

# PERANCANGAN DATA WAREHOUSE DAN PENERAPAN DATA MINING UNTUK MENDUKUNG SISTEM INFORMASI AKADEMIK UNIVERSITAS KUNINGAN

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

*\*Fakultas Ilmu Komputer Universitas Kuningan*

*Jalan Tjut Nyak Dhien Cijoho Kuningan Telepon (0232) 873696 Fax. (0232) 874 824*

*[Aah.Sumiah@gmail.com](mailto:Aah.Sumiah@gmail.com)<sup>1</sup>, [dede.pendekar@gmail.com](mailto:dede.pendekar@gmail.com)<sup>2</sup>, [fauziah.cdc@uniku.ac.id](mailto:fauziah.cdc@uniku.ac.id)<sup>3</sup>*

## *Abstrak*

*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.*

*Tujuan penelitian ini adalah untuk menghasilkan model data warehouse Sistem Informasi Akademik Universitas Kuningan dan menghasilkan pola data mining pada data penerimaan mahasiswa baru untuk mengetahui kemungkinan calon mahasiswa melakukan registrasi ulang atau tidak sehingga jika pengunduran diri calon mahasiswa baru dapat diketahui lebih dini diharapkan akan membantu pihak manajemen kampus dalam mengambil tindakan-tindakan yang diperlukan untuk mempertahankan calon mahasiswa tersebut.*

*Penelitian ini menggunakan dua teknik pemodelan yaitu model schema star untuk data warehouse dan model classification dengan teknik decision tree untuk data mining.*

*Keyword : Data warehouse, data mining, classification, dan 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 faktor 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.

Sejak berdiri pada tahun 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. Rumusan Masalah

Untuk mencapai tujuan dan hasil akhir yang sesuai dengan latar belakang masalah sebagaimana tercantum diatas, dan juga berdasarkan hasil analisis penulis mengidentifikasi beberapa kategori permasalahan sebagai berikut :

1. Di Universitas tempat penulis menimba ilmu belum adanya aplikasi Streaming Online.
2. Tutorial atau modul yang digunakan untuk pembelajaran di Universitas hanya berupa dokumen yang berupa tulisan.
3. Dokumentasi suatu acara di Universitas jarang dipublikasikan melalui video.

## 4. Batasan Masalah

Ada beberapa permasalahan yang penulis temui pada system informasi akademik yang ada di universitas kuningan yaitu :

1. Sistem informasi yang kurang baik sehingga pihak manajemen kesulitan untuk mendapatkan informasi.
2. Belum adanya data warehouse sehingga dipandang perlu adanya pembangunan datawarehouse untuk mendukung system informasi akademik

3. Berdasarkan data PMB taun 2003-2010 di temukan banyak sekali mahasiswa yang melakukan pendaftaran tetapi tidak melakukan registrasi ulang.
4. Pada saat perkuliahan berjalan ada beberapa mahasiswa yang cenderung pindah jurusan atau bahkan pindah fakultas.
5. Banyak mahasiswa yang ketika lulus mempunyai IPK di bawah standar kelulusan yang di harapkan

Berdasarkan rumusan masalah di atas maka peneliti membatasi masalah pada dua masalah yaitu :

- a. Perancangan data warehouse sebagai bagian dari proses persiapan penempatan data yang disesuaikan untuk keperluan bidang akademik
- b. Analisa data mining untuk mencari pola kemungkinan calon mahasiswa baru melakukan registrasi ulang atau tidak.

## 5. Metode Penelitian

### Tinjauan Pustaka

#### Data, Informasi dan Database

Sebelum kita membahas tentang data warehouse, hal yang harus dipahami terlebih dahulu yaitu pengertian tentang data, sistem informasi dan database.

##### A. Data

Menurut Steven Alter, data merupakan fakta, gambar atau suara yang mungkin atau tidak berhubungan atau berguna bagi tugas tertentu.

Menurut McLeod, data terdiri dari fakta-fakta dan angka yang secara relatif tidak berarti bagi pemakai.

Data merupakan suatu bentuk keterangan-keterangan yang belum diolah atau dimanipulasi sehingga belum begitu berarti bagi sebagian pemakai.

