

DS 4002 Case Study

Predicting Air Quality and Pollutant Concentration

Air pollution poses one of the most significant environmental and public health challenges worldwide, contributing to millions of premature deaths annually due to respiratory and cardiovascular diseases [1]. In densely populated regions, air quality regularly exceeds recommended limits for pollutants like nitrogen dioxide and particulate matter, especially during the winter months when temperatures decrease and energy demand increases [2]. Monitoring and predicting fluctuations in these pollutants are essential for guiding public health advisories and informing environmental policy.

You are a climatologist studying the fluctuating levels of five types of air pollutants in an Italian city. You have set up sensors around the city to detect the concentrations of the pollutants in the atmosphere, and left them for a year to gather data. Now the year is up, and you have 12 months' worth of pollutant data to analyze. You would like to use the data from the past year to predict the pollutant concentrations of the five different chemicals for the next month. You can do this by creating a regression model that is trained on the past 12 months of data. Your code should produce two graphs for each pollutant: one to demonstrate that the model can accurately predict pollutant concentrations from a previous month (validation), and one to predict pollutant concentrations for a future month (a true prediction).

You will be able to find all the necessary information, data, and code to help you complete this project at the following GitHub link: [jessiebailey/DS4002-CS3: DS 4002 Case Study 3](https://github.com/jessiebailey/DS4002-CS3)

Your work will enable scientists and weather analysts to determine which days will have the most hazardous pollutant conditions, to better inform populations who may be at high risk of negative health impacts from poor air quality. Good luck!

References:

- [1] State of Global Air, “Health Impacts of Air Pollution | State of Global Air,” www.stateofglobalair.org/hap, 2021. <https://www.stateofglobalair.org/hap>
- [2] “Air pollution fluctuations over the Po Valley,” [www.esa.int](https://www.esa.int/Applications/Observing_the_Earth/Air_pollution_fluctuations_over_the_Po_Valley). https://www.esa.int/Applications/Observing_the_Earth/Air_pollution_fluctuations_over_the_Po_Valley

GitHub: [jessiebailey/DS4002-CS3: DS 4002 Case Study 3](https://github.com/jessiebailey/DS4002-CS3)