SpaceX Falcon Launch Analysis

Keerthi

7-07-2023

OUTLINE

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

IBM **Developer**

SKILLS NETWORK

EXECUTIVE SUMMARY

- This report contains an analysis of the success rate of landing the first stage of the falcon9 rocket by SpaceX.

- We collected data and performed the following :
  - Data cleaning
  - Descriptive statistics
  - Data visualisation
  - Feature Extraction

- Later, We built a classification model using various algorithms

- Model evaluation and feedback.

INTRODUCTION

- Problem Statement: SpaceX released the news of launching their new falcon9 with a budget of 65 million, which usually takes 168 million. We want to predict the success rate of the first stage of the rocket landing safely.

- Key questions to be answered :
  - What are the factors that influence the target variable?
  - Which launching site has more success rate of first-stage recovery
  - What is the criterion for selecting a launching site?
  - Effect of payload?
  - Which booster version and booster version category is much more suitable?

METHODOLOGY
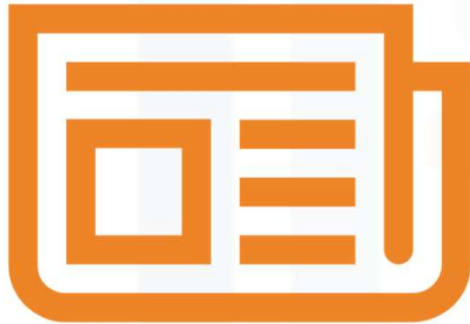
- Collected the data from Wikipedia by web scraping and spaceX rest API.

- The data initially contained unnecessary data, which was removed later during data wrangling.

- The problem requires a binary event's outcome so that the classification model would do the task.

- The analysis consists of looking into the map of the launch sites and the reason behind their selection.

- Analysis also includes which combination of booster version, launch site, and payload range is best suited.

RESULTS

- Used libraries such as Beautiful Soup to extract data from a website.
- We pandas, numpy to manipulate data and to study the descriptive statistics of data.
- Used SQL to query the data and gain insights.
- Used seaborn, matplotlib libraries to visualise the data and folium for geospatial data.
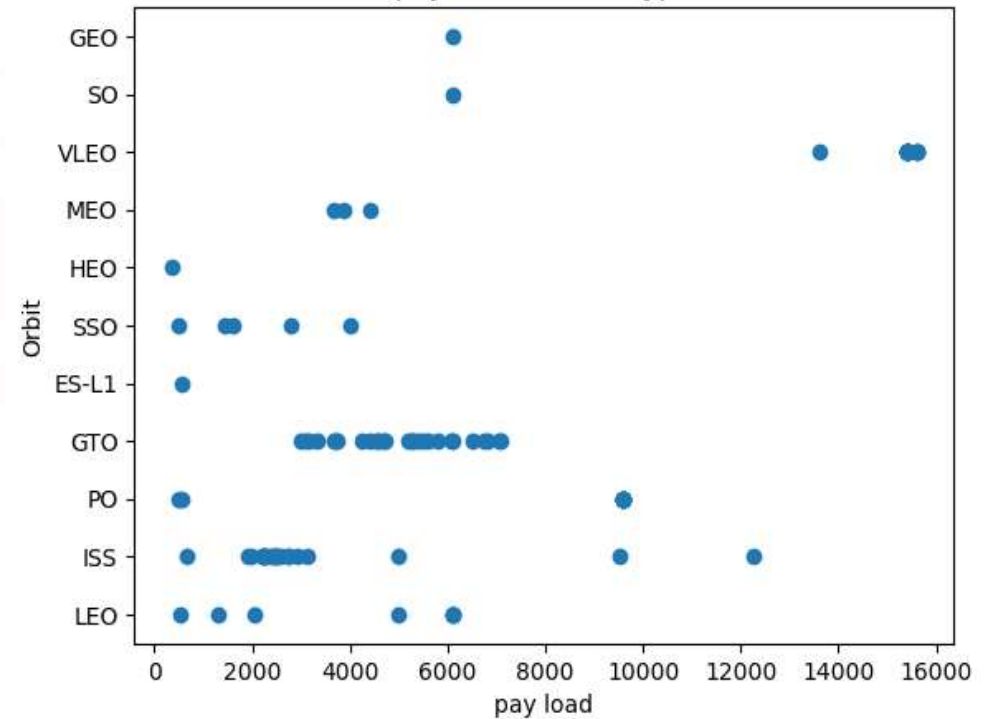- Used dash for the interactive dashboard.
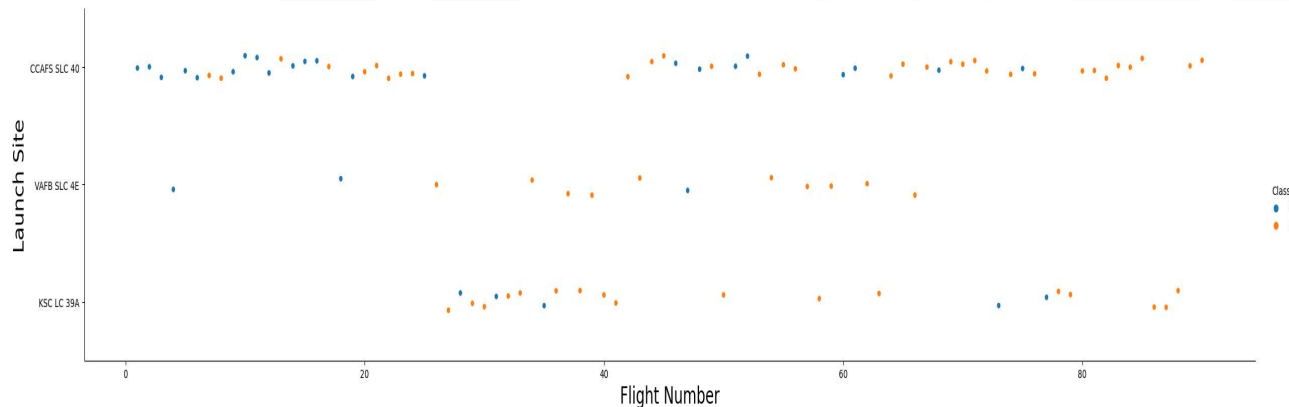
This graph shows the distribution of flights in different launch sites and their success rates.

