Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

Repository: https://github.com/farischaudhry/heroku-demo

# Toy Data Set, Model Training, Flask Code (From Week 4)

Diabetes data set (https://scikit-learn.org/stable/datasets/toy_dataset.html)

442 samples; 10 features (described below) along with target value.

Features are scaled by mean and std deviation; no null entries or wrong data.

| | |
|---|---|
| **Number of Instances:** | 442 |
| **Number of Attributes:** | First 10 columns are numeric predictive values |
| **Target:** | Column 11 is a quantitative measure of disease progression one year after baseline |
| **Attribute Information:** | • age age in years<br>• sex<br>• bmi body mass index<br>• bp average blood pressure<br>• s1 tc, total serum cholesterol<br>• s2 ldl, low-density lipoproteins<br>• s3 hdl, high-density lipoproteins<br>• s4 tch, total cholesterol / HDL<br>• s5 ltg, possibly log of serum triglycerides level<br>• s6 glu, blood sugar level |

The MEANS Procedure

| Variable | N | Mean | Std Dev |
|---|---|---|---|
| age | 442 | 48.5180995 | 13.1090278 |
| sex | 442 | 1.4683258 | 0.4995612 |
| bmi | 442 | 26.3757919 | 4.4181216 |
| bp | 442 | 94.6470136 | 13.8312834 |
| s1 | 442 | 189.1402715 | 34.6080517 |
| s2 | 442 | 115.4391403 | 30.4130810 |
| s3 | 442 | 49.7884615 | 12.9342022 |
| s4 | 442 | 4.0702489 | 1.2904499 |
| s5 | 442 | 4.6414109 | 0.5223906 |
| s6 | 442 | 91.2601810 | 11.4963347 |
| y | 442 | 152.1334842 | 77.0930045 |

Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

Standard training parameters used.

70% used as training data, 30% as testing data.

Model saved using pickle.

Saved as ‘model.pkl’

```
import pandas as pd
import pickle
from sklearn.datasets import load_diabetes
from sklearn.model_selection import train_test_split
from xgboost import XGBClassifier
from sklearn.preprocessing import LabelEncoder



X, y = load_diabetes(return_X_y=True, as_frame=True)
X.head()

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)
model = XGBClassifier(random_state=0)
le = LabelEncoder()
y_train = le.fit_transform(y_train)
model.fit(X_train, y_train)

XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0, gpu_id=-1,
              importance_type='gain', interaction_constraints='',
              learning_rate=0.300000012, max_delta_step=0, max_depth=6,
              min_child_weight=1, monotone_constraints='()',
              n_estimators=100, n_jobs=0, num_parallel_tree=1,
              objective='multi:softprob', random_state=0, reg_alpha=0,
              reg_lambda=1, scale_pos_weight=None, subsample=1,
              tree_method='exact', validate_parameters=1, verbosity=None)

print(X_train.head())

pickle.dump(model, open('.\model.pkl', 'wb'))
```

Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

app.py (scale factors are from the MEANS approach above)

Port restriction had to be removed in app.run().

```python
from flask import Flask, request, render_template
import numpy as np
import pandas as pd
import math
import pickle

scale_factor = math.sqrt(442)
app = Flask(__name__)
model = pickle.load(open('./model.pkl', 'rb'))

# home endpoint
@app.route('/')
def home():
    return render_template('index.html')

# prediction endpoint
@app.route('/predict', methods=['POST'])
def predict():
    int_features = [float(x) for x in request.form.values()]
    final_features = [np.array(int_features)]

    # each feature has to be scaled by some specific variable found here:
    # https://www4.stat.ncsu.edu/~boos/var.select/diabetes.read.tab.out.txt
    df = pd.DataFrame({ 'age': (final_features[0][0] -  48.5180995) / ( 13.1090278 * scale_factor),
                        'sex': (final_features[0][1] -  1.4683258) / (0.4995612 * scale_factor),
                        'bmi': (final_features[0][2] - 26.3757919) / (4.4181216 * scale_factor),
                        'bp': (final_features[0][3] - 94.6470136) / (13.8312834 * scale_factor),
                        's1': (final_features[0][4] - 189.1402715) / (34.6080517 * scale_factor),
                        's2': (final_features[0][5] - 115.4391403) / (30.4130810 * scale_factor),
                        's3': (final_features[0][6] - 49.7884615) / (12.9342022 * scale_factor),
                        's4': (final_features[0][7] - 4.0702489) / (1.2904499 * scale_factor),
                        's5': (final_features[0][8] - 4.6414109) / (0.5223906 * scale_factor),
                        's6': (final_features[0][9] - 152.1334842) / (77.093004 * scale_factor)},
                       index=[0])

    prediction = model.predict(df)
    return render_template('index.html', prediction_text='Regression value is {}'.format(prediction))

if __name__ == '__main__':
    app.run()
```

