

Prediksi Dropout Mahasiswa: *Early-Warning* Berbasis *Enrollment* dengan *Machine Learning*

Febri Andika Putra^{1*}, Syahisro Mirajdandi², Nandra³, Bisma Okmarizal⁴, Sandy Mulyanda⁵

¹Program Studi Teknik Informatika, STMIK Citra Mandiri Padangsidempuan, Indonesia

^{2,3}Program Studi Manajemen Informatika, Politeknik Lembaga Pendidikan dan Pengembangan Profesi Indonesia, Indonesia

⁴Program Studi Pendidikan Agama Islam, Universitas Islam Negeri Jurai Siwo Lampung, Indonesia

⁵Program Studi sistem informasi, Institute Bisnis dan Teknologi Pelita Indonesia

¹febriandika9725@gmail.com*, ²syahisromirajdandi@gmail.com, ³nandra997@gmail.com,

⁴bisma.okmarizal@uinjusila.ac.id, ⁵sandy.mulyanda@lecturer.pelitaindonesia.ac.id

Abstract

Dropout among university students remains a major challenge in higher education because it affects study continuity, institutional performance, and the efficiency of academic service planning. This study develops a machine learning-based Early Warning System (EWS) that leverages data available at enrollment and is updated after the first semester. Using the public dataset "Predict Students' Dropout and Academic Success" (n = 4,424), the original three-class outcome (Dropout, Enrolled, Graduate) is simplified into a binary target, with dropout treated as the positive class. Two feature scenarios are evaluated: (1) enrollment-only for pre-entry screening and (2) enrollment plus first-semester indicators to update risk scores. Three models are compared: class-balanced Logistic Regression, class-balanced Random Forest, and Gradient Boosting. Model performance is assessed using accuracy, precision/recall/F1score for the dropout class, balanced accuracy, and ROC-AUC. Under the enrollment-only setting, Logistic Regression achieves the best early-warning performance (recall = 0.697; F1 score = 0.651). After incorporating first-semester features, performance improves (recall = 0.792; F1 score = 0.779). Beyond model comparison, this study adds an operational perspective through confusion-matrix simulation and probability-threshold analysis to balance missed at-risk cases and follow-up workload.

Keywords: student dropout; early warning system; machine learning; enrollment data; binary classification

Abstrak

*Dropout mahasiswa masih menjadi tantangan penting di perguruan tinggi karena berdampak pada keberlanjutan studi, kinerja institusi, serta efisiensi perencanaan layanan akademik. Penelitian ini mengembangkan *Early Warning System (EWS)* berbasis *machine learning* yang memanfaatkan data yang tersedia sejak pendaftaran (*enrollment*) dan diperbarui setelah semester pertama. Menggunakan *dataset* publik "*Predict Students' Dropout and Academic Success*" (n=4.424), keluaran tiga kelas (*Dropout, Enrolled, Graduate*) disederhanakan menjadi target biner dengan *dropout* sebagai kelas positif. Dua skenario fitur diuji: (1) *enrollment-only* untuk skrining pra-kuliah dan (2) *enrollment + indikator semester 1* untuk pembaruan skor risiko. Tiga model dibandingkan, yaitu *Logistic Regression (class-balanced)*, *Random Forest (class-balanced)* dan *Gradient Boosting*. Evaluasi menggunakan *accuracy, precision/recall/F1score* untuk kelas *dropout, balanced accuracy*, dan *ROC-AUC*. Pada *enrollment-only*, *Logistic Regression* memberikan performansi *early warning* terbaik (*recall* 0,697; *F1 score* 0,651). Setelah penambahan fitur semester 1, performa meningkat (*recall* 0,792; *F1 score* 0,779). Selain membandingkan model, studi ini menambahkan perspektif operasional melalui simulasi *confusion matrix* dan analisis *threshold* probabilitas untuk menyeimbangkan kasus berisiko yang terlewat dan beban tindak lanjut.*

Kata kunci: dropout mahasiswa; early warning system; machine learning; data enrollment; klasifikasi biner

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License

1. Pendahuluan

*Dropout mahasiswa merupakan permasalahan strategis pada pendidikan tinggi karena memengaruhi keberlanjutan studi, capaian indikator kinerja institusi (retensi dan kelulusan), serta efisiensi perencanaan layanan akademik. Dampak *dropout* tidak hanya dirasakan oleh mahasiswa (biaya, waktu, dan konsekuensi sosial), tetapi juga oleh institusi karena berpengaruh pada evaluasi mutu, akreditasi, dan perencanaan kapasitas layanan (dosen wali, remedial,*

*beasiswa, serta konseling) [1] [2]. Dalam banyak kasus, risiko *dropout* bersifat multifaktor dan saling berinteraksi, mencakup faktor akademik (kinerja awal, keterlibatan belajar), administrasi (status registrasi, kepatuhan administratif), kondisi finansial, dan faktor sosial/psikologis [1] [2]. Namun, pada banyak institusi, identifikasi mahasiswa berisiko masih dilakukan secara manual dan cenderung terlambat, sehingga intervensi (konseling, remedial, atau dukungan finansial) sering diberikan ketika gejala sudah memburuk. Penelitian ini berangkat dari kebutuhan menyediakan mekanisme*

skrining awal berbasis data yang sudah tersedia sejak pendaftaran (*enrollment*) untuk menandai mahasiswa berisiko *dropout* secara lebih dini.

Early Warning System (EWS) dibutuhkan untuk mengubah data akademik dan administratif menjadi sinyal risiko yang mudah dipahami sehingga pembimbing akademik dapat menentukan prioritas pendampingan secara lebih cepat dan objektif. Berbeda dengan evaluasi “tradisional” yang sering reaktif (baru bertindak setelah masalah besar terlihat), EWS bersifat proaktif: sistem berupaya mendeteksi gejala awal sebelum mahasiswa benar-benar berhenti studi [2] [3]. Pada konteks ini, performa model tidak cukup dinilai dari *accuracy*, karena *accuracy* dapat “menipu” pada data tidak seimbang (kelas *dropout* lebih sedikit). EWS lebih menekankan kemampuan menangkap kasus *dropout* (*recall*) sekaligus mengelola beban tindak lanjut yang timbul akibat *false alarm* (*false positive*). Oleh karena itu, evaluasi perlu menggunakan metrik yang berorientasi pada kelas *dropout* serta mempertimbangkan *confusion matrix* sebagai gambaran biaya operasional [2] [3]. Dalam studi ini, *confusion matrix* digunakan bukan sebagai “metrik pengganti”, melainkan sebagai ringkasan hasil klasifikasi yang menjadi dasar perhitungan *precision/recall/F1score* sekaligus menerjemahkan performa model menjadi konsekuensi operasional: berapa mahasiswa yang perlu ditindaklanjuti (TP+FP) dan berapa yang berpotensi terlewat (FN).

