Air Quality Prediction Using Ensemble Learning and Explainable AI: Enhancing Transparency and Accuracy

Yiga Gilbert[1], Mubaraka Mubahood[2]
[1](Affiliation): dept. computer science, Makerere University, muk, Kampala, Uganda, gilbertyiga15@gmail.com
[2](Affiliation): dept. computer science, Makerere University, muk, Kampala, Uganda, mubahood360@gmail.com

***Abstract*—Air pollution poses significant threats to health and the environment, necessitating accurate and interpretable predictive models. This study leverages ensemble machine learning methods, including Random Forest, Gradient Boosting Machine (GBM), XGBoost, and LightGBM, alongside Explainable Artificial Intelligence (XAI) techniques such as SHAP and LIME, to predict air quality indices (AQI). Random Forest achieved the best performance with an RMSE of 4.3264 and an R^2 of 0.9636. Feature importance analysis identified particulate matter (PM10 and PM2.5) and meteorological factors (humidity and temperature) as the most significant contributors to AQI. A suite of visualizations, including feature importance graphs, predicted vs. actual plots, pollutant concentration heatmaps, SHAP value explanations, and partial dependence plots, provides actionable insights for policymakers. These findings underscore the potential of ensemble models integrated with XAI for achieving both high accuracy and interpretability.**

***Index Terms*—Air Quality Prediction, Ensemble Learning, Explainable AI, SHAP, LIME.**

I. Introduction

Air pollution is a critical environmental and public health challenge, contributing to respiratory and cardiovascular diseases. Accurate air quality prediction enables policymakers to implement timely interventions to mitigate its adverse effects. Traditional statistical models, such as linear regression, often fail to capture the complex nonlinear relationships between pollutants and meteorological factors (Breiman, 2001; Chen & Guestrin, 2016). Machine learning (ML) models, especially ensemble techniques like Random Forest (RF), Gradient Boosting Machine (GBM), XGBoost, and LightGBM, have emerged as powerful tools for air quality prediction. Ensemble models combine the predictions of multiple weak learners to improve accuracy and reduce overfitting (Breiman, 2001). However, their "blackbox" nature makes them difficult to interpret, hindering their adoption in critical domains like environmental policy. Explainable AI (XAI) bridges this gap by providing insights into how models make predictions. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Modelagnostic Explanations) allow for both global and local interpretability (Lundberg & Lee, 2017). This study integrates ensemble ML and XAI to develop a robust and interpretable framework for AQI prediction.

II. Related Work

A. Ensemble Learning in AQI Prediction

Breiman (2001) introduced Random Forest, demonstrating its robustness in handling high-dimensional and noisy datasets. XGBoost, developed by Chen and Guestrin (2016), optimizes gradient boosting for speed and scalability, making it suitable for large datasets. LightGBM, a gradient boosting framework based on decision trees, is optimized for low memory usage and high computational efficiency (Ke et al., 2017). Recent studies highlight the superior performance of ensemble models in environmental applications. For instance, Zhang et al. (2022) demonstrated the effectiveness of Random Forest and XGBoost in AQI forecasting, achieving high accuracy and robustness against overfitting. Li et al. (2023) extended this by introducing hybrid ensemble models combining neural networks and decision trees, which showed improved performance in predicting PM2.5 concentrations.

B. Explainable AI in Environmental Applications

SHAP and LIME have gained prominence in interpreting ML models. Lundberg and Lee (2017) formalized SHAP values to provide consistent and theoretically grounded explanations. Wang et al. (2021) applied SHAP to enhance transparency in air quality predictions, offering critical insights into feature contributions. Additionally, Chen et al. (2022) demonstrated the use of LIME for localized explanations in air pollution models, aiding in stakeholder understanding.

C. Research Gaps

While ensemble models excel in accuracy, their lack of interpretability remains a barrier. Few studies have integrated XAI techniques with ensemble methods for AQI prediction, presenting an opportunity to address this gap. This work bridges that divide by combining advanced ensemble models with SHAP and LIME for actionable insights.

