

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Biaya Pendidikan merupakan sejumlah biaya atau uang yang dibayarkan oleh peserta didik kepada Lembaga Pendidikan yang diikutinya. Selain itu, biaya Pendidikan juga dibutuhkan untuk memenuhi kebutuhan perlengkapan sekolah seperti halnya seragam, buku pelajaran, alat tulis, transportasi dan lain sebagainya. Jika kita lihat kondisi saat ini masih ditemukan beberapa warga negara yang belum mendapatkan hak untuk merasakan pendidikan sebagaimana mestinya. Mengutip data jumlah siswa putus sekolah menurut tingkat provinsi, yang dapat diakses melalui laman website *statistik.data.kemdikbud.go.id* diketahui bahwa jumlah anak putus sekolah tersebar di seluruh provinsi di Indonesia. Jumlah anak putus sekolah paling banyak berada di Jawa Barat sebanyak 10.884 siswa. Jumlah tersebut meliputi anak putus sekolah di tingkat SD, SMP, SMA dan SMK, baik negeri maupun swasta.

SMA Negeri 2 Garut merupakan salah satu sekolah yang menerima bantuan dana Pendidikan dari pemerintah. Sekolah ini berada di Kabupaten Garut, tepatnya di Kecamatan Leles. SMA Negeri 2 Garut bekerja sama dengan Dinas Pendidikan Provinsi Jawa Barat untuk menyediakan solusi dalam mengurangi angka putus sekolah yang disebabkan faktor ekonomi tersebut, salah satunya dengan memberikan beasiswa Program Indonesia Pintar.

Dari pemaparan diatas maka dapat ditarik rumusan masalahnya untuk memberikan solusi membangun proses rekomendasi penerimaan beasiswa dengan menggunakan atribut atau kriteria yang sudah ditentukan. Dalam proses penerimaan beasiswa diperlukan suatu sistem yang dapat memprediksi penerimaan beasiswa untuk mengurangi subyektifitas dalam pengambilan keputusan dan efektivitas proses penentuan penerima beasiswa dengan memproses atribut tersebut menggunakan Algoritma C4.5.

Algoritma C4.5 merupakan metode data mining yang banyak digunakan untuk melakukan pengklasifikasian data, dikarenakan algoritma C4.5 akan didapatkan sebuah pohon keputusan yang mudah dipahami dan mudah dimengerti. Algoritma C4.5 menghasilkan pohon keputusan menggunakan struktur pohon yang memiliki node dan setiap node digunakan dalam mempresentasikan atribut, kemudian cabang yang dimiliki mempresentasikan nilai dari atribut, sedangkan untuk daunnya mempresentasikan kelas. Sehingga data berikutnya akan terbentuk rule-rule solusi permasalahan (Nur Wulandari dkk., 2022).

Dalam penelitian sebelumnya terkait pengujian kelayakan algoritma C4.5 oleh (Kasus & Maal, 2018) sebagai pendukung keputusan dalam pengajuan penerimaan beasiswa dengan lima faktor penentu pada STIKOM Artha Buana menghasilkan tingkat akurasi 93.0556%, sedangkan dengan pengujian kasus yang sama, menggunakan algoritma *Naïve Bayes* menghasilkan tingkat akurasi sebesar 90.2778%.

Penelitian lain dilakukan oleh (Wibowo et al., 2019) dalam melakukan komparasi algoritma *Naïve Bayes* dengan *decision tree* untuk memprediksi lama studi mahasiswa berdasarkan perhitungan yang telah ditentukan menggunakan 100 data sample yang sama dari masing-masing metode sebagai bahan perhitungan maka didapatkan hasil akhir yaitu perhitungan menggunakan metode Decision Tree mendapatkan hasil dengan tingkat akurasi 55%, dan perhitungan menggunakan Metode *Naïve Bayes* mendapatkan tingkat akurasi sebesar 30%. Dari perhitungan tersebut dapat disimpulkan bahwa tingkat akurasi metode *decision tree* lebih tinggi dibandingkan tingkat akurasi metode *Naïve Bayes*.

Untuk menghasilkan tingkat akurasi yang jauh lebih baik, diperlukan sebuah solusi perbaikan dalam mengoptimalkan kinerja Algoritma C4.5. Salah satu solusi yang bisa di terapkan adalah dengan melakukan optimasi algoritma menggunakan metode *Particle Swarm Optimization (PSO)*. PSO memiliki beberapa kelebihan yaitu sedikit parameter, mudah diterapkan, konvergensi yang cepat, dan sederhana sehingga PSO banyak diterapkan pada optimasi fungsi, optimasi metode konvensional dan klasifikasi pola.

Pada penelitian sebelumnya yang dilakukan oleh (Taufik, 2017) membahas tentang optimalisasi fitur menggunakan metode *Particle Swarm Optimization (PSO)* untuk meningkatkan akurasi pengklasifikasian *Naïve Bayes* pada review hotel dari situs [www.tripadvisor.com](http://www.tripadvisor.com) dengan hasil peningkatan akurasi dari 90,50% menjadi 96,92%.

Selain itu, pada penelitian optimasi algoritma C4.5 untuk mengukur keputusan pembelajaran daring berbasis *Particle Swarm Optimization* (PSO) dengan model pengujian 80:20, didapat nilai akurasi sebesar 91,76% setelah dilakukan optimasi menggunakan PSO, sehingga dapat disimpulkan bahwa penerapan algoritma PSO dapat berpengaruh terhadap nilai akurasi pada setiap kriteria splitting algoritma C4.5 (Nur Wulandari dkk., 2022a).

