



IBM Data Science Capstone Project

Jumbalo Dobbins

September 22, 2023

OUTLINE



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

EXECUTIVE SUMMARY



- The objective of this project was to predict if a Falcon 9 first stage will land successfully. Based on research, the following points have been determined:
 - Launch Site KSC LC-39A has the highest launch success rate. However launch site CCAFS LC-40 has the highest quantity of successful launches.
 - Payload masses between 2,000 and 4,000 kg have the highest success rate.
 - Payload masses between 6,000 and 8,000 kg have the lowest success rate.
 - The Falcon F9 booster version FT has the highest launch success rate.
 - The SVM machine learning method provides the highest accuracy and AUC when determining if a land will be successful or not.

INTRODUCTION



- The objective was to determine if the Falcon 9 first stage will land successfully.
 - Being able to predict if the first stage will land will determine the cost of a launch and if an alternate company should bid against SpaceX
- The nature of this problem is classification
 - i.e. will a launch be successful or unsuccessful

METHODOLOGY



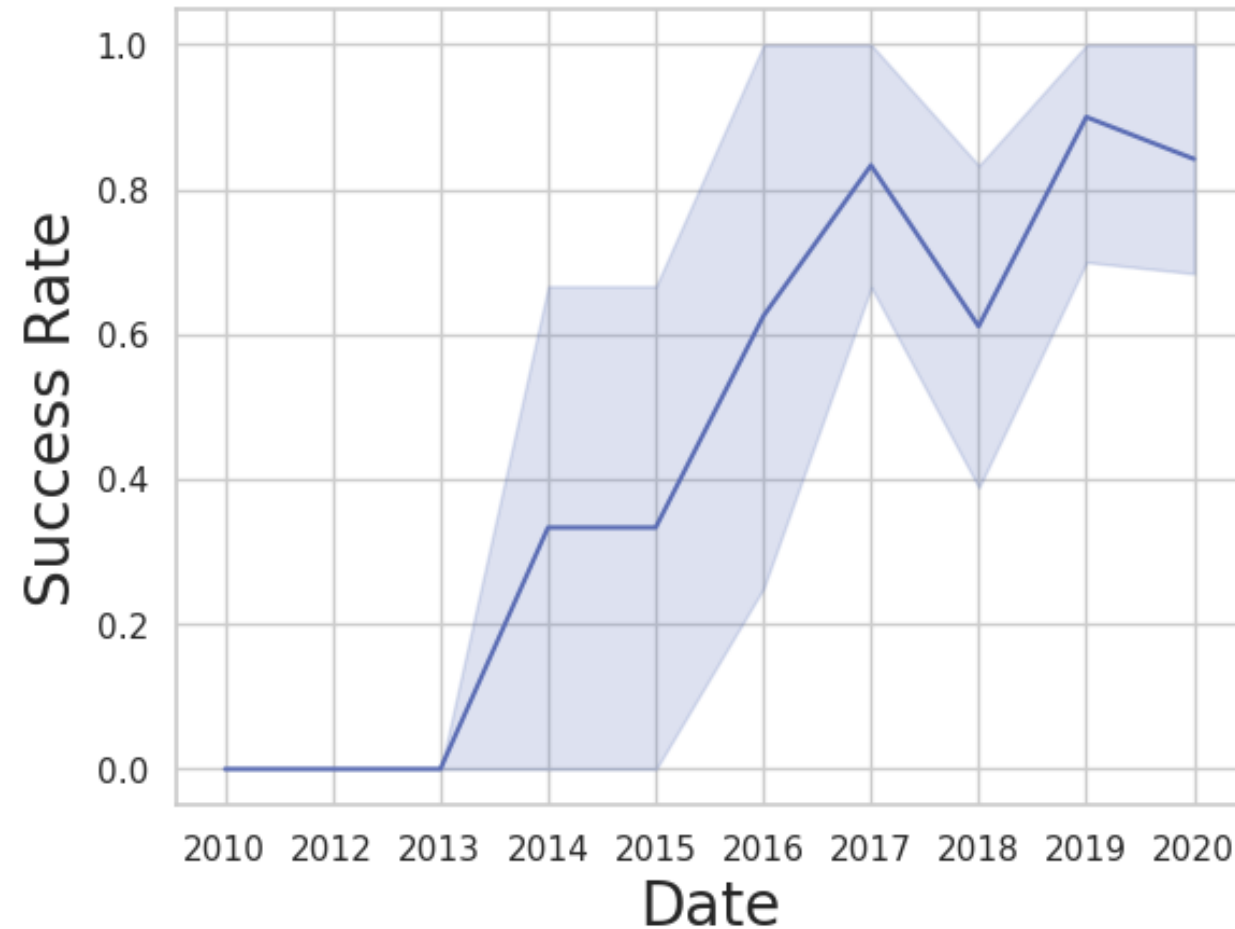
- Data sources include:
 - Open source Falcon 9 launch data from <https://api.spacexdata.com/v4/launches/past>
 - Falcon 9 historical launch records obtained through webscraping from a Wikipedia page titled "List of Falcon 9 and Falcon Heavy launches" (https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches)
- Classification was used to analyze the data by creating a "Outcome" column that classified if the launch would be a success or not.

