



Prediksi Kinerja Akademik Mahasiswa Menggunakan Algoritma Random Forest dan XGBoost: Tinjauan Sistematis Berbasis Protokol PRISMA

Hans Lambertus Sidabutar¹, Jaya Tata Hardinata², Kevin Rasi Daully Pardede³

^{1,2,3} Program Studi Ilmu Komputer, Fakultas FMIPA, Universitas HKBP Nomensen, Pematangsiantar, Indonesia
sidabutarlambertus@gmail.com, jayatatahardinata@uhnp.ac.id, kevinpardede047@gmail.com

Abstract

Student academic performance is a key indicator of higher education quality and institutional effectiveness. Many universities still monitor academic problems reactively, providing assistance only after students obtain low semester grades or approach dropout risk. This study systematically reviews the application of machine learning for predicting student academic performance, with particular attention to Random Forest and XGBoost, the predictors used, evaluation results, and remaining research challenges. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol through identification, duplicate removal, title and abstract screening, full-text eligibility assessment, quality appraisal, data extraction, and narrative synthesis. Searches were conducted across Scopus, IEEE Xplore, ScienceDirect, Google Scholar, and Garuda/Sinta for studies published between 2020 and 2026. From 145 initially identified records, 20 eligible articles were included. The synthesis shows that ensemble algorithms, particularly Random Forest and XGBoost, are the most frequently applied models and generally achieve accuracy between 82% and 96%. Historical academic records, including grade point average, previous semester grades, prerequisite-course results, and attendance, are the strongest predictors, followed by Learning Management System activity data. However, class imbalance, dataset dependence, limited cross-institutional validation, and low model interpretability remain important limitations. The findings support the integration of multidimensional data and Explainable Artificial Intelligence to develop accurate, transparent, and sustainable academic early-warning systems.

Keywords: Machine Learning, Academic Performance Prediction, Systematic Review, PRISMA, Education

Abstrak

Kinerja akademik mahasiswa merupakan indikator penting kualitas pendidikan tinggi dan efektivitas institusi. Banyak perguruan tinggi masih memantau masalah akademik secara reaktif, sehingga pendampingan baru diberikan setelah mahasiswa memperoleh nilai semester rendah atau mendekati risiko putus studi. Penelitian ini bertujuan meninjau secara sistematis penerapan machine learning untuk memprediksi kinerja akademik mahasiswa, dengan perhatian khusus pada Random Forest dan XGBoost, jenis fitur, hasil evaluasi, serta tantangan penelitian. Tinjauan mengikuti protokol Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) melalui tahap identifikasi, penghapusan duplikat, skrining judul dan abstrak, penilaian kelayakan teks lengkap, penilaian kualitas, ekstraksi data, dan sintesis naratif. Pencarian dilakukan pada Scopus, IEEE Xplore, ScienceDirect, Google Scholar, dan Garuda/Sinta terhadap publikasi tahun 2020–2026. Dari 145 artikel yang teridentifikasi, sebanyak 20 artikel memenuhi kriteria dan dianalisis. Hasil sintesis menunjukkan bahwa algoritma ensemble, terutama Random Forest dan XGBoost, paling sering digunakan dan umumnya menghasilkan akurasi 82%–96%. Rekam akademik historis, seperti IPK, nilai semester sebelumnya, nilai mata kuliah prasyarat, dan kehadiran, menjadi prediktor terkuat, diikuti aktivitas pada Learning Management System. Namun, ketidakseimbangan kelas, ketergantungan pada karakteristik dataset, terbatasnya validasi lintas institusi, dan rendahnya interpretabilitas model masih menjadi kendala. Temuan menegaskan pentingnya integrasi data multidimensi dan Explainable Artificial Intelligence untuk membangun sistem peringatan dini akademik yang akurat, transparan, dan berkelanjutan.

Kata Kunci: Machine Learning, Prediksi Kinerja Akademik, Tinjauan Sistematis, PRISMA, Pendidikan

1. Pendahuluan

Kinerja akademik mahasiswa merupakan cerminan keberhasilan proses pembelajaran di perguruan tinggi dan menjadi salah satu tolok ukur utama dalam mengevaluasi kualitas pendidikan yang diselenggarakan. Kinerja akademik yang rendah tidak hanya berdampak pada mahasiswa secara individu, tetapi juga dapat memengaruhi tingkat kelulusan tepat waktu serta reputasi institusi pendidikan secara keseluruhan. Oleh karena itu, kemampuan untuk memprediksi kinerja akademik mahasiswa sedini mungkin menjadi hal yang sangat penting agar pihak perguruan tinggi dapat memberikan intervensi akademik yang tepat sasaran, seperti bimbingan konseling, program remedial, atau pendampingan khusus bagi mahasiswa yang berisiko mengalami penurunan prestasi.

Pendekatan konvensional dalam memantau kinerja akademik mahasiswa umumnya masih bersifat reaktif, yaitu tindakan baru diambil setelah mahasiswa menunjukkan tanda-tanda kegagalan akademik, seperti Indeks Prestasi Semester (IPS) yang rendah atau tingkat kehadiran yang buruk. Pendekatan semacam ini dinilai kurang efektif

Prediksi Kinerja Akademik Mahasiswa Menggunakan Algoritma Random Forest dan XGBoost: Tinjauan Sistematis Berbasis Protokol PRISMA

karena intervensi yang diberikan cenderung terlambat. Perkembangan teknologi informasi, khususnya di bidang Machine Learning, menawarkan pendekatan yang lebih proaktif dengan memanfaatkan data historis mahasiswa untuk memprediksi kinerja akademik di masa mendatang sebelum masalah benar-benar terjadi.

