Jurnal Penelitian Ilmu Pendidikan Indonesia
Volume 5, Nomor 3, Tahun 2026, Halaman 2210 – 2222
https://jpion.org/index.php/jpi
E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

Analisis Penerapan *Machine Learning* Berbasis Data Historis untuk Prediksi Hasil Belajar Siswa: Kajian Studi Literatur

Andreas Damanik[1)*], Mikhael Batubara[2)], Ariel Trifandi Zalukhu[3)], Fransiskus Jul Rikar Gea[4)], Afnel Yoniva Marbun[5)], Putri Adelia br. Lumbantobing[6)], Elsa Sabrina[7)]

[1,2,3,4,5,6,7])Pendidikan Teknologi Informasi dan Komputer, Fakultas Teknik, Universitas Negeri Medan, Indonesia

* Andreas Damanik
Email : andreasdmk.5243551034@mhs.unimed.ac.id,
mikhaelbb.5243151045@mhs.unimed.ac.id
arielz.5243151042@mhs.unimed.ac.id
fransiskus.5243151002@mhs.unimed.ac.id
putriadelia.5243151009@mhs.unimed.ac.id
afnel.5243151046@mhs.unimed.ac.id

Abstrak

Prediksi hasil belajar diperlukan untuk membantu mengidentifikasi peserta didik yang berisiko mengalami kesulitan akademik dan mendukung pemberian intervensi secara lebih awal. Penelitian ini bertujuan menganalisis penerapan machine learning berbasis data historis dalam prediksi hasil belajar peserta didik dengan mengidentifikasi algoritma yang digunakan, variasi data historis, performa model, serta kontribusinya terhadap early warning system. Penelitian menggunakan metode Systematic Literature Review dengan kerangka PICOC dan pedoman PRISMA. Pencarian dilakukan melalui Google Scholar dan Crossref terhadap publikasi tahun 2021-2026 menggunakan bantuan Publish or Perish. Dari proses seleksi diperoleh 12 jurnal utama yang dianalisis secara naratif dan deskriptif. Hasil sintesis menunjukkan bahwa Naïve Bayes merupakan algoritma yang paling sering digunakan, diikuti Random Forest dan Decision Tree. Data historis mencakup data akademik, kehadiran, aktivitas pembelajaran atau log LMS, serta data sosial-demografis. Performa model bervariasi bergantung pada karakteristik data, preprocessing, dan metode pengujian sehingga nilai akurasi tidak dapat dibandingkan secara langsung. Kontribusi terhadap early warning system masih terbatas, dengan implementasi yang paling jelas berupa sistem berbasis web untuk membantu guru mengidentifikasi peserta didik berisiko.

***Kata kunci**: Machine Learning, Data Historis, Prediksi Hasil Belajar, Early Warning System*

Abstract

Learning outcome prediction is needed to help identify students who are at risk of experiencing academic difficulties and to support earlier intervention. This study aims to analyze the application of historical data-based machine learning for predicting student learning outcomes by identifying the algorithms used, variations in historical data, model performance, and their contribution to an early warning system. This study employed a Systematic Literature Review using the PICOC framework and PRISMA guidelines. Literature searches were conducted through Google Scholar and Crossref for publications from 2021 to 2026 with the assistance of Publish or Perish. The selection process resulted in 12 primary articles that were analyzed using narrative and descriptive synthesis. The results show that Naïve Bayes was the most frequently used algorithm, followed by Random Forest and Decision Tree. Historical data included academic data, attendance, learning activities or LMS logs, and socio-demographic data. Model performance varied depending on data characteristics, preprocessing, and evaluation methods; therefore, accuracy values could not be directly compared across studies. The contribution to early warning systems remains limited, with the clearest implementation involving a web-based system to assist teachers in identifying students at academic risk.

***Keywords**: Machine Learning, Historical Data, Learning Outcome Prediction, Early Warning System*

PENDAHULUAN

Keberhasilan belajar merupakan tujuan akhir pembelajaran, dan capaiannya terlihat pada hasil belajar sebagai luaran perubahan tingkah laku setelah mengikuti pembelajaran (Rajagukguk, 2021). Prediksi hasil belajar penting sebagai dasar intervensi dini bagi peserta didik yang berpotensi gagal dan penyesuaian strategi pembelajaran (Rajagukguk, 2021), serta

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

membantu institusi pendidikan tinggi menyusun rencana intervensi sejak awal (Albreiki et al., 2021). Prediksi tersebut hanya dapat dilakukan jika tersedia informasi yang menunjukkan pola pencapaian akademik, sehingga diperlukan pemanfaatan data yang sudah dimiliki lembaga pendidikan.

Lembaga pendidikan menghasilkan data dalam jumlah besar, terutama sejak penggunaan sistem manajemen pembelajaran (*learning management system*) (Alalawi et al., 2023). Data historis, yaitu data masa lalu yang diproyeksikan pada suatu model untuk memperkirakan hasil di masa depan (Sghir et al., 2023), mencakup nilai pada mata kuliah, semester, atau tahun sebelumnya dan menjadi jenis data yang paling sering digunakan bersama data sistem pembelajaran dan data demografi (Alalawi et al., 2023). Platform daring juga merekam aktivitas belajar, seperti frekuensi akses dan lama waktu belajar, sehingga data dapat bersifat statis, seperti data pendaftaran dan demografi, atau dinamis, seperti catatan interaksi dengan sistem pembelajaran (Albreiki et al., 2021). Data yang beragam dan berjumlah besar ini membutuhkan metode yang mampu mengenali pola di dalamnya.

Metode tersebut adalah *machine learning*, yaitu pendekatan yang belajar dari data untuk menemukan pola dan memprediksi hasil (Albreiki et al., 2021), dan menurut (Sandra et al., 2021) mengisi kekurangan pemrograman eksplisit untuk data pendidikan yang besar dan terus berubah. Dalam prediksi hasil belajar, klasifikasi paling sering digunakan dengan algoritma seperti *artificial neural network*, *support vector machine*, *decision tree*, *naïve Bayes*, dan *logistic regression*, dengan hasil umumnya dua atau tiga kategori, misalnya lulus dan gagal (Sandra et al., 2021; Zhang et al., 2021) menambahkan bahwa prediksi nilai dapat pula dirumuskan sebagai regresi atau klasterisasi. Metode ini membantu guru memperkirakan capaian siswa lebih awal (Sandra et al., 2021) dan dinilai berperan penting dalam memprediksi siswa berisiko dan putus studi (Albreiki et al., 2021). Penerapan tersebut telah dikaji dalam banyak penelitian dengan fokus berbeda sehingga perlu dibandingkan.

