

PENERAPAN EXTREME GRADIENT BOOSTING DALAM KLASIFIKASI TINGKAT PROKRASTINASI AKADEMIK MAHASISWA TINGKAT AKHIR BERBASIS DATA PSIKOMETRIK

Hapsoro Wisnu Murti¹, Faisal Muttaqin^{2*}, Fetty Tri Anggraeny³

^{1,2,3}Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, Indonesia

*Penulis korespondensi: faisalmuttaqin.if@upnjatim.ac.id

ABSTRAK

Prokrastinasi akademik merupakan perilaku penundaan dalam menyelesaikan tugas akademik yang masih banyak dialami mahasiswa, terutama mahasiswa tingkat akhir yang sedang menghadapi penyelesaian tugas akhir. Kondisi ini dapat berdampak pada keterlambatan studi, penurunan produktivitas akademik, serta tekanan psikologis. Penelitian ini bertujuan menerapkan algoritma *Extreme Gradient Boosting (XGBoost)* untuk mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir berdasarkan data psikometrik. Data penelitian diperoleh melalui kuesioner yang memuat instrumen *Academic Procrastination Scale-Short Form*, *Emotion Regulation Questionnaire*, dan *Perceived Stress Scale-10*. Tahapan penelitian meliputi skoring dan labeling data untuk membentuk kelas prokrastinasi akademik rendah, sedang, dan tinggi, pembagian data menggunakan *stratified split* dengan rasio 80:20, pra-pemrosesan data serta pelatihan model XGBoost. Ketidakseimbangan kelas pada data latih ditangani menggunakan *sample weight*. Hyperparameter tuning dilakukan menggunakan *GridSearchCV* dengan *Stratified K-Fold Cross Validation* dan pemilihan parameter terbaik berdasarkan *F1-score macro*. Berdasarkan hasil pengujian pada data uji, model memperoleh akurasi sebesar 0,9114, *precision macro* sebesar 0,9084, *recall macro* sebesar 0,9007, dan *F1-score macro* sebesar 0,9025. Hasil *confusion matrix* menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data uji dengan benar pada setiap kelas, terutama pada kelas rendah dan tinggi, meskipun masih terdapat beberapa kesalahan klasifikasi pada kelas sedang. Dengan demikian, XGBoost dapat diterapkan sebagai metode klasifikasi tingkat prokrastinasi akademik mahasiswa tingkat akhir berbasis data psikometrik dengan performa yang baik.

Kata kunci: XGBoost, Prokrastinasi Akademik, Klasifikasi, Data Psikometrik, Mahasiswa.

1 PENDAHULUAN

Mahasiswa tingkat akhir berada pada fase akademik yang menuntut kemandirian tinggi dalam menyelesaikan tugas akhir atau skripsi. Pada tahap ini, mahasiswa tidak hanya dituntut memahami fokus penelitian, tetapi juga mengatur waktu, menjaga konsistensi, serta mengambil keputusan akademik selama proses penyusunan (Raharjani dkk., 2023). Kondisi dalam mengelola tuntutan tersebut dapat mendorong munculnya perilaku penundaan tugas akademik. Dalam konteks akademik, prokrastinasi akademik dapat dipahami sebagai perilaku menunda untuk memulai atau menyelesaikan tugas akademik, seperti pengerjaan skripsi (Aisyah dkk., 2021; Huang dkk., 2025). Perilaku ini dapat terlihat dari adanya kesenjangan antara rencana dan pelaksanaan, serta kecenderungan memilih aktivitas lain yang lebih menyenangkan dibandingkan mengerjakan tugas akademik (Tri Yulia & Armianti, 2025). Prokrastinasi akademik dapat memberikan dampak terhadap aspek akademik maupun psikologis mahasiswa seperti, penurunan performa akademik, meningkatnya tekanan waktu, berkurangnya ketepatan dan ketelitian dalam penyelesaian tugas, serta kesejahteraan yang lebih rendah (Rad dkk., 2025).

Fenomena ini masih cukup banyak ditemukan pada lingkungan perguruan tinggi. Tingkat kejadian prokrastinasi akademik pada mahasiswa dilaporkan berada pada kisaran 81% hingga 90%, dengan sekitar 40% mahasiswa termasuk dalam kelompok yang sering melakukan prokrastinasi (Huang dkk., 2025). Dalam konteks mahasiswa di Indonesia, prokrastinasi akademik juga masih ditemukan pada tingkat yang cukup tinggi. Pada mahasiswa Fakultas Ekonomi dan Bisnis Universitas Negeri Padang, prokrastinasi akademik dalam menyelesaikan tugas perkuliahan memperoleh rata-rata skor 3,08 dengan tingkat capaian responden sebesar 61,62% termasuk dalam kategori tinggi (Tri Yulia & Armianti, 2025). Temuan tersebut mengindikasikan bahwa prokrastinasi akademik masih menjadi permasalahan yang relevan dalam lingkungan perguruan tinggi, baik pada konteks umum maupun pada penyelesaian tugas skripsi ('Aisyah dkk., 2021). Tingginya kecenderungan penundaan ini tidak hanya berdampak pada performa akademik, tetapi juga berhubungan dengan kemampuan mahasiswa dalam mengelola diri dan stres akademik. Regulasi diri berperan dalam membantu mahasiswa mengarahkan perilaku dan mempertahankan strategi belajar sesuai tujuan akademik. Di sisi lain, stres akademik dapat mendorong kecenderungan menghindari tugas yang dianggap berat atau menekan. Keterkaitan tersebut menunjukkan bahwa perilaku penundaan tidak hanya berkaitan dengan manajemen waktu, tetapi juga dengan kontrol diri dan tekanan psikologis selama proses penyelesaian kewajiban akademik (Arwina dkk., 2022; Elizondo dkk., 2024).

