

Implementasi Algoritma C4.5 untuk Prediksi Penerima Beasiswa Program Indonesia Pintar

Tesa Vausia Sandiva[✉], Sarjon Defit, Gunadi Widi Nurcahyo

Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK, Padang, 25221, Indonesia

sandivatesa@gmail.com

Abstract

Education is the main foundation in human resource development and plays an important role in improving the quality of life. However, financial constraints often become a barrier for many underprivileged families to continue their education to a higher level. The Smart Indonesia Program (PIP) is an Indonesian government initiative designed to provide financial assistance to students who meet certain criteria, so that they can continue their education without economic constraints. This study aims to develop a prediction model for PIP scholarship recipients using the C4.5 Algorithm, which has been proven effective in dealing with classification problems. This study used 50 student data, obtained from the Elementary School Dapodik System. This data was analyzed to find relevant patterns in determining the eligibility of scholarship recipients, which were then used to build a prediction model. The main pattern found showed that the status of students as recipients of the Smart Indonesia Card (KIP) greatly determines the prediction of the C4.5 model KIP recipient students tend to be predicted to receive PIP scholarships. The results of the study showed that the resulting classification model had a very high level of accuracy, which was 96.00%. In addition, the precision and recall for the scholarship recipient category reached 95.65%, while the precision and recall for the non-scholarship recipient category were recorded at 96.30%. These findings indicate that the C4.5 Algorithm-based prediction model can function as a reliable method to support decision-making in the education sector, and has the potential to be applied in various contexts that require accurate data-based predictions. Thus, this study not only makes a significant contribution to the development of science, but also has a real impact on efforts to equalize education in Indonesia.

Keywords: Smart Indonesia Program (PIP), Prediction, C4.5 Algorithm, Classification, Smart Indonesia Card (KIP)

Abstrak

Pendidikan merupakan fondasi utama dalam pengembangan sumber daya manusia dan berperan penting dalam meningkatkan kualitas hidup. Namun, hambatan finansial sering kali menjadi penghalang bagi banyak keluarga kurang mampu untuk melanjutkan pendidikan ke jenjang yang lebih tinggi. Program Indonesia Pintar (PIP) adalah inisiatif pemerintah Indonesia yang dirancang untuk memberikan bantuan keuangan kepada siswa yang memenuhi kriteria tertentu, sehingga mereka dapat melanjutkan pendidikan tanpa kendala ekonomi. Penelitian ini bertujuan untuk mengembangkan model prediksi penerima beasiswa PIP dengan menggunakan Algoritma C4.5, yang telah terbukti efektif dalam menangani masalah klasifikasi. Penelitian ini menggunakan 50 data siswa, yang diperoleh dari Sistem Dapodik Sekolah Dasar. Data ini dianalisis untuk menemukan pola-pola yang relevan dalam menentukan kelayakan penerima beasiswa, yang kemudian digunakan untuk membangun model prediksi. Pola utama yang ditemukan menunjukkan bahwa status siswa sebagai penerima Kartu Indonesia Pintar (KIP) sangat menentukan prediksi model C4.5 siswa penerima KIP cenderung diprediksi menerima beasiswa PIP. Hasil penelitian menunjukkan bahwa model klasifikasi yang dihasilkan memiliki tingkat akurasi yang sangat tinggi, yaitu sebesar 96.00%. Selain itu, *precision* dan *recall* untuk kategori penerima beasiswa masing-masing mencapai 95.65%, sementara *precision* dan *recall* untuk kategori bukan penerima beasiswa tercatat sebesar 96.30%. Temuan ini menunjukkan bahwa model prediksi berbasis Algoritma C4.5 dapat berfungsi sebagai metode yang dapat diandalkan untuk mendukung pengambilan keputusan di sektor pendidikan, serta memiliki potensi untuk diterapkan dalam berbagai konteks yang memerlukan prediksi berbasis data yang akurat. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi signifikan terhadap pengembangan ilmu pengetahuan, tetapi juga memiliki dampak nyata dalam upaya pemerataan pendidikan di Indonesia.

Kata kunci: Program Indonesia Pintar (PIP), Prediksi, Algoritma C4.5, Klasifikasi, Kartu Indonesia Pintar (KIP)

KomtekInfo is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



1. Pendahuluan

Pengetahuan yang ditemukan dalam basis data, yang juga dikenal sebagai Knowledge Discovery in Databases (KDD), merupakan pendekatan yang memanfaatkan Data Mining untuk mengeksplorasi

pengetahuan dari data berukuran besar. KDD menjadi proses yang penting dalam upaya mengungkap informasi tersembunyi dari data. Meskipun sering kali istilah Data Mining dan KDD digunakan secara bergantian dalam konteks ini, keduanya sebenarnya

memiliki fokus yang berbeda [1]. Data *Mining* sendiri adalah metode analisis yang bertujuan untuk mengidentifikasi korelasi, pola, dan tren yang tersembunyi dalam kumpulan data besar. Algoritma-Algoritma seperti klasifikasi, pengelompokan, dan asosiasi adalah beberapa yang paling umum digunakan dalam Data *Mining*, dengan fleksibilitas yang tidak dimiliki oleh pendekatan statistik tradisional yang lebih kaku [2]. Salah satu teknik yang paling sering digunakan dalam Data *Mining* adalah Machine Learning [3]. Machine Learning, yang merupakan bagian penting dari bidang *Artificial Intelligence* (AI), telah berkembang pesat dan secara bertahap masuk ke berbagai industri [4].

Algoritma C4.5 adalah salah satu algoritma yang dikenal luas dalam pembuatan pohon keputusan, yang digunakan untuk melakukan klasifikasi berdasarkan data latih [5]. Pohon keputusan ini banyak digunakan dalam Data *Mining* dan Machine Learning karena strukturnya yang sederhana namun efektif, serta kemampuannya untuk menghasilkan hasil klasifikasi yang akurat dan dapat diandalkan [6]. Dalam proses ini, Algoritma C4.5 menggunakan rasio perolehan informasi untuk menentukan atribut terbaik yang akan digunakan sebagai dasar pembagian dalam pohon keputusan. Algoritma ini juga dikenal dengan kemampuannya untuk menangani atribut diskrit dan kontinu serta data yang tidak lengkap [7].

