

PENERAPAN ALGORITMA C4.5 UNTUK PREDIKSI PENERIMA BEASISWA SISWA BERPRESTASI DI MI AL-ISHLAH CIGANITRI

Risma Fardiani¹⁾, Fitriyani²⁾

^{1,2}Teknik Informatika, Universitas Adhirajasa Reswara Sanjaya, Antapani, Jl. Terusan Sekolah No.1-2, Cicaheum, Kec. Kiaracondong, Kota Bandung, Jawa Barat

Co Responden Email: rismafardiani27@gmail.com

Abstract

Article history

Received 27 Jul 2024

Revised 26 Sep 2024

Accepted 10 Oct 2024

Available online 31 Oct 2024

Keywords

Education,

Prediction,

Scholarship,

C4.5 Algorithm.

Education aims to produce quality and characterful people who have a big vision in realizing their dreams. Scholarships are the answer to realizing equal education. This scholarship is carried out agreeing to the sort of scholarship given. Each type of scholarship has different criteria or weights, so the school chooses based on the final grade and other achievements, if the grades and requirements are met they can get a scholarship. This research was conducted at MI Al-Ishlah Ciganitri, for the classification of students who received scholarships with the classification strategy utilizing the C4.5 calculation. By studying the MI Al-Ishlah Ciganitri scholarship data pattern, a model can be created from the data to predict students who will receive the next scholarship. The accuracy results obtained were 93.33% with the Good Classification category. The results of this research can be classified into two categories, namely receiving scholarships and not receiving scholarships. Thus, the C4.5 Algorithm can be a reference in classifying scholarship recipients properly so as to increase performance productivity.

Abstrak

Riwayat

Diterima 27 Jul 2024

Revisi 26 Sep 2024

Disetujui 10 Okt 2024

Terbit online 31 Okt 2024

Kata Kunci

Pendidikan,

Prediksi,

Beasiswa,

Algoritma C4.5.

Pendidikan bertujuan melahirkan manusia berkualitas dan berkarakter yang mempunyai visi besar dalam mewujudkan impiannya. Beasiswa jawaban untuk mewujudkan pemerataan pendidikan. Beasiswa ini dilakukan secara selektif sesuai jenis beasiswa yang diberikan. Setiap jenis beasiswa mempunyai kriteria atau bobot yang berbeda, maka pihak sekolah memilih berdasarkan nilai akhir dan prestasi lainnya, jika nilai dan persyaratan terpenuhi mereka bisa mendapatkan beasiswa. Penelitian ini dilakukan di MI Al-Ishlah Ciganitri, untuk klasifikasi siswa yang menerima beasiswa dengan teknik classification menggunakan algoritma C4.5. Dengan mempelajari pola data beasiswa MI Al-Ishlah Ciganitri, model dapat dibuat dari data tersebut untuk memprediksi siswa yang akan menerima beasiswa selanjutnya. Hasil akurasi yang diperoleh sebesar 93,33% dengan kategori Class Good Clasification. Hasil dari penelitian ini dapat mengklasifikasi kedalam dua kategori, yaitu menerima beasiswa dan tidak menerima beasiswa. Dengan demikian, Algoritma C4.5 dapat menjadi referensi dalam melakukan klasifikasi penerima beasiswa dengan baik sehingga meningkatkan produktifitas kinerja.

PENDAHULUAN

Pendidikan, yang sering disebut sebagai *long life education*, sangat penting bagi kehidupan manusia. Pendidikan secara umum mengacu pada proses kehidupan dalam mengembangkan diri setiap orang sehingga mereka dapat hidup dan menjalani kehidupan mereka dengan baik, sehingga menjadi seseorang yang terdidik. untuk menjadi orang yang bermanfaat bagi negara, bangsa, dan bangsa (Widiastuti et al., 2023).

Pendidikan adalah tabungan masa depan yang sangat penting bagi setiap orang. Dari zaman ke zaman, orang menyadari bahwa pendidikan dan guru sangat penting untuk setiap orang yang sukses. Pendidikan yang baik menghasilkan orang yang ramah dan sejahtera di masyarakat. Yang dimana, Pendidikan sangat penting untuk melahirkan manusia-manusia berbakat yang mampu bersaing dengan orang lain (Prasetyo et al., 2021).

Pendidikan berusaha untuk menghasilkan sumber daya manusia yang kaya dan berkualitas tinggi dan berkepribadian, berwawasan luas dan mampu mewujudkan impiannya. Untuk mencapai tujuan ini, pendidikan harus diberikan secara merata di seluruh masyarakat. Beasiswa merupakan sebuah solusi dalam mewujudkan pemerataan Pendidikan (Kristania, 2023).

Beasiswa merupakan suatu bentuk dukungan finansial kepada individu atau organisasi atas hasil penelitian yang dilakukan. Dukungan tersebut dapat berupa dukungan finansial untuk kegiatan pembelajaran atau dukungan finansial untuk pembelajaran (Khaliq et al., 2023).

