



# Airline Delay

Morgan A. Keith



## Business understanding

- This project is to build a model that predicts an Airline has a flight departure delay. The model will represent the Airlines American Airlines (AA), United Airlines (UA), and American Eagle Airlines (MQ).
- The model helps a customer determine if their flight will be delayed so they might plan to reschedule.
- The models used are Decision Tree, Random Forest, and XGBoost. Each model has a baseline, then applied random search, then SMOTE. So we can observe each model's predictions.
- Departure delay will be defined as any delay more than 15 minutes.



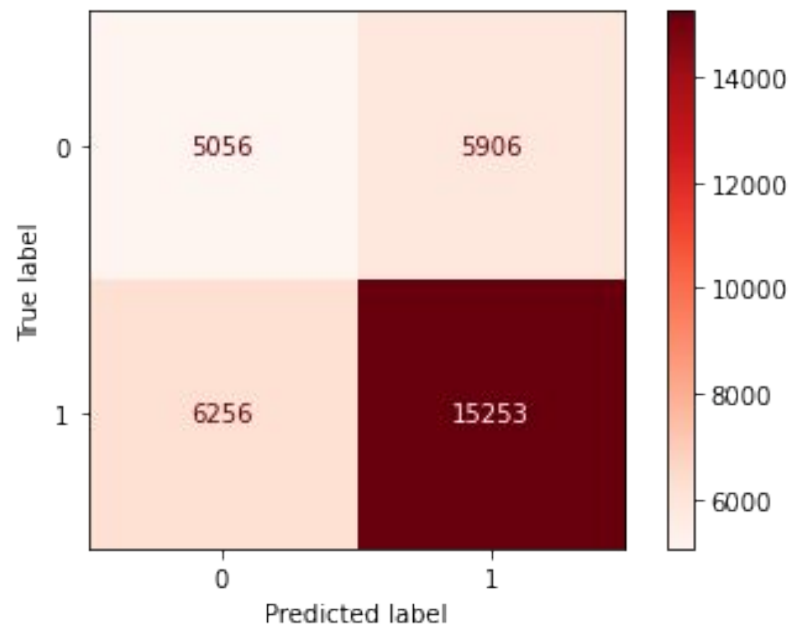
# Model Features

The features used to predict delay:

- Arrival Time of the flight
- Distance of the Flight
- Day of Month
- Month
- Day of Week
- Destination:
  - Dallas Fort-Worth
  - Denver
  - Orlando
  - Los Angeles
  - Atlanta

