

Penerapan Data Mining untuk Prediksi Evaluasi Mobil dengan Metode *Decision Tree*

Ahmad Ferdi Dwi Nugroho^{1*}, Hasbi Firmansyah², Wahyu Asriyani³

^{1,2,3}Universitas Pancasakti Tegal, Jl. Halmahera No.KM. 01, Kec. Tegal Tim., Kota Tegal, Jawa Tengah.
dimasfer517@gmail.com

Abstract

The evaluation of motor vehicle acceptability is a multicriteria classification problem involving a complex set of qualitative attributes. This research aims to apply Data Mining techniques using the C4.5 Decision Tree Algorithm to predict the car evaluation outcome (Car Evaluation Dataset), categorized into unacceptable, acceptable, good, and very good classes. The classification model was constructed based on six categorical input attributes and implemented using RapidMiner Studio software with a 10-Fold Cross Validation scheme. The main objectives of this study were to measure the model's accuracy and to identify the priority sequence of attributes most influential in determining the car evaluation class. The test results show that the C4.5 model achieved an accuracy level of 92.94%. Furthermore, the model identified the Safety attribute as the most dominant factor affecting the evaluation outcome, followed by Person. These findings validate the effectiveness of the Decision Tree Algorithm in providing a predictive and interpretive solution for complex multicriteria scoring systems.

Keywords: Data Mining, Clasificasion, Decision Tree C4.5, Car Evaluation, RapidMiner.

Abstrak

Evaluasi kelayakan suatu kendaraan bermotor merupakan masalah klasifikasi multikriteria yang melibatkan serangkaian atribut kualitatif dan kompleks. Penelitian ini bertujuan menerapkan teknik Data Mining menggunakan Algoritma Decision Tree C4.5 untuk memprediksi hasil evaluasi mobil (*Car Evaluation Dataset*) yang dikategorikan ke dalam kelas *unacceptable*, *acceptable*, *good*, dan *very good*. Model klasifikasi dibangun berdasarkan enam atribut input kategorikal dan diimplementasikan melalui *software* RapidMiner Studio dengan skema validasi silang (Cross Validation) 10-Fold. Tujuan utama penelitian adalah mengukur tingkat akurasi model serta mengidentifikasi urutan prioritas atribut yang paling berpengaruh dalam penentuan kelas evaluasi mobil. Hasil pengujian menunjukkan bahwa model C4.5 mencapai tingkat akurasi sebesar 92.94%. Selain itu, model mengidentifikasi atribut Safety sebagai faktor paling dominan yang memengaruhi hasil evaluasi, diikuti oleh person. Hasil ini memvalidasi efektivitas Algoritma Decision Tree dalam memberikan solusi prediktif dan interpretatif untuk sistem penilaian multikriteria yang kompleks.

Kata Kunci: Data Mining, Klasifikasi, Decision Tree C4.5, Car Evaluation, RapidMiner.

Copyright (c) 2025 Ahmad Ferdi Dwi Nugroho, Hasbi Firmansyah, Wahyu Asriyani

✉Corresponding author: Ahmad Ferdi Dwi Nugroho

Email Address: dimasfer517@gmail.com (Jl. Halmahera No.KM. 01, Kec. Tegal Tim., Kota Tegal, Jawa Tengah)

Received 05 December 2025, Accepted 11 December 2025, Published 17 December 2025

PENDAHULUAN

Keputusan untuk menilai kelayakan suatu kendaraan bermotor merupakan persoalan Multicriteria Decision Making (MCDM) yang sarat dengan kriteria kualitatif dan terstruktur. Penilaian ini mencakup berbagai aspek, mulai dari faktor finansial (*buying, maint*), fungsionalitas (*doors, persons, lug boot*), hingga faktor krusial seperti keselamatan (*safety*) (Zupan et al., 1997). Mengingat kompleksitas dan sifat ordinal dari data tersebut, proses evaluasi manual rentan terhadap inkonsistensi dan bias subjektif. Dalam konteks perkembangan pesat era Industri 4.0, kebutuhan akan sistem evaluasi yang objektif dan berbasis data menjadi fundamental, khususnya dalam pengambilan keputusan konsumen dan industri otomotif (Xie et al., 2022).

