

Pengembangan Model Prediksi Keberhasilan Mahasiswa Menggunakan Algoritma Machine Learning Dalam *Learning Management System*

Edi Tohidi¹, Irfan Ali^{2*}

¹Program Studi Komputerisasi Akuntansi, STMIK IKMI Cirebon, Kota Cirebon, Indonesia

²Program Studi Rekayasa Perangkat Lunak, STMIK IKMI Cirebon, Kota Cirebon, Indonesia

Email: ¹editohidi.ikmi@gmail.com, ^{2*}Irfanali.ikmi@email.com

(* : coresponding author)

Abstrak – Perkembangan teknologi digital dalam dunia pendidikan mendorong pemanfaatan sistem manajemen pembelajaran (*Learning Management System/LMS*) sebagai media pembelajaran daring yang efektif. Penelitian ini bertujuan untuk mengembangkan model prediksi keberhasilan mahasiswa dalam LMS menggunakan algoritma machine learning. Keberhasilan mahasiswa diklasifikasikan berdasarkan parameter seperti tingkat partisipasi, frekuensi akses, hasil evaluasi, dan ketepatan waktu dalam mengumpulkan tugas. Berbagai algoritma machine learning seperti Decision Tree, *Random Forest*, *Support Vector Machine*, dan K-Nearest Neighbors digunakan untuk membangun model prediksi. Evaluasi performa model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* memberikan performa terbaik dengan akurasi mencapai 89%, diikuti oleh *Support Vector Machine* dan Decision Tree. Model yang dikembangkan diharapkan dapat membantu pengajar dan institusi pendidikan dalam mengidentifikasi mahasiswa yang berpotensi mengalami kesulitan belajar sejak dini, sehingga dapat diberikan intervensi yang tepat. Penelitian ini memberikan kontribusi dalam penerapan machine learning untuk mendukung proses pembelajaran adaptif dan peningkatan kualitas pendidikan berbasis data.

Kata Kunci: Machine Learning, Prediksi Keberhasilan, Mahasiswa, *Learning Management System*, Klasifikasi

Abstract – The advancement of digital technology in education has driven the widespread adoption of *Learning Management Systems (LMS)* as effective platforms for online learning. This study aims to develop a predictive model for student success in LMS environments using machine learning algorithms. Student success is classified based on parameters such as participation levels, access frequency, assessment results, and punctuality in assignment submissions. Several machine learning algorithms, including Decision Tree, *Random Forest*, *Support Vector Machine*, and K-Nearest Neighbors, are employed to build the prediction model. The performance of each model is evaluated using metrics such as accuracy, precision, recall, and F1-score. The results show that the *Random Forest* algorithm achieved the best performance with an accuracy of 89%, followed by *Support Vector Machine* and Decision Tree. The developed model is expected to assist educators and academic institutions in identifying students who may face learning difficulties at an early stage, allowing for timely and targeted interventions. This research contributes to the application of machine learning in supporting adaptive learning processes and enhancing data-driven educational quality.

Keywords: Machine Learning, Student Success Prediction, *Learning Management System*, Classification, Education Analytics

1. PENDAHULUAN

Dalam dunia pendidikan yang semakin berkembang, pemantauan dan evaluasi keberhasilan mahasiswa menjadi aspek penting untuk meningkatkan kualitas pembelajaran. Salah satu tantangan utama yang dihadapi oleh institusi pendidikan adalah bagaimana memprediksi dan menganalisis potensi keberhasilan akademik mahasiswa secara lebih tepat dan berbasis data. Dengan berkembangnya teknologi digital, *Learning Management System (LMS)* menjadi platform yang sering digunakan untuk mendukung proses belajar mengajar. LMS tidak hanya menyediakan materi pembelajaran tetapi juga mengumpulkan data aktivitas mahasiswa, seperti frekuensi login, interaksi dengan konten pembelajaran, serta hasil evaluasi tugas yang dapat digunakan untuk menganalisis pola keberhasilan akademik.

Namun, meskipun LMS menyimpan data yang kaya, pengolahan dan analisis data tersebut secara manual sangat memakan waktu dan sering kali tidak dapat memberikan wawasan yang akurat. Oleh karena itu, diperlukan suatu sistem yang dapat secara otomatis memprediksi keberhasilan mahasiswa dengan memanfaatkan algoritma machine learning. Algoritma ini dapat menganalisis

data dari LMS untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap keberhasilan akademik mahasiswa dan memberikan rekomendasi bagi pengelola pendidikan.

Metode yang dapat digunakan untuk memprediksi keberhasilan akademik mahasiswa adalah algoritma machine learning, seperti *Random Forest*, *Support Vector Machine* (SVM), atau *Deep Neural Networks* (DNN). Dengan menggunakan algoritma tersebut, sistem dapat dilatih dengan data mahasiswa dan memberikan prediksi yang akurat tentang keberhasilan mereka. Melalui pendekatan ini, diharapkan mahasiswa yang berisiko dapat segera teridentifikasi dan diberikan intervensi lebih awal untuk meningkatkan peluang keberhasilan akademik mereka.

Penelitian ini bertujuan untuk mengembangkan model prediksi keberhasilan mahasiswa dengan menggunakan algoritma machine learning dalam *Learning Management System* (LMS). Diharapkan hasil dari penelitian ini dapat memberikan wawasan yang lebih baik mengenai bagaimana data yang tersedia di LMS dapat digunakan untuk mendukung keputusan pendidikan yang lebih efektif dan meningkatkan kualitas pembelajaran secara keseluruhan.

2. METODOLOGI PENELITIAN

2.1 Data dan Sumber

Data yang digunakan dalam penelitian ini berasal dari log aktivitas mahasiswa yang diperoleh melalui sistem *Learning Management System* (LMS) pada salah satu perguruan tinggi swasta di Indonesia. LMS yang digunakan merupakan platform pembelajaran daring yang mencatat berbagai aktivitas akademik mahasiswa selama satu semester penuh. Data dikumpulkan dari mata kuliah umum yang diikuti oleh mahasiswa dari berbagai program studi, sehingga memberikan keragaman perilaku belajar. Jumlah total data mencakup 350 mahasiswa dengan rekaman aktivitas mingguan selama 16 minggu perkuliahan.

Jenis data yang dihimpun dan digunakan dalam pemodelan meliputi:

- 1) Frekuensi Akses Harian ke LMS
Mencatat jumlah login yang dilakukan mahasiswa setiap harinya, sebagai indikator keterlibatan dalam pembelajaran.
- 2) Tingkat Kehadiran pada Sesi Sinkronus
Merekam keikutsertaan mahasiswa dalam pertemuan virtual melalui LMS atau platform konferensi video yang terintegrasi.
- 3) Pengumpulan Tugas
Meliputi jumlah tugas yang dikumpulkan, keterlambatan pengumpulan, serta keutuhan file tugas yang diunggah.
- 4) Nilai Kuis dan Ujian Tengah/Akhir Semester
Data numerik yang merepresentasikan tingkat penguasaan materi mahasiswa.
- 5) Aktivitas dalam Forum Diskusi
Termasuk jumlah postingan, jumlah komentar, dan tingkat partisipasi dalam diskusi daring yang diadakan oleh dosen.
- 6) Waktu Akses terhadap Materi Pembelajaran
Informasi ini menunjukkan seberapa aktif mahasiswa membuka dan membaca materi pembelajaran dalam LMS.
- 7) Indikator Keberhasilan Akademik (Label Kelas)
Data ini merupakan label kelas yang dibentuk berdasarkan capaian akhir mahasiswa, diklasifikasikan ke dalam dua kategori: *Berhasil* (nilai akhir ≥ 70) dan *Tidak Berhasil* (nilai akhir < 70).

- 8) Seluruh data diolah secara anonim, dan hanya digunakan untuk kepentingan penelitian ini dengan tetap menjaga etika dan privasi pengguna. Sebelum dilakukan analisis, data telah melalui proses validasi dan pembersihan untuk memastikan keakuratan serta kelengkapan informasi yang digunakan dalam proses pemodelan machine learning.

2.2 Pra-pemrosesan Data

Dalam membangun model prediksi keberhasilan mahasiswa, pemilihan fitur (fitur input atau *independent variables*) menjadi aspek krusial karena kualitas fitur akan sangat menentukan performa model. Fitur-fitur yang digunakan dalam penelitian ini berasal dari log aktivitas mahasiswa dalam LMS dan telah melalui tahap seleksi berdasarkan korelasi dan kontribusinya terhadap variabel target.

Berikut adalah deskripsi masing-masing fitur yang digunakan dalam penelitian:

- 1) Frekuensi Login Mingguan
Jumlah rata-rata login mahasiswa ke LMS per minggu. Fitur ini merepresentasikan tingkat konsistensi mahasiswa dalam mengakses platform pembelajaran.
- 2) Jumlah Tugas yang Dikumpulkan
Total tugas yang berhasil dikumpulkan oleh mahasiswa selama satu semester. Termasuk juga data keterlambatan dalam pengumpulan tugas.
- 3) Skor Rata-Rata Kuis
Nilai rata-rata dari seluruh kuis daring yang diikuti mahasiswa. Fitur ini mencerminkan pemahaman materi secara periodik.
- 4) Kehadiran dalam Pertemuan Sinkronus
Persentase kehadiran mahasiswa dalam perkuliahan daring (live session) baik melalui Zoom, Google Meet, maupun fitur conference dalam LMS.
- 5) Aktivitas Forum Diskusi
Jumlah kontribusi mahasiswa dalam forum diskusi LMS baik berupa inisiasi topik maupun komentar terhadap diskusi rekan sejawat.
- 6) Waktu Interaksi Terakhir (Days Since Last Access)
Mengukur seberapa lama waktu yang telah berlalu sejak terakhir kali mahasiswa mengakses LMS. Semakin kecil nilainya semakin aktif mahasiswa tersebut.
- 7) Durasi Akses Materi
Estimasi waktu yang dihabiskan mahasiswa dalam membaca atau membuka file materi yang tersedia di LMS.
- 8) Nilai Akhir Mata Kuliah (*Target / Label*)
 - a. Nilai akhir diklasifikasikan menjadi dua kategori:
 - b. 1 (Berhasil) jika nilai akhir ≥ 70
 - c. 0 (Tidak Berhasil) jika nilai akhir < 70

Fitur-fitur ini dipilih karena mewakili dimensi penting dalam pembelajaran daring, yaitu keterlibatan, keteraturan, pemahaman, dan kedisiplinan mahasiswa. Penggunaan fitur-fitur ini diharapkan mampu memberikan gambaran yang cukup akurat terhadap kecenderungan keberhasilan akademik mahasiswa.

2.4 Algoritma Machine Learning

Empat algoritma yang digunakan untuk pengujian model adalah:

- 1) Decision Tree : algoritma berbasis pohon keputusan yang mudah diinterpretasikan.

- 2) *Random Forest* : metode ensemble yang menggabungkan beberapa pohon keputusan.
- 3) *Support Vector Machine (SVM)* : algoritma klasifikasi yang efektif untuk data berdimensi tinggi.
- 4) *K-Nearest Neighbors (KNN)* : algoritma yang mengklasifikasikan berdasarkan kedekatan fitur antar data.

2.5 Evaluasi Model

Setelah proses pelatihan dilakukan, performa masing-masing model dievaluasi untuk mengetahui tingkat akurasi dan efektivitasnya dalam memprediksi keberhasilan mahasiswa. Evaluasi dilakukan dengan menggunakan beberapa metrik evaluasi yang umum digunakan dalam klasifikasi biner, yaitu:

1. Akurasi(Accuracy)

Persentase prediksi yang benar dari keseluruhan jumlah data uji. Akurasi mengukur seberapa sering model membuat prediksi yang benar.

2. Presisi(Precision)

Rasio antara jumlah prediksi *positif benar* (true positive) dengan total prediksi *positif* (true positive + false positive). Presisi penting untuk mengukur ketepatan model dalam memprediksi keberhasilan.

3. Recall(Sensitivity)/

Rasio antara jumlah prediksi *positif benar* dengan total jumlah *positif sebenarnya* (true positive + false negative). Recall mengukur kemampuan model dalam menangkap semua kasus yang sebenarnya berhasil.

4. F1-Score

Merupakan rata-rata harmonis antara presisi dan recall. F1-score memberikan ukuran performa keseluruhan, terutama berguna saat terdapat ketidakseimbangan antar kelas.

5. Cross-Validation(10-Fold)

Untuk menghindari overfitting dan memastikan generalisasi model, evaluasi dilakukan menggunakan teknik 10-fold cross-validation. Data dibagi ke dalam 10 bagian, 9 bagian digunakan untuk pelatihan dan 1 bagian untuk pengujian, lalu proses ini diulang hingga seluruh data digunakan untuk pengujian.

Setiap model kemudian dibandingkan berdasarkan hasil metrik di atas, dan model dengan performa terbaik akan diusulkan sebagai model utama yang dapat digunakan untuk diintegrasikan ke dalam sistem LMS.

3. ANALISA DAN PEMBAHASAN

3.1 Hasil Pelatihan dan Pengujian

Setelah dilakukan proses pelatihan, berikut hasil evaluasi terhadap masing-masing model:

Algoritma	Akurasi	Presisi	Recall	F1-Score
Decision Tree	84%	82%	81%	81.5%
<i>Random Forest</i>	89%	87%	88%	87.5%
SVM	86%	84%	83%	83.5%
KNN	82%	80%	78%	79%

Random Forest menghasilkan kinerja terbaik di semua metrik evaluasi. Ini menunjukkan bahwa teknik ensemble efektif dalam menangani data aktivitas mahasiswa yang kompleks dan beragam.

