

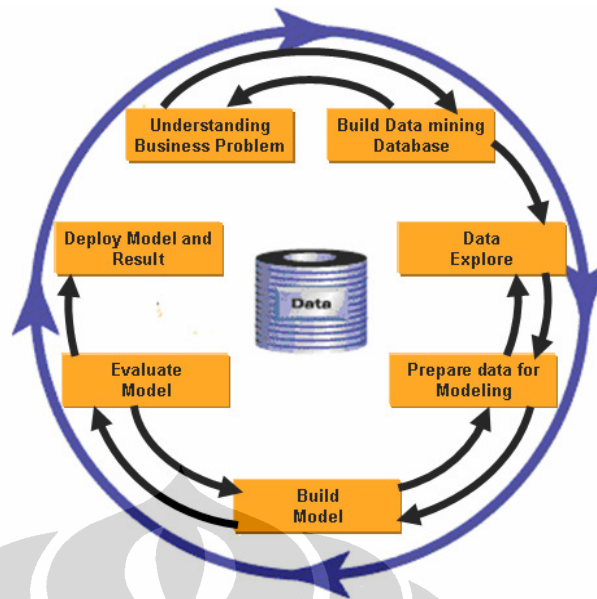
### BAB 3

## METODOLOGI PENELITIAN

Metodologi penelitian *data mining* pada prinsipnya merupakan kegiatan pencarian pengetahuan atau lebih dikenal dengan *Knowledge Discovery*. Dalam penelitian ini menggunakan menggunakan metodologi *data mining two crows-corporation*. Adapun pertimbangan menggunakan metodologi ini berkaitan dengan adanya tahapan yang tidak dimiliki oleh metodologi lainnya. Tahapan-tahapan yang membedakannya adalah:

- a. Adanya tahapan *Business Understanding* yang menjadi tahapan awal dalam proses KDD, tahapan ini sangat penting berkaitan dengan pemahaman terhadap latar belakang, permasalahan, tujuan dan manfaat KDD bagi bisnis atau organisasi. Karena hal itu menjadi fundamental sekali apakah KDD diperlukan atau tidak.
- b. Adanya tahapan *data explore* atau penyelidikan data sebelum dilakukan *data mining*. Pada tahapan ini berusaha menemukan sejak awal fakta-fakta yang terjadi dengan melakukan proses *summaries and visualization*. Dengan harapan dapat memunculkan/mengungkapkan *hidden information* yang masih ada. Dalam tahapan ini dibutuhkan alat bantu OLAP untuk melakukan *explore* terhadap *database* yang telah dibuat berupa *data warehouse*.

Dapat dikatakan metodologi *Two Crows Corp* dapat menjawab tantangan suatu organisasi dalam upaya menemukan potensi *hidden information* yang didasarkan atas latarbelakang dari kepentingannya dengan memanfaatkan potensi data yang dimiliki. Adapun tahapan metodologi ini ditunjukkan dalam gambar 3.1.



Gambar 3.1 Tahapan *Data Mining Two Crows-Corporation* (Adaptasi *CRISP-DM, 2005*).

Selanjutnya tahapan *Two Crows-Corporation for Data Mining* dapat dijelaskan sebagai berikut:

1. *Understanding Bussiness Problem*

Pada tahapan awal ini, perlu dilakukan pemahaman mengenai data dan bisnis Industri Telekomunikasi khususnya untuk produk dan layanan berbasis *Wireline* meliputi kebijakan perusahaan, proses bisnis dan informasi-informasi yang terkait.

Aktifitas yang dilakukan di tahapan ini dimulai dari pemahaman kebijakan dan regulasi bidang telekomunikasi yang telah ditetapkan oleh pemerintah Indonesia. Kemudian dilanjutkan dengan mempelajari kebijakan internal seperti ketentuan-ketentuan yang mengatur produk *Wireline* berupa Keputusan Direksi, Keputusan *Executive General Manager*, Dokumen *Bussiness Plan*, Dokumen Proses Bisnis, Serta kebijakan-kebijakan berkaitan dengan Program Promosi Produk, Pengelolaan *Network* dan *Customer Care/Retention*. Di samping itu melakukan pembicaraan dan diskusi dengan pejabat/karyawan dari unit-unit yang terkait khususnya pengelola produk *wireline* seperti Unit *Billing and Collection*, *Fixed Phone Sales*, *Customer Care*, *Access Network Maintenance*, *Access Network Operation*, *Bussiness Performance* dan lain-lain.

## 2. *Build Data Mining Database*

Tujuan tahapan ini adalah membuat *database* yang digunakan dalam proses *data mining*. Adapun hal-hal yang perlu dipenuhi adalah *data collection*, *data description*, *data selection*, *data quality assesment* dan *data cleansing*. Serta perlu diperhatikan juga *Consolidation* dan *Integration*.

Kegiatan yang dilakukan di tahapan ini yaitu pemahaman terhadap bisnis model dan logika model dari sumber data yang ada di Operator Telekomunikasi berupa data *billing* ataupun data *Call Duration Record*, data kinerja *Network* dan lain-lain, yang dilanjutkan dengan pembuatan model dimensi (*Star Schema*) dan Proses *extraction*, *transforming* dan *loading* (ETL). Pada tahapan ini perlu dilakukan diskusi dengan pejabat/karyawan di unit terkait sebagai konfirmasi dan validasi data untuk menjaga keakuratan dari *database data mining* yang telah dibuat, yaitu dengan Unit *Billing and Collection*, *Access Network Maintenance/Operation* dan *IS/IT Centre Area*.

## 3. *Data Explore*

Pada tahapan ini dilakukan penyelidikan dan mengidentifikasi informasi penting mengenai data yang akan digunakan pada saat pencarian pengetahuan di proses *data mining*. Pengamatan grafik dan tabulasi data hasil dari pengolahan secara multidimensional data dengan menggunakan teknologi *On line Analytical Processing* (OLAP). Hasil dari pengerjaan di tahapan ini adalah berupa *interface* program OLAP yang dapat digunakan sebagai pemenuhan kebutuhan informasi yang diperlukan oleh unit *Customer Care* dan *Bussiness Performance*.

Data yang digunakan di proses OLAP adalah data dari pengolahan *data billing* pelanggan dan perfomansi *Network* yang berada di masing-masing *Switching*. Masing-masing *switching* melayani suatu wilayah/area, yang mana antara wilayah satu dengan yang lainnya memiliki karakteristik perilaku pelanggan yang berbeda.

Sedangkan periode data yang digunakan untuk proses ini meliputi data *billing* dan kinerja *Network* tahun 2007 sampai tahun 2008.

#### 4. *Prepare Data for Modelling*

Pemanfaatan data secara menyeluruh merupakan langkah yang ideal, akan tetapi kurang praktis dalam pengerjaan *data mining*. Untuk itu pemilihan produk *wireline* yang digunakan dalam proses *data mining* dapat dikatakan sudah memenuhi. Karena dari produk *wireline* ini layanan yang diberikan cukup beragam mulai dari percakapan Lokal, Sambungan Langsung Jarak Jauh, Sambungan Langsung Internasional, dan Akses Internet.

Data yang dipersiapkan sebagai *data training* dalam proses pemodelan *data mining* menggunakan data *billing* dan kinerja *network* dari bulan januari 2007 sampai dengan Desember 2008 . Adapun lokasi pelanggan berasal dari Sepuluh *switching* yang berada di wilayah kantor Unit Bisnis Telekomunikasi.

#### 5. *Build Model*

Dalam tahapan ini dilakukan pembuatan model. Hal yang terpenting dalam tahapan ini adalah proses *iterative* dalam pemodelan. Dengan melakukan teknik dan algoritma *data mining* yang berbeda untuk mendapatkan penyelesaian permasalahan bisnis yang sesuai. Adapun pemodelan *data mining* dengan teknik *Segmentasi* , *Classification* atau teknik lainnya. Untuk Algoritma *Segmentasi* menggunakan Algoritma *K-Means*.

Proses pengerjaan *data mining* ini dilakukan secara berulang dengan menggunakan *data training* yang telah disiapkan. Penggunaan pemodelan *data mining* dan penampilan output hasil *data mining* dalam bentuk grafik sangat membantu dalam menguji hasil model yang diperoleh sehingga hasil dari pemodelan *data mining* dapat dianalisa secara lebih menyeluruh.

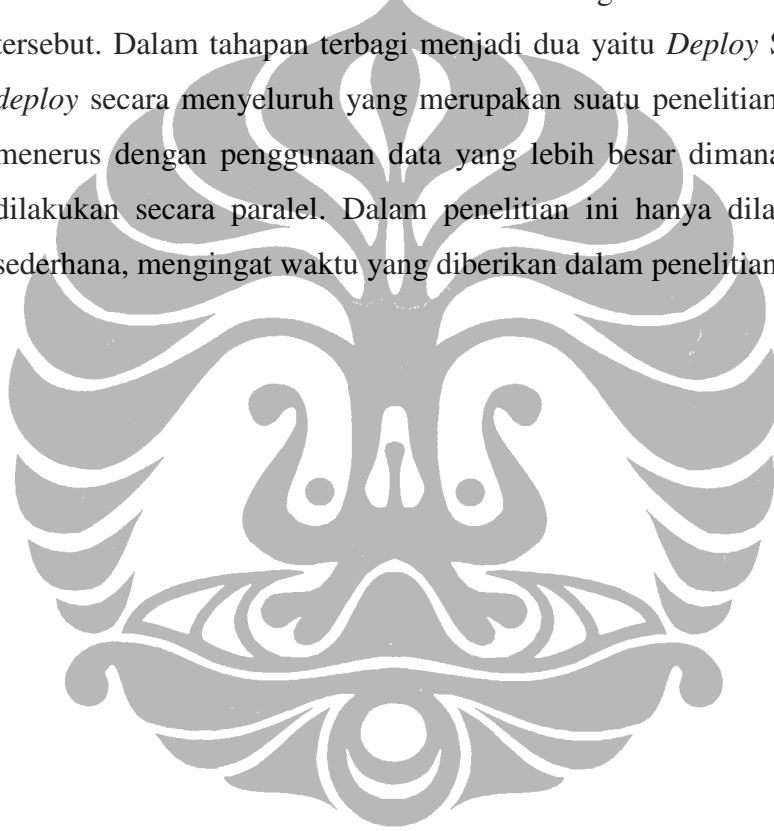
#### 6. *Evaluate Model*

Kemudian setelah model dibuat, tahapan selanjutnya adalah melakukan evaluasi dari hasil model-model yang telah terbentuk berupa pola-pola *usage* pelanggan *wireline* selama periode pengamatan. Model-model ini kemudian diterjemahkan sebagai informasi yang penting atau bernilai tinggi.

Bahwa hasil model ini merupakan hasil penemuan dari pengujian data yang disiapkan untuk pembuatan model. Maka diperlukan evaluasi dan validasi terhadap pola awal yang terbentuk, guna menguji kualitas dan efektifitas model sebelum diputuskan bahwa model yang ditemukan layak untuk dilakukan *deploy*.

#### 7. *Deploy Model and Result*

Tahapan ini dapat dilakukan ketika model *data mining* yang sudah dibuat sudah divalidasi. Dari model-model yang dihasilkan, perlu diuji dengan data yang berbeda dan dilakukan evaluasi kembali mengenai keakuratan dari hasil model tersebut. Dalam tahapan terbagi menjadi dua yaitu *Deploy* Secara sederhana dan *deploy* secara menyeluruh yang merupakan suatu penelitian intensif secara terus menerus dengan penggunaan data yang lebih besar dimana proses *data mining* dilakukan secara paralel. Dalam penelitian ini hanya dilakukan *deploy* secara sederhana, mengingat waktu yang diberikan dalam penelitian ini terbatas.



## BAB 4

### PROSES Pengerjaan DATA MINING

Dalam penelitian ini terdiri dari 7 (tujuh) tahapan seperti halnya yang sudah dijelaskan di Bab 3 (tiga) mengenai metodologi penelitian. Semua tahapan dimasukkan di dalam bab ini karena merupakan satu rangkaian aktifitas yang utuh. Adapun tahapan tahapan pengerjaan proses *data mining* seperti di bawah ini.

#### 4.1 Pemahaman Permasalahan Bisnis

Penyediaan layanan *wireline* masih merupakan kegiatan bisnis utama industri telekomunikasi. Karena masih menyumbangkan kontribusi pendapatan cukup besar di samping itu juga menjadi *costumer base* atau *entry point* untuk penetrasi produk-produk baru lainnya. Sehingga pengelolaan pelanggan *wireline* dilakukan secara komprehensif mulai dari bagaimana upaya peningkatan *sales*, peningkatan *usage*, penanganan komplain dan sebagainya.

##### 4.1.1 Latar Belakang Permasalahan.

Tingkat persaingan industri telekomunikasi semakin meningkat semenjak Pemerintah menggulirkan kebijakan penghapusan monopoli sektor telekomunikasi. Sehingga agar tetap bertahan menghadapi persaingan, pelaku industri ini membutuhkan strategi dan segala upaya agar layanannya lebih unggul dibanding pesaing dan bahkan dapat meningkatkan skala bisnisnya.

Banyaknya pilihan produk layanan yang diberikan oleh operator telekomunikasi dan adanya kebebasan pelanggan memilih sesuai kebutuhannya, menjadikan penyediaan layanan yang beragam dan lengkap merupakan strategi yang diterapkan sebagai bentuk diferensiasi dari pesaing. Sehingga para operator berlomba-lomba mengemas produk dan layanan semenarik mungkin dengan mengkombinasikannya berupa paket produk dan tarif yang lebih kompetitif. Di samping itu kualitas layanan menjadi perhatian pula, gangguan yang sering terjadi berakibat *brand image* bisa menurun, bahkan opportunity mendapatkan *revenue* hilang. Sehingga kinerja layanan *network* juga menjadi penting.

