

Klasifikasi Tingkat Kedisiplinan Siswa Menggunakan Algoritma Machine Learning: Decision Tree, KNN, dan Naive Bayes

Damri Mulia Hutabalian^{1,*}, Pebruarianto Hutabarat², Mhd Prasetyo³, Mhd Agung Irnanda⁴, Naufal Dhiya Putra Dalimunthe⁵, Rika Rosnelly⁶

^{1,2,3,4,5,6} Fakultas Teknik dan Ilmu Komputer, Ilmu Komputer, Universitas Potensi Utama, Kota Medan, Indonesia

Email: ^{1*}damrinaingolan23@gmail.com, ²rintohutabarat20@gmail.com, ³dundarsihite@gmail.com,

⁴agungirnanda16@gmail.com, ⁵naufaldhiyaputradalimunthe@gmail.com, ⁶rikarosnelly@gmail.com

(* Email Corresponding Author: damrinaingolan23@gmail.com)

Received: January 8, 2026 | Revision: January 10, 2026 | Accepted: January 10, 2026

Abstrak

Kedisiplinan merupakan faktor penting yang memengaruhi keberhasilan proses pembelajaran dan kualitas lulusan di sekolah kejuruan. SMK Swasta RK Bintang Timur Pematangsiantar memiliki data absensi dan prestasi akademik yang berpotensi untuk dianalisis sebagai indikator perilaku kedisiplinan siswa. Namun, data tersebut selama ini belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan untuk deteksi dini siswa yang berisiko mengalami penurunan kedisiplinan. Penelitian ini bertujuan untuk mengembangkan model prediksi kedisiplinan siswa dengan mengenali pola absensi dan prestasi akademik menggunakan pendekatan data mining. Metode penelitian menggunakan kerangka kerja CRISP-DM yang meliputi tahapan business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Dataset penelitian terdiri atas data absensi harian, nilai hasil belajar setiap mata pelajaran per semester, serta catatan kedisiplinan siswa yang dijadikan sebagai label kelas. Beberapa algoritma klasifikasi seperti Decision Tree (C4.5), KNN, dan Naive Bayes digunakan untuk membandingkan performa model. Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score melalui teknik k-fold cross validation. Hasil penelitian menunjukkan bahwa pola absensi dan prestasi akademik memiliki pengaruh signifikan dalam memprediksi tingkat kedisiplinan siswa. Variabel yang paling berkontribusi terhadap prediksi antara lain persentase kehadiran, jumlah ketidakhadiran tanpa keterangan, dan rata-rata nilai akademik. Model yang dihasilkan kemudian diimplementasikan ke dalam bentuk prototype dashboard sistem pendukung keputusan untuk memantau potensi pelanggaran dan memberikan rekomendasi intervensi kepada guru BK dan wali kelas. Penelitian ini diharapkan mampu menjadi dasar pengembangan sistem monitoring kedisiplinan berbasis data di lingkungan sekolah, serta memberikan manfaat praktis dalam mendukung tindakan preventif dan peningkatan kualitas perilaku siswa di SMK Swasta RK Bintang Timur Pematangsiantar.

Kata Kunci: Data Mining, Prediksi Kedisiplinan, Absensi Siswa, Prestasi Akademik, CRISP-DM, SMK.

Abstract

Discipline is a crucial factor influencing the effectiveness of learning processes and the quality of graduates in vocational education. SMK Swasta RK Bintang Timur Pematangsiantar maintains records of student attendance and academic performance that have the potential to be analyzed as indicators of student discipline. However, these data have not been optimally utilized as a basis for decision-making to provide early detection of students who are at risk of declining discipline. This research aims to develop a predictive model of student discipline by identifying patterns of attendance and academic achievement using a data mining approach. The study employs the CRISP-DM framework, consisting of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The dataset includes daily attendance records, semester academic grades, and documented disciplinary behavior used as class labels. Several classification algorithms—Decision Tree (C4.5), KNN, Naive Bayes were implemented to compare model performance. Model evaluation was conducted using confusion matrix, accuracy, precision, recall, and F1-score, with k fold cross-validation. The results show that attendance and academic performance patterns significantly influence the prediction of student discipline levels. The Random Forest algorithm produced the highest performance results, with consistent F1-scores for at-risk student categories. The most influential features include attendance percentage, the number of unexcused absences, and average academic scores. The resulting model is implemented as a decision support prototype dashboard to assist counseling teachers and homeroom teachers in monitoring potential disciplinary violations and planning early intervention. This research is expected to support the development of data-driven discipline monitoring systems in schools and provide practical benefit in preventive actions to improve student behavior quality at SMK Swasta RK Bintang Timur Pematangsiantar

