CAPSTONE DATA SCIENCE

SPACEX: THE FALCON 9, UPWARDS AND BEYOND: A DATA SCIENCE STORY

ORANE SOUTH

April 07,2024

IBM Developer

SKILLS NETWORK

EXECUTIVE SUMMARY

Summary and Methodologies:

Falcon 9 the most cost-effective and reliable space deployment rocket currently.

- Data Collection.
  - Data cleaning, collection, Data wrangling
- EDA and Interactive visual analytics methodology
  - Predictive Analysis Methodology
- EDA with visualization results
- EDA with SQL results
- Interactive Map with Folium results
- Plotly Dash dashboard results
- Predictive Analysis Results

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

INTRODUCTION

- SpaceX advertises Falcon 9 rocket launches with a cost of 62 million dollars each which can be reused for the first stage.
- Others provide rockers at a cost of upward of 165 million dollars each.
- If we can determine the first stage will land successfully, we can determine the cost of a launch.
- There is a great advantage to predicting a successful Falcon 9 mission.
- SpaceX Falcon 9 Rockets changes the game, and this is confirmed via exploratory and predictive analysis of the data.

METHODOLOGY

Data Collection:	Exploring Data Analysis:	Exploratory Data:	Interactive Visual Analytics:	Machine Learning:
Request SpaceX past launch data. Data collection API Data collection API with Webscraping	Complete the EDA with SQL Exploring the Data using SQL	Analysis for Data Visualization.	Interactive Visual Analytics with Folium. Dashboard with Plotly Dash.	Predictive Analysis.
Collection	Perform Data Wrangling with SQL	Analyze and Present	Present Folium, Plotly	Present Results: Decision Tree, KNN.

DATA COLLECTION METHODOLOGY

Data Collection Goal: Using the data to predict whether SpaceX will attempt to land a rocket or not.

- Using the SpaceX REST API

Data Wrangling Goal: Transform the data into a clean dataset which provides meaningful and actionable data.

- Transform the SpaceX API using JSON objects.
- Convert the JSON objects to a data frame.

IBM Developer

SKILLS NETWORK

DATA COLLECTION AND DATA WRANGLING METHODOLGY

Dataframe Launch Data (All Boosters)

```python
# Show the head of the dataframe
data.head()
```

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCount	Serial	Longitude	Latitude
0	1	2006-03-24	Falcon 1	20	LEO	Kwajalein Atoll	None None	1	False	False	False	None	None	0	Merlin1A	167.743129	9.047721
1	2	2007-03-21	Falcon 1	None	LEO	Kwajalein Atoll	None None	1	False	False	False	None	None	0	Merlin2A	167.743129	9.047721
2	4	2008-09-28	Falcon 1	165	LEO	Kwajalein Atoll	None None	1	False	False	False	None	None	0	Merlin2C	167.743129	9.047721
3	5	2009-07-13	Falcon 1	200	LEO	Kwajalein Atoll	None None	1	False	False	False	None	None	0	Merlin3C	167.743129	9.047721
4	6	2010-06-04	Falcon 9	None	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1	0	B0003	-80.577366	28.561857

Data Wrangled and Clean Falcon 9

```python
#df= pd.DataFrame({ key:pd.Series(value) for key, value in launch_dict.items() })


df=pd.DataFrame(launch_dict)
df.head()
```

	Flight No.	Launch site	Payload	Payload mass	Orbit	Customer	Launch outcome	Version Booster	Booster landing	Date	Time
0	1	CCAFS	Dragon Spacecraft Qualification Unit	0	LEO	[[SpaceX], \n]	Success\n	F9 v1.0B0003.1	Failure	4 June 2010	18:45
1	2	CCAFS	Dragon	0	LEO	[[.mw-parser-output .plainlist ol,.mw-parser-o...	Success	F9 v1.0B0004.1	Failure	8 December 2010	15:43
2	3	CCAFS	Dragon	525 kg	LEO	[[NASA], ([COTS],)\n]	Success	F9 v1.0B0005.1	No attempt\n	22 May 2012	07:44
3	4	CCAFS	SpaceX CRS-1	4,700 kg	LEO	[[NASA], ([CRS],)\n]	Success\n	F9 v1.0B0006.1	No attempt	8 October 2012	00:35
4	5	CCAFS	SpaceX CRS-2	4,877 kg	LEO	[[NASA], ([CRS],)\n]	Success\n	F9 v1.0B0007.1	No attempt\n	1 March 2013	15:10

IBM Developer

SKILLS NETWORK

INTERACTIVE VISUAL ANALYTICS METHODOLOGY

Folium Map

- Focusing on analyzing launch site geo and proximities.
- Mapping launch side locations and their proximities on the Folium interactive map.
- Main goal is to look for patterns and relevant relationships between launch sites.

