

ANALISIS KERANJANG BELANJA DENGAN ALGORITMA APRIORI PADA PERUSAHAAN RETAIL

A.A. Gede Bagus Ariana¹⁾, I Made Dwi Putra Asana²⁾

¹⁾STMIK STIKOM INDONESIA

Jln. Tukad Pakerisan No. 97 Denpasar (80225)

Telp : (0361) 256995, Fax : (0361) 246875

E-mail : gungariana@yahoo.com¹⁾

Abstrak

Sering terjadinya kekosongan salah satu stok barang yang sering dibeli secara bersamaan oleh pelanggan merupakan akibat dari tidak adanya informasi mengenai kebiasaan belanja pelanggan. Sehingga perlu dilakukan penggalian informasi pada data transaksi dengan teknik association rule mining (aturan asosiasi) untuk mengetahui barang yang sering dibeli secara bersamaan oleh pelanggan. association rule mining juga sering dinamakan market basket analysis (analisis keranjang belanja). Algoritma yang digunakan dalam association rule mining pada penelitian ini adalah algoritma apriori. Kombinasi barang dicari berdasarkan nilai minimum support, nilai minimum confidence, dan rentang data transaksi penjualan yang dimasukkan. Salah satu informasi yang dihasilkan adalah Jika membeli Cakra Kembar 25 Kg maka akan membeli Segitiga Biru 25 Kg dengan nilai support 11,76 % dan nilai confidence 43,24 %. Berdasarkan beberapa hasil uji coba dapat diketahui bahwa nilai minimum support yang dimasukkan mempengaruhi banyak frequent itemset yang dihasilkan. Nilai minimum support dan minimum confidence yang dimasukkan mempengaruhi waktu proses sistem dalam penggalian informasi.

Kata kunci: Association Rule Mining, Analisis Keranjang Belanja, Algoritma Apriori

Abstract

The vacancy stock of one product that happen when the customers are purchasing more than one item at the same time is as a result from lacking information of the customers shopping habit. Therefore, information extraction on data transaction through Association Rule Mining technique is necessarily needed in order to find out what products are usually bought at the same time by the customers. Association rule mining is also called the Market Basket Analysis. Algorithms which is used in Association Rule Mining in this study is Apriori Algorithm. The products combination is based on minimum support value, minimum confidence value, and sales transaction data range. One of the resulting information is if the customers buy 25 kilograms of Cakra Kembar then they would buy 25 kilograms of Segitiga Biru with support value 11.76 % and confidence value 43.24 %. Based on some test results, it can be seen that the minimum support value affected the number of frequent item set which is generated. The minimum support value and minimum confidence value affected the time process of system in the information extraction.

Key word: Association Rule Mining, Market Basket Analysis, Apriori Algorithm

1. PENDAHULUAN

Sistem informasi yang digunakan pada sebagian besar perusahaan hanya mampu memberikan informasi berdasarkan aktifitas transaksi dari sisi perusahaan. Sedangkan dengan memanfaatkan data yang telah tersimpan dapat digali beberapa informasi tersembunyi berdasarkan aktifitas transaksi dari sisi pelanggan, sehingga pihak pengambil keputusan dapat melakukan usaha untuk meningkatkan penjualan berdasarkan informasi aktifitas belanja pelanggan.

Perusahaan X adalah sebuah perusahaan retail bahan baku kue yang telah berdiri sejak tahun 2006. 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 10.000 item barang dan 3.000.000 transaksi penjualan. Sistem informasi yang digunakan tidak mampu memberikan informasi kebiasaan belanja pelanggan dalam membeli barang yang sering dibeli secara bersamaan. Sehingga sering terjadi kekosongan stok barang yang biasa dibeli bersamaan dengan barang lainnya.

