

PENINGKATAN KINERJA XGBOOST DALAM KLASIFIKASI PEUBAH PENTING INDEKS PRESTASI MAHASISWA SNBP IPB UNIVERSITY BERBASIS *EASYENSEMBLE*

SALSABILA FAYIZA



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Peningkatan Kinerja XGBoost dalam Klasifikasi Peubah Penting Indeks Prestasi Mahasiswa SNBP IPB University Berbasis *EasyEnsemble*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dalam penyusunan karya ini penulis menggunakan bantuan kecerdasan buatan berupa ChatGPT berbasis GPT-5.5 untuk membantu penyuntingan bahasa, penyesuaian kalimat, dan peninjauan paragraf. Seluruh hasil penggunaan layanan tersebut telah ditinjau dan disunting kembali oleh penulis sesuai kebutuhan, sehingga penulis bertanggung jawab penuh atas isi karya tugas akhir ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juni 2026

Salsabila Fayiza
G1401221036



ABSTRAK

SALSABILA FAYIZA. Peningkatan Kinerja XGBoost dalam Klasifikasi Peubah Penting Indeks Prestasi Mahasiswa SNBP IPB University Berbasis *EasyEnsemble*. Dibimbing oleh UTAMI DYAH SYAFITRI dan AULIA RIZKI FIRDAWANTI.

Implementasi Kurikulum Merdeka menekankan pembelajaran yang fleksibel dan berbasis kompetensi, sehingga menghasilkan mahasiswa dengan latar belakang akademik yang semakin beragam. Kondisi ini berkaitan erat dengan skema penerimaan SNBP yang menyeleksi mahasiswa berdasarkan prestasi akademik tanpa melalui ujian tertulis. Penelitian ini bertujuan mengklasifikasikan mahasiswa program sarjana IPB yang diterima melalui jalur SNBP berdasarkan kategori IP semester 1 menggunakan metode XGBoost dengan penanganan ketidakseimbangan data menggunakan pendekatan *EasyEnsemble*. Pada metode XGBoost dilakukan *hyperparameter tuning* dengan pendekatan *Fractional Factorial Design* 3^{4-1} . Data IP dikategorikan ke dalam skema tiga kelas dan empat kelas, kemudian dievaluasi menggunakan *K-Fold Cross Validation*. Kinerja model diukur menggunakan metrik akurasi, *G-Mean*, presisi, sensitivitas, dan skor F1, sedangkan *feature importance* berdasarkan nilai *gain* XGBoost digunakan untuk mengidentifikasi peubah yang berpengaruh. Hasil penelitian menunjukkan bahwa model XGBoost dengan *EasyEnsemble* pada skema tiga kelas memberikan performa klasifikasi yang paling seimbang dengan nilai *G-Mean* sebesar 58.18% dan sensitivitas sebesar 47.44%, sehingga lebih mampu menangani ketidakseimbangan kelas, terutama pada kelas minoritas. Peubah yang paling berpengaruh meliputi nilai rata-rata keseluruhan, rata-rata nilai mata pelajaran yang menjadi syarat, program studi, provinsi sekolah, dan klaster PPKU. Penerapan *EasyEnsemble* pada XGBoost dengan *hyperparameter tuning* menggunakan *Fractional Factorial Design* 3^{4-1} menghasilkan model klasifikasi terbaik pada skema tiga kelas.

Kata kunci: *EasyEnsemble*, klasifikasi, ketidaksimbangan kelas, peubah penting, XGBoost



ABSTRACT

SALSABILA FAYIZA. Improving XGBoost Performance for Classifying Importance Variables of SNBP Student GPA at IPB Using EasyEnsemble. Supervised by UTAMI DYAH SYAFITRI and AULIA RIZKI FIRDAWANTI.

The implementation of the Merdeka Curriculum emphasizes flexible and competency-based learning, resulting in students with a more diverse range of academic backgrounds. This is closely related to the SNBP admission scheme, which selects students based on academic performance without a written examination. This study aims to classify undergraduate students at IPB admitted through the SNBP based on their first-semester GPA categories using XGBoost, with EasyEnsemble employed to address class imbalance. Hyperparameter tuning of XGBoost was conducted using a 3^{4-1} Fractional Factorial Design (FFD). GPA data are categorized into three-class and four-class schemes and evaluated using K-Fold Cross Validation. Model performance is measured using accuracy, G-Mean, precision, recall, and F1-score, while feature importance is assessed using XGBoost gain values to identify influential variables. The results showed that the three-class XGBoost model with EasyEnsemble achieves the most balanced classification performance, with a G-Mean of 58.18% and a sensitivity of 47.44%. These results indicate that the proposed model was more effective in handling class imbalance, particularly for the minority classes. The most influential variables include overall average grades, average required grades, study program, school province, and “PPKU” cluster, which directly affect the model’s results and first-semester GPA. The implementation of EasyEnsemble on XGBoost with hyperparameter tuning using a 3^{4-1} FFD resulted in the best-performing classification model for the three-class scheme.