Plotly Dash: Dashboard

- Initiated using the Python Plotly Dash Package
- Key tool to visualize a interactive dashboard for all launch site locations.
- Excellent way to present finding in a non-static way and find patterns and insights

PREDICTIVE ANALYSIS METHODOLOGY

Building a Machine Learning Pipeline

Purpose: Predict First Stage of the Falcon 9 Successfully Landings.

Using:

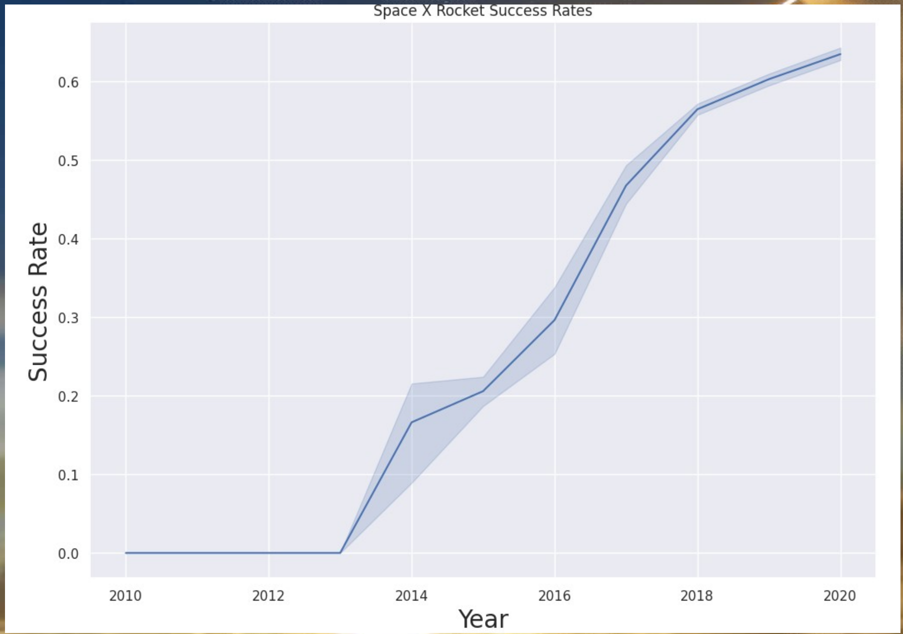
- Logistic Regression
- Support Vector Machines
- Decision Tree Classifier
- K nearest neighbors
- Output Confusion Matrix.

IBM Developer

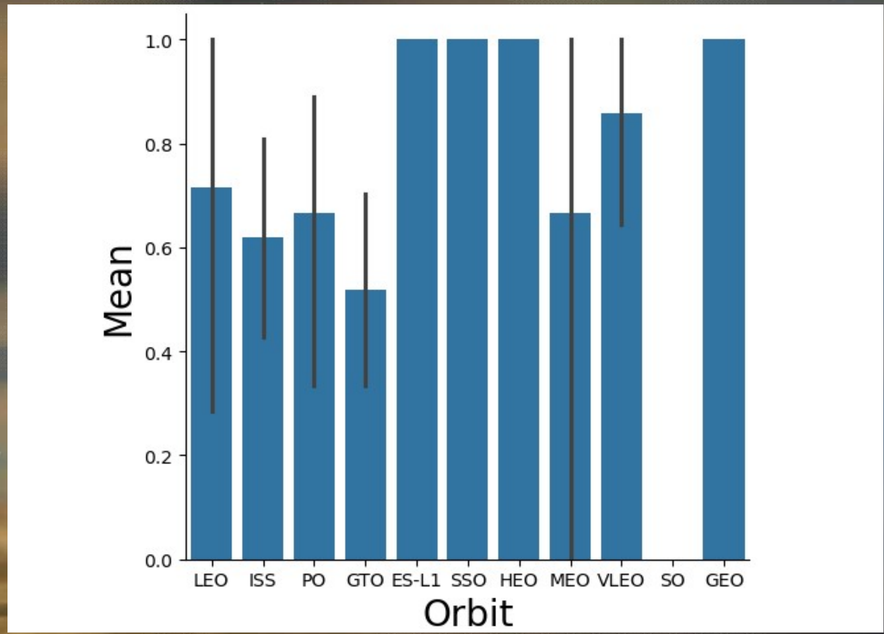
SKILLS NETWORK

EDA WITH VISUALIZATION RESULTS (ONE) SUCCESS RATE ORBIT TYPE

Launch Success Rate Over Time



ES-L1,SSO, HEO,GEO No Failed First Stage Landing
LEO, ISS,PO,GTO and MEO Some Failed First Stage Landing
SO No Successful First Stage Landing



