

**ANALISIS ALGORITMA RANDOM FOREST PADA DETEKSI
KECURANGAN TRANSAKSI PERBANKAN DIGITAL**

SKRIPSI

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



disusun oleh

WILDAN ABRORI AZIZ

22.11.5186

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

ANALISIS ALGORITMA RANDOM FOREST PADA DETEKSI KECURANGAN TRANSAKSI PERBANKAN DIGITAL

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
WILDAN ABRORI AZIZ
22.11.5186

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS ALGORITMA RANDOM FOREST PADA DETEKSI
KECURANGAN TRANSAKSI PERBANKAN DIGITAL**

yang disusun dan diajukan oleh

WILDAN ABRORI AZIZ

22.11.5186

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 06 Mei 2026

Dosen Pembimbing,



Subektiningsih, S.Kom., M.Kom.
NIK. 190302413

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS ALGORITMA RANDOM FOREST PADA DETEKSI
KECURANGAN TRANSAKSI PERBANKAN DIGITAL

yang disusun dan diajukan oleh

WILDAN ABRORI AZIZ

22.11.5186

Telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Uyock Anggoro Saputro, S.Kom., M.Kom.
NIK. 190302419

Muhammad Tofa Nurcholis, S.Kom., M.Kom.
NIK. 190302281

Subektiningsih, S.Kom., M.Kom.
NIK. 190302413

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 20 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Wildan Abrori Aziz
NIM : 22.11.5186

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Algoritma Random Forest Pada Deteksi Kecurangan Transaksi Perbankan Digital

Dosen Pembimbing : Subektiningsih, 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, 20 Mei 2026

Yang Menyatakan,



Wildan Abrori Aziz

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT atas segala rahmat, nikmat, dan kemudahan yang telah diberikan, karya ini saya persembahkan kepada:

Kedua orang tua saya tercinta, yang selalu memberikan doa, dukungan, kasih sayang, serta semangat yang tidak pernah berhenti dalam setiap langkah yang saya jalani. Terima kasih atas segala pengorbanan, kesabaran, dan kepercayaan yang telah diberikan hingga saya dapat menyelesaikan skripsi ini.

Dosen pembimbing, yang telah memberikan arahan, bimbingan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini.

Teman-teman seperjuangan, yang telah memberikan bantuan, semangat, dan kebersamaan selama masa perkuliahan hingga penyusunan skripsi ini selesai.

Terakhir untuk diri saya sendiri, terima kasih telah bertahan, berusaha, dan tidak menyerah dalam menghadapi setiap proses, kesulitan, dan tantangan hingga akhirnya karya ini dapat diselesaikan. Semoga karya sederhana ini dapat menjadi bentuk kecil dari rasa terima kasih saya kepada semua pihak yang telah mendukung saya hingga sampai pada tahap ini.

KATA PENGANTAR

Puji Syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala berkat, rahmat, hidayah dan limpahan karunia-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul **“Analisis Algoritma Random Forest Pada Deteksi Kecurangan Transaksi Perbankan Digital”**. Skripsi ini ditulis untuk memenuhi persyaratan untuk mencapai gelar Sarjana Komputer dalam Program Studi Informatika Fakultas Ilmu Komputer.

Skripsi ini dapat diselesaikan atas doa, dukungan bantuan serta bimbingan dari berbagai pihak, baik bantuan berupa moril maupun materil. Dalam kesempatan ini dengan segala kerendahan hati, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Kedua Orang Tua serta seluruh keluarga yang tidak pernah berhenti mendoakan dan memberikan dukungannya selama ini.
2. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas AMIKOM Yogyakarta.
3. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
4. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
5. Ibu Subektiningsih, M.Kom., selaku dosen pembimbing yang telah memberikan waktu, saran, kritik, motivasi dan bimbingan dalam skripsi ini sehingga skripsi ini dapat selesai dengan baik.
6. Seluruh Staff Pengajar Khususnya Dosen yang mengajar di Program Studi Informatika, Universitas Amikom Yogyakarta yang telah memberikan bantuan, masukan, dan informasi serta ilmunya kepada penulis selama mengikuti perkuliahan.
7. Terakhir, diri sendiri yang sudah bertahan selama ini.

