

MEMBANGUN DATA MINING DENGAN SQL SERVER 2005 BUSINESS INTELLIGENCE

Emha Taufiq Luthfi
Dosen STMIK AMIKOM Yogyakarta

Abstraksi

Data mining melakukan proses penggalian terhadap data untuk mendapatkan pola tersembunyi. Terdapat berbagai teknik atau algoritma yang digunakan dalam melakukan data mining. Artikel ini membahas cara melakukan data mining dengan algoritma association menggunakan SQL Server 2005 Business Intelligence.

Kata Kunci : *data mining, association, SQL Server Business Intelligence*

Pendahuluan

Data mining merupakan proses analisis data dan penemuan pola tersembunyi secara otomatis maupun semi otomatis. Data mining merupakan konsep utama dari Business Intelligence, Online Analytical Processing (OLAP), Enterprise Reporting. Berbagai business software yang ada seperti Financial Application, Enterprise Resource Planner (ERP), Customers Relation Management (CRM) telah membuat perusahaan memiliki atau kaya akan data akan tetapi belum banyak pengetahuan atau manfaat yang didapat dengan data yang ada tersebut. Data mining dapat menyelesaikan berbagai masalah bisnis antara lain :

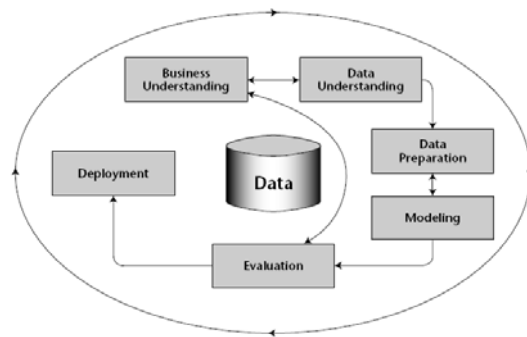
- Churn Analysis
Mengenali kemungkinan pelanggan berpindah ke pesaing
- Cross-Selling
Mengenali produk yang banyak dipesan pelanggan

- Fraud Detection
Mengenali klaim asuransi yang ilegal
- Risk Management
Mengurangi resiko, persetujuan kredit nasabah
- Customer Segmentation
Siapakah pelanggan kita ?
- Targeted Ads
Banner apa yang harus ditampilkan thd pengunjung tertentu
- Sales Forecast
Berapa besar penjualan minggu ini
- Classification
Memasukan data ke dalam suatu kategori berdasar attribute tertentu
- Clustering / Segmentation
Mengidentifikasi kelompok-kelompok kasus berdasar sekumpulan attribute
- Association / Market Basket Analysis
Mengidentifikasi produk-produk yang mungkin dibeli bersama dengan pembelian suatu produk
- Regression
Serupa dengan classification dengan attribute berupa nilai continue
- Forecasting
Memperkirakan nilai untuk masa yang akan datang

Hal tersebut dapat dilaksanakan dengan berbagai teknik yang ada dalam data mining seperti :

- Regression
- Time Series
- Decision Trees
- Neural Network
- Genetic Algorithm

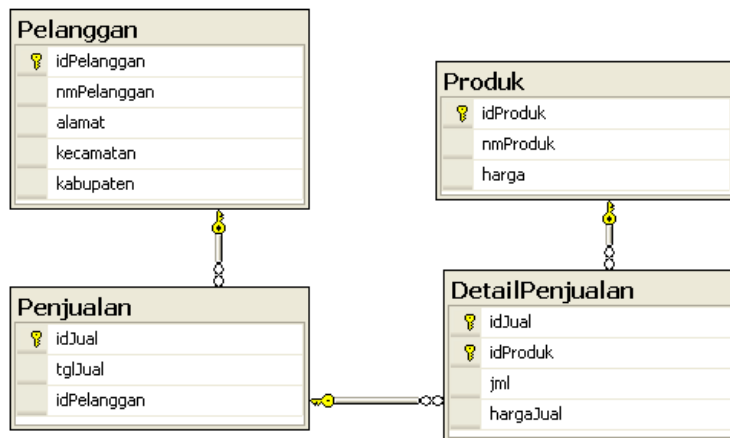
Siklus pengembangan data mining menurut CRISP-DM terlihat dalam gambar 1.