Keywords: Data Mining, Discipline Prediction, Student Attendance, Academic Performance, CRISP-DM, Vocational High School.

1. PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) merupakan satuan pendidikan yang memiliki peran strategis dalam menyiapkan sumber daya manusia yang kompeten, terampil, dan siap memasuki dunia kerja[1]. Tidak hanya dituntut memiliki penguasaan kompetensi akademik dan keahlian vokasional, lulusan SMK juga diharapkan memiliki sikap, karakter, dan kedisiplinan yang baik. Kedisiplinan menjadi salah satu aspek fundamental dalam proses pendidikan karena berkaitan langsung dengan sikap tanggung jawab, kepatuhan terhadap aturan, manajemen waktu, serta konsistensi

perilaku siswa dalam kegiatan belajar mengajar[2]. Dalam konteks dunia industri dan kerja, kedisiplinan merupakan indikator utama profesionalisme, sehingga pembinaan disiplin sejak di bangku sekolah menjadi hal yang sangat penting[3].

Namun, pada praktiknya, permasalahan kedisiplinan siswa masih menjadi tantangan yang sering dihadapi oleh pihak sekolah, termasuk di lingkungan Sekolah Menengah Kejuruan[4]. Bentuk permasalahan tersebut antara lain keterlambatan hadir di sekolah, ketidakhadiran tanpa keterangan (alfa), pelanggaran tata tertib, rendahnya partisipasi dalam kegiatan pembelajaran, hingga penurunan prestasi akademik. Jika tidak ditangani sejak dini, permasalahan kedisiplinan dapat berdampak negatif terhadap iklim belajar di sekolah, prestasi siswa secara individu, serta citra dan kualitas lulusan sekolah secara keseluruhan[5].

SMK Swasta RK Bintang Timur Pematangsiantar sebagai salah satu institusi pendidikan kejuruan memiliki jumlah data siswa yang cukup besar, terutama data absensi harian, nilai akademik, serta catatan perilaku siswa[6]. Data-data tersebut pada dasarnya menyimpan informasi penting yang dapat dimanfaatkan untuk mengidentifikasi pola perilaku siswa, khususnya yang berkaitan dengan tingkat kedisiplinan[7]. Akan tetapi, selama ini data tersebut umumnya hanya digunakan untuk keperluan administratif, seperti rekap kehadiran dan pelaporan nilai, tanpa dilakukan analisis lebih lanjut untuk mendukung pengambilan keputusan strategis. Akibatnya, potensi data sebagai sumber informasi untuk deteksi dini siswa berisiko belum dimanfaatkan secara optimal[8].

Perkembangan teknologi informasi dan komputasi, khususnya dalam bidang data mining dan machine learning, membuka peluang besar untuk mengolah dan menganalisis data pendidikan secara lebih sistematis dan mendalam[8], [9], [10]. Data mining memungkinkan proses ekstraksi pengetahuan dari kumpulan data dalam jumlah besar untuk menemukan pola, hubungan, dan kecenderungan yang sebelumnya tidak terlihat secara langsung. Dalam konteks pendidikan, pendekatan ini dikenal sebagai *educational data mining*, yang bertujuan untuk meningkatkan kualitas proses pembelajaran, manajemen sekolah, serta pengambilan keputusan berbasis data[11], [12].

