

Analysis Of Student Performance Determination Using K-Means Clustering And Perceptron At Private Madrasah Tsanawiyah Koto Anau

Zakiya Hikmi ¹⁾; Musli Yanto ²⁾; Syafri Arlis ³⁾

^{1,2,3)} Informatics Engineering, Faculty of Computer Science, Universitas Putra Indonesia YPTK Padang

Email: ¹⁾ zakiyahikmizaki@gmail.com

How to Cite :

Hikmi, Z., Yanto, M., Arlis, S. (2026). Analysis Of Student Performance Determination Using K-Means Clustering And Perceptron At Private Madrasah Tsanawiyah Koto Anau Jurnal Media Computer Science, 5(4)

ARTICLE HISTORY

Received [03 Agustus 2026]

Revised [25 Oktober 2026]

Accepted [27 Oktober 2026]

KEYWORDS

Machine Learning, K-Means Clustering, Perceptron, Student Achievement, Classification, Prediction, Madrasah Tsanawiyah.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Metode konvensional untuk menilai hasil belajar siswa di MTsS Koto Anau belum optimal karena tidak dapat mengidentifikasi pola karakteristik siswa selama pembelajaran. Penelitian ini menawarkan pendekatan hibrida K-Means Clustering dan Perceptron untuk melakukan analisis objektif terhadap penilaian kinerja siswa. Metode hibrida ini bekerja dalam dua tahapan. Pertama, mengelompokkan siswa dengan K-Means ke dalam tiga kluster performa berdasarkan kemiripan kesenjangan pada atribut akademis. Kemudian, menggunakan hasil kluster sebagai label untuk melatih model klasifikasi Perceptron. Data diperoleh dari 113 siswa di Kantor Tata Usaha MTsS Koto Anau. Data ini melibatkan tiga variabel, yaitu nilai rata-rata rapor, tingkat partisipasi, dan skor ujian akhir semester. Hasil penilaian model menunjukkan bahwa akurasi total mencapai 99% dengan nilai precision, recall, dan F1-score rata-rata tertimbang semua sebesar 0,99 pada 113 data uji. Ini mengindikasikan model sangat dapat diandalkan dalam mengklasifikasikan tiga kategori prestasi. Temuan utama menunjukkan bahwa kombinasi K-Means dan Perceptron bekerja dengan baik dalam menggambarkan profil belajar siswa serta memprediksi kategori prestasi siswa baru dengan akurat. Penelitian ini menghadirkan model analitik yang sederhana, mudah dimengerti, dan relevan bagi madrasah dalam merencanakan intervensi pembelajaran yang lebih personal dan terfokus.

ABSTRACT

The way we normally evaluate how well students are doing at MTsS Koto Anau is not good enough. This is because it does not show us the hidden patterns of how students learn. This study wants to use a way to evaluate how well students are doing. It uses K-Means Clustering and Perceptron together to get an idea of what helps students achieve. We used a two step method. First K-Means Clustering puts students into three groups based on how similar they're in terms of school work. Then we use these groups to teach the Perceptron model. The information we used came from the office at MTsS Koto Anau. It had details about 113 students. Included things like their average grades how often they came to school and their final exam scores. When we tested the model it was very good at putting students into the group. It was 99 percent of the time. This shows that the model is very good, at figuring out which group a student belongs to. The main thing we found out is that using K-Means Clustering and Perceptron is a good way to understand how students learn and predict how well new students will do. This study helps because it gives madrasahs an useful way to analyze student data. This

means they can make learning plans that're just right for each student. The MTsS Koto Anau can use this to make sure students get the help they need to do their best.

PENDAHULUAN

Pendidikan merupakan pilar fundamental pembangunan sumber daya manusia karena kualitas sistem pendidikan secara langsung menentukan daya saing bangsa dan kemajuan peradaban (Dihe & Wangdra, 2023). Prestasi siswa menjadi indikator vital mengukur keberhasilan proses belajar mengajar sekaligus merefleksikan pemahaman siswa dan efektivitas metode pengajaran (Royhanuddin et al., 2024). Namun, penilaian prestasi di MTsS Koto Anau masih mengandalkan metode konvensional yang hanya berfokus pada nilai ujian akhir dan peringkat kelas, sehingga mengabaikan faktor multidimensi seperti motivasi, kecerdasan, kesehatan, dan dukungan keluarga yang terbukti signifikan memengaruhi hasil belajar (Islam & Sunan, 2023).

Perkembangan teknologi informasi dan ilmu data melahirkan paradigma baru melalui Educational Data Mining (EDM) yang menerapkan statistik dan penambangan data untuk menggali wawasan pendidikan secara holistik (Seminar et al., 2024). Machine learning, sebagai cabang kecerdasan buatan, memungkinkan sistem belajar otomatis dan adaptif dari data tanpa pemrograman eksplisit, sehingga mendeteksi pola kompleks yang terlewat analisis deskriptif konvensional (Widya et al., 2024). Predictive analytics in education memungkinkan identifikasi dini siswa berisiko, rekomendasi personalisasi pembelajaran, dan intervensi berbasis bukti untuk peningkatan kualitas pendidikan berkelanjutan (Irvan & Annur, 2024).

Penelitian terdahulu mengeksplorasi machine learning untuk klasifikasi dan clustering secara terpisah dengan beragam hasil yang cukup menjanjikan. Perbandingan antara algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM), dan Artificial Neural Network (ANN) untuk memprediksi kelulusan tepat waktu mahasiswa menunjukkan bahwa SVM memberikan akurasi tertinggi sebesar 0,94 pada dataset dengan batas pemisah yang jelas (Junaidi et al., 2024). Studi komparasi lebih lanjut mengonfirmasi keunggulan SVM hingga mencapai 94,4% pada dataset dengan margin data yang jelas, sementara Random Forest menawarkan konsistensi akurasi sebesar 85% untuk dataset yang lebih besar dan kompleks (Azis, 2024). Penerapan algoritma Decision Tree, Random Forest, dan Extreme Gradient Boosting (XGBoost) untuk memprediksi status kelulusan siswa Sekolah Menengah Pertama melaporkan bahwa Random Forest memberikan performa terbaik dengan akurasi 99,5% serta mengidentifikasi variabel nilai matematika sebagai prediktor paling dominan (Frananda et al., 2023). Pada ranah clustering, eksplorasi algoritma K-Means untuk penilaian prestasi siswa dengan menggabungkan variabel akademik dan non-akademik seperti nilai rata-rata, kehadiran, sikap, dan kegiatan ekstrakurikuler menunjukkan bahwa evaluasi menggunakan Silhouette Score dan Davies-Bouldin Index menghasilkan konfigurasi 6 cluster sebagai segmentasi paling optimal (Gunandaru et al., 2025). Secara lebih luas, integrasi antara supervised dan unsupervised learning dinilai sangat potensial untuk personalisasi konten pembelajaran, analisis performa siswa secara real-time, serta deteksi dini kesulitan belajar yang dapat mendukung pengambilan keputusan akademik di tingkat sekolah (Halimah et al., 2025).