RESULTS

- The data was organized into a pandas dataframe and visualized via seaborn plots and folium maps, and interactive plotly dashboards.
- Launch Site KSC LC-39A has the highest success rate of launches. However launch site CCAFS LC-40 has the highest quantity of successful launches.
- GEO orbits have the highest success rate, however VLEO orbits have the highest quantity of successful launches
- Payload masses between 2,000 and 4,000 kg have the highest success rate.
 - Payload masses between 6,000 and 8,000 kg have the lowest success rate.
- The Falcon F9 booster version FT has the highest launch success rate.
- The SVM machine learning method provides the highest accuracy and AUC when determining if a land will be successful or not.

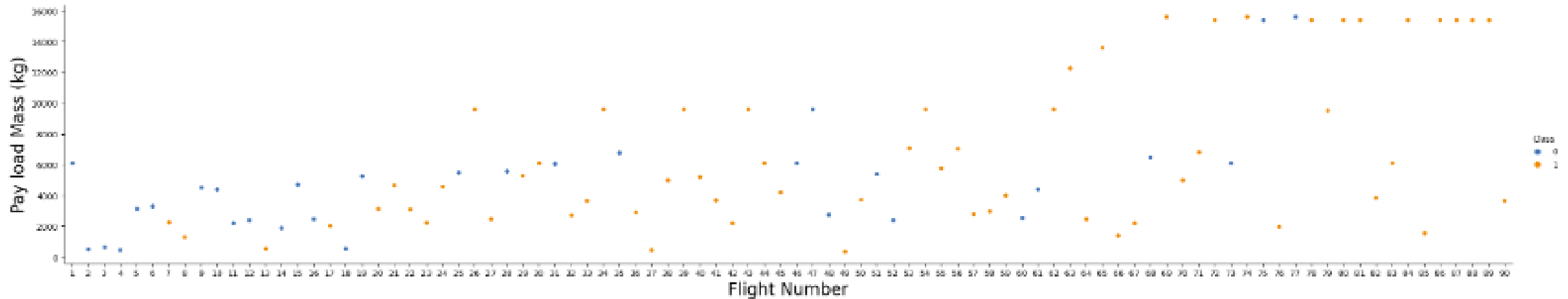
RESULTS

The success rate of launches unsurprisingly increased with time



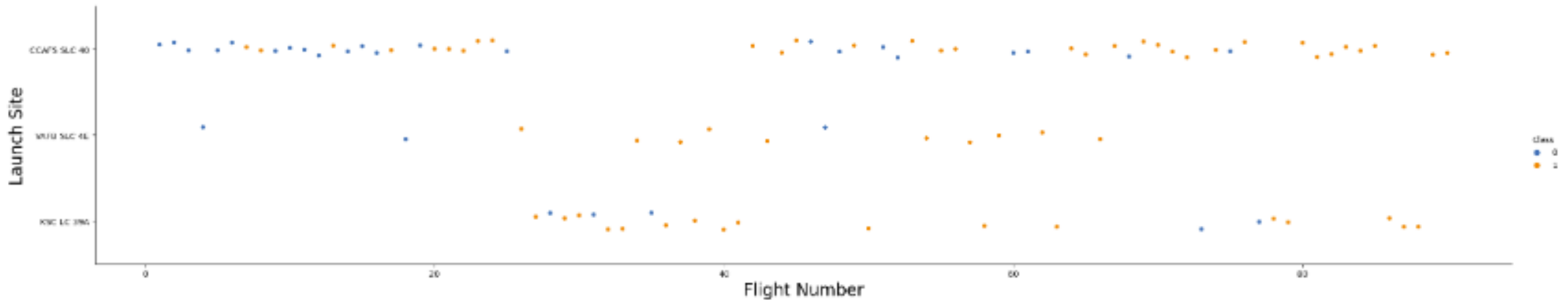
RESULTS

The Payload Mass of the launches increased with time.
The 2,000 to 4,000 kg range has the highest launch success rate



