

**IMPLEMENTASI DATA WAREHOUSE DAN DATA MINING
UNTUK MENEMUKAN POLA PEMAKAIAN LAYANAN PADA
INDUSTRI TELEKOMUNIKASI**

KARYA AKHIR

Diajukan sebagai salah satu syarat untuk memperoleh
gelar Magister Teknologi Informasi

INDRA SUYITNO

0706194293



**UNIVERSITAS INDONESIA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
JULI 2009**

HALAMAN PERNYATAAN ORISINALITAS

Karya Akhir ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.

Nama : Indra Suyitno

NPM : 0706194293

Tanda tangan :

Tanggal :

HALAMAN PENGESAHAN

Karya Akhir ini diajukan oleh :

Nama : Indra Suyitno

NPM : 0706194293

Program Studi : Magister Teknologi Informasi

Judul Karya Akhir : IMPLEMENTASI DATA WAREHOUSE DAN DATA MINING UNTUK MENEMUKAN POLA PEMAKAIAN LAYANAN PADA INDUSTRI TELEKOMUNIKASI

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Teknologi Informasi pada Program Studi Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Indonesia.

DEWAN PENGUJI

Pembimbing : Yudho Giri Sucahyo, PhD..... (.....)

Pembimbing : Yova Ruldeviyani, M.Kom.....(.....)

Penguji : Indra Budi, Dr..... (.....)

Penguji : Rizal Fathoni Aji, M.Kom..... (.....)

Ditetapkan di : Jakarta

Tanggal :

KATA PENGANTAR

Segala puji bagi Allah, Tuhan sekalian alam. Atas segala limpahan rahmat dan berkah, kami tidak henti-hentinya mengucapkan syukur. Sesungguhnya Allah SWT Maha Terpuji dan Maha Kuasa tiada daya dan kekuatan melainkan atas pertolongan Allah SWT. Sehingga saya dapat menyelesaikan tesis ini yang berjudul : IMPLEMENTASI DATA WAREHOUSE DAN DATA MINING UNTUK MENEMUKAN POLA PEMAKAIAN LAYANAN PADA INDUSTRI TELEKOMUNIKASI.

Upaya penyelesaian tesis ini secara langsung atau tidak langsung terbantu juga oleh orang-orang yang telah memberi dorongan semangat, inspirasi atau dalam bentuk lain. Oleh karena itu saya mengucapkan terima kasih kepada :

1. Bapak Yudho Giri Sucahyo, PhD dan Ibu Yova Ruldeviyani, M.Kom sebagai pembimbing tesis yang selalu memberikan arahan, motivasi dan dorongan agar segera menyelesaikannya.
2. Kedua orang tua saya yang selalu memberi dukungan semangat dan do'a.
3. Istriku, Sri Murhandjati dan Anakku, Ahmad Shafy Irfanto yang telah memberikan inspirasi dan penggugah semangat dalam percepatan penyelesaian tesis.
4. Seluruh staf pengajar dan non pengajar yang telah banyak membantu dalam proses pembelajaran selama mengikuti program perkuliahan MTI UI.
5. Rekan-rekan seluruh angkatan 2007 MTI-UI yang telah membantu memberikan berbagai macam *literature* dan aplikasi *software*.
6. Para pejabat struktural maupun non struktural dan rekan-rekan kerja di PT.XYZ yang telah banyak membantu dalam pemenuhan data dan informasi lainnya.

**HALAMAN PERNYATAAN PERSETUJUAN
PUBLIKASI
KARYA AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Indra Suyitno.
NPM : 0706194293
Program Studi : Magister Teknologi Informasi
Departemen :
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi/Tesis/Disertasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul :

**IMPLEMENTASI DATA WAREHOUSE DAN DATA MINING UNTUK
MENEMUKAN POLA PEMAKAIAN LAYANAN PADA INDUSTRI
TELEKOMUNIKASI**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-ekskutif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*). Merawat, dan mempublikasikan karya akhir saya tanpa meminta izin dari saya selama tetap mencantumkan saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal :

Yang menyatakan

(Indra Suyitno)

ABSTRACT

Name : Indra Suyitno

Study Program : Master of Information Technology

Title : Implementation of Data Warehouse and Data Mining: Finding pattern of service usage in Industry of Telecommunication

Competition on business industry of telecommunication is very tighten. Consequently strategy and new idea should be thought by corporate executives, so that their business can be successful. Customer satisfaction is ones of important factor that must be noticed in order to customer will have loyalty using their product.

