



Evaluation of Generative Augmentation and Oversampling Across Model Architectures for Credit Card Fraud Detection

Evaluasi Augmentasi Generatif dan Oversampling pada Berbagai Arsitektur Model untuk Deteksi Penipuan Kartu Kredit

**Gilang Bagus Ramadhan¹, Riska Dhenabayu^{2*}, Achmad Kautsar³,
Achmad Fitro⁴, Nurul Fajrin Ariyani⁵**

^{1,2,3,4}Program Studi Bisnis Digital, Universitas Negeri Surabaya, Indonesia

⁵Program Studi Teknik Informatika, Institut Teknologi Sepuluh Nopember, Indonesia

E-Mail: ¹gilang.23045@mhs.unesa.ac.id, ²riskadhenabayu@unesa.ac.id,
³achmadkautsar@unesa.ac.id, ⁴achmadfitro@unesa.ac.id, ⁵nf.ariyani@its.ac.id

Received Apr 23th 2026; Revised May 20th 2026; Accepted Jun 23th 2026; Available Online Jun 26th 2026

Corresponding Author: Riska Dhenabayu

Copyright © 2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

Credit card fraud detection in highly imbalanced datasets remains a critical challenge because classification models tend to be biased toward the majority class, reducing their ability to identify fraudulent transactions. Traditional oversampling techniques, such as Synthetic Minority Over-sampling Technique (SMOTE), are widely used but often fail to capture the complex distribution of tabular data. In contrast, generative approaches such as Conditional Tabular Generative Adversarial Network (CTGAN) offer the potential to generate more realistic synthetic data, although their effectiveness across different classification models has not been comprehensively evaluated. This study aims to compare the effectiveness of four data augmentation methods SMOTE, Borderline-SMOTE, SMOTE-ENN, and CTGAN when applied to Extreme Gradient Boosting (XGBoost), TabTransformer, and Deep Autoencoder models. Performance was evaluated using metrics sensitive to the minority class, including Receiver Operating Characteristic–Area Under the Curve (ROC-AUC), Area Under the Precision–Recall Curve (AUPRC), and other supporting metrics. The results show that XGBoost combined with Borderline-SMOTE achieved the best precision–recall balance, with an F1-score of 0.9269 and an AUPRC of 0.8306, while TabTransformer combined with CTGAN demonstrated the highest discriminative capability, achieving a ROC-AUC of 0.9846 and a precision of 0.9795. These findings indicate that the effectiveness of data augmentation depends on the compatibility between the characteristics of synthetic data and the learning mechanism of the classification model. This study provides a comprehensive comparative evaluation of data augmentation approaches and offers practical insights for developing robust fraud detection systems in highly imbalanced environments.

Keywords: Class Imbalance, Credit Card Fraud Detection, CTGAN, SMOTE, XGBoost, TabTransformer

Abstrak

Deteksi penipuan kartu kredit pada data dengan ketidakseimbangan kelas ekstrem merupakan tantangan karena model klasifikasi cenderung bias terhadap kelas mayoritas sehingga menurunkan kemampuan mendeteksi transaksi penipuan. Metode oversampling tradisional seperti Synthetic Minority Over-sampling Technique (SMOTE) banyak digunakan, tetapi sering belum mampu merepresentasikan kompleksitas distribusi data tabular. Sebaliknya, pendekatan generatif seperti Conditional Tabular Generative Adversarial Network (CTGAN) berpotensi menghasilkan data sintetis yang lebih realistis, meskipun evaluasinya pada berbagai model klasifikasi masih terbatas. Penelitian ini bertujuan membandingkan efektivitas empat metode augmentasi data, yaitu SMOTE, Borderline-SMOTE, SMOTE-ENN, dan CTGAN, pada model Extreme Gradient Boosting (XGBoost), TabTransformer, dan Deep Autoencoder. Evaluasi dilakukan menggunakan metrik yang sensitif terhadap kelas minoritas, meliputi Receiver Operating Characteristic–Area Under the Curve (ROC-AUC), Area Under the Precision–Recall Curve (AUPRC), serta metrik pendukung lainnya. Hasil menunjukkan bahwa kombinasi XGBoost dan Borderline-SMOTE memberikan keseimbangan precision–recall terbaik dengan F1-score 0,9269 dan AUPRC 0,8306, sedangkan TabTransformer dengan CTGAN mencapai kemampuan diskriminasi tertinggi dengan ROC-AUC 0,9846 dan precision 0,9795. Temuan ini menunjukkan bahwa efektivitas augmentasi bergantung pada kesesuaian karakteristik data sintetis dengan mekanisme pembelajaran model serta memberikan implikasi praktis bagi pengembangan sistem deteksi penipuan yang andal pada data tidak seimbang.

Kata Kunci: CTGAN, Deteksi Penipuan Kartu Kredit, Ketidakseimbangan Kelas, SMOTE, XGBoost, TabTransformer



1. PENDAHULUAN

Digitalisasi masif di sektor perbankan memicu lonjakan volume transaksi elektronik sekaligus eskalasi ancaman keamanan siber, terutama penipuan kartu kredit. Laporan industri menunjukkan kerugian global akibat penipuan kartu kredit mencapai 33,83 miliar dolar per tahun [1]. Kondisi ini mendorong kebutuhan mendesak akan sistem deteksi penipuan mutakhir yang adaptif [2][3]. Dalam konteks tersebut, ketidakseimbangan kelas yang sangat tinggi menjadi masalah krusial. Rasio transaksi penipuan yang sangat rendah, di bawah 0,2% dari total transaksi, menyebabkan model konvensional terpacu pada kelas mayoritas. Model dengan mudah mencapai akurasi tinggi pada klasifikasi normal namun gagal mendeteksi penipuan minoritas sehingga menjadi fenomena *accuracy paradox*. Dominasi kelas mayoritas ini memaksa model mengabaikan pola minoritas, sehingga meningkatkan jumlah *false negative* yang sangat merugikan secara operasional [4][5]. Oleh karena itu, riset terkini berfokus pada meningkatkan sensitivitas model terhadap anomali minoritas tanpa mengorbankan presisi terhadap transaksi normal [6].

Merespons problematika tersebut, literatur terkini menyoroti konvergensi tiga paradigma komputasional utama. Pertama, *Extreme Gradient Boosting* (XGBoost) hadir sebagai algoritma tangguh pengolah data tabular berkat kemampuannya memformulasikan pemisahan nonlinier melalui mekanisme *gradient boosting*, yang terbukti empiris mampu mempertahankan *recall* dan presisi tinggi pada distribusi timpang [7]. Kedua, TabTransformer memperkenalkan arsitektur *deep learning* dengan mekanisme *self-attention* guna mengurai keterkaitan laten antarvariabel melalui representasi *embedding* kontekstual. Konsep revolusioner ini efektif mengekstraksi fitur esensial dalam dimensi dinamis [8]. Ketiga, arsitektur *Deep Autoencoder* berbasis *unsupervised learning* disertakan sebagai fondasi *benchmarking* komparatif. Model ini mempelajari distribusi alami transaksi normal secara mandiri untuk mendeteksi anomali melalui deviasi rekonstruksi tanpa manipulasi data. Ketiga pendekatan ini secara harmonis mewakili trikotomi metodologis komprehensif: eksplorasi batas spasial deterministik, abstraksi probabilitas laten, dan deteksi deviasi berbasis rekonstruksi [9].

Strategi penanganan dataset tidak seimbang idealnya terbagi menjadi tiga paradigma utama: tingkat algoritma, metrik evaluasi, dan data. Karena modifikasi tingkat algoritma seringkali kaku terhadap arsitektur model tertentu, strategi tingkat data, khususnya *oversampling*, jauh lebih direkomendasikan. Pendekatan ini terbukti sangat fleksibel dan efektif mencegah hilangnya informasi krusial [2]. Dalam praktiknya, teknik *oversampling* tradisional seperti *Synthetic Minority Over-sampling Technique* (SMOTE) banyak digunakan guna memperluas *manifold* minoritas melalui interpolasi spasial [4]. Namun, metode konvensional ini belum mampu menangkap kompleksitas hubungan antarvariabel pada data tabular nyata. Sebagai alternatif, metode generatif seperti *Conditional Tabular Generative Adversarial Network* (CTGAN) menawarkan kemampuan memodelkan distribusi multivariat yang lebih dinamis [10]. Meskipun demikian, mayoritas studi saat ini masih mengevaluasi teknik augmentasi dan algoritma secara terpisah. Interaksi kausal antara karakteristik data sintetis dengan bias arsitektur model seperti XGBoost, TabTransformer, dan Deep Autoencoder belum pernah diungkap secara mendalam [4]. Terkait dengan paradigma evaluasi, penggunaan metrik Akurasi dan *Receiver Operating Characteristic* (ROC) - *Area Under the Curve* (AUC) dapat menyesatkan pada ketidakseimbangan ekstrem [11]. Di sisi lain, perbandingan langsung antara teknik augmentasi *supervised* SMOTE, CTGAN, dengan pendekatan deteksi anomali *unsupervised* Deep Autoencoder masih minim dieksplorasi [12].

