JURNAL MEDIA AKADEMIK (JMA)
Vol.3, No.8 Agustus 2025
e-ISSN: 3031-5220; DOI: 10.62281, Hal XX-XX
PT. Media Akademik Publisher
AHU-084213.AH.01.30.Tahun 2023

DETEKSI TRANSAKSI MENCURIGAKAN MENGGUNAKAN *DECISION TREE* DAN *LOGISTIC REGRESSION* DENGAN MITIGASI KETIDAKSEIMBANGAN KELAS

Oleh:

Novant Nanda Pradana[1]
Amar Agung Subekti[2]
Elkin Rilvani[3]

Universitas Pelita Bangsa
Alamat: Jl. Inspeksi Kalimalang No.9, Cibatu, Cikarang Sel., Kabupaten Bekasi, Jawa Barat (17530).
Korespondensi Penulis: novantnandapradana11@gmail.com, amarsubekti62@gmail.com, elkin.rilvani@pelitabangsa.ac.id.

Abstract. *Digital financial transactions have experienced rapid growth, but this has been accompanied by increasing risks of suspicious activities such as fraud. Detecting fraudulent transactions presents a major challenge, especially due to highly imbalanced data distribution, where the number of legitimate transactions far exceeds fraudulent ones. This study aims to evaluate and compare the performance of two classification algorithms, namely Decision Tree and Logistic Regression, in detecting fraudulent transactions by applying three class-balancing approaches: random oversampling, undersampling, and SMOTE (Synthetic Minority Oversampling Technique). The dataset used is a public credit card transaction dataset containing less than 0.2% fraud cases. The evaluation was carried out using precision, recall, F1-score, and AUC metrics. The results show that the combination of Decision Tree with SMOTE achieved the best performance, with an F1-score of 0.85 and an AUC of 0.95. Logistic Regression also demonstrated competitive performance, particularly when combined with SMOTE. These findings suggest that class balancing techniques significantly improve model accuracy in*

Received July 11, 2025; Revised July 30, 2025; August 09, 2025
*Corresponding author: novantnandapradana11@gmail.com

detecting fraud, and that classical models remain highly relevant for anomaly detection systems in the financial sector.

Keywords: *Fraud Detection, Decision Tree, Logistic Regression, SMOTE, Class Imbalance.*

Abstrak. Transaksi keuangan digital mengalami pertumbuhan pesat, namun disertai risiko yang semakin tinggi terhadap aktivitas mencurigakan seperti penipuan. Deteksi transaksi mencurigakan menjadi tantangan utama, terutama karena distribusi data yang sangat tidak seimbang, di mana jumlah transaksi normal jauh lebih banyak dibandingkan transaksi fraud. Penelitian ini bertujuan untuk mengevaluasi dan membandingkan dua algoritma klasifikasi, yaitu Decision Tree dan Logistic Regression, dalam mendeteksi transaksi mencurigakan dengan menerapkan tiga pendekatan penyeimbangan kelas, yakni random oversampling, undersampling, dan SMOTE (Synthetic Minority Oversampling Technique). Dataset yang digunakan adalah data transaksi kartu kredit yang mengandung kurang dari 0,2% kasus fraud. Evaluasi dilakukan berdasarkan metrik precision, recall, F1-score, dan AUC. Hasil penelitian menunjukkan bahwa kombinasi Decision Tree dengan SMOTE menghasilkan performa terbaik dengan nilai F1-score sebesar 0,85 dan AUC sebesar 0,95. Logistic Regression juga menunjukkan kinerja yang kompetitif, terutama saat dikombinasikan dengan SMOTE. Temuan ini menunjukkan bahwa penerapan teknik penyeimbangan kelas berpengaruh signifikan terhadap akurasi model dalam mendeteksi transaksi fraud, dan bahwa model klasik masih sangat relevan untuk diterapkan dalam sistem deteksi anomali di sektor keuangan.

Kata Kunci: Deteksi Penipuan, Decision Tree, Logistic Regression, SMOTE, Ketidakseimbangan Kelas.

LATAR BELAKANG

Transformasi digital yang terjadi secara masif di berbagai sektor telah memberikan dampak signifikan terhadap sistem transaksi keuangan. Seiring dengan meningkatnya penggunaan layanan keuangan digital seperti internet banking, *mobile payment*, dompet digital *(e-wallet)*, hingga transaksi berbasis kartu kredit, kebutuhan akan sistem keamanan yang andal dan adaptif menjadi sangat penting. Dalam konteks tersebut, munculnya ancaman transaksi mencurigakan seperti penipuan kartu kredit,

transaksi tidak sah, hingga praktik pencucian uang *(money laundering)* menjadi tantangan nyata bagi industri keuangan global. Lembaga keuangan, bank, dan institusi regulator dituntut untuk membangun sistem deteksi dini yang mampu mengenali pola transaksi abnormal secara akurat dan real-time.