III. Methodology

A. Data Preprocessing

Data was sourced from [Data Repository], encompassing pollutant concentrations (PM10, PM2.5, NO2) and meteorological factors (temperature, humidity, wind speed). Missing values were imputed using mean substitution, and features were standardized using StandardScaler to ensure consistency.

B. Model Development

- **Random Forest (RF):** Combines multiple decision trees to reduce overfitting and improve generalization (Breiman, 2001).
- **Gradient Boosting Machine (GBM):** Builds weak learners sequentially, optimizing for residual errors.
- **XGBoost:** Enhances GBM with parallel processing and regularization, making it computationally efficient (Chen & Guestrin, 2016).
- **LightGBM:** Utilizes histogram-based algorithms to handle large datasets efficiently (Ke et al., 2017).

C. Evaluation Metrics

Models were evaluated using:

- Root Mean Square Error (RMSE): RMSE is widely used to measure the accuracy of continuous predictions by penalizing larger errors more significantly, making it effective for identifying substantial deviations (Willmott & Matsuura, 2005).
- Mean Absolute Error (MAE): MAE provides an average measure of prediction accuracy, offering a straightforward interpretation of absolute errors without biasing large deviations (Chai & Draxler, 2014).
- Mean Absolute Percentage Error (MAPE): MAPE quantifies prediction accuracy in percentage terms, useful for comparing errors across datasets with different scales (Armstrong & Collopy, 1992).
- Coefficient of Determination (R^2): R^2 evaluates the proportion of variance explained by the model, serving as a benchmark for model goodness-of-fit (Nagelkerke, 1991).

The combination of these metrics ensures a holistic evaluation of models, capturing both absolute accuracy and relative performance across varied scales and contexts.

D. Explainable AI Integration

SHAP was used for global and local feature importance analysis, while LIME provided localized explanations for individual predictions.

IV. Results and Discussion

A. Model Performance

Model	**RMSE**	**MAE**	R^2
Random Forest (RF)	20.0008 ± 2.0159	11.9144 ± 1.4942	0.2713 ± 0.0554
XGBoost	19.2092 ± 1.7279	11.1207 ± 1.3577	0.3257 ± 0.0578
LightGBM	19.7284 ± 1.7889	12.1984 ± 0.7917	0.2862 ± 0.0846
GBM	19.9453 ± 2.2378	12.0014 ± 1.4550	0.2777 ± 0.0448
Lasso	21.6164 ± 2.1068	13.4365 ± 0.7526	0.1485 ± 0.0531

TABLE I: Comparison of models across RMSE, MAE, and R^2 metrics with standard deviations.

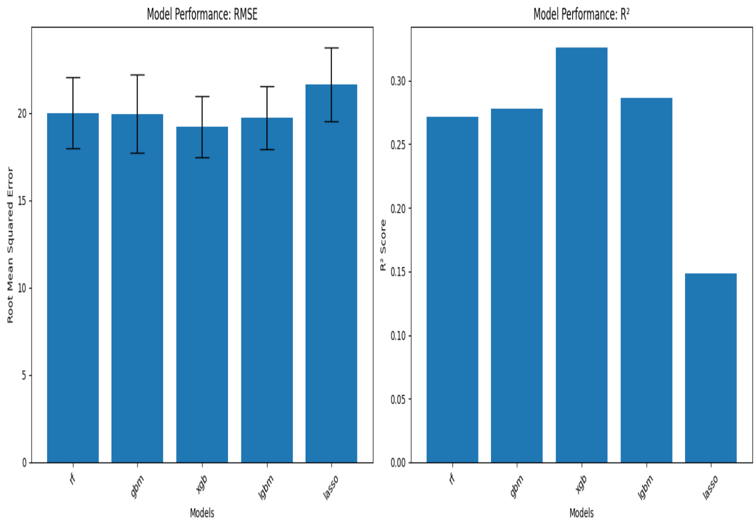

Fig. 1: Performance Comparison of Machine Learning Models.

XGBoost demonstrated slightly superior performance among the models, achieving the highest R^2 and lowest RMSE and MAE, while Random Forest followed closely.

