

PENGUNAAN *DATA WAREHOUSE* DAN *DATA MINING* UNTUK DATA AKADEMIK SEBUAH STUDI KASUS PADA UNIVERSITAS NASIONAL

Ariana Azimah

Jurusan Sistem Informasi
Fakultas Teknologi Komunikasi & Informatika
Universitas Nasional
ariana@unas.ac.id

Yudho Giri Sucahyo

Program Studi Magister Teknologi Informasi
Fakultas Ilmu Komputer
Universitas Indonesia
yudho@cs.ui.ac.id

Abstrak

Kegiatan evaluasi, perencanaan, dan pengambilan keputusan akan dapat dilakukan dengan lebih baik jika sebuah organisasi memiliki informasi yang lengkap, cepat, tepat, dan akurat. Untuk sebuah perguruan tinggi yang sudah memiliki sistem informasi, informasi yang dibutuhkan dapat diekstrak dari data operasional yang tersimpan dalam database yang terintegrasi. Penelitian ini mengkaji ekstraksi data operasional ke dalam sebuah data warehouse untuk kemudian dilanjutkan dengan kegiatan analisis data menggunakan teknik data mining. Hasil dari penelitian ini adalah sebuah data warehouse lengkap dengan aplikasi pelaporan informasi berbasis web. Selanjutnya dilakukan analisa data menggunakan teknik data mining terhadap data warehouse yang sudah dibangun. Hasil dari penerapan data mining adalah dihasilkannya pola karakteristik mahasiswa yang mengambil program peminatan tertentu.

Keywords-Data Warehouse, Data Mining, Data Akademik

I. PENDAHULUAN

Perguruan tinggi saat ini dituntut untuk memiliki keunggulan bersaing dengan memanfaatkan semua sumber daya yang dimiliki. Selain sumber daya sarana, prasarana, dan manusia, sistem informasi adalah salah satu sumber daya yang dapat digunakan untuk meningkatkan keunggulan bersaing. Sistem informasi dapat digunakan untuk mendapatkan, mengolah, dan menyebarkan informasi untuk menunjang kegiatan operasional sehari-hari sekaligus menunjang kegiatan pengambilan keputusan strategis. Hambatan yang sering ditemui adalah kenyataan bahwa sistem informasi yang ada belum terintegrasi secara sempurna dan mengakibatkan informasi yang tersaji kurang lengkap atau kurang akurat.

Pembangunan *data warehouse* merupakan salah satu cara untuk mengekstrak informasi penting dari data yang tersebar di beberapa sistem informasi [1]. Data yang sudah terintegrasi selanjutnya dapat dimanfaatkan untuk kegiatan penyampaian informasi yang dapat ditinjau dari berbagai dimensi dan dapat diatur tingkatan rinciannya. Pemanfaatan lebih lanjut dari informasi yang ada dalam *data warehouse* adalah kegiatan analisa data dengan menggunakan teknik-teknik *data mining*.

Dalam penelitian ini, dibangun sebuah *data warehouse* untuk Universitas Nasional (UNAS). UNAS adalah Perguruan Tinggi Swasta (PTS) di Jakarta yang didirikan pada tanggal 15 Oktober 1949. UNAS memiliki tujuh Fakultas untuk jenjang pendidikan Strata Satu (S1) dengan 17 Program Studi, tiga Program Studi untuk jenjang pendidikan Magister (S2), serta tiga Akademi dengan enam Program Studi Diploma Tiga (DIII).

UNAS menyadari bahwa untuk memanfaatkan data yang ada di dalam sistem informasi untuk menunjang kegiatan pengambilan keputusan, tidak cukup hanya mengandalkan data operasional saja, diperlukan suatu analisis data untuk menggali potensi-potensi informasi yang ada. Untuk itulah perlu dibangun sebuah *data warehouse* yang selanjutnya dapat dianalisis menggunakan teknik *data mining*. Penggunaan teknik *data mining* diharapkan dapat memberikan pengetahuan-pengetahuan yang tadinya tersembunyi di dalam gudang data.

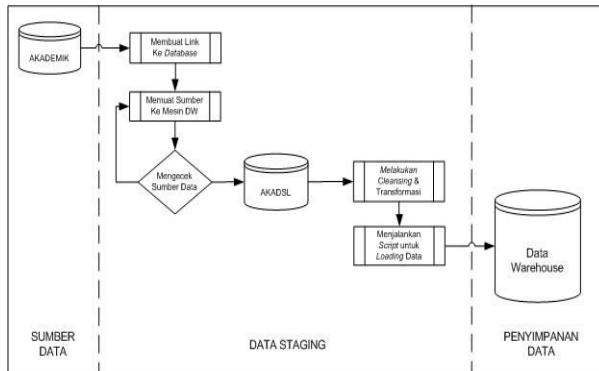
II. PERANCANGAN *DATA WAREHOUSE*

Tahap perancangan ini meliputi perancangan arsitektur *data warehouse*, perencanaan sumber data, serta pemodelan data dimensional.

