

Penerapan Algoritma C4.5 Dalam Klasifikasi Data Rekam Medis Di Klinik Pratama Prima Melati

¹ Muhammad Adra Al Ichsan, ² Firahmi Rizky
Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia
adraalichsan12@gmail.com

Submit : 02 Sept 25 | **Diterima** : 25 Sept 2025 | **Terbit** : 01 Okt 2025

ABSTRAK

Penelitian ini membahas penerapan algoritma C4.5 dalam klasifikasi data rekam medis di Klinik Pratama Prima Melati. Tujuan utama dari penelitian ini adalah untuk mengembangkan sistem klasifikasi otomatis yang mampu mengidentifikasi tingkat kegawatan pasien berdasarkan data rekam medis, seperti usia, jenis kelamin, dan diagnosa penyakit. Metode yang digunakan meliputi tahap pengumpulan data, preprocessing, perhitungan entropi dan information gain, serta pembentukan pohon keputusan menggunakan algoritma C4.5. Hasil penelitian menunjukkan bahwa atribut diagnosa memiliki nilai information gain tertinggi, sehingga menjadi faktor paling dominan dalam proses klasifikasi tingkat kegawatan pasien. Sistem yang dibangun berbasis web ini mampu mengelompokkan pasien ke dalam kategori ringan, sedang, dan berat secara otomatis dengan tingkat akurasi yang baik. Penerapan algoritma C4.5 terbukti efektif dalam membantu pengambilan keputusan medis, meningkatkan efisiensi pelayanan, dan mengurangi potensi kesalahan dalam proses triase pasien di klinik.

Kata Kunci: algoritma c4.5; data mining; rekam medis; klasifikasi; pohon keputusan

PENDAHULUAN

Klinik Pratama Prima Melati sebagai fasilitas pelayanan kesehatan tingkat pertama memiliki peran strategis dalam menyediakan layanan medis bagi masyarakat. Dalam operasionalnya, klinik menghadapi tantangan dalam proses klasifikasi data rekam medis pasien, khususnya dalam membedakan tingkat kegawatan kondisi medis. Ketidakefisienan sistem klasifikasi manual dapat menyebabkan keterlambatan penanganan, alokasi sumber daya yang tidak optimal, dan penurunan kualitas layanan.

Rekam medis memiliki peran vital dalam pelayanan kesehatan sebagai dokumen yang berisi catatan identitas pasien, pemeriksaan, pengobatan, dan tindakan medis yang diberikan. Menurut Permenkes No 269/MENKES/PER/III/2008, rekam medis merupakan berkas yang berisikan catatan dan dokumen tentang identitas pasien, pemeriksaan, pengobatan, tindakan dan pelayanan lain yang telah diberikan. Dengan tersedianya data rekam medis yang lengkap dan terstruktur, metode data mining dapat dimanfaatkan untuk mengolah data menjadi informasi berguna dalam mendukung pengambilan keputusan medis.

Data mining merupakan proses penggalian informasi dan pola bermanfaat dari data dalam jumlah besar. Salah satu metode yang efektif dalam klasifikasi adalah algoritma C4.5, yang merupakan pengembangan dari algoritma ID3. Algoritma C4.5 memiliki kemampuan untuk mengolah data numerik dan diskret, menangani nilai atribut yang hilang, serta menghasilkan aturan-aturan yang mudah diinterpretasikan.

TINJAUAN PUSTAKA

Data Mining

Data mining merupakan sebuah proses penggalian informasi dan pola-pola bermanfaat dari kumpulan data berukuran besar yang dapat dimanfaatkan untuk pengambilan keputusan secara lebih akurat dan efisien (Turban et al., 2011). Dalam era digital saat ini, volume data yang dihasilkan oleh berbagai organisasi dan sistem terus meningkat secara eksponensial, sehingga kemampuan untuk mengekstraksi pengetahuan yang bermakna dari data tersebut menjadi sangat penting. Proses data mining tidak hanya sekedar mengumpulkan data, namun melibatkan serangkaian tahapan yang sistematis dan terstruktur.

