

Implementasi Hybrid Machine Learning (Random Forest dan XGBoost) Untuk Prediksi Kinerja Akademik Mahasiswa Universitas Dipa Makassar

Khaerunnisa.S^{1*}, Fitriah Nabila Putri Tamar¹, Indra Samsie², Andi Irmayana²

¹Teknik Informatika, Universitas Dipa Makassar, Makassar, Indonesia

²Universitas Dipa Makassar, Makassar, Indonesia

e-mail: ¹*khaerunnisa250403@gmail.com, ²fitrihanabila127@gmail.com,

³indrasamsie@dipanegara.ac.id, ⁴irmayana180985@gmail.com

Abstrak

Penelitian ini bertujuan mengklasifikasikan kinerja akademik mahasiswa menggunakan pendekatan hybrid machine learning berbasis stacking ensemble. Dataset terdiri dari data akademik mahasiswa angkatan 2023 Universitas Dipa Makassar berupa total SKS lulus selama tiga semester (SKSL) sebagai variabel input, dengan label kinerja akademik dibentuk berdasarkan IPK yang dikategorikan menjadi Tidak Baik (IPK < 2.00), Baik (IPK 2.00–2.99), dan Sangat Baik (IPK 3.00–4.00). Model hybrid mengombinasikan Random Forest dan XGBoost sebagai base learners, serta Logistic Regression sebagai meta-learner. Tahapan penelitian meliputi preprocessing data, pembagian data stratified split 80:20, hyperparameter tuning, dan pengujian 5-fold cross-validation. Hasil menunjukkan model hybrid stacking ensemble mencapai akurasi 91%, precision 90%, recall 86%, dan F1-score 88%. Model Random Forest dan XGBoost masing-masing mencapai akurasi 92%, precision 91%, recall 87%, dan F1-score 89%. Pendekatan hybrid machine learning dengan stacking ensemble efektif mendukung evaluasi kinerja akademik mahasiswa berbasis data dan dapat digunakan untuk deteksi dini mahasiswa berisiko.

Kata kunci—Kinerja Akademik, Hybrid Machine Learning, Stacking Ensemble, Random Forest, XGBoost.

Abstract

This study aims to classify students' academic performance using a hybrid machine learning approach based on stacking ensemble. The dataset consists of academic records from the 2023 cohort at Universitas Dipa Makassar, using total passed credits over three semesters (SKSL) as input variable, with performance labels based on IPK categorized as Poor (IPK < 2.00), Good (IPK 2.00–2.99), and Excellent (IPK 3.00–4.00). The hybrid model combines Random Forest and XGBoost as base learners with Logistic Regression as meta-learner. Research stages include data preprocessing, stratified split 80:20, hyperparameter tuning, and 5-fold cross-validation. Results show the hybrid stacking ensemble achieved accuracy of 91%, precision 90%, recall 86%, and F1-score 88%. Random Forest and XGBoost models each achieved accuracy of 92%, precision 91%, recall 87%, and F1-score 89%. The hybrid machine learning approach with stacking ensemble is effective for data-driven evaluation of academic performance and can be used for early detection of at-risk students.

Keywords—Academic Performance, Hybrid Machine Learning, Stacking Ensemble, Random Forest, XGBoost.

1. PENDAHULUAN

Perkembangan machine learning telah memberikan manfaat besar di berbagai bidang, termasuk pendidikan tinggi. Penerapan machine learning dalam konteks akademik terbukti mampu memprediksi hasil pembelajaran dengan akurasi tinggi, sehingga memungkinkan intervensi dini terhadap mahasiswa yang berisiko mengalami penurunan prestasi [1], [2], [3]. Kemampuan prediksi ini memungkinkan perguruan tinggi untuk mengevaluasi kinerja akademik secara sistematis, mengoptimalkan alokasi sumber daya, serta meningkatkan kualitas pengambilan keputusan berbasis data [4].

Meskipun demikian, penerapan sistem prediksi akademik di banyak perguruan tinggi, khususnya di kawasan Asia, masih terbatas dan cenderung bergantung pada mekanisme bimbingan akademik konvensional [3]. Kondisi serupa juga ditemukan di Universitas Dipa Makassar, di mana pemantauan perkembangan akademik mahasiswa sebagian besar masih mengandalkan konsultasi dengan dosen Pembimbing Akademik (PA). Dalam praktiknya, tidak seluruh mahasiswa aktif melakukan konsultasi rutin, sehingga perubahan kinerja akademik sering kali teridentifikasi setelah berada pada kondisi yang cukup kritis.

Berdasarkan rekapitulasi data akademik mahasiswa angkatan 2023 yang diperoleh dari bagian ICT Universitas Dipa Makassar, terdapat 793 mahasiswa yang tercatat dalam sistem akademik. Dari jumlah tersebut, sebanyak 652 mahasiswa (82,2%) pernah mengalami penurunan IPK antarsemester, 156 mahasiswa tercatat memiliki IPK di bawah 2,00 pada periode tertentu, serta 318 mahasiswa tidak mampu menyelesaikan seluruh SKS yang diprogramkan pada salah satu semester. Data ini menunjukkan bahwa fluktuasi dan penurunan kinerja akademik merupakan fenomena yang cukup signifikan, sehingga diperlukan pendekatan prediksi dini untuk membantu proses identifikasi mahasiswa berisiko secara lebih cepat dan terukur.