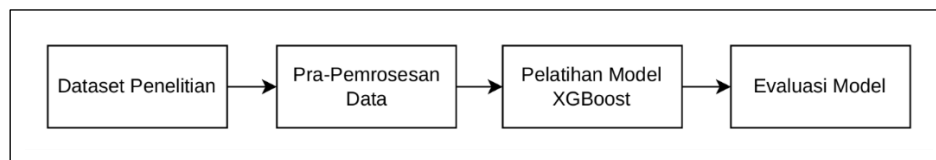
Perkembangan *machine learning* dalam bidang pendidikan memungkinkan perilaku akademik mahasiswa dianalisis melalui data yang lebih terstruktur. Pendekatan ini dapat digunakan untuk mengenali pola perilaku mahasiswa dan membantu proses identifikasi dini terhadap mahasiswa yang berisiko mengalami hambatan akademik (Boicu dkk., 2023). Penelitian terdahulu yang relevan telah mengkaji perilaku prokrastinasi akademik melalui pendekatan berbasis data oleh Sartana dkk, memanfaatkan data log Moodle mahasiswa dengan K-Means clustering untuk mengelompokkan pola prokrastinasi, kemudian membandingkan beberapa model *machine learning* untuk memprediksi performa akademik, dengan hasil Random Forest menjadi model paling unggul yang ditunjukkan oleh nilai R^2 sebesar 0,812, MAE sebesar 6,248, dan RMSE sebesar 8,456 (Sartana dkk., 2025). Penelitian lain oleh Memon dkk., menganalisis perilaku prokrastinasi mahasiswa pada Virtual Learning Environment menggunakan pendekatan *machine learning*, dengan Logistic Regression dan Artificial Neural Network-Multi Layer Perceptron sebagai metode *supervised learning*. Hasil penelitian tersebut menunjukkan bahwa ANN-MLP memperoleh performa terbaik, dengan akurasi 83,5% pada data latih dan 82,9% pada data uji (Memon dkk., 2020).

Temuan dari penelitian terdahulu menunjukkan bahwa data aktivitas e-learning dan log pembelajaran telah dimanfaatkan dalam analisis perilaku penundaan akademik berbasis *machine learning*. Selain itu, penelitian terdahulu juga menyarankan pengembangan metode *ensemble learning*, termasuk pendekatan *boosting*, untuk meningkatkan analisis perilaku akademik mahasiswa (Imhof dkk., 2022; Sartana dkk., 2025). Berdasarkan hal tersebut, penelitian ini menggunakan metode *Extreme Gradient Boosting* (XGBoost) untuk mengklasifikasikan tingkat prokrastinasi akademik. XGBoost merupakan metode pembelajaran dalam *supervised learning* berbasis *gradient-boosted decision tree* (Wiens dkk., 2025). Algoritma ini umumnya digunakan untuk menyelesaikan permasalahan regresi dan klasifikasi. XGBoost bekerja dengan menggabungkan beberapa *weak learner*, seperti pohon keputusan secara bertahap. Setiap model baru dibangun untuk memperbaiki kesalahan prediksi yang dihasilkan oleh model sebelumnya (Fathu & Faulina, 2026). Algoritma ini diketahui memiliki keunggulan dalam mempelajari hubungan pola kompleks antar fitur serta kemampuan regularisasi untuk mengurangi risiko *overfitting* (Arif Ali dkk., 2023; Wiens dkk., 2025). Penggunaan data psikometrik dalam penelitian ini diarahkan untuk merepresentasikan kondisi internal mahasiswa yang berkaitan dengan perilaku prokrastinasi akademik (Arwina dkk., 2022). Berbeda dengan penelitian terdahulu yang memanfaatkan data aktivitas e-learning

dan log pembelajaran menggunakan metode K-Means, Random Forest, Logistic Regression, atau ANN-MLP untuk menganalisis perilaku prokrastinasi atau memprediksi performa akademik, penelitian ini menggunakan XGBoost untuk mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir berbasis data psikometrik. Hasil yang diharapkan dari penelitian ini adalah diperolehnya model klasifikasi yang mampu membedakan tingkat prokrastinasi akademik ke dalam kategori rendah, sedang, dan tinggi. Dengan demikian, penelitian ini bertujuan untuk menerapkan dan mengevaluasi XGBoost dalam mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir berbasis data psikometrik.

2 METODE

Penelitian ini menggunakan pendekatan kuantitatif berbasis *machine learning* untuk mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir. Alur penelitian dimulai dari penggunaan dataset penelitian, pra-pemrosesan data (Muttaqin dkk., 2025), penanganan ketidakseimbangan kelas menggunakan *sample weight*, pelatihan dan tuning model XGBoost, serta evaluasi model.



Gambar 1. Alur Penelitian

2.1 Dataset Penelitian

Data penelitian diperoleh melalui kuesioner daring dengan teknik *purposive sampling* dan *convenience sampling*. Responden penelitian merupakan mahasiswa tingkat akhir minimal semester 7 yang sedang menyusun tugas akhir atau skripsi. Lingkup pengumpulan data difokuskan pada mahasiswa UPN “Veteran” Jawa Timur dan diperluas secara daring kepada mahasiswa dari perguruan tinggi lain yang memenuhi kriteria responden. Dataset yang digunakan dalam penelitian ini berjumlah 394 data responden dan telah berada dalam bentuk siap analisis.

