Verzeo Major Project Submission

Create a streamlit web app showing all the classification models and deploy this web app using Heroku

By- Suryaansh Jaiswal

Project review

I have made a ML project to demonstrate multiple machine learning models to classify a sample as benign - 0 or malware – 1 using a dataset

Malicious program or malware is an intentionally written program to indulge in various malicious activities, ranging from user's information stealing to cyber-espionage. The behavioral dynamism exposed by the malware is dependent on various factors such as, nature of the attack, sophisticated technology and the rapid increase in exploitable vulnerabilities. Malware attacks also increased along with the rapid growth in the use of digital devices and the internet. The exponential increase in the creation of new malware in the last five years, made malware detection as a challenging research issue. Problem Statement Malware is one of the top most obstructions for expansion and growth of digital acceptance among the users. Both enterprises and common users are struggling to get protected from the malware in cyberspace, which emphasizes the importance of developing efficient methods of malware detection. Malware detection is the technique for identifying malware in the end devices or networks.

Tools Used

Python

Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notableuse of significant indentation.

Google colab

Colab notebooks allow you to combine executable code and richtext in a single document, along with images, HTML, LaTeX and more.

Algorithm

KNN-

The k-nearest neighbors (KNN) algorithm is used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.

It is one of the fundamental algorithms in machine learning. Machine learning models use a set of input values to predict output values. KNN is one of the simplestforms of machine learning algorithms mostly used for classification. It classifies the data point on how its neighbor is classified.

Logistic Regression-

In statistics, the logistic model is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick. This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc.

Naïve Bayes-

Naive Bayes is a classification algorithm that is suitable for binary and multiclass classification. It is a supervised classification technique used to classify future objects by assigning class labels to instances/records using conditional probability.

SVC-

The objective of a Linear SVC (Support Vector Classifier) is to fit to the data you provide, returning a "best fit" hyperplane that divides, or categorizes, your data. From there, after getting the hyperplane, you can then feed some features to your classifier to see what the "predicted" class is

Conclusion-

The F1 score calculated after using Logistic Regression is 0.918

F1 score

```
from sklearn.metrics import f1_score
print(f1_score(y_test,y_pred))
```

```
0.9182692307692308
```

The F1 score calculated after using KNN is 0.938

F1 Score

```
[59] from sklearn.metrics import f1_score
     print(f1_score(y_test,y_pred))
```

```
0.9386281588447652
```

Hence the F1 Score from KNN comes out to be 0.938

The F1 score calculated after using SVC is 0.918

F1 Score

```
[43] from sklearn.metrics import f1_score
     print(f1_score(y_test,y_pred))
```

```
0.9182692307692308
```

Hence the F1 Score from SVC comes out to be 0.918

The F1 score calculated after using Naïve Bayes is 0.807

F1 Score

```
[84] from sklearn.metrics import f1_score
     print(f1_score(y_test,y_pred))
```

```
0.8077709611451943
```