Berfokus pada metode inti yang diusulkan dalam penelitian ini, yaitu K-Means Clustering dan algoritma klasifikasi Perceptron, sejumlah studi terdahulu telah menguji kedua metode ini secara individual dengan hasil yang cukup meyakinkan. Algoritma K-Means bekerja dengan mengelompokkan data di sekitar centroid yang diinisialisasi secara iteratif, sehingga mampu membagi siswa berdasarkan kemiripan karakteristik belajar yang diukur melalui atribut-atribut akademik tertentu. Optimalisasi pengelompokan indeks prestasi akademik siswa menggunakan K-Means berhasil menghasilkan dua klaster utama berdasarkan evaluasi Davies-Bouldin Index terendah, yang memudahkan sekolah dalam menyusun strategi pembelajaran terdiferensiasi sesuai kebutuhan masing-masing kelompok siswa secara lebih terarah (Suarna et al., 2025). Prediksi nilai akhir mahasiswa menggunakan K-Means berdasarkan nilai tugas percabangan, perulangan, dan

array pada mata kuliah pemrograman dasar mencapai akurasi 93% dengan precision 86% serta recall 100%, membuktikan bahwa K-Means tidak hanya berguna untuk eksplorasi data tanpa label tetapi juga dapat diarahkan untuk keperluan prediksi awal yang cukup andal (Nasional et al., 2023). Penerapan Explainable Machine Learning dengan algoritma Gradient Boosting, Random Forest, dan Regresi Logistik untuk memprediksi penurunan minat belajar dan prestasi akademik siswa Sekolah Menengah Atas membuktikan bahwa pendekatan prediktif sangat efektif untuk sistem peringatan dini (early warning system) di lingkungan sekolah, dengan Gradient Boosting mencapai AUC 0,84 untuk prediksi minat dan 0,82 untuk prediksi prestasi (Yusup, 2025). Meskipun kedua metode inti ini telah diuji secara terpisah dalam berbagai konteks dan menghasilkan performa yang baik, kajian mendalam terhadap penelitian terdahulu menunjukkan bahwa belum ada studi yang secara khusus mengintegrasikan K-Means dan Perceptron dalam satu pipeline hibrida yang bersifat eksploratif sekaligus prediktif untuk penentuan prestasi siswa di tingkat madrasah tsanawiyah yang memiliki karakteristik kurikulum terpadu antara pendidikan umum dan keagamaan.

Penegasan masalah berakar pada kesenjangan antara potensi analitik data dan praktik konvensional di MTsS Koto Anau, di mana kurangnya sistem teknologi menyebabkan intervensi pembelajaran umum, tidak tepat sasaran, dan tidak personal. Tujuan penelitian adalah merancang kerangka hibrida dua tahap: K-Means mengelompokkan siswa ke tiga kluster prestasi berdasarkan nilai rata-rata rapor, persentase kehadiran, dan nilai ujian akhir, sedangkan Perceptron mengklasifikasi siswa baru secara akurat dan objektif. Kontribusi penelitian meliputi tiga aspek: teoritis, memperkaya literatur Educational Data Mining dengan bukti empiris efektivitas hibrida K-Means dan Perceptron pada dataset madrasah; praktis, menyediakan model ringan komputasi, mudah diinterpretasi pendidik, dan dapat diadopsi sebagai alat keputusan berbasis data; serta transformasi digital madrasah dengan solusi teknologi relevan dan kontekstual untuk sistem pendidikan adaptif, inklusif, dan berkeadilan sesuai semangat Kurikulum Merdeka yang menekankan pembelajaran memerdekakan dan memenuhi kebutuhan unik setiap peserta didik.

LANDASAN TEORI

Kecerdasan buatan (AI) memungkinkan sistem melakukan tugas yang memerlukan kecerdasan manusia, seperti pengenalan pola dan pengambilan keputusan (Widya et al., 2024). Dalam dunia pendidikan, AI menawarkan potensi besar untuk mentransformasi pembelajaran melalui personalisasi dan analisis data skala besar guna mengungkap pola tersembunyi (Irvan & Annur, 2024). Educational Data Mining (EDM) menerapkan teknik machine learning pada data pendidikan untuk meningkatkan kualitas pengajaran secara berkelanjutan (Anastassia et al., 2024). Penelitian ini memanfaatkan AI melalui machine learning untuk membangun model analitik guna mendukung pengambilan keputusan akademik yang objektif di MTsS Koto Anau.

