

## PENERAPAN ALGORITMA RANDOM FOREST SEBAGAI EARLY WARNING SYSTEM PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU

### IMPLEMENTATION OF THE RANDOM FOREST ALGORITHM AS AN EARLY WARNING SYSTEM TO PREDICTION STUDENT GRADUATION ON TIME

Azzumardi Mubarak Romadhoni<sup>1</sup>, Widyasari Puspa Permata Witra<sup>2</sup>, Indra Gita Anugrah<sup>3</sup>  
Program Studi Sistem Informasi, Fakultas Teknik, Universitas Muhammadiyah Gresik, Indonesia<sup>1,2,3</sup>  
[azzumardimubarak@umg.ac.id](mailto:azzumardimubarak@umg.ac.id)<sup>1</sup>, [widyasarippw@umg.ac.id](mailto:widyasarippw@umg.ac.id)<sup>2</sup>, [indragitaanuugrah@umg.ac.id](mailto:indragitaanuugrah@umg.ac.id)<sup>3</sup>

#### ABSTRACT

*On-time graduation is a key indicator of higher education performance and reflects the effectiveness of academic processes. The low rate of on-time graduation in the Faculty of Engineering at a university in Gresik has encouraged the development of an early prediction mechanism to identify students at risk of delayed graduation. This study develops an Early Warning System (EWS) model using the Random Forest algorithm based on academic data from 1,558 students enrolled between 2018 and 2021. The variables used include semester grade point averages (GPA) from semesters 1–4, cumulative GPA at semester 4, and employment status. Model evaluation was conducted using combinations of 5-fold and 10-fold cross-validation with train-test split ratios of 70:30, 80:20, and 90:10. The best-performing model was obtained from the cv5\_split70 scenario with  $n\_estimators = 50$  and  $max\_depth = 5$ , achieving an accuracy of 77.56% and an F1-score of 0.7764. The fourth-semester GPA was identified as the most influential predictor. The results demonstrate that the Random Forest algorithm is effective as a foundation for developing an EWS to support more targeted academic interventions.*

**Keywords:** *Early warning System, Random Forest, Graduation Prediction, Higher education performance*

#### ABSTRAK

Ketepatan waktu kelulusan merupakan indikator penting dalam evaluasi kinerja perguruan tinggi. Rendahnya tingkat kelulusan tepat waktu pada Fakultas Teknik di salah satu universitas di Gresik mendorong pengembangan mekanisme prediksi dini untuk mengidentifikasi mahasiswa berisiko terlambat lulus. Penelitian ini membangun model Early Warning System (EWS) menggunakan algoritma Random Forest berdasarkan data akademik 1.558 mahasiswa angkatan 2018 - 2021. Variabel yang digunakan meliputi IPS semester 1 - 4, IPK semester 4, dan status pekerjaan. Pengujian dilakukan dengan kombinasi K-Fold Cross Validation (5 dan 10) serta pembagian data 70:30, 80:20, dan 90:10. Model terbaik diperoleh pada skenario cv5\_split70 dengan parameter  $n\_estimators=50$  dan  $max\_depth=5$ , menghasilkan akurasi 77,56% dan F1-Score 0,7764. IPS semester 4 menjadi prediktor paling dominan. Hasil penelitian menunjukkan Random Forest efektif digunakan sebagai dasar pengembangan EWS untuk mendukung intervensi akademik yang lebih tepat sasaran.

**Kata Kunci:** Early Warning System, Random Forest, Prediksi Kelulusan, Kinerja pendidikan tinggi

#### PENDAHULUAN

Perguruan tinggi memiliki tanggung jawab strategis untuk menghasilkan lulusan yang berkualitas. Kelulusan tepat waktu menjadi salah satu indikator utama dalam menilai keberhasilan perguruan tinggi karena mencerminkan efektivitas proses akademik yang berjalan secara optimal dan terstruktur. Selain berdampak pada citra institusi, tingginya persentase kelulusan tepat waktu juga berkontribusi terhadap penilaian akreditasi perguruan tinggi (Ayyub & Iddrus, 2024). Dalam standar akreditasi IAPS 4.0, Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT)

menjadikan tingkat kelulusan tepat waktu sebagai salah satu indikator penting dalam penilaian mutu program studi dan institusi perguruan tinggi (Hasibuan & Mahdiana, 2023). Oleh karena itu, rendahnya tingkat kelulusan tepat waktu dapat memengaruhi hasil evaluasi akreditasi suatu perguruan tinggi. Lebih jauh, tantangan akademik yang tidak tertangani sejak dini tidak hanya berpotensi menunda kelulusan, tetapi juga dapat berujung pada meningkatnya risiko mahasiswa mengalami putus studi (drop out).

Permasalahan putus studi (drop out) dan keterlambatan kelulusan masih menjadi

tantangan dalam dunia pendidikan tinggi. Isu mahasiswa drop out telah mendapat perhatian luas karena berkaitan dengan keberhasilan perguruan tinggi dalam menjamin kelulusan mahasiswa dan menjaga kualitas akademik (Wan Yaacob et al., 2020). Berdasarkan data Statistik Pendidikan Tinggi tahun 2020, jumlah mahasiswa yang mengalami drop out pada tahun 2019 mencapai 602.208 mahasiswa dari total mahasiswa terdaftar di Indonesia (Handini et al., 2020) Tingginya angka tersebut menunjukkan bahwa perguruan tinggi memerlukan mekanisme pemantauan akademik yang mampu mendeteksi potensi keterlambatan studi maupun risiko putus kuliah secara lebih dini agar intervensi akademik dapat dilakukan secara tepat sasaran.

