

PENGUNAAN ALGORITMA FP-GROWTH UNTUK MENEMUKAN ATURAN ASOSIASI PADA DATA TRANSAKSI PENJUALAN OBAT DI APOTEK (Studi Kasus : APOTEK UAD)

¹Ratih Rifaatul Mahmudah (08018301) , ²Eko Aribowo (0006027001)

^{1,2}Program Studi Teknik Informatika
Universitas Ahmad Dahlan

Prof.Dr.Soepomo,S.H.,Janturan,Umbulharjo,Yogyakarta 55164

¹Email: ratih.rifa21@gmail.com

²Email: ekoab@tif.uad.ac.id

ABSTRAK

Apotek merupakan salah satu sarana penyedia obat. Setiap hari di apotek terjadi puluhan transaksi penjualan obat. Data transaksi tersebut akan bertambah setiap harinya, sehingga semakin banyak data transaksi penjualan yang tersimpan yang menyebabkan penyimpanan yang sangat besar, biasanya data transaksi tersebut hanya di simpan sebagai arsip. Data transaksi tersebut dapat digunakan untuk mencari pola penjualan obat di apotek. Berdasarkan permasalahan tersebut, maka dibuat suatu penelitian untuk mengembangkan aplikasi komputer untuk mencari pola atau aturan asosiasi antar item obat di dalam transaksi penjualan apotek.

Dalam penelitian ini yang menjadi subyek adalah aplikasi untuk menemukan aturan asosiasi pada data transaksi penjualan obat. Aturan asosiasi merupakan salah satu teknik data mining untuk mencari pola data transaksi penjualan obat. Salah satu algoritma yang dapat digunakan untuk mencari aturan asosiasi adalah algoritma FP-Growth dengan menggunakan dua parameter yaitu minimum support dan minimum confidence. Pembangunan implementasi menggunakan visual basic .net dan database menggunakan sql server. Untuk pengujiannya menggunakan Alpha Test dan Black box Test.

Dari penelitian yang dilakukan menghasilkan sebuah perangkat lunak dengan menggunakan algoritma FP-Growth untuk menemukan aturan asosiasi pada data transaksi penjualan obat dimana hasil pola penjualan obat yang dilakukan pembeli sangat bervariasi.

Kata Kunci : *Datamining, FP-Growth, Aturan Asosiasi*

1. PENDAHULUAN

Apotek merupakan salah satu sarana atau tempat penyedia obat. Setiap harinya di apotek tersebut terjadi puluhan transaksi penjualan obat. Data transaksi penjualan akan terus bertambah setiap harinya. Semakin banyak data transaksi yang tersimpan menyebabkan penyimpanan data yang sangat besar. Biasanya data-data transaksi

penjualan tersebut hanya digunakan sebagai arsip dan tidak diketahui manfaat kedepannya.

Data mining merupakan suatu proses yang interaktif untuk menemukan pola (*pattern*) data tersebut dan memprediksi perilaku dimasa mendatang berdasarkan pola tersebut. Pada dasarnya data mining berhubungan dengan analisa data dan penggunaan teknik-teknik perangkat lunak untuk mencari pola dalam himpunan data yang sifatnya tersembunyi. Dengan memperoleh data-data yang ada, hubungan antar item dalam transaksi obat, selanjutnya dapat di ekstrak atau dianalisa dan diteliti lebih lanjut..

Penelitian yang akan dilakukan berupa analisis terhadap data transaksi penjualan, dimana untuk melakukan analisa terhadap data transaksi tersebut akan digunakan suatu metode yaitu metode *Assosiation rule mining* yang digunakan untuk menemukan pola-pola tertentu. Pada penelitian ini juga akan digunakan algoritma FP-Growth sebagai proses utama dari market basket analysis. *Market basket analysis* merupakan sebuah analisis terhadap kebiasaan konsumen dalam berbelanja. Hasil analisisnya adalah menemukan pola keterkaitan antar produk yang dibeli oleh konsumen dalam bentuk rules.

