An Explainable Machine Learning Framework for Air Quality Classification

Muskanpreet
Department of Computer Science & Engineering, Swami Vivekanand Institute of Engineering and Technology, Ramnagar, Punjab, India

***Abstract*—Air pollution caused by pollutants such as PM2.5, PM10, NO_2, and CO has become a major environmental concern. Accurate air quality classification is important for environmental monitoring and public health. This study proposes an explainable machine learning framework for air quality classification using a Decision Tree classifier.**

A global air quality dataset was used to compute Air Quality Index (AQI) values, which were categorized into Good, Moderate, and Poor classes. Data preprocessing, correlation analysis, and feature importance evaluation were performed to improve model performance and interpretability. The proposed model was also compared with a Random Forest classifier.

Experimental results show that the Decision Tree model achieved 91.5% accuracy, outperforming the Random Forest model with 89.3% accuracy. In addition, the Decision Tree provided transparent IF–THEN decision rules for classification.

The main contribution of this work is a comparative study of interpretable and ensemble machine learning models for AQI classification with emphasis on explainability.

***Index Terms*—Air Quality Classification, Explainable Machine Learning, Decision Tree, Random Forest, Air Quality Index**

I. INTRODUCTION

Air pollution is one of the most serious environmental problems affecting human health and ecological systems worldwide. Pollutants such as PM2.5, PM10, nitrogen dioxide (NO_2), and carbon monoxide (CO) contribute significantly to poor air quality and increase the risk of respiratory and cardiovascular diseases [1]. Therefore, accurate monitoring and classification of air quality are essential for environmental management and public awareness.

Traditional air quality prediction methods such as regression and ARIMA models have been widely used; however, these approaches often fail to capture complex nonlinear relationships in environmental data [2]. With the advancement of machine learning, algorithms such as Support Vector Machines, Random Forest, and Neural Networks have been applied for AQI prediction and classification [3]. Although these models achieve high accuracy, many of them operate as black-box systems and provide limited interpretability [4].

Recently, Explainable Artificial Intelligence (XAI) has gained attention for improving transparency in machine learning systems. Explainable models provide understandable reasoning behind predictions and help users interpret the influence of input features on classification outcomes [5]. Among various machine learning techniques, Decision Trees are considered highly interpretable because they generate clear IF–THEN decision rules.

This research proposes an explainable machine learning framework for air quality classification using a Decision Tree classifier. A weighted AQI was computed using pollutant concentrations and categorized into Good, Moderate, and Poor classes. To evaluate the effectiveness of the proposed framework, a comparative analysis with Random Forest classification was also performed.

The proposed Decision Tree model achieved an accuracy of 91.5%, outperforming the Random Forest model while also providing better interpretability. The study demonstrates that explainable machine learning approaches can effectively balance prediction performance and transparency for environmental monitoring applications.

II. LITERATURE REVIEW

Numerous studies have explored air quality prediction using machine learning and statistical approaches. Soundari et al. applied machine learning techniques for air quality analysis and achieved promising prediction results using pollutant datasets [1].

Recent studies have also focused on hybrid and ensemble learning methods for AQI prediction. Bhusan et al. proposed a hybrid framework combining machine learning and deep learning techniques to improve prediction performance [2].

Traditional statistical approaches such as regression analysis and ARIMA models have also been widely used for AQI forecasting. However, these methods are less effective in capturing nonlinear relationships in environmental data [3].

With the advancement of machine learning, models such as Decision Trees, Support Vector Machines, and Random Forest classifiers have been applied for AQI prediction and classification. Liu et al. demonstrated that machine learning algorithms improve prediction accuracy compared to conventional approaches [4].

Castelli et al. proposed a machine learning framework for air quality prediction and highlighted the effectiveness of data-driven models for large environmental datasets [5]. Similarly, Halsana explored supervised learning algorithms for AQI prediction and demonstrated their effectiveness in environmental monitoring applications [7].

Although many existing studies achieve strong prediction performance, most models focus mainly on accuracy and provide limited interpretability. Complex ensemble and deep learning techniques often behave as black-box systems, making it difficult to understand how predictions are generated. Therefore, there is a need for explainable machine learning frameworks that provide both reliable performance and transparent decision-making for air quality classification.

III. METHODOLOGY

3.1. Overview

The methodology proposed in this study focuses on developing an explainable machine learning framework for air quality classification. The overall process includes dataset collection, preprocessing, AQI computation, feature selection, model training, testing, and performance evaluation. Two machine learning models, Decision Tree and Random Forest, were implemented and compared to analyze both classification performance and interpretability.