Dalam konteks pendidikan, terutama di Indonesia, biaya sering kali menjadi hambatan utama bagi keluarga berpenghasilan rendah untuk melanjutkan pendidikan anak-anak mereka. Beasiswa merupakan solusi penting dalam situasi ini, yang dirancang untuk mengurangi beban finansial dan membantu siswa berprestasi dari kalangan kurang mampu agar dapat terus belajar. Salah satu program beasiswa yang dilaksanakan oleh pemerintah Indonesia adalah Program Indonesia Pintar [8]. Program Indonesia Pintar adalah sebuah inisiatif yang diperkenalkan oleh pemerintah Indonesia di bawah kepemimpinan Presiden Joko Widodo, sesuai dengan Peraturan Presiden No. 7 Tahun 2014. Salah satu elemen kunci dari program ini adalah Kartu Indonesia Pintar, yang bertujuan untuk memberikan bantuan keuangan kepada siswa yang berasal dari keluarga kurang mampu dengan tujuan meningkatkan akses mereka terhadap pendidikan yang berkualitas [9]. Meskipun Program Indonesia Pintar telah memberikan dampak positif yang signifikan, tantangan utama terletak pada proses seleksi penerima bantuan. Variabel-variabel seperti pekerjaan orang tua, penghasilan, jenis tempat tinggal, serta status penerima KPS dan KIP harus dianalisis secara cermat untuk memastikan bantuan ini diberikan kepada mereka yang paling membutuhkan [10].

Penelitian yang sama juga menjelaskan tentang penerapan Algoritma C4.5 dalam proses klasifikasi penerimaan bantuan sosial dengan menggunakan *Feature Selection*. Hasil penelitian ini menunjukkan

bahwa akurasi dalam pemilihan penerima KIP mencapai 98,21%, dengan presisi yang konsisten dan sebesar 99,48% [11]. Penerapan Algoritma C4.5 untuk menentukan jurusan siswa baru di Sekolah Menengah Kejuruan juga menunjukkan hasil yang signifikan, dengan akurasi sebesar 98,02%, presisi mencapai 98,73%, dan nilai *recall* 98,73% [12]. Penelitian lainnya membahas tentang konstruksi dan optimalisasi sistem pendukung keputusan mesin penjual otomatis berdasarkan pohon keputusan C4.5, di mana hasil penelitian ini menunjukkan bahwa akurasi Algoritma ini bervariasi, dengan nilai tertinggi mencapai 87% dan terendah 68% [13].

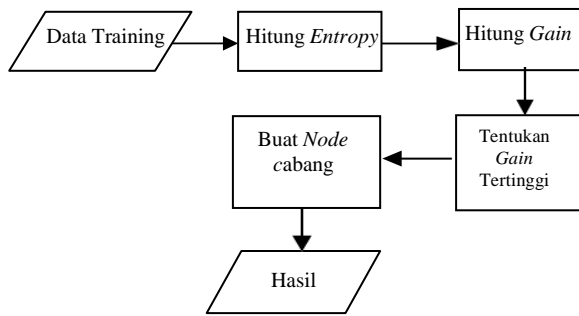
Selanjutnya Penelitian lain yang meneliti penerapan Algoritma Decision Tree C4.5 untuk menentukan penerima beasiswa Program Indonesia Pintar (PIP) di SMA Negeri 3 Timang Gajah menunjukkan hasil yang bervariasi. Tingkat akurasi untuk menentukan siswa yang layak dan tidak layak menerima PIP adalah sebesar 72%. Dari seluruh siswa yang diprediksi layak menerima PIP, sebanyak 85% benar-benar memenuhi kriteria kelayakan (*Precision*). Namun, dari total siswa yang sebenarnya layak menerima PIP, hanya 69% yang berhasil diidentifikasi dengan benar (*Recall*). Sementara itu, tingkat *Specificity*, yang mengukur ketepatan dalam memprediksi siswa yang tidak layak menerima PIP dibandingkan dengan seluruh siswa yang memang tidak layak, mencapai 78% [14].

Berdasarkan beberapa penjelasan sebelumnya, penelitian ini akan menyajikan proses implementasi Algoritma C4.5 untuk prediksi penerima beasiswa Program Indonesia Pintar. Proses analisis tersebut menggunakan konsep Data *Mining* dengan fokus pada kinerja Algoritma C4.5 dalam menyajikan hasil prediksi yang tepat dan akurat. Proses prediksi dilakukan dengan tujuan untuk menganalisis kinerja Algoritma C4.5 dalam mengidentifikasi faktor-faktor yang mempengaruhi kelayakan penerimaan beasiswa pada periode waktu berikutnya. Hasil penelitian ini diharapkan dapat berkontribusi dalam menyajikan rekomendasi yang tepat untuk menentukan siapa yang layak menerima beasiswa dari Program Indonesia Pintar. Dengan penerapan Algoritma yang tepat, diharapkan dapat meminimalkan kesalahan dalam penentuan penerima beasiswa, sehingga bantuan pendidikan dapat diberikan secara lebih efektif dan efisien kepada siswa yang benar-benar membutuhkan.

2. Metodologi Penelitian

Proses prediksi menggunakan algoritma C4.5 bertujuan untuk mencapai hasil yang optimal dalam memprediksi penerima bantuan dari Program Indonesia Pintar. Proses prediksi menggunakan algoritma C4.5 bertujuan untuk mencapai hasil yang optimal dalam memprediksi penerima bantuan dari Program Indonesia Pintar. Analisis ini dilakukan pada data penerima beasiswa yang diperoleh dari sumber resmi yang relevan.

Tahapan penelitian tersebut dapat dilihat pada Gambar 1.



Gambar 1. Tahap Penelitian

Untuk memahami lebih lanjut tahapan yang ditampilkan pada Gambar 1 di atas, proses pertama adalah menghitung *entropy* yang bertujuan untuk mengukur tingkat variasi dalam data pelatihan, sehingga dapat menentukan seberapa baik suatu atribut memisahkan data. Selanjutnya, nilai *gain* dihitung untuk menilai efektivitas pemisahan berdasarkan atribut tersebut, dan atribut dengan *gain* tertinggi dipilih sebagai titik pemisah dalam pembuatan *node* cabang.

2.1 Data Training

Data yang dikumpulkan adalah dari Dapodik siswa penerima beasiswa Program Indonesia Pintar di Sekolah Dasar selama tahun 2024. Data yang digunakan dalam pengolahan mencakup 50 sampel. Proses input Data Training ini merupakan tahap awal untuk menghitung nilai *entropy* dan *gain*. Sampel data training dapat disajikan pada Tabel 1.

