

INTISARI

Pengaplikasian Regresi Logistik dan *Decision Tree* Dalam Mendeteksi *Fraud* Asuransi Kesehatan

Oleh

Ignatius Baskoro Wijoyo

19/442452/PA/19201

Ada banyak algoritma *machine learning* untuk mendeteksi *fraud* (penipuan) dalam asuransi kesehatan, dua diantaranya adalah regresi logistik dan *Decision Tree*. Pada dua algoritma *machine learning* ini data pemegang polis akan dibagi menjadi dua bagian dimana 80% data digunakan sebagai pembuat model dan 20% data digunakan untuk *testing* model. Model yang terbentuk kemudian akan digunakan untuk mendeteksi profil para pemegang polis, terutama yang baru, apakah para pemegang polis memiliki kecenderungan tinggi untuk melakukan *fraud* atau tidak. Dengan kata lain model yang terbentuk ini bisa digunakan sebagai *Early Warning System* bagi perusahaan asuransi sehingga perusahaan bisa menjadi lebih waspada terhadap upaya-upaya *fraud* dari para pemegang polis. Meskipun regresi logistik lebih baik dibandingkan *Decision Tree* dalam mendeteksi kemungkinan *fraud* akan tetapi penggabungan kedua algoritma tersebut lebih disarankan agar kemampuan mendeteksi upaya-upaya *fraud* dari para pemegang polis menjadi lebih lengkap.

ABSTRACT

Application of Logistic Regression and Decision Tree in Detecting Health Insurance Fraud

By

Ignatius Baskoro Wijoyo

19/442452/PA/19201

There are many machine learning algorithms to detect fraud in health insurance, two of which are Logistic Regression and Decision Tree. In these two machine learning algorithms, the policyholder data will be divided into two parts where 80% of the data is used for model building and 20% of the data is used for model testing. The model will then be used to detect the profiles of policyholders, especially new ones, whether the policyholders have a high tendency to commit fraud or not. In other words, the model can be used as an Early Warning System for insurance companies so that companies can be more aware of fraud attempts from policyholders. Although Logistic Regression is better than Decision Tree in detecting the possibility of fraud, combining the two algorithms is recommended so that the ability to detect fraud attempts from policyholders becomes more complete.