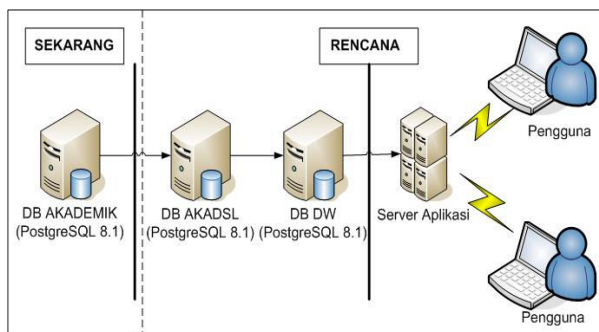
2.1. Perancangan Arsitektur *Data Warehouse*

Sumber data operasional yang digunakan adalah *database AKADEMIK* yang merupakan *database* berisi semua data akademik UNAS. Dari sumber data tersebut, secara periodik dilakukan pemilihan data dan selanjutnya dimuatkan ke *database* terpisah yaitu AKADSL agar *database* operasional tidak terganggu kinerja sistemnya. Dalam teori *data warehouse*, AKADSL disebut juga sebagai *data staging*. Dalam proses pemuatan data ke *database* AKADSL, selalu dilakukan proses pengecekan terlebih dahulu untuk memastikan bahwa data *valid* dan dapat dimuat ke *database* AKADSL. Setelah itu dilakukan proses pembersihan dan transformasi data pada *database* AKADSL. Hasil dari proses pembersihan dan transformasi inilah yang kemudian disimpan ke dalam *data warehouse*. Gambar 1 memperlihatkan rancangan arsitektur *logical* dari *data warehouse* UNAS yang sekaligus menggambarkan proses pengisian data ke *data warehouse*.

Rancangan arsitektur fisik dari *data warehouse* UNAS dapat dilihat pada Gambar 2. Pengguna mengakses *data warehouse* melalui *server* aplikasi. *Database* AKADSL dan *data warehouse* berada pada mesin yang sama sehingga proses Ekstraksi, Transformasi dan Loading dilakukan di mesin *data warehouse* dan tidak mengganggu kerja mesin operasional. Baik *database* AKADEMIK, AKADSL, maupun *data warehouse* menggunakan RDBMS PostgreSQL 8.1.



Gambar 1. Arsitektur Logical DW UNAS



Gambar 2. Arsitektur Fisik DW UNAS.

2.2. Perencanaan Sumber Data

UNAS telah mengembangkan beberapa aplikasi untuk menunjang pelaksanaan kegiatan akademis yang dapat digunakan baik oleh mahasiswa, dosen maupun staf administrasi. Aplikasi-aplikasi yang dikembangkan merupakan aplikasi yang berbasis Intranet (hanya dapat diakses di lingkungan kampus dengan teknik *client-server*) dan juga aplikasi yang berbasis *web* yang dapat diakses dari Internet. Aplikasi-aplikasi tersebut adalah:

- Sistem Informasi Akademik (SIA) yang merupakan aplikasi akademik yang utama. Dengan aplikasi ini dosen dapat melakukan pengisian nilai secara *online*.
- Layanan Akademik Mahasiswa merupakan aplikasi yang digunakan oleh mahasiswa. Dengan aplikasi ini mahasiswa dapat melakukan pengisian KRS (Kartu Rencana Studi) secara *online*. Selain itu juga dapat melihat nilai dan jadwal kuliah dari masing-masing mata kuliah.
- Aplikasi Pendaftaran Mahasiswa Baru yang merupakan aplikasi untuk menangani proses penerimaan mahasiswa baru.
- *Visual Scheduling Sistem* (VSS) yang merupakan aplikasi untuk mengelola jadwal ruangan untuk tiap fakultas di setiap semester.

2.3. Pemodelan Data Dimensional

Skema yang digunakan untuk pemodelan data adalah *star schema* dimana terdapat satu tabel fakta dan beberapa tabel dimensi. Penggunaan *star schema* memungkinkan proses *query* yang lebih ringan dan memudahkan penjelajahan terhadap data dimensinya [2]. Tabel fakta yang terbentuk dari perancangan *data warehouse* ini merupakan tabel yang berhubungan dengan evaluasi kegiatan akademik mahasiswa mulai dari kegiatan pendaftaran mahasiswa baru sampai kelulusan mahasiswa. Tabel-tabel fakta yang ada dalam *data warehouse* UNAS adalah sebagai berikut:

Tabel *camaba_fact*

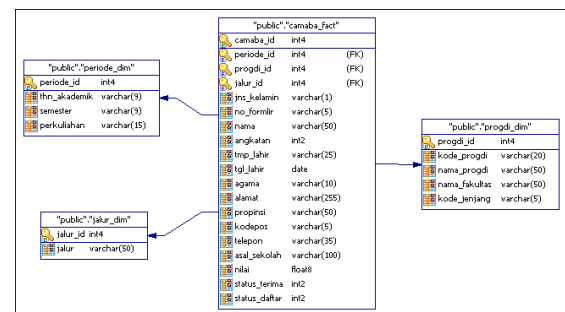
Tabel ini berisi semua data yang berhubungan dengan penerimaan mahasiswa baru yang meliputi data pribadi, nilai yang diperoleh, status penerimaan dan status pendaftaran calon mahasiswa. *Star schema* untuk penerimaan mahasiswa baru dapat dilihat pada gambar 3.

Tabel *nilaimhs_fact*

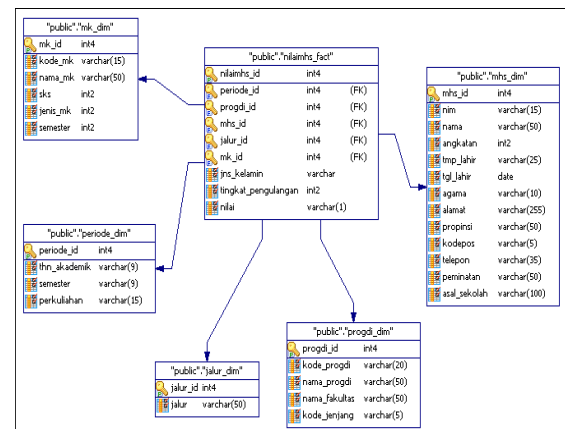
Tabel ini berisi data nilai yang telah diperoleh mahasiswa yang terdiri dari matakuliah yang dipilih, tingkat pengulangan, dan nilai yang diperoleh. *Star schema* untuk Nilai Mahasiswa dapat dilihat pada gambar 4.