Dengan adanya beasiswa, terbuka lebih banyak peluang bagi individu untuk mengakses pendidikan berkualitas, merintis jalan menuju peluang untuk mendapatkan pendidikan berkualitas tinggi, yang berdampak positif pada kemajuan masyarakat dan bangsa (Kalua et al., 2024).

Data mining adalah serangkaian proses yang secara manual mengekstraksi nilai tambahan yang berasal dari informasi yang sebelumnya tidak diketahui dari sekumpulan data (Sari et al., 2021).

Data mining merupakan bagian dari proses KDD (Penemuan Pengetahuan dalam Basis Data) dan mencakup beberapa Langkah-langkahnya seperti pemilihan data, preprocessing, transformasi, penambangan data, dan evaluasi hasil. KDD juga biasa dikenal dengan database (Zai, 2022).

Proses data mining terbagi menjadi beberapa tahapan, dan tahapan-tahapan proses data mining, diantaranya (Saputra & Sibarani, 2020) :

1. Pembersihan data (*Data cleaning*)
2. Integrasi data (*Data integration*)
3. Seleksi data (*Data selection*)
4. Transformasi data (*Data transformation*)
5. Proses mining Merupakan suatu proses
6. Evaluasi pola (*Pattern evaluation*)
7. Presentasi pengetahuan (*Knowledge presentation*)

Klasifikasi dalam data mining bertujuan untuk membuat pengelompokan berdasarkan atribut tertentu. Pengelompokan dilakukan berdasarkan tujuan kelas data serta nilai atribut masing-masing data. Dalam proses klasifikasi data diperlukan adanya tujuan kelas agar proses pengelompokannya mudah

tercapai. Dalam klasifikasi data mining, ada banyak jenis algoritma yang berbeda (Kushartanto & Aldisa, 2023).

Klasifikasi didefinisikan sebagai metode pembelajaran terawasi yang memerlukan data pelatihan berlabel untuk membangun model dari kelompok atau kelas data pengujian tertentu (Setio et al., 2020).

Algoritma C4.5 merupakan solusi permasalahan yang berkaitan dengan teknik klasifikasi. Untuk memeriksa data, khususnya untuk mencari hubungan antar kriteria atau atribut keputusan dari variabel input atau target (Sihombing et al., 2021).

Rapid Miner adalah platform perangkat lunak yang menyediakan berbagai alat untuk persiapan data, pemodelan, evaluasi, dan penerapan. RapidMiner dirancang untuk memungkinkan pengguna membuat dan menguji model yang berbeda dengan mudah tanpa pengalaman pemrograman apa pun. RapidMiner menyediakan antarmuka drag-and-drop yang memungkinkan pengguna membuat alur kerja untuk pemrosesan dan analisis data. Perangkat lunak ini juga mencakup banyak operator yang telah ditentukan sebelumnya, yang membentuk blok penyusun alur kerja, yang mencakup semua langkah dalam proses penambangan data, seperti pembersihan data, pemilihan fitur, dan pemodelan (Nahjan et al., 2023).

Beberapa contoh dijadikan acuan dalam penelitian yang sudah dilakukakn, penelitian pertama (Massulloh & Fitriyani, 2020), Implementasi algoritma c4.5 Untuk klasifikasi anak berkebutuhan Khusus di ibnu sina stimulasi center. Jika menggunakan aplikasi RapidMiner dan melakukan pengujian model menggunakan cross validation, maka akurasi pengujian menggunakan metode cross validation pada aplikasi RapidMiner adalah nilai akurasi sebesar 96,18%.

Penelitian lainnya dilakukan oleh (Yunita & Alaeysda, 2022), Penerapan Algoritma C4.5 Untuk Memprediksi Penerimaan Beasiswa di SD 4 Pelangian. Studi ini memberikan aturan untuk membantu sekolah menerapkan proses klasifikasi untuk menentukan kelayakan siswa untuk mendapatkan beasiswa. Metode algoritma C.45 dipilih karena membantu dalam klasifikasi dengan melakukan prediksi yang mencapai nilai akurasi yang relatif

tinggi. Berdasarkan hasil penelitian, akurasi adalah 85,36%.

Selanjutnya penelitian dilakukan oleh (Fitriyani & Arifin, 2020), Implementasi *Greedy Forward Selection* Untuk Prediksi Metode Penyakit Kutil Menggunakan *Decision Tree*. Dua dataset digunakan dalam penelitian ini, dataset cryotherapy dan dataset imunoterapi. Kedua dataset tersebut dapat diakses melalui website repositori UCI. Hasil penelitian pada dataset imunoterapi dengan menggunakan pohon keputusan tanpa pemilihan fitur menunjukkan nilai akurasi sebesar 78,89%, sedangkan ketika menggunakan pemilihan fitur “*Greedy Forward Selection*” nilai akurasi sebesar 82,22%. Hasil tersebut menunjukkan bahwa pemilihan fitur *Greedy Forward Selection* dapat meningkatkan performa model.