##### B. Informasi

Informasi adalah data yang sudah diolah sehingga menjadi lebih berguna bagi pemakai.

Informasi dapat diibaratkan sebagai darah yang mengalir di dalam tubuh manusia, seperti halnya informasi di dalam sebuah perusahaan yang sangat penting untuk mendukung kelangsungan perkembangannya,

sehingga terdapat alasan bahwa informasi sangat dibutuhkan bagi sebuah perusahaan. Akibat kurangnya informasi dalam waktu tertentu perusahaan akan mengalami ketidakmampuan mengontrol sumber daya, sehingga dalam mengambil keputusan strategis sangat terganggu. Kualitas Informasi tergantung dari tiga hal yang dominan yaitu keakuratan, tepat waktu dan relevan. (Andi Kristanto :11) Ketiga hal tersebut dapat dijelaskan sebagai berikut :

##### a. Akurat

Informasi yang dihasilkan harus bebas dari kesalahan-kesalahan yang menyesatkan bagi orang-orang yang menerima informasi tersebut

##### b. Tepat waktu

Informasi yang di terima harus tepat waktu, sebab kalau informasi yang di terima terlambat maka informasi tersebut sudah tidak berguna

##### c. Relevan

Informasi harus mempunyai manfaat bagi si penerima, sebab informasi ini akan digunakan untuk pengambilan keputusan dalam pemecahan suatu masalah

#### B. Basis Data (data base)

##### a. Pengertian Basis Data

Basis data merupakan sekelompok tabel data berisi informasi yang saling berelasi atau berhubungan. Data didalam database diproses untuk mengidentifikasi berbagai alternatif dan mengevaluasi potensinya untuk memecahkan masalah, setelah teridentifikasi solusi terbaik dilaporkan ke organisasi. Database bahkan dapat menyimpan solusi terbaik untuk suatu masalah tertentu sehingga bila dimasa depan organisasi menghadapi masalah yang sama, tersedia suatu solusi tanpa perlu memproses ulang data

##### b. Bahasa Basis Data

Bahasa basis data yang umum dipakai yaitu Microsoft acces, mysql dan sql server. Sql server merupakan salah satu database engine terpopuler dan terbaik saat ini yang dikeluarkan oleh Microsoft. Sql server telah diakui berbagai pihak dan mampu mendukung berbagai macam bahasa pemrograman berbasis windows seperti visualbasic, Delphi, maupun untuk bahasa pemrograman berbasis web seperti asp.

### C. Data Warehouse

#### 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.

Menurut Vidette Poe, *data warehouse* merupakan database yang bersifat analisis dan *read only* yang digunakan sebagai fondasi dari sistem penunjang keputusan.

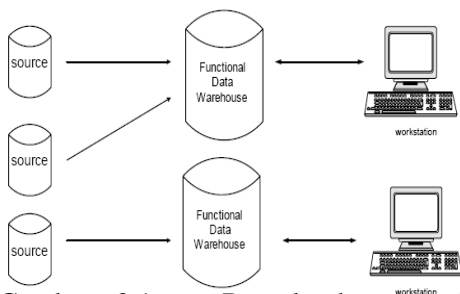
Menurut Paul Lane, *data warehouse* merupakan database relasional yang didesain lebih kepada *query* dan analisa dari pada proses transaksi, biasanya mengandung *history* data dari proses transaksi dan bisa juga data dari sumber lainnya. *Data warehouse* memisahkan beban kerja analisis dari beban kerja transaksi dan memungkinkan organisasi menggabung/konsolidasi data dari berbagai macam sumber.

Dari definisi-definisi yang dijelaskan tadi, dapat disimpulkan *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.

Berikut ini adalah tiga jenis Arsitektur *Data Warehouse* :

#### a. Functional Data Warehouse (*Data Warehouse Fungsional*)

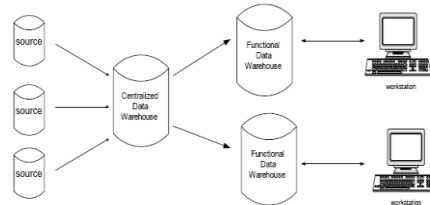
Kata operasional disini merupakan database yang diperoleh dari kegiatan sehari-hari. *Data warehouse* dibuat lebih dari satu dan dikelompokkan berdasar fungsi-fungsi yang ada di dalam perusahaan seperti fungsi keuangan (*financial*), marketing, personalia dan lain-lain.