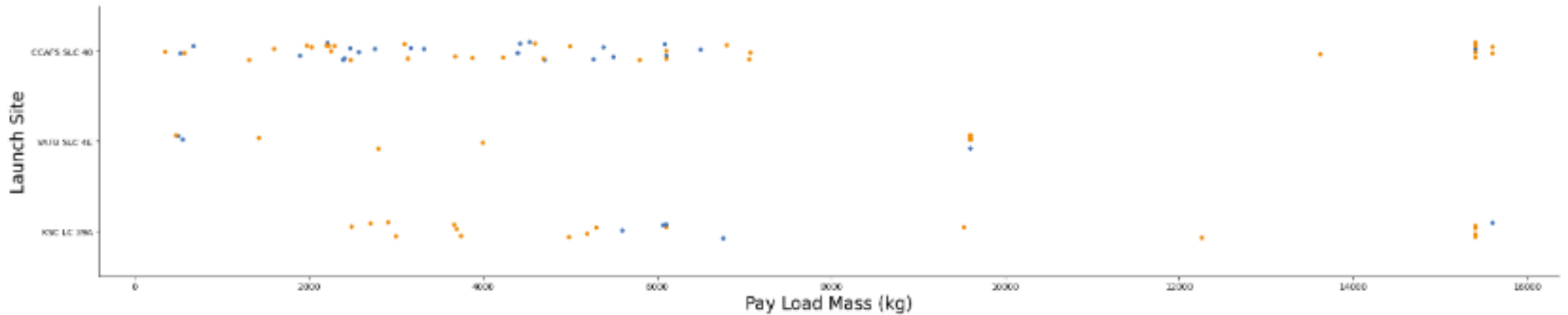
RESULTS

Amount of successful and unsuccessful launches from each launch site.



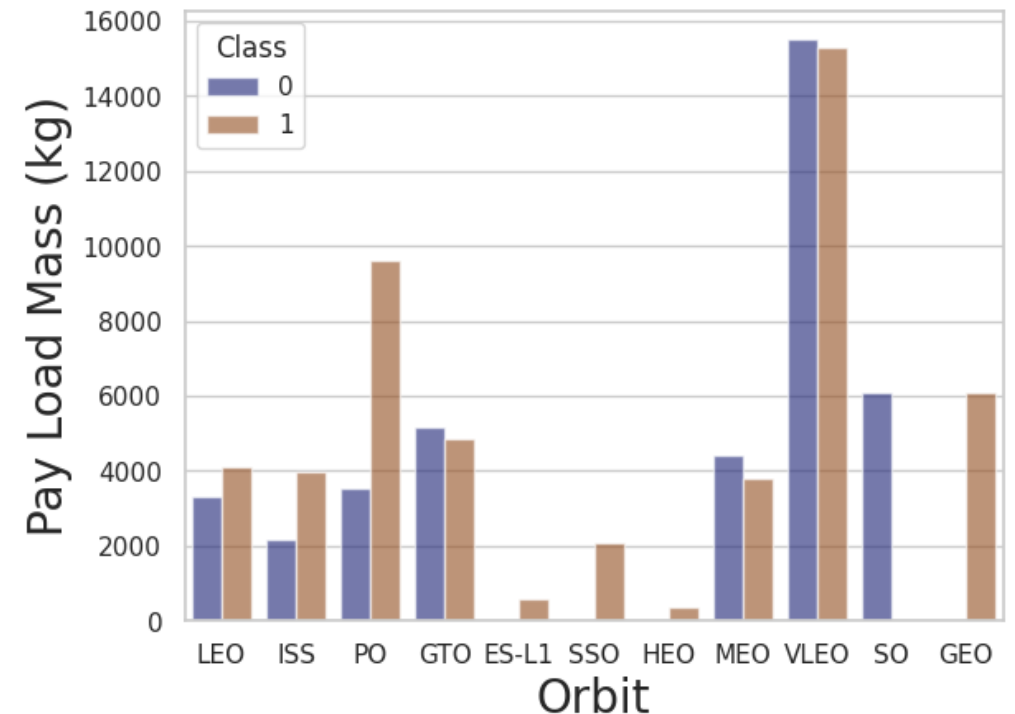
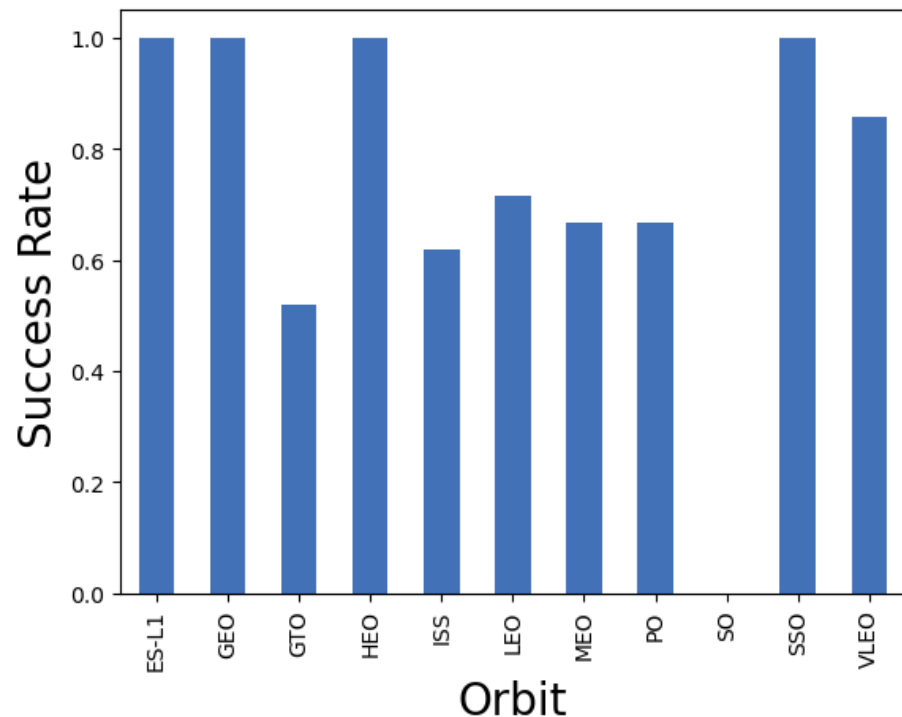
RESULTS

