

## ABSTRAK

Dalam penelitian ini dua percobaan berbeda akan dilakukan untuk mengklasifikasikan dataset *loan default* (Gagal Bayar Pinjaman) menggunakan algoritma klasifikasi *Adaptive Boosting (Adaboost)* dan *Random Forest*. Kedua Percobaan menggunakan skenario pengujian *Cross Validation* dan *n-estimators* (jumlah pohon). Tujuannya adalah mengetahui hasil akurasi terbaik dan waktu yang diperlukan dari algoritma *Random Forest* dan *Adaboost* dalam mengklasifikasi risiko gagal bayar pinjaman pada data tidak seimbang dan data yang seimbang. Pada percobaan pertama, dilakukan pengujian pada kedua algoritma klasifikasi dengan menggunakan dataset yang belum seimbang, sedangkan percobaan kedua, dilakukan pengujian dengan menggunakan dataset yang telah diseimbangkan menggunakan teknik *Oversampling Adasyn*. Hasil analisis menunjukkan bahwa pada data tidak seimbang, kedua algoritma memiliki tingkat akurasi yang serupa sebesar 88%, ini menunjukkan kemampuan baik dalam mengatasi ketidakseimbangan data. Sedangkan, pada data seimbang, *Random Forest* menghasilkan akurasi tertinggi, mencapai 94.2% dengan skenario menggunakan *k-fold = 10* dan *n-estimators = 200* meskipun dengan waktu komputasi yang lebih lama.

Kata Kunci : *Loan Default, Adaboost, Random Forest, Gagal Pinjaman, Unbalance Data, Balance Data.*

## ABSTRACT

*In this study, two different experiments will be conducted to classify a loan default dataset using the Adaptive Boosting (Adaboost) and Random Forest classification algorithms. Both experiments utilize Cross Validation testing scenarios and n-estimators (number of trees). The objective is to determine the best accuracy and the time required by the Random Forest and Adaboost algorithms in classifying loan default risk on both imbalanced and balanced data. In the first experiment, testing is conducted on both classification algorithms using the imbalanced dataset, while in the second experiment, testing is done using the balanced dataset achieved through Adasyn Oversampling technique. The analysis results indicate that on imbalanced data, both algorithms achieve similar accuracy rates of 88%, demonstrating their effectiveness in handling data imbalance. On balanced data, Random Forest achieves the highest accuracy, reaching 94.2% with a k-fold scenario of 10 and n-estimators of 200, albeit with longer computation time.*

*Keywords : Loan Default, Adaboost, Random Forest, Loan Failure, Unbalanced Data, Balanced Data.*

