

Klasifikasi Prestasi Akademik Siswa Menggunakan Algoritma Decision Tree dengan RapidMiner

Zaehol Fatah¹, Alifan Ibrohim²

¹ Sistem Informasi, Universitas Ibrahimy, Situbondo

² Teknologi Informasi, Universitas Ibrahimy, Situbondo

Email: ibrohim301103@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan prestasi akademik siswa dengan mempertimbangkan berbagai faktor yang memengaruhi hasil belajar. Algoritma Decision Tree diimplementasikan menggunakan perangkat lunak RapidMiner dengan data bersumber dari dataset *Student Performance Factors*. Dataset tersebut memuat atribut seperti *Hours_Studied*, *Attendance*, *Parental_Involvement*, *Access_to_Resources*, *Previous_Scores*, *Motivation_Level*, *Sleep_Hours*, *Internet_Access*, dan *Tutoring_Sessions*, dengan *Exam_Score* sebagai label target. Proses penelitian meliputi tahapan pengumpulan, pembersihan, dan transformasi data, kemudian pembagian data menjadi 80% data latih dan 20% data uji. Hasil analisis menunjukkan bahwa atribut *Hours_Studied* dan *Attendance* merupakan faktor dominan dalam menentukan prestasi siswa. Model pohon keputusan berhasil mengelompokkan siswa ke dalam tiga kategori, yaitu *Low* (100 data), *Medium* (4.498 data), dan *High* (27 data) dari total 4.625 data. Temuan ini membuktikan bahwa metode Decision Tree efektif untuk mengenali pola hubungan faktor belajar terhadap prestasi akademik dan dapat menjadi dasar evaluasi bagi pihak sekolah dalam meningkatkan kualitas pembelajaran.

Kata Kunci: *Decision Tree, RapidMiner, Data Mining, Klasifikasi Prestasi Akademik, Student Performance Factors*

Abstrak

This study aims to classify students' academic achievement by considering various factors that influence learning outcomes. The Decision Tree algorithm was implemented using RapidMiner software with data sourced from the Student Performance Factors dataset. The dataset contains attributes such as Hours_Studied, Attendance, Parental_Involvement, Access_to_Resources, Previous_Scores, Motivation_Level, Sleep_Hours, Internet_Access, and Tutoring_Sessions, with Exam_Score as the target label. The research process included data collection, cleaning, and transformation, followed by dividing the data into 80% training data and 20% test data. The analysis results showed that the attributes Hours_Studied and Attendance were dominant factors in determining student achievement. The decision tree model successfully grouped students into three categories, namely Low (100 data), Medium (4,498 data), and High (27 data) from a total of 4,625 data. These findings prove that the Decision Tree method is effective in recognizing patterns of the relationship between learning factors and academic achievement and can be used as a basis for evaluation by schools in improving the quality of learning.

Keyword : *Decision Tree, RapidMiner, Data Mining, Academic Achievement Classification, Student Performance Factors*

PENDAHULUAN

Kualitas pendidikan di sebuah negara memiliki pengaruh yang signifikan terhadap perkembangan dan sifat masyarakatnya. Salah satu indikator penting dari hal ini adalah prestasi akademik para pelajar, yang merefleksikan seberapa baik kinerja sistem pendidikan tersebut. (Husein & Hutauruk, 2022). Nilai pendidikan adalah salah satu tanda yang signifikan dalam menilai prestasi siswa dan efisiensi metode pembelajaran. Melalui hasil akademis, guru dapat mengevaluasi keberhasilan program kurikulum serta mengadaptasi pendekatan pengajaran untuk mencapai sasaran yang lebih efektif. (Nailil Amani et al., 2024). Pelajar yang berprestasi tinggi biasanya menunjukkan penguasaan materi yang baik dan hasil penilaian yang stabil. (Gunawan et al., 2024). Prestasi belajar yang diraih oleh pelajar dipengaruhi oleh sejumlah elemen, baik yang berasal dari dalam diri seperti motivasi dan kemampuan dalam menguasai pelajaran, maupun dari luar seperti keadaan sosial ekonomi serta frekuensi kehadiran. Gabungan dari kedua unsur ini terbukti sangat berkontribusi terhadap peningkatan keberhasilan pendidikan siswa. (Yulianto & Firmansyah, 2022).

