

Egg Sensor NO2 Modelling - Week 4.5

Jonathan Levine, June 4th 2020

I started to model the Egg sensor data with respect to the Airpointer sensor. Below is a correlation matrix with all the elements I have so far in the dataset, including both the Egg and Airpointer features. I have also graphed both the time-series and the values plot for the Airpointer vs. Egg sensor; both plots are below.

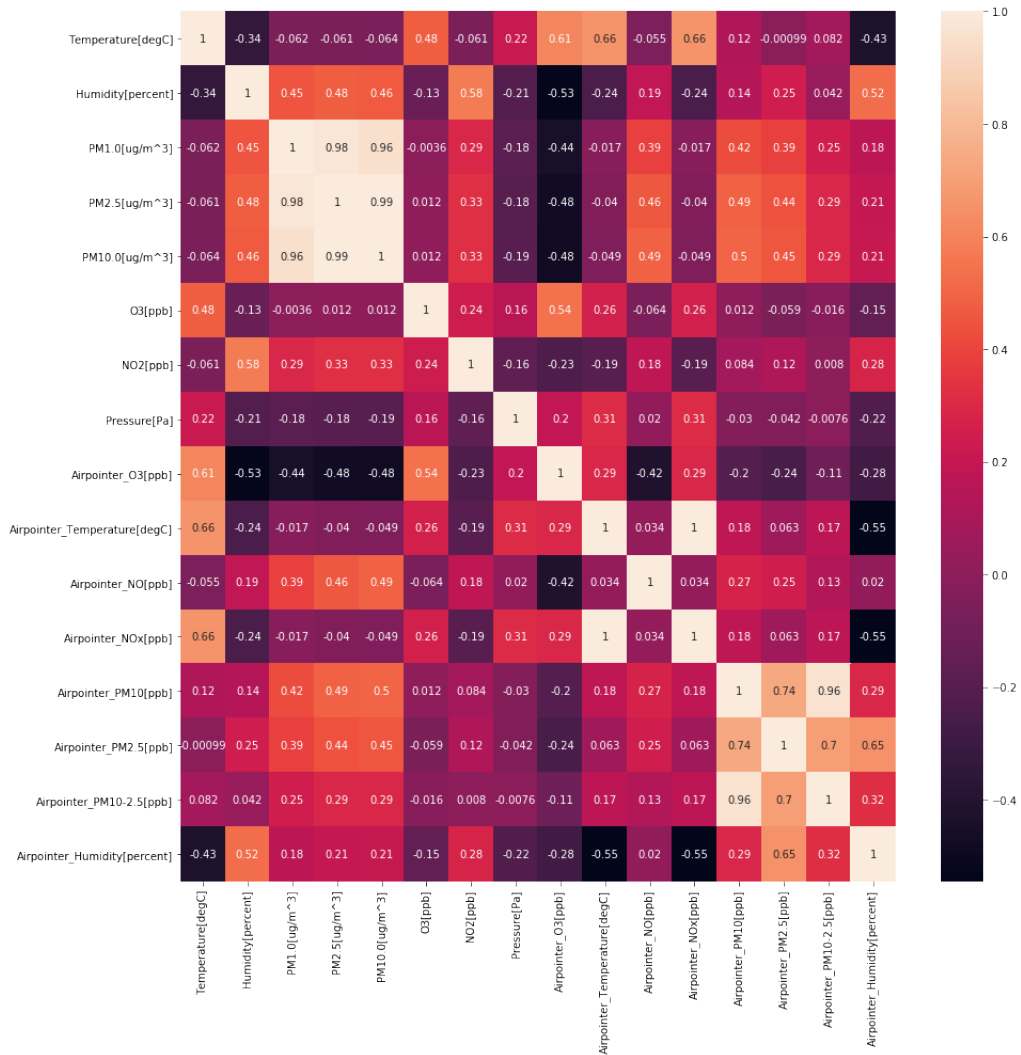


Figure 1: Correlation Matrix

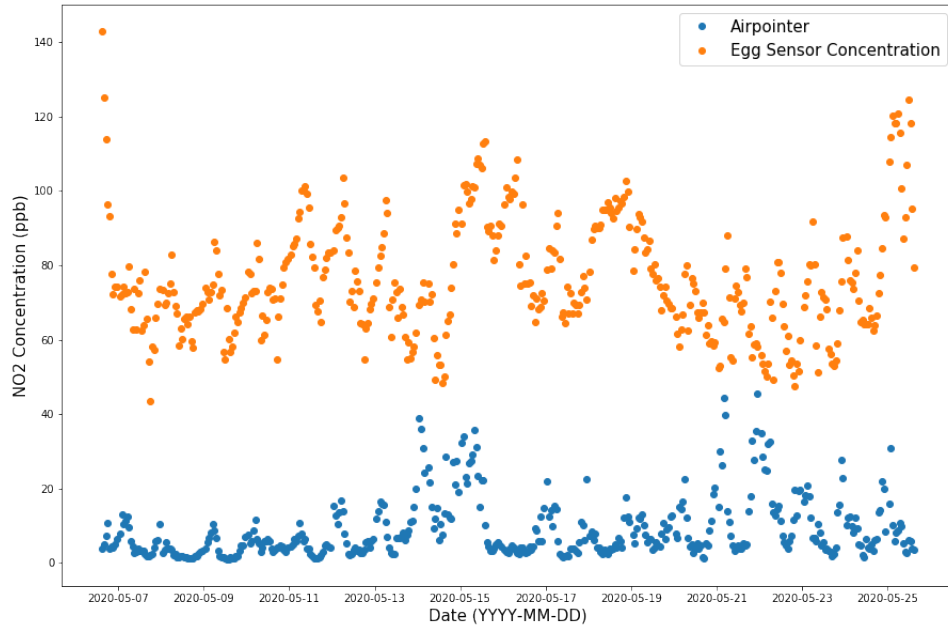


Figure 2: Time Series Plot

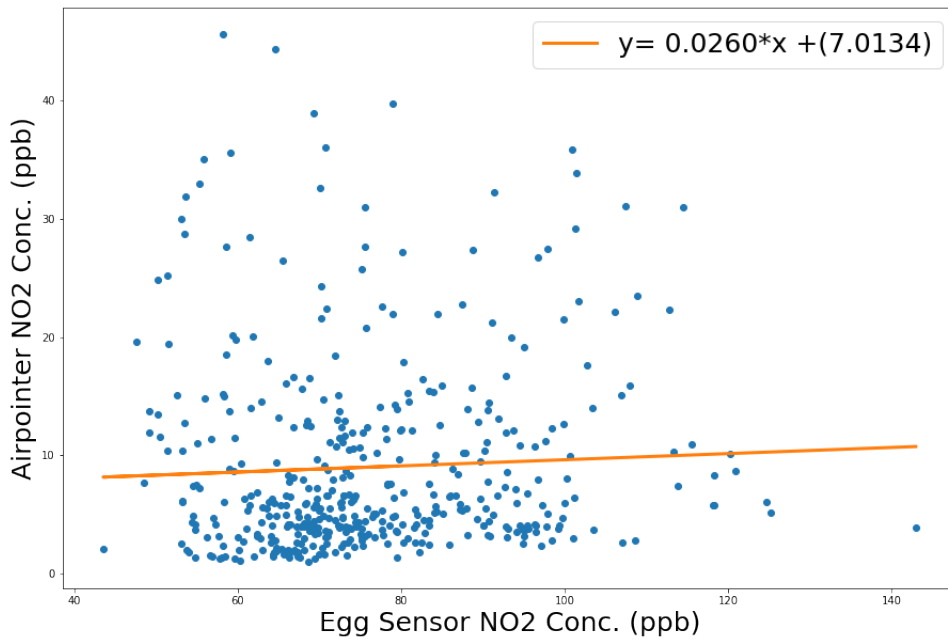


Figure 3: Airpointer vs. Egg Sensor (R2: 0.002476)

For my preliminary models, I split the dataset into a 75%/25% random train/test split. This means that training and testing points were randomly selected throughout the dataset

and used in the models. I did this because the current size of the dataset is quite small (437 points separated by an hour), and thus if I split it so that each model was trained on the first 90% of the dataset and tested on the last 10%, I am not sure how representative the results would be in terms of model performance.

The two modelling approaches I used were an **ElasticNet** and a **Decision Tree** regression. The features used in the models were all the measurements pulled from the egg sensor: temperature, humidity, PM1.0, PM2.5, PM10, O3, NO2, pressure. Note that the temperature reading is the temperature inside the box that the Egg sensors are housed in.

