JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Optimasi Hyperparameter XGBoost Berbasis Bayesian Optimization dan SHAP untuk Prediksi Dropout Mahasiswa Perguruan Tinggi

Herwis Gultom*, Ahmad Fauzi, Ria Ester
Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: [1,*]dosen02535@unpam.ac.id, [2]dosen02621@unpam.ac.id, [3]dosen02665@unpam.ac.id
Email Penulis Korespondensi: dosen02535@unpam.ac.id
Submitted **30-07-2026**; Accepted **22-08-2026**; Published **31-08-2026**

Abstrak

Putus studi (*dropout*) mahasiswa merupakan masalah pendidikan tinggi yang merugikan secara finansial dan sosial, terutama karena proporsi kelas *dropout* yang tidak seimbang serta minimnya model prediksi yang transparan dan teruji stabil. Penelitian ini menerapkan algoritma Extreme Gradient Boosting (XGBoost) yang hyperparameter-nya dioptimasi melalui Bayesian Optimization berbasis Gaussian Process sebanyak 50 iterasi, dilengkapi ambang keputusan berbasis skor F2, yaitu metrik yang memberi bobot lebih besar pada *recall* dibandingkan *precision*, untuk mengatasi ketidakseimbangan kelas, serta interpretasi Shapley Additive exPlanations (SHAP) guna menjelaskan kontribusi tiap fitur terhadap prediksi model. Penelitian ini bertujuan mengembangkan kerangka kerja prediksi *dropout* yang akurat, transparan, dan tervalidasi secara statistik pada dataset 4.424 mahasiswa dengan 34 atribut akademik, demografis, dan sosioekonomi. Hasil sementara menunjukkan bahwa model XGBoost teroptimasi mencapai luas area di bawah kurva *Receiver Operating Characteristic* (ROC-AUC) sebesar 0,930 dan luas area di bawah kurva presisi-recall (PR-AUC) sebesar 0,904 pada data uji, dengan recall 90,49% pada ambang F2 optimal sebesar 0,2251. Validasi tambahan melalui *nested cross-validation* menghasilkan ROC-AUC rata-rata 0,924, yang mengonfirmasi stabilitas performa model. Interpretasi SHAP mengungkapkan bahwa jumlah mata kuliah yang lulus pada semester kedua merupakan prediktor paling dominan terhadap risiko *dropout*. Temuan ini menegaskan bahwa integrasi optimasi hyperparameter dan *explainable AI* mampu menghasilkan model prediksi *dropout* yang andal dan akuntabel bagi kebijakan intervensi akademik dini.

Kata Kunci: Bayesian Optimization ; Dropout Mahasiswa; Explainable AI; SHAP ; XGBoost

Abstract

Student *dropouts* are a problem in higher education that is financially and socially detrimental, especially due to the unbalanced proportion of dropout classes and the lack of a transparent and stable predictive model. This study applied the Extreme Gradient Boosting (XGBoost) algorithm whose hyperparameters were optimized through Bayesian Optimization based on Gaussian Process as many as 50 iterations, equipped with a decision threshold based on F2 score, which is a metric that gives more weight to *recall* than *precision*, to overcome class imbalances, and the interpretation of Shapley Additive exPlanations (SHAP) to explain the contribution of each feature to prediction model. This study aims to develop an accurate, transparent, and statistically validated dropout prediction framework on a dataset of 4,424 students with 34 academic, demographic, and socioeconomic attributes. Provisional results show that the optimized XGBoost model achieves an area area below the *Receiver Operating Characteristic* curve (ROC-AUC) of 0.930 and an area under the precision-recall curve (PR-AUC) of 0.904 in the test data, with a 90.49% recall at the optimal F2 threshold of 0.2251. Additional validation via *nested cross-validation* resulted in an average ROC-AUC of 0.924, which confirmed the stability of the model's performance. The SHAP interpretation reveals that the number of courses passed in the second semester is the most dominant predictor of dropout risk. These findings confirm that the integration of hyperparameter optimization and *explainable AI* is able to produce a reliable and accountable dropout prediction model for early academic intervention policies.

Keywords: Bayesian Optimization ; Dropout Mahasiswa; Explainable AI; SHAP ; XGBoost

1. PENDAHULUAN

Pendidikan tinggi berperan strategis dalam pembentukan sumber daya manusia yang kompetitif, namun putus studi (*dropout*) tetap menjadi persoalan mahal dan persisten bagi perguruan tinggi di seluruh dunia[1]. Selain merugikan mahasiswa secara individu, *dropout* membebani institusi, keluarga, dan masyarakat melalui menurunnya efisiensi anggaran pendidikan[2]. Tingginya angka *failure, withdrawal,* dan *dropout* bahkan pada mata kuliah dasar menegaskan urgensi deteksi dini mahasiswa berisiko[3].

Berbagai algoritma *machine learning* telah diterapkan untuk memprediksi risiko *dropout* dari data demografis, sosioekonomi, hingga akademik. Putra dkk. membandingkan Random Forest dan XGBoost pada dataset *dropout* Indonesia (34 fitur) dan mendapati Random Forest sedikit unggul pada konfigurasi baku[4]. Kim dkk. menggabungkan data akademik, demografis, sosioekonomi, dan makroekonomi hingga ROC-AUC 0,935, dengan fitur akademik paling dominan[5]. Goren dkk. mengevaluasi Neural Network dan XGBoost lintas bidang studi[6], sementara Song dkk. menunjukkan *gradient boosting* berbasis rata-rata nilai riwayat memberi generalisasi terbaik[7]. Pendekatan lain turut dieksplorasi, seperti *Retrieval-Augmented Generation* [8] dan *text mining* esai motivasi[9]. Namun sebagian besar masih mengandalkan hyperparameter baku; Al Hakim dkk. menerapkan *Grid dan Random Search* pada XGBoost, AdaBoost, dan Random Forest (91,18%) tanpa optimasi global seperti *Bayesian Optimization[10]*.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Karena XGBoost bersifat *black-box*, kebutuhan transparansi mendorong penerapan *Explainable AI* (XAI) pada domain prediksi *dropout*. Nagy dan Molontay memanfaatkan *permutation importance, partial dependence plot*, LIME, dan SHAP untuk intervensi personalisasi[11]; Nti dan Ramanayake mengembangkan kerangka XAI untuk platform pembelajaran daring[12]. SHAP konsisten menunjukkan interpretabilitas baik: Al Hashmi dkk. mencapai AUC-ROC di atas 0,91 pada mahasiswa STEM MENA[13]; Zirekgur dkk. mengombinasikan SMOTE-ENN, CatBoost, SHAP, dan LIME hingga akurasi 98%[14]; Liu dkk. mengembangkan EASE-Predict berbasis ansambel dan SHAP pada 4.424 mahasiswa[15].

Pada dataset yang sama persis dengan penelitian ini *"Predict Students' Dropout and Academic Success"* (4.424 mahasiswa, 34–36 atribut) beberapa studi Indonesia turut menerapkan XGBoost dan SHAP. Istiwana dkk. menjelaskan variasi IPK mahasiswa *dropout* (R^2 = 0,820), dengan SKS tempuh dan status SPP paling berkontribusi[16]. Kurniasari dkk. membandingkan Random Forest (77,40%) dan XGBoost (76,05%), dan menemukan *curricular_units_2nd_sem_approved* serta *tuition_fees_up_to_date* sebagai faktor paling dominan[17]. Wardani dkk. melaporkan akurasi 95,2% dengan SHAP pada IPK kumulatif dan kehadiran[18]. Dominasi dua fitur pada temuan Kurniasari dkk. selaras dengan hasil SHAP penelitian ini (Bagian 3.5), memperkuat validitas keduanya sebagai prediktor *dropout* yang universal pada dataset ini.