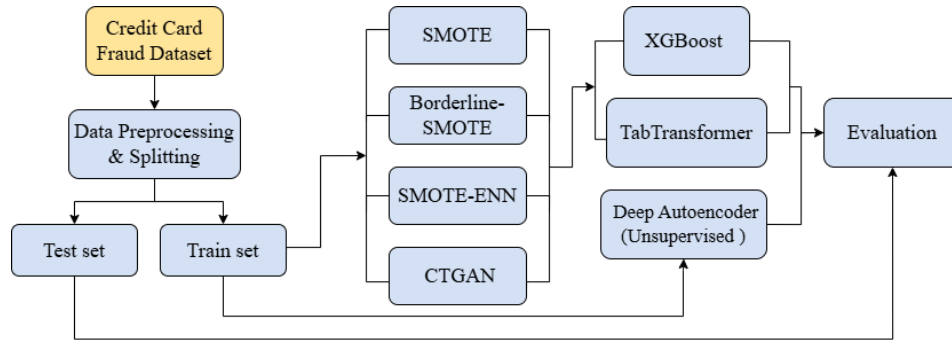
Dari uraian di atas, terlihat bahwa belum ada kajian terintegrasi yang menganalisis augmentasi spasial versus generatif dengan beragam model klasifikasi terkini, disertai metrik evaluasi yang tepat untuk minoritas (AUPRC). Kesenjangan ilmiah ini penting karena efektivitas augmentasi tergantung pada kesesuaian topologi data sintetis dengan paradigma pembelajaran model. Oleh karena itu, tujuan penelitian ini adalah untuk mengetahui apakah augmentasi generatif (CTGAN) lebih efektif dibandingkan *oversampling* tradisional (SMOTE dan variannya) dalam meningkatkan deteksi fraud pada model XGBoost, TabTransformer, dan Deep Autoencoder. Selain itu, penelitian ini mengeksplorasi bagaimana karakteristik data sintetis memengaruhi performa tiap model dalam konteks kegagalan minoritas [12].

Mengacu pada serangkaian kesenjangan tersebut, studi ini menyusun kerangka *benchmarking* terpadu yang membandingkan efek berbagai teknik augmentasi (SMOTE, Borderline-SMOTE, SMOTE-ENN, CTGAN) pada tiga arsitektur model berbeda (XGBoost, TabTransformer, Deep Autoencoder). Evaluasi kinerja melibatkan metrik yang sensitif terhadap kelas minoritas, yaitu *Receiver Operating Characteristic* (ROC) - *Area Under the Curve* (AUC) dan *Area Under Precision-Recall Curve* (AUPRC), beserta metrik pendukung lainnya yaitu Akurasi, Presisi, Recall dan F1-Score. Protokol eksperimen juga mempertimbangkan trade-off konsekuensi *false positive/false negative* dalam konteks operasional perbankan [13].

Kontribusi penelitian ini meliputi analisis komparatif komprehensif lintas pendekatan augmentasi (*oversampling* tradisional vs generatif) dan arsitektur model (*tree ensemble* vs *deep learning*) untuk deteksi penipuan pada data dengan ketidakseimbangan kelas yang tinggi. Kontribusi selanjutnya adalah identifikasi fenomena *distribution - model alignment*, yaitu kecocokan topologi data sintetis dengan bias model, sebagai faktor kunci efektivitas augmentasi. Serta penyusunan panduan praktis untuk merancang sistem deteksi penipuan yang *robust* dengan menggabungkan sensitivitas terhadap penipuan dan minimisasi gangguan pada transaksi normal.

2. METODOLOGI PENELITIAN

Penelitian ini disusun menggunakan pendekatan eksperimental yang terstruktur melalui integrasi tahapan utama yang mencakup prapemrosesan data, augmentasi berbasis oversampling dan generatif, pelatihan model, serta evaluasi kinerja secara komprehensif.



Gambar 1. Diagram alir yang diusulkan

Berdasarkan kerangka kerja komputasional yang diilustrasikan pada Gambar 1, prosedur diawali dengan akuisisi dataset transaksi kartu kredit, yang kemudian dilanjutkan oleh analisis data eksploratif guna mengidentifikasi topologi distribusi dan pola anomali spasial. Dataset selanjutnya dipartisi menggunakan teknik stratified sampling untuk menghasilkan partisi data latih (*Train set*) dan data uji (*Test set*) guna mempertahankan distribusi probabilitas kelas aslinya. Transformasi normalisasi fitur dieksekusi untuk mengunci stabilitas numerik sebelum intervensi sintesis data dilakukan. Empat skenario augmentasi independen meliputi *Synthetic Minority Oversampling Technique* (SMOTE), Borderline-SMOTE, SMOTE-ENN, serta *Conditional Tabular Generative Adversarial Network* (CTGAN) diaplikasikan secara eksklusif pada partisi data latih. Ruang laten termodifikasi tersebut selanjutnya diintegrasikan pada fase pelatihan model klasifikasi XGBoost dan TabTransformer, untuk dikomparasikan dengan arsitektur *unsupervised Deep Autoencoder*. Pada tahap akhir, kapabilitas deteksi dari seluruh konfigurasi divalidasi menggunakan data uji murni (*Test set*) dengan mengandalkan *Area Under the Precision-Recall Curve* (AUPRC) sebagai metrik utama penentu kinerja pada kondisi ketidakseimbangan kelas ekstrem.

2.1 Eksplorasi dan Prapemrosesan Data

Dataset yang digunakan merupakan dataset publik transaksi kartu kredit yang telah banyak diadopsi sebagai *benchmark* dalam penelitian deteksi fraud berbasis *machine learning*. Keseluruhan data (N_{total}) mencakup 284.807 transaksi dengan hanya 492 transaksi fraud (N_{fraud}), sehingga tingkat ketimpangan kelas direpresentasikan melalui rasio (R) yang dapat dilihat pada persamaan 1.

$$R = \frac{N_{fraud}}{N_{total}} = \frac{492}{284807} \approx 0.0017 \quad (1)$$

Nilai rasio tersebut menunjukkan bahwa probabilitas kemunculan fraud berada pada tingkat yang sangat rendah, menyebabkan representasi geometris kelas minoritas terdistribusi secara terisolasi di dalam manifold data berdimensi tinggi. Sebagian besar atribut numerik telah ditransformasikan menggunakan *Principal Component Analysis* (PCA) menjadi variabel laten V1–V28, sementara fitur *Time* dan *Amount* dipertahankan dalam bentuk asli. Transformasi PCA diterapkan untuk menjaga kerahasiaan informasi finansial sekaligus mengurangi redundansi linier melalui pembentukan ruang fitur ortogonal. Konsekuensinya, proses klasifikasi tidak lagi bergantung pada hubungan korelatif eksplisit antarvariabel, melainkan pada kemampuan model dalam mengenali pola laten yang tersembunyi pada ruang representasi multivariat. [14].

Karakteristik statistik dataset selanjutnya dievaluasi melalui tahapan *Exploratory Data Analysis* (EDA) guna mengidentifikasi stabilitas distribusi numerik sebelum proses pemodelan dilakukan. Distribusi fitur *Amount* memperlihatkan tingkat *skewness* yang tinggi akibat dominasi transaksi bernilai kecil disertai keberadaan sejumlah kecil transaksi ekstre. Kondisi tersebut mengindikasikan variansi yang tidak homogen serta sensitivitas distribusi terhadap *outlier* [15], sehingga normalisasi utama diterapkan menggunakan pendekatan *Robust Scaling*. Pada formulasi ini, nilai fitur awal (X) dinormalisasi menggunakan median (Q_2), kuartil pertama (Q_1), dan kuartil ketiga (Q_3), sehingga nilai hasil normalisasi (X_{scaled}) yang dapat dilihat pada persamaan 2.

$$X_{scaled} = \frac{X - Q_2}{Q_3 - Q_1} \quad (2)$$

Sebagai pembanding, standardisasi berbasis *Z-score* turut diimplementasikan untuk menjaga konsistensi skala fitur dan meningkatkan stabilitas optimisasi algoritma berbasis gradien. Setelah tahapan normalisasi diselesaikan, dataset dipartisi menggunakan *Stratified Sampling* dengan rasio 80:20 guna mempertahankan distribusi probabilitas kelas pada data pelatihan dan pengujian [15]. Pemisahan dilakukan sebelum augmentasi diterapkan untuk mencegah kebocoran data, sehingga seluruh proses sintesis hanya dijalankan pada subset pelatihan, sedangkan data uji dipertahankan sebagai representasi distribusi nyata sebelum memasuki tahap augmentasi sintesis.