Kuesioner yang digunakan memuat data pendukung responden dan instrumen psikometrik, yaitu *Academic Procrastination Scale-Short Form* (APS-S), *Emotion Regulation Questionnaire* (ERQ), dan *Perceived Stress Scale-10* (PSS-10). Target klasifikasi pada penelitian ini berupa tingkat prokrastinasi akademik mahasiswa yang terdiri atas tiga kategori, yaitu rendah, sedang, dan tinggi. Kategori tersebut diperoleh dari hasil total skor APS berdasarkan batas kategori yang telah ditentukan menggunakan norma ideal. Proses penentuan kategori dan penyusunan data penelitian ditinjau melalui pendampingan akademisi bidang psikologi untuk memastikan kesesuaian tahapan pengolahan data dengan konteks pengukuran prokrastinasi akademik. Dalam proses pemodelan, atribut psikometrik dan data pendukung responden digunakan sebagai fitur masukan, sedangkan tingkat prokrastinasi akademik digunakan sebagai target klasifikasi.

2.2 Pra-Pemrosesan Data

2.2.1 Split Data

Tahap awal dalam pra-pemrosesan data dilakukan dengan membagi dataset menjadi data latih dan data uji. Pembagian data dilakukan menggunakan *stratified split* dengan rasio 80:20, yaitu 80% data digunakan sebagai data latih dan 20% data digunakan sebagai data uji. Penggunaan *stratified split* bertujuan untuk menjaga proporsi kelas rendah, sedang, dan tinggi agar tetap seimbang secara proporsional pada data latih dan data uji. Setelah proses pembagian

data latih dan data uji , kemudian data dipisahkan menjadi data fitur dan target . Data fitur yang digunakan dalam penelitian ini meliputi usia, jenis kelamin, semester, item APS-S, item ERQ, dan item PSS-10. Sementara itu, target klasifikasi yang digunakan adalah label tingkat prokrastinasi akademik mahasiswa, yang terdiri atas kelas rendah, sedang, dan tinggi.

2.2.2 Encoding dan Normalisasi Data

Tahap berikutnya dalam pra-pemrosesan data adalah encoding dan normalisasi data. Encoding dilakukan untuk mengubah variabel kategorikal ke dalam bentuk numerik agar dapat diproses pada tahap pelatihan model XGBoost. Pada penelitian ini, variabel kategorikal seperti jenis kelamin dikonversi ke dalam bentuk angka. Proses encoding diterapkan pada data latih dan data uji dengan format transformasi yang sama agar struktur data tetap konsisten.

Selanjutnya, normalisasi dilakukan pada variabel numerik untuk menyamakan skala antar fitur. Hal ini diperlukan karena atribut psikometrik memiliki rentang nilai yang berbeda sesuai dengan skala instrumen yang digunakan. Normalisasi dilakukan menggunakan StandardScaler dan hanya diterapkan pada variabel numerik, sedangkan variabel kategorikal hasil encoding tidak ikut dinormalisasi. Normalisasi dilakukan dengan menghitung nilai rata-rata dan standar deviasi dari data latih, kemudian nilai tersebut digunakan untuk mentransformasikan data latih dan data uji. Tahapan ini dilakukan agar informasi dari data uji tidak memengaruhi proses pembentukan skala pada data latih. Rumus StandardScaler ditunjukkan sebagai berikut.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Keterangan :

x = Nilai asli fitur

μ = Nilai rata – rata fitur

σ = Nilai standar deviasi fitur

z = Nilai hasil normalisasi

2.3 Pelatihan Model XGBoost

Tahap pelatihan model dilakukan menggunakan algoritma Extreme Gradient Boosting (XGBoost) pada data latih yang telah melalui proses pra-pemrosesan. Sebelum proses pelatihan dilakukan, sample weight diterapkan untuk menangani ketidakseimbangan jumlah data pada setiap kelas. Pembobotan ini diberikan pada data latih agar model tidak terlalu cenderung mempelajari kelas dengan jumlah data yang lebih besar, serta tetap memperhatikan kelas dengan jumlah data yang lebih sedikit (Bakirara & Elhan, 2023). Bobot kelas dihitung berdasarkan jumlah data pada setiap kelas dalam data latih. Secara umum, perhitungan bobot kelas dapat dituliskan sebagai berikut.

$$w_c = \frac{N}{K \times N_c} \quad (2)$$

Keterangan :

w_c = bobot kelas ke-c

N = jumlah seluruh data latih

K = jumlah kelas target

N_c = jumlah data latih pada kelas ke-c

Berdasarkan rumus tersebut, kelas dengan jumlah data lebih sedikit akan memperoleh bobot yang lebih besar, sedangkan kelas dengan jumlah data lebih banyak akan memperoleh bobot yang lebih kecil. Nilai bobot tersebut kemudian diberikan pada setiap data latih sesuai dengan kelasnya masing-masing pada proses pelatihan XGBoost.

Setelah bobot sampel ditentukan, model dilatih menggunakan algoritma XGBoost yang bekerja dengan membangun model secara bertahap melalui penambahan beberapa pohon keputusan. Pada setiap iterasi, pohon keputusan baru ditambahkan untuk memperbaiki kesalahan prediksi dari model sebelumnya (Pramesti dkk., 2025). Proses pembaruan prediksi pada setiap iterasi XGBoost ditunjukkan pada rumus berikut.

