

## INTISARI

### Penentuan Premi Murni Asuransi Kendaraan Bermotor berdasarkan Jarak Tempuh (Pay-As-You-Drive Insurance) dengan Tree-Based Machine Learning

Oleh

Dhestar Bagus Wirawan  
18/427694/PA/18654

Sistem penentuan premi asuransi kendaraan bermotor di Indonesia saat ini masih belum cukup adil bagi nasabah, terutama terkait jarak tempuh. Terdapat subsidi silang antara nasabah yang jarang menggunakan kendaraannya dengan yang sering menggunakan. Penelitian yang dilakukan oleh Ferreira dan Minikel (2012) menargetkan permasalahan ini dan menggunakan *Generalized Linear Models* (GLM) untuk menghitung premi murni asuransi kendaraan bermotor berdasarkan jarak tempuh atau biasa disebut *pay-as-you-drive insurance*. Meskipun GLM sering digunakan dalam pemodelan di dunia asuransi, ketergantungan metode GLM terhadap asumsi-asumsi yang harus terpenuhi dan ketidakmampuan dalam menangkap pola non-linear menjadi kelemahan utamanya. Penelitian ini akan memanfaatkan *tree-based machine learning* yaitu *Random Forest* dan *Gradient Boosting Machine* dalam perhitungan premi murni *pay-as-you-drive insurance* dan membuka *black-box machine learning* sehingga memiliki kemampuan interpretasi yang sama dengan GLM. Diperoleh kesimpulan bahwa algoritma *Gradient Boosting Machine* mampu menghasilkan model yang memiliki RMSE terendah, baik itu untuk pemodelan frekuensi klaim maupun *severity* klaim. Selain itu, diperoleh kesimpulan juga bahwa asuransi PAYD lebih baik daripada asuransi kendaraan bermotor tradisional.

Kata-kata kunci : Gradient Boosting Machine, Generalized Linear Models, Pay-as-you-drive Insurance, Random Forest.

## ABSTRACT

### **Determining Auto Insurance Pure Premium based on Mileage (Pay-As-You-Drive Insurance using Tree-Based Machine Learning)**

By

Dhestar Bagus Wirawan  
18/427694/PA/18654

The current system for determining auto-insurance premium in Indonesia is still not fair enough for the customer, especially related to mileage. There are cross-subsidized between the low-mileage customers and high-mileage customers. In 2012, Ferreira and Minikel target this problems and used Generalized Linear Models (GLM) to calculate the pure premium based on mileage or known as pay-as-you-drive insurance. Although GLM is often used when modelling in insurance, the dependence of GLM on assumptions and its inability to capture non-linear pattern is the main weaknesses of GLM. This research would use tree-based machine learning, e.g. Random Forest and Gradient Boosting Machine, in calculating the pure premium of pay-as-you-drive insurance and opening the black-box of machine learning so that it has the same interpretation capabilities as the GLM models. It is concluded that the Gradient Boosting Machine algorithm is able to produce a model that has the lowest RMSE, both for modeling claim frequency and claim severity. In addition, it is also concluded that PAYD insurance is better than traditional motor vehicle insurance.

Keywords : Gradient Boosting Machine, Generalized Linear Models, Pay-as-you-drive Insurance, Random Forest.