2.2 Skenario Augmentasi Data

Ketimpangan distribusi transaksi finansial yang sangat ekstrem menyebabkan representasi geometris kelas minoritas tereduksi secara signifikan pada ruang fitur berdimensi tinggi. Untuk mengatasi kondisi tersebut, skenario augmentasi diterapkan melalui pendekatan oversampling berbasis kedekatan spasial yang mencakup Synthetic Minority Oversampling Technique (SMOTE), Borderline-SMOTE, dan SMOTE-ENN [4]. Pada mekanisme SMOTE, sampel sintetis baru (x_{new}) dibangkitkan melalui interpolasi linier antara sampel minoritas (x_i) dan tetangga terdekatnya (x_{nn}), dengan faktor interpolasi acak (λ) yang berada pada rentang $[0, 1]$. Formulasi tersebut dapat dilihat pada persamaan 3.

$$x_{new} = x_i + \lambda(x_{nn} - x_i), \lambda \in [0, 1] \quad (3)$$

Proses tersebut memungkinkan perluasan manifold minoritas secara kontinu sehingga kepadatan distribusi fraud meningkat pada ruang pembelajaran model. Meskipun demikian, mekanisme interpolasi linier berpotensi menghasilkan titik sintetis pada wilayah transisi antarkelas yang ambigu akibat tidak mempertimbangkan struktur batas keputusan secara eksplisit. Konsekuensinya, perluasan distribusi minoritas berisiko membentuk konektivitas spasial yang terlalu mekanistik apabila sensitivitas terhadap zona klasifikasi kritis tidak dikendalikan secara memadai selama proses sintesis data [16].

Untuk mempertahankan sensitivitas spasial terhadap area klasifikasi yang paling rentan, pendekatan Borderline-SMOTE diterapkan dengan memusatkan pembangkitan sampel sintetis pada titik minoritas yang berada di sekitar *decision boundary*. Tingkat kerawanan spasial suatu titik dihitung menggunakan rasio r , yang dapat dilihat pada persamaan 4.

$$r = \frac{N_{majority}}{k}, \quad (4)$$

dengan $N_{majority}$ merepresentasikan jumlah tetangga mayoritas dari total k tetangga terdekat. Melalui mekanisme tersebut, distribusi minoritas diperluas secara lebih terarah pada wilayah yang memiliki probabilitas kesalahan klasifikasi tertinggi sehingga batas diskriminatif antarkelas dapat dipertahankan secara lebih stabil [17]. Di sisi lain, pendekatan SMOTE-ENN diintegrasikan untuk mengombinasikan proses oversampling dengan mekanisme pembersihan distribusi lokal berbasis *Edited Nearest Neighbors*. Sampel yang memiliki inkonsistensi label terhadap mayoritas tetangganya dieliminasi guna mereduksi noise dan outlier lokal [18]. Integrasi kedua mekanisme tersebut memungkinkan distribusi minoritas diperkaya tanpa menghasilkan distorsi manifold yang berlebihan serta tetap mempertahankan kontinuitas spasial data asli.

Selain pendekatan interpolatif berbasis kedekatan geometris, augmentasi generatif diterapkan menggunakan Conditional Tabular Generative Adversarial Network (CTGAN) untuk merekonstruksi distribusi probabilistik data tabular secara lebih adaptif. Arsitektur ini terdiri atas dua komponen utama, yaitu *Generator* (G) yang bertugas menghasilkan data sintetis dan *Discriminator* (D) yang berfungsi membedakan data asli ($x \sim p_{data}$) dan data sintetis yang dibangkitkan dari variabel acak ($z \sim p_z$). Proses optimisasi kedua komponen tersebut dilakukan melalui skema permainan minimaks yang dapat dilihat pada persamaan 5.

$$\min_G \max_D E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p_z} [\log (1 - D(G(z)))] \quad (5)$$

Berbeda dengan oversampling konvensional, CTGAN tidak sekadar membentuk titik sintetis melalui interpolasi lokal, melainkan merekonstruksi estimasi distribusi multivariat pada manifold finansial berdimensi tinggi [10]. Untuk mempertahankan kontrol terhadap kelas minoritas, mekanisme *conditional vector* diterapkan sehingga proses generasi diarahkan secara spesifik pada label fraud [19]. Melalui kombinasi pendekatan interpolatif dan generatif tersebut, ruang fitur minoritas tidak hanya diperluas secara kuantitatif, tetapi turut direstrukturisasi agar memiliki kepadatan geometris dan kontinuitas probabilistik yang lebih stabil sebelum memasuki tahap pemodelan klasifikasi.

2.3 Pemodelan Klasifikasi

Tahap klasifikasi dilaksanakan menggunakan Extreme Gradient Boosting (XGBoost) sebagai representasi pendekatan *ensemble* berbasis pohon keputusan untuk menangani distribusi data finansial

berdimensi tinggi. Fungsi objektif algoritma dibangun melalui integrasi komponen kerugian prediktif $l(y_i, \hat{y}_i)$ antara nilai aktual (y_i) dan hasil prediksi ($\hat{y}_i$), serta komponen regularisasi kompleksitas model $\Omega(f_k)$ pada setiap pohon keputusan (f_k) sebanyak K iterasi. Formulasi fungsi objektif dapat dilihat pada persamaan 6.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (6)$$

Mekanisme *boosting* iteratif diterapkan melalui pembaruan residual secara bertahap sehingga kesalahan klasifikasi pada kelas minoritas dapat direduksi secara progresif. Parameter *scale_pos_weight* digunakan untuk menyeimbangkan kontribusi kelas fraud terhadap fungsi kerugian global, sedangkan *learning rate* diterapkan guna menjaga stabilitas pembentukan struktur *split* selama proses pembelajaran [20]. Pendekatan ini dipertahankan sebagai baseline utama karena mekanisme eksplorasi spasial lokal pada XGBoost dinilai lebih adaptif terhadap karakteristik data tabular finansial yang telah mengalami transformasi PCA dan augmentasi sintesis [12]. Sebagai pembanding berbasis *deep learning*, pemodelan turut dilakukan menggunakan TabTransformer untuk mengevaluasi efektivitas representasi laten pada data tabular. Setiap fitur diproyeksikan ke dalam ruang *embedding* berdimensi tinggi sebelum diproses melalui beberapa lapisan *Transformer encoder* [21]. Hubungan antarfitur dipelajari menggunakan mekanisme pada persamaan 7.

$$\text{self-attention: } Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

dengan Q , K , dan V masing-masing merepresentasikan *query*, *key*, dan *value* dari *embedding* fitur. Mekanisme tersebut memungkinkan dependensi nonlinier antaratribut dipelajari secara kontekstual sehingga pola penipuan tersembunyi dapat direpresentasikan melalui *manifold laten multivariat*. Berbeda dengan pendekatan berbasis pohon keputusan yang menitikberatkan pada pemisahan spasial deterministik, TabTransformer memanfaatkan pembentukan relasi global antarfitur untuk mengevaluasi pola anomali pada ruang fitur finansial ortogonal hasil transformasi PCA [19].

Selain pendekatan *supervised*, deteksi anomali berbasis *unsupervised reconstruction learning* juga dapat menjadi alternatif untuk diterapkan menggunakan Deep Autoencoder [12]. Tingkat anomali dihitung berdasarkan *Reconstruction Error* berbasis *Mean Squared Error* (MSE), dengan y_i sebagai nilai aktual pada atribut ke- i , $\hat{y}_i$ sebagai nilai hasil rekonstruksi model, dan n sebagai jumlah atribut yang dievaluasi. Formulasinya dapat dilihat pada persamaan 8.

$$MSE(y) = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (8)$$

Semakin besar deviasi rekonstruksi yang dihasilkan, semakin tinggi probabilitas suatu transaksi diklasifikasikan sebagai anomali. Seluruh model dilatih menggunakan kombinasi berbagai metode augmentasi pada subset pelatihan, sedangkan evaluasi dilakukan sepenuhnya pada data uji asli untuk menjaga validitas estimasi generalisasi [9]. Optimalisasi *hyperparameter* dilaksanakan melalui kombinasi *Grid Search* dan *Random Search* yang diintegrasikan dengan skema *k-fold cross-validation* guna meminimalkan bias estimasi dan menjaga stabilitas performa pada distribusi minoritas. Dengan demikian, probabilitas prediksi dan *reconstruction error* yang dihasilkan memerlukan instrumen evaluasi yang sensitif terhadap *extreme class imbalance* sebelum kualitas deteksi fraud dianalisis lebih lanjut [22].

