



Implementasi Metode Decision Tree Pada Tingkat Prestasi Belajar Siswa di SMK Swasta Anak Bangsa

Nurhayati*, Saifullah, Riki Winanjaya

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1,*}yatiee650@gmail.com, ²saifullah@amiktunasbangsa.ac.id, ³riki@amiktunasbangsa.ac.id

INFORMASI ARTIKEL

Article History

Received : Mar 31, 2021

Accepted : Mar 31, 2021

Published : Mar 11, 2021

KORSPONDENSI

Email: yatiee650@gmail.com

A B S T R A K

Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Pengelompokan ini bertujuan untuk mengetahui tingkat keberhasilan siswa terhadap proses belajar yang telah dilakukan. Pendekatan yang digunakan adalah kuantitatif. Subjek dari pengelompokan ini adalah Kelas X (Sepuluh) Tahun Ajaran 2018 sampai dengan 2020. Teknik pengumpulan data yang digunakan adalah dengan test hasil belajar. Pengelompokan ini dilakukan dengan menggunakan software Data Mining Rapid Miner 5.3, dimana hasilnya akan membuktikan bahwa hasil evaluasi prestasi belajar dilakukan dengan menerapkan Algoritma C4.5. Hasil yang didapatkan nilai akurasi sebesar 71,43%, artinya bahwa rule yang dihasilkan tingkat kebenarannya mendekati 100%. Dimana hasil dari *class precision* label Berprestasi sebesar 63,89% dan label Tidak Berprestasi sebesar 92,31%. Sesuai dengan ketentuan tersebut maka hasil perhitungan manual dengan pengujian *RapidMiner* menghasilkan sebanyak 11 model aturan atau *rule* Tingkat Prestasi Siswa.

Kata Kunci: Siswa; Prestasi; *Decision Tree*; Algoritma

A B S T R A C T

Data Mining is a series of processes to explore added value in the form of knowledge that has not been known manually from a data set. This grouping aims to determine the level of student success in the learning process that has been carried out. The approach used is quantitative. The subjects of this grouping are Class X (Ten) Academic Years 2018 to 2020. The data collection technique used is the learning outcome test. This grouping is done using Data Mining Rapid Miner 5.3 software, where the results will prove that the results of the evaluation of learning achievement are carried out by applying the C4.5 Algorithm. The results obtained are an accuracy value of 71.43%, meaning that the resulting rule is close to 100% correctness. Where the results of the Class Achieving precision label is 63.89% and the label Not Achieving is 92.31%. In accordance with these provisions, the results of manual calculations with Rapid Miner testing produce 11 models of rules or rules for Student

Keywords: Student, Achievement, Decision Tree, Algorithm

1. PENDAHULUAN

Pendidikan merupakan aspek penting bagi setiap negara terutama bagi negara berkembang seperti Indonesia. Selain itu, pendidikan merupakan suatu kebutuhan yang penting bagi kehidupan setiap individu. Melalui pendidikan, individu dapat mengembangkan potensi diri dan kepribadian yang dimilikinya[1]. Hal ini sesuai dengan tujuan pendidikan nasional sebagaimana yang telah dirumuskan dalam undang-undang (UU) Republik Indonesia (RI) nomor 20 tahun 2003 pasal 3 tentang Sistem Pendidikan Nasional yaitu, "Pendidikan Nasional berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri dan menjadi warga negara yang demokratis serta bertanggungjawab".

Salah satu lembaga pendidikan formal yang mendidik para siswa menuju jenjang pendidikan perguruan tinggi adalah Sekolah Menengah Kejuruan (SMK). SMK Swasta Anak Bangsa merupakan Sekolah Menengah Kejuruan (SMK) yang memiliki 100 siswa. Untuk meningkatkan proses pendidikan yang menghasilkan lulusan yang berkualitas, juga dibutuhkan kerjasama yang baik antara pemerintah, manajemen sekolah, kemitraan dengan dunia industri atau dunia usaha, kompetensi guru, orangtua dan siswa itu sendiri. Beberapa elemen tersebut sangat

berpengaruh baik secara langsung maupun tidak langsung terhadap peningkatan mutu siswa yang pada akhirnya akan berimbas pada peningkatan kualitas pendidikan. Tingkat keberhasilan siswa sangat berpengaruh terhadap kualitas proses pembelajaran yang dilakukan di sekolah. Untuk mampu bersaing dan memiliki kualitas sekolah yang baik, upaya yang dilakukan adalah meminimalisir tingkat kegagalan siswa dengan cara melakukan proses penerapan data hasil evaluasi belajar siswa.