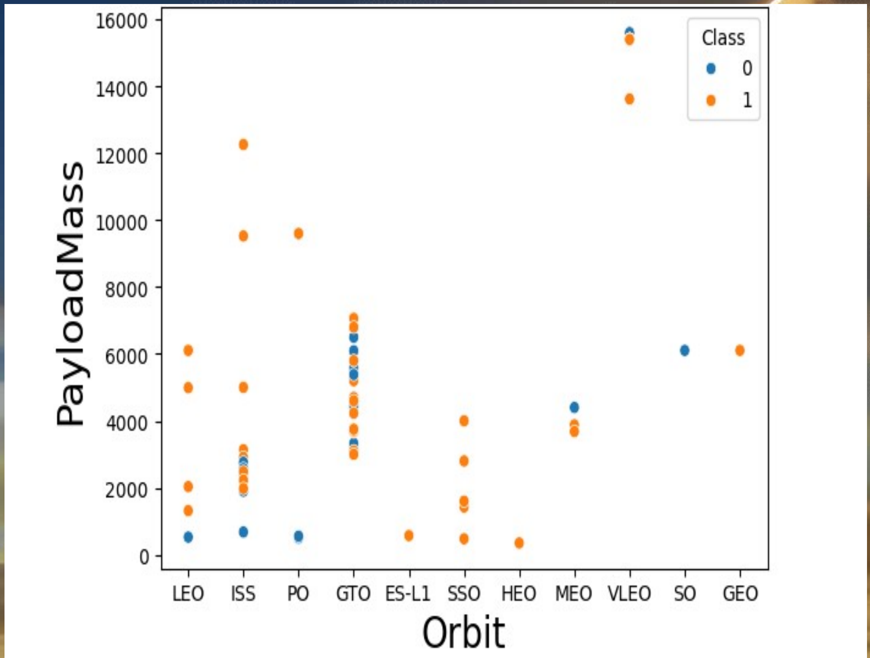
IBM Developer

SKILLS NETWORK

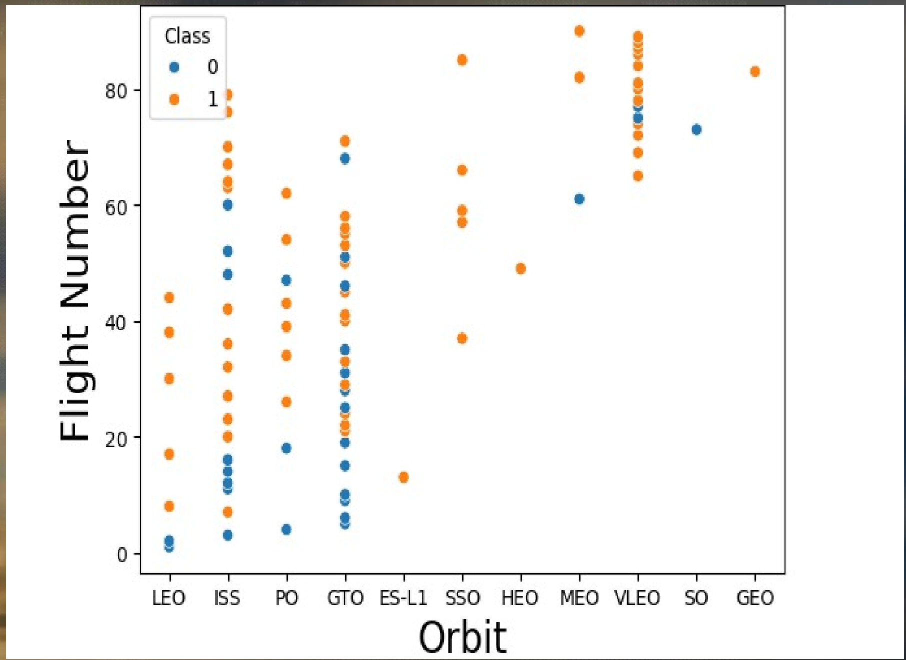
EDA WITH VISUALIZATION RESULTS (TWO)

Visualization of Success Regardless of Payload Mass or Flight Number. Note GTO and ISS.

