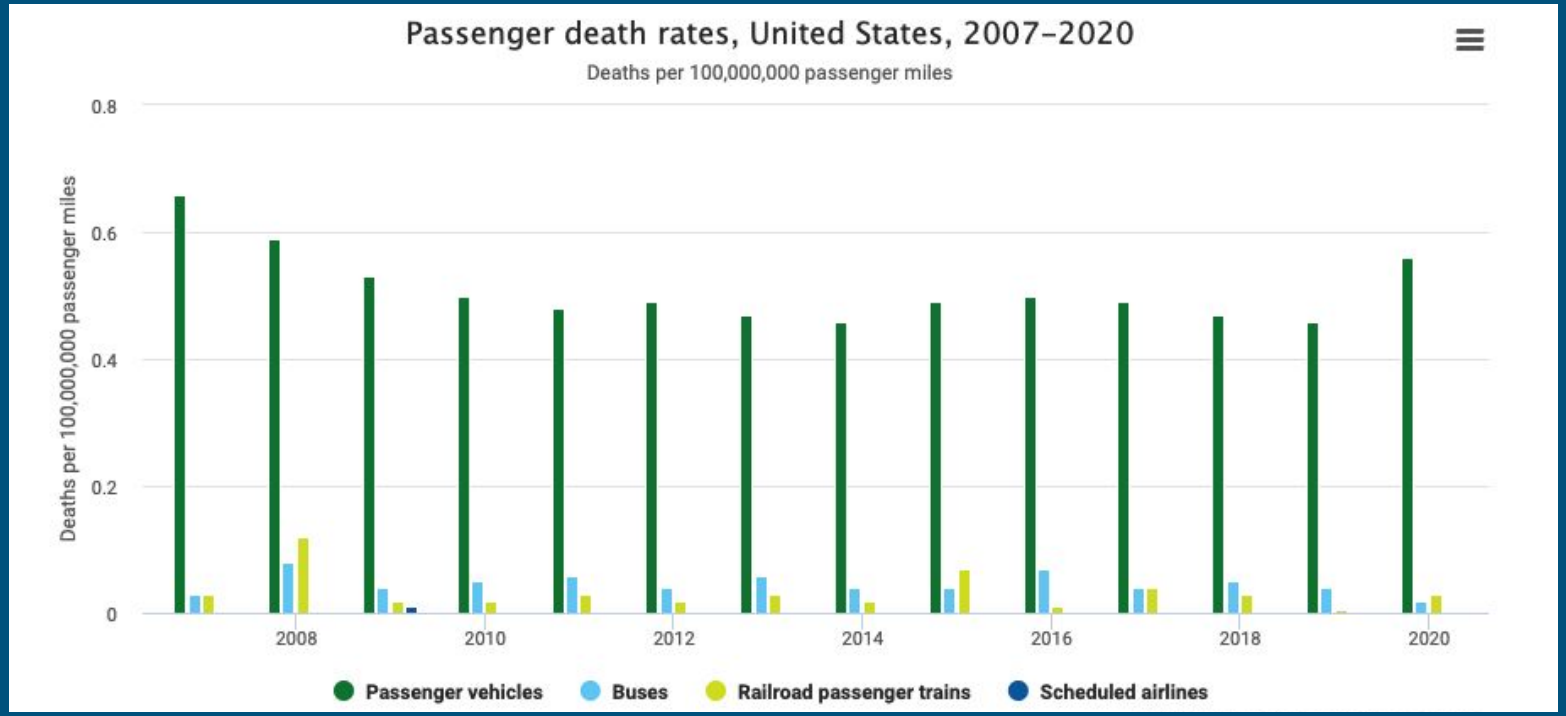
Air Crashes, How Severe Would Be?

Andy Xu

Why people pay more attention to air crash?



Source: https://injuryfacts.nsc.org/home-and-community/safety-topics/deaths-by-transportation-mode/

Problem Statement and Result

- In order to help publics to understand more about the plane crashes, this study utilizes major features in the historical airplane crash records and explores the severeness of crash, whether there is survivor or not.