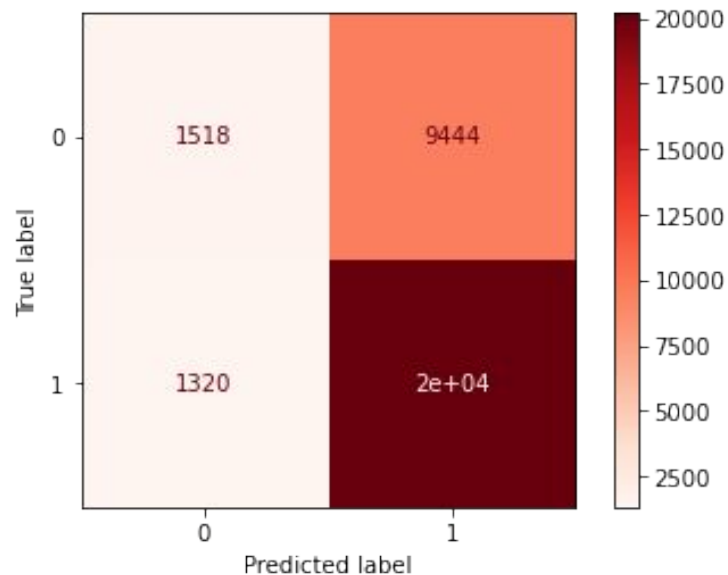
## Baseline Decision Tree

	precision	recall	f1-score
On Time	0.45	0.46	0.45
Delay	0.72	0.71	0.71



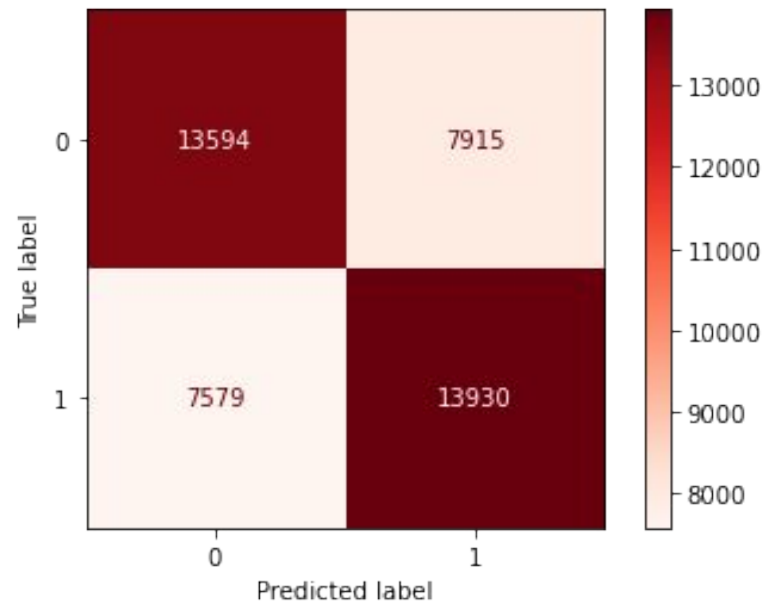
## Random Search Decision Tree

	precision	recall	f1-score
On Time	0.53	0.14	0.28
Delay	0.68	0.94	0.79



## SMOTE Decision Tree

	precision	recall	f1-score
On Time	0.64	0.63	0.64
Delay	0.64	0.65	0.64



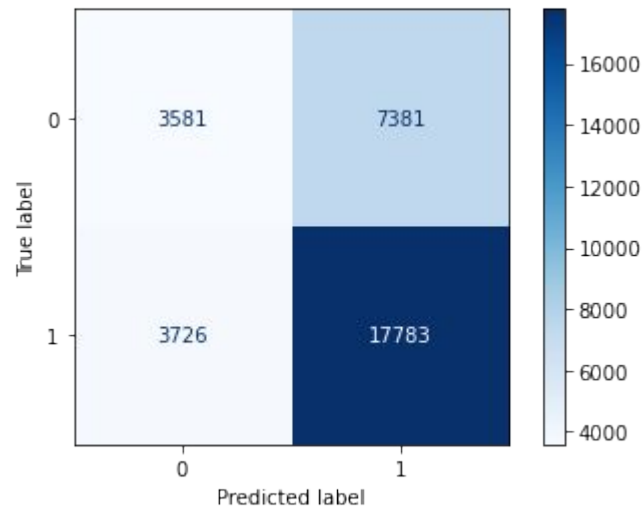


# Decision Tree

Random search varies its results on its ability to better predict delay. Though, SMOTE creates a consistent model. Neither model providing a significant prediction.

