

# Implementation of FP-Growth Algorithm in Determining Food Package Recommendation in Sunan Giri Ribs Meatball Restaurant

Vera Munfarijah  
ST  
Gunadarma University  
Depok, Indonesia

D. Lucia Crispina Pardede, PhD  
DEA  
Gunadarma University  
Depok, Indonesia

## ABSTRACT

Recommendation packages or menu packs are found in many restaurants or dining houses in order to help the restaurant manager in both developing and promoting menu innovations. Using sales transaction data, process of data mining is assisted by FP-Growth algorithm in which also called a market basket analysis. This analysis has the capability to perform customers shopping pattern by finding associations between several different items bought. The FP-Growth algorithm has three stages in processing the analysis, i.e. Conditional Pattern, Conditional FP-Tree and Frequent Item Set. In this study, the FP-Growth algorithm established an association in customers shopping pattern of Sunan Giri Ribs Meatballs Restaurant by using a minimum support value of 0.2 and the highest minimum confidence as a parameter for the FP-Growth algorithm to produce recommendation packages. This research article gives technical explanation in the utilisation of FP-Growth algorithm in executing the analysis and providing recommendation packages. In line to the e-search findings, it is proved that FP-Growth algorithm is capable to be considered as a tool in making a new innovation.

## General Terms

Association Rule, Data Mining, Market Basket Analysis, Frequent Pattern Growth

## Keywords

FP-Growth Algorithm, Association Rule, Recommendation.

## 1. INTRODUCTION

In the era of consumerism and globalism, restaurant still be a thriving business which is also marked as one of the highly competitive markets in the field of food and beverages. Thus, running this business requires an innovative strategy to compete with other restaurant and café. Therefore, developing ideas and innovations to improve the value of products, services, and increase profitability while also reduce risk of loss is needed by the managerial. Nevertheless, they have to meet the needs and capacity of every restaurant. Sunan Giri Ribs Meatballs Restaurant (SGRMR) for instance is one of the leading meatball restaurant in Jakarta, Indonesia. Established in 2018, they opened their first branch in Kelapa Gading Business District area. One of their trademark ingredient from which they get their unique delicacy is ribs from young cow.

Until today, they have already had 11 branches, nine of them located around Jakarta, and two other in Bogor. They offer several varieties of menu including fine ribs meatball, spicy meatball, egg meatball, and all of them are served with their special ribs. Even though SGRMR has become one of the

leading meatball restaurant in Jakarta, they still have to deal with the fast growing food industry. This reason require them to always developing new food innovation and ideas in terms of promotional and marketing strategy. For further strategy planning, this research article utilize FP-Growth algorithm to analyse every possible package menu recommendations from their previous sales data. This algorithm generate package menu recommendation from establishment of association rule and look for the association rule by specifying the minimum support value and the minimum confidence value on the data selection.

## 2. LITERATURE REVIEW

In conducting the research, review of several literatures in the same topic is needed. Table below shows three literature used in the research as the base line to help the researchers in collecting and analyzing data to finally draw conclusion from research findings as well as providing some strategy in further research on the same topic.

Table.1 Related studies

Year Published	Author	Title	Results
2017	Lintang Pahridil, Muhammad Iqbal and Ade Pujianto	Food Package recommendation system with FP-Growth algorithm at Restaurant Seafood XYZ	The FP-Growth algorithm covers the weakness of a priori algorithm in determining a frequent itemset. The data mining performance is influenced greatly by the stage of determining frequent itemset, especially if the itemset candidate is very large. In this study, FP-growth algorithm scan only twice the database from the beginning of the iteration to determine the support count and the first TID reading. The performance of data mining will increase compared to a priori algorithm that will scan the database as much as K-Itemset.
2018	Muhammad Fauriah	Application of Data	Implementing Data Mining method with FP-Growth algorithm into

	Fajrin and Algifanri Maulana	Mining for analysis of consumer purchase patterns with FP-Growth algorithm on Motor Spare part Sales Transaction Data	application for analysis of consumers buying pattern is very beneficial for the company, cause it helps TB-Damar to track which spare part is bought and sort the data out to be used by head office.
2019	Kana Fitria Rahmawati	Implementation of FP-Growth algorithm in Erafone counter sales system to maximize sales and minimize losses	The Erafone counter application has able to implement the data mining method of the Association Rule FP-Growth algorithms well. The application in this industry helps them sorting the popularity of the mobile phones as well as accessories based on customers from previous sales data. The algorithm can also used to combine which products are bought the most and least to finally help managerial in making decisions on stocking products.

