Jurnal Informatika Terpadu

https://journal.nurulfikri.ac.id/index.php/JIT
ISSN ONLINE : 2460-8998

DETEKSI *FRAUD* KARTU KREDIT DENGAN *LOGISTIC REGRESSION, RANDOM FOREST*, DAN *GRADIENT BOOSTING*

Herlambang Awan Irawan[1]

[1] Sains Data, Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya, Jawa Timur, Indonesia 60294
22083010101@student.upnjatim.ac.id

Abstract

This study aims to develop a credit card transaction fraud detection model using machine learning approaches, namely Logistic Regression, Random Forest, and Gradient Boosting Classifier. The dataset used is sourced from real credit card transactions with a fraud proportion of 0.17%, which reflects the problem of class imbalance. To overcome this, the Synthetic Minority Over-sampling Technique (SMOTE) and feature transformation using Principal Component Analysis (PCA) were applied. The evaluation was carried out using accuracy, precision, recall, F1-score, and ROC-AUC metrics. The results show that Random Forest and Gradient Boosting Classifier produced the best performance with near-perfect accuracy and ROC-AUC values (ROC-AUC > 0.999), while Logistic Regression gave very good results but slightly below the other two models. However, the near-perfect ROC-AUC value may indicate potential overfitting, requiring further validation on different datasets. Unlike previous studies that only used one algorithm, this study compared three models simultaneously and integrated SMOTE and PCA to improve detection performance. The practical implication of this study is that the proposed model can be implemented in digital financial systems to assist banking institutions in detecting fraud in real time and reducing potential financial losses.

Keywords: *credit card, fraud detection, gradient boosting, machine learning, SMOTE*

Abstrak

Penelitian ini bertujuan untuk membangun model deteksi penipuan transaksi kartu kredit menggunakan pendekatan *machine learning*, yaitu *Logistic Regression*, *Random Forest*, dan *Gradient Boosting Classifier*. *Dataset* yang digunakan bersumber dari transaksi kartu kredit asli dengan proporsi *fraud* sebesar 0,17%, yang mencerminkan permasalahan ketidakseimbangan kelas. Untuk mengatasi hal ini, diterapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) dan transformasi fitur menggunakan *Principal Component Analysis* (PCA). Evaluasi dilakukan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan ROC-AUC. Hasil penelitian menunjukkan bahwa *Random Forest* dan *Gradient Boosting Classifier* menghasilkan performa terbaik dengan akurasi dan nilai ROC-AUC mendekati sempurna (ROC-AUC > 0.999), sedangkan *Logistic Regression* memberikan hasil yang sangat baik namun sedikit di bawah dua model lainnya. Meskipun demikian, nilai ROC-AUC yang hampir sempurna dapat mengindikasikan potensi *overfitting*, sehingga diperlukan validasi lebih lanjut pada *dataset* berbeda. Berbeda dengan penelitian sebelumnya yang hanya menggunakan satu algoritma, penelitian ini membandingkan tiga model sekaligus serta mengintegrasikan SMOTE dan PCA untuk meningkatkan performa deteksi. Implikasi praktis dari penelitian ini adalah model yang diusulkan dapat diimplementasikan pada sistem keuangan digital untuk membantu lembaga perbankan dalam mendeteksi penipuan secara *real-time* dan mengurangi potensi kerugian finansial.

Kata kunci: deteksi penipuan, kartu kredit, *gradient boosting*, *machine learning*, SMOTE

1. PENDAHULUAN

Perkembangan pesat teknologi digital telah meningkatkan penggunaan kartu kredit sebagai alat transaksi utama di berbagai sektor ekonomi [1]. Namun, seiring dengan kemajuan ini, kasus penipuan kartu kredit (*credit card fraud*) juga mengalami peningkatan signifikan, menimbulkan kerugian finansial yang besar bagi lembaga keuangan dan konsumen. Penipuan kartu kredit menjadi ancaman serius yang memerlukan perhatian khusus dalam sistem keuangan digital [2].

