Managing Flight Delays

By: Dong Zhen

About Delta Airlines

Delta is **one of the largest airlines** in the world, transporting nearly **200 million passengers to over 300 destinations** every year.

The Problem

Commercial airlines in the United States incur **billions of dollars in losses annually** due to flight delays.

What is a delay?

Departure from the gate **15+ minutes** after the scheduled departure time.

Top Contributors to Flight Delays

1. Air Carrier
2. Extreme Weather
3. National Aviation System
4. Security

The Real Cost of Flight Delays

Delta incurs **additional operational costs**

Loyal customers **use other airlines**

Damaged reputation

The Opportunity

There is an opportunity to **reduce the costs** incurred by unforeseen delays.

Impact

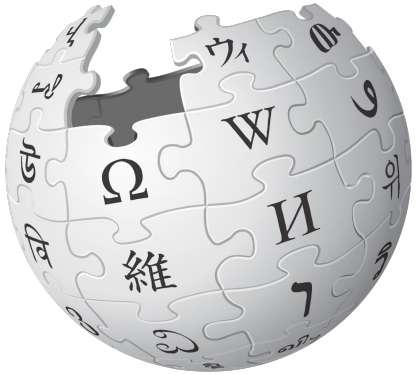
Delta can increase customer service personnel at the gate to help customers manage the delays.

Impact Hypothesis

The **costs incurred by unforeseen delays can be reduced** by **predicting which Delta flights is likely to be delayed** in order for Delta to allocate customer service personnel and better prepare for the delay.

Datasource

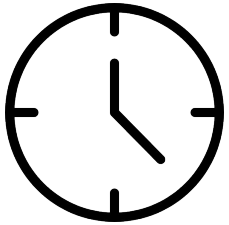
Bureau of Transportation Statistics



Wikipedia

Data from 2019

Data Features



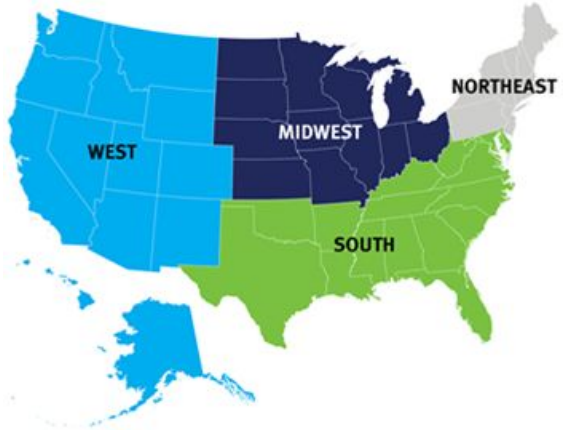
Engineered Features

Average Minutes Delayed Due to Weather Yesterday Region, Total Number of Scheduled Flights for the Day, Time Names

Classification Model

Baseline Model	Features Used	Score
Logistic Regression, Random Forest, XGBoost	Average Delayed Due to NAS, Delta, and Weather Yesterday	**Precision: 72%** **Recall: 4%**

Attempts to Improve Score



Used features by region only

Precision 59%
Recall 2 %

Used a combination of region and flight features

Precision 75%
Recall 1 %

Feature Importance

Regional Weather Yesterday .2

Schedule Departure Time Evening . 12

Schedule Departure Time Early Morning .12

Schedule Departure Time Name Morning .1

Schedule Departure Time Name_Night .07

Delayed Due To NAS Yesterday .05

Conclusion

The current modelling is **only able to pick up a small percentage of delays and predict relatively precisely using general features.** Some features always score at the top of importance such as time of day, NAS, and weather.

Future Work

I would like to **use more current data** instead of historical like weather at the origin and destination airports, look into crew schedule and air traffic control data.

Appendix

https://www.sheffield.com/articles/how-air-traffic-control-systems-work
https://www.faa.gov/data_research/aviation_data_statistics/media/cost_delay_estimates.pdf
https://transtats.bts.gov/ONTIME/Departures.aspx
https://www.bts.gov/topics/airlines-and-airports/understanding-reporting-causes-flight-delays-and-cancellations