Sejumlah studi menunjukkan bahwa informasi akademik awal (misalnya capaian semester pertama) berkontribusi besar untuk prediksi *dropout*. Namun, pendekatan yang mengandalkan data beberapa semester sering dinilai kurang “dini” untuk intervensi, karena ketika sinyal akademik yang kuat muncul, sebagian mahasiswa sudah berada pada tahap kritis [1] [3] [4]. Di sisi lain, data pendaftaran (*enrollment*) tersedia sejak awal kuliah, tetapi sinyal prediktifnya cenderung lebih lemah sehingga dibutuhkan pendekatan yang realistis tentang *trade-off* antara ketepatan prediksi dan waktu ketersediaan data [1] [3] [4]. Studi terkini bahkan menekankan pentingnya prediksi sejak hari pertama perkuliahan berbasis informasi pendaftaran [5], sementara kajian empiris lain menunjukkan bahwa tambahan indikator akademik dan keterlibatan belajar dapat memperkuat prediksi seiring berjalannya semester [3]. Namun, sebagian pendekatan terdahulu baru memanfaatkan indikator performa setelah beberapa semester, sehingga sinyal risiko muncul ketika peluang intervensi dini semakin sempit. Fokus pada skenario *enrollment-only* yakni memanfaatkan data yang tersedia saat mahasiswa baru masuk masih relatif lebih jarang, padahal justru dibutuhkan untuk skrining pra-kuliah.

Meskipun literatur prediksi *dropout* terus berkembang, masih terdapat dua celah yang relevan untuk implementasi EWS di perguruan tinggi. Pertama, studi yang benar-benar membatasi fitur pada data yang tersedia saat pendaftaran relatif lebih sedikit, padahal

jenis data ini paling mudah diintegrasikan ke SIAKAD untuk skrining awal. Kedua, banyak penelitian berhenti pada metrik statistik tanpa mengaitkan keluaran model dengan konsekuensi operasional, misalnya berapa banyak mahasiswa yang harus dihubungi (*workload*), berapa banyak *false alarm* yang muncul, dan bagaimana pengaturan *threshold* dapat menyesuaikan daftar prioritas dengan kapasitas layanan pendampingan. Berangkat dari rekomendasi literatur tentang kebutuhan model yang dapat diterapkan dan dievaluasi secara menyeluruh [3] [4], penelitian ini menguji alur kerja dua fase (pra-kuliah dan pasca-semester 1) dengan pra-pemrosesan yang konsisten serta menambahkan simulasi beban intervensi berbasis *confusion matrix* dan analisis *threshold* sebagai jembatan antara performa model dan keputusan operasional kampus. Dengan demikian, gap yang hendak dijawab adalah bagaimana merancang EWS yang tetap berguna pada fase sangat awal (*enrollment-only*), sekaligus menunjukkan dampak praktisnya melalui pengaturan *threshold* dan estimasi beban tindak lanjut berbasis *confusion matrix*.

Permasalahan yang diselesaikan pada penelitian ini adalah bagaimana memprediksi risiko *dropout* sejak fase pendaftaran menggunakan *machine learning*, dan bagaimana keluaran model tersebut diterjemahkan menjadi daftar prioritas intervensi yang realistis sesuai kapasitas kampus.

Berdasarkan latar belakang tersebut, penelitian ini merumuskan prediksi *dropout* sebagai klasifikasi biner untuk kebutuhan *early warning* dan menyusun EWS dua fase: (1) skrining pra-kuliah berbasis *enrollment-only* dan (2) pembaruan skor risiko setelah semester 1 menggunakan indikator akademik awal. Secara khusus, tujuan penelitian adalah: (a) membandingkan *Logistic Regression*, *Random Forest*, dan *Gradient Boosting* pada kedua fase dengan pipeline yang sama agar perbandingan adil; (b) mengukur besarnya peningkatan kinerja ketika data semester 1 tersedia; dan (c) menunjukkan cara menerjemahkan keluaran model ke keputusan operasional melalui penetapan *threshold* serta simulasi beban tindak lanjut. Dengan demikian, hasil penelitian diharapkan tidak hanya memberikan angka performa, tetapi juga panduan praktis untuk mengoperasikan EWS sesuai kapasitas layanan dan kebutuhan intervensi kampus.

Kontribusi utama penelitian ini adalah:

1. Menyajikan desain EWS dua fase yang selaras dengan ketersediaan data (pra-kuliah dan pembaruan setelah semester 1).
2. Membandingkan tiga model yang umum dipakai pada data tabular dengan fokus pada kelas *dropout* sebagai target *early warning*.
3. Menambahkan perspektif operasional melalui *threshold* dan *confusion matrix* untuk menghubungkan keluaran model dengan beban intervensi kampus. Menekankan interpretabilitas dan *monitoring* sederhana (data *drift* dan potensi bias) agar EWS lebih transparan dan berkelanjutan.

2. Metode Penelitian

2.1 Dataset dan transformasi target dataset ini tersedia secara terbuka melalui UCI *Machine Learning Repository* sehingga dapat direplikasi. Tidak ada proses seleksi atau pembuangan baris data; seluruh 4.424 rekam digunakan. Transformasi target dilakukan dengan menggabungkan *Enrolled* dan *Graduate* sebagai *non-dropout* (0) dan mempertahankan *dropout* sebagai kelas positif (1) agar sesuai kebutuhan operasional EWS.

Penelitian menggunakan dataset publik “*Predict Students’ Dropout and Academic Success*” (n=4.424) yang tersedia dipublikasikan dalam studi realinho dkk. [1]. Label pada dataset awalnya terdiri dari tiga kelas, yaitu *Dropout*, *Enrolled*, dan *Graduate*. Agar selaras dengan kebutuhan *early warning* yang operasional (prioritas risiko), label tersebut ditransformasikan menjadi target biner dengan menetapkan *dropout* sebagai kelas positif (1), sedangkan *non-dropout* (0) merupakan gabungan kelas *enrolled* dan *graduate*. Dengan pemetaan ini, keluaran model dapat langsung dipakai untuk skrining “berisiko *dropout*” versus “tidak berisiko *dropout*” pada fase awal masuk kuliah. Ketidakseimbangan kelas berpotensi membuat model tampak “baik” jika hanya mengoptimalkan *accuracy*. Untuk itu, penelitian ini menggunakan *stratified split* agar proporsi kelas terjaga pada data latih dan uji, serta menekankan metrik yang sensitif pada kelas positif (*recall*, *F1*, dan *AUC/PR*) dan analisis *threshold*. Pendekatan lain seperti pembobotan kelas (*class weight*) atau resampling dapat dipertimbangkan pada studi lanjutan untuk mengurangi bias terhadap kelas minoritas.

Distribusi kelas menunjukkan 1421 data *dropout* (32,12%) dan 3003 data *non-dropout* (67,88%), sehingga dataset bersifat tidak seimbang. Ketidakseimbangan ini penting karena model yang hanya mengoptimalkan *accuracy* berisiko menurunkan sensitivitas terhadap kelas *dropout* dan meningkatkan *false negative* (FN), yaitu mahasiswa yang seharusnya terdeteksi berisiko tetapi tidak masuk daftar tindak lanjut. Untuk mengurangi bias terhadap kelas mayoritas, penelitian ini menerapkan pembobotan kelas pada tahap pelatihan, yaitu *class_weight='balanced'* pada *Logistic Regression* dan *Random Forest*, sedangkan *Gradient Boosting* menggunakan pengaturan default. Penerapan *class_weight='balanced'* membuat proses pelatihan tidak terlalu “nyaman” pada kelas *non-dropout*, karena kesalahan pada kasus *dropout* diberi bobot lebih besar sehingga model terdorong lebih peka mendeteksi mahasiswa berisiko. Selain itu, evaluasi tidak hanya mengandalkan *accuracy*, tetapi juga menekankan *Recall*, *Precision*, *F1-score*, serta *confusion matrix* agar *trade-off FP/FN* dapat diamati secara jelas dalam konteks operasional *early warning system*.

2.2 Skenario Fitur untuk *Early Warning*

Untuk mensimulasikan prediksi bertahap berdasarkan waktu ketersediaan data di perguruan tinggi, penelitian menggunakan dua skenario fitur. Skenario pertama adalah *enrollment-only*, yaitu hanya menggunakan atribut yang tersedia saat pendaftaran, sehingga seluruh atribut yang berkaitan dengan semester 1 dan semester 2 dikeluarkan. Skenario kedua adalah *enrollment + semester 1*, yaitu menggabungkan atribut *enrollment* dengan indikator akademik semester 1, sementara atribut semester 2 tetap dikeluarkan agar tidak terjadi kebocoran informasi masa depan.

Skenario *enrollment-only* merepresentasikan skrining pra-kuliah yang bertujuan memberikan sinyal risiko sedini mungkin, sedangkan *enrollment + semester 1* merepresentasikan pembaruan skor risiko setelah semester pertama ketika sinyal akademik awal mulai terlihat. Desain dua skenario ini digunakan untuk menilai *trade-off* antara ketepatan prediksi dan waktu ketersediaan data dalam implementasi EWS [3] [4].

2.3 Pra-pemrosesan dan Pembagian Data

Pra-pemrosesan dilakukan secara konsisten di dalam pipeline agar perbandingan antarmodel tetap adil: (1) pemeriksaan nilai hilang dan imputasi sederhana bila diperlukan (median untuk fitur numerik dan modus untuk fitur kategorikal), (2) pengkodean fitur kategorikal dengan *one-hot encoding*, (3) standarisasi fitur numerik (*StandardScaler*) untuk *Logistic Regression*, sedangkan model berbasis pohon tidak memerlukan standarisasi, dan (4) seluruh transformasi di-fit hanya pada data latih untuk mencegah kebocoran informasi, kemudian diterapkan pada data uji.

Data diperlakukan sebagai data tabular dan diproses melalui pipeline pra-pemrosesan yang konsisten agar perbandingan antar model tetap adil. Sebelum pelatihan, dilakukan pemeriksaan kualitas data (nilai hilang dan duplikasi). *Dataset* tidak memiliki nilai hilang sehingga tidak diperlukan imputasi; duplikasi yang ditemukan (jika ada) dihapus agar tidak mengganggu evaluasi. Variabel kategorikal pada dataset sudah disandikan sebagai kode numerik sesuai dokumentasi, sehingga diperlakukan sebagai fitur tabular pada eksperimen ini. Pembagian data dilakukan menggunakan *stratified split* dengan proporsi 80% data latih dan 20% data uji. Parameter *random_state=42* digunakan sebagai *seed* acak agar pembagian dapat direplikasi; nilai *seed* bersifat konvensi (arbitrer) dan dapat diganti, namun seluruh hasil pada makalah ini merujuk pada konfigurasi tersebut. Stratifikasi menjaga proporsi kelas *dropout* dan *non-dropout* pada data latih maupun data uji sehingga evaluasi tidak bias akibat perubahan distribusi kelas.

Nilai *random_state=42* dipilih sebagai penanda (*seed*) yang bersifat arbitrer untuk memastikan pembagian data dapat direplikasi. Jika nilai *seed* diubah, susunan data latih–uji dapat berbeda sehingga nilai metrik berpotensi sedikit berubah; oleh karena itu, validasi silang (*cross-validation*) atau pengulangan beberapa

seed disarankan pada pengembangan berikutnya untuk memperoleh estimasi performa yang lebih stabil. Standardisasi *z-score* diterapkan pada *Logistic Regression* karena model ini sensitif terhadap skala fitur, sehingga perbedaan rentang nilai antar fitur tidak mendistorsi proses pembelajaran. Skaler dipasang (fit) hanya pada data latih lalu diterapkan pada data uji untuk mencegah kebocoran informasi. Sementara itu, pada *random forest* dan *gradient boosting* standardisasi tidak bersifat wajib karena kedua model tersebut relatif lebih robust terhadap perbedaan skala fitur. Namun, alur pemrosesan tetap dijaga konsisten agar evaluasi perbandingan lebih valid dan perbedaan kinerja antarmodel dapat diatribusikan pada karakteristik model, bukan pada perbedaan pra-pemrosesan data.

2.4 Model Klasifikasi yang Dibandingkan

Penelitian ini berfokus pada pembangunan *Early Warning System* (EWS) berbasis data *enrollment* untuk memprediksi risiko *dropout* sebagai klasifikasi biner. Untuk menjawab pertanyaan penelitian tentang model yang paling sesuai pada fase skrining awal dan fase pembaruan risiko setelah semester 1, studi ini melakukan komparasi tiga model yang lazim digunakan pada data tabular, yaitu *Logistic Regression*, *Random Forest*, dan *Gradient Boosting*. Komparasi dilakukan bukan hanya untuk mencari akurasi tertinggi, tetapi untuk memilih model yang seimbang antara kemampuan menangkap mahasiswa berisiko (*recall*), beban tindak lanjut (*precision*), dan kemudahan implementasi di lingkungan kampus.

