



Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Penentuan Penerima Beasiswa

**Mahardhika Putra*, Muhammad Vajri Al Kautsar, Agung Pratama, Fuad
Naufal Ryan, Adi Suhendi, Annida Purnamawati**

Universitas Bina Sarana Informatika, Indonesia

Email: mahardhikaputra168@gmail.com* , vajrialkautsar21@gmail.com,
pratamaagung542@gmail.com, naufal.ryan@gmail.com,
pratamaagung542@gmail.com, annida.npr@bsi.ac.id

Abstrak:

suatu lembaga pendidikan terutama universitas ataupun sekolah tinggi, Mahasiswa yang memiliki IPK yang bagus dan mengikuti setiap keorganisasian ataupun kegiatan kampus lainnya sering kali ditawarkan beasiswa terutama untuk mahasiswa dari keluarga yang kurang mampu dalam urusan keuangan. Dimana IPK menjadi kriteria utama dalam pemberian beasiswa tersebut. Algoritma C4.5 dapat membantu menentukan tingkat keakurasian mahasiswa yang lebih layak mendapatkan ataupun tidak layak untuk menerima beasiswa. Algoritma C4.5 mempunyai kelebihan utama yang dapat menghasilkan suatu model berupa decision tree yang mudah diimplementasikan untuk menentukan suatu keputusan dan juga untuk mengetahui tingkat keakurasian dalam suatu data. Dalam melakukan analisa untuk menentukan penerima beasiswa. Maka dari itu terdapat 6 atribut yang digunakan yaitu Ikut Organisasi, Ikut Ukm, IPK, Pekerjaan Orang Tua, Tanggungan dan Status beasiswa yang menjadi label. Hasil implementasi dilakukan menggunakan software Data Mining Rapid Miner, Dimana hasil implementasi didapatkan tingkat keakurasian data sebesar 80,10%.

Kata Kunci: Algoritma C4.5, Beasiswa, Mahasiswa

Abstract:

In educational institutions, especially universities or colleges, students with high GPAs and participation in organizations or other campus activities are often offered scholarships, especially for students from financially disadvantaged families. GPA is the primary criterion for awarding scholarships. The C4.5 algorithm can help determine the level of accuracy of students who are more deserving or not worthy of receiving a scholarship. The C4.5 algorithm has the main advantage of being able to produce a model in the form of a decision tree that is easy to implement to determine a decision and also to determine the level of accuracy in a data. In conducting analysis to determine scholarship recipients, there are 6 attributes used: Joining an Organization, Joining a UKM, GPA, Parent's Occupation, Dependents and Scholarship Status which serve as labels. The implementation results were carried out using Rapid Miner Data Mining software, where the implementation results obtained a data accuracy level of 80.10%.

Keywords: C4.5 Algorithm, Scholarship, Collage Student

PENDAHULUAN

Pendidikan di tingkat universitas adalah salah satu fondasi utama dalam kemajuan sebuah negara (Khasanah et al., 2025). Banyak mahasiswa yang mengalami masalah dalam menyelesaikan studi tepat waktu, bahkan ada yang sampai memutuskan untuk berhenti belajar. Hal ini disebabkan oleh beberapa faktor seperti prestasi akademik yang kurang memuaskan, keadaan sosial ekonomi yang kurang mendukung dan banyaknya faktor lain (Tabuni et al., 2024). Mendapatkan beasiswa merupakan harapan bagi setiap mahasiswa sedang menjalani pendidikan. Beasiswa merupakan suatu dukungan finansial yang tidak berasal dari dana pribadi atau orang tua melainkan beasiswa disediakan oleh pemerintah, perusahaan swasta, atau lembaga pendidikan dan penelitian (Apdian et al., 2024). Tidak semua bantuan pendidikan yang ada bisa diakses oleh semua mahasiswa. Untuk itu, institusi pendidikan selalu melakukan proses pemilihan atau penyeleksian di antara para mahasiswa yang berminat untuk menerima beasiswa yang disediakan (Halifia Dalam Jumadi, 2021).

