Design Studies and Intelligence Engineering
L.C. Jain et al. (Eds.)

doi:10.3233/FAIA231451

Explainable Artificial Intelligence (XAI) for Air Quality Assessment

Sayan CHAKRABORTY[1], Bitan MISRA[2], Nilanjan DEY[3]

[1,2,3] *Department of Computer Science and Engineering*
Techno International New Town, West Bengal, India

Abstract- Accurate air quality analysis is essential for comprehending the reasons for and consequences of air pollution, which is a serious environmental concern. Understanding the underlying causes contributing to pollution levels is challenging when using traditional methodologies for air quality analysis since they frequently lack transparency and interpretability. This work examines the integration of XAI with deep learning to enhance air quality prediction. Explainable AI provides a solution by illuminating the ways in which AI models make decisions. It emphasizes the requirement for clear and understandable AI models to win stakeholders' trust and adoption. Utilizing explainable AI makes it feasible to improve the readability and transparency of air quality studies, allowing stakeholders to comprehend and verify the predictions and suggestions made by AI systems. Rough woodland. To classify the data, XGBoost and KNN are used. SHAP and LIME are then applied to discover the major characteristics and variables that affect air quality predictions. These findings can help to improve decision-making and the creation of efficient plans for the management and mitigation of air quality.
Keywords: Explainable artificial intelligence (XAI), Artificial intelligence (AI), Air quality analysis, SHAP, LIME

1. Introduction

The practice of predicting the amount of air pollution in a certain place or region at a specific future time is known as air quality prediction. It entails analyzing and interpreting data pertaining to air pollutants, meteorological conditions, pollutant emission sources, and other pertinent elements using various models and approaches. The process of predicting air quality often entails gathering historical or real-time information on pollutant concentrations, weather factors (such as temperature, humidity, wind speed, and direction), emissions inventories, and other pertinent variables. To predict future air quality levels, these data are then examined using mathematical models and statistical techniques [1-3]. The evaluation and administration of the effects of air pollution on public health, the ecosystem, and different industries, such as agriculture, transportation, and manufacturing, depend heavily on-air quality forecasts. Authorities and politicians can protect the public's health by making decisions based on accurate air quality predictions, putting suitable mitigation measures in place, and issuing alerts or warnings. Depending on the requirements and capabilities of the system, air quality prediction systems can offer short-term forecasts (a few hours to a few days ahead) or longer-term projections (weeks to months ahead). To illustrate these projections to the public and convey the degree of air pollution and related health hazards, color-coded maps or air quality indices can be used. It is crucial to remember that the geography, local infrastructure, and specific activities taking place can all affect the causes and severity of air pollution [2, 3].

Depending on the situation and the specific classification scheme being employed, the major air contaminants can change. However, several contaminants are widely acknowledged as being significant air pollution culprits. These "criteria air pollutants" have been determined to be threatening for both public health and the ecosystem. Particulate matter (PM) is used to describe the suspension of liquid droplets and solid particles in air. There are two size fractions that can be distinguished: PM 10 (particles with diameters of 10 micrometers or less) and PM 2.5 (particles with diameters of 2.5 micrometers or less). PM can come from a number of sources, including combustion, industrial emissions, automobile exhaust, and environmental dust and pollen. Sulfur-containing fossil fuels in particular release sulfur dioxide as a gas when they are burned. In the presence of sunshine, nitrogen oxides (NOx) and volatile organic compounds (VOCs) mix to form ground-level ozone. Vehicle emissions, industrial activities, and chemical reactions involving VOCs are the main sources. Specifically, in cars and power plants, the burning of fossil fuels produces nitrogen dioxide [2-4]. The negative health impacts of exposure to air pollution can be both temporary and permanent. Vulnerable groups that are particularly at risk include children, elderly individuals, and persons with preexisting respiratory or cardiovascular problems.