Namun ketiga penelitian [18], [19] tersebut belum menerapkan optimasi hyperparameter secara sistematis, belum melaporkan pemilihan ambang keputusan yang mempertimbangkan ketidakseimbangan kelas *dropout* (±32% pada dataset ini), serta belum melengkapi evaluasi dengan interval kepercayaan *bootstrap* maupun *nested cross-validation* untuk memastikan hasil bebas dari bias optimis.

Pemetaan bibliometrik oleh Radulescu dan Balas terhadap 352 dokumen Scopus bidang prediksi *dropout* (2014–2025) mengidentifikasi empat celah utama interpretabilitas, transferabilitas, *fairness*, dan kesiapan implementasi serta mengusulkan kerangka X-EWS lima lapis[1]. Namun belum ditemukan penelitian yang sekaligus (i) menerapkan *Bayesian Optimization*, bukan sekadar *Grid/Random Search[10]*; (ii) melaporkan ambang keputusan yang sadar ketidakseimbangan kelas, bukan ambang baku 0,5; (iii) mengintegrasikan SHAP global-lokal berkuantifikasi ketidakpastian *bootstrap* dan *nested cross-validation*; serta (iv) mengaudit keadilan subgrup[20] pada dataset *"Higher Education Predictors of Student Retention"* yang banyak dijadikan acuan.

Mayoritas penelitian terdahulu unggul pada satu atau dua dimensi saja, namun belum ada yang mengintegrasikan optimasi Bayesian, evaluasi robust berlapis, dan audit keadilan subgrup sekaligus. Dataset penelitian ini terdiri atas 4.424 mahasiswa dengan proporsi *dropout* ±32%, sehingga ketidakseimbangan kelas perlu ditangani eksplisit baik pada optimasi model maupun pemilihan ambang keputusan agar model tidak bias terhadap kelas mayoritas.

Berdasarkan celah tersebut, penelitian ini mengembangkan model prediksi *dropout* berbasis XGBoost[21] yang hyperparameter-nya dioptimasi sistematis melalui *Bayesian Optimization* berpendekatan Gaussian Process dan fungsi akuisisi Expected Improvement pada sepuluh ruang hyperparameter. Berbeda dari penelitian terdahulu yang umumnya berhenti pada akurasi dan SHAP global semata[17], [18], [19], penelitian ini menerapkan pemilihan ambang keputusan berbasis skor F2, interpretasi SHAP TreeExplainer global-lokal, kuantifikasi ketidakpastian melalui *bootstrap* dan *nested cross-validation*, serta audit performa pada subgrup demografis dan sosioekonomi[20], [22]. Penelitian ini juga mempertimbangkan pendekatan *glass-box* seperti regresi logistik terklaster[23], pemodelan interpretatif[24], serta audit *outlier* dan rekayasa fitur[25]. Kontribusinya adalah kerangka kerja prediksi *dropout* yang unggul secara teknis, transparan, terukur ketidakpastiannya, dan akuntabel terhadap keadilan antar subgrup, sehingga menjadi referensi lebih komprehensif bagi pengelola perguruan tinggi dalam merancang intervensi dini risiko putus studi.

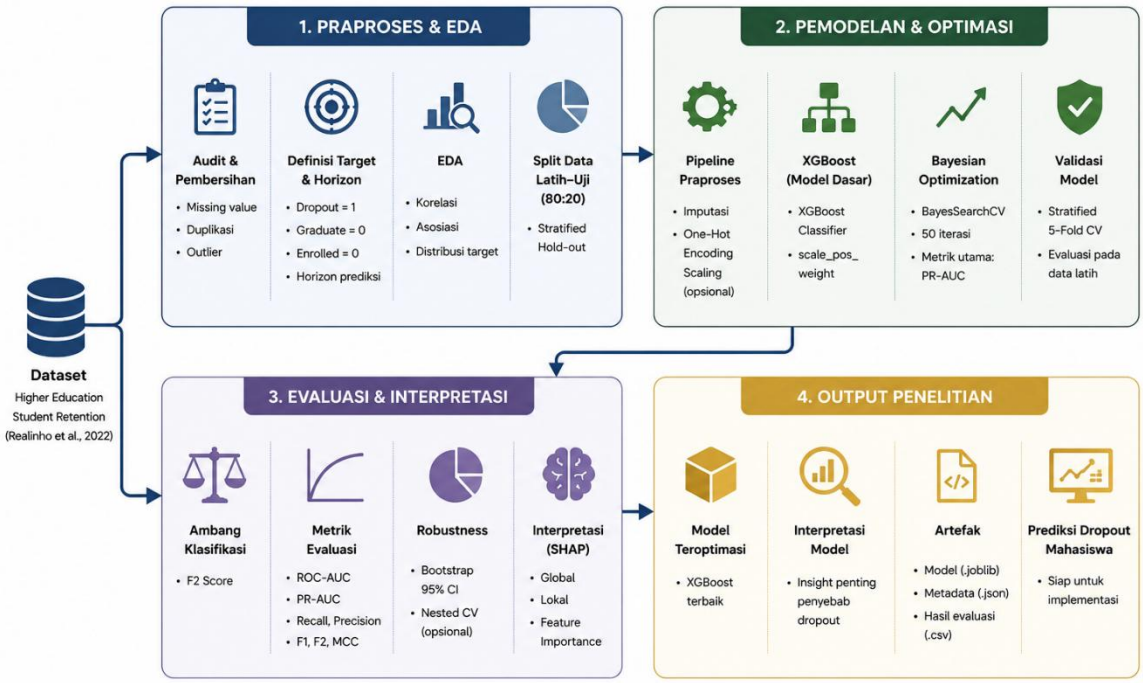
2. METODOLOGI PENELITIAN

2.1 Alur Penelitian

Penelitian ini terdiri atas empat tahapan utama sebagaimana ditunjukkan pada Gambar 1. Alur desain penelitian, yaitu Praproses & EDA, Pemodelan & Optimasi, Evaluasi & Interpretasi, dan Output Penelitian. Tahap Praproses & EDA meliputi audit dan pembersihan data, definisi target dan horizon prediksi, eksplorasi data, serta pembagian data latih dan uji secara *stratified*. Tahap Pemodelan & Optimasi mencakup penyusunan pipeline praproses, pembangunan model dasar XGBoost, optimasi hyperparameter menggunakan Bayesian Optimization, dan validasi dengan *Stratified 5-Fold Cross-Validation*. XGBoost bekerja secara bertahap dengan membentuk pohon keputusan untuk memperbaiki kesalahan prediksi sebelumnya, sedangkan Bayesian Optimization mencari kombinasi hyperparameter terbaik secara efisien berdasarkan hasil evaluasi sebelumnya. Selanjutnya, tahap Evaluasi & Interpretasi dilakukan melalui penentuan ambang klasifikasi, pengukuran metrik performa, pengujian robustness, serta interpretasi menggunakan SHAP untuk mengetahui

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

kontribusi setiap fitur terhadap prediksi *dropout*. Tahap terakhir, Output Penelitian, menghasilkan model XGBoost teroptimasi, interpretasi faktor-faktor penting penyebab *dropout*, artefak model, dan prediksi risiko *dropout* mahasiswa.