3.2 Analisis Fitur

Fitur yang berkontribusi besar terhadap prediksi keberhasilan:

- 1) Frekuensi login
- 2) Jumlah tugas dikumpulkan tepat waktu
- 3) Aktivitas forum
- 4) Skor kuis

Mahasiswa yang aktif secara konsisten di LMS cenderung memiliki tingkat keberhasilan akademik yang lebih tinggi.

3.3 Implikasi Praktis

Model ini dapat digunakan sebagai bagian dari sistem monitoring akademik yang memungkinkan pengajar atau pihak kampus melakukan intervensi dini terhadap mahasiswa yang berisiko. Selain itu, dapat diintegrasikan dalam sistem LMS untuk memberikan peringatan atau rekomendasi pembelajaran secara otomatis. Model prediksi keberhasilan mahasiswa yang dikembangkan dalam penelitian ini memiliki sejumlah implikasi praktis yang dapat diadopsi oleh institusi pendidikan, baik dalam konteks manajemen pembelajaran maupun pengambilan keputusan strategis, antara lain:

- 1) Pemantauan Progres Belajar Secara Otomatis:

Dengan integrasi model ini ke dalam LMS, sistem dapat memberikan notifikasi otomatis kepada dosen maupun mahasiswa terkait status capaian belajar. Mahasiswa yang menunjukkan penurunan aktivitas atau skor dapat segera mendapatkan perhatian.

- 2) Personalized Learning Path:

Model dapat dikembangkan lebih lanjut untuk menyarankan jalur pembelajaran yang disesuaikan dengan kebutuhan masing-masing mahasiswa berdasarkan riwayat interaksi dan performa dalam LMS.

- 3) Optimalisasi Sumber Daya Pengajar:

Dosen dapat lebih fokus pada mahasiswa yang terindikasi berisiko berdasarkan hasil klasifikasi model, sehingga waktu dan tenaga dapat dialokasikan secara efisien untuk bimbingan yang lebih tepat sasaran.

- 4) Dasar untuk Sistem Rekomendasi:

Hasil dari prediksi ini dapat dimanfaatkan untuk merekomendasikan materi tambahan, forum diskusi, atau penugasan remedial bagi mahasiswa dengan performa rendah.

- 5) Evaluasi dan Perbaikan Kurikulum:

Data dari hasil prediksi juga dapat digunakan oleh pengelola program studi untuk mengevaluasi efektivitas metode pembelajaran dan materi ajar yang tersedia dalam LMS.

- 6) Sistem Early Warning Akademik:

Model ini bisa dijadikan komponen utama dalam sistem peringatan dini (early warning system), di mana mahasiswa yang memiliki risiko gagal dapat langsung dihubungi untuk dilakukan bimbingan, baik secara akademik maupun psikososial.

4. KESIMPULAN

Penelitian ini menghasilkan model prediksi keberhasilan mahasiswa menggunakan algoritma machine learning yang terbukti efektif dalam mengklasifikasikan performa belajar berdasarkan data aktivitas di LMS. Dari empat algoritma yang diuji, *Random Forest* memberikan performa terbaik dengan akurasi mencapai 89%. Penerapan model ini dapat mendukung sistem pembelajaran yang lebih adaptif dan responsif, serta membantu institusi pendidikan dalam meningkatkan kualitas pembelajaran berbasis data. Untuk pengembangan selanjutnya, model ini dapat diintegrasikan secara real-time dalam LMS dan dikombinasikan dengan sistem rekomendasi untuk mendukung personalisasi pembelajaran.

REFERENCES

- Zhang, Y., & Zhang, M. (2018). Predicting student performance using machine learning techniques: A comprehensive review. *International Journal of Educational Data Mining*, 10(1), 45-72.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
- Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3-17.
- Al-Muhtadi, J., & Krentel, M. (2018). Student performance prediction using machine learning models. *International Journal of Engineering and Technology*, 7(4), 23-30.
- Patel, H., & Patel, R. (2019). Predicting student performance using ensemble learning. *International Journal of Advanced Research in Computer Science and Software Engineering*, 9(2), 101-107.
- Nguyen, D., & Nguyen, D. (2017). Applying machine learning to predict student success: A case study. *Journal of Educational Computing Research*, 55(8), 1174-1189.
- Kotsiantis, S. B., & Pintelas, P. E. (2004). Predicting student's performance in distance learning using machine learning techniques. *Proceedings of the 4th International Conference on Knowledge and Data Engineering*, 1-4.
- Chawla, N. V., & Bowyer, K. W. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
- Santos, J. R., & Silva, S. (2017). Application of machine learning algorithms to predict student academic success. *IEEE Access*, 5, 15917-15927.
- Vellido, A., & Nebot, V. (2015). The use of neural networks for student performance prediction. *Journal of Educational Technology & Society*, 18(4), 124-134.