

PENERAPAN DATA MINING PADA PENJUALAN TAN'S BAKERY MENGGUNAKAN ALGORITMA APRIORI

Desak Made Dwi Utami Putra¹⁾ Setio Budiono²⁾

Program Studi Teknik Informatika^{1) 2)}

STMIK STIKOM Indonesia, Bali.^{1) 2)}

desak.utami@stiki-indonesia.ac.id¹⁾ setiobudiono1992@gmail.com²⁾

ABSTRACT

PT.Tan's Bakery is a manufacture company of food that provides a variety of bread products (Bakery) and cakes (Pastry) to supply the needs of Hotels, Villas, and Restaurants in Bali in particular. There are a lot of products that are produced and offered to its customers ranging from European bakery products to international-standard pastry products, frequent daily activities at PT.Tan's Bakery among others, order receipt, product manufacturing instructions, product stock provision, procurement raw materials, customer additions, etc. In this study developed data mining analysis system using a priori algorithm is expected to give solution because by applying data analysis using a priori algorithm at Tan's Bakery can generate useful information for company, in analyzing sales transaction besides saving time, high, from the results of testing conducted obtained the results of the developed system is running as expected.

Keywords: Tan's Bakery, Apriori, Data Mining, Support & Confident.

ABSTRAK

PT.Tan's Bakery merupakan suatu perusahaan yang bergerak dalam bidang pembuatan makanan yang menyediakan berbagai produk roti (Bakery) dan kue (Pastry) untuk menyuplai kebutuhan Hotel, Villa, dan Restoran yang ada di Bali khususnya. Ada banyak sekali produk-produk yang diproduksi dan ditawarkan ke pelanggannya mulai dari produk bakery Eropa hingga produk pastry yang bertaraf International, kegiatan yang sering dilakukan setiap hari pada PT.Tan's Bakery diantara lain, penerimaan pesanan, intruksi pembuatan produk, penyediaan stok produk, pengadaan bahan baku, penambahan pelanggan,dll.Pada penelitian ini dikembangkan sistem analisis data mining menggunakan algoritma apriori diharapkan dapat memeberikan solusi karena dengan menerapkan analisis data mining menggunakan algoritma apriori pada Tan's Bakery dapat menghasilkan informasi yang berguna bagi perusahaan, dalam menganalisis transaksi penjualan selain menghemat waktu, tingkat keakuratan informasi yang dihasilkan menjadi lebih tinggi, dari hasil pengujian yang dilakukan didapat hasil sistem yang dikembangkan sudah berjalan sesuai dengan yang diharapkan.

Kata Kunci: Tan's Bakery, Apriori, Data Mining, Support & Confident.

PENDAHULUAN

PT.Tan's Bakery merupakan suatu perusahaan yang bergerak dalam bidang pembuatan makanan yang menyediakan berbagai produk roti (Bakery) dan kue (Pastry) untuk menyuplai kebutuhan Hotel, Villa, dan Restoran yang ada di Bali khususnya. Ada banyak sekali produk-produk yang diproduksi dan ditawarkan ke pelanggannya mulai dari produk bakery Eropa

hingga produk pastry yang bertaraf International, kegiatan yang sering dilakukan setiap hari pada PT.Tan's Bakery diantara lain, penerimaan pesanan, intruksi pembuatan produk, penyediaan stok produk, pengadaan bahan baku, penambahan pelanggan,dll.

Berdasarkan hasil wawancara dengan bapak Endik Purwanto selaku operasional supervisor PT Tan's Bakery, dari awal berdiri perusahaan ini telah menggunakan sistem

informasi berbasis komputer untuk menangani point of sale (POS) dan inventory, yang hingga kini telah mencatat lebih dari 125 item barang dan 3250 lebih transaksi penjualan. Sistem informasi yang digunakan oleh perusahaan ini belum dapat memberi informasi tentang kebiasaan berbelanja pelanggan, sistem tidak dapat memberi informasi hubungan antar item yang dibeli oleh pelanggan. Sehingga sering terjadi kekosongan salah satu stok barang yang sering dibeli secara bersamaan dengan barang lain oleh pelanggan. Penggunaan metode manual dalam menganalisis kebiasaan belanja pelanggan sangat sulit dilakukan, karena perkembangan data yang semakin pesat. Hal ini disebabkan karena data transaksi penjualan yang terus bertambah biasanya hanya digunakan sebagai arsip untuk data-data kegiatan Tan's Bakery, baik itu tentang data pemesanan bahan baku, data produksi, data penjualan produk Tan's Bakery, gaji karyawan (payroll) hingga pelaporan akuntansi keseluruhan Tan's Bakery. Sehingga terjadi penumpukan data yang tidak diketahui apa manfaatnya.

Pada dasarnya kumpulan data tersebut memiliki informasi-informasi yang bermanfaat, yang bisa digunakan untuk mengambil suatu keputusan dan untuk memperoleh pengetahuan yang baru (knowledge), yang berguna untuk peningkatan kemajuan perusahaan, seperti peningkatan penjualan produk, penentuan produk yang akan diproduksi terlebih dahulu, promosi dll. Penerapan data mining dapat membantu untuk menganalisa data yang diperoleh dari transaksi pada sistem penjualan, sehingga dapat menggali pola-pola kebiasaan belanja pelanggan di perusahaan Tan's Bakery.

