Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

Performa Random Forest dan XGBoost pada Deteksi Penipuan *E-Commerce* Menggunakan Augmentasi Data CGAN

Sarmini[1,*], Sunardi[2], Abdul Fadlil[2]

[1] Fakultas Teknologi Industri, Program Studi Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[2] Fakultas Teknologi Industri, Program Studi Teknik Elektro, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
Email: [1,*]2437083013@webmail.uad.ac.id, [2]sunardi@mti.uad.ac.id, [3]fadlil@mti.uad.ac.id
Email Penulis Korespondensi: 2437083013@webmail.uad.ac.id
Submitted: **10/12/2024**; Accepted: **25/12/2024**; Published**: 26/12/2024**

Abstrak−Deteksi penipuan pada e-*commerce* menghadapi tantangan besar akibat ketidakseimbangan data karena transaksi sah jauh lebih banyak daripada transaksi penipuan. Penelitian ini mengeksplorasi penggunaan *Conditional Generative Adversarial Network* (CGAN) untuk menghasilkan data transaksi penipuan sintetis guna mengatasi masalah ketidakseimbangan tersebut. Penambahan jumlah data pada kelas minoritas pada penelitian ini bertujuan untuk meningkatkan kinerja dua algoritma *machine learning* yang banyak digunakan, yaitu Random Forest dan XGBoost. Dataset yang digunakan sebanyak 23.634 dengan 22.412 transaksi non-penipuan dan 1.222 transaksi penipuan. Metrik *accuracy*, *precision*, *recall*, dan F1-*score* digunakan untuk mengevaluasi kinerja model dalam mendeteksi penipuan pada dataset yang tidak seimbang dan yang telah diaugmentasi. Hasil penelitian menunjukkan bahwa augmentasi data dengan CGAN secara signifikan meningkatkan kinerja kedua model, terutama dalam meningkatkan *recall* untuk transaksi penipuan. Pada dataset asli yang tidak seimbang, Random Forest dan XGBoost menunjukkan *recall* yang rendah (masing-masing 12,81% dan 13,08%), dengan *accuracy* masing-masing 95,35% dan 95,32%. Penerapan augmentasi mendapatkan *recall* meningkat menjadi 95,15% untuk Random Forest dan 95,22% untuk XGBoost, dengan F1-*score* masing-masing 97,47% dan 97,42%, serta *accuracy* masing-masing 97,50% dan 97,42%. XGBoost menunjukkan sedikit keunggulan dalam *precision* dan *recall* dibandingkan Random Forest, terutama pada dataset yang telah diaugmentasi. Temuan ini menegaskan efektivitas CGAN sebagai metode augmentasi data dalam meningkatkan performa deteksi penipuan dan menawarkan solusi yang kuat untuk mengatasi ketidakseimbangan data di sektor keuangan.

Kata Kunci: Augmentasi Data; CGAN; Deteksi Penipuan; E-commerce; Pembelajaran Mesin

Abstract−Fraud detection in e-commerce faces great challenges due to data imbalance, where legitimate transactions far outnumber fraudulent transactions. This research explores the use of Conditional Generative Adversarial Network (CGAN) to generate synthetic fraudulent transaction data to address the imbalance problem. By increasing the amount of data in the minority class, this research aims to improve the performance of two widely used machine learning algorithms, namely Random Forest and XGBoost. The dataset used of 23,634 transactions with 22,412 non-fraud transactions and 1,222 fraudulent transactions. Accuracy, precision, recall, and F1-score metrics were conducted to assess the performance of the model in detecting fraud on the imbalanced and augmented datasets. The results show that augmentation of data with CGAN significantly improves the performance of both models, especially in improving recall for fraudulent transactions. On the original unbalanced dataset, Random Forest and XGBoost showed low recall (12.81% and 13.08%), with accuracy of 95.35% and 95.32% respectively. However, after augmentation, recall improved to 95.15% for Random Forest and 95.22% for XGBoost, with F1-score of 97.47% and 97.42% respectively, and accuracy of 97.50% for Random Forest and 97.42% for XGBoost. XGBoost showed a slight advantage in precision and recall over Random Forest, especially on the augmented dataset. These findings confirm the effectiveness of CGAN as a data augmentation method in improving fraud detection performance and offer a robust solution to address data imbalance in the financial sector.

Keywords: Data Augmentation; CGAN; Fraud Detection; E-commerce; Machine Learning

1. PENDAHULUAN

Pertumbuhan e-*commerce* yang meningkat pesat terutama didorong oleh pandemi COVID-19 menyebabkan peningkatan signifikan dalam transaksi daring. Pada awal tahun 2020, transaksi e-*commerce* di Amerika Serikat meningkat 110% dibandingkan tahun sebelumnya, mencerminkan tren transformasi digital yang meluas di berbagai sektor. Pertumbuhan yang cepat ini telah meningkatkan kenyamanan konsumen dan sekaligus menciptakan peluang bagi aktivitas penipuan sehingga membutuhkan mekanisme deteksi penipuan yang kuat [1]. Dengan berkembangnya e-*commerce*, tantangan dalam mendeteksi penipuan menjadi semakin rumit. Salah satu kendala utama adalah ketidakseimbangan data transaksi, yaitu jumlah transaksi penipuan jauh lebih sedikit dibandingkan transaksi yang sah. Ketidakseimbangan ini menyulitkan pelatihan model *machine learning* yang seringkali berakibat rendahnya akurasi dalam deteksi terhadap aktivitas penipuan [2][3]. Agoritma *machine learning* tradisional sering mengalami kesulitan untuk mempelajari pola dari data yang dapat menyebabkan tingginya tingkat *false-negative* [4]. Selain itu, kualitas dan ketersediaan data juga menjadi tantangan karena banyak dataset yang digunakan untuk deteksi penipuan sering kali memiliki kualitas rendah sehingga mengurangi efektivitas algoritma yang digunakan [5].

Kemajuan terbaru dalam *machine learning* dan kecerdasan buatan (*Artificial Intelligent*, AI) telah membuka cara-cara baru untuk meningkatkan deteksi penipuan pada e-*commerce*. Berbagai teknik seperti *deep learning*, metode ansambel, dan *reinforcement learning* sedang dikembangkan untuk meningkatkan akurasi dan efisiensi dalam deteksi penipuan [6][7]. *Convolutional Neural Networks* (CNNs) telah digunakan untuk meningkatkan kemampuan prediksi penipuan dengan mengatasi masalah seperti *overfitting* dan tingginya dimensi data. Selain itu, analitik *big data* memungkinkan pengembangan sistem deteksi penipuan yang dapat ditingkatkan dan mampu memproses data transaksi dalam waktu nyata sehingga meningkatkan kecepatan respons dalam mendeteksi penipuan [6]. Taktik penipuan yang terus berkembang menuntut strategi deteksi yang selalu

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

beradaptasi. Para pelaku penipuan semakin cerdas menggunakan teknik canggih untuk menghindari sistem deteksi [7][8][9]. Hal ini menuntut implementasi sistem pembelajaran adaptif yang dapat berkembang seiring dengan munculnya pola penipuan baru. Pendekatan *transfer learning* telah diusulkan untuk memanfaatkan pengetahuan dari berbagai konteks sehingga meningkatkan ketahanan model deteksi penipuan di beragam lingkungan transaksi [10][11].

Ketidakseimbangan data menjadi tantangan utama dalam deteksi penipuan terutama di bidang e-*commerce* dan transaksi keuangan. Ketidakseimbangan ini terjadi ketika jumlah transaksi sah jauh lebih banyak daripada transaksi penipuan sehingga menyulitkan pelatihan model *machine learning* secara efektif. Model sering kali bias terhadap kelas mayoritas (transaksi sah) sehingga mengurangi kemampuan dalam mendeteksi kelas minoritas (transaksi penipuan) [7][12][13]. Ketidakseimbangan ini menyebabkan model *machine learning*, meskipun memiliki akurasi tinggi, sering gagal mengidentifikasi aktivitas penipuan dengan tepat, yang pada akhirnya mengurangi efektivitas sistem deteksi penipuan. Salah satu dampak utama dari ketidakseimbangan data adalah kecenderungan algoritma *machine learning* untuk mencapai akurasi tinggi terutama memprediksi kelas mayoritas. Hal ini menyebabkan tingginya tingkat *false-negative* atau transaksi penipuan diklasifikasikan sebagai transaksi sah, sehingga memungkinkan pelaku penipuan untuk mengeksploitasi kelemahan dalam sistem [14]. Kesalahan klasifikasi semacam ini memiliki dampak finansial yang besar karena dapat menyebabkan kerugian signifikan bagi bisnis dan merusak kepercayaan konsumen terhadap platform *e-commerce* [15][16]. Hal ini menyoroti pentingnya mengatasi ketidakseimbangan data sebagai langkah utama untuk meningkatkan efektivitas sistem deteksi penipuan.

Berbagai teknik telah diusulkan untuk mengurangi dampak ketidakseimbangan data. Salah satu pendekatan umum adalah metode *oversampling*, seperti *Synthetic Minority Over-sampling Technique* (SMOTE) dan Conditional Generative Adversarial Networks (CGAN), yang menghasilkan sampel sintetis dari kelas minoritas untuk menyeimbangkan dataset [15][17][18][19][20]. Teknik ini telah meningkatkan deteksi transaksi penipuan dengan menyediakan lebih banyak contoh bagi model untuk dipelajari. Selain itu, metode pembelajaran ansambel, yang menggabungkan beberapa klasifier, mampu menangani data yang tidak seimbang dengan efektif dan memanfaatkan kekuatan berbagai model untuk meningkatkan kinerja keseluruhan [7]. Penelitian terbaru juga berfokus pada optimalisasi *neural networks* khusus untuk dataset yang tidak seimbang. Teknik seperti *cost-sensitive learning* menerapkan biaya kesalahan klasifikasi berbeda pada kelas mayoritas dan minoritas dapat membantu model mencapai kinerja yang lebih baik pada kelas minoritas [21]. Selain itu, pendekatan hibrida yang menggabungkan metode *oversampling* dan *under-sampling* telah diuji untuk menciptakan dataset pelatihan yang lebih seimbang yang pada akhirnya meningkatkan deteksi transaksi penipuan [20][22][23][24]. Penanganan ketidakseimbangan data secara efektif sangat penting untuk mengembangkan sistem deteksi penipuan yang kuat dan andal sehingga mampu melindungi platform e-*commerce* dari ancaman penipuan yang terus meningkat.

Solusi ketidakseimbangan data sangat penting untuk meningkatkan akurasi sistem deteksi penipuan terutama dalam sektor keuangan. SMOTE telah terbukti efektif dalam meningkatkan kinerja model dengan menyeimbangkan dataset sehingga mengurangi risiko *overfitting* dan hilangnya informasi yang sering terjadi pada metode tradisional seperti *Random Undersampling* (RUS) [15][16][25]. Dataset yang tidak seimbang dapat menyebabkan kinerja yang kurang optimal pada model *machine learning*, terutama dalam mengidentifikasi kelas minoritas, yang sering kali merupakan kasus penipuan dalam skenario deteksi penipuan [7][26]. Teknik *resampling* hibrida yang menggabungkan berbagai metode dapat mengoptimalkan proses deteksi untuk meningkatkan akurasi model dalam aplikasi dunia nyata [16]. Integrasi algoritma *machine learning* lanjutan dengan teknik pengelolaan ketidakseimbangan kelas sangat penting untuk membangun sistem deteksi penipuan yang kuat dan adaptif terhadap perubahan pola penipuan. Selain itu, CGAN memungkinkan pembuatan data sintetis yang realistis dan relevan dengan mengkondisikan fitur tertentu, seperti jenis transaksi atau demografi pengguna, sehingga efektif dalam menangani ketidakseimbangan kelas. Dengan menghasilkan sampel minoritas yang bervariasi dan berkualitas, CGAN memperkuat kemampuan model pembelajaran mesin untuk mendeteksi pola unik pada data dunia nyata, seperti transaksi penipuan.

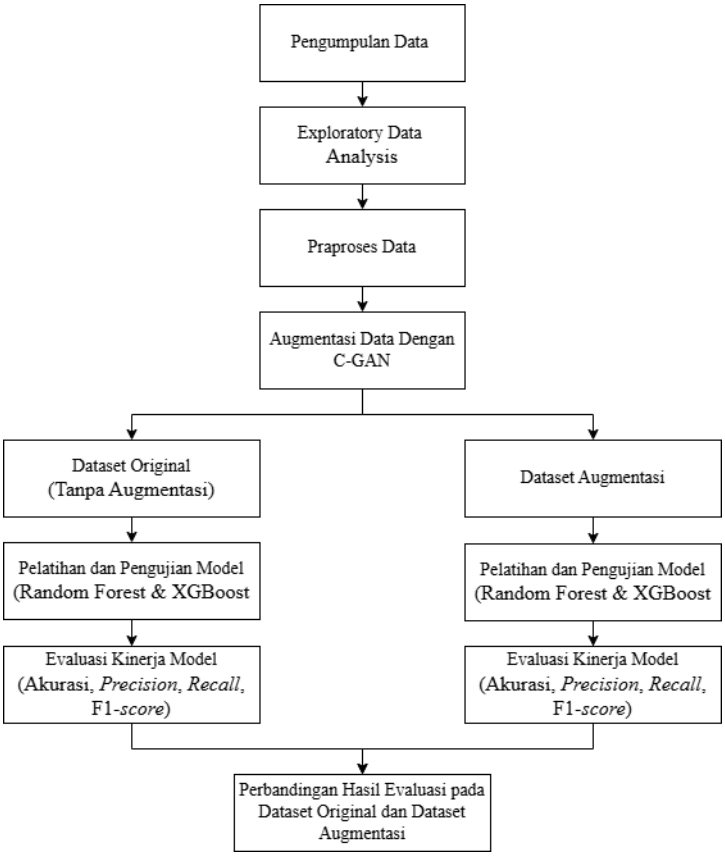
Penelitian ini bertujuan untuk mengevaluasi dan membandingkan efektivitas CGAN dalam augmentasi data guna meningkatkan deteksi penipuan dalam transaksi e-*commerce*. Masalah ketidakseimbangan data yang sering terjadi dalam deteksi penipuan menjadi tantangan pada penelitian ini untuk menghasilkan data sintetis yang menyerupai kelas minoritas khususnya transaksi penipuan yang dikondisikan pada label penipuan. Augmentasi ini diharapkan dapat memperkaya dataset pelatihan dan meningkatkan kinerja model *machine learning* dalam mendeteksi aktivitas penipuan. Penelitian berfokus pada beberapa sasaran spesifik. Studi ini menyoroti dampak ketidakseimbangan data terhadap kinerja model deteksi penipuan. Data penipuan sintetis menggunakan CGAN bertujuan untuk menyeimbangkan dataset dan mengurangi ketimpangan yang cenderung menguntungkan transaksi sah. Efektivitas augmentasi data berbasis CGAN dievaluasi dengan membandingkan kinerja dua algoritma klasifikasi yang banyak digunakan, yaitu Random Forest dan XGBoost, pada dataset yang tidak seimbang dan yang telah di-augmentasi. Terakhir, penelitian ini bertujuan untuk mengidentifikasi model yang paling efektif untuk deteksi penipuan dengan mengevaluasi metrik kinerja utama seperti *recall*, *precision*, dan F1-*score*.

2. METODOLOGI PENELITIAN

Metodologi penelitian dalam studi ini mengikuti alur kerja terstruktur yang mencakup beberapa tahapan utama, seperti yang ditunjukkan pada Gambar 1. Proses dimulai dengan pengumpulan data, yaitu mengambil dataset dari Kaggle. Setelah itu, dilakukan *Exploratory Data Analysis* (EDA) untuk mendapatkan wawasan mengenai fitur, distribusi, dan potensi korelasi dalam dataset. Langkah selanjutnya adalah praproses data, yang mencakup penanganan nilai yang hilang, pengkodean variabel kategoris, dan normalisasi fitur numerik untuk memastikan kesesuaian dengan model *machine learning*. Selanjutnya, dilakukan augmentasi data menggunakan CGAN untuk mengatasi ketidakseimbangan kelas yang signifikan dalam dataset dengan

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

menghasilkan transaksi penipuan sintetis. Dataset yang seimbang ini kemudian digunakan dalam pelatihan model, yaitu algoritma Random Forest dan XGBoost diterapkan untuk mengevaluasi dampak augmentasi. Terakhir, model-model ini dievaluasi menggunakan berbagai metrik, termasuk *presisi, recall,* dan F1-*score* untuk mengukur efektivitas dalam mendeteksi aktivitas penipuan dengan akurat.



Gambar 1. Diagram Alur Penelitian

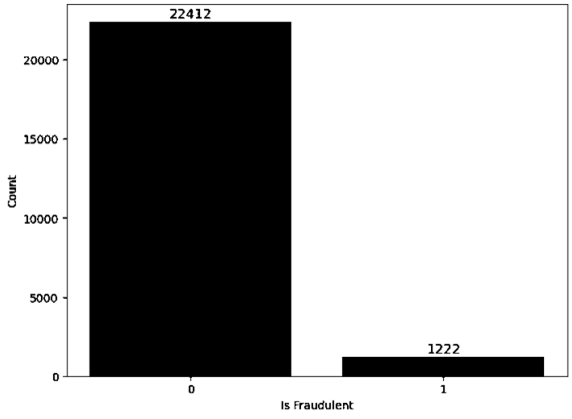
2.1 Pengumpulan Data

Dataset yang digunakan dalam studi ini adalah *Fraudulent E-Commerce Transactions* yang diperoleh dari Kaggle [27]. Dataset berisi total 23.634 catatan dengan 22.412 diantaranya adalah transaksi bukan penipuan dan 1.222 transaksi penipuan, seperti ditunjukkan pada Tabel 1.

Tabel 1. Distribusi Kelas Pada Dataset Original

Kelas	Jumlah Transaksi
Penipuan	1.222
Bukan Penipuan	22.412
Jumlah	23.634

Ketidak seimbangan kelas yang signifikan ini divisualisasikan pada Gambar 2 sebagai berikut.



Gambar 2. Distribusi Kelas Transaksi Penipuan dan Bukan Penipuan Pada Dataset Original

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

Setiap catatan mewakili transaksi individu dengan berbagai fitur yang relevan untuk deteksi penipuan. Fitur-fitur utama mencakup atribut seperti 'jumlah transaksi' (*transaction amount*), 'usia pelanggan' (*customer age*), 'usia akun dalam hari' (*account age days*), 'metode pembayaran' (*payment method*), 'kategori produk' (*product category*), dan 'apakah penipuan' (*is fraudulent*), selengkapnya ditampilkan pada Tabel 2.

Tabel 2. Fitur pada Dataset *Fraudulent E-Commerce Transactions*

Fitur	Penjelasan
Transaction ID	Identifikasi unik untuk setiap transaksi.
Customer ID	Identifikasi unik untuk setiap pelanggan.
Transaction Amount	Jumlah total uang yang dipertukarkan dalam transaksi.
Transaction Date	Tanggal dan waktu ketika transaksi terjadi.
Payment Method	Metode yang digunakan untuk menyelesaikan transaksi (misalnya, kartu kredit, PayPal, dll.).
Product Category	Kategori produk yang terlibat dalam transaksi.
Quantity	Jumlah produk yang terlibat dalam transaksi.
Customer Age	Usia pelanggan yang melakukan transaksi.
Customer Location	Lokasi geografis pelanggan.
Device Used	Jenis perangkat yang digunakan untuk melakukan transaksi (misalnya, ponsel, desktop).
IP Address	Alamat IP dari perangkat yang digunakan untuk transaksi.
Shipping Address	Alamat tempat produk dikirim.
Billing Address	Alamat yang terkait dengan metode pembayaran.
Is Fraudulent	Indikator biner apakah transaksi merupakan penipuan (1 untuk penipuan, 0 untuk sah).
Account Age Days	Usia akun pelanggan dalam hari pada saat transaksi terjadi.
Transaction Hour	Jam saat transaksi terjadi.

