

Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ)

Annisa Paramitha Fadillah

Sistem Informasi, UNIKOM

Bandung

annisap9@gmail.com

Abstract- XYZ University as the leading private universities in Indonesia is required to improve its quality of education to produce quality graduates. To produce quality graduates, students' success in their study is very influential. Student study success can be judged from several aspects of assessments in the lecture. For information systems courses, student success can be viewed in several student assessment aspects such as assignments, Mid-Semester Exam, and End-Semester Exam. More in-depth analysis on the value of the portion of the pattern of student needs in order to know the hidden patterns and to discover the knowledge of students will predict course passing success. This pattern is expected to help increase the success of students study. This research used CRISP-DM methods as standard processes for data mining that can be applied to the general problem-solving strategies on business or to other research units. Election algorithm as an algorithm C4.5 decision tree is used to facilitate the establishment. The results of this research study is the prediction of the pattern of graduation students in taking the course. This pattern is generated from the decision tree is expected to be a reference for graduate studies students take a course in information systems courses XYZ University.

Keywords - CRISP - DM, C4.5 algorithm, graduation studies, student, and course

I. PENDAHULUAN

Perguruan tinggi merupakan salah satu organisasi yang memberikan jasa pelayanan pendidikan kepada masyarakat. Dalam perkembangannya perguruan tinggi Baik negeri maupun swasta bersaing untuk menjadi perguruan tinggi terbaik dan untuk menghasilkan lulusan terbaik.

Universitas XYZ adalah salah satu perguruan tinggi swasta yang sedang berkembang di Bandung. Sebagai perguruan tinggi Universitas XYZ swasta ternama di Indonesia, Universitas XYZ dituntut untuk meningkatkan mutu pendidikan dengan menghasilkan lulusan yang berkualitas, inovatif dan adaptif terhadap perkembangan

yang ada di Indonesia. Tidak dipungkiri untuk menghasilkan lulusan berkualitas, keberhasilan studi mahasiswa itu sendiri sangat berpengaruh.

Keberhasilan studi mahasiswa dapat terlihat dari beberapa aspek penilaian dalam perkuliahan, untuk program studi sistem informasi di Universitas XYZ aspek penilaian mahasiswa antara lain, aspek penilaian dalam pengerjaan tugas, penilaian dalam Ujian Tengah Semester (UTS), dan penilaian Ujian Akhir Semester.

Setiap mahasiswa memiliki pola atau strategi sendiri untuk meraih kelulusan dalam setiap bidang studi. Ada mahasiswa yang lebih fokus pada penilaian UTS dan UAS, tanpa terlalu memperhatikan penilaian terhadap tugas yang diberikan pada saat perkuliahan. Ada pula mahasiswa yang fokus pada ketiga aspek penilaian sehingga aspek penilaian ketiga aspek tersebut relatif seimbang satu sama lain. Pola tersebut diketahui melalui hasil kuesioner awal, yang dilakukan pada 40 responden acak yang terdapat di prodi sistem informasi. Hasil kuesioner tersebut dapat dilihat pada Tabel I

TABEL I
HASIL KUESIONER AWAL

No	Aspek Penilaian	Persentase
1	Tugas – UTS	2,50%
2	Tugas – UAS	15%
3	UAS – UTS	20%
4	UAS - UTS – TUGAS	62,5%

Analisis terhadap pola porsi nilai mahasiswa perlu dilakukan lebih mendalam, agar dapat diketahui pola yang tersembunyi dan menemukan pengetahuan tentang prediksi kelulusan mahasiswa menempuh mata kuliah. Pola ini diharapkan dapat membantu meningkatkan keberhasilan studi mahasiswa.

Metode yang tepat perlu digunakan untuk melakukan tahapan penelitian ini. CRISP-DM menyediakan standar proses baku untuk *data mining* yang dapat diterapkan ke dalam strategi pemecahan masalah umum pada bisnis atau pada unit penelitian. CRISP-DM membandingkan metodologi *data mining* lain lebih lengkap dan terdokumentasi dengan baik. Setiap fase terstruktur dan terdefinisi dengan jelas sehingga mudah diaplikasikan bahkan bagi pemula sekalipun [1].

Penggunaan metodologi CRISP-DM pada *data mining* sudah banyak diterapkan dalam penelitian, salah satunya pada jurnal “*Business and Data Understanding* dalam rangka pembentukan model tata letak dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan menggunakan metode *association rule* dan *decision tree*” jurnal ini membahas mengenai masalah pemahaman tujuan bisnis dan data yang akan digunakan pada proses *data mining* untuk menentukan tata letak strategis dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan. Penelitian ini menggunakan metodologi CRISP-DM sebagai tahapan penelitian.

Selain penggunaan metode yang tepat, pemilihan algoritma yang tepat adalah hal yang penting dilakukan dalam yang berhubungan dengan *data mining*. Pemilihan algoritma C4.5 sebagai algoritma untuk mempermudah pembentukan pohon keputusan, didasarkan pada survei yang dipublikasikan Springer “*Top 10 Algorithm in Data Mining*” [3], algoritma C4.5 ditempatkan pada posisi pertama sebagai algoritma paling banyak digunakan dalam *data mining*.

Hasil dari penelitian ini adalah prediksi pola kelulusan studi mahasiswa dalam menempuh mata kuliah, pola ini dihasilkan dari pohon keputusan yang diharapkan dapat menjadi acuan untuk kelulusan studi mahasiswa menempuh mata kuliah pada Universitas XYZ.

Berdasarkan uraian tersebut, maka penelitian diberi judul “Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus pada Universitas XYZ)”

II. KAJIAN PUSTAKA

Perguruan tinggi merupakan salah satu organisasi yang memberikan jasa pelayanan pendidikan kepada masyarakat. Dengan perkembangan teknologi yang semakin pesat pemanfaatan berbagai sebuah sistem informasi pada perguruan tinggi sangat penting. Salah satu pemanfaatan sistem informasi pada perguruan tinggi adalah sistem informasi akademik.

Sistem informasi merupakan sistem yang dibuat oleh manusia yang terdiri dari komponen-komponen dalam organisasi untuk mencapai tujuan yaitu menyajikan informasi. Tugas dari sistem informasi adalah untuk melakukan siklus pengolahan data, yaitu mengolah data untuk dijadikan informasi yang berguna lewat suatu siklus.

Database adalah kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan di

perangkat keras komputer dan digunakan perangkat lunak untuk memanipulasinya.

Data mining menurut Larose pada buku ‘*Algoritma Data Mining*’[2], *Data Mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar.

Metodologi CRISP-DM adalah standarisasi *data mining* yang disusun oleh tiga penggagas *data mining market* yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Kemudian dikembangkan pada berbagai *workshops* (antara 1997-1999). Lebih dari 300 organisasi yang berkontribusi dalam proses *modelling* ini dan akhirnya CRISP-DM 1.0 dipublikasikan pada 1999 [1].

