Product Bundling Application to Support Implementation of Association Rule Mining Using FP-Growth

Iin Ernawati[1*], Nurhafifah Matondang[2]
[1,2]Educational Information System, Universitas Pembangunan Nasional Veteran Jakarta, Indonesia
Jl. RS Fatmawati No.1, Pondok Labu, Jakarta Selatan
Corresponding Author: iinerti@gmail.com

INTISARI

Media online telah terbukti memberikan dampak yang baik khususnya dalam bidang bisnis, untuk membuka dan memperluas akses masyarakat dan masyarakat, sehingga memberikan peluang lebih luas dan mudahnya akses bisnis melalui perdagangan elektronik (*e-commerce*). Hal inilah yang menjadi dasar pemikiran perlunya peningkatan jangkauan bisnis toko tanaman hias di wilayah Tangerang. Pola transaksi asosiasi sebagai hasil dari proses data mining dengan mengimplementasikan Fp-Growth Algorithm yang disuntikkan ke aplikasi mobile dengan tujuan untuk mendapatkan pola belanja pembeli, yang juga dikemas sebagai media promosi bundling produk dan akhirnya membuka peluang untuk memperluas jangkauan bisnis dan informasi. Eksperimen yang dilakukan dengan menggunakan nilai minimum support dan confidence yang berbeda yaitu (1% dengan 10%, 2% dengan 20%, 3% dengan 25%, dan 4% dengan 30%) memberikan hasil berupa tujuh aturan asosiasi yang diperoleh dari eksperimen dengan dukungan minimum 4% dan kepercayaan minimum 30%). Aturan asosiasi yang dihasilkan menunjukkan bahwa pembeli (pelanggan/non-pelanggan) cenderung membeli tanaman dan media tanamnya.

Kata kunci: E-Commerce, Association Pattern, FP-Growth, Product Bundling

ABSTRACT

Online media has been proven to gave good impact especially in the business field, for opening and widen access to people and society, and so giving more opportunity to any business broader and easy access through electronic commerce (e-commerce). This is the basic idea for the need to improve business reach for a decorative plants store in Tangerang Area. The association transaction pattern as the outcome of data mining process by implementing the Fp-Growth Algorithm that is injected to a mobile application with the purpose to get the buyer's shopping pattern, which is also packed as the products bundling promotion media and finally open the opportunity to broaden reach for business and information. Experiments carried out using different minimum support and confidence values, which are (1% with 10%, 2% with 20%, 3% with 25%, and 4% with 30%) is giving results in the form of seven association rules obtained from experiments with a minimum support 4% and minimum confidence 30%). The resulting association rules show that buyers (customers / non-customers) tend to buy plants and their growing media.

Keywords: E-Commerce, Association Pattern, FP-Growth, Product Bundling

INTRODUCTION

Every company has data stored in its database. The transaction data is increasing day by day. With the increasing amount of data in the company, the role of the analyst to analyze data manually needs to be replaced with computer-based applications. So that the analysis process can be carried out precisely and accurately. This study applies the FP-Growth algorithm in an application that can find out consumer buying patterns at each different branch with different characteristics. From the resulting pattern, information will be obtained.

Association analysis is also known as one of the data mining techniques which is the basis of various other data mining techniques. In particular, one of the stages of association analysis called high frequency pattern analysis (frequent pattern mining) has attracted the attention of many researchers to produce efficient algorithms. The importance of an associative rule can be determined by two parameters, support (supporting value), namely the percentage of the combination of these items in the database and confidence [5].

Ecommerce

E-commerce is a term that is often used or heard today that is related to the internet, where no one clearly knows the meaning of e-commerce. Electronic commerce or also known as e-commerce, is how people use communication networks and computers for business processes to be done. A popular outlook for e-commerce is the use of the internet and a computer to buy and sell products using web browser [6].

The benefits of using E-commerce in a company as a transaction system are:

a. Can increase market exposure (market share)
b. Reducing operating costs (operating cost)
c. Extending reach (global reach)
d. Increase customer loyalty
e. Improve supply management

A brief overview of the advantages of e-commerce is as follows:

a. For consumers: lower prices, shopping in one place.
b. For managers: efficiency, error-free, and on time.

