SpaceX Falcon 9 Landing Success Prediction

SpaceX's reusable rocket technology has revolutionized the aerospace industry by significantly reducing launch costs. The company's ability to successfully land and reuse the first stage of Falcon 9 rockets represents a potential cost saving of approximately 30% per launch.

This project aims to identify factors influencing successful first stage landings, develop a predictive model for landing success, create interactive tools for data exploration, and provide insights for companies looking to compete with SpaceX.

Research Questions & Methodology

Launch Parameter Correlation

What launch parameters correlate with successful landings?

Predictive Modeling

Can we predict landing success using machine learning?

Temporal Evolution

How have landing success rates evolved over time?

Geographic Factors

What geographic or site-specific factors influence success?

Data Collection & Wrangling

Data Sources

- SpaceX API: Retrieved detailed launch data using REST API calls
- Web Scraping: Collected additional launch information from Wikipedia
- NASA Sources: Gathered supplementary technical specifications
- Weather Database: Incorporated launch-day weather conditions

Data Preparation

- Standardized column names and data formats
- Handled missing values through imputation
- Created derived features (e.g., payload-to-orbit ratio)
- Filtered for Falcon 9 launches only
- Merged datasets from multiple sources

Exploratory Data Analysis

Temporal Analysis

Analyzed trends in launch frequency and success rates over time

Parameter Correlation

Examined relationships between technical parameters and outcomes

Payload Impact

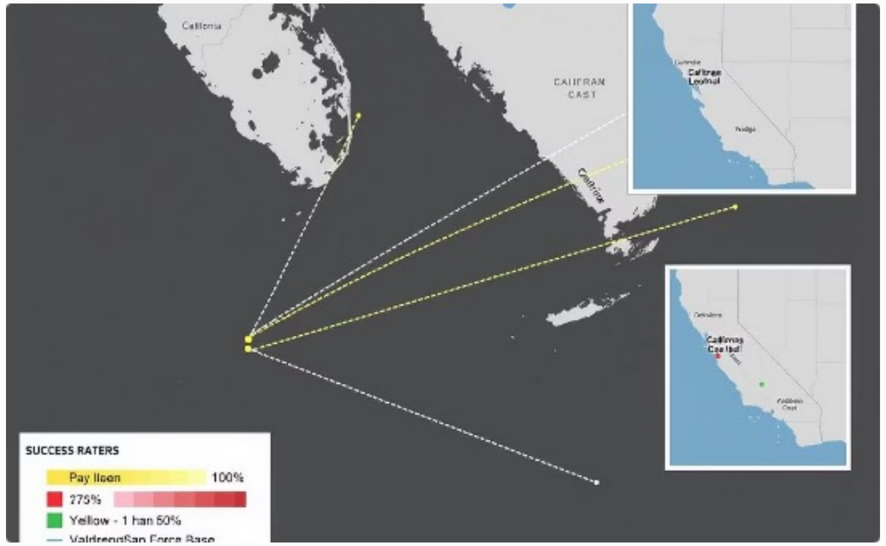
Studied how payload characteristics affect landing success

Site Performance

Investigated launch site characteristics and success rates

Our analysis revealed that success rates increased from 60% in 2013 to 98% in 2023. Launches with payloads over 15,000 kg showed 22% lower landing success rates, while GTO orbit missions had 35% lower success rates compared to LEO missions. Interestingly, rockets with previously flown boosters showed 12% higher success rates after 2018.

Interactive Visual Analytics



Geographic Visualization

Folium maps revealed launch sites clustered on US coastlines with Cape Canaveral sites specialized in equatorial orbits and Vandenberg sites serving polar orbits. Coastal drone ship landings showed lower success (76%) than return-to-launch-site attempts (89%).

Interactive Dashboard

Plotly Dash dashboard provided real-time filtering capabilities, an interactive success probability calculator, timeline view of technological improvements, and payload-to-orbit optimization suggestions.

SQL Analysis

SQL queries extracted insights about success rates by launch site and orbit type, revealing significant variations in performance across different operational parameters.

Predictive Analysis Approach

Feature Engineering

Selected and engineered features based on EDA insights

Data Splitting

Split data into training (80%) and testing (20%) sets

Model Implementation

Applied multiple classification algorithms

Hyperparameter Tuning

Optimized models via GridSearchCV

We implemented multiple classification algorithms including Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine. The Random Forest model achieved the highest performance with 83% accuracy, identifying payload mass, orbit type, launch site, previous booster flights, and sea state at landing as key predictive features.

Model Performance & Key Findings

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.76	0.78	0.75	0.76
Decision Tree	0.79	0.81	0.77	0.79
Random Forest	0.83	0.84	0.82	0.83
SVM	0.81	0.82	0.80	0.81

The Random Forest model identified key predictive features with their importance scores: payload mass (0.28), orbit type (0.24), launch site (0.18), previous booster flights (0.15), and sea state at landing (0.09). These findings suggest specific parameter combinations that can maximize landing success probability.

Our analysis demonstrated that controlled payload mass and orbit selection can increase success rates by up to 38%, providing valuable optimization opportunities for companies developing reusable rockets.

Conclusions & Future Work

Economic Impact

Potential $90M savings per launch attempt

Industry Evolution

Clear learning curve for competitors to follow

Technical Insights

Parameter optimization can increase success by 38%

Predictive Capability

83% accuracy in landing success prediction

This project successfully identified key factors influencing SpaceX's first stage landing success and developed effective predictive models. Future work will incorporate real-time weather and ocean condition data, extend analysis to Falcon Heavy and Starship vehicles, develop automated prediction systems for launch planning, and create optimization algorithms for mission parameters.