The Data Incubator

Project Proposal:
Flight Departure Delay Prediction

Presented by:
Dooman Akbarian

Oct 2020

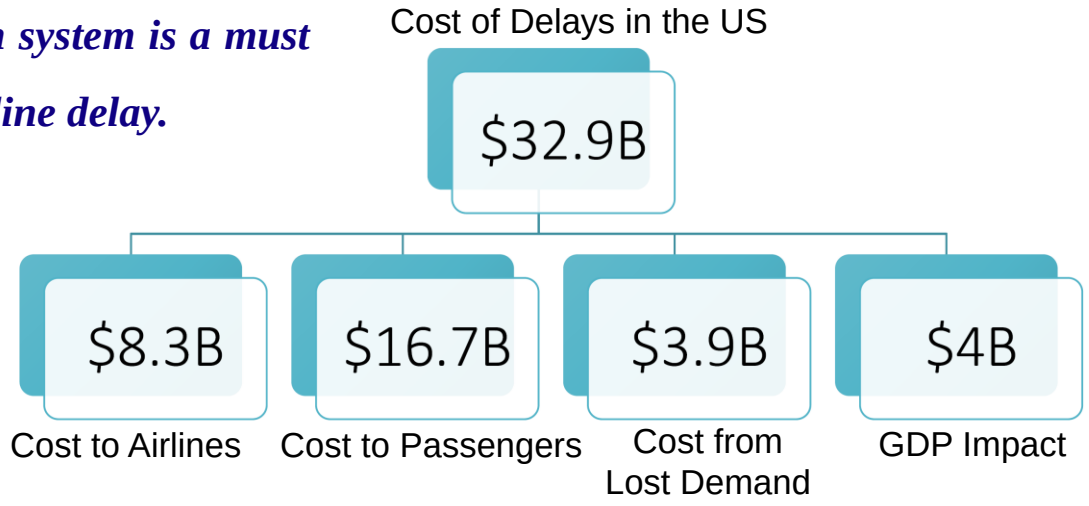
Business Problem Overview

- Flight delay is a challenging problem for passengers and airline companies, which leads to[1]:
  - Financial losses.
  - Negative impact on their business reputation.

"You look like you need a holiday."

- Airline industry incurs an average cost of about $11,300 per delayed flight[1].
  - 61,000 delayed flights per month on average.
  - Excluding costs to passengers and lost demand.

❖ ***An intelligent and automated prediction system is a must in this case that can predict possible airline delay.***



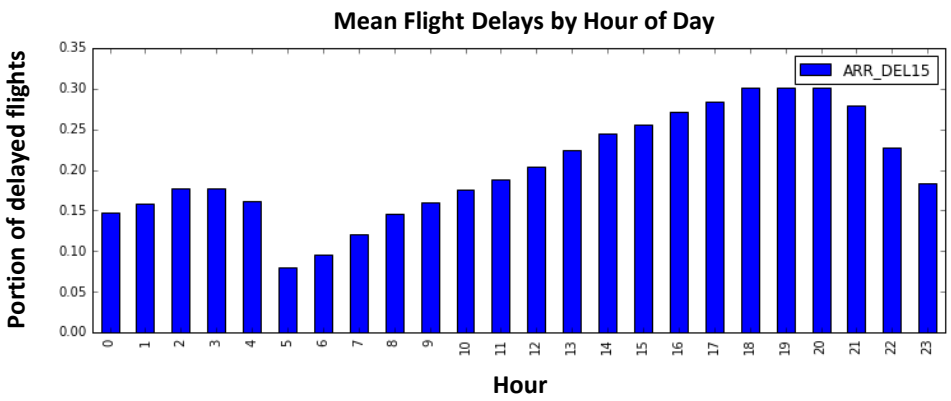
1. Ferguson, J.,et al., 2013. Trans Res C, 33, 311-323.

