JPPI Vol 9 No 2 (2019) 151-159

Jurnal Penelitian Pos dan Informatika

32a/E/KPT/2017

e-ISSN 2476-9266
p-ISSN: 2088-9402

Doi:10.17933/jppi.2019.090206

DESIGNING AN APPLICATION FOR ANALYZING CONSUMER SPENDING PATTERNS USING THE FREQUENT PATTERN GROWTH ALGORITHM

PERANCANGAN APLIKASI UNTUK ANALISA POLA BELANJA KONSUMEN MENGGUNAKAN METODE FREQUENT PATTERN GROWTH ALGORITHM

[1]Wisda, [2]Mashud

[1]Program Studi Sistem Inforrmasi , [2]Program Studi Komputerisasi Akuntansi STMIK AKBA
Jalan Perintis kemerdekaan Km. 9 No.75 Tamalanrea Makassar 90245, Indonesia
wisda@akba.ac.id

Naskah Diterima: 29 August 2019; Direvisi : 26 November 2019; Disetujui : 17 December 2019

Abstract

In this modern era, the market has been growing rapidly which can be seen from the navel shopping that is lined up in the hearts of big cities such as supermarkets, grocery stores and others that are provided to meet people's needs for primary goods that are always needed at all times. One of them is Giant Express Tamalanrea, a supermarket in the city of Makassar that serves the sale of household goods and general needs. With the use of customer data analysis to determine the customers' purchasing patterns, Giant Express can optimize the collation of goods, by positioning goods at closer shelves based on the level of frequency of goods purchased together by customers. Therefore, this study suggests the creation of an application to analyze consumer spending patterns using the frequent pattern growth algorithm method to ensure appropriate placement of goods to increase sales at Giant Express Tamalanrea. The purpose of this study is to develop an application that can analyze consumer spending patterns to increase sales by positioning goods based on consumer shopping patterns, as well as implementing the Frequent Pattern Growth Algorithm method to determine customer spending patterns to increase sales. Stages of research methods conducted begin with data collection at the study site, system requirements analysis, system design with UML, and system testing with the Black Box method.

Keywords: Sales, Patterns, Consumers, Applications, Web

INTRODUCTION

In companies engaged within the retail industry, transaction data is the most important data that can store information on top selling goods. Giant Express Tamalanrea was no exception. Its daily transaction data which reaches up to 300 customers can be used to analyze customers' spending patterns. Results of the analysis can be used to optimise layouts for types and categories of goods.

Products positioning at this grocery store is arranged based on their categories. Items of kitchen utensils groups, for example, are placed close together. Meanwhile, on the adjacent shelves are goods of women's need. By analyzing the shopping patterns of Giant Express customers, the item sets that are bought frequently can be identified. Customers' preferences on purchasing goods of the same category can be seen by analyzing the customer's shopping basket

Data mining is used to classify existing transaction data into good categories data, while the customers purchasing patterns were analysed using the frequent pattern growth algorithm. The frequent pattern growth algorithm is developed from Apriori method, which is one of the alternatives in determining the most frequently appearing data set (frequent itemset) in a data set by generating a Tree data structure or called the Frequent Pattern Tree.

Several previous related studies have been conducted concerning this matter. In order to develop this research, a literature review was conducted.

Frequent pattern growth algorithm (FPGA) was applied in previous researches to examine the relationship between sales data of chemical products and inventory. Therefore, sales patterns of chemical products were established so that set of chemicals products purchased together were identified. This allowed the establishment of patterns of set of chemical products frequently purchased together which was useful in ensuring supply of related set of goods and reducing purchase of chemicals that are not frequently purchased together. In this case, more effective optimization of chemical products supply can be achieved. {Erwansyah, 2019}

Another study in PDIndo Sarana Utama investigated the case of recommendations for the purchase of goods. The study resulted in a valid AssociationRule output of 16 with minimum support (minsupp) of 0.03 and minimum confidence (mincof) of 0.2. In addition, the initial association Rule produced consisted of 35 rules and were tested based on a life value of> 1. The study resulted 16 valid rules which can be used as a strategy for sales and recommendations for purchasing goods in PDIndo Sarana Utama.

