 ISSN:2302-8149
e-ISSN:2540-9719

Prediksi Kelulusan Calon Mahasiswa dengan *Stacking Ensemble Learning*

Graduation Prediction for Prospective University Students using Stacking Ensemble Learning

[1]Claudia Swastikawati*, [2]Ema Utami
[1,2]Informatika Program Magister, Universitas Amikom Yogyakarta
[1,2]Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281, Indonesia
*e-mail: claudia.swastikawati@students.amikom.ac.id

(***received*:** 15 July 2025, ***revised*:** 26 July 2025, ***accepted*:** 27 July 2025)

Abstrak

Kelulusan mahasiswa merupakan indikator penting dalam akreditasi dan menjadi bagian dari strategi pengelolaan mutu di perguruan tinggi. Sehingga, prediksi kelulusan calon mahasiswa sejak dini diperlukan untuk meningkatkan efektivitas dalam pengambilan keputusan penerimaan mahasiswa berbasis data. Perbedaan tingkat kelulusan mahasiswa dipengaruhi oleh kombinasi faktor akademik, demografis, ekonomi, dan keluarga. Penelitian ini menerapkan metode *Stacking Ensemble Learning* dengan menggabungkan *Random Forest*, *K-Nearest Neighbors*, dan *Support Vector Machine* menggunakan XGBoost sebagai *meta-learner*. Dataset yang digunakan merupakan gabungan data pendaftaran mahasiswa baru dan laporan status kelulusan dari sistem NeoFeeder PDDikti, mencakup 16 fitur variabel akademik dan non-akademik. Model diuji menggunakan evaluasi berbasis akurasi, *precision*, *recall*, *F1-score*, dan *Area Under Curve* (AUC). Hasil penelitian menunjukkan bahwa model *stacking ensemble* memberikan performa terbaik dibandingkan model tunggal, dengan akurasi 82%, *weighted F1-score* 80%, dan AUC pada data uji 87,15%. Dengan hasil ini, penelitian memberikan kontribusi berupa fitur yang digunakan serta penerapan model *ensemble* dalam membangun sistem prediksi berbasis *machine learning*, khususnya dalam menghadapi ketidakseimbangan data dan memperbaiki akurasi klasifikasi.

Kata kunci: stacking, multi-model classification, prediksi kelulusan, machine learning

Abstract

Student graduation is an important indicator in accreditation and serves as part of quality management strategies in higher education. Therefore, early prediction of student graduation is necessary to improve the effectiveness of data-driven admission decision-making. Differences in student graduation rates are influenced by a combination of academic, demographic, economic, and family factors. This study applies the Stacking Ensemble Learning method by combining Random Forest, K-Nearest Neighbors, and Support Vector Machine, with XGBoost serving as the meta-learner. The dataset used integrates student admission records and graduation status reports from the NeoFeeder PDDikti system, covering 16 academic and non-academic feature variables. The model was evaluated using accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). The results show that the stacking ensemble model outperformed single models, achieving 82% accuracy, a weighted F1-score of 80%, and an AUC of 87.15% on the test data. These findings contribute both the selected feature set and the implementation of an ensemble model for building a machine learning–based prediction system, particularly in addressing data imbalance and improving classification accuracy.

Keywords: *stacking, multi-model classification, graduation prediction, machine learning*

 ISSN:2302-8149
e-ISSN:2540-9719

1 Pendahuluan

Keberhasilan akademik mahasiswa merupakan salah satu indikator utama dalam menilai efektivitas institusi pendidikan tinggi[1]. Perguruan tinggi tidak hanya bertanggung jawab dalam meningkatkan kualitas program pendidikan, tetapi juga dalam memastikan bahwa mahasiswa yang diterima memiliki potensi akademik yang kuat untuk menyelesaikan studi tepat waktu. Oleh karena itu, prediksi kinerja akademik mahasiswa menjadi aspek yang semakin penting dalam mendukung kebijakan pendidikan, seperti seleksi penerimaan mahasiswa, pemberian beasiswa, serta intervensi akademik yang berbasis data[2], [3]. Dengan meningkatnya jumlah pendaftar setiap tahun, institusi pendidikan menghadapi tantangan dalam mengelola dan menganalisis data yang dihasilkan selama proses penerimaan mahasiswa baru (PMB). Data ini mencakup faktor akademik, demografis, dan sosioekonomi yang dapat menjadi indikator keberhasilan mahasiswa dalam menyelesaikan studinya[4].

Penerapan metode *machine learning* dalam analisis prediksi akademik telah banyak dilakukan pada penelitian sebelumnya, dengan berbagai algoritma yang diterapkan, seperti *Random Forest* (RF), *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN), *Neural Networks*, dan *Decision Trees[2], [5], [6], [7]*. Studi-studi tersebut menunjukkan bahwa setiap algoritma memiliki keunggulan dan keterbatasannya masing-masing[8]. Sebagai contoh, algoritma RF mampu meningkatkan akurasi dengan bertambahnya jumlah data, sementara SVM menunjukkan performa tinggi dalam menangani data dengan dimensi yang kompleks[6]. Namun, masing-masing metode juga memiliki kelemahan, seperti ketidakefisienan pada dataset yang besar, waktu pelatihan yang lama, serta risiko *overfitting* pada model tertentu. Oleh karena itu, diperlukan pendekatan yang lebih robust untuk mengatasi keterbatasan ini dan meningkatkan akurasi prediksi kelulusan mahasiswa[9].

Salah satu pendekatan yang menjanjikan dalam meningkatkan akurasi prediksi adalah dengan menggunakan teknik *Stacking Ensemble Learning* [2], [9], [10]. Teknik ensemble learning telah mendapatkan perhatian karena kemampuannya dalam meningkatkan akurasi prediksi. Metode ini memungkinkan penggabungan beberapa model dasar (base learners) untuk menghasilkan prediksi yang lebih akurat dan stabil [8]. Dalam penelitian ini, algoritma *Random Forest, K-Nearest Neighbors*, dan *Support Vector Machine* digunakan sebagai *base models* untuk mempelajari pola dalam data akademik mahasiswa. Sementara itu, model *XGBoost* diterapkan sebagai *meta-learner* untuk mengoptimalkan hasil prediksi yang diperoleh dari ketiga model dasar tersebut. Pendekatan ini diharapkan dapat mengatasi keterbatasan masing-masing algoritma individu serta menghasilkan model prediksi yang lebih andal dalam memproyeksikan status kelulusan mahasiswa.

Penelitian ini berfokus pada dua tujuan utama, yaitu mengidentifikasi atribut demografis, akademik, dan sosioekonomi yang berperan signifikan dalam memengaruhi status kelulusan mahasiswa baik yang lulus tepat waktu, terlambat, maupun yang dikeluarkan serta mengevaluasi performa pendekatan *Stacking Ensemble* dengan *XGBoost* sebagai meta model dalam memprediksi kelulusan calon mahasiswa baru. Melalui pendekatan ini, penelitian tidak hanya berkontribusi dalam membangun model prediktif yang lebih adaptif dan akurat, tetapi juga menawarkan perspektif baru bagi institusi pendidikan tinggi dalam merumuskan kebijakan akademik berbasis data. Secara akademik, studi ini memperkaya wacana keilmuan di bidang *machine learning* dan analitik prediktif dengan menguji efektivitas teknik ensemble dalam konteks pendidikan tinggi[11]. Sementara itu, dari sisi praktis, temuan penelitian ini dimasa mendatang dapat dimanfaatkan untuk merancang sistem cerdas yang mendukung proses seleksi mahasiswa serta merumuskan intervensi akademik yang lebih tepat sasaran [1], [2], [3], [4]. Dengan demikian, pemanfaatan teknologi prediktif dalam penelitian ini diharapkan dapat menjadi landasan strategis bagi peningkatan efisiensi tata kelola akademik dan penguatan kebijakan berbasis data di lingkungan perguruan tinggi.

2 Tinjauan Literatur

Penelitian sebelumnya, model pembelajaran mesin digunakan untuk memprediksi potensi *dropout* mahasiswa di pendidikan tinggi, dimulai dari fase pendaftaran hingga semester lanjutan. Penelitian menerapkan pendekatan *Stacking ensemble* dengan menggabungkan Random Forest, SVM, dan Gradient Boosting, menghasilkan prediksi yang lebih akurat terhadap mahasiswa yang berisiko tidak menyelesaikan studinya. Hasil penelitian menunjukkan bahwa metrik *recall* menjadi indikator penting dalam menilai kinerja model untuk mengidentifikasi sebanyak mungkin mahasiswa yang

 ISSN:2302-8149
e-ISSN:2540-9719

berisiko *dropout*. Model *ensemble* dipilih pada penelitian tersebut untuk tahap pendaftaran mencapai *recall* (74,79%), namun *precision* rendah (54,27%) selanjutnya *recall* meningkat pada setiap semesternya hingga mencapai *recall* 91,55% pada semester empat perkuliahan. Sehingga, Model *ensemble* menjadi model yang sesuai karena memiliki *recall* cukup tinggi dan semakin akurat seiring bertambahnya data mahasiswa. Penggunaan data akademik yang komprehensif memperkuat model, namun fokus penelitian tersebut masih berkutat pada fase keberlangsungan studi, bukan pada prediksi status kelulusan akhir [12]. Dengan demikian, ruang eksplorasi masih terbuka untuk memperluas penerapan *ensemble* dalam konteks kelulusan tepat waktu dengan fokus pada data pendaftaran.