Berbagai penelitian terdahulu telah memanfaatkan algoritma machine learning seperti Random Forest dan XGBoost dalam konteks prediksi kinerja akademik maupun manajemen kinerja lainnya. Random Forest terbukti efektif dalam memprediksi performa akademik dan konteks manajemen kinerja lainnya, serta memiliki stabilitas prediksi yang baik melalui mekanisme bagging yang mampu mengurangi varians dan risiko overfitting [5], [6], [2], sedangkan XGBoost unggul dalam efisiensi komputasi, regularisasi, serta efektif menangani data berukuran besar, dengan kemampuan optimalisasi hyperparameter melalui teknik grid search yang dapat meningkatkan akurasi model secara signifikan [7], [8], [9], [10]. Berbagai studi juga menunjukkan bahwa model hibrida yang menggabungkan lebih dari satu algoritma menghasilkan performa yang lebih baik daripada model tunggal, hal ini dibuktikan melalui pendekatan stacking ensemble yang mengungguli model tunggal dengan akurasi 82% dan AUC 87,15%, serta terbukti efektif dalam menghadapi ketidakseimbangan data [11], [12], [13]. Namun, penerapan pendekatan hibrida dalam prediksi kinerja akademik mahasiswa masih terbatas, dan sebagian besar penelitian sebelumnya hanya memanfaatkan satu algoritma.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan pendekatan hybrid stacking ensemble learning dengan mengombinasikan algoritma Random Forest dan XGBoost sebagai base learners, serta Logistic Regression sebagai meta-learner. Pemilihan Logistic Regression sebagai meta-learner didasarkan pada kemampuannya menggabungkan output dari berbagai base learners secara efektif sekaligus memberikan interpretabilitas yang baik terhadap hasil prediksi [14]. Penelitian ini diharapkan dapat mendukung sistem pengambilan keputusan akademik serta memperkuat sistem pemantauan kinerja mahasiswa di Universitas Dipa Makassar secara lebih adaptif dan berbasis data.

2. METODE PENELITIAN

Penelitian ini menggunakan jenis penelitian deskriptif kuantitatif dengan analisis prediktif berbasis machine learning. Pendekatan ini dipilih karena penelitian memanfaatkan data akademik numerik mahasiswa untuk menggambarkan karakteristik kinerja akademik serta

mengembangkan model prediksi menggunakan algoritma Random Forest, XGBoost, dan model hybrid stacking ensemble.

2.1 Dataset dan Sampel penelitian

Dataset yang digunakan merupakan data akademik mahasiswa angkatan 2023 Program Studi Teknik Informatika, Sistem Informasi, dan Rekayasa Perangkat Lunak Universitas Dipa Makassar yang diperoleh dari bagian ICT. Dari total 793 mahasiswa angkatan 2023, sebanyak 670 mahasiswa termasuk dalam populasi penelitian setelah difokuskan pada tiga program studi tersebut. Teknik pengambilan sampel menggunakan metode total sampling dengan kriteria mahasiswa telah menempuh tiga semester dan memiliki data akademik lengkap. Setelah proses verifikasi, diperoleh 609 mahasiswa sebagai sampel akhir penelitian.

2.2 Deskripsi Variabel Penelitian

Pada data awal yang telah didapatkan, memuat beberapa atribut akademik, antara lain STB, semester, SKSL, SKSP, IPS, dan IPK. Dalam penelitian ini, variabel input (X) yang digunakan adalah total SKS lulus kumulatif selama tiga semester (SKSL), karena dianggap merepresentasikan progres akademik mahasiswa.

Variabel target (Y) berupa kategori kinerja akademik yang diturunkan dari nilai IPK kumulatif tiga semester. IPK diklasifikasikan menjadi tiga kategori, yaitu Tidak Baik ($IPK < 2,00$), Baik ($2,00-2,99$), dan Sangat Baik ($\geq 3,00$). Variabel IPS tidak digunakan sebagai fitur karena berpotensi menyebabkan data leakage, mengingat IPS merupakan komponen pembentuk IPK yang dijadikan dasar pembentukan label.

Pemilihan SKSL sebagai satu-satunya variabel input juga didasarkan pada hasil eksplorasi awal yang menunjukkan bahwa penggunaan variabel IPS menghasilkan akurasi model yang sangat tinggi (mendekati 100%). Hal ini mengindikasikan adanya potensi data leakage, karena IPS merupakan komponen langsung dalam perhitungan IPK yang digunakan sebagai dasar pembentukan label. Oleh karena itu, penggunaan IPS dapat menyebabkan model mempelajari hubungan yang bersifat deterministik terhadap label, bukan pola prediktif yang sebenarnya.

Meskipun penggunaan satu fitur membatasi kompleksitas model, pendekatan ini dipilih untuk menjaga validitas model serta menghindari bias akibat kebocoran data. Selain itu, SKSL dinilai cukup representatif dalam menggambarkan progres akademik mahasiswa pada tahap awal perkuliahan.

2.3 Preprocessing Data

Data awal berbentuk vertikal, di mana satu mahasiswa memiliki beberapa baris sesuai semester. Data kemudian ditransformasi menjadi format horizontal sehingga setiap mahasiswa direpresentasikan dalam satu baris.

Tahap preprocessing meliputi seleksi mahasiswa dengan data lengkap tiga semester, transformasi struktur data dari format vertikal ke horizontal, pemeriksaan nilai kosong, penghapusan duplikasi, dan penyesuaian format nilai numerik. Dari total 609 sampel yang memenuhi kriteria kelengkapan tiga semester, seluruh data dinyatakan valid setelah melalui proses pemeriksaan dan tidak ditemukan duplikasi maupun nilai kosong yang signifikan. Hasil preprocessing menghasilkan dataset terstruktur dengan satu fitur utama (SKSL) dan satu label klasifikasi.

2.4 Pembagian Data dan Validasi

Dataset dibagi menggunakan metode train-test split dengan proporsi 80% data latih dan 20% data uji, menghasilkan 487 data latih dan 122 data uji. Pembagian dilakukan menggunakan teknik stratified sampling untuk mempertahankan proporsi masing-masing kelas pada data latih dan data uji. Distribusi kelas pada data latih adalah Sangat Baik (330 sampel), Baik (92 sampel), dan Tidak Baik (65 sampel), sedangkan pada data uji adalah Sangat Baik (83 sampel), Baik (23 sampel), dan Tidak Baik (16 sampel). Parameter `random_state=42` digunakan untuk memastikan reproduktibilitas hasil. Untuk menjaga stabilitas model dan meminimalkan overfitting, pelatihan dilakukan menggunakan 5-fold cross-validation pada data latih.

Dalam penelitian ini, variabel input dan target berasal dari periode waktu yang sama, yaitu hingga akhir semester ketiga. Variabel input berupa SKSL mencerminkan jumlah SKS yang berhasil diselesaikan selama tiga semester, sedangkan variabel target berupa IPK kumulatif juga dihitung pada periode yang sama. Meskipun terdapat korelasi antara kedua variabel tersebut, SKSL tidak secara langsung mengandung informasi nilai akademik yang membentuk IPK, sehingga tidak menyebabkan data leakage secara langsung.

Namun demikian, pendekatan ini lebih merepresentasikan klasifikasi berbasis data historis dibandingkan prediksi masa depan. Oleh karena itu, penelitian selanjutnya disarankan menggunakan pendekatan temporal berbeda, seperti menggunakan data hingga semester ke-2 untuk memprediksi kinerja pada semester ke-3.

2.5 Pembangunan Model

Penelitian ini membangun model prediksi menggunakan pendekatan hybrid stacking ensemble yang mengombinasikan Random Forest dan XGBoost sebagai base learner serta Logistic Regression sebagai meta-learner.

2.5.1 Random Forest

Random Forest adalah salah satu algoritma ensemble learning berbasis decision tree yang diperkenalkan oleh Breiman pada tahun 2001. Model random forest terdiri dari banyak pohon keputusan dan penentuan klasifikasi dilakukan berdasarkan hasil voting dari pohon-pohon keputusan yang terbentuk [15].

Implementasi dilakukan menggunakan RandomForestClassifier. Optimasi hyperparameter dilakukan menggunakan GridSearchCV dengan 5-fold cross-validation pada data latih. Parameter yang diuji meliputi jumlah pohon (`n_estimators`), kedalaman maksimum pohon (`max_depth`), `min_samples_split`, dan `min_samples_leaf`. Model dengan kombinasi parameter terbaik digunakan sebagai base learner pertama dalam arsitektur stacking.

2.5.2 XGBoost

XGBoost (Extreme Gradient Boosting) merupakan pengembangan dari algoritma Gradient Boosting Machine (GBM) yang dirancang untuk memberikan performa prediksi tinggi dengan efisiensi komputasi yang lebih baik [16]. XGBoost bekerja dengan membangun model secara bertahap (stage-wise), di mana setiap model baru difokuskan untuk memperbaiki kesalahan yang dihasilkan model sebelumnya.

Model diimplementasikan menggunakan XGBClassifier dengan parameter awal `use_label_encoder=False`, `eval_metric='mlogloss'`, dan `random_state=42`. Optimasi hyperparameter dilakukan menggunakan RandomizedSearchCV dengan 5-fold cross-validation. Parameter yang diuji meliputi `n_estimators`, `max_depth`, `learning_rate`, dan `subsample`. Model terbaik hasil proses tuning digunakan sebagai base learner kedua dalam arsitektur stacking.

2.5.3 Hybrid Stacking Ensemble

Model hybrid dibangun menggunakan pendekatan stacking ensemble dengan Random Forest dan XGBoost sebagai base learner. Prediksi probabilitas dari kedua model tersebut digunakan sebagai meta-features yang menjadi input bagi meta-learner.

