

# PERANCANGAN DATA WAREHOUSE DAN PENERAPAN DATA MINING UNTUK MENDUKUNG SISTEM INFORMASI AKADEMIK UNIVERSITAS KUNINGAN (DESIGN DATA WAREHOUSE AND IMPLEMENTATION OF DATA MINING FOR ACADEMIC SUPPORT SYSTEMS INFORMATION KUNINGAN UNIVERSITY )

Aah Sumiah<sup>1</sup>, Dede Irawan<sup>2</sup>, Fauziah<sup>3</sup>

## ABSTRAK

*Good quality education is a goal to be achieved by any educational institution. In the world of education dynamically evolving the required educational institutions to improve the quality of education and graduates. To achieve this goal each competing institution to produce a model of innovation and education. Institutions can utilize the availability of appropriate information very influential in increasing the quality of service to students and help produce the right business decisions.*

*The purpose of this study was to produce a model of the data warehouse Brass University Academic Information Systems and generate patterns of data mining to the data admission to prospective students determine possible re-register or not so if the resignation of incoming freshmen are expected to be known early will help the management campus to take the actions necessary to maintain the prospective student. This Tstudy used two models that modeling techniques for data warehouse star schema model and decision tree classification technique for data mining*

*Keyword: Data warehouse, data mining, classification, and decision tree*

### 1. Pendahuluan

Kualitas pendidikan yang baik merupakan tujuan yang ingin dicapai oleh setiap lembaga pendidikan. Dalam dunia pendidikan yang terus berkembang secara dinamis maka lembaga pendidikan diharuskan untuk meningkatkan kualitas pendidikan dan lulusan. Untuk mencapai tujuan tersebut masing masing lembaga berlomba untuk menghasilkan inovasi dan model pendidikan yang baik.

Lembaga pendidikan yang dapat memanfaatkan ketersediaan informasi yang tepat sangat berpengaruh pada meningkatnya kualitas layanan kepada mahasiswa dan membantu menghasilkan keputusan bisnis yang tepat.

Salah satu factor yang mempengaruhi arah perkembangan pendidikan di masing-masing lembaga pendidikan ialah proses pembuatan kebijakan. Dalam sebuah lembaga pendidikan, baik itu yang kecil ataupun yang besar Selalu dibutuhkan informasi-informasi untuk menunjang proses pengambilan keputusan. Dapat dikatakan tidak mungkin seorang decision maker membuat sebuah kebijakan tanpa didukung oleh informasi yang tepat. Kata tepat disini tidak berarti bahwa informasi harus selalu up to date, tetapi bisa berupa informasi yang sudah lama, berkas-berkas, rangkuman-rangkuman dan sebagainya.

Keputusan yang akan diambil seringkali harus didukung oleh tersedianya informasi yang cepat, ini sering terjadi bila keputusan harus segera diambil. Dengan menggunakan komputer maka perolehan informasi yang cepat dan tepat dapat direalisasikan. Menggunakan komputer tidak berarti semua informasi pasti dapat diperoleh dari komputer, tetapi harus ada data-data yang disimpan di dalam komputer dan ada prosedur-prosedur yang dapat

memperbaharui dan mengolah data-data tersebut menjadi sebuah informasi yang dibutuhkan.

Universitas kuningan adalah sebuah lembaga pendidikan yang berdiri pada tahun 2003 di kabupaten kuningan jawa barat yang merupakan merger dari 4 sekolah tinggi yang ada di kabupaten kuningan yaitu sekolah tinggi keguruan dan ilmu pendidikan (STKIP), sekolah tinggi ilmu ekonomi (STIE), sekolah tinggi manajemen informatika komputer (STMIK) dan sekolah tinggi ilmu kehutanan(STIKU)

Sejak berdiri pada taun 2003 animo masyarakat untuk memasuki jenjang pendidikan tinggi universitas kuningan setiap tahun meningkat. Sehingga dengan kondisi ini sudah sewajarnya jika dibangun sebuah system informasi akademik guna memberikan pelayanan yang maksimal kepada mahasiswa. Universitas kuningan menyadari bahwa untuk memanfaatkan data yang ada dalam system informasi akademik harus di perlukan analisis data untuk menggali potensi-potensi informasi yang ada, untuk itu perlu dibangun sebuah data warehouse yang selanjutnya dapat di analisis menggunakan teknik data mining. Penggunaan teknik data mining diharapkan dapat memberikan pengetahuan-pengetahuan yang tadinya tersembunyi didalam gudang data.

## 2. Tinjauan Pustaka

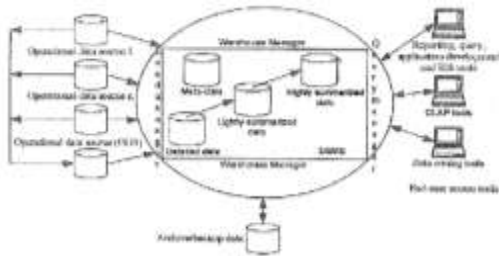
### A. Data Warehouse

#### 1. Sistem Data Warehouse

Menurut W.H. Inmon dan Richard D.H., *data warehouse* adalah koleksi data yang mempunyai sifat berorientasi subjek, terintegrasi, time-variant, dan bersifat tetap dari koleksi data dalam mendukung proses pengambilan keputusan management.