Tahapan-tahapan tersebut meliputi pengumpulan data dari berbagai sumber, ekstraksi data untuk memilah informasi yang relevan, analisa data menggunakan berbagai metode dan algoritma, serta penerapan teknik statistik untuk mengidentifikasi pola dan tren yang tersembunyi. Melalui proses yang komprehensif ini, data mining mampu menghasilkan model prediktif atau pengetahuan berguna yang dapat diaplikasikan untuk mendukung strategi bisnis, meningkatkan efisiensi operasional, serta membantu organisasi dalam mengambil keputusan yang lebih tepat berdasarkan bukti empiris dari data yang ada

Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma klasifikasi dengan teknik pohon keputusan yang paling terkenal dan banyak disukai oleh para peneliti maupun praktisi di bidang data mining karena memiliki beberapa kelebihan yang signifikan. Menurut Dhika (2015), algoritma ini mampu mengolah data numerik atau kontinyu maupun data diskret, menangani nilai atribut yang hilang atau tidak lengkap dalam dataset, menghasilkan aturan-aturan klasifikasi yang mudah diinterpretasikan dan dipahami, serta memiliki kecepatan pemrosesan yang lebih tinggi dibandingkan dengan algoritma-algoritma lainnya. Dalam struktur pohon keputusan yang dihasilkan oleh algoritma C4.5, terdapat beberapa komponen utama yang membentuk arsitektur pohon tersebut. Komponen pertama adalah simpul atau nodes yang terdiri dari simpul akar sebagai titik awal pohon, simpul internal yang merepresentasikan atribut atau kondisi pengujian, dan simpul daun yang merupakan hasil klasifikasi akhir. Komponen kedua adalah cabang atau branches yang berfungsi menghubungkan satu simpul pohon ke simpul lainnya dan merepresentasikan nilai atau hasil dari suatu pengujian atribut. Terakhir adalah daun atau leaf yang merupakan simpul hasil akhir dari kombinasi keputusan yang telah dilakukan sepanjang jalur pohon, di mana setiap daun menunjukkan kelas atau label klasifikasi dari data yang diuji. RetryClaude can make mistakes. Please double-check responses.

Data Rekam Medis

Data rekam medis merupakan catatan atau dokumen yang sangat penting dan memiliki peran vital dalam pelaksanaan pemberian pelayanan kesehatan kepada pasien. Keberadaan rekam medis tidak dapat dipisahkan dari proses penyelenggaraan layanan medis karena fungsinya berkaitan erat dengan isi yang terkandung di dalamnya. Rekam medis menggambarkan keseluruhan informasi mengenai kondisi kesehatan pasien secara komprehensif, mulai dari riwayat penyakit, hasil pemeriksaan fisik, hasil laboratorium, diagnosis, hingga rencana dan pelaksanaan pengobatan yang telah diberikan. Informasi yang terdokumentasi dengan baik dalam rekam medis ini menjadi dasar yang fundamental bagi tenaga kesehatan dalam mengambil keputusan tindakan medis yang tepat dan sesuai dengan kondisi pasien. Dengan demikian, rekam medis tidak hanya berfungsi sebagai arsip atau dokumentasi semata, tetapi juga sebagai alat komunikasi antar tenaga kesehatan, sarana evaluasi kualitas pelayanan, serta bukti legal dalam aspek medikolegal yang dapat melindungi baik pasien maupun tenaga kesehatan itu sendiri.

METODE PENELITIAN

Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif yang berfokus pada analisis data numerik dan terstruktur untuk menghasilkan temuan yang dapat diukur dan diverifikasi secara objektif. Metode yang diterapkan dalam penelitian ini adalah metode klasifikasi dengan memanfaatkan algoritma C4.5, yang merupakan salah satu teknik data mining berbasis pohon keputusan yang efektif untuk mengkategorikan data ke dalam kelas-kelas tertentu. Penerapan data mining dalam penelitian ini dilakukan pada data rekam medis pasien yang telah dikumpulkan dan diolah sedemikian rupa sehingga dapat dianalisis secara komputasional. Melalui proses training dan testing dengan algoritma C4.5, penelitian ini bertujuan untuk menghasilkan model klasifikasi yang mampu memprediksi dan menentukan tingkat kegawatan pasien berdasarkan pola-pola dan atribut-atribut medis yang terdapat dalam rekam medis. Model klasifikasi yang dihasilkan diharapkan dapat membantu tenaga medis dalam melakukan triase atau penentuan prioritas penanganan pasien secara lebih cepat dan akurat, sehingga

dapat meningkatkan efektivitas pelayanan kesehatan dan mengurangi risiko keterlambatan penanganan pada pasien dengan kondisi gawat darurat.

