Spacex First Stage Success Lunching

Arezoo

10/17/2023

IBM Developer

SKILLS NETWORK

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
- Conclusion

SKILLS NETWORK

EXECUTIVE SUMMARY

- To estimate the cost of each SpaceX launching, we can determine if the first stage will land successfully. If so, SpaceX can reuse the first stage and deduct the cost.
- Falcon 9 launch data from the SpaceX REST API, and web scraping related Wiki pages including information about the rocket used, payload delivered, launch specifications, landing specifications, and landing outcome were analyzed

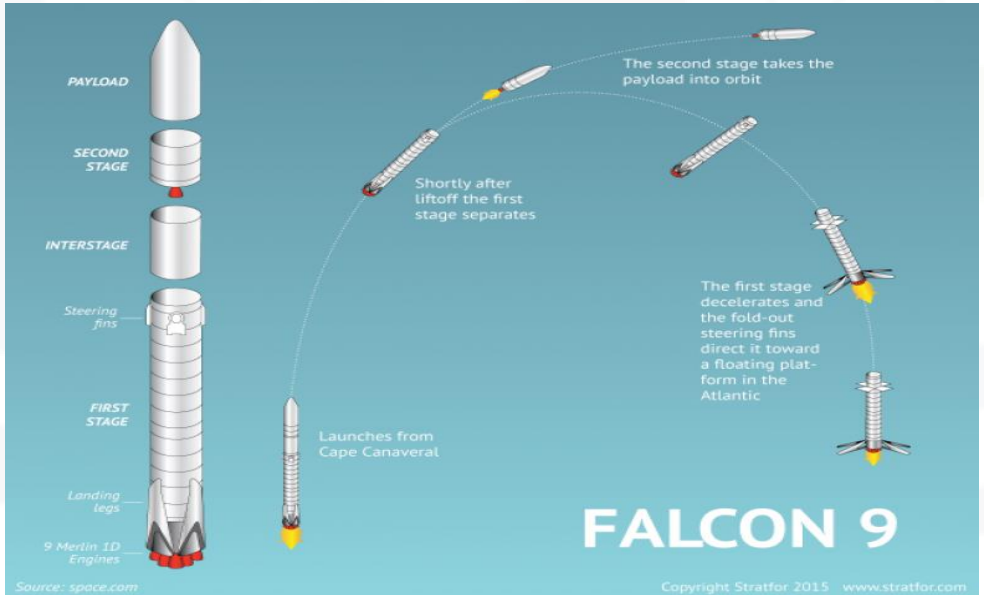
❖ The purpose of this study is to use this data and apply machine learning algorithms to predict whether SpaceX will attempt to land a rocket or not. The data were analyzed by Python

❖ The findings reveal that applying classification methods such as logistic regression, Support vector machine, decision tree, and k nearest neighbor, can classify Falcon 9 launch accurately (0.83).

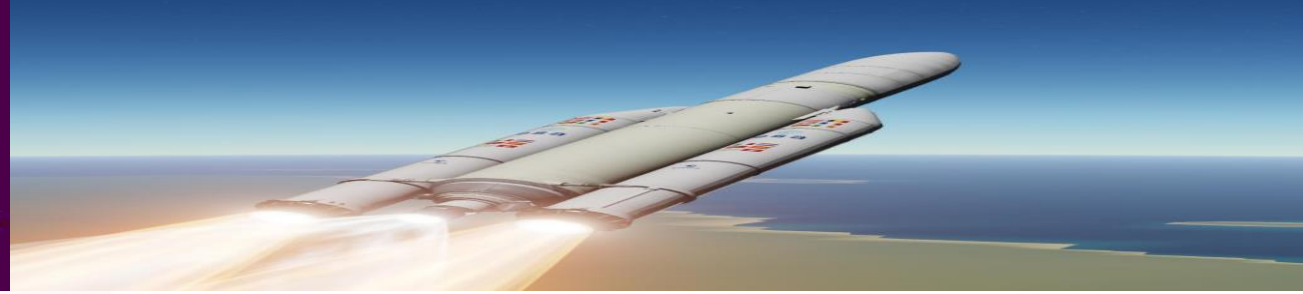
❖ According to the results of the study, There is a very high probability that the next launch of Falcon 9 will be successful.

