

BAB I

PENDAHULUAN

1.1 Latar Belakang

Universitas Amikom merupakan sebuah instansi yang bergerak dibidang akademik yang setiap tahun mahasiswanya semakin bertambah, namun tidak semua mahasiswa dapat lulus tepat waktu sesuai dengan masa studi yang ditempuh sehingga mengakibatkan penumpukan data mahasiswa yang semakin banyak.

Melihat kondisi tersebut diperlukan penelitian untuk menggali data yang dimiliki oleh prodi informatika Universitas Amikom Yogyakarta. Data yang akan dimanfaatkan dalam penelitian ini adalah data mahasiswa alumni berupa NIM, jenis kelamin, jumlah sks, indeks prestasi semester 1 sampai semester 6, asal sekolah, dan tahun kelulusan. Kemudian data tersebut akan digunakan sebagai data *training* dan data *testing* untuk memprediksi masa studi mahasiswa. Penelitian ini dirasa perlu karena jika masa studi mahasiswa dapat diketahui sejak dini, maka pihak prodi dapat melakukan tindakan yang dirasa perlu supaya mahasiswa dapat lulus tepat waktu sekaligus dapat meningkatkan kualitas prodi itu sendiri. Dari data masalah dan historis yang sudah terjadi dibutuhkan suatu teknik dan metode klasifikasi untuk menganalisa masalah tersebut. Algoritma C4.5 dan Naive Bayes adalah salah satu metode klasifikasi yang dapat di uji untuk memproses dan memprediksi kelulusan mahasiswa.

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Algoritma C4.5 memiliki beberapa kelebihan yaitu dapat mengolah data numerik (*kontinyu*) dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan- aturan yang mudah diinterpretasikan dan tercepat diantara algoritma- algoritma yang lain [1].

Naïve Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas[2].

Pada kesempatan kali ini peneliti menggunakan metode klasifikasi dengan algoritma C4.5 dan Naive Bayes untuk menganalisa dan membandingkan akurasi dua metode pada sejumlah data sets.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan:

1. Berapa tingkat akurasi algoritma C4.5 dan Naive Bayes dalam memprediksi kelulusan mahasiswa Universitas Amikom Yogyakarta?
2. Algoritma manakah yang lebih akurat dalam memprediksi kelulusan mahasiswa Universitas Amikom Yogyakarta?

1.3 Batasan Masalah

Beberapa batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Data mahasiswa berasal dari Direktorat Innovation Center Universitas Amikom Yogyakarta.
2. Data mahasiswa yang digunakan adalah data mahasiswa alumni angkatan 2014 dan 2015 prodi informatika dengan 200 mahasiswa.

3. Kriteria kelulusan tepat waktu yaitu lama studi 4 tahun dihitung dari tahun lulus dikurangi tahun masuk mahasiswa. Bila lebih dari 4 tahun maka termasuk pada klasifikasi kasus tidak lulus tepat waktu.
4. Metode yang digunakan dalam proses prediksi adalah metode C4.5 dan Naive Bayes.

1.4 Maksud dan Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah diatas, tujuan yang ingin dicapai dari penulisan penelitian ini adalah :

1. Menentukan mahasiswa yang lulus tepat waktu dan terlambat.
2. Membandingkan nilai akurasi prediksi kelulusan mahasiswa jurusan Teknik Informatika dengan menggunakan dua metode yaitu algoritma C4.5 dan Naive Bayes.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini adalah :

1. Dapat memberikan informasi prediksi kelulusan mahasiswa mahasiswa yang tepat waktu dan terlambat.
2. Dapat membantu mengevaluasi kelulusan mahasiswa dan pihak akademik.
3. Dapat dijadikan referensi bagi peneliti berikutnya.

1.6 Metode Penelitian

Dalam penelitian ini ada beberapa metode yang digunakan dalam pengumpulan data, mengolah data, dan analisis data. Metode yang digunakan yaitu:

1.6.1 Metode Pengumpulan Data

1.6.1.1 Metode Studi Dokumentasi

Metode ini dilakukan dengan mengambil data dari Direktorat Innovation Center Universitas Amikom Yogyakarta. Data yang diambil yaitu data mahasiswa alumni mengenai NIM, jenis kelamin, jumlah SKS yang diambil pada semester 6, IPK Semester 6, asal sekolah, Tanggal kelulusan.

1.6.1.2 Metode Kepustakaan

Metode ini dilakukan dengan melakukan pengumpulan data – data yang berkaitan dengan algoritma C4.5 dan Naive Bayes yang dapat dijadikan suatu landasan dalam melakukan analisa dan perbandingan. Pengumpulan data dilakukan dengan menggunakan buku, jurnal, e-book (electronic book), maupun informasi dari berbagai situs internet dengan melakukan browsing pada aplikasi browser yang ada di setiap sistem operasi, maupun segala informasi yang dapat mendukung dan menjadi penunjang dalam analisis dan perbandingan serta dapat menjadi landasan dalam penulisan penelitian ini.

1.6.2 Preprocessing Data

Pada tahap ini berisi pemrosesan data mulai dari data yang mentah yang diolah dan dilakukan konversi data sampai data tersebut siap diujikan.

1.6.3 Metode Implementasi dan Pengujian

Implementasi dilakukan untuk menyelesaikan masalah sesuai dengan perancangan proses pembentukan algoritma C4.5 dan Naive Bayes, dan dilakukan pengujian dengan menginputkan data uji mahasiswa prodi teknik informatika Universitas Amikom Yogyakarta angkatan 2014 dan 2015. Pengujian ini menerapkan metode C4.5 dan Naive Bayes dalam menentukan studi mahasiswa. Nilai akhir implementasi dan pengujian data

yaitu nilai perbandingan dan akurasi 2 metode menggunakan Confusion Matrix.

1.6.4 Analisis

Metode ini digunakan untuk memastikan bahwa implementasi yang dilakukan sudah sesuai dengan proses pembentukan algoritma C4.5 dan Naive Bayes serta memastikan bahwa proses perbandingan dan analisa sudah sesuai.

1.7 Sistematika Penulisan

Sistematikan penulis dalam penyusunan bab dalam penelitian ini :

BAB I PENDAHULUAN

Pada bab ini membahas mengenai gambaran umum yang menyajikan latar belakang masalah, rumusan masalah, maksud dan tujuan penelitian, manfaat, metode pengumpulan data, dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini menjelaskan tentang algoritma C4.5 dan Naive Bayes, serta tinjauan umum yang didapat dan segala teori yang berkaitan dengan topik pada penulisan skripsi ini.

BAB III METODE PENELITIAN

Pada bab ini tahap perancangan proses pembentukan algoritma C4.5 dan Naive Bayes serta analisa perbandingan serta menentukan nilai akurasi prediksi kelulusan mahasiswa Teknik Informatika Universitas Amikom Yogyakarta.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini menjelaskan tahap pengolahan data menggunakan metode C4.5 dan Naive Bayes serta analisis hasil perhitungan manual dan nilai akurasi dua metode.

BAB V KESIMPULAN

Pada bab ini berisi hasil akurasi pada setiap metode.

