SpaceX Falcon 9 First Stage Landing Prediction: A Data Analysis and Machine Learning Report

Author: Prajit Bhalala

Date: July 20, 2025

2. Executive Summary

- This report details a comprehensive project aimed at predicting successful first-stage landings of the SpaceX Falcon 9 rocket.
- The ability to reuse the first stage is a cornerstone of SpaceX's cost-efficiency, making accurate landing prediction vital for competitive bidding.
- The project involved three key phases: **data acquisition from the SpaceX API** , thorough **exploratory data analysis (EDA)** and **feature engineering**, and the application of **machine learning models**.
- Key EDA insights revealed a strong upward trend in launch success rates over the years, alongside varying success probabilities across different launch sites and orbit types.
- In the machine learning phase, **Logistic Regression, SVM, Decision Tree, and K-Nearest Neighbors (KNN)** models were developed and optimized. These models all achieved an **83.33%** accuracy on the test set, demonstrating their effectiveness in distinguishing between successful and unsuccessful landing outcomes.
- This analysis provides valuable predictive capabilities for future SpaceX missions and competitive market strategies.

3. Table of Contents

4. Introduction

4.1 Problem Statement

The core problem addressed in this project is to accurately predict whether the first stage of the SpaceX Falcon 9 rocket will land successfully. This predictive capability is vital for assessing the cost-effectiveness of a launch, as the reusability of the first stage significantly reduces overall mission expenses compared to other providers. Such information can be leveraged by alternative companies bidding against SpaceX for rocket launch contracts.