Deteksi penipuan secara manual tidak lagi efektif dalam menghadapi volume transaksi yang besar dan kompleksitas pola penipuan yang terus berkembang [3]. Oleh karena itu, pendekatan berbasis *machine learning* (ML) telah menjadi solusi yang menjanjikan untuk mengidentifikasi transaksi

Diterima 3 Juni 2025, Direvisi 27 Agustus 2025, Diterima untuk publikasi 26 September 2025

mencurigakan secara *real-time*. Studi sebelumnya menunjukkan bahwa integrasi *Neural Network* dengan teknik SMOTE dapat meningkatkan akurasi deteksi penipuan pada data yang tidak seimbang [4].

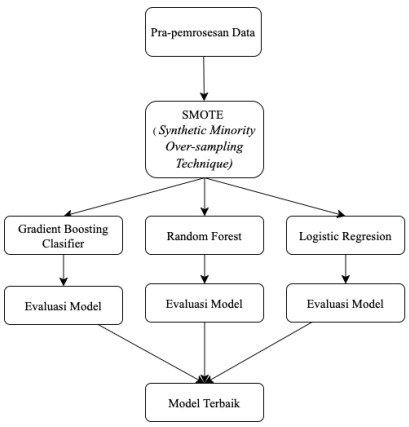
Berbagai algoritma ML telah diterapkan dalam deteksi penipuan kartu kredit, termasuk *Logistic Regression, Decision Trees, Random Forest,* dan XGBoost. Penelitian sebelumnya membahas penggunaan berbagai metode ini dalam mendeteksi penipuan dan menunjukkan bahwa pendekatan *ensemble learning* dapat meningkatkan performa model [5]. Selain itu, penggunaan teknik *deep learning* seperti *Convolutional Neural Networks* (CNN) dan *Gated Recurrent Units* (GRU) juga telah dieksplorasi untuk meningkatkan akurasi deteksi [6].

Meskipun telah banyak penelitian yang mengaplikasikan ML dalam deteksi penipuan kartu kredit, masih terdapat beberapa kendala, seperti kurangnya *dataset* yang representatif, tantangan dalam interpretabilitas model, dan kebutuhan akan sistem yang dapat beradaptasi dengan pola penipuan baru. Hal ini menekankan pentingnya pengembangan model yang tidak hanya akurat tetapi juga dapat diinterpretasikan dengan baik oleh pengguna akhir [7].

Berbeda dengan penelitian sebelumnya, penelitian ini tidak hanya menguji satu algoritma, tetapi juga membandingkan tiga model populer sekaligus dengan teknik SMOTE dan PCA untuk meningkatkan performa deteksi, serta menganalisis pentingnya fitur-fitur tertentu dalam proses deteksi. Dengan pendekatan ini, diharapkan dapat dikembangkan sistem deteksi penipuan yang lebih efektif dan efisien, serta dapat diimplementasikan dalam lingkungan keuangan digital di Indonesia.

2. METODE PENELITIAN

Alur penelitian ini mengacu diagram alir pada Gambar 1, peneltiain dimulai dari Pra-pemrosesan Data, penanganan *over sampling* dengan menggunakan SMOTE, pemilihan tiga model yaitu *Gradient Boosting Classifier, Random Forest,* dan *Logistic Regression*, selanjutnya dievaluasi dan didapatkan model terbaik.



