

SISTEM INFORMASI KLASIFIKASI LATAR BELAKANG MAHASISWA BERDASARKAN PRESTASI AKADEMIKNYA MENGGUNAKAN TEKNIK DECISION TREE

STUDI KASUS : UNIVERSITAS KUNINGAN

Aah Sumiah¹, Siti maesyaroh², Fauziah³

Salah Satu proses akademik untuk masuk ke sebuah perguruan tinggi adalah melalui proses Penerimaan Mahasiswa Baru. Dalam proses ini diberikan ujian dan proses tes potensi akademik yang digunakan untuk mengukur kemampuan mahasiswa dalam berbagai bidang ilmu pengetahuan. Jika calon mahasiswa bisa melewati ujian dengan hasil yang bagus maka calon mahasiswa akan diterima sebagai mahasiswa sesuai dengan fakultas yang pilihnya dan jika tidak maka calon mahasiswa dinyatakan gagal dalam ujian. Proses seleksi calon mahasiswa baru yang dilakukan oleh universitas kuningan meliputi tes bahasa inggris, tes kemampuan potensi akademik, dan tes wawancara. Yang menjadi pertanyaan adalah apakah ketika mahasiswa mampu mengikuti ujian tes yang diselenggarakan pada awal Penerimaan Mahasiswa Baru dengan hasil bagus apakah kemudian berarti mahasiswa tersebut juga mampu menyelesaikan studinya dengan hasil yang bagus juga. Hal ini lah yang kemudian mendasari penelitian ini.

Tujuan Penelitian ini adalah untuk menghasilkan pola data mining pada Penerimaan Mahasiswa Baru untuk memprediksi prestasi akademik mahasiswa berdasarkan latar belakangnya dan mengimplementasikannya ke dalam sebuah *sistem informasi* berupa program aplikasi yang dapat digunakan pihak lembaga untuk mengambil keputusan menerima atau tidaknya calon mahasiswa.

Penelitian ini menggunakan *data mining* model *classification* dengan teknik *decision tree* menggunakan *algoritma C.45* dan implementasi program menggunakan *visual basic.net*.

Keyword : *sistem informasi, data mining, classification, decision tree, algoritma C.45, visual basic.net*

1. Pendahuluan

Salah Satu proses akademik untuk masuk ke sebuah perguruan tinggi adalah melalui proses Penerimaan Mahasiswa Baru. Dalam proses ini diberikan ujian dan proses tes potensi akademik yang digunakan untuk mengukur kemampuan mahasiswa dalam berbagai bidang ilmu pengetahuan. Jika calon mahasiswa bisa melewati ujian dengan hasil yang bagus maka calon mahasiswa akan diterima sebagai mahasiswa sesuai dengan fakultas yang pilihnya dan jika tidak maka calon mahasiswa dinyatakan gagal dalam ujian. Proses seleksi calon mahasiswa baru yang dilakukan oleh universitas kuningan meliputi tes bahasa inggris, tes kemampuan potensi akademik, dan tes wawancara.

Yang menjadi pertanyaan adalah apakah ketika mahasiswa mampu mengikuti ujian tes yang diselenggarakan pada awal Penerimaan Mahasiswa Baru dengan hasil bagus apakah kemudian berarti mahasiswa tersebut juga mampu menyelesaikan studinya dengan hasil yang bagus juga. Hal ini lah yang kemudian mendasari penelitian ini.

2. Tinjauan Pustaka

A. 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 elarning 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.

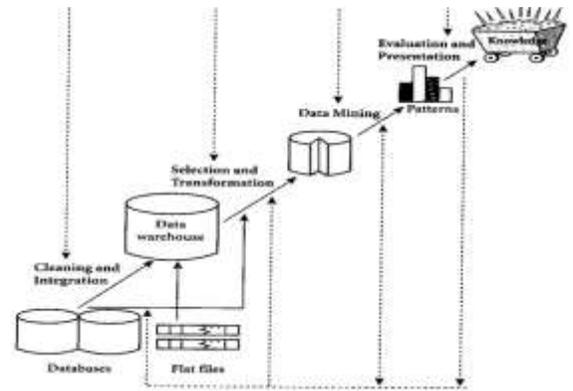
3. Pengelompokan data mining

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 klasifikasi 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 lainya 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

B. SQL Server

Sql server 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.