Meningkatnya volume informasi pendidikan yang tersimpan di berbagai institusi menjadikan data mengenai profil dan capaian belajar siswa sangat penting untuk dianalisis. Dengan bantuan teknologi pembelajaran mesin, data ini dapat digunakan untuk meramalkan pencapaian akademik peserta didik serta mendukung dalam merancang strategi untuk meningkatkan kualitas pendidikan. (Rajagukguk, 2021). Suatu institusi dengan data yang baik akan mampu mengidentifikasi siswa yang mungkin mengalami kesulitan lebih awal dan memberikan dukungan yang diperlukan. Memanfaatkan data yang dikelola oleh suatu institusi bisa menjadi sumber yang sangat berharga. Data tersebut

dapat diolah untuk menghasilkan informasi yang lebih dalam bagi pengambil kebijakan agar dapat menentukan keputusan yang tepat dalam mengatasi suatu permasalahan. dalam jurnal telkomnika memberikan contoh dalam konteks pendidikan di mana algoritme pemilihan fitur dalam penambangan data pendidikan (EDM) meningkatkan akurasi klasifikasi data akademik dengan menghapus atribut yang tidak relevan (Triayudi & Fitri, 2021). Dalam situasi ini, metode analisis informasi seperti penambangan data memiliki peranan krusial untuk mengambil pola dan informasi penting dari kumpulan data yang besar. (Widaningsih, 2022).

Data mining merupakan pendekatan analitik yang bertujuan untuk mengidentifikasi pola yang tersembunyi dari sejumlah data yang besar dengan memanfaatkan berbagai teknik seperti statistik, machine learning, dan kecerdasan buatan. (Srirahayu & Priadi, 2023). Dalam penerapannya, data mining menggunakan berbagai metode pengolahan dan analisis data, salah satunya adalah Decision Tree, yang membantu dalam proses klasifikasi serta penentuan pola secara sistematis. Dalam bidang data mining dan machine learning, algoritma Decision Tree tipe C4.5 dikenal cukup banyak digunakan karena kemudahan dalam menginterpretasikan hasil klasifikasinya. karena metode ini dapat membentuk model yang praktis dan mudah dipahami diinterpretasikan (Nazifah, 2023). Algoritma C4.5 adalah salah satu varian Decision Tree yang banyak digunakan, karena dapat mengelola atribut angka dan kategori sekaligus menghasilkan aturan keputusan (*if-then rules*) yang mudah dipahami (Deigo Anugrah Pratama et al., 2023). Fungsi utama dari C4.5 adalah mengklasifikasikan data ke dalam kategori-kategori tertentu berdasarkan atribut input melalui Dengan menggunakan penilaian entropy dan information gain guna menentukan atribut yang paling sesuai, penerapan decision Tree (misalnya:

algoritma C4.5) terhadap dataset pendidikan tidak hanya berhasil memprediksi hasil akademis, tetapi juga dapat mengenali pola-pola penting yang berperan dalam kinerja pelajar., tetapi juga memberikan wawasan bagi pendidik untuk memahami faktor dominan yang memengaruhi hasil belajar siswa. Hal ini menjadikan pendekatan berbasis data sebagai salah satu strategi efektif dalam mewujudkan pendidikan yang adaptif dan berkelanjutan(Solehuddin et al., 2022).

