

ABSTRAK

Pengelompokan perusahaan berdasarkan sektor industri merupakan aspek penting dalam analisis investasi, namun klasifikasi yang dilakukan secara manual di Bursa Efek Indonesia (BEI) masih kurang optimal dalam memanfaatkan data laporan keuangan sebagai dasar prediksi. Oleh karena itu, penelitian ini mengembangkan model prediksi sektor industri berdasarkan laporan keuangan menggunakan metode *Linear Discriminant Analysis* (LDA) dan membandingkannya dengan *Extreme Gradient Boosting* (XGBoost). Data laporan keuangan perusahaan yang terdaftar di BEI dalam rentang 2010–2022 digunakan sebagai dataset utama, dengan proses *pre-processing*, normalisasi, dan *oversampling* menggunakan *Borderline-SMOTE* untuk mengatasi ketidakseimbangan kelas. Evaluasi model dilakukan dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta dilakukan analisis fitur menggunakan *Permutation Importance* untuk menentukan variabel yang paling berpengaruh. Hasil penelitian menunjukkan bahwa metode LDA memiliki akurasi 27,51%, sedangkan XGBoost mencapai 63,87%, yang menunjukkan bahwa pendekatan non-linear XGBoost lebih unggul dalam mengklasifikasikan sektor industri berdasarkan laporan keuangan. Selain itu, fitur total aset, total pendapatan, dan inventaris ditemukan sebagai variabel paling berpengaruh dalam prediksi. Penelitian ini memberikan kontribusi dalam pengembangan metode otomatis untuk klasifikasi sektor industri, yang dapat digunakan oleh investor dan analis dalam mendukung pengambilan keputusan investasi yang lebih akurat.

Kata Kunci: bursa efek indonesia, linear discriminant analysis, xgboost, laporan keuangan, prediksi sektor industri.