Fitur-fitur ini memberikan gambaran menyeluruh terhadap transaksi dan membantu mengidentifikasi pola-pola yang mungkin menunjukkan aktivitas penipuan. Dataset ini juga mencakup variabel kategoris yang menunjukkan perilaku pelanggan dan variabel numerik yang mengukur berbagai aspek transaksi. Salah satu tantangan utama dalam dataset ini adalah ketidakseimbangan pada atribut 'Apakah Penipuan' (*Is Fraudulent*), yang berfungsi sebagai variabel target dalam studi ini. Ketidakseimbangan ini menjadi tantangan signifikan bagi algoritma *machine learning* karena sering berkinerja buruk dalam mengidentifikasi kelas minoritas akibat kecenderungan alami mereka untuk memprioritaskan kelas mayoritas.

2.2 *Exploratory Data Analysis* (EDA)

EDA dilakukan untuk lebih memahami dataset dan fitur-fiturnya yang sangat penting untuk pembangunan model yang efektif. EDA melibatkan analisis statistik dan visualisasi untuk mengidentifikasi tren, *outliers*, dan potensi korelasi diantara atribut-atribut utama. Proses ini membantu menentukan kecocokan data untuk *machine learning* dan mengidentifikasi langkah-langkah praproses yang diperlukan. Atribut-atribut utama yang dianalisis selama EDA mencakup 'jumlah transaksi' (*transaction amount*), 'usia pelanggan' (*customer age*), dan 'usia akun dalam hari' (*account age days*). Statistik ringkasan seperti rata-rata, median, simpangan baku, dan kuartil dihitung untuk fitur-fitur numerik ini guna memahami distribusinya dan mengidentifikasi adanya ketidakteraturan.

EDA juga mencakup visualisasi untuk memberikan gambaran yang lebih jelas tentang distribusi data. Diagram batang dibuat untuk memvisualisasikan distribusi antara transaksi penipuan dan transaksi bukan penipuan yang menunjukkan ketidakseimbangan yang signifikan antara kedua kelas. Selain itu, histogram digunakan untuk menampilkan distribusi fitur numerik seperti 'jumlah transaksi' dan 'usia pelanggan'. Histogram-histogram ini membantu mengidentifikasi adanya *skewness* dan *outlier* yang dapat mempengaruhi kinerja model *machine learning* jika tidak ditangani dengan baik.

2.3 Praproses Data

Praproses data adalah langkah penting untuk mempersiapkan dataset dalam pelatihan model *machine learning*. Langkah ini meliputi penanganan nilai hilang, pengkodean data kategoris, dan normalisasi atribut numerik. Nilai hilang pada fitur numerik (seperti jumlah transaksi, usia pelanggan, dan usia akun) diimputasi menggunakan median untuk mengurangi dampak *outliers*, sedangkan fitur kategoris (seperti metode pembayaran dan kategori produk) diimputasi dengan nilai yang paling sering muncul.

Fitur kategoris kemudian diubah menjadi format numerik dengan *one-hot encoding* untuk menghindari asumsi hubungan ordinal antar kategori. Fitur numerik dinormalisasi dengan pengurangan rata-rata dan pembagian dengan standar deviasi untuk memastikan skala yang seragam, yang penting bagi algoritma seperti XGBoost. Praproses ini memastikan dataset bersih dan siap untuk pelatihan sehingga meningkatkan kinerja model deteksi penipuan.

2.4 Augmentasi Data dengan CGAN

CGAN digunakan untuk menghasilkan transaksi penipuan sintetis guna mengatasi ketidakseimbangan data. CGAN terdiri dari dua jaringan saraf, yaitu *generator* dan *discriminator*, yang bekerja dalam proses pelatihan adversarial dengan kondisi tertentu, misalnya label penipuan atau bukan penipuan. *Generator* menciptakan data sintetis yang menyerupai transaksi penipuan nyata berdasarkan label, sementara *discriminator* membedakan data asli dan sintetis dengan mempertimbangkan label tersebut. Selama

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

pelatihan, *generator* berusaha menghasilkan sampel penipuan yang realistis untuk menipu *discriminator*, sementara *discriminator* belajar mengidentifikasi data palsu. Proses ini berlanjut hingga *generator* dapat menghasilkan transaksi penipuan yang sangat realistis dan sulit dibedakan dari data asli.

Arsitektur CGAN dalam penelitian ini menggunakan jaringan saraf *feedforward* untuk *generator* dan *discriminator* dengan *input* tambahan berupa label. *Generator* menerima vektor acak dan label, lalu mengubahnya menjadi transaksi penipuan sintetis dengan karakteristik yang mirip data asli. *Discriminator* juga diberi label kondisi untuk membedakan data nyata dan sintetis, memberikan umpan balik kepada *generator* untuk meningkatkan kualitas data. *Generator* dan *discriminator* dilatih bergantian, dengan *loss* dari *discriminator* membantu *generator* menghasilkan sampel sintetis yang lebih baik. Mekanisme pengondisian ini memungkinkan CGAN menghasilkan sampel penipuan yang lebih realistis dan sesuai dengan label yang ditargetkan.

Setelah data penipuan sintetis dihasilkan, data tersebut digabungkan dengan dataset asli untuk menciptakan dataset yang lebih seimbang. Integrasi data sintetis pada dataset yang awalnya tidak seimbang antara transaksi sah dan penipuan kemudian lebih representatif terhadap kedua kelas. Dataset yang seimbang ini kemudian digunakan untuk melatih model *machine learning*, memastikan kedua kelas terwakili dengan baik selama pelatihan. Augmentasi data ini bertujuan meningkatkan kemampuan model mendeteksi penipuan, mengurangi bias terhadap kelas mayoritas, dan meningkatkan kinerja deteksi secara keseluruhan. Mekanisme pengondisian CGAN juga memastikan bahwa data penipuan yang dihasilkan lebih realistis dan memenuhi kriteria yang diinginkan.

2.5 Pelatihan Model

Proses pelatihan dan pengujian model dilakukan secara terpisah pada dua pipeline yang berbeda, yaitu pipeline untuk dataset asli (*imbalanced*) dan pipeline untuk dataset augmentasi (*balanced*). Tujuan dari pemisahan ini adalah untuk mengevaluasi secara independen kinerja model pada dataset asli yang memiliki ketidakseimbangan data dan dataset yang telah diaugmentasi menggunakan CGAN untuk menciptakan data penipuan sintetis sehingga lebih seimbang.

Dataset asli dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian, dengan parameter random_state=42 untuk memastikan hasil yang konsisten. Dataset asli terdiri dari 23.634 catatan, di mana setelah dibagi, data pelatihan berjumlah 16.543 catatan, dan data pengujian 7.091 catatan. Dataset ini memiliki distribusi yang tidak seimbang, di mana jumlah transaksi sah (bukan penipuan) jauh lebih banyak dibandingkan transaksi penipuan. Model Random Forest dan XGBoost dilatih secara terpisah pada data pelatihan dari dataset asli. Penggunaan random_state=42 membantu dalam menghasilkan hasil yang konsisten, baik saat membagi data maupun melatih model. Dalam inisialisasi XGBoost, *parameter use_label_encoder=False* memastikan kompatibilitas dengan versi terbaru pustaka, dan eval_metric='*logloss*' digunakan karena relevan untuk klasifikasi biner. Random Forest bekerja dengan membangun sejumlah besar pohon keputusan (*decision trees*) secara acak selama proses pelatihan. Setiap pohon dilatih pada subset data yang diambil secara acak (*bootstrapping*), dengan subset fitur yang dipilih secara acak pada setiap pembagian simpul (*node*). Prediksi akhir dihasilkan dengan menggabungkan hasil dari semua pohon menggunakan *voting* mayoritas untuk klasifikasi. Random Forest memiliki keunggulan dalam menangani data yang memiliki hubungan kompleks antar fitur dan mengurangi risiko *overfitting* dibandingkan dengan pohon keputusan tunggal. Dalam konteks penelitian ini, model Random Forest memanfaatkan data pelatihan untuk belajar membedakan transaksi penipuan berdasarkan pola-pola yang ada dalam dataset. Proses pelatihan mencakup penyesuaian hiperparameter seperti jumlah pohon (n_estimators) dan kedalaman maksimum pohon (max_depth) untuk mengoptimalkan performa model, dengan metrik F1-*score* sebagai fokus utama untuk menangani ketidakseimbangan data. XGBoost (*Extreme Gradient Boosting*), di sisi lain, adalah algoritma *boosting* yang lebih kompleks. XGBoost bekerja dengan membangun model secara iteratif, di mana setiap model baru mencoba untuk memperbaiki kesalahan dari model sebelumnya. Pada setiap iterasi, XGBoost menghitung gradien dari fungsi loss untuk mengidentifikasi area di mana model saat ini lemah, lalu membangun pohon keputusan baru yang fokus pada area tersebut. Proses ini berlanjut hingga jumlah pohon tertentu tercapai atau kriteria penghentian lainnya terpenuhi. Dalam penelitian ini, XGBoost menggunakan metrik *logloss* untuk mengukur performa pada setiap iterasi, yang relevan untuk masalah klasifikasi biner. Hiperparameter seperti tingkat pembelajaran (*learning_rate*), kedalaman maksimum (*max_depth*), dan jumlah estimator (n_*estimators*) disesuaikan menggunakan GridSearchCV untuk mencapai performa terbaik. Proses pelatihan melibatkan optimasi hiperparameter menggunakan GridSearchCV, dengan metrik F1-*score* sebagai fokus utama untuk menangani ketidakseimbangan data. Setelah pelatihan, model diuji pada data pengujian dari dataset asli. Evaluasi dilakukan menggunakan metrik seperti *accuracy*, *precision*, *recall*, F1-*score*, dan ROC-AUC.