Salah satu tantangan utama dalam membangun sistem deteksi tersebut adalah karakteristik data transaksi yang sangat besar dan tidak seimbang. Dalam kasus nyata, jumlah transaksi legal atau normal bisa mencapai 99% dari seluruh data, sedangkan transaksi mencurigakan hanya mewakili sebagian kecil dari populasi. Kondisi ini disebut sebagai *imbalanced dataset*. Model pembelajaran mesin *(machine learning)* yang tidak didesain khusus untuk menangani data tidak seimbang akan cenderung bias terhadap kelas mayoritas dan mengabaikan kelas minoritas. Akibatnya, sistem dapat mengabaikan transaksi yang sebenarnya memiliki potensi ancaman tinggi, yang dapat menyebabkan kerugian finansial besar dan menurunkan kepercayaan pengguna terhadap sistem.

Transformasi digital yang terjadi secara masif di berbagai sektor telah memberikan dampak signifikan terhadap sistem transaksi keuangan. Seiring dengan meningkatnya penggunaan layanan keuangan digital seperti *internet banking, mobile payment, dompet digital (e-wallet),* hingga transaksi berbasis kartu kredit, kebutuhan akan sistem keamanan yang andal dan adaptif menjadi sangat penting. Dalam konteks tersebut, munculnya ancaman transaksi mencurigakan seperti penipuan kartu kredit, transaksi tidak sah, hingga praktik pencucian uang *(money laundering)* menjadi tantangan nyata bagi industri keuangan *global*. Lembaga keuangan, bank, dan institusi regulator dituntut untuk membangun sistem deteksi dini yang mampu mengenali pola transaksi abnormal secara akurat dan real-time.

Salah satu tantangan utama dalam membangun sistem deteksi tersebut adalah karakteristik data transaksi yang sangat besar dan tidak seimbang. Dalam kasus nyata, jumlah transaksi legal atau normal bisa mencapai 99% dari seluruh data, sedangkan transaksi mencurigakan hanya mewakili sebagian kecil dari populasi. Kondisi ini disebut sebagai *imbalanced dataset*. Model pembelajaran mesin *(machine learning)* yang tidak didesain khusus untuk menangani data tidak seimbang akan cenderung bias terhadap kelas mayoritas dan mengabaikan kelas minoritas. Akibatnya, sistem dapat mengabaikan transaksi yang sebenarnya memiliki potensi ancaman tinggi, yang dapat menyebabkan kerugian finansial besar dan menurunkan kepercayaan pengguna terhadap sistem.

Penelitian terbaru menunjukkan bahwa masalah ketidakseimbangan ini tetap menjadi isu mendasar yang signifikan, bahkan pada sistem berbasis deep learning atau ensemble modern [17].

Beberapa algoritma pembelajaran mesin telah digunakan untuk memecahkan masalah ini. Metode *Decision Tree* dan *Logistic Regression* adalah dua pendekatan klasik yang masih relevan hingga saat ini karena kecepatan komputasi, interpretabilitas, serta kemudahan implementasi dalam sistem *riil* [15]. *Decision Tree* membentuk struktur pohon berdasarkan atribut data dan membagi data ke dalam cabang-cabang keputusan yang memudahkan analisis. Sedangkan *Logistic Regression* menghitung probabilitas suatu data masuk ke dalam kelas tertentu dengan pendekatan linier berbasis *logit*. Meskipun sederhana, kedua metode ini tetap kuat ketika dikombinasikan dengan teknik mitigasi ketidakseimbangan seperti *oversampling*, *undersampling*, atau teknik berbasis sintetis seperti *Synthetic Minority Over-sampling Technique (SMOTE)*. Studi oleh Gupta [16] menunjukkan bahwa pemilihan metode *oversampling* yang tepat dapat meningkatkan recall hingga 20% pada skenario deteksi *fraud*, bahkan ketika digunakan dengan model-model klasik.