Yogyakarta, 20 Mei 2026

Wildan Abrori Aziz

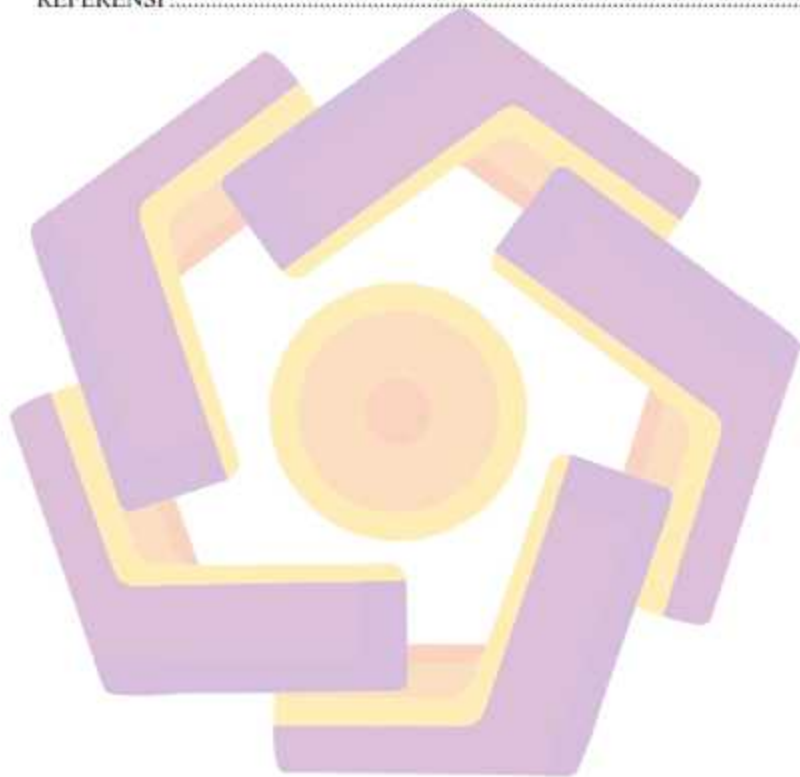
DAFTAR ISI

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.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xv
INTISARI.....	xviii
ABSTRACT.....	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5

2.2	Dasar Teori.....	10
2.2.1	Machine Learning	10
2.2.2	Transaksi Perbankan Digital.....	10
2.2.3	Validitas Transaksi Digital.....	11
2.2.4	Algoritma Random Forest.....	12
2.2.5	Ketidakseimbangan Data dan Metode SMOTE.....	14
2.2.6	Principal Component Analysis (PCA).....	14
2.2.7	Data Pribadi dan Privasi.....	16
2.2.8	Evaluasi Model	16
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian.....	19
3.2	Alur Penelitian	20
3.2.1	Pengumpulan Data	21
3.2.2	Subsampling Non-fraud.....	21
3.2.3	Seleksi dan Prapemrosesan Data.....	22
3.2.4	StandardScaler	22
3.2.5	Train-test Split	23
3.2.6	SMOTE	23
3.2.7	GridSearchCV.....	24
3.2.8	Random Forest	24
3.2.9	Evaluasi Model	24
3.2.10	Pengumpulan Data Dummy	25
3.2.11	Simpan Model dan Data Pendukung.....	25
3.2.12	Integrasi Model dan Sistem.....	26
3.2.13	Tampilan dan Hasil Prediksi.....	26

3.2.14	Pengujian Aplikasi	27
3.3	Alat dan Bahan	27
3.3.1	Perangkat Keras	27
3.3.2	Perangkat Lunak	28
BAB IV HASIL DAN PEMBAHASAN		29
4.1	Subsampling Non-fraud	29
4.2	Seleksi dan Prapemrosesan Data	30
4.3	StandardScaler	31
4.4	Train-test Split	32
4.5	SMOTE	33
4.6	GridSearchCV	35
4.7	Random Forest	37
4.8	Evaluasi Model	38
4.8.1	Confusion Matrix	39
4.8.2	Metrik Evaluasi	40
4.8.3	Analisis Kinerja Model	42
4.9	Integrasi Model dan Sistem	43
4.9.1	Implementasi Backend	43
4.9.2	Penetapan Threshold Klasifikasi	45
4.9.3	Arsitektur Pipeline Prediksi	45
4.9.4	Rule-Based Layer pada Mode Input Data Mentah	47
4.10	Implementasi Antarmuka Pengguna	49
4.10.1	Halaman Utama	49
4.10.2	Halaman Pengecekan Transaksi	50
4.10.3	Hasil Prediksi Sistem	52