3.2. Dataset Description

The dataset used in this research is a Global Air Quality dataset collected from Kaggle. It contains air pollution measurements collected from different regions worldwide. The dataset includes important pollutant and environmental parameters such as:

- PM2.5 (Particulate Matter ≤ 2.5 µm)
- PM10 (Particulate Matter ≤ 10 µm)
- NO_2 (Nitrogen Dioxide)
- CO (Carbon Monoxide)
- SO_2 (Sulfur Dioxide)
- Temperature
- Humidity
- Wind Speed

Since the dataset did not contain predefined Air Quality Index (AQI) values, AQI values were calculated using pollutant concentrations and later used for classification.

3.3 Data Preprocessing

Data preprocessing was performed to improve dataset quality and ensure reliable model performance. Environmental datasets often contain missing values, duplicate entries, and inconsistent records that can negatively affect classification accuracy.

The following preprocessing steps were applied:

- Removal of missing and null values
- Elimination of duplicate records
- Cleaning of invalid or inconsistent data entries
- Selection of relevant pollutant and environmental features

Feature selection was carried out to identify variables that significantly influence AQI classification. Pollutants such as PM2.5 and PM10 were found to contribute strongly to air quality prediction.

3.4 Correlation Analysis

Correlation analysis was performed to study the relationship between pollutant concentrations and environmental variables. A correlation heatmap was generated to identify highly related features affecting AQI classification.

The analysis showed that PM2.5 and PM10 had strong positive correlation with AQI values, indicating their major contribution to air pollution levels.

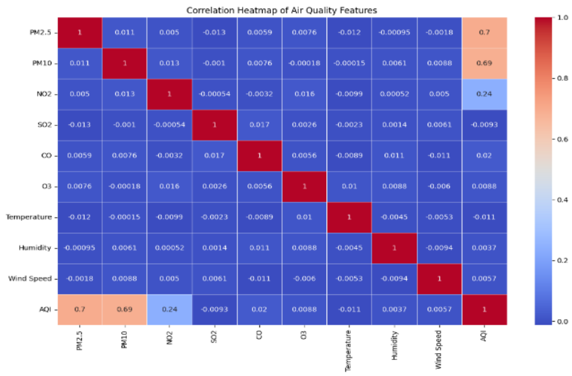

Figure 1: Correlation Heatmap of Air Quality Features

3.5 AQI Computation

As the dataset did not contain AQI values, a weighted AQI score was computed using pollutant concentrations. The AQI was calculated using the following formula:

$$AQI = 0.4 \times PM2.5 + 0.3 \times PM10 + 0.2 \times NO_2 + 0.1 \times CO$$

Higher weights were assigned to PM2.5 and PM10 because of their significant impact on human health and environmental quality.

3.6 AQI Classification

The computed AQI values were categorized into three air quality classes:

- Good: AQI ≤ 50
- Moderate: 50 < AQI ≤ 100
- Poor: AQI > 100

This categorization simplified the classification process and improved the interpretability of model outputs.

3.7 Data Splitting

The dataset was divided into training and testing sets using an 80:20 ratio.

- Training Set (80%): Used for model training
- Testing Set (20%): Used for model evaluation

This approach ensured that the models were tested on unseen data to evaluate generalization performance.

3.8 Model Development

Two machine learning models were implemented in this study:

A. Decision Tree Classifier

The Decision Tree classifier was selected as the primary model because of its explainability and ability to generate transparent decision rules. The model classifies AQI categories by recursively splitting the dataset based on feature thresholds.

The following parameters were used:

- Maximum depth = 5
- Minimum samples split = 10

B. Random Forest Classifier

A Random Forest classifier was implemented for comparative analysis. Random Forest combines multiple decision trees to improve prediction performance and reduce overfitting. The comparison helped evaluate the balance between classification accuracy and interpretability.

3.9 Mathematical Background

A. Entropy

$$\text{Entropy}(S) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

Entropy measures the impurity or randomness in a dataset. Higher entropy indicates greater uncertainty in classification.

B. Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Accuracy represents the proportion of correctly classified instances.

C. Precision and Recall

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Precision measures the correctness of positive predictions, while recall measures the ability of the model to identify actual positive instances.

3.10 Model Training

The models were trained using selected pollutant and environmental features including:

- PM2.5
- PM10
- NO_2
- CO
- Temperature
- Humidity
- Wind Speed

The algorithms learned classification patterns by analyzing feature importance and pollutant thresholds from the training dataset.

3.11 Model Testing
The trained models were tested using unseen testing data. Predictions generated by the models were compared with actual AQI categories to evaluate classification performance.

