

Pengembangan Model Performa Akademik Mahasiswa Dengan Early Warning System (EWS) Berbasis Data LMS

Dosen Pembimbing :

M. Faris Al Hakim, S.Pd., M.Cs.



Disusun Oleh :

Muhammad Inzaghi Rizqullah	(2404130020)
Shofwan Rafiurachman	(2404130038)
Fathan Nabil Rahman	(2404130140)

DIVISI PENAMBANGAN DATA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI SEMARANG

TAHUN 2025/2026

Pengembangan Model Performa Akademik Mahasiswa Dengan Early Warning System (EWS) Berbasis Data LMS

Muhammad Inzaghi Rizqullah¹, Shofwan Rafiurachman², Fathan Nabil Rahman³

^{1,2,3}Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Negeri Semarang

Abstrak

Kegagalan akademik dan *drop out* mahasiswa masih menjadi tantangan serius di pendidikan tinggi. Metode identifikasi risiko yang konvensional umumnya bersifat reaktif, sehingga intervensi sering terlambat dilakukan. Penelitian ini bertujuan mengembangkan dan mengevaluasi Early Warning System (EWS) berbasis Learning Analytics untuk memprediksi risiko kegagalan studi (*Fail/Withdrawn*) mahasiswa secara proaktif dalam 30 hari pertama perkuliahan. Menggunakan dataset Open University Learning Analytics Dataset (OULAD) dengan jumlah sampel 32.593 mahasiswa, model prediksi dikembangkan menggunakan Extreme Gradient Boosting (XGBoost) yang memanfaatkan 22 fitur gabungan dari data aktivitas Learning Management System (LMS) awal—meliputi skor awal, total klik, durasi partisipasi, dan frekuensi hari aktif—serta data demografis. Model dioptimasi menggunakan GridSearchCV dan dibandingkan dengan model pembanding Random Forest (RF). Hasil evaluasi menunjukkan bahwa model XGBoost memberikan performa terbaik dengan akurasi 71,04%, AUC 0,71, dan F1-Score 0,71, sedikit lebih unggul dari RF (akurasi 70,23%). Analisis interpretabilitas menggunakan SHAP (SHapley Additive exPlanations) mengidentifikasi fitur *avg_score_early* (rata-rata skor awal), *durasi_partisipasi_early* (durasi partisipasi awal), dan *total_hari_aktif_early* (jumlah hari aktif awal) sebagai prediktor paling berpengaruh terhadap risiko kegagalan studi. Temuan ini menegaskan bahwa prediksi risiko akademik sangat dini—bahkan dalam 30 hari pertama—dapat dicapai secara akurat menggunakan data LMS dan model berbasis XGBoost. EWS yang dihasilkan bersifat prediktif sekaligus interpretatif, memberikan landasan kuat bagi institusi pendidikan untuk menerapkan strategi intervensi dini yang lebih efektif.

Kata Kunci: Early Warning System, Learning Analytics, Prediksi Kinerja Akademik, XGBoost, SHAP, Data LMS, OULAD.

Daftar Isi

BAB 1.....	4
PENDAHULUAN.....	4
1.1 Latar Belakang.....	4
1.2 Tujuan Penelitian.....	5
1.3 Manfaat Penelitian.....	6
1.4 Batasan Masalah.....	7
BAB 2.....	9
KAJIAN TERKAIT.....	9
2.1 Pendahuluan.....	9
2.2 Tinjauan Penelitian Terkait EWS Berbasis Data LMS.....	9
2.3 Posisi Penelitian.....	10
BAB 3.....	11
METODE PENELITIAN.....	11
3.1 Deskripsi Solusi Usulan.....	11
3.2 Dataset.....	11
3.3 Metode yang digunakan.....	13
3.3.1 Arsitektur Model.....	13
3.3.2 Penanganan Ketidakseimbangan Data.....	14
3.3.3 Interpretabilitas Model.....	14
3.4 Perbedaan dengan solusi sebelumnya.....	15
3.5 Metrik Evaluasi.....	15
BAB 4.....	17
HASIL DAN PEMBAHASAN.....	17
4.1 Deskripsi Data dan Hasil Pra-pemrosesan.....	17
4.2 Hasil Optimasi Model (Hyperparameter Tuning).....	18
4.3 Evaluasi Kinerja Model.....	18
4.4 Pembahasan dan Analisis Interpretabilitas (SHAP).....	20
4.5 Analisis Hasil Prediksi, Kesalahan Model, dan Implikasi Intervensi.....	23
BAB 5.....	27
KESIMPULAN DAN SARAN.....	27
5.1 Kesimpulan.....	27
5.2 Saran.....	27
Daftar Pustaka.....	29

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Keberhasilan mahasiswa dalam menyelesaikan studi merupakan salah satu parameter esensial dalam evaluasi kualitas institusi pendidikan tinggi. Akan tetapi, fenomena kegagalan akademik serta pengunduran diri mahasiswa (drop out) masih menjadi permasalahan signifikan yang dihadapi oleh banyak perguruan tinggi secara global, termasuk di Indonesia. Berdasarkan data tahun 2020, tercatat **602.208 mahasiswa** di Indonesia mengundurkan diri dari total 8,4 juta mahasiswa terdaftar, dengan persentase tertinggi (79,5%) terjadi di perguruan tinggi swasta. [6] Kondisi ini menimbulkan kerugian bagi mahasiswa—dari sisi waktu dan finansial—serta berdampak pada reputasi dan efisiensi sumber daya institusi. Oleh karena itu, pengembangan mekanisme proaktif untuk identifikasi dini mahasiswa yang berpotensi mengalami kesulitan akademik menjadi kebutuhan mendesak. [4]

Seiring kemajuan teknologi informasi di sektor pendidikan, pemanfaatan Learning Management System (LMS) atau Virtual Learning Environment (VLE) telah menjadi praktik standar. Platform digital ini secara otomatis merekam berbagai jejak aktivitas pembelajaran mahasiswa, seperti frekuensi akses sistem, interaksi dengan materi, partisipasi forum diskusi, waktu pengumpulan tugas, dan skor kuis awal. Data log LMS menyajikan sumber informasi ekstensif yang berpotensi dianalisis untuk mendeteksi pola perilaku belajar dan memprediksi kinerja akademik pada lingkungan daring maupun blended learning. [3][5]

Metode konvensional untuk mengidentifikasi mahasiswa berisiko—misalnya pemantauan Indeks Prestasi Kumulatif (IPK) pada akhir semester atau penilaian subjektif dosen wali—cenderung bersifat reaktif dan sering terlambat memberikan peringatan. Ketika risiko baru terdeteksi melalui IPK, permasalahan akademik sering sudah mengakar sehingga intervensi menjadi kurang efektif. Hal ini menegaskan pentingnya pengembangan Early Warning System (EWS) berbasis data yang mampu mendeteksi tanda awal risiko sehingga memungkinkan intervensi proaktif yang lebih tepat sasaran. Studi sebelumnya menunjukkan efektivitas pendekatan berbasis aktivitas online untuk prediksi awal performa mahasiswa. [1][5]

Berbagai penelitian juga mengindikasikan bahwa fitur aktivitas awal di LMS—seperti rerata skor tugas/kuis awal (avg_score_early), jumlah hari keaktifan (total_hari_aktif_early), total interaksi klik (total_clicks_early), serta durasi partisipasi pada periode awal perkuliahan

(durasi_partisipasi_early)—memiliki nilai prediktif terhadap capaian akademik akhir. Selain itu, penggabungan fitur perilaku (learning analytics) dengan variabel demografis umumnya meningkatkan kemampuan prediktif model dalam mengidentifikasi mahasiswa berisiko. [3][5]

Dari sisi algoritmik, penerapan teknik machine learning—khususnya metode ensemble dan boosting—telah menunjukkan hasil menjanjikan untuk klasifikasi risiko akademik dan penanganan pola non-linear yang sulit ditangkap metode tradisional. Selain performa prediksi, interpretabilitas model (mis. teknik SHAP) penting agar pemangku kepentingan memahami faktor penentu risiko dan merancang intervensi yang sesuai. [2][4]

Meskipun demikian, kajian yang secara khusus menargetkan **prediksi sangat dini**—yaitu dalam **30 hari pertama** masa studi—dengan memanfaatkan kombinasi data clickstream awal dan data demografis, serta menekankan interpretasi faktor risiko, masih relatif terbatas. Oleh karena itu, penelitian ini bertujuan merancang dan mengevaluasi sebuah EWS berbasis XGBoost menggunakan data aktivitas 30 hari pertama dari LMS serta data demografis mahasiswa. Penelitian juga akan mengidentifikasi fitur-fitur prediktor paling signifikan dan mengusulkan rekomendasi intervensi berbasis tingkat risiko. Hasil penelitian diharapkan menjadi landasan empiris bagi institusi pendidikan dalam menyusun strategi intervensi dini yang lebih efektif dan berbasis bukti.