#### 4.1.2 Perumusan Masalah.

Data pola pemakaian produk/layanan dan pola gangguan saat ini belum dimanfaatkan secara maksimal untuk penetapan strategi dan kebijakan pelayanan maupun pemasaran. padahal dengan menggunakan teknologi *data mining* akan diperoleh informasi-informasi yang tersembunyi yang kemungkinan dapat digunakan untuk menentukan strategi dan kebijakan tersebut.

Proses Pencarian pengetahuan dengan menggunakan teknologi *data mining* dengan menggunakan teknik *asosiation rules* diharapkan akan ditemukan informasi-informasi tersembunyi terkait dengan pola-pola sebagai berikut :

- Pola *usage* untuk layanan lokal, sambungan langsung jarak jauh (SLJJ), pemakaian internet dan Sambungan Langsung Internasional (SLI).
- Pola gangguan seperti jumlah gangguan dan lamanya gangguan.
- Karakteristik pemakaian di suatu wilayah pelanggan berdasarkan boundery/area catuan *switching* atau berdasarkan wilayah pemerintahan seperti Kelurahan.
- Pola pemakaian berdasarkan bulan kalender.
- Pola pemakaian berdasarkan *customer category*.

#### 4.1.3 Tujuan dan Manfaat

Berdasarkan perumusan permasalahan yang dihadapi tujuan penulisan ini adalah menemukan *Association Rules* pada beberapa pola data seperti dijelaskan dalam bagian 4.1.2 dengan menerapkan teknologi *data mining*. Hasil interpretasi terhadap temuan *Association Rules* nantinya dapat dimanfaatkan sebagai bahan analisis untuk membantu manajemen PT. XYX dalam pengambilan keputusan berkaitan dengan peningkatan pelayanan, Upaya penetrasi pemasaran produk baru dan upaya meretensi pelanggan *existing*.

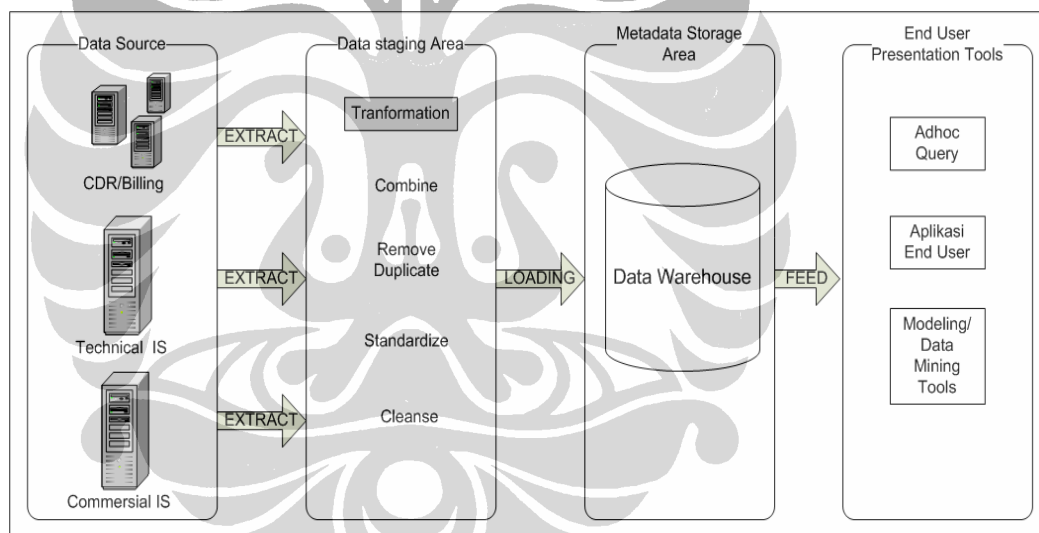
Adapun manfaat adanya Teknologi *data mining*, proses analisis terhadap data akan menjadi lengkap yaitu analisis dilakukan lebih terstruktur dan lebih mendalam. Diharapkan informasi-informasi yang tersembunyi dapat terungkap dengan jelas. Sehingga diharapkan output berupa strategi yang akan diterapkan tepat sasaran dan dapat dilaksanakan secara efektif.

## 4.2 Pembuatan *Database Data Mining*

Berdasarkan proses *knowledge discovery in database* dan metodologi yang digunakan, *database* yang digunakan adalah *data warehouse*. Adapun Pertimbangan pemilihan *data warehouse* karena hal-hal sebagai berikut sudah dipenuhi yaitu :

- *Data collection*
- *Data data Description*
- *Data Quality Assessment dan data Cleansing*
- *Consolidation dan Integration.*

Selanjutnya dengan melakukan adaptasi arsitektur *data warehouse* dari Hoffer, Prescott, McFadden maka dalam penelitian ini arsitektur *data warehouse* yang dibuat seperti gambar 4.1. berikut ini.



Gambar 4.1 Arsitektur *Data Warehouse* (Adaptasi dari Hoffer, Prescott, McFadden, 2005)

Penjelasan proses *extract*, *transformation* dan *loading* akan dijelaskan pada subbagian 4.2.6. Kemudian tahapan-tahapan pembuatan *data warehouse* berdasarkan arsitektur *data warehouse* dapat dijelaskan sebagai berikut:

### 4.2.1 Sumber Data

Sumber data merupakan data transaksional yang tercatat secara harian. Adapun sumber data yang digunakan untuk pembuatan *data mining* berasal dari



*billing* bulanan pelanggan PT XYZ di Unit Bisnis yang menangani pelanggan *wireline* dari sejumlah *switching* yang ditanganinya. Data tersebut terkompulasi dari proses *billing* tiap *switching* selama 2 (dua) tahun dari tahun 2007 sampai dengan tahun 2008. Data tersebut tersimpan di data *center* yang dikelola oleh Sistem Informasi *Centre*. Data *Legacy* ini tersimpan dalam *database oracle 8i*.

Kemudian sumber data lain adalah berasal data operasional pengelolaan *network* yaitu *technical information system*, di dalam system ini tersimpan informasi alat produksi berupa *access network* dan *fault handling*. Setiap proses bisnis berkaitan dengan pengelolaan *access network* harus menggunakan *tools* ini. Adapun proses bisnis yang dimaksud adalah proses *new installation*, *fault handling* dan *access network maintenance*.

Selanjutnya sumber data lain yang digunakan adalah berkaitan dengan data dari pengelolaan *customer* yaitu *Commercial Information System*. System IS ini merupakan mencatat semua data transaksional *customer* meliputi proses *new installation*, pindah alamat, ganti nomor, mutasi fitur layanan, berhenti berlangganan dan lain-lain.

Adapun secara ringkas sumber data yaitu database yang dimanfaatkan dalam penelitian ini beserta spesifikasinya adalah sebagai berikut :

No	Nama Database	Nama Tabel	Jumlah Kolom	Jumlah Row	Size	Sumber Data
1	Commercial IS	Client	37	462,719	390 byte/row	Oracle 8i
		Dossier	83	637,600	446 byte/row	Oracle 8i
2	Technical IS	Paire	29	570,863	188 byte/row	Oracle 8i
		Fault	9	313,193	116 byte/row	Oracle 8i
3	Billing	Usage	7	13,173,609,930	89 byte/row	Oracle 8i

Tabel 4.1 Daftar *database* dan *table* sebagai sumber data

#### 4.2.2 Pemahaman Bisnis

Untuk pemahaman bisnis, kegiatan yang dilakukan adalah mencari informasi dan data yang dibutuhkan berkaitan dengan pengelolaan produk *wireline* mulai dari proses *new intallation* sampai dengan terbitnya *billing*. Disini dilakukan pembahasan dan diskusi dengan karyawan atau pejabat berwenang di bagian/dinas yang berbeda seperti : *Unit Sales*, *Unit Customer care*, *Unit Bussiness Performance*, *Unit Access Network*, *Unit IS/IT* serta *Unit Billing and Collection*. Adapun data yang didiskusikan adalah sebagai berikut:

1. Dokumen Proses bisnis *New installation* dan *Fault handling* termasuk form-form yang digunakan dalam proses tersebut.
2. Dokumen Proses bisnis *Billing Processing*, proses *loading* dari tiap *switching* ke *data center* sampai dengan terbentuknya *billing* yang ditagihkan ke pelanggan.
3. Format laporan kinerja Unit Bisnis yang terbentuk dari *data billing* maupun dari *data fault* (gangguan) yang dialami oleh pelanggan.
4. Pemahaman *data billing*, *fault handling* yang ada di *data centre*.
5. Dokumen-dokumen berkaitan dengan permintaan end *user* akan kebutuhan data yang digunakan untuk pengelolaan unit-unit terkait dan berhubungan dengan data *wireline*.

Kemudian dari semua dokumen yang diperoleh dilakukan validasi kebutuhan data dan pemahaman kebutuhan informasi berkaitan dengan produk *wireline* khususnya data *billing* dan *fault handling*. Di samping itu juga dilakukan konfirmasi pemahaman kepada unit terkait terhadap dokumen yang kurang memadai.

#### **4.2.3 Pemahaman Logika**

Dalam tahapan ini proses yang dilakukan adalah memahami keterkaitan antara data yang satu dengan data yang lainnya. Prinsip-prinsip relasi database sudah digunakan sebagai landasan memahami hubungan antar data.

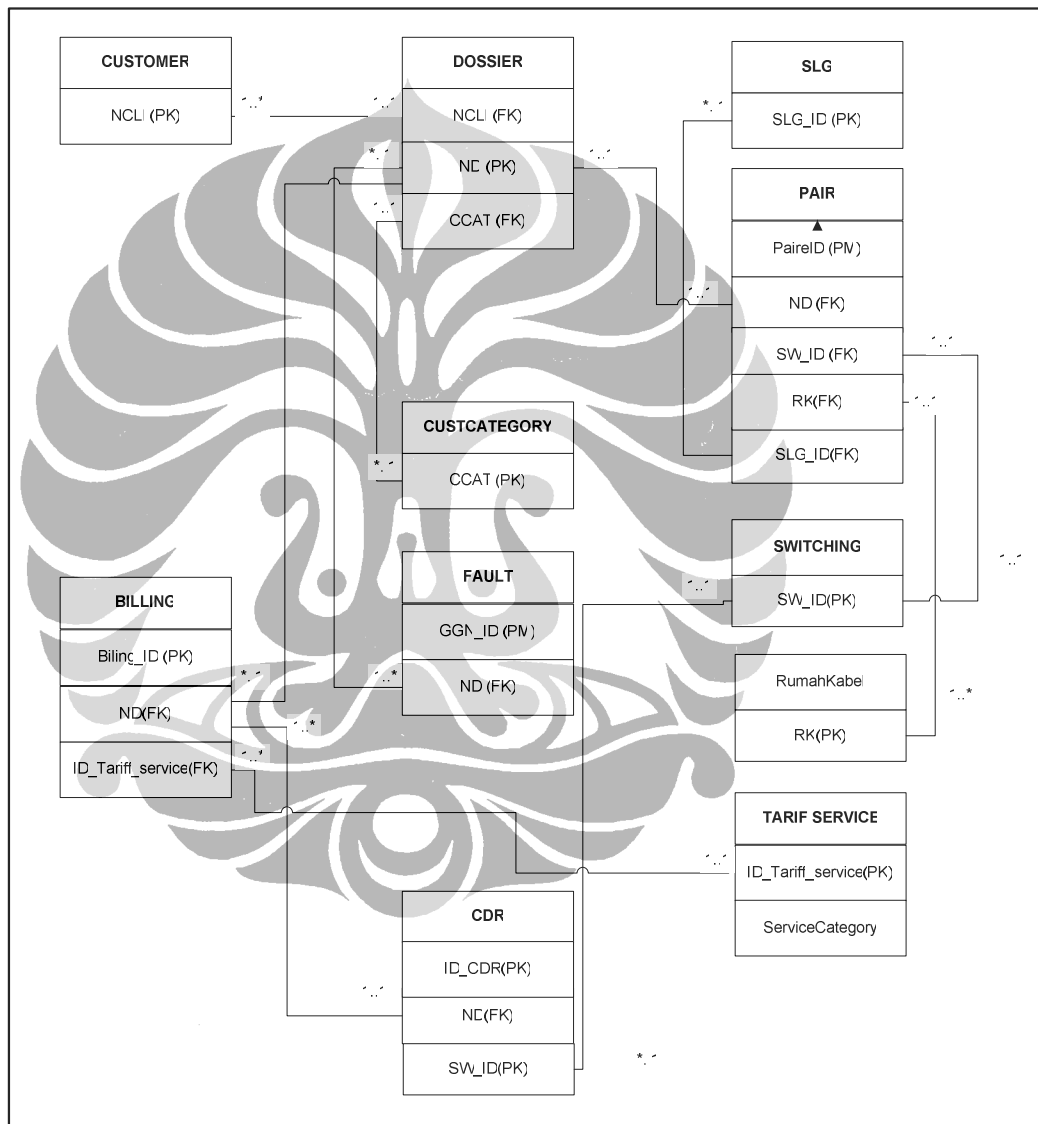
Pada gambar 4.2 ditampilkan hubungan antara *customer* dengan dossier, dimana setiap *customer* dapat memiliki satu atau banyak nomor telepon *wireline*. *Customer* dikelompokkan dalam beberapa kategori berdasarkan penggunaannya seperti residensial, bisnis atau *public phone*.