Dalam penerapannya *data mining* dengan menggunakan metodologi CRISP-DM telah banyak dilakukan, salah satu contohnya dalam jurnal “*Business and Data Understanding* dalam rangka pembentukan model tata letak dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan menggunakan metode *association rule* dan *decision tree* (studi kasus di kota Bandung)” Jurnal ini membahas mengenai masalah pemahaman tujuan bisnis dan data yang akan digunakan pada proses *data mining* untuk menentukan tata letak strategis dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan. Penelitian ini menggunakan metodologi CRISP-DM sebagai tahapan penelitian. Pada penelitian ini menggunakan metode *association rule* dan *decision tree*.

Metodologi CRISP-DM juga digunakan dalam jurnal “Prediksi Kinerja Penjualan Karya Musik Menggunakan *Framework* CRISP-DM” [1]. Jurnal ini membahas tentang prediksi kinerja penjualan karya musik dengan memanfaatkan proses *data mining* yang mengikuti langkah-langkah dalam *framework* CRISP-DM. CRISP-DM dapat membantu analisis kinerja secara lebih menyeluruh dan terstruktur.

Penelitian lain yang menggunakan CRISP-DM sebagai metodologi dan proses *clustering* dengan menggunakan algoritma K-Means adalah penelitian “*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan Tridharma” [2]. Penelitian ini menggunakan tahapan CRISP-DM sebagai metodologi dan proses *clustering* dengan menggunakan algoritma K-Means.

Selain penelitian menggunakan metodologi CRISP-DM, terdapat penelitian yang menerapkan algoritma C4.5 sebagai algoritma yang membantu dalam pembuatan pohon keputusan, judul penelitian tersebut adalah “Implementasi Algoritma C4.5 dalam Pembuatan Aplikasi untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa Baru” [1]. Penelitian ini membahas mengenai pembuatan aplikasi untuk menganalisis kemungkinan pengunduran diri calon mahasiswa baru. Analisis dilakukan dengan memanfaatkan teori penalaran berbasis kasus, yaitu membandingkan kasus calon mahasiswa baru dengan kasus-kasus yang pernah terjadi sebelumnya.

Selain itu terdapat juga penelitian lain yang menggunakan algoritma C4.5 penelitian tersebut berjudul "Penerapan *Data Mining* untuk Memprediksi Fluktuasi Harga Saham Menggunakan Metode *Classification* dengan Teknik *Decision Tree*" [8]. Penelitian ini menghasilkan aplikasi yang dapat melakukan keseluruhan proses baik *preparation data* dalam hal pengolahan data awal sampai dengan analisis data menggunakan teknik di atas sehingga membuat efisien dan efektif dalam menganalisa dengan jumlah data yang besar untuk dapat menghasilkan prediksi fluktuasi harga saham lebih optimal.

Penelitian yang akan dilakukan terinspirasi dengan oleh penelitian dengan judul "*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan Tridharma" [5], akan tetapi tema yang diangkat membahas mengenai analisis prediksi kelulusan studi mahasiswa menempuh mata kuliah, tahapan yang dilakukan mengacu pada penelitian '*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma' [5], namun perbedaan pada penelitian yang dilakukan ialah objek penelitian di mana penelitian akan dilakukan di Universitas XYZ, perbedaan juga terdapat pada variabel penelitian, dan juga pada penelitian ini proses modeling tidak menggunakan tahap *clustering* dan menggunakan algoritma k-means, melainkan menggunakan algoritma C4.5 yang digunakan untuk membangun pohon keputusan.

III. METODE PENELITIAN

Alur pada penelitian ini, disesuaikan dengan tahapan pada metode CRISP-DM. Tahapan yang digunakan mengacu pada penelitian yang berjudul '*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma'. Pemahaman tahapan mengacu pada penelitian tersebut akan tetapi terdapat perbedaan pada objek serta variabel penelitian yang dilakukan.

Terdapat tiga tahapan penelitian, pertama merupakan tahapan pengumpulan data. Pada tahapan ini menjelaskan mengenai cara pengumpulan data yang dilakukan dalam penelitian. Pengumpulan data dilakukan selain dengan melakukan wawancara dan kuesioner, data juga dikumpulkan terkait dengan suasana akademik mahasiswa pada universitas XYZ. Pengumpulan data tersebut dilakukan dengan melakukan observasi terkait dengan suasana akademik pada Universitas XYZ, sumber data tersebut didapatkan dari sistem informasi akademik yang telah diimplementasikan pada Universitas XYZ

Tahapan kedua adalah tahapan pemahaman bisnis dan pemahaman data. Proses pada tahapan ini merupakan bagian dari Metodologi CRISP-DM. Proses pada tahapan ini yaitu, pemahaman bisnis, pemahaman data, dan proses persiapan data. Tahapan ketiga merupakan tahapan pemodelan dan evaluasi. Proses pada tahapan ini juga merupakan bagian dari metodologi CRISP-DM. Proses

pada tahapan ini yaitu, proses pemodelan, proses evaluasi dan proses penyebaran.

1. Tahapan pengumpulan data awal:
 - a. Studi Pustaka
 - b. Observasi
 - c. Wawancara
 - d. Kuesioner
2. Tahapan Pemahaman bisnis dan pemahaman data:
 - a. Pemahaman bisnis
Pada tahapan pemahaman bisnis ada beberapa hal yang dilakukan antara lain, menentukan tujuan bisnis, menilai situasi, dan menentukan tujuan *data mining*.
 - b. Pemahaman data
Pada tahapan pemahaman data, ada beberapa hal yang dilakukan antara lain, mengumpulkan data awal, mendeskripsikan data, mengeksplorasi data, dan memverifikasi kualitas data.
 - c. Persiapan Data
Pada tahapan persiapan data, ada beberapa hal yang dilakukan antara lain, deskripsi data set, memilih data, membangun data, mengintegrasikan data dan membersihkan data. Tahapan persiapan data yang dilakukan berdasarkan penelitian yang berjudul '*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma' [2].
3. Tahap Pemodelan dan Evaluasi:
 - a. Pemodelan
Pada penelitian yang berjudul '*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma' [2], didefinisikan 'tahapan pemodelan merupakan tahapan yang secara langsung melibatkan *data mining*. Pemilihan teknik *data mining*, algoritma dan menentukan parameter dengan nilai yang optimal'. Pada tahapan pemodelan, ada beberapa hal yang dilakukan antara lain, memilih teknik pemodelan, membangun model, dan menilai model. Tahapan yang dilakukan mengacu pada penelitian tersebut.
 - b. Evaluasi
Pada penelitian yang berjudul '*Data Clustering* menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma' [5], didefinisikan tahapan evaluasi adalah 'fase interpretasi terhadap hasil *data mining*. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap *modelling* sesuai dengan sasaran yang ingin dicapai dalam tahap *business understanding*'. Pada tahapan evaluasi, ada beberapa hal yang dilakukan antara lain:
 - 1) Mengevaluasi hasil (*Evaluate Results*)
 - 2) Meninjau Proses (*Review Process*)
 - 3) Menentukan Langkah selanjutnya (*Determine Next Steps*)

IV. PEMBAHASAN DAN HASIL PENELITIAN

Pada penelitian ini memanfaatkan proses yang terdapat pada metode CRISP-DM, akan tetapi tidak melakukan proses terakhir yakni proses implementasi. Tahapan yang dilakukan untuk melakukan penerapan metode CRISP-DM mengacu pada penelitian yang berjudul ‘Data Clustering menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan tridharma’ [5].