*data warehouse* adalah database yang saling bereaksi yang dapat digunakan untuk *query* dan analisis, bersifat orientasi subjek, terintegrasi, time-variant, tidak berubah yang digunakan untuk membantu para pengambil keputusan.

#### 2. Arsitektur komponen data warehouse



Gambar 2.1 Arsitektur komponen data warehouse Conolly, T.M., Begg

### B. Data Mining

#### 1. Definisi data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [Turban 2005].

Menurut gartner group, data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dan sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistic dan matematika [Larose 2005]

#### 2. Komponen Data Mining

Secara alami, material data mining sebenarnya sudah terbentuk karena factor rutinitas dan waktu seraya perusahaan melakukan aktivitasnya. Tanpa disadari perusahaan berinvestasi dengan menggunakan budgetnya untuk penggunaan teknologi informasi atau computer. Teknologi *data mining* mulai muncul karena akumulasi data yang besar dan pesat pertumbuhannya sehingga menimbulkan apa

yang disebut *rich of data but poor information*. Tumpukan data ini tidak dapat digunakan pada aplikasi yang ada sehingga menjadi gudang data.

#### 3. Pengelompokan data mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu [larose, 2005]

##### a. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

##### b. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable target. Estimasi lebih kearah numeric dari pada kearah kategori.

##### c. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa datang.

##### d. Klasifikasi

Dalam kalsifikasi terdapat target variable kategori, sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

##### e. Pengklusteran

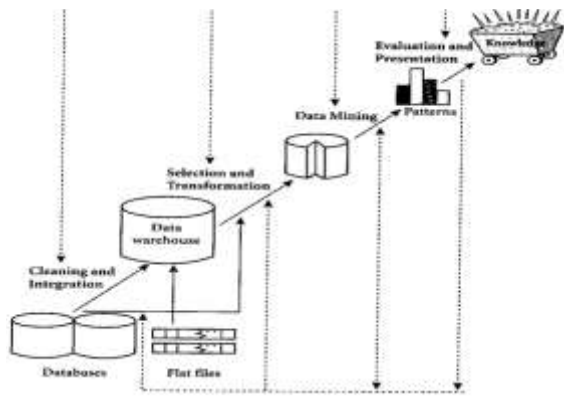
Pengklusteran merupakan pengelompokan record, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan record-record dalam cluster.

##### f. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang biasa.

#### 4. Tahapan Data Mining

Tahapan *data mining* dilakukan melalui pembentukan *data warehouse* karena hanya struktur data dari *data warehouse* yang dapat digunakan untuk *data mining*. Dengan melaksanakan ekstraksi, pembersihan dan transformasi (*Extraction, Cleansing, Transformation*) atas data transaksional yang disebut dengan data OnLine Transactional Processing (OLTP) ke OnLine Analytical Processing (OLAP) System untuk membuat sebuah *data warehouse* maka akan didapat suatu kumpulan data yang luas dan besar serta bersifat statis dan historis yang berasal dari data-data transaksi OLTP selama perusahaan menjalankan bisnisnya [Vieira 2000].



Gambar 2.2 Tahapan Data Mining [Pramudiono 2003]

### 5. Algoritma Data Mining

Algoritma C.45 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Data didalam pohon keputusan biasanya dinyatakan dalam bentuk table dengan atribut dan record. atribut menyatakan suatu parameter yang dibuat sebagai criteria dalam pembentukan pohon. Misalnya untuk menentukan main tennis, criteria yang di perhatikan adalah cuaca, angin, temperature. Salah satu atribut merupakan atribut yang menyatakan data solusi per item data yang disebut target atribut. Atribut memiliki nilai-nilai yang dinamakan dengan instance. Misalkan atribut cuaca mempunyai instance berupa cerah, berawan, dan hujan [Basuki & Syarif, 2003].

Secara umum algoritma C.45 untuk membangun pohon keputusan adalah sebagai berikut :

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama

### C. SQL Server 2008

Sql server 2008 merupakan database engine yang mendukung penggunaan arsitektur client server. Penggunaan client server bisa digambarkan bahwa aplikasi dibagi menjadi dua bagian yaitu bagian client dan server. Bagian client atau user bisa mendapatkan keuntungan bahwa sql server bisa dijalankan di system operasi personal seperti windows 95 ataupun window ME, dan juga bisa di koneksikan dengan berbagai aplikasi atau bahasa pemrograman dengan bantuan ODBC, sedangkan bagian server atau database administrator (DBA) akan mempunyai database yang handal, konsisten, metode locking dan control secara baik.