Kemudian setiap nomor telepon memiliki akses ke *switching* yang disebut *access network* berupa kabel *single pair* yang menghubungkan antara pelanggan dengan *switching*. Untuk Layanan penanganan gangguan (*fault handling*) dikelompokkan dalam beberapa segmen layanan berupa *service level guarantee* (SLG).

Setiap *customer* dapat melakukan berbagai macam panggilan yaitu berupa panggilan Lokal, SLJJ, SLI atau akses Internet. Panggilan semua itu tercatat

dalam *Call Data Recording* (CDR). Kemudian setiap akhir bulan CDR diproses di *billing processing* dengan menggunakan beberapa parameter tarif yang berlaku.

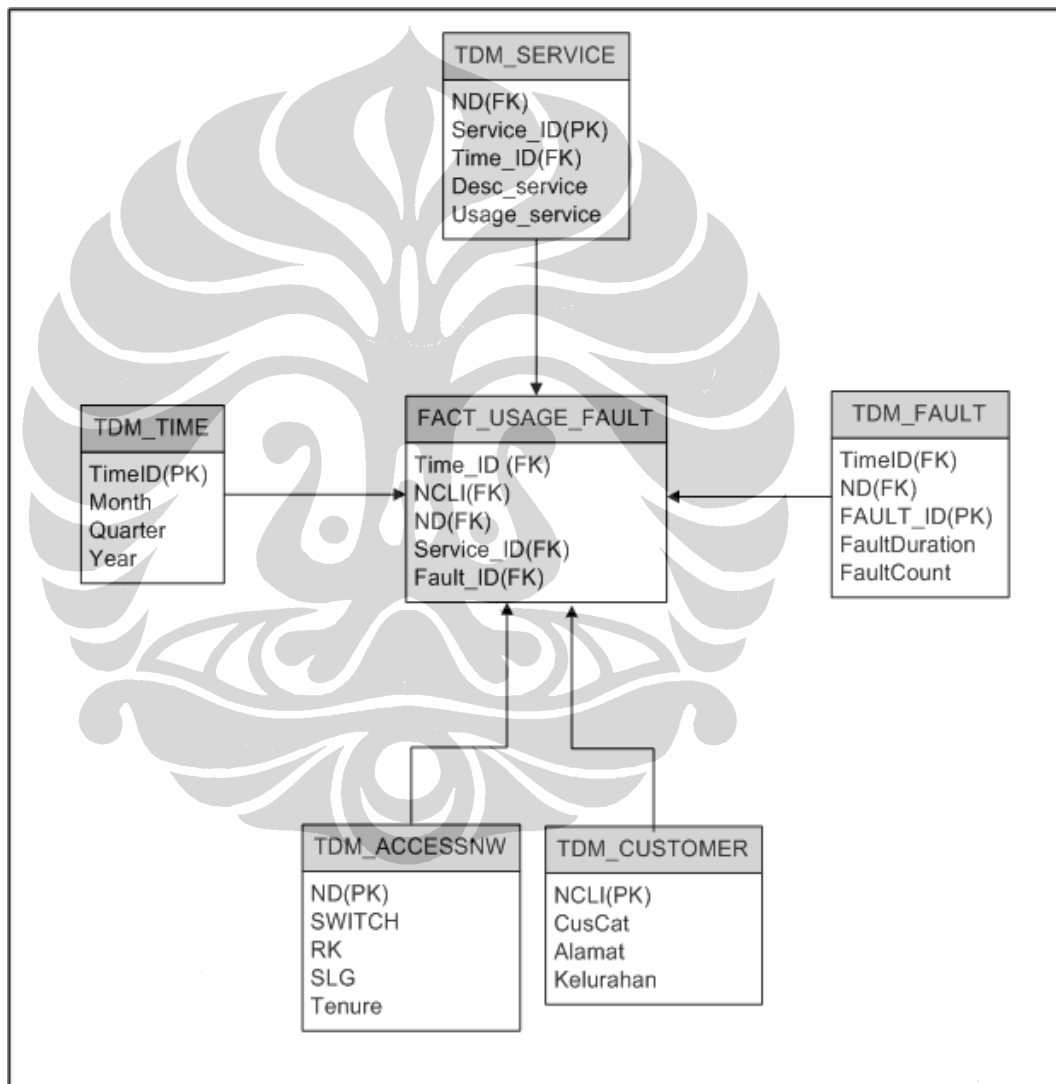
Bila pelanggan mengalami gangguan akan tercatat di *fault handling*, adapun yang tercatat kapan pelanggan lapor, kapan penyelesaiannya, lokasi gangguan dan sebagainya. Seberapa cepat penanganan/penyelesaian gangguan tercatat dalam SLG, berdasarkan SLG ini petugas menentukan skala prioritas penyelesaian gangguan.



Gambar 4.2 ERD Customer Wireline

#### 4.2.4 Perancangan Dimensi Model dan *Star Schema*

Berdasarkan *Entity Relationship Diagram* yang ada selanjutnya dilakukan proses pembuatan model *star schema* yang terdiri dari beberapa tabel dimensi dan tabel fakta. Dalam proses ini yang terpenting adalah relasi antara tabel dimensi dan tabel fakta, agar semua informasi yang penting berkaitan dengan pelanggan *wireline* dapat dipenuhi untuk kebutuhan analisa selanjutnya. Sehingga bentuk dari model *Star Schema* yang dapat dibuat seperti yang ditunjukkan di gambar 4.3



Gambar 4.3 *Star schema Usage&Fault*, tabel dimensi dan tabel fakta

#### 4.2.5 Pendefinisian *Physical Model*

Dalam tahapan ini dijelaskan deskripsi dari tiap-tiap tabel dimensi, tabel fakta dan sumber data untuk dibuat secara fisik di *database*. Di tahapan ini dilakukan juga perhitungan berapa besar data yang disimpan. Berikut ini adalah pendefinisian dari masing-masing tabel dimensi dan fakta yang dibuat sebagai *database data mining*.

No	Nama Tabel	Jumlah Kolom	Jumlah Row	Size	Sumber Data	Tools Database
1	Dim_TIME	5	730	21 byte/row	Generate System	Oracle11g
2	Dim_Customer	8	289,888	124 byte/row	Legacy Oracle8i	Oracle11g
3	Dim_AccessNW	7	411,657	65 byte/row	Legacy Oracle8i	Oracle11g
4	Dim_UsageService	8	11,719,874,790	118 byte/row	Legacy Oracle8i	Oracle11g
5	Dim_Fault	3	183,209	39 byte/row	Legacy Oracle8i	Oracle11g
6	Fact_Usage_Fault	13	15,741,781,211	140 byte/row	Legacy Oracle8i	Oracle11g

Tabel 4.2 tabel dimensi dan fakta

#### 4.2.6 Proses *Extraction, Transformation dan Loading (ETL)*

Tahapan selanjutnya adalah proses *extraction, transformation* dan *loading* (ETL). Pada tahapan ini *data legacy oracle8i* dilakukan proses ETL menggunakan *tools database Oracle11g*. Proses Pengerjaan ETL menggunakan algoritma memanfaatkan fasilitas PL/SQL yang ada di *tools database* tersebut.

##### 4.2.6.1 *Extraction*

Proses Extract adalah proses pengambilan data dari sumber data. Disebut *extract*, karena proses pengambilan data ini tidak mengambil keseluruhan data yang ada di *database* operasional, melainkan hanya mengambil data matang saja (*non volatile*). Proses *Extract* ini harus mengakomodir berbagai macam teknologi yang digunakan oleh sumber data dan diintegrasikan ke dalam *database* tunggal. Sesuai dengan arsitektur data *warehouse* yang dibuat sesuai gambar 4.1 sebelumnya bahwa sumber data berasal dari tiga *database* yaitu *database billing, database technical* dan *database commercial*. Tidak semua kolom/*field* yang ada diambil, tentunya hanya kolom/*field* tertentu yang dibutuhkan yang diambil sesuai dengan kebutuhan tabel-tabel dimensi yang digunakan. Adapun *field-field* yang diekstrak adalah sesuai dengan kebutuhan data *warehouse*. Maka bila dikaitkan dengan *field-field* sumber data yang berasal dari tabel-tabel dari *database Commercial IS, Technical IS* maupun *Billing* maka *field-field* yang

dibutuhkan untuk pembentukan tabel dimensi atau data *warehouse* adalah sebagai berikut:

DATA WAREHOUSE		DATA SOURCE		
Nama Table dimensi	Field	Database Source	Table Source	Field Source
Dim_Customer	NCLI	Commercial IS	Dossier	NCLI
Dim_Customer	ND	Commercial IS	Dossier	ND
Dim_Customer	CUSTCAT	Commercial IS	client	CCAT
Dim_Customer	GELAR	Commercial IS	client	Cprof
Dim_Customer	NAMA	Commercial IS	client	NOM
Dim_Customer	ALAMAT	Commercial IS	client	address
Dim_Customer	RT_RW	Commercial IS	client	NREGC
Dim_Customer	KELURAHAN	Commercial IS	client	LCOM_REGC
Dim_AccessNW	ND	Technical IS	paire	ND
Dim_AccessNW	NCLI	Technical IS	paire	NCLI
Dim_AccessNW	SWITCH	Technical IS	paire	MDF
Dim_AccessNW	RK	Technical IS	paire	Cabinet
Dim_AccessNW	SLG	Technical IS	paire	prio
Dim_AccessNW	KRING	Technical IS	paire	DATE_ETAT
Dim_AccessNW	Tenure	Technical IS	paire	DATE_ETAT
Dim_UsageService	ND	billing	usage	a number
Dim_UsageService	B_number	billing	usage	b number
Dim_UsageService	Time_id	billing	usage	time_id
Dim_UsageService	Start	billing	usage	Start
Dim_UsageService	Duration	billing	usage	Duration
Dim_UsageService	Tariff	billing	usage	Tariff
Dim_UsageService	UsageCat	billing	usage	UsageCat
Dim_UsageService	Amount	billing	usage	Amount
Dim_Fault	ND	Technical IS	fault	ND
Dim_Fault	Time_ID	Technical IS	fault	Time_ID
Dim_Fault	Fault_count	Technical IS	fault	Fault_count
Dim_Fault	Fault_Dura	Technical IS	fault	Fault_Dura

Tabel 4.3 *Field Data Warehouse dan Field Sumber data*

Karena arsitektur *data warehouse* yang dibuat menggunakan *staging area*, maka dibuat *tabel* atau *view* yang berfungsi sebagai *hold store* untuk mengumpulkan *field-field* hasil *extract* tersebut.

#### 4.2.6.2 Transformation

Proses *transformation* dibutuhkan bilamana *data legacy* belum memenuhi kebutuhan struktur data untuk *data mining*. Berikut ini terdapat beberapa proses *transformation* yang dilakukan.

1. Pembentukan *Field*/kolom *usage\_id*

*Transformation* pembentukan *usage\_id* dilakukan dengan menggunakan teknik menggabungkan beberapa *field/kolom* yang ada menjadi satu *field/kolom*. Prosesnya adalah sebagai berikut:

Field/Kolom	ND	UsageCat	Bln_tag
Data	0214353XXXX	LOKAL	2007.01

Menjadi

Field/Kolom	UsageID
Data	0214353XXXXLOKAL2007.01

Gambar 4.4 Proses transformasi pada *field/kolom usage\_id*

Adapun *script query* pembentukan *field/ kolom usageid* sebagai berikut:

```
SELECT ND, Usagecat, bln_tag,
       ND || " " || Usagecat || " " || Bln_TAG as Usageid
FROM usage;
```

## 2. Pembentukan *Field/kolom Tenure*

*Transformation* pembentukan *field/kolom tenure* dilakukan dengan menghitung lamanya suatu *customer* berlangganan telepon (dalam tahun). Mulai dihitung sejak tanggal *kring* sampai dengan saat ini. Prosesnya adalah

Field/Kolom	Kring	sysdate
Data	19-May-01	20-May-09

Menjadi

Field/Kolom	Kring	sysdate	Tenure
Data	19-May-01	20-May-09	8.00822

sebagai berikut:

Gambar 4.5 Proses transformasi pada *field/kolom Tenure*

Adapun *script query* pembentukan *field/ kolom Tenure* adalah sebagai berikut:

```
SELECT kring, sysdate,  
       (sysdate-to_date(kring))/365 as Tenure  
FROM accessnw;
```

Di dalam proses ini juga dilakukan proses pengecekan data untuk menjaga kualitas dan konsistensi data yang dihasilkan. Selain itu dilakukan juga proses *data cleansing* dengan menggunakan *query-query* yang dijalankan secara manual untuk menjaga kualitas data yang sudah terbentuk. Misalnya penulisan format tanggal yang berbeda harus disamakan.

#### 4.2.6.3 Loading

Setelah dilakukan proses *transformation* di *staging area* maka dilakukan proses *loading* ke database data warehouse. Adapun *mode loading* yang digunakan untuk proses ini adalah *update mode* melakukan *insert* terhadap target data yang berubah secara periodik. Jadi data yang tidak mengalami perubahan tidak dilakukan *loading*. Sedangkan proses *update* ini sendiri dilakukan pada saat beban kerja *database* operasional pada kondisi rendah, biasanya dibuat *schedule* di malam hari atau di saat di luar jam kerja.