Tabel 1. Data Penerima Program Indonesia Pintar

No	Pekerjaan Orang Tua	Penghasilan Orang Tua	PKIP	PKPS	Prediksi
1	Petani	1,000,000 – 1,999,999	Ya	Tidak	Ya
2	Petani	<500,000	Ya	Tidak	Ya
3	Petani	<500,000	Ya	Ya	Ya
4	Wiraswasta	5,000,000 – 20,000,000	Tidak	Tidak	Tidak
5	Karyawan Swasta	<500,000	Tidak	Tidak	Ya
..
45	Petani	<500,000	Ya	Ya	Ya
46	Wiraswasta	500,000 – 999,999	Ya	Ya	Ya
47	Petani	500,000 – 999,999	Tidak	Ya	Ya
48	Petani	<500,000	Ya	Ya	Ya
49	PNS/TNI/POLRI	5,000,000 – 20,000,000	Tidak	Tidak	Tidak
50	PNS/TNI/POLRI	5,000,000 – 20,000,000	Tidak	Tidak	Tidak

Tabel 1 menggambarkan hasil prediksi kelayakan penerimaan beasiswa Program Indonesia Pintar berdasarkan beberapa variabel seperti, pekerjaan orang

tua, penghasilan orang tua, penerimaan Kartu Indonesia Pintar (PKIP), dan penerimaan Kartu Perlindungan Sosial (PKPS). Setiap baris mewakili data seorang siswa, dengan prediksi apakah siswa tersebut dianggap layak menerima beasiswa. Misalnya, siswa dengan latar belakang keluarga petani dengan penghasilan rendah dan memiliki PKIP umumnya diprediksi layak menerima beasiswa. Sebaliknya, siswa yang orang tuanya memiliki pekerjaan dengan penghasilan tinggi, seperti PNS atau TNI/POLRI, cenderung diprediksi tidak layak menerima beasiswa.

2.2 Hitung Entropy

Pada tahap ini melakukan proses perhitungan dengan nilai atribut sehingga menghasilkan nilai *entropy*. Pencarian *entropy* total dilakukan dengan cara mengelompokkan data dengan benar dan tepat. Kemudian menghitung data menggunakan rumus *entropy* dengan Persamaan 1.

$$Entropy(A) = (S) - \sum_{i=1}^n - Pi * \log_2 Pi \tag{1}$$

Dimana:

S : Himpunan Kasus

A : Fitur

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

2.3 Hitung Gain

Dalam melakukan perhitungan *gain*, harus mendapatkan nilai *entropy* terlebih dahulu. *Gain* adalah perbedaan antara *entropy* data awal dan rata-rata *entropy* setiap atribut. Perhitungan *Gain* dapat dilakukan dengan menggunakan Persamaan 2.

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|Si|}{|S|} * Entropy(Si) \tag{2}$$

Dimana:

S : Himpunan Kasus

A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

2.4 Tentukan Gain Tertinggi

Pada tahap ini, sangat penting untuk menentukan akar dari pohon keputusan, karena akar akan menjadi dasar dari seluruh struktur pohon. Pemilihan atribut yang tepat sebagai akar akan memengaruhi efektivitas dan akurasi klasifikasi. Jika atribut dengan *gain* tertinggi dipilih dengan benar, pohon keputusan yang dihasilkan akan lebih optimal dalam melakukan klasifikasi data secara akurat dan konsisten.

2.4.1 Membuat Node

Pada tahap ini, cabang dibuat untuk menghasilkan daun dari pohon keputusan, yang merupakan langkah penting dalam proses klasifikasi. Daun mewakili keputusan akhir atau hasil dari setiap jalur pada pohon, yang berarti setiap daun menunjukkan kategori atau label

yang telah ditentukan berdasarkan atribut yang dipilih sebelumnya. Proses ini memastikan bahwa setiap data dalam set pelatihan dapat diklasifikasikan dengan benar sesuai dengan aturan yang dihasilkan oleh pohon keputusan.

2.5 Transformasi Data

Sebelum melangkah ke tahap hasil dan pembahasan, dilakukan transformasi data terlebih dahulu. Dari Tabel 1 yang berisi Data Penerima Program Indonesia Pintar, transformasi dilakukan pada atribut penghasilan orang tua (SO) dan pekerjaan orang tua (POT) untuk mempermudah proses pengolahan data. Atribut pendukung yang digunakan dapat dilihat pada Tabel 2.

Tabel 2. Indikator Pekerjaan Orang Tua (SO)

Pekerjaan Orang Tua (SO)	Transformasi
Petani, Wiraswasta, Lainnya	Rendah
Wirasaha, Karyawan BUMN, Karyawan Swasta	Menengah
PNS/TNI/Polri	Tinggi

Tabel 2 di atas menunjukkan klasifikasi pekerjaan orang tua berdasarkan tingkat transformasi sosial ekonomi. Pekerjaan seperti petani, wiraswasta, dan lainnya dikategorikan sebagai kelompok dengan tingkat transformasi rendah. Sementara itu, pekerjaan sebagai PNS, TNI, atau Polri ditempatkan dalam kategori transformasi tinggi, sedangkan wirasaha, karyawan BUMN, dan karyawan swasta berada pada tingkat menengah pada Tabel 3.

Tabel 3. Indikator Penghasilan Orang Tua (SO)

Penghasilan Orang Tua (SO)	Transformasi
<1,000,000 – 1,999,999	Rendah
2,000,000 – 4,999,999	Menengah
5,000,000 – 20,000,000	Tinggi

Tabel 3 transformasi memiliki beberapa atribut, antara lain Pekerjaan Orang Tua (POT), Penghasilan Orang Tua (SO), Penerima Kartu Indonesia Pintar (PKIP), Kartu Perlindungan Sosial (PKPS), dan Jenis Tinggal (JT). Data transformasi ini akan digunakan sebagai dataset dalam perhitungan *node* untuk pohon keputusan pada Tabel 4.

Tabel 4. Data Transformasi

No	POT	SO	PKIP	PKPS	Prediksi
1	Rendah	Rendah	Ya	Tidak	Ya
2	Rendah	Rendah	Ya	Tidak	Ya
3	Rendah	Rendah	Ya	Ya	Ya
4	Rendah	Tinggi	Tidak	Tidak	Tidak
5	Menengah	Rendah	Tidak	Tidak	Ya
..
45	Rendah	Rendah	Ya	Ya	Ya
46	Rendah	Rendah	Ya	Ya	Ya
47	Rendah	Rendah	Tidak	Ya	Ya
48	Rendah	Rendah	Ya	Ya	Ya
49	Tinggi	Tinggi	Tidak	Tidak	Tidak
50	Tinggi	Tinggi	Tidak	Tidak	Tidak

Tabel 4 di atas menunjukkan hasil prediksi kelayakan penerimaan beasiswa Program Indonesia Pintar

berdasarkan beberapa atribut, yaitu pekerjaan orang tua (POT), penghasilan orang tua (SO), penerimaan Kartu Indonesia Pintar (PKIP), dan penerimaan Kartu Perlindungan Sosial (PKPS). Setiap baris dalam table mewakili data seorang siswa, dengan kolom Prediksi menunjukkan apakah siswa tersebut diprediksi layak menerima beasiswa. Sebagian besar siswa dengan penghasilan dan pekerjaan orang tua yang rendah serta memiliki KIP, diprediksi layak menerima beasiswa, sementara siswa dengan penghasilan dan pekerjaan orang tua yang tinggi lebih cenderung tidak layak menerima bantuan tersebut.