1.2 Tujuan Penelitian

1. Merancang dan mengembangkan model *Early Warning System (EWS)* berbasis algoritma XGBoost untuk memprediksi status kelulusan mahasiswa (Lulus vs. Putus Kuliah) sebagai langkah preventif terhadap risiko keterlambatan studi dan kegagalan akademik.
2. Mengintegrasikan data aktivitas mahasiswa selama 30 hari pertama perkuliahan yang terekam dalam *Learning Management System (LMS)* dengan data demografis sebagai fitur prediktif utama yang merepresentasikan pola perilaku belajar awal.
3. Mengevaluasi performa model secara komprehensif menggunakan metrik akurasi, presisi, recall, F1-score, dan AUC, serta membandingkannya dengan *baseline models* untuk memastikan keunggulan model yang dikembangkan.
4. Menganalisis faktor-faktor prediktor paling berpengaruh terhadap status kelulusan mahasiswa dengan memanfaatkan teknik interpretasi model SHAP (SHapley Additive exPlanations) guna memperoleh pemahaman yang lebih mendalam

5. Merumuskan rekomendasi strategis bagi pihak kampus atau dosen pembimbing akademik berdasarkan hasil analisis model, sehingga EWS dapat dimanfaatkan sebagai dasar intervensi dini terhadap mahasiswa berisiko.

1.3 Manfaat Penelitian

Penelitian ini mengembangkan sebuah *Early Warning System* (EWS) berbasis **XGBoost** yang secara **inovatif** mampu memprediksi risiko **Putus Kuliah** (*Fail/Withdrawn*) mahasiswa pada **30 hari pertama** perkuliahan. Model ini menunjukkan **kinerja solid** (**AUC=0.71, F1=0.71**) pada data uji (N=32.593) dan **terbukti mengungguli *baseline*** Random Forest. Kontribusi utama EWS ini adalah **kemampuan prediksi sangat dini** untuk intervensi proaktif serta **interpretabilitas** melalui **SHAP** yang menyediakan dasar justifikasi rekomendasi pedagogis.

1. Untuk Institusi Pendidikan / Akademik:

- **Deteksi Risiko Dini & Optimalisasi Sumber Daya:** EWS memungkinkan identifikasi mahasiswa berisiko **pada 30 hari pertama** perkuliahan, menyediakan dasar (*insight* dari kinerja model & SHAP) untuk **intervensi preventif** dan **optimalisasi alokasi program dukungan** (remedial, bimbingan) sebelum terjadi eskalasi masalah akademik.
- **Potensi Peningkatan Metrik Kinerja:** Implementasi EWS berpotensi **menurunkan angka *dropout*** dan **meningkatkan persentase kelulusan tepat waktu**, berkontribusi positif pada akreditasi dan reputasi.

2. Untuk Dosen Wali / Pembimbing Akademik:

- **Prioritisasi Bimbingan Efektif:** EWS menyediakan **daftar prioritas mahasiswa** berbasis probabilitas risiko, memungkinkan **fokus intervensi** pada kasus paling mendesak (dapat divisualisasikan dalam *dashboard*).
- **Wawasan Bimbingan Berbasis Bukti:** Analisis SHAP memberikan penjelasan faktor risiko utama (mis., *avg_score_early*, *durasi_partisipasi_early*, *total_hari_aktif_early*), mendukung strategi bimbingan yang lebih terarah (contoh: rekomendasi remedial dini berdasarkan skor awal). Selain itu, analisis *False Negative* (848 kasus) menyoroti perlunya perhatian tambahan pada mahasiswa lulus yang terdeteksi berisiko.

3. Untuk Mahasiswa:

- **Memperoleh Dukungan Akademik Dini:** Melalui **deteksi risiko awal** oleh EWS, mahasiswa yang teridentifikasi dapat memperoleh **intervensi suportif**

(misalnya bimbingan atau remedial) secara **tepat waktu**, sehingga meningkatkan peluang retensi dan keberhasilan studi mereka.

4. Kontribusi Ilmiah & Metodologis:

Penelitian ini memvalidasi kelayakan **prediksi EWS sangat dini (30 hari)** menggunakan XGBoost (di-*tuning*, 22 fitur), yang berkinerja solid (AUC/F1=0.71) dan unggul atas *baseline* RF. **Interpretabilitas model** dicapai melalui **SHAP**, yang efektif mengaitkan fitur kunci dengan rekomendasi pedagogis. *Pipeline* analisis disajikan secara **reproducible** dalam **artefak notebook Jupyter** untuk studi lanjutan.

Klaim Kinerja:

EWS berbasis **XGBoost** (22 fitur LMS+demografis, data 30 hari) mencapai **AUC=0.71** dan **F1-Score=0.71** untuk prediksi risiko **gagal/mundur**, **mengungguli baseline** Random Forest (Akurasi XGBoost 71.04% vs RF 70.23%). Analisis **SHAP** menyoroti **skor awal (avg_score_early)**, **durasi partisipasi (durasi_partisipasi_early)**, dan **hari aktif (total_hari_aktif_early)** sebagai faktor risiko kunci. Sistem ini mendukung **intervensi proaktif yang terinformasi dan dapat dijustifikasi** sejak awal masa studi.

1.4 Batasan Masalah

Untuk menjaga fokus dan kejelasan ruang lingkup, penelitian ini dibatasi oleh beberapa aspek sebagai berikut:

1. Konteks Dataset:

Studi ini menggunakan *Open University Learning Analytics Dataset* (OULAD) sebagai **studi kasus**. Penggunaan dataset standar ini memungkinkan **analisis mendalam dan perbandingan**, namun **adaptasi model** ke institusi pendidikan lain dengan karakteristik mahasiswa atau sistem LMS yang berbeda mungkin memerlukan **penyesuaian dan validasi tambahan**.

2. Fokus Temporal:

Ruang lingkup temporal analisis **difokuskan secara strategis** pada data aktivitas LMS selama **30 hari pertama** perkuliahan. Pilihan ini **sejalan dengan tujuan utama EWS** untuk memungkinkan **intervensi sedini mungkin**. Analisis mengenai dinamika perilaku mahasiswa dalam jangka waktu yang lebih panjang menjadi **area potensial untuk penelitian selanjutnya**.

3. Pemilihan Variabel:

Variabel prediktor yang digunakan **dipilih berdasarkan ketersediaan** dalam dataset OULAD, mencakup metrik aktivitas LMS awal dan data demografis dasar.

Faktor-faktor laten lain yang dapat memengaruhi performa (seperti motivasi intrinsik atau kondisi sosio-ekonomi rinci) **diakui keberadaannya**, namun **berada di luar variabel terukur** yang dianalisis dalam studi ini.

4. Pendekatan Algoritmik:

Evaluasi algoritmik terkonsentrasi pada *Extreme Gradient Boosting* (XGBoost) sebagai model utama dan *Random Forest* sebagai pembanding (*baseline*), mengingat **kinerja efektif kedua model ini** dalam berbagai tugas klasifikasi. Eksplorasi menggunakan arsitektur model lain (misalnya, *deep learning*) **disimpan sebagai peluang** pengembangan di masa depan.

5. Tahapan Studi:

Fokus utama penelitian ini adalah pada **pengembangan, evaluasi kinerja, dan interpretasi model prediktif** EWS. Tahap **implementasi teknis** EWS dalam lingkungan operasional serta **evaluasi dampak** program intervensi yang dihasilkan **merupakan langkah lanjutan** yang tidak termasuk dalam cakupan studi ini.

BAB 2

KAJIAN TERKAIT

2.1 Pendahuluan

Bab ini menyajikan tinjauan sistematis terhadap literatur ilmiah yang relevan dengan pengembangan model prediksi performa akademik mahasiswa melalui pendekatan *Early Warning System* (EWS) berbasis data *Learning Management System* (LMS). Fokus utama kajian meliputi urgensi penerapan EWS dalam pendidikan tinggi, potensi data LMS sebagai sumber prediksi dini, penerapan metode *machine learning* dalam klasifikasi risiko akademik, serta pentingnya interpretabilitas model untuk mendukung intervensi yang berbasis bukti. Kajian ini bertujuan untuk memetakan perkembangan riset terkini serta memposisikan kontribusi penelitian ini dalam lanskap *learning analytics* modern.

2.2 Tinjauan Penelitian Terkait EWS Berbasis Data LMS

Identifikasi dini terhadap mahasiswa yang berisiko mengalami penurunan performa atau *drop out* menjadi isu strategis dalam manajemen pendidikan tinggi [1][2]. Penerapan EWS memberikan mekanisme prediktif untuk mendeteksi potensi risiko secara proaktif dan memungkinkan lembaga melakukan intervensi suportif secara tepat waktu. Sejalan dengan itu, pemanfaatan data digital yang terekam dalam LMS membuka peluang signifikan dalam *learning analytics*, karena setiap aktivitas mahasiswa seperti frekuensi akses, pola klik, durasi interaksi, partisipasi forum, hingga waktu pengumpulan tugas merepresentasikan perilaku belajar yang dapat dianalisis secara kuantitatif [3][4][7].

Dataset publik berskala besar seperti *Open University Learning Analytics Dataset* (OULAD) menjadi acuan utama dalam banyak penelitian karena menyediakan data komprehensif yang mendukung replikasi dan perbandingan hasil model [13]. Beberapa studi menunjukkan bahwa metrik keterlibatan awal seperti *average score early*, *total clicks*, dan jumlah hari aktif merupakan indikator utama kesuksesan akademik [5][14][17]. Integrasi data perilaku dengan faktor demografis seperti usia, jenis kelamin, dan riwayat pendidikan juga terbukti memperkuat akurasi prediksi [8].