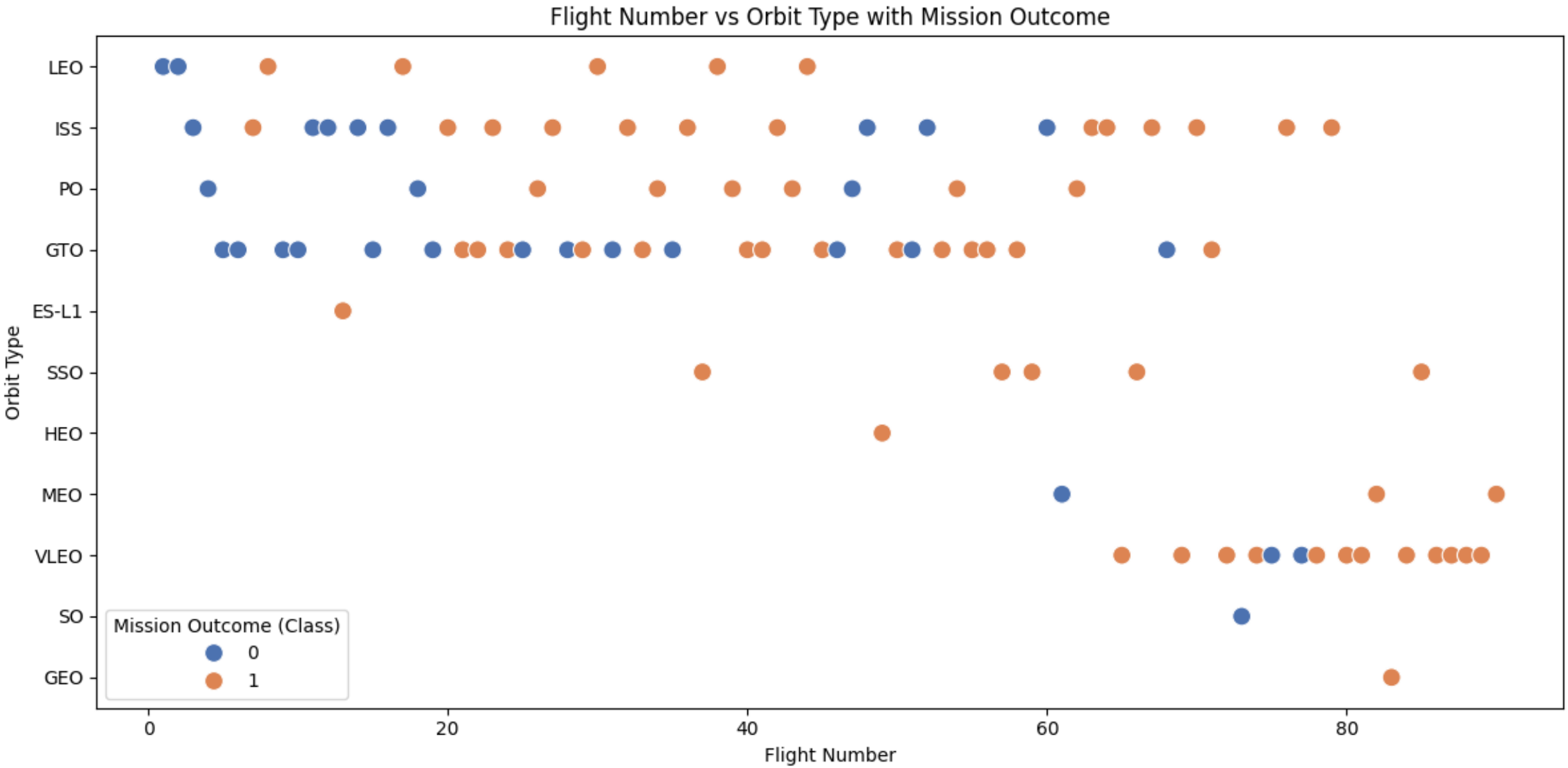
4.2 Background

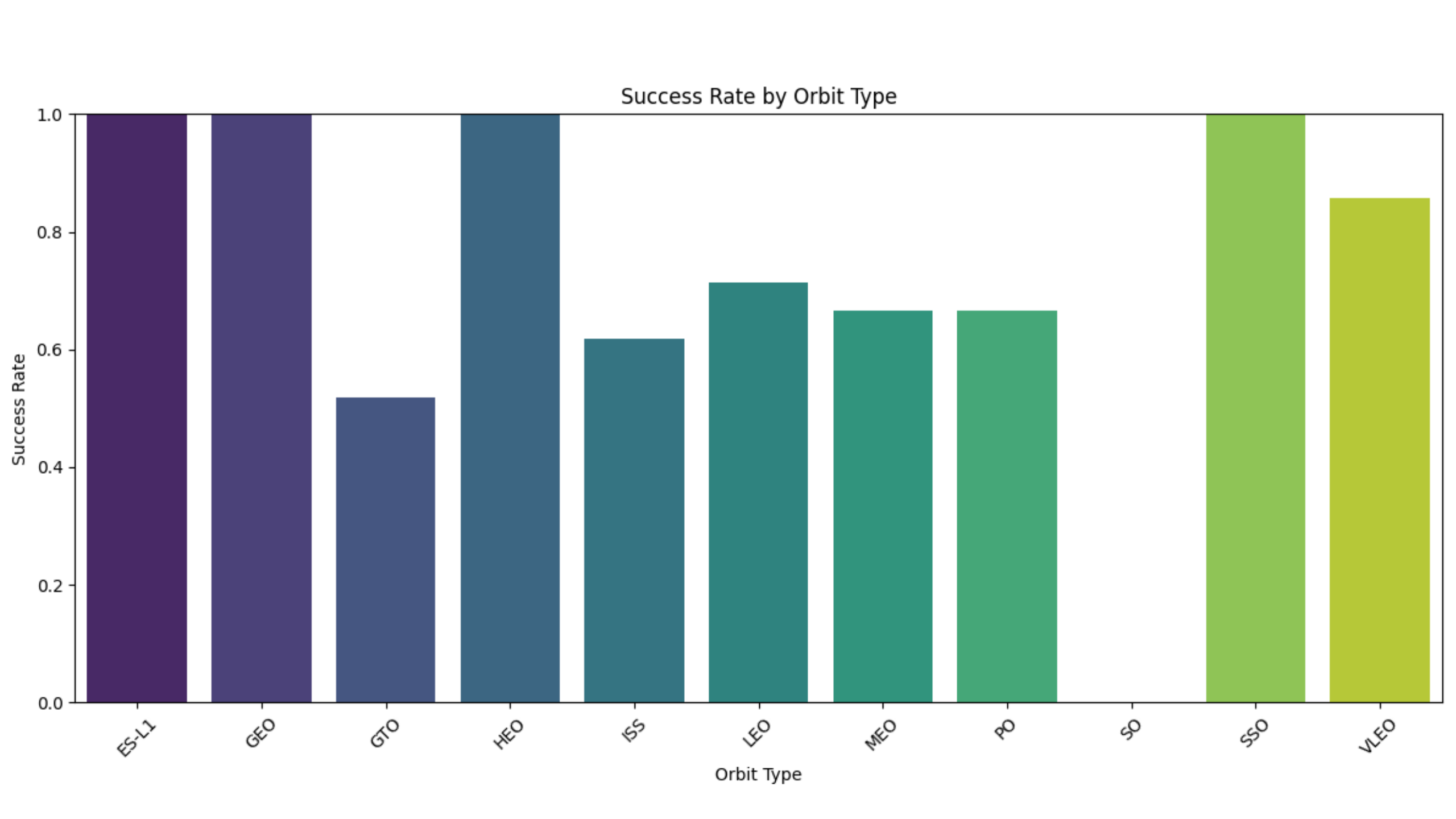
SpaceX has garnered global attention for its achievements in space exploration, notably its pioneering work in reusing the first stage of its Falcon 9 rockets. The ability to return the first stage from low-Earth orbit, first accomplished in December 2010, represents a monumental shift in space launch economics. While successful landings are ideal, some missions involve planned unsuccessful landings, often controlled landings in the ocean. Visual examples illustrate both the precision of successful landings and the outcomes of unsuccessful attempts, including crashes.