Pada penelitian sebelumnya Syaifullah Muhammad Afif (2010), pada analisis sistem penjualan dengan algoritma apriori menyimpulkan teknik Data mining dengan algoritma apriori dapat di implementasikan pada sistem penjualan, dengan aplikasi yang berbasis teknologi informasi dihasilkan sebuah metode yang bisa meningkatkan penjualan dengan cara memberikan saran kepada konsumen, dan keterkaitan suatu barang yang dibeli oleh konsumen bisa dihitung dengan teknik algoritma apriori. Serta pada penelitian Gunadi (2012), menemukan sejumlah aturan asosiasi dari basis data transaksi penjualan produk buku di percetakan PT. Gramedia dan dapat digunakan sebagai pertimbangan dalam pembuatan strategi pemasaran dan penjualan. Data

mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengolahan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar.

Hubungan yang dicari dalam data mining dapat berupa hubungan antara dua atau lebih dalam satu dimensi. Misalnya dalam dimensi produk dapat melihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan juga dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek. Algoritma apriori termasuk jenis aturan asosiasi pada data mining, algoritma apriori yang bertujuan untuk menemukan frequent item sets dijalankan pada sekumpulan data. Analisis apriori di definisikan suatu proses untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk support dan syarat minimum untuk confidence.

Oleh karena itu dalam penelitian ini dicoba untuk menyelesaikan permasalahan yang ada di PT. Tan's Bakery dengan memanfaatkan salah satu jenis data yang ada yaitu data penjualan, dengan menggunakan salah satu teknik yang digunakan dalam pengolahan data tersebut dengan menggunakan metode algoritma Apriori yang merupakan salah satu teknik dalam data mining. Hal ini bertujuan agar nantinya dengan menggunakan teknik ini dapat dihasilkan informasi baru seperti pengetahuan tingkat kecenderungan atau kebiasaan pelanggan dalam membeli barang-barang secara bersamaan, menentukan tata letak barang, dan menentukan antrian produk-produk yang akan diproduksi lebih dulu, yang berguna untuk meningkatkan kinerja operasional perusahaan serta sebagai referensi para sales atau marketing Tan's Bakery untuk meningkatkan keuntungan perusahaan.

TINJAUAN PUSTAKA

Data Mining

Data mining adalah salah satu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Data mining* adalah proses 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, dkk, 2005).

Menurut Gartner Group *data mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan

memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan seperti teknik statistik dan matematika (Larose, 2005).

Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2005) :

a. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

b. Estimasi

Hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah *numeric* daripada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasar nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya. Contoh lain yaitu estimasi nilai indeks prestasi kumulatif mahasiswa program pascasarjana dengan melihat nilai indeks prestasi mahasiswa tersebut pada saat mengikuti program sarjana.

c. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Contoh prediksi dalam bisnis dan penelitian adalah :

1. Prediksi harga beras dalam tiga bulan yang akan datang.
2. Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

d. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah .

Contoh lain klasifikasi dalam bisnis dan penelitian adalah :

1. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
2. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
3. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa.

e. Klaster

Klaster merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam klaster lain.

Klaster berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam klaster. Klaster tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma klaster mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

Contoh klaster dalam bisnis dan penelitian adalah :

1. Mendapat kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
2. Untuk tujuan audit akunTan'si, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
3. Melakukan klaster terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

f. Asosiasi

Tugas Asosiasi dalam *data mining* adalah menentukan atribut yang muncul

dalam satu waktu. Dalam dunia bisnis lebih umum disebut Data Analisis.

Contoh asosiasi dalam bisnis dan penelitian adalah :

1. Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran *upgrade* layanan yang diberikan.
2. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

Analisis Asosiasi

a. Association Rule Mining

Menurut Kusriani dan Luthfi (2009), "Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*". Contoh aturan asosiatif dari analisis pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli *Toast white* bersamaan dengan *Burger*. Dengan pengetahuan tersebut, pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Aturan asosiasi (*association rules*) atau analisis afinitas (*affinity analysis*) berkenaan dengan studi tentang "apa bersama apa". Ini bisa berupa studi transaksi di supermarket, misalnya seseorang yang membeli *Burger* bayi juga membeli sabun mandi. Disini berarti *Burger* bayi bersama dengan sabun mandi. Karena awalnya berasal dari studi tentang *database* transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama produk apa, maka aturan asosiasi juga sering dinamakan *market basket analysis* (Santosa, 2007).

Market basket analysis (Data Analisis) adalah suatu metodologi untuk melakukan analisis pola belanja konsumen dengan menemukan asosiasi antar beberapa item yang berbeda, yang diletakkan konsumen dalam *shopping basket* yang dibeli pada suatu transaksi tertentu. Tujuan dari market basket adalah untuk mengetahui produk-produk mana yang mungkin akan dibeli secara bersamaan. Analisis data transaksi dapat menghasilkan pola pembelian produk yang sering terjadi. Informasi ini dapat digunakan bagi para penjual dalam mengembangkan strategi dan pengambilan keputusan dengan melihat beberapa item mana saja yang sering

dibeli secara bersamaan oleh konsumen, misalnya dalam pengaturan peletakan produk di toko, produk yang sering dibeli secara bersamaan diletakkan secara berdekatan. teknik ini telah digunakan oleh banyak toko grosir maupun retail (Olson, Young, 2006).

Analisis asosiasi didefinisikan suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk *support* (*minimum support*) dan syarat minimum untuk *confidence* (*minimum confidence*). *Support* (nilai penunjang) adalah persentase kombinasi *item* tersebut tersebut dalam *database*, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antar-*item* dalam aturan asosiasi (Kusriani & Luthfi, 2009).

Salah satu contoh bentuk aturan asosiasi sebagai berikut (Handojo, 2007) :

Toast white -> *keju* [*support* = 2%, *confidence* = 60%] Seorang konsumen yang membeli *Toast white* punya kemungkinan 60% untuk juga membeli keju. Aturan ini cukup signifikan karena mewakili 2 % dari catatan transaksi selama ini.