Berdasarkan kajian yang telah dipaparkan, dapat dilihat bahwa penerapan *Particle Swarm Optimization* (PSO) dapat meningkatkan nilai akurasi yang jauh lebih baik. Maka dari itu tujuan penelitian ini adalah untuk mengukur tingkat akurasi algoritma C4.5 dengan metode *Particle Swarm Optimization* (PSO) untuk memprediksi kelayakan siswa dalam menerima bantuan dana Pendidikan guna mengurangi subyektifitas serta meningkatkan efektivitas dalam pengambilan keputusan.

## 1.2 Rumusan Masalah

1. Bagaimana mengimplementasikan algoritma C4.5 pada prediksi kelayakan siswa dalam menerima bantuan dana pendidikan di SMA Negeri 2 Garut?
2. Bagaimana performa *accuracy*, *precision* dan *recall* yang dihasilkan dari algoritma C4.5 untuk memprediksi kelayakan siswa dalam menerima bantuan dana pendidikan di SMA Negeri 2 Garut?
3. Bagaimana tingkat akurasi setelah menambahkan metode optimalisasi algoritma menggunakan *Particle Swarm Optimization* (PSO) pada hasil implementasi algoritma C4.5?

### 1.3 Batasan Masalah

1. *Dataset* yang digunakan dalam penelitian ini bersumber dari data primer yang diperoleh dari SMA Negeri 2 Garut.
2. *Dataset* siswa yang digunakan merupakan data siswa-siswa kelas X tahun pelajaran 2022/2023.
3. Algoritma yang digunakan adalah Algoritma C4.5 dan *Particle Swarm Optimization*
4. *Particle Swarm Optimization (PSO)* digunakan untuk mengatasi masalah fitur yang tidak relevan dan berlebihan.
5. Hasil yang diukur dalam penelitian ini merupakan tingkat akurasi dari prediksi kelayakan siswa dalam memperoleh bantuan dana Pendidikan di SMA Negeri 2 Garut.

### 1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu:

1. Mengimplementasikan algoritma C4.5 pada klasifikasi prediksi kelayakan siswa dalam menerima bantuan dana pendidikan di SMA Negeri 2 Garut.
2. Mengukur performa *accuracy*, *precision* dan *recall* dari algoritma C4.5 untuk memprediksi kelayakan siswa dalam menerima bantuan dana pendidikan di SMA Negeri 2 Garut.
3. Mengukur tingkat akurasi yang dihasilkan Algoritma C4.5 dengan penerapan *Particle Swarm Optimization (PSO)* untuk menghasilkan nilai akurasi yang baik.

## 1.5 Manfaat Penelitian

Dengan adanya penelitian prediksi kelayakan siswa dalam menerima bantuan beasiswa terhadap algoritma C4.5, diharapkan algoritma ini dapat membantu penelitian dalam pemilihan algoritma untuk membantu proses pengembangan penemuan algoritma dalam pemberian beasiswa kepada siswa yang membutuhkan sesuai kriteria yang berlaku.

## 1.6 Metodologi Penelitian

Dalam penelitian ini dilakukan beberapa langkah untuk memenuhi proses penelitian tersebut, diantaranya:

### 1. Pra Penelitian

#### a. Studi Pendahuluan

Data yang digunakan merupakan data primer. Data yang digunakan adalah data siswa-siswa kelas X SMA Negeri 2 Garut per tahun 2022.

#### b. Studi Pustaka

Dalam mencari sumber referensi pada penelitian ini, pencarian sumber referensi mencakup terkait Data Mining dan Prediksi. Referensi yang didapat bersumber dari jurnal, buku, dan tugas akhir mahasiswa lain yang penelitiannya berhubungan dengan data mining khususnya prediksi.

### 2. Analisis Data Mining

Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting

yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya (Angga Ginanjar Mabur, 2012).

### **1.7 Sistematika Penulisan**

Sistematika dalam penulisan ini bertujuan untuk mempermudah pembacaan dan gambaran umum tentang penelitian yang akan dilakukan dengan sistematika penulisan penelitian sebagai berikut:

#### **BAB I PENDAHULUAN**

Pada BAB I membahas diantaranya mengenai latar belakang dari penelitian, rumusan masalah, batasan masalah, tujuan penelitian yang ingin dicapai, manfaat penelitian, metodologi penelitian yang merupakan langkah untuk menyelesaikan masalah, dan sistematika penulisan.

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

Pada BAB II membahas mengenai berbagai konsep dasar teori-teori yang berkaitan dengan penelitian, yaitu teori mengenai data mining, algoritma C4.5, *Decision Tree* serta Rapidminer.

#### **BAB III METODOLOGI**

Pada BAB III membahas mengenai penjelasan kebutuhan perangkat lunak yang akan digunakan serta metode-metode yang akan digunakan dalam penelitian ini, dengan penjabaran hasil analisis. Hasil analisis dari ini digunakan untuk melakukan proses penelitian.

#### **BAB IV HASIL DAN PEMBAHASAN**

Pada BAB IV membahas terkait hasil dari proses yang telah dilakukan dari penelitian ini. Lebih tepatnya adalah membahas mengenai implementasi dari tools dan metode yang digunakan didalam proses penelitian, serta proses pengolahan data set yang selanjutnya akan ditampilkan hasil dari proses pengolahan data tersebut, disertai dengan pengujian dari data set yang telah diolah untuk mengetahui apakah algoritma yang digunakan dalam penelitian ini bisa menghasilkan prediksi yang baik atau tidak.

#### **BAB V KESIMPULAN DAN SARAN**

Pada BAB V membahas mengenai penarikan kesimpulan dari hasil penelitian yang telah dilakukan. Terlebih terkait masalah dan tujuan dari penelitian ini. Dan juga saran yang bersifat membangun untuk penelitian selanjutnya terkait penelitian ini agar dapat menutupi dan memperbaiki kekurangan-kekurangan yang ada.