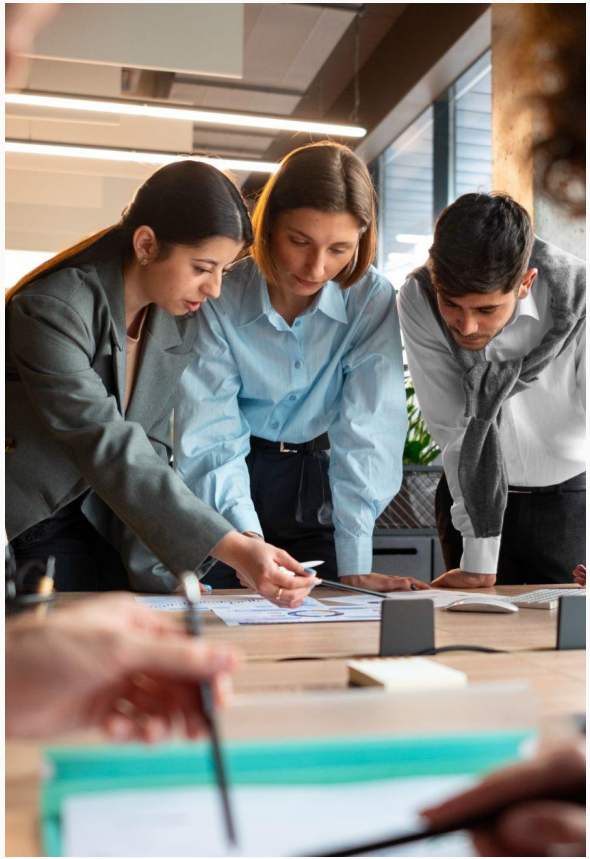
Predict the success of SpaceX's Falcon 9 first-stage launches

Mini Project | Mason Phung

Table of contents



01

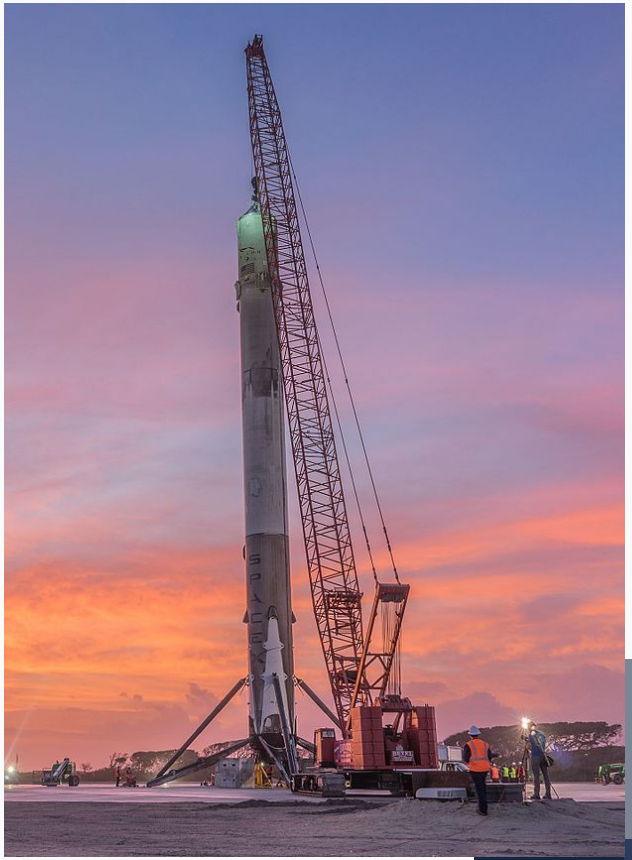
Introduction and Objectives

SpaceX ‘s falcon 9

SpaceX is popular for her innovations and leading positions in the space industry. Many of their accomplishments include sending humans and spacecrafts to the Space and developing the world’s most powerful rocket.

During 2013 and 2016, SpaceX conducted a series of controlled-descent flight tests for its Falcon 9 rocket.

SpaceX’s Falcon 9 rocket was well-known for its ability to recover the first stage, which could save a lot of money for the company.



