ABSTRAK

Kualitas air merupakan faktor penting dalam keberhasilan budidaya ikan karena berpengaruh terhadap pertumbuhan, kesehatan, dan kelangsungan hidup ikan. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma Decision Tree, Random Forest, dan K-Nearest Neighbors dalam mengklasifikasikan kesesuaian kualitas air untuk budidaya ikan Bandeng, Kakap Putih, dan Kerapu Macan. Laporan ini bermanfaat sebagai referensi penerapan Machine Learning dalam analisis kualitas air serta sebagai dasar awal pengembangan sistem pendukung keputusan pada bidang akuakultur.

Data penelitian diperoleh dari tiga sumber, yaitu ACEH, BCO-DMO, dan PANGAEA, kemudian diselaraskan berdasarkan lima parameter kualitas air, yaitu pH, suhu, Dissolved Oxygen (DO), salinitas, dan Total Dissolved Solids (TDS). Dari 582 data awal, proses pembersihan menghasilkan 309 data bersih, kemudian ditambahkan 1.691 data sintetis sehingga dataset akhir berjumlah 2.000 data. Label kesesuaian dibentuk secara biner untuk masing-masing jenis ikan berdasarkan parameter pH, suhu, DO, dan salinitas, sedangkan TDS digunakan sebagai fitur pendukung. Dataset dibagi menjadi data latih dan data uji dengan rasio 80:20, kemudian model dioptimasi menggunakan GridSearchCV dengan 5-fold cross-validation.

Hasil penelitian menunjukkan bahwa Random Forest menghasilkan kinerja terbaik dibandingkan Decision Tree dan KNN. Random Forest memperoleh Macro F1-Score sebesar 0,9937 dan Hamming Loss sebesar 0,0031, sedangkan Decision Tree memperoleh Macro F1-Score sebesar 0,9803 dan KNN sebesar 0,9569. Random Forest juga menunjukkan performa yang stabil pada seluruh target ikan, dengan tingkat kesalahan prediksi paling rendah. Sementara itu, KNN masih mampu menghasilkan performa yang baik, tetapi memiliki kecenderungan kesalahan lebih tinggi, terutama pada prediksi False Positive.

Berdasarkan hasil pengujian, Random Forest menjadi algoritma yang paling optimal untuk klasifikasi kesesuaian kualitas air pada dataset penelitian ini. Algoritma berbasis pohon, khususnya Random Forest, lebih mampu menangkap pola batas kesesuaian kualitas air dibandingkan pendekatan berbasis jarak seperti KNN. Penelitian selanjutnya disarankan menggunakan data empiris dari sistem monitoring kualitas air berbasis IoT, melakukan validasi label bersama ahli budidaya perikanan, serta membandingkan algoritma lain seperti SVM, XGBoost, atau Artificial Neural Network agar hasil klasifikasi semakin kuat dan siap dikembangkan pada kondisi lapangan.

Kata kunci*: Decision Tree, K-Nearest Neighbors, kualitas air, Random Forest*, sistem pendukung keputusan.