

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dalam dunia pendidikan, aspek penjaminan mutu merupakan hal yang sangat penting. Perguruan tinggi menerapkan sistem penjaminan mutu pendidikan dalam rangka meningkatkan kualitas dan daya saing di tengah kompetisi baik di tingkat regional maupun global (Kandil dkk., 2010). Di dalam proses pendidikan sendiri, terdapat unsur masukan yang berupa peserta didik potensial, dan unsur keluaran yang berupa peserta didik setelah menjalani proses pembelajaran. Pembelajaran pada satu tingkat pendidikan dibagi menjadi beberapa satuan pembelajaran yang biasa disebut dengan semester. Di dalam proses pembelajaran terdapat beberapa subproses pembelajaran dengan unsur masukan dan keluaran yang lebih spesifik. Baik pada masukan maupun keluaran biasa ditandai dengan adanya ujian. Ujian masuk merupakan masalah krusial di berbagai negara (Ebrey, 2010). Ujian masuk diterapkan sekali pada calon peserta didik. Namun, peserta didik harus melalui ujian keluaran beberapa kali sesuai dengan tahapan yang ditentukan di dalam kurikulum pembelajaran. Hasil-hasil ujian dapat dijadikan salah satu tolok ukur kualitas proses pembelajaran pada perguruan tinggi tersebut (Leeuwenkamp dkk., 2017).

Penerapan teknologi untuk analisis data merupakan hal penting dalam institusi pendidikan. Dengan *data mining*, pihak institusi dapat mengeksplorasi data dengan tujuan mengekstraksi pengetahuan yang berguna dalam proses pembelajaran sehingga lebih efektif dan efisien. Hasil penambangan data berguna untuk meningkatkan praktik pendidikan dan membantu manajemen dalam proses administrasi. Secara umum, informasi yang dihasilkan oleh praktik penambangan data memiliki tujuan perbaikan metode pendidikan seperti penilaian, pemantauan, dan personalisasi proses pembelajaran. *Data mining* juga dapat digunakan untuk memodifikasi pendekatan pembelajaran (Rodrigues dkk., 2018).

Piranti *data mining* menunjukkan kemampuan untuk memprediksi tingkat keberhasilan mahasiswa di institusi pendidikan tinggi. *Data mining* dapat digunakan untuk menyelidiki karakteristik-karakteristik khusus mahasiswa yang dapat dikaitkan dengan tingkat keberhasilannya. Selain itu, juga dilakukan penyelidikan apakah data mahasiswa yang relevan tersebut tersedia untuk pendidikan tinggi yang dapat digunakan sebagai dasar untuk memprediksi tingkat keberhasilan mahasiswa (Natek dan Zwilling, 2014).

Untuk analisis data secara mendalam diperlukan teknik *data mining*. Sebagai salah satu teknik dalam *data mining*, klasifikasi menjadi penting dalam analisis karena klasifikasi dapat digunakan untuk memetakan suatu *record* informasi ke dalam kategori yang telah didefinisikan berdasarkan nilai tertentu dari atribut *record* tersebut (Han dkk., 2011).

Pembentukan pohon keputusan merupakan mekanisme alami bagi manusia dalam menghadapi suatu persoalan untuk kemudian melakukan pengambilan keputusan terbaik berdasarkan beberapa pilihan yang tersedia. Hal ini memiliki kinerja prediktif yang tinggi dengan usaha komputasi yang relatif kecil dan dapat menghasilkan model klasifikasi yang komprehensif. *Decision tree* telah terbukti menjadi salah satu pendekatan yang paling kuat dan populer untuk menemukan pola yang berguna dalam *data science* (Guan dkk., 2017).