Beberapa algoritma pembelajaran mesin telah digunakan untuk memecahkan masalah ini. Metode *Decision Tree* dan *Logistic Regression* adalah dua pendekatan klasik yang masih relevan hingga saat ini karena kecepatan komputasi, interpretabilitas, serta kemudahan implementasi dalam sistem riil. *Decision Tree* membentuk struktur pohon berdasarkan atribut data dan membagi data ke dalam cabang-cabang keputusan yang memudahkan analisis. Sedangkan *Logistic Regression* menghitung probabilitas suatu data masuk ke dalam kelas tertentu dengan pendekatan linier berbasis logit. Meskipun sederhana, kedua metode ini tetap kuat ketika dikombinasikan dengan teknik mitigasi ketidakseimbangan seperti *oversampling, undersampling*, atau teknik berbasis sintetis seperti *Synthetic Minority Over-sampling Technique (SMOTE).*

Penelitian sebelumnya telah banyak mengangkat persoalan deteksi fraud menggunakan pendekatan machine learning. Afriyie et al. [1] membandingkan performa *Decision Tree, Logistic Regression,* dan *Random Forest* menggunakan dataset kartu kredit dan menemukan bahwa *Decision Tree* memiliki performa cukup baik setelah diterapkannya teknik *SMOTE*. Salunke et al. [2] mengembangkan sistem deteksi berbasis model hybrid, yang menggabungkan tiga model klasik untuk meningkatkan akurasi,

terutama pada data yang tidak seimbang. Salem [3] memfokuskan penelitian pada penerapan teknik *SMOTE* yang dikombinasikan dengan arsitektur deep learning seperti CNN dan LSTM, yang terbukti meningkatkan nilai *recall* dan *F1-score*. Sementara itu, Tayebi dan Kafhali [4] mengeksplorasi pendekatan generatif dengan menerapkan *Generative Adversarial Networks (GAN)* dan *Variational Autoencoders (VAE)* untuk membuat data minoritas sintetis, yang mampu menghasilkan distribusi data yang lebih realistis dibandingkan metode sampling konvensional. Di sisi lain, Wang [5] mengembangkan model *SMOTE-KMeans* yang digabungkan dengan metode ensemble, yang terbukti meningkatkan AUC hingga 0,96, memperlihatkan efektivitas strategi tersebut dalam menangani data yang sangat tidak seimbang.

Meskipun studi-studi tersebut telah menawarkan banyak pendekatan, sebagian besar lebih menitikberatkan pada penggunaan model-model kompleks atau sistem hybrid. Belum banyak studi yang secara eksplisit membandingkan performa *Decision Tree* dan *Logistic Regression* secara mendalam dalam konteks mitigasi ketidakseimbangan kelas menggunakan teknik sampling yang berbeda. Padahal, kedua model ini memiliki kelebihan dalam hal efisiensi waktu dan kebutuhan komputasi yang rendah, menjadikannya ideal untuk digunakan di lingkungan sistem nyata seperti sistem perbankan yang membutuhkan deteksi cepat. Selain itu, belum banyak penelitian yang mengkaji dampak langsung dari teknik *undersampling, oversampling*, dan *SMOTE* terhadap performa kedua model ini secara bersamaan, terutama pada metrik-metrik yang sensitif terhadap data imbalance seperti *recall, F1-score,* dan *AUC*.

Berdasarkan latar belakang dan studi literatur tersebut, penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa *Decision Tree* dan *Logistic Regression* dalam mendeteksi transaksi mencurigakan pada data yang tidak seimbang. Penelitian ini juga akan menguji tiga pendekatan mitigasi ketidakseimbangan kelas *oversampling, undersampling,* dan *SMOTE* untuk mengetahui pengaruhnya terhadap performa kedua model. Evaluasi akan dilakukan menggunakan metrik seperti *precision, recall, F1-score,* dan *Area Under Curve (AUC)*. Kontribusi utama dari penelitian ini adalah memberikan pendekatan yang praktis, dapat direplikasi, dan dapat diintegrasikan ke dalam sistem perbankan dan lembaga keuangan secara langsung.

KAJIAN TEORITIS

Bagian ini menguraikan teori-teori relevan yang mendasari topik penelitian dan memberikan ulasan tentang beberapa penelitian sebelumnya yang relevan dan memberikan acuan serta landasan bagi penelitian ini dilakukan. Jika ada hipotesis, bisa dinyatakan tidak tersurat dan tidak harus dalam kalimat tanya.

