

## INTISARI

### PERBANDINGAN MODEL ARIMAX DAN LSTM UNTUK PERAMALAN DATA RUNTUN WAKTU DATA KONSENTRASI PM2.5

Oleh

Veren Oktavia Halim

21/474844/PA/20504

Peningkatan konsentrasi PM2.5 sebagai polutan udara berbahaya telah menjadi perhatian penting, khususnya di kota-kota besar seperti Pekanbaru. Oleh karena itu, diperlukan metode peramalan yang akurat untuk memantau dan memprediksi tingkat polusi udara sebagai dasar pengambilan keputusan. Penelitian ini bertujuan untuk membandingkan performa dua pendekatan peramalan data runtun waktu, yaitu model statistik *Autoregressive Integrated Moving Average with Exogenous Variables* (ARIMAX) dan model *deep learning Long Short-Term Memory* (LSTM), dalam memprediksi konsentrasi PM2.5 berdasarkan data harian tahun 2023–2025. Model ARIMAX dirancang dengan mempertimbangkan stasioneritas data serta penggunaan variabel eksogen seperti suhu, kelembapan, dan kecepatan angin. Sementara itu, model LSTM memanfaatkan kemampuan dalam mempelajari pola non-linear jangka panjang dan dilakukan *hyperparameter tuning* terhadap unit, *batch size*, *learning rate*, dan *dropout*. Evaluasi performa dilakukan menggunakan metrik MAE, RMSE, dan MAPE. Hasil penelitian menunjukkan bahwa model LSTM memberikan hasil prediksi yang kompetitif, terutama dalam memproyeksikan nilai PM2.5 satu minggu ke depan dengan rentang prediksi yang masih wajar ( $10\text{--}30 \mu\text{g}/\text{m}^3$ ). Dengan demikian, model LSTM dinilai lebih adaptif untuk digunakan dalam sistem peringatan dini kualitas udara di Pekanbaru.

**Kata Kunci:** PM2.5, ARIMAX, LSTM, Peramalan, Kualitas Udara, *Deep Learning*.

## ABSTRACT

### COMPARISON OF ARIMAX AND LSTM MODELS FOR TIME SERIES FORECASTING OF PM2.5 CONCENTRATION DATA

By

Veren Oktavia Halim

21/474844/PA/20504

The increase in PM2.5 concentration, a hazardous air pollutant, has become a significant concern, especially in urban areas such as Pekanbaru. Therefore, accurate forecasting methods are essential for monitoring and predicting air pollution levels to support decision-making processes. This study aims to compare the performance of two time series forecasting approaches: the statistical Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) model and the deep learning-based Long Short-Term Memory (LSTM) model, in predicting PM2.5 concentrations using daily data from 2023 to 2025. The ARIMAX model was constructed by ensuring data stationarity and incorporating exogenous variables such as temperature, humidity, and wind speed. Meanwhile, the LSTM model leveraged its ability to learn long-term nonlinear dependencies, with hyperparameter tuning applied to units, batch size, learning rate, and dropout rate. The model performances were evaluated using MAE, RMSE, and MAPE metrics. The results show that the LSTM model produced competitive forecasts, particularly in projecting PM2.5 values for the next seven days with a reasonable prediction range ( $10\text{--}30 \mu\text{g}/\text{m}^3$ ). Thus, the LSTM model is considered more adaptive for implementation in early warning systems for air quality monitoring in Pekanbaru.

**Keyword:** PM2.5, ARIMAX, LSTM, Forecasting, Air Quality, Deep Learning