PERANCANGAN DATAWAREHOUSE DAN PENERAPAN METODE ASSOCIATION UNTUK EKSTRAKSI DATA PENJUALAN PADA PT. CNI PALEMBANG

NurArifHidayat (Ndsarif@gmail.com), Nurtriana (Nurtriian@gmail.com)
Mardiani, S.Si, M.T.I. (Mardiani@stmik-mdp.net)
JurusanTeknikInformatika
STMIK MDP

Abstrak: Data warehouse merupakan sebuah cara yang digunakan untuk menampung data dalam jumlah besar, untuk membangun sebuah data warehouse dibutuhkan sebuah tools microsft SQL server 2008, microsoft SQL business intelligence development dan WEKA (Waikato Environment for Knowledge Analysis). Dengan menggunakan tools microsft SQL server 2008, microsoft SQL business intelligence development dan WEKA (Waikato Environment for Knowledge Analysis) dapat dibangun sebuah data warehouse penjualan dengan mengumpulkan data-data penjualan meliputi payment, customers id, gender, age, occupation, product id, product name, product cost, dan agency id yang dimasukan kedalam sebuahMs. Excel kemudian hasildari Ms. Excel dimasukan ke dalam database microsft SQL server 2008, setelah data masuk ke dalam database kemudian data tersebut di load kedalam microsoft SQL business intelligence development untuk dianalisis kemudian dimining menggunakan WEKA (Waikato Environment for Knowledge Analysis). Hasil analisis data warehouse dan data mining yang dibuat dapat memberikan informasi yang lebih jelas untuk pihak PT. CNI dan mempermudah dalam menganalisis barang-barang yang berpasangan danassociation rule dari pasangan barang tersebut.

Kata kunci: Data warehouse, data mining, microsoft SQL business intelligence development, microsoft SQL server 2008, WEKA, database, penjualan.

1 PENDAHULUAN

Perkembangan dunia bisnis yang sangat pesat, mendorong terbentuknya suatu timbunan data-data yang berukuran sangat besar. Informasi yang didapat dari data dalam bentuk tabel-tabel tersebut kurang dapat dimanfaatkan oleh pihak manajemen perusahaan dalam pengambilan keputusan untuk kemajuan perusahaan. Oleh karena itu perlu adanya aktivitas penggalian (ekstraksi) tersembunyi yang masih selanjutnya diolah menjadi pengetahuan yang bermanfaat dalam pengambilan keputusan.

PT.CNI Palembang memiliki volume produk dan transaksi yang cukup besar, perusahaan ini juga memiliki beberapa cabang lokasi sehingga menyulitkan perusahaan untuk mengetahui produk mana yang lebih cepat laku dan kurang diminati pelanggan per daerah.

2 LANDASAN TEORI

2.1 Data Warehouse

Menurut Feri Sulianta dan Dominikus dalam buku *Data Mining* Meramalkan Bisnis Perusahaan (2010, hal. 32), *data warehouse* adalah kumpulan data dari berbagai sumber yang ditempatkan menjadi satu dalam tempat penyimpanan berukuran besar lalu diproses menja dibentuk penyimpanan multidimensional dan didesain untuk *querying* dan *reporting*.

Karakteristik Data Warehouse:

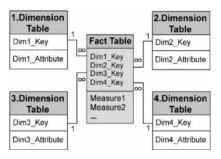
- a. Subject-oriented (Berorientasi Subjek)
- b. Integrated (Terintegrasi)
- c. Time Variant (Rentang Waktu)
- d. Nonvolatile

2.2 Dimensional Modelling

Beberapa konsep pemodelan data Warehouse pada dimensionality modeling yang dikenal pada umumnya, konsep-konsep tersebut adalah star schema, snowflake dan fact constellation schema.

2.2.1 Star Schema (Skema Bintang)

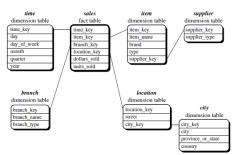
Skema bintang adalah sebuah logical struktur yang mempunyai sebuah table fakta berisi data terbaru di tengah, yang dikelilingi table dimensi yang berisi data referensi.



