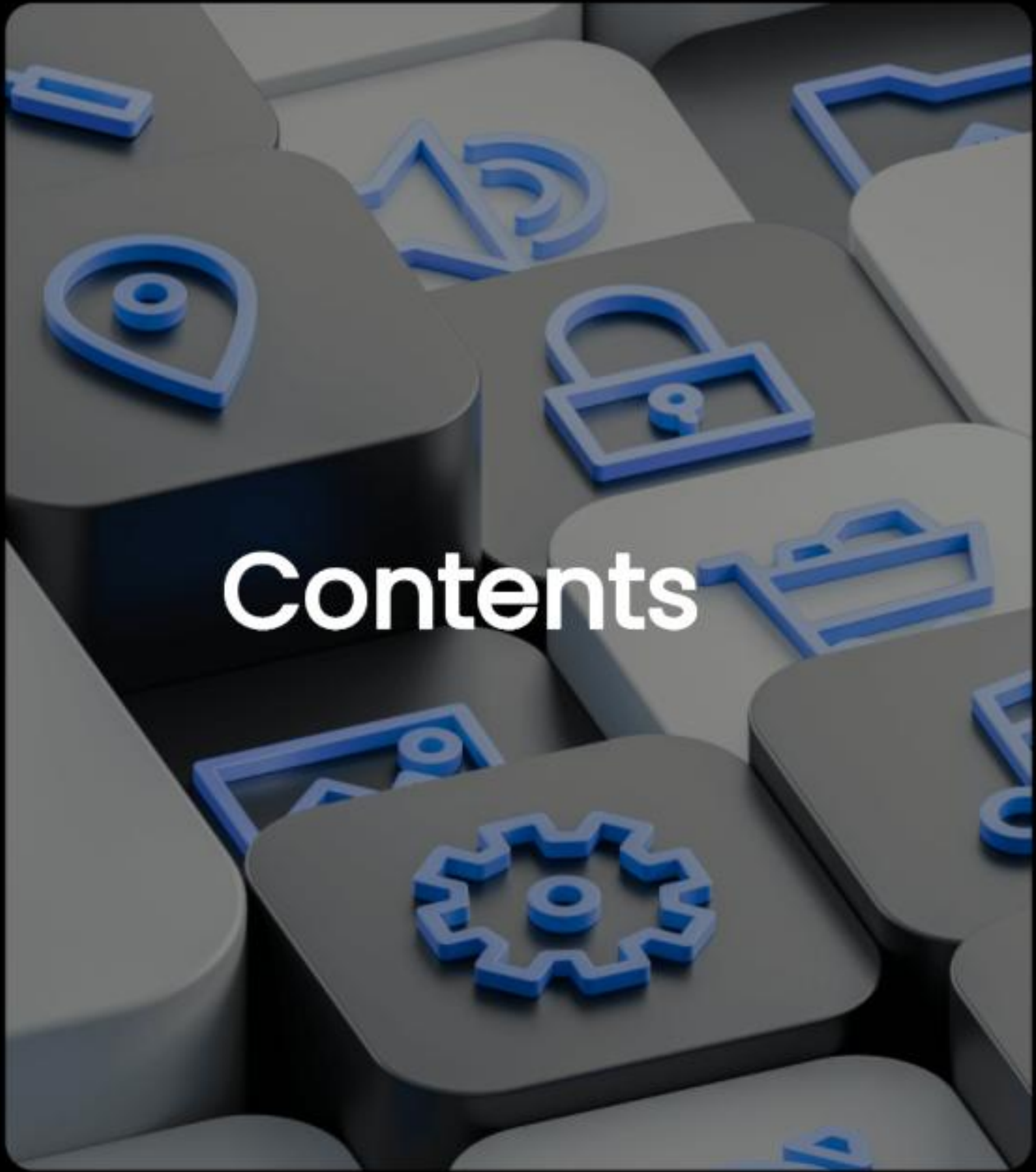
https://github.com/heetmdave/SpaceX_Data_analysis