All of these three literature use the FP-Growth algorithm in analyzing research data to make a useful strategy planning not only in the food industry but also in other sales industry. From the table above, results found that FP-Growth algorithm help managerial to make decision based on the package recommendation generated by the algorithm. Related to the previous studies, this research use the same FP-Growth algorithm to analyze sales data from the SGRMR data pool to provide strategy planning in generating menu package recommendation.

## 2.1 Data Mining

Data mining can be defined as a process of extracting or collecting useful information from a collection of data. This information is usually hidden patterns contained in the data, the relationship of data elements, or the model of the application for the purposes of data forecasting (Adinugroho & Sari, 2018:3).

Moreover, data mining is also defined as a series of processes for digging up an added value from an unknown information from a database manually. The information is produced by extracting and identifying important patterns or interesting data contained in the database. This process of data mining is used to generate knowledge contained in large databases so it is often referred to as Knowledge Discovery Databases (KDD) (Vulandari, 2017).

While according to Susanto & Suryadi (2011), data mining is explained as the act of finding, digging, or mining knowledge

from some sets of data or informations. To conclude, in this research, data mining is determined as the process inside the FP-Growth algorithm in which hundreds or even thousands sets of data is sorted to make desired output.

## 2.2 Algorithm Frequent Pattern Growth (FP-Growth)

The FP-Growth algorithm generates frequent item sets data from FP-Tree using Divide and Conquer method. This process is to find the frequent item set without creating an item candidate. Built using two datasets, the first path scans the transaction database and finds support value for each item. By this, the support value will gradually increase and items without enough frequency or below minimum support will be considered unused automatically. After this, the data above minimum support will be sorted based on the highest frequency. In the second path, the FP-Growth reads and tracks transactions to the path in the same time. To allocate each vertex of an item in a hierarchy, each item can be linked with each vertex based on a single linked list. Items within high frequencies will be used even though in this process the items will likely to be overlapped on each path (Mayilvate and Kalpanadevi, 2018).

The FP-Growth method is divided into three phases (Huda, 2006):

1. First phase is generating conditional pattern base which is a database part that contains the prefix path and suffix pattern. The Pattern base is obtained by processing the FP-Tree.
2. Second phase includes generating FP-tree conditional by summing the support count of each item on each conditional pattern base. For each item that has a support count greater than or equal to minimum support will be raised within this phase..
3. Third phase is the mining of frequent item set whereas if the FP-tree conditional is a single trajectory then the frequent item set is obtained by combining the items for each FP-tree conditional. Otherwise, other FP-Growth will be generated.

## 2.3 Rule Association Mining

The Rule Association Mining is also known as affinity analysis or analysis of relation. In this case it is also defined as the study of 'what is with what' or 'something has a connection to something' (Muflikhah & Ratnawati, 2018). Other study found that rule association mining is a technique or process of data mining to find the rules of association between a combination of items (Kusrini & Taufiq, 2009).

To define the rule association mining, interestingness measure (trust size) is used. This measure has two main parameters:

1. Support (supporting value) is a measure that shows how large the level of dominance of an item/itemset of the entire transaction.
2. Confidence (value of certainty) is a measure that shows the relationship between two items in conditional based. This Parameter helps to determine the interesting association rules, in order to be compared with the threshold specified by the user. This limitation is better known as min\_support and min\_confidence.

## 3. METHODS

In conducting the research, method is used to make sure it is within the line. Figure below explained the method of this

research.