Seiring perkembangan teknologi analitik, algoritma *machine learning* semakin mendominasi implementasi EWS karena kemampuannya mengenali pola non-linear dan menangani data berdimensi tinggi [4][9]. Algoritma seperti *Logistic Regression*, *Support Vector Machine* (SVM), *Random Forest*, dan *Neural Network* banyak digunakan, dengan pendekatan *ensemble learning* seperti *XGBoost* menunjukkan performa unggul dalam

prediksi performa akademik [10][11][15]. Keunggulan ini sering dikaitkan dengan kemampuannya dalam menangani data tidak lengkap (*missing values*), memiliki *regularization* bawaan untuk mencegah *overfitting*, serta efisiensi komputasi pada data tabular [9][10]. Penelitian terkini juga menyoroti pentingnya *hyperparameter optimization* dan *feature selection* untuk meningkatkan *generalization performance* model [9][15].

Namun, peningkatan akurasi model sering kali disertai dengan berkurangnya interpretabilitas. Model yang bersifat *black box* sulit diimplementasikan dalam konteks pendidikan karena tidak memberikan pemahaman yang jelas mengenai penyebab risiko akademik [12][16]. Oleh karena itu, pendekatan interpretasi pasca-hoc seperti *SHapley Additive Explanations* (SHAP) dan *Local Interpretable Model-Agnostic Explanations* (LIME) menjadi krusial dalam menjelaskan kontribusi setiap fitur terhadap hasil prediksi [12][16]. Pendekatan ini memungkinkan pendidik memahami faktor-faktor yang paling berpengaruh terhadap performa mahasiswa, sehingga intervensi dapat disesuaikan secara individual.

2.3 Posisi Penelitian

Secara keseluruhan, literatur yang ada menunjukkan efektivitas penggunaan data LMS dalam pengembangan sistem peringatan dini berbasis *machine learning* [1][3][5][7]. Namun, sebagian besar penelitian masih berfokus pada periode pembelajaran menengah atau akhir semester. Penelitian ini berupaya mengisi kesenjangan tersebut dengan memusatkan perhatian pada prediksi sangat dini, yakni dalam 30 hari pertama perkuliahan, menggunakan algoritma *XGBoost* yang dioptimasi [10][15].

Selain itu, penelitian ini menekankan aspek interpretabilitas melalui analisis SHAP untuk menelusuri pengaruh relatif setiap fitur perilaku terhadap risiko akademik individu [12][16]. Dengan menggunakan dataset OULAD sebagai studi kasus [13], penelitian ini diharapkan mampu menghasilkan model EWS yang tidak hanya unggul secara prediktif, tetapi juga transparan dan aplikatif bagi pengambil kebijakan akademik. Kontribusi yang diharapkan adalah tersusunnya kerangka metodologis yang mampu mendukung implementasi sistem peringatan dini yang akurat, interpretatif, dan dapat diadaptasi oleh institusi pendidikan tinggi lainnya.

BAB 3 METODE PENELITIAN

3.1 Deskripsi Solusi Usulan

Penelitian ini bertujuan mengembangkan sebuah *Early Warning System* (EWS) yang memanfaatkan teknik *machine learning* untuk melakukan prediksi dini terhadap performa akademik mahasiswa. Sistem ini dirancang menggunakan data aktivitas pembelajaran yang terekam dalam *Learning Management System* (LMS).

Tujuan utama dari EWS yang diusulkan adalah untuk menyediakan alat bantu bagi pihak akademik (dosen atau pengelola program studi) dalam mengidentifikasi mahasiswa yang berpotensi mengalami risiko akademik pada tahap awal perkuliahan. Dengan deteksi dini ini, diharapkan intervensi pembelajaran yang tepat dapat segera dilakukan sebelum terjadi penurunan performa yang lebih signifikan pada mahasiswa tersebut.

Solusi yang dikembangkan mengintegrasikan algoritma *Extreme Gradient Boosting* (XGBoost) sebagai metode klasifikasi utama. Model XGBoost ini akan memprediksi risiko akademik berdasarkan analisis pola perilaku pembelajaran mahasiswa selama **30 hari pertama** masa studi mereka. Untuk memastikan hasil prediksi dapat dipahami dan ditindaklanjuti oleh para pengajar, sistem ini juga dilengkapi dengan komponen interpretabilitas menggunakan metode SHAP (*SHapley Additive exPlanations*).

Pendekatan penelitian ini didasarkan pada temuan sebelumnya, seperti yang dikemukakan oleh Macfadyen dan Dawson [7], yang menunjukkan adanya korelasi kuat antara data interaksi mahasiswa di LMS (misalnya, frekuensi login, aktivitas forum, dan pengumpulan tugas) dengan hasil akademik akhir mereka. Selain itu, prinsip desain sistem ini juga mengacu pada model EWS berbasis *machine learning* yang telah dikembangkan oleh Chung dan Lee [8] untuk deteksi dini potensi *dropout* mahasiswa.

3.2 Dataset

Penelitian ini memanfaatkan *Open University Learning Analytics Dataset* (OULAD), sebuah dataset publik yang berisi data komprehensif mengenai aktivitas pembelajaran daring mahasiswa. Dataset ini mencakup data dari tujuh modul (mata kuliah) yang berbeda selama dua periode perkuliahan (B = Februari, J = Oktober). OULAD sering digunakan dalam penelitian *learning analytics* karena kelengkapan datanya yang mencakup karakteristik demografis mahasiswa, interaksi mereka dengan materi pembelajaran di LMS (*Learning Management System*), serta hasil asesmen yang diperoleh.

Dataset OULAD terdiri dari tujuh berkas data CSV utama yang saling terhubung, yaitu:

- Courses.csv: Berisi informasi mengenai daftar modul dan durasinya.
- Assessments.csv: Menyajikan detail asesmen, termasuk jenis (TMA, CMA, Exam), tanggal batas pengumpulan, dan bobot nilainya.
- Vle.csv: Merinci daftar materi atau konten pembelajaran yang tersedia dalam *Virtual Learning Environment* (VLE) atau LMS.
- StudentInfo.csv: Memuat data profil demografis mahasiswa, seperti jenis kelamin, tingkat pendidikan terakhir, wilayah geografis, kelompok usia, status disabilitas, serta hasil akhir studi (final_result). File ini menjadi basis utama untuk data mahasiswa.
- StudentRegistration.csv: Mencatat informasi waktu registrasi mahasiswa pada suatu modul dan tanggal pembatalan registrasi (jika ada).
- StudentAssessment.csv: Berisi data nilai yang diperoleh setiap mahasiswa pada setiap asesmen yang dikerjakan.
- StudentVle.csv: Mencatat log interaksi harian setiap mahasiswa dengan materi pembelajaran di VLE, diukur berdasarkan jumlah klik (sum_click).

Tahapan Pengolahan Data

Proses pengolahan data dalam penelitian ini mengikuti tahapan berikut:

1. Penggabungan Data: Tabel-tabel utama (studentInfo, studentAssessment, studentVle) digabungkan berdasarkan kunci identifikasi mahasiswa (id_student), kode modul (code_module), dan kode presentasi (code_presentation) untuk membentuk dataset awal.
2. Rekayasa Fitur (*Feature Engineering*): Berdasarkan data aktivitas selama 30 hari pertama perkuliahan, fitur-fitur baru berikut ini dibentuk untuk menangkap pola perilaku awal mahasiswa:
 - avg_score_early: Rata-rata skor asesmen yang dikumpulkan mahasiswa dalam 30 hari pertama.
 - total_clicks_early: Jumlah total interaksi (klik) mahasiswa dengan materi VLE selama 30 hari pertama.
 - total_hari_aktif_early: Jumlah hari unik mahasiswa aktif berinteraksi dengan VLE selama 30 hari pertama.
 - durasi_partisipasi_early: Rentang waktu (dalam hari) antara interaksi pertama dan terakhir mahasiswa dengan VLE selama 30 hari pertama.

3. Pra-pemrosesan Data:

- Encoding Fitur Kategorikal: Fitur-fitur demografis kategorikal (seperti gender, region, highest_education, disability) diubah menjadi representasi numerik menggunakan teknik *One-Hot Encoding* (pd.get_dummies).
 - Penanganan Nilai Hilang: Nilai hilang (NaN) yang muncul pada fitur-fitur hasil rekayasa (misalnya, jika mahasiswa tidak memiliki skor atau klik dalam 30 hari pertama) diisi dengan nilai 0. *Scaling* atau normalisasi pada fitur numerik tidak diterapkan pada tahap ini.
4. Pembentukan Target Variabel: Kolom final_result digunakan untuk membuat variabel target biner (target_lulus), di mana nilai 1 merepresentasikan 'Lulus' ('Pass' atau 'Distinction') dan nilai 0 merepresentasikan 'Gagal/Mundur' ('Fail' atau 'Withdrawn').
5. Pembagian Data: Dataset akhir yang telah diproses kemudian dibagi menjadi dua bagian: 80% sebagai data latih (*training set*) dan 20% sebagai data uji (*test set*) menggunakan pemisahan acak (train_test_split).