A falcon 9 rocket on a ground pad at Cape Canaveral AF Station

(source: wikipedia)

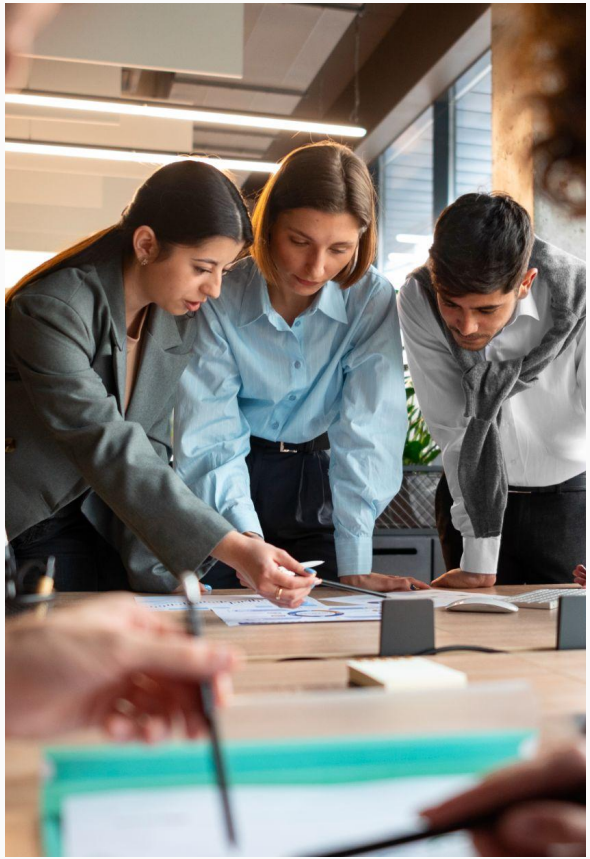
Objectives

Our aim

The aim in the project is to determine the price of each rocket launch by gathering information about SpaceX and creating dashboards for the team.

The goal

Successfully train a machine learning model to predict **if SpaceX will reuse the first stage - which will determine the cost of each launch**



02

Methodology

Data source

	Description	Link
SpaceX REST API	Open source REST API for launch, rocket, core, capsule, starlink, launchpad and landing pad data	*http://github.com/r-spacex/SpaceX-API*
Falcon 9 Launch record	Wikipedia launch record data, include launch outcomes, sites, customers, version and date	*https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922*

Tools

Python

Most of the work will be conducted by using Python.

Several python libraries will be used, such as pandas, numpy, matplotlib, dash, scikit-learn, ...

SQL

Using SQL queries to quickly find important exploratory insights

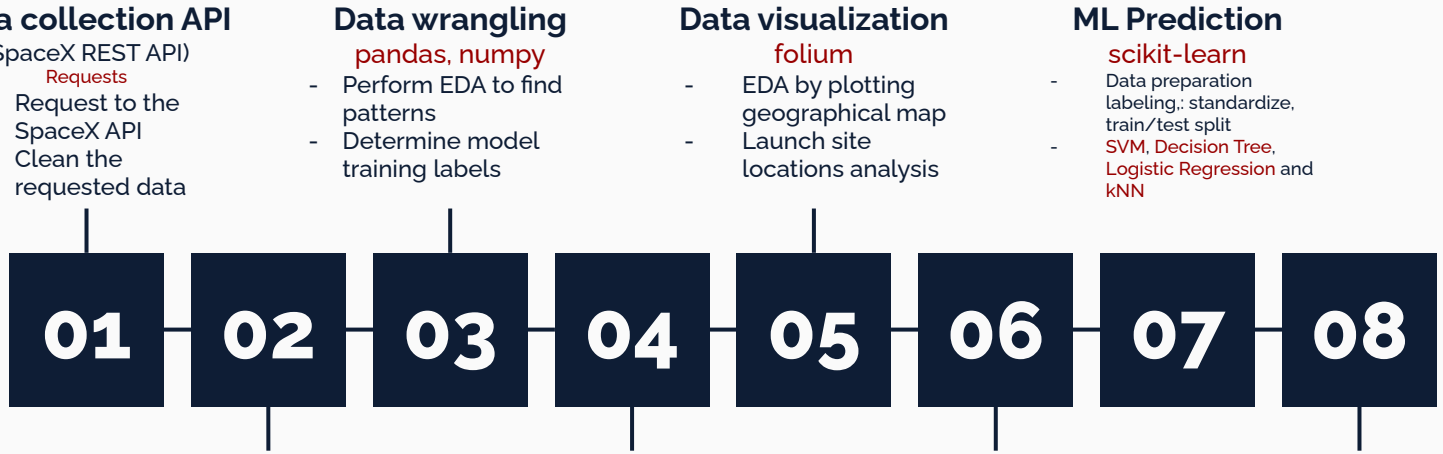
SQL syntax will be used directly in Python environment using *sqlalchemy*

