Dataset for Predicting Air Quality Levels Using Advanced Machine Learning Algorithms

1. UCI Machine Learning Repository - Beijing Multi-Site Air-Quality Data
   - Link: https://archive.ics.uci.edu/ml/datasets/Beijing+Multi-Site+Air-Quality+Data
   - Features: Hourly air pollutants, meteorological data, multiple monitoring stations.
   - Period: 2013 to 2017.

2. OpenAQ Air Quality Data
   - Link: https://openaq.org/
   - Features: Real-time and historical air quality data from many cities worldwide.
   - API Access: Available for integration into ML pipelines.

3. Air Quality Data Set - Kaggle
   - Link: https://www.kaggle.com/datasets
   - Examples:
     - "Air Quality in India"
     - "Daily Air Quality Data in California"
     - "Real-time Air Pollution in Bangkok"

4. European Environment Agency (EEA)
   - Link: https://www.eea.europa.eu/en

These datasets include pollutant concentrations (PM2.5, PM10, NO2, SO2, CO, O3), weather conditions, timestamps, and geographic locations, which are essential for machine learning-based air quality prediction.