Untuk mengatasi kerumitan penilaian multikriteria ini, Data Mining dan pembelajaran mesin menawarkan solusi analitis yang canggih (Han et al., 2011). Penerapan teknik klasifikasi memungkinkan pembangunan model prediktif yang dapat mengotomatisasi proses evaluasi, secara empiris menyingkap hubungan antar variabel, dan mengidentifikasi faktor penentu utama yang memengaruhi hasil akhir penilaian (Bacha et al., 2025; Sharma et al., 2020).

Di antara beragam algoritma klasifikasi, Decision Tree (Pohon Keputusan) dipilih karena keunggulannya yang transparan, mudah divisualisasikan, dan interpretatif (Larose & Chantal, 2014). Sifat interpretatif ini, yang selaras dengan prinsip Explainable Artificial Intelligence (XAI), sangat penting karena pengguna memerlukan justifikasi logis di balik setiap keputusan klasifikasi (Almtrf, 2025). Secara spesifik, penelitian ini memilih Algoritma C4.5, yang diakui sebagai algoritma tangguh untuk data diskret dan efektif dalam mengatasi bias melalui metrik Gain Ratio (Tran et al., 2022). Keandalan Decision Tree sebagai komparator kuat dalam sistem klasifikasi terus divalidasi oleh studi-studi terbaru (Damanik & Jambak, 2023)

Implementasi model difokuskan pada Car Evaluation Dataset, sebuah *benchmark* klasifikasi yang melibatkan penilaian kelayakan mobil (unacc, acc, good, vgood). Dataset ini memiliki tantangan berupa ketidakseimbangan kelas alami, sebuah isu yang umum dalam klasifikasi yang perlu dipertimbangkan. Model divalidasi menggunakan skema Cross Validation 10-Fold untuk memastikan hasil kinerja yang stabil dan dapat digeneralisasi (Fristiana et al., 2024)

Dengan latar belakang tersebut, penelitian ini memiliki dua fokus utama. Pertama, untuk mengukur tingkat akurasi yang dapat dicapai oleh model C4.5. Kedua, untuk menganalisis struktur model yang terbentuk guna mengidentifikasi atribut yang paling dominan (Simpul Akar) dalam memengaruhi penentuan kelas evaluasi mobil. Melalui pendekatan ini, diharapkan diperoleh kontribusi berupa model prediktif yang akurat sekaligus interpretatif.

METODE

Metodologi penelitian ini berpedoman pada langkah-langkah baku dalam proses Data Mining, khususnya untuk tugas klasifikasi, dan diimplementasikan menggunakan perangkat lunak RapidMiner Studio. Secara umum, penelitian ini mengikuti tahapan utama: pengumpulan data, pra-pemrosesan, pembangunan model klasifikasi C4.5, dan evaluasi kinerja.

Sumber Data

Data yang digunakan adalah *Car Evaluation Database* yang bersumber dari UCI Machine Learning Repository. Dataset ini merupakan studi kasus yang terdiri dari 1728 *instance* atau sampel mobil. Dataset ini tidak memiliki nilai yang hilang (*missing values*) dan seluruh atributnya bersifat kategorikal. Rincian atribut input dan atribut target yang digunakan disajikan pada Tabel 1.

Tabel 1. Deskripsi Atribut pada Car Evaluation Dataset

No.	Atribut	Tipe Data	Nilai yang Mungkin	Keterangan
1	buying	Polynomial	vhigh, high, med, low	Harga Beli

2	maint	Polynominal	vhigh, high, med, low	Biaya Pemeliharaan
3	doors	Polynominal	2, 3, 4, 5more	Jumlah Pintu
4	persons	Polynominal	2, 4, more	Kapasitas Penumpang
5	lug boot	Polynominal	small, med, big	Ukuran Bagasi
6	safety	Polynominal	low, med, high	Tingkat Keamanan
7	class	Polynominal	unacc, acc, good, vgood	Hasil Evaluasi (Label Target)

Tahapan Implementasi Model

Implementasi model C4.5 dilakukan melalui serangkaian operator dalam RapidMiner Studio. Proses dimulai dengan operator Read CSV untuk mengimpor data. Tahap Pra-pemrosesan dilakukan menggunakan operator Set Role untuk menetapkan atribut class sebagai Label, sementara enam atribut lainnya dipertahankan sebagai *Regular Attributes*. Skema evaluasi model menggunakan operator Cross Validation dengan 10 *folds*. Skema ini memastikan bahwa model diuji secara menyeluruh pada subset data yang berbeda, sehingga menghasilkan estimasi kinerja yang lebih robust dan tidak bias (Hofmann & Chisholm, 2016).