Gambar 1. Alur Penelitian

Rincian setiap tahapan pada Gambar 1 dijelaskan sebagai berikut:

a. Dataset dan Sumber Data

Dataset yang digunakan bersumber dari data akademik mahasiswa jenjang pendidikan tinggi yang mencakup atribut saat pendaftaran, akhir semester pertama, dan akhir semester kedua[26]. Dataset terdiri atas 4.424 data mahasiswa dengan 35 atribut awal yang mencakup atribut demografis, sosial-ekonomi, makroekonomi, dan akademik. Variabel target asli dalam dataset terbagi menjadi tiga kategori, yaitu Dropout, Enrolled, dan Graduate.

b. Audit, Pembersihan, dan Definisi Variabel Target

Sebelum digunakan untuk pemodelan, dataset diaudit melalui penghapusan kolom bersifat indeks, konversi berbagai token nilai kosong menjadi *missing value* secara konsisten, konversi tipe data numerik pada atribut yang secara dominan (≥95%) valid, dan penghapusan data duplikat yang identik. Variabel target selanjutnya diformulasikan sebagai klasifikasi biner, dengan kategori Dropout ditetapkan sebagai kelas positif (label 1), sedangkan Graduate dan Enrolled digabungkan sebagai kelas *non-dropout* (label 0). Penelitian ini turut mendefinisikan tiga skenario horizon prediksi berdasarkan ketersediaan informasi pada saat prediksi dilakukan, yaitu *enrollment*, *after_first_semester*, *dan after_second_semester*, dengan skenario terakhir digunakan sebagai konfigurasi acuan utama.

c. Analisis Eksplorasi Data (EDA)

Analisis eksplorasi data dilakukan untuk memahami distribusi kelas target, kelengkapan data, sebaran nilai ekstrem, dan keterkaitan antar-atribut. Deteksi *outlier* pada atribut numerik menggunakan metode interquartile range (IQR), sedangkan keterkaitan atribut numerik dengan target dianalisis menggunakan korelasi Spearman dan keterkaitan atribut kategorikal dengan target dianalisis menggunakan koefisien asosiasi Cramér's V.

d. Pembagian Data dan Pipeline Praproses

Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan teknik *stratified split* terhadap variabel target. Seluruh transformasi fitur numerik (imputasi median) dan fitur kategorikal (imputasi modus dan *one-hot encoding*) diimplementasikan dalam satu pipeline praproses terpadu yang di-fit khusus pada data latih, guna mencegah kebocoran informasi statistik ke proses pelatihan. Ketidakseimbangan kelas ditangani melalui parameter *scale_pos_weight* pada algoritma XGBoost.

e. Pemodelan XGBoost dan Optimasi Hyperparameter

Algoritma klasifikasi yang digunakan adalah Extreme Gradient Boosting (XGBoost)[21]. Validasi model dilakukan menggunakan Stratified K-Fold Cross-Validation dengan k=5, sedangkan pencarian hyperparameter optimal dilakukan

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

menggunakan Bayesian Optimization berbasis Gaussian Process dengan fungsi akuisisi Expected Improvement. Sebagai pembanding, dibangun pula model baseline berupa Dummy Classifier dan XGBoost dengan hyperparameter default.

f. Penentuan Ambang Klasifikasi dan Evaluasi Model

Ambang klasifikasi optimal ditentukan berdasarkan probabilitas prediksi out-of-fold pada data latih, dengan kriteria *F2-score* maksimum. Evaluasi akhir dilakukan pada data uji yang tidak dilibatkan dalam pelatihan maupun optimasi, meliputi metrik ROC-AUC, PR-AUC, *recall*, *specificity*, *precision*, F1/F2-score, *balanced accuracy*, *Matthews Correlation Coefficient* (MCC)[27], *Brier score*, dan *log loss*, dilengkapi interval kepercayaan 95% melalui *bootstrap resampling* [8] serta *nested cross-validation* [9] sebagai validasi tambahan atas stabilitas optimasi.

g. Interpretasi Model dengan SHAP

Interpretasi model dilakukan menggunakan SHAP (*Shapley Additive exPlanations*) melalui TreeExplainer, mencakup interpretasi global (*beeswarm dan bar chart*) untuk mengidentifikasi atribut paling berkontribusi secara keseluruhan, serta interpretasi lokal (*waterfall plot*) pada kasus risiko tertinggi dan kasus *false negative*.

h. Audit Kinerja Subgrup dan Reproduksibilitas

Audit performa subgrup dilakukan pada atribut jenis kelamin, status penerima beasiswa, status internasional, status displaced, dan status debitur untuk menilai konsistensi kinerja antar kelompok. Seluruh eksperimen menerapkan *random state* tetap dan menyimpan model akhir, konfigurasi hyperparameter, ringkasan metrik, serta metadata eksperimen dalam format terstruktur guna mendukung reproduksibilitas.

2.2 Konfigurasi Ruang Pencarian Hyperparameter

Penelitian ini menerapkan proses optimasi hyperparameter menggunakan Bayesian Optimization, berbeda dengan pendekatan yang hanya mengandalkan parameter bawaan pustaka. Pencarian dilakukan pada ruang sepuluh hyperparameter XGBoost menggunakan strategi Gaussian Process dengan fungsi akuisisi *Expected Improvement*, sebanyak 50 iterasi, dievaluasi melalui Stratified 5-Fold Cross-Validation dengan metrik optimasi Average Precision. Guna menjaga konsistensi dan reproduksibilitas hasil, seluruh proses stokastik menggunakan nilai *random_state*=42 secara seragam. Rincian ruang pencarian hyperparameter disajikan pada Tabel 1.

Tabel 1. Ruang Pencarian Hyperparameter Bayesian Optimization

Hyperparameter	Rentang Pencarian	Skala Prior
n_estimators	200 – 1.400	uniform
learning_rate	0,01 – 0,20	log-uniform
max_depth	2 – 8	uniform
min_child_weight	1,0 – 15,0	log-uniform
gamma	1×10^{-8} – 5,0	log-uniform
subsample	0,60 – 1,00	uniform
colsample_bytree	0,60 – 1,00	uniform
reg_alpha	1×10^{-8} – 10,0	log-uniform
reg_lambda	1×10^{-3} – 30,0	log-uniform
scale_pos_weight	0,6r – 1,4r *	log-uniform

* r = rasio jumlah kelas *non-dropout* terhadap kelas *dropout* pada data latih.

2.3 Machine Learning

Machine Learning (ML) adalah proses pembelajaran otomatis di mana algoritma mempelajari pola langsung dari data untuk menemukan representasi input dan memprediksi output tanpa pengkodean eksplisit [28]. ML dimanfaatkan secara luas dalam analitik prediktif di bidang pendidikan (*educational data mining*), termasuk untuk mendeteksi risiko putus studi (*dropout*) mahasiswa secara dini berdasarkan pola data akademik dan demografis.