Data mining

Data mining is an automatic search process for useful information in large data storage areas. Other terms that are often used include knowledge discovery (mining) in databases (KDD), knowledge extraction, data / pattern analysis, data archeology, data dredging, information harvesting, and business intelligence. Data mining techniques are used to examine large databases as a way to find new and useful patterns. Not all information search task are stated as data mining. For example, searching for individual records using a database management system or search for a specific web pages through queries to all search engines is an information search task that is closely related to

information retrieval. Data mining techniques can be used to improve capabilities of the information retrieval systems [4].
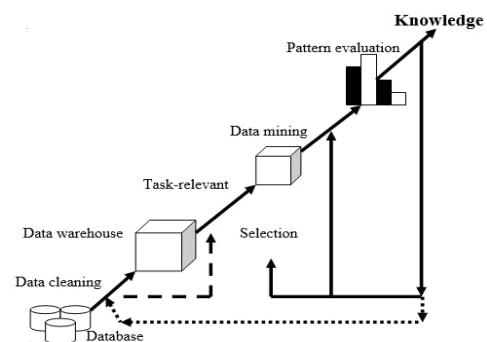

Figure 1. *Data mining* as one phase in the process of *Knowledge Discovery*

Association Rule

The definition of association rule is a convention that has an interrelated relationship between attributes used to obtain associations between data. There are several terms of association, that is the antecedent representing "if" and consequent representing "then". In the association analysis, these two things are groups of items that are not related together [4].

$$Support\ (A \cap B)$$
$$= \frac{\text{Transaction Amount contain item A and B}}{\text{Total Amount Transaction}}$$

Min. support Count = Min. support x Total Amount Transaction

$$Confidence\ (A \rightarrow B)$$
$$= \frac{Transaction\ Amount\ with\ item\ A\ \&\ B}{Transaction\ Amount\ with\ item\ A}$$

FP-Growth

Frequent Pattern Growth (FP-Growth) is an alternative algorithm that can be used to determine the most frequently occurring data set (frequent itemset) in a data set [2]. The FP-Growth algorithm is a development of the Apriori algorithm. So that the shortcomings of the Apriori algorithm are corrected by the FP-Growth algorithm [2].

FP-Growth uses the concept of tree construction in searching for frequent itemsets. This is what causes the FP-Growth algorithm to be faster than the Apriori algorithm. The characteristic of the FP-Growth algorithm is that the data structure used is a tree called FP-Tree. By using FP-Tree, FP-Growth algorithm can directly extract frequent itemset from FP-Tree.

FP-tree is a compressed data storage structure. 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 paths allow for overlapping. The more transaction data that has the same item, the more effective the compression process with the FP-tree data structure will be.

Mining of frequent itemsets using the FP-Growth algorithm will be carried out by generating a tree data structure (FP-Tree). The FP-Growth method can be divided into 3 main stages, namely as follows [4]:

a. Generation Phase conditional pattern base Conditional Pattern Base is a subdatabase that contains prefix path (prefix path) and suffix pattern (suffix pattern). The conditional pattern base generation is obtained through the previously built FP-tree.
b. FP-Tree conditional generation stage At this stage, the support count of each item in each conditional pattern base is summed, then each item that has a greater support count equal to the minimum support count will be generated with the conditional FP-tree
c. Frequent itemset search stage If the conditional FP-tree is a single path, then the frequent itemset is obtained by combining items for each conditional FP-tree. If it is not a single path, then the FP-growth generation is done recursively.

Table 1. Sample of Association Implementation

Transaction ID	Goods item sold
T1	{CJP Kayu 820 Putih 25 cm, Media Hijau }
T2	{CJP Kayu putih 25 cm, NKT 20 PINK}
T3	{Grace 30 Polos Hitam, Rajawali}
T4	{Media Hijau, Grace 30 Polos Hitam }
T5	{NKT 20 PINK, Pohon 40}

Android

Android is an operating system for mobile devices that includes an operating system, middleware and applications. Android provides an open platform for developers to create their applications. Android as "the first complete, open, and free mobile platform" Complete (Complete Platform): android is a safe operating system and provides many tools in building software and allows for open application development opportunities (Open Source Platform) : developers can freely to develop Free applications (Free Platform): android

is a platform / application that is free to develop. There is no license or royalty fee to be developed on the android platform.

METHOD

Formation of a Header Table for the set of items that meet the minimum support value requirements. Here are some of them:

a. Minimum support 3%, there are 18 items and 5 of them are :

Table 2. Minsup 3%

No	Id item	Jns barang	N(A)
1	1075	Pakis gantung	214
2	1893	Rajawali	528
3	1692	Lavenia 32 warna	216
4	0746	Sekam Bakar	225
5	0724	Pupuk kambing halus	224

The minimum support count of 94.59 is rounded up to 95.

b. Minimum support 4%, there are 14 items and 5 of them are :

Tabel 3. Minsup 4%

no	Id item	Jns barang	N(A)
1	1075	Pakis gantung	214
2	1893	Rajawali	528
3	1692	Lavenia 32 warna	216
4	0746	Sekam Bakar	225
5	0724	Pupuk kambing halus	224

The minimum support count of 126.12 rounded down to 126.

