Data Science Intern AT Data Glacier

Week 5: Cloud API development

Name: Rayan Yassminh

Batch Code: LISUM13: 30

Summitted to: Data Glacier

Model Deployment using Heroku

Table of Contents:

1.Introduction

In this project I will deploy the Logistic regression model using Flask. My ML model predicts if someone is diabetic or not. The goal of this project is to create API for the ML model using Flask. Flask is a framework for building a web application.

2.Data information

Researchers at the Bio-Solutions lab want to get better understanding of diabetes disease among women and are planning to use machine learning models that will help them to identify patients who are at risk of diabetes. The "Pima Indians Diabetes.csv" dataset was collected. All patients here are females at least 21 years old of Pima Indian heritage.
Our primarily ML model shows the most important features to make the prediction; Therefore, I filtered the dataset to include only the important features then we rebuild the model on the filtered Dataset.

Pregnancies; 768 non-null int64
Glucose : 768 non-null float64
BMI : 768 non-null float64
Pedigree: 768 non-null float64
Class : 768 non-null int64
dtypes: float64(3), int64(2)
memory usage: 30.1 KB

Data Description:

- Pregnancies: Number of times pregnant
- Glucose: Plasma glucose concentration over 2 hours in an oral glucose tolerance test
- BMI: Body mass index (weight in kg/(height in m)^2)
- Pedigree: Diabetes pedigree function - A function that scores likelihood of diabetes based on family history.
- Class: Class variable (0: the person is not diabetic or 1: the person is diabetic)

Index	*Pregnancies*	*Glucose*	*BMI*	*Pedigree*	*Class*
566	*1*	*99.0*	*38.6*	*0.412*	*0*
207	*5*	*162.0*	*37.7*	*0.151*	*1*
58	*0*	*146.0*	*40.5*	*1.781*	*0*
686	*3*	*130.0*	*23.1*	*0.314*	*0*

751	*1*	*121.0*	*39.0*	*0.261*	*0*

3. The Machine learning model

Import the libraries:

```
# To filter the warnings
import warnings
warnings.filterwarnings("ignore")
# Libraries to help with reading and manipulating data
import pandas as pd
import numpy as np
# libaries to help with data visualization
import matplotlib.pyplot as plt
import seaborn as sns
# Library to split data
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
# To build linear model for statistical analysis and prediction
# To get diferent metric scores
from sklearn import metrics
from sklearn.metrics import accuracy_score, roc_curve, confusion_matrix, roc_auc_score,recall_score
import pickle
```

Read the dataset

```
data = pd.read_csv("pima-indians-diabetes.csv")
# defining the most important columns and replace 0 with NaN
cols = ["Glucose", "BMI", "Pedigree"]
# replacing 0 with NaN
data[cols] = data[cols].replace(0, np.nan)
# Let's impute missing values using mean value
data[cols] = data[cols].fillna(data[cols].mean())
#Target and independent features.
X = data.drop(["Class"], axis=1)
Y = data["Class"]
# split data into training and test
X_train,X_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=1,stratify=Y)
```

Logistic Regression model

```
lg=LogisticRegression(penalty='elasticnet',l1_ratio=0.9,max_iter=500,solver='saga',class_weight='balanced')
lg.fit(X_train,y_train)
# convert the model pickle.
pickle.dump(lg, open('model.pkl','wb'))
```

4. Applying ML Model to flask framework

A. Creating application files

We create a folder for the application. The directory tree inside this folder as follow:

```
week4
│   app.py
│    logistic_regression_diabetes.py
│   Logistic_Regression_Hands-On.ipynb
│   model.pkl
│   pima-indians-diabetes.csv
│   request.py
│
├───static
│       script.js
│       style.css
│
└───templates
        index.html
```

- APP.Py

App.py contains the main code that will be run by python. It includes the Machine leaning model and route the Index.html. the route decorator used to map the URL to a return value that means connect url to return value of a function.

```
from webbrowser import get
import numpy as np
import pandas as pd
import pickle
from flask import Flask,request,render_template,jsonify

app=Flask(__name__)
model=pickle.load(open('model.pkl','rb'))
@app.route('/') # take to the out page
def home():
    return render_template('index.html')
@app.route('/predict',methods=['post','get'])
def predict():
    '''
    For rendering results on HTML GUI
    '''
    float_features = [float(x) for x in request.form.values()]
    final_features = [np.array(float_features)]
    prediction = model.predict(final_features)

    output = round(prediction[0], 2)
    if output == 1:
        output_text='Diabetic'
    elif output == 0:
        output_text='Not Diabetic'

    return render_template('index.html',prediction_text=f'You are {output_text}')
if __name__ == "__main__":
    app.run(debug=True)
```

