



Jurnal Review Pendidikan dan Pengajaran
<http://journal.universitaspahlawan.ac.id/index.php/jrpp>
 Volume 8 Nomor 4, 2025
 P-2655-710X e-ISSN 2655-6022

Submitted : 29/11/2025
 Reviewed : 11/12/2025
 Accepted : 19/12/2025
 Published : 29/12/2025

Kristin Lourensi
 Sitompul^{1*}
 Petti Indrayati Sijabat²
 Puteri Fajar Addini³
 Rossya Diva Anwar⁴
 Icha Nurhidayah⁵

IDENTIFIKASI FAKTOR DOMINAN YANG MEMPENGARUHI INDEKS PRESTASI KUMULATIF MENGGUNAKAN ALGORITMA CLASSIFICATION AND REGRESSION TREES

Abstrak

Prestasi akademik mahasiswa yang diukur melalui Indeks Prestasi Kumulatif (IPK) merupakan salah satu indikator utama keberhasilan proses pembelajaran di perguruan tinggi. Data akademik mahasiswa tersedia dalam jumlah besar namun belum dimanfaatkan secara optimal untuk menggali pola-pola tersembunyi yang mempengaruhi keberhasilan studi. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor dominan yang mempengaruhi capaian IPK mahasiswa serta membangun model prediksi risiko akademik menggunakan teknik Data Mining. Metode yang digunakan adalah Decision Tree (Pohon Keputusan) dengan algoritma Classification and Regression Trees (CART). Data penelitian dikumpulkan melalui kuesioner terhadap 50 responden mahasiswa Teknik Informatika dengan variabel yang mencakup faktor internal (motivasi, gaya belajar) dan eksternal (lingkungan, ekonomi). Proses analisis dilakukan dengan menghitung nilai Gini Impurity dan Gini Gain untuk menentukan atribut pemecah terbaik (splitting attribute). Hasil penelitian menunjukkan bahwa algoritma CART mampu mengklasifikasikan profil mahasiswa ke dalam kategori risiko akademik. Ditemukan bahwa faktor kesesuaian jurusan memiliki pengaruh signifikan dalam menentukan peluang mahasiswa memperoleh IPK di atas rata-rata, sementara faktor motivasi menunjukkan tingkat redundansi akibat variasi data yang rendah. Penelitian ini menyimpulkan bahwa metode CART efektif digunakan sebagai alat bantu pengambilan keputusan bagi institusi dalam merancang strategi bimbingan akademik yang lebih tepat sasaran.

Kata Kunci: Data Mining, Decision Tree, Algoritma CART, Indeks Prestasi Kumulatif, Prediksi Akademik.

Abstract

Student academic achievement, as measured by the Cumulative Grade Point Average (GPA), is one of the main indicators of the success of the learning process in higher education. A large amount of student academic data is available but has not been optimally utilized to uncover hidden patterns that influence academic success. This study aims to identify the dominant factors influencing student GPA achievement and to develop an academic risk prediction model using data mining techniques. The method used is a Decision Tree with the Classification and Regression Trees (CART) algorithm. Data were collected through questionnaires from 50 Informatics Engineering students, with variables covering internal (motivation, learning style) and external (environmental, economic) factors. The analysis process was carried out by calculating the Gini Impurity and Gini Gain values to determine the best splitting attribute. The results show that the CART algorithm is capable of classifying student profiles into academic risk categories. It was found that the major suitability factor had a significant influence on students' chances of achieving an above-average GPA, while the motivation factor showed a level of redundancy due to low data variation. This study concludes that the CART method is effective as a decision-making tool for institutions in designing more targeted academic guidance strategies.

^{1,2,3,4,5}) Teknik Informatika, STMIK Pelita Nusantara
 email: Sriwijaya11121987@gmail.com

Keywords: Data Mining, Decision Tree, CART Algorithm, Grade Point Average, Academic Prediction.

PENDAHULUAN

Kemajuan teknologi informasi yang pesat telah mendorong transformasi signifikan di berbagai sektor kehidupan, termasuk dalam dunia pendidikan tinggi. Perguruan tinggi tidak hanya berperan sebagai institusi penyelenggara pendidikan, tetapi juga sebagai pengelola berbagai data akademik yang kompleks. Data tersebut mencakup aktivitas mahasiswa mulai dari kehadiran, nilai akademik, hingga pola perilaku belajar yang semuanya tersimpan dalam sistem informasi kampus.