### D. Weka

Weka merupakan aplikasi data mining yang berbasis opens source (GPL) dan berengine java. Aplikasi ini dikembangkan pertama kali oleh sebuah universitas selandia baru yang bernama Universitas Waikato sebelum menjadi bagian di pentaho. Software yang mulai dikembangkan sejak tahun 1994 ini telah menjadi software data mining open source yang paling populer. Banyaknya algoritma data mining dan machine learning, kemudahan dalam penggunaan ditambah lagi selalu up to date dengan algoritma-algoritma baru yang muncul menjadikan software ini banyak digunakan, tidak saja di gunakan untuk akademik namun banyak juga dipakai untuk meramalkan bisnis perusahaan.

Weka terdiri dari koleksi algoritma machine learning yang dapat digunakan untuk melakukan generalisasi atau formulasi dari sekumpulan data sampling. Kekuatan weka sebenarnya terletak pada algoritma yang makin lengkap dan canggih , kesuksesan data mining tetap terletak pada factor pengetahuan manusianya sendiri. Hal tersebut bisa dilakukan jika adanya pengumpulan data yang berkualitas sehingga akan menjamin keakuratan formulasi yang diharapkan.

### E. Tinjauan Studi

1. Penelitian yang dilakukan oleh sutedi dengan judul “Perancangan Data warehouse dan penerapan data mining untuk mendukung bidang akademik pada Institut Informatika dan Bisnis Darmajaya “. Pada penelitian ini : hasil penelitian ini berupa data warehouse yang dapat menyajikan informasi guna mendukung proses evaluasi dan perencanaan promosi penerimaan mahasiswa baru ke daerah-daerah dan sekolah-sekolah, proses evaluasi dan perencanaan media promosi yang akan digunakan, proses pengamatan trend mahasiswa berdasarkan status akademiknya, proses evaluasi dan perencanaan studi mahasiswa, serta proses evaluasi kinerja program studi dari segi kualitas lulusan dan masa studinya serta penerapan data mining untuk mencari rule-rule yang digunakan dalam mengarahkan peminatan mahasiswa dan mengarahkan pemilihan program studi bagi calon mahasiswa baru.
2. Penelitian yang dilakukan oleh Ariana azimah dan yudho giri sucahyo dengan judul “penggunaan data warehouse dan data mining untuk data akademik. Sebuah studi kasus pada Universitas Nasional” Pada penelitian ini : hasil penelitian adalah pembuatan data warehouse akademik dan analisis data mining untuk mengukur peminatan mahasiswa terhadap mata kuliah tertentu.

3. Penelitian yang dilakukan oleh Aah Sumiah, M.kom dengan judul “Model data mining pada system informasi akademik. Studi kasus Universitas Kuningan”

Pada penelitian ini : hasil penelitian adalah model data mining menggunakan metode classification menggunakan pendekatan decision tree untuk memprediksi tingkat registrasi calon mahasiswa baru.

## F. Tinjauan Objek Penelitian

### 1. Sejarah Universitas Kuningan

Universitas Kuningan (UNIKU) didirikan sebagai perwujudan dari idealisme dan komitmen Yayasan Pendidikan Sang Adipati Kuningan untuk terus menerus berkarya khususnya dalam bidang peningkatan sumber daya manusia menuju peningkatan mutu kehidupan masyarakat pada umumnya melalui pendidikan tinggi. Gagasan tentang pendirian Universitas di Kuningan sebetulnya telah ada sejak tahun 1979 ketika Yayasan ini didirikan. Namun karena keterbatasan sumber daya, baik sumber daya manusia maupun sumber daya keuangan, maka niat itu tidak bisa langsung diwujudkan sekaligus. Mengingat keterbatasan itu yayasan menggunakan strategi bertahap melalui pendirian sekolah tinggi sebagai cikal bakal berdirinya universitas.

Sekolah tinggi pertama kali lahir pada tahun 1985 yaitu Sekolah Tinggi Keguruan dan Ilmu Pendidikan (STKIP) Kuningan. Sekolah tinggi ini sebelum menjadi Universitas Kuningan mempunyai tiga program Studi yaitu Pendidikan Bahasa dan Sastra Indonesia (S1), Pendidikan Biologi (S1), dan Pendidikan Ekonomi (S1).

Sekolah tinggi kedua lahir tahun 1995 yaitu Sekolah Tinggi Ilmu Ekonomi (STIE) Kuningan, dengan 2 (dua) program studi yaitu program Studi Manajemen (S1) dan Program Studi Akuntansi (S1).

Sekolah tinggi ketiga lahir pada tahun 2001; yaitu Sekolah Tinggi Ilmu Kehutanan (STIK) Kuningan dan Sekolah Tinggi Kehutanan (STIK) Kuningan dengan 2 (dua) Program Studi yaitu Program Studi Budidaya Hutan (S1). Karena dirasakan adanya tuntutan perkembangan teknologi terutama computer maka pada waktu yang bersamaan lahir pula Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Kuningan dengan 4 program studi yaitu Program Studi Teknik Informatika (S1), Sstem Informasi (S1), Teknik (DIII) dan Manajemen Informatika (DIII).