Index.html

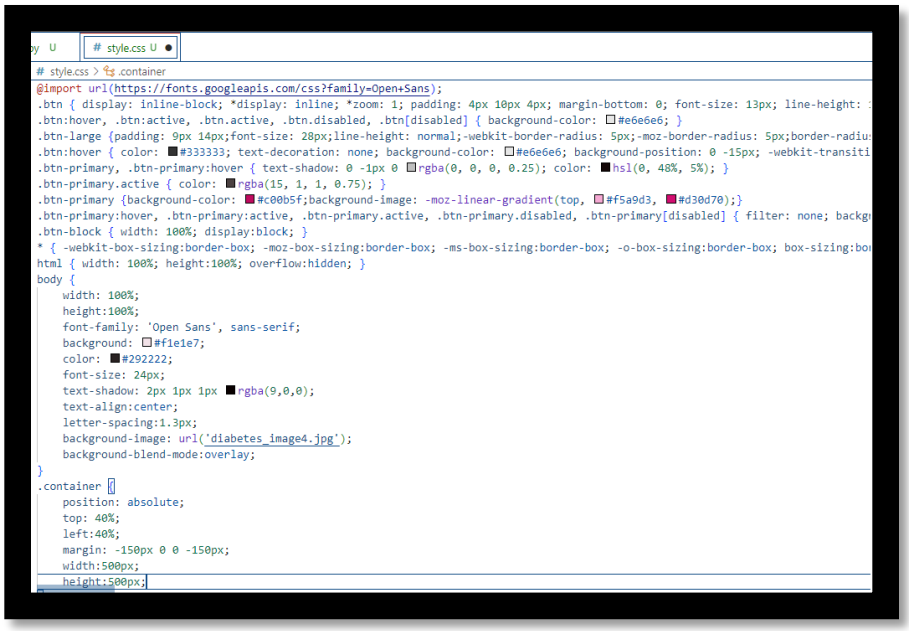
The index file is a html file. Index.html contains HTML code or tags with some CSS and JavaScript code. It should be in templates folder. When the web site is requested, by default index.html file is returned.

templates > index.html

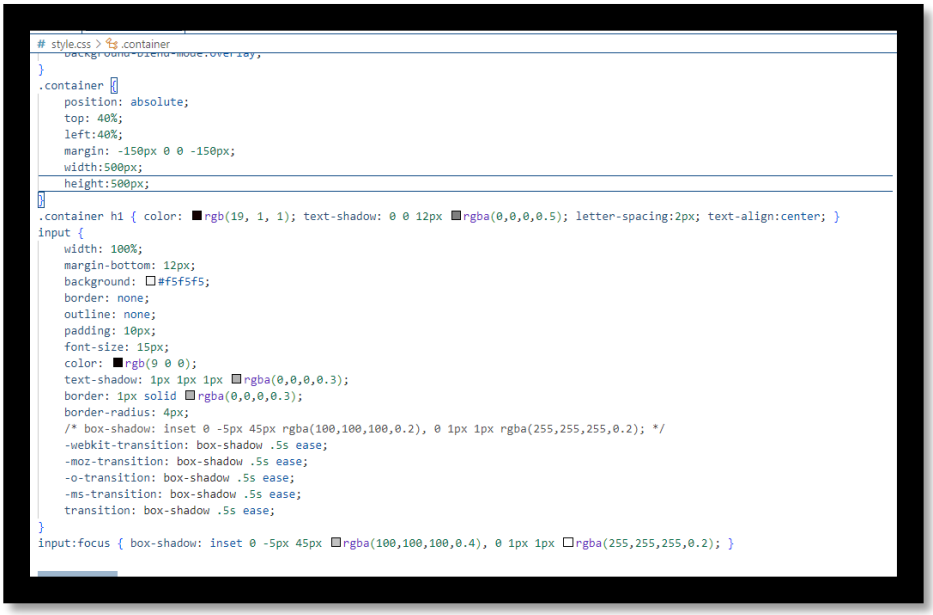
```
<!DOCTYPE html>
<html >
<head>

 <title>Machine Learning Model</title>
<link rel="stylesheet" href="{{ url_for('static', filename='css/style.css') }}"></link>
<script type="test/javascript" src="{{ url_for('static', filename='script/script.js') }}">
</script>
</head>

<body>
 <div class="container">
  <h1>Predict Diabetes</h1>

    <!-- Main Input For Receiving Query to our ML -->
    <form action="{{ url_for('predict')}}"method="post">
     <input type="text" name="Pregnancies" placeholder="Pregnancies" required="required" />
      <input type="text" name="Glucose" placeholder="Glucose" required="required" />
    <input type="text" name="BMI" placeholder="BMI" required="required" />
      <input type="text" name="Pedigree" placeholder="Pedigree" required="required" />

      <button type="submit" class="btn btn-primary btn-block btn-large">Predict</button>
    </form>

   <br>
   <br>
   {{ prediction_text }}

 </div>


</body>
```