Namun demikian, banyak data tersebut hanya digunakan sebatas kebutuhan administratif tanpa diolah lebih lanjut untuk menghasilkan informasi yang strategis bagi pengambilan keputusan institusi. Kondisi ini membuka peluang untuk menerapkan pendekatan analitik berbasis data guna memperoleh wawasan yang lebih mendalam terhadap perilaku dan performa mahasiswa. Salah satu pendekatan yang relevan adalah data mining, yaitu proses penemuan pola tersembunyi atau pengetahuan baru dari kumpulan data dalam jumlah besar.

Dalam konteks pendidikan, teknik data mining telah digunakan untuk berbagai tujuan, seperti memprediksi prestasi belajar, mengidentifikasi potensi putus studi, serta mengevaluasi faktor internal yang paling memengaruhi keberhasilan mahasiswa. Hal ini didukung oleh penelitian (Rafika et al., 2025) yang memanfaatkan data mining untuk mengevaluasi dan memprediksi kinerja akademik mahasiswa Teknik Informatika agar penanganan akademik dapat dilakukan lebih dini. Selain itu, pentingnya analisis faktor risiko juga dikemukakan oleh (Rolansa et al., 2020), yang menyatakan bahwa identifikasi dini risiko gagal studi melalui data akademik dan non-akademik sangat krusial untuk merancang intervensi yang tepat.

Fenomena tersebut juga ditemukan di lingkungan STMIK Pelita Nusantara Medan, di mana terdapat tantangan dalam memantau fluktuasi prestasi akademik mahasiswa yang berujung pada keterlambatan penyelesaian studi. Kondisi ini dipengaruhi oleh berbagai faktor perilaku dan eksternal, seperti kemampuan manajemen waktu yang lemah, tingkat stres yang tinggi, keterbatasan kondisi keuangan, kurangnya pemanfaatan sumber belajar, serta ketidaksesuaian jurusan dengan minat mahasiswa. Pengaruh signifikan faktor eksternal terhadap keberhasilan studi ini sejalan dengan temuan (Pambudi & Setiawan, 2018), yang menunjukkan bahwa faktor di luar kemampuan internal siswa memiliki dampak besar terhadap penyelesaian studi. Fokus pada faktor-faktor yang bersifat "variabel perilaku" menjadi krusial karena faktor inilah yang paling dinamis dan dapat diintervensi melalui kebijakan bimbingan akademik.

Untuk mencapai akurasi prediksi yang interpretatif, dibutuhkan metode analisis yang mampu menghasilkan model yang mudah dipahami oleh pengambil kebijakan. Metode Decision Tree (Pohon Keputusan) menjadi solusi yang tepat karena mampu mengklasifikasikan data berdasarkan atribut tertentu dalam bentuk struktur yang hierarkis. Di antara berbagai algoritma, CART (Classification and Regression Tree) menjadi salah satu algoritma yang paling unggul karena kemampuannya menghasilkan model klasifikasi biner yang efisien dan memiliki tingkat akurasi tinggi dalam memetakan profil risiko mahasiswa secara kontras. Keunggulan ini dibuktikan dalam penelitian (Rafika et al., 2025) di mana algoritma CART berhasil mencapai akurasi 85% dalam mengklasifikasikan risiko studi mahasiswa. Selain itu, juga menemukan bahwa meskipun algoritma lain seperti C4.5 memiliki kelebihan, CART tetap menjadi metode yang efektif dan valid dalam menyeleksi variabel yang paling signifikan untuk prediksi akademik.

Berdasarkan uraian di atas, penelitian ini dilakukan untuk menerapkan metode Decision Tree menggunakan algoritma CART dalam mengidentifikasi faktor dominan dan memprediksi prestasi mahasiswa di STMIK Pelita Nusantara Medan dengan menyederhanakan variabel ke dalam kategori biner guna mendapatkan wawasan yang lebih tajam. Oleh karena itu, penelitian ini diberi judul: "Identifikasi Faktor Dominan Yang Mempengaruhi Indeks Prestasi Kumulatif Menggunakan Classification And Regression Trees."

1.1 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan algoritma CART dalam mengolah data perilaku dan eksternal

mahasiswa untuk menghasilkan model klasifikasi prestasi akademik ?

