



Prediksi Tingkat Pemahaman Siswa SMP Pada Pembelajaran Digital Menggunakan Algoritma Machine Learning

Munadi Fauzan^{1(*)}, Dinda Permata Sukma¹, Eriwandi¹, Diana Kemala Odang¹

¹Program Studi Informatika, Universitas Alifiah Padang, Indonesia

Abstract

With the proliferation of digital learning platforms—including Learning Management Systems (LMS)—data on student learning activities has been generated that is useful for supporting the evaluation of student learning. A total of 200 data points were analyzed, with key variables including login frequency, study duration, quiz scores, assignment scores, timeliness in submitting assignments, attendance, and engagement during learning. These variables can be analyzed using machine learning to predict students' comprehension levels early on, enabling teachers to identify which students require additional support. This study focuses on the application of machine learning to predict students' comprehension levels at SMP Negeri 18 Padang as an effort to support more accurate decision-making in the learning process. This study utilized three algorithms: Random Forest, Decision Tree, and Naïve Bayes. The results of the testing revealed that the Decision Tree algorithm achieved the highest accuracy rate at 88.52%, followed by the Random Forest algorithm at 83.61%, and the Naïve Bayes algorithm at 62.30%. Of the three algorithms, the Decision Tree algorithm achieved the highest accuracy. This algorithm can be utilized as an early warning system for teachers to identify which students have high, moderate, or low levels of understanding, allowing teachers to provide appropriate support.

Kata Kunci: Machine Learning, Predicting Student, Learning Management System (LMS), Naïve Bayes.

Informasi Artikel:

Dikirim : 16 Juli 2026
Direvisi : 22 Juli 2026
Diterima : 24 Juli 2026
Diterbitkan : 26 Juli 2026

Januari – Juni 2026, Vol 7 (1) : hlm 23-30
Institut Teknologi dan Bisnis Ahmad Dahlan

(*) Korespondensi: munadifauzan@gmail.com (Munadi Fauzan)

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi digital pada berbagai bidang, termasuk pada bidang pendidikan. Bentuk implementasinya adalah penggunaan *Learning Management System* (LMS) yang mampu merekap berbagai aktivitas belajar siswa, seperti frekuensi login, durasi belajar, penyelesaian tugas, hasil kuis, absensi, serta interaksi terhadap materi pembelajaran. Data yang didapatkan dari hasil *Learning Management System* (LMS) bukan hanya sebagai dokumentasi proses belajar mengajar, tetapi juga menjadi sumber informasi yang dapat dimanfaatkan untuk mengevaluasi dan meningkatkan kualitas pembelajaran di dalam kelas.

Peningkatan penggunaan *Learning Management System* (LMS) menghasilkan *educational big data* yang memiliki potensi yang besar untuk menganalisis menggunakan pendekatan *Educational Data Mining* (EDM) dan *Machine Learning* (ML) dalam mendukung pengambilan keputusan pada bidang pendidikan. Melalui pemanfaatan data ini, sekolah dapat memperoleh informasi yang lebih objektif tentang perkembangan akademik siswa dibandingkan dengan evaluasi yang hanya didasarkan pada hasil ujian akhir saja. Alalawi, Athauda and Chiong (2023).

Machine learning adalah salah satu cabang dari kecerdasan buatan yang mampu mempelajari pola dari data historis untuk menghasilkan prediksi terhadap data yang baru. Dalam pendidikan, teknologi dimanfaatkan untuk memprediksi prestasi nilai akademik siswa mengidentifikasi siswa yang beresiko mengalami kesulitan dalam belajar serta mendukung agar pembelajaran lebih adaptif. *Machine learning* dapat membantu pendidikan dalam mengenali apa saja faktor – faktor yang mempengaruhi keberhasilan akademik siswa sehingga tindakan dapat dilakukan sebelum terjadinya penurunan terhadap prestasi siswa. Ahmed (2024). Pengembangan model prediksi bukan saja untuk memperoleh nilai akurasi yang tinggi, tetapi juga mampu menghasilkan sistem peringatan dini (*early warning system*) yang dapat membantu para guru memberikan umpan balik dan memberikan pendampingan lebih tepat waktu. Alalawi, Athauda, and Chiong (2023).