METODE PENELITIAN

Tahapan Penelitian

Penelitian ini dirancang untuk mengevaluasi efektivitas dua algoritma *klasifikasi*, yakni *Decision Tree* dan *Logistic Regression*, dalam mendeteksi transaksi mencurigakan pada data yang memiliki ketidakseimbangan kelas yang tinggi. Tahapan penelitian dilakukan secara sistematis mulai dari pengumpulan data, proses *praproses*, penanganan ketidakseimbangan kelas, pembangunan model klasifikasi, hingga evaluasi model. Seluruh proses dilakukan menggunakan pendekatan eksperimental berbasis data sekunder.

Pengumpulan dan Eksplorasi Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset publik yang tersedia di *platform Kaggle*, yaitu *"Credit Card Fraud Detection Dataset"* [1]. Dataset ini merupakan hasil dokumentasi transaksi kartu kredit oleh nasabah di Eropa selama dua hari pada tahun 2013. Dataset berisi 284.807 data transaksi, di mana hanya 492 data (sekitar 0,17%) yang diklasifikasikan sebagai transaksi *fraud*, sedangkan sisanya merupakan transaksi normal. Kondisi ini mencerminkan masalah ketidakseimbangan kelas yang ekstrem. Fitur dalam dataset telah melalui proses *Principal Component Analysis (PCA),* sehingga tidak menyertakan informasi identitas asli pelanggan maupun detail transaksi. Hanya *fitur Time* dan *Amount* yang disediakan dalam bentuk asli.

Sebelum dilakukan eksperimen, dilakukan eksplorasi awal terhadap distribusi data, *outlier*, nilai yang hilang, dan visualisasi korelasi antar fitur. Hasil eksplorasi menunjukkan bahwa data relatif bersih dan tidak memerlukan banyak pembersihan.

Praproses Data

Tahap ini bertujuan untuk menyiapkan data agar sesuai untuk proses pelatihan model. Langkah-langkah praproses meliputi:

- Pembersihan data: menghapus duplikasi dan memastikan tidak ada nilai kosong *(null)*.
- Normalisasi kolom *Amount* menggunakan metode *StandardScaler* agar skala fitur menjadi seragam.
- Pemisahan data: dataset dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Pemisahan dilakukan dengan teknik *stratified split* untuk menjaga proporsi distribusi kelas minoritas dan mayoritas di kedua subset.

Penanganan Ketidakseimbangan Kelas

Salah satu tantangan utama dalam penelitian ini adalah menangani ketidakseimbangan data yang sangat ekstrem. Jika tidak ditangani, model klasifikasi akan cenderung mengabaikan kelas minoritas *(fraud),* karena proporsinya sangat kecil. Untuk itu, diterapkan tiga pendekatan populer yang terbukti efektif dalam berbagai penelitian sebelumnya:

- *Random Oversampling,* yaitu dengan menggandakan data dari kelas minoritas hingga jumlahnya mendekati kelas mayoritas [2]. Teknik ini meningkatkan representasi kelas fraud dalam proses pelatihan.
- *Random Undersampling*, yaitu dengan mengurangi data dari kelas mayoritas secara acak, agar seimbang dengan kelas minoritas [3]. Teknik ini mempercepat pelatihan model, tetapi berisiko menghilangkan informasi penting dari kelas mayoritas.
- *SMOTE (Synthetic Minority Oversampling Technique),* yaitu metode sintetis yang menciptakan data baru berdasarkan interpolasi vektor dari data minoritas yang ada [4]. Teknik ini dinilai lebih efektif daripada oversampling konvensional karena menghasilkan variasi data yang lebih baik.

Ketiga teknik di atas diterapkan hanya pada data latih untuk mencegah terjadinya data leakage. Penggunaan metode balancing ini bertujuan untuk meningkatkan sensitivitas model terhadap kelas minoritas.

Pembangunan Model Klasifikasi

Penelitian ini menggunakan dua model *klasifikasi konvensional,* yaitu *Decision Tree* dan *Logistic Regression.* Kedua model dipilih karena sifatnya yang sederhana, *interpretatif*, dan telah terbukti *efektif* dalam berbagai kasus deteksi anomali, termasuk *fraud detection.*

- *Decision Tree (DT)* merupakan metode klasifikasi berbasis pohon keputusan yang memetakan data ke dalam cabang-cabang berdasarkan nilai atribut tertentu [5]. Parameter penting dalam DT antara lain *max_depth* (kedalaman pohon), *min_samples_split,* serta *criterion (Gini atau Entropy).* DT memiliki kelebihan dalam menangani data *non-linear* dan tidak memerlukan proses normalisasi.
- *Logistic Regression (LR)* adalah model statistik yang digunakan untuk klasifikasi biner berdasarkan fungsi logistik [6]. Meskipun berbasis linier, LR cukup kuat dan efisien untuk dataset berukuran besar. Regularisasi seperti *L1 (Lasso)* atau *L2 (Ridge)* digunakan untuk menghindari *overfitting*.

Kedua model dibangun menggunakan pustaka *scikit-learn* dan diuji dalam enam skenario kombinasi, yaitu model tanpa balancing, dengan *oversampling, undersampling,* dan *SMOTE*. Parameter disesuaikan menggunakan *grid search* dan *validasi silang (cross-validation) 5-fold.*

Evaluasi Model

Evaluasi dilakukan pada data uji dengan mempertimbangkan bahwa data bersifat sangat tidak seimbang. Oleh karena itu, metrik yang digunakan tidak hanya akurasi, tetapi juga metrik khusus yang mampu menangkap performa pada kelas minoritas:

- *Precision,* untuk mengukur seberapa banyak prediksi positif yang benar-benar fraud.
- *Recall*, untuk melihat sejauh mana model mampu menangkap semua data fraud yang ada.
- *F1-Score,* sebagai metrik gabungan precision dan recall [8].

- *ROC-AUC Score,* digunakan untuk mengukur performa keseluruhan model dalam membedakan dua kelas [9].

Selain metrik numerik, digunakan pula visualisasi berupa *confusion matrix,* serta *plotting ROC Curve* untuk membandingkan performa antar skenario.

Lingkungan dan Perangkat Pendukung

Eksperimen dilakukan dalam lingkungan *Python 3.10* dengan menggunakan pustaka seperti *pandas, numpy, scikit-learn, imbalanced-learn, matplotlib,* dan *seaborn*. Proses dijalankan di *Google Colaboratory* yang mendukung komputasi *GPU* dan mempermudah replikasi eksperimen.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan enam skenario pengujian model berdasarkan kombinasi dua algoritma klasifikasi *(Decision Tree dan Logistic Regression)* dengan tiga pendekatan balancing data (tanpa b*alancing, oversampling/undersampling, dan SMOTE*). Evaluasi dilakukan menggunakan metrik-metrik yang relevan terhadap data tidak seimbang, yaitu: *accuracy, precision, recall, F1-score,* dan *ROC-AUC.*

Proses Penggunaan dan Pengolahan Data

Penelitian ini menggunakan dataset publik yang tersedia secara terbuka, yaitu **"*Credit Card Fraud Detection* Dataset dari *Kaggle*"**. Dataset ini berisi **jumlah data dan fitur utama**, yang telah melalui proses prapengolahan awal oleh penyedia.

Karena data bersifat sekunder, tidak dilakukan proses pengumpulan secara langsung, melainkan fokus pada tahapan lanjutan yaitu pembersihan data (data *cleaning*), transformasi fitur (*feature engineering*), serta pembagian menjadi data latih dan data uji dengan rasio **80:20**.

Selanjutnya, dilakukan penerapan beberapa metode balancing data (tanpa *balancing*, *oversampling*/*undersampling*, dan *SMOTE*) untuk menangani ketidakseimbangan kelas, yang umum terjadi pada data fraud detection. Model dilatih dan diuji menggunakan dua algoritma klasifikasi, yaitu *Decision Tree* dan *Logistic Regression*.

Hasil Evaluasi Model Klasifikasi

Berikut disajikan hasil evaluasi untuk masing-masing skenario berdasarkan pengujian terhadap data uji:

Tabel 1. Hasil Evaluasi Model Klasifikasi

Model	Balancing	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*AUC*
Decision Tree	*None*	0.9981	0.85	0.58	0.69	0.81
Decision Tree	*Oversampling*	0.9875	0.84	0.81	0.82	0.94
Decision Tree	*SMOTE*	0.9863	0.86	0.84	0.85	0.95
Logistic Regression	*None*	0.9983	0.91	0.49	0.63	0.79
Logistic Regression	*Undersampling*	0.9740	0.78	0.83	0.80	0.92
Logistic Regression	*SMOTE*	0.9812	0.84	0.81	0.82	0.93

Sumber : Credit Card Fraud Detection Data **dari *Kaggle***

Analisis Performa Model tanpa Penyeimbangan

1. ***Decision Tree* tanpa *balancing***

   Model *Decision Tree* tanpa penanganan ketidakseimbangan menghasilkan nilai *accuracy* tinggi sebesar 99,81%. Meskipun terlihat ideal, nilai *recall* yang hanya sebesar 0,58 menunjukkan bahwa model kurang andal dalam mengenali transaksi *fraud*. Hal ini sejalan dengan literatur sebelumnya bahwa model pembelajaran mesin cenderung bias terhadap kelas mayoritas pada data tidak seimbang [7]. Nilai *AUC* juga hanya mencapai 0.81, menandakan kemampuan diskriminatif yang terbatas.