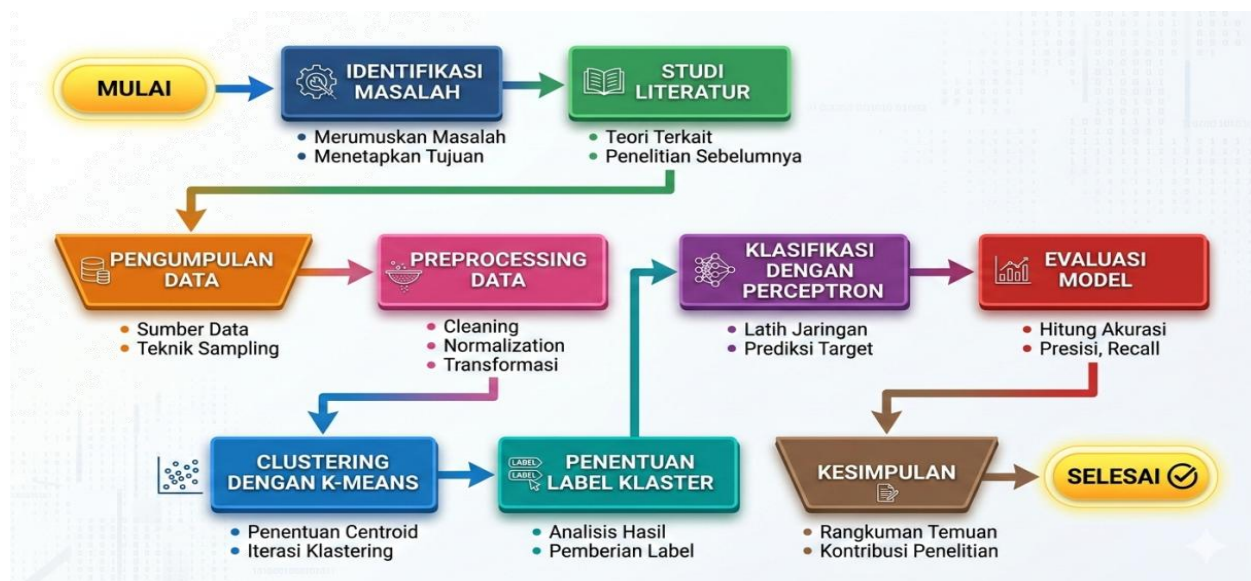
Machine learning adalah sub-bidang AI yang memungkinkan sistem belajar dari data tanpa pemrograman eksplisit, dengan kualitas model sangat bergantung pada kualitas dan kuantitas data pelatihan (Sari et al., 2025). Dalam konteks pendidikan, machine learning digunakan untuk memprediksi kinerja akademik, mengelompokkan siswa berdasarkan gaya belajar, serta mendeteksi siswa berisiko putus sekolah (Seminar et al., 2024). Penelitian ini mengadopsi pendekatan hibrida yang menggabungkan unsupervised learning dan supervised learning untuk mengeksplorasi pola alami dalam data sekaligus membangun model prediktif (Mujiyono et al., 2025). Pendekatan ini dipilih karena kemampuannya menangani kompleksitas data pendidikan secara lebih komprehensif dan terukur.

Unsupervised learning bertujuan menemukan struktur tersembunyi dalam data tanpa label berdasarkan kemiripan antar data (Sinaga & Passarella, 2025). Dalam pendidikan, metode ini sangat berguna untuk mengelompokkan siswa berdasarkan karakteristik akademik guna mengidentifikasi segmen dengan kebutuhan belajar yang berbeda (Safitri et al., 2025). Algoritma K-Means merupakan clustering berbasis partisi yang sederhana dan efisien, bekerja dengan membagi data ke dalam K cluster berdasarkan kedekatan dengan centroid secara iteratif hingga konvergen (Sukmaningtyas et

al., 2024). Penelitian ini menggunakan K-Means dengan K=3 untuk mengelompokkan siswa ke dalam kategori prestasi tinggi, sedang, dan rendah berdasarkan nilai rapor, kehadiran, dan nilai ujian akhir semester. Supervised learning melatih algoritma menggunakan data berlabel untuk mempelajari pemetaan antara input dan output guna memprediksi data baru (Nurhalizah & Ardianto, 2024). Algoritma Perceptron adalah bentuk paling sederhana dari jaringan saraf tiruan yang menghitung penjumlahan terboboti input dengan bias dan menerapkan fungsi aktivasi step untuk menghasilkan keluaran biner (Aulia et al., 2025). Proses pembelajaran dilakukan secara iteratif dengan memperbarui bobot jika terjadi kesalahan klasifikasi sesuai rumus $y = f(w_1x_1 + \dots + w_nx_n + b)$ (Directions, 2023). Penelitian ini menerapkan Perceptron pada tahap klasifikasi setelah pelabelan dari K-Means, sehingga model hibrida yang dihasilkan mampu mengeksplorasi data secara alami sekaligus memprediksi kategori prestasi siswa baru secara akurat.

METODE PENELITIAN

Metodologi penelitian merupakan prosedur ilmiah sistematis untuk mengumpulkan, mengolah, dan menganalisis data guna menghasilkan pemecahan masalah yang valid dan objektif (Halimah et al., 2025). Kerangka penelitian dirancang sebagai struktur logis yang mengintegrasikan pendekatan hibrida antara unsupervised learning (K-Means Clustering) dan supervised learning (Perceptron) dengan cakupan meliputi identifikasi masalah, studi literatur, pengumpulan data, preprocessing, clustering, penentuan label, klasifikasi, dan evaluasi model (Nasional et al., 2023). Alur kerja operasional penelitian disajikan pada Gambar 1.



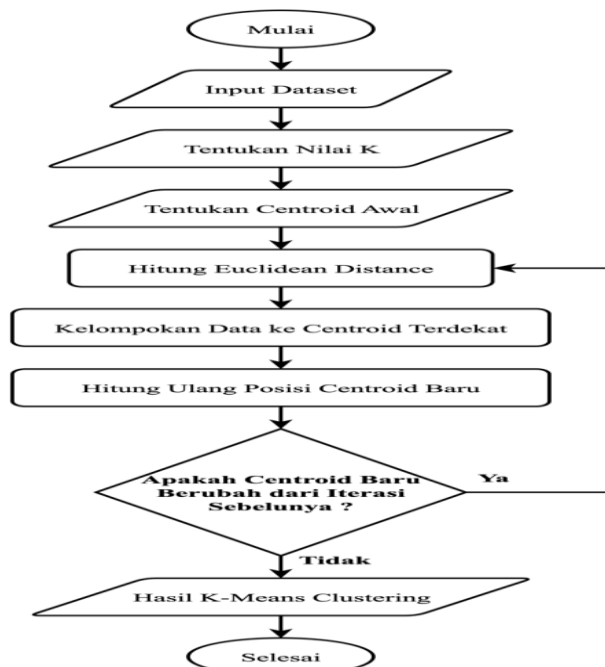
Gambar 1. Kerangka Penelitian

Gambar 1 menggambarkan alur penelitian yang dimulai dari identifikasi masalah melalui observasi di MTsS Koto Anau, dilanjutkan studi literatur untuk memetakan state of the art. Pengumpulan data dilakukan dengan purposive sampling pada 113 siswa tahun ajaran 2024/2025-2025/2026 dengan tiga variabel: nilai rata-rata rapor (X1), persentase kehadiran (X2), dan nilai ujian akhir (X3). Data melalui preprocessing (cleaning, normalisasi Min-Max Scaling, transformasi .csv), kemudian clustering K-Means (K=3) mengelompokkan siswa ke kluster tinggi, sedang, rendah melalui iterasi jarak Euclidean hingga konvergen dan WCSS minimum, lalu diberi label berdasarkan analisis centroid.