The models performed as follows,

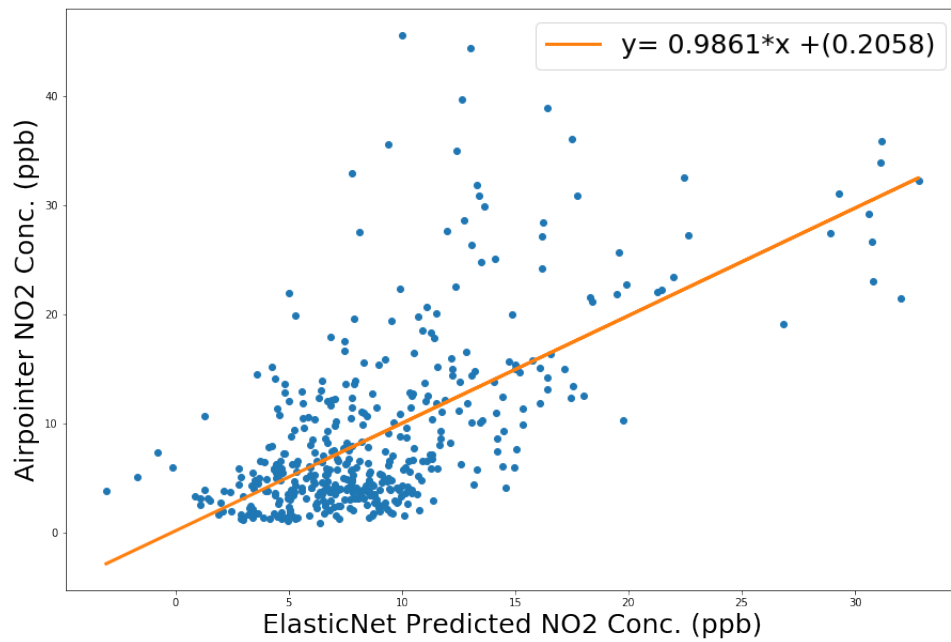


Figure 4: ElasticNet Regression (R2: 0.4037)

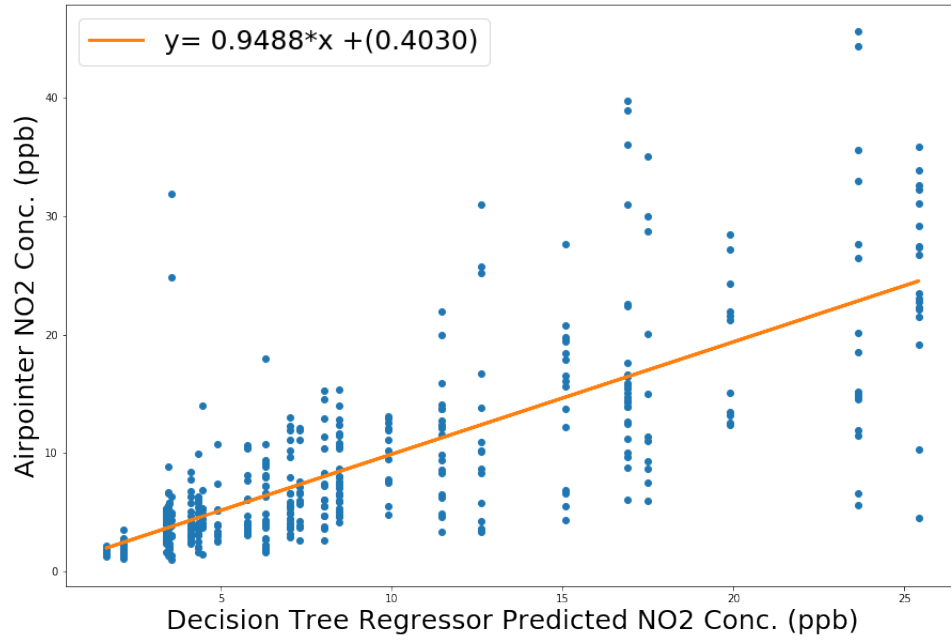


Figure 5: Decision Tree Regression (R2: 0.5734)

From this we can see a major improvement in the accuracy of the predicted NO2 concentration. If I also included environmental measurements from the Airpointer (air temperature, NO, NO_x, PM10, PM2.5, PM10-2.5, humidity), the models improve even more.