Logistic Regression dipilih sebagai model baseline karena sederhana, cepat, dan relatif mudah dijelaskan kepada pemangku kepentingan, sehingga sering digunakan pada konteks *early warning* yang menuntut interpretabilitas. *Random Forest* digunakan untuk menangkap relasi non-linear dan interaksi antarfitur yang mungkin muncul pada data administrasi maupun akademik, sementara *Gradient Boosting* dipilih karena dikenal kompetitif pada data tabular melalui mekanisme pembelajaran bertahap yang memperbaiki kesalahan model sebelumnya. Untuk mengurangi dampak ketidakseimbangan kelas, penyeimbangan berbasis bobot kelas (*class weight*) diterapkan pada model yang mendukung sehingga model lebih peka terhadap kelas *dropout*. Dengan demikian, fokus penelitian ini memang bersifat komparatif, dan hasil komparasi digunakan untuk merekomendasikan konfigurasi EWS dua fase yang paling realistis dioperasikan di kampus. [6] [7] [8]. Studi pada Jurnal Fasilkom juga melaporkan penerapan machine learning untuk prediksi kelulusan mahasiswa, sehingga pendekatan prediktif relevan sebagai dasar pemantauan akademik. [9] [10].

2.5 Metrik Evaluasi

Evaluasi performa model dilakukan menggunakan *accuracy*, *precision*, *recall*, dan *F1-score* untuk kelas positif *dropout*, serta *balanced accuracy* dan *ROC-AUC*. *Accuracy* memberi gambaran umum, *balanced*

accuracy mengurangi bias akibat dominasi kelas mayoritas, sedangkan *ROC-AUC* bersifat tidak bergantung pada threshold sehingga menggambarkan kemampuan model dalam melakukan pemeringkatan risiko. Namun, karena keputusan EWS bersifat operasional pada data yang tidak seimbang, interpretasi utama tetap menekankan trade-off antara *precision* dan *recall* (serta *F1-score*) pada kelas *dropout*.

Seluruh metrik tersebut diturunkan dari *confusion matrix* (TP, FP, TN, FN). Dalam studi ini, *confusion matrix* tidak diposisikan sebagai metrik yang berdiri sendiri, melainkan sebagai ringkasan hasil klasifikasi yang menjadi dasar perhitungan metrik evaluasi dan sekaligus jembatan interpretasi operasional: FP merepresentasikan beban tindak lanjut yang tidak perlu (*false alarm*), sedangkan FN merepresentasikan kasus *dropout* yang berpotensi terlewat.

Secara formal, untuk kelas positif *dropout*, definisi metrik adalah sebagai berikut: $Accuracy = (TP + TN) / (TP + TN + FP + FN)$; $Precision = TP / (TP + FP)$; $Recall = TP / (TP + FN)$; $F1-score = 2 \times (Precision \times Recall) / (Precision + Recall)$. *Balanced accuracy* dihitung sebagai $(TPR + TNR) / 2$, dengan $TPR = Recall$ dan $TNR = TN / (TN + FP)$. *ROC-AUC* dihitung sebagai luas di bawah kurva ROC yang memplot TPR terhadap $FPR = FP / (FP + TN)$ pada berbagai nilai threshold.

2.6 Pengaturan Threshold Probabilitas

Selain memilih model terbaik, sistem EWS dapat menyesuaikan threshold keputusan pada probabilitas *dropout*. Pada pelaporan metrik utama penelitian ini, threshold 0,50 (*default*) digunakan untuk menjaga konsistensi perbandingan antar model. Namun, secara operasional threshold tidak harus 0,50: menurunkan threshold cenderung meningkatkan *recall* (lebih banyak mahasiswa terdeteksi berisiko) dengan konsekuensi *false positive* meningkat, sedangkan menaikkan threshold cenderung meningkatkan *precision* tetapi berisiko menaikkan *false negative* (kasus berisiko terlewat).

Acuan penentuan threshold dapat ditetapkan berdasarkan kebutuhan institusi, misalnya: (1) berbasis kapasitas pendampingan (memilih threshold sehingga jumlah mahasiswa yang ter-flag sesuai kuota tindak lanjut per periode); (2) berbasis target kinerja minimum (misalnya menetapkan *recall* minimum dengan batas *precision* tertentu); atau (3) berbasis fungsi biaya/utility yang membobotkan konsekuensi FP dan FN sesuai kebijakan kampus. Praktik ini dapat dipandu melalui kurva *precision-recall* atau evaluasi beberapa threshold (misalnya 0,30–0,70) untuk melihat *trade-off* cakupan deteksi dan beban intervensi.

2.7 Rancangan Implementasi Early Warning dan Replikasi Eksperimen

Penelitian dirancang agar *pipeline* eksperimen mudah direplikasi dan dapat diintegrasikan ke Sistem Informasi Akademik sebagai modul EWS. Keluaran

model berupa probabilitas *dropout* diperlakukan sebagai skor risiko yang dihitung secara otomatis dari data mahasiswa yang telah diproses. Skor ini dapat dihasilkan sejak pra-kuliah menggunakan skenario *enrollment-only* dan kemudian diperbarui setelah semester 1 menggunakan skenario *enrollment + semester 1* [3] [4]. Dengan rancangan dua fase tersebut, institusi dapat mengoperasikan EWS sebagai proses berkelanjutan yang mengikuti alur kehidupan akademik mahasiswa, mulai dari tahap awal masuk hingga penyelesaian studi.

Dalam implementasi praktis, skor risiko dapat dipetakan ke dalam kategori tertentu, misalnya risiko tinggi, sedang, dan rendah, yang kemudian dikaitkan dengan prosedur tindak lanjut yang berbeda tingkat intensitasnya. Risiko tinggi dapat memicu intervensi berupa konseling intensif dan pemantauan berkala, risiko sedang dapat diarahkan pada program penguatan akademik atau pengingat administratif, sedangkan risiko rendah cukup dipantau secara pasif. Dokumentasi pipeline dan parameter yang digunakan juga disusun sedemikian rupa sehingga eksperimen dapat direplikasi di institusi lain atau diperbarui ketika tersedia data angkatan baru, sehingga EWS tidak hanya berhenti pada tahap penelitian tetapi dapat berkembang menjadi sistem pendukung keputusan yang berkelanjutan [3] [4].

3. Hasil dan Pembahasan

3.1 Distribusi Kelas dan Implikasi Ketidakseimbangan

Dataset terdiri dari 4.424 rekam mahasiswa dengan proporsi *dropout* 32,12% dan *non-dropout* 67,88%. Ketidakseimbangan kelas ini berpotensi membuat model memperoleh *accuracy* tinggi hanya dengan sering memilih kelas mayoritas, sehingga *accuracy* saja tidak cukup untuk merepresentasikan kinerja EWS yang bertujuan menangkap mahasiswa berisiko.

Untuk mengurangi dampak ketidakseimbangan kelas, penelitian menerapkan pembagian data secara *stratified* agar proporsi kelas terjaga pada data latih dan uji, serta menggunakan penyeimbangan bobot kelas (*class_weight*) pada model yang mendukung. Evaluasi difokuskan pada metrik yang sensitif terhadap kelas minoritas, terutama *recall* dan *F1-score* kelas *dropout*, sementara *precision* juga dipertimbangkan karena berkaitan langsung dengan beban tindak lanjut (FP) dan efisiensi intervensi. Dengan kombinasi metrik ini, model dapat dievaluasi secara lebih adil pada data yang tidak seimbang.

