ABSTRAK

MUHAMAD NUR AMIN. Analisis Prediksi Polutan PM2.5 dengan Algoritma *Differential Evolution* dan *XGBoost*. Dibimbing oleh HENDRA KURNIAWAN dan NOLAN EFRANDA.

PM2.5 merupakan polutan udara paling berbahaya bagi kesehatan manusia, terutama di kawasan tropis yang rentan lonjakan episodik akibat kabut asap lintas batas. Penelitian ini bertujuan membandingkan akurasi prediksi PM2.5 antara model *XGBoost* standar (*Baseline*) dan model *hibrida* DE-*XGBoost* yang dioptimalkan dengan *Differential Evolution* (DE) serta diperkaya *fitur lag* historis, menggunakan data deret waktu AQMS Serpong, Tangerang Selatan (17.520 observasi, interval 30 menit, tahun 2023). DE mengoptimalkan seleksi fitur dan *hyperparameter XGBoost* secara simultan, mereduksi 15 fitur kandidat menjadi 4 fitur terpilih: kecepatan angin, suhu udara, radiasi matahari, dan PM2.5 30 menit sebelumnya. Hasil menunjukkan DE-*XGBoost* mengungguli *Baseline* pada ketiga metrik evaluasi: RMSE turun 4,55% (8,6918→8,2964 µg/m³), MAE turun 4,26% (6,0770→5,8181 µg/m³), dan R^2 meningkat dari 0,8602 menjadi 0,8726. Uji *Wilcoxon Signed-Rank* (p=0,0086) mengonfirmasi keunggulan ini signifikan secara statistik, dengan perbaikan paling nyata pada kategori polusi tinggi (MAE turun 10,9%). Penelitian ini juga menghasilkan prototipe aplikasi *web* berbasis *Streamlit* sebagai demonstrasi sistem peringatan dini kualitas udara.

Kata kunci: *Differential Evolution*, *XGBoost*, PM2.5, optimasi *hyperparameter*, prediksi kualitas udara.

ABSTRACT

MUHAMAD NUR AMIN. Prediction Analysis of PM2.5 Pollutant Using Differential Evolution and XGBoost Algorithm. Supervised by HENDRA KURNIAWAN and NOLAN EFRANDA.

PM2.5 is one of the most dangerous air pollutants to human health, especially in tropical areas prone to episodic surges from cross-border haze. This study aims to compare PM2.5 prediction accuracy between a standard XGBoost model (Baseline) and a hybrid DE-XGBoost model optimized with Differential Evolution (DE) and enriched with historical lag features, using time series data from the Serpong AQMS, South Tangerang (17,520 observations, 30-minute interval, 2023). DE simultaneously optimizes feature selection and XGBoost hyperparameters, reducing 15 candidate features to 4 selected features: wind speed, air temperature, solar radiation, and PM2.5 30 minutes prior. Results show DE-XGBoost outperforms the Baseline on all three evaluation metrics: RMSE decreased 4.55% (8.6918→8.2964 µg/m³), MAE decreased 4.26% (6.0770→5.8181 µg/m³), and R^2 increased from 0.8602 to 0.8726. The Wilcoxon Signed-Rank Test (p=0.0086) confirmed this improvement is statistically significant, with the most notable gain in the high-pollution category (10.9% MAE reduction). This study also produces a Streamlit-based web application prototype as a demonstration of an air quality early warning system.

Keywords: Differential Evolution, XGBoost, PM2.5, hyperparameter optimization, air quality prediction.