Quality Assurance

Q1/6: The accuracy of the predictions about the live data are not as good as the results that you showed during the training of the model. Why do you think this is?

1. The model is underfit

Q2/6: Can you suggest a way to improve the predictions on the live data?

2. Collect more rows of data to train the model with

Q3/6: The client has offered some additional datasets that may be useful. Can you let us know which two datasets you think would be best to include in the model?

3. Weather, deliveries

Q4/6: We would like to explore the possibility of using a more complex machine learning algorithm to see how it compares to the current Random Forest. Can you suggest one to try?

4. Neural Network

Q5/6: What would be a disadvantage of using the more complex model from the previous question, against the current Random Forest?

5. More difficult to explain the algorithms results

Q6/6: Can you suggest a way that we can optimize the performance of the current Random Forest algorithm? In particular, we want to know how we can improve the MAE of the current algorithm.

6. Tune hyperparameters