2.4 Extreme Gradient Boosting (XGBoost)

XGBoost merupakan algoritma *ensemble learning* berbasis *gradient boosting* yang menggabungkan beberapa model pohon keputusan lemah secara aditif untuk meningkatkan akurasi prediksi, efisiensi komputasi pada data berskala besar, serta ketahanan terhadap overfitting melalui regularisasi eksplisit. Guna menghasilkan akurasi tinggi sekaligus mempertahankan kemampuan generalisasi, XGBoost meminimalkan fungsi objektif formal yang mengintegrasikan fungsi kerugian terdiferensiasi (*loss function*) dan suku regularisasi pada iterasi ke-t, sebagaimana dirumuskan pada Persamaan (1):

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (1)$$

Di mana $\mathscr{L}$ melambangkan fungsi kerugian terdiferensiasi yang mengukur selisih antara target aktual y_i dengan hasil prediksi kumulatif ŷ_i pada iterasi sebelumnya ditambah proyeksi fungsi pohon baru f_t(x_i). Komponen Ω(f_t) merupakan fungsi regularisasi yang dirumuskan sebagai Ω(f) = γT + ½λΣw_j², yang mengontrol kompleksitas struktural pohon melalui penalti terhadap jumlah daun (T) dan bobot skor daun (w). Penambahan fungsi regularisasi ini memberi XGBoost ketahanan tinggi terhadap overfitting sekaligus kemampuan menangani pola data akademik non-linier yang kompleks secara efisien [5].

2.5 Bayesian Optimization

Bayesian Optimization merupakan pendekatan pencarian hyperparameter yang memodelkan fungsi objektif sebagai proses Gaussian (Gaussian Process), sehingga pencarian dapat diarahkan secara adaptif ke wilayah ruang parameter yang paling menjanjikan tanpa memerlukan evaluasi pada seluruh kombinasi parameter seperti pada *grid search* . Pada setiap iterasi, kandidat hyperparameter berikutnya dipilih dengan memaksimalkan fungsi akuisisi Expected Improvement (EI), sebagaimana dirumuskan pada Persamaan (2) [13]:

$$EI(x) = (\mu(x) - f(x^+) - \xi)\Phi(Z) + \sigma(x)\phi(Z) \quad (2)$$

Di mana μ(x) dan σ(x) masing-masing merepresentasikan nilai rata-rata dan simpangan baku prediksi model surrogate Gaussian Process pada titik x, f(x⁺) merupakan nilai terbaik yang telah ditemukan sejauh ini, Φ dan φ berturut-turut adalah fungsi distribusi kumulatif dan fungsi densitas probabilitas normal standar, sedangkan ξ adalah parameter eksplorasi-eksploitasi. Pendekatan ini memungkinkan pencarian hyperparameter yang lebih efisien dibandingkan *grid search* maupun *random search*, khususnya pada ruang pencarian berdimensi tinggi seperti pada algoritma XGBoost.

2.6 Explainable AI (SHAP)

Explainable Artificial Intelligence (XAI) merupakan kumpulan teknik yang dirancang untuk memberikan penjelasan transparan terhadap keputusan model *machine learning* guna meningkatkan akuntabilitas dan kepercayaan pengguna. Salah satu metode XAI yang banyak digunakan adalah Shapley Additive exPlanations (SHAP), yang mengadopsi konsep Shapley Value dari teori permainan kooperatif untuk mengonversi model prediktif menjadi model penjelasan aditif. Formulasi dasar nilai Shapley untuk mengukur atribusi kontribusi fitur i dinyatakan pada Persamaan (3) [10]:

$$\phi_i(x) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f_x(S \cup \{i\}) - f_x(S)] \quad (3)$$

Di mana φ_i(x) melambangkan kontribusi marjinal fitur i terhadap hasil prediksi pada sampel x, F merupakan himpunan seluruh fitur, dan S adalah subset fitur yang tidak mengandung fitur i. Pada penelitian ini, interpretasi dilakukan menggunakan TreeExplainer yang secara eksklusif dioptimalkan untuk model berbasis pohon seperti XGBoost, sehingga perhitungan nilai SHAP dapat dilakukan secara eksak dan efisien tanpa aproksimasi permutasi.

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik dan Pembagian Data

Audit data sebelum pemodelan meliputi pemeriksaan nilai hilang, duplikasi, dan distribusi *outlier* berdasarkan aturan *interquartile range* (IQR). Tidak ditemukan nilai hilang maupun data duplikat. Proporsi *outlier* tertinggi terdapat pada *curricular_units_2nd_sem_grade* (19,82%) dan *curricular_units_1st_sem_grade* (16,41%), sedangkan fitur makroekonomi tidak menunjukkan *outlier*. Karena XGBoost berbasis pohon keputusan dan tidak sensitif terhadap skala maupun nilai ekstrem, seluruh *outlier* dipertahankan tanpa transformasi agar distribusi asli data terjaga.

Dataset dibagi menggunakan skema *stratified split* 80:20 agar proporsi kelas dropout pada data latih dan data uji tetap konsisten (Tabel 2). Proporsi *dropout* yang konsisten (±32%) pada kedua subset mengonfirmasi bahwa pembagian data menjaga representativitas kelas minoritas sehingga evaluasi pada data uji dapat dianggap adil.

Tabel 2. Pembagian Data Latih dan Data Uji

Subset	n	Jumlah Dropout	Proporsi Dropout
Data Latih (Train)	3.539	1.137	0,3213
Data Uji (Test)	885	284	0,3209

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Analisis korelasi Spearman menunjukkan jumlah mata kuliah lulus semester kedua (*curricular_units_2nd_sem_approved*) berkorelasi negatif terkuat terhadap *dropout* (ρ = -0,582), diikuti mata kuliah lulus semester pertama (ρ = -0,523), sementara usia saat pendaftaran berkorelasi positif (ρ = 0,280) (Tabel 3).

Tabel 3. Lima Fitur Numerik dengan Korelasi Spearman Tertinggi terhadap *Dropout*

Fitur	Korelasi Spearman (ρ)
curricular_units_2nd_sem_approved	-0,5815
curricular_units_1st_sem_approved	-0,5229
curricular_units_2nd_sem_grade	-0,5100
curricular_units_1st_sem_grade	-0,4426
age_at_enrollment	0,2800

Untuk fitur kategorikal, Cramér's V menunjukkan status pembayaran uang kuliah (*tuition_fees_up_to_date*) memiliki asosiasi terkuat dengan *dropout* (V = 0,429), diikuti jalur masuk, program studi, status beasiswa, dan status tunggakan (Tabel 4), menegaskan faktor akademik dan finansial sebagai penanda risiko *dropout* paling menonjol.

