0012}
7	KD0017	{KD0017} {KD0017, KD0012}
8	KD0014	{KD0014} {KD0014, KD0012}
9	KD0012	{KD0012}

Stages carried out after Frequent Item, then the next stage is to add up from each item or product that has a support count with a minimum support value of> 17% and mimimum confidence is 60%, will be generated in the Conditional FP-Tree. The following is the generation of Conditional FP-Tree on 2 items, 3 items and 4 items. The explanation obtained from Frequent Item can be seen in the following table[6] :

$$\text{Suppor} = \frac{\text{Number of Transaction cases Contataining A and B}}{\text{Total Transactions}} \times 100$$

$$\text{Confidence} = \frac{\text{Number of Transactions Containing A and B}}{\text{Number of Transactions Containing A}} \times 100$$

Table 6 Support and Confidence 2 Items

If Antecedent then Consequent	*Support*	*Confidence*
{KD0016, KD0014}	$= \frac{4}{23} \times 100 = 17{,}39\%$	$= \frac{4}{4} \times 100 = 100\%$
{KD0016, KD0011}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{4} \times 100 = 25\%$
{KD0016, KD0010}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{4} \times 100 = 25\%$
{KD0016, KD0002}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{4} \times 100 = 25\%$
{KD0010, KD0012}	$= \frac{2}{23} \times 100 = 8{,}69\%$	$= \frac{2}{4} \times 100 = 50\%$
{KD0010, KD0014}	$= \frac{2}{23} \times 100 = 8{,}69\%$	$= \frac{2}{4} \times 100 = 50\%$
{KD0010, KD0017}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{4} \times 100 = 25\%$
{KD0010, KD0011}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{4} \times 100 = 25\%$
{KD0008, KD0012}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{5} \times 100 = 20\%$
{KD0008, KD0017}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{5} \times 100 = 20\%$
{KD0008, KD0006}	$= \frac{1}{23} \times 100 = 4{,}34\%$	$= \frac{1}{5} \times 100 = 20\%$
{KD0008, KD0002}	$= \frac{2}{23} \times 100 = 8{,}69\%$	$= \frac{2}{5} \times 100 = 40\%$
{KD0011, KD0012}	$= \frac{3}{23} \times 100 = 13{,}04\%$	$= \frac{3}{5} \times 100 = 60\%$
{KD0011, KD0014}	$= \frac{2}{23} \times 100 = 8{,}69\%$	$= \frac{2}{5} \times 100 = 40\%$
{KD0011, KD0006}	$= \frac{2}{23} \times 100 = 8{,}69\%$	$= \frac{2}{5} \times 100 = 40\%$

If Antecedent then Consequent	*Support*	*Confidence*
{KD0011, KD0017}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{5} \times 100$ $= 20\%$
{KD0011, KD0002}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{5} \times 100$ $= 20\%$
{KD0006, KD0012}	$= \frac{4}{23} \times 100$ $= 17{,}39\%$	$= \frac{4}{6} \times 100$ $= 66{,}66\%$
{KD0006, KD0017}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{6} \times 100$ $= 16{,}66\%$
{KD0002, KD0006}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{6} \times 100$ $= 16{,}66\%$
{KD0002, KD0014}	$= \frac{3}{23} \times 100$ $= 13{,}04\%$	$= \frac{3}{6} \times 100$ $= 50\%$
{KD0002, KD0012}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{6} \times 100$ $= 16{,}66\%$
{KD0002, KD0017}	$= \frac{1}{23} \times 100$ $= 4{,}34\%$	$= \frac{1}{6} \times 100$ $= 16{,}66\%$
{KD0017, KD0012}	$= \frac{7}{23} \times 100$ $= 30{,}43\%$	$= \frac{7}{9} \times 100$ $= 77{,}77\%$
{KD0014, KD0012}	$= \frac{2}{23} \times 100$ $= 8{,}69\%$	$= \frac{2}{9} \times 100$ $= 22{,}22\%$

After all the suffix searches above, it can produce a Conditional FP-Tree that is obtained in accordance with the minimum support> 17% and minimum confidence 60% and items that are eliminated or discarded are not written again in Table 7 below:

Table 7 Results of Conditional FP-Tree

No	*Suffix*	*Conditional FP-Tree*
1	KD0016	{KD0016 = 4} {KD0016, KD0014 = 4}
2	KD0010	{KD0010 = 4}
3	KD0011	{KD0011 = 5}
4	KD0006	{KD0006 = 6} {KD0006, KD0012 = 4}
5	KD0002	{KD0002 = 6}
6	KD0017	{KD0017 = 9} {KD0017, KD0012 = 8}
7	KD0014	{KD0014 = 10}
8	KD0012	{KD0012 = 11}

Of the 23 Conditional FP-Tree, not all are calculated. Because in generating a minimum Association Rule Conditional FP-Tree that is calculated there are 2 items or products where if you buy a KD0016 product, you will buy a KD0014 product. So that what is worth calculating the confidence is 3 subsets, namely {KD0016, KD0014}, {KD0006, KD0012}, {KD0017, KD0012}.

Association Rule can be known by two parameters, namely support with a minimum support of 17% and confidence with a minimum confidence of 60%.

After obtaining the frequent itemset, then create a rule by calculating the confidence of each rule.