Data mining merupakan cara untuk mengidentifikasi pola atau informasi yang bermanfaat dari sekumpulan data yang besar (Syahriani, 2022). Dalam dunia pendidikan, data mining biasanya dikenal sebagai educational data mining (EDM) Salah satu algoritma yang umum diterapkan dalam klasifikasi di data mining adalah C4. 5, yang merupakan algoritma pohon keputusan dan dirancang oleh Ross Quinlan (Mantopani et al., 2024). Metode klasifikasi dapat digunakan juga dalam proses seleksi calon mahasiswa penerima beasiswa (Khotimah, 2021). C4.5 dapat mengelola atribut yang berupa angka dan kategori, mengatasi data yang hilang, serta menciptakan model pohon yang mudah dipahami . Dalam penelitian ini digunakan dataset sekunder dari website kaggle untuk mengklasifikasikan dan mengetahui tingkat ke akurasian data menggunakan metode C4.5 untuk data mahasiswa sebanyak 1042 data. Data tersebut di olah menggunakan perangkat lunak Rapid Miner yang merupakan tools data mining berbasis GUI yang sering digunakan dalam studi klasifikasi serta ramalan (Budiarto et al., 2022; Hasmin & Aisa, 2019).

Penelitian ini menawarkan kebaruan dalam hal penerapan algoritma C4.5 pada sistem penentuan penerima beasiswa di lembaga pendidikan, dengan menggunakan atribut-atribut yang relevan seperti IPK, kegiatan organisasi, pekerjaan orang tua, dan jumlah tanggungan. Kebaruan lainnya adalah penerapan sistem ini dengan menggunakan perangkat lunak RapidMiner, yang memungkinkan analisis yang lebih efisien dan akurat. Penelitian ini juga memberikan wawasan baru mengenai pengolahan dan klasifikasi data mahasiswa dalam konteks beasiswa, yang dapat diadaptasi oleh berbagai lembaga pendidikan untuk meningkatkan proses seleksi beasiswa mereka.

Penelitian ini bertujuan untuk mengimplementasikan algoritma C4.5 dalam sistem penentuan penerima beasiswa di lembaga pendidikan, dengan memanfaatkan data mahasiswa yang meliputi berbagai atribut seperti IPK, kegiatan organisasi, pekerjaan orang tua, dan jumlah tanggungan. Penelitian ini juga bertujuan untuk menganalisis tingkat akurasi dalam menentukan kelayakan penerima beasiswa berdasarkan data yang ada, serta mengevaluasi efektivitas penggunaan algoritma C4.5 dalam meningkatkan akurasi keputusan pemberian beasiswa (Suweleh et al., 2020). Selain itu, penelitian ini bertujuan untuk memberikan pemahaman tentang penerapan metode data mining dalam

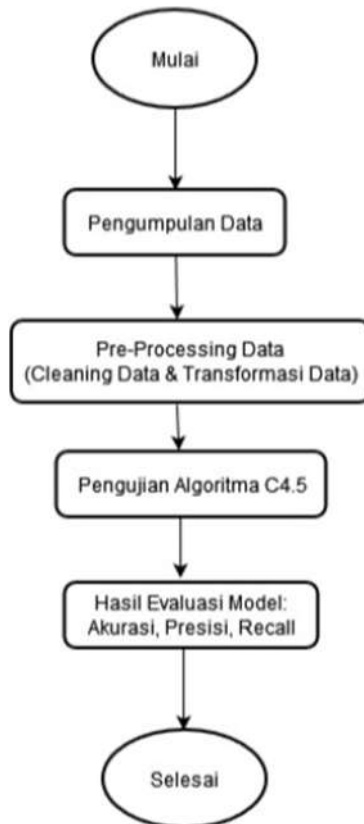
konteks pendidikan dan membantu lembaga pendidikan dalam membuat keputusan yang lebih objektif dan berbasis data (Rismayanti, 2016; Sandiva et al., 2024).

