

ANALISIS KINERJA ALGORITMA *FREQUENT PATTERN GROWTH* (FP-GROWTH) PADA PENAMBANGAN POLA ASOSIASI DATA TRANSAKSI

Putri Kurnia Handayani¹, Nanik Susanti²

¹Fakultas Teknik Universitas Muria Kudus

²Fakultas Teknik Universitas Muria Kudus

Email: ¹putri.kurnia@umk.ac.id, ²nanik.susanti@umk.ac.id

Abstrak

Data transaksi penjualan yang setiap hari bertambah menyebabkan banjir data dalam database. Data transaksi tersebut hanya digunakan sebagai laporan penjualan yang dicetak setiap bulannya. Data mining merupakan kegiatan menambang/menggali data untuk mengenali pola atau aturan tertentu dari sejumlah dataset yang sangat besar dan mempunyai dimensi tinggi. Asosiasi adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Pola asosiasi yang berhasil diketahui dapat membantu pihak manajemen untuk mendukung pengambilan keputusan berkaitan dengan strategi penjualan, promosi produk, *reward* bagi pelanggan dan kendali stok. Penggalan pola asosiasi menggunakan algoritma FP-Growth melalui 3 tahap, yaitu pembangkitan *conditional pattern base*, *conditional pattern tree* dan pencarian *frequent itemset*. Metode perancangan sistem menggunakan UML. Tujuan penelitian ini adalah untuk menghasilkan sebuah sistem yang dapat mengenali pola asosiasi produk pada database.

Kata kunci: *data mining, asosiasi, PR-Growth*

PERFORMANCE ANALYSIS OF FREQUENT PATTERN GROWTH (FP-GROWTH) ALGORITHM IN MINING ASSOCIATION RULES OF TRANSACTION DATA

Abstract

Sales transaction data which every day increases causing data flooding in the database. The transaction data is only used as a sales report that is printed every month. Data mining is an activity of mining / digging data to recognize certain patterns or rules from a number of very large and high dimensional datasets. Association is a data mining technique for finding the rules of a combination of items. The pattern of successful associations is known to help management to support decision making related to sales strategies, product promotions, customer rewards and stock control. The association pattern collection uses the FP-Growth Algorithm through 3 stages, namely the generation of a base conditional pattern, a conditional pattern tree and frequent itemset searches. The system design method uses UML. The purpose of this research is to produce a system that can recognize patterns of product associations in the database.

Keywords: *data mining, association, FP-Growth*

1. PENDAHULUAN

Banjir data yang terjadi hampir disemua instansi membuat data semakin menumpuk tanpa ada proses lanjutan untuk mengolah tumpukan data tersebut. Data yang tersimpan semakin banyak, sehingga memenuhi media penyimpanan tanpa ada manfaat.

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban, 2005). Definisi lain

data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Pramudiono, 2006).

Asosiasi adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Pola asosiasi yang berhasil dikenali melalui penambangan data dari sebuah dataset akan bermanfaat bagi manajemen (*top level management*) dalam mendukung pengambilan keputusan.

Ketersediaan database mengenai catatan transaksi pembelian para pelanggan suatu supermarket ataupun bidang usaha lain, telah

mendorong pengembangan teknik-teknik yang secara otomatis menemukan asosiasi produk atau item-item yang tersimpan dalam database tersebut (Listriani, et al., 2016). Salah satu penerapan *association rules* adalah *market basket analysis*, dimana bertujuan untuk menemukan bagaimana item yang dibeli pelanggan dalam supermarket atau toko saling berhubungan. Pencarian *association rules* dilakukan melalui dua tahap, yaitu pencarian *frequent itemset* dan penyusunan *rules*. Penting tidaknya *association rules* dapat diketahui dengan dua parameter, yaitu *support* (nilai penunjang) dan *confidence* (nilai kepastian). Support adalah ukuran yang menunjukkan tingkat dominasi itemset dari keseluruhan transaksi.