Tabel *hasilstudy_fact*

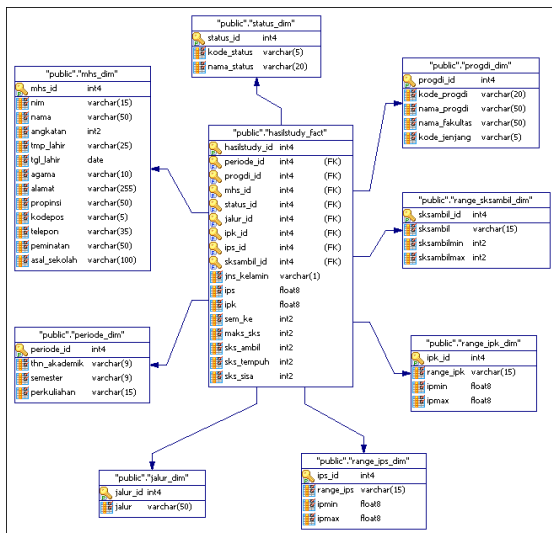
Tabel ini berisi catatan akademik mahasiswa yang terdiri dari indeks prestasi semester (IPS), indeks prestasi kumulatif (IPK), sistem kredit semester (SKS) yang diambil dan SKS yang sudah ditempuh pada periode tersebut. Selain itu juga disimpan data status akademik mahasiswa pada periode tersebut. *Star Schema* untuk Catatan Akademik Mahasiswa dapat dilihat pada gambar 5.



Gambar 3. Star Schema Mahasiswa Baru.



Gambar 4. Star Schema Nilai Mahasiswa.



Gambar 5. Star Schema Catatan Akademik Mahasiswa.

Tabel dimensi merupakan tabel yang berisi data yang menunjukkan tinjauan dari berbagai perspektif. Adanya dimensi memungkinkan pembuatan laporan dinamis yang bisa ditinjau dari berbagai perspektif sekaligus tingkat perinciannya bisa diatur. Tabel-tabel dimensi yang ada dalam *data warehouse* UNAS adalah sebagai berikut:

Tabel periode_dim

Data yang termasuk dimensi periode adalah tahun akademik, semester (ganjil dan genap) dan perkuliahan (reguler dan pendek).

Tabel mhs_dim

Data yang termasuk dalam dimensi mahasiswa adalah data pribadi seperti nama, tanggal lahir, alamat dan lain sebagainya.

Tabel prodi_dim

Dimensi jurusan mengandung data tentang jenjang pendidikan, fakultas dan jurusan, dimana setiap fakultas terdiri dari satu jurusan atau lebih. Jenjang pendidikan terdiri dari D3 (Diploma III), S1 (Sarjana) dan S2 (Magister).

Tabel status_dim

Dimensi status menunjukkan informasi tentang status akademik yang terdiri dari Aktif, Cuti, Lulus, Cuti Tidak Lapor, Mengundurkan Diri dan Drop Out.

Tabel jalur_dim

Dimensi jalur menunjukkan informasi tentang jalur masuk. Ada dua Jalur penerimaan mahasiswa baru di UNAS yaitu jalur reguler dan jalur pindahan.

Tabel mk_dim

Dimensi matakuliah menunjukkan informasi tentang data matakuliah yang terdiri dari nama matakuliah, kode matakuliah, sks dan jenis matakuliah.

Tabel range_ipk_dim

Data yang termasuk range IPK adalah IPK antara 0 dan 2.74, IPK antara 2.75 dan 3.00, IPK antara 3.01 dan 3.50, IPK antara 3.51 dan 4.00.

Tabel range_ips_dim

Data yang termasuk Range IPS adalah IPS antara 0 dan 2.74, IPS antara 2.75 dan 3.00, IPS antara 3.01 dan 3.50, IPS antara 3.51 dan 4.00.

Tabel range_sks_dim

Data yang termasuk Range SKS adalah SKS antara 0-18, SKS antara 19-21, SKS antara 22-26.

III. KAPABILITAS DATA WAREHOUSE

Kemampuan penyediaan informasi dari *data warehouse* ini digunakan untuk melakukan analisa terhadap kegiatan akademik yang berlangsung di UNAS baik dari pendaftaran sampai kelulusan mahasiswa sehingga pimpinan universitas dapat menentukan langkah apa saja yang perlu diambil untuk meningkatkan kualitas kegiatan akademik.

3.1. Presentasi Data Warehouse

Kemampuan *data warehouse* menyediakan informasi kepada pengguna merupakan hal terpenting dari *data warehouse*. Platform database yang digunakan adalah PostgreSQL. Sedangkan tools yang digunakan untuk presentasi data kepada pengguna adalah:

- Mondrian, merupakan OLAP engine yang menggunakan bahasa pemrograman Java [3].
- Jpivot, merupakan JSP (Java Server Pages) yang dapat menampilkan OLAP table dan chart.

Media yang digunakan untuk mempresentasikan data adalah menggunakan aplikasi berbasis web dengan bahasa pemrograman Java.

3.2. Penyediaan Informasi

Data warehouse UNAS dibangun agar dapat menyediakan informasi yang cepat, tepat dan akurat sehingga pimpinan universitas dapat mengambil keputusan yang tepat dalam menjalankan proses kegiatan akademik.

Karena menggunakan aplikasi berbasis web maka informasi yang tersedia ini dapat diakses dari komputer manapun yang mempunyai koneksi ke Internet. Selain itu kemampuan lain yang cukup bermanfaat dari adanya *data warehouse* adalah:

- Kemampuan *roll-up* dan *drill-down* untuk memudahkan pemetaan data dan mempertajam analisis. *Roll-up* adalah kemampuan untuk menampilkan data dengan tingkat rincian yang lebih rendah. *Drill-down* adalah kemampuan menampilkan data dengan tingkat rincian yang lebih tinggi.
- Kemampuan membuat *query* sendiri sesuai dengan kebutuhan.
- Kemampuan *report customization* sesuai dengan kebutuhan informasi.
- Kemampuan untuk membuat *chart* atau grafik sesuai dengan laporan yang diinginkan.
- Kemampuan membuat laporan yang kemudian dapat disimpan dalam format Excel dan PDF.