2. KAJIAN PUSTAKA

2.1. Kajian Terdahulu

Penelitian ini dilakukan mengacu pada penelitian terdahulu dengan judul “*Market Basket Analysis* dengan algoritma apriori pada data transaksi penjualan di swalayan” (Dwi jayanti, 2009). Penelitian ini menjelaskan tentang *market basket analysis* yang digunakan untuk mengetahui produk – produk yang sering dibeli secara bersamaan oleh konsumen. Algoritma yang digunakan salah satunya yaitu algoritma apriori dan menghasilkan informasi berupa pola-pola pembelian yang dilakukan oleh konsumen. Informasi yang telah ditemukan dapat digunakan untuk keperluan pemasaran salah satunya desain *layout* barang.

Penelitian mengenai pencarian pola pengunjung perpustakaan yang meminjam buku perpustakaan, dengan menggunakan aturan asosiasi pada data mining dan memanfaatkan data peminjaman buku perpustakaan (Prittahastari, 2007). Penelitian ini bertujuan untuk menemukan kecenderungan pengunjung perpustakaan yang meminjam buku perpustakaan dengan membangun aplikasi yang digunakan untuk mencari aturan asosiasi dari dalam data peminjaman buku perpustakaan.

2.2. Landasan Teori

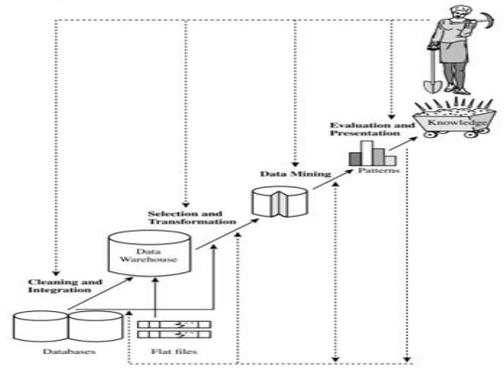
2.2.1. Data Mining

Data mining merupakan serangkaian proses untuk mencari nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak di ketahui sebelumnya..

2.2.2. Tahapan KDD (*Knowledge Discovery in Database*)

Data mining merupakan salah satu rangkaian dari proses pencarian pengetahuan pada database (*Knowledge Discovery in Database/KDD*). KDD

adalah keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. Serangkaian proses tersebut yang memiliki tahap yaitu pada gambar 1 :



Gambar 1: Tahap – tahap data mining

2.2.3. Assosiation Rules Mining

Assosiation rules mining digunakan untuk mencari hubungan korelasi atau asosiasi antar item pada sebuah set data. Analisis asosiasi dikenal sebagai salah satu teknik data mining yang menjadi dasar dari berbagai teknik data mining lainnya. *Support* (nilai penunjang) yaitu persentase kombinasi *item* tersebut dalam database.

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ jumlah\ transaksi} \quad [Novrina]$$

Confidence (nilai kepastian) merupakan rasio antara jumlah transaksi yang di dalamnya terdapat *item-item* yang ada dalam aturan asosiasi, dengan jumlah transaksi yang di dalamnya terdapat *item* yang ada dalam kondisi asosiatif.

$$Confidence\ P(B|A) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Total\ transaksi} \quad [Novrina]$$

2.2.4. Struktur FP-Tree

FP-tree dibangun dengan memetakan setiap data transaksi kedalam setiap lintasan tertentu dalam FP-tree, karena dalam setiap transaksi yang dipetakan mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data FP-tree semakin efektif. Kelebihan dari FP-tree adalah hanya memerlukan dua kali pemindaian data transaksi yang terbukti sangat efisien. (Samuel) FP-tree adalah sebuah pohon dengan definisi sebagai berikut :

1. FP-tree dibentuk oleh sebuah akar yang diberi label *null*, sekumpulan pohon yang beranggotakan item-item tertentu dan sebuah tabel *frequent header*.
2. Setiap simpul dalam FP-tree mengandung tiga informasi penting yaitu label item menginformasikan jenis item yang direpresentasikan simpul tersebut, *support count* merepresentasikan jumlah lintasan transaksi yang melalui simpul tersebut, dan pointer penghubung yang menghubungkan simpul-simpul dengan label item sama antar lintasan ditandai dengan garis panah putus-putus.