(Sekeroglu et al., 2021) mengkaji model, metode evaluasi, dan validasi dalam prediksi hasil belajar, lalu mengusulkan standardisasi evaluasi dan validasi. (Alalawi et al., 2023) menemukan bahwa hanya sedikit studi yang menindaklanjuti hasil prediksi bagi siswa berisiko, sedangkan (Turkmenbayev et al., 2025) memusatkan kajian pada mahasiswa program teknik dan menemukan validasi eksternal yang terbatas. Kesenjangan penelitian (*research gap*) terletak pada belum adanya sintesis yang secara khusus menghubungkan jenis data historis, pendekatan *machine learning*, dan praktik evaluasinya. Kebaruan (*novelty*) artikel ini adalah menempatkan data historis sebagai penghubung ketiga unsur tersebut. Karena itu, penelitian ini bertujuan menganalisis penerapan *machine learning* berbasis data historis untuk prediksi hasil belajar siswa melalui studi literatur.

Untuk itu, *educational data mining* (EDM) sebagai dasar kajian ini adalah bidang yang mempelajari pola tersembunyi dalam berbagai situasi Pendidikan (Zhang et al., 2021) dengan memakai metode *machine learning* dan statistik (Albreiki et al., 2021). *Machine learning* membangun model matematis yang dapat menyesuaikan diri dengan perubahan data, dan terbagi menjadi *supervised learning* yang memakai data berlabel serta *unsupervised learning* yang tidak memakai label (Sandra et al., 2021). Prediksi kinerja siswa (*student performance prediction*) adalah upaya memperkirakan nilai yang akan diperoleh siswa, dengan tahapan pengumpulan data, perumusan masalah, pemodelan, prediksi, dan penerapan (Zhang et al., 2021). Konsep ini menjadi dasar untuk menelaah bagaimana penelitian terdahulu menerapkan metode prediksi pada data pendidikan.

Pemilihan algoritma bergantung pada jenis data dan tujuan prediksi (Rajagukguk, 2021), dan peneliti umumnya mencoba beberapa algoritma lalu membandingkan hasilnya (Sghir et al., 2023). Pada data historis, (Zhang et al., 2021) menemukan bahwa penambahan nilai mata kuliah sebelumnya meningkatkan kinerja model, sedangkan data latar belakang berpengaruh kecil. Jenis data yang dominan berbeda antarkajian, yaitu data historis nilai pada (Alalawi et al., 2023) dan data perilaku daring pada (Sghir et al., 2023), yang kemungkinan berkaitan dengan perbedaan

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

cakupan kedua kajian. Jaringan saraf tiruan banyak digunakan dan akurat, tetapi membutuhkan komputasi besar dan sulit dijelaskan (Sghir et al., 2023), dengan keterbatasan berulang berupa data kecil dan generalisasi terbatas (Sghir et al., 2023), sampel satu institusi (Turkmenbayev et al., 2025), serta indikator penilaian yang berbeda antarkelas dan Lembaga (Rajagukguk, 2021). Dengan demikian, kajian ini difokuskan pada hubungan antara data historis, pendekatan *machine learning*, dan kualitas evaluasi dalam prediksi hasil belajar siswa.

METODE PENELITIAN

Penelitian ini menerapkan metode *Systematic Literature Review* (SLR) (Page et al., 2021). Untuk menentukan batasan kajian, perumusan masalah dibantu oleh kerangka PICOC (*Population, Intervention, Comparison, Outcome, Context*) (Kitchenham, B., & Charters, 2007). Kerangka ini menentukan kelompok yang terdiri dari peserta didik pada berbagai jenjang, intervensi adalah penerapan *Machine Learning* berdasarkan data historis, perbandingan dilakukan untuk mengevaluasi berbagai metode klasifikasi dan regresi, hasil yang diharapkan mengarah pada jenis kemampuan dan akurasi prediksi, serta lingkungannya adalah dalam konteks pendidikan resmi (Kitchenham, B., & Charters, 2007; Page et al., 2021).

Penggunaan PICOC menghasilkan empat pertanyaan penelitian (*Research Question*/RQ) yang membantu alur sintesis (Kitchenham, B., & Charters, 2007). Keempat pertanyaan tersebut meliputi:

1) RQ1: Algoritma *Machine Learning* apa saja yang paling sering digunakan untuk memprediksi hasil belajar peserta didik?
2) RQ2: Bagaimana pengaruh variasi data historis (data akademik statis, log aktivitas dinamis LMS, dan data sosio-demografi) terhadap pemodelan prediksi hasil belajar?
3) RQ2: Bagaimana pengaruh variasi data historis (data akademik statis, log aktivitas dinamis LMS, dan data sosio-demografi) terhadap pemodelan prediksi hasil belajar?
4) RQ3: Seberapa tinggi tingkat performa dan keandalan (berdasarkan metrik seperti *Accuracy*, *Precision*, *Recall*, AUC, atau RMSE) dari masing-masing algoritma *Machine Learning* yang dievaluasi?

Untuk menjawab pertanyaan dari penelitian tersebut, dilakukan pencarian literatur dengan basis data Google Scholar serta Crossref dalam kurun waktu 2021 hingga 2026 dengan bantuan Publish or Perish. Pencarian menggunakan kata kunci (*keyword*) kombinasi ("*Machine Learning*, *Naïve Bayes*, *Random Forest*, Data historis, Hasil belajar") (Page et al., 2021). Kemudian hasil dari pencarian disaring menggunakan kriteria inklusi dan eksklusi yang diperlihatkan pada Tabel 1

