

MITIGASI KETIDAKSEIMBANGAN KELAS EKSTREM PADA DETEKSI PENIPUAN KARTU KREDIT MENGGUNAKAN PENDEKATAN HYBRID OVERSAMPLING DAN ENSEMBLE LEARNING

Zakaria¹, Dian Ade Kurnia², Yudhistira Arie Wijaya³, Rosidin⁴.

Program Studi Teknik Informatika¹
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi^{3,4}

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
zakaria.bajrie2509@gmail.com

(*) Corresponding Author : zakaria.bajrie2509@gmail.com
Published : 30 Januari 2026

Abstract—This study addresses the challenge of extreme class imbalance in credit card transaction data, which often causes fraud detection models to favour legitimate transactions and miss rare fraudulent cases. We investigate a hybrid oversampling approach using SMOTE-ENN combined with the Random Forest ensemble classifier on a large public credit card fraud dataset from Kaggle. The research pipeline covers exploratory data analysis, feature transformation, stratified train-test splitting, and the application of SMOTE-ENN strictly on the training set to avoid data leakage. Model performance is evaluated using metrics recommended for imbalanced classification, including class-wise precision, recall, F1-score, macro-averaged scores, and ROC-AUC, complemented with confusion matrix analysis. Experimental results show that the proposed pipeline preserves perfect performance on the majority (normal) class while substantially improving the minority (fraud) class, achieving precision of 0.79, recall of 0.83, and F1-score of 0.81 for fraud, with an overall accuracy of 1.00 and a macro F1-score of 0.90. These findings indicate that the combination of SMOTE-ENN and Random Forest is effective for mitigating class imbalance and enhancing the reliability of credit card fraud detection systems, and can serve as a reproducible baseline for future operational deployments. The implementation is carried out using Python, scikit-learn, and imbalanced-learn.

Keywords: credit card fraud detection, class imbalance, SMOTE-ENN, Random Forest, imbalanced classification.

Abstrak—Ketidakseimbangan kelas yang ekstrem pada data transaksi kartu kredit menyebabkan model deteksi penipuan cenderung bias terhadap transaksi normal sehingga transaksi fraud sulit terdeteksi secara andal. Penelitian ini bertujuan memitigasi masalah tersebut dengan menggabungkan teknik hybrid oversampling SMOTE-ENN dan algoritma ensemble Random Forest pada dataset penipuan kartu kredit berskala besar. Dataset publik kredit kartu yang digunakan memiliki proporsi kelas yang sangat timpang antara transaksi normal dan fraud. Pra-pemrosesan dilakukan melalui eksplorasi data, transformasi fitur, pembagian data secara stratified, dan penerapan SMOTE-ENN hanya pada data latih untuk menghindari data leakage. Model Random Forest kemudian dilatih pada data hasil resampling dan dievaluasi menggunakan metrik yang relevan untuk data tidak seimbang, yaitu precision, recall, F1-score, PR-based metrics, serta ROC-AUC. Hasil eksperimen menunjukkan bahwa kombinasi SMOTE-ENN dan Random Forest mampu mempertahankan kinerja sempurna pada kelas normal sekaligus meningkatkan kemampuan deteksi kelas fraud dengan nilai precision 0,79, recall 0,83, dan F1-score 0,81, disertai ROC-AUC yang sangat tinggi. Temuan ini mengindikasikan bahwa pendekatan hybrid oversampling dan ensemble learning efektif sebagai fondasi pengembangan pipeline deteksi penipuan kartu kredit yang lebih akurat dan siap diterapkan secara operasional.

Kata Kunci: penipuan kartu kredit, ketidakseimbangan kelas, SMOTE-ENN, Random Forest, deteksi penipuan berbasis machine learning.

INTRODUCTION

Kemajuan teknologi finansial dan meningkatnya transaksi daring telah membawa kemudahan

signifikan bagi pengguna, namun juga membuka peluang baru bagi pelaku kejahatan siber untuk melakukan penipuan kartu kredit. Deteksi penipuan kartu kredit menggunakan pembelajaran mesin (*machine learning*) menjadi fokus penelitian utama selama dekade terakhir, karena sistem tradisional berbasis aturan sering gagal mengikuti pola penipuan yang berubah cepat [1]. Meski demikian, pendekatan berbasis *machine learning* masih menghadapi sejumlah tantangan mendasar yang memengaruhi kinerja dan reliabilitasnya dalam skenario dunia nyata.

