

# Algoritma Keputusan C4.5 Untuk Klasifikasi Data Medis Pasien

Zaehol Fatah<sup>1</sup>, Rifki Dwi Saputra<sup>2</sup>  
<sup>1,2</sup> Universitas Ibrahimy, Situbondo  
Email : riefkysyaputra14@gmail.com

## Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan status kesehatan pasien berdasarkan data rekam medis menggunakan algoritma pohon keputusan C4.5. Dataset yang digunakan terdiri dari 1000 entri pasien dengan atribut seperti usia, kadar gula, tekanan darah, gejala, dan riwayat keluarga. Studi ini dilakukan menggunakan metode kuantitatif dengan karakter deskriptif eksploratif. Pemodelan dilakukan menggunakan perangkat lunak RapidMiner dengan tahapan: praproses data, penentuan atribut target, pembangunan model klasifikasi, dan evaluasi hasil. Hasil pemodelan menunjukkan bahwa atribut Kadar\_Gula memiliki bobot tertinggi dan menjadi akar dari pohon keputusan, disusul oleh Usia dan Riwayat\_Keluarga. Model yang dihasilkan bersifat logis dan sistematis, serta menunjukkan jalur klasifikasi yang mudah ditelusuri. Proses pruning otomatis yang dilakukan oleh algoritma mencegah overfitting, sehingga model memiliki performa generalisasi yang baik. Model dievaluasi memakai teknik 10-fold cross-validation, kemudian hasilnya dipresentasikan dalam bentuk visualisasi melalui struktur pohon serta grafik bobot atribut. Analisis data mengindikasikan bahwa implementasi algoritma C4.5 memberikan performa yang efisien dalam pengolahan data rekam medis untuk mendukung sistem pengambilan keputusan medis yang berbasis data.

**Kata Kunci:** Medical Data, Algoritma C4.5, Decision Tree, RapidMine, Klasifikasi.

## Abstrac

*The study is conducted to determine patient health categories through the utilization of the C4.5 decision tree model on medical data. The dataset used consists of 1000 patient entries with attributes such as age, blood sugar level, blood pressure, symptoms, and family history. This work is conducted using a quantitative method that emphasizes descriptive and exploratory analysis. Modeling is carried out using RapidMiner software with the following stages: data preprocessing, determining target attributes, building a classification model, and evaluating results. The modeling results show that the Blood Sugar Level attribute has the highest weight and is the root of the decision tree, followed by Age and Family History. The resulting model is logical and systematic, and shows an easy-to-track classification path. The automatic pruning process carried out by the algorithm prevents overfitting, so that the model has good generalization performance. Employing the 10-fold cross-validation procedure, the model evaluation results were depicted via a decision tree layout and an attribute importance graph. These results indicate that the C4.5 algorithm is effective in processing medical record data to support a data-based medical decision-making system.*

**Keywords:** Medical Data, C4.5 Algorithm, Decision Tree, RapidMiner, Classification.

## PENDAHULUAN

Perkembangan teknologi informasi telah memicu pertumbuhan data dalam skala besar pada berbagai bidang, termasuk kesehatan. Data yang melimpah ini menuntut adanya metode analisis cerdas agar dapat dimanfaatkan menjadi

pengetahuan yang bermakna. Di sinilah peran data mining menjadi sangat penting sebagai proses untuk mengekstraksi pola dan informasi tersembunyi dari kumpulan data besar (Fiandra et al., 2017). Melalui pendekatan ini, informasi yang semula tidak tampak dapat diubah menjadi dasar

pengambilan keputusan yang bermanfaat, baik dalam penelitian maupun praktik profesional.

Dalam konteks kesehatan, data mining berperan signifikan dalam menganalisis data rekam medis untuk mengidentifikasi pola penyakit, faktor risiko, dan potensi diagnosis (Budiarti, 2025). Rekam medis merupakan sumber data kaya informasi yang berisi catatan identitas, pemeriksaan, tindakan medis, serta riwayat penyakit pasien. Sayangnya, sebagian besar data ini hanya disimpan tanpa diolah lebih lanjut, sehingga potensinya untuk mendukung pengambilan keputusan medis belum dimanfaatkan secara optimal (Fiandra et al., 2017).