Meta-learner yang digunakan adalah Logistic Regression dengan parameter `max_iter=1000` untuk memastikan proses optimasi mencapai konvergensi. Proses stacking dilakukan dengan mekanisme 5-fold cross-validation untuk membentuk meta-features secara objektif dan mencegah terjadinya data leakage. Model stacking ensemble yang dihasilkan kemudian digunakan sebagai model utama dalam penelitian ini untuk melakukan prediksi kategori kinerja akademik mahasiswa.

Pada proses stacking ensemble, pembentukan meta-features dilakukan menggunakan pendekatan out-of-fold (OOF) prediction pada data latih. Setiap base learner dilatih menggunakan sebagian data latih dan menghasilkan prediksi pada fold validasi yang tidak digunakan dalam proses pelatihan. Prediksi tersebut kemudian dikumpulkan sebagai meta-features yang menjadi input bagi meta-learner.

Pendekatan ini bertujuan untuk mencegah terjadinya data leakage antara base learner dan meta-learner. Selain itu, data uji sepenuhnya dipisahkan dan tidak digunakan dalam proses pelatihan maupun pembentukan meta-features, sehingga evaluasi model tetap objektif.

2.6 Evaluasi Model

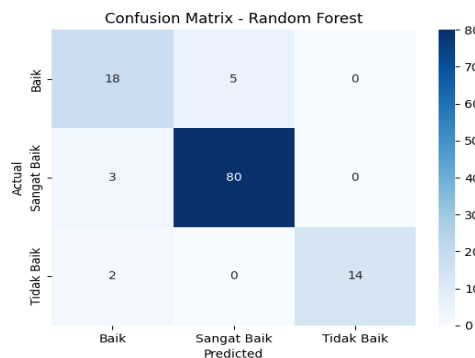
Evaluasi performa model dilakukan menggunakan data uji sebanyak 20% dari total dataset. Metrik evaluasi yang digunakan meliputi accuracy, precision, recall, dan F1-score. Selain itu, confusion matrix digunakan untuk menganalisis distribusi hasil klasifikasi pada masing-masing kategori kinerja akademik, sehingga dapat memberikan gambaran lebih rinci mengenai kemampuan model dalam mengklasifikasikan setiap kelas.

3. HASIL DAN PEMBAHASAN

3.1 Evaluasi Model Random Forest

Model Random Forest dilatih menggunakan data latih dengan parameter terbaik hasil proses hyperparameter tuning. Selanjutnya model diuji menggunakan data uji sebanyak 122 sampel untuk mengevaluasi performa klasifikasi. Evaluasi model dilakukan menggunakan confusion matrix serta metrik accuracy, precision, recall, dan F1-score.

Hasil klasifikasi model Random Forest dapat dilihat pada confusion matriks yang ditunjukkan pada gambar 1.



Gambar 1 Confusion Matrix Random Forest

Berdasarkan confusion matrix tersebut, sebagian besar data berhasil diklasifikasikan dengan benar oleh model. Hasil pengujian, model Random Forest memperoleh nilai accuracy sebesar 0,92 (91,8%), yang menunjukkan bahwa sebagian besar data uji berhasil diklasifikasikan dengan benar. Selanjutnya performa model dianalisis menggunakan metrik evaluasi yang disajikan pada tabel 1.

Tabel 1 Metrik Evaluasi Random Forest

Kategori	Precision	Recall	F1-Score
Baik	0,78	0,76	0,78
Sangat Baik	0,94	0,96	0,95
Tidak Baik	1,00	0,88	0,93

Model menunjukkan performa terbaik pada kategori Sangat Baik dengan F1-score sebesar 0,95. Sementara itu kategori Tidak Baik memiliki precision sempurna sebesar 1,00, yang berarti seluruh prediksi pada kategori tersebut tepat sasaran, meskipun masih terdapat beberapa data aktual yang tidak terdeteksi. Nilai macro average F1-score sebesar 0,89 menunjukkan bahwa performa model relatif seimbang antar kelas dan tidak terlalu dipengaruhi oleh dominasi kelas mayoritas.

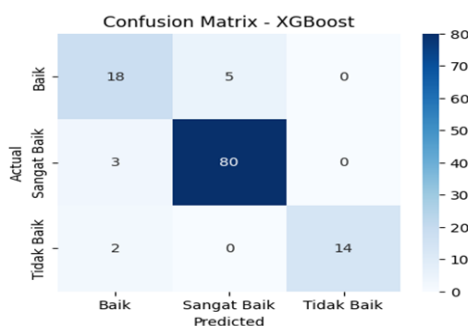
Perlu diperhatikan bahwa distribusi kelas pada dataset cenderung tidak seimbang, dengan dominasi pada kategori *Sangat Baik*. Kondisi ini berpotensi memengaruhi performa model, khususnya dalam mendeteksi kelas minoritas seperti *Tidak Baik*. Meskipun demikian, penggunaan metrik macro average F1-score dalam evaluasi menunjukkan bahwa model tetap mampu mempertahankan keseimbangan performa antar kelas.

Selain itu, nilai recall sebesar 0,88 pada kategori *Tidak Baik* menunjukkan bahwa sebagian besar mahasiswa yang berpotensi mengalami kinerja akademik rendah berhasil terdeteksi oleh model, meskipun masih terdapat sebagian kecil yang belum teridentifikasi.