Gambar 1: Star Schema

2.2.2 SnowflakeSchema

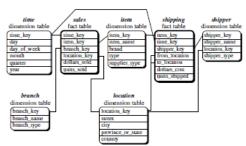
Skema kepingan salju adalah metode normalisasi table dimensi dalam skema bintang ketika anda benar-benar menormanlkan semua table dimensi, struktur yang dihasilkan menyerupai kepingan salju dengan table fakta ditengah.Dapat dilihat pada gambar 2.5 (Ponniah 2011, h.259).



Gambar 2: Snowflake Schema

2.2.3 Fact Constellation Schema

Fact constellation schema adalah skema multi dimensional yang berisikan lebih dari satu tabel fakta yang saling berbagi tabel dimensi.



Gambar 3: Constellation Schema

2.3 Data Mining

Menurut Vercellis (2009, h.77) *Data mining* adalah sebuah proses berulang bertujuan untuk menganalisa *database* yang besar dengan tujuan penggalian informasi dan pengetahuan yang berkaitan dengan pengambilan keputusan dan pemecahan masalah.

2.4 ETL (Extract, Transfrom, Load)

a. Extraction

Menurut Golfarelli dan Rizzi (2009), Extraction adalah pengambilan data yang relevant atau berkaitan dari sumber data. Extract merupakan proses yang pertama kali dilakukan dalam pengisian data warehouse.

b. Cleansing

Proses *cleansing* sangatlah penting dalam sistem data *warehouse* untuk meningkatkan kualitas data, karena normalnya data dari sumber data berkualitas rendah.

c. Transformation

Transformation mengubah format data dari sumber data operasional menjadi format data warehouse yang lebih spesifik.

2.5 Algoritma*Apriori*

Analisis asosiasi atau*association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*. Contoh aturan asosiatif dari analisis pembelian disuatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu.

Adapun tahap mencari kombinasi item yang memenuhi syarat minimum *support* dari nilai *support* dalam *database*. Nilai *support* sebuah item diperoleh dengan rumus berikut:

$$Support(A) = \frac{JumlahtransaksimengandungA}{TotalTransaksi}$$

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

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

$$Support(A, B)$$

$$= \frac{\sum TransaksimengandungAdanB}{\sum Transaksi}$$

Aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A→B.

Nilai confidence dari aturan A B di peroleh dari rumus beikut:

$$\begin{aligned} & Confidence &= P(B|A) \\ &= \frac{\sum TransaksimengandungAdanB}{\sum TransaksimengandungA} \end{aligned}$$

2.6 Weka (Waikato Environment for Knowledge Analysis)

Menurut (Feri Sulianta, hal.63) Weka aplikasi merupakan datamining berbasis open source (GPL) dan berengine Java. Pengertian Weka menurut situs resminva adalah sebagai berikut Weka adalah koleksi dari algoritma pembelajaran mesin untuk tugas-tugas data mining. Algoritma dapat diterapkan secara langsung ke data set atau disebut kode Java. Weka berisi alat untuk data pre-processing,

classification, regression, clustering, association rules, and visualization. Hal ini juga cocok untuk mengembangkan skema pembelajaran mesin baru.

2.7 Nine-step Methodology

Nine-step Methodology(Connolly dan Begg, 2005, h.1187). Kesembilan tahap itu yaitu:

- 1. Pemilihan Proses
- 2. Pemilihan *Grain*
- 3. Identifikasi dan penyesuaian
- 4. Pemilihan Fakta
- 5. Penyimpanan *pre-calculation* di tabel
- 6. Memastikan tabel dimensi
- 7. Pemilihan durasi database
- 8. Melacak perubahan dari dimensi secara perlahan
- 9. Penentuan prioritas dan model query

2.8 CRISP-DM (Cross Industry Standard Process for Data Mining)

Dalam penerapan *Data Mining* digunakan metodologi *CRISP-DM (Cross Industry Standard Process for Data Mining)* (Olson & Delen, 2008).

- 1. Business understanding
- 2. Data understanding
- 3. Data preparation
- 4. Modeling
- 5. Evaluation
- 6. Deployment

3 PERANCANGAN WAREHOUSE

DATA

3.1 Perancangan Data Warehouse

Perancangan *Data Warehouse* pada PT. CNI Palembang menggunakan metodologi sembilan tahapan *(nine-step methodology)* atau biasa disebut dengan metode Ralph Kimbal, metode ini meliputi :

a. Pemilihan Proses

Proses ini meliputi pendataan tiap-tiap pelanggan yang membeli produk CNI dari setiap agen-agen CNI. Adapun data – data atribut yang digunakan antara lain: Tanggal transaksi penjualan, kelompok transaksi, ID

pelanggan, kelompok kelamin, umur dan pekerjaan dari pelanggan, ID agen, nama agen, kota penjualan agen, ID produk, nama produk, harga produk, kelompok produk, tingkatan kelompok produk, jumlah barang yang dibeli, jumlah pembayaran, pajak pertambahan nilai dan total pembayaran dari transaksi pembelian yang dilakukan.