SpaceX Falcon 9 Launch Analysis and Prediction

Data： 2025.6.22

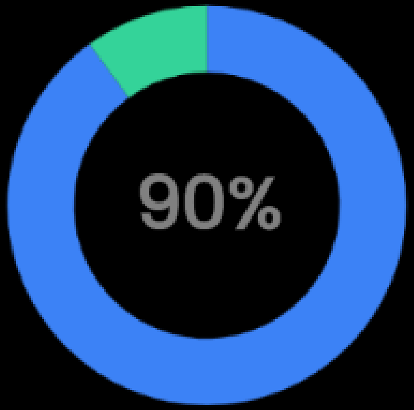
Presenter: Heet Dave

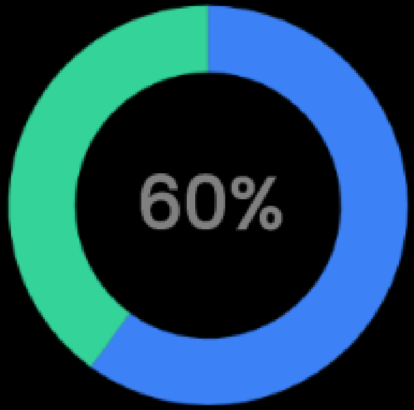


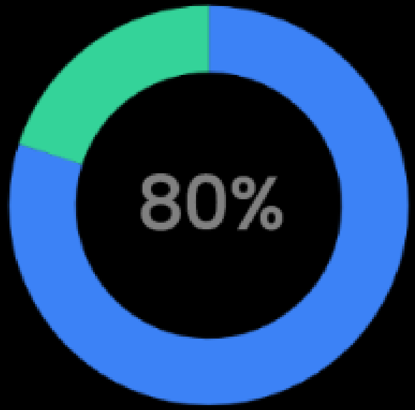
Contents

01
Executive Summary

Overview of Analysis







Key objectives and goals

The analysis aims to assess Falcon 9 launch performance, identify key factors influencing success rates, and predict future launch outcomes. Key objectives include enhancing operational efficiency, reducing costs, and informing strategic decisions for upcoming missions.

Brief project description

The SpaceX Falcon 9 Launch Analysis examines historical launch data, identifies performance trends, and predicts future mission success probabilities. This project aims to enhance understanding of Falcon 9 reliability and optimize future launch strategies through comprehensive data analytics.

Significance of study

The significance of this study lies in its potential to enhance understanding of Falcon 9 launch performance, inform future mission design, and aid in improving reliability and cost efficiency in spaceflight operations, ultimately contributing to advancements in commercial space exploration.

Your logo

02
Introduction

Background

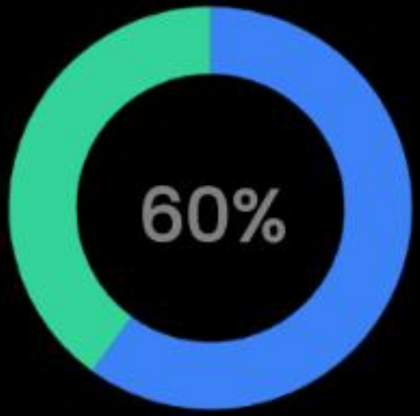
SpaceX and the Falcon 9 overview

SpaceX, founded by Elon Musk in 2002, aims to reduce space transportation costs. The Falcon 9, a two-stage rocket, is designed for reliability and reusability, making it a cornerstone of SpaceX's mission to enable human life on Mars and facilitate satellite launches.

Importance of launch analysis

Launch analysis is crucial for ensuring mission success, optimizing performance, and enhancing safety. It provides insights into vehicle behavior, trajectory dynamics, and potential anomalies, enabling engineers to make data-driven decisions that improve future launches and contribute to the overall advancement of space exploration.