3. Hasil dan Pembahasan

Pada bagian hasil dan pembahasan, data yang telah diproses dianalisis untuk mengevaluasi kinerja model prediksi yang digunakan. Hasil dari analisis ini memberikan wawasan tentang seberapa akurat model dalam memprediksi penerima beasiswa, serta mengidentifikasi faktor-faktor yang paling berpengaruh dalam menentukan kelayakan penerima.

3.1 Perhitungan *Entropy* Total

Proses pencarian *entropy* total dan *gain* dilakukan dengan mengelompokkan data secara benar dan tepat. Setelah data dikelompokkan, perhitungan *entropy* dilakukan menggunakan rumus yang telah ditentukan untuk setiap atribut. Rumus *gain* kemudian diaplikasikan pada atribut yang sama guna menentukan atribut dengan pemisahan terbaik dalam proses klasifikasi. Setelah itu, perhitungan dilakukan menggunakan Persamaan 1 dan Persamaan 2 pada setiap atribut data.

$$\begin{aligned} &\text{Hasil Perhitungan Entropy Total} \\ &Entropy\ Total = 0.995378439 \end{aligned}$$

Nilai *Entropy Total* yang dihitung adalah 0,995378439. Nilai ini menunjukkan bahwa data memiliki tingkat keragaman yang cukup tinggi. Nilai ini dihitung dengan menggunakan rumus perhitungan *entropy* sesuai dengan Persamaan 1. Dengan nilai *entropy* yang mendekati 1, diperlukan pemisahan data yang lebih baik agar klasifikasi bisa dilakukan dengan lebih akurat.

3.2 Perhitungan *Entropy* Tiap Atribut

Perhitungan nilai *entropy* dilakukan pada setiap atribut, yaitu Pekerjaan Orang Tua (POT), Penghasilan Orang Tua (SO), Kartu Indonesia Pintar (KIP), Kartu Perlindungan Sosial (KPS), dan Jenis Tinggal (JT). Perhitungan ini menggunakan persamaan (1) yang telah dijelaskan sebelumnya. Hasil dari perhitungan *entropy* tersebut dapat dilihat pada tabel berikut.

3.2.1 Atribut POT

Atribut Pekerjaan Orang Tua (POT) merupakan salah satu variabel penting dalam proses klasifikasi. Pekerjaan orang tua dibagi menjadi beberapa kategori seperti, petani, wiraswasta, karyawan, dan PNS.

Kategori-kategori ini akan memengaruhi bagaimana data tersebut dikelompokkan dalam perhitungan *entropy* dan *gain* untuk pembuatan pohon keputusan.

Hasil Perhitungan *Entropy* kategori POT = Rendah
 $Entropy = 0,797326$

Hasil perhitungan *entropy* untuk kategori Pekerjaan Orang Tua (POT) dengan status Rendah. Nilai *entropy* yang dihasilkan adalah 0,797326, yang dihitung dengan menggunakan rumus *entropy* sesuai dengan Persamaan 1. Nilai ini digunakan dalam proses pembuatan pohon keputusan untuk menentukan seberapa baik kategori ini dapat memisahkan data ke dalam berbagai kelompok.

Hasil Perhitungan *Entropy* Kategori POT = Menengah
 $Entropy = 0,413816$

Hasil perhitungan *entropy* untuk kategori Pekerjaan Orang Tua (POT) dengan status Menengah. Nilai *entropy* yang dihasilkan adalah 0,413816, yang dihitung dengan menggunakan rumus *entropy* sesuai dengan Persamaan 1. Nilai ini digunakan dalam proses pembuatan pohon keputusan untuk menentukan seberapa baik kategori ini dapat memisahkan data ke dalam berbagai kelompok.

Hasil Perhitungan *Entropy* Kategori POT = Tinggi
 $Entropy = 0$

Hasil perhitungan *entropy* untuk kategori Pekerjaan Orang Tua (POT) dengan status Tinggi adalah 0. Nilai ini diperoleh menggunakan rumus entropi berdasarkan Persamaan 1. Nilai *entropy* sebesar 0 menunjukkan bahwa dalam kategori Tinggi tidak terdapat keragaman, sehingga data sudah terkelompokkan secara sempurna tanpa perlu pemisahan lebih lanjut.

3.2.2 Atribut SO

Atribut Penghasilan Orang Tua (SO) digunakan untuk mengukur tingkat kesejahteraan ekonomi keluarga. Penghasilan orang tua dikelompokkan dalam beberapa tingkatan, seperti rendah, menengah, dan tinggi. Tingkat penghasilan ini memengaruhi kualifikasi siswa untuk menerima bantuan atau program dukungan pendidikan.

Hasil Perhitungan *Entropy* Kategori SO = Rendah
 $Entropy = 0,797326$

Hasil perhitungan *entropy* untuk kategori Penghasilan Orang Tua (SO) dengan status Rendah adalah 0,797326. Nilai ini dihitung menggunakan rumus *entropy* yang dijelaskan dalam Persamaan 1. Nilai *entropy* yang mendekati 1 menunjukkan bahwa terdapat tingkat keragaman yang cukup tinggi dalam data kategori Rendah, sehingga masih diperlukan pemisahan lebih lanjut untuk klasifikasi yang lebih optimal.

Entropy Kategori SO = Menengah
 $Entropy = 0,353359$

Hasil perhitungan *entropy* untuk kategori Penghasilan Orang Tua (SO) dengan status Menengah adalah 0,353359. Nilai ini dihitung menggunakan rumus *entropy* Persamaan 1. Nilai *entropy* yang lebih rendah ini menunjukkan bahwa data dalam kategori Menengah memiliki tingkat keragaman yang lebih kecil dibandingkan kategori lainnya, namun masih memerlukan pemisahan lebih lanjut untuk klasifikasi yang optimal.

Hasil Perhitungan *Entropy* Kategori SO = Tinggi
 $Entropy = 0$

Hasil perhitungan *entropy* untuk kategori Penghasilan Orang Tua (SO) dengan status Tinggi adalah 0. Nilai ini dihitung menggunakan rumus *entropy* Persamaan 1. Nilai ini menunjukkan bahwa data dalam kategori tersebut tidak memiliki variasi. Oleh karena itu, tidak diperlukan pemisahan lebih lanjut dalam proses klasifikasi.