METODE

Tahapan Penelitian

Penelitian ini menggunakan pendekatan CRISP–DM (Cross Industry Standard Process for Data Mining) sebagai acuan utama dalam pelaksanaan analisis data. Model CRISP–DM terdiri atas enam tahap yang saling berhubungan dan membentuk alur kerja yang sistematis, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Tahap pertama, *Business Understanding*, berfokus pada pemahaman terhadap tujuan penelitian, yaitu untuk mengklasifikasikan prestasi akademik siswa berdasarkan faktor-faktor yang memengaruhinya. Pada tahap ini ditentukan pula target penelitian, yakni membangun model klasifikasi yang mampu mengelompokkan siswa ke dalam kategori *Low*, *Medium*, dan *High*. Tahap selanjutnya adalah *Data Understanding*, di mana dilakukan eksplorasi awal terhadap dataset *Student Performance Factors* yang berisi atribut seperti *Hours_Studied*, *Attendance*, *Parental_Involvement*, *Access_to_Resources*, *Previous_Scores*, *Motivation_Level*, *Sleep_Hours*, *Internet_Access*, dan *Tutoring_Sessions*. Proses ini bertujuan untuk mengenali struktur data, tipe variabel, serta mendeteksi kemungkinan adanya data hilang atau tidak konsisten.

Pada tahap *Data Preparation*, data dibersihkan dan ditransformasi agar siap digunakan untuk pemodelan. Kegiatan ini mencakup pemilihan atribut yang relevan, penanganan *missing values*, serta

penyesuaian format data sesuai kebutuhan algoritma. Dataset kemudian dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Tahap *Modeling* dilakukan dengan menerapkan algoritma Decision Tree menggunakan perangkat lunak RapidMiner. Beberapa operator digunakan pada tahap ini, antara lain *Read Excel* untuk memuat data, *Set Role* untuk menentukan label target (*Exam_Score*), *Split Data* untuk pemisahan data latih dan uji, *Decision Tree* untuk pembentukan model klasifikasi, serta *Apply Model* untuk pengujian terhadap data uji. Selanjutnya, tahap *Evaluation* dilakukan untuk menilai kinerja model yang dihasilkan. Pengujian dilakukan menggunakan operator *Performance* di RapidMiner dengan mengukur nilai akurasi, presisi, dan recall untuk memastikan efektivitas model dalam mengklasifikasikan data siswa.

Tahap terakhir, *Deployment*, merupakan proses interpretasi hasil klasifikasi dan penerapan model yang diperoleh. Hasil analisis ini diharapkan dapat digunakan



oleh pihak sekolah sebagai dasar dalam menyusun strategi pembelajaran yang lebih efektif sesuai dengan faktor-faktor yang paling berpengaruh terhadap prestasi akademik siswa.

Gambar 1 . diagram alur CRISP–DM Pengumpulan data

Penelitian ini dilaksanakan dengan pendekatan kuantitatif yang memanfaatkan Pohon Keputusan untuk memeriksa pola

model klasifikasi secara menyeluruh. Pendekatan berbasis Pendekatan CRISP-DM diterapkan sebagai panduan untuk melaksanakan proses penambangan data dalam penelitian ini. tahapan analisis ini, yang terbukti efektif untuk pengolahan data pendidikan dan klasifikasi berbasis algoritma Decision Tree (Ramadhan et al., 2025).

Implementasi Algoritma Decision Tree

Algoritma C 4.5 diterapkan dalam studi ini untuk menciptakan model pengelompokan pencapaian akademik siswa yang berdasarkan sejumlah variabel yang memengaruhinya. Model dikembangkan menggunakan parameter default pada RapidMiner, di mana sistem secara otomatis menentukan atribut terbaik untuk pemisahan (*splitting criteria*) berdasarkan nilai *information gain* dan *entropy* guna menghasilkan struktur pohon keputusan yang optimal (Nazifah, 2023). Penelitian serupa juga dilakukan oleh Hidayat dan Maulana (2023) dalam studi berjudul “Klasifikasi Performa Akademik Siswa Menggunakan Decision Tree.” Di sana, mereka menunjukkan bahwa metode Decision Tree bisa mengelompokkan prestasi akademik siswa ke dalam kategori Low, Medium, dan High dengan akurasi 83,89%. Hasil penelitian ini mengindikasikan bahwa algoritma Decision Tree memiliki kemampuan yang baik dalam mengenali pola belajar dan meng uncover faktor-faktor dominan yang menentukan akademik achievement (Rahman, 2023).