PayloadMass Vs.Orbit



Flight Number Vs. Orbit



IBM Developer

SKILLS NETWORK

EDA With SQL Results: (One)

Data Illustration of attempts, successes and failures for the Falcon 9

DATE	time_utc_	booster_version	launch_site	payload	payload_mass_kg_	orbit	customer	mission_outcome	landing_outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

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

Month_of_2015	Landing _Outcome	Booster_Version	Launch_Site
01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

Landing_Outcome	Count
Success (drone ship)	8
Success (ground pad)	6

IBM Developer

SKILLS NETWORK

EDA With SQL Results: (Two)

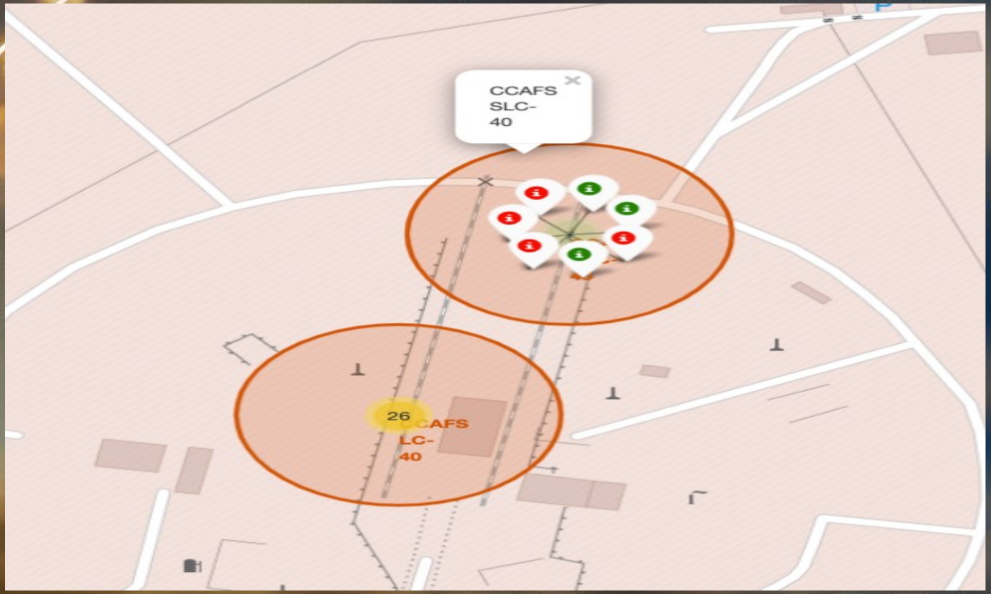
Data Illustration of attempts, successes and failures for the Falcon 9

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing _Outcome
04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
08-12-2010	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

Folium Results: (One)

Visual illustration via Folium of the Falcon 9 Site Locations

	Launch Site	Lat	Long	class
46	KSC LC-39A	28.573255	-80.646895	1
47	KSC LC-39A	28.573255	-80.646895	1
48	KSC LC-39A	28.573255	-80.646895	1
49	CCAFS SLC-40	28.563197	-80.576820	1
50	CCAFS SLC-40	28.563197	-80.576820	1
51	CCAFS SLC-40	28.563197	-80.576820	0
52	CCAFS SLC-40	28.563197	-80.576820	0
53	CCAFS SLC-40	28.563197	-80.576820	0
54	CCAFS SLC-40	28.563197	-80.576820	1
55	CCAFS SLC-40	28.563197	-80.576820	0