INTRODUCTION

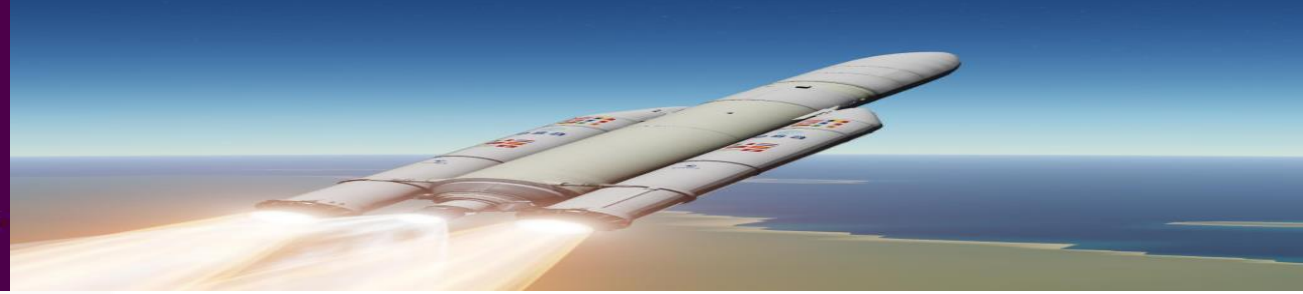
- Blue Origin manufactures sub-orbital and orbital reusable rockets. Perhaps the most successful is SpaceX. It can reuse the first stage and deduct the cost of launching. By predicting the rate of its success landing, we can predict the cost as well.
- The purpose of this study is to use Falcon 9 SpaceX launch to predict whether SpaceX will attempt to land a rocket or not. There are a significant amount of research devoting to this issue but none of them used machine learning algorithms to predict the success rate of the first stage landing of Falcon 9.



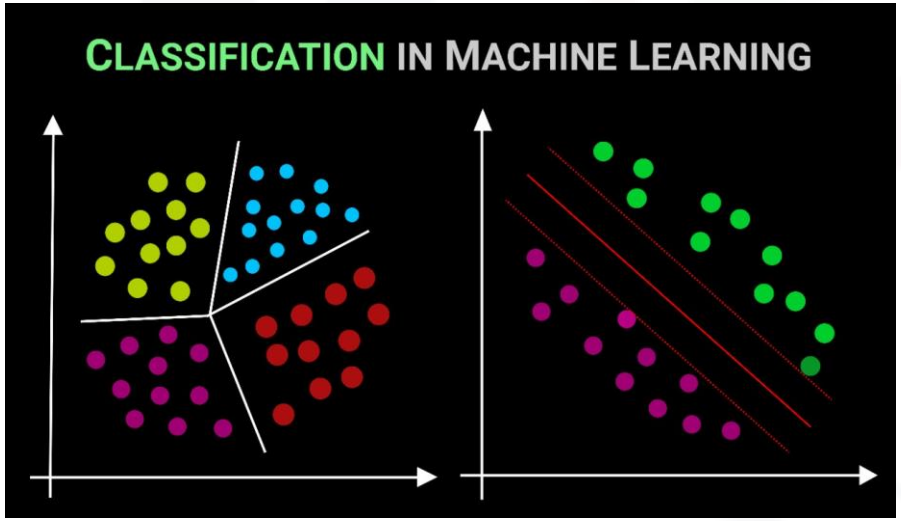
Methodology



- In this study, SpaceX launch data is gathered from an API, specifically the SpaceX REST API are used. This API will give us data about launches, including information about the rocket used, payload delivered, launch specifications, landing specifications, and landing outcome.
- Get request using the requests library in Python can obtain the launch data, which we will use to get the data from the API in Json form. we can use the json_normalize function to “normalize” the structured json data into a flat table.
- Another popular data source for obtaining Falcon 9 Launch data is web scraping-related Wiki pages. We will use the Python BeautifulSoup package to web scrape some HTML tables that contain valuable Falcon 9 launch records. Then we need to parse the data from the tables and convert them into a Pandas data frame for further visualization and analysis.