3.12 Performance Evaluation
The performance of both models was evaluated using standard classification metrics:

- Accuracy
- Precision
- Recall
- F1-score

Confusion matrices and classification reports were also generated to analyze prediction performance and classification errors.

3.13 Interpretability
One of the major advantages of the Decision Tree model is its interpretability. The model generates understandable IF–THEN decision rules that explain how pollutant levels influence AQI classification.
For example:

- IF PM2.5 value is high → Poor air quality
- IF PM10 and NO_2 values are low → Good air quality

This transparency makes the proposed framework more suitable for environmental monitoring applications where prediction reasoning is important.

3.14 Software Tools
The implementation was carried out using Python and the following libraries:

- Scikit-learn (Machine Learning Models)
- Pandas (Data Handling)
- NumPy (Numerical Computation)
- Matplotlib (Visualization)
- Seaborn (Graphical Analysis)

IV. RESULTS

4.1 Model Performance
The proposed framework was evaluated using Decision Tree and Random Forest classifiers. Experimental results show that the Decision Tree model achieved an accuracy of 91.5%, while the Random Forest model achieved 89.3% accuracy.
The higher accuracy of the Decision Tree model indicates that explainable machine learning approaches can provide strong classification performance without relying on complex black-box techniques.

4.2 Confusion Matrix Analysis
The confusion matrices shown in Figure 2 and Figure 3 illustrate the classification performance of both models on the testing dataset. Most AQI samples were correctly classified, with minor misclassification occurring between Moderate and Poor categories.

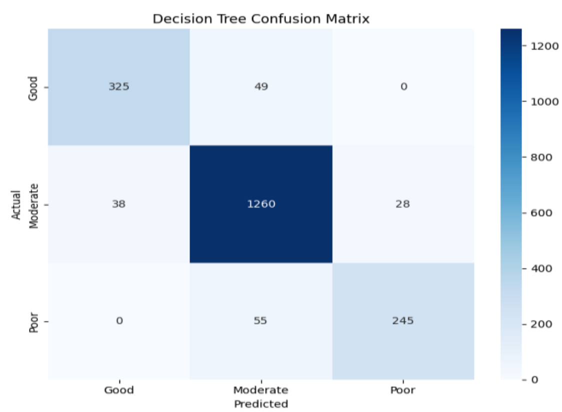

Figure 2: Confusion Matrix of Decision Tree Classifier

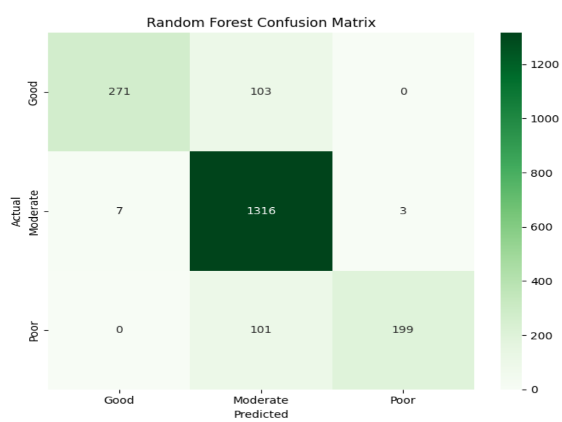

Figure 3: Confusion Matrix of Random Forest Classifier

4.3 Feature Importance Analysis
Feature importance analysis was performed to identify the most influential pollutants affecting AQI classification. As shown in Figure 4, PM2.5 and PM10

contributed most significantly to the classification process.

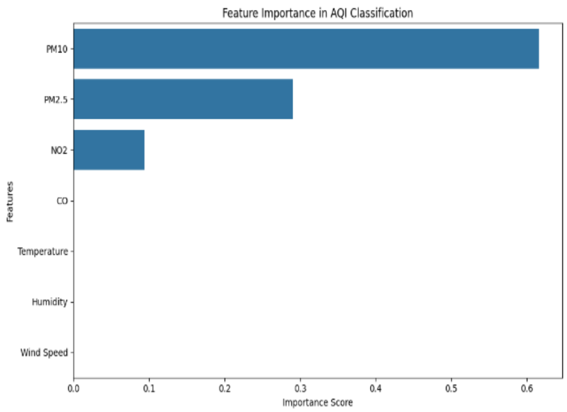

Figure 4: Feature Importance Analysis of AQI Parameters

4.4 Model Comparison

Figure 5 presents the accuracy comparison between Decision Tree and Random Forest classifiers. The Decision Tree model outperformed the Random Forest classifier while also providing better interpretability.