Kondisi ini juga terjadi di Universitas XYZ yang berada di wilayah Gresik khususnya Fakultas Teknik dimana tingkat kelulusan tepat waktu mahasiswa adalah  $\leq 50\%$  angka tersebut masih tergolong rendah. Kondisi tersebut menunjukkan bahwa proses pemantauan dan pengelolaan perkembangan akademik mahasiswa belum berjalan secara optimal. Di sisi lain, perguruan tinggi memiliki berbagai data akademik mahasiswa, seperti Indeks Prestasi Semester (IPS), Indeks Prestasi Kumulatif (IPK), serta status akademik mahasiswa. Namun, data tersebut umumnya belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan akademik. Akibatnya, dosen wali maupun ketua program studi mengalami kesulitan dalam mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan kelulusan secara dini.

Berdasarkan kondisi tersebut, diperlukan pendekatan berbasis data yang mampu menghasilkan informasi strategis untuk mendukung pengambilan keputusan akademik. Salah satu solusi yang dapat diterapkan adalah *Early Warning System* (EWS), yaitu sistem yang dirancang untuk mendeteksi secara dini mahasiswa yang berpotensi mengalami keterlambatan atau

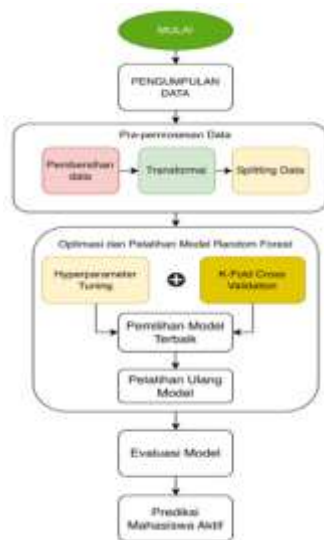
putus studi berdasarkan pola data akademik yang tersedia. Dalam implementasinya, EWS dapat dikembangkan menggunakan *machine learning*, yaitu metode kecerdasan buatan yang memungkinkan sistem mempelajari pola dari data historis untuk menghasilkan prediksi secara otomatis. Penerapan EWS berbasis machine learning telah terbukti efektif dalam membantu perguruan tinggi melakukan pemantauan akademik secara lebih proaktif (Bhutoria, 2022; Meeûs et al., 2023).

Salah satu algoritma *machine learning* yang banyak digunakan untuk membangun EWS dan prediksi kelulusan adalah Random Forest. Random Forest merupakan metode ensemble yang membangun sejumlah pohon keputusan (*decision tree*) secara acak dan menentukan hasil klasifikasi melalui mekanisme voting mayoritas (Breiman, 2001). Algoritma ini memiliki keunggulan dalam menangani data berdimensi tinggi, tahan terhadap *overfitting*, dan mampu menghasilkan informasi *feature importance* yang berguna untuk interpretasi model. Beberapa penelitian telah membuktikan efektivitasnya dalam memprediksi kelulusan mahasiswa dengan tingkat akurasi yang baik (Budi Kurniawan & Farokhah, 2025; Satrio Junaidi et al., 2024)

Dengan demikian, penelitian ini bertujuan untuk membangun model Early Warning System menggunakan algoritma Random Forest guna memprediksi status kelulusan mahasiswa tepat waktu di Fakultas Teknik. Hasil prediksi diharapkan dapat menjadi dasar intervensi akademik yang lebih dini dan tepat sasaran oleh pihak fakultas.

## **METODE**

Penelitian ini dilakukan berbagai tahapan yang dapat di lihat pada Gambar 1.



**Gambar 1. Metodologi Penelitian**

### Pengumpulan Data

Data akademik mahasiswa diperoleh dari sistem informasi akademik pada salah satu Fakultas Teknik Universitas XYZ. Dataset yang digunakan terdiri atas data historis mahasiswa angkatan tahun 2018–2021 sebanyak 1.558 data yang digunakan sebagai dataset pelatihan. Pemilihan tahun 2018-2021 dilakukan karena mahasiswa pada periode tersebut telah memiliki status kelulusan yang jelas, sehingga data dapat digunakan sebagai dasar pembelajaran model dalam mengenali pola kelulusan mahasiswa. Selain itu, digunakan pula data mahasiswa angkatan tahun 2022–2023 sebanyak 677 data sebagai data prediksi mahasiswa aktif. Variabel yang digunakan dalam penelitian ini disajikan pada Tabel 1.

**Tabel 1. Variabel Prediksi**

No	Variabel	Keterangan
1	IPS	Performa akademik mahasiswa pada semester 1-4.
2	IPK	IPK akhir pada semester 4
3	Status Pekerjaan	Mengetahui apakah mahasiswa bekerja/tidak selama masa studi.
4	Status Kelulusan	Variabel target klasifikasi, yaitu tepat waktu dan terlambat.