Tabel 1. Distribusi kelas target (*dropout* vs *non-dropout*).

Kelas	Jumlah	Persentase
Dropout (1)	1421	32.12%
Non-dropout (0)	3003	67.88%
Total	4424	100.00%

3.2 Kinerja Model pada Skenario *Enrollment-only*

Skenario *enrollment-only* mengevaluasi prediksi pra-kuliah menggunakan informasi yang tersedia sejak

pendaftaran. Hasil menunjukkan *Logistic Regression* memiliki profil *early warning* paling kuat untuk fase awal, terutama pada *recall* dan F1 kelas *dropout*.

Pada skenario *enrollment-only*, *Logistic Regression* mencapai *recall dropout* 0,697 dan F1 0,651, sementara *Random Forest* memiliki *precision* lebih tinggi (0,767) tetapi *recall* lebih rendah (0,475). *Gradient Boosting* menunjukkan pola yang mirip dengan *Random Forest* (*precision* 0,762; *recall* 0,486). Secara operasional, temuan ini menunjukkan adanya trade-off: *Logistic Regression* cenderung menghasilkan cakupan deteksi lebih luas (lebih sedikit kasus berisiko yang terlewat), namun dengan false alarm yang lebih banyak; sedangkan model ensemble lebih selektif (*precision* lebih tinggi) tetapi berisiko melewatkan sebagian kasus *dropout*.

Nilai *accuracy* dan *ROC-AUC* pada ketiga model relatif tinggi, yang mengindikasikan model mampu melakukan pemeringkatan risiko secara cukup baik. Namun, pada data tidak seimbang, *accuracy* dapat terlihat baik meskipun sensitivitas terhadap kelas *dropout* rendah, sehingga interpretasi utama tetap berbasis *recall/F1-score* dan dilengkapi *precision* untuk mengontrol beban intervensi. Secara intuitif, pada fase pra-kuliah sinyal prediktif masih terbatas pada variabel administrasi awal; dengan kondisi ini, model linear yang teratur (dan dibantu penyeimbangan bobot kelas) dapat lebih stabil dalam menangkap pola umum risiko, sedangkan model non-linear cenderung lebih konservatif dalam memberi label positif ketika informasi yang tersedia belum kaya.

Tabel 2. Kinerja model pada skenario *enrollment-only* (kelas positif: *dropout*)

Model	Accur acy	Precis ion	Rec all	F1	Balance dAcc	RO C- AU C
<i>LogReg</i> (<i>balanced</i>)	0.760	0.611	0.6 97	0.6 51	0.744	0.8 14
<i>RandomFo</i> <i>rest</i> (<i>balanced</i>)	0.785	0.767	0.4 75	0.5 87	0.704	0.8 22
<i>GradientBo</i> <i>osting</i>	0.786	0.762	0.4 86	0.5 94	0.707	0.8 35

3.3 Kinerja Model pada Skenario *Enrollment + Semester 1*

Skenario kedua menambahkan indikator akademik semester 1 sehingga sinyal risiko akademik awal mulai terlihat. Pada fase ini, semua model mengalami peningkatan kinerja, yang menguatkan pentingnya pembaruan skor risiko setelah semester 1.

Pada skenario *enrollment + semester 1*, seluruh model mengalami peningkatan kinerja karena penambahan indikator akademik awal memperkaya sinyal risiko (misalnya capaian nilai, beban studi, atau indikator keterlibatan belajar pada semester pertama). Pada fase ini, *Logistic Regression* mencapai *recall dropout* 0,792 dan F1 0,779. *Random Forest* memberikan *precision* tertinggi (0,850) yang berarti false alarm lebih sedikit,

sedangkan *ROC-AUC* ketiga model berada pada kisaran tinggi (LogReg 0,907; RF 0,910; GB 0,909). Peningkatan ini konsisten dengan intuisi EWS: setelah semester 1, variabel akademik mulai merekam dinamika risiko, dan model non-linear (seperti *Random Forest* dan *Gradient Boosting*) berpotensi lebih unggul karena mampu menangkap interaksi antarfitur (misalnya kombinasi beban SKS tinggi dengan performa rendah) yang tidak selalu linear.

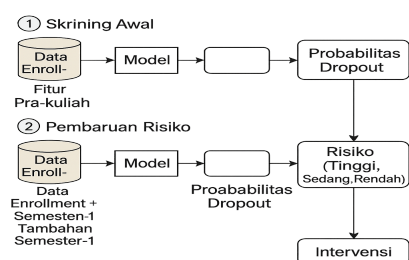
Tabel 3. Kinerja model pada skenario *enrollment + semester 1* (kelas positif: *dropout*)

Model	Accur acy	Preci sion	Rec all	F1	Balance dAcc	RO C- AU C
LogReg (balanced)	0.855	0.765	0.7 92	0.7 79	0.839	0.9 07
RandomFo rest (balanced)	0.858	0.850	0.6 76	0.7 53	0.810	0.9 10
GradientBo osting	0.859	0.822	0.7 15	0.7 65	0.821	0.9 09

3.4 Rekomendasi Implementasi *Early Warning*

Hasil komparasi menunjukkan bahwa pemilihan model untuk EWS sebaiknya disesuaikan dengan tujuan operasional institusi. Pada fase pra-kuliah (*enrollment-only*), *Logistic Regression* cenderung lebih sesuai untuk skrining awal karena recall yang lebih tinggi sehingga mengurangi risiko kasus berisiko yang terlewat. Sebaliknya, *Random Forest/Gradient Boosting* lebih selektif (precision lebih tinggi) dan dapat dipilih ketika kampus memiliki keterbatasan kapasitas tindak lanjut sehingga *false alarm* perlu ditekan. Dengan kata lain, pemilihan model adalah keputusan *trade-off* antara cakupan deteksi (*recall*) dan efisiensi intervensi (*precision*).

Penerapan EWS disarankan menggunakan pendekatan dua fase: (1) skrining pra-kuliah dengan *enrollment-only* untuk membentuk daftar prioritas awal, dan (2) pembaruan skor risiko setelah semester 1 (*enrollment + semester 1*) ketika sinyal akademik mulai tersedia sehingga rekomendasi intervensi dapat dibuat lebih tepat. Selain itu, *threshold* probabilitas perlu ditetapkan sebagai kebijakan operasional; misalnya memilih *threshold* yang menghasilkan jumlah kasus ter-flag sesuai kuota pendampingan, atau menetapkan target minimal recall dengan batas precision tertentu. Pengaturan *threshold* ini membuat EWS fleksibel dan dapat disesuaikan dengan sumber daya institusi.



Gambar 1. Ilustrasi alur kerja EWS dua fase (skrining pra-kuliah dan pembaruan setelah semester 1).