Dataset augmentasi yang telah dihasilkan menggunakan CGAN juga dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian, dengan parameter random_state=42 untuk menjaga konsistensi. Dataset ini telah seimbang, di mana jumlah data penipuan dan non-penipuan dibuat setara, sehingga memberikan lebih banyak informasi bagi model untuk mempelajari pola transaksi penipuan. Untuk dataset augmentasi yang memiliki distribusi seimbang, data pelatihan berjumlah 33.024 catatan, dan data pengujian 14.160 catatan. Model Random Forest dan XGBoost dilatih secara terpisah pada data pelatihan dari dataset augmentasi. Proses pelatihan serupa dengan *pipeline* sebelumnya, menggunakan GridSearchCV untuk optimasi hiperparameter. Data yang seimbang memungkinkan model untuk lebih efektif dalam belajar membedakan pola transaksi penipuan dan non-penipuan. Setelah pelatihan, model diuji pada data pengujian dari dataset augmentasi.

2.6 Metrik Evaluasi

Evaluasi model deteksi penipuan menggunakan beberapa metrik utama untuk memberikan gambaran menyeluruh tentang kinerjanya. Metrik-metrik ini meliputi *accuracy*, *precision*, *recall*, F1-*score*, ROC-AUC, dan PR-AUC. Akurasi mengukur proporsi prediksi yang benar terhadap total jumlah data. Rumus akurasi adalah:

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Di mana TP (*True Positive*) adalah transaksi penipuan yang berhasil dideteksi sebagai penipuan, TN (*True Negative*) adalah transaksi sah yang berhasil dideteksi sebagai bukan penipuan, FP (*False Positive)* adalah transaksi sah yang salah dideteksi sebagai penipuan dan FN (*False Negative*) adalah transaksi penipuan yang salah dideteksi sebagai bukan penipuan. Akurasi memberikan gambaran umum kinerja model, tetapi pada dataset tidak seimbang, metrik ini sering menyesatkan karena model dapat mencapai akurasi tinggi hanya dengan memprediksi kelas mayoritas. *Precision* mengukur akurasi prediksi positif. Rumusnya adalah:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Precision penting untuk meminimalkan false positives, terutama dalam konteks deteksi penipuan. *Precision* tinggi menunjukkan bahwa sebagian besar transaksi yang diprediksi sebagai penipuan memang benar-benar penipuan. *Recall* mengukur kemampuan model untuk mendeteksi semua kasus yang sebenarnya positif (penipuan). Rumusnya adalah:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

Recall sangat penting dalam konteks deteksi penipuan untuk memastikan tidak ada kasus penipuan yang terlewatkan (*False Negatives*). *Recall* tinggi menunjukkan bahwa model berhasil mendeteksi sebagian besar transaksi penipuan yang sebenarnya. F1-*score* adalah rata-rata harmonik dari *precision* dan *recall*. Rumusnya adalah:

$$\text{F1-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

F1-*score* digunakan untuk menyeimbangkan *precision* dan *recall*, terutama ketika keduanya memiliki nilai yang sangat berbeda. Nilai F1-*score* yang tinggi menunjukkan model yang efektif dalam mendeteksi penipuan tanpa terlalu banyak menghasilkan *false positives* atau *false negatives.*

ROC-AUC (*Receiver Operating Characteristic - Area Under Curve*) mengukur kemampuan model untuk membedakan antara kelas positif (penipuan) dan negatif (bukan penipuan) di berbagai ambang batas. Kurva ROC adalah grafik yang memplot *True Positive Rate* (TPR) terhadap *False Positive Rate* (FPR):

$$\text{TPR } (True\ Positive\ Rate) = \frac{TP}{TP+FN} \quad (5)$$

$$\text{FPR } (False\ Positive\ Rate) = \frac{FP}{FP+TN} \quad (6)$$

Nilai AUC (*Area Under Curve*) menunjukkan seberapa baik model dapat membedakan antara dua kelas, dengan nilai 1 menunjukkan performa sempurna dan 0,5 menunjukkan performa acak. PR-AUC (*Precision-Recall Area Under Curve*) lebih relevan pada dataset tidak seimbang karena fokus pada *precision* dan *recall*, yang menjadi perhatian utama dalam mendeteksi kelas minoritas (penipuan). Kurva *Precision-Recall* memplot precision terhadap *recall* di berbagai ambang batas. PR-AUC memberikan gambaran menyeluruh tentang kemampuan model dalam mendeteksi transaksi penipuan tanpa terlalu banyak menghasilkan *false positives.*

3. HASIL DAN PEMBAHASAN

3.1 Hasil

3.1.1 Hasil awal dataset tidak seimbang

Berdasarkan hasil klasifikasi menggunakan Random Forest pada dataset asli yang tidak seimbang, terlihat performa model dalam mengklasifikasikan transaksi penipuan dan bukan penipuan. Dataset asli ini memiliki distribusi yang tidak seimbang, di mana jumlah transaksi sah (non-penipuan) jauh lebih banyak dibandingkan dengan transaksi penipuan. Sebelum pelatihan, dataset dibagi menjadi 70% untuk data pelatihan dan 30% untuk data pengujian, dengan menggunakan parameter random_state=42 untuk memastikan hasil yang konsisten. Dari total 23,634 transaksi, data pelatihan terdiri dari 16.543 catatan, dan data pengujian mencakup 7.091 catatan.

Tabel 3. *Confusion Matrix* Model Random Forest pada Dataset Asli (Tidak Seimbang)

	Predicted Positive (Penipuan)	*Predicted Negative* (Bukan Penipuan)
Actual Positive (Penipuan)	47	320
Actual Negative (Bukan Penipuan)	6	6718

Tabel 3 menunjukkan model menghasilkan 6718 *True Negatives* (TN), yaitu jumlah transaksi yang sebenarnya bukan penipuan dan diklasifikasikan dengan benar oleh model sebagai bukan penipuan. Nilai ini menunjukkan bahwa model mampu mengenali transaksi yang tidak mencurigakan sebagai kelas mayoritas dalam dataset. Sebaliknya, terdapat 320 *False Negatives* (FN), yaitu transaksi penipuan yang salah diklasifikasikan sebagai transaksi bukan penipuan. Hal ini menunjukkan bahwa model

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

sering kali gagal mendeteksi transaksi penipuan, yang menyebabkan penipuan lolos dari pendeteksian. Jumlah FN yang cukup tinggi ini menegaskan bahwa model mengalami kesulitan dalam mempelajari karakteristik kelas minoritas (penipuan), yang sering kali terjadi pada dataset yang sangat tidak seimbang seperti ini. Selain itu, terdapat 6 *False Positives* (FP), yaitu transaksi yang sebenarnya bukan penipuan, namun salah diklasifikasikan oleh model sebagai penipuan. Meskipun jumlah FP relatif kecil, hal ini menunjukkan adanya kesalahan dalam mengidentifikasi transaksi yang sah sebagai penipuan, yang dapat mengganggu pengalaman pelanggan karena memerlukan verifikasi lebih lanjut. *True Positive* (TP), yaitu jumlah transaksi penipuan yang berhasil diklasifikasikan dengan benar sebagai penipuan, hanya sebanyak 47. Hal ini berarti bahwa model hanya mampu mendeteksi sebagian kecil dari keseluruhan transaksi penipuan yang mengindikasikan rendahnya kemampuan model dalam mengidentifikasi kasus penipuan dengan tepat.

Tabel 4. Hasil Evaluasi Model Random Forest pada Dataset Asli (Tidak Seimbang)

Metrik	Nilai (%)
Accuracy	95,35
Precision	88,68
Recall	12,81
F1-*Score*	22,36

Tabel 4 menunjukkan hasil evaluasi model Random Forest pada dataset asli yang tidak seimbang. *Accuracy* dari model adalah 95,35%, yang terlihat tinggi. Namun, metrik ini tidak cukup untuk mengevaluasi performa dalam konteks deteksi penipuan, karena dataset sangat tidak seimbang sehingga model cenderung lebih sering memprediksi transaksi sebagai bukan penipuan untuk mencapai tingkat akurasi tinggi. *Precision* untuk kelas penipuan adalah 88,68% yang menunjukkan bahwa dari semua transaksi yang diprediksi sebagai penipuan, sebagian besar benar-benar merupakan penipuan. Hal ini berarti model cukup baik dalam mengurangi *false positive*, namun karena jumlah prediksi penipuan yang benar (TP) sangat kecil, nilai ini masih belum cukup merepresentasikan performa keseluruhan. *Recall* untuk kelas penipuan hanya 12,81%, yang sangat rendah. Hal ini mengindikasikan bahwa model gagal mendeteksi sebagian besar transaksi penipuan yang sebenarnya, yang sangat berbahaya dalam sistem deteksi penipuan karena dapat menyebabkan kerugian finansial. F1-*score* untuk kelas penipuan adalah 22,36%, yang merupakan rata-rata harmonis antara *precision* dan *recall*. Nilai F-*score* yang rendah menunjukkan bahwa model tidak seimbang dalam mendeteksi kasus penipuan karena *recall* sangat rendah meskipun *precision* cukup tinggi.

Selanjutnya, hasil klasifikasi menggunakan XGBoost pada dataset asli yang tidak seimbang, terlihat performa model dalam mengklasifikasikan transaksi penipuan dan bukan penipuan. *Confusion matrix* pada Tabel 5 memberikan informasi tentang kinerja model dalam memprediksi transaksi yang benar dan salah untuk masing-masing kelas.

Tabel 5. *Confusion Matrix* Model XGboost pada Dataset Asli (Tidak Seimbang)

	Predicted Positive (Penipuan)	*Predicted Negative* (Bukan Penipuan)
Actual Positive (Penipuan)	48	319
Actual Negative (Bukan Penipuan)	13	6711

Tabel 5 menunjukkan 6711 TN, yaitu jumlah transaksi yang sebenarnya bukan penipuan dan diklasifikasikan dengan benar oleh model sebagai bukan penipuan. Nilai ini menunjukkan bahwa model cukup baik dalam mengenali transaksi yang tidak mencurigakan sebagai kelas mayoritas dalam dataset. Sebaliknya, terdapat 319 FN, yaitu transaksi penipuan yang salah diklasifikasikan sebagai transaksi bukan penipuan. Hal ini menunjukkan bahwa model sering kali gagal mendeteksi transaksi penipuan yang menyebabkan penipuan lolos dari pendeteksian. Jumlah FN yang tinggi ini menegaskan bahwa model mengalami kesulitan dalam mempelajari karakteristik kelas minoritas (penipuan), yang sering kali terjadi pada dataset yang sangat tidak seimbang seperti ini. Selain itu, terdapat 13 FP yaitu transaksi yang sebenarnya bukan penipuan, namun salah diklasifikasikan oleh model sebagai penipuan. Meskipun jumlah FP relatif kecil, hal ini menunjukkan adanya kesalahan dalam mengidentifikasi transaksi yang sah sebagai penipuan, yang dapat mengganggu pengalaman pelanggan karena memerlukan verifikasi lebih lanjut. TP, yaitu jumlah transaksi penipuan yang berhasil diklasifikasikan dengan benar sebagai penipuan, hanya sebanyak 48. Hal ini berarti bahwa model hanya mampu mendeteksi sebagian kecil dari keseluruhan transaksi penipuan yang mengindikasikan rendahnya kemampuan model dalam mengidentifikasi kasus penipuan dengan tepat.

Tabel 6. Hasil Evaluasi Model XGboost pada Dataset Asli (Tidak Seimbang)

Metrik	Nilai (%)
Accuracy	95,32
Precision	78,69
Recall	13,08
F1-*score*	22,42

Tabel 6 menunjukkan hasil evaluasi model XGBoost pada dataset asli yang tidak seimbang. *Accuracy* dari model adalah 95,32%, yang terlihat tinggi. Namun, metrik ini tidak cukup untuk mengevaluasi performa dalam konteks deteksi penipuan karena dataset sangat tidak seimbang sehingga model cenderung lebih sering memprediksi transaksi sebagai bukan penipuan untuk mencapai tingkat akurasi tinggi. *Precision* untuk kelas penipuan adalah 78,69% yang menunjukkan bahwa dari semua transaksi yang diprediksi sebagai penipuan, sebagian besar benar-benar merupakan penipuan. Hal ini berarti model cukup baik dalam mengurangi *false positive*, namun karena jumlah prediksi penipuan yang benar (TP) sangat kecil, nilai ini belum cukup

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

merepresentasikan performa keseluruhan. *Recall* untuk kelas penipuan hanya 13,08%, yang sangat rendah. Hal ini mengindikasikan bahwa model gagal mendeteksi sebagian besar transaksi penipuan yang sebenarnya, yang sangat berbahaya dalam sistem deteksi penipuan karena dapat menyebabkan kerugian finansial. F1-*score* untuk kelas penipuan adalah 22,42% yang merupakan rata-rata harmonis antara *precision* dan *recall*. Nilai F1-*score* yang rendah menunjukkan bahwa model tidak seimbang dalam mendeteksi kasus penipuan, karena *recall* sangat rendah meskipun *precision* tinggi.

Hasil klasifikasi menggunakan Random Forest dan XGBoost pada dataset asli yang tidak seimbang menunjukkan bahwa kedua model memiliki akurasi tinggi secara keseluruhan, namun performa dalam mendeteksi transaksi penipuan sangat buruk. Tingginya nilai FN menunjukkan bahwa model sering kali tidak mendeteksi transaksi penipuan, sedangkan rendahnya nilai TP mengindikasikan kesulitan dalam mempelajari karakteristik kelas minoritas (penipuan). Hal ini disebabkan oleh ketidakseimbangan data, yaitu jumlah transaksi bukan penipuan jauh lebih banyak daripada transaksi penipuan. Oleh karena itu, diperlukan strategi untuk mengatasi ketidakseimbangan ini seperti *data augmentation* atau *oversampling* guna meningkatkan performa model dalam mendeteksi transaksi penipuan dan mengurangi risiko gagal mendeteksi kasus penipuan.

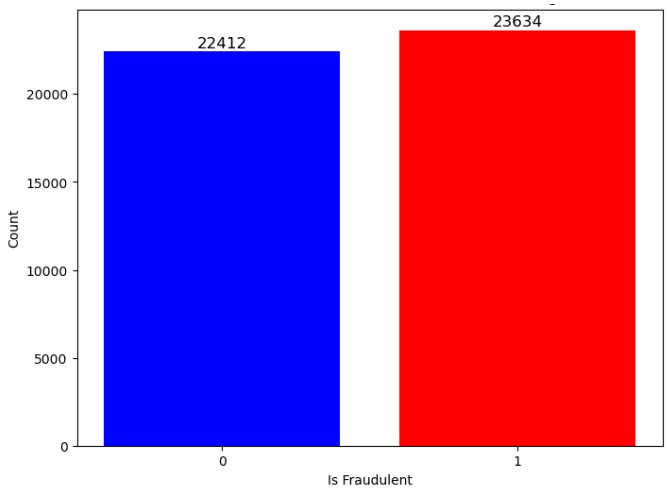
3.1.2 Hasil Setelah Augmentasi Data (Dataset Seimbang)

Untuk mengatasi ketidakseimbangan kelas, data diaugmentasi menggunakan CGAN untuk menghasilkan transaksi penipuan sintetis. Proses augmentasi ini menghasilkan dataset yang lebih seimbang, di mana jumlah transaksi penipuan menjadi setara dengan transaksi bukan penipuan. Augmentasi data ini bertujuan memberikan representasi yang lebih baik dari kelas minoritas (penipuan) dalam dataset, sehingga model dapat mempelajari karakteristik transaksi penipuan dengan lebih efektif. Dataset augmentasi dibagi menjadi 70% untuk data pelatihan dan 30% untuk data pengujian, serupa dengan dataset asli, dengan parameter random_state=42 untuk memastikan konsistensi pembagian. Dari total 46.046 catatan, data pelatihan terdiri dari 32.232 transaksi, sementara data pengujian mencakup 13.814 transaksi. Distribusi kelas dalam dataset augmentasi menjadi jauh lebih seimbang, dengan 23.634 transaksi penipuan dan 22.412 transaksi bukan penipuan pada data pelatihan, serta proporsi serupa pada data pengujian. Dataset yang telah diaugmentasi ini memberikan lebih banyak contoh transaksi penipuan, yang sebelumnya hanya terdiri dari 1.222 transaksi pada dataset asli, seperti ditunjukkan pada Tabel 7.

Tabel 7. Distribusi Kelas Pada Dataset Augmentasi

Kelas	Jumlah Transaksi
Penipuan	23.634
Bukan Penipuan	22.412
Jumlah	46.046

Dataset dengan kelas yang sudah seimbang ini divisualisasikan pada Gambar 3 sebagai berikut.



Gambar 3. Distribusi Kelas Transaksi Penipuan dan Bukan Penipuan Pada Dataset Setelah Augmentasi

Berdasarkan hasil klasifikasi menggunakan Random Forest pada dataset yang telah diimbangi menggunakan *data augmentation* berbasis CGAN (*balanced dataset*), terlihat performa model dalam mengklasifikasikan transaksi penipuan dan bukan penipuan. *Confusion matrix* pada Tabel 8 memberikan informasi tentang kinerja model dalam memprediksi transaksi yang benar dan salah untuk masing-masing kelas (penipuan dan bukan penipuan).

Tabel 8. *Confusion Matrix* Model Random Forest pada Dataset *Augmented* (Seimbang)

	Predicted Positive **(Penipuan)**	***Predicted Negative*** **(Bukan Penipuan)**
Actual Positive (Penipuan)	6746	344
Actual Negative (Bukan Penipuan)	1	6723

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

Tabel 8 menunjukkan model menghasilkan 6723 TN, yaitu jumlah transaksi yang sebenarnya bukan penipuan dan diklasifikasikan dengan benar oleh model sebagai bukan penipuan. Nilai ini menunjukkan bahwa model sangat baik dalam mengenali transaksi yang tidak mencurigakan sebagai kelas mayoritas dalam dataset. Sebaliknya, terdapat 344 FN, yaitu transaksi penipuan yang salah diklasifikasikan sebagai transaksi bukan penipuan. Meskipun nilai ini tidak terlalu tinggi, hal ini menunjukkan bahwa model masih dapat mengalami kesulitan dalam mendeteksi semua transaksi penipuan, walaupun sudah dilakukan *data augmentation*. Selain itu, terdapat 1 FP, yaitu transaksi yang sebenarnya bukan penipuan, namun salah diklasifikasikan oleh model sebagai penipuan. Jumlah FP yang sangat kecil ini menunjukkan bahwa model memiliki akurasi yang sangat baik dalam membedakan antara transaksi non-penipuan dan penipuan, dengan kata lain bahwa kesalahan yang terjadi dalam mengidentifikasi transaksi sah sebagai penipuan sangat minim. TP, yaitu jumlah transaksi penipuan yang berhasil diklasifikasikan dengan benar sebagai penipuan, sebanyak 6746. Hal ini menunjukkan bahwa model mampu mendeteksi sebagian besar dari keseluruhan transaksi penipuan dengan baik, yang mengindikasikan meningkatnya kemampuan model dalam mengidentifikasi kasus penipuan setelah dilakukan *data augmentation*.

Tabel 9. Hasil Evaluasi Model Random Forest pada Dataset *Augmented* (Seimbang)

Metrik	**Nilai (%)**
Accuracy	97,50
Precision	99,99
Recall	95,15
F1-*score*	97,47

Tabel 9 menunjukkan hasil evaluasi model Random Forest pada dataset yang telah diimbangi menggunakan CGAN. *Accuracy* dari model adalah 97,50%,yang menunjukkan bahwa model memiliki performa yang sangat baik secara keseluruhan dalam mengklasifikasikan transaksi penipuan dan bukan penipuan. *Precision* untuk kelas penipuan adalah 99,99% yang menunjukkan bahwa hampir semua transaksi yang diprediksi sebagai penipuan benar-benar merupakan penipuan. Hal ini berarti model sangat efektif dalam mengurangi *false positive* sehingga hampir tidak ada transaksi sah yang salah diklasifikasikan sebagai penipuan. *Recall* untuk kelas penipuan adalah 95,15% yang cukup tinggi dan menunjukkan bahwa model berhasil mendeteksi sebagian besar transaksi penipuan. Hal ini penting untuk sistem deteksi penipuan karena rendahnya *false negative* berarti semakin sedikit transaksi penipuan yang tidak terdeteksi. F1-*score* untuk kelas penipuan adalah 97,47% yang merupakan rata-rata harmonis antara *precision dan recall*. Nilai F1-*score* yang tinggi menunjukkan bahwa model seimbang dalam mendeteksi kasus penipuan, dengan *precision* dan *recall* yang sama-sama sangat baik.

Kemudian, berdasarkan hasil klasifikasi menggunakan XGBoost pada dataset yang telah diaugmentasi menggunakan CGAN (seimbang), terlihat bahwa model menunjukkan peningkatan performa dalam mengklasifikasikan transaksi penipuan dan bukan penipuan. *Confusion matrix* pada Tabel 10 memberikan informasi tentang kinerja model dalam memprediksi transaksi yang benar dan salah untuk masing-masing kelas (penipuan dan bukan penipuan).

Tabel 10. *Confusion Matrix* Model XGboost pada Dataset *Augmented* (Seimbang)

	Predicted Positive (Penipuan)	*Predicted Negative* (Bukan Penipuan)
Actual Positive (Penipuan)	6751	339
Actual Negative (Bukan Penipuan)	18	6706

Model pada Tabel 10 menghasilkan 6706 TN, yaitu jumlah transaksi yang sebenarnya bukan penipuan dan diklasifikasikan dengan benar oleh model sebagai bukan penipuan. Nilai ini menunjukkan bahwa model sangat baik dalam mengenali transaksi yang tidak mencurigakan sebagai kelas mayoritas dalam dataset. Sebaliknya, terdapat 339 FN, yaitu transaksi penipuan yang salah diklasifikasikan sebagai transaksi non-penipuan. Hal ini menunjukkan bahwa meskipun model sudah lebih baik, masih ada transaksi penipuan yang tidak terdeteksi. Namun, jumlah FN ini sudah berkurang dibandingkan dengan sebelum dilakukan data *augmentation*. Selain itu, terdapat 18 FP, yaitu transaksi yang sebenarnya bukan penipuan namun salah diklasifikasikan oleh model sebagai penipuan. Meskipun jumlah FP relatif kecil, hal ini menunjukkan adanya kesalahan dalam mengidentifikasi transaksi yang sah sebagai penipuan, yang dapat mengganggu pengalaman pelanggan karena memerlukan verifikasi lebih lanjut. TP, yaitu jumlah transaksi penipuan yang berhasil diklasifikasikan dengan benar sebagai penipuan sebanyak 6751. Hal ini menunjukkan bahwa model mampu mendeteksi sebagian besar dari keseluruhan transaksi penipuan yang mengindikasikan peningkatan kemampuan model dalam mengidentifikasi kasus penipuan dengan tepat.

Tabel 11. Hasil Evaluasi Model XGboost pada Dataset *Augmented* (Seimbang)

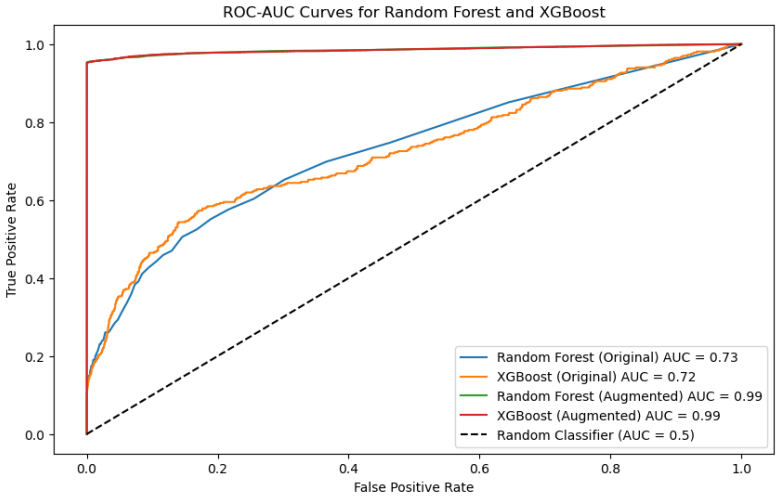
Metrik	**Nilai (%)**
Accuracy	97,42
Precision	99,73
Recall	95,22
F1-*score*	97,42

Tabel 11 menunjukkan hasil evaluasi model XGBoost pada dataset yang telah diaugmentasi dengan CGAN dan seimbang. *Accuracy* dari model adalah 97,42% yang terlihat sangat tinggi. Hal ini menunjukkan bahwa model mampu mengklasifikasikan transaksi dengan baik secara keseluruhan. *Precision* untuk kelas penipuan adalah 99,73% yang menunjukkan bahwa dari semua transaksi yang diprediksi sebagai penipuan hampir semuanya benar-benar merupakan penipuan. Hal ini berarti

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

model sangat baik dalam mengurangi *false positive*. *Recall* untuk kelas penipuan adalah 95,22% yang cukup tinggi. Hal ini mengindikasikan bahwa model mampu mendeteksi sebagian besar transaksi penipuan yang sebenarnya, yang penting dalam konteks deteksi penipuan untuk mengurangi risiko kerugian finansial. F1-*score* untuk kelas penipuan adalah 97,42% yang merupakan rata-rata harmonis antara *precision* dan *recall*. Nilai F1-*score* yang tinggi menunjukkan bahwa model seimbang dalam mendeteksi kasus penipuan, dengan *precision* dan *recall* yang sama-sama tinggi. Hasil klasifikasi menggunakan XGBoost pada dataset yang telah diaugmentasi menunjukkan peningkatan performa yang signifikan dalam mendeteksi penipuan. Tingginya nilai TP dan rendahnya FN mengindikasikan bahwa model memiliki kemampuan yang lebih baik dalam mempelajari karakteristik kelas penipuan setelah *data augmentation*. Oleh karena itu, strategi *data augmentation* menggunakan CGAN terbukti efektif dalam meningkatkan performa model dalam mendeteksi transaksi penipuan.

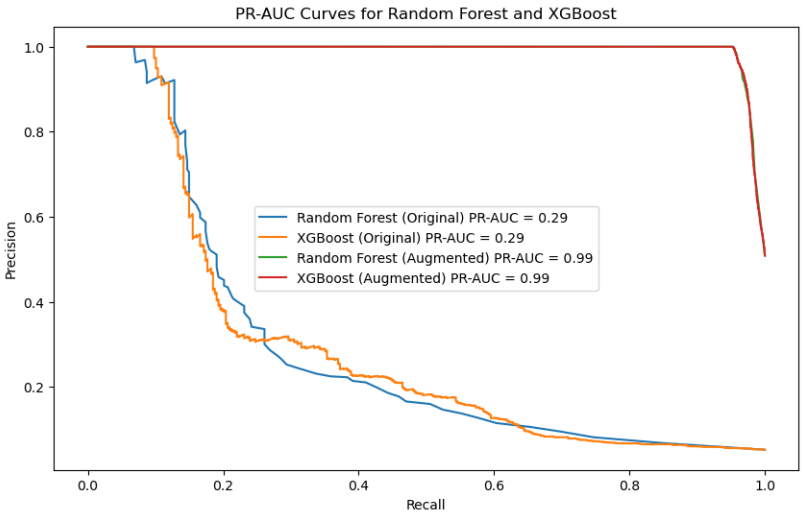
Berdasarkan hasil evaluasi dari model Random Forest dan XGBoost pada dataset yang telah diaugmentasi menggunakan CGAN dan seimbang, keduanya menunjukkan peningkatan performa yang signifikan dalam mendeteksi transaksi penipuan. Model Random Forest mencapai *accuracy* sebesar 97,50%, *precision* 99,99%, *recall* 95,15%, dan F1-*score* 97,47%, sedangkan model XGBoost memiliki *accuracy* 97,42%, *precision* 99,73%, *recall* 95,22%, dan F1-*score* 97,42%. Meskipun kedua model menunjukkan performa yang sangat baik, Random Forest sedikit unggul dalam hal *precision* dan F1-*score*. *Precision* yang sangat tinggi pada kedua model menunjukkan bahwa sebagian besar prediksi penipuan adalah benar-benar penipuan yang mengurangi kemungkinan *false positive*. *Recall* yang tinggi juga menunjukkan bahwa model mampu mendeteksi sebagian besar transaksi penipuan yang sebenarnya sehingga risiko penipuan yang tidak terdeteksi menjadi lebih kecil. F1-*score* yang tinggi pada kedua model menunjukkan keseimbangan yang sangat baik antara *precision* dan *recall*,. Hal ini mengindikasikan performa yang optimal dalam mendeteksi penipuan pada dataset yang telah seimbang. Secara keseluruhan, hasil ini menunjukkan bahwa *data augmentation* dengan CGAN efektif dalam meningkatkan kemampuan model dalam mendeteksi transaksi penipuan. Random Forest maupun XGBoost dapat diandalkan dalam mendeteksi penipuan dengan akurasi yang tinggi.



Gambar 4. Grafik ROC-AUC

Gambar 4 menunjukkan kurva ROC-AUC untuk dua model *machine learning*, yaitu Random Forest dan XGBoost yang diterapkan pada dataset asli dan tidak seimbang serta dataset setelah diaugmentasi dengan CGAN sehingga dataset menjadi seimbang. Pada dataset asli yang tidak seimbang, transaksi bukan penipuan jauh lebih banyak daripada transaksi penipuan, kedua model menunjukkan performa yang lebih rendah. Kurva ROC untuk kedua model pada dataset asli menunjukkan AUC yang lebih rendah, yang mengindikasikan bahwa model kesulitan dalam membedakan antara transaksi penipuan dan bukan penipuan. Hal ini terjadi karena dominasi kelas mayoritas (bukan penipuan) yang membuat model lebih cenderung memprediksi transaksi sebagai bukan penipuan meskipun ada transaksi penipuan yang perlu dideteksi. Setelah data diaugmentasi menggunakan CGAN yang menambahkan transaksi penipuan sintetis untuk menyeimbangkan dataset maka kedua model menunjukkan peningkatan yang signifikan pada ROC-AUC. Kurva ROC pada data yang diaugmentasi terlihat lebih mendekati sudut kiri atas (*True Positive Rate* (TPR) tinggi dan *False Positive Rate* (FPR) rendah). Hal inimenunjukkan bahwa model sekarang lebih mampu membedakan transaksi penipuan dari bukan penipuan. Peningkatan ini tercermin pada AUC yang lebih tinggi yang menunjukkan bahwa dengan dataset yang lebih seimbang, model dapat mempelajari karakteristik transaksi penipuan dan bukan penipuan dengan lebih baik. Secara keseluruhan, XGBoost menunjukkan performa yang sedikit lebih baik dibandingkan Random Forest pada dataset asli maupun data diaugmentasi. Hal ini tercermin dalam AUC yang lebih tinggi. Namun, Random Forest juga mengalami peningkatan yang signifikan setelah augmentasi data. Hal ini menunjukkan bahwa kedua model mendapatkan manfaat yang jelas dari teknik augmentasi data berbasis CGAN. Data yang seimbang dapat meningkatkan kinerja model dalam membedakan transaksi penipuan dari transaksi yang sah, mengurangi bias terhadap kelas mayoritas, dan meningkatkan kemampuan deteksi penipuan secara keseluruhan.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430



Gambar 5. Grafik PR-AUC

Grafik PR-AUC di Gambar 5 menunjukkan performa dua model, yaitu Random Forest dan XGBoost yang diterapkan pada dataset asli dan tidak seimbang serta dataset setelah diaugmentasi dengan CGAN menjadi dataset yang seimbang. Pada dataset asli yang tidak seimbang, terlihat bahwa *precision* (akurasi prediksi penipuan) dan *recall* (kemampuan model untuk mendeteksi penipuan) untuk kedua model cukup rendah. Hal ini menunjukkan bahwa model kesulitan dalam mendeteksi transaksi penipuan karena jumlah transaksi non-penipuan jauh lebih banyak dan model lebih cenderung untuk memprediksi transaksi sebagai bukan penipuan. Hal ini mengarah pada F1-*score* yang rendah karena meskipun *precision* mungkin tidak terlalu buruk didapatkan *recall* sangat rendah, artinya banyak transaksi penipuan yang terlewat. Setelah dilakukan data augmentasi menggunakan CGAN yang menyeimbangkan dataset dengan menambah transaksi penipuan sintetis didapatkan peningkatan yang signifikan dalam *precision* dan *recall* untuk kedua model. Setelah augmentasi, kedua model (Random Forest dan XGBoost) mengalami peningkatan yang jelas dalam kemampuan untuk mendeteksi transaksi penipuan tanpa mengorbankan terlalu banyak dari transaksi bukan penipuan yang salah diklasifikasikan. F1-*score* juga meningkat, mencerminkan perbaikan keseimbangan antara *precision* dan *recall*. Peningkatan yang signifikan setelah augmentasi data ini menunjukkan bahwa dengan dataset yang lebih seimbang, model dapat belajar dengan lebih baik untuk mengidentifikasi transaksi penipuan. Teknik augmentasi data berbasis CGAN terbukti efektif untuk meningkatkan kemampuan model dalam mendeteksi penipuan, mengurangi bias terhadap transaksi bukan penipuan, dan memberikan performa yang lebih optimal. Secara keseluruhan, meskipun XGBoost sedikit lebih unggul dibandingkan Random Forest dalam hal *precision* dan *recall*, kedua model menunjukkan peningkatan yang sangat signifikan setelah augmentasi dataset. Hal ini menegaskan pentingnya data *balancing* dalam meningkatkan kemampuan model deteksi penipuan, terutama ketika berhadapan dengan dataset yang sangat tidak seimbang.

3.2 Pembahasan

Berdasarkan hasil evaluasi model Random Forest dan XGBoost setelah menggunakan augmentasi data berbasis CGAN pada dataset yang seimbang, terlihat bahwa penggunaan teknik ini memberikan dampak yang signifikan terhadap peningkatan kinerja deteksi penipuan oleh kedua model. Pada dataset asli yang tidak seimbang, kedua model menunjukkan kesulitan dalam mendeteksi transaksi penipuan dengan efektif, yang disebabkan oleh dominasi besar dari transaksi bukan penipuan. Setelah menerapkan augmentasi data dengan CGAN, kedua model menunjukkan peningkatan performa yang signifikan pada semua metrik kinerja utama, seperti yang dapat dilihat dari hasil evaluasi pada Tabel 7 dan Tabel 9.

Pada model Random Forest, setelah augmentasi data, *accuracy* meningkat menjadi 97,50%, *precision* mencapai 99,99%, *recall* sebesar 95,15%, dan F1-*score* mencapai 97,47%. Peningkatan *precision* yang sangat tinggi menunjukkan bahwa model mampu mengurangi jumlah *false positive* dengan sangat baik, sementara peningkatan *recall* mengindikasikan kemampuan model dalam mendeteksi sebagian besar transaksi penipuan yang sebenarnya. F1-*score* yang tinggi menggambarkan keseimbangan yang baik antara *precision* dan *recall*, menunjukkan bahwa model Random Forest setelah augmentasi data mampu memberikan performa yang sangat baik dalam mendeteksi transaksi penipuan.

Sementara itu, model XGBoost pada dataset yang sama juga menunjukkan peningkatan performa yang signifikan setelah menggunakan augmentasi data berbasis CGAN. *Accuracy* model mencapai 97,42%, *precision* sebesar 99,73%, *recall* sebesar 95,22%, dan F1-*score* sebesar 97,42%. Hasil ini menunjukkan bahwa XGBoost, seperti halnya Random Forest, mampu meningkatkan performanya dalam mendeteksi penipuan setelah dataset diseimbangkan. *Precision* yang tinggi mengindikasikan bahwa sebagian besar prediksi penipuan benar-benar merupakan penipuan, sementara *recall* yang meningkat menunjukkan bahwa model dapat mendeteksi lebih banyak transaksi penipuan dibandingkan dengan saat menggunakan dataset yang tidak seimbang.

Saat membandingkan kedua model, Random Forest dan XGBoost menunjukkan performa yang hampir seimbang setelah augmentasi data. Random Forest sedikit lebih unggul dalam hal *precision* dan F1-*score*, sedangkan XGBoost menunjukkan keunggulan dalam *recall*. Hal ini menunjukkan bahwa kedua model mampu memanfaatkan data sintetis yang dihasilkan oleh

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

CGAN dengan sangat baik untuk meningkatkan performa. Meskipun XGBoost memiliki performa yang lebih baik pada dataset asli yang tidak seimbang, Random Forest mampu menunjukkan peningkatan yang lebih signifikan setelah augmentasi data, yang berarti model ini lebih responsif terhadap peningkatan variasi dan jumlah contoh penipuan yang dihasilkan oleh CGAN.

Secara keseluruhan, hasil ini menekankan bahwa augmentasi data berbasis CGAN sangat efektif dalam meningkatkan performa deteksi penipuan untuk model Random Forest dan XGBoost. Teknik ini berhasil menyeimbangkan dataset yang awalnya sangat tidak seimbang sehingga memungkinkan kedua model untuk mempelajari karakteristik transaksi penipuan dengan lebih baik. Peningkatan performa ini tidak hanya ditunjukkan oleh peningkatan *accuracy*, tetapi juga oleh peningkatan *precision*, *recall*, dan F1-*score* yang signifikan. Oleh karena itu, augmentasi data berbasis CGAN dapat dianggap sebagai solusi efektif untuk mengatasi masalah ketidakseimbangan kelas dalam sistem deteksi penipuan dan membantu meningkatkan keandalan model dalam mendeteksi transaksi penipuan.

4. KESIMPULAN

Temuan penelitian ini menegaskan efektivitas Conditional Generative Adversarial Network (CGAN) sebagai alat augmentasi data untuk mengatasi ketidakseimbangan kelas dalam deteksi penipuan e-commerce. Hasil evaluasi menunjukkan bahwa transaksi penipuan sintetis yang dihasilkan oleh CGAN secara signifikan meningkatkan kinerja model Random Forest dan XGBoost dalam mendeteksi transaksi penipuan. Pada dataset asli yang tidak seimbang, kedua model mengalami kesulitan dalam mendeteksi transaksi penipuan secara efektif, dengan nilai recall yang sangat rendah. Namun, setelah menerapkan augmentasi data dengan CGAN, peningkatan performa yang signifikan terlihat pada semua metrik kinerja utama. Pada model Random Forest, setelah augmentasi data, accuracy meningkat menjadi 97,50%, precision mencapai 99,99%, recall sebesar 95,15%, dan F1-score mencapai 97,47%. Hasil ini menunjukkan bahwa augmentasi data tidak hanya meningkatkan kemampuan model untuk mendeteksi transaksi penipuan (recall yang tinggi) tetapi juga mengurangi false positives secara signifikan, sebagaimana ditunjukkan oleh precision yang hampir sempurna. Model XGBoost juga menunjukkan peningkatan performa yang signifikan setelah augmentasi data. Pada dataset yang telah diseimbangkan, model ini mencapai accuracy sebesar 97,42%, precision sebesar 99,73%, recall sebesar 95,22%, dan F1-score sebesar 97,42%. Meskipun nilai precision dan F1-score sedikit lebih rendah dibandingkan Random Forest, XGBoost tetap menunjukkan performa yang sangat baik, dengan kemampuan mendeteksi penipuan yang lebih akurat dibandingkan dataset yang tidak seimbang. Secara keseluruhan, XGBoost menunjukkan kinerja yang sedikit lebih unggul dalam hal keseimbangan antara precision dan recall dibandingkan Random Forest. Peningkatan ini menunjukkan bahwa augmentasi data berbasis CGAN memberikan manfaat yang signifikan, memungkinkan kedua model untuk belajar dari dataset yang lebih representatif terhadap transaksi penipuan. Pendekatan ini memiliki implikasi praktis bagi platform e-commerce, memungkinkan sistem deteksi penipuan yang lebih akurat untuk meminimalkan kerugian finansial akibat transaksi yang tidak terdeteksi dan mengurangi gangguan bagi pelanggan akibat false positives. Namun, penelitian ini juga memiliki keterbatasan, seperti biaya komputasi yang tinggi untuk melatih CGAN dan kompleksitas implementasi. Penelitian di masa depan dapat berfokus pada pengembangan arsitektur CGAN yang lebih efisien, mengeksplorasi teknik augmentasi data alternatif, atau mengaplikasikan metode ini pada dataset lain untuk menggeneralisasi temuan. Pendekatan ini berpotensi menghasilkan sistem pencegahan penipuan yang lebih andal, meningkatkan kepercayaan pelanggan dalam transaksi online.

REFERENCES

[1] A. Saputra and Suharjito, "Fraud detection using machine learning in e-commerce," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 9, pp. 332–339, 2019, doi: 10.14569/ijacsa.2019.0100943.
[2] 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. September, p. 100597, 2024, doi: 10.1016/j.mlwa.2024.100597.
[3] I. de Zarzà, J. de Curtò, and C. T. Calafate, "Optimizing Neural Networks for Imbalanced Data," *Electron.*, vol. 12, no. 12, pp. 1–26, 2023, doi: 10.3390/electronics12122674.
[4] T. Karthikeyan, M. Govindarajan, and V. Vijayakumar, "An effective fraud detection using competitive swarm optimization based deep neural network," *Meas. Sensors*, vol. 27, no. December 2022, p. 100793, 2023, doi: 10.1016/j.measen.2023.100793.
[5] H. A. Gameng, B. D. Gerardo, and R. P. Medina, "A modified adaptive synthetic smote approach in graduation success rate classification," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 6, pp. 3053–3057, 2019, doi: 10.30534/ijatcse/2019/63862019.
[6] A. Mutemi and F. Bacao, "E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review," *Big Data Min. Anal.*, vol. 7, no. 2, pp. 419–444, 2024, doi: 10.26599/BDMA.2023.9020023.
[7] A. Cherif, H. Ammar, M. Kalkatawi, S. Alshehri, and A. Imine, "Encoder–decoder graph neural network for credit card fraud detection," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 3, p. 102003, 2024, doi: 10.1016/j.jksuci.2024.102003.
[8] R. Damayanti and Z. Adrianto, "Machine Learning For E-Commerce Fraud Detection," *J. Ris. Akunt. Dan Bisnis Airlangga*, vol. 8, no. 2, pp. 1562–1577, Nov. 2023, doi: 10.20473/jraba.v8i2.48559.
[9] E. Khan, M. Zia Ur Rehman, F. Ahmed, S. A. Alsuhibany, M. Zulfiqar Ali, and J. Ahmad, "An Automated Classification Technique for COVID-19 Using Optimized Deep Learning Features," *Comput. Syst. Sci. Eng.*, vol. 46, no. 3, pp. 3799–3814, 2023, doi: 10.32604/csse.2023.037131.
[10] B. Lebichot, T. Verhelst, Y.-A. Le Borgne, L. He-Guelton, F. Oble, and G. Bontempi, "Transfer Learning Strategies for Credit Card Fraud Detection," *IEEE Access*, vol. 9, pp. 114754–114766, 2021, doi: 10.1109/ACCESS.2021.3104472.
[11] D. Sisodia and D. S. Sisodia, "A transfer learning framework towards identifying behavioral changes of fraudulent publishers

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1919−1931
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6430

in pay-per-click model of online advertising for click fraud detection," *Expert Syst. Appl.*, vol. 232, p. 120922, Dec. 2023, doi: 10.1016/j.eswa.2023.120922.

[12] M. Azim Mim, N. Majadi, and P. Mazumder, "A soft voting ensemble learning approach for credit card fraud detection," *Heliyon*, vol. 10, no. 3, p. e25466, Feb. 2024, doi: 10.1016/j.heliyon.2024.e25466.

[13] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," *IEEE Access*, vol. 10, pp. 39700–39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[14] Y. Bing Chu, Z. Min Lim, B. Keane, P. Hao Kong, A. Rafat Elkilany, and O. Hisham Abusetta, "Credit Card Fraud Detection on Original European Credit Card Holder Dataset Using Ensemble Machine Learning Technique," *J. Cyber Secur.*, vol. 5, no. 0, pp. 33–46, 2023, doi: 10.32604/jcs.2023.045422.

[15] N. Mqadi, N. Naicker, and T. Adeliyi, "A SMOTe based Oversampling Data-Point Approach to Solving the Credit Card Data Imbalance Problem in Financial Fraud Detection," *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 277–286, Feb. 2021, doi: 10.12785/ijcds/100128.

[16] R. Bounab, K. Zarour, B. Guelib, and N. Khlifa, "Enhancing Medicare Fraud Detection Through Machine Learning: Addressing Class Imbalance With SMOTE-ENN," *IEEE Access*, vol. 12, pp. 54382–54396, 2024, doi: 10.1109/ACCESS.2024.3385781.

[17] J. Lee, D. Jung, J. Moon, and S. Rho, "Advanced R-GAN: Generating anomaly data for improved detection in imbalanced datasets using regularized generative adversarial networks," *Alexandria Eng. J.*, vol. 111, no. September 2024, pp. 491–510, 2025, doi: 10.1016/j.aej.2024.10.084.

[18] M. J. Madhurya, H. L. Gururaj, B. C. Soundarya, K. P. Vidyashree, and A. B. Rajendra, "Exploratory analysis of credit card fraud detection using machine learning techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 31–37, 2022, doi: 10.1016/j.gltp.2022.04.006.

[19] M. Â. L. Moreira *et al.*, "Exploratory analysis and implementation of machine learning techniques for predictive assessment of fraud in banking systems," *Procedia Comput. Sci.*, vol. 214, no. C, pp. 117–124, 2022, doi: 10.1016/j.procs.2022.11.156.

[20] 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, 2022, doi: 10.1016/j.procs.2023.01.231.

[21] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, "A Neural Network Ensemble With Feature Engineering for Improved Credit Card Fraud Detection," *IEEE Access*, vol. 10, pp. 16400–16407, 2022, doi: 10.1109/ACCESS.2022.3148298.

[22] J. K. Afriyie *et al.*, "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," *Decis. Anal. J.*, vol. 6, no. December 2022, p. 100163, 2023, doi: 10.1016/j.dajour.2023.100163.

[23] L. Cao and H. Shen, "Imbalanced data classification based on hybrid resampling and twin support vector machine," *Comput. Sci. Inf. Syst.*, vol. 14, no. 3, pp. 579–595, 2017, doi: 10.2298/CSIS161221017L.

[24] H. R. Sneha and B. Annappa, "Exploratory Analysis of Methods, Techniques, and Metrics to Handle Class Imbalance Problem," *Procedia Comput. Sci.*, vol. 235, pp. 863–877, 2024, doi: 10.1016/j.procs.2024.04.082.

[25] S. R. Byrapu Reddy, P. Kanagala, P. Ravichandran, D. R. Pulimamidi, P. V. Sivarambabu, and N. S. A. Polireddi, "Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics," *Meas. Sensors*, vol. 33, no. April, p. 101138, 2024, doi: 10.1016/j.measen.2024.101138.

[26] Kanika, J. Singla, A. K. Bashir, Y. Nam, N. U. I. Hasan, and U. Tariq, "Handling class imbalance in online transaction fraud detection," *Comput. Mater. Contin.*, vol. 70, no. 2, pp. 2861–2877, 2022, doi: 10.32604/cmc.2022.019990.

[27] S. Jagtap, *Fraudulent E-Commerce Transactions*, Kaggle, Apr. 2024. [Online]. Available: https://www.kaggle.com/datasets/shriyashjagtap/fraudulent-e-commerce-transactions.