

**PENGARUH RANDOMIZEDSEARCHCV PADA OPTIMASI
RANDOM FOREST DAN XGBOOST DALAM
MEMPREDIKSI PERILAKU KONSUMEN
PADA PLATFORM E-COMMERCE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi SI-Informatika



disusun oleh

DWI KURNIAWAN SAPUTRA
21.11.3856

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

**PENGARUH RANDOMIZEDSEARCHCV PADA OPTIMASI
RANDOM FOREST DAN XGBOOST DALAM
MEMPREDIKSI PERILAKU KONSUMEN
PADA PLATFORM E-COMMERCE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

DWI KURNIAWAN SAPUTRA

21.11.3856

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**PENGARUH RANDOMIZEDSEARCHCV PADA OPTIMASI RANDOM
FOREST DAN XGBOOST DALAM MEMPREDIKSI PERILAKU
KONSUMEN PADA PLATFORM E-COMMERCE**

yang disusun dan diajukan oleh

Dwi Kurniawan Saputra

21.11.3856

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 8 Agustus 2025

Dosen Pembimbing,



Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

HALAMAN PENGESAHAN

SKRIPSI

**PENGARUH RANDOMIZEDSEARCHCV PADA OPTIMASI RANDOM
FOREST DAN XGBOOST DALAM MEMPREDIKSI PERILAKU
KONSUMEN PADA PLATFORM E-COMMERCE**

yang disusun dan diajukan oleh

Dwi Kurniawan Saputra

21.11.3856

Telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Nafiatun Sholihah, S.Kom., M.Cs
NIK. 190302524

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

Tanda Tangan



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 20 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Dwi Kurniawan Saputra

NIM : 21.11.3856

Menyatakan bahwa Skripsi dengan judul berikut:

Pengaruh RandomizedSearchCV Pada Optimasi Random Forest Dan Xgboost Dalam Memprediksi Perilaku Konsumen Pada Platform E-Commerce

Dosen Pembimbing : Ike Vrawati, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan **gagasan, rumusan dan penelitian SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 8 Agustus 2025

Yang Menyatakan,



Dwi Kurniawan Saputra

HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat, taufik, dan hidayah-Nya. Berkat karunia-Nya, penulis diberi kekuatan, kesehatan, serta kemudahan sehingga skripsi ini dapat disusun dan diselesaikan dengan baik.

Penulis juga ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah memberikan dukungan, bantuan, serta motivasi selama proses penyusunan skripsi ini, antara lain:

1. Kedua orang tua tercinta serta anggota keluarga, yang selalu memberikan doa, kasih sayang, dukungan, serta semangat yang tiada henti.
2. Dosen pembimbing, Ike Vrawati, S.Kom., M.Kom. yang dengan sabar memberikan bimbingan, masukan, serta arahan dari awal hingga skripsi ini terselesaikan.
3. Teman-teman seperjuangan di kelas IF01 yang telah memberikan semangat, bantuan, dan kebersamaan selama proses perkuliahan hingga penyusunan skripsi ini.

Semoga segala bantuan, dukungan, dan doa yang telah diberikan mendapatkan balasan yang setimpal dari Allah SWT.

KATA PENGANTAR

Alhamdulillah, segala puji atas kehadiran Allah SWT yang telah memberikan Rahmat dan Hidayah serta pertolongan kepada penulis sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pengaruh RandomizedSearchCV Pada Optimasi Random Forest Dan Xgboost Dalam Memprediksi Perilaku Konsumen Pada Platform E-Commerce” dengan sebaik-baiknya untuk memenuhi persyaratan untuk mencapai gelar Sarjana Komputer (S.Kom). Sholawat serta salam selalu tercurahkan kepada junjungan Nabi besar Muhammad SAW, keluarga, sahabat, yang senantiasa mengikutinya hingga akhir zaman. Dalam melakukan penyusunan skripsi ini penulis banyak mengalami hambatan, namun berkat bantuan dan bimbingan dari berbagai pihak akhirnya skripsi ini dapat terselesaikan dengan baik. Oleh karena itu penulis ingin menyampaikan rasa hormat serta rasa terima kasih kepada banyak pihak yang telah membantu penulis dalam memberikan dukungan, bimbingan terkhususnya kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr. Kusri, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom. selaku ketua program studi S1 Informatika.
4. Ibu Ike Vrawati, S.Kom., M.Kom. selaku dosen pembimbing skripsi.
5. Bapak dan Ibu dosen Universitas Amikom Yogyakarta yang telah mengajar dan memberikan ilmu pengetahuan selama di bangku perkuliahan. Dalam melakukan penulisan skripsi ini masih banyak kekurangan yang harus disempurnakan, oleh sebab itu penulis mengharapkan kritik dan saran yang membangun terhadap skripsi ini agar membantu jauh lebih baik.

Yogyakarta, 8 Agustus 2025

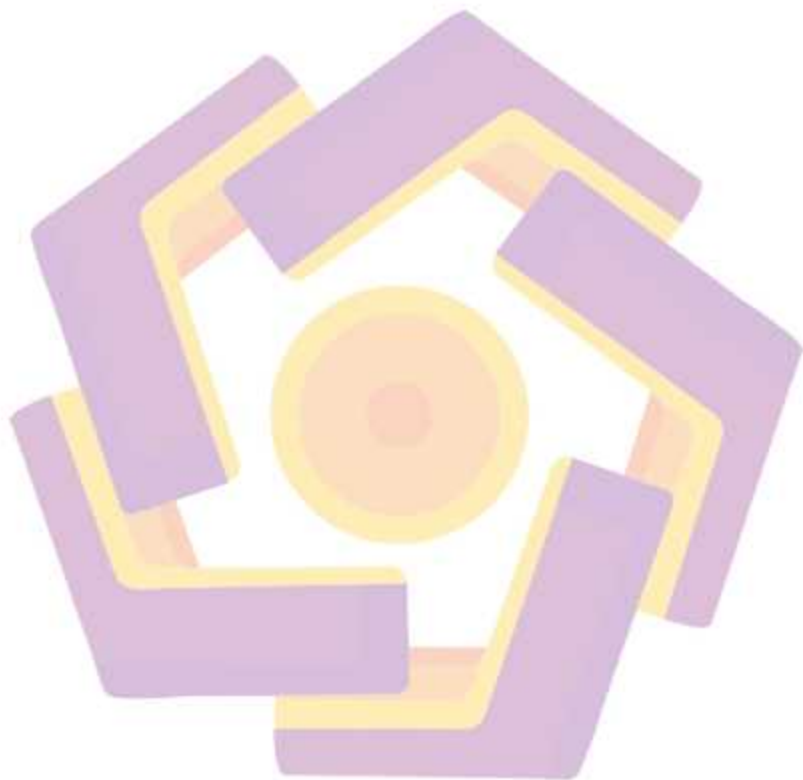
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6. Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1. Studi Literatur	4
2.2. Dasar Teori	10
2.2.1. E-Commerce	10
2.2.2. Perilaku Konsumen	10
2.2.3. Prediksi	10
2.2.4. Machine Learning	11
2.2.5. Data Preprocessing	11
2.2.6. SMOTE	11
2.2.7. Random Forest	12

2.2.8. XGBoost	14
2.2.9. RandomizedSearchCV	16
2.2.10. Evaluasi	16
BAB III METODE PENELITIAN	19
3.1. Objek Penelitian	19
3.2. Alur Penelitian	19
3.2.1. Pengumpulan Data	21
3.2.2. Eksplorasi Data	21
3.2.3. Preprocessing Data	21
3.2.4. Feature Selection	21
3.2.5. Data Splitting	21
3.2.6. Synthetic Minority Oversampling Technique (SMOTE)	22
3.2.7. Implementasi Algoritma	22
3.2.8. Hyperparameter Tuning	22
3.2.9. Evaluasi	22
3.3. Alat dan Bahan	22
3.3.1. Alat	22
3.3.2. Bahan	29
BAB IV HASIL DAN PEMBAHASAN	31
4.1. Pengumpulan Data	31
4.2. Eksplorasi Data	33
4.3. Data Preprocessing	35
4.4. Feature Selection	37
4.5. Data Splitting	38
4.6. Data Sampling	39
4.7. Pembuatan Model	41
4.8. Hyperparameter Tuning	42
4.9. Hasil Pelatihan Model	43
4.10. Evaluasi	56
BAB V PENUTUP	63
5.1. Kesimpulan	63

