

INTISARI

CREDIT SCORING MENGGUNAKAN REGRESI LOGISTIK LASSO

Oleh:

Amalia Afifah
11/319999/PA/14313

Credit scoring merupakan suatu metode berbasis analisis statistika yang digunakan untuk mengukur besaran resiko kredit. Metode klasifikasi yang paling populer digunakan untuk *credit scoring* adalah regresi logistik. Regresi logistik digunakan untuk memprediksi variabel respon yang biner dengan satu set variabel penjelas (prediktor). Regresi logistik mempunyai keterbatasan yaitu jika terdapat multikolinieritas (korelasi yang tinggi antar variabel bebas) membuat model regresi yang didapat menjadi tidak lagi efisien karena nilai standar error koefisien regresi menjadi sangat besar (*overestimate*) atau dengan kata lain mengurangi akurasi dari estimasi. Oleh karena itu, diusulkan metode Least Absolute Shrinkage and Selection Operator (LASSO) untuk mengatasi hal tersebut. LASSO akan menyusutkan koefisien (parameter β) yang berkorelasi, menjadi nol atau mendekati nol. Sehingga menghasilkan model akhir yang lebih representatif. Pada akhirnya, performa model regresi logistik LASSO dibandingkan dengan model regresi logistik. Dengan melihat nilai presentase ketepatan model, model regresi logistik LASSO dianggap lebih baik daripada model regresi logistik karena memiliki nilai presentase ketepatan model yang lebih besar.

Kata kunci : *credit scoring*, multikolinearitas, regresi logistik, metode LASSO

ABSTRACT

CREDIT SCORING USING LASSO LOGISTIC REGRESSION

By:

Amalia Afifah
11/319999/PA/14313

Credit scoring is a method based on a statistical analysis is used to measure the amount of credit risk. The most popular method of classification used for credit scoring is logistic regression. Logistic regression was used to predict the response variable is a binary with a set of explanatory variables (predictors). Logistic regression have limitations, if there is multicollinearity (a high correlation between independent variables) obtained makes regression model becomes more not efficient because the value of the standard error of regression coefficient becomes very large (overestimate), or in other words reduce the accuracy of the estimate. Therefore, Least Absolute Shrinkage and Selection Operator (LASSO) method proposed to overcome it. LASSO would shrink coefficient (parameter β), which correlated, to zero or close to zero. Resulting in a final model more representative. In the end, the performance of the logistic regression model LASSO compared with logistic regression model. By looking at the value of the percentage of accuracy from the model, the logistic regression model LASSO considered better than the logistic regression model because it has larger percentage of of models accuracy.

Keywords: credit scoring, multikolinearitas, logistic regression, LASSO methods