Data Glacier

Your Deep Learning Partner

Week 4: Deployment on Flask

Name: Kseniya Kudzelich

Batch Code: LISUM34

Submission Date: 27 June 2024

Submitted to: Data Glacier

Introduction:

I would like to remind you that during the second and third weeks of the "G2M insight for Cab Investment firm" project, I conducted an in-depth Exploratory Data Analysis (EDA) of the data, tested hypotheses, and also developed recommendations on choosing a company to invest in, and then presented the results.

This week, based on this data, I trained a Machine Learning model to predict Profit per Trip, taking into account various features (like "City", "Payment Mode", "Gender", "KM Travelled", "Date", etc.), and then deployed this model using the Flask framework.

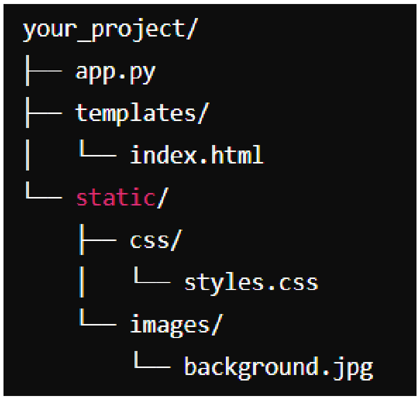
Stages of model deployment:

1. Prepare the model:

- Preprocess the previously available data (*'df.csv'*): train and test splitting, scaling, encoding.
- Train the models: Linear Regression, Ridge, Lasso, Random Forest, XGBoost, XGBoost with GridSearchCV.
- Choose the best model and save it together with the preprocessor using ***pickle.***

2. Create a directory structure for the project:

- Organize the project files into a directory structure that includes a Flask script, HTML template, static resources (CSS, images), a model, and a preprocessor.



```
your_project/
├── app.py
├── templates/
│   └── index.html
└── static/
    ├── css/
    │   └── styles.css
    └── images/
        └── background.jpg
```

3. Create a Flask app:

- Write a Flask script ('*app.py*') that loads the model and the preprocessor, extracts data from the from, processes the input data from the web form, makes predictions and displays the results on the web page.

4. Create an HTML template:

- Create an template ('*index.html*') for a web form that allows users to enter data for prediction. You need to select an option from the list for the "City", "Payment Method", "Gender" and "Company". All other elements of the form should be entered manually. After submitting the form, the template displays the predicted result.

5. Create CSS styles:

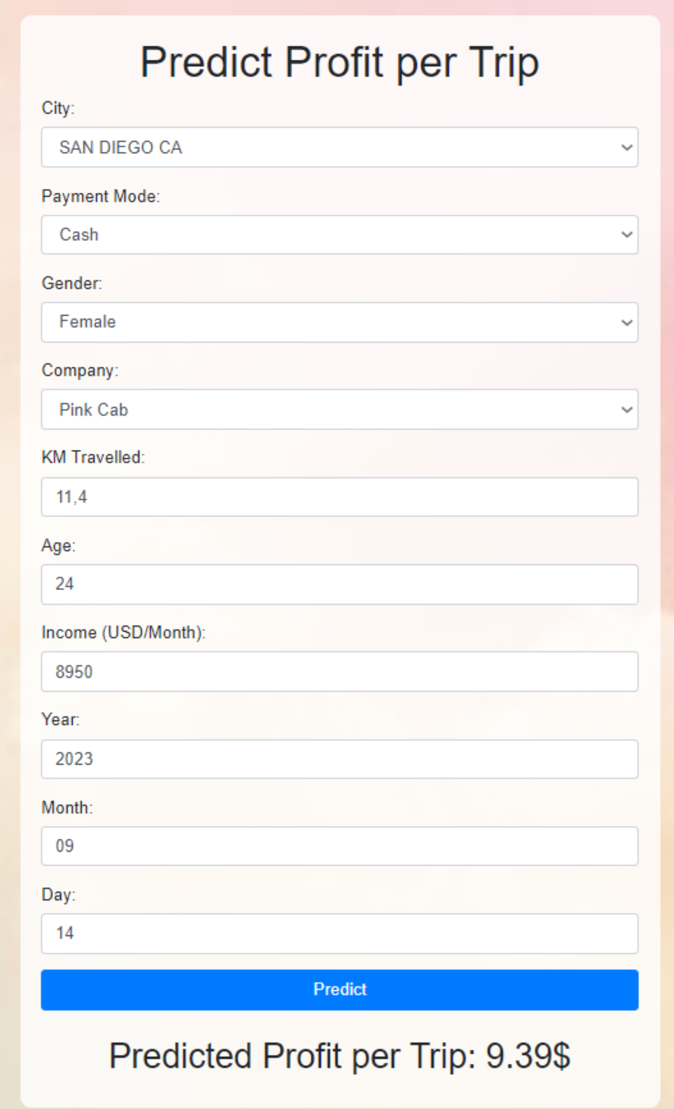
- Add styles to enhance the appearance of the web page, including the design of the form and background, to create a more user-friendly and visually appealing interface.