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i) \quad (3)$$

Keterangan :

$\hat{y}_i^{(t)}$ = hasil prediksi data ke-i pada iterasi ke-t

$\hat{y}_i^{(t-1)}$ = hasil prediksi pada iterasi sebelumnya

η = learning rate

$f_t(x_i)$ = pohon keputusan yang ditambahkan pada iterasi ke-t

Nilai learning rate digunakan untuk mengatur besar kecilnya kontribusi pohon baru, sehingga proses pembelajaran model dapat berlangsung lebih terkontrol.

Setelah seluruh pohon keputusan dibangun, hasil prediksi akhir XGBoost diperoleh dari akumulasi keluaran setiap pohon keputusan. secara umum, rumus prediksi akhir ditunjukkan sebagai berikut.

$$\hat{y}_i = \sum_{t=1}^T f_t(x_i) \quad (4)$$

Keterangan :

$\hat{y}_i$ = hasil prediksi pada data ke-i

T = jumlah pohon keputusan

$f_t(x_i)$ = keluaran pohon keputusan ke-t terhadap data ke-i

Selain membentuk prediksi secara bertahap, XGBoost juga mengoptimalkan fungsi objektif yang terdiri atas fungsi loss untuk meminimalkan kesalahan prediksi dan komponen regularisasi digunakan untuk mengurangi risiko overfitting. Fungsi objektif tersebut ditunjukkan pada rumus berikut.

$$Obj = \sum_{i=1}^N l(y_i, \hat{y}_i) + \sum_{t=1}^T \Omega(f_t) \quad (5)$$

Keterangan :

Obj = fungsi objektif model

$l(y_i, \hat{y}_i)$ = fungsi loss antara nilai aktual dan hasil prediksi

N = jumlah data latih

T = jumlah pohon Keputusan

$\Omega(f_t)$ = komponen regularisasi pada pohon keputusan

Untuk memperoleh model dengan performa terbaik, dilakukan *hyperparameter tuning* menggunakan *GridSearchCV*. Proses validasi pada *GridSearchCV* dilakukan menggunakan *Stratified K-Fold Cross Validation* agar proporsi kelas rendah, sedang, dan tinggi tetap terjaga pada setiap fold. Metode ini dipilih karena grid search mengevaluasi secara menyeluruh setiap kombinasi *hyperparameter* yang telah ditentukan. Kinerja setiap kombinasi kemudian diestimasi melalui *cross-validation* yang membagi data ke dalam beberapa fold dan menggabungkan hasil evaluasinya, sehingga konfigurasi terbaik ditentukan dari rata-rata performa validasi (Bischl dkk., 2023). Parameter yang diuji meliputi *learning_rate*, *max_depth*, *n_estimators*, dan *subsample*. Pemilihan kombinasi hyperparameter terbaik didasarkan pada nilai *F1-score macro* karena metrik ini mampu mengevaluasi performa model

pada setiap kelas secara seimbang. Ruang pencarian hyperparameter yang digunakan ditampilkan pada Tabel 1.

Tabel 1. Ruang pencarian Hyperparameter

Hyperparameter	Nilai diuji
Learning_rate	0.01, 0.1
Max_depth	2, 3
N_estimators	50, 100
Subsample	0.8 , 1.0

Pemilihan ruang pencarian dilakukan dengan menetapkan dua nilai pada setiap hyperparameter untuk mewakili variasi konfigurasi model sekaligus menjaga jumlah kombinasi tetap terukur. Nilai *learning_rate* 0,01 dan 0,1 dipilih untuk menguji laju pembelajaran yang lebih bertahap dan moderat. Nilai *max_depth* 2 dan 3 dipilih untuk menjaga kedalaman pohon tetap terbatas karena pohon yang semakin dalam dapat meningkatkan kompleksitas model dan risiko overfitting (Johnsen dkk., 2021). Pembatasan kedalaman tersebut juga mempertimbangkan jumlah data latih yang tersedia. Sementara itu, nilai *n_estimators* 50 dan 100 digunakan untuk menguji dua jumlah putaran boosting dalam proses pencarian konfigurasi model. Nilai *subsample* 0,8 dan 1,0 digunakan untuk menguji penggunaan sebagian dan seluruh data latih pada setiap putaran pembentukan pohon (González-Castro dkk., 2024; Mehdary dkk., 2024). Setelah kombinasi *hyperparameter* terbaik diperoleh, model XGBoost dilatih kembali menggunakan data latih dengan parameter terbaik dan *sample weight* yang telah ditentukan. Model terbaik yang dihasilkan kemudian digunakan untuk melakukan prediksi pada data uji.

2.4 Evaluasi Model

Evaluasi model dilakukan menggunakan data uji yang tidak digunakan pada proses pelatihan. Tahap ini bertujuan untuk mengetahui kemampuan model XGBoost dalam mengklasifikasikan tingkat prokrastinasi akademik mahasiswa ke dalam kategori rendah, sedang, dan tinggi. Evaluasi dilakukan dengan menggunakan confusion matrix dan beberapa metrik evaluasi, yaitu accuracy, precision, recall, dan F1-score. Perhitungan metrik evaluasi dilakukan berdasarkan nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) yang diperoleh dari confusion matrix. Berikut rumus metrik evaluasi yang digunakan ditunjukkan sebagai berikut.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

Pada penelitian ini, nilai F1-score macro digunakan sebagai perhatian utama karena mampu mengevaluasi performa model pada setiap kelas secara seimbang tanpa dipengaruhi oleh jumlah data pada masing-masing kelas. Selain itu, nilai F1-score weighted juga digunakan untuk melihat performa model dengan mempertimbangkan proporsi data pada setiap kelas.