Algoritma klasifikasi merupakan salah satu teknik data mining yang banyak digunakan untuk melakukan prediksi dan pengelompokan data berdasarkan karakteristik tertentu. Beberapa algoritma klasifikasi yang populer dan terbukti efektif antara lain Decision Tree (C4.5), k-Nearest Neighbor (KNN), dan Naïve Bayes. Algoritma Decision Tree memiliki keunggulan dalam kemudahan interpretasi hasil model, KNN efektif dalam menangani pola data non-linear, sedangkan Naïve Bayes unggul dalam kecepatan komputasi dan kinerja yang stabil pada dataset tertentu. Penerapan algoritma-algoritma tersebut dalam bidang pendidikan telah menunjukkan hasil yang menjanjikan, termasuk dalam prediksi prestasi akademik, kelulusan siswa, dan perilaku belajar[13], [14].

Berdasarkan latar belakang tersebut, penelitian ini berfokus pada pengembangan model prediksi tingkat kedisiplinan siswa dengan memanfaatkan pola absensi dan prestasi akademik menggunakan pendekatan data mining. Penelitian dilakukan pada SMK Swasta RK Bintang Timur Pematangsiantar dengan menerapkan kerangka kerja CRISP-DM (Cross Industry Standard Process for Data Mining), yang mencakup tahapan pemahaman masalah, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi. Pemilihan CRISP-DM bertujuan untuk memastikan proses penelitian berjalan secara sistematis dan terstruktur[15].

Penelitian ini menggunakan beberapa algoritma klasifikasi, yaitu Decision Tree (C4.5), KNN, dan Naïve Bayes, untuk membandingkan kinerja masing-masing model dalam memprediksi tingkat kedisiplinan siswa. Evaluasi model dilakukan menggunakan confusion matrix serta metrik kinerja seperti accuracy, precision, recall, dan F1-score. Dengan pendekatan ini, diharapkan dapat diperoleh model klasifikasi yang paling optimal dan dapat diandalkan sebagai dasar sistem pendukung keputusan.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi praktis bagi pihak sekolah, khususnya guru Bimbingan Konseling (BK), wali kelas, dan manajemen sekolah, dalam melakukan pemantauan dan pembinaan kedisiplinan siswa secara lebih proaktif. Model prediksi yang dihasilkan dapat berfungsi sebagai *early warning system* untuk mengidentifikasi siswa yang berpotensi mengalami penurunan kedisiplinan sehingga intervensi dapat dilakukan lebih cepat dan tepat sasaran. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi akademik dalam pengembangan kajian *educational data mining*, khususnya dalam penerapan algoritma machine learning untuk analisis perilaku dan kedisiplinan siswa di lingkungan sekolah kejuruan.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan bentuk penambangan data yang digunakan untuk mengekstraksi pengetahuan dari sejumlah besar data. Data mining diperlukan dalam membuat prediksi terhadap hubungan-hubungan yang memiliki makna, pola, dan kecenderungan dengan menganalisis kumpulan data besar yang tersimpan dalam penyimpanan menggunakan teknik pengenalan pola statistik atau matematis. Dalam Data Mining, proses menemukan pola atau informasi berguna dari data yang telah dipilih atau diproses disebut sebagai Penemuan Pengetahuan dalam Basis Data (Knowledge Data Discovery/KDD).

2.2 Confusion Matrix

Confusion Matrix adalah sebuah tabel yang menyatakan klasifikasi dari jumlah uji yang benar dan jumlah uji yang salah. Dari definisi confusion matrix tersebut, beberapa poin dalam confusion matrix digunakan untuk menghitung precision, recall, dan f1-score. Precision merupakan perbandingan antara True Positive (TP) dengan jumlah data yang diprediksi positif, secara matematis dapat dilihat di bawah ini [20]:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Adapun recall didefinisikan sebagai perbandingan antara true positive (TP) terhadap jumlah data yang sesungguhnya positif. Secara matematis, rumus recall dapat dituliskan sebagai berikut:

$$Precision = \frac{TP}{TP+FN} \quad (2)$$

Sementara itu, F1-Score merepresentasikan nilai harmonik rata-rata dari precision dan recall. Nilai ideal F1-Score adalah 1, sedangkan nilai terendah adalah 0. Secara matematis, perhitungan F1-Score dinyatakan sebagai berikut:

$$\frac{1}{F1} = \frac{1}{2} \left(\frac{1}{Precision} + \frac{1}{Recall} \right) \quad (3)$$

