

BAB 1

PENDAHULUAN

Bab ini menguraikan penjelasan umum mengenai tugas akhir yang dikerjakan. Penjelasan tersebut meliputi latar belakang masalah, tujuan tugas akhir, lingkup tugas akhir, metodologi yang digunakan dan sistematika penulisan laporan.

1.1 Latar Belakang Masalah

Mahasiswa merupakan salah satu aspek penting dalam evaluasi keberhasilan penyelenggara program studi pada suatu perguruan tinggi [RID13]. Prediksi kinerja mahasiswa dalam setiap perguruan tinggi telah menjadi salah satu kebutuhan yang paling penting untuk meningkatkan kualitas pengajaran di lembaga tersebut. Melalui proses ini perguruan tinggi dapat mengetahui kebutuhan mahasiswa untuk mendapatkan hasil yang lebih baik. Mahasiswa yang membutuhkan perhatian khusus juga dapat diidentifikasi melalui proses ini. Sejumlah algoritma tersedia untuk memprediksi kinerja mahasiswa. Beberapa diantaranya *Artificial Neural Network* (ANN), *Pohon Keputusan*, *Clustering*, algoritma *Naive Bayes* [RAM13]. Oleh karena itu, evaluasi kinerja adalah salah satu basis untuk memantau perkembangan kinerja mahasiswa di perguruan tinggi.

Setiap lembaga mempunyai sistem operasional yang setiap transaksi kegiatan operasinya selalu dicatat dan didokumentasikan. Pendokumentasian setiap transaksi sangat berguna bagi lembaga tersebut untuk berbagai keperluan. Data-data tersebut tersimpan dalam sebuah basis data berkapasitas besar [PRA12]. Teknik Informatika Universitas Pasundan Bandung merupakan penyelenggara pendidikan tinggi teknik bidang informatika. Teknik Informatika Universitas Pasundan Bandung sebagai penyelenggara pendidikan tinggi, tentunya memiliki banyak data yang terus bertambah jumlahnya. Data-data yang ada bisa digunakan untuk membantu pihak jurusan dalam meningkatkan layanan kepada mahasiswa.

Data mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (hidden pattern) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam sebuah basis data, data warehouse, atau penyimpanan data lainnya (Tan dkk, 2006). *Educational data mining* adalah cara pendekatan yang muncul berkaitan dengan pengembangan metode untuk menjelajahi jenis yang unik dari data yang berasal dari lingkungan pendidikan dan menggunakan metode tersebut untuk lebih memahami para siswa [RAM13]. Salah satu teknik data mining yang bisa diterapkan pada lingkungan pendidikan adalah klasifikasi. Klasifikasi merupakan bentuk analisis data yang menghasilkan sebuah model untuk menjelaskan kelas data penting [HAN12].

Beberapa tahun terakhir, telah terjadi peningkatan minat dalam penggunaan Data Mining di bidang pendidikan. *Educational Data Mining* (EDM) berkaitan dengan bagaimana mengembangkan metode dan menganalisis konten pendidikan untuk memungkinkan pemahaman yang lebih baik terhadap kinerja siswa. Hal ini juga penting untuk meningkatkan proses belajar mengajar. Ada berbagai metode EDM dan algoritma yang digunakan untuk menemukan pola dan hubungan yang meliputi

prediksi, *clustering* dan asosiasi. Tujuan dari prediksi adalah untuk mengembangkan model yang dapat menyimpulkan satu aspek data (diprediksi parameter) dari beberapa kombinasi data lain (parameter prediktor). Pola yang ditemukan dengan metode *Data Mining* dari data pendidikan dapat digunakan untuk meningkatkan pengambilan keputusan dalam hal mengidentifikasi siswa yang berisiko, penurunan tingkat *drop-out* mahasiswa, meningkatkan keberhasilan siswa, dan meningkatkan hasil belajar siswa. Pola yang diekstrak juga dapat menesumkan perilaku belajar siswa dan karakteristik yang mempengaruhi kinerja akademis mereka [AZI13].