This research analyses the most significant pollutant utilizing explainable artificial intelligence (XAI) [5, 6] to overcome the aforementioned issues. To combat the black-box aspect of artificial intelligence (AI) models, XAI strategies attempt to offer insights into the decision-making process and the underlying variables that affect predictions. These discoveries can be used to create prediction models that are more precise, to pinpoint the variables that have the greatest influence on air quality and to guide the processes used to decide on air pollution reduction tactics. It is possible to obtain deep learning architectures with improved credibility and accuracy for predicting air quality. As a result, its practical use is improved. This paper's contributions can be summarized as follows.

1. To predict air quality, the random forest, XGBoost, and K-nearest neighbor (KNN) models are employed.

2. Using the SHapley Additive exPlanations (SHAP) and local interpretable model-agnostic explanations (LIME) techniques, the impact of air pollutant parameters on air quality prediction is highlighted, which helps to increase the precision of air quality prediction.

The article's remaining sections are structured as follows. In Section 2, a discussion of the relevant literature is offered. Section 3 provides a description of the dataset and methodology applied in this study. In Section 4, the outcome of the air quality prediction is analyzed, and the SHAP and LIME techniques are used to assess the explainability of the deep learning architectures. Section 5 contains the article's conclusion.

2. Literature review

To improve the understanding of AI models employed for air quality analysis, researchers have investigated various machine learning interpretation techniques. To identify the variables affecting air quality predictions, methodologies such as feature significance analysis, rule-based designs, and local explanatory techniques are being used. Deep learning approaches are employed for a variety of big data forecasting issues. Long short-term memory (LSTM) networks are used in a deep learning model for predicting future air quality values in smart cities [4]. To calculate the air quality index, a thorough examination [7] of the effectiveness of various machine learning models, including linear regression, decision tree, random forest, artificial neural network, and support vector machine models, is conducted. Article [8] discusses a thorough examination of studies that have been done on air pollution forecasting using machine learning algorithms that use sensor data. One can draw the conclusion from this study [8] that the authors currently use complex and advanced methodologies rather than straightforward machine learning methods. In addition, the primary prediction goal was PM 2.5, and China accounted for the majority of the case studies. In the Jing-Jin-Ji region, which has the worst air pollution in China, a neural network with a temporal sliding long short-term memory extended model [9] is used to forecast the following 24 hours' average PM2.5 concentration. A multidisciplinary approach is used in the research on explainable AI for accurate air quality analysis, combining knowledge from areas including air quality research, machine learning, and meteorology [10]. Article [11] presents a visual analytics method to assist specialists in validating and confirming the learning of the ML model with their domain knowledge. The system consists of numerous coordinated views that may be used to display the impact of input characteristics at different degrees of aggregation in both the temporal and geographic dimensions. Additionally, it presents an analysis of the efficacy of ML and traditional architectures in terms of accuracy and geographic map effectiveness, as well as an animation of the raw wind trajectory data for the input time. User engagement can help with exploration, while data visualizations can offer overviews of various parts of the data [12]. Article [12] presents a visual analytics platform for building unique interactive visual analytics dashboards with clearly defined components that are connected to one another. In article [13], scientists provide a new architecture that incorporates a highly stacked 1-dimensional CNN with a residual connection and attention mechanism that provides state-of-the-art results in PM 2.5 and PM 10 prediction through extensive tests using Seoul air pollution data and public benchmarks. From the literature review, it is identified that different air pollutants have different amounts of effect in different areas. Hence, it is important to identify the pollutant that contributes the most to air pollution in that area so that necessary steps can be taken. Our study aims to bridge this gap by applying explainable AI to identify the pollutant that has the highest contribution in determining air quality.

3. Proposed Method

The current framework analyzed the dataset and performed preprocessing to clean the data, as there were many null and redundant data. Following preprocessing, the data are further analyzed using three popular supervised learning models: random forest, KNN and XGBoost. The key reason behind choosing these models is their distinct advantages. High prediction accuracy and scalability are provided by XGBoost, nonlinear decision boundaries are a strength of KNN's classification jobs, and high-dimensional data are successfully handled by random forest's robustness against overfitting. The dataset is split into an 80-10-10 ratio in training-testing-validation data. Then, each of the machine learning architectures is applied to the data.