3.5 Keterbatasan

Keterbatasan utama penelitian ini terletak pada strategi evaluasi yang masih menggunakan satu skema *hold-out*, sehingga estimasi performa dapat sensitif terhadap komposisi data latih dan data uji. Walaupun pembagian dilakukan secara *stratified* dan menggunakan *random_state* untuk replikasi, konfigurasi pembagian tertentu dapat menghasilkan estimasi yang lebih optimistis atau lebih pesimistis dibanding kondisi populasi. Untuk mengurangi bias pembagian dan variabilitas, evaluasi selanjutnya disarankan menggunakan *k-fold cross-validation* (misalnya *stratified 5-fold*) atau *multi-seed hold-out*, sehingga metrik dirata-rata pada beberapa lipatan/skenario dan lebih stabil.

Keterbatasan berikutnya adalah ruang lingkup data yang masih mengandalkan satu dataset publik. Perbedaan kebijakan akademik, sistem penilaian, dan karakteristik mahasiswa di institusi lain dapat memengaruhi pola risiko, sehingga generalisasi model perlu diuji melalui validasi eksternal. Selain itu, penelitian ini belum mengukur dampak intervensi secara prospektif; keluaran model baru disimulasikan sebagai daftar prioritas, belum dievaluasi dalam praktik layanan nyata. Karena itu, pemantauan jangka panjang terhadap perubahan pola data (*concept drift*) dan evaluasi dampak intervensi terhadap retensi mahasiswa menjadi agenda penting pada implementasi berikutnya.

Secara ringkas, keterbatasan penelitian ini mencakup: (1) estimasi kinerja yang masih bergantung pada satu pembagian data; (2) keterbatasan cakupan fitur, terutama faktor non-akademik yang berpotensi relevan; (3) belum adanya uji generalisasi lintas institusi/angkatan; dan (4) evaluasi operasional yang belum dilakukan secara eksperimental, sehingga indikator efektivitas intervensi masih bersifat konseptual dan berbasis simulasi.

3.6 Simulasi Beban Intervensi Berdasarkan *Confusion Matrix*

Untuk menerjemahkan kinerja model ke konsekuensi layanan, penelitian ini menggunakan *confusion matrix* sebagai ringkasan jumlah kasus benar-salah pada kelas *dropout*. Pada skenario *enrollment-only* dengan *Logistic Regression*, diperoleh TN=475, FP=126, FN=86, dan TP=198. Artinya terdapat 324 mahasiswa yang ter-flag sebagai berisiko (TP+FP), dengan precision daftar prioritas sebesar 61,1% (198/324) dan *false alarm* sebesar 38,9% (126/324). Secara operasional, 126 *false positive* menggambarkan beban tindak lanjut yang berpotensi tidak perlu; misalnya apabila satu tindak lanjut awal memerlukan 10–15 menit, maka *false positive* dapat menyerap waktu layanan yang seharusnya dialokasikan untuk kasus lain. Sementara itu, 86 *false negative* adalah kasus berisiko yang luput sehingga berpotensi terlambat menerima dukungan. Dengan demikian, *confusion matrix* tidak hanya membantu menghitung *precision/recall*, tetapi

juga memetakan kebutuhan sumber daya dan risiko layanan.

Pada skenario *enrollment + semester 1, confusion matrix* menjadi TN=532, FP=69, FN=59, dan TP=225. Jumlah mahasiswa yang ter-flag berkurang menjadi 294, sementara *false alarm* turun dari 126 ke 69 (sekitar 45%). Precision daftar prioritas meningkat menjadi 76,5% (225/294), dan kasus berisiko yang terlewat (FN) turun 31,4%. Perubahan ini menunjukkan bahwa fitur semester 1 tidak hanya meningkatkan performa statistik, tetapi juga meningkatkan efisiensi layanan: lebih sedikit mahasiswa yang salah sasaran, sehingga kapasitas pendampingan dapat difokuskan pada kasus yang lebih relevan.

Perlu dicatat bahwa beban intervensi sangat dipengaruhi oleh ambang keputusan (*threshold*) yang digunakan. Penurunan *threshold* cenderung menaikkan *recall* (lebih banyak kasus berisiko tertangkap), tetapi juga memperbesar jumlah mahasiswa yang perlu ditindaklanjuti karena *false positive* meningkat. Karena itu, penetapan *threshold* sebaiknya mengikuti kapasitas layanan (misalnya jumlah konselor/dosen wali) dan risiko yang bersedia ditoleransi, sehingga daftar prioritas yang dihasilkan tetap realistis dan dapat ditangani.

3.7 Evaluasi Operasional dan *Monitoring Model*

Keberhasilan EWS tidak hanya diukur dari metrik model, tetapi juga dari indikator operasional yang dapat diaudit. Contoh indikator yang praktis adalah: (1) *coverage rate* yaitu proporsi mahasiswa yang ter-flag sebagai prioritas (TP+FP) dibanding total mahasiswa; (2) *follow-up rate* yaitu proporsi mahasiswa ter-flag yang benar-benar dihubungi/didampingi; (3) tingkat penyelesaian layanan (misalnya konseling/kelas penguatan) dan jenis intervensi yang diberikan; serta (4) perubahan luaran akademik yang relevan, seperti retensi ke semester berikutnya, akumulasi SKS, atau perbaikan IP. Indikator-indikator ini membantu menilai apakah skor risiko yang dihasilkan model benar-benar menghasilkan tindakan nyata dan berdampak.

Dalam konteks layanan, *waktu respons* menjadi kunci karena EWS bertujuan memberi sinyal sedini mungkin. Praktik yang dapat diterapkan adalah menetapkan standar waktu, misalnya kontak awal dilakukan dalam 3-7 hari sejak mahasiswa ter-flag, kemudian tindak lanjut dilakukan sesuai tingkat risiko. Agar pengukuran ini konsisten, kampus perlu menyiapkan mekanisme pencatatan terstruktur: setiap *flag* menyimpan tanggal deteksi, skor risiko, unit penanggung jawab, jenis tindak lanjut, status (selesai/belum), serta tanggal penyelesaian. Dengan log tersebut, evaluasi operasional dapat dilakukan secara rutin melalui laporan bulanan/semesteran tanpa membebani petugas secara administratif.

Selain indikator operasional, model perlu dimonitor terhadap perubahan pola data (*data drift*) antar angkatan atau perubahan kebijakan akademik yang

dapat menggeser distribusi fitur dan label. Monitoring berkala dapat mencakup pemantauan stabilitas metrik, kalibrasi probabilitas, serta pemeriksaan performa per kelompok (misalnya program studi atau kategori penerimaan) untuk mengurangi risiko bias layanan. Jika ditemukan penurunan performa atau ketimpangan yang konsisten, pembaruan model, penyesuaian *threshold*, atau penyempurnaan fitur perlu dilakukan agar EWS tetap relevan, adil, dan berkelanjutan.