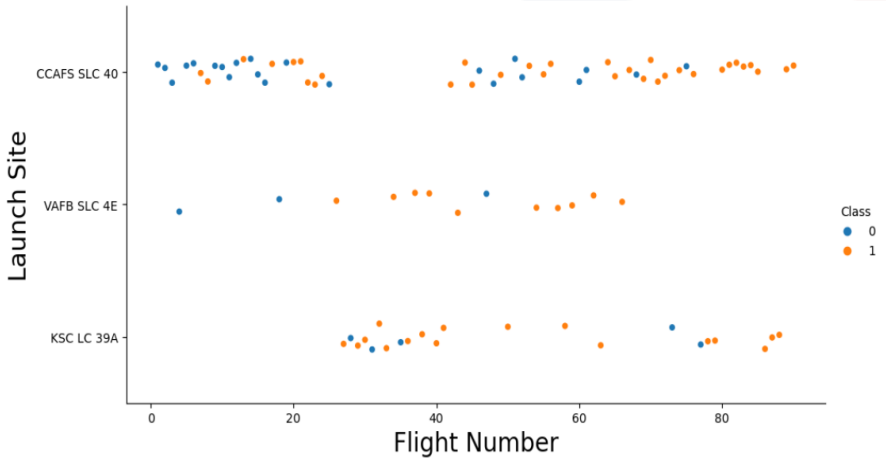


- We use some of the preprocessing methods such as missing value imputation and data wrangling to make data ready for analyzing.
- we use some of exploratory data analysis for the first steps of analyzing such as using different kinds of graphs for data visualization and SQL
- Comparing classification algorithms as machine learning methods to predict Falcon 9 success launching.



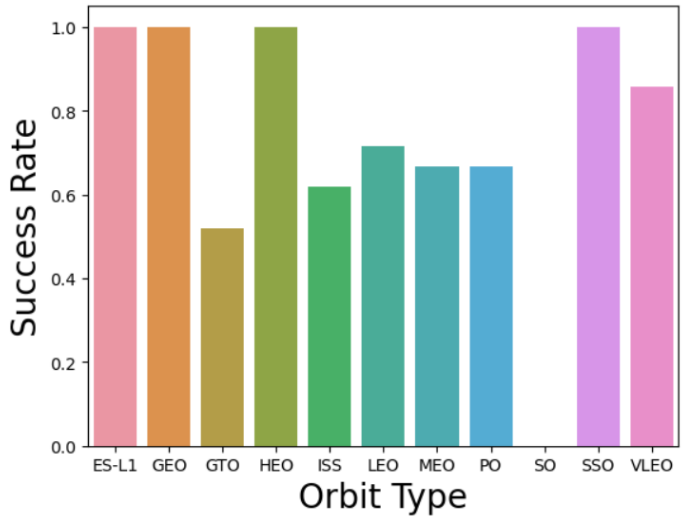
RESULTS: Visualization

Scatter plot of flight number vs launch site



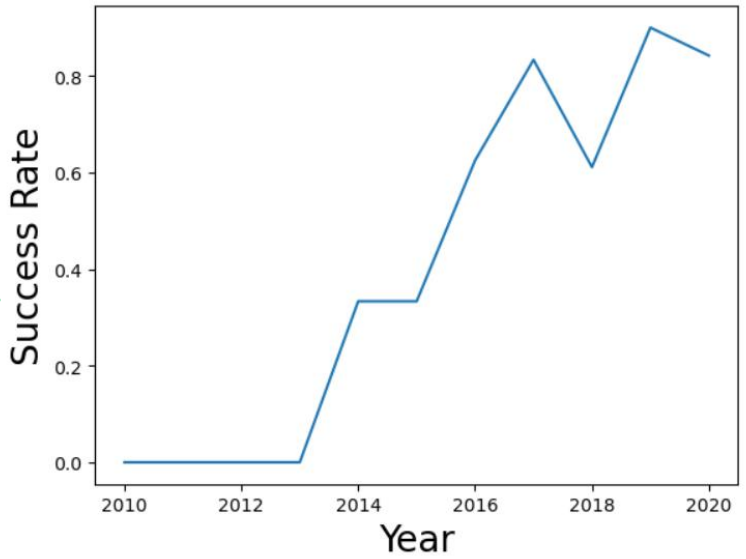
The number of failed flights in the launch site of CCAFS 40 is more.

Bar chart of orbit type vs success rate



The full success rate is for orbit types of ES-L1, GEO, and HEO.

Line chart of yearly success rate



Since 2013, the yearly success rate increased.

RESULTS: SQL

Names of the unique launch sites in the space mission

[5]:

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

5 main unique launch sites exist.

5 records where launch sites begin with the string 'CCA'

Done.

[7]:

DATE	time_utc_	booster_version	launch_site	payload	payload_mass_kg_	orbit	customer	mission_
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	None	0	LEO	SpaceX	
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	None	0	LEO (ISS)	NASA (COTS) NRO	
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	None	525	LEO (ISS)	NASA (COTS)	
2012-10-08	00:35:00	F9 v1.0 B0006	CCAFS LC-40	None	500	LEO (ISS)	NASA (CRS)	

The information on the 5 launches where the launch site begins with ‘CCA’.