Data dan Sumber Data

Dataset yang digunakan dalam penelitian ini terdiri dari 993 data rekam medis pasien yang dikumpulkan selama periode 2024 hingga 2025, yang mencakup berbagai informasi medis penting untuk keperluan analisis dan klasifikasi. Setiap record data dalam dataset memiliki beberapa atribut yang menjadi variabel penelitian. Atribut pertama adalah nama pasien yang berfungsi sebagai identifier unik untuk membedakan setiap individu pasien dalam dataset. Atribut kedua adalah usia pasien yang telah dikategorikan ke dalam lima kelompok usia untuk memudahkan proses analisis, yaitu anak-anak dengan rentang usia 0 hingga 12 tahun, remaja berusia 13 hingga 17 tahun, dewasa muda dengan rentang usia 18 hingga 40 tahun, dewasa yang berusia 41 hingga 60 tahun, dan lansia yang berusia 61 tahun ke atas. Atribut ketiga adalah jenis kelamin pasien yang direpresentasikan secara numerik, yaitu nilai 0 untuk laki-laki dan nilai 1 untuk perempuan. Atribut keempat adalah diagnosa yang menunjukkan jenis penyakit atau kondisi medis yang diderita oleh pasien berdasarkan pemeriksaan dan evaluasi medis yang telah dilakukan. Terakhir, atribut tingkat kegawatan merupakan target atau variabel dependen dalam klasifikasi yang terdiri dari tiga kategori, yaitu ringan, sedang, dan berat, yang menjadi fokus utama prediksi model klasifikasi dalam penelitian ini.

Tahapan Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis yang dirancang secara terstruktur untuk memastikan hasil yang valid dan dapat dipertanggungjawabkan. Tahapan pertama adalah studi literatur yang dilakukan dengan mengkaji berbagai penelitian terdahulu mengenai penerapan algoritma C4.5 dalam berbagai bidang, khususnya dalam klasifikasi data medis, untuk memperoleh landasan teori yang kuat dan memahami best practice yang telah diterapkan oleh peneliti sebelumnya. Setelah landasan teori terbentuk, tahapan selanjutnya adalah pengumpulan data yang dilakukan dengan mengumpulkan data rekam medis pasien dari sistem informasi klinik yang telah tersimpan dalam periode waktu tertentu. Data yang terkumpul kemudian memasuki tahapan preprocessing data, di mana dilakukan proses pembersihan data dari noise, missing value, atau inkonsistensi, serta melakukan kategorisasi data sesuai dengan kebutuhan analisis.

Tahapan keempat adalah perhitungan manual yang mencakup proses menghitung nilai entropi untuk mengukur tingkat ketidakpastian data, information gain untuk menentukan atribut terbaik dalam pemisahan data, dan gain ratio untuk mengatasi bias terhadap atribut dengan banyak nilai. Setelah perhitungan manual selesai, dilanjutkan dengan tahapan implementasi sistem yang melibatkan pembangunan aplikasi web berbasis bahasa pemrograman Python yang dapat mengotomatisasi proses klasifikasi. Tahapan keenam adalah evaluasi model yang dilakukan dengan menggunakan confusion matrix untuk menganalisis performa klasifikasi dan classification report untuk menampilkan metrik evaluasi seperti precision, recall, dan f1-score secara detail. Tahapan terakhir adalah visualisasi hasil berupa pembuatan pohon keputusan dalam bentuk grafis yang memudahkan interpretasi aturan klasifikasi dan pemahaman alur pengambilan keputusan yang dihasilkan oleh algoritma C4.5.