Gambar 1 Siklus Data Mining CRISP-DM

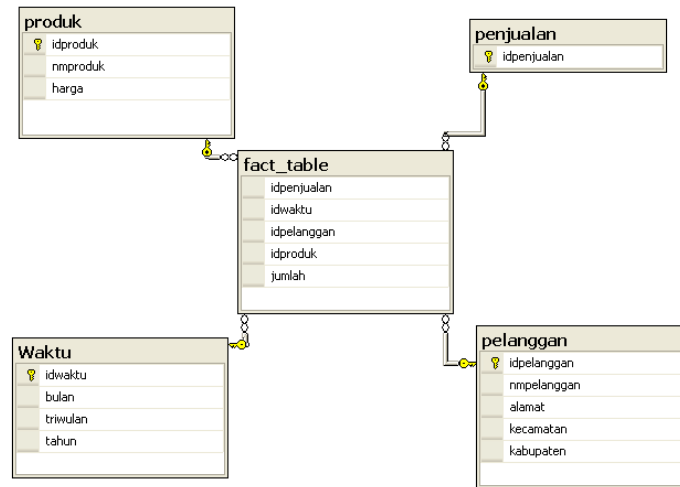
Data Mining Dengan Sql Server

Akan coba dilakukan data mining association untuk database penjualan dengan menggunakan SQL Server Business Intelligence.



Gambar 2 Design Database Penjualan

Sebelum dapat dilakukan data mining, dilakukan proses data warehouse untuk membangun star schema dari database penjualan yang ada.



Gambar 3 Star Schema Database Penjualan

Selanjutnya dilaksanakn proses ETL untuk membawa data dari database transaksi penjualan ke dalam database warehouse model star schema penjualandw. Dalam SQL Server nantinya dapat di buat dalam DTS dengan script sbb :

```
delete from [PenjualanDW].[dbo].[Fact_table]
go

delete from [PenjualanDW].[dbo].[Pelanggan]
go

Insert into [PenjualanDW].[dbo].[Pelanggan] SELECT [idPelanggan]
,[nmPelanggan],[alamat],[kecamatan],[kabupaten] FROM
```

```

[Penjualan].[dbo].[Pelanggan]
go

delete from [PenjualanDW].[dbo].[Produk]
go

Insert into [PenjualanDW].[dbo].[Produk]
SELECT [idProduk],[nmProduk],[harga] FROM
[Penjualan].[dbo].[Produk]
go

delete from [PenjualanDW].[dbo].[Penjualan]
go

Insert into [PenjualanDW].[dbo].[Penjualan]
SELECT [idJual] FROM [Penjualan].[dbo].[Penjualan]
go

delete from [PenjualanDW].[dbo].[Waktu]
go
DBCC CHECKIDENT ('[PenjualanDW].[dbo].[Waktu]',RESEED,0)
go

Insert into [PenjualanDW].[dbo].[Waktu]
SELECT distinct month([tglJual]),
               Triwulan = CASE month([tglJual])
                WHEN '1' THEN '1'
                WHEN '2' THEN '1'
                WHEN '3' THEN '1'
                WHEN '4' THEN '1'
                WHEN '5' THEN '2'
                WHEN '6' THEN '2'
                WHEN '7' THEN '2'
                WHEN '8' THEN '2'
                WHEN '9' THEN '3'

```

```

        WHEN '10' THEN '3'
        WHEN '11' THEN '3'

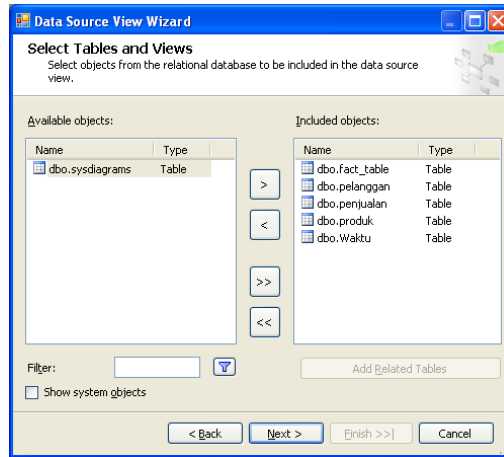
        WHEN '12' THEN '3'
    END,
    year([tglJual]) FROM [Penjualan].[dbo].[Penjualan]
go

delete from [PenjualanDW].[dbo].[fact_table]
go

insert into [PenjualanDW].[dbo].[fact_table] SELECT A.idJual,
(select idwaktu from [PenjualanDW].[dbo].[Waktu]
 where bulan=month(A.tglJual) and tahun=year(A.tglJual)) as
IdWaktu,
A.idPelanggan, B.idProduk, B.jml
FROM [Penjualan].[dbo].[Penjualan] A
inner join [Penjualan].[dbo].[DetailPenjualan] B
on A.idJual=B.idJual
go

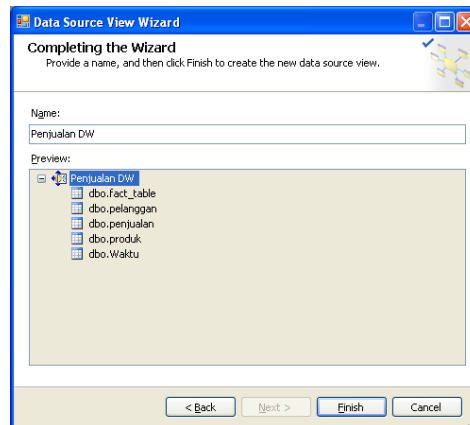
```

