# ABSTRAK

Provinsi Jawa Barat menghadapi tantangan besar dalam menekan angka stunting akibat besarnya populasi balita. Fluktuasi tren kasus secara spasial dan temporal menyulitkan pemerintah dalam mengalokasikan sumber daya secara presisi. Penelitian ini bertujuan memprediksi jumlah balita stunting di 27 kabupaten dan kota menggunakan pemodelan regresi berbasis *machine learning*. Sistem mengeksekusi komparasi antara algoritma *Random Forest* dan *Extreme Gradient Boosting* (XGBoost) menggunakan matriks data panel historis. Kebaruan penelitian ini terletak pada integrasi rekayasa fitur (*Lag-1* dan dimensi ekonomi), penerapan *Time-Series Cross Validation*, serta optimasi *hyperparameter* melalui *RandomizedSearchCV*. Hasil komputasi membuktikan bahwa rekayasa fitur sukses mendongkrak performa kedua algoritma. *Random Forest* terbukti sebagai model paling superior dengan skor $R^2$ sebesar 0,85, mengungguli XGBoost yang hanya mencapai 0,77. *Random Forest* juga mencatatkan tingkat kesalahan terendah dengan nilai MAE sebesar 1.552 dan RMSE sebesar 2.278. Kesimpulannya, model *Random Forest* terbukti sangat presisi dalam menangkap kompleksitas tren stunting di Jawa Barat. Arsitektur model final ini kemudian berhasil diintegrasikan ke dalam antarmuka aplikasi berbasis web sebagai instrumen proyeksi interaktif bagi pengguna.

**Kata kunci:** analitik prediktif, ensemble learning, optimasi hyperparameter, random forest, rekayasa fitur, stunting

UBP KARAWANG

# ABSTRACT

West Java Province faces significant challenges in reducing stunting due to its massive toddler population. Spatial and temporal fluctuations complicate precise government resource allocation. This study predicts stunting cases across 27 regencies and cities using machine learning regression. The system compares Random Forest (RF) and Extreme Gradient Boosting (XGBoost) using historical panel data. Novelty involves integrating feature engineering (Lag-1 and economic dimensions), Time-Series Cross Validation, and RandomizedSearchCV hyperparameter optimization. Results demonstrate that feature engineering significantly improves both models. Random Forest emerged as superior with an $R^2$ of 0.85, outperforming XGBoost's 0.77. Random Forest achieved lower error rates with a Mean Absolute Error (MAE) of 1,552 and a Root Mean Squared Error (RMSE) of 2,278. In conclusion, Random Forest is highly precise in capturing stunting trends in West Java. The final architecture was successfully integrated into a web-based application for interactive projections.

***Keywords*****:** *ensemble learning, feature engineering, hyperparameter optimization, predictive analytics, random forest, stunting*

UBP KARAWANG