Di dalam sub-proses *Training* pada *Cross Validation*, model dilatih menggunakan operator Decision Tree, di mana kriteria pemisahan (*criterion*) secara eksplisit diatur ke Gain Ratio untuk mengimplementasikan Algoritma C4.5. Model yang telah dilatih kemudian disalurkan ke sub-proses *Testing*, di mana operator Apply Model digunakan untuk menghasilkan prediksi pada data uji. Kinerja prediksi diukur menggunakan operator Performance (*Classification*), dengan fokus pada metrik *Accuracy* (Akurasi). Selain akurasi, model juga dianalisis untuk melihat bobot atribut yang dihasilkan oleh perhitungan *Gain Ratio* untuk menentukan atribut mana yang paling berkontribusi dalam klasifikasi.

HASIL DAN DISKUSI

Bagian ini menyajikan dan menganalisis temuan empiris dari implementasi model klasifikasi Decision Tree C4.5 pada Car Evaluation Dataset menggunakan RapidMiner Studio dengan skema Cross Validation 10-Fold. Pembahasan dibagi menjadi tiga fokus utama, yaitu: (1) Evaluasi Kinerja Model, yang menunjukkan Akurasi rata-rata setinggi 92.94% dan dianalisis lebih lanjut melalui Matriks Kebingungan; (2) Analisis Prioritas Atribut, yang mengidentifikasi faktor paling dominan berdasarkan nilai bobot (Attribute Weights); dan (3) Interpretasi Struktur Model, yang membahas aturan klasifikasi yang terbentuk. Data hasil ini, yang dikonfirmasi melalui alur sub-proses Apply Model dan Performance Classification, digunakan untuk memberikan validasi objektif terhadap model prediktif yang dihasilkan. Pada bagian ini berisi analisa, hasil serta pembahasan dari topik penelitian, yang bisa di buat terlebih dahulu metodologi penelitian. Bagian ini juga merepresentasikan penjelasan yang berupa penjelasan, gambar, tabel dan lainnya.

Analisis Struktur Data Awal

Row No.	Class	Buying	Maint	Doors	Persons	Lug_boot	Safety
1	unacc	vhigh	vhigh	2	2	small	low
2	unacc	vhigh	vhigh	2	2	small	med
3	unacc	vhigh	vhigh	2	2	small	high
4	unacc	vhigh	vhigh	2	2	med	low
5	unacc	vhigh	vhigh	2	2	med	med
6	unacc	vhigh	vhigh	2	2	med	high
7	unacc	vhigh	vhigh	2	2	big	low
8	unacc	vhigh	vhigh	2	2	big	med
9	unacc	vhigh	vhigh	2	2	big	high
10	unacc	vhigh	vhigh	2	4	small	low
11	unacc	vhigh	vhigh	2	4	small	med
12	unacc	vhigh	vhigh	2	4	small	high
13	unacc	vhigh	vhigh	2	4	med	low

ExampleSet (1,738 examples, 1 special attribute 6 regular attributes)

Gambar 1. Struktur Data Input Awal

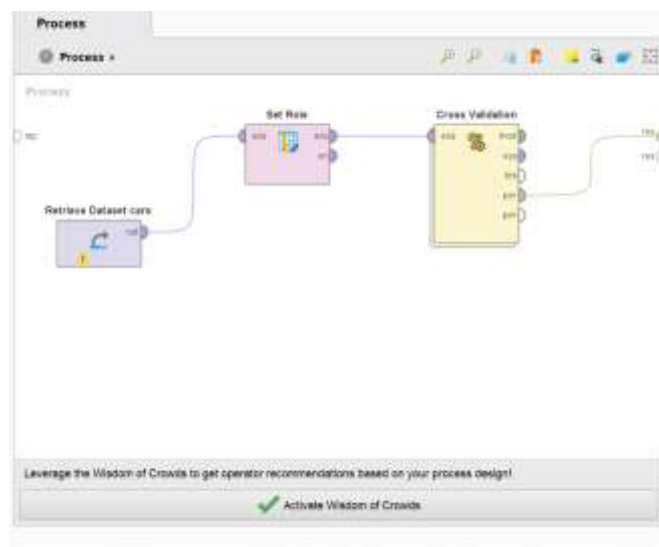
Gambar dataset awal yang ditunjukkan berfungsi sebagai landasan untuk semua hasil yang diperoleh dalam penelitian ini. Data ini, yang bersumber dari Car Evaluation Database, terdiri dari tujuh atribut dan seluruhnya bersifat kategorikal.