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [1]. Teknik data mining yang digunakan untuk menemukan hubungan antar item yang dibeli pelanggan adalah analisis asosiasi (association rule mining). Tugas asosiasi dalam data mining adalah menentukan atribut yang muncul dalam satu waktu. [2]

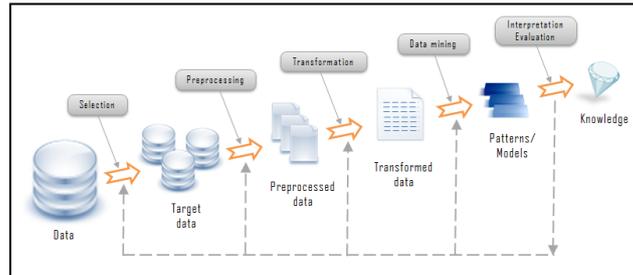
Pada penelitian ini menggunakan algoritma apriori untuk menemukan pola frekuensi tinggi. Algoritma apriori adalah algoritma analisis keranjang pasar yang digunakan untuk menghasilkan aturan asosiasi, dengan pola “*if-then*”. Algoritma apriori menggunakan pendekatan iteratif yang dikenal dengan *level-wise search*, dimana k-kelompok produk digunakan untuk mengeksplorasi (k+1)-kelompok produk atau (k+1)-itemset [3].

2. TELAAH PUSTAKA

Pada bab ini akan dijelaskan konsep dasar *data mining*, *association rule mining*, dan algoritma apriori

2.1 Data Mining

“*Data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual” [1].



Gambar 1. Tahapan Proses KDD

Istilah *data mining* dan *knowledge discovery in database* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* [3]. Proses KDD secara garis besar dapat dilihat pada Gambar 1.

2.2 Analisis Asosiasi (Association Rule Mining)

“Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*” [1]. Contoh aturan asosiatif dari analisis pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut, pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Menurut [2], Metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

- Analisis Pola Frekuensi Tinggi (*Frequent Itemset*)

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

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \quad (1)$$

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

$$Support(A, B) = \frac{\sum Jumlah\ transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi} \quad (2)$$

- Pembentukan Aturan Asosiasi (*Association Rule*)

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut.

$$Confidence = P(AB | A) = \frac{\sum Jumlah\ transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi\ mengandung\ A} \quad (3)$$

Salah satu contoh bentuk aturan asosiasi sebagai berikut [4] :

Roti tawar \rightarrow keju [*support* = 2%, *confidence* = 60%]

Seorang konsumen yang membeli roti tawar punya kemungkinan 60% untuk juga membeli keju. Aturan ini cukup signifikan karena mewakili 2 % dari catatan transaksi selama ini [2].

2.3 Algoritma Apriori

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 [4]:

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

3. METODE PERANCANGAN

Untuk mempermudah dalam penggalian data, dilakukan proses *denormalisasi* antara tabel transaksi penjualan dengan tabel barang ke dalam satu tabel *Data Mart*.

Tabel 1. Tabel tblDM

Nama Kolom	Tipe Data	Keterangan
NoNota	Varchar(10)	-
Tgl	Date	-
kodeBarang	Varchar(8)	-
namaBarang	Text	-

Kemudian untuk menyimpan itemset yang telah memenuhi minimum support diperlukan sebuah tabel, untuk mempermudah menemukan calon kandidat 1+k-itemset. Berikut adalah struktur tabel yang digunakan untuk menyimpan k-itemset.

Tabel 2. Tabel tblItemSet

Nama Kolom	Tipe Data	Keterangan
Kombinasi	Integer	-
Support	Float	-
Item1	Varchar(8)	-
Item2	Varchar(8)	-
ItemN	Varchar(8)	-

Setelah ditemukan *frequent itemset*, kemudian dilakukan pencarian nilai *confidence*. Nilai *confidence* yang sesuai dengan minimum *confidence* yang dimasukkan maka disimpan pada tabel hasil.

Tabel 3. Tabel tblHasilAnalisis

Nama Kolom	Tipe Data	Keterangan
Kombinasi	Integer	-
Asosiasi	Text	-
Confidence	Float	-
Support	Float	-

Perusahaan X hingga kini tercatat memiliki 10.000 item barang dan 500 kategori barang, adapun barang-barang yang dijual adalah bahan baku untuk membuat kue, alat-alat pengolah kue dan aksesoris kue.

