

INTISARI

Pencatatan akumulasi penyusutan Barang Milik Daerah (BMD) di Kabupaten Banyumas menghadapi risiko salah saji akibat heterogenitas nilai aset yang ekstrem serta kompleksitas kodefikasi hierarkis pada sistem SIMBADAMAS. Verifikasi audit manual berbasis sampling dinilai kurang memadai untuk menapisi populasi data secara objektif dan efisien, sehingga diperlukan model prediktif berbasis *machine learning* sebagai sistem deteksi dini anomali. Penelitian ini bertujuan mengevaluasi kinerja komparatif XGBoost dan CatBoost pada tiga jenjang pemodelan (*Baseline*, *Random Search*, dan optimasi *Bayesian/Optuna*) untuk mengestimasi rasio akumulasi penyusutan serta menerapkan model terbaik sebagai alat bantu deteksi anomali audit.

Metodologi mengadaptasi kerangka *Knowledge Discovery in Databases* (KDD) yang diintegrasikan dengan *equal-frequency quantile stratification*. Evaluasi dilakukan pada *hold-out test set* menggunakan metrik R^2 , MAE, dan WAPE. Dari 98.526 data mentah diperoleh 42.241 observasi bersih pada tiga stratum.

Hasil menunjukkan bahwa XGBoost-Optuna teroptimasi *Bayesian* konsisten mengungguli CatBoost-Optuna dalam meminimalkan kesalahan prediksi, di mana XGBoost menghasilkan WAPE lebih rendah pada Stratum 1 (0,5683% berbanding 0,5911%), Stratum 2 (0,7182% berbanding 0,8005%), dan Stratum 3 (1,4442% berbanding 1,6716%). Optimasi XGBoost juga sekitar 54 kali lebih efisien (89,23 detik berbanding 4.868,76 detik). Dengan ambang materialitas audit 10%, model mengisolasi 98 aset berisiko tinggi (*red flag*) dari 8.450 data uji. Temuan ini mendukung transisi dari audit manual menuju verifikasi komputasional yang lebih presisi bagi Badan Keuangan dan Aset Daerah (BKAD) Kabupaten Banyumas.

Kata kunci: Audit Komputasional, XGBoost, CatBoost, Optimasi *Bayesian*, Deteksi Anomali.

ABSTRACT

The recording of accumulated depreciation for Regional Property (Barang Milik Daerah/BMD) in Banyumas Regency faces the risk of misstatement due to extreme heterogeneity in asset values and the complexity of the hierarchical codification in the SIMBADAMAS system. Sampling-based manual audit verification is considered inadequate to screen the data population objectively and efficiently, necessitating a machine learning-based predictive model as an early anomaly detection system. This study aims to evaluate the comparative performance of XGBoost and CatBoost across three modeling stages (Baseline, Random Search, and Bayesian optimization/Optuna) to estimate the accumulated depreciation ratio and to implement the best model as an audit anomaly detection tool.

The methodology adapts the Knowledge Discovery in Databases (KDD) framework integrated with equal-frequency quantile stratification. Evaluation was conducted on a hold-out test set using R^2 , MAE, and WAPE metrics. From 98,526 raw data, 42,241 clean observations were obtained across three strata.

The results indicate that the Bayesian-optimized XGBoost-Optuna consistently outperforms CatBoost-Optuna in minimizing prediction errors, where XGBoost produced a lower WAPE in Stratum 1 (0.5683% versus 0.5911%), Stratum 2 (0.7182% versus 0.8005%), and Stratum 3 (1.4442% versus 1.6716%). XGBoost optimization was also approximately 54 times more efficient (89.23 seconds versus 4,868.76 seconds). Utilizing a 10% audit materiality threshold, the model successfully isolated 98 high-risk assets (red flags) out of 8,450 test instances. These findings support the transition from manual audits to more precise computational verification for the Regional Financial and Asset Agency (BKAD) of Banyumas Regency.

Keywords: Computational Audit, XGBoost, CatBoost, Bayesian Optimization, Anomaly Detection.