Jatilima : Jurnal Multimedia Dan Teknologi Informasi
Volume 06, Number 02, 2024, doi.org/10.54209/jatilima.v6i02.681
E.SSN : 2721-1800 (Online)
https://journal.cattleyadf.org/index.php/jatilima/index

Determining Product Category Sales Patterns to Maintain Inventory Stability Using the fp-Growth Algorithm

[1]Nelisa, [2]Khairunnisa, [3]Ika Yusnita Sari, [4]Elvika Rahmi, [5]Fajar Maulana
[1] STIKOM 22 Januari Kendari
[2,3,4] Universitas Imelda Medan
[5]Universitas Adzkia Padang

Article Info
Keywords:
Data Mining,
Product,
Fp-Growth,
Weka

ABSTRACT
Managing a very large collection of databases requires a method or technique that can convert piles of data into useful information, one of the data that can be processed is sales data. Anoa Mart mini market is one of the mini markets that participates in meeting consumer needs, so a good way of managing goods is needed to meet consumer needs. One of the data that can be processed is sales transactions at Anoa Mart. Mini Market will provide new information to increase sales. This research aims to determine consumer purchasing habits in order to maintain inventory stability. In this research, the data processing used is Anoa Mart Mini Market sales transaction data. Transaction data is processed using the Fp-Growth consortium Data Mining technique with a confidence value of 70% and a minimum support level of 30%. Processing transaction data produces new knowledge in the form of rules for each item. Each item purchased simultaneously constitutes an association rule that is derived from the value of trust and later becomes knowledge for the small market owner. The model resulting from these rules can be used as a reference to maintain inventory stability and increase sales. This method can be used by small markets to convey information more quickly and accurately, so that sales levels increase and are well controlled.

Corresponding Author:
Nelisa
STIKOM 22 Januari Kendari
Jl.MT.Haryono No.79 Kendari, Sulawesi Tenggara
nelisanasution21@gmail.com

INTRODUCTION

Knowledge Discovery in Databases (KDD) is the process of retrieving useful information and hidden patterns from data in large databases. Knowledge Discovery in Database (KDD). Knowledge Discovery in Database (KDD) consists of several steps that aim to find valuable knowledge from raw data. the main steps in the KDD process are Selection, Preprocessing, Transformation, Data Mining, Interpretation/Evaluation, Knowledge Presentation. Research on the Fp-Growth algorithm using Data Mining has been carried out before, research using Data Mining is used to extract new knowledge from large chunks of data, stored data is stored in storage using pattern recognition techniques. Data mining techniques can find correlations between different attributes to detect specific data.

The association rule is a data mining technique that is part of other data mining methods. Association rules are used to determine the correlation of elements in a predefined dataset . Currently, combinations have been applied to classification based on correlations between itemsets. A search to obtain association rules can be performed from transactional databases to obtain support regarding the number of items that are usually obtained. One of the algorithm steps that eliminates all frequently used elements in a data file is the frequent sample growth (Fp-Growth) algorithm. This element represents the combination of data sets that are often brought together in data with many links. The main research steps on the Fp-Growth algorithm are building a model tree, building a conditional model base, and exploring models that often appear simultaneously. The Fp-Growth Algorithm search is used to help determine the relationships between elements.

Anoa Mart Mini Market is a convenience store that will meet the needs of the community. To improve service efficiency, it is necessary to establish a stable work method that can provide information quickly, accurately and on time according to the needs of consumers or society. To facilitate the management of small market data, a system is needed that is able to make decisions to determine which products are most requested and needed by consumers and predict the availability of inventory. Therefore, a lot of data can be used optimally.

From the above problems, alternative solutions can be applied using data mining by developing the Fp-Growth method to find out the sales pattern at the Anoa Mart Mini Market. can calculate different values based on predefined items and data.

