

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Salah satu fakta yang terjadi di lembaga pendidikan STIKOM Dinamika Bangsa Jambi adalah ledakan pertumbuhan data mahasiswa. Ini terjadi setiap tahun dan setiap penerimaan mahasiswa baru. Data menentukan bahwa jumlah mahasiswa baru lebih banyak dibandingkan dengan mahasiswa yang lulus. Waktu studi mahasiswa merupakan salah satu faktor yang menentukan kualitas dari suatu lembaga pendidikan tinggi. Lembaga pendidikan tinggi perlu melakukan analisa terhadap hasil belajar mahasiswa sehingga dapat memprediksi mahasiswa yang mungkin bermasalah. Hasil prediksi dapat digunakan sebagai peringatan dini sehingga dapat dicari solusi untuk membantu mahasiswa lulus tepat waktu. Namun pada saat ini, tumpukan data tersebut belum dimanfaatkan secara maksimal. Padahal tumpukan data tersebut dapat menjadi sebuah informasi yang bermanfaat dengan menggunakan suatu teknik yaitu teknik *data mining*.

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. tujuan utama *data mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki, Mujib Ridwan et al. (2013 : 60).

Metode klasifikasi algoritma *C5.0* merupakan salah satu metode pengklasifikasian data mining yang khususnya diterapkan pada teknik *Decision Tree*. Dalam algoritma *C5.0* menghasilkan *tree* dengan jumlah cabang per node bervariasi. *C5.0* memperlakukan variabel kontinyu dan pemilihan atribut yang akan diproses menggunakan *information gain* tertinggi sebagai *parent* bagi node selanjutnya. *C5.0* adalah versi komersial dari *C4.5* yang secara luas digunakan di banyak pemaketan data mining seperti *Clementine and RuleQuest*. Tidak seperti *C4.5*, penggunaan algoritma yang tepat untuk *C5.0* belum terungkap. Hasil menunjukkan bahwa *C5.0* meningkatkan pada penggunaan memori sekitar 90%, lebih cepat daripada *C4.5*, Kurniarti Putri Wirdhaningsih, dkk (2013:2). Algoritma *C5.0* tetap dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi dapat diperoleh dengan jelas baik dalam bentuk struktur pohon keputusan maupun aturan *if-then*, sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan (Susanto & Sudiyatno, 2014).

Hal inilah yang melatar belakangi penulis untuk melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul : **"PENERAPAN DATA MINING UNTUK MEMPREDIKSI LAMA MASA STUDI MAHASISWA SISTEM INFORMASI PADA STIKOM DB JAMBI MENGGUNAKAN METODE *C5.0*"**. Hal ini peneliti lakukan untuk memprediksi lama masa studi mahasiswa dengan harapan hasil dari penelitian dapat memberikan manfaat bagi STIKOM Dinamika Bangsa Jambi dalam

meminimalisir dan memberikan penanganan lebih awal terhadap mahasiswa yang kemungkinan terlambat lulus.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang diatas, maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana mengklasifikasi dan menganalisis prediksi lama masa studi mahasiswa jurusan sistem informasi STIKOM Dinamika Bangsa Jambi menggunakan metode *C5.0* ?.
2. Bagaimana mengevaluasi hasil dari prediksi kelulusan mahasiswa jurusan Sistem Infomasi STIKOM Dinamika Bangsa Jambi dengan menggunakan metode *C 5.0* ?.

1.3 BATASAN MASALAH

Agar penelitian ini dapat berjalan dengan baik, Terarah, dan tidak keluar dari topik pembahasan maka penulis menetapkan batasan-batasan masalah yaitu :

1. Penelitian ini hanya mengklasifikasi dan menganalisis prediksi lama masa studi Mahasiswa Jurusan Sistem Informasi STIKOM DB pada tahun 2011 – 2013.
2. Analisis menggunakan teknik data mining klasifikasi dengan metode *C 5.0*.
3. Data *training* menggunakan 60 % dari data mahasiswa, sedangkan data *testing* Menggunakan 40 % dari data mahasiswa.

4. Atribut yang digunakan yaitu Jenis Kelamin, Sekolah Asal, Jurusan Sekolah Asal, IPK, Status Kuliah, Kelas Lulus (Cepat (3,5 Tahun), Tepat (4 Tahun), Lambat (Lebih dari 4 Tahun)) , dan jumlah keseluruhan nilai matakuliah.
5. Informasi yang dihasilkan berupa data pengklasifikasian prediksi lama masa studi mahasiswa dengan tiga kategori yaitu CEPAT, TEPAT dan LAMBAT.
6. Pengujian hasil analisis menggunakan *tools WEKA* dan *RapidMiner*.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Adapun tujuan penelitian yang dilakukan oleh penulis, yaitu :

1. Mengklasifikasi dan menganalisis data-data mahasiswa STIKOM DB Jambi dengan metode *C 5.0* sehingga dapat memprediksi lama masa studinya.
2. Mengevaluasi hasil dari prediksi kelulusan mahasiswa jurusan Sistem Infomasi STIKOM DB Jambi dengan menggunakan metode *C 5.0*.

1.4.2 Manfaat Penelitian

Manfaat dilakukannya penelitian ini adalah sebagai berikut :

1. Dapat mengetahui prediksi lama masa studi masing-masing Mahasiswa Sistem Informasi.
2. Dapat menggunakan hasil analisis untuk menentukan strategi dalam meningkatkan kualitas dan kuantitas kelulusan mahasiswa pada tahun-tahun berikutnya.

3. Dapat Mengetahui metode terbaik mengklasifikasikan data untuk lama masa studi mahasiswa dan hasil penelitian dapat digunakan sebagai bahan acuan untuk melakukan penelitian selanjutnya.

1.5 SISTEMATIKA PENULISAN

Adapun sistematika penulisan tugas akhir ini, penulis menguraikan dalam beberapa bab yaitu :

BAB I : PENDAHULUAN

Pada bab ini akan diuraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penulisan, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang dianalisis. Teori-teori yang digunakan antara lain mengenai data mining, klasifikasi, prediksi, lama masa studi, *naïve bayes*, dan WEKA.

BAB III : METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode klasifikasi, serta alat bantu yang digunakan pada penelitian ini.

BAB IV : ANALISIS

Pada bab ini dilakukan perhitungan analisis menggunakan metode naïve bayes terhadap data-data nilai siswa yang tersedia.

BAB V : HASIL ANALISIS DAN VISUALISASI

Pada bab ini akan ditampilkan hasil dari analisis dan bentuk visualisasi analisis dari tools WEKA yang digunakan.

BAB VI : PENUTUP

Bab ini yang berisikan kesimpulan-kesimpulan yang diambil dari hasil analisis serta saran-saran yang mencakup keseluruhan dari hasil penelitian.