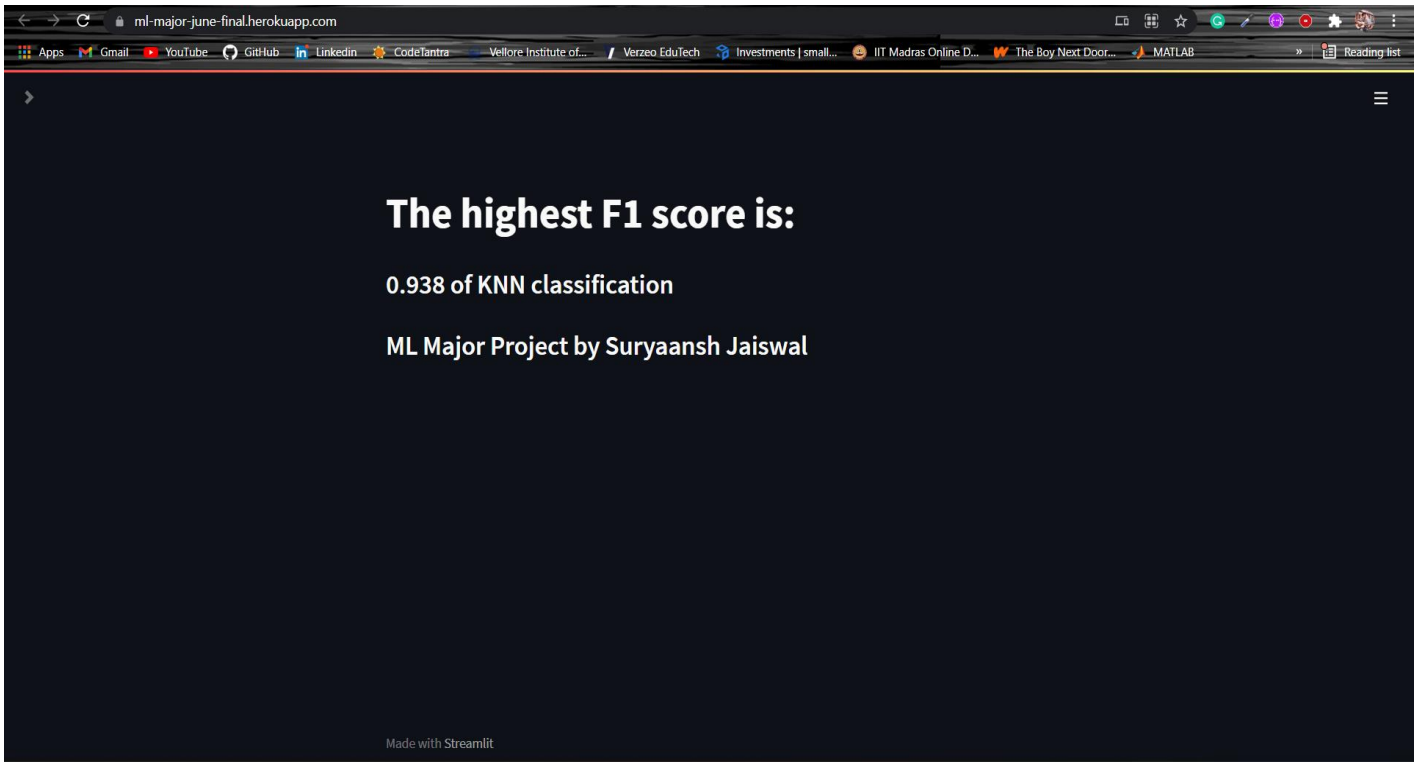
Hence the F1 Score from Naive Bayes comes out to be 0.8077

