SpaceX Falcon 9 First Stage Landing Prediction

Forecasting rocket landings for reusable spaceflight

Executive Summary

Project Highlights

Landing Prediction Purpose

The project predicts Falcon 9 first stage landing success to optimize costs and improve aerospace planning.

Machine Learning Model

Logistic Regression model achieved a high R^2 score of 0.94 after tuning for accurate landing forecasts.

Deployment and Accessibility

The solution was deployed as an interactive web app on Streamlit Cloud for easy user access.

Strategic Impact

Predictive insights help competitors estimate launch costs and plan missions more effectively.

Project Overview

Cost Advantage and Strategic Importance

Launch Cost Reduction

Reusable rocket stages reduce launch costs dramatically, making space missions more affordable and viable.

Economic Impact of Reusability

Recovering and reusing rocket parts accounts for 50-60% cost savings in total rocket expenses.

Predictive Analytics Role

Predicting landing success is crucial for cost estimation and strategic decision-making in aerospace missions.

Competitive Strategy

Accurate predictions enable companies to benchmark, optimize planning, and allocate resources effectively.

Business Problem & Objective

Challenges and Goals

Rocket Reusability Impact

Reusing the first rocket stage reduces costs and increases mission frequency in the space industry.

Complexity of Prediction

Landing success prediction depends on parameters like orbit type, payload mass, and launch conditions.

Business Problem and Objective

Accurate predictions optimize launch costs and mission planning to reduce financial risks.

Strategic Benefits

Predictive modeling supports informed decisions and enhances operational efficiency in space missions.

Methodology

Data Collection and Modeling Approach

Data Collection and Storage

Data was gathered from APIs and web scraping, then stored efficiently in a structured SQLite database.

Data Wrangling and Feature Engineering

Data cleaning fixed types, removed irrelevant info, and created a target 'Class' column for success classification.

Exploratory Data Analysis (EDA)

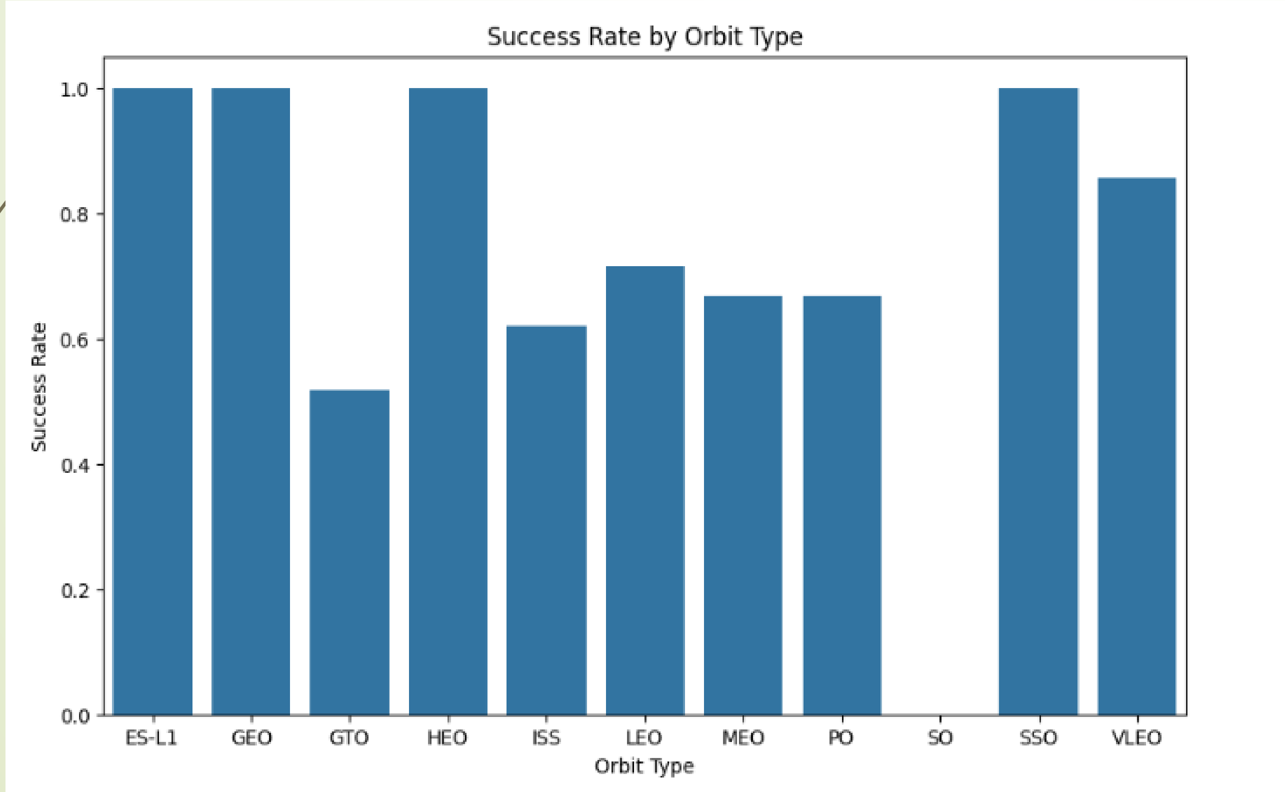
Visualization of trends and launch locations was performed using Seaborn, Matplotlib, and Folium tools.