Salah satu penelitian mengenai prediksi peningkatan kemampuan siswa menggunakan *machine learning* terus berkembang dalam beberapa tahun terakhir. Dari beberapa artikel ilmiah, metode klasifikasi adalah metode yang paling banyak digunakan dalam memprediksi tingkat kemampuan siswa. Dalam penelitian ini menunjukkan lima algoritma yang paling sering diterapkan adalah *Decision Tree*, *Random Forest*, *Naïve Bayes*, *Artificial Neural Network* (ANN), dan *Support Vector Machine* (SVM). Alalawi, athauda and Chiong, 2023. Algoritma yang berbasis *Artificial Neural Network* mempunyai kemampuan yang baik dalam memanfaatkan data aktifitas LMS untuk meningkatkan akurasi dalam memprediksi kemampuan siswa, terutama dalam pembelajaran daring (*online*). Xiao and Hu, 2023. Hasil – hasil penelitian tersebut menunjukkan bahwa pemilihan algoritma dan kualitas merupakan faktor yang penting dalam menghasilkan model prediksi.

Meskipun berbagai algoritma sudah memiliki hasil yang baik, namun implementasi *machine learning* pada lingkungan sekolah masih menghadapi

beberapa kendala yang meliputi ketersediaan data, proses prapemrosesan data, pemilihan atribut dan kesiapan sekolah dalam memanfaatkan hasil prediksi yang menjadi faktor dalam mempengaruhi keberhasilan penerapan model. Akurasi model sangat dipengaruhi oleh kualitas data masukan dan proses seleksi. Hasil prediksi dapat memberikan manfaat yang lebih besar apabila diintegrasikan dalam pembelajaran melalui pendekatan *learning analytics*, sehingga guru bukan hanya memperoleh hasil klasifikasi, tetapi juga mendapatkan rekomendasi tindakan yang dapat dilakukan terhadap siswa yang diprediksi mengalami kesulitan belajar.

Penelitian ini bertujuan menerapkan algoritma *machine learning* untuk memprediksi tingkat pemahaman siswa di SMPN 18 Padang menggunakan data nilai tugas, nilai kuis, kehadiran dan aktifitas pembelajaran digital. Hasil dari penelitian ini diharapkan dapat menjadi sistem pendukung keputusan bagi guru dalam proses belajar mengajar dan menjadi lebih efektif, efisien dan berbasis kan data. Selain itu, manfaat dari penelitian ini juga mendukung dalam peningkatan kualitas pembelajaran di sekolah.

METODE

Tahapan 1 : Perumusan Masalah

Langkah awal dalam metode ini adalah merumuskan masalah yang ingin diselesaikan sebelum melakukan proses selanjutnya. Pada tahapan ini dilakukan observasi di SMPN 18 Padang untuk melihat kondisi pembelajaran digital yang sedang berjalan dan platform apa yang digunakan kemudian data apa saja yang sudah tersedia dan seperti apa kebutuhan para guru dalam memantau tingkat pemahaman siswa sebelum ujian akhir dilaksanakan.

Selain observasi, dilakukan juga wawancara dengan guru mata pelajaran dan wali kelas untuk mengetahui variabel mana yang paling dominan dalam menentukan tingkat pemahaman seorang siswa dalam memahami pelajaran. Tahapan ini merupakan tahapan dasar untuk menentukan target prediksi dan variabel input yang akan dimasukkan ke dalam model.

Tahapan 2 : Pengumpulan data

Setelah perumusan masalah, langkah selanjutnya adalah mengumpulkan data yang dibutuhkan. Dalam penelitian ini memiliki 2 sumber data. Pertama, *platform* pembelajaran yang digunakan oleh SMPN 18 Padang. Kedua, dari data akademik siswa yang disimpan oleh bagian administrasi sekolah.

Variabel yang dikumpulkan mencakup frekuensi login, durasi belajar, nilai kuis, nilai tugas, ketepatan waktu dalam mengumpulkan tugas, kehadiran dan keaktifan ketika belajar, nilai akhir sebagai label target (tinggi, sedang atau rendah). Setelah data terkumpulkan, dilakukan eksplorasi awal untuk memahami distribusi data, mengidentifikasi nilai yang hilang (*missing value*), data yang duplikat dan melihat apakah ada ketimpangan jumlah antar kelas (*imbalanced data*). Ini adalah

tahapan penting karena kualitas data sangat menentukan kualitas prediksi yang akan keluar.