Informasi Penerimaan Mahasiswa Baru

Informasi tentang penerimaan mahasiswa baru digunakan untuk melihat:

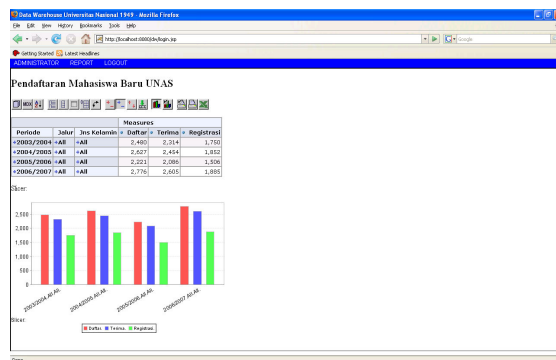
1. Jumlah calon mahasiswa yang mendaftar, yang diterima, dan yang melakukan registrasi untuk tiap jurusan per periode.
2. Jurusan mana yang mengalami penurunan jumlah calon mahasiswa yang mendaftar, diterima dan registrasi. Sehingga dapat dicari penyebabnya dan dicari solusinya.
3. Komposisi jumlah mahasiswa baru pada periode tertentu.

Contoh tampilan informasi pendaftaran mahasiswa baru dapat dilihat pada gambar 6 dan gambar 7.

The screenshot shows a web application titled 'Pendaftaran Mahasiswa Baru UNAS'. It features a table with columns: Jenjang, Periode, Jumlah, and Measures (Jns Kalamis, Daftar, Tertima, Registrasi). The table lists data for various academic levels (S1, S2) and periods (2003/2004 to 2006/2007). Below the table, there is a bar chart titled 'Bar:' showing the distribution of data across different categories.

Jenjang	Periode	Jumlah	Jns Kalamis	Daftar	Tertima	Registrasi
S1	2003/2004	10,118	9,471	7,001		
	2004/2005	1,422	1,422	1,422		
	2005/2006	440	440	343		
	2006/2007	704	666	511		
S2	2003/2004	8,058	7,529	5,450		
	2004/2005	1,966	1,925	1,343		
	2005/2006	2,027	1,901	1,427		
	2006/2007	2,246	2,101	1,423		

Gambar 6. Laporan Pendaftaran Mahasiswa Baru.



Gambar 7. Grafik Pendaftaran Mahasiswa Baru.

Informasi Status Akademik Mahasiswa

Status akademik terdiri dari Aktif, Cuti, Lulus, Cuti Tidak Lapor, Mengundurkan Diri, dan Drop Out. Informasi tentang status akademik ini sangat diperlukan untuk melihat komposisi jumlah mahasiswa berdasarkan status akademiknya dalam satu periode tertentu.

Informasi Distribusi Jumlah SKS Ambil Mahasiswa

Jumlah SKS yang dapat diambil oleh seorang mahasiswa tergantung dari IPS (Indeks Prestasi Semester) yang diperoleh di semester sebelumnya. Sehingga semakin bagus IPS mahasiswa tersebut semakin banyak juga jumlah SKS yang bisa diambil, semakin banyak SKS

yang sudah diambil maka semakin cepat juga masa studi mahasiswa tersebut. Informasi tentang distribusi jumlah SKS ambil digunakan untuk melihat banyaknya mahasiswa yang mengikuti perkuliahan berdasarkan range jumlah SKS yang diambil. Selain itu juga digunakan untuk melihat rata-rata jumlah SKS yang diambil mahasiswa dalam satu periode tertentu.

Informasi Distribusi Indeks Prestasi Semester (IPS) Mahasiswa

IPS merupakan salah satu indikator keberhasilan mahasiswa dalam satu semester. Semakin banyak mahasiswa yang mendapat IPS yang bagus dalam suatu jurusan, semakin baik penilaian terhadap jurusan tersebut. Informasi tentang distribusi IPS mahasiswa digunakan oleh pimpinan UNAS untuk melihat jumlah mahasiswa berdasarkan range IPS yang sudah ditentukan. Selain itu juga dapat digunakan untuk melihat rata-rata IPS mahasiswa dalam satu periode tertentu.

Informasi Distribusi Indeks Prestasi Kumulatif (IPK) Mahasiswa

IPK merupakan salah satu indikator keberhasilan studi mahasiswa selama periode yang sudah ditempuh. IPK dihitung berdasarkan nilai yang diperoleh selama masa studi mahasiswa tersebut. Informasi tentang distribusi IPK mahasiswa digunakan oleh pimpinan universitas untuk melihat jumlah mahasiswa berdasarkan range IPK yang sudah ditetapkan. Selain itu juga digunakan untuk melihat rata-rata IPK yang diperoleh mahasiswa.

Informasi Profil Lulusan

Salah satu kriteria laporan akademik yang menjadi kebutuhan universitas adalah profil lulusan. Informasi tentang profil lulusan digunakan untuk melihat:

1. Profil lulusan berdasarkan jenjang studi dalam satu periode tertentu.
2. Jumlah mahasiswa lulusan dalam satu periode tertentu. Semakin banyak jumlah lulusan yang tepat waktu maka semakin baik penilaian terhadap jurusan tersebut.
3. Jumlah mahasiswa lulusan berdasarkan masa studinya.

Informasi Distribusi IPK Lulusan

Informasi mengenai distribusi IPK lulusan digunakan oleh pimpinan universitas untuk melihat:

1. Kualitas dari mahasiswa lulusan.
2. Jumlah mahasiswa lulusan berdasarkan range IPK yang sudah ditentukan.
3. Rata-rata IPK dari mahasiswa lulusan.