Model Development and Deployment

Multiple algorithms were tested; Logistic Regression tuned with GridSearchCV was deployed via a Streamlit web app.

EDA Insights

Key Observations from Data Analysis



ORBIT	SUCCESS RATE (%)
ES-L1	100
GEO	100
HEO	100
SSO	100

Key Observations from Data Analysis

Maximum payload carried by Falcon 9: **5000 kg**

```
%sql select max(Payload_mass) as MAX_Payload_mass from SPACEXTBL;
```

```
 * sqlite:///my_data1.db
Done.
```

MAX_Payload_mass
~5,000 kg

NASA is a major client with **32** launches

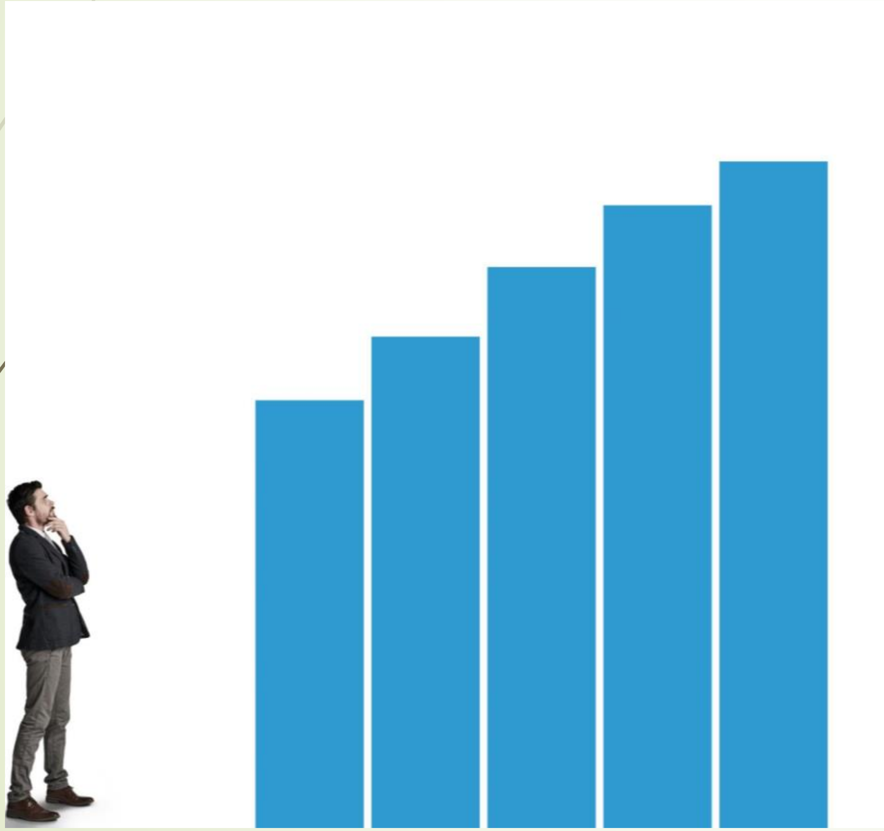
```
%sql select count(*) as Count from SPACEXTBL where Customer= 'NASA';
```

```
 * sqlite:///my_data1.db
Done.
```