- style.css

CSS file is necessary to determine how the HTML API looks and it must be saved in static folder.



```
@import url(https://fonts.googleapis.com/css?family=Open+Sans);
.btn { display: inline-block; *display: inline; *zoom: 1; padding: 4px 10px 4px; margin-bottom: 0; font-size: 13px; line-height:
.btn:hover, .btn:active, .btn.active, .btn.disabled, .btn[disabled] { background-color: #e6e6e6; }
.btn-large {padding: 9px 14px;font-size: 28px;line-height: normal;-webkit-border-radius: 5px;-moz-border-radius: 5px;border-radiu
.btn:hover { color: #333333; text-decoration: none; background-color: #e6e6e6; background-position: 0 -15px; -webkit-transiti
.btn-primary, .btn-primary:hover { text-shadow: 0 -1px 0 rgba(0, 0, 0, 0.25); color: hsl(0, 48%, 5%); }
.btn-primary.active { color: rgba(15, 1, 1, 0.75); }
.btn-primary {background-color: #c00b5f;background-image: -moz-linear-gradient(top, #f5a9d3, #d30d70);}
.btn-primary:hover, .btn-primary:active, .btn-primary.active, .btn-primary.disabled, .btn-primary[disabled] { filter: none; backg
.btn-block { width: 100%; display:block; }
* { -webkit-box-sizing:border-box; -moz-box-sizing:border-box; -ms-box-sizing:border-box; -o-box-sizing:border-box; box-sizing:bo
html { width: 100%; height:100%; overflow:hidden; }
body {
    width: 100%;
    height:100%;
    font-family: 'Open Sans', sans-serif;
    background: #f1e1e7;
    color: #292222;
    font-size: 24px;
    text-shadow: 2px 1px 1px rgba(9,0,0);
    text-align:center;
    letter-spacing:1.3px;
    background-image: url('diabetes_image4.jpg');
    background-blend-mode:overlay;
}
.container {
    position: absolute;
    top: 40%;
    left:40%;
    margin: -150px 0 0 -150px;
    width:500px;
    height:500px;
```



```
}
.container {
    position: absolute;
    top: 40%;
    left:40%;
    margin: -150px 0 0 -150px;
    width:500px;
    height:500px;
}
.container h1 { color: rgb(19, 1, 1); text-shadow: 0 0 12px rgba(0,0,0,0.5); letter-spacing:2px; text-align:center; }
input {
    width: 100%;
    margin-bottom: 12px;
    background: #f5f5f5;
    border: none;
    outline: none;
    padding: 10px;
    font-size: 15px;
    color: rgb(9 0 0);
    text-shadow: 1px 1px 1px rgba(0,0,0,0.3);
    border: 1px solid rgba(0,0,0,0.3);
    border-radius: 4px;
    /* box-shadow: inset 0 -5px 45px rgba(100,100,100,0.2), 0 1px 1px rgba(255,255,255,0.2); */
    -webkit-transition: box-shadow .5s ease;
    -moz-transition: box-shadow .5s ease;
    -o-transition: box-shadow .5s ease;
    -ms-transition: box-shadow .5s ease;
    transition: box-shadow .5s ease;
}
input:focus { box-shadow: inset 0 -5px 45px rgba(100,100,100,0.4), 0 1px 1px rgba(255,255,255,0.2); }
```

