BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Pendidikan merupakan faktor utama dalam pembentukkan pribadi manusia. Pendidikan sangat berperan dalam membentuk baik atau buruknya pribadi manusia. Dengan hal tersebut, pemerintah sangat serius menangani bidang pendidikan, sebab dengan sistem pendidikan yang baik diharapkan muncul generasi penerus bangsa yang berkualitas dan mampu menyesuaikan diri untuk hidup bermasyarakat, berbangsa dan bernegara.

SD N 134/IV KOTA JAMBI adalah Sekolah yang berlokasi di Jln. Adityawarman Kelurahan Thehok Kecamatan Jambi Selatan yang memiliki berbagai macam data siswa seperti data profil, dan data hasil akademik siswa selama menempuh proses kegiatan belajar mengajar . SD N 134/IV KOTA JAMBI adalah salah satu sekolah di Kota Jambi yang memiliki program beasiswa. Data-data siswa semakin bertambah setiap tahunnya dan tidak ada tindak lanjut manfaat dari data-data yang tersedia. Padahal data - data tersebut dapat di manfaatkan dan diolah kembali untuk menjadi sebuah pengetahuan dan informasi yang bermanfaat sebagai bahan pertimbangan dalam memprediksi penentuan beasiswa.

Hingga saat ini pemberian program Beasiswa masih belum optimal dan belum tepat sasaran dikarenakan jumlah penerima Beasiswa jumlahnya tidak selalu sama

setiap tahun, ada yang meningkat dan ada juga yang menurun. Menentukan penerimaan beasiswa tersebut masih ditentukan oleh pihak sekolah tampa menggunakan sebuat metode atau aplikasi . Mengacu pada permasalahan tersebut maka diperlukan sebuah metode yang dapat digunakan untuk mengurangi tingkat kesalahan dan kecurangan yaitu dengan menerapkan teknik data mining. Penerapan data mining dinilai tepat karena memudahkan dalam melakukan penggalian informasi terhadap data di SD N 134/IV KOTA JAMBI. Sehingga data yang telah didapat dapat diolah untuk mengetahui informasi-informasi yang tersembunyi dari data-data tersebut.

Data *mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. tujuan utama data *mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki, (Mujib Ridwan et al, 2013 : 60).

Terdapat beberapa algoritma dalam metode klasifikasi, diantaranya yaitu : *Decision Tree*, *Iterative Dichotomiser 3 (ID3)*, *Naive Bayes, K-Nearest Neighbor, Support Vector Machine*

Menurut Dewi Sartika dan Dana Indra Sensuse dalam jurnalnya yang berjudul "Perbandingan Algoritma Klasifikasi *Naive Bayes, K-Nearest Neigh* dan *Decision Tree* pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian" (2017 : 153), menjelaskan bahwa :

"Terdapat 5 kategori klasifikasi yaitu berbasis *statistic,* berbasis jarak, berbasis pohon keputusan, berbasis jaringan syaraf, dan berbasis aturan. Hasil perbandingan menunjukkan bahwa metode *Iterative Dichotomiser 3 (ID3)* memiliki tingkat akurasi tertinggi dibandingkan algoritma *Naive Bayes, K-Nearest Neigh* dan *Decision Tree*, algoritma *Iterative Dichotomiser 3 (ID3)* yang digunakan algoritma *J48* atau *C4.5".*

Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan (*Decision tree*), Antara lain *ID3, CART*, dan *C4.5*. Algoritma *C4.5* merupakan pengembangan dari algoritma *ID3*, (larose, 2005).

Algoritma *C4.5* adalah algoritma yang digunakan untuk menghasilkan sebuah pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai *gain* tertinggi berdasarkan nilai *entropy* atribut tersebut sebagai poros atribut klasifikasi, (Erni Ermawati, 2016).

Berdasarkan uraian permasalahan dari penelitian-penelitian sebelumnya, maka peneliti tertarik untuk mengklasifikasikan siswa yang termasuk dalam kategori keluarga miskin atau hamper miskin, untuk klasifikasi tersebut dibutuhkan sebuah algoritma klasifikasi yang tepat. Penulis menggunakan algoritma *C4.5* dikarenakan algoritma ini sering digunakan dalam klasifikasi data dan dinilai berpotensi baik dalam mengklasifikasikan data. algoritma *C4.5* yang digunakan untuk menghasilkan sebuah pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai *gain* tertinggi berdasarkan nilai *entropy* atribut tersebut sebagai poros atribut klasifikasi, dari pohon keputusan tersebut kita

bisa menentukan hasilnya. *C4.5* juga dinilai tepat dalam pengolahan data yang sudah ada sebagai data *training* dan data yang akan diuji sebagai data *testing*.