A. Pembahasan

1) Pemahaman Bisnis (Business Understanding)

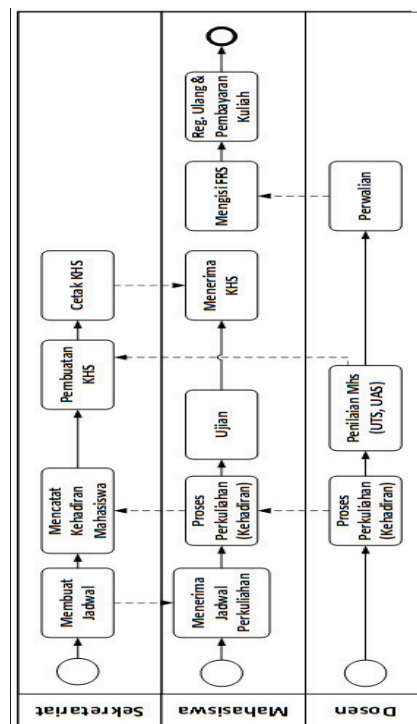
Tahapan pemahaman bisnis berfokus pada pemahaman tujuan kebutuhan berdasarkan penilaian bisnis. Kemudian pemahaman tersebut diubah menjadi sebuah rencana awal *data mining* yang dirancang untuk mencapai tujuan. Pemahaman bisnis mengacu pada aturan penilaian Mahasiswa (Nilai Akhir) yang ditentukan oleh universitas XYZ. Pada tahap ini diperlukan pemahaman tentang latar belakang dan tujuan pada proses bisnis yang berhubungan dengan nilai mahasiswa:

1. Menentukan Tujuan Bisnis (Determine Business Objectives)

Tujuan bisnis dilakukannya penelitian mengenali pola nilai mahasiswa untuk mengetahui prediksi kelulusan studi mahasiswa dalam menempuh mata kuliah.

2. Menilai Situasi (Assess the Situation)

Sistem informasi akademik ini berkaitan dengan suasana akademik mahasiswa pada universitas XYZ. Sistem informasi akademik merupakan sebuah sistem yang dikembangkan untuk melakukan manajemen dan pengolahan data terhadap jalannya suatu proses administrasi dan operasional akademik pada program studi sistem informasi data yang diolah pada sistem informasi akademik pada program studi yang berkaitan antara lain, data mahasiswa, data nilai, data penjadwalan, dan lain-lain.



Gambar 1 Alur Proses Bisnis

3. Menentukan Tujuan Data Mining (Determine the Data Mining Goals)

Tujuan data mining atau tujuan penelitian ini adalah menggali pengetahuan (*discovering knowledge*) tentang pola (*pattern*) porsi nilai mahasiswa yang berkaitan dengan keberhasilan studi mahasiswa dalam sebuah mata kuliah.

2) Pemahaman Data (Data Understanding)

Pada penelitian yang berjudul ‘Business and Data Understanding dalam rangka pembentukan model tata letak dan tata ruang pasar tradisional dalam meningkatkan laba penjualan menggunakan metode *association rule dan decision tree* (studi kasus di kota Bandung) [3]’, dijelaskan bahwa ‘tahapan pemahaman data dimulai dengan pengumpulan data awal dan hasil kegiatan dalam rangka untuk membiasakan diri dengan data untuk mengidentifikasi masalah data, untuk menentukan wawasan pertama ke data atau mendeteksi *subset* menarik untuk membentuk hipotesis untuk informasi yang tersembunyi’.

1. Mengumpulkan Data Awal (Collect the Initial Data)

Pengumpulan data awal dilakukan dengan studi pustaka, wawancara observasi, dan kuesioner. Kuesioner digunakan untuk memperoleh data awal yang menunjang penelitian yang berhubungan dengan pola porsi nilai mahasiswa dalam menempuh mata kuliah. Kuesioner dilakukan pada 40 responden acak yang disebar di lingkungan Universitas XYZ. Penentuan 40 responden berdasarkan asumsi bahwa jumlah rata-rata mahasiswa setiap kelas adalah 40

orang. Kuesioner yang disebarakan pada responden dapat dilihat pada Lampiran.

TABEL II
HASIL KUESIONER

No	Aspek Penilaian	Persentase
1	Tugas – UTS	2,50%
2	Tugas – UAS	15%
3	UAS – UTS	20%
4	UAS - UTS – TUGAS	62,5%

2. Mendeskripsikan Data (*Describe the Data*)

TABEL III
TAHAPAN MENDESKRIPSIKAN DATA

Tahapan Pemahaman Data (<i>Data Understanding</i>)	
Nama Fungsi	Mendeskripsikan Data (<i>Describe the Data</i>)
Input	KHS, Data Absen, Hasil Kuesioner, Database SIAKAD
Output	Database SIAKAD prodi SI
Tujuan	Melakukan analisis dan memahami data yang didapatkan dari hasil pengumpulan data awal.

3. Mengeksplorasi Data (*Explore the Data*)

TABEL IV
TAHAPAN MENGEKSPLORASI DATA

Tahapan Pemahaman Data (<i>Data Understanding</i>)	
Nama Fungsi	Mengeksplorasi Data (<i>Explore the Data</i>)
Input	Database SIAKAD prodi SI
Output	Tabel Dosen, Tabel FRS, Tabel Login, Tabel Mahasiswa, Tabel Mata kuliah
Tujuan	Melakukan eksplorasi data dan analisis struktur tabel yang ada pada SIAKAD SI

4. Memverifikasi Kualitas Data (*Verify Data Quality*)

TABEL V
TAHAPAN MEMVERIFIKASI KUALITAS DATA

Tahapan Pemahaman Data (<i>Data Understanding</i>)	
Nama Fungsi	Memverifikasi Kualitas Data (<i>Verify Data Quality</i>)
Input	Tabel Dosen, Tabel FRS, Tabel Login, Tabel Mahasiswa, Tabel Matakuliah
Output	Tabel Dosen, Tabel FRS, Tabel Login, Tabel Mahasiswa, Tabel Matakuliah (Hasil Verifikasi)
Tujuan	Melakukan proses verifikasi data, Verifikasi data dilakukan pada struktur tabel, data yang terdapat pada tabel dan integrasi antar tabel.