2.4 Metrik Evaluasi Model

Evaluasi performa awal dilakukan menggunakan metrik berbasis *Confusion Matrix* yang mencakup *Precision*, *Recall*, dan *F1-Score* untuk mengukur sensitivitas model terhadap transaksi fraud serta kestabilan prediksi pada distribusi kelas mayoritas dan minoritas. Pada formulasi tersebut, *True Positive* (TP) merepresentasikan jumlah transaksi *fraud* yang berhasil diidentifikasi secara benar, *False Positive* (FP) menunjukkan jumlah transaksi normal yang keliru diklasifikasikan sebagai *fraud*, sedangkan *False Negative* (FN) menyatakan jumlah transaksi *fraud* yang gagal terdeteksi oleh model. Ketiga metrik tersebut dirumuskan pada persamaan 9.

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}, \quad F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (9)$$

Meskipun *Accuracy* turut dihitung sebagai indikator umum performa klasifikasi, metrik tersebut tidak dijadikan landasan utama evaluasi karena dominasi *True Negative* pada distribusi data yang sangat timpang berpotensi menghasilkan estimasi performa yang bias [13]. Pada kondisi *extreme class imbalance*, model masih dapat memperoleh nilai akurasi sangat tinggi meskipun sebagian besar transaksi fraud gagal diidentifikasi secara tepat. Konsekuensinya, interpretasi performa tidak dapat direduksi hanya pada keberhasilan klasifikasi kelas mayoritas, melainkan harus mempertimbangkan kemampuan model mempertahankan sensitivitas terhadap pola anomali tanpa memicu peningkatan kesalahan prediksi yang berlebihan pada transaksi normal [23].

Untuk mengatasi distorsi evaluasi tersebut, *Area Under the Precision-Recall Curve* (AUPRC) digunakan sebagai parameter utama dalam pengujian model. Secara matematis, perumusan AUPRC dapat dilihat pada persamaan 10.

$$AUPRC = \int_0^1 \text{Precision}(\text{Recall}) d\text{Recall} \quad (10)$$

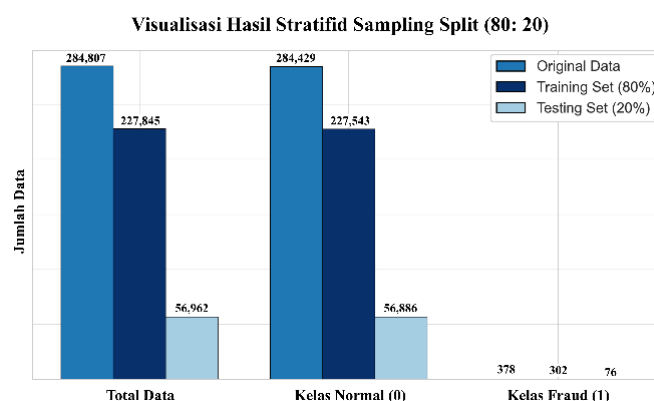
Berbeda dengan ROC-AUC yang masih dipengaruhi dominasi kelas mayoritas, AUPRC secara langsung merepresentasikan hubungan antara kemampuan mempertahankan *Recall* dan kestabilan *Precision* pada distribusi *extreme class imbalance* [11]. Pendekatan tersebut memungkinkan kualitas deteksi fraud diukur secara lebih realistis karena sensitivitas model terhadap transaksi anomali dapat dievaluasi bersamaan dengan kemampuan menekan *False Positive*. Dalam konteks operasional perbankan digital, keseimbangan tersebut menjadi sangat krusial karena peningkatan *False Negative* berpotensi memicu kebocoran aset finansial, sedangkan eskalasi *False Positive* dapat mengganggu pengalaman transaksi nasabah serta menurunkan kepercayaan terhadap sistem keamanan yang diterapkan [24].

3. HASIL DAN PEMBAHASAN

Bagian ini menguraikan luaran evaluasi arsitektur klasifikasi pada dataset penipuan kartu kredit. Pembahasan mencakup analisis distribusi data, topologi spasial, komparasi metrik kinerja, matriks kebingungan, hingga implikasi praktisnya bagi operasional perbankan.

3.1 Analisis Distribusi Dataset

Transformasi *Principal Component Analysis* (PCA) diterapkan untuk menjaga konfidensialitas atribut finansial sekaligus membentuk ruang fitur ortogonal melalui reduksi multikolinearitas antarvariabel. Konsekuensinya, hubungan linier eksplisit antarfitur menjadi sangat minimal sehingga proses klasifikasi tidak lagi bergantung pada korelasi langsung, melainkan pada kemampuan model dalam mengenali pola spasial laten pada manifold berdimensi tinggi. Kondisi tersebut menciptakan tantangan metodologis karena konsistensi struktur distribusi harus tetap dipertahankan selama proses augmentasi maupun pembelajaran model agar distorsi *decision boundary* tidak terbentuk pada ruang representasi sintetis.



Gambar 2. Hasil partisi data menggunakan *stratified sampling*

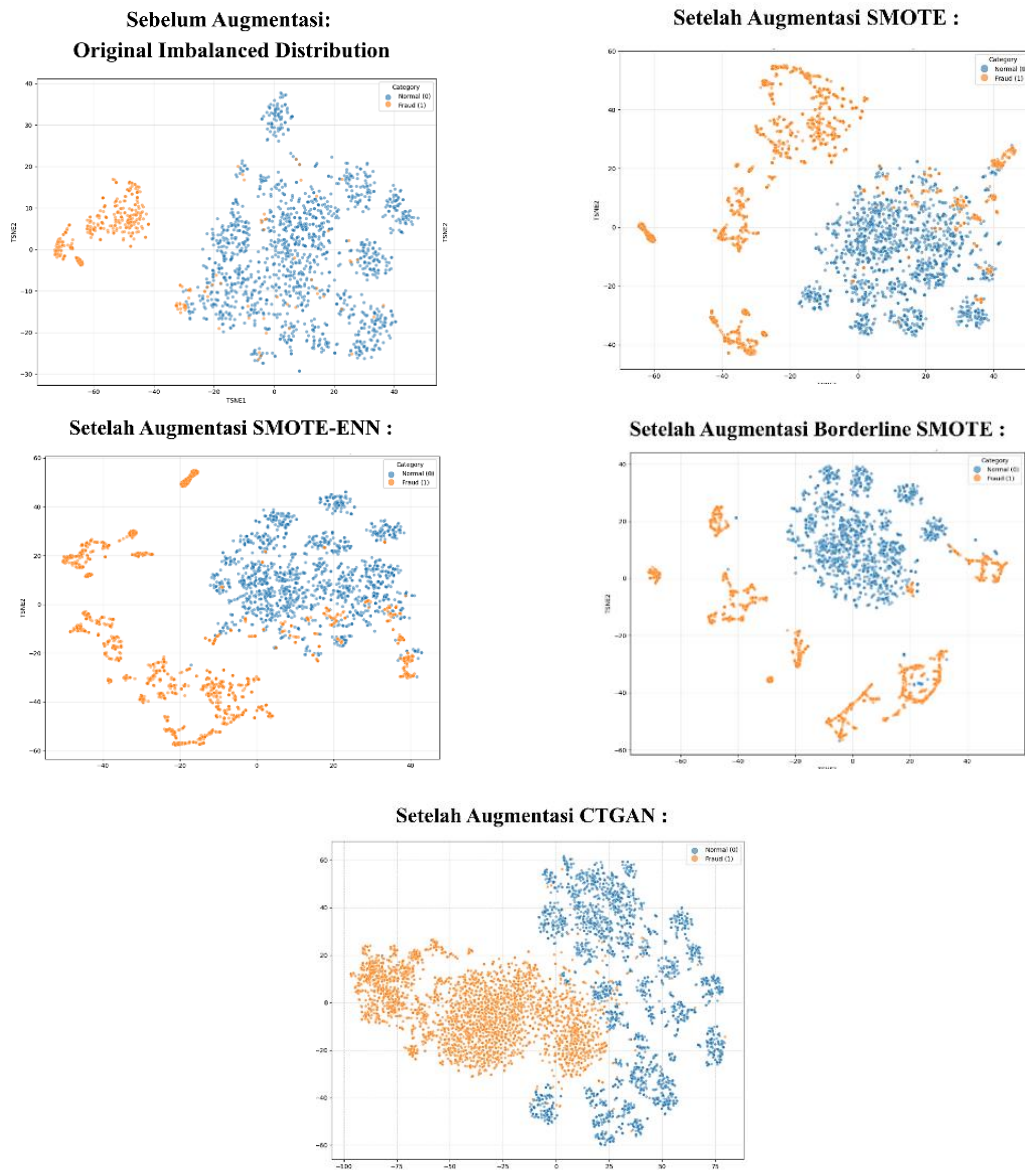
Untuk menjaga validitas eksperimental, partisi dataset dilakukan menggunakan *Stratified Random Sampling* dengan rasio 80:20 sebagaimana divisualisasikan pada Gambar 2. Strategi tersebut diterapkan guna mempertahankan proporsi distribusi kelas pada subset pelatihan dan pengujian sehingga risiko *data leakage* dapat diminimalkan secara sistematis. Distribusi yang diperoleh memperlihatkan dominasi absolut kelas normal, di mana transaksi fraud hanya merepresentasikan fraksi yang sangat kecil dibandingkan keseluruhan populasi transaksi sah. Ketimpangan struktural tersebut secara inheren mendorong fungsi objektif model menjadi bias mayoritas, sehingga sensitivitas terhadap pola anomali berpotensi mengalami degradasi signifikan. Oleh sebab itu, struktur distribusi yang sangat terisolasi tersebut mengharuskan dilakukannya pembedahan visual terhadap hubungan antarvariabel dan dinamika manifold laten sebelum proses klasifikasi dilaksanakan.