Gambar 2.1 Bentuk data warehouse fungsional

#### b. Centralized datawarehouse (*data warehouse terpusat*)

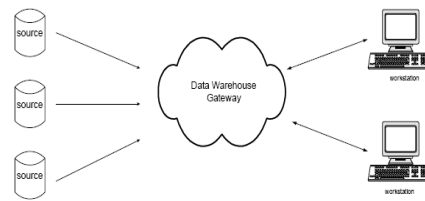
Bentuk ini terlihat seperti bentuk data warehouse fungsional, namun terlebih dahulu sumber data dikumpulkan dalam satu tempat terpusat, kemudian data disebar ke dalam fungsinya masing-masing, sesuai kebutuhan perusahaan. *Data warehouse terpusat* ini, biasa digunakan oleh perusahaan yang belum memiliki jaringan eksternal.



Gambar 2.2 Bentuk data warehouse data warehouse terpusat

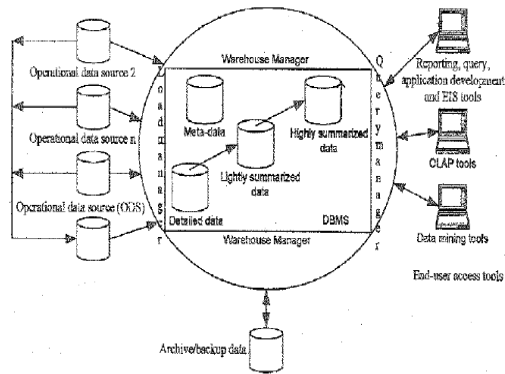
#### c. Distributed data warehouse (*data warehouse terdistribusi*)

Pada *data warehouse terdistribusi* ini, digunakan gateway yang berfungsi sebagai jembatan penghubung antara *data warehouse* dengan workstation yang menggunakan sistem beraneka ragam. Dengan sistem terdistribusi seperti ini memungkinkan perusahaan dapat mengakses sumber data yang berada diluar lokasi perusahaan (eksternal).



Gambar 2.3 Bentuk data warehouse terdistribusi

## 1. Arsitektur komponen data warehouse



Gambar 2.4 Arsitektur komponen data warehouse Conolly, T.M., Begg

Keterangan :

### a. Operational Data

Sumber data dari *data warehouse* dapat diambil langsung dari *mainframe*, basis data relasional seperti Oracle, Ms SQL server dan sebagainya. Selain itu dapat melalui *Operational Data Source (ODS)*. ODS menampung data yang diekstraksi dari sistem utama atau sumber-sumber data yang ada dan kemudian data hasil ekstraksi tersebut dibersihkan.

### b. Load manager

*Load manager* juga disebut sebagai komponen *front-end* yang bertugas melakukan seluruh operasi yang berhubungan dengan ekstraksi dan me-load data ke *warehouse*.

### c. Warehouse manager

*Warehouse manager* melakukan seluruh operasi-operasi yang berhubungan dengan kegiatan manajemen data di dalam *warehouse*. Operasi-operasi tersebut meliputi :

- Analisis terhadap data untuk memastikan konsistensi
- Transformasi dan penggabungan sumber data dari tempat penyimpanan sementara menjadi tabel-tabel *data warehouse*.
- Penciptaan indeks-indeks dan view berdasarkan tabel-tabel dasar
- Melakukan denormalisasi dan agregasi jika diperlukan
- *Backing-Up* dan mengarsipkan data

### d. Query manager

*Query manager* juga disebut komponen *back-end*, melakukan operasi-operasi yang berhubungan dengan manajemen *user queries*. Operasi-operasi yang dilakukan oleh komponen ini termasuk mengarahkan *query* kepada tabel-tabel yang tepat dan menjadwalkan eksekusi dari *query* tersebut.