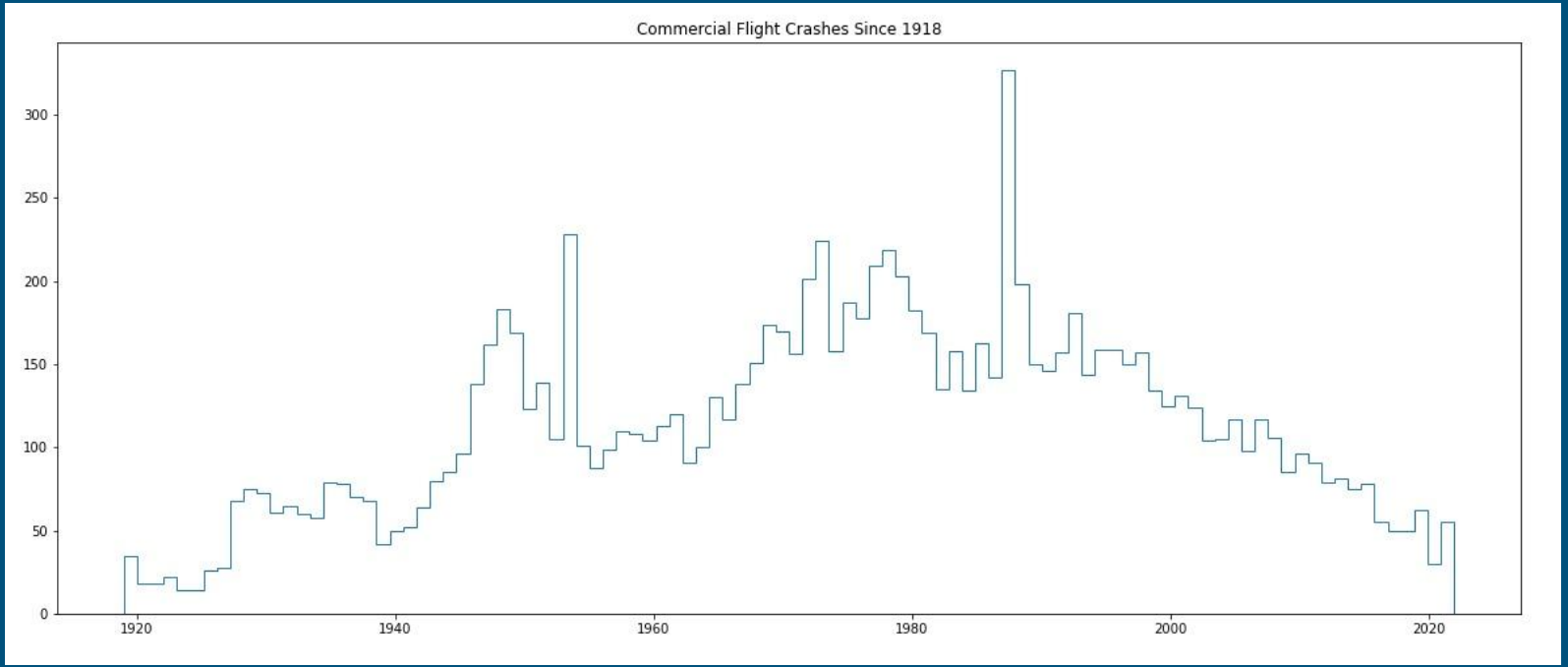
- **Result:** **73.5% accuracy**

Background and Approach

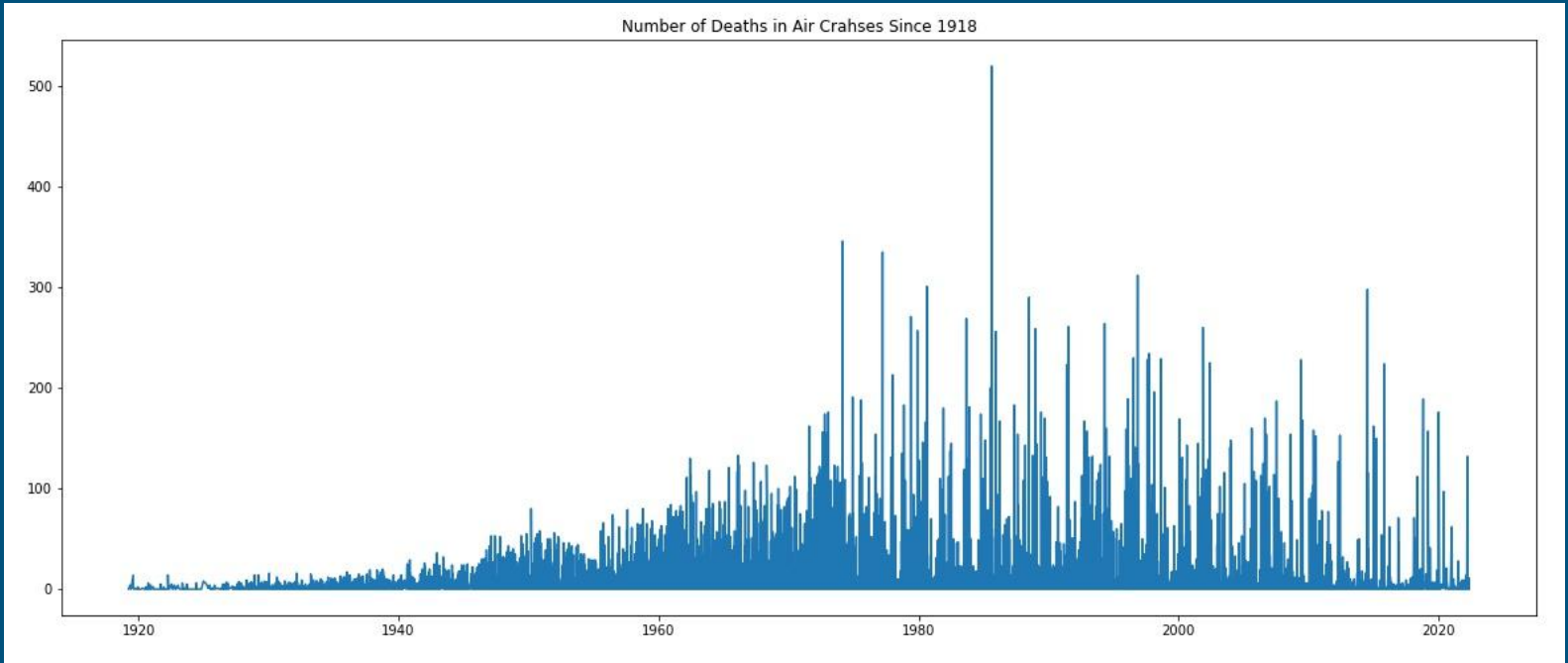
- Data includes all flight crashes since 1918.
- Based on ICAO, there are two flight types
  - Commercial flights
  - General aviation
- Method used: Logistic Regression, Knn, Random Forest, Neural Network
- Issued faced: missing values, data conflict, too many categories of aircraft.