Count
32

Model Development

Algorithm Selection and Performance



Algorithm Testing

Several algorithms including Logistic Regression, SVC, DecisionTreeClassifier AND KNeighborsClassifier were evaluated for model development and classification performance.

Logistic Regression Selection

Logistic Regression was selected for its high accuracy, interpretability, and strong predictive capability.

Hyperparameter Tuning

GridSearchCV optimized parameters such as C, penalty, and solver to improve model performance.

Model Validation and Deployment

The tuned model achieved an R^2 score of 0.94 through cross-validation, ideal for real-time applications.

Model Comparison

Evaluation of Tested Models

MODEL	ACCURACY
Logistic Regression	94%
SVC	88%
Decision Tree	88%
KNN	93%

Deployment

Interactive Application and Hosting

Interactive Web Application

Streamlit was used to build an interactive app offering real-time mission prediction based on input parameters.

Model Serialization and Storage

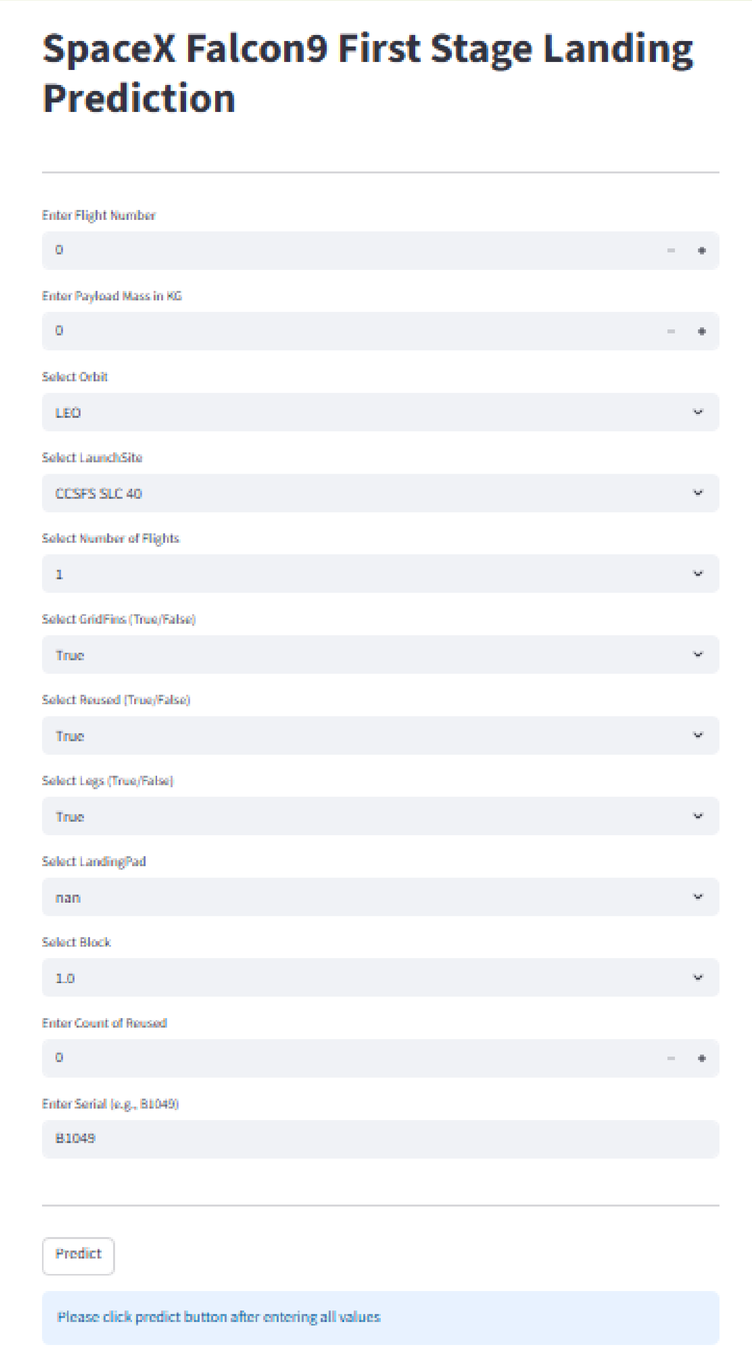
The predictive model was serialized with joblib into a .pkl file for efficient loading and use within the app.

