Space X Falcon 9 First Stage Landing Prediction

Christopher Nindyo

10 Oktober 2022

IBM Developer

SKILLS NETWORK

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

EXECUTIVE SUMMARY

- This report analyzes how data science can predict the SpaceX landing
- To answer this problem, we look it as problem that can be solved with collected data from the past and then solve them with machine learning
- Our results showed that we can predict with machine learning model with accuracy 83% to tell which the launch is success or not
- From safety reason, this study also showed that the launch site far from city but people can still reach the site.

INTRODUCTION

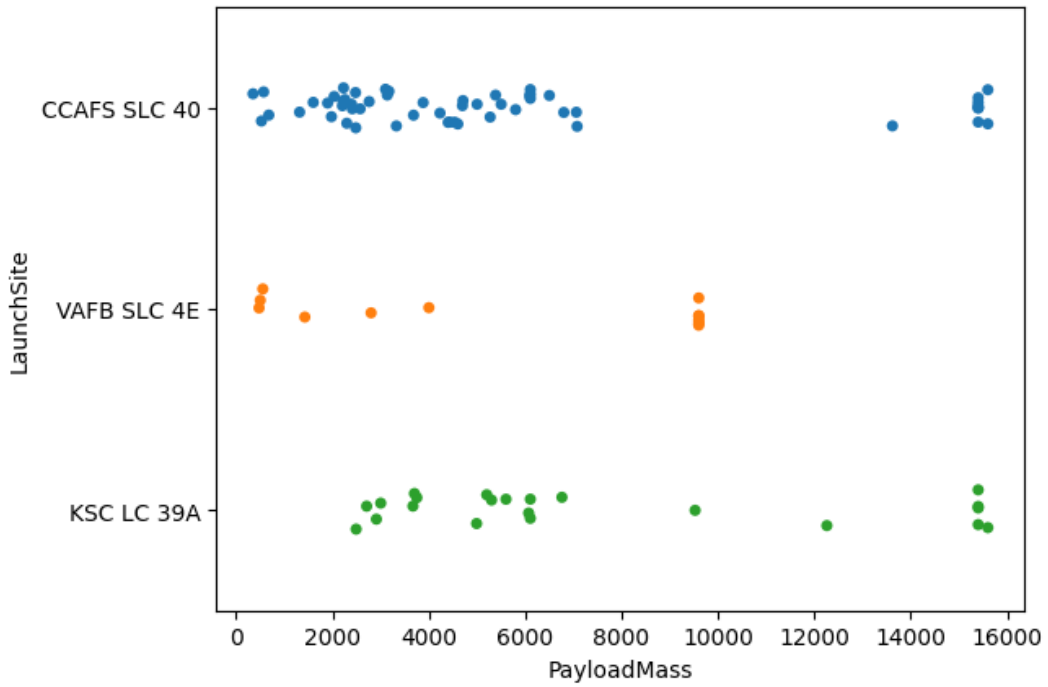
- We will predict if the Falcon 9 first stage will land successfully
- If we can determine if the first stage will land, we can determine the cost of a launch

METHODOLOGY

- Data Collection from SpaceX REST API
- Cleansing the data to be ready analyzed
- Exploratory Data Analysis (EDA) to find some patterns in the data and determine what would be the label for training supervised
- Visualization of the relationship between the variables of Space X Dataset
- Finding an optimal location for building a launch site
- Determine the best model for Space X Falcon 9 First Stage Landing Prediction

RESULTS

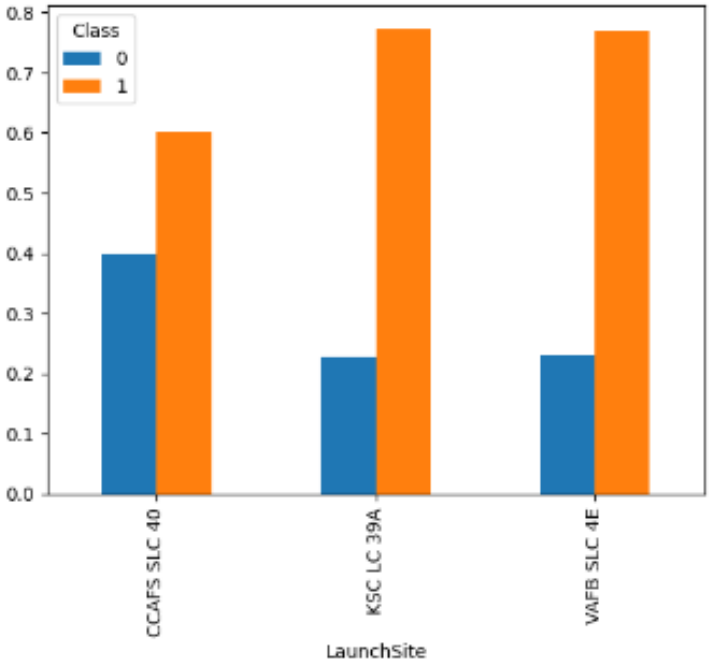
The relationship between Payload and Launch Site