Survivors	Crash site	Crash location	Country	Region	Circumstances	Crash cause	aircraft_age	people_on_board	fatal_on_board	num_survivor
0.0	Plain, Valley	Ostrogozhsk Voronezh oblast	Russia	Asia	The aircraft crashed in unknown circumstances ...	Unknown	43.0	0.0	0.0	0.0
1.0	Plain, Valley	Cobán Alta Verapaz	Guatemala	Central America	Crashed in unknown circumstances in bushes nea...	Human factor	NaN	0.0	0.0	0.0

Number of Commercial Flights Crashes

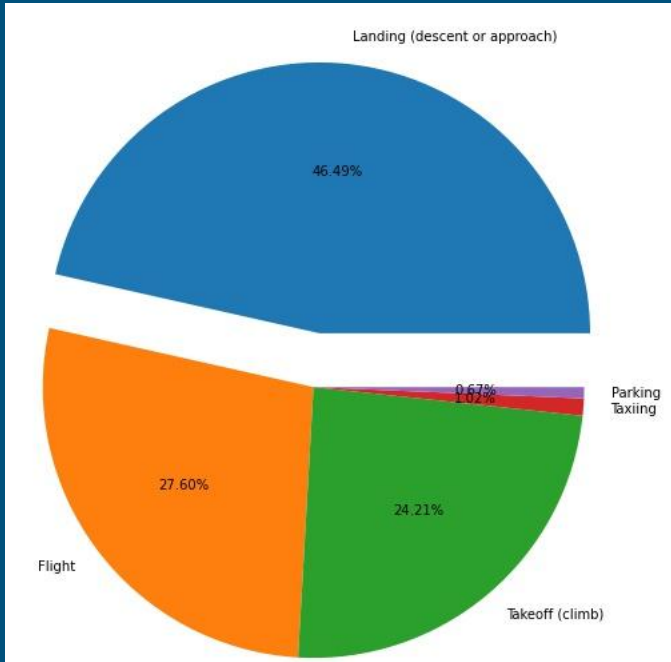


Number of Deaths in Commercial Flight Crashes

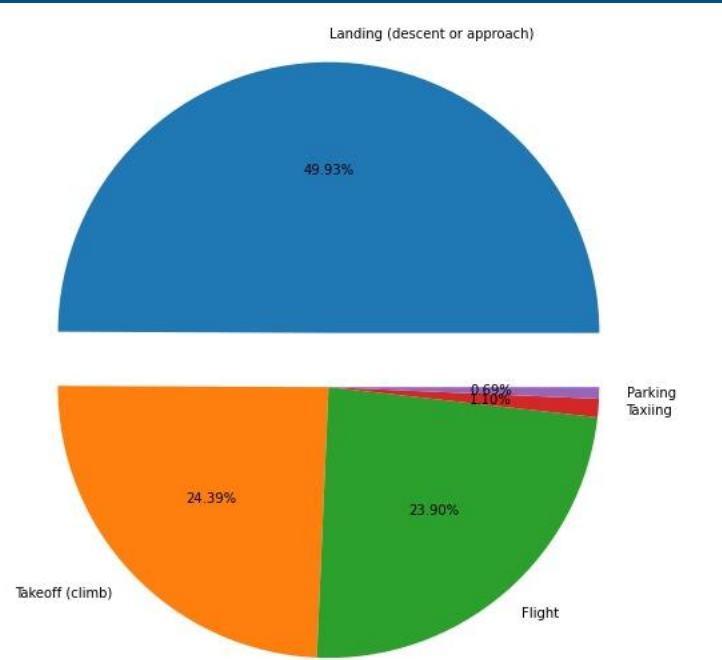


Distributions by Flight Phases

All Commercial Flights

