

Automobile insurance fraud detection using supervised learning classifiers = Deteksi kecurangan pada asuransi kendaraan bermotor dengan supervised learning classifiers.

Iffa Maula Nur Prasasti, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20504728&lokasi=lokal>

Abstrak

Asuransi mobil adalah produk asuransi yang banyak digunakan di Indonesia. Namun, asuransi mobil memiliki potensi untuk kecurangan klaim yang menyebabkan kerugian bagi perusahaan dan pemegang polis. Penelitian ini bertujuan untuk merancang model prediksi deteksi kecurangan asuransi mobil di Indonesia menggunakan pendekatan *machine learning*. *Supervised classifiers* adalah salah satu teknik *machine learning* yang memiliki kemampuan untuk memprediksi kasus-kasus anomali. *Supervised classifiers* yang digunakan pada penelitian ini adalah *Multilayer Perceptron* (MLP), *Decision Tree C4.5*, dan *Random Forest* (RF). Penelitian ini menggunakan data *real-world* pada perusahaan asuransi mobil di Indonesia. Dataset memiliki distribusi tidak seimbang yang sangat tinggi antara data pemegang polis yang melakukan kecurangan dan pemegang polis yang sah. Penelitian ini menangani masalah dataset yang tidak seimbang dengan menggunakan *Synthetic Minority Oversampling Technique* (SMOTE) dan metode *undersampling*. Kinerja model dievaluasi melalui *confusion matrix*, Kurva ROC, dan parameter seperti sensitivitas. Penelitian ini menemukan bahwa *Random Forest* memberikan hasil terbaik dibandingkan dengan MLP dan *Decision Tree C4.5*.

.....Automobile insurance is widely used insurance product in Indonesia. However, automobile insurance has the potential for fraudulent claim that leads to several consequences for the company and policyholder. This research aims to design a prediction model of automobile insurance fraud detection in Indonesia using a machine learning approach. Supervised classifiers is one of machine learning techniques that has the ability to predict anomaly cases. The proposed supervised classifiers are Multilayer Perceptron (MLP), Decision Tree C4.5, and Random Forest(RF). This research used real-world data on an automobile insurance company in Indonesia. The dataset has a high imbalanced distribution between the data of policyholders who commit fraud and legitimate. This study handles the imbalanced dataset problem by using the Synthetic Minority Oversampling Technique (SMOTE) and undersampling methods. The performance of models is evaluated through the confusion matrix, ROC Curve, and parameters such as sensitivity. This research found that Random Forest outperformed the results comparing to other classifiers.