# Capstone Presentation

SpaceX Falcon 9 Launch Success Prediction

Data Science Capstone Project

Author: Ashish Anand

# Executive Summary

This project analyzes SpaceX Falcon 9 launch data to predict successful landings.

Data was collected using API and web scraping, followed by exploratory data analysis and machine learning.

Final classification accuracy achieved: 83.33% using SVM (linear kernel).

# Introduction

SpaceX aims to reduce launch costs with reusable rockets.

Predicting landing outcomes helps optimize resources.

We use data science techniques to forecast success of first stage landings.

# Data Collection and Wrangling

Launch data was collected using the SpaceX REST API and Wikipedia (using BeautifulSoup).

Missing and inconsistent data were cleaned.

Multiple dataframes were merged and processed for further analysis.

# EDA & Visual Analytics Methodology

Performed EDA using pandas, matplotlib, seaborn, and plotly.

Analyzed relationships between payload, site, orbit, booster version, and success.

Created charts like scatter plots, bar charts, box plots, and heatmaps.

# Predictive Analysis Methodology

Selected features: payload mass, orbit, site, reuse count, etc.

Applied one-hot encoding and feature scaling.

Used Logistic Regression, SVM, Decision Tree, and KNN classifiers.

Hyperparameter tuning used GridSearchCV.

# EDA with Visualization Results

- Scatter plot: Payload vs Success

- Bar plot: Launch Site vs Success Rate

- Box plot: Booster Version vs Outcome

- Heatmap: Feature correlations

- Flight number trends and class distribution

# EDA with SQL Results

- Total launches per site

- Max and min payload mass

- Launch counts by booster version

- Mission outcome counts per site

- SQL queries used: SELECT, GROUP BY, ORDER BY, WHERE, LIMIT

# Interactive Map with Folium

Created an interactive map showing launch locations using Folium.

Used MarkerCluster to group close coordinates.

Each marker displays site name and launch outcome.

# Plotly Dash Dashboard

Built a dashboard using Plotly Dash.

Interactive dropdown filters for launch site and payload range.

Live-updating graphs show success rate and class distribution.

# Predictive Analysis Results

Support Vector Machine (linear kernel) performed best with 83.33% accuracy.

Included confusion matrix and accuracy scores for all models.

SVM was chosen based on validation results and simplicity.

# Conclusion

This project showed that data science can predict rocket landing outcomes.

SVM model achieved high accuracy using limited features.

Future work: include weather, flight time, and additional metadata.

## Creativity and Innovation

Enhanced visuals and structure beyond the template.

Used interactive maps, dashboards, and customized charts.

Uncovered insights like low-payload rockets having higher success.