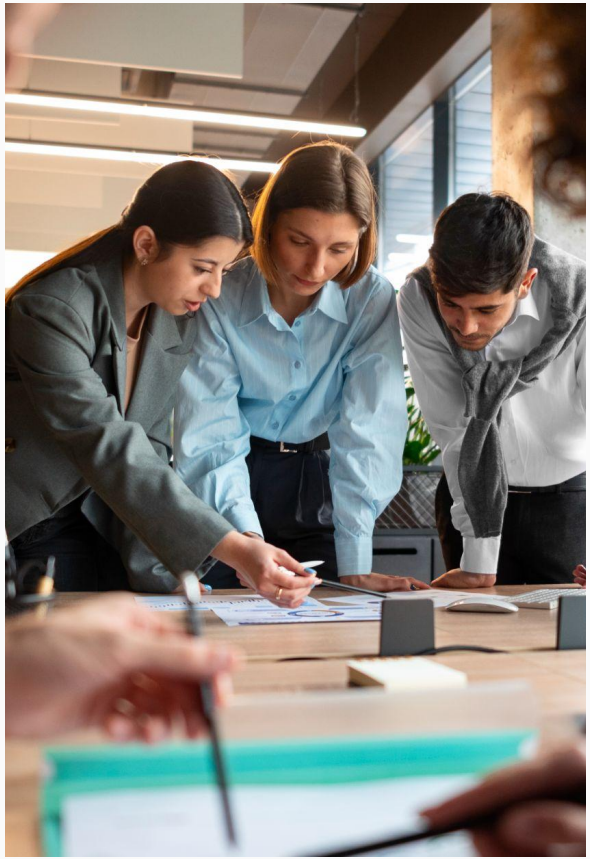
Visualization libraries

To support the project's plotting, multiple Python visualization libraries will be used. Such as:

matplotlib, seaborn, folium, plotly

Data analysis process

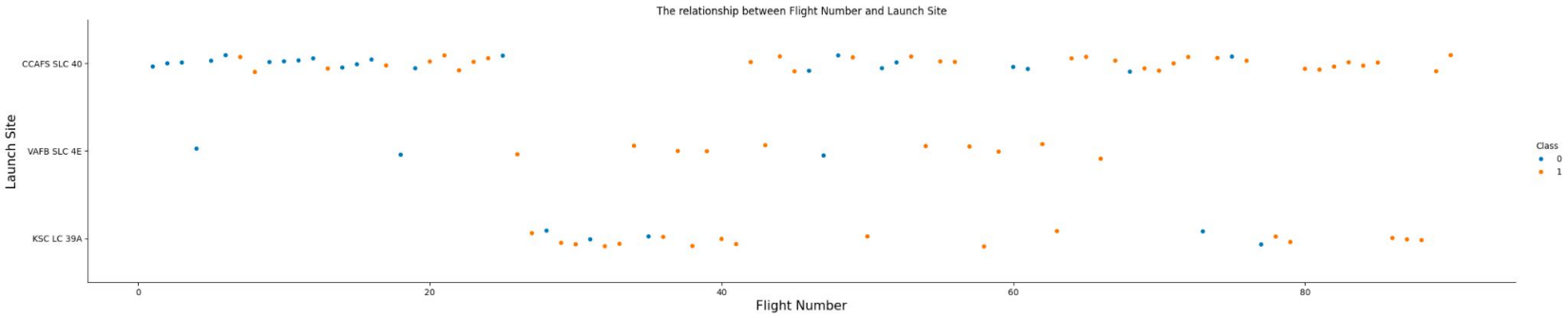




03

Findings

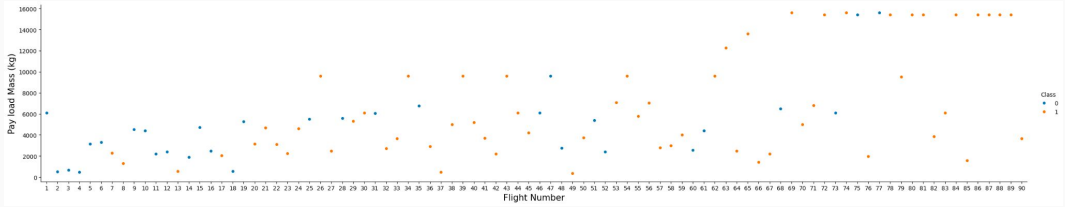
Exploratory research: Visualization

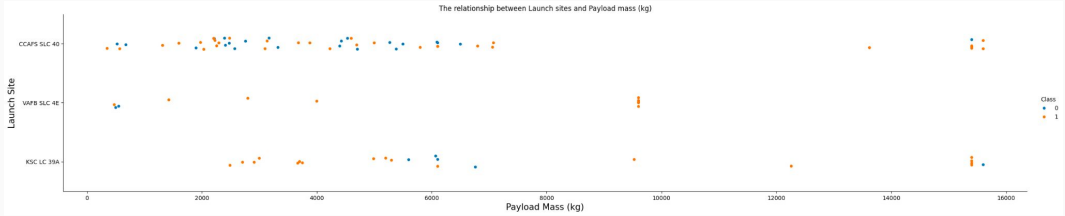


Later flights have higher success rate

- *The later the flight, the higher the chance of success*
- *From flight 78th, the success rate is 100%*

Exploratory research: Visualization



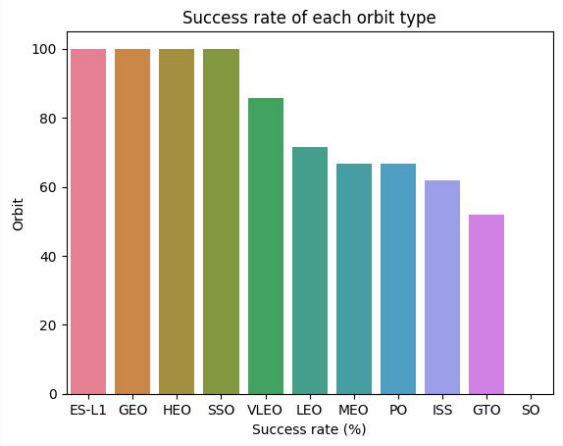


Higher payload mass results in better result

*By observing the **success rate** and the **relationship between Payload Mass and Number of Flights + Launch Site***

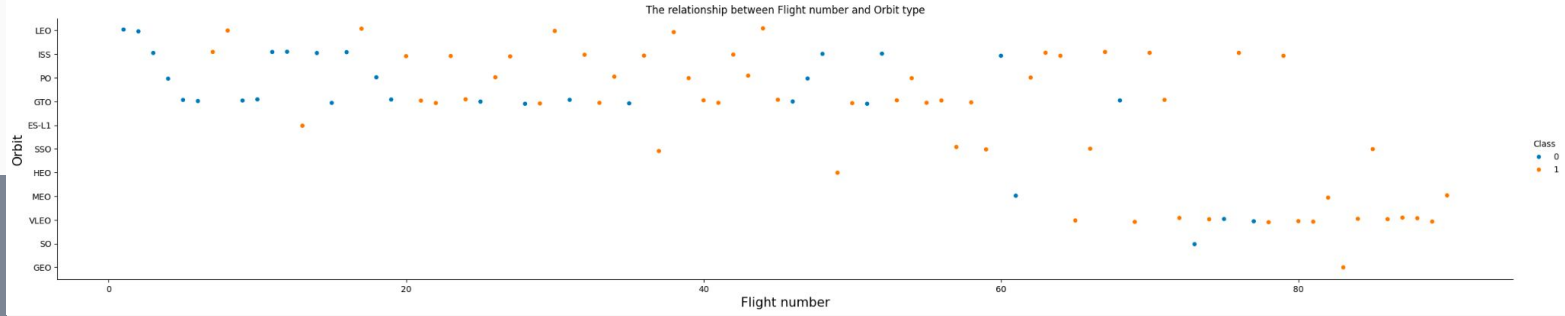
- Success rate for launches from >8000kg payload is very high (~87%)
- Two sites KSC LC-39A and VAFB SLC 4E has success rate of 77% while the other has 60%

Exploratory research: Visualization



Orbit types

- *For orbits that have at least 5 landings, **SSO** is the only orbit that has a success rate of **100%***
- ***VLEO** and **LEO** has significantly high success rate **~70-80%**, and their success rate seems to correlate to the number of flights*
- ***GTO** orbit has a low success rate of **50%***



Exploratory research: SQL

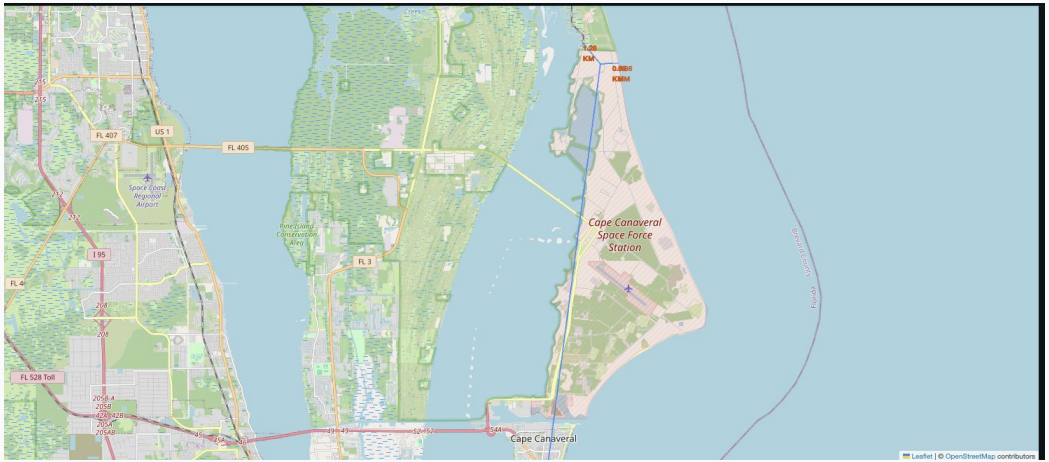
Mission_Outcome	Count
Success	98
Failure (in flight)	1
Success (payload status unclear)	1
Success	1

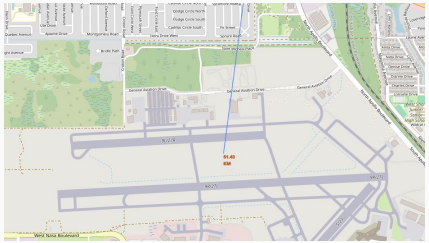
Landing_Outcome	Count
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

Various landing outcomes, but there's always a successful mission

- ***Many landing outcomes**: There are 8 types of landing outcomes*
- ***No attempt** happened the most (10 times).*
- *However, landing outcomes related to **drone ship** happened the most (11 times)*
- ***High mission success rate** (100/101), include the mission with payload status unclear*

Exploratory research: Mapping



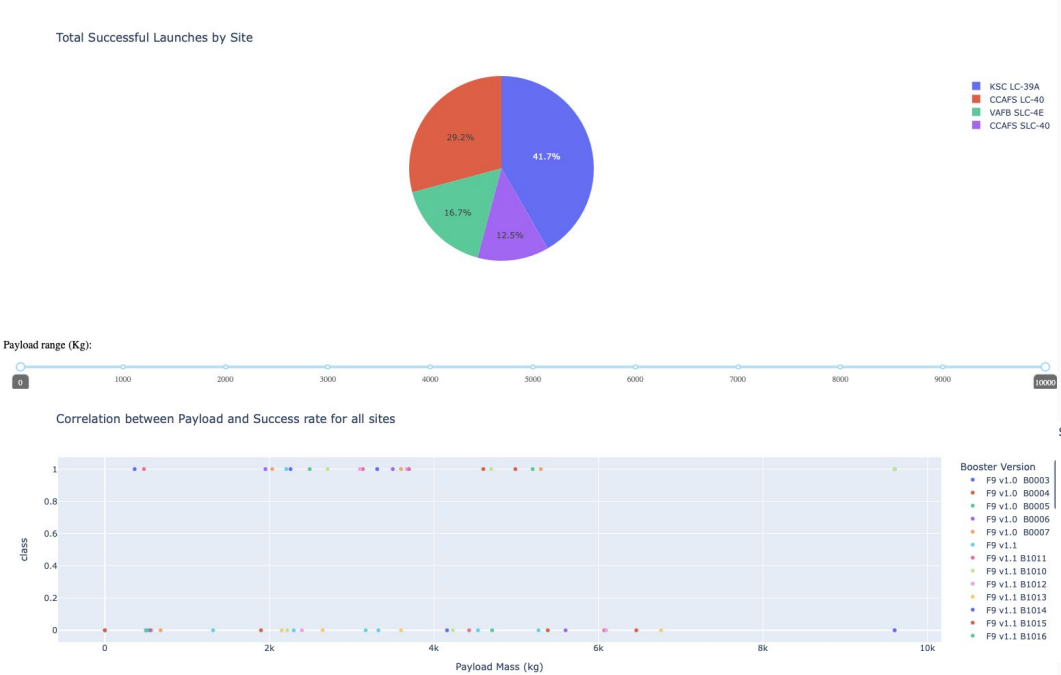


Launch sites has notable geographical characteristics

By observing:

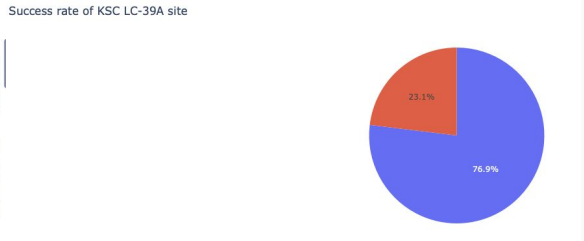
- Launch sites are **close to railways, highways and coastline**, which is important and convenient for transportation and logistics
- Launch sites are **far away from cities**, which is necessary for safety and security reasons