Manfaat dari penelitian ini adalah untuk memberikan kontribusi dalam pengembangan sistem informasi yang dapat mempermudah lembaga pendidikan dalam proses seleksi penerima beasiswa, yang lebih adil dan objektif, berdasarkan data yang relevan. Secara praktis, hasil penelitian ini dapat menjadi referensi bagi lembaga pendidikan dalam mengadopsi teknologi data mining untuk mempercepat proses seleksi dan meningkatkan akurasi keputusan. Secara teoritis, penelitian ini akan memperkaya studi mengenai penerapan algoritma C4.5 dalam data mining, khususnya dalam konteks pendidikan, dan memberikan wawasan baru tentang pemanfaatan algoritma pohon keputusan dalam menentukan kelayakan penerima beasiswa.

METODE

Pengumpulan Data

Tahapan pertama untuk penelitian ini ialah pengambilan dataset sekunder yang berasal dari kaggle. Dataset yang digunakan terdiri dari 1042 entri. Setiap entri mempresentasikan satu individu dari mahasiswa dengan atribut seperti: Nama lengkap, Jenis kelamin, Ikut organisasi, ikut ukm, IPK, Pekerjaan orang tua, Penghasilan dan Status beasiswa. Atribut Status beasiswa (Terima atau Tidak) ditetapkan sebagai label. Penelitian ini menggunakan metode C4.5 model classification decision tree yang diharapkan mampu untuk membantu dalam mengklasifikasikan mahasiswa penerima beasiswa dengan akurat dan efisien.



Gambar 1. Flowchart algoritma C4.5

Gambar 1. Menggambarkan urutan – urutan proses kegiatan dimulai dari pengumpulan data hingga terakhir proses evaluasi hasil model. Penjelasan pada gambar diatas sangat penting untuk memahami proses langkah-langkah yang dilakukan supaya proses penelitian berjalan sistematis.

Pre-Processing Data

Pre-processing data dalam penelitian adalah langkah dimana data disiapkan terlebih dahulu sebelum diproses dan dimodelkan. Oleh karena itu, diperlukan beberapa tahap agar data dapat diproses ke langkah berikutnya dengan tujuan agar data menjadi lebih teratur (Muriyatmoko et al, dalam Sugiyono, 2024) Adapun tahap-tahap tersebut :

1. Data Cleaning

Data yang sudah terkumpul kadang-kadang belum dalam keadaan siap untuk digunakan. Untuk itu, perlu dilakukan langkah-langkah pembersihan data agar integritas dan kualitas data terjaga sebelum dilakukan analisis (Santoso & Priyadi, 2024).

2. Transformasi Data

Transformasi data dilaksanakan terutama untuk mengubah tingkat pengukuran data yang ada ke format baru supaya data dapat memenuhi syarat yang mendasari analisis varians dan selaras dengan perhitungan algoritma C4.5 (Dandi et al., 2024) .

3. Pemisahan data

Spit data merupakan metode yang diterapkan untuk membagi dataset dan ini adalah salah satu diantara faktor yang penting terhadap seberapa efektif model klasifikasi dalam algoritma pembelajaran mesin (Oktafiani et al., 2023).

4. Membangun model

Proses membangun model dilaksanakan dengan memilih algoritma yang terhubung dengan tujuan dari penelitian menyesuaikan variabel input dan output yang digunakan. Tahap tersebut terlibat dalam menentukan paramater dan struktur model yang akan dipelajari pada saat pelatihan

5. Pelatihan Model

Tahapan pelatihan model bertujuan untuk data pelatihan agar sistem dapat mengenali pola dan suatu hubungan antara variabel. Model menyesuaikan dengan parameter internal untuk mengurangi kesalahan dan meningkatkan nilai akurasi dari klasifikasi

6. Tahap Prediksi

Tahap Prediksi merupakan cara untuk menilai suatu keadaan secara teratur dengan menggunakan data yang tersedia dari waktu lalu dan masa kini. Hal ini dapat terkait dengan kejadian yang akan datang, sehingga kemungkinan terjadinya kesalahan (selisih antara realitas dan ramalan) dapat dikurangi (Akbar et al., 2022) .