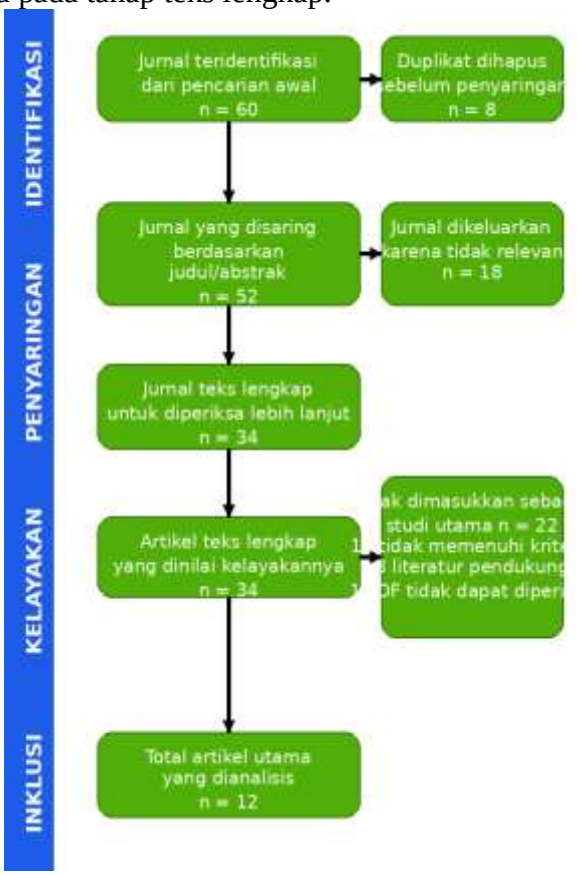
Tabel 1. Matrik Inklusi dan Eksklusi

Kriteria	Inklusi	Eksklusi
Relevansi topik	Artikel yang membahas penerapan *Machine Learning* untuk memprediksi atau menganalisis hasil belajar maupun peforma siswa.	Artikel tidak membahas prediksi hasil belajar atau performa akademik siswa menggunakan *Machine Learning*.
Penggunaan data	Artikel menggunakan data historis atau data akademik siswa sebagai dasar analisis atau prediksi.	Artikel hanya menggunakan metode statistik atau metode lain tanpa penerapan *Machine Learning*.

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

Metode penelitian	Artikel menjelaskan penggunaan metode atau algoritma *Machine Learning*, *Decision Tree*, *Random Forest*, *Naive Bayes*, atau algoritma lainnya.	Artikel tidak menggunakan data historis atau data akademik yang relevan untuk prediksi hasil belajar.
Jenis publikasi	Artikel berupa jurnal ilmiah atau artikel penelitian yang relevan dengan topik kajian.	Dokumen berupa skripsi, tesis, buku, artikel opini, atau publikasi lain di luar jenis sumber yang ditetapkan.
Ketersediaan artikel	Artikel memiliki teks lengkap yang dapat diakses untuk dianalisis.	Artikel tidak memiliki teks lengkap yang dapat diakses sehingga isi dan metodenya tidak dapat diperiksa.
Periode publikasi	Artikel diterbitkan dalam rentang tahun yang telah ditetapkan dalam strategi pencarian penelitian.	Artikel diterbitkan di luar rentang tahun yang telah ditetapkan.

Untuk menjamin kualitas dan relevansi literatur yang dikaji, kriteria inklusi dan eksklusi yang telah ditetapkan mengadopsi protokol *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) (Page et al., 2021). Proses seleksi diawali dengan 60 artikel yang teridentifikasi dari pencarian. Setelah penghapusan 8 duplikat, sebanyak 52 artikel disaring berdasarkan judul dan abstrak dan 18 artikel dikeluarkan karena tidak relevan. Sebanyak 34 artikel kemudian diperiksa pada tahap teks lengkap.



Gambar 1. Diagram PRISMA

Berdasarkan alur seleksi PRISMA pada Gambar 1, sebanyak 34 artikel mencapai tahap penilaian teks lengkap. Dari jumlah tersebut, 22 artikel tidak dimasukkan sebagai studi utama,

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

yang terdiri atas 13 artikel yang tidak memenuhi kriteria inklusi, 8 artikel yang digunakan hanya sebagai literatur pendukung pendahuluan, dan 1 artikel yang dikeluarkan karena PDF tidak dapat diperiksa. Dengan demikian, 12 jurnal utama digunakan untuk ekstraksi dan sintesis data dalam menjawab RQ1 hingga RQ4.

Setelah penetapan studi utama, dilakukan ekstraksi data terhadap 12 jurnal penelitian utama. Formulir ekstraksi digunakan untuk mencatat penulis dan tahun, jenis data historis, algoritma *machine learning*, metrik dan hasil performa, serta kontribusi terhadap *early warning system*. Struktur ekstraksi disajikan pada Tabel 2.

Tabel 2. Ekstraksi Data Artikel Utama

Penulis & Tahun	**Jenis Data Historis (RQ2)**	**Algoritma ML (RQ1)**	**Metrik dan Hasil Peforma (RQ3)**	**Kontribusi pada *Early Warning System* (RQ4)**
Mifta Ardianti, Oky Dwi Nurhayati, Budi Warsito (2023)	Log LMS Kalboard 360; perilaku/interaksi siswa; 480 data setelah pemrosesan.	*Bagging* (*Random Forest*), *Boosting* (*AdaBoost*), *Voting* (*Logistic Regression* + SVM).	*Bagging*: *accuracy* 81,25%; *precision* 0,810; *recall* 0,812; *F-measure* 0,809.	Prediksi kinerja *high, medium, low* dari log LMS untuk deteksi dini; tidak mengembangkan EWS mandiri.
Slamet Kusworo, Nugroho Adhi Santoso, Rifki Dwi Kurniawan (2024)	100 siswa; nilai tugas, ulangan harian, UTS, UAS, kehadiran, partisipasi diskusi, atribut pribadi/non-akademik.	*Random Forest*.	*Accuracy* 93,33%; *precision* 93,33%; *recall* 100%; *F1-score* 96,55%.	Prediksi nilai akhir sebelum periode penilaian untuk mengenali siswa berisiko dan memungkinkan intervensi.
Septiar Maulana, Agyztia Premana, Bambang Irawan (2025)	519 data historis; nilai semester 5, prestasi akademik, keaktifan, kehadiran.	*Decision Tree*.	*Accuracy* 97,80%; *precision* 98,38% dan 94,55% per kategori; *recall* dan *F1-score* dilaporkan per kategori.	Mendukung seleksi siswa terbaik secara berbasis data; tidak dilaporkan sebagai EWS siswa berisiko.
Yetti Fitriani, Sarjon Defit, Gunadi Widi Nurcahyo (2021)	61 siswa; absensi, tugas, ulangan harian, ujian.	*Decision Tree* C4.5.	*Accuracy* 83,33%; *precision* 100% dan 61,54% per kelas; *recall* 77,27% dan 100% per kelas.	Prediksi hasil belajar untuk mendukung kebijakan sekolah dan pencegahan risiko belajar.
Arham Rahim, Mustakim (2025)	87 siswa; nilai, kehadiran, lingkungan sosial, ekonomi, motivasi, dukungan keluarga dan faktor eksternal lain.	*Naïve Bayes*, *K-Nearest Neighbor*, *Support Vector Machine*, *Neural Network*.	Skenario S3: SVM mencapai *accuracy* 100%; *Naïve Bayes* 93% direkomendasikan karena lebih stabil terhadap risiko *overfitting*.	Mendeteksi siswa yang berpotensi tidak mencapai KKM untuk mendukung intervensi pembelajaran.
Agusti Frananda Alfonsus Naibaho, Amalia Zahra (2023)	Data kelulusan siswa; data pribadi, akademik, dan pekerjaan orang tua.	*Decision Tree*, *Random Forest*, *XGBoost*; GridSearchCV dan *5-fold cross-validation*.	*Random Forest* 99,5%; *XGBoost* 99,0%; *Decision Tree* 96,6%.	Prediksi status kelulusan dan identifikasi faktor yang dapat membantu sekolah menangani risiko kelulusan lebih dini.
Takhamo Gori, Andi Sunyoto, Hanif Al Fatta (2024)	*Dataset* Matematika 395 dan Bahasa Portugis 649 dari dua sekolah menengah di Portugal; nilai G1/G2, absensi,	*Naïve Bayes* dengan *Correlation-Based Feature Selection* (CFS).	Sebelum CFS: *accuracy* 89,27%; *recall* 89,27%; *precision* 89,86%; *F1* 89,47%. Setelah CFS: *accuracy* 91,22%; *recall*	Mendukung tindakan pencegahan terhadap kinerja rendah dan risiko putus sekolah; belum berupa EWS operasional.