## Baseline Random Forest

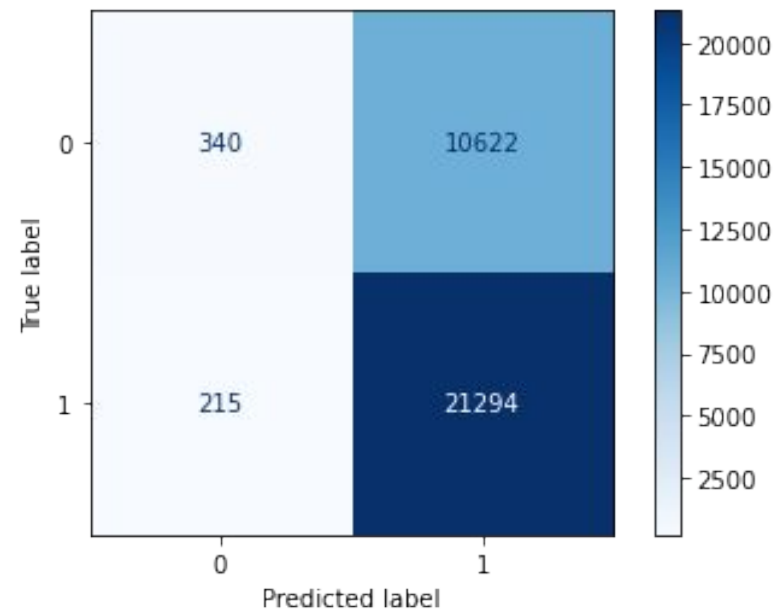
	precision	recall	f1-score
On Time	0.49	0.33	0.39
Delay	0.71	0.83	0.76





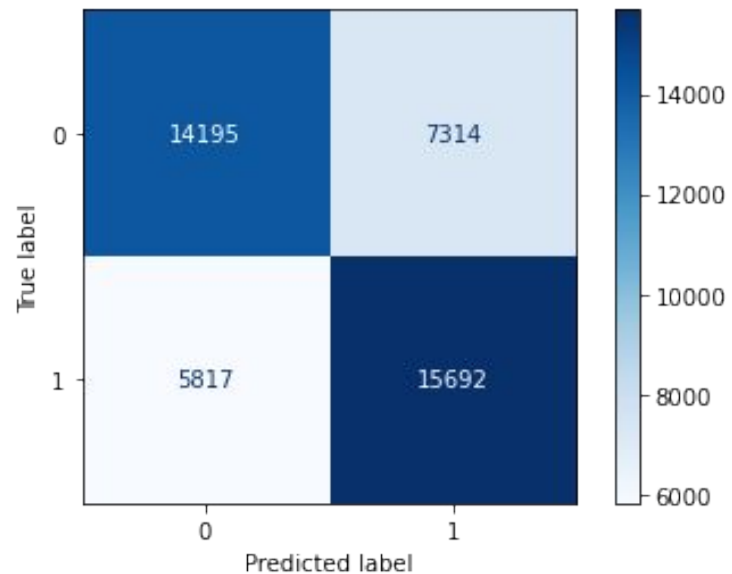
## Random Search Random Forest

	precision	recall	f1-score
On Time	0.61	0.15	0.23
Delay	0.67	0.99	0.80



## SMOTE Random Forest

	precision	recall	f1-score
On Time	0.71	0.66	0.68
Delay	0.68	0.73	0.71



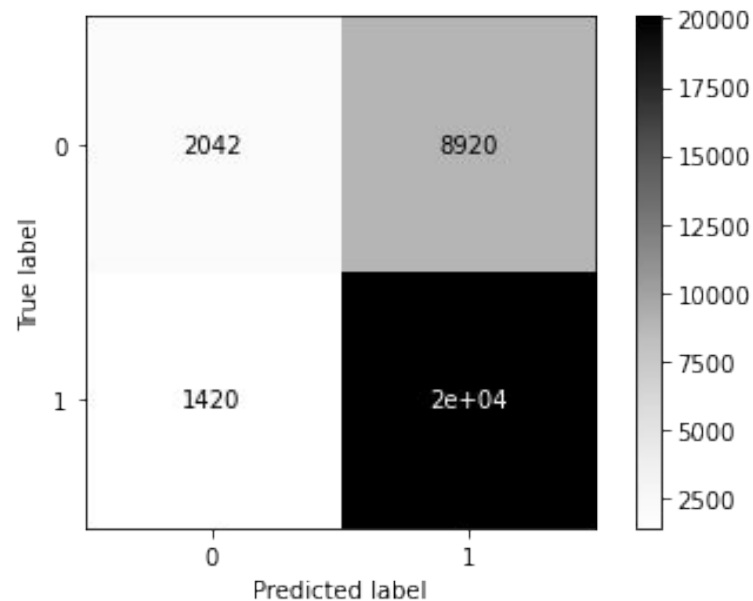


# Random Forest

Random search improves Delay recall and On Time precision. SMOTE creates a more consistent predictor for precision and recall. Random Forest does improve precision and recall on the decision tree model.