Tahapan 3 : Persiapan Data

Pada tahapan ini adalah tahap persiapan data yang mencakup pembersihan data, pra-pemrosesan, dan seleksi fitur sebagai bagian dari alur penelitian sebelum masuk ke dalam tahap pemodelan. Pembersihan data dilakukan untuk menangani nilai data yang kosong diisi menggunakan nilai rata-rata (untuk data numerik) atau modus (untuk data kategorikal) atau menghapus jika proporsinya kecil. Setelah data di *cleaning* dan siap untuk dijalankan. Data dibagi menjadi dua bagian : data latih (*training set*) dan data uji (*testing set*). Pembagian data dilakukan dengan proporsi 80:20, dimana 80% data digunakan untuk melatih model dan 20% sisanya digunakan untuk menguji seberapa baik model dapat bekerja pada data yang belum pernah dilihat sebelumnya

Tahapan 4 : Evaluasi Model

Setelah model dilatih dan diuji dengan menggunakan data testing sudah dipisahkan sebelumnya. *Confusion matrix* digunakan sebagai alat utama untuk mengevaluasi kinerja model klasifikasi, karena dapat membandingkan hasil prediksi model dengan nilai yang sebenarnya dalam dataset pengujian.

Selain *confusion matrix*, digunakan juga *k-fold cross validation* untuk memastikan hasil evaluasi tidak bergantung pada satu pembagian data tertentu. Dari hasil evaluasi ini, algoritma yang menghasilkan performa terbaik akan dipilih sebagai model final yang akan digunakan dalam sistem prediksi. Kriteria yang dipilih bukan hanya berdasarkan akurasi tertinggi namun juga mempertimbangkan keseimbangan antara *precision* dan *recall*, terutama untuk memprediksi siswa dengan pemahaman yang rendah, karena jika salah dalam memprediksi kelompok maka bisa berakibat siswa tidak mendapat bantuan yang seharusnya.

Tahapan 6 : Implementasi dan Pendampingan

Model yang sudah dipilih kemudian diimplementasikan ke dalam sistem website sederhana yang dapat diakses oleh guru. Rekomendasi pengembangan sistem interaktif berbasis data menjadi langkah yang penting agar hasil dan prediksi model *machine learning* dapat dimanfaatkan oleh para guru.

Setelah sistem selesai dibangun, dilakukan uji coba dengan para guru di SMPN 18 Padang dalam pelatihan untuk para guru, selanjutnya guru diminta untuk mencoba sistem dengan data kelas mereka sendiri. Kemudian memberikan umpan balik tentang kemudahan penggunaan dan kegunaan informasi yang ditampilkan sebagai acuan untuk mengambil keputusan.

HASIL DAN PEMBAHASAN

Penelitian dilaksanakan di SMPN 18 padang dengan memanfaatkan data aktivitas pembelajaran digital siswa yang diperoleh dari platform *Learning Management System* (LMS) dan data nilai akademik sekolah. Beberapa variabel yang digunakan dalam kegiatan penelitian ini adalah durasi penggunaan *Learning Management System* (LMS), durasi belajar pada setiap sesinya, jumlah akses mengulang materi, absensi pembelajaran digital, ketepatan dalam pengumpulan tugas, nilai tugas, nilai kuis, keaktifan dan kategori tingkat pemahaman siswa. Setelah dilakukan proses *cleaning* di peroleh sebanyak 200 data yang layak untuk digunakan sebagai dataset penelitian.

Selanjutnya tahapan preprocessing yang meliputi, menghapus data duplikat, penanganan *missing value*, Normalisasi menggunakan *Min-Max Scaling*, Label *Encoding* pada variabel kategori, Pembagian data *training* dan *testing* sebesar 80:20. Dari tahapan tersebut di maka di peroleh *dataset Training* sebanyak 163 data dan *dataset Testing* sebanyak 41 data.

accuracy: 88.52%

	true Tinggi	true Sedang	true Rendah	class precision
pred. Tinggi	29	2	2	87.88%
pred. Sedang	1	19	1	90.48%
pred. Rendah	1	0	6	85.71%
class recall	93.55%	90.48%	66.67%	

