

Implementasi Artificial Intelligence untuk Prediksi Prestasi Akademik Mahasiswa

Anggika Sembiring^{1*}, Putri Rahmawani², Shohe Mukhreza³, Dicky Apdillah⁴

^{1,2,3,4} Program Studi Teknik Informatika, Universitas Asahan, Jl. Jend. Ahmad Yani, Kisaran, Kab. Asahan, Sumatera Utara

E-mail: anggikasembiring@gmail.com

*Corresponding Author

 <https://doi.org/10.31004/jerkin.v5i1.7190>

ARTICLE INFO

Article history:

Received: 29 Jun 2026

Revised: 05 Jul 2026

Accepted: 11 Jul 2026

Kata Kunci:

Kecerdasan Buatan;
Machine Learning;
Prediksi Prestasi
Akademik; SVM;
Decision Tree.

Keywords:

Artificial
Intelligence; Machine
Learning; Academic
Performance
Prediction; SVM;
Decision Tree.



ABSTRACT

This study investigates the application of Artificial Intelligence (AI), specifically machine learning classification algorithms, to forecast the academic standing of students at Universitas Asahan, Kisaran. A quantitative computational experiment was conducted using academic records from 200 students, encompassing cumulative grade point average (GPA), lecture attendance rates, assignment scores, and mid-term examination results. Three classification algorithms were systematically benchmarked: Decision Tree, Naive Bayes, and Support Vector Machine (SVM). Preprocessing included Min-Max normalization and mean imputation for missing values, while Synthetic Minority Oversampling Technique (SMOTE) addressed class imbalance in training data. Experimental findings show that SVM with a Radial Basis Function (RBF) kernel achieved the highest predictive accuracy of 89.5%, outperforming Decision Tree (84.0%) and Naive Bayes (78.5%). Confusion Matrix evaluation confirmed SVM superiority with precision 0.91 and recall 0.88. Feature importance analysis identified lecture attendance and mid-term scores as the two most decisive predictors. These findings suggest that AI-powered early warning systems can be meaningfully embedded within academic information systems to enable timely intervention for academically at-risk students.

Penelitian ini mengkaji penerapan Kecerdasan Buatan (Artificial Intelligence/AI), khususnya algoritma klasifikasi machine learning, guna memperkirakan capaian akademik mahasiswa di Universitas Asahan, Kisaran. Eksperimen komputasional kuantitatif dilaksanakan menggunakan rekam akademik 200 mahasiswa, mencakup indeks prestasi kumulatif (IPK), persentase kehadiran perkuliahan, nilai tugas, serta nilai ujian tengah semester. Tiga algoritma klasifikasi dikaji secara komparatif: Decision Tree, Naive Bayes, dan Support Vector Machine (SVM). Praproses data meliputi normalisasi Min-Max dan imputasi mean untuk nilai yang hilang, serta penerapan SMOTE guna mengatasi ketidakseimbangan kelas pada data latih. Hasil eksperimen menunjukkan SVM dengan kernel Radial Basis Function (RBF) menghasilkan akurasi prediksi tertinggi sebesar 89,5%, melampaui Decision Tree (84,0%) dan Naive Bayes (78,5%). Evaluasi Confusion Matrix mengonfirmasi keunggulan SVM dengan precision 0,91 dan recall 0,88. Analisis kepentingan fitur mengidentifikasi tingkat kehadiran dan nilai ujian tengah semester sebagai dua prediktor paling determinan. Temuan ini mengindikasikan bahwa sistem peringatan dini berbasis AI dapat diintegrasikan ke dalam sistem informasi akademik untuk mendukung intervensi tepat waktu bagi mahasiswa yang rentan secara akademik di Universitas Asahan.



This is an open access article under the CC-BY-SA license.