METHODS

The method used in the study to manage the data that needs to be used as a sample is the Fp-Growth algorithm method. This algorithm is a commonly used development, namely the A priori algorithm. Fp-Growth is an alternative algorithm used to identify frequent itemsets in a data set. Based on previous work by Sun and Chen (2023), Fp-Growth is a priori-based algorithm that stores data in a frequency pattern tree (FP tree) to extract frequency patterns and association rules. A pattern is determined in frequency if the number of hits is greater than or equal to a predetermined support threshold [9]. When finding patterns that occur frequently with a minimum support limit using the FP tree algorithm, the Fp-Growth algorithm is used. If the support of the pattern is not less than a predetermined constant (minimum support limit), then the result is called a general pattern (often occurring) [10]. The steps performed on the Fp-Growth algorithm can be seen in Figure 1.

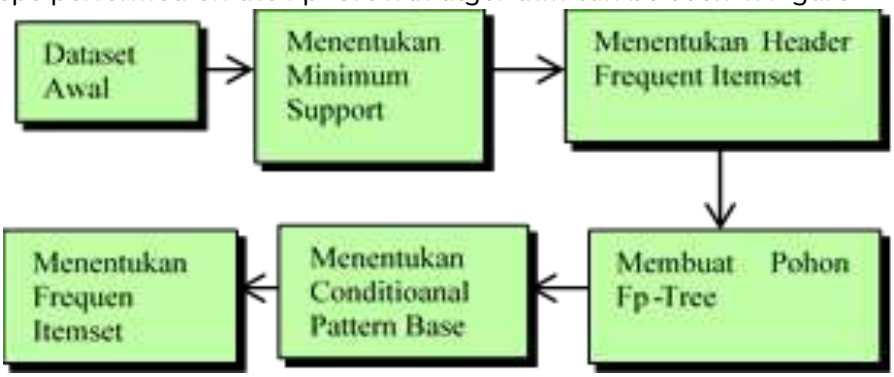

Figure 1. Stages of the Fp-Growth Algorithm Diagram

The Fp-Growth algorithm is a fairly old algorithm, the Fp-Growth algorithm can reduce a large number of items and only runs the database twice to find a collection of items that appear frequently. The steps of the Fp-Growth method are:.

a. The conditional pattern base,
   In this step, a conditional model base is created. The subdata contains the prefix path (initial path) and the suffix pattern (suffix pattern) that form the basis of the conditional pattern. From the pre-built FP Tree, a basic generation of conditional models can be obtained.
b. FP-Tree conditional generation stage
   After performing a basic conditional model search, the next step is to sum each element, then any element with a support number greater than or equal to the minimum support number will be generated using the FP Tree with conditions.
c. Frequent itemset search stage.
   At this stage, the determination of the FP-Tree conditional. After obtaining the results of the condotional pattern base, it is continued with the search for items for conditional FP-Tree to obtain frequent item sets.

RESULTS AND DISCUSSION

During the implementation process, the first thing to do is analyze the data. The data analysis process is one of the important stages of a process in a research, because at the data analysis stage the results will also be determined. . Such as analyzing the data obtained, where the data is used as documentation. To create an application system, the data used is data on sales transactions of a product that helps determine consumer purchase patterns for a product. The types of consumer buying behaviors are so diverse that it can sometimes be difficult to estimate what type of products consumers need when shopping. Therefore, it is necessary to analyze the data obtained.

Transaction data collection The data used in this study is sales transaction data in 2023 by limiting the 20 most purchased products among 20 transaction data and sorting their priority values. and encoding these elements as study objects in Table 1.

Table 1. Research data table

Id Item	Item	Priority
1	Ultra Milk, Malboro Putih, Gudang Garam, Big Cola	{7,10,13,17}
2	Esse Change Juicy, Frestea	{9,14}
3	Esse Change Juicy, Top Coffe, Big Cola	{8,9,17}
4	Ultra Milk, Malboro Putih	{7,10}
5	Ultra Milk, Malboro Putih	{7,10}
6	Mencis, Frestea, Ades	{1,4,14}
7	Esse Change Juicy, Top Coffe	{8,9}
8	Mencis, Malboro Putih, Frestea, Gudang Garam	{4,10,13,14}
9	Mencis, Malboro Putih	{4,10}
10	Esse Change Juicy, Frestea	{9,14}
11	Ultra Milk, Top Coffe, Nabati, Ades	{1,7,8,16}