3.3. Pengaruh Perubahan

Dengan adanya *data warehouse*, proses penyusunan laporan menjadi lebih sederhana, karena pengguna bisa melakukan *customization report* sesuai dengan yang diinginkan, sehingga tercipta efisiensi waktu dari yang sebelumnya memerlukan waktu satu bulan untuk membuat program lagi atau satu minggu untuk mengerjakan secara manual menjadi hanya satu hari. Dengan demikian UNAS dapat mengurangi jumlah sumber daya manusia yang sebelumnya dialokasikan

untuk melakukan penyusunan laporan untuk lebih fokus pada hal-hal lain yang lebih penting untuk meningkatkan kualitas kegiatan akademik. Selain itu *data warehouse* ini juga menyimpan data *history* status akademik mahasiswa sehingga laporan yang dahulu sulit disediakan menjadi lebih mudah.

Pemenuhan kebutuhan laporan ini juga menjadi lebih fleksibel untuk para pimpinan universitas karena dapat diakses melalui Internet sehingga meskipun tidak bisa datang ke kampus, pimpinan tetap bisa mendapatkan laporan untuk mendukung proses pengambilan keputusan.

IV. PENERAPAN TEKNIK DATA MINING

Pada penelitian ini penerapan teknik *data mining* dilakukan pada data akademik yang diekstrak dari *data warehouse*. Proses *cleansing* tidak perlu dilakukan karena sudah dilakukan saat pembentukan *data warehouse*. Dengan melihat pada kondisi tabel fakta dan dimensi yang telah dibentuk di tahap sebelumnya, fokus pada penelitian ini adalah "Bagaimana karakteristik mahasiswa yang mengambil program peminatan tertentu?"

Adanya pengetahuan tersebut diharapkan dapat membantu mahasiswa yang mengalami kesulitan dalam menentukan program peminatan mana yang harus dipilih. Dengan mengetahui pola karakteristik mahasiswa yang berhasil dari suatu program peminatan, dosen pembimbing dapat memberikan saran yang lebih tepat dalam pemilihan program peminatan sesuai dengan minat dan karakteristik mahasiswa.

4.1. Pemilihan Teknik Data Mining

Teknik *data mining* yang digunakan untuk menjawab pertanyaan diatas adalah teknik *classification* dan *association rules* [4]. Dengan menggunakan kedua teknik diatas akan dapat dilihat kelompok pola mahasiswa yang mengambil program peminatan tertentu. Untuk kemudian ditelusur lebih rinci fakta-fakta pendukungnya.

Perangkat lunak yang digunakan untuk melakukan proses *data mining* pada penelitian ini adalah:

1. CBA. *Data mining tool* yang dikembangkan oleh School of Computing, National University of Singapore [5]. Algoritma yang digunakan adalah *Classification Based Association*.
2. Microsoft SQL Server 2005. Algoritma yang digunakan adalah *Association Rules*.

4.2. Pemodelan Data Mining

Pemodelan *data mining* yang dirancang pada penelitian ini menyesuaikan dengan kebutuhan informasi dan karakteristik data di dalam *data warehouse*. Program peminatan yang dipilih pada penelitian ini adalah program peminatan pada program studi Manajemen Perusahaan UNAS. Karena penelitian ini bertujuan untuk mencari pola karakteristik dari mahasiswa yang berhasil dari suatu program peminatan, maka *dataset* yang digunakan adalah data mahasiswa program studi Manajemen Perusahaan yang berstatus lulus dengan IPK lebih besar atau sama dengan 2.75. Alasan pemilihan

program peminatan tersebut adalah karena datanya mempunyai jumlah yang paling banyak dan lengkap dibanding dengan program peminatan pada program studi yang lain.

Program studi Manajemen Perusahaan UNAS mempunyai empat program peminatan yaitu:

1. Manajemen Keuangan (MK).
2. Manajemen Pemasaran (MP).
3. Manajemen Sumber Daya Manusia (MSDM).
4. Manajemen Operasional (MO).

Dari keempat program studi tersebut, program peminatan MO tidak diikutsertakan karena jumlah datanya sedikit (hanya 33 data). Total jumlah data yang digunakan pada penelitian ini adalah 701 data.

Karakteristik data yang digunakan untuk proses *data mining* terlihat di gambar 8. Dari data tersebut kemudian dicari hubungan antara nilai-nilai matakuliah yang diambil dengan peminatan yang dipilihnya sehingga didapatkan pola karakteristik mahasiswa yang mengambil suatu program peminatan tertentu. Oleh karena itu program peminatan dipilih sebagai *output* sedangkan nilai-nilai matakuliah dipilih sebagai *input*.

No	Deskripsi	Type	Ket
1	Program Peminatan	Discrete	Output
2	Bahasa Inggris I	Discrete	Input
3	Matematika Ekonomi	Discrete	Input
4	Pengantar Akuntansi I	Discrete	Input
5	Pengantar Bisnis	Discrete	Input
6	Pengantar Manajemen	Discrete	Input
7	Teori Ekonomi Mikro I	Discrete	Input
8	Bahasa Inggris II	Discrete	Input
9	Manajemen SDM I	Discrete	Input
10	Pengantar Akuntansi II	Discrete	Input
11	Pengantar Aplikasi Komputer	Discrete	Input
12	Statistik I	Discrete	Input
13	Teori Ekonomi Makro I	Discrete	Input
14	Teori Ekonomi Mikro II	Discrete	Input
15	Akuntansi Biaya	Discrete	Input
16	Aspek Hukum dalam Ekonomi	Discrete	Input
17	Bahasa Inggris Bisnis I	Discrete	Input
18	Bank dan Lembaga Keuangan Lainnya	Discrete	Input
19	Manajemen Pemasaran I	Discrete	Input
20	Manajemen SDM II	Discrete	Input
21	Statistik II	Discrete	Input
22	Teori Ekonomi Makro II	Discrete	Input
23	Aplikasi Komputer Akuntansi	Discrete	Input
24	Bahasa Inggris Bisnis II	Discrete	Input
25	Ekonomi Manajerial	Discrete	Input
26	Manajemen Keuangan I	Discrete	Input
27	Manajemen Operasional I	Discrete	Input
28	Manajemen Pemasaran II	Discrete	Input
29	Riset Operasi	Discrete	Input
30	Sistem Informasi Akuntansi	Discrete	Input
31	Ekonomi Pembangunan	Discrete	Input
32	Manajemen Resiko/Asuransi	Discrete	Input
33	Ekonomi Internasional	Discrete	Input
34	Perpajakan	Discrete	Input
35	Perilaku Organisasi	Discrete	Input
36	Manajemen Operasional II	Discrete	Input
37	Bahasa Inggris Bisnis III	Discrete	Input
38	Manajemen Keuangan II	Discrete	Input

