

BAB II

LANDASAN TEORI

2.1. Bantuan Beasiswa

Bantuan beasiswa merupakan salah satu bentuk program pemerintah maupun lembaga pendidikan yang diberikan kepada peserta didik untuk membantu memenuhi kebutuhan pendidikan, terutama bagi siswa yang berasal dari keluarga kurang mampu (Iriani et al., 2022). Pada jenjang Sekolah Dasar Negeri (SDN), bantuan beasiswa diberikan agar siswa tetap memiliki kesempatan yang sama dalam memperoleh pendidikan tanpa terkendala oleh kondisi ekonomi keluarga. Melalui bantuan tersebut, siswa diharapkan dapat mengikuti proses belajar secara optimal sehingga tidak mengalami hambatan dalam menyelesaikan pendidikan dasar.

Bantuan beasiswa yang diberikan kepada siswa Sekolah Dasar Negeri umumnya berupa bantuan dana pendidikan yang dapat digunakan untuk memenuhi berbagai kebutuhan sekolah. Dana tersebut dapat dimanfaatkan untuk membeli perlengkapan belajar seperti buku tulis, alat tulis, tas sekolah, sepatu, seragam, maupun perlengkapan pendukung lainnya yang dibutuhkan selama proses pembelajaran (Usnaini et al., 2021). Selain itu, bantuan juga dapat digunakan untuk membantu biaya transportasi menuju sekolah atau kebutuhan pendidikan lainnya sehingga dapat mengurangi beban ekonomi orang tua dan meningkatkan semangat belajar siswa.

Dalam proses penentuan penerima bantuan beasiswa, pihak sekolah perlu mempertimbangkan berbagai kriteria agar bantuan diberikan kepada siswa yang

benar-benar membutuhkan. Beberapa kriteria yang umum digunakan antara lain pendapatan orang tua, jumlah tanggungan keluarga, nilai rapor, serta tingkat kehadiran siswa di sekolah. Apabila proses seleksi masih dilakukan secara manual, maka berpotensi menimbulkan kesalahan penilaian, membutuhkan waktu yang lebih lama, serta memungkinkan terjadinya subjektivitas dalam pengambilan keputusan. Oleh karena itu, diperlukan suatu metode yang mampu membantu sekolah melakukan proses seleksi secara lebih objektif, cepat, dan tepat sasaran.

Pemanfaatan teknologi analisis data menjadi salah satu solusi yang dapat diterapkan untuk meningkatkan kualitas proses seleksi penerima bantuan beasiswa. Algoritma C4.5 mampu mengolah data berdasarkan berbagai kriteria yang telah ditentukan sehingga menghasilkan model pohon keputusan yang dapat digunakan sebagai dasar dalam menentukan kelayakan penerima bantuan. Dengan penerapan metode tersebut, proses seleksi menjadi lebih transparan, terstruktur, dan mudah dipahami. Hasil klasifikasi yang dihasilkan diharapkan dapat membantu pihak SD Negeri 100540 dalam menentukan siswa yang layak menerima bantuan beasiswa secara lebih adil, objektif, dan sesuai dengan kondisi yang sebenarnya.

2.2. SD Negeri 100540

SD Negeri 100540 merupakan salah satu sekolah dasar yang berkomitmen memberikan pendidikan berkualitas bagi seluruh siswanya. Sekolah ini berlokasi di Huta Ibaru Sim, Desa/Kelurahan Huta Ibaru Sim, Kecamatan Dolok Sigompulon, Kabupaten Padang Lawas Utara, Provinsi Sumatera Utara, dan memiliki lingkungan belajar yang cukup kondusif serta didukung oleh tenaga

pendidik berpengalaman. Berbagai program pembinaan juga dijalankan untuk mendorong perkembangan akademik, karakter, dan kedisiplinan siswa. Keberagaman latar belakang sosial ekonomi siswa menjadikan SD Negeri 100540 aktif dalam mewujudkan pemerataan kesempatan belajar, termasuk melalui program bantuan beasiswa.