Setelah proses *loading* selesai dilakukan proses *integrity data*, yaitu dengan dengan melakukan pengujian *star schema* dan konsistensi jumlah data, proses ini menggunakan *script query* yang dijalankan secara manual. Data yang sudah masuk dalam tabel-tabel dimensi dan fakta digunakan sebagai sumber data untuk proses *Online Analytical Processing* (OLAP) dan *data mining*.

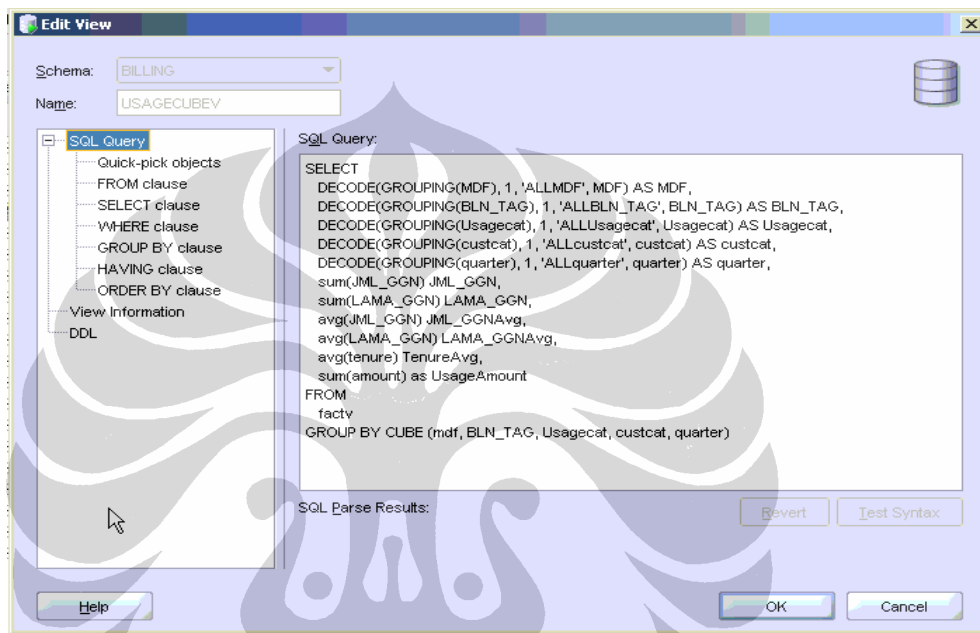
### 4.3 Penyelidikan Data

Di dalam tahapan penyelidikan data, dilakukan analisa terhadap tabulasi-tabulasi dan grafik-grafik yang dihasilkan dari tabel-tabel dimensi dan fakta yang ada. Dengan teknologi OLAP, tabel/view *cube* fakta yang telah dibuat, dapat dilakukan proses *slicing*, *drilldown* dan *rollup* dengan memanfaatkan *query* yang ada dan aplikasi yang dibuat untuk memudahkan *end user* melakukan penyelidikan data.



### 4.3.1 Pembuatan Tabel/view Cube OLAP

Dengan memanfaatkan fasilitas OLAP dari *Oracle 11g* serta tabel dimensi dan fakta (*star schema*) yang sudah ada, maka dapat dimulai membuat *Query Grouping SQL* berupa *view cube* seperti gambar 4.6 Dengan menggunakan fasilitas tersebut diharapkan data yang terbentuk dari *star schema* yang berjumlah puluhan juta baris dapat diakses/*diquery* dengan cepat berdasarkan kriteria-kriteria yang diinginkan.



Gambar 4.6 Pembuatan SQL Cube

Kemudian bila sql *cube* tersebut dijalankan menghasilkan tabel/view seperti gambar 4.7 sebagai berikut:

MDF	BLN_TAG	USAGECAT	CUSTCAT	QUARTER	JML GGN	LAMA GGN	JML GGN/AVG	LAMA GGN/AVG	TENURE/AVG	
1	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	ALLquarter	51579035	725296313.5	4.3984345952...	61.85009853598...	10.3345073...
2	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2006.Q3	6151225	86352305.65	4.3643816760...	61.26818974539...	10.7124547...
3	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2006.Q4	6282305	86296592.45	4.3946954222...	61.76660169147...	10.5791699...
4	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2007.Q1	6366845	89500549.75	4.4146142752...	62.05746758857...	10.5349716...
5	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2007.Q2	6471795	91176148.15	4.4224979157...	62.30517579166...	10.3770642...
6	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2007.Q3	6534760	92082618.15	4.4151844682...	62.21525280984...	10.2530850...
7	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2007.Q4	6566130	92414235.55	4.4051134800...	61.99925904184...	10.1836023...
8	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2008.Q1	6596275	92774288.6	4.3956265472...	61.82294186832...	10.1095173...
9	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2008.Q2	2204250	30967167.7	4.3848219614...	61.60188629401...	10.0507789...
10	ALLMDF	ALLBLN_TAG	ALLUsagecat	ALLcustcat	2008.Q4	4405450	61732407.5	4.3697708212...	61.23244458992...	9.88675072...
11	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	ALLquarter	68930	477959.9	5.7513558614...	39.87984146850...	18.2147171...
12	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2006.Q3	8565	59502.9	5.8265306122...	40.47816326530...	18.5625030...
13	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2006.Q4	8565	59502.9	5.8265306122...	40.47816326530...	18.4288849...
14	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2007.Q1	8525	59403.7	5.8390410958...	40.68746575342...	18.6197348...
15	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2007.Q2	8525	59403.7	5.8390410958...	40.68746575342...	18.7347508...
16	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2007.Q3	8585	59628.6	5.8006756756...	40.28989459459...	18.3863802...
17	ALLMDF	ALLBLN_TAG	ALLUsagecat	100	2007.Q4	8715	60161.7	5.6407766990...	38.93961165048...	18.0439228...

Gambar 4.7 Hasil Running SQL Cube

#### 4.3.2 Pembuatan Query OLAP

Kemudian setelah tabel/view cube dibuat, maka dilakukan query-query sesuai tabulasi data yang diinginkan. query-query tersebut dibuat dengan sederhana, karena menggunakan fasilitas OLAP yang ada, dapat menghasilkan data sesuai kebutuhan analisa atau penyelidikan. Berikut ini query-query yang dijalankan untuk mengambil data dari view cube fakta yang telah dibuat.

1. Misalnya kita ingin melihat jumlah usage masing-masing category untuk semua switching, semua customer category selama 12 bulan tahun 2007. Maka dibuat query sederhana yang dapat melakukan drilldown terhadap data tahun 2007 dan usage category, sedangkan measure yang ditampilkan adalah usage amount. Adapun query dan hasilnya adalah sebagai berikut :

Enter SQL Statement:

```

select
  bln_tag, usagecat, usageamount
from FACTCUBEV
where
  year='2007' AND SWITCH='ALLSWITCH' AND
  custCat='ALLCUSTCAT' AND quarter='ALLQUARTER'

```

Results:

	BLN_TAG	USAGECAT	USAGEAMOUNT
7	2007.01	ALLUSAGE	112266638078
8	2007.01	SLJJ	28970494347
9	2007.01	TENI	10808550160
10	2007.01	LOKAL	63028670412
11	2007.01	SLI001	452122888
12	2007.01	TIC007	9006800271
13	2007.02	ALLUSAGE	117789043454
14	2007.02	SLJJ	31126417684
15	2007.02	TENI	10459218263
16	2007.02	LOKAL	66319907402
17	2007.02	SLI001	625548235

Gambar 4.8 hasil *query* bulan tagihan, *usage category* dan *amount*

- Melihat jumlah *usage* dan gangguan untuk di lokasi masing-masing *switching* pada tahun 2007. Maka dibuat *query* yang melakukan *rollup* data tahun 2007 dan *drilldown* terhadap data *total switching*, sedangkan *measure* yang ditampilkan adalah *usage amount* dan Jumlah gangguan. Adapun *query* dan hasilnya adalah sebagai berikut:

Enter SQL Statement:

```

select
  switch, usageamount, jml_ggn
from FACTCUBEV
where
  year='2007' AND usagecat='ALLUSAGE' AND
  custCat='ALLCUSTCAT' AND quarter='ALLQUARTER' AND bln_tag='ALLBLN_TAG'

```

Results:

	SWTCH	USAGEAMOUNT	JML_GGN
1	ALLSWTCH	2123915029533	1157439
2	CID	294303031106	112497
3	CIL	105971181896	86091
4	KMY	53165876531	8450
5	KTA	158627086121	4075
6	MBS	493759244521	95584
7	MKR	388343214294	344475
8	MRD	16143392725	10530

Gambar 4.9 *Drilldown switching*, *usage amount* dan jumlah gangguan

- Melihat lama gangguan dan pemakaian lokal untuk masing-masing customer *category* untuk lokasi *switching* di area5. Maka dibuat *query* yang melakukan *slicing* terhadap data *usage* lokal dan *switching* area5, serta *drilldown* terhadap data *customer category*. Maka hasil *query* adalah sebagai berikut:

Enter SQL Statement:

```

select
  Custcat, usageamount, lama_ggn
from FACTCUBEV
where
  year='2007' AND usagecat='LOKAL' AND SWITCH='MBS' AND
  quarter='ALLQUARTER' AND bln_tag='ALLBLN_TAG'

```

Results:

	CUSTCAT	USAGEAMOUNT	LAMA_GGN
1	ALLCUSTCAT	196654021271	441254.71
2	100	76479000	(null)
3	102	266548500	238.28
4	113	250	(null)
5	117	134816000	210.6
6	118	21869174000	245374.82
7	300	167643856500	192202.52
8	400	333032750	37.96
9	401	19274750	(null)
10	402	3364000	(null)
11	403	18725250	(null)
12	405	298750	(null)
13	411	24041000	(null)
14	480	1946000	(null)
15	601	4467481200	(null)

Gambar 4.10 Slicing *switching*, *Drilldown Customer Category*.

### 4.3.3 Pembuatan Aplikasi OLAP

Untuk memudahkan *end user* dalam memanfaatkan data OLAP maka dibuat aplikasi yang gampang digunakan dalam menjalankan *query-query* yang telah dibuat sebelumnya. Agar *user* dengan mudah menjalankan aplikasi ini, maka aplikasi dibuat berbasis web. Pengembangan web ini menggunakan bahasa pemrograman *html* dan *php*, karena lebih mudah diintegrasikan dengan *database oracle* dan lebih mudah pengerjaannya. Adapun fungsi-fungsi utama yang perlu ditampilkan adalah sebagai berikut:

#### A. *Drilldown dan rollup*

Aplikasi ini memiliki kemampuan *drilldown* dan *rollup* terhadap data area *switching* dan data *usage service*. Seperti contoh pada gambar 4.11 status area *switching* dan *usage service* di *drilldown* ke *level* yang lebih rendah. Kemudian

contoh *rollup* seperti yang ditunjukkan dalam gambar 4.12, masing-masing *area switching* di *rollup* ke atas. Sehingga yang tampak adalah jumlah total / *all* dari seluruh *area switching*.

The screenshot shows a web browser window titled 'http://10.44.12.102/olap/olap.htm - Microsoft Internet Explorer'. The page content is titled 'APLIKASI OLAP'. It features a navigation area with 'MONTH' (JAN, FEB, MAR), 'YEAR' (2005, 2006, 2007, 2008), and 'MEASURE' (Usage(%)). Below this is a table with the following data:

SWITCHING	Usage Service	- ALL	- LOKAL	- PONSEL	- SLJJ	- SLI007	- SLI001	- INET
- ALL		100.00%	33.18%	40.40%	15.51%	4.88%	0.27%	5.70%
- Area1		13.58%	4.64%	5.55%	1.97%	0.43%	0.04%	0.94%
- Area2		4.97%	1.44%	2.31%	0.62%	0.40%	0.00%	0.18%
- Area3		2.52%	0.65%	1.22%	0.34%	0.18%	0.00%	0.12%
- Area4		7.26%	2.23%	3.05%	1.31%	0.35%	0.03%	0.27%
- Area5		22.82%	9.09%	8.20%	3.37%	0.85%	0.06%	1.25%
- Area6		19.26%	6.34%	7.27%	3.30%	0.91%	0.07%	1.36%
- Area7		0.75%	0.22%	0.36%	0.08%	0.05%	0.00%	0.04%
- Area8		6.33%	2.03%	2.59%	0.89%	0.27%	0.02%	0.47%
- Area9		13.40%	3.51%	5.81%	2.29%	1.04%	0.04%	0.71%
- Area10		9.11%	2.97%	4.04%	1.34%	0.40%	0.01%	0.36%

At the bottom of the table, there are 'Print' and 'Export' buttons. The interface also includes a 'Cust Category' filter with options: ALL, 100, 102, 104, 118.