Berdasarkan penelitian yang dilakukan oleh [2] dengan judul “*Authoritative Parenting Practices* dan Prestasi Belajar Siswa” Atribut yang digunakan yaitu *Authoritative Parenting Practices*, prestasi belajar dan siswa dan [3] dengan judul “Prediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternalatribut yang digunakan yaitu jenis kelamin, umur, pendidikan ibu, pendidikan ayah, pekerjaan ibu, pekerjaan ayah, alasan memilih sekolah, waktu perjalanan ke sekolah, lama waktu belajar, kegagalan pada kelas sebelumnya, dukungan sekolah, kesehatan, absensi, nilai pertama, nilai kedua, dan G3 sebagai output.

Data mining adalah sebuah bidang ilmu yang berupaya menemukan pola, kaidah, aturan dan informasi berharga yang menarik dan belum diketahui sebelumnya dari sekumpulan besar data[4] [5] [6]. Data mining berisi pencarian *trens* atau pola yang diinginkan dalam *database* besar untuk membantu pengambilan keputusan di waktu yang akan datang. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk *decision tree* berdasarkan training data [7] [8] [9]. C4.5 diperkenalkan Quinlan (1996) sebagai versi terbaru dari ID3. Dalam induksi tree hanya bisa dilakukan pada fitur bertipe katagorikal, sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan. Perbaikan yang membedakan algoritma C4.5 dan ID3 adalah dapat menangani fitur dengan tipe numerik, melakukan pemotongan *decision tree* [10]. Salah satu metode Data Mining yang umum digunakan adalah *decision tree*. *Decision Tree* adalah struktur *flowchart* yang mempunyai *tree* (pohon), dimana setiap simpul *internal* menandakan suatu tes atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun mempresentasikan kelas atau distribusi kelas [7] [11] [12] [13].

2. METODOLOGI PENELITIAN

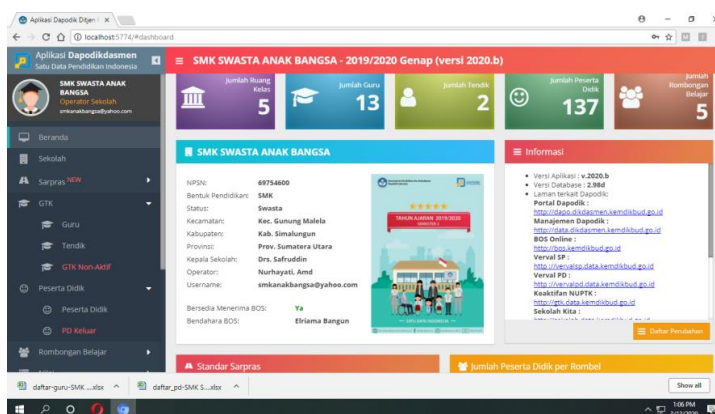
2.1 Pengumpulan Data

Metode penelitian merupakan langkah yang dimiliki dan dilakukan oleh peneliti dalam rangka untuk mengumpulkan informasi atau data serta melakukan investigasi pada data yang telah didapatkan. Adapun metode penyelesaian yang digunakan dalam penelitian ini yaitu metode *datamining* dengan *Decision Tree*. Penelitian dilakukan di Sekolah Menengah Kejuruan (SMK) Swasta Anak Bangsa, Jalan Utama No. 1 Nagori Bandar Siantar Kecamatan Gunung Malela Kabupaten Simalungun. Lamanya waktu yang dibutuhkan untuk menyelesaikan penelitian ini selama 2 minggu yang dimulai pada 08 Juni 2020 sampai dengan 22 Juni 2020.

Rancangan Penelitian ini pertama kali dilakukan dengan melakukan pengamatan (*observasi*) untuk mempelajari klasifikasi data Nilai. Hasil pengamatan kemudian dibuat percobaan mendukung, selanjutnya dilakukan eksperimen data dengan menggunakan *RapidMiner* yang merupakan *software open source* untuk membuat model aturan data yang diambil dari database pendidikan akademik dan mensurvei siswa SMK pada Tahun Pelajaran 2018 sampai dengan 2019 dan Tahun Pelajaran 2019 sampai dengan 2020 Sekolah Menengah Kejuruan (SMK) Swasta Anak Bangsa Nagori Bandar Siantar.