Hence the F1 Score calculated by KNN is the Largest

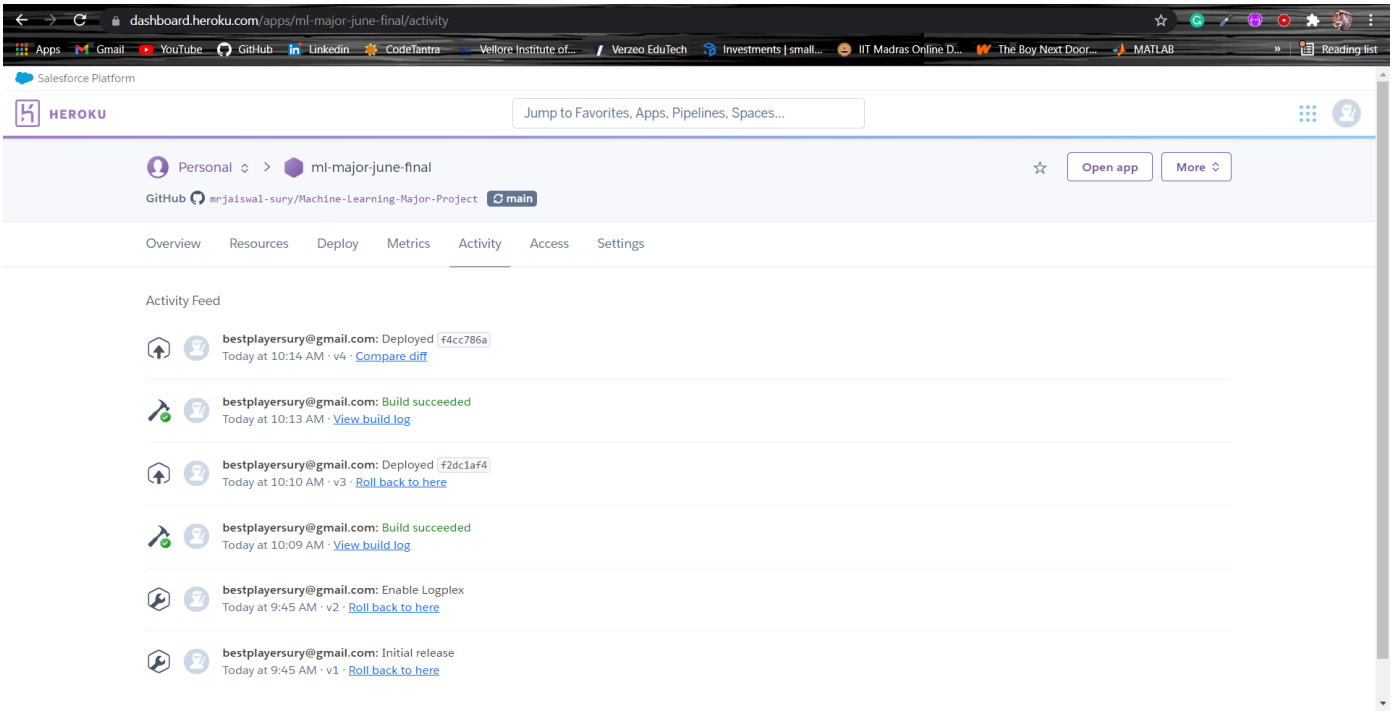
Heroku App Link-

https://ml-major-june-final.herokuapp.com/