Pemanfaatan teknologi dalam pengelolaan data siswa mulai menjadi kebutuhan penting di SD Negeri 100540, terutama karena jumlah siswa yang cukup besar dan variatif. Pengambilan keputusan terkait pemberian beasiswa membutuhkan pertimbangan banyak kriteria yang sulit dikelola secara manual, sehingga sering terjadi ketidaktepatan atau keterlambatan dalam proses penilaian. Ketersediaan data seperti prestasi akademik, kondisi ekonomi keluarga, serta indikator pendukung lainnya membuka peluang bagi sekolah untuk mengadopsi metode analisis berbasis data yang lebih sistematis, objektif, dan efisien dalam menentukan siswa yang layak menerima bantuan.

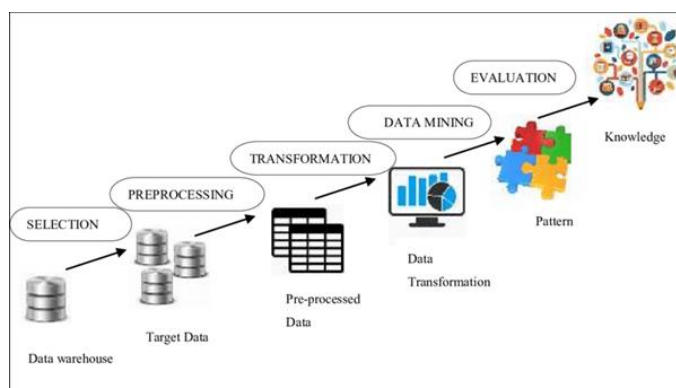
Peningkatan kualitas layanan pendidikan di SD Negeri 100540 akan semakin optimal dengan dukungan sistem yang membantu guru dan pihak sekolah membuat keputusan penting secara lebih cepat dan akurat. Pengelolaan data yang baik dapat meningkatkan transparansi, akuntabilitas, serta efektivitas alur kerja sekolah. Dengan memaksimalkan potensi data yang tersedia, sekolah tidak hanya dapat meningkatkan kepercayaan orang tua.

2.3. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan suatu proses sistematis untuk menemukan pengetahuan baru dari kumpulan data yang besar

dan kompleks (Kharis et al., 2023). KDD terdiri dari beberapa tahap utama, seperti seleksi data, praproses data, transformasi, data mining, dan interpretasi hasil. Melalui tahapan ini, data yang semula tidak terstruktur atau sulit dianalisis dapat diolah menjadi informasi yang lebih jelas, terarah, dan memiliki nilai strategis (Purwati et al., 2023). KDD banyak digunakan dalam berbagai bidang seperti pendidikan, kesehatan, bisnis, dan pemerintahan karena kemampuannya mengungkap pola tersembunyi yang tidak terlihat melalui analisis manual.

Pemanfaatan KDD menjadi sangat relevan ketika sebuah institusi memiliki data yang cukup besar dan memerlukan mekanisme analisis yang lebih sistematis untuk mendukung pengambilan Keputusan (Xu et al., 2023). Setiap tahap dalam KDD membantu mengolah data siswa menjadi lebih siap untuk dianalisis, mulai dari pembersihan data agar lebih akurat, pemilihan atribut yang sesuai, hingga proses data mining dengan algoritma tertentu untuk menemukan pola yang dapat dijadikan dasar keputusan. Pola yang dihasilkan dari proses tersebut dapat mempermudah pihak sekolah dalam menilai kelayakan siswa berdasarkan kriteria yang telah ditetapkan, sehingga keputusan yang diambil lebih objektif dan terstruktur.



Gambar 2. 1. Knowledge Discovery in Database (KDD)

KDD tidak hanya berfungsi sebagai alat analisis, tetapi juga menjadi pendekatan yang mampu meningkatkan kualitas pengelolaan informasi dalam jangka Panjang (Atalya Angelus Leza et al., 2024). Dengan menerapkan proses KDD secara berkelanjutan, institusi dapat terus memperbarui pemahaman terhadap perilaku atau karakteristik data yang dimiliki seiring perubahan waktu. Selain itu, metode ini juga mendukung pengembangan sistem yang lebih adaptif, transparan, dan efisien dalam berbagai aktivitas operasional. KDD memberikan dasar yang kuat bagi penggunaan teknologi berbasis data, yang semakin penting seiring meningkatnya kebutuhan akan keputusan yang cepat dan akurat.