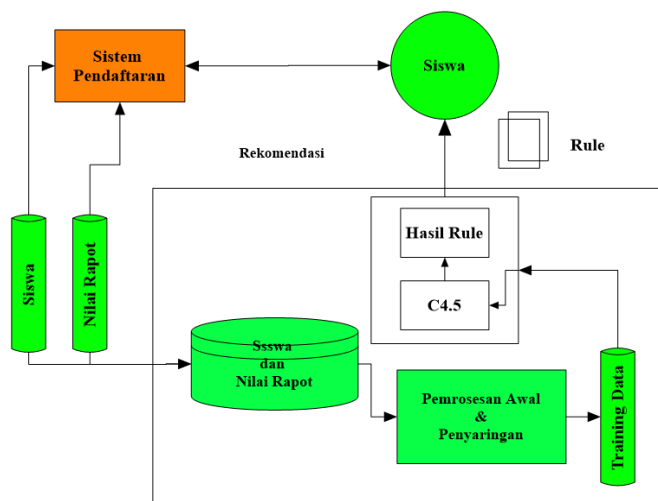
Dalam melakukan klasifikasi data dari database pendidikan akademik SMK Swasta Anak Bangsa penulis melakukan tahapan, antara lain :

- a) Mengambil data dari database pendidikan akademik, seperti pada Gambar 1



Gambar 1. Database Pendidikan Akademik

- b) Mengambil data Nilai Rapot dari pendidikan akademik.
- c) Melakukan *Data Transformation Service* (DTS) dari beberapa excel diubah ke 1 bentuk excel (.xls).
- d) Untuk mendapatkan aturan dari data tingkat prestasi siswa yang diambil dari sekumpulan data siswa dan nilai rapot, penulis menggunakan program *Rapidminer* dengan rancangan penelitiannya dapat dilihat pada Gambar 2



Gambar 2. Rancangan Penelitian Aturan tingkat prestasi siswa

Pada Gambar 2 dapat dijelaskan urutan pertama siswa melakukan sistem pendaftaran ke Sekolah Menengah Kejuruan, kemudian siswa mengikuti ujian berdasarkan mata pelajaran yang telah diajarkan oleh guru. Hasil ujian disimpan kedalam database, data nilai hasil ujian diklasifikasikan dengan melakukan penyaringan data kemudian data yang sudah konsisten dilakukan training data, apabila sampel data sudah *valid* maka dilakukan pembuatan model aturan dengan algoritma C4.5. Hasil dari model aturan ini yang dapat menunjukkan tingkat prestasi siswa dan mempermudah sekolah mencari siswa berprestasi serta meningkatkan keberhasilan mengajar guru.

2.2 Validitas dan Reliabilitas (Keakuratan) Data

Keakuratan instrumen yang digunakan pada penelitian ini adalah penting. Keakuratan mengacu apakah instrumen yang digunakan mengukur secara konsisten setiap waktu dan populasi (Gall et al., 1996). Survei dalam studi ini diuji dalam jangka waktu dan ukuran internal yang terpercaya yang memiliki keterkaitan antara bagian bagian tes (Brown and Alexander, 1991). Hal ini menjamin apakah pengukuran instrumen secara akurat dimaksudkan untuk mengukur.

2.3 Preprocessing Data

Untuk mendapatkan *input* yang lebih baik dari teknik *data mining*, penulis melakukan beberapa *preprocessing* untuk data yang akan dikumpulkan. Data disusun dalam tabel yang berbeda yang dihubungkan dalam tabel tunggal. Tabel yang digabungkan adalah :

- a) tabel Siswa
- b) tabel Nilai
- c) tabel Ekstrakurikuler
- d) table Pendidikan Ayah
- e) table Pendidikan Ib

Dari ke 5 tabel di atas, dipilih data yang digunakan untuk data penelitian yang terdiri dari data Siswa, Nilai, Ekstrakurikuler, Pendidikan Ayah dan Pendidikan Ibu.

2.4 Alat Analisis Data

RapidMiner adalah platform perangkat lunak data ilmu pengetahuan yang dikembangkan oleh perusahaan dengan nama yang sama, yang menyediakan lingkungan terpadu untuk pembelajaran mesin (*machine learning*), pembelajaran mendalam (*deep learning*), penambangan teks (*text mining*), dan analisis prediktif (*predictive analytics*). Aplikasi ini digunakan untuk aplikasi bisnis dan komersial serta untuk penelitian, pendidikan, pelatihan, pembuatan *prototype* dengan cepat, dan pengembangan aplikasi serta mendukung semua langkah proses pembelajaran mesin termasuk persiapan data, visualisasi hasil, validasi dan pengoptimalan. RapidMiner dikembangkan dengan model *open core* [14] [15].

Rapidminer 5.3 digunakan pada studi ini untuk menampilkan kelompok siswa pada kumpulan data dan memperlihatkan matriks presentasi yang tersebar dari kelompok data.

2.5 Instrument Penelitian

Dalam penelitian ini, penulis melakukan beberapa data set file dalam bentuk *spreadsheet* file excel 2003 (xls). Transformasi data ini diperlukan sebagai masukan untuk perangkat lunak yang akan digunakan dalam penelitian ini. Dalam pengolahan data transformasi, penulis menggunakan perangkat lunak rapidminer versi 5.3 Instrumen penelitian.