Menurut Kusriani & Luthfi (2009), Metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

Analisis Pola Frekuensi Tinggi

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah *Item* diperoleh dengan rumus berikut.

$$Support(A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total Transaksi}}$$

Sementara itu, nilai *support* dari 2 item diperoleh dari rumus 2 berikut.

$$Support(A,B) = P(A \cap B)$$

$$Support(A,B) = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi}}$$

Berikut merupakan contoh data transaksi sederhana yang ditunjukkan pada Table 2.1 dibawah ini.

Tabel 2.1 Data Transaksi

Transaksi	Item yang dibeli
1	<i>Burger, Baguette, Sourdough</i>
2	<i>Baguette, Sourdough, Toast white</i>
3	<i>Baguette, Sourdough</i>
4	<i>Burger, Toast white</i>
5	<i>Burger, Sourdough, Toast white</i>

6	<i>Baguette, Sourdough</i>
7	<i>Sourdough, Long bread, Burger</i>
8	<i>Sourdough, Long bread, Burger</i>
9	<i>Burger, Toast white, Long bread</i>
10	<i>Sourdough, Baguette, Long bread</i>

Sebagai contoh, ada *database* dari transaksi belanja pasar swalayan seperti ditunjukkan dalam Tabel 2.1

Data pada Tabel 2.1 dalam *database* transaksional biasa direpresentasikan seperti Tabel 2.2 dibawah ini.

Tabel 2.2 Representasi Data Transaksi dalam Database Transaksional

Transaksi	Item yang dibeli
1	<i>Burger</i>
1	<i>Baguette</i>
1	<i>Sourdough</i>
2	<i>Baguette</i>
2	<i>Sourdough</i>
2	<i>Toast white</i>
3	<i>Baguette</i>
3	<i>Sourdough</i>
Transaksi	Item yang dibeli
4	<i>Burger</i>
4	<i>Toast white</i>
5	<i>Burger</i>
5	<i>Sourdough</i>
5	<i>Toast white</i>
6	<i>Baguette</i>
6	<i>Sourdough</i>
7	<i>Sourdough</i>
7	<i>Long bread</i>
7	<i>Burger</i>
8	<i>Sourdough</i>
8	<i>Long bread</i>
8	<i>Burger</i>
9	<i>Burger</i>
9	<i>Toast white</i>
9	<i>Long bread</i>
10	<i>Sourdough</i>
10	<i>Baguette</i>
10	<i>Long bread</i>

Dan bila kita bentuk dalam bentuk tabular, data transaksi akan tampak seperti Tabel 2.3 berikut

Tabel 2.3 Format Tabular Data Transaksi

Transaksi	<i>Toast white</i>	<i>Sourdough</i>	<i>Long bread</i>	<i>Burger</i>	<i>Baguette</i>
1	1	1	0	1	0
2	1	1	0	0	1
3	1	1	0	0	0

4	0	0	0	1	1
5	0	1	0	1	1
6	1	1	0	0	0
7	0	1	1	1	0
8	0	1	1	1	0
9	0	0	1	1	1
10	1	1	1	0	0
Jumlah	5	8	4	6	4

Misalkan D adalah himpunan transaksi yang dipresentasikan dalam Tabel 2.3, dimana setiap transaksi T dalam D merepresentasikan himpunan item yang berada dalam I. I adalah himpunan item yang dijual {*Baguette, Sourdough, Long bread, Burger, Toast white*}. Misalkan kita memiliki himpunan item A misal (*Burger* dan *Sourdough*) dan himpunan item lain B (misal *Long bread*). Kemudian aturan asosiasi akan berbentuk :

Jika A, maka B (A->B)

Dimana *antecedent* A dan *consequent* B merupakan subset dari I, dan A dan B merupakan *mutually exclusive* dimana :

Jika A, Maka B

Tidak berarti :

Jika B, maka A.

Definisi ini tidak berlaku untuk aturan *trivial* seperti :

Jika beans dan Squash, maka beans

Seorang analisis mungkin hanya akan mengambil aturan yang memiliki *support* dan/atau *confidence* yang tinggi. Aturan yang kuat adalah aturan-aturan yang melebihi kriteria *support* dan/atau *confidence* minimum. Misalnya seorang analis menginginkan aturan yang memiliki *support* lebih dari 20% dan *confidence* lebih dari 35%.

Sebuah *itemset* adalah himpunan *item-item* yang ada dalam I, dan *k-itemset* adalah *itemset* yang berisi k item. Misalnya {*Baguette, Sourdough*} adalah sebuah *2-itemset* dan {*Baguette, Sourdough, Toast white*} merupakan *3-itemset*.

Frequent Itemset menunjukkan *itemset* yang memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan (ϕ). Misalkan $\phi = 2$, maka semua *itemset* yang frekuensi kemunculannya lebih dari atau sama dengan 2 kali disebut *frequent*. Himpunan dari *frequent k-itemset* dilambangkan dengan F_k .

Tabel 2.4 berikut ini menunjukkan calon *2-itemset* dari data transaksi pada Tabel 2.1.