3 HASIL DAN PEMBAHASAN

Pada bagian ini disajikan hasil penelitian berdasarkan tahapan pemodelan yang telah dilakukan. Pembahasan difokuskan pada interpretasi hasil yang diperoleh dari proses penerapan XGBoost dalam klasifikasi tingkat prokrastinasi akademik mahasiswa. Hasil yang ditampilkan meliputi distribusi data hasil split, hasil hyperparameter tuning, perbandingan performa data latih dan data validasi, serta evaluasi model pada data uji.

3.1 Distribusi Hasil Split Data

Hasil pembagian dataset menggunakan stratified split menghasilkan 315 data latih dan 79 data uji. Pembagian ini dilakukan agar proporsi kelas pada data latih dan data uji tetap terjaga. Distribusi data pada masing-masing kelas ditampilkan pada Tabel 2.

Tabel 2. Distribusi Hasil Split Data

Kelas	Data Latih	Data Uji	Total Data
Rendah	123	31	154
Sedang	88	22	110
Tinggi	104	26	130
Total	315	79	394

Berdasarkan Tabel 2, Kelas rendah memiliki jumlah data paling banyak, yaitu 154 data, diikuti oleh kelas tinggi sebanyak 130 data dan kelas sedang sebanyak 110 data. Kondisi tersebut menunjukkan bahwa distribusi kelas pada dataset tidak sepenuhnya seimbang. Distribusi kelas yang tidak sepenuhnya seimbang tersebut menjadi dasar penggunaan *sample weight* pada proses pelatihan model, agar model tetap memperhatikan kelas dengan jumlah data yang lebih sedikit.

3.2 Hasil Hyperparameter Tuning

Hyperparameter tuning dilakukan menggunakan *GridSearchCV* dengan *Stratified K-Fold Cross Validation*. Pemilihan parameter terbaik didasarkan pada nilai *F1-score macro* karena metrik ini dapat mengevaluasi performa model pada setiap kelas secara seimbang. Hasil hyperparameter tuning model XGBoost ditampilkan pada Tabel 3.

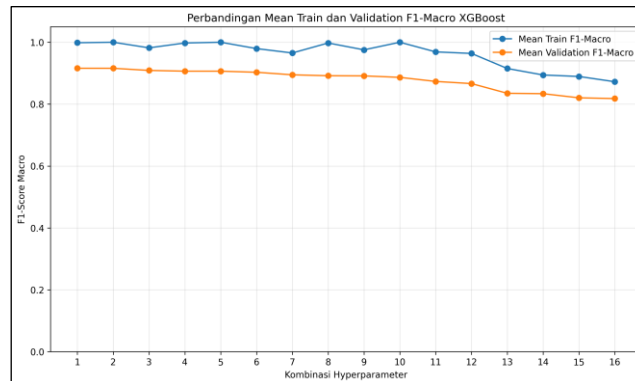
Tabel 3. Hasil *Hyperparameter* XGBoost

Parameter	Nilai Terbaik
Learning_rate	0.1
Max_depth	3
N_estimators	50
Subsample	0.8
Best CV F1-Score Macro	0.9158

Berdasarkan Tabel 3, kombinasi hyperparameter terbaik diperoleh pada *learning_rate* 0,1, *max_dept* 3, *n_estimators* 50, dan *subsample* 0,8. Hasil tersebut menunjukkan bahwa pada proses tuning, kombinasi parameter dengan kedalaman pohon dan jumlah pohon yang relatif terbatas memberikan nilai *F1-score macro* terbaik. Dengan demikian, model yang dipilih tidak hanya mempertimbangkan kemampuan dalam mempelajari pola data latih, tetapi juga performanya pada data validasi. Hasil tuning menunjukkan nilai best CV *F1-score macro* sebesar 0.9159. Nilai ini menunjukkan bahwa parameter tersebut memberikan performa validasi terbaik dibandingkan kombinasi *hyperparameter* lainnya pada proses *GridSearchCV*.

Pemilihan *max_depth* 3 menunjukkan bahwa kedalaman pohon yang relatif terbatas sudah mampu memberikan performa validasi terbaik pada data psikometrik yang digunakan. Nilai *n_estimators* 50 juga menunjukkan bahwa jumlah pohon yang tidak terlalu banyak telah

cukup untuk membentuk model dengan performa yang baik pada proses tuning. Sementara itu, *learning_rate* 0,1 berperan dalam mengatur kontribusi setiap pohon baru secara bertahap, sedangkan *subsample* 0,8 menunjukkan bahwa pembentukan pohon dilakukan dengan menggunakan sebagian data latih.



Gambar 2. Grafik Train-Validation F1-Score Macro

Pada Gambar 1. menunjukkan perbandingan nilai *F1-score macro* pada data latih dan data validasi untuk melihat perbedaan performa model selama proses tuning. Perbandingan ini digunakan untuk mengetahui apakah model cenderung terlalu menyesuaikan diri pada data latih atau masih mampu mempertahankan performa yang baik pada data validasi.