3.2 Evaluasi Model XGBoost

Model XGBoost selanjutnya diuji menggunakan data uji yang sama untuk menjaga konsistensi perbandingan performa model. Hasil evaluasi menunjukkan bahwa model XGBoost juga memperoleh nilai accuracy sebesar 0,92 (91,8%).

Hasil klasifikasi model XGBoost dapat dilihat pada confusion matrix pada gambar 2.



Gambar 2 Confusion Matrix XGBoost

Distribusi hasil klasifikasi menunjukkan bahwa dari 23 data kategori Baik, sebanyak 18 data berhasil diprediksi dengan benar, sedangkan 5 data lainnya salah diklasifikasikan sebagai Sangat Baik. Pada kategori Sangat Baik, sebanyak 80 dari 83 data berhasil diprediksi dengan benar. Sementara itu, pada kategori Tidak Baik, sebanyak 14 dari 16 data berhasil diklasifikasikan dengan tepat. Nilai metrik evaluasi XGBoost dapat dilihat pada tabel 2.

Tabel 2 Metrik Evaluasi XGBoost

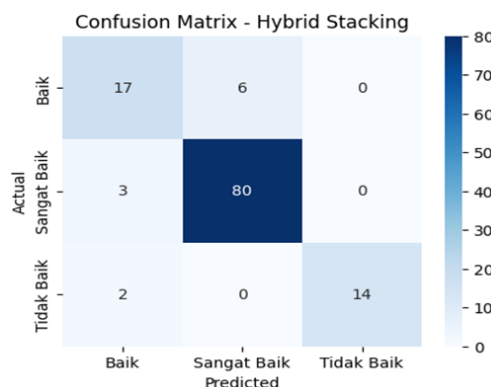
Kategori	Precision	Recall	F1-Score
Baik	0,78	0,76	0,78
Sangat Baik	0,94	0,96	0,95
Tidak Baik	1,00	0,88	0,93

Hasil ini menunjukkan bahwa model XGBoost memiliki pola performa yang sangat mirip dengan Random Forest, baik dari sisi accuracy maupun distribusi metrik evaluasi. Hal ini mengindikasikan bahwa kedua model cenderung menangkap pola yang serupa dalam data, sehingga perbedaan kontribusi informasi yang dihasilkan relatif kecil.

3.3 Evaluasi Model Hybrid Stacking Ensemble

Setelah pengujian model tunggal, penelitian ini juga menguji model hybrid menggunakan metode stacking ensemble yang menggabungkan Random Forest dan XGBoost sebagai base learner serta Logistic Regression sebagai meta-learner. Pengujian dilakukan menggunakan data uji yang sama sebanyak 122 sampel.

Hasil klasifikasi model Hybrid Stacking dapat dilihat pada confusion matrix pada gambar 3.



Gambar 3 Confusion Matrix Hybrid Stacking Ensemble

Berdasarkan confusion matrix tersebut, dari 23 data kategori Baik, sebanyak 17 data berhasil diprediksi dengan benar, sedangkan 6 data lainnya salah diklasifikasikan sebagai Sangat Baik. Pada kategori Sangat Baik, sebanyak 80 dari 83 data berhasil diprediksi dengan benar, sementara 3 data lainnya salah diprediksi sebagai Baik. Pada kategori Tidak Baik, sebanyak 14 dari 16 data berhasil diprediksi dengan benar, sedangkan 2 data lainnya salah diprediksi sebagai Baik.

Hasil evaluasi selanjutnya dianalisis menggunakan metrik accuracy, precision, recall, dan F1-score yang disajikan pada tabel 3.

Tabel 3 Metrik Evaluasi Hybrid Stacking Ensemble

Kategori	Precision	Recall	F1-Score
Baik	0,77	0,74	0,76
Sangat Baik	0,93	0,96	0,95
Tidak Baik	1,00	0,88	0,93

Secara keseluruhan, model Hybrid Stacking Ensemble memperoleh nilai accuracy sebesar 0,91 (90,98%) dengan macro average F1-score sebesar 0,88. Kategori Sangat Baik menunjukkan performa terbaik dengan F1-score sebesar 0,95, sedangkan kategori Baik mengalami sedikit penurunan performa dibandingkan model tunggal dengan F1-score sebesar 0,76.

Penurunan performa pada model stacking ini menguatkan temuan pada tahap metode bahwa keberagaman informasi antar base learner menjadi faktor penting dalam keberhasilan ensemble. Dalam penelitian ini, penggunaan satu fitur utama serta kemiripan pola prediksi antara Random Forest dan XGBoost menyebabkan meta-learner tidak memperoleh informasi tambahan yang cukup untuk meningkatkan akurasi secara signifikan.

3.4 Perbandingan Kinerja Model

Untuk memahami perbedaan performa antar model yang digunakan dalam penelitian ini, dilakukan analisis perbandingan terhadap metrik evaluasi yang dihasilkan oleh masing-masing algoritma yang disajikan pada tabel 4 berikut.

