



**PERBANDINGAN TEKNIK SAMPLING SMOTE, ADASYN,
DAN SMOTE-TOMEK PADA DETEKSI FRAUD TRANSAKSI
KEUANGAN MENGGUNAKAN CATBOOST DAN XGBOOST**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

BIMA AJI NUGRAHA

NIM 2211016210003

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JUNI 2026

SKRIPSI

PERBANDINGAN TEKNIK SAMPLING SMOTE, ADASYN, DAN SMOTE-TOMEK PADA DETEKSI FRAUD TRANSAKSI KEUANGAN MENGUNAKAN CATBOOST DAN XGBOOST

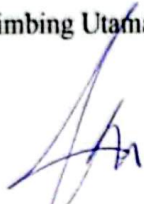
Oleh:

BIMA AJI NUGRAHA

NIM. 2211016210003

Telah dipertahankan di depan Dosen Penguji pada tanggal 10 Juni 2026.
Susunan Penguji:

Pembimbing Utama



Muhammad Itqan Mazdadi, S.Kom., M.Kom
NIP. 199006122019031013

Penguji

1.



Fatma Indriani, S.T., M.I.T., Ph.D
NIP. 198404202008122004

Pembimbing Pendamping



Triando Hamonangan Saragih, S.Kom., M.Kom
NIP. 199308242019031012

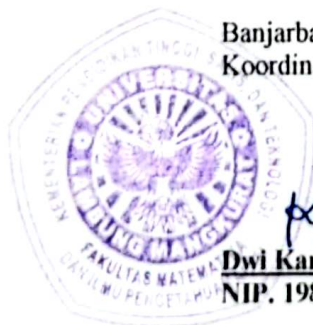
2.



Andri Farmadi, S.Si., M.T
NIP. 197307252008011006

Banjarbaru, 25 Juni 2026

Koordinator Program Studi Ilmu Komputer



Dwi Kartini, S.Kom., M.Kom
NIP. 198704212012122003

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 10 Juni 2026

Yang Menyatakan,



Bima Aji Nugraha

NIM. 2211016210003

ABSTRAK

PERBANDINGAN TEKNIK SAMPLING SMOTE, ADASYN, DAN SMOTE-TOMEK PADA DETEKSI FRAUD TRANSAKSI KEUANGAN MENGGUNAKAN CATBOOST DAN XGBOOST

(Oleh: Bima Aji Nugraha; Pembimbing: Muhammad Itqan Mazdadi, S.Kom., M.Kom. dan Triando Hamonangan Saragih, S.Kom., M.Kom.; 2026; 73 halaman)

Meningkatnya penggunaan sistem pembayaran digital menyebabkan volume transaksi keuangan terus bertambah, sehingga risiko terjadinya *fraud* semakin tinggi. Salah satu tantangan utama dalam deteksi *fraud* adalah ketidakseimbangan kelas, yaitu jumlah transaksi *fraud* yang jauh lebih sedikit dibandingkan transaksi *legitimate*. Kondisi ini dapat menyebabkan model klasifikasi bias terhadap kelas mayoritas dan kurang optimal dalam mengenali transaksi *fraud*. Penelitian ini menggunakan tiga teknik penyeimbangan data, yaitu *Synthetic Minority Oversampling Technique* (SMOTE), *Adaptive Synthetic Sampling* (ADASYN), dan SMOTE-Tomek, yang dikombinasikan dengan algoritma CatBoost dan XGBoost. *Dataset* yang digunakan adalah *Credit Card Fraud Detection Dataset* dengan tahapan *preprocessing*, seleksi fitur, optimasi *hyperparameter* menggunakan Optuna, serta optimasi *threshold* keputusan menggunakan data validasi. Hasil penelitian menunjukkan bahwa teknik *sampling* mampu meningkatkan kemampuan model dalam mendeteksi *fraud*, terutama pada metrik *recall* saat menggunakan *threshold default* 0,5. Setelah optimasi *threshold*, kombinasi SMOTE dan CatBoost menghasilkan performa terbaik pada skenario *sampling* dengan *precision* sebesar 0,8859, *recall* sebesar 0,7455, *F1-score* sebesar 0,8096, dan *threshold* optimal sebesar 0,9407. Nilai *recall* tertinggi diperoleh oleh ADASYN CatBoost sebesar 0,7650.