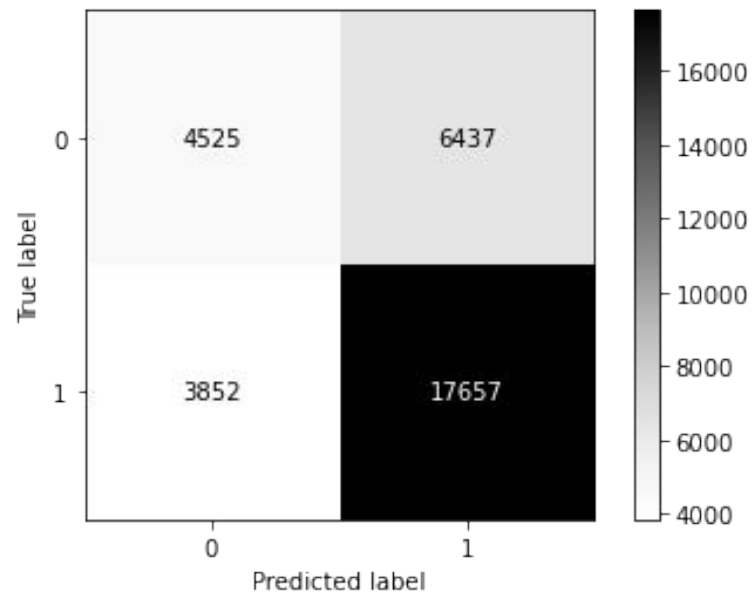
## Baseline XGBoost

	precision	recall	f1-score
On Time	0.59	0.19	0.28
Delay	0.69	0.93	0.80



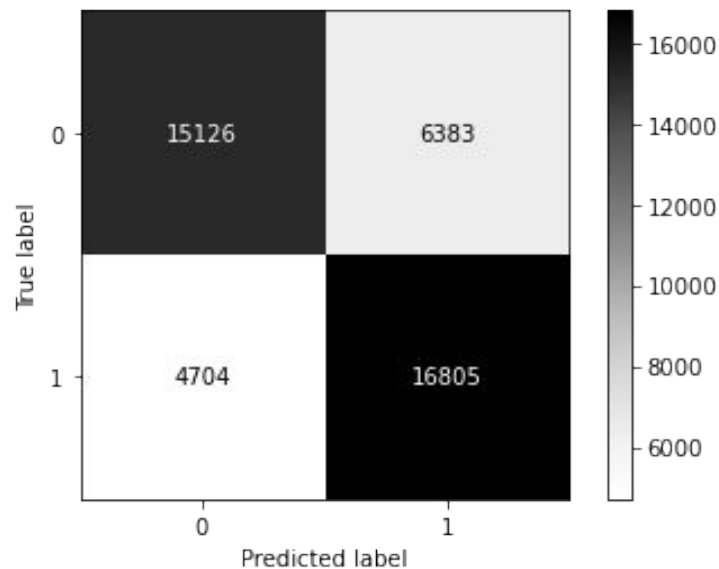
## Random Search XGBoost

	precision	recall	f1-score
On Time	0.54	0.41	0.47
Delay	0.73	0.82	0.77



## SMOTE XGBoost

	precision	recall	f1-score
On Time	0.76	0.70	0.73
Delay	0.72	0.78	0.75





# XGBoost

Random Search improves On Time recall while other measurement are the same. SMOTE improves all scores for On Time classification while keeping Delay the same.

XGBoost slightly improves on Random Forest as the best model for predicting a delay. XGBoost with random search and SMOTE create a model able to predict On Time and Delay flights over 70%.