Analisis Algoritma Image Generator for Tabular Data untuk Prediksi Klaim Asuransi Kompensasi Pekerja Menggunakan Convolutional Neural Network = Analysis of Image Generator for Tabular Data Algorithm for Workers's Compensation Insurance Claim Prediction Using Convolutional Neural Network

Reri Nandar Munazat, author

Abstrak

Seiring meningkatnya tren kecelakaan kerja selama periode 2007-2017 serta berjalannya kembali kegiatan usaha secara normal pascapandemi COVID-19, lini usaha asuransi kompensasi pekerja menjadi sangat potensial untuk dikembangkan. Sebagai komponen penting dalam model bisnis asuransi, severitas klaim perlu diprediksi seakurat mungkin karena berpengaruh terhadap penetapan tarif premi bagi tertanggung serta bermanfaat dalam mekanisme pengamatan klaim selama proses penyelesaian klaim. Proses prediksi ini dikategorikan sebagai masalah regresi yang biasanya ditangani oleh model-model pembelajaran mesin untuk data tabular. Namun dalam perkembangan studi pembelajaran mesin, terdapat upaya untuk memanfaatkan model Convolutional Neural Network (CNN) untuk melakukan prediksi terhadap data tabular dengan cara mentransformasikan data tersebut ke dalam representasi gambarnya, salah satunya melalui algoritma Image Generator for Tabular Data (IGTD). Penelitian ini bertujuan untuk menguji akurasi model CNN berbasis algoritma IGTD dalam memprediksi klaim asuransi kompensasi pekerja serta membandingkan performa model tersebut dengan model Multi-Layer Perceptron, Random Forest, serta eXtreme Gradient Boosting. Hasil simulasi dengan metode repeated holdout sebanyak lima iterasi menunjukkan bahwa model CNN dapat memprediksi klaim dengan baik meskipun secara umum belum mampu menyaingi model-model non-CNN secara signifikan.

.....Along with the increasing trend of work accidents during 2007-2017 period as well as the resumption of business activities normally after the COVID-19 pandemic, the workers’ compensation insurance business line has great potential to be developed. As an important component in the insurance business model, the claim severity needs to be predicted as accurate as possible because it affects the determination of premium rates for the insured and is useful in the claim watching mechanism during the claim settlement process. This prediction process is categorized as a regression problem which is usually handled by machine learning models for tabular data. However, in the development of machine learning studies, there are emerging efforts to utilize the Convolutional Neural Network (CNN) model to predict tabular data by transforming the data into its image representation, one of which is through Image Generator for Tabular Data (IGTD) algorithm. This study aims to test the accuracy of the CNN model based on the IGTD algorithm in predicting workers’ compensation insurance claims and to compare the model performance with the Multi-Layer Perceptron, Random Forest, and eXtreme Gradient Boosting models. The simulation result using the repeated holdout method for five iterations shows that the CNN model can well predict the claims, although in general, it has not been able to significantly compete with non-CNN models.