3) *Persiapan Data (Data Preparation)*

Pada penelitian yang berjudul ‘*Business and Data Understanding*’ dalam rangka pembentukan model tata letak dan tata ruang pasar tradisional dalam meningkatkan laba penjualan menggunakan metode *association rule* dan *decision tree* (studi kasus di kota Bandung) [6], dijelaskan bahwa tahapan persiapan data mencakup semua kegiatan yang membangun *dataset* akhir (data yang akan dimasukkan ke dalam pemodelan) dari data mentah awal. *Data preparation* mencakup semua kegiatan untuk membangun *data set* yang akan dimasukkan ke dalam alat pemodelan dari data mentah awal atau membuat *database* baru untuk *set up data mining*. *Data preparation* mencakup semua kegiatan untuk membangun *data set* yang akan diolah

dalam proses pemodelan menggunakan algoritma C4.5 untuk membangun pohon keputusan.

1. Mendeskripsikan Data Set (*Data Set Description*)

TABEL VI
TAHAPAN MENDESKRIPSIKAN DATA SET

Tahapan Persiapan Data (<i>Data Preparation</i>)	
Nama Fungsi	Mendeskripsikan Data Set (<i>Data Set Description</i>)
Input	Tabel Dosen, Tabel FRS, Tabel Login, Tabel Mahasiswa, Tabel Matakuliah (Tabel yang sudah diverifikasi)
Output	Tabel Mahasiswa, tabel FRS, tabel mata kuliah,
Tujuan	Melakukan rancangan set data yang sesuai dengan tujuan data mining. Pada tahap ini tabel yang berhubungan dipilih untuk mempermudah proses pemilihan data.

2. Memilih Data (*Select Data*)

TABEL VII
TAHAPAN MENDESKRIPSIKAN DATA

Tahapan Persiapan Data (<i>Data Preparation</i>)	
Nama Fungsi	Memilih Data (<i>Select Data</i>)
Input	Tabel FRS, tabel mata kuliah.
Output	Tabel Mahasiswa dan tabel nilai
Tujuan	Tabel yang dipilih berhubungan dengan pengenalan pola nilai mahasiswa. Tabel nilai merupakan gabungan dari tabel FRS dan tabel matakuliah.

3. Membangun Data (*Construct Data*)

TABEL VIII
TAHAPAN MEMBANGUN DATA

Tahapan Persiapan Data (<i>Data Preparation</i>)	
Nama Fungsi	Membangun Data (<i>Construct Data</i>)
Input	Tabel Nilai
Output	Tabel Nilai
Struktur Tabel Awal	kode_mk, Mtkuliah, Sks, na, nim, smstr, nip
Struktur Tabel Akhir	kode_mk, Mtkuliah, Sks, na, tugas, uts, uas, nim, smstr, nip
Tujuan	Melakukan proses pembangunan data (tabel). Pembangunan data dapat dilakukan dengan penambahan <i>query</i> .

4. Mengintegrasikan Data (*Integrate Data*)

Menggabungkan dua atau lebih tabel yang memiliki informasi yang berbeda tentang objek yang sama ke dalam set data baru yang telah disiapkan pada tahap awal *data preparation*. Tabel yang diintegrasikan berupa hasil transformasi data pada tahap sebelumnya.

5. Membersihkan Data (*Clean Data*)

TABEL IX
TAHAPAN MEMBERSIHKAN DATA

Tahapan Persiapan Data (<i>Data Preparation</i>)	
Nama Fungsi	Membersihkan Data (<i>Clean Data</i>)
Input	Tabel Mahasiswa
Output	Tabel Mahasiswa (setelah melakukan tahap pembersihan data)
Struktur tabel awal	NIM, nama, wali, thnmasuk, tahunlulus, judulta, tempat, prog, kelas, jurusan, tgl, ipk, yudisium,

	status, IP1, IP2, IP3, IP4, IP5, IP6, IP7, IP8, IPK1, totsks, sks1, sks2, sks3, sks4, sks5, sks6, sks7, sks8, alamat_bd, kota_bd, kodepos_bd, notelp_bd, nama_ortu, pekerjaan, alamat_asal, kota_asal, prov_asal, notelp_asal, kodepos_asal, aktif, ta_ta, smt_ta, pembimbing, Jk, Email, judulta2
Struktur tabel akhir	NIM, nama, wali, thnmasuk, kelas, jurusan, wali_dosen
Tujuan	Proses pembersihan data pada tabel mahasiswa membuat tabel mahasiswa lebih sederhana dari tabel mahasiswa sebelumnya dan lebih difokuskan hanya pada data yang berhubungan untuk mengetahui hasil dari pembelajaran mahasiswa di pada saat perkuliahan berlangsung, tidak termasuk tahapan pembuatan tugas akhir.

6. Format Data

TABEL X
TAHAPAN FORMAT DATA

Tahapan Persiapan Data (Data Preparation)	
Nama Fungsi	Format Data
Input	Tabel Mahasiswa dan Tabel Nilai
Output	3 aspek penilaian: UTS, UAS, TUGAS
Tujuan	Pemilihan nilai atau field yang akan digunakan pada tahapan pemodelan dengan menggunakan algoritma C4.5

4) Pemodelan (Modeling)

Pemilihan teknik data mining, algoritma dan menentukan parameter dengan nilai yang optimal. Pada tahapan pemodelan, ada beberapa hal yang dilakukan antara lain, memilih teknik pemodelan, membangun model, dan menilai model:

1. Memilih Teknik Pemodelan (Select Modelling Technique)

Teknik data mining yang dipilih adalah *decision tree* dengan menggunakan algoritma C4.5. *Decision tree* dan algoritma C4.5 sangat tepat digunakan untuk mencapai tujuan awal penelitian ini yaitu menggali pengetahuan tentang prediksi mahasiswa menempuh matakuliah. Pemodelan *data mining* diawali dengan membuat *rule* untuk pembentukan pohon keputusan.

2. Membangun Model (Build Model)

Sesuai ketentuan umum yang telah diterapkan, kriteria penilaian mahasiswa menjadi salah satu tolak ukur kelulusan mahasiswa dalam setiap mata kuliah. Terdapat 3 aturan penilaian yang diterapkan yaitu:

- a. NA = 10% Quiz + 20% Tugas + 30% UTS + 40% UAS
- b. NA = 20% Tugas + 30% UTS + 50% UAS.
- c. NA = 40% UTS + 60% UAS.

