

**PERANCANGAN *DATA WAREHOUSE* DAN PENERAPAN
DATA MINING ANGGOTA *COBRAND* PADA SISTEM
FREQUENT FLYER :
STUDI KASUS PT. GARUDA INDONESIA**

KARYA AKHIR

**MUJOKO
0706193813**



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

**PERANCANGAN *DATA WAREHOUSE* DAN PENERAPAN
DATA MINING ANGGOTA *COBRAND* PADA SISTEM
FREQUENT FLYER :
STUDI KASUS PT. GARUDA INDONESIA**

KARYA AKHIR

**Diajukan sebagai salah satu syarat
untuk memperoleh gelar Magister Teknologi Informasi**

**MUJOKO
0706193813**



UNIVERSITAS INDONESIA

FAKULTAS ILMU KOMPUTER

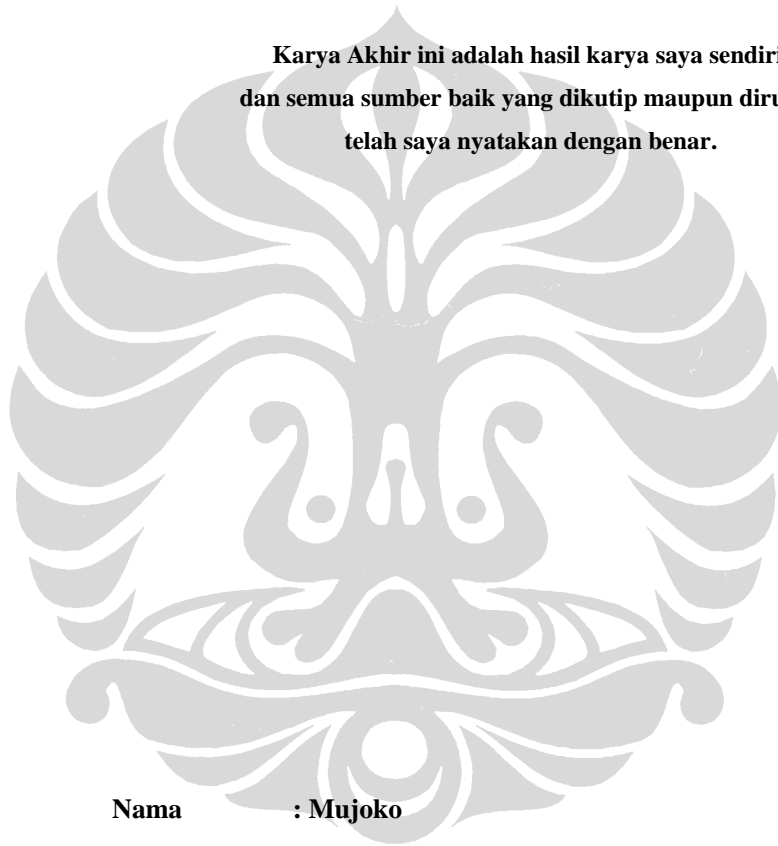
PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI

JAKARTA

JANUARI 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 : Mujoko

NPM : 0706193813

Tanda tangan :

Tanggal : 09 Januari 2009

HALAMAN PENGESAHAN

Karya Akhir ini diajukan oleh :

Nama : Mujoko
NPM : 0706193813
Program Studi : Magister Teknologi Informasi
Judul Karya Akhir : Perancangan *Data Warehouse* dan Penerapan
Data Mining Anggota Cobrand Pada Sistem
Frequent Flyer : Studi Kasus PT. Garuda
Indonesia

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

Pembimbing : Yudho Giri Sucahyo, Ph.D, CISA (.....)

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

Penguji : Dr. Achmad Nizar H (.....)