3.3 Metode yang digunakan

Metodologi yang diterapkan dalam penelitian ini mencakup pemilihan dan pelatihan model *machine learning*, serta teknik untuk menjelaskan hasil prediksi model tersebut.

3.3.1 Arsitektur Model

Penelitian ini membandingkan kinerja dua algoritma klasifikasi:

1. *Extreme Gradient Boosting* (XGBoost): Digunakan sebagai model utama karena kemampuannya yang unggul dalam menangani dataset dengan fitur kompleks, hubungan non-linear antar fitur, serta performa prediktifnya yang tinggi di berbagai aplikasi [9], [10]. XGBoost merupakan algoritma berbasis pohon (*tree-based*) yang menerapkan teknik *gradient boosting*.
2. *Random Forest* (RF): Digunakan sebagai model pembanding. RF adalah algoritma *ensemble* klasik yang juga berbasis pohon dan dikenal cukup robust.

Kedua model ini dilatih menggunakan data latih (80% dari dataset) untuk memprediksi variabel target target_lulus. Untuk model XGBoost, dilakukan proses *hyperparameter tuning* guna mencari kombinasi parameter optimal. Tuning ini menggunakan pendekatan *Grid Search* (GridSearchCV) dengan validasi silang 3-fold (cv=3) di dalam prosesnya. Parameter yang dioptimalkan meliputi:

- `n_estimators`: Jumlah pohon dalam *ensemble*.
- `learning_rate`: Laju pembelajaran untuk setiap pohon.
- `max_depth`: Kedalaman maksimum setiap pohon.

Pendekatan optimasi ini bertujuan meningkatkan akurasi model, sebagaimana ditunjukkan dalam penelitian Tarwidi et al. [10] mengenai pengaruh parameter optimal pada XGBoost. Model RF dilatih dengan parameter default (`n_estimators=100`) sebagai pembandingan kinerja.

3.3.2 Penanganan Ketidakseimbangan Data

Observasi awal pada distribusi variabel target `target_lulus` dalam dataset OULAD menunjukkan adanya potensi ketidakseimbangan kelas, di mana jumlah mahasiswa yang 'Lulus' ('Pass' atau 'Distinction') mungkin tidak seimbang dengan yang 'Gagal/Mundur' ('Fail' atau 'Withdrawn'). Meskipun teknik seperti SMOTE (*Synthetic Minority Oversampling Technique*), sebagaimana diterapkan oleh Airlangga [11], dapat digunakan untuk mengatasi masalah ini dengan menghasilkan data sintetis bagi kelas minoritas, teknik penyeimbangan data tersebut tidak diterapkan dalam implementasi notebook ini. Model dilatih langsung pada data asli yang telah diproses.

3.3.3 Interpretabilitas Model

Dalam konteks EWS pendidikan, kemampuan untuk menjelaskan mengapa suatu prediksi dibuat (transparansi model) sangatlah penting agar hasil prediksi dapat dipercaya dan ditindaklanjuti oleh pengajar. Untuk memenuhi kebutuhan ini, digunakan metode SHAP (*SHapley Additive exPlanations*) guna menganalisis kontribusi relatif setiap fitur terhadap prediksi akhir yang dihasilkan oleh model XGBoost.

SHAP, yang didasarkan pada konsep nilai Shapley dari teori permainan (*game theory*), menghitung dampak rata-rata setiap fitur pada output model. Metode ini cocok untuk model berbasis pohon seperti XGBoost. Sejalan dengan temuan Hooshyar dan Yang [12] mengenai potensi variabilitas interpretasi *post-hoc* seperti SHAP dan LIME pada model kompleks, penelitian ini **menitikberatkan pada analisis interpretabilitas global** (menggunakan *summary plot* SHAP) untuk memahami pentingnya fitur secara keseluruhan, bukan hanya pada penjelasan prediksi individual, demi menjaga konsistensi analisis.

3.4 Perbedaan dengan solusi sebelumnya

Meskipun beberapa penelitian sebelumnya telah berfokus pada pengembangan EWS berbasis data LMS [7], [8], pendekatan tersebut seringkali menggunakan dataset yang lebih terbatas cakupannya atau mengandalkan algoritma klasifikasi tradisional. Penelitian ini membedakan diri melalui beberapa aspek utama:

1. Penggunaan Dataset Skala Besar dan Terbuka: Memanfaatkan dataset OULAD yang komprehensif dan tersedia secara publik, memungkinkan potensi generalisasi hasil yang lebih kuat dibandingkan dataset institusional tunggal.
2. Penerapan Algoritma Boosting Tingkat Lanjut: Mengimplementasikan XGBoost dengan proses optimasi *hyperparameter* menggunakan *Grid Search*. Pendekatan ini ditujukan untuk mencapai performa prediktif yang lebih tinggi dibandingkan model *ensemble* klasik seperti Random Forest atau model linear [9], [10].
3. Fokus pada Interpretabilitas: Menambahkan analisis interpretabilitas model menggunakan SHAP. Hal ini memungkinkan pemangku kepentingan (pendidik) untuk memahami faktor-faktor utama yang memengaruhi prediksi risiko mahasiswa, sehingga hasil model tidak hanya akurat tetapi juga dapat dijelaskan secara edukatif [12].

Dengan kombinasi penggunaan dataset yang kaya, algoritma *boosting* yang dioptimalkan, dan penekanan pada interpretabilitas, penelitian ini diharapkan dapat menghasilkan model EWS yang akurat, dapat digeneralisasi, serta transparan sebagai alat bantu pengambilan keputusan bagi para pendidik.

3.5 Metrik Evaluasi

Kinerja model klasifikasi yang dikembangkan dalam penelitian ini dievaluasi menggunakan serangkaian metrik standar untuk klasifikasi biner. Metrik-metrik utama yang digunakan meliputi:

1. Akurasi (*Accuracy*): Mengukur proporsi keseluruhan prediksi yang benar (baik prediksi Lulus maupun Gagal/Mundur) terhadap total jumlah data uji.
2. Presisi (*Precision*): Mengukur tingkat ketepatan model ketika memprediksi suatu kelas positif (Lulus). Secara spesifik, ini adalah rasio prediksi Lulus yang benar terhadap total prediksi Lulus yang dibuat model.
3. Recall (Sensitivitas atau *Sensitivity*): Mengukur kemampuan model untuk mengidentifikasi semua kasus positif (Lulus) yang sebenarnya. Ini adalah rasio

prediksi Lulus yang benar terhadap total mahasiswa yang sebenarnya Lulus.

4. F1-Score: Merupakan rata-rata harmonik dari Presisi dan Recall. Metrik ini memberikan ukuran keseimbangan antara kemampuan model untuk tidak salah memprediksi Lulus (Presisi) dan kemampuannya menemukan semua yang Lulus (Recall).
5. AUC-ROC (*Area Under the Receiver Operating Characteristic Curve*): Mengukur kemampuan diskriminatif model secara keseluruhan dalam membedakan antara kelas Lulus dan Gagal/Mundur di berbagai ambang batas probabilitas. Nilai AUC yang lebih tinggi menunjukkan kemampuan pemisahan kelas yang lebih baik.
6. Confusion Matrix: Sebuah tabel visualisasi yang merangkum hasil prediksi (True Positive, True Negative, False Positive, False Negative), memberikan gambaran detail tentang jenis kesalahan yang dibuat model.

Pemilihan metrik-metrik ini didasarkan pada praktik umum dalam evaluasi model prediksi akademik dan sistem peringatan dini [7]–[10], yang menekankan pentingnya mengevaluasi tidak hanya akurasi keseluruhan tetapi juga kemampuan model dalam menangani setiap kelas secara spesifik, terutama dalam konteks yang mungkin memiliki konsekuensi berbeda untuk prediksi yang salah (misalnya, *False Negative*).

BAB 4

HASIL DAN PEMBAHASAN

Bab ini menyajikan temuan utama dari penelitian, yang diperoleh melalui eksekusi metodologi yang diuraikan sebelumnya. Bagian ini mencakup hasil pra-pemrosesan data, hasil optimasi hyperparameter, evaluasi kuantitatif kinerja model pembandingan, serta analisis interpretasi mendalam terhadap model EWS yang diusulkan menggunakan SHAP.

4.1 Deskripsi Data dan Hasil Pra-pemrosesan

Penelitian ini menggunakan dataset OULAD yang telah melalui tahap pembersihan dan rekayasa fitur. Total dataset akhir (`df_expert`) yang digunakan untuk eksperimen terdiri dari 32.593 data mahasiswa.

Sesuai dengan tujuan penelitian untuk memprediksi status kelulusan (Lulus vs. Putus Kuliah), data target (`target_lulus`) dibagi menjadi dua kelas berdasarkan kolom `final_result`:

- Kelas 0 (Gagal/Mundur): 17.208 mahasiswa (sekitar 52,79%)
- Kelas 1 (Lulus/Distinction): 15.385 mahasiswa (sekitar 47,21%)

Dataset ini menunjukkan adanya ketidakseimbangan kelas (*class imbalance*) minor, di mana kelas 'Gagal/Mundur' sedikit lebih dominan.

