

Air Quality Index Prediction Model

-Bhargavi Jadhav

Problem Statement:

Predict various parameters of air quality using Python.

Approach:

1. Understanding the data.
2. Data processing:
This is done by making relevant changes to the data frame in order to get data that is fit for modelling, like:
 - i) Changing the decimal
 - ii) Slicing the data frame to get rid of unwanted columns
 - iii) Dropping the unwanted rows
 - iv) Handling the missing values and converting them to the mean values of the column

	Date	Time	CO(GT)	PT08.S1(CO)	NMHC(GT)	C6H6(GT)	PT08.S2(NMHC)	NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)	PT08.S5(O3)	T	RH
0	10/03/2004	18.00.00	2.60000	1360.0	150.0	11.9	1046.0	166.000000	1056.0	113.000000	1692.0	1268.0	13.6	48.9
1	10/03/2004	19.00.00	2.00000	1292.0	112.0	9.4	955.0	103.000000	1174.0	92.000000	1559.0	972.0	13.3	47.7
2	10/03/2004	20.00.00	2.20000	1402.0	88.0	9.0	939.0	131.000000	1140.0	114.000000	1555.0	1074.0	11.9	54.0
3	10/03/2004	21.00.00	2.20000	1376.0	80.0	9.2	948.0	172.000000	1092.0	122.000000	1584.0	1203.0	11.0	60.0
4	10/03/2004	22.00.00	1.60000	1272.0	51.0	6.5	836.0	131.000000	1205.0	116.000000	1490.0	1110.0	11.2	59.6
5	10/03/2004	23.00.00	1.20000	1197.0	38.0	4.7	750.0	89.000000	1337.0	96.000000	1393.0	949.0	11.2	59.2

3. Time series analysis is done using the FB Prophet model. The following steps were followed:
 - i) Creating a data frame with respect to the Prophet model
 - ii) Using an appropriate format for date and time

	ds	y
0	2004-03-10 18:00:00	48.9
1	2004-03-10 19:00:00	47.7
2	2004-03-10 20:00:00	54.0
3	2004-03-10 21:00:00	60.0
4	2004-03-10 22:00:00	59.6

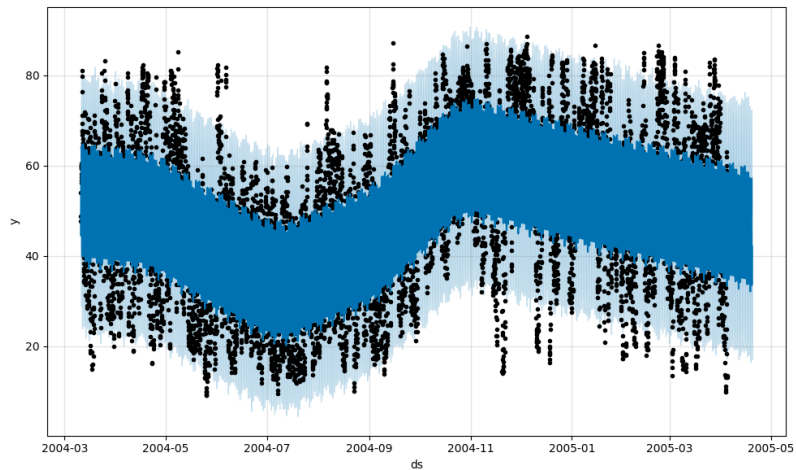
- iii) Fitting the model to the data frame
- iv) Making future predictions based on the model

Results:

The prediction model was used to predict the values of relative humidity for the upcoming periods =365. A similar prediction model can be used to predict other parameters by changing the column of y and over a different stretch of periods.

Outcomes:

	ds	yhat	yhat_lower	yhat_upper
9717	2005-04-19 15:00:00	33.553208	17.106383	49.303663
9718	2005-04-19 16:00:00	34.036737	19.329992	49.934142
9719	2005-04-19 17:00:00	35.670056	19.476000	51.412489
9720	2005-04-19 18:00:00	38.565679	22.224756	54.238524
9721	2005-04-19 19:00:00	42.119881	25.764562	57.701013



Trend and seasonality:

