Research On

SpaceX

HARSH RAGHUVANSHI

DATE :- 21/02/2023

Executive Summary

- In this report we are discuss about SpaceX. SpaceX is a company which sending spacecraft to International Space Station. SpaceX rocket launches are relatively inexpensive then other companies.

- SpaceX costing of falcon9 rocket launch is 62 million dollars , and other providers cost upwards of 165 million dollars. This is very big difference between SpaceX and other providers.

- In this report we see the research on SpaceX, how they can be inexpensive then other provider. Which method they are using to reducing the cost of rocket launch.

Table of Contents

Introduction

Space Exploration Technologies Corp. (SpaceX) is an American spacecraft , launchers and satellite communications corporation headquartered in Hawthorne, California. It was founded in 2002 by Elon Musk with stated goal of reducing space transportation costs. The company manufactures the Falcon 9, Falcon Heavy, and Starship launch vehicles, several rocket engines, Cargo Dragon and Crew Dragon spacecraft, and Starlink communications satellites.

Falcon 9 is orbital launch vehicle made by SpaceX. Falcon 9 is the reusable heavier lift vehicle. SpaceX's Falcon 9 launch like regular rockets. Falcon 9 between in two stages, second stage helps bring payload(payload is enclosed in the fairings) to orbit but most of the work is done by the first stage. First stage does most of the work and is much larger than the second stage. First stage is quite large and expensive. Unlike other rocket providers, SpaceX's Falcon 9 can recover the first stage. Sometimes the first stage does not land. Sometimes it will crash. Other times, SpaceX

will sacrifice the first stage due to mission parameters like payload, orbit and customer. By this feature SpaceX charges is cheaper than other providers (approximately 1/3 of their cost) because they reuse first stage which price is 60%-70% of the rocket.

SpaceX have 4 launching Cape Canaveral Space Launch Complex 40**(CCAFS SLC 40)**, Vandenberg Air Force Base Space Launch Complex 4E **(VAFB SLC-4E)**, Kennedy Space Center Launch Complex 39A **(KSC LC 39A).**

Methodology

In this research, I am using several methods to find the result. Thes methods are :-

- **Collection of Data -**

We use API to extract the information of SpaceX. We requesting (by using Requests method) rocket launch data from SpaceX API with URL. Then we use **.json()** method to decode the response content as a json and **.json_normalize()** method to turn it into a pandas dataframe. In this dataframe we have two three type of rockets but we want only Falcon 9 data so we create a new dataframe only with falcon 9 data. When we check the data we found that there are some null values in PayLoad column so we have to give some value in it so we calculate total mean of PayLoad column and replace null values with it. Then I extract Falcon 9 launch records HTML table from Wikipedia by web scrap with **BeautifulSoup** then parse the table and convert into a pandas dataframe.

❖ **Data Wrangling -**

In this method, we perform exploratory Data Analysis and determine Training labels. We load SpaceX dataset and then check the type of data with respect to columns. Then we calculate number of launches on each sites. After this we calculate the number and occurrence of each orbit. Then we calculate the number and occurrence of mission outcome per orbit type. Finally we create a Landing Outcome label from Outcome column.

❖ **EDA with SQL -**

In this method, we perform exploratory Data Analysis with SQL. We use SQL queries for analysis like knowing unique launch sites , average payload mass carried by Falcon 9 and so on.

❖ **EDA with Visualization Lab -**

In this method, we perform exploratory Data Analysis with Visualization Libraries. We visualize the relationship between Flight Number and Launch Site , we visualize the launch success yearly trend and so on.

❖ **Launch Sites Location Analysis With Folium -**

In this section, we mark all sites on a map. We mark the success and failure launches for each sites on the map and calculate the distance between a launch site to its proximities.

❖ **Build a Dashboard Application With Plotly Dash -**

In this section, we create a dashboard where we input all the data related to launch sites, payload, and so on.