### e. End-user Access Tools

Prinsip atau tujuan utama dari dibangunnya *data warehouse* adalah untuk menyediakan informasi bisnis kepada *user-user* untuk dapat melakukan pengambilan keputusan secara cepat dan tepat. *User* ini berinteraksi dengan *warehouse* melalui *end-user access tools*. *Data warehouse* harus secara efisien mendukung secara khusus kebutuhan *user* serta secara rutin melakukan analisis. Performa yang baik dapat dicapai dengan merencanakan dahulu keperluan-keperluan untuk melakukan *joins, summations* dan laporan-laporan per periode dengan *end-users*.

Keterangan:

#### 1. System Engineering

Melakukan pengumpulan data dan penetapan kebutuhan semua elemen untuk pembangunan sistem.

#### 2. Analysis

Melakukan analisis terhadap permasalahan yang dihadapi dan menetapkan kebutuhan perangkat lunak, fungsi performansi dan *interfacting*.

#### 3. Design

Menetapkan domain informasi untuk perangkat lunak, fungsi dan *Interfacting*.

#### 4. Coding

Pengkodean yang mengimplementasi hasil desain kedalam kode atau bahasa yang dimengerti oleh mesin komputer dengan menggunakan bahasa pemrograman tertentu.

#### 5. Testing

Melakukan pengujian kebenaran logic dan fungsionalitis. Disinilah akan diketahui kekurangan-kekurangan yang menyebabkan kerusakan.

#### 6. Maintenance

Menangani perngkat lunak yang sudah selesai supaya dapat berjalan dan terhindar dari gangguan-gangguan yang menyebabkan kerusakan.

## Operasi OLAP

Operasi yang dapat dilakukan oleh OLAP adalah sebagai berikut :

### 1. Roll UP dan Drill down

Drill down mengacu pada kemampuan untuk menampilkan data detail penyusun suatu *summerization* data, dengan memandu pengguna untuk memperoleh data yang lebih detail, pengguna dapat memahami informasi *summerize* lebih baik. Drill down ini bisa digunakan untuk menjawab pertanyaan atas suatu kasus tertentu. Misalnya untuk menjawab pertanyaan ketika sebuah summery number (rata-rata atau jumlah) dibawah atau diatas harapan. Roll up adalah operasi untuk melihat data global di sepanjang level hirarky, agar user dapat melihat data rangkuman

### 2. Slicing dan Dicing

Slicing dan dicing adalah operasi untuk melihat data sebagai visualisasi dari kubus. Dengan slicing dan dicing pengguna dapat melihat data dari beberapa perspektif. Pengguna dapat mengambil informasi dasar, lalu mengelompokanya sesuai keinginan dan analisa ulang dari data agregat berdasarkan dimensi-dimensi yang diinginkan ([immon 2005], 234 dengan perubahan). Data aggregated merupakan data praperhitungan (precalculated) dalam bentuk rangkuman data (data summerized ) sehingga query dalam kubus (cube) lebih cepat.

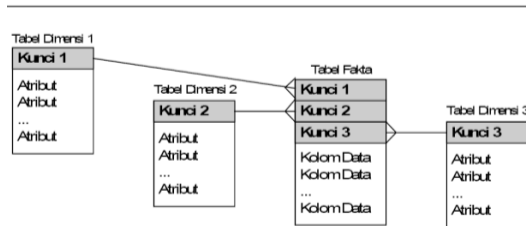
Slicing memotong kubus sehingga dapat memfokuskan pada perspektif yang spesifik (pada suatu dimensi). Sedangkan dicing memberikan kemampuan untuk melihat pemilihan data pada dimensi atau lebih yaitu dengan merotasi cube pada perspektif yang lain sehingga pengguna dapat melihat labih spesifik terhadap data yang dianalisa

### Teknik Pemodelan Data Warehouse

Terdapat dua bentuk pemodelan (schema), yaitu :

#### 1. Star

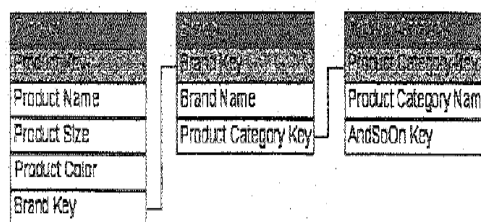
Adalah bentuk model yang mempunyai bentuk bintang, dimana tabel fakta terdapat di tengah dan tabel dimensi mengelilinginya