Table 8 Results of Support and Confidence

If Antecedent then Consequent	*Support*	*Confidence*
If Customer Buys KD0014 Then Customer Also Buys KD0016	$= \frac{4}{23} \times 100$ $= 17{,}39\%$	$= \frac{4}{10} \times 100$ $= 40\%$
If Customer Buys KD0016 Then Customer Also Buys KD0014	$= \frac{4}{23} \times 100$ $= 17{,}39\%$	$= \frac{4}{4} \times 100$ $= 100\%$
If the customer buys KD0012 then the customer also buys KD0016.	$= \frac{4}{23} \times 100$ $= 17{,}39\%$	$= \frac{4}{11} \times 100$ $= 36{,}36\%$
If the customer purchases KD0016 then the customer also purchases KD0012	$= \frac{4}{23} \times 100$ $= 17{,}39\%$	$= \frac{4}{6} \times 100$ $= 66{,}67\%$
If the customer purchases KD0012 then the customer also purchases KD0017.	$= \frac{8}{23} \times 100$ $= 34{,}78\%$	$= \frac{8}{11} \times 100$ $= 72{,}73\%$
If the customer purchases KD0017 then the customer also purchases KD0012.	$= \frac{8}{23} \times 100$ $= 34{,}78\%$	$= \frac{8}{9} \times 100$ $= 88{,}89\%$

Of the many rules obtained, multiply the Support with Confidence, where Confidence is taken around 60% and above Like the table

Table 9 Results of ruls

If Antecedent then Consequent	*Support*	*Confidence*	*Support * Confidence*
Jika Pelanggan Membeli KD0016 Maka Pelanggan Juga Membeli KD0014	0,1739	100	$0{,}1739 \times 100 = 17{,}39$
Jika Pelanggan Membeli KD0002 Maka Pelanggan Juga Membeli KD0014	0,1739	0,6667	$0{,}1739 \times 0{,}666 = 0{,}116$
Jika Pelanggan Membeli KD0006 Maka Pelanggan Juga Membeli KD0012	0,3478	0,7273	$0{,}3478 \times 0{,}727 = 0{,}253$
Jika Pelanggan Membeli KD0012 Maka Pelanggan Juga Membeli	0,3478	0,8889	$0{,}3478 \times 0{,}8889 = 0{,}309$

If Antecedent then Consequent	*Support*	*Confidence*	*Support * Confidence*
Jika Pelanggan Membeli KD0016 Maka Pelanggan Juga Membeli KD0014 KD0017	0,1739	100	0,1739 × 100 = 17,39

Minimum support and confidene results. The result of the related rule is the one with the largest Confidene, namely the gray-marked case. After obtaining a rule that has confidene ≥60% then determining the rule that meets the needs, it can be concluded that the related rule is if the customer buys a 1TB External HDD then the customer also buys a Logitech 186 Wireless Mouse, if the customer buys a Gamepet Singel Black then the customer buys a Gamepet Double Black, if the buyer buys a Gameped Double Black then the buyer also buys a Logitech 186 Wireless Mouse.

4. Conclusion

From a series of research processes carried out, conclusions can be drawn regarding the development of IT Bussines Management using the FP-Growth algorithm method, with the existence of IT Bussines Management by determining the purchase association pattern built using the FP-Growth algorithm will help the owner of the Expert Computer Store in analyzing sales transaction data and further utilization to obtain useful information for the Expert Computer Store.

References

[1] L. Y. Siregar and M. I. P. Nasution, "Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online," *HIRARKI J. Ilm. Manaj. dan Bisnis*, vol. 02, no. 01, pp. 71–75, 2020, [Online]. Available: http://journal.upp.ac.id/index.php/Hirarki%0APERKEMBANGAN

[2] U. Nusamandiri, "Penerapan Algoritma C4.5 Untuk Kepuasan Pelanggan Toko," vol. 3, no. 1, pp. 10–21, 2022.

[3] W. Wulandari, "Implementasi Sistem Peramalan Persediaan Barang Menggunakan Metode Moving Average," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 707, 2020, doi: 10.30865/mib.v4i3.2199.

[4] D. A. Silitonga and A. P. Windarto, "Implementasi Market Basket Analysis Menggunakan Assocation Rule Menerapkan Algoritma FP-Growth," *J. Inf. Syst. Res.*, vol. 3, no. 2, pp. 101–109, 2022, doi: 10.47065/josh.v3i2.1239.

[5] S. Suhada, D. Ratag, G. Gunawan, D. Wintana, and T. Hidayatulloh, "Penerapan Algoritma Fp-Growth Untuk Menentukan Pola Pembelian Konsumen Pada Ahass Cibadak," *Swabumi*, vol. 8, no. 2, pp. 118–126, 2020, doi: 10.31294/swabumi.v8i2.8077.

[6] H. E. Simanjuntak and W. Windarto, "Analisa Data Mining Menggunakan Frequent Pattern Growth pada Data Transaksi Penjualan PT Mora Telematika Indonesia untuk Rekomendasi Strategi Pemasaran Produk Internet," *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 914–923, 2020, doi: 10.30865/mib.v4i4.2300.

[7] A. Chailes, A. Hermawan, and D. Kurnaedi, "Penerapan Metode Data Mining Untuk Menentukan Pola Pembelian Dengan Menggunakan Algoritma," *J. Algor*, vol. 1, no. 2, pp. 1–8, 2020.