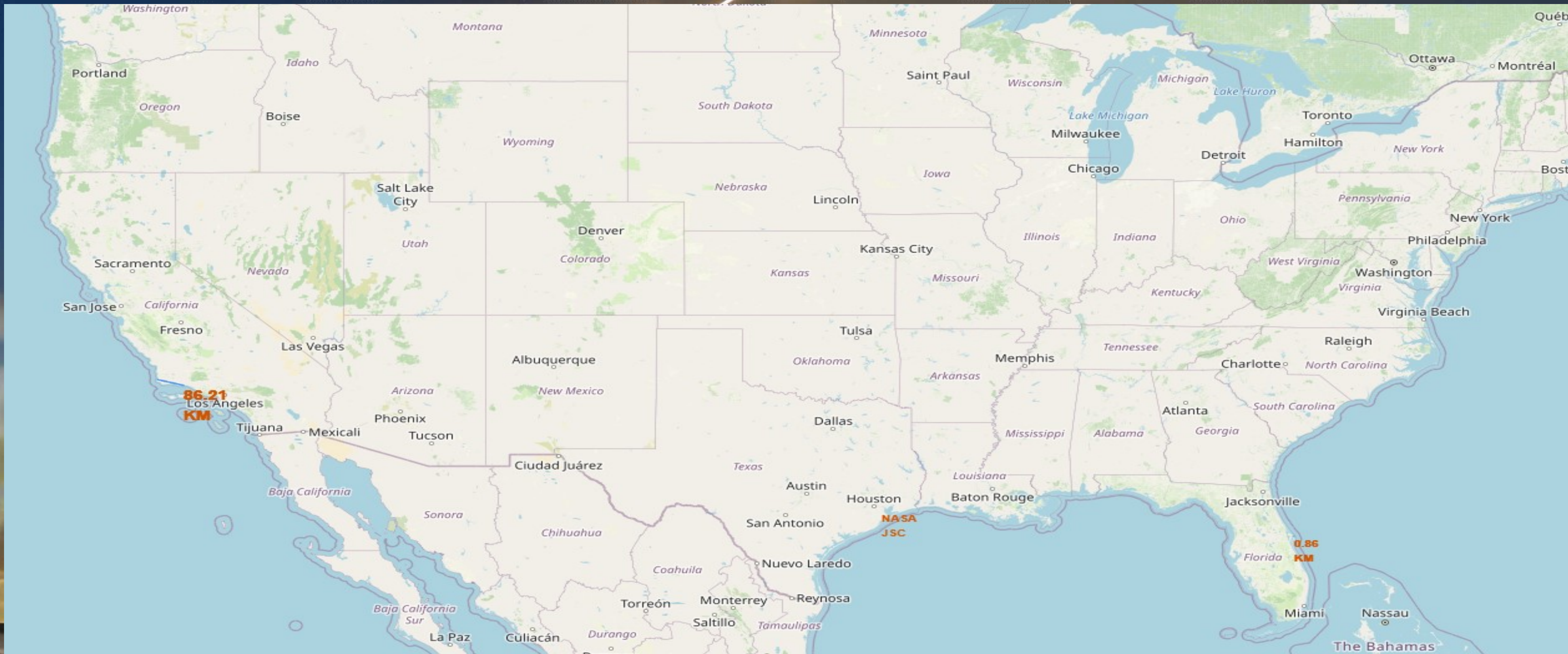


IBM Developer

SKILLS NETWORK

Folium results: (Two)

Visual illustration via Folium of the Falcon 9 Site Geo-Locations

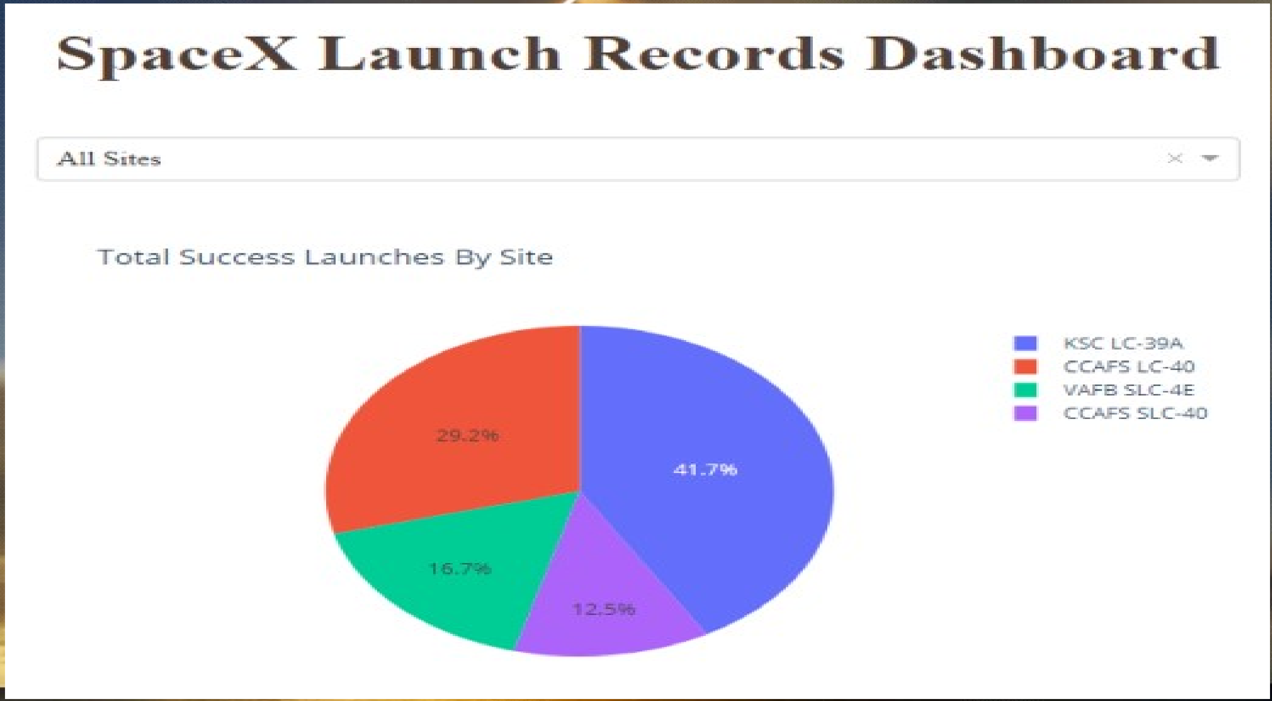


IBM Developer

SKILLS NETWORK

Plotly Dash Dashboard: (One)