Klasifikasi termasuk metode pokok dalam data mining yang digunakan untuk menentukan kategori data berdasarkan sifat atau atribut yang melekat padanya. Metode klasifikasi banyak diterapkan pada kasus medis untuk mendiagnosis penyakit seperti diabetes, paru-paru, jantung, dan penyakit kuning (Budiarti, 2025; Mansiz et al., 2025). Di antara berbagai algoritma klasifikasi, C4.5 merupakan algoritma populer karena mampu menghasilkan model berbentuk pohon keputusan (decision tree) yang mudah dipahami oleh pengguna non-teknis (Han, Kamber, & Pei, 2012; Wulandari, 2025). Metode klasifikasi seperti C4.5 telah banyak digunakan dalam diagnosis medis karena kemampuannya menghasilkan model yang mudah diinterpretasi oleh tenaga kesehatan (Lestari & Wibowo, 2023; Setiawan & Ramadhani, 2024). Keunggulan algoritma ini terletak pada kemampuannya mengelola data kategorikal, menangani nilai hilang, serta menghasilkan aturan keputusan yang transparan dan mudah diinterpretasikan (Mansiz et al., 2025).

Penelitian-penelitian terdahulu menunjukkan efektivitas C4.5 dalam menganalisis data medis. Misalnya, penerapan algoritma C4.5 pada klasifikasi penyakit paru-paru mampu mencapai tingkat akurasi hingga 94% (Mansiz et al., 2025), sedangkan pada prediksi diabetes

menghasilkan akurasi sebesar 88,46% (Pamungkas et al., 2025). Selain itu, algoritma ini juga digunakan untuk klasifikasi penyakit berdasarkan kode ICD-10, yang membantu rumah sakit mengelompokkan penyakit sesuai kategori internasional (Fiandra et al., 2017). Implementasi tersebut membuktikan bahwa data mining berperan signifikan dalam proses pengambilan keputusan medis yang didasarkan pada data empiris.

Rekam medis pasien cenderung menumpuk tanpa dimanfaatkan sebagai dasar pengambilan keputusan untuk peningkatan pelayanan. Solusi yang dapat dikembangkan melibatkan penerapan metode C4.5 dimanfaatkan dalam proses pengelompokan penyakit yang mengacu pada sistem kode internasional ICD-10. Pada penelitian Fiandra dkk. (2017), diterapkan C4.5 dengan 21 kategori ICD-10 (A00–Z99). Hasilnya, algoritma C4.5 berhasil merepresentasikan 14 dari 21 kategori, dengan persentase data terbaca lebih dari 66%. Ini menunjukkan potensi algoritma C4.5 dalam mengungkap pola penyakit umum dalam data rekam medis.

Penelitian selanjutnya memperluas studi menggunakan empat atribut—kode ICD-10, jenis kelamin, umur, dan bulan kunjungan—untuk menyederhanakan model, dan berhasil mengklasifikasikan lima jenis kategori penyakit yang dikodekan dalam rentang A00–B99, I00–I99, J00–J99, O00–O99, hingga Z00–Z99. Temuan menarik meliputi pola demografis: penyakit A00–B99 dominan pada laki-laki muda dan dewasa, perempuan tua, dengan distribusi bulan tertentu; penyakit J00–J99 dominan pada bayi/anak laki-laki di November; dan lainnya dengan variasi usia serta jenis kelamin menurut bulan spesifik. Namun, karena label kelas terlalu banyak (21 kelompok), hanya 5 kelas yang berhasil diklasifikasikan, dengan persentase data terbaca kurang dari 50% (Elektro, 2013).