Salah satu permasalahan utama adalah ketidakseimbangan kelas yang ekstrem (*extreme class imbalance*), di mana transaksi sah jauh lebih banyak dibanding transaksi penipuan. Ketidakseimbangan ini menyebabkan model cenderung bias terhadap kelas mayoritas, menghasilkan akurasi dan *ROC-AUC* yang tampak tinggi namun menutupi performa deteksi kelas minoritas yang buruk [2], [3]. Dalam kondisi ini, metrik evaluasi tradisional seperti akurasi tidak lagi relevan; diperlukan metrik yang lebih representatif seperti *Precision*, *Recall*, *F1-score*, dan *Precision-Recall AUC (PR-AUC)* [4], [5]. Penilaian berbasis *PR-AUC* dan pendekatan pembobotan biaya (*cost-sensitive evaluation*) menjadi penting karena kesalahan mendeteksi penipuan (*false negative*) dapat menimbulkan kerugian finansial yang besar bagi institusi keuangan.

Selain itu, data transaksi kartu kredit sering kali bersifat dinamis dan mengalami *concept drift*, di mana pola penipuan, perilaku pengguna, dan jenis transaksi berubah seiring waktu [1], [6]. Model deteksi yang statis akan mengalami degradasi performa apabila tidak disertai dengan mekanisme pemantauan *drift* dan pembaruan berkala. Tantangan lain adalah karakteristik data yang berdimensi tinggi, bersifat *sparsely encoded*, serta keterbatasan akses terhadap dataset publik yang representatif [7]. Dataset yang digunakan oleh industri umumnya bersifat tertutup, sementara label fraud sering kali tertunda (*delayed*) atau bahkan mengandung *noise*, sehingga menyulitkan proses pelatihan model secara terawasi (*supervised learning*) dan menghambat reproduksibilitas hasil penelitian [1].

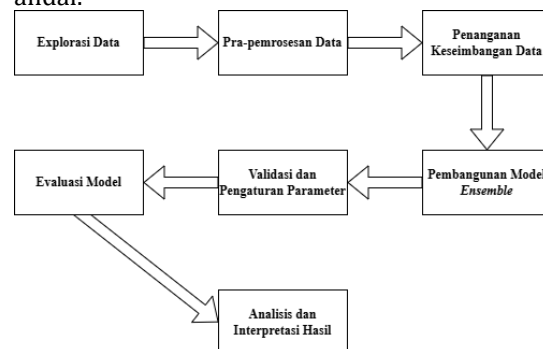
Dari sisi operasional, persyaratan *real-time* dan skalabilitas menambah kompleksitas penerapan model deteksi penipuan. Sistem harus mampu memproses jutaan transaksi per detik dengan latensi minimum tanpa mengorbankan akurasi. Selain itu, isu *explainability* dan auditabilitas regulatif juga

menjadi perhatian penting agar sistem deteksi penipuan dapat memberikan alasan yang transparan dalam pengambilan keputusan terutama dalam konteks kebijakan *adverse action* di lembaga keuangan [1], [3].

Berdasarkan tantangan tersebut, pendekatan yang menggabungkan *hybrid oversampling* dan *ensemble learning* menjadi semakin relevan. Teknik *hybrid* seperti kombinasi *SMOTE*, *Tomek Links*, dan *SMOTEENN* memungkinkan pembentukan data yang lebih seimbang tanpa kehilangan karakteristik penting kelas minoritas, sementara *ensemble learning* seperti *Random Forest* dan *XGBoost* mampu meningkatkan generalisasi serta mengurangi varians model. Integrasi kedua pendekatan ini dapat memperkuat kemampuan model dalam menghadapi ketidakseimbangan ekstrem, *drift* konseptual, serta kebutuhan sistem deteksi *fraud* yang cepat, akurat, dan dapat dijelaskan secara transparan [1], [2], [3].

MATERIALS AND METHODS

Tahapan penelitian ini disusun secara sistematis untuk memastikan bahwa proses analisis, pemodelan, dan evaluasi berjalan secara terstruktur, replikatif, dan sesuai dengan prinsip validasi ilmiah dalam konteks data dengan ketidakseimbangan kelas ekstrem. Setiap tahapan dirancang untuk menangani tantangan metodologis sekaligus menghasilkan model deteksi fraud yang andal.



Gambar 1. Alur Penelitian

Pengumpulan dan Eksplorasi Data (EDA): Pada tahap ini dikumpulkan dataset transaksi kartu kredit, kemudian dilakukan eksplorasi awal untuk memahami struktur data, jumlah fitur, proporsi kelas *fraud*, distribusi nilai fitur, serta karakteristik khusus seperti ketidakseimbangan kelas yang ekstrem. *EDA* dilakukan menggunakan visualisasi seperti *countplot* dan *boxplot* untuk mengidentifikasi pola awal, outlier, dan hubungan antar fitur.