2. Variabel perilaku manakah (Manajemen Waktu, Stres, Keuangan, Sumber Belajar, Motivasi, Kesesuaian Jurusan) yang paling dominan menjadi penentu pencapaian IPK mahasiswa berdasarkan skor Gini Gain?
3. Bagaimana bentuk struktur pohon keputusan biner yang dihasilkan dalam memetakan aturan (If-Then rules) terkait potensi prestasi akademik mahasiswa?

1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Membangun model klasifikasi yang mampu memetakan potensi prestasi mahasiswa menggunakan algoritma CART dengan pendekatan atribut biner.
2. Mengidentifikasi dan mengurutkan faktor-faktor perilaku yang paling berpengaruh terhadap capaian prestasi akademik berdasarkan hasil perhitungan Gini Gain.
3. Menghasilkan aturan keputusan (Decision Rules) yang dapat digunakan sebagai acuan bimbingan akademik bagi pihak kampus.

1.3 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan dasar pengambilan keputusan dalam merancang program pembinaan mahasiswa berdasarkan faktor risiko yang teridentifikasi.
2. Menjadi referensi dalam penerapan algoritma CART untuk penelitian data mining pendidikan, khususnya mengenai efektivitas penyederhanaan data biner.
3. Memberikan gambaran mengenai aspek perilaku (seperti manajemen waktu dan stres) yang perlu diperbaiki guna meningkatkan prestasi akademik.

1.4 Batasan Masalah

Agar penelitian ini lebih terarah, maka ruang lingkup dibatasi sebagai berikut:

1. Penelitian difokuskan pada analisis data mahasiswa dengan menggunakan enam variabel perilaku/eksternal: Motivasi Belajar, Manajemen Waktu, Tingkat Stres, Kondisi Keuangan, Pemanfaatan Sumber Belajar, dan Kesesuaian Jurusan.
2. Atribut target adalah klasifikasi IPK yang dibagi menjadi dua kategori: Tinggi (≥ 3.5) dan Rendah (< 3.5).
3. Seluruh atribut independen (input) disederhanakan ke dalam bentuk biner (skala 1 dan skala > 1) untuk mendapatkan hasil klasifikasi yang lebih kontras.
4. Variabel demografis dan administratif seperti Jenis Kelamin, Program Studi, dan Semester tidak diikutsertakan dalam pemodelan utama guna memfokuskan analisis pada variabel perilaku.

METODE

Penelitian ini termasuk dalam kategori penelitian terapan (applied research) karena bertujuan untuk mengimplementasikan metode Decision Tree dengan algoritma Classification and Regression Tree (CART) dalam kasus nyata, yaitu prediksi kegagalan studi mahasiswa. Pendekatan yang digunakan adalah pendekatan kuantitatif, karena penelitian ini melibatkan analisis data numerik untuk menghasilkan model klasifikasi yang dapat diukur secara statistik (Rafika et al., 2025).

Selain itu, penelitian ini juga memanfaatkan konsep data mining sebagai pendekatan analitik. Pendekatan ini digunakan untuk menggali pola dan hubungan tersembunyi dari data akademik dan non-akademik mahasiswa yang dapat membantu dalam pengambilan keputusan akademik.

1.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data primer.

Data primer diperoleh melalui survei dan pengumpulan data internal mahasiswa yang mencakup aspek motivasi belajar, manajemen waktu, tingkat stres, kondisi keuangan, sumber belajar, dan kesesuaian jurusan.

1.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari:

1. Variabel Dependen (Y)
 - Kegagalan Studi Mahasiswa (dinyatakan dalam dua kelas: Berpotensi Gagal Studi dan

Tidak Berpotensi Gagal Studi).

Variabel Independen (X)

- X_1 = IPK terakhir
- X_2 = Motivasi belajar
- X_3 = Manajemen waktu
- X_4 = Tingkat stres
- X_5 = Kondisi keuangan
- X_6 = Sumber belajar
- X_7 = Kesesuaian jurusan

Masing-masing variabel di atas memiliki skala data kuantitatif yang diperoleh melalui hasil survei dan data akademik mahasiswa.

1.3 Metode Pengumpulan Data

Metode pengumpulan data dalam penelitian ini dilakukan melalui beberapa teknik, yaitu:

1. Observasi, untuk mengamati secara langsung kondisi mahasiswa dan faktor-faktor yang dapat memengaruhi keberhasilan studi.
2. Kuesioner, untuk memperoleh data non-akademik seperti motivasi belajar, manajemen waktu, tingkat stres, kondisi keuangan, sumber belajar, dan kesesuaian jurusan.
3. Dokumentasi, untuk memperoleh data akademik mahasiswa seperti IPK terakhir dari sistem akademik kampus.

