SPACEX LAUCH DATA ANALYSIS CAPSTONE PROJECT

DATA COLLECTION

- Data sets were collected from Space X API (https://api.spacexdata.com/v4/rockets/) and from Wikipedia (https://en.wikipedia.org/wiki/List_of_Falcon/_9/_and_Falcon_Heavy_launches), using web scraping technics.

EDA with Visualization

- Initially some Exploratory Data Analysis (EDA) was performed on the dataset.
- Then the summaries launches per site, occurrences of each orbit and occurrences of mission outcome per orbit type were calculated.

Finally, the landing outcome label was created from Outcome column.

EDA with SQL

The following SQL queries were performed:

- Names of the unique launch sites in the space mission;
- Top 5 launch sites whose name begin with the string 'CCA’;
- Total payload mass carried by boosters launched by NASA (CRS);
- Average payload mass carried by booster version F9 v1.1;
- Date when the first successful landing outcome in ground pad was achieved;
- Names of the boosters which have success in drone ship and have payload mass between 4000 and 6000 kg;

DATA VISUALIZATION(Folium)

Markers, circles, lines and marker clusters were used with Folium Maps:

- Markers indicate points like launch sites;
- Circles indicate highlighted areas around specific coordinates, like NASA Johnson Space Center;
- Marker clusters indicates groups of events in each coordinate, like launches in a launch site;

DATA VISUALIZATION(Plotly Dash)

The following graphs and plots were used to visualize data:

- Percentage of launches by site
- Payload range

This combination allowed to quickly analyze the relation between payloads and launch sites, helping to identify where is best place to launch according to payloads.

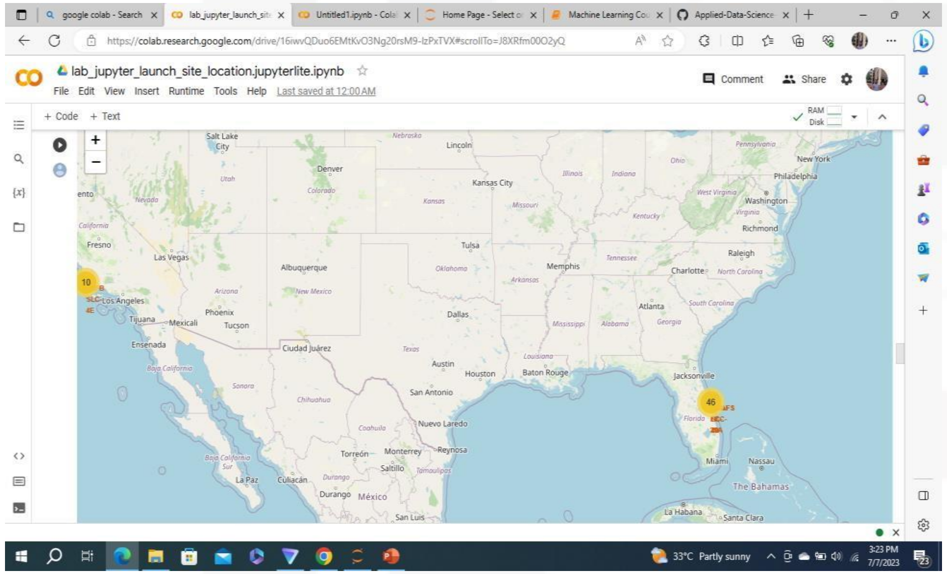
RESULTS

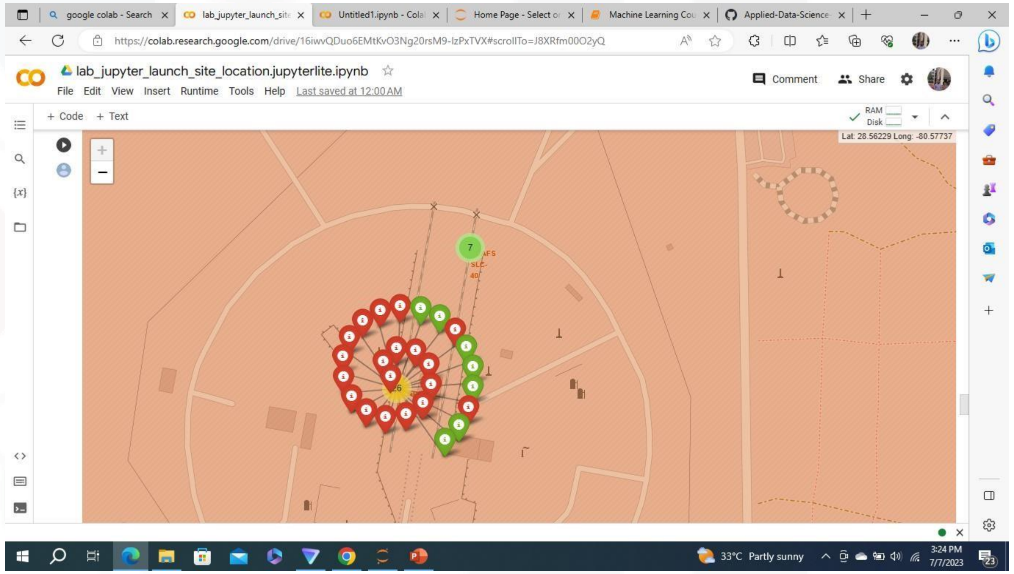
Exploratory data analysis results:

Space X uses 4 different launch sites:

- The first launches were done to Space X itself and NASA;
- The average payload of F9 v1.1 booster is 2,928 kg;
- The first success landing outcome happened in 2015 fiver year after the first launch;
- Many Falcon 9 booster versions were successful at landing in drone ships having payload above the average;
- Almost 100% of mission outcomes were successful;
- Two booster versions failed at landing in drone ships in 2015: F9 v1.1 B1012 and F9 v1.1 B1015;
- The number of landing outcomes became as better as years passed.