3.2 Korelasi dan Topologi Spasial Augmentasi

3.2.1. Analisis Topologi t-SNE

Validasi kualitas distribusi sintetis dilakukan menggunakan *t-Distributed Stochastic Neighbor Embedding* (t-SNE) sebagaimana ditunjukkan pada Gambar 3 untuk mengevaluasi integritas manifold data sebelum tahap klasifikasi. Proyeksi nonlinier ini tidak hanya dimanfaatkan sebagai instrumen eksploratif, tetapi

juga sebagai mekanisme verifikasi struktural dalam menilai kontinuitas spasial, kepadatan lokal, dan koherensi distribusi antara data sintetis dan ruang laten transaksi asli. Pada kondisi baseline tanpa augmentasi, kluster transaksi normal tampak sangat dominan, sedangkan titik-titik fraud tersebar sporadis pada beberapa wilayah manifold. Fenomena *spatial isolation* tersebut menyebabkan representasi geometris kelas minoritas melemah dan mendorong pembentukan fungsi keputusan yang cenderung bias terhadap kelas mayoritas.



Gambar 3. Visual Manifold Learning (t-SNE)

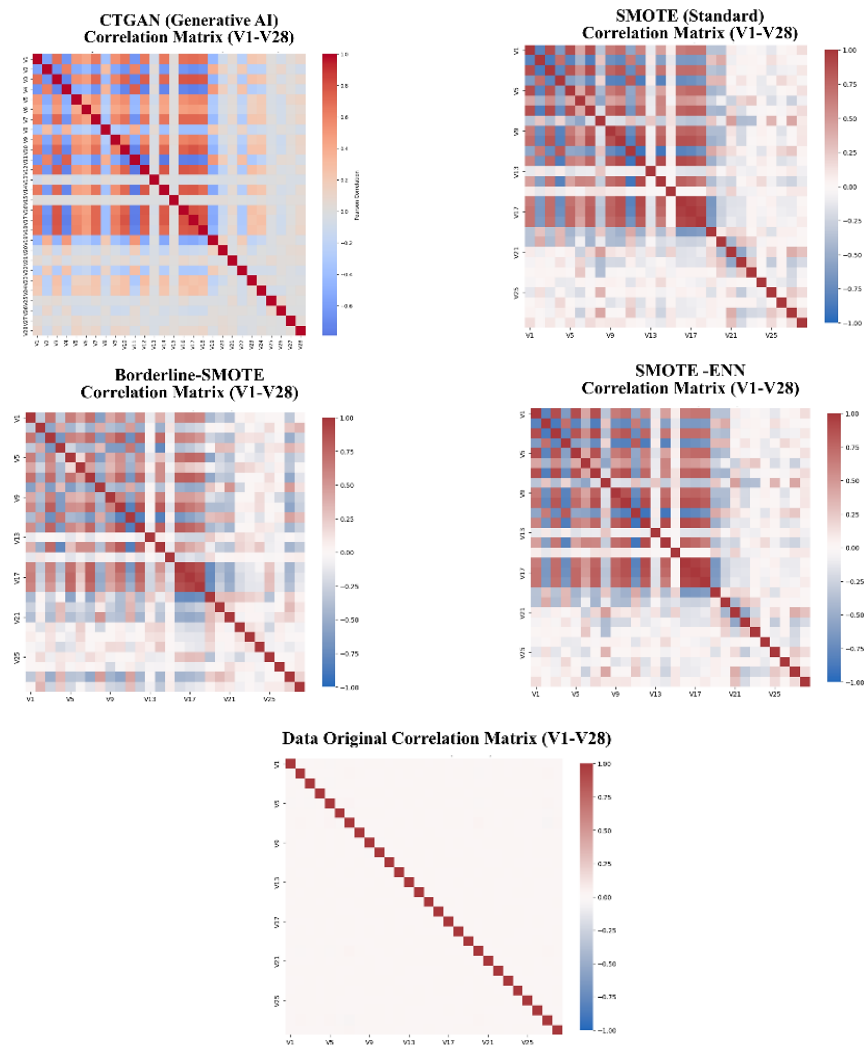
Setelah augmentasi diterapkan, evolusi topologi manifold memperlihatkan karakteristik spasial yang berbeda pada setiap pendekatan. SMOTE membentuk distribusi sintetis melalui interpolasi linier antar-*neighbor* minoritas sehingga muncul kecenderungan *linear bridging* pada ruang fitur. Meskipun cakupan wilayah fraud berhasil diperluas dan kepadatan lokal meningkat, konektivitas yang terbentuk masih bersifat mekanistik serta belum sepenuhnya merepresentasikan kompleksitas distribusi probabilistik asli. Di sisi lain, Borderline-SMOTE memperlihatkan kepadatan sintetis yang lebih selektif pada area *decision boundary*, sehingga perluasan manifold berlangsung lebih terarah. Sementara itu, SMOTE-ENN menghasilkan fenomena *clear moat effect* berupa pemisahan spasial yang lebih tegas akibat eliminasi sampel ambigu melalui mekanisme *Edited Nearest Neighbors*. Struktur tersebut meningkatkan separasi antarkelas, meskipun pada beberapa area turut membentuk batas keputusan yang relatif agresif terhadap variasi distribusi fraud nyata.

Berbeda dari pendekatan berbasis interpolasi geometris, CTGAN pada Gambar 3 memperlihatkan distribusi manifold yang lebih kontinu dan adaptif terhadap struktur laten data finansial. Kepadatan sintetis dibangun melalui estimasi densitas multivariat sehingga heterogenitas spasial dapat direpresentasikan secara lebih natural tanpa menghasilkan distorsi topologi yang signifikan terhadap distribusi transaksi normal. Oleh sebab itu, perubahan geometris manifold yang terbentuk memerlukan verifikasi lanjutan terhadap derajat

interaksi antarvariabel pembentuknya guna memastikan bahwa kontinuitas spasial yang dihasilkan tetap mempertahankan integritas statistik pada ruang fitur finansial.

3.2.2 Analisis Korelasi Fitur

Sebagai validasi struktural lanjutan, kualitas data sintetis dievaluasi melalui visualisasi matriks *correlation heatmap* terhadap fitur laten V1–V28 hasil transformasi *Principal Component Analysis* (PCA) sebagaimana ditunjukkan pada Gambar 4. Secara teoritis, transformasi PCA membentuk ruang ortogonal sehingga kovarians antarvariabel bernilai mendekati nol. Oleh karena itu, independensi linier diposisikan sebagai indikator utama dalam menilai retensi karakteristik statistik data. Untuk mengakomodasi masukan kritis, *heatmap* dataset asli sebelum augmentasi sengaja disertakan sebagai *baseline* acuan guna melacak kemunculan dependensi artifisial yang berpotensi merusak struktur manifold finansial asli.