5.2. Saran	64
REFERENSI	65
LAMPIRAN	69



DAFTAR TABEL

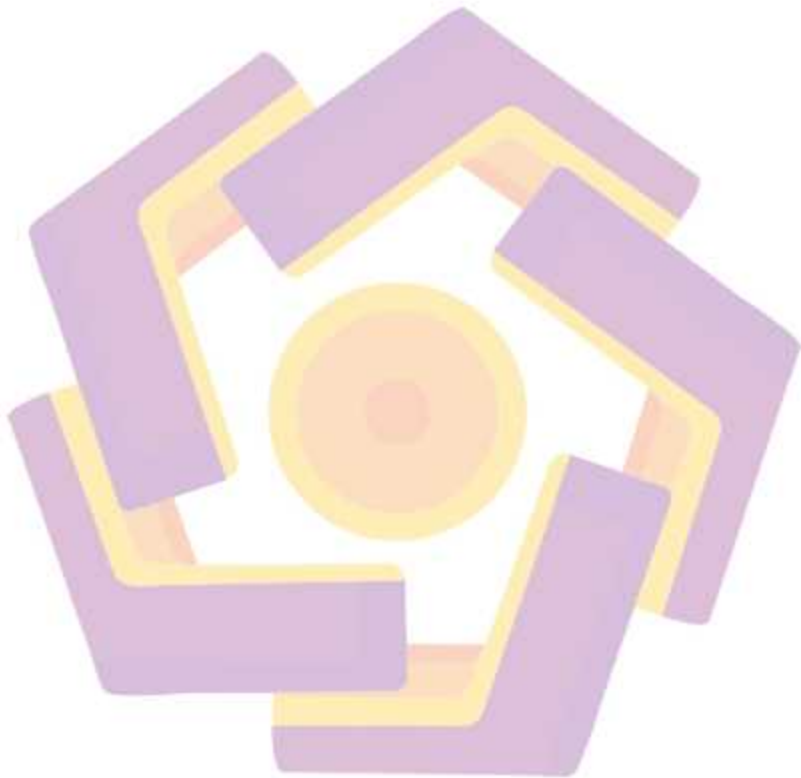
Tabel 2.1 Keaslian Penelitian	6
Tabel 2.2 Tabel Confusion Matrix	17
Tabel 3.1 Sampel dataset random forest	23
Tabel 3.2 Contoh data bootstrap	23
Tabel 3.3 Sampel dataset xgboost	26
Tabel 3.4 Nilai residual	26
Tabel 3.5 Sampel Dataset	29
Tabel 4.1 Deskripsi Dataset	31
Tabel 4.2 Jumlah sampel data splitting menggunakan RFECV	39
Tabel 4.3 Jumlah sampel data splitting tanpa RFECV	39
Tabel 4.4 Jenis Model Training	41
Tabel 4.5 Parameter untuk Random Forest	42
Tabel 4.6 Parameter untuk XGBoost	42
Tabel 4.7 Hasil evaluasi Model 1	43
Tabel 4.8 Hasil evaluasi Model 2	45
Tabel 4.9 Hasil evaluasi Model 3	47
Tabel 4.10 Hasil evaluasi Model 4	48
Tabel 4.11 Hasil evaluasi optimasi Model 1	50
Tabel 4.12 Hasil evaluasi optimasi Model 2	52
Tabel 4.13 Hasil evaluasi Model 3	53
Tabel 4.14 Hasil evaluasi optimasi Model 4	55
Tabel 4.15 Evaluasi Random Forest sebelum optimasi	56
Tabel 4.16 Evaluasi XGBoost sebelum optimasi	57
Tabel 4.17 Evaluasi Random Forest sesudah optimasi	58
Tabel 4.18 Evaluasi XGBoost sesudah optimasi	58