2.3.1. Machine Learning

Machine learning merupakan cabang kecerdasan buatan yang berfokus pada pengembangan algoritma yang mampu belajar dari data untuk membuat prediksi atau mengambil keputusan secara otomatis (Lakhdari et al., 2023). Teknologi ini bekerja dengan menganalisis pola, hubungan, dan struktur dalam data untuk menghasilkan model yang dapat digunakan kembali tanpa harus diprogram secara eksplisit. Dengan kemampuan untuk mengolah data dalam jumlah besar dan menghasilkan keputusan berbasis pola, machine learning telah menjadi salah satu teknologi penting yang mendukung berbagai bidang, mulai dari pendidikan, kesehatan, hingga industri.

Pemanfaatan machine learning memberikan peluang besar dalam membantu institusi pendidikan membuat keputusan yang lebih cepat, akurat, dan objektif (Al Fathir As et al., 2023). Ketika sebuah sekolah harus menentukan penerima beasiswa berdasarkan banyak kriteria seperti kondisi ekonomi, prestasi akademik,

dan kehadiran siswa, machine learning mampu mengolah berbagai variabel tersebut secara sistematis (Esteban et al., 2022). Pendekatan algoritmik seperti C4.5, yang merupakan bagian dari metode klasifikasi dalam machine learning, dapat menghasilkan pohon keputusan yang mudah dipahami dan dijadikan dasar penentuan kelayakan bantuan (Hasibuan & Allistair, 2022). Dengan cara ini, proses seleksi menjadi lebih terstruktur, transparan, dan mengurangi risiko subjektivitas manusia.

Kemajuan machine learning juga mendorong lahirnya berbagai inovasi dalam pengelolaan data Pendidikan (Adnan et al., 2023). Model analisis yang dihasilkan tidak hanya membantu dalam pemilihan penerima beasiswa, tetapi juga dapat digunakan untuk memprediksi potensi siswa, menilai risiko ketertinggalan belajar, serta merancang strategi pembelajaran yang lebih efektif. Kehadiran teknologi ini memberikan peluang besar bagi sekolah untuk meningkatkan kualitas pengambilan keputusan dan menghadirkan layanan pendidikan yang lebih adaptif terhadap kebutuhan siswa.

2.3.2. Database dan Data Processing

Database dan data processing merupakan dua komponen penting dalam sistem informasi modern yang berfungsi untuk menyimpan, mengelola, dan mengolah data secara terstruktur. Database memungkinkan data disimpan dengan rapi, mudah diakses, serta dapat diperbarui kapan saja tanpa menghilangkan konsistensi dan integritas informasi. Sementara itu, data processing berperan dalam mengubah data mentah menjadi informasi yang bermakna melalui tahap pembersihan, pengelompokan, analisis, hingga penyajian hasil. Kombinasi antara

database yang baik dan proses pengolahan data yang tepat akan menghasilkan informasi yang akurat dan mendukung proses pengambilan keputusan.

Pemanfaatan database dan data processing menjadi sangat penting ketika sekolah perlu mengelola data siswa yang jumlahnya terus meningkat setiap tahun. Dalam proses seleksi beasiswa, berbagai data seperti kondisi ekonomi, nilai akademik, dan tingkat kehadiran harus disimpan secara rapi dalam database agar mudah diakses dan dianalisis. Ketika data tersedia dalam format yang terstruktur, proses pengolahan menggunakan metode seperti algoritma C4.5 dapat berjalan lebih efisien dan menghasilkan model keputusan yang lebih akurat. Hal ini membantu sekolah menggunakan data yang valid dan terorganisir sehingga keputusan yang diambil tidak hanya cepat, tetapi juga lebih objektif.

Penerapan database yang terkelola dengan baik juga berkontribusi pada transparansi dan akuntabilitas lembaga pendidikan. Sistem yang tersusun rapi mampu meminimalkan kesalahan pencatatan, mengurangi duplikasi data, serta mempercepat pelaporan ketika dibutuhkan. Selain itu, data dapat dianalisis kembali di masa mendatang untuk melihat tren, mengevaluasi kebijakan, atau meningkatkan strategi pendidikan. Dengan pengelolaan database dan pengolahan data yang maksimal, kualitas layanan sekolah meningkat dan proses administrasi menjadi lebih modern serta responsif terhadap kebutuhan siswa dan lingkungan pendidikan.