Tabel 2.4 Calon 2-itemset

Kombinasi	Jumlah
Toast white, Sourdough	5
Toast white, Long bread	1
Toast white, Burger	1
Toast white, Baguette	1
Sourdough, Long bread	3
Sourdough, Burger	4
Sourdough, Baguette	2
Kombinasi	Jumlah
Long bread, Burger	3
Long bread, Baguette	1
Burger, Baguette	3

Dari data tersebut di atas jika ditetapkan nilai $\phi = 2$, maka :

$$F2 = \{\{Toast\ white, Sourdough\}, \{Sourdough, Long\ bread\}, \{Sourdough, Burger\}, \{Sourdough, Baguette\}, \{Long\ bread, Burger\}, \{Burger, Baguette\}\}$$

Tabel 2.5 Calon 3-itemset

Kombinasi	Jumlah
Toast white, Sourdough, Long bread	1
Toast white, Sourdough, Burger	1
Sourdough, Burger, Long bread	2
Sourdough, Burger, Baguette	1
Sourdough, Long bread, Baguette	0
Long bread, Burger, Baguette	1

Kombinasi dari *itemset* dalam F2 dapat kita gabungkan menjadi calon 3-*itemset*. *Itemset-itemset* dari F2 yang dapat digabungkan adalah *itemset-itemset* yang memiliki kesamaan dalam *k-1* item pertama. Calon 3-*itemset* yang dapat dibentuk dari F2 tampak pada tabel 2.5

Dengan demikian F3 = $\{\{Sourdough, Burger, Long\ bread\}\}$, karena hanya kombinasi inilah yang memiliki frekuensi kemunculan $\geq \phi$.

a. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A-> B.

Nilai *confidence* dari aturan A-> B diperoleh dari rumus berikut.

$$Confidence = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} = P(B|A)$$

Dari F3 yang telah ditemukan, dapat dilihat besarnya nilai *support* dan *confidence* dari calon aturan asosiasi seperti tampak pada Tabel 2.6.

Tabel 2.6 Calon Aturan Asosiasi dari F3

Aturan	Confidence	
Jika membeli Sourdough dan Burger, maka akan membeli Long bread.	2/4	50%
Jika membeli Sourdough dan Long bread, maka akan membeli Burger	2/3	67%
Jika membeli Long bread dan Burger, maka akan membeli Sourdough	2/3	67%

Misalkan ditetapkan nilai *confidence* minimal adalah 60%, maka aturan yang bisa terbentuk adalah aturan dengan dua *antecedent* berikut.

Jika membeli Sourdough dan Long bread, maka akan membeli Burger.

Jika membeli Long bread dan Burger, maka akan membeli Sourdough.

Calon aturan asosiasi dari F2 bisa dilihat pada Tabel 2.7.

Tabel 2.7. Aturan Asosiasi

Aturan	Confidence	
Jika membeli Teh, maka akan membeli gula.	5/5	100%
Jika membeli gula, maka akan membeli teh.	5/8	62.5%
Jika membeli gula, maka akan membeli kopi.	3/8	37.5%
Jika membeli kopi, maka akan membeli gula.	3/4	75%
Jika membeli gula, maka akan membeli susu.	4/8	50%
Jika membeli susu, maka akan membeli gula.	4/6	67%
Jika membeli gula, maka akan membeli roti.	2/8	25%
Jika membeli roti, maka akan membeli gula.	2/4	50%
Jika membeli kopi, maka akan membeli susu.	3/4	75%
Jika membeli susu, maka akan membeli kopi.	3/6	50%
Jika membeli susu, maka akan membeli roti.	3/6	50%
Jika membeli roti, maka akan membeli susu.	3/4	75%

Algoritma Apriori

Menurut Moertini dan Marsela (2007), Apriori adalah suatu algoritma yang sudah sangat dikenal dalam melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule*.

Menurut Han dan Kamber (2001) Algoritma Apriori adalah algoritma analisis keranjang pasar yang digunakan untuk menghasilkan aturan asosiasi, dengan pola "if-then". Algoritma Apriori menggunakan pendekatan iteratif yang dikenal dengan *level-wise search*, dimana k-kelompok produk digunakan untuk mengeksplorasi (k+1)-kelompok produk atau (k+1)-itemset.

Menurut Han dan Kamber (2006) Algoritma Apriori menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Pada Algoritma Apriori untuk menentukan kandidat-kandidat yang mungkin muncul dengan cara memperhatikan minimum *support*. Adapun dua proses utama yang dilakukan dalam Algoritma Apriori, yaitu :

1. *Join* (penggabungan). Pada proses ini setiap item dikombinasikan dengan *item* yang lainnya sampai tidak terbentuk kombinasi lagi.
2. *Prune* (pemangkasan). Pada proses ini, hasil dari item yang telah dikombinasikan tadi lalu dipangkas dengan menggunakan minimum *support* yang telah ditentukan oleh *user*.

Dua proses utama tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*, yang dapat dilihat pada Gambar 2.3 berikut ini.

Input :

D, a *database* of a transactions;
Min_support, the minimum support count threshold

Output : L, frequent itemsets in D

Method :

```

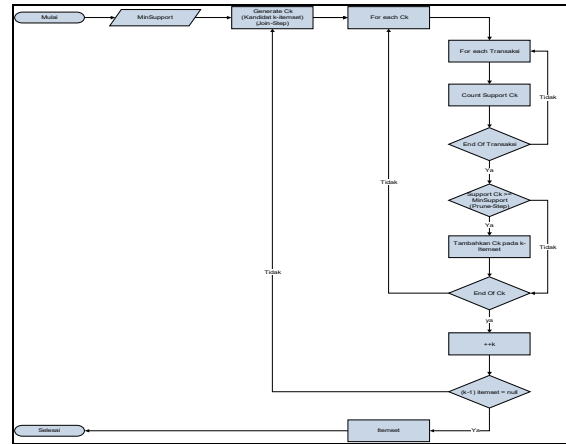
L1 = find_frequent_1_itemsets(D);
for (k = 2; Lk-1 ≠ ∅, k | |) {
    Ck = Apriori_gen(Lk-1);
    for each transaction t ∈ D {
        //scan D for counts
        Ct = subset(Ck,t); //get the
        subsets of t that are candidates
        for each candidate c ∈ Ct
            c.count | |;
    }
}
    
```

$$L_k = \{c \mid C_k \mid c.counts \geq$$

min_sup}

Return L = $\bigcup_k L_k$;

