

## INTISARI

### **Pengaplikasian *Tree Based Machine Learning* pada Prediksi Frekuensi Klaim Asuransi Kendaraan Bermotor Komersial**

Oleh

Emil Valentino

19/445627/PA/19451

Asuransi kendaraan di Indonesia berkembang cukup pesat. Tidak bisa dipungkiri, jumlah kendaraan bermotor di Indonesia sangat besar, sehingga banyak orang memilih untuk mentransfer risiko kepada perusahaan asuransi. Hal ini tidak terkecuali bagi perusahaan-perusahaan yang memiliki kendaraan operasional. Kendaraan operasional biasanya akan diasuransikan secara kolektif. Hal inilah yang mendorong terbentuk asuransi kendaraan bermotor komersial. Penentuan premi (*pricing*) pada asuransi kendaraan bermotor komersial di Indonesia masih menggunakan metode tradisional. Penggunaan metode tradisional seperti dengan regresi tradisional, GLM memiliki kelemahan, di antaranya yaitu ketidakmampuan GLM menangkap pola-pola non linear. Pada penelitian ini, akan dihadirkan solusi untuk mengatasi ketidakmampuan GLM tersebut dengan menggunakan *tree based machine learning*. *Tree based machine learning* yang digunakan untuk memprediksi frekuensi klaim, yaitu *Random forest* dan *Extreme Gradient Boosting*. *Tree based machine learning* mampu mengatasi ketidakmampuan GLM dalam menangkap pola non linear sekaligus dapat beradaptasi dengan data, sehingga tidak perlu memenuhi uji asumsi apapun. Diperoleh kesimpulan bahwa penggunaan *tree based machine learning* masih lebih baik untuk memprediksi frekuensi klaim karena menghasilkan model dengan RMSE yang lebih rendah.

Kata kunci : frekuensi klaim, *machine learning*, GLM, *Random Forest*, *Extreme Gradient Boosting*

## **ABSTRACT**

### **The application of Tree-Based Machine Learning in Predicting Commercial Motor Vehicle Insurance Claim Frequency**

By

Emil Valentino

19/445627/PA/19451

Motor vehicle insurance in Indonesia has been growing rapidly. It is undeniable that the number of motor vehicles in Indonesia is very large, prompting many people to transfer the risk to insurance companies. This is also true for companies that have operational vehicles. Operational vehicles are usually collectively insured, which has led to the formation of commercial motor vehicle insurance. The determination of premiums in commercial motor vehicle insurance in Indonesia still relies on traditional methods. The use of traditional methods such as traditional regression and GLM has limitations, including the inability of GLM to capture non-linear patterns. This research presents a solution to overcome the limitations of GLM by using tree-based machine learning. The tree-based machine learning methods used to predict claim frequency are Random Forest and Extreme Gradient Boosting. Tree-based machine learning can overcome the inability of GLM to capture non-linear patterns and can adapt to data without requiring any assumption tests. It is concluded that the use of tree-based machine learning is better for predicting claim frequency as it produces models with lower RMSE.

**Keywords:** claim frequency, machine learning, GLM, Random Forest, Extreme Gradient Boosting