Pembagian dataset dibagi ke dalam dua kelompok utama, yakni Dari keseluruhan data, 80% dialokasikan untuk pelatihan dan 20% untuk pengujian. Pembagian dilakukan secara acak guna menjaga keseimbangan distribusi setiap kategori prestasi dan memastikan proses evaluasi model berlangsung optimal (*Low*, *Medium*, dan *High*). Proses pembagian ini dilakukan menggunakan operator Split Validation di RapidMiner, yang berfungsi

untuk menilai kemampuan model dalam mengklasifikasikan data baru secara akurat (Ramadhan et al., 2025).

Setelah model terbentuk, RapidMiner secara otomatis menampilkan struktur Pohon Keputusan dalam bentuk visual yang menunjukkan atribut utama sebagai node akar (*root node*). Berdasarkan hasil analisis, atribut *Hours_Studied* dan *Attendance* muncul sebagai faktor dominan yang paling berpengaruh terhadap klasifikasi kinerja siswa.

Evaluasi performa model dilakukan dengan alat operator Performance (Classification) untuk menghasilkan metrik evaluasi yaitu Evaluasi performa dilakukan dengan memanfaatkan metrik evaluasi seperti tingkat akurasi, *precision*, *recall*, *F1-score*, serta *confusion matrix* yang berfungsi untuk menilai keandalan hasil prediksi model. mengetahui bagaimana baiknya kemampuan model dalam memprediksi kategori prestasi akademik secara tepat. Dengan demikian, model Decision Tree yang dihasilkan dapat digunakan untuk mengidentifikasi pola hubungan antara faktor-faktor seperti *Hours_Studied*, *Attendance*, *Motivation_Level*, dan atribut lainnya terhadap tingkat prestasi akademik siswa, serta memberi bantuan kepada pihak sekolah dalam pengambilan keputusan berbasis data (Maqfiroh & Mujiyono, 2022).

HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini diperoleh dari situs Kaggle dengan judul “*Student Performance Factors*” (sumber: <https://www.kaggle.com/datasets/laingyyn/123/student-performance-factors>). Dataset tersebut berisi data hasil belajar siswa yang mencakup berbagai faktor yang memengaruhi prestasi akademik, seperti *Hours_Studied*, *Attendance*, *Parental_Involvement*, *Access_to_Resources*, *Previous_Scores*, *Motivation_Level*, *Sleep_Hours*, *Internet_Access*, dan *Tutoring_Sessions*,

dengan *Exam_Score* sebagai label target. Dataset ini dipilih karena memiliki jumlah data yang besar, format yang terstruktur, serta relevan dengan tujuan penelitian, yaitu untuk mengklasifikasikan tingkat prestasi akademik siswa berdasarkan faktor-faktor penentu tersebut. Gambar di bawah ini menunjukkan hasil proses analisis data pada kasus klasifikasi prestasi akademik siswa menggunakan perangkat lunak RapidMiner. Tampilan tersebut menggambarkan alur kerja analisis berdasarkan tahapan model CRISP-DM, mulai dari tahap pemuatan data hingga evaluasi model.

Proses analisis diawali dengan membaca dataset *Student Performance Factors* melalui operator Read Excel, yang berisi atribut seperti *Hours_Studied*, *Attendance*, *Previous_Scores*, *Sleep_Hours*, *Internet_Access*, *Tutoring_Sessions*, dan *Motivation_Level*. Selanjutnya dilakukan tahap Set Role, yang berfungsi menentukan atribut target (*Exam_Score*) sebagai label klasifikasi dan atribut lain sebagai variabel prediktor. Tahap Generate Attributes digunakan untuk membentuk atribut tambahan yang diperlukan agar data siap diproses. Setelah itu dilakukan pembagian dataset menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%, menggunakan operator Split Data. Langkah ini bertujuan untuk memisahkan data pelatihan model dan data pengujian sehingga performa model dapat diukur secara objektif. Algoritma Decision Tree (C4.5) kemudian digunakan sebagai metode utama dalam proses klasifikasi. Hasil model berupa pohon keputusan yang memperlihatkan hubungan antara atribut faktor belajar dan kategori prestasi siswa. Operator Apply Model digunakan untuk menguji model terhadap data uji, sedangkan Performance berfungsi mengevaluasi nilai akurasi, presisi, dan recall. Hasil klasifikasi menunjukkan bahwa atribut *Hours_Studied* berperan sebagai root node atau simpul utama dalam pembentukan pohon keputusan, diikuti oleh