Results

- We count the number of launches on each site :

```
[12]: # Apply value_counts() on column LaunchSite
      df.iloc[:,5].value_counts()

[12]: CCAFS SLC 40    55
      KSC LC 39A      22
      VAFB SLC 4E     13
      Name: LaunchSite, dtype: int64
```

- We calculate the number and occurrence of each orbit :

```
[14]: # Apply value_counts on Orbit column
      df.iloc[:,4].value_counts()

[14]: GTO      27
      ISS      21
      VLEO     14
      PO        9
      LEO       7
      SSO       5
      MEO       3
      ES-L1     1
      HEO       1
      SO        1
      GEO       1
      Name: Orbit, dtype: int64
```

- The names of the unique launch sites in the space mission :

```
[8]: %sql SELECT DISTINCT Launch_Site FROM SPACEXTBL
```

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

[8]:

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

- Average payload mass carried by F9 v1.1 :

```
[11]: %sql SELECT AVG(PAYLOAD_MASS__KG_) AS payload_mass FROM SPACEXTBL WHERE Booster_Version = 'F9 v1.1'
```

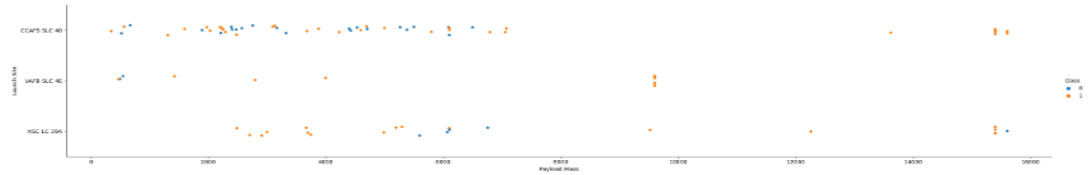
```
 * sqlite:///my_data1.db
Done.
```

[11]:

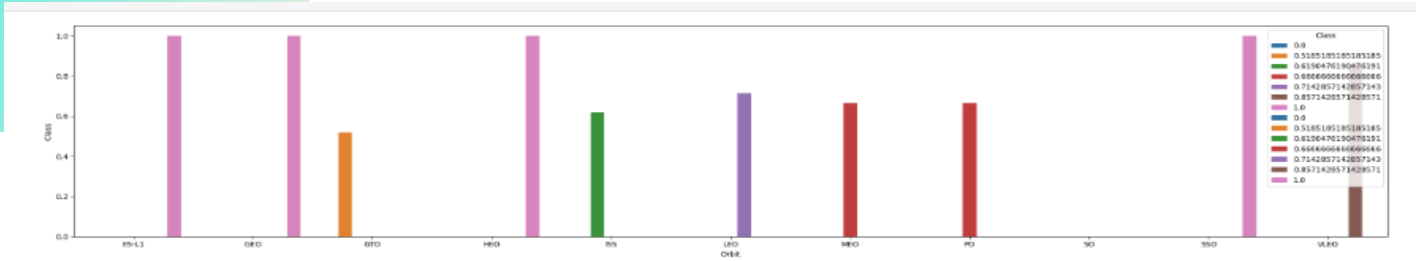
payload_mass
2928.4

- Relationship between payload and launch site :

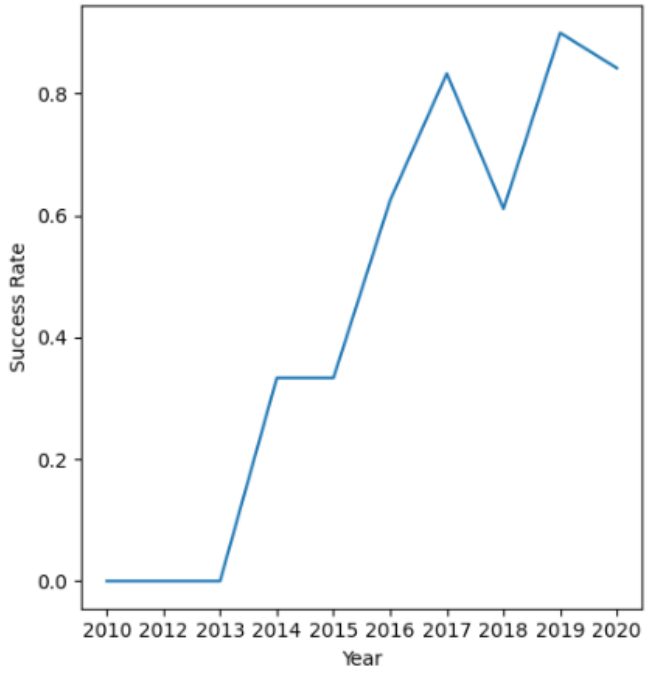
```
# Plot a scatter point chart with x axis to be Pay Load Mass (kg) and y axis to be the launch site, and hue
sns.catplot(y="LaunchSite",x="PayloadMass", hue="Class", data=df, aspect=5)
plt.xlabel("Payload Mass")
plt.ylabel("Launch Site")
plt.show()
```