Tahap klasifikasi Perceptron menggunakan data berlabel yang dipecah menjadi data latih dan uji, dengan masukan X_1 , X_2 , X_3 diproses melalui bobot sinaptik, bias, dan fungsi aktivasi hingga model mengenali pola pemisah linear antar-kelas. Evaluasi model menggunakan Confusion Matrix untuk menghitung akurasi, presisi, recall, dan F1-score pada setiap kategori prestasi guna memastikan performa optimal, yang menjadi dasar kelayakan model hibrida sebagai alat pendukung keputusan akademik di MTsS Koto Anau.

K-Means Clustering

K-Means merupakan algoritma clustering berbasis partisi yang paling sederhana dan banyak digunakan karena efisiensinya dalam menangani data dalam jumlah besar serta kemudahan implementasinya (Ardianti et al., 2023). Algoritma ini bertujuan membagi data ke dalam sejumlah K cluster yang telah ditentukan, di mana setiap data menjadi anggota cluster berdasarkan kedekatannya dengan pusat cluster (centroid) dengan meminimalkan jarak intra-cluster dan memaksimalkan jarak inter-cluster (Dewi et al., 2022). Algoritma bekerja secara iteratif hingga mencapai konvergensi, yaitu ketika posisi centroid tidak lagi berubah secara signifikan. Secara visual, alur kerja algoritma K-Means Clustering disajikan pada Gambar 2.



Gambar 2 Flowchart Algoritma K-Means Clustering

Gambar 2 menggambarkan alur prosedur K-Means yang dimulai dari input dataset, penentuan nilai K , penentuan centroid awal secara acak, perhitungan jarak Euclidean, pengelompokan data ke centroid terdekat, pembaruan posisi centroid, dan pengulangan proses hingga centroid tidak berubah yang menandakan konvergensi, kemudian menghasilkan output berupa hasil clustering. Secara matematis, algoritma ini memiliki dua persamaan utama, yaitu rumus Euclidean Distance untuk menghitung jarak data ke centroid pada Persamaan (1):

$$d(A_i, B_i) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2} \quad (1)$$

Di mana:

$d(A_i, B_i)$: Jarak antara objek data dan centroid

A_i : Nilai objek data ke -i

B_i : Nilai centroid ke -i

n : Jumlah fitur atau dimensi data

Persamaan kedua merupakan rumus untuk memperbarui posisi centroid baru berdasarkan rata-rata seluruh anggota cluster, seperti ditunjukkan pada Persamaan (2):

$$C_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i \quad (2)$$

Di mana:

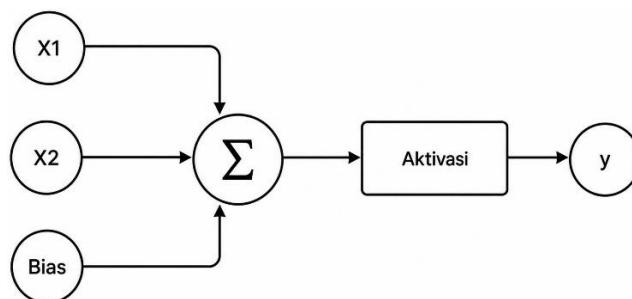
C_j : Centroid baru pada cluster ke -j

N_j : Jumlah data yang tergabung dalam cluster ke -j

X_i : Data ke -i yang menjadi anggota cluster

Algoritma Perceptron

Perceptron merupakan model klasifikasi linear berbasis supervised learning dan merupakan bentuk paling sederhana dari jaringan saraf tiruan yang ditemukan oleh Frank Rosenblatt pada tahun 1958 (Junaidi et al., 2024). Algoritma ini menerima vektor masukan dengan bobot dan bias, menghitung jumlah terboboti, lalu menerapkan fungsi aktivasi step untuk menghasilkan keluaran biner, sehingga efektif untuk klasifikasi data yang bersifat linearly separable (Aulia et al., 2025). Proses pembelajaran dilakukan secara iteratif menggunakan perceptron learning rule, di mana bobot diperbarui jika terjadi kesalahan klasifikasi (Saraf et al., 2024). Secara visual, alur perhitungan Perceptron disajikan pada Gambar 3 di bawah ini:



Gambar 3. Alur Perhitungan Perceptron

Gambar 3 menggambarkan alur komputasi Perceptron di mana input X_1 dan X_2 melalui proses perkalian dengan bobot sinaptik, kemudian dijumlahkan dan ditambahkan dengan bias, lalu hasilnya dilewatkan ke fungsi aktivasi untuk menghasilkan keluaran y berupa kelas prediksi. Secara matematis, penjumlahan terboboti pada Perceptron dinyatakan dalam Persamaan (3):

$$y = f(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) \quad (3)$$

Di mana:

y : Keluaran model.

f : Fungsi aktivasi.

w : Bobot.

x : Input.

b : Bias.

HASIL DAN PEMBAHASAN

Hasil

Bagian ini merupakan hasil eksperimen dan pembahasan implementasi metode hybrid K-Means Clustering dan Perceptron dalam menganalisis penentuan prestasi siswa di MTsS Koto Anau. Penelitian menggunakan data 113 siswa dengan tiga variabel utama: nilai rata-rata 5 mata pelajaran wajib (X1), persentase kehadiran (X2), dan rata-rata nilai rapor (X3). Seluruh proses dimulai dari pengumpulan data, preprocessing, clustering, klasifikasi, hingga evaluasi model secara menyeluruh.

Pembahasan

Dataset Penelitian

Dataset diperoleh dari bagian Tata Usaha MTsS Koto Anau tahun ajaran 2024/2025 dan 2025/2026. Data terdiri dari 113 siswa dengan variabel X1 (rata-rata rapor 5 mapel inti), X2 (persentase kehadiran), dan X3 (rata-rata ujian akhir semester). Seluruh data telah melalui verifikasi untuk memastikan keakuratan sebelum dianalisis. Adapun 5 sampel data yang ditampilkan pada Table 1.