There is no particular strong relationship between launch sites and their payload mass. If we see the graph above :

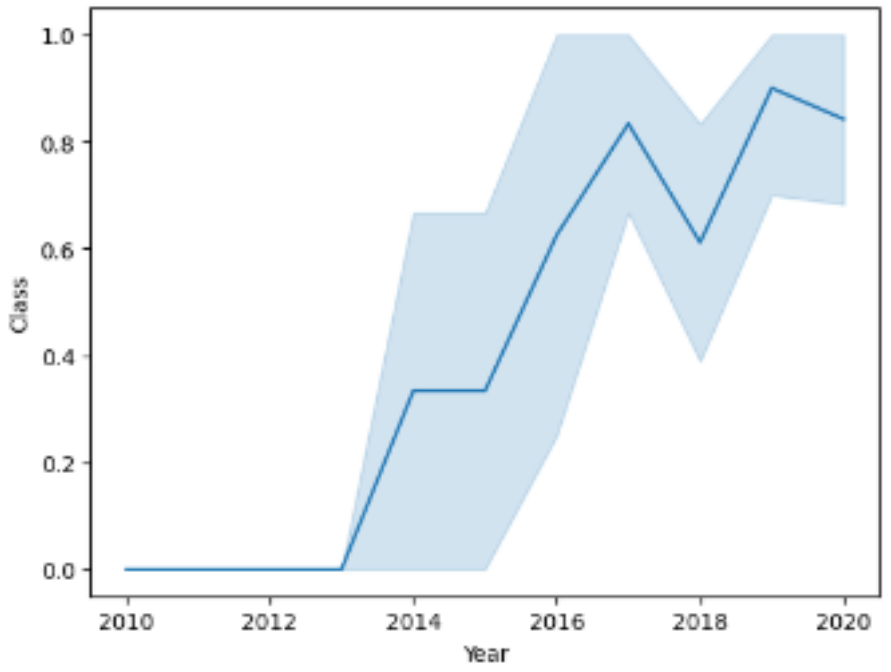
- CCAFS SLC 40 : payload mass around 0 - 8000 and 13000 - 16000
- VAFB SLC 4E : payload mass around 0 - 5000 and 9000 - 10000
- KSC LC 39A : payload mass around 2000 - 16000