Code Management and Collaboration

Application code was managed in GitHub to ensure version control and facilitate team collaboration.

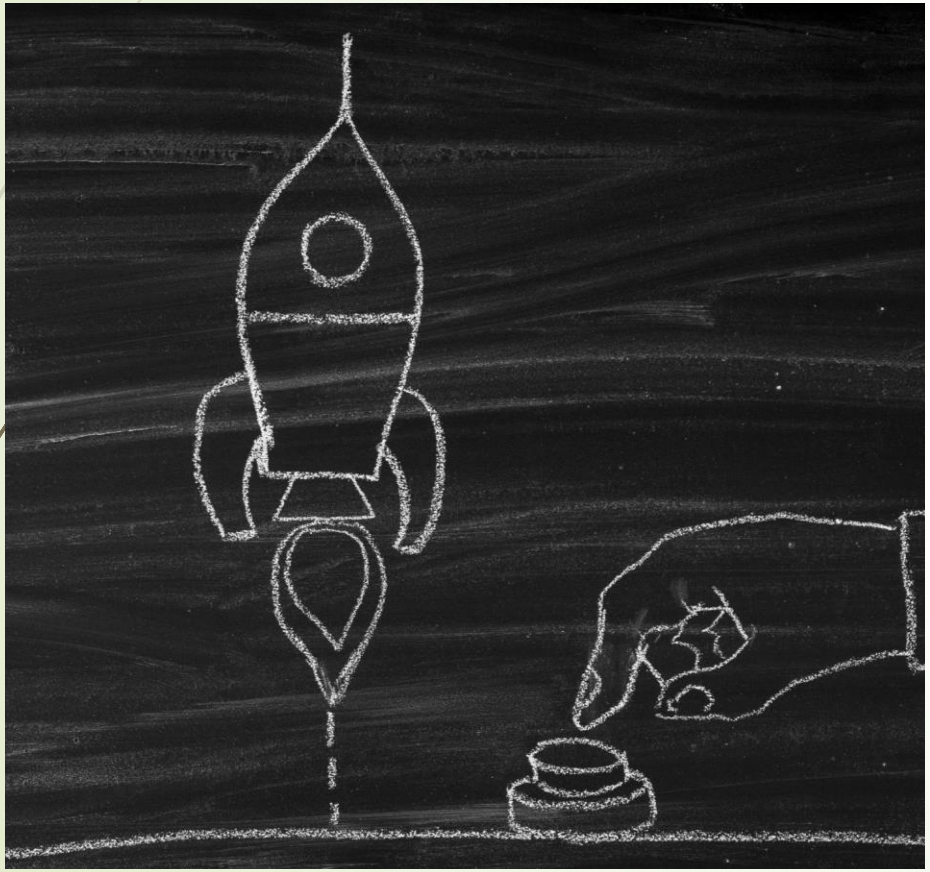
Cloud Hosting Deployment

Streamlit Cloud hosted the app, providing easy access for stakeholders and seamless deployment.



