

ABSTRAK

Memburuknya kualitas udara di Indonesia menuntut kami untuk meninjau dan meningkatkan pendeteksian dan prediksi kualitas udara untuk menghasilkan indikator yang lebih akurat untuk manajemen kualitas udara, khususnya dalam sistem peringatan kesehatan. Oleh karena itu, penelitian ini bertujuan untuk membandingkan dua metode peramalan yaitu SARIMA dan LSTM dalam memprediksi indeks kualitas udara (AQI) di Kemang, Jakarta. Menurut penelitian sebelumnya, metode deep learning LSTM mengungguli SARIMA dalam memprediksi AQI di wilayah metropolitan. Namun, metode SARIMA yang lebih sederhana dapat memberikan prediksi yang lebih baik di wilayah pesisir. Dengan itu, peneliti melakukan riset untuk menemukan metode peramalan yang tepat untuk daerah Kemang, Jakarta Selatan. Prosedur penelitian ini dimulai dengan memastikan data time-series stasioner, kemudian melakukan time-series train-test split, dilanjutkan dengan tuning hiperparameter dengan menggunakan metode Bayes atau random search kemudian membandingkan kedua metode prediksi tersebut berdasarkan nilai RMSE dari masing-masing metode untuk menarik kesimpulan. Ditemukan bahwa metode LSTM lebih cocok untuk data dengan granularitas yang lebih tinggi seperti data per jam, namun jika menggunakan data per hari metode seperti SARIMA/ARIMA bisa digunakan untuk meramal indeks kualitas udara di Kemang.

Kata kunci: Peramalan Kualitas Udara, Indeks Kualitas Udara, LSTM, SARIMA, Neural Network

ABSTRACT

The worsening air quality in Indonesia warrants us to review and improve the detection and prediction of air quality to produce a more accurate indicator for air quality management, particularly in health warning systems. Therefore, the purpose of this research is to compare two forecasting methods, namely SARIMA and LSTM, in predicting the air quality index (AQI) in Kemang, South Jakarta. According to prior studies, the LSTM deep learning method outperforms SARIMA in predicting AQI in metropolitan areas. However, the simpler SARIMA method gave a better prediction in coastal areas. Therefore, researchers conducted research to find which method is suitable to be implemented in Kemang, South Jakarta. This research's procedure is started by making sure the time-series data is stationary, then conducting a time-series train-test split, followed by tuning the hyperparameters using either random search or Bayesian method and then comparing the two methods based on their respective RMSE values to draw conclusions. It was found that LSTM is indeed a better method in forecasting data with higher granularity such as hourly data. However, the SARIMA/ARIMA method can still provide reliable results when using daily data to forecast AQI in Kemang, South Jakarta.

Keywords: Forecasting, Air Quality Index, LSTM, SARIMA, Neural Network

