ABSTRAK

Penelitian ini membahas performa algoritma Decision Tree C 4.5 dan Random Forest dalam klasifikasi data yang tidak seimbang menggunakan metode Non-Balancing, Balancing, dan Hyperparameter. Hasil menunjukkan bahwa pada metode NoniBalancing, Random Forest memiliki akurasi lebih tinggi dibandingkan Decision Tree C 4.5. Penerapan metode Balancing secara signifikan meningkatkan performa kedua algoritma, dengan Random Forest mencapai akurasi tertinggi sebesar 84.19%. Pengaturan Hyper Parameter juga memberikan peningkatan performa, terutama pada Random Forest yang mencapai akurasi hingga 84.85%. Kesimpulannya, Random Forest dengan metode Balancing dan Hyperparameter Tuning menunjukkan performa terbaik dalam hal akurasi, precision, recall, dan F1-Score.

Kata Kunci : Decision Tree C 4.5, Random Forest, Non-Balancing, Balancing, Hyper Parameter, Klasifikasi Data Tidak Seimbang

ABSTRACT

This study examines the performance of Decision Tree C 4.5 and Random Forest algorithms in classifying imbalanced data using Non-Balancing, Balancing, and Hyperparameter methods. The results indicate that in the Non-Balancing method, Random Forest has higher accuracy compared to Decision Tree C 4.5. The application of Balancing methods significantly enhances the performance of both algorithms, with Random Forest achieving the highest accuracy of 84.19%. Hyper Parameter tuning also improves performance, especially for Random Forest, which reaches an accuracy of up to 84.85%. In conclusion, Random Forest with Balancing and Hyper Parameter methods shows the best performance in terms of accuracy, precision, recall, and F1-Score.

Keywords: Decision Tree C 4.5, Random Forest, Non-Balancing, Balancing, Hyper Parameter, Imbalanced Data Classification