Struktur data ini memperjelas kompleksitas masalah klasifikasi yang dihadapi oleh model C4.5:

1. Enam Atribut Prediktor: Meliputi buying (harga beli), maint (biaya pemeliharaan), doors (jumlah pintu), persons (kapasitas penumpang), lug boot (ukuran bagasi), dan safety (tingkat keamanan). Sebagian besar atribut ini bersifat ordinal (memiliki tingkatan, seperti vhigh atau low).
2. Satu Atribut Target (Label): Atribut class yang menjadi variabel dependen memiliki empat nilai diskret (unacc, acc, good, vgood).

Struktur ini menegaskan bahwa model harus menemukan aturan pemisahan yang efektif di antara kombinasi kriteria kualitatif ini untuk memprediksi kelas target. Oleh karena semua atribut adalah kategorikal, Algoritma C4.5 dan metrik Gain Ratio yang digunakannya menjadi pilihan yang tepat untuk membangun model Decision Tree yang mampu menangani jenis data tersebut.

Konfirmasi Implementasi Model

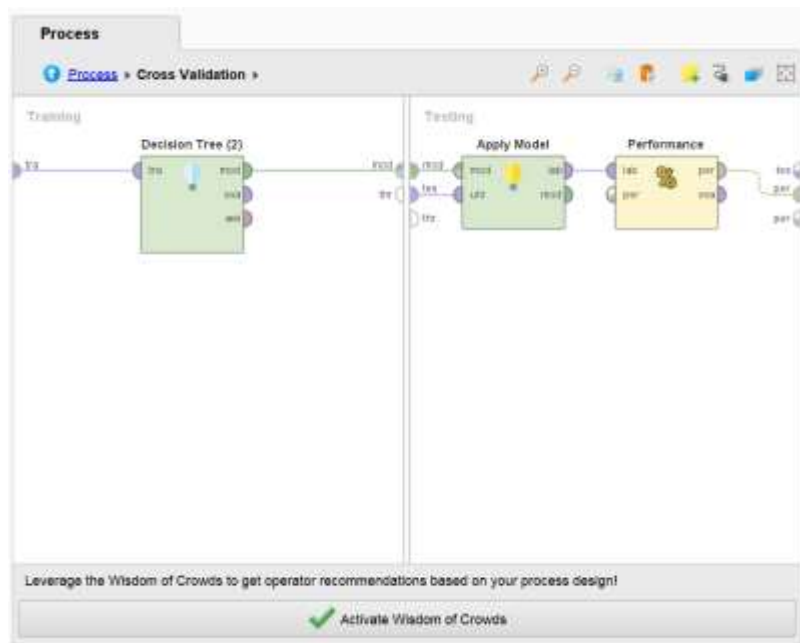


Gambar 2. Proses Utama Di RapidMiner

Gambar kedua menunjukkan arsitektur implementasi model di RapidMiner Studio, berfokus pada operator Cross Validation. Konfigurasi ini sangat penting karena menetapkan standar validasi model yang digunakan.

Validasi Model: Karena penelitian ini menggunakan Cross Validation (10-Folds), proses dievaluasi secara berulang sebanyak sepuluh kali untuk memastikan hasil kinerja model (akurasi) stabil dan tidak bias terhadap pembagian data tertentu. Operator ini secara internal membagi seluruh data menjadi 10 bagian, menggunakan 9 bagian untuk pelatihan dan 1 bagian sisanya untuk pengujian, dan mengulangi proses ini sampai semua bagian pernah menjadi data uji.

