

Title Slide

Predicting Falcon 9 First Stage Landing Success

Joe Bryant

June 11, 2024

Outline

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

Executive Summary

- Predict the success of Falcon 9 first stage landings using machine learning.
- Steps include data collection, wrangling, EDA, data visualization, and machine learning prediction.
- Decision tree algorithm may be the best for prediction.

Introduction

- SpaceX advertises Falcon 9 launches at \$62 million.
- Predicting landing success helps determine launch costs.
- Main question: Given features of a launch, will the first stage land successfully?

Data Collection and Wrangling Methodology

- Data Collection:
 - - SpaceX API: <https://api.spacexdata.com/v4/rockets/>
 - - Web Scraping: Wikipedia for Falcon 9 launches
- Data Wrangling:
 - - Replaced missing values with column means.
 - - One-hot encoding for categorical features.
 - - Added 'Class' column (0 for failure, 1 for success).
 - - Final dataset: 90 rows, 83 columns.

Exploratory Data Analysis and Visualization Methodology

- EDA:
 - - Pandas and NumPy for basic data statistics.
 - - SQL for querying data insights.
- Data Visualization:
 - - Matplotlib and Seaborn for scatterplots, bar charts, line charts.
 - - Folium for interactive maps.
 - - Dash for interactive dashboard analytics.

Predictive Analysis Methodology

- Standardizing data.
- Splitting data into training and test sets.
- Models used: Logistic Regression, SVM, Decision Tree, KNN.
- Evaluated models based on accuracy and confusion matrix.

EDA with Visualization Results

- Flight number vs. launch site.
- Payload mass vs. launch site.
- Success rate by orbit type.
- Flight number vs. orbit type.
- Payload mass vs. orbit type.
- Yearly launch success trend.

EDA with SQL Results

- Unique launch sites.
- Payload mass by NASA (CRS).
- Average payload mass for F9 v1.1.
- First successful landing on ground pad.
- Booster success on drone ship with payload 4000-6000 kg.
- Total mission outcomes.
- Boosters with maximum payload mass.
- 2015 failed drone ship landings.
- Landing outcomes (2010-2017).

Interactive Map with Folium

- Folium map showing launch sites and outcomes.
- Interactive features highlighting successful and failed landings.

Plotly Dash Application

- Dash dropdown for selecting launch sites.
- Range slider for selecting payload mass.
- Interactive pie chart for success rate by site.
- Interactive scatter plot for payload vs. success rate.

Predictive Analysis Results

- Logistic Regression:
 - - Best score: 0.846
 - - Accuracy: 0.833
- SVM:
 - - Best score: 0.848
 - - Accuracy: 0.833
- Decision Tree:
 - - Best score: 0.889
 - - Accuracy: 0.833
- KNN:
 - - Best score: 0.848
 - - Accuracy: 0.833
- Model comparison and ranking.

Conclusion

- Predicting first stage landing success helps determine launch costs.
- Each feature impacts the mission outcome.
- Decision Tree is the best model for prediction.

Creativity and Insights

- Application of interactive visual analytics.
- Innovative use of Folium and Dash for real-time data analysis.
- Insights beyond required tasks: comparison of multiple machine learning models.

Questions

- Thank you!