Data of billing and network performance are very important datas in the industry of telecommunication, even one valueable asset in business and management of telecommunication. Be related to that matter has used technology to explore new knowledge to help their business strategy planning in the future. Technology of data mining is one of the solution can be applied.

This research will explore Implementation of data mining for finding association rules with apriori algorithm, inside of industry of telecommunication. Association rules can be interpreted to be new knowledge, its about either characteristic of wireline product of operator telecommunication services. Better, the new knowledge can be used to analize of matter to increase performance services so that customer still believe and use their product.

Keyword : Data Mining, Association Rules, Apriori, Telecommunication

xv + 77 pages ; 19 figure; 5 graphic; 7 tabel ; 4 attachment

ABSTRAK

Nama : Indra Suyitno

Program Studi : Magister Teknologi Informasi

Judul : Implementasi *Data Warehouse* dan *Data Mining* untuk Menemukan *Pola Pemakaian Layanan* pada Industri Telekomunikasi

Persaingan dalam dunia bisnis khususnya Industri Telekomunikasi semakin ketat membuat para pelaku harus memikirkan strategi-strategi atau terobosan yang dapat menjamin keberlangsungan bisnis mereka. Kepuasan pelanggan merupakan salah satu faktor yang sangat perlu diperhatikan untuk menjaga pelanggan agar tetap setia pada produk atau layanan yang ditawarkan.

Dalam Industri Telekomunikasi data *billing* dan data kinerja *Network* dapat dikatakan sebagai data produksi yang sangat penting, bahkan merupakan asset yang berharga di dalam pengelolaan Industri Telekomunikasi. Hal ini menciptakan sebuah kebutuhan akan adanya teknologi yang dapat memanfaatkannya untuk menggali pengetahuan-pengetahuan baru, yang dapat membantu dalam perencanaan strategi bisnis di masa depan. Dalam hal ini teknologi *data mining* merupakan salah satu solusi yang dapat diterapkan.

Dalam penelitian ini akan dibahas Implementasi *Data Mining* dan *Data Warehouse* untuk Menemukan Pola Pemakaian Layanan pada Industri Telekomunikasi. Model data mining menggunakan *association rules* algoritma *apriori*. *Association rules* yang dihasilkan yang dapat diinterpretasikan menjadi pengetahuan baru mengenai karakteristik obyek produk atau layanan berbasis *wireline* salah satu operator telekomunikasi. Pengetahuan baru nantinya dapat digunakan sebagai bahan analisis dalam menentukan rencana kebijakan strategis di masa yang akan datang dalam rangka meningkatkan kinerja layanan agar keberadaan pelanggan dapat dipertahankan dan ditingkatkan.

Kata kunci : *Data Mining*, *Association Rules*, *Apriori*, Telekomunikasi

xv + 77 halaman ; 19 gambar; 5 grafik; 7 tabel ; 4 lampiran

DAFTAR ISI

JUDUL	i
HALAMAN PERNYATAAN ORISINALITAS	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRACT	vi
ABSTRAK	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR GRAFIK	xiii
DAFTAR TABEL	xiv
BAB 1 : PENDAHULUAN	1
1.1 Latar belakang	1
1.2 Permasalahan	2
1.3 Pembatasan masalah	3
1.4 Tujuan dan manfaat.....	4
BAB 2 : LANDASAN TEORI	5
2.1 <i>Data warehouse</i>	5
2.1.1 <i>Arsitektur data warehouse</i>	5
2.1.2 <i>Extraction, transformation and loading</i>	8
2.1.3 <i>Skema data warehouse</i>	9
2.2 <i>On-line analytical procesing (OLAP)</i>	10
2.3 <i>Konsep data mining</i>	11
2.4 <i>Metodologi knowledge discovery in database</i>	12
2.5 <i>Teknik teknik data mining</i>	17
2.5.1 <i>Association rules</i>	18
2.5.2 <i>Algoritma apriori</i>	18
2.5.3 <i>Classification</i>	22
2.5.3 <i>Segmentasi</i>	23