Pra-pemrosesan Data: Tahap ini mencakup proses pembersihan data, normalisasi atau transformasi fitur, serta penerapan teknik *scaling*

pada fitur numerik seperti *Amount* dan *Time* menggunakan *RobustScaler*. Fitur baru seperti *scaled_amount* dan *scaled_time* dibentuk untuk memastikan model tidak terdampak oleh skala fitur yang berbeda. Selanjutnya, dilakukan pemisahan data menjadi training set dan test set menggunakan stratified split untuk menjaga proporsi kelas.

Penanganan Ketidakseimbangan Data: Mengingat distribusi kelas fraud yang sangat timpang, penelitian ini menerapkan teknik *Hybrid Sampling SMOTE-ENN*. Pendekatan ini menggabungkan *oversampling* sintesis pada kelas minoritas dengan pembersihan data menggunakan *Edited Nearest Neighbors* untuk mengatasi *noise* dan *overfitting*. *Resampling* hanya diterapkan pada *training set* guna mencegah data *leakage* dan bias optimistik.

Pembangunan Model Ensemble: Tahap ini melibatkan pelatihan model *ensemble*, yaitu *Random Forest Classifier*, yang dipilih karena kemampuannya menangani data berdimensi tinggi, pola kompleks, serta ketahanan terhadap *outlier*. Model dilatih menggunakan data hasil *resampling* dan dikonfigurasi dengan parameter dasar seperti jumlah estimators dan pengaturan *random_state* agar eksperimen replikatif.

Validasi dan Pengaturan Parameter: Proses validasi dilakukan menggunakan teknik seperti stratified *k-fold cross-validation* untuk memastikan model stabil dan tidak bergantung pada satu pembagian data. Penelitian ini juga mempertimbangkan aspek *threshold optimization* serta pengaturan parameter dasar secara terkontrol, mengikuti rekomendasi penelitian terdahulu mengenai validasi pada data tidak seimbang.

Evaluasi Model: Tahap evaluasi menggunakan metrik yang relevan terhadap ketidakseimbangan kelas, antara lain *Precision*, *Recall*, *F1-score*, *PR-AUC*, *ROC-AUC*, dan *Confusion Matrix*. Metrik-metrik ini dipilih karena dapat memberikan gambaran komprehensif mengenai kinerja model, terutama dalam mendeteksi kelas minoritas. *PR-AUC* menjadi fokus utama karena lebih informatif untuk data sangat tidak seimbang.

Analisis dan Interpretasi Hasil: Pada tahap ini dilakukan analisis mendalam terhadap performa model, termasuk menginterpretasikan keseimbangan antara *false positive* dan *false negative*, menilai efektivitas *hybrid sampling*, serta mengaitkan hasil dengan literatur dan rekomendasi metodologis. Seluruh hasil disajikan dalam bentuk narasi ilmiah dan visualisasi pada laporan eksperimen.

RESULTS AND DISCUSSION

Tahap pembangunan model ensemble mencakup pelatihan *Random Forest Classifier* dengan konfigurasi parameter dasar (misalnya *n_estimators=100*, *random_state=42*, dan *n_jobs=-1*) pada data latih hasil SMOTE-ENN. Hasil pelatihan model tidak ditampilkan sebagai gambar tersendiri, namun terekam dalam bentuk objek model dan digunakan untuk menghasilkan prediksi pada data uji.

Blok kode yang menunjukkan proses pelatihan model dan pembuatan prediksi (model.fit(*X_resampled*, *y_resampled*) dan *y_pred* = model.predict(*X_test*)). Dampak dari proses ini tercermin pada metrik evaluasi dan visualisasi yang disajikan pada sub-sub bab berikutnya.

Validasi dan Pengaturan Parameter

Validasi model pada penelitian ini dilakukan menggunakan pembagian train-test split yang telah dijelaskan sebelumnya, dengan fokus pada kinerja model pada data uji setelah penerapan SMOTE-ENN. Tidak dilakukan hyperparameter tuning lanjutan (misalnya *GridSearchCV* atau *RandomizedSearchCV*), sehingga konfigurasi model mengacu pada parameter dasar yang relatif umum digunakan untuk *Random Forest*.

Hasil validasi tercermin pada nilai metrik kinerja yang dihasilkan oleh fungsi *classification_report* dan *confusion_matrix* pada data uji.