Abi Prakasa & Milkhatun (2019) Dalam penelitian ini, algoritma C4.5

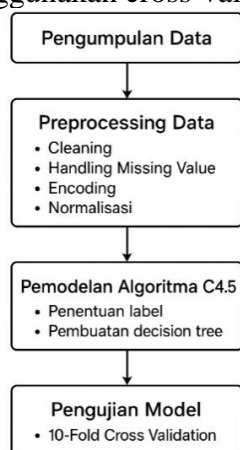
diaplikasikan pada studi kasus pasien gangguan proses pikir waham di RSJD Atma Husada Mahakam Samarinda, dengan total 54 data rekam medis dijadikan sebagai sampel penelitian. Temuan analisis mengindikasikan bahwa proporsi terbesar pasien adalah laki-laki berusia antara 36 sampai 64 tahun. Dengan gejala dominan berupa gangguan emosi (Treise et al., 2019). Temuan penelitian ini memperkuat bukti bahwa C4.5 efektif dalam mengekstraksi informasi klinis bahkan pada dataset dengan jumlah terbatas.

### METODE

Rancangan penelitian ini menggunakan pendekatan kuantitatif dengan jenis deskriptif eksploratif. Pendekatan ini bertujuan untuk mendeskripsikan pola klasifikasi pada data rekam medis pasien menggunakan algoritma C4.5. Data yang digunakan merupakan data sekunder berupa dataset rekam medis pasien dalam format Excel. Dataset terdiri atas atribut:

1. Usia
2. Tekanan darah
3. Kadar gula
4. Gejala
5. Riwayat keluarga
6. Status kesehatan (sebagai label klasifikasi)

Tahapan penelitian meliputi: pengumpulan data, preprocessing data, pemodelan algoritma C4.5, dan pengujian model menggunakan cross-validation.



Gambar 1. Diagram Alur Penelitian

### Deskripsi Umum Diagram Alur Penelitian

1. Pengumpulan Data  
Mengambil dataset rekam medis dari sumber yang telah disediakan. Data digunakan dalam bentuk Excel dan telah dianonimkan untuk menjaga kerahasiaan pasien.
2. Preprocessing Data  
Meliputi pembersihan data, penanganan missing value, pengkodean atribut kategorikal, dan normalisasi apabila diperlukan. Tujuannya adalah menyiapkan data agar siap diproses oleh algoritma.
3. Pemodelan Algoritma C4.5  
Data diolah menggunakan RapidMiner untuk membangun model pohon keputusan. Atribut *Status* ditetapkan sebagai label, dan algoritma C4.5 digunakan dengan metode Gain Ratio serta mekanisme pruning otomatis.
4. Pengujian Model  
Model dievaluasi menggunakan teknik 10-fold cross-validation untuk mengukur akurasi dan performa generalisasi.

### Pengumpulan Data

Pengumpulan data dilakukan dengan metode studi dokumentasi, yaitu mengakses dataset rekam medis yang telah dikompilasi sebelumnya. Populasi penelitian terdiri atas 1000 entri pasien, dan karena seluruh data memenuhi kriteria kelayakan, penelitian ini menggunakan metode total sampling. Dataset yang digunakan telah dianonimkan sehingga tidak memuat identitas pribadi pasien, sehingga aman ditinjau dari aspek etika penelitian.

### Sumber Dataset

Dataset diperoleh dari: Bagian Rekam Medis Rumah Sakit X (disamakan untuk menjaga kerahasiaan institusi), berupa data rekam medis tahun 2023 yang telah dikompilasi dan dianonimkan oleh pihak rumah sakit.

### Praproses Data

Tahapan pengolahan awal dilakukan

menggunakan Microsoft Excel sebagai upaya menyiapkan dataset agar siap diproses dalam model klasifikasi. Praproses mencakup:

1. Pembersihan Data (Cleaning) untuk menghilangkan data duplikat dan inkonsistensi.
2. Penanganan Missing Value dengan pengisian atau penghapusan sesuai kebutuhan.
3. Pengkodean Atribut Kategorikal seperti gejala dan riwayat keluarga menjadi format numerik.
4. Normalisasi Data Numerik jika diperlukan untuk meningkatkan performa algoritma.