Ditetapkan di : Jakarta

Tanggal : 9 Januari 2009

KATA PENGANTAR

Segala puji syukur kepada Allah SWT, karena hanya dengan berkat dan karunia-Nya penulis dapat menyelesaikan tesis yang berjudul “PERANCANGAN *DATA WAREHOUSE* DAN PENERAPAN *DATA MINING* ANGGOTA *COBRAND* PADA SISTEM *FREQUENT FLYER* STUDI KASUS PT. GARUDA INDONESIA” ini sesuai dengan yang direncanakan. Tesis ini dibuat guna melengkapi persyaratan kelulusan pada Program Studi Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Indonesia. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan karya akhir ini, sangatlah sulit bagi saya untuk menyelesaikannya. Oleh karena itu, saya mengucapkan terima kasih kepada:

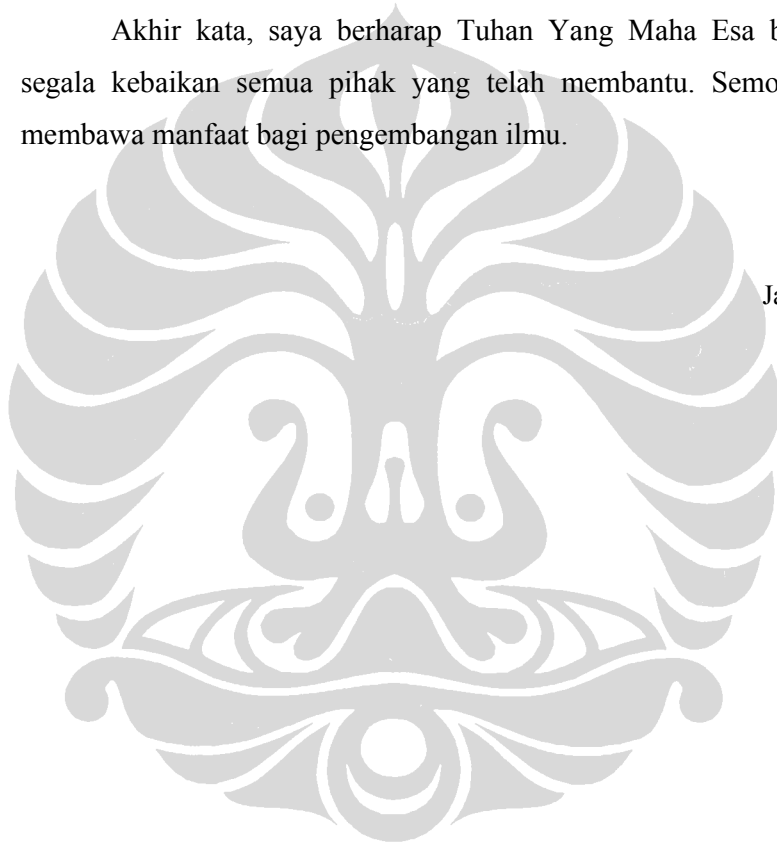
- (1) Orang tua, istri dan ananda Muhammad Ihsan yang telah memberikan dukungan bantuan berupa api semangat.
- (2) Ibu Yova Ruldeviyani, M.Kom sebagai dosen pembimbing yang telah membimbing penulis selama mengerjakan Tesis ini.
- (3) Bapak Yudho Giri Suchahyo, Ph.D, CISA, selaku dosen pembimbing akademik dan ketua program studi Studi Magister Teknologi Informasi, yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini mulai dari penyusunan proposal hingga Tesis dan jabatannya berakhir
- (4) Bapak Dr. Indra Budi selaku penguji dalam sidang tesis
- (5) Bapak Dr. Achmad Nizar H. selaku penguji dalam sidang tesis, sekaligus ketua program studi Magister Teknologi Informasi mendatang dan selamat atas jabatannya.
- (6) Bapak K. Budiyanto, Ibu Lusi, Ibu Cita, Pak Sonny dan segenap staff karyawan Garuda yang telah bersedia meluangkan waktu untuk menjawab pertanyaan-pertanyaan penulis dan menyediakan data, semoga Allah membalas kebaikan rekan semua.
- (7) Ibu Connie dan Pak Heru selaku atasan saya yang memberi kesempatan pada saya untuk mengeksplorasi FFP sebagai bahan kajian.
- (8) Sahabat-sahabat saya di FFP yang telah berjuang siang dan malam, Bank Dukun/Mukhlis, Pak Dipo, Iin, Vera, Jonny dan Gardiary sang pejection, yang telah banyak membantu saya dalam menyelesaikan karya akhir ini.