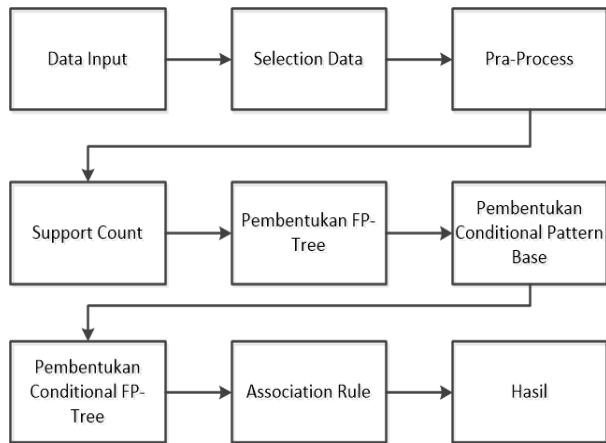


Fig 1. Research Framework

Figure above explained the method of the research using FP-Growth algorithm. First step is data collection and input followed by data selection which is together involved in the pra-process. After that, the system will make a support count to finally generate FP-Tree and conditional pattern base. Lastly, FP-Tree Conditional will be generated as well as the association rule to make output.

### 3.1 Support Count

An indication of how often item set appears in the sales transaction data is called the support count. Below is the formula to make support count in this research.

$$\text{Support (A)} = \frac{\sum \text{trans } A}{\sum \text{trans}} \times 100 \% \quad (1)$$

As for the search, support value of two items obtained by calculating using formula as follows:

$$\text{Support (A} \cap \text{B)} = \frac{\sum \text{trans } A \& B}{\sum \text{trans}} \times 100 \% \quad (2)$$

Example

$$\text{Support (2)} = \frac{2}{10} \times 100 \% = 0.2$$

### 3.2 FP-Tree Formation

FP-Tree creation uses transaction sample data in which already sorted from items within the largest number of support count to the smallest in one invoice.

Table 2. Sample Data Sales Transaction Support

No	No Invoices	Item Name
1	1905002276	Mie Campur (8), Bakso Urat + Halus (3), Mie Kuning (9), Air Mineral Biasa (1)
2	1905000156	Bakso Halus (6pcs) (2), Mie Ayam Bakso (6), Lemon Tea Panas (5)
3	1905000032	Mie Campur (8), Kerupuk (4), Bakso Halus (6pcs) (2), Mie Ayam Polos (7)
4	1905000362	Mie Campur (8), Bakso Urat + Halus (3), Kerupuk (4)
5	1905000636	Bakso Urat + Halus (3), Mie Putih (10)

6	1905000033	Kerupuk (4), Mie Putih (10)
7	1905000400	Mie Ayam Bakso (6), Air Mineral Biasa (1)
8	1905000271	Mie Campur (8), Bakso Urat + Halus (3), Bakso Halus (6pcs) (2), Mie Kuning (9), Lemon Tea Panas (5)
9	1905000602	Mie Campur (8), Bakso Halus (6pcs) (2)
10	1905000050	Bakso Urat + Halus (3), Kerupuk (4), Mie Kuning (9), Mie Ayam Polos (7)

### 3.3 Conditional Pattern Base

At this stage, calculation will be made by way of up-tree or determined by branches of trees that have the smallest number of passes. The Conditional pattern base is a sub database containing the prefix and suffix pattern.

### 3.4 Conditional FP-Tree

The Conditional pattern base is a sub database containing the prefix and suffix pattern. At this stage the support count of each item on the conditional pattern base is added, then each item that has larger number of support counts equals the minimum support count will be raised with a FP-Tree Conditional.

### 3.5 Association Rule

In this stage use the confidence formula to create the rule within confidence pattern of  $\geq 0.75$  is used. Here is the formula to find the confidence pattern:

$$\text{Confidence (A U B)} = \frac{\text{support count } A \cup B}{\text{support count } A} \quad (3)$$

## 4. RESULTS AND DISCUSSION

Based on the methods and calculation made, results are as follows.

### 4.1 Support Count

Table 3. Results of Support Count

Item Name	Frequency	Support Value
Bakso Halus (6pcs) (2)	4	0.4
Bakso Urat + Halus (3)	5	0.5
Mie Ayam Polos (7)	2	0.2
Mie Ayam Bakso (6)	2	0.2
Kerupuk (4)	4	0.4
Mie Kuning (9)	3	0.3
Mie Putih (10)	2	0.2
Mie Campur (8)	5	0.5
Air Mineral Biasa (1)	2	0.2
Lemon Tea Panas (5)	2	0.2

As shown by the table above, once the support count value is known, data will be sorted from the biggest support value to the smallest in the invoice number.

### 4.2 FP-Tree Formation

FP-Tree is formed from the transaction table that has been sorted by the result of the support count.

