



LITERATURE REVIEW: PERBANDINGAN ALGORITMA ID3 DAN CART DALAM DATA MINING

Ferri Rizaldi ¹, Rijki Hidayat ², Elkin Rilvani ³^{1,2,3}Program Studi Teknik Informatika , Fakultas Teknik, Universitas Pelita Bangsa, CikarangE-mail: rizaldi.feri1@gmail.com ¹, rijki18hidayat@gmail.com ², elkin.rilvani@pelitabangsa.ac.id ³**ABSTRACT**

Classification algorithms play a significant role in data mining, particularly in supporting decision-making processes based on data patterns. Two commonly used decision tree algorithms are ID3 and CART, each with distinct advantages and limitations in terms of speed, accuracy, and flexibility toward different data types. This study aims to conduct a systematic review of recent studies that compare the performance of ID3 and CART algorithms in classification tasks within the context of data mining. The review process follows the PRISMA protocol, starting with literature identification through academic databases such as Google Scholar, ResearchGate, and Garuda. Out of 12 articles identified, only 1 article explicitly compares both algorithms and meets the inclusion criteria. The results of this review indicate that CART generally outperforms ID3 in terms of accuracy and model stability, especially when applied to large-scale numerical datasets. In contrast, ID3 remains relevant for smaller datasets with categorical attributes due to its simplicity and ease of interpretation. This review contributes to the body of knowledge by providing insights and guidance for researchers and practitioners in selecting appropriate classification algorithms based on the characteristics of their data and analytical objectives. Additionally, the findings emphasize the need for further empirical studies that statistically compare decision tree algorithms across various data mining applications.

Keywords: Data Mining, Classification Algorithms, ID3, CART**ABSTRAK**

Algoritma klasifikasi merupakan bagian penting dalam data mining, khususnya dalam mendukung proses pengambilan keputusan berbasis pola data. Dua algoritma yang sering digunakan adalah ID3 dan CART, yang keduanya memiliki keunggulan dan kekurangan dalam hal kecepatan, akurasi, dan fleksibilitas terhadap jenis data. Penelitian ini bertujuan untuk melakukan systematic review terhadap studi-studi terkini yang membahas dan membandingkan algoritma ID3 dan CART dalam konteks klasifikasi data mining. Metode yang digunakan mengacu pada protokol PRISMA, dimulai dari identifikasi literatur melalui database daring seperti Google Scholar, ResearchGate, dan Garuda. Dari total 12 artikel yang ditemukan, hanya 1 artikel yang secara eksplisit membahas perbandingan langsung antara ID3 dan CART dan memenuhi kriteria inklusi. Hasil review menunjukkan bahwa CART memiliki

Article History

Received : Agustus 2025

Reviewed: Agustus 2025

Published: Agustus 2025

Plagiarism Checker No

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)**Copyright : Author****Publish by : Kohesi**

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



keunggulan dalam akurasi dan stabilitas model, terutama pada dataset berskala besar dan numerik, sementara ID3 tetap relevan untuk dataset kecil dengan atribut kategorikal karena mudah diimplementasikan dan diinterpretasikan. Review ini berkontribusi dalam memberikan arah bagi peneliti dan praktisi data mining dalam memilih algoritma klasifikasi yang sesuai dengan karakteristik data dan tujuan analisis. Selain itu, hasil kajian ini diharapkan dapat mendorong penelitian lanjutan yang lebih mendalam dan kuantitatif dalam membandingkan performa algoritma pohon keputusan di berbagai domain.

Kata Kunci: *Data Mining, Algoritma Klasifikasi, ID3, CART*

1. PENDAHULUAN

Semua toko, termasuk toko di Indonesia, harus membuat keputusan yang tepat tentang strategi marketing mereka karena peningkatan persaingan di industri bisnis dan kemajuan teknologi. [1]. Data mining merupakan bagian penting dalam ilmu komputer yang bertujuan untuk mengekstraksi informasi tersembunyi dari kumpulan data besar. Salah satu teknik paling vital dalam proses data mining adalah klasifikasi, yaitu proses memetakan data ke dalam kelas-kelas tertentu berdasarkan atribut yang dimilikinya. Dalam berbagai bidang seperti kesehatan, pemasaran, dan pendidikan, klasifikasi berperan dalam mendukung pengambilan keputusan berbasis data [2].