Gambar 2.5 Schema model star

### 2. Snowflaked

Model ini menyerupai model star, jika pada model star memiliki tabel fakta, maka disini jumlah tabel faktanya lebih dari satu dan berbagi dimensi lainnya



Gambar 2.6 Schema model snowflaked

## Data Mining

### 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 elarning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban,dkk.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 (Iarose, 2005)

Data mining adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaanya. (pramudiono, 2006)

Dari definisi yang telah disampaikan, hal penting yang terkait dengan data mining adalah:

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada
2. Data yang di proses berupa data yang sangat besar
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Hubungan yang dicari dalam data mining dapat berupa hubungan antara dua atau lebih dalam satu dimensi. Misalnya dalam dimensi produk, kita dapat melihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan juga dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek (ponniah, 2001). Sementara itu, penemuan pola merupakan keluaran lain dari data mining, misalkan sebuah perusahaan yang akan meningkatkan fasilitas kartu kredit dari pelanggan, maka perusahaan akan mencari pola dari pelanggan yang ada untuk mengetahui pelanggan yang potensial dan pelanggan yang tidak potensial.

## M. Implementasi Antarmuka

### Weka

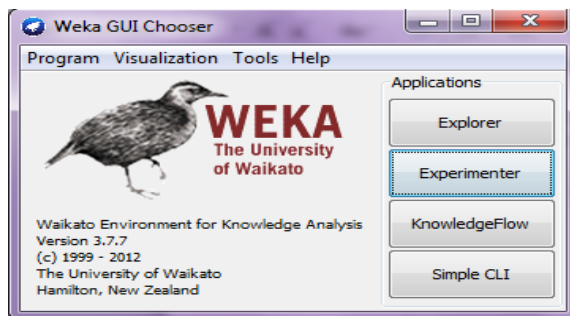
#### A. Classifier RJ48

Aplikasi weka bersifat open source, kita bisa mengunduhnya secara gratis melalui alamat <http://www.cs.waikato.ac.nz/ml/weka/>.

Aplikasi weka digunakan untuk membentuk pohon keputusan. Software yang penulis gunakan dalam kasus ini menggunakan weka versi 3.7.7.

Adapun langkah-langkah yang digunakan untuk menjalankan aplikasi weka adalah sebagai berikut:

1. Klik Start → All program → weka 3.7.7 maka akan menghasilkan tampilan Weka GUI chooser seperti di berikut ini :

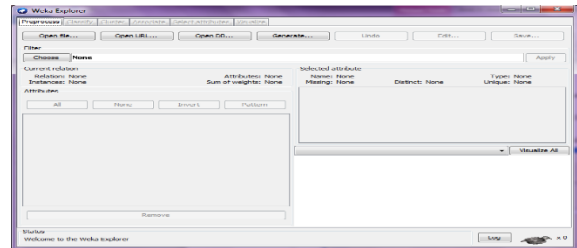


Gambar 4.24 Aplikasi weka

Weka mempunyai tiga buah aplikasi yang bisa digunakan yaitu explorer, experimenter, KnowledgeFlow, dan simple CLI.

Dalam laporan ini penulis lebih dominan menggunakan Explorer.

2. Langkah selanjutnya adalah membuka file

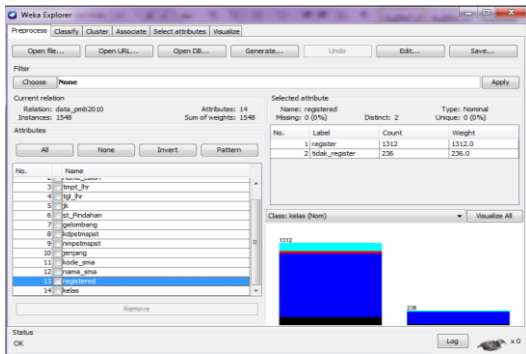


Gambar 4.25 Weka Explorer

Format file data yang digunakan dalam weka mempunyai format \*.CVS (comma separated value), \*. Arff (Attribute-Relation File Format), file format C4.5 \*.data, dan \*.bsi (binary serialized instances (BSI)).

