Predicting air quality levels using advanced machine leaming algorithms for environmental insights

Student Name: PANDIYARAJAN G

Register Number: 732323205307

Institution: ssm college of engineering

Department: Information Technology

Date of Submission:23/04/2025

Abstract

Air pollution poses a significant threat to human health and the environment, making accurate air quality prediction a critical task for sustainable urban management and public health protection. This study explores the

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

Predicting air quality has become increasingly important due to the health and environmental impacts of air pollution. Advanced machine learning (ML) algorithms offer powerful tools for analyzing large volumes of environmental data and making accurate air quality forecasts. These

1. Hardware Requirements

- **Processor (CPU):**

Intel i5/i7 or AMD Ryzen 5/7 (Minimum Quad-core)
For deep learning: Intel Xeon or AMD Threadripper (Recommended)

2. Software Requirements

- **Operating System:**
Windows 10/11, Ubuntu 20.04+, macOS Monterey or later

- **Programming Language:**
Python 3.7 or later

Objectives:

1. **Develop predictive models** using advanced machine learning algorithms to accurately forecast air quality levels based on historical and real-time environmental data.

Here is a **flowchart description** for the project titled **"Predicting Air Quality Levels Using Advanced Machine Learning Algorithms for Environmental Insights"**. You can use this to create a visual flowchart or include it as a structured explanation in your report:

NO2, O3, temperature, humidity, wind).

3. **Data Preprocessing**

→ Clean the data (handle missing values, outliers)

→ Normalize/standardize features

→ Feature engineering (create new variables, encode categorical data)

Dataset Description:

For this project, an air quality dataset is used, consisting of historical and real-time environmental measurements collected from air quality monitoring stations. The dataset includes both pollutant concentrations and meteorological parameters that influence air quality.

1. Data Sources:

- Government environmental agencies (e.g., EPA, CPCB, AQICN)
- Open-access repositories (e.g., UCI Machine Learning Repository, Kaggle)
- Real-time IoT-based sensor networks (optional for advanced use)