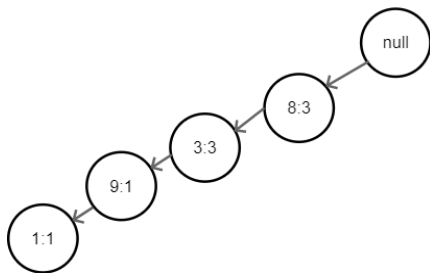


Fig 2. FP-Tree Invoice 1905002276

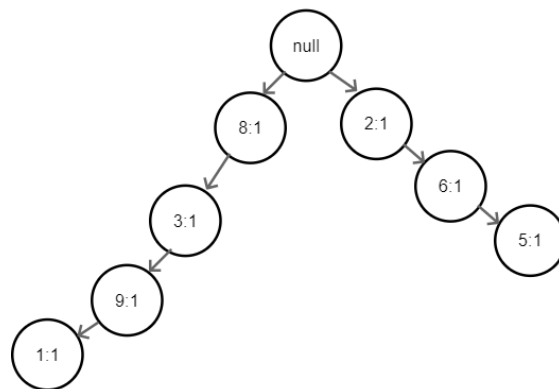


Fig 3. FP-Tree Invoice 1905000156

FP-Tree is generated until the last invoice data, following the final result of FP-Tree sales transaction of SGRMR.

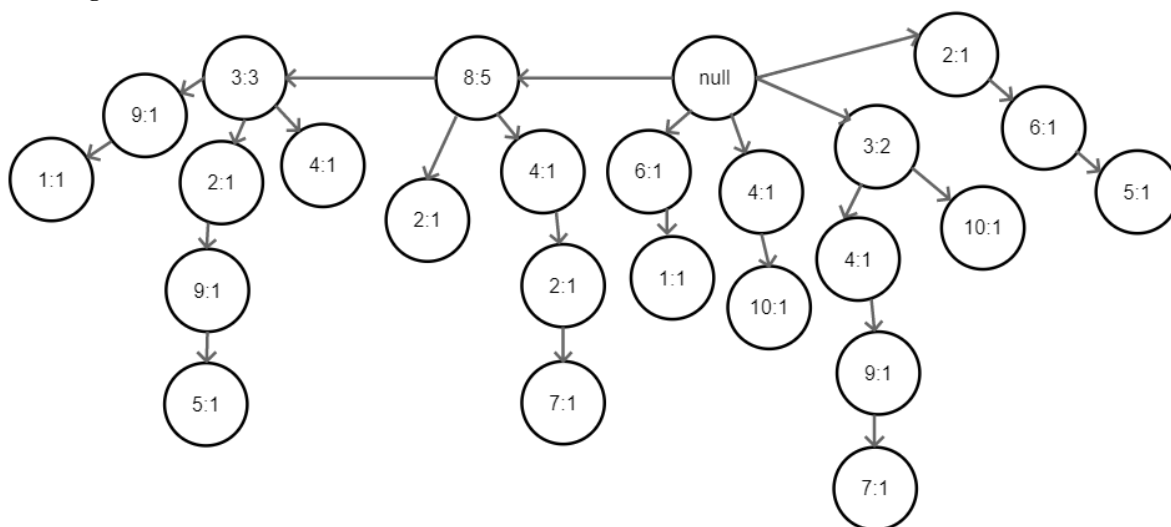


Fig 4. FP-Tree End Results

### 4.3 Conditional Pattern Base

The conditional pattern base is resulted from the FP-tree scanned by the smallest support values which is 5, 9, 2, 4, 3 and 8.

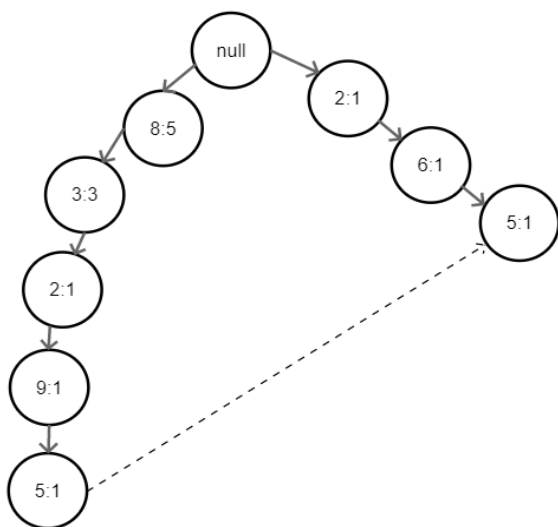


Fig 5. End of vertex 5 trajectory

Table below shows the final result of the conditional pattern base.