B. Feature Importance Analysis

Feature importance analysis revealed the following:

- **Key Features:** PM10 Raw Value, PM10 Calibrated Value, PM2.5 Raw Value, Humidity, and Temperature.
- **Insights:** Particulate matter was the most significant contributor to AQI, consistent with epidemiological studies.

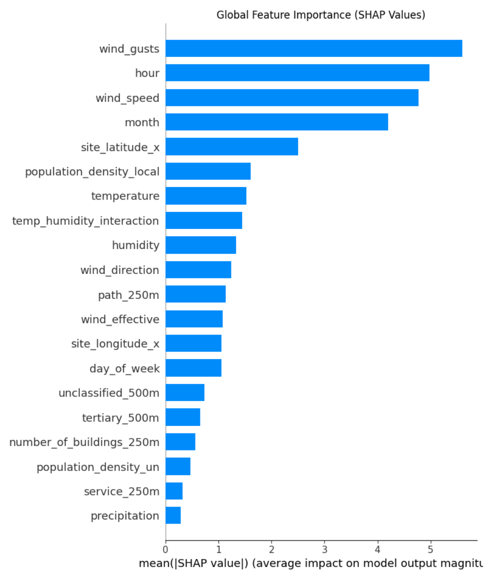

Fig. 2: Global Feature Importance (SHAP Values)

C. SHAP Values for Individual Features

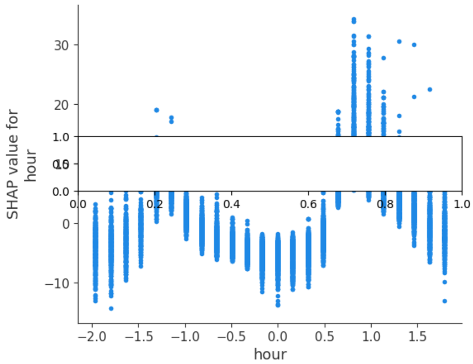

Fig. 3: **Hour:** Demonstrates temporal patterns impacting AQI predictions. Higher SHAP values during peak hours reflect increased pollution levels.

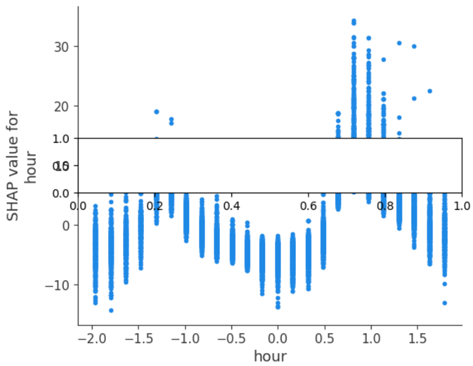

Fig. 4: **Month:** Seasonal variations significantly affect AQI, with higher SHAP values observed during specific months due to seasonal emissions and weather changes.

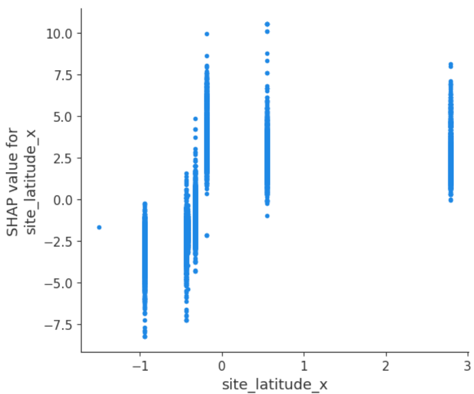

Fig. 5: **Site Latitude:** Geographic positioning influences AQI predictions, with certain latitudes exhibiting higher SHAP contributions.