Project Scope



Objectives of the capstone project

This capstone project aims to analyze the launch performance of SpaceX Falcon 9, assess its reliability and efficiency, and develop predictive models to enhance future launch outcomes, contributing valuable insights to the aerospace industry.

Methodological approaches

This analysis employs quantitative and qualitative methodologies to examine Falcon 9 launch data. Techniques include statistical modeling, trend analysis, and simulation to predict future launch outcomes and identify operational efficiencies, providing insights into the launch vehicle's performance and reliability.

03

Data Collection and Wrangling Methodology

Data Sources

spaceX Rest API

Retrieved structured launch data including flight number, mission name, payload mass, orbit, launch site, and landing outcome. Used requests library to get data from API . Then filtered data set for falcon9 data.

Web scraping

Scraped Wikipedia pages for additional details on payload mass and mission outcomes where API data was incomplete.

Data Wrangling Techniques

Cleaning and preprocessing

Data cleaning involved identifying and correcting inaccuracies in launch datasets, including removing duplicates, handling missing values, and standardizing formats. Preprocessing ensured data consistency, enabling more reliable analyses and predictions regarding Falcon 9 launch outcomes. Techniques included outlier detection and normalization for improved model performance.

Converted tagetoutcomes to indicator variale(0/1)

Handling missing values

Handling missing values involved using imputation techniques, such as mean, median, and mode substitution, to fill gaps in the dataset.

imported missing values of payload mass by its mean value.

04

Exploratory Data Analysis and Interactive Visual Analytics Methodology

EDA Techniques

Trend and corellation analysis

Analysis of trend and corellation done by visualizing scatter plot of different attributes with taking outcome as hue , mapping bar charts among atributes .

01

Data transformation

Data transformation has done by converting categorical variables to numerical variable by using technique called one hot encoding

02

Interactive Visual Analytics

Overview of visualization tools used

visualization tools used to visualize Folium lab and ploty Dash libraryes . these libraries has utilised in python sctipt.

Benefits of interactivity in analysis

Interactive visual analytics enhance data exploration by allowing users to dynamically manipulate datasets, uncover patterns, and gain insights in real-time. This engagement fosters a deeper understanding of the data, enabling more informed decisions and facilitating effective communication of findings.

05

Pridictive Analysis Methodology

Pridictive Analysis

Predictive Analysis refers to trainig model to inference outcome based on attributes .
predictive model maps the relationship between attributes and target variables .
In project predictive models utilized are svm , decision trees and logistic regression.
GreedsearcCV is utilised to finetune hyperparameters.

06

EDA with Visualization Results

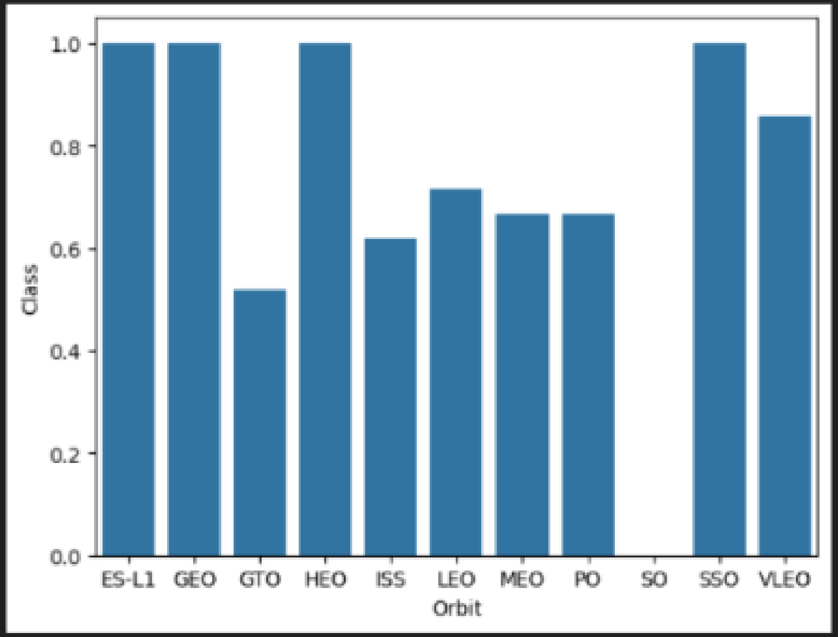
Charts and Graphs

- **Yearly trend**

  sucess rate increse from 2013 to 2017 decrese in 2018 than increase in 2019 .