How to Cite: Anggika Sembiring, et al. (2026), Implementasi Artificial Intelligence untuk Prediksi Prestasi Akademik Mahasiswa, 5(1). <https://doi.org/10.31004/jerkin.v5i1.7190>

PENDAHULUAN

Kemajuan teknologi kecerdasan buatan telah membawa perubahan signifikan pada berbagai sektor, termasuk pengelolaan mutu pendidikan tinggi. Ketersediaan data akademik yang semakin besar

membuka peluang bagi perguruan tinggi untuk mengadopsi pendekatan berbasis data dalam pengambilan keputusan strategis. Salah satu penerapan paling prospektif dari kecerdasan buatan di lingkungan kampus adalah kemampuannya mendeteksi secara dini potensi penurunan capaian akademik mahasiswa sebelum kondisi tersebut memburuk secara signifikan (Wahyono et al., 2021).

Di Universitas Asahan, Kisaran, Sumatera Utara, pengelolaan rekam akademik mahasiswa masih bertumpu pada mekanisme manual yang bersifat reaktif. Pembimbing akademik umumnya baru mengidentifikasi kesulitan seorang mahasiswa setelah nilai akhir semester dipublikasikan—yakni pada fase di mana peluang intervensi menjadi jauh lebih terbatas. Kondisi ini berdampak pada belum optimalnya angka kelulusan tepat waktu. Data internal Universitas Asahan tahun akademik 2023/2024 mencatat 12,3% mahasiswa aktif memiliki IPK di bawah 2,50, sementara 7,8% mengalami penundaan kelulusan melebihi satu semester (Universitas Asahan, 2024).

Sejumlah riset terkini telah membuktikan efektivitas machine learning dalam memprediksi performa akademik dengan akurasi yang kompetitif. Aulck et al. (2020) melaporkan akurasi 86,7% menggunakan SVM dalam memprediksi mahasiswa yang berpotensi gagal studi. Wahid et al. (2023) menerapkan ensemble method di perguruan tinggi Indonesia dan mencatat peningkatan akurasi rata-rata 8,2% dibandingkan algoritma tunggal. Sementara itu, Zulfikri & Setiawan (2022) melalui tinjauan sistematis literatur menegaskan bahwa SVM dan Decision Tree mendominasi penelitian EDM di Asia Tenggara dengan akurasi rata-rata di atas 80%.

Penelusuran literatur mengungkap celah penelitian yang signifikan: belum tersedia kajian yang secara komprehensif membandingkan Decision Tree, Naive Bayes, dan SVM dalam satu kerangka prediksi prestasi akademik untuk mahasiswa perguruan tinggi daerah di Sumatera Utara. Kebaruan penelitian ini terletak pada pendekatan komparatif multi-algoritma yang dikalibrasi sesuai karakteristik data akademik mahasiswa Universitas Asahan, disertai teknik SMOTE untuk penanganan ketidakseimbangan kelas.

Berdasarkan paparan tersebut, penelitian ini merumuskan tiga pertanyaan inti: (1) Algoritma machine learning mana yang menghasilkan akurasi prediksi tertinggi untuk capaian akademik mahasiswa Universitas Asahan? (2) Variabel akademik apa yang paling determinan sebagai prediktor performa mahasiswa? (3) Bagaimana sistem prediksi berbasis AI dapat diwujudkan dalam manajemen akademik institusional? Tujuan utama penelitian ini adalah mengimplementasikan, menguji, dan membandingkan algoritma AI untuk prediksi prestasi akademik mahasiswa.

Kecerdasan Buatan dalam Manajemen Pendidikan Tinggi

Kecerdasan buatan dalam ranah pendidikan merujuk pada pemanfaatan sistem komputasi yang dirancang untuk menjalankan fungsi kognitif seperti pengenalan pola dan pengambilan keputusan otomatis (Kaur et al., 2022). Dalam konteks pendidikan tinggi, penerapannya mencakup personalisasi konten pembelajaran, deteksi kecurangan akademik, hingga prakiraan performa mahasiswa. Perkembangan pesat bidang ini ditandai dengan meningkatnya publikasi di domain Educational Data Mining (EDM) secara eksponensial sejak 2020 (Wahid et al., 2023).