DAFTAR GAMBAR

Gambar 2.1 Alur kerja Random Forest.....	13
Gambar 3.1 Alur Penelitian	20
Gambar 4.1 Deskripsi Dataset	31
Gambar 4.2 Informasi Dataset	33
Gambar 4.3 Histogram data numerikal	34
Gambar 4.4 Histogram data kategorikal	34
Gambar 4.5 Boxplot data numerikal	35
Gambar 4.6 Boxplot data numerikal setelah outlier ditangani.....	36
Gambar 4.7 Data kategorikal sebelum dirubah.....	36
Gambar 4.8 Data kategorikal setelah dirubah.....	37
Gambar 4.9 Hasil feature selection RFECV	38
Gambar 4.10 Dataset Sebelum SMOTE	40
Gambar 4.11 Dataset Sesudah SMOTE.....	41
Gambar 4.12 Confusion Matrix Random Forest.....	44
Gambar 4.13 Confusion Matrix XGBoost	44
Gambar 4.14 Confusion Matrix Random Forest.....	45
Gambar 4.15 Confusion Matrix XGBoost	46
Gambar 4.16 Confusion Matrix Random Forest.....	47
Gambar 4.17 Confusion Matrix XGBoost	48
Gambar 4.18 Confusion Matrix Random Forest.....	49
Gambar 4.19 Confusion Matrix XGBoost	49
Gambar 4.20 Confusion Matrix Random Forest.....	51
Gambar 4.21 Confusion Matrix XGBoost	51
Gambar 4.22 Confusion Matrix Random Forest.....	52
Gambar 4.23 Confusion Matrix XGBoost	53
Gambar 4.24 Confusion Matrix Random Forest.....	54
Gambar 4.25 Confusion Matrix XGBoost	54
Gambar 4.26 Confusion Matrix Random Forest.....	55
Gambar 4.27 Confusion Matrix XGBoost	56
Gambar 4.28 Perbandingan F1-Score pada penggunaan SMOTE.....	60

Gambar 4.29 Perbandingan F1-Score pada penggunaan RFECV61

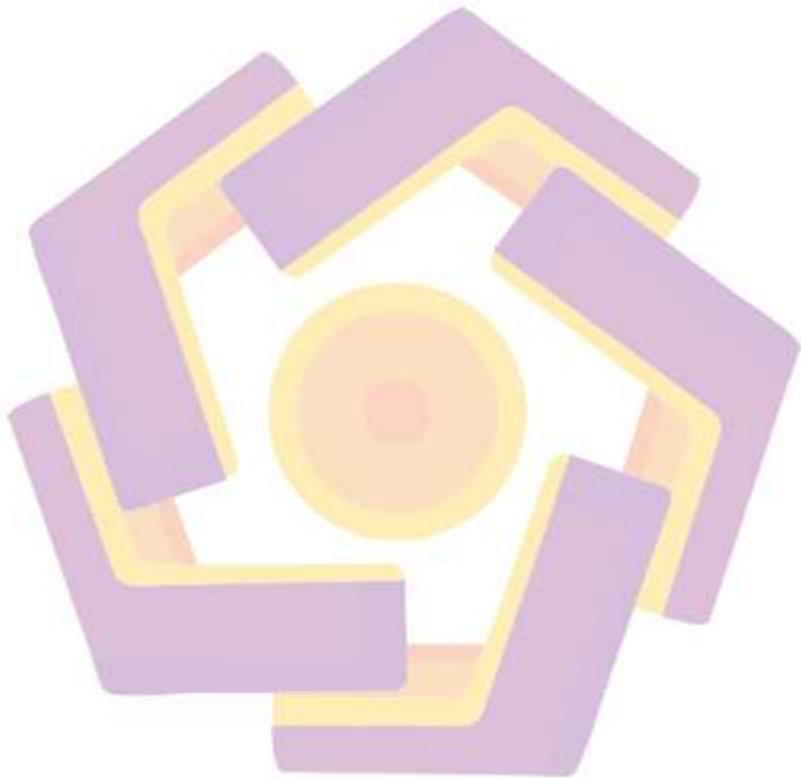
Gambar 4.30 Perbandingan F1-Score pada penggunaan SMOTE dan RFECV62