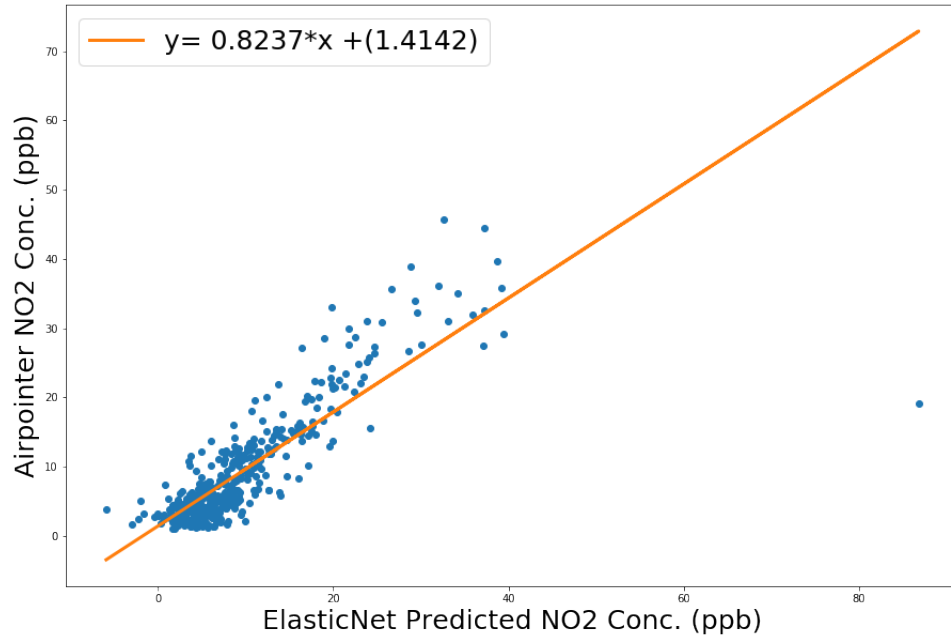


Figure 6: ElasticNet Regression (R2: 0.6724)

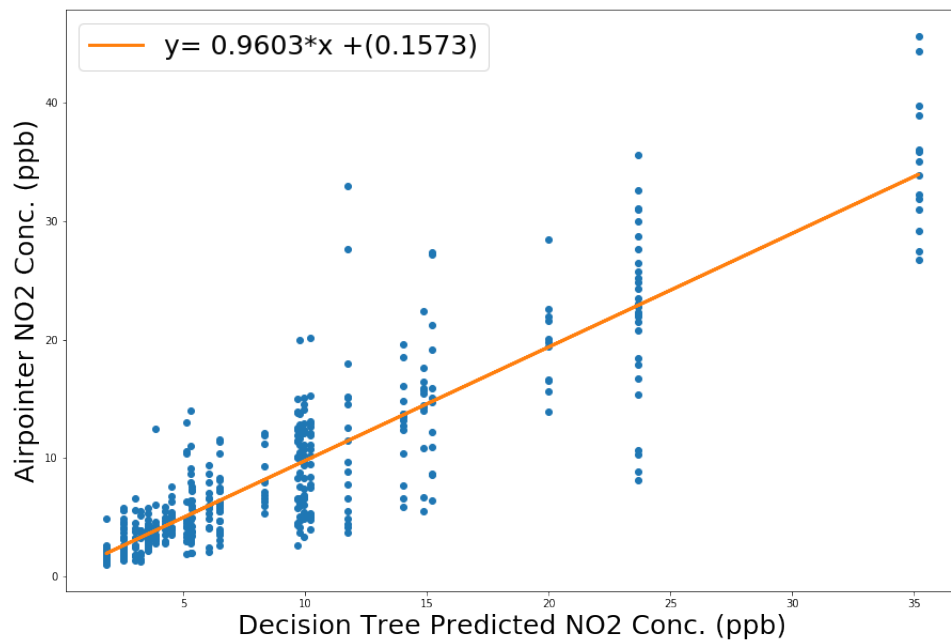


Figure 7: Decision Tree Regression (R2: 0.7827)

I am not sure how to interpret this improvement. More analysis is definitely needed.