Kata Kunci: deteksi *fraud*, ketidakseimbangan kelas, *sampling*, CatBoost, XGBoost.

ABSTRACT

COMPARISON OF SMOTE, ADASYN, AND SMOTE-TOMEK SAMPLING TECHNIQUES FOR FINANCIAL TRANSACTION FRAUD DETECTION USING CATBOOST AND XGBOOST

(By: Bima Aji Nugraha; Supervisor: Muhammad Itqan Mazdadi, S.Kom., M.Kom. and Triando Hamonangan Saragih, S.Kom., M.Kom.; 2026; 73 pages)

The increasing use of digital payment systems has led to continuous growth in financial transaction volume, which also increases the risk of transaction fraud. One of the main challenges in fraud detection is class imbalance, where fraudulent transactions are far fewer than legitimate transactions. This condition may cause classification models to become biased toward the majority class and reduce their ability to identify fraudulent transactions. This study applies three data balancing techniques, namely Synthetic Minority Oversampling Technique (SMOTE), Adaptive Synthetic Sampling (ADASYN), and SMOTE-Tomek, combined with CatBoost and XGBoost algorithms. The Dataset used in this study is the Credit Card Fraud Detection Dataset, with research stages including data preprocessing, feature selection, hyperparameter optimization using Optuna, and decision threshold optimization using Validation data. The results show that sampling techniques improve the model's ability to detect fraud, particularly in terms of recall when using the default threshold of 0.5. After threshold optimization, the combination of SMOTE and CatBoost achieved the best performance in the sampling scenario, with a precision of 0.8859, recall of 0.7455, F1-score of 0.8096, and an optimal threshold of 0.9407. The highest recall was achieved by ADASYN CatBoost at 0.7650.

Keywords: *fraud detection, class imbalance, sampling, CatBoost, XGBoost*

PRAKATA

Puji syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Perbandingan Teknik Sampling SMOTE, ADASYN, dan SMOTE-Tomek pada Deteksi Fraud Transaksi Keuangan Menggunakan CatBoost dan XGBoost". Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan S1 pada Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Selawat dan salam senantiasa tercurah kepada Nabi Muhammad SAW., beserta keluarga, sahabat, dan pengikut beliau hingga akhir zaman.

Penulis menyampaikan terima kasih kepada:

1. Keluarga tercinta yang selalu memberikan doa, semangat, dukungan, dan kepercayaan kepada penulis.
2. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku dosen pembimbing utama dan Bapak Triando Hamonangan Saragih, S.Kom., M.Kom. selaku dosen pembimbing pendamping.
3. Bapak Radityo Adi Nugroho, S.T., M.Kom. selaku Dosen Pembimbing Akademik dan Ibu Dwi Kartini, S.Kom., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM.
4. Seluruh dosen dan staf Program Studi Ilmu Komputer FMIPA ULM.
5. Semua pihak yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan perkembangan ilmu pengetahuan.

Banjarbaru, 10 Juni 2026



Bima Aji Nugraha

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
PRAKATA	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kajian Terdahulu.....	6
2.2 Landasan Teori	10
2.2.1 Audit Transaksi Keuangan	10
2.2.2 <i>Fraud</i> Transaksi	11
2.2.3 Optuna	11
2.2.4 <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	12