Another research conducted related to the sale of goods with the apriori method. results of the apriori algorithm can be used as a consideration in developing marketing strategies, optimizing sales, as well as consideration for restocking. {Aditya et al., 2016}.

By analysing customer data to determine the customers' purchasing patterns, Giant Express can optimize the positioning of goods. Goods can be displayed in closer shelves based on the item sets frequently brought together by displaying goods in closer shelves based on the item sets frequently brought together. For example, when an analysis on customers' shopping basket revealed that the percentage for spice products to be purchased together with toiletry products categories is higher, the two categories of products shall be placed on adjacent shelves.

To achieve the objective of this research to develop an applications that can read consumers' spending patterns, the Frequent Pattern Growth algorithm method is applied. The method is used to determine the customers' preferences in buying certain item set together. When customers' preferences are identified, item positioning can be arranged accordingly in Giant Express to increase sales..

Spending Patterns

Spending Pattern means a definite working system or structure form. Consumer spending pattern can be interpreted as a definite form of structure of consumer spending activities. The consumer spending pattern can be used as a consideration for decision makers in developing a more effective strategy. In this case, the Pattern comprises 2 parts, namely the Left Hand Side (LHS) and the Right Hand Side (RHS). LHS is a collection of one or more attribute values or items, while RHS is an attribute or a single item.

Data Mining

Data mining is a term used to describe the step of knowledge discovery in databases. Data mining is a semi-automatic process that uses statistical, mathematical, artificial intelligence, and machine learning to extract and identify useful and beneficial potential knowledge information that is stored in large databases. {Kusrini, 2009}.

According to Gartner Group, data mining is a process of finding meaningful relationships, patterns, and trends by examining in a large set of data stored in storage by using pattern recognition techniques such as statistical and mathematical techniques. Data mining is an analysis of observing data sets to find relationships that are unexpected and summarize data in a different way than before, which can be understood and useful for the data owner. {Mardi, Y. 2017}.

Frequent Pattern Growth Algorithm

Frequent Pattern Growth (FP-Growth) is one of the algorithms that can be used to determine the set of data that most frequently appears (frequent itemset) in a data set. The FP-Growth algorithm is a development of the Apriori algorithm. So the weaknesses of the Apriori algorithm are corrected by the FP-Growth algorithm. {Gunadi and Sensuse, 2016}.

FP-Growth uses the concept of tree development in searching for frequent itemsets. Therefore, the FP-Growth algorithm are faster than the 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.

METHODOLOGY

Research method used in this study is the applied method. Meanwhile, to create a web-based applications, Frequent Pattern Growth Algorithm method is used.

The stages of FP-Tree development can be describes as follows.

1. Data sets are analyzed to determine the support count of each item. Infrequent items are discarded, whereas frequent items are sequenced decreasingly based on support count. In figure 2.25, the sequence is a, b, c, d, and e.
2. After reading the first transaction, {a, b}, nodes a and b are created. Path created from null → a → b to encode transactions. Each node that passes

through this path has a frequency calculation of one.

3. After reading the second transaction, {b, c, d}, nodes b, c, and d are created. The path was created from null → b → c → d to encode transactions. Each node that passes through this path has a frequency calculation of one. Although node B is present in the first transaction and the second transaction, the path is not joined because it does not have a similar prefix.
4. The third transaction, {a, c, d, e}, has a prefix similar to the first transaction. Therefore, the path for the third transaction, null → a → c → d → e, overlaps with the first transaction path. The frequency calculation at node a becomes two, while the frequency calculation at node c, d, and e is one.
5. This process continues until all transactions are mapped into the path.

Data Collection Technique

In data collection stage, information or data related to research problems are collected. Two following approaches were used to obtain such information:

1. *Literature review*

Gathering sources of literature as references to support the topic and serves as a more convincing theoretical foundation. The reading sources that can be used as references are in the form of text books, final assignments, programming learning manuals, and soft copy reading sources obtained from internet.

2. Experiments and observations

Conducting experiments on web-based applications by filling in the data collected and observing the results of the calculation process using the frequent pattern growth method.

