Final Project: Building Basic predictive models over the NYC Taxi Trip dataset.

You are provided with the NYC Taxi Trip Dataset. This dataset contains information about the taxi trips that took place in different parts of New York City and how much time did that trip take to complete.

In this project, the following are the tasks you must complete and submitted.

1. Choose the most suitable evaluation metric and state why you chose it.
2. Build a benchmark model for the given dataset.
3. Build a K-Nearest neighbours' model for the given dataset and find the best value of K.
4. Build a Linear model for the given dataset with regularisation. Attempt to interpret the variable coefficients of the Linear Model.
5. Build a Decision tree model for the given dataset. Attempt to interpret the variable importance.
6. Plot the following Bar plots:
    0. train score of all the above models.
    1. test **(not validation!)** score of all the above models.
    2. Attempt to explain the observations from the plots (optional)

Guidelines:

The code should be submitted in the form of pdf format of the jupyter notebook. The option could be found in the "*File > Download as"* section in the notebook.

1. Make sure the pdf format notebook contains all the relevant output of the code cells. Notebooks that have not been run for displaying the outputs will be marked as incomplete.
2. The code should be well commented on and documented.
3. All project files should be added in a folder with your name and then compressed. (*Right-click on the folder> Send to > Compressed (zipped) folder*) and give the file your name. Then you may upload this zipped file on the project submission page.

Click here to download the Problem Statement & Dataset