Data yang diperoleh kemudian diolah menggunakan perangkat lunak pengolah data seperti Python dan Microsoft Excel untuk kebutuhan analisis model CART.

1.4 Metode Analisis Data

Metode analisis data dalam penelitian ini menggunakan pendekatan data mining dengan algoritma Classification and Regression Tree (CART). Tahapan analisis meliputi:

1. Data Preparation (Persiapan Data)

Meliputi pembersihan data, normalisasi, dan penghapusan data yang tidak relevan.

2. Model Building

Proses pembentukan Decision Tree menggunakan algoritma CART berdasarkan atribut dengan nilai Gini Index terkecil.

3. Interpretation

Menarik kesimpulan dari hasil model untuk mengidentifikasi faktor-faktor dominan yang memengaruhi potensi kegagalan studi.

1.5 Tahapan Penelitian

Langkah-langkah dalam penelitian ini dapat dijelaskan sebagai berikut:

1. Identifikasi Masalah

Menentukan permasalahan utama, yaitu belum adanya sistem prediksi kegagalan studi mahasiswa berbasis data.

2. Studi Literatur

Mengumpulkan teori dan referensi terkait algoritma CART, data mining, dan faktor-faktor yang memengaruhi kegagalan studi.

3. Pengumpulan Data

Melakukan pengumpulan data akademik dan non-akademik mahasiswa.

4. Pengolahan Data

Melakukan pembersihan, transformasi, dan pemilihan variabel yang relevan.

5. Penerapan Algoritma CART

Membangun model Decision Tree untuk klasifikasi potensi kegagalan studi.

6. Evaluasi Model

Mengukur akurasi hasil prediksi dan mengidentifikasi variabel paling berpengaruh.

7. Kesimpulan dan Rekomendasi

Menyimpulkan hasil penelitian serta memberikan rekomendasi bagi pihak kampus.

HASIL DAN PEMBAHASAN

Berdasarkan data penelitian terhadap 50 responden mahasiswa, terdapat sebaran nilai IPK

yang menunjukkan bahwa mayoritas mahasiswa (44 orang atau 88%) memiliki $IPK \geq 3.5$, sementara sebagian kecil (6 orang atau 12%) memiliki $IPK < 3.5$.

Ada pun keseluruhan data yang akan diolah menggunakan CART ini adalah sebagai berikut.

Table 1. Dataset Mahasiswa untuk Klasifikasi

No	Prodi	Sem	IPK	Waktu	Motivasi	Stres	Keuangan	Sumber	Jurusan
1.	1	7	3.80	2	2	2	2	1	1
2.	1	7	3.85	3	3	2	2	2	2
3.	1	7	3.85	2	3	2	2	2	3
4.	1	7	3.70	2	2	2	2	2	2
5.	2	5	3.70	3	2	2	2	2	2
6.	1	7	3.57	3	3	3	2	2	3
7.	1	7	3.64	2	2	2	2	2	2
8.	1	7	3.86	2	2	2	2	1	2
9.	1	7	3.86	3	3	2	3	3	2
10.	1	7	3.43	2	2	3	2	2	2
50.	2	3	4.00	3	3	3	3	3	3

Kondisi awal ini menghasilkan nilai Gini Impurity sebesar 0.2112, yang menjadi dasar bagi algoritma CART untuk melakukan pembagian data (splitting). Langkah-langkah dalam pengerjaan menggunakan Algoritma CART adalah sebagai berikut.

Langkah 1: Hitung Gini Awal (Root Node)

Tahap pertama dalam algoritma CART adalah menganalisis kondisi awal seluruh data penelitian sebelum dilakukan pembagian (splitting)

- **IPK Tinggi (1):** 44 orang
- **IPK Rendah (0):** 6 orang
- **Total (n):** 50 orang

$$Gini = 1 - \sum(p^2)$$

$$Gini_{Total} = 1 - \left[\left(\frac{44}{50} \right)^2 + \left(\frac{6}{50} \right)^2 \right]$$

$$Gini_{Total} = 1 - [0.7744 + 0.0144] = 0.2112$$

Langkah 2: Uji Coba Split (Contoh: Faktor "Waktu")

Setelah menentukan nilai Gini Index awal, langkah selanjutnya dalam algoritma CART adalah mengevaluasi setiap variabel independen untuk menentukan kemampuannya dalam mereduksi ketidakmurnian data. Sebagai ilustrasi proses teknis, dilakukan pengujian terhadap variabel Manajemen Waktu.