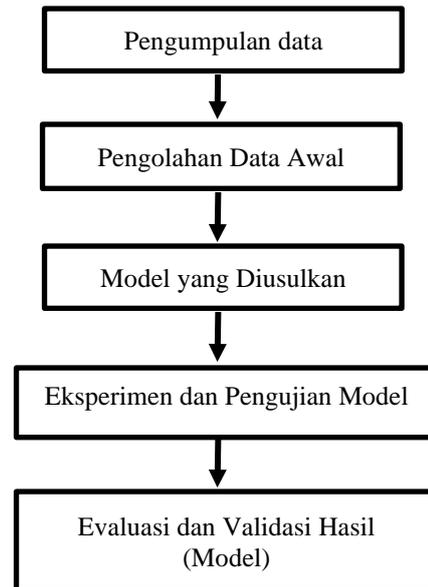
Selanjutnya adalah penelitian (Hendri, 2021), Implementasi Data Mining Dengan Metode C4.5 Untuk memprediksi penerima beasiswa. Hasil survei menunjukkan bahwa dari 100 mahasiswa, 9 orang (9%) dan 91 orang 91 % tidak menerima beasiswa.

Terakhir penelitian (Susanto & Mulyani, 2022), Analisa Algoritma C4.5 Terhadap Penentuan Rekomendasi Penerima Beasiswa. Hasil Penelitian Pohon keputusan dibuat melalui pengolahan data mining dari data pelatihan. Metode evaluasi yang dilakukan dalam pengujian ini didapatkan data nilai akurasi 90 %, hal ini bisa menjadi bukti bahwa algoritma tersebut cukup akurat untuk memberikan rekomendasi beasiswa.

METODE PENELITIAN

Tahapan Penelitian

Bagian ini menggunakan pendekatan kuantitatif sebagai metode penelitiannya. Penelitian langsung pada sekolah MI Al-Ishlah Ciganitri, terdapat beberapa tahapan penelitian yang umum dilakukan. Langkah-langkah yang dilakukan dalam penelitian ini ditunjukkan pada gambar di bawah ini.



Gambar 1. Tahapan Penelitian

Pengumpulan Data

Proses pengumpulan data menggunakan dataset yang diambil langsung dari MI Al-Ishlah Ciganitri data siswa kelas 4 dan kelas 5. Dalam pengumpulan data dilakukan dengan data primer. Data primer adalah data yang diperoleh langsung dari subjek penelitian. Dalam penelitian ini, proses pengumpulan data dilakukan melalui survei langsung ke sekolah.

Data yang dikumpulkan selama penelitian mencakup variabel yang diolah, memperoleh sebanyak 150 data dengan 15 atribut dalam Ms. Excel, terdapat 15 atribut yang diantaranya : Nama siswa, Nomor Induk, Jenis Kelamin, NISN, NIK, Tempat Lahir, Tanggal Lahir, No Hp, Nama orang tua, Pekerjaan orang tua, Penghasilan orang tua, Nilai akhir siswa (Nilai Raport), Prestasi Akademik, Kepribadian, dan Status Beasiswa.

Pengolahan Data Awal

Pada tahap ini, dibutuhkan pendalaman terhadap Dataset. Pendalaman ini dilakukan dengan pembersihan data (*cleaning data*) untuk menghilangkan beberapa data yang tidak dipakai, dan untuk memeriksa data yang perlu di perbaiki agar dapat dipastikan semua atribut dan kelas dalam dataset tersebut valid, sehingga data siap untuk dijadikan objek penelitian (Resmiati & Arifin, 2021). Dengan begitu, tujuan untuk mengetahui hasil klasifikasi terbaik. Transformasi data kemudian dilakukan dan diproses langsung

menggunakan template yang ditentukan di Microsoft Excel.

1. Pembersihan Data

Proses pembersihan data yang tidak sesuai dan memperbaiki kesalahan data pada saat penulisan. Total data awal mencakup 150 data dan 15 atribut. Pada proses pembersihan data ini terdapat 7 atribut yang tidak digunakan dan akan dihapus yaitu : Nomor Induk Kependudukan, NISN, NIK, Tempat Lahir, Tanggal Lahir, Nomor Handphone dan Nama Ayah Ibu.

2. Seleksi data

Seleksi data merupakan langkah penentuan atribut-atribut yang akan digunakan pada langkah data mining. Variabel independen yang akan digunakan antara lain nama siswa, jenis kelamin (JK), nilai rapor, prestasi akademik, kepribadian, pendapatan orang tua, dan pekerjaan orang tua. Sedangkan informasi status beasiswa dimasukkan ke dalam variabel dependen dengan hasil YA atau TIDAK.