2.2.5. Algoritma FP-Growth (Frequent Pattern Growth)

FP-Growth merupakan salah satu algoritma yang termasuk dalam *association rule mining*. (Samuel) Algoritma FP-Growth dibagi menjadi tiga langkah utama, yaitu :

- a. Tahap pembangkitan *Conditional Pattern Base Conditional Pattern Base* merupakan subdatabase yang berisi *prefix path* (lintasan prefix) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui FP-tree yang telah dibangun sebelumnya.
- b. Tahap pembangkitan *Conditional FP-tree* Pada tahap ini *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah *support count* lebih besar sama dengan minimum *support count* akan dibangkitkan dengan *conditional FP-tree*.
- c. Tahap pencarian frequent itemset apabila *Conditional FP-tree* merupakan lintasan tunggal (*single path*), maka didapatkan frequent itemset dengan melakukan kombinasi item untuk setiap *conditional FP-tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-growth secara rekursif.

3. METODE PENELITIAN

3.1. Subyek dan Alat Penelitian

Dalam penelitian ini yang menjadi subyek penelitian adalah Penggunaan Algoritma FP-growth untuk menemukan aturan asosiasi pada Transaksi Penjualan di Apotek. Adapun alat yang digunakan dalam penelitian adalah Visual Basic.net dan untuk database menggunakan SQL Server.

3.2. Metode Pengumpulan Data

Untuk pengumpulan data yaitu melalui studi pustaka dengan mencari, membaca dan mengumpulkan dokumen-dokumen sebagai referensi seperti buku, artikel, dan literatur-literatur tugas akhir yang berhubungan dengan algoritma FP-growth dalam datamining pada transaksi penjualan di Apotek, termasuk pengumpulan dokumen melalui *browsing*. Kemudian melalui Wawancara yang dilakukan dengan cara tanya jawab atau wawancara pada orang yang mempunyai kapasitas dan informasi yang dibutuhkan seperti pegawai apotek. Terakhir dengan Observasi seperti melihat secara langsung kegiatan-kegiatan yang dilakukan oleh

Apotek maupun konsumen dalam hal transaksi penjualan obat dan obat-obat apa saja yang dibeli konsumen.

4. HASIL DAN PEMBAHASAN

4.1. Analisis Data Mining

Analisis Kebutuhan fungsional pada sistem yaitu :

- Transformasi atau perubahan data yang akan diolah oleh user.
- Dapat melakukan proses data mining untuk mencari aturan asosiasi.
- Dapat menampilkan hasil aturan asosiasi dengan algoritma FP-Growth berdasarkan *minimum support* dan *minimum confidence*.

4.2. Implementasi

Contoh kasus dari data transaksi penjualan diberikan tabel data transaksi sebagai berikut, dengan minimum support count = 2.

Tabel 1 : Data Transaksi

Id transaksi	Nama item
1	SP Trochess, woods antitusif
2	woods antitusif, vipro G, betadine, decolsin, bodrex
3	SP Trochess, Vipro G, betadine, Hansaplast, kalpanax
4	SP Trochess, betadine, hansaplast
5	Woods antitusif, SP Trochess, minyak tawon, vipro G
6	Vipro G, betadine, SP Trochess, woods antitusif
7	SP Trochess, antangin cair
8	Woods antitusif, vipro G, SP Trochess
9	SP Trochess, woods antitusif, betadine
10	Woods antitusif, vipro G, hansaplast

Frekuensi kemunculan tiap item dapat dilihat pada tabel 2 berikut :

Tabel 2 : Kemunculan tiap item

Item	Frekuensi
SP Trochess	8
Woods antitusif	7
Vipro G	6
Betadine	5
Hansaplast	3
Kalpanax	1
Antangin cair	1
Minyak tawon	1
Decolsin	1
Bodrex	1

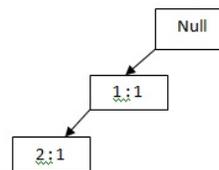
Setelah dilakukan pemindaian pertama didapat item yang memiliki frekuensi diatas support count = 2 . Kelima item inilah yang akan berpengaruh dan akan dimasukkan kedalam FP-tree, selebihnya dibuang karena tidak berpengaruh signifikan. Berikut tabel kemunculan item diurut berdasarkan yang frekuensinya paling tinggi.