3.2.3 Atribut PKIP

Atribut Penerima Kartu Indonesia Pintar (PKIP) digunakan untuk menentukan apakah seorang siswa terdaftar dalam program bantuan pendidikan dari pemerintah. Siswa yang menerima Kartu Indonesia Pintar (KIP) biasanya berasal dari keluarga dengan kondisi ekonomi yang kurang mampu. Atribut ini penting dalam proses klasifikasi karena memengaruhi kelayakan siswa untuk menerima bantuan pendidikan lebih lanjut.

Hasil Perhitungan *Entropy* Kategori PKIP = Ya
 $Entropy = 0$

Hasil perhitungan *entropy* untuk kategori Penerima Kartu Indonesia Pintar (PKIP) dengan status Ya adalah 0. Nilai ini dihitung menggunakan rumus *entropy* Persamaan 1. Nilai *entropy* 0 ini menunjukkan bahwa data dalam kategori ini tidak memiliki variasi, yang berarti semua data dalam kategori Ya bersifat sama. Oleh karena itu, tidak diperlukan pemisahan lebih lanjut untuk kategori ini dalam proses klasifikasi.

Hasil Perhitungan *Entropy* Kategori PKIP = Tidak
 $Entropy = 0,684038$

Hasil perhitungan *entropy* untuk kategori Penerima Kartu Indonesia Pintar (PKIP) dengan status "Tidak" adalah 0,684038. Nilai ini dihitung menggunakan rumus *entropy* Persamaan 1. Nilai *entropy* ini menunjukkan bahwa ada perbedaan dalam data pada kategori ini. Oleh karena itu, masih diperlukan pemisahan lebih lanjut dalam proses klasifikasi untuk mendapatkan hasil yang lebih baik.

3.2.4 Atribut PKPS

Atribut Penerima Kartu Perlindungan Sosial (PKPS) digunakan untuk menunjukkan apakah seseorang menerima bantuan sosial dari pemerintah. Kartu ini

diberikan kepada keluarga dengan kondisi ekonomi yang kurang mampu sebagai bentuk perlindungan sosial. Atribut ini penting dalam proses klasifikasi karena dapat memengaruhi kelayakan individu untuk menerima bantuan lebih lanjut atau program dukungan lainnya.

Hasil Perhitungan *Entropy* Kategori PKPS = Ya
 $Entropy = 0$

Hasil perhitungan *entropy* untuk kategori Penerima Kartu Perlindungan Sosial (PKPS) dengan status Ya adalah 0. Nilai *entropy* 0 ini menunjukkan bahwa semua data dalam kategori ini sama. Oleh karena itu, tidak diperlukan pemisahan lebih lanjut dalam proses klasifikasi.

Hasil Perhitungan *Entropy* Kategori PKPS = Tidak
 $Entropy = 0,890491$

Hasil perhitungan *entropy* untuk kategori Penerima Kartu Perlindungan Sosial (PKPS) dengan status "Tidak" adalah 0,890491. Nilai *entropy* ini menunjukkan adanya perbedaan yang cukup besar dalam data pada kategori ini. Oleh karena itu, pemisahan lebih lanjut diperlukan dalam proses klasifikasi untuk mencapai hasil yang lebih baik.

3.3 Menghitung Nilai *Gain*

Perhitungan nilai *gain* dilakukan untuk setiap atribut, yaitu Pekerjaan Orang Tua (POT), Penghasilan Orang Tua (SO), Kartu Indonesia Pintar (KIP), Kartu Perlindungan Sosial (KPS), dan Jenis Tinggal (JT). Informasi lebih lanjut mengenai perhitungan *gain* dapat dilihat pada Persamaan (1) dan (2). Hasil perhitungan nilai *gain* untuk masing-masing atribut akan dijelaskan dalam bagian berikut.

3.3.1 Atribut POT

Atribut Pekerjaan Orang Tua (POT) merupakan salah satu variabel penting dalam proses klasifikasi. Pekerjaan orang tua dibagi menjadi beberapa kategori seperti, petani, wiraswasta, karyawan, dan PNS. Kategori-kategori ini akan memengaruhi bagaimana data tersebut dikelompokkan dalam perhitungan *entropy* dan *gain* untuk pembuatan pohon keputusan.

Hasil Perhitungan *Gain* Kategori POT
 $Gain = 0,433613019$

Hasil perhitungan *gain* untuk kategori Pekerjaan Orang Tua (POT) adalah 0,433613019. Perhitungan *gain* ini dilakukan dengan menggunakan rumus Persamaan 2. Nilai *gain* tersebut menunjukkan seberapa besar atribut Pekerjaan Orang Tua berkontribusi dalam memisahkan data dalam proses klasifikasi, di mana semakin tinggi nilai *gain*, semakin efektif atribut tersebut dalam menentukan keputusan klasifikasi.

3.3.2 Atribut SO

Atribut Penghasilan Orang Tua (SO) digunakan untuk mengukur tingkat kesejahteraan ekonomi keluarga. Penghasilan orang tua dikelompokkan dalam beberapa tingkatan, seperti rendah, menengah, dan tinggi, tingkat penghasilan ini memengaruhi kualifikasi siswa untuk menerima bantuan atau program dukungan pendidikan.

Hasil Perhitungan *Gain* Kategori SO
 $Gain = 0,426921262$

Hasil perhitungan *gain* untuk kategori Penghasilan Orang Tua (SO) adalah 0,426921262. Perhitungan *gain* ini dilakukan menggunakan rumus Persamaan 2. Nilai ini menunjukkan seberapa besar kontribusi atribut Penghasilan Orang Tua dalam memisahkan data untuk proses klasifikasi, dengan semakin tinggi nilai *gain*, semakin baik atribut tersebut dalam membantu pengambilan keputusan klasifikasi.

3.3.3 Atribut PKIP

Atribut Penerima Kartu Indonesia Pintar (PKIP) digunakan untuk menentukan apakah seorang siswa terdaftar dalam program bantuan pendidikan dari pemerintah. Siswa yang menerima Kartu Indonesia Pintar (KIP) biasanya berasal dari keluarga dengan kondisi ekonomi yang kurang mampu. Atribut ini penting dalam proses klasifikasi karena memengaruhi kelayakan siswa untuk menerima bantuan pendidikan lebih lanjut.

Hasil Perhitungan *Gain* Kategori PKIP
 $Gain = 0,543913071$

Hasil perhitungan *gain* untuk kategori Penerima Kartu Indonesia Pintar (PKIP) adalah 0,543913071. Perhitungan *gain* ini dilakukan dengan menggunakan rumus Persamaan 2. Nilai *gain* yang tinggi ini menunjukkan bahwa atribut PKIP berperan penting dalam memisahkan data dan membantu proses klasifikasi dengan lebih tepat.