3. Transformasi Data

Langkah ini melibatkan konversi data asli ke dalam format yang sesuai untuk memfasilitasi penambahan atau pemrosesan data. Atribut alfabet, seperti pekerjaan orang tua dan pendapatan orang tua, dalam data mining dikonversi terlebih dahulu ke dalam bentuk numerik atau numerik. Berikut tabel konversi data yang ditunjukkan pada Tabel 1.

Tabel 1.

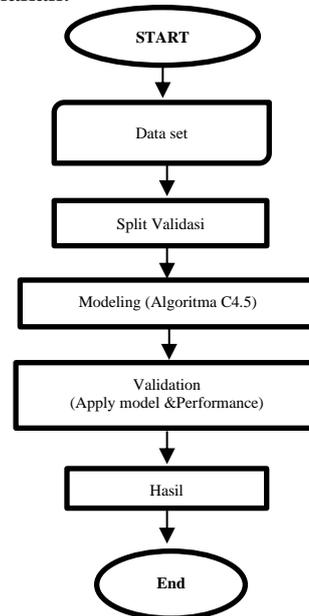
Data Transformasi

No	Atribut	Keterangan	Nilai Variabel
1	Pekerjaan Orangtua	karyawan Sawasta	1
		Wiraswasta	2
		PNS/TNI/POLRI	3
		Karyawan BUMN	4
		Pedagang	5
		Buruh	6
		Guru	7
2	Penghasilan Orangtua	1.000.000-1.999.999	1
		2.000.000-2.999.999	2
		3.000.000-4.999.000	3
		4.999.000	3

		>5.000.000	4
3	Nilai Rapor	86-90	A
		81-85	B
		76-80	C
		70-75	D

Model yang Diusulkan

Pada tahap ini, data dianalisis dan dikelompokkan berdasarkan variabel yang relevan. Setelah menganalisis data, diterapkan model yang sesuai dengan jenis data yang digunakan.



Gambar 2. Kerangka Penelitian

Dalam kerangka penelitian pada Gambar 2, dataset beasiswa terbagi dua bagian yaitu data pelatihan dan data pengujian yang dilakukan split validation secara otomatis pada *RapidMiner*. Kemudian Setelah data di validasi selanjutnya data akan diuji. Kemudian fitur dan record akan diproses menggunakan Modeling algoritma C4.5 untuk mendapatkan hasil dari kinerja model yang dibuat.

Eksperimen dan Pengujian Model

Langkah-langkah pengujian model pada penelitian ini adalah sebagai berikut.

- Langkah pertama menyiapkan data yang akan diuji.
- Langkah kedua adalah proses pembersihan data. Ini termasuk menghapus data yang tidak diperlukan atau memperbaiki data yang kosong.

3. Langkah ketiga dilakukan proses seleksi data pemilihan data yang diperlukan untuk digunakan.
4. Langkah keempat adalah transformasi atribut. Ini mengubah data alfabet menjadi bentuk numerik sehingga dapat diproses menggunakan data mining.
5. Langkah kelima adalah validasi terpisah. Ini membagi kumpulan data menjadi data pelatihan dan data pengujian, Ini dilakukan secara otomatis di Rapidminer.
6. Langkah keenam data tersebut akan diuji dengan menggunakan algoritma *Decision Tree C4.5*.
7. Langkah dilakukan validasi *performance* untuk mendapatkan performa pada model yang diterapkan.

Evaluasi dan Validasi Hasil (Model)

Saat ini data diolah menggunakan algoritma C4.5, langkah selanjutnya adalah mengevaluasi dan memvalidasi hasilnya. Evaluasi dilakukan dengan menggunakan tabel konfusi matriks yang digunakan untuk mengukur kinerja model sebagai kinerja nilai Accuracy yang dihasilkan. Validasi dilakukan dengan menggunakan validasi terpisah. Kumpulan data RapidMiner secara otomatis dibagi menjadi dua bagian data pelatihan dan data pengujian.

Saat menguji kinerja klasifikasi menggunakan algoritma C4.5, matriks konfusi sering digunakan. Confusion Matrix merupakan teknik yang digunakan untuk menghitung keakuratan data mining, bentuk dari metode ini adalah bentuk tabel (Ayudhitama & Pujianto, 2020).

Tabel 2.

Confusion Matrix

	Relevan	Non Relevan
Retrive	True Positive (TP)	False Positive (FP)
Non Retrive	False Negative (FN)	True Negative (TN)

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$PPV = \frac{TP}{TP+FP} \quad (5)$$

$$NPV = \frac{TN}{TN+FN} \quad (6)$$

Dengan Keterangan :

TP = True Positive

FP = False Positive

FN = False Negative

TN = True Negative

HASIL DAN PEMBAHASAN

Bagian ini menjelaskan hasilnya dari pengukuran kinerja pada model, untuk hasil dengan menggunakan dataset beasiswa dan akan dijelaskan cara pengolahan dataset tersebut dengan menggunakan *tools RapidMiner* dengan algoritma C4.5.

Pengumpulan data

Dalam penelitian ini, pengumpulan data untuk keperluan penelitian dilakukan oleh MI Al-Ishlah Ciganitri. Data yang dikumpulkan adalah tentang penerima beasiswa kelas 4 dan 5 sekolah dasar. Data yang diperoleh terdiri dari 15 atribut dan 150 record, dalam format Ms Excel sehingga memudahkan pengolahan data.

Pengolahan data

Pada tahap pengumpulan data, terkumpul 150 data dengan jumlah awal 15 atribut. Data tersebut kemudian diproses terlebih dahulu sebagai berikut:

- a. Pembersihan data (Cleaning data)

Proses pembersihan data yang buruk dan perbaikan kesalahan data pada saat ditulis. Pembersihan data ini mencakup total 150 data dan 15 atribut, serta atribut yang tidak terpakai yang dihapus adalah Nomor Induk, NISN, NIK, Tempat Lahir, Tanggal Lahir, Nomor HP, dan nama orang tua.

- b. Seleksi Atribut

Langkah pemilihan data menentukan atribut yang dipilih untuk digunakan dalam langkah penambahan data. Variabel independen yang digunakan yaitu nama siswa, jenis kelamin (JK), nilai rapor, prestasi akademik, kepribadian, pendapatan orang tua, dan pekerjaan orang tua. Sedangkan informasi status beasiswa dimasukkan ke dalam

variabel terikat dan outputnya adalah YA atau TIDAK.

c. Penyusunan dan Transformasi Data

Pada tahapan ini melakukan konversi atribut data dalam format alfabet ke format numerik atau numerik berdasarkan konversi data yang ditentukan dengan menggunakan perangkat lunak Microsoft Excel diantaranya seperti pekerjaan orang tua, penghasilan orang tua dan nilai raport. Sehingga dapat diperoleh hasil pada Gambar 3 berikut:

NO	STATUS	KEPERAWAN	PENGHASILAN	PEKERJAAN				
1	TIDAK	Andi Rizki	L	C	1	4	1	1
2	YA	Adhika Rizki	P	A	2	4	2	6
3	YA	Adhika Rizki	L	B	3	4	4	2
4	TIDAK	Alya Rizki	L	B	2	4	2	1
5	YA	Alya Rizki	P	B	1	5	1	6
6	TIDAK	Alya Rizki	L	C	2	4	4	4
7	TIDAK	Alya Rizki	P	C	2	3	1	1
8	YA	Alya Rizki	P	B	1	4	3	2
9	TIDAK	Andi Rizki	L	C	1	4	2	1
10	TIDAK	Andi Rizki	L	C	1	4	2	1
11	TIDAK	Alya Rizki	P	C	1	3	4	2
12	TIDAK	Alya Rizki	L	C	1	3	3	1
13	TIDAK	Alya Rizki	P	C	1	4	2	6
14	TIDAK	Alya Rizki	P	D	1	4	2	6
15	YA	Alya Rizki	L	B	3	4	1	1

Gambar 3.

Hasil penyusunan dan Transformasi Data

Perhitungan Menggunakan Algoritma C4.5

5 langkah berikut untuk membuat pohon keputusan:

1. Temukan nilai Entropy menggunakan kriteria-kriteria.

$$Entropy(S) = -\sum_{i=1}^n p_i * \log_2 p_i \quad (7)$$

2. Carilah nilai Gain setiap atribut.

$$Gain = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (8)$$

3. Bentuklah atribut-atribut tersebut sebagai dasar berdasarkan Gain tertinggi.
4. Membangun cabang berdasarkan setiap nilai.
5. Ulangi proses untuk setiap cabang.

Pengujian Algoritma C4.5 menggunakan Rapidminer

Langkah pertama dalam melakukan penelitian ini yaitu mempersiapkan data yang telah didapatkan dari MI Al-ishlah ciganitri. Data yang diambil yaitu dataset beasiswa

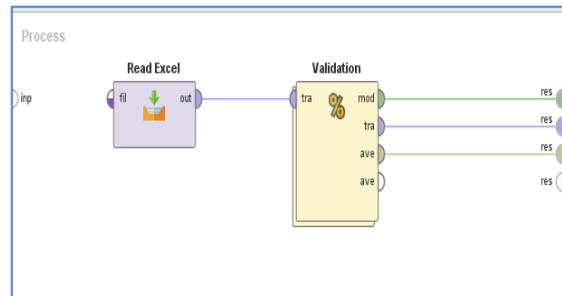
yang merupakan data siswa penerima beasiswa, dengan jumlah data 150 dan memiliki 8 atribut yang dapat dilihat pada Gambar 4 berikut.