Tabel 4 Perbandingan Hasil Model

Model	Accuracy	Precision (Macro Avg)	Recall (Macro Avg)	F1-score (Macro Avg)
Random Forest	0,92	0,91	0,87	0,89
XGBoost	0,92	0,91	0,87	0,89
Hybrid Stacking Ensemble	0,91	0,90	0,86	0,88

Jika dibandingkan dengan penelitian terdahulu yang menggunakan pendekatan machine learning dalam klasifikasi kinerja akademik mahasiswa, hasil yang diperoleh dalam penelitian ini tergolong kompetitif. Beberapa penelitian sebelumnya melaporkan tingkat akurasi pada rentang 80% hingga 90% tergantung pada jumlah fitur dan kompleksitas model yang digunakan.

Dengan capaian akurasi sebesar 91%–92%, model yang dikembangkan dalam penelitian ini menunjukkan bahwa meskipun hanya menggunakan satu fitur utama, performa yang dihasilkan tetap mampu bersaing dengan model lain yang menggunakan lebih banyak variabel. Hal ini menunjukkan bahwa SKSL sebagai indikator progres akademik memiliki kontribusi yang signifikan dalam proses klasifikasi kinerja mahasiswa.

Hasil penelitian menunjukkan bahwa algoritma Random Forest dan XGBoost memiliki kemampuan yang sangat baik dalam melakukan klasifikasi kinerja akademik mahasiswa. Kedua algoritma tersebut mampu mencapai tingkat akurasi yang tinggi serta menunjukkan stabilitas performa pada berbagai metrik evaluasi.

Random Forest memiliki keunggulan dalam mengurangi variansi melalui mekanisme bagging yang membangun banyak pohon keputusan secara paralel. Pendekatan ini membuat model lebih stabil dan mampu mengurangi risiko overfitting pada data pelatihan. Sementara itu, XGBoost menggunakan pendekatan boosting yang membangun model secara bertahap dengan memperbaiki kesalahan prediksi dari model sebelumnya. Mekanisme optimisasi berbasis gradien serta fitur regularisasi pada XGBoost memungkinkan model menangkap pola kompleks dalam data secara lebih efektif.

Model hybrid stacking ensemble dibangun dengan tujuan menggabungkan kekuatan kedua algoritma tersebut melalui meta-learner Logistic Regression. Namun, hasil penelitian menunjukkan bahwa performa stacking belum mampu melampaui model tunggal. Hal ini kemungkinan disebabkan oleh pola prediksi yang dihasilkan oleh Random Forest dan XGBoost yang relatif mirip, sehingga meta-learner tidak memperoleh informasi tambahan yang cukup untuk meningkatkan performa klasifikasi secara signifikan. Kondisi ini juga dapat dipengaruhi oleh jumlah fitur yang terbatas pada dataset penelitian. Meskipun demikian, pendekatan hybrid stacking ensemble tetap menunjukkan performa yang kompetitif dan dapat digunakan sebagai alternatif model prediksi dalam konteks analisis kinerja akademik mahasiswa.

Untuk memperkuat analisis perbandingan, dilakukan komparasi dengan beberapa penelitian terdahulu yang relevan yang disajikan pada Tabel 5.

Tabel 5 Perbandingan Hasil Penelitian Terdahulu (State of The Art)

No	Penelitian	Metode	Dataset	Akurasi	Keterangan
1	Inyang & Johnson (2025)[2]	Random Forest, XGBoost	Data akademik mahasiswa	98,88%	Menggunakan banyak fitur akademik
2	Kuswanto & Hakim (2025)[6]	Random Forest	Data performa akademik	85%	Dataset terbatas
3	Swastikawati & Utami (2025)[13]	Stacking Ensemble	Data prediksi kelulusan mahasiswa	82%	Fokus pada prediksi kelulusan
4	Osman et al. (2025)[12]	Hybrid Machine Learning	Dataset klasifikasi umum	99,89%	Bukan dataset akademik

Jika dibandingkan dengan penelitian terdahulu, hasil yang diperoleh dalam penelitian ini tergolong kompetitif. Beberapa penelitian sebelumnya melaporkan tingkat akurasi yang tinggi, bahkan mencapai di atas 95%, namun umumnya menggunakan jumlah fitur yang lebih banyak serta dataset dengan karakteristik yang berbeda.

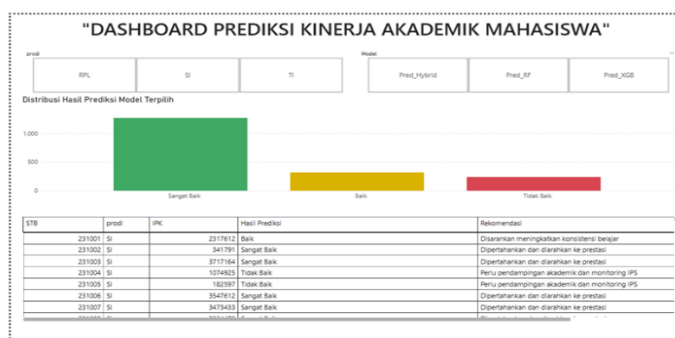
Dalam penelitian ini, penggunaan satu fitur utama berupa SKSL secara sengaja dipilih untuk menghindari potensi data leakage dan menjaga validitas model. Meskipun sederhana, hasil akurasi sebesar 91%–92% menunjukkan bahwa fitur tersebut tetap memiliki kontribusi yang signifikan dalam proses klasifikasi kinerja akademik mahasiswa.