Setelah memiliki empat sekolah tinggi, yayasan merasa punya cukup pengalaman untuk mendirikan universitas. Oleh karena itu pengembangan dan penggabungan keempat sekolah tinggi lantas dilaksanakan. Berdasarkan Surat Keputusan Menteri Pendidikan Nasional Nomor :

62/D/O/2003 tanggal 6 Juni 2003, berdirilah Universitas Kuningan diharapkan yang merupakan penggabungan keempat sekolah tinggi yang ada dibawah naungan Yayasan Pendidikan Sang Adipati Kuningan. Universitas ini diresmikan oleh Menteri Pendidikan Nasional Prof. A. Malik Fajar, M.Sc. pada tanggal 17 Juni 2003.

Pada tahun 2006 Universitas Kuningan membuka dua program studi baru, yaitu Program Studi Pendidikan Bahasa Inggris Jenjang S1, dan Program Studi Pendidikan Ekonomi Jenjang S2 (Magister).

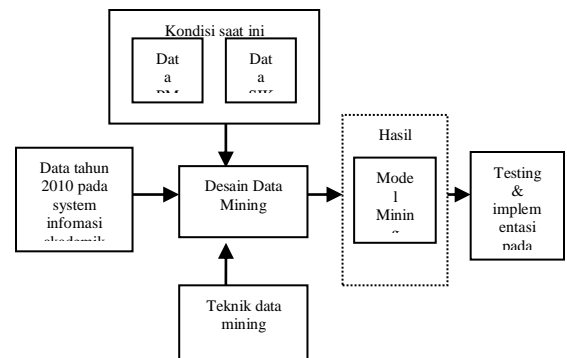
### 2. Fasilitas Layanan

Berdasarkan penelitian, Universitas Kuningan sudah memiliki sistem informasi akademik yang diberi nama SIKA (sistem informasi akademik) menggunakan database MYSQL dengan aplikasi berbasis web. SIKA melakukan pengolahan data untuk semua fakultas yaitu: Fakultas ilmu computer, Fakultas kehutanan, Fakultas ekonomi, dan Fakultas keguruan ilmu pendidikan

### 3. Kerangka Konsep

Sesuai dengan tujuan penelitian yang telah di tetapkan maka peneliti ingin melakukan analisa dan perancangan data warehouse sistem informasi akademik yang ada dengan konsep yang menurut peneliti lebih baik dengan dukungan data mining Sehingga pihak manajemen universitas dapat mengambil keputusan dengan cepat.

Kerangka konsep di atas dapat di gambarkan seperti berikut ini :



Gambar 2.3 Kerangka pemikiran

Kerangka Konsep Diatas Dapat di jelaskan sebagai berikut :

Perancangan data warehouse dilandasi atas kebutuhan data warehouse yang akan digunakan sebagai sumber data untuk system informasi akademik. Data warehouse ini di rancang dengan menganalisis data operasional yang sedang berjalan berdasarkan data PMB (Penerimaan Mahasiswa Baru) dan Data SIKA (Sistem Informasi Akademik) serta memasukan model data warehouse dan data mining yang sesuai. Hasil

dari perancangan diharapkan akan terbentuk sebuah prototype data warehouse akademik dan data mining melalui tahap testing dan implementasi.

### 3. Desain Penelitian

#### 1. Metode Penelitian

Agar penelitian yang dilakukan lebih terarah maka peneliti menggunakan suatu metode deskriptif kualitatif yang bertujuan untuk mendapatkan gambaran lengkap dari proses dan subjek penelitian. Dalam penelitian ini metode yang digunakan adalah *waterfall*. Langkah yang dilakukan dimulai dengan identifikasi dan analisis kebutuhan pengguna, yang dilakukan bersama sama dengan mengamati kondisi sistem akademik universitas kuningan yang saat ini digunakan. Setelah itu dilanjutkan dengan desain data warehouse dan desain data mining. Desain-desain tersebut kemudian diaplikasikan dan dievaluasi agar didapatkan aplikasi yang benar-benar bermanfaat bagi pengguna

#### 2. Metode Pengumpulan Data

Data untuk penelitian ini diambil dari data primer dan data sekunder. Data primer di dapat dari wawancara dan observasi lapangan. Data sekunder di dapat dari data objek penelitian (database), studi literature dan tulisan ilmiah tentang data mining.

##### a. Wawancara

Penulis melakukan wawancara dengan pengembang system tentang system yang sedang berjalan. Sehingga di ketahui kebutuhan dari system.

##### b. Observasi

Penulis melakukan observasi langsung dengan melihat system informasi Penerimaan mahasiswa baru dan system informasi akademik yang sedang berjalan.

##### c. Literatur

Penulis mengumpulkan literature dan sumber-sumber yang mendukung penelitian.