Keywords: classification, class imbalance, EasyEnsemble, feature importance, XGBoost



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENINGKATAN KINERJA XGBOOST DALAM KLASIFIKASI PEUBAH PENTING INDEKS PRESTASI MAHASISWA SNBP IPB UNIVERSITY BERBASIS EASYENSEMBLE

SALSABILA FAYIZA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
1. Eka Putri Nur Utami, S.Stat., M.Si.

Perpustakaan IPB University



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Peningkatan Kinerja XGBoost dalam Klasifikasi Peubah Penting
Indeks Prestasi Mahasiswa SNBP IPB University Berbasis
EasyEnsembe

Nama : Salsabila Fayiza
NIM : G1401221036

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:

Dr. Utami Dyah Syafitri, S.Si., M.Si.

Pembimbing 2:

Aulia Rizki Firdawanti, S.Stat., M.Si.

Diketahui oleh

Ketua Program Studi:

Dr. Bagus Sartono, S.Si., M.Si.
NIP. 19780411 200501 1002

Tanggal Ujian:
29 Juni 2026

Tanggal Lulus:

(tanggal penandatanganan oleh Dekan
Fakultas/Sekolah ...)

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2025 sampai bulan Juni 2026 ini ialah Indeks Prestasi Mahasiswa, dengan judul “Peningkatan Kinerja XGBoost dalam Klasifikasi Peubah Penting Indeks Prestasi Mahasiswa SNBP IPB University Berbasis *EasyEnsemble*”.

Terselesaikannya karya ilmiah ini tidak terlepas dari dukungan dan peran dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Dr. Utami Dyah Syafitri, S.Si., M.Si. dan Ibu Aulia Rizki Firdawanti, S.Stat., M.Si. selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, motivasi, serta masukan yang membangun selama proses penyusunan karya ilmiah ini. Penulis juga mengucapkan terima kasih atas kesabaran, ketulusan, dan dedikasi yang diberikan, termasuk kesediaan untuk tetap membimbing di luar jam kerja, serta respons yang cepat terhadap berbagai kendala yang dihadapi penulis hingga karya ilmiah ini dapat terselesaikan.
2. Ibu Eka Putri Nur Utami, S.Stat., M.Si. selaku dosen penguji luar komisi, Ibu Laily Nissa Atul Mualifah, S.Si., M.Si. selaku moderator seminar hasil, dan Bapak Prof. Dr. Ir. I Made Sumertajaya M.Si. selaku moderator kolokium atas saran dan masukan untuk penulisan karya ilmiah ini.
3. Direktorat Administrasi Pendidikan dan Penerimaan Mahasiswa Baru (DAPPMB) IPB University yang telah memberikan penulis kesempatan untuk menggunakan data mahasiswa dalam penelitian ini.
4. Bapak Helmy Ahmad, S.H yang senantiasa berjuang untuk kehidupan dan pendidikan penulis, memberikan dukungan, motivasi, dan doa yang tidak pernah putus di setiap langkah penulis, yang selalu menjadi kekuatan serta sumber semangat.
5. Almarhumah Ibu Citra Norisya, A.Md.Keb. yang telah lebih dahulu berpulang pada 10 Maret 2024. Skripsi ini merupakan wujud dari doa-doa yang beliau titipkan, meskipun beliau tidak sempat menyaksikan penulis menyelesaikan pendidikan. Terima kasih atas kasih sayang, doa, dukungan, serta kesediaan untuk selalu hadir dan mendengarkan setiap keluh kesah penulis.
6. Ananda Myratama Ahmad, Asyifa Fayiza Ahmad, dan keluarga besar Syafrudin.Y, yang senantiasa memberikan doa, dukungan, dan semangat sehingga penulis bisa menyelesaikan perjalanan akademik ini.
7. Delia Fitri Audina, Salsa Rifda Erira, dan Raihana Asma Amani sebagai sahabat penulis yang selalu menemani, membantu, memberikan dukungan, serta memberikan masukan dalam mempersiapkan kolokium, seminar hasil, hingga sidang akhir.
8. Deswita Nur Alpharofi, Sandra Berliana Putri, Yulianti Nurdzanah, serta teman-teman Statistika 59 yang telah memberikan dukungan, semangat, dan dorongan positif kepada penulis selama masa kuliah dan penyusunan karya ilmiah ini.