Gambar 1. Alur Penelitian

2.1. *Dataset* dan Pra-pemrosesan

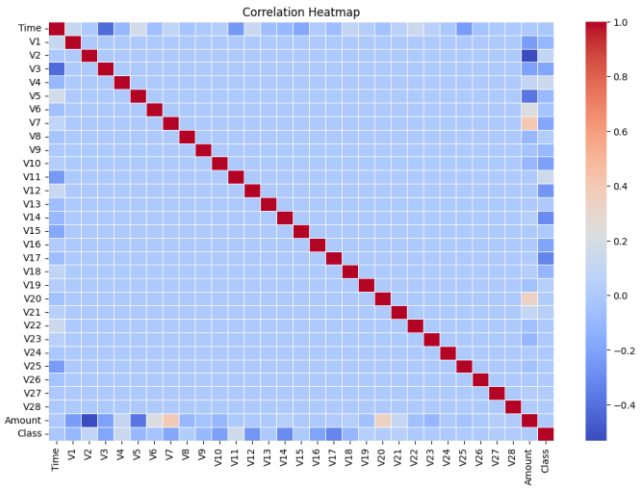
Penelitian ini menggunakan *dataset* transaksi kartu kredit yang diperoleh dari sumber terbuka, seperti *dataset* transaksi kartu kredit Eropa yang tersedia di Kaggle. *Dataset* tersebut terdiri dari 284.807 transaksi, dengan proporsi transaksi *fraud* sebesar 0,17%, mencerminkan ketidakseimbangan kelas yang signifikan [7]. Pra-pemrosesan data meliputi beberapa langkah penting:

a) Pembersihan Data, digunakan untuk menghapus duplikat dan menangani nilai yang hilang untuk memastikan kualitas data yang optimal [8].

b) Normalisasi, digunakan untuk menyesuaikan skala fitur numerik agar memiliki rentang yang seragam, sehingga algoritma *machine learning* dapat memproses data dengan lebih efektif [9].

c) Transformasi Fitur, menggunakan teknik *Principal Component Analysis* (PCA) digunakan untuk mengurangi dimensi data dan menghilangkan redundansi antar fitur [10].

Sebagai bagian dari pra-pemrosesan data, dilakukan analisis korelasi antar fitur menggunakan *heatmap* untuk memahami hubungan linear antar variabel dalam *dataset*.



Gambar 2. *Heatmap* Korelasi

Visualisasi pada Gambar 2 menunjukkan tingkat korelasi positif maupun negatif antar fitur, yang dapat membantu dalam identifikasi fitur-fitur redundan atau yang sangat berpengaruh terhadap target. Fitur dengan korelasi sangat tinggi satu sama lain dapat dipertimbangkan untuk pengurangan dimensi guna menghindari multikolinearitas yang dapat mempengaruhi performa model [11].

2.2. Penanganan Ketidakseimbangan Data

Ketidakseimbangan kelas antara transaksi normal dan *fraud* merupakan tantangan utama dalam deteksi penipuan kartu kredit. Untuk mengatasi hal ini, digunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) yang bertujuan untuk menyeimbangkan distribusi kelas dengan

cara menghasilkan sampel sintetis dari kelas minoritas [12]. Pendekatan ini telah terbukti efektif dalam meningkatkan performa model deteksi *fraud*.

2.3. Pemilihan dan Pelatihan Model

Beberapa algoritma *machine learning* digunakan dalam penelitian ini untuk membangun model deteksi penipuan, antara lain:

a. *Random Forest* (RF)

Random Forest merupakan metode *ensemble learning* yang membangun sejumlah pohon keputusan (*decision trees*) dan menggabungkannya untuk mendapatkan hasil klasifikasi yang lebih stabil dan akurat. Proses prediksi dilakukan melalui voting mayoritas dari semua pohon untuk klasifikasi, seperti pada formula (1):

$$\hat{y} = mode\big(h_1(x), h_2(x), h_3(x) \dots, h_T(x)\big) \qquad (1)$$

Dengan:

$h_T(x)$: adalah hasil prediksi dari pohon ke-t

T : adalah jumlah total pohon

$\hat{y}$: adalah hasil prediksi akhir dari *Random Forest*

Setiap pohon dibangun dengan memilih subset acak dari data pelatihan dan subset acak dari fitur. Tujuan utama pendekatan ini adalah mengurangi varian model sambil tetap menjaga bias tetap rendah, sehingga performa model lebih stabil dan tidak mudah *overfitting* [13].