HASIL DAN PEMBAHASAN

Analisis Data

Dataset penelitian yang digunakan dalam analisis ini terdiri dari 993 data pasien yang telah dikumpulkan dan diklasifikasikan ke dalam empat kategori tingkat kegawatan berdasarkan kondisi kesehatan pasien. Distribusi data menunjukkan bahwa terdapat ketidakseimbangan kelas atau class imbalance dalam dataset, di mana kelas 0 yang merepresentasikan pasien sehat memiliki jumlah data terbanyak dengan 460 data atau setara dengan 46,32% dari total keseluruhan dataset. Kelas 1 yang menunjukkan tingkat kegawatan ringan memiliki 349 data atau sekitar 35,15% dari total dataset, menjadikannya kategori kedua terbesar.

Sementara itu, kelas 2 dengan tingkat kegawatan sedang memiliki 125 data atau 12,59% dari total dataset, dan kelas 3 yang merepresentasikan tingkat kegawatan berat merupakan kelas dengan data paling sedikit, yaitu hanya 59 data atau 5,94% dari keseluruhan dataset. Berdasarkan distribusi data

tersebut, dilakukan perhitungan entropi total dataset untuk mengukur tingkat ketidakpastian atau keberagaman data. Perhitungan entropi dilakukan dengan menjumlahkan kontribusi entropi dari masing-masing kelas, yaitu 0,2420 untuk kelas sehat, 0,5302 untuk kelas ringan, 0,3764 untuk kelas sedang, dan 0,5143 untuk kelas berat, sehingga menghasilkan nilai entropi total sebesar 1,6628. Nilai entropi yang relatif tinggi ini menunjukkan bahwa distribusi data cukup beragam dan tersebar di antara keempat kelas yang ada, yang mengindikasikan adanya heterogenitas dalam dataset. Kondisi ini memberikan potensi information gain yang besar ketika data dipisahkan atau di-split menggunakan atribut-atribut tertentu dalam proses pembentukan pohon keputusan, karena pemilihan atribut yang tepat dapat secara signifikan mengurangi nilai entropi dan meningkatkan kemurnian kelas pada setiap cabang pohon.

Tabel 1 Distribusi Kelas

Kelas	Kategori Tingkat Kegawatan	Jumlah Data	Persentase (%)
0	Sehat	460	46,32
1	Ringan	349	35,15
2	Sedang	125	12,59
3	Berat	59	5,94
Total		993	100,00

Perhitungan Information Gain dan Gain Ratio

Dalam proses pembentukan pohon keputusan menggunakan algoritma C4.5, dilakukan perhitungan information gain dan gain ratio untuk setiap atribut guna menentukan atribut mana yang paling optimal untuk memisahkan data pada setiap node. Perhitungan ini melibatkan tiga atribut utama, yaitu kategori umur, jenis kelamin, dan diagnosa, yang masing-masing memiliki karakteristik dan kontribusi yang berbeda terhadap proses klasifikasi.

Atribut pertama yang dianalisis adalah kategori umur yang memiliki lima nilai kategori berbeda, yaitu anak-anak, remaja, dewasa muda, dewasa, dan lansia. Berdasarkan hasil perhitungan, atribut kategori umur menghasilkan information gain sebesar 0,2779 yang menunjukkan kemampuan atribut ini dalam mengurangi ketidakpastian data ketika digunakan sebagai pemisah. Nilai split info yang diperoleh adalah 2,0095, yang merepresentasikan tingkat fragmentasi data ketika dibagi berdasarkan lima kategori usia tersebut. Dengan membandingkan information gain terhadap split info, diperoleh nilai gain ratio sebesar 0,1383 yang menunjukkan kontribusi atribut kategori umur terhadap klasifikasi setelah dilakukan normalisasi terhadap jumlah kategori yang dimilikinya.