Penelitian selanjutnya, menyoroti pentingnya penanganan data tidak seimbang dalam prediksi performa mahasiswa menggunakan *Decision Tree* (J48), *Naïve Bayes* (NB), *Logistic Regression* (LR), RF, KNN, dan SVM. Dengan menerapkan teknik SMOTE, akurasi model meningkat secara signifikan. Meski demikian, penelitian ini terbatas pada prediksi nilai akhir mahasiswa dan belum menguji kombinasi algoritma dalam kerangka *ensemble* lanjutan [13]. Sedangkan pada penelitian lainnya, teknik *stacking* terbukti mampu meningkatkan performa klasifikasi pada dataset tidak seimbang dibandingkan model tunggal [14]. *Stacking ensemble* (dengan XGBoost sebagai meta-learner) memiliki performa unggul bahkan tanpa SMOTE, hanya dengan stratifikasi dan validasi silang untuk menjaga evaluasi tetap akurat pada data tidak seimbang[15]. Penelitian saat ini menerapkan bangun model *stacking ensemble* menggunakan *XGBoost* sebagai *meta-learner*, yang diharapkan dapat mengintegrasikan keunggulan algoritma dasar dan mengatasi tantangan distribusi data tidak merata.

Kemudian terdapat penelitian pengembangan pendekatan prediktif yang juga mengkhuskan pembahasan pada model ensembel menggunakan XGBoost dan Extra Trees, dilengkapi dengan interpretasi berbasis SHAP untuk memahami kontribusi setiap atribut. Peningkatan akurasi sebesar 20,3% dibandingkan model dasar menunjukkan efektivitas gabungan algoritma dan teknik penyeimbangan data. Penelitian juga menguraikan variabel yang mempengaruhi kinerja siswa adalah domain keluarga, personal, akademis, ekonomi, sosial, dan kelembagaan (family factor, academic factor, economic factor)[5]. Faktor internal seperti usia dan jenis kelamin merupakan yang paling sering digunakan karena mudah didefinisikan dan memiliki ukuran yang jelas untuk penelitian prediksi[4]. Faktor akademis seperti nilai akademis dan skor indeks prestasi komulatif (IPK) menjadi atribut terpenting untuk memprediksi masa depan pendidikan mahasiswa. Faktor Ekonomi mengacu pada kemampuan finansisal orang tua dalam memfasilitasi anak dalam membentuk karir di masa depan. Faktor keluarga yang juga banyak digunakan untuk penelitian adalah menggunakan atribut pendidikan terakhir orang tua dan pekerjaan orangtua. Meski memberikan pemetaan faktor yang luas, belum ada pendekatan prediktif yang secara khusus mengintegrasikan faktor-faktor ini dalam satu model *ensemble* prediksi kelulusan [16]. Meski demikian, fokus pada interpretabilitas belum sepenuhnya diadopsi dalam model prediksi kelulusan, terutama di tahap awal perkuliahan. Hal ini memperkuat relevansi penelitian saat ini dalam mengeksplorasi variabel yang berdampak pada proses pendaftaran untuk memprediksi kelulusan.

Setelah itu penelitian selanjutnya mengevaluasi performa Naive Bayes, KNN, dan RF dalam memprediksi apakah calon mahasiswa akan diterima atau mengundurkan diri. Random Forest menunjukkan akurasi tertinggi, dengan akurasi mencapai 73,61%. Hal ini menunjukkan bahwa RF memberikan kinerja yang baik dalam klasifikasi yang melibatkan data kompleks seperti profil mahasiswa baru. Penelitian ini juga menunjukkan bahwa pengoptimalan hyperparameter dapat meningkatkan akurasi, namun memerlukan waktu komputasi yang lebih lama pada RF [17]. Penelitian serupa, yang menggunakan analisis latar belakang pendidikan menengah, juga menunjukkan hal yang sama dalam efektivitas algoritma RF dan terbukti efisien dalam mengidentifikasi pola data kompleks dalam identifikasi dini terhadap mahasiswa yang berisiko putus sekolah [18]. Penelitian-penelitian tersebut belum dikembangkan ke arah model prediksi komprehensif yang mempertimbangkan aspek non-akademik. Penelitian ini menjembatani kesenjangan tersebut dengan menggabungkan faktor akademik dan sosial ekonomi dalam model *stacking* yang dibangun.

Selanjutnya terdapat penelitian yang menunjukkan bahwa akurasi prediksi dengan seleksi fitur menggunakan Information Gain dalam model RF terbukti lebih unggul dibandingkan algoritma klasifikasi lain seperti Decision Tree dan SVM. Selain itu, validasi silang dengan skema k-fold juga untuk mengonfirmasi konsistensi performa dari model RF. Penelitian ini menemukan variabel seperti

 ISSN:2302-8149
e-ISSN:2540-9719

indeks prestasi, pekerjaan dan pendidikan orang tua, kehadiran di kelas, dan beban ekonomi keluarga sebagai prediktor penting prestasi akademik siswa [19]. Meski demikian, eksplorasi korelasi antar variabel belum dikembangkan lebih lanjut, serta dibutuhkan penambahan variabel baru yang relevan untuk meningkatkan kualitas prediksi.

Penelitian lainnya menekankan keunggulan pendekatan *stacking ensemble* dibandingkan beberapa metode *ensemble* lain seperti *bagging*, *boosting*, dan *voting* dalam memprediksi performa mahasiswa berdasarkan data historis dari sistem manajemen pembelajaran [9], [20]. Fokus pada pendekatan *stacking ensemble learning* melalui integrasi metode SMOTE, seleksi fitur, serta penerapan *meta learner* XGBoost terbukti dapat menangani permasalahan ketidakseimbangan data antara kelas normal dan default yang kerap menurunkan akurasi model klasifikasi. Melalui proses *preprocessing* yang meliputi pembersihan data, penghilangan *missing* values dan *outlier*, serta pengkodean data non-numerik, penelitian sebelumnya sukses menghasilkan data bersih yang kemudian diseimbangkan menggunakan SMOTE. Model *stacking ensemble* memanfaatkan kombinasi dari berbagai *base learner* yang selanjutnya digabungkan melalui meta-learner XGBoost. Hasil evaluasi menunjukkan bahwa model *Stacking* XGBoost mampu meningkatkan performa secara signifikan dibandingkan model klasifikasi tunggal maupun *ensemble* lain tanpa optimasi fitur. Penelitiannya menyimpulkan bahwa kombinasi strategi penyeimbangan data, seleksi fitur terfokus, serta arsitektur *ensemble stacking* yang tepat, dapat meningkatkan ketepatan prediksi [14], [21]. Penelitian ini menjadi dasar pengembangan model saat ini, yang mana penelitian sebelumnya tidak secara eksplisit mengintegrasikan variabel sosial ekonomi dan keluarga. Penelitian ini menanggapi kekurangan tersebut dengan memperluas cakupan variabel, sehingga model prediksi memiliki dimensi yang lebih komprehensif dengan tetap mempertahankan pendekatan *stacking ensemble* sebagai kerangka utama dan pemanfaatan ketangguhan SVM, fleksibilitas KNN, dan stabilitas RF seperti pada penelitian terdahulunya.

Seperti pada pembahasan penelitian sebelumnya, salah satu tantangan dalam pemodelan klasifikasi di bidang pembelajaran mesin adalah ketidakseimbangan kelas (imbalance class), sehingga penggunaan metrik evaluasi seperti akurasi saja menjadi kurang representatif. Beberapa peneliti merekomendasikan penggunaan metrik *Area Under the Receiver Operating Characteristic Curve* (AUC-ROC) karena metrik ini dinilai lebih robust dalam menilai performa model pada data dengan distribusi tidak seimbang [14], [22], [23]. AUC-ROC dapat memberikan gambaran yang lebih adil tentang kemampuan klasifikasi, karena metrik ROC-AUC mengukur performa model berdasarkan perbandingan antara *True Positive Rater* (TPR) dan *False Positive Rate* (FPR), yang dihitung terhadap masing-masing populasi kelas dan tidak dipengaruhi oleh frekuensi absolut kelas minoritas [23], [24].