Sesuai metodologi, data dibagi menjadi data latih (80%) dan data uji (20%) menggunakan `train_test_split`:

- Data Latih (*Train Set*): 26.074 sampel.
- Data Uji (*Test Set*): 6.519 sampel.

Model dilatih langsung menggunakan data latih yang memiliki distribusi kelas asli (minor imbalance). Proses evaluasi kinerja model dilakukan pada data uji (6.519 sampel) untuk merefleksikan performa model pada data baru yang tidak terlihat saat pelatihan



```
Berhasil membuat kolom 'target_lulus'. (0: 17208, 1: 15385)
```

Gambar 4.1: Distribusi Kelas Target ('`target_lulus`') pada Keseluruhan Dataset (Kelas 0: Gagal/Mundur = 17.208; Kelas 1: Lulus = 15.385)

4.2 Hasil Optimasi Model (Hyperparameter Tuning)

Model utama yang diusulkan, XGBoost, dioptimasi menggunakan GridSearchCV untuk menemukan kombinasi *hyperparameter* terbaik dari *grid* yang telah ditentukan (`n_estimators`: [100, 200], `learning_rate`: [0.1, 0.05], `max_depth`: [3, 5]). Proses *tuning* ini bertujuan untuk memaksimalkan kinerja prediktif model pada data latih melalui validasi silang (`cv=3`). Berdasarkan hasil eksekusi dari notebook (`[execution_count: 12]`), *hyperparameter* optimal (terbaik) yang ditemukan untuk model XGBoost adalah sebagai berikut:

- `learning_rate`: 0.1
- `max_depth`: 3
- `n_estimators`: 200

Kombinasi parameter ini kemudian digunakan untuk melatih model final XGBoost (`model_xgb_expert`) pada keseluruhan data latih, yang selanjutnya dievaluasi pada data uji.

```
--- F.4: Uji Model XGBoost Expert + Tuning ---  
✓ Selesai Tuning! Parameter Terbaik: {'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 200}  
Model XGBoost Expert (Final) siap.
```

Gambar 4.2: Hasil Optimasi Hyperparameter XGBoost Menggunakan GridSearchCV (Parameter Terbaik: `learning_rate=0.1`, `max_depth=3`, `n_estimators=200`)

4.3 Evaluasi Kinerja Model

Kinerja model EWS yang diusulkan (XGBoost dengan 22 fitur yang dioptimasi) dievaluasi secara komprehensif pada data uji (Test Set). Performanya dibandingkan dengan model *baseline*, yaitu Random Forest (RF) yang juga menggunakan 22 fitur (*expert*) namun dengan parameter *default*. Evaluasi ini menggunakan metrik standar yang dijelaskan dalam Bab 3.5.

Tabel 4.1: Perbandingan Metrik Kinerja Model pada Data Uji (Test Set)

Model	Akurasi	AUC-ROC	F1-Score (Weighted)
Random Forest (Baseline - 22 Fitur)	0.7023	-	0.70*

XGBoost (Model Usulan - 22 Fitur)	0.7104	0.7111	0.71
--	---------------	---------------	-------------

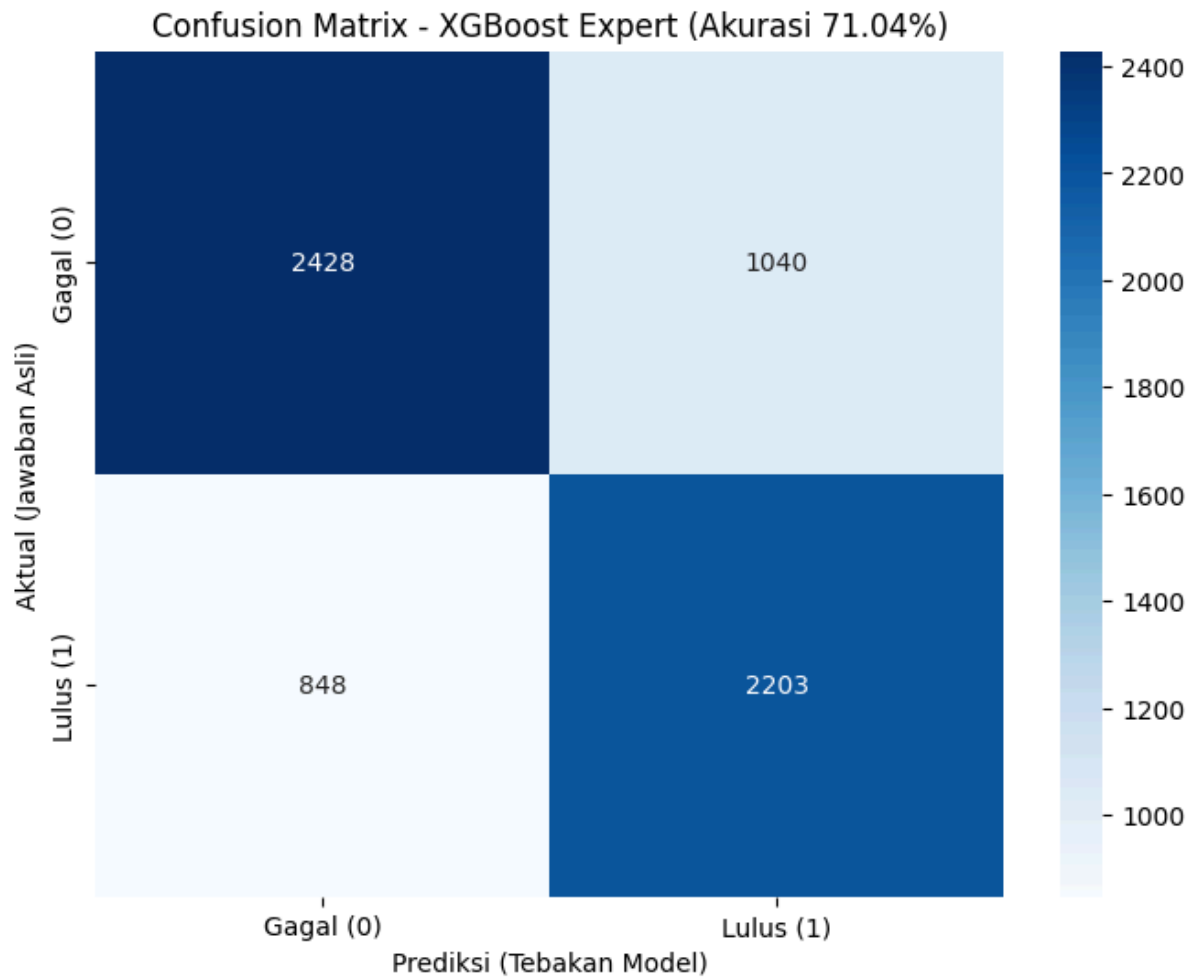
Tabel 4.2: Rincian Classification Report XGBoost (Model Usulan) pada Data Uji

Kelas	Precision	Recall	F1-Score	Support
0 (Gagal/Mundur)	0.74	0.70	0.72	3468
1 (Lulus)	0.68	0.72	0.70	3051
Weighted Avg	0.71	0.71	0.71	6519

Dari hasil evaluasi, dapat disimpulkan:

1. **Keunggulan Model Usulan:** Model XGBoost yang dioptimasi secara konsisten sedikit mengungguli model *baseline* Random Forest (keduanya menggunakan 22 fitur) dalam hal Akurasi (71.04% vs 70.23%). Model XGBoost juga menunjukkan kemampuan diskriminasi antar kelas yang baik dengan AUC-ROC 0.7111. Hal ini mengkonfirmasi klaim kinerja yang disebutkan di Bab 1.
2. **Keseimbangan Model:** Berdasarkan *Classification Report* (Tabel 4.2), model XGBoost menunjukkan kinerja yang relatif seimbang dalam memprediksi kedua kelas, dengan F1-Score 0.72 untuk kelas 'Gagal/Mundur' dan 0.70 untuk kelas 'Lulus'. Kemampuan model untuk mengidentifikasi kelas 'Gagal/Mundur' (Recall = 0.70) cukup baik untuk tujuan EWS dalam mendeteksi mahasiswa berisiko.

