Attention: All the .py file should be put in the same work folder with the original data. The code will generate many files while running and the 'submission.csv' is the exact file that the project needs.

The code has several important parts. The 'pollution_data.py' and 'weather.py' do the data cleaning work, which is dropping the duplicate data and filling in the missing data. The 'feature.py' generates features for our model, the 'modelconstruct.py' builds models for all pollutants and stations. Finally, 'prediction.py' predicts the 48 hours' concentration and 'getresult.py' makes the final csv file for submission. You just need to **put the original data file in the same fold with them and run the 'main.py'**.

More details will be shown in report and codes.