AN ABSTRACT ON

PREDICTION OF AIR POLLUTION USING MACHINE LEARNING

By,

Ankush Farmania	16900318125
Jagriti Nahata	16900318096
Deepshikha Das	16900318106
Aamanuddin Ahmed	16900318143
Abhimanyu Kumar Roy	16900318142

Under the guidance of

Prof. Rakhi Neogi

Submitted in partial fulfilment of

BACHELOR OF TECHNOLOGY

In

ELECTRONICS AND COMMUNICATION ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING ACADEMY OF TECHNOLOGY

ACADEMIC YEAR 2020-2021

Objective of the Project

To collect data from websites hosting air pollutants data, mainly PM2.5 and predict future pollutant levels and Air Quality Index using Linear Regression Model from Machine Learning.

Abstract

Air quality of a certain region can be used as one of the major factors determining pollution index also how well the city's industries and population is managed. In our study of making air pollution prediction system, the main focus will be in exploring the suitable Machine Learning techniques that will help in better forecasting of the Air Quality Index (AQI). We are focused on predicting the air pollution level of a specific region by using certain parameters like PM10, PM2.5, SO₂, NO₂, Benzene, CO and O₃ (ozone). This will help individuals and government to take steps to mitigate air pollution and will also help people to choose to avoid polluted areas and environmentalists may study patterns in data to arrive at scientific conclusions about how certain environmental factors affect pollution.

Plan of Work

The data is collected from various sources. The data is analysed, processed and incomplete data are completed by using various algorithms. We, also made a comparative analysis on the results predicted using various parameters and actual results by the use of suitable algorithms. This will help in the prediction of air quality in different areas around the globe and this could serve as an important reference for local government agencies in evaluating present and making future air pollution policies.

Block Diagram

First data is collected from various online sources.

The data is pre-processed cleaned and incomplete data is completed using various python functions.

Linear Regression model is used in this project to predict future air pollution levels. The data set is divided between 2 parts. The model will be trained on one set and the second set (test set) will be used to verify results produced.

The predicted and actual values of the outcome variable in the test set and measure accuracy using various measures.