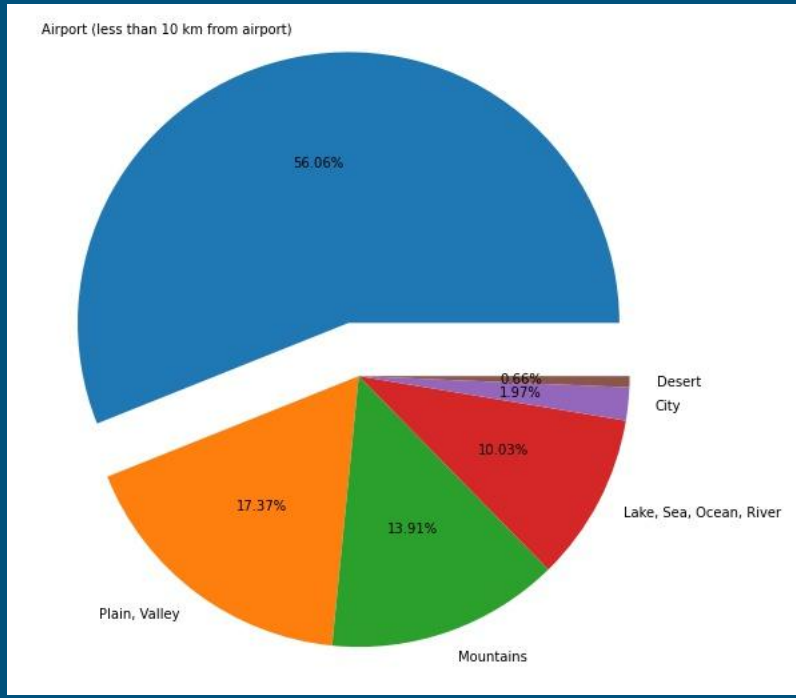


Commercial Flights in Jet Age

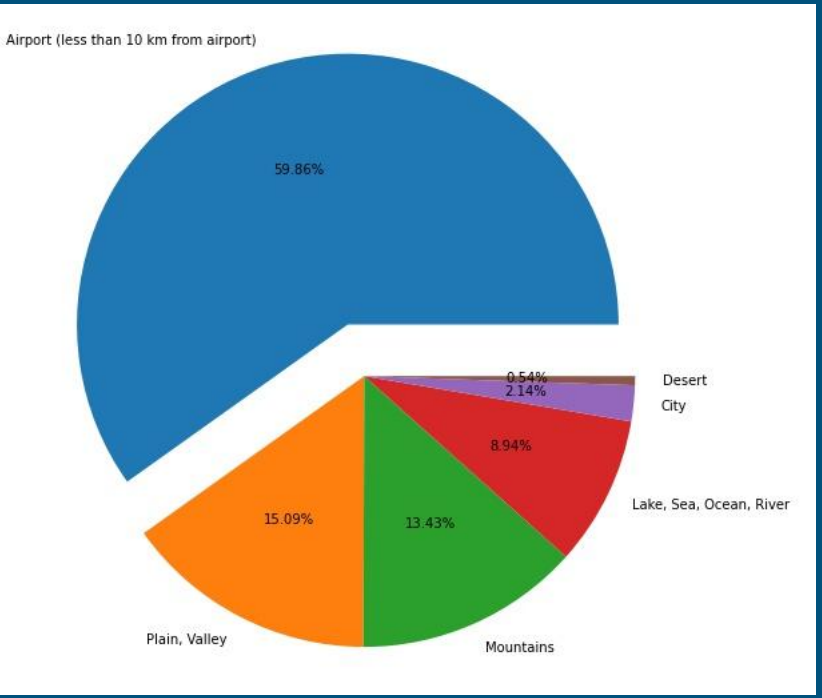


Distributions by Crash Locations

All Commercial Flights

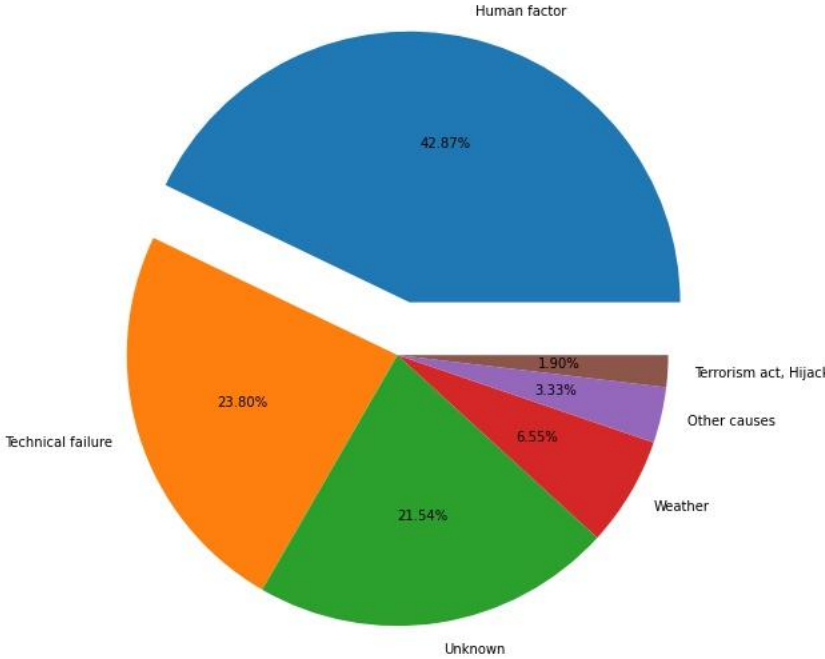


Commercial Flights in Jet Age

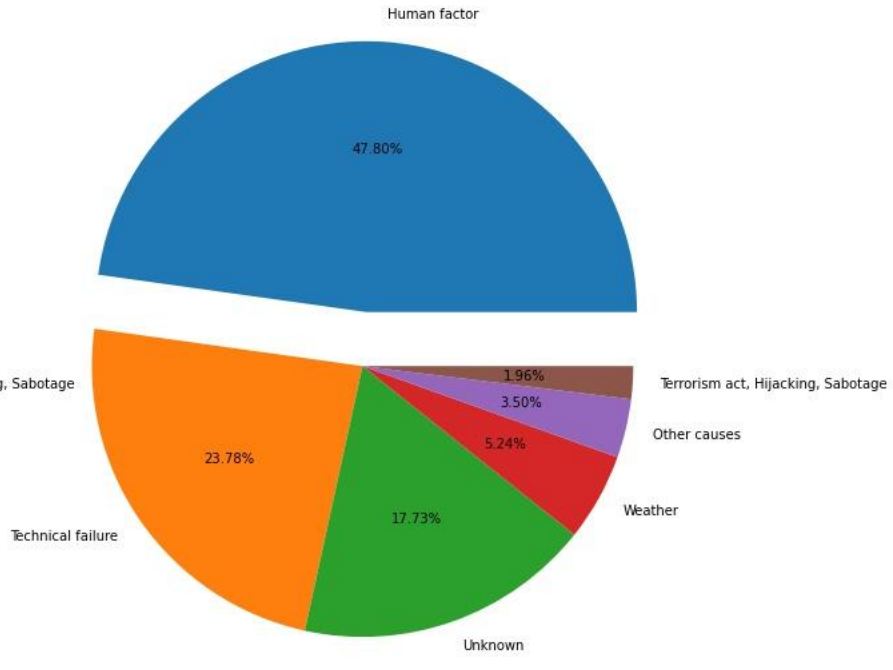


Distributions by Crash Causes