Gambar 4. Matriks Korelasi (Heatmap)

Merujuk pada Gambar 4, *heatmap* dataset asli didominasi warna putih di luar diagonal utama, yang secara empiris mengonfirmasi keberhasilan PCA dalam menghilangkan korelasi linier. Ketika dibandingkan dengan *baseline* tersebut, keluarga metode interpolasi spasial memperlihatkan karakteristik visual yang merepresentasikan mekanisme kerjanya. Algoritma SMOTE konvensional melakukan interpolasi acak antar sampel minoritas sehingga membentuk blok korelasi linier samar yang tersebar merata. Sementara itu, Borderline-SMOTE yang memfokuskan sintesis pada sampel minoritas di sekitar *decision boundary* (*danger zone*) menciptakan penguatan korelasi lokal yang lebih padat dan terlokalisasi pada fitur tertentu. Di sisi lain, SMOTE-ENN menunjukkan intensitas warna yang relatif lebih bersih karena setelah fase interpolasi spasial dilakukan, algoritma ini menerapkan *Edited Nearest Neighbors* (ENN) untuk menghapus sampel bising atau tumpang tindih yang berada pada wilayah kelas mayoritas. Pola visual ketiga metode tersebut memperlihatkan bahwa manipulasi geometris berbasis *nearest neighbors* secara inheren memperkenalkan dependensi lokal baru pada ruang dimensi laten.

Berbeda dari pendekatan interpolatif tersebut, visualisasi *correlation heatmap* pada model CTGAN memperlihatkan pola korelasi dinamis yang jauh lebih mendekati *baseline* asli. Karena didasarkan pada mekanisme *conditional generation*, CTGAN mampu memodelkan distribusi probabilitas multivariat secara fleksibel dan adaptif tanpa menciptakan blok multikolinearitas artifisial yang masif di luar diagonal. Dominasi warna pada spektrum netral tetap terlihat pada keseluruhan matriks, yang mengindikasikan retensi independensi ruang laten hasil transformasi PCA yang optimal. Konsistensi antara visualisasi *heatmap* dan manifold *t-SNE* ini menegaskan bahwa kualitas augmentasi tidak hanya bertumpu pada perluasan kuantitas kelas minoritas, melainkan juga pada kapasitas mempertahankan integritas geometris data. Stabilitas struktural tersebut pada akhirnya menjamin validitas pengujian kapabilitas diskriminatif model klasifikasi pada tahapan evaluasi berikutnya.

3.3 Komparasi Kinerja Klasifikasi

3.3.1 Evaluasi Metrik performa

Evaluasi komputasional terhadap sebelas konfigurasi model yang mencakup *ensemble tree*, transformer tabular, dan *unsupervised anomaly detection* (Tabel 1) menunjukkan bahwa mayoritas konfigurasi. Namun, metrik ini sangat menyedapkan pada distribusi *extreme class imbalance*. Dominasi transaksi normal membuat kontribusi *True Negative* teramat besar, sehingga kegagalan intersepsi *fraud* tertutupi oleh tingginya proporsi prediksi benar kelas mayoritas. Fenomena mematikan ini memicu *accuracy paradox*, di mana model tampak superior secara numerik padahal sensitivitasnya terdegradasi parah. Pada konfigurasi *baseline* tanpa augmentasi, XGBoost mencatat akurasi (0,9994) dan *F1-Score* (0,8940), sedangkan TabTransformer hanya memperoleh *F1-Score* (0,8130). Penurunan sensitivitas ini membuktikan secara empiris bahwa sparsitas distribusi minoritas menyebabkan ruang representasi *fraud* kekurangan kepadatan lokal untuk membentuk batas keputusan diskriminatif yang stabil. Keterbatasan representasi minoritas ini tereskalasi menjadi jauh lebih parah pada pendekatan *Deep Autoencoder* berbasis *reconstruction learning*. Model *unsupervised* ini menghasilkan akurasi (0,9983), tetapi *F1-Score* anjlok drastis ke angka (0,7058).

Tabel 1. Performance Metrics

Algorithm	Acc	Pre	Recall	F1	ROC-AUC	AUPRC
XGBoost (Baseline)	0.9994	0.9464	0.8526	0.8940	0.9377	0.7566
TabTransformer (Baseline)	0.9989	0.8658	0.7735	0.8130	0.8795	0.5800
XGBoost + SMOTE	0.9993	0.8876	0.8998	0.8936	0.9611	0.8169
XGBoost + Borderline-SMOTE	0.9995	0.9577	0.8999	0.9269	0.9783	0.8306
XGBoost + SMOTE-ENN	0.9992	0.8665	0.9050	0.8848	0.9639	0.8006
XGBoost + CTGAN	0.9995	0.9669	0.8736	0.9151	0.9681	0.8144
TabTransformer + SMOTE	0.9747	0.5275	0.9243	0.5455	0.9650	0.6223
TabTransformer + Borderline-SMOTE	0.9893	0.5590	0.9106	0.6009	0.9351	0.6312
TabTransformer + SMOTE-ENN	0.9747	0.5275	0.9243	0.5455	0.9655	0.6114
TabTransformer + CTGAN	0.9995	0.9795	0.8737	0.9200	0.9846	0.8003
Deep Autoencoder	0.9983	0.7656	0.6681	0.7058	0.9661	0.3158

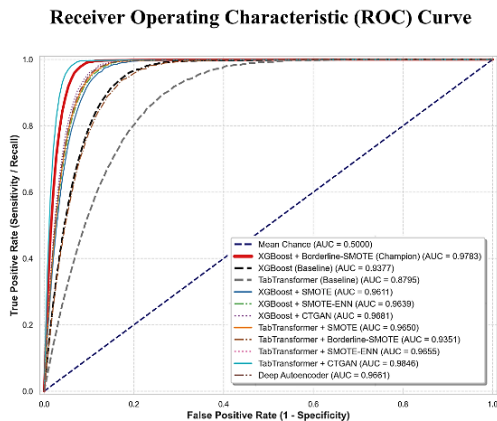
Secara konseptual, pendekatan rekonstruksi beroperasi dengan cara mempelajari distribusi dominan transaksi sah untuk mengidentifikasi anomali via deviasi *error*. Sayangnya, pola transaksi *fraud* modern sengaja dimanipulasi menyerupai transaksi normal sehingga distribusi *error* antarkelas mengalami tumpang tindih tinggi. Kegagalan separasi ini menegaskan bahwa tanpa intervensi augmentasi tingkat data, model kesulitan memisahkan kelas pada ruang fitur ortogonal hasil PCA. Oleh karena itu, sangat diperlukan keselarasan absolut antara karakteristik topologi data sintesis dengan mekanisme arsitektur model klasifikasi.

Perbaikan performa yang amat signifikan terbukti secara nyata saat karakteristik data sintesis dicocokkan dengan algoritma klasifikasi yang sangat relevan. Kombinasi XGBoost dan Borderline-SMOTE sukses membukukan *F1-Score* tertinggi yang mencapai (0,9269). Keberhasilan luar biasa ini terjadi karena Borderline-SMOTE mensintesis sampel di area kritis perbatasan batas keputusan, menciptakan kepadatan spasial deterministik yang sangat ideal bagi XGBoost yang membelah ruang fitur secara asimetris berbasis aturan pohon. Secara paralel, TabTransformer dan CTGAN sukses mendongkrak *F1-Score* menjadi (0,9200) dengan presisi superior (0,9795). Sinergi ini optimal karena CTGAN tidak sekadar melakukan interpolasi linier, melainkan memodelkan distribusi probabilitas multivariat yang kaya akan heterogenitas tersembunyi. Karakteristik probabilitas dinamis dari data ini merupakan bahan bakar yang tepat bagi mekanisme *self-attention* TabTransformer untuk mengurai relasi laten kompleks secara sangat akurat dan presisi.

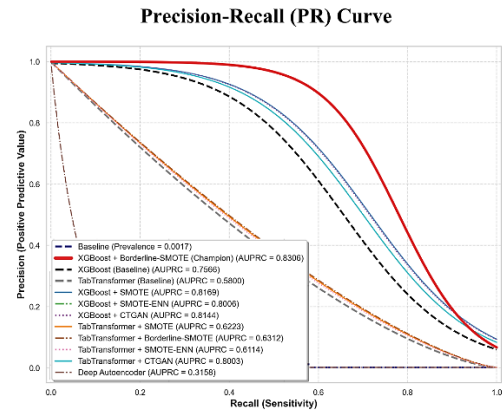
3.3.2 Analisis Kurva Kinerja

Visualisasi pada Gambar 5 menunjukkan bahwa hampir seluruh konfigurasi model menghasilkan lintasan Receiver Operating Characteristic (ROC) yang terkonsentrasi pada wilayah superior dengan nilai AUC di atas (0,90), termasuk beberapa model baseline tanpa augmentasi. Secara metodologis, kondisi tersebut merefleksikan keterbatasan ROC-AUC pada distribusi *extreme class imbalance*, karena dominasi transaksi

normal menyebabkan kontribusi *True Negative* menjadi sangat besar sehingga degradasi sensitivitas terhadap fraud tersamarkan oleh optimisme performa semu. Oleh sebab itu, interpretasi utama dialihkan menuju Kurva Precision-Recall (PR) pada Gambar 6 yang lebih sensitif dalam mengevaluasi kestabilan hubungan antara *precision* dan *recall* pada kelas minoritas. Berdasarkan lintasan PR Curve, konfigurasi XGBoost + Borderline-SMOTE memperlihatkan *plateau precision* paling stabil pada rentang *recall* menengah hingga tinggi dengan nilai AUPRC sebesar (0,8306). Stabilitas tersebut mengindikasikan bahwa sensitivitas terhadap fraud mampu dipertahankan tanpa menghasilkan eskalasi *false positive* secara agresif. Secara struktural, kondisi ini menunjukkan bahwa Borderline-SMOTE berhasil memfokuskan sintesis minoritas pada area *decision boundary*, sehingga mekanisme *recursive partitioning* pada XGBoost memperoleh batas pemisahan yang lebih diskriminatif terhadap zona transisi antarkelas.



