ML Major June
ML Major Project
Report

Problem Statement : You need to build a machine learning model and Webapp (Streamlit) to accurately predict whether or not the patients in the dataset have diabetes or not?

Dataset : Diabetes

Support Vector Classifier :

SVC is a nonparametric clustering algorithm that does not make any assumption on the number or shape of the clusters in the data. In our experience it works best for low-dimensional data, so if your data is high-dimensional, a preprocessing step, e.g. using principal component analysis, is usually required.

Accuracies

When kernel = “linear” : 82.4%

When kernel = “rbf” : 79.2%

```
SVC with kernel = linear Test Accuracy: 0.8246753246753247
SVC with kernel = rbf Test Accuracy: 0.7922077922077922
```

Applied a Support Vector Classifier as accuracy of SVC is high.
Considered kernel=”linear” for building webapp

Pickle:

Used pickle to extract classifier from model and used it in streamlit.

Streamlit :

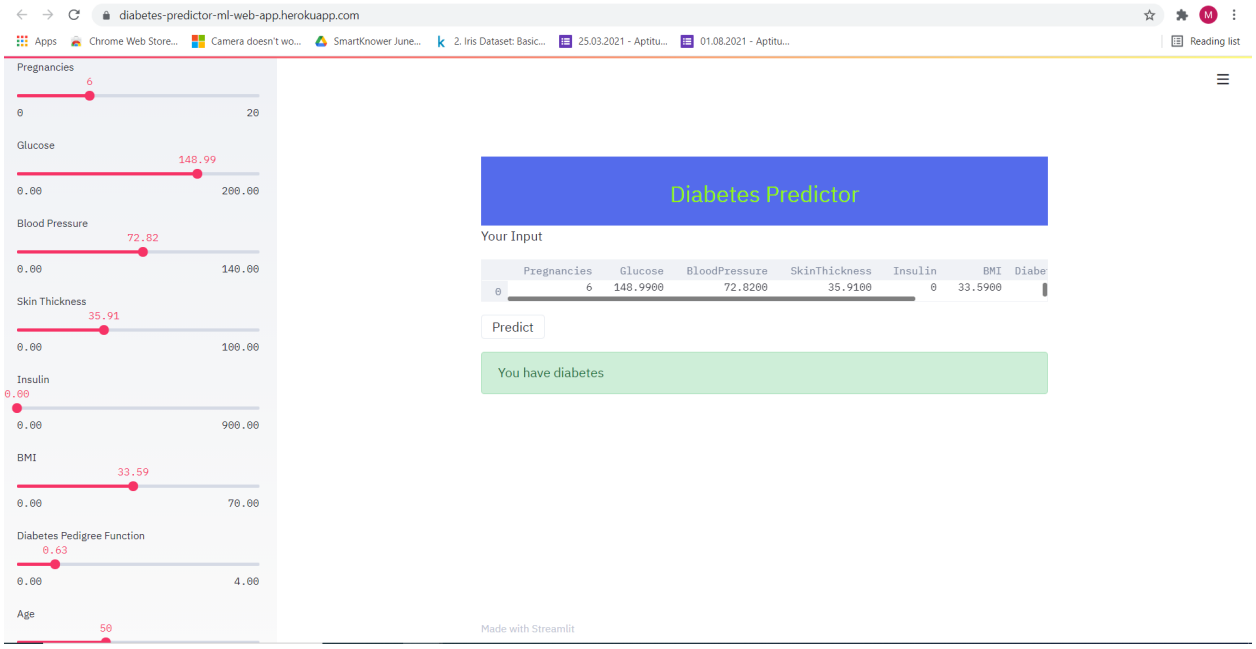
Stream is used to build a webapp for machine learning models.

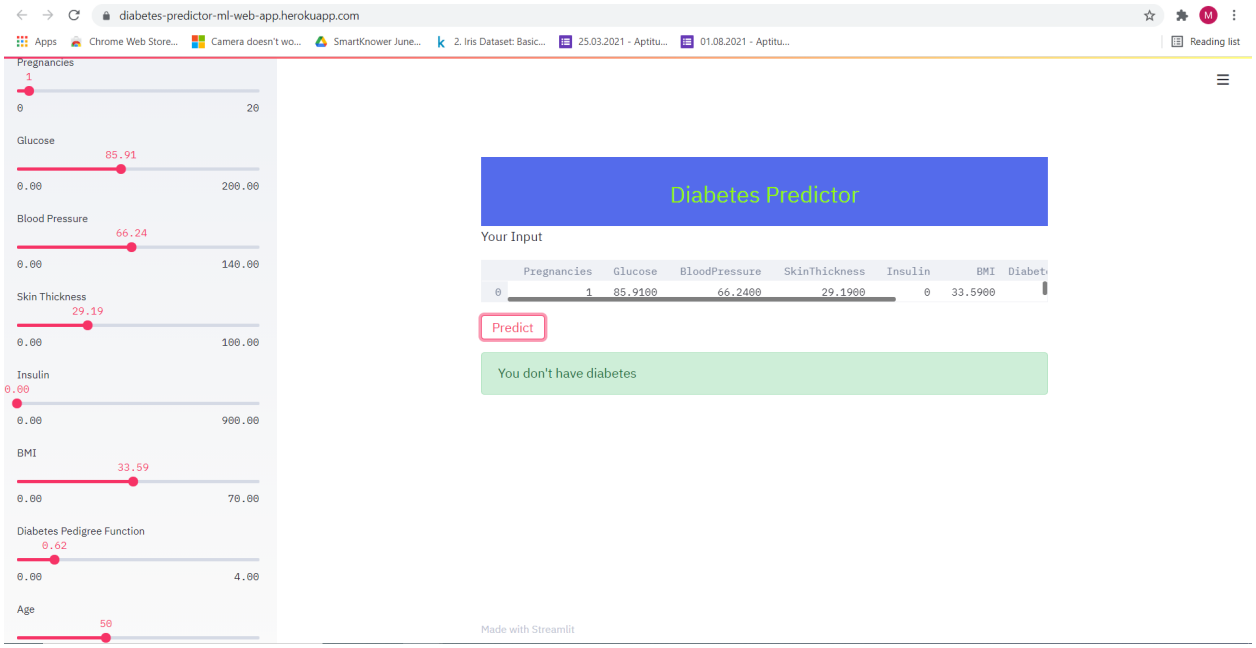
Pyngrok is used for generating url for web app.

Heroku :

Machine learning model is deployed in heroku using github(files uploaded in github). After being deployed in heroku the web app is live. Anyone can access. Anyone can check if they have diabetes or not by giving respective inputs.

Web app





Github Link :
https://github.com/midhun999/Diabetes-Predictor-ML-Web-App
Heroku Link : https://diabetes-predictor-ml-web-app.herokuapp.com/