3.8 Analisis Sensitivitas *Threshold* dan Skenario Kapasitas Intervensi

Pada implementasi EWS, daftar mahasiswa berisiko yang dihasilkan model akan menjadi beban kerja nyata bagi pembimbing akademik atau unit kemahasiswaan. Karena itu, selain memilih model dengan metrik yang baik, institusi perlu menetapkan strategi *threshold* untuk menyesuaikan cakupan deteksi dengan kapasitas layanan. Jika kapasitas intervensi besar (misalnya program pendampingan massal atau reminder administratif), *threshold* dapat diturunkan untuk meningkatkan *recall*. Sebaliknya, jika kapasitas intervensi terbatas (misalnya konseling intensif), *threshold* dapat dinaikkan agar daftar prioritas lebih selektif. Praktik yang lazim adalah menetapkan target kuota *workload* (misalnya hanya 15–20% mahasiswa dengan skor risiko tertinggi) lalu memilih *threshold* yang menghasilkan proporsi *flagged* mendekati kuota tersebut, sambil memastikan *recall dropout* tetap berada pada tingkat yang dianggap aman.

Untuk menggambarkan efek *threshold*, Tabel 4 menyajikan perubahan *precision*, *recall*, F1, jumlah false positive (FP), false negative (FN), serta proporsi mahasiswa yang ter-flag sebagai berisiko pada *Random Forest (enrollment-only)*. Secara umum, *threshold* yang lebih rendah meningkatkan *recall* namun menaikkan jumlah mahasiswa yang harus ditangani.

Tabel 4. Dampak perubahan *threshold* pada *Random Forest (enrollment-only)*

Thresh old	Precis ion	Rec all	F1	F P	F N	Flag ged	Flagged (%)
0.60	0.837	0.38	0.5	21	17	129	14.6
		0	23		6		
0.50	0.756	0.46	0.5	43	15	176	19.9
		8	78		1		
0.40	0.693	0.62	0.6	78	10	254	28.7
		0	54		8		
0.30	0.551	0.76	0.6	17	67	394	44.5
		4	40		7		
0.20	0.445	0.90	0.5	32	27	577	65.2
		5	97		0		
0.10	0.369	0.97	0.5	47	8	747	84.4
		2	35		1		

Dari Tabel 4 terlihat bahwa penurunan *threshold* memperluas cakupan deteksi (*recall* meningkat), tetapi juga memperbesar jumlah mahasiswa yang ter-flag. Artinya, peningkatan *recall* “dibayar” dengan bertambahnya *workload* dan potensi *false alarm*. Dalam konteks kebijakan, kampus dapat menerapkan strategi bertingkat: *threshold* lebih rendah untuk intervensi ringan (misalnya reminder, kelas penguatan,

pemetaan kendala), dan *threshold* lebih tinggi untuk intervensi intensif (misalnya konseling rutin atau bantuan finansial). Dengan cara ini, precision dan recall tidak dipandang sebagai kompetisi semata, tetapi sebagai parameter desain layanan.

3.9 Interpretabilitas dan Faktor Dominan Risiko Dropout

Selain performa, sistem EWS membutuhkan penjelasan yang dapat dipahami pemangku kepentingan agar hasil prediksi dapat ditindaklanjuti secara tepat. Salah satu pendekatan yang relatif mudah adalah menggunakan model yang interpretatif seperti *Logistic regression*, kemudian meninjau fitur dengan koefisien absolut terbesar sebagai indikasi faktor dominan yang berasosiasi dengan peningkatan/penurunan risiko *dropout*.

Pada fase *enrollment-only*, faktor administratif-finansial dan karakteristik awal mahasiswa cenderung menjadi sinyal utama. Pada fase *enrollment + semester I*, indikator capaian akademik awal menjadi lebih dominan. Ringkasan 10 fitur teratas (berdasarkan |koefisien|) disajikan pada Tabel 5 dan Tabel 6.

Tabel 5. Sepuluh fitur dominan *Logistic Regression (enrollment-only)* berdasarkan |koefisien|

Fitur	Koefisien	Arah
<i>Tuition fees up to date</i>	-0.825	↓ risiko
<i>Scholarship holder</i>	-0.487	↓ risiko
<i>Age at enrollment</i>	0.361	↑ risiko
<i>Gender</i>	0.294	↑ risiko
<i>International</i>	-0.269	↓ risiko
<i>Nacionality</i>	0.229	↑ risiko
<i>Debtor</i>	0.215	↑ risiko
<i>Admission grade</i>	-0.208	↓ risiko
<i>Mother's occupation</i>	-0.189	↓ risiko
<i>Application mode</i>	0.157	↑ risiko

Tabel 6. Sepuluh fitur dominan *Logistic Regression (enrollment + semester I)* berdasarkan koefisien

Fitur	Koefisien	Arah
<i>Curricular units 1st sem (approved)</i>	-2.288	↓ risiko
<i>Curricular units 1st sem (enrolled)</i>	0.766	↑ risiko
<i>Tuition fees up to date</i>	-0.762	↓ risiko
<i>Curricular units 1st sem (credited)</i>	0.712	↑ risiko
<i>Age at enrollment</i>	0.380	↑ risiko
<i>Mother's occupation</i>	-0.380	↓ risiko
<i>International</i>	-0.347	↓ risiko
<i>Course</i>	0.337	↑ risiko
<i>Nacionality</i>	0.300	↑ risiko
<i>Scholarship holder</i>	-0.273	↓ risiko

Secara praktis, interpretasi fitur dapat membantu kampus merancang intervensi yang lebih actionable. Sebagai contoh, jika indikator administrasi/keuangan muncul dominan pada fase pra-kuliah, maka intervensi dapat diarahkan pada pendampingan administrasi atau skema keringanan. Jika indikator akademik semester I dominan, maka intervensi dapat diarahkan pada remedial, tutoring, atau konseling belajar. Namun, pemanfaatan fitur demografis perlu disikapi hati-hati:

model sebaiknya dipakai untuk membuka akses dukungan (supportive), bukan untuk memberi stigma. Karena itu, audit sederhana terhadap distribusi FP/FN pada subkelompok dan evaluasi kebijakan tindak lanjut perlu menjadi bagian dari tata kelola EWS.

3.10 Rekomendasi Integrasi Sistem dan Monitoring Dampak Intervensi

Untuk penerapan di kampus, EWS dapat diintegrasikan ke SIAKAD melalui alur ETL terjadwal: pengambilan data enrollment, pembaruan data semester I, scoring probabilitas *dropout*, dan penyajian hasil pada dashboard pembimbing akademik. Dashboard sebaiknya menampilkan kategori risiko, alasan ringkas (misalnya 3 fitur dominan per mahasiswa), serta status tindak lanjut (sudah dihubungi, konseling, remedial, dsb.).