Algoritma pohon keputusan (*decision tree*) menjadi salah satu pendekatan yang banyak digunakan karena kemampuannya dalam menghasilkan model yang mudah diinterpretasikan dan cepat dalam proses pelatihan. Diantara algoritma pohon keputusan yang populer, *Iterative Dichotomiser 3* (ID3) dan *Classification and Regression Tree* (CART) merupakan dua algoritma yang sering menjadi pilihan utama. ID3, dikembangkan oleh J. R. Quinlan, dikenal karena kesederhanaannya dalam membangun pohon keputusan berdasarkan gain informasi. Sementara itu, CART dikenal karena fleksibilitasnya dalam menangani data numerik dan kategorikal, serta kemampuannya melakukan proses pruning untuk menghindari *overfitting* [3].

Data mining dapat digunakan dalam banyak hal, seperti manajemen hubungan pelanggan (CRM), deteksi penipuan, evaluasi kredit, perawatan medis, dan deteksi penyakit. Penelitian ini dilakukan dengan menganalisis perbandingan algoritma seperti *Decision tree*, *K- Nearest Neighbor* dan *Random Forest* dengan memperhatikan kriteria seperti keakuratan, kecepatan dan ketangguhan dalam penggunaan beberapa jenis data (Dogan & Tanrikulu, 2013). Pada penelitian ini menggunakan data yang diberikan oleh UPT Balai Latihan Kerja Surabaya dengan menggunakan metode klasifikasi dan membandingkan beberapa algoritma yaitu, Algoritma



Decision Tree, Algoritma *K-Nearest Neighbor*, Algoritma *Random Forest* untuk menganalisis tingkat kepuasan masyarakat terhadap sarana prasarana UPT Balai Latihan Kerja Surabaya [4].

Algoritma CART (*Classification and Regression Tree*) merupakan suatu algoritma pohon keputusan yang diperkenalkan oleh empat ilmuwan pada tahun 1984, yaitu Leo Breiman, Jerome H. Friedman, Richard A. Olshen, dan Charles J. Stone. CART dikembangkan untuk memperoleh pohon keputusan baik regresi ataupun klasifikasi. Algoritma C5.0 dan CART memiliki kesamaan dalam hal kecepatan dan menangani data bertipe numerik maupun kategorik. Amalda dkk (2022) menyatakan algoritma C5.0 memiliki kecepatan dalam membuat model dan mampu mengatasi data dalam jumlah besar. C5.0 adalah salah satu algoritma klasifikasi yang paling efektif untuk mengolah data dengan variabel numerik dan kategorik [5]. Pada algoritma CART, interpretasinya lebih mudah, lebih akurat dan lebih cepat dalam perhitungannya dan secara efisien menangani dataset yang besar, baik data tersebut berupa variabel numerik maupun kategorik [3].

Dalam kategorisasi data, setiap algorithm memiliki keuntungan dan kerugian. *Decision trees* sangat mudah digunakan dan digunakan, tetapi mereka sering *overfitting* besar dataset. Random forest, yang menggabungkan banyak decision trees, dapat menghasilkan classifications yang lebih akurat serta mengurangi kemungkinan overfitting, tetapi memerlukan lebih banyak waktu komputasi [6].

2. TINJAUAN PUSTAKA

a) Data Mining

Secara umum, Data Mining mencakup dua elemen utama:

- Data: Merupakan himpunan informasi atau fakta yang telah tercatat, namun kerap tidak dimanfaatkan secara optimal.
- Mining: Proses menambang data untuk menghasilkan pengetahuan.

Untuk memeriksa besar databases dan menemukan patterns baru yang berguna, data mining adalah teknik yang digunakan untuk mengekstrak data berguna dari repositori data besar. Namun, beberapa aktivitas pengumpulan data tidak semuanya dapat dikategorikan sebagai data mining. Contoh berikut menunjukkan apa yang dianggap data mining dan apa yang tidak:

- Bukan Data Mining: Menggunakan kata kunci untuk mencari informasi tertentu di internet. Dalam konteks Data Mining, informasi yang sejenis dapat dikelompokkan dari hasil pencarian mesin pencari dengan mempertimbangkan variabel-variabel tertentu seperti status akreditasi, total mahasiswa, dan reputasi lulusan.
- Bukan Data Mining: Menganalisis kenaikan permintaan dengan membandingkan laporan keuangan antara bulan Januari dan Februari. Data Mining: Memprediksi pengeluaran dan pembelian di masa mendatang dengan menggunakan laporan keuangan saat ini.

**b) Algoritma CART**

CART (*Classification and Regression Tree*) adalah metode yang digunakan untuk menyusun pohon keputusan dalam bentuk pohon klasifikasi. Pada prosesnya, setiap simpul induk (*parent node*) akan dipecah menjadi dua simpul anak (*child node*). Setiap *child node* selanjutnya dapat berperan sebagai *parent node* baru dalam proses pembentukan berikutnya. Proses ini akan terus berlangsung secara rekursif sampai tidak ditemukan lagi kemungkinan pemisahan yang optimal.

c) Algoritma ID3

ID3 (*Iterative Dichotomiser 3*) merupakan salah satu algoritma pengembangan pohon keputusan yang dikembangkan oleh Quinlan, dengan pendekatan pemilihan atribut berdasarkan nilai informasi tertinggi (*information gain*) pada setiap node. Algoritma ini membagi data pada setiap proses pembentukan pohon dengan memilih fitur yang memberikan informasi paling signifikan. ID3 cukup efektif diterapkan pada data berukuran kecil, namun cenderung kurang akurat saat digunakan pada dataset berukuran besar. [2].

3. METODE PENELITIAN

Metodologi yang diterapkan dalam penelitian ini adalah *Systematic Literature Review* (SLR). Tinjauan literatur merupakan jenis penelitian yang melibatkan pengumpulan dan ulasan data dari 12 artikel jurnal, untuk mengumpulkan informasi yang diperlukan, yang kemudian akan dianalisis lebih lanjut. Studi ini menerapkan metode yang dikenal sebagai *Systematic Literature Review* (SLR). Metodologi ini digunakan untuk mengintegrasikan atau menganalisis penelitian yang ada mengenai topik tertentu. SLR bertujuan untuk menganalisis, mengidentifikasi, menyelidiki, dan mengevaluasi penelitian terkait dengan topik dan pertanyaan penelitian.

4. HASIL PENELITIAN

Berdasarkan hasil pencarian dan telaah terhadap sejumlah artikel penelitian sebelumnya, ditemukan 12 artikel yang relevan dengan topik Perbandingan Algoritma ID3 Dan Cart Dalam Data Mining. Artikel-artikel ini dipilih karena secara khusus membahas topik Perbandingan Algoritma ID3 Dan Cart Dalam Data Mining. tinjauan ini juga bertujuan untuk mengevaluasi kontribusi masing-masing penelitian dalam memberikan wawasan strategis bagi pengambilan keputusan bisnis berbasis data.



No	Judul	Hasil Penelitian
1	Klasifikasi Data Penjualan	Hasil penelitian menunjukkan bahwa, meskipun metode pemisahan data berbeda, nilai akurasi regresi dan pohon keputusan untuk kelas percabangan memiliki nilai yang sama. Sebaliknya, untuk kelas tipe pelanggan, nilai akurasi regresi dan pohon keputusan untuk kedua metode pemisahan data berbeda. Untuk metode pemisahan data hold-out, akurasi regresi adalah 0,5014285714285714, dan akurasi pohon keputusan adalah 0,5. Untuk metode pemisahan data k-fold, akurasi regresi adalah 0,466, dan akurasi pohon keputusan adalah 0,492. Untuk kelas percabangan, akurasi pohon keputusan adalah 0,5 [7].
2	Klasifikasi Penjualan Produk	Hasil dari optimasi hiperparameter menunjukkan bahwa hasil yang diperoleh tidak jauh berbeda dengan model dasar meskipun menggunakan kombinasi parameter terbaik. <i>Random Forest</i> menunjukkan performa yang konsisten dan tingkat generalisasi yang unggul dalam berbagai skenario klasifikasi, dengan akurasi 99% dibandingkan 97,5% <i>Decision Tree</i> . [4].
3	Penilaian Kepuasan Penumpang Maskapai Pesawat	Hasil dari perbandingan performa ketiga algoritma menunjukkan bahwa algoritma <i>Random Forest</i> memiliki kelebihan nilai akurasi yang paling besar di antara dua algoritma pembandingnya. Algoritma <i>Decision Tree</i> memiliki kelebihan nilai akurasi yang lebih besar daripada algoritma <i>naif Bayes</i> , tetapi kekurangannya adalah nilai akurasi yang dihasilkannya masih lebih rendah daripada algoritma <i>Random Forest</i> [8].
4	Klasifikasi Hasil Indeks Kepuasan Masyarakat Terhadap UPT Balai Latihan Kerja Surabaya	Hasil penelitian dengan menggunakan <i>Jupyter Annaconda 3</i> ini menunjukan algoritma <i>Random Forest</i> dengan performa terbaik yang mendapat nilai akurasi sebesar 0.95. Algoritma <i>Decision Tree</i> dengan nilai akurasi sebesar 0.77. Algoritma <i>KNearest Neighbor</i> mendapat nilai akurasi sebesar 0.82. Nilai akurasi



		tersebut terdapat dengan menggunakan data Indeks Kepuasan Masyarakat terhadap sarana prasarana [9].
5	Klasifikasi Penyakit Diabetes Mellitus	Proses klasifikasi dalam penelitian ini dilakukan dengan menerapkan algoritma ID3 dan <i>Naive Bayes</i> , menggunakan bahasa pemrograman Python melalui platform open-source berbasis web, yakni Jupyter Notebook. Tujuan utama dari studi ini adalah membangun model klasifikasi serta mengimplementasikan teknik data mining dalam konteks tersebut. Hasil Algoritma ID3 memiliki akurasi 74% dengan nilai AUC 0,788, sedangkan algoritma <i>Naive Bayes</i> memiliki 76% dengan nilai AUC 0,794. Klasifikasi data dievaluasi menggunakan matriks konfusi dan kurva ROC. Akibatnya, dapat dikatakan bahwa algoritma <i>Naive Bayes</i> memiliki hasil prediktor diabetes melitus yang baik [10].
6	Klasifikasi Penjualan Barang di Swalayan Dutalia	Untuk menemukan pola penjualan, algoritma C4.5 memproses banyak data. Pokok keputusan dibuat dan prediksi untuk catatan penjualan baru dibuat. Ini memberi Dutalia Supermarket kemampuan untuk menetapkan tingkat persediaan minimum. Jenis produk, harga, kuantitas, waktu penjualan, dinamika penjualan, dan status penjualan adalah enam atribut yang digunakan sebagai atribut target. Aplikasi RapidMiner Studio versi 9.7 memungkinkan pengujian data pelatihan dan hasil prediksi dengan akurasi 100%. Ini memungkinkan sistem untuk memprediksi penjualan di masa mendatang [1].
7	Prediksi Serangan Jantung	Berdasarkan eksperimen dan pengujian yang telah dilakukan menggunakan metode komputasi algoritma ID3, <i>Naive Bayes</i> , dan SVM berbasis <i>Particle Swarm Optimization</i> (PSO) dalam memprediksi potensi serangan jantung, ditemukan bahwa algoritma ID3



		menunjukkan tingkat akurasi sebesar 75.20% dan AUC = 0.735. ketika Agoritma ID3 dioptimasi menggunakan PSO, hasilnya meningkat menjadi 80.49% dan AUC = 0.815. Selanjutnya pada pengujian algoritma Naïve Bayes menunjukkan tingkat akurasi sebesar 81.52% dan AUC = 0.890. ketika Agoritma Naïve Bayes ini dioptimasi menggunakan PSO, hasilnya meningkat menjadi 83.94% dan AUC = 0.901. Dan pada pengujian terakhir yang menggunakan algoritma SVM menunjukkan tingkat akurasi sebesar 82.13% dan AUC = 0.895 sementara ketika algoritma SVM ini dioptimasi menggunakan PSO, hasilnya meningkat menjadi 84.83% dan AUC = 0.900 dan menjadi algoritma dengan <i>accuracy</i> tertinggi dibandingkan 2 algoritma sebelumnya, tetapi untuk AUC algoritma <i>Niave Bayes</i> + PSO lebih tinggi 1 angka dari pada algoritma SVM + PSO [2].
8	Menyeleksi Calon Karyawan	Dibandingkan dengan algoritma C4.5, algoritma CART menunjukkan tingkat akurasi dan keberhasilan yang lebih baik dalam proses seleksi calon karyawan. Kedua algoritma tersebut pada dasarnya tidak terlalu dipengaruhi oleh variasi jumlah data dalam proses pemilihan. Namun, khusus pada algoritma CART, terdapat kecenderungan bahwa semakin besar jumlah data yang dianalisis, maka semakin tinggi pula tingkat keberhasilan dan ketepatan hasil seleksi yang dihasilkan oleh sistem. [11].
9	Prediksi Penyakit Stroke	Hasil penelitian algoritma dengan memeriksa akurasi dengan matriks konfusi di atas, kita dapat menentukan perbandingan jumlah hasil <i>TRUE POSITIVE</i> , <i>TRUE NEGATIF</i> , <i>FALSE POSITIVE</i> , dan <i>FALSE NEGATIF</i> dari kedua model tersebut. Untuk data tidak seimbang, algoritma penelitian dengan akurasi tertinggi adalah regresi logistik, pohon random, SVM, dan KNN, dengan akurasi 75,65% [9].



10	<i>Stroke Classification</i>	Penanganan data hilang pada variabel BMI dan status merokok dilakukan dengan melakukan penghapusan amatan yang memiliki data hilang. Analisis data menggunakan CART dan C5.0 dapat menangani data yang bertipe numerik dan kategorik, namun pada analisis menggunakan <i>python</i> data kategorik tidak dapat diolah sehingga harus dilakukan transformasi data. Oleh karena itu, dilakukan transformasi data kategorik menjadi numerik pada variabel jenis kelamin, status pernikahan, jenis pekerjaan, jenis tempat tinggal, dan status merokok. Selanjutnya, dilakukan undersampling untuk mengatasi data tak seimbang pada variabel stroke sehingga jumlah data sampel mayoritas diturunkan sampai sama dengan jumlah kelas minoritas dan diperoleh jumlah amatan kategori mengalami stroke dan tidak mengalami stroke masing-masing sebesar 180 amatan [12].
11	Klasifikasi Kasus Penderita Kanker Payudara	Dalam algoritma CART, variabel yang paling penting adalah bulan bertahan hidup, simpul regional positif, dan usia; dalam algoritma C4.5, variabel yang paling penting adalah bulan bertahan hidup, level 6, dan kelas. Ada satu kesamaan antara kedua algoritma ini: variabel simpul akar, yang diwakili oleh variabel "bulan bertahan hidup". Nilai akurasi, perolehan kembali, presisi, dan tingkat kesalahan model pohon yang dihasilkan sangat bergantung dan dipengaruhi oleh bagaimana set data dalam set data pelatihan dan uji disusun [3].
12	Klasifikasi Faktor-Faktor yang Mempengaruhi Karier Mahasiswa Ilmu Komputer	Berdasarkan hasil evaluasi, algoritma Random Forest menunjukkan tingkat akurasi prediksi yang lebih tinggi dibandingkan Decision Tree dan ID3, terutama dalam menangani dataset berskala besar dan kompleks dengan risiko overfitting yang lebih rendah. Selain itu, ditemukan bahwa mahasiswa yang terlibat dalam proyek-proyek kecerdasan buatan atau pembelajaran



		mesin cenderung memilih karier di bidang data science, sedangkan mereka yang fokus pada pengembangan perangkat lunak lebih banyak menempuh jalur karier sebagai software engineer. Kemampuan teknis dalam bahasa pemrograman seperti Python dan SQL, serta keterlibatan aktif dalam proyek perangkat lunak, menjadi faktor penting yang memengaruhi arah karier mahasiswa [6].
--	--	--

Berdasarkan hasil analisis terhadap 12 artikel yang membahas perbandingan algoritma klasifikasi dalam data mining, dapat disimpulkan bahwa algoritma CART menunjukkan performa yang lebih unggul dibandingkan ID3 dalam hal akurasi dan stabilitas model, khususnya pada dataset yang besar dan memiliki atribut numerik. Kemampuan CART dalam menangani data kompleks dan melakukan pruning menjadikannya pilihan yang lebih andal dalam banyak studi yang direview.

Sementara itu, algoritma ID3 tetap memiliki kelebihan pada dataset kecil dengan atribut kategorikal, karena proses pembentukan pohonnya yang sederhana dan hasil yang mudah diinterpretasikan. Namun, ID3 cenderung kurang stabil dan akurat ketika diterapkan pada data berskala besar.

Dengan demikian, pemilihan algoritma klasifikasi dalam data mining sangat bergantung pada karakteristik data dan tujuan analisis. Hasil review ini diharapkan dapat menjadi acuan bagi peneliti dan praktisi dalam menentukan pendekatan yang paling tepat dan efisien dalam proses klasifikasi data, serta mendorong adanya penelitian lanjutan yang lebih terarah dan berbasis eksperimen kuantitatif.

5. KESIMPULAN

Berdasarkan hasil literature review terhadap 12 artikel yang relevan, dapat disimpulkan bahwa algoritma CART cenderung memiliki kinerja lebih baik dibandingkan dengan algoritma ID3, terutama dalam hal akurasi dan stabilitas model pada dataset yang besar dan bertipe numerik. CART juga memiliki keunggulan dalam melakukan proses pruning untuk menghindari overfitting.

Algoritma ID3 masih relevan digunakan untuk dataset kecil dengan atribut kategorikal, karena algoritma ini lebih sederhana dan mudah diinterpretasikan. Namun, ID3 kurang mampu



menghasilkan model yang akurat dan stabil ketika dihadapkan pada data yang kompleks dan berukuran besar.

Dalam praktiknya, pemilihan algoritma klasifikasi sangat tergantung pada karakteristik data yang digunakan, seperti ukuran dataset, jenis atribut (numerik atau kategorikal), serta tujuan analisis data. Studi ini memberikan gambaran bahwa tidak ada satu algoritma yang selalu unggul dalam semua kondisi, melainkan pemilihan harus disesuaikan dengan kebutuhan dan konteks penggunaannya.

6. SARAN

Berdasarkan hasil kajian literatur yang telah dilakukan, disarankan agar penelitian-penelitian selanjutnya dapat melakukan eksplorasi yang lebih mendalam dan bersifat kuantitatif dalam membandingkan algoritma ID3 dan CART. Penelitian lanjutan diharapkan tidak hanya terbatas pada aspek akurasi, tetapi juga mempertimbangkan faktor lain seperti waktu komputasi, kompleksitas model, dan kemampuan generalisasi terhadap berbagai jenis data serta konteks penerapannya di dunia nyata.

Bagi para praktisi di bidang data mining, pemilihan algoritma klasifikasi sebaiknya didasarkan pada karakteristik data yang dimiliki, baik dari segi ukuran dataset, jenis atribut, hingga kebutuhan interpretasi hasil. CART dapat menjadi pilihan utama ketika menghadapi data berskala besar dan kompleks, sedangkan ID3 tetap layak digunakan pada dataset yang lebih kecil dan sederhana, khususnya ketika kemudahan interpretasi menjadi prioritas utama.

Pengembangan sistem berbasis klasifikasi seperti sistem pendukung keputusan atau sistem intelijen bisnis sebaiknya juga memperhatikan keseimbangan antara performa algoritma dengan kebutuhan pengguna akhir. Dengan demikian, pemanfaatan algoritma klasifikasi dalam data mining tidak hanya menjadi solusi teknis, tetapi juga mampu menjawab tantangan strategis dalam pengambilan keputusan berbasis data.

DAFTAR PUSTAKA

- [1] S. Dutalia, A. K. Lalo, P. Batarius, Y. Carmeneja, and H. Siki, "Implementasi Algoritma C4 . 5 Untuk Klasifikasi Penjualan," vol. 06, pp. 1-12, 2021.
- [2] M. B. Prayogi, I. Irawan, and Y. I. Fajar, "Perbandingan Algoritma ID3 , Naive Bayes , SVM Berbasis PSO Untuk Prediksi Serangan Jantung," *Mdp Student Conf.*, pp. 22-28, 2024.
- [3] F. Melani and S. Sulastri, "Analisis Perbandingan Klasifikasi Algoritma CART dengan Algoritma C 4.5 Pada Kasus Penderita Kanker Payudara," *J. Tekno Kompak*, vol. 17, no. 1, p. 171, 2023, doi: 10.33365/jtk.v17i1.2379.
- [4] Putri Ayu Firnanda, Litasya Shofwatillah, Fauziah Rahma, and Fatkhurokhman Fauzi, "Analisis Perbandingan Decision Tree dan Random Forest dalam Klasifikasi Penjualan



- Produk pada Supermarket,” *Emerg. Stat. Data Sci. J.*, vol. 3, no. 1, pp. 445-461, 2025, doi: 10.20885/esds.vol3.iss.1.art2.
- [5] L. I. Prahartiwi, “Implementasi Algoritma Fp-Growth Untuk Menemukan Pola Pembelian Konsumen Pada Analisis Keranjang Pasar,” *IJIS - Indones. J. Inf. Syst.*, vol. 7, no. 1, pp. 71-78, 2022, doi: 10.36549/ijis.v7i1.208.
- [6] D. Hartama and N. Amalya, “Perbandingan Algoritma Decision Tree, ID3, dan Random Forest dalam Klasifikasi Faktor-Faktor yang Mempengaruhi Karier Mahasiswa Ilmu Komputer,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 1, pp. 72-80, 2025, doi: 10.35870/jimik.v6i1.1113.
- [7] A. W. Wardhana, E. Patimah, A. I. Shafarindu, Y. M. Siahaan, B. V. Haekal, and D. S. Prasvita, “Klasifikasi Data Penjualan pada Supermarket dengan Metode Decision Tree,” *Senamika*, vol. 2, no. 1, pp. 660-667, 2021, [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/1389>
- [8] A. . Rahmat, M. . Ladjamuddin, and T. . Awaludin, “Perbandingan Algoritma Decision Tree, Random Forest Dan Naive Bayes Pada Prediksi Penilaian Kepuasan Penumpang Maskapai Pesawat Menggunakan Dataset Kaggle,” *J. Rekayasa Inf.*, vol. 12, no. 2, pp. 150-159, 2023, [Online]. Available: www.kaggle.com,
- [9] S. A. Wicaksono and I. K. D. Nuryana, “Perbandingan Klasifikasi Hasil Indeks Kepuasan Masyarakat Terhadap Upt Balai Latihan Kerja Surabaya Menggunakan Algoritma Decision Tree, Random Forest, K-Nearest Neighbor,” *J. Informatics Comput. Sci.*, vol. 5, no. 04, pp. 552-228, 2024, doi: 10.26740/jinacs.v5n04.p552-228.
- [10] N. Nurdiana and A. Algifari, “Studi Komparasi Algoritma Id3 Dan Algoritma Naive Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus,” *INFOTECH J.*, vol. 6, no. 2, pp. 18-23, 2020, [Online]. Available: <https://ejournal.unma.ac.id/index.php/infotech/article/view/816>
- [11] N. P. Wong, F. N. S. Damanik, C. -, E. S. Jaya, and R. Rajaya, “Perbandingan Algoritma C4.5 dan Classification and Regression Tree (CART) Dalam Menyeleksi Calon Karyawan,” *J. SIFO Mikroskil*, vol. 20, no. 1, pp. 11-18, 2019, doi: 10.55601/jsm.v20i1.622.
- [12] Indah Lestari, Dina Fitria, Syafriandi Syafriandi, and Admi Salma, “Comparison of the C5.0 Algorithm and the CART Algorithm in Stroke Classification,” *UNP J. Stat. Data Sci.*, vol. 2, no. 1, pp. 90-98, 2024, doi: 10.24036/ujsds/vol2-iss1/144.