Tabel 4. Contoh item barang yang dijual

Kode	Nama	Jenis	Satuan
010003	Cakra Kembar 1 Kg	Tepung Kereta Kencana	Kg
600828	Soft Pandan Mung Beans Pasta 18x1Kg 1 Crt	Soft Pandan	Karton
010007	Segitiga Biru 1 Kg LOS	Gunung Bromo	Kg
600669	PalmVita (13 Kg) 1 Crt	Palmvita	Karton
010015	Kunci Biru 1 Kg	Roda Biru	Kg
600737	Bali Asli Lime Marmalade 330 ml	Accessories	PCS
010028	Maizena RRC 1 Kg	Meizena Korea	Kg
010030	Maizena Korea 1/2Kg	Meizena Korea	PCS

010034	Maizena Hawaii 200 gr	MZ Hawaii	PCS
010039	Maizena Honig 750 gr	MZ Honig	PCS

4. HASIL DAN PEMBAHASAN

Aplikasi dibangun menggunakan Visual Basic 2005 dan sistem pengolah basis data SQL Server 2008 Express Edition. Pada aplikasi ini pengguna dapat mencari hubungan antar barang berdasarkan tanggal, waktu, triwulan dan tahun transaksi penjualan, tampilan menu dapat dilihat sebagai berikut.

Setelah dilakukan proses *mining* berdasarkan rentang data transaksi yang telah dipilih sebelumnya, berikut adalah informasi aturan asosiasi barang yang dihasilkan.

Kombinasi	Asosiasi	Support	Confidence
2	Jika membeli Cakra Kembar 25 Kg 1 Zak. 17/5 maka akan membeli Segitiga Biru 25 Kg 1 Zak. 17/5	11.7647058823529	43.2432432432
2	Jika membeli Segitiga Biru 25 Kg 1 Zak. 17/5 maka akan membeli Cakra Kembar 25 Kg 1 Zak. 17/5	11.7647058823529	45.7142857142857

Gambar 2 Hasil Analisis Tanggal 1 sampai 10 Mei 2009 Dengan Minimum Support 10 % dan Minimum Confidence 25 %

Berikut adalah contoh hasil pengujian, ditemukan 2 aturan asosiasi dengan kombinasi masing-masing 2 barang. Aturan pertama adalah :

- Jika membeli Cakra Kembar 25 Kg maka akan membeli Segitiga Biru 25 Kg dengan nilai *support* 11,76 % dan nilai *confidence* 43,24 %. Aturan tersebut juga dapat diartikan sebagai berikut : pada tanggal 1 sampai 10 Mei 2009 pelanggan yang membeli Cakra Kembar 25 Kg punya kemungkinan 43,24 % juga membeli Segitiga Biru 25 Kg. Aturan ini cukup signifikan karena mewakili 11,76 % dari catatan transaksi penjualan pada tanggal 1 sampai 10 Mei 2009 .
- Jika membeli Segitiga Biru 25 Kg maka akan membeli Cakra Kembar 25 Kg dengan nilai *support* 11,76 % dan nilai *confidence* 45,71%. Aturan tersebut dapat diartika sebagai berikut : pada tanggal 1 sampai 10 Mei 2009 Pelanggan yang membeli Segitiga Biru 25 Kg punya kemungkinan 45,71 % juga membeli Cakra Kembar 25 Kg. Aturan ini cukup signifikan karena mewakili 11,76 % dari catatan transaksi penjualan pada tanggal 1 sampai 10 Mei 2009

Berdasarkan hasil analisis tersebut, kemungkinan pelanggan membeli Segitiga Biru 25 Kg juga membeli Cakra Kembar 25 Kg lebih tinggi dibandingkan dengan kemungkinan pelanggan membeli Cakra Kembar 25 Kg juga membeli Segitiga Biru 25 Kg, hal tersbut dapat dilihat dari nilai *confidence* yang ditemukan.

Kombinasi	Asosiasi	Support	Confidence
2	Jika membeli Cakra Kembar 25 Kg 1 Zak. 17/5 maka akan membeli Segitiga Biru 25 Kg 1 Zak. 17/5	10.9125117991722	40.8738546279432
2	Jika membeli Segitiga Biru 25 Kg 1 Zak. 17/5 maka akan membeli Cakra Kembar 25 Kg 1 Zak. 17/5	10.9125117991722	44.869123879461
2	Jika membeli Cakra Kembar 25 Kg 1 Zak. 17/5 maka akan membeli Gula Lokal 50 Kg 1 Zak. 14/8/10	8.92796349952963	33.4390415765765
2	Jika membeli Gula Lokal 50 Kg 1 Zak. 14/8/10 maka akan membeli Cakra Kembar 25 Kg 1 Zak. 17/5	8.92796349952963	30.146124632007
2	Jika membeli Segitiga Biru 25 Kg 1 Zak. 17/5 maka akan membeli Gula Lokal 50 Kg 1 Zak. 14/8/10	8.5606772381609	35.1894818252127
2	Jika membeli Gula Lokal 50 Kg 1 Zak. 14/8/10 maka akan membeli Segitiga Biru 25 Kg 1 Zak. 17/5	8.5606772381609	28.9072426937738