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

	kegagalan, waktu belajar, demografi, dan perilaku.		91,22%; *precision* 92,24%; *F1* 91,48%.	
Uswatun Hasanah, Fatmawati (2022)	78 data kuesioner; waktu belajar, pemahaman materi, tugas, lingkungan belajar.	*Decision Tree* C4.5 dan *Naïve Bayes*.	C4.5 *accuracy* 90,00%; *Naïve Bayes* 86,13%; AUC C4.5 0,596.	Mengklasifikasikan penyebab penurunan hasil belajar dan memberi dasar peringatan dini terkait masalah pembelajaran.
Rizky Adinda Azizah, Fitra A. Bachtiar, Sigit Adinugroho (2022)	480 data Kalboard 360; aktivitas LMS, demografi, dan atribut campuran; 80% latih dan 20% uji.	*Neighbor Weighted K-Nearest Neighbor* (NWKNN) dengan seleksi fitur *Information Gain*.	*Accuracy* 0,604; *precision* 0,719; *recall* 0,676; *F-measure* 0,661.	Pelacakan aktivitas LMS untuk klasifikasi kinerja dan penyesuaian pembelajaran; tidak menerapkan EWS secara langsung.
Randi Estian Pambudi, Hendri Purnomo, Rendi Irawan (2024)	Pola kehadiran dan aktivitas belajar siswa.	*Naïve Bayes* dan *Logistic Regression*; artikel juga menyebut kombinasi kedua metode.	Abstrak: NB *accuracy* 85,7%, LR 88,9%, kombinasi 90,3%. Bagian hasil: *Logistic Regression accuracy* 0,98, *F1* 0,98, *precision* 0,981, *recall* 0,98; nilai NB ditulis 940/950 tanpa skala jelas.	Mendukung pengambilan keputusan intervensi berdasarkan pola kehadiran dan aktivitas belajar; tidak ada EWS operasional.
Saeful Anwar, Dian Ade Kurnia, Ahmad Faqih, Siti Rini Sari (2022)	1.231 data hasil belajar *Hybrid* Learning; 12 atribut terkait kompetensi, karakteristik mahasiswa, interaksi, strategi *online*, dan evaluasi.	*Artificial Neural Network* dengan *Multilayer Perceptron*.	*Accuracy* 99,35%; menggunakan validasi silang pada proses pemodelan.	Prediksi hasil belajar *hybrid*; artikel tidak melaporkan penerapan EWS atau intervensi siswa berisiko.
Munadi Fauzan, Dinda Permata Sukma, Eriwandi, Diana Kemala Odang (2026)	200 data; frekuensi *login*, durasi belajar, nilai kuis, nilai tugas, ketepatan waktu pengumpulan, kehadiran, dan keaktifan.	*Decision Tree*, *Random Forest*, *Naïve Bayes*.	*Decision Tree* 88,52%; *Random Forest* 83,61%; *Naïve Bayes* 62,30%.	EWS berbasis web untuk guru; memprediksi tingkat pemahaman tinggi, sedang, dan rendah sebelum ujian akhir serta mendukung pendampingan.

Teknik analisis dan sintesis data menggunakan *narrative synthesis* yang dikombinasikan dengan analisis deskriptif terhadap 12 jurnal utama. Temuan dibandingkan, dikelompokkan, dan dipetakan berdasarkan empat pertanyaan penelitian sehingga hasil yang disajikan tidak berupa ringkasan artikel secara terpisah, tetapi berupa sintesis tematik.

1) Sintesis Algoritma *Machine Learning* (RQ1): menghitung frekuensi penggunaan algoritma pada 12 jurnal utama. Jika satu artikel menggunakan beberapa algoritma, setiap algoritma dihitung satu kali pada artikel tersebut, sedangkan kombinasi atau teknik *ensemble* tidak dihitung sebagai algoritma tambahan apabila hanya merupakan cara penggabungan model.
2) Sintesis Variasi Data Historis (RQ2): mengelompokkan data menjadi data akademik, data kehadiran, data aktivitas atau log pembelajaran, serta data sosial dan demografis. Pengelompokan digunakan untuk membandingkan karakteristik data yang menjadi masukan model.
3) Sintesis Performa Algoritma (RQ3): membandingkan metrik yang dilaporkan, meliputi *accuracy*, *precision*, *recall*, *F1-score*, AUC, dan metrik lain yang tersedia. Angka dari *dataset*

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

dan skenario pengujian yang berbeda tidak diperlakukan sebagai ukuran performa yang dapat dibandingkan secara langsung.

4) Sintesis Implikasi *Early Warning System* (RQ4): mengelompokkan kontribusi model menjadi penerapan EWS secara langsung, dukungan terhadap EWS melalui identifikasi siswa berisiko atau rekomendasi intervensi, dan prediksi saja tanpa penerapan EWS yang jelas.

HASIL DAN PEMBAHASAN

Analisis dilakukan terhadap 12 jurnal utama yang memenuhi fungsi sebagai studi primer dalam sintesis. Ke-12 artikel terbit pada rentang 2021-2026 dan mencakup prediksi kinerja, hasil belajar, pemahaman siswa, serta status kelulusan. Karakteristik data yang digunakan cukup beragam, mulai dari data nilai dan kehadiran hingga log aktivitas *Learning Management System* (LMS), data demografis, serta faktor lingkungan dan keluarga. Perbedaan karakteristik tersebut membuat hasil performa perlu dibaca dalam konteks *dataset*, algoritma, *preprocessing*, dan metode pengujian masing-masing penelitian.