Tabel 1. Dataset Nilai Siswa

No	Nama	Niali Rata-Rata 5 Mapel Wajib	Kehadiran	Rata-Rata Nilai Rapor
1	ADEFRYAN RAHMAT	83	93%	83,5
2	ADZKADINA FEBBY SYAHIRA	88	100%	86,9
3	AFIFAH NAZIFA HAZYAH	85,4	100%	86,7
4	AGUNG PUTRA EFENDI	79,4	93%	79,6
5	AHMAD FAIZ ROHAIDZAT	80	93%	81,5

Tabel 1 memperlihatkan bahwa nilai rata-rata 5 mapel wajib siswa berada pada rentang 79,4 hingga 88, Kehadiran pada rentang 93% hingga 100%, dan rata-rata nilai rapor pada rentang 79,6 hingga 86,9. Variasi nilai yang cukup beragam ini menunjukkan bahwa siswa MTsS Koto Anau memiliki tingkat capaian akademik yang bervariasi, sehingga diperlukan pendekatan analitik untuk mengelompokkan dan memprediksi prestasi mereka secara lebih objektif.

Analisis Variabel

Setelah data dikumpulkan, variabel dikonversi ke bentuk kode untuk memudahkan proses komputasi. Adapun 5 data setelah pengkodean variable disajikan pada table 2 berikut:

Tabel 2. Dataset dalam bentuk Variabel

No	Nama	X1	X2	X3
1	ADEFRYAN RAHMAT	83	93%	83,5
2	ADZKADINA FEBBY SYAHIRA	88	100%	86,9
3	AFIFAH NAZIFA HAZYAH	85,4	100%	86,7
4	AGUNG PUTRA EFENDI	79,4	93%	79,6
5	AHMAD FAIZ ROHAIDZAT	80	93%	81,5

Tabel 2 menunjukkan hasil konversi data ke kode variabel X1, X2, dan X3 yang siap digunakan dalam proses clustering dan klasifikasi.

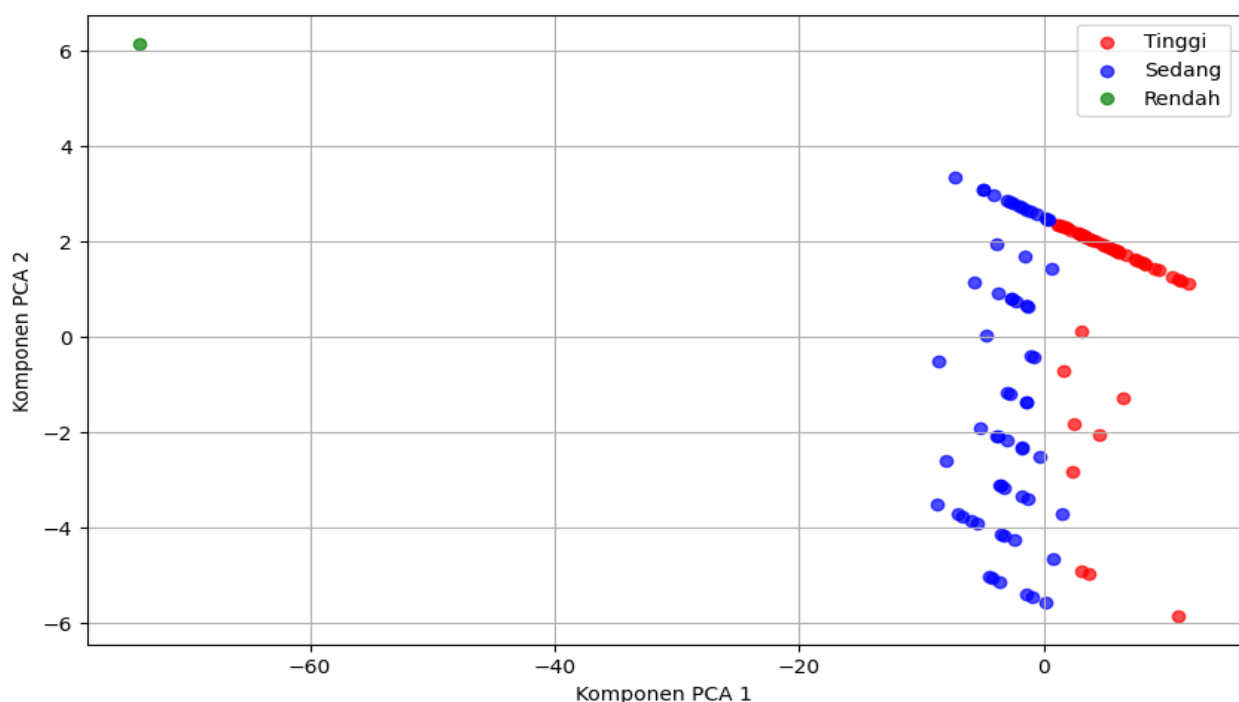
Hasil K-Means Clustering

Algoritma K-Means diterapkan dengan K=3 untuk mengelompokkan siswa ke kategori Tinggi, Sedang, dan Rendah. Proses clustering melalui 8 iterasi hingga konvergen, di mana posisi centroid iterasi ke-8 sama dengan iterasi ke-7 dan tidak ada siswa berpindah cluster. Hasil pengelompokan disajikan pada Tabel 3:

Tabel 3. Hasil Pengujian dengan K-Means

No	Nama	X1	X2	X3	Klaster	Status_Prestasi
1	ADEFRIAN RAHMAT	83	93%	83,5	2	Sedang
2	ADZKADINA FEBBY SYAHIRA	88	100%	86,9	1	Tinggi
3	AFIFAH NAZIFA HAZYAH	85,4	100%	86,7	1	Tinggi
4	AGUNG PUTRA EFENDI	79,4	93%	79,6	2	Sedang
5	AHMAD FAIZ ROHAIDZAT	80	93%	81,5	2	Sedang

Tabel 3 menunjukkan hasil akhir klasterisasi dengan distribusi: 50 siswa Tinggi, 62 siswa Sedang, dan 1 siswa Rendah. Centroid akhir yang diperoleh: C0 (Rendah) = (28.40, 95.00, 37.90), C1 (Tinggi) = (88.93, 99.16, 88.81), dan C2 (Sedang) = (83.16, 96.23, 83.67). Visualisasi PCA 2D memperlihatkan tiga kelompok dengan karakteristik berbeda yang akan disajikan pada Gambar 4.