4.11 Pengujian Sistem.....	53
BAB V PENUTUP	55
5.1 Kesimpulan	55
5.2 Saran	55
REFERENSI	58



DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 3.1 Daftar Perangkat Keras	29
Tabel 3.2 Daftar Perangkat Lunak	29
Tabel 4.1 Distribusi Data Sebelum dan Sesudah Subsampling	30
Tabel 4.2 Distribusi Data Latih Sebelum dan Sesudah SMOTE	35
Tabel 4.3 Pencarian Parameter GridSearchCV	37
Tabel 4.4 Hasil Pencarian Hyperparameter GridSearchCV	37
Tabel 4.5 Confusion Matrix Hasil Evaluasi Model	40
Tabel 4.6 Hasil Metrik Evaluasi Model	42
Tabel 4.7 Daftar Endpoint API Sistem	45
Tabel 4.8 Perbandingan Pipeline Prediksi Per Mode Input	47
Tabel 4.9 Aturan Rule-Based Layer Berdasarkan Nominal Transaksi	48



DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	21
Gambar 4.1 Kode Subsampling	30
Gambar 4.2 Kode Seleksi dan Prapemrosesan Data	31
Gambar 4.3 Kode StandardScaler	32
Gambar 4.4 Kode Train-test Split	33
Gambar 4.5 Kode SMOTE	34
Gambar 4.6 Kode GridSearchCV	36
Gambar 4.7 Kode Random Forest	38
Gambar 4.8 Visualisasi Heatmap Confusion Matrix	40
Gambar 4.9 Classification Report	41
Gambar 4.10 Visualisasi Kurva AUC-ROC	42
Gambar 4.11 Kode Load Model	45
Gambar 4.12 Kode Endpoint predict_pca_csv	47
Gambar 4.13 kode fungsi adjust_risk	49
Gambar 4.14 Tampilan Halaman Utama	51
Gambar 4.15 Tampilan Halaman Pengecekan Transaksi	52
Gambar 4.16 Tampilan Formulir Input Manual	53
Gambar 4.17 Tampilan Hasil Prediksi	54
Gambar 4.18 Hasil Pengujian Upload PCA CSV dengan Data Bervariasi	55
Gambar 4.19 Output Jika Data yang Diinput Tidak Sesuai	55

