

Implementasi Metode Binary Logistic Regression dan Gradient Boosting dalam Pemodelan Credit Scoring = Implementation of Binary Logistic Regression and Gradient Boosting Methods for Credit Scoring Model

Assyifa Ulhusna, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20527330&lokasi=lokal>

Abstrak

Credit scoring adalah sebuah sistem yang digunakan kreditor seperti bank dan perusahaan asuransi untuk menentukan apakah pemohon kredit termasuk dalam grup good credit yakni grup yang kemungkinan besar akan membayar utangnya tepat waktu atau bad credit yang merupakan grup dengan kemungkinan besar tidak membayar utangnya tepat waktu. Salah satu metode yang paling sering digunakan dalam pembuatan model credit scoring adalah binary logistic regression. Namun, seiring dengan kemajuan komputasi, banyak metode lain yang berkembang saat ini untuk dipakai dalam pembuatan model credit scoring yakni, metode gradient boosting. Pada skripsi ini dilakukan implementasi metode binary logistic regression dan gradient boosting dalam pemodelan credit scoring. Hasil yang didapatkan dengan menggunakan data 537.667 debitur dengan rincian 535.705 good credits dan 1.962 bad credits adalah pada data train penggunaan gradient boosting memberikan nilai tingkat akurasi 79,65%, uji KS 0,5389 dan AUROC/AUC 0,8393. Sementara pada data test penggunaan gradient boosting memberikan nilai tingkat akurasi 79,92%, uji KS 0,5345 dan AUROC/AUC 0,8313. Nilai-nilai tersebut lebih tinggi dibandingkan dengan penggunaan binary logistic regression baik pada data train maupun data test. Berdasarkan nilai uji AUC, metode gradient boosting tergolong klasifikasi yang baik, sedangkan metode binary logistic regression > tergolong klasifikasi yang cukup. Hasil simulasi ini menunjukkan untuk data yang digunakan, metode gradient boosting memberikan hasil yang lebih baik dari sisi akurasi, uji KS, dan AUROC/AUC daripada binary logistic regression. Dengan kata lain, metode gradient boosting dapat meningkatkan discriminant power, yakni kemampuan untuk membandingkan target yang lebih baik dibandingkan dengan metode binary logistic regression.

.....Credit scoring is a system used by creditors such as banks and insurance companies to determine whether credit applicants are included in the good credit group, namely the group that is most likely to pay its debts on time or the bad credit group which is the group that is most likely to not pay its debts on time. One of the most frequently used methods in making credit scoring models is binary logistic regression. However, along with the progress of computation, many other methods are currently being developed to be used in making credit scoring models, namely, the gradient boosting method. In this thesis, we will compare the binary logistic regression and gradient boosting methods in credit scoring model. The results obtained using data from 537,667 debtors with details of 535,705 good credits and 1,962 bad credits are the train data using gradient boosting gives an accuracy rate of 79.65%, KS test 0.5389 and AUROC/AUC 0.8393. Meanwhile, the test data using gradient boosting gives an accuracy rate of 79.92%, KS test 0.5345, and AUROC/AUC 0.8313. These values are higher than the use of binary logistic regression in both the train and test data. Based on the AUC test value, the gradient boosting method is a good classifier, while the binary logistic regression method is an acceptable classifier. The results of this simulation show that for the data used, the gradient boosting method gives better results in terms of accuracy, KS test, and AUROC/AUC than binary logistic regression. In other words, the gradient boosting method can increase discriminant

power or the ability to compare targets better than the binary logistic regression method.