Berikut ini adalah *flowchart* dari algoritma Apriori yang diterangkan pada Gambar 2.4 dibawah ini



Gambar 1. Flowchart Algoritma Apriori

Awal dari proses pencarian pola frekuensi tinggi (*frequent pattern mining*) dengan Algoritma Apriori adalah dengan memasukkan nilai *minimum support*, misalnya 20 (dalam bentuk persen). Kemudian dilakukan proses penggabungan (*join step*) yaitu menggabungkan setiap barang menjadi kandidat *k-itemset*, misalnya jika $k=2$, maka dilakukan penggabungan 2 barang untuk dijadikan kandidat *2-itemset*. Setelah dilakukan proses penggabungan, selanjutnya dilakukan perulangan untuk mencari nilai *support* dari masing-masing kandidat pada transaksi yang telah terjadi. Nilai *support* yang ditemukan kemudian diuji pada proses pemangkasan (*prune-step*) yaitu mencocokkan nilai *support* terhadap *minimum support* yang telah ditentukan. Jika nilai *support* kandidat sama atau melebihi nilai *minimum support*, maka kandidat tersebut ditambahkan pada tabel *k-itemset*. Setelah semua kandidat *k-itemset* selesai diuji, berikutnya nilai *k* ditambah dengan 1 untuk melanjutkan pencarian kandidat *k+1-itemset*. Sebelum melanjutkan pencarian kandidat, dilakukan pengujian jumlah jumlah *itemset* sebelum kandidat baru ($k-1$), apabila *itemset* $k-1$ kosong maka proses pencarian berhenti, jika *itemset* $k-1$ tidak kosong maka dilanjutkan untuk mencari kandidat selanjutnya.

METODOLOGI PENELITIAN

Metode Pengumpulan Data Primer

a. Wawancara

Metode wawancara adalah teknik pengumpulan data yang dilakukan dengan cara langsung, yaitu menemui narasumber secara langsung dan melakukan tanya jawab terkait dengan permasalahan yang akan dibahas dalam laporan penelitian. Wawancara dilakukan secara langsung kepada pihak *Tan's Bakery* yaitu Manager Operasional. Wawancara bertujuan untuk mengetahui sistem penjualan yang digunakan di PT. *Tan's Bakery* dan permasalahan yang sering menjadi kendala pada sistem yang digunakan tersebut. Dari hasil wawancara diketahui, Untuk itu metode wawancara ini dirangkum sebagai berikut :

1. Sistem apa yang digunakan pada PT. *Tan's Bakery* dalam menjalankan operational setiap harinya ?
2. Kenapa PT. *Tan's Bakery* menggunakan sistem tersebut?
3. Adakah suatu kemampuan khusus dalam sistem tersebut?
4. Apa kelebihan dan kekurangan sistem tersebut?

b. Observasi

Metode observasi adalah teknik pengumpulan data dengan pengamatan dan pencatatan secara langsung melihat kondisi sebenarnya di *Tan's Bakery*.

Data yang digunakan berasal dari basis data penjualan produk *Tan's Bakery*. Data yang akan dianalisis adalah *sample* data penjualan *Tan's Bakery* periode 2016 selama tiga hari dibulan Juni yaitu, tanggal 7,8,9 Juni 2016.

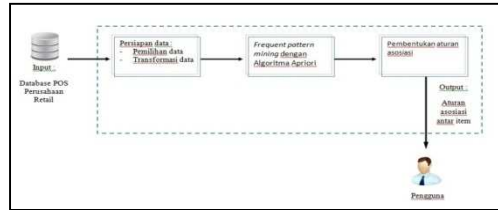
Metode Pengumpulan Data Sekunder

Dalam pengumpulan data dengan metode sekunder, penulis menggunakan sebuah teknik, yaitu:

- a. Studi Kepustakaan/Literatur
Studi kepustakaan dilakukan dengan mencari segala informasi yang berkaitan dengan masalah yang dihadapi pada *Tan's Bakery*. Seperti halnya dengan membaca buku-buku dan membuka website-website menyangkut masalah yang perlu di analisa.

Gambaran Umum Sistem

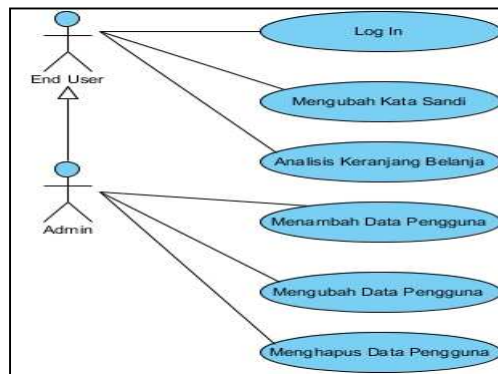
Dibawah ini akan menjelaskan mengenai gambaran umum system yang akan dibuat yang dapat dilihat pada Gambar 3.5 dibawah ini