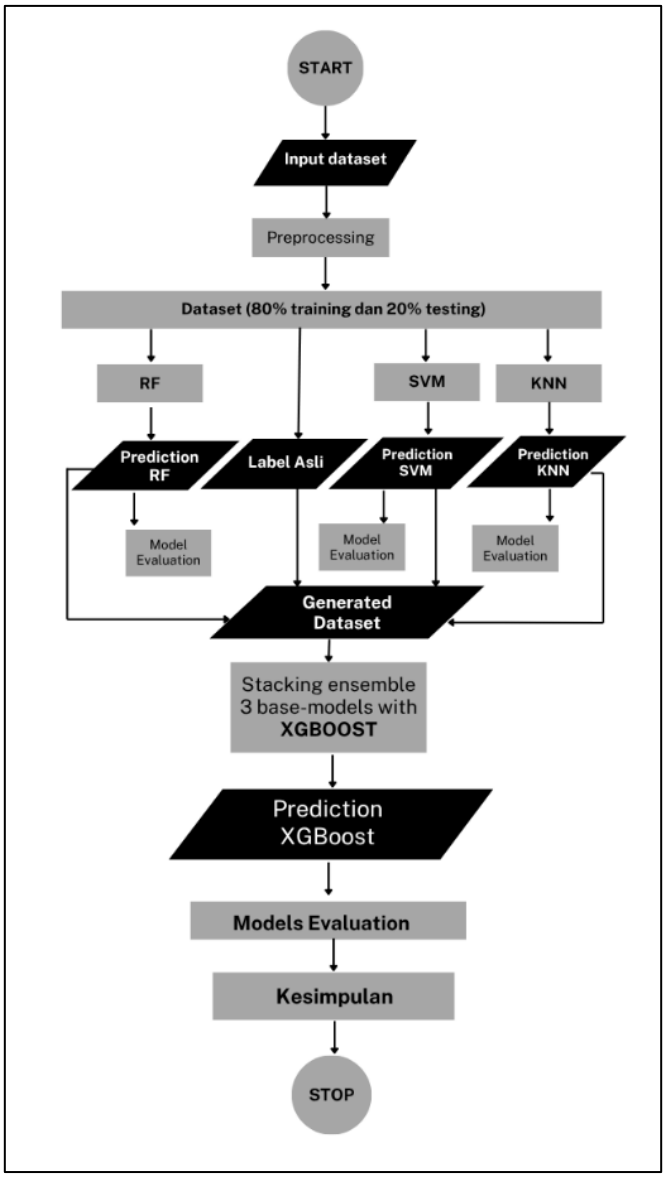
Namun demikian, mayoritas penelitian terdahulu masih terbatas pada prediksi performa berdasarkan data akademik parsial, seperti nilai ujian tengah semester atau IPK semata, dan belum banyak yang memfokuskan pada penggunaan data historis mahasiswa di tahap pendaftaran dan data kelulusan untuk pengembangan model prediksi performa mahasiswa sejak dini. Selain itu, meskipun model *stacking* telah diimplementasikan, kombinasi spesifik antara SVM, KNN, dan RF sebagai model dasar dengan XGBoost sebagai meta-model masih jarang dieksplorasi, khususnya dalam konteks prediksi kelulusan tepat waktu dengan penggunaan data historis mahasiswa baru yang mencakup aspek akademik, sosial, dan ekonomi. Oleh karena itu, penelitian ini berupaya mengisi celah tersebut dengan mengembangkan model prediksi performa calon mahasiswa menggunakan pendekatan *stacking ensemble* secara optimal, mengombinasikan kekuatan masing-masing algoritma dan memperluas ruang lingkup variabel yang digunakan. Keunggulan lain dari penelitian yang dikerjakan saat ini adalah perhatian khusus terhadap distribusi variabel dalam proses pembagian dataset, yang dilakukan secara eksplisit guna menjaga proporsi dan representasi kelas minoritas. Hal ini bertujuan untuk meningkatkan kemampuan generalisasi model sekaligus meminimalkan bias prediksi terhadap kelas mayoritas. Tidak hanya menitikberatkan pada akurasi keseluruhan seperti pendekatan sebelumnya, studi ini juga mengevaluasi performa antar kelas serta ketepatan model dalam mengidentifikasi calon mahasiswa yang berisiko tidak lulus tepat waktu. Dengan demikian, kontribusi utama dari penelitian ini terletak pada integrasi strategis antara teknik *stacking* serta pengelolaan distribusi variabel dalam proses pelatihan dalam konteks prediksi kelulusan berbasis data historis pendaftaran mahasiswa, yang secara strategis penting dalam perencanaan akademik dan efisiensi institusi pendidikan tinggi.

 ISSN:2302-8149
e-ISSN:2540-9719

3 Metode Penelitian

Penelitian ini dikembangkan dalam kerangka kuantitatif eksperimental komputasi dengan pendekatan analisis prediktif. Fokus utamanya diarahkan pada pengujian kinerja model pembelajaran mesin berbasis *stacking ensemble* dalam memprediksi kelulusan calon mahasiswa program sarjana. Kombinasi tiga model dasar yaitu *Random Forest* (RF), *Support Vector Machine* (SVM), dan K-*Nearest Neighbor* (KNN) digunakan sebagai model dasar, sementara model XGBoost difungsikan sebagai *meta-learner* untuk mengintegrasikan hasil prediksi ketiga model sebelumnya [21].

Secara teknis, proses analisis dilakukan menggunakan platform Google *Colaboratory* (Colab) dengan bahasa Python, didukung oleh pustaka *scikit-learn*, *NumPy*, Pandas, dan XGBoost. Perangkat yang digunakan adalah laptop dengan prosesor Intel Core i7 dan RAM 16GB, serta browser Google Chrome pada sistem operasi Windows yang berfungsi untuk mendukung aktivitas komputasi dan *multitasking* selama proses analisis data. Tahapan penelitian secara singkat dapat dijelaskan melalui Gambar 1 dimana alur sistematis mulai dari input dataset, *preprocessing* data (data *cleaning*, konversi kategori ke numerik, normalisasi), selanjutnya pembagian data menjadi data latih (80%) dan data uji (20%), pelatihan model dasar (RF, SVM, KNN) secara paralel, penggabungan *output* ketiga model ke dalam meta-data, pelatihan model XGBoost sebagai *meta-learner*, evaluasi model, dan terakhir penarikan kesimpulan dari model yang digunakan.



 ISSN:2302-8149
e-ISSN:2540-9719

Gambar 1 Tahapan penelitian

Dataset yang digunakan bersumber dari Universitas Slamet Riyadi, mencakup data mahasiswa angkatan 2019, diperoleh dari portal pendaftaran, sistem akademik, dan informasi kelulusan dari NeoFeeder PDDikti periode 2019 - 2024 dengan pengecualian data tidak lengkap dan data berstatus mutasi. Proses validasi instrumen dataset ini dilakukan melalui validasi internal (academic department approval) di Biro Administrasi Akademik dan *cross-check* ke database resmi (administrative database) di Unit Pelaksana Teknis Teknologi Informasi Komunikasi dan Komputer (UPT TIKK) sebagai bentuk validasi isi serta untuk memastikan integritas, dan akurasi data privat, sebagaimana diadopsi dalam penelitian sebelumnya [12], [14]. Selain itu, validasi statistik dilakukan dengan analisis korelasi antar fitur, Gambar 2, dan penghapusan atribut yang tidak relevan, selaras dengan metode yang diterapkan di penelitian sebelum ini [12], [25]. Tiga file yang berisi data pendaftaran, data akademik, dan data kelulusan digabungkan berdasarkan atribut identitas, hingga menghasilkan dataset awal. Dataset awal bersifat eksklusif untuk kepentingan ilmiah dan telah tersedia secara publik melalui repositori berikut:
https://kaggle.com/datasets/ee6d06734c34e2b0583ed657f467b8721b7a6dfca7acdb858ed10f03379aaa59

Proses *preprocessing* yang dilakukan mencakup tahap *cleaning*, validasi statistik melalui analisis korelasi antar fitur, penghapusan atribut redundan, serta penerapan teknik *feature engineering* yang tepat, seperti transformasi variabel [12]. Proses *cleaning* dilakukan dengan membersihkan data duplikat, data dengan nilai kosong, dan status tidak diketahui[12], [14]. Kemudian validasi statistik melalui matriks korelasi dihitung menggunakan *Cramér's* V yang mengukur hubungan antar fitur diskret (nilai 0 - 1) [12]. Setelah melalui tahap *preprocessing*, variabel-variabel yang tersisa didefinisikan secara operasional sebagai berikut: (1) Status kelulusan sebagai variabel dependen yang diklasifikasikan menjadi tiga kelas: 0 = Dikeluarkan, 1 = Lulus Tepat Waktu, dan 2 = Lulus Terlambat; (2) fitur input independen yang terdiri dari data demografis (jenis kelamin, pekerjaan orang tua, penghasilan orang tua, jumlah tanggungan/anak kandung), dan (3) atribut akademik awal (rata-rata nilai rapor SMA), serta data administratif lainnya[15].