Payload masses of both successful and unsuccessful launches from each launch site.



RESULTS

Success rate of launches conducted in each orbit and the amount of successful and unsuccessful launches in each orbit with their payload mass. The smaller the payload mass the higher the success rate. Geo orbits have the highest success rate, however VLEO orbits have the highest quantity of successful launches



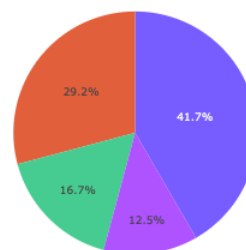
DASHBOARD – ALL SITES

SpaceX Launch Records Dashboard

All Sites

×

Success Count for all launch sites

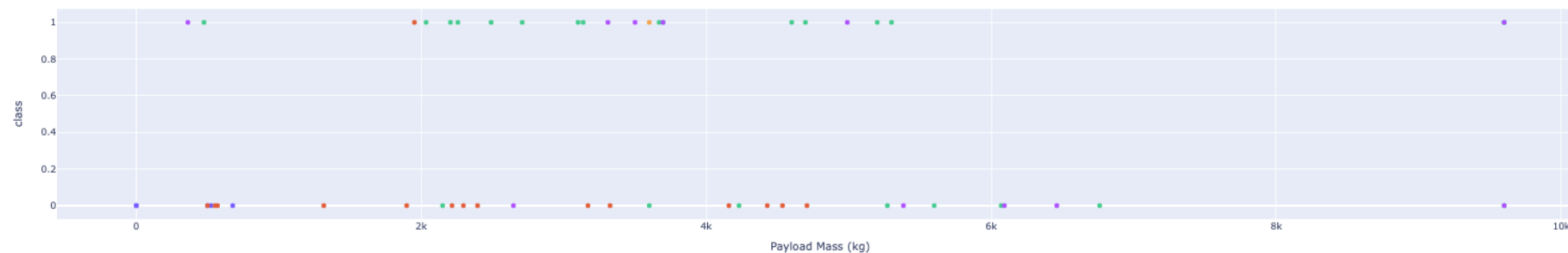


KSC LC-39A
CCAFS LC-40
VAFB SLC-4E
CCAFS SLC-40

Payload range (Kg):



Success count on Payload mass for all sites



Booster Version Category
v1.0
v1.1
FT
B4
B5

DASHBOARD – KSC LC-39A

SpaceX Launch Records Dashboard

KSC LC-39A

✕

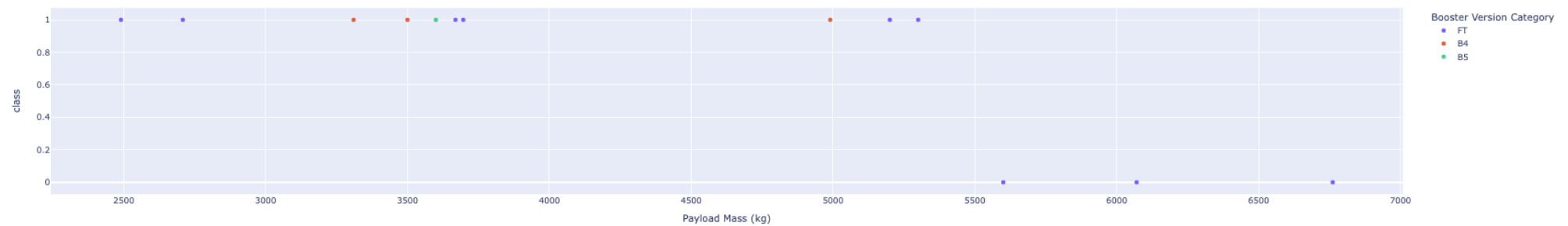
Total Success Launches for site KSC LC-39A



Payload range (Kg):