The total payload mass carried by boosters launched by NASA (CRS)

[9]:

1
107010

RESULTS: SQL

Average payload mass carried by booster version F9 v1.1

[11]:

avg_payload
2928

Date when the first successful landing outcome in the ground pad was achieved

[13]:

first_success_gp
2015-12-22

List the names of the boosters which have success in drone ships and have payload mass greater than 4000 but less than 6000

[14]:

booster_version
F9 FT B1021.2
F9 FT B1031.2
F9 FT B1022
F9 FT B1026

List the total number of successful and failed mission outcomes

[15]:

mission_outcome	qty
Failure (in flight)	1
Success	99
Success (payload status unclear)	1

RESULTS: SQL

List the names of the booster versions which have carried the maximum payload mass. Use a subquery

[23]:

booster_version
F9 B5 B1048.4
F9 B5 B1048.5
F9 B5 B1049.4
F9 B5 B1049.5
F9 B5 B1049.7
F9 B5 B1051.3
F9 B5 B1051.4
F9 B5 B1051.6
F9 B5 B1056.4
F9 B5 B1058.3
F9 B5 B1060.2
F9 B5 B1060.3

List the failed landing outcomes in drone ships, their booster versions, and launch site names for in year 2015

[24]:

booster_version	launch_site
F9 v1.1 B1012	CCAFS LC-40
F9 v1.1 B1015	CCAFS LC-40

Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the dates 2010-06-04 and 2017-03-20, in descending order

[25]:

landing__outcome	qty
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

RESULTS: Folium

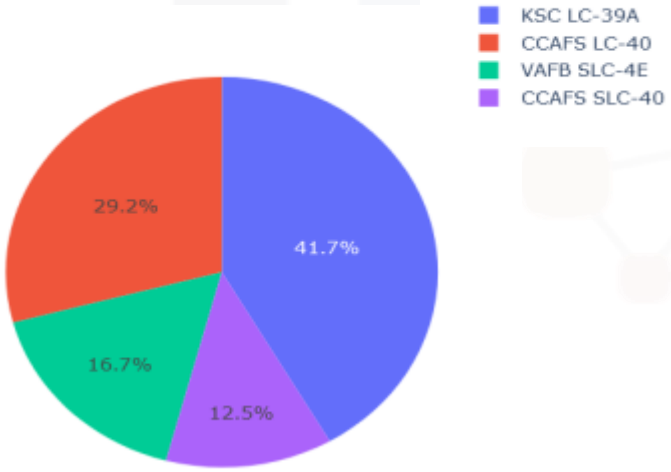
East Coast: 2

West Coast: 1

USA: 3

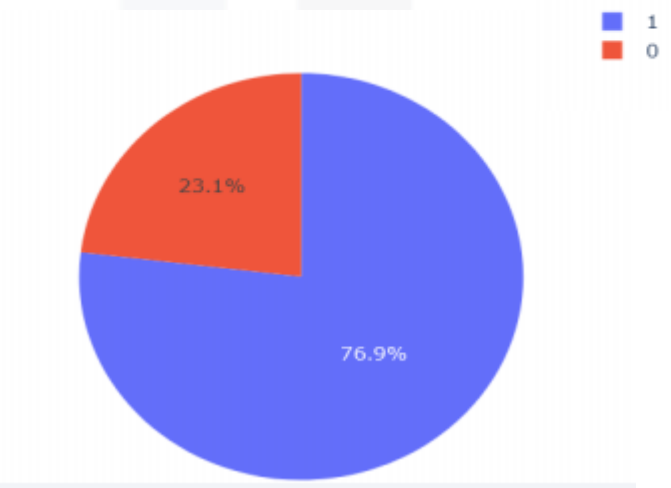
RESULTS: Dashboard

Pie chart of total success launch by site



KSC LC-39A site had the most number of successful launch

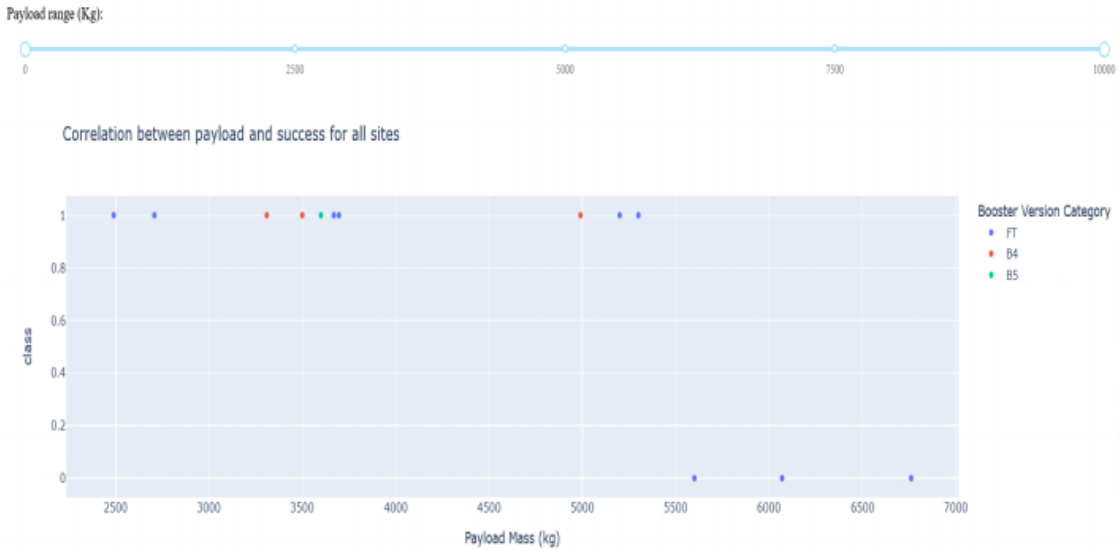
Pie chart of The success rate of the most successful site



The success rate of this site is 76.9 %.

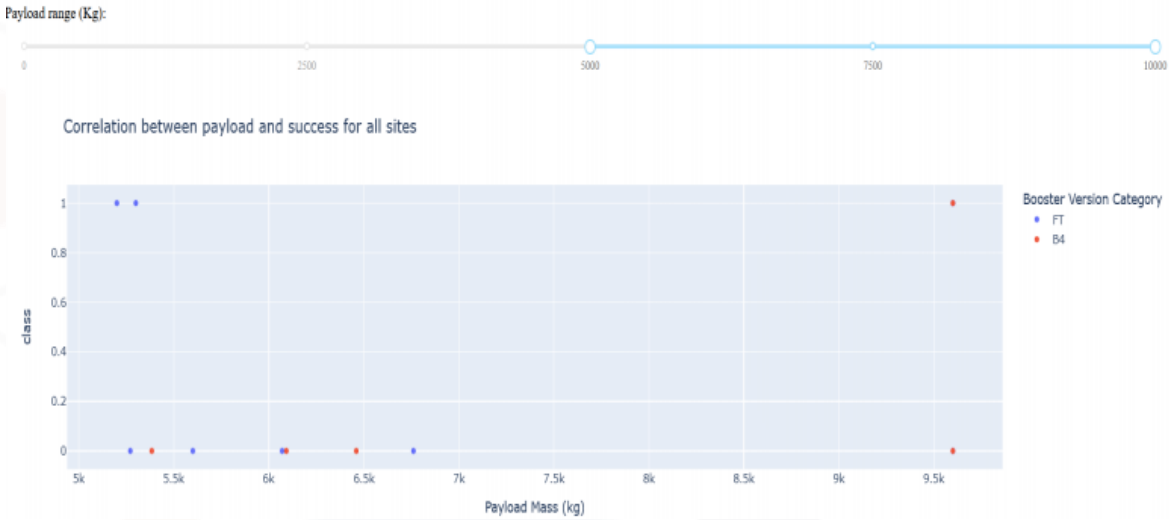
RESULTS: Dashboard

Correlation between success rate and the payload mass



The chance of launch failure for FT booster version was higher for payload mass over 5500 kg