Gambar 5. Kurva Receiver Operating Characteristic (ROC)



Gambar 6. Kurva Precision-Recall (PR)

Di sisi lain, integrasi CTGAN dengan XGBoost menghasilkan AUPRC sebesar (0,8144) yang masih berada di bawah kombinasi Borderline-SMOTE. Meskipun CTGAN mampu membentuk distribusi sintetis yang lebih kontinu dan adaptif terhadap manifold laten, kontinuitas probabilistik yang terlalu halus justru mengurangi ketegasan batas spasial yang dibutuhkan oleh mekanisme pemisahan berbasis pohon keputusan. Distribusi sintetis yang terlalu membaur dengan manifold asli menyebabkan orientasi optimal pembentukan *decision boundary* mengalami reduksi. Observasi lanjutan terhadap Kurva PR juga memperlihatkan bahwa seluruh varian TabTransformer tetap berada di bawah performa keluarga XGBoost meskipun telah memperoleh augmentasi CTGAN maupun SMOTE. Kondisi tersebut mengindikasikan keterbatasan mekanisme *Self-Attention* pada ruang fitur PCA yang telah mengalami dekorelasi ekstrem antarvariabel, sehingga representasi laten berbasis embedding tidak memperoleh sinyal relasional yang cukup kaya untuk dipelajari secara optimal. Dengan demikian, superioritas lintasan AUPRC pada kombinasi XGBoost + Borderline-SMOTE perlu diverifikasi lebih lanjut melalui dekomposisi empiris terhadap distribusi kesalahan prediksi aktual pada setiap kuadran klasifikasi.

3.4 Analisis Confusion Matrix

Interpretasi Confusion Matrix pada Tabel 2 menunjukkan bahwa *False Positive* (FP) dan *False Negative* (FN) tidak lagi dapat diposisikan sebagai residu kesalahan klasifikasi semata, melainkan sebagai indikator strategis yang secara langsung memengaruhi stabilitas operasional institusi finansial. Peningkatan *False Positive* merepresentasikan aktivasi alarm keamanan terhadap transaksi sah yang berpotensi memicu penolakan pembayaran, pembatasan akun sementara, hingga degradasi pengalaman pengguna secara sistemik. Dampak berantai dari kondisi tersebut dapat meningkatkan *customer churn*, menurunkan loyalitas nasabah, serta mengikis *digital trust* yang menjadi fondasi utama layanan perbankan digital berbasis transaksi real-time. Sebaliknya, *False Negative* merepresentasikan ancaman kebocoran aset yang jauh lebih destruktif karena transaksi fraud yang gagal diintersepsi dapat membuka ruang eskalasi serangan finansial secara berulang sekaligus meningkatkan risiko penalti kepatuhan dan penurunan kredibilitas institusi di hadapan regulator.

Tabel 2. Confusion Matrix

Algorithm	TN	FP	FN	TP
XGBoost (Baseline)	56643	8	28	67
TabTransformer (Baseline)	56632	19	43	52
XGBoost + SMOTE	56629	22	19	76
XGBoost + Borderline-SMOTE	56644	7	19	76
XGBoost + SMOTE-ENN	56623	28	18	77

Algorithm	TN	FP	FN	TP
XGBoost + CTGAN	56646	5	24	71
TabTransformer + SMOTE	55230	1421	12	83
TabTransformer + Borderline-SMOTE	56062	589	16	79
TabTransformer + SMOTE-ENN	55229	1421	13	83
TabTransformer + CTGAN	56648	3	24	71
Deep Autoencoder	56835	29	65	33

Berdasarkan hasil evaluasi empiris, konfigurasi TabTransformer + CTGAN menghasilkan tingkat *False Positive* paling rendah dengan hanya mencatatkan 3 kasus alarm palsu, sehingga kontinuitas transaksi sah mampu dipertahankan secara lebih stabil dibandingkan konfigurasi lainnya. Kondisi tersebut mengindikasikan bahwa distribusi sintesis berbasis CTGAN berhasil membentuk representasi probabilistik yang lebih adaptif terhadap pola transaksi normal, sehingga proses klasifikasi menjadi lebih selektif dalam memberikan label fraud. Akan tetapi, reduksi alarm palsu tersebut tetap diikuti oleh nilai *False Negative* yang relatif lebih tinggi dibandingkan konfigurasi XGBoost + Borderline-SMOTE. Temuan ini menegaskan adanya *trade-off* empiris antara kenyamanan transaksi dan agresivitas proteksi finansial, sehingga diperlukan formulasi kebijakan mitigasi yang mampu menjaga keseimbangan antara pengalaman pengguna dan ketahanan sistem deteksi fraud secara simultan.

3.5 Implikasi Operasional Bisnis

Sebagai implikasi operasional akhir, konfigurasi XGBoost + Borderline-SMOTE dikukuhkan sebagai arsitektur dengan *business equilibrium* paling optimal. Model komputasional ini sukses mengintersepsi 76 transaksi *fraud* secara presisi dengan hanya memicu 7 *False Positive*, yang didukung oleh capaian ROC-AUC (0,9783) serta AUPRC (0,8306). Temuan esensial ini secara empiris mengonfirmasi bahwa keselarasan sistemis antara mekanisme *boundary-focused oversampling* dan algoritma deterministik pohon keputusan mampu memahat *decision boundary* yang sangat tajam dan presisi pada ruang dimensi fitur PCA. Keunggulan ini tidak semata-mata merepresentasikan kestabilan performa statistik, melainkan juga membuktikan kapabilitas riil sistem dalam menjaga ekuilibrium krusial antara proteksi agresif terhadap aset finansial dan kontinuitas pengalaman transaksi nasabah tanpa menimbulkan friksi operasional yang berarti.

Sebagai model superior lainnya, integrasi CTGAN pada TabTransformer tampil sangat kompetitif dengan perolehan ROC-AUC tertinggi (0,9846) dan AUPRC (0,8003). Visualisasi *confusion matrix* membuktikan arsitektur probabilistik ini berhasil mendeteksi 56648 *True Positive* dan menekan *False Positive* di angka 71. Hal ini memvalidasi kemampuan *deep learning* mereduksi alarm palsu via pengayaan minoritas yang adaptif terhadap *manifold* laten. Eksistensi kedua konfigurasi ini sukses menjawab urgensi perbankan modern: mengamankan aset maksimal tanpa mengorbankan kenyamanan nasabah.

4. KESIMPULAN

Penelitian ini menyimpulkan bahwa efektivitas deteksi *fraud* pada distribusi kelas sangat timpang bergantung secara absolut pada keselarasan antara topologi data sintesis dengan arsitektur model. Evaluasi komprehensif berhasil mengonfirmasi dua konfigurasi unggulan yang tangguh. Kombinasi XGBoost dengan Borderline-SMOTE tampil superior dalam mengintersepsi *fraud* secara agresif, dibuktikan lewat capaian metrik akurasi (0,9995), F1-Score (0,9269), dan AUPRC (0,8306) karena mampu meningkatkan *recall* secara signifikan. Secara paralel, sinergi TabTransformer dengan CTGAN juga menunjukkan performa superior yang efektif meminimalkan alarm palsu, direpresentasikan oleh pencapaian nilai tertinggi pada akurasi (0,9995), ROC-AUC (0,9846), presisi (0,9795), serta F1-Score (0,9200). Sebaliknya, arsitektur Deep Autoencoder yang beroperasi secara tunggal berbasis *unsupervised reconstruction learning* justru menunjukkan performa paling rendah dan kalah bersaing dibandingkan kedua model klasifikasi *supervised* tersebut. Meskipun mencatat akurasi sebesar (0,9983), model ini mengalami degradasi sensitivitas parah dengan nilai AUPRC yang hanya menyentuh angka (0,3158) akibat kegagalan memisahkan *error* rekonstruksi yang tumpang tindih.