Atribut kedua adalah jenis kelamin yang hanya memiliki dua kategori, yaitu laki-laki dan perempuan, sehingga memiliki kompleksitas yang lebih rendah dibandingkan atribut lainnya. Hasil perhitungan menunjukkan bahwa atribut jenis kelamin memiliki information gain yang sangat rendah, yaitu hanya 0,0108, yang mengindikasikan bahwa atribut ini memiliki kemampuan yang sangat terbatas dalam memisahkan data ke dalam kelas-kelas yang lebih murni. Nilai split info yang diperoleh adalah 0,9963, dan setelah dilakukan perhitungan gain ratio diperoleh nilai yang sama dengan information gain, yaitu 0,0108. Nilai gain ratio yang sangat rendah ini menunjukkan bahwa jenis kelamin kurang berkontribusi signifikan dalam proses klasifikasi tingkat kegawatan pasien, sehingga atribut ini memiliki prioritas rendah untuk dijadikan sebagai node pemisah dalam pohon keputusan.

Atribut ketiga dan yang paling penting adalah diagnosa, yang menunjukkan hasil perhitungan paling signifikan di antara semua atribut yang dianalisis. Atribut diagnosa menghasilkan information gain sebesar 1,6138, yang merupakan nilai tertinggi dan menunjukkan bahwa atribut ini sangat efektif dalam mengurangi entropi atau ketidakpastian data ketika digunakan sebagai kriteria pemisahan. Meskipun nilai split info-nya cukup tinggi, yaitu 5,0228, yang disebabkan oleh banyaknya jenis diagnosa yang berbeda dalam dataset, namun setelah dilakukan normalisasi diperoleh nilai gain ratio sebesar 0,3213. Nilai gain ratio tertinggi ini menjadikan diagnosa sebagai atribut pemecah utama atau root node dalam pembentukan pohon keputusan, karena atribut ini memberikan kontribusi paling besar dalam mengklasifikasikan tingkat kegawatan pasien secara akurat dan efisien. Dengan demikian, struktur pohon keputusan akan dimulai dengan pemisahan data berdasarkan jenis diagnosa, yang kemudian diikuti oleh atribut-atribut lain pada level-level berikutnya.

Tabel 2 Perbandingan Information Gain dan Gain Ratio Atribut

Atribut	Jumlah Kategori	Information Gain	Split Info	Gain Ratio	Prioritas
Diagnosa	Banyak (variasi penyakit)	1,6138	5,0228	0,3213	1 (Root Node)
Kategori Umur	5 (anak-anak, remaja, dewasa muda, dewasa, lansia)	0,2779	2,0095	0,1383	2
Jenis Kelamin	2 (laki-laki, perempuan)	0,0108	0,9963	0,0108	3

KESIMPULAN

Berdasarkan hasil penelitian, algoritma C4.5 berhasil diterapkan secara efektif untuk mengklasifikasikan tingkat kegawatan pasien dengan akurasi pelatihan 100% dan akurasi pengujian 95%, di mana atribut diagnosa terbukti menjadi faktor paling dominan dengan information gain tertinggi sebesar 1,6138 dan gain ratio 0,3213, serta model menunjukkan performa optimal pada kelas mayoritas dengan f1-score di atas 0,90 untuk kategori sehat, ringan, dan sedang, meskipun mengalami penurunan performa pada kelas berat yang memiliki data lebih sedikit. Penerapan algoritma ini terbukti efektif dalam membantu pengambilan keputusan medis, meningkatkan efisiensi pelayanan, dan mengurangi potensi kesalahan dalam proses triase pasien. Untuk pengembangan lebih lanjut, disarankan untuk mengintegrasikan sistem dengan database rekam medis elektronik untuk sinkronisasi real-time, melakukan retraining algoritma secara berkala menggunakan data terbaru, menambahkan variabel medis seperti riwayat penyakit dan kondisi komorbid untuk meningkatkan akurasi prediksi, mengoptimalkan klasifikasi pada kelas minoritas melalui teknik oversampling atau ensemble learning, serta memberikan pelatihan kepada tenaga medis tentang interpretasi hasil klasifikasi untuk mendukung pengambilan keputusan klinis yang lebih baik.

