

# **BAB I**

## **PENDAHULUAN**

### **1.1 LATAR BELAKANG**

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang sedang ditempuh [1]. Pemberian beasiswa dilakukan untuk membantu siswa dalam meraih cita-cita dan dengan harapan siswa tersebut memperoleh prestasi yang baik, selain itu juga untuk membantu siswa yang kurang mampu. Dalam memberikan beasiswa harus dilakukan proses secara selektif agar sesuai dengan jenis beasiswa yang didapat. Setiap beasiswa memiliki syarat – syarat atau ketentuan yang harus dipenuhi [2]. Setiap sekolah ada beasiswa yang diberikan kepada siswa-siswa, salah satu sekolah yang memberikan beasiswa yaitu SDN 206/IV Kota Jambi.

Beasiswa yang ada di SDN 206/IV Kota Jambi Beasiswa Kurang Mampu. Beasiswa kurang mampu dilakukan melalui program PIP yaitu Program Indonesia Pintar. Proses seleksi penerimaan beasiswa di SDN 206/IV Kota Jambi sejauh ini masih dilakukan secara manual yaitu dengan menginput satu persatu data siswa ke dalam satu file kemudian dilakukan *sorting* data siswa yang seringkali menimbulkan beberapa permasalahan, antara lain membutuhkan waktu yang lama dan ketelitian yang tinggi. Prosedur pengolahan data yang dilakukan meliputi kegiatan pengumpulan data, pengelompokan, pencocokan data dengan biodata siswa, dan perkiraan siswa penerima. Pemberian beasiswa dilakukan cukup lama

dalam perhitungan dan memiliki kekurangan dalam keakuratan hasil sehingga hasilnya ada yang tidak tepat sasaran, ada siswa yang memang kurang mampu namun tidak diberikan beasiswa oleh pihak sekolah [3] Berdasarkan hasil observasi yang peneliti lakukan.

Mengacu pada permasalahan tersebut maka diperlukan sebuah metode yang dapat digunakan untuk mengurangi tingkat kesalahan dan kecurangan yaitu dengan menerapkan teknik *data mining* [4]. *Data mining* adalah serangkaian proses untuk memecahkan masalah dengan menganalisis data yang telah ada dalam *database* [5]. Penerapan data mining dinilai tepat karena memudahkan dalam melakukan penggalian informasi terhadap data SDN 206/IV Kota Jambi. Sehingga data yang telah diperoleh dapat diolah untuk mengetahui informasi-informasi yang tersembunyi dari data-data tersebut. *Data mining* sendiri memiliki beberapa metode salah satunya klasifikasi.

Metode Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui [5]. Terdapat beberapa kelebihan dan kekurangan algoritma yang dapat digunakan dalam klasifikasi *data mining* yaitu : (1) *k- Nearest Neighbor* (KNN) memiliki beberapa kelebihan, diantaranya adalah pelatihan sangat cepat, sederhana dan mudah dipelajari, tahan terhadap data pelatihan yang memiliki derau, dan efektif jika data pelatihan besar. Sedangkan, kekurangan dari KNN adalah nilai K bias, komputasi kompleks, keterbatasan memori, dan mudah tertipu dengan atribut yang tidak relevan [6]; (2) *Support Vector Machine* (SVM) kelebihanannya yaitu, dapat

diimplementasikan dengan mudah, kekurangannya berupa sulit dipakai dalam masalah berskala besar; (3) Jaringan Syaraf Tiruan (JST) ) memiliki beberapa kelebihan, diantaranya bisa memetakan berdasarkan *inputan* dan *outputan*, sangat fleksibel dengan data yang *noisy*, kekurangannya tahap dalam model JST sangat panjang, menggunakan analisis yang sangat kompleks sehingga sangat susah untuk diinterpretasikan; (4) *Naïve Bayes* mempunyai kelebihan yaitu, Kinerja *naïve bayes* masih tetap unggul ketika pengujian dilakukan pada tipe data kategori, Semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas, performa Baik, kekurangannya ukuran dari vektor fitur yang dihasilkan cukup besar dan butuh teknik untuk memperkecil ukuran vektor tersebut; (5) Algoritma C4.5 mempunyai beberapa kelebihan yaitu, dapat membuat pohon keputusan (*decision tree*) yang efisien menangani atribut tipe diskrit dan tipe diskrit- numerik, mudah untuk diinterpretasikan dan memiliki tingkat akurasi yang dapat diterima, adapun kelebihannya berupa terdapat di skalabilitas yaitu *data training* hanya dapat digunakan dan disimpan secara keseluruhan pada waktu yang bersamaan di memori [6].

Pada penelitian tentang penerimaan beasiswa sebelumnya yang berjudul “Penerapan Algoritma *Naïve Bayes* Untuk Penentuan Calon Penerima Beasiswa Pada SMK Pasim Plus Sukabumi”. Pada penelitian ini menggunakan analisa data pada siswa SMK Pasim Plus Sukabumi dengan nilai rata-rata siswa lulus seleksi 25 orang dan tidak lulus seleksi 33 orang, hasil pengujian menggunakan algoritma *Naïve Bayes* selanjutnya menggunakan rule *Confusion Matrix* dan kurva *Receiver Operating Characteristic* (ROC) dengan hasil akurasi yaitu 96,67%. Dengan

demikian, metode *Naïve Bayes* merupakan metode yang cukup baik dalam menentukan calon penerima beasiswa secara lebih efektif dan efisien [7].