- (9) Bang Haikal, yang telah membantu saya selama saya sakit atas dukungannya berupa materi dan material. Semoga Allah membalas kebaikan abang.
- (10) Rekan kelas B dan kelompok ABUDOMM, terima kasih atas kebersamaan dan kenangan indah yang tak terlupakan selama di MTI.
- (11) Kang Asep dan Dodick dengan tumpangan kendaraan setiap saya pulang selama kuliah.
- (12) Mas Ganda yang telah menyempahi saya agar mendapat nilai yang terbaik.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga karya akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 9 Januari 2009

Penulis



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

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

Nama : Mujoko
NPM : 0706193813
Program Studi : Magister Teknologi Informasi
Departemen : Magister Teknologi Informasi
Fakultas : Ilmu Komputer
Jenis Karya : Tesis

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

Perancangan *Data Warehouse* dan Penerapan *Data Mining* Anggota *Cobrand* Pada Sistem *Frequent Flyer* : Studi Kasus PT. Garuda Indonesia

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 : 09 Januari 2009

Yang menyatakan

(.....Mujoko.....)

ABSTRACT

Globalization is now inevitable, which causes airlines to face a major problem nowadays. Time becomes major influence in decision making. If the management takes a lot of time in decision making, then company will be stagnant or even decline.

There are needs to execute business decision with little time consuming, so that more opportunities can be captured with less cost. In addition to time consuming, accuracy is also a vital necessity for mature analysis and rational calculation.

Any information from the insignificant one, as long as it is relevance and accurate can be used as an opportunity for increasing revenue. The main principle is how to optimize aircraft load to its maximum capacity as often as possible. In the context of winning a competition, the combination of a flexible management style and utilization of Information Technology become a vital element.

Any business entity can survive depends on the existence of its customer. In order to gain more customer, business player has to execute strategy and innovation. This is necessary to influence people in choosing services provider. In this case, the context is airline service business. Any kind of Loyalty Program such as Garuda Frequent Flyer (GFF), has important role in targeting core customer, i.e. customer that give the highest revenue to the company.

GFF operational system that keeps information of customer profile and flight transaction, needs to be collected into a single form of data warehouse and updated periodically, so that reporting function can be executed without affecting operational performance. The extensive use of data with datamining technique could assist management in making the right decision based on historical data. Additionally, datamining has to be relevance and being part of business process in order to improve business performance. The focus of this research is to develop initial phase of datamining that can be scaled up in future, not a practical solution that can be directly used.

Keyword : Data Warehouse, Data Mining, Frequent Flyer, Garuda Frequent Flyer

xii+109 pages; 35 figures; 10 tables; 8 attachments; 0 technical documentation

Bibliography: 16 (2001- 2008)

ABSTRAK

Globalisasi memang arus yang tidak bisa ditangkal. *Airlines* menghadapi masalah besar pada saat ini dan hari-hari mendatang. Manajemen yang terbiasa lambat mengambil keputusan dipastikan hanya akan membawa perusahaannya *stall*. Diperlukan kecepatan dalam mengeksekusi bisnis, agar kesempatan dapat diraih lebih banyak. Selain kecepatan, diperlukan juga ketepatan pengambilan keputusan, berdasarkan analisa perhitungan yang rasional dan matang.

Persaingan mengkondisikan bisnis *airlines* dalam kondisi perang total. Informasi yang sekecil apapun tetapi relevan dimanfaatkan sebagai sebuah peluang untuk meningkatkan pendapatan. Prinsipnya adalah bagaimana memenuhi kapasitas angkut pesawat secara maksimal dan sesering mungkin. Dalam konteks memenangkan persaingan ini perpaduan antara gaya manajemen yang fleksibel dan peran teknologi informasi menjadi amat vital.