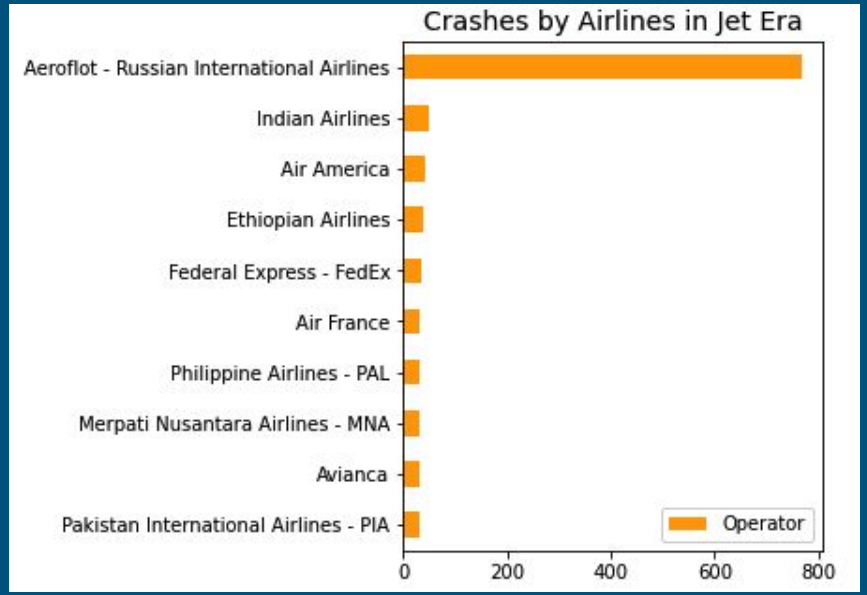
All Commercial Flights

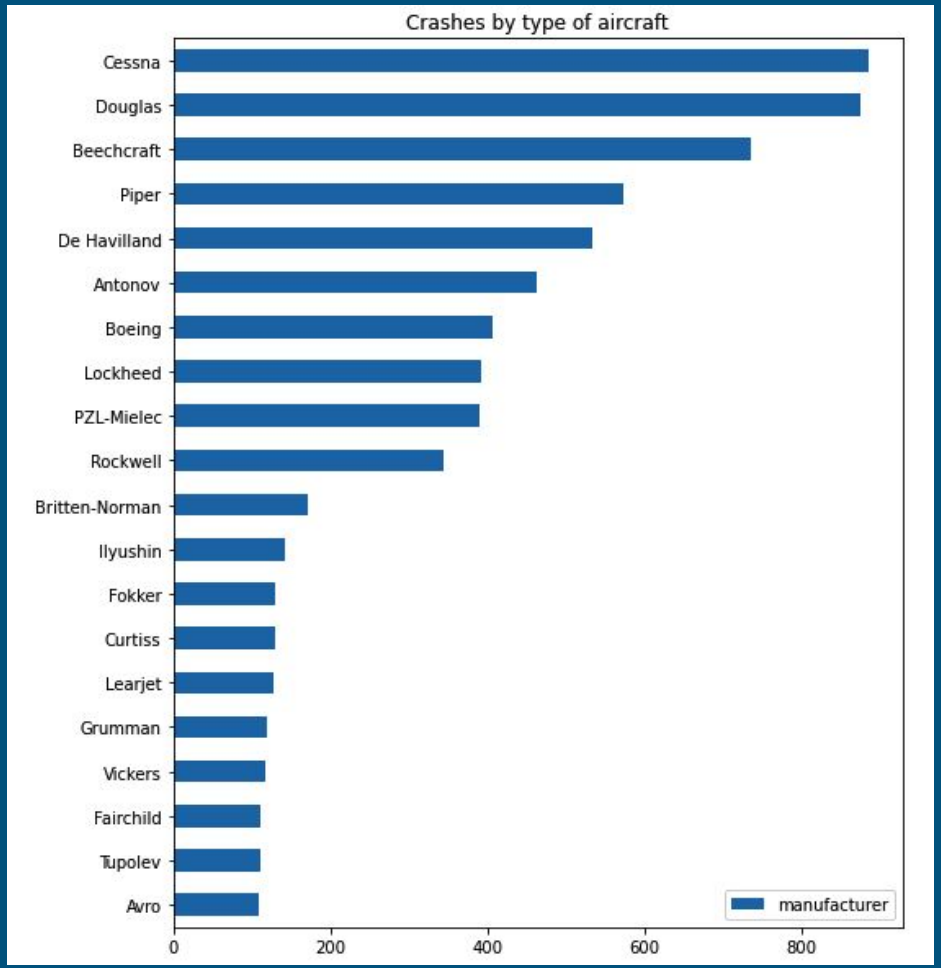


Commercial Flights in Jet Age

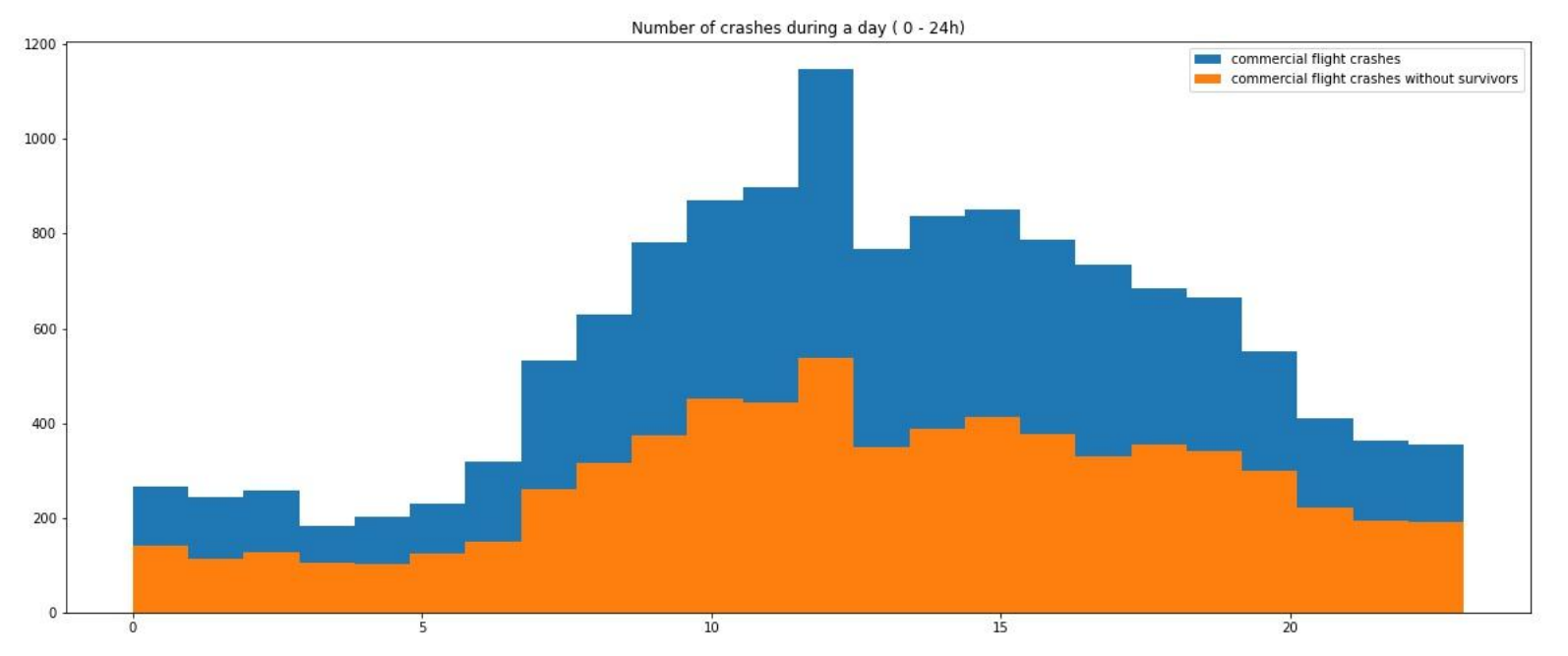


Crashes by Airlines and Aircraft Manufacturer



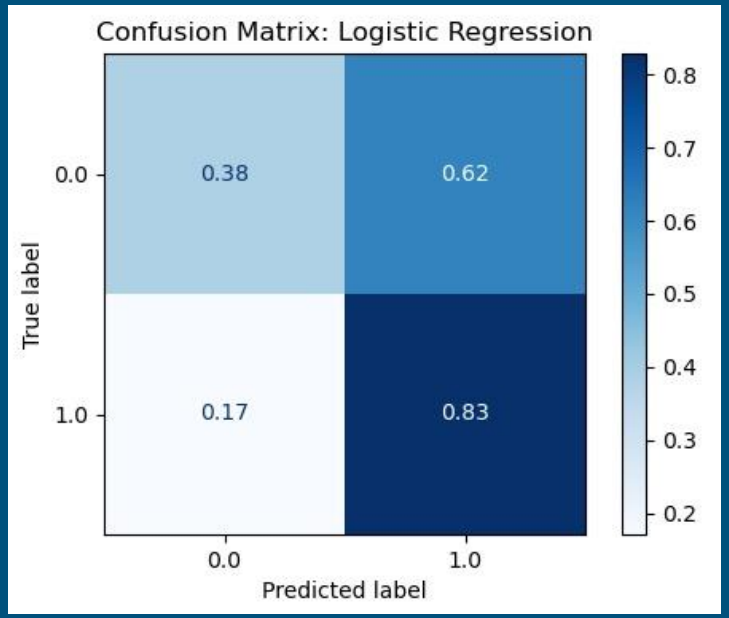


Crash Distributions by Day Range

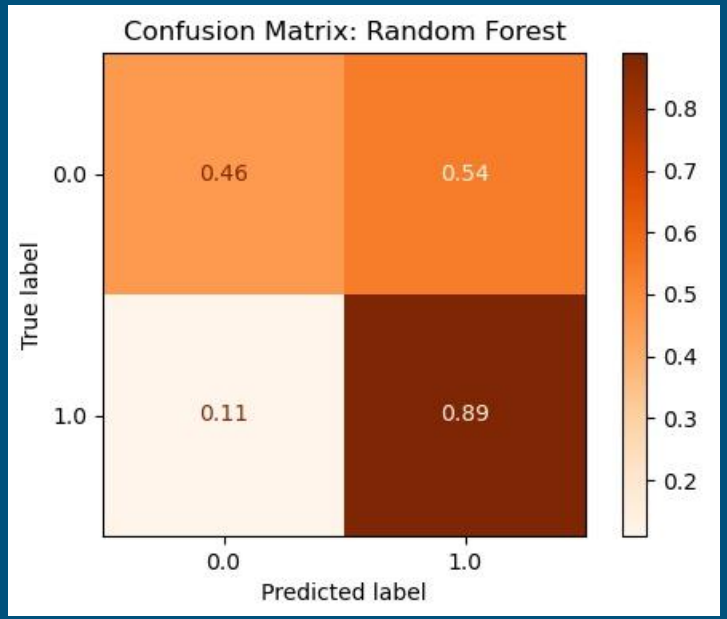