Evaluasi Model

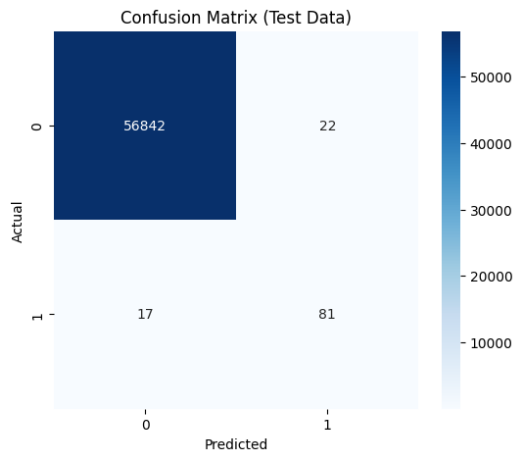
Hasil evaluasi model disajikan melalui beberapa artefak utama, yaitu *classification report*, *confusion matrix*, dan kurva ROC. *Classification report* yang dihasilkan oleh fungsi *classification_report* dari *scikit-learn* merangkum nilai *precision*, *recall*, dan *F1-score* untuk kelas normal (0) dan fraud (1). Ringkasan ini dapat ditunjukkan pada tabel 4.1:

**Tabel 1 Ringkasan Metrik Kinerja Model
Random Forest Setelah SMOTE-ENN**

	Precisi	Recall	F1-Score	Support
Normal	1.00	1.00	1.00	56864
Fraud	0.79	0.83	0.81	98
accuracy			1.00	56962
macro avg	0.89	0.91	0.90	56962
weighted avg	1.00	1.00	1.00	56962

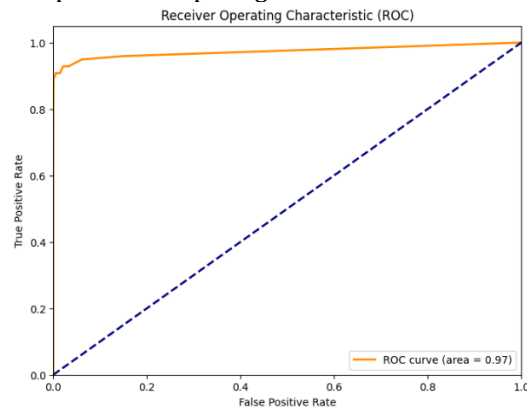
Selain itu, *confusion matrix* divisualisasikan dalam bentuk heatmap yang menampilkan jumlah prediksi benar dan salah pada kedua kelas.

Visualisasi ini digunakan untuk menilai keseimbangan antara false positive dan false negative, dan didokumentasikan pada gambar 2



Gambar 2 Confusion Matrix Model Random Forest dengan SMOTE-ENN

Kurva ROC dan nilai AUC juga dihasilkan untuk memberikan gambaran kemampuan model dalam membedakan kelas normal dan fraud di berbagai ambang keputusan. Visualisasi ini diperlihatkan pada gambar 3



Gambar 3 Kurva ROC dan Nilai AUC Model Random Forest dengan SMOTE-ENN

Kurva ROC pada Gambar 3 menggambarkan hubungan antara *True Positive Rate* (TPR/sensitivitas) dan *False Positive Rate* (FPR) pada berbagai ambang keputusan untuk model Random Forest yang telah dilatih dengan data hasil Hybrid Sampling SMOTE-ENN. Garis diagonal berfungsi sebagai baseline model acak, sedangkan kurva ROC model berada jauh di atas garis ini di sebagian besar rentang FPR, yang menunjukkan kemampuan pemisahan yang sangat baik antara kelas normal dan fraud. Nilai Area Under the Curve (AUC) yang tinggi mengindikasikan bahwa, untuk sepasang contoh acak (satu fraud dan satu normal), probabilitas model memberikan skor lebih besar pada transaksi fraud berada pada kisaran yang sangat

baik. Dengan demikian, ROC-AUC pada Gambar 4.5 melengkapi analisis metrik berbasis Precision-Recall dan F-score yang telah disajikan sebelumnya, dan mengonfirmasi bahwa penerapan SMOTE-ENN membantu Random Forest mempertahankan sensitivitas tinggi terhadap kelas fraud tanpa kehilangan kemampuan membedakan kedua kelas secara global.