Selanjutnya melaksanakan data mining menggunakan SQL Server Business Intelligence diawali dengan membangun data source. Tentukan tabel dan view yang akan digunakan, pilih : **fact_table**, **penjualan**, **waktu**, **produk**, **pelanggan**

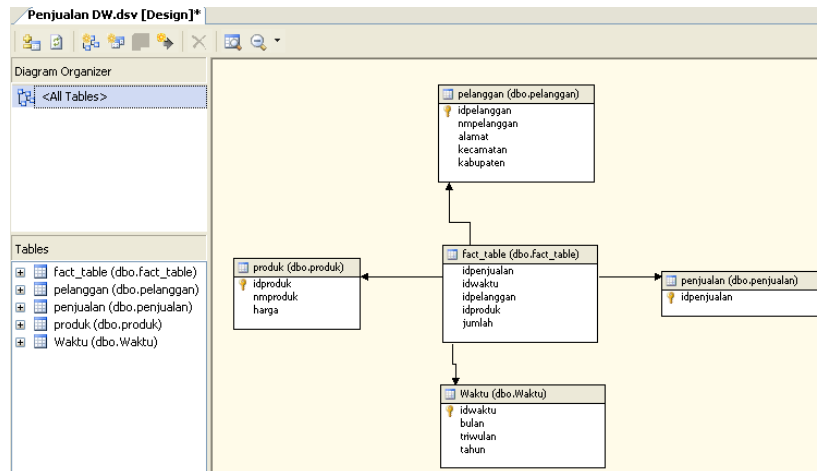


Gambar 4 Data Source

Selanjutnya tentukan untuk data source view.



Gambar 5 Data Source View



Gambar 6 Schema Star

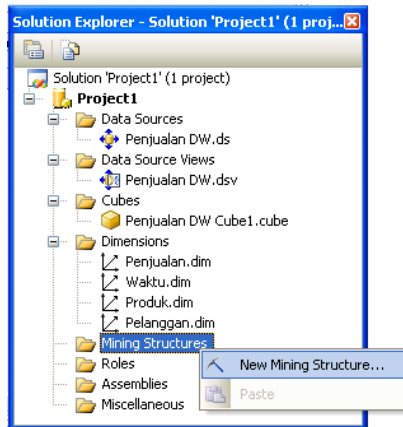
Selanjutnya membuat Cube merupakan suatu kubus yang biasa digunakan untuk model analisis dalam OLAP (Online Analytical Processing). Cube mengandung sekumpulan Dimensi dan Measure. Dimensi merupakan acuan pembuatan agregasi, misal dalam penjualan data source view terdapat dimensi : Pelanggan, Produk, Penjualan, Waktu. Measure merupakan nilai yang akan diagregasikan.

	Bulan ▼		Triwulan ▼		Tahun ▼		Grand Total
	9		Total		10		
	3		Total		3		
	2007	Total	Jumlah	Jumlah	2007	Total	
Nmproduk ▼	Jumlah	Jumlah	Jumlah	Jumlah	Jumlah	Jumlah	Jumlah
Coklat	1	1	1	1	1	1	2
Ice Cream	2	2	2				2
Kecap	1	1	1				1
Keju	3	3	3	1	1	1	4
Korek	1	1	1				1
Mie Instant	14	14	14	12	12	12	26
Minuman Soda	11	11	11	1	1	1	12
Rokok	2	2	2				2
Roti Tawar	4	4	4	2	2	2	6
Saos	1	1	1	1	1	1	2
Snack	7	7	7				7
Grand Total	47	47	47	18	18	18	65