3.3.4 Atribut PKPS

Atribut Penerima Kartu Perlindungan Sosial (PKPS) digunakan untuk menunjukkan apakah seseorang menerima bantuan sosial dari pemerintah. Kartu ini diberikan kepada keluarga dengan kondisi ekonomi yang kurang mampu sebagai bentuk perlindungan sosial. Atribut ini penting dalam proses klasifikasi karena dapat memengaruhi kelayakan individu untuk menerima bantuan lebih lanjut atau program dukungan lainnya.

Hasil Perhitungan *Gain* Kategori PKPS
 $Gain = 0,300794959$

Hasil perhitungan *gain* untuk kategori Penerima Kartu Perlindungan Sosial (PKPS) adalah 0,300794959. Perhitungan *gain* ini dilakukan dengan menggunakan rumus Persamaan 2. Nilai *gain* ini menunjukkan bahwa atribut PKPS berkontribusi dalam memisahkan data,

namun kontribusinya tidak sebesar atribut lain dalam proses klasifikasi.

Hasil perhitungan *entropy* dan *gain* yang telah dilakukan dapat ditemukan dalam Tabel 8 yang terletak di bawah ini. Data ini dapat dijadikan untuk perhitungan *Node* selanjutnya.

3.4 Perhitungan *Node*

Nilai *entropy* dan *gain* dapat dihitung menggunakan persamaan (1) dan (2). Informasi lebih detail mengenai nilai-nilai tersebut dapat dilihat pada Tabel 5 di bawah ini. Tabel 5 menyajikan data hasil perhitungan *entropy* dan *gain*.

Tabel 5. Perhitungan *Node* 1

Node 1	Jumlah (S)	Ya	Tidak	Entropy	Gain
Total	50	23	27	0,995378	
POT					0,43361
Rendah	29	22	7	0,797326	
Menengah	11	1	11	0,413816	
Tinggi	9	0	9	0	
SO					0,42692
Rendah	29	22	7	0,797326	
Menengah	15	1	14	0,353359	
Tinggi	6	0	6	0	
PKIP					0,54391
Ya	17	17	0	0	
Tidak	33	6	27	0,684038	
PKPS					0,30079
Ya	11	11	0	0	
Tidak	39	12	27	0,890491	

Tabel 5 menunjukkan hasil perhitungan *node* 1, termasuk jumlah total data, nilai *entropy*, dan *gain* untuk setiap atribut. Atribut Pekerjaan Orang Tua (POT) memiliki *gain* tertinggi sebesar 0,43361, yang berarti atribut ini paling signifikan dalam memisahkan data pada *node* 1. Atribut PKPS memiliki *gain* terendah sebesar 0,30079, yang menunjukkan kontribusi pemisahan yang lebih rendah dibandingkan atribut lainnya.

3.4.1 Perhitungan *Node* 1.1

Perhitungan *node* 1.1 menggunakan Persamaan 1 dan 2 untuk menentukan *node* berikutnya, yang dapat dilihat pada Tabel 6. Hasil dari perhitungan ini akan menjadi dasar dalam pengembangan pohon keputusan selanjutnya. Langkah ini penting dalam memastikan pohon keputusan dapat memisahkan data secara optimal untuk proses klasifikasi berikutnya

Tabel 6. Perhitungan *Node* 1.1

Node 1.1	Jumlah (S)	Ya	Tidak	Entropy	Gain
Total (Tidak)	33	6	27	0,684038	
POT					0,17724
Rendah	12	5	7	0,979868	
Menengah	12	1	11	0,413816	
Tinggi	9	0	9	0	
SO					0,16710

Rendah	12	5	7	0,979868	
Menengah	15	1	14	0,353359	
Tinggi	6	0	6	0	
PKPS					0,25767
Ya	3	3	0	0	
Tidak	30	3	27	0,468995	

Tabel 6 menampilkan hasil perhitungan *node* 1.1, termasuk jumlah total data, nilai *entropy*, dan *gain* untuk setiap atribut. Atribut PKPS memiliki nilai *gain* tertinggi sebesar 0,25767, yang menunjukkan bahwa atribut ini paling efektif dalam memisahkan data pada *node* 1.1. Atribut Pekerjaan Orang Tua (POT) memiliki nilai *gain* sebesar 0,17724, yang menandakan kontribusinya dalam memisahkan data lebih rendah dibandingkan PKPS.

3.4.2 Perhitungan *Node* 1.1.1

Perhitungan *node* 1.1.1 menggunakan Persamaan 1 dan 2 untuk menentukan nilai *entropy* dan *gain* dari setiap atribut. Proses ini dilakukan guna menentukan *node* selanjutnya dalam pembentukan pohon keputusan. Hasil perhitungan untuk *node* 1.1.1 dapat dilihat secara rinci pada Tabel 7.

Tabel 7. Perhitungan *Node* 1.1.1

Node 1.1.1	Jumlah (S)	Ya	Tidak	Entropy	Gain
Total (Tidak)	30	3	27	0,468995	
POT					0,07420
Rendah	9	2	7	0,764204	
Menengah	12	1	11	0,413816	
Tinggi	9	0	9	0	
SO					0,06305
Rendah	9	2	7	0,764204	
Menengah	15	1	14	0,353359	
Tinggi	6	0	6	0	

Tabel 7 memperlihatkan hasil perhitungan *node* 1.1.1, dengan total jumlah data, nilai *entropy*, dan *gain* untuk masing-masing atribut. Atribut Pekerjaan Orang Tua (POT) memiliki nilai *gain* tertinggi sebesar 0,07420, yang berarti atribut ini paling berperan dalam memisahkan data pada *node* ini. Sementara itu, atribut Penghasilan Orang Tua (SO) memiliki nilai *gain* sebesar 0,06305, yang menunjukkan kontribusi pemisahan yang lebih rendah dibandingkan POT.

3.4.3 Perhitungan *Node* 1.1.1.1

Perhitungan *node* 1.1.1.1 dilakukan menggunakan Persamaan 1 dan 2 untuk menentukan *node* selanjutnya. Hasil perhitungan ini dapat dilihat pada Tabel 7. Proses ini penting dalam pembentukan pohon keputusan untuk memisahkan data secara lebih baik dan akurat.