Dari ketiga aturan tersebut dipilih aturan kedua yang mewakili hasil kuesioner yang telah disebarkan kepada 40 responden pada tahap pengumpulan data awal. Sehingga diketahui nilai mahasiswa memiliki

tiga aspek penilaian yaitu tugas, uts, dan uas. Tiga aspek tersebut disebut dengan atribut. Atribut Atribut memiliki nilai-nilai yang dinamakan *instance*.

3. Menilai Model (Asses Model)

Pemodelan dilakukan dengan pembentukan pohon keputusan menggunakan algoritma C.45. Dengan aturan (*rule*) yang telah ditentukan sebelumnya langkah selanjutnya memilih atribut sebagai akar.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut yang ada. Untuk menghitung *gain* digunakan rumus sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S : Himpunan Kasus

A : Atribut

n : jumlah partisi atribut A

|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam SS

Sementara untuk menghitung entropi dapat dilihat pada persamaan 2 berikut:

$$Entropy(S) = \sum_{i=1}^n - pi * \log_2 pi$$

Keterangan:

S : Himpunan Kasus

A : Fitur

n : jumlah partisi S

pi : proposi dari Si terhadap S

5) Evaluasi

Evaluasi pada penelitian ini lebih difokuskan pada model atau pola yang dihasilkan oleh algoritma C4.5. Model yang dihasilkan dianalisa untuk mengetahui apakah pola yang dihasilkan sudah sesuai dengan standar yang terdapat di universitas XYZ. Jika pola yang dihasilkan belum sesuai, maka analisa lebih lanjut terhadap pola yang dihasilkan perlu dilakukan sehingga dapat menghasilkan rekomendasi perbaikan dalam suasana akademik, yang diharapkan dapat membantu keberhasilan studi mahasiswa dalam menempuh mata kuliah.

6) Penyebaran

Berdasarkan buku algoritma dan *data mining* tahapan penyebaran merupakan pembuatan laporan, dan tahapan penyebaran juga bisa didefinisikan sebagai penerapan proses *data mining* secara paralel pada departemen lain, dalam hal ini tahapan penyebaran *data mining* dapat dilakukan pada seluruh program studi yang ada di Universitas XYZ. Pada penelitian ini tahapan *deployment* atau tahap penyebaran tidak dilakukan.

B. Hasil penelitian

Untuk melakukan pengujian model tersebut diuji pada sampel data nilai mahasiswa untuk mata kuliah pengelolaan instalasi komputer yang diajarkan pada semester VII pada

tahun ajaran 2006/2007. Pada data tersebut akan dibuat pohon keputusan untuk menentukan aspek manakah yang sangat berpengaruh dalam penilaian mahasiswa yang berkaitan dengan keberhasilan studi mahasiswa untuk mata kuliah Pengelolaan Instalasi Komputer dengan menggunakan sampel sebanyak 78 mahasiswa.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang

Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Pembentukan pohon keputusan dengan menggunakan algoritma C4.5 untuk menyelesaikan permasalahan. hal pertama yang perlu dilakukan adalah menghitung jumlah kasus, jumlah kasus untuk keputusan LULUS, jumlah kasus untuk keputusan TIDAK LULUS, dan *entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut, Tugas, UTS, dan UAS. Kemudian lakukan perhitungan *gain* untuk setiap atribut.

Berikut langkah-langkah pembentukan pohon keputusan dari 78 sampel *data record* sampai pembentukan pohon keputusan awal.

1. Perhitungan *Entropy* Total

Menentukan *Entropy* (S). Untuk mencari *entropy* himpunan kasus maka dilakukan perhitungan dengan menggunakan persamaan 2 sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad \dots (2)$$

$$Entropy(S) = \left(-\frac{18}{78} * \log_2\left(\frac{18}{78}\right)\right) + \left(-\frac{59}{78} * \log_2\left(\frac{59}{78}\right)\right)$$

$$Entropy(S) = 0,7928382$$

2. Perhitungan *Entropy* Tugas

$$\begin{aligned} Entropy \text{ Tugas} &= \left(-\frac{0}{13} * \log_2\left(\frac{0}{13}\right)\right) + \left(-\frac{13}{13} * \log_2\left(\frac{13}{13}\right)\right) \\ Entropy \text{ Tugas Sangat Baik} &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ Tugas Baik} &= \left(-\frac{13}{60} * \log_2\left(\frac{13}{60}\right)\right) + \left(-\frac{47}{60} * \log_2\left(\frac{47}{60}\right)\right) \\ &= 0,7540341 \end{aligned}$$

$$\begin{aligned} Entropy \text{ Tugas Cukup} &= \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ Tugas Kurang} &= \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ Tugas Sangat Kurang} &= \left(-\frac{5}{5} * \log_2\left(\frac{5}{5}\right)\right) + \left(-\frac{0}{5} * \log_2\left(\frac{0}{5}\right)\right) \\ &= 0 \end{aligned}$$

3. Perhitungan *Entropy* UTS

$$Entropy \text{ UTS Sangat} = \left(-\frac{1}{27} * \log_2\left(\frac{1}{27}\right)\right) + \left(-\frac{26}{27} * \log_2\left(\frac{26}{27}\right)\right)$$

$$\begin{aligned} \text{Baik} &= 0,2285381 \\ Entropy \text{ UTS Baik} &= \left(-\frac{0}{15} * \log_2\left(\frac{0}{15}\right)\right) + \left(-\frac{15}{15} * \log_2\left(\frac{15}{15}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Cukup} &= \left(-\frac{2}{6} * \log_2\left(\frac{2}{6}\right)\right) + \left(-\frac{4}{6} * \log_2\left(\frac{4}{6}\right)\right) \\ &= 0,9182958 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Kurang} &= \left(-\frac{1}{11} * \log_2\left(\frac{1}{11}\right)\right) + \left(-\frac{10}{11} * \log_2\left(\frac{10}{11}\right)\right) \\ &= 0,439497 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Sangat Kurang} &= \left(-\frac{15}{19} * \log_2\left(\frac{15}{19}\right)\right) + \left(-\frac{4}{19} * \log_2\left(\frac{4}{19}\right)\right) \\ &= 0,7424876 \end{aligned}$$

4. Perhitungan *Entropy* UAS

$$\begin{aligned} Entropy \text{ UTS Sangat Baik} &= \left(-\frac{0}{1} * \log_2\left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * \log_2\left(\frac{1}{1}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Baik} &= \left(-\frac{0}{5} * \log_2\left(\frac{0}{5}\right)\right) + \left(-\frac{5}{5} * \log_2\left(\frac{5}{5}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Cukup} &= \left(-\frac{0}{7} * \log_2\left(\frac{0}{7}\right)\right) + \left(-\frac{7}{7} * \log_2\left(\frac{7}{7}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Kurang} &= \left(-\frac{0}{21} * \log_2\left(\frac{0}{21}\right)\right) + \left(-\frac{21}{21} * \log_2\left(\frac{21}{21}\right)\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy \text{ UTS Sangat Kurang} &= \left(-\frac{19}{44} * \log_2\left(\frac{19}{44}\right)\right) + \left(-\frac{25}{44} * \log_2\left(\frac{25}{44}\right)\right) \\ &= 0,9865446 \end{aligned}$$