Tabel 4. Lima Fitur Kategorikal dengan Asosiasi Cramér's V Tertinggi terhadap Dropout

Fitur	Cramér's V
tuition_fees_up_to_date	0,4289
application_mode	0,2939
course	0,2526
scholarship_holder	0,2449
debtor	0,2289

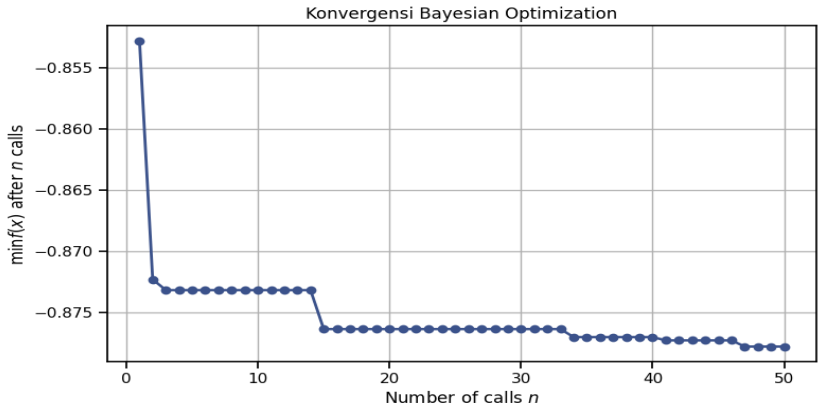
3.2 Perbandingan Baseline dan Hasil Bayesian Optimization

Performa XGBoost konfigurasi dasar dibandingkan dengan *DummyClassifier* melalui *Stratified 5-Fold Cross-Validation*. Tabel 5 menunjukkan XGBoost baseline unggul jauh di atas model dummy pada seluruh metrik, dengan ROC-AUC meningkat dari 0,500 menjadi 0,917 dan PR-AUC dari 0,321 menjadi 0,878, membuktikan daya prediktif fitur yang kuat.

Tabel 5. Perbandingan Baseline Dummy dan XGBoost (Cross-Validation 5-Fold pada Data Latih)

Metrik	Dummy	XGBoost Baseline
ROC-AUC	0,5000	0,9167
PR-AUC	0,3213	0,8777
Balanced Accuracy	0,5000	0,8432
Precision	0,0000	0,7872
Recall	0,0000	0,7871
Specificity	1,0000	0,8993
F1-Score	0,0000	0,7870
MCC	0,0000	0,6864

Optimasi hyperparameter dilakukan menggunakan *Bayesian Optimization* (Gaussian Process, akuisisi Expected Improvement) sebanyak 50 iterasi pada sepuluh hyperparameter, dengan target Average Precision karena proporsi kelas tidak seimbang (±32%). Kurva konvergensi (Gambar 2) menurun tajam pada iterasi awal dan stabil setelah iterasi ke-30. Konfigurasi terbaik diperoleh pada iterasi ke-46 dengan skor CV Average Precision 0,8777 (Tabel 6).



Gambar 2. Kurva Konvergensi Bayesian Optimization

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Tabel 6. Hyperparameter Optimal Hasil Bayesian Optimization

Hyperparameter	Nilai Optimal
n_estimators	1.391
learning_rate	0,0434
max_depth	2
min_child_weight	1,4908
gamma	$9{,}41 \times 10^{-8}$
subsample	0,9364
colsample_bytree	0,7465
reg_alpha (L1)	0,0371
reg_lambda (L2)	2,8517
scale_pos_weight	1,4851

Pola hyperparameter terpilih mencerminkan regularisasi implisit: kedalaman pohon dangkal (*max_depth* = 2), estimator besar (1.391 pohon), dan *learning_rate* kecil (0,0434) menekan risiko *overfitting*. Nilai *scale_pos_weight* optimal (1,4851) lebih rendah dari rasio kelas alami (2,1126), menunjukkan penyeimbangan agresif justru menurunkan presisi. Evaluasi ulang cross-validation menunjukkan interval kepercayaan sempit (ROC-AUC 0,9084–0,9257), mengonfirmasi stabilitas model (Tabel 7) sebelum pengujian pada data uji independen.

Tabel 7. Hasil Cross-Validation Model Teroptimasi (Data Latih, 5-Fold)

Metrik	Rata-rata	Std. Dev.	CI 95% Bawah	CI 95% Atas
ROC-AUC	0,9171	0,0099	0,9084	0,9257
PR-AUC	0,8777	0,0151	0,8645	0,8909
Balanced Accuracy	0,8410	0,0122	0,8303	0,8517
Precision	0,8093	0,0208	0,7910	0,8276
Recall	0,7678	0,0190	0,7511	0,7844
Specificity	0,9142	0,0104	0,9051	0,9233
F1-Score	0,7879	0,0172	0,7728	0,8029
MCC	0,6919	0,0249	0,6701	0,7138

3.3 Pemilihan Ambang Keputusan (Threshold) Berbasis Skor F2

Karena kegagalan mendeteksi mahasiswa berisiko (*false negative*) berkonsekuensi lebih besar daripada *false alarm*, ambang keputusan dioptimasi menggunakan skor F2 (bobot *recall* dua kali *precision*) berdasarkan prediksi *out-of-fold*. Ambang optimal diperoleh sebesar 0,2251, dengan *precision* 0,6386, *recall* 0,8874, dan F2 0,8233.

3.4 Evaluasi Akhir pada Data Uji (Holdout Test)

Model final diuji pada 885 sampel data uji independen; Tabel 8 membandingkan ambang *default* 0,5 dan ambang optimal F2. ROC-AUC 0,930 dan PR-AUC 0,904 yang *threshold-independent* menunjukkan diskriminasi model sangat baik. Pada ambang optimal F2, *recall* meningkat dari 0,8204 menjadi 0,9049 (*false negative* turun dari 51 menjadi 27), meski *precision* menurun dari 0,8351 menjadi 0,6624 trade-off yang sejalan dengan prioritas peringatan dini karena biaya intervensi tambahan lebih rendah dibandingkan biaya kehilangan mahasiswa.

Tabel 8. Perbandingan Metrik Evaluasi pada Data Uji Berdasarkan Skema Threshold

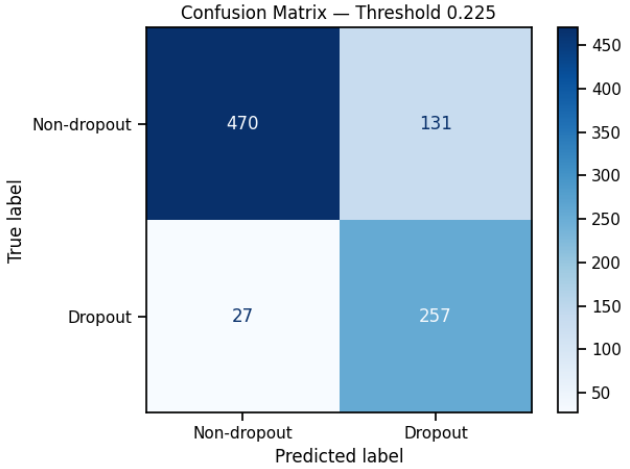
Metrik	Threshold 0,5	Threshold F2-Optimal (0,2251)
Accuracy	0,8904	0,8215
Balanced Accuracy	0,8719	0,8435
Precision	0,8351	0,6624
Recall (Sensitivity)	0,8204	0,9049
Specificity	0,9235	0,7820
F1-Score	0,8277	0,7649
F2-Score	0,8233	0,8432
MCC	0,7474	0,6463
ROC-AUC	0,9300	0,9300
PR-AUC	0,9040	0,9040
Brier Score	0,0862	0,0862
Log Loss	0,2920	0,2920
TP / TN / FP / FN	233 / 555 / 46 / 51	257 / 470 / 131 / 27

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Laporan klasifikasi pada ambang F2-optimal disajikan pada Tabel 9; confusion matrix (Gambar 3) menunjukkan model mengenali 257 dari 284 mahasiswa *dropout* aktual (*recall* 90,5%), dengan 27 kasus terlewat.