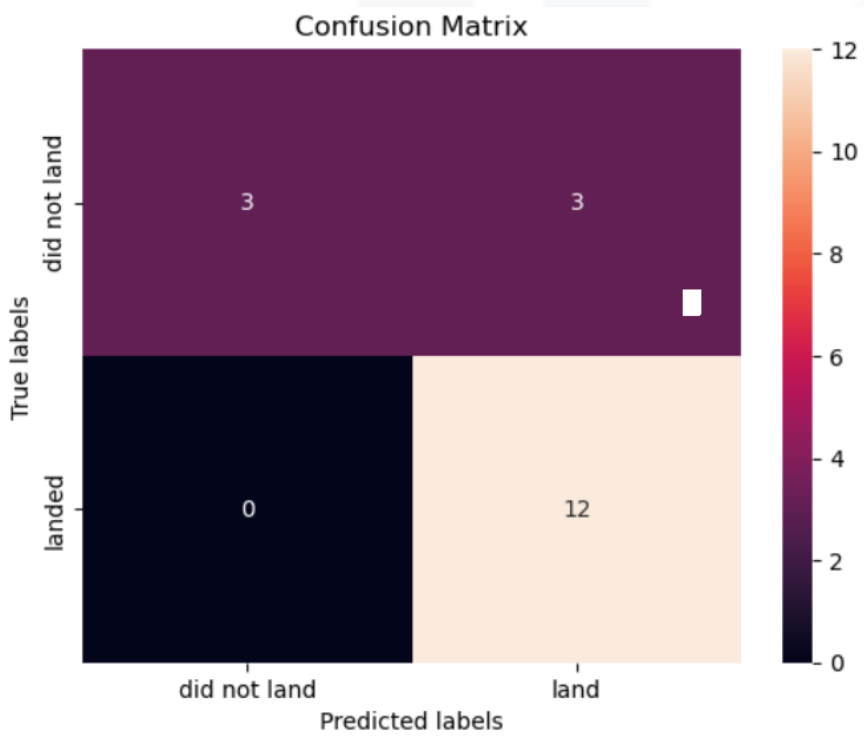
Correlation between success rate and the payload mass over 5000 kg



It is hard to conclude which booster version had more success rate.