Dampak EWS perlu dipantau menggunakan indikator operasional dan hasil (*outcome*). Indikator operasional meliputi jumlah mahasiswa yang ter-flag, waktu respons tindak lanjut, dan tingkat penyelesaian pendampingan. Indikator *outcome* meliputi perubahan retensi/DO dibanding periode sebelumnya. Jika memungkinkan, evaluasi dampak dapat dilakukan melalui desain kuasi-eksperimen (misalnya perbandingan antar angkatan) atau uji A/B pada program intervensi untuk memastikan bahwa EWS tidak hanya akurat, tetapi juga efektif meningkatkan retensi.

4. Kesimpulan

Penelitian ini menegaskan bahwa *Early Warning System* (EWS) dapat dibangun sejak tahap awal masuk kuliah dengan memanfaatkan data *enrollment-only* untuk memprediksi risiko *dropout* secara lebih dini. Inovasi utama yang ditawarkan adalah rancangan EWS dua fase: skrining pra-kuliah berbasis *enrollment* untuk membentuk daftar prioritas awal, lalu pembaruan skor risiko setelah semester I ketika sinyal akademik mulai terlihat. Pada kedua fase tersebut, penelitian juga membandingkan tiga model yang umum dan mudah direplikasi pada data tabular, yakni *logistic regression*, *Random forest*, dan *gradient boosting*.

Hasil eksperimen menunjukkan bahwa pada fase *enrollment-only*, *Logistic regression* cenderung lebih sesuai untuk skrining awal karena mampu menjaga cakupan deteksi (*recall*) dan F1 pada kelas *dropout*. Setelah fitur semester I ditambahkan, seluruh model meningkat; model berbasis pohon menjadi lebih selektif (*precision* lebih tinggi) sehingga bermanfaat ketika institusi ingin menekan *false alarm*. Temuan ini menguatkan bahwa pemilihan model perlu diselaraskan dengan tujuan layanan: meminimalkan kasus berisiko yang terlewat, atau mengendalikan beban tindak lanjut agar tetap realistis.

Selain metrik statistik, penelitian ini menunjukkan nilai tambah pendekatan operasional melalui *confusion matrix* dan pengaturan *threshold* probabilitas.

Confusion matrix memungkinkan keluaran model diterjemahkan menjadi estimasi jumlah mahasiswa yang harus ditindaklanjuti (TP+FP) serta potensi konsekuensi jika ada kasus yang terlewat (FN). Dengan demikian, EWS tidak berhenti pada prediksi, tetapi dapat dihubungkan dengan perencanaan sumber daya dan desain intervensi bertingkat (misalnya pengingat administratif untuk risiko rendah hingga konseling intensif untuk risiko tinggi).

Sebagai arah pengembangan, penelitian lanjutan disarankan melakukan evaluasi yang lebih robust melalui *k-fold cross-validation* atau evaluasi multi-seed untuk mengurangi variabilitas akibat satu pembagian data. Penguatan model juga dapat dilakukan melalui *hyperparameter tuning* (misalnya jumlah pohon, kedalaman maksimum, dan *learning rate* pada model *ensemble*), serta kalibrasi probabilitas agar skor risiko lebih konsisten saat digunakan sebagai dasar kebijakan threshold. Di sisi implementasi, validasi eksternal pada data institusi/angkatan lain dan evaluasi prospektif (pra-pasca intervensi) diperlukan untuk memastikan model benar-benar meningkatkan retensi mahasiswa. Terakhir, aspek keberlanjutan dan keadilan layanan perlu dijaga melalui tata kelola data, monitoring bias per kelompok, dan memastikan intervensi yang diprioritaskan tetap bersifat suportif dan dapat diakses secara setara.

Dalam praktik, EWS sebaiknya diposisikan sebagai alat bantu pengambilan keputusan, bukan mekanisme pelabelan yang bersifat menghukum. Kampus dapat menetapkan SOP pencatatan tindak lanjut, menjaga kerahasiaan data, menyediakan jalur umpan balik bagi dosen wali/konselor, dan memastikan distribusi layanan tidak timpang antar program studi atau kelompok mahasiswa. Dengan penguatan tersebut, EWS berpotensi menjadi sistem yang akurat,

transparan, dan berkelanjutan untuk mendukung peningkatan retensi mahasiswa secara berkelanjutan.

Daftar Rujukan

- [1] V. Realinho, J. Machado, L. Baptista, and M. V Martins, "Predicting Student Dropout and Academic Success," *Data*, vol. 7, no. 11, p. 146, 2022.
- [2] J. Kabáthová and M. Drlík, "Towards predicting student's dropout in university courses using different machine learning techniques," *Appl. Sci.*, vol. 11, no. 7, p. 3130, 2021.
- [3] M. V Martins, L. Baptista, J. Machado, and V. Realinho, "Multi-Class Phased Prediction of Academic Performance and Dropout in Higher Education," *Appl. Sci.*, vol. 13, no. 8, p. 4702, 2023.
- [4] R. da Silva, V. B. Realinho, L. M. Baptista, and M. V Martins, "Forecasting Students Dropout: A UTAD University Study," *Futur. Internet*, vol. 14, no. 3, p. 76, 2022.
- [5] M. G. Carballo-Mendivil, M. M. Inzunza-González, and others, "Predicting Student Dropout from Day One: XGBoost-Based Early Detection Model Using First-Year Data," *Appl. Sci.*, 2025.
- [6] J. Niyogisubizo, L. L. Niyigena, and A. G. Lopez, "Predicting student's dropout in university classes using ensemble machine learning approaches," *Comput. Educ. Artif. Intell.*, vol. 3, 2022.
- [7] E. F. Villegas-Ch, X. A. Palacios-Pacheco, and W. L. Tipán-Pachacama, "Applying Machine Learning and Academic Analytics to Predict Dropout and Performance in Higher Education," *Sustainability*, vol. 15, no. 19, 2023.
- [8] S. Kim, Y. Kim, and M. Cho, "Using Machine Learning and Deep Learning Models to Predict Dropout Risk in College Students," *Appl. Sci.*, vol. 13, no. 10, 2023.
- [9] F. B. Kurniawan and L. Farokhah, "Aplikasi Cerdas Prediksi Kelulusan Mahasiswa Berbasis Website Menggunakan Metode Support Vector Machine (SVM)," *J. FASILKOM (teknologi Inf. dan Ilmu KOMputer)*, vol. 15, no. 1, pp. 155–162, 2025, doi: 10.37859/jf.v15i1.8767.
- [10] F. I. Rachman, S. Mujadilah, T. Wahyuni, and L. Anas, "Prediksi Tingkat Kelulusan Menggunakan K-Means Pada Program Studi Informatika Unismuh Makassar," *J. FASILKOM (teknologi Inf. dan Ilmu KOMputer)*, vol. 13, no. 3, pp. 504–510, 2023, doi: 10.37859/jf.v13i3.6061.