Untuk memvisualisasikan kinerja ini secara lebih rinci, *confusion matrix* dari model XGBoost disajikan pada **Gambar 4.3**



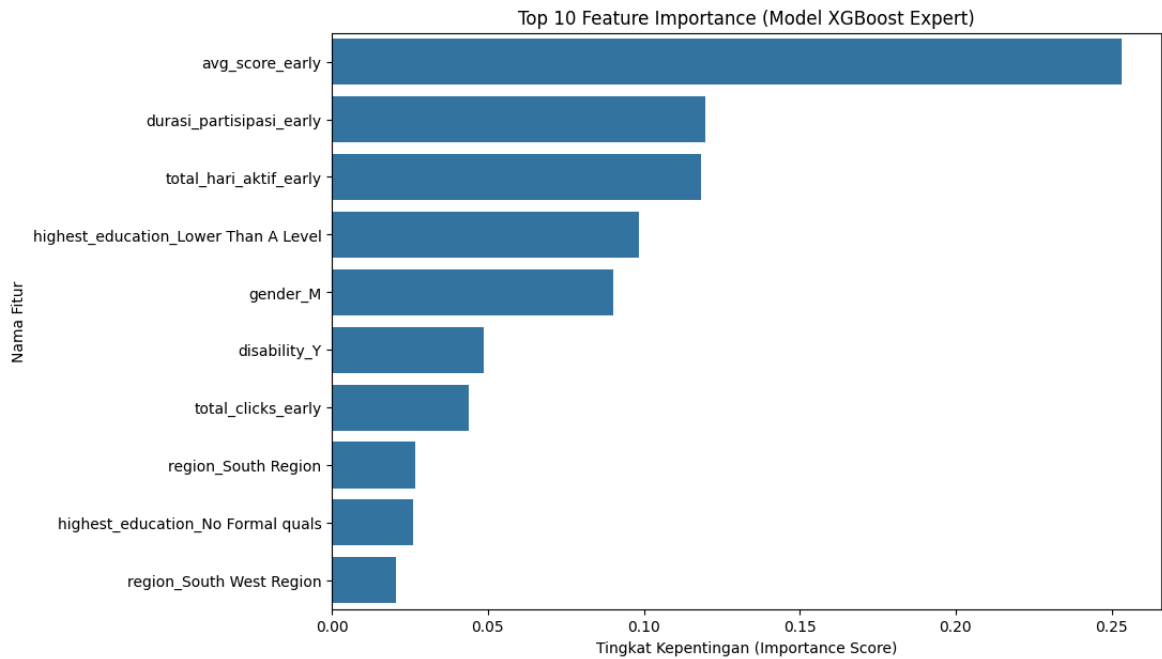
Gambar 4.3: *Confusion Matrix Model XGBoost pada Data Uji (Keterangan: Gambar ini menunjukkan 2428 True Negative (Prediksi Gagal Benar), 848 False Negative (Prediksi Gagal Salah), 1040 False Positive (Prediksi Lulus Salah), dan 2203 True Positive (Prediksi Lulus Benar))*

Visualisasi pada **Gambar 4.2** memperjelas hasil pada Tabel 4.2, menunjukkan secara spesifik jumlah prediksi yang benar dan salah untuk setiap kelas.

4.4 Pembahasan dan Analisis Interpretabilitas (SHAP)

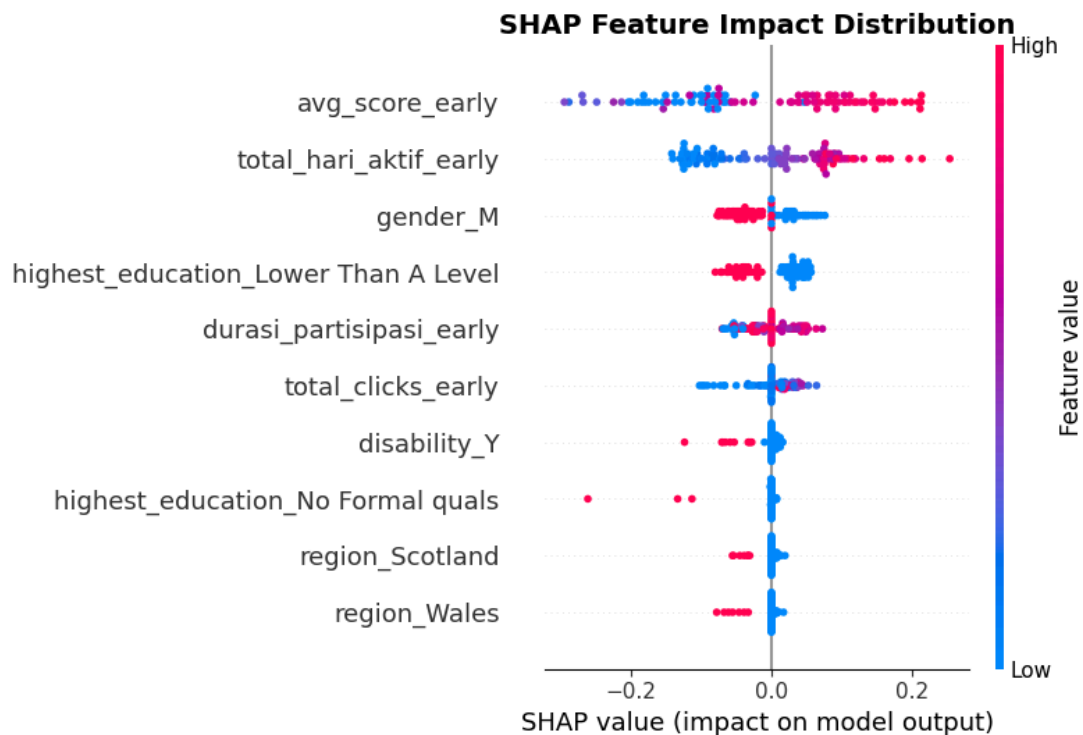
Untuk memenuhi tujuan penelitian ke-4, yaitu memahami faktor-faktor prediktor yang paling berpengaruh, analisis interpretabilitas model dilakukan menggunakan SHAP (*SHapley Additive exPlanations*) pada model XGBoost final. Analisis ini krusial untuk memahami kontribusi fitur-fitur dari 30 hari pertama studi terhadap prediksi risiko kelulusan mahasiswa

Gambar 4.4 menyajikan visualisasi pentingnya fitur secara global, diurutkan berdasarkan rata-rata dampak absolut nilai SHAP pada output model.



Gambar 4.4: SHAP Global Feature Importance (Bar Plot) Model XGBoost
(Keterangan: Plot batang ini menunjukkan rata-rata nilai absolut SHAP untuk 10 fitur teratas. Fitur dengan dampak prediksi terbesar secara keseluruhan berada di paling atas.)

Hasil analisis SHAP (*Feature Importance*) pada Gambar 4.3 secara jelas mengidentifikasi avg_score_early (rata-rata skor awal) sebagai prediktor paling kuat, diikuti oleh fitur keterlibatan awal seperti durasi_partisipasi_early (durasi partisipasi) dan total_hari_aktif_early (total hari aktif), serta beberapa fitur demografis (highest_education_Lower Than A Level, gender_M). Untuk memahami arah dan distribusi pengaruh fitur-fitur kunci ini terhadap prediksi, **Gambar 4.4** (SHAP Summary Plot) memberikan visualisasi yang lebih rinci.



Gambar 4.5: SHAP Summary Plot (Beeswarm Plot) Model XGBoost

(Keterangan: Setiap titik mewakili satu mahasiswa dari data uji sampel. Posisi horizontal menunjukkan dampak fitur pada output model (nilai SHAP positif mendorong prediksi Lulus, negatif mendorong prediksi Gagal/Mundur). Warna menunjukkan nilai fitur asli (merah=tinggi, biru=rendah).

Berdasarkan Gambar 4.4 dan 4.5, temuan utama dari analisis interpretabilitas dapat didiskusikan sebagai berikut:

1. **avg_score_early (Rata-rata Skor Awal):** Konsisten sebagai prediktor paling dominan. Gambar 4.4 menunjukkan tren yang jelas: nilai skor awal yang tinggi (titik merah) secara kuat berkontribusi positif terhadap prediksi 'Lulus' (nilai SHAP > 0), sementara skor awal yang rendah (titik biru, terutama mendekati 0) memberikan kontribusi negatif terkuat, mendorong prediksi ke arah 'Gagal/Mundur' (nilai SHAP < 0).
2. **durasi_partisipasi_early (Durasi Partisipasi):** Fitur ini, yang mengukur rentang hari aktif dalam 30 hari pertama, juga signifikan. Durasi partisipasi yang lebih panjang (titik merah) cenderung memiliki nilai SHAP positif, meningkatkan kemungkinan prediksi 'Lulus'.
3. **total_hari_aktif_early (Total Hari Aktif):** Jumlah hari unik mahasiswa mengakses LMS dalam periode awal menunjukkan pola serupa. Aktivitas yang lebih konsisten