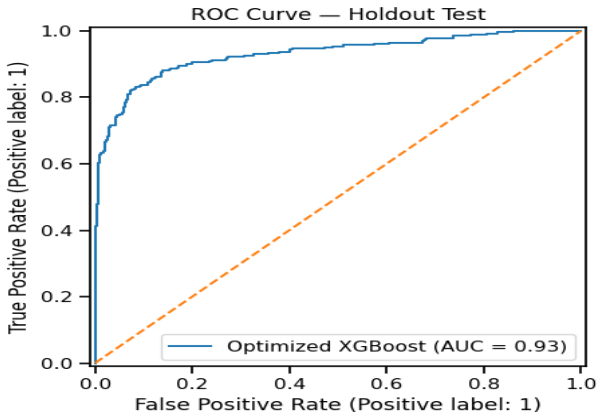
Tabel 9. Laporan Klasifikasi pada Threshold F2-Optimal (0,2251)

Kelas	Precision	Recall	F1-Score	Support
Non-dropout	0,9457	0,7820	0,8561	601
Dropout	0,6624	0,9049	0,7649	284
Macro avg	0,8040	0,8435	0,8105	885
Weighted avg	0,8548	0,8215	0,8268	885



Gambar 3. Confusion Matrix pada Threshold F2-Optimal (Data Uji)

Kurva ROC (Gambar 4) berada jauh di atas garis diagonal *random guess* dengan AUC 0,93, mengonfirmasi model memisahkan kelas *dropout* dan non-dropout secara konsisten. *Bootstrap resampling* 2.000 iterasi menghasilkan interval kepercayaan 95% pada Tabel 10.



Gambar 4. Kurva ROC pada Data Uji

Tabel 10. Interval Kepercayaan 95% Bootstrap (2.000 Resample, Data Uji)

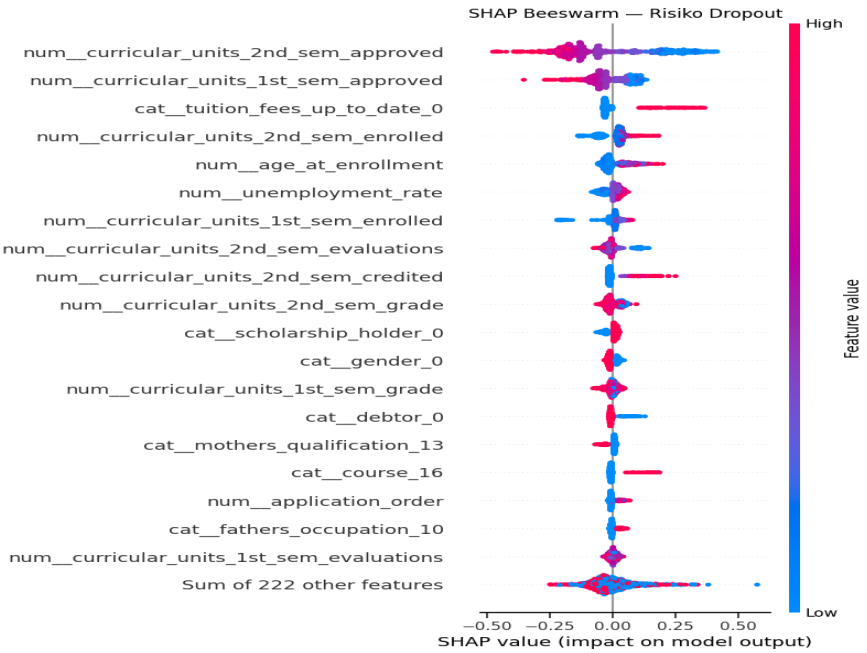
Metrik	Rata-rata Bootstrap	CI 95% Bawah	CI 95% Atas
ROC-AUC	0,9299	0,9080	0,9496
PR-AUC	0,9038	0,8757	0,9285
Balanced Accuracy	0,8437	0,8188	0,8671
Recall	0,9049	0,8691	0,9388
Specificity	0,7824	0,7491	0,8156
F1-Score	0,7645	0,7282	0,7994
MCC	0,6465	0,5981	0,6961

Rentang interval kepercayaan yang sempit menegaskan bahwa performa model bukan hasil kebetulan satu pembagian data, melainkan estimasi stabil yang mewakili kemampuan generalisasi model pada data mahasiswa baru.

3.5 Interpretasi Global Model dengan SHAP

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Interpretasi model menggunakan SHAP TreeExplainer dihitung terhadap 885 sampel data uji dan 241 fitur hasil *one-hot encoding*, diagregasi ke 34 atribut asal. Gambar 5 menyajikan *beeswarm plot* arah dan besar kontribusi tiap fitur terhadap prediksi *dropout*.



Gambar 5. SHAP Beeswarm Plot

Sepuluh fitur dengan kontribusi SHAP tertinggi disajikan pada Tabel 11.

Tabel 11. Sepuluh Fitur dengan Kontribusi SHAP (Mean |SHAP Value|) Tertinggi

Peringkat	Fitur	Mean |SHAP Value|
1	curricular_units_2nd_sem_approved	0,1798
2	curricular_units_1st_sem_approved	0,0682
3	course	0,0679
4	tuition_fees_up_to_date	0,0608
5	curricular_units_2nd_sem_enrolled	0,0435
6	age_at_enrollment	0,0374
7	mothers_qualification	0,0371
8	fathers_occupation	0,0345
9	mothers_occupation	0,0321
10	application_mode	0,0294

curricular_units_2nd_sem_approved merupakan prediktor paling dominan, dengan kontribusi SHAP jauh melampaui fitur lain (0,1798, hampir tiga kali peringkat kedua), selaras dengan korelasi Spearman dan mengonfirmasi performa akademik aktual sebagai sinyal risiko *dropout* terkuat. Nilai rendah pada *curricular_units_2nd_sem_approved* dan *curricular_units_1st_sem_approved* konsisten bergeser ke arah probabilitas *dropout* lebih tinggi. Program studi dan status pembayaran uang kuliah tetap menjadi kontributor utama, memperkuat temuan Cramér's V bahwa beban finansial berperan penting. Fitur sosiodemografi seperti usia, kualifikasi pendidikan ibu, dan pekerjaan orang tua turut muncul pada sepuluh besar dengan kontribusi lebih kecil.

3.6 Interpretasi Lokal dengan SHAP Waterfall Plot

Analisis lokal dilakukan pada dua kasus representatif: mahasiswa dengan probabilitas dropout tertinggi dan kasus *false negative* dengan probabilitas tertinggi. Kasus dengan probabilitas *dropout* tertinggi (0,9991) merupakan mahasiswa yang secara aktual *dropout*, didorong terutama oleh rendahnya jumlah mata kuliah lulus pada kedua semester. Analisis kasus *false negative* mengungkap profil mahasiswa yang tampak aman secara administratif namun tetap berhenti studi karena faktor yang tidak terekam pada dataset, menegaskan bahwa keputusan intervensi individu tetap perlu penilaian akademik, bukan semata output model.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 4, Agustus 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i4.10055
Hal 1189-1200
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

3.7 Audit Performa Model pada Subgrup Demografis dan Sosioekonomi

Sebagai bagian akuntabilitas model, performa prediksi diuji pada beberapa subgrup atribut sensitif untuk memeriksa potensi disparitas kinerja (Tabel 12).