2. METODE PENELITIAN

Langkah-langkah yang dilakukan dalam penelitian ini adalah:

- a. Pengumpulan data
- b. Pre-processing data
- c. Perancangan
- d. Penerapan algoritma FP-Growth

Penggalian itemset yang frequent dengan menggunakan algoritma FP-Growth akan dilakukan dengan cara membangkitkan struktur data tree atau disebut dengan FP-Tree (Han & Micheline, 2000). Algoritma FP-Growth memiliki tiga tahapan utama yaitu:

- a. pembangkitan conditional pattern base
- b. pembangkitan conditional pattern tree
- c. pencarian frequent itemset. Pseudocode dari algoritma FP- Growth sebagai berikut:

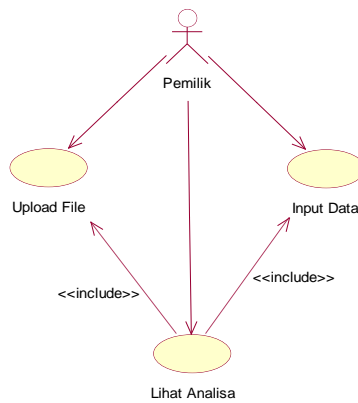
```

Procedure FP_growth(Tree,  $\alpha$ )
(1) if tree contains a single path P then
(2) for each combination (denoted as  $\beta$ ) of the nodes in the path P
(3) generate pattern  $\beta \cup \alpha$  with support_count = minimum support count of nodes in  $\beta$ 
(4) else for each  $a_i$  in the header of Tree {
(5) generate pattern  $\beta = a_i \cup \alpha$  with support_count =  $a_i.support\_count$ ;
(6) construct  $\beta$ 's conditional pattern base and then  $\beta$ 's conditional FP_tree  $Tree_{\beta}$ ;
(7) if  $Tree_{\beta} \neq \emptyset$  then
(8) call FP_growth( $Tree_{\beta}, \beta$ );}
    
```

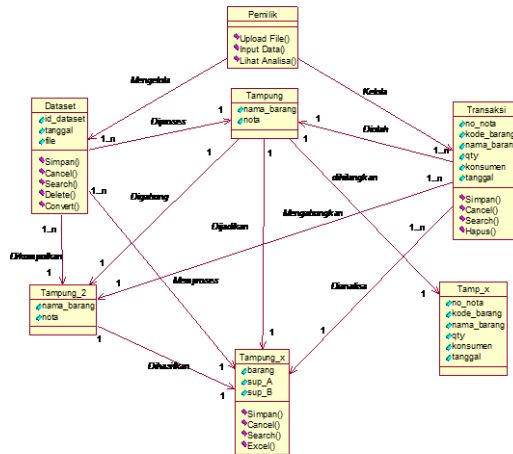
Gambar 1. Pseudocode algoritma FP-Growth

3. PERANCANGAN

Perancangan sistem menggunakan tools *Unified Modelling Language* (UML) yang terdiri dari *use case diagram*, *class diagram*, *sequence diagram* dan *activity diagram*. *Use case diagram* memberikan gambaran urutan aksi yang dilakukan oleh aktor terhadap sistem. *Class diagram* memberikan gambaran objek/kelas yang ada di dalam sistem dan hubungan asosiasi yang terjadi antar kelas.



Gambar 2. Sistem Use Case Diagram



Gambar 3. Class Diagram

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini data private, berupa data transaksi penjualan selama 4 tahun. Data transaksi tersebut akan diolah menggunakan algoritma *FP-Growth* untuk dapat dikenali pola asosiasi yang terjadi antar item. Sampel data transaksi dapat dilihat pada gambar 4.

NO	NO. NOTA	KODE BARANG	NAMA BARANG	QTY	KONSUMEN	TANGGAL
1	K1/1406/38294	KBSX441B	CD BLANK HP	10	PELANG KOMPUTER	Friday, June 27, 2014
		KBT2055	CD BLANK SPC	10		
2	K1/1406/38295	KBT1645B	PROCE G180X TRAY	1	KHOLIQ	Monday, June 30, 2014
3	K1/1406/38296	KME2022	MS NEW G41 REFUR	1	MENYING	Monday, June 30, 2014
		KST1960	OBENG ALL IN 1 KERUCUT HS-6036B	1		
4	K1/1406/38297	L3609PC	HEADSET BUFTech 208MV	1	ERFAN	Monday, June 30, 2014
5	K1/1406/38298	TNT10	CPU I3 HP TWR D3 2GB/320-500	1	KARSA COMPUTER	Monday, June 30, 2014
6	K1/1407/38299	TV798FM	ADAPTOR ACER 19V/2,1A (B)	1	KIRANA JAYA COMP	Tuesday, July 1, 2014
		UVC35D83	ADAPTOR ACER 19V/2,1A	1		
		ADH118535	ADAPTOR DELL PRISMA 19V/3,3A	1		
		ARLB20325	ADAPTOR HP JARUM 19V/3,3A	1		
7	K1/1407/38300	AHDKST2	BT HP DV2	1	MH COMP	Tuesday, July 8, 2014
		AHDOMN2001	BT LENOVO G480-8480	1		
		CBAK01	BT OEM HP DV2	1		
		CBL851F	BT ORI ACER A725/756/V5 11"	1		
		CBLR61M	BT ORI DELL 3421	1		
		CBL551M	BT ORI FUJITSU AH530	1		
8	K1/1407/38301	CBLV51M	CPU C2D DELL DESK D3 2GB/250	1	MEDIA GRAFIKA	Friday, July 11, 2014
9	K1/1407/38302	CBLV51R	CPU C2D ION TWR D2 2GB/160	1		
10	K1/1407/38303	CBLV61M	KABEL ATA	1	MOHAMMAD FADLI	Thursday, July 17, 2014
		CBL851F	KABEL CCTV SPC R659 1M	1		
11	K1/1407/38304	CBL851R	KABEL CCTV SPC R659 1ROL (300M)	1	KRIS CCTV	Saturday, July 19, 2014
12	K1/1407/38305	CBV15W	OBENG ALL IN 1 KERUCUT HS-6036B	1	MS ANAS	Saturday, July 19, 2014
13	K1/1407/38306	CD801	HEADSET APOLLON FMV	1	MERKAN INTERNET	Sunday, July 27, 2014
		H161NGP	HEADSET BUFTech 208MV	1		
14	K1/1407/38307	HDD25500	HDD ENCLD 2.5 HAYABUSA 2.0	3	PANDU COMP	Sunday, July 27, 2014
15	K1/1407/38308	KBC214B	KEY ASUS X441B	1	PRAMUKA COMP	Monday, July 28, 2014
		KB1	KEY TOSHIBA N8205 SILVER	1		
		KBSA43B	KEY TOSHIBA C640/L645 DOFF	1		
		KBS5400B	KEY +MOUSE EPAIRAZER	1		
16	K1/1407/38309	KBPJM	CAM WIFI SPC CLEVER KST1-960P	1	PASSNET	Tuesday, July 29, 2014
17	K1/1407/38310	LBE-45-33	PRINTER EPSON 1360 PSC	1	PAK YANTO	Tuesday, July 29, 2014
		NP29CM	TINTA SL BLACK	1		

Gambar 4. Data Sampel

4.2. Perhitungan Algoritma FP-Growth

Berdasarkan data transaksi yang telah dikumpulkan maka proses *preprocessing* dilakukan terhadap data sampel yang terdiri dari 100 record dari 76 transaksi pada tahun 2014 dengan minimum support 2. Hasil pembersihan data dapat dilihat pada gambar 5.

