

ANALISIS KELAYAKAN KREDIT MENGGUNAKAN CLASSIFICATION TREE DENGAN TEKNIK RANDOM OVERSAMPLING

INTISARI

Kredit adalah kegiatan pemberian uang atau tagihan berdasarkan suatu perjanjian antara bank dan pihak lainnya. Kredit yang diajukan oleh debitur dapat berisiko sehingga menyebabkan kredit macet. Upaya untuk mengurangi kredit macet adalah dengan melakukan analisis kredit terhadap debitur sebelum menyetujui pinjaman yang diajukan debitur untuk menentukan apakah permohonan kredit akan disetujui. Penelitian ini menggunakan data sekunder, yaitu data status debitur salah satu bank di Kalimantan Barat. Data penelitian berjumlah 800 sampel data terdiri dari variabel kolektibilitas sebagai variabel target dan 10 variabel independen yaitu *limit*, *rate*, *tenor*, *total angsuran*, *usia*, *gaji*, *premi* dan *admin*, *instansi*, *jenis kredit* dan *jenis kebutuhan*. Metode yang digunakan adalah metode *Classification Tree* dengan teknik *Random Oversampling* untuk mengatasi *imbalanced data*. Tujuan penelitian ini yaitu untuk menangani *imbalanced data* menggunakan *Random Oversampling* dan menentukan variabel yang berpengaruh dalam klasifikasi menggunakan *Classification Tree*. Pengklasifikasian diawali dengan *preprocessing data*, lalu data dibagi menjadi data latih dan uji dengan proporsi 70:30, 80:20 dan 90:10 untuk masing-masing perlakuan tanpa *Random Oversampling* dan dengan *Random Oversampling*. Selanjutnya, model klasifikasi dibentuk menggunakan data latih dan validasi model klasifikasi menggunakan data uji. Setelah itu, dilakukan evaluasi keseluruhan model untuk mengetahui model terbaik yang digunakan dalam proses klasifikasi. Berdasarkan hasil penelitian, model terbaik adalah model dari *Classification Tree* dengan teknik *Random Oversampling* proporsi 70:30, dengan nilai *specificity* sebesar 75,00%, nilai *recall* sebesar 89,66% dan nilai akurasi sebesar 89,17%. Dapat diartikan, model dapat digunakan untuk mengklasifikasikan data lancar dan non lancar debitur dengan variabel paling berpengaruh dalam pengklasifikasian status debitur adalah variabel total angsuran.

Kata Kunci: *Imbalanced Data, Classification Tree, Random Oversampling*

CREDITWORTHINESS ANALYSIS USING CLASSIFICATION TREE WITH RANDOM OVERSAMPLING TECHNIQUE

ABSTRACT

Credit is the activity of providing money or bills based on an agreement between a bank and another party. Loans submitted by debtors can be risky, causing bad credit. An effort to reduce bad debts is to conduct a credit analysis of the debtor before approving the loan submitted by the debtor to determine whether the credit application will be approved. This study uses secondary data, namely data on the status of debtors from one bank in West Kalimantan. The research data amounted to 800 data samples consisting of collectibility variables as target variables and 10 independent variables, namely limit, rate, tenor, total installments, age, salary, premium and admin, agency, type of credit and type of need. The method used is the Classification Tree method with Random Oversampling technique to overcome imbalanced data. The purpose of this research is to handle imbalanced data using Random Oversampling and determine variables that influence classification using Classification Tree. The classification begins with data preprocessing, then the data is divided into training and test data with a proportion of 70:30, 80:20 and 90:10 for each treatment without Random Oversampling and with Random Oversampling. Next, a classification model is formed using the training data and validation of the classification model using the test data. After that, an overall evaluation of the model is carried out to determine the best model used in the classification process. Based on the research results, the best model is a model from Classification Tree with Random Oversampling technique in proportion 70:30, with a specificity value of 75.00%, a recall value of 89.66% and an accuracy value of 89.17%. It can be interpreted that the model can be used to classify current and non-current debtor data with the most influential variable in classifying debtor status is the total installment variable.

Keywords: *Imbalanced Data, Classification Tree, Random Oversampling*