Setiap bisnis akan tetap eksis jika dia memiliki *customer*. Untuk mendapatkan *customer*, pelaku bisnis melakukan berbagai strategi dan inovasi yang akan mempengaruhi ketertarikan orang dan memutuskan menggunakan layanan yang ditawarkan oleh penyedia layanan jasa/barang, dalam hal ini layanan angkutan penerbangan. *Loyalty program* semacam *Garuda Frequent Flyer* (GFF), memegang peranan penting untuk membidik *core customer*, *customer* yang paling banyak memberikan *benefit* bagi perusahaan.

Sistem operasional GFF yang menyimpan data anggota dan transaksi penerbangan, perlu dikumpulkan dalam satu bentuk data *warehouse* dan secara periodik di-*update*, sehingga *reporting* dapat dieksekusi dari data *warehouse* tanpa mempengaruhi kinerja operasional. Penggunaan data yang ekstensif dengan menggunakan teknik data mining dapat membantu manajemen dalam mengambil kebijakan yang tepat berdasar data-data historical. Agar *data mining* dapat mempengaruhi bisnis, maka *data mining* sendiri harus relevan dan menjadi bagian dari bisnis proses.

Kata Kunci : *Data Warehouse, Data Mining, Frequent Flyer, Garuda Frequent Flyer*

xii+109 halaman; 35 gambar; 10 table; 8 lampiran; 0 technical documentation

Bibliography: 16 (2001- 2008)