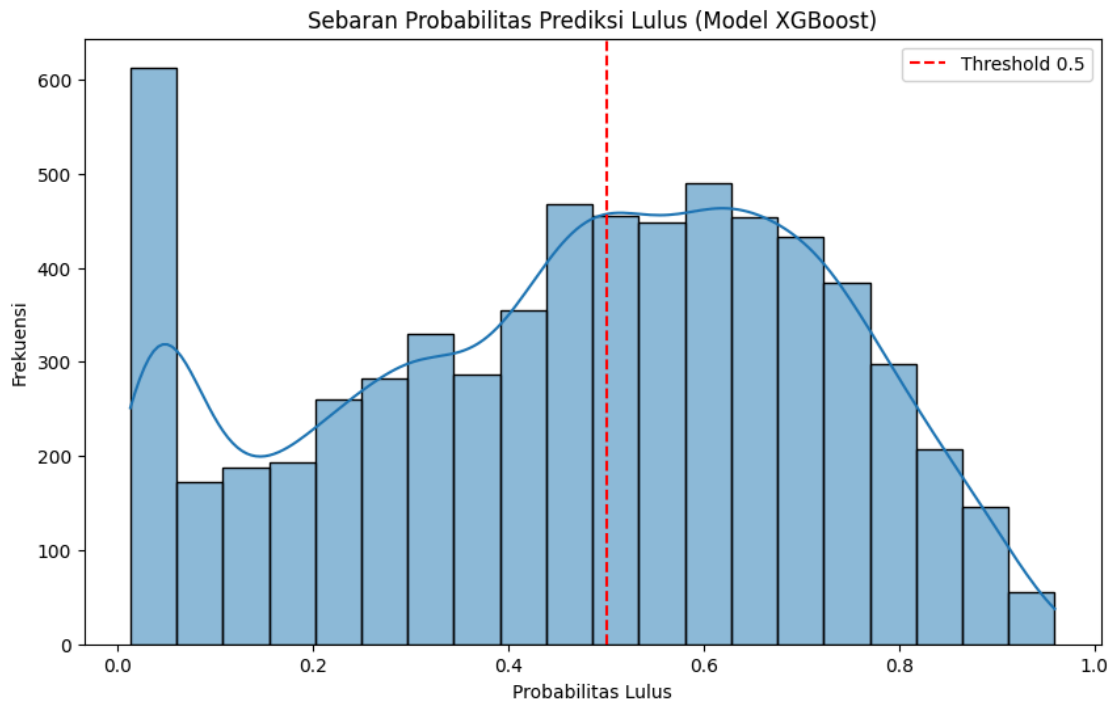
(jumlah hari aktif tinggi, titik merah) berkorelasi dengan nilai SHAP positif dan prediksi 'Lulus'.

4. **Fitur Keterlibatan Lainnya (total_clicks_early):** Meskipun tidak berada di 3 besar *importance*, jumlah klik awal juga menunjukkan tren positif: klik yang lebih banyak (merah) umumnya berkontribusi pada prediksi 'Lulus' .
5. **Fitur Demografis:** Beberapa fitur demografis seperti highest_education_Lower Than A Level (pendidikan terakhir di bawah A Level) dan gender_M (laki-laki) juga muncul sebagai faktor yang cukup berpengaruh dalam model ini, meskipun interpretasi dampaknya (positif/negatif) mungkin lebih kompleks dan memerlukan konteks dataset lebih lanjut.

Diskusi Temuan SHAP: Temuan ini memberikan bukti empiris yang kuat bahwa kinerja akademik awal (avg_score_early) dan tingkat keterlibatan (*engagement*) awal (terutama durasi_partisipasi_early dan total_hari_aktif_early) adalah pilar utama dalam memprediksi keberhasilan studi pada tahap sangat dini. Hasil ini sejalan dengan literatur (Bab 2) namun secara spesifik mengkuantifikasi dampak relatif fitur-fitur ini dalam model XGBoost pada periode 30 hari pertama. Hal ini membuktikan bahwa intervensi tidak perlu menunggu hingga tengah semester; sinyal risiko atau potensi keberhasilan sudah dapat diukur secara kuantitatif dan dijelaskan faktor pendorongnya sejak bulan pertama perkuliahan .

4.5 Analisis Hasil Prediksi, Kesalahan Model, dan Implikasi Intervensi

Model XGBoost yang telah dilatih menghasilkan prediksi probabilitas kelulusan untuk setiap mahasiswa dalam data uji. **Gambar 4.5** menyajikan distribusi dari probabilitas prediksi ini, menunjukkan sebaran keyakinan model terhadap prediksi 'Lulus' di seluruh data uji. Terlihat adanya konsentrasi prediksi pada probabilitas rendah (mendekati 0) dan juga pada rentang probabilitas menengah ke atas (sekitar 0.4 hingga 0.8), dengan garis *threshold* 0.5 yang memisahkan prediksi akhir menjadi 'Gagal/Mundur' atau 'Lulus'.



Gambar 4.6: Sebaran Probabilitas Prediksi Lulus oleh Model XGBoost pada Data Uji (*Keterangan: Histogram menunjukkan frekuensi mahasiswa pada setiap rentang probabilitas prediksi Lulus. Garis merah putus-putus menandakan threshold klasifikasi 0.5.*)

Untuk melihat contoh konkret hasil prediksi, **Gambar 4.6** menampilkan 10 mahasiswa yang diprediksi Lulus dengan probabilitas tertinggi dan 10 mahasiswa yang diprediksi Gagal/Mundur dengan probabilitas lulus terendah (keyakinan gagal tertinggi). Terlihat bahwa model dapat memberikan prediksi dengan keyakinan tinggi (probabilitas mendekati 1 atau 0) pada beberapa kasus. Contohnya, mahasiswa ID 622374 diprediksi Lulus dengan probabilitas 0.959, sementara mahasiswa ID 577673 diprediksi Gagal/Mundur dengan probabilitas Lulus hanya 0.013.


```

=====
✅ HASIL PREDIKSI MAHASISWA PADA DATA UJI
=====

--- A. 10 Mahasiswa yang DIPREDIKSI LULUS (Probabilitas Tertinggi) ---
id_student final_result Hasil_Aktual Probabilitas_Lulus
11915 622374 Pass LULUS (1) 0.959394
9020 420935 Distinction LULUS (1) 0.954417
24177 335644 Fail GAGAL/MUNDUR (0) 0.948968
13549 476964 Distinction LULUS (1) 0.946803
25487 595106 Distinction LULUS (1) 0.945735
25246 586781 Distinction LULUS (1) 0.944996
22222 1674799 Distinction LULUS (1) 0.944691
23700 560431 Distinction LULUS (1) 0.942756
24239 380435 Distinction LULUS (1) 0.942334
27520 634479 Fail GAGAL/MUNDUR (0) 0.940661

--- B. 10 Mahasiswa yang DIPREDIKSI GAGAL/MUNDUR (Probabilitas Terendah) ---
id_student final_result Hasil_Aktual Probabilitas_Lulus \
25029 577673 Withdrawn GAGAL/MUNDUR (0) 0.012740
19612 502479 Withdrawn GAGAL/MUNDUR (0) 0.012740
10493 2355744 Withdrawn GAGAL/MUNDUR (0) 0.013420
21994 683170 Fail GAGAL/MUNDUR (0) 0.013420
20667 589083 Withdrawn GAGAL/MUNDUR (0) 0.013420
3063 502391 Fail GAGAL/MUNDUR (0) 0.013420
...
12104 635410 Withdrawn 0.511903 74.0
=====
✅ Data hasil prediksi berhasil diekspor ke: hasil_prediksi_mahasiswa_ews_final.csv
=====

```

Gambar 4.7: Contoh Hasil Prediksi Mahasiswa pada Data Uji (*Keterangan: Menampilkan ID mahasiswa, hasil akhir aktual, hasil prediksi, dan probabilitas prediksi Lulus untuk kasus dengan keyakinan prediksi tertinggi (Lulus) dan terendah (Gagal/Mundur).*)

Namun, model tidak selalu sempurna dalam prediksinya. Analisis lebih lanjut terhadap kasus-kasus di mana model melakukan kesalahan prediksi, sebagaimana dirangkum dalam *confusion matrix* (**Gambar 4.3**), penting untuk merumuskan rekomendasi intervensi yang lebih efektif.

1. Kasus False Negative (Kritis):

- **Definisi:** Mahasiswa yang sebenarnya (aktual) Lulus, namun diprediksi 'Gagal/Mundur' oleh EWS.
- **Jumlah:** Terdapat **848** kasus *False Negative* dalam data uji (lihat **Gambar 4.2**).
- **Analisis:** Analisis sampel kasus ini dari notebook ([execution_count: 17], bagian C) menunjukkan bahwa sebagian mahasiswa dalam kelompok ini memang menunjukkan indikator risiko awal, seperti `avg_score_early` yang rendah (termasuk 0.0). Hal ini menyebabkan model mengklasifikasikan mereka sebagai berisiko tinggi, meskipun pada akhirnya mereka berhasil lulus.

- **Implikasi Intervensi:** Kelompok ini dapat dianggap sebagai "Kritis" atau *at-risk* yang berhasil *rebound*. Meskipun prediksi hasil akhir keliru, EWS berhasil mengidentifikasi sinyal bahaya awal (misalnya, skor rendah). Ini mengindikasikan bahwa dosen wali atau sistem pendukung **harus memprioritaskan** perhatian dan bimbingan kepada mahasiswa yang terdeteksi dengan profil serupa di awal perkuliahan. Keberhasilan akhir mereka mungkin disebabkan oleh intervensi yang sudah ada, ketahanan pribadi, atau faktor lain yang tidak terukur di 30 hari pertama. EWS dalam hal ini berfungsi memvalidasi kebutuhan intervensi dini bagi mereka.