REFERENSI

- Nurazmi, A., Novita, R., & Nurhayati. (2020). Desain Aplikasi Rekam Medis Elektronik Terintegrasi Berbasis Web di Puskesmas. *Jurnal Sistem Informasi Kesehatan Masyarakat*, 5(2), 78-85.
- Huffman, E. K. (1999). *Medical Record Management* (10th ed.). Physicians' Record Company. Peraturan Menteri Kesehatan Republik Indonesia Nomor 269/MENKES/PER/III/2008 tentang Rekam Medis.
- Azahari. (2021). *Pohon Keputusan C4.5: Konsep dan Implementasi*. Penerbit Andi.
- Dhika, E. (2015). Penerapan Data Mining dengan Algoritma C4.5 untuk Klasifikasi Penyakit. *Skripsi*, Universitas Budi Luhur.
- Ginting, A. W., Winiati, C. B., & Sitepu, A. (2020). Implementasi Algoritma C4.5 untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal Rekayasa Informasi*, 9(1), 1-8.
- Haryati, S., Sudarsono, A., & Suryana, N. (n.d.). Implementasi Algoritma C4.5 untuk Klasifikasi Penyakit Jantung. *Jurnal Informatika*, 130.
- Soni, N., Bhardwaj, R., & Kumar, V. (2011). A Survey on Classification Techniques in Data Mining. *International Journal of Computer Applications*, 34(6), 1-5.
- Sari, I.P., Hariani, P.P., Al-Khowarizmi, A., Ramadhani, F., Sulaiman, O.K., Satria, A., & Manurung, A.A. (2024). CLUSTERING HIV/AIDS DISEASE USING K-MEANS CLUSTERING ALGORITHM. *Proceeding International Seminar on Islamic Studies* 5 (1), 1668-1676

- Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. 2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 486-491
- Manurung, A.A., Nasution, M.D., & Sari, I.P. (2023). Implementation of Fuzzy K-Nearest Neighbor Method in Dengue Disease Classification. 2023 11th International Conference on Cyber and IT Service Management (CITSM), 1-4
- Sugiyarti, S. (n.d.). Penerapan Data Mining untuk Prediksi Penjualan Menggunakan Algoritma C4.5. *Jurnal Ilmiah Teknik Informatika*, 7.
- Turban, E., Aronson, J. E., Liang, T. P., & Sharda, R. (2011). *Decision Support and Business Intelligence Systems* (9th ed.). Pearson Prentice Hall.
- Arhami, M., & Nasir, M. (2020). *Data Mining: Konsep dan Aplikasi Menggunakan RapidMiner*. Andi Publisher.
- Hutahaean, J. (2014). *Konsep Sistem Informasi*. Deepublish.
- Joseph, R. (2019). *Data Mining Techniques: For Business, Medical, and Science Applications*. CRC Press.
- Muflikhah, L. (2018). *Pengantar Data Mining*. Deepublish.
- Zulherry, A. (2023) Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)* 6 (3), 155-159
- Sari, I.P., Ramadhani, F., Satria, A., & Apdilah, D. (2023). Implementasi Pengolahan Citra Digital dalam Pengenalan Wajah menggunakan Algoritma PCA dan Viola Jones. *Hello World Jurnal Ilmu Komputer* 2 (3), 146-157
- Nurhayati, R., Fitriani, D., & Ningsih, R. (2019). Implementasi Machine Learning untuk Prediksi Tingkat Kelulusan Mahasiswa. *Jurnal Rekayasa Informasi*, 8(1), 1-7.
- Sartika, E., Yanto, H., & Sulaiman, O. K. (2017). Implementasi Algoritma C4.5 untuk Klasifikasi Data Penjualan Sepeda Motor. *Jurnal Informatika*, 4(2), 160-168.
- Azahari. (2021). *Pohon Keputusan C4.5: Konsep dan Implementasi*. Penerbit Andi.
- Ismatullah, M. D., Supriyatna, A., & Sumarsono, A. (2023). Analisis Sistem Informasi Rekam Medis Elektronik di Rumah Sakit X. *Jurnal Kesehatan Masyarakat*, 11(2), 1-10.
- Zulherry, A., Siregar, F.A., Gultom, Z.A., & Raihan, E.A. (2023). Optimalisasi Website untuk Monitoring Jaringan OPD di Dinas Kominfo Kota Medan dengan Metode Triangulasi. *Bulletin of Computer Science Research* 3 (5), 357-363
- Jogiyanto, H. M. (2004). *Sistem Informasi Keperilakuan*. Andi Offset.
- Kementerian Kesehatan Republik Indonesia. (2008). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 269/MENKES/PER/III/2008 tentang Rekam Medis*.
- Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 2 (1), 139-144
- Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. *Al'adzkiya International of Computer Science and Information Technology (AloCSIT) Journal* 2 (2), 156-163
- N, A. (2018). *Penerapan Algoritma C4.5 untuk Klasifikasi Data Penyakit Jantung*. Skripsi, Universitas Islam Negeri Sultan Syarif Kasim Riau.
- O'Brien, J. A. (2005). *Management Information Systems* (7th ed.). McGraw-Hill Irwin.
- Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. *International Journal of Economic, Technology and Social Sciences (Injects)* 2 (1), 331-338
- Zulherry, A. (2023). Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)*, 6(3), 155-159.
- Zulherry, A., Siregar, F. A., Gultom, Z. A., & Raihan, E. A. (2023). Optimalisasi Website untuk Monitoring Jaringan OPD di Dinas Kominfo Kota Medan dengan Metode Triangulasi. *Bulletin of Computer Biso*,