Pemodelan dilakukan untuk membangun sistem prediksi status kelulusan calon mahasiswa dengan pendekatan *stacking ensemble*. Strategi ini menggabungkan kekuatan beberapa algoritma dasar dan satu algoritma meta untuk meningkatkan performa klasifikasi. Tiga model dasar (KNN, RF, SVM) dipilih karena masing-masing mewakili pendekatan yang berbeda*: ensemble decision trees* [4], *margin-based learning* [3], dan *instance-based learning* [17]. Masing-masing dilatih secara independen untuk mengenali pola dalam data pelatihan, kemudian hasil prediksi awal dari ketiga model dikombinasikan menjadi meta-fitur. Fitur-fitur ini selanjutnya menjadi input bagi model meta-learner, yaitu XGBoost, yang bertugas mengintegrasikan dan mengoptimalkan prediksi akhir [21]. Pendekatan ini dirancang untuk mengatasi kelemahan masing-masing model dasar dan meningkatkan stabilitas prediksi, terutama pada distribusi kelas yang tidak seimbang [14][15], Selanjutnya, proses prediksi akhir dilakukan pada data uji dengan alur yang sama dengan prediksi awal oleh RF, SVM, dan KNN, dilanjutkan dengan prediksi akhir oleh XGBoost berdasarkan *meta-features* hasil model dasar.

Performa prediksi model dinilai berdasarkan metrik evaluasi klasik seperti akurasi, presisi, *recall*, serta metrik spesifik seperti AUC-ROC. Khususnya AUC-ROC diaggap cocok untuk menganalisa performa model pada distribusi kelas tidak seimbang, sehingga dapat mengetahui prediksi yang lebih stabil [15]. Performa model dievaluasi tidak hanya dengan data uji, tetapi juga melalui 5 *fold cross-validation* untuk menguji sejauh mana model dapat digeneralisasi ke data baru. Pengujian model dilihat berdasarkan skor AUC-ROC dan *confusion matrix* untuk menilai kekuatan klasifikasi yang digunakan. Hasil akhir penelitian ini diharapkan tidak hanya menghasilkan model prediksi kelulusan yang akurat, tetapi juga dapat menjadi rujukan dalam pengembangan sistem pendukung keputusan penerimaan mahasiswa berbasis data.

4 Hasil dan Pembahasan

Penelitian ini bertujuan untuk membangun model prediktif yang dapat mengestimasi kelulusan calon mahasiswa baru dengan memanfaatkan data historis yang telah tersedia sejak

 ISSN:2302-8149
e-ISSN:2540-9719

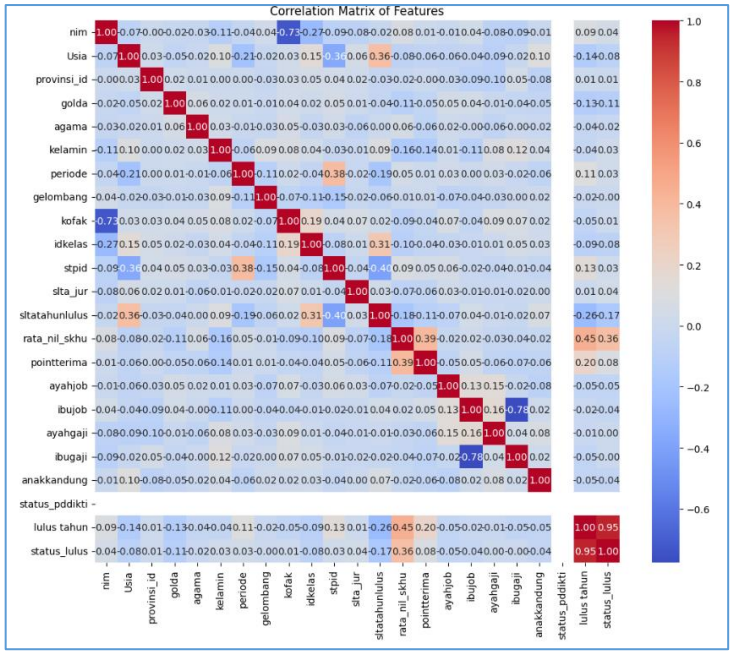
awal pendaftaran mahasiswa [12]. Rangkaian analisis dimulai dari tahap *input* data, dengan dataset awal yang terdiri atas 1.418 entri mahasiswa. Data tersebut disajikan pada Tabel 1 untuk memberikan gambaran umum komposisi data dan fitur-fitur yang direkam dari sistem akademik, mencakup variabel demografis, akademik, administratif, serta status kelulusan mahasiswa. Tabel ini berisi data awal sebelum dilakukan proses seleksi dan transformasi lanjutan.

Tabel 1 Dataset Awal

nim	Usia	…	rata_nil_skhu	pointterima	ayahjob	ibujob	ayahgaji	ibugaji	anakkandung	status_pddikti	lulustahun	status_lulus
1	21	...	74	41.25	6	6	1	1	2		2023	1
2	18	...	84	55.6	6	1	1	0	2	Lulus	2023	1
3	18	...	85	40.75	6	2	2	2	3	Lulus	2023	1
4	25	...	77	36.5	6	1	2	0	2	Lulus	2023	1
5	18	...	80	37.5	6	6	1	1	1	Lulus	2023	1
6	20	...	82	18.75	10	6	2	2	3	Lulus	2023	1
7	40	...	70	31.6	1	2	0	2	5			
8	18	...	89	55.75	6	1	1	0	2	Lulus	2023	1
9	18	...	85	65.55	8	1	1	0	3	Lulus	2023	1
…	…	...	…	…	…	…	…	…	…	…	…	…
1418	22	...	80	37.75	2	2	2	2	2	Lulus	2024	2

Dataset pada file yang diimport pada sistem berawal dari 23 atribut, di mana dua kolom pertama berisi informasi Nomer Induk Mahasiswa (NIM) dan usia mahasiswa, sementara label target berupa status kelulusan terdapat pada kolom terakhir. Setelah melalui tahapan *cleaning*, *preprocessing*, dan analisis menggunakan matriks korelasi, penelitian ini hanya menggunakan 15 atribut sebagai fitur dan 1 atribut sebagai target. Salah satu keputusan awal dalam proses pembersihan data adalah menghapus fitur numerik NIM karena bersifat identitas unik yang tidak mengandung informasi umum yang berguna bagi proses prediksi. Selain NIM juga terdapat fitur "statuspddikti" dan "lulus tahun" dari dataset yang dihapus karena fitur tersebut berkontribusi serta mewakili untuk perhitungan "status_lulus". Pada tahap *feature selection*, fitur-fitur yang kurang relevan dihilangkan berdasarkan hasil analisa menggunakan matriks korelasi [12]. Hasil analisis menggunakan matriks korelasi, Gambar 2, penelitian ini menunjukkan korelasi rendah antara fitur "golda", "sltatahunlulus", "Usia", dan "idkelas" terhadap label "status_lulus", sehingga keempat atribut tersebut turut dihapus. Hasil akhir dari tahapan ini adalah 16 variabel (termasuk target) dengan total 875 baris data yang memenuhi kriteria dari 1.418 entri awal.

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 2 Visualisasi *cramér's v correlation matrix* terhadap status_lulus

Untuk memberikan gambaran yang lebih jelas mengenai karakteristik data yang digunakan, Tabel 2 menyajikan deskripsi dari masing-masing fitur yang telah diseleksi. Sebanyak 15 fitur hasil seleksi berdasarkan matriks korelasi terhadap fitur target dikodekan ke dalam format numerik agar dapat diolah secara optimal oleh algoritma klasifikasi. Transformasi nilai pada setiap fitur ke dalam kelas numerik bertujuan untuk memudahkan model dalam mempelajari pola data, meminimalkan kesalahan prediksi, serta meningkatkan efisiensi dan keandalan model dalam proses klasifikasi [16], [17]. Pada tahap *feature transformation* dan normalisasi yang dilakukan berguna untuk menyesuaikan jenis data dengan kebutuhan algoritma. Pertama penggunaan *frequency encoding* untuk karakteristik data kategorikal dengan data berjenis label, selanjutnya fitur berkarakteristik kategorikal ordinal atau fitur yang terdapat urutan penting antar ketegori menggunakan *OrdinalEncoder*, dan terakhir fitur numerik seperti rata_nil_skhu, pointterima, dan anakkandung dinormalisasi menggunakan *MinMaxScaler* agar relasi antar nilai terjaga dan nilai berada dalam skala yang seragam, sehingga mendukung stabilitas dalam proses pelatihan model klasifikasi.