9. Tiara Ardita, Bintang Rheinjani N.M, Ayu Bagus Larasati, dan Laura Alivia Hermawan sebagai teman KKN penulis yang memberikan dukungan, semangat, dan motivasi selama satu tahun terakhir.
10. Fauzan Ilhamda Saputra, Nindya Lyana Prasya, dan Syakila Pricilia selaku sahabat yang telah menemani penulis sejak masa Sekolah Menengah Pertama, yang senantiasa hadir dalam setiap fase kehidupan dengan kebersamaan, cerita, dukungan, dan semangat.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juni 2026

Salsabila Fayiza



DAFTAR ISI

DAFTAR TABEL	xii
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xii
I PENDAHULUAN	13
1.1 Latar Belakang	13
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 <i>eXtreme Gradient Boosting (XGBoost) Classifier</i>	3
2.2 <i>Fractional Factorial Design (FFD)</i>	5
2.3 <i>EasyEnsemble</i>	5
2.4 Evaluasi Kebaikan Model	6
2.5 <i>Feature Importance</i>	9
III METODE	10
3.1 Data	10
3.2 Prosedur Analisis Data	11
IV HASIL DAN PEMBAHASAN	14
4.1 Praproses Data	14
4.2 Eksplorasi Data	14
4.3 <i>Hyperparameter Tuning</i> dengan 3^{4+1} <i>Fractional Factorial Design</i>	16
4.4 Pengkategorian Peubah Target dan Pembagian Data dengan <i>Cross-Validation</i>	18
4.5 Penanganan Ketidakseimbangan Data dengan <i>EasyEnsemble</i>	18
4.6 Perbandingan Model XGBoost dan Evaluasi	19
4.7 Identifikasi <i>Feature Importance</i>	20
V SIMPULAN DAN SARAN	22
5.1 Simpulan	22
5.2 Saran	22
DAFTAR PUSTAKA	23
LAMPIRAN	25
RIWAYAT HIDUP	29



DAFTAR TABEL

1	<i>Confusion matrix</i> klasifikasi biner	6
2	<i>Confusion matrix</i> multikelas	7
3	Definisi TP, FN, FP, dan TN untuk kelas ke-k	7
4	Peubah yang digunakan pada penelitian	10
5	Kategorisasi IP skema pertama (3 kelas)	11
6	Kategorisasi IP skema kedua (4 kelas)	11
7	Nilai <i>hyperparameter</i> yang digunakan	12
8	Hasil signifikansi <i>hyperparameter</i>	17
9	Kombinasi <i>hyperparameter</i> terbaik	17
10	Komposisi data latih sebelum dan sesudah <i>EasyEnsemble</i>	18
11	Metrik evaluasi	19
12	<i>Confusion matrix</i> hasil prediksi dari model 2	20

DAFTAR GAMBAR

1	Diagram alir tahapan analisis data	13
2	Pie Chart kurikulum sekolah dan jurusan sekolah	14
3	Histogram sebaran IP semester 1 mahasiswa S1 SNBP angkatan 2025	15
4	Sebaran IP semester 1 berdasarkan Fakultas	15
5	Sebaran IP semester 1 berdasarkan Fakultas	16
6	Peubah penting dalam klasifikasi IP Semester 1	21

DAFTAR LAMPIRAN

1	Lampiran 1 Keterangan kode program studi	26
2	Lampiran 2 Keterangan Klaster Sekolah	27
3	Lampiran 3 Tabel ANOVA untuk skema 1	27
4	Lampiran 4 Tabel ANOVA untuk skema 2	27
5	Lampiran 5 <i>Confusion matrix</i> model 1	27
6	Lampiran 6 <i>Confusion matrix</i> model 3	28
7	Lampiran 7 <i>Confusion matrix</i> model 4	28
8	Lampiran 8 Nilai <i>gain</i> dari peubah yang digunakan	28