



INTISARI

Penelitian ini menyelidiki apakah pendekatan *Machine Learning* (ML) dengan menggunakan model *Random Forest* (RF), *Support Vector Machine* (SVM), *Neural Network* (NN), dan *Decision Tree* (DT) memiliki pengaruh dalam mendeteksi potensi kecurangan keuangan daerah. Peneliti menggunakan tiga belas fitur (x) dan dua target (y) yang bersumber dari data keuangan daerah dan non-keuangan daerah. Data ini diperoleh melalui *website* resmi pemerintah daerah dan analisis sentimen berita daerah dari 514 daerah untuk tahun 2022-2023. Hasil menunjukkan bahwa model ML (RF, NN, dan DT) memiliki pengaruh signifikan dalam mendeteksi potensi kecurangan, kecuali SVM. Nilai negatif dan tidak signifikan SVM bukan menandakan buruknya model, melainkan karena ketidakcocokan model dengan jenis penelitian. Melalui ML, peneliti menemukan adanya peningkatan potensi kecurangan daerah yang terjadi pada tahun 2023. Sebanyak 143 kabupaten/kota yang sebelumnya (2022) di kategorikan “Tidak Berpotensi”, dan “Agak Berpotensi”, berubah status menjadi daerah yang “Sangat Berpotensi” dalam melakukan kecurangan. Akhirnya, studi ini berkontribusi pada literatur dengan memberikan wawasan tentang ketepatan penggunaan model ML (RF, NN, dan DT) dalam mendeteksi potensi kecurangan keuangan disektor pemerintah daerah Indonesia berbasis big data regional.

Kata Kunci: Kecurangan daerah, machine learning, Indonesia



ABSTRACT

This study investigates whether Machine Learning (ML) approaches using Random Forest (RF), Support Vector Machine (SVM), Neural Network (NN), and Decision Tree (DT) models are effective in detecting potential regional financial fraud. Researchers used thirteen features (x) and two targets (y) sourced from regional financial and non-financial data. This data was obtained through official regional government websites and sentiment analysis of regional news from 514 regions for 2022-2023. The results show that the ML models (RF, NN, and DT) have a significant effect in detecting potential fraud, except for SVM. The significant negative value of SVM does not indicate a poor model but rather a mismatch between the model and the research type. Researchers found an increase in the potential for regional fraud in 2023. A total of 143 districts/cities previously (in 2022) categorized as "No Potential" and "Moderate Potential" changed their status to "High Potential" for fraud. Finally, this study contributes to the literature by providing insights into the accuracy of using ML models (RF, NN, and DT) in detecting potential financial fraud in the Indonesian local government sector based on regional big data.

Keyword: Regional fraud, machine learning, Indonesia