Gambar 4. Visualisasi Klaster Prestasi Siswa (PCA 2D)

Berdasarkan Gambar 4, proses pengelompokan siswa selesai dengan distribusi Klaster 2 (Sedang) sebanyak 62 siswa, Klaster 1 (Tinggi) 50 siswa, dan Klaster 0 (Rendah) 1 siswa. Visualisasi PCA 2D memperlihatkan sebaran data yang terbentuk menjadi tiga kelompok dengan karakteristik berbeda, di mana kelompok Tinggi dan Sedang terpisah secara jelas, sedangkan kelompok Rendah terisolasi karena hanya memiliki satu anggota. Seluruh data hasil pengelompokan digunakan sebagai basis data utama pengujian Perceptron.

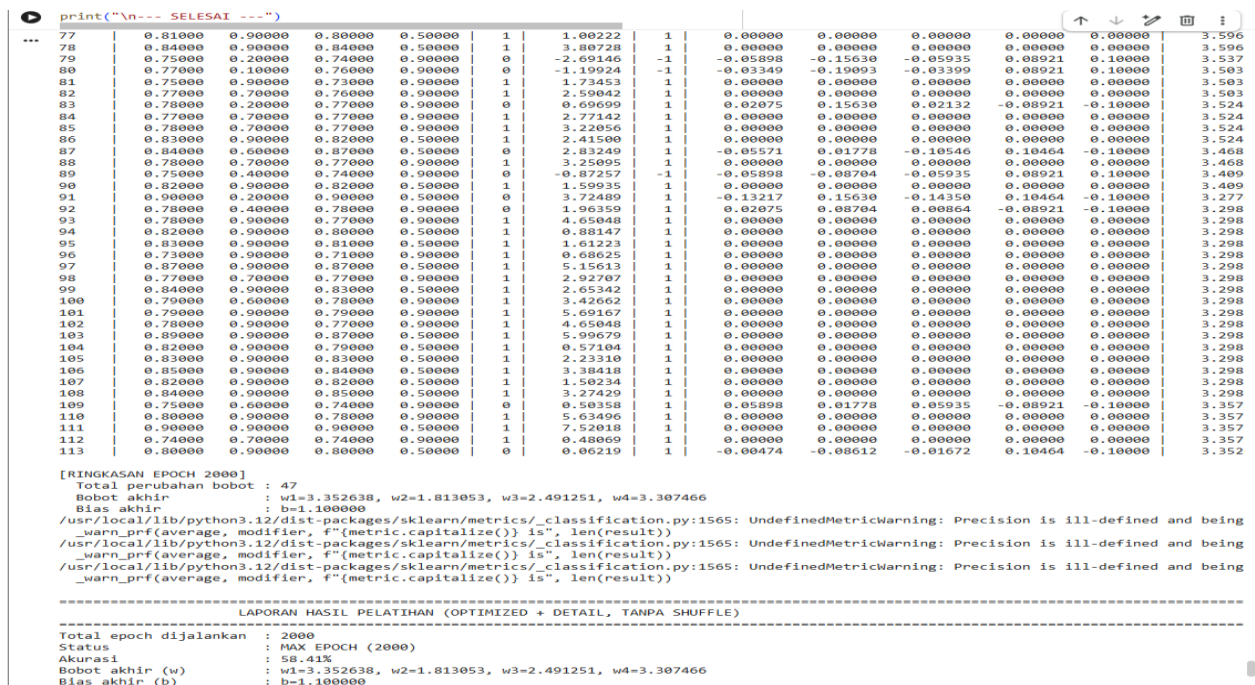
Hasil Algoritma Perceptron

Algoritma Perceptron diterapkan untuk klasifikasi prestasi siswa berdasarkan label hasil K-Means dengan 4 fitur (X1, X2, X3, X4). Parameter pelatihan: learning rate 0.1, threshold 0.0, dan maksimum epoch 2000. hasil pelatihan disajikan pada Tabel 4:

Tabel 4. Hasil Pelatihan Perceptron

Metrik	Nilai
Total epoch	2000
Status	Maksimum epoch – belum konvergen
Akurasi	58.41%
Bobot akhir	w1=3.3526, w2=1.8131, w3=2.4913, w4=3.3075
Bias akhir	b = 1.100000

Tabel 4 menunjukkan pelatihan dihentikan pada epoch 2000 karena belum konvergen, mengindikasikan data tidak sepenuhnya linear separable. Bobot akhir seluruhnya positif dengan fitur x4 dan x1 paling dominan. Bias positif menyebabkan kecenderungan prediksi ke kelas Tinggi. Adapun Hasil Perceptron menggunakan google colab ditampilkan pada Gambar 5.



Gambar 5. Hasil Perceptron

Gambar 5 pada epoch ke-2000 proses pelatihan perceptron dengan 4 fitur dan target 3 kelas (-1, 0, 1) masih menunjukkan fluktuasi yang signifikan dengan total perubahan bobot mencapai 47 dari 113 sampel (41.59%). Grafik perubahan bobot per epoch menunjukkan penurunan yang tidak stabil, mengindikasikan model belum mencapai konvergensi penuh. Dominasi kelas mayoritas (65

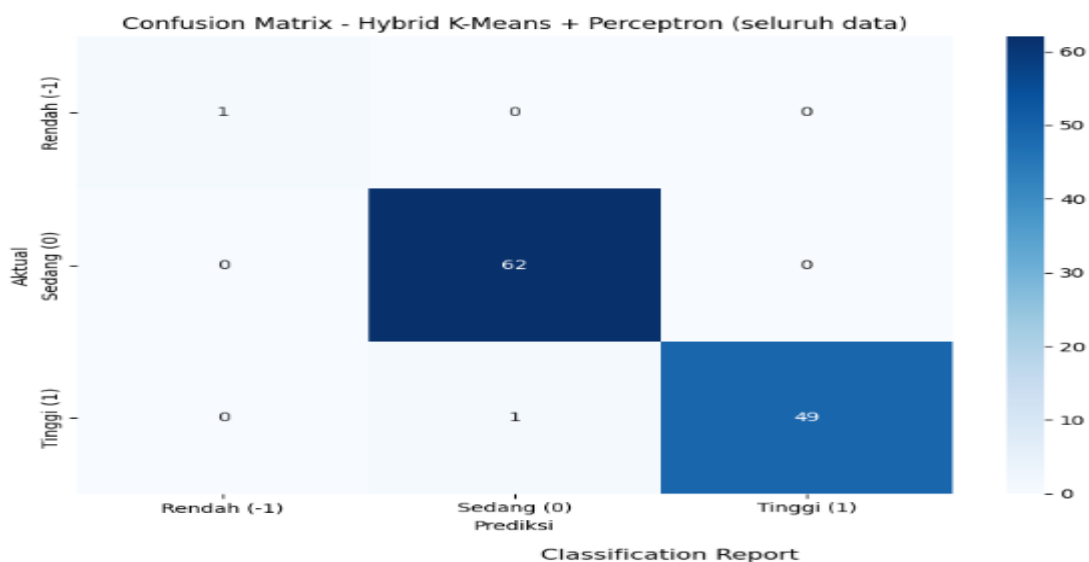
siswa Tinggi) menyebabkan model bias terhadap kelas tersebut, di mana data dengan target 0 (Sedang) sering salah diklasifikasikan sebagai kelas 1 karena nilai y_{in} yang positif. Untuk mengatasi masalah ini, diperlukan penyesuaian parameter atau pendekatan yang lebih kompleks seperti Multilayer Perceptron (MLP).