2.3.3. Visualisasi

Visualisasi merupakan teknik penyajian data dalam bentuk grafis yang memungkinkan informasi kompleks dapat dipahami dengan lebih cepat dan

intuitif. Melalui visualisasi, pola, hubungan, serta kecenderungan dalam data dapat terlihat lebih jelas dibandingkan jika disajikan dalam bentuk tabel atau teks semata. Penggunaan grafik, diagram, maupun peta data membantu pengguna dalam menginterpretasikan informasi secara lebih akurat dan efisien, sehingga visualisasi menjadi bagian penting dalam proses analisis data modern.

Pemanfaatan visualisasi sangat mendukung proses analisis terkait seleksi penerima beasiswa karena hasil pengolahan data dapat ditampilkan secara lebih terstruktur dan mudah dipahami. Grafik atau diagram yang dihasilkan dari penerapan algoritma C4.5 dapat menunjukkan distribusi kriteria siswa, tingkat kelayakan, serta pola-pola yang muncul dalam proses pengambilan keputusan. Penyajian visual tersebut memudahkan pihak sekolah dalam menilai hasil model, melakukan evaluasi, serta memastikan bahwa keputusan yang diambil benar-benar berdasarkan data yang akurat. Selain itu, visualisasi membantu meminimalkan kesalahan interpretasi dan mempersingkat waktu pemahaman terhadap hasil analisis.

Penggunaan visualisasi juga memberikan manfaat tambahan berupa peningkatan transparansi dan akuntabilitas dalam proses pengambilan keputusan. Ketika hasil analisis dapat ditampilkan secara grafis dan mudah dipahami, pihak sekolah, orang tua, maupun pemangku kepentingan lainnya dapat melihat dasar pertimbangan secara lebih jelas.

2.3.4. Statistik

Statistik merupakan ilmu yang berfungsi untuk mengumpulkan, mengolah, menganalisis, dan menyajikan data agar dapat memberikan gambaran yang jelas

mengenai suatu kondisi atau fenomena. Melalui statistik, data mentah yang awalnya sulit dipahami dapat diubah menjadi informasi yang lebih terstruktur dan mudah diinterpretasikan. Statistik juga berperan penting dalam pengambilan keputusan karena mampu menunjukkan pola, kecenderungan, serta hubungan antarvariabel yang tidak terlihat secara langsung. Dengan demikian, statistik menjadi fondasi utama dalam berbagai bidang seperti penelitian, pendidikan, ekonomi, kesehatan, dan proses evaluasi.

Pemanfaatan statistik menjadi sangat penting ketika sebuah proses pengambilan keputusan membutuhkan dasar yang objektif, termasuk dalam penentuan penerima beasiswa. Data siswa yang terdiri dari nilai akademik, kondisi ekonomi, tingkat kehadiran, serta indikator lainnya memerlukan pengolahan statistik agar dapat dibandingkan secara akurat. Analisis statistik membantu mengidentifikasi karakteristik data, menemukan variabel yang berpengaruh, serta menyediakan gambaran awal sebelum diterapkan metode seperti algoritma C4.5. Pemahaman statistik memberikan landasan kuat untuk menghasilkan model keputusan yang lebih valid dan dapat dipertanggungjawabkan.

Peran statistik tidak hanya berhenti pada tahap analisis, tetapi juga membantu dalam mengevaluasi hasil dari suatu metode atau sistem yang diterapkan. Melalui statistik, efektivitas model dapat diukur, tingkat akurasi dapat diuji, dan hasil prediksi dapat dibandingkan dengan kondisi nyata. Selain itu, statistik memungkinkan pengembang sistem atau peneliti untuk melihat sejauh mana data yang tersedia mencukupi, apakah ada ketidakseimbangan data, serta

bagaimana kualitas data tersebut memengaruhi hasil akhir. Statistik memberikan kejelasan dan arah yang lebih pasti dalam upaya pengembangan sistem berbasis data di berbagai bidang.

2.4. Algoritma C4.5

2.4.1. Pengertian Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma klasifikasi yang digunakan untuk membangun pohon keputusan berdasarkan data yang memiliki atribut dan nilai tertentu (Pascalina et al., 2022). Algoritma ini bekerja dengan cara memilih atribut terbaik sebagai akar pohon menggunakan nilai *gain* dan *gain ratio*, kemudian membagi data ke dalam cabang-cabang berdasarkan nilai atribut tersebut (Arif et al., 2022). Hasil akhirnya adalah model berupa struktur pohon yang mudah dipahami karena setiap jalur keputusan merepresentasikan aturan-aturan logis yang dibangun dari pola yang ditemukan dalam data (Sikumbang et al., 2022). Keunggulan C4.5 terletak pada kemampuannya menangani atribut kontinu, mengatasi data yang hilang, serta menghasilkan model yang cukup stabil untuk berbagai jenis kasus klasifikasi.