Success count on Payload mass for site KSC LC-39A

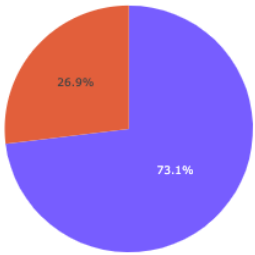


DASHBOARD – CCAFS LC-40

SpaceX Launch Records Dashboard

CCAFS LC-40

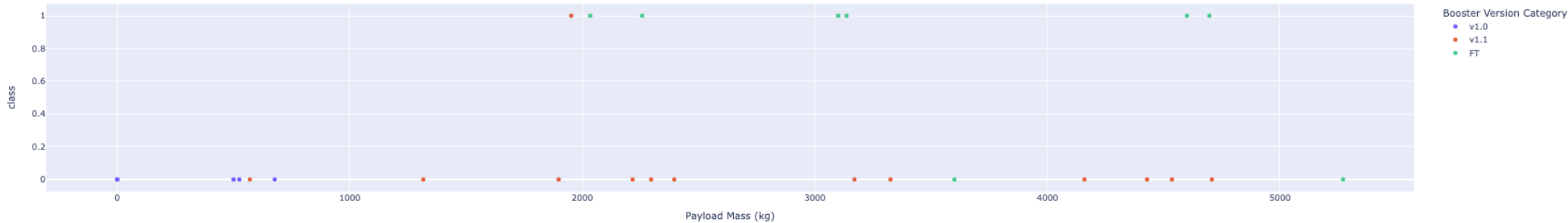
Total Success Launches for site CCAFS LC-40



Payload range (Kg):



Success count on Payload mass for site CCAFS LC-40



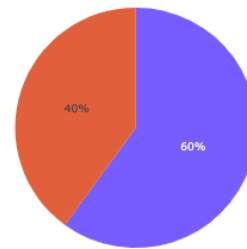
DASHBOARD – VAFB SLC-4E

SpaceX Launch Records Dashboard

VAFB SLC-4E

X

Total Success Launches for site VAFB SLC-4E

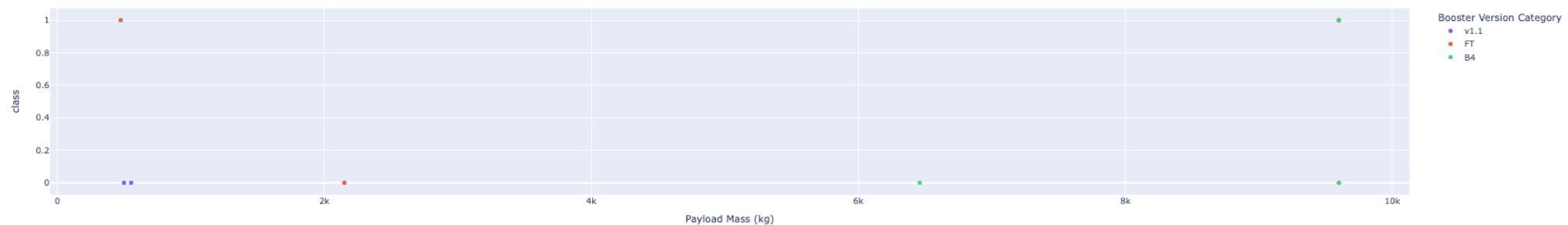


0
1

Payload range (Kg):



Success count on Payload mass for site VAFB SLC-4E



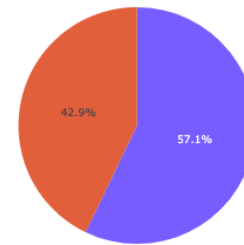
DASHBOARD – CCAFS SLC-40

SpaceX Launch Records Dashboard

CCAFS SLC-40

X

Total Success Launches for site CCAFS SLC-40

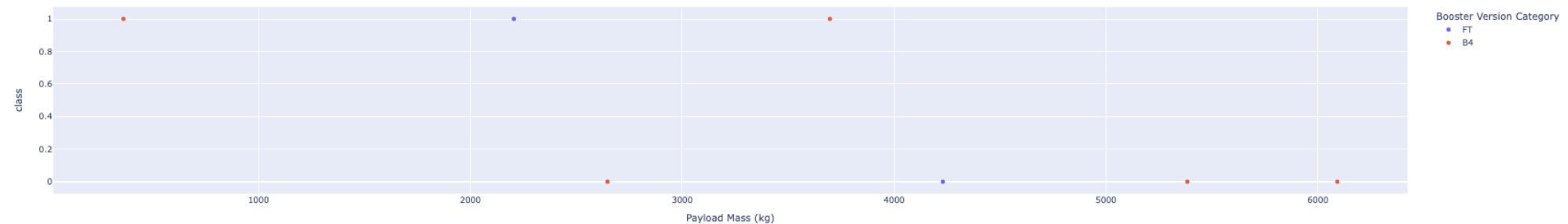


0
1

Payload range (Kg):



Success count on Payload mass for site CCAFS SLC-40



DISCUSSION



- Based on the conditions of a proposed SpaceX launch, it can be reasonably predicted if the launch will be successful or not.
- Using the SVM machine learning method, predictions can be made with an accuracy of 83.3% and precision of 80%.
 - SVM also have an AUC of 95.83%, F1-score of 88.89%, and Recall of 100%
- If the model predicts the SpaceX launch will be unsuccessful given the proposed conditions, a counterbid can be made.