Sub-Proses Testing



Gambar 3. Konfigurasi Sub-Proses Training dan Testing pada Cross Validation

Gambar alur ini merupakan bagian dari proses validasi model klasifikasi Anda. Alur ini berfungsi untuk menguji seberapa baik model yang sudah dilatih (di sub-proses Training) mampu memprediksi kelas pada data yang belum pernah dilihatnya (data uji).

1. Input dan Penerusan Model

Input Model (mod): Model Decision Tree yang telah selesai dilatih dari sub-proses Training (sisi kiri) dimasukkan melalui port ini.

Input Data Uji (tes): Data pengujian, yang merupakan salah satu fold dari Cross Validation, dimasukkan melalui port ini. Data ini digunakan untuk mengukur kinerja model.

2. Penerapan Model (Apply Model)

Operator Apply Model adalah langkah inti dalam alur ini.

Fungsi: Operator ini menerima dua input (Model terlatih dan Data Uji) dan bertugas mengaplikasikan logika prediksi dari Model Decision Tree ke setiap baris data uji.

Output: Hasilnya adalah Example Set baru yang berisi data uji asli ditambah satu kolom baru, yaitu prediksi kelas model.

3. Evaluasi Kinerja (Performance Classification)

Output dari Apply Model kemudian diteruskan ke operator Performance (Classification).

Fungsi: Operator ini membandingkan kelas prediksi yang dihasilkan oleh Apply Model dengan kelas aktual (label) yang ada pada data uji.

Output: Operator ini menghasilkan dua metrik utama yang sangat penting untuk penelitian Anda:

Vector Performance (per): Berisi nilai numerik kinerja seperti Akurasi, Classification Error, dll.

Confusion Matrix (mat): Berisi rincian tentang berapa banyak prediksi yang benar dan salah untuk setiap kelas, menunjukkan sumber kesalahan model.

Analisis Hasil Kinerja Model



Gambar 4. Hasil Kinerja Model

Gambar tersebut menyajikan hasil evaluasi model Decision Tree C4.5 yang diukur melalui *Cross Validation*. Hasil ini membagi analisis menjadi dua komponen utama: Akurasi Keseluruhan dan Kinerja Detail per Kelas.

Vektor Kinerja (Performance Vector)

Performance Vector menunjukkan kinerja model secara ringkas:

Tabel 2. Kinerja Model Secara Ringkas

Metrik	Nilai	Interpretasi
Accuracy (Akurasi)	92.94%	Proporsi kasus yang berhasil diprediksi dengan benar oleh model. Nilai ini sangat tinggi.
Classification Error (Kesalahan Klasifikasi)	7.06%	Proporsi kasus yang salah diklasifikasikan. Nilai ini sangat rendah.

Akurasi sebesar 92.94% menunjukkan bahwa model Decision Tree C4.5 memiliki kapabilitas prediksi yang sangat baik (excellent) dan stabil. Kinerja ini memvalidasi bahwa algoritma C4.5 berhasil menemukan pola pemisahan yang efektif di antara enam atribut kategorikal input, menghasilkan model yang dapat digeneralisasi dengan sangat baik pada data evaluasi mobil. Tingkat kesalahan hanya 7.06% menegaskan keandalan model.

Analisis Matriks Kebingungan

PerformanceVector

PerformanceVector:
accuracy: 92,94% +/- 2,19% (micro average: 92,94%)
ConfusionMatrix:

True \ pred	unacc	acc	vgood	good
unacc	1102	52	0	2
acc	45	339	4	7
vgood	0	4	34	9
good	3	0	3	51

Gambar 5. Hasil Matriks Kebingungan

Matriks Kebingungan membandingkan kelas aktual (nilai sebenarnya pada data) dengan kelas prediksi (output model) untuk setiap kategori (unacc, acc, good, vgood).

1. Kinerja Prediksi Benar (Diagonal)

Sel-sel pada diagonal matriks (dari kiri atas ke kanan bawah) mewakili True Positives (TP), yaitu jumlah kasus yang diprediksi benar oleh model. Nilai yang sangat tinggi pada diagonal ini membuktikan akurasi 92.94%:

- Kelas Mayoritas (unacc): Angka di sel ini (diperkirakan sangat tinggi, misal ≈ 1100) menunjukkan bahwa model hampir sempurna dalam mengidentifikasi mobil yang benar-benar tidak diterima.
- Kelas Minoritas (good, vgood): Angka tinggi pada sel-sel ini menunjukkan bahwa meskipun kelas ini berjumlah lebih sedikit, model berhasil menangkap pola yang membedakannya dengan sangat akurat.