Algoritma induksi pohon keputusan yang populer antara lain ID3 dan C4.5. Kriteria pemilahan ID3 berdasarkan pada entropi informasi. Namun, ID3 memiliki tiga kekurangan yaitu kecenderungan memilih atribut yang memiliki lebih banyak nilai, kemampuan *anti-noise* yang buruk, dan ketidakmampuan untuk menangani nilai-nilai yang hilang. C4.5 sebagai versi pengembangan dari ID3 dapat menangani atribut berkelanjutan dan nilai-nilai yang hilang. C4.5 menggunakan rasio perolehan informasi alih-alih perolehan informasi sebagai kriteria terpisah untuk memperoleh partisi data yang lebih baik, serta menggunakan metode pemangkasan kesalahan pesimis untuk mengurangi *overfitting* (Guan dkk., 2017). Bahkan, algoritma C4.5 dapat digunakan untuk melakukan prediksi atas fluktuasi harga saham (Al-Radaideh dkk., 2013).

Dengan kemampuan handal, *data mining* seharusnya dapat digunakan untuk menganalisis data ujian. Di tengah perkembangan teknologi informasi di mana sistem ujian dapat dilakukan menggunakan komputer, masih terdapat model ujian menggunakan kertas (Kim dkk., 2018). Meskipun proses validasi dan penilaian model ujian berbasis kertas dilakukan dengan komputer, masih terdapat beberapa persoalan terkait pengisian data ujian oleh peserta. Salah satu persoalan tersebut adalah kesalahan pengisian komponen data pada lembar jawab ujian. Kesalahan tersebut antara lain isian kosong, isian tidak sama dengan data yang seharusnya, atau isian terlalu tipis sehingga tidak terekam dalam sistem komputer. Hal ini tentu berakibat terhambatnya proses penilaian karena diperlukan proses pencocokan data secara manual. Namun di sisi lain, kesalahan tersebut tentu tidak dialami oleh peserta ujian berbasis komputer karena jawaban peserta dapat langsung ke proses penilaian (Karim dkk., 2016). Hal ini menyebabkan perbedaan kesempatan di antara kedua jenis model peserta meskipun dalam sebuah sistem ujian yang sama.

Penelitian ini dimotivasi adanya wacana kesetaraan hak antara peserta ujian berbasis kertas dengan peserta ujian berbasis komputer. Persoalan yang diangkat adalah bagaimana cara mengukur dan menganalisis akurasi pengisian data peserta pada lembar jawab ujian berbasis kertas dengan menggunakan *data mining*. Selain itu, penting juga untuk mengetahui komponen data peserta apa saja yang paling menentukan dalam sebuah lembar jawab. Untuk itu, penelitian ini dibatasi pada faktor kesalahan pengisian komponen nomor tes, tanggal lahir, dan kode soal. Berdasarkan analisis data kemudian diusulkan rekomendasi untuk perbaikan sistem ujian.

1.2. Tujuan Penelitian

Penelitian ini dilaksanakan dengan tujuan sebagai berikut :

1. Mengimplementasikan *data mining* untuk mengukur tingkat ketelitian pengisian data peserta pada lembar jawab ujian.
2. Menganalisis data dan model pohon keputusan yang dihasilkan dari penggunaan algoritma C4.5.

3. Menentukan komponen data peserta yang paling penting dalam lembar jawab ujian berbasis kertas.

1.3. Manfaat Penelitian

Hasil yang diperoleh dalam penelitian ini dapat digunakan oleh para pemangku kepentingan, antara lain penyelenggara ujian, pengawas ujian berbasis kertas, dan peserta ujian. Penyelenggara ujian dapat mempertimbangkan analisis akurasi dalam rangka memutuskan kebijakan penyetaraan kesempatan peserta ujian berbasis komputer dengan ujian berbasis kertas serta untuk perbaikan sistem ujian. Bagi pengawas ujian, dapat meningkat kesadarannya dalam membantu pengecekan secara manual ketelitian peserta dalam mengisi lembar jawab. Bagi peserta ujian berbasis kertas dapat meningkatkan ketelitian saat mengisi lembar jawab ujian, mengingat pentingnya ketepatan isian data untuk proses penilaian.