12	Top Coffe	{8}
13	Esse Change Juicy, Mencis, Top Coffe	{4,8,9}
14	Ultra Milk, Gudang Garam, Nabati, Big Cola, Big Cola, Ades	{1,7,13,16,17}
15	Top Coffe	{8}
16	Esse Change Juicy, Mencis, Top Coffe	{4,8,9}
17	Ultra Milk, Frestea, Nabati	{7,14,16}
18	Esse Change Juicy, Mencis	{4,9}
19	Ultra Milk, Malboro Putih, Gudang Garam, Big Cola	{7,10,13,17}
20	Ultra Milk, Nabati	{7,16}

A. Formation of Fp_Tree

After identifying high frequency values or elements that meet the minimum support level, the next step is to form an Fp Tree by searching using the decision tree.

a. FP tree to sort from transaksi_id 1:{7,10,13,17}. The resulting values are as shown in Figure 2.

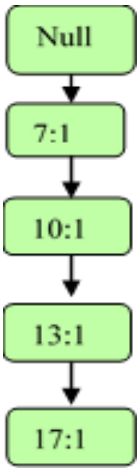

Figure 2. FP-Tree at id_transaksi 1

b. FP-Tree follows the order, specifically transaksi_id 2: { 9.14 }. The resulting values are as seen in Figure 3.

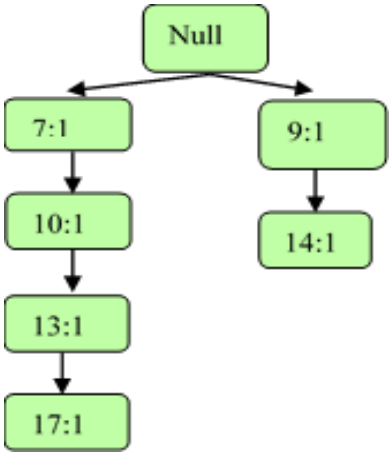

Figure 3. FP-Tree at id_transaksi 2

c. FP-Tree then follows the sequence, specifically Transaction_id 3: {8,9,17} The resulting values are as shown in Figure 4

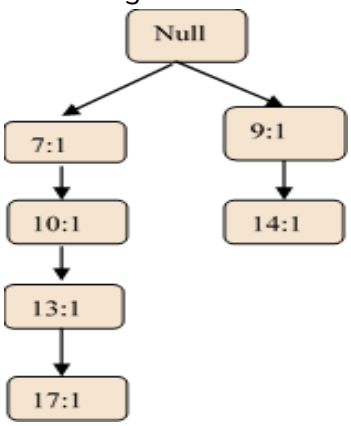

Figure 4. FP-Tree at id_transaksi 3

d. FP Tree is created up to Transaction_id 20 until the final value is formed as shown in Figure 5:

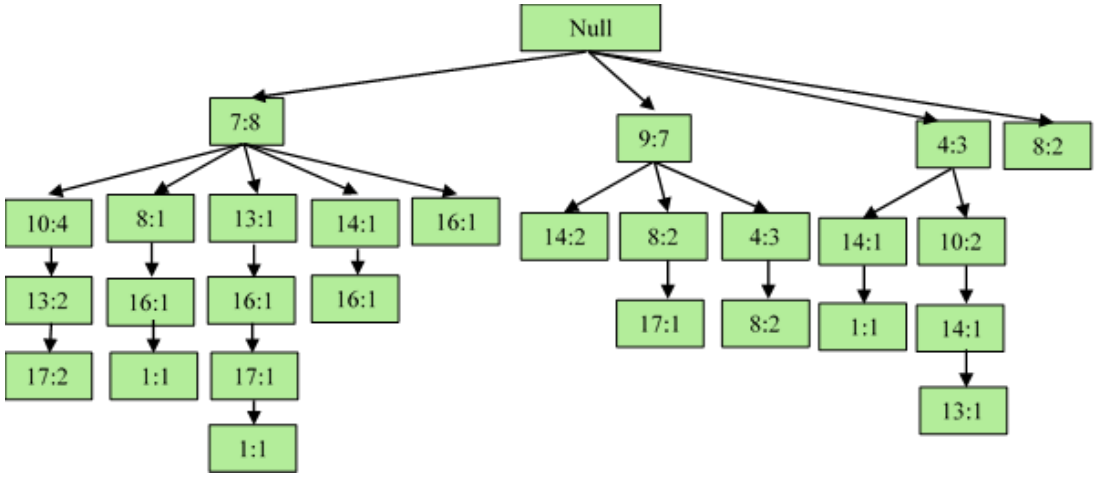

Figure 5. FP-Tree Final Results

B. Stages of the Fp-Growth Algorithm (Frequent Pattern Growth)

Once the Fp-Tree decision tree search is complete, the next step is to perform a search using the Fp-Growth algorithm procedure to find frequent and meaningful itemsets. Fp-Growt has 3 main steps in the research process.