Row No.	STATUS BE.	MAMA SISWA	JIK	Nilai Raport	PRESTASI.	KEBERSIKAPAN	PENGHASILAN.	PEKERJAAN.
1	TIDAK	Andi Rizki	L	C	1	4	1	1
2	YA	Adhika Rizki	P	A	2	4	2	6
3	YA	Adhika Rizki	L	B	3	4	4	2
4	TIDAK	Alya Rizki	L	B	2	4	2	1
5	YA	Alya Rizki	P	B	1	5	1	6
6	TIDAK	Alya Rizki	L	C	2	4	4	4
7	TIDAK	Alya Rizki	P	C	2	3	1	1
8	YA	Alya Rizki	P	B	1	4	3	2
9	TIDAK	Andi Rizki	L	C	1	4	2	1
10	TIDAK	Andi Rizki	L	C	1	4	2	1
11	TIDAK	Alya Rizki	P	C	1	3	4	2
12	TIDAK	Alya Rizki	L	C	1	3	3	1
13	TIDAK	Alya Rizki	P	C	1	4	2	6
14	TIDAK	Alya Rizki	P	D	1	4	2	6
15	YA	Alya Rizki	L	B	3	4	1	1

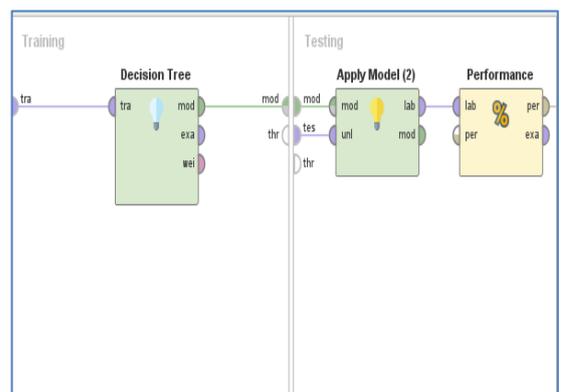
Gambar 4.
Data Pengujian

Setelah data siap, langkah selanjutnya adalah mengujinya menggunakan aplikasi Rapidminer. Bagian ini berisi tampilan sistem dengan algoritma C4.5 menggunakan tool RapidMiner sebagai berikut:

1. Rancangan Model dengan menggunakan Algoritma C4.5 :



Gambar 5. Rancangan Model Algoritma C4.5 di RapidMiner



Gambar 6. Rancangan Model di dalam Operator X-Validation

2. Hasil Algoritma C4.5

Setelah membuat model algoritma C4.5 di tool RapidMiner, kami mendapatkan hasil klasifikasi calon penerima beasiswa di MI Al-Ishlah Ciganitri. Hasil aturan ditunjukkan di bagian bawah Gambar 7 :

Tree	
PRESTASI AKADEMIK > 2.500:	YA (TIDAK=0, YA=11)
PRESTASI AKADEMIK ≤ 2.500	
Nilai Raport = A:	YA (TIDAK=0, YA=3)
Nilai Raport = B	
PRESTASI AKADEMIK > 1.500:	YA (TIDAK=1, YA=8)
PRESTASI AKADEMIK ≤ 1.500:	TIDAK (TIDAK=8, YA=3)
Nilai Raport = C:	TIDAK (TIDAK=77, YA=3)
Nilai Raport = D:	TIDAK (TIDAK=36, YA=0)

Gambar 7. Hasil Aturan atau Rule

Tampilan teks hasil pohon keputusan pada Gambar 7 menunjukkan hasil rule model dalam format teks berdasarkan Gambar 7 dengan uraian sebagai berikut: Jika Prestasi Akademik > 2.500, maka hasilnya = YA (Tidak = 0, YA=11).

- Jika Prestasi Akademik ≤ 2.500, Nilai Raport = A, maka hasilnya = YA (Tidak=0, Ya=3).
- Jika Nilai Raport = B, Prestasi Akademik > 1.500, maka hasilnya = YA (Tidak=1, Ya=8).
- Jika Nilai Raport = B, Prestasi Akademik ≤ 1.500, maka hasilnya = TIDAK (Tidak=8, Ya=3).
- Jika Prestasi Akademik ≤ 2.500, Nilai Raport = C, maka hasilnya = TIDAK (Tidak=77, Ya=3).
- Jika Prestasi Akademik ≤ 2.500, Nilai Raport = D, maka hasilnya = TIDAK (Tidak=36, Ya=0).