Jadi dapat disimpulkan bahwa kenapa peneliti menggukan algoritma *C4.5* karna algoritma ini sering digunakan dalam mengklasifikasi penerapan data *mining* dan memiliki akurasi tertinggi dibandingkan algoritma *Naive Bayes, K-Nearest Neigh* dan *Decision Tree*.

Hal inilah yang melatar belakangi penulis untuk melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul: **"KLASIFIKASI KELAYAKAN PENERIMA BEASISWA SD N 134/IV KOTA JAMBI MENGGUNAKAN ALGORITMA *C4.5*".**

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang diatas masalah yang dapat dirumuskan adalah sebagai berikut :

1. Bagaimana mengklasifikasi kelayakan penerima beasiswa menggunakan algoritma *C4.5*?
2. Bagaimana mengukur nilai akurasi klasifikasi kelayakan penerima beasiswa?

1.3 BATASAN MASALAH

Untuk menghindari pembahasan yang meluas pada penelitian ini, maka penulis memberikan pembatasan masalah yaitu :

1. Pada penelitian ini peneliti menggunakan metode klasifikasi dengan algoritma *C4.5*.

2. Data yang digunakan merupakan data penerima beasiswa di SD N 134/IV KOTA JAMBI pada tahun 2019 dengan jumlah data 227 data siswa.
3. Keluaran yang dihasilkan berupa Layak" atau "Tidak Layak" sebuah siswa dalam menerima beasiswa.
4. Alat bantu (*tools*) yang digunakan pada penelitian ini adalah *WEKA*.
5. Seleksi fitur atau seleksi atribut menggunakan metode *Classifier Attribute Eval*.
6. Atribut yang digunakan dalam penelitian ini yaitu NISN, Nama, Jenis Kelamin, Alat Transportasi, Pekerjaan Ayah, Penghasilan Ayah, Pekerjaan Ibu, Penghasilan Ibu, Penerima KIP (Kartu Indonesia Pintar), Penerima PIP (Program Indonesia Pintar), Penerima KPS (Kartu Perlindungan Sosial), Status Penerima.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Berdasarkan perumusan masalah diatas, maka tujuan dari penelitian ini adalah :

1. Untuk mengklasifikasi kelayakan penerima beasiswa menggunakan algoritma *C4.5*.
2. Untuk mengukur tingkat akurasi dalam klasifikasi kelayakan penerima beasiswa menggunakan algoritma *C4.5*.

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Proses penyaluran beasiswa bisa tepat sasaran.

2. Dapat membantu pemerintah dan pihak sekolah dalam mengklasifikasikan kelayakan penerima beasiswa**.**

1.5 SISTEMATIKA PENULISAN

Laporan penelitian ini dibuat dalam sistematika yang sesuai dengan kaidah penulisan ilmiah yang benar dan dibagi dalam bab-bab sebagai berikut .

BAB I : PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian dan manfaat penelitian, serta sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini penulis membahas serta mengurai beberapa definisi yang digunakan dalam analisa *data mining* yang meliputi pengertian Data *Mining*, Klasifikasi, *C4.5*, Beasiswa, Data, dan *Tools Data Mining*.

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 penduduk dengan melakukan *preprocessing* data menjadi *prepocessed* data. Transformasi data tersebut kemudian diolah dengan teknik *data mining* untuk mengklasifikasi kelayakan penerima beasiswa menggunakan metode klasifikasi algoritma *C4.5*. Hasil dari analisis yang telah dilakukan diinterpretasikan menjadi sebuah informasi mengenai kelayakan penerima beasiswa.

BAB V : HASIL ANALISIS DAN VISUALISASI

Pada bab ini membahas tentang hasil analisis dari data siswa yang telah dihitung dengan menggunakan metode klasifikasi algoritma *C4.5*, serta memberikan klasifikasi kelayakan penerima beasiswa.

BAB VI : PENUTUP

Pada bab ini merupakan bab penutup dari laporan penelitian yang berisi kesimpulan dan saran dari penelitian yang telah dilakukan oleh penulis.