Tahapan ini bertujuan memastikan bahwa data telah terstruktur dan siap diproses oleh perangkat lunak pemodelan.

**Pembangunan Model Klasifikasi**

Setelah praproses selesai, data dianalisis menggunakan algoritma pohon keputusan C4.5 yang diimplementasikan melalui perangkat lunak RapidMiner Studio. RapidMiner dipilih karena memiliki antarmuka visual yang memudahkan penyusunan alur kerja (*workflow*) menggunakan operator-operator modular.

Tahapan pembangunan model meliputi:

1. Retrieve – mengimpor dataset ke dalam workspace RapidMiner.
2. Set Role – menentukan atribut *Status* sebagai *label* (target klasifikasi).
3. Decision Tree – membangun model klasifikasi berbasis algoritma C4.5 dengan pendekatan *Gain Ratio*.
4. Validation – melakukan evaluasi

kinerja model menggunakan teknik *cross-validation*.

Algoritma C4.5 dipilih karena mampu menangani data numerik dan kategorikal, serta dilengkapi dengan mekanisme pruning otomatis untuk mengurangi kompleksitas model sehingga terhindar dari *overfitting*.

**Perhitungan Entropy**

Perhitungan *entropy* digunakan untuk menentukan tingkat ketidakpastian pada atribut yang dianalisis. Rumus entropy yang digunakan adalah:

$$Entropy(S) = - \sum_{i=1}^n p_i \text{LOG}_2(p_i)$$

Dimana :

*S* : Himpunan Kasus

*N* : Total partisi dari himpunan tersebut

*P<sub>i</sub>* : Menunjukkan nilai proporsional antara *S<sub>i</sub>* dan *S* secara keseluruhan.

**HASIL DAN PEMBAHASAN**

Pada bagian ini merupakan hasil dan pembahasan dari penelitian yang dilakukan menggunakan Metode Decision Tree.

**Hasil**

Bagian ini menyajikan temuan penelitian secara terstruktur melalui tabel, gambar, dan deskripsi verbal berdasarkan proses klasifikasi data rekam medis pasien menggunakan algoritma C4.5 pada RapidMiner. Dataset terdiri dari 1000 entri pasien dengan enam atribut utama: usia, tekanan darah, kadar gula, riwayat keluarga, gejala, dan status kesehatan (label).

Tabel 1. Contoh Dataset Rekam Medis Pasien

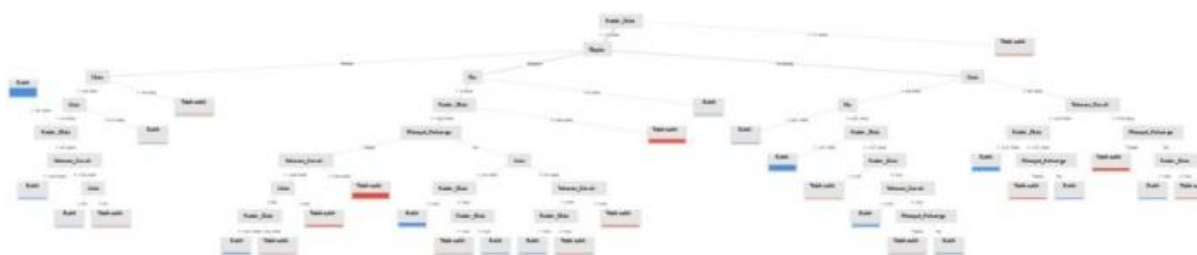
Usia	Tekanan_Darah	Kadar_Gula	Riwayat_Keluarga	Gejala	Status
40	147	98	Tidak	Parah	Sakit
55	131	194	Ya	Ringan	Sakit
69	163	235	Ya	Ringan	Sakit
65	98	178	Ya	Ringan	Sakit
34	95	247	Tidak	Sedang	Tidak sakit
56	104	156	Tidak	Ringan	Tidak sakit
73	167	199	Tidak	Parah	Sakit
77	140	155	Ya	Ringan	Sakit
21	145	137	Tidak	Ringan	Tidak sakit

Berdasarkan Tabel 1, data direpresentasikan secara ringkas dan menjadi dasar pembangunan model klasifikasi. Atribut-atribut tersebut kemudian diproses menggunakan RapidMiner untuk menentukan pola klasifikasi antara pasien “Sakit” dan “Tidak Sakit”.