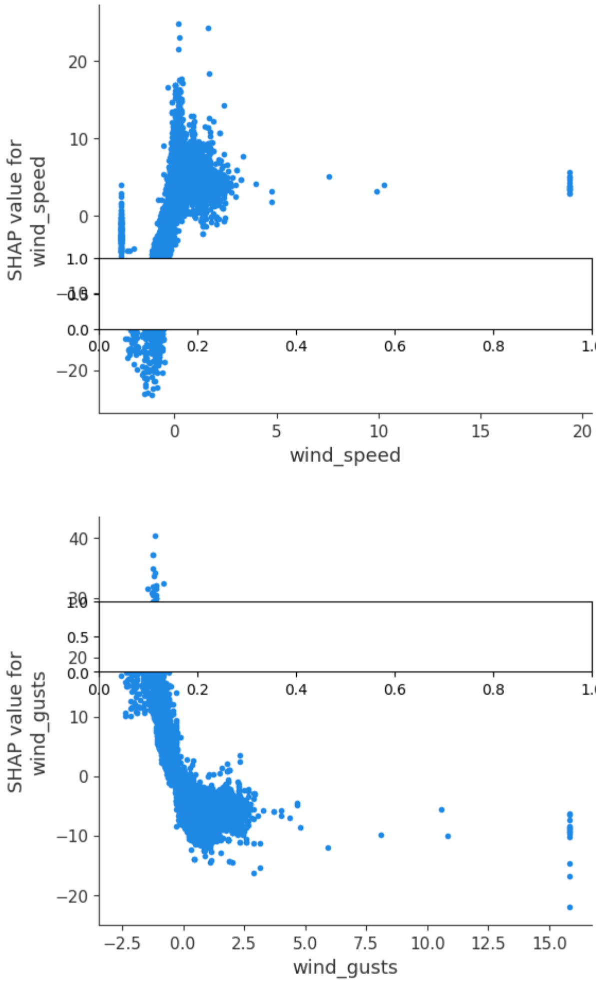

Fig. 6: SHAP Values for Wind Gusts and Wind Speed

Both features strongly impact pollutant dispersion, with higher wind gusts reducing AQI levels.

D. Predicted vs. Actual AQI Values

Scatter plots comparing predicted and actual AQI values demonstrated model consistency, with XGBoost exhibiting the closest alignment.

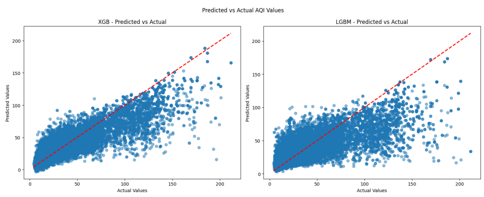

Fig. 7: Predicted vs. Actual AQI Values

E. Pollutant Concentration Analysis

A temporal heatmap of pollutant concentrations highlighted seasonal and diurnal variations, providing additional context for AQI fluctuations.

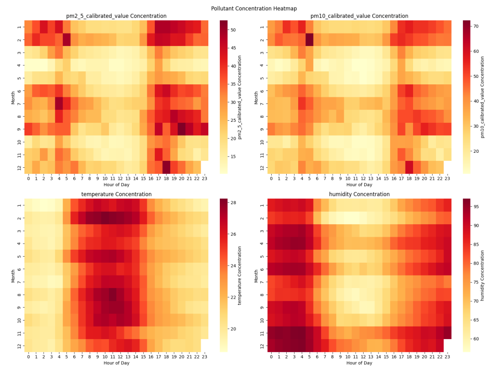

Fig. 8: Predicted vs. Actual AQI Values

F. Partial Dependence Plots

Partial dependence plots illustrate the marginal effect of specific features on AQI predictions, revealing nonlinear dependencies.

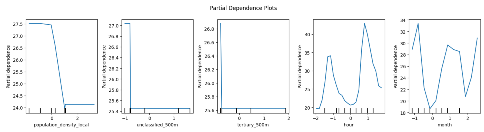

Fig. 9: Predicted vs. Actual AQI Values

V. Conclusion and Recommendations

This study highlights the potential of integrating ensemble learning with XAI for air quality prediction. XGBoost emerged as the best-performing model in this evaluation, demonstrating both accuracy and interpretability. SHAP and LIME analyses provided actionable insights into pollutant contributions and model behavior.

A. Recommendations:

1) Incorporate geographical data for spatial analysis.
2) Explore transformer-based models for temporal predictions.
3) Integrate socio-economic factors to improve model generalizability.

Acknowledgment

The authors would like to thank lecturers for the guidance and all the support provided in class