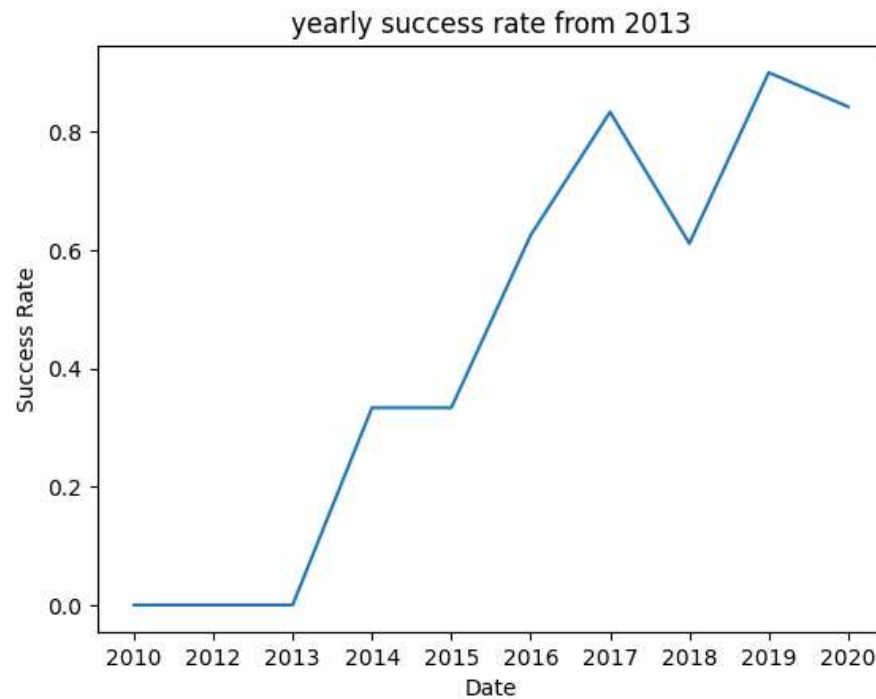# INFORMATION OF DIFFERENT ORBITS (DESTINATION OF ROCKETS)