Untuk nilai *Gain* dihitung dengan menggunakan persamaan satu sebagai berikut:

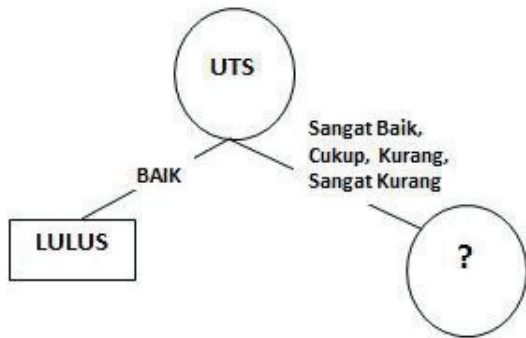
$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

$$\begin{aligned} \text{Gain Tugas} &= 0,7928382 - \left(\frac{13}{78} * 0\right) + \left(\frac{60}{78} * 0,7540341\right) + \\ &\quad \left(\frac{0}{78} * 0\right) + \left(\frac{0}{78} * 0\right) + \left(\frac{5}{78} * 0\right) \\ &= 0,212812 \end{aligned}$$

$$\begin{aligned} \text{Gain UTS} &= 0,7928382 - \left(\frac{1}{78} * 0\right) + \left(\frac{5}{78} * 0\right) + \left(\frac{7}{78} * 0\right) \\ &\quad \left(\frac{21}{78} * 0\right) + \left(\frac{19}{78} * 0,7424876\right) \\ &= 0,479357 \end{aligned}$$

$$\begin{aligned} \text{Gain UAS} &= 0,7928382 - \left(\frac{0}{78} * 0\right) + \left(\frac{0}{78} * 0\right) + \left(\frac{0}{78} * 0\right) + \\ &\quad \left(\frac{0}{78} * 0\right) + \left(\frac{44}{78} * 0,9865446\right) \\ &= 0,236326 \end{aligned}$$

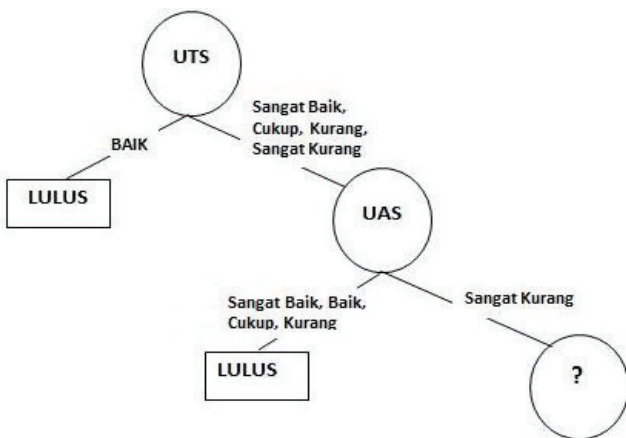
Dari hasil tersebut dapat diketahui bahwa atribut dengan *gain* tertinggi adalah UTS, yaitu sebesar 0,479357. Dengan demikian, UTS dapat menjadi node akar. Dari hasil tersebut dapat digambarkan pohon keputusan sementara yang dapat dilihat pada Gambar 2.



Gambar 2 Pohon keputusan hasil perhitungan Node 1

Entropy dan *gain* dihitung kembali untuk mengetahui atribut mana yang akan menjadi *node* akar dari atribut UTS, perhitungan dilakukan sesuai dengan persamaan satu dan dua, seperti yang telah dijelaskan sebelumnya. Menghitung jumlah kasus untuk LULUS, jumlah kasus untuk keputusan TIDAK LULUS, dan *entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut Tugas, dan UAS yang dapat menjadi *node* akar dari nilai atribut Sangat Baik, Cukup, Kurang dan Sangat Kurang. Setelah itu, lakukan perhitungan *gain* untuk tiap-tiap atribut. Hasil perhitungan dapat dilihat pada tabel 4.27

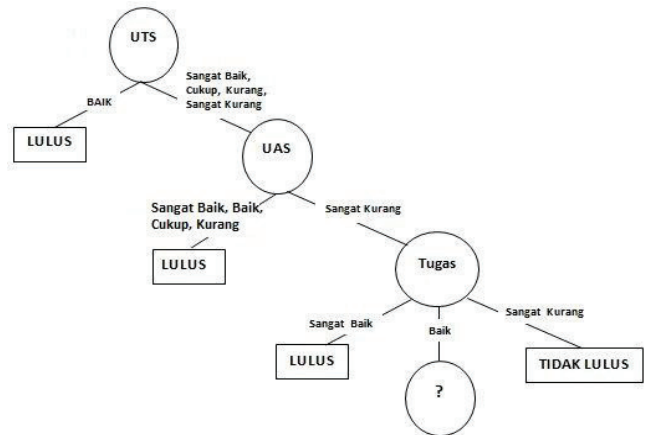
Dari hasil pada tabel 4.27 dapat diketahui bahwa atribut dengan *gain* tertinggi untuk *node* akar UTS dengan atribut nilai Sangat Baik, Cukup, Kurang dan Sangat Kurang adalah UAS, yaitu sebesar 0,291067. Dengan demikian UAS dapat menjadi *node* akar *node* UTS. Dari hasil tersebut dapat digambarkan pohon keputusan sementara yang dapat dilihat pada Gambar 3.



Gambar 3 Pohon keputusan hasil perhitungan Node 1.2

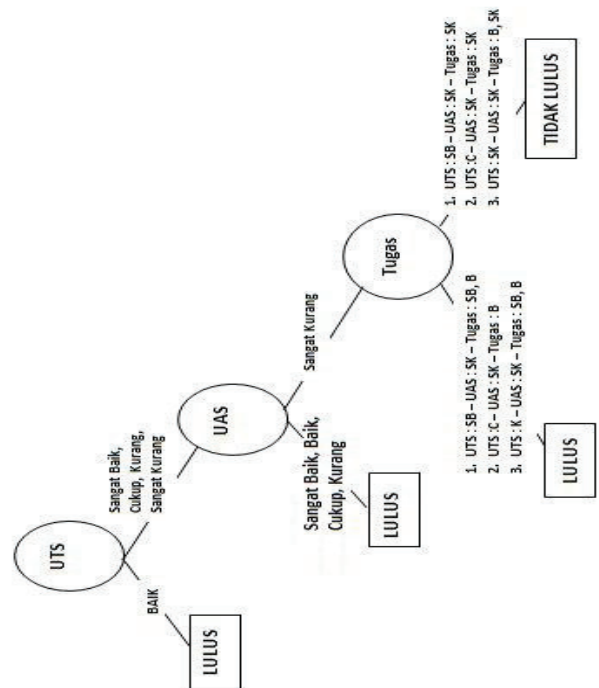
Secara tidak langsung dengan diketahui, *node* cabang dari UTS adalah UAS maka kemudian diketahui juga *node*