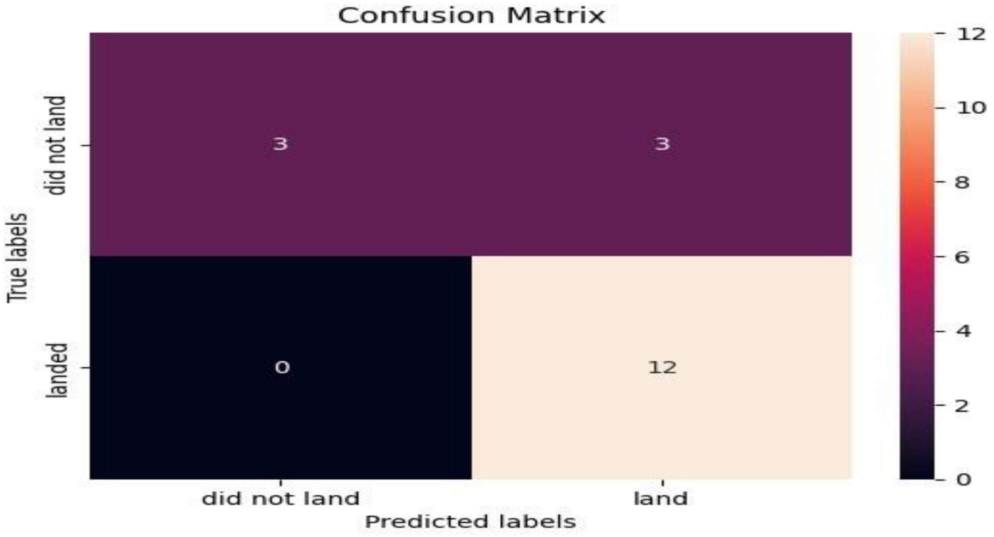
- Using interactive analytics was possible to identify that launch sites use to be in safety places, near sea, for example and have a good logistic infrastructure around.
- Most launches happens at east cost launch sites.