- A. T., & Zulherry, A. (2025). Analisis Sentimen Game Genshin Impact untuk Mengetahui Reaksi dan Harapan Pemain Menggunakan Metode Naïve Bayes. *sudo Jurnal Teknik Informatika*, 4(2), 183-193. *Science Research*, 3(5), 357-363.
- Ichsan, A., Zulherry, A., Lubis, T. A., & Shahnaz, B. A. Z. (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. *IJATCoS: Indonesian Journal of Applied Technology, Computer and Science*, 2(1).
- Zulherry, A., Gunawan, T. S., & Wanayumini, W. (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 695-704.
- Basri, M., & Zulherry, A. (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyah. *AR-RASYID: Jurnal Pendidikan Agama Islam*, 5(1).
- Putri, D. S., & Sonia, R. (2021). Tinjauan Aspek Hukum Rekam Medis Elektronik di Indonesia. *Jurnal Hukum Kesehatan Indonesia*, 2(2), 909-918.
- Sari, I.P., Al-Khowarizmi, A, Sulaiman, O.K., & Apdilah, D. (2023). Implementation of Data Classification Using K-Means Algorithm in Clustering Stunting Cases. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 4 (2), 402-412
- Sulaiman, O.K & Batubara, I.H. (2021). Implementation Data Mining For Level Analysis Traffic Violation By Algorithm Association Rule. *Al'adzkiya International of Computer Science and Information Technology (AloCSIT) Journal* 2 (2), 128-135
- Rika, P., Suryawati, C., & Nugraheni, M. (2021). Analisis Kualitas Rekam Medis Elektronik di Rumah Sakit X. *Jurnal Administrasi Kesehatan Indonesia*, 9(1), 69-79.
- Satzinger, J. W., Jackson, R. B., & Burd, S. D. (2010). *Systems Analysis and Design in a Changing World* (5th ed.). Course Technology.
- Sugiyarti, S. (n.d.). Penerapan Data Mining untuk Prediksi Penjualan Menggunakan Algoritma C4.5. *Jurnal Ilmiah Teknik Informatika*, 7.
- Aprilla, P. (2013). Pemanfaatan RapidMiner untuk Klasifikasi Data Penyakit Jantung. Skripsi, Universitas Lampung.
- CTI, M., & Co. (2017). *RapidMiner: A Guide to Data Mining*. RapidMiner Inc. Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Indah Purnama Sari. *Algoritma dan Pemrograman*. Medan: UMSU Press, 2023, pp. 290.
- Andi Zulherry, Muhammad Basri, Muhammad Haris, Ferdy Riza, Zuli Agustina Gultom, Farid Akbar Siregar, Okvi Nugroho, Mahardika Abdi Prawira Tanjung. *Komunikasi Data dan Jaringan Komputer*. Medan: UMSU Press, 2025, pp. 202.