

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Knowledge Discovery in Database (KDD) adalah proses menganalisis terstruktur untuk memperoleh informasi yang benar, baru, dan menemukan pola dari data yang besar dan kompleks. Data mining menjadi inti dari proses Knowledge Discovery in Database (KDD) yaitu dengan menggunakan algoritma tertentu untuk mengeksplorasi data, membangun model, dan menemukan pola yang belum diketahui. Proses text mining dengan Knowledge Discovery in Database (KDD) sama dengan pada data mining (Aldiyansyah et al., 2024). Knowledge Discovery in Databases (KDD) juga adalah aktivitas yang mencakup serta memakai data historis dengan tujuan mendapatkan keteraturan, pola serta korelasi pada perpaduan data yang besar (Pebdika et al., 2023).

Data mining merupakan proses menemukan informasi dari suatu data yang tersimpan dalam suatu database atau datasheet. Pembuatan model dilakukan dengan proses menggunakan algoritma atau rumus tertentu. Proses data mining menggunakan berbagai teknik seperti teknik dalam proses statistik, matematika, dan machine learning yang digunakan dalam melakukan identifikasi dan mengolah berbagai data menjadi informasi yang bermanfaat (Sholeh et al., 2023).

Data mining merupakan proses menemukan korelasi baru yang bermanfaat, pola dan trend dengan menambang sejumlah repository data dalam jumlah besar, menggunakan teknologi pengenalan pola seperti statistik dan teknik matematika. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan. Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer

(machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis (Bachtiar & Mahradianur, 2023)

Decision Tree C.45 merupakan sebuah algoritma yang termasuk pada teknik klasifikasi pada data mining. Proses pada decision tree bertujuan untuk mengelompokkan atau mengklasifikasikan data terhadap kelasnya masing – masing. Pada decision tree hasil dari proses yang dilakukan berupa sebuah pohon keputusan dengan rule yang didapatkan dari pohon keputusan tersebut. Pohon keputusan pada decision tree terbentuk berdasarkan dengan perhitungan nilai gain dan entropy. Atribut yang memiliki nilai gain tersebut yang didapatkan dari proses perhitungan pada pengolahan data (Arfyanti, Ita, Muhammad Fahmi, 2022)

Menurut Madhyastha dkk Tahun 2024 Pohon keputusan adalah metode pembelajaran non parametrik yang diawasi untuk klasifikasi dan regresi. Pohon keputusan memprediksi nilai variabel target dengan mempelajari aturan keputusan yang disimpulkan dari fitur data. Pada pohon klasifikasi, daun mewakili label kelas dan cabang mewakili gabungan dari fitur- fitur yang mengarah ke label kelas tersebut. Dalam pohon regresi, daun adalah target nilai nyata. Pohon dilatih dengan memisahkan secara rekursif himpunan semua pengamatan pelatihan sampai kriteria penghentian terpenuhi. Kriteria pemisahan, seperti koefisien Gini, entropi silang, dan varians menentukan bagaimana observasi dalam node tertentu dipartisi. Pengamatan yang serupa (ditentukan oleh kriteria pemisahan) tetap bersama dalam simpul yang sama.

Berdasarkan hasil eksperimen Rahmadden Tahun 2024 tentang “Implementasi Algoritma SVM dan C4.5 dalam Klasifikasi Calon Penerimaan Beasiswa”, hasil eksperimen menunjukkan bahwa kedua algoritma mampu memberikan prediksi yang cukup akurat. Akurasi model C4.5 mencapai akurasi sebesar 78.27%, yang lebih tinggi dibandingkan model SVM yaitu sebesar 75.72%. Ini menunjukkan bahwa model C4.5 lebih sering membuat prediksi yang benar dibandingkan model SVM. Presisi model C4.5 adalah 60.27%, menunjukkan bahwa model ini cukup efektif dalam meminimalkan prediksi positif palsu sedangkan model svm hanya 57.45% . Recall model C4.5 adalah 53.01%, jauh lebih tinggi dibandingkan model SVM yaitu 32.53%. Ini berarti model C4.5 lebih baik dalam menangkap semua kasus positif yang sebenarnya, yang sangat penting dalam

konteks penerimaan beasiswa. F1 Score dari model C4.5 adalah 56.41%, menggambarkan keseimbangan yang lebih di antara presisi dan recall dibandingkan dengan model SVM 41.54%. Model C4.5 menunjukkan kinerja yang lebih baik dalam semua metrik evaluasi dibandingkan model SVM. Oleh karena itu, model C4.5 disarankan untuk digunakan dalam implementasi sistem klasifikasi calon penerima beasiswa.

