

Model Deployment Documentation

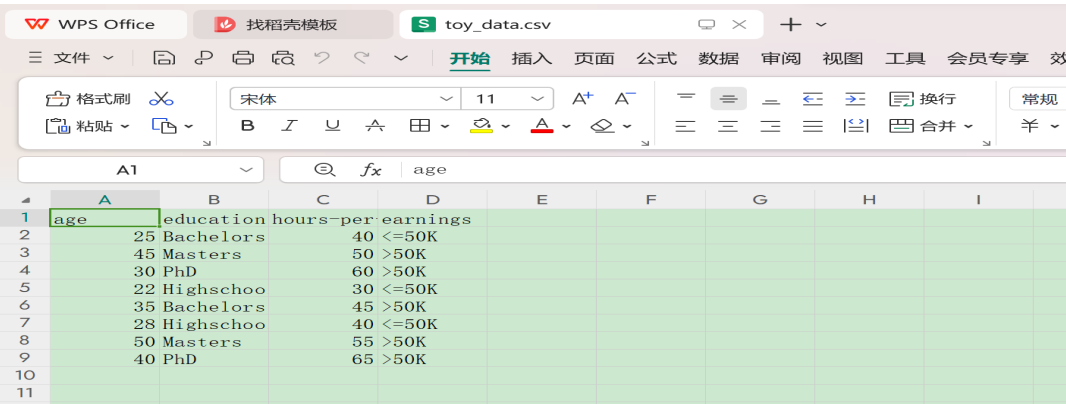
Name: Haipei Xu

Batch Code: LISUM35

Submission Date: 2024-08-21

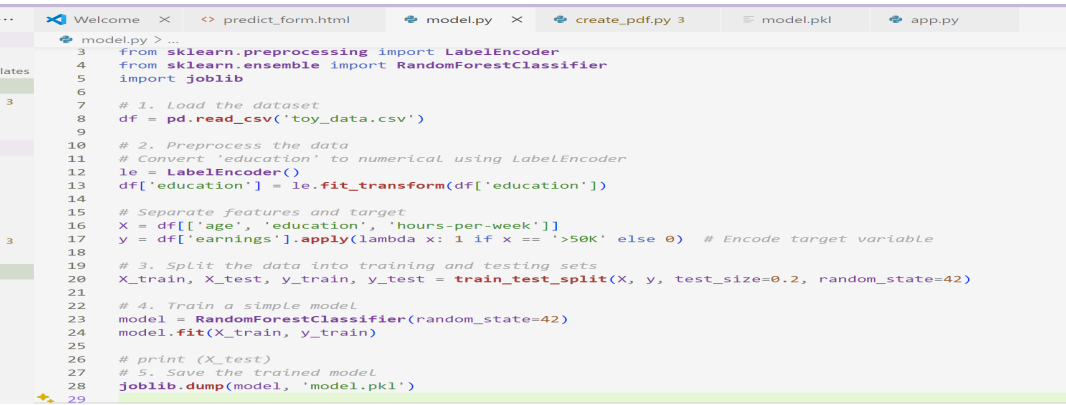
Submitted To: Data_Glacier

Snapshot of Data Preparation:



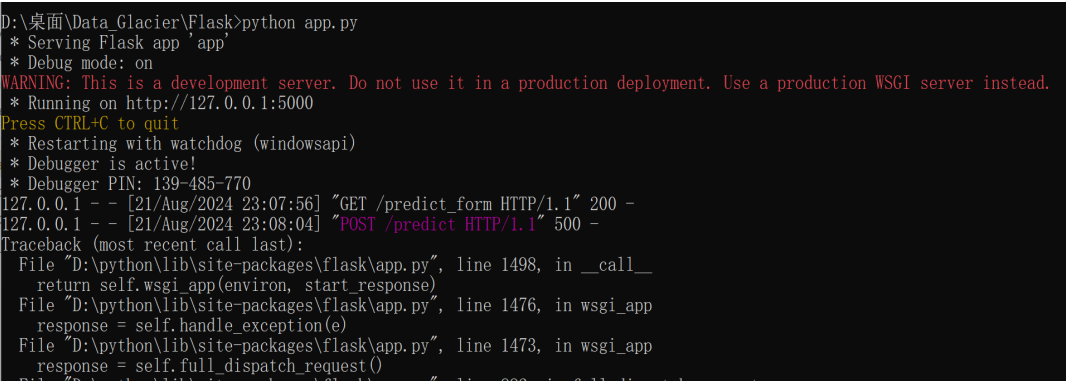
	A	B	C	D	E	F	G	H	I
1	age	education	hours-per-week	earnings					
2	25	Bachelors	40	<=50K					
3	45	Masters	50	>50K					
4	30	PhD	60	>50K					
5	22	Highschool	30	<=50K					
6	35	Bachelors	45	>50K					
7	28	Highschool	40	<=50K					
8	50	Masters	55	>50K					
9	40	PhD	65	>50K					
10									
11									