Success rate of different orbitsThe success
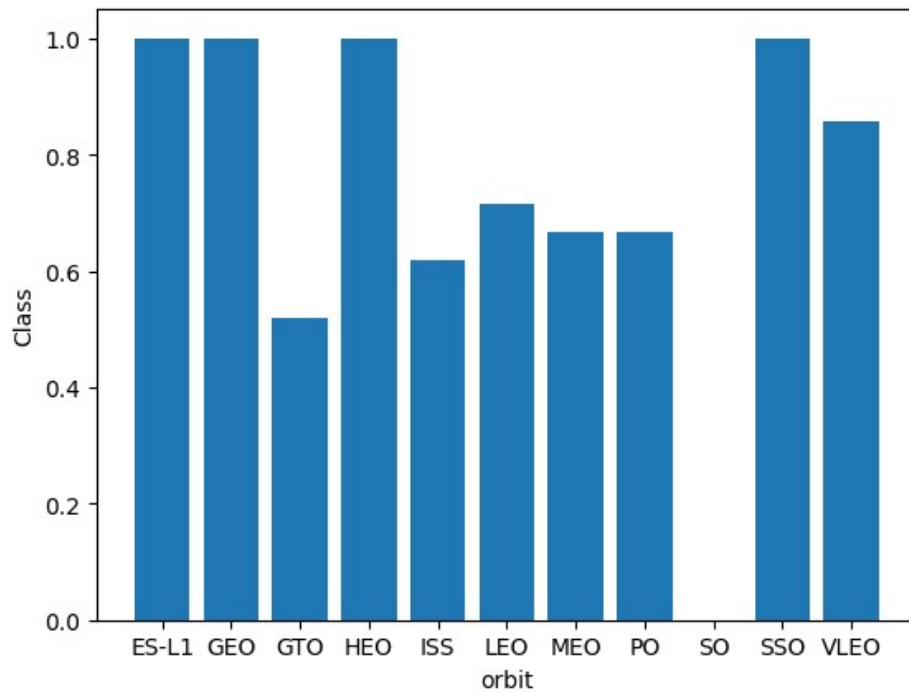
payLoad vs orbit type

SOME MORE FINDINGS FROM THE DATA

yearly success rate from 2013

Location of the launch sites

Most of the sites are located near the sites.

DATABASE TRENDS

The launch sites are close to the railways and the highways.

Most the launching sites are located near the coast and far from the city.

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings :

- From the year 2013, the success rate has rapidly increased.

- Most of the launch sites are near the coast sites.

- ES-L1, GEO, HEO, SSO orbits have a high success rate.

- FT booster version releases more payload rockets.

- The features that influence the target variable are booster version, booster version Category, payload mass, Launch site.

DASHBOARD

https://sankojukeert-8050.theiadocker-2-labs-prod-theiak8s-4-tor01.proxy.cognitiveclass.ai/
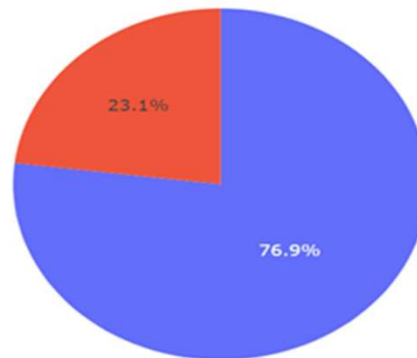
The above link contains dashboard link

DASHBOARD TAB 1

SpaceX Launch Records Dashboard
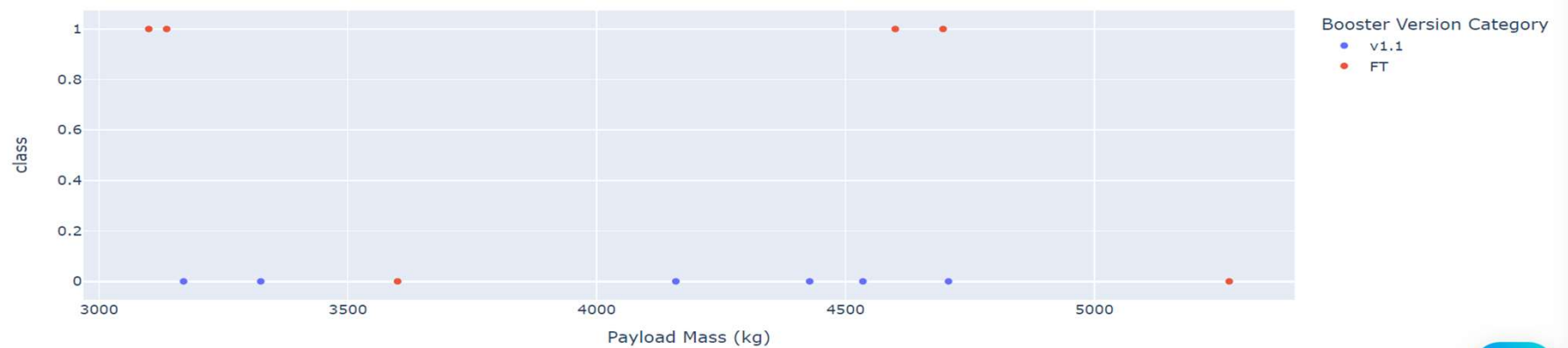
KSC LC-39A

Success and Failed Launches - KSC LC-39A

- Success
- Failed

23.1%

76.9%

O. Ask

DASHBOARD TAB 2

DASHBOARD TAB 3

Screenshot of dashboard tab 3 goes here

DISCUSSION

Initially we have to select a perfect launch site

Then having a suitable payload increases the probability of landing the first stage.

