ABSTRACT

Air temperature is an important climatic variable that affects various sectors, including public health, agriculture, infrastructure, and urban activities. Rising daily air temperatures may lead to various climate-related risks; therefore, accurate forecasting methods are needed to support climate risk mitigation. However, air temperature data exhibit nonlinear patterns and strong temporal dependencies, requiring machine learning approaches that can capture these patterns effectively. This study applies the Light Gradient Boosting Machine (LightGBM) algorithm with time series features, namely lag and rolling mean features, to forecast daily air temperature in Semarang City. The study also analyzes the model mechanism through leafwise growth and histogram-based learning, which are used by LightGBM to improve training efficiency. This study used 803 daily climate observations from Semarang City. Model performance was evaluated using RMSE, MAE, and MAPE. The results show that the addition of *time series* features improved the model's predictive performance, achieving an RMSE of 0.5664, an MAE of 0.4395, and a MAPE of 1.57%. These results outperformed the model without *time series* features, which achieved an RMSE of 0.5827, an MAE of 0.4544, and a MAPE of 1.62%.

Keywords: Daily air temperature, climate risk, machine learning, LightGBM, time series, leaf wise growth, histogram-based learning, forecasting