Table 4. Result Conditional Pattern Base

Item	Conditional Pattern Base
5	{8, 3, 2, 9 : 1}, {2, 6 : 1}
9	{8, 3 : 1}, {8, 3, 2 : 1}, {3, 4 : 1}
2	{8, 3 : 1}, {8 : 1}, {8, 4 : 1}
4	{8, 3 : 1}, {8 : 1}, {3 : 1}
3	{8 : 3}

Table 4 is the end result of a conditional pattern base that contains an item's final value and any vertex value passed from the final item to the starting point.

#### 4.4 Conditional FP-Tree

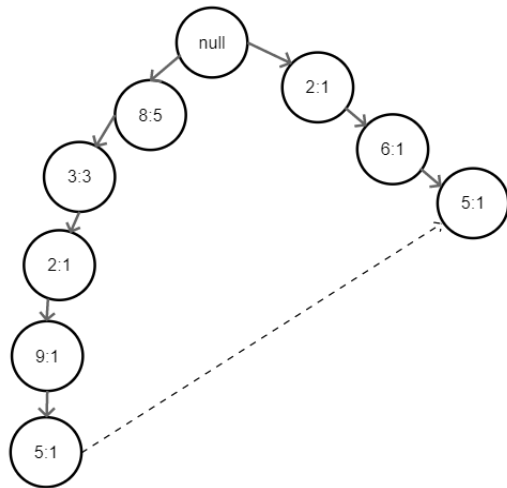


Fig 6. Vertex 5

FP-Tree conditionals are performed by removing the end value from Vertex 5.

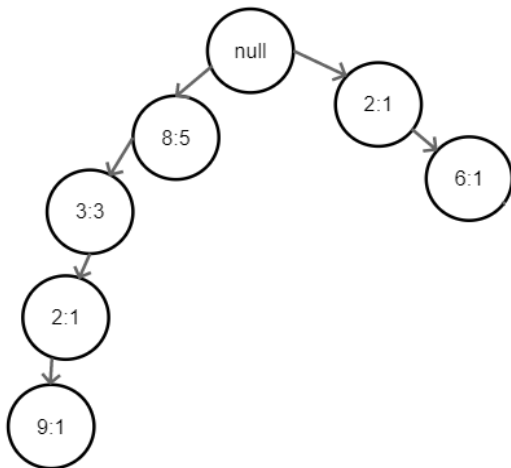


Fig 7. FP-Tree Conditional 5

After Vertex 5 is eliminated then recalculated will be made to the support value above vertex 5, corresponding to the occurrence value of transaction value with item 5 at the same time.

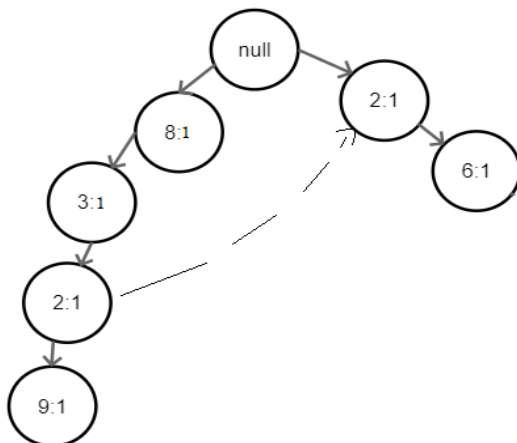


Fig 8. FP-Tree Conditional results of knot 5

Value of support count in each conditional pattern base is collected and compared with the minimum value of support which has the value of 2. If the value in the item is less than 2 then the item can be eliminated. Items that are not frequent are 8, 3 and 9. Thus, for FP-Tree conditional, the vertex 5 is the value of item 2. Here is the final FP-tree conditional result for the restaurant sales transaction data.