Dari grafik menunjukkan bahwa nilai *train F1-score macro* cenderung lebih tinggi dibandingkan nilai *validation F1-score macro* pada setiap kombinasi hyperparameter, yang menandakan bahwa model mampu mempelajari pola data latih dengan baik. Pada kombinasi parameter terbaik, model memperoleh *mean train F1-score macro* sebesar 0,9984 dan *mean validation F1-score macro* sebesar 0,9159, dengan *gap train-validation* sebesar 0,0825. Selisih tersebut menunjukkan adanya perbedaan performa antara data latih dan validasi, tetapi nilai validasi yang tetap pada rentang cukup tinggi menunjukkan bahwa model masih memiliki kemampuan generalisasi yang baik.

3.3 Hasil Evaluasi Model

Setelah kombinasi hyperparameter terbaik diperoleh, model XGBoost dievaluasi menggunakan data uji. Evaluasi dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan tingkat prokrastinasi akademik mahasiswa ke dalam tiga kelas, yaitu rendah, sedang, dan tinggi. Hasil evaluasi model berdasarkan accuracy, precision, recall, dan F1-score ditampilkan pada Tabel 4.

Tabel 4. Hasil Evaluasi Model pada Data Uji

Metrik Evaluasi	Nilai
Accuracy	0.9114
Precision Macro	0.9084
Recall Macro	0.9007
F1-Score Macro	0.9025
Precision Weighted	0.9019
Recall Weighted	0.9114
F1-Score Weighted	0.9094

Hasil pada Tabel 4 menunjukkan bahwa model XGBoost memperoleh accuracy sebesar 0,9114 dengan 72 dari 79 data uji berhasil diklasifikasikan dengan benar oleh model. Hal ini menunjukkan bahwa model mampu dalam membedakan tingkat prokrastinasi akademik berdasarkan data psikometrik dengan baik. Nilai precision macro sebesar 0,9084, recall macro

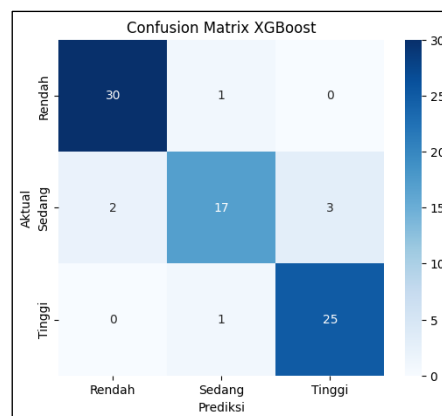
sebesar 0,9007, dan F1-score macro sebesar 0,9025 menunjukkan bahwa model memiliki rata - rata ketepatan klasifikasi yang baik dan kemampuan pengenalan kelas yang cukup seimbang pada setiap kelas. Nilai F1-score macro menjadi perhatian utama karena tidak hanya melihat performa model secara umum, tetapi juga menggambarkan performa model pada setiap kelas secara lebih seimbang. Selain itu, nilai F1-score weighted sebesar 0,9094 yang mendekati F1-score macro menunjukkan bahwa performa model tetap baik meskipun ada perbedaan jumlah data pada masing-masing kelas.

Ditinjau dari hasil penelitian terdahulu oleh Memon dkk menggunakan ANN-MLP yang diterapkan pada data aktivitas mahasiswa dalam Virtual Learning Environment memperoleh akurasi data uji sebesar 82,9%, sementara itu penelitian Sartana dkk menggunakan data log Moodle dan memperoleh Random Forest sebagai model terbaik dengan nilai R^2 sebesar 0,812. Hasil tersebut menunjukkan bahwa XGBoost memberikan capaian klasifikasi yang lebih tinggi daripada ANN-MLP dengan akurasi sebesar 91,14% dan F1-score macro sebesar 90,25%. Sementara itu, Random Forest menunjukkan performa yang baik dalam memprediksi performa akademik, sedangkan XGBoost pada penelitian ini menunjukkan performa yang baik dalam mengklasifikasikan tingkat prokrastinasi secara langsung. Untuk melihat performa model pada masing-masing kelas, hasil classification report ditampilkan pada Gambar 3.

CLASSIFICATION REPORT XGBOOST				
	precision	recall	f1-score	support
Rendah	0.94	0.97	0.95	31
Sedang	0.89	0.77	0.83	22
Tinggi	0.89	0.96	0.93	26
accuracy			0.91	79
macro avg	0.91	0.90	0.90	79
weighted avg	0.91	0.91	0.91	79

Gambar 3. Classification Report XGBoost

Berdasarkan classification report, kelas rendah memperoleh *precision* sebesar 0,94, *recall* sebesar 0,97, dan *F1-score* sebesar 0,95, sedangkan kelas tinggi memperoleh *precision* sebesar 0,89, *recall* sebesar 0,96, dan *F1-score* sebesar 0,93. Nilai recall yang tinggi pada kedua kelas tersebut menunjukkan bahwa sebagian besar data aktual kelas rendah dan tinggi berhasil dikenali dengan benar oleh model. Sementara itu, kelas sedang memperoleh *precision* sebesar 0,89, *recall* sebesar 0,77, dan *F1-score* sebesar 0,83. Nilai recall yang lebih rendah dibandingkan kelas rendah dan tinggi menunjukkan bahwa kelas sedang relatif lebih sulit dikenali. Kondisi tersebut dapat terjadi karena kelas sedang berada di antara kelas rendah dan tinggi, sehingga karakteristiknya berpotensi beririsan dengan dua kelas lainnya. Untuk melihat sebaran prediksi benar dan salah pada setiap kelas, selanjutnya ditampilkan melalui confusion matrix pada Gambar 4.



