

## ABSTRAK

Karena banyaknya nasabah melakukan pinjaman dan risiko gagal membayar pinjaman dapat menyebabkan kerugian maka pihak yang memberikan pinjaman seperti bank perlu memprediksi kriteria nasabah yang gagal dalam membayar pinjaman, maka penelitian ini bertujuan untuk menerapkan algoritma *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF) dalam membangun model klasifikasi guna memprediksi nasabah yang gagal membayar pinjaman bank, serta membandingkan pengaruh metode balancing data *Random Oversampling* (ROS) dan *Near Miss* terhadap akurasi klasifikasi. Dataset yang digunakan terdiri dari data nasabah dengan ketidakseimbangan kelas antara nasabah yang gagal bayar (Default) dan yang tidak gagal bayar (Not Default). Metode *Random Oversampling* menyeimbangkan kelas dengan menambah jumlah sampel minoritas, sementara metode *Near Miss* mengurangi jumlah sampel mayoritas yang dekat dengan kelas minoritas. Hasil penelitian menunjukkan bahwa model SVM yang dilatih tanpa *balancing* memiliki akurasi tertinggi 88.49%, namun dengan *recall* yang rendah untuk kelas minoritas 10%. Setelah menggunakan ROS, akurasi model sedikit menurun 83.52%, tetapi *recall* untuk kelas minoritas meningkat signifikan hingga 94%. Sebaliknya, dengan *Near Miss*, akurasi model menurun lebih jauh 65.29%, namun menghasilkan presisi dan *recall* yang lebih seimbang antara kedua kelas. Meskipun *Random Oversampling* sedikit menurunkan akurasi keseluruhan, metode ini berhasil meningkatkan kinerja model dalam memprediksi kelas minoritas secara signifikan, sementara *Near Miss* memberikan keseimbangan yang lebih baik antara presisi dan *recall*, meskipun dengan penurunan akurasi keseluruhan yang lebih besar.

Kata kunci : *Support Vector Machine*, Kegagalan membayar pinjaman, *Random Oversampling*, *Near Miss*, *Radial Basis Function*

## ABSTRACT

Because many customers make loans and the risk of defaulting on loans can cause losses, lenders such as banks need to predict the criteria for customers who fail to pay loans, so this study aims to apply the Support Vector Machine (SVM) algorithm with the Radial Basis Function (RBF) kernel in building a classification model to predict customers who default on bank loans, and compare the effect of Random Oversampling (ROS) and Near Miss data balancing methods on classification accuracy. The dataset used consists of customer data with class imbalance between customers who default and those who do not default. The Random Oversampling method balances the classes by increasing the number of minority samples, while the Near Miss method reduces the number of majority samples close to the minority class. The results show that the SVM model trained without balancing has the highest accuracy of 88.49%, but with low recall for the 10% minority class. After using ROS, the model accuracy slightly decreased to 83.52%, but the recall for the minority class significantly increased to 94%. In contrast, with Near Miss, the model accuracy decreased further by 65.29%, but resulted in more balanced precision and recall between the two classes. Although Random Oversampling slightly decreased the overall accuracy, it managed to significantly improve the model's performance in predicting the minority class, while Near Miss provided a better balance between precision and recall, but with a considerable decrease in overall accuracy.

Keywords: Support Vector Machine, Loan Default, Random Oversampling, Near Miss, Radial Basis Function