2.2.5	<i>Adaptive Synthetic Sampling (ADASYN)</i>	13
2.2.6	SMOTE-Tomek.....	14
2.2.7	Categorical <i>Boosting</i> (CatBoost).....	15
2.2.8	Extreme Gradient <i>Boosting</i> (XGBoost)	17
2.2.9	Metrik Evaluasi	19
2.2.10	<i>Threshold Tuning</i>	21
BAB III METODE PENELITIAN		22
3.1	Alat Penelitian	22
3.2	Bahan Penelitian.....	22
3.3	Variabel Penelitian	22
3.4	Prosedur Penelitian.....	23
BAB IV HASIL DAN PEMBAHASAN.....		29
4.1	Hasil	29
4.1.1	Pengumpulan <i>Dataset</i>	29
4.1.2	<i>Preprocessing</i>	31
4.1.3	Seleksi Fitur.....	34
4.1.4	Penanganan Ketidakseimbangan Kelas (<i>Sampling</i>).....	36
4.1.5	<i>Hyperparameter Tuning</i> dengan Optuna.....	39
4.1.6	Klasifikasi.....	44
4.1.7	Optimasi <i>Threshold</i> Keputusan	49
4.2	Pembahasan.....	54
4.2.1	Pengaruh Teknik <i>Sampling</i> SMOTE.....	54
4.2.2	Pengaruh Teknik <i>Sampling</i> ADASYN.....	58
4.2.3	Pengaruh Teknik <i>Sampling</i> SMOTE-Tomek	63
4.2.4	Perbandingan Ketiga Teknik <i>Sampling</i>	68

BAB V PENUTUP.....	72
5.1 Kesimpulan.....	72
5.2 Saran.....	72
DAFTAR PUSTAKA	74
LAMPIRAN.....	77
RIWAYAT PENULIS	102

DAFTAR TABEL

Tabel 1. Keaslian Penelitian.....	8
Tabel 2. Perancangan Penelitian	9
Tabel 3. Daftar Fitur pada <i>Dataset</i>	30
Tabel 4. Contoh Hasil Ekstraksi Fitur Waktu	32
Tabel 5. Contoh Pembentukan Fitur Usia	33
Tabel 6. Contoh Pembentukan Fitur Jarak Lokasi Transaksi	33
Tabel 7. Contoh Transformasi Fitur <i>Category</i>	34
Tabel 8. Fitur Akhir yang Digunakan dalam Model	35
Tabel 9. Distribusi Data Setelah <i>Stratified Hold-Out</i> 80:20	36
Tabel 10. Distribusi Data <i>Train Inner</i> Setelah <i>Sampling</i>	37
Tabel 11. Contoh data <i>sampling</i>	38
Tabel 12. Ruang Pencarian <i>Hyperparameter</i> CatBoost dengan Optuna.....	40
Tabel 13. Ruang Pencarian <i>Hyperparameter</i> XGBoost dengan Optuna	40
Tabel 14. <i>Hyperparameter</i> Terbaik CatBoost Optimasi Optuna	43
Tabel 15. <i>Hyperparameter</i> Terbaik XGBoost Hasil Optimasi Optuna.....	43
Tabel 16. Hasil Klasifikasi <i>Threshold</i> 0.5.....	44
Tabel 17. <i>Confusion Matrix</i> pada <i>Threshold</i> 0,5.....	45
Tabel 18. Hasil Klasifikasi dengan <i>Threshold</i> Optimal	49
Tabel 19. <i>Confusion Matrix</i> pada <i>Threshold</i> Optimal	50

DAFTAR GAMBAR

Gambar 1. Pohon CatBoost dari <i>Dataset</i> Iris.....	15
Gambar 2. Pohon XGBoost dari <i>Dataset</i> Iris.....	17
Gambar 3. Prosedur penelitian.....	23
Gambar 4. Grafik Optuna CatBoost.....	41
Gambar 5. Grafik Optuna XGBoost.....	42
Gambar 6. Komposisi Prediksi <i>Fraud</i> pada <i>Threshold</i> 0,5.....	48
Gambar 7. Contoh Grafik Penentuan <i>Threshold</i> Optimal.....	50
Gambar 8. Komposisi Prediksi <i>Fraud</i> pada <i>Threshold</i> Optimal	53
Gambar 9. Perbandingan Hasil Kombinasi CatBoost SMOTE	55
Gambar 10. Perbandingan Hasil Kombinasi XGBoost SMOTE	55
Gambar 11. Perbandingan Hasil Kombinasi CatBoost ADASYN	59
Gambar 12. Perbandingan Hasil Kombinasi XGBoost ADASYN	59
Gambar 13. Perbandingan Hasil Kombinasi CatBoost SMOTE-Tomek.....	64
Gambar 14. Perbandingan Hasil Kombinasi XGBoost SMOTE-Tomek.....	64
Gambar 15. Perbandingan Seluruh <i>F1-score</i>	68