Classification Methods

Baseline Accuracy:
63.7%

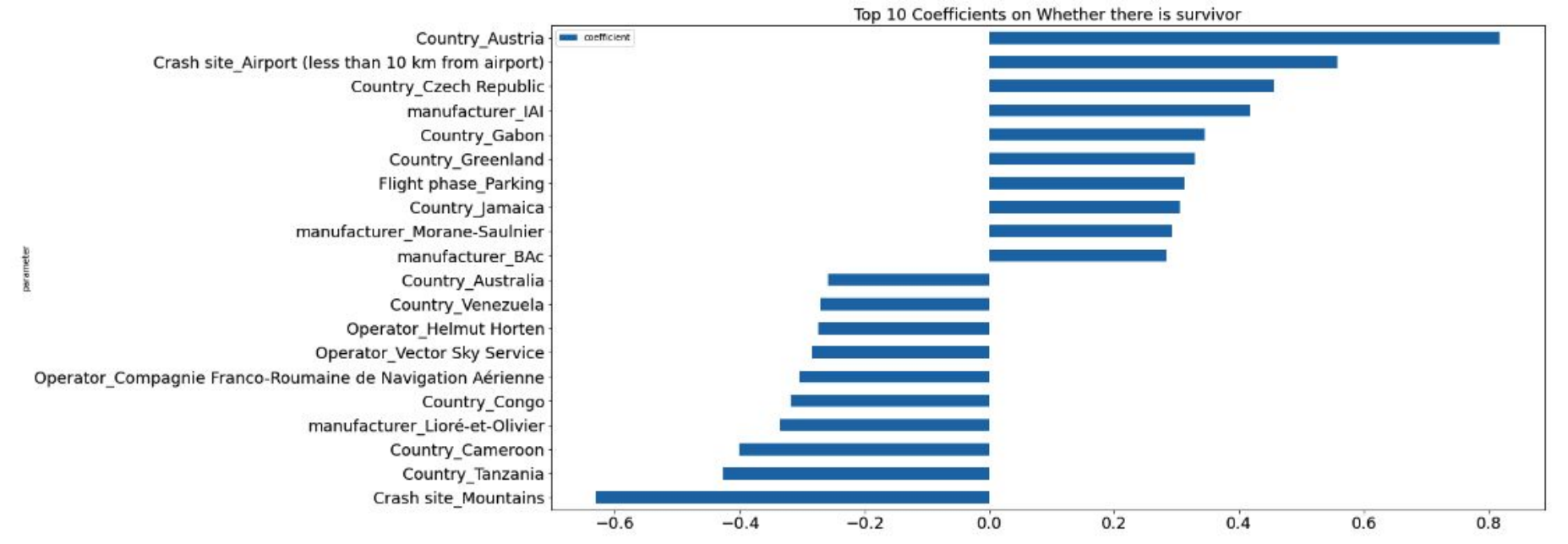


Accuracy: **66.7%**
Precision: **70.2%**

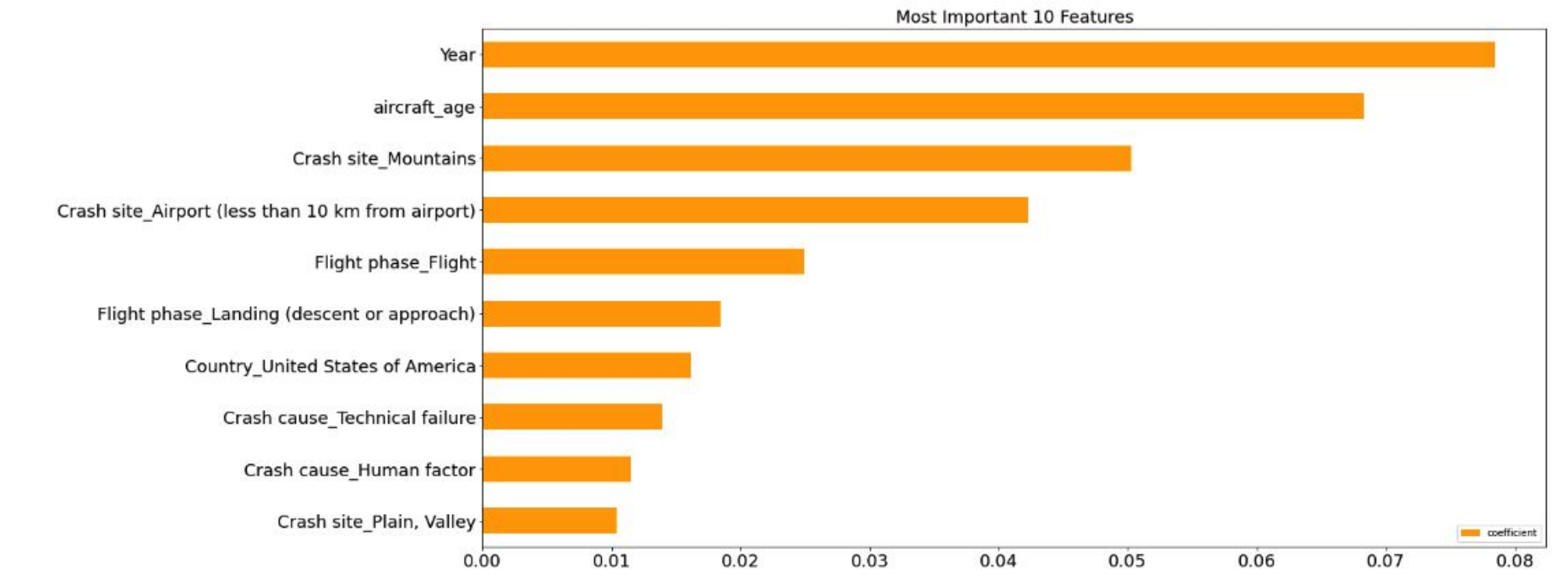


Accuracy: **73.5%**
Precision: **74.3%**

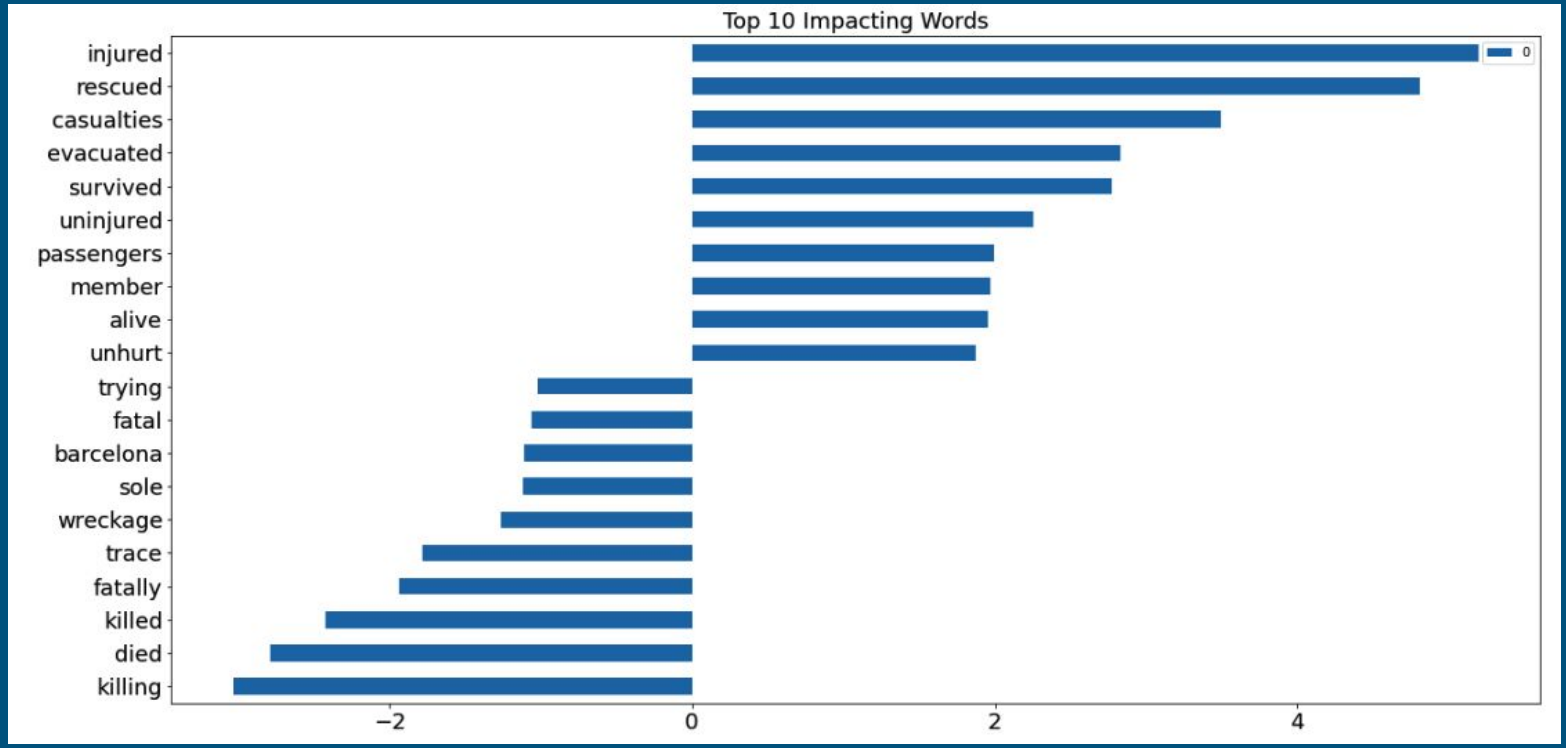
Top Coefficients



Feature Importance



NLP



Accuracy: **94.8%,** Precision: **96.4%**

Conclusions

- Air travel is still one of the safest transportation.
- It is difficult summarizing the major causes of one accident.
- It is difficult predicting whether an air crash would lead to no survivors.
- It is impossible to predict the safety level of any future flight.

Thank you