

Problem Statement

What factors cause airline flight delays in commercial operations and can those factors be used to predict flight delays within 24 hours with an accuracy of at least 90% enabling air traffic to compensate and recover from said delays improving passenger's experience?

SMART Problem Checklist

S: Specific

- Flights within commercial operations.

M: Measurable

- Predict flights with an accuracy of at least 90%.

A: Action-oriented

- Air traffic compensate and recover from flight delays.

R: Relevant

- Improving passenger experience.

T: Time-bound

- Predict flight delays within 24 hours.
-

Context

Flight delays are unavoidable major issues in the airline industry. Once a flight has been delayed, it begins a chain reaction that ultimately reflects poorly on the business. The ability to predict flight delays will give both the business and the customer the advantage to make changes to their itinerary before the delay causes an inconvenience where the customer is relying on arriving at their destination in a timely manner. The solution is preemptive actions based on predictions to identify and adjust for delays accordingly.

Criteria for Success

Create a prediction model that will be able to predict airline flight delays within 24 hours at least within 90% accuracy.

Scope of Solution Space

Research within airport, flight, and weather data will be used to find correlations relating to flight delays. These correlations will be modeled providing predictions on future data.

Constraints within Solution Space

There is a limit to predicting outliers and out of context patterns especially pertaining to weather. Thunderstorms within high traffic air space often halts the surrounding airports and data Relevant to that event may skew accuracy when training the model.

Stakeholders

- Airlines
 - Airports
 - Investors
-

Key Data Sources

- [On Time Performance 2017](#)
 - [Severe Weather Data in 2017](#)
-

Overview

This project will be done for Springboard's Data Science Career Track as capstone two. My background in aeronautics will help me gain a better understanding on selecting delay factors when choosing my model. I plan to use the 2017 weather data to create filters on severe weather in locations using latitude and longitude. These filters will omit the severe weather delays since those factors are more difficult to predict. I am more concerned with the subtle factors that create flight delays on a day to day basis instead large weather related flight delays. From there I will be able to explore the data and find correlations between factors and flight delays. This information will be used to train my model which will predict flight delays within the next 24 hours. Theoretically, even though the data is from 2017 the model should still be able to predict flight delays on current data if executed correctly.

GitHub repo: <https://github.com/KalenWillits/CapstoneTwo.git>

Author: Kalen Willits