- **Comparison across different launch sites**

  The comparison across different launch sites reveals distinct patterns in launch success rates and payload capacities. Site-specific factors such as geographical location and historical data influence performance metrics, emphasizing the need for tailored strategies in mission planning for maximized efficiency and reliability.



Best Orbits

Es-L1,GEO,HEO,SSO

Best orbits for hig payload mass

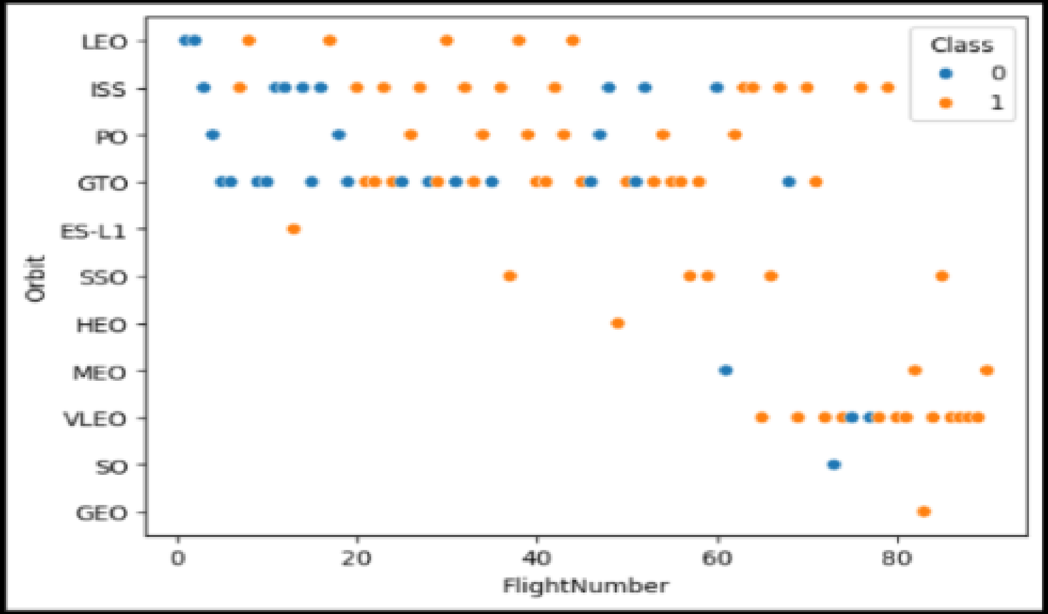
ISS, PO, VLEO

Year of highest sucess rate

2019

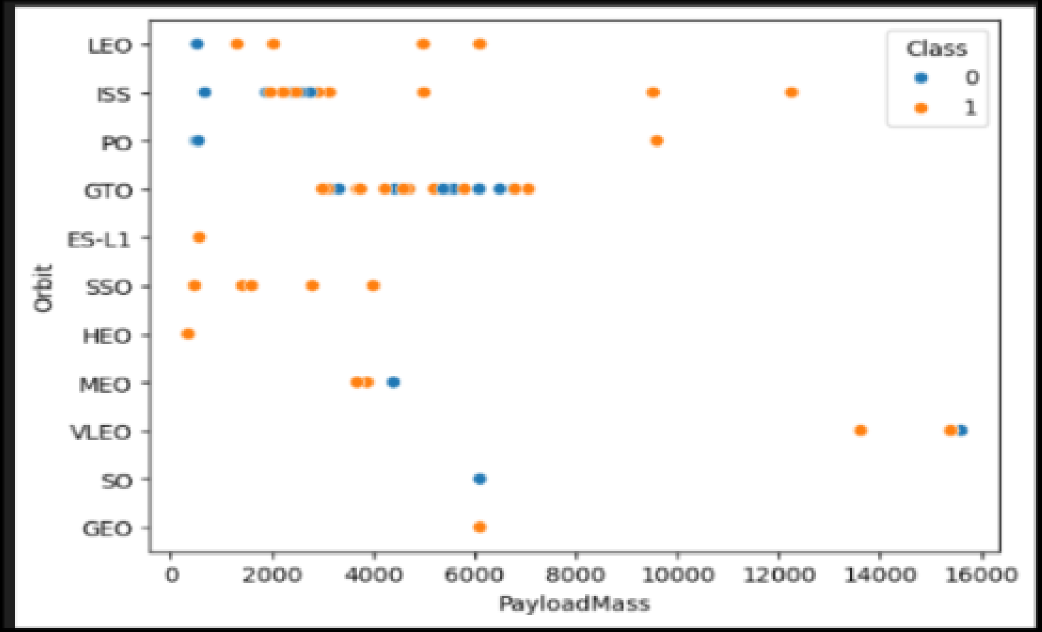
Plot Explanations

Orbit vs time(representing by fight number) of visual data



In VLEO orbit almost all lunches got sucess and in PO after some time intervel thel got able to sucessful landing. there are no such relations in iss

Orbit vs payloadmass with outcome as hue



lunch ISS and PO range got good resultes in high payload . for low payload sso gives good results . it seems like sso does not capable to carry high payload.

07

EDA with SQL Results

SQL Queries and Results

Display the total payload mass carried by boosters launched by NASA (CRS): 619967 KG

Display average payload mass carried by booster version F9 v1.1 : 2534.66 KG

List the date when the first succesful landing outcome in ground pad was acheived.: 2010-06-04

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