- **Waktu = 1:** 1 Rendah, 1 Tinggi (Total 2)
- **Waktu = 2 dan 3:** 5 Rendah, 43 Tinggi (Total 48) Kita hitung Gini untuk masing-masing cabang:

1. Cabang Kiri (Waktu = 1):

$$Gini_{Kiri} = 1 - \left[\left(\frac{1}{2} \right)^2 + \left(\frac{1}{2} \right)^2 \right] = 1 - [0.25 + 0.25] = 0.500$$

2. Cabang Kanan (Waktu = 2 atau 3):

$$Gini_{Kanan} = 1 - \left[\left(\frac{5}{48} \right)^2 + \left(\frac{43}{48} \right)^2 \right] = 1 - [0.0108 + 0.8034] = 0.1858$$

Langkah 3: Hitung Gini Akhir (Weighted Gini)

Setelah dibagi, kita hitung rata-rata bobot Gini dari kedua cabang tersebut:

$$Gini_{Split} = \left(\frac{n_{kiri}}{n_{total}} \times Gini_{kiri} \right) + \left(\frac{n_{kanan}}{n_{total}} \times Gini_{kanan} \right)$$

$$Gini_{Split} = \left(\frac{2}{50} \times 0.500 \right) + \left(\frac{48}{50} \times 0.1858 \right)$$

$$Gini_{Split} = 0.02 + 0.1783 = 0.1983$$

Langkah 4: Hitung Gini Gain (Skor Kepentingan)

Tahap akhir dalam evaluasi variabel pada algoritma CART adalah menghitung nilai Gini Gain. Nilai ini merepresentasikan efektivitas atau besarnya informasi yang diperoleh setelah data awal dipecah berdasarkan variabel tertentu.

$$Gain = Gini_{Total} - Gini_{Split}$$

$$Gain = 0.2112 - 0.1983 = \mathbf{0.0129}$$

Selanjutnya, nilai Gain dari faktor-faktor lainnya dirangkum ke dalam tabel berikut. Table 2. Hasil Perhitungan Nilai Gini Split dan Gini Gain Tiap Faktor

Faktor	Cabang Kiri (≤ 1)	Cabang Kanan (> 1)	Gini Split	Gain
Waktu	1 Tinggi, 1 Rendah	43 Tinggi, 5 Rendah	0.1983	0.0129
Faktor	Cabang Kiri (≤ 1)	Cabang Kanan (> 1)	Gini Split	Gain
Kesesuaian Jurusan	5 Tinggi, 0 Rendah	39 Tinggi, 6 Rendah	0.2080	0.0032
Stres	3 Tinggi, 1 Rendah	41 Tinggi, 5 Rendah	0.2083	0.0029
Sumber	3 Tinggi, 0 Rendah	41 Tinggi, 6 Rendah	0.2094	0.0018
Kuangan	2 Tinggi, 0 Rendah	42 Tinggi, 6 Rendah	0.2100	0.0012
Motivasi	1 Tinggi, 0 Rendah	43 Tinggi, 6 Rendah	0.2106	0.0006
Prodi	35 Tinggi, 6 Rendah	9 Tinggi, 0 Rendah	0.2049	0.0063

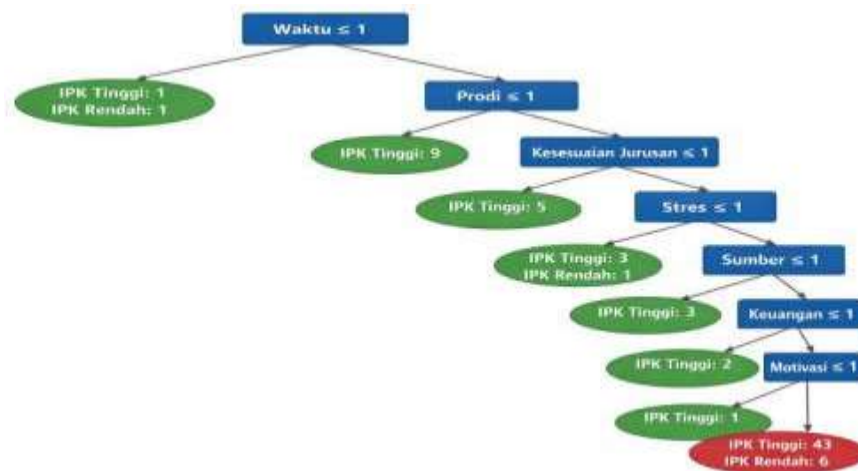
4.2 Penentuan Faktor Paling Berpengaruh (Feature Importance)