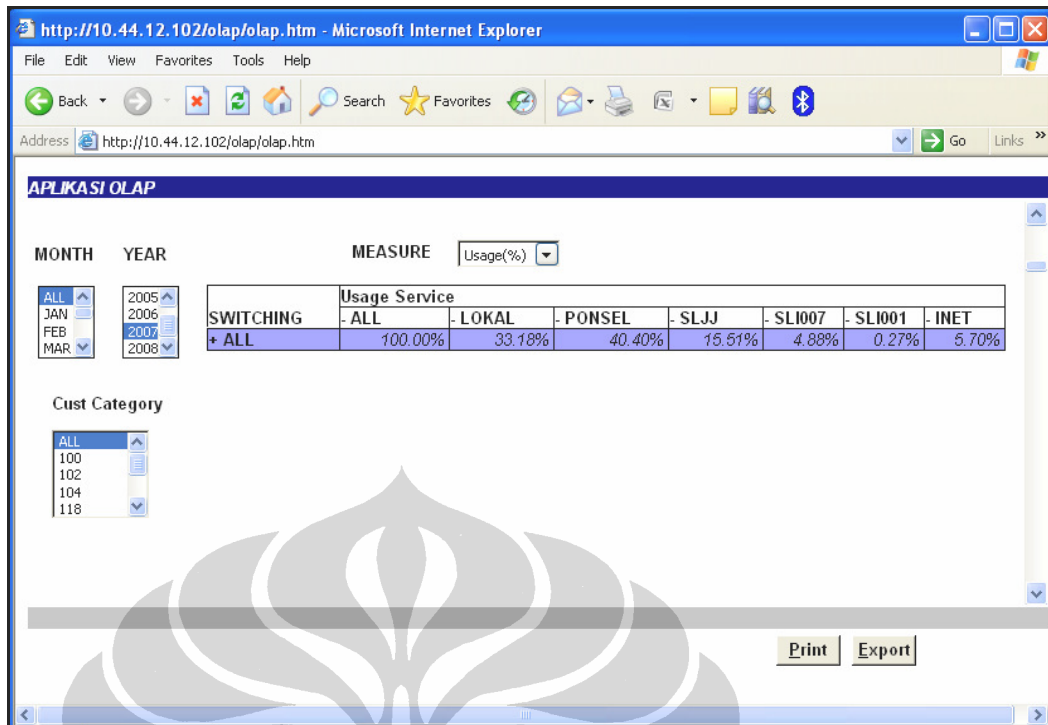
Gambar 4.11 Aplikasi OLAP *drilldown*

#### B. Pilihan *Combo Box*

Sedangkan untuk *customer category*, *month-year*, *measure* dibuat pilihan dalam bentuk *combo box*. Data *field* tersebut dibuat *combo box* agar memudahkan pencarian informasi sesuai dengan kriteria yang dikehendaki. Kalau dibuat dalam bentuk *drilldown/rollup* untuk *customer category* akan panjang atau lebar karena jenis *customer category* sebanyak tiga puluh jenis.

#### C. *Print* dan *Export*.

Kemampuan lain yang dibuat untuk aplikasi ini adalah dapat melakukan pencetakan hasil *query* ke media printer yang tersedia. Kemudian ditampilkan kemampuan *export* data dalam format *excel* maupun teks. Hal ini perlu dimunculkan, karena adanya kebutuhan untuk pengolahan data lebih lanjut misalnya digunakan untuk pembuatan grafik.



Gambar 4.12 Aplikasi OLAP rollup

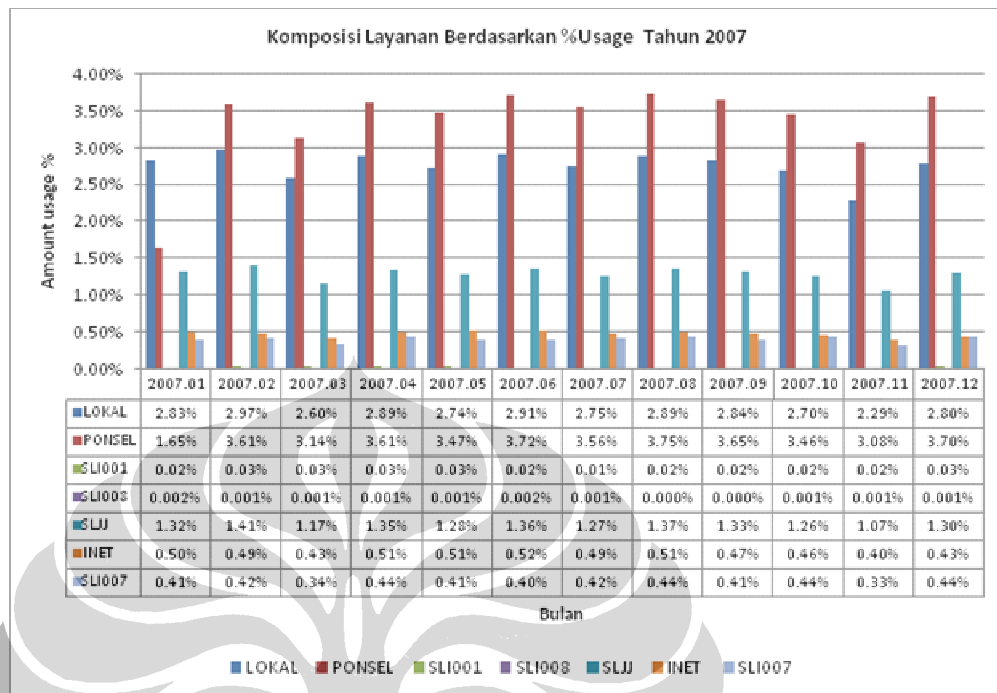
#### 4.3.4 Data Hasil OLAP

Untuk Penyelidikan lebih lanjut dapat dilakukan dengan membuat grafik dari hasil-hasil *query* sebelumnya. Dengan menggunakan grafik diharapkan dapat membantu menggali informasi-informasi jika menggunakan tabulasi belum muncul. Kemudian untuk membuat grafik memanfaatkan kemampuan Excel2007 yang mampu membuat grafik secara *customize* dan mengakses *query/view* yang telah dibuat sebelumnya tanpa harus melakukan proses export/import. Berikut ini adalah grafik hasil proses OLAP untuk menjelaskan hasil tabulasi yang sudah dibuat.

##### A. Komposisi *User* dan *Usage* dari tiap Kategori *Service* di semua Area *Switching*

Angka *usage* ditampilkan dalam bentuk prosentasi karena data ini di PT.XYZ termasuk data *confidential*. Angka prosentase ini berasal dari jumlah tagihan/*billing* dari tiap pelanggan dibagi dengan jumlah total *billing* pelanggan selama tahun 2007. Sehingga secara individual angka prosentasi ini identik dengan angka *usage*. Walaupun dalam bentuk prosentase masih dapat digunakan untuk

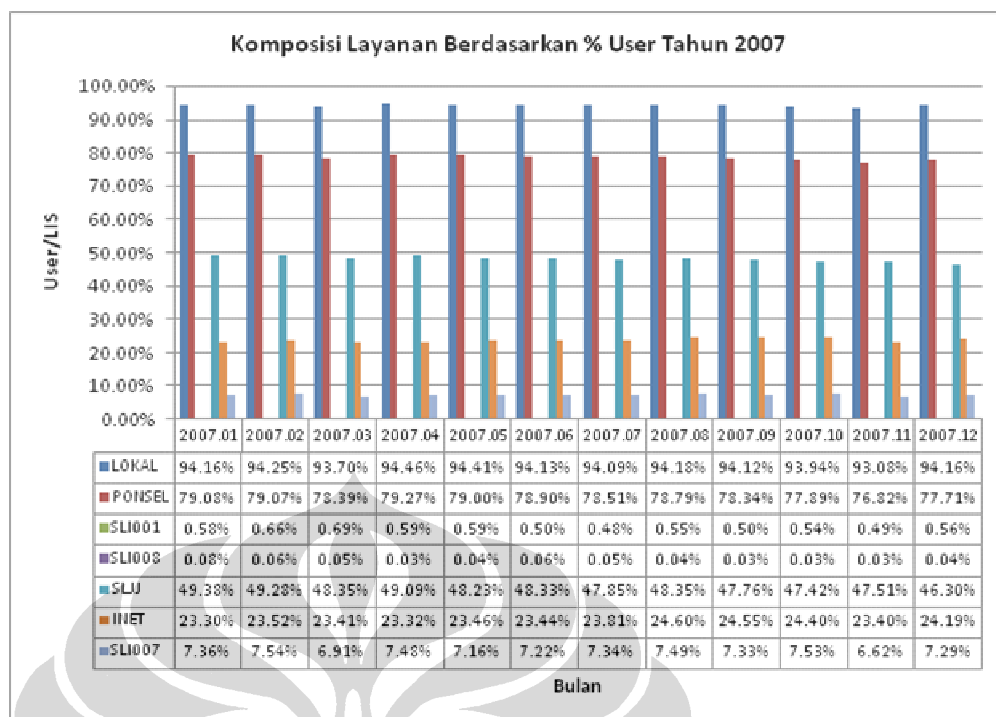
membandingkan besaran masing-masing data *usage*. Adapun data yang ditampilkan adalah periode tahun 2007.



Grafik 4.1 Komposisi Layanan Berdasarkan Prosentase *Usage* Tahun 2007

Sesuai grafik 4.1 ditemukan fakta bahwa komposisi *usage* layanan mulai yang terbesar sampai dengan yang terkecil adalah : PONSEL (1.65%-3.75%), Lokal (2.29%-2.91%), SLJJ (1.07%-1.41%), INET (0.40%-0.52%), SLI007(0.33%-0.44%), SLI001(0.02%-0.03%) dan SLI008(0.001%-0.002%).

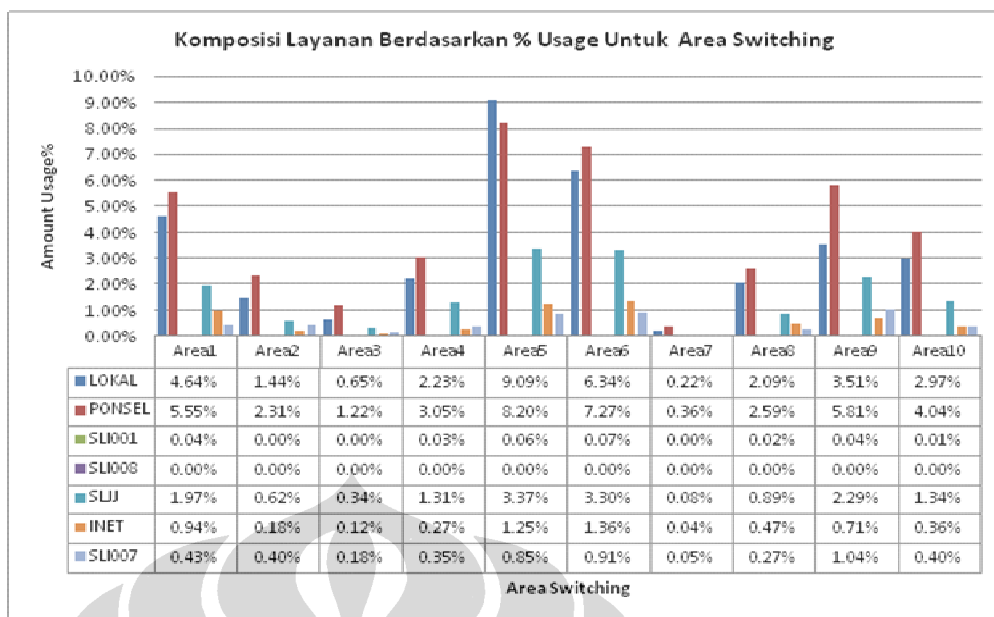
Kemudian kategori layanan dipetakan berdasarkan jumlah *user* (dalam bentuk prosentase) sesuai grafik 4.2. Angka prosentasi berasal dari jumlah *user* dibagi *Line In Service* (LIS). Angka LIS adalah jumlah nomor pelanggan yang masih aktif berlangganan. Adapun komposisi prosentase *user* mulai yang terbesar sampai dengan yang terkecil hampir sama dengan komposisi *usage*, namun yang berbeda adalah besaran prosentasinya saja. Adapun komposisinya adalah sebagai berikut : Lokal (93.08%-94.41%), Ponsel (76.82%-79.27%), SLJJ (46.30%-49.38%), INET (23.30%-24.60%), SLI007(6.62%-7.54%), SLI001(0.49%-0.69%) dan SLI008(0.03%-0.08%). Dari uraian tersebut Layanan Lokal, Ponsel, SLJJ, INET kontribusinya masih dominan dibandingkan layanan lainnya seperti SLI007, SLI001 dan SLI008.



Grafik 4.2 Komposisi Layanan Berdasarkan Prosentase *User* tahun 2007

Kemudian kalau dipetakan berdasarkan area *switching* untuk melihat kontribusi masing area masing-masing *switching* masih untuk tahun yang sama. Maka grafik yang dapat ditampilkan seperti grafik 4.3. Di dalam grafik tersebut tampak bahwa kontribusi area *switching* lima besar mulai dari yang terbesar sampai dengan yang terkecil untuk layanan yang dominan (Lokal,Ponsel,SLJJ) adalah : Area5 (lokal 9.09%, Ponsel 8.20%, SLJJ 3.37%), Area6 (lokal 6.34%, Ponsel 7.27%, SLJJ 3.30%), Area9 (lokal 2.09% , Ponsel 5.81%, SLJJ 2.29%), Area1 (lokal 4.64% , Ponsel 5.55%, SLJJ 1.97%) dan Area10 (lokal 2.97% , Ponsel 4.04%, SLJJ 1.34%). Dari data tersebut Area *switching* yang dominan adalah area5 dan komposisi layanan yang dominan adalah *usage* PONSEL, lokal dan SLJJ.