Designing Stage

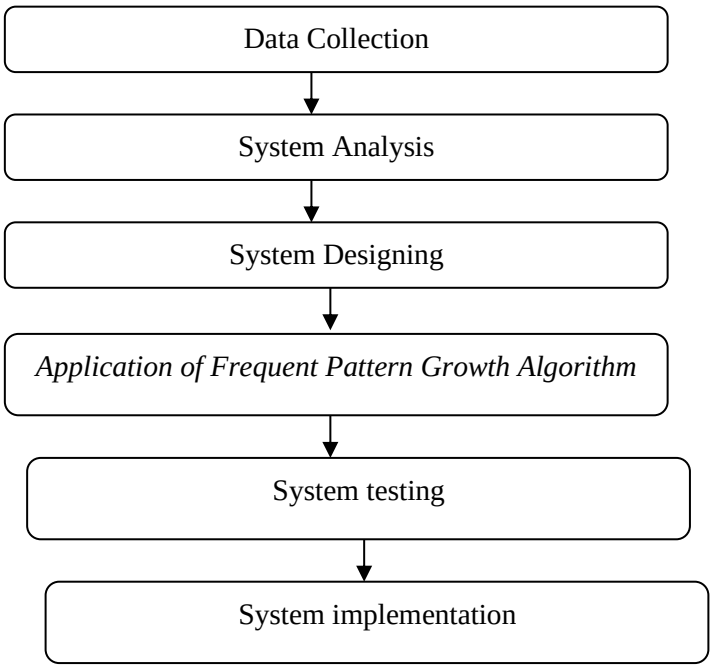

Figure I: Designing Stage

1. Data collection means gathering all data relating to the required research topic.

2. System Analysis means analyzing the system to find out what is required by the system to be designed, such as software, hardware or operating systems and the likes.

3. System designing means how to design a system that will be built based on the results of previous analysis. At this stage, required concepts and functions are designed.

4. Application of the Frequent Pattern Growth Algorithm Method which will process categorized transaction data to produce information about the monthly sales data.

5. System testing means testing the performance of a system that has been designed by using the Black Box testing method to determine whether the system designed is feasible to use.

6. System Implementation is an abstraction from the implementation of a software system.

System designing

The application is designed using UML

(Unifield Modeling Language) which contains an activity carried out to design an application that has a logically arranged algorithm, starting from collecting data needed for designing to the development of application.

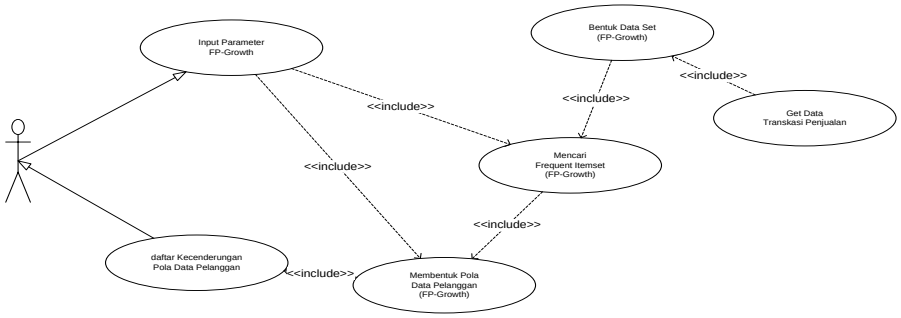

Figure 1 :*Use Case Diagram*

Figure 1 shows the use case diagram where the administrator can log in by inputting user and password, after successfully logging in, the administrator enters the main page. In this case, the administrator can input the data categories of goods, data of goods, transaction data. And can see a graph of consumer spending patterns per month and do the configuration.

.

RESULTS AND DISCUSSION

1. Table 1 shows transaction data to be processed. from this data, item set will be generated from the categories of goods in transactions occured.