Key Findings

Insights Driving Strategic Decisions



Key Factors for Landing Success

Orbit type and payload mass are critical factors influencing successful spacecraft landings.

Predictive Model Performance

Logistic Regression shows strong predictive accuracy, suitable for operational use in mission planning.

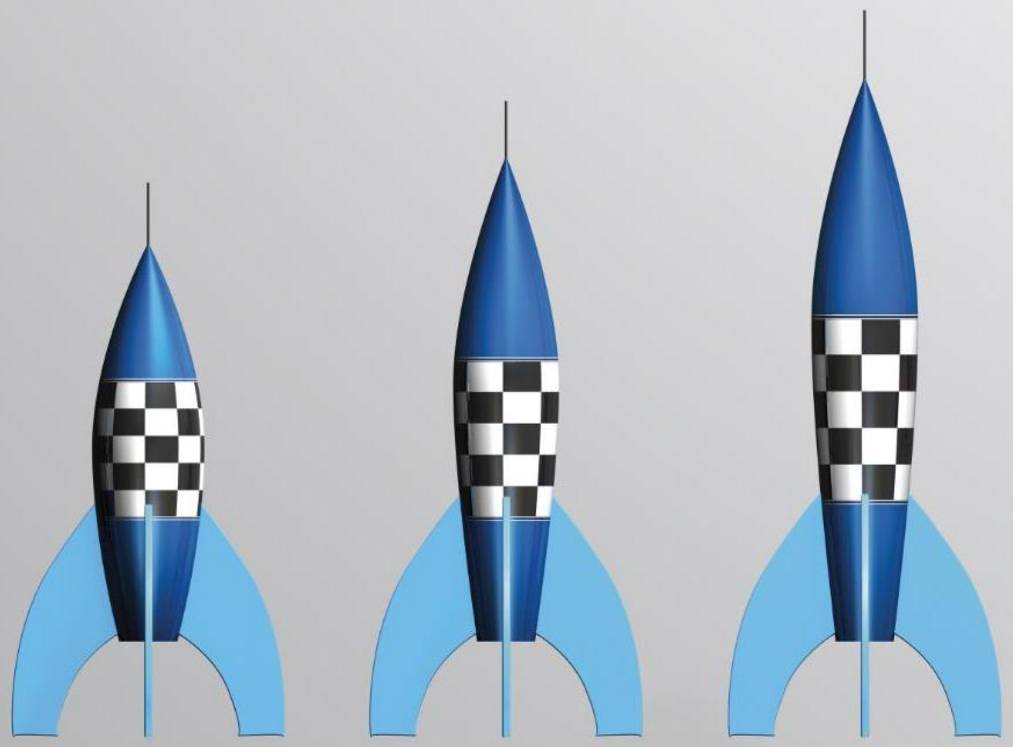
Institutional Partnerships

NASA’s role as a primary client highlights strategic partnerships that support frequent launches and revenue.

Optimizing Mission Planning

Insights guide stakeholders to reduce costs and boost competitiveness in the evolving space industry.

Conclusion & Future Scope



Summary and Next Steps

Project Achievement

Developed a predictive model with 0.94 R^2 score using Logistic Regression for Falcon 9 landing success.

Strategic Impact

Model insights aid cost estimation and mission planning, enhancing the advantage of reusability in space missions.

Future Enhancements

Plans include adding weather features, advanced algorithms, and real-time data for improved predictions.

Appendix

Technical Details and Tools

Tech Stack Overview

The project used Python and key libraries like Pandas, NumPy, and BeautifulSoup for data processing and modeling.

Application Development

Streamlit facilitated app creation with hosting on Streamlit Cloud for user accessibility.

Collaboration Tools

Google Colab and GitHub enabled collaborative coding, version control, and project deployment.

Model Evaluation Metrics

Metrics like R^2 score, accuracy, confusion matrix, and cross-validation ensured thorough model assessment.