b. Pemilihan Grain

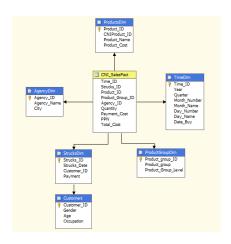
Grain merupakan proses untuk menentukan apa yang digambarkan oleh record di dalam tabel fakta. Terdapat empat perhitungan dalam tabel fakta yang akan dibuat di dalam data warehouse CNI ini yaitu Quantity atau jumlah barang, payment cost atau jumlah pembayaran tanpa PPN, adapun PPN atau Pajak Pertambahan Nilai dan total cost atau total pembayaran. Berikut adalah grain yang ada dalam perancangan Data Warehouse pada CNI Palembang.

c. Identifikasi dan Penyesuaian Dimensi

Pada tahap ini dilakukan identifikasi dan penyesuaian dimensi yang terkait dengan fakta yang ditampilkan dalam bentuk matriks. Dimensi yang digunakan dalam tabel fakta penjualan pada CNI Palembang adalah dimensi *time* (waktu), dimensi *strucks* (transaksi) dimensi *agency* (agen penjualan), dimensi *product* dan dimensi *product group*.

d. Pemilihan Fakta

Pada tahap ini dilakukan pemilihan fakta yang akan digunakan pada tabel fakta penjualan CNI. Adapun kolom – kolom pada tabel fakta penjualan CNI meliputi: Time_ID, Strucks_ID, Product_ID, Product_Group_ID, Agency_ID, Quantity, Payment_Cost, PPN, dan Total_Cost.



Gambar 4: Fakta Skema Snowflake

e. Penyimpanan *Pre Calculation* Ditabel Fakta

Dalam table fakta terdapat data yang merupakan kalkulasi awal. Hasil dari kalkulasi awal ini disimpan dalam tabel-tabel fakta. Adapun kalkulasi awal yang disimpan pada table CNI_SalesFact.

f. Memastikan Tabel Dimensi

Dalam tahap ini, kembali pada tabel dimensi dan menambahkan gambaran teks terhadap dimensi yang memungkinkan. Gambaran teks harus mudah digunakan dan dimengerti oleh user.

g. Pemilihan Durasi Database

Data yang dimasukan kedalam *Data Warehouse* ialah data transaksi penjualan produk pada PT.CNI Palembang selama tiga tahun.

h. Melacak Perubahan dari Dimensi Secara Perlahan

Dimensi-dimensi dan atribut yang mungkin dapat berubah dapat dilihat pada tabel.

Tabel 1: Kolom Dimensi yang Dapat Berubah

Nama Dimensi	Atribut yang Mungkin Berubah
AgencyDim	Agency_Name City
ProductDim	Product_Cost Product_Name
StrucksDim	Age Occupation

i. Penentuan Prioritas dan Model Query

Dalam tahapan ini yang dilakukan adalah mempertimbangkan pengaruh pada perancangan fisikal, seperti keberadaan dari ringkasan (summaries) dan penjumlahan (aggregate). Dalam perancangan data warehouse, kapasitas media penyimpanan merupakan salah satu aspek yang perlu dipertimbangkan.

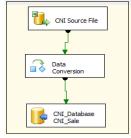
3.2 Extracting Data

Sumber data yang digunakan dalam perancangan *data warehouse* dan sumber data yang digunakan dalam perancangan *data warehouse* pada CNI Palembang ialah data *MS. Excel*.

Tabel 2: Sumber Data Penjualan CNI dalam Bentuk*Ms. Excel*

Payment	Customer_ID	Gender	Age	Occupation	Product Id	Product Name	Product Cost	Agency_II
Cash	CNIC135	Р	42	Pegawai Swasta	HF025	CNI VITASIGI F RASA APEL	42000	CNIA221
Cash	CNIC135	Р	42	Pegawai Swasta	N_07	PHYTO LITE SACHET	163000	CNIA221
Cash	CNIC135	P	42	Pegawai Swasta	HF024	CNI VITASIGI F RASA STRAWBERRY	42000	CNIA221
Cash	CNIC135	Р	42	Pegawai Swasta	PC091	FIRE SENS NEW LABEL	23000	CNIA221
Cash	CNIC132	L	32	Wiraswasta	FD 38S	SEHATI KECAP MANIS	14000	CNIA221
Cash	CNIC132	L	32	Wiraswasta	N_06	NARISH SACHET	175000	CNIA221
Cash	CNIC128	L	27	Wiraswasta	FD 03	PROGENE	610000	CNIA221

Kemudian data yang telah dikonversi akan dimasukkan kedalam Tabel CNI_Sale. Berikut gambar dari proses *Extracting Data*.



Gambar 5: Proses Ekstrak Data Ms. Excel

3.3 Transformation

Setelah data-data berhasil dimasukkan ke dalam tabel CNI_sale pada *database* CNI, isi dari tabel CNI_sale akan diisi ke beberapa tabel baru yang dibuat di dalam *database* CNI yaitu : tabel Agency, Customer, Products, Strucks, ProductGroup dan TimeTable.



Gambar 6: Paket Transformasi pada SSIS

3.4 Loading Tabel Dimensi dan Tabel Fakta

Proses ini merupakan tahapan pembuatan tabel-tabel dimensi pada database DW_CNI yang bertujuan untuk mengisi tiap tabel-tabel dimensi, setiap isi tabel dimensi berasal dari tabel database CNI yang sudah terpisah. Berikut gambar proses loading tabel dimensi dan tabel fakta



Gambar 7: Proses *Loading* Dimensi dan Fakta

4 ANALISIS DATA WAREHOUSE DAN HASIL DATA MINING

4.1 Presentasi DataWarehouse

Adapun hasil informasi yang didapatkan pada *data warehouse* PT. CNI Palembang yang berisikan data-data penjualan produk selama 3 tahun terakhir dari tahun 2011-2013 untuk setiap periode waktu, antara lain:

a. Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Kelamin, Umur dan Pekerjaan Pelanggan

Total pembayaran dan jumlah barang berdasarkan jenis kelamin, umur dan pekerjaan pelanggan perbulan dapat dilihat pada tabel.

Tabel 3:Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Kelamin, Umur dan Pekerjaan Pelanggan Perbulan

			Year ▼	Month Name												
			☐ 2011													
			January		February		March		April		May					
Gender ▼	Age ▼	Occupation ▼	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost				
ÐΙ	FI 19	Mahasiswa			54	9954900.6335	49	3946936.4	32	5465972.1875	20	5972199.2				
						Pegawai Swasta										
			PNS	28	2834100.5	16	880850.3			11	833191.7499	15	667451.4			
				Wiraswasta	73	17085394.6221	43	4934600	76	11370294.2375	97	12808005.0307	56	11510914.0384		
		Tetal	101	1001040E 1221	112	15770250 0225	125	15317230 6375	140	10107169 0691	01	10100004 6004				

Pada tabel 4.1 PT.CNI Palembang dapat melihat jumlah penjualan produk dari pelanggan pengendara dengan status mahasiswa yang berumur 19 tahun dan berjenis kelamin laki-laki sejumlah 54 item produk dengan total pembeliannya sebesar Rp. 9.954.900,6335 untuk bulan Februari 2011.

b. Informasi *Data Warehouse* CNI Berdasarkan Nama dan Kota Agen Penjualan

Tabel 4: Total Pembayaran dan Jumlah Barang Berdasarkan Nama dan Kota Agen Penjualan Pertahun

			Year ▼							
			2011	2011		2012		2013		ital
City	•	Agency Name	 Quantity 	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost
⊞ Belitung			1520	262773121.2392	1272	204308391.1018	772	159759131.7658	3564	626840644.10680
⊞ Bengkulu			13264	2005987855.442	6610	999967237.1161	8531	1237840018.9516	28405	4243795111.5097
□ Jambi		BERMAWAN SIHOTANG	1468	213706985.9793	1349	231955848.5776	645	91418869.8054	3462	537081704.3623
_		CV. ANEKA CEMERLANG	2984	475855172.2963	3441	563256603,986	1382	194407544.4229	7807	1233519320.7052
		PT.SUKSES TUGU MANDIRI	4469	655032032.2771	3776	544509987,9699	3150	502074875.6054	11395	1701616895.8524
	ŀ	Total	8921	1344594190.5527	8566	1339722440.5335	5177	787901289.8337	22664	3472217920.9199