### Pre-Processing

Tahap pre-processing pada penelitian ini terdiri atas tiga tahapan utama, yaitu *data cleaning*, transformasi data, dan

pembagian dataset (*data splitting*). Tahapan pre-processing merupakan bagian penting dalam pipeline penelitian berbasis data karena kualitas data yang digunakan secara langsung memengaruhi performa model yang dihasilkan (Romadloni & Anugrah, 2021).

Tahap pre-processing pada penelitian ini terdiri atas tiga tahapan utama, yaitu data cleaning, transformasi data, dan pembagian dataset (*data splitting*).

Tahap pertama adalah data cleaning, yang dilakukan dengan melakukan pengecekan terhadap nilai kosong (*missing value*) pada dataset. Data yang memiliki nilai kosong akan dihilangkan agar tidak mengganggu dalam proses pelatihan model.

Tahap kedua adalah transformasi data, yaitu mengubah variabel target Status Kelulusan dari bentuk kategorikal menjadi numerik menggunakan metode label encoding. Pada proses ini, label Lulus Tepat Waktu dikonversi menjadi nilai 1, sedangkan label Terlambat dikonversi menjadi nilai 0.

Tahap ketiga adalah pembagian dataset, yaitu membagi data historis mahasiswa ke menjadi 2 yaitu data latih dan juga data prediksi. Data latih digunakan dalam proses pelatihan model dan cross validation, sedangkan data prediksi digunakan untuk memprediksi mahasiswa aktif.

### Optimasi dan Pelatihan Model Random Forest

Tahap ini mencakup dua proses yang berjalan secara bersamaan sebagaimana ditunjukkan pada Gambar 1, yaitu hyperparameter tuning dan K-Fold Cross Validation, yang bersama-sama digunakan untuk menentukan model terbaik sebelum dilakukan pelatihan ulang menggunakan seluruh data pelatihan. Parameter yang digunakan dalam proses pengujian disajikan pada Tabel 2.

**Tabel 2. Kombinasi Parameter**

HyperParameter	Simbol	Nilai uji
n_estimators	n	100, 150, 200

max_depth	md	5, 10, 15
min_samples_split	mss	2, 5
min_samples_leaf	msl	1, 2, 4, 8
K-fold CV	k	5,10
Data Split	-	70:30, 80:20, 90:10

**Pelatihan Model Random Forest**

Setelah kombinasi parameter terbaik diperoleh, model Random Forest dilatih ulang menggunakan seluruh data pelatihan. Proses pembentukan model diawali dengan *bootstrap sampling*, yaitu pengambilan sampel secara acak dengan pengembalian (*sampling with replacement*) untuk menghasilkan subset data yang berbeda pada setiap pohon keputusan. Selanjutnya, setiap *decision tree* dibangun secara paralel dengan mempertimbangkan sejumlah fitur yang dipilih secara acak pada setiap node pemisahan. Mekanisme pemilihan fitur secara acak ini bertujuan untuk mengurangi korelasi antar pohon sehingga meningkatkan keragaman model dan performa klasifikasi secara keseluruhan.



**Gambar 2. Flowchart Random Forest**

$$f(x) = \text{majority vote}(T_1, T_2, \dots, T_n)$$

Proses pemisahan node dilakukan berdasarkan kriteria impuritas terbaik, seperti Gini impurity atau entropy, hingga memenuhi kondisi penghentian tertentu

yang ditentukan oleh parameter seperti *max\_depth* dan *min\_samples*. Setelah seluruh pohon keputusan terbentuk, masing-masing pohon menghasilkan prediksi kelas secara independen terhadap data input. Tahap akhir dilakukan menggunakan mekanisme *majority voting*, yaitu kelas yang paling banyak diprediksi oleh seluruh pohon dipilih sebagai hasil klasifikasi akhir. Sebagaimana diilustrasikan pada Gambar 2. Mekanisme ini menjadikan Random Forest lebih robust terhadap *overfitting* dibandingkan penggunaan *decision tree* tunggal (Breiman, 2001).

**Pemilihan Model Terbaik**

Model terbaik dipilih berdasarkan nilai F1-Score tertinggi yang diperoleh dari proses optimasi dan pelatihan model melalui pencarian kombinasi *hyperparameter* terbaik pada data pelatihan. Pada proses ini, dataset dibagi menjadi beberapa bagian (*fold*). Setiap iterasi menggunakan sebagian data sebagai data latih dan satu bagian lainnya sebagai data validasi, kemudian proses dilakukan secara bergantian hingga seluruh bagian pernah digunakan sebagai data validasi. Pendekatan ini sejalan dengan praktik yang digunakan pada penelitian berbasis pola data, di mana pemilihan model yang optimal dilakukan berdasarkan konsistensi evaluasi lintas subset data (Setiawan et al., 2019).

Metode F1-Score digunakan sebagai metrik evaluasi utama karena mampu memberikan keseimbangan antara nilai *precision* dan *recall*, sehingga lebih efektif dalam mengevaluasi performa model klasifikasi. Perhitungan F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

**Evaluasi Model**

Model yang telah dilatih kemudian diuji menggunakan data *holdout testing*. Evaluasi dilakukan menggunakan confusion matrix dengan metrik accuracy,

precision, recall, dan F1-score. Rumus masing-masing metrik adalah:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

### Prediksi Mahasiswa Aktif

Model terbaik yang telah dievaluasi kemudian digunakan untuk memprediksi status kelulusan mahasiswa angkatan tahun 2022–2023 sebanyak 677 data. Hasil prediksi digunakan sebagai bagian dari implementasi *Early Warning System* untuk mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan kelulusan sehingga dapat dilakukan tindakan preventif secara lebih dini.

### HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penerapan algoritma Random Forest sebagai model *Early Warning System* dalam memprediksi status kelulusan mahasiswa tepat waktu. Pembahasan mencakup analisis data, proses pembentukan model, perbandingan seluruh skenario eksperimen, evaluasi model terbaik, serta interpretasi hasil prediksi mahasiswa aktif.

#### Distribusi Data Training

Data training yang digunakan terdiri dari 1.558 mahasiswa angkatan 2018–2021 dengan distribusi label sebagaimana ditunjukkan pada Tabel 3.

**Tabel 3. Distribusi Data**

	Jumlah Mahasiswa	Persentase (%)
Terlambat	869	55,70%
Lulus Tepat Waktu	689	44,30%
Total	1.558	100%

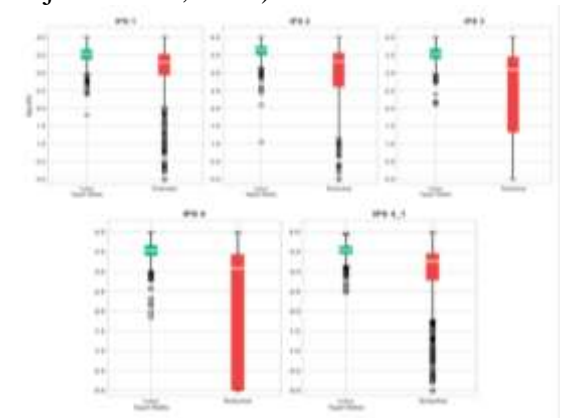
Distribusi label menunjukkan ketidakseimbangan yang moderat antara kelas Terlambat (55,7%) dan Lulus Tepat Waktu (44,3%). Selisih sebesar 11,4% masih tergolong dalam batas yang dapat ditangani algoritma Random Forest tanpa

teknik penyeimbangan tambahan. Ketidakseimbangan yang ekstrem umumnya didefinisikan jika rasio antar kelas melebihi 3:1 (Nurmalitasari et al., 2023). sedangkan pada dataset ini rasionya hanya sekitar 1,26:1. Dengan kondisi ini, model Random Forest dapat belajar secara proporsional dari kedua kelas sehingga menghasilkan prediksi yang tidak bias terhadap salah satu kelas.

#### Distribusi Nilai IPS per Status Kelulusan

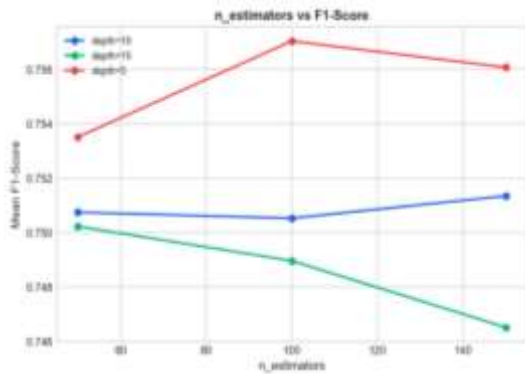
Analisis distribusi nilai IPS per semester berdasarkan status kelulusan disajikan dalam bentuk boxplot pada Gambar 3. Hasil analisis boxplot menunjukkan pola yang konsisten di seluruh semester: mahasiswa dengan status Lulus Tepat Waktu cenderung memiliki distribusi nilai IPS yang lebih tinggi dan lebih stabil (nilai median sekitar 3,5) dibandingkan mahasiswa dengan status Terlambat.

Selain itu, kelompok Terlambat menunjukkan sebaran (*variance*) yang jauh lebih besar dengan banyak outlier ke arah nilai rendah, terutama pada IPS 3 dan IPS 4. Pola ini mengonfirmasi bahwa performa akademik pada semester awal hingga pertengahan studi memiliki relevansi kuat terhadap ketepatan waktu kelulusan, sejalan dengan temuan penelitian sebelumnya yang menyatakan bahwa kinerja akademik semester awal merupakan prediktor signifikan kelulusan (Guanin-Fajardo et al., 2024).



**Gambar 3. Distribusi Nilai IPS per Status Kelulusan**





**Gambar 6. Pengaruh n\_estimator terhadap Mean F1-Score**

Hal ini menunjukkan bahwa pohon yang terlalu dalam cenderung overfitting pada data pelatihan sehingga menurunkan kemampuan generalisasi model (Breiman, 2001). Dari Gambar 6, penambahan n\_estimators dari 50 ke 100 tidak memberikan peningkatan F1-Score yang berarti, sehingga n\_estimators=50 dipilih sebagai nilai yang efisien secara komputasi. Parameter terbaik dan nilai Mean F1-Score CV dari seluruh eksperimen dirangkum pada Tabel 5.