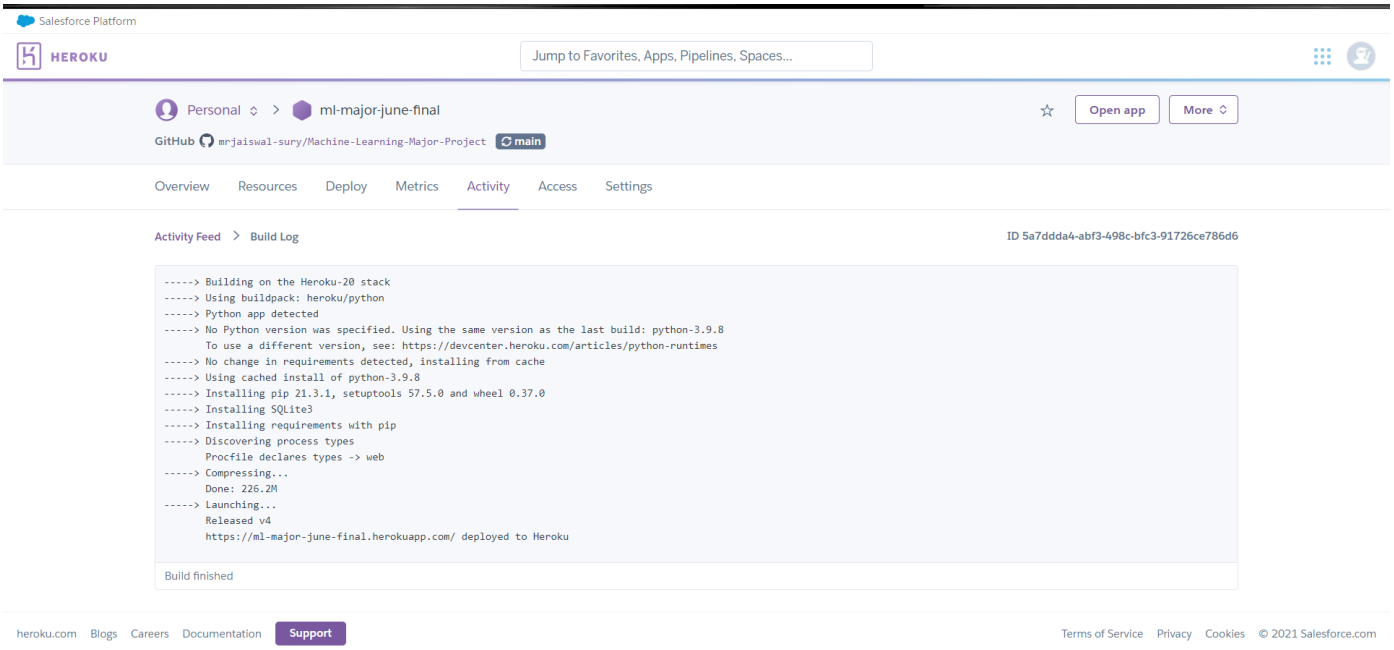
Output Screen-



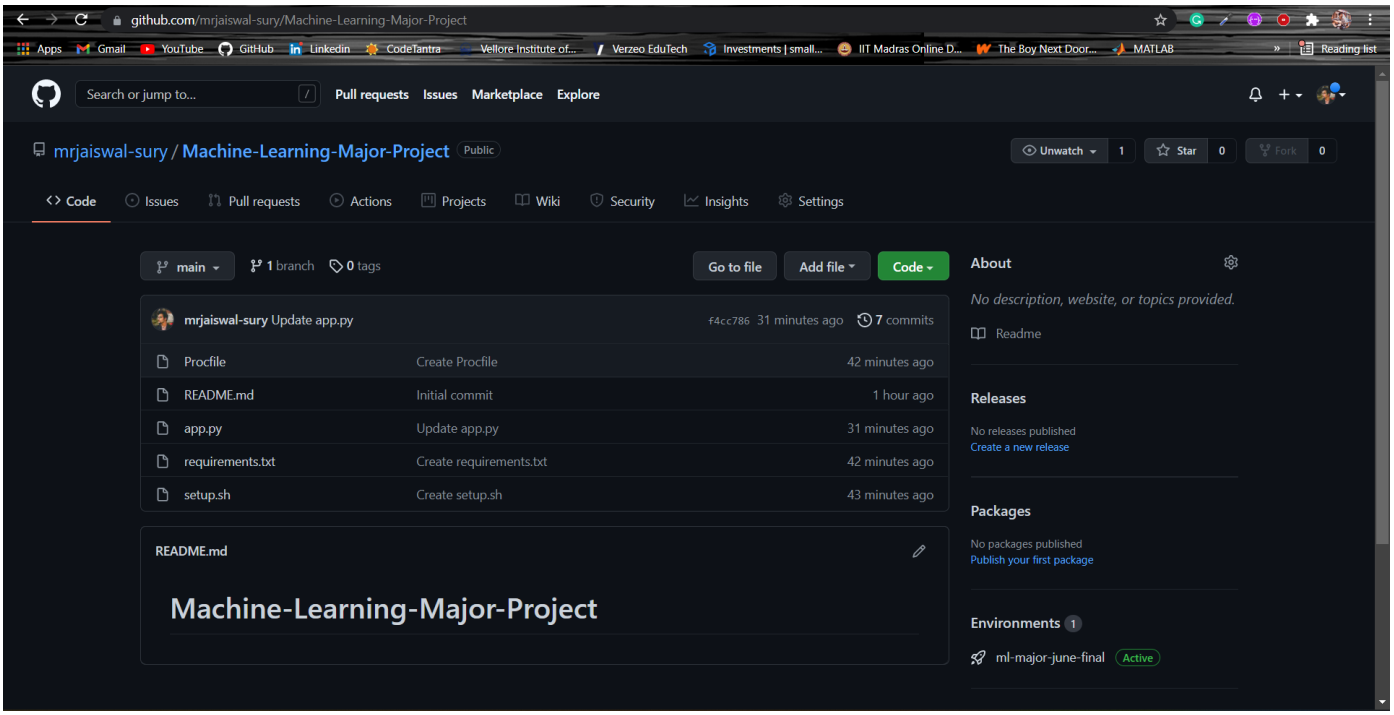
Activity section of the Heroku Cloud-