7. Evaluasi Model

Evaluasi model dilaksanakan untuk mengukur perform model dengan menggunakan metrik antara lain akurasi, presisi dan recall. Dengan tahapan ini peneliti mengenali baik model tentang pola dan menentukan penyempurnaan atau pelatihan kembali

Klasifikasi

Klasifikasi ialah satu metode pada data mining. Metode klasifikasi adalah pendekatan untuk memproyeksikan data, menciptakan suatu asumsi nilai dari sebuah data dengan hasil yang telah diberitahu berasal dari data lainnya. Sasaran dari pendekatan tersebut adalah untuk memperkirakan nilai dari satu variabel yang tidak dikenal berdasarkan variabel lain yang sudah ada (Nugraha et al., 2022).

Rapid Miner

Rapid Miner adalah aplikasi yang diciptakan oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Ralf Klinkenberg dari rapid-i. com, yang memiliki antarmuka grafis (GUI) yang membuat pengguna lebih mudah dalam mengoperasikannya. Software ini bersifat sumber terbuka dan dikembangkan menggunakan bahasa pemrograman Java dengan lisensi GNU Public License, serta kompatibel dengan berbagai sistem operasi. Dalam menggunakan Rapid Miner, pengguna tidak perlu memiliki keahlian pemrograman khusus karena semua alat yang diperlukan telah disediakan.

Rapid Miner difokuskan pada aktivitas pengolahan data. Terdapat berbagai model yang ditawarkan, seperti Model Bayesian, Modelling, Tree Induction, Neural Network, dan lain- lain. Rapid Miner juga menyajikan berbagai metode, termasuk klasifikasi, kluster, asosiasi, dan masih banyak yang lainnya. Jika ada model atau algoritma tertentu yang tidak tersedia di Weka, pengguna diizinkan untuk menambahkan modul lain, karena

Weka bersifat sumber terbuka sehingga siapa pun dapat berkontribusi dalam pengembangan aplikasi ini. (Budiarto, 2022).

Decision Tree

Decision Tree merupakan suatu metode yang umum diterapkan untuk membuat keputusan. Metode ini berusaha mencari penyelesaian atas suatu permasalahan dengan memakai kriteria sebagai titik- titik yang saling terhubung, membentuk sebuah struktur yang mirip seperti pohon. Model Decision Tree digunakan untuk meramalkan keputusan dengan memanfaatkan bentuk hierarkis atau pohon. Setiap pohon memiliki bagian cabang, di mana masing-masing cabang mewakili atribut yang harus dipenuhi agar bisa bergerak ke cabang selanjutnya hingga sampai ke daun. Dalam Decision Tree, data disajikan dalam bentuk tabel yang terdiri dari atribut dan entri. (Pamuji & Ramadhan, 2021).

HASIL DAN PEMBAHASAN

Data

Dataset diambil dari www.kaggle.com dalam bentuk excel csv dan diubah ke xlsx yang bernama DATASET 1000 beasiswa dengan jumlah entry sebanyak 1042 terdiri atas 8 atribut.

Selection Data

Data yang akan digunakan dalam penelitian ini merupakan data sekunder yang diambil dari www.kaggle.com kriteria-kriteria yang digunakan untuk perhitungan, yaitu pada mahasiswa calon penerima beasiswa yang akan digunakan untuk perhitungan alternatif penentuan mahasiswa penerima beasiswa. Metode yang akan digunakan untuk proses permasalahan diatas adalah metode klasifikasi dengan algoritma yang digunakan adalah algoritma C4.5 dengan kriteria-kriteria sebagai berikut :

Tabel 1. Atribut Terpilih

No	Nama Atribut	Keterangan
1	Jenis Kelamin	Atribut
2	Ikut Organisasi	Atribut
3	Ikut UKM	Atribut
4	IPK	Atribut
5	Pekerjaan Orang Tua	Atribut
6	Penghasilan	Atribut
7	Tanggungan	Atribut
8	Status Beasiswa	Label

Pada tabel 1. Attribute terpilih adalah atribut-atribut yang digunakan untuk nantinya ditransformasikan kedalam rapid miner.