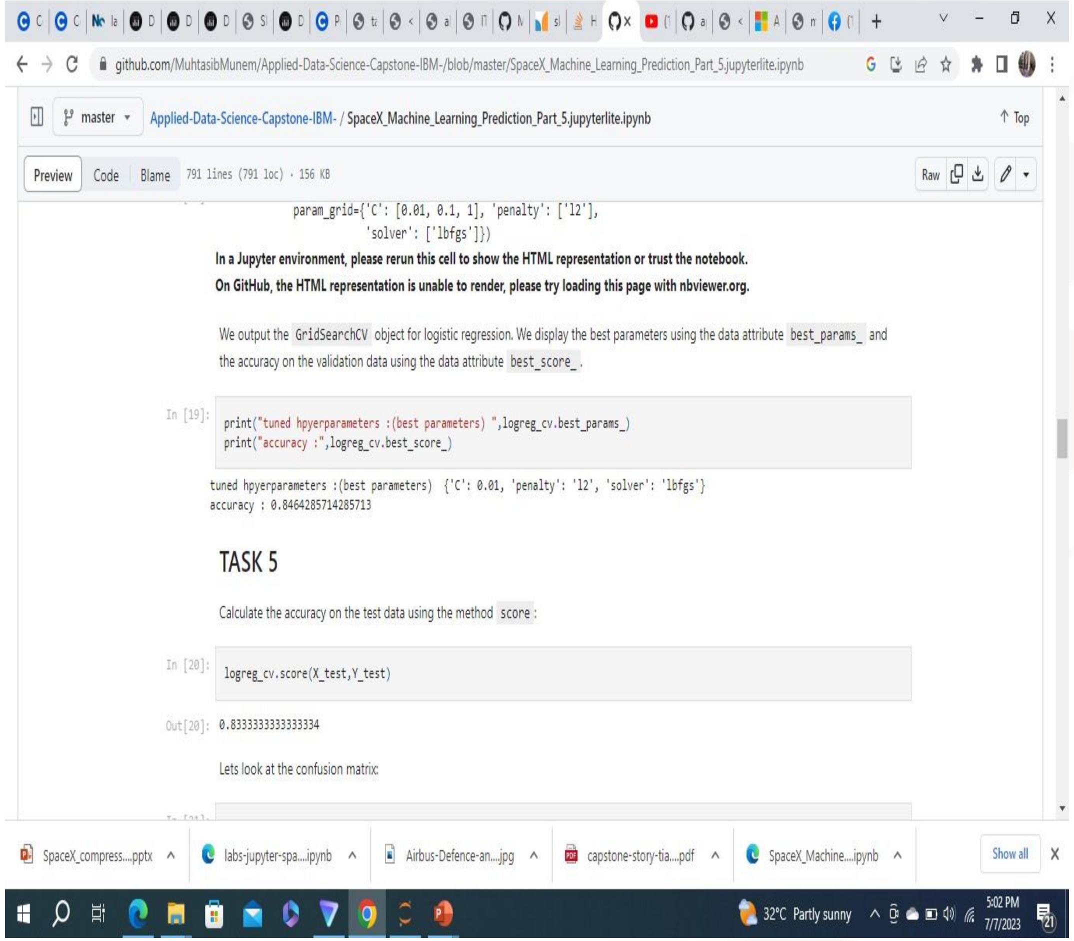


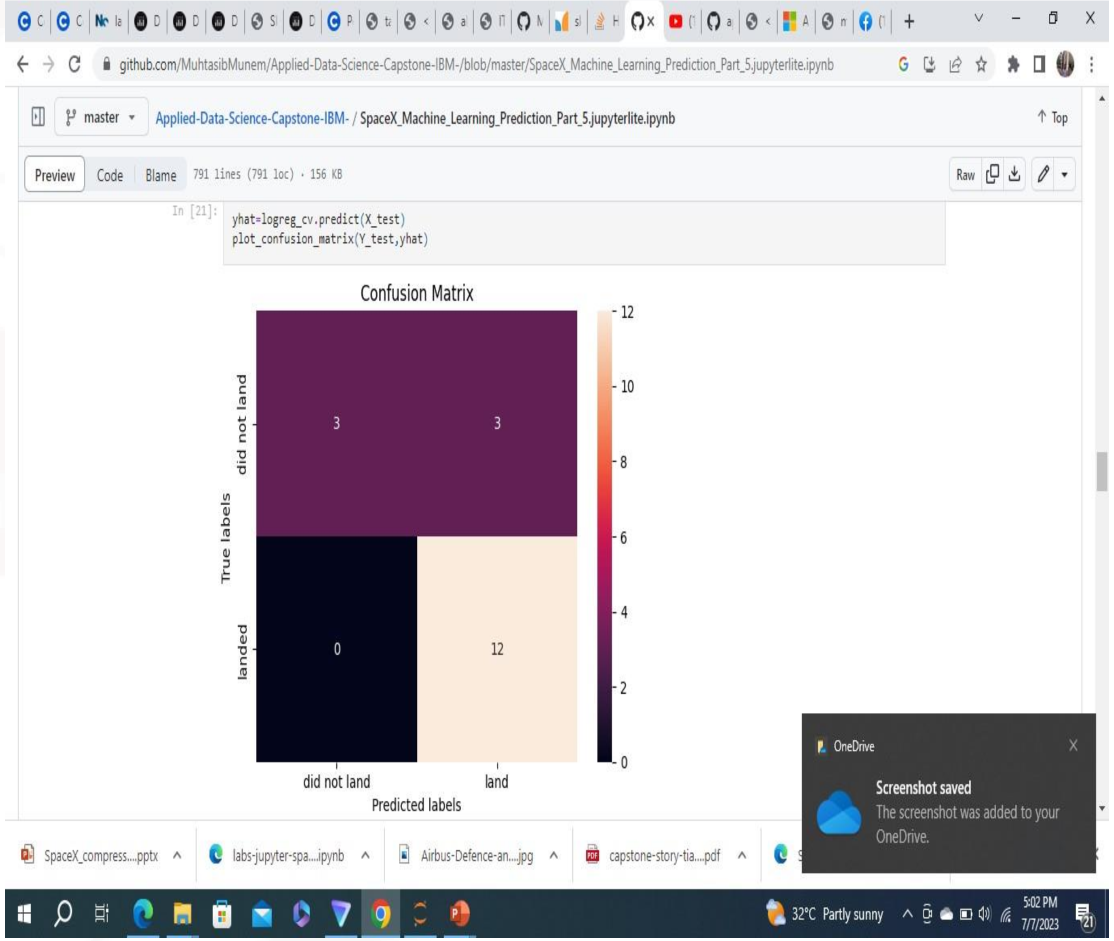
PREDICTIVE ANALYSIS RESULTS

- Predictive Analysis showed that Decision Tree Classifier is the best model to predict successful landings, having accuracy over 87% and accuracy for test data over 88%.
- The following picture is the confusion matrix of DTC:

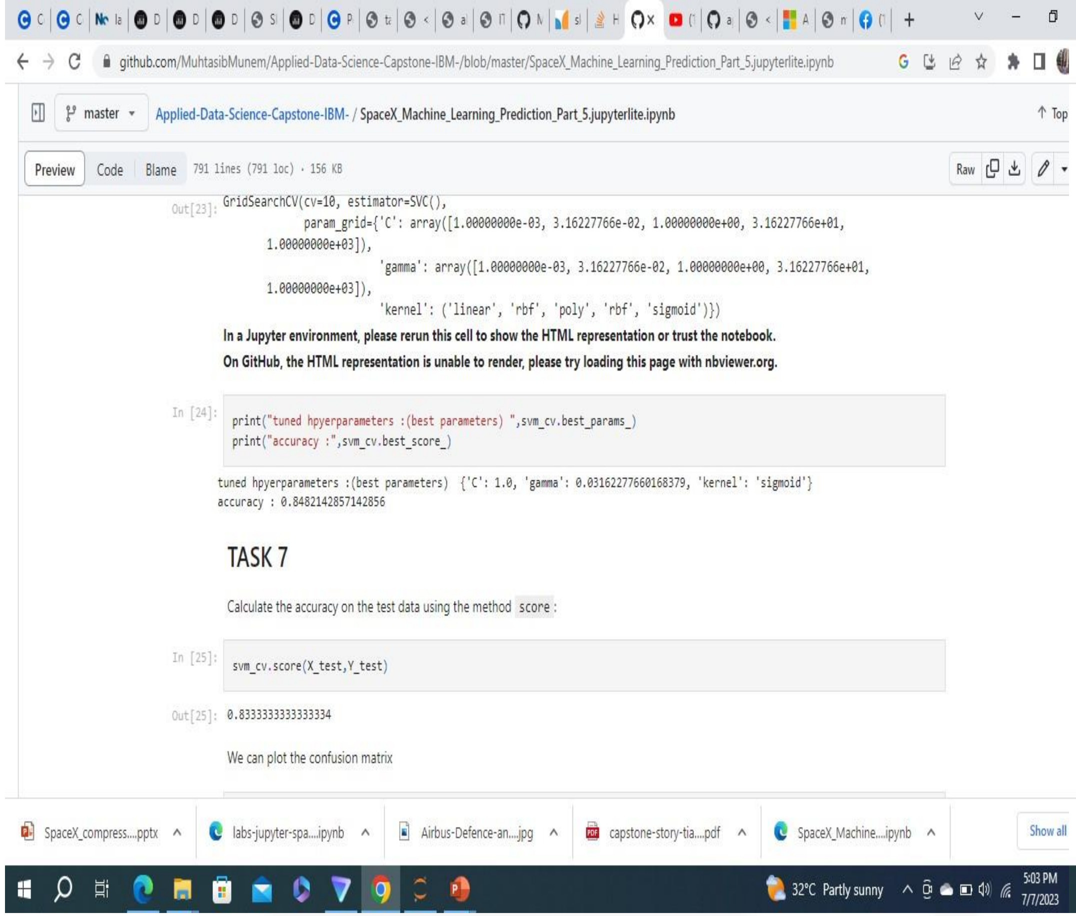


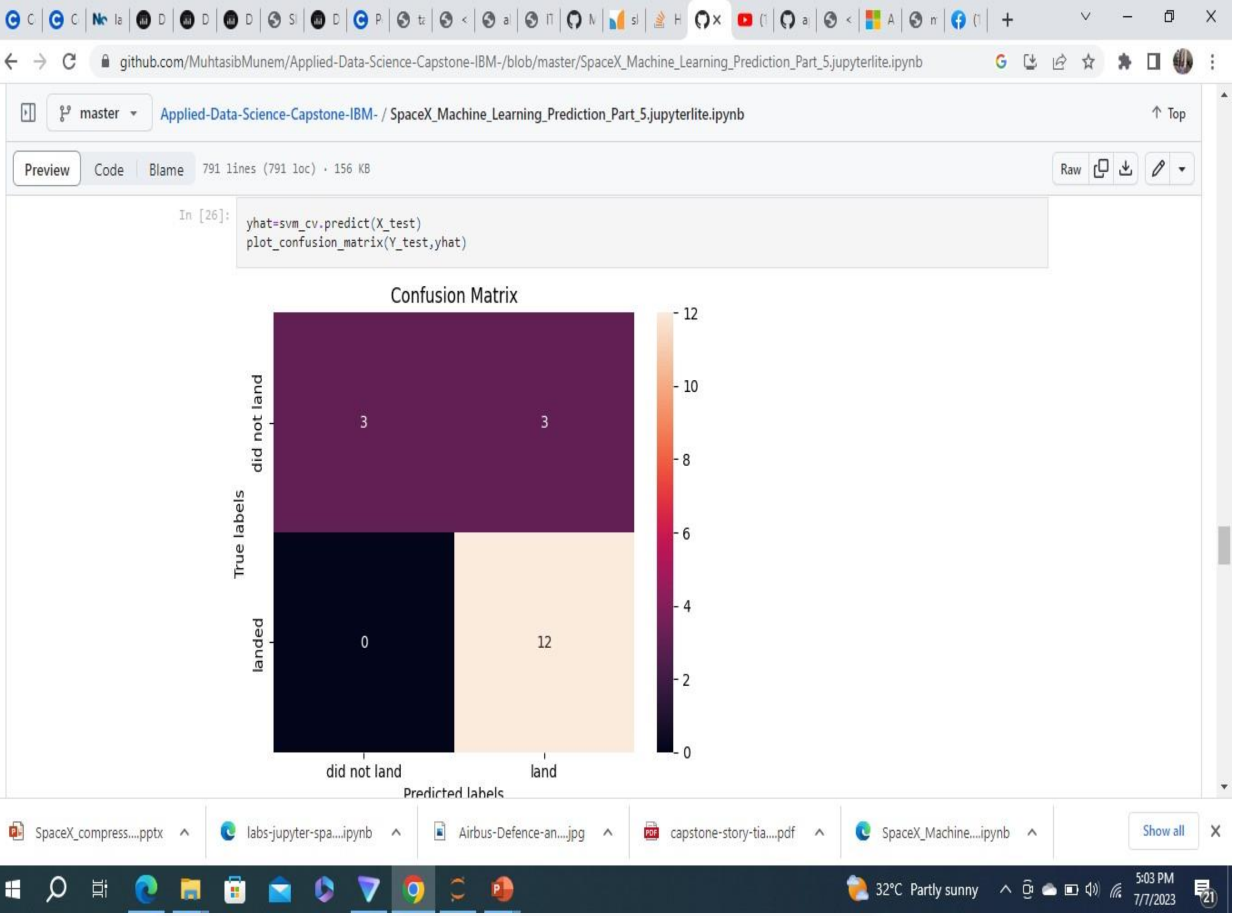
Accuracy of Logistic Regression



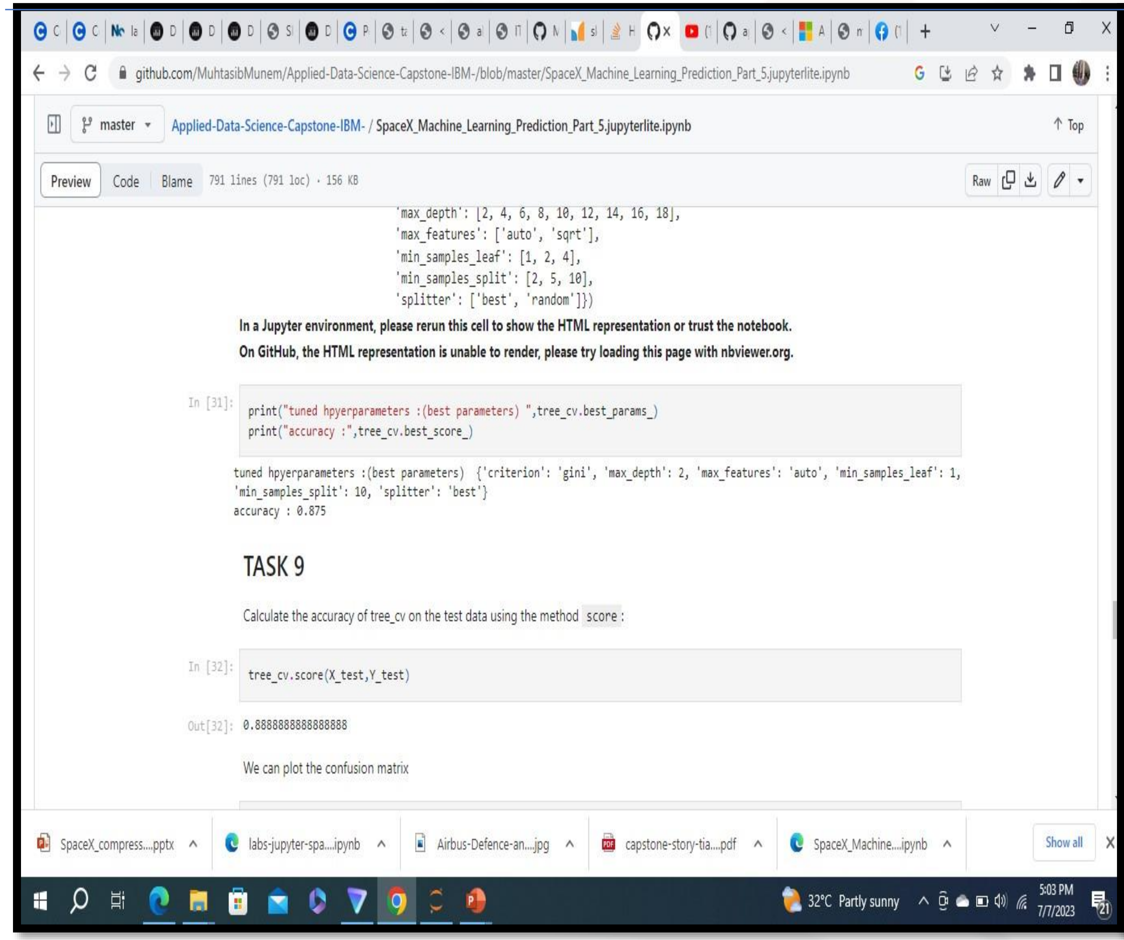