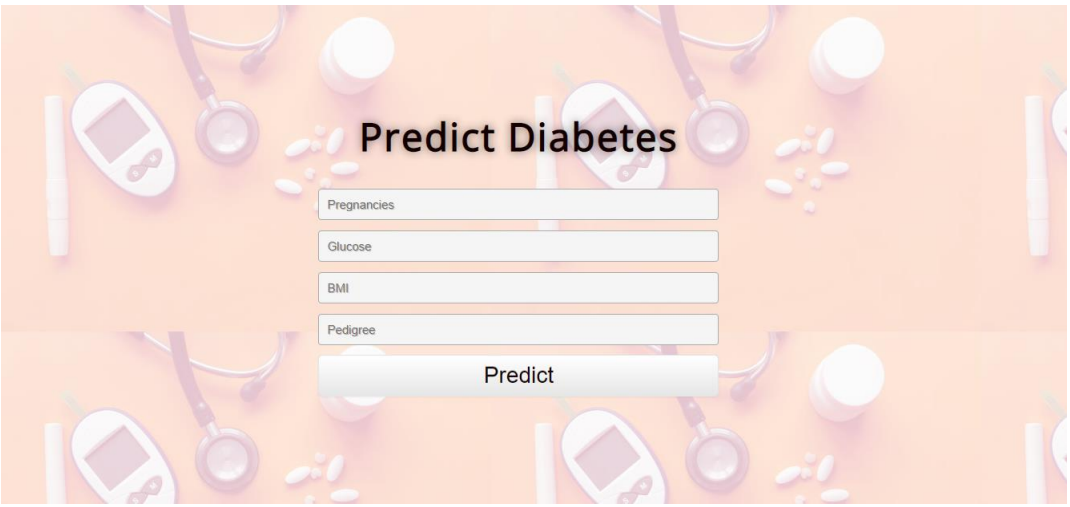
B. Running App.py

We can run the application either by double click and app.py or by run the command in terminal.

Running with debugger shows the following results.

```
 * Serving Flask app 'app' (lazy loading)
 * Environment: production
   WARNING: This is a development server. Do not use it in a production deployment.
   Use a production WSGI server instead.
 * Debug mode: on
 * Restarting with stat
 * Debugger is active!
 * Debugger PIN: 808-315-324
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```

- The output looks like this:



- We fill the input values, and the output will be like this.

6. Model deployment using Heroku

After building the machine leaning model and applicaiton using flask, we are ready to deploy the model using Heroku. Heroku is a cloud platform supporting several programming languages. One of the first cloud platforms and it support most of the well known programing languages.

To Deploy python code in Heroku. We need two main files: Requirements.txt/setup.txt and Profile.

- **requirements file**

it contains all the python packages required to execute the application.

➢ **Profile file**.

We need first to install gunicorn library then build the Profile contains the application file and programing language. Gunicorn is used to maintain the web application multiple instance executions, distributing incoming requests across those instances,and communicate with the web server.
App is the flask application file name. in our case is app.py, could be different name.

web: gunicorn app:app

Steps to deploy Ml model in Heroku:

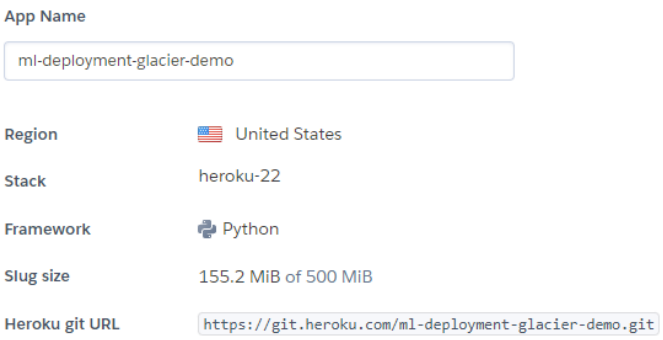
First step: We must create an account in Heroku and create an app as follow.
Second step: Connecting the Heroku app to Github repository containing the application files.
Third step: Choosing the language and then deploy the application. The output of this step as follows.

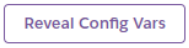
```
-----> Building on the Heroku-22 stack
-----> Determining which buildpack to use for this app
-----> Python app detected
-----> No Python version was specified. Using the buildpack default: python-3.10.7
       To use a different version, see: https://devcenter.heroku.com/articles/python-runtimes
-----> Installing python-3.10.7
-----> Installing pip 22.2.2, setuptools 63.4.3 and wheel 0.37.1
-----> Installing SQLite3
-----> Installing requirements with pip
       Collecting Flask==1.1.1
```

```
     Downloading Flask-1.1.1-py2.py3-none-any.whl (94 kB)
   Collecting gunicorn==19.9.0
     Downloading gunicorn-19.9.0-py2.py3-none-any.whl (112 kB)
   Collecting itsdangerous==1.1.0
     Downloading itsdangerous-1.1.0-py2.py3-none-any.whl (16 kB)
   Collecting Jinja2==2.10.1
     Downloading Jinja2-2.10.1-py2.py3-none-any.whl (124 kB)
   Collecting MarkupSafe==1.1.1
     Downloading MarkupSafe-1.1.1.tar.gz (19 kB)
     Preparing metadata (setup.py): started
     Preparing metadata (setup.py): finished with status 'done'
   Collecting Werkzeug==0.15.5
     Downloading Werkzeug-0.15.5-py2.py3-none-any.whl (328 kB)
   Collecting numpy>=1.9.2
     Downloading numpy-1.23.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.1
MB)
   Collecting scipy>=0.15.1
     Downloading scipy-1.9.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (43.9 MB)
   Collecting scikit-learn>=0.18
     Downloading scikit_learn-1.1.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(30.5 MB)
   Collecting matplotlib>=1.4.3
     Downloading matplotlib-3.6.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(11.8 MB)
   Collecting pandas>=0.19
     Downloading pandas-1.5.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.1
MB)
   Collecting click>=5.1
     Downloading click-8.1.3-py3-none-any.whl (96 kB)
   Collecting threadpoolctl>=2.0.0
     Downloading threadpoolctl-3.1.0-py3-none-any.whl (14 kB)
   Collecting joblib>=1.0.0
     Downloading joblib-1.2.0-py3-none-any.whl (297 kB)
   Collecting fonttools>=4.22.0
     Downloading fonttools-4.37.4-py3-none-any.whl (960 kB)
   Collecting contourpy>=1.0.1
     Downloading contourpy-1.0.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (295
kB)
   Collecting kiwisolver>=1.0.1
     Downloading kiwisolver-1.4.4-cp310-cp310-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (1.6
MB)
   Collecting pyparsing>=2.2.1
     Downloading pyparsing-3.0.9-py3-none-any.whl (98 kB)
   Collecting python-dateutil>=2.7
     Downloading python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
   Collecting cycler>=0.10
     Downloading cycler-0.11.0-py3-none-any.whl (6.4 kB)
   Collecting pillow>=6.2.0
```

```
         Downloading Pillow-9.2.0-cp310-cp310-manylinux_2_28_x86_64.whl (3.2 MB)
       Collecting packaging>=20.0
         Downloading packaging-21.3-py3-none-any.whl (40 kB)
       Collecting pytz>=2020.1
         Downloading pytz-2022.4-py2.py3-none-any.whl (500 kB)
       Collecting six>=1.5
         Downloading six-1.16.0-py2.py3-none-any.whl (11 kB)
       Building wheels for collected packages: MarkupSafe
         Building wheel for MarkupSafe (setup.py): started
         Building wheel for MarkupSafe (setup.py): finished with status 'done'
         Created wheel for MarkupSafe: filename=MarkupSafe-1.1.1-cp310-cp310-linux_x86_64.whl
size=17161 sha256=5489f15cc59c53fc957dfc5536c588cb6f4863b1cbab52f9646cd4a6de3922ff
         Stored in directory: /tmp/pip-ephem-wheel-cache-
qog_ykty/wheels/a6/81/81/3fcafa7c24e4b4e25bcf383c792b343e53c38e6196f44bc3e3
       Successfully built MarkupSafe
       Installing collected packages: pytz, Werkzeug, threadpoolctl, six, pyparsing, pillow, numpy,
MarkupSafe, kiwisolver, joblib, itsdangerous, gunicorn, fonttools, cycler, click, scipy, python-
dateutil, packaging, Jinja2, contourpy, scikit-learn, pandas, matplotlib, Flask
       Successfully installed Flask-1.1.1 Jinja2-2.10.1 MarkupSafe-1.1.1 Werkzeug-0.15.5 click-8.1.3
contourpy-1.0.5 cycler-0.11.0 fonttools-4.37.4 gunicorn-19.9.0 itsdangerous-1.1.0 joblib-1.2.0
kiwisolver-1.4.4 matplotlib-3.6.0 numpy-1.23.3 packaging-21.3 pandas-1.5.0 pillow-9.2.0 pyparsing-
3.0.9 python-dateutil-2.8.2 pytz-2022.4 scikit-learn-1.1.2 scipy-1.9.1 six-1.16.0 threadpoolctl-3.1.0
-----> Discovering process types
       Procfile declares types -> web
-----> Compressing...
       Done: 155.2M
-----> Launching...
       Released v3
       https://ml-deployment-glacier-demo.herokuapp.com/ deployed to Heroku
Starting November 28th, 2022, free Heroku Dynos, free Heroku Postgres, and free Heroku Data for
Redis® will no longer be available.
If you have apps using any of these resources, you must upgrade to paid plans by this date to ensure
your apps continue to run and to retain your data. For students, we will announce a new program by
the end of September. Learn more at https://blog.heroku.com/next-chapter
```

The application information and setting:





The application can be found at

https://ml-deployment-glacier-demo.herokuapp.com/