cabang dari UAS adalah Tugas, dengan nilai atribut Sangat baik diklasifikasikan menjadi keputusan LULUS, dan nilai atribut Sangat Kurang diklasifikasikan menjadi TIDAK LULUS. Akan tetapi terjadi keganjilan pada keputusan yang dihasilkan. Pada *node* cabang TUGAS dengan atribut BAIK masih memiliki keputusan LULUS dan TIDAK LULUS, sehingga klasifikasi pohon menjadi tidak konsisten pada nilai UTS dengan atribut BAIK. Pohon keputusan yang terbentuk dapat dilihat pada Gambar 4.



Gambar 4 Pohon keputusan hasil perhitungan Node 1.2.2 sementara

Untuk memastikan klasifikasi pohon keputusan akhir yang dihasilkan maka klasifikasi *node* 3 dipecah kembali berdasarkan atribut UTS yaitu: Sangat Baik, Cukup, Kurang dan Sangat Kurang, UAS dengan atribut Sangat Kurang. Dari hasil pemecahan beberapa tabel tersebut menghasilkan pohon keputusan seperti yang ditunjukkan pada Gambar 5.



Gambar 5 Pohon Keputusan Akhir

Keterangan:

- SB : Sangat Baik
- B : Baik
- C : Cukup
- K : Kurang
- SK : Sangat Kurang

C. Tahap Interpretasi Pola Pohon Keputusan

Berdasarkan gambar maka ditemukan pola porsi akademik berdasarkan sampel data sebanyak 78 data, dengan mata kuliah yang diajarkan Pengelolaan Instalasi Komputer tahun ajaran 2006/2007 yang diajarkan semester VII angkatan IV. Terdapat 8 pola yang dihasilkan oleh pohon keputusan yang terdapat pada Gambar 4, dapat dilihat pada Tabel XI.

TABEL XI
HASIL POLA POHON KEPUTUSAN

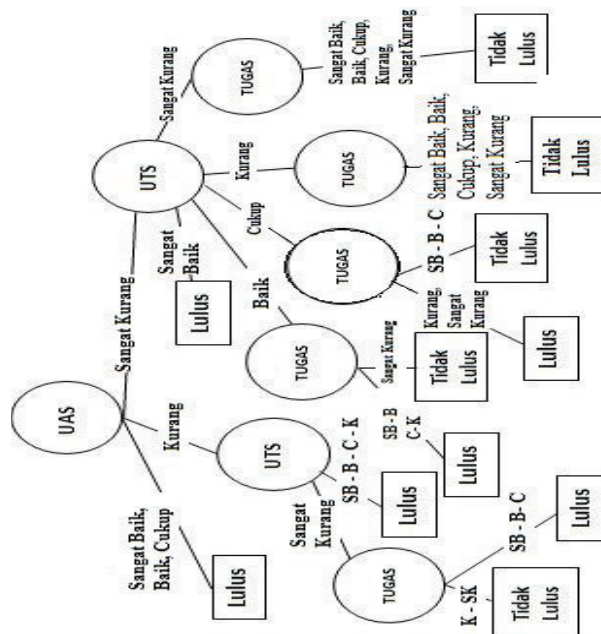
No	Nilai	Atribut	Status
1	Jika uts	Baik	Lulus
2	Jika uts	Sangat Baik atau Cukup atau Kurang atau Sangat Kurang	Lulus
	Dan uas	Sangat Baik atau Baik atau Cukup atau kurang	
3	Jika uts	Sangat Baik	Lulus
	Dan uas	Sangat Baik atau Baik atau Cukup atau kurang	
4	Jika uts	Cukup	Lulus
	Dan uas	Sangat Kurang	
	Dan tugas	Sangat Baik atau Baik	
5	Jika uts	Kurang	Lulus
	Dan uas	Sangat Kurang	
	Dan tugas	Baik	
6	Jika uts	Sangat Baik	Tidak lulus
	Dan uas	Sangat Kurang	
	Dan tugas	Sangat Kurang	
7	Jika uts	Cukup	Tidak lulus
	Dan uas	Sangat Kurang	
	Dan tugas	Sangat Kurang	
8	Jika uts	Sangat Kurang	Tidak lulus
	Dan uas	Sangat Kurang	
	Dan tugas	Baik atau Sangat Kurang	

Dari pola-pola yang terbentuk pada pohon keputusan, ada dua kategori aturan yaitu, aturan yang menghasilkan klasifikasi tertentu (LULUS dan TIDAK LULUS). Pola yang dihasilkan oleh data sampel diketahui atribut UTS yang paling berpengaruh pertama, UAS menjadi atribut paling berpengaruh kedua serta Tugas merupakan atribut paling berpengaruh ketiga untuk pembentukan pola pada pohon keputusan.

Diketahui pola yang dihasilkan oleh data sampel memiliki perbedaan dengan aturan normal yang telah ditentukan di universitas, pola pohon keputusan yang dihasilkan sangat tergantung dari penilaian yang diberikan

dosen masing-masing pada mata kuliah, hal ini dapat dilihat dari gambar aturan normal pada Gambar 6.

Pola aturan (Gambar 6) ini berdasarkan ketentuan yang berlaku di Universitas XYZ seperti yang telah dijelaskan pada subbab 4.1.4 yaitu tahapan pemodelan. Pola ini merupakan hasil dari penjabaran aturan pada tabel aturan awal. Aturan yang digunakan berdasarkan standar aturan penilaian yang ditentukan yaitu: NILAI AKHIR = 50% UAS + 30% UTS + 20% TUGAS



Gambar 6 Pola aturan

Keterangan:

- SB : Sangat Baik
- B : Baik
- C : Cukup
- K : Kurang
- SK : Sangat Kurang

Pola pohon keputusan yang dihasilkan dari data sampel memiliki perbedaan dengan pola aturan yang telah ditentukan. Perbedaan yang terdapat antara kedua pola tersebut antara lain:

1. Pada Pola Pohon keputusan berdasarkan data sampel diketahui aspek penilaian paling berpengaruh pertama, dalam kelulusan studi mahasiswa menempuh mata kuliah adalah UTS, sedangkan pada pola aturan yang diterapkan, nilai yang berpengaruh adalah UAS hal ini diketahui berdasarkan perhitungan nilai akhir yang diterapkan oleh universitas yaitu: NILAI AKHIR = 50% UAS + 30% UTS + 20% TUGAS
2. Pola yang dihasilkan oleh pohon keputusan memiliki perbedaan dengan aturan yang diterapkan, perbedaan tersebut dapat dilihat pada Tabel XII.