The relationship between success rate of each orbit type



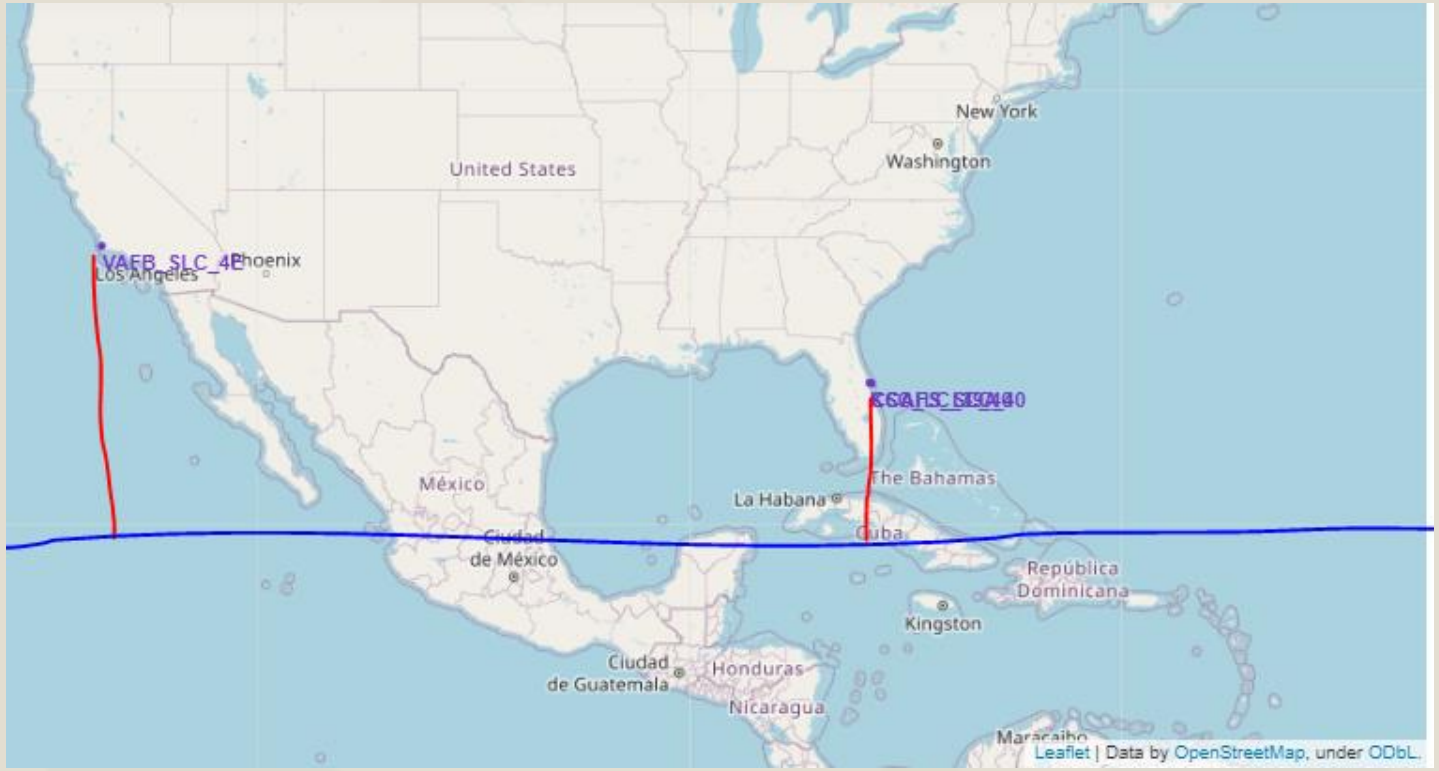
KSC LC 39A and VAFB SLC 4E have higher success rate than CCAFS SLC 40

The launch success yearly trend



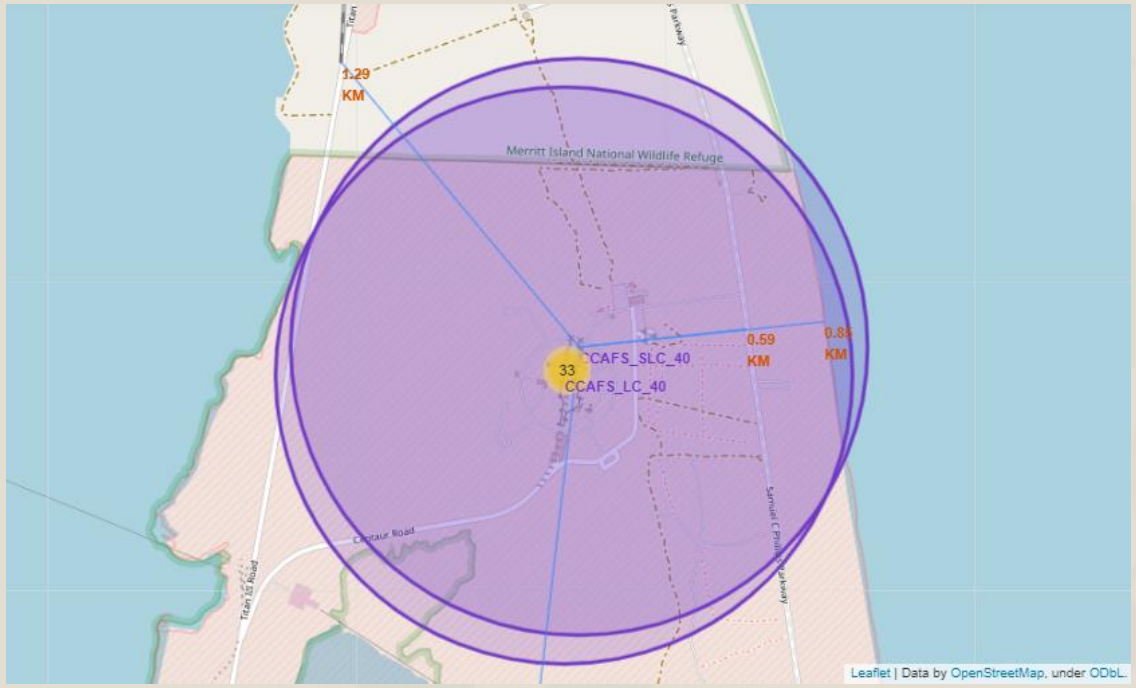
The Succeess rate since 2013 kept increasing untull 2020. The most improve in success rate is in 2013 - 2017 period.

