

Air Quality Prediction

Problem Statement

- 1) The objective is to read historical data from two different sources and join them to predict Air Quality
- 2) Historical weather report is available here: <https://en.tutiempo.net/climate/03-2013/ws-421820.html>. The web page shows the data for one calendar month at a time. You need to compile the data for all the months from 2013 to 2015
- 3) There is another csv file given to you that contains the Air Quality Index on an hourly basis for all days from 2013 to 2015. The file is available in Data/AQI/aqi/2013.csv, aqi2014.csv, aqi2015.csv
- 4) Merge the above files date wise to have the weather data and PM2.5 value together in the same row.
- 5) The weather data are the independent variables and PM2.5 value is the Target Variable that we need to predict
- 6) Build Machine Learning Models for both Regression and Classification to predict PM2.5 value when weather data is provided (regression) or Classify the Air Quality as good, moderate, bad, hazardous (Classification) and come up with the best model for both Classification and Regression