Hasil Evaluasi

Evaluasi performa model hybrid K-Means + Perceptron dilakukan untuk mengukur keandalan sistem dalam mengklasifikasikan prestasi siswa MTsS Koto Anau. Proses evaluasi menggunakan data uji yang belum pernah dilihat oleh model dengan membandingkan hasil prediksi terhadap label aktual menggunakan Confusion Matrix dan Classification Report. Adapun tampilannya pada Gambar 6 dan 7.

	Accuracy	Precision	Recall	F1-Score	Support
Rendah (-1)	1.00	1.00	1.00	1.00	1
Sedang (0)	0.99	0.98	1.00	0.99	62
Tinggi (1)	0.99	1.00	0.98	0.99	50
macro avg	0.99	0.99	0.99	0.99	113
weighted avg	0.99	0.99	0.99	0.99	113

Gambar 6. Hasil Evaluasi Model Hybrid pada Google Colab



Gambar 7 Hasil Evaluasi Model Hybrid pada Google Colab

Gambar 5 dan 6 menunjukkan hasil evaluasi model hybrid yang sangat baik dengan akurasi global mencapai 99%. Metrik precision, recall, dan F1-score untuk masing-masing kelas menunjukkan nilai mendekati sempurna, yaitu 0.98 hingga 1.00. Kelas Rendah (-1) memiliki precision, recall, dan F1-score sempurna (1.00). Kelas Sedang (0) mencapai precision 0.98, recall 1.00, dan F1-score 0.99. Kelas Tinggi (1) memperoleh precision 1.00, recall 0.98, dan F1-score 0.99. Nilai macro avg dan weighted avg yang mencapai 0.99 menegaskan model tidak mengalami bias terhadap kelas mayoritas dan mampu mempertahankan performa tinggi secara konsisten di seluruh kategori prestasi siswa.

Analisis lebih lanjut melalui Confusion Matrix mengonfirmasi keunggulan model ini dalam mengklasifikasikan prestasi siswa. Dari total 113 data uji, hanya terjadi 1 kesalahan klasifikasi yaitu pada kelas Tinggi yang salah diprediksi sebagai Sedang, sementara kelas Rendah dan Sedang berhasil diklasifikasikan sempurna masing-masing 1 dan 62 siswa. Tingkat kesalahan yang sangat rendah ini membuktikan bahwa pipeline hybrid K-Means + Perceptron efektif dan dapat diandalkan. Hasil ini sekaligus menjawab tujuan penelitian pada pendahuluan, yaitu merancang kerangka hibrida dua tahap yang mampu mengelompokkan siswa ke dalam tiga kluster prestasi (Tinggi,

Sedang, Rendah) dan memprediksi kategori prestasi siswa baru secara akurat, objektif, dan efisien, serta memberikan kontribusi praktis berupa model analitik ringan yang dapat diadopsi sekolah sebagai alat bantu pengambilan keputusan akademik berbasis data untuk perbaikan kualitas pembelajaran di MTsS Koto Anau.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa implementasi metode hybrid K-Means Clustering dan algoritma Perceptron berhasil menganalisis penentuan prestasi siswa di MTsS Koto Anau secara objektif dan akurat. Proses clustering dengan K-Means berhasil mengelompokkan 113 siswa ke dalam tiga klaster prestasi, yaitu Tinggi (50 siswa), Sedang (62 siswa), dan Rendah (1 siswa) berdasarkan kemiripan karakteristik akademik yang diukur melalui nilai rata-rata rapor, persentase kehadiran, dan nilai ujian akhir semester. Selanjutnya, model klasifikasi Perceptron yang dilatih menggunakan label hasil clustering menunjukkan performa yang sangat baik dengan akurasi global mencapai 99%, precision, recall, dan F1-score masing-masing kelas mendekati sempurna (0.98–1.00), serta hanya terjadi 1 kesalahan klasifikasi dari total 113 data uji. Dengan demikian, pendekatan hybrid yang diusulkan terbukti efektif dan andal dalam memetakan serta memprediksi kategori prestasi siswa, sehingga dapat menjadi solusi berbasis data yang mendukung pengambilan keputusan akademik di lingkungan madrasah secara lebih tepat, personal, dan terukur.

Saran

Berdasarkan keterbatasan yang ditemukan selama penelitian, disarankan bagi penelitian selanjutnya untuk menambahkan variabel non-akademik seperti motivasi belajar, minat siswa, dukungan orang tua, atau aktivitas ekstrakurikuler guna memperkaya fitur dan meningkatkan generalisasi model. Selain itu, mengingat algoritma Perceptron dasar memiliki keterbatasan pada data yang tidak linear separable, pengembangan lebih lanjut dapat menggunakan arsitektur Multilayer Perceptron (MLP) atau algoritma klasifikasi lain seperti Support Vector Machine (SVM) dan Random Forest untuk menangani kompleksitas data yang lebih tinggi. Bagi pihak MTsS Koto Anau, model ini dapat diimplementasikan sebagai sistem pendukung keputusan berbasis web atau aplikasi dashboard interaktif yang mudah digunakan oleh guru dan wali kelas dalam mengidentifikasi siswa berprestasi maupun siswa yang memerlukan intervensi pembelajaran secara dini, sehingga kualitas pendidikan di madrasah dapat terus ditingkatkan secara berkelanjutan dan berkeadilan.

