

INTISARI

PERBANDINGAN PERFORMA AUGMENTASI DATA SMOTE-WGAN DAN SMOTE-WCGAN PADA DETEKSI PENIPUAN KARTU KREDIT

Oleh
Naafi Rofiiqoh Makaarimah
23/525192/PPA/06595

Penipuan kartu kredit menghadirkan ancaman yang serius terhadap sistem keuangan, dan menuntut mekanisme deteksi yang akurat. Tantangan utama terletak pada ketidakseimbang antar kelas, yang menyebabkan model cenderung bias dan kinerja model menjadi tidak optimal. Pendekatan konvensional seperti Synthetic Minority Over-Sampling (SMOTE) seringkali menghasilkan sampel dengan variabilitas terbatas, sementara Generative Adversarial Network (GAN) menghadapi permasalahan mode collapse dan ketidakstabilan dalam pelatihan. Penelitian ini mengusulkan pendekatan hibrida dengan mengkombinasikan SMOTE dengan Wasserstein GAN – Gradient Penalty (WGAN) dan Conditional WGAN – Gradient Penalty (WCGAN) untuk menghasilkan sampel minoritas yang lebih representatif. Evaluasi dilakukan pada dua dataset yaitu, CCF dan FFSD dengan menggunakan algoritma XGBoost dan LGBM. Hasil penelitian menunjukkan bahwa metode augmentasi SMOTE-WGAN yang paling efektif. Metode ini konsisten meningkatkan performa klasifikasi pada kedua dataset, pada dataset CCF dengan algoritma XGBoost, metrik MCC meningkat 0.8209 menjadi 0.8524. Sementara pada dataset FFSD, MCC meningkat dari 0.6698 menjadi 0,6832, dan precision meningkatkan signifikan dari 0.9026 menjadi 0.9789. Sementara itu, SMOTE-WCGAN tidak memberikan peningkatan yang sebanding karena konflik antara sampel interpolatif SMOTE dan mekanisme kondisional pada WCGAN yang membatasi keragaman data sintesis. Temuan ini menegaskan bahwa kombinasi SMOTE dan WGAN menghasilkan data sintesis dengan utility yang lebih tinggi, sehingga lebih optimal untuk menangani ketidakseimbangan kelas dalam deteksi penipuan kartu kredit.

Kata Kunci: Deteksi Penipuan Kartu Kredit, SMOTE, WGAN, WCGAN, XGBoost, LightGBM

ABSTRACT

PERFORMANCE COMPARISON OF SMOTE-WGAN AND SMOTE-WCGAN DATA AUGMENTATION IN CREDIT CARD FRAUD DETECTION

By

Naafi Rofiiqoh Makaarimah

23/525192/PPA/06595

Credit card fraud poses a serious threat to the financial system and requires accurate detection mechanism. The main challenge in class imbalance, which causes models to be biased and their performance to be suboptimal. Conventional approaches like Synthetic Minority Over-sampling Technique (SMOTE) often generate insufficiently diverse samples, while Generative Adversarial Network (GAN) suffer from mode collapse and training instability. This study proposes a hybrid approaches by combining SMOTE with Wasserstein GAN – Gradient Penalty (WGAN) and SMOTE with Conditional WGAN Gradient Penalty (WCGAN) to generate more representative minority samples. The evaluation was conducted on two datasets, CCF and FFSD, using the XGBoost and LGBM algorithms. The results of the study show that the SMOTE-WGAN augmentation method is the most effective. This method consistently improves classification performance on both datasets. On the CCF dataset with the XGBoost algorithm, the MCC metric increased from 0.8209 to 0.8524. Meanwhile, on the FFSD dataset, the MCC increased from 0.6698 to 0.6832, and precision increased significantly from 0.9026 to 0.9789. Meanwhile, SMOTE-WCGAN did not provide a comparable improvement due to conflicts between SMOTE interpolative samples and the conditional mechanism in WCGAN, which limited the diversity of synthetic data. These findings confirm that the combination of SMOTE and WGAN produces synthetic data with higher utility, making it more optimal for handling class imbalance in credit card fraud detection.

Keywords: Credit Card Fraud Detection, SMOTE, WGAN, WCGAN, XGBoost, LightGBM