Pihak CNI palembang dapat melihat total jumlah penjualan produk dari 5 kota penjualan dan nama agen di wilayah Sumatera. Dapat dilihat pada agen 'CV. ANEKA CEMERLANG', pada kota 'Jambi' terjual 1382 item produk dengan total pembayaran sebesar Rp. 194.407.544,4229 jumlahnya menurun drastis yang dibandingkan hasil dari tahun dua sebelumnya.

c. Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Transaksi dan Nama Produk

Tabel 5: Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Transaksi dan Nama Produk

		Year ▼	Year ▼								
		2011	2011			2013					
Payment *	Product Name	me ▼ Quantity		Quantity	Total Cost	Quantity	Total Cost				
	Super SC 88 Laundry Compound - sachet	38	1669145.6876	40	1757457.1544	16	699680.3096				
	SUPER SOFT	6	145200			5	121000				
	TAS WANITA 2 MERAH	16	220000	12	165000	3	41250				
	TAS WANITA PRODUKTIF BIRU	20	319000	8	127600	10	159500				
	TRAY - UNTUK IC009	4	40000.4	7	70000.7						
	TUTUP ATAS	6	759000	9	1138500						

PT. CNI Palembang dapat melihat dari tabel 4.3 berdasarkan nama produk Super SC 88 Laundry Compound – Sachet yang jumlahnya ada 38 unit dan total pembayaran.

d. Total Pembayaran dan Jumlah Barang Berdasarkan Kelompok dan Tingkatan Kelompok Barang

Tabel 6: Total Pembayaran dan Jumlah Barang Berdasarkan Kelompok dan Tingkatan Kelompok Barang

		1	rear ▼	Month Name ▼				
		E	⊞ 2011		⊞ 2012			
Product Group Level ▼	Product Group	- 0	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost
⊞3	HSHCDG88 - DETERGEN	1	159	8347070.1097	126	6462489.2006	90	4604595.7707
	HSHCSR - SOFTENER	2	2002	40607600	1717	34468500	1172	23532300
	PMPEOT - PROMO EVENT OTHERS	5	9	21285000	23	11561000	26	9350000
	PMPRFB - PROMO FOOD & BEVERAGES	1	122	11316263.2	179	17410951.8	65	6273443
	PMPRHF - PROMO HEALTH FOOD	5	52	20020000	20	7700000	3	1155000

PT. CNI Palembang dapat melihat tingkatan kelompok barang yang ke-3 pada tahun 2011 dan kelompok produk HSHCDG88-DETERGEN dari tahun 2011 sebanyak 159 item dengan total pembayaran sebesar Rp. 8.347.070,1097.

e. Analisis Pertumbuhan Jumlah Penjualan Produk dan Total Pembayaran Penjualan Produk

Tampilan analisis pertumbuhan jumlah penjualan produk dan total pembayaran transaksi perbulan yang tercatat dalam database PT. CNI Palembang.

Tabel 7: Pertumbuhan Data Perbulan

	rear ·	rional name .	IVIIUI Naii	ic tillic billi calci	JIGUVIIS *								
	□ 2013												
	∃ Januar	-у	☐ Februa	iry	⊞ March								
			Current Time Dim		Month Ov	er Month Growth %	1						
City •	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost					
Belitung							61	17164038.1					
Bengkulu	651	79741624.6	1193	178705836.3203	83.26%	124.11%	743	101640700.7607					
Jambi	317	54824954.5438	445	64498180.0263	40.38%	17.64%	723	117261303.2093					
Muara Bungo													
Palembang	1360	166371160.0287	628	104857437.7627	-53.82%	-36.97%	904	146274907.0315					
Grand Total	2328	300937739.1725	2266	348061454.1093	-2.66%	15.66%	2431	382340949.1015					

Dapat dilihat dari tabel 7 terjadi peningkatan jumlah penjualan produk pada kota penjualan Bengkulu sebesar 83.26 % dari bulan Januari ke bulan Februari 2013 dan total pembayaran transaksi sebanyak 124,11 %.

