Predicting Air Quality Levels using Advanced Machine Learning Algorithms for Environmental Insights

This project leverages machine learning techniques to predict air quality levels using environmental and pollutant data.

The objective is to build accurate, interpretable models that can assist in proactive environmental monitoring and public health safety.

Table of Contents:

- Introduction

- Dataset

- Features

- Technologies Used

- Modeling Approach

- Results

- How to Run

- Project Structure

- Contributing

- License

Introduction:

Air pollution is a pressing global concern with direct impacts on health and the environment. Predicting air quality levels enables

better planning and response to pollution events. This project explores multiple ML algorithms to forecast Air Quality Index (AQI)

using real-world data.

Dataset:

- Source: [Insert dataset source or URL]

- Attributes: Includes concentrations of PM2.5, PM10, NO2, SO2, CO, O3, temperature, humidity, wind speed, and more.

- Target: Air Quality Index (AQI) level.

Features:

- Data preprocessing and cleaning

- Feature selection and engineering

- Multiple machine learning models (Random Forest, XGBoost, etc.)

- Model evaluation using metrics like MAE, RMSE, R2

- Visualizations for analysis and interpretation

Technologies Used:

- Python

- Pandas, NumPy

- Scikit-learn

- XGBoost / LightGBM

- Matplotlib, Seaborn

- Jupyter Notebooks

Modeling Approach:

1. Exploratory Data Analysis (EDA)

2. Preprocessing: Handling missing values, outliers, encoding

3. Model Training: Tried various regression models

4. Hyperparameter Tuning

5. Evaluation & Visualization

Results:

- Best-performing model: [e.g., Random Forest with R2 = 0.89]

- Key factors influencing AQI identified

- Graphical insights into model predictions vs actual values

How to Run:

1. Clone this repository:

   git clone https://github.com/yourusername/air-quality-ml.git

   cd air-quality-ml

2. Install dependencies:

    pip install -r requirements.txt

3. Run the notebook or script:

    jupyter notebook notebooks/air_quality_prediction.ipynb

Project Structure:

air-quality-ml/

  data/   # Raw and processed datasets

  notebooks/   # Jupyter notebooks

  models/   # Saved model files

  src/   # Source code for preprocessing, training, etc.

  visuals/   # Generated plots and graphs

  requirements.txt   # Python dependencies

  README.md   # Project overview

Contributing:

Contributions are welcome! Please open issues or pull requests for suggestions and improvements.

License:

This project is licensed under the MIT License. See LICENSE for details.