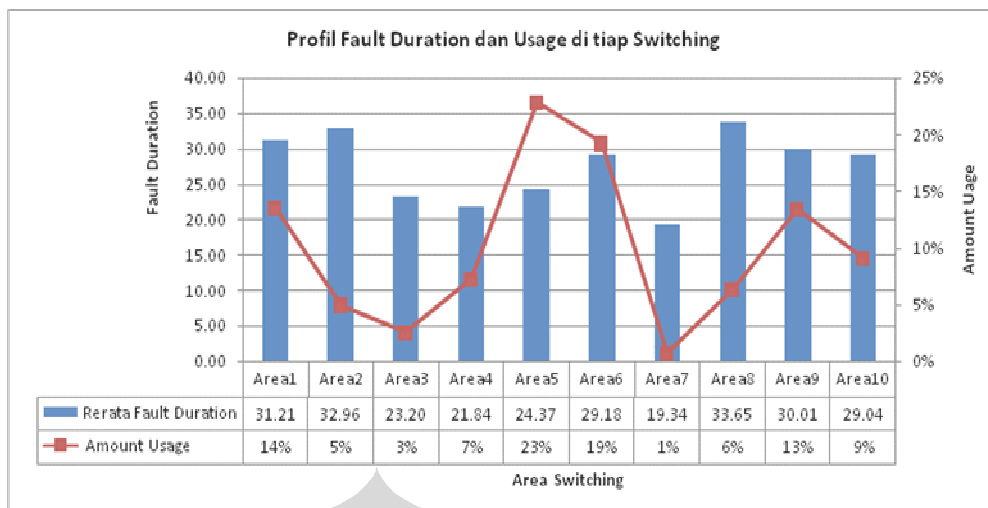


Grafik 4.3 Komposisi layanan berdasarkan prosentase *usage* untuk area *switching*

#### B. Profil *Fault duration* dan *Usage* semua *Switching*

Selanjutnya dipetakan juga profil *fault duration* dan *Usage* di tiap-tiap *switching* seperti Grafik 4.4. *Fault duration* adalah lama gangguan yang berlangsung dari suatu nomor telepon dengan kata lain pada saat tersebut telepon tidak dapat digunakan mengakses layanan yang ada. *Fault duration* yang baik adalah semakin kecil angkanya semakin baik. Angka *Fault duration* untuk tiga terbaik dari terkecil sampai yang terbesar secara berurutan adalah Area7 (19.34 jam), Area4 (21.84 jam) dan area3 (23.20 jam). Sedangkan untuk *amount usage* untuk tiga terbaik dari yang terbesar sampai yang terkecil secara berurutan adalah Area5 (23%), area6 (19%) dan area1 (14%).

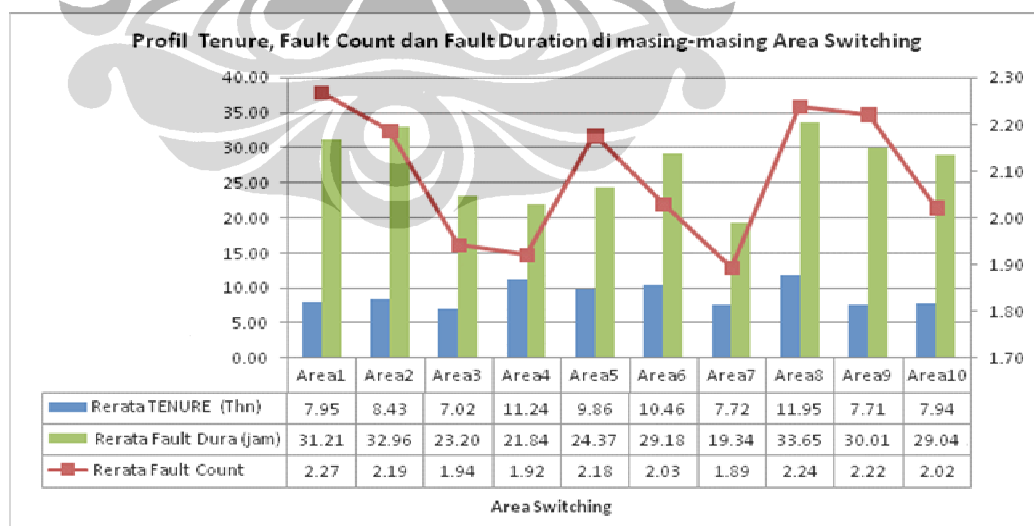
Dari fakta tersebut secara agregat antara data *fault duration* dan *usage* tidak tampak adanya korelasi. Karena masih ada faktor lain yang menentukan besaran *usage* yaitu jumlah *user* yang berada di suatu area *Switching*. Sehingga dengan fakta-fakta tersebut belum dapat diambil kesimpulan secara utuh.



Grafik 4.4 Profil *Fault Duration* dan *Usage* di tiap Area *Switching*

### C. Profil *Tenure* dan *Fault Count* Pelanggan di masing-masing *Switching*

*Tenure* merupakan angka seberapa lama pelanggan sudah berlangganan, angka *Tenure* dalam grafik 4.5 merupakan angka rata-rata untuk tiap *switching* (satuan tahun). Sedangkan *Fault count* merupakan seberapa banyak atau berapa kali pelanggan mengalami gangguan. Profil masing-masing *switching* seperti yang digambarkan pada grafik 4.5. Untuk angka rata-rata *tenure* tiga terlama adalah Area8 (11.95 tahun), area4 (11.24 tahun) dan area6 (10.46 tahun). Sedangkan yang paling baru adalah area3 (7.02 tahun).



Grafik 4.5 Profil *Tenure*, *Fault Count* dan *Fault Duration*

Kemudian *Fault Count* adalah jumlah rata-rata gangguan yang terjadi selama periode satu tahun. Semakin kecil angkanya akan makin baik. Adapun

profil masing-masing *switching* tiga terbaik adalah area7 (1.89 kali), area4 (1.92 kali) dan Area3 (1.94 kali). Sedangkan yang terbanyak adalah area1 (2.27 kali).

Fakta-fakta tersebut di atas menggambarkan secara agregat antara *tenure* dan *fault count*. Kalau diperhatikan menggunakan grafik yang ada, tidak ditemukan adanya korelasi yang cukup signifikan. Kemungkinan masih ada factor lainnya yang mempengaruhi besaran *fault count* selain *Tenure*.

#### **4.4 Persiapan Data untuk Pemodelan *Data Mining***

Sesuai dengan tujuan penulisan tesis ini adalah mencari *asosiation rules* dari data *billing* dan *fault*, maka data maupun fakta-fakta yang muncul di OLAP yang sudah dibuat disiapkan sebagai input dalam proses *Data Mining*. Tentunya data yang diambil relevan dengan permasalahan bisnis yang ada. Agar proses *data mining* lebih fokus, terarah dan tepat sasaran.

##### **4.4.1. Pemilihan Data yang Relevan sebagai Pemodelan *Data Mining***

Permasalahan bisnis PT.XYZ saat ini yang ada adalah Layanan *wireline* masih menjadi tulang punggung kegiatan bisnis utama. Tentunya layanan/produk yang disediakan dan kualitas jaringan akan menjadi perhatian utama dalam upaya meningkatkan pendapatan dan memenangkan persaingan. Oleh karena itu data yang relevan faktor-faktor tersebut akan dipergunakan untuk proses *data mining*.

Pemilihan *data mining* juga mengacu pada data OLAP yang sudah dibuat. Dalam proses OLAP juga ditemukan fakta-fakta yang relevan dengan permasalahan bisnis yang ada. Adapun data dan fakta-fakta yang relevan adalah sebagai berikut:

- Berkaitan dengan Layanan *wireline* seperti layanan panggilan Lokal, PONSEL, SLJJ , SLI dan layanan akses internet. Adapun alasannya adalah semua layanan tersebut masih mendatangkan pendapatan bagi PT.XYZ, hal ini terlihat dalam proses OLAP masih terlihat pemakaiannya (*usage*).
- Berkaitan dengan kinerja jaringan seperti jumlah gangguan (*Fault Count*) dan lama gangguan (*Fault Duration*) di suatu area *switching*. Karena

secara individual kinerja yang jaringan yang rendah akan mengurangi opportunity pemakaian telepon.

- Berkaitan dengan Data pelanggan seperti *Customer Category* dan *Tenure*

#### 4.4.2 Pemilihan *Field*/kolom *Data Training* untuk *Data Mining*.

Pemilihan *field* / kolom *data training* untuk *data mining* berdasarkan data OLAP yang sudah dibuat. Serta informasi lainnya untuk mengakomodasi analisa selanjutnya agar relevan dengan permasalahan bisnis yang ada. *Field* / kolom yang telah dipilih, nantinya dipergunakan sebagai *variabel* dalam proses *data mining*. Adapun *field*/kolom yang dijadikan sebagai *data training* adalah sebagai berikut:

NO	Nama <i>Field</i> /kolom	Deskripsi
1	ND	Kode Nomor Telpon
2	Bln_Tag	Kode Bulan tagihan
3	<i>UsageCat</i>	Kode <i>Usage Service Category</i>
4	<i>Custcat</i>	Kode <i>Customer Category</i>
5	<i>Switch</i>	Kode area switch
6	<i>Amount</i>	Besarnya pemakaian/ <i>Usage</i>
7	<i>Fault_count</i>	Banyaknya gangguan
8	<i>Fault_duration</i>	Lama waktu terjadi gangguan
9	<i>Tenure</i>	Lama Berlangganan

Tabel 4.4 *Field*/kolom *Data Training*

Kemudian data *field* tersebut dibuat dalam sebuah format *data training* agar dapat diproses menggunakan algoritma *association rules*. Khusus untuk data *field* yang berkarakter data *range* seperti *amount usage category*, *fault count*, dan *fault duration* dilakukan proses *segmentasi* dan pengkodean yang dijelaskan pada bagian 4.5.1 dan 4.5.2. Adapun format *data training* / data set yang harus dibuat terdapat di lampiran 2.

#### 4.5. Pembuatan Model Data Mining

Dalam proses pengerjaan pemodelan *data mining* ini, menggunakan *tools data mining clementine* versi 12. Adapun pertimbangan menggunakan *tools* ini adalah memiliki kemampuan mengakses sumber data secara langsung dari *database oracle* tanpa harus melalui proses ekspor-impor. Selain itu memiliki kemampuan untuk memroses puluhan juta *row data*. serta fitur yang disediakan terdiri dari berbagai macam model analisis seperti *classification*, *segmentation* dan *association*.

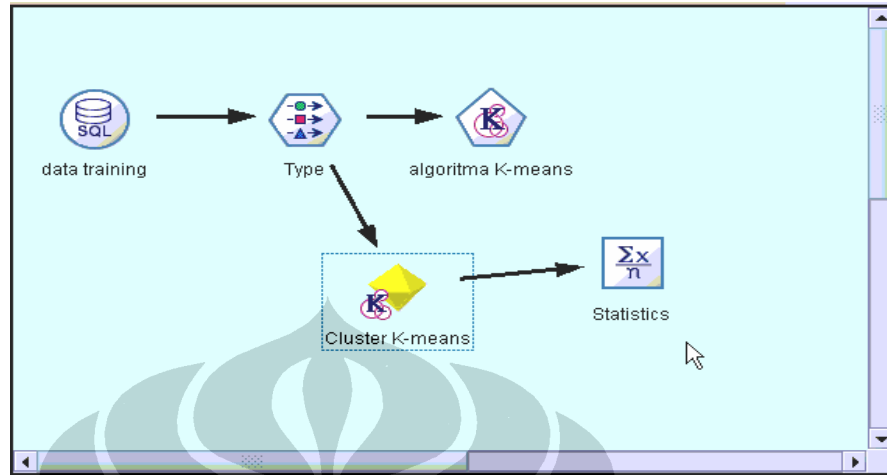
Kemudian dalam proses pengerjaan *data mining* penulisan ini menggunakan proses *segmentation* yaitu Algoritma *segmentasi* K-Means dan proses *Association Rules* yaitu menggunakan Algoritma *Apriori*. Adapun pertimbangan menggunakan kedua proses *data mining* tersebut adalah proses penyiapan datanya mudah, sederhana dan parameter tambahan untuk pemodelan tidak terlalu banyak, serta diakui di lingkungan riset sebagai metode yang cukup handal.

##### 4.5.1 Pengelompokan (Segmentasi)

Dari data yang disajikan *OLAP* masih belum cukup dijadikan sebagai variabel untuk proses pencarian *Association Rules*, masih dibutuhkan proses *segmentasi* untuk data yang memiliki rentang atau *range*. Dalam penulisan ini data yang berkarakter seperti itu adalah data *service usage* (Lokal, *PONSEL*, *SLJJ*, *SLI*, Akses *Internet*), *tenure*, *fault duration* dan *fault count*. Hasil proses pengelompokan/*segmentasi* ini nantinya dipergunakan sebagai *variabel* untuk proses *Association Rules*.

Seperti penjelasan sebelumnya bahwa proses *segmentasi* menggunakan metode/algoritma *segmentasi K-Means*. Adapun data yang dibuat *segmentasinya* secara detil adalah *Usage Lokal*, *SLJJ*, *SLI001*, *SLI007*, *SLI008*, *INET*, *fault duration* dan *fault count*. Sebagai contoh berikut ini adalah proses pemodelan *segmentasi* K-means untuk data *usage* lokal seperti yang ditunjukkan dalam gambar 4.13 Dalam menentukan jumlah *segmen* penulis membandingkan *error* dari *misclassification* yang muncul dari hasil running model dengan jumlah *segmen* sebanyak 4 (empat), 5 (lima) dan 6 (enam) kemudian dipilih *error* yang

terkecil. Selain itu prinsip penting dari *segmentasi* adalah memaksimalkan kesamaan yang ada antara anggota di dalam *segmen* dan meminimalkan kesamaan antar *segmen* (Two Cross Corporation, 2006).