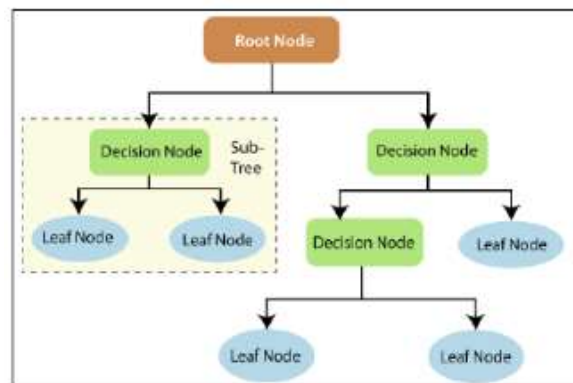
Nilai F1-Score yang tinggi mengindikasikan bahwa model klasifikasi yang dikembangkan memiliki kinerja yang seimbang antara precision dan recall.

2.3 Algoritma Klasifikasi dalam Prediksi Kedisiplinan

Beberapa algoritma yang umum digunakan dalam prediksi berbasis data mining antara lain:

2.3.1 Decision Tree (C4.5)

Algoritma klasifikasi dalam pembelajaran mesin terdiri dari beberapa jenis. Pada penelitian ini, algoritma Decision Tree (DT) dipilih sebagai fokus utama. Gambar 1 memperlihatkan representasi visual dari struktur DT secara umum.



Gambar 1. Ilustrasi struktur DT

DT adalah model berurutan yang menyatukan serangkaian tes dasar secara efisien dan kohesif dimana fitur numerik dibandingkan dengan nilai ambang batas disetiap pengujian. Dari masing-masing statement akan menghasilkan algoritma baru sesuai data yang kita inginkan.

2.3.2 K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) adalah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Untuk pemilihan atribut terdiri dari n neighbors (biasa disebut k) [9]. Mengklasifikasikan objek berdasarkan kedekatan data lain dalam ruang multidimensi. Cocok untuk dataset dengan pola non-linear. Penghitungan jarak umumnya menggunakan jarak Euclidean Distance sebagai berikut.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (4)$$

Dimana $d(x_i, x_j)$ merupakan jarak Euclidean, x_i dan x_j adalah dua titik yang akan dihitung jaraknya, dengan x_i merupakan record data ke i dan x_j adalah record data ke j, serta a_r adalah data ke-r dengan i, j bernilai 1,2,3,...n.

2.3.3 Naive Bayes

Metode klasifikasi naive bayes adalah salah satu algoritma dalam teknologi klasifikasi yang mudah diimplementasikan dan cepat dalam kecepatan pemrosesan [3]. Metode Naive Bayes menggunakan model statistik untuk melakukan proses klasifikasi data. Metode ini menghitung nilai probabilitas data uji berdasarkan data kasus yang sudah pernah terjadi [3]. Berikut adalah simulasi perhitungan pada probabilitas naive bayes yang bisa dilihat pada Persamaan 1 berikut:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5)$$

$P(A|B)$ adalah peluang terjadinya A jika B telah terjadi, $P(B|A)$ adalah peluang terjadinya B jika A diketahui benar, $P(B)$ adalah probabilitas marginal B tanpa mempertimbangkan bukti lain, dan $P(A)$ adalah probabilitas marginal A.

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Dataset Penelitian

Penelitian ini bertujuan untuk mengembangkan model prediksi tingkat kedisiplinan siswa berdasarkan pola absensi dan prestasi akademik dengan menerapkan pendekatan data mining menggunakan algoritma klasifikasi. Dataset yang digunakan diperoleh dari SMK Swasta RK Bintang Timur Pematangsiantar dan terdiri atas 711 data siswa dari kelas X, XI, dan XII. Data tersebut mencakup empat jurusan, yaitu Rekayasa Perangkat Lunak, Teknik Busana, Teknologi Farmasi, dan Layanan Kesehatan, yang dikumpulkan selama dua semester pembelajaran.