Selecting the right booster version

Recover the first stage
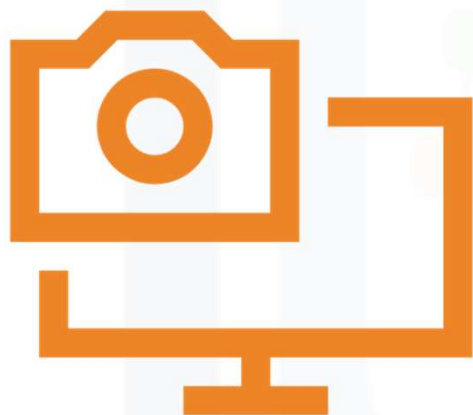
Launching the rocket
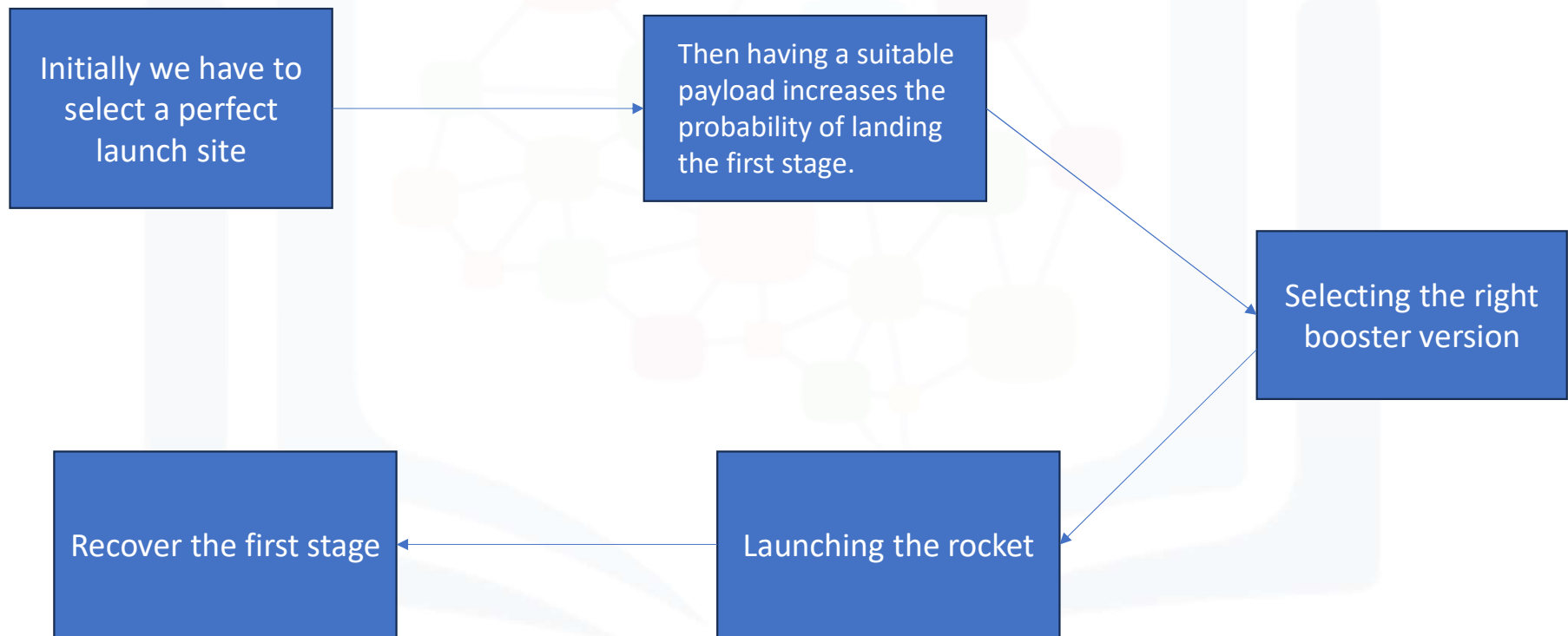
CONCLUSION

- We have analysed the data and built a machine-learning model that predicts whether or not the newly launched Falcon 9 with recover its first phase.

- From this information, we can plan our upcoming events to give tough competition to spaceX.

- Necessary steps to be taken by the concerned teams.

APPENDIX

- Model scores:
  - Logistic regression – 0.8333
  - Decision Tree classifier – 0.7777
  - KNN – 0.61111