Gambar 8. Model Data Mining.

4.3. Analisa Hasil Uji Coba

Pada uji coba ini *minimum support* yang digunakan adalah 10%. Hasil penerapan teknik *classification* dengan menggunakan CBA dapat dilihat pada gambar 9 sedangkan hasil penerapan teknik *association rules* dengan menggunakan Microsoft SQL Server dapat dilihat pada gambar 10.

Dari hasil percobaan ditemukan karakteristik mahasiswa yang mengambil program peminatan sebagai berikut:

1. Pola karakteristik mahasiswa yang mengambil program peminatan MP antara lain:
 - Manajemen Resiko/Asuransi=B dan Pengantar Akuntansi II=C
 - Ekonomi Manajerial=B dan Teori Ekonomi Mikro II=C
 - Manajemen Resiko/Asuransi=B dan Manajemen Operasional I=B dan Manajemen Pemasaran I=B
 - Manajemen Keuangan II=B dan Pengantar Akuntansi I=C
 - Manajemen Pemasaran I=B dan Riset Operasi =B
 - Teori Ekonomi Mikro II=C dan Manajemen Pemasaran II=B

Rule	Cover %	Conf %	Cover Count	Sup Count	Sup %
Manajemen Resiko/Asuransi = B Pengantar Akuntansi II = C -> class = MP	18.26%	56.25 %	128	72	10.27 %
Ekonomi Manajerial = B Teori Ekonomi Mikro II = C ->class = MP	21.39%	55.33 %	150	83	11.84 %
Manajemen Resiko/Asuransi = B Manajemen Operasional I = B Manajemen Pemasaran I = B -> class = MP	22.25%	55.12 %	156	86	12.26 %
Manajemen Resiko/Asuransi = B Pengantar Akuntansi I = C -> class = MP	19.82%	54.67 %	139	76	10.84 %
Manajemen Keuangan II = B Pengantar Akuntansi I = C -> class = MP	19.25%	54.07 %	135	73	10.41 %
Teori Ekonomi Mikro II = C Pengantar Aplikasi Komputer = B -> class = MP	20.54%	53.47 %	144	77	10.98 %
Statistik II = B Aspek Hukum Dalam Ekonomi = B Bahasa Inggris II = B -> class = MP	19.54%	52.55 %	137	72	10.27 %
Manajemen Resiko/Asuransi = B Bahasa Inggris Bisnis III = B Bank & Lembaga Keuangan Lainnya = B -> class = MP	21.54%	52.31 %	151	79	11.27 %
Pengantar Akuntansi II = A -> class = MK	24.67%	50.28 %	173	87	12.41 %
Teori Ekonomi Mikro II = B Pengantar Manajemen = A ->class = MK	22.11%	49.67 %	155	77	10.98 %
Manajemen Keuangan I = B Pengantar Bisnis = A ->class = MK	23.82%	47.30 %	167	79	11.27 %
Akuntansi Biaya = A -> class = MK	24.82%	47.12 %	174	82	11.69 %
Riset Operasi = A -> class = MK	22.11%	47.09 %	155	73	10.41 %
Bahasa Inggris Bisnis II = B Manajemen Operasional I = A -> class = MK	24.82%	45.40 %	174	79	11.27 %

Gambar 9. Hasil Uji Coba menggunakan CBA.