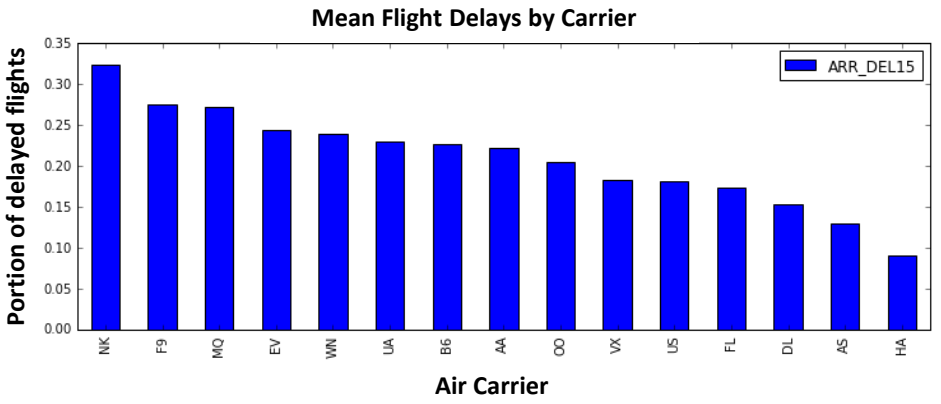
Data Analysis

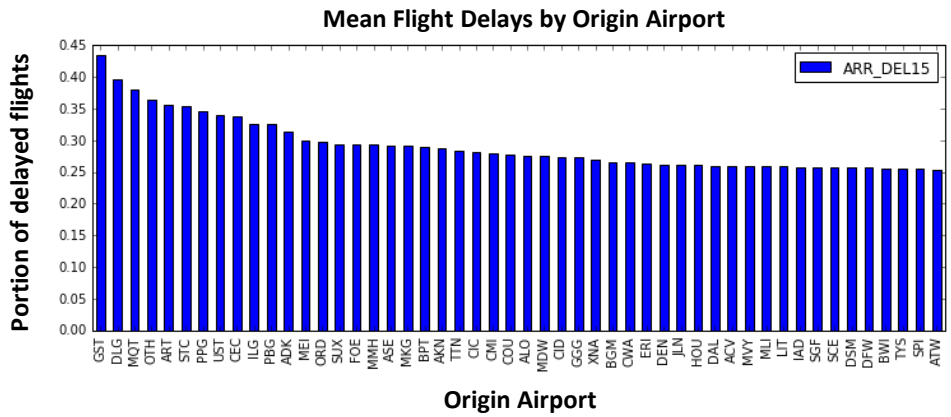
US Airline On-Time Performance Data provided by BTS

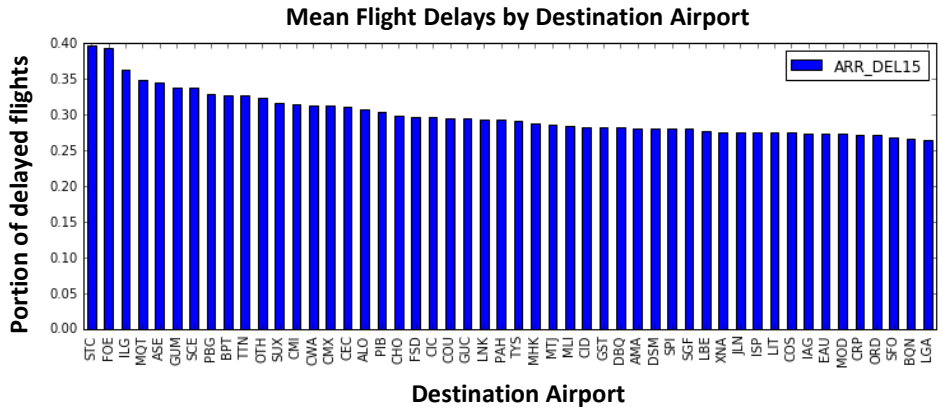
- Number of Flights
- Scheduled Arrivals/Departure Date and time
- Delay of flights
- Air Carrier
- Origin and Destination Airport