Tabel 3 : Data Transaksi

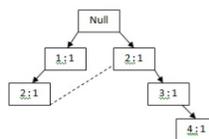
TID	Item
-----	------

1	{1,2}
2	{2,3,4}
3	{1,3,4,7}
4	{1,4,7}
5	{1,2,3}
6	{1,2,3,4}
7	{1}
8	{1,2,3}
9	{1,2,4}
10	{2,3,7}

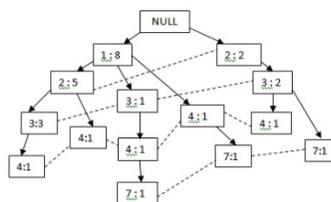
Gambar 2 menunjukkan ilustrasi mengenai pembentukan FP-Tree setelah pembacaan TID 1.



Gambar 2 : Hasil pembentukan FP-Tree setelah pembacaan TID 1



Gambar 3 : Hasil Pembentukan FP-Tree setelah pembacaan TID 2

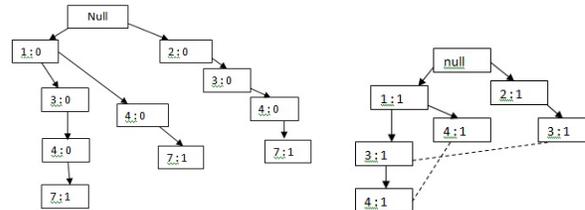


Gambar 4 : Hasil Pembentukan FP-Tree setelah pembacaan TID 10

Gambar 4 menunjukkan proses terbentuknya FP-tree setiap TID dibaca. Setiap simpul pada FP-tree mengandung nama sebuah item dan counter support yang berfungsi untuk menghitung frekuensi kemunculan item tersebut dalam tiap lintasan transaksi.

Selanjutnya yaitu mencari semua subset yang memungkinkan dengan membangkitkan *conditional FP Tree* dan mencari *frequent itemset*, sesuai urutan *frequent list* dari yang paling kecil jumlah kemunculannya.

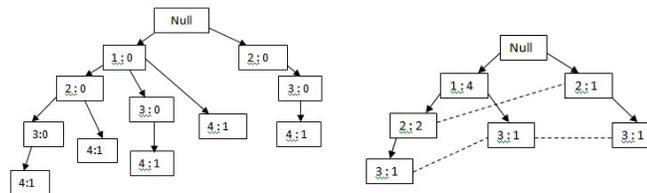
a. Kondisi FP-Tree untuk item 7



Gambar 5 : Kondisi FP-Tree untuk item 7

Ekstrak semua lintasan yang berakhir 7 dan beri nilai 0 untuk path yang selain 7. Ini untuk mengetahui informasi berapa kali item yang lain dibeli bersamaan dengan item 7 dan mengetahui frequent itemset mana yang memenuhi syarat minimum *support*. Setelah itu, naikan satu persatu *path* 7 sampai ke *null* dan nilai *path* 7 dimasukkan ke setiap *path* yang dilewati sampai ke *null*. Seperti gambar 7 item 1 dan 2 nilai kemunculan bersamaan dengan item 7 hanya 1 sehingga item 1 dan 2 dibuang. Sedangkan item 3 dan 4 nilai kemunculannya 2 maka subset yang dihasilkan adalah $\{7\}$, $\{4,7\}$, $\{3,4,7\}$, $\{3,7\}$.

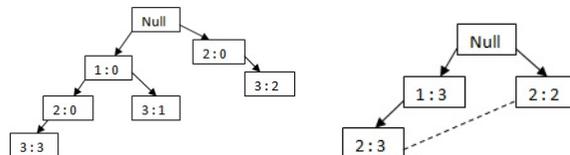
b. Kondisi FP-Tree untuk item 4



Gambar 6 : Kondisi FP-Tree untuk item 4

Maka subset yang dihasilkan adalah $\{1,4\}$, $\{1,2,4\}$, $\{1,3,4\}$, $\{2,3,4\}$, $\{2,4\}$, $\{3,4\}$, $\{4\}$.