Task 9

List the records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_sit for the months in year 2015.

Note: SQLLite does not support monthnames. So you need to use substr(Date, 6,2) as month to get the months and substr(Date,0,5)='2015' for year.

```
%sql select substr(Date, 6,2),Booster_Version,Launch_Site,Landing_Outcome from SPACEXTABLE where substr(Date,0,5)='2015'
```

Python

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

substr(Date, 6,2)	Booster_Version	Launch_Site	Landing_Outcome
01	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
02	F9 v1.1 B1013	CCAFS LC-40	Controlled (ocean)
03	F9 v1.1 B1014	CCAFS LC-40	No attempt
04	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)
04	F9 v1.1 B1016	CCAFS LC-40	No attempt
06	F9 v1.1 B1018	CCAFS LC-40	Precluded (drone ship)
12	F9 FT B1019	CCAFS LC-40	Success (ground pad)

Task 10

Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order.

```
%sql select count(Landing_Outcome),Landing_Outcome from SPACEXTABLE where Date>'2010-06-04' and Date<'2017-03-20' group by Landing_Outc
```

Python

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

count(Landing_Outcome)	Landing_Outcome
1	Failure (parachute)
1	Precluded (drone ship)
2	Uncontrolled (ocean)
3	Controlled (ocean)
3	Success (ground pad)
5	Failure (drone ship)
5	Success (drone ship)
10	No attempt

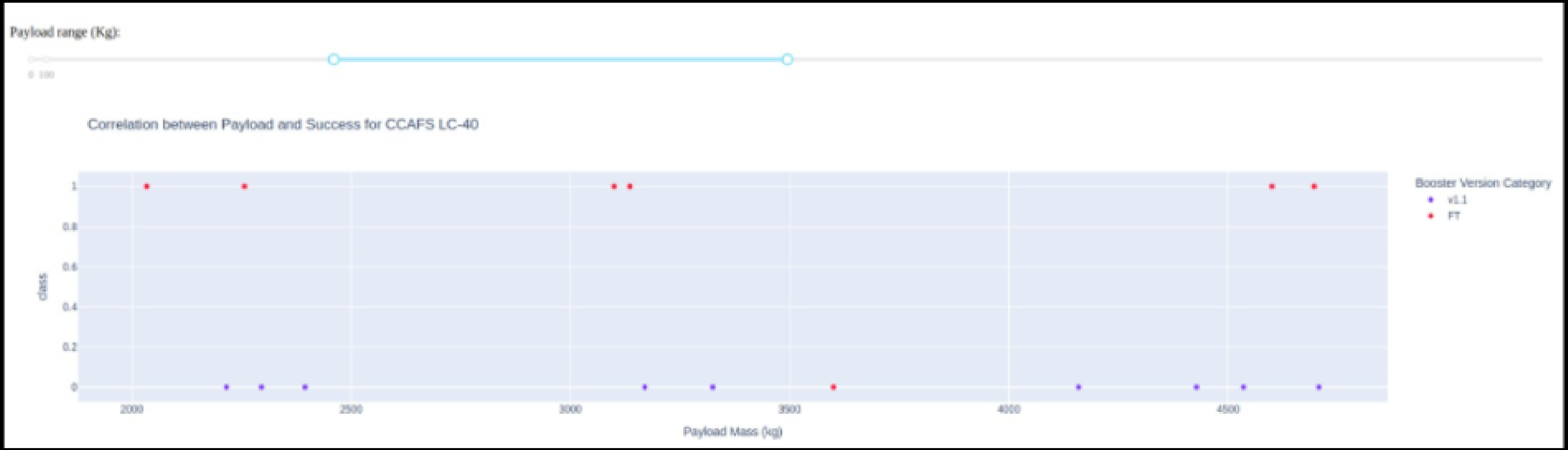
08

Interactive Visualizations and Predictive Analysis Results

Interactive Map with Folium Results

lable based on sucess

polyline to coast



for FT booster version 2000 to 2500 , 3000 to 3500 and 4500 to 5000 KG are range of sucess