Penelitian ini menggunakan kriteria biner (nilai ≤ 1 dan > 1) untuk menguji kekuatan setiap variabel dalam memprediksi kategori IPK. Berdasarkan perhitungan Gini Gain, didapatkan urutan faktor dari yang paling berpengaruh sebagai berikut:

1. Manajemen Waktu (Gini Gain: 0.0129): Menjadi faktor utama (Akar).
2. Program Studi (Gini Gain: 0.0063): Menjadi penentu kedua.
3. Kesesuaian Jurusan (Gini Gain: 0.0032): Menjadi penentu ketiga.
4. Tingkat Stres (Gini Gain: 0.0029): Menjadi penentu keempat.

Faktor lain seperti Sumber Belajar, Keuangan, dan Motivasi memiliki nilai gain yang sangat rendah (≤ 0.0018). Hal ini menunjukkan bahwa setelah variabel Jenis Kelamin dan Semester dihilangkan, manajemen waktu menjadi "pintu masuk" utama dalam membedakan prestasi akademik mahasiswa.

Pohon keputusan yang terbentuk melalui proses CART ini dapat dilihat sebagai berikut.



Gambar 1. Visualisasi Struktur Pohon Keputusan

Aturan IF-THEN dari Pohon yang sudah dibuat

1. **IF Waktu ≤ 1 THEN**
IPK Tinggi = 1, IPK Rendah = 1
2. **IF Waktu > 1 AND Prodi ≤ 1 THEN**
IPK Tinggi = 9
3. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan ≤ 1 THEN**
IPK Tinggi = 5
4. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan > 1 AND Stres ≤ 1 THEN**
IPK Tinggi = 3, IPK Rendah = 1
5. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan > 1 AND Stres > 1 AND Sumber ≤ 1 THEN**
IPK Tinggi = 3
6. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan > 1 AND Stres > 1 AND Sumber > 1 AND Keuangan ≤ 1 THEN**
IPK Tinggi = 2
7. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan > 1 AND Stres > 1 AND Sumber > 1 AND Keuangan > 1 AND Motivasi ≤ 1 THEN**
IPK Tinggi = 1
8. **IF Waktu > 1 AND Prodi > 1 AND Kesesuaian Jurusan > 1 AND Stres > 1 AND Sumber > 1 AND Keuangan > 1 AND Motivasi > 1 THEN**
IPK Tinggi = 43, IPK Rendah = 6

4.3 Pembahasan Logika Pohon Keputusan (CART)

Struktur pohon keputusan yang dihasilkan memberikan gambaran nyata mengenai pola perilaku akademik mahasiswa sebagai berikut:

1. Dominasi Manajemen Waktu sebagai Akar (Root Node)

Algoritma memilih Manajemen Waktu sebagai pemisah pertama. Secara manual, dapat dilihat bahwa pada kelompok mahasiswa dengan manajemen waktu rendah (≤ 1), tingkat ketidakpastian (Gini) sangat tinggi (0.500). Namun, variabel ini berhasil mengisolasi kelompok mayoritas (48 orang) dengan manajemen waktu yang baik, di mana 43 di antaranya memiliki IPK tinggi.

2. Pengaruh Program Studi dan Kesesuaian Jurusan

Setelah melewati penyaringan manajemen waktu, Program Studi muncul sebagai faktor pendukung. Mahasiswa pada program studi tertentu cenderung memiliki konsistensi nilai yang lebih stabil. Hal ini diperkuat dengan faktor Kesesuaian Jurusan, di mana mahasiswa yang merasa jurusannya sangat sesuai (skala > 1) memiliki risiko lebih rendah untuk mendapatkan IPK di bawah 3.5.

3. Redundansi Faktor Motivasi dan Sumber Belajar

Hal menarik ditemukan pada faktor Motivasi. Meskipun secara teori motivasi sangat penting, dalam data ini Gini Gain-nya hanya 0.0006 (terendah). Hal ini disebabkan oleh sebaran data yang terlalu seragam, di mana hampir seluruh responden (49 dari 50) mengisi skor motivasi > 1 ,

sehingga algoritma tidak dapat menggunakan variabel ini untuk membedakan antara mahasiswa IPK tinggi dan rendah.

