

INTISARI

PERBANDINGAN ANALISIS KLASIFIKASI EXTREME GRADIENT BOOSTING DAN LIGHT GRADIENT BOOSTING MACHINE PADA DATA STUNTING DI NUSA TENGGARA TIMUR

Oleh

DESI NATALIA MUSKANANFOLA

23/511924/PPA/06494

Stunting merupakan masalah kesehatan serius di Indonesia yang memengaruhi pertumbuhan fisik dan kognitif anak. Prevalensi stunting di Nusa Tenggara Timur (NTT) pada tahun 2021 mencapai 37,8%, tertinggi di Indonesia. Penelitian ini membandingkan algoritma *Extreme Gradient Boosting* (XGBoost) dan *Light Gradient Boosting Machine* (LightGBM) untuk mendeteksi stunting pada anak di NTT. Data telah melalui proses *preprocessing*, termasuk *data cleaning*, *label encoding*, *handling imbalance class* dengan SMOTE, dan *splitting data*. Evaluasi dilakukan pada skenario pembagian data 80:20, 70:30, dan 60:40 menggunakan metrik *accuracy*, *precision*, *recall*, *F1-Score*, dan *Area Under the Curve* (AUC). Hasil menunjukkan XGBoost unggul dalam performa klasifikasi, dengan *accuracy* 93,11%, *precision* 91,70%, *recall* 95,61%, *F1-Score* 93,62%, dan AUC 0,9848 pada pembagian data 80:20. Sementara itu, LightGBM lebih efisien dengan waktu komputasi rata-rata sepuluh kali lebih cepat meskipun memiliki performa yang lebih rendah (*accuracy* 82,96%, AUC 0,9072). Penelitian juga mengevaluasi kesesuaian data stunting NTT dengan standar Peraturan Menteri Kesehatan dan WHO, menunjukkan tingkat kesesuaian sebesar 83,57%. Kesimpulannya, XGBoost unggul dalam akurasi, sementara LightGBM lebih efisien dalam waktu komputasi, memungkinkan kedua algoritma digunakan secara komplementer dalam analisis stunting.

Kata Kunci Stunting, Nusa Tenggara Timur, XGBoost, LightGBM, SMOTE.

ABSTRACT

COMPARATIVE ANALYSIS OF EXTREME GRADIENT BOOSTING AND LIGHT GRADIENT BOOSTING MACHINE CLASSIFICATION ON STUNTING DATA IN EAST NUSA TENGGARA

By

DESI NATALIA MUSKANANFOLA

23/511924/PPA/06494

Stunting is a significant public health issue in Indonesia, affecting children's physical and cognitive growth. In 2021, Nusa Tenggara Timur (NTT) recorded the highest stunting prevalence in Indonesia at 37.8%. This study compares the performance of *Extreme Gradient Boosting* (XGBoost) and *Light Gradient Boosting Machine* (LightGBM) algorithms in detecting stunting among children in NTT. The data underwent *preprocessing*, including *data cleaning*, *label encoding*, handling imbalance with SMOTE, and *splitting data*. Evaluations were conducted using data splits of 80:20, 70:30, and 60:40, with metrics such as *accuracy*, *precision*, *recall*, *F1-Score*, and *Area Under the Curve* (AUC). Results show XGBoost outperformed LightGBM in classification performance, achieving *accuracy* of 93.11%, *precision* of 91.70%, *recall* of 95.61%, *F1-Score* of 93.62%, and AUC of 0.9848 on the 80:20 split. Conversely, LightGBM demonstrated superior computational efficiency with nearly ten times faster training time, though with lower performance (*accuracy* of 82.96%, AUC of 0.9072). Furthermore, this study evaluated the alignment of stunting data in NTT with the Ministry of Health and WHO standards, finding an agreement rate of 83.57%. In conclusion, XGBoost excels in classification accuracy, while LightGBM is more efficient in computation time, making them complementary tools for stunting analysis.

Keywords Stunting, East Nusa Tenggara, XGBoost, LightGBM, SMOTE.