**Gambar 2. Alur Proses Pembangunan Model Di RapidMiner**

Proses pemodelan dimulai dengan mengimpor dataset menggunakan operator Retrieve, kemudian menentukan peran label melalui Set Role untuk menetapkan atribut *Status* sebagai target klasifikasi. Langkah berikutnya adalah menggunakan operator Decision Tree, yaitu implementasi algoritma C4.5 di RapidMiner, untuk menghasilkan model pohon keputusan berdasarkan perhitungan *Gain Ratio*.



**Gambar 3. Pohon Keputusan (Decision Tree)**



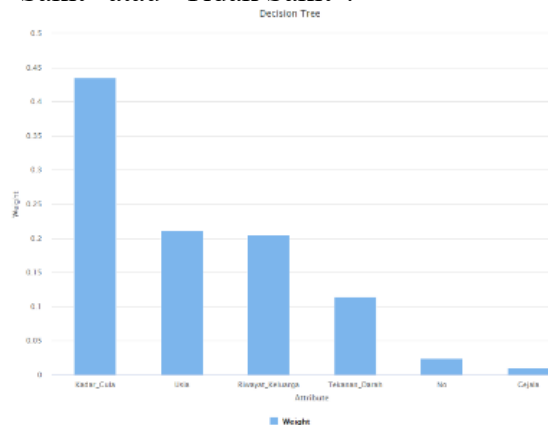
**Gambar 4. Deskripsi Node Pohon Keputusan**

Berdasarkan hasil pemodelan, atribut *Kadar\_Gula* menjadi akar utama pohon keputusan. Hal ini menunjukkan bahwa atribut tersebut memiliki nilai *information gain* paling tinggi dibanding atribut lain, sehingga paling menentukan dalam pemisahan data awal.

Selanjutnya, beberapa cabang terbentuk berdasarkan atribut:

1. Gejala, membedakan tingkat keparahan kondisi pasien.
2. Usia, sebagai faktor umum pembeda risiko penyakit.
3. Tekanan Darah, yang turut menyumbang pengelompokan risiko.
4. Riwayat Keluarga, menentukan ada tidaknya faktor keturunan terhadap penyakit tertentu.

Kombinasi variabel tersebut membentuk simpul-simpul yang menghasilkan klasifikasi akhir berupa “Sakit” atau “Tidak Sakit”.



**Gambar 5. Grafik Bobot Atribut**

Grafik bobot atribut menunjukkan kontribusi masing-masing atribut terhadap keputusan model:

1. Kadar\_Gula: bobot tertinggi (0.44) → atribut paling dominan.
2. Usia & Riwayat\_Keluarga: bobot  $\pm 0.20$  → berpengaruh signifikan.
3. Tekanan\_Darah: bobot 0.12 → pengaruh moderat.
4. Gejala dan No (ID Data):  $< 0.05$  → kontribusi rendah.

Temuan ini menegaskan bahwa kadar gula, usia, dan riwayat keluarga merupakan penentu utama status kesehatan pasien.

### Pembahasan

Pembahasan bertujuan menjelaskan makna dan implikasi hasil penelitian berdasarkan teori dan temuan terdahulu. Model klasifikasi C4.5 terbukti mampu mengolah atribut rekam medis secara efektif untuk menghasilkan pola keputusan yang logis dan mudah ditafsirkan.