1. Hasil Seleksi dan Karakteristik Artikel

Proses seleksi menghasilkan 12 jurnal utama untuk dianalisis secara mendalam. Jurnal-jurnal tersebut terdiri atas penelitian yang menggunakan data sekolah, data LMS, *dataset* terbuka, maupun data hasil pembelajaran *hybrid*. Rentang jumlah data yang dilaporkan juga berbeda, misalnya 61 siswa pada (Fitriani et al., 2021), 519 data pada (Maulana et al., 2025), 1.044 data awal pada (Gori et al., 2024), dan 1.231 data pada (Anwar et al., 2022). Variasi tersebut menunjukkan bahwa ukuran *dataset* bukan satu-satunya dasar untuk menilai model karena setiap penelitian menggunakan tujuan, fitur, dan skenario evaluasi yang berbeda.

Tabel 3. Karakteristik 12 jurnal utama

Karakteristik	Temuan sintesis
Tahun publikasi	2021–2026
Fokus prediksi	Hasil belajar, kinerja akademik, pemahaman siswa, dan status kelulusan
Jenis data dominan	Nilai/akademik, kehadiran, aktivitas/LMS, serta data sosial-demografis
Algoritma	*Naïve Bayes*, *Random Forest*, *Decision Tree*/C4.5, SVM, *Logistic Regression*, ANN/MLP, KNN/NWKNN, *XGBoost*, *AdaBoost*, dan *ensemble*
Evaluasi	*Accuracy* paling sering dilaporkan; sebagian artikel melaporkan *precision*, *recall*, *F1-score*, AUC, dan validasi silang
EWS	1 artikel menerapkan EWS berbasis web secara langsung; artikel lain terutama memberikan dukungan prediktif atau rekomendasi intervensi

Secara umum, penelitian primer memperlihatkan bahwa prediksi hasil belajar telah diterapkan pada konteks yang berbeda. (Ardianti et al., 2023) memanfaatkan log LMS dan *ensemble*, sedangkan (Azizah et al., 2022) menggunakan aktivitas LMS dengan NWKNN. Penelitian lain lebih menekankan data akademik dan kehadiran, seperti (Fitriani et al., 2021; Kusworo et al., 2024; Maulana et al., 2025). (Rahim & Mustakim, 2025) memperluas masukan model dengan faktor internal dan eksternal, sementara (Anwar et al., 2022) menggunakan atribut pembelajaran *hybrid* yang lebih beragam.

2. RQ1: Algoritma *Machine Learning* yang Digunakan

Frekuensi algoritma dihitung berdasarkan kemunculan algoritma pada setiap artikel utama. Jika sebuah artikel menguji beberapa algoritma, masing-masing algoritma dihitung satu kali. Teknik penggabungan seperti *ensemble* tidak dihitung sebagai algoritma tambahan ketika algoritma pembentuknya sudah dicatat, sehingga perhitungan tetap berfokus pada model yang benar-benar diuji.

Tabel 4. Distribusi penggunaan algoritma pada 12 jurnal utama

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

Algoritma/famili	Jumlah artikel	Artikel utama
Naïve Bayes	5	Rahim & Mustakim; Gori et al.; Hasanah & Fatmawati; Pambudi et al.; Fauzan et al.
Random Forest	4	Ardianti et al.; Kusworo et al.; Naibaho & Zahra; Fauzan et al.
Decision Tree	3	Maulana et al.; Naibaho & Zahra; Fauzan et al.
C4.5	2	Fitriani et al.; Hasanah & Fatmawati
Support Vector Machine	2	Ardianti et al.; Rahim & Mustakim
Logistic Regression	2	Ardianti et al.; Pambudi et al.
KNN/NWKNN	2	Rahim & Mustakim; Azizah et al.
Neural Network/ANN/MLP	2	Rahim & Mustakim; Anwar et al.
XGBoost	1	Naibaho & Zahra
AdaBoost	1	Ardianti et al.

Hasil perhitungan menunjukkan bahwa *Naïve Bayes* merupakan algoritma yang paling sering muncul, diikuti *Random Forest* dan *Decision Tree*. *Naïve Bayes* digunakan pada lima artikel dengan karakteristik data yang berbeda, sedangkan *Random Forest* digunakan pada empat artikel dan beberapa penelitian membandingkannya dengan model lain. *Decision Tree* dan C4.5 juga cukup dominan karena digunakan untuk klasifikasi tingkat hasil belajar atau kelulusan yang bersifat kategorikal. Temuan ini sejalan dengan kecenderungan dalam kajian pendahuluan yang menunjukkan dominasi metode klasifikasi, tetapi pada 12 studi utama tidak terdapat satu algoritma yang konsisten menjadi model terbaik pada seluruh *dataset* (Alalawi et al., 2023; Rajagukguk, 2021).

Penelitian (Hasanah & Fatmawati, 2022) juga menerapkan algoritma C4.5 dan *Naïve Bayes* untuk memprediksi penurunan hasil belajar siswa secara daring. Hasil pengujian menggunakan *RapidMiner* menunjukkan bahwa algoritma C4.5 memperoleh akurasi sebesar 90,00%, sedangkan algoritma *Naïve Bayes* memperoleh akurasi sebesar 86,13%. Hasil tersebut menunjukkan bahwa C4.5 memiliki akurasi yang lebih tinggi dibandingkan *Naïve Bayes* pada penelitian tersebut, sehingga memperkuat temuan bahwa performa algoritma dapat berbeda meskipun digunakan untuk tujuan prediksi hasil belajar yang sama.

Perbedaan performa antarmodel memperlihatkan bahwa frekuensi penggunaan tidak identik dengan keunggulan universal. (Naibaho & Zahra, 2023) memperoleh 99,5% dengan *Random Forest*, sedangkan (Maulana et al., 2025) memperoleh 97,80% dengan *Decision Tree*. (Rahim & Mustakim, 2025) memperoleh 100% dengan SVM pada satu skenario *preprocessing*, tetapi artikel tersebut juga mengingatkan risiko *overfitting* dan merekomendasikan *Naïve Bayes* yang mencapai 93% karena lebih stabil. Karena itu, pemilihan algoritma lebih tepat dipahami sebagai keputusan yang berkaitan dengan karakteristik data, *preprocessing*, target prediksi, dan skenario evaluasi.

3. RQ2: Variasi Data Historis dan Pemodelan

Data historis pada 12 penelitian dapat dikelompokkan menjadi data akademik, kehadiran, aktivitas pembelajaran atau log LMS, serta data sosial-demografis. Data akademik muncul dalam bentuk nilai ujian, tugas, nilai semester, atau capaian sebelumnya. Data aktivitas mencakup frekuensi *login*, interaksi LMS, durasi belajar, pengumpulan tugas, dan aktivitas pembelajaran lain. Beberapa penelitian menggabungkan beberapa kelompok data untuk memperkaya masukan model.