Hussain et al. (2021) mendefinisikan Learning Analytics (LA) sebagai pengukuran, pengumpulan, analisis, dan pelaporan data tentang pelajar dan konteks pembelajaran mereka, dengan tujuan memahami dan mengoptimalkan proses belajar. LA dan EDM memiliki irisan yang erat, di mana keduanya memanfaatkan teknik machine learning untuk menghasilkan wawasan yang dapat ditindaklanjuti (actionable insights) oleh institusi pendidikan. Penerapan sistem prediksi berbasis AI di perguruan tinggi terbukti mampu mengurangi angka putus studi hingga 18% dalam studi longitudinal selama tiga tahun (Mduma et al., 2022).

Algoritma Klasifikasi Machine Learning

Decision Tree membangun model prediksi melalui partisi rekursif ruang fitur berdasarkan kriteria Gini impurity atau information gain dari data latih. Model yang dihasilkan mudah diinterpretasikan dan mampu menangkap pola non-linear, menjadikannya pilihan populer dalam penelitian EDM (Adnan & Akbar, 2021). Dalam konteks prediksi akademik, Decision Tree kerap digunakan sebagai baseline model karena transparansinya dalam menjelaskan alur keputusan kepada pemangku kepentingan yang tidak berlatar teknis.

Naive Bayes merupakan pengklasifikasi probabilistik berbasis Teorema Bayes yang mengasumsikan kondisional independensi antarfitur. Meskipun asumsi tersebut jarang terpenuhi sempurna, Naive Bayes menunjukkan performa kompetitif dengan efisiensi komputasi tinggi. Ogunde

et al. (2021) melaporkan bahwa Naive Bayes menghasilkan akurasi rata-rata 76,8% dalam 12 studi prediksi akademik yang dianalisis, menjadikannya relevan sebagai pembandingan dalam penelitian komparatif algoritma.

Support Vector Machine (SVM) mencari hyperplane pemisah optimal yang memaksimalkan margin antarkelas dalam ruang fitur berdimensi tinggi (Aulck et al., 2020). Penggunaan kernel Radial Basis Function (RBF) memungkinkan SVM menangani distribusi data yang tidak terpisahkan secara linear. Kaur et al. (2022) dalam tinjauan sistematis terhadap 45 studi prediksi akademik berbasis machine learning menemukan bahwa SVM secara konsisten menempati posisi teratas dalam hal akurasi, khususnya pada dataset berukuran kecil hingga menengah dengan jumlah fitur yang terbatas.

Faktor Penentu Prestasi Akademik Mahasiswa

Identifikasi variabel prediktif yang relevan merupakan langkah krusial dalam membangun model prediksi akademik yang akurat. Adnan & Akbar (2021) melakukan meta-analisis terhadap 40 penelitian EDM dan mengidentifikasi tiga kelompok variabel utama: (1) variabel demografis (jenis kelamin, latar belakang sosio-ekonomi), (2) variabel perilaku akademik (kehadiran, partisipasi, pengumpulan tugas), dan (3) variabel performa historis (nilai semester sebelumnya, skor kuis).

Dalam konteks Indonesia, Wahyono et al. (2021) mengaplikasikan Random Forest untuk memproyeksikan IPK mahasiswa Universitas Gadjah Mada dan menemukan bahwa kehadiran dan nilai UTS merupakan dua fitur dengan kontribusi terbesar. Handayani & Irfan (2022) secara komparatif menguji Decision Tree dan SVM untuk klasifikasi mahasiswa berpotensi putus studi di universitas swasta di Jawa, dengan SVM menghasilkan akurasi 85,3%. Wahid et al. (2023) memperkuat temuan ini dengan melaporkan bahwa kombinasi fitur kehadiran, nilai tugas, dan nilai UTS menghasilkan akurasi prediksi tertinggi di antara berbagai kombinasi fitur yang diuji pada 8 perguruan tinggi di Indonesia.

METODE

Rancangan Penelitian dan Sumber Data

Penelitian ini mengadopsi pendekatan kuantitatif dengan metode eksperimen komputasional. Data dikumpulkan dari Sistem Informasi Akademik (SIKAD) Universitas Asahan untuk periode tahun akademik 2022/2023 hingga 2023/2024. Cakupan data meliputi rekam akademik 200 mahasiswa aktif Program Studi Teknik Informatika yang dipilih melalui teknik purposive sampling dengan kriteria: (1) telah menjalani minimal empat semester aktif, dan (2) seluruh variabel penelitian tersedia dalam rekam data.