Dalam menentukan penerima beasiswa di sekolah secara manual, kesalahan sering terjadi menyebabkan manajemen data beasiswa yang tidak efisien dalam hal waktu dan tidak adanya kriteria yang jelas tentang bagaimana siswa dapat memperoleh beasiswa. Untuk mengantisipasi agar kesalahan tidak terjadi, diperlukan Sistem Pendukung Keputusan. Sistem Pendukung Keputusan (SPK) merupakan salah satu model yang dibangun untuk menyelesaikan masalah yang terstruktur. Untuk itulah, diperlukan suatu metode seleksi yang terstruktur, sintesis, transparan dan dapat dipertanggungjawabkan. Proses pengambilan keputusan pada dasarnya adalah memilih sesuatu yang alternatif (Kartiko, 2021)

SD Negeri 09 Berok mengalami kesulitan dalam penentuan beasiswa, dikarenakan jumlah siswa-siswi SD Negeri 09 Berok yang mengajukan permohonan 80 siswa-siswi dari 100 siswa-siswi pada tahun 2023. Dan yang diberi kesempatan mendapat beasiswa tidak menentu, biasanya berkisar kurang lebih 50 sampai 60 siswa, tergantung jumlah keluarga kurang mampu. Biasanya guru melihat dari jumlah penghasilan orangtua dan jumlah tanggungan orangtua di biodata siswa-siswinya. Dalam proses pemberian beasiswa, sekolah masih menggunakan angket dan rapat guru. Prosedur pengolahan data yang dilakukan meliputi kegiatan pengumpulan data, pengelompokan, pengurutan, perhitungan manual atau perkiraan sendiri, dan pada akhirnya menyusun dalam sejumlah bentuk laporan. Sehingga pemberian beasiswa dilakukan cukup lama yaitu kurang lebih mencapai 3 sampai 4 minggu. Kemudian saat pemberian beasiswa diberikan ke siswa-siswi banyak yang protes dikarenakan tidak tepat sasaran, ada yang memang kurang mampu tapi tidak diberikan beasiswa oleh guru. Sehingga untuk mempermudah proses pemberian beasiswa, maka perlu dibangun suatu sistem pendukung keputusan yang dapat membantu memberikan rekomendasi penerima beasiswa.

Berdasarkan pertimbangan di atas, penulis akan mencoba menggunakan metode dengan pendekatan data mining dan juga menerapkan metode Decision Tree Algoritma C4.5 untuk penentuan pemberian beasiswa di SD Negeri 09 Berok, karena metode Decision Tree Algoritma C4.5 mempunyai kelebihan yaitu dapat menggali informasi tersembunyi dalam suatu data yang besar, membagi kumpulan data yang besar menjadi himpunan-himpunan yang lebih kecil dan hasil analisa berupa diagram pohon yang mudah dimengerti.

Dengan menggunakan metode klasifikasi yaitu menilai objek data yang dimasukan ke beberapa kelas tertentu dari sejumlah kelas yang tersedia. Untuk menunjang hal tersebut maka model algoritma Decision Tree dipilih untuk dapat membantu dalam memberikan keputusan siapa yang layak untuk mendapatkan beasiswa agar tepat sasaran sehingga harus melewati seleksi dengan proses yang mudah dan cepat. Adapun metode yang diterapkan dalam proses seleksi ini dengan pemberian kriteria pada nilai siswa untuk menjadi syarat konversi nilai yang sudah ditentukan sebelumnya yang kemudian dianalisis untuk mendapatkan keputusan akhir (Andriani Hartanto, 2023). Klasifikasi dapat didefinisikan sebagai pekerjaan yang melakukan pelatihan/pembelajaran terhadap fungsi target yang memetakan setiap set atribut (fitur) ke satu jumlah label kelas yang tersedia (I Made Agus Oka Gunawan et al, 2023)