Table 1 : Transaction Data

No	No. of receipt	Barcode	Item Name	Category	Total
1	1247306	383812473065	IndomieAyamBawang 69g	G001	2
2	1247306	899290711001	Sari Roti Tawar	G001	1
3	1247306	899270200001	IndomilkSkmPth 370 Gr	G003	1
4	2606638	899290711001	Sari Roti Tawar	G001	3
5	2606638	899516340030	Gt GulaPasirLkl Pls 1kg	G001	2
6	2606638	955600122756	Kit Kat 2f Green Tea 17G	G005	2
7	2375833	899290711001	Sari Roti Tawar	G001	1
8	2375833	899516340030	Gt GulaPasirLkl Pls 1kg	G001	2
9	2375833	899886620030	Sedaap Mie Goreng 90GR	G001	1
10	2375833	899275310120	BenderaKental Manis	G003	3
11	2375833	899269640444	Bear Brand 189ML	G005	2

No	No. of receipt	Barcode	Item Name	Category	Total
12	1660727	899290711001	Sari Roti Tawar	G001	2
13	1660727	899886620030	Sedaap Mie Goreng 90GR	G001	2
14	1660727	899275310120	BenderaKental Manis	G003	1
15	1660727	899294652811	Dancow 1Thn 400GR	G010	2

2. generating item set of category of goods

Tabel 2 : Item set of category of goods

No	Kode	Itemset
1	G001	Stapple-food
2	G002	fresh & frozen
3	G003	beverages
4	G004	Breads
5	G005	Snacks
6	G006	bodycare
7	G007	Housecare materials and utensils
8	G008	Laundry and washing materials and utensils
9	G009	Stationary and Office supplies
10	G010	formula

Table 3: Item set generation

TID	ITEM	No. receipt
1	G001,G003	1247306
2	G001,G005	2606638
3	G001,G003,G005	2375833
4	G001,G003,G010	1660727
5	G002,G005	1776442
6	G002,G004	2100832
7	G001,G003,G007	2381592
6	G006,G007	3037782
8	G006,G008	1108752
9	G001,G003,G008	2463936
10	G001,G004,G009	2239967

3. Formation of FP - Tree Category of Goods

3. Formation of FP - Tree Category of Goods based on the above itemset table.

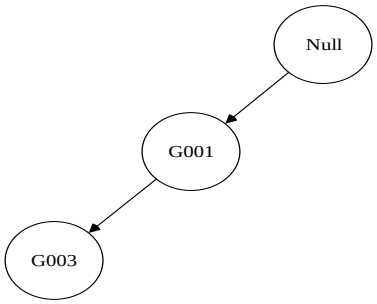

Figure 4 formation of FP-Tree TID 1

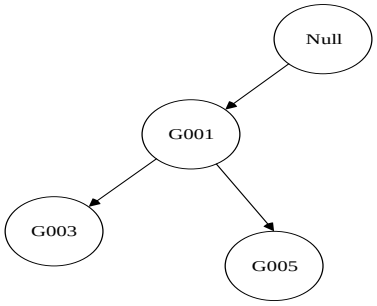

Figure 5 Formation of FP-Tree TID 2

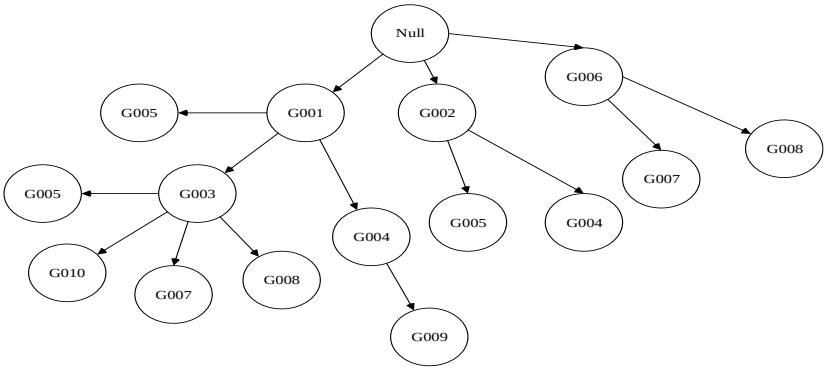

Figure 6. Formation of FP-Tree TID 10

From the figure of the formation of FP-Tree 10, the following data is acquired

Table 3.Number of Item in itemset

No	Code	Itemset	Number
1	G001	Stapple-food	7
2	G002	fresh & frozen	2
3	G003	beverages	5
4	G004	Breads	2
5	G005	Snacks	3
6	G006	bodycare	2
7	G007	Housecare materials and utensils	2
8	G008	Laundry and washing materials and utensils	2