#### 3. Teknik Analisa data

Teknik analisis di bagi menjadi dua yaitu perancangan data warehouse dan analisis data mining

##### a. Desain Data Warehouse

1. Merancang database untuk data warehouse  
Membuat perancangan database untuk data warehouse system Informasi akademik
2. Membuat model (schema) data warehouse  
Model data yang penulis buat yaitu menggunakan model data dimension star schema

##### b. Analisa data mining

Salah Salah satu cara untuk melakukan analisis kemungkinan pengunduran diri seorang calon mahasiswa adalah dengan melakukan klasifikasi data dan variabel dari datawarehouse PMB yang sudah terbentuk, kemudian membuat atribut

sebagai akar, membuat cabang untuk tiap-tiap nilai dengan menggunakan algoritma C4.5 dan ini di aplikasikan dengan menggunakan tools microsoft SQL Server 2008 dan Weka 3.7.7

#### 4. Instrumen

Instrument yang digunakan peneliti untuk menunjang proses penelitian adalah menggunakan Microsoft access 2010, Microsoft excel 2010, Microsoft word 2010, Microsoft Sql Server 2008 ,Weka 3.7.7 (2012) dan beberapa software tambahan.

#### 4. Hasil Penelitian

##### a. Sumber Data

Universitas Kuningan Memiliki Database yang diberi nama SikaV menggunakan database MYSQL 5.5 yang bisa di akses melalui tools phpMyAdmin SQL Dump --version 3.3.10deb1. Isi Database SikaV mencakup banyak table yang terdiri dari data Akademik, data kepegawaian dan data keuangan. Selain itu data di ambil dari data registrasi penerimaan mahasiswa baru tahun 2008 , tahun 2009, dan tahun 2010 dimana data masih berbentuk file excel.

##### b. Desain Data Warehouse

###### 1. Pemilihan Proses

Seperti disebutkan diatas bahwa Data yang di gunakan dalam penelitian ini adalah data akademik yang berasal dari Database SikaV, untuk itu data akademik yang ada dalam database SikaV kemudian di pilih dan dipisahkan untuk proses pembentukan data warehouse.

Adapun proses tersebut adalah sebagai berikut:

1. Proses DTS ke database yang diberi nama "Data sumber"  
Table-tabel yang terdapat didalam Database SikaV di Transformasi ke dalam Database yang di beri nama "Data\_sumber".
2. Membuat Database "Data source".  
Langkah selanjutnya adalah memilih tabel-tabel yang berhubungan dengan data akademik yang terdapat di dalam database "Data\_Sumber" untuk kemudian di pindahkan ke dalam database terpisah yang di beri nama "Data Source".
3. Membuat Tabel staging  
Dalam Database "Data Source" dibentuk Tabel-tabel Staging yang berfungsi untuk menampung data dari sumber data yang berbeda. Di sini dilakukan proses pemilihan struktur dan penyeragaman struktur table sehingga menghasilkan struktur table yang sama yang akan digunakan untuk keperluan data warehouse.
4. Melakukan proses transformasi table staging ke data warehouse.

Pada proses ini struktur tabel dan data yang sudah sama kemudian di simpan dan di transformasi ke dalam database baru yang di beri nama "DW\_AKADEMIK". Database Ini digunakan sebagai Data warehouse yang nantinya di gunakan sebagai acuan untuk keperluan analisis dan pembentukan laporan.

## 2. Pemilihan Grain

Grain merupakan data dari calon fakta yang dapat di analisis. Grain yang di gunakan untuk merancang data warehouse Sistem Informasi Akademik adalah sebagai berikut :

1. Jumlah mahasiswa baru berdasarkan fakultas
2. Jumlah mahasiswa baru berdasarkan program studi
3. Jumlah mahasiswa baru berdasarkan asal sekolah
4. Jumlah mahasiswa berdasarkan tahun masuk dan gelombang
5. Jumlah mahasiswa berdasarkan Jenjang Studi
6. Nilai Mahasiswa Per mata kuliah
7. Nilai IPS dan IPK Mahasiswa untuk per semester
8. Jumlah Mahasiswa yang melakukan cuti
9. Jumlah Mahasiswa yang melakukan Pindahan

## 3. Menentukan Measure

Grain dari table fakta nantinya akan digunakan untuk menentukan fakta-fakta yang di tampilkan. Pada tahap ini hal yang dilakukan adalah menentukan measure yang di butuhkan oleh table fakta.

Berikut ini adalah measure yang di butuhkan table fakta :

1. Measure untuk fakta calon mahasiswa baru berdasarkan program studi, jenis kelamin, gelombang, dan asal sekolah adalah jumlah calon mahasiswa
2. Measure untuk fakta nilai mahasiswa adalah data nilai mahasiswa dan mata kuliah adalah data nilai
3. Measure untuk fakta hasil studi berdasarkan KRS adalah KRS
4. Measure untuk fakta Cuti mahasiswa adalah jumlah cuti
5. Measure untuk menentukan jumlah mahasiswa pindahan adalah jumlah pindahan

## 4. Pemilihan Table Fakta

Table fakta adalah table yang nantinya digunakan sebagai sumber dalam analisis data. Berdasarkan proses analisis maka terbentuk table fakta seperti berikut ini:

1. Table Fact\_Data\_PMB
2. Tabel Fact\_Data\_Nilai
3. Tabel Fact\_KRS
4. Tabel Fact\_Pindah
5. Tabel Fact\_Cuti

## 5. Mengidentifikasi Dimensi

Berdasarkan kebutuhan tersebut maka dimensi yang dibentuk adalah sebagai berikut ini:

- a. Tabel Dim\_info\_daftar
- b. Table Dim\_prodi
- c. Tabel Dim\_JK
- d. Tabel Dim\_Registrasi
- e. Tabel Dim\_gelombang
- f. Tabel Dim\_SMA
- g. Tabel Dim\_Matakuliah
- h. Tabel Dim\_mhs
- i. Tabel\_nilai
- j. Tabel Dim\_mhs
- k. Tabel Dim Tr\_KRS
- l. Tabel Dim\_Nilai
- m. Tabel Dim\_Tr\_Akd
- n. Tabel Dim\_mhs
- o. Tabel Dim\_alasan

## 6. Perancangan Schema Bintang (Star Schema)

Berdasarkan hasil rancangan di peroleh star schema sebagai berikut :

### a. Star Schema Data PMB

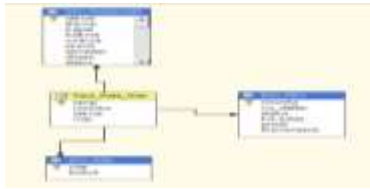


Gambar 4.1 Star Schema Data PMB

Star Schema data PMB akan menghasilkan Informasi tentang penerimaan mahasiswa baru yang digunakan untuk melihat :

- a. Berapa jumlah calon mahasiswa yang mendaftar, jumlah mahasiswa yang registrasi dan yang tidak registrasi, program studi yang banyak dipilih dan jenjang untuk tiap jurusan per periode
- b. Jumlah Mahasiswa baru berdasarkan asal SMA
- c. Komposisi jumlah mahasiswa baru pada periode tertentu

**b. Star Schema Data Nilai**



Gambar 4.2 Star Schema Data Nilai

Star schema data nilai akan menghasilkan informasi tentang nilai mahasiswa yang dapat digunakan untuk :

- a. Melihat data nilai mahasiswa per matakuliah
- b. Melihat nilai mahasiswa per mahasiswa dan per program studi

**c. Star Schema Data KRS**



Gambar 4.3 Star Schema Data KRS

Star Schema data KRS akan menghasilkan informasi yang dapat digunakan untuk melihat :

- a. Jumlah Sks yang diambil
- b. Jumlah Nilai IPS dan Nilai IPK mahasiswa

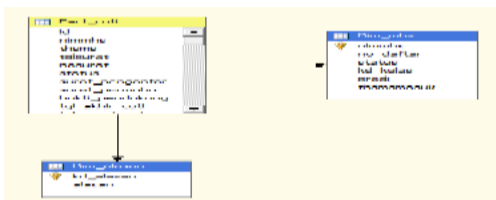
**d. Star Schema Pindahan**



Gambar 4.4 Star Schema Mahasiswa Pindahan

Informasi mahasiswa pindahan dapat digunakan untuk melihat data jumlah mahasiswa pindahan per program studi beserta alasannya

**e. Star Schema Cuti**



Gambar 4.5 Star Schema Cuti Mahasiswa

Informasi mahasiswa Cuti dapat digunakan untuk melihat data jumlah mahasiswa yang cuti per program studi beserta alasannya]

**7. Transformasi data**

Merupakan sekumpulan operasi yang dilakukan pada sumber data sebelum di copy ke tempat tujuan. Pada data warehouse peran transformasi sangat penting untuk proses pemindahan data dan penyeragaman data kedalam data warehouse.

Sumber data yang berasal dari SikaV di transformasi menggunakan Navicat software, dilakukan proses Batch file kemudian di ekspor ke microsoft access. Setelah itu dilakukan proses ekspor dan import data dari Microsoft access tersebut ke sql server 2008. Didalam sql server kemudian dilakukan pemfilteran data dan pemilihan data yang berhubungan dengan data akademik.

**c. Analisa Data Mining**

Berdasarkan data yang di peroleh dari data warehouse Data Penerimaan Mahasiswa Baru diperoleh informasi sebagai berikut:

**a. Informasi Jumlah Pendaftar Dan Informasi Jumlah Calon Mahasiswa Yang Melakukan registrasi**

Table Rekapitulasi Pendaftaran Per Fakultas Tahun Akademik 2010/2011 adalah sebagai berikut :

Tabel 4.1 Rekapitulasi Pendaftaran Mahasiswa Tahun 2010/2011

Pendaftar			Daftar Ulang		
No	Fakultas	Jumlah	No	Fakultas	Jumlah
1	FE	272	1	FE	237
2	FHUT	36	2	FHUT	31
3	FKIP	934	3	FKIP	789
4	FKOM	306	4	FKOM	255
Total		1548	Total		1312

