

ABSTRAK

Permasalahan status gizi balita di Indonesia masih tergolong tinggi sehingga diperlukan metode klasifikasi yang akurat, terutama pada data yang tidak seimbang. Penelitian ini bertujuan mengimplementasikan teknik *hybrid sampling* SMOTE-ENN pada algoritma C5.0 dengan metode AdaBoost untuk meningkatkan performa klasifikasi status gizi. Dataset yang digunakan berjumlah 2.759 data dengan enam fitur, yaitu usia, jenis kelamin, tinggi badan, berat badan, lingkaran lengan atas (LiLA), dan status gizi sebagai variabel target. Tahapan penelitian meliputi imputasi *missing value*, transformasi usia, normalisasi data numerik, *label encoding*, penanganan outlier (dihapus, dibiarkan, atau diimputasi), serta penyeimbangan data menggunakan SMOTE, ENN, dan SMOTE-ENN. Pengujian dilakukan pada tiga skenario pembagian data (70:30, 80:20, dan 90:10) dengan variasi parameter dan kondisi data, yang dievaluasi menggunakan akurasi, precision, recall, specificity, F1-score, dan G-Mean. Hasil terbaik diperoleh pada skenario 90:10 dengan SMOTE-ENN dan C5.0–AdaBoost ($k_neighbors=3$, $n_neighbors=5$, $max_depth=7$, $n_estimator=100$) dengan akurasi 97,68%, precision 96,37%, recall 97,76%, specificity 98,46%, F1-score 96,98%, dan G-Mean 97,74%. Penerapan AdaBoost meningkatkan akurasi sebesar 0,37% hingga 12,68%. Namun, performa SMOTE-ENN dipengaruhi oleh karakteristik data dan konfigurasi parameter, serta belum sepenuhnya melampaui SMOTE.

Kata Kunci : Klasifikasi status gizi, *Hybrid Sampling*, SMOTE-ENN, algoritma C5.0, AdaBoost.