Tabel 7. Perhitungan *Node* 1.1.1.1

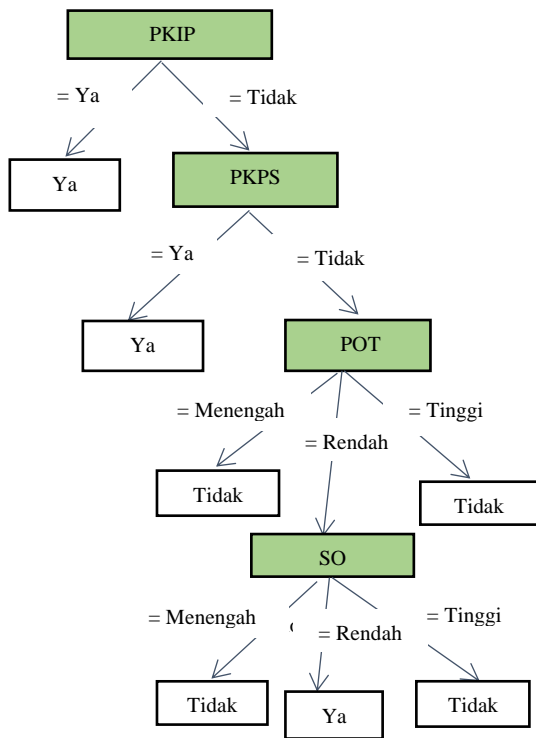
Node 1.1.1	Jumlah (S)	Ya	Tidak	Entropy	Gain
Total (Rendah)	9	2	7	0,764204	
SO					0,45810
Rendah	3	2	1	0,918295	

Menengah	4	0	4	0
Tidak	2	0	2	0

Tabel 7 menunjukkan hasil perhitungan *node* 1.1.1.1, termasuk jumlah data, nilai *entropy*, dan *gain* untuk atribut Penghasilan Orang Tua (SO). Atribut SO memiliki nilai *gain* sebesar 0,45810, yang berarti atribut ini memberikan kontribusi signifikan dalam memisahkan data. Selain itu, nilai *entropy* total adalah 0,764204, yang menandakan adanya variasi dalam data pada kategori ini.

3.5 Hasil Pohon Keputusan

Pohon keputusan merupakan hasil akhir dari kinerja metode C45. Hasil tersebut digambarkan berdasarkan pembahasan sebelumnya. Adapun hasil keluaran pohon keputusan dapat disajikan pada Gambar 2.

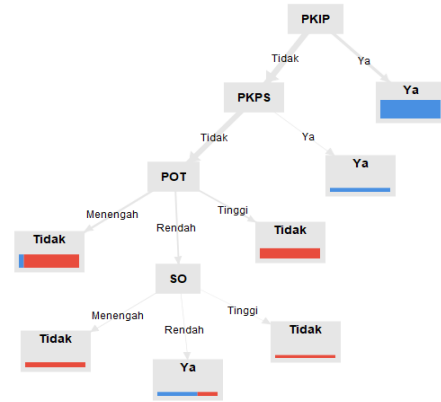


Gambar 2. Pohon Keputusan

Berdasarkan Gambar 2 di atas, aturan yang dihasilkan adalah sebagai berikut. Jika seorang siswa memiliki PKIP, maka siswa tersebut diprediksi layak menerima bantuan. Jika siswa tidak memiliki PKIP tetapi memiliki PKPS, maka siswa tersebut juga diprediksi layak menerima bantuan. Namun, jika siswa tidak memiliki PKIP dan PKPS, maka keputusan akan tergantung pada Pekerjaan Orang Tua (POT) dan Penghasilan Orang Tua (SO). Jika pekerjaan orang tua adalah rendah dan penghasilan orang tua juga rendah, maka siswa tersebut diprediksi layak menerima bantuan. Sebaliknya, jika pekerjaan orang tua menengah atau tinggi, atau penghasilan orang tua tinggi, maka siswa diprediksi tidak layak menerima bantuan.

3.6 Hasil Pengujian Dengan Aplikasi RapidMiner

Setelah dilakukan input Data *Training* ke dalam aplikasi RapidMiner sesuai dengan langkah-langkah yang telah ditentukan, maka diperoleh hasil pohon keputusan seperti yang ditunjukkan pada Gambar 3. Pohon keputusan ini menggambarkan struktur klasifikasi yang dihasilkan berdasarkan data yang dimasukkan dan aturan yang diterapkan. Dengan visualisasi ini, dapat diidentifikasi pola-pola penting yang digunakan untuk memprediksi kelayakan penerima beasiswa.



Gambar 3. Hasil Pohon Keputusan RapidMiner

Gambar 3 menunjukkan hasil kinerja metode yang divisualisasikan dalam bentuk pohon keputusan. Berdasarkan pohon keputusan yang telah dibentuk, aturan atau *rule* dapat diturunkan dari setiap cabang pohon tersebut. Aturan ini membantu dalam memahami bagaimana setiap variabel berkontribusi terhadap keputusan akhir, sehingga mempermudah proses klasifikasi. Selain itu, *rule* yang dihasilkan juga dapat diimplementasikan untuk memprediksi data baru dan memastikan bahwa keputusan yang diambil konsisten dengan pola yang teridentifikasi.

- IF* Penerima Kartu Indonesia Pintar (PKIP) = Ya, *THEN* Prediksi = Ya.
- IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Ya, *THEN* Prediksi = Ya.
- IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Tidak, *AND* Pekerjaan Orang Tua (POT) = Tinggi, *THEN* Prediksi = Tidak.
- IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Tidak, *AND* Pekerjaan Orang Tua (POT) = Rendah, *AND* Penghasilan Orang Tua (SO) = Tinggi, *THEN* Prediksi = Tidak.
- IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Tidak, *AND* Pekerjaan Orang Tua (POT) = Menengah, *THEN* Prediksi = Tidak.
- IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Tidak, *AND* Pekerjaan Orang Tua (POT) = Rendah, *AND* Penghasilan Orang Tua (SO) = Menengah, *THEN* Prediksi = Tidak.

- g. *IF* Penerima Kartu Indonesia Pintar (PKIP) = Tidak, *AND* Penerima Kartu Perlindungan Sosial (PKPS) = Tidak, *AND* Pekerjaan Orang Tua (POT) = Rendah, *AND* Penghasilan Orang Tua (SO) = Rendah, *THEN* Prediksi = Ya.

Berdasarkan aturan yang terbentuk, terdapat 7 aturan yang berhasil diidentifikasi dari data penerima beasiswa Program Indonesia Pintar. Aturan-aturan ini menggambarkan pola dan pengetahuan yang diekstraksi dari data yang digunakan dalam klasifikasi. Informasi tersebut membantu dalam memahami faktor-faktor yang memengaruhi kelayakan penerimaan beasiswa berdasarkan berbagai atribut yang dianalisis.