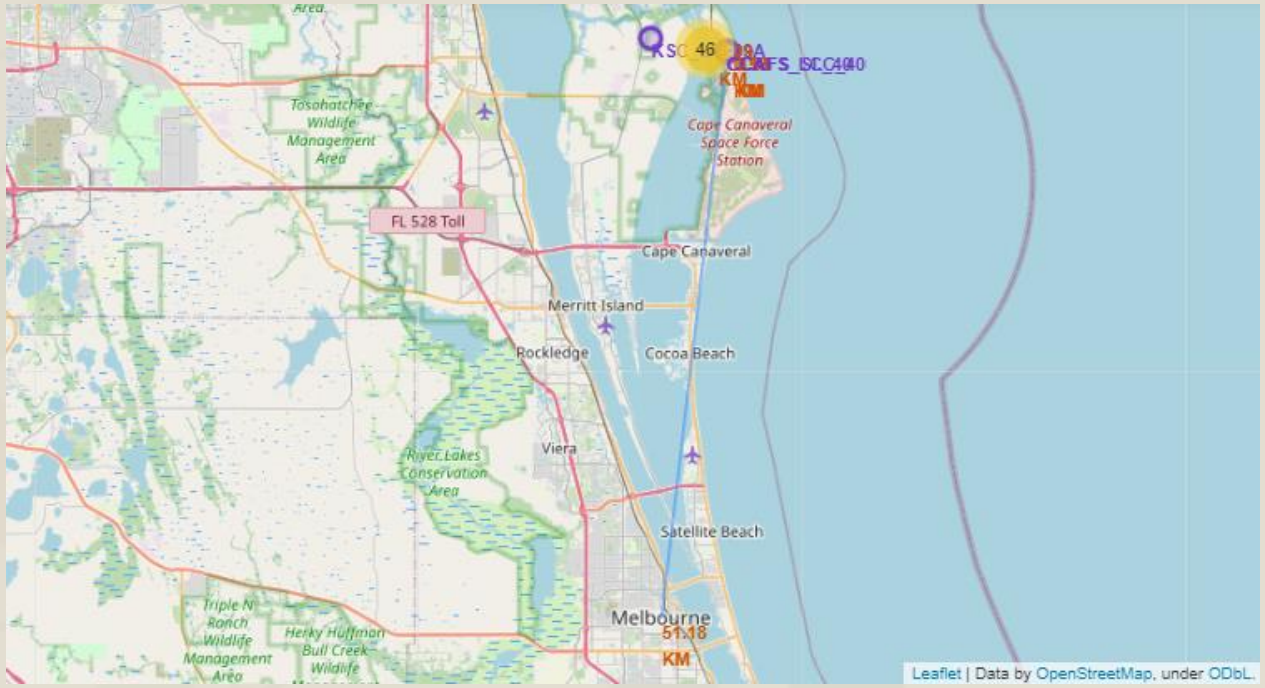
Location of Launch Sites



- From image above, we can see that all launch sites are not far from Equator Line. So, the distance from earth to space is actually the closest as they can be. That makes the fuel needed is fewer.
- They also are close to the coast. All the launch sites in close to the coast is a good choice for safety reasons, it's far from where a lot of people live.

The Closest to Railways, Highways, Coastline, and City






```
launch sites to railways : 0.6 km
launch sites to highways : 1.3 km
launch sites to coastline : 0.85 km
launch sites to the closest city is 51.2 km.
```

DASHBOARD

Link for interactive dashboard :

https://github.com/ChristoNindyo/SpaceX-Interactive-Dashboard/blob/main/Build%20an%20Interactive%20Dashboard%20with%20Ploty%20Dash.ipynb


jupyter Build an Interactive Dashboard with Ploty Dash Last Checkpoint: 6 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Payload range (Kg):
Success count on Payload mass for site KSC LC-39A
class
1
0.8
0.6
0.4
0.2
0
3000
4000
5000
6000
7000
Payload Mass (kg)
Booster Version Category
FT
B4
B5


jupyter Build an Interactive Dashboard with Ploty Dash Last Checkpoint: 4 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
SpaceX Launch Records Dashboard
All Sites
Success Count for all launch sites
29.2%
41.7%
16.7%
12.5%
KSC LC-39A
CCAFS LC-40
VAFB SLC-4E
CCAFS SLC-40

