

Predicting Flight Delays



Drew Hibbard

Prepare for the Worst

- Flight delays are relatively common
- Understand the factors that affect delays



Obtaining the Data

- Kaggle - all 2015 US flights
- Airfleets.net - aircraft info
- NOAA - weather data
- FlightAware - real-time flight data

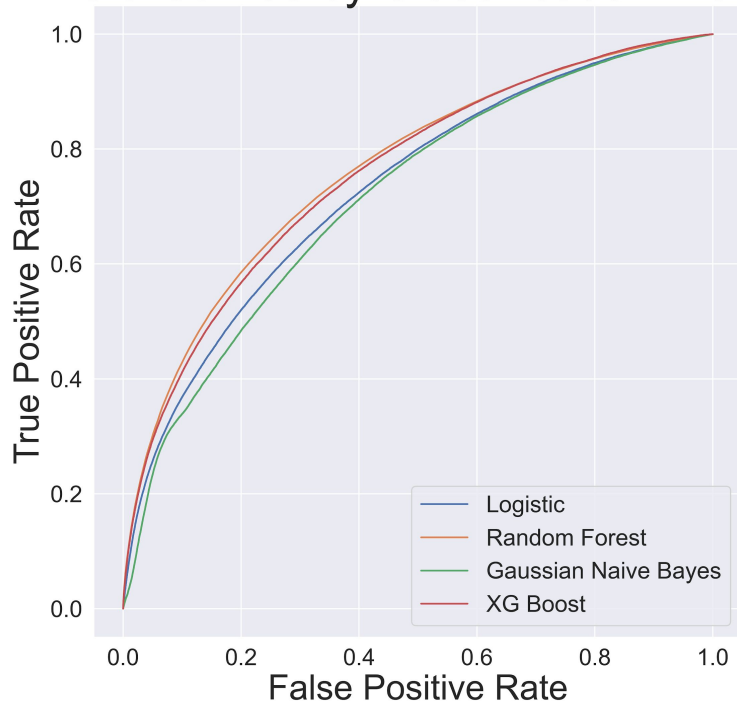




Exploratory Data Analysis

Model Testing and Results

ROC Curves by Classification Model



Model	AUC	F1 (delays)	F1 (non-delays)
Logistic Regression	0.74	0.43	0.86
Random Forest	0.76	0.49	0.85
Gaussian Naive Bayes	0.71	0.40	0.86
XG Boost	0.76	0.48	0.85

Threshold = 0.25



Recorded Demo



Questions?

Next Steps

- Use more recent data
- Update web app to allow users to change each predictive feature
at will