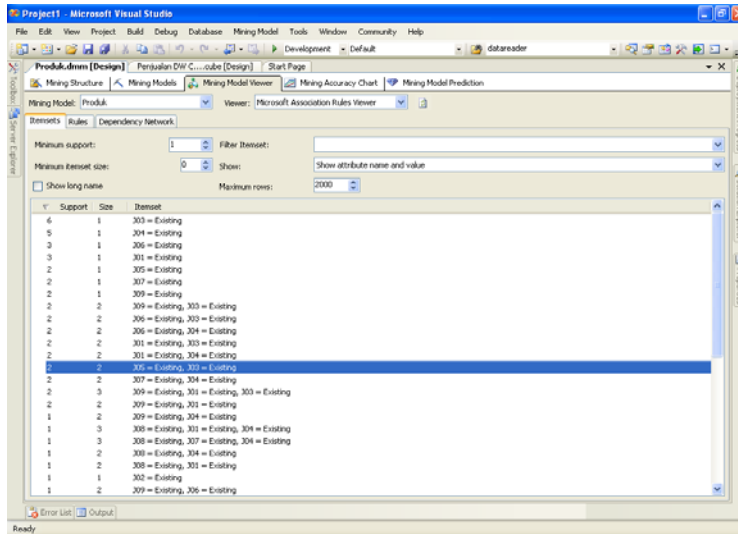
Penjualan	Kecamatan		Kabupaten		wirocrajan		Grand Total
	condong catur		mergangsan		kota yogyakarta		
	jumlah	jumlah	jumlah	jumlah	jumlah	jumlah	
J01			13	13			13
J02	2	2					2
J03			20	20			20
J04	8	8					8
J05					4	4	4
J06			3	3			3
J07	11	11					11
J08	2	2					2
J09			2	2			2
Grand Total	23	23	38	38	4	4	65

Gambar 7 Cube Penjualan

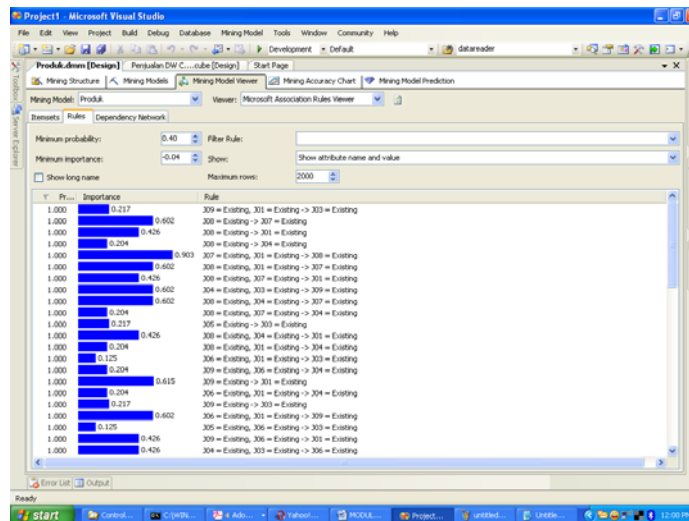
Setelah semua langkah tersebut tinggal membuat data mining analisis asosiasi.



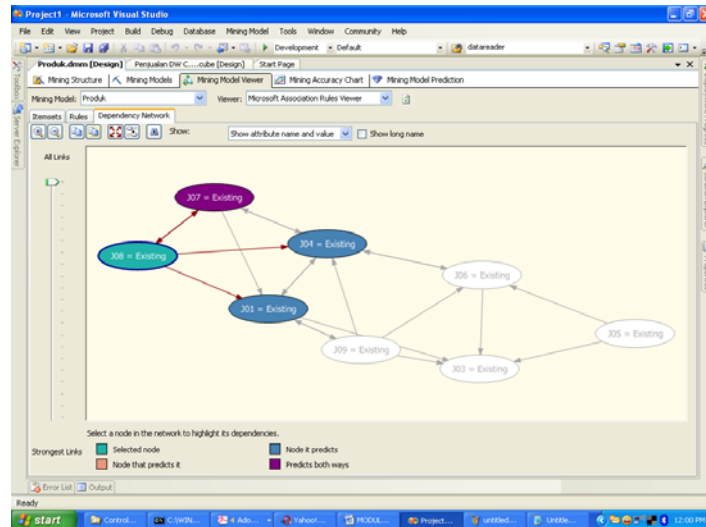
Gambar 8 New Mining Structure



Gambar 9 Association Itemsets



Gambar 9 Association Rule



Gambar 10 Hasil Association Dependency Network

Penutup

Data Mining merupakan proses penggalian data untuk menemukan pola tersembunyi dari data. Analisis Association mencari hubungan antar kemunculan data. Data Mining Analisis Association dapat dilaksanakan dengan mudah menggunakan SQL Server 2005 Business Intelligence

Daftar Pustaka

-