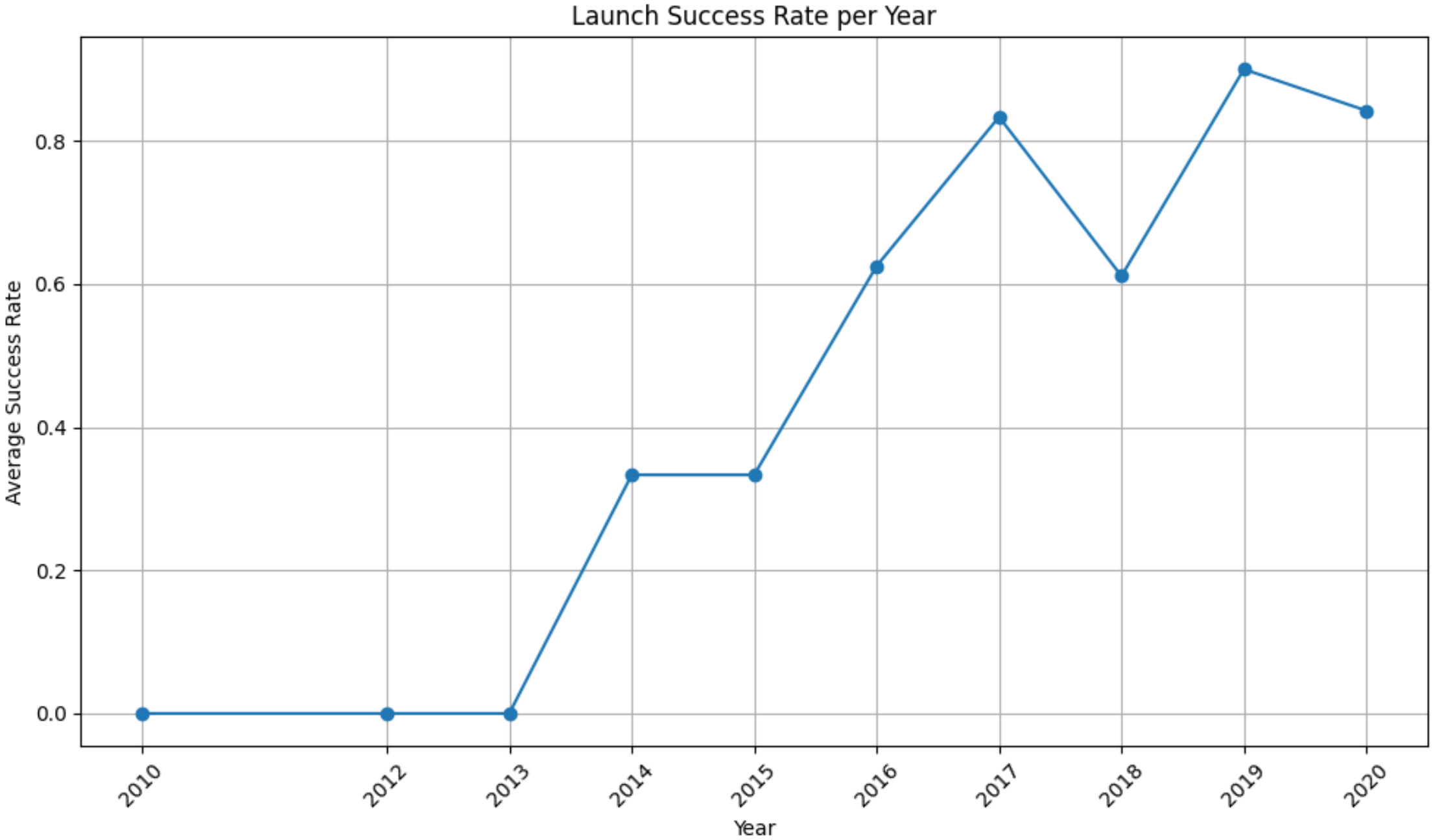
4.3 Objectives

This report outlines a systematic approach to build a machine learning pipeline for predicting Falcon 9 first-stage landing outcomes. The project's objectives are structured into three main phases:

- **Data Collection:** To gather comprehensive historical launch records for Falcon 9 and Falcon Heavy rockets from various online sources, including the SpaceX API and Wikipedia. This involves making HTTP requests and performing web scraping to extract relevant data.
- **Exploratory Data Analysis (EDA) & Feature Engineering:** To analyze the collected data to identify patterns, understand the relationships between different variables and landing success, and prepare the data for machine learning. This includes handling missing values, creating new features, and encoding categorical variables.
- **Machine Learning Prediction:** To develop and evaluate several classification models (Logistic Regression, SVM, Decision Tree, and KNN) capable of predicting the landing outcome based on the engineered features. The aim is to identify the best-performing model for this prediction task.


Flight Number vs Orbit Type with Mission Outcome
LEO
ISS
PO
GTO
ES-L1
SSO
HEO
MEO
VLEO
SO
GEO
Orbit Type
0
20
40
60
80
Flight Number
Mission Outcome (Class)
0
1


Success Rate by Orbit Type
Success Rate
1.0
0.8
0.6
0.4
0.2
0.0
ES-L1
GEO
GTO
HEO
ISS
LEO
MEO
PO
SO
SSO
VLEO
Orbit Type


Launch Success Rate per Year
Average Success Rate
0.0
0.2
0.4
0.6
0.8
2010
2012
2013
2014
2015
2016
2017
2018
2019
2020
Year

5. Research Methodology

The project's predictive modeling for SpaceX Falcon 9 first-stage landings was built upon a robust research methodology, encompassing data acquisition, rigorous data preparation, and the application of diverse machine learning techniques.

5.1 Data Sources

Comprehensive historical launch information was gathered from two primary sources:

- **SpaceX API:** Historical launch data was retrieved directly from the SpaceX API, specifically from the /v4/launches/past endpoint. Python's requests and pandas libraries were used to process the JSON responses into a DataFrame. Helper functions extracted key details like BoosterVersion, LaunchSite (including geo-coordinates), PayloadMass, Orbit, and various Core data attributes such as landing success, type, and reuse counts. The initial DataFrame was filtered to include only single-core, single-payload Falcon 9 launches within a specific date range (up to November 13, 2020), and missing PayloadMass values were imputed with the mean.
- **Wikipedia Web Scraping:** Additional historical launch records for Falcon 9 and Falcon Heavy were scraped from a Wikipedia page titled "List of Falcon 9 and Falcon Heavy launches" (snapshot from June 9, 2021). requests and BeautifulSoup were employed for HTML parsing, with custom helper functions to accurately extract data from the complex HTML table structures.

5.2 Data Preparation & Feature Engineering

Once collected, the raw data underwent several critical transformations for machine learning:

- **Handling Missing Values:** NaN entries in LandingPad were retained, indicating no landing pad use, while missing PayloadMass values were imputed with their mean.

•**Creation of Class Label:** A binary Class variable (0 for unsuccessful, 1 for successful) was created by transforming the Outcome column (e.g., 'False ASDS', 'None None' mapped to 0; 'True ASDS' mapped to 1).
•**Feature Selection & Encoding:** Relevant features including FlightNumber, PayloadMass, Orbit, LaunchSite, Flights, GridFins, Reused, Legs, LandingPad, Block, ReusedCount, and Serial were selected. Categorical features (Orbit, LaunchSite, LandingPad, Serial) were then converted into numerical representations using One-Hot Encoding via pandas.get_dummies(). All resulting numeric columns were cast to float64 for consistency.

5.3 Machine Learning Methods

The prepared dataset was then used to train and evaluate several supervised machine learning classification models:

Data Partitioning: The dataset was split into training and testing sets using sklearn.model_selection.train_test_split. A test_size of 0.2 and random_state of 2 were applied to ensure reproducibility and an 80/20 train-test split.

Data Standardization: Feature scaling was applied to the training and testing data using sklearn.preprocessing.StandardScaler to normalize the feature values. This step is important for algorithms sensitive to feature magnitudes, ensuring fair weighting during model training.

Model Training and Hyperparameter Tuning: For each classification algorithm, sklearn.model_selection.GridSearchCV was employed to systematically search for the best hyperparameters using 10-fold cross-validation (cv=10).

Logistic Regression: Tuned parameters included C (inverse of regularization strength), penalty (l2), and solver (lbfgs).

Support Vector Machine (SVM): Parameters tuned were kernel (linear, rbf, poly, sigmoid), C, and gamma.

Decision Tree Classifier: Parameters included criterion (gini, entropy), splitter (best, random), max_depth, max_features, min_samples_leaf, and min_samples_split.

K-Nearest Neighbors (KNN): Parameters tuned were n_neighbors, algorithm (auto, ball_tree, kd_tree, brute), and p (power parameter for Minkowski metric).

Model Evaluation: The performance of the best-tuned model for each algorithm was assessed by calculating its accuracy on the unseen test data using the score() method. Confusion matrices were also generated and visualized for each model to provide a detailed breakdown of true positives, true negatives, false positives, and false negatives.

6. Findings and Results

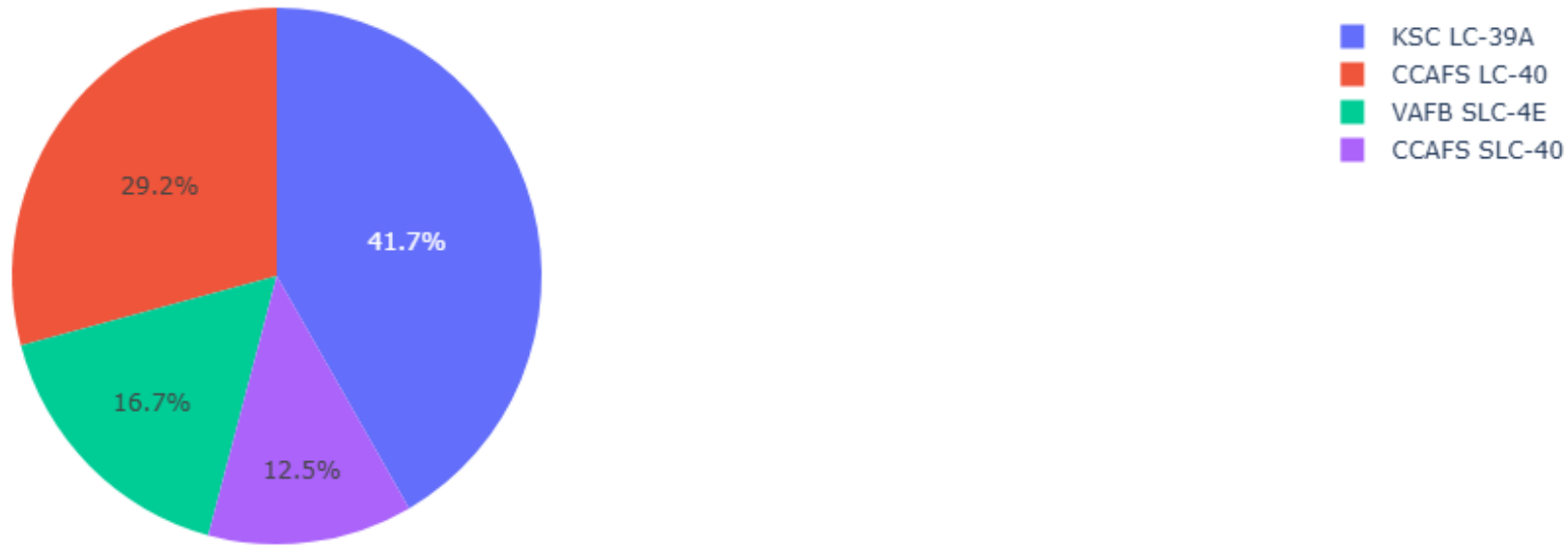
The project's analytical phase yielded significant insights into SpaceX Falcon 9 landing outcomes.