Visual Dashboard of Successful Launches for all Locations as a Percentage of All Launches



IBM Developer

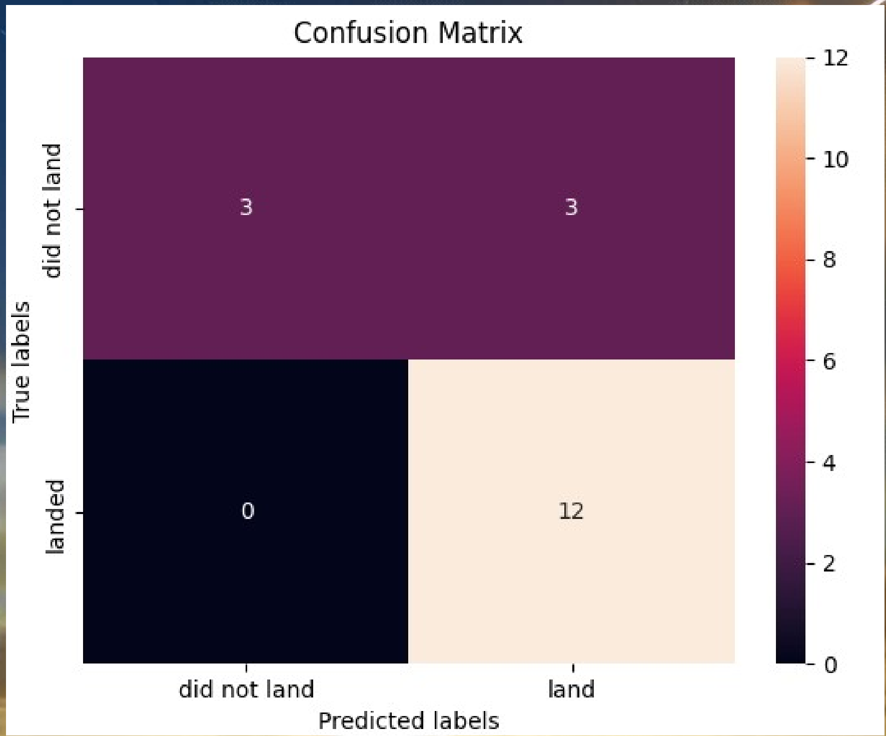
SKILLS NETWORK

Plotly Dash Dashboard: (Two)

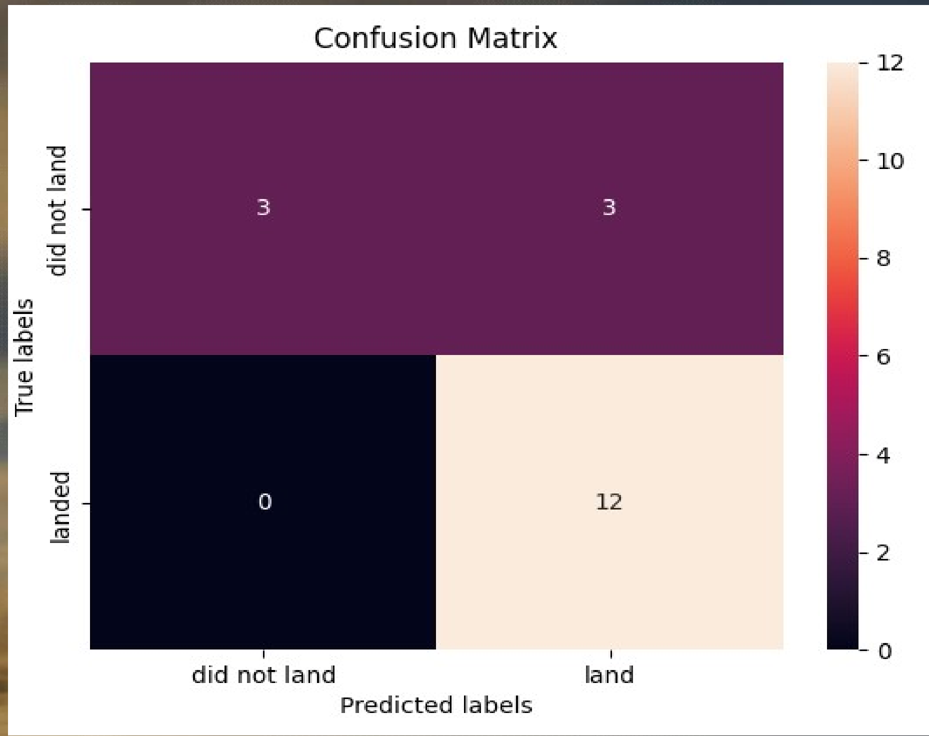
Correlation between Payload and Successes for each Launch Site.
2,000 kg to 5,000 kg highest success rate. FT Booster Version largest success rate.