Ada dua fasilitas SQL yang digunakan dalam penelitian ini yaitu Enterprise Manager, dan Query Analyzer.

- a. Enterprise manager merupakan utility bantu utama dalam SQL Server. Enterprise manager merupakan

interface terintegrasi di dalam SQL Server yang mampu melakukan hampir semua fungsi dan perintah yang kita butuhkan dalam SQL Server.

- b. Query analyzer merupakan wadah dimana data penelitian ini disimpan. Fasilitas ini juga menyediakan fungsi-fungsi untuk melakukan proses-proses yang berhubungan dengan pembuatan table, atribut, dan tipe data.

C. 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.

D. Tinjauan Studi

Beberapa peneliti telah melakukan proses penelitian yang sama Diantaranya:

1. Penelitian yang dilakukan oleh Gede Rasben Dantes dan ketut agustini, Universitas Pendidikan Ganesha dengan judul “ Sistem Informasi Kalasifikasi Tingkat Prestasi akademik Mahasiswa Berdasarkan Seleksi Ujian masuk Perguruan Tinggi dengan algoritma K-means”. Pada penelitian ini : hasil penelitian ini berupa system informasi klasifikasi tingkat prestasi mahasiswa berdasarkan seleksi ujian masuk ke perguruan tinggi dengan teknik clustering..
2. Penelitian yang dilakukan oleh Ni Made Pritty wirati dan P.H Prima Rosa dengan judul “klasifikasi Latar belakang mahasiswa Berdasarkan Prestasi akademiknya Dengan Metode Pohon Keputusan” Pada penelitian ini : hasil penelitian adalah klasifikasi latar belakang mahasiswa berdasarkan prestasi akademiknya dengan

mengimplementasikan teknik penambangan data mining mempergunakan metode pohon keputusan ID3.

3. Penelitian yang dilakukan oleh Eko Nur Wahyudi, Arif Jananto dan Narwati dengan judul “Analisa Fropil data mahasiswa baru terhadap program studi yang di pilih di perguruan tinggi swasta Jawa tengah dengan menggunakan teknik data mining” Pada penelitian ini : hasil penelitian adalah data mining untuk memprediksi peminatan trend mahasiswa baru berdasarkan bidang ilmu dan program studi yang dipilihnya.

E. 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).

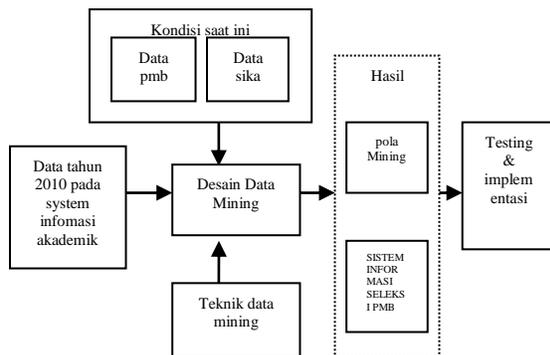
F. Fasilitas Layanan

Berdasarkan penelitian, Universitas Kuningan sudah memiliki memiliki system 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

G. Kerangka Konsep

Sesuai dengan tujuan penelitian yang telah di tetapkan maka peneliti ingin melakukan analisa data mining pada sistem informasi akademik yang ada dengan konsep yang menurut penulis cocok digunakan di Universitas Kuningan. Sehingga pihak manajemen kampus dapat mengambil keputusan dengan cepat.

kerangka konsep di atas dapat di gambarkan seperti berikut ini :



Gambar 2.3 Kerangka pemikiran

Kerangka Konsep Diatas Dapat jelaskan sebagai berikut :

Desain data mining di buat berdasarkan data pada tahun 2010 yang diambil dari system informasi

akademik. Data mining ini di desain dengan menganalisis data operasional yang sedang berjalan berdasarkan data PMB (Penerimaan Mahasiswa Baru) dan Data SIKA (Sistem Informasi Akademik) serta memasukan teknik data mining yang sesuai. Hasil dari perancangan diharapkan akan terbentuk sebuah pola data mining yang kemudian pola tersebut di implementasikan menjadi sebuah sistem informasi berupa program aplikasi seleksi penerimaan mahasiswa baru.

