



Predictive Analysis of SpaceX Launch Outcomes

Applied Data Science Capstone Project

Presented by: Tuomas V

Date: 16th July 2024

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Discussion
- Conclusion

Executive Summary

- **Summary of methodologies:** Data collection from API and webscraping, Exploratory analysis with SQL and visualization, Machine learning prediction
- Objective was to analyze and predict the outcomes of SpaceX Falcon 9 rocket launches.
- Key factors influencing launch success were identified, and predictive models were developed.
- Successful prediction of launch outcomes with insights for future improvements were found.

Introduction

- **Background:** SpaceX aims to make space travel more cost-effective and reliable.
- **Problem Statement:** Predicting the success of Falcon 9 launches based on historical data.
- **Goals:** Improve prediction accuracy and identify factors correlating with launch success.

Methodology

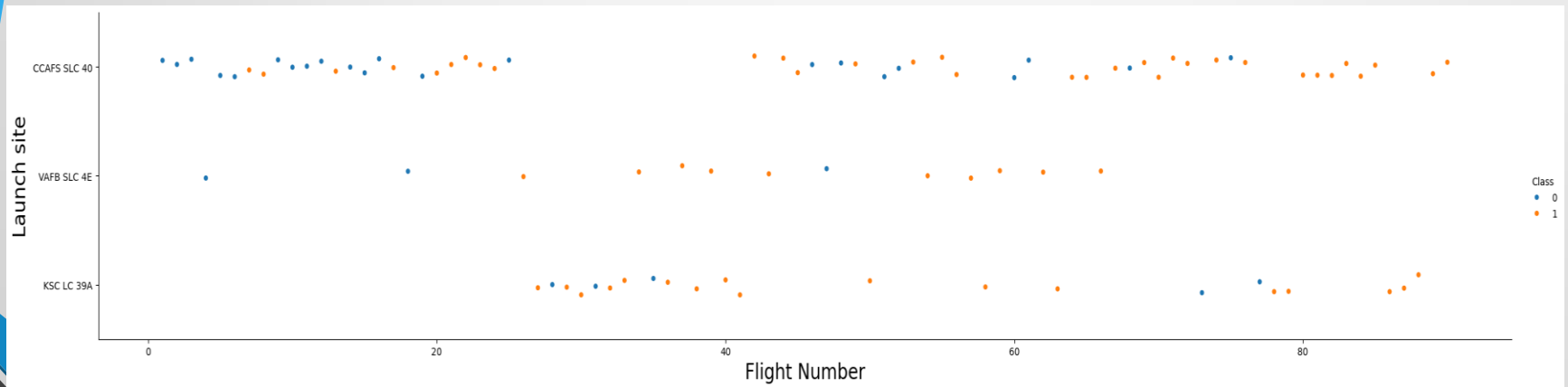
- **Data Collection:** SpaceX API and web scraping.
- **Data Wrangling:** Cleaning and preprocessing using Python libraries.
- **Exploratory Data Analysis (EDA):** Visualizing data trends and distributions.
- **Machine Learning Models:** Logistic Regression, Decision Trees, Support Vector Machines (SVM).
- **Tools:** Jupyter Notebooks, Pandas, Matplotlib, Seaborn, Scikit-Learn, SQL, Github.

Results

- **Model Performance:** Evaluated using accuracy, precision, recall, and F1 score.
- Best Model could not be reliably determined as all models tested gave the same accuracy result, likely due to the relatively small sample size.
- Significant factors correlating to success include launch site, payload mass, and booster version.

Launch success by Launch Site

- From this graph we can see Falcon 9 launches by launch site. The orange dots indicate successful launches and blue ones unsuccessful ones.



Further results

- Further results can be found on Github:

<https://github.com/hmllss/Applied-Data-Science-Capstone>

Discussion

- Launch site and payload mass appear to be critical factors for predicting success.
- **Challenges:** Dealing with the sample size being relatively small and handling missing data and model overfitting.
- In the future repeating the steps with a larger sample size would be useful and integrating additional data sources and enhancing model complexity.

Conclusion

- Success rate of Falcon 9 launches has increased over time.
- KSC LC-39A is the most successful launch site.
- Heavier payloads have a higher success rate though it may be due to payload being increased alongside other development in reliability.