OVERALL FINDINGS & IMPLICATIONS

Launches the contain the following have the highest chance of success:

- Launch is conducted at either KSC LC-39A or CCAFS LC-40
- Launch is into GEO or VLEO orbits
- Payload mass is between 2,000 and 4,000 kg
- The Falcon F9 booster version FT is being used.

Implications

- If a proposed launch is not being conducted following these parameters, a counter bid can be submitted that predicts a SpaceX launch failure.

CONCLUSION



- Launch Site KSC LC-39A has the highest success rate of launches. However launch site CCAFS LC-40 has the highest quantity of successful launches.
- Geo orbits have the highest success rate, however VLEO orbits have the highest quantity of successful launches.
- The highest rate of successful launches have a payload mass between 2,000 and 4,000 kg
- The Falcon F9 booster version FT has the highest launch success rate.
- The SVM machine learning method provides the highest accuracy and AUC when determining if a land will be successful or not

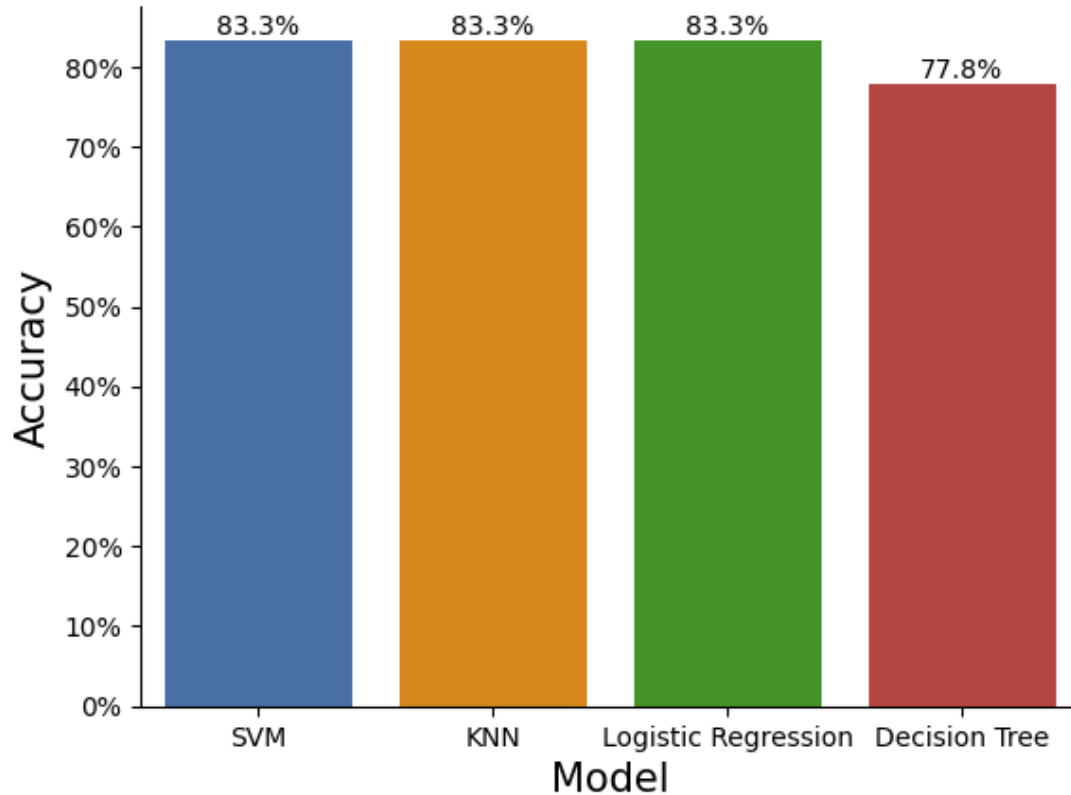
APPENDIX

- IBM Y'all

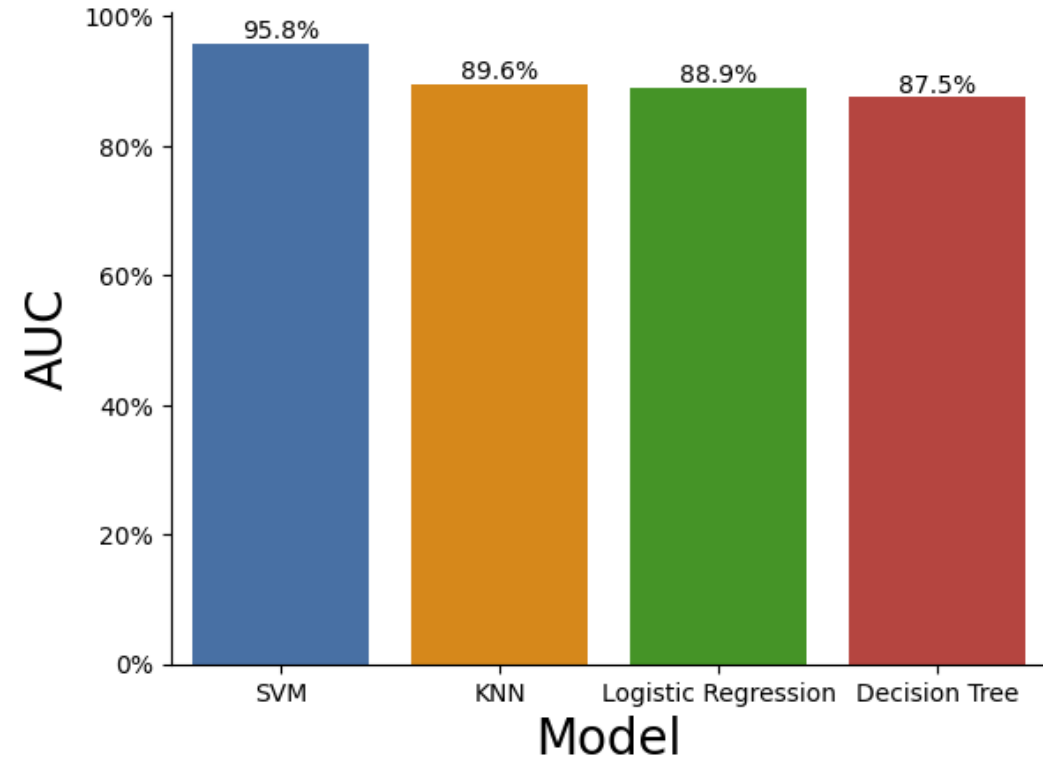