Transformasi Data

Transformasi data ialah tahapan yang memiliki tujuan untuk menyesuaikan format data supaya dapat diproses dalam kegiatan data mining. Pada tahapan ini proses

transformasi dilakukan menggunakan software Rapid Miner dengan mengubah tipe data pada atribut yang digunakan sebagaimana ditunjukkan pada gambar dibawah ini

Import Data - Format your columns:

Format your columns.

☐ Replace errors with missing values ⓘ

	Jenis Kelamin ◯ binominal	Ikut Organis... ◯ binominal	Ikut UKM ◯ binominal	IPK ◯ real	Pekerjaan O... ◯ polynominal	Penghasilan ◯ polynominal	Tanggungan ◯ integer	Status Beasi... binominal label
1	L	Ikut	Ikut	3.570	Wiraswasta	Sedang	4	Terima
2	L	Tidak	Ikut	2.950	Buruh	Sedang	2	Tidak
3	P	Tidak	Ikut	3.670	Petani	Sedang	4	Terima
4	L	Tidak	Ikut	3.190	Wiraswasta	Tinggi	2	Tidak
5	P	Tidak	Ikut	3.190	Wiraswasta	Sedang	2	Tidak
6	P	Tidak	Ikut	3.780	Wiraswasta	Rendah	3	Terima
7	L	Tidak	Ikut	3.330	Petani	Sedang	3	Tidak
8	L	Tidak	Ikut	3.290	Petani	Rendah	1	Tidak
9	P	Tidak	Ikut	3.480	Wiraswasta	Sedang	3	Tidak
10	P	Tidak	Ikut	3.380	Pedagang	Sedang	1	Tidak
11	P	Tidak	Ikut	3.520	Petani	Rendah	2	Tidak
12	P	Tidak	Ikut	2.810	Pedagang	Rendah	2	Tidak
13	P	Tidak	Ikut	3.700	Karyawan Swasta	Sedang	1	Tidak
14	L	Tidak	Ikut	3.710	Petani	Rendah	2	Tidak
15	L	Tidak	Ikut	3.380	Wiraswasta	Tinggi	1	Tidak
16	P	Tidak	Ikut	3.570	Petani	Rendah	2	Tidak
17	P	Tidak	Ikut	3.290	Pedagang	Rendah	2	Tidak
18	P	Tidak	Ikut	3.000	Wiraswasta	Rendah	1	Tidak

3.000

no problems

Previous Next Cancel

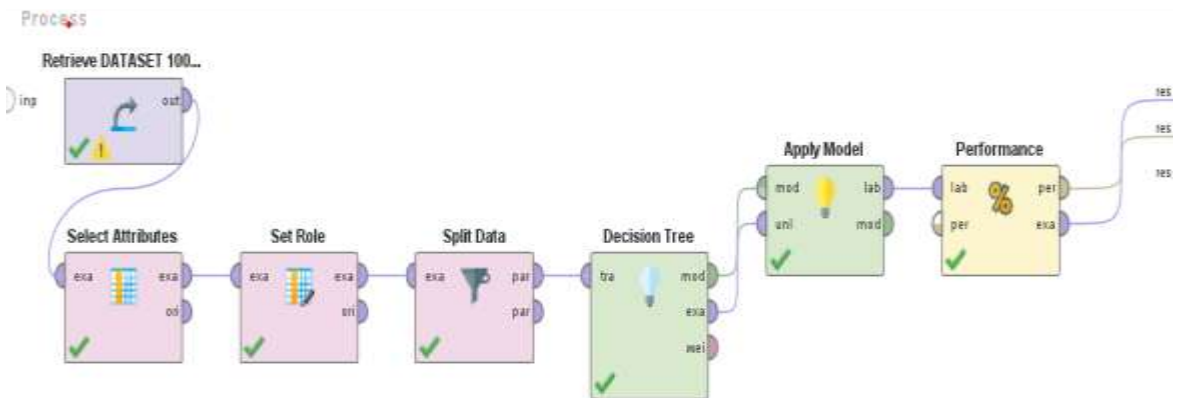
Gambar 2. Transformasi dataset 1000 BEASISWA

Pada gambar 2. Merupakan atribut – atribut yang sudah ditransformasikan kedalam rapid miner terdapat 8 atribut yang sudah dipilih dari dataset yang diambil sebelumnya atribut tersebut terdiri dari : jenis kelamin, Ikut Organisasi, Ikut UKM, IPK, Pekerjaan Orang Tua, Penghasilan , Tanggungan dan Status beasiswa yang diberi nilai label.