Perbedaan performa antara model tunggal dan model hybrid relatif kecil, yaitu sekitar 1%. Model Random Forest dan XGBoost menunjukkan akurasi yang sedikit lebih tinggi (92%) dibandingkan model hybrid stacking ensemble (91%). Hal ini mengindikasikan bahwa dalam kondisi fitur yang terbatas dan pola data yang relatif sederhana, model tunggal sudah mampu menangkap pola data secara optimal.

Di sisi lain, model hybrid stacking ensemble tetap memiliki nilai penting dari sisi fleksibilitas dan potensi generalisasi, karena menggabungkan karakteristik dua algoritma yang berbeda. Namun, pada penelitian ini, kemiripan pola prediksi antara Random Forest dan XGBoost menyebabkan meta-learner tidak memperoleh informasi tambahan yang cukup untuk meningkatkan performa secara signifikan.

Dalam konteks implementasi praktis, khususnya untuk deteksi dini mahasiswa berisiko, metrik recall pada kategori Tidak Baik menjadi aspek yang penting. Nilai recall yang tinggi menunjukkan kemampuan model dalam mengidentifikasi mahasiswa yang membutuhkan intervensi akademik. Oleh karena itu, pemilihan model tidak hanya bergantung pada nilai akurasi, tetapi juga mempertimbangkan trade-off antara kompleksitas model, stabilitas prediksi, dan kebutuhan implementasi dalam konteks nyata.

3.5 Implementasi Dashboard Power BI



Gambar 4 Dashboard Prediksi Kinerja Akademik Mahasiswa

Dashboard yang ada pada gambar 4 tersebut, dikembangkan dalam satu halaman interaktif untuk memudahkan eksplorasi dan analisis hasil prediksi kinerja akademik mahasiswa. Pada bagian atas dashboard disediakan fitur slicer untuk memilih Program Studi (RPL, SI, dan TI) serta model prediksi yang digunakan, yaitu Random Forest, XGBoost, atau Hybrid. Pemilihan model secara otomatis memperbarui seluruh visualisasi yang ditampilkan.

Visualisasi utama berupa grafik batang yang menunjukkan distribusi hasil prediksi kinerja akademik mahasiswa ke dalam tiga kategori, yaitu Sangat Baik, Baik, dan Tidak Baik. Grafik ini bersifat dinamis sehingga memungkinkan pengguna melakukan analisis dan perbandingan hasil prediksi antar model dan program studi.

Selain itu, dashboard juga menampilkan tabel detail mahasiswa yang memuat informasi STB, Program Studi, IPK, hasil prediksi, serta rekomendasi akademik yang dihasilkan berdasarkan aturan logika tertentu. Dashboard ini berfungsi sebagai alat bantu visualisasi dan analisis untuk mendukung pemantauan kinerja akademik serta pengambilan keputusan berbasis data.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi kinerja akademik mahasiswa menggunakan pendekatan machine learning berbasis data akademik tiga semester awal. Model yang dibangun mampu mengidentifikasi kategori kinerja akademik mahasiswa berdasarkan total SKS lulus (SKSL) sebagai indikator progres akademik. Pemilihan SKSL sebagai fitur tunggal dilakukan secara sengaja untuk menghindari potensi data leakage, sehingga model yang dihasilkan mencerminkan kemampuan prediksi yang lebih objektif dan realistis.

Hasil penelitian menunjukkan bahwa model Random Forest dan XGBoost sebagai model tunggal memberikan performa terbaik dengan akurasi sebesar 91,8%, sementara model hybrid stacking ensemble memperoleh akurasi sebesar 90,98%. Meskipun performa model hybrid sedikit lebih rendah, pendekatan stacking tetap memberikan nilai tambah dari sisi fleksibilitas dan potensi generalisasi dengan menggabungkan karakteristik dua algoritma yang berbeda.

Temuan ini menunjukkan bahwa dalam kondisi fitur yang terbatas dan pola data yang relatif sederhana, model tunggal sudah mampu memberikan performa yang optimal, sehingga peningkatan melalui metode ensemble tidak selalu signifikan. Dengan demikian, pemilihan model dalam konteks prediksi kinerja akademik tidak hanya bergantung pada tingkat akurasi, tetapi juga mempertimbangkan kompleksitas model, stabilitas, serta tujuan implementasi dalam praktik.

Secara keseluruhan, penelitian ini menunjukkan bahwa pendekatan machine learning berbasis fitur sederhana tetap mampu menghasilkan performa yang kompetitif, serta berpotensi dimanfaatkan sebagai dasar dalam pengembangan sistem deteksi dini mahasiswa berisiko secara berbasis data.

5. SARAN

Penelitian selanjutnya disarankan untuk menambahkan variabel akademik maupun non-akademik sebagai fitur input, seperti Indeks Prestasi Semester (IPS), tingkat kehadiran, aktivitas pembelajaran, serta faktor sosial-ekonomi mahasiswa, dengan tetap memperhatikan potensi data leakage melalui pengaturan waktu penggunaan data.