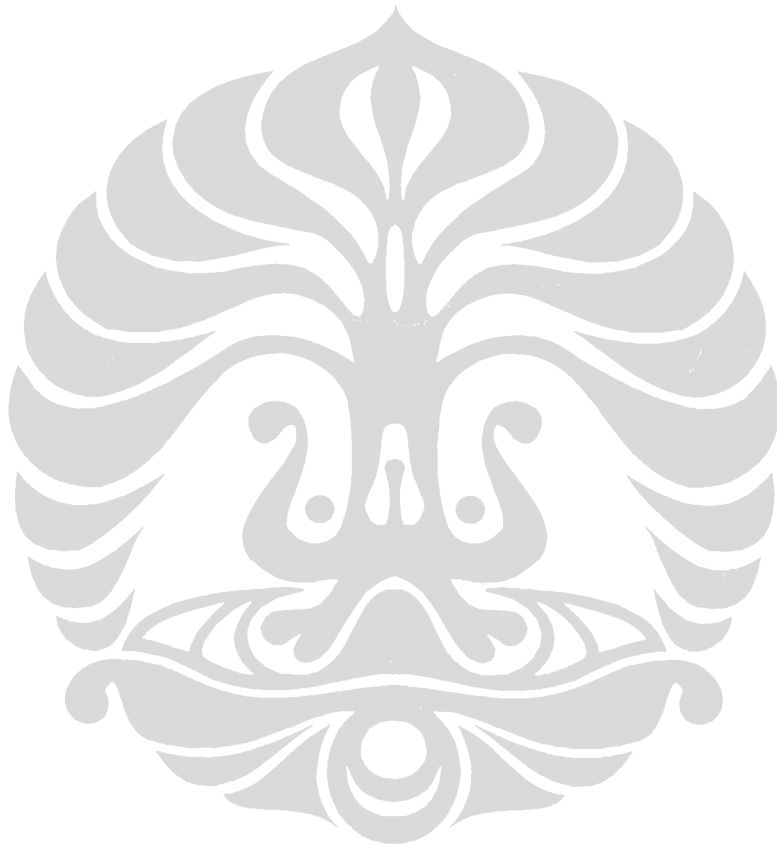
DAFTAR ISI

KATA PENGANTAR	I
ABSTRACT	IV
ABSTRAK	V
DAFTAR ISI	VI
DAFTAR GAMBAR	X
DAFTAR TABEL	XII
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 PERMASALAHAN	2
1.3 PEMBatasan MASALAH.....	3
1.4 TUJUAN DAN MANFAAT PENELITIAN.....	4
BAB II LANDASAN TEORI	5
2.1 <i>DATA WAREHOUSE</i>	5
2.1.1 <i>Definisi Data Warehouse</i>	5
2.1.2 <i>Penelitian Implementasi Data Warehouse</i>	6
2.1.3 <i>Sistem Operasional dan Sistem Pendukung Pengambilan Keputusan</i>	7
2.1.4 <i>Keuntungan Data Warehouse</i>	8
2.1.5 <i>Kategori Data pada Data Warehouse</i>	10
2.1.6 <i>Arsitektur Data Warehouse</i>	10
2.1.7 <i>Tahapan Data Warehouse</i>	12
2.1.8 <i>Desain Data Warehouse</i>	15
2.2 <i>DATA MINING</i>	17
2.2.1 <i>Definisi Data Mining</i>	17
2.2.2 <i>Teknik Data Mining</i>	18
2.2.2.1 <i>Predictive Modeling</i>	19
2.2.2.2 <i>Database Segmentation dengan Clustering</i>	21
2.2.2.3 <i>Link Analysis</i>	22
2.2.2.4 <i>Deviation Detection</i>	23
2.2.3 <i>Tahap-Tahap Data Mining</i>	24
2.3 <i>BISNIS AIRLINE DAN DATA MINING</i>	27

2.3.1 <i>Loyalty Program dan Frequent Flyer Program</i>	28
2.4 <i>TOOLS DEVELOPMENT</i>	30
2.4.1 <i>Pentaho Data Integration (PDI/Kettle)</i>	30
2.4.2 <i>Mondrian</i>	31
2.5 <i>KEBUTUHAN DATA MINING DALAM FREQUENT FLYER PROGRAM</i>	31
BAB III METODOLOGI PENELITIAN	32
3.1 <i>PENGUMPULAN DATA</i>	32
3.2 <i>KEBUTUHAN BISNIS DAN INFORMASI</i>	32
3.3 <i>PERANCANGAN DATA WAREHOUSE</i>	33
3.4 <i>IMPLEMENTASI DATA MINING</i>	33
BAB IV PERANCANGAN DATA WAREHOUSE	36
4.1 <i>PROFIL PERUSAHAAN</i>	36
4.1.1 <i>Latar Belakang Perusahaan</i>	37
4.1.2 <i>Visi dan Misi Perusahaan</i>	37
4.1.3 <i>Struktur Organisasi Perusahaan</i>	38
4.2 <i>PEMAHAMAN TERHADAP PROSES BISNIS</i>	40
4.2.1 <i>Pendaftaran (Enrollment)</i>	42
4.2.2 <i>Upgrade/Downgrade Tier</i>	43
4.2.3 <i>Aktifitas Earning Anggota</i>	44
4.2.4 <i>Aktifitas Redeem Anggota</i>	44
4.2.5 <i>Keseimbangan Accrual dan Reedem</i>	45
4.2.6 <i>Kerja Sama Partner</i>	45
4.3 <i>PERANCANGAN ARSITEKTUR</i>	45
4.3.1 <i>Arsitektur logical</i>	46
4.3.2 <i>Arsitektur Fisik</i>	47
4.3.4 <i>Sumber Data</i>	48
4.3.5 <i>Data Staging</i>	49
4.3.6 <i>Proses ETL</i>	50
4.3.6.1 <i>Proses Extract</i>	50
4.3.6.2 <i>Proses Cleansing</i>	51
4.3.6.3 <i>Proses Transformasi</i>	52
4.3.6.4 <i>Proses Loading</i>	53
4.3.7 <i>Model Data Warehouse</i>	55
4.3.6.1 <i>Model Customer</i>	55
4.3.6.2 <i>Model Air Transaction (Earning)</i>	56
4.3.6.3 <i>Model Non Air Transaction (Earning)</i>	57
4.3.6.4 <i>Model Certificate Transaction (Spending)</i>	58