Gambar 4. Confusion Matrix XGBoost

Model berhasil mengklasifikasikan kelas rendah dan tinggi dengan baik, yaitu 30 dari 31 data kelas rendah dan 25 dari 26 data kelas tinggi diprediksi secara benar. Kesalahan pada kedua kelas tersebut hanya terjadi pada masing-masing 1 data yang diklasifikasikan sebagai kelas sedang. Sementara itu, pada kelas sedang, model berhasil mengklasifikasikan 17 dari 22 data dengan benar, tetapi masih terdapat 2 data yang diklasifikasikan sebagai kelas rendah dan 3 data sebagai kelas tinggi. Hasil ini menunjukkan bahwa kesalahan klasifikasi lebih banyak terjadi pada kelas sedang, yang mengindikasikan bahwa kategori ini memiliki karakteristik yang relatif lebih sulit dibedakan dibandingkan kelas rendah dan tinggi. Temuan tersebut juga sejalan dengan nilai recall kelas sedang yang lebih rendah pada classification report.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa model XGBoost mampu mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir dengan performa yang baik. Nilai accuracy sebesar 0.9114 dan F1-score macro sebesar 0.9025 menunjukkan bahwa model tidak hanya menghasilkan prediksi yang tepat secara umum, tetapi juga mampu menjaga keseimbangan performa pada ketiga kelas target.

4 KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma XGBoost dapat diterapkan untuk mengklasifikasikan tingkat prokrastinasi akademik mahasiswa tingkat akhir berbasis data psikometrik. Hasil evaluasi pada data uji menunjukkan bahwa model memperoleh *accuracy* sebesar 0,9114 dan *F1-score macro* sebesar 0,9025. Nilai tersebut menunjukkan bahwa model mampu memberikan performa klasifikasi yang baik dan relatif seimbang pada tiga kelas target, yaitu rendah, sedang, dan tinggi. Hasil penelitian menunjukkan bahwa kelas rendah dan tinggi dapat dikenali dengan lebih baik oleh model, sedangkan kelas sedang masih menjadi kelas yang relatif lebih sulit diklasifikasikan. Hal ini terlihat dari nilai recall kelas sedang yang lebih rendah dibandingkan kelas lainnya serta adanya kesalahan klasifikasi kelas sedang ke kelas rendah dan tinggi. Dengan demikian, penerapan XGBoost pada data psikometrik dapat digunakan sebagai pendekatan awal untuk mengidentifikasi kecenderungan tingkat prokrastinasi akademik mahasiswa. Hasil klasifikasi yang diperoleh dapat menjadi informasi pendukung bagi pengelola atau pihak terkait dalam mengenali mahasiswa yang memerlukan perhatian dan pendampingan lebih lanjut selama penyusunan tugas akhir, dengan tetap mempertimbangkan informasi akademik lainnya dan tidak menjadikannya sebagai satu-satunya dasar dalam menentukan bentuk pendampingan.

Penelitian selanjutnya disarankan untuk memperluas jumlah dan variasi data responden, meninjau kembali fitur yang digunakan, serta membandingkan performa XGBoost dengan algoritma machine learning lainnya. Selain itu, perhatian khusus perlu diberikan pada peningkatan kemampuan model dalam membedakan kelas sedang, karena kelas ini masih menunjukkan pola yang lebih beririsan dengan kelas rendah dan tinggi.

UCAPAN TERIMA KASIH

Penelitian ini dapat terlaksana berkat dukungan, arahan dan keterlibatan dari berbagai pihak. Penulis menyampaikan penghargaan dan terima kasih kepada Dr. Faisal Muttaqin, S.Kom., M.T. serta Fetty Tri Anggraeny, S.Kom., M.Kom. atas bimbingan, masukan, evaluasi, dan dukungan yang diberikan sejak tahap perencanaan penelitian hingga penyusunan artikel. Ucapan terima kasih juga disampaikan kepada seluruh responden yang telah meluangkan waktu dan memberikan data melalui pengisian kuesioner, sehingga penelitian ini dapat terlaksana sesuai dengan tujuan yang telah ditetapkan. Selain itu, apresiasi diberikan kepada Program Studi Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur, atas dukungan akademik dan lingkungan pembelajaran yang menunjang proses penelitian serta penyelesaian artikel ini.