Tabel 2 Deskripsi fitur

Atribut	Karakteristik	Type Data	Keterangan
provinsi_id	Kategorikal	*Float*	Provinsi domisili
agama	Kategorikal	*Float*	Agama pendaftar
kelamin	Kategorikal	*Float*	Jenis kelamin pendaftar Pria atau Wanita
periode	Kategorikal	*Float*	Periode pendaftaran Gasal atau Genap
gelombang	Kategorikal ordinal	*Float*	Gelombang pendaftaran : dini, I, II
kofak	Kategorikal	*Float*	Kode program studi di Unisri
stpid	Kategorikal	*Float*	Status pendaftaran: Mahasiswa Baru, Pindahan
slta_jur	Kategorikal	*Float*	Jurusan di SLTA : IPA, IPS
rata_nil_skhu	Numerik	*Float*	Rata-rata nilai pada SKHU saat SLTA
pointterima	Numerik	*Float*	Score / nilai ujian CBT
Ayahjob	Kategorikal	*Float*	Pekerjaan ayah
ibujob	Kategorikal	*Float*	Pekerjaan ibu
ayahgaji	Kategorikal ordinal	*Integer*	Penghasilan ayah: rendah, sedang, tinggi

 ISSN:2302-8149
e-ISSN:2540-9719

Atribut	Karakteristik	Type Data	Keterangan
ibugaji	Kategorikal ordinal	*Integer*	Penghasilan ibu: rendah, sedang, tinggi
anakkandung	Numerik	*Float*	Jumlah saudara kandung calon mahasiswa
status_lulus	Kategorikal	*Float*	Status kelulusan: dikeluarkan, lulus tepat waktu, lulus terlambat

Tahapan awal dalam pemodelan dimulai dengan membagi dataset menjadi dua bagian, yaitu data *training* dan data *testing*, yang masing-masing memiliki 15 fitur input. Proporsi distribusi target variabel status_lulus pada dataset utuh, 875 baris data, menunjukkan bahwa mayoritas mahasiswa lulus tepat waktu (69,37%), diikuti oleh mahasiswa yang lulus terlambat (22,29%), dan sisanya dikeluarkan (8,34%). Tabel 3 menunjukkan bahwa proporsi tetap konsisten pada data pelatihan dan pengujian, sehingga memastikan tidak terjadi bias distribusi akibat pembagian dataset. Adapun komposisi jumlah data *training* dan data *testing* disajikan pada Tabel 3 berikut.

Tabel 3 Komposisi pembagian data *training* dan data *testing*

No	Jenis Data	Jumlah	Distribusi kelas target		
			(1)	(2)	(0)
1	Data *Training (80%)*	700 Data	69.43%	22.29%	8.34%
2	Data *Testing* (20%)	175 Data	69.14%	22.29%	8.57%

Setelah membagi data menjadi data latih dan data uji, proses pemodelan dapat dilanjutkan dengan membangun model dasar sebagai fondasi pembelajaran [12], [14], [21]. Tiga algoritma dasar yaitu, RF, KNN, dan SVM dilatih secara terpisah menggunakan data pelatihan, dengan skema validasi silang 5-*fold* dalam upaya menjaga kestabilan evaluasi pada dataset yang memiliki distribusi kelas tidak merata. Hasil evaluasi, Tabel 4, model dasar pada data latih menunjukkan bahwa masing-masing algoritma menghasilkan prediksi awal berdasarkan pola data yang teridentifikasi dari fitur input. Secara keseluruhan, RF menunjukkan performa terbaik dibandingkan KNN dan SVM, dengan akurasi tertinggi sebesar 0,76 dan *weighted* F1-*score* sebesar 0,71. Pada kelas 0 (mahasiswa dikeluarkan), RF mencatat *precision* tertinggi sebesar 0,89, meskipun *recall*-nya masih rendah (0,43), menunjukkan keterbatasan dalam mendeteksi kasus *dropout* secara konsisten. Selanjutnya model SVM unggul dalam *recall* kelas 1 (lulus tepat waktu) dengan nilai 0,98, namun performanya menurun drastis pada kelas 2 (lulus terlambat), yang ditandai dengan F1-*score* sebesar 0,00. Hal serupa terjadi pada model KNN yang memiliki *recall* rendah pada kelas 0 (0,31) dan kelas 2 (0,15), menunjukkan sensitivitas yang terbatas terhadap kategori minoritas. Hasil klasifikasi model dasar terhadap data latih ditunjukkan pada Tabel 4 berikut.

Tabel 4 Laporan klasifikasi model dasar

Kelas	Metrik	RF	KNN	SVM
Class 0	Precision	0.89	0.78	0.62
	Recall	0.43	0.31	0.50
	F1-Score	0.58	0.44	0.55
	Support	58	58	58
Class 1	Precision	0.77	0.75	0.73
	Recall	0.97	0.95	0.98
	F1-Score	0.86	0.84	0.84
	Support	486	486	486
Class 2	Precision	0.54	0.40	0.00

 ISSN:2302-8149
e-ISSN:2540-9719

Kelas	Metrik	RF	KNN	SVM
	Recall	0.22	0.15	0.00
	F1-Score	0.31	0.22	0.00
	Support	156	156	156
Akurasi		**0.76**	0.72	0.72
Avg	Weighted Precision	0.73	0.67	0.56
	Weighted Recall	0.76	0.72	0.72
	Weighted F1-Score	**0.71**	0.67	0.63

Selanjutnya untuk mengatasi keterbatasan model tunggal sebelumnya, perlu membentuk *meta features* dari *output* (hasil prediksi) model dasar yang kemudian dikombinasikan dengan fitur awal sehingga membentuk 18 fitur total. Meta data tersebut selanjutnya digunakan untuk melatih model *stacking* pada tingkat meta model [12]. Dimana untuk mengatasi keterbatasan model tunggal, pendekatan dengan skema *stacking ensemble* diterapkan menggunakan XGBoost sebagai *meta learner [14], [15], [16], [21]*. Hasil evaluasi meta model terhadap meta *feature*, Tabel 5, menunjukkan semua kelas memiliki performa yang lebih merata, dengan F1-*score* mencapai 0,84 untuk kelas 0, 0,92 untuk kelas 1, dan 0,67 untuk kelas 2. Model XGBoost sebagai meta model mencapai akurasi keseluruhan 0,86, dengan *weighted* F1-*score* 0,85, yang menunjukkan peningkatan performa signifikan dibandingkan model tunggal.

Tabel 5 Laporan evaluasi meta model

Kelas	Metrik	XGBoost (Meta-Model)
Class 0	Precision	0.93
	Recall	0.74
	F1-Score	0.83
	Support	58
Class 1	Precision	0.86
	Recall	0.98
	F1-Score	0.92
	Support	486
Class 2	Precision	0.86
	Recall	0.55
	F1-Score	0.67
	Support	156
Akurasi		0.86
Avg	Weighted Precision	0.87
	Weighted Recall	0.86
	Weighted F1-Score	0.85

Untuk lebih jelasnya tersedia tabel rekap *classification report* (Tabel 6) dari proses pelatihan model dasar serta meta model terhadap data latih. Hasil evaluasi terhadap data latih menunjukkan *Random Forest* terbukti sebagai *base learner* yang memberikan kontribusi paling kuat dan konsisten memiliki keunggulan pada masalah klasifikasi data akademik di antara model dasar lain [12], [13], [21]. Hal tersebut dibuktikan berdasarkan evaluasi model RF dibandingkan dengan penelitian terdahulu [12], dengan menggunakan jenis atribut yang

 ISSN:2302-8149
e-ISSN:2540-9719

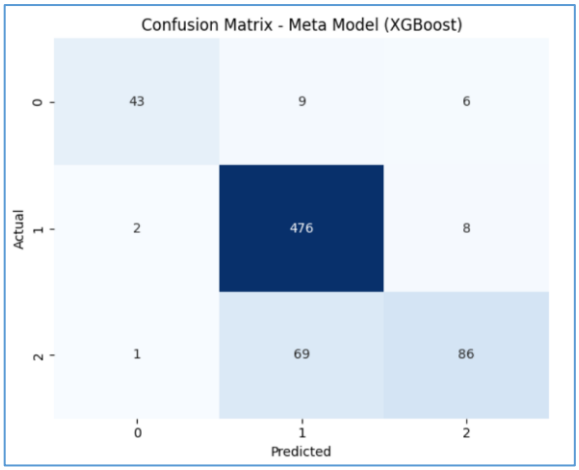
sama, penelitian saat ini menghasilkan akurasi 0,76 dan *weighted avg* F1-*score* 0,71, menunjukkan kemampuan yang lebih stabil dalam menangkap pola pada dataset dengan menambahkan teknik 5-*fold cross validation* dan *feature selection* dalam mempersiapkan dataset penelitian. Sementara itu, SVM tercatat mengungguli *score* evaluasi dari penelitian sebelumnya [12], dengan akurasi 0,72 mengindikasikan performa yang relatif mencapai *score* tidak buruk namun tetap penting sebagai *base learner* untuk mendeteksi pola *non-linear*, meskipun rentan pada distribusi kelas tidak seimbang. Selain itu terlihat bahwa meta model XGBoost menunjukkan performa terbaik dibandingkan model dasar lainnya pada seluruh metrik evaluasi, termasuk akurasi, *precision*, *recall*, dan F1-*score*. Peningkatan performa ini memperkuat temuan pada penelitian sebelumnya yang menegaskan bahwa pendekatan *ensemble* dengan meta *learner* yang kuat mampu mengoptimalkan kelemahan masing-masing base model, serta mendukung pengendalian ketidakseimbangan antar kelas [12]. Peningkatan ini terutama terlihat pada metrik *weighted average F1-score* yang mencapai 0,85, dibandingkan dengan skor tertinggi dari model dasar yang hanya sebesar 0,71 pada *Random Forest*. Selain itu hasil evaluasi juga menunjukkan peningkatan akurasi yang signifikan menjadi 0,86 dibandingkan dengan model dasar sebelumnya dan penelitian terdahulunya yang hanya mencapai akurasi 0.67 pada tahap ensembel dengan proses awal data pendaftaran [12].