Dashboard: Mapping

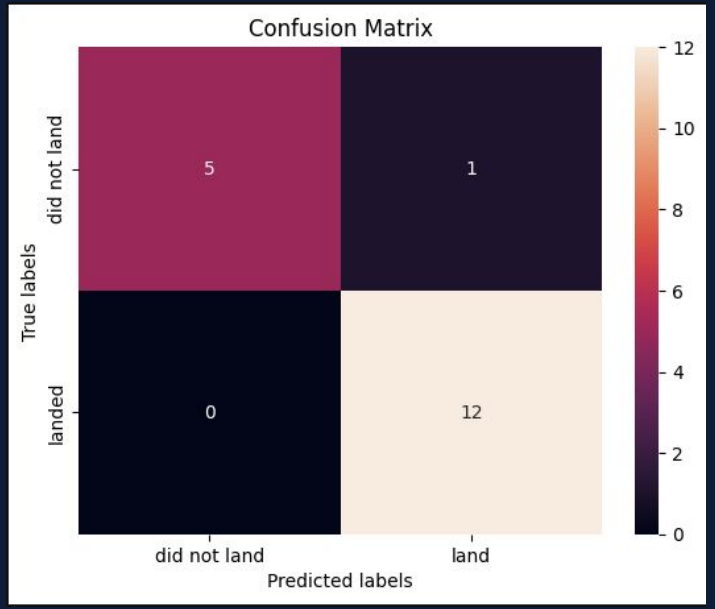


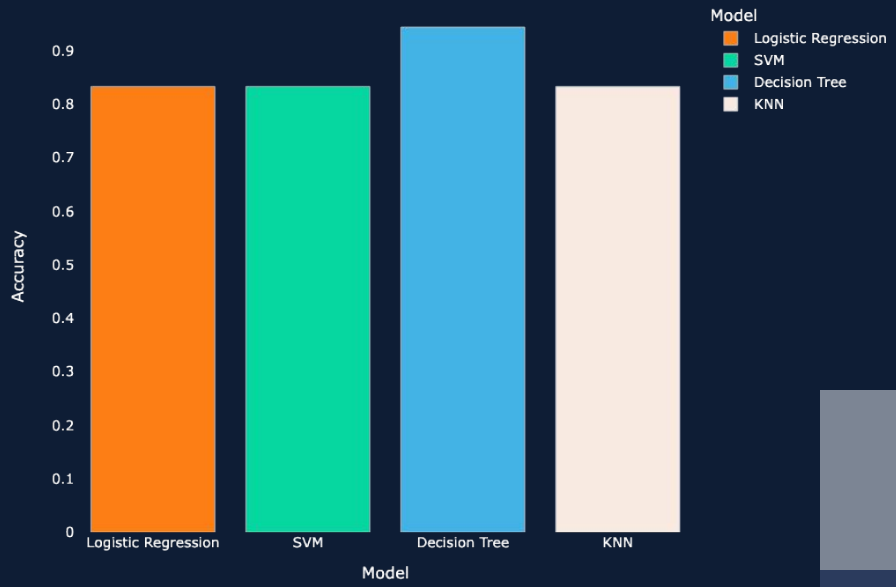
KSC LC-39A site dominates

- ***41%** of successful launches are from **KSC LC-39A** site*
- ***KSC LC-39A**, with **76.9%** success rate, is the only site that has a success rate more than 50%*



ML Predicting: Decision Tree

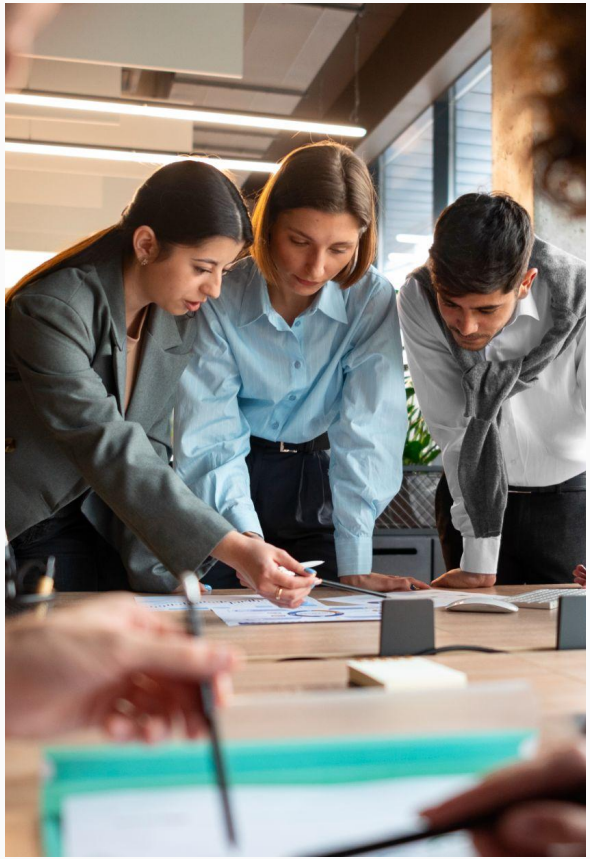




ML Predicting: Decision Tree

Classification Tree has the highest predicting accuracy

- Out of 18 samples, the model predicted 17 correct with a success rate of 94%
- In the wrong prediction:
  - *The model predicted the positive class incorrectly, which was a False Positive (FP case)*
  - *The model predicted a successfully landed case, but the actual value was a 'did not land' case*
- The other 3 methods: SVM, logistic regression and k nearest neighbor has a similar success rate of 83.3%



04

Conclusions

Conclusions

Prediction method: Decision Tree

The test shows an outstanding result of 94% accuracy for decision tree classification. Which should be used for future prediction.

Geographical aspects

Geography shows an essential impact to the launch sites. Most sites were placed close to transportations for logistics and were far from urban areas.

Promising future

Recent testing results show positive outcomes, which is an improvement compared to the past. Most missions were completed successfully.