H. Hipotesis

Hipotesis dibuat berdasarkan dasar teori dan kerangka konsep berpikir yang telah dikemukakan diatas. Beberapa prediksi hasil penelitian adalah sebagai berikut

Hipotesis Pertama :

Teknik *decision tree* dapat memprediksi prestasi akademik mahasiswa berdasarkan latar belakangnya.

Hipotesis Kedua :

Teknik *decision tree* tidak dapat memprediksi prestasi akademik mahasiswa berdasarkan latar belakangnya.

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

2. Metode Pengumpulan Data

Data untuk penelitian ini diambil dari data primer dan data skunder. 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.

3. Teknik Analisis Data

Teknik analisa yang dilakukan yaitu Analisis Data Mining Dan hasilnya diimplementasikan menjadi sebuah Program Aplikasi penilaian Penerimaan mahasiswa baru.

4. Hasil Penelitian

4.1 Sumber Data

Data yang di ambil untuk penelitian ini terdiri dari data akademik mahasiswa, data penerimaan mahasiswa baru dan data wisuda mahasiswa. Adapun data yang diambil yaitu data mahasiswa angkatan 2008-2009 dan data penerimaan mahasiswa baru tahun 2008-2009. Data tersebut diambil dengan asumsi bahwa angkatan tahun 2008 -2009 mahasiswa sudah menempuh pendidikan dan lulus sesuai waktunya.

Selain itu data di ambil dari data nilai tes PMB 2008-2009 dan data wisuda mahasiswa tahun 2011 dimana data tersebut masih berbentuk file excel.

Adapun table yang diambil dari database sikaV yang digunakan oleh peneliti untuk proses perancangan data adalah table-table berikut ini :

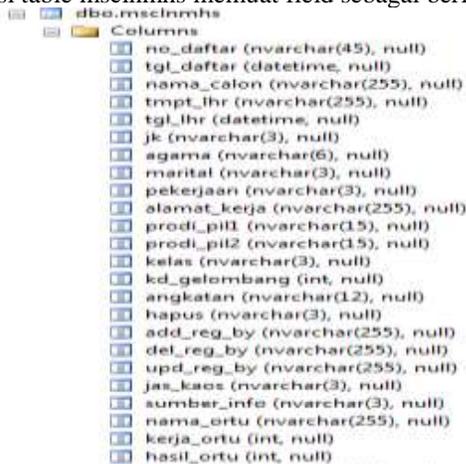
Tabel 4.1 Tabel Akademik

Nama Tabel	Primary Key	Keterangan
Tabel msclnmhs	No_daftar	Berisi data calon mahasiswa dan registrasinya
Table msmhs	Nimmhs	Berisi data mahasiswa
Table v asal_sma_s1_d3_full	No_daftar	Berisi data mahasiswa dan nama SMU asal mahasiswa
Table msfak	Id_fak	Berisi nama fakultas
Table mspst	Prodi_pil	Berisi nama program studi

Untuk table akademik yang sudah ada di dalam database sikaV tidak perlu lagi melakukan transformasi data. tabel tersebut adalah sebagai berikut :

A. Tabel Msclnmhs

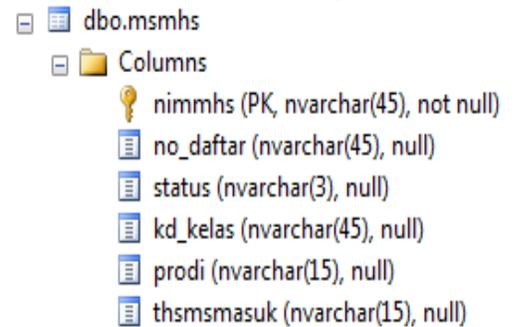
Isi table msclnmhs memuat field sebagai berikut :



Gambar 4. 1 Desain tabel msclnmhs

B. Data msmhs

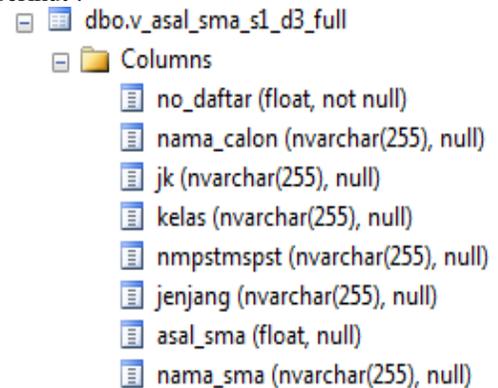
Data mahasiswa memuat field seperti berikut ini :