Gambar 2. Gambaran Umum Sistem

Use Case Diagram Aplikasi Data Analisis

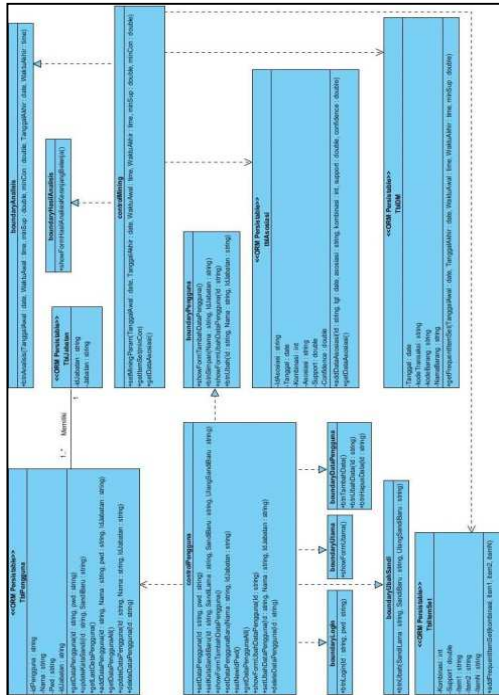
Use case diagram yang menggambarkan bagaimana orang-orang berinteraksi dengan sistem tersebut. Sebuah *use case* merepresentasikan sebuah interaksi antara aktor dengan sistem. *Use case* adalah teknik untuk merekam persyaratan fungsional sebuah sistem. *Use case* mendeskripsikan interaksi tipikal antara para pengguna sistem dengan sistem itu sendiri, dengan , memberi sebuah narasi tentang bagaimana sistem tersebut digunakan (fowler, 2005). Berikut ini adalah *use case diagram* aplikasi Data Analisis.



Gambar 3. Use Case Diagram Aplikasi Data Analisis

Class Diagrams Aplikasi Data Analisis

Class diagrams (diagram kelas) merupakan kumpulan kelas-kelas objek. Secara teknis, Pender (2003) mengartikan sebuah kelas sebagai suatu definisi sumber daya yang termasuk di dalamnya informasi-informasi yang menggambarkan fitur suatu entitas dan bagaimana penggunaannya. Sedangkan objek adalah entitas yang bersifat unik yang mengikuti aturan-aturan yang sudah didefinisikan dalam kelasnya. Berikut ini adalah diagram kelas aplikasi Data Analisis :



Gambar 4. Class diagram Aplikasi keranjang belanja

Class Diagram Aplikasi Keranjang Belanja memiliki 14 kelas. Kelas TblPegawai memiliki hubungan asosiasi (*association relationship*) dengan kelas TblJabatan, yaitu setiap pengguna hanya memiliki satu jabatan dan setiap jabatan dapat berada pada satu atau lebih data pengguna. Kelas control Pengguna memiliki hubungan *dependency* ke kelas Tbl Pengguna. Kelas *control Mining* memiliki hubungan *dependency* dengan kelas TblAsosiasi, TblItemSet, dan TblDM. Kelas control Pengguna memiliki hubungan *realization* dengan kelas *boundary LogIn*, *boundary Utama*, *boundary DataPengguna*, *boundary Pengguna*, dan *boundary Ubah Kata Sandi*. Kelas *control Mining* memiliki hubungan *realization* dengan kelas *boundary Analisis*, dan *boundary HasilAnalisis*.

IMPLEMENTASI

Form Log In

Form Log In merupakan *form* yang pertama kali muncul ketika aplikasi dijalankan. Pada *form* ini pengguna memasukan kode pengguna dan kata sandi untuk dapat menggunakan aplikasi, untuk lebih jelasnya akan dapat dilihat pada Gambar 4.1 dibawah ini.



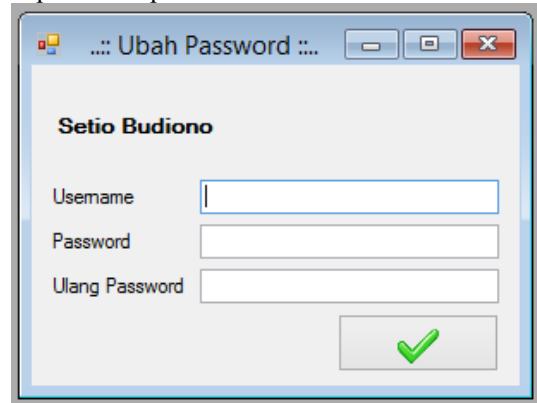
Gambar 5. Form Log In

Berbeda dengan *form login* pada umumnya, form login ini memiliki dua fungsi:

1. Admin: merupakan pengguna aplikasi yang dapat memiliki akses atau wewenang untuk melakukan penambahan, mengubah, dan menghapus daftar user yang menggunakan aplikasi ini.
2. Enduser: merupakan pengguna aplikasi yang melakukan kegiatan analisis penjualan menggunakan aplikasi analisis data transaksi ini.

Form Ubah kata sandi

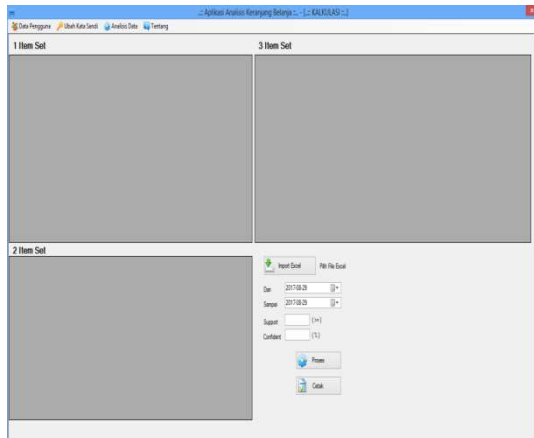
Fasilitas ini berguna untuk mengubah kata sandi pengguna yang telah masuk ke sistem. Pengguna harus memasukan kata sandi lama, kata sandi baru, dan mengulang kata sandi baru lagi. Untuk melanjutkan proses mengubah kata sandi dilanjutkan dengan menekan tombol ubah agar lebih jelas dapat dilihat pada Gambar 4.3 dibawah ini.