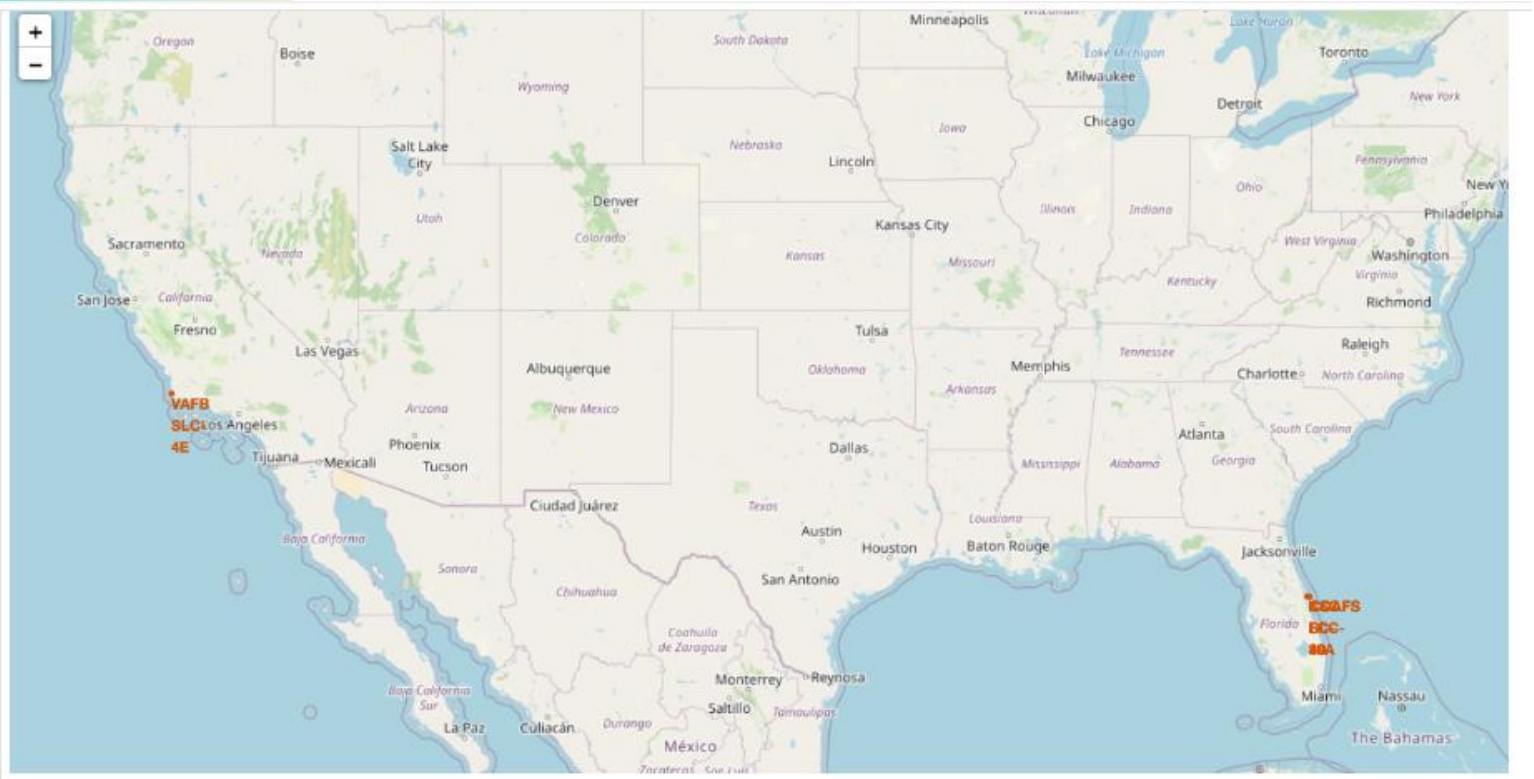
- Relationship between Flight number and Orbit type :



- The launch success yearly trend :



- Mark all site on map :