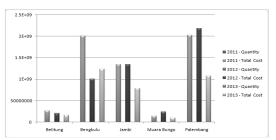
4.2 Laporan Hasil Analisis Data Warehouse

Berikut ini merupakan hasil laporan jumlah penjualan produk dan total Total Pembayaran untuk setiap nama agen per tahun dalam bentuk tabel pada *Microsoft Office Excel*.

Tabel 8: Laporan Jumlah Penjualan Produk dan Total Pembayaran CNI Palembang Selama Tiga Tahun dalam Bentuk Tabel

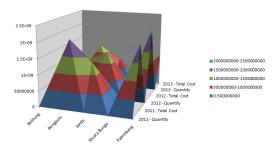
	Column Labels 💌							
	2011		2012		2013		Total Quantity	Total Total Cost
Row Labels	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost		
BERMAWAN SIHOTANG	1468	213706986	1349	231955848.6	645	91418869.81	3462	537081704.4
CV. ANEKA CEMERLANG	2984	475855172.3	3441	563256604	1382	194407544.4	7807	1233519321
CV. MAMINTA NAULI	3266	488306308.3	2673	431405306	1605	244612641.8	7544	1164324256
CV. MANDIRI BERSAMA	1430	207788033	2387	363937769.7	861	116215272.9	4678	687941075.7
CV.BANGUN JAYA	1393	232164302.1	1564	227429359.1	753	99002451.96	3710	558596113.2
CV.GLOBALINDO INSAN MANDIRI	3115	473454921.4	2569	456347408.8	2271	316948049.8	7955	1246750380
CV.JAYA SUWANDI PERSADA	9879	1470034692	4482	685263159.9	6739	969852193.2	21100	3125150045
HADI OKTORA LUBIS	4052	619935886.8	4493	702981261.8	1944	297346651.2	10489	1620263800
IR.ANTONIUS MARBUN	3385	535953163.4	2128	314704077.2	1792	267987825.8	7305	1118645066
PT.SUKSES TUGU MANDIRI	4469	655032032.3	3776	544509988	3150	502074875.6	11395	1701616896
RACHMAT BASUKI	848	139679988.9	1393	234490543.6	657	99396522.13	2898	473567054.6
TIORLI SIMBOLON	1520	262773121.2	1272	204308391.1	772	159759131.8	3564	626840644.1
Grand Total	37809	5774684608	31527	4960589718	22571	3359022030	91907	14094296356

Berikut ini tampilan beberapa laporan dalam bentuk grafik batang dan *surface*. Laporan dalam bentuk grafik batang dapat dilihat pada gambar yang menunjukkan jumlah Total Pembayaran untuk setiap kota penjualan selama 3 tahun.



Gambar 8: Laporan Jumlah Penjualan Produk dan Total Pembayaran Per Kota Penjualan dalam Bentuk Grafik Batang

Laporan jumlah Penjualan produk untuk setiap Kota Penjualan perbulan selama tiga tahun dalam bentuk diagram *surface*.



Gambar 9: Laporan Jumlah Penjualan Produk Per Kota Penjualan Per Bulan dalam Bentuk Diagram *Surface*

Pihak PT. CNI Palembang juga dapat melihat grafik informasi jumlah Penjualan produk dan Total Pembayaran dari berbagai bentuk grafik lainnya ataupun dapat juga mengkombinasikan antara dimensi dengan dimensi lainnya, dengan kategori yang dipilih sesuai kebutuhan.

4.3 Penerapan Data Mining

Pada Proses penerapan data miningpada CNI Palembang, metode yang digunakan adalah metodologi CRISP-DM (Cross Industry Standard Process for Data Mining).

a. Business Understanding (Pemahaman Bisnis)

Tahap pertama dari penerapan data mining ini ialah menenentukan tujuan bisnis dan tujuan *Data Mining* serta rencana proyek lebih lanjut. Adapun tujuan bisnis yang ingin dicapai dari penerapan data mining pada PT. CNI Palembang ialah peningkatan keuntungan dan kinerja pada perusahaan.

b. Data Understanding (Pemahaman Data)

Adapun data yang digunakan untuk dapat memenuhi tujuan bisnisdantujuan *data mining* yang telah ditetapkan adalah tabel transaksi pada CNI_DataWarehouse dalam bentuk format *CSV file* yang dapat dibentuk melalui *MS. excel*. Tabel transaksi penjualan yang ada pada *database* CNI_Data Warehouse akan diubah sesuai kebutuhan.

