

Analisis akurasi model XGBoost untuk klasifikasi multikelas: studi kasus prediksi tingkat klaim risiko pemohon pada asuransi jiwa =
Analyzing accuracy of XGBoost model for multiclass classification: a case study of the applicant level claim risk prediction for life insurance

Widya Fajar Mustika, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20493284&lokasi=lokal>

Abstrak

Penilaian tingkat klaim risiko pemohon asuransi merupakan bagian penting dalam asuransi jiwa, sehingga perlu untuk diklasifikasikan. Penentuan tingkat klaim risiko pada asuransi jiwa didasarkan pada data historis pemohon. Pengajuan untuk menjadi anggota suatu asuransi jiwa membutuhkan waktu yang tidak singkat. Namun pengaplikasian suatu model machine learning dapat membantu mengklasifikasikan calon pemohon asuransi berdasarkan tingkat risiko dengan cepat. Salah satu model machine learning yaitu Extreme Gradient Boosting (XGBoost) yang merupakan suatu model berbasis decision tree. Model ini digunakan untuk memprediksi risiko pada asuransi jiwa. Adanya missing values pada data yang digunakan diatasi dengan beberapa strategi pada proses prapengolahan data untuk meningkatkan nilai akurasi model XGBoost. Hasil penelitian ini diperoleh bahwa akurasi model XGBoost sebesar 0,60730 dengan satuan kappa yang menunjukkan bahwa model XGBoost sangat baik dan dapat diterapkan pada masalah prediksi tingkat klaim risiko pemohon asuransi jiwa. Jika dibandingkan dengan model decision tree, random forest dan Bayesian ridge, kinerja model XGoost masih tetap unggul dalam memproses missing values pada data yang digunakan.

<hr>

Risk level assessment for insurance applicants is an important part of life insurance, so it needs to be classified. Determination of the level of risk claims on life insurance is based on the applicants historical data. Submission to become a member of a life insurance requires a short time. But the application of a machine learning model can help classify prospective insurance applicants based on the level of risk quickly. One machine learning model is Extreme Gradient Boosting (XGBoost) which is a decision tree based model. This model is used to predict risk in life insurance. The missing values in the data used are overcome by several strategies in the data processing process to increase the accuracy value of the XGBoost model. The results of this study show that the accuracy of the XGBoost model is 0.60730 with kappa units which indicates that the XGBoost model is very good and can be applied to the problem of predicting the level of risk claims for life insurance applicants. When compared to the decision tree, random forest and Bayesian ridge models, the performance of the XGoost model still excels in processing missing values in the data used.