**Tabel 5. Perbandingan Parameter terbaik dan Mean F1-score cv**

Eksperimen	n	md	mss	msl	Mean F1 cv	Std F1-CV
cv5_s90	100	5	5	1	0,758	0,0383
cv5_s80	150	10	2	2	0,763	0,0309
cv10_s90	100	5	5	2	0,756	0,0411
cv10_s80	50	15	2	4	0,759	0,0271
cv5_s70	50	5	5	1	0,759	0,0380
cv10_s70	50	10	2	2	0,753	0,0409

Terdapat perbedaan yang menarik pada parameter terbaik yang dipilih oleh masing-masing eksperimen. Eksperimen dengan split 90:10 dan 70:30 secara konsisten memilih *max\_depth*=5, sedangkan eksperimen dengan split 80:20 cenderung memilih *max\_depth* yang lebih besar (10 dan 15).

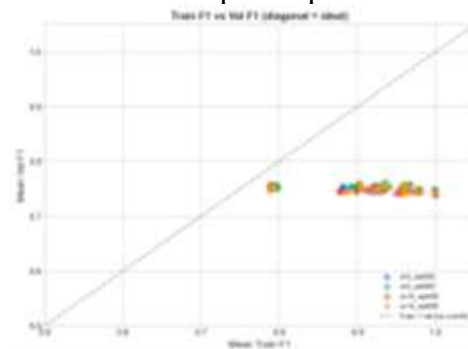
Perbedaan ini kemungkinan disebabkan oleh jumlah data pelatihan yang berbeda: split 90:10 menghasilkan 1.402 data latih, split 80:20 sebanyak 1.246, dan split 70:30 sebanyak 1.090 data latih. Dengan data latih yang lebih sedikit, model pada split 80:20 cenderung membutuhkan pohon yang lebih dalam untuk menangkap

pola yang sama, namun hal ini sekaligus meningkatkan risiko *overfitting*.

Meskipun cv5\_split80 menghasilkan Mean F1-CV tertinggi (0,7627) di antara eksperimen yang diuji, nilai ini belum dapat dijadikan patokan final karena pengujian pada data holdout yang benar-benar baru diperlukan untuk menilai kemampuan generalisasi yang sesungguhnya.

### Analisis Overfitting Lintas Eksperimen

Sebelum menentukan model terbaik, dilakukan analisis overfitting secara menyeluruh untuk memastikan bahwa model yang dipilih benar-benar andal dan tidak hanya bagus pada data pelatihan. Hasil analisis ditampilkan pada Gambar 7



**Gambar 7. Analisis Overfit Lintas Eksperimen ( cv-fold X splitt)**

Titik-titik dari seluruh eksperimen terkonsentrasi pada area yang sama dan tidak ada satu eksperimen pun yang memiliki pola berbeda secara signifikan, menunjukkan bahwa perbedaan strategi CV (5-Fold vs 10-Fold) maupun rasio split tidak menyebabkan perbedaan karakteristik model yang signifikan.

Pasangan eksperimen cv5 dan cv10 pada setiap rasio split menghasilkan distribusi yang hampir identik, mengonfirmasi bahwa penggunaan 10-Fold tidak memberikan keuntungan tambahan dibandingkan 5-Fold pada dataset ini.

### Evaluasi Model pada Holdout Test Set

Tahap ini merupakan pengujian final terhadap data yang sama sekali tidak pernah dilihat oleh model selama proses pelatihan maupun cross validation. Perbandingan performa seluruh eksperimen pada holdout test set disajikan pada Tabel 6.

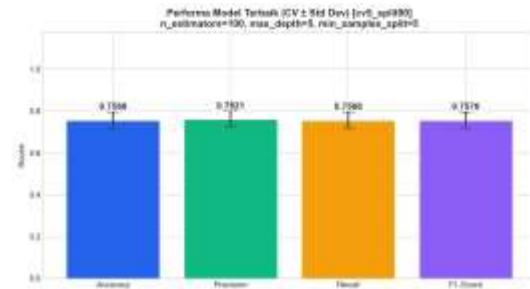
**Tabel 6. Performa Seluruh Eksperimen pada Holdout Test Set**

Eksperimen	Acc	F1-T	F1-L	F1 Weight	OOB
cv5_s90	76,3%	0,773	0,752	0,764	0,751
cv10_s90	76,3%	0,773	0,752	0,764	0,752
cv10_s80	76,0%	0,789	0,719	0,759	0,744
cv5_s80	74,7%	0,773	0,713	0,747	0,737
cv10_s70	76,9%	0,791	0,743	0,770	0,737
cv5_s70	<b>77,6%</b>	<b>0,790</b>	<b>0,760</b>	<b>0,776</b>	<b>0,742</b>

Berdasarkan Tabel 6, terdapat beberapa temuan penting. Pertama, eksperimen cv5\_split70 menghasilkan performa holdout tertinggi dengan akurasi 77,56% dan F1 Weighted 0,7764, diikuti cv10\_split70 (0,7695), cv5\_split90 dan cv10\_split90 (0,7636), cv10\_split80 (0,7586), dan cv5\_split80 (0,7467). Kedua, meskipun cv5\_split80 memiliki Mean F1-CV tertinggi pada tahap cross validation (0,7627), performanya pada holdout justru paling rendah (0,7467). Penurunan ini mengindikasikan bahwa model cv5\_split80 sedikit overfit terhadap data CV karena jumlah data pelatihan yang lebih terbatas. Ketiga, perbandingan antara cv5 dan cv10 pada setiap rasio split menunjukkan hasil yang hampir identik, mengonfirmasi bahwa penambahan fold dari 5 menjadi 10 tidak memberikan peningkatan performa yang berarti pada dataset ini, sehingga 5-Fold CV sudah mencukupi secara efisiensi komputasi. Keempat, OOB Score seluruh eksperimen konsisten mendekati akurasi holdout, yang menunjukkan bahwa OOB Score merupakan estimator yang cukup andal untuk dataset ini.