Gambar 3 Hasil Analisis Tahun 2009 Dengan Minimum Support 8 % dan Minimum Confidence 25 %

Hasil uji coba dengan *minimum support* 8 % dan *minimum confidence* 25 % menghasilkan 6 aturan asosiasi dengan kombinasi masing-masing 2 barang.

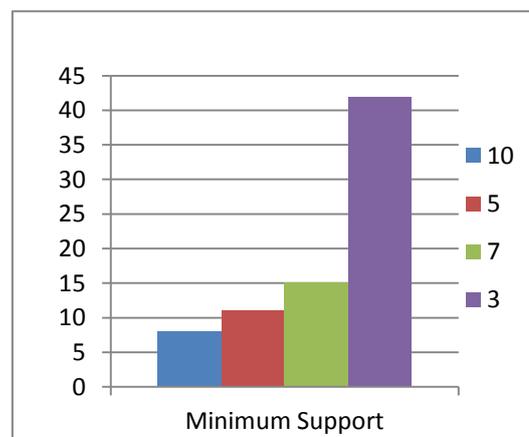
- Jika membeli Cakra Kembar 25 Kg maka akan membeli Segitiga Biru 25 Kg dengan nilai *support* 10,91 % dan nilai *confidence* 40,87 %. Aturan tersebut juga dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang membeli Cakra Kembar 25 Kg punya kemungkinan 40,87 % juga membeli Segitiga Biru 25 Kg. Aturan ini cukup signifikan karena mewakili 10,91 % dari catatan transaksi penjualan pada tahun 2009.
- Jika membeli Segitiga Biru 25 Kg maka akan membeli Cakra Kembar 25 Kg dengan nilai *support* 10,91 % dan nilai *confidence* 44,86 %. Aturan tersebut dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang membeli Segitiga Biru 25 Kg punya kemungkinan 44,86 % juga membeli Cakra Kembar 25 Kg. Aturan ini cukup signifikan karena mewakili 10,91 % dari catatan transaksi penjualan pada tahun 2009.
- Jika membeli Cakra Kembar 25 Kg maka akan membeli Gula Lokal 50 Kg dengan nilai *support* 8,93 % dan *confidence* 33,44 %. Aturan tersebut dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang membeli Segitiga Biru 25 Kg punya kemungkinan 44,86 % juga membeli Cakra Kembar 25 Kg. Aturan ini cukup signifikan karena mewakili 10,91 % dari catatan transaksi penjualan pada tahun 2009.
- Jika membeli Gula Lokal 50 Kg maka akan membeli Cakra Kembar 25 Kg dengan nilai *support* 8,93 % dan *confidence* 30,15 %. Aturan tersebut dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang

- membeli Gula Lokal 50 Kg punya kemungkinan 30,15 % juga membeli Cakra Kembar 25 Kg. Aturan ini cukup signifikan karena mewakili 8,93 % dari catatan transaksi penjualan pada tahun 2009.
- e. Jika membeli Segitiga Biru 25 Kg maka akan membeli Gula Lokal 50 Kg dengan nilai *support* 8,56 % dan *confidence* 35,19 %. Aturan tersebut dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang membeli Segitiga Biru 25 Kg punya kemungkinan 35,19 % juga membeli Gula Lokal 50 Kg. Aturan ini cukup signifikan karena mewakili 8,56 % dari catatan transaksi penjualan pada tahun 2009.
- f. Jika membeli Gula Lokal 50 Kg maka akan membeli Segitiga Biru 25 Kg dengan nilai *support* 8,56 % dan *confidence* 28,91 %. Aturan tersebut dapat diartikan sebagai berikut : pada tahun 2009 pelanggan yang membeli Gula Lokal 50 Kg punya kemungkinan 28,91 % juga membeli Segitiga Biru 25 Kg. Aturan ini cukup signifikan karena mewakili 8,56 % dari catatan transaksi penjualan pada tahun 2009.