Table 5. FP-Tree Conditional Results

Item	Conditional FP-Tree
5	{2 : 2}
9	{8 : 2}, {3 : 3}
2	{8 : 3}
4	{8 : 2}, {3 : 2}
3	{8 : 3}
8	{8:5}

#### 4.5 Association Rule

This process is for calculating the confidence value  $\geq 0.75$  for the frequent item set. Here is the calculation of the patterns obtained by the frequent item set:

- Frequent item set {5 → 2}  
 $\{5 \rightarrow 2\} = \frac{2}{2} \times 100\% = 1$   
 $\{2 \rightarrow 5\} = \frac{2}{4} \times 100\% = 0.5$
- Frequent itemset {8 → 9}  
 $\{8 \rightarrow 9\} = \frac{2}{5} \times 100\% = 0.4$   
 $\{9 \rightarrow 8\} = \frac{2}{3} \times 100\% = 0.67$
- Frequent item set {3 → 9}  
 $\{3 \rightarrow 9\} = \frac{3}{5} \times 100\% = 0.6$   
 $\{9 \rightarrow 3\} = \frac{3}{3} \times 100\% = 1$

After calculating all confidence on the item then the next process is to check each item unmeet the criteria of confidence value. For items that is unmeet the confidence value will be deleted.

Table 6. Rule Results formed

Rule	Confidence
{Lemon Tea Panas} → {Bakso Halus (6pcs)}	1.0%
{Mie Kuning} → {Bakso Urat + Halus}	1.0%

From confidence calculation is obtained two consumer habits patterns in purchasing goods that meet the requirements of confidence  $\geq 0.75$  namely: {5 → 2} and {9 → 3}.

## **5. CONCLUSIONS AND SUGGESTIONS**

### **5.1 Conclusions**

1. Based on the use of data mining with the algorithm method FP-Growth selected data successfully processed and generate a frequent item set, from the value of the frequent item set generated formed association rule.
2. Selected data that has form of frequent item set is given minimum support value and minimum confidence value to generate association rule.
3. Menu package recommendation successfully formed with the search minimum value of support 2 and minimum confidence of 75% and produce a mixture of package menu between hot lemon tea with fine meatballs (6 pieces).

### **5.2 Suggestions**

1. Further development is expected to recommend a menu package based on the sales transaction data for each customer.
2. Future development is expected to reduce the price of package menu created using the time series analysis method.

## **6. REFERENCES**

- [1] Adinugroho, S. and Sari, Y. A. (2018). Implementasi Data Mining menggunakan Weka. Malang: UB Press.
- [2] Astrina, I., Arifin, M. Z., and Pujianto, U. (2019). Penerapan Algoritma FP-Growth dalam Penentuan Pola Pembelian Konsumen Pada Kain Tenun Medali Mas. *Jurnal Matrix*, 9(1).
- [3] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases" (PDF). Retrieved 17 December 2008.
- [4] Huda, M. 2006. *Algoritma Data Mining: Analisis Data Dengan Komputer*. Bisakimia.
- [5] Kusriani (2007). *Strategi Perancangan Dan Pengelolaan Basis Data*. Yogyakarta : Penerbit Andi
- [6] Larose Daniel T (2005). *Discovering knowledge in data : an introduction to data mining*. Canada : Wiley Inter science
- [7] Mayilvaganan, M. and Kalpanadevi, D. (2018). Comparison of Apriori, FP-Tree Growth and Fuzzy FP-Tree Growth Algorithm for Generating Association Rule Mining of Cognitive Skill. *Engineering Research and General Science*, 6(2). 2091-2730.
- [8] Mei, Rizki. 2010. Perbandingan Algoritma Apriori dan Algoritma Fp-Growth Untuk Perekomendasi Pada Transaksi Peminjaman Buku di Perpustakaan Universitas Dian Nuswantoro. Skripsi, Prodi Teknik Informatika : Universitas Dian Nuswantoro Semarang.
- [9] Muflikhah, L, Ratnawari, D, E, and Rekyan, R. 2018. *Data Mining*. Universitas Brawijaya Press
- [10] Susanto, Sani and Dedy Suryadi. 2010. *Pengantar data mining menggali pengetahuan dari bongkahan data*. Yogyakarta: Andi Offset.
- [11] Vuldari, R. T. (2017). *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Gava Media.