Secara keseluruhan, dataset awal terdiri dari 39 fitur yang merepresentasikan karakteristik akademik, absensi, perilaku, serta faktor pendukung lainnya. Fitur-fitur tersebut meliputi data identitas siswa, latar belakang keluarga, nilai akademik, kehadiran, keterlambatan, keterlibatan dalam kegiatan sekolah, hingga skor dan kategori kedisiplinan. Variabel kedisiplinan digunakan sebagai label kelas yang menjadi target prediksi dalam penelitian ini.

3.2 Hasil Pra-Pemrosesan dan Seleksi Fitur

Sebelum dilakukan pemodelan, dataset melalui tahap pra-pemrosesan untuk meningkatkan kualitas data. Proses ini meliputi data cleaning dengan menghapus data yang tidak lengkap, tidak konsisten, serta duplikasi yang berpotensi memengaruhi hasil klasifikasi. Tahapan ini menghasilkan dataset yang lebih bersih dan representatif terhadap kondisi sebenarnya.

Selanjutnya dilakukan proses seleksi fitur (feature selection) untuk mengidentifikasi atribut yang paling relevan terhadap tingkat kedisiplinan siswa. Dari 39 fitur awal, diperoleh sembilan fitur utama yang dianggap paling berpengaruh, yaitu prestasi nilai, persentase absensi, keterlambatan, ketepatan pengumpulan tugas, jumlah pelanggaran, partisipasi kelas, nilai sikap, motivasi belajar, dan kedisiplinan. Pengurangan jumlah fitur ini bertujuan untuk meningkatkan efisiensi pemodelan serta mengurangi risiko overfitting.

3.3 Normalisasi dan Pembagian Data

Dataset hasil seleksi fitur selanjutnya dinormalisasi menggunakan metode Min-Max Scaling. Normalisasi ini bertujuan untuk menyamakan skala nilai antar fitur sehingga tidak terjadi dominasi atribut tertentu dalam proses perhitungan, khususnya pada algoritma berbasis jarak dan probabilitas.

Setelah proses normalisasi, data dibagi menjadi dua bagian, yaitu 80% sebagai data latih (training data) dan 20% sebagai data uji (testing data). Pembagian ini memungkinkan evaluasi performa model secara objektif terhadap data yang belum pernah digunakan pada tahap pelatihan.

3.4 Hasil Pemodelan Klasifikasi

Tabel 1. Perbandingan Komprehensif Model Klasifikasi

Model	Akurasi	Precision	Recall	F1-Score
Decision Tree	0.6713	0.6592	0.6713	0.6490
KNN	0.7692	0.7679	0.7692	0.7557
Naive Bayes	0.7762	0.7788	0.7762	0.7541

Pada tahap pemodelan, tiga algoritma klasifikasi digunakan untuk membangun model prediksi kedisiplinan siswa, yaitu Decision Tree (C4.5), K-Nearest Neighbor (KNN), dan Naive Bayes. Ketiga algoritma tersebut dipilih karena umum digunakan dalam penelitian educational data mining dan memiliki karakteristik yang berbeda, sehingga memungkinkan perbandingan kinerja yang komprehensif.

Proses pelatihan dan pengujian model dilakukan menggunakan platform Google Colab. Setiap algoritma dilatih menggunakan data latih yang sama dan diuji menggunakan data uji yang identik untuk memastikan keadilan dalam perbandingan performa. Hasil pemodelan kemudian dianalisis menggunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score untuk menilai kemampuan masing-masing algoritma dalam mengklasifikasikan tingkat kedisiplinan siswa.