Berdasarkan permasalahan dan beberapa klasifikasi diatas penulis memilih menggunakan algoritma *Naïve Bayes* dikarenakan algoritma ini sering digunakan dalam klasifikasi data dan dinilai berpotensi baik dalam mengklasifikasi data. *Naïve Bayes* juga dinilai tepat dalam pengolahan data yang sudah ada sebagai data *training* dan data yang akan diuji sebagai data *testing*.

Hal ini yang melatar belakangi penulis untuk melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul **“Penerapan *Data Mining* untuk Klasifikasi Kelayakan Penerimaan Beasiswa SDN 206/IV Kota Jambi Menggunakan Metode *Naïve Bayes*”**

## **1.2 RUMUSAN MASALAH**

Berdasarkan latar belakang masalah yang telah diuraikan diatas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana penerapan *Data Mining* untuk klasifikasi kelayakan penerimaan beasiswa SDN 206/IV Kota Jambi menggunakan metode *Naïve Bayes*?
2. Bagaimana analisa dan evaluasi *Data Mining Naïve Bayes* untuk menentukan rekomendasi kelayakan Beasiswa berdasarkan data yang ada?

## **1.3 BATASAN MASALAH**

Dalam penelitian ini penulis membatasi masalah yang terjadi dalam perusahaan hanya dalam beberapa masalah yang meliputi :

1. Objek yang menjadi sasaran penelitian adalah data siswa kelas 1 - 6 SDN 206/IV Kota Jambi, mulai dari tahun ajaran 2018 – 2022.
2. Metode yang digunakan adalah metode *Naïve Bayes* dalam aturan algoritma klasifikasi.
3. Mengklasifikasikan dan memprediksi berdasarkan data siswa dan nilai akademik siswa.
4. Data siswa yang di pakai yaitu nama siswa, tanggal lahir, jenis kelamin, alat transportasi, tingkat kelas, pekerjaan ayah, pekerjaan ibu, penghasilan orang tua, agama, tanggungan orang tua dan status penerima beasiswa.
5. Data Nilai siswa yang di pakai adalah nilai mata pelajaran yang akan di UAN kan yaitu, bahasa indonesia, matematika, dan ilmu pengetahuan alam.

## **1.4 TUJUAN DAN MANFAAT PENELITIAN**

### **1.4.1 Tujuan**

Berdasarkan rumusan masalah diatas, maka tujuan yang ingin dicapai pada tugas akhir ini adalah :

1. Menerapkan Data Mining untuk klasifikasi kelayakan penerimaan beasiswa SDN 206/IV Kota Jambi.
2. Menganalisa dan evaluasi Data Mining *Naïve Bayes* untuk menentukan rekomendasi kelayakan Beasiswa berdasarkan data yang ada.

### **1.4.2 Manfaat**

Manfaat yang diperoleh dari penelitian ini adalah :

1. Dapat menerapkan konsep *Data Mining* dengan metode *Naïve Bayes*.

2. Dapat menambah pengetahuan peneliti tentang bagaimana mencari informasi penting yang tersembunyi dalam suatu data menggunakan teknik algoritma klasifikasi.
3. Dapat mengetahui rekomendasi penerima beasiswa dengan tingkat akurasi yang tinggi.
4. Agar dapat membantu mempermudah pihak sekolah dalam menentukan kelayakan penerima Beasiswa.

## **1.5 SISTEMATIKA PENULISAN**

Sistematika penulisan menggambarkan secara umum tentang apa yang dibahas dalam setiap bab dari penulisan tugas akhir ini. Dimana penulisannya terdiri dari 5 bab, adapun isi pokok masing-masing bab sebagai berikut :

### **BAB I : PENDAHULUAN**

Pada bab ini menguraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

### **BAB II : LANDASAN TEORI**

Pada bab ini terdiri dari penjelasan mengenai studi sebelumnya yang mendukung tugas akhir dan konsep atau teori-teori yang memiliki keterkaitan terhadap topik penelitian, yaitu klasifikasi, penerimaan beasiswa, *data mining*, metode *naïve bayes*, dan memprediksi beasiswa dengan mengolah data siswa menggunakan aplikasi *Rapidminer*

**BAB III : METODOLOGI PENELITIAN**

Pada bab ini menjelaskan mengenai kerangka kerja penelitian, metode atau teknik pengumpulan data serta alat-alat dan bahan-bahan pendukung selama melakukan penelitian.

**BAB IV : ANALISIS**

Pada bab ini menjelaskan tentang gambaran umum objek penelitian, serta menganalisis dan seleksi data Siswa dengan melakukan *preprocessing* data menjadi *preprocessed* data. Transformasi data tersebut kemudian diolah dengan teknik data mining untuk mengklasifikasi kelayakan Siswa penerima program Beasiswa menggunakan metode klasifikasi algoritma *naïve bayes*. Hasil dari analisis yang telah dilakukan diinterpretasikan menjadi sebuah informasi mengenai kelayakan penerima beasiswa.

**BAB V : KESIMPULAN**

Bab ini merupakan penutup yang berisi kesimpulan yang merupakan jawaban dari masalah-masalah yang diperoleh selama penelitian dan saran-saran yang ditunjukkan kepada pihak terkait dengan hasil penelitian