SIMPULAN

Berdasarkan hasil implementasi algoritma Classification and Regression Trees (CART) untuk mengidentifikasi faktor dominan yang mempengaruhi Indeks Prestasi Kumulatif (IPK) mahasiswa, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Efektivitas Algoritma CART Algoritma CART terbukti mampu memetakan profil risiko akademik mahasiswa STMIK Pelita Nusantara Medan dengan baik. Dengan menggunakan pendekatan atribut biner (skala ≤ 1 dan > 1), model ini berhasil membentuk pohon keputusan yang memisahkan mahasiswa ke dalam kategori risiko IPK Tinggi dan IPK Rendah secara hierarkis.
2. Faktor Paling Dominan Berdasarkan perhitungan nilai Gini Gain, faktor perilaku yang paling dominan dalam menentukan keberhasilan akademik mahasiswa adalah Manajemen Waktu dengan nilai Gini Gain tertinggi sebesar 0.0129. Faktor ini menjadi root node (akar) dalam pohon keputusan, yang artinya kemampuan mengatur waktu menjadi indikator awal yang paling krusial dalam membedakan mahasiswa berprestasi tinggi dan rendah.
3. Hierarki Faktor Pendukung Setelah manajemen waktu, faktor berikutnya yang berpengaruh secara berurutan adalah Program Studi (Gain: 0.0063) dan Kesesuaian Jurusan (Gain: 0.0032). Hal ini menunjukkan bahwa selain perilaku disiplin, kenyamanan mahasiswa terhadap lingkungan akademik dan pilihan jurusannya memegang peranan penting dalam pencapaian IPK di atas 3.5.
4. Faktor motivasi belajar teridentifikasi sebagai variabel dengan pengaruh paling rendah, ditunjukkan oleh nilai Gain terendah sebesar 0,0006. Temuan ini tidak menunjukkan bahwa motivasi belajar bersifat tidak penting, melainkan disebabkan oleh rendahnya variasi data, di mana hampir seluruh responden (98%) melaporkan tingkat motivasi yang tinggi. Kondisi tersebut menyebabkan algoritma tidak mampu memanfaatkan variabel ini sebagai pembeda (discriminator) yang efektif.

DAFTAR PUSTAKA

- Algoritma, M., Dan, C., & Algoritma, C. (n.d.). UNSIKA. 13(3).
- Alvian Setiono, S., & Purwanto, E. (2025). Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree. Prosiding Seminar Nasional Teknologi Informasi Dan Bisnis, 401–406. <https://doi.org/10.47701/4q3z9j41>
- Indah Prabawati, N., Widodo, & Ajie, H. (2019). Kinerja Algoritma Classification And Regression Tree (Cart) dalam Mengklasifikasikan Lama Masa Studi Mahasiswa yang Mengikuti Organisasi di Universitas Negeri Jakarta. PINTER : Jurnal Pendidikan Teknik Informatika Dan Komputer, 3(2), 139–145. <https://doi.org/10.21009/pinter.3.2.9>
- Kasus, S., Alumni, D., Teknik, F., & Bengkulu, U. (2022). Analisis Prediksi Masa Studi Mahasiswa Menggunakan Metode Decision Tree Dengan Penerapan Algoritme Cart (Classification And Regression Trees). 10(1).
- Kpt, E. (2024). Abstrak. 16(2), 334–339.
- Muriyatmoko, D., Musthafa, A., & Wijaya, M. H. (2024). Klasifikasi Profil Kelulusan Nilai AKPAM Dengan Metode Decision Tree. Seminar Nasional Sains Dan Teknologi 2024 Fakultas, April, 448–453.
- Pambudi, R. H., & Setiawan, B. D. (2018). Penerapan Algoritma C4 . 5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal. 2(7), 2637– 2643.
- Rafika, C. S., Reynanada, R. M., & Sari, A. P. (2025). Penerapan Decision Tree CART untuk Klasifikasi Risiko Gagal Studi Mahasiswa. 4, 37–46.
- Rolansa, F., Yunita, Y., Suheri, S., Studi, P., Informatika, T., & Negeri, P. (2020). Jurnal Pendidikan Informatika dan Sains. 9(1),