



Stock Industry Sector Prediction Based on Financial Reports using Random Forest Method

Kamil Elian Zhafran*, Deni Saepudin

School of Computing, Telkom University, Bandung, Indonesia

Email: ¹*zhafrankamil@student.telkomuniversity.ac.id, ²denisaepudin@telkomuniversity.ac.id

Correspondence Author Email: zhafrankamil@student.telkomuniversity.ac.id

Submitted: 99/99/9999; Accepted: 99/99/9999; Published: 99/99/9999

Abstrak– Penelitian ini bertujuan untuk memprediksi sektor industri saham di Bursa Efek Indonesia (BEI) berdasarkan laporan keuangan menggunakan metode Random Forest. Penerapan pendekatan machine learning ini sangat penting karena kompleksitas data keuangan yang menuntut metode yang kuat dan adaptif untuk prediksi yang akurat. Dataset tersebut terdiri dari data keuangan dari perusahaan di 10 sektor industri di BEI, yang mencakup tahun 2010-2022, dan mencakup 17 fitur dari setiap laporan keuangan. Khususnya, terdapat ketidakseimbangan jumlah perusahaan per sektor, dengan sektor B mewakili 14,76% dan sektor G hanya 1,98%. Ketidakseimbangan ini menimbulkan bias dalam analisis data, sehingga memerlukan penerapan metode oversampling SMOTE untuk mengatasinya. Proses penelitian melibatkan pembersihan data, pemisahan data menjadi 80% set pelatihan dan 20% set pengujian, penerapan teknik oversampling SMOTE, dan membandingkan prediksi dari dataset yang tidak seimbang dan seimbang. Metode Random Forest dipilih karena kemampuannya menangani dataset yang kompleks untuk klasifikasi sektor industri. Hasil evaluasi menunjukkan bahwa tanpa oversampling, model mencapai akurasi 73,57%, presisi 74,29%, recall 73,57%, dan F1-score 73,51%. Dengan oversampling, metrik ini meningkat menjadi akurasi 80,21%, presisi 81,34%, recall 80,21%, dan F1-score 80,45%.

Kata kunci: Prediksi Sektor Industri; Bursa Efek Indonesia; Laporan Keuangan; Random Forest; SMOTE