Gambar 4.2 Desain Tabel msmhs

C. Data v_asal_sma_s1_d3_full

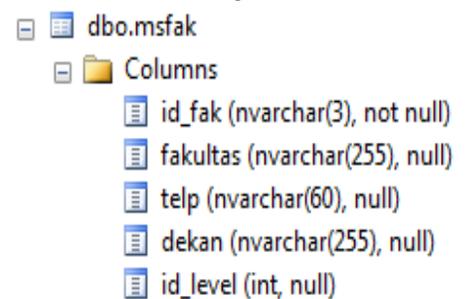
Isi table v_asa_sma_s1d3_full memuat field sebagai berikut :



Gambar 4.3 Desain Tabel v_asal_sma_s1_d3_full

D. Data Msfak

Data msfak memuat field sebagai berikut :



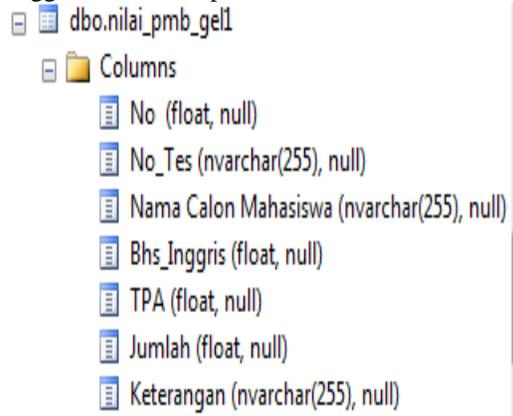
Gambar 4.4 Desain Tabel msfak

Untuk data-data yang masih berbentuk file excel, Sebelum data diolah menggunakan sistem yang akan dibuat, terlebih dahulu dilakukan pemrosesan data awal dengan melakukan transformasi data dari excel ke dalam database.

Adapun table tersebut adalah sebagai berikut :

a. Data Nilai tes PMB

Data nilai tes pmb memuat field sebagai berikut : no, nomor tes, nama calon mahasiswa, nilai bhs inggris, nilai TPA, jumlah dan keterangan. Field tersebut di transformasikan ke dalam database sehingga berbentuk seperti berikut ini :

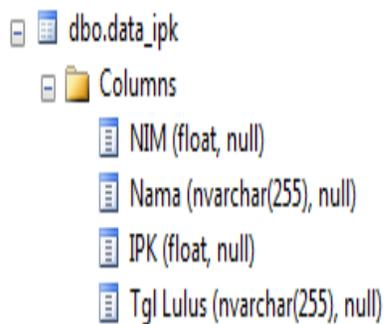


Gambar 4.5 Desain Tabel nilai PMB

b. Data Ipk

Data wisuda memuat field nim, nama, IPK, dan tgl lulus

Field tersebut di transformasikan ke dalam database sehingga berbentuk seperti berikut ini:



Gambar 4.6 Desain Tabel data IPK

A. Pembersihan data

Data yang tersimpan dalam database dilakukan pembersihan data. Pada tahap ini dilakukan pembersihan data terhadap semua data dan field yang akan digunakan. Pembersihan data dilakukan untuk data yang tidak lengkap dan bernilai null. Hal ini dilakukan karna jika data ada yang bernilai null maka akan terjadi kesalahan dalam penyimpanan data. Field-field yang tidak digunakan untuk penelitian dihapus dan tidak akan dipakai untuk tahap selanjutnya. Pada tahap ini tabel-tabel yang sudah dibersihkan di rename guna keperluan penelitian.

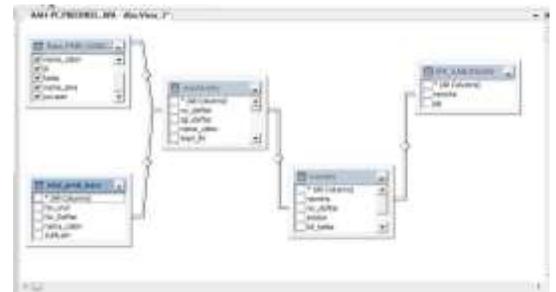
Tabel-tabel tersebut adalah sebagai berikut ini :

Tabel 4.2 Tabel yang di rename

No	Nama Tabel awal	Rename New Tabel
1	Data IPK	IPK_gabungan
2	Nilai_PMB_gel1	Nilai_PMB_baru
3	v asal_sma_s1_d3_full	Data_Pmb_Gabungan

B. Integrasi data

Pada tahap ini dilakukan proses penggabungan data clnmhs, msmhs, data nilai_pmb_baru dan IPK_gabungan. Data yang didapat di satukan ke dalam satu table dengan membuat table baru bernama Data_gabungan melalui proses view pada database.



Gambar 4.7 tabel View

C. Penyeleksian Data

Pada tahap ini dilakukan penyeleksian data yang kurang relevan dengan penelitian. Berdasarkan data yang di dapat data yang diperoleh sebanyak 531 data saja. Hal ini disebabkan tidak lengkapnya data sekolah calon mahasiswa baru, adanya mahasiswa yang belum lulus dan mahasiswa yang tidak melakukan registrasi atau mengundurkan diri pada saat PMB, Sehingga setelah dilakukan penggabungan data hanya 531 record mahasiswa yang mempunyai data yang lengkap

D. Transformasi Data

Pada tahap ini akan dilakukan proses perubahan data mentah menjadi data yang mudah dikelola. berdasarkan data yang didapat, sebagian data berupa nilai dan angka sehingga untuk memudahkan proses penambangan data (data mining) maka data tersebut dikelompokkan bedasarkan jangkauan tertentu. Hal ini untuk memudahkan pengkategorian nilai. Proses transformasi data dilakukan untuk IPK_gabungan dan data nilai_PMB_baru

Untuk aturan penilaian IPK adalah sebagai berikut :

Table 4.3 Aturan penilain IPK

No	Aturan	Kategori
1	IPK \geq 3.00	Baik
2	IPK $<$ 3.00	Kurang

Untuk aturan penilaian nilai PMB adalah sebagai berikut :

Tabel 4.4 Aturan nilai PMB

No	Aturan	Kategori
1	Jumlah_tes \geq 70	A
2	Jumlah_tes $<$ 70	B

4.1. Data mining

Proses selanjutnya dilakukan proses penambangan data

A. Pemilihan Variabel

Berdasarkan hasil pengolahan data diatas, maka di peroleh informasi data_gabungan sebagai berikut :

Table 4.5 Tabel Informasi data_gabungan

No	Aturan	Kategori	Keterangan
1	No_daftar	Berisi nomor daftar calon mahasiswa	
2	Nama_calon	Berisi nama calon mahasiswa	
3	Fakultas	Berisi fakultas mahasiswa	FKIP, FKOM, FHUT, dan FE
4	Jk	Berisi Jenis kelamin	P, L
5	Kelas	Berisi Kelas Mahasiswa	K, R
6	Asal_sma	Memuat Asal sma calon mahasiswa baru	Swasta, Negeri
7	Jenis_sma	Memuat jenis_sma	SMA, SMK, MA, atau lainnya
8	Jurusan_Sma	Memuat jurusan sma mahasiswa	IPA, IPS, lainnya
9	Jumlah_tes	Memuat jumlah nilai tes PMB	Jumlah \geq 70 Jumlah $<$ 70
10	IPK	Memuat IPK mahasiswa	IPK \geq 3.00 IPK $<$ 3.00

Dari data-data tersebut, kolom yang diambil sebagai variable keputusan adalah kolom IPK, sedangkan kolom yang diambil sebagai variable penentu dalam pembentukan pohon keputusan adalah kolom jk , kelas, asal_sma, jurusan_sma, jns_sma, fakultas, dan jumlah. variable-variabel tersebut dipilih dengan pertimbangan

bahwa jumlah nilai variabelnya tidak banyak sehingga diharapkan calon mahasiswa yang masuk dalam suatu klasifikasi nilai variable tersebut cukup banyak. Sedangkan no_daftar dan nama calon tidak dipilih karena mempunyai nilai yang cukup banyak.

B. Melakukan Preprocessing

Berdasarkan pemilihan variable-variabel di atas, maka format data akan dibuat menjadi seperti berikut ini :

Tabel 4.6 Format data preprocessing

Jk	Kelas	Asal_sma	Jurusan_sma	Jns_sma	Jumlah_tes	Fakultas	IPK
L	K	Swasta	IPA	SMA	A	FKOM	Baik
L	R	Negeri	IPS	SMK	B	FKIP	Kurang
P	K	Swasta	IPS	MA	B	FHUT	Baik
P	R	Negeri	Lainya	Lainya	A	FE	Baik

Berdasarkan data yang sudah di lakukan pra proses, langkah selanjutnya adalah menghitung nilai entropy dan gain untuk membentuk pohon keputusan.

C. Perhitungan Nilai Entropy dan Gain

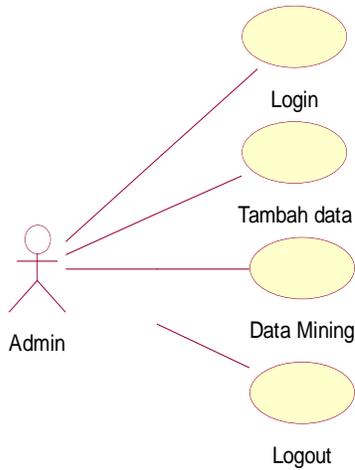
Untuk menghitung nilai gain, hal pertama yang harus dilakukan adalah menghitung total jumlah kasus, jumlah kasus untuk keputusan IPK \geq 3.00 (baik) dan jumlah kasus untuk keputusan IPK $<$ 3.00 (kurang) dan entropy dari semua kasus , dan kasus yang dibagi berdasarkan atribut.

E. Perancangan Sistem

Dalam melakukan proses perancangan ini, peneliti menggunakan diagram-diagram UML untuk memberikan gambaran terintegrasi terhadap sistem yang akan dibangun. Beberapa diagram UML yang penulis gunakan yaitu use case diagram, Aktivitas Diagram, Sekuensial Diagram dan Kelas Diagram. Adapun proses Perancangannya adalah sebagai berikut ini :

4.6.1 Diagram Use Case

Use case adalah fungsionalitas atau persyaratan-persyaratan sistem yang harus di penuhi oleh sistem yang akan dikembangkan tersebut menurut pandangan pemakai sistem. Sedangkan actor bisa berupa orang, peralatan, atau sistem lain yang berinteraksi terhadap sistem yang akan dibangun. [Sholiq 2010]. Dibawah ini merupakan penggambaran umum mengenai use case yang terdapat dalam sistem informasi klasifikasi latar belakang mahasiswa berdasarkan prestasi akademiknya .



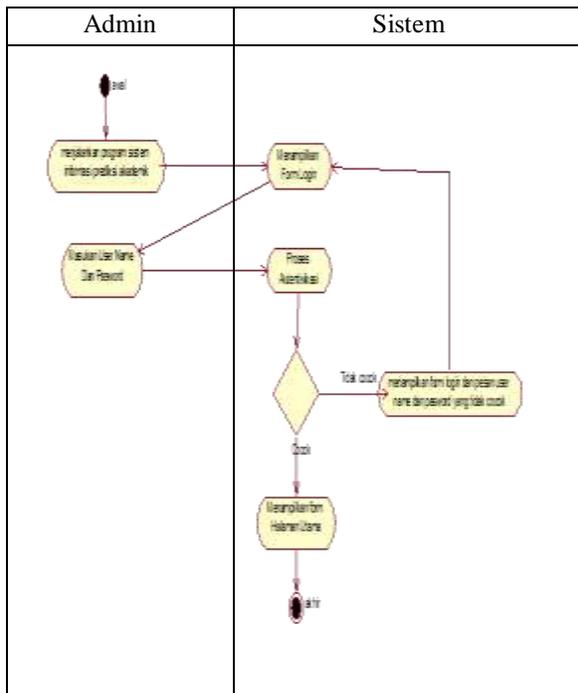
Gambar 4.16 Use case prediksi prestasi mahasiswa

Dibawah ini merupakan penggambaran umum mengenai use case yang terdapat dalam sistem:

4.6.3 Diagram Aktivitas

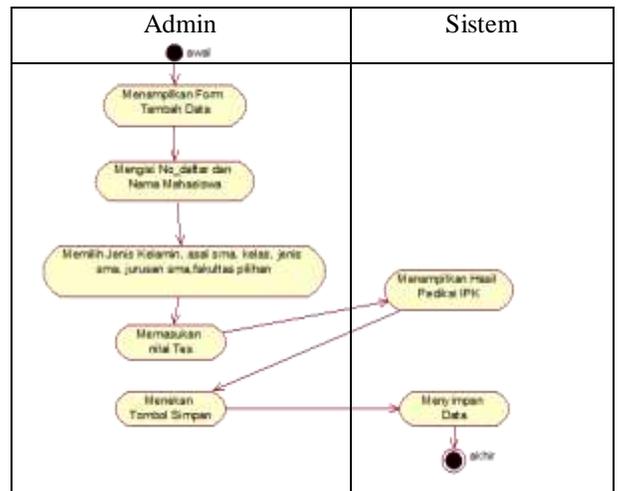
Diagram aktivitas merupakan diagram yang menjelaskan aktivitas pengguna dengan program. Berikut ini adalah diagram aktivitas pada pengolahan data akademik dan PMB

a. Diagram aktivitas login



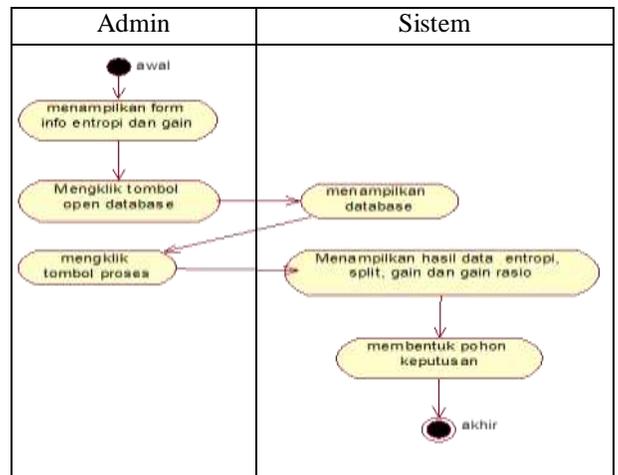
Gambar 4.17 Diagram Aktivitas login

b. Diagram aktivitas tambah data



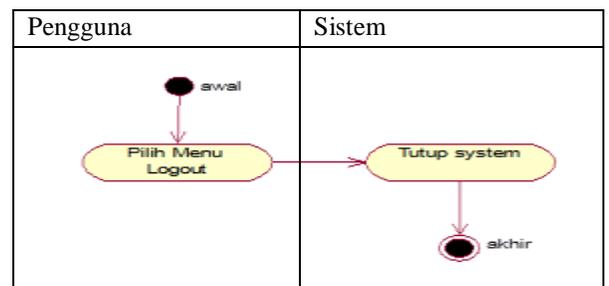
Gambar 4.18 Diagram Aktivitas Tambah Data

c. Diagram aktivitas Data mining



Gambar 4.19 Diagram Aktivitas Data Mining

d. Diagram aktivitas logout



Gambar 4.20 Diagram Aktivitas Logout

a. Menu Utama



Gambar 4.32 Tampilan Menu Utama

Untuk mengendalikan jalanya program peneliti membuat sebuah form induk yang dibuat dengan menggunakan MDIform dengan nama halaman_utama. Form ini terdiri dari 3 buah menu utama yang terdiri dari File, Data mining, dan Exit. Menu file terdiri dari sub menu tambah data sedangkan data mining terdiri dari sub menu info entropi , gain rasio serta sub menu pohon keputusan. Dan exit terdiri dari sub menu logout. Tugas utama dari form ini adalah mengendalikan seluruh jalanya program dan melakukan pemanggilan pada form-form anak.

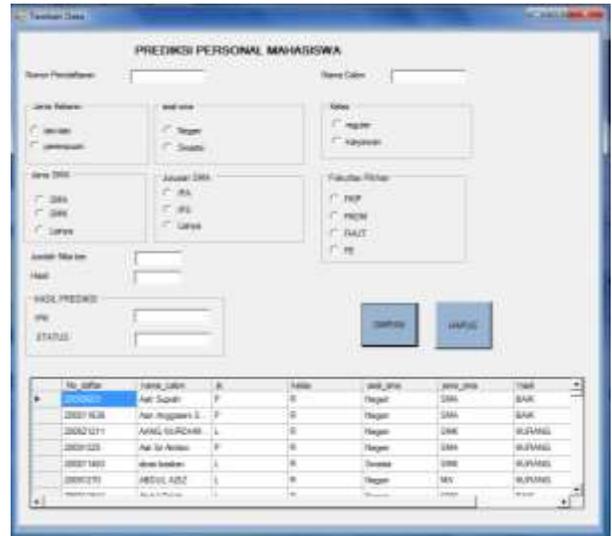
b. Form Login



Gambar 4.33 Tampilan Form Login

Form ini merupakan form yang digunakan untuk dapat memasuki program Sistem informasi prediksi prestasi akademik. Form ini terdiri atas input user name dan password. Hanya administrator yang telah diberi hak akses yang dapat memasuki program ini

