

ABSTRAK

ANALISIS PERBANDINGAN ALGORITMA CATBOOST DAN LIGHTGBM DALAM PREDIKSI *MULTI-LABEL* SERANGAN HAMA PADA TEBU

Oleh

MUTIARA CINTIA RAINY

Penelitian ini bertujuan membangun model prediksi *multi-label* tingkat keparahan serangan hama tebu (penggerek batang, penggerek pucuk, dan kutu perisai) menggunakan algoritma CatBoost dan LightGBM. *dataset* pengamatan lapangan diperkaya dengan fitur temporal berbasis *lag* (1–3) serta fitur kategorikal yang diproses secara native oleh masing-masing algoritma. Pembagian data dilakukan berbasis blok untuk mencegah *data leakage* dan merepresentasikan skenario prediksi pada blok yang belum diamati. Evaluasi menggunakan *F1-score macro average* dan *AUC macro* pada tiga skema pembagian data menunjukkan bahwa kedua model memiliki performa yang relatif sebanding, dengan *F1-score* pada kisaran 0,59–0,69 dan *AUC macro* 0,80–0,86. Pada tingkat kelas, kemampuan deteksi terhadap kelas keparahan tinggi konsisten dengan *AUC* di atas 0,90. LightGBM unggul dalam waktu pelatihan, sedangkan CatBoost lebih stabil dan memiliki waktu prediksi lebih cepat. Model yang direkomendasikan diharapkan mendukung sistem pemantauan risiko serangan hama secara dini dan pengambilan keputusan yang lebih terarah.

Kata kunci: Klasifikasi *Multi-label*, LightGBM, CatBoost, Serangan Hama Tebu.

ABSTRACT

COMPARATIVE ANALYSIS OF CATBOOST AND LIGHTGBM ALGORITHMS FOR MULTI-LABEL PREDICTION OF PEST INFESTATION IN SUGARCANE

By

MUTIARA CINTIA RAINY

This study aims to develop a multi-label prediction model for the severity levels of sugarcane pest infestations (stem borer, top borer, and scale insects) using the CatBoost and LightGBM algorithms. Field observation data were enriched with temporal lag-based features (1–3 periods) and categorical features processed natively by each algorithm. Data splitting was performed at the block level to prevent data leakage and to represent prediction scenarios for previously unseen blocks. Evaluation using macro-averaged F1-score and macro AUC across three data-splitting schemes indicates that both models achieve comparable performance, with F1-scores ranging from 0,59 to 0,69 and macro AUC between 0,80 and 0,86. At the class level, detection performance for the high-severity class is consistently strong, with AUC values above 0,90. LightGBM demonstrates superior training speed, while CatBoost shows greater stability and faster inference time. The recommended model is expected to support early pest risk monitoring systems and enable more targeted decision-making in sugarcane cultivation.

Keywords: Multi-label Classification, LightGBM, CatBoost, Sugarcane Pest Infestation.