DAFTAR LAMPIRAN

Lampiran 1. Hasil deploy menggunakan streamlit cloud

66



INTISARI

Pertumbuhan e-commerce yang pesat menuntut kemampuan untuk memprediksi perilaku konsumen secara akurat guna meningkatkan strategi penawaran produk dan diskon. Machine learning menjadi salah satu solusi yang efektif untuk menganalisis data dalam skala besar. Penelitian ini memfokuskan pada perbandingan kinerja dua algoritma populer, yaitu Random Forest dan XGBoost, dalam memprediksi perilaku konsumen pada platform e-commerce. Untuk meningkatkan performa model, diterapkan teknik optimasi hyperparameter menggunakan metode *RandomizedSearchCV*. Tujuan penelitian ini adalah untuk mengevaluasi pengaruh optimasi terhadap kinerja kedua algoritma dan membandingkan performa keduanya. Eksperimen dilakukan pada dataset perilaku pengguna e-commerce yang memiliki karakteristik *imbalanced class*. Evaluasi kinerja model meliputi metrik akurasi, presisi, recall, dan *f1-score*. Hasil pengujian menunjukkan bahwa tidak terdapat peningkatan signifikan pada akurasi kedua algoritma setelah optimasi. Akurasi Random Forest sebelum dan sesudah optimasi berada pada kisaran 88%–90%, sedangkan XGBoost berkisar 87%–90%. Nilai presisi, recall, dan *f1-score* untuk kelas mayoritas (kelas 0) tetap tinggi, yaitu presisi 92%–96%, recall 90%–97%, dan *f1-score* 92%–94%. Sebaliknya, untuk kelas minoritas (kelas 1), nilai presisi hanya 57%–74%, recall 51%–78%, dan *f1-score* 60%–67%, tanpa peningkatan berarti setelah optimasi. Secara keseluruhan, Random Forest menunjukkan kinerja lebih unggul dibanding XGBoost, terutama pada prediksi kelas 1.

Kata kunci: E-commerce, Machine Learning, Random Forest, XGBoost, *RandomizedSearchCV*.

ABSTRACT

The rapid growth of e-commerce demands the ability to accurately predict consumer behavior to improve strategies for product offerings and discounts. Machine learning has emerged as an effective solution for analyzing large-scale data. This study focuses on comparing the performance of two popular algorithms, Random Forest and XGBoost, in predicting consumer behavior on e-commerce platforms. To enhance model performance, hyperparameter optimization using the RandomizedSearchCV method was applied. The objective of this research is to evaluate the impact of optimization on the performance of both algorithms and to compare their results. Experiments were conducted on an e-commerce user behavior dataset characterized by class imbalance. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The experimental results indicate that there was no significant improvement in the accuracy of either algorithm after optimization. The accuracy of Random Forest before and after optimization remained in the range of 88%–90%, while XGBoost ranged from 87%–90%. Precision, recall, and F1-score for the majority class (class 0) remained high, with precision between 92%–96%, recall between 90%–97%, and F1-score between 92%–94%. Conversely, for the minority class (class 1), precision was only 57%–74%, recall 51%–78%, and F1-score 60%–67%, with no meaningful improvement after optimization. Overall, Random Forest demonstrated superior performance compared to XGBoost, particularly in predicting class 1.

Keyword: E-commerce, Machine Learning, Random Forest, XGBoost, RandomizedSearchCV