Pada penelitian sebelumnya oleh Mujib, dkk (2010) penerapan data mining digunakan untuk membuat sebuah sistem untuk mengklasifikasikan kelulusan mahasiswa dengan cara mengevaluasi kinerja pada tahun pertama dan tahun kedua menggunakan Algoritma *Naïve Bayes Classifier* [RID13]. Marselia, dkk (2010) data mining digunakan untuk mencari dan menemukan pola yang terdapat pada data mahasiswa berdasarkan atribut-atribut tertentu untuk memprediksi mahasiswa yang lulus atau tidak lulus sesuai dengan waktu studi dengan menggunakan algoritma *Naïve Bayes* dan C4.5, kemudian membandingkan hasil dan akurasi kedua algoritma tersebut [SUH10]. Sedangkan, Endorgan, dkk (2005) menggunakan teknik data mining untuk menggali berbagai informasi dari sebuah database mahasiswa, seperti untuk analisis hubungan antara jalur masuk mahasiswa baru, prestasi, dan kelulusannya menggunakan algoritma *K-Means* [TIM05].

Algoritma C4.5 merupakan salah satu algoritma modern untuk melakukan *data mining*, algoritma C4.5 disebut juga dengan pohon keputusan (*decision tree*) yaitu merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon dan pada setiap *node* merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas, Konsep dari pohon keputusan ini adalah dengan mengumpulkan data selanjutnya dibuatkan *decision tree* yang kemudian akan dihasilkan *rule-rule* sebagai hasil dari solusi permasalahan.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas dapat dilakukan identifikasi masalah, sebagai berikut :

- a. Bagaimana menerapkan konsep data mining dalam menemukan pola lama studi mahasiswa pada *database* mahasiswa?
- b. Bagaimana mengolah data dari suatu database mahasiswa untuk penerapan data mining dalam menemukan pola lama studi mahasiswa?
- c. Bagaimana data mining dapat memberikan pengetahuan mengenai pola lama studi mahasiswa?

1.3 Tujuan Tugas Akhir

Tujuan yang ingin dicapai dari tugas akhir ini adalah :

- a. Mengetahui penerapan konsep data minig pada *database* mahasiswa dalam memberikan informasi mengenai pola lama studi mahasiswa

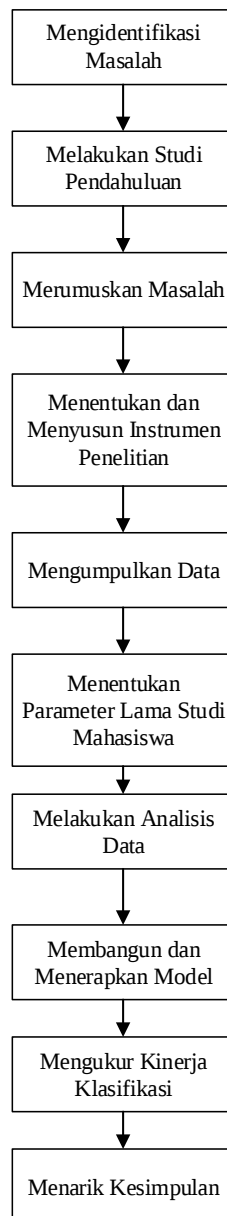
- b. Mengetahui cara mengolah data dari sebuah *database* mahasiswa untuk penerapan data mining
- c. Mengetahui pengetahuan yang dihasilkan dari penerapan data mining sehingga diharapkan dapat membantu lembaga dalam pengambilan keputusan

1.4 Lingkup Tugas Akhir

Berikut lingkup tugas akhir untuk menghindari pembahasan yang meluas :

- a. Data yang digunakan dalam penerapan data mining adalah data mahasiswa program studi Teknik Informatika Universitas Pasundan Bandung kelas Reguler
- b. Metode data mining yang digunakan adalah klasifikasi menggunakan *Algoritma C4.5*
- c. Kakas data mining yang digunakan adalah Rapid Miner