Gambar 1. Hasil akurasi algoritma *Decision Tree*

accuracy: 83.61%

	true Tinggi	true Sedang	true Rendah	class precision
pred. Tinggi	28	4	0	87.50%
pred. Sedang	1	16	2	84.21%
pred. Rendah	2	1	7	70.00%
class recall	90.32%	76.19%	77.78%	

Gambar 2. Hasil akurasi algoritma *Random Forest*

accuracy: 62.30%

	true Tinggi	true Sedang	true Rendah	class precision
pred. Tinggi	22	7	0	75.86%
pred. Sedang	7	8	1	50.00%
pred. Rendah	2	6	8	50.00%
class recall	70.97%	38.10%	88.89%	

Gambar 3. Hasil akurasi algoritma Naïve Bayes

Hasil akurasi dari pemodelan *machine learning* dengan membandingkan 3 algoritma *machine learning* yaitu *Decision Tree*, *Random Forest* dan *Naive Bayes*. Hasil akurasi dari ketiga algoritma menunjukkan bahwa algoritma *Decision Tree* mendapatkan hasil yang tinggi dibandingkan algoritma yang lain.

Berdasarkan hasil pengujian, algoritma *Decision Tree* menghasilkan tingkat akurasi sebesar 88,52%, lebih tinggi dibandingkan dengan algoritma *Random Forest* sebesar 83,61% dan algoritma *Naive Bayes* 62,30%. Hal ini menunjukkan bahwa algoritma *Decision Tree* memiliki hasil yang cukup tinggi dan mampu mengenali pola aktifitas pembelajaran digital dengan tingkat pemahaman siswa.

Analisis *feature importance* menunjukkan bahwa nilai kuis, nilai tugas, dan frekuensi akses *Learning Management System* (LMS) merupakan tiga variabel yang paling berpengaruh dalam menentukan tingkat pemahaman siswa. Hal ini mengindikasikan bahwa siswa yang aktif mengakses materi pembelajaran serta memperoleh nilai evaluasi yang baik cenderung memiliki tingkat pemahaman yang lebih tinggi.

Dari hasil model prediksi tersebut dapat dimanfaatkan oleh guru sebagai sistem pendukung keputusan dalam proses pembelajaran digital sehingga guru dapat mengidentifikasi siswa yang diprediksi memiliki tingkat pemahaman yang rendah sebelum ujian akhir dilaksanakan dan dapat dilakukan tindakan remedial, pendampingan belajar serta pemberian materi tambahan. Dengan demikian proses pembelajaran lebih adaptif dan berbasis data.

KESIMPULAN

Kesimpulan dari penelitian yang sudah dilakuakn adalah variabel yang paling berpengaruh terhadap tingkat pemahaman siswa pada pembelajaran digital menggunakan algoritma Machine Learning adalah nilai kuis, nilai tugas dan frekuensi akses Learning Management System (LMS). Dari ketiga algoritma yang dibandingkan. Algoritma *Decision Tree* menghasilkan performa yang terbaik dibandingkan algoritma *Random Forest* dan *Naïve Bayes*. Model yang dibangun dapat mengklasifikasikan tingkat pemahaman siswa yang di bagi kedalam kategori tinggi, sedang, dan rendah dengan tingkat ketepatan yang cukup tinggi. Dan sistem prediksi tingkat pemahaman siswa berbasis machine learning dapat dimanfaatkan sebagai alat bantu bagi guru di SMPN 18 Padang dalam mendeteksi siswa yang memerlukan pendampingan pembelajaran lebih awal sebelum ujian akhir dilaksanakan sehingga pembelajaran dapat berlangsung lebih efektif.

DAFTAR PUSTAKA

Amor, A. F., Ihsanudin, M. R., Yulistianto, S. O., & Hidayat, H. (2025). *Prediksi Kinerja Siswa SMP Berdasarkan Data Akademik dan Perilaku Menggunakan Machine Learning. Teknomatika: Jurnal Informatika dan Komputer*, 18(2), 67–76.

