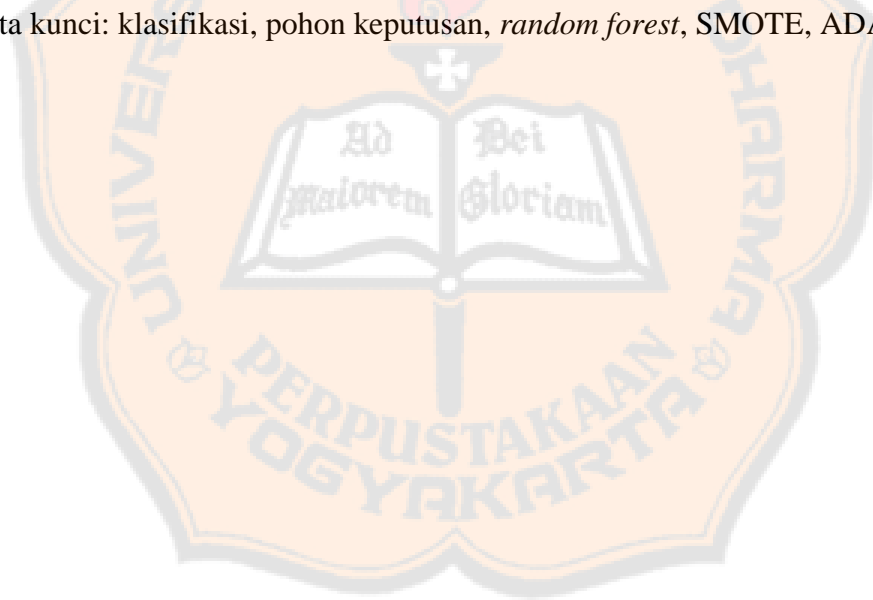


## ABSTRAK

Model klasifikasi dengan *random forest* dapat digunakan untuk memprediksi tanggapan pelanggan terhadap promosi bank tentang deposito berjangka. Pemodelan *random forest* menggunakan data promosi dari salah satu bank di Portugal. Data tersebut memiliki dua kelas data dan jumlah data dari kedua kelas data tidak seimbang. Data akan diseimbangkan dengan teknik *oversampling* SMOTE dan ADASYN. Tujuan penelitian ini adalah membandingkan akurasi, presisi, *recall*, dan *f1-score* dari model *random forest* yang menerapkan penyeimbangan data SMOTE dan ADASYN. Penelitian ini juga membandingkannya dengan model *random forest* tanpa penyeimbangan data. Pembagian data menjadi data latih dan data uji dilakukan dengan validasi silang *k*-lipatan. Akurasi akan dihitung menggunakan *confusion matrix*. Penelitian ini juga akan melakukan beberapa percobaan dengan jumlah lipatan dan jumlah pohon yang berbeda. Model klasifikasi ini berguna untuk meningkatkan efisiensi dan efektivitas bank dalam melakukan promosi deposito berjangka kepada pelanggan bank. Berdasarkan hasil penelitian, nilai akurasi, presisi, *recall*, dan *f1-score* tertinggi didapatkan dari penyeimbangan data dengan metode SMOTE, 10 lipatan, dan 500 pohon. Parameter tersebut menghasilkan nilai akurasi sebesar 93,57%, nilai presisi sebesar 91,35%, nilai *recall* sebesar 96,25%, dan nilai *f1-score* sebesar 93,73%.

Kata kunci: klasifikasi, pohon keputusan, *random forest*, SMOTE, ADASYN



## ABSTRACT

A classification model with random forest can be used to predict customer responses to bank promotions on time deposit. Random forest modeling using promotion data from a bank in Portugal. The data has two classes and the number of data from both classes is imbalanced. The data will be balanced using the SMOTE and ADASYN oversampling techniques. The aim of this research is to compare the accuracy, precision, recall, and f1-score of the random forest model that apply SMOTE and ADASYN data balancing. This research also compares it with the random forest model without data balancing. The data splitting into training data and test data is performed using k-fold cross-validation. Accuracy will be calculated using the confusion matrix. This research will also conduct several experiments through different number of folds and number of trees. This classification model is useful for improving the efficiency and effectiveness of banks in promoting time deposit to bank customers. Based on the research results, the highest accuracy, precision, recall, and f1-score values are obtained from data balancing with the SMOTE method, 10 folds, and 500 trees. These parameters result in an accuracy of 93.57%, precision of 91.35%, recall of 96.25%, and an f1-score of 93.73%.

Keywords: classification, decision tree, random forest, SMOTE, ADASYN

