

INTISARI

ANALISIS DETEKSI KECURANGAN PADA KLAIM ASURANSI KENDARAAN MENGGUNAKAN ALGORITMA *MACHINE LEARNING*

Oleh

Almarizti Hidayani

19/442447/PA/19196

Seiring dengan meningkatnya jumlah kendaraan bermotor di Indonesia, risiko terjadinya kecurangan (*fraud*) dalam klaim asuransi juga semakin tinggi. Kecurangan ini tidak hanya berdampak negatif pada keuangan perusahaan asuransi, tetapi juga pada kredibilitas industri tersebut. Oleh karena itu, diperlukan sistem deteksi kecurangan yang efektif. Penelitian ini bertujuan untuk menganalisis deteksi kecurangan pada klaim asuransi kendaraan menggunakan algoritma machine learning, yaitu *Naive Bayes*, *K-Nearest Neighbors* (KNN), *Decision Tree*, dan *Logistic Regression*. Data yang digunakan adalah dataset klaim asuransi kendaraan bermotor roda empat dari Kaggle, yang menunjukkan ketidakseimbangan kelas antara klaim *fraud* dan *non-fraud*. Untuk mengatasi hal ini, teknik oversampling SMOTE diterapkan. Model dievaluasi menggunakan metrik seperti akurasi, presisi, *recall*, *F1-score*, dan ROC AUC. Dari hasil penelitian dapat ditunjukkan bahwa algoritma *Decision Tree* memiliki performa yang terbaik setelah oversampling, dengan *recall* dan *F1-score* tertinggi dibandingkan algoritma lainnya. Studi ini memberikan kontribusi dalam pengembangan model deteksi kecurangan yang dapat membantu perusahaan asuransi mengurangi risiko kerugian finansial.

ABSTRACT

ANALYSIS OF FRAUD DETECTION IN VEHICLE INSURANCE CLAIM USING MACHINE LEARNING ALGORITHMS

By

Almarizti Hidayani

19/442447/PA/19196

As the number of motorized vehicles in Indonesia continues to rise, the risk of fraud in insurance claims also increases. This fraud negatively impacts not only the financial stability of insurance companies but also the credibility of the industry. Therefore, an effective fraud detection system is essential. This study aims to analyze fraud detection in vehicle insurance claims using machine learning algorithms, namely Naive Bayes, K-Nearest Neighbors (KNN), Decision Tree, and Logistic Regression. The data used is a vehicle insurance claim dataset from Kaggle, highlighting class imbalance between fraud and non-fraud claims. To address this issue, the SMOTE oversampling technique was applied. Models were evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC AUC. The results indicate that the Decision Tree algorithm achieved the best performance after oversampling, with the highest recall and F1-score compared to other algorithms. This study contributes to the development of fraud detection models that can help insurance companies mitigate financial risks.