Data Mining

Tahapan modeling penelitian ini menggunakan metode Algoritma C4.5 Decision Tree untuk melakukan pengklasifikasi. Implementasi menggunakan Rapid Miner dengan menggunakan beberapa operator yaitu Retrieve dataset untuk melakukan import data ke repository, Select attribut

, Set role untuk memberikan label disini adalah Status Beasiswa, Split data sebagai pemisah .



Gambar 3. Proses langkah *Data Mining*

Pada gambar 3. merupakan proses training digunakan modeling tree Decision Tree dan Apply Model untuk menguji data dan Performance untuk mendapatkan hasil evaluasi dari kinerja algoritma

Decision Tree



Gambar 4. Tampilan *Decision Tree*

Pada gambar 4. Merupakan hasil dari penggunaan modeling decision tree.

Tree

```

Ikut UKM = Ikut
|   Tanggungan > 4.500
|   |   IPK > 3.055
|   |   |   Ikut Organisasi = Ikut
|   |   |   |   IPK > 3.405: Terima {Terima=4, Tidak=0}
|   |   |   |   IPK ≤ 3.405
|   |   |   |   |   Penghasilan = Rendah
|   |   |   |   |   |   Jenis Kelamin = L: Tidak {Terima=2, Tidak=3}
|   |   |   |   |   |   Jenis Kelamin = P: Terima {Terima=3, Tidak=1}
|   |   |   |   |   |   Penghasilan = Sedang
|   |   |   |   |   |   |   IPK > 3.255: Terima {Terima=2, Tidak=0}
|   |   |   |   |   |   |   IPK ≤ 3.255: Tidak {Terima=0, Tidak=2}
|   |   |   |   |   |   |   Penghasilan = Tinggi: Terima {Terima=2, Tidak=0}
|   |   |   |   |   |   Ikut Organisasi = Tidak
|   |   |   |   |   |   |   Jenis Kelamin = L: Terima {Terima=6, Tidak=2}
|   |   |   |   |   |   |   Jenis Kelamin = P
|   |   |   |   |   |   |   |   IPK > 3.130
|   |   |   |   |   |   |   |   |   IPK > 3.155
|   |   |   |   |   |   |   |   |   |   IPK > 3.450: Tidak {Terima=0, Tidak=3}
|   |   |   |   |   |   |   |   |   |   IPK ≤ 3.450
|   |   |   |   |   |   |   |   |   |   |   Penghasilan = Rendah: Terima {Terima=2, Tidak=1}
|   |   |   |   |   |   |   |   |   |   |   Penghasilan = Tinggi: Tidak {Terima=0, Tidak=2}
|   |   |   |   |   |   |   |   |   |   |   IPK ≤ 3.155: Tidak {Terima=0, Tidak=5}
|   |   |   |   |   |   |   |   |   |   |   IPK ≤ 3.130: Terima {Terima=2, Tidak=1}
|   |   |   |   |   |   |   |   |   IPK ≤ 3.055: Tidak {Terima=0, Tidak=2}
|   |   |   |   |   |   Tanggungan ≤ 4.500
|   |   |   |   |   |   |   Ikut Organisasi = Ikut
|   |   |   |   |   |   |   |   IPK > 3.855: Terima {Terima=2, Tidak=0}
|   |   |   |   |   |   |   |   IPK ≤ 3.855
|   |   |   |   |   |   |   |   |   IPK > 2.950
|   |   |   |   |   |   |   |   |   |   Tanggungan > 3.500
|   |   |   |   |   |   |   |   |   |   |   Penghasilan = Rendah: Terima {Terima=4, Tidak=2}
|   |   |   |   |   |   |   |   |   |   |   Penghasilan = Sedang
|   |   |   |   |   |   |   |   |   |   |   |   Jenis Kelamin = L: Terima {Terima=2, Tidak=2}