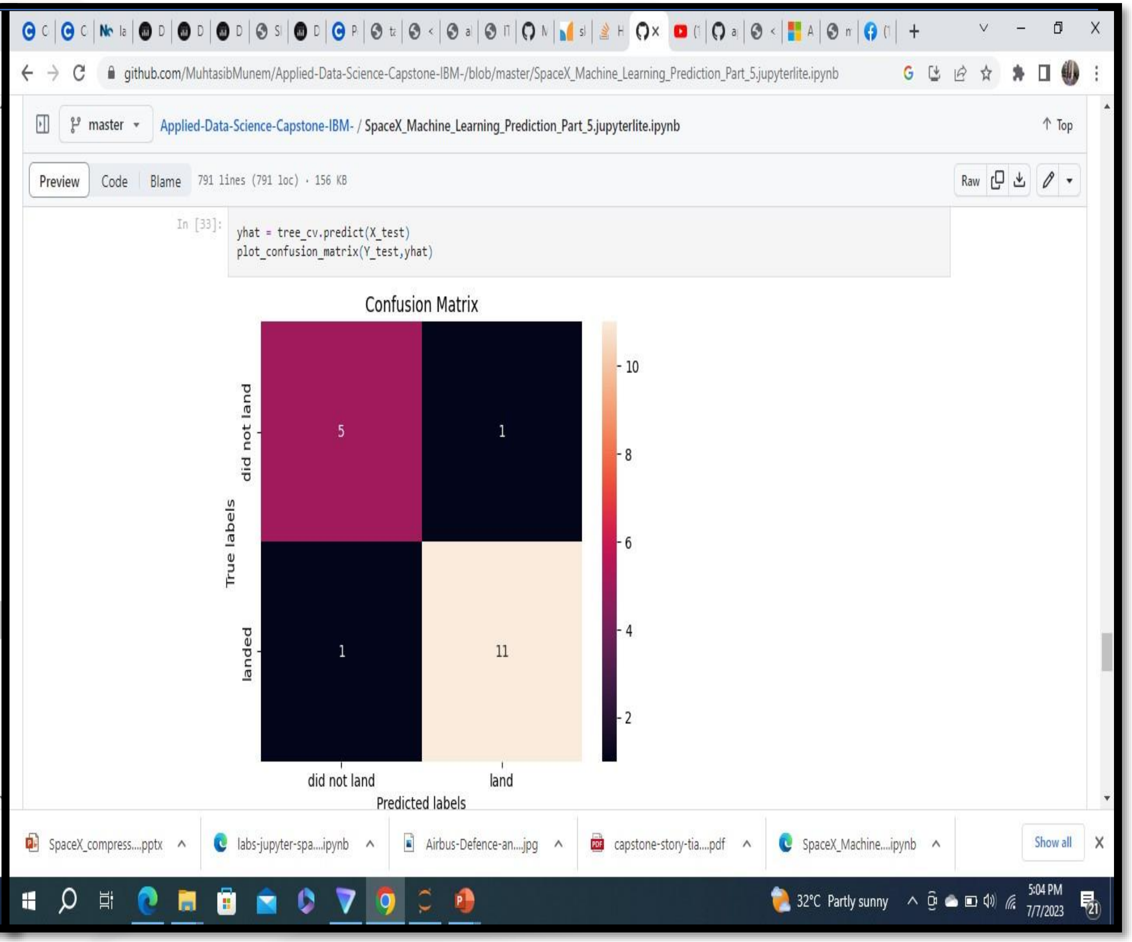
Accuracy of SVM



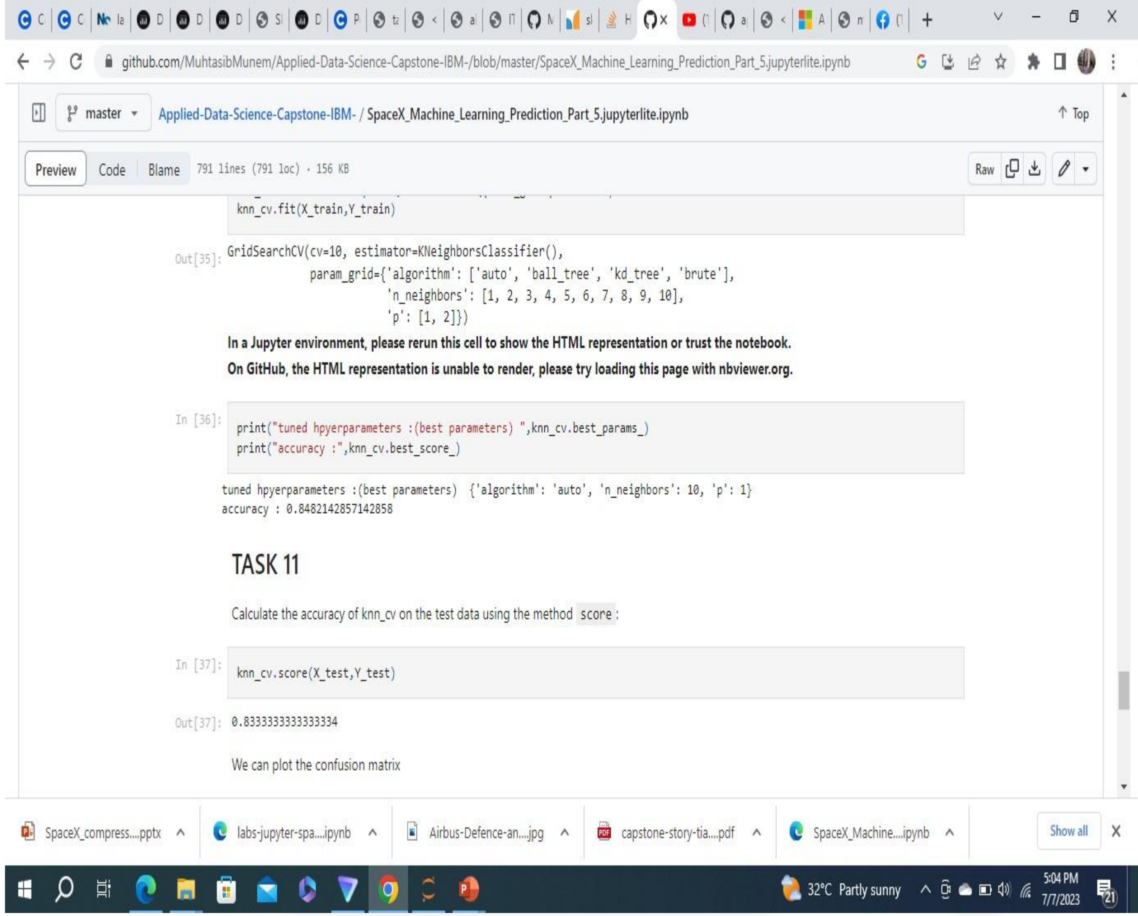


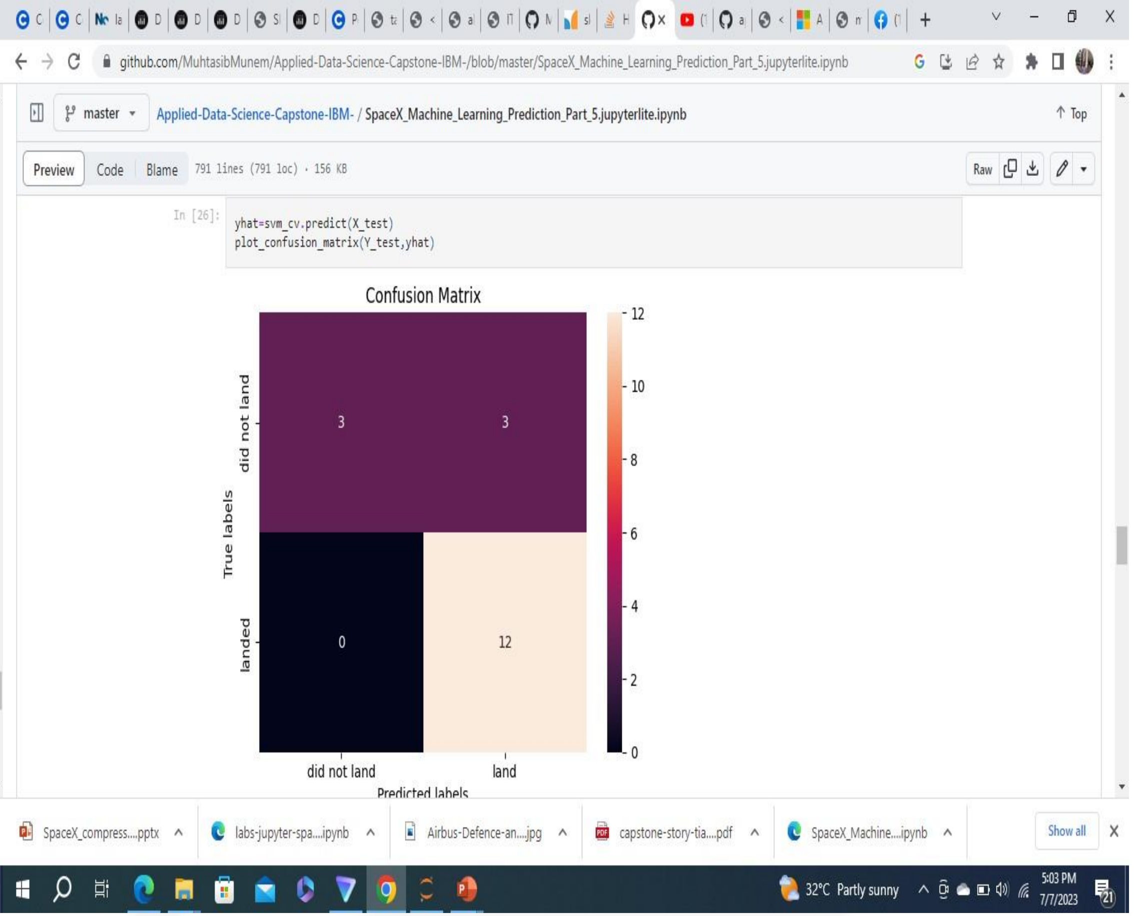