Predictive Analysis: (One)

Accuracy on the test data using the method score: Logreg Cv. Accuracy

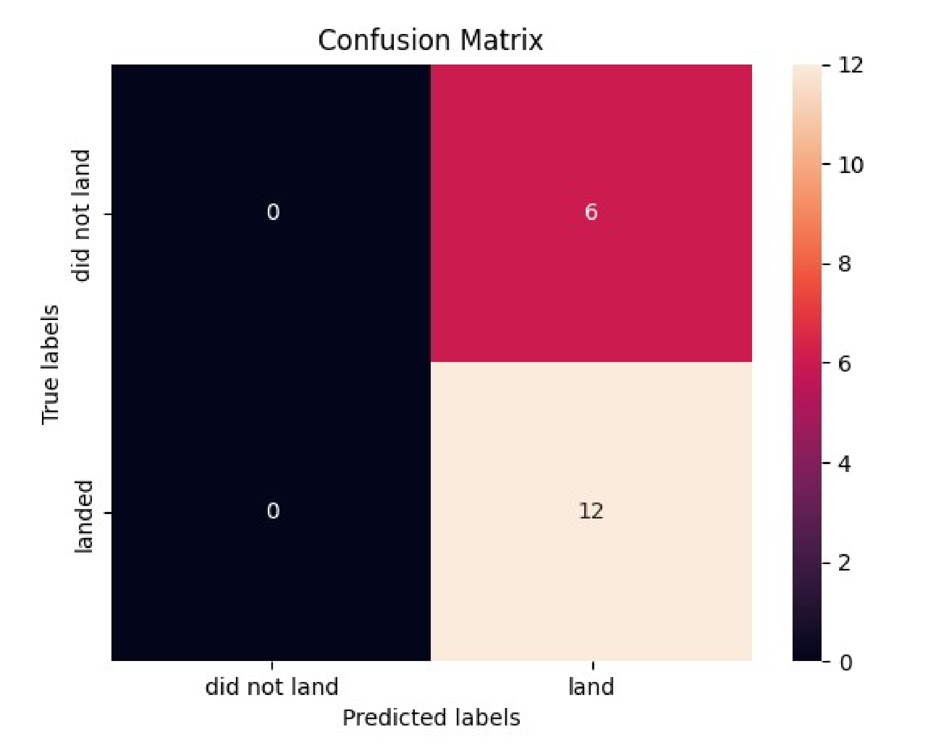


Accuracy on the test data using the method score: Decision Tree Cv.Accuracy

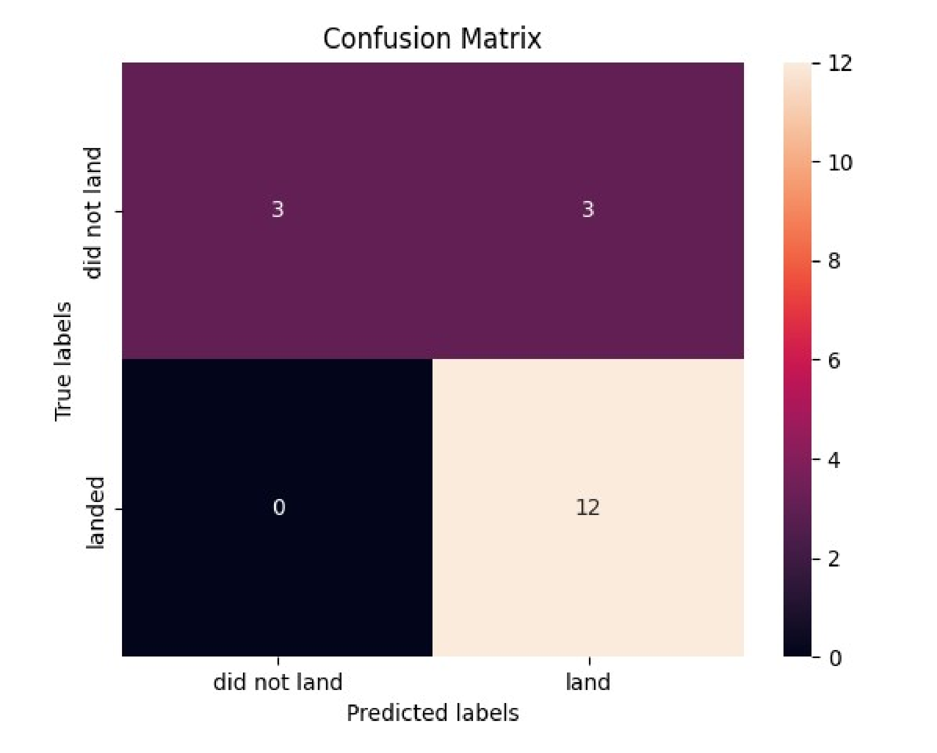


Predictive Analysis: (Two)