6.1 Exploratory Data Analysis (EDA)

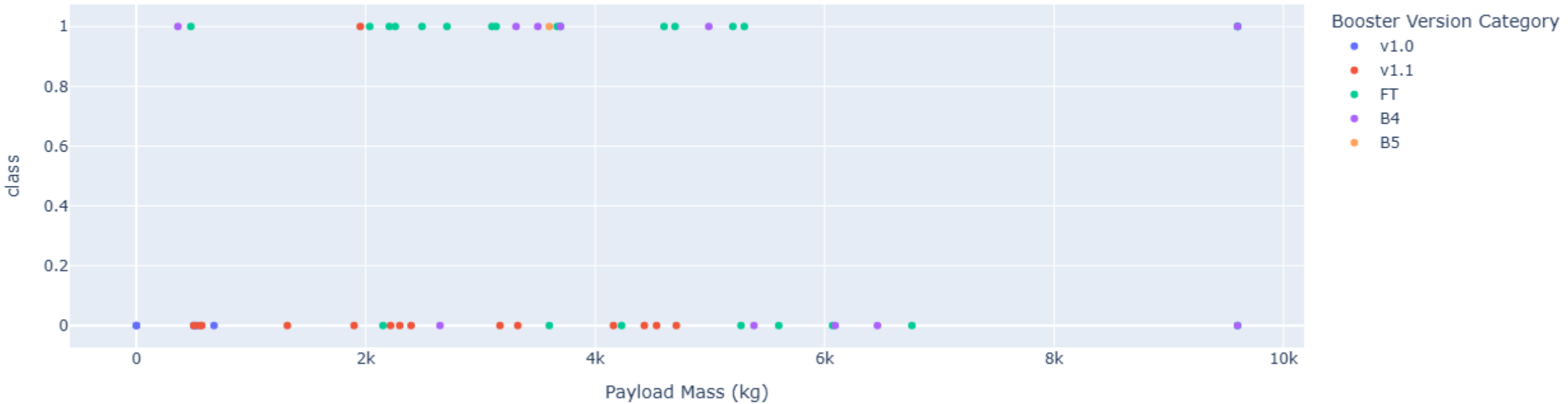
Initial data inspection of 90 Falcon 9 launches revealed 5 missing PayloadMass values and 26 LandingPad missing values, with PayloadMass imputed by its mean, while LandingPad NaNs were retained as valid indicators of no pad use.

- **Launch Site Analysis:** Three unique launch sites were identified: CCAFS SLC 40 (55 launches), KSC LC 39A (22 launches), and VAFB SLC 4E (13 launches).
- **Relationship between Flight Number and Launch Site:** A scatter plot of FlightNumber vs. LaunchSite showed that increasing flight numbers correlated with higher landing success across different sites. Notably, VAFB-SLC 4E recorded no successful landings for heavy payloads (>10,000 kg).
- **Relationship between Payload Mass and Launch Site:** The PayloadMass vs. LaunchSite scatter plot further highlighted VAFB-SLC 4E's consistent unsuccessful landings for heavy payloads (>10,000 kg), while other sites showed success across varied payload masses.
- **Orbit Type Analysis:** Eleven unique orbit types were present, with GTO (27 launches) and ISS (21 launches) being the most frequent. A bar chart of success rates per orbit indicated 100% success for HEO, GEO, and ES-L1 (though with limited launches), and high rates for ISS and VLEO.
- **Relationship between Flight Number and Orbit Type:** Success in LEO orbits positively correlated with higher FlightNumber, whereas GTO orbits showed no clear relationship, with mixed outcomes across flight numbers.