Setelah dilakukan beberapa pengujian terhadap data penjualan pada perusahaan X, dapat diketahui bahwa nilai *minimum support* yang dimasukkan mempengaruhi jumlah *frequent item set* yang ditemukan. *Frequent itemset* adalah kombinasi frekuensi tinggi yaitu kombinasi barang yang memenuhi syarat persentase kemunculan pada transaksi penjualan (*minimum support*). Hubungan tersebut dapat dilihat pada salah satu tabel hasil uji coba berikut :

Tabel 5 Analisis Pukul 12:00 Sampai 18:00 WITA Tahun 2010 Pada Transaksi Penjualan

Rentang Data	Min. Supp	Min. Conf	Waktu Proses	Asosiasi	
12:00 – 18:00 Tahun 2010 Jumlah Transaksi 3510 Jumlah Barang 404	10	50	3 detik	Kandidat	0
		25	3 detik	425	2
		0	6 detik	Frequent	2
	7	50	4 detik	Kandidat	0
		25	4 detik	433	5
		0	6 detik	Frequent	6
	5	50	7 detik	Kandidat	1
		25	7 detik	463	6
		0	12 detik	Frequent	8
	3	50	43 detik	Kandidat	2
		25	44 detik	626	16
		0	45 detik	Frequent	25
42					



Gambar 4 Grafik Hubungan Nilai Minimum Support Dengan Frequent Item Set yang Ditemukan

Tabel 4 merupakan hasil uji coba analisis keranjang belanja pada rentang pukul 12:00 sampai 18:00 WITA pada tahun 2010. Berdasarkan Tabel 4 semakin kecil nilai *minimumsupport* yang dimasukkan maka akan menghasilkan lebih banyak *frequent itemset*. Setiap kombinasi barang dicari nilai persentasenya

pada data transaksi penjualan berdasarkan rentang yang dimasukkan. Kemudian setiap kombinasi yang memiliki nilai persentase kemunculan sama atau lebih besar dari nilai *minimum support* yang dimasukkan maka kombinasi tersebut adalah *frequent itemset* (kombinasi frekuensi tinggi). Berikut ini grafik hubungan nilai *minimum support* dengan *frequent itemset* yang ditemukan :

5. SIMPULAN DAN SARAN

Berdasarkan hasil uji coba dari aplikasi analisis keranjang belanja dengan algoritma apriori, dapat diambil beberapa kesimpulan sebagai berikut :

- a. Aplikasi analisis keranjang belanja dengan algoritma apriori dapat menentukan kecenderungan kebiasaan pelanggan dalam membeli barang-barang secara bersamaan.
- b. Dengan mengetahui informasi kebiasaan belanja pelanggan pihak perusahaan dapat melakukan beberapa tindakan untuk meningkatkan penjualan. Contoh tindakan yang dapat dilakukan seperti promosi harga untuk barang yang berpasangan, mengatur letak barang yang sering dibeli secara bersamaan pada satu rak, dan menjaga stok anggota barang yang sering dibeli secara bersamaan.
- c. Semakin kecil nilai *minimum support* yang dimasukkan maka akan menghasilkan lebih banyak *frequent itemset*.
- d. Semakin kecil nilai *minimum support* dan *minimum confidence* yang dimasukkan maka semakin lama waktu proses analisis.

6. DAFTAR RUJUKAN

- [1] Pramudiono, I. 2006. *Apa Itu Data Mining?* Dalam <http://datamining.japati.net/cgi-bin/indodm.cgi?bacaarsip&1155527614&artikel>. Diakses tanggal 28 Februari 2013.
- [2] Kusriani, & Luthfi, Emha Taufiq. 2009. *Algoritma Data Mining*. Yogyakarta : Penerbit Andi.
- [3] Fayyad, Usama. 1996. *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- [4] Handojo, Andreas, Satia, Gregorius Budhi, & Rusly, Hendra. 2007. *Aplikasi Data Mining Untuk Meneliti Asosiasi Pembelian Item Barang di Supermarket Dengan Metode Market Basket Analysis*. Jurusan Teknik Informatika, Universitas Kristen Petra.