TABEL XII
TABEL PERBANDINGAN POLA POHON KEPUTUSAN DENGAN POLA AWAL

No	Pola Pohon Keputusan (berdasarkan data sampel)	Pola aturan (aturan penilaian)
1	Jika UTS: SANGAT KURANG Dan UAS: KURANG Status: LULUS (Pola ke 2)	Jika UAS: KURANG Dan UTS: SANGAT KURANG Dan Tugas: SB atau Baik atau Cukup Status: LULUS Jika UAS: KURANG Dan UTS: SANGAT KURANG Dan Tugas: Kurang atau Sangat Kurang Status: TIDAK LULUS
2	Jika UTS: CUKUP Dan UAS: KURANG Dan TUGAS: SANGAT BAIK atau BAIK Status: LULUS (Pola ke 4)	Jika UAS: KURANG Dan UTS: CUKUP Dan TUGAS: SANGAT BAIK atau BAIK Status: TIDAK LULUS
3	Jika UTS: KURANG Dan UAS: SANGAT KURANG Dan TUGAS: BAIK Status: LULUS (Pola ke 5)	Jika UAS: SANGAT KURANG Dan UTS: KURANG Dan TUGAS: BAIK Status: TIDAK LULUS
4	Jika UTS: SANGAT BAIK Dan UAS: SANGAT KURANG Dan TUGAS: SANGAT KURANG Status: TIDAK LULUS (Pola ke 6)	Jika UAS: SANGAT KURANG Dan UTS: SANGAT BAIK Dan TUGAS: SANGAT KURANG Status: LULUS
5	Jika UTS: CUKUP Dan UAS: SANGAT KURANG Dan TUGAS: SANGAT KURANG Status: TIDAK LULUS (Pola ke 7)	Jika UAS: SANGAT KURANG Dan UTS: CUKUP Dan TUGAS: SANGAT KURANG Status: LULUS

V. SIMPULAN DAN SARAN

A. Simpulan

Berdasarkan hasil penelitian yang dilakukan, dapat diambil simpulan sebagai berikut:

1. Metodologi CRISP-DM dapat diterapkan untuk melakukan prediksi kelulusan studi mahasiswa menempuh mata kuliah.
2. Algoritma C4.5 digunakan untuk melakukan pembentukan pohon keputusan. Pohon keputusan yang terbentuk menghasilkan pola nilai mahasiswa, untuk mengklasifikasikan nilai mahasiswa dengan kesimpulan hasil tidak lulus dan lulus.
3. Pola pohon keputusan yang dihasilkan dari Algoritma C4.5 pada jurusan sistem informasi sangat tergantung dari penilaian yang diberikan dosen masing-masing pada mata kuliah.

B. Saran

Berdasarkan simpulan yang dipaparkan, hal yang diharapkan adalah sebagai berikut:

1. Pengolahan *database* sebaiknya dikembangkan pada seluruh program studi yang ada di universitas.
2. Pada penelitian selanjutnya diharapkan tahapan *Deployment* pada metodologi CRISP-DM dapat dilakukan.
3. Adanya penggunaan algoritma pembentuk pohon keputusan lain, untuk membandingkan hasil yang didapatkan dari algoritma C4.5 yang digunakan pada penelitian ini.
4. Penggunaan data sampel diharapkan, dapat menggunakan berbagai data pada mata kuliah yang diajarkan di universitas.

DAFTAR PUSTAKA

- [1] A. A. Prajitno dan A. Purwarianti, "Prediksi Kinerja Penjualan Karya Musik Menggunakan Framework CRISP-DM (Studi Kasus: X Music Indonesia)," *Jurnal Sarjana Institut Teknologi Bandung bidang Teknik Elektro dan Informatika*, 2011.
- [2] Kusrini, *Algoritma Data Mining*, Yogyakarta: Andi, 2009.
- [3] X. Wu dan V. Kumar, "Top 10 Algorithms in Data Mining," *Survey Paper*.
- [4] R. G. Gelar, "Implementasi Algoritma C4.5 dalam Pembuatan Aplikasi untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa Baru," [Online]. Available: http://elib.unikom.ac.id/gdl.php?mod=browse&op=read&id=jbptuni_kompp-gdl-ranggagela-26808. [Diakses 27 Feb 2013].
- [5] I. Budiman, T. Prahasto dan Y. Christyono, "Data Clustering Menggunakan Metodologi CRISP-DM untuk Pengenalan Pola Proporsi Pelaksanaan Tridharma," 2012. [Online]. Available: <http://journal.uui.ac.id/index.php/Snati/article/viewFile/2901/2662>. [Diakses 27 Feb 2013].
- [6] Kania S, Mira., et al., (27/02/2013), "*Business dan Data Understanding* dalam Rangka Pembentukan Model Tata Letak dan Tata Ruang Pasar Tradisional dalam Rangka Meningkatkan Laba Penjualan Menggunakan Metode *Association Rule dan Decision Tree* (studi Kasus di Kota Bandung)" [Online], 2012, Tersedia: http://komputa.if.unikom.ac.id/_s/data/jurnal/volume-1-2/7-komputa-1-2-business-understanding-tata-letak-pasar-association-rule-mira.pdf/pdf/7-komputa-1-2-business-understanding-tata-letak-pasar-association-rule-mira.pdf
- [7] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*, Yogyakarta: Andi, 2012
- [8] D. R. Shavkat, "Penerapan Dta Mining untuk Memprediksi Fluktuasi Harga Saham Menggunakan Metode Classification dengan Teknik Decision Tree," [Online]. Available: http://elib.unikom.ac.id/gdl.php?mod=browse&op=read&id=jbptuni_kompp-gdl-dadanshavk-26780. [Diakses 1 Mar 2013].
- [9] Susanto, Sani, Ph.D., dan Suryadi, Dedi., S.T.,M.S, "Pengantar Data Mining - Menggali Pengetahuan Dari Bongkahan Data", Yogyakarta: Andi, Shavkat R, Dadan., 2011, Penerapan *Data Mining* untuk memprediksi fluktuasi harga saham menggunakan metode *Classification* dengan teknik *Decision Tree*, Bandung.
- [10] A. Kadir, *Dasar Perancangan dan Implementasi Database Relasional*, Yogyakarta: Andi, 2008.
- [11] S. Gurtino, *Theory and Application of IT Research*, Yogyakarta: Andi.
- [12] R. Aditama, *Sistem Informasi Akademik Berbasis Web dengan PHP*, Yogyakarta: Lokomedia, 2012.