3. HASIL DAN PEMBAHASAN

Dalam implementasi hasil akhir dari penerapan Algoritma C4.5 dibagi ke dalam dua tahap yaitu proses perhitungan manual menggunakan Algoritma C4.5 dan penyesuaian hasil perhitungan manual melalui pengujian menggunakan *software RapidMiner 5.3*. Data yang digunakan adalah DAPODIK SMK Swasta Anak Bangsa sebagai *input* untuk membuat model aturan menggunakan Algoritma C4.5. Pohon keputusan digunakan untuk membuat model aturan yang akan dipilih dalam pengambilan keputusan. Berikut ini tahapan pengolahan data dengan Algoritma C4.5 untuk memperoleh model aturan pohon keputusan tingkat prestasi belajar siswa sesuai data hasil kuesioner yang diperoleh.

3.1 Proses Perhitungan Algoritma C4.5

Perhitungan Algoritma C4.5 untuk memperoleh model aturan pohon keputusan dapat diuraikan sebagai berikut :
Langkah 1: Menghitung jumlah kasus, jumlah kasus untuk keputusan Berprestasi, jumlah kasus untuk keputusan Tidak Berprestasi.

Langkah 2: Menghitung *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan kelas atribut dengan persamaan (1). Selanjutnya dilakukan penghitungan *Gain* untuk masing-masing atribut dengan persamaan (2). Berikut ini adalah perhitungan nilai *entropy* dan *gain*.

Menghitung entropy total :

$$\begin{aligned} Entropy \text{ [Total]} &= \left(-\frac{47}{100} \times \log_2\left(\frac{47}{100}\right)\right) + \left(-\frac{53}{100} \times \log_2\left(\frac{53}{100}\right)\right) \\ &= 0,997401589 \end{aligned}$$

Menghitung entropy dan gain Kehadiran :

$$Entropy \text{ [Kehadiran – Hadir]} = \left(-\frac{47}{100} \times \log_2\left(\frac{47}{100}\right)\right) + \left(-\frac{53}{100} \times \log_2\left(\frac{53}{100}\right)\right) = 0,997401589$$

$$Entropy \text{ [Kehadiran – Tidak Hadir]} = 0$$

$$Gain \text{ [Kehadiran]} = 0,997401589 - \left(\left(\frac{100}{100} \times 0,997401589\right) + \left(\frac{0}{100} \times 0\right)\right) = 0$$

Menghitung entropy dan gain Nilai :

$$Entropy \text{ [Nilai – Tinggi]} = 0$$

$$Entropy \text{ [Nilai – Sedang]} = \left(-\frac{38}{60} \times \log_2\left(\frac{38}{60}\right)\right) + \left(-\frac{22}{60} \times \log_2\left(\frac{22}{60}\right)\right) = 0,948078244$$

$$Entropy \text{ [Nilai – Rendah]} = \left(-\frac{9}{40} \times \log_2\left(\frac{9}{40}\right)\right) + \left(-\frac{31}{40} \times \log_2\left(\frac{31}{40}\right)\right) = 0,769192829$$

$$Gain \text{ [Nilai]}$$

$$= 0,997401589 - \left(\left(\frac{0}{100} \times 0\right) + \left(\frac{60}{100} \times 0,948078244\right) + \left(\frac{40}{100} \times 0,769192829\right)\right) = 0,120877511$$

Menghitung entropy dan gain Ekstrakurikuler :

$$Entropy \text{ [Ekstrakurikuler – Ya]} = \left(-\frac{41}{64} \times \log_2\left(\frac{41}{64}\right)\right) + \left(-\frac{23}{64} \times \log_2\left(\frac{23}{64}\right)\right) = 0,942163169$$

$$Entropy \text{ [Ekstrakurikuler - Tidak]} = \left(-\frac{6}{36} \times \log_2\left(\frac{6}{36}\right)\right) + \left(-\frac{30}{36} \times \log_2\left(\frac{30}{36}\right)\right) = 0,650022422$$

$$Gain \text{ [Ekstrakurikuler]} = 0,997401589 - \left(\left(\frac{64}{100} \times 0,942163169\right) + \left(\frac{36}{100} \times 0,650022422\right)\right) = 0,160409089$$