a. Conditional Pattern Base Phase

The conditional pattern base is a subdatabase containing the initial path and suffix pattern and the frequent itemset is the result, this generation is achieved through the result of the previous FP Tree, which can be shown in Table 2.

Tabel 2. Conditional Pattern Base

Priority	Item	Condition pattern base
10	1	{7,8,16: 1} {7,13,16,17 : 1} {4,14 : 1}

9	17	{7,10,13 : 2} {7,13,16 : 1} {9,8 : 2}
8	16	{7,8 :1} {7,13 : 1} {7,14 : 1} {7:1}
7	13	{7,10: 2} {7:1} {4,10,14:1}
6	14	{7:1} {9:2} {4:1} {4,10:1}
5	18	{7:1} {9:2}{9,4:2}
4	10	{7:4} {4:2}
3	4	{9:3}
2	9	-
1	1	-

b. FP-Tree Conditional Generation Stage

At this point, the number of supports of each element in each conditional model base is added, and then any element with a support number greater than or equal to a minimum number of supports of two will be created using the Conditional FP Tree. The results of the FP-Tree conditional search can be seen in Table 3.

Table 3. Conditional Fp-Tree

Priority	Item	Conditional Fp-Tree
10	1	-
9	17	<7:3> <13:3>
8	16	<7:4>
7	13	<7:3> <10:3>
6	14	-
5	18	<9:4>
4	10	<7:4>
3	4	<9:3>
2	9	-
1	1	-

c. Frequent Itemset Search Stage

A Conditional FP Tree is a single path, so the generated frequent itemset or also called a frequent itemset is created by combining items for each Conditional FP Tree. If this is not a unique path then Fp-Growth will be created recursively (this process will call itself). The results of the frequent itemset search can be seen in Table 4.

Tabel 4. Frequent Pattern Generated

Priority	Item	Frequent Pattern Generated
10	1	
9	17	{7,17:3} {13,17:3} {7,13,17:3}
8	16	{7,16:4}
7	13	{7,13:3} {10,13:3} {7,10,13:3}
6	14	-
5	18	{9,8:4}
4	10	{7,10:4}

3	4	{9,4:3}
2	9	-
1	1	-

The minimum support requirement is 30% and the minimum reliability is 70%. Thus, there are 10 rules that are included in the strong association rule. It can be seen in Table 5.

Tabel 5. Strong Association Rule

If Buying	So Buy	Support	Confidance
13	7 dan 17	15%	75%
17	7 dan 13	15%	75%
7	17 dan 13	15%	100%
7	13 dan 17	15%	100%
17	13 dan 7	15%	100%
16	7	20%	100%
13	7 dan 10	15%	75%
7	13 dan 10	15%	100%
7	10 dn 13	15%	100%
13	10 dan 7	15%	100%

Table 5 presents the 10 strong association rules obtained, specifically the new form of knowledge with the highest reliability value derived from consumer purchasing habits from sales transactions with a total of 20 transactions in 2023 data at the Anoa Mart mini supermarket. Below are the complete results of the model or rules generated:

a. If you buy Gudang Garam, you will buy Ultra Milk and Big Cola with an accuracy level of 75% with a support of 15% of the overall data is Rule 1.
b. If you buy Big Cola, you will buy Ultra Milk and Gudang Garam with an accuracy level of 75% with a support of 15% of the overall data is Rule 2.
c. If you buy Ultra Milk, you will buy Big Cola and Gudang Garam with a level of 100% accuracy with 15% support of the overall data is Rule 3.
d. If you buy Ultra Milk, you will buy Gudang Garam and Big Cola with an accuracy level of 100% with a support of 15% of the overall data is Rule 4.
e. If you buy Big Cola, you will buy Gudang Garam and Ultra Milk also with an accuracy level of 100% with a support of 15% of the overall data is Rule 5.
f. If you buy Nabati, you will buy Ultra Milk with an accuracy level of 100 with support of 20% of the overall data as a Rule 6.
g. If you buy Gudang Garam, you will buy Ultra Milk and Malboro Putih with an accuracy level of 75% with a support of 15% of the overall data is Rule 7.
h. If you buy Ultra Milk, you will buy Gudang Garam and Malboro Putih with an accuracy level of 100% with a support of 15% of the overall data is Rule 8.
i. If you buy Ultra Milk, you will buy Malboro Putih and Gudang Garam with an accuracy level of 100% with a support of 15% of the overall data is Rule 9.
j. If you buy Gudang Garam, you will buy Malboro White and Ultra Milk with an accuracy rate of 100% with a support of 15% of the overall data is Rule 10.

CONCLUSION

Sales transactions at the Anoa Mart mini supermarket use the Fp-Growth data mining technique with a sample of 20 transaction data resulting in 10 sales models of processed products. From these models, we can introduce them to the owners of the Anoa Mart mini market. Thanks to that, the Anoa Mart Mini Market can use this model as information to predict product availability and prepare product inventory to be able to control the sale of products that are in great demand by consumers to increase sales results.

REFERENCE

Dewi, A. O. P. (2023). Big Data di Perpustakaan dengan Memanfaatkan Data Mining. Anuva: Jurnal Kajian Budaya, Perpustakaan, dan Informasi, 4(2), 223-230. https://doi: 10.14710/anuva.4.2.223-230

N, A. P., & Dr., M. P. (2023). Bounded Memory Based Frequent Pattern Growth Approach With Deep Neural Network And Decision Tree For Road Accident Prediction. Indian Journal of Computer Science and Engineering, 11(5), 623–633. https://doi.org/10.21817/indjcse/2023/v11i5/201105189

Noorkholid, M. I., Hidayat, M. A., & Fajarianto, G. W. (2023). Sistem Informasi Penentuan Paket Pembelian Produk Menggunakan Algoritma Frequent Pattern-Growth pada KPRI Jember. BERKALA SAINSTEK, 8(2), 59. https://doi.org/10.19184/bst.v8i2.11848

Wang, L., Zhu, H., & Huang, R. (2018). Association Rule Classification and Regression Algorithm Based on Frequent Itemset Tree. Proceedings of the 2018 3rd International Conference on Modelling, Simulation and Applied Mathematics (MSAM 2018). https://doi.org/10.2991/msam- 18.2018.30

Li, Y., & Yin, S. (2019). Mining Algorithm for Weighted FP- Growth Frequent Item Sets based on Ordered FP-Tree. International Journal of Engineering and Management Research, 09(05),154–158. https://doi.org/10.31033/ijemr.9.5.22

G., A. (2018). A Data Mining Analysis of ERP System using Frequent Pattern Growth Algorithm. International Journal of Computer Applications, 182(26), 30–35. https://doi.org/10.5120/ijca2018918154

Yu, W., Yi, M., & Li, Z. (2019). Research on Constructing Technology of Implicit Hierarchical Topic Network Based on FP-Growth. Artificial Intelligence and Security, 260–272. https://doi.org/10.1007/978-3-030-24274-9_23

Wang, C.-S., & Chang, J.-Y. (2019). MISFP-Growth: Hadoop- Based Frequent Pattern Mining with Multiple Item Support. Applied Sciences, 9(10), 2075. https://doi.org/10.3390/app9102075

Sun, Y., & Chen, X. (2023). An Improved Frequent Pattern Growth Based Approach to Intrusion Detection System Alert Aggregation. Journal of Physics: Conference Series, 1437, 012070. https://doi.org/10.1088/1742-6596/1437/1/012070

Setyo, W. N., & Wardhana, S. (2019). Implementasi Data Mining pada Penjualan Produk di CV Cahaya Setya Menggunakan Algoritma Fp-Growth. Petir, 12(1). https://doi.org/10.33322/petir.v12i1.416