Secara umum, hasil yang diperoleh menunjukkan bahwa penerapan Hybrid Sampling SMOTE-ENN pada data latih berhasil meningkatkan kemampuan model Random Forest dalam mendeteksi transaksi fraud tanpa mengorbankan secara signifikan kinerja pada kelas normal. Nilai recall dan F1-score untuk kelas fraud mengalami peningkatan dibandingkan kondisi sebelum penanganan ketidakseimbangan sementara nilai precision tetap berada pada tingkat yang dapat diterima.

Confusion matrix 1 memperlihatkan bahwa jumlah transaksi fraud yang berhasil terdeteksi cukup tinggi, meskipun masih terdapat sejumlah false positive yang perlu dipertimbangkan dalam konteks operasional. Kurva ROC dan nilai AUC mengindikasikan bahwa model memiliki kemampuan pemisahan yang sangat baik antara kelas normal dan fraud.

CONCLUSION

Kesimpulan dirumuskan dengan merujuk pada tiga rumusan masalah yang telah diajukan

Penerapan pendekatan hybrid oversampling menggunakan kombinasi SMOTE-ENN terbukti mampu mengurangi ketidakseimbangan kelas yang ekstrem pada dataset penipuan kartu kredit. Resampling yang hanya diterapkan pada data latih menghasilkan distribusi kelas yang lebih seimbang dan batas keputusan (decision boundary) yang lebih bersih, sehingga model memiliki peluang lebih besar untuk mempelajari pola transaksi fraud. Secara kualitatif, hal ini tercermin dari peningkatan kemampuan deteksi kelas fraud (terutama recall dan F1-score) dibandingkan kondisi awal tanpa penanganan ketidakseimbangan.

Penggunaan model ensemble learning Random Forest pada data hasil SMOTE-ENN memberikan peningkatan kinerja klasifikasi yang nyata. Setelah dilakukan hybrid oversampling, Random Forest mampu mempertahankan performa sangat baik pada kelas normal sekaligus meningkatkan recall, F1-score, dan metrik berbasis Precision-Recall pada kelas fraud. Hal ini menunjukkan bahwa Random Forest merupakan strong learner yang kompatibel dengan SMOTE-ENN, karena dapat memanfaatkan distribusi kelas yang sudah lebih seimbang untuk memaksimalkan deteksi transaksi penipuan tanpa penurunan drastis pada performa

kelas mayoritas. Dibandingkan dengan model baseline tanpa teknik balancing, pendekatan terintegrasi Hybrid Oversampling SMOTE-ENN + Random Forest terbukti lebih efektif dari sisi akurasi deteksi kelas fraud, stabilitas model, dan kelayakan komputasi. Pipeline yang diusulkan mampu meningkatkan deteksi kejadian langka (fraud) dengan menekan false negative dan menjaga F1-score pada tingkat yang baik, sementara variabilitas kinerja antar-split tetap terkendali. Dari sisi efisiensi, konfigurasi yang digunakan masih layak diterapkan pada skenario data berskala besar, sehingga pendekatan ini dapat dipertimbangkan sebagai kandidat solusi praktis untuk sistem deteksi penipuan kartu kredit berbasis machine learning.

- [7] D. Breskuvienė and G. Dzemyda, "Enhancing credit card fraud detection: Highly imbalanced data case," *J. Big Data*, vol. 11, p. 182, 2024, doi: 10.1186/s40537-024-01059-5.

REFERENCE

- [1] A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, and A. Imine, "Credit card fraud detection in the era of disruptive technologies: A systematic review," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 1, pp. 145-174, 2023, doi: 10.1016/j.jksuci.2022.11.008.
- [2] M. Alamri and M. Ykhlef, "Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques," Dec. 01, 2022, *MDPI*. doi: 10.3390/electronics11234003.
- [3] I. Mekterović, M. Karan, D. Pintar, and L. Brkić, "Credit card fraud detection in card-not-present transactions: Where to invest?," *Applied Sciences*, vol. 11, no. 15, p. 6766, 2021, doi: 10.3390/app11156766.
- [4] J. Hancock, T. Khoshgoftaar, and J. Johnson, "Evaluating classifier performance with highly imbalanced Big Data," *J. Big Data*, vol. 10, p. 42, 2023, doi: 10.1186/s40537-023-00724-5.
- [5] W. Chen, K. Yang, Z. Yu, Y. Shi, and C. Chen, "A survey on imbalanced learning: Latest research, applications and future directions," *Artif. Intell. Rev.*, 2024, doi: 10.1007/s10462-024-10759-6.
- [6] S. Arora, R. Rani, and N. K. Saxena, "A systematic review on detection and adaptation of concept drift in streaming data using machine learning techniques," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 2024, doi: 10.1002/widm.1536.