4. Kesimpulan

Berdasarkan implementasi Algoritma C4.5 untuk prediksi penerima beasiswa Program Indonesia Pintar di lingkungan pendidikan terkait. Hasil penelitian menunjukkan bahwa model klasifikasi yang dihasilkan memiliki tingkat akurasi yang sangat tinggi, yaitu sebesar 96.00%. Selain itu, *precision* dan *recall* untuk kategori penerima beasiswa masing-masing mencapai 95.65%, sementara *precision* dan *recall* untuk kategori bukan penerima beasiswa tercatat sebesar 96.30%. Hasil tersebut didapat berdasarkan data penerima beasiswa sebanyak 50 data sampel yang menghasilkan 7 *rule*. Proses kinerja metode dilakukan dengan pengujian menggunakan aplikasi RapidMiner, yang menghasilkan aturan untuk memprediksi penerima beasiswa Program Indonesia Pintar. Hasil ini dapat digunakan sebagai acuan untuk memprediksi penerima beasiswa secara lebih tepat sasaran, sehingga membantu instansi pendidikan terkait dalam menyalurkan beasiswa dengan lebih efektif dan efisien Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi penggunaan teknik atau model lain yang dapat dibandingkan dengan pendekatan ini guna melihat perbedaan dalam performa. Selain itu, penggunaan data yang lebih besar dan beragam akan membantu menguji generalisasi model yang dihasilkan.

Daftar Rujukan

- [1] Aulia, N., Suarna, N., & Prihartono, W. (2024). Klasifikasi Penentuan Penerima Program Indonesia Pintar Di Krwilbidikcam Greded Menggunakan Algoritma C4.5. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3913–3919. <https://doi.org/10.36040/jati.v7i6.8294>
- [2] Chen, X. (2023). Design and research of MOOC teaching system based on TG-C4.5 algorithm. *Systems and Soft Computing*, 5(August), 200064. <https://doi.org/10.1016/j.sasc.2023.200064>
- [3] Gezer, M., Selc, S., Demirkol, D., Kas, M., Bucak, Y., Erol, C., & Yag y z, A. (2022). Evaluasi Faktor Risiko Yang Berhubungan Dengan Keberhasilan Pengobatan Antihipertensi Menggunakan Teknik Penambangan Data Nermin Yasemin Perkenalan Metode. 0, 1–11. <https://doi.org/10.1177/10742484221136758>
- [4] Husain, S., Rahmadanti, R., Husain, E. S., & ... (2023). Kepuasan terhadap Tata Kelola Penerimaan Kartu Indonesia Pintar Kuliah di Perguruan Tinggi. *Jurnal Pendidikan ...*, 7, 1965–1973. <https://jptam.org/index.php/jptam/article/download/6112/5117>
- [5] Isra, M. (2022). Behavior Analysis and Prediction of Civil Services Staff in Occupational Functional Positions Using C4.5 Algorithm. *Jurnal Informasi Dan Teknologi*, 4(Eselon V), 58–63. <https://doi.org/10.37034/jidt.v4i1.186>
- [6] Kadja, M. T. G., Rumlaklak, N. D., & Djahi, B. S. (2023). Penerapan Metode Fuzzy C-Means Dalam Penentuan Penerima Beasiswa Program Indonesia Pintar (Pip). *Jurnal Komputer Dan Informatika*, 11(1), 37–43. <https://doi.org/10.35508/jicon.v11i1.9846>
- [7] Li, P., Xiong, F., Huang, X., & Wen, X. (2024). Construction and optimization of vending machine decision support system based on improved C4.5 decision tree. *Heliyon*, 10(3), e25024. <https://doi.org/10.1016/j.heliyon.2024.e25024>
- [8] Metode, P., Bayes, N., Penjualan, M., Ban, P., Aprihartha, M. A., Astutik, F., & Sulistianingsih, N. (2024). Comparison of Naive Bayes , CART , dan CART Adaboost Methods in Predicting Tire Product Sales. 20(3), 596–605. <https://doi.org/10.20956/j.v20i3.33187>
- [9] Oktafiqurahman, A., Kusri, K., & Nasiri, A. (2023). Implementasi Algoritma C4.5 Untuk Tes Kepribadian Penerimaan Karyawan Di Dinas Perhubungan Provinsi Sulawesi Tengah. *Jurnal Teknologi Informasi Dan Komunikasi (TIKOMSiN)*, 11(1), 11. <https://doi.org/10.30646/tikomsin.v11i1.719>
- [10] Simanullang, R., Hartama, D., Poningsih, P., Parlina, I., & Lubis, M. R. (2022). Model Aturan dalam Menentukan Prestasi Nilai Siswa di SMK GKPS 1 Raya Menggunakan Algoritma C4.5. *Jurnal Ilmu Komputer Dan Informatika*, 2(1), 1–24. <https://doi.org/10.54082/jiki.16>
- [11] Sipayung, S. P., Sihaloho, T. P., Purba, A., Tarigan, R., Letjend, J., Ginting, D., Bulan, P., Baru, M., Medan, K., & Indonesia, S. U. (2023). Analisa Algoritma C . 45 Terhadap Penentuan Rekomendasi Penerima Beasiswa SMP Swasta Methodist-8 Medan. 3(1), 20–24. <https://doi.org/10.58918/lofian.v3i1.217>
- [12] Surya Maulana, A., Nazir, A., Handayani, L., & Afrianty, I. (2023). KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma Fuzzy C-Means untuk Melihat Pola Penerima Beasiswa Bank Indonesia. *Media Online*, 3(6), 670–679. <https://doi.org/10.30865/klik.v3i6.788>
- [13] Ubaedi, I., & Djaksana, Y. M. (2022). Optimasi Algoritma C4.5 Menggunakan Metode Forward Selection Dan Stratified Sampling Untuk Prediksi Kelayakan Kredit. *JSii (Jurnal Sistem Informasi)*, 9(1), 17–26. <https://doi.org/10.30656/jsii.v9i1.3505>
- [14] Wang, R., Xin, T., Jia, S., Ren, D., & Li, M. (2024). Production line balance problem identification and improvement based on decision tree: A case study of commercial air conditioner production line. *Science Progress*, 107(1), 1–30. <https://doi.org/10.1177/00368504241238612>
- [15] Yanasari, C., & Arifin, T. (2023). Implementasi Algoritma K-Nearest Neighbor Untuk Klasifikasi Penerimaan Beasiswa Program Indonesia Pintar. *Jurnal Sistem Informasi Dan Ilmu Komputer*, 1(4), 178–194. <https://doi.org/10.59581/jusiik-widyakarya.v1i4.1862>