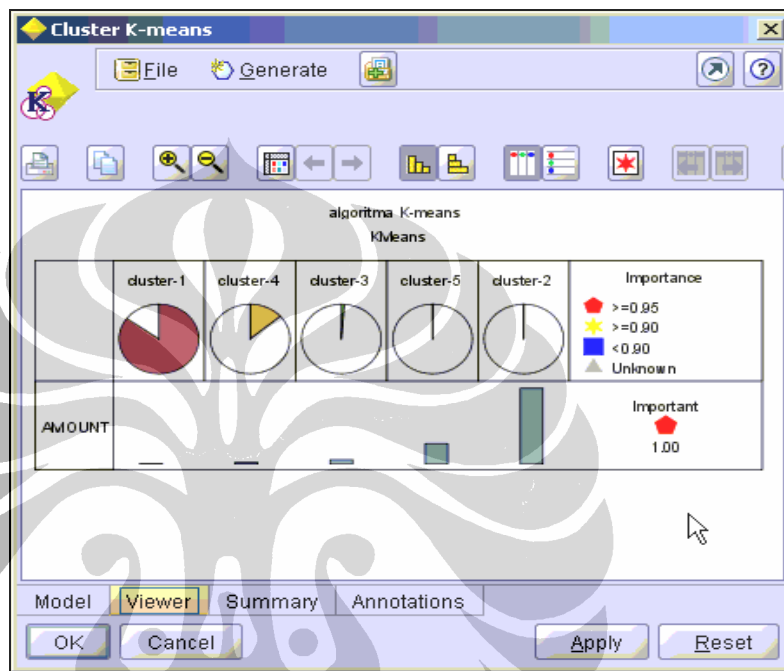
Gambar 4.13 Pemodelan *Segmentasi K-means Usage Lokal*

Setelah model *segmentasi* tersebut dijalankan menghasilkan output berupa *segmen-segmen* seperti yang ditunjukkan dalam gambar 4.14. Dari model yang dihasilkan tampak jumlah *segmen* yang dihasilkan lima *segmen*, jumlah *record* dari masing-masing *segmen*, Nilai rata-rata (*mean*) dan standard deviasi dari tiap *segmen*.



Gambar 4.14 Hasil *Segmen K-means* dari *Usage Lokal*

*Output model* juga dapat ditampilkan dengan gambar *pie diagram* dan *histogram* seperti yang ditunjukkan dalam gambar 4.15 *Pie diagram* menunjukkan jumlah *record*/elemen dari masing-masing *segmen*. *Histogram* menunjukkan nilai *mean* dalam hal ini adalah *amount usage* dari tiap *segmen*. Selain itu muncul tingkat *importance* dari variabel –variabel yang dijadikan input dalam proses *segmentasi*.



Gambar 4.15 *Pie Diagram Segmentasi Usage Lokal.*

Proses pembentukan *segmen* juga dilakukan untuk data *service usage* SLJJ, SLI001, SLI007, SLI008, INET, *fault duration* dan *fault count*. Kemudian masing-masing pemakaian dibuat segmen sejumlah 4 (empat), 5 (lima) dan 6 (enam). Selanjutnya dibuat perbandingan error terjadi, lalu dipilih error yang terkecil. Jumlah segmen yang memberikan *error misclassification* terkecil inilah dianggap jumlah segmen yang paling baik.

NO	PEMAKAIAN	ERROR MISCLASSIFICATION		
		4 segmen	5 segmen	6 segmen
1	LOKAL	0.0410	0.0150	0.0200
2	SLJJ	0.0070	0.0100	0.0130
3	PONSEL	0.0020	0.0090	0.0100
4	SLI007	0.0070	0.0120	0.0200
5	SLI001	0.0020	0.0010	0.0011
6	SLI008	0.0010	0.0001	0.0010
7	INET	0.0020	0.0060	0.0130
8	TENURE	0.0000	0.0001	0.0001
9	FAULT COUNT	0.0520	0.0000	0.0001
10	FAULT DURATION	0.0060	0.0140	0.0100

Tabel 4.5 Perbandingan *Error Misclassification*.

Setelah jumlah segmen yang paling baik ditentukan, maka dibuat segmentasi untuk masing-masing pemakaian. Adapun segmentasi tersebut seperti yang ditunjukkan di dalam lampiran 3.

#### 4.5.2 Pengkodean Hasil *Segmentasi*

Setelah dilakukan pembentukan *segmen* dari masing-masing data yang berkarakter *range*, langkah selanjutnya adalah membuat pengkodean terhadap hasil *segmen* tersebut. Dari *segmen segmen* yang terbentuk dibuat batasan-batasan *range*. Batasan *segmen* terbentuk dari *mean* dan *standard deviasi* yang dihasilkan dari proses *segmentasi*. Batasan berupa *start value* dan *end value* dari masing-masing *segmen* terbentuk dipergunakan sebagai identifikasi kode-kode nama variabel. Nama kode diawali dengan nama variabel kemudian diikuti dengan *value mean* yang terbentuk dari proses *segmentasi*. Adapun kode-kode variabel yang sudah terbentuk seperti yang ditampilkan dalam tabel 4.6 sebagai berikut .

Hasil pengkodean merupakan variabel yang digunakan dalam proses *association rules* yaitu merupakan *field* dari *data training*. Data dari *field* ini berbentuk data *flaging true* yang disingkat dengan notasi 't' dan *flaging false* yang disingkat dengan 'f'.



DATA RANGE	CLUSTER	MEAN	STANDARD DEVIASI	START VALUE	END VALUE	KODE	DESKRIPSI
LOKAL	cluster-1	26,565	23,024	3,541	49,589	LOK02rb	Pemakaian Lokal rata-rata Rp 26.565
LOKAL	cluster-4	142,995	44,415	98,580	187,410	LOK142rb	Pemakaian Lokal rata-rata Rp 142.995
LOKAL	cluster-3	354,937	89,311	265,626	444,248	LOK354rb	Pemakaian Lokal rata-rata Rp 354.937
LOKAL	cluster-5	870,674	355,952	514,722	1,226,626	LOK870rb	Pemakaian Lokal rata-rata Rp 870.674
LOKAL	cluster-2	8,196,857	3,492,885	4,703,972	11,689,742	LOK8_2jt	Pemakaian Lokal rata-rata Rp 8.196.857
PONSEL	cluster-1	40,654	50,713	-	91,367	SEL40rb	Pemakaian Ponsel rata-rata Rp 40.654
PONSEL	cluster-4	418,465	161,733	256,732	580,198	SEL418rb	Pemakaian Ponsel rata-rata Rp 418.465
PONSEL	cluster-3	1,291,451	371,781	919,670	1,663,233	SEL1_2jt	Pemakaian Ponsel rata-rata Rp 1.291.451
PONSEL	cluster-2	3,493,448	1,395,916	2,097,532	4,889,364	SEL3_4jt	Pemakaian Ponsel rata-rata Rp 3.493.448
INTERNET	cluster-1	30,271	59,620	-	89,891	INET30rb	Pemakaian Internet rata-rata Rp 30.271
INTERNET	cluster-4	724,229	444,117	280,112	1,168,346	INET724rb	Pemakaian Internet rata-rata Rp 724.229
INTERNET	cluster-3	7,654,149	3,243,394	4,410,756	10,897,543	INET7_6jt	Pemakaian Internet rata-rata Rp 7.654.149
INTERNET	cluster-2	28,390,368	6,159,521	22,230,846	34,549,889	INET28_3jt	Pemakaian Internet rata-rata Rp 28.390.368
SLJJ	Cluster-1	31,442	38,417	-	69,859	SLJJ31rb	Pemakaian SLJJ rata-rata Rp 31.442
SLJJ	Cluster-4	322,179	128,294	193,885	450,473	SLJJ322rb	Pemakaian SLJJ rata-rata Rp 322.179
SLJJ	Cluster-3	1,061,081	331,800	729,281	1,392,881	SLJJ1_1jt	Pemakaian SLJJ rata-rata Rp 1.061.081
SLJJ	Cluster-2	3,287,120	1,205,681	2,081,439	4,492,800	SLJJ3_2jt	Pemakaian SLJJ rata-rata Rp 3.287.120
SLI007	cluster-1	74,511	91,398	-	165,909	SLI07_74rb	Pemakaian SLI007 rata-rata Rp 74.511
SLI007	cluster-4	785,680	324,551	461,129	1,110,232	SLI07_785rb	Pemakaian SLI007 rata-rata Rp 785.680
SLI007	cluster-3	2,761,123	951,510	1,809,612	3,712,633	SLI07_2_7jt	Pemakaian SLI007 rata-rata Rp 2.761.123
SLI007	cluster-2	9,485,778	3,645,745	5,840,033	13,131,523	SLI07_9_4jt	Pemakaian SLI007 rata-rata Rp 9.485.778
SLI001	cluster-1	51,782	41,686	10,096	93,469	SLI01_51rb	Pemakaian SLI001 rata-rata Rp 51.782
SLI001	cluster-4	302,725	103,211	199,514	405,936	SLI01_302rb	Pemakaian SLI001 rata-rata Rp 302.725
SLI001	cluster-3	843,045	230,113	612,932	1,073,158	SLI01_843rb	Pemakaian SLI001 rata-rata Rp 843.045
SLI001	cluster-5	2,157,694	730,812	1,426,882	2,888,506	SLI01_2_1jt	Pemakaian SLI001 rata-rata Rp 2.157.694
SLI001	cluster-2	8,731,209	1,862,658	6,848,551	10,613,867	SLI01_8_7jt	Pemakaian SLI001 rata-rata Rp 8.731.209
FAULT COUNT	cluster-1	1	0	1	2	FC1X	Jumlah Gangguan rata-rata 1 kali
FAULT COUNT	cluster-4	4	1	3	5	FC4X	Jumlah Gangguan rata-rata 4 kali
FAULT COUNT	cluster-3	9	1	7	10	FC9X	Jumlah Gangguan rata-rata 9 kali
FAULT COUNT	cluster-5	16	2	13	18	FC16X	Jumlah Gangguan rata-rata 16 kali
FAULT COUNT	cluster-2	38	-	38	38	FC38X	Jumlah Gangguan rata-rata 38 kali
FAULT DURATION	cluster-1	11	9	2	20	FD11jam	Lama Gangguan rata-rata 11 jam
FAULT DURATION	cluster-4	47	11	35	58	FD47jam	Lama Gangguan rata-rata 47 jam
FAULT DURATION	cluster-3	105	22	82	127	FD105jam	Lama Gangguan rata-rata 105 jam
FAULT DURATION	cluster-5	224	65	159	288	FD224jam	Lama Gangguan rata-rata 224 jam
TENURE	cluster-4	5	2	2	7	TEN5thn	Lama Berlangganan rata-rata 5 tahun
TENURE	cluster-1	13	3	11	16	TEN13thn	Lama Berlangganan rata-rata 13 tahun
TENURE	cluster-5	26	4	22	30	TEN26thn	Lama Berlangganan rata-rata 26 tahun
TENURE	cluster-3	52	5	47	57	TEN52thn	Lama Berlangganan rata-rata 52 tahun

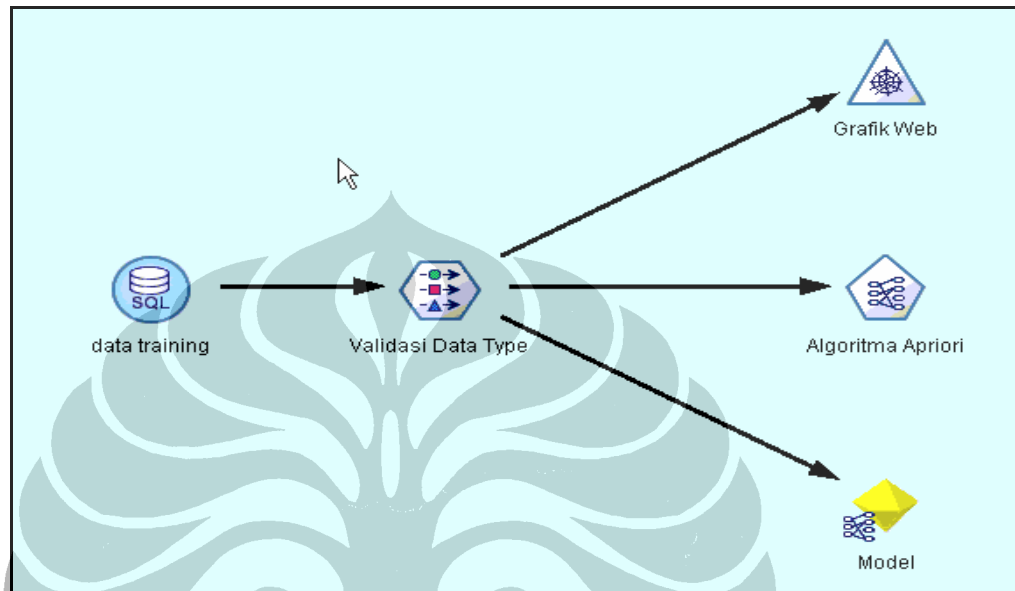
Tabel 4.6 Pengkodean dari hasil *Segmentasi*