Tabel 12. Performa Model pada Subgrup Demografis dan Sosioekonomi (Data Uji, Threshold F2-Optimal)

Atribut	Kelompok	n	Dropout Rate	Precision	Recall	Specificity	F1	ROC-AUC
gender	0	564	0,2447	0,6321	0,8841	0,8333	0,7372	0,9256
gender	1	321	0,4548	0,6923	0,9247	0,6571	0,7918	0,9180
scholarship_holder	1	231	0,1255	0,4762	0,6897	0,8911	0,5634	0,8607
scholarship_holder	0	654	0,3899	0,6850	0,9294	0,7268	0,7887	0,9333
international	0	859	0,3213	0,6658	0,9022	0,7856	0,7662	0,9283
displaced	1	500	0,2840	0,6127	0,8803	0,7793	0,7225	0,9004
displaced	0	385	0,3688	0,7174	0,9296	0,7860	0,8098	0,9586
debtor	0	795	0,2918	0,6426	0,8836	0,7975	0,7441	0,9224
debtor	1	90	0,5778	0,7536	1,0000	0,5526	0,8595	0,9504

Model mempertahankan ROC-AUC di atas 0,86 pada seluruh subgrup, menunjukkan diskriminasi tetap baik lintas kelompok. Subgrup penerima beasiswa memiliki *precision* dan F1-score jauh lebih rendah (0,4762 dan 0,5634), kemungkinan akibat ukuran subgrup kecil (n = 231) dan prevalensi dropout rendah (12,55%). Subgrup *debtor* menunjukkan prevalensi *dropout* tertinggi (57,78%) dengan *recall* sempurna (1,0000) namun *specificity* rendah (0,5526), menandakan model cenderung menandai hampir seluruh mahasiswa kelompok ini berisikomenegaskan pentingnya audit subgrup dalam tata kelola model prediktif pendidikan.

3.8 Validasi Tambahan dengan Nested Cross-Validation

Untuk memastikan performa tinggi bukan akibat bias optimis pemilihan hyperparameter, dilakukan *nested cross-validation* (5-fold *outer loop*, 3-fold *inner loop*, 25 iterasi Bayesian Optimization per outer fold) pada seluruh 4.424 sampel, dengan hasil rata-rata dirangkum pada Tabel 13.

Tabel 13. Rata-rata Hasil Nested Cross-Validation (5 Outer Fold, 3 Inner Fold)

Metrik	ROC-AUC	PR-AUC	Balanced Acc.	Recall	Specificity	F1
Rata-rata ± SD	0,9240 ± 0,0073	0,8877 ± 0,0139	0,8492 ± 0,0113	0,7966 ± 0,0421	0,9017 ± 0,0287	0,7949 ± 0,0125

Rata-rata ROC-AUC nested cross-validation sebesar 0,9240 ± 0,0073 dan PR-AUC 0,8877 ± 0,0139 konsisten dengan hasil cross-validation maupun evaluasi data uji holdout, menjadi bukti bahwa performa model bukan hasil *overfitting* melainkan kemampuan generalisasi yang sesungguhnya.

3.9 Perbandingan Hasil dengan Penelitian Terdahulu

Dibandingkan penelitian terdahulu, Putra dkk. melaporkan Random Forest sedikit unggul dibandingkan XGBoost pada konfigurasi baku, sementara Bayesian Optimization pada penelitian ini mengangkat ROC-AUC XGBoost hingga 0,930. Dibandingkan Al Hakim dkk. (Grid Search dan Random Search, akurasi 91,18%), penelitian ini mencapai akurasi kompetitif (89,04%) dengan evaluasi jauh lebih hemat. Temuan *curricular_units_2nd_sem_approved* dan *tuition_fees_up_to_date* sebagai prediktor dominan konsisten dengan Kurniasari dkk. dan Istiwana dkk. pada dataset identik. Berbeda dengan Zirekgur dkk. (akurasi 98% melalui SMOTE), penelitian ini sengaja tidak menerapkan oversampling agar distribusi kelas alami terjaga sehingga metrik lebih jujur dan general. Penambahan interval bootstrap, nested cross-validation, dan audit subgrup menjawab celah interpretabilitas, transferabilitas, dan fairness yang diidentifikasi Radulescu dan Balas, sekaligus melengkapi kerangka XAI dengan audit keadilan sebagaimana ditekankan Yu dan Lee.

Secara keseluruhan, kombinasi optimasi hyperparameter XGBoost berbasis Bayesian Optimization dengan ambang keputusan berbasis skor F2 menghasilkan model prediksi *dropout* berdiskriminasi sangat baik (ROC-AUC 0,930; PR-AUC 0,904) dan *recall* tinggi (90,5%). Interpretasi SHAP membuktikan performa akademik semester pertama dan kedua sebagai sinyal risiko *dropout* paling dominan, diikuti program studi, status finansial, dan latar belakang sosiodemografi. Konsistensi hasil pada bootstrap, nested cross-validation, dan literatur terdahulu menegaskan kelayakan kerangka ini sebagai acuan sistem peringatan dini *dropout* yang transparan dan akuntabel.

4. KESIMPULAN

Penelitian ini berhasil mengatasi kesenjangan metodologis dalam prediksi *student dropout* melalui pengembangan kerangka kerja yang mengintegrasikan Bayesian Optimization, pemilihan ambang keputusan berbasis F2-score, interpretasi model menggunakan SHAP, kuantifikasi ketidakpastian melalui *bootstrap*, serta audit keadilan antar subgrup. Model XGBoost yang dioptimasi menunjukkan performa prediksi yang sangat baik dengan ROC-AUC 0,930 dan PR-AUC 0,904 pada data uji, didukung interval kepercayaan *bootstrap* yang sempit (0,908–0,950) dan hasil nested cross-validation yang konsisten (ROC-AUC 0,924), sehingga mengindikasikan kemampuan generalisasi yang baik. Penggunaan ambang keputusan berbasis F2 (0,2251) meningkatkan *recall* menjadi 90,49%, sehingga mampu meminimalkan kasus *dropout* yang tidak terdeteksi dan mendukung implementasi sistem peringatan dini akademik. Analisis SHAP mengidentifikasi jumlah mata kuliah lulus pada semester kedua sebagai prediktor paling berpengaruh, diikuti performa semester pertama, program studi, dan status pembayaran uang kuliah, yang konsisten dengan temuan penelitian sebelumnya. Audit keadilan mengungkap adanya disparitas performa pada beberapa subgrup, sehingga evaluasi aspek *fairness* perlu menjadi bagian dari implementasi model. Keterbatasan penelitian meliputi penggunaan data dari satu institusi, peningkatan *false positive* akibat optimasi *recall*, serta tingginya beban komputasi. Penelitian selanjutnya disarankan melakukan validasi lintas institusi, menguji prediksi pada horizon yang lebih dini, dan mengeksplorasi strategi mitigasi disparitas subgrup serta integrasi umpan balik intervensi akademik untuk meningkatkan efektivitas model dalam lingkungan nyata.