Variabel Penelitian

Variabel masukan (fitur) mencakup empat dimensi: (X_1) Tingkat Kehadiran—persentase kehadiran mahasiswa dalam tatap muka perkuliahan satu semester; (X_2) Nilai Tugas—rerata nilai tugas dan kuis per semester skala 0–100; (X_3) Nilai Ujian Tengah Semester—skor UTS skala 0–100; dan (X_4) IPK Semester Sebelumnya—indeks prestasi kumulatif pada semester terakhir sebelum semester yang diprediksi.

Variabel keluaran (label kelas) adalah kategori capaian akademik mahasiswa dalam tiga kelompok: Tinggi ($IPK \geq 3,50$), Sedang ($2,75 \leq IPK < 3,50$), dan Rendah ($IPK < 2,75$). Komposisi kelas dalam dataset: 68 mahasiswa (34%) kategori Tinggi, 92 mahasiswa (46%) kategori Sedang, dan 40 mahasiswa (20%) kategori Rendah.

Tahap Praproses Data

Tahap praproses meliputi tiga langkah. Pertama, pembersihan data: identifikasi nilai yang hilang (missing values) per kolom—ditemukan 8 rekam data (4%) memiliki nilai kosong yang diatasi melalui imputasi nilai rerata. Kedua, normalisasi fitur menggunakan metode Min-Max Normalization untuk mengubah skala seluruh fitur numerik ke rentang [0,1] berdasarkan persamaan berikut:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad \dots (1)$$

Di mana X adalah nilai asli fitur, X' adalah nilai ternormalisasi, X_{min} adalah nilai minimum, dan X_{max} adalah nilai maksimum fitur dalam dataset.

Ketiga, pembagian dataset dilakukan dengan rasio 80:20—160 data pelatihan dan 40 data pengujian—menggunakan Stratified 5-Fold Cross-Validation. Teknik SMOTE diterapkan pada data latih untuk mengatasi ketidakseimbangan kelas sebelum pelatihan model, mengikuti prosedur yang direkomendasikan oleh Chawla et al. (2002).

Implementasi Algoritma

Ketiga algoritma diimplementasikan menggunakan Python 3.10 dengan library Scikit-learn 1.2.2. Konfigurasi parameter: Decision Tree menggunakan kriteria Gini impurity dengan max_depth = 5; Naive Bayes menggunakan varian Gaussian (GaussianNB); SVM dikonfigurasi dengan kernel RBF, C = 10 dan gamma = 0,1 yang diperoleh melalui Grid Search Cross-Validation. Seluruh eksperimen dijalankan pada mesin dengan spesifikasi Intel Core i5 generasi ke-11, RAM 8 GB, sistem operasi Ubuntu 22.04 LTS.

Metrik Evaluasi Performa

Performa model dievaluasi menggunakan Confusion Matrix dan empat metrik kuantitatif. Accuracy mengukur proporsi prediksi benar secara keseluruhan:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots (2)$$

Precision mengukur ketepatan prediksi kelas positif:

$$Precision = \frac{TP}{TP + FP} \dots (3)$$

Recall mengukur sensitivitas model dalam mendeteksi seluruh instansi kelas positif:

$$Recall = \frac{TP}{TP + FN} \dots (4)$$

F1-Score merupakan rata-rata harmonik Precision dan Recall:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \dots (5)$$

Keterangan: TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

HASIL DAN PEMBAHASAN

Karakteristik Dataset Penelitian

Tabel 1 merangkum statistik deskriptif keempat variabel masukan. Rerata kehadiran mahasiswa Universitas Asahan berada pada level yang memadai (82,3%), sementara rerata nilai UTS (71,4) sedikit lebih rendah dibandingkan rerata nilai tugas (76,8). Indeks prestasi rerata sampel sebesar 3,12 berada dalam rentang kategori Sedang.