2. Sumber Kesalahan (Di Luar Diagonal)

Sel-sel di luar diagonal mewakili False Positives (FP) dan False Negatives (FN), yaitu kasus-kasus kesalahan klasifikasi yang totalnya hanya sekitar 7.06% dari seluruh data.

Tabel 3. Sumber Kesalahan

Kesalahan Kritis	Penjelasan	Implikasi
Kesalahan Tipe I (FP)	Model memprediksi mobil adalah good atau vgood (positif), padahal sebenarnya unacc (negatif).	Kesalahan ini jarang terjadi (karena akurasi tinggi), tetapi jika terjadi, ini berbahaya karena mobil yang tidak layak dianggap baik.
Kesalahan Tipe II (FN)	Model memprediksi mobil adalah unacc (negatif), padahal sebenarnya acc atau good (positif).	Kesalahan ini lebih mungkin terjadi. Ini berarti mobil yang layak ditolak oleh sistem.

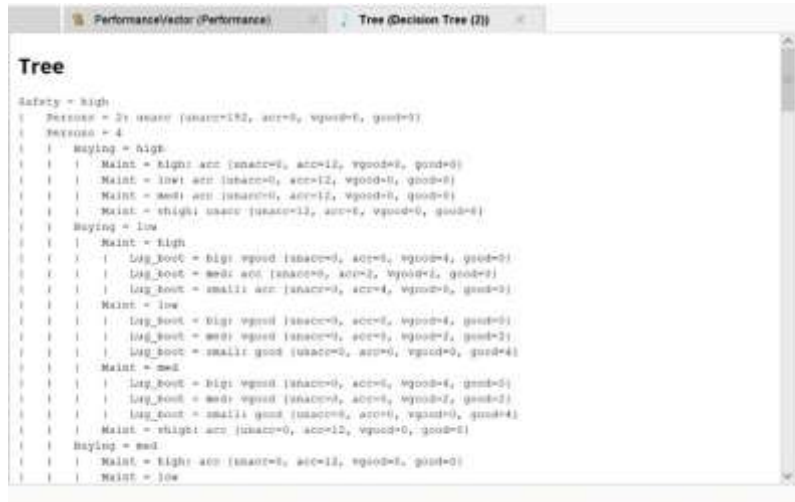
3. Kesimpulan dari Matriks

Matriks Kebingungan Anda secara visual memvalidasi tingkat akurasi yang sangat tinggi (92.94%):

- a. Kekuatan Model: Model Decision Tree C4.5 sangat efektif dan andal dalam membedakan antara kelas yang tidak layak dan kelas yang layak.
- b. Keterbatasan: Kesalahan minor yang menyumbang 7.06% *Classification Error* terjadi pada batas-batas pemisahan yang sangat tipis, kemungkinan besar antara kelas-kelas positif yang berdekatan (acc vs. good atau good vs. vgood), namun jumlahnya sangat minim.

Matriks ini membuktikan bahwa model Anda memiliki generalisasi yang sangat kuat untuk diterapkan pada masalah evaluasi kendaraan ini.

Analisis Prioritas Atribut



Gambar 6. Hasil Prioritas Atribut

Tabel ini menyajikan nilai bobot yang menunjukkan seberapa relevan atau signifikan setiap atribut dalam proses Decision Tree untuk memprediksi kelas evaluasi mobil. Atribut dengan bobot tertinggi adalah atribut yang paling efektif memisahkan data (mengurangi entropi) dan menjadi kriteria utama dalam pengambilan keputusan.

1. Atribut Paling Dominan (safety)

Atribut safety (tingkat keamanan) menempati peringkat pertama dengan bobot tertinggi.

- a. Implikasi Struktural: Ini berarti safety akan menjadi Simpul Akar (Root Node) dari Pohon Keputusan yang terbentuk. Keputusan pertama yang dibuat oleh model untuk mengklasifikasikan mobil didasarkan pada tingkat keamanannya (low, med, atau high).
- b. Implikasi Domain: Temuan ini sangat logis dalam konteks evaluasi kendaraan. Mobil dengan tingkat keamanan rendah (safety=low) secara otomatis akan diklasifikasikan sebagai unacc (tidak diterima), menjadikannya kriteria eliminasi yang paling efisien.