APPENDIX

Model Accuracy on the Test Set



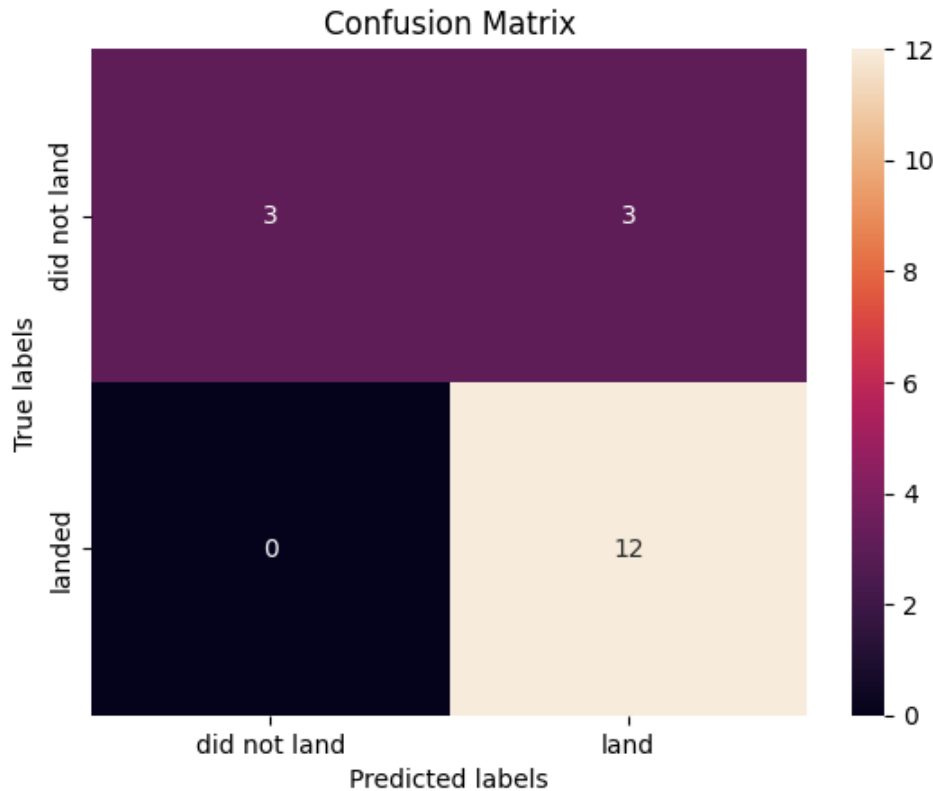
Model AUC on the Test Set



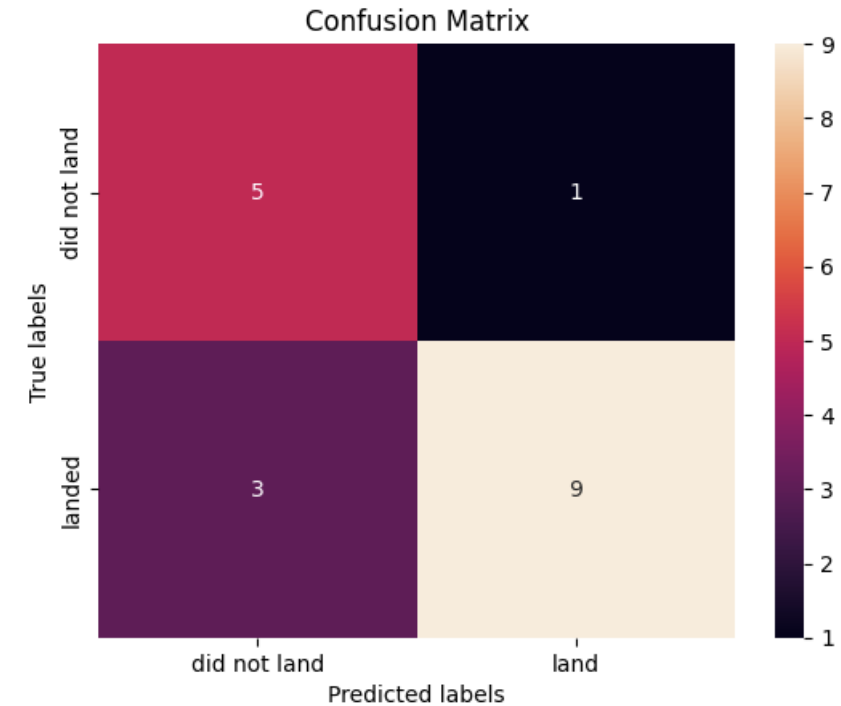
APPENDIX

	AUC	F1	Accuracy	Recall	Precision
SVM	0.958333	0.888889	0.833333	1.00	0.8
KNN	0.895833	0.888889	0.833333	1.00	0.8
Logistic Regression	0.888889	0.888889	0.833333	1.00	0.8
Decision Tree	0.875000	0.818182	0.777778	0.75	0.9

APPENDIX

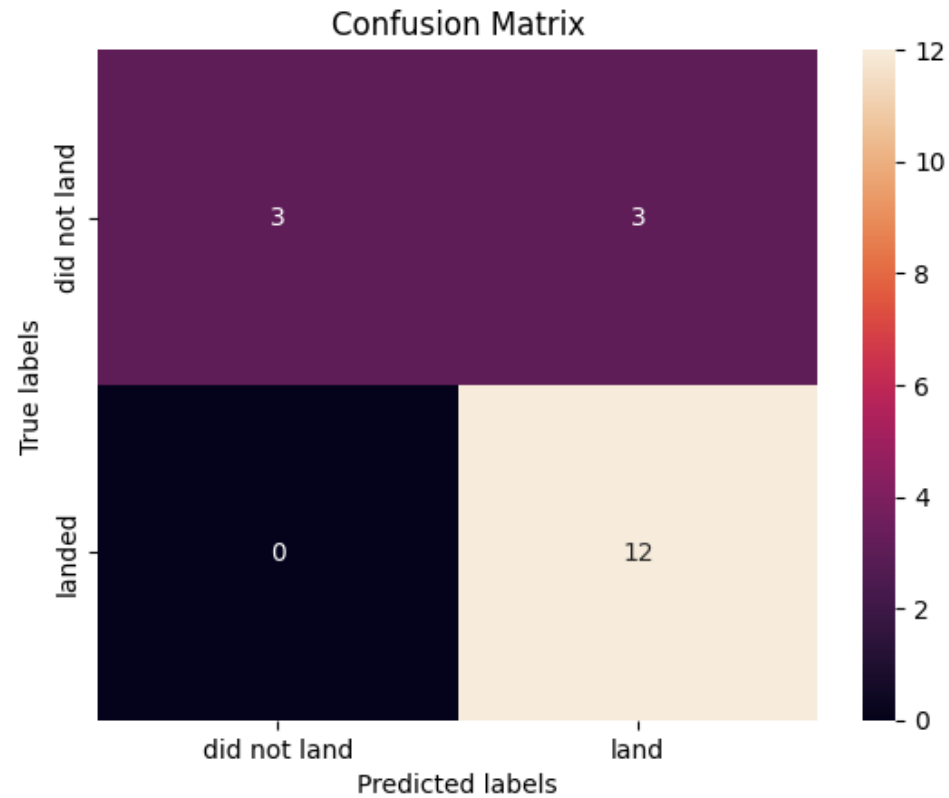


Model Evaluation Metrics for KNN:



Model Evaluation Metrics for DecisionTree:

APPENDIX



Model Evaluation Metrics for SVM:

