Plight or Flight

A view into SpaceX launch data

Brian Cohan

2023-11-07

EXECUTIVE SUMMARY

A series of Falcon 9 launches were analyzed to determine the likelihood of recovering the first stage booster. The analysis considers the effects of launch site, orbit trajectory, payload mass, and other factors.

Several Python libraries were used to perform this analysis. Scikit-Learn was used to train a model to predict the likelihood of successful recovery on future launches.

Based on the training data, the models were able to predict the correct outcome of booster recover 80% of the time.

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion

INTRODUCTION

Space flight is an expensive endeavor. Historically, rockets have been single-use until the Space Shuttle program allowed for reusable parts. However, the shuttles did not produce the intended savings and the boosters were still single use.

SpaceX revolutionized space exploration by finding way to make the first stage booster, which does most of the heavy lifting, recoverable.

This analysis reviews the trends of how successful the recovery has been which can help in determining a reasonable cost for future flights.

METHODOLOGY - Technology

A number of Python packages were used to support this analysis:

- Requests – Get API data and web scraping
- BeautifulSoup – Process HTML from web scraping
- Pandas – Data analysis and manipulation
- Scikit-Learn – Machine Learning
- Plotly – Graphing Library
- Folium – Geospatial Graphing Library
- Dash – Dashboarding tool used with Plotly

METHODOLOGY – Data Sources

SpaceX provides an open API where users may get information about launches. Information was gathered from the API as well as other sources.

- SpaceX API
  - https://api.spacexdata.com/v4/launches/past
  - https://api.spacexdata.com/v4/rockets/<rocket_id>
  - https://api.spacexdata.com/v4/launchpads/<launchpad_id>
  - https://api.spacexdata.com/v4/payloads/<payload_id>
  - https://api.spacexdata.com/v4/cores/<core_id>
- Other Sources
  - https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922

METHODOLOGY – Data Wrangling

Several steps were necessary in cleaning up the data for this analysis:

- Combine data from various sources
- Limit analysis to launches with a single rocket core
- Limit analysis to launches with a single payload
- Limit analysis to Falcon 9 launches
- Fill missing payload mass data with average
- Convert categorical data to dummy variables

METHODOLOGY – Exploratory Data Analysis

An exploratory data analysis (EDA) was performed. The purpose of the EDA is to get a sense of what is in the data. This will allow us to gain an intuition used to gauge if the predictive models are making sense. Some topics explored are:

- Launch pad usage
- Paylaod orbits
- Payload weight distribution
- Trend in successful recoveries
- Failure modes

METHODOLOGY – Predictive Analysis

The goal of the analysis is to predict if a booster recovery would be successful or not. This is a binary classification problem and the following methods were evaluated:

- Logistic Regression
- Support Vector Machine
- Decision Trees
- K-Nearest Neighbors

In each case, a Grid Search with a 10 fold cross validation was used to find the best parameters.

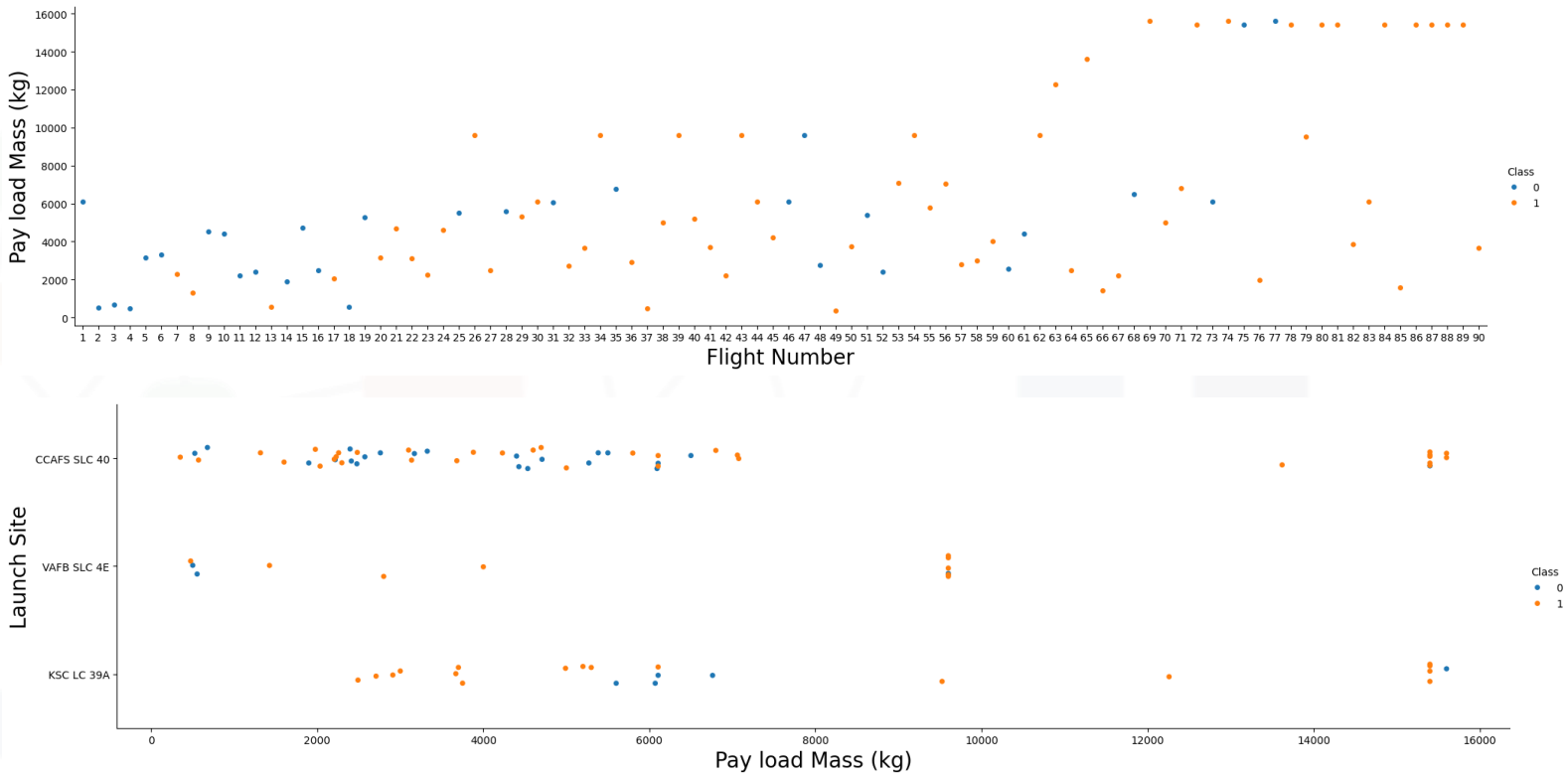
RESULTS – EDA Summary Info

- SpaceX uses 4 launch sites:
  - CCAFS LC-40
  - CCAFS SLC-40
  - VAFB SLC-4E
  - KSC LC-39A
- Average payload mass: 2534 kg
- First successful recovery
  - 2015-12-22

- Summary of recoveries:
  - 10 - No attempt
  - 5 - Success (ground pad)
  - 5 - Success (drone ship)
  - 5 - Failure (drone ship)
  - 3 - Controlled (ocean)
  - 2 - Uncontrolled (ocean)
  - 1 - Precluded (drone ship)
  - 1 - Failure (parachute)

RESULTS – EDA Graphical Insights

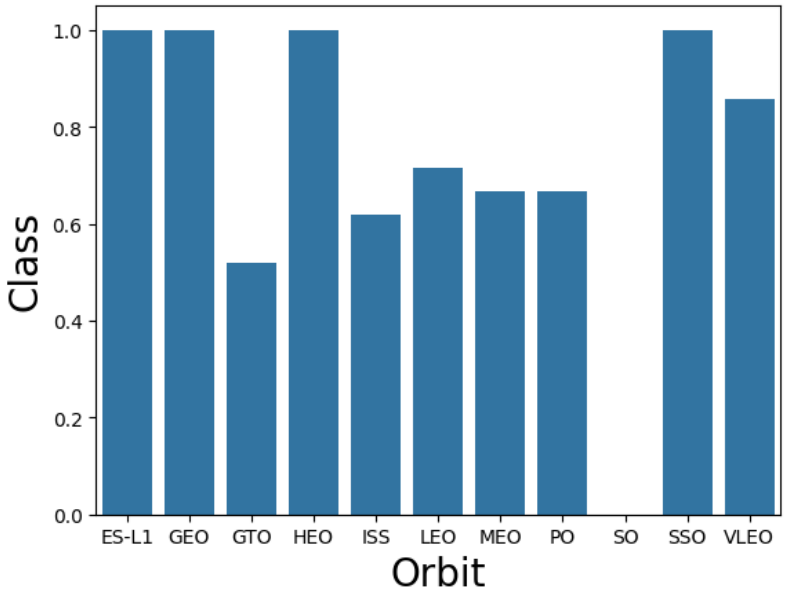
As the program matured, it's evident that not only were they more successful in recovering the first stage (indicted by orange data points) but SpaceX was willing to take on heavier payloads. The heavier payloads were also launched from either CCAFS or KSC.

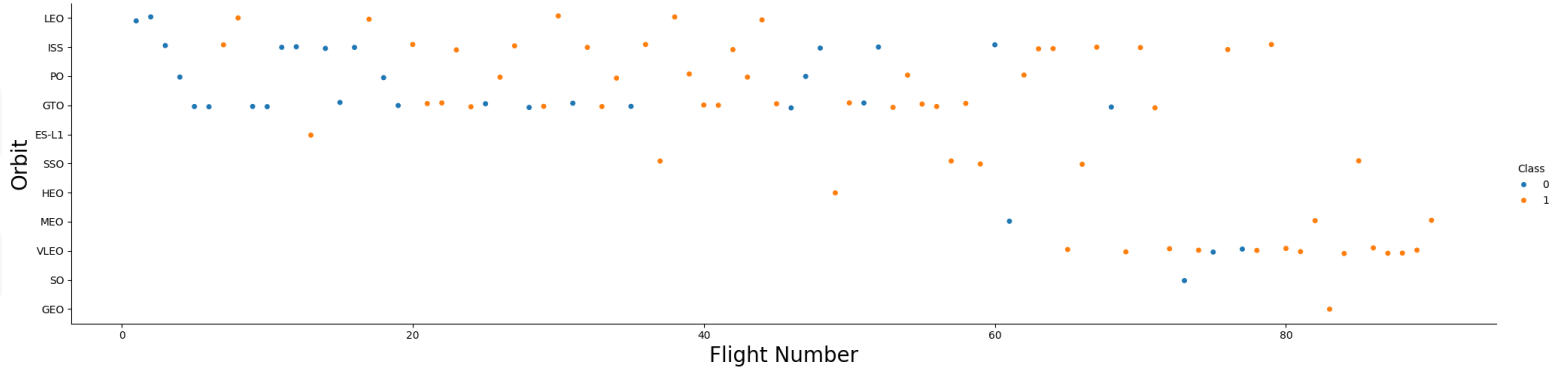


RESULTS – EDA Graphical Insights

Certain orbits have a better success rate. For example, ES-L1, GEO, HEO, and SSO have never had a failure in recovering the booster.

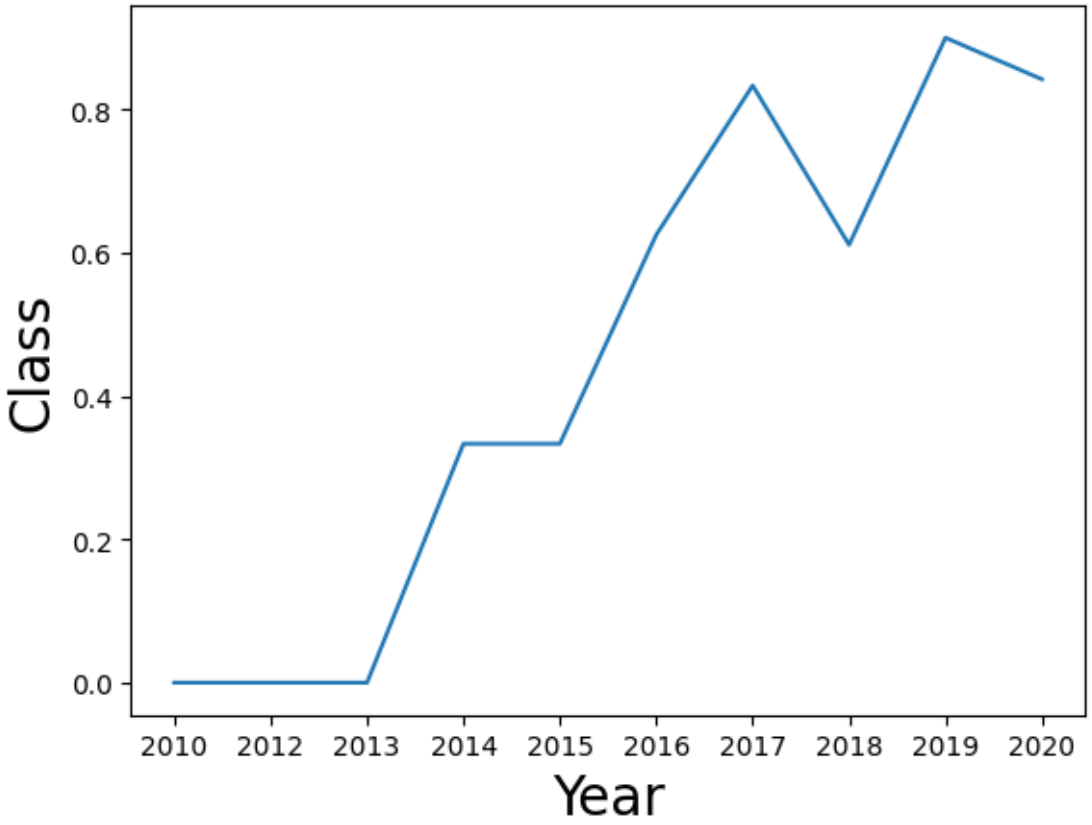
While these orbits have the highest success rate, they also are the least common and were not part of the initial launches where there were more booster recovery failures.





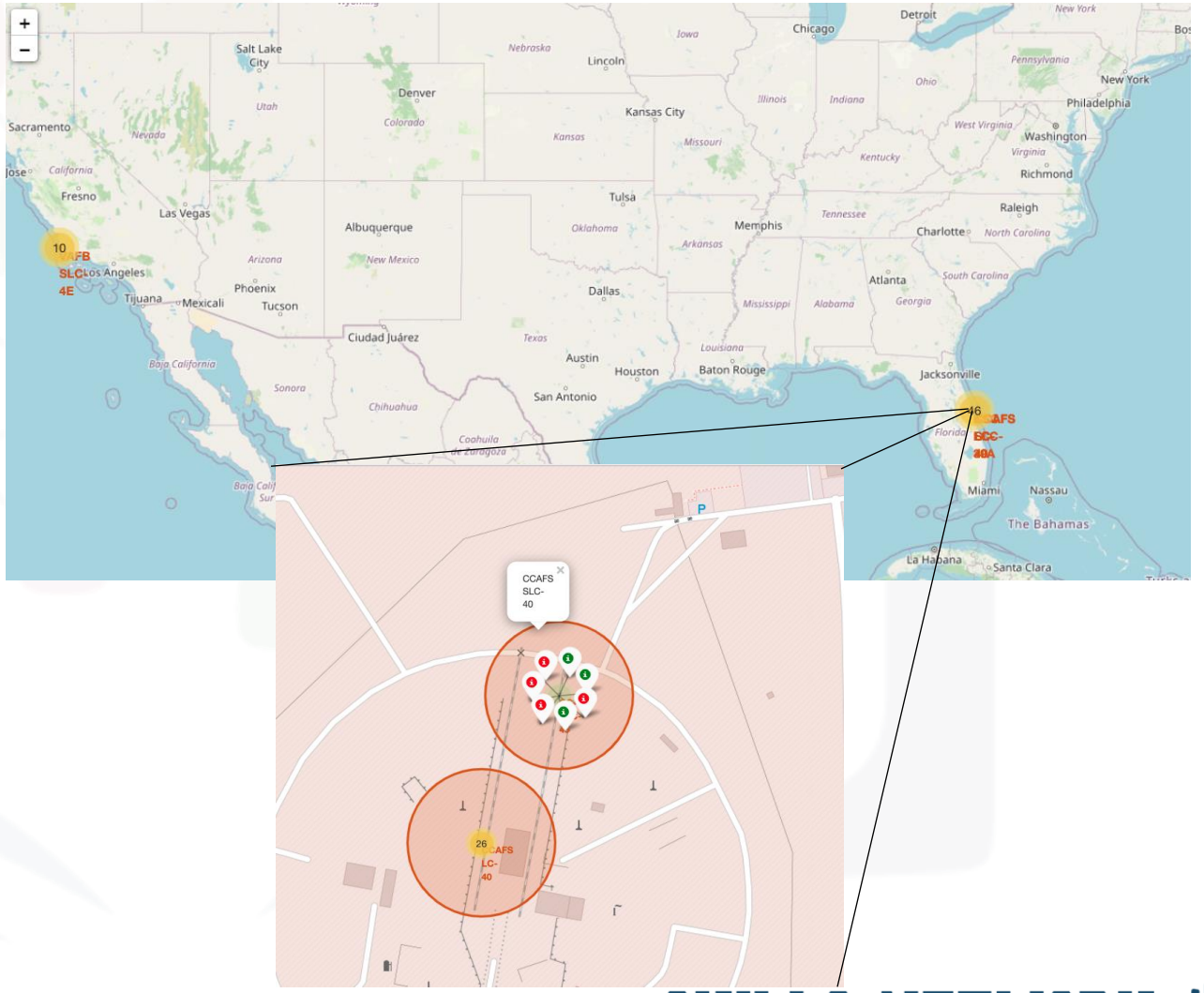
RESULTS – EDA Graphical Insights

This graph shows that as the program matured, the portion of successful recoveries has steadily increased with the exception of a slight dip in 2018.