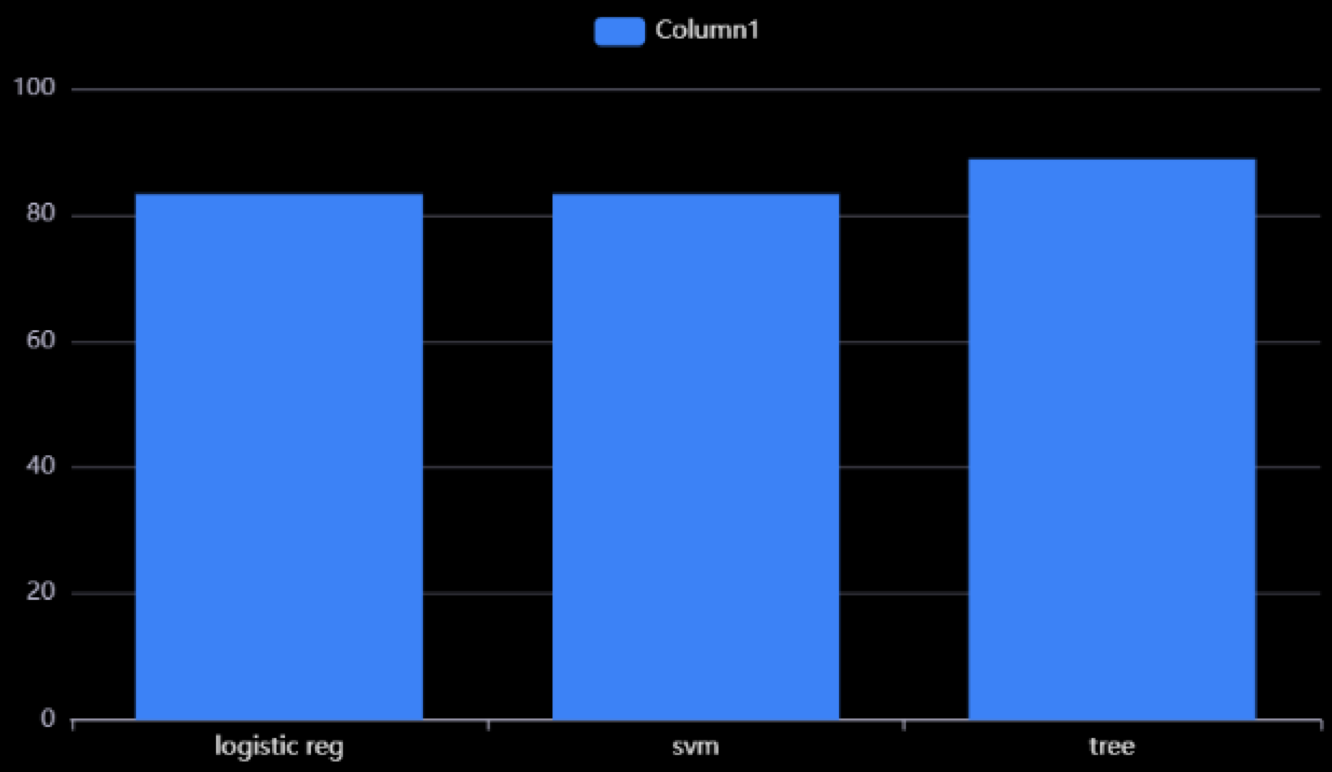
Predictive Analysis Outcomes

Classification model performance

accuracy	Model
83.3%	logistic reg
88.9%	tree
83.3%	svm

Best classifire

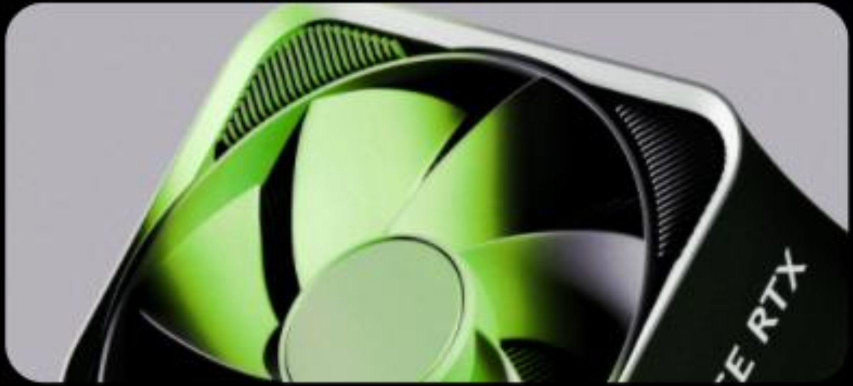
Decision tree providing best outputs among all tested



Summary of Findings

Recap of key insights

The analysis of Falcon 9 launches reveals consistent improvements in launch reliability and turnaround times. Key insights emphasize the importance of innovative reusable technology and efficient operational strategies, which contribute to SpaceX's competitive edge in the aerospace industry.



Implications for future launches

The analysis of Falcon 9 launches reveals a trend toward increased reliability and cost efficiency. These findings suggest that future launches could leverage reusable technology to enhance accessibility and sustainability in space exploration, potentially revolutionizing the space industry.

Recommendations

Suggestions for SpaceX

To enhance Falcon 9's efficiency, SpaceX should invest in advanced predictive analytics for launch patterns, streamline supply chain processes, and focus on environmental sustainability initiatives to strengthen public support and compliance with regulatory standards. This will boost long-term mission viability.

Future research directions

Future research should explore advanced materials for enhanced rocket durability, improved AI algorithms for flight path optimization, and comprehensive studies on reusability impacts on launch costs. Collaborative efforts with international space agencies may also yield innovative strategies for regular missions and deeper space exploration.

Data： 2025.6.22

Presenter: Heet Dave