b. *Gradient Boosting Classifier* (GBC)

Gradient Boosting Classifier yaitu algoritma ensemble yang membangun model prediktif secara bertahap dengan menggabungkan sejumlah pohon keputusan yang lemah untuk menghasilkan prediktor yang kuat. GBC dikenal memiliki kemampuan generalisasi yang baik serta efisien dalam menangani data tidak seimbang, sehingga sering digunakan dalam tugas-tugas klasifikasi seperti deteksi penipuan [14]. Model boosting bertujuan membentuk fungsi prediksi *F(x)* sebagai jumlah dari *M* model lemah *hm(x)*, yang biasanya berupa pohon keputusan dalam, dirumuskan sebagai formula (2):

$$F_M(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \qquad (2)$$

Dengan:

$h_m(x)$: pohon keputusan ke-*m*

γ_m : *learning rate* atau kontribusi pohon ke-*m*

M : jumlah total iterasi (jumlah pohon)

Pada setiap iterasi, algoritma menghitung residual dari prediksi sebelumnya, lalu menyesuaikan model baru berdasarkan arah negatif dari gradien fungsi *loss*.

c. *Logistic Regression* (LR)

Logistic Regression adalah model statistik yang digunakan untuk memprediksi probabilitas kejadian *fraud* berdasarkan fitur-fitur *input* [15]. Berbeda dengan regresi linear, LR memodelkan probabilitas suatu peristiwa (misalnya, *fraud* atau tidak) menggunakan fungsi sigmoid (*logistic function*) yang membatasi *output* antara 0 dan 1. Fungsi dasarnya dituliskan sebagai formula (3):

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n)}} \qquad (3)$$

Dengan:

$P(y = 1|x)$: probabilitas bahwa label *y* bernilai 1 (*fraud*) diberikan fitur *x*

β_0 : *intercept* (bias)

β_1 : koefisien regresi untuk fitur ke-i

x_1 : nilai dari fitur ke-i

e : bilangan eksponensial

Model ini meminimalkan fungsi *log-loss* atau *cross-entropy* sebagai fungsi kerugiannya pada formula (4):

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_1 \log(\hat{y_i}) + (1 - y_1)\log(1 - \hat{y_i})] \qquad (4)$$

Dengan:

$\hat{y_i}$: probabilitas prediksi dari model

y_i : label aktual (0 atau 1)

N : jumlah sampel

Keunggulan LR adalah interpretabilitas koefisiennya yang merepresentasikan pengaruh setiap fitur terhadap probabilitas output. Namun, LR memiliki keterbatasan dalam menangani hubungan *non-linear* antar fitur [15].

2.4. Evaluasi Model

Evaluasi kinerja model dilakukan menggunakan metrik-metrik berikut:

a. Akurasi

Akurasi digunakan untuk mengukur seberapa sering model membuat prediksi yang benar dari keseluruhan prediksi yang dilakukan. Nilai akurasi dihitung dengan membandingkan jumlah prediksi benar (baik positif maupun negatif) dengan jumlah total data. Meskipun akurasi mudah dihitung dan dipahami, metrik ini tidak ideal jika data sangat tidak seimbang, seperti pada kasus deteksi penipuan, di mana model bisa terlihat "akurat" hanya karena mendominasi prediksi kelas mayoritas (*non-fraud*).

b. Presisi

Presisi adalah rasio antara jumlah prediksi positif yang benar (*True Positive*) terhadap semua prediksi yang diklasifikasikan sebagai positif (*True Positive + False Positive*). Metrik ini mengukur ketepatan model dalam mengidentifikasi kasus positif (*fraud*). Presisi menjadi penting ketika biaya kesalahan positif tinggi, misalnya memblokir transaksi yang sah.

c. *Recall* (Sensitivitas)