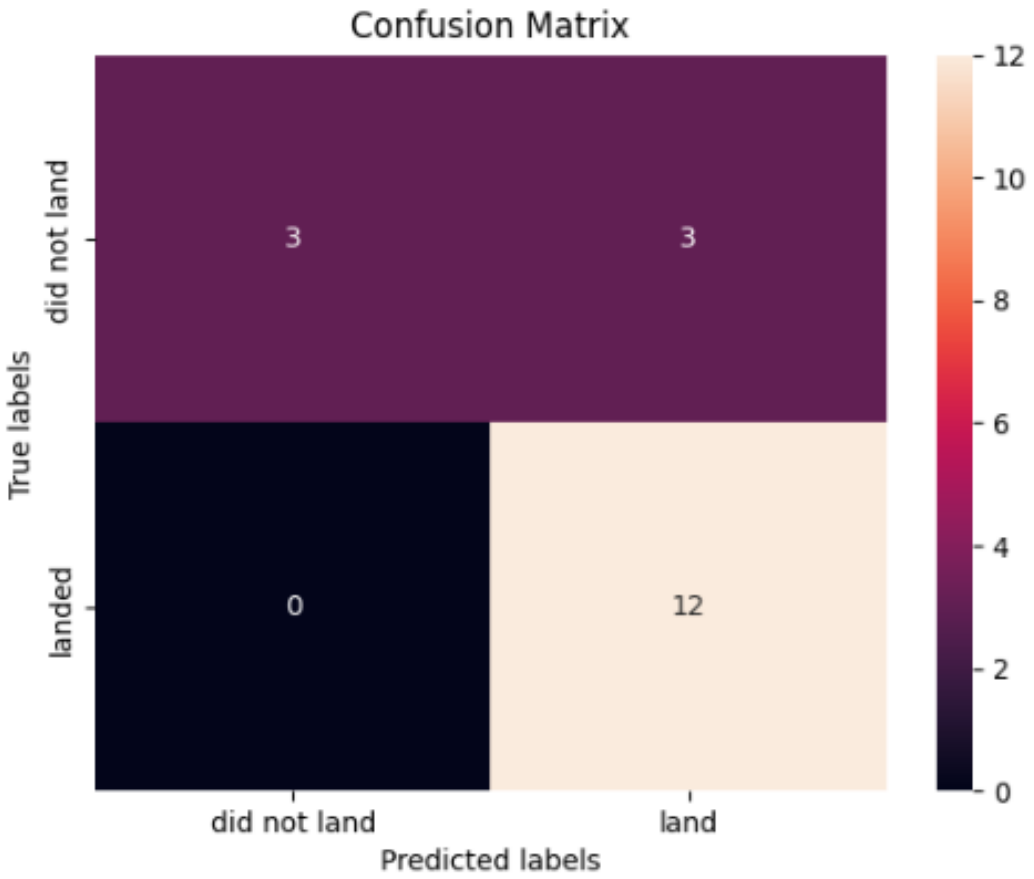


- We split the data in two parts for training and testing where 80% of data is use for training and 20% of data is use for testing, train it using following methods:

- Logistic Regression :

```
tuned hpyerparameters :(best parameters)  {'C': 0.01, 'penalty': 'l2', 'solver': 'lbfgs'}
accuracy : 0.8464285714285713
```

