

ABSTRAK

Tantangan utama dalam memanfaatkan model *Long Short-Term Memory* (LSTM) untuk prediksi kualitas udara adalah menemukan setelan *hyperparameter* yang tepat, karena performa model ini sangat sensitif terhadap konfigurasinya. Berangkat dari masalah tersebut, penelitian ini berfokus untuk mencari kombinasi paling optimal untuk kasus prediksi di wilayah DKI Jakarta, di mana kualitas udara merupakan isu yang krusial.

Untuk menemukan setelan terbaik, pendekatan *Grid Search* digunakan untuk "menyisir" berbagai kemungkinan konfigurasi. Metode ini diterapkan untuk menguji 12 skenario berbeda dengan memvariasikan tiga *hyperparameter* kunci: *learning rate*, *batch size*, dan *epoch*. Semua proses pengujian ini menggunakan data historis kualitas udara harian yang dikumpulkan dari lima stasiun pemantau di Jakarta selama periode 2023-2024.

Dari serangkaian pengujian tersebut, berhasil ditemukan satu konfigurasi yang sangat optimal, yakni dengan *learning rate* 0.0001, *batch size* 32, dan 25 *epoch*. Model dengan setelan ini mampu mencapai akurasi validasi 91,3%. Lebih dari itu, performanya terbukti sangat stabil dan bebas *overfitting*, terlihat dari selisihnya yang hanya 0,2% dari akurasi data latih. Model final yang andal ini kemudian menjadi inti dari aplikasi web prediksi yang fungsional, sekaligus membuktikan bahwa *Grid Search* adalah pendekatan yang ampuh untuk optimasi LSTM.

Kata kunci: *long short-term memory*, prediksi, kualitas udara. *grid search*

ABSTRACT

The primary challenge in using the Long Short-Term Memory (LSTM) model for air quality prediction is identifying the correct hyperparameter settings, given that the model's performance is highly sensitive to its configuration. Addressing this, the research focuses on finding the most optimal combination for predictions in the DKI Jakarta area, where air quality is a critical issue.

To find the best settings, the Grid Search approach was used to explore various possible configurations. This method was applied to test 12 different scenarios by varying three key hyperparameters: learning rate, batch size, and epoch. All these tests were conducted using historical daily air quality data collected from five monitoring stations in Jakarta over the 2023-2024 period.

From the series of tests, a highly optimal configuration was identified: a learning rate of 0.0001, a batch size of 32, and 25 epochs. The model with these settings achieved a validation accuracy of 91.3%. More importantly, its performance was proven to be very stable and free of overfitting, evidenced by a mere 0.2% difference from the training accuracy. This reliable final model became the core of a functional web prediction application, ultimately demonstrating that Grid Search is a powerful approach for LSTM optimization.

Keywords: Long Short-Term Memory, Prediction, Air Quality, Grid Search