Gambar 6. Form Ubah Kata Sandi

Form Analisis Data

Adapun beberapa langkah-langkah untuk melakukan analisis keranjang belanja adalah sebagai berikut :



Gambar 7. Form Awal Analisis

Pada Gambar 7 form ini terdapat tiga tampilan jendela analisis:

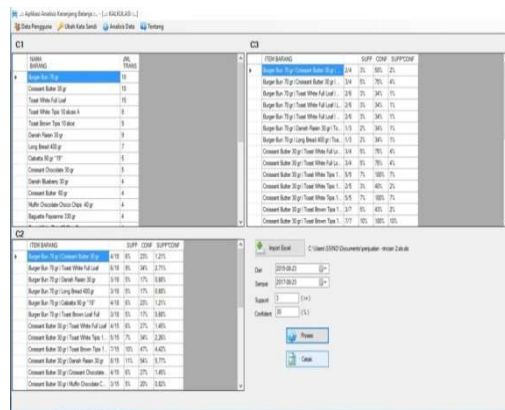
1. 1ItemSet merupakan lembar kerja yang menunjukkan analisis transaksi yang ada berdasarkan urutan jumlah transaksi barang terbanyak hingga terendah sebagai hasil laporan K-Item Set 1(1 Item set) saat dilakukannya proses analisis.
2. 2Itemset merupakan bentukan atau gabungan antar 2 item set berdasarkan analisis transaksi k-item set 2 (2 Item Set) yang terbentuk yang menunjukkan presentase tingkat dukungan dan keyakinan apabila 2 item set terbentuk saat dilakukannya analisis.
3. 3ItemSet merupakan bentukan atau gabungan transaksi barang yang terbentuk dari hasil gabungan 3 Item set dari analisis k-item set 2 (2 Item Set) sebelumnya, yang bisa menjadi pemicu untuk membentuk laporan yang menunjukkan seberapa besar pengaruh nilai nilai suport dan confidenya dalam membentuk k-itemset 3 (3 Item Set).

Untuk melanjutkan pengguna menekan tombol import file yang berbentuk dokumen excel kemudian memilih file yang akan dianalisis berdasarkan alamat directory dimana file tsb tersimpan, agar lebih jelasnya dapat diperhatikan pada Gambar 8



Gambar 8. Tampilan Inputan Tanggal Dan Nilai Support Confident

Pada form ini pengguna harus memasukan tanggal kegiatan transaksi yang akan dianalisis serta nilai *minimum support* dan nilai *minimum confidence*. Setelah memasukan kedua nilai tersebut kemudian untuk melanjutkan tekan tombol proses untuk menunjukkan hasil analisis data yang sedang di analisis, form anlisis akan tampak seperti Gambar 9 dibawah ini



Gambar 9. Hasil Pemrosesan Data Transaksi

Form Hasil Laporan Analisis Data Transaksi Penjualan

Setelah form hasil analisis berubah seperti pada Gambar 9 pada form tersebut juga dapat disesuaikan kembali nilai support dan *confident* sesuai keinginan. Setelah puas dengan hasil analisis maka dapat dilakukan pencetakan hasil analisis yang memiliki tingkat hubungan antar barang 3 Item Set yang diinginkan. Setelah tombol cetak ditekan maka hasil laporan analisis akan tercetak seperti Gambar 10 dibawah ini.

LAPORAN HASIL PERHITUNGAN TINGKAT PRESENTASE KETERKAITAN ANTAR PRODUK PADA PENJUALAN		
No	Itemset Terbentuk	Support Conf
1	Jika membeli Burger Bun 70 gr dan Croissant Butter 30 gr maka akan membeli Toast White Full Loaf	3% 50%
2	Jika membeli Burger Bun 70 gr dan Croissant Butter 30 gr maka akan membeli Danish Raisin 30 gr	3% 75%
3	Jika membeli Burger Bun 70 gr dan Toast White Full Loaf maka akan membeli Danish Raisin 30 gr	3% 34%
4	Jika membeli Burger Bun 70 gr dan Toast White Full Loaf maka akan membeli Long Bread 400 gr	3% 34%
5	Jika membeli Burger Bun 70 gr dan Toast White Full Loaf maka akan membeli Toast Brown Loaf Full	3% 34%
6	Jika membeli Burger Bun 70 gr dan Danish Raisin 30 gr maka akan membeli Toast Brown Loaf Full	2% 34%
7	Jika membeli Burger Bun 70 gr dan Long Bread 400 gr maka akan membeli Toast Brown Loaf Full	2% 34%
8	Jika membeli Croissant Butter 30 gr dan Toast White Full Loaf maka akan membeli Danish Raisin 30 gr	5% 75%
9	Jika membeli Croissant Butter 30 gr dan Toast White Full Loaf maka akan membeli Croissant Chocolate 30 gr	5% 75%
10	Jika membeli Croissant Butter 30 gr dan Toast White Tips 10 slices A maka akan membeli Toast Brown Tips 10 slice	7% 100%
11	Jika membeli Croissant Butter 30 gr dan Toast White Tips 10 slices A maka akan membeli Danish Raisin 30 gr	3% 40%
12	Jika membeli Croissant Butter 30 gr dan Toast White Tips 10 slices A maka akan membeli Toast Brown Tips 10 Slice	7% 100%
13	Jika membeli Croissant Butter 30 gr dan Toast Brown Tips 10 slice maka akan membeli Danish Raisin 30 gr	5% 43%
14	Jika membeli Croissant Butter 30 gr dan Toast Brown Tips 10 slice maka akan membeli Toast Brown Tips 10 Slice	10% 100%
15	Jika membeli Croissant Butter 30 gr dan Danish Raisin 30 gr maka akan membeli Croissant Chocolate 30 gr	5% 38%
16	Jika membeli Croissant Butter 30 gr dan Danish Raisin 30 gr maka akan membeli Muffin Chocolate Choco Chips 40 gr	5% 38%
17	Jika membeli Croissant Butter 30 gr dan Danish Raisin 30 gr maka akan membeli Toast Brown Tips 10 Slice	5% 38%
18	Jika membeli Croissant Butter 30 gr dan Muffin Chocolate Choco Chips 40 gr maka akan membeli Toast Brown Tips 10 Slice	2% 34%
19	Jika membeli Toast White Full Loaf dan Danish Raisin 30 gr maka akan membeli Croissant Chocolate 30 gr	5% 100%
20	Jika membeli Toast White Full Loaf dan Danish Raisin 30 gr maka akan membeli Toast Brown Loaf Full	2% 34%
21	Jika membeli Toast White Full Loaf dan Long Bread 400 gr maka akan membeli Toast Brown Loaf Full	2% 34%
22	Jika membeli Toast White Full Loaf dan Croissant Chocolate 30 gr maka akan membeli Toast Brown Loaf Full	2% 34%
23	Jika membeli Toast White Tips 10 slices A dan Toast Brown Tips 10 slice maka akan membeli Toast Brown Tips 10 Slice	9% 100%
24	Jika membeli Toast Brown Tips 10 slice dan Danish Raisin 30 gr maka akan membeli Toast White Tips 10 Slice B	2% 34%
25	Jika membeli Toast Brown Tips 10 slice dan Danish Raisin 30 gr maka akan membeli Toast Brown Tips 10 Slice	5% 100%
26	Jika membeli Toast Brown Tips 10 slice dan Toast White Tips 10 Slice B maka akan membeli Toast Brown Tips 10 Slice	5% 100%
27	Jika membeli Danish Raisin 30 gr dan Croissant Chocolate 30 gr maka akan membeli Muffin Chocolate Choco Chips 40 gr	2% 34%
28	Jika membeli Danish Raisin 30 gr dan Muffin Chocolate Choco Chips 40 gr maka akan membeli Toast Brown Tips 10 Slice	2% 34%