+5 Million flights in 2015 were analyzed







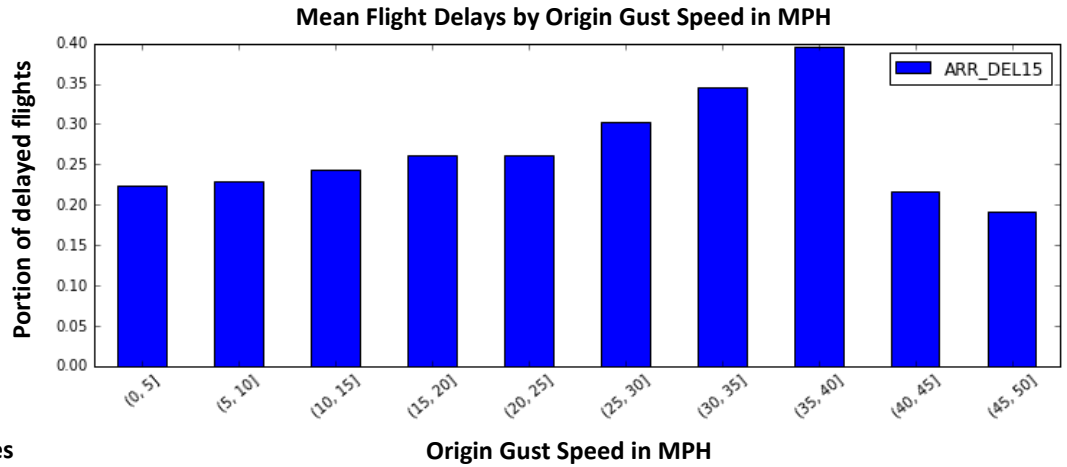


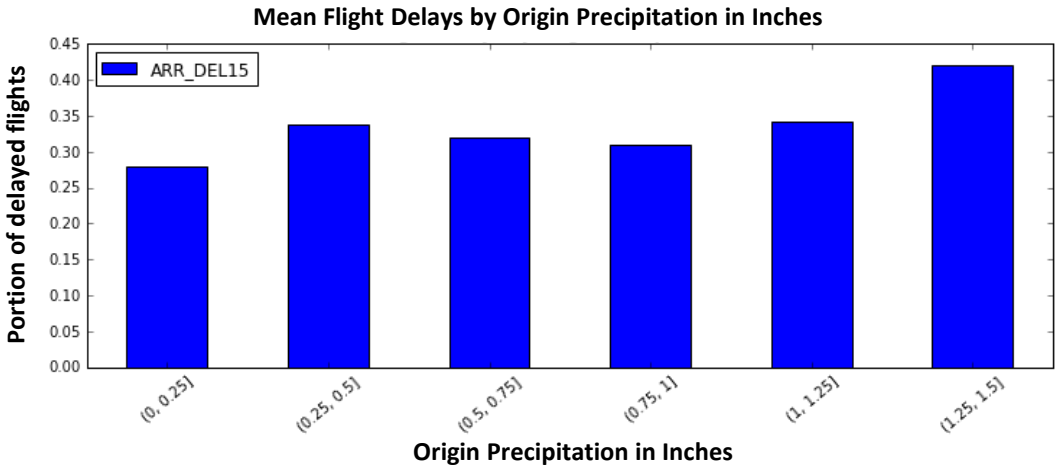
https://www.transtats.bts.gov/Tables.asp?DB_ID=120

Data Analysis

Weather data

Temperature	Precipitation	Clear Cloudy Fog Rain Snow
Wind Speed	Humidity	

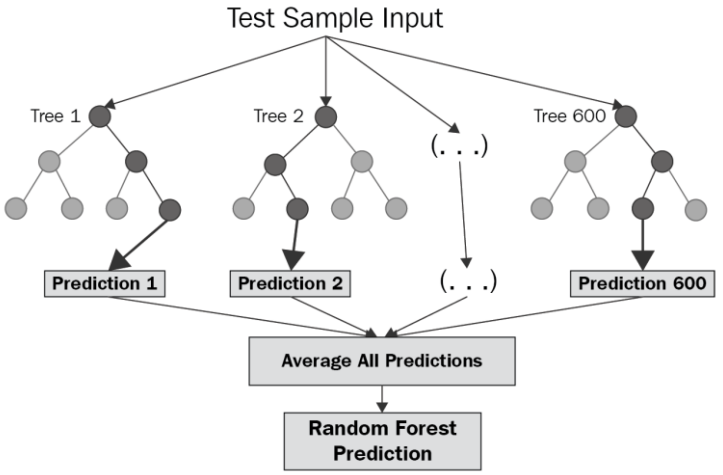




https://mesonet.agron.iastate.edu/ASOS/

ML based prediction webapp using Flask and Heroku

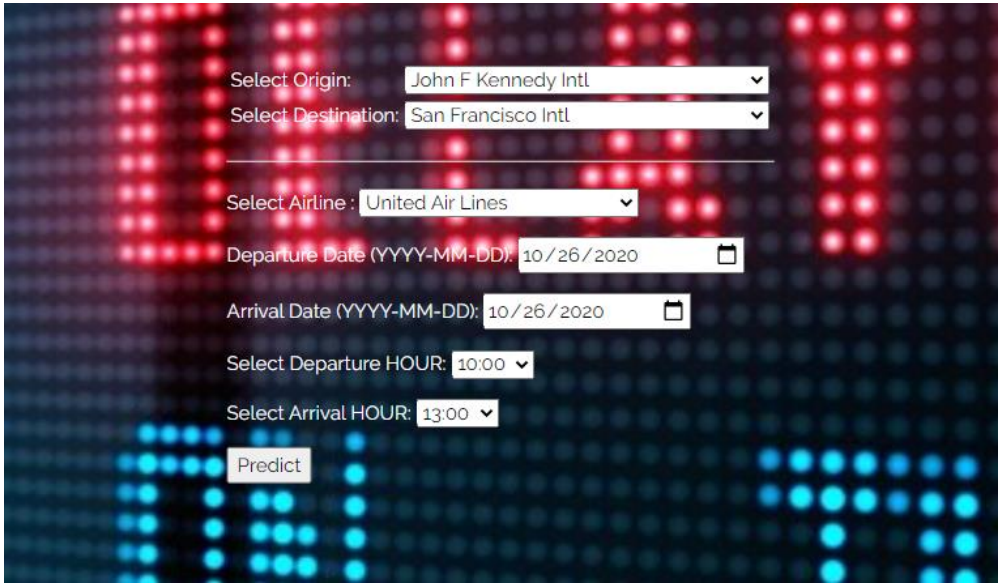
➢Using random forest to predict flight departure delay based on features such as: arrival/departure month, day of the week, hour of the day, origin/destination airports as well as the airline.