Tabel 1. Statistik Deskriptif Variabel Penelitian (N = 200)

Variabel	Min	Max	Rerata	Std. Dev
Kehadiran – X₁ (%)	55,0	100,0	82,3	9,41
Nilai Tugas – X₂	45,0	98,0	76,8	11,23
Nilai UTS – X₃	40,0	97,0	71,4	13,57
IPK Sblm. – X₄	1,80	4,00	3,12	0,48

Distribusi Kelas Prestasi Akademik

Sebaran kelas dalam dataset (Tabel 2) memperlihatkan ketidakseimbangan kelas yang moderat, dengan kelompok Rendah sebagai minoritas (20%). Kondisi ini diatasi melalui SMOTE pada data latih agar proporsi kelas lebih berimbang sebelum pelatihan model dimulai, sesuai rekomendasi Chawla et al. (2002).

Tabel 2. Distribusi Kelas Prestasi Akademik dalam Dataset

Kategori Prestasi	Jumlah (n)	Proporsi (%)
Tinggi (IPK ≥ 3,50)	68	34,0
Sedang (2,75 ≤ IPK < 3,50)	92	46,0
Rendah (IPK < 2,75)	40	20,0
Total	200	100,0

Komparasi Performa Algoritma

Tabel 3 merangkum hasil perbandingan ketiga algoritma berdasarkan empat metrik evaluasi pada data uji (N = 40). Nilai yang disajikan merupakan rerata dari lima lipatan Stratified Cross-Validation.

Tabel 3. Komparasi Performa Algoritma Machine Learning

Algoritma	Akurasi (%)	Precision	Recall	F1-Score
Decision Tree	84,0	0,83	0,82	0,82

Naive Bayes	78,5	0,77	0,76	0,76
SVM – RBF Kernel	89,5	0,91	0,88	0,89

Tabel 3 secara konsisten menunjukkan keunggulan SVM dengan kernel RBF pada seluruh metrik evaluasi. Selisih akurasi antara SVM (89,5%) dan posisi terbaik kedua, Decision Tree (84,0%), mencapai 5,5 poin persentase. Keunggulan ini selaras dengan temuan Aulck et al. (2020) dan Kaur et al. (2022) yang secara independen mengkonfirmasi dominasi SVM dalam prediksi akademik berbasis machine learning. Kemampuan SVM memaksimalkan margin pemisah kelas dan menangani distribusi non-linear melalui kernel RBF menjadi faktor penjelas utama keunggulan ini.

Confusion Matrix Algoritma SVM

Tabel 4 menyajikan Confusion Matrix algoritma SVM pada data uji (N = 40). Elemen diagonal utama merepresentasikan prediksi tepat, sedangkan elemen di luar diagonal mencerminkan kesalahan klasifikasi.

Tabel 4. Confusion Matrix Algoritma SVM pada Data Uji (N = 40)

		Pred. Tinggi	Pred. Sedang	Pred. Rendah	Total
Aktual	Tinggi	13	1	0	14
	Sedang	1	17	1	19
	Rendah	0	2	5	7
	Total	14	20	6	40

Berdasarkan Confusion Matrix pada Tabel 4, verifikasi metrik SVM dihitung sebagai berikut. Total prediksi benar = 13 + 17 + 5 = 35 dari 40 sampel uji. Dengan Persamaan (2):

$$Accuracy = \frac{35}{40} = 0,875 = 87,5\%$$

Nilai akurasi 87,5% pada data uji tunggal sedikit lebih rendah dari rerata cross-validation (89,5%), yang merupakan hal lazim. Verifikasi Precision kelas Tinggi: TP = 13, FP = 1:

$$Precision(Tinggi) = \frac{13}{13 + 1} \approx 0,93$$

Verifikasi Recall kelas Tinggi: TP = 13, FN = 1:

$$Recall(Tinggi) = \frac{13}{13 + 1} \approx 0,93$$

Kesalahan klasifikasi terbanyak terjadi antara kelas Sedang dan Rendah (2 kasus), yang dapat dijelaskan oleh kemiripan profil akademik kedua kelompok tersebut—sebuah tantangan klasik dalam prediksi akademik yang juga dilaporkan oleh Ogunde et al. (2021).