6. Run app:

- Launch the Flask app and open it in a web browser on a local server http://127.0.0.1:5000 to test the data entry and predictions, ensuring that all elements are working correctly.

Below you can see 6 examples of using my deployed Flask app:

Example 1:



Predict Profit per Trip

City: SAN DIEGO CA

Payment Mode: Cash

Gender: Female

Company: Pink Cab

KM Travelled: 11,4

Age: 24

Income (USD/Month): 8950

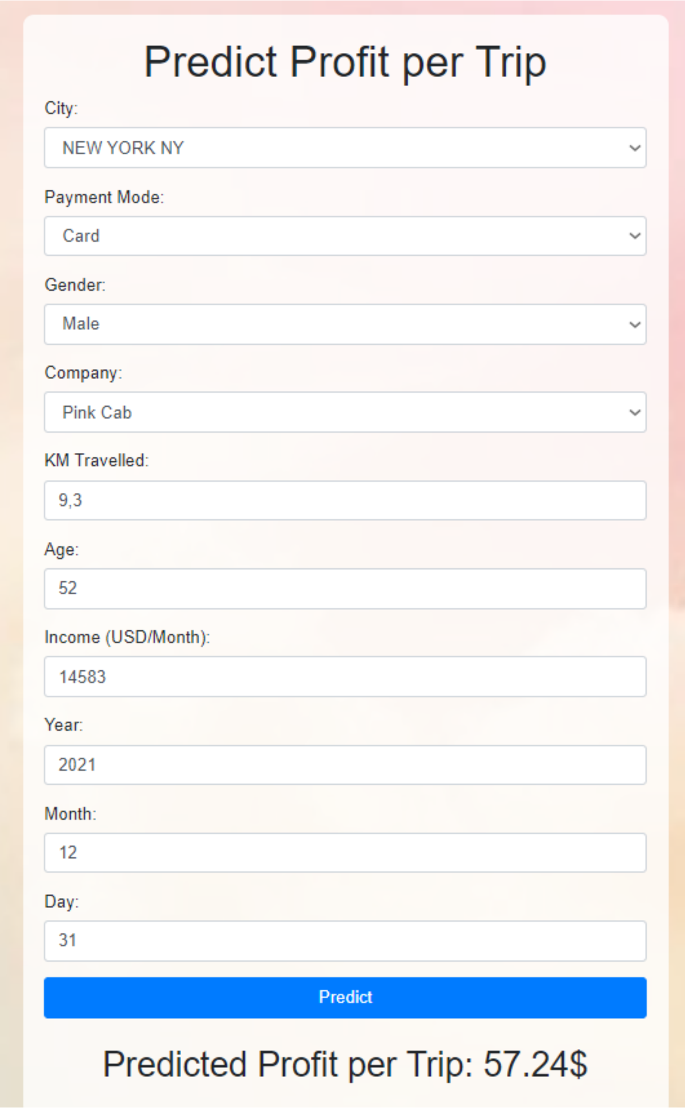
Year: 2023

Month: 09

Day: 14

Predicted Profit per Trip: 9.39$

Example 2:



Predict Profit per Trip

City: NEW YORK NY

Payment Mode: Card

Gender: Male

Company: Pink Cab

KM Travelled: 9,3

Age: 52

Income (USD/Month): 14583

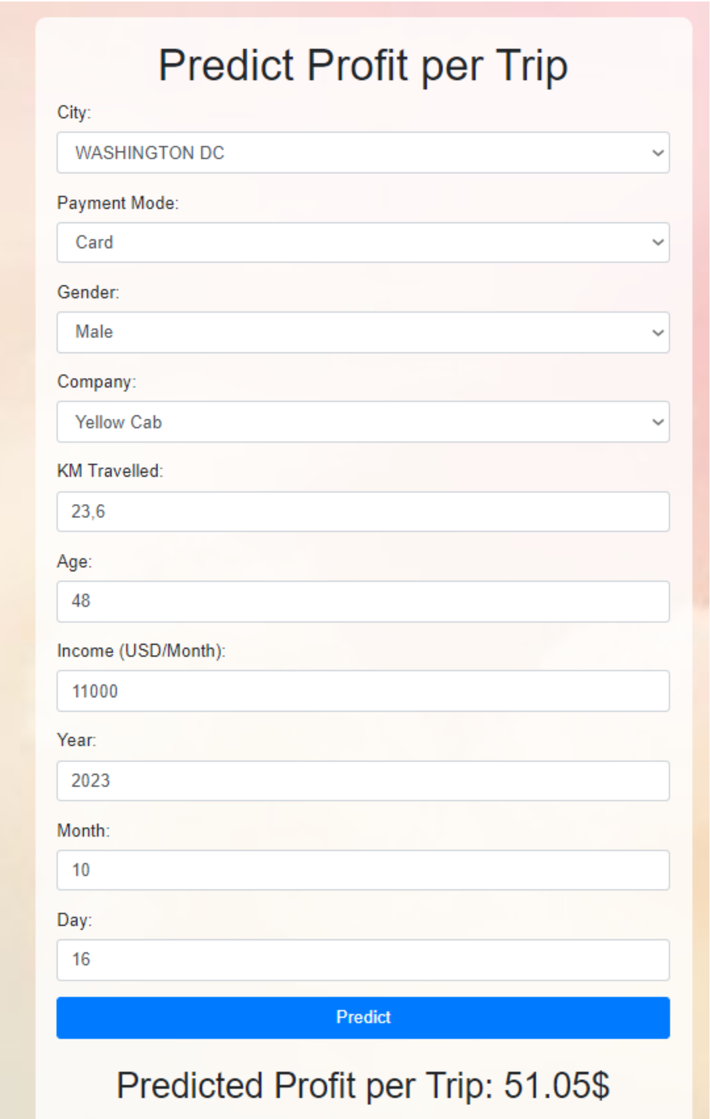
Year: 2021

Month: 12

Day: 31

Predicted Profit per Trip: 57.24$

Example 3:



Predict Profit per Trip

City: WASHINGTON DC

Payment Mode: Card

Gender: Male

Company: Yellow Cab

KM Travelled: 23,6

Age: 48

Income (USD/Month): 11000

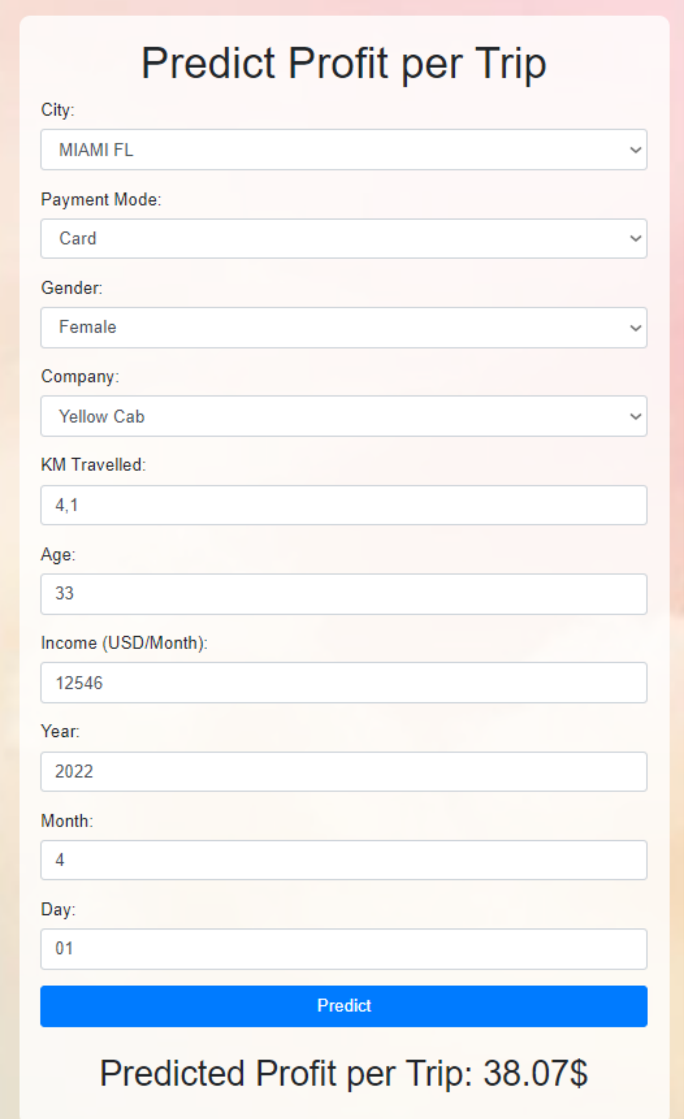
Year: 2023

Month: 10

Day: 16

Predicted Profit per Trip: 51.05$

Example 4:



Predict Profit per Trip

City: MIAMI FL

Payment Mode: Card

Gender: Female

Company: Yellow Cab

KM Travelled: 4,1

Age: 33

Income (USD/Month): 12546

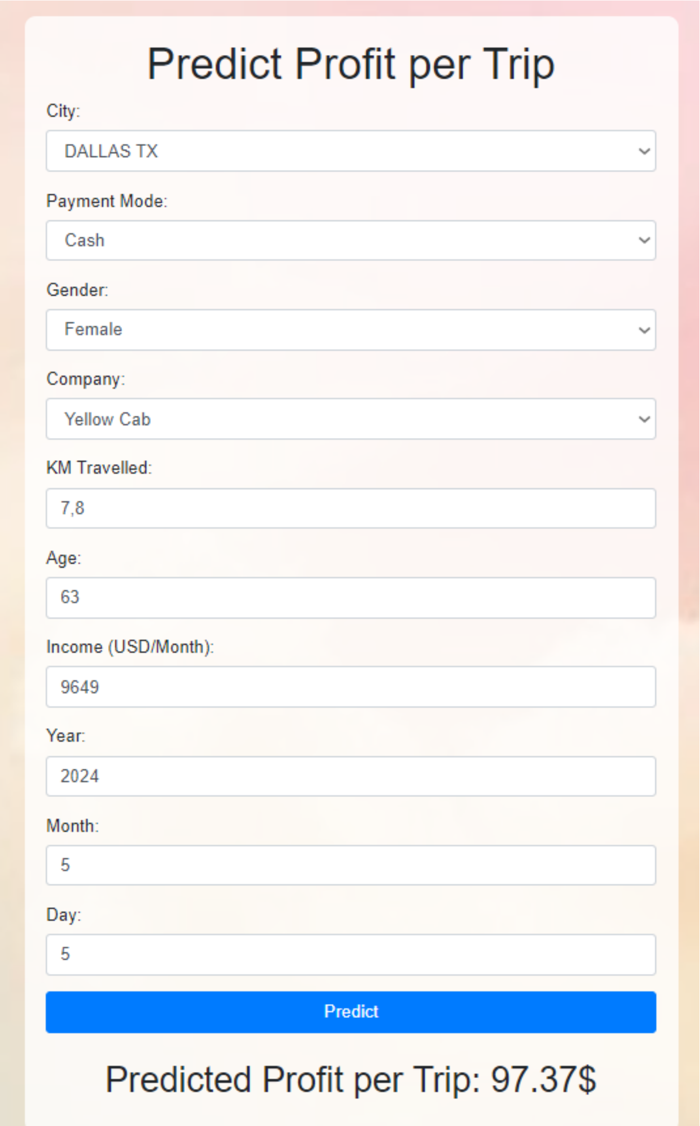
Year: 2022

Month: 4

Day: 01

Predicted Profit per Trip: 38.07$

Example 5:



Predict Profit per Trip

City: DALLAS TX

Payment Mode: Cash

Gender: Female

Company: Yellow Cab

KM Travelled: 7,8

Age: 63

Income (USD/Month): 9649

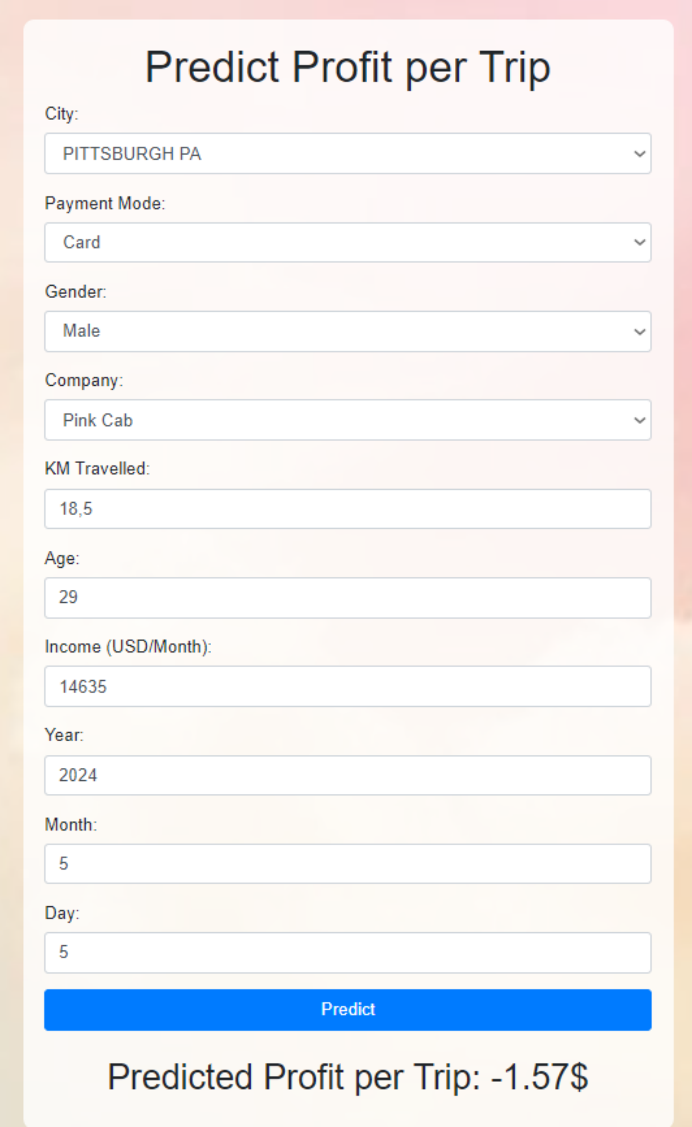
Year: 2024

Month: 5

Day: 5

Predicted Profit per Trip: 97.37$

Example 6:



Predict Profit per Trip

City: PITTSBURGH PA

Payment Mode: Card

Gender: Male

Company: Pink Cab

KM Travelled: 18,5

Age: 29

Income (USD/Month): 14635

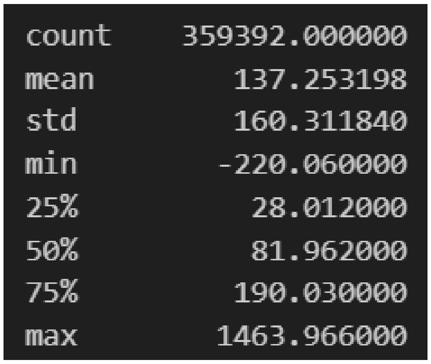
Year: 2024

Month: 5

Day: 5

Predicted Profit per Trip: -1.57$

The results are acceptable, even those with a negative Profit per Trip, as the Profit column in the training data ranges from -220.06 to +1463.96.



```
count    359392.000000
mean        137.253198
std         160.311840
min        -220.060000
25%          28.012000
50%          81.962000
75%         190.030000
max        1463.966000
```