Accuracy of Decision Tree





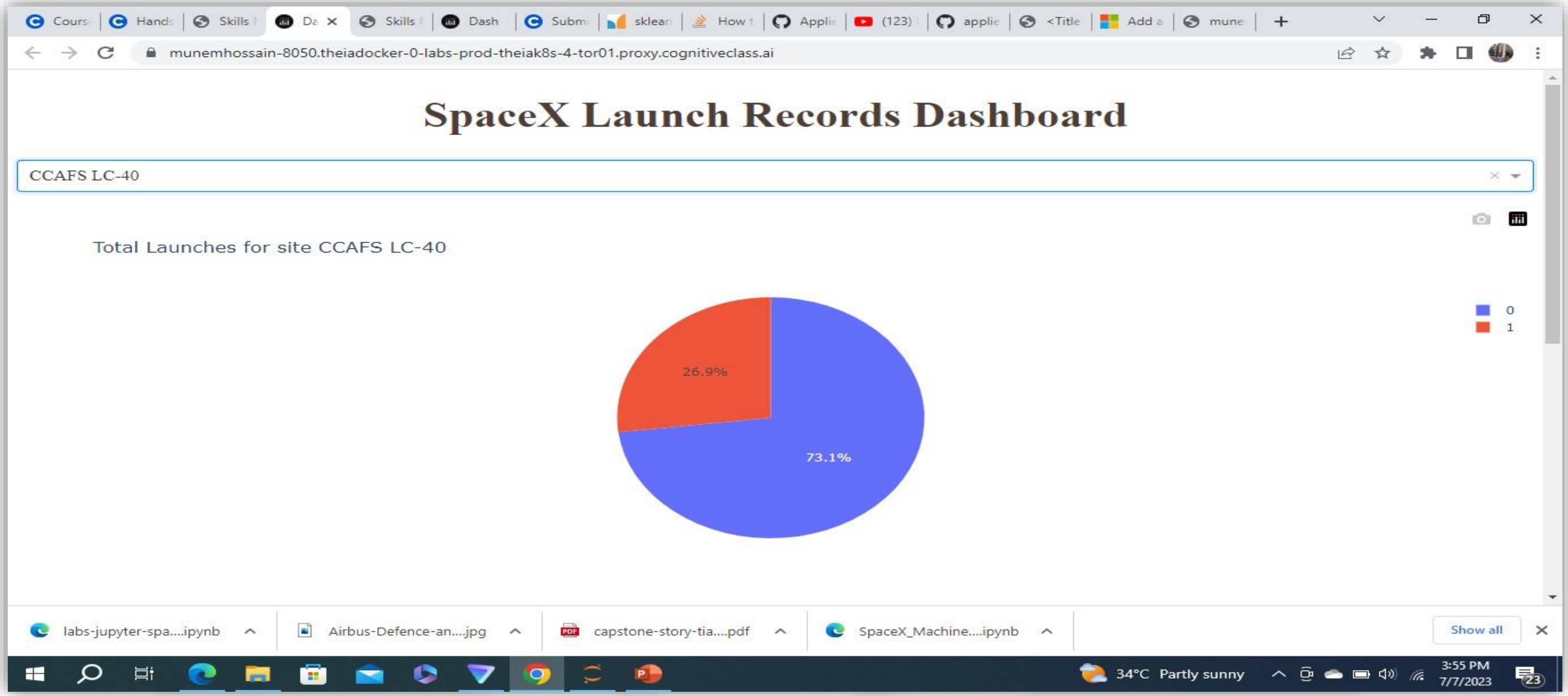
Accuracy of KNN

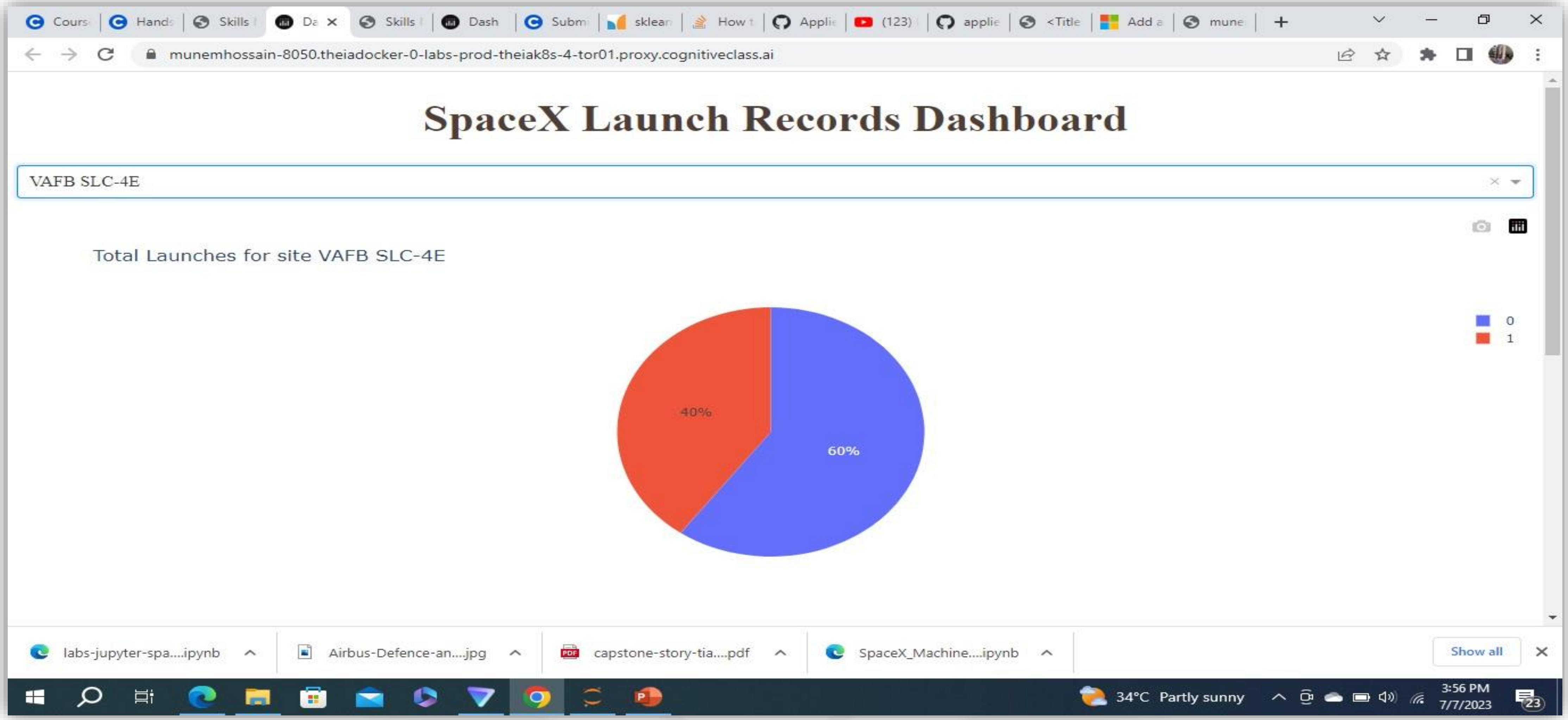