Tabel 5. Sintesis jenis data historis pada 12 jurnal utama

Kelompok data	Contoh pada studi utama	Pola penggunaan
Akademik	Nilai tugas/ujian, nilai semester, G1/G2, nilai rapor, capaian KKM	Paling banyak digunakan sebagai indikator capaian sebelumnya dan target prediksi.
Kehadiran	Absensi, frekuensi kehadiran	Digabung dengan nilai atau

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

		aktivitas untuk memperkuat representasi perilaku belajar.
Aktivitas/LMS	*Login*, kunjungan materi, durasi belajar, interaksi, pengumpulan tugas	Menangkap perilaku belajar yang berubah selama proses pembelajaran.
Sosial-demografis	Jenis kelamin, latar belakang, pekerjaan orang tua, ekonomi, lingkungan, dukungan keluarga	Digunakan sebagai fitur tambahan untuk merepresentasikan kondisi siswa.
Kombinasi	Akademik + kehadiran + aktivitas + sosial-demografis	Memberikan masukan lebih beragam, tetapi meningkatkan kompleksitas *preprocessing* dan kebutuhan validasi.

Penelitian (Ardianti et al., 2023) menunjukkan pemanfaatan log LMS untuk merepresentasikan perilaku dan interaksi siswa, sedangkan (Azizah et al., 2022) menggunakan fitur aktivitas LMS seperti kunjungan sumber belajar, absensi, *raised hands*, pengumuman, dan diskusi. (Fauzan et al., 2026) menggabungkan frekuensi *login*, durasi belajar, nilai kuis, tugas, ketepatan pengumpulan, kehadiran, dan keaktifan sehingga model tidak hanya bergantung pada nilai akademik. Pada sisi lain, (Naibaho & Zahra, 2023; Rahim & Mustakim, 2025) memperlihatkan bahwa data pribadi, pekerjaan orang tua, lingkungan, ekonomi, motivasi, dan dukungan keluarga dapat digunakan sebagai fitur tambahan. Temuan tersebut menunjukkan bahwa variasi data historis terutama memperluas representasi kondisi siswa, tetapi tidak cukup untuk menyatakan adanya hubungan sebab-akibat terhadap hasil belajar.

Pengaruh *preprocessing* juga terlihat jelas pada pemodelan. (Gori et al., 2024) melaporkan peningkatan performa *Naïve Bayes* setelah *Correlation-Based Feature Selection*, sedangkan (Rahim & Mustakim, 2025) menunjukkan perubahan performa yang besar antar-skenario normalisasi dan seleksi fitur. (Azizah et al., 2022) juga memperoleh parameter dan fitur terbaik setelah seleksi fitur *Information Gain*. Hal ini menunjukkan bahwa jenis data dan cara data diproses merupakan bagian yang tidak dapat dipisahkan dari interpretasi performa algoritma.

4. RQ3: Performa dan Keandalan Algoritma

Performa model pada 12 penelitian menunjukkan rentang yang luas. *Accuracy* merupakan metrik yang paling sering dilaporkan, sedangkan *precision*, *recall*, dan *F1-score* tersedia pada sebagian studi. Sebagian penelitian menggunakan pembagian data latih-uji, sedangkan penelitian lain menggunakan *k-fold cross-validation* atau validasi silang pada proses pemodelan. Perbedaan ini perlu diperhatikan karena angka *accuracy* dari *dataset*, target, dan skenario evaluasi yang berbeda tidak dapat dianggap sebagai perbandingan langsung.

Tabel 6. Performa model pada 12 jurnal utama

No	Penelitian	Model utama/terbaik	Performa yang dilaporkan
1	Ardianti et al. (2023)	*Bagging*	*Accuracy* 81,25%; *Precision* 0,810; *Recall* 0,812; *F1* 0,809
2	Kusworo et al. (2024)	*Random Forest*	*Accuracy* 93,33%; *Precision* 93,33%; *Recall* 100%; *F1* 96,55%
3	Maulana et al. (2025)	*Decision Tree*	*Accuracy* 97,80%; *Precision* 98,38% dan 94,55% per kategori
4	Fitriani et al. (2021)	C4.5	*Accuracy* 83,33%; *precision* dan *recall* dilaporkan per kelas
5	Rahim & Mustakim (2025)	SVM / NB	S3: SVM 100%; NB 93% direkomendasikan karena lebih stabil
6	Naibaho & Zahra (2023)	*Random Forest*	*Accuracy* 99,5%
7	Gori et al. (2024)	*Naïve Bayes* + CFS	*Accuracy* 91,22%;

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

			Precision 92,24%; *Recall* 91,22%; *F1* 91,48%
8	Hasanah & Fatmawati (2022)	C4.5	*Accuracy* 90,00%; NB 86,13%; AUC C4.5 0,596
9	Azizah et al. (2022)	NWKNN	*Accuracy* 0,604; *Precision* 0,719; *Recall* 0,676; *F1* 0,661
10	Pambudi et al. (2024)	*Logistic Regression* / kombinasi	Abstrak: kombinasi 90,3%; hasil: *Logistic Regression* 0,98 *accuracy*, *F1* 0,98, *precision* 0,981, *recall* 0,98
11	Anwar et al. (2022)	ANN-MLP	*Accuracy* 99,35%
12	Fauzan et al. (2026)	*Decision Tree*	*Accuracy* 88,52%; *Random Forest* 83,61%; NB 62,30%

Nilai tertinggi tidak dapat langsung digunakan untuk menetapkan model terbaik secara universal. (Naibaho & Zahra, 2023) memperoleh 99,5% dengan *Random Forest*, (Anwar et al., 2022) memperoleh 99,35% dengan ANN-MLP, dan (Rahim & Mustakim, 2025) memperoleh 100% dengan SVM pada satu skenario *preprocessing*. Namun, penelitian Rahim dan Mustakim sendiri mengingatkan adanya risiko *overfitting* pada *accuracy* 100% dan merekomendasikan *Naïve Bayes* yang mencapai 93% karena lebih stabil. Sebaliknya, (Azizah et al., 2022) memperoleh *accuracy* 60,4% pada NWKNN meskipun menggunakan seleksi fitur, menunjukkan bahwa seleksi fitur tidak otomatis menghasilkan *accuracy* tinggi.

Perbedaan metode evaluasi juga membatasi perbandingan. (Naibaho & Zahra, 2023) menggunakan GridSearchCV dan *5-fold cross-validation*, (Azizah et al., 2022) menggunakan *K-Fold Cross Validation*, (Maulana et al., 2025) menggunakan pembagian 70:30, sedangkan (Kusworo et al., 2024) menggunakan pembagian 70:30 dengan *stratified sampling*. Karena itu, keandalan model dalam SLR ini lebih tepat dilihat melalui konsistensi metrik, desain pengujian, dan keterbatasan *dataset* daripada hanya melalui satu angka *accuracy*. Temuan (Pambudi et al., 2024) juga perlu diberi catatan khusus karena terdapat ketidakkonsistenan antara angka pada abstrak dan bagian hasil, sehingga angka tersebut tidak digunakan sebagai dasar untuk menyatakan keunggulan model secara kuat.

5. RQ4: Kontribusi terhadap *Early Warning System*

Kontribusi terhadap *early warning system* (EWS) tidak sama pada seluruh penelitian. Sebagian artikel benar-benar mengembangkan sistem, sebagian menghasilkan prediksi yang dapat menjadi dasar identifikasi siswa berisiko, dan sebagian hanya melakukan prediksi tanpa menjelaskan mekanisme peringatan atau intervensi. Pengelompokan ini penting agar setiap model prediksi tidak otomatis disebut sebagai EWS.

Tabel 7. Pemetaan kontribusi 12 jurnal terhadap *Early Warning System*

Kategori	Studi	Sintesis
EWS diterapkan langsung	Fauzan et al. (2026)	Model diimplementasikan sebagai sistem berbasis web untuk guru dan digunakan untuk mendeteksi tingkat pemahaman siswa sebelum ujian akhir.
Mendukung EWS	Ardianti et al. (2023); Kusworo et al. (2024); Fitriani et al. (2021); Rahim & Mustakim (2025); Naibaho & Zahra (2023); Gori et al. (2024); Hasanah & Fatmawati (2022); Pambudi et al. (2024)	Prediksi digunakan atau diarahkan untuk identifikasi risiko, pencegahan kegagalan/putus sekolah, kebijakan, atau intervensi pembelajaran, tetapi tidak seluruhnya menerapkan sistem peringatan operasional.
Prediksi saja	Maulana et al. (2025); Azizah et al.	Model menghasilkan prediksi atau

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

	(2022); Anwar et al. (2022)	klasifikasi yang berguna, tetapi artikel tidak melaporkan EWS operasional atau evaluasi intervensi.

Temuan paling jelas terdapat pada (Fauzan et al., 2026), yang mengimplementasikan hasil prediksi sebagai EWS berbasis web bagi guru. Penelitian lain lebih banyak berhenti pada tahap prediksi dan menyatakan bahwa hasilnya dapat digunakan untuk identifikasi siswa berisiko atau pemberian intervensi. (Alalawi et al., 2023), yang digunakan sebagai literatur pendukung, juga menemukan bahwa hanya sedikit penelitian yang benar-benar menindaklanjuti hasil prediksi dengan tindakan remedial dan mengevaluasi dampaknya. Artinya, celah antara model prediksi dan tindakan intervensi masih menjadi bagian penting yang belum kuat dalam penelitian primer yang dianalisis.

6. Pemetaan 12 Jurnal terhadap RQ

Tabel 8. Pemetaan penggunaan 12 jurnal utama terhadap RQ

No	Artikel	RQ1	RQ2	RQ3	RQ4
1	Ardianti et al. (2023)	✓	✓	✓	✓
2	Kusworo et al. (2024)	✓	✓	✓	✓
3	Maulana et al. (2025)	✓	✓	✓	-
4	Fitriani et al. (2021)	✓	✓	✓	✓
5	Rahim & Mustakim (2025)	✓	✓	✓	✓
6	Naibaho & Zahra (2023)	✓	✓	✓	✓
7	Gori et al. (2024)	✓	✓	✓	✓
8	Hasanah & Fatmawati (2022)	✓	✓	✓	✓
9	Azizah et al. (2022)	✓	✓	✓	-
10	Pambudi et al. (2024)	✓	✓	✓	✓
11	Anwar et al. (2022)	✓	✓	✓	-
12	Fauzan et al. (2026)	✓	✓	✓	✓

Pemetaan menunjukkan bahwa seluruh 12 jurnal memberikan kontribusi terhadap RQ1, RQ2, dan RQ3 karena setiap studi menyediakan informasi tentang algoritma, data masukan, dan hasil evaluasi. Kontribusi terhadap RQ4 lebih selektif karena tidak semua penelitian membahas identifikasi risiko atau tindakan setelah prediksi. (Anwar et al., 2022; Azizah et al., 2022; Maulana et al., 2025) tetap relevan untuk RQ1-RQ3, tetapi tidak digunakan untuk menyatakan adanya EWS yang tidak dilaporkan dalam artikel. Pembatasan ini menjaga agar sintesis tidak melebih-lebihkan fungsi model.

7. Sintesis Keseluruhan, Keterbatasan, dan Kesenjangan

Secara keseluruhan, 12 penelitian menunjukkan pola bahwa data historis berfungsi sebagai dasar utama pembentukan model prediksi, sedangkan algoritma digunakan untuk mengenali pola dari data tersebut. *Naïve Bayes* menjadi algoritma yang paling sering digunakan, tetapi *Random Forest*, *Decision Tree*, C4.5, dan jaringan saraf juga memiliki peran penting. Data akademik dan kehadiran tetap menjadi sumber utama, sementara log LMS dan data sosial-demografis

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

memperluas informasi mengenai perilaku serta kondisi siswa. Performa yang tinggi cenderung muncul ketika fitur relevan dan *preprocessing* disesuaikan dengan karakteristik *dataset*, tetapi hasil antarpenelitian tidak dapat dibandingkan hanya berdasarkan angka *accuracy*.

Keterbatasan utama yang terlihat adalah perbedaan ukuran *dataset*, sumber data, karakteristik target, pembagian data, dan metrik evaluasi. Sebagian penelitian juga berfokus pada satu sekolah atau institusi sehingga generalisasi ke konteks lain masih terbatas. Kasus (Pambudi et al., 2024) memperlihatkan pentingnya konsistensi pelaporan karena angka pada abstrak dan bagian hasil tidak sama. Selain itu, hanya (Fauzan et al., 2026) yang secara jelas menerapkan hasil prediksi sebagai EWS berbasis web, sedangkan sebagian besar penelitian berhenti pada prediksi atau menyatakan kemungkinan intervensi.

Kesenjangan penelitian yang terlihat dari sintesis ini adalah belum kuatnya hubungan antara model prediksi, validasi lintas konteks, dan tindakan intervensi yang terukur. Penelitian lanjutan perlu memperhatikan evaluasi yang lebih konsisten, validasi pada data dari lebih dari satu konteks pendidikan, pengujian risiko *overfitting*, serta evaluasi terhadap dampak intervensi setelah siswa berisiko berhasil diidentifikasi. Dengan demikian, pengembangan sistem prediksi tidak hanya berorientasi pada peningkatan nilai metrik, tetapi juga pada keterterapan hasil prediksi dalam proses pendampingan siswa.

KESIMPULAN

Analisis terhadap 12 jurnal utama menunjukkan bahwa *Naïve Bayes* merupakan algoritma yang paling sering digunakan, diikuti *Random Forest* dan *Decision Tree*, tetapi tidak terdapat algoritma yang dapat dinyatakan terbaik secara universal karena performa dipengaruhi oleh karakteristik data dan *preprocessing*. Data historis yang digunakan mencakup data akademik, kehadiran, aktivitas atau log LMS, serta data sosial-demografis, dengan kombinasi beberapa jenis data pada sejumlah penelitian. Performa model bervariasi dan tidak dapat dibandingkan secara langsung tanpa mempertimbangkan *dataset* dan metode evaluasi, sedangkan kontribusi terhadap *early warning system* masih didominasi fungsi pendukung dan hanya satu studi yang melaporkan penerapan EWS berbasis web secara langsung. Hasil sintesis menunjukkan bahwa pengembangan penelitian selanjutnya perlu memperkuat konsistensi evaluasi, validasi lintas konteks, pengendalian *overfitting*, dan pengukuran dampak intervensi setelah prediksi siswa berisiko.

REFERENSI

Alalawi, K., Athauda, R., & Chiong, R. (2023). Contextualizing the current state of research on the use of machine learning for student performance prediction: A systematic literature review. *Engineering Reports*, *5*(12), 1–25. https://doi.org/10.1002/eng2.12699

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090552

Anwar, S., Kurnia, D. A., Faqih, A., & Sari, S. R. (2022). Prediksi Hasil Belajar Hybrid Menggunakan Artificial Neural Network Dengan Multilayer Perceptron. *JURIKOM (Jurnal Riset Komputer)*, *9*(5), 1591. https://doi.org/10.30865/jurikom.v9i5.5024

Ardianti, M., Nurhayati, O. D., & Warsito, B. (2023). Model Prediksi Kinerja Siswa Berdasarkan Data Log LMS Menggunakan Ensemble Machine Learning. *JST (Jurnal Sains Dan Teknologi)*, *12*(3), 562–571. https://doi.org/10.23887/jstundiksha.v12i3.59816

Azizah, R. A., Bachtiar, F. A., & Adinugroho, S. (2022). Klasifikasi Kinerja Akademik Siswa Menggunakan Neighbor Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *9*(3), 605–614. https://doi.org/10.25126/jtiik.2022935751

Fauzan, M., Sukma, D. P., Eriwandi, E., & Odang, D. K. (2026). Prediksi Tingkat Pemahaman Siswa SMP Pada Pembelajaran Digital Menggunakan Algoritma Machine Learning. *Jurnal Teknologi*

E-ISSN : 2961 - 9386
P-ISSN : 2963 - 1742

Informasi (JUTECH), *7*(1), 23–30. https://doi.org/10.32546/jutech.v7i1.3543

Fitriani, Y., Defit, S., & Nurcahyo, G. W. (2021). Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4.5. *Jurnal Sistim Informasi Dan Teknologi*, *3*, 120–127. https://doi.org/10.37034/jsisfotek.v3i3.54

Gori, T., Sunyoto, A., & Al Fatta, H. (2024). Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *11*(1), 215–224. https://doi.org/10.25126/jtiik.20241118074

Hasanah, U., & Fatmawati. (2022). Prediksi Penurunan Hasil Belajar Daring Siswa Dengan Algoritma C4.5 Dan Naïve Bayes. *Inti Nusa Mandiri*, *17*(1), 8–15. https://doi.org/10.33480/inti.v17i1.3139

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical Report, Ver. 2.3 EBSE Technical Report. EBSE*, (January 2007), 1–57.

Kusworo, S., Santoso, N. A., & Kurniawan, R. D. (2024). Prediksi Nilai Akhir Semester Siswa Menggunakan Algoritma Random Forest. *Jurnal Sains Dan Ilmu Terapan*, *7*(2), 246–256. https://doi.org/10.59061/jsit.v7i2.918

Maulana, S., Premana, A., & Irawan, B. (2025). Prediksi Prestasi Akademik Siswa Terbaik Menggunakan Algoritma Decision Tree Berbasis Data Historis. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(5), 7890–7897. https://doi.org/10.36040/jati.v9i5.14936

Naibaho, A. F. A., & Zahra, A. (2023). Prediksi Kelulusan Siswa Sekolah Menengah Pertama Menggunakan Machine Learning. *Jurnal Informatika Dan Teknik Elektro Terapan*, *11*(3), 267–281. https://doi.org/10.23960/jitet.v11i3.3056

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Bmj*, *372*. https://doi.org/10.1136/bmj.n71

Pambudi, R. E., Purnomo, H., & Irawan, R. (2024). Pemanfaat Data Mining Untuk Prediksi Prestasi Akademik Siswa Berdasarkan Pola Kehadiran, Aktivitas Belajar Menggunakan Naive Bayes Logistic Regression. *Jurnal Teknologi Informasi Mura*, *16*(2), 132.

Rahim, A., & Mustakim, M. (2025). Peningkatan Akurasi Model Untuk Prediksi KKM Siswa Sekolah Dasar Menggunakan Supervised Machine Learning dengan Integrasi Faktor Internal dan Eksternal. *Euler : Jurnal Ilmiah Matematika, Sains Dan Teknologi*, *13*(3), 372–380. https://doi.org/10.37905/euler.v13i3.34577

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), 22–32. https://doi.org/10.20885/snati.v1i1.4

Sandra, L., Lumbangaol, F., & Matsuo, T. (2021). Machine Learning Algorithm to Predict Student's Performance: A Systematic Literature Review. *TEM Journal*, *10*(4), 1919–1927. https://doi.org/10.18421/TEM104-56

Sekeroglu, B., Abiyev, R., Ilhan, A., Arslan, M., & Idoko, J. B. (2021). Systematic literature review on machine learning and student performance prediction: Critical gaps and possible remedies. *Applied Sciences (Switzerland)*, *11*(22), 2–23. https://doi.org/10.3390/app112210907

Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022). In *Education and Information Technologies* (Vol. 28, Number 7). Springer US. https://doi.org/10.1007/s10639-022-11536-0

Turkmenbayev, A., Abdykerimova, E., Nurgozhayev, S., Karabassova, G., & Baigozhanova, D. (2025). The application of machine learning in predicting student performance in university engineering programs: a rapid review. *Frontiers in Education*, *10*(September), 1–14. https://doi.org/10.3389/feduc.2025.1562586

Zhang, Y., Yun, Y., An, R., Cui, J., Dai, H., & Shang, X. (2021). Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis. *Frontiers in Psychology*, *12*(December), 1–19. https://doi.org/10.3389/fpsyg.2021.698490