**b. Melakukan Pre Processing**

Tabel 4.2 Format Data Calon Mahasiswa Setelah Pemilihan Variabel

St_pindahan	gelombang	Jenjang	registered
B	1	S1	Y
B	1	S1	Y
B	1	S1	Y
B	1	S1	Y

**c. Perhitungan Nilai Entropy dan Gain**

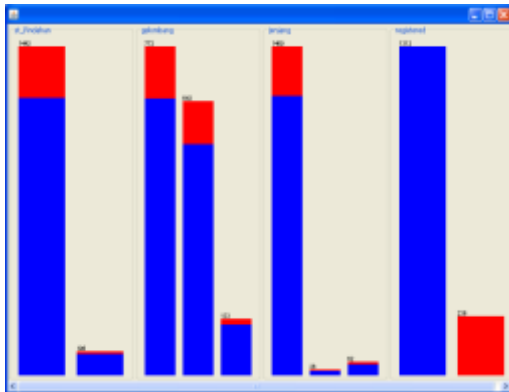
Untuk menghitung nilai gain, hal pertama yang harus dilakukan adalah menghitung total jumlah kasus, jumlah kasus untuk keputusan registrasi, dan jumlah kasus untuk keputusan tidak\_registrasi dan entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut.

Tabel 4.3 Perhitungan Node 1.0

Node		Jumlah kasus (S)	Tidak registrasi (S1)	Registrasi (S2)	Entropy	Gain
1	Total	1548	236	1312	0.61590453	
	St_pindahan					0.0010675
	baru	1442	225	1217	0.624734894	
	Pindahan	106	11	95	0.480844694	
	Gelombang					0.00174509
	Gelombang1	772	122	650	0.629578416	
	Gelombang2	643	101	542	0.6272688	
	Gelombang3	133	13	120	0.461804332	
	Jenjang					0.00619397
	S1	1460	219	1241	0.609840305	
	S2	62	5	50	0.543199147	
	D3	26	12	21	0.763703647	

**d. Pemrosesan Data Menggunakan Weka**

Berikut ini adalah gambar Visualisasi yang memuat informasi atribut dengan tampilan grafik diagram batang dari keseluruhan atribut :



Gambar 4.6 Grafik Diagram Batang Untuk Keseluruhan Atribut.

Pada data tersebut dapat diperoleh informasi sebagai berikut :

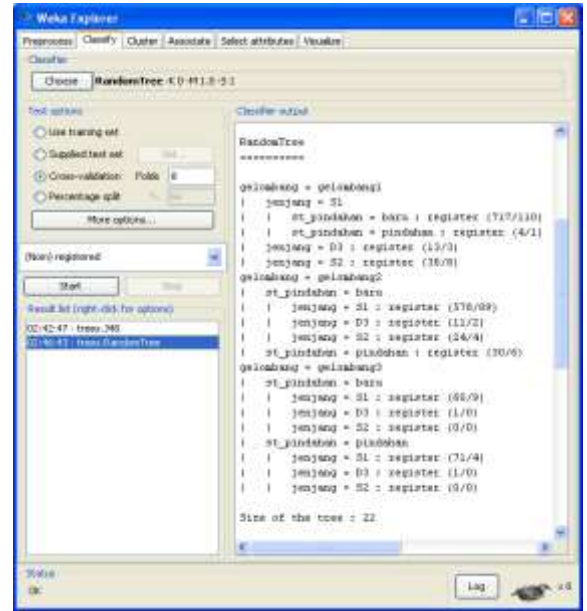
Tabel 4.4 Informasi Data Mahasiswa Berdasarkan Kategori Tertentu

No	Atribut	Kategori	Jumlah
1	St_pindahan	Baru	1442
		Pindahan	106
2	gelombang	Gelombang1	772
		Gelombang2	643
		Gelombang3	133
3	Jenjang	S1	1460
		S2	62
		D3	26

4	registered	register	1312
		Tidak_register	236

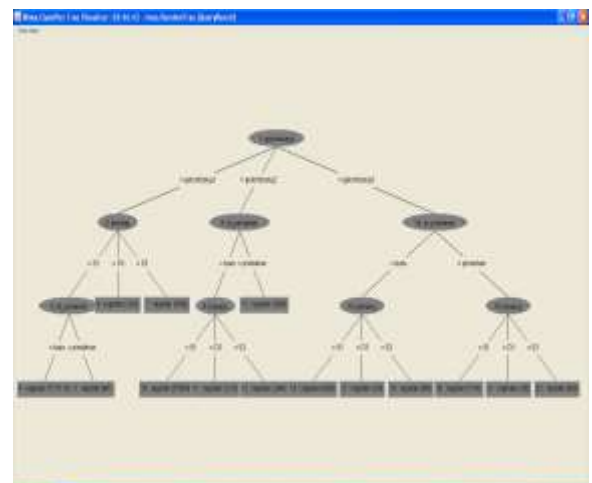
Hasil diatas akan di coba dengan menggunakan metode random tree sebagai pembanding. Dimana metode ini dilakukan dengan membentuk pohon keputusan acak pada tiap node nya tanpa melakukan pemangkasan sehingga hasil *decision tree* dapat di gambarkan.