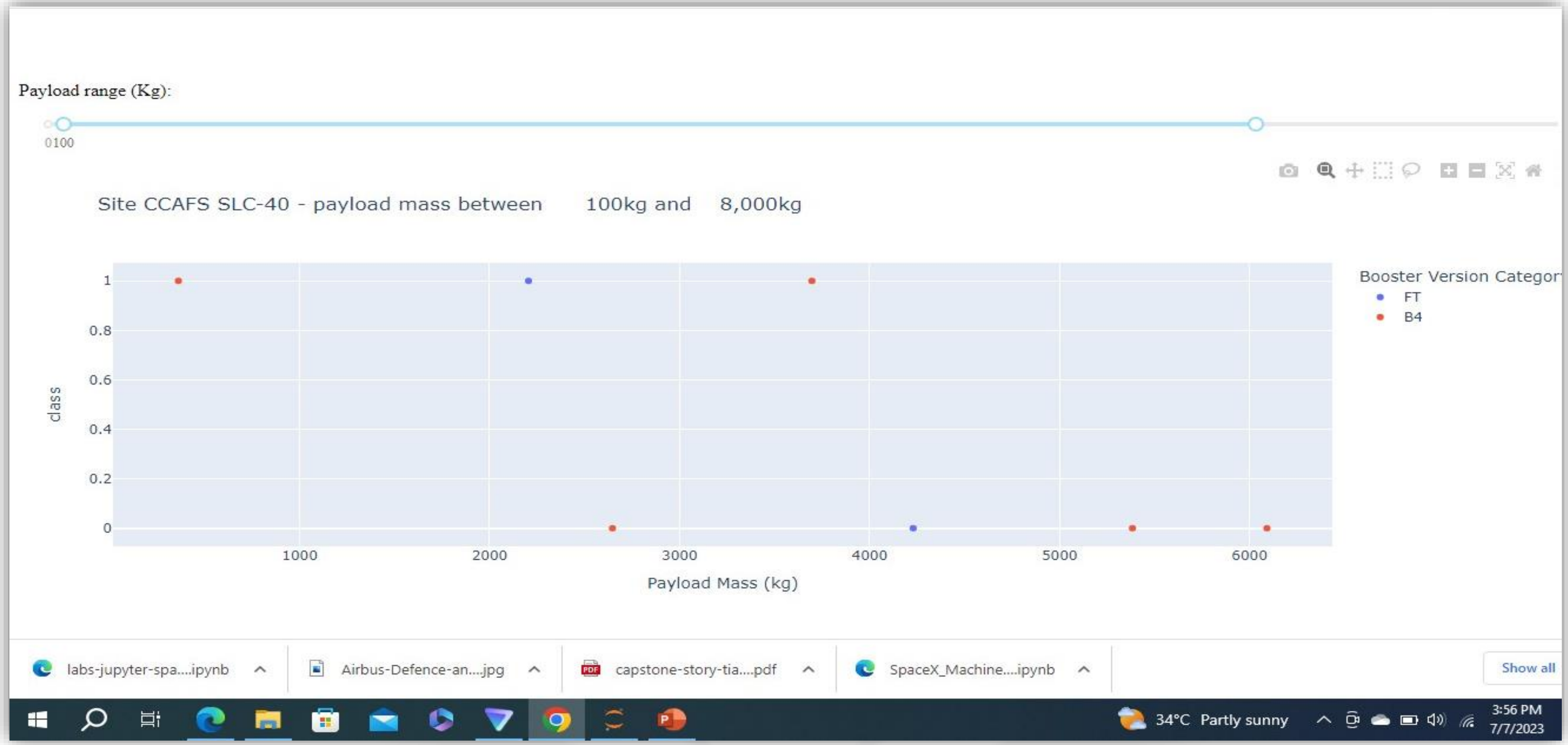
DASHBOARD

https://munemhossain-8050.theiadocker-0-labs-prodtheiak8s-4-tor01.proxy.cognitiveclass.ai/