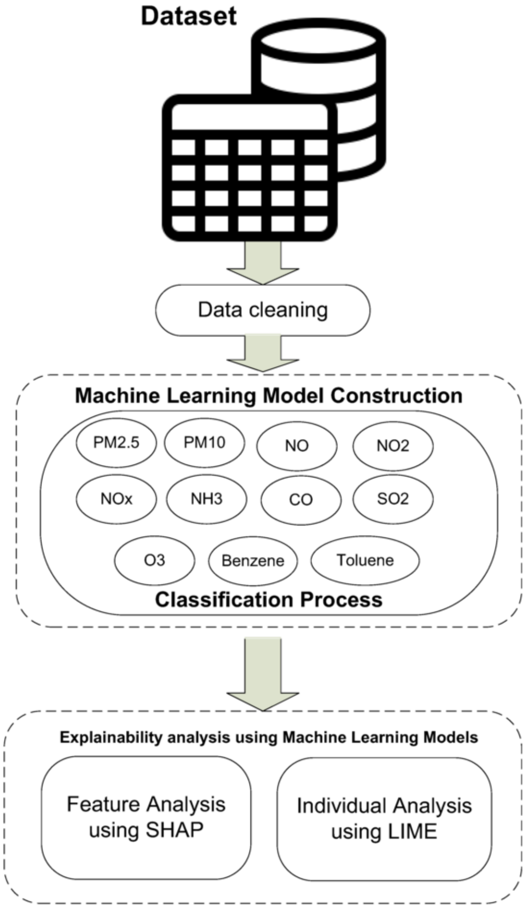

Figure 1. Architectural diagram of the proposed method

The proposed approach's block diagram, as detailed in the section, is shown in Figure 1. Two popular methods for analyzing machine learning architectures are LIME and SHAP, as depicted in the diagram. They are used with machine learning models to offer an understanding of the predictions made by the model and the significance of various characteristics. Tools for model interpretation and explainability include LIME and SHAP. They can assist in debugging models, gaining confidence in forecasts, understanding how machine learning models make decisions, and spotting bias or unfairness in predictions. These techniques are frequently applied in a variety of fields, such as healthcare, finance, and natural language processing, where the capacity to comprehend data is essential for making decisions and adhering to regulations. Machine learning models are given interpretability and explainability using LIME and SHAP. They aid in understanding the reasons behind the predictions made by models as well as the features that influence those forecasts the most. Stakeholders may enhance the transparency and accountability of machine learning models by employing these strategies to acquire insights into how they make decisions.

4. Results and Discussion

In this work, machine learning frameworks such as random forest, KNN and XGBoost were applied using the Python environment in the Windows 10 operating system. The work was tested on an Intel i3 8th generation-based computer system. The dataset (https://www.kaggle.com/datasets/rohanrao/air-quality-data-in-india?resource=download) contains data on the AQI (air quality index) at hourly and daily rates from a large number of sites dispersed throughout various Indian cities. For the current framework, data from 2015 to 2020 were examined. The metric used to assess classification accuracy is accuracy. Accuracy is typically a key component of accurate forecasting. The aspect of successful forecasts that characterizes good predictions is precision. The portion of the total number of already retrieved real-world samples is called recall. The F1 score is a suitable metric for profiting from the occasion. For viewing the F1 score, an analogy between precision and recall is discovered, although there is still a jagged class propagation. The obtained results of the classification models are presented in Table 1.

