

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Data mining adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning yang bertujuan untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang besar [1]. Data mining juga dikatakan suatu proses untuk menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan disajikan dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar untuk pengambilan keputusan [2].

Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang terdiri dari prodi D3 Manajemen Informatika, D3 Teknik Informatika, S1 Informatika, S1 Sistem Informasi, S1 Teknik Komputer, dan S1 Teknologi Informasi. Diantara prodi tersebut salah satunya yaitu prodi S1 Informatika menjadi salah satu prodi dengan jumlah mahasiswa paling banyak diminati dan lulusan terbanyak dari tahun ke tahun, tentu dengan banyaknya minat tersebut bisa dijadikan acuan sebagai bahan prediksi ketepatan kelulusan sekaligus untuk menguji kinerja algoritma C4.5.

Pada penelitian ini peneliti akan menggunakan teknik data mining dengan algoritma algoritma C4.5 yang digunakan dalam memprediksi ketepatan kelulusan Mahasiswa S1 Informatika Universitas Amikom Yogyakarta. Pohon keputusan merupakan metode yang umum digunakan untuk melakukan klasifikasi data mining. Klasifikasi merupakan suatu teknik menemukan suatu pola atau fungsi yang nantinya dapat mendeskripsikan serta memisahkan kelas data yang satu

dengan kelas data yang lainnya untuk menyatakan objek tersebut masuk pada kategori tertentu dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Banyak algoritma yang dapat digunakan untuk membangun suatu decision tree, salah satunya yaitu algoritma C4.5 [3].

Dalam berbagai penelitian telah banyak yang menggunakan algoritma data mining dan banyak juga yang telah membandingkan kinerja dari beberapa algoritma data mining, seperti pada penelitian tentang Analisis Kinerja Algoritma C4.5 Pada Sistem Pendukung Keputusan Penentuan Jenis Pelatihan [4], penelitian tentang Analisis Kinerja Decision Tree C4.5 dalam Prediksi Potensi Pelunasan Kredit Calon Debitur [5], dan penelitian tentang Perbandingan Kinerja Algoritma C4.5 dan Naïve Bayes untuk ketepatan pemilihan konsentrasi mahasiswa yang menyebutkan bahwa hasil uji kinerja algoritma klasifikasi untuk kasus ketepatan pemilihan konsentrasi mahasiswa untuk algoritma C4.5 diperoleh nilai akurasi sebesar 84,43% dari tingkat akurasi algoritma Naive Bayes [3]. Dengan demikian dapat disimpulkan bahwa metode ini cocok untuk nantinya dijadikan sebagai referensi dalam melakukan analisis kinerja algoritma c4.5 untuk prediksi ketepatan kelulusan mahasiswa SI Informatika Universitas Amikom Yogyakarta.

Oleh karena itu dalam penelitian ini peneliti akan membuat suatu Aplikasi dengan memanfaatkan algoritma data mining C4.5 sebagai algoritma yang akan di analisis hasil kerjanya yang nantinya digunakan untuk memprediksi ketepatan kelulusan Mahasiswa SI Informatika Universitas Amikom Yogyakarta. Nantinya, peneliti juga akan mencocokkan dari hasil yang telah didapat dari uji menggunakan tools rapidminer dengan program yang telah dibuat untuk membuktikan bahwa aplikasi yang dibuat telah sesuai proses perhitungannya. Rapidminer dinilai

cocok untuk proses mencocokkan hasil karena berdasarkan beberapa penelitian sebelumnya telah banyak yang menggunakan tools ini sekaligus menggunakan metode confusion matrix, dan menghasilkan tingkat Akurasi yang lebih tinggi dibanding tools lain seperti weka[3][4][7].