Menghitung entropy dan gain Pendidikan Ayah:

$$Entropy \text{ [Pendidikan Ayah – S1]} = 0$$

$$Entropy \text{ [Pendidikan Ayah – SMA]} = \left(-\frac{22}{29} \times \log_2\left(\frac{22}{29}\right)\right) + \left(-\frac{7}{29} \times \log_2\left(\frac{7}{29}\right)\right) = 0,79732651$$

$$Entropy \text{ [Pendidikan Ayah – SMP]} = \left(-\frac{22}{49} \times \log_2\left(\frac{22}{49}\right)\right) + \left(-\frac{27}{49} \times \log_2\left(\frac{27}{49}\right)\right) = 0,992476004$$

$$Entropy \text{ [Pendidikan Ayah – SD]} = \left(-\frac{2}{21} \times \log_2\left(\frac{2}{21}\right)\right) + \left(-\frac{19}{21} \times \log_2\left(\frac{19}{21}\right)\right) = 0,453716339$$

$$Gain \text{ [Ekstrakurikuler]}$$

$$\begin{aligned} &= 0,997401589 - \left(\left(\frac{2}{100} \times 0\right) + \left(\frac{29}{100} \times 0,79732651\right) + \left(\frac{41}{100} \times 0,992476004\right) + \left(\frac{21}{100} \times 0,453716339\right)\right) \\ &= 0,184583227 \end{aligned}$$

Menghitung entropy dan gain Pendidikan Ibu :

$$Entropy \text{ [Pendidikan Ibu– S1]} = 0$$

$$Entropy \text{ [Pendidikan Ibu – SMA]} = \left(-\frac{23}{27} \times \log_2\left(\frac{23}{27}\right)\right) + \left(-\frac{4}{27} \times \log_2\left(\frac{4}{27}\right)\right) = 0,605186577$$

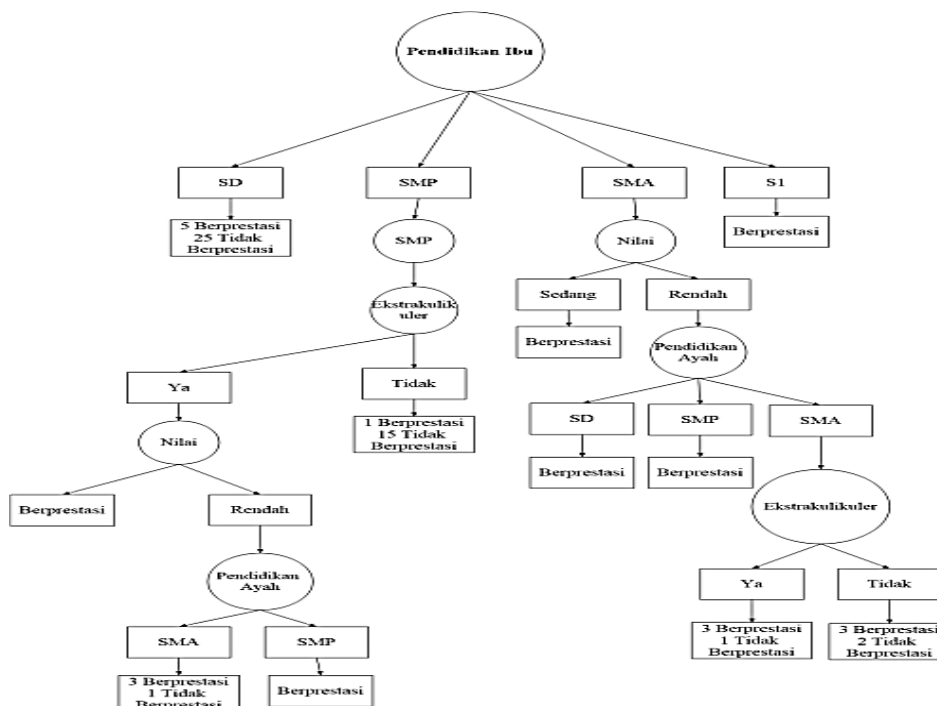
$$Entropy \text{ [Pendidikan Ibu– SMP]} = \left(-\frac{17}{41} \times \log_2\left(\frac{17}{41}\right)\right) + \left(-\frac{24}{41} \times \log_2\left(\frac{24}{41}\right)\right) = 0,978869851$$

$$Entropy \text{ [Pendidikan Ibu – SD]} = \left(-\frac{5}{30} \times \log_2\left(\frac{5}{30}\right)\right) + \left(-\frac{25}{30} \times \log_2\left(\frac{25}{30}\right)\right) = 0,650022422$$

$$\text{Gain [Pendidikan Ibu]} = 0,997401589 - \left(\left(\frac{2}{100} \times 0 \right) + \left(\frac{27}{100} \times 0,605186577 \right) + \left(\frac{41}{100} \times 0,978869851 \right) + \left(\frac{30}{100} \times 0,650022422 \right) \right) = 0,237657848$$

3.2 Pohon Keputusan

Berikut ini adalah proses pemodelan pohon keputusan berdasarkan hasil perhitungan Algoritma C4.5 :