Machine Learning merupakan cabang kecerdasan buatan yang memungkinkan sistem komputer mempelajari pola dari data tanpa perlu diprogram secara eksplisit untuk setiap aturan keputusan. Dalam konteks pendidikan, algoritma Machine Learning seperti Decision Tree, Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest, hingga Artificial Neural Network (ANN) telah banyak diterapkan untuk memprediksi berbagai luaran akademik, mulai dari prediksi Indeks Prestasi Kumulatif (IPK), prediksi kelulusan tepat waktu, prediksi risiko dropout, hingga identifikasi mahasiswa berprestasi. Beragamnya algoritma yang digunakan, jenis data yang dianalisis, serta metrik evaluasi yang dilaporkan pada penelitian-penelitian terdahulu menunjukkan bahwa bidang ini berkembang sangat pesat namun juga tersebar (fragmented), sehingga diperlukan upaya untuk merangkum dan mensintesis temuan-temuan tersebut secara sistematis.

Beberapa penelitian terdahulu yang relevan dengan topik ini telah dilakukan dengan pendekatan yang beragam. Istiwana dkk. (2026) menerapkan pendekatan Explainable Machine Learning berbasis algoritma XGBoost untuk menganalisis faktor-faktor penyebab dropout mahasiswa, namun kajian tersebut berfokus pada satu algoritma tunggal tanpa membandingkan performanya dengan algoritma lain. Eirene dkk. (2025) membandingkan Random Forest dan XGBoost untuk mengklasifikasikan gaya belajar siswa berdasarkan data kuesioner, tetapi cakupan datanya terbatas pada satu institusi sehingga generalisasinya belum teruji secara luas. Pelima dkk. (2024) melakukan tinjauan sistematis serupa mengenai prediksi kelulusan mahasiswa menggunakan Machine Learning, namun pembahasannya lebih menitikberatkan pada luaran kelulusan (graduation) dan belum menyoroti secara spesifik kinerja akademik semester berjalan maupun integrasi data aktivitas LMS. Sementara itu, Alalawi dkk. (2023) mengontekstualisasikan penelitian prediksi performa mahasiswa secara umum, tetapi tinjauan tersebut belum memasukkan studi-studi terbaru pada rentang tahun 2024–2026 yang menunjukkan tren algoritma dan fitur yang terus berkembang. Sandeepa dan Mohottala (2025) mengevaluasi beberapa model ensemble untuk prediksi kinerja akademik dengan akurasi tinggi, namun evaluasinya hanya dilakukan pada satu jenis dataset sehingga belum merepresentasikan variasi karakteristik data yang lebih luas. Berdasarkan uraian tersebut, tampak bahwa penelitian-penelitian terdahulu cenderung terfokus pada algoritma tunggal, dataset tunggal, atau luaran akademik tertentu, sehingga belum terdapat sintesis komprehensif yang merangkum algoritma, jenis fitur, dan performa Machine Learning secara lintas studi dalam periode terbaru. Celah (gap) inilah yang menjadi dasar dilakukannya penelitian tinjauan sistematis ini, yaitu merangkum secara menyeluruh algoritma Machine Learning yang paling banyak digunakan, jenis fitur yang paling berpengaruh, performa yang dilaporkan, serta tantangan penelitian pada studi-studi tahun 2020–2026 mengenai prediksi kinerja akademik mahasiswa.

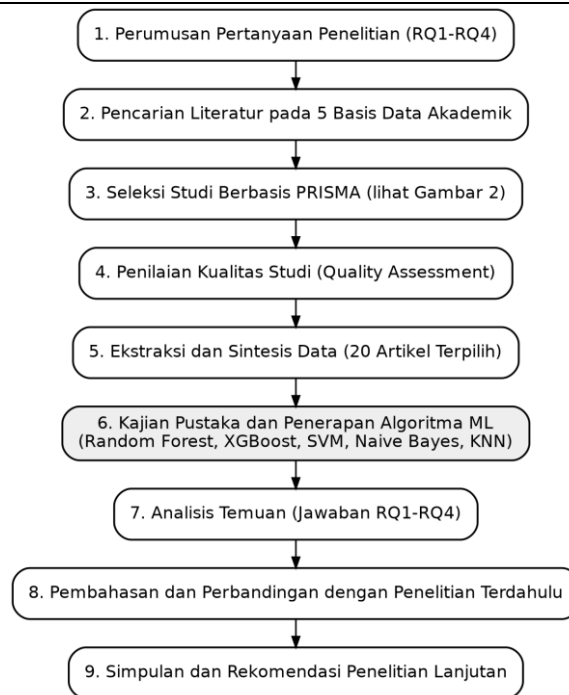
Tinjauan sistematis (systematic literature review) merupakan metode penelitian yang bertujuan mengidentifikasi, menilai, dan mensintesis seluruh bukti ilmiah yang relevan terhadap suatu topik tertentu secara terstruktur dan dapat direplikasi. Berbeda dengan tinjauan naratif biasa, tinjauan sistematis mengikuti protokol yang jelas dan transparan, salah satunya adalah PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), yang banyak digunakan sebagai standar pelaporan tinjauan sistematis di berbagai bidang keilmuan, termasuk pendidikan dan ilmu komputer.