Tombol open file digunakan untuk membuka file dengan extension yang disebutkan diatas. Tombol open database digunakan untuk membuka data yang terdapat di dalam database.

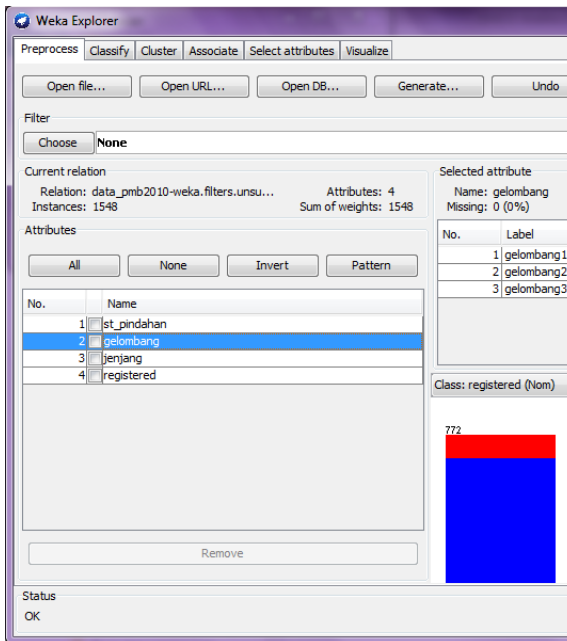
3. Langkah selanjutnya adalah klik open file. Disini penulis menggunakan data \*.csv. dikarenakan beberapa type data dalam Sql server 2008 tidak dikenali didalam weka 3.7.7. sehingga ketika mencoba melakukan open database terjadi error. Adapun untuk membuat file dalam bentuk \*.csv caranya adalah dengan mengkonversi data PMB 2010/2011 yang ada didalam sql 2008 ke dalam Microsoft excel terlebih dahulu kemudian di save ke dalam file \*.csv. Kemudian klik open file kemudian pilih file yang sudah berbentuk .csv tadi. Maka akan muncul tampilan seperti berikut ini :



Gambar 4.26 Sql viewer

Pilih variable st\_pindahan, gelombang, panjang, registered yang akan di analisis. Untuk data lain yang tidak diperlukan klik remove untuk membuang data

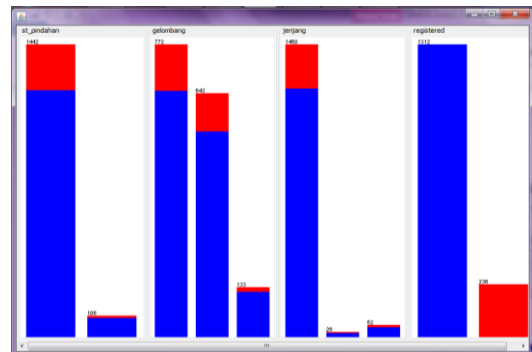
- Hasilnya dapat dilihat pada weka seperti berikut :



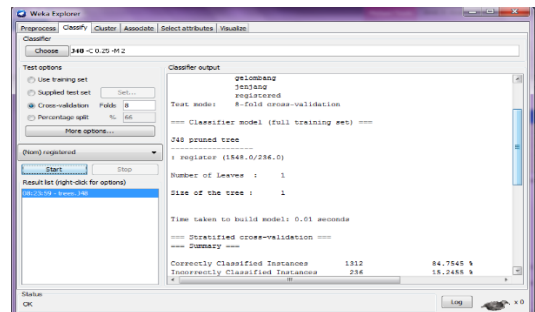
Gambar 4.27 Hasil open file

Informasi yang dapat diperoleh dari gambar diatas adalah jumlah data untuk masing-masing atribut ditunjukkan dengan grafik diagram batang dan selected atribut sebagai detailnya. Pada current relation data menampilkan instance (jumlah data) yang akan diproses yaitu sebanyak 1548. pada data diatas dapat dilihat bahwa mahasiswa yang melakukan pendaftaran pada gelombang 1 sebanyak 772 orang, gelombang 2 sebanyak 643 orang dan pada gelombang 3 sebanyak 133 orang.

Visualisasi memuat informasi atribut dengan tampilan grafik diagram batang dari keseluruhan atribut :



Gambar 4.28 Grafik diagram batang untuk keseluruhan atribut.