4.3.6.5	<i>Tabel Dimensi</i>	60
4.3.6.5.1	<i>Dimensi Airport</i>	60
4.3.6.5.2	<i>Dimensi Location</i>	60
4.3.6.5.3	<i>Dimensi Gender</i>	60
4.3.6.5.4	<i>Dimensi Customer Status</i>	61
4.3.6.5.5	<i>Dimensi Member Tier Composition</i>	61
4.3.6.5.6	<i>Dimensi Booking Class</i>	61
4.3.6.5.7	<i>Dimensi Time</i>	61
4.3.6.5.8	<i>Dimensi Operating Flight Number</i>	62
4.3.6.5.9	<i>Dimensi Partner</i>	62
BAB V IMPLEMENTASI DATA WAREHOUSE		63
5.1	PRESENTASI <i>DATA WAREHOUSE</i>	63
5.2	PENYEDIAAN INFORMASI	64
5.2.1	<i>Informasi Customer</i>	65
5.2.2	<i>Informasi Earning/Accrual Traffic</i>	69
5.2.3	INFORMASI <i>REDEEM TRAFFIC</i>	70
5.3	PENGARUH PERUBAHAN	71
BAB VI IMPLEMENTASI DATA MINING		74
6.1	PEMAHAMAN TERHADAP BISNIS	74
6.2	SUMBER DATA	75
6.3	PEMILIHAN TEKNIK DAN PERANGKAT LUNAK DATA MINING	76
6.4	PERSIAPAN DATA	77
6.4.1	<i>Persiapan Data Teknik Association Rule</i>	77
6.4.2	<i>Persiapan Data Teknik Decision Tree</i>	81
6.5	ANALISA HASIL UJI COBA	82
6.5.1	<i>Analisa Hasil Uji Coba Dengan Teknik Association Rule</i>	82
6.5.2	<i>Analisa Hasil Uji Coba Dengan Teknik Decision Tree</i>	85
BAB VII KESIMPULAN		89
7.1	KESIMPULAN	89
7.2	SARAN	90
DAFTAR PUSTAKA		91
LAMPIRAN-LAMPIRAN		93
LAMPIRAN 1		93
LAMPIRAN 2		97
LAMPIRAN 3		102

LAMPIRAN 4.....	103
LAMPIRAN 5.....	104
LAMPIRAN 6.....	107
LAMPIRAN 7.....	108
LAMPIRAN 8.....	109



DAFTAR GAMBAR

Gambar 2.1 Arsitektur data <i>warehouse</i> (Ponniah, 2001).....	11
Gambar 2.2 – Contoh transformasi <i>single-field</i> (Zain, 2008)	14
Gambar 2.3 – Contoh transformasi <i>multi-field</i> (Zain, 2008)	14
Gambar 2.4 <i>Star Schema</i> (Ponniah, 2001).....	16
Gambar 2.5 <i>Snow flake Schema</i> (Ponniah, 2001)	17
Gambar 2.6 <i>Clasification</i> menggunakan <i>tree induction</i> (http://www.cse.unsw.edu.au/~billw/cs9414/notes/ml/06prop/id3/id3.html).....	20
Gambar 2.7 <i>Clasification</i> menggunakan <i>neural network</i> (Connonly and Begg, 2005)	20
Gambar 2.9 Tahap-tahap <i>data mining</i> (Zein, 2008).....	24
Gambar 3.1 Tahap-tahap metodologi penelitian.....	34
Gambar 4.1-Struktur Organisasi <i>Executive Vice President</i>	38
Gambar 4.2- Struktur Organisasi <i>Customer Relation Management</i>	39
Gambar 4.3- Arsitektur logical <i>Garuda Frequent Flyer</i>	47
Gambar 4.4- Arsitektur fisik <i>data warehouse Garuda Frequent Flyer</i>	48
Gambar 4.5- Alur transformasi transaksi penerbangan anggota GFF.....	52
Gambar 4.6- Alur transformasi tabel <i>AIRPORT</i>	53
Gambar 4.7- Skema job pada Spoon yang mengurutkan pengerjaan transformasi yang telah dibuat sebelumnya	54
Gambar 4.8- Model skema bintang <i>enrollment</i> anggota GFF.....	55
Gambar 4.9- Model skema bintang aktifitas penerbangan anggota	57
Gambar 4.10- Model skema bintang aktifitas bukan penerbangan (<i>non air activity</i>) anggota GFF.....	58
Gambar 4.11- Model skema bintang <i>redeem</i> (penggunaan <i>mileage</i>) anggota GFF	59
Gambar 5.1- Contoh report <i>Customer</i> per <i>branch office</i> pada tahun 2007	66
Gambar 5.2- Contoh report <i>Customer</i> per <i>Tier Composition</i> pada tahun 2007	67
Gambar 5.3 - Contoh report <i>Customer</i> per yang melakukan aktivasi pada tahun 2007.....	67