Probability	Importance	Rule
0.503	0.255	Pengantar Akuntansi II = A -> Peminatan = MK
0.497	0.236	Pengantar Manajemen = A, Teori Ekonomi Mikro II = B -> Peminatan = MK
0.471	0.212	Akuntansi Biaya = A -> Peminatan = MK
0.473	0.211	Pengantar Bisnis = A, Manajemen Keuangan I = B -> Peminatan = MK
0.467	0.205	Bank Dan Lembaga Keuangan Lainnya = A, Teori Ekonomi Mikro II = B -> Peminatan = MK
0.469	0.203	Bank Dan Lembaga Keuangan Lainnya = A, Bahasa Inggris I = B -> Peminatan = MK
0.471	0.203	Riset Operasi = A -> Peminatan = MK
0.459	0.198	Pengantar Akuntansi I = A -> Peminatan = MK
0.439	0.189	Manajemen Operasional I = A -> Peminatan = MK
0.454	0.188	Manajemen Operasional I = A, Bahasa Inggris Bisnis II = B -> Peminatan = MK
0.455	0.182	Pengantar Manajemen = A, Matematika Ekonomi = B -> Peminatan = MK
0.426	0.180	Teori Ekonomi Mikro II = B, Bahasa Inggris Bisnis I = B -> Peminatan = MK
0.442	0.177	Pengantar Manajemen = A, Bahasa Inggris Bisnis I = B -> Peminatan = MK
0.423	0.177	Bahasa Inggris I = B, Manajemen Keuangan I = B -> Peminatan = MK
0.424	0.173	Pengantar Bisnis = A -> Peminatan = MK
0.422	0.172	Teori Ekonomi Mikro II = B, Manajemen Keuangan I = B -> Peminatan = MK
0.440	0.172	Pengantar Bisnis = A, Perpajakan = A -> Peminatan = MK
0.438	0.171	Pengantar Bisnis = A, Bahasa Inggris Bisnis II = B -> Peminatan = MK
0.415	0.169	Pengantar Manajemen = A -> Peminatan = MK
0.438	0.167	Pengantar Manajemen = A, Bahasa Inggris I = B -> Peminatan = MK
0.427	0.164	Teori Ekonomi Mikro II = B, Statistik II = B -> Peminatan = MK
0.439	0.163	Pengantar Manajemen = A, Statistik II = B -> Peminatan = MK
0.435	0.158	Bank Dan Lembaga Keuangan Lainnya = A, Manajemen SDM I = B -> Peminatan = MK
0.428	0.154	Sistem Informasi Manajemen = A -> Peminatan = MK
0.426	0.151	Pengantar Manajemen = A, Ekonomi Pembangunan = B -> Peminatan = MK
0.427	0.150	Pengantar Manajemen = A, Manajemen Keuangan I = B -> Peminatan = MK
0.429	0.149	Aplikasi Komputer Akuntansi = A -> Peminatan = MK
0.425	0.149	Pengantar Akuntansi I = B, Teori Ekonomi Mikro II = B -> Peminatan = MK
0.404	0.149	Manajemen Operasional II = B, Manajemen Keuangan I = B -> Peminatan = MK
0.423	0.148	Ekonomi Manajerial = A -> Peminatan = MK
0.427	0.148	Pengantar Bisnis = A, Bahasa Inggris Bisnis I = B -> Peminatan = MK
0.425	0.144	Pengantar Bisnis = A, Ekonomi Pembangunan = B -> Peminatan = MK
0.416	0.144	Pengantar Manajemen = A, Bahasa Inggris Bisnis II = B -> Peminatan = MK
0.424	0.142	Pengantar Manajemen = A, Manajemen SDM I = B -> Peminatan = MK
0.412	0.141	Teori Ekonomi Mikro II = B, Matematika Ekonomi = B -> Peminatan = MK
0.414	0.141	Manajemen Pemasaran I = A -> Peminatan = MK
0.411	0.139	Bank Dan Lembaga Keuangan Lainnya = A, Bahasa Inggris Bisnis II = B -> Peminatan = MK
0.421	0.139	Bank Dan Lembaga Keuangan Lainnya = A, Statistik II = B -> Peminatan = MK
0.415	0.138	Pengantar Manajemen = A, Perpajakan = A -> Peminatan = MK
0.419	0.136	Pengantar Manajemen = A, Pengantar Aplikasi Komputer = B -> Peminatan = MK
0.415	0.135	Bank Dan Lembaga Keuangan Lainnya = A, Manajemen Keuangan I = B -> Peminatan = MK
0.404	0.135	Matematika Ekonomi = B, Manajemen Keuangan I = B -> Peminatan = MK
0.406	0.133	Teori Ekonomi Mikro II = B, Bahasa Inggris I = B -> Peminatan = MK
0.405	0.132	Manajemen SDM II = A -> Peminatan = MK
0.415	0.131	Pengantar Bisnis = A, Manajemen Keuangan II = B -> Peminatan = MK
0.409	0.131	Manajemen Pemasaran II = A -> Peminatan = MK
0.410	0.130	Bank Dan Lembaga Keuangan Lainnya = A, Ekonomi Pembangunan = B -> Peminatan = MK
0.403	0.130	Teori Ekonomi Mikro II = B, Manajemen SDM I = B -> Peminatan = MK
0.414	0.128	Pengantar Manajemen = A, Sistem Informasi Manajemen = B -> Peminatan = MK
0.406	0.128	Teori Ekonomi Mikro I = B, Bahasa Inggris Bisnis I = B -> Peminatan = MK
0.412	0.128	Pengantar Manajemen = A, Ekonomi Internasional = B -> Peminatan = MK
0.406	0.125	Manajemen Resiko Asuransi = A -> Peminatan = MK
0.405	0.124	Pengantar Akuntansi I = B, Manajemen Keuangan I = B -> Peminatan = MK
0.404	0.120	Teori Ekonomi Mikro I = B, Teori Ekonomi Mikro II = B -> Peminatan = MK
0.401	0.111	Pengantar Manajemen = A, Bahasa Inggris Bisnis III = B -> Peminatan = MK
0.563	0.157	Pengantar Akuntansi II = C, Manajemen Resiko Asuransi = B -> Peminatan = MP
0.553	0.155	Teori Ekonomi Mikro II = C, Ekonomi Manajerial = B -> Peminatan = MP
0.547	0.144	Pengantar Akuntansi I = C, Manajemen Resiko Asuransi = B -> Peminatan = MP
0.541	0.137	Pengantar Akuntansi I = C, Manajemen Keuangan II = B -> Peminatan = MP
0.532	0.133	Teori Ekonomi Mikro II = C, Manajemen Pemasaran II = B -> Peminatan = MP
0.535	0.133	Teori Ekonomi Mikro II = C, Pengantar Aplikasi Komputer = B -> Peminatan = MP
0.523	0.128	Teori Ekonomi Mikro I = C -> Peminatan = MP
0.515	0.125	Pengantar Akuntansi I = C -> Peminatan = MP
0.521	0.124	Teori Ekonomi Mikro I = C, Ekonomi Manajerial = B -> Peminatan = MP
0.521	0.123	Teori Ekonomi Mikro II = C, Riset Operasi = B -> Peminatan = MP
0.502	0.122	Manajemen Pemasaran I = B, Manajemen Resiko Asuransi = B -> Peminatan = MP
0.526	0.121	Teori Ekonomi Mikro II = C, Aspek Hukum Dalam Ekonomi = B -> Peminatan = MP
0.526	0.121	Teori Ekonomi Mikro II = C, Manajemen Pemasaran I = B -> Peminatan = MP
0.515	0.117	Pengantar Akuntansi II = C -> Peminatan = MP
0.522	0.117	Teori Ekonomi Mikro I = C, Bahasa Inggris Bisnis II = B -> Peminatan = MP
0.522	0.117	Teori Ekonomi Mikro II = C, Manajemen Keuangan I = B -> Peminatan = MP
0.500	0.114	Teori Ekonomi Mikro II = C -> Peminatan = MP
0.512	0.114	Teori Ekonomi Mikro II = C, Manajemen Resiko Asuransi = B -> Peminatan = MP
0.518	0.113	Teori Ekonomi Mikro II = C, Bahasa Inggris II = B -> Peminatan = MP
0.488	0.110	Manajemen Operasional I = B, Manajemen Resiko Asuransi = B -> Peminatan = MP
0.514	0.110	Teori Ekonomi Mikro II = C, Ekonomi Internasional = B -> Peminatan = MP
0.492	0.108	Manajemen Pemasaran I = B, Riset Operasi = B -> Peminatan = MP
0.511	0.105	Teori Ekonomi Mikro II = C, Perpajakan = A -> Peminatan = MP
0.507	0.103	Pengantar Akuntansi I = C, Bahasa Inggris Bisnis II = B -> Peminatan = MP
0.491	0.103	Statistik II = B, Bahasa Inggris II = B -> Peminatan = MP
0.503	0.102	Teori Ekonomi Mikro I = C, Manajemen Operasional I = B -> Peminatan = MP
0.507	0.101	Pengantar Akuntansi I = C, Aplikasi Komputer Akuntansi = B -> Peminatan = MP