3.5 Pembahasan Hasil Klasifikasi

Hasil pengujian menunjukkan bahwa penerapan metode data mining dengan algoritma klasifikasi mampu mengidentifikasi pola hubungan antara absensi, prestasi akademik, dan tingkat kedisiplinan siswa. Fitur-fitur seperti persentase absensi, keterlambatan, dan jumlah pelanggaran terbukti memiliki pengaruh signifikan terhadap hasil prediksi. Siswa dengan tingkat kehadiran tinggi dan prestasi akademik yang baik cenderung diklasifikasikan dalam kategori kedisiplinan tinggi, sedangkan siswa dengan tingkat alfa dan keterlambatan yang tinggi lebih banyak masuk dalam kategori kedisiplinan rendah. Temuan ini menunjukkan bahwa model prediksi yang dikembangkan memiliki potensi untuk digunakan sebagai sistem pendukung keputusan dalam pemantauan dan pembinaan kedisiplinan siswa di lingkungan sekolah kejuruan.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai Pengenalan Pola Absensi dan Prestasi Akademik Siswa SMK Swasta RK Bintang Timur Pematangsiantar Menggunakan Data Mining untuk Prediksi Kedisiplinan, dapat disimpulkan bahwa penerapan metode data mining mampu membantu sekolah dalam menganalisis pola kedisiplinan siswa secara lebih objektif dan terukur. Hasil pengolahan data menggunakan algoritma classification menunjukkan bahwa variabel absensi dan prestasi akademik memiliki pengaruh signifikan terhadap tingkat kedisiplinan siswa. Model prediksi yang dibangun dapat mengelompokkan siswa ke dalam kategori disiplin tinggi, sedang, dan rendah dengan tingkat akurasi yang memadai sehingga dapat dijadikan dasar rekomendasi kebijakan dalam pembinaan siswa. Temuan ini membuktikan bahwa teknologi data mining memberikan kontribusi penting bagi manajemen sekolah dalam proses pengambilan keputusan, terutama dalam upaya meningkatkan mutu pendidikan dan membentuk karakter kedisiplinan peserta didik. Penelitian ini diharapkan menjadi referensi bagi sekolah untuk mengembangkan sistem monitoring kedisiplinan berbasis data secara berkelanjutan dan memperluas cakupan variabel penelitian pada aspek-aspek lain yang relevan, seperti faktor lingkungan, motivasi belajar, dan keterlibatan orang tua.

REFERENCES

- [1] F. S. A. Safitri and E. Sutadji, "Strategi Pengembangan Kompetensi Lulusan Pendidikan Kejuruan Guna Meningkatkan Daya Saing Global," *Didakt. J. Kependidikan*, vol. 14, no. 1, pp. 1507–1522, 2025, [Online]. Available: <https://jurnaldidaktika.org>
- [2] N. A. Pristi, A. V. Sjafri, and G. Suprayitno, "Strategi Peningkatan Daya Saing Usaha Kecil Menengah Melalui Peningkatan Mutu Sekolah Menengah Kejuruan," *Manaj. IKM J. Manaj. Pengemb. Ind. Kecil Menengah*, vol. 18, no. 1, pp. 73–83, 2023, doi: 10.29244/mikm.18.1.73-83.
- [3] S. Bidol, "Analisis ketersediaan sumber daya dan proses pengembangan kurikulum terhadap kebutuhan industri dimediasi oleh mutu pendidikan di SMK Negeri 8 Samarinda ...," *J. Din. Manaj. Berkelanjutan*, vol. 6, no. 4, 2024, [Online]. Available: <https://ejournals.com/ojs/index.php/jdmb/article/view/1320%0Ahttps://ejournals.com/ojs/index.php/jdmb/article/download/1320/1571>
- [4] A. Rifai, "Strategi Kepala Sekolah Dalam Implementasi Pendidikan Nilai Di Sekolah," *TA'DIBUNA J. Pendidik. Agama Islam*, vol. 1, no. 1, p. 1, 2018, doi: 10.30659/jpai.1.1.1-10.
- [5] A. Widiyanto and M. Ali, "Efektivitas Pendekatan Segitiga Restitusi dalam Pendidikan Agama Islam untuk Meningkatkan Disiplin Siswa Sekolah Menengah Kejuruan di Boyolali (Studi Kasus SMK N 1 Sawit dan SMK Muhammadiyah 1 Boyolali)," *J. Islam. Educ. Stud.*, vol. 10, no. 3, pp. 544–567, 2025, [Online]. Available: <http://ejournal.uniramalang.ac.id/index.php/alwijdan>
- [6] E. Alzizah, D. S. Wulandari, R. Rahim, J. Wulansyah, and Duriska, "Pengelolaan Ketidakhadiran Siswa Berbasis Aplikasi di SMK 1 Muhammadiyah Sangatta," *J. Adm. Pendidik. Islam*, vol. 4, no. 2, pp. 190–200, 2022, doi: 10.15642/japi.2022.4.2.190-200.
- [7] Y. A. Rindri, D. Yolanda, and I. Irwan, "Sistem Monitoring dan Penilaian Siswa Terbaik Praktik Kerja Industri pada SMK Negeri 1 Bakam," *Manutech J. Teknol. Manufaktur*, vol. 16, no. 02, pp. 135–144, 2024, doi: 10.33504/manutech.v16i02.345.
- [8] I. Malah, H. Sumual, and I. Rianto, "Perancangan Sistem Absensi, Tracking Guru dan Siswa di Sekolah Menengah