Recall atau sensitivitas mengukur seberapa banyak dari total kasus positif (*fraud*) yang berhasil ditangkap oleh model. Ini adalah rasio antara *True Positive* terhadap jumlah seluruh kasus aktual positif (*True Positive + False Negative*). Metrik ini sangat penting ketika tujuan utama adalah mendeteksi sebanyak mungkin kasus penipuan**,** walaupun mungkin mengorbankan beberapa kesalahan dalam bentuk *false positive*.

d. F1-Score

F1-Score merupakan rata-rata harmonik antara presisi dan *recall*, yang memberikan keseimbangan antara keduanya. Nilai *F1-Score* yang tinggi menunjukkan bahwa model tidak hanya tepat dalam prediksi positif, tetapi juga sensitif terhadap kasus-kasus positif yang nyata. *F1-Score* sangat berguna ketika kita menghadapi *dataset* yang tidak seimbang dan ingin menghindari penilaian yang bias hanya berdasarkan satu metrik.

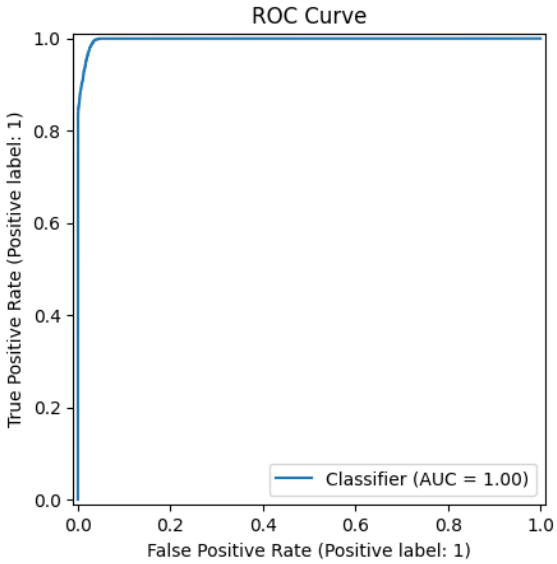
e. *Area Under the Curve* (AUC)

AUC adalah metrik yang berasal dari kurva ROC (*Receiver Operating Characteristic*), yang menggambarkan *trade-off* antara *True Positive Rate* (*recall*) dan *False Positive Rate* pada berbagai ambang batas klasifikasi. Nilai AUC berkisar antara 0 hingga 1. Semakin dekat ke 1, semakin baik kemampuan model dalam membedakan antara kelas *fraud* dan *non-fraud*. AUC sangat berguna karena tidak bergantung pada *threshold* tertentu, sehingga memberikan gambaran yang lebih komprehensif terhadap performa model. Metrik-metrik ini memberikan gambaran menyeluruh mengenai performa model dalam mendeteksi penipuan kartu kredit.

3. HASIL DAN PEMBAHASAN

3.1 Evaluasi Model

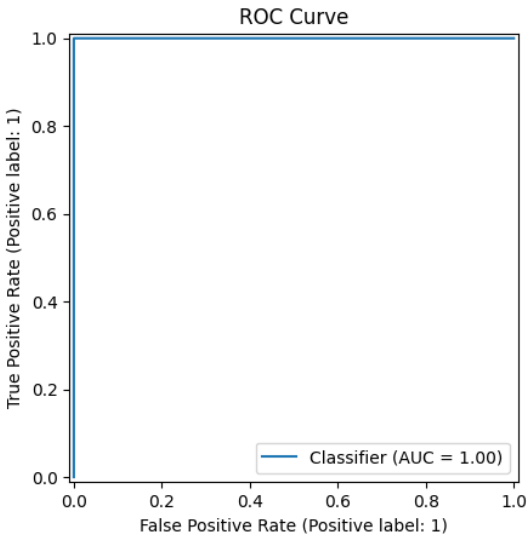
a. Logistic Regression



Gambar 3. ROC *Curve Logistic Regression*

Berdasarkan Gambar 3, dapat disimpulkan bahwa model *Logistic Regression* memiliki kemampuan klasifikasi yang sangat baik kurva terlihat sangat dekat dengan titik (0,1), menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membedakan antara transaksi *fraud* dan *non-fraud*. AUC sebesar 1.00 menandakan prediksi yang hampir sempurna.

b. Random Forest

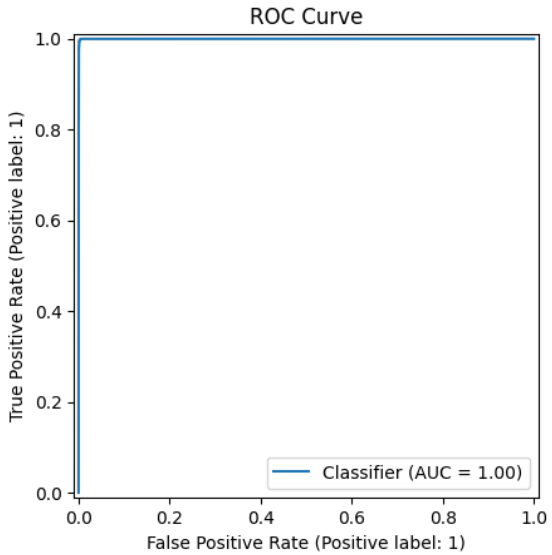


Gambar 4. *ROC Curve Random Forest*

Berdasarkan Gambar 4, dapat disimpulkan bahwa model *Random Forest* menunjukkan performa klasifikasi yang sangat optimal dengan nilai Area *Under the Curve* (AUC) sebesar 1.00. Kurva tampak sangat tajam dan langsung naik ke titik maksimum, menandakan bahwa model ini hampir tidak membuat kesalahan dalam klasifikasi. Ini sesuai dengan karakteristik *Random Forest* yang kuat dalam menangani variabel *non-linear* dan memiliki mekanisme internal yang membuatnya tahan terhadap *overfitting*. Hasil

ini mendukung hasil metrik evaluasi, di mana akurasi dan *F1-Score*-nya juga berada pada angka maksimal.

c. Gradient Boosting Classifier



Gambar 5. ROC *Curve Gradient Boost Classifier*

Berdasarkan Gambar 5, dapat disimpulkan bahwa model *Gradient Boosting Classifier* memiliki kinerja klasifikasi yang sangat tinggi. Sama seperti dua model sebelumnya, kurva menunjukkan nilai AUC = 1.00, mencerminkan performa klasifikasi yang sangat presisi. GBC unggul karena proses *boosting* memungkinkan model terus memperbaiki kesalahan sebelumnya secara iteratif, sehingga sangat efektif dalam menangani data tidak seimbang. Kemampuan GBC untuk menyesuaikan dengan data membuatnya unggul dalam banyak kasus *real-world fraud detection*.

Penelitian ini menguji tiga model *machine learning*, yaitu *Logistic Regression*, *Random Forest*, dan *Gradient Boosting Classifier*, untuk mendeteksi penipuan kartu kredit.

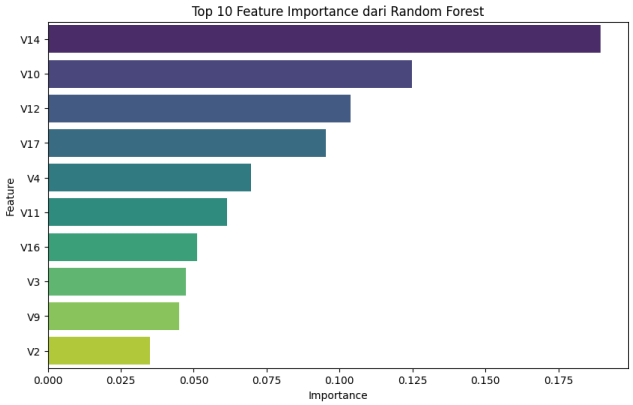
Tabel 1. *Evaluation Model*

Matrix	***Logistic Regression***	***Random Forest***	***Gradient Boosting***
Accuracy	0.98	1.00	1.00
Precision	0.97	1.00	1.00
Recall	0.98	1.00	1.00
F1-Score	0.98	1.00	1.00
ROC-AUC Score	0.9977	0.9999	0.9997

Berdasarkan hasil evaluasi pada tabel 1, ketiga model menunjukkan performa tinggi dengan akurasi di atas 98%. *Random Forest* memberikan hasil terbaik dengan akurasi dan ROC-AUC mendekati 100%, diikuti oleh *Gradient Boosting* dan *Logistic Regression* yang juga menunjukkan performa sangat baik.

3.2 Analisis *Feature Importance*

Gambar 6 merupakan hasil *feature importance* pada model *Random Forest* dan *Gradient Boosting*.



Gambar 6. *Feature Importance*

Pada Gambar 6, dapat disimpulkan bahwa fitur V14 merupakan variabel paling berpengaruh dalam mendeteksi transaksi *fraud*, diikuti oleh V10, V12, dan V17. Hal ini menunjukkan bahwa fitur tersebut memegang peranan penting dalam proses klasifikasi dan bisa menjadi fokus utama dalam pengembangan model selanjutnya.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa ketiga model *Logistic Regression*, *Random Forest*, dan *Gradient Boosting Classifier* mampu mendeteksi transaksi penipuan dengan performa sangat tinggi. *Logistic Regression* memberikan hasil yang baik dengan ROC-AUC 0,9977, sementara *Random Forest* dan *Gradient Boosting* mencapai nilai hampir sempurna dengan ROC-AUC di atas 0,999.

Dominasi dua model ensemble ini menegaskan bahwa pendekatan *ensemble learning* lebih unggul dalam menangani ketidakseimbangan data. Namun, hasil yang terlalu sempurna juga mengindikasikan potensi *overfitting*, terutama akibat penggunaan SMOTE yang menghasilkan data sintetis ideal serta kompleksitas model *ensemble* yang cenderung menyesuaikan data pelatihan secara berlebihan. Oleh karena itu, validasi silang dan pengujian pada *dataset* berbeda perlu dilakukan untuk memastikan kemampuan generalisasi model. Analisis *feature importance* memperlihatkan bahwa variabel V14, V10, V12, dan V17 berperan penting dalam membedakan transaksi *fraud* dan *non-fraud*. Hal ini dapat menjadi fokus dalam pengembangan model selanjutnya.

Dengan demikian, penelitian ini menegaskan efektivitas *Random Forest* dan *Gradient Boosting* untuk deteksi penipuan kartu kredit, sekaligus menawarkan *novelty* berupa perbandingan tiga algoritma dengan integrasi SMOTE dan PCA. Implikasi praktisnya, model berbasis *ensemble learning* berpotensi diimplementasikan pada sistem keuangan digital guna membantu lembaga perbankan

mendeteksi *fraud* secara *real-time*, meskipun evaluasi lanjutan tetap diperlukan

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam pelaksanaan penelitian ini. Secara khusus, penulis menyampaikan apresiasi kepada pengelola platform Kaggle atas penyediaan *dataset* transaksi kartu kredit yang menjadi dasar penelitian ini. Ucapan terima kasih juga disampaikan kepada dosen pembimbing dan rekan-rekan akademisi yang telah memberikan masukan serta motivasi selama proses penyusunan penelitian ini berlangsung. Tidak lupa, penulis berterima kasih kepada keluarga dan semua pihak yang tidak dapat disebutkan satu per satu atas dukungan moral dan spiritual yang telah diberikan.

DAFTAR PUSTAKA

[1] E. Constancio and K. D. Tania, "Penerapan Metode Supervised Learning dan Teknik Resampling untuk Prediksi Penipuan Transaksi Keuangan," Dec. 2024. [Online]. Available: https://repository.unsri.ac.id/162912/

[2] L. Theodorakopoulos, A. Theodoropoulou, A. Tsimakis, and C. Halkiopoulos, "Big Data-Driven Distributed Machine Learning for Scalable Credit Card Fraud Detection Using PySpark, XGBoost, and CatBoost," *Electronics (Basel)*, vol. 14, no. 9, p. 1754, 2025, doi: 10.3390/electronics14091754.

[3] S. Jiang, R. Dong, J. Wang, and M. Xia, "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," *Systems*, vol. 11, no. 6, pp. 234–241, 2023.

[4] M. Zhu, Y. Zhang, Y. Gong, C. Xu, and Y. Xiang, "Enhancing Credit Card Fraud Detection: A Neural Network and SMOTE Integrated Approach," *arXiv preprint arXiv:2405.00026*, 2024.

[5] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, "Credit Card Fraud Detection using Machine Learning: A Study," *arXiv preprint arXiv:2108.10005*, 2021.

[6] N. K. A. Dewi and L. P. Mahyuni, "Pemetaan bentuk dan pencegahan penipuan e-commerce," *E-Jurnal Ekonomi Dan Bisnis Universitas Udayana*, vol. 9, pp. 851–878, 2020.

[7] P. T. S. Ningsih, M. Gusvarizon, and R. Hermawan, "Analisis Sistem Pendeteksi Penipuan Transaksi Kartu Kredit dengan Algoritma Machine Learning," *Jurnal Teknologi Informatika Dan Komputer*, vol. 8, no. 2, pp. 386–401, 2022.

[8] S. Beigi and M.-R. Amin-Naseri, "Credit Card Fraud Detection using Data mining and Statistical Methods," *Journal of AI and Data Mining*, vol. 8, no. 2, pp. 149–160, 2020.

[9] M. Sholeh, D. Andayati, R. Yuliana Rachmawati, P. Studi Informatika, and F. Teknologi Informasi dan Bisnis, "Data Mining Model Klasifikasi Menggunakan Algoritma K-Nearest Neighbor Dengan Normalisasi Untuk Prediksi Penyakit Diabetes Data Mining Model Classification Using Algorithm K-Nearest Neighbor With Normalization For Diabetes Prediction," 2022.

[10] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *J Big Data*, vol. 9, no. 1, p. 24, 2022, doi: 10.1186/s40537-022-00573-8.

[11] V. B. Nguyen, K. G. Dastidar, M. Granitzer, and W. Siblini, "The Importance of Future Information in Credit Card Fraud Detection," in *Lecture Notes in Computer Science*, vol. 151, Springer, 2022, pp. 234–241.

[12] P. Gupta, A. Varshney, M. R. Khan, R. Ahmed, M. Shuaib, and S. Alam, "Unbalanced Credit Card Fraud Detection Data: A Machine Learning-Oriented Comparative Study of Balancing Techniques," *Procedia Comput Sci*, vol. 218, pp. 2575–2584, 2023.

[13] A. H. M. Aburbeian and H. I. Ashqar, "Credit Card Fraud Detection Using Enhanced Random Forest Classifier for Imbalanced Data," in *International Conference on Advances in Computing Research*, Springer, Cham, 2023, pp. 234–241.

[14] Y. F. Zhang, H. L. Lu, H. F. Lin, X. C. Qiao, and H. Zheng, "The Optimized Anomaly Detection Models Based on an Approach of Dealing with Imbalanced Dataset for Credit Card Fraud Detection," *Mobile Information Systems*, vol. 2022, pp. 234–241, 2022.

[15] M. M. Mijwil and I. E. Salem, "Credit Card Fraud Detection in Payment Using Machine Learning Classifiers," *Asian Journal of Computer and Information Systems*, vol. 8, no. 4, pp. 234–241, 2020.

.