Tabel 6 Rekap *classification report* pelatihan *base model* dan meta model

Metrics	Algorithms			
	Base Model			Meta Model
	KNN	RF	SVM	XGBoost
Accuracy	0,72	0,76	0,72	**0,86**
Wighted avg Precision	0,67	0,73	0,56	0,87
Wighted avg Recall	0,72	0,76	0,72	0,86
Wighted avg F1-Score	0.67	0.71	0.63	**0.85**

Dengan demikian, kombinasi model melalui teknik *stacking ensemble* tidak hanya meningkatkan akurasi keseluruhan, tetapi juga membantu mengatasi ketidakseimbangan antar kelas, sebagaimana juga diperkuat oleh analisis *confusion matrix* pada Gambar 3. Performa meta model yang ditampilkan dalam *confusion matrix[12], [16]*, Gambar 3, menunjukkan penurunan kesalahan klasifikasi, terutama dalam mengidentifikasi mahasiswa yang berisiko lulus tidak tepat waktu atau terlambat. Hal ini membuktikan bahwa kombinasi model memberikan kontribusi positif dalam menangani variasi kompleks antar kelas. Meski demikian, terlihat bahwa sebagian besar kesalahan prediksi terjadi pada kelas 2 (lulus terlambat), di mana model masih cukup sering menyamakannya dengan kelas 1. Hal ini menjadi catatan bahwa meskipun *stacking* memberikan perbaikan, perbedaan antar kelas dalam segi fitur input kemungkinan masih belum cukup kuat untuk memberikan pemisahan yang tegas di ruang vektor prediktif.

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 3 *Confusion matrix* meta model

Setelah seluruh model dasar dilatih dan dikombinasikan melalui pendekatan *stacking ensemble*, langkah berikutnya adalah menguji kinerja model. Untuk memperoleh gambaran yang objektif terhadap kemampuan model dalam menghadapi data baru, diperlukan pengujian performa menggunakan data uji yang sepenuhnya terpisah dari proses pelatihan [12], [20]. Pengujian ini menjadi bagian penting dalam memastikan bahwa model tidak hanya bekerja baik pada data yang telah dipelajari, tetapi juga mampu mengenali pola yang belum pernah ditemui sebelumnya [16]. Evaluasi terhadap data uji bertujuan untuk mengukur sejauh mana model yang dibangun mampu mempertahankan performa stabilitas model prediksi di luar data latih, serta mendeteksi potensi *overfitting* yang tersembunyi [14]. Oleh karena itu, analisis performa terhadap data uji memberikan dasar yang lebih kuat dalam menilai efektivitas model secara praktis.

Hasil pengujian pada data uji ditampilkan dalam Tabel 7. Model *stacking ensemble* menunjukkan akurasi sebesar 0,8229 dan *weighted* F1-*score* sebesar 0,80. Kelas “lulus tepat waktu” (class 1) mencatat performa tertinggi dengan *recall* 0,98 dan F1-*score* 0,90. Sementara itu, kelas “dikeluarkan” (class 0) menunjukkan stabilitas moderat dengan *precision* 0,64, *recall* 0,60, dan F1-*score* 0,62. Performa terhadap kelas “lulus terlambat” (class 2) masih menjadi tantangan, dengan F1-*score* sebesar 0,58, meskipun menunjukkan peningkatan dibandingkan model dasar.

Capaian ini menunjukkan keunggulan dibandingkan hasil model dalam penelitian sebelumnya untuk klasifikasi status lulus [12], [14] yang sama-sama menggunakan data formulir pendaftaran mahasiswa baru. Hasil akurasi pada penelitian saat ini melampaui performa model pada salah satu penelitian terdahulu yang tercatat mencapai akurasi maksimal 68.55% pada penerapan model *cascade ensemble* menggunakan dataset pendaftaran [12]. Dalam penelitian tersebut tidak berhenti pada *ensemble* awal dengan data pendaftaran saja. Pendekatan *cascade ensemble* terus dilanjutkan dengan menyertakan tambahan fitur capaian akademik setiap semesternya hingga menghasilkan peningkatan performa model menjadi 88.82% dengan menggunakan tambahan data akademik sampai semester 4 [12]. Namun, penelitian ini secara konsisten membatasi ruang lingkup pada pemodelan berbasis data pendaftaran awal, sehingga relevan sebagai pendekatan prediksi dini sebelum data akademik tersedia. Meskipun tidak menggunakan data berjenjang seperti model *ensemble* sebelumnya, pendekatan yang digunakan penelitian saat ini mampu mencapai akurasi cukup tinggi 82%, dengan nilai *weighted* F1-*score* 80% yang menunjukkan performa model cukup konsisten meskipun dihadapkan pada distribusi kelas yang tidak seimbang. Recall 98% pada kelas 'lulus tepat waktu' menunjukkan keunggulan model dalam mengenali kategori mayoritas, yang secara umum melampaui capaian model pada penelitian sebelumnya dalam memprediksi status kelulusan berbasis data pendaftaran awal [12].

 ISSN:2302-8149
e-ISSN:2540-9719

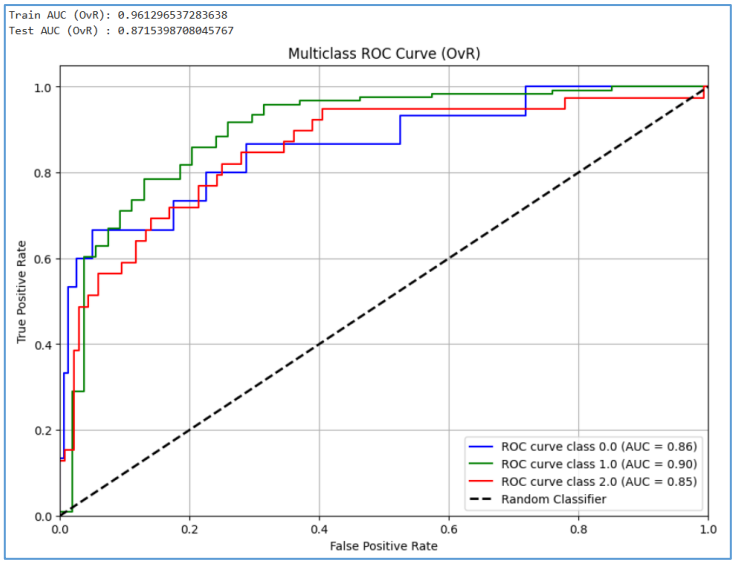
Selain itu, Tabel 7 menampilkan selisih akurasi antara data latih (0,86) dan data uji (0,82) tidak lebih dari 5%, yang mengindikasikan bahwa model tidak mengalami indikasi *overfitting* yang serius. Selisih yang relatif kecil ini menjadi indikator bahwa model memiliki kemampuan generalisasi yang baik terhadap data baru. Dengan demikian, pendekatan *stacking ensemble* yang dikembangkan menunjukkan prediksi kelulusan secara dini sejak tahap awal pendaftaran tetap memungkinkan dilakukan secara andal, bahkan sebelum memperoleh data akademik lanjutan.

Tabel 7 Perbandingan *classification report* antara *train set* dan *test set*

Kelas	Metrik	*Train* Set	*Test* Set
Class 0	Precision	0.93	0.64
	Recall	0.74	0.60
	F1-Score	0.83	0.62
	Support	58	15
Class 1	Precision	0.86	0.84
	Recall	0.98	0.98
	F1-Score	0.92	0.90
	Support	486	121
Class 2	Precision	0.86	0.85
	Recall	0.55	0.44
	F1-Score	0.67	0.58
	Support	156	39
Akurasi		0.86	0.82
Avg	Weighted Precision	0.87	0.82
	Weighted Recall	0.86	0.82
	Weighted F1-Score	0.85	0.80

Selain metrik evaluasi akurasi, presisi, dan *recall*, dalam mengukur performa prediksi model, penelitian ini juga menggunakan metrik AUC-ROC yang cocok untuk mengevaluasi performa model dengan distribusi kelas tidak seimbang [14], [15]. Adapun skor AUC dan visualisasi ROC disajikan pada Gambar 4. Kurva ROC mendukung penelitian dengan AUC sebesar 0,9613 pada data latih dan 0,8715 pada data uji. Masing-masing kelas mencatat skor AUC yang tinggi pada kelas data uji: 0,86 (kelas 0), 0,90 (kelas 1), dan 0,85 (kelas 2). Hal ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang kuat dan terdistribusi seimbang di seluruh kelas target, bahkan pada kondisi data minoritas. Setelah memastikan performa klasifikasi model melalui nilai AUC yang tinggi dan konsisten di seluruh kelas, langkah selanjutnya adalah menganalisa kontribusi masing-masing fitur terhadap prediksi yang dihasilkan.