1.5 Metodologi Tugas Akhir



Gambar 1. 1 Metodologi Tugas Akhir

1.6 Langkah-langkah Pengerjaan

Langkah-langkah pengerjaan tugas akhir adalah sebagai berikut :

a. Mengidentifikasi Masalah

Pada tahap ini dilakukan identifikasi mengenai isu yang menarik mengenai perkembangan di bidang teknologi informasi dan mengamati masalah yang menarik untuk diamati.

b. Melakukan Studi Pendahuluan

Pada tahap ini dilakukan studi pendahuluan yang berkaitan dengan masalah yang diamati dan dianggap relevan mendukung penelitian yang akan dilakukan.

c. Merumuskan Masalah

Pada tahap ini dilakukan perumusan terhadap masalah yang diamati, sehingga diperoleh rumusan-rumusan masalah yang akan diselesaikan melalui penelitian.

d. Menentukan dan Menyusun Instrumen Penelitian

Pada tahap ini dilakukan penentuan dan penyusunan instrumen penelitian seperti kebutuhan data, metode data mining yang digunakan untuk membuat model, metode pengujian dan evaluasi model serta kaskas data mining.

e. Mengumpulkan Data

Pada tahap ini dilakukan pengumpulan data yang berkaitan dengan kebutuhan data yang sudah didefinisikan sebelumnya.

f. Menentukan Parameter Lama Studi Mahasiswa

Pada tahap ini dilakukan penentuan parameter mengenai lama studi berdasarkan data yang telah diperoleh sebelumnya. Parameter tersebut akan digunakan sebagai masukan untuk proses data mining.

g. Melakukan Analisis Data

Pada tahap ini dilakukan analisis terhadap data beserta parameternya agar data yang akan digunakan dalam proses mining.

h. Membangun dan Menerapkan Model

Pada tahap ini dilakukan pembangunan model pohon keputusan menggunakan algoritma C4.5 yang merepresentasikan lama studi mahasiswa dan model tersebut dapat digunakan untuk memprediksi lama studi mahasiswa.

i. Mengukur Kinerja Klasifikasi

Pada tahap ini dilakukan pengukuran kinerja klasifikasi yang dalam penelitian ini algoritma yang digunakan adalah algoritma C4.5. Pengukuran kinerja terdiri dari pengujian model dan evaluasi model.

j. Menarik Kesimpulan dan Saran

Pada tahap ini dilakukan penarikan kesimpulan terhadap penelitian yang telah dilakukan dan pemberian saran untuk penelitian selanjutnya.

1.7 Sistematika Penulisan Laporan

Sistematika penulisan tugas akhir ini terdiri dari beberapa bab, yang akan dijelaskan sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menguraikan penjelasan umum mengenai tugas akhir yang dikerjakan. Penjelasan tersebut meliputi latar belakang masalah, tujuan tugas akhir, lingkup tugas akhir, metodologi yang digunakan dan sistematika penulisan laporan.

BAB 2 LANDASAN TEORI

Bab ini menguraikan penjelasan mengenai landasan teori yang digunakan dalam pengerjaan tugas akhir. Pada bab ini dikemukakan definisi-definisi, teori-teori dan konsep yang diperlukan sebagai alat untuk membantu penelitian.

BAB 3 SKEMA PENELITIAN

Bab ini menguraikan penjelasan mengenai tahapan-tahapan dalam menyelesaikan tugas akhir dan analisis terkait penelitian tugas akhir.

BAB 4 IMPLEMENTASI

Bab ini menguraikan penjelasan mengenai tahapan-tahapan pembentukan dataset untuk penelitian tugas akhir beserta sumber datanya, pembentukan model dan pengukuran kinerja model.

BAB 5 KESIMPULAN DAN SARAN

Bab ini menguraikan kesimpulan dari penerapan data mining terkait penelitian tugas akhir dan saran secara umum