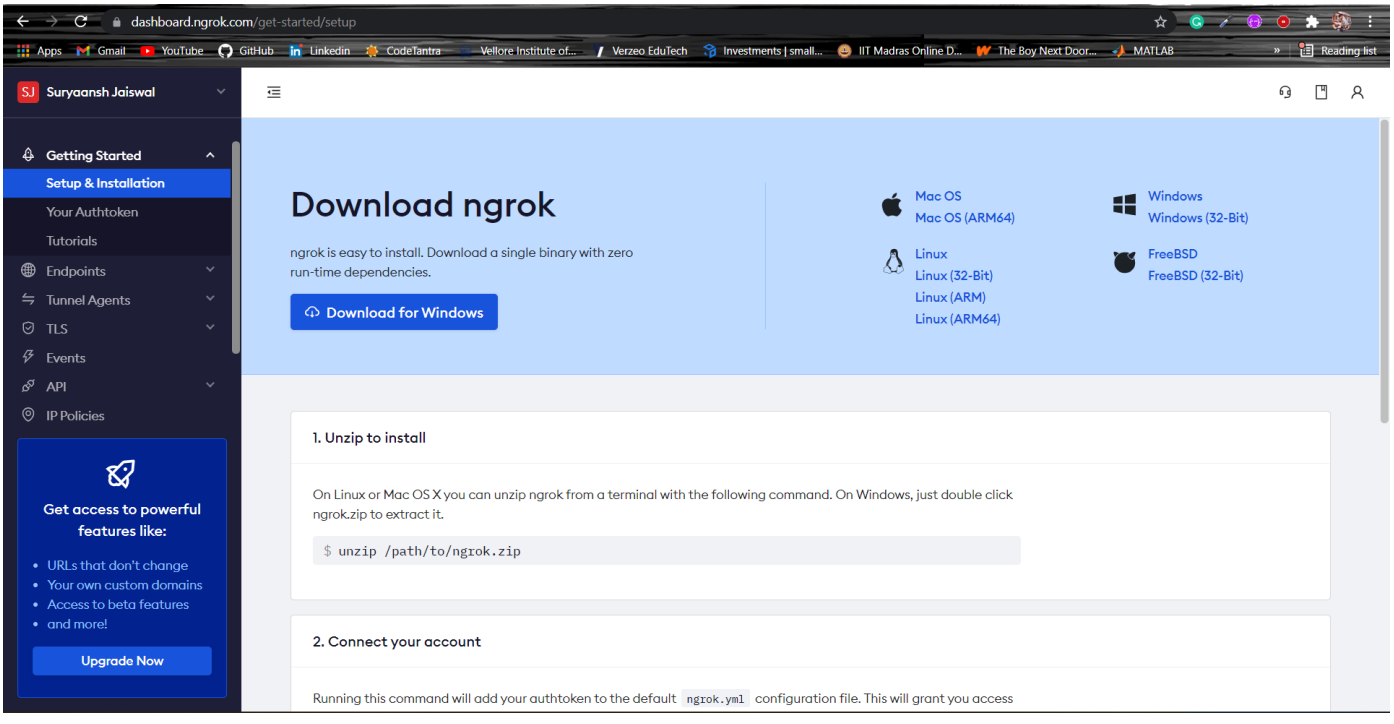
Build Log of the App-



Github Page for the Project



Ngrok Account for Streamlit-



Thank You