

ABSTRACT

Yosua Efraim Young (01679210013)

DETECTION OF FRAUDULENT FINANCIAL STATEMENT USING ANN AND SMOTE UNDER IMBALANCED DATA

(xii + 50 pages; 3 figures; 9 tables, 1 appendix)

This study aims to create a model to detect fraudulent financial statements. Previous study has been conducted and showcased the ability of supervised learning in identifying fraudulent financial statements. However, those studies generally show its effectiveness on a balanced dataset, where it is frequently found that fraud datasets are presented as imbalanced. This can potentially lead to a model that is inadequately trained with fraud characteristics, which is the minority class.

This study employs artificial neural networks, specifically feedforward neural networks to perform binary classification. Using previous research dataset that shows severe imbalance dataset, which fraud samples only represented by 0.67% from the population, SMOTE is employed. SMOTE is used to oversample the minority class, hence generating additional samples and creating a balance dataset.

This study finds that artificial neural networks is able to identify fraudulent financial statements in a severe imbalance dataset. Found in literature as benchmark, the best results of the proposed model show an AUC score and precision rate of 70.6% and 2.89%, respectively, in comparable to existing models, and sensitivity rate of 83% outperforms all the existing models. The proposed model offers reliable assurance for regulators or auditors with a low risk of false negatives, hence, ensure practical utility and maintain confidence.

Keywords: Artificial Neural Networks, Fraudulent Financial Statements, SMOTE, Supervised Learning

References 40 (1997 – 2022)

ABSTRAK

Yosua Efraim Young (01679210013)

DETECTION OF FRAUDULENT FINANCIAL STATEMENT USING ANN AND SMOTE UNDER IMBALANCED DATA

(xii + 50 halaman; 3 gambar; 9 tabel; 1 lampiran)

Penelitian ini bertujuan untuk membuat sebuah model yang dapat mendeteksi kecurangan pelaporan keuangan. Penelitian sebelumnya menunjukkan kemampuan pembelajaran terarah (*supervised learning*) dalam mendeteksi kecurangan pelaporan keuangan. Namun, penelitian tersebut umumnya menunjukkan efektivitasnya pada dataset yang seimbang, dimana seringkali ditemukan bahwa dataset terkait kecurangan pelaporan keuangan disajikan dalam bentuk tidak seimbang. Hal ini dapat berpotensi menyebabkan model tidak cukup terlatih menghadapi karakteristik kecurangan pelaporan keuangan, yang dalam kasus dimana data tidak seimbang, kecurangan termasuk ke dalam kelas minoritas.

Penelitian ini menggunakan jaringan saraf tiruan (*artificial neural networks*), khususnya *feedforward networks* untuk melakukan klasifikasi biner. Menggunakan dataset dari penelitian sebelumnya yang menunjukkan ketidakseimbangan yang tinggi, dimana sampel kecurangan pelaporan keuangan hanya diwakili oleh 0,67% dari populasi, *SMOTE* digunakan. *SMOTE* digunakan untuk meningkatkan jumlah sampel dari kelas minoritas, sehingga menghasilkan sampel tambahan dan menciptakan dataset yang seimbang.

Penelitian ini menemukan bahwa jaringan saraf tiruan mampu mengidentifikasi kecurangan pelaporan keuangan dalam dataset yang tidak seimbang. Menggunakan literatur sebagai tolak ukur, hasil terbaik dari model yang diusulkan menunjukkan skor *AUC* dan tingkat *precision* sebesar 70.6% dan 2.89%, berturut-turut, dibandingkan model yang ada, dan tingkat *sensitivity* sebesar 83% melebihi semua model yang ada. Model yang diusulkan menawarkan asurans yang dapat diandalkan untuk regulator dan auditor dengan risiko rendah terjadinya *false negative*, dengan demikian, memastikan kegunaan praktis dan mampu menjaga kepercayaan.

Kata Kunci: Jaringan Saraf Tiruan, Kecurangan Pelaporan Keuangan, SMOTE, Pembelajaran Terarah
References 40 (1997– 2022)