

ABSTRACT

Air pollution is one of the most pressing environmental issues in many urban areas, especially in rapidly growing cities such as Jakarta, Indonesia. Harmful pollutants like particulate matter (PM), including PM10 and PM2.5, have been linked to severe health effects such as respiratory and cardiovascular diseases, contributing to thousands of premature deaths and substantial economic losses. Increasing industrial activities, vehicle emissions, and coal-powered plants have contributed to Jakarta's worsening air quality, with PM2.5 levels reaching critical levels over the past two decades. This underscores the urgent need for effective air pollution forecasting methods to guide policymakers and evaluate the impact of existing regulations.

This research aims to assess and compare the effectiveness of two advanced forecasting methods, XGBoost and Long Short-Term Memory (LSTM) networks, in predicting particulate matter (PM10) levels in Jakarta. The study leverages historical air quality data from 2010 to 2023 to evaluate the predictive performance of these models, providing a direct comparison between traditional machine learning and deep learning approaches.

The expected output of this research includes a comprehensive analysis of the predictive capabilities of XGBoost and LSTM in handling air quality datasets, as well as insights into PM10 pollution trends in Jakarta. The findings could inform data-driven public policy creation, assisting climate activists and policymakers in devising more effective strategies to mitigate air pollution in the city. Additionally, the study contributes to the academic field by comparing the strengths and limitations of machine learning and deep learning models in environmental forecasting.

Keywords: air pollution forecasting, Jakarta air pollution, particulate matter (PM10), machine learning, deep learning, XGBoost, Long Short-Term Memory (LSTM)