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 4 Test AUC dan kurva ROC

Analisis interpretabilitas model *stacking ensemble* dilakukan dengan memanfaatkan teknik SHAP (SHapley Additive exPlanations) [16], [26]. SHAP digunakan untuk mengungkap dan memvisualisasikan logika prediksi dari model *ensemble* yang diterapkan. SHAP bekerja dengan menghitung kontribusi marjinal rata-rata dari setiap fitur dalam semua kemungkinan kombinasi fitur yang mungkin, sehingga hasilnya bersifat adil dan dapat dijelaskan secara matematis [26]. Hal ini memungkinkan model yang kompleks seperti XGBoost untuk dapat dijelaskan dengan cara yang intuitif dan visual. SHAP mendukung interpretasi secara global (misalnya, fitur mana yang paling mempengaruhi kelulusan secara keseluruhan) maupun secara lokal (misalnya, mengapa seorang mahasiswa tertentu diprediksi tidak lulus) [16]. Pendekatan ini pertama kali diperkenalkan oleh Lloyd Shapley pada tahun 1953 dalam konteks teori permainan kooperatif, dan oleh karena itu nilai kontribusi yang dihitung dengan metode ini dikenal sebagai Shapley values.

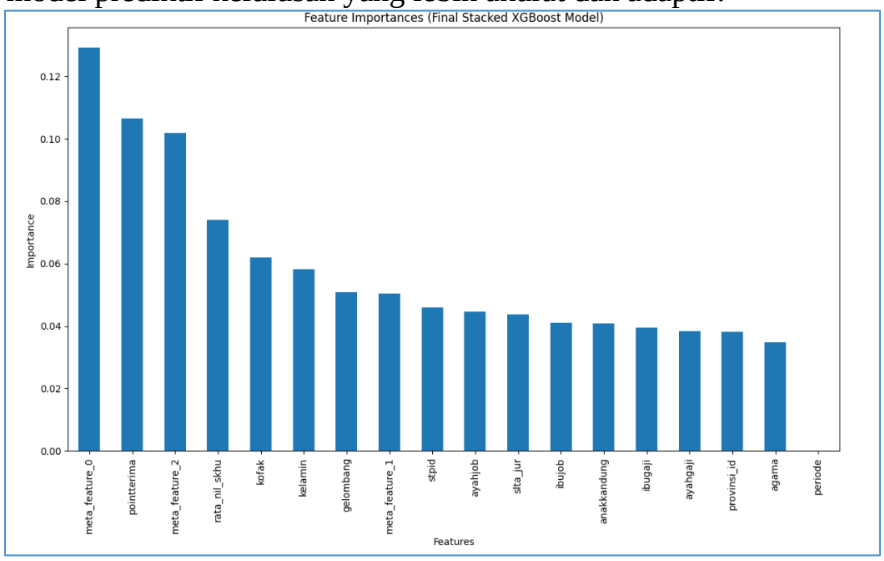
Dalam konteks *machine learning*, metode ini digunakan untuk mengukur seberapa besar kontribusi setiap fitur *input* terhadap *output* prediksi dari model kompleks (seperti ensemble atau neural networks). Misalnya, diberikan suatu fungsi prediksi p dari model, nilai Shapley untuk fitur tertentu i (dari total n fitur) dihitung dengan cara mengambil rata-rata kontribusi marjinal fitur i terhadap hasil prediksi, dihitung dari semua kemungkinan subset fitur lain S (yang tidak mengandung i). Secara matematis, teknik SHAP dirumuskan sebagaimana persamaan 1 berikut ini [16]:

$$\phi_i(p) = \sum_{S \subseteq \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} [p(S \cup \{i\}) - p(S)] \qquad (1)$$

Hasil analisis teknik visualisasi, menggunakan nilai SHAP untuk memberikan tampilan yang jelas terhadap fitur-fitur yang memengaruhi performa calon mahasiswa, disajikan pada Gambar 5. Analisis interpretabilitas model *stacking ensemble* dengan memanfaatkan teknik SHAP (SHapley Additive exPlanations), menunjukkan bahwa fitur-fitur akademik seperti skor seleksi masuk dan rata-rata nilai SKHU memiliki kontribusi paling besar dalam menentukan status kelulusan mahasiswa. Di samping itu, sejumlah atribut demografis seperti jenis kelamin, pekerjaan orang tua, dan jumlah anak kandung juga turut memberikan pengaruh yang cukup signifikan, mengindikasikan adanya keterkaitan antara latar belakang sosial dengan keberhasilan studi. Sementara itu, beberapa atribut administratif seperti periode pendaftaran, agama, dan asal provinsi tercatat memiliki pengaruh yang sangat rendah terhadap hasil prediksi. Temuan ini menguatkan bahwa kombinasi antara kemampuan

 ISSN:2302-8149
e-ISSN:2540-9719

akademik awal dan kondisi sosial ekonomi mahasiswa menjadi faktor utama dalam membangun model prediktif kelulusan yang lebih akurat dan adaptif.



Gambar 5 Feature importance pada model final *stacking* XGBoost

Berdasarkan hasil penelitian, menunjukkan bahwa pendekatan *stacking ensemble* yang mengombinasikan RF, SVM, dan KNN dengan XGBoost sebagai *meta-learner* mampu memberikan peningkatan akurasi dan stabilitas prediksi, terutama untuk kelas mayoritas. Keunikan dari pendekatan ini dibandingkan penelitian terdahulu terletak pada penggunaan data historis pendaftaran dalam memprediksi kelulusan serta implementasi *stacking* dalam konteks prediksi multikelas di lingkungan perguruan tinggi. Penelitian ini juga membuktikan bahwa penggabungan kekuatan model klasik dengan metode ensembel mampu menghasilkan prediksi yang lebih robust dan aplikatif untuk kebutuhan akademik berbasis data. Namun demikian, terdapat beberapa batasan yang perlu dicermati untuk pengembangan ke depan.

Batasan pertama adalah model yang dikembangkan dalam penelitian ini dibangun berdasarkan data pendaftaran dari satu institusi, sehingga potensi generalisasinya terhadap perguruan tinggi lain masih perlu diuji lebih lanjut. Meskipun fitur-fitur seperti latar belakang demografis dan nilai akademik awal umumnya digunakan dalam proses seleksi mahasiswa baru di berbagai kampus, perbedaan dalam karakteristik institusional, sistem pelaporan, serta skema penilaian antar perguruan tinggi dapat memengaruhi performa model secara signifikan. Selanjutnya, penting untuk dipahami bahwa model ini dirancang untuk memprediksi status kelulusan sejak tahap awal pendaftaran, sebelum mahasiswa menjalani proses akademik di perguruan tinggi. Oleh karena itu, cakupan prediksi terbatas pada fitur-fitur statis yang tersedia saat seleksi masuk, dan belum mencakup dinamika yang mungkin terjadi setelah mahasiswa diterima, seperti perubahan motivasi belajar, adaptasi terhadap lingkungan akademik, atau penyesuaian terhadap perubahan kurikulum. Penelitian selanjutnya dapat diarahkan pada pengembangan model yang lebih adaptif dan dapat diperbarui secara berkala, guna merespons perubahan data atau kebijakan institusi. Selain itu, pengujian model secara lintas institusi perlu dilakukan untuk menguji robusta dan stabilitas performa model dalam skala yang lebih luas.

5 Kesimpulan

Penelitian ini berhasil mengembangkan model prediksi kelulusan calon mahasiswa baru menggunakan pendekatan *stacking ensemble* yang menggabungkan *Random Forest*, SVM, dan KNN sebagai model dasar dengan XGBoost sebagai *meta-learner*. Hasil evaluasi menunjukkan kinerja yang lebih akurat dan stabil dibandingkan model tunggal, dengan akurasi 0,82, *weighted F1-score* 0,80, dan AUC 0,87 pada data uji, serta peningkatan signifikan dalam mengenali kelas minoritas

 ISSN:2302-8149
e-ISSN:2540-9719

seperti mahasiswa yang dikeluarkan dan lulus terlambat. Pendekatan *ensemble* terbukti mampu mengurangi ketimpangan klasifikasi akibat distribusi data tidak merata sekaligus memperkuat kemampuan generalisasi model. Analisis fitur menunjukkan bahwa skor akademik, seperti nilai seleksi masuk dan rata-rata nilai SKHU, menjadi faktor dominan, sedangkan atribut demografis berperan sebagai pendukung, dan variabel administratif memiliki kontribusi terbatas. Integrasi informasi akademik dan sosial ekonomi, disertai penerapan stratifikasi dalam validasi silang serta pemilihan parameter dalam meta model, menjadi kunci untuk mencegah *overfitting* pada kelas mayoritas dalam klasifikasi multikelas, sehingga meningkatkan kualitas dan keandalan prediksi menggunakan *stacking ensemble*. Untuk pengembangan selanjutnya, validasi lintas institusi dan penyesuaian terhadap perubahan sistem seleksi menjadi langkah penting untuk memperluas generalisasi model, sekaligus meningkatkan efektivitasnya sebagai alat pendukung keputusan akademik berbasis data.

Referensi

[1] A. F. Mohamed Nafuri, N. S. Sani, N. F. A. Zainudin, A. H. A. Rahman, and M. Aliff, "*Clustering Analysis for Classifying Student Academic Performance in Higher Education*," *Applied Sciences (Switzerland)*, Vol. 12, No. 19, Oct. 2022, DOI: 10.3390/app12199467.

[2] T. R. Noviandy *et al.*, "*Machine Learning for Early Detection of Dropout Risks and Academic Excellence: A Stacked Classifier Approach*," *Journal of Educational Management and Learning*, Vol. 2, No. 1, pp. 28–34, 2024, DOI: 10.60084/jeml.v2i1.191.

[3] H. Karalar and C. Kapucu, "Akses terbuka memprediksi Siswa Berisiko Gagal Akademik menggunakan Model Ansambel pada Masa Pandemi dalam Sistem Pembelajaran Jarak Jauh," 2021.

[4] Z. Sun, Y. Yuan, X. Xiong, S. Meng, Y. Shi, and A. Chen, "*Predicting Academic Achievement from the Collaborative Influences of Executive Function, Physical Fitness, and Demographic Factors among Primary School Students in China: Ensemble Learning Methods*," *BMC Public Health*, Vol. 24, No. 1, pp. 1–13, 2024, DOI: 10.1186/s12889-024-17769-7.

[5] F. Ouatik and M. Eritali, "*Machine Translated by Google memprediksi Keberhasilan Siswa menggunakan Big Data dan Mesin Algoritma Pembelajaran Machine Translated by Google*," pp. 236–251.

[6] M. Yaÿcÿ, "Penambangan Data Pendidikan : Prediksi Kinerja Akademik Siswa menggunakan Algoritma Pembelajaran Mesin," 2022.

[7] H. Karalar, C. Kapucu, and H. Gürüler, "*Predicting Students at Risk of Academic Failure using Ensemble Model during Pandemic in a Distance Learning System*," *International Journal of Educational Technology in Higher Education*, Vol. 18, No. 1, 2021, DOI: 10.1186/s41239-021-00300-y.

[8] R. Shintabella, C. Edi Widodo, and A. Wibowo, "*Loss of Life Transformer Prediction based on Stacking Ensemble Improved by Genetic Algorithm By IJISRT*," *International Journal of Innovative Science and Research Technology (IJISRT)*, Vol. 9, No. 3, pp. 1061–1066, 2024, DOI: 10.38124/ijisrt/ijisrt24mar1125.

[9] M. R. Alzahrani, "*Predicting Student Performance using Ensemble Models and Learning Analytics Techniques*," *Preprints.org*, p. 202406.1100.v1, 2024, DOI: 10.20944/preprints202406.1100.v1.

[10] K. Mahboob, Sarfaraz Abdul Sattar Natha, Syed Saood Zia, Priha Bhatti, Abeer Javed Syed, and Samra Mehmood, "*An Ensemble Modeling Approach to Enhance Grade Prediction in Academic Engineering Programming Courses*," *VFAST Transactions on Software Engineering*, Vol. 11, No. 4, pp. 01–14, 2023, DOI: 10.21015/vtse.v11i4.1641.

[11] L. Yan and Y. Liu, "*An Ensemble Prediction Model for Potential Student Recommendation using Machine Learning*," *Symmetry (Basel)*, Vol. 12, No. 5, pp. 1–17, 2020, DOI: 10.3390/SYM12050728.

[12] A. J. Fernandez-Garcia, J. C. Preciado, F. Melchor, R. Rodriguez-Echeverria, J. M. Conejero, and F. Sanchez-Figueroa, "*A Real-Life Machine Learning Experience for Predicting University Dropout at Different Stages using Academic Data*," *IEEE Access*, Vol. 9, pp. 133076–133090, 2021, DOI: 10.1109/ACCESS.2021.3115851.

 ISSN:2302-8149
e-ISSN:2540-9719

[13] S. D. A. Bujang *et al.*, "*Multiclass Prediction Model for Student Grade Prediction using Machine Learning*," *IEEE Access*, Vol. 9, pp. 95608–95621, 2021, DOI: 10.1109/ACCESS.2021.3093563.

[14] Herianto, B. Kurniawan, Z. H. Hartomi, Y. Irawan, and M. K. Anam, "*Machine Learning Algorithm Optimization using Stacking Technique for Graduation Prediction*," *Journal of Applied Data Sciences*, Vol. 5, No. 3, pp. 1272–1285, 2024, DOI: 10.47738/jads.v5i3.316.

[15] A. Ghasemieh, A. Lloyed, P. Bahrami, P. Vajar, and R. Kashef, "*A Novel Machine Learning Model with Stacking Ensemble Learner for Predicting Emergency Readmission of Heart-Disease Patients*," *Decision Analytics Journal*, Vol. 7, No. February, p. 100242, 2023, DOI: 10.1016/j.dajour.2023.100242.

[16] H. Sahlaoui, E. A. A. Alaoui, A. Nayyar, S. Agoujil, and M. M. Jaber, "*Predicting and Interpreting Student Performance using Ensemble Models and Shapley Additive Explanations*," *IEEE Access*, Vol. 9, pp. 152688–152703, 2021, DOI: 10.1109/ACCESS.2021.3124270.

[17] P. Sejati, Munawar, M. Pilliang, and H. Akbar, "Studi Komparasi *Naive Bayes* , *K-Nearest Neighbor*, dan *Random Forest* untuk Prediksi Calon Mahasiswa yang Diterima atau *Comparative Study of Naive Bayes* , *K-Nearest Neighbor* , *and Random Forest for the Prediction of Prospective Students*," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, Vol. 9, No. 7, pp. 1341–1348, 2022, DOI: 10.25126/jtiik.202296737.

[18] M. Nachouki and M. A. Naaj, "*Predicting Student Performance to Improve Academic Advising using the Random Forest Algorithm*," *International Journal of Distance Education Technologies*, Vol. 20, No. 1, pp. 1–17, 2022, DOI: 10.4018/IJDET.296702.

[19] I. Vol and M. Gusnina, "Machine Translated by Google Prediksi Kinerja Mahasiswa Universitas Sebelas Maret Berdasarkan Random Forest Algoritma Machine Translated by Google," Vol. 27, No. 3, pp. 495–501, 2022.

[20] N. A. Butt, Z. Mahmood, K. Shakeel, S. Alfarhood, M. Safran, and I. Ashraf, "*Performance Prediction of Students in Higher Education using Multi-Model Ensemble Approach*," *IEEE Access*, Vol. 11, pp. 136091–136108, 2023, DOI: 10.1109/ACCESS.2023.3336987.

[21] M. A. Muslim *et al.*, "*New Model Combination Meta-Learner to Improve Accuracy Prediction P2P Lending with Stacking Ensemble Learning*," *Intelligent Systems with Applications*, Vol. 18, No. February, p. 200204, 2023, DOI: 10.1016/j.iswa.2023.200204.

[22] F. Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, "*Learning from Imbalanced Data Sets*," *Springer*, 2018, DOI: https://doi.org/10.1007/978-3-319-98074-4.

[23] G. Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, "*Learning from Class-Imbalanced Data: Review of Methods and Applications. Expert Systems with Applications*," Vol. 73, pp. 220–239, 2017, DOI: https://doi.org/10.1016/j.eswa.2016.12.035.

[24] E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, "*Pr ep rin t n pe er r ed Pr ep rin t n er ed*".

[25] S. S. Yadav* and G. P. Bhole, "*Learning from Imbalanced Data in Classification*," *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 8, No. 5, pp. 1907–1016, 2020, DOI: 10.35940/ijrte.e6286.018520.

[26] S. M. Lundberg and S. I. Lee, "*A Unified Approach to Interpreting Model Predictions*," *Adv Neural Inf Process Syst*, Vol. 2017-Decem, No. Section 2, pp. 4766–4775, 2017.