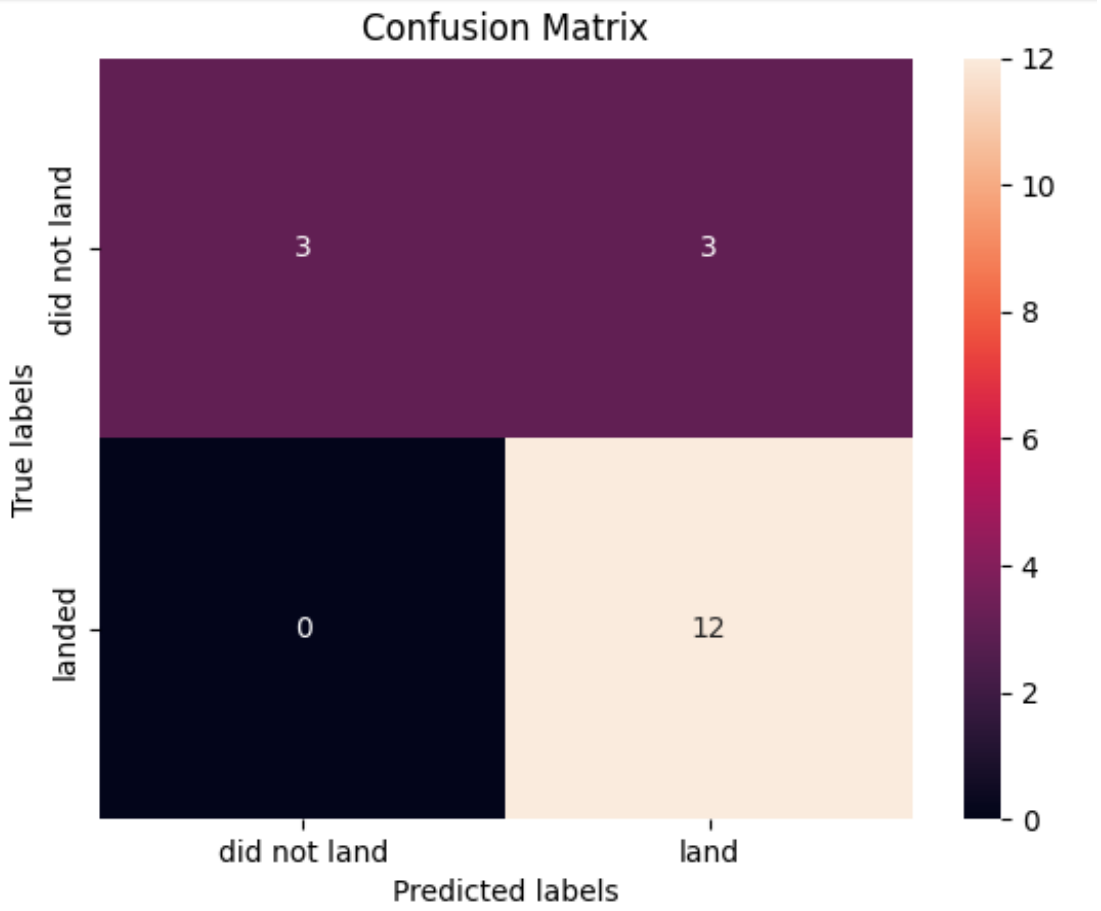
Confusion matrix:



- Support Vector Machine(SVM) :

```
tuned hpyerparameters :(best parameters)  {'C': 1.0, 'gamma': 0.03162277660168379, 'kernel': 'sigmoid'}
accuracy : 0.8482142857142856
```

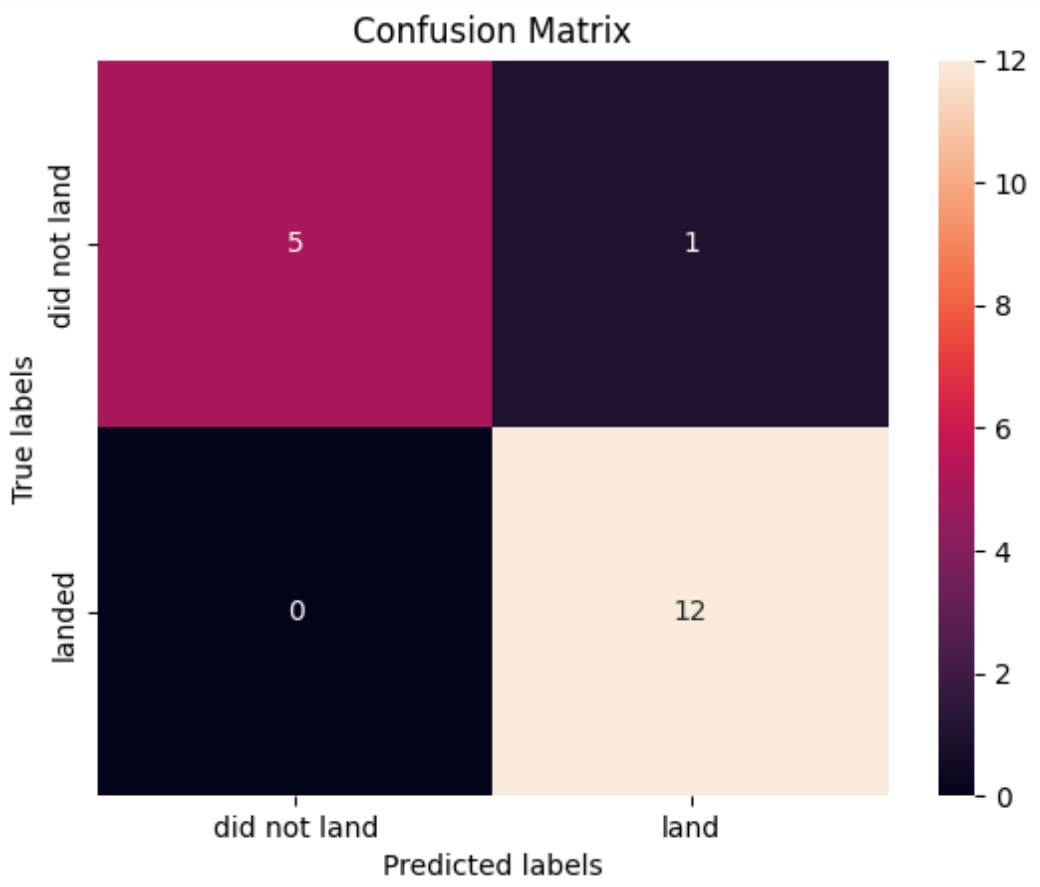
Confusion Matrix :