## **1.2 Perumusan Masalah**

Dalam penelitian ini penulis mengemukakan beberapa rumusan masalah :

1. Bagaimana menerapkan metode decision tree dapat mengkalsifikasikan pemberian beasiswa pada SDN 09 Berok?
2. Bagaimana merancang aplikasi Sistem pendukung keputusan dengan metode decision tree dapat menentukan pemberian beasiswa?
3. Bagaimana menerapkan dan menguji aplikasi Sistem pendukung keputusan dengan metode decision tree dapat menentukan pemberian beasiswa ?

### **1.3 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai oleh penulis dalam penelitian ini adalah sebagai berikut :

1. Mengidentifikasi faktor-faktor atau kriteria penting yang mempengaruhi kelayakan siswa dalam menerima beasiswa
2. Mengembangkan aplikasi yang dapat mengotomatiskan proses seleksi penerima beasiswa, sehingga mengurangi beban kerja manual dan meningkatkan efisiensi.
3. Mengevaluasi sejauh mana aplikasi ini mudah digunakan oleh pihak sekolah atau pengguna lain, termasuk kemudahan dalam input data, interpretasi hasil, dan navigasi antar fitur

### **1.4 Batasan Masalah**

Agar penelitian ini lebih terarah, maka perlu adanya batasan terhadap masalah yang akan dijadikan pokok pembahasan, adapun batasan tersebut adalah :

1. Data difokuskan untuk pemberian beasiswa dengan atribut yang digunakan meliputi nama, nis, jenis kelamin, jumlah penghasilan orang tua, rata rata nilai rapor, prestasi.
2. Data difokuskan pada siswa kelas 1-5 Sekolah dasar Negeri 09 Berok Tahun Ajaran 2024/2025.

### **1.5 Manfaat Penelitian**

Adapun manfaat penelitian ini diantara lain yaitu:

1. Dapat mempermudah pengambilan keputusan dalam pemilihan penerima beasiswa dalam rangka peningkatan kualitas pendidikan.
2. Memotivasi untuk melakukan penelitian selanjutnya baik pada permasalahan serupa atau berbeda dengan menggunakan metode yang sama.

3. Menambah pengetahuan dalam memahami cara pembuatan dan implementasi Sistem Pengambil Keputusan Penerima Beasiswa untuk SDN 09 Berok.

## **1.6 Sistematika Penulisan**

Sistematika Penulisan merupakan gambaran secara umum tentang apa yang akan dibahas setiap bab pada tesis ini yang terdiri dari 6 (enam) bab, antara lain:

### **BAB I : PENDAHULUAN**

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

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

Di sini penulis akan menguraikan tentang landasan teoritis yang berkaitan dengan pembuatan tesis dan bertujuan untuk menjelaskan mengenai prosedur, teori- teori, pengertian dan definisi serta pendapat para ahli.

### **BAB III : METODOLOGI PENELITIAN**

Membahas mengenai kerangka kerja penelitian dan uraian, ruang lingkup masalah, analisa masalah, analisa sistem, pengumpulan data, perancangan sistem dan membuat aplikasi perangkat lunak serta menjelaskan metode penelitian, identifikasi, pengujian sistem dan implementasi.

### **BAB IV : ANALISA DAN PERANCANGAN**

Menguraikan dan menerangkan mengenai tahapan-tahapan penyelesaian masalah, mendisain arsitektur sistem, cara kerja sistem, disain database, perancangan program dan interface.

### **BAB V : IMPLEMENTASI DAN HASIL**

Membahas mengenai implementasi sistem dan user interface,

bahasa pemrograman, pengolahan data, tampilan menu utama dan tampilan hasil program.

## **BAB VI : PENUTUP**

Memberikan kesimpulan dari keseluruhan pembahasan tesis mengenai hasil analisa, cara penggunaan aplikasi yang telah dirancang dan saran kepada pemakai aplikasi serta yang akan mengembangkan sistem ini dimasa mendatang.