Selain itu, penggunaan pendekatan temporal, seperti memanfaatkan data hingga semester ke-2 untuk memprediksi kinerja pada semester ke-3, dapat dipertimbangkan untuk meningkatkan validitas prediksi secara prospektif. Penelitian lanjutan juga dapat mengeksplorasi metode ensemble lain, seperti boosting atau voting, serta variasi meta-learner pada stacking untuk mengevaluasi potensi peningkatan performa model.

Lebih lanjut, model yang dihasilkan dalam penelitian ini dapat dikembangkan menjadi sistem pendukung keputusan yang membantu pihak akademik dalam melakukan pemantauan dan intervensi terhadap mahasiswa secara lebih terarah dan berbasis data. Selain itu, integrasi model ke dalam sistem dashboard interaktif seperti Power BI dapat dikembangkan lebih lanjut untuk mendukung implementasi nyata di lingkungan perguruan tinggi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan bantuan dalam penyusunan penelitian ini. Terima kasih disampaikan kepada dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penelitian hingga penyusunan artikel ini. Selain itu, penulis turut menyampaikan apresiasi kepada keluarga dan rekan-rekan yang telah memberikan dukungan moral dan motivasi sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] H. N. T. Cam, A. Sarlan, and N. I. Arshad, "A hybrid model integrating recurrent neural networks and the semi-supervised support vector machine for identification of early student dropout risk," *PeerJ Comput. Sci.*, vol. 10, pp. 1–31, 2024, doi: 10.7717/peerj-cs.2572.
- [2] U. P. Inyang and E. A. Johnson, "Performance Comparison of Xgboost and Random Forest for The Prediction of Students Academic Performance," *Eur. J. Comput. Sci. Inf. Technol.*, vol. 13, no. 2, pp. 1–21, 2025, doi: 10.37745/ejsit.2013/vol13n2121.
- [3] T. Mazhar, T. Shahzad, W. Waheed, A. Waheed, and H. Hamam, "Predictive analytics in education-enhancing student achievement through machine learning," *Soc. Sci. Humanit. Open*, vol. 12, p. 101824, 2025.
- [4] R. Guevara-reyes, R. Andrade, F. Cox-riqueti, and W. Villegas-ch, "Machine learning models for academic performance prediction: interpretability and application in educational," no. August, pp. 1–27, 2025, doi: 10.3389/feduc.2025.1632315.
- [5] S. Basnet, "A Data-Driven Revolution in Performance Management : Harnessing the Power

- of Random Forest Algorithm,” vol. 12, no. May, pp. 546–562, 2025.
- [6] J. Kuswanto and L. Hakim, “Penerapan Algoritma Random Forest untuk memprediksi Performa Akademik Mahasiswa,” vol. 5, no. 1, pp. 262–270, 2025.
- [7] A. Hermawan, A. Saputra, M. D. Rafi, and S. Basmallah, “Implementing XGBoost Model for Predicting Customer Churn in E-Commerce Platforms,” vol. 3, no. 2, 2025.
- [8] H. Wijaya, D. P. Hostiadi, and E. Triandini, “Optimation Of Xgboost Algorithm Using Parameter Tunning In Retail Sales Prediction Jurnal Nasional Pendidikan Teknik Informatika : Janapati | 770,” vol. 13, no. 3, pp. 769–786, 2024.
- [9] M. Imani, A. Beikmohammadi, and H. R. Arabnia, “Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels,” *Technologies*, vol. 13, no. 3, pp. 1–40, 2025, doi: 10.3390/technologies13030088.
- [10] A. Sutou and J. Wang, “Influence-Balanced XGBoost: Improving XGBoost for Imbalanced Data Using Influence Functions,” *IEEE Access*, vol. 12, no. December, pp. 193473–193486, 2024, doi: 10.1109/ACCESS.2024.3520159.
- [11] F. Alfawaid *et al.*, “Penerapan metode multiple machine learning (hybrid model) untuk mendeteksi link phishing sebagai upaya,” vol. 17, 2023.
- [12] B. M. Osman, M. Sheikh, A. Jirow, and D. A. Aser, “A Novel Hybrid Approach to Machine Learning for Enhanced Model Precision and Classification,” vol. 12, no. 2, pp. 91–101, 2025.
- [13] C. Swastikawati and E. Utami, “Graduation Prediction for Prospective University Students Using Stacking Ensemble Learning,” *Sist. J. Sist. Inf.*, vol. 14, no. 6, pp. 2707–2723, 2025.
- [14] T. Tong and Z. Li, “Predicting learning achievement using ensemble learning with result explanation,” *PLoS One*, vol. 20, no. 1, pp. 1–25, 2025, doi: 10.1371/journal.pone.0312124.
- [15] A. Baita, I. A. Prasetyo, and N. Cahyono, “HYPERPARAMETER TUNING ON RANDOM FOREST FOR DIAGNOSE,” vol. 6, no. 2, pp. 138–143, 2023, doi: 10.33387/jiko.v6i2.6389.
- [16] T. Chen, “XGBoost : A Scalable Tree Boosting System”.