Berdasarkan pertimbangan F1 Weighted tertinggi pada holdout test set, cv5\_split70 dipilih sebagai model terbaik untuk tahap selanjutnya. Detail evaluasi model terbaik ini ditampilkan pada Gambar 7 dan Tabel 7.

Model terbaik menghasilkan Mean Accuracy 0,7578, Mean Precision 0,7641, Mean Recall 0,7578, dan Mean F1-Score 0,7585 pada evaluasi CV dengan standar deviasi sekitar 0,038.

**Gambar 8. Performa Model Terbaik (CV ± Std Dev) [cv5\_split70]**

Standar deviasi yang kecil mengindikasikan model stabil di seluruh fold dan tidak bergantung pada komposisi data tertentu. Overfit gap sebesar 0,0334 berada di bawah ambang 0,05, menunjukkan model mampu menggeneralisasi pola dengan baik.

Hasil confusion matrix pada holdout test set ditampilkan pada Tabel 8 dan dirangkum pada Tabel 7.

**Tabel 7. Detail Hasil Evaluasi Model Terbaik [cv5\_split70]**

Metrik	Terlambat	Lulus Tepat Waktu	Rata-rata
Precision	0,8277	0,7217	0,7747
Recall	0,7548	0,8019	0,7783
F1-Score	0,773	0,7597	0,7747
Accuracy	—	—	77,56%

**Tabel 8. Hasil Confusion Matrix**

	Prediksi Terlambat	Prediksi Lulus Tepat Waktu
Aktual Terlambat	197 (TN)	64 (FP)
Aktual Lulus Tepat Waktu	41 (FN)	166 (TP)

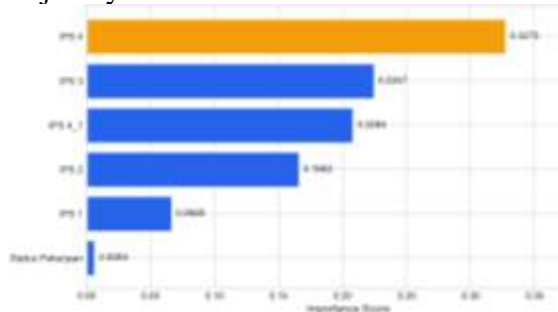
Model berhasil mengklasifikasikan 197 dari 261 mahasiswa Terlambat dengan benar dan 166 dari 207 mahasiswa Lulus Tepat Waktu dengan benar. Nilai False Positive (64) yang lebih besar dari False Negative (41) menunjukkan model lebih konservatif dalam memprediksi keterlambatan, yang dalam konteks EWS justru menguntungkan karena meminimalkan jumlah mahasiswa berisiko yang luput dari deteksi. Precision kelas Terlambat yang tinggi (0,8277) juga menunjukkan bahwa ketika model

memprediksi seorang mahasiswa akan terlambat, prediksi tersebut memiliki tingkat kebenaran yang cukup tinggi sehingga dapat dijadikan dasar intervensi akademik yang tepat sasaran.

### Analisis Feature Importance

Salah satu keunggulan Random Forest adalah kemampuannya menghasilkan informasi feature importance yang menunjukkan kontribusi relatif setiap variabel terhadap prediksi model. Hasil analisis ditampilkan pada Gambar 9 untuk model terbaik (cv5\_split70). Secara keseluruhan, IPS semester 3 dan 4 menyumbang lebih dari 53% pengaruh prediksi model, yang menandakan bahwa performa akademik pada akhir tahun kedua menjadi indikator paling kuat terhadap ketepatan waktu kelulusan.

Kondisi ini relevan dengan karakteristik proses akademik di Universitas XYZ, di mana pada semester 3 dan 4 mahasiswa mulai memasuki mata kuliah inti program studi dengan tingkat kompleksitas yang lebih tinggi, sehingga penurunan performa akademik pada fase tersebut dapat menjadi sinyal awal terjadinya keterlambatan studi.



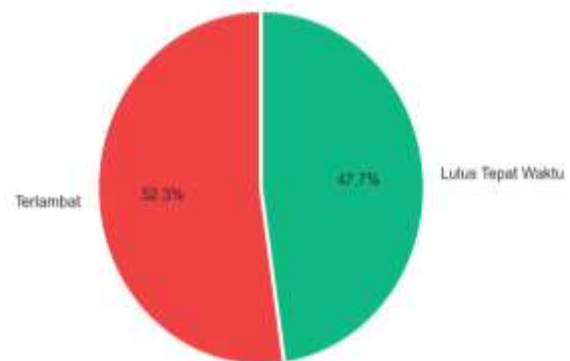
**Gambar 9. Feature Importance**

Temuan ini juga sejalan dengan penelitian Widya Dharma Sidi (2026) yang menyatakan IPS Semester 4 sebagai prediktor utama keterlambatan kelulusan. Sementara itu, IPS 1 memiliki variasi kontribusi yang lebih besar antar eksperimen, sedangkan Status Pekerjaan memberikan pengaruh yang sangat kecil. Oleh karena itu, perguruan tinggi perlu memfokuskan pemantauan akademik pada mahasiswa yang mengalami penurunan IPS

pada semester 3 dan 4 sebagai langkah deteksi dini risiko keterlambatan kelulusan.