Tabel 9: Data Transaksi dalam Format CSV File



c. Data Preparation (Persiapan Data)

Setelah sumber data telah tersedia untuk diidentifikasi, pada tahap ini akan dilakukan persiapan. Pembentukan tabel yang dibutuhkan dalam proses *mining* menggunakan *apriori* pada Aplikasi *WEKA*.

d. Modeling

Langkah keempat dari proses *data mining* adalah membangun sebuah model *data mining*. Adapun langkah-langkah pembentukan model mining menggunakan algoritma Apriori pada *Mining Tools WEKA* antara lain:

- 1. Memilih.file format csv yang akan *dimining*.
- 2. Memilih algoritma yang digunakan dalam hal ini dipilih *algoritma apriori*.
- 3. Mengatur minimal *support*, *confidence* dan *rules* yang dihasilkan.

e. Evaluation

Pada tahap ini akan dilakukan pengujian untuk menentukan frequent item set dan association rules dengan minimal support 0.1 dan minimal confidence 0.25.Pengujian dilakukan menggunakan data penjualan bulan januari 2011 dari table CNISales_fact. Berikut hasil 10 rules teratas yang dihasilkan.



Gambar 10: Frequent Item Set yang Dihasilkan



Gambar 11: *Rules* yang Dihasilkan dari Pengujian

f. Deployment

Pada tahap ini akan dijalankan mining seluruh transaksi penjualan produk selama tiga tahun pada PT. CNI Palembang. Adapun *file CSV* yang dibentuk terdiri dari 8443 baris transaksi dan 293 kolom nama produk.

Dengan persetujuan pihak PT. CNI Palembang, pembentukan association rules menggunakan algoritma apriori pada aplikasi WEKA dijalankan dengan minimum support sebesar 0,1 atau 84 kali kemunculan barang dan nilai minimum confidence sebesar 0,25 yang artinya aturan dengan tingkat kepercayaan rendah akan ditampilkan. Dari mining yang dilakukan dihasilkan 10 assosiation rules.

4.4 Laporan Hasil Analisis Data Mining

Setelah analisis selesai dilaksanakan maka dapat ditarik beberapa informasi yang telah didapatkan diantaranya:

- 1. Berdasarkan pengujian yang dilakukan menggunakan data penjualan tiga tahun dari tabel CNISales_fact dengan min *support* 0,1 dan min *confidence* 0,25, berikut hasil 1000 *rules* teratas yang dihasilkan sebanyak 13 item produk yang sering muncul, 5 pasang item produk yang berkaitan, dan 10 pasang item produk *best rules*.
- 2. Berdasarkan hasil yang telah di prediksi dengan menggunakan WEKA, item produk yang saling berkaitan yaitu:

 confidence 0,80, "Jika membeli Sehati Kecap Manis maka membeli Narish Sachet".

confidence 0,75, "Jika membeli Super MPC Sachet new maka membeli Narish Sachet".

confidence 0,74, "Jika membeli Bio Plus 1.2 gr isi 10 stick maka membeli Progene". confidence 0,73, "Jika membeli Sigi-F Hand & Body Lotion (new bottle) maka membeli CNI Vitasigi F Rasa Strawberry". confidence 0,72, "Jika membeli CNI Sarung Sejahtera maka membeli Mie Ginseng Soto Ayam – Ekonomis", "Jika membeli Mie Ginseng Soto Ayam – Ekonomis maka membeli CNI Sarung Sejahtera", dan "Jika membeli CNI Vitasigi F Rasa Strawberry maka membeli Sigi-F Hand & Body Lotion (New Bottle)".

confidence 0,68, "Jika membeli Progene (1strip) maka membeli Bio Plus 1.2 Gr Isi 10 Stick".

