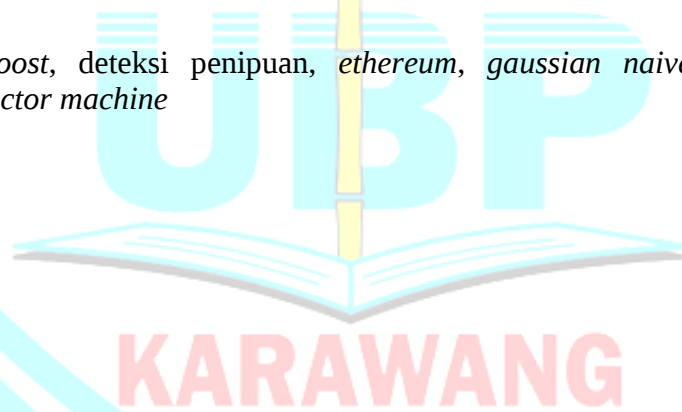


ABSTRAK

Penipuan pada transaksi *Ethereum* menjadi tantangan besar dalam ekosistem *blockchain*, sehingga diperlukan metode deteksi yang akurat. Penelitian ini bertujuan untuk membandingkan performa lima algoritma klasifikasi, yaitu *LightGBM*, *Extra Trees Classifier*, *Support Vector Machine* (SVM), *Gaussian Naive Bayes* (GNB), dan *CatBoost*, dalam mendeteksi penipuan pada dataset *Ethereum*. Metodologi penelitian meliputi pengumpulan data transaksi, pembersihan data, seleksi fitur, pembagian *dataset* (12,240 data latih, 1,969 data uji), penyeimbangan kelas dengan SMOTE, transformasi data (*MinMaxScaler* untuk SVM, *PowerTransformer* untuk GNB), pelatihan model dengan parameter *default*, dan optimasi *tuning* parameter menggunakan *GridSearchCV* dengan validasi silang 10-fold. Hasil menunjukkan *LightGBM* dan *CatBoost* unggul dengan *accuracy* 99,34% pada parameter *default*, meningkat menjadi 99,39% setelah *hyperparameter tuning* pada data uji. GNB mengalami peningkatan *accuracy* dari 89,18% menjadi 94,92%, sementara SVM menurun dari 98,93% menjadi 98,73%. *LightGBM* direkomendasikan sebagai model terbaik karena performa stabil dan seimbang.

Kata Kunci: *catboost*, deteksi penipuan, *ethereum*, *gaussian naive bayes*, *lightgbm*, *support vector machine*



ABSTRACT

Fraud in Ethereum transactions poses a significant challenge in the blockchain ecosystem, necessitating accurate detection methods. This study aims to compare the performance of five classification algorithms LightGBM, Extra Trees Classifier, Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), and CatBoost in detecting fraud on an Ethereum dataset. The methodology involves collecting transaction data, cleaning the dataset, performing feature selection, splitting the dataset (12,240 training samples, 1,969 test samples), balancing classes using SMOTE, applying data transformation (MinMaxScaler for SVM, PowerTransformer for GNB), training models with default parameters, and optimizing tuning params using GridSearchCV with 10-fold cross-validation. Results show that LightGBM and CatBoost achieved the highest accuracy of 99.34% with default parameters, improving to 99.39% after hyperparameter tuning on the test set. GNB improved significantly from 89.18% to 94.92%, while SVM slightly declined from 98.93% to 98.73%. LightGBM is recommended as the best model due to its stable and balanced performance.

Keyword: catboost, ethereum, fraud detection, gaussian naive bayes, lightgbm, support vector machine

