RESTAURANT RATING PREDICTION

DETAILED PROJECT REPORT

BY:
HARSHITHA AV
SPOORTHI CHANDRA.P

ABSTRACT

The basic idea of analyzing the Zomato dataset is to get a fair idea about the factors affecting the establishment of different types of restaurants at different places in Bengaluru, aggregate rating of each restaurant.

Bengaluru being one such city has more than 50,000 restaurants with restaurants serving dishes from all over the world. With each day new restaurants opening the industry hasn't been saturated yet and the demand is increasing day by day. Bengaluru being an IT capital of India, most of the people here are dependent mainly on the restaurant food as they don’t have time to cook for themselves. With such an overwhelming demand for new restaurants, it has become important to study the ratings of restaurants.

Introduction

Why this HLD Document?

The main purpose of this HLD document is to feature the required details of the project and supply the outline of the Model Creation, Evaluation and Deployment.This additionally provides the careful description on however the complete project has been designed end-to-end.

The HLD will:

- Present of the design aspects and define them in detail.
- Describe the user interface being implemented.
- Describe the hardware and software interfaces.
- Describe the performance requirements.

•Include design features and architectural design of the project.
• List and describe the non - functional attributes like:
o Security
o Reliability
o Maintainability
o Portability
o Reusability
o Resource Utilization.

Scope :
The HLD documentation presents the structure of the system, such as database design, architectural design, application flow and technology architecture. The HLD uses non-technical terms to technical terms that can be understandable to the administrator of the system.

DEFINITIONS

Term	Description
FFP	Restaurant Rating Prediction
Database	Collection of all the information used by the System.
Jupyter-Notebook	It is an interactive computational environment, in which you can combine code execution, rich text, mathematics, plots and rich media.

General Description

Problem Perspective: The Restaurant Rating Prediction may be a machine learning model that helps users to predict the rating of the restaurant and help them to understand about the price, quality, location etc. about the restaurant.

Problem Statement :

The main goal of this project is to perform extensive Exploratory Data Analysis (EDA) on the Zomato Dataset and build an appropriate Machine Learning Model that will help various Zomato Restaurants to predict their respective Ratings based on certain features.

Proposed Solution :

To solve the problem, we have created a User interface for taking the input from the user to predict the Restaurant Rating using our trained ML model after processing the input and at last the output (predicted value) from the model is communicated to the User.

Further Improvements :

We also analysis the data used for training the ML model by considering different occasions such as Weekday, Season or any Social reasons, considering different angle of business. If we method such information and predict the Restaurant Rating, it will bring some loss to the restaurant but user can get benefit from that.

Technical Requirements:

As technical requirements, we don't need any specialized hardware for virtualization of the application. The user should have the device that has the access to the web and the fundamental understanding of providing the input

Tools Used :

Python 3.9 is employed because the programming language and frameworks like NumPy, Pandas, Scikit - learn and alternative modules for building the model.
Jupyter-Notebook is employed as IDE.
For Data visualizations, seaborn and components of matplotlib are getting used. For information assortment prophetess info is getting used
Front end development is completed victimization HTML/CSS.
Flask is employed for each information and backend readying
GitHub is employed for version management.