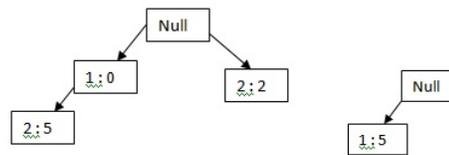
c. Kondisi FP-Tree untuk item 3



Gambar 7 : Kondisi FP Tree untuk item 3

Maka subset yang dihasilkan adalah $\{3\}$, $\{2,3\}$, $\{1,2,3\}$, $\{1,3\}$.

d. Kondisi FP-Tree untuk item 2



Gambar 8 : Kondisi FP-Tree untuk item 2

Maka subset yang dihasilkan adalah {1,2}, {2}.

e. Kondisi FP-Tree untuk item 1



Gambar 9 : Kondisi FP-Tree untuk item 1

Pada kasus ini, lintasan yang berakhiran 1 merupakan lintasan tunggal yang berdiri sendiri dan memiliki nilai frekuensi 8 sehingga frequent itemsetnya hanya {1}.

Setelah memeriksa kondisi FP Tree didapat 18 frequent itemset yang hasilnya yaitu :

Tabel 4 : Hasil frequent item set

Suffix	Frequent item set
5	{7}, {4,7}, {3,7}, {3,4,7}
4	{4}, {3,4}, {2,3,4}, {1,3,4}, {2,4}, {1,2,4}, {1,4}
3	{3}, {2,3}, {1,2,3}, {1,3}
2	{2}, {1,2}
1	{1}

Dari 18 frequent itemset tidak semua dihitung karena *rule* yang dihasilkan adalah jika kita membeli barang A, maka akan membeli barang B, maka subset yang dihitung minimal berisi dua item. Maka yang akan dihitung *confidence* nya adalah 13 subset, yaitu : {4,7}, {3,7}, {3,4,7}, {3,4}, {2,3,4}, {1,3,4}, {2,4}, {1,2,4}, {1,4}, {2,3}, {1,2,3}, {1,3}, {1,2}.

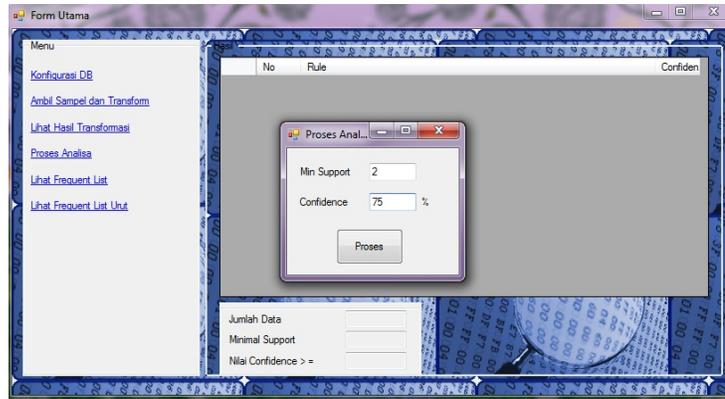
- Contoh itemset {2,3,4} yaitu :
- $2 \rightarrow 3^4 = \frac{2,3,4}{2} = \frac{2}{7} = 0.28 \rightarrow 0.28 \times 100 = 28\%$
 - $3^4 \rightarrow 2 = \frac{3,4,2}{3,4} = \frac{2}{3} = 66.7\%$
 - $2^4 \rightarrow 3 = \frac{2,4,3}{2,4} = \frac{2}{2} = 100\%$
 - $2 \rightarrow 3 = \frac{2,3}{2} = \frac{5}{7} = 71.4\%$
 - $3 \rightarrow 2 = \frac{3,2}{3} = \frac{5}{6} = 83.3\%$

Berikut hasil lengkap pola-pola atau rules yang dihasilkan dari 13 subset.