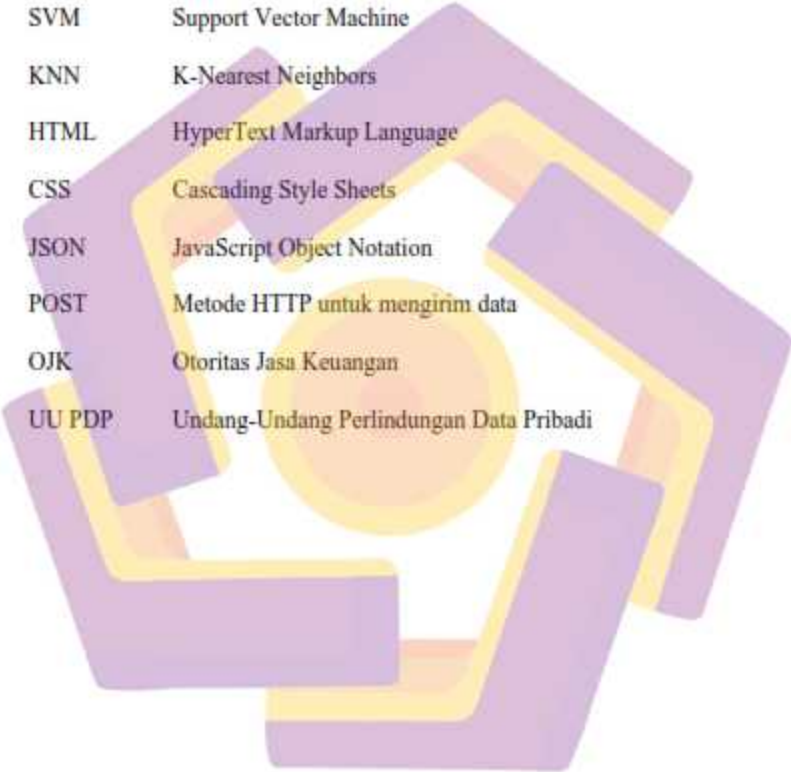
DAFTAR LAMBANG DAN SINGKATAN

Lambang

n	Jumlah observasi / data
μ	Rata-rata (mean)
σ	Simpangan baku (standard deviation)
λ	Eigenvalue pada PCA
v	Vektor eigen (eigenvector)
X	Matriks data awal
Z	Data hasil transformasi PCA
W	Matriks eigenvector (proyeksi PCA)
T	Jumlah pohon pada Random Forest

Singkatan

TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
F1	F1-Score
AUC	Area Under the Curve
ROC	Receiver Operating Characteristic
RF	Random Forest
SMOTE	Synthetic Minority Over-sampling Technique



PCA	Principal Component Analysis
API	Application Programming Interface
CSV	Comma-Separated Values
ML	Machine Learning
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
HTML	HyperText Markup Language
CSS	Cascading Style Sheets
JSON	JavaScript Object Notation
POST	Metode HTTP untuk mengirim data
OJK	Otoritas Jasa Keuangan
UU PDP	Undang-Undang Perlindungan Data Pribadi

DAFTAR ISTILAH



Accuracy	Metrik evaluasi yang mengukur proporsi prediksi benar dari seluruh data uji.
Algorithm	Serangkaian langkah terstruktur yang digunakan untuk menyelesaikan suatu masalah.
Anomali	Pola data yang menyimpang dari distribusi normal.
Antarmuka Pengguna	Bagian sistem yang berinteraksi langsung dengan pengguna.
Backend	Lapisan sistem yang berjalan di sisi server dan mengelola logika aplikasi.
Bootstrap Sampling	Teknik pengambilan sampel acak dengan pengembalian.
Class Imbalance	Kondisi ketidakseimbangan jumlah data antar kelas.
Confusion Matrix	Tabel evaluasi yang membandingkan hasil prediksi dengan data sebenarnya.
Cross-Validation	Teknik evaluasi model dengan membagi data menjadi beberapa bagian.
Dataset	Kumpulan data yang digunakan untuk pelatihan dan pengujian model.

Decision Tree

Algoritma berbasis pohon untuk klasifikasi.

Eigenvalue

Nilai yang menunjukkan besarnya variansi dalam PCA.

Eigenvector

Vektor arah dalam transformasi PCA.

Ensemble Learning

Metode yang menggabungkan beberapa model.

Feature

Variabel input dalam model pembelajaran mesin.

F1-Score

Rata-rata harmonis antara precision dan recall.

False Positive

Data normal yang salah diprediksi sebagai fraud.

False-Negative

Data fraud yang tidak terdeteksi.

GridSearchCV

Metode pencarian parameter terbaik.

Hyperparameter

Parameter yang ditentukan sebelum pelatihan.

Inferensi

Proses prediksi menggunakan model.

Machine Learning

Metode pembelajaran dari data.

Model

Representasi matematis untuk prediksi.

Normalisasi

Proses penyesuaian skala data.

Overfitting

Model terlalu menyesuaikan data latih.

Oversampling

Teknik menambah data pada kelas minoritas.

Pipeline

Rangkaian proses pengolahan data.

Precision

Ketepatan prediksi positif.

Probabilitas

Tingkat keyakinan model terhadap prediksi.

Random Forest

Algoritma berbasis banyak decision tree.

Recall

Kemampuan mendeteksi seluruh data positif.

Reduksi Dimensi

Proses mengurangi jumlah fitur.