Table 1. Comparative analysis of the results using the random forest, KNN and XGBoost models

Methods	Precision	Recall	F1-score	Accuracy
Random Forest	0.81	0.81	0.81	0.81
KNN	0.8	0.8	0.8	0.80
XGBoost	0.84	0.83	0.84	0.83

Table 1 shows that the XGBoost-based framework outperformed the random forest and KNN-based classification models in terms of accuracy. High performance and accuracy are hallmarks of XGBoost. To construct a strong learner, XGBoost combines the predictions of several weak learners (decision trees). It builds the ensemble of trees iteratively while using a gradient descent optimization algorithm to optimize a certain objective function. Random Forest is a similar ensemble learning technique that aggregates the predictions of various decision trees. However, unlike XGBoost, random forest creates every decision tree on its own. KNN is a nonparametric method that categorizes fresh data points in accordance with the training data's k nearest neighbors' dominant class. KNN can function well when the decision boundary is nonlinear and is straightforward to construct. Although XGBoost is known for its great prediction performance and accuracy, tuning and parameter optimization may be more difficult for XGBoost than for random forest and KNN. Conversely, random forest is frequently chosen when interpretability and feature relevance are crucial.

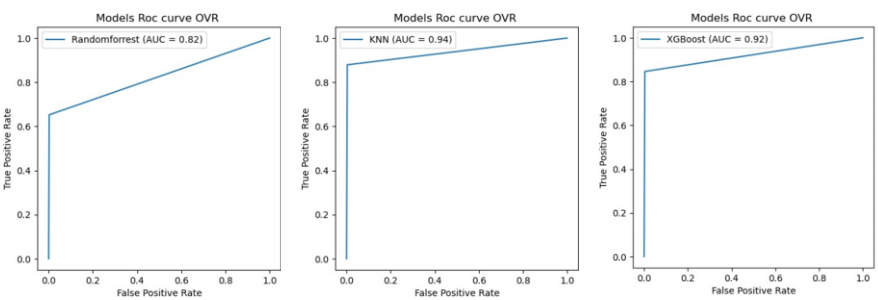

Figure 2. RoC Curves of the Random Forest-, KNN- and XGBoost-based Frameworks

The RoC curves for the random forest-, KNN-, and XGBoost-driven architectures are shown in Figure 2. The RoC and AUC score provide an indication of the model's efficacy. The model performs better at differentiating between positive and negative classifications with a higher AUC. The performance of two or more diagnostic tests was compared using the RoC curve to evaluate a test's overall diagnostic performance. It is also used to establish the proper cutoff value for determining whether a possibility even exists. The KNN algorithm, bases its predictions on the majority class of its nearest neighbors. KNN was able to acquire a higher AUC score since the dataset had a clear structure or clusters. KNN is a distance-based algorithm that depends on feature value similarity. Features in the dataset were quite pertinent. As a result, KNN has excelled and received a higher AUC rating. In further processing, the XGBoost-based framework was chosen due to its higher accuracy in classification compared to KNN- and random forest-based models.