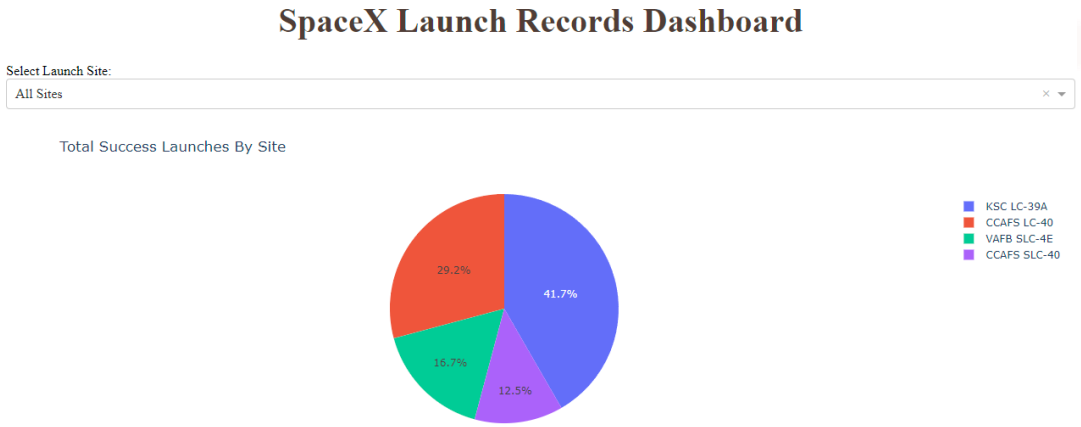
RESULTS – EDA Geospatial Insights

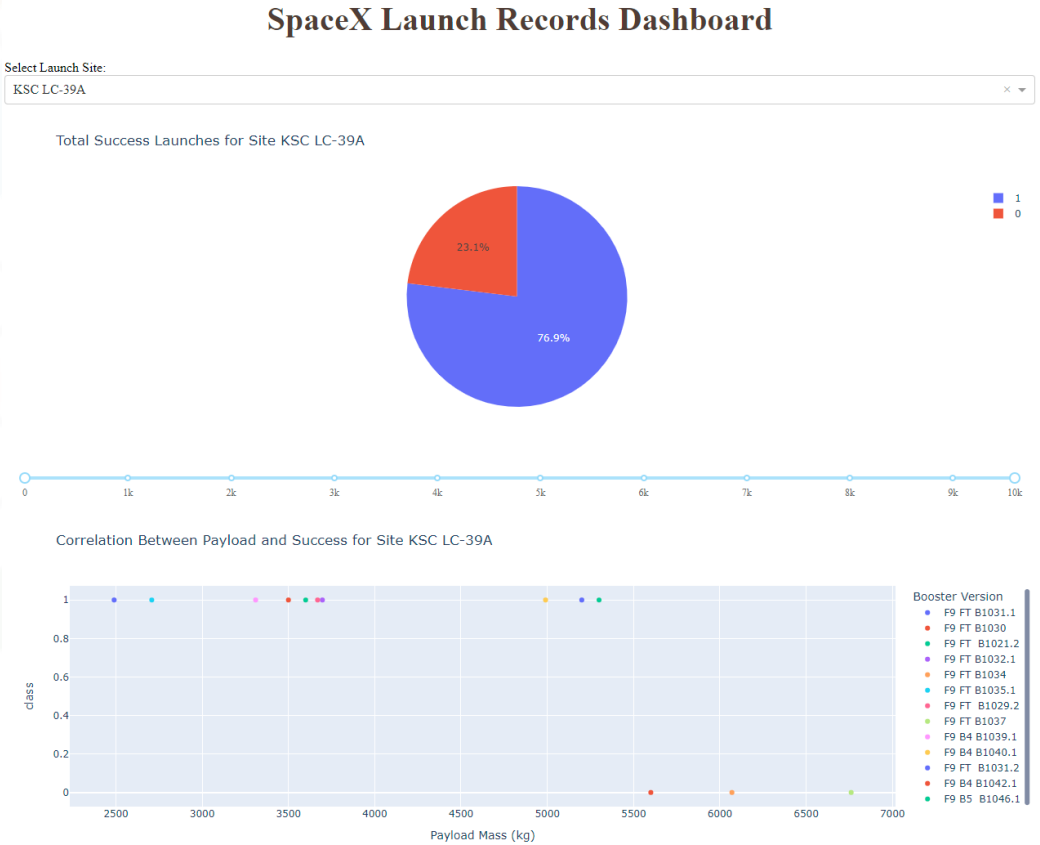
All the launches have taken place in either CA or FL. The majority have favored the east coast with a mix of successes (green markers) and failures (red markers)



RESULTS – EDA Dashboard

A dashboard was developed to examine success rates. It appears that KSC LC-39A has the highest success rate, though the failure included launches associated with the heavier payloads



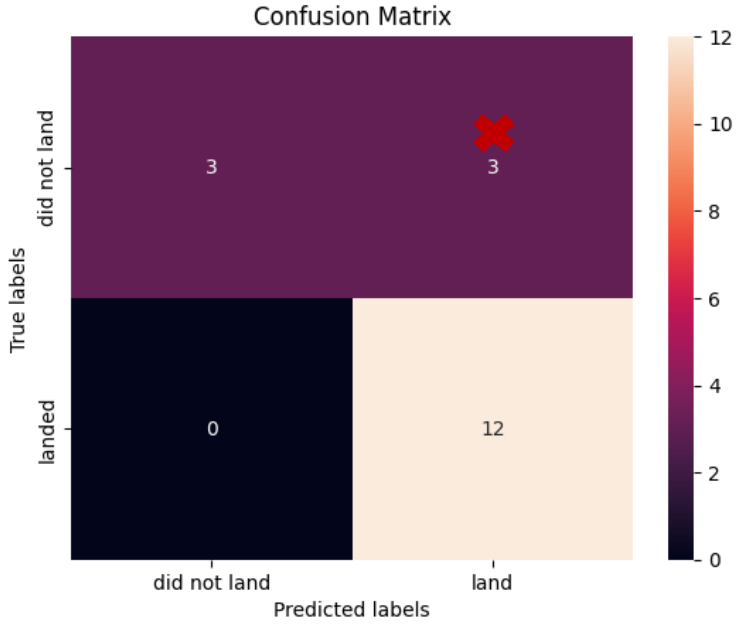


RESULTS – Predictive Analysis

Four different classification models were evaluated

- Logistic Regression
- Support Vector Machine
- Decision Tree
- K-Nearest Neighbors

All four models yielded the same results with a high level of precision. There were three cases where each model predicted a successful landing where the boosters did not land.



	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

CONCLUSION

- As the booster recovery program matured, the successful recovery rate increased – This appeared to be the biggest factor in success
- Successes appeared to be independent of launch site, orbit, or payload when viewed from
- The classification models all yielded the same results, likely due to the small dataset
- Based on visible trends, future launches can be priced assuming successful recoveries of the first stage booster