```

Gambar 5. Description Decision Tree

Penjelasan *Description tree* :

- 1 Tanggungan > 4.500
 - a) Jika IPK > 3.055 maka Jika ikut Organisasi IPK > 3.405 Maka diterima
 - b) IPK < 3.405 lihat faktor penghasilan & jenis kelamin jika penghasilan rendah laki-laki lebih banyak tidak diterima sedangkan perempuan lebih banyak diterima ,
 - c) jika penghasilan sedang maka IPK diatas > 3.255 diterima < 3.255 tidak diterima,
 - d) Jika tidak mengikuti organisasi laki-laki sebagian besar diterima, perempuan jika IPK < 3,130 sebagian besar diterima

- e) IPK antara 3,130-3,450 hasil bervariasi tergantung penghasilan rendah lebih banyak diterima sedangkan tinggi tidak jika IPK <3,055 tidak diterima
- 2. Tanggungan <4.500
 - a) Jika ikut organisasi IPK >3,855 diterima
 - b) IPK antara 2,950 – 3,855 dipengaruhi oleh penghasilan dan jenis kelamin , Penghasilan rendah masih banyak diterima, Penghasilan sedang hasil seimbang, Penghasilan tinggi banyak yang tidak diterima
 - c) Jika tidak ikut organisasi IPK > 3,715 Penghasilan rendah hampir semua diterima, Penghasilan sedang bervariasi bergantung tanggungan , Penghasilan tinggi hampir semua tidak diterima
 - d) IPK < 3,715 sebagian besar tidak diterima (38 vs 403)
- 3. Tidak Ikut UKM
 - a) Jika IPK >3,825 ikut organisasi, tanggungan < 3,500 semua diterima
 - b) tanggungan >3,500 tidak diterima , tidak ikut organisasi semua diterima,
 - c) IPK antara 2,755-3,825 tanggungan >4,500 penghasilan rendah , IPK > 3,405 tidak diterima
 , IPK < 3,405 tergantung organisasi dan jenis kelamin, ikut organisasi sebagian besar diterima, tidak ikut perempuan lebih banyak tidak diterima, penghasilan sedang ikut organisasi semua diterima, tidak ikut perempuan lebih banyak diterima, penghasilan tinggi semua diterima,
 - d) tanggungan <4,500 sebagian besar tidak diterima, jika IPK <2,755 tidak diterima

Evaluasi

accuracy: 80.10%

	true Terima	true Tidak	class precision
pred. Terima	70	18	79.55%
pred. Tidak	148	598	80.16%
class recall	32.11%	97.08%	

Gambar 6. Hasil dari Evaluasi untuk mengetahui hasil keakurasian data

Berdasarkan hasil pengujian menggunakan Rapid Miner diperoleh hasil dengan Algoritma C4.5 dimana dari total entry sebanyak 1042 data diperoleh sebagai berikut

1. Tingkat akurasi sebesar 80,10%
2. Sebanyak 70 data prediksi terima yaitu masuk kedalam class terima dan masuk kedalam class tidak sebanyak 18
3. Sebasnyak 148 data prediksi tidak masuk kedalam class terima dan masuk kedalam class tidak sebanyak 598
4. Nilai true terima sebesar 32,11% sedangkan nilai true tidak sebesar 97,08%
- e) Nilai *class precision* pred terima sebesar 79,55% sedangkan pred tidak sebesar 80,16%

KESIMPULAN

Berdasarkan hasil penelitian dan implementasi menggunakan rapid miner pada klasifikasi calon penerima beasiswa didapatkan hasil berupa tingkat keakurasian data sebesar 80,10% dapat dilihat pada Gambar 5. Hasil Evaluasi untuk mendapatkan hasil akurasi Data. Mahasiswa aktif di UKM dan organisasi, serta memiliki IPK tinggi atau tanggungan besar, memiliki peluang lebih besar pada peluang penerimaan beasiswa, Sebanyak 70 data prediksi diterima berhasil diidentifikasi sedangkan 18 tidak sebanyak 148 data prediksi tidak, tidak masuk kedalam kelas terima dan sebanyak 598 data masuk ke dalam kelas tidak dapat dilihat pada penjelasan description tree

DAFTAR PUSTAKA

- Akbar, F., Saputra, H. W., & Maulaya, A. K. (2022). *Implementation of decision tree algorithm C4.5 and support vector regression for stroke disease prediction*. **2**, 61–67.
- Apdian, D., Tiarna, M., & Hutabarat, B. (2024). Sistem penunjang keputusan penerima beasiswa SMK Ristek Karawang berbasis web menggunakan metode SMART. **17**, 17–24.
- Budiarto, E., Rino, R., Hariyanto, S., & Susilawati, D. (2022). Penerapan data mining untuk rekomendasi beasiswa pada SD Maria Mediatrix menggunakan algoritma C4.5. *ALGOR*, **3**(2), 23–34.
- Hasmin, E., & Aisa, S. (2019). Penerapan algoritma C4.5 untuk penentuan penerima beasiswa mahasiswa. *CogITo Smart Journal*, **5**(2), 308–320.
- Kegiatan, I. (2022). Penerapan data mining dengan metode klasifikasi menggunakan algoritma C4.5. **9**(4).
- Khasanah, N., Uki, D., Saputri, E., Hidayat, T., & Aziz, F. (2025). Prediksi kelulusan mahasiswa menggunakan algoritma C4.5 dengan RapidMiner: Studi kasus data akademik perguruan tinggi XYZ. **4**(2), 100–107.
- Khatib, J. (2021). *Indonesian Journal of Computer Science*, **10**(1).
- Khotimah, K. (2021). Teknik data mining menggunakan algoritma decision tree (C4.5)

untuk prediksi seleksi beasiswa jalur KIP pada Universitas Muhammadiyah Kotabumi. *4*(2), 145–152.

- Mantopani, R., Ramadhan, I. L., Samsara, R. P., & Samsinar, R. (2024). Analisis perbandingan decision tree dan algoritma C4.5 untuk mengklasifikasikan penerimaan mahasiswa elektro. 295–302.
- Muriyatmoko, D., Musthafa, A., & Wijaya, M. H. (2024). Klasifikasi profil kelulusan nilai AKPAM dengan metode decision tree. April.
- Oktafiani, R., Hermawan, A., & Avianto, D. (2023). Pengaruh komposisi split data terhadap performa klasifikasi penyakit kanker payudara menggunakan algoritma machine learning. *Jurnal Sistem Informasi*, *9*(1), 19–28. <https://doi.org/10.34128/jsi.v9i1.622>
- Pamuji, F. Y., & Ramadhan, V. P. (2021). Komparasi algoritma random forest dan decision tree untuk memprediksi keberhasilan immunotherapy. *Jurnal Teknologi dan Manajemen Informatika*, *7*(1), 46–50.
- Rismayanti, R. (2016). Implementasi algoritma C4.5 untuk menentukan penerima beasiswa di STT Harapan Medan. *Jurnal Media Infotama*, *12*(2).
- Sandiva, T. V., Defit, S., & Nurcahyo, G. W. (2024). Implementasi algoritma C4.5 untuk prediksi penerima beasiswa Program Indonesia Pintar. *Jurnal KomtekInfo*, 354–362.
- Syahriani, S. (2022). Penerapan data mining untuk menentukan pola penjualan sepatu menggunakan metode algoritma apriori. *9*(1), 43–52.
- Tabuni, N., Satyawati, S. T., & Ismanto, B. (2024). Evaluasi program beasiswa bagi mahasiswa Kabupaten Pegunungan Bintang menggunakan model goal oriented evaluation.