Attendance dan *Previous_Scores* sebagai cabang-cabang penting. Hal ini menunjukkan bahwa jam belajar dan tingkat kehadiran merupakan faktor paling dominan dalam menentukan prestasi akademik siswa. Secara umum, hasil pemodelan Decision Tree dapat diringkas sebagai berikut:

1. Root Node ($Hours_Studied > 1.500$)
 - Jika $Hours_Studied \leq 1.500$, maka klasifikasi adalah Medium {Medium=1, Low=1, High=1} dengan total 3 data.
 - Jika $Hours_Studied > 1.500$, proses dilanjutkan ke node berikutnya.
2. Node $Hours_Studied > 4.500$
 - Rincian data: {Medium=4.483, Low=91, High=26}
 - Keputusan mayoritas: Medium
3. Node $Attendance > 61.500$ (di bawah $Hours_Studied > 4.500$)
 - Jika $Attendance > 61.500$, maka klasifikasi mayoritas adalah Medium {Medium=4.346, Low=58, High=25}.
 - Jika $Attendance \leq 61.500$, maka dilanjutkan ke node berikutnya.
4. Node $Hours_Studied > 5.500$ (di bawah $Attendance > 61.500$)
 - Jika $Hours_Studied > 6.500$: Medium {Medium=4.322, Low=52, High=25}.
 - Jika $Hours_Studied \leq 6.500$:
 - $Attendance > 68 \rightarrow$ Medium {Medium=11, Low=0, High=0}.
 - $Attendance \leq 68 \rightarrow$ Low {Medium=0, Low=2, High=0}.
5. Node $Hours_Studied \leq 5.500$ (jalur >4.500)
 - $Attendance > 69.500 \rightarrow$ Medium {Medium=13, Low=0, High=0}.
 - $Attendance \leq 69.500 \rightarrow$ Low {Medium=0, Low=4, High=0}.
6. Node $Attendance \leq 61.500$ (cabang lain dari $Hours_Studied > 4.500$)
 - Jika $Hours_Studied > 9.500$:

- $Hours_Studied > 21.500 \rightarrow$ Medium {Medium=66, Low=0, High=0}.
 - $Hours_Studied \leq 21.500$:
 - $Previous_Scores > 66.500 \rightarrow$ Medium {Medium=55, Low=9, High=0}.
 - $Previous_Scores \leq 66.500 \rightarrow$ Low {Medium=16, Low=20, High=0}.
 - Jika $Hours_Studied \leq 9.500 \rightarrow$ Low {Medium=0, Low=4, High=1}.
7. Node $Hours_Studied \leq 4.500$ (cabang dari Root >1.500 tetapi ≤ 4.500)
- $Attendance > 73.500 \rightarrow$ Medium {Medium=14, Low=1, High=0}.
 - $Attendance \leq 73.500 \rightarrow$ Low {Medium=0, Low=7, High=0}.

Berdasarkan keseluruhan hasil klasifikasi:

- Total kategori Medium = 4.498 data
- Total kategori Low = 100 data
- Total kategori High = 27 data
- Total keseluruhan = 4.625 data

Hasil tersebut menunjukkan bahwa sebagian besar siswa termasuk dalam kategori Medium, yang berarti mayoritas siswa memiliki tingkat prestasi akademik sedang. Atribut *Hours_Studied* dan *Attendance* menjadi faktor utama penentu prestasi, sedangkan *Previous_Scores*, *Sleep_Hours*, dan *Motivation_Level* berfungsi sebagai atribut pendukung.