### 4.5.3 Pemodelan *Association Rules*.

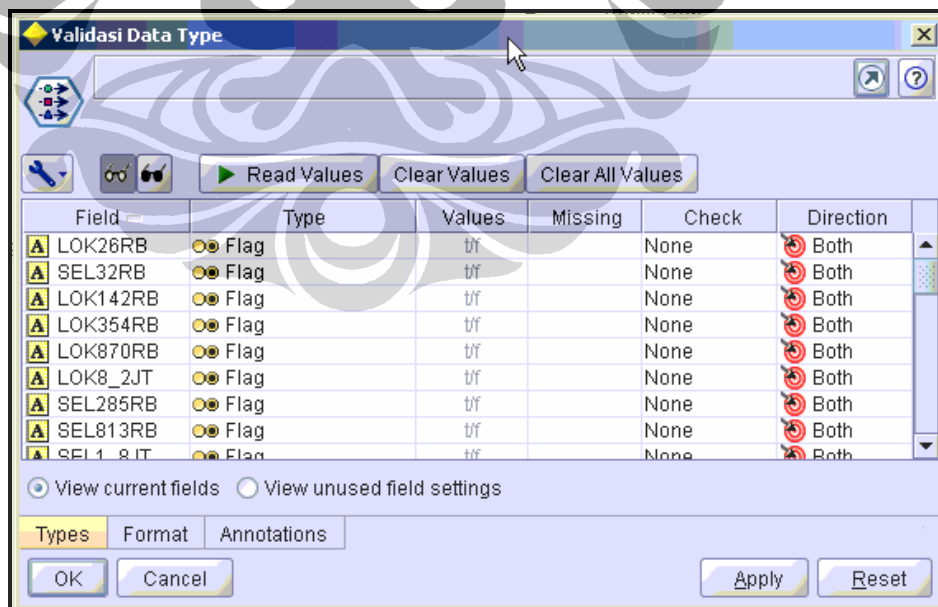
Setelah dilakukan penyiapan *data training*, termasuk pembentukan *segmentasi* terhadap data yang berkarakteristik *range*, maka selanjutnya adalah membuat model *association rules*. Model ini dibuat dengan menggunakan *tools* dari *clementine* yang menyediakan berbagai macam algoritma. Di dalam penulisan ini, pembuatan model *Association rules* menggunakan Algoritma *Apriori*. Seperti yang dijelaskan sebelumnya bahwa pencarian *association rules* dilakukan terhadap *data training usage service category, customer category, area switching, tenure, fault count* dan *fault duration*. Setelah dilakukan pengkodean sesuai subbab 4.5.2 maka jumlah variabel/*field* input model sebanyak empat puluh lima variabel.

Proses pemodelan digambarkan seperti gambar 4.16. *Data training* yang disiapkan menggunakan *database oracle* dikoneksikan ke *tools clementine*, dalam hal ini didefinisikan dalam obyek *data training*. Kemudian obyek ini diteruskan

kepada obyek *validasi type data*, untuk memastikan *type* data yang diproses sudah sesuai dengan kebutuhan model. Adapun *type* data model association rules ini adalah dalam bentuk *flag* yaitu *true* atau *false* seperti yang ditunjukkan dalam gambar 4.17. Dari obyek *validasi type data* dilanjutkan ke proses *data mining* yaitu algoritma *apriori* dan obyek *grafik web*.



Gambar 4.16 Model Association Rules



Gambar 4.17 Validasi Type Data

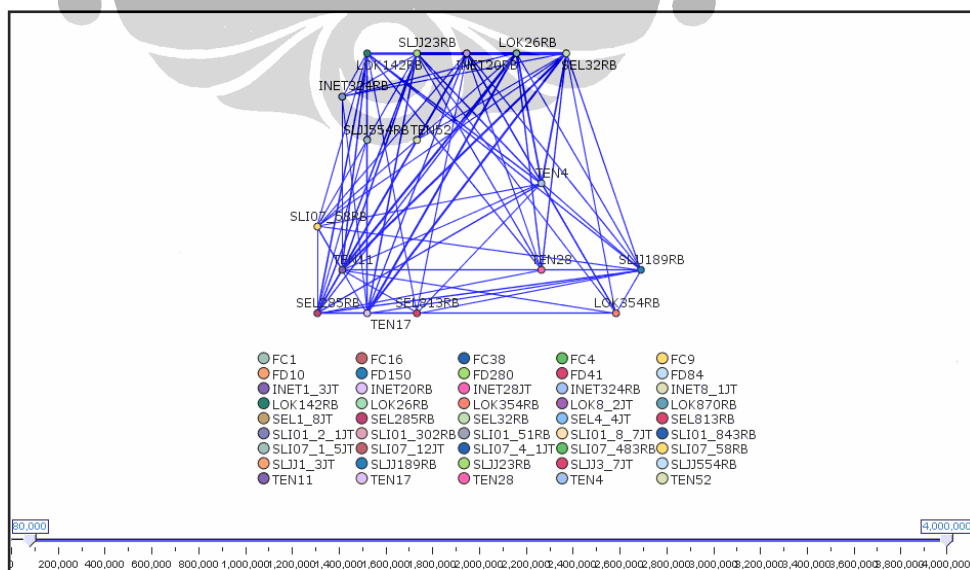
Setelah model algoritma *apriori* dijalankan menghasilkan model *association rules* seperti gambar 4.18 sebagai berikut:

Consequent	Antecedent	Support %	Confidence %
SEL40RB	LOK26RB	50.669	69.466
LOK26RB	SEL40RB	49.315	71.372
SEL40RB	SLJJ31RB	28.082	62.368
SLJJ31RB	SEL40RB	49.315	35.515
LOK26RB	TEN13THN	32.078	53.199
TEN13THN	LOK26RB	50.669	33.68
TEN13THN	SEL40RB	49.315	32.364
SEL40RB	TEN13THN	32.078	49.755
SLJJ31RB	LOK26RB	50.669	31.281
LOK26RB	SLJJ31RB	28.082	56.441
LOK26RB	TEN26THN	27.954	54.782
TEN26THN	LOK26RB	50.669	30.223
SEL40RB	TEN26THN	27.954	54.406
TEN26THN	SEL40RB	49.315	30.84
LOK26RB	TEN5THN	26.459	45.678
TEN5THN	LOK26RB	50.669	23.852
SEL40RB	TEN13THN LOK26RB	17.065	69.837

Gambar 4.18 Hasil Model *Association Rules*

Secara lengkap hasil pemodelan *Association Rules* terdapat di lampiran 4

Kemudian secara visual hubungan antar *field*/variabel dapat digambarkan dengan grafik *web*. Grafik ini dijelaskan hubungan antara beberapa variabel yang dominan maupun yang tidak. Semakin tebal garis hubungan / *link* menunjukkan bahwa hubungannya semakin dominan. Adapun grafik web seperti yang ada pada gambar 4.19



Gambar 4.19 *Grafik web*

## 4.6 Evaluasi Model

Setelah model algoritma *apriori* dijalankan menghasilkan *association rules* yang berkaitan dengan pola belanja *customer* terhadap *usage* layanan telekomunikasi, lama berlangganan (*tenure*) dan adanya gangguan (*fault duration*). Adapun pengetahuan-pengetahuan baru yang muncul dengan tingkat *support* dan *confidence* relatif tinggi adalah sebagai berikut:

### A. Berkaitan dengan pola pemakaian (*usage*).

- 1) *Association Rules* : Lok26rb → Sel40rb (49%,71%)

Hal ini berarti bahwa *customer* yang menggunakan *usage* lokal rata-rata Rp 26.000,- dan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian dari *customer* menggunakan *usage* lokal rata-rata sebanyak Rp 26.000 kemungkinan menggunakan seluler rata-rata Rp 40.000, sebesar 71%.

- 2) *Association Rules*: Sel40rb → Lok26rb (50%,69%).

Hal ini berarti bahwa *customer* yang menggunakan *usage* seluler rata-rata Rp 40.000,- dan lokal rata-rata Rp 26.000 sebanyak 50%. Kemudian *customer* yang menggunakan seluler rata-rata Rp 40.000 kemungkinan menggunakan lokal rata-rata Rp 26.000, sebesar 69%.

- 3) *Association Rules*: Sel40rb → Sljj31rb (28%,62%)

Hal ini berarti bahwa *customer* yang menggunakan *usage* seluler rata-rata Rp 40.000,- dan SLJJ rata-rata Rp 31.000 sebanyak 29%. Kemudian *customer* yang menggunakan seluler rata-rata Rp 40.000,- kemungkinan menggunakan SLJJ rata-rata Rp 31.000,- sebesar 62%.

- 4) *Association Rules*: Sljj31rb → Sel40rb (49%,36%)

Hal ini berarti bahwa *customer* yang menggunakan *usage* SLJJ rata-rata Rp 31.000,- dan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian *customer* yang menggunakan SLJJ rata-rata Rp 31.000,- kemungkinan menggunakan seluler rata-rata Rp 40.000, sebesar 36%

**Dari pola-pola *association rules* yang ditemukan tersebut, diperoleh pengetahuan bahwa umumnya *customer* dengan pemakaian salah satu layanan rendah maka untuk pemakaian layanan lainnya juga relatif rendah.**

**Dengan kata lain total belanja komunikasi untuk tiap bulannya dibatasi dengan jumlah tertentu.**

B. Kemudian yang berkaitan dengan *tenure* dan *usage Association Rules* yang diperoleh adalah sebagai berikut :

- 1) *Association Rules*: Ten13thn → Lok26rb (51%,34%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 13 tahun dan menggunakan *usage* lokal rata-rata Rp 26.000 sebanyak 51%. Kemudian yang sudah berlangganan sebelas tahun dengan kemungkinan pemakaian lokal Rp 26.000 sebesar 34%.

- 2) *Association Rules*: Ten13thn → Sel40rb (49%,32%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 13 tahun dan menggunakan Seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian yang sudah berlangganan sebelas tahun dengan kemungkinan pemakaian seluler Rp 40.000 sebanyak 32%.

- 3) *Association Rules*: Ten26thn → lok26rb (50%,30%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 17 tahun dan menggunakan lokal rata-rata Rp 26.000 sebanyak 50%. Kemudian yang sudah berlangganan tujuh belas tahun dengan kemungkinan pemakaian lokal Rp 26.000, sebesar 30%.

- 4) *Association Rules*: Ten26thn → Sel40rb (49%,31%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 26 tahun dan menggunakan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian yang sudah berlangganan selama tujuh belas tahun dengan kemungkinan pemakaian seluler Rp 40.000, sebesar 31%.

**Dari pola-pola tersebut dapat diambil pengetahuan baru bahwa *customer* yang sudah berlangganan cukup lama, pemakaian masing-masing layanan rata-rata relatif rendah.**

C. Kemudian bila *usage* dikaitkan dengan *fault duration* maka *Association Rules* yang didapatkan adalah sebagai berikut :

- 1) *Association Rules*: FD11JAM → SEL40RB (49%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* seluler rata-rata Rp 40.000 sebanyak 49%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 40.000, hanya sebesar 1%.

2) *Association Rules*: FD11JAM → SLJJ31RB (28%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* SLJJ rata-rata Rp 31.000 sebanyak 28%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 40.000, hanya sebesar 1%.

3) *Association Rules*: FD11JAM → LOK142RB (24%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* Lokal rata-rata Rp 142.000 sebanyak 24%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 142.000, hanya sebesar 1%.

**Dari pola-pola tersebut tampak jelas bahwa *fault duration* cukup mempengaruhi *opportunity* pelanggan yang akan menggunakan telepon. Dengan kata lain bila telepon terganggu tentunya tidak dapat digunakan untuk berkomunikasi, kemungkinan untuk menggunakan layanan rendah.**

Secara lengkap *association rules* yang ada terdapat dalam lampiran 4.

#### 4.7 Penerapan Model

Proses pengerjaan *data mining* tahapan terakhir adalah *deploy* atau penerapan model. Dalam penelitian ini dipilih *deploy* sederhana dengan menggunakan data lima bulan berikutnya yaitu data bulan januari sampai dengan bulan juni 2008. Dengan menggunakan *field*/kolom yang sama dengan *data training*, tetapi dengan bulan berbeda akan menghasilkan *deploy* model yang tidak jauh berbeda untuk angka *support level* maupun *confidence level*-nya, seperti yang ditunjukkan dalam tabel 4.7 berikut ini.

Consequent	Antecedent	Model		Deploy	
		Support %	Confidence %	Support %	Confidence %
lok26rb	sel40rb	49	71	50	71
sel40rb	lok26rb	50	69	50	70
sel40rb	sljj31rb	28	62	29	63
sljj31rb	sel40rb	49	36	49	34
ten13thn	lok26rb	51	34	52	30
ten13thn	sel40rb	49	32	49	31
ten26thn	lok26rb	51	30	52	30
ten26thn	sel40rb	50	31	49	31
fd11jam	sel40rb	49	1	49	1
fd11jam	sljj31rb	28	1	26	1
fd11jam	lok142rb	24	1	22	1

Tabel 4.7 Association Rules Model dan deploy

#### 4.8 Pemanfaatan Pengetahuan *Data Mining*

Dari hasil pengetahuan yang diperoleh dari association rules maka kemungkinan-kemungkinan pemanfaatan yang dapat diterapkan oleh PT. XYZ khususnya pola-pola yang terbentuk dari data *usage* maupun data fault adalah:

- a. Pembuatan paket *product* berupa *cross selling*, khusus untuk pelanggan dengan pola *usage* rendah untuk lokal, SLJJ. Dengan membuat paket harga *bundling* murah dan pemakaian sepuasnya.
- b. Melakukan program *customer retention* khususnya untuk pelanggan yang sudah lama berlangganan dengan *usage* rendah, dengan penawaran paket produk/layanan baru lainnya yang lebih menarik.
- c. Penyusunan kebijakan *Fault handling*, khususnya percepatan penyelesaian gangguan khusus untuk pelanggan yang memiliki *usage* tertentu. Agar kegagalan *customer* menggunakan layanan dapat ditekan.