Evaluasi dan Validasi Model

Matriks konfusi digunakan dalam penelitian ini memverifikasi validitas kinerja nilai accuracy, dalam dilihat nilai akurasi sebagai berikut ini :

Tabel 3. Performance Vector

	True TIDAK	True YA	Class Precision
Pred TIDAK	37	3	92.50%

Pred	0	5	100.00%
YA			
Class	100.00%	62.50%	
recall			

Tabel di atas menunjukkan hasil vektor kinerja, sehingga dengan menggunakan algoritma C4.5 kita dapat menutupnya dengan 150 data. Terlihat prediksi no pada label no adalah 37 dan label yes adalah 3. Oleh karena itu, nilai akurasi kelas tersebut adalah 92,50%. Terdapat 0 prediksi ya berlabel “tidak” dan 5 prediksi ya, sehingga akurasi kelasnya 100%. Berikut ini nilai akurasi yang diperoleh, yaitu :

$$Accuracy = \frac{37 + 5}{37 + 3 + 0 + 5} = 0.9333$$

$$= 93.33\%$$

$$Precision = \frac{37}{37 + 3} = 0.925 = 92.5\%$$

$$Recall = \frac{37}{37 + 0} = 1 = 100\%$$

$$Specificity = \frac{5}{5 + 3} = 0.625 = 62.5\%$$

$$PPV = \frac{37}{37 + 3} = 0.925 = 92,5\%$$

$$NPV = \frac{5}{5 + 0} = 1 = 100\%$$

Dapat disimpulkan perolehan akurasi dari Algoritma C4.5 pada tools RapidMiner sebesar 93,33%.

KESIMPULAN

Hasil dari penelitian yang dilakukan menyimpulkan data mining dapat diklasifikasikan menggunakan algoritma C4.5 calon penerima Beasiswa di MI Al-Ishlah Ciganitri, dapat disimpulkan sebagai berikut :

- Hasil akurasi yang diperoleh penelitian ini sebesar 93,33%. Berdasarkan hasil yang didapatkan, penelitian dengan kasus klasifikasi calon penerima Beasiswa di MI Al-Ishlah Ciganitri dengan menggunakan Algoritma C4.5

mendapatkan akurasi yang baik dengan *Class Good Clasification*.

2. Hasil dari analisis dan klasifikasi menggunakan Algoritma C4.5 diperoleh 6 rule model dan dapat disimpulkan dalam 150 data, sebanyak 122 siswa tidak menerima beasiswa dan 28 siswa menerima beasiswa. Hasil dari penelitian ini dapat mengklasifikasi kedalam dua kategori, yaitu menerima beasiswa dan tidak menerima beasiswa. Dengan demikian, Algoritma C4.5 dapat menjadi referensi dalam melakukan klasifikasi penerima beasiswa dengan baik sehingga meningkatkan produktifitas kinerja.

REFERENSI

- Abidin, Z., Nurhana, E., Permata, & Ulum, F. (2023). Eka Nurhana 2) , Permata 3) , Faruq Ulum 4). *Analisis Perbandingan Algoritma Decission Tree C4.5 dan C5.0 pada Data Karyawan Berpotensi Promosi Jabatan*. Jurnal TEKNOINFO, Vol. 17, Issue 2.
- Adila, F., & Fitriyani, F. (2024). *Pengembangan Sistem Informasi Kasih Berbasis Website pada Biya Salon Muslimah*. JIKA (Jurnal Informatika), 8(2), 171–179. <https://doi.org/10.31000/jika.v8i2.9990>
- Ayudhitama, A. P., & Pujianto, U. (2020.) *Analisa 4 Algoritma Dalam Klasifikasi Penyakit Liver Menggunakan RAPIDMINER*. JIP (Jurnal Informatika Polinema), 6(2).
- Fitriyani, F., & Arifin, T. (2020a). Implementasi Greedy Forward Selection Untuk Prediksi Metode Penyakit Kutil Menggunakan Decossion Tree. *Sains Dan Teknologi, Jurnal*. Vol 9 (1).
- Fitriyani, F., & Arifin, T. (2020b). *Implementasi Greedy Forward Selection Untuk Prediksi Metode Penyakit Kutil Menggunakan Decision Tree*. JST (Jurnal Sains Dan Teknologi), Vol 9(1), 76–85. <https://doi.org/10.23887/jstundiksha.v9i1.24896>
- Haryati, N. T., Negara, E. S., & Kurniawan, T. B. (2023). *Klasifikasi Pemberian Beasiswa Berprestasi Menggunakan Perbandingan Tiga Algoritma*, Vol. 17, Issue 1.
- Hendri, H. (2021). Implementasi Data Mining Dengan Metode C4.5 Untuk Prediksi Mahasiswa Penerima Beasiswa. *Indonesian Journal of Computer Science Attribution-ShareAlike*, 4(2), 2021–2312.
- Kalua, A. L., Mantiri, R., Rumondor, C., & Mogogibung, E. (2024). Sistem Informasi Pendaftaran Beasiswa dan Jadwal Legalisir Berbasis Website Responsive. *Journal of Information Technology, Software Engineering and Computer Science (ITSECS)*, 2(2). <https://doi.org/10.58602/itsecs.v2i2.108>
- Khaliq, N. A., Josi, A., & Fujiyanti, L. (2023). *Sistem Informasi Pendukung Keputusan Seleksi Beasiswa Menggunakan Metode SAW*. JSITIK: Jurnal Sistem Informasi Dan Teknologi Informasi Komputer, 1(2), 94–108. <https://doi.org/10.53624/jsitik.v1i2.162>
- Kristania, Y. M. (2023). Penerapan Combined Compromise Solution Method Dalam Penentuan Penerima Beasiswa. *CHAIN: Journal of Computer Technology, Computer Engineering*, 1(2). <https://doi.org/10.58602/chain.v1i2.27>
- Kushartanto, A. I., & Aldisa, R. T. (2023). Data Mining Perbandingan Algoritma K-Nearest Neighbor dan Naïve Bayes dalam Prediksi Penerimaan Beasiswa. *Journal of Computer System and Informatics (JoSYC)*, 5(1), 196–207. <https://doi.org/10.47065/josyc.v5i1.4566>
- Massulloh, I., & Fitriyani, F. (2020). *Implementasi Algoritma C4.5 untuk Klarifikasi Anak Berkebutuhan Khusus di Ibnu Sina Stimulasi Center*. Sistem Informasi (POTENSI), 1(1), 136–144
- Nahjan, M. R., Heryana, N., & Voutama, A. (2023). *Implementasi RAPIDMINER Dengan Metode Clustering K-MEANS untuk Analisa Penjualan pada Toko OJ CELL*. In Jurnal Mahasiswa Teknik Informatika. Vol 7, Issue 1.