- Train-test split: 80-20%
- 5 Fold Cross-validation
- ***R^2 (Test score): =0.65***

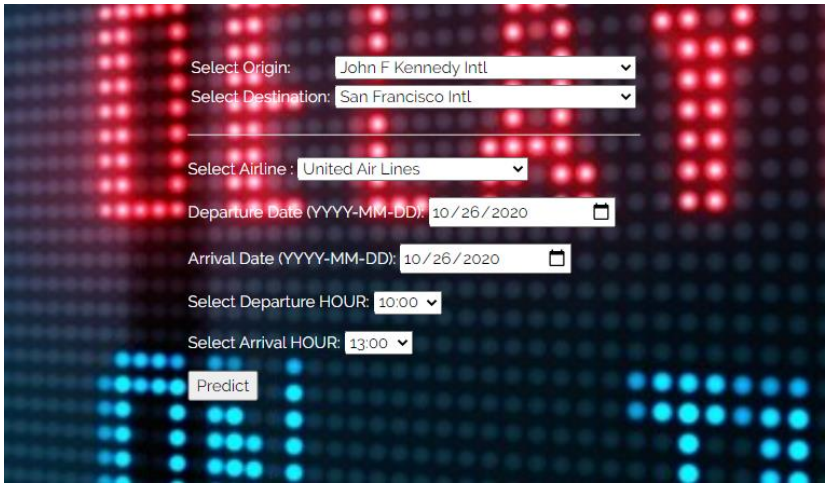
➢ The model was saved using Pickle and then a Flask app was created and finally deployed on Heroku:

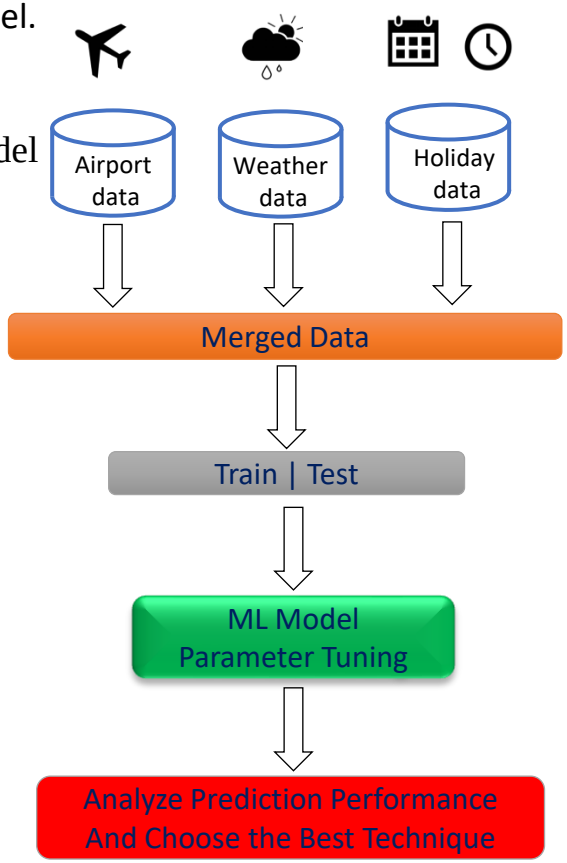
https://dooman-data-incubator.herokuapp.com



Future Work

- Holiday data as well as weather data will be considered as well to build the ML model.
- Several ML models will be investigated, and after hyperparameters tunning, the model with best performance will be chosen.
- Development of an interactive app for users :





	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Data wrangling/Analysis								
ML model development								
App platform								
Testing/Report								

Thank You For Your Attention