Total Success Launches by Site



Payload vs. Outcome for All Sites



- **Launch Success Yearly Trend:** A line chart demonstrated a consistent increase in launch success rates from 2013, peaking in 2020, reflecting ongoing improvements in SpaceX's capabilities.
- **Mission Outcome Analysis:** The Outcome column was used to create a binary Class variable (0 for failure, 1 for success). The overall landing success rate was approximately 66.67%.

6.2 Machine Learning Model Performance

Four classification models were trained and evaluated on standardized data, split into 80% training and 20% testing sets. Hyperparameter tuning was performed using GridSearchCV with 10-fold cross-validation.

- **Logistic Regression:** Best hyperparameters: {'C': 0.01, 'penalty': 'l2', 'solver': 'lbfgs'}. Validation Accuracy: 0.8464. Test Accuracy: 0.8333. The confusion matrix showed 12 true positives, 3 false positives, 3 false negatives, and 0 true negatives.
- **Support Vector Machine (SVM):** Best hyperparameters: {'C': 1.0, 'gamma': 0.0316, 'kernel': 'sigmoid'}. Validation Accuracy: 0.8482. Test Accuracy: 0.8333. Confusion matrix results were identical to Logistic Regression.
- **Decision Tree Model:** Best hyperparameters: {'criterion': 'gini', 'max_depth': 4, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 2, 'splitter': 'random'}. Validation Accuracy: 0.8893. Test Accuracy: 0.8333. Confusion matrix results mirrored the other models.
- **K-Nearest Neighbors (KNN) Model:** Best hyperparameters: {'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}. Validation Accuracy: 0.8482. Test Accuracy: 0.8333. Confusion matrix results were consistent across all models.

Comparative Model Performance: All four models achieved an identical test accuracy of 83.33%. While the Decision Tree showed a higher validation accuracy, its test set performance aligned with the others, suggesting comparable predictive power on this dataset.

7. Discussion

- The exploratory data analysis provided crucial insights into factors influencing Falcon 9 first-stage landing success. An increasing success rate over the years highlights SpaceX's continuous improvements in reusability.
- Launch site analysis revealed distinct patterns; VAFB SLC 4E, for instance, showed no successful heavy payload landings (>10,000 kg), suggesting specific operational limitations. Orbit type also proved significant, with LEO, ISS, HEO, GEO, and ES-L1 demonstrating high success rates, while GTO presented more complexity.
- In the machine learning phase, Logistic Regression, SVM, Decision Tree, and KNN models all achieved an identical test accuracy of 83.33%. Consistently, confusion matrices showed 12 true positives, 3 false positives, 3 false negatives, and 0 true negatives, indicating strong success prediction but some false positives/negatives.
- The Decision Tree model exhibited the highest validation accuracy (88.93%), suggesting potential for better generalization with more data. Overall, the models offer a robust framework, with EDA insights informing predictions.

8. Conclusion

- This report successfully demonstrated a comprehensive approach to predicting SpaceX Falcon 9 first-stage landing outcomes, crucial for optimizing rocket reusability and launch costs.
- Through meticulous data collection from the SpaceX API and Wikipedia, a rich dataset was compiled. Exploratory data analysis revealed significant trends: SpaceX's landing success rate has shown a clear upward trajectory since 2013, and different launch sites and orbit types exhibit varying success patterns.
- Four machine learning models—Logistic Regression, SVM, Decision Tree, and KNN—were implemented, optimized, and evaluated. All achieved an identical test accuracy of 83.33%, demonstrating strong predictive capability for landing success. While test accuracies were uniform, the Decision Tree showed a promising higher cross-validation score.
- Future work includes increasing data volume, advanced feature engineering, and exploring deep learning for enhanced accuracy and interpretability. This project establishes a solid foundation for understanding and predicting Falcon 9 landing outcomes.