RESULTS: Machine learning algorithms

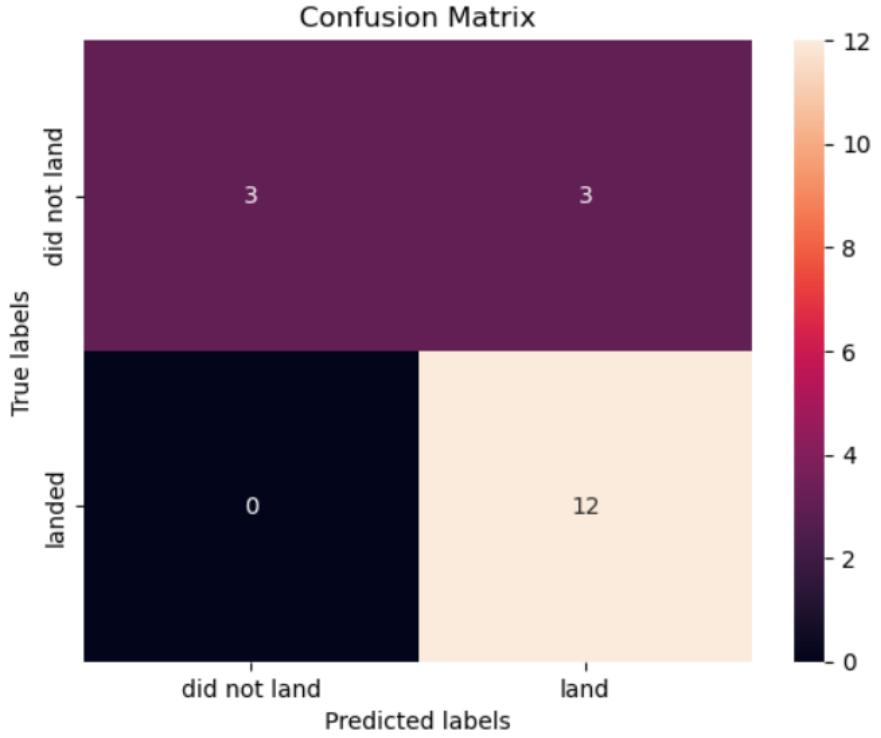
Logistic regression



Accuracy on test data:0.833

Accuracy:0.846

Support Vector Machine



Accuracy on test data:0.833

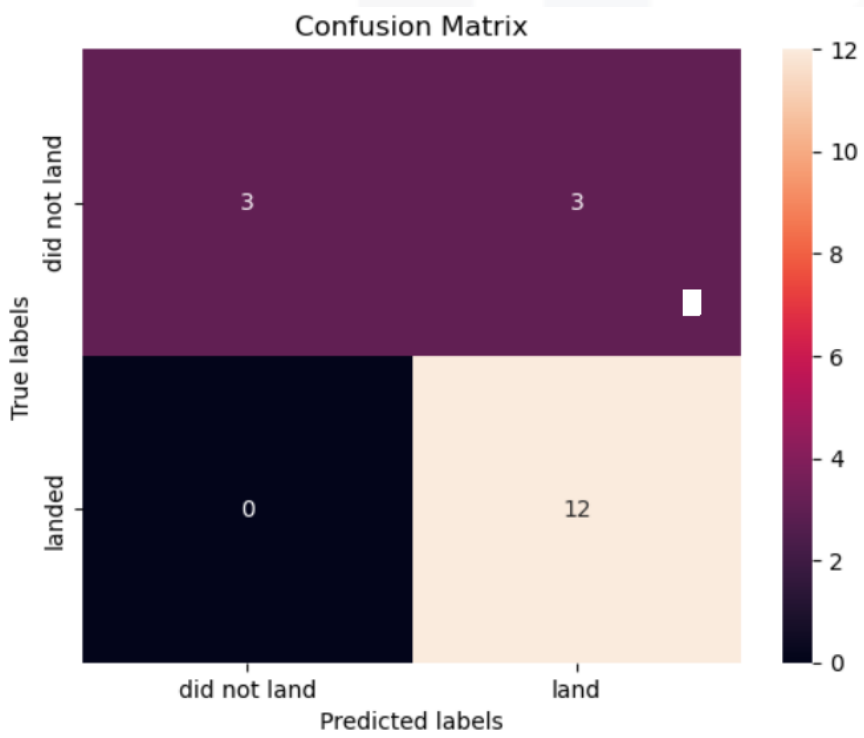
Accuracy:0.848

IBM Developer

SKILLS NETWORK

RESULTS: Machine learning algorithms

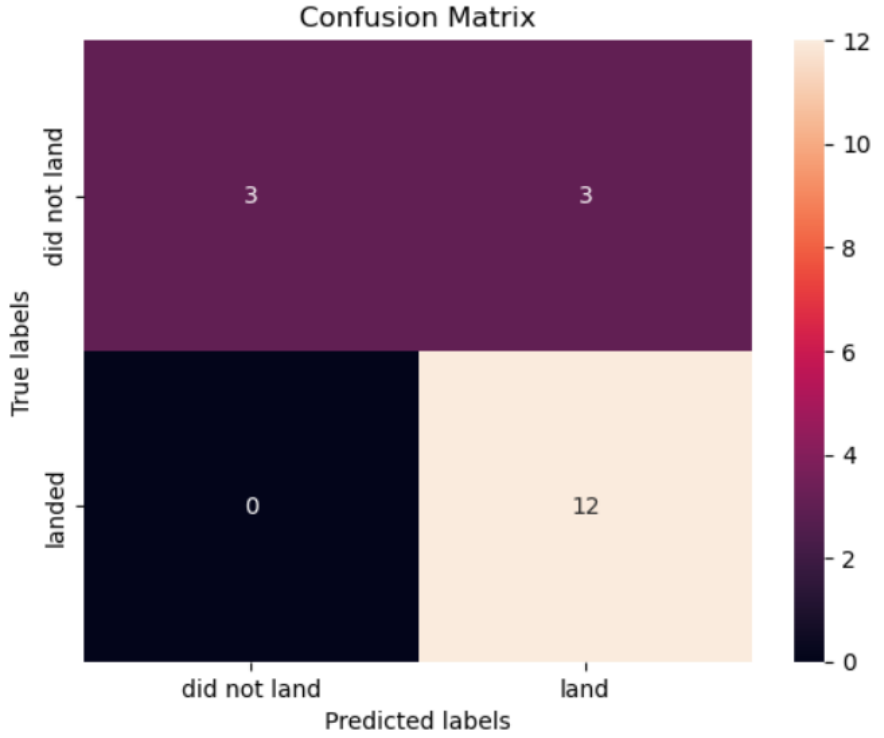
Decision Tree



Accuracy on test data:0.833

Accuracy:0.875

K Nearest Neighbors



Accuracy on test data:0.833

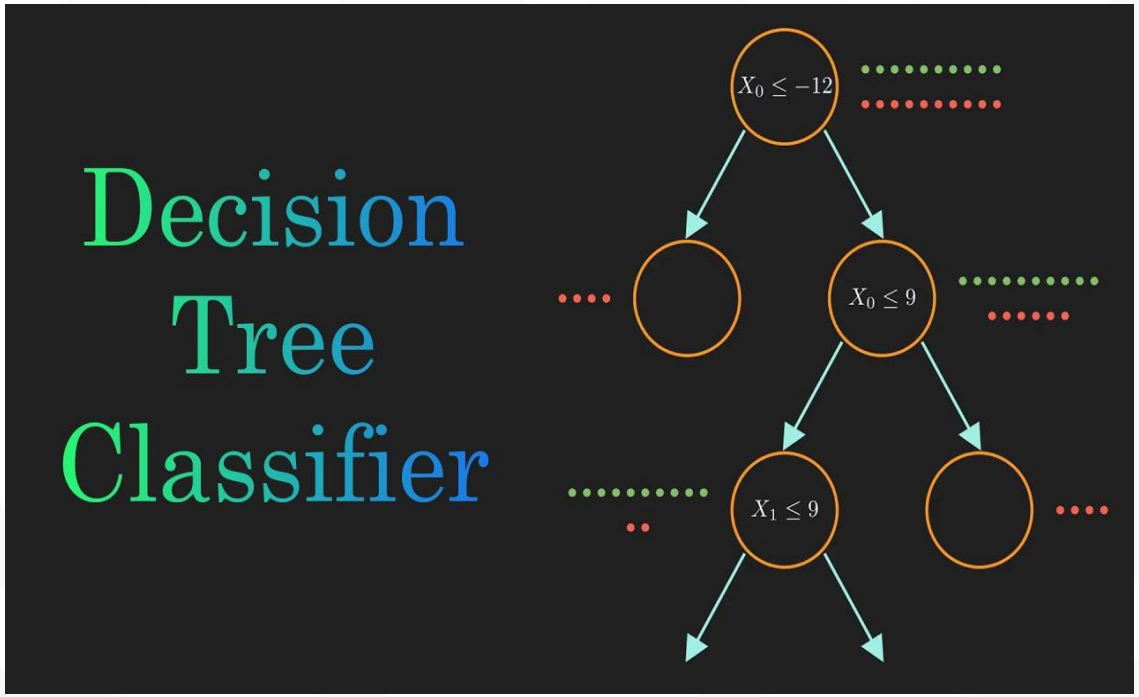
Accuracy:0.848

IBM Developer

SKILLS NETWORK

Discussions

The best model according to the accuracy of training data is the Decision Tree (0.875), it is vital to mention that the accuracy of all methods on test data was similar (0.833)



Conclusions

- This study aimed to use Falcon 9 launch data from the SpaceX REST API, and web scraping related Wiki pages including information about the rocket used, payload delivered, launch specifications, landing specifications, and landing outcome and apply machine learning algorithms to predict whether SpaceX will attempt to land a rocket.

- Since the dependent variable in this study was binary (Class=0, failure, Class=1, success), different classification methods such as logistic regression, support vector machine, decision tree, and k nearest neighbor were compared.

