

OPTIMALISASI PARAMETER SCALE POS WEIGTH PADA XGBOOST BERBASIS ALGORITMA GENETIKA UNTUK DETEKSI FRAUD YANG PRESISI

Azril Ardiansyah¹, Martanto², Raditya Danar Dana³, Riri Narasati⁴.

Program Studi Teknik Informatika¹⁴
Program Studi Manajemen Informatika²³

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
azrilardnsyh11@gmail.com

(*) Corresponding Author : azrilardnsyh11@gmail.com
Published : 17 Juli 2026

Abstract—Fraud detection in financial transactions is a critical issue in digital security systems, particularly due to the highly imbalanced nature of the data, where the number of normal transactions significantly exceeds fraudulent ones. This imbalance causes classification models to be biased toward the majority class, thereby reducing their ability to detect fraud. This study aims to optimize the `scale_pos_weight` parameter in the XGBoost algorithm to improve the precision of fraud detection performance. The research questions addressed in this study are: (1) What is the performance of the model before applying the `scale_pos_weight` parameter? (2) What is the performance after applying the `scale_pos_weight` parameter? (3) What is the performance after optimization using a Genetic Algorithm? (4) How does the model performance compare before applying `scale_pos_weight`, after applying it, and after optimization using a Genetic Algorithm? The methods employed include Exploratory Data Analysis (EDA), modeling using XGBoost without parameters, with the `scale_pos_weight` parameter, and parameter optimization using a Genetic Algorithm. Model evaluation is conducted using AUPRC, precision, recall, and F1-score, which are suitable metrics for imbalanced datasets. The results based on the confusion matrix indicate significant performance differences across scenarios. In the baseline model (without `scale_pos_weight`), the values obtained were TN=56,850 (normal data), FP=14 (normal data misclassified as fraud), FN=23 (fraud data misclassified as normal), and TP=75 (fraud data correctly detected). This indicates high precision but limited recall due to undetected fraud cases. In the model with `scale_pos_weight`, recall improved as FN decreased to 9 and TP increased to 89; however, FP sharply increased to 1,044, causing a significant drop in precision. After optimization using a Genetic Algorithm with `scale_pos_weight` = 40.9, a more balanced result was achieved with TN=56,807, FP=57, FN=11, and TP=87. This model maintains high recall while significantly reducing false positives compared to the initial weighted model. Therefore, the optimized model demonstrates more stable and balanced performance between precision and recall, making it more suitable for real-world fraud detection applications. In conclusion, optimizing the `scale_pos_weight` parameter using a Genetic Algorithm effectively improves the performance of the XGBoost model in fraud detection, resulting in a more balanced and practically applicable model.

Keywords: Imbalanced data, fraud detection, XGBoost, `scale_pos_weight`, genetic algorithm.

Abstrak—Deteksi fraud pada transaksi keuangan merupakan permasalahan penting dalam sistem keamanan digital, terutama karena karakteristik data yang sangat tidak seimbang (imbalanced), dimana jumlah transaksi normal jauh lebih besar dibandingkan transaksi fraud. Ketidakseimbangan ini menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas sehingga menurunkan kemampuan dalam mendeteksi fraud. Penelitian ini bertujuan untuk mengoptimalkan parameter `scale_pos_weight` pada algoritma XGBoost guna meningkatkan performa deteksi fraud yang lebih presisi. Adapun rumusan masalah dalam penelitian ini adalah: (1) Apa hasil kinerja model sebelum menggunakan parameter `scale_pos_weight`? (2) Apa hasil kinerja model setelah menggunakan parameter `scale_pos_weight`? (3) Apa hasil kinerja model setelah dilakukan optimasi menggunakan algoritma genetika? (4) Bagaimana perbandingan kinerja model sebelum menggunakan `scale_pos_weight`, setelah menggunakan `scale_pos_weight`, dan setelah dilakukan optimasi menggunakan algoritma genetika? Metode yang digunakan meliputi eksplorasi data (Exploratory Data Analysis), pemodelan menggunakan XGBoost

tanpa parameter, dengan parameter *scale_pos_weight*, serta optimasi parameter menggunakan Algoritma Genetika (Genetic Algorithm). Evaluasi model dilakukan menggunakan metrik AUPRC, precision, recall, dan F1-score yang sesuai untuk dataset tidak seimbang. Hasil penelitian berdasarkan confusion matrix menunjukkan perbedaan kinerja yang signifikan pada setiap skenario. Pada model baseline sebelum memakai *scale_pos_weight*, diperoleh nilai TN=56850 (data Normal), FP=14 (data normal namun ditandai tidak normal), FN=23 (data fraud namun ditandai normal), dan TP=75 (data fraud), yang menunjukkan precision tinggi namun recall masih terbatas karena masih terdapat sejumlah fraud yang tidak terdeteksi. Pada model dengan parameter *scale_pos_weight*, terjadi peningkatan recall dengan FN menurun menjadi 9 dan TP meningkat menjadi 89, namun diikuti lonjakan FP menjadi 1044 yang menyebabkan precision menurun drastis. Setelah dilakukan optimasi menggunakan Algoritma Genetika dengan nilai bobot *scale_pos_weight*=40.9, diperoleh hasil yang lebih seimbang dengan TN=56807, FP=57, FN=11, dan TP=87. Model ini mampu mempertahankan recall yang tinggi sekaligus menekan jumlah false positive secara signifikan dibandingkan model berbobot awal. Dengan demikian, hasil optimasi menunjukkan performa yang lebih stabil dan seimbang antara precision dan recall, sehingga lebih optimal untuk implementasi pada deteksi fraud di dunia nyata. Kesimpulannya, optimasi parameter *scale_pos_weight* berbasis Algoritma Genetika efektif dalam meningkatkan kinerja model XGBoost pada deteksi fraud, sehingga menghasilkan model yang lebih seimbang dan layak diterapkan pada kasus nyata.

Kata kunci: Data tidak seimbang, deteksi fraud, XGBoost, *scale_pos_weight*, algoritma genetika.

INTRODUCTION

Perkembangan teknologi informasi dan komunikasi telah mendorong transformasi signifikan dalam sistem keuangan digital, khususnya pada layanan perbankan online, e-commerce, dan sistem pembayaran elektronik. Peningkatan volume transaksi digital yang masif diiringi dengan kompleksitas pola transaksi telah memunculkan tantangan baru dalam menjaga keamanan sistem keuangan. Salah satu tantangan utama yang semakin krusial adalah meningkatnya aktivitas kecurangan (*fraud*) yang semakin kompleks dan sulit dideteksi secara konvensional. Oleh karena itu, sistem deteksi fraud berbasis data menjadi komponen penting dalam menjaga integritas dan kepercayaan terhadap ekosistem keuangan digital [1].

Dalam konteks fraud detection, permasalahan utama yang dihadapi bukan hanya pada kompleksitas data, tetapi terutama pada karakteristik data yang tidak seimbang (*imbalanced data*). Pada umumnya, jumlah transaksi normal jauh lebih besar dibandingkan transaksi fraud, sehingga menciptakan distribusi kelas yang sangat timpang. Kondisi ini menyebabkan algoritma machine learning cenderung bias terhadap kelas mayoritas dan mengabaikan pola dari kelas minoritas yang justru menjadi fokus utama deteksi [2]. Namun, penelitian yang secara khusus mengkaji optimalisasi parameter *scale_pos_weight* pada algoritma XGBoost untuk menangani ketidakseimbangan data dalam deteksi fraud masih terbatas. Selain itu, pendekatan yang digunakan dalam menentukan nilai parameter tersebut umumnya masih bersifat manual atau

heuristik sederhana, sehingga belum mampu menghasilkan performa model yang optimal. Oleh karena itu, diperlukan suatu pendekatan optimasi yang lebih adaptif, seperti Algoritma Genetika, untuk memperoleh keseimbangan yang lebih baik antara precision dan recall dalam mendeteksi transaksi fraud. Akibatnya, model dapat menunjukkan nilai akurasi yang tinggi, namun memiliki kemampuan deteksi fraud yang rendah, terutama pada metrik recall dan precision.

Permasalahan data tidak seimbang ini menjadi isu fundamental dalam pengembangan model deteksi fraud modern. Model yang dilatih pada data imbalanced cenderung menghasilkan tingkat *false negative* yang tinggi, di mana transaksi fraud tidak terdeteksi oleh sistem. Hal ini sangat berbahaya karena dapat menyebabkan kerugian finansial yang signifikan bagi institusi keuangan. Selain itu, penggunaan metrik evaluasi konvensional seperti akurasi menjadi tidak relevan dalam konteks ini, sehingga diperlukan metrik alternatif seperti AUPRC, F1-score, dan recall untuk memberikan gambaran performa yang lebih representatif [3].

Berbagai pendekatan telah dikembangkan untuk mengatasi permasalahan ketidakseimbangan data, seperti teknik resampling, cost-sensitive learning, dan penyesuaian bobot kelas (*class weighting*). Salah satu pendekatan yang efektif adalah penggunaan parameter *scale_pos_weight* pada algoritma XGBoost, yang memungkinkan model memberikan bobot lebih besar pada kelas minoritas selama proses pelatihan [4]. Meskipun demikian, penentuan nilai parameter yang optimal masih menjadi tantangan karena sangat bergantung pada karakteristik distribusi data.

Sejumlah penelitian sebelumnya telah mengkaji penerapan machine learning dalam fraud detection

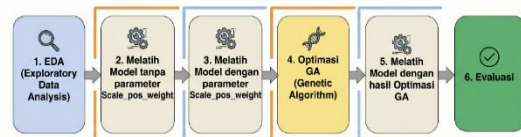
dengan mempertimbangkan isu data tidak seimbang. Studi oleh [5] menunjukkan bahwa algoritma berbasis ensemble seperti XGBoost memiliki performa yang baik, namun tetap memerlukan penyesuaian khusus untuk menangani class imbalance. Penelitian oleh [2] menegaskan bahwa performa model sangat dipengaruhi oleh strategi penanganan data tidak seimbang. Sementara itu, [6] menunjukkan bahwa metrik AUPRC lebih relevan dalam mengevaluasi model pada kondisi tersebut. Namun, sebagian besar penelitian masih terbatas pada pendekatan statis dalam penentuan parameter dan belum banyak mengeksplorasi metode optimasi adaptif.

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada optimalisasi parameter *scale_pos_weight* pada algoritma XGBoost untuk meningkatkan performa deteksi fraud pada data tidak seimbang. Untuk mencapai hal tersebut, digunakan pendekatan optimasi berbasis algoritma genetika (*Genetic Algorithm*) yang mampu mengeksplorasi ruang parameter secara global dan adaptif [7]. Pendekatan ini diharapkan dapat menemukan kombinasi parameter terbaik yang mampu meningkatkan sensitivitas model terhadap kelas minoritas tanpa mengorbankan performa keseluruhan.

Dengan demikian, penelitian ini memiliki kontribusi penting dalam mengatasi permasalahan fundamental pada fraud detection, yaitu ketidakseimbangan data. Hasil penelitian diharapkan dapat meningkatkan kemampuan model dalam mendeteksi transaksi fraud secara lebih akurat dan robust, serta memberikan solusi praktis yang dapat diterapkan pada sistem keuangan digital. Selain itu, penelitian ini juga diharapkan dapat memperkaya literatur terkait optimasi model machine learning pada data imbalanced, khususnya dalam konteks penggunaan XGBoost dan algoritma genetika.

MATERIALS AND METHODS

Dalam pengembangan sistem klasifikasi berbasis pembelajaran mesin, diperlukan alur metodologi yang sistematis agar model yang dihasilkan memiliki kinerja optimal dan dapat digeneralisasikan dengan baik. Proses ini tidak hanya melibatkan tahap pelatihan model, tetapi juga mencakup eksplorasi data, penyesuaian parameter, serta optimasi untuk memperoleh kombinasi terbaik. Gambar berikut menggambarkan tahapan lengkap yang digunakan dalam penelitian ini, mulai dari analisis awal data hingga evaluasi akhir model.



Gambar 1. Alur Penelitian

Gambar tersebut menunjukkan alur kerja sistem yang terdiri dari enam tahapan utama. Tahap pertama adalah *Exploratory Data Analysis (EDA)* yang bertujuan untuk memahami karakteristik data, mendeteksi pola, serta mengidentifikasi potensi masalah seperti missing values atau ketidakseimbangan data. Selanjutnya, pada tahap kedua dilakukan pelatihan model dengan parameter awal, yaitu *scale_pos_weight*, sebagai baseline untuk mengetahui performa awal model. Tahap ketiga masih berfokus pada pelatihan model, namun dengan penyesuaian parameter guna melihat pengaruh perubahan tersebut terhadap kinerja model.

Tahap keempat merupakan inti optimasi menggunakan *Genetic Algorithm (GA)*, di mana algoritma ini digunakan untuk mencari kombinasi parameter terbaik secara evolusioner melalui proses seleksi, crossover, dan mutasi. Hasil dari optimasi ini kemudian digunakan pada tahap kelima, yaitu pelatihan ulang model dengan parameter terbaik yang diperoleh dari GA. Terakhir, tahap keenam adalah evaluasi model untuk mengukur performa akhir menggunakan metrik yang relevan, sehingga dapat diketahui seberapa efektif optimasi yang telah dilakukan dalam meningkatkan akurasi dan stabilitas model.

RESULTS AND DISCUSSION

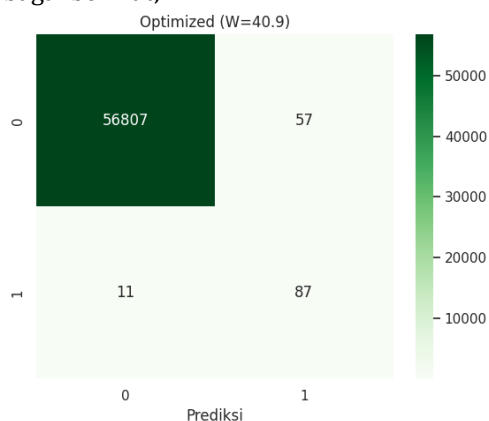
Pada tahapan ini, dilakukan pencarian parameter terbaik untuk model XGBoost menggunakan Algoritma Genetika (GA). Parameter yang dioptimasi adalah *scale_pos_weight* (bobot untuk kelas positif/fraud) dan *max_depth* (kedalaman maksimum pohon) untuk meningkatkan performa model. Proses GA berjalan selama 5 generasi dengan populasi 8, dan pada setiap generasi dilaporkan *scale_pos_weight* dan *max_depth* terbaik yang ditemukan beserta AUPRC Validasi-nya. Adapun hasilnya adalah sebagai berikut :

Tabel 1 Optimasi GA

Gen 1	Best Weight = 73.79	Depth = 3	Val AUPRC = 0.7327
Gen 2	Best Weight = 38.26	Depth = 3	Val AUPRC = 0.7577
Gen 3	Best Weight = 38.26	Depth = 3	Val AUPRC = 0.7577
Gen 4	Best Weight = 43.73	Depth = 3	Val AUPRC = 0.7614
Gen 5	Best Weight = 40.85	Depth = 3	Val AUPRC = 0.7647

Dalam tabel, kita bisa melihat bagaimana `scale_pos_weight` bervariasi dari 73.79 menjadi 38.26, lalu 43.73, dan akhirnya 40.85 pada Gen 5, sementara `max_depth` tetap stabil di 3. Peningkatan Val AUPRC dari 0.7327 di Gen 1 menjadi 0.7647 di Gen 5 menunjukkan bahwa GA berhasil menemukan kombinasi parameter yang secara progresif lebih baik. Kombinasi `scale_pos_weight` = 40.85 dan `max_depth` = 3 inilah yang kemudian menjadi nilai terbaik.

Pada tahap pemodelan ini, model XGBoost dilatih menggunakan parameter terbaik yang ditemukan pada tahap optimasi Genetic Algorithm (GA) di tahap sebelumnya. Dalam hal ini, `scale_pos_weight` diatur ke 40.85 dan `max_depth` 3. Tahapan ini fokus pada pembangunan model akhir dengan konfigurasi yang sudah disempurnakan. Hasil performa dari model ini disajikan dalam confusion matriks sebagai berikut;

**Gambar 2 Confussion Matriks 3**

Pada tahap pemodelan ini dihasilkan;

True Negative (TN): 56.807 Ini berarti 56.807 transaksi yang sebenarnya bukan fraud (normal) berhasil diprediksi dengan benar oleh model sebagai transaksi normal. Ini adalah

indikator yang sangat baik bahwa model tidak memberikan terlalu banyak 'alarm palsu' untuk transaksi yang sah.

False Positive (FP): 57 Ini menunjukkan bahwa ada 57 transaksi yang sebenarnya normal, namun salah diprediksi oleh model sebagai transaksi fraud. Ini adalah 'alarm palsu' yang, meskipun relatif kecil, dapat menimbulkan ketidaknyamanan bagi pengguna yang transaksinya salah terdeteksi sebagai fraud.

False Negative (FN): 11 Angka ini sangat penting dalam deteksi fraud. Ini berarti ada 11 transaksi yang sebenarnya adalah fraud, tetapi model gagal mendeteksinya dan memprediksinya sebagai transaksi normal. Ini adalah fraud yang 'lolos' dari deteksi model, dan idealnya angka ini harus sekecil mungkin.

True Positive (TP): 87 Ini adalah 87 transaksi yang sebenarnya fraud, dan berhasil dideteksi dengan benar oleh model sebagai fraud. Angka ini menunjukkan seberapa baik model dalam menangkap aktivitas penipuan yang sebenarnya.

Secara keseluruhan, confusion matrix untuk model yang dioptimasi GA menunjukkan keseimbangan yang baik antara mendeteksi fraud (TP yang tinggi yaitu 87, FN yang rendah yaitu 11) dan meminimalkan alarm palsu (FP yang rendah dengan nilai 57). Ini adalah hasil dari proses optimasi menggunakan Genetic Algorithm yang berhasil menemukan bobot (`scale_pos_weight`) dan kedalaman pohon (`max_depth`) yang paling efektif untuk dataset yang digunakan.

Tabel 2 Evaluasi Model

Metrik	Tanpa Parameter Scale_pos_weight	Menggunakan Parameter Scale_pos_weight	Hasil Optimasi
AUPRC	0.845895	0.716107	0.811396
Recall (Fraud)	0.765306	0.908163	0.887755
Precisi on (Fraud)	0.842697	0.078553	0.604167
F1-Score (Fraud)	0.802139	0.144598	0.719008

AUPRC (Area Under the Precision-Recall Curve): Metrik ini sangat cocok untuk dataset yang tidak seimbang (imbalanced), seperti kasus deteksi

fraud. AUPRC mengukur kualitas model dalam mengidentifikasi kelas minoritas (fraud) tanpa menghasilkan terlalu banyak positif palsu (False Positive). Nilai yang lebih tinggi menunjukkan kinerja yang lebih baik.

CONCLUSION

Berdasarkan hasil penelitian yang telah dilakukan mengenai optimalisasi parameter *scale_pos_weight* pada algoritma XGBoost menggunakan Genetic Algorithm untuk deteksi fraud, maka diperoleh beberapa kesimpulan utama sebagai berikut:

Pada model awal (baseline) tanpa penggunaan parameter *scale_pos_weight*, diperoleh kinerja model yang menunjukkan nilai precision yang relatif tinggi yaitu 0.842697, namun memiliki keterbatasan dalam mendeteksi kasus fraud yang ditunjukkan oleh nilai false negative yang masih cukup besar yaitu 23. Hal ini mengindikasikan bahwa model masih bias terhadap kelas mayoritas sehingga belum optimal dalam mengidentifikasi transaksi fraud.

Penggunaan parameter *scale_pos_weight* pada model XGBoost terbukti mampu meningkatkan kemampuan model dalam mendeteksi kelas minoritas, yang ditunjukkan dengan meningkatnya nilai recall sebesar 0.908163 dan berkurangnya jumlah false negative menjadi 9. Namun demikian, peningkatan tersebut diikuti dengan kenaikan jumlah false positive yang cukup signifikan sebesar 1.044, sehingga menyebabkan penurunan nilai precision menjadi 0.078553. Kondisi ini menunjukkan adanya trade-off antara recall dan precision dalam penanganan data tidak seimbang.

Hasil optimasi parameter *scale_pos_weight* menggunakan algoritma genetika (GA) menunjukkan peningkatan performa yang lebih seimbang dibandingkan kedua model sebelumnya. Nilai parameter optimal yang diperoleh mampu meningkatkan recall menjadi 0.8878 sekaligus menekan jumlah false positive ke 57, sehingga menghasilkan keseimbangan yang lebih baik antara precision (0.604167) dan recall (0.8878). Hal ini menunjukkan bahwa pendekatan optimasi berbasis algoritma genetika efektif dalam menemukan nilai parameter yang optimal secara adaptif.

Berdasarkan hasil perbandingan kinerja model, dapat disimpulkan bahwa model XGBoost yang dioptimasi menggunakan algoritma genetika memiliki performa terbaik dibandingkan model baseline maupun model dengan *scale_pos_weight* tanpa optimasi. Model

ini mampu memberikan hasil evaluasi yang lebih stabil dan representatif berdasarkan metrik AUPRC, precision, recall, dan F1-score, sehingga lebih layak untuk diterapkan dalam sistem deteksi fraud pada kondisi data tidak seimbang.

REFERENCE

- [1] M. Ali, "Fraud detection using machine learning in financial transactions," *Journal of Financial Crime*, 2021.
- [2] L. Hernandez Aros, L. X. Bustamante Molano, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, "Financial fraud detection through the application of machine learning techniques: A literature review," *Humanities and Social Sciences Communications*, vol. 11, pp. 1–22, 2024, doi: 10.1057/s41599-024-03606-0.
- [3] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLOS ONE*, 2020.
- [4] C. Wang, C. Deng, and S. Wang, "Imbalance-XGBoost: Leveraging weighted and focal losses for binary label-imbalanced classification," *Pattern Recognition Letters*, vol. 136, pp. 190–197, 2020.
- [5] A. Ali *et al.*, "Financial fraud detection based on machine learning: A systematic literature review," *Applied Sciences*, vol. 12, no. 19, p. 9637, 2021, doi: 10.3390/app12199637.
- [6] G. Velarde, A. Sudhir, S. Deshmane, A. Deshmunkh, K. Sharma, and V. Joshi, "Evaluating XGBoost for balanced and imbalanced data: Application to fraud detection," *arXiv preprint*, 2023, doi: 10.48550/arXiv.2303.15218.
- [7] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, pp. 8091–8126, 2021.