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan adalah bagaimana kinerja algoritma C4.5 untuk memprediksi ketepatan kelulusan Mahasiswa S1 Informatika, dan bagaimana pola yang dihasilkan dari pohon keputusan algoritma c4.5 untuk data Prediksi ketepatan kelulusan mahasiswa s1 informatika.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat diketahui rumusan masalahnya sebagai berikut:

1. Bagaimana kinerja suatu algoritma prediksi yang dibuat dengan menggunakan algoritma c4.5?
2. Bagaimana pola yang dihasilkan dari pohon keputusan algoritma c4.5 untuk data Prediksi ketepatan kelulusan mahasiswa s1 informatika ?

1.3 Batasan Masalah

Dalam perancangan dan pembuatan aplikasi ini memiliki cakupan yang cukup luas, oleh karena itu penulis membuat beberapa batasan masalah, diantaranya sebagai berikut:

1. Pembuatan dan hasil akhir aplikasi berbasis Web.
2. Pembuatan aplikasi menggunakan Bahasa *PHP* 7.4.3.
3. Memakai database *MySQL*

1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan penelitian adalah sebagai berikut:

1. Untuk mengetahui kinerja suatu algoritma prediksi yang dibuat dengan mengimplementasikan algoritma *c4.5*.
2. Untuk mengetahui pola yang dihasilkan dari pohon keputusan algoritma *c4.5* untuk data Prediksi ketepatan kelulusan mahasiswa s1 informatika.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat untuk mengetahui kinerja algoritma *C4.5*, bermanfaat untuk mengetahui pola yang dihasilkan dari pohon keputusan algoritma *c4.5* untuk data Prediksi ketepatan kelulusan mahasiswa s1 informatika.

1.6 Metode Penelitian

1.6.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Observasi

Metode pengumpulan data dengan melakukan pengamatan secara langsung ke Universitas Amikom Yogyakarta untuk memperoleh data yang dibutuhkan.

2. Wawancara

Metode pengumpulan data dengan mengadakan wawancara langsung dengan pihak yang berkaitan langsung dengan permasalahan yang sedang

dibahas pada penelitian ini untuk memperoleh gambaran dan penjelasan secara mendasar.

3. Studi Pustaka

Penulis mengumpulkan informasi yang relevan dengan topik skripsi dari berbagai referensi atau literatur. Adapun referensi tersebut adalah jurnal ilmiah, buku-buku, dan artikel-artikel yang ada di internet.

1.6.2 Metode Analisis

Metode analisis yang akan digunakan oleh penulis antara lain adalah analisis kebutuhan sistem yang diperlukan, analisis kebutuhan perangkat keras dan perangkat lunak.

1.6.3 Metode Perancangan

Aplikasi yang akan dibuat sebelumnya akan dirancang seperti menyiapkan gambaran alur pembuatan program, memilih bahasa pemrograman yang akan dipakai, dan memilih metode yang tepat untuk menganalisis hasil kinerja dari algoritma *C4.5* untuk prediksi ketepatan kelulusan mahasiswa s1 informatika universitas amikom yogyakarta.

1.6.4 Metode Pengujian

Algoritma *C4.5* yang sudah di *implementasikan* ke dalam program atau aplikasi, nantinya akan diuji dengan metode *Confusion Matrix* untuk mengetahui tingkat akurasinya.

1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini terdiri dari lima bab, yang terdiri dari sebagai berikut:

BAB I: PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, metode penelitian, metode pengumpulan data dan sistematika penulisan.

BAB II: LANDASAN TEORI

Bab ini membahas berbagai konsep dan dasar-dasar teori yang berkaitan dengan penelitian kali ini.

BAB III: METODE PENELITIAN

Bab ini menjelaskan secara detail tentang perangkat dan *software* yang digunakan dalam melakukan penelitian, serta menjelaskan secara rinci alur dari penelitian ini.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang penerapan dari alur penelitian yang telah dibuat sebelumnya, dan melihat hasil uji coba kinerja algoritma yang telah dipakai.

BAB V: PENUTUP

Bab ini berisi kesimpulan dan saran yang merupakan jawaban yang melatar belakangi skripsi ini.