2. Kasus False Positive (Kepercayaan Semu):

- **Definisi:** Mahasiswa yang sebenarnya (aktual) Gagal/Mundur, namun diprediksi 'Lulus' oleh EWS.
- **Jumlah:** Terdapat **1040** kasus *False Positive* dalam data uji (lihat **Gambar 4.2**).
- **Analisis:** Analisis sampel kasus ini dari notebook ([*execution_count*: 17], bagian D) seringkali menunjukkan *avg_score_early* yang relatif tinggi. Performa awal yang baik ini kemungkinan membuat model "terkecoh" dan memprediksi mereka akan lulus, padahal kenyataannya mereka gagal atau mengundurkan diri di kemudian hari.
- **Implikasi Intervensi:** Kelompok ini mewakili tantangan "Kepercayaan Semu" (*false sense of security*). Risiko akademik mereka tidak terdeteksi oleh EWS berbasis data 30 hari pertama. Kegagalan mereka kemungkinan besar dipicu oleh faktor-faktor yang muncul setelah periode awal tersebut, seperti kesulitan memahami materi tingkat lanjut, masalah pribadi, penurunan motivasi, atau *burnout*.
- **Rekomendasi:** Hasil ini menunjukkan keterbatasan EWS yang hanya berfokus pada periode sangat awal. Untuk menangkap risiko pada kelompok ini, **pemantauan kinerja dan keterlibatan mahasiswa secara berkelanjutan** sangat diperlukan. Institusi sebaiknya mempertimbangkan implementasi EWS *multi-tahap*, misalnya dengan melakukan prediksi ulang pada titik waktu strategis berikutnya (contoh: akhir kuartal pertama atau pertengahan semester) menggunakan data aktivitas yang lebih baru untuk mengidentifikasi mahasiswa yang awalnya berkinerja baik namun mulai menunjukkan tren penurunan.

BAB 5

KESIMPULAN DAN SARAN

Bab ini menyajikan rangkuman kesimpulan yang ditarik dari hasil analisis dan pembahasan yang telah dipaparkan, serta merumuskan saran praktis dan rekomendasi untuk penelitian selanjutnya berdasarkan temuan dan batasan studi.

5.1 Kesimpulan

Berdasarkan analisis data dan evaluasi model yang telah dilakukan, dapat ditarik beberapa kesimpulan utama sebagai berikut:

1. **Prediksi Risiko Sangat Dini yang Feasible:** Penelitian ini berhasil menunjukkan bahwa prediksi risiko kegagalan studi mahasiswa (status Gagal/Mundur) secara *sangat dini*, yaitu dalam 30 hari pertama perkuliahan, adalah *feasible* dan dapat dicapai dengan tingkat akurasi yang memadai.
2. **Keunggulan Model XGBoost:** Model *Early Warning System* (EWS) yang dikembangkan menggunakan algoritma XGBoost, setelah melalui proses optimasi *hyperparameter*, menunjukkan kinerja terbaik dibandingkan model *baseline* Random Forest pada dataset OULAD. Model XGBoost mencapai akurasi 71,04% dan AUC-ROC 0.7111 pada data uji, dengan kemampuan identifikasi kelas 'Gagal/Mundur' (Recall) sebesar 0.70.
3. **Identifikasi Faktor Risiko Kunci:** Analisis interpretabilitas menggunakan SHAP secara signifikan mengidentifikasi bahwa rata-rata skor awal (*avg_score_early*), durasi partisipasi awal (*durasi_partisipasi_early*), dan jumlah hari aktif awal (*total_hari_aktif_early*) merupakan faktor prediktor paling berpengaruh terhadap kemungkinan mahasiswa diprediksi Gagal/Mundur dalam 30 hari pertama. Kinerja dan keterlibatan awal terbukti menjadi indikator kunci.
4. **EWS yang Interpretable dan Aplikatif:** Kombinasi model XGBoost dengan analisis SHAP menghasilkan EWS yang tidak hanya prediktif tetapi juga *interpretable*. Kemampuan untuk menjelaskan faktor-faktor pendorong prediksi memberikan landasan berbasis bukti yang kuat bagi institusi pendidikan untuk merancang dan membenarkan strategi intervensi proaktif.

5.2 Saran

Berdasarkan kesimpulan dan analisis kesalahan model, berikut adalah beberapa saran yang dapat dipertimbangkan:

1. Untuk Institusi Pendidikan / Akademik:

- Implementasi EWS Dini: Institusi disarankan mempertimbangkan pengembangan dan implementasi sistem EWS serupa yang fokus pada data 30 hari pertama sebagai alat bantu strategis untuk identifikasi proaktif mahasiswa berisiko.
- Prioritisasi Intervensi Berbasis Bukti: Hasil SHAP menunjukkan pentingnya skor dan keterlibatan awal. Intervensi (seperti bimbingan akademik intensif, program tutorial, atau *workshop* strategi belajar) sebaiknya diprioritaskan bagi mahasiswa yang terdeteksi berisiko oleh EWS, terutama yang memiliki nilai rendah pada fitur-fitur kunci tersebut. Perhatian juga perlu diberikan pada kelompok *False Negative* yang mungkin membutuhkan dukungan awal.
- Pemantauan Berkelanjutan & EWS Multi-Tahap: Mengingat adanya kasus *False Positive* (mahasiswa yang awalnya tampak baik namun akhirnya gagal), EWS berbasis 30 hari sebaiknya dilengkapi dengan sistem pemantauan berkelanjutan. Pertimbangkan implementasi EWS multi-tahap (misalnya, prediksi ulang di pertengahan semester) untuk menangkap dinamika perubahan risiko.

2. Untuk Penelitian Selanjutnya:

- Eksplorasi Algoritma dan Fitur: Penelitian mendatang dapat mengeksplorasi algoritma lain (misal, *deep learning*) atau melakukan *feature engineering* yang lebih mendalam, termasuk integrasi data non-LMS (jika tersedia) untuk potensi peningkatan akurasi.
- Analisis Jendela Waktu: Menganalisis dampak penggunaan data dari jendela waktu yang berbeda (misal, 15 hari, 60 hari) terhadap akurasi dan *lead time* intervensi.
- Validasi Eksternal: Menguji generalisasi model pada dataset dari institusi pendidikan lain dengan konteks dan karakteristik mahasiswa yang berbeda.
- Studi Dampak Intervensi: Melakukan studi lanjutan untuk mengukur efektivitas intervensi yang dirancang berdasarkan output EWS terhadap tingkat retensi dan keberhasilan mahasiswa.
- Penyebab Risiko Lanjutan: Menyelidiki lebih lanjut faktor-faktor penyebab kegagalan studi pada kelompok *False Positive* yang tidak terdeteksi oleh EWS tahap awal.

Daftar Pustaka

- [1] Hu, Y. H., Lo, C. L., & Shih, S. P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, 36, 469-478.
- [2] Aguiar, E., et al. (2015). A machine learning approach to identify students with risk of dropping out: A case study. *2015 International Conference on Learning and Teaching in Computing and Engineering*.
- [3] Zacharis, N. Z. (2015). A multivariate approach to predicting student performance in distance education. *The International Review of Research in Open and Distributed Learning*, 16(2).
- [4] Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- [5] Lu, O. H., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying learning analytics for the early prediction of Students' academic performance in blended learning. *Educational Technology & Society*, 21(2), 220-232.
- [6] Open Data Jabar. (2021). *Alasan Mahasiswa Drop Out atau Putus Kuliah*. Diakses dari <https://opendata.jabarprov.go.id/id/infografik/alasan-mahasiswa-drop-out-atau-putus-kuliah,-apakah-gara-gara-skripsi-susah> Diakses 22 Okt. 2025.
- [7] L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an 'early warning system' for educators: A proof of concept," *Computers & Education*, vol. 54, pp. 588–599, 2010.
- [8] J. Y. Chung and S. Lee, "Dropout early warning systems for high school students using machine learning," *Children and Youth Services Review*, vol. 96, pp. 346–353, 2019.
- [9] B. S. C. Putra, I. Tahyudin, B. A. Kusuma, and K. N. Isnaini, "Efektivitas Algoritma Random Forest, XGBoost, dan Logistic Regression dalam Prediksi Penyakit Paru-paru," *Techno.COM*, vol. 23, no. 4, pp. 909–922, 2024.
- [10] D. Tarwidi, S. R. Pudjaprasetya, D. Adytia, and M. Apri, "An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach," *MethodsX*, vol. 10, 102119, 2023.
- [11] G. Airlangga, "Comparative Study of XGBoost, Random Forest, and Logistic Regression Models for Predicting Customer Interest in Vehicle Insurance," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 4, pp. 2542–2545, 2024.
- [12] D. Hooshyar and Y. Yang, "Problems With SHAP and LIME in Interpretable AI for Education: A Comparative Study of Post-Hoc Explanations and Neural-Symbolic Rule Extraction," *IEEE Access*, vol. 12, pp. 137472–137484, 2024.
- [13] Kuzilek, J., Hlostá, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data*, 4, art. no. 170171.
- [14] Lakkaraju, H., et al. (2015). A machine learning framework to identify students at risk of dropping out. *ACM Trans. Intell. Syst. Technol.*
- [15] Hakkal, S. (2024). XGBoost To Enhance Learner Performance Prediction. *International Journal / Elsevier article*.

[16] Gunasekara, S. (2024). Explainability in Educational Data Mining and Learning Analytics: an umbrella review. *EDM Conf. Proc. / poster*.

[17] Abdullah, M. (2023). Predicting students' academic performance using e-learning logs. *Int. J. Artificial Intelligence / conference/journal*.