- Ardianti, M. (2023). Model Prediksi Kinerja Siswa Berdasarkan Data Log LMS Menggunakan Ensemble Machine Learning. *Jurnal Sains dan Teknologi*, 12(3), 562–571. <https://doi.org/10.23887/jstundiksha.v12i3.59816>
- 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>
- 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>
- 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>
- Learning. *Teknomatika: Jurnal Informatika dan Komputer*, 18(2), 67–76. <https://doi.org/10.36040/teknomatika.v18i2.1639>
- Muzakir, A., Adi, K., & Kusumaningrum, R. (2024). Inovasi Prediksi Hasil Belajar Siswa Berbasis Machine Learning untuk Mendukung Keputusan Pembelajaran yang Tepat di SMKN 2 Semarang. *Jurnal Majemuk*, 4(2), 420–432.
- Nashrullah, M., Syaiful Rahman, Abdul Majid, Hariyati, N., & Budiyanto. (2025). Transformasi Digital dalam Pendidikan Indonesia: Analisis Kebijakan dan Implikasinya terhadap Kualitas Pembelajaran. *Mudir: Jurnal Manajemen Pendidikan*, 7(1), 52–59.
- Pranata, B. D., Mahdiyah, U., & Kasih, P. (2025). Transformasi Digital dalam Pendidikan: Penerapan Algoritma Machine Learning dan Deep Learning untuk Meningkatkan Efektivitas Pembelajaran SMP. *Jurnal Ekonomi Manajemen Sistem Informasi*. <https://doi.org/10.38035/jemsi>
- Pranata, B. D., Mahdiyah, U., & Kasih, P. (2025). Transformasi Digital dalam Pendidikan: Penerapan Algoritma Machine Learning dan Deep Learning untuk Meningkatkan Efektivitas Pembelajaran SMP. *Jurnal Ekonomi Manajemen Sistem Informasi*. <https://doi.org/10.38035/jemsi.v6i4.6569>
- Purnasari, P. D., Sadewo, Y. D., Santosa, D. S. S., & Sanoto, H. (2024). Analisis Digitalisasi Pembelajaran Sekolah Dasar Wilayah Perbatasan. *Scholaria: Jurnal Pendidikan dan Kebudayaan*, 14(2), 198–205. <https://doi.org/10.24246/j.js.2024.v14.i2.p198-205>
- 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>
- Salim, et al. (2024). Analisis Pengaruh Tingkat Literasi Digital Guru dan Siswa terhadap Kualitas Pembelajaran di Era Digital di Indonesia. *Jurnal Pendidikan Digital*, 2(2), 50–60.

- Sembiring, D., & Kona, M. (2024). *Inovasi Teknologi Pembelajaran*. Solok: Mafy Media Literasi.
- Setiadi, D., & Husni, H. (2025). *Metode Machine Learning dan Deep Learning dalam Prediksi Kinerja Siswa: Tinjauan Sistematis*. ResearchGate. <https://doi.org/10.13140/RG.2.2.xxxxx>
- Setiadi, D., & Husni, H. (2025). *Metode Machine Learning dan Deep Learning dalam Prediksi Kinerja Siswa: Tinjauan Sistematis*. ResearchGate. <https://www.researchgate.net/publication/388549021>
- Shafa, A. (2024). *Implementasi Learning Management System dalam Meningkatkan Efektivitas Pembelajaran*. *Jurnal Teknologi Pendidikan*, 1(4). <https://doi.org/10.47134/jtp.v1i4.658>
- Sugandi, Z. A. W., & Ababil, M. I. (2025). *Inovasi Prediksi Hasil Belajar Siswa Berbasis Machine Learning untuk Mendukung Keputusan Pembelajaran yang Tepat di SMKN 2 Semarang*. *Jurnal Majemuk*. <https://jurnalilmiah.org/journal/index.php/majemuk/article/view/1123>
- Wahyuni, R., et al. (2024). *Prediksi Performa Akademik Siswa Berdasarkan Kehadiran dan Aktivitas E-Learning Menggunakan Algoritma Decision Tree*. *RIGGS: Journal of Artificial Intelligence and Digital Business*. <https://doi.org/10.33020/riggs.v1i1.1352>
- Wiryanto, W., Kumala, I., & Adhiyati, U. P. (2023). *Efek Metode Pembelajaran Berbasis E-Learning pada Motivasi Belajar Peserta Didik SMKN 1 Lelea Kabupaten Indramayu*. *Jurnal Pendidikan dan Pembelajaran Indonesia*.