DAFTAR PUSTAKA

- Anastassia, S., Kharis, A., Haqqi, A., Zili, A., Matematika, P. S., Terbuka, U., Matematika, P. S., & Indonesia, U. (2024). *Tantangan dan peluang educational data mining dan learning analytics pada pendidikan jarak jauh di indonesia*. 1(2), 218–225.
- Ardianti, M., Nurhayati, O. D., & Warsito, B. (2023). *Model Prediksi Kinerja Siswa Berdasarkan Data Log LMS Menggunakan Ensemble Machine Learning*. 12(3), 562–571.
- Aulia, M. F., Studi, P., Komputer, I., Islam, U., & Sumatera, N. (2025). *Klasifikasi Biner Bipolar Pada Data Kusioner Pelamar Asisten Laboratorium Menggunakan Model Hebbian Dan Perceptron*. 8(2), 137–150.
- Dewi, F. P., Aryni, P. S., & Umaidah, Y. (2022). *Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran*. 7(2), 111–121.
- Dihe, L., & Wangdra, Y. (2023). *Pendidikan Adalah Faktor Penentu Daya Saing Bangsa*. September, 84–90.
- Directions, F. (2023). *Understanding of Machine Learning with Deep Learning* :

- Frananda, A., Naibaho, A., & Zahra, A. (2023). *PREDIKSI KELULUSAN SISWA SEKOLAH MENENGAH PERTAMA MENGGUNAKAN MACHINE LEARNING*. 11(3).
- Gunandar, H., Nerisafitra, P., Belakang, A. L., & Sistem, A. M. P. (2025). *Eksplorasi Teknik Berbasis K-Means Clustering Untuk Penilaian Prestasi Siswa*. 06, 1076–1082.
- Halimah, L. N., Riyadi, S., & Jurjani, A. F. (2025). *IMPLEMENTASI PENGGUNAAN MACHINE LEARNING DALAM PEMBELAJARAN : SUATU TELAAH DESKRIPTIF*. 1(1), 1–10.
- Irvan, I. A., & Annur, S. (2024). *ENHANCING PERSONALIZED LEARNING THROUGH ARTIFICIAL INTELLIGENCE (AI) IN EDUCATION 5 . 0 : A FRAMEWORK FOR ADAPTIVE LEARNING*. 1, 670–680.
- Islam, U., & Sunan, N. (2023). *Pengaruh Motivasi Belajar Terhadap Peningkatan Prestasi Akademik Siswa*. 7(2).
- Junaidi, S., Anggela, R. V., & Kariman, D. (2024). *JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST) Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes , Random Forest , Support Vector Machine (SVM) dan Artificial Neural Network (ANN)*. 5(1), 109–119.
- Mujiyono, S., Sanjaya, U. P., Wibisono, I. S., & Setyowati, H. (2025). *Prediksi Fluktuasi Berat Badan Berdasarkan Pola Hidup Menggunakan Model XGBoost dan Deep Learning*. 221–233. <https://doi.org/10.33364/algoritma/v.22-1.2253>
- Nasional, J., Informasi, S., Pasek, M., Ariawan, A., Bagus, I., Peling, A., & Brahupadhy, G. (2023). *Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus : Matakuliah Pemrograman Dasar)*. 02, 122–131.
- Nurhalizah, R. S., & Ardianto, R. (2024). *Analisis Supervised dan Unsupervised Learning pada Machine Learning : Systematic Literature Review*. 4(1), 61–72.
- Royhanuddin, F., Ari, W., & Harahap, A. (2024). *Evaluasi Pembelajaran Pada Mata Pelajaran Pendidikan Agama Islam dan Implikasinya Terhadap Motivasi Belajar Siswa MAN 1 Padangsidempuan*. 2(2), 17–25.
- Safitri, A. D., Hendrik, M., & Pratama, Y. B. (2025). *Implementasi Data Science untuk Analisis Efektivitas Media Pembelajaran Etnobook Digital dalam Mendukung Literasi Siswa Sekolah Dasar TIN : Terapan Informatika Nusantara*. 5(11), 922–932. <https://doi.org/10.47065/tin.v5i11.7182>
- Saraf, J., Dengan, T., & Perceptron, A. (2024). *Pengenalan pola bunga berbasis citra menggunakan jaringan saraf tiruan dengan algoritma*. 09, 365–373.
- Sari, A. W., Arta, D. J., Studi, P., Bahasa, P., Keguruan, F., Pendidikan, I., Muhammadiyah, U., & Utara, S. (2025). *Implementasi Deep Learning : Suatu Inovasi Pendidikan*. 13(01).
- Seminar, P., Sanata, N., Berbagi, D., & Depan, P. M. (2024). *Learning Analytics*. 2.
- Sinaga, A. O., & Passarella, R. (2025). *Analysis of Passenger and Bus Numbers Using Unsupervised Learning Analisis Hubungan Antara Jumlah Bus dan Jumlah Penumpang Menggunakan Unsupervised Learning*. 5(July), 1029–1036.
- Suarna, N., Rahaningsih, N., & Suarna, A. A. (2025). *OPTIMALISASI PRESTASI AKADEMIK SISWA MELALUI PENGELOMPOKAN INDEKS PRESTASI DENGAN K-MEANS CLUSTERING OPTIMIZATION OF STUDENTS ' ACADEMIC PERFORMANCE THROUGH GRADE POINT INDEX CLUSTERING USING K-MEANS CLUSTERING*. 4(2), 198–207.
- Sukmaningtyas, Y. N., Akbar, R. M., & Rohma, G. (2024). *Penerapan Predictive Analytics untuk Analisis Faktor-faktor yang Mempengaruhi Performa Akademik Siswa menganalisis faktor-faktor yang mempengaruhi performa akademik siswa telah*. 4(2), 127–145.
- Widya, L., Utami, T., Kapoyos, P. B., & Rafikri, M. D. (2024). *Pengenalan Konsep Kecerdasan Buatan (AI) Dalam Simulasi dan Kasus Kehidupan Sehari – Hari Bagi Siswa – Siswi SMK Nusantara 1 Ciputat*. 2(2), 6–15.
- Yusup, M. (2025). *Sistem Peringatan Dini Penurunan Prestasi dan Minat Belajar pada Siswa Sekolah Menengah di Medan : Pendekatan Explainable Machine Learning*. 4(2), 1423–1431.