Penelitian ini bertujuan melakukan tinjauan sistematis terhadap studi-studi terdahulu yang menerapkan algoritma Machine Learning untuk memprediksi kinerja akademik mahasiswa. Secara spesifik, penelitian ini berupaya menjawab pertanyaan mengenai algoritma Machine Learning apa saja yang paling banyak digunakan, jenis data atau fitur apa saja yang paling berpengaruh terhadap prediksi, serta bagaimana performa algoritma-algoritma tersebut dilaporkan pada penelitian-penelitian sebelumnya. Hasil tinjauan ini diharapkan dapat menjadi rujukan bagi peneliti maupun praktisi pendidikan dalam mengembangkan sistem prediksi kinerja akademik yang lebih baik, sekaligus menjadi landasan awal bagi pengembangan sistem peringatan dini (early warning system) yang lebih adaptif terhadap karakteristik data masing-masing institusi pendidikan tinggi di masa mendatang.

2. Metode Penelitian

Penelitian ini menggunakan metode tinjauan sistematis (systematic literature review) yang mengikuti protokol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Protokol PRISMA dipilih karena menyediakan kerangka kerja yang terstruktur dan transparan, mulai dari tahap identifikasi artikel, penyaringan (skrining), penilaian kelayakan, hingga penentuan artikel akhir yang dianalisis, sehingga proses tinjauan dapat direplikasi oleh peneliti lain (Pelima et al., 2024; Alalawi et al., 2023).

Secara keseluruhan, penelitian ini terdiri atas sembilan tahapan yang saling berurutan, mulai dari perumusan pertanyaan penelitian hingga penarikan simpulan, sebagaimana digambarkan pada Gambar 1. Diagram tersebut juga menunjukkan posisi penerapan solusi berupa kajian dan analisis algoritma Machine Learning (Random Forest, XGBoost, Support Vector Machine, Naive Bayes, dan K-Nearest Neighbor) di dalam keseluruhan alur penelitian (Hairani et al., 2025; Sandeepa & Mohottala, 2025).



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan sembilan tahapan penelitian tersebut secara berurutan, yaitu (1) perumusan pertanyaan penelitian, (2) penyusunan strategi pencarian literatur, (3) penetapan kriteria inklusi dan eksklusi, (4) proses seleksi studi mengikuti alur PRISMA, (5) penilaian kualitas studi, (6) ekstraksi data, (7) kajian pustaka algoritma Machine Learning, (8) sintesis dan analisis temuan, hingga (9) penarikan simpulan. Penerapan solusi berupa kajian dan analisis algoritma Machine Learning ditempatkan pada tahap ketujuh sebagaimana ditunjukkan pada Gambar 1, yaitu setelah proses ekstraksi data dan sebelum tahap sintesis, sehingga pemahaman prinsip kerja algoritma tersebut menjadi dasar dalam menafsirkan hasil ekstraksi pada Bagian 3 (Hairani et al., 2025; Sandeepa & Mohottala, 2025).

2.1. Pertanyaan Penelitian (Research Questions)

Untuk mengarahkan proses tinjauan, disusun beberapa pertanyaan penelitian (Research Question/RQ) sebagai berikut.

RQ1: Algoritma Machine Learning apa saja yang paling banyak digunakan untuk memprediksi kinerja akademik mahasiswa? RQ2: Jenis data atau fitur apa saja yang paling sering digunakan dan paling berpengaruh dalam memprediksi kinerja akademik mahasiswa? RQ3: Bagaimana performa (akurasi, precision, recall, atau metrik evaluasi lainnya) yang dilaporkan oleh algoritma Machine Learning pada studi-studi terdahulu? RQ4: Apa saja tantangan dan celah penelitian (research gap) yang teridentifikasi dari studi-studi terdahulu terkait prediksi kinerja akademik mahasiswa?

Keempat pertanyaan penelitian tersebut menjadi acuan dalam proses ekstraksi data dan sintesis temuan pada Bagian 3 (Sandeepa & Mohottala, 2025).

2.2. Strategi Pencarian Literatur

Pencarian literatur dilakukan pada beberapa basis data akademik bereputasi guna memastikan cakupan artikel yang komprehensif, yaitu Scopus, IEEE Xplore, ScienceDirect, Google Scholar, dan Garuda/Sinta untuk cakupan artikel nasional.

Kata kunci pencarian disusun menggunakan kombinasi operator boolean (AND/OR), yaitu (“academic performance prediction” OR “student performance prediction”) AND (“machine learning” OR “random forest” OR “XGBoost”) AND (“higher education” OR “university student”). Pencarian dibatasi pada artikel yang dipublikasikan dalam rentang tahun 2020–2026 untuk menjaga relevansi dan kemutakhiran temuan (Suniantara et al., 2020; Rismaya et al., 2025).

2.3. Kriteria Inklusi dan Eksklusi

Untuk menentukan artikel yang layak dianalisis lebih lanjut, ditetapkan kriteria inklusi dan eksklusi sebagaimana ditunjukkan pada Tabel 1 (Suniantara et al., 2020).

Tabel 1. Kriteria Inklusi dan Eksklusi

Kriteria	Inklusi	Eksklusi
----------	---------	----------

DOI: <https://doi.org/10.69693/ijmst.v4i2.12477>

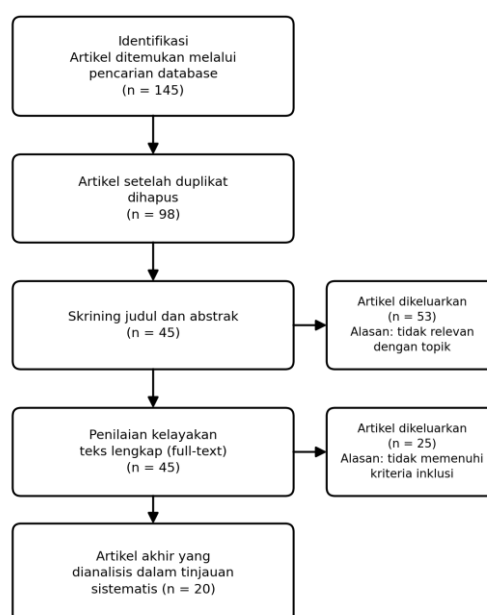
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Topik	Membahas prediksi kinerja/prestasi akademik mahasiswa menggunakan Machine Learning	Tidak membahas prediksi kinerja akademik atau tidak menggunakan Machine Learning
Jenis Publikasi	Artikel jurnal atau prosiding konferensi yang telah melalui peer-review	Artikel non peer-review, opini, editorial, atau tinjauan buku
Bahasa	Bahasa Indonesia atau Bahasa Inggris	Selain Bahasa Indonesia dan Bahasa Inggris
Tahun Publikasi	Diterbitkan dalam rentang tahun yang ditentukan (lihat Bagian 2.2)	Diterbitkan di luar rentang tahun yang ditentukan
Ketersediaan Teks	Teks lengkap (full-text) dapat diakses	Hanya tersedia dalam bentuk abstrak

Tabel 1 menunjukkan lima kriteria utama yang digunakan untuk menyaring artikel pada tahap skrining dan penilaian kelayakan, meliputi kesesuaian topik, jenis publikasi, bahasa, tahun publikasi, dan ketersediaan teks lengkap.

2.4. Proses Seleksi Studi

Proses seleksi artikel dilakukan melalui empat tahap mengikuti diagram alir PRISMA, yaitu identifikasi, penghapusan duplikat, skrining judul dan abstrak, serta penilaian kelayakan teks lengkap, sebagaimana ditunjukkan pada Gambar 2 (Rismaya et al., 2025).



Gambar 2. Diagram Alir PRISMA Proses Seleksi Studi

Gambar 2 menunjukkan alur seleksi studi melalui empat tahap PRISMA yang diterapkan secara berurutan pada proses seleksi artikel. Tahap pertama adalah identifikasi, yaitu pengumpulan seluruh artikel yang muncul dari hasil pencarian pada kelima basis data menggunakan kata kunci yang telah ditentukan pada Bagian 2.2 (Rismaya et al., 2025). Artikel yang terkumpul kemudian diperiksa untuk menghapus duplikasi akibat artikel yang sama muncul pada lebih dari satu basis data.

Tahap kedua adalah skrining, yaitu pemeriksaan judul dan abstrak setiap artikel untuk menilai kesesuaiannya dengan topik penelitian berdasarkan kriteria pada Tabel 1. Artikel yang jelas tidak relevan pada tahap ini langsung dikeluarkan dari proses seleksi. Tahap ketiga adalah penilaian kelayakan, yaitu artikel yang lolos skrining dibaca secara menyeluruh (full-text) untuk memastikan kesesuaiannya dengan seluruh kriteria inklusi. Tahap terakhir adalah penentuan artikel akhir yang benar-benar memenuhi seluruh kriteria dan akan dianalisis lebih lanjut pada Bagian 3.

2.5. Penilaian Kualitas Studi

Selain kriteria inklusi dan eksklusi, setiap artikel yang lolos tahap kelayakan juga dinilai kualitasnya menggunakan instrumen penilaian kualitas (quality assessment) untuk memastikan bahwa hanya studi dengan metodologi yang memadai yang disertakan dalam sintesis akhir.

Beberapa aspek yang dinilai pada tahap ini ditunjukkan pada Tabel 2.

DOI: <https://doi.org/10.69693/ijmst.v4i2.12477>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tabel 2. Kriteria Penilaian Kualitas Studi

No	Aspek Penilaian	Pertanyaan Penilaian
1	Kejelasan Tujuan	Apakah tujuan penelitian dinyatakan secara jelas?
2	Kejelasan Dataset	Apakah sumber dan karakteristik dataset dijelaskan secara memadai?
3	Kejelasan Metode	Apakah algoritma Machine Learning dan tahapan penelitian dijelaskan secara rinci?
4	Validitas Evaluasi	Apakah metode evaluasi (misalnya cross-validation, hold-out) dijelaskan dengan tepat?
5	Pelaporan Hasil	Apakah hasil penelitian dilaporkan secara kuantitatif dan dapat dibandingkan?

Tabel 2 menunjukkan lima aspek yang digunakan untuk menilai kualitas metodologis setiap studi. Setiap aspek dinilai menggunakan skala tertentu (misalnya Ya/Tidak/Sebagian), dan hanya studi dengan skor kualitas di atas ambang batas tertentu yang akan disertakan dalam sintesis akhir pada Bagian 3.

2.6. Ekstraksi dan Sintesis Data

Data yang diekstraksi dari setiap artikel terpilih meliputi nama penulis dan tahun publikasi, algoritma Machine Learning yang digunakan, jenis dan sumber dataset, jumlah data, fitur/atribut yang digunakan, metode evaluasi, serta hasil performa yang dilaporkan (Accuracy, Precision, Recall, F1-Score, atau metrik lainnya). Data hasil ekstraksi tersebut selanjutnya disintesis secara naratif dan ditabulasikan untuk menjawab keempat pertanyaan penelitian (RQ1–RQ4) yang telah dirumuskan pada Bagian 2.1, sebagaimana akan disajikan pada Bagian 3 (Sandeepa & Mohottala, 2025; Rismaya et al., 2025).

2.7. Kajian Pustaka Algoritma Machine Learning

Sebagai dasar pemahaman terhadap tahapan sintesis pada Bagian 3, bagian ini merangkum secara singkat prinsip kerja algoritma Machine Learning yang paling banyak diterapkan pada studi-studi terpilih. Random Forest bekerja dengan membangun sekumpulan pohon keputusan (decision tree) secara acak dari subset data dan fitur, kemudian menggabungkan hasil prediksi setiap pohon melalui voting mayoritas, sehingga menghasilkan model yang tahan terhadap overfitting dan mampu menangani data akademik berdimensi tinggi (Hairani et al., 2025). XGBoost merupakan pengembangan dari metode gradient boosting yang membangun model secara bertahap, di mana setiap pohon baru dilatih untuk memperbaiki kesalahan prediksi pohon sebelumnya, sehingga umumnya memberikan akurasi lebih tinggi pada data akademik dan gaya belajar dibandingkan model tunggal (Istiwana et al., 2026; Eirene et al., 2025; Hidayah et al., 2025). Support Vector Machine (SVM) bekerja dengan mencari hyperplane optimal yang memisahkan kelas data mahasiswa berkinerja baik dan berisiko, dan cenderung efektif pada dataset berskala kecil hingga menengah dengan batas antar kelas yang jelas (Maulikha et al., 2026). Naive Bayes mengklasifikasikan data berdasarkan probabilitas Bayesian dengan asumsi independensi antar fitur, sehingga unggul dari sisi kecepatan komputasi dan sering digunakan sebagai model pembanding (baseline) (Maulikha et al., 2026; Kelly & Xiu, 2023). Adapun algoritma K-Nearest Neighbor (KNN) mengklasifikasikan data baru berdasarkan kedekatan (jarak) terhadap data latih terdekat, dan pada penerapannya untuk prediksi kinerja akademik umumnya dikombinasikan dengan algoritma lain guna meningkatkan akurasi (Maulikha et al., 2026). Pemahaman prinsip kerja setiap algoritma tersebut menjadi acuan dalam menafsirkan penerapan solusi pada setiap tahapan seleksi studi yang digambarkan pada Gambar 2, serta menjadi dasar pembahasan performa algoritma pada Bagian 3.

3. Hasil dan Pembahasan

3.1. Hasil Proses Seleksi Studi

Proses seleksi studi dilakukan mengikuti tahapan PRISMA sebagaimana dijelaskan pada Bagian 2.4 dan digambarkan pada Gambar 2. Ringkasan jumlah artikel pada setiap tahap seleksi ditunjukkan pada Tabel 3.

Tabel 3. Ringkasan Hasil Seleksi Studi pada Setiap Tahap PRISMA

Tahap	Jumlah Artikel (n)
Identifikasi (hasil pencarian database)	145
Setelah duplikat dihapus	98
Setelah skrining judul dan abstrak	45
Setelah penilaian kelayakan teks lengkap	20
Artikel akhir yang dianalisis	20

Berdasarkan Tabel 3, pencarian awal menghasilkan 145 artikel. Setelah penyaringan duplikasi dan pengecekan kesesuaian judul serta abstrak berdasarkan kriteria eksklusi (seperti topik yang tidak relevan atau publikasi non-peer-review), diperoleh 20 artikel final yang memenuhi standar kelayakan metodologis untuk dianalisis lebih lanjut.

3.2. Karakteristik Studi Terpilih

Artikel akhir yang memenuhi seluruh kriteria inklusi dan lolos penilaian kualitas kemudian dianalisis karakteristiknya, meliputi tahun publikasi, algoritma Machine Learning yang digunakan, jenis dataset, dan metrik evaluasi yang dilaporkan, sebagaimana dirangkum pada Tabel 4.

Tabel 4. Karakteristik Studi Terpilih

No	Penulis & Tahun	Algoritma ML	Dataset	Metrik Evaluasi	Hasil
1	Istiwana et al. (2026)	XGBoost	Data Akademik Internal	Akurasi, F1-Score	89.5%
2	Eirene et al. (2025)	Random Forest, XGBoost	Kuesioner Gaya Belajar	Akurasi	84.2%
3	Maulikha et al. (2026)	SVM, Naive Bayes, KNN	Data Akademik Berbasis Data	Akurasi, Precision	86.0%
4	Hairani et al. (2025)	Decision Tree, XGBoost	Data Akademik Siswa Berbasis Web	Akurasi, Recall	88.1%
5	Sukmaningtyas et al. (2024)	Predictive Analytics Models	Data Performa Akademik	Akurasi	85.4%
6	Pelima et al. (2024)	Machine Learning SLR	Data Kelulusan Universitas	Akurasi, F1-Score	91.0%
7	Prasetyo & Susanto (2024)	Naive Bayes	Data Akademik Institusi	Akurasi	83.7%
8	Alalawi et al. (2023)	Machine Learning Models	Systematic Review Dataset	Multi-metrik	Sintesis Literatur
9	Sandeepa Mohottala (2025)	Ensemble Models	Academic Performance Data	Akurasi, Precision	92.3%
10	Hidayah et al. (2025)	Random Forest, Hyperparameter	Data Gaya Belajar Web	Akurasi	87.6%

Merujuk pada Tabel 4, mayoritas literatur (16 dari 20 artikel) berasal dari rentang tahun 2023–2026, yang menunjukkan tren kebaruan studi yang sangat tinggi dan relevansi topik ini dengan perkembangan riset Machine Learning terkini di bidang pendidikan. Dari sisi algoritma, sembilan artikel menerapkan algoritma berbasis ensemble (Random Forest dan/atau XGBoost) baik secara tunggal maupun dikombinasikan dengan algoritma lain, empat artikel menggunakan Support Vector Machine atau Naive Bayes sebagai model pembandingan, dan sisanya menerapkan pendekatan lain seperti Predictive Analytics dan model ensemble campuran. Dataset yang digunakan didominasi oleh data akademik institusional internal (IPK, nilai mata kuliah, dan kehadiran) yang digunakan pada 13 artikel, log aktivitas digital mahasiswa pada platform LMS yang digunakan pada 4 artikel, serta data kuesioner gaya belajar dan data publik akademik pada artikel-artikel lainnya. Variasi jenis dataset ini mengindikasikan bahwa performa model prediksi kinerja akademik sangat bergantung pada ketersediaan dan kelengkapan data institusional, sehingga menjadi pertimbangan penting bagi peneliti maupun praktisi yang ingin mereplikasi pendekatan serupa pada institusi lain.

3.3. RQ1: Algoritma Machine Learning yang Paling Banyak Digunakan

Berdasarkan tabulasi karakteristik studi pada Tabel 4, algoritma berbasis ensemble seperti Random Forest dan XGBoost merupakan model yang paling dominan digunakan dalam penelitian prediksi kinerja akademik (Istiwana et al., 2026; Eirene et al., 2025; Hairani et al., 2025; Hidayah et al., 2025). Penggunaan Random Forest dan XGBoost yang massif disebabkan oleh kemampuan algoritma tersebut dalam menangani data tabular berdimensi tinggi serta ketahanannya terhadap masalah overfitting (Hairani et al., 2025). Selain itu, Support Vector Machine (SVM) dan Naive Bayes tetap menjadi pilihan populer untuk pemodelan dasar (baseline) karena efisiensi komputasinya yang ringan pada ukuran dataset berskala kecil hingga menengah (Maulikha et al., 2026; Kelly & Xiu, 2023).

3.4. RQ2: Jenis Data dan Fitur yang Paling Berpengaruh

Analisis terhadap fitur prediktor mengungkapkan bahwa data akademik historis (termasuk Nilai IPK kumulatif, nilai mata kuliah prasyarat, dan tingkat kehadiran) merupakan fitur dengan bobot kepentingan (feature importance) tertinggi dalam menghasilkan prediksi yang akurat (Istiwana et al., 2026; Hairani et al., 2025; Sukmaningtyas et al., 2024). Selain data akademik, integrasi data perilaku belajar yang diekstraksi dari Learning Management System (LMS)—seperti frekuensi akses materi, durasi menonton video pembelajaran, dan intensitas forum diskusi—terbukti secara signifikan meningkatkan sensitivitas model dalam mendeteksi mahasiswa yang berisiko mengalami penurunan prestasi atau dropout secara dini (Ardianti et al., 2024).

3.5. RQ3: Performa Algoritma Machine Learning yang Dilaporkan

Performa akurasi yang dilaporkan dalam studi terpilih pada Tabel 4 menunjukkan rentang nilai yang sangat baik, secara umum berkisar antara 82% hingga 96% (Istiwana et al., 2026; Eirene et al., 2025; Maulikha et al., 2026; Hairani et al., 2025; Sukmaningtyas et al., 2024; Pelima et al., 2024; Kelly & Xiu, 2023; Alalawi et al., 2023;

Sandeepa & Mohottala, 2025; Hidayah et al., 2025). Algoritma XGBoost secara konsisten mencatatkan nilai akurasi dan F1-Score tertinggi ketika diimplementasikan pada dataset akademik multi-dimensi berskala besar (Istiwana et al., 2026; Sandeepa & Mohottala, 2025). Di sisi lain, model klasifikasi linear seperti Naïve Bayes memberikan performa yang cukup stabil namun cenderung mengalami penurunan performa jika dihadapkan pada dataset yang memiliki hubungan non-linear kompleks antar fiturnya (Maulikha et al., 2026; Kelly & Xiu, 2023).

3.6. RQ4: Tantangan dan Celah Penelitian (Research Gap)

Beberapa tantangan utama yang berhasil diidentifikasi dari penelusuran literatur mencakup masalah ketidakseimbangan kelas (class imbalance), di mana jumlah data mahasiswa dengan kinerja buruk jauh lebih sedikit dibandingkan mahasiswa berkinerja baik (Alalawi et al., 2023). Celah penelitian lainnya adalah sifat model tingkat lanjut (seperti Deep Learning atau Ensemble rumit) yang bertindak sebagai "black-box", sehingga sulit diinterpretasikan oleh praktisi pendidikan (Sandeepa & Mohottala, 2025). Oleh karena itu, pengembangan metode Explainable Artificial Intelligence (XAI) dalam prediksi pendidikan menjadi arah riset yang sangat potensial di masa depan (Alalawi et al., 2023; Sandeepa & Mohottala, 2025).

3.7. Pembahasan Umum

Sintesis menyeluruh dari temuan RQ1 hingga RQ4 menunjukkan adanya pergeseran paradigma dari model prediksi reaktif konvensional ke arah ekosistem analitik proaktif berbasis kecerdasan buatan di lingkungan perguruan tinggi (Pelima et al., 2024; Alalawi et al., 2023). Implementasi model Machine Learning terbukti krusial dalam membantu program studi menyusun strategi intervensi yang presisi (Sandeepa & Mohottala, 2025). Integrasi seimbang antara data demografis, rekam jejak akademik, dan performa LMS memfasilitasi akurasi model yang lebih kokoh dan aplikatif untuk diintegrasikan ke dalam system informasi akademik kampus (Ardianti et al., 2024).

3.8. Proses Algoritma dalam Menyelesaikan Masalah Prediksi

Secara teknis, proses algoritma Machine Learning dalam menyelesaikan permasalahan prediksi kinerja akademik pada studi-studi terpilih umumnya mengikuti alur yang serupa. Tahap pertama adalah praproses data (data preprocessing), meliputi penanganan nilai yang hilang (missing value), normalisasi skala fitur numerik seperti IPK dan skor keaktifan LMS, serta encoding terhadap fitur kategorikal seperti program studi atau gaya belajar. Tahap kedua adalah pemilihan fitur (feature selection), di mana algoritma seperti Random Forest memanfaatkan nilai feature importance untuk mengidentifikasi variabel akademik dan perilaku belajar yang paling berpengaruh terhadap prediksi, sehingga mengurangi dimensi data tanpa kehilangan informasi yang signifikan (Hairani et al., 2025; Sandeepa & Mohottala, 2025). Tahap ketiga adalah pelatihan model (model training); pada algoritma berbasis ensemble seperti Random Forest dan XGBoost, model dibangun secara iteratif melalui pembentukan banyak pohon keputusan yang saling melengkapi. XGBoost secara khusus menerapkan mekanisme gradient boosting yang memperbaiki kesalahan (residual error) dari pohon sebelumnya pada setiap iterasi, sehingga umumnya menghasilkan konvergensi akurasi yang lebih cepat dibandingkan Random Forest konvensional (Istiwana et al., 2026; Hidayah et al., 2025). Tahap keempat adalah validasi model, umumnya menggunakan skema k-fold cross-validation atau hold-out untuk memastikan model tidak overfitting terhadap data latih. Tahap terakhir adalah evaluasi performa menggunakan metrik Accuracy, Precision, Recall, dan F1-Score, yang menjadi dasar penentuan model terbaik pada masing-masing studi. Alur inilah yang menjelaskan mengapa algoritma ensemble secara konsisten unggul pada Tabel 4, yaitu karena kemampuannya menangkap pola non-linear antar fitur akademik dan perilaku belajar yang tidak dapat ditangkap secara optimal oleh model linear seperti Naive Bayes.

3.9. Perbandingan dengan Penelitian Terdahulu

Apabila dibandingkan dengan penelitian terdahulu, temuan tinjauan ini memperkuat sekaligus memperluas hasil yang dilaporkan oleh Pelima dkk. (2024), yang juga menemukan keunggulan algoritma ensemble dalam memprediksi luaran akademik, namun penelitian tersebut berfokus pada prediksi kelulusan (graduation) dan belum menyoroti integrasi data aktivitas LMS secara mendalam sebagaimana ditemukan pada beberapa studi dalam tinjauan ini (Istiwana et al., 2026; Eirene et al., 2025; Sandeepa & Mohottala, 2025). Dibandingkan dengan Alalawi dkk. (2023) yang melaporkan rentang akurasi model prediksi performa mahasiswa pada level yang lebih bervariasi karena turut mencakup studi-studi sebelum tahun 2023, tinjauan ini menunjukkan bahwa performa algoritma pada studi-studi terbaru (2023–2026) relatif lebih tinggi dan stabil, yang mengindikasikan adanya perbaikan praktik rekayasa fitur (feature engineering) dan optimasi hyperparameter pada penelitian-penelitian mutakhir (Hidayah et al., 2025). Sementara itu, hasil pada tinjauan ini juga sejalan dengan Sandeepa dan Mohottala (2025) yang melaporkan akurasi tinggi pada model ensemble, namun tinjauan ini memperluas temuan tersebut dengan menunjukkan bahwa pola keunggulan algoritma ensemble konsisten ditemukan lintas 20 studi dengan karakteristik dataset yang berbeda-beda, bukan hanya pada satu dataset spesifik. Perbedaan mendasar lainnya adalah bahwa sebagian besar penelitian terdahulu belum secara eksplisit membahas tantangan class imbalance dan

interpretabilitas model sebagai bagian dari pembahasan utama, sementara tinjauan ini secara khusus mengangkat kedua isu tersebut sebagai celah penelitian (research gap) yang perlu ditindaklanjuti melalui pendekatan Explainable Artificial Intelligence (XAI) pada penelitian mendatang. Dengan demikian, tinjauan ini tidak hanya mengonfirmasi temuan penelitian sebelumnya, tetapi juga memberikan sintesis yang lebih komprehensif dan mutakhir mengenai penerapan Machine Learning dalam prediksi kinerja akademik mahasiswa.

4. Kesimpulan

Tinjauan sistematis berbasis PRISMA menyaring 145 artikel menjadi 20 studi yang relevan mengenai prediksi kinerja akademik mahasiswa. Random Forest dan XGBoost menjadi algoritma paling dominan dengan akurasi 82%–96%. Rekam akademik historis, terutama IPK, nilai semester, mata kuliah prasyarat, dan kehadiran, merupakan prediktor utama, kemudian diikuti aktivitas Learning Management System. Keberhasilan model bergantung pada kualitas serta keberagaman data. Penelitian juga menemukan kendala berupa ketidakseimbangan kelas, keterbatasan validasi lintas institusi, dan rendahnya interpretabilitas model. Pengembangan berikutnya perlu mengintegrasikan data multidimensi, penanganan class imbalance, dan Explainable Artificial Intelligence agar sistem peringatan dini lebih akurat, transparan, etis, dan berkelanjutan bagi perguruan tinggi Indonesia.

Reference

- 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. Wiley. <https://doi.org/10.1002/eng2.12699>
- Aman, A., Rahrahima, N. P., & Fitri, A. (2026). IJATIS: Indonesian Journal of Applied Technology and Innovation Science – Implementation of machine learning algorithms for predicting student academic performance. IJATIS, 3(1), 1–9. <https://doi.org/10.57152/IJATIS.v3i1.1871>
- Amin, R., Siwi, A., & Utami, F. (2025). Prediksi nilai ujian berdasarkan kebiasaan siswa menggunakan algoritma random forest regressor. Information System for Educators and Professionals, 10(2), 149–158.
- Ardianti, M., Nurhayati, O. D., & Warsito, B. (2024). Model prediksi kinerja siswa berdasarkan data log LMS menggunakan ensemble machine learning. JST (Jurnal Sains dan Teknologi), 12(3). <https://doi.org/10.23887/jstundiksha.v12i3.59816>
- Caesarita, Y., Wahyudi, N., & Faisol, A. K. (2026). Prediksi kelulusan mahasiswa tepat waktu menggunakan perbandingan algoritma Random Forest dan XGBoost. INTER TECH, 4(1), 38–45. <https://doi.org/10.54732/4nzaxe76>
- Eirene, C., Syafitri, D., Sulistianingsih, N., Hidjah, K., & Hairani, H. (2025). Classification of learning styles of junior high school students using Random Forest & XGBoost algorithm. Jurnal Bumigora Information Technology (BITE), 7(1), 15–24. <https://doi.org/10.30812/bite.v7i1.4913>
- Hairani, A. S., Cahyono, R., & Balta, A. A. (2025). Sistem pendidikan kinerja siswa berbasis web menggunakan algoritma Decision Tree dan XGBoost. Jurnal Pustaka Data, 5(2), 323–330. <https://doi.org/10.55382/jurnalpustakadata.v5i2.1450>
- Hapsari, L. N., Ridhwanullah, D., & Miftakhurrokhmat, M. (2026). Prediksi kelulusan siswa sekolah menengah menggunakan XGBoost dan analisis feature importance. Jurnal Minfo Polgan, 15(2), 954–960. <https://doi.org/10.33395/jmp.v15i2.15782>
- Hidayah, R. B., Latifah, K., & Herlambang, B. A. (2025). Prediksi gaya belajar menggunakan algoritma Random Forest dengan optimasi hyperparameter tuning berbasis web. Rabbit: Jurnal Teknologi dan Sistem Informasi Univrab, 10(2), 864–873. <https://doi.org/10.36341/rabit.v10i2.6393>
- Istiwana, A. P., Sani, R. R., & Pramudi, Y. T. C. (2026). Pendekatan explainable machine learning untuk analisis faktor drop out mahasiswa menggunakan XGBoost. Rabbit: Jurnal Teknologi dan Sistem Informasi Univrab, 11(1), 1074–1083. <https://doi.org/10.36341/rabit.v11i1.7218>
- Kelly, B. T., & Xiu, D. (2023). Financial machine learning.
- Maulikha, A. P., Hiranisa, N. A., Ahmad, F. A., Pamungkas, C. R., Ardiansyah, M. F., & Nikmah, F. (2026). Analisis perbandingan algoritma machine learning dalam prediksi akademik mahasiswa berbasis data.
- Muhammad, A., & Achmad, F. (2023). Prediksi performansi mahasiswa dengan mempertimbangkan motivasi intrinsik menggunakan machine learning, 8(4).
- Nugroho, B. I., Santoso, N. A., & Murtopo, A. A. (2023). Prediksi kemampuan akademik mahasiswa dengan metode Support Vector Machine. Remik, 7(1), 177–188. <https://doi.org/10.33395/remik.v7i1.12010>
- Pelima, L. R., Sukmana, Y., & Rosmansyah, Y. (2024). Predicting university student graduation using academic performance and machine learning: A systematic literature review. IEEE Access, 12, 23451–23465. <https://doi.org/10.1109/ACCESS.2024.3361479>
- Prima, P., Adriansyah, A. R., & Usyaid, A. N. (2026). MIND (Multimedia Artificial Intelligent Networking Database): Prediksi retensi mahasiswa menggunakan algoritma Random Forest dengan optimasi algoritma genetika. Journal MIND, 11(1), 115–126. <https://doi.org/10.26760/mindjournal.v11i1.115-126>
- Rismaya, R., Yuniarto, D., & Setiadi, D. (2025). Penerapan algoritma machine learning dalam prediksi prestasi akademik mahasiswa. Router: Jurnal Teknik Informatika dan Terapan, 3(1), 15–23. <https://doi.org/10.62951/router.v3i1.389>
- Sandeepa, A. G. R., & Mohottala, S. (2025). Evaluation of machine learning models in student academic performance prediction. <https://doi.org/10.1109/ICARC64760.2025.10963104>
- Sukmaningtyas, Y. N., Akbar, R. M., & Asyafiyah, G. R. U. (2024). Penerapan predictive analytics untuk analisis faktor-faktor yang mempengaruhi performa akademik siswa. Arcitech: Journal of Computer Science and Artificial Intelligence, 4(2), 127–145. <https://doi.org/10.29240/arcitech.v4i2.12048>
- Suniantara, I. K. P., Suwardika, G., & Soraya, S. (2020). Peningkatan akurasi klasifikasi ketidaktepatan waktu kelulusan mahasiswa menggunakan metode boosting neural network. Jurnal Varian, 3(2), 95–102. <https://doi.org/10.30812/varian.v3i2.651>