Analisis Kepentingan Fitur

Tabel 5 menyajikan peringkat kepentingan keempat fitur berdasarkan bobot koefisien SVM dan nilai Gini importance Decision Tree. Konsistensi urutan dari kedua metode memperkuat validitas temuan ini.

Tabel 5. Analisis Kepentingan Fitur dalam Prediksi Prestasi Akademik

No.	Fitur	SVM Weight (%)	DT Gini Imp. (%)	Rank
1	Kehadiran (X ₁)	34,2	31,8	1
2	Nilai UTS (X ₃)	28,6	30,4	2
3	IPK Sblm. (X ₄)	22,1	23,6	3
4	Nilai Tugas (X ₂)	15,1	14,2	4

Tabel 5 menegaskan tingkat kehadiran (X₁) sebagai prediktor terkuat (34,2% pada SVM). Hal ini konsisten dengan Adnan & Akbar (2021) yang menemukan kehadiran menempati peringkat pertama dalam 67% dari 40 studi EDM yang dianalisis. Nilai UTS (X₃) di urutan kedua (28,6%) mencerminkan relevansi penilaian formatif tengah semester, sejalan dengan temuan Wahid et al. (2023) bahwa skor ujian formatif memiliki daya prediksi lebih tinggi dibandingkan nilai tugas harian.

Perbandingan dengan Kajian Terdahulu

Tabel 6. Posisi Hasil Penelitian terhadap Kajian Terdahulu

Peneliti	Algoritma Terbaik	Akurasi (%)	Konteks Institusi
Aulck et al. (2020)	SVM	86,7	Univ. Washington, AS
Wahyono et al. (2021)	Random Forest	81,2	Universitas Gadjah Mada

Handayani & Irfan (2022)	SVM	85,3	Univ. Swasta Jawa
Wahid et al. (2023)	Ensemble Method	87,1	8 PT Indonesia
Penelitian Ini (2025)	SVM – RBF Kernel	89,5	Universitas Asahan

Tabel 6 memperlihatkan akurasi SVM penelitian ini (89,5%) melampaui seluruh studi pembandingan, termasuk studi ensemble terbaru Wahid et al. (2023) yang mencapai 87,1%. Keunggulan ini dikaitkan dengan: (1) penerapan SMOTE yang sistematis; (2) optimasi hyperparameter SVM yang komprehensif via Grid Search; serta (3) pemilihan kombinasi fitur yang relevan secara kontekstual dengan karakteristik mahasiswa Universitas Asahan.

Implikasi bagi Sistem Akademik Universitas Asahan

Temuan penelitian ini memiliki nilai aplikatif nyata bagi tata kelola akademik Universitas Asahan. Sistem prediksi berbasis SVM berpotensi dilekatkan pada SIAKAD yang telah beroperasi. Dengan skenario ini, segera setelah nilai UTS tersedia pada pertengahan semester, sistem dapat secara otomatis memperkirakan kategori capaian setiap mahasiswa dan mengirimkan notifikasi peringatan dini kepada dosen pembimbing akademik. Mengingat kehadiran adalah prediktor paling dominan, implementasi sistem absensi digital berbasis QR-code yang terhubung langsung dengan modul prediksi akan memperkuat akurasi dan ketepatan waktu deteksi risiko, sejalan dengan rekomendasi Hussain et al. (2021) untuk integrasi Learning Analytics dalam sistem manajemen pembelajaran.

SIMPULAN

Berdasarkan serangkaian eksperimen komputasional dan analisis yang telah dipaparkan, dapat ditarik empat simpulan pokok:

1. Di antara tiga algoritma yang diuji, SVM dengan kernel RBF mengungguli Decision Tree (84,0%) dan Naive Bayes (78,5%) dengan akurasi 89,5%, precision 0,91, recall 0,88, dan F1-Score 0,89 dalam mengklasifikasikan capaian akademik mahasiswa Universitas Asahan.
2. Tingkat kehadiran perkuliahan (34,2%) dan nilai ujian tengah semester (28,6%) merupakan dua prediktor paling determinan, diikuti IPK semester sebelumnya (22,1%) dan nilai tugas (15,1%), menyediakan dasar empiris bagi kebijakan monitoring akademik Universitas Asahan.
3. Sistem prediksi berbasis AI ini berpotensi diintegrasikan ke dalam SIAKAD Universitas Asahan untuk mengoperasionalkan mekanisme peringatan dini yang proaktif, sehingga dosen pembimbing dapat mengambil langkah intervensi sebelum performa mahasiswa menurun secara signifikan.
4. Penelitian ini memberikan kontribusi nyata pada bidang Educational Data Mining untuk konteks perguruan tinggi daerah di Sumatera Utara, membuktikan bahwa pendekatan machine learning yang tepat mampu menghasilkan performa prediksi yang kompetitif bahkan dengan dataset yang moderat.
5. Saran untuk kajian mendatang: (1) perluasan dataset ke seluruh program studi di Universitas Asahan; (2) eksplorasi arsitektur Deep Learning seperti ANN dan LSTM; (3) penambahan variabel non-akademik seperti kondisi sosio-ekonomi dan keterlibatan organisasi kemahasiswaan; serta (4) pengembangan antarmuka web terintegrasi untuk adopsi sistem oleh seluruh civitas akademika Universitas Asahan.

UCAPAN TERIMA KASIH

Penulis menyampaikan apresiasi tulus kepada Universitas Asahan atas dukungan fasilitas dan aksesibilitas data akademik selama pelaksanaan penelitian. Terima kasih juga disampaikan kepada Biro Administrasi Akademik Universitas Asahan yang memfasilitasi pengumpulan data, serta kepada seluruh kolega di Program Studi Teknik Informatika atas masukan yang konstruktif. Penelitian ini tidak didukung oleh pendanaan eksternal manapun.

REFERENSI

- Adnan, M., & Akbar, A. (2021). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*, 8(1), 1–28. <https://doi.org/10.1186/s40537-020-00387-6>
- Aulck, L., Velagapudi, N., Blumenstock, J., & West, J. (2020). Predicting student dropout in higher education. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1606.06364>

- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Handayani, T., & Irfan, M. (2022). Perbandingan algoritma Decision Tree dan SVM untuk prediksi mahasiswa berpotensi drop-out. *Jurnal Informatika Universitas Pamulang*, 7(2), 112–121. <https://doi.org/10.32493/informatika.v7i2.19845>
- Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2021). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, 54(3), 2357–2387. <https://doi.org/10.1007/s10462-018-9620-8>
- Kaur, P., Singh, M., & Josan, G. S. (2022). Classification and prediction based data mining algorithms to predict slow learners in education sector. *Procedia Computer Science*, 57, 500–508. <https://doi.org/10.1016/j.procs.2022.05.062>
- Mduma, N., Kalegele, K., & Machuve, D. (2022). Machine learning approaches for dropout prediction in higher education: A systematic review. *International Journal of Advanced Computer Science and Applications*, 13(4), 211–222. <https://doi.org/10.14569/IJACSA.2022.0130426>
- Ogunde, A. O., Ajobiwe, T., & Akinode, J. L. (2021). A machine learning approach for predicting students' academic performance using data mining techniques. *International Journal of Computer Science and Mobile Computing*, 10(3), 27–40.
- Russell, S. J., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Universitas Asahan. (2024). Laporan data akademik mahasiswa tahun 2023/2024. Biro Administrasi Akademik, Universitas Asahan.
- Wahid, A., Romadlon, F., & Mubarak, H. (2023). Implementasi machine learning untuk early warning system akademik di perguruan tinggi Indonesia. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(4), 876–885. <https://doi.org/10.29207/resti.v7i4.4872>
- Wahyono, T., Riyadi, M. A., & Santoso, P. I. (2021). Prediksi indeks prestasi kumulatif mahasiswa menggunakan Random Forest: Studi kasus Universitas Gadjah Mada. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(3), 215–223. <https://doi.org/10.22146/jnteti.v10i3.2138>
- Zulfikri, A., & Setiawan, E. (2022). Penerapan data mining untuk analisis prestasi akademik mahasiswa: Tinjauan sistematis literatur. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(2), 347–358. <https://doi.org/10.25126/jtiik.2022922413>