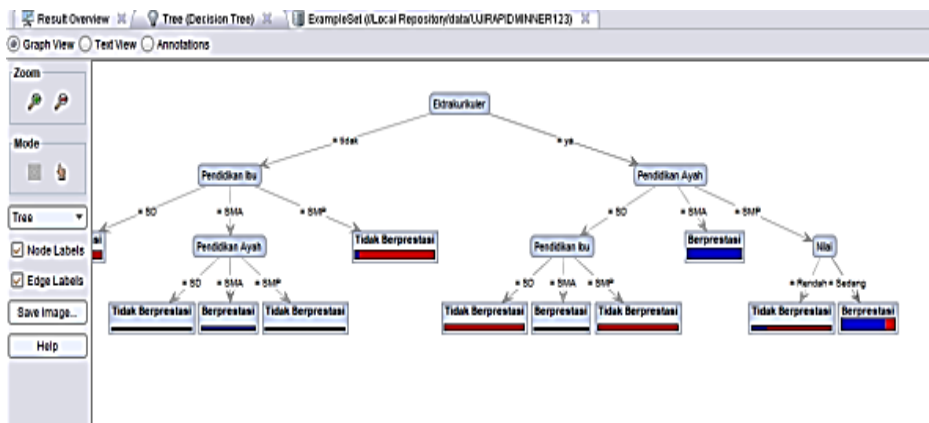
Gambar 3. Pohon Keputusan Hasil Perhitungan Algoritma C4.5

Dari perhitungan diatas terdapat 11 (Sebelas) *rules* yang dapat di jadikan sebagai referensi dalam menentukan tingkat prestasi siswa. Adapun aturan atau *rule* yang terbentuk berdasarkan pohon keputusan pada Gambar 3.1. di atas yaitu 7 (Tujuh) *rules* keputusan berprestasi dan 4 (empat) *rules* keputusan tidak berprestasi dijelaskan melalui teks narasi sebagai berikut:

- 1) Jika Ekstrakurikuler = Ya, Pendidikan Ayah = SD dan Pendidikan Ibu = SD, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 0, Tidak Berprestasi = 7}
- 2) Jika Ekstrakurikuler = Ya, Pendidikan Ayah = SD dan Pendidikan Ibu = SMA, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 1, Tidak Berprestasi = 1}
- 3) Jika Ekstrakurikuler = Ya, Pendidikan Ayah = SD dan Pendidikan Ibu = SMP, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 0, Tidak Berprestasi = 7}
- 4) Jika Ekstrakurikuler = Ya dan Pendidikan Ayah = SMA, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 20, Tidak Berprestasi = 0}
- 5) Jika Ekstrakurikuler = Ya, Pendidikan Ayah = SMA dan Nilai = Rendah maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 1, Tidak Berprestasi = 4}
- 6) Jika Ekstrakurikuler = Ya, Pendidikan Ayah = SMA dan Nilai = Sedang maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 19, Tidak Berprestasi = 4}
- 7) Jika Ekstrakurikuler = Tidak dan Pendidikan Ibu = SD, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 0, Tidak Berprestasi = 12}
- 8) Jika Ekstrakurikuler = Tidak, Pendidikan Ibu = SMA dan Pendidikan Ayah = SD, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 0, Tidak Berprestasi = 2}
- 9) Jika Ekstrakurikuler = Tidak, Pendidikan Ibu = SMA dan Pendidikan Ayah = SMA, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 4, Tidak Berprestasi = 0}
- 10) Jika Ekstrakurikuler = Tidak, Pendidikan Ibu = SMA dan Pendidikan Ayah = SMP, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 1, Tidak Berprestasi = 1}
- 11) Jika Ekstrakurikuler = Tidak dan Pendidikan Ibu = SMP, maka hasilnya Berprestasi dan Tidak Berprestasi {Berprestasi = 1, Tidak Berprestasi = 15}

3.3 Hasil Percobaan

Pada tahap akhir penerapan Algoritma C4.5 dilakukan penyesuaian hasil perhitungan manual melalui pengujian menggunakan *software RapidMiner 5.3*. Hasil pengolahan data dengan model pohon keputusan sesuai dengan *software RapidMiner*, dapat dilihat pada Gambar 4 sebagai berikut :



Gambar 4. Decision Tree Pada Rapidminer

Gambar 4 diatas merupakan pohon keputusan yang dihasilkan pada Rapidminer dengan aturan atau rule yang dapat dilihat pada text view pada Gambar 5 berikut :

```

Tree
Ekstrakurikuler = tidak
| Pendidikan Ibu = SD: Tidak Berprestasi {Berprestasi=0, Tidak Berprestasi=12}
| Pendidikan Ibu = SMA
| | Pendidikan Ayah = SD: Tidak Berprestasi {Berprestasi=0, Tidak Berprestasi=2}
| | Pendidikan Ayah = SMA: Berprestasi {Berprestasi=4, Tidak Berprestasi=0}
| | Pendidikan Ayah = SMP: Tidak Berprestasi {Berprestasi=1, Tidak Berprestasi=1}
| Pendidikan Ibu = SMP: Tidak Berprestasi {Berprestasi=1, Tidak Berprestasi=15}
Ekstrakurikuler = ya
| Pendidikan Ayah = SD
| | Pendidikan Ibu = SD: Tidak Berprestasi {Berprestasi=0, Tidak Berprestasi=7}
| | Pendidikan Ibu = SMA: Berprestasi {Berprestasi=1, Tidak Berprestasi=1}
| | Pendidikan Ibu = SMP: Tidak Berprestasi {Berprestasi=0, Tidak Berprestasi=7}
| Pendidikan Ayah = SMA: Berprestasi {Berprestasi=20, Tidak Berprestasi=0}
| Pendidikan Ayah = SMP
| | Nilai = Rendah: Tidak Berprestasi {Berprestasi=1, Tidak Berprestasi=4}
| | Nilai = Sedang: Berprestasi {Berprestasi=19, Tidak Berprestasi=4}
    
```

Gambar 5. Rule Decision Tree Pada Rapidminer

Hasil penerapan Algoritma C4.5 menggunakan software RapidMiner dengan operator Split Validation diperoleh nilai akurasi yaitu sebesar 71,43%. Hasil akurasi tersebut diperoleh dengan pengaturan pada operator split validation dengan nilai split ratio = 0,5 dan sampling type = linear sampling. Berikut ini adalah hasil akurasi yang diperoleh.

	true Berprestasi	true Tidak Berprestasi	class precision
pred. Berprestasi	23	13	63.89%
pred. Tidak Berprestasi	1	12	92.31%
class recall	95.83%	48.00%	

Gambar 6. Nilai Akurasi Algoritma C4.5

PerformanceVector

PerformanceVector:
 accuracy: 71.43%

ConfusionMatrix:
 True: Berprestasi Tidak Berprestasi
 Berprestasi: 23 13
 Tidak Berprestasi: 1 12

classification_error: 28.57%

ConfusionMatrix:
 True: Berprestasi Tidak Berprestasi
 Berprestasi: 23 13
 Tidak Berprestasi: 1 12

Gambar 7. Performance Vektor Algoritma C4.5

Dari gambar diatas nilai Accuracy sebesar 71,43%, artinya aturan atau rule yang dihasilkan mendekati 100%, dimana untuk Class Precision pada prediksi label Prestasi sebesar 63,89% dan prediksi label Tidak Berprestasi sebesar 92,31%.

3.4 Pembahasan

Hasil yang dilakukan peneliti dalam perhitungan Algoritma C4.5 diperoleh 11 model aturan atau *rule* tingkat prestasi siswa. Model aturan dalam bentuk pohon keputusan yang diperoleh oleh peneliti dapat dilihat pada Gambar 5. Berdasarkan pengolahan data menggunakan *software RapidMiner* didapat nilai akurasi sebesar 71,43%, artinya bahwa *rule* yang dihasilkan tingkat kebenarannya mendekati 100%. Dimana hasil dari *class precision* label berprestasi sebesar 63,89% dan label Tidak berprestasi sebesar 92,31%. *Parameters* yang digunakan pada *decision tree* telah dilakukan penyesuaian terhadap kriteria *decision tree* yang digunakan pada *RapidMiner* yaitu *information gain*, *maksimal depth*= 20, *confidence*= 0,25, *minimal gain*= 0,1, *minimal leaf size*= 1, *minimal size for split*= 2, *number of pruning alt*= 3. Sesuai dengan ketentuan tersebut maka hasil perhitungan manual dengan pengujian *RapidMiner* menghasilkan sebanyak 11 model aturan atau *rule* tingkat prestasi siswa. Model aturan dalam bentuk pohon keputusan yang dihasilkan dari *Rapidminer*. Artinya hasil proses yang dilakukan peneliti pada perhitungan Algoritma C4.5 dan *Rapidminer* diperoleh hasil yang sama dan sesuai. Sehingga pengujian dengan *RapidMiner* dapat dikatakan berhasil dan dapat menemukan pohon keputusan pada kasus tingkat prestasi siswa.

4. KESIMPULAN

Berdasarkan pengolahan data menggunakan *software RapidMiner* didapat nilai akurasi sebesar 71,43%, artinya bahwa *rule* yang dihasilkan tingkat kebenarannya mendekati 100%. Dimana hasil dari *class precision* label Berprestasi sebesar 63,89% dan label Tidak Berprestasi sebesar 92,31%. *Parameters* yang digunakan pada *decision tree* telah dilakukan penyesuaian terhadap kriteria *decision tree* yang digunakan pada *RapidMiner* yaitu *information gain*, *maksimal depth*= 20, *confidence*= 0,25, *minimal gain*= 0,1, *minimal leaf size*= 1, *minimal size for split*= 2, *number of pruning alt*= 3. Sesuai dengan ketentuan tersebut maka hasil perhitungan manual dengan pengujian *RapidMiner* menghasilkan sebanyak 11 model aturan atau *rule* Tingkat Prestasi Siswa. Model aturan dalam bentuk pohon keputusan yang dihasilkan dari *Rapidminer*. Artinya hasil proses yang dilakukan peneliti pada perhitungan Algoritma C4.5 dan *Rapidminer* diperoleh hasil yang sama dan sesuai. Sehingga pengujian dengan *RapidMiner* dapat dikatakan berhasil dan dapat menemukan pohon keputusan pada kasus penerima Beasiswa.

REFERENCES

- [1] M. D. Mayangsari, "Motivasi Berprestasi Mahasiswa Ditinjau Dari Penerimaan Orangtua," *J. Ecopsy*, vol. 1, no. 1, pp. 21–27, 2016, doi: 10.20527/ecopsy.v1i1.480.
- [2] Z. Anggraini and U. Ridha, "Authoritative Parenting Practices Dan Prestasi Belajar Siswa Sma Di Banda Aceh," *J. Psikol. Undip*, vol. 16, no. 1, p. 20, 2017, doi: 10.14710/jpu.16.1.20-31.
- [3] R. H. Pambudi, B. D. Setiawan, and Indriati, "Penerapan Algoritma C4 . 5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2637–2643, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id5>.
- [4] P. G. S. C. Nugraha, I. W. Aribawa, I. P. O. Priyana, and G. Indrawan, "Penerapan Metode Decision Tree(Data Mining) Untuk Memprediksi Tingkat Kelulusan Siswa Smpn1 Kintamani," *Semin. Nas. Vokasi dan Teknol.*, pp. 35–44, 2016.
- [5] Y. S. Luvia, A. P. Windarto, S. Solikhun, and D. Hartama, "Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Keberhasilan Mahasiswa Di Amik Tunas Bangsa," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 1, no. 1, p. 75, 2017, doi: 10.30645/jurasik.v1i1.12.
- [6] E. Hasmin, S. Aisa, and T. Informatika, "Penerapan Algoritma C4 . 5 Untuk Penentuan Penerima Beasiswa Mahasiswa," vol. 5, no. 2, pp. 308–320, 2019.
- [7] Rismayanti, "Implementasi Algoritma C4.5 Untuk Menentukan Penerima Beasiswa Di Stt Harapan Medan," *Media Infotama*, vol. 12, no. 2, pp. 116–120, 2016.
- [8] R. Winanjaya, F. Amir, and R. Doni, "Penerapan Data Mining Untuk Memprediksi Penerimaan Peserta Didik Baru Menggunakan Algoritma C4.5," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 1, 2019, doi: 10.30645/senaris.v1i0.1.
- [9] N. L. Ratniasih, "Optimasi Data Mining Menggunakan Algoritma Naïve Bayes Dan C4.5 Untuk Klasifikasi Kelulusan Mahasiswa," *J. Teknol. Inf. dan Komput.*, vol. 5, no. 1, pp. 28–34, 2019, doi: 10.36002/jutik.v5i1.634.
- [10] K. Malik and M. Faid, "Prediksi Prestasi Siswa Smp Nurul Jadid Menggunakan Algoritma C4.5," *Nusant. J. Comput. its Appl.*, vol. 2, no. 4, pp. 2–5, 2016.
- [11] P. Kasih, "Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara," *Innov. Res. Informatics*, vol. 1, no. 2, pp. 63–69, 2019, doi: 10.37058/innovatics.v1i2.918.
- [12] M. Hamidah, H. Fitriyah, and I. Arwani, "Implementasi Decision Tree pada Penentuan Kondisi Ruang Berasap Menggunakan Multi-Sensor Berbasis Arduino Uno," vol. 3, no. 4, pp. 3845–3854, 2019.
- [13] I. Sutoyo, "IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI DATA PESERTA DIDIK," *J. Pilar Nusa Mandiri*, vol. 14, no. 2, p. 217, 2018, doi: 10.33480/pilar.v14i2.926.
- [14] "365-690-1-PB.pdf," *JURTEKSI 9Jurnal Teknol. dan Sist. Inf.*, vol. V, no. 2, pp. 199–204, 2019.