Untuk melihat gambaran hasil prediksi yang telah dilakukan dengan menggunakan WEKA dapat dilihat pada gambar dibawah ini.

```
Best rules found:

1. STEAII METCAP MANIS-t 1284 => NABISE SACKET=t 1037 conf:(0.8)
2. STEPA NFC SACKET new=1 1560 => NABISE SACKET=t 108 conf:(0.75)
3. BIO FURS 12.6 RISI 10 STOR(0.1565 => PROBER (SIREF)=1655 conf:(0.75)
4. SIG1-F BAND 4 BODY LOTION (NEW BOTTLE)-t 1173 => CHI VITASIGI F RASA STRAMBERRY=t 861 conf:(0.72)
5. CHI SARUNG SELRATERA=t 1439 => MIE SIRESNG SUD AIMA - EMERGENE=t 102 conf:(0.72)
7. CHI VITASIGI F RASA STRAMBERRY=t 1193 => SIG1-F BAND 6 BODY LOTION (NEW BOTTLE)=t 861 conf:(0.72)
8. HOGORDE (ISTRE)=t 1177 -> BID FURS 1.2 GR 131 10 SIGNED SALUNG SACKET-1042 conf:(0.88)
9. NABISE SACKET=t 2635 => STEAII METCAP MANIS-t 1097 conf:(0.39)
```

Gambar 12: Rules hasil Pengujian

5 PENUTUP

5.1 Kesimpulan

- a) Terjadi penurunan drastis dua tahun sebelumnya pada agen 'CV.ANEKA CEMERLANG' Kota 'Jambi' dengan 1382 item produk terjual.
- b) Pada tahun 2011 dapat dilihat berdasarkan dari nama produk Super SC 88 Laundry Compound Sachet yang jumlahnya ada 38 unit dengan total pembayaran sebesar Rp. 1.669.145.6876.

- c) Dapat dilihat tingkatan kelompok barang yang ke 3 pada tahun 2011 dan kelompok produk dari tahun 2011 sebanyak 159 item dengan total pembayaran sebesar Rp. 8.347.070,1097.
- d) Terjadi peningkatan jumlah penjualan produk pada Kota 'Bengkulu' sebesar 83.26% dari bulan Januari ke bulan Februari 2013 dan total pembayaran transaksi sebanyak 124,11%.
- e) Transaksi penjualan produk selama tiga tahun pada PT.CNI Palembang terdiri dari 8443 baris transaksi dan 293 kolom nama produk.
- f) Pembentukan Association Rules menggunakan Algoritma Apriori pada aplikasi WEKA dijalankan dengan support 0,1 dan confidence 0,25.
- g) Dapat dilihat jumlah penjualan produk dari pelanggan PT.CNI Palembang dengan status pekerjaan, umur, jenis kelamin perbulannya.

5.2 Saran

- a) Diharapkan data yang digunakan dalam pembentukan datawarehouse ini bukan hanya pada wilayah Sumatera Selatan dan sekitarnya saja, melainkan seluruh cabang penjualan PT. CNI yang ada di Indonesia.
- b) PT. CNI Palembang agar menambah proses lain selain penjualan barang, misalnya pembelian barang ataupun proses lainnya.
- c) Digunakan software data mining yang lain selain Weka misalnya Ms. Visual Studio, Rapidminer ataupun software lainnya dengan algoritma yang sama ataupun berbeda sehingga pihak PT. CNI dapat membandingkan association rules yang dihasilkan.

DAFTAR PUSTAKA

- [1] Connolly, Thomas, C. B 2010, Database System: A Practical Approach to Design, Implementation, and Management Fifth Edition, Pearson Education Inc.
- [2]Cuzzocrea, Alfredo 2011, Data Warehouseing and Knowledge Discovery.

 London: Springer-Verlag Berlin Heidelberg.
- [3] Indrajani 2009, Sistem Basis Data dalam Paket Five In One, PT. Elex Media Komputindo, Jakarta
- [4] Kimball, R, Margy R, Warren T, Joy M and Bob B 2008. *The Data Warehouse Lifecycle Toolkit*, Wiley Publishing Inc, Canada
- [5] Komputer, Wahana 2010, Panduan Belajar MySQL Database Server, Media Kita, Jakarta Selatan
- [6] Kusrini 2009, *Algoritma Data Mining*, Andi Offset, Yogyakarta
- [7] Nuraida, Ida 2008, *Manajemen Administrasi Perkantoran*, Kanisius,
 Yogyakarta
- [8] Ponniah, Paulraj 2011, *Data Warehouseing*, John Wiley & Sons Inc, Canada
- [9] Rainaldi, Vincent 2008, Building A Data Warehouse With Examples In SQL Server, Apress
- [10] Sulianta, Feri dan Dominikus Juju 2010, Data Mining Meramalkan Bisnis Perusahaan, Elex Media Komputindo, Jakarta
- [11] Tantra, Rudi 2012, Manajemen Proyek Sistem Informasi, Andi Offset, Yogyakarta