2.6	Produk <i>wireline</i>	24
2.7	Loyalitas pelanggan	26
BAB 3 : METODOLOGI PENELITIAN		28
BAB 4 : TAHAPAN Pengerjaan <i>DATA MINING</i>		33
4.1	Pemahaman permasalahan bisnis.....	33
4.1.1	Latar belakang masalah.....	33
4.1.2	Perumusan masalah.....	34
4.1.3	Tujuan dilakukan proses <i>data mining</i>	34
4.2	Pembuatan <i>database data mining</i>	35
4.2.1	Sumber data	35
4.2.2	Pemahaman bisnis.....	36
4.2.3	Pemahaman logika model	37
4.2.4	Perancangan dimensi model (<i>star schema</i>)	39
4.2.5	Pendefinisian <i>physical model</i>	40
4.2.6	Proses <i>extraction, transformation</i> dan <i>loading</i>	40
4.2.6.1	<i>Extraction</i>	40
4.2.6.2	<i>Transformation</i>	41
4.2.6.3	<i>Loading</i>	43
4.3	Penyelidikan data	43
4.3.1	Pembuatan tabel/ <i>view cube</i> OLAP.....	44
4.3.2	Pembuatan <i>query</i> OLAP	44
4.3.3	Pembuatan aplikasi OLAP	47
4.3.4	Data hasil OLAP	49
4.4	Persiapan data untuk pembuatan model <i>data mining</i>	54
4.4.1	Pemilihan data yang relevan untuk model <i>data mining</i>	54
4.4.2	Pemilihan <i>field/kolom data training</i> untuk <i>data mining</i>	55
4.5	Pembuatan model <i>data mining</i>	56
4.5.1	Pengelompokan (segmentasi)	56
4.5.2	Pengkodean hasil segmentasi.....	59
4.5.3	Pemodelan <i>association rules</i>	60
4.6	Evaluasi model	63
4.7	Penerapan model.....	65

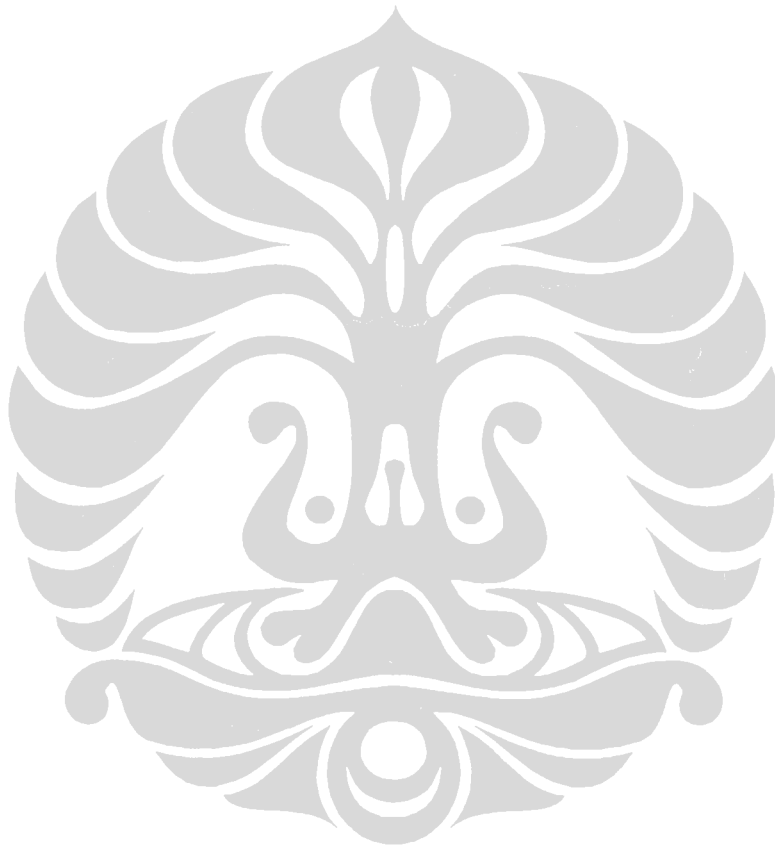
4.8 Pemanfaatan pengetahuan <i>data mining</i>	66
BAB 5 : PEMBAHASAN	67
5.1 Pertanyaan dan tujuan dari penelitian	67
5.1.1 Pertanyaan pertama	67
5.1.2 Pertanyaan kedua	67
5.1.3 Pertanyaan ketiga	68
5.1.4 Tujuan penelitian	71
BAB 6 : PENUTUP	72
6.1. Kesimpulan	72
6.2 Kemungkinan penelitian selanjutnya	74
DAFTAR PUSTAKA	75
LAMPIRAN	
Lampiran 1 : <i>Field-field</i> Dimensi	
Lampiran 2 : Daftar <i>field/kolom data training</i>	
Lampiran 3 : Hasil pengelompokkan	
Lampiran 4 : Hasil pencarian <i>association rules</i>	

DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>data warehouse</i> (<i>Simitsis,Theodoratos,2006</i>).....	6
Gambar 2.2 Arsitektur <i>data warehouse</i> (<i>Hoffer, Prescott, McFadden, 2005</i>)	8
Gambar 2.3 <i>Star schema</i> proses pemesanan(<i>Adam Son, 2006</i>)	10
Gambar 2.4 Bentuk pemodelan data <i>cube</i> dan tampilan data sisi dari <i>cube</i>	11
Gambar 2.5 Tahapan-tahapan dalam proses KDD (<i>Fayyad et.al, 1996</i>)	13
Gambar 2.6 Proses KDD Osmar	15
Gambar 2.7 Proses tahapan KDD CRISP-DM	16
Gambar 2.8 <i>Pseudocode</i> dari algoritma <i>Apriori</i>	20
Gambar 2.9 <i>Pseudocode</i> pembentukan kandidat <i>itemset</i>	20
Gambar 2.10 Contoh aplikasi algoritma <i>apriori</i>	21
Gambar 2.11 <i>Decision tree</i> dengan pengujian pada atribut X dan Y	22
Gambar 2.12 <i>Neural network diagram</i>	23
Gambar 3.1 Tahapan <i>data mining two crows-corporation</i>	29
Gambar 4.1 Arsitektur data warehouse adaptasi dari <i>Hoffer, Prescott, McFadden</i>	35
Gambar 4.2 ERD <i>customer wireline</i>	38
Gambar 4.3 <i>Star schema usage</i> dan <i>fault</i> , tabel dimensi dan tabel fakta.....	39
Gambar 4.4 Proses transformasi pada <i>field/kolom usageid</i>	42
Gambar 4.5 Proses transformasi pada <i>field/kolom tenure</i>	42
Gambar 4.6 Pembuatan <i>SQL cube</i>	44
Gambar 4.7 Hasil <i>Running SQL cube</i>	45
Gambar 4.8 Hasil <i>query</i> bulan tagihan, <i>usage category</i>	46
Gambar 4.9 <i>Drilldown switching, usage amount</i> dan jumlah gangguan	46
Gambar 4.10 <i>Slicing switching, drilldown customer category</i>	47
Gambar 4.11 Aplikasi <i>OLAP drilldown</i>	48
Gambar 4.12 Aplikasi <i>OLAP rollup</i>	49
Gambar 4.13 Pemodelan <i>Segmentasi K-means usage</i> lokal	57
Gambar 4.14 Hasil <i>segmen K-means</i> dari <i>usage</i> lokal	57

DAFTAR GAMBAR

	lanjutan
Gambar 4.15 <i>Pie diagram segmentasi usage lokal</i>	58
Gambar 4.16 Model <i>association rules</i>	61
Gambar 4.17 Validasi <i>type data</i>	61
Gambar 4.18 Hasil model <i>association rules</i>	62
Gambar 4.19 Grafik <i>web</i>	62



DAFTAR GRAFIK

Grafik 4.1 Komposisi Layanan Berdasarkan Prosentase <i>Usage</i> Tahun 2007	50
Grafik 4.2 Komposisi Layanan Berdasarkan Prosentase <i>User</i> tahun 2007	51
Grafik 4.3 Komposisi layanan berdasarkan prosentase <i>usage</i> untuk area <i>Switching</i>	52
Grafik 4.4 Profil <i>Fault Duration</i> dan <i>Usage</i> di tiap Area <i>Switching</i>	53
Grafik 4.5 Profil <i>Tenure</i> , <i>Fault Count</i> dan <i>Fault Duration</i>	53



DAFTAR TABEL

Tabel 4.1 Daftar database dan table sebagai sumber data	36
Tabel 4.2 Tabel dimensi dan fakta	40
Tabel 4.3 Field Data Warehouse dan Field Sumber data.....	41
Tabel 4.4 <i>Field/kolom data training</i>	55
Tabel 4.5 Perbandingan <i>Error Misclassification</i>	59
Tabel 4.6 Pengkodean dari hasil <i>segmentasi</i>	60
Tabel 4.7 <i>Association rules</i> model dan <i>deploy</i>	66

