MIDI 2025 Conference

Air Quality Prediction in Urban Recreational Zones Using Machine Learning Techniques

Abstract

Urban recreational zones in rapidly growing and densely populated cities such as Dhaka are increasingly vital for promoting public health and well being. Parks and lakesides are crucial environmental breakers for city inhabitants in this regard, as they present cleaner microenvironments and rehabilitation space for both body and mind.

In this study, a machine learning (ML) model was developed to predict outdoor air quality in selected recreational areas located in Dhaka, Bangladesh. It mainly assesses temperature, humidity and carbon dioxide (CO_2) to predict air quality. The study was conducted during two months of the summer season (12:00 to 17:00) based on smart meter and user questionnaire data. This data was used to train three regression-based ML models: Random Forest (RF), Decision Tree (DT), and Extreme Gradient Boosting (XGBoost). For this assessment, RF demonstrated the best R2 values for the prediction of temperature (0.96), humidity (0.86) and CO_2 (0.69). It also obtained the lowest Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) of 0.83 and 0.91 respectively, in temperature prediction. Conversely, the highest MSE and RMSE values indicated significant errors in the CO_2 predictions. Additionally, RF, DT, and XGBoost were selected for sensitivity analysis due to their effectiveness in modeling non-linear data, with temperature, humidity, and CO_2 levels used as key air quality indicators.

While the study is predictive, it establishes a foundation for future deployment of intelligent air quality control systems capable of real-time monitoring and automated decision-making in urban outdoor environments. The findings can inform adaptive environmental management strategies and smart urban design initiatives.