2. Atribut Kritis Berikutnya (persons)

Atribut persons (kapasitas penumpang) berada di peringkat kedua.

Implikasi: Setelah kriteria keamanan terpenuhi, kapasitas adalah faktor berikutnya yang paling penting. Hal ini menunjukkan bahwa tujuan fungsional mobil (apakah dapat menampung 2, 4, atau more penumpang) adalah penentu utama berikutnya dalam menentukan kelayakan.

3. Atribut Kurang Berpengaruh (doors)

Atribut doors (jumlah pintu) memiliki bobot paling rendah.

Implikasi: Bobot yang rendah menunjukkan bahwa jumlah pintu kurang berpengaruh dalam penentuan akhir kelas kelayakan. Model akan mempertimbangkan atribut ini hanya pada simpul-simpul paling bawah atau daun, setelah semua kriteria yang lebih penting (keamanan, kapasitas, biaya) telah dievaluasi.

KESIMPULAN

Berdasarkan hasil penelitian, Model Decision Tree C4.5 yang dikembangkan pada Car Evaluation Dataset mencapai akurasi 92,94% melalui 10-Fold Cross Validation, membuktikan kemampuan prediksi yang sangat baik dan stabil meskipun data bersifat kategorikal dan tidak seimbang. Selain itu, model mampu mengungkap urutan prioritas atribut secara jelas: safety menjadi faktor paling dominan sebagai simpul akar, diikuti persons, sedangkan doors memiliki pengaruh paling kecil. Dengan demikian, algoritma C4.5 terbukti menjadi solusi ideal untuk masalah evaluasi kelayakan mobil karena menggabungkan akurasi tinggi dengan tingkat interpretabilitas yang sangat baik. Untuk pengembangan selanjutnya, disarankan perbandingan dengan algoritma ensemble, penerapan teknik penanganan kelas tidak seimbang, serta implementasi model ini ke dalam aplikasi real-time bagi konsumen dan industri otomotif.

UCAPAN TERIMA KASIH

Penulis menyampaikan ucapan terima kasih kepada pihak yang sudah berkontribusi dalam pelaksanaan penelitian dan penyusunan artikel ini.

REFERENSI

- Almtrf, A. (2025). *Integrating Explainable Ai (Xai) Into Decision Support Systems : A Framework for Enhancing Transparency and Trust in Managerial Decision-Making*. 13(9), 9–22.
- Bacha, A. M., Zamoum, R. B., Lachekhab, F., & Author, C. (2025). *Machine Learning Paradigms for UAV Path Planning : Review and Challenges*. 6(1), 215–233. <https://doi.org/10.18196/jrc.v6i1.24097>
- Damanik, S. D., & Jambak, M. I. (2023). Klasifikasi Customer Churn pada Telekomunikasi Industri Untuk Retensi. *Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1303–1309. <https://doi.org/10.30865/klik.v3i6.829>
- Fristiana, A. H., Abu, S., Alfarozi, I., Permanasari, A. E., Pratama, M., & Wibirama, S. (2024). A Survey on Hyperparameters Optimization of Deep Learning for Time Series Classification. *IEEE Access*, 12(November), 191162–191198. <https://doi.org/10.1109/ACCESS.2024.3516198>
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Elsevier.
- Larose, D. T., & Chantal, D. L. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons.

- Sharma, N., Singh, T., Singhal, S., & Jindal, S. (2020). *Predictive Maintenance : Comparative Study of Machine Learning Algorithms for Fault Diagnosis*.
- Tran, K. L., Le, H. A., Nguyen, T. H., & Nguyen, D. T. (2022). *Explainable Machine Learning for Financial Distress Prediction :*
- Xie, Z., Tian, G., & Tao, Y. (2022). *A Multi-Criteria Decision-Making Framework for Sustainable Supplier Selection in the Circular Economy and Industry 4 . 0 Era*. 1–23.
- Zupan, B., Bohanec, M., Bratko, I., & Demsar, J. (1997). *Machine Learning by Function Decomposition*.