ROC Curve

Grafik hubungan TPR dan FPR.

Scaler

Alat untuk normalisasi data.

StandardScaler

Metode normalisasi berbasis z-score.

Supervised Learning

Pembelajaran dengan data berlabel.

Threshold

Nilai batas untuk klasifikasi.

Train-Test Split

Pembagian data latih dan uji.

Transformasi PCA

Proses reduksi dimensi menggunakan PCA.

INTISARI

Pertumbuhan transaksi perbankan digital di Indonesia yang mencapai Rp87 kuadriliun mendorong meningkatnya risiko penipuan (fraud) yang merugikan nasabah maupun institusi keuangan. Kerugian akibat scam dan fraud perbankan telah menembus angka triliunan rupiah, sementara sistem deteksi konvensional kerap gagal mengimbangi pola kecurangan yang terus berkembang. Kondisi ini menuntut pendekatan berbasis kecerdasan buatan yang mampu mengenali pola anomali transaksi secara otomatis, cepat, dan akurat.

Penelitian ini menerapkan algoritma Random Forest pada dataset Credit Card Fraud Detection yang memuat 284.807 rekaman transaksi dengan rasio ketidakseimbangan kelas 19:1. Untuk mengatasi ketimpangan tersebut, digunakan teknik Synthetic Minority Over-sampling Technique (SMOTE) guna menyeimbangkan data latih sebelum model dibangun. Konfigurasi model optimal diperoleh melalui GridSearchCV dengan metrik F1-score dan validasi silang tiga lipatan, menghasilkan kombinasi terbaik $n_estimators=100$, max_depth tanpa pembatasan, dan $class_weight$ balanced. Nilai threshold klasifikasi diturunkan menjadi 0,25 untuk memprioritaskan kemampuan deteksi fraud.

Model yang dihasilkan mencapai akurasi 99%, precision kelas fraud 94%, recall 84%, F1-score 89%, dan AUC-ROC 0,97 pada 1.999 data uji. Penelitian ini juga menghasilkan sistem aplikasi berbasis web yang mengintegrasikan model ke dalam antarmuka tiga mode input: unggah CSV berformat PCA, unggah CSV data mentah, dan pengisian formulir manual. Hasil penelitian ini dapat dimanfaatkan oleh analis risiko perbankan, tim keamanan siber institusi keuangan, serta pengembang sistem deteksi kecurangan yang membutuhkan solusi berbasis machine learning yang terukur dan dapat dioperasikan langsung.

Kata kunci: Perbankan digital, pembelajaran mesin, Random Forest, deteksi fraud, SMOTE.

ABSTRACT

The rapid growth of digital banking transactions in Indonesia, reaching Rp87 quadrillion, has heightened the risk of fraud that harms both customers and financial institutions. Losses from banking scams and fraud have surpassed trillions of rupiah, while conventional detection systems frequently fail to keep pace with increasingly sophisticated fraud patterns. This situation calls for an artificial intelligence-based approach capable of automatically, swiftly, and accurately identifying anomalous transaction patterns.

This study applies the Random Forest algorithm to the Credit Card Fraud Detection dataset, comprising 284,807 transaction records with a class imbalance ratio of 19:1. To address this imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was employed to balance the training data prior to model construction. The optimal model configuration was determined through GridSearchCV using F1-score as the evaluation metric with three-fold cross-validation, yielding the best-performing combination of $n_{\text{estimators}}=100$, unbounded max_depth , and balanced class_weight . The classification threshold was lowered to 0.25 to prioritize fraud detection sensitivity.

The resulting model achieved 99% accuracy, 94% precision for the fraud class, 84% recall, 89% F1-score, and an AUC-ROC of 0.97 on 1,999 test samples. This research also produced a web-based application integrating the model into a three-mode input interface: PCA-formatted CSV upload, raw CSV upload, and a manual entry form. The findings of this study can be utilized by banking risk analysts, cybersecurity teams at financial institutions, and fraud detection system developers seeking a scalable, machine learning-based solution that can be deployed directly in operational environments.

Keyword: Digital banking, machine learning, Random Forest, fraud detection, SMOTE.