c. Form Tambah Data



Gambar 4.34 Tampilan Form Tambah Data

Form ini digunakan untuk melakukan input data mahasiswa, melakukan prediksi prestasi mahasiswa, penyimpanan data, pencarian data serta melakukan penghapusan data. Secara umum Form ini bertugas untuk menyimpan data mahasiswa dan prediksi nilai IPK mahasiswa pada data survey kedalam data base.

d. Form info Entropi Dan Gain



Gambar 4.35 Tampilan Form Info entropi Dan Gain

Form ini digunakan untuk menampilkan database dan melakukan penghitungan entropi, split dan gain rasio serta menampilkan pohon keputusan. Data yang muncul akan ditampilkan jumlah data nya kemudian ditampilkan hasil perhitungan entropi , gain, split dan gain rasio dengan mengklik tombol proses.

BAB V PENUTUP

5.1 Kesimpulan.

1. Informasi *knowledge discovery* dari *data mining* dengan metode *classification* yang menggunakan pendekatan *decision tree* dapat menjawab apa yang menjadi kebutuhan manajemen yaitu memprediksi tingkat prestasi mahasiswa dengan akurasi yang cukup baik serta menunjukkan *pattern* yang selama ini tidak tampak pada system yang sudah ada.
2. 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.

5.2 Saran

1. Mengingat Universitas Kuningan adalah lembaga pendidikan yang sedang berkembang, maka untuk memperoleh informasi yang lebih optimal, sebaiknya pihak lembaga menggunakan data mining pada system informasi akademik untuk menunjang proses pengambilan kebijakan.
2. Setelah melakukan penelitian pada data PMB dan korelasi data IPK pada data akademik mahasiswa, akhirnya peneliti berhasil membuat sebuah program yang diberi nama *Sistem Informasi prediksi prestasi akademik* . Bagi pembaca yang ingin mempelajari program yang peneliti buat diharapkan juga dapat mengembangkannya sehingga lebih baik dari program yang penulis buat .

DAFTAR PUSTAKA

- [Azimah 2007] Azimah Ariana Dan Suchayo Yudho Giri, “Penggunaan Data Warehouse Dan Data Mining Untuk Data Akademik. Sebuah Studi Kasus Pada Universitas Nasional”, 2007
- [Gede 2008] Gede Rasben Dantes, Ketut Agustini, “ Sistem Informasi Klasifikasi Tingkat Prestasi Akademik Mahasiswa Berdasarkan Seleksi Ujian Masuk Perguruan Tinggi Dengan algoritma K-means “, 2008
- [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 in Data :An Introduction to data mining. John Willey & Sons.inc, 2005
- [McLeod 2004] McLeod,Jr. Raymond , “Sistem Informasi Manajemen”., Indeks., 2004
- [Pritty 2008] Ni Made Pritty dan P.H. Prima Rosa, “ Klasifikasi Latar Belakang Mahasiswa Berdasarkan Prestasi akademiknya Dengan Metode pohon Keputusan”. 2008
- [Ramakrishnan 2003] Ramakrishnan, Raghu and Gehrke Johannes, “Database Management Systems 3rd Edition”, 2003, McGraw-Hill, New York, 2003.
- [Rizky 2004] Rizky, Soetam, “Panduan Belajar SQL Server”, Prestasi Pustaka Publisher , Jakarta 2004
- [Sholih 2010] sholih,dan prof.dr.Ir. Imam robandi,MT, “Analisis dan Perancangan Berorientasi Obyek”, Muara Indah, Bandung 2010
- [Sulianta 2010] Sulianta, Feri dan Juju, Dominikus, “Data Mining, Meramalkan Bisnis Perusahaan ” , Elex Media Komputindo, 2010
- [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.