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i)$$

Penggunaan algoritma C4.5 sangat relevan ketika data yang dianalisis memiliki banyak kriteria yang harus dipertimbangkan secara bersamaan. Setiap atribut, seperti kondisi ekonomi keluarga, prestasi akademik, dan tingkat kehadiran, dapat diolah untuk menghasilkan aturan yang membantu proses

pengambilan keputusan secara lebih objektif. Dengan demikian, algoritma ini memberikan struktur analisis yang jelas dan terukur, sehingga pemilihan siswa yang layak menerima beasiswa dapat dilakukan berdasarkan pola dan hubungan antarvariabel yang terbentuk dari data aktual.

Daya tarik utama dari algoritma C4.5 adalah interpretabilitas hasil akhirnya yang mudah dipahami oleh pengguna non-teknis. Pohon keputusan yang dihasilkan dapat langsung diterjemahkan menjadi aturan-aturan sederhana, sehingga memudahkan pihak sekolah, staf administrasi, atau pengambil kebijakan untuk memahami dasar penilaian yang dihasilkan sistem. Selain itu, metode ini mampu meningkatkan kualitas keputusan karena setiap rekomendasi terbentuk dari analisis statistik terhadap data, bukan berdasarkan intuisi semata.

2.4.2. Konsep Dasar Algoritma C4.5

Algoritma C4.5 merupakan salah satu metode populer dalam klasifikasi yang digunakan untuk membentuk pohon keputusan berdasarkan data yang tersedia. Algoritma ini bekerja dengan memilih atribut yang memiliki nilai gain tertinggi sebagai akar atau simpul utama, kemudian membagi data ke dalam cabang-cabang berdasarkan nilai atribut tersebut. Proses ini berlanjut secara rekursif hingga seluruh data terbagi ke dalam kelas tertentu. Keunggulan C4.5 terletak pada kemampuannya menangani atribut kontinu, menangani missing value, serta menghasilkan model keputusan yang mudah dipahami karena disajikan dalam bentuk pohon yang terstruktur.

Penerapan algoritma C4.5 sangat relevan dalam proses analisis yang membutuhkan pengambilan keputusan berdasarkan banyak kriteria, termasuk

dalam penentuan kelayakan beasiswa. Setiap atribut siswa, seperti prestasi akademik, kondisi ekonomi, dan tingkat kehadiran, dapat diolah menjadi percabangan keputusan yang jelas dan objektif. Melalui perhitungan gain dan entropy, algoritma ini mampu menunjukkan atribut mana yang paling berpengaruh dalam menentukan hasil akhir, sehingga proses seleksi dapat dilakukan secara lebih sistematis. Penggunaan metode ini memberikan gambaran yang lebih transparan mengenai alasan pemilihan atau penolakan seorang siswa dalam proses seleksi beasiswa.

Pengembangan model pohon keputusan menggunakan algoritma C4.5 juga memberikan manfaat tambahan berupa kemampuan evaluasi dan validasi hasil analisis secara lebih terukur. Model yang dihasilkan dapat diuji kembali menggunakan data baru untuk menilai tingkat akurasi dan memastikan bahwa keputusan yang diambil tetap konsisten. Selain itu, struktur pohon yang dihasilkan dapat menjadi dasar bagi pengembangan sistem pendukung keputusan yang lebih otomatis, sehingga mempermudah pihak sekolah atau lembaga lain dalam melakukan analisis secara berulang tanpa harus melakukan perhitungan manual.

2.4.3. Kelebihan Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma klasifikasi yang populer karena kemampuannya dalam membentuk pohon keputusan yang mudah dipahami. Algoritma ini bekerja dengan menghitung nilai gain dan gain ratio untuk menentukan atribut terbaik pada setiap percabangan, sehingga keputusan yang dihasilkan menjadi lebih akurat. Selain itu, C4.5 mampu menangani data kontinu maupun diskrit, serta dapat mengatasi nilai data yang hilang tanpa

mengurangi kualitas model. Keunggulan ini menjadikan C4.5 sebagai metode yang fleksibel, efisien, dan sering digunakan dalam berbagai permasalahan pengambilan keputusan.