Hasil dari Random tree adalah sebagai berikut :



Gambar 4.7 Hasil decision Tree

Hasil Decision Tree



Gambar 4.8 Hasil Decision tree



## 5. Penutup

### A. Kesimpulan.

1. Dengan adanya data warehouse proses penyusunan laporan menjadi lebih sederhana dan tidak membutuhkan waktu yang lama seperti halnya proses manual yang membutuhkan waktu ber minggu-minggu.
2. Informasi *knowledge discovery* dari *data mining* dengan metode *classification* yang menggunakan pendekatan *decision tree* dapat menjawab apa yang menjadi kebutuhan manajemen yaitu memprediksi tingkat registrasi mahasiswa dengan akurasi yang cukup baik serta menunjukkan *pattern* yang selama ini tidak tampak pada system yang sudah ada.
3. Metodologi yang digunakan penulis dalam melakukan pengukuran akurasi terhadap model data mining cukup mudah untuk diimplementasikan, namun kesulitan terbesar berada pada pemilihan variable serta penghitungan nilai entropi dan gain sehingga semakin banyak variabel yang ingin diprediksi akan semakin banyak waktu yang dibutuhkan untuk mencapai sasaran data mining yang ingin dicapai.

### B. Saran

Mengingat Universitas Kuningan adalah lembaga pendidikan yang sedang berkembang, maka untuk memperoleh informasi yang lebih optimal, sebaiknya pihak lembaga menggunakan data warehouse dan data mining pada system informasi akademik untuk menunjang proses pengambilan kebijakan.

## 6. Ucapan Terima Kasih

Penulis mengucapkan terima kasih yang sedalam-dalamnya kepada :

1. Direktorat penelitian dan pengabdian kepada masyarakat Ditjen Dikti selaku penyumbang dana.
2. Bapak Fahmi Yusuf MMSI selaku Dekan Fakultas Ilmu Komputer Universitas Kuningan.
3. Bapak Agus Yadi Ismail, M.Si selaku ketua lembaga penelitian Universitas Kuningan.
4. Bapak Oya Suryana SE selaku staf bagian akademik Universitas Kuningan.
5. Seluruh karyawan/ti Fakultas Ilmu Komputer Universitas Kuningan.
6. Teman-teman yang selalu memberikan motivasi dan dorongan sehingga laporan penelitian ini bisa diselesaikan tepat waktu.

## Daftar Pustaka

- [Aah 2011] Aah Sumiah, M.Kom, “ Model Data Mining Pada Sistem Informasi Akademik: Studi Kasus Universitas Kuningan”, 2011
- [Azimah 2007] Azimah Ariana Dan Suchyo Yudho Giri, “Penggunaan Data Warehouse Dan Data Mining Untuk Data Akademik. Sebuah Studi Kasus Pada Universitas Nasional”, 2007
- [Edward 2010] Edward, “Model Data Mining Sebagai Sarana Untuk Mendukung Pengambilan Keputusan manajemen Operasional: Studi Kasus PT. Teguh Aman Sejahtera”, 2010
- [Jogiyanto 2008] Prof. Jogiyanto HM, Akt, MBA, Ph.D, “Metodologi Penelitian Sistem Informasi”, Penerbit Andi, Yogyakarta 2008.
- [Kusrini 2009] Kusrini, dan Luthfi Emha Taufiq, “Algoritma Data Mining”, Penerbit Andi, Yogyakarta 2009
- [Larose 2005] Larose, Daniel T. “Discovering Knowledge ind Data :An Introduction to data mining. John Willey & Sons.inc, 2005
- [McLeod 2004] McLeod,Jr. Raymond , “Sistem Informasi Manajemen”., Indeks., 2004
- [Ponniah 2001] Ponniah Paulraj, “Data Warehousing Fundamental : A Comprehensive Guide For IT Professional.”.,John Willey Interscience Publication, 2001
- [Ramakrishnan 2003] Ramakrishnan, Raghu and Gehrke Johannes, “Database Management Systems 3<sup>rd</sup> Edition”, 2003, McGraw-Hill, New York, 2003.
- [Rizky 2004] Rizky, Soetam, “Panduan Belajar SQL Server”, Prestasi Pustaka Publisher , Jakarta 2004
- [Sulianta 2010] Sulianta, Feri dan Juju, Dominikus, “Data Mining, Meramalkan Bisnis Perusahaan ” , Elex Media Komputindo, 2010
- [Sutedi 2005] Sutedi, “Perancangan Data Warehouse Dan Penerapan Data Mining Untuk Mendukung Bidang Akademik Pada Institut Informatika Dan Bisnis Darmajaya, 2005
- [Turban 2005] Turban, E.,dkk , “Decision Support System And Inteligent System”, Andi , Yogyakarta,2005
- [Vieira 2000] Vieira, Robert, CMA, MCDBA, “Professional SQL Server 2000 Programming”, Wrox Press Ltd, Birmingham 2000.
- [Written 2005] Written, Ian H and Frank, Eibe, ‘Data Mining, Practical Machine Learning Tools And Technique’ Morgan Kaufman Publisher, San Francisco, 2005.