

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi merupakan salah satu aspek dari perkembangan zaman. Hadirnya teknologi membuat pekerjaan manusia semakin dipermudah, perkembangan suatu teknologi juga menghasilkan suatu data. Dengan berkembangnya suatu teknologi maka semakin bertambah banyak pula data yang disimpan. Hampir setiap perusahaan maupun instansi menggunakan teknologi yang berbasis database, hal tersebut akan membuat data tersimpan dengan aman.

Data yang tersimpan dalam database dapat berlangsung selama bertahun-tahun, sehingga perusahaan akan dengan mudah mencari data pada kurun waktu yang sangat lama. Namun semakin banyaknya data muncul permasalahan-permasalahan tersendiri, selain masalah *storage* atau media penyimpanan, data yang telah tersimpan kurang dapat dimanfaatkan secara maksimal, sehingga muncul cabang ilmu baru di bidang statistika yaitu data mining atau disebut juga menambang data.

Data Mining 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 (Sulastri, 2017:1). Diharapkan dengan data mining data yang telah disimpan dalam kurun waktu yang lama dan sangat banyak mampu dimanfaatkan secara maksimal. Data yang diolah dengan data mining akan menghasilkan suatu kesimpulan atau ilmu pengetahuan yang dapat diambil. Ilmu pengetahuan tersebut dapat berupa suatu prediksi atau suatu hal yang bersifat membantu para pemangku kebijakan untuk membuat kebijakan baru dengan lebih bijak.

Salah satu teknik data mining yaitu metode Naive Bayes. Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan (Saleh, 2015:3). Metode Naive Bayes sering digunakan untuk

menentukan suatu kebijakan-kebijakan atau untuk mendukung suatu keputusan-keputusan yang awalnya bersifat abstrak menjadi lebih jelas pola penyebaran datanya.

SMK Bustanul Ulum Al-ghazali merupakan Sekolah Menengah Kejuruan dibawah yayasan Pondok Pesantren Bustanul Ulum Al-ghazali. Selain memiliki SMK, yayasan tersebut juga memiliki perguruan tinggi serta banyak menjalin kerjasama dengan para pengusaha dan beberapa perusahaan. Dengan begitu SMK Bustanul Ulum Al-ghazali membutuhkan suatu sistem yang mampu mengklasifikasikan para alumni untuk melanjutkan ke jenjang pendidikan yang lebih tinggi (kuliah) atau jenjang karir (bekerja) secara tepat, konsisten, objektif serta hasil evaluasi dari klasifikasi tersebut dapat diketahui keakuratannya. Proses klasifikasi tersebut memanfaatkan data alumni yang sebelumnya hanya tersimpan tanpa dioalah lebih lanjut, selain itu hasil dari klasifikasi dapat dijadikan sebagai landasan untuk pengambilan kebijakan-kebijakan demi meningkatkan mutu para alumni.

Penelitian ini menggunakan metode Naive Bayes karena terdapat atribut-atribut yang bersifat independen atau berdiri sendiri dan tidak terkait satu dengan yang lainnya. Performa Naive Bayes yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi keindepedenan atribut (tidak ada kaitan antar atribut) (Shadiq, 2009:1). Metode Naive Bayes juga dinilai berpotensi baik dalam mengklasifikasi dokumen dibandingkan metode pengklasifikasian yang lain dalam hal akurasi dan efisiensi komputasi (Saleh, 2015:208). Pada penelitian sebelumnya juga menerapkan metode Naive Bayes untuk menentukan pilihan jurusan yang akan ditempuh untuk mahasiswa baru.

Dengan menerapkan metode Naive Bayes dimana data tahun 2015 hingga tahun 2018 dapat dimanfaatkan secara maksimal sebagai data *training* dan data testing. Data *training* dan data testing dari alumni SMK Bustanul Ulum al-ghozali dapat diolah dan menghasilkan klasifikasi berupa status alumni secara tepat, konsisten, objektif serta hasil evaluasi dari klasifikasi tersebut dapat diketahui keakuratannya. Seluruh proses hingga hasil dari pengolahan akan disajikan dalam bentuk aplikasi berbasis website agar mudah diakses dan digunakan.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas dirumuskan permasalahan sebagai berikut.

- 1) Bagaimana hasil penerapan metode Naive Bayes untuk klasifikasi status alumni pada SMK Bustanul Ulum Al-ghazali?
- 2) Berapa nilai evaluasi kinerja klasifikasi Naive Bayes ditinjau dari presisi, *recall*, dan akurasi?

1.3 Tujuan

Berdasarkan rumusan masalah diatas dapat disusun beberapa tujuan yaitu.

- 1) Mengetahui hasil penerapan metode Naive Bayes untuk klasifikasi status alumni pada SMK Bustanul Ulum Al Ghazali.
- 2) Mengetahui nilai evaluasi kinerja klasifikasi Naive Bayes ditinjau dari presisi, *recall*, dan akurasi.

1.4 Manfaat

Manfaat yang dapat diambil dari penelitian ini adalah.

- 1) Bagi pihak sekolah hasil penerapan metode Naive Bayes yang sudah diketahui dapat digunakan sebagai landasan untuk pengambilan kebijakan-kebijakan demi meningkatkan mutu lulusan SMK Bustanul Ulum Al Ghazali.
- 2) Dari nilai evaluasi kinerja klasifikasi Naive Bayes dapat dijadikan sebagai referensi untuk penelitian-penelitian selanjutnya yang berkaitan dengan metode klasifikasi.

1.5 Batasan Masalah

Agar penelitian yang dilaksanakan dapat fokus pada ruang lingkup yang dibahas, maka perlu adanya batasan masalah, batasan masalah pada penelitian ini adalah.

- 1) Dataset yang digunakan untuk *training* dan testing adalah data alumni pada tahun 2015 hingga tahun 2018 sebanyak 292 *record* data, sebanyak 262 *record* untuk data *training* yang sudah diketahui kelasnya yaitu kuliah atau bekerja dan 30 *record* untuk data testing.

Objek penelitian ini berfokus pada alumni SMK Bustanul Ulum Al-Ghazali.