- Decision Tree Classifier :

```
tuned hpyerparameters :(best parameters)  {'criterion': 'gini', 'max_depth': 12, 'max_features': 'auto', 'min_samples
_leaf': 1, 'min_samples_split': 5, 'splitter': 'best'}
accuracy : 0.8767857142857143
```

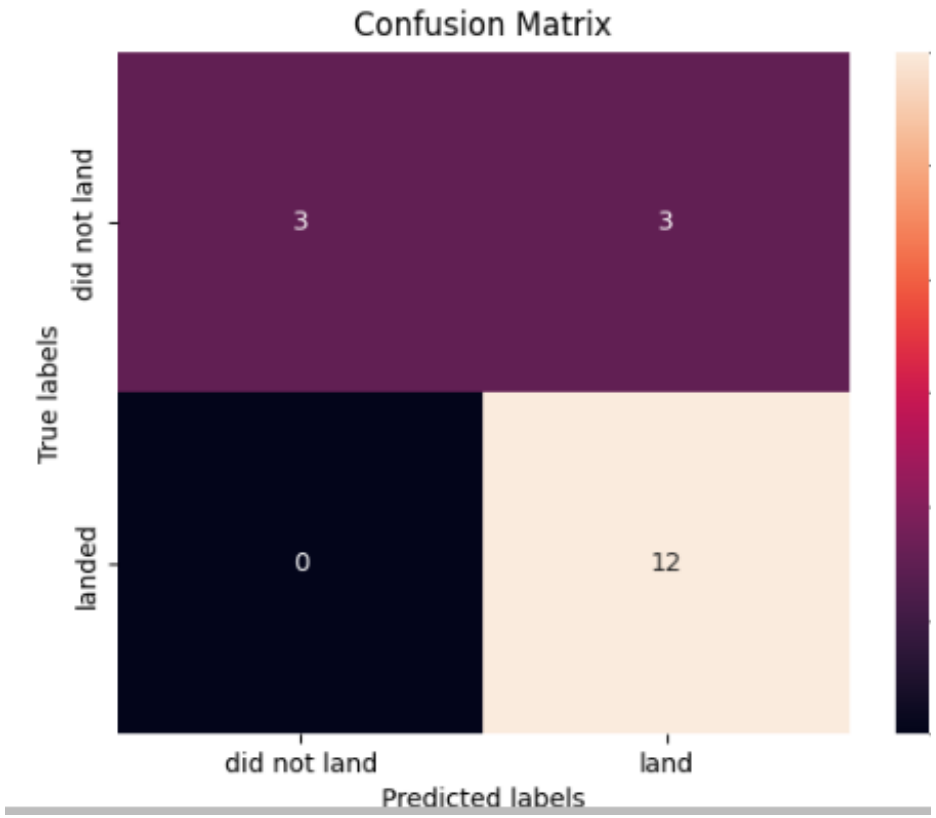
Confusion Matrix :



- K Nearest Neighbors :

```
tuned hpyerparameters :(best parameters)  {'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}
accuracy : 0.8482142857142858
```

Confusion Matrix :



Discussion

In this project, we understand that SpaceX main feature is reusability of stage 1. The success rate of landing of stage 1 is increase year by year. SpaceX have 4 launch sites and the most used site is CCAFS LC-40. They launch most rockets for GTO orbit. In this project we are focused on Booster Version Falcon 9.

Conclusion

- The Decision Tree Classifier model achieve highest accuracy rate 87.6% , Logistic Regression, SVM and KNN model achieve accuracy rate 84.8%.
- The success rates for SpaceX launches is increasing every year, they will eventually perfect the launches.
- KSC LA 39A had the most successful launches from all the sites.
- Orbit GEO, HEO, SSO,ES L1 has the best success rate.
- Performance is inversely proportional to Payload Weight.

Appendix

- IBM DEVELOPER SKILLS NETWORK
- COURSERA
- Wikipedia

Thank you!