REFERENCES

[1] D. A. Radulescu and V. E. Balas, "Explainable Early Warning Systems for Student *Dropout* Prediction: Insights from a Bibliometric Analysis," *International Journal of Computers, Communications and Control*, vol. 21, no. 4, Jul. 2026, doi: 10.15837/ijccc.2026.4.7599.

[2] J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, "Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization," *Computers and Education: Artificial Intelligence*, vol. 3, Jan. 2022, doi: 10.1016/j.caeai.2022.100066.

[3] S. A. Tazehkand and M. C. Wang, "Leveraging XGBoost to Reduce Failure, Withdrawal, and Dropout Rates in Undergraduate Level Mathematics/Statistics Education," in *Frontiers in Artificial Intelligence and Applications*, IOS Press BV, Jul. 2024, pp. 396–404. doi: 10.3233/FAIA240276.

[4] L. G. R. Putra, D. D. Prasetya, and M. Mayadi, "Student Dropout Prediction Using Random Forest and XGBoost Method," *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 9, no. 1, pp. 147–157, Feb. 2025, doi: 10.29407/intensif.v9i1.21191.

[5] S. Kim, E. Yoo, and S. Kim, "Why Do Students Drop Out? University Dropout Prediction and Associated Factor Analysis Using Machine Learning Techniques," Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.10987

[6] O. Goren, L. Cohen, and A. Rubinstein, "Early Prediction of Student Dropout in Higher Education using Machine Learning Models," in *Proceedings of the International Conference on Educational Data Mining*, International Educational Data Mining Society, 2024, pp. 349–359. doi: 10.5281/zenodo.12729834.

[7] Z. Song, S. H. Sung, D. M. Park, and B. K. Park, "All-Year Dropout Prediction Modeling and Analysis for University Students," *Applied Sciences (Switzerland)*, vol. 13, no. 2, Jan. 2023, doi: 10.3390/app13021143.

[8] M. Mihoubi, M. Zerkouk, and B. Chikhaoui, "Beyond classical and contemporary models: a transformative AI framework for student dropout prediction in distance learning using RAG, Prompt engineering, and Cross-modal fusion," Jul. 2025, doi: 10.21125/edulearn.2025.0815.

[9] K. F. B. Soppe, A. Bagheri, S. Nadi, I. G. Klugkist, T. Wubbels, and L. D. N. V Wijngaards-De Meij, "Predicting First-Year Dropout from Pre-Enrolment Motivation Statements Using Text Mining."

[10] M. Faris Al Hakim *et al.*, "Optimization of Machine Learning Model using Grid and Random Search Algorithms for Predicting Student Dropout," *Jurnal Teknik Informatika (JUTIF)*, vol. 7, no. 3, 2026, doi: 10.52436/1.jutif.2026.7.3.5627.

[11] M. Nagy and R. Molontay, "Interpretable Dropout Prediction: Towards XAI-Based Personalized Intervention," *Int. J. Artif. Intell. Educ.*, vol. 34, no. 2, pp. 274–300, Jun. 2024, doi: 10.1007/s40593-023-00331-8.

[12] I. K. Nti and S. Ramanayake, "Explainable machine learning for student dropout prediction and tailored interventions in online personalized education," *Discover Artificial Intelligence*, vol. 6, no. 1, Dec. 2026, doi: 10.1007/s44163-026-01016-6.

[13] R. A. M. Al Hashmi, P. D. Zervopoulos, H. M. Elmehdi, and I. Ozturk, "Predicting Dropout in MENA STEM Higher Education Using Explainable AI: A Machine Learning Approach," *Emerging Science Journal*, vol. 9, no. Special Issue, pp. 268–286, May 2025, doi: 10.28991/ESJ-2025-SIED1-016.

[14] M. Zirekgür, B. Karakaya, and D. Avcı, "Predicting Student Dropout Using Explainable Artificial Intelligence: The Impact of SMOTE-Based Class Balancing Techniques," *Firat University Journal of Experimental and Computational Engineering*, vol. 5, no. 2, pp. 574–593, Jun. 2026, doi: 10.62520/fujece.1887893.

[15] Z. Liu, X. Zhou, and Y. Liu, "Student Dropout Prediction Using Ensemble Learning with SHAP-Based Explainable AI Analysis," *Journal of Social Systems and Policy Analysis*, vol. 2, no. 3, pp. 111–132, Aug. 2025, doi: 10.62762/jsspa.2025.321501.

[16] M. Dutta, K. M. Mehedi Hasan, A. Akter, M. H. Rahman, and M. Assaduzzaman, "An interpretable machine learning-based breast cancer classification using XGBoost, SHAP, and LIME," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 6, 2024, doi: 10.11591/eei.v13i6.7866.

[17] N. D. Kurniasari, T. B. Rohman, A. Widianto, A. Shobikah, and G. Triono, "ANALISIS FAKTOR DOMINAN PREDIKSI DROPOUT MAHASISWA MENGGUNAKAN RANDOM FOREST, XGBOOST, DAN XAI."

[18] S. Wardani, S. Mandasari, and M. Riandini, "Predicting Student Dropout Risk Using XGBoost and Explainable AI", doi: 10.12345/xxxxx.

[19] A. P. Istiwana, R. R. Sani, and Y. T. C. Pramudi, "PENDEKATAN EXPLAINABLE MACHINE LEARNING UNTUK ANALISIS FAKTOR DROP OUT MAHASISWA MENGGUNAKAN XGBOOST," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1074–1083, Jan. 2026, doi: 10.36341/rabit.v11i1.7218.

[20] R. Yu, H. Lee, and R. F. Kizilcec, "Should College Dropout Prediction Models Include Protected Attributes?," in *L@S 2021 - Proceedings of the 8th ACM Conference on Learning @ Scale*, Association for Computing Machinery, Inc, Jun. 2021, pp. 91–100. doi: 10.1145/3430895.3460139.

[21] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[22] S. Yang, "Who will dropout from university? Academic risk prediction based on interpretable machine learning", doi: 10.3969/j.issn.1673-4807.2012.01.018.

[23] A. Nigri, M. Bilancia, B. Cafarelli, and S. Magro, "Modelling higher education dropouts using sparse and interpretable post-clustering logistic regression," May 2025, [Online]. Available: http://arxiv.org/abs/2505.07582

[24] S. Wang and B. Luo, "Academic achievement prediction in higher education through interpretable modeling," *PLoS One*, vol. 19, no. 9 September, Sep. 2024, doi: 10.1371/journal.pone.0309838.

[25] E. Alzabidi, Ö. Yakut, S. Saad, and O. Fındık, "SOD-FE: a supervised outlier detection and feature engineering approach for student dropout prediction," *Sci. Rep.*, Jun. 2026, doi: 10.1038/s41598-026-56591-6.

[26] V. Realinho, J. Machado, L. Baptista, and M. V. Martins, "Predicting Student Dropout and Academic Success," *Data (Basel).*, vol. 7, no. 11, Nov. 2022, doi: 10.3390/data7110146.

[27] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.

[28] G. Di Franco and M. Santurro, "Machine learning, artificial neural networks and social research," *Qual. Quant.*, vol. 55, no. 3, pp. 1007–1025, Jun. 2021, doi: 10.1007/s11135-020-01037-y.