Gambar 10. Hasil Uji Coba menggunakan Microsoft SQL Server.

2. Pola karakteristik mahasiswa yang mengambil program peminatan MK antara lain :

- Pengantar Akuntansi I = A
- Pengantar Akuntansi II = A
- Akuntansi Biaya = A

- Manajemen Keuangan I = B dan Pengantar Bisnis = A
- Pengantar Manajemen = A dan Manajemen Keuangan I = B

Berdasarkan pola-pola diatas didapatkan fakta bahwa mahasiswa yang mempunyai karakteristik nilai matakuliah Pengantar Akuntansi=A akan cenderung memilih program peminatan MK, sedangkan mahasiswa yang mempunyai nilai matakuliah Pengantar Akuntansi=C akan cenderung memilih program peminatan MP. Hal ini menunjukkan hubungan yang relevan mengingat matakuliah Pengantar Akuntansi merupakan salah satu matakuliah yang menjadi dasar pada program peminatan MK.

Fakta lainnya, mahasiswa yang mempunyai karakteristik nilai matakuliah Akuntansi Biaya=A cenderung memilih program peminatan MK. Hal ini menunjukkan hubungan yang relevan antara nilai matakuliah dengan program peminatan yang dipilih karena matakuliah Akuntansi Biaya merupakan salah satu matakuliah yang menjadi dasar pada program peminatan MK.

Dari kedua percobaan yang dilakukan tidak terdapat informasi berkaitan dengan program peminatan MSDM. Hal ini disebabkan kombinasi item yang dihasilkan untuk program peminatan MSDM tidak memenuhi *minimum support* yang sudah ditentukan.

V. KESIMPULAN

Kesimpulan yang dapat diambil dari perancangan *data warehouse* dan penerapan teknik *data mining* pada data akademik di UNAS antara lain:

1. Adanya *data warehouse* membuat proses penyusunan laporan di UNAS menjadi lebih sederhana. Kemampuan untuk melakukan *customization report* sesuai dengan yang diinginkan pengguna membuat terciptanya efisiensi waktu pelaporan dari yang sebelumnya memerlukan waktu satu bulan jika dikerjakan dengan membuat program baru atau satu minggu jika dikerjakan secara manual menjadi hanya satu hari. Keadaan ini tentunya memberikan penghematan kepada UNAS akan sumber daya manusia dan biaya untuk kebutuhan penyusunan laporan.
2. Dengan memanfaatkan informasi yang ada pada *data warehouse* UNAS, telah berhasil dilakukan analisis data lebih lanjut dengan menggunakan teknik *data mining* untuk mengenali pola karakteristik mahasiswa yang mengambil program peminatan tertentu di program studi Manajemen Perusahaan UNAS. Hasilnya, ditemukan pola karakteristik mahasiswa yang mengambil program peminatan MP dan MK. Untuk program peminatan MSDM tidak didapatkan hasil, hal ini disebabkan kombinasi item yang dihasilkan pada program peminatan MSDM tidak memenuhi *minimum support* yang sudah ditentukan. Sedangkan untuk MO, karena jumlah data yang kecil yaitu hanya 33 data maka tidak digunakan dalam penelitian ini.

DAFTAR PUSTAKA

- [1]. Paulraj Ponniah. 2001. *Data Warehousing Fundamentals*. Wiley-Interscience Publication.
- [2]. Jeffrey A. Hoffer, Mary B. Prescott, dan Fred R. McFadden. 2005. *Modern Database Management*. Seventh Edition. Prentice Hall.
- [3]. Kathleen Polo, Juan Gonzales dan Edwin Rojas. 2006. *Quick Tutorial about Mondrian Datamart*. International Potato Center (CIP). Research Informatics Unit (RIU).
- [4]. Witten, Ian H dan Eibe Frank, 2005. *Data Mining : Practical Machine Learning Tools and Technique*. Second Edition. Morgan Kaufmann.
- [5]. <http://www.comp.nus.edu.sg>. Diakses pada tanggal 07 Mei 2007. *Data Mining II. Classification Based Association* (CBA).