### Hasil Prediksi Mahasiswa Baru

Model terbaik (cv5\_split70) digunakan untuk memprediksi status kelulusan 677 mahasiswa aktif angkatan 2022–2023. Hasil prediksi ditampilkan pada Gambar 10. Dari 677 mahasiswa yang diprediksi, sebanyak 321 mahasiswa (47,4%) diprediksi Lulus Tepat Waktu dan 356 mahasiswa (52,6%) diprediksi Terlambat. Proporsi ini konsisten dengan distribusi kelas pada data training (Tabel 3), yang menunjukkan model tidak mengalami bias prediksi yang signifikan terhadap salah satu kelas.



**Gambar 10. Hasil Prediksi Status Kelulusan Mahasiswa Aktif [cv5\_split70]**

Selain itu, dominasi prediksi pada kategori terlambat mengindikasikan bahwa lebih dari separuh mahasiswa aktif masih berada pada kelompok yang memerlukan perhatian akademik lebih lanjut. Temuan ini sejalan dengan hasil *feature importance* sebelumnya yang menunjukkan IPS semester 3 dan IPS semester 4 sebagai prediktor utama dalam model, sehingga performa akademik pada pertengahan masa studi menjadi indikator penting terhadap ketepatan waktu kelulusan. Berdasarkan hasil tersebut, model *Early Warning System* yang dibangun dapat dimanfaatkan oleh program studi untuk mengidentifikasi mahasiswa berisiko lebih awal serta mendukung pemberian intervensi akademik yang lebih terarah dan tepat waktu.

## SIMPULAN

Penelitian ini berhasil membangun model *Early Warning System* berbasis algoritma Random Forest untuk memprediksi status kelulusan mahasiswa tepat waktu di salah satu Fakultas Teknik di wilayah Gresik. Dari enam skenario eksperimen yang diuji melalui kombinasi K-Fold CV (5 dan 10) dengan rasio data split (70:30, 80:20, dan 90:10), eksperimen *cv5\_split70* dipilih sebagai model terbaik dengan parameter *n\_estimators=50*, *max\_depth=5*, *min\_samples\_split=5*, dan *min\_samples\_leaf=1*. Model ini menghasilkan F1-Score (*weighted average*) tertinggi sebesar 0,7764 dan akurasi 77,56% pada *holdout test set*, dengan *overfit gap* sebesar 0,0334 yang berada di bawah ambang batas 0,05, menunjukkan kemampuan generalisasi yang baik.

Nilai akurasi 77,56% dan F1-Score 0,7764 yang dihasilkan mencerminkan kapabilitas prediktif model yang cukup baik, namun sekaligus mengindikasikan adanya keterbatasan informasi pada fitur yang tersedia. Model hanya menggunakan enam variabel input berupa IPS semester 1–4, IPK semester 4, dan status pekerjaan, sehingga faktor-faktor lain yang secara empiris turut memengaruhi kelulusan seperti kehadiran, keterlibatan organisasi, kondisi ekonomi, maupun riwayat pembayaran yang belum tertangkap dalam model. Keterbatasan cakupan fitur ini menjadi salah satu faktor yang membatasi performa model untuk mencapai akurasi yang lebih tinggi. Selain itu, distribusi kelas yang tidak seimbang antara mahasiswa terlambat (869 data) dan lulus tepat waktu (689 data) juga turut berkontribusi pada nilai F1-Score yang belum optimal, karena model cenderung lebih mudah mengenali kelas mayoritas dibandingkan kelas minoritas.

Perbandingan lintas enam skenario eksperimen menunjukkan pola yang konsisten dan dapat dijelaskan secara sistematis. Eksperimen dengan split 70:30 menghasilkan performa *holdout* tertinggi (*cv5\_split70*: akurasi 77,56%, F1 0,7764),

meskipun menggunakan data latih paling sedikit (1.090 data). Hal ini terjadi karena split 70:30 menyisakan porsi data uji yang lebih besar dan lebih representatif, sehingga evaluasi pada *holdout* lebih mencerminkan kemampuan generalisasi model yang sesungguhnya. Sebaliknya, split 80:20 justru menghasilkan performa *holdout* terendah (*cv5\_split80*: F1 0,7467), meskipun memiliki nilai Mean F1-CV tertinggi pada tahap *cross validation* (0,7627). Pola ini mengindikasikan bahwa model pada split 80:20 mengalami sedikit *overfitting* terhadap distribusi data CV, sehingga performa aktualnya pada data baru tidak sebaik yang ditunjukkan saat validasi. Sementara itu, split 90:10 menghasilkan performa *holdout* yang stabil di tengah (F1  $\pm$ 0,764), namun dengan *holdout* hanya 156 data, evaluasinya kurang representatif karena sangat sensitif terhadap komposisi sampel yang terpilih. Adapun perbandingan antara 5-Fold dan 10-Fold CV pada setiap rasio split secara konsisten menghasilkan performa yang hampir identik, mengonfirmasi bahwa penambahan jumlah *fold* dari 5 menjadi 10 tidak memberikan peningkatan performa yang berarti pada dataset ini, sehingga 5-Fold CV sudah mencukupi secara efisiensi komputasi.