DAFTAR PUSTAKA

- 'Aisyah, S., Adi, E. P., & Wedi, A. (2021). Studi Analisis Prokrastinasi Akademik Mahasiswa dalam Mengerjakan Skripsi. *JKTP: Jurnal Kajian Teknologi Pendidikan*, 4(4), 358–367. <https://doi.org/10.17977/um038v4i42021p358>
- Arif Ali, Z., H. Abduljabbar, Z., A. Tahir, H., Bibo Sallow, A., & Almufti, S. M. (2023). eXtreme Gradient Boosting Algorithm with Machine Learning: a Review. *Academic Journal of Nawroz University*, 12(2), 320–334. <https://doi.org/10.25007/ajnu.v12n2a1612>
- Arwina, A., Fadilah, R., & Murad, A. (2022). Hubungan Regulasi Diri dan Stres Akademik dengan Prokrastinasi Akademik pada Mahasiswa Fakultas Psikologi Universitas Medan Area. *Journal of Education, Humaniora and Social Sciences (JEHSS)*, 4(4), 1981–1991. <https://doi.org/10.34007/jehss.v4i4.988>
- Bakirara, B., & Elhan, A. H. (2023). Class Weighting Technique to Deal with Imbalanced Class Problem in Machine Learning: Methodological Research. *Turkiye Klinikleri Journal of Biostatistics*, 15(1), 19–29. <https://doi.org/10.5336/biostatic.2022-93961>
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A. L., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Dalam *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (Vol. 13, Nomor 2). John Wiley and Sons Inc. <https://doi.org/10.1002/widm.1484>
- Boicu, M., Lalwani, M. J., High, T. J., Daga, A., & Lalwani, J. (2023). Analyzing Student Procrastination to Identify At-Risk Behavior. *American Society for Engineering Education (ASEE) Southeastern Section Conference*. <https://doi.org/https://doi.org/10.18260/1-2--44985>
- Elizondo, K., Valenzuela, R., Pestana, J. V., & Codina, N. (2024). Self-regulation and procrastination in college students: A tale of motivation, strategy, and perseverance. *Psychology in the Schools*, 61(3), 887–902. <https://doi.org/10.1002/pits.23088>
- Fathu, A. N., & Faulina, R. (2026). KOMPARASI ALGORITMA XGBOOST DAN RANDOM FOREST DALAM MENGLASIFIKASIKAN SENTIMEN PENGGUNA IPUSNAS. *Prosiding Seminar Nasional Sains dan Teknologi Seri V Fakultas Sains dan Teknologi, Universitas Terbuka*, 3(1). <https://conference.ut.ac.id/index.php/saintek/article/view/7957>
- González-Castro, L., Chávez, M., Duflot, P., Bleret, V., Del Fiol, G., & López-Nores, M. (2024). Impact of Hyperparameter Optimization to Enhance Machine Learning Performance: A Case Study on Breast Cancer Recurrence Prediction. *Applied Sciences (Switzerland)*, 14(13). <https://doi.org/10.3390/app14135909>
- Huang, S., Li, Z., Li, J., & Ding, X. (2025). A network analysis of academic procrastination, psychological and environmental factors among medical students. *BMC Psychology*, 13(1). <https://doi.org/10.1186/s40359-025-02916-5>
- Imhof, C., Comsa, I.-S., Hlosta, M., Parsaeifard, B., Moser, I., & Bergamin, P. (2022). Prediction of Dilatory Behavior in e-Learning: A Comparison of Multiple Machine Learning Models. *Arxiv*. <https://doi.org/https://doi.org/10.48550/arXiv.2206.15079>
- Johnsen, P. V., Riemer-Sørensen, S., DeWan, A. T., Cahill, M. E., & Langaas, M. (2021). A new method for exploring gene–gene and gene–environment interactions in GWAS with

- tree ensemble methods and SHAP values. *BMC Bioinformatics*, 22(1). <https://doi.org/10.1186/s12859-021-04041-7>
- Mehdary, A., Chehri, A., Jakimi, A., & Saadane, R. (2024). Hyperparameter Optimization with Genetic Algorithms and XGBoost: A Step Forward in Smart Grid Fraud Detection. *Sensors*, 24(4). <https://doi.org/10.3390/s24041230>
- Memon, M. M., Hashmani, M. A., Muslim Jameel, S., Junejo, S., & Raza, K. (2020). Analysis of Student Procrastinatory Behavior in Virtual Learning Environments Using Machine Learning. *Journal of Hunan University (Natural Sciences)*, 47(10). <https://jonuns.com/index.php/journal/article/view/459>
- Muttaqin, F., Rahardjo, P., Zaim Chilmi, M., Priharyoto Bayuseno, A., Indah Winarni, T., Rizal Isnanto, R., & Jamari. (2025). A Combination Method of ROI, CLAHE, and DenseNet-169 for Hip Osteoarthritis Detection. *Engineering, Technology and Applied Science Research*, 15(3), 22690–22697. <https://doi.org/10.48084/etasr.10576>
- Pramesti, B. T., Anggraeny, F. T., Syaifullah, W., & Saputra, J. (2025). Klasifikasi Data Tidak Seimbang Menggunakan SMOTE-ENN dengan Optimasi Bayesian. *JiIP (Jurnal Ilmiah Ilmu Pendidikan)*, 8(10), 12207. <https://doi.org/https://doi.org/10.54371/jiip.v8i10.9456>
- Rad, H. F., Bordbar, S., Bahmaei, J., Vejdani, M., & Yusefi, A. R. (2025). Predicting academic procrastination of students based on academic self-efficacy and emotional regulation difficulties. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-87664-7>
- Raharjani, A., Gede Dhika Widarnandana, I., & Nyoman Ari Indra Dewi, N. (2023). Prokrastinasi Akademik pada Mahasiswa Tingkat Akhir yang Menyusun Skripsi. *Jurnal Kesehatan, Sains, dan Teknologi (JAKASAKTI)*, 2(3). <https://doi.org/https://doi.org/10.36002/js.v2i3.2695>
- Sartana, B. T., Nugroho, S. M. S., Yuhana, U. L., & Purnomo, M. H. (2025). Multimodel Prediction Score Based on Academic Procrastination Behavior in E-Learning. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 14(1), 1–12. <https://doi.org/10.23887/janapati.v14i1.85880>
- Tri Yulia, R., & Armianti, A. (2025). Analisis Prokrastinasi Akademik Mahasiswa Fakultas Ekonomi dan Bisnis Universitas Negeri Padang dalam Menyelesaikan Tugas Perkuliahan. *Jurnal Pendidikan Tambusai*, 9(1). <https://doi.org/https://doi.org/10.31004/jptam.v9i1.25073>
- Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J. T., & Burton, J. (2025). A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. *Clinical and Translational Science*, 18(3). <https://doi.org/10.1111/cts.70172>