Atribut Kadar\_Gula menjadi indikator paling kuat dalam penentuan status kesehatan pasien. Hal ini sejalan dengan temuan Andini & Prasetyo (2022) yang menegaskan bahwa kadar gula darah berperan penting dalam deteksi penyakit kardiovaskular dan metabolik. Sementara itu, kontribusi signifikan dari Usia dan Riwayat\_Keluarga juga mendukung penelitian Lestari & Wibowo (2023) bahwa faktor usia dan faktor keturunan merupakan variabel determinan dalam diagnosis penyakit klinis.

Struktur pohon keputusan yang dihasilkan cukup sederhana namun komprehensif, ditunjang oleh mekanisme *pruning* otomatis pada algoritma C4.5. Mekanisme ini mencegah terjadinya *overfitting*, sehingga model memiliki generalisasi yang baik terhadap data baru. Hal ini konsisten dengan pernyataan Rohman & Supriyanto (2021) bahwa teknik *pruning* efektif meningkatkan akurasi model C4.5 pada data medis.

Selain itu, penggunaan RapidMiner mempermudah proses analisis karena menyediakan alur kerja visual, sehingga

keseluruhan proses mulai dari praproses data, penentuan label, pembangunan model, hingga evaluasi dapat dilakukan secara sistematis.

Temuan penelitian ini menguatkan hasil-hasil penelitian terdahulu yang menggunakan algoritma C4.5 untuk diagnosis penyakit (Mansiz et al., 2025; Fiandra et al., 2017). Kombinasi atribut numerik dan kategorikal yang dianalisis menunjukkan bahwa algoritma C4.5 tetap relevan untuk diterapkan dalam lingkungan data medis modern yang membutuhkan model interpretatif sebagai dasar pengambilan keputusan.

### SIMPULAN (PENUTUP)

Berdasarkan hasil penelitian dan pembahasan mengenai penerapan algoritma C4.5 untuk klasifikasi data medis pasien, dapat disimpulkan bahwa model pohon keputusan yang dibangun mampu memberikan gambaran yang jelas dan terstruktur mengenai faktor-faktor yang mempengaruhi status kesehatan pasien. Atribut Kadar\_Gula merupakan variabel yang paling dominan dalam proses klasifikasi, disusul oleh Usia, Riwayat\_Keluarga, dan Tekanan\_Darah. Struktur pohon keputusan yang terbentuk menunjukkan alur pengambilan keputusan yang logis, mudah ditelusuri, serta dapat diinterpretasikan dengan baik oleh tenaga medis maupun peneliti.

Penggunaan perangkat lunak RapidMiner mampu mempermudah seluruh tahapan pemodelan, mulai dari impor data, penetapan label, pembangunan model klasifikasi, hingga evaluasi performa model. Fitur *pruning* otomatis yang terdapat pada algoritma C4.5 membantu mengurangi kompleksitas model sehingga tidak terjadi *overfitting*, sekaligus meningkatkan kemampuan generalisasi model terhadap data baru.

Secara praktis, model yang dihasilkan dapat dimanfaatkan oleh rumah sakit atau penyedia layanan kesehatan sebagai sistem pendukung keputusan untuk mendeteksi potensi penyakit secara

lebih dini berdasarkan karakteristik pasien. Model ini juga berpotensi mempercepat proses analisis medis dan membantu pengambilan keputusan berbasis data secara lebih akurat.

#### Saran

Penelitian selanjutnya dapat mengembangkan model dengan menambahkan lebih banyak atribut klinis, membandingkan performa algoritma C4.5 dengan metode lain seperti Random Forest atau Naïve Bayes, serta menerapkan dataset yang lebih beragam agar hasil klasifikasi semakin akurat dan dapat diaplikasikan pada berbagai kondisi medis yang berbeda.