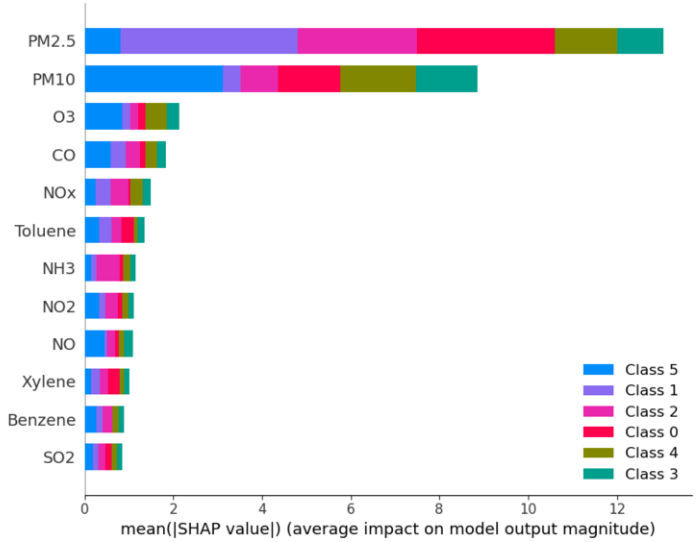

Figure 3. XGBoost-based SHAP summary plot

Figure 3 depicts a SHAP summary plot that gives a machine learning model's feature importance a visual representation. The top features are displayed in order of relevance, along with the matching SHAP values, which demonstrate the influence each feature has on predictions. As Fig. 3 shows, PM 10 and PM 2.5 are the primary features in determining air quality in the XGBoost-based framework. Different size fractions of airborne particulate matter are referred to as PM10 and PM2.5. Monitoring these particles aids in obtaining the extent of air pollution and directing environmental strategies. They are crucial components of air quality evaluation because when inhaled, they may induce respiratory and cardiovascular problems.

The summary plot of SHAP values for the label "0" is shown in Figure 4. The feature names are shown on the Y-axis from top to bottom in order of importance. The SHAP value, which denotes the degree of change in log odds, is represented on the X-axis. To show the value of the associated property, each point on the graph is colored, with red suggesting high values and blue denoting low values. Each point represents a row of data from the original dataset.

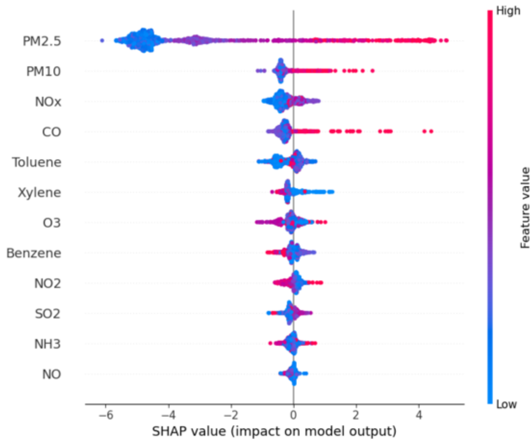

Figure 4. XGBoost-based SHAP summary plot for label 0.

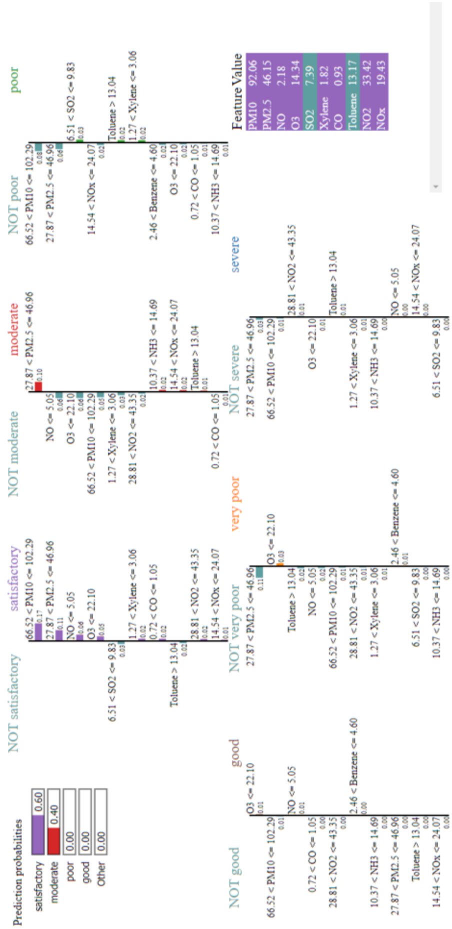

Figure 5. XGBoost-based LIME

The output of LIME is a list of justifications showing how each attribute contributed to the prediction of a data sample. As a result, it is easy to determine which feature changes will have the most impact on the prediction, which provides local interpretation. An accessible model produces an explanation by localizing the underlying model. Decision trees and linear models with extensive regularization are examples of comprehensible models. The interpretable models, which are trained on few alterations of the original instance, should only offer a good local approximation. An interpretable paradigm produces an explanation by localizing the underlying model. Decision trees and linear models with extensive regularization are examples of interpretable models. Fig. 5 shows the LIME-based interpretability of the XGBoost-based model. Different decisions, such as 'severe', 'very poor', 'poor', 'moderate', 'satisfactory', and 'good', were analyzed using LIME, taking different features into consideration.