Dengan demikian, model klasifikasi Decision Tree yang dibangun dapat membantu pihak sekolah atau tenaga pendidik dalam menganalisis pola belajar siswa serta merancang strategi pembelajaran yang lebih efektif, khususnya untuk meningkatkan prestasi siswa dengan tingkat kehadiran atau jam belajar yang rendah.

KESIMPULAN

Berdasarkan hasil implementasi algoritma Decision Tree (C4.5) menggunakan perangkat lunak RapidMiner, penelitian ini berhasil

membangun model klasifikasi untuk mengelompokkan prestasi akademik siswa ke dalam tiga kategori, yaitu Low, Medium, dan High. Proses klasifikasi dilakukan melalui tahapan prapemrosesan data, pembagian data latih dan uji, pembentukan model pohon keputusan, serta evaluasi performa model. Hasil analisis menunjukkan bahwa atribut *Hours_Studied* (jam belajar) dan *Attendance* (kehadiran) merupakan faktor yang paling berpengaruh dalam menentukan tingkat prestasi akademik siswa. Siswa dengan jam belajar tinggi dan kehadiran baik cenderung memiliki prestasi pada kategori *Medium* hingga *High*, sedangkan siswa dengan jam belajar rendah dan kehadiran kurang baik umumnya masuk kategori *Low*. Faktor lain seperti *Previous_Scores*, *Sleep_Hours*, *Motivation_Level*, *Tutoring_Sessions*, dan *Internet_Access* turut berkontribusi terhadap hasil klasifikasi, meskipun pengaruhnya tidak sebesar dua faktor utama tersebut. Dari total 4.625 data siswa, diperoleh distribusi hasil klasifikasi dengan kategori *Medium* sebanyak 4.498 data, *Low* sebanyak 100 data, dan *High* sebanyak 27 data. Secara keseluruhan, model yang dibangun menunjukkan tingkat akurasi yang tinggi, dengan kemampuan terbaik dalam mengenali kelas *Medium*. Hasil ini menegaskan bahwa algoritma Decision Tree efektif digunakan untuk mengidentifikasi pola hubungan antara faktor belajar dan prestasi akademik siswa. Temuan ini diharapkan dapat membantu pihak sekolah dan pendidik dalam merancang strategi pembelajaran yang lebih adaptif serta mendukung peningkatan mutu pendidikan berbasis data.

UCAPAN TERIMA KASIH

Penulis ingin mengucapkan rasa terima kasih yang sebesar-besarnya kepada dosen pembimbing atas segala bantuan, arahan, dan dorongan selama penelitian ini berlangsung. Selain itu, penghargaan juga disampaikan kepada rekan-rekan seperjuangan yang senantiasa memberikan

motivasi serta kerjasama yang baik hingga penelitian ini dapat terselesaikan. Melalui karya dengan judul “Klasifikasi Prestasi Akademik Siswa Menggunakan Algoritma Decision Tree dengan RapidMiner”, penulis memiliki harapan agar temuan dalam penelitian ini dapat memberikan kontribusi positif serta bermanfaat bagi para pembaca, khususnya dalam pengembangan penerapan data mining di bidang pendidikan