Accuracy on the test data using the method score: SVM.Cv



Accuracy on the test data using the method score: Knn_Cv

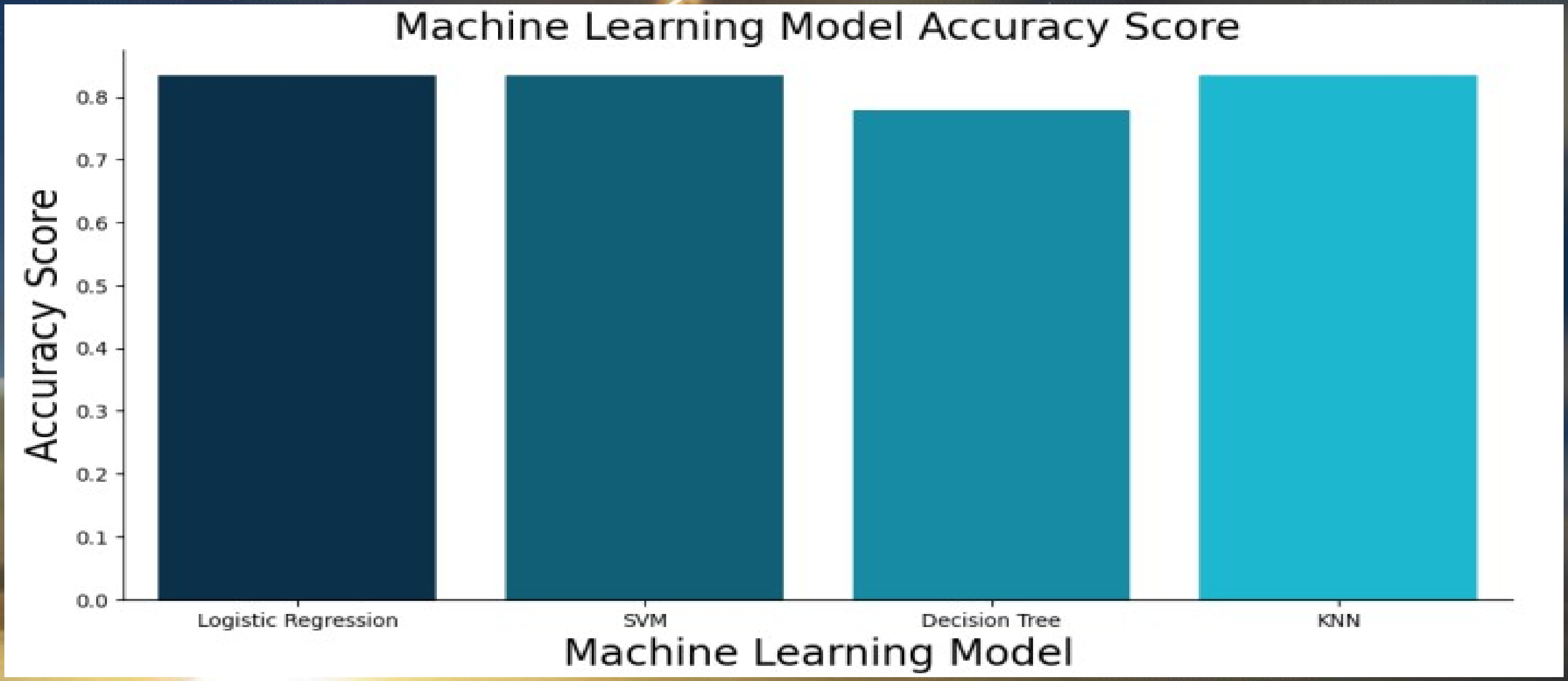


IBM Developer

SKILLS NETWORK

Predictive Analysis: (Three)

All four models performed well with only Decision Tree performing at less than 80%.



CONCLUSION

- Point 1: Machine Learning models can be used to predict Falcon 9 first stage landing outcomes,effectively.
- Point 2: While not perfect Spacex Falcon 9 Rockets have become more and more effective over time.
- Point 3: As a business model reusing first stage falcon 9 rockets is cost effective and keeps SpaceX ahead of the curve with its competitors.

APPENDIX

- https://github.com/oasouth/Capstone-Project-IBM/blob/main/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/edadataviz.ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/jupyter-labs-eda-sql-coursera_sqllite%20(3).ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/jupyter-labs-spacex-data-collection-api.ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/jupyter-labs-webscraping%20(1).ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/lab_jupyter_launch_site_location%20(1).ipynb
- https://github.com/oasouth/Capstone-Project-IBM/blob/main/labs-jupyter-spacex-Data%20wrangling%20(2).ipynb