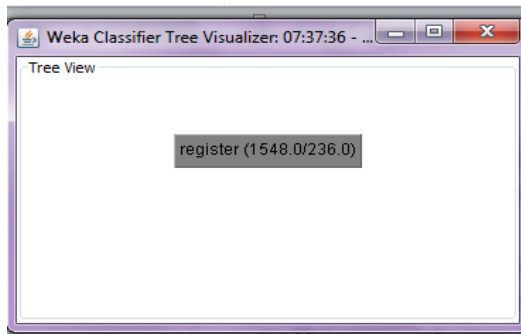


Gambar 4.29 hasil classify J48

Teknik decision tree yang akan digunakan adalah C4.5 maka Pada classifier klik tombol coose lalu pilih tree J48. pilih cross-validation folds 8 kemudian klik tombol start maka akan muncul informasi dari pembentukan pola decision tree. dari hasil diatas di dapat informasi bahwa mahasiswa yang melakukan registrasi adalah 1548, artinya semuanya di prediksi melakukan registrasi ulang, tetapi sebagian tidak sesuai dengan prediksi pola yang sudah dibuat yaitu sebanyak 236 orang.

- Klik kanan pada result (right-click for option) pilih visualize tree maka akan muncul hasil visualisasi decien tree seperti berikut :

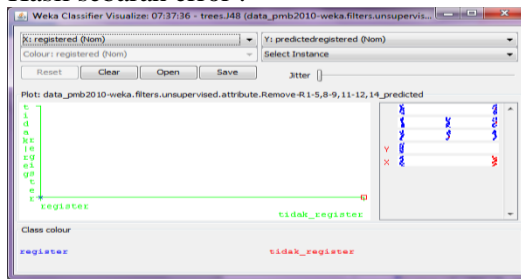




Gambar 4.30 tree viewer

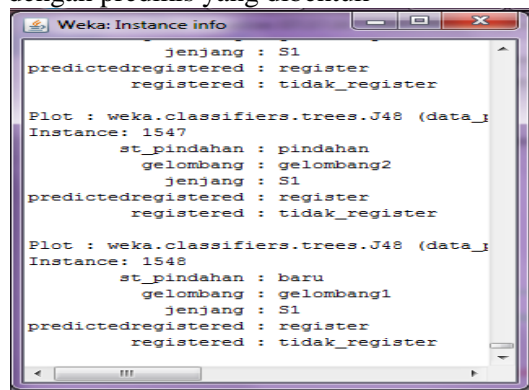
Pada data tersebut tidak terdapat pembentukan cabang pohon karena di prediksi semua mahasiswa akan melakukan registrasi ulang. Dari jumlah 1548 mahasiswa yang di duga akan melakukan registrasi ulang, 236 diantaranya tidak sesuai dengan hasil prediksi. Untuk melihat sebaran data yang error dapat dilakukan dengan klik kanan pada result (right-click for option) lalu pilih visualize tree error.

Hasil sebaran error :



Gambar 4.31 Visualize tree error

klik pada data tidak register, maka akan muncul sebaran data yang tidak sesuai dengan predikis yang dibentuk



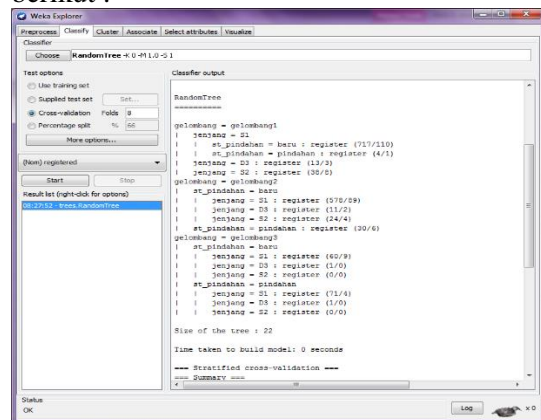
Gambar 4.32 instance info

## B. Classifier Random Tree

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.

Hasil dari Random tree adalah sebagai berikut :



Gambar 4.33 hasil random tree

## 1. Penutup 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.

## 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 Sucahyo 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 Intelegent 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.