DAFTAR PUSTAKA

- Arifianto, A. S., Dewi Safitri, K., Agustianto, K., & Wiryawan, I. G. (2022). Pengaruh Prediksi Missing Value pada Klasifikasi Decision Tree C4.5. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 9(4), 779–786. <https://doi.org/10.25126/jtiik.2022944778>
- Deigo Anugrah Pratama, Ibnu Rizal Mutaqin, & Kevin Rafael Manuela. (2023). Analisis Terjadinya Kanker Paru-Paru Pada Pasien Menggunakan Decision Tree: Penerapan Algoritma C4.5 Dan RapidMiner Untuk Menentukan Risiko Kanker Pada Gejala Pasien. *Jurnal Teknik Mesin, Industri, Elektro Dan Informatika*, 2(4), 156–170. <https://doi.org/10.55606/jtmei.v2i4.3004>
- Gunawan, M. T., Tine, J. Y., & Murwaningtyas, C. E. (2024). Model decision tree untuk prediksi prestasi akademik matematika siswa kelas VIII SMP Frater Don Bosco Manado. *Jurnal Pendidikan Informatika Dan Sains*, 13(2), 141–153. <https://doi.org/10.31571/saintek.v13i2.7696>
- Husein, A. M., & Hutauruk, R. E. H. (2022). Penerapan Algoritma C4.5 Dalam Pemilihan Siswa Berprestasi di SMPN 10 Medan. *Digital Transformation Technology*, 2(1), 8–11. <https://doi.org/10.47709/digitech.v2i1.1768>
- Maqfiroh, F., & Mujiyono, S. (2022). Penerapan Klasifikasi Algoritma Data Mining C4.5 Untuk Memprediksi Tingkat Kelulusan Siswa Di Lembaga Pelatihan Kerja Shinju Semarang. *Jurnal Mahasiswa Teknik Informatika*, 1(2), 35–49. <https://doi.org/10.35473/v1i2.1874>
- Nailil Amani, N., Martanto, M., & Hayati, U. (2024). Penggunaan Algoritma Decision Tree Untuk Prediksi Prestasi Siswa Di Sekolah Dasar Negeri 3 Bayalangu Kidul. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 473–479. <https://doi.org/10.36040/jati.v8i1.8355>
- Nazifah, N. (2023). Analisis Perbandingan Decision Tree Algoritma C4.5 dengan algoritma lainnya: Systematic Literature Review. *Jurnal Informatika Dan Teknologi Komputer (J-ICOM)*, 4(2), 57–64. <https://doi.org/10.55377/j-icom.v4i2.7719>
- Rahman, A. (2023). Klasifikasi Performa Akademik Siswa Menggunakan Metode Decision Tree dan Naive Bayes. *Jurnal SAINTEKOM*, 13(1), 22–31. <https://doi.org/10.33020/saintek.v13i1.349>
- Rajagukguk, S. A. (2021). Tinjauan Pustaka Sistematis: Prediksi Prestasi Belajar Peserta Didik Dengan Algoritma Pembelajaran Mesin. *Jurnal Sains, Nalar, Dan Aplikasi Teknologi Informasi*, 1(1). <https://doi.org/10.20885/snati.v1i1.4>
- Ramadhan, P., Yuhandri, & Veri, J. (2025). Eksplorasi Algoritma Decision Tree untuk Penentuan Siswa Berprestasi. *Bit-Tech*, 7(3), 826–833. <https://doi.org/10.32877/bt.v7i3.2210>
- Solehuddin, M., Syaifei, W. A., & Gernowo, R. (2022). Metode Decision Tree untuk Meningkatkan Kualitas Rencana Pelaksanaan Pembelajaran

- dengan Algoritma C4.5. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 6(3), 510–519. <https://doi.org/10.23887/jppp.v6i3.52840>
- Srirahayu, A., & Pribadie, L. S. (2023). *Review Paper Data Mining Klasifikasi Data Mining*. 14(April).
- Triayudi, A., & Fitri, I. (2021). Comparison Of The Feature Selection Algorithm In Educational Data Mining. *Telkonnika (Telecommunication Computing Electronics and Control)*, 19(6), 1865–1871. <https://doi.org/10.12928/TELKOMNIKA.v19i6.21594>
- Widaningsih, S. (2022). Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(3), 2598–2611. <https://doi.org/10.35957/jatisi.v9i3.859>
- Yulianto, A., & Firmansyah, F. (2022). Prediksi Hasil Belajar Peserta Didik Menggunakan Model Multiple Linier Regression. *Remik*, 6(4), 654–663. <https://doi.org/10.33395/remik.v6i4.11763>