Project Charter

Problem Statement

As of 2017, about 1 in 5 US flights were delayed by more than 15 minutes which resulted in an aggregate economic toll of $26.6 billion, almost half of which was charged to passengers.

Of the various causes for delays, about 39% were attributed to weather.

Passengers suffer loss of time, missed business and leisure opportunities.

Airlines suffer loss of business, increased fuel costs and have a negative impact on overall environment.

Business Case & Benefits

A new ML based solution would help airlines as well as passengers with timely re-bookings which would minimize losses due to cancellations and detours and improve operational efficiencies, user experience and customer satisfaction.

A new ML based solution can be used for more reliable, closer to real-time decision-making in air traffic management.

This solution could potentially be used to augment the current FAA system in operation which is based on demand-capacity models.

Goal Statement

To develop and implement a machine learning model using individual flight and location specific meteorological data to predict departure delays that exceed 15 minutes.

The prediction timeline is for 2 hrs prior to flight departure.

The baseline goal for accuracy and recall for this model should be close to 70% (based on current research).

The execution time for this model must not exceed

Milestones

Week 1: EDA - due July 5th
Week 2: Feature Engineering - due July 12
Week 3: Base Algorithm - due July 19
Mid-Project Presentation: due July 23
Week 4 Algorithm Optimization & Fine-tuning: due July 26
Week 5: Novel Approaches due
Week 6: Write-up due Aug 5
Final Presentation: due Aug 6

Scope

Approaches to predicting delays can be classified into 2 groups:

1 - **Delay Propagation Based Methods** - which study the delay propogation mechanisms in the air-traffic network. These methods are based on understanding causation.

2 - **Data Driven Methods** - which is based on data mining, statistical techniques and ML. These methods are based on leveraging correlation.

In this project,
- we shall attempt to produce an ML based solution (group 2).
- we shall be considering weather information

Team Members

Sirak Ghebremusse: Software/ML Engineer
Madhukar Reddy: ML Research Data Scientist
Radhika Satapathy: Data Science Product Manager
Max Ziff: ML Data Engineer