- According to the accuracy results, the decision tree is considered as the best algorithm.

References

- https://www.coursera.org

- https://www.bing.com/images/search?view=detailV2&ccid=CbGbYzD%2b&id=B9AFB1320FA77A165C449D646866FEA482956031&thid=OIP.CbGbYzD-Fth2S_kRP0Tt8wHaEK&mediaurl=https%3a%2f%2fi.ytimg.com%2fvi%2fZVR2Way4nwQ%2fmaxresdefault.jpg&cdnurl=https%3a%2f%2fth.bing.com%2fth%2fid%2fR.09b19b6330fe16d8764bf9113f44edf3%3frik%3dMWCVgqT%252bZmhknQ%26pid%3dImgRaw%26r%3d0&exph=720&expw=1280&q=desicion+tree&simid=608032249941085952&FORM=IRPRST&ck=CAA10AD8EA286D8DB475739ED4978FF4&selectedIndex=43&ajaxhist=0&ajaxserp=0

- https://www.bing.com/images/search?view=detailV2&ccid=o%2byf9KYw&id=C1280FD4565D965FEC952E8D65D1B137D0561687&thid=OIP.o-yf9KYwFw6Eo3nr4kkvvwHaHa&mediaurl=https%3a%2f%2fi.pinimg.com%2f736x%2f81%2f0a%2f07%2f810a074a58e75c06388c44ad5259c2de--acquisition-data.jpg&cdnurl=https%3a%2f%2fth.bing.com%2fth%2fid%2fR.a3ec9ff4a630170e84a379ebe2492fbf%3frik%3dhxZW0Dex0WWNLg%26pid%3dImgRaw%26r%3d0&exph=720&expw=720&q=data+gathering&simid=608053621780208477&FORM=IRPRST&ck=C1D8C3BD07EB21D437321DAE929799CF&selectedIndex=1&ajaxhist=0&ajaxserp=0


Thank you