Menyikapi keterbatasan arsitektur Deep Autoencoder konvensional tersebut, saran pengembangan riset lanjutan harus diarahkan untuk mengevolusikan potensi model rekonstruksi ini ke dalam kerangka hibrida yang lebih adaptif. Mengacu pada temuan terbaru Alharbi dkk tahun 2026 mengenai integrasi *multi-autoencoders* dan *generative ensemble learning*, penelitian masa depan sangat direkomendasikan untuk melebur kapabilitas *representation learning* dengan teknik augmentasi sintesis probabilistik menjadi satu kesatuan sistem hibrida terpadu [25]. Strategi inovatif ini diproyeksikan sangat mampu mengatasi kelemahan model tunggal dalam memetakan anomali laten, memperluas distribusi *manifold* minoritas secara akurat, serta menciptakan batas keputusan yang jauh lebih kokoh pada seluruh ekosistem transaksi finansial.

REFERENSI

- [1] The Nilson Report, "Payment card fraud losses approach \$34 billion.," 2025, [Online]. Available: <https://www.globenewswire.com/news-release/2025/01/06/3004931/0/en/%0APayment-Card-Fraud-Losses-Approach-34-Billion.html>.

- [2] I. Y. Hafez, A. Y. Hafez, A. Saleh, A. A. Abd El-Mageed, and A. A. Abohany, "A systematic review of AI-enhanced techniques in credit card fraud detection," *J. Big Data*, vol. 12, p. 6, 2025, doi: 10.1186/s40537-024-01048-8.
- [3] S. M. Darwish, S. EL-Naggar, and S. M. Elkaffas, "Securing financial transactions: exploring the role of lightweight blockchain-enabled deep learning for fraud detection in FinTech systems," *Cybersecurity*, vol. 9, p. 8, 2026, doi: 10.1186/s42400-025-00436-8.
- [4] E. Ileberi, "Performance Analysis of Cost-Sensitive Oversampling Mechanisms for Credit Card Fraud Detection," *IEEE Access*, vol. 13, no. December, pp. 202655–202664, 2025, doi: 10.1109/ACCESS.2025.3637132.
- [5] F. Movahedi, R. Padman, and J. F. Antaki, "Adult : Mechanical Circulatory Support Limitations of receiver operating characteristic curve on imbalanced data : Assist device mortality risk scores," *J. Thorac. Cardiovasc. Surg.*, vol. 165, no. 4, pp. 1433-1442.e2, 2021, doi: 10.1016/j.jtcvs.2021.07.041.
- [6] H. Yin, Y. Teng, X. Zhang, and L. Yu, "Adaptive online bagging with hybrid sampling for non-stationary imbalanced data streams: A case study on credit card fraud detection," *Inf. Process. Manag.*, vol. 63, p. 104841, 2026, doi: 10.1016/j.ipm.2026.104841.
- [7] M. Tayebi and S. El Kafhali, "A novel approach based on XGBoost classifier and Bayesian optimization for credit card fraud detection," *Cyber Secur. Appl.*, vol. 3, p. 100093, 2025, doi: 10.1016/j.csa.2025.100093.
- [8] S. Lai, J. Wu, W. Wei, and C. Ye, "MiKTr: Multiscale knowledge fusion transformer for consumer fraud detection," *Inf. Fusion*, vol. 127, p. 103779, 2025, doi: 10.1016/j.inffus.2025.103779.
- [9] H. Fanai and H. Abbasimehr, "A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection," *Expert Syst. Appl.*, vol. 217, p. 119562, 2023, doi: 10.1016/j.eswa.2023.119562.
- [10] O. Habibi, M. Chemmakha, and M. Lazaar, "Imbalanced tabular data modelization using CTGAN and machine learning to improve IoT Botnet attacks detection," *Eng. Appl. Artif. Intell.*, vol. 118, p. 105669, 2023, doi: 10.1016/j.engappai.2022.105669.
- [11] M. Imani, M. Joudaki, A. Bagheri, and H. R. Arabnia, "Why ROC-AUC Is Misleading for Highly Imbalanced Data: In-Depth Evaluation of MCC, F2-Score, H-Measure, and AUC-Based Metrics Across Diverse Classifiers," *Technologies*, 2026, doi: 10.3390/technologies14010054.
- [12] M. Zavvar *et al.*, "A hybrid deep learning framework using synthetic oversampling, autoencoder, convolutional neural networks, and an attention mechanism for credit card fraud detection," *J. Big Data*, vol. 13, p. 21, 2026, doi: 10.1186/s40537-025-01331-2.
- [13] A. Al-dulaimi and N. Nnamoko, "Efficient treatment of class imbalance for credit card fraud detection Efficient treatment of class imbalance for credit card fraud detection," 2025, doi: 10.36227/techrxiv.175693752.21183675/v1.
- [14] J. Teng, "ScienceDirect Financial Data Reduction and Information Retention Strategy based on Principal Component Analysis (PCA) Algorithm," *Procedia Comput. Sci.*, vol. 262, pp. 218–226, 2025, doi: 10.1016/j.procs.2025.05.047.
- [15] M. Kurniawan, H. Putra, C. A. Mintaraga, and Hidayaturrehman, "When Two are Better Than One: Synthesizing Heavily Unbalanced Data," *Procedia Comput. Sci.*, vol. 269, pp. 485–501, 2025, doi: 10.1016/j.procs.2025.08.301.
- [16] S. Matharaarachchi, M. Domaratzki, and S. Muthukumarana, "Machine Learning with Applications Enhancing SMOTE for imbalanced data with abnormal minority instances," *Mach. Learn. with Appl.*, vol. 18, no. December 2023, 2024.
- [17] R. J. Kuo, K.-W. Zheng, F. E. Zulvia, and T. Kuo, "Impurity-based borderline SMOTE with affinity propagation for imbalanced data classification," *Comput. Ind. Eng.*, vol. 212, p. 111720, 2026, doi: https://doi.org/10.1016/j.cie.2025.111720.
- [18] I. Aruleba and Y. Sun, "Machine Learning with Applications Enhanced credit risk prediction using deep learning and SMOTE-ENN resampling," *Mach. Learn. with Appl.*, vol. 21, no. April, p. 100692, 2025, doi: 10.1016/j.mlwa.2025.100692.
- [19] J. Chen, Y. Liang, J. Liu, and M. Zhou, "Temporal Transformer with Conditional Tabular GAN for Credit Card Fraud Detection: A Sequential Deep Learning Approach," *Mathematics*, 2026, doi: 10.1016/j.aei.2026.104514.
- [20] R. Lingeswari and S. Brindha, "Online payments fraud prediction using optimized genetic algorithm based feature extraction and modified loss with XG boost algorithm for classification," *Swarm Evol. Comput.*, vol. 95, no. April, p. 101934, 2025, doi: 10.1016/j.swevo.2025.101934.
- [21] M. Duan, R. Du, and S. Ma, "Temporal Hypergraph Attentive Transformer for regional financial risk prediction: Integrating risk contagion dynamics and temporal pattern mining," *Alexandria Eng. J.*, vol. 133, pp. 137–156, 2025, doi: 10.1016/j.aej.2025.11.005.
- [22] S. Alharbi, "Aware Credit Card Fraud Detection Using Autoencoders and Generative Ensemble Learning," *Expert Syst.*, pp. 1–25, 2026, doi: 10.1111/exsys.70256.

- [23] R. H. Egami, A. A. A. El-mageed, M. Gafar, and A. A. Abohany, "An enhanced secretary bird-inspired multi-classifier framework for high-performance Credit Card Fraud Detection in imbalanced data," *Alexandria Eng. J.*, vol. 146, no. May, pp. 184–217, 2026, doi: 10.1016/j.aej.2026.04.061.
- [24] E. Richardson, R. Trevizani, J. A. Greenbaum, and H. Carter, "Article The receiver operating characteristic curve accurately assesses imbalanced datasets The receiver operating characteristic curve accurately assesses imbalanced datasets," *Patterns*, vol. 5, no. 6, p. 100994, 2024, doi: 10.1016/j.patter.2024.100994.
- [25] S. Alharbi, K. Alahmadi, and X. Wang, "Imbalance-Aware Credit Card Fraud Detection Using Multi-Autoencoders and Generative Ensemble Learning," *Expert Syst.*, 2026, doi: 10.1111/exsy.70256.