Gambar 10. Hasil Laporan Data Transaksi

SIMPULAN

Berdasarkan penelitian yang dilakukan dengan merancang, membangun, menguji dan menganalisis, maka dapat diambil kesimpulan sebagai berikut:

- Dengan adanya penelitian ini, algoritma apriori pada proses analisis data *mining* terbukti dapat digunakan pada data penjualan produk Tan's Bakery sehingga dapat mengetahui hubungan pola frekuensi penjualan produk Tan's Bakery yang paling sering dibeli oleh konsumen melalui hasil analisis.

- Dengan menggunakan teknik data *mining* terbukti dapat membantu perusahaan dalam menemukan informasi yang sangat penting seperti keterkaitan antar produk melalui nilai support suatu produk terhadap produk yang lain, sehingga mampu memberikan informasi ketergantungan suatu produk terhadap produk yang lain.

- Melalui analisis ini, kita dapat mengetahui Asosiasi Data tertinggi yang terbentuk dari analisis data *mining* penjualan produk Tan's Bakery yaitu: *Croissant Butter 30 gr, Toast white slice 10 tipis, dan Toast Brown Slice 10 tipis* dengan nilai *support* adalah 6 *fk* dan *confident* 100%.

DAFTAR PUSTAKA

- [1] Erwin. 2009. *Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth*. Jurnal Generic, Vol. 4 No. 2.

- [2] Fayyad & Usama. 1996. *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- [3] Handojo, dkk. 2007. *Aplikasi Data Mining Untuk Meneliti Asosiasi Pembelian Item Barang di Supermarket Dengan Metode Market Basket Analysis*. Jurusan Teknik Informatika, Universitas Kristen Petra.
- [4] Jiawei dkk. 2006. *Data Mining : Concepts and Techniques*. United States of America : The Morgan Kaufmann.
- [5] Jogyanto HM, MBA, Akt, Prof. Dr. 1989. *Analisis & Desain Sistem Informasi : Pendekatan Terstruktur, Teori, dan Praktik Aplikasi Bisnis*. Yogyakarta: Penerbit Andi.
- [6] Kusriani, dkk. 2009. *Algoritma Data Mining*. Yogyakarta : Penerbit Andi.
- [7] Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons, Inc.
- [8] Moertini, dkk. 2007. *Analisis Keranjang Pasar Dengan Algoritma Hash-Based Pada Data Transaksi Penjualan Apotek*. Jurusan Ilmu Komputer, Universitas Katolik Parahyangan, Bandung.
- [9] Ponniah, P. 2001. *Datawarehouse Fundamentals: A Comprehensive Guide for IT Profesional*. John Willey & Sons, Inc.
- [10] Putra Asana Dwi. 2013. *Analisis Keranjang Belanja Dengan Algoritma Apriori pada perusahaan Retail*. Teknik Informatika, STMIK STIKOM Indonesia
- [11] Pramudiono, I. 2006. *Apa Itu Data Mining ?* Dalam http://datamining.japati.net/cgi/bin/in_dodm.cgi?bacaarsip&1155527614&a_rtikel. Diakses tanggal 28 Februari 2012.
- [12] Syaifullah Muhammad Afif. 2010. *Implementasi Data Mining Algoritma Apriori pada Sistem Penjualan*. Gava Media, Yogyakarta.
- [13] Turban, E., dkk. 2005. *Decision Support System and Intelligent Systems*. Yogyakarta: Andi Offset.
- [14] Widodo, dkk. 2011. *Menggunakan UML*. Bandung : Penerbit Informatika.