

ABSTRACT

The high rate of child stunting in West Java Province represents a serious issue requiring a data-driven approach for effective prevention and intervention. This study aims to develop a prediction and classification system for stunting levels in children using the Logistic regression algorithm, while also presenting the results visually through a web-based information system. The data used includes the number of live births and stunting cases across 27 regencies/cities from 2020 to 2024, sourced from Open Data Jabar. The research process involved Exploratory Data Analysis (EDA), data aggregation from district to city/regency level, stunting percentage calculation, and binary classification modeling based on a 10% threshold to categorize low and high stunting levels.

Evaluation results show that the model achieved an accuracy of 96.3%, with f1-scores of 0.97 for the low category and 0.96 for the high category. The model also demonstrated strong identification of high-priority areas with a recall score of 1.00 in the high category. The developed system includes interactive visualizations showing annual stunting trends by region and offers predictions for the year 2025 based on the average stunting percentages from the previous five years. These findings indicate that the application of machine learning can serve as an effective tool to support data-driven stunting intervention policy decisions.

Keywords: *stunting, machine learning, logistic regression, prediction, information system.*

ABSTRAK

Tingginya angka *stunting* balita di Provinsi Jawa Barat menjadi permasalahan serius yang memerlukan pendekatan berbasis data dalam upaya pencegahan dan penanganannya. Penelitian ini bertujuan untuk membangun sistem prediksi dan klasifikasi tingkat *stunting* balita menggunakan algoritma *Logistic regression*, serta menyajikan hasil analisis secara visual melalui sistem informasi berbasis web. Data yang digunakan berupa jumlah bayi lahir dan jumlah balita *stunting* di 27 kabupaten/kota selama periode 2020–2024, yang diperoleh dari portal Open Data Jabar. Penelitian ini mencakup tahapan *Exploratory Data Analysis* (EDA), agregasi data dari tingkat kecamatan ke kabupaten/kota, perhitungan persentase *stunting*, serta pembangunan model klasifikasi biner berdasarkan ambang batas 10% untuk membedakan kategori *stunting* rendah dan tinggi.

Hasil evaluasi model menunjukkan akurasi sebesar 96,3%, dengan nilai *f1-score* sebesar 0,97 untuk kategori rendah dan 0,96 untuk kategori tinggi. Model juga memiliki kemampuan identifikasi wilayah prioritas dengan *recall* 1,00 pada kategori tinggi. Sistem yang dibangun dilengkapi fitur visualisasi interaktif yang menampilkan tren *stunting* per kabupaten/kota dari tahun ke tahun, dan juga menyediakan prediksi kondisi tahun 2025 berdasarkan rata-rata persentase *stunting* lima tahun sebelumnya. Temuan ini menunjukkan bahwa penerapan *machine learning* dapat menjadi alat bantu yang efektif dalam mendukung pengambilan Keputusan untuk pencegahan *stunting* yang lebih terarah dan berbasis data.

Kata kunci: *stunting*, *machine learning*, *logistic regression*, prediksi, sistem informasi, klasifikasi.