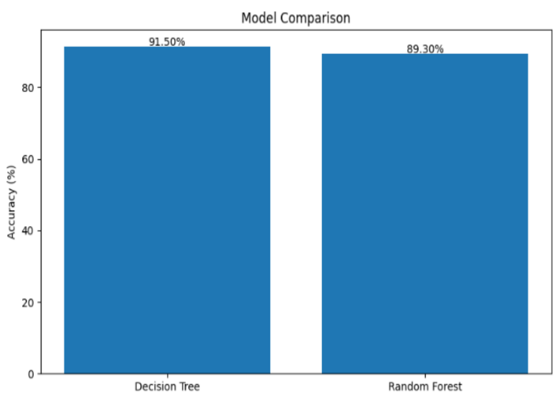

Figure 5: Accuracy Comparison Between Models

Table 1: Performance Comparison of Models

Model	Accuracy	Interpretability
Decision Tree	91.5%	High
Random Forest	89.3%	Moderate

The comparative analysis indicates that the Decision Tree classifier achieved better classification accuracy while also providing higher interpretability compared to the Random Forest model.

4.5 Classification Report Analysis

Classification report heatmaps were generated to evaluate precision, recall, and F1-score for each AQI category. The Decision Tree model demonstrated balanced classification performance across all classes compared to the Random Forest model.

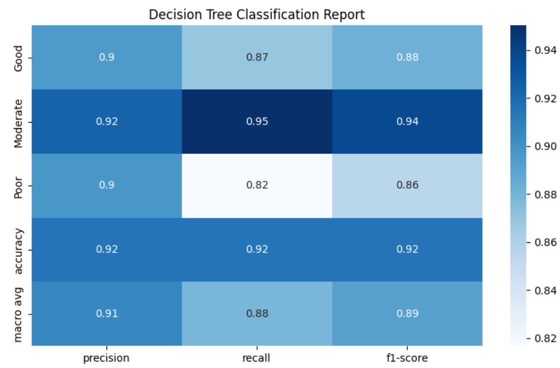

Figure 6: Classification Report of Decision Tree Model

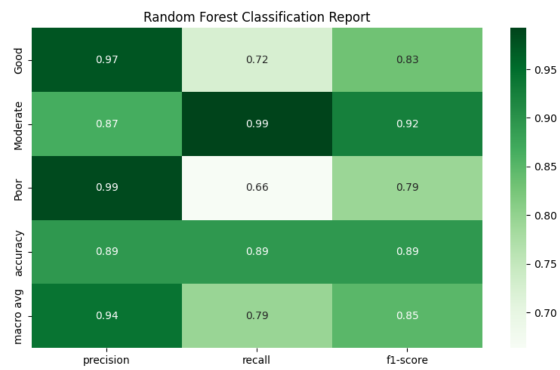

Figure 7: Classification Report of Random Forest Model

4.6 Comparison with Existing Research

The proposed framework was also compared with existing research studies related to air quality prediction and classification. Most previous studies focused mainly on improving prediction accuracy using machine learning, hybrid learning, and deep learning approaches. However, many of these techniques provided limited interpretability and behaved as black-box systems.

In contrast, the proposed framework focuses on both classification performance and explainability. The Decision Tree classifier achieved higher classification accuracy while also generating transparent IF–THEN decision rules for AQI classification. Table 2 presents the comparison between the proposed framework and existing research studies.

Table 2: Comparison with Existing Research Studies

Ref.	Technique	Interpretability
Soundari et al. [1]	ML-based AQI Prediction	Limited
Bhusan et al. [2]	Hybrid ML & DL	Low
Liu et al. [4]	ML Algorithms	Moderate
Halsana [7]	Supervised Learning	Moderate
Proposed	DT & RF Classification	High

The comparative analysis shows that the proposed explainable framework provides a better balance between classification performance and transparency. This makes the framework more suitable for practical environmental monitoring and decision-making applications.

V. EVALUATION

The performance of the proposed framework was evaluated using Decision Tree and Random Forest classifiers with standard classification metrics including accuracy, precision, recall, and F1-score. The comparison was performed to analyze both prediction performance and model interpretability.

The Decision Tree classifier achieved an accuracy of 91.5%, whereas the Random Forest classifier achieved 89.3% accuracy. This indicates that the Decision Tree model performed better on the selected air quality dataset.