- Kejuruan," *Eduatik J. Pendidik. Teknol. Inf. dan Komun.*, vol. 2, no. 2, pp. 159–171, 2022, doi: 10.53682/edutik.v2i2.4431.
- [9] S. A. A. Kharis and A. H. A. Zili, "Learning Analytics dan Educational Data Mining pada Data Pendidikan," *J. Ris. Pembelajaran Mat. Sekol.*, vol. 6, no. 1, pp. 12–20, 2022, doi: 10.21009/jrpms.061.02.
- [10] A. R. Azis, "Analisis Komparasi Algoritma Machine Learning dalam Prediksi Performa Akademik Mahasiswa: Literature Review," *J. Ilmu Komput. dan Inform.*, vol. 4, no. 2, pp. 143–148, 2025, doi: 10.54082/jiki.212.
- [11] S. Anastassia Amellia Kharis and A. Haqqi Anna Zili, "Tantangan Dan Peluang Educational Data Mining Dan Learning Analytics Pada Pendidikan Jarak Jauh Di Indonesia," in *Prosiding Seminar Nasional Sains dan ...*, 2024, pp. 218–225. [Online]. Available: <https://conference.ut.ac.id/index.php/saintek/article/view/2536%0Ahttps://conference.ut.ac.id/index.php/saintek/article/download/2536/1049>
- [12] Nuari Anisa Sivi, Rudi Hartono, and Putra Hanafi, "Penerapan Algoritma C4.5 untuk Prediksi Kelulusan Mahasiswa berdasarkan Data Akademik," *Polyg. J. Ilmu Komput. dan Ilmu Pengetah. Alam*, vol. 1, no. 5, pp. 01–17, 2023, doi: 10.62383/polygon.v1i5.855.
- [13] H. Pratiwi, M. I. Sa'ad, and Salmon, "Strategi Manajemen Pendidikan Berbasis Machine Learning untuk Prediksi Prestasi Siswa," *Borneo Educ. Manag. Res. J.*, vol. 6, no. 1, pp. 21–30, 2025, [Online]. Available: <https://jurnal.fkip.unmul.ac.id/index.php/bedu/article/view/5016%0Ahttps://jurnal.fkip.unmul.ac.id/index.php/bedu/article/download/5016/1989>
- [14] T. Gori, A. Sunyoto, and H. Al Fatta, "Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 1, pp. 215–224, 2024, doi: 10.25126/jtiik.20241118074.
- [15] D. A. Triani, M. Aldi, N. H. P. Fauzi, and R. N. Safitri, "Curriculum innovation at SMK PGRI 2 Cimahi: Preparing students for the workforce," *Hipkin J. Educ. Res.*, vol. 2, no. 1, pp. 23–36, 2025, doi: 10.64014/hipkin-jer.v2i1.40.