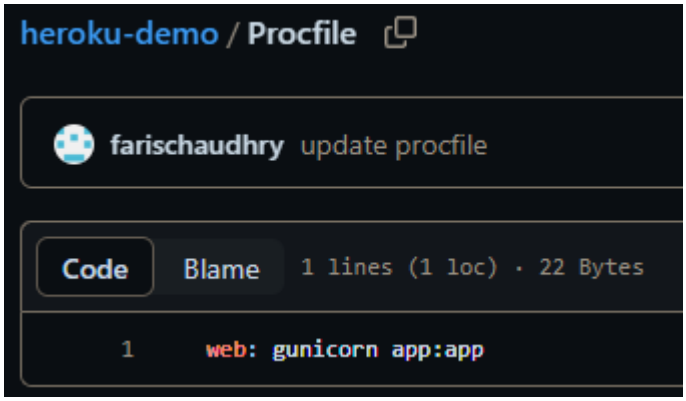
Index.html and style.css modified from https://www.w3docs.com/learn-html/html-form-templates.html (free to copy and use)

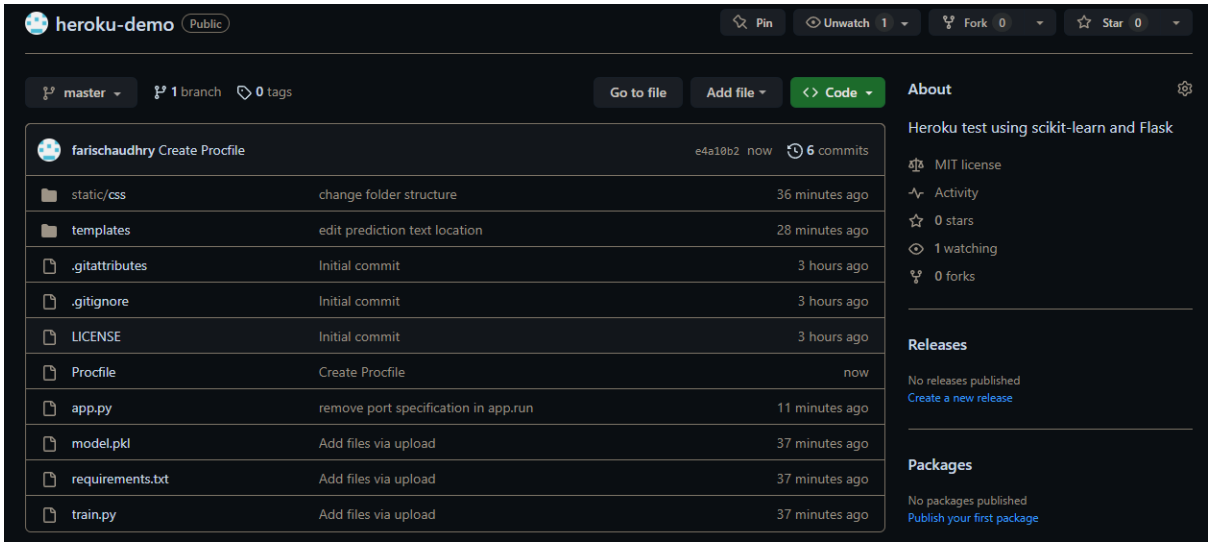
CSS: https://github.com/farischaudhry/heroku-demo/blob/master/static/css/style.css

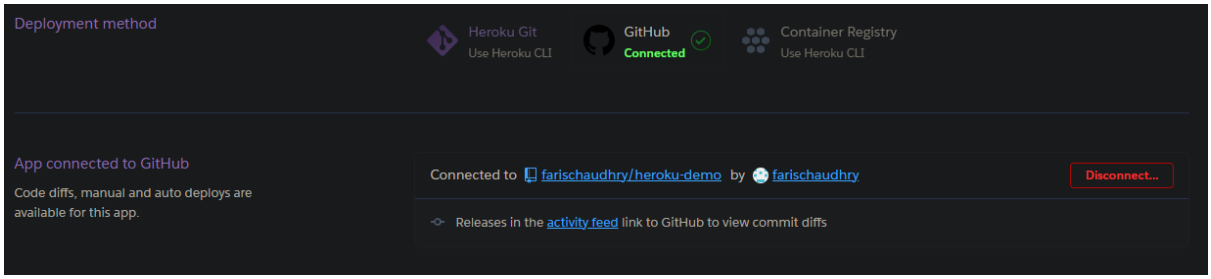
HTML: https://github.com/farischaudhry/heroku-demo/blob/master/templates/index.html

Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

# Heroku Deployment



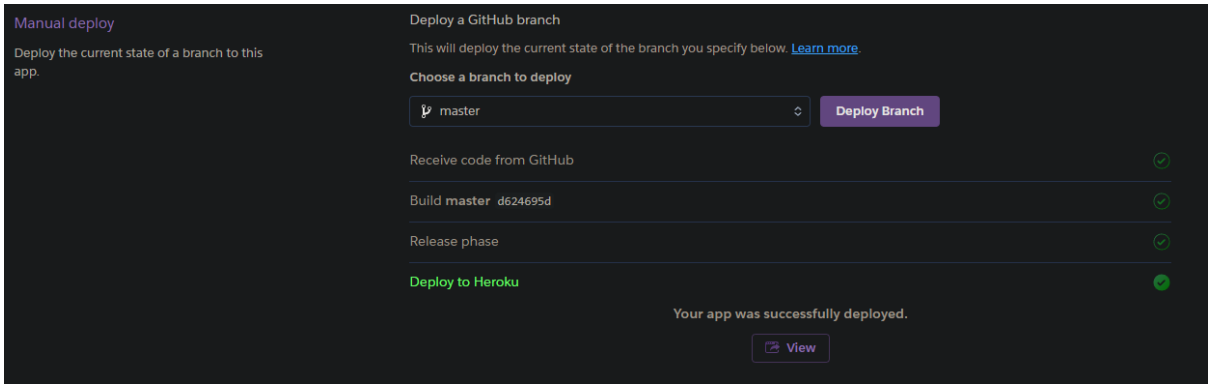
Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

heroku-demo / requirements.txt

farischaudhry add xgboost to reqs

Code | Blame | 12 lines (12 loc) · 198 Bytes | Code 55% faster with GitHub Copilot

```
1   Flask>2.2.2
2   gunicorn==19.9.0
3   itsdangerous>=2.1.2
4   Jinja2>=3.0
5   MarkupSafe>=2.0
6   Werkzeug>=2.2.2
7   numpy>=1.23.5
8   scipy>=1.11.2
9   scikit-learn>=0.18
10  matplotlib>=1.4.3
11  pandas>=0.19
12  xgboost>=1.7.6
```




```
PS C:\Users\faris\Documents\projects\heroku-demo> heroku ps:scale web=1 --app heroku-demo-app2
 »   Warning: heroku update available from 7.53.0 to 8.4.1.
Scaling dynos... done, now running web at 1:Basic
```

```
2023-08-30T15:22:58.993202+00:00 app[api]: Release v15 created by user faris.chaudhry@outlook.com
2023-08-30T15:23:12.318962+00:00 heroku[web.1]: Starting process with command `gunicorn app:app`
2023-08-30T15:23:12.977318+00:00 app[web.1]: [2023-08-30 15:23:12 +0000] [2] [INFO] Starting gunicorn 19.9.0
2023-08-30T15:23:12.977544+00:00 app[web.1]: [2023-08-30 15:23:12 +0000] [2] [INFO] Listening at: http://0.0.0.0:59750 (2)
2023-08-30T15:23:12.977578+00:00 app[web.1]: [2023-08-30 15:23:12 +0000] [2] [INFO] Using worker: sync
2023-08-30T15:23:12.977799+00:00 app[web.1]: <frozen os>:1031: RuntimeWarning: line buffering (buffering=1) isn't supported in binary mode, the default buffer size will be used
2023-08-30T15:23:12.979355+00:00 app[web.1]: [2023-08-30 15:23:12 +0000] [7] [INFO] Booting worker with pid: 7
2023-08-30T15:23:13.069291+00:00 app[web.1]: [2023-08-30 15:23:13 +0000] [8] [INFO] Booting worker with pid: 8
2023-08-30T15:23:13.555074+00:00 heroku[web.1]: State changed from starting to up
2023-08-30T15:23:40.000000+00:00 app[api]: Build succeeded
```

Name: Faris Chaudhry
Batch Code: LISUM25
Submission Date: 08/31/23

# Final Site