Kode Transaksi	Kode Barang	Nama Barang (Item)	Pelanggan	Tanggal
K1/1401/38188	D21GND	DDR2 1GB 2 nd	BUDI LESTARI	Wednesday, Janua 1, 2014
K1/1401/38189	ADC19158	ADAPTOR ACER 19V/1,58A	77COMP	Wednesday, Janua 1, 2014
	ADC1921B	ADAPTOR ACER 19V/2,1A (B)		
	ADMO165365	ADAPTOR MACBOOK 16,5V/3,65A		
K1/1401/38190	BTOH14H504	BT ORI HP 14 H504	ABADI KOMPUTER	Wednesday, Janua 1, 2014
K1/1401/38191	C2DT22160	CPU C2D DELL TWR D2 2GB/160	BUDI	Wednesday, Janua 1, 2014
K1/1401/38192	CBPO15	KABEL POWER PRINTER	ADHA COMPUTER	Thursday, January 2014
	BNC02	CONECTOR BNC RG6		
	AHGO20	ADAPTOR DELL JARUM 19V/3,3A		
K1/1401/38193	KBCV3571B	CAM 2MP GLENZ INDO	AD ELEKTRONIK	Saturday, January 2014
K1/1401/38194	KBCV5121B	KEY ACER V3 571 B	ADI S	Saturday, January 2014
K1/1401/38195	KBCV5471B	KEY ACER V5-121B 11.6"	BUDI SUKUN	Saturday, January 2014
K1/1401/38196	AHDKST2	KEY ACER V5-471 B	EKO PRASETYO	Saturday, January 2014
K1/1401/38197	TNT04	ADAPTOR HP BIASA 18,5V/3,5A	EL IHWAN	Saturday, January 2014
K1/1401/38198	CBL851R	CAM WIFI SPC EXPERT KST2	CEMBAN CELL	Sunday, January 5, 2014
K1/1401/38199	BTS1025	TINTA KOREA 100ML YELLOW	CAKRAWALA COMPUTER	Sunday, January 5, 2014
K1/1401/38200	AGM5HP23	KABEL LAN BELDEN CAT5 1ROL	CAHAYA NET	Sunday, January 5, 2014
K1/1401/38201	BTOSX550E	BT ASUS 1025	CERIA COMPUTER	Monday, January 6 2014
K1/1401/38202	CNVRTFLX	ADAPTOR ACER 19V/2,1A (B)	ERFAN	Monday, January 6 2014
	ADC1921B	ADAPTOR ACER 19V/2,1A (B)		
K1/1401/38203	CR4	ADAPTOR LENOVO 20V/3,25A USB	ERWIN UNGU	Monday, January 6 2014
K1/1401/38204	D34GV1012	ADAPTOR LENOVO 20V/3,25A USB	KAMILA	Tuesday, January 7 2014
K1/1401/38205	FC8T	UBIQUITI AIRGRID M5HP 23DBI	ENHA PRINT	Tuesday, January 7 2014
K1/1401/38206	HDI0102	BT ORI ASUS X550E TANAM	KAYIS	Tuesday, January 7 2014
K1/1401/38207	HDS25250	CONVERTER FLEXIBLE	CHELSEI COMPUTER	Wednesday, Janua 15, 2014
K1/1401/38208	HDS25320	CARD READER 4 SLOT	CAHAYA COMPUTER	Wednesday, Janua 15, 2014
K1/1401/38209	KBH4420	DDR3 4GB VGEN NEW	KLINIK LAPTOP	Wednesday, Janua 15, 2014

Gambar 5. Data Cleaning

Tahap selanjutnya adalah menentukan *header frequent itemset*. Untuk mendapatkan *header frequent itemset*, scan itemset dan frekuensi masing-

masing itemset. Itemset yang tidak memenuhi minimum support akan dieliminasi dari data. *Header frequent itemset* dapat dilihat pada tabel 1.

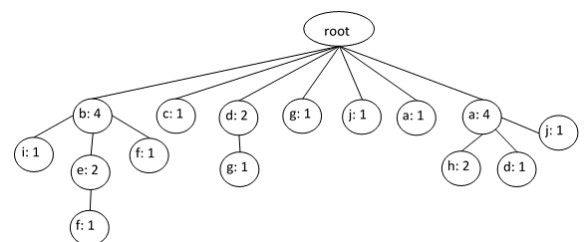
Tabel 1. Header Frequent Itemset

Item	Alias	Frequency
BRACKET PROYEKTOR GALAXY	a	5
ADAPTOR ACER 19V/2,1A (B)	b	4
CPU C2D DELL TWR D2 2GB/160	c	3
ADAPTOR LENOVO 20V/3,25A USB	d	3
ADAPTOR ACER 19V/1,58A	e	2
ADAPTOR MACBOOK 16,5V/3,65A	f	2
UBIQUITI AIRGRID M5HP 23DBI	g	2
CONECTOR BNC RG6	h	2
BT ACER D255W	i	2
CPU AMD DELL TWR D2 2GB/160	j	3

Berikutnya adalah pembentukan *FP Tree* (tabel 2) dengan cara mencari item sesuai urutan pada item yang *frequent*. Data transaksi tidak perlu diurutkan, setiap item dapat langsung dimasukkan dalam *FP Tree* (gambar 6). Jika item yang ditemukan sudah ada maka nilai *supportitem* ditambahkan. Namun jika path belum ada maka dibuat node baru.