Snapshot of Model Training:



```
... Welcome X predict_form.html model.py X create_pdf.py 3 model.pkl app.py
model.py > ...
3 from sklearn.preprocessing import LabelEncoder
4 from sklearn.ensemble import RandomForestClassifier
5 import joblib
6
7 # 1. Load the dataset
8 df = pd.read_csv('toy_data.csv')
9
10 # 2. Preprocess the data
11 # Convert 'education' to numerical using LabelEncoder
12 le = LabelEncoder()
13 df['education'] = le.fit_transform(df['education'])
14
15 # Separate features and target
16 X = df[['age', 'education', 'hours-per-week']]
17 y = df['earnings'].apply(lambda x: 1 if x == '>50K' else 0) # Encode target variable
18
19 # 3. Split the data into training and testing sets
20 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
21
22 # 4. Train a simple model
23 model = RandomForestClassifier(random_state=42)
24 model.fit(X_train, y_train)
25
26 # print (X_test)
27 # 5. Save the trained model
28 joblib.dump(model, 'model.pkl')
29
```

Snapshot of Flask Deployment:



```
D:\桌面\Data Glacier\Flask>python app.py
* Serving Flask app 'app'
* Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on http://127.0.0.1:5000
Press CTRL+C to quit
* Restarting with watchdog (windowsapi)
* Debugger is active!
* Debugger PIN: 139-485-770
127.0.0.1 - - [21/Aug/2024 23:07:56] "GET /predict_form HTTP/1.1" 200 -
127.0.0.1 - - [21/Aug/2024 23:08:04] "POST /predict HTTP/1.1" 500 -
Traceback (most recent call last):
  File "D:\python\lib\site-packages\flask\app.py", line 1498, in __call__
    return self.wsgi_app(environ, start_response)
  File "D:\python\lib\site-packages\flask\app.py", line 1476, in wsgi_app
    response = self.handle_exception(e)
  File "D:\python\lib\site-packages\flask\app.py", line 1473, in wsgi_app
    response = self.full_dispatch_request()
  File "D:\python\lib\site-packages\flask\app.py", line 1525, in full_dispatch_request
    rv = self.handle_user_exception(e)
  File "D:\python\lib\site-packages\flask\app.py", line 1522, in full_dispatch_request
    rv = self.ensure_sync(self.dispatch_request)()
  File "D:\python\lib\site-packages\flask\app.py", line 1593, in dispatch_request
    return self.ensure_sync(self.view_functions[rule.endpoint])(**view_args)
  File "D:\桌面\Data Glacier\Flask\app.py", line 10, in <lambda>
    return predict_form(**kwargs)
  File "D:\桌面\Data Glacier\Flask\app.py", line 10, in predict_form
    model = joblib.load('model.pkl')
  File "D:\python\lib\site-packages\joblib\load.py", line 122, in load
    with open(filename, 'rb') as f:
FileNotFoundError: [Errno 2] No such file or directory: 'model.pkl'
```

Snapshot of Webpage:

← → ↻ 🏠 ⓘ 127.0.0.1:5000/predict_form

Enter Data for Prediction

Age:

Education:

Hours per week:

← ↻ 🏠 ⓘ 127.0.0.1:5000/predict

```
1 {  
2   "earnings": "<=50K"  
3 }
```