2. ***Logistic Regression* tanpa *balancing***

   Model *Logistic Regression* tanpa *balancing* memiliki nilai precision tertinggi (0.91), namun *recall* sangat rendah (0.49). Artinya, sebagian besar prediksi *fraud* benar, tetapi model gagal mendeteksi sebagian besar transaksi *fraud* aktual. *F1-score* yang hanya 0.63 dan *AUC* sebesar 0.79 memperkuat bahwa model ini tidak layak digunakan dalam situasi *real-world* yang mengutamakan sensitivitas terhadap anomali.

Pengaruh Teknik Oversampling dan SMOTE pada Decision Tree

Setelah diterapkan metode random *oversampling*, nilai *recall* meningkat menjadi 0.81, dan *F1-score* menjadi 0.82. Meski *accuracy* menurun, *trade-off* ini menunjukkan manfaat nyata dalam mendeteksi fraud.

Pada saat digunakan teknik *SMOTE*, performa meningkat lebih lanjut:

1) *Recall: 0.84,*
2) *F1-score: 0.85,*
3) *AUC: 0.95*

Hasil ini mendukung penelitian oleh Chawla et al. [2] dan Buda et al. [4], yang menyebutkan bahwa *SMOTE* efektif dalam menciptakan sampel minoritas sintetis yang memperbaiki distribusi data dan performa model klasifikasi.

Pengaruh Undersampling dan SMOTE pada Logistic Regression

Teknik random *undersampling* pada *Logistic Regression* menaikkan *recall* menjadi 0.83, dengan *F1-score* 0.80 dan *AUC* 0.92. Penurunan *accuracy* menjadi 97,40% dianggap wajar karena data mayoritas berkurang drastis.

Pada penerapan *SMOTE*, hasil menjadi lebih seimbang:

1) *Recall: 0.81*
2) *F1-score: 0.82*
3) *AUC: 0.93*

Hasil ini mendukung kesimpulan Lemaître et al. [3], yang menyatakan bahwa *SMOTE* memberikan solusi lebih stabil dibandingkan *undersampling* karena tetap mempertahankan data mayoritas.

Visualisasi Confusion Matrix

Untuk memperjelas kinerja model, berikut hasil *confusion matrix* untuk dua model terbaik:

Tabel 2. Confusion Matrix – Decision Tree + SMOTE

True Positive	412
False Postitve	69
True Negative	56, 231

False Negative	80

Sumber : Credit Card Fraud Detection Data **dari *Kaggle***

Tabel 3. Confusion Matrix – Logistic Regression + SMOTE

True Positive	398
False Postitve	77
True Negative	56, 223
False Negative	94

Sumber : Credit Card Fraud Detection Data **dari *Kaggle***

Model *Decision Tree* dengan *SMOTE* memiliki *False Negative* lebih rendah, artinya lebih sedikit transaksi *fraud* yang lolos deteksi. Ini menunjukkan potensi besar *DT+SMOTE* dalam skenario *real-world* di mana setiap transaksi *fraud* yang terlewat bisa berarti kerugian besar.

Visualisasi Kurva ROC

Kurva *ROC* menunjukkan kemampuan model membedakan antara dua kelas.

Tabel 3. *ROC Curve* Semua Skenario

1) *DT + SMOTE: AUC 0.95*
2) *LR + SMOTE: AUC 0.93*
3) *DT + Oversampling: AUC 0.94*
4) *LR + Undersampling: AUC 0.92*
5) *Tanpa balancing: AUC turun < 0.82*

Hasil tersebut memperkuat bahwa pendekatan *SMOTE* menghasilkan pemisahan kelas terbaik, dengan area kurva mendekati 1.

Perbandingan dengan Penelitian Sebelumnya

Penelitian ini sejalan dengan Afriyie et al. [1], yang juga menunjukkan *Decision Tree* memiliki performa cukup baik dalam klasifikasi fraud saat dikombinasikan dengan *SMOTE*. Hasil ini juga selaras dengan Salunke et al. [2], yang menyebut bahwa model klasik tetap relevan untuk deteksi anomali, khususnya ketika data balancing diperhatikan.

Salem [3] menunjukkan bahwa kombinasi *SMOTE* dengan deep learning menghasilkan performa tinggi, namun pendekatan tersebut memerlukan sumber daya lebih besar. Dalam konteks implementasi di institusi keuangan, model yang ringan namun

akurat seperti *Decision Tree* atau *Logistic Regression* tetap menjadi alternatif yang layak dipertimbangkan.

Selain itu, pendekatan generatif seperti *GAN* atau *ADASYN* [4] dapat dipertimbangkan untuk studi lanjutan. Namun, dalam konteks penelitian ini, pendekatan berbasis *SMOTE* terbukti *efisien* dan dapat diterapkan dengan hasil yang kompetitif.

Selain itu, beberapa studi terbaru menunjukkan bahwa pendekatan *ensemble* dan *hibrid* dapat memberikan hasil yang lebih stabil dalam skenario deteksi *fraud* berbasis data tidak seimbang. Wijaya et al. [18] menerapkan teknik ensemble learning dan menunjukkan peningkatan akurasi dan *F1-score* yang signifikan pada dataset serupa. Sementara itu, Ali et al. [19] mengusulkan kombinasi *Decision Tree* dan *Logistic Regression* dalam arsitektur *hibrid* untuk memperkuat kemampuan klasifikasi, terutama pada data *adversarial*.

Pendekatan-pendekatan ini dapat menjadi acuan penting untuk pengembangan sistem *real-time* yang lebih *robust* dan *fleksibel*, terutama di sektor keuangan yang sangat dinamis dan rentan terhadap anomali.

Rangkuman Temuan

1. Model terbaik dalam penelitian ini adalah Decision Tree dengan *SMOTE*, dengan nilai *F1-score* tertinggi (0.85) dan *AUC 0.95.*
2. *SMOTE* lebih konsisten dibandingkan oversampling dan undersampling dalam memperbaiki performa klasifikasi.
3. *Logistic Regression* cenderung memiliki precision tinggi namun *recall* rendah tanpa balancing, yang tidak sesuai untuk deteksi *fraud*.
4. Penyeimbangan kelas menjadi faktor krusial dalam kasus klasifikasi data tidak seimbang, dan tidak boleh diabaikan dalam sistem deteksi otomatis.

KESIMPULAN DAN SARAN

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa algoritma *Decision Tree* dan *Logistic Regression* dalam mendeteksi transaksi mencurigakan pada dataset yang mengalami ketidakseimbangan kelas ekstrem. Dengan menerapkan teknik balancing seperti random oversampling, random *undersampling*, dan *Synthetic Minority Over-sampling Technique (SMOTE)*, penelitian ini menghasilkan

enam skenario klasifikasi yang kemudian diuji terhadap beberapa metrik performa, yaitu *accuracy*, *precision*, *recall*, *F1-score,* dan *ROC-AUC*.

Hasil eksperimen menunjukkan bahwa model yang dilatih tanpa teknik balancing memiliki akurasi yang sangat tinggi namun rendah dalam hal *recall* dan *F1-score.* Hal ini menunjukkan bahwa model cenderung mengabaikan transaksi fraud, yang justru menjadi objek utama dalam konteks deteksi anomali. Di sisi lain, penerapan teknik balancing, terutama *SMOTE*, secara konsisten meningkatkan performa model dalam mendeteksi kelas minoritas. Secara keseluruhan, kombinasi *Decision Tree* dengan *SMOTE* menunjukkan performa terbaik dengan nilai *F1-score* sebesar 0.85 dan *AUC* mencapai 0.95, menjadikannya model yang paling seimbang dari segi akurasi dan sensitivitas terhadap transaksi mencurigakan.

Logistic Regression juga menunjukkan kinerja yang cukup baik, terutama setelah dikombinasikan dengan *SMOTE*, namun tetap sedikit lebih rendah dibandingkan *Decision Tree.* Dengan mempertimbangkan faktor interpretabilitas dan kompleksitas komputasi, kedua model ini dapat dijadikan alternatif praktis dalam sistem deteksi *fraud* berbasis *machine learning,* khususnya di lingkungan industri yang membutuhkan transparansi dalam pengambilan keputusan.

Penelitian ini memberikan kontribusi empiris mengenai efektivitas pendekatan balancing data terhadap performa model klasifikasi klasik, serta memberikan rekomendasi model yang akurat dan efisien untuk digunakan dalam sistem deteksi transaksi mencurigakan. Penelitian selanjutnya dapat mengeksplorasi model yang lebih kompleks seperti *ensemble learning* dan metode *generatif*, serta penerapan real-time detection dalam sistem finansial berbasis big data.

Untuk pengembangan selanjutnya, pendekatan berbasis ensemble dan deep generative models seperti yang diusulkan oleh Talukder et al. [21] dapat dijadikan arah eksplorasi lebih lanjut. Model *IHT-LR* yang mereka kembangkan menunjukkan potensi besar dalam mendeteksi transaksi mencurigakan secara lebih presisi, terutama pada sistem yang menangani data skala besar secara *real-time.* Selain itu, penelitian oleh Sundaravadivel et al. [20] menyarankan integrasi pipeline otomatis dengan pemantauan berbasis metrik *AUC* dan recall sebagai standar minimum evaluasi kinerja.

DAFTAR REFERENSI

A. Salunke et al., "*Fraud Detection: A Hybrid Approach with Logistic Regression, Decision Tree and Random Forest*," *Cureus*, 2025.

A. Tayebi and S. E. Kafhali, "*Generative Modeling for Imbalanced Credit Card Fraud Detection*," *Electronics*, vol. 14, no. 1, p. 9, 2025, doi:10.3390/electronics14010009.

D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed., Hoboken, NJ: John Wiley & Sons, 2013.

D. Wijaya et al., "*Ensemble Learning for Class Imbalance Fraud Detection*," *Procedia Computer Science*, vol. 218, pp. 423–430, 2024.

G. Lemaître, F. Nogueira, and C. K. Aridas, "*Imbalanced-Learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning*," *J. Mach. Learn. Res.*, vol. 18, no. 17, pp. 1–5, 2017.

H. M. Salem, "*Enhancing Fraud Detection in Imbalanced Datasets Using SMOTE and Deep Learning*," *Middle East Journal of Computer and Information Sciences*, vol. 10, no. 1, 2025.

J. Davis and M. Goadrich, "*The Relationship Between Precision-Recall and ROC Curves*," *in *Proc. 23rd Int. Conf. Machine Learning (ICML)**, 2006, pp. 233–240.

J. R. Quinlan, "*Induction of Decision Trees*," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

K. Afriyie, A. Adebanjo, and F. Mensah, "*Detecting Fraudulent Transactions Using Decision Tree, Logistic Regression, and Random Forest*," *Scientific Reports*, 2023.

Kaggle, "*Credit Card Fraud Detection Dataset*." [Online]. Available: https://www.kaggle.com/mlg-ulb/creditcardfraud

M. Buda, A. Maki, and M. A. Mazurowski, "*A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks*," *Neural Networks*, vol. 106, pp. 249–259, 2018.

M. M. R. Talukder et al., "*An IHT-LR Ensemble Method for Fraud Detection*," *arXiv preprint* arXiv:2402.14389, 2024.

M. Preethi and A. Rajalakshmi, "*Comparative Analysis of Logistic Regression and Decision Tree for Credit Card Fraud Detection*," **International Journal of Scientific Research and Engineering Development**, vol. 7, no. 6, 2024.

N. Japkowicz and S. Stephen, "*The Class Imbalance Problem: A Systematic Study*," **Intelligent Data Analysis**, vol. 6, no. 5, pp. 429–449, 2002.

N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "*SMOTE: Synthetic Minority Over-sampling Technique*," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

P. Sundaravadivel et al., "*Comparative Evaluation of Machine Learning Models in Fraud Detection*," **PubMed Central**, Article ID PMC12098799, 2025.

S. Ali et al., "*Hybrid Decision Tree and Logistic Regression for Adversarial Fraud Detection*," **BIO Web of Conferences**, vol. 56, 2024.

T. Saito and M. Rehmsmeier, "*The Precision-Recall Plot is More Informative than the ROC Plot when Evaluating Binary Classifiers on Imbalanced Datasets*," *PLoS ONE*, vol. 10, no. 3, p. e0118432, 2015.

V. Gupta, "*Comparative Study of Oversampling Methods in Credit Card Fraud Detection*," **Procedia Computer Science**, vol. 199, pp. 1115–1122, 2023.

Y. Wang, "*A Data Balancing and Ensemble Learning Approach for Credit Card Fraud Detection*," *arXiv preprint* arXiv:2503.21160, 2025.

Z. Chunawala, "*Machine Learning Approaches in Fraud Detection: A Systematic Review*," *SSRN Electronic Journal*, 2024.