No	Code	Itemset	Number
9	G009	Stationary and Office supplies	1
10	G010	formula	1

3. Calculation of *Support* and *Confidence* for category of goods

Table 4: Number of Item in Itemset

Category	Number of transaction	Support
G001→ 7	10	0.7
G002→ 2	10	0.2
G003→ 5	10	0.5
G004→ 2	10	0.2
G005→ 3	10	0.3
G006→ 2	10	0.2
G007→ 2	10	0.2
G008→ 2	10	0.2
G009→ 1	10	0.1
G010→ 1	10	0.1

4. Calculation of *Confidence* value to G001

Table 5. G001 *Confidence* value

category	Number of G001	Support
G003→ 5	7	0.72
G004→ 2	7	0.28
G005→ 3	7	0.42
G007→ 2	7	0.28
G008→ 2	7	0.28
G009→ 1	7	0.14
G010→ 1	7	0.14

From the results of calcualtion of *confidence* in table 5, it is concluded that Consumers' preferences in purchasing G003 (beverages) together with G001 (stapple-food) with support value = 0.5 and confidence value of 0.72, which is the highest compared to other categories of goods. Therefore, the study suggests that "**stapple-food**" items are positioned in adjacent shelves with "**beverages**" category.

CONCLUSIONS

The implementation of the frequent pattern growth method can identify the consumers' spending patterns . This pattern can be used as a reference in recommending the shelving arrangement of categories of goods at Giant Express Tamalanrea. Frequent pattern growth algorithm method can produce a pattern of consumer spending that can be used as a guideline for the positioning of items so as to increase sales.

Development can be made by implementing other methods in providing recommendations for consumer spending patterns. To improve this system, following research may include other criteria or variables as well as adding more categories of analyzed goods of more than two categories of goods.

ACKNOWLEDGEMENTS

Researches extend their appreciations to chairman of P3M STMIK AKBA for their support in the study. Appreciation also goes to colleagues who have supported the writings process of this paper.

REFFERENCES

Aditya, A., Marisa, F., & Purnomo, D. (2016). Penerapan Algoritma Apriori Terhadap Data Penjualan di Toko Gudang BM. *Journal of Information Technology and Computer Science*, *1*(1).

EkoPrasetyo. (2014_)“ Data *Mining Mengolah Data Manjadi Informasi Dengan Matlab”*. Penerbit Andi, Yogyakarta.

Erwansyah, K. (2019). Implementasi Data Mining Untuk Menganalisa Hubungan Data Penjualan Produk Bahan Kimia Terhadap Persedian Stok Barang Menggunakan Algoritma FP (Frequent Pattern) Growth Pada PT. Grand Multi Chemicals. *J-SISKO TECH (Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD)*, *2*(2), 30-40.

Gunadi, G., & Sensuse, D. I. (2016). Penerapan metode data mining market basket analysis terhadap data penjualan produk buku dengan menggunakan algoritma apriori dan frequent pattern growth (fp-growth): studi kasus percetakan pt. Gramedia. *Telematika MKOM*, *4*(1), 118-132.

Kusrini, dan EmhaTaufikLuthfi, (2009), *Algoritma Data Mining*”, Penerbit Andi,Yogyakarta.

Mardi, Y. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4. 5. *Edik Informatika*, *2*(2), 213-219.

Sasono, A., Amalia, A. E., & Ramadhani, R. D. (2018). Implementasi Algoritma Frequent Pattern Growth (FPGrowth) untuk Rekomendasi Pembelian Barang (Studi Kasus: PD. Indoapi Sarana Utama). In *Conference on Electrical Engineering, Telematics, Industrial technology, and Creative Media (CENTIVE)* (pp. 150-156).

Sumangkut, K., Lumenta, A. S., & Tulenan, V. (2016). Analisa Pola Belanja Swalayan Daily Mart Untuk Menentukan Tata Letak Barang Menggunakan Algoritma FP-Growth. *Jurnal Teknik Informatika*, *8*(1).

Yuyun Dwi Lestari, 2015, “*Penerapan Data Mining Menggunakan Algoritma Fp-Tree Dan Fp-Growth Pada Data Transaksi Penjualan Obat* “, Snastikom, Hal 61 – 65.