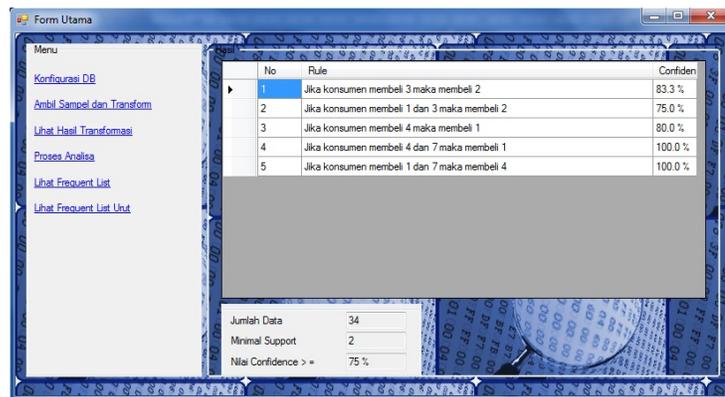
Tabel 5 : Hasil Assosiation Rules

Jika Membeli	Maka akan Membeli	Confidence
1^7 (SP Trochess dan hansaplast)	4 (betadine)	100 %
3 (vipro G)	2 (woods antitusif)	83.8 %
4^7 (betadine dan hansaplast)	1 (SP Trochess)	100%
4 (betadine)	1 (SP Trochess)	80%
1^3 (SP troches dan vipro G)	2 (woods antitusif)	75%

Form proses analisa terdapat parameter inputan untuk melakukan analisis data mining algoritma fp-growth seperti berikut:



Gambar 10 : Form Proses Analisa



Gambar 11 : Form Hasil Analisa

4.3. Evaluasi

Tahap evaluasi yaitu melakukan percobaan dengan menggunakan jumlah data transaksi yang berbeda. Data transaksi yang digunakan berasal dari bulan januari 2011 sampai maret 2011 dimana masing-masing bulan mempunyai jumlah data transaksi yang berbeda. Bulan Januari 2011 jumlah data transaksi yang digunakan ada 337 transaksi dan rule yang dihasilkan ada 40 rules. Bulan Februari 2011 jumlah data transaksi yang digunakan ada 768 transaksi dan rule yang dihasilkan ada 32 rules. Bulan Maret 2011 jumlah data transaksi yang digunakan ada 1142 transaksi dan rule yang dihasilkan ada 44 rules.

5. PENUTUP

5.1. Kesimpulan

Hasil analisis dan aplikasi Datamining menggunakan Algoritma Frequent Pattern Growth dapat disimpulkan bahwa :

- a. Melakukan perhitungan dalam proses Datamining menggunakan Algoritma Frequent Pattern Growth yang menghasilkan *association rule*.
- b. Menampilkan hasil mencari association rule dengan Algoritma Frequent Pattern Growth.
- c. Telah berhasil melakukan analisa data transaksi obat di apotek yang dilakukan pada bulan Januari 2011.

5.2. Saran

Saran yang dapat diberikan untuk pengembangan lebih lanjut yaitu :

- a. Menggunakan algoritma selain algoritma FP Growth untuk mencari aturan asosiasi dalam transaksi penjualan obat.
- b. Algoritma Frequent Pattern Growth masih mungkin dikembangkan lagi untuk penelitian lebih lanjut.

6. DAFTAR PUSTAKA

- [1] Pramudiono, Iko. 2003. Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data. Ilmukomputer.com.
- [2] Kusriani, Luthfi, Emha Taufiq. 2009. *Algoritma Data Mining*, Andi, Yogyakarta.
- [3] Prithahastari, M., 2007, *Penerapan Teknik Aturan Asosiasi dalam Data Mining untuk pencarian pola peminjam Buku Perpustakaan (Studi Kasus : Perpustakaan FMIPA UGM)*, Skripsi, Fakultas Matematika dan Ilmu Pengetahuan Alam-UGM, Yogyakarta.
- [4] Dwi Jayanti, Aprilia., 2009, *Market Basket Analysis Dengan Algoritma Apriori Pada Data Transaksi Penjualan Di Swalayan*, Skripsi, Fakultas Matematika dan Ilmu Pengetahuan Alam-UGM, Yogyakarta.
- [5] Samuel, David, *Penerapan Struktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset*, Jurnal, Program Studi Teknik Informatika, Sekolah Teknologi Elektro dan Informatika, ITB, Bandung.
- [6] Kusriani, Luthfi, Emha Taufiq. 2009. Definisi Data Mining, Jurnal, Jurusan Teknik Informatika, STMIK AMIKOM, Yogyakarta.
- [7] Novrina, Konsep Data Mining. online tersedia : "Konsep Data mining"novrina.staff.gunadarma.ac.id/Downloads/files/.../Association+Rule.pdf [25 desember 2013 pukul 11.00]