Formation of frequent itemset
a. Minimum support 3% and minimum confidence 25%

Table 4. Minsup 3% & mincof 25%

item	Antecedent	count	support	confidence
NKT 20 Pink ->media hijau	430	121	0.04	0.28
Pohon40 -> media hijau	353	97	0.03	0.27
Pohon40 -> NKT 20 Pink	353	105	0.03	0.30

b. Minimum support 4% and minimum confidence 30%

Table 5. Minsup 4% & mincof 30%

item	Antecedent	count	support	confidence
CJP Kayu 820 putih 25 cm -> media hijau	397	143	0.05	0.36

The experiment was carried out 16 times with several minimum support values and

minimum confidence values, the following table shows the results obtained.

Table 6. The result of association rules

Minsup	Mincof			
	10%	20%	25%	30%
1%	28 rules	22 rules	17 rules	10 rules
2%	21 rules	17 rules	14 rules	8 rules
3%	13 rules	9 rules	7 rules	2 rules
4%	2 rules	1 rule	1 rule	1 rule

The results obtained from the experiments carried out are when the minimum support value and minimum confidence value are greater, the resulting association rules will be fewer and when the minimum support value and minimum confidence value are smaller, the resulting association rules will be more and more. Experiments with a minsup value of 3% and a mincof value of 25% can be used as a reference to determine product bundling because the resulting association rules are ideal with a lift ratio value > 1 so it has the potential to be used.

CONCLUSION

Experiments carried out using different minimum support and confidence values, which are (1% with 10%, 2% with 20%, 3% with 25%, and 4% with 30%) is giving results in the form of seven association rules obtained from experiments with a minimum support 4% and minimum confidence 30%). The resulting association rules show that buyers (customers / non-customers) tend to buy plants and their growing media.

SUGGESTION

In the future, it is hoped that similar research can be carried out and supported by more transaction data, to obtain information and patterns that may arise with respect to certain events or moments that occur so that existing applications can be developed to promote events or moments. this so that it can maximize application utilization and increase more sales opportunities.

REFERENCES

[1] Faisal, Edi, dkk. "Pola Beli Konsumen Menggunakan Algoritma FP-Growth untuk Rekomendasi Promosi Penjualan pada Batik Nadya Pekalongan". *Makalah Seminar SeNTIK 2017 – STMIK JAKARTA STI&K*.

[2] Gunadi, Goldie, dan Dana Indra Sensuse. "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 Jakarta": *Jurnal Telematika MKom* Volume 4 No.1 ISSN: 2085-725X. 2012.

[3] Kurniawan, S., Gata, W. and Wiyana, H. "Analisis Algoritma FP-Growth Untuk Rekomendasi Produk Pada Data Retail Penjualan Produk Kosmetik (Studi Kasus: MT Shop Kelapa Gading)". *Seminar Nasional Teknologi Informasi dan Komunikasi 2018 (SENTIKA 2018)*, 2018(8), pp. 61–69.

[4] Han, Jiawei, dkk. "Data Mining: Concept and Techniques, Third Edition". *Waltham: Morgan Kaufmann Publishers*. 2012.

[5] Alfannisa Annurullah Fajrin, Algifanri Maulana. "PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA FPGROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR". Kumpulan jurnaL Ilmu Komputer (KLIK), 2018.

[6] Raymond McLeod. "Information Management System". Pearson. 2008.

[7] Russy Amelia, Dito Putro Utomo. "ANALISA POLA PEMESANAN PRODUK MODERN TRADE INDEPENDENT DENGAN MENEREPAKAN ALGORITMA FP. GROWTH (STUDI KASUS: PT. ADAM DANI LESTARI)". KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer) Volume 3, Nomor 1, Oktober 2019.

[8] Dyah Pramesthi Larasati , Muhammad Nasrun, S.Si., MT. , Umar Ali Ahmad, ST., MT. " Analisis Dan Implementasi Algoritma Fp- Growth Pada Aplikasi Smart Untuk Menentukan Market Basket Analysis Pada Usaha Retail (Studi Kasus : Pt.X)". e‑Proceeding of Engineering: Vol.2, No.1 April 2015.

[9] Yuyun Dwi Lestari. "Penerapan Data Mining Menggunakan Algoritma Fp‑Tree Dan Fp-Growth Pada Data Transaksi Penjualan Obat". Seminar Nasional Teknologi Informasi dan Komunikasi (SNASTIKOM 2015).

[10] Rizky Fitria, Warnia Nengsih, Dini Hidayatul Qudsi. "Implementasi Algoritma Fp‑Growth Dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas". Jurnal Sistem Informasi (Journal of Information Systems) 2017.