- Accuracy represents the overall correctness of classification results. The higher accuracy of the Decision Tree model shows that it classified AQI categories more effectively compared to the Random Forest model.
- Precision measures how reliably the models predicted AQI categories. The Decision Tree classifier produced slightly higher precision values, indicating fewer incorrect classifications than the Random Forest model.
- Recall evaluates the ability of the models to correctly identify actual AQI classes. Both models performed well for the Moderate category; however, the Decision Tree classifier showed better recall for Good and Poor air quality classes.
- F1-score provides a balance between precision and recall. The Decision Tree model achieved more balanced F1-score values across all AQI categories, demonstrating stable classification performance.

The confusion matrices and classification report heatmaps further confirmed that the Decision Tree classifier generated fewer misclassifications compared to the Random Forest model. Most classification errors occurred between Moderate and Poor categories because of overlapping pollutant concentration ranges. Feature importance analysis revealed that PM2.5 and PM10 were the most influential parameters in both models. However, the Decision Tree model provided clearer understanding of feature influence through transparent IF–THEN decision rules.

Although Random Forest is generally considered a powerful ensemble learning technique, it behaves more like a black-box model and provides limited interpretability. In contrast, the Decision Tree classifier offers simple and understandable decision paths, making the prediction process easier to interpret and validate.

5.1 Limitations

Despite achieving good classification performance, the proposed framework has some limitations:

- Limited dataset size may affect model generalization.
- Slight overlap between Moderate and Poor AQI categories may cause minor misclassification.
- Decision Tree models are sensitive to feature threshold values.
- Random Forest provides lower interpretability compared to Decision Tree classification.
- Environmental conditions may vary across different geographical regions.
- External environmental factors such as rainfall and atmospheric pressure were not fully considered.
- AQI values were computed using a weighted formula rather than official government AQI standards.

VI. CONCLUSION

This research presented an explainable machine learning framework for air quality classification using Decision Tree and Random Forest classifiers. The study focused on balancing classification performance with model interpretability for environmental monitoring applications.

A weighted Air Quality Index (AQI) was computed using pollutant concentrations such as PM2.5, PM10, NO_2, and CO, and the AQI values were classified into Good, Moderate, and Poor categories. Data preprocessing, correlation analysis, and feature importance evaluation were performed to improve prediction reliability and understand pollutant influence on air quality classification.

Experimental results showed that the Decision Tree classifier achieved an accuracy of 91.5%, outperforming the Random Forest model, which achieved 89.3% accuracy. In addition to better classification performance, the Decision Tree model provided transparent IF–THEN decision rules that improved explainability and made the prediction process easier to interpret.

The comparative analysis demonstrates that explainable machine learning approaches can provide reliable air quality classification while maintaining transparency and interpretability. The proposed framework can therefore be effectively used in practical environmental monitoring and decision-making systems.

6.1 Future Work

Future improvements to this research may include:

- Using larger and more diverse environmental datasets
- Comparing additional machine learning models such as Gradient Boosting and XGBoost
- Incorporating environmental factors such as atmospheric pressure, rainfall, and seasonal variations
- Applying advanced explainable AI (XAI) techniques for better model interpretation
- Improving AQI computation using standard government AQI formulas
- Developing real-time air quality monitoring and prediction systems

REFERENCES

[1] G. Soundari, J. Gnana, and A. C. Akshaya, “Indian air quality prediction and analysis using machine learning,” International Journal of Applied Engineering Research, 2019.

[2] Bhusan, S. Singh, and R. Kumar, “Enhancing AQI prediction using hybrid machine and deep learning models,” Environmental Science and Pollution Research, 2022.

[3] G. Mani, J. K. Viswanadhapalli, and A. A. Stonie, “Prediction and forecasting of air quality index using regression and ARIMA models,” Journal of Engineering Research, 2021.

[4] H. Liu, Q. Li, D. Yu, and Y. Gu, “Air quality index and air pollutant concentration prediction based on machine learning algorithms,” Applied Sciences, vol. 9, no. 19, Art. no. 4069, 2019.

[5] M. Castelli, F. M. Clemente, A. Popovic, S. Silva, and L. Vanneschi, “A machine learning approach to predict air quality in California,” Complexity, vol. 2020, pp. 1–15, 2020.

[6] N. Srinivasa Gupta, Y. Mohta, K. Heda, R. Armaan, B. Valarmathi, and G. Arulkumaran, “Prediction of Air Quality Index Using Machine Learning Techniques: A Comparative Analysis,” Journal of Environmental and Public Health, vol. 2023, Art. no. 4916267, 2023, doi: 10.1155/2023/4916267.

[7] S. Halsana, “Air quality prediction using supervised machine learning algorithms,” International Journal of Scientific Research in Computer Science, 2020.