Tabel 2. Pembentukan FP Tree

TID	Item bought (alias)	(ordered) frequent items
1	{e,b,f}	{b, e, f}
2	{c}	{c}
3	{e,b}	{b, e}
4	{d}	{d}
5	{d,g}	{d,g}
6	{g}	{g}
7	{j}	{j}
8	{h}	{h}
9	{a,d}	{a, d}
10	{a,h}	{a, h}
11	{b,i}	{b, i}
12	{b,f}	{b, f}



Gambar 6. FP Tree

Langkah selanjutnya adalah membuat *conditional pattern base* (tabel 3) berdasarkan *FP Tree*. Setelah

conditional pattern base terbentuk, berikutnya adalah menentukan frequent itemset (tabel 4).

Tabel 3. Conditional Pattern Base

head items	conditional pattern base
i	b: 1
e	b: 2
f	b,e: 2; b: 1
c	c: 1
d	d: 1; a: 1
g	d: 1; g: 1
j	j: 1; a: 1
a	a: 1
h	a: 2

Tabel 4. Frequent Itemset

head item	condition pattern base	conditional FP-Tree	frequency item
h	a: 2	<a: 2>	ah: 2
e	b: 2	<b: 2>	be: 2
f	b,e: 2; b: 1	<b: 3; e: 2; b, be: 2>	bf: 3; ef: 2; bef: 2
c	c: 1	∅	∅
d	d: 1; a: 1	∅	∅
g	d: 1; g: 1	∅	∅
j	j: 1; a: 1	∅	∅
a	a: 1	∅	∅
i	b: 1	∅	∅

4.3. Implementasi Sistem

Implementasi sistem merupakan tahapan penerapan hasil desain/perancangan sistem ke dalam bahasa pemrograman untuk menghasilkan aplikasi yang dapat mengenali pola belanja konsumen dengan menggunakan algoritma FP-Growth. Aplikasi/sistem yang dihasilkan terdiri dari tiga form, yaitu form upload file, input data, dan lihat analisa. Form upload file digunakan untuk mengimpor data yang bersumber dari database sistem lama yang berupa file dengan ekstensi .xls atau .csv. Data ini nanti yang akan di mining dengan algoritma FP-Growth untuk dapat dikenali pola asosiasi yang terjadi antaritem. Form input data berfungsi untuk menambahkan data baru tanpa harus mengimpor data dari database, sedangkan form lihat analisa digunakan untuk melihat pola asosiasi yang dihasilkan dari proses mining data menggunakan algoritma FP-Growth (gambar 7). Pengetahuan yang dihasilkan dari form lihat analisa berupa pola asosiasi antaritem, data pelanggan terloyal (tabel 5) dan data produk terlaris (tabel 6).



Gambar 7. Form Lihat Analisa

Tabel 5. Pelanggan Terloyal

No	Konsumen	Total
1	Abadi Komputer	11
2	MH Comp	6
3	Mlati Net	5
4	Power Computer	5
5	MM Comp	4
6	77 Comp	3
7	Passnet	3
8	Extrim Net	3
9	Activ Comp	3
10	Cahaya Computer	3

Tabel 6. Produk Terlaris

No	Nama Barang	Total
1	Tinta Korea 100ml Yellow	40
2	Kabel LAN SPC Cat5 1 meter	25
3	Headset Buftech 208 MV	20
4	Headset Appolion F9 MV	19
5	Kabel LAN Belden Cat5 1 Feet	16
6	Kabel LAN SPC Cat6 1 meter	15
7	Adaptor MacBook 16.5V/3.65A	12
8	Adaptor Acer 19V/1.58A	11
9	AdaptorAcer 19V/2.1A (B)	7
10	Key HP Mini 110-1 Soket Panjang	6

DAFTAR PUSTAKA

Han, J. & Micheline, K., 2000. *Data Mining: Concept and Techniques*. Simon Fraser University: Morgan Kaufmann.

Listriani, D., Setyaningrum, A. H. & M.A, F. E., 2016. Penerapan Metode Asosiasi Menggunakan Algoritma Apriori pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro). *Jurnal Teknik Informatik*, Volume Vol. 9 No. 2, pp. 120-127.

Pramudiono, 2006. Indo Datamining, online. 21 June, p. Retrieved 21 June 2010 From <http://datamining.japati.net>.

Turban, E., 2005. *Decision Support System and Intellegent Systems*. Yogyakarta: Andi Offset.