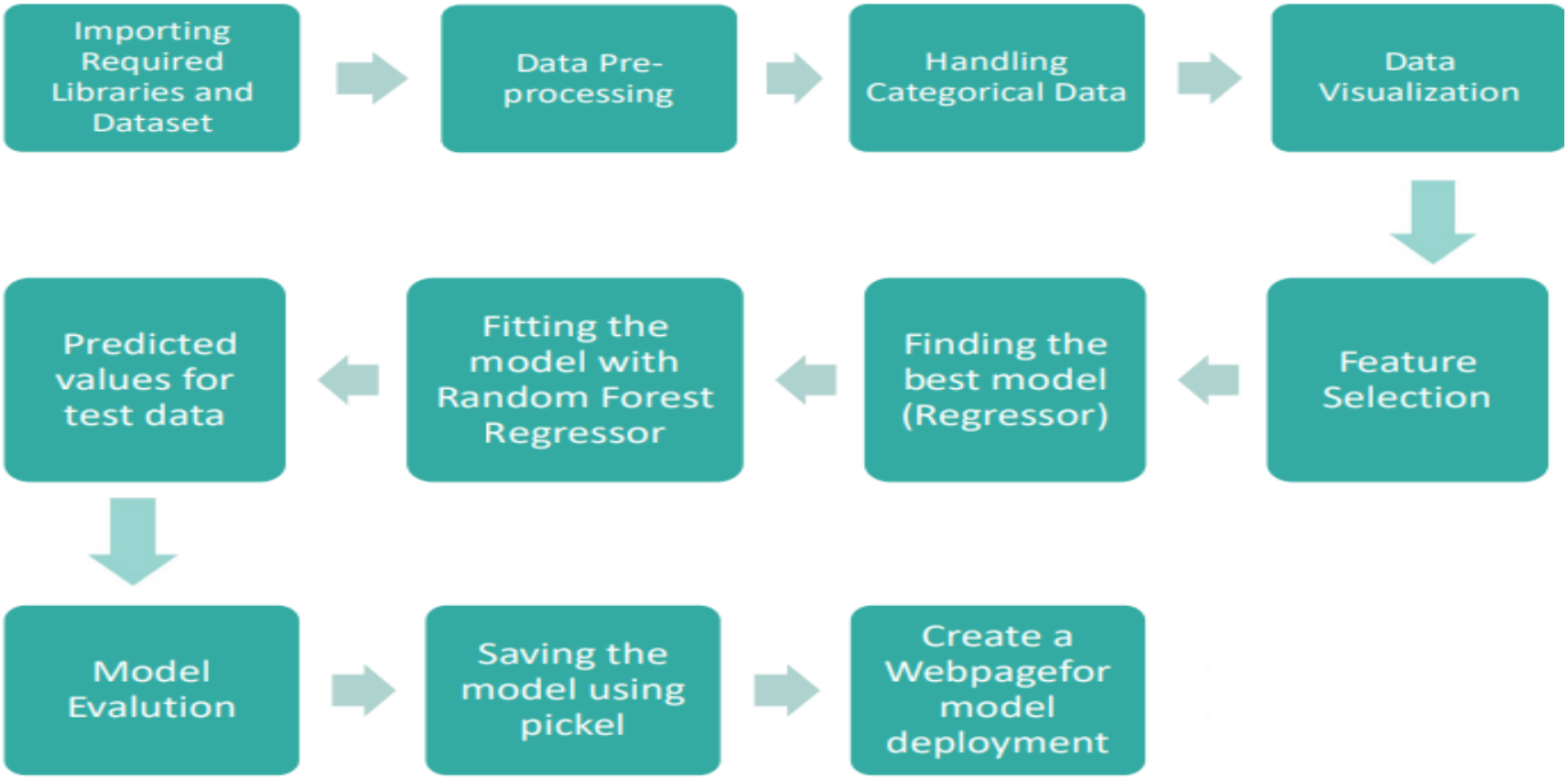
Constraints:

The Restaurant rating prediction answer should be user friendly, as automatic as attainable and also the user should not be needed to understand any of the operating

Assumptions:

The main objective of the project is to implement the utility cases as for the new dataset that provides the user the ability to predict Restaurant rating . Machine learning model is employed for process the user input for prediction. It additionally assumed that each one aspects of this project have the flexibility to figure along within the approach the designer is expecting.

DATA FLOW:



Performance Evaluation

Reusability:

The code written and the components used should have the ability to be reused with no problems.

Application Compatibility:

The different parts of the system are communicating or using Python as an interface between them. All the components have its own tasks to perform and it is a job of a Python to ensure proper transfer of data.

Resource Utilization:

When ant task is performed, it'll doubtless use all the process power offered till the process is finished.

Data Description

The dataset contains 17 variables all of which were scrapped from the Zomato website. The dataset contains details of more than 50,000 restaurants in Bengaluru in each of its neighborhood. The total size of dataset is approximately 547 MB.

url: contains the url of the restaurant in the zomato website

address: contains the address of the restaurant in Bengaluru

name: contains the name of the restaurant

online_order: whether online ordering is available in the restaurant or not

book_table: table book option available or not

rate: contains the overall rating of the restaurant out of 5

votes: contains total number of rating for the restaurant as of the above mentioned date.
phone: contains the phone number of the restaurant.
location: contains the neighborhood in which the restaurant is located.
rest_type: restaurant type.
dish_liked: dishes people liked in the restaurant.
cuisines: food styles, separated by comma.
approx_cost(for two people): contains the approximate cost for meal for two people.
reviews_list: list of tuples containing reviews for the restaurant, each tuple.
menu_item: contains list of menus available in the restaurant.
listed_in(type): type of meal.
listed_in(city): contains the neighborhood in which the restaurant is listed.

Importing dataset and libraries:

We have imported certain libraries such as numpy,panda,matplot and many more.

The ZOMATO dataset is imported in the form of csv.

Data Preprocessing

Checked for info of the Dataset, to verify the correct datatype of the Columns.

Checked for Null values, because the null values can affect the accuracy of the model.

Converted all the illegal values into legal values.

Checking the distribution of the columns to interpret its importance.

Prepared the relevant data from the dataset.

Now, the info is prepared to train a Machine Learning Model.

Modelling Creation

After preprocessing the data, we visualize our data to gain insights and then these insights are randomly spread and split into two parts, train and test data. After splitting the data, we use Random Forest Regressor to model our data to predict the Restaurant Rating.

UI Integration

Both CSS and HTML files are being created and are being integrated with the created machine learning model. All the required files are then integrated to the app.py file and tested locally.

Data from User

The data from the user is retrieved from the created HTML web page.

Data Validation

The data provided by the user is then being processed by app.py file and validated. The validated data is then sent to the prepared model for