Based on the dashboard

If we want the highest rate of success , this is what we need :

- Launch Site is KSC LC-39A
- The payload range 0 - 5000 kg
- Use F9 Booster version FT

MACHINE LEARNING

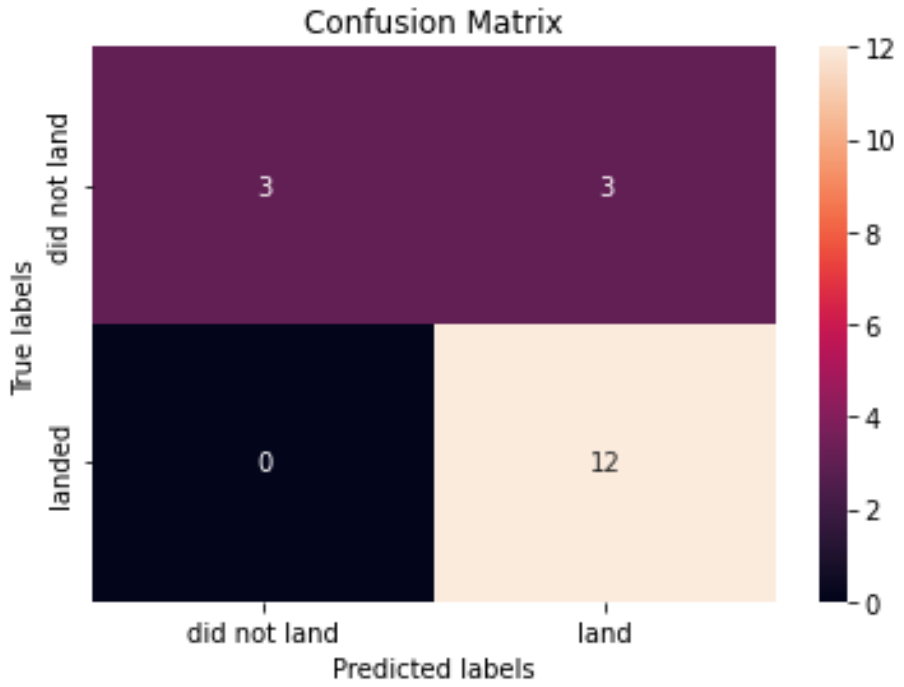
Calculate the accuracy of tree_cv on the test data using the method `score` :

```
yhat_tree = tree_cv.predict(X_test)
```

```
print('Test set Accuracy:', metrics.accuracy_score(Y_test, yhat_tree))
```

```
Test set Accuracy: 0.8333333333333334
```

```
plot_confusion_matrix(Y_test,yhat)
```



The Best Model we got

Decision Tree

Accuracy : 83,34 %

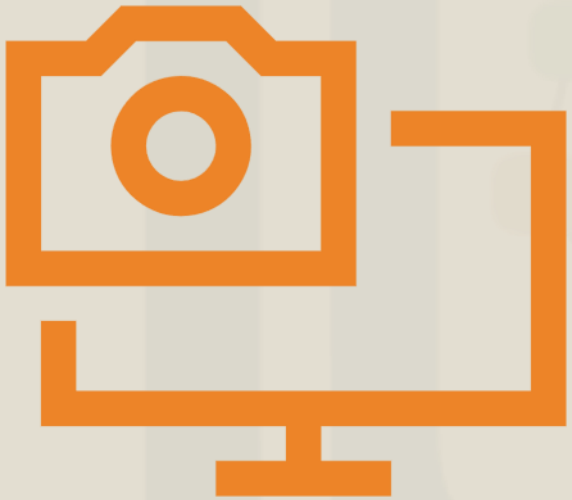
DISCUSSION

- Do Data Science actually help in this problem ?
- Is accuracy of our model is good enough ? If not, how can we improve ?
- How business environment we should make to get the best result ?

CONCLUSION

- In this particular problem, Data Science help to deconstruct the main problem to sub problem. So, we can analyze easier and faster
- We can solve this problem without complex rocket science but with data analysis & machine learning
- We get what parameter that make the launch get most success
- We get model to deploy to determine launch which success or not

APPENDIX



- Coursera Data Science Program by IBM
- Data Science for Business book by Provost and Fawcett
- Business Analytics book by Camm, Cochran, Fry, Ohlmann, Anderson, Sweeney, and Williams

THANK YOU