Penerapan algoritma C4.5 menjadi sangat relevan ketika proses pengambilan keputusan melibatkan banyak atribut yang saling berpengaruh dan membutuhkan analisis objektif berbasis data. Algoritma ini dapat menghasilkan model keputusan yang jelas melalui struktur pohon yang memperlihatkan urutan evaluasi dari setiap kriteria yang digunakan. Dengan demikian, hasil klasifikasi dapat dipertanggungjawabkan karena setiap cabang pohon mencerminkan hubungan logis antara data dan keputusan akhir. Kejelasan model yang dihasilkan juga memudahkan pihak pengguna dalam memahami dasar penilaian, sehingga proses keputusan menjadi lebih transparan.

Kelebihan lainnya dari C4.5 adalah kemampuannya menghasilkan model yang tetap stabil meskipun data yang digunakan memiliki karakteristik yang tidak seragam atau mengandung noise. Mekanisme pruning yang dimiliki algoritma ini membantu mengurangi risiko overfitting, sehingga model tetap mampu melakukan generalisasi yang baik ketika diterapkan pada data baru. Fleksibilitas dan ketahanan tersebut menjadikan C4.5 sebagai pilihan yang tepat untuk berbagai kebutuhan analisis klasifikasi, terutama yang membutuhkan hasil final yang mudah dibaca, cepat diproses, dan memiliki tingkat akurasi yang dapat diandalkan.

2.4.4. Kekurangan Algoritma C4.5

Algoritma C4.5 dikenal sebagai metode yang efektif dalam membangun pohon keputusan, namun tetap memiliki sejumlah kekurangan yang perlu diperhatikan. Salah satu kelemahannya adalah sensitivitas terhadap data yang tidak seimbang, di mana kategori dengan jumlah data lebih banyak cenderung lebih dominan dalam hasil keputusan. Selain itu, C4.5 dapat menghasilkan pohon yang terlalu kompleks jika data memiliki banyak atribut dengan nilai beragam, sehingga mengurangi interpretabilitas. Proses komputasinya juga bisa menjadi lebih berat ketika berhadapan dengan dataset berukuran besar atau atribut yang memiliki banyak nilai kontinu.

Penerapan C4.5 pada kebutuhan seleksi beasiswa membuat kekurangan tersebut perlu dipahami dengan baik. Data siswa yang beragam, mulai dari kondisi ekonomi, prestasi, hingga kehadiran, dapat menimbulkan kemungkinan pohon keputusan menjadi terlalu bercabang jika tidak dilakukan praproses data yang tepat. Ketidakseimbangan jumlah siswa dalam kategori tertentu—misalnya lebih banyak siswa dengan kondisi ekonomi menengah dibandingkan ekonomi rendah—juga berpotensi memengaruhi hasil pembentukan model. Oleh karena itu, pengaturan atribut, pembersihan data, dan normalisasi perlu diperhatikan agar model yang dihasilkan lebih representatif dan tetap mudah dipahami oleh pihak sekolah.

Meskipun memiliki sejumlah keterbatasan, C4.5 tetap menjadi algoritma yang banyak digunakan karena sifatnya yang jelas, mudah ditafsirkan, dan mampu mengolah data campuran baik kategori maupun numerik. Pohon keputusan yang dihasilkan dapat mempermudah pemangku kebijakan untuk melihat pola secara

langsung, terutama jika dilakukan proses penyederhanaan (pruning) agar struktur pohon lebih ringkas. Dengan optimasi yang tepat, kekurangan C4.5 dapat diminimalkan sehingga tetap memberikan hasil analisis yang bermanfaat dan dapat diandalkan.

2.5. Model Klasifikasi

Model klasifikasi merupakan salah satu teknik dalam data mining yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan pola atau karakteristik yang terdapat pada data tersebut (Di et al., 2023). Proses klasifikasi melibatkan pembelajaran dari data historis untuk kemudian digunakan dalam memprediksi kelas atau label dari data baru secara otomatis. Model ini dapat menggunakan berbagai algoritma seperti Decision Tree, Naive Bayes, K-Nearest Neighbor, maupun metode berbasis aturan lainnya. Keunggulan utama model klasifikasi adalah kemampuannya dalam memberikan hasil keputusan yang terstruktur, mudah dipahami, dan dapat digunakan untuk berbagai kebutuhan analisis prediktif.

Penerapan model klasifikasi menjadi sangat relevan ketika terdapat sejumlah data dengan kriteria tertentu yang perlu diproses untuk menghasilkan keputusan yang akurat dan objektif (Harahap et al., 2023). Kebutuhan ini muncul ketika proses pemilihan penerima beasiswa melibatkan banyak variabel seperti prestasi akademik, kondisi ekonomi keluarga, dan tingkat kehadiran yang tidak mudah dianalisis secara manual. Dengan menggunakan model klasifikasi, data siswa dapat diolah secara sistematis sehingga menghasilkan pemetaan kategori kelayakan secara lebih jelas. Proses ini memungkinkan pihak sekolah menentukan

siapa yang layak menerima bantuan berdasarkan pola yang dihasilkan dari data sebelumnya, sehingga keputusan lebih transparan dan dapat dipertanggungjawabkan.

Penggunaan model klasifikasi tidak hanya membantu proses pengambilan keputusan menjadi lebih cepat dan efisien, tetapi juga meningkatkan kualitas analisis melalui identifikasi pola yang mungkin tidak terlihat secara manual. Model ini memberikan peluang bagi institusi untuk mengoptimalkan data yang dimiliki sebagai dasar evaluasi dan perbaikan kebijakan di masa mendatang. Selain itu, hasil klasifikasi dapat digunakan untuk mengembangkan sistem pendukung keputusan yang lebih terintegrasi sehingga mampu memberikan manfaat jangka panjang dalam pengelolaan berbagai program sekolah.

2.6. Alat Bantu Analisis Data

Alat bantu analisis data menjadi komponen penting dalam proses pengolahan informasi karena mampu mempermudah pengguna dalam mengolah, memvisualisasikan, dan mengekstraksi pola dari kumpulan data. Kehadiran berbagai platform analitik membuat pengguna dapat melakukan pengolahan data tanpa harus membangun algoritma secara manual. Salah satu alat yang banyak digunakan adalah **Orange Data Mining**, yaitu aplikasi analisis data berbasis visual yang menyediakan antarmuka sederhana sehingga memudahkan pengguna dalam merancang alur analisis melalui metode drag-and-drop. Orange Data Mining mendukung berbagai teknik analisis seperti klasifikasi, clustering, regresi, visualisasi data, serta evaluasi model, sehingga sangat fleksibel digunakan dalam berbagai kebutuhan penelitian maupun pengambilan keputusan.

Orange Data Mining memberikan kemudahan dalam penerapan algoritma C4.5 karena menyediakan widget yang dapat digunakan untuk membentuk pohon keputusan berdasarkan data yang dimasukkan. Data siswa dapat diolah dengan lebih cepat, mulai dari proses seleksi data, pembersihan data, pemilihan atribut, hingga pembangunan model klasifikasi tanpa memerlukan kemampuan pemrograman yang kompleks. Alur kerja yang terstruktur pada Orange Data Mining juga memudahkan proses evaluasi model melalui berbagai fitur pengujian yang tersedia, sehingga hasil klasifikasi dapat dianalisis secara lebih objektif untuk memastikan ketepatan keputusan dalam penentuan kelayakan penerima bantuan beasiswa. Selain itu, kemampuan visualisasi pohon keputusan membantu pihak sekolah memahami pola data dan aturan keputusan yang dihasilkan secara lebih mudah.

Pemanfaatan Orange Data Mining tidak hanya mempercepat proses analisis, tetapi juga meningkatkan akurasi dan konsistensi hasil pengolahan data. Aplikasi ini memungkinkan integrasi berbagai jenis data, pengujian model secara berulang, serta penyajian hasil analisis dalam bentuk visual yang mudah dipahami. Dengan fitur yang lengkap, antarmuka yang sederhana, dan dukungan terhadap berbagai metode data mining, Orange Data Mining menjadi pilihan yang efektif untuk membantu proses analisis data dalam penelitian penentuan penerima bantuan beasiswa menggunakan algoritma C4.5.