Gambar 5.4 Informasi <i>Customer</i> yang terdistribusi dalam beberapa <i>tier</i> yang di <i>slice</i> hanya tahun 2007 dan <i>drill-down</i> berdasar usianya.....	68
Gambar 5.5 Informasi <i>Enrolment Customer</i> yang terdistribusi dalam dalam tahun pendaftaran dilengkapi dengan <i>grafik trend enrollment</i>	68
Gambar 5.6- Earning mileage customer yang di- <i>slice</i> berdasarkan <i>Origin Destination airport</i> penerbangan	69
Gambar 5.7- Earning mileage customer yang di <i>slice</i> berdasarkan Origin Destination airport penerbangan yang dilengkap dengan grafik <i>pie chart</i>	70
Gambar 5.8- Contoh <i>report redeem activity per branch office</i> pada tahun 2007 .	71
Gambar 5.9- Langkah-langkah implementasi jenis <i>report</i> baru.	72
Gambar 6.1 <i>Building block</i> menggunakan Rapid miner Community 4.2 dengan teknik <i>data mining association rule</i>	78
Gambar 6.2 <i>Building block</i> metode <i>Decision tree</i> menggunakan <i>Rapidminer Community 4.2</i>	82
Gambar 6.3 Bagian dari diagram pohon karakter anggota <i>Cobrand</i> yang mendaftar melalui <i>channel Cobrand</i> atau <i>Citibank</i>	85
Gambar 6.4 Bagian diagram pohon karakter anggota <i>Cobrand</i> yang mendaftar melalui <i>Customer Self Service (CSS)</i>	86
Gambar 8.1 Pola anggota GFF <i>Cobrand</i> yang melakukan pendaftaran melalui jalur BO.....	102
Gambar 8.2 Pola anggota GFF <i>Cobrand</i> yang melakukan pendaftaran melalui jalur BO, CSS dan jalur <i>Citibank</i> atau disebut <i>channel Cobrand</i>	103

DAFTAR TABEL

Tabel 2.1 Tabel operasi <i>data mining</i> dan teknik yang digunakannya (Zein, 2008)	19
Tabel 3.1 Alur Pengerjaan.....	35
Tabel 4.1- <i>Platform database</i> sistem GFF operasional dan penambahan data....	49
Tabel 4.2-Perbandingan sumber data GFF dan DatabaseTemp.....	50
Tabel 4.3-Perbedaan format data	51
Tabel 4.4-Ketidak konsistenan data penerbangan anggota GFF karena duplikasi data.....	51
Tabel 6.1 Tabel sebelum transformasi	79
Tabel 6.2 Tabel setelah <i>transformasi/preprocessing</i>	80
Tabel 6.3 <i>Model Rule</i> dengan kesimpulan masa keanggotaan <i>Cobrand</i> dalam mendaftar melalui <i>channel BO</i>	83
Tabel 6.4 <i>Model Rule</i> dengan kesimpulan masa keanggotaan <i>Cobrand</i> dalam <i>range1, range2, dan range2</i>	84