- Nuraeni, O., & Fitriyani, F. (2023b). *Sistem Pakar Diagnosa Kondisi Gigi Tiruan Menggunakan Metode Naive Bayes Classifier*. *JIKA (Jurnal Informatika)*, 7(1), 79–88. <https://doi.org/10.31000/jika.v7i1.7257>
- Prasetyo, I. A., Mustofa, H., Supriadi, S., Prasetyo, R. H., Yasri, M. L., & Fitriyani, F. (2021). *Metode Decision Tree Dalam Pemilihan Gaya Belajar Pada Siswa Sekolah Dasar*. In *dkk. Saintek* (Vol. 5, Issue 1).
- Resmiati, R., & Arifin, T. (2021). *SISTEMASI: Jurnal Sistem Informasi Klasifikasi Pasien Kanker Payudara Menggunakan Metode Support Vector Machine dengan Backward Elimination*. <http://sistemasi.ftik.unisi.ac.id>
- Safitri, T. N. (2020). *Penerapan Algoritma Naive Bayes untuk Penentuan Calon Penerimaan Beasiswa pas SD Negeri 6 Ketapang*. *JUISI (Jurnal Informatika dan Sistem Informasi)*, Vol 6(01).
- Saputra, R., & Sibarani, A. J. P. (2020). *Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat*. *JUISI (Jurnal Informatika dan Sistem Informasi)*, Vol 7(2)
- Sari, M., Windarto, A. P., & Okprana, H. (2021). *BEES: Bulletin of Electrical and Electronics Engineering Penerapan Data Mining Klasifikasi C4.5 Pada Penerima Beasiswa di SMK Swasta Anak Bangsa*. *Media Online*, 1(3), 115–121.
- Setio, P. B. N., Saputro, D. R. S., & Winarno, B. (2020). *PRISMA, Prosiding Seminar Nasional Matematika Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5*. 3, 64–71. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Sihombing, I. A., Hartama, D., Parlina, I., Gunawan, I., Kirana, I. O., Tunas, S., Pematangsiantar, B., Sudirman Blok, J. J., No, A., Pematangsiantar, I., & Tunas, A. (2021). *Analisis Keberhasilan Pembelajaran Daring pada Masa Pandemi Covid-19 menggunakan Algoritma C4.5 dan Naive Bayes*. *JUKI (Jurnal Komputer dan Informatika)*, Vol 3(2).
- Susanto, W., & Mulyani, A. (2022). *Analisa Algoritma C4.5 Terhadap Penentuan Rekomendasi Penerima Beasiswa*. <https://journal.mediapublikasi.id/index.php/oktal>
- Topadang, A., Irwansyah, I., & Safruddin, S. (2020). *Penerapan Metode Simple Additive Weighting(SAW) Untuk Pemilihan Beasiswa Kurang Mampu pada Sekolah Dasa Katolik Hati Kudus Samarinda*. *Just TI (Jurnal Sains Terapan Teknologi Informasi)*, 12(2), 66. <https://doi.org/10.46964/justti.v12i2.377>
- Widiastuti, N. A., Azhar, M., & Mulyo, H. (2023). *Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Jurusan pada Peserta Didik Baru*. *Jurnal SIMETRIS*, 14(2).
- Yunita, S., & Alaeyda, V. N. (2022). *Penerapan Algoritma C4.5 Untuk Prediksi Penerimaan Beasiswa di SD 4 Pelangsan*. *Jurnal ICIT Universitas Raharja*, Vol. 8(2).
- Zai, C. (2022). *Implementasi Data Mining Sebagai Pengolahan Data*. *Jurnal Portal Data*, Vol. 2, Issue 3.