ML Project Module 1

Model a two stage predictive machine learning engine that forecasts the on-time performance of flight from the given datasets.

DATA PREPROCESSING

Module 1: The weather folder consists of the weather data at 15 different airports in the USA between 2013 and 2017. The flight data of all the flights that flew inside the USA is available in the url provided below. You can use scripts to download them. Merge these two data appropriately such that each record of the flight should have the corresponding weather data available.

Preprocess the data only for the airports mentioned below.

Airport : codes
ATL
CLT
DEN
DFW
EWR
IAH
JFK
LAS
LAX
MCO
MIA
ORD
PHX
SEA
SFO

Using Jupyter notebook/jupyterLab is highly encouraged as it makes it easier for us to help you out.
To install all the packages you will be requiring to do ML, visit here.

Link to flight data: here
Link to weather data: here

Lookups: Dask ,Numpy, Pandas, datetime, os, json, wget

Deadline: 15 days