SpaceX Launch Records Dashboard
CCAFS LC-40
Total Launches for site CCAFS LC-40
26.9%
73.1%
0
1

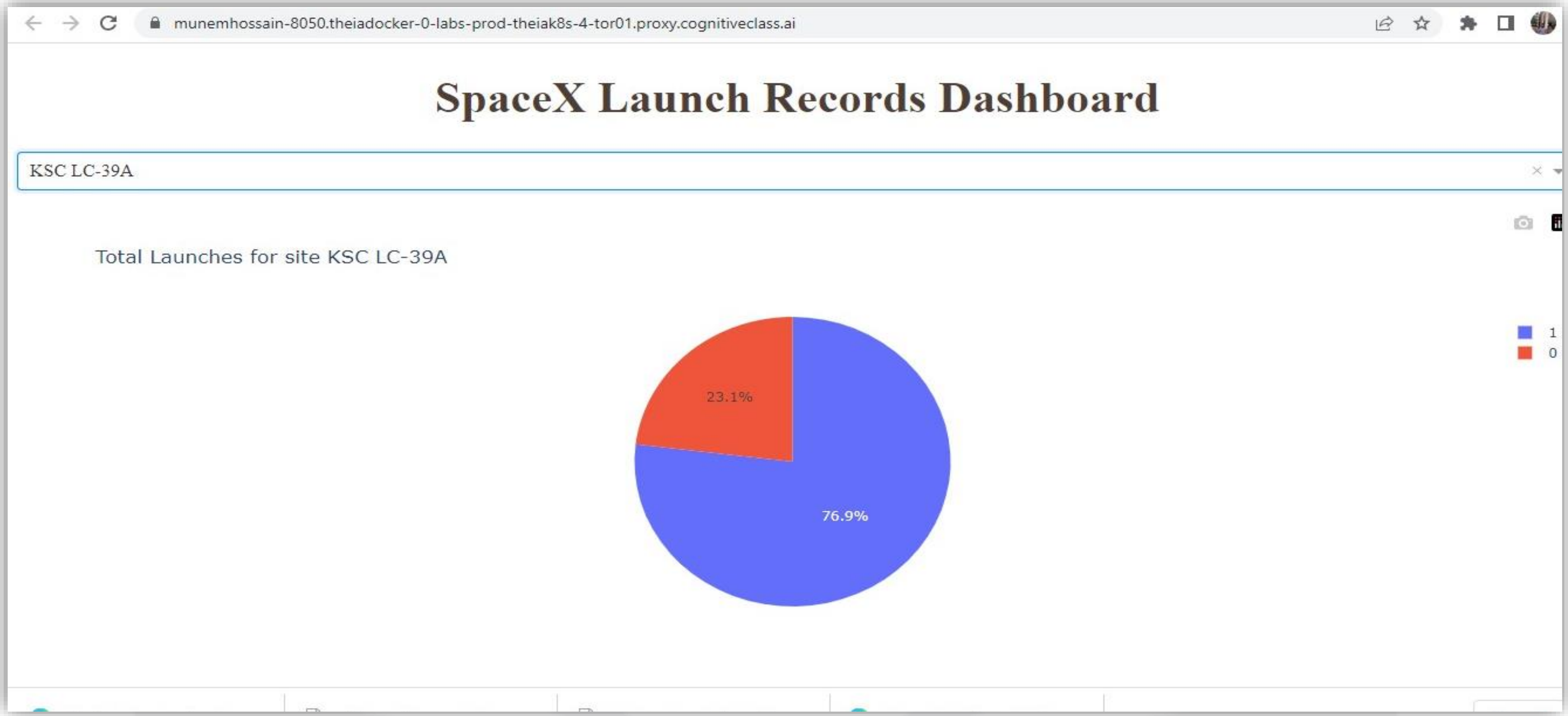


munemhossain-8050.theiadocker-0-labs-prod-theiak8s-4-tor01.proxy.cognitiveclass.ai
SpaceX Launch Records Dashboard
VAFB SLC-4E
Total Launches for site VAFB SLC-4E
40%
60%
0
1
labs-jupyter-spa....ipynb
Airbus-Defence-an....jpg
capstone-story-tia....pdf
SpaceX_Machine....ipynb
Show all
34°C Partly sunny
3:56 PM
7/7/2023


Payload range (Kg):
0100
Site CCAFS SLC-40 - payload mass between 100kg and 8,000kg
class
1
0.8
0.6
0.4
0.2
0
1000
2000
3000
4000
5000
6000
Payload Mass (kg)
Booster Version Categor
FT
B4
labs-jupyter-spa....ipynb
Airbus-Defence-an....jpg
capstone-story-tia....pdf
SpaceX_Machine....ipynb
Show all
34°C Partly sunny
3:56 PM
7/7/2023

DISCUSSION

- The place from where launches are done seems to be a very important factor of success of missions.
- 76.9% of launches are successful for KSC LC-39A which is the highest.


munemhossain-8050.theiadocker-0-labs-prod-theiak8s-4-tor01.proxy.cognitiveclass.ai
SpaceX Launch Records Dashboard
KSC LC-39A
Total Launches for site KSC LC-39A
23.1%
76.9%
1
0

OVERALL FINDINGS & IMPLICATIONS

Findings

- The best launch site is KSC LC-39A;
- Launches above 7,000kg are less risky;
- Decision Tree Classifier can be the best classifier predict successful landings and increase profits.

Implications

- The launch site is the most important factor to determine.
- Payload mass significantly effect the success of launching.
- Classification algorithm depends largely on the data quality, so understanding the data is very important.

CONCLUSION

- Different data sources were analysed, refining conclusions along the process;
- Although most of mission outcomes are successful, successful landing outcomes seem to improve over time, according the evolution of processes and rockets;
- Success will not be achieved if someone ignores the data driven insights. So, data analysis should be the first priority.