

ABSTRAK

Vinny Ramayani Saragih, NIM 4203250007 (2020). Perbandingan Metode Supervised Machine Learning dalam Prediksi Prevalensi Stunting di Provinsi Sumatera Utara.

Stunting adalah gangguan tumbuh kembang anak yang diakibatkan oleh kekurangan gizi kronis serta infeksi yang berulang. Stunting memiliki dampak jangka pendek, dan panjang yang signifikan dan merupakan salah satu masalah di bidang kesehatan yang saat ini sedang dihadapi oleh masyarakat Indonesia. Provinsi Sumatera Utara memiliki nilai prevalensi stunting sebesar 18,9% Pemerintah Provinsi Sumatera Utara memiliki target untuk menurunkan prevalensi stunting pada tahun 2024 menjadi 14%. Penelitian ini bertujuan untuk membandingkan kinerja beberapa metode *supervised machine learning* dalam memprediksi prevalensi stunting di Provinsi Sumatera Utara. Data yang digunakan adalah data sekunder dari tahun 2021 hingga 2023 yang mencakup 33 kabupaten/kota di provinsi tersebut. Penelitian ini mengevaluasi tiga model machine learning, yaitu *Support Vector Regression* (SVR), *Decision Tree*, dan *Random Forest*, dengan menggunakan metrik evaluasi seperti *Mean Squared Error* (MSE), dan *Mean Absolute Percentage Error* (MAPE). Hasil penelitian menunjukkan bahwa SVR unggul pada wilayah dengan data yang terstruktur, seperti Kabupaten Karo (MSE: 1.24E-05, MAPE: 0.014%), sementara Random Forest menunjukkan performa terbaik pada wilayah dengan kompleksitas data yang tinggi, seperti Kabupaten Nias Barat (MSE: 1.26E-29, MAPE: 1.23E-14%). Di sisi lain, Decision Tree memberikan hasil baik pada wilayah dengan pola data sederhana, seperti Kabupaten Simalungun (MSE: 0.09, MAPE: 1.69%). Namun, beberapa wilayah seperti Kabupaten Padang Lawas menunjukkan nilai MSE yang sangat tinggi (MSE: 13,972.31, menggunakan SVR, yang disebabkan oleh distribusi data yang tidak merata antar tahun. Penelitian ini memberikan kontribusi pada literatur terkait penggunaan model pembelajaran mesin dalam prediksi stunting dan menawarkan implikasi praktis untuk mendukung pengambilan kebijakan kesehatan berbasis data. Penelitian lanjutan disarankan untuk mengeksplorasi variabel tambahan, model pembelajaran mesin yang lebih canggih, serta analisis spasial dan temporal untuk meningkatkan akurasi prediksi.

Kata kunci: Stunting, Prevalensi Stunting, Sumatera Utara, Machine Learning, Supervised Learning.

ABSTRACT

Vinny Ramayani Saragih, NIM 4203250007 (2020). Perbandingan Metode Supervised Machine Learning dalam Prediksi Prevalensi Stunting di Provinsi Sumatera Utara.

Stunting is a growth and development disorder in children caused by chronic malnutrition and repeated infections. Stunting has significant short-term and long-term impacts and is currently one of the Indonesian population's major health issues. North Sumatra Province has a stunting prevalence rate of 18.9%. The provincial government has targeted reducing the stunting prevalence rate to 14% by 2024. This study aims to compare the performance of several *supervised machine learning* methods in predicting stunting prevalence in North Sumatra Province. The data used are secondary data from 2021 to 2023 covering 33 districts/cities in the province. This research evaluates three machine learning models: *Support Vector Regression* (SVR), *Decision Tree*, and *Random Forest*, using evaluation metrics such as *Mean Squared Error* (MSE), and *Mean Absolute Percentage Error* (MAPE). The research results show that SVR excels in areas with structured data, such as Karo Regency (MSE: 1.24E-05, MAPE: 0.014%), while Random Forest performs best in areas with high data complexity, such as West Nias Regency (MSE: 1.26E-29, MAPE: 1.23E-14%). On the other hand, Decision Tree yields good results in areas with simple data patterns, such as Simalungun Regency (MSE: 0.09, MAPE: 1.69%). However, some areas like Padang Lawas Regency show very high MSE values (MSE: 13,972.31, using SVR), caused by uneven data distribution across years. This research contributes to the literature on the use of machine learning models in predicting stunting and offers practical implications to support data-driven health policy-making. Further research is recommended to explore additional variables, more advanced machine learning models, as well as spatial and temporal analyses to improve prediction accuracy.

Keywords: Stunting, Stunting Prevalence, North Sumatra, Machine Learning, Supervised Learning.