Analisis *feature importance* mengungkapkan bahwa IPS 4 merupakan prediktor paling dominan dengan rata-rata skor 0,3164 di seluruh eksperimen, diikuti IPS 3 (0,2157) dan IPS 4\_1 (0,1990). Secara keseluruhan, nilai IPS semester 3 dan 4 berkontribusi dominan terhadap keputusan prediksi model, mengimplikasikan bahwa performa akademik pada akhir tahun kedua masa studi merupakan titik kritis yang paling perlu dipantau. Sebaliknya, variabel Status Pekerjaan konsisten memiliki kontribusi sangat kecil (rata-rata 0,0118) di seluruh eksperimen, menunjukkan bahwa faktor pekerjaan tidak secara signifikan membedakan kedua kelas pada dataset ini. Penerapan model pada 677 mahasiswa aktif angkatan 2022–2023 menghasilkan

prediksi 356 mahasiswa (52,6%) berisiko terlambat dan 321 mahasiswa (47,4%) diprediksi lulus tepat waktu.

Untuk penelitian selanjutnya, disarankan untuk menambahkan variabel non-akademik seperti data kehadiran dan riwayat pembayaran, serta membandingkan performa Random Forest dengan algoritma lain seperti XGBoost atau *Gradient Boosting* guna memperoleh model prediksi yang lebih akurat.

#### DAFTAR PUSTAKA

- Ayyub, A., & Iddrus, I. (2024). Prediksi Kelulusan Mahasiswa Tepat Waktu atau Tidak Menggunakan Metode Naive Bayes Classifier (Studi kasus : Mahasiswa Fakultas Teknik). *Journal of Advanced Research in Informatics*, 3(1), 18–28. <https://doi.org/10.24929/jars.v3i1.3787>
- Bhutoria, A. (2022). Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. *Computers and Education: Artificial Intelligence*, 3, 100068. <https://doi.org/10.1016/j.caeai.2022.100068>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Budi Kurniawan, F., & Farokhah, L. (2025). Aplikasi Cerdas Prediksi Kelulusan Mahasiswa Berbasis Website Menggunakan Metode Support Vector Machine (SVM). *JURNAL FASILKOM*, 15(1), 155–162. <https://doi.org/10.37859/jf.v15i1.8767>
- Guanin-Fajardo, J. H., Guña-Moya, J., & Casillas, J. (2024). Predicting Academic Success of College Students Using Machine Learning Techniques. *Data*, 9(4), 60. <https://doi.org/10.3390/data9040060>
- Handini, D., Hidayat, F., Attamimi, A. N., Rouf, F., & Anjani, N. (2020). Statistik pendidikan tinggi tahun 2020 (higher education statistics 2020). *Kemendikdasmen Repositori Institusi*, 5.
- Hasibuan, T. H., & Mahdiana, D. (2023). PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN ALGORITMA C4.5 PADA UIN SYARIF HIDAYATULLAH JAKARTA. *SKANIKA*, 6(1), 61–74. <https://doi.org/10.36080/skanika.v6i1.2976>
- Meeûs, J., Dewulf, W., & Macário, R. (2023). Management Systems in Aviation: Challenges and Opportunities to Upgrade to an Integrated Management System. *Sustainability*, 15(13), 10424. <https://doi.org/10.3390/su151310424>
- Nurmalitasari, Awang Long, Z., & Faizuddin Mohd Noor, M. (2023). Factors Influencing Dropout Students in Higher Education. *Education Research International*, 2023, 1–13. <https://doi.org/10.1155/2023/7704142>
- Romadloni, M. D., & Anugrah, I. G. (2021). Implementation of EM Algorithm in Opinion Mining Movies Review Case Studies. *Journal of Development Research*, 5(2), 94–105. <https://doi.org/10.28926/jdr.v5i2.149>
- Satrio Junaidi, Valicia Anggela, R., & Kariman, D. (2024). Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Network (ANN). *Journal of Applied Computer Science and Technology*, 5(1), 109–119. <https://doi.org/10.52158/jacost.v5i1.489>
- Setiawan, A., Indra, D., & Anugrah, G. (2019). Penentuan Pola Pembelian

Konsumen pada Indomaret GKB Gresik dengan Metode FP-Growth. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 2(2).

Wan Yaacob, W. F., Mohd Sobri, N., Nasir, S. A. M., Wan Yaacob, W. F., Norshahidi, N. D., & Wan Husin, W. Z. (2020). Predicting Student Drop-Out in Higher Institution Using Data Mining Techniques. *Journal of Physics: Conference Series*, 1496, 012005.

<https://doi.org/10.1088/1742-6596/1496/1/012005>

Widya Dharma Sidi. (2026). Dominansi Fitur Perilaku Keuangan dan Tren Akademik pada Deteksi Dini Kelulusan Mahasiswa dengan XGBoost. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 7(3), 1039–1049.

<https://doi.org/10.30865/json.v7i3.9591>.