5. Conclusions

Researchers from all across the world are working to resolve the problem of air pollution on a worldwide scale. Estimating pollution levels is inherently challenging due to the data's instability, dynamicity, and variability in both place and time. In this study, we use the SHAP and LIME explainable deep learning methods to demonstrate how air contaminants affect air quality. KNN, XGBoost, and random forest are the three machine learning models used, and XGBoost provides the highest accuracy. PM 10 and PM 2.5 are shown to be the most significant factors in determining air quality in the XGBoost-based framework using XAI models. However, the dataset used in this study only contains AQI data for 5 years (2015-2020). Additional features of XAI model incorporation can improve the quality of the result. Future research will concentrate on analyzing the various reliable ML forecasting techniques. This makes it easier to analyze and comprehend in-depth the deep learning models used to predict air quality, and it increases their reliability for use in accurate air quality prediction.

References

[1] Kang, G. K., Gao, J. Z., Chiao, S., Lu, S., & Xie, G. (2018). Air quality prediction: Big data and machine learning approaches. Int. J. Environ. Sci. Dev, 9(1), 8-16.

[2] Yi, X., Zhang, J., Wang, Z., Li, T., & Zheng, Y. (2018, July). Deep distributed fusion network for air quality prediction. In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 965-973).

[3] Zhu, D., Cai, C., Yang, T., & Zhou, X. (2018). A machine learning approach for air quality prediction: Model regularization and optimization. Big data and cognitive computing, 2(1), 5.

[4] Kök, İ., Şimşek, M. U., & Özdemir, S. (2017, December). A deep learning model for air quality prediction in smart cities. In 2017 IEEE international conference on big data (big data) (pp. 1983-1990). IEEE.

[5] Kamal, M. S., Northcote, A., Chowdhury, L., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2021). Alzheimer’s patient analysis using image and gene expression data and explainable-AI to present associated genes. IEEE Transactions on Instrumentation and Measurement, 70, 1-7.

[6] Kamal, M. S., Dey, N., Chowdhury, L., Hasan, S. I., & Santosh, K. C. (2022). Explainable AI for glaucoma prediction analysis to understand risk factors in treatment planning. IEEE Transactions on Instrumentation and Measurement, 71, 1-9.

[7] Madan, T., Sagar, S., & Virmani, D. (2020, December). Air quality prediction using machine learning algorithms–a review. In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) (pp. 140-145). IEEE.

[8] Iskandaryan, D., Ramos, F., &Trilles, S. (2020). Air quality prediction in smart cities using machine learning technologies based on sensor data: a review. Applied Sciences, 10(7), 2401.

[9] Mao, W., Wang, W., Jiao, L., Zhao, S., & Liu, A. (2021). Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. Sustainable Cities and Society, 65, 102567.

[10] Yang, Y., Mei, G., & Izzo, S. (2022). Revealing influence of meteorological conditions on air quality prediction using explainable deep learning. IEEE Access, 10, 50755-50773.

[11] Palaniyappan Velumani, R., Xia, M., Han, J., Wang, C., LAU, A. K., & Qu, H. (2022, March). AQX: Explaining Air Quality Forecast for Verifying Domain Knowledge using Feature Importance Visualization. In 27th International Conference on Intelligent User Interfaces (pp. 720-733).

[12] Kalamaras, I., Xygonakis, I., Glykos, K., Akselsen, S., Munch-Ellingsen, A., Nguyen, H. T., ... &Tzovaras, D. (2019, November). Visual analytics for exploring air quality data in an AI-enhanced IoT environment. In Proceedings of the 11th international conference on management of Digital EcoSystems (pp. 103-110).

[13] Choi, J., Kim, J., & Jung, K. (2021, January). Air quality prediction with 1-dimensional convolution and attention on multimodal features. In 2021 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 196-202). IEEE.