

BAB I

PENDAHULUAN

1.1 Latar Belakang

Knowledge Discovery In Database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar serta hubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data (Elisa, 2017).

Data Mining yang merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. *Data mining* dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data (Sulastri, Gufroni, 2017).

Data Mining itu sendiri memiliki banyak jenis algoritma, salah satu nya yaitu algoritma C4.5. Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal, bisa mengatasi *missing data*, bisa mengatasi *data continue* dan *pruning* (Elisa, 2017).

Teknologi informasi berkembang begitu pesat dan menyentuh pada semua sisi kehidupan. Agar kegiatan manusia lebih mudah dan efisien, banyak peneliti melakukan kolaborasi antara disiplin ilmu komputer dan disiplin lain (Suryani, dkk, 2018). Termasuk salah satu kolaborasi berlangsung antara ilmu komputer dan ilmu pendidikan.

Pendidikan adalah bagian terpenting dalam kehidupan manusia karena pendidikan dapat mengubah sikap dan pola pikir manusia itu sendiri (Pertiwi, dkk,

2017). Pemerintah Indonesia telah menerapkan wajib belajar 12 tahun untuk siswa, yang terdiri dari SD 6 tahun, SMP 3 tahun dan SMA 3 tahun. Dari SD hingga SMP, setiap siswa akan mengambil mata pelajaran yang sama. Namun, saat memasuki Sekolah Menengah Atas (SMA), setiap siswa akan perlu memilih jurusan (Hidayat, dkk, 2018)

Pemilihan jurusan oleh siswa sesuai minat ketika mendaftar di sekolah menengah atas memiliki tujuan untuk memberikan kesempatan bagi siswa untuk mengembangkan sikap kompetensi dan keterampilan kompetensi siswa sesuai dengan minat, bakat, dan keterampilan akademik mereka dalam kelompok mata pelajaran dalam ilmu pengetahuan. Penentuan jurusan akan mempengaruhi tingkat akademik berikutnya dan akan mempengaruhi bidang ilmu pengetahuan atau studi untuk siswa yang akan melanjutkan ke tingkat universitas sehingga pemilihan jurusan yang salah dapat membahayakan siswa dan masa depan mereka (Nugraha, dkk, 2018).

Dalam hal ini algoritma C4.5 digunakan dalam membantu panitia Penerimaan Peserta Didik Baru (PPDB) dalam mengklasifikasikan penjurusan bagi siswa yang baru masuk di jenjang Sekolah Menengah Atas (SMA), algoritma C4.5 sendiri merupakan algoritma yang digunakan untuk membentuk pohon keputusan (*Decision Tree*). Pohon keputusan merupakan metode klasifikasi dan prediksi yang terkenal. Pohon keputusan berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target.

Diharapkan dengan menggunakan metode ini dapat menyelesaikan penelitian ini, dan dapat membantu panitia PPDB dalam melakukan tugasnya dalam menentukan jurusan bagi siswa baru dan juga dapat meminimalisir kesalahan penentuan jurusan bagi siswa, agar proses pengembangan sikap kompetensi dan keterampilan kompetensi siswa sesuai dengan minat, bakat, dan keterampilan akademik mereka dalam kelompok mata pelajaran ilmu pengetahuan memiliki hasil yang baik untuk dapat dilanjutkan ke perguruan tinggi.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah di atas, maka dapat di jabarkan perumusan masalah sebagai berikut :

1. Bagaimana mengimplementasikan algoritma C4.5 agar dapat menyelesaikan masalah dan memberikan pengetahuan untuk pengklasifikasian jurusan siswa baru pada SMA Pertiwi 2 Padang ?
2. Bagaimana mendapatkan sebuah informasi baru yang bermanfaat untuk menunjang kegiatan pada SMA Pertiwi 2 Padang dalam proses pengklasifikasian jurusan siswa baru?
3. Bagaimana algoritma C4.5 dapat melakukan pengklasifikasian data jurusan siswa secara cepat, tepat dan akurat pada BAZNAS di Kabupaten Solok?

1.3 Batasan Masalah

Agar penelitian ini tidak mengambang dari tujuan penelitian, maka penulis menyampaikan batasan masalah yang akan di gunkakan dalam tesis ini, yaitu :

1. Sistem ini memberikan rekomendasi kepada guru dalam mengambil keputusan untuk menetapkan jurusan siswa.
2. Sistem ini hanya untuk mengklasifikasikan penjurusan siswa Sekolah Menengah Atas (SMA) yang baru masuk di kelas sepuluh (X), bukan untuk jurusan untuk jenjang perguruan tinggi.
3. Penelitian ini hanya menggunakan satu metode yaitu Algoritma C4.5.

1.4 Tujuan Penelitian

Yang menjadi tujuan dari penelitian ini adalah :

1. Mengimplementasikan algoritma C4.5 untuk menentukan penjurusan siswa baru pada SMA Pertiwi 2 Padang.
2. Mengetahui apakah algoritma C4.5 dapat meningkatkan kinerja untuk pengambilan keputusan dalam proses pencarian informasi bru.
3. Mengimplementasikan algoritma C4.5 kedalam aplikasi agar dapat memberikan informasi yang cepat, tepat dan akurat kepada pihak SMA Pertiwi 2 Padang.

1.5 Manfaat Penelitian

Berdasarkan analisis permasalahan di atas, adapun manfaat penelitian ini adalah sebagai berikut:

- a. Dapat membantu memudahkan kerja panitia Penerimaan Peserta Didik Baru (PPDB) dalam menentukan jurusan siswa.
- b. Dapat mengurangi kesalahan dalam menentukan penjurusan bagi siswa baru.
- c. Dapat mempercepat proses klasifikasi penjurusan bagi siswa baru.

1.6 Sistematika Penulisan

Didalam tugas akhir ini, penulis menyusun pembahasan materi penelitian menjadi beberapa bab. Materi tersebut disusun dengan sistematika sebagai berikut :

BAB I PENDAHULUAN

Pada bab pendahuluan berisi tentang latar belakang masalah, perumusan masalah, pembatasan masalah, tujuan penelitian, metodologi penelitian yang di pakai, serta sistematika penulisan penelitian ini.

BAB II TINJAUAN PUSTAKA

Pada bab tinjauan pustaka di bahas definisi sistem, definisi data mining, sejarah data mining, penerapan data mining, keuntungan dan kerugian data mining, jenis pilihan jurusan.

BAB III METODOLOGI PENELITIAN

Pada bab metodologi penelitian membahas tentang analisa dan perancangan serta implementasi dari klasifikasi penjurusan pada sekolah menengah atas (SMA) dengan menggunakan teorama perbandingan eksponensial yaitu kerangka kerja dan uraian kerangka kerja.

BAB IV ANALISIS DAN PERANCANGAN

Pada bab ini membahas tentang analisis dan perancangan, yaitu analisis penerapan algoritma C4.5 guna membantu dalam mengklasifikasikan penjurusan bagi siswa SMA, pemrosesan data dan arsitektur.

BAB V IMPLEMENTASI DAN HASIL

Pada bab Implementasi dan hasil ini menguraikan cara implementasi dan hasil pengujian sistem yang dibangun.

BAB VI KESIMPULAN

Pada bab ini memuat kesimpulan hasil penelitian tentang sistem pendukung keputusan, penilaian resiko keberhasilan usaha, serta beberapa saran untuk pengembangan dan aplikasi lebih lanjut