#### UCAPAN TERIMA KASIH

Penulis menyampaikan ucapan terima kasih kepada Universitas Ibrahimy Situbondo yang telah memberikan dukungan dalam pelaksanaan penelitian ini. Penghargaan juga disampaikan kepada pihak-pihak yang telah menyediakan data rekam medis sehingga penelitian ini dapat terlaksana dengan baik. Selain itu, penulis berterima kasih kepada para dosen pembimbing, rekan peneliti, serta semua pihak yang telah memberikan masukan, bantuan teknis, dan dukungan selama proses penyusunan artikel ini. Semoga kontribusi tersebut menjadi bagian dari pengembangan ilmu pengetahuan di bidang sistem informasi dan data mining medis.

#### DAFTAR PUSTAKA

Budiarti, L. (2025). *Sistem Pendukung Keputusan untuk Klasifikasi Penyakit Kuning dengan Algoritma Decision Tree*. 2, 43–48.

Elektro, J. T. (2013). Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit Internasional Menggunakan Algoritma C4.5 Wenefrida Tulit Ina. *Jurnal Media Elektro*, 1(3), 105–110.

Fiandra, Y. A., Defit, S., & Yuhandri, Y. (2017). Penerapan Algoritma C4.5 untuk Klasifikasi Data Rekam Medis berdasarkan International

Classification Diseases (ICD-10). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(2), 82–89. <https://doi.org/10.29207/resti.v1i2.48>

Mansiz, M. I., Homaidi, A., Prasetyo, J. D., Informasi, T., Sains, F., & Ibrahimy, U. (2025). *Implementasi Algoritma Klasifikasi C4 . 5 Untuk Memprediksi Resiko Terkena Penyakit Paru-Paru*. 4(4), 1164–1171.

Pamungkas, W. A., Statistika, P. S., Terbuka, U., & Selatan, T. (2025). *Penerapan klasifikasi data mining untuk prediksi dengan metode algoritma decision tree*. 2(2), 655–678.

Treise, C., Brown, R. J., & Perez, J. (2019). Towards a Multi-Level Phenomenology of Delusional Disorder: The Dissociative Thought-Script. *Psychopathology*, 52(1), 50–58. <https://doi.org/10.1159/000499596>

Wulandari, D. (2025). *Analisis Penerapan Metode Decision Tree C4.5 dalam Klasifikasi Penyakit Jantung*. *Journal of Informatics and Applied Computing*, 7(2), 101–109.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Sucipto, H., & Rachman, F. (2020). *Penerapan Algoritma C4.5 untuk Klasifikasi Penyakit Jantung Berdasarkan Data Rekam Medis Pasien*. *Jurnal Teknologi dan Sistem Komputer*, 8(3), 223–230.

Yulia, R., & Santoso, B. (2019). *Penerapan Data Mining untuk Klasifikasi Penyakit Kanker Serviks Menggunakan Algoritma C4.5*. *Jurnal Rekayasa Sistem dan Teknologi Informasi (RESTI)*, 3(2), 235–242.

Andini, L., & Prasetyo, J. D. (2022). Implementasi Data Mining Menggunakan Algoritma C4.5 untuk Prediksi Penyakit Jantung. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(1), 123–130.

Lestari, N., & Wibowo, D. (2023). Analisis Prediksi Risiko Stroke

- Menggunakan Algoritma Decision Tree C4.5. *Jurnal Sistem Cerdas*, 5(2), 201–210.
- Setiawan, B., & Ramadhani, E. (2024). Penerapan Metode C4.5 untuk Klasifikasi Penyakit Demam Berdarah Berdasarkan Gejala Klinis. *Jurnal Sains Komputer dan Informatika (JSAKTI)*, 9(1), 33–41.
- Rohman, M., & Supriyanto, A. (2021). Penerapan Algoritma C4.5 untuk Klasifikasi Penyakit Diabetes Melitus. *Jurnal Teknologi dan Sistem Informasi (JTISI)*, 9(2), 87–94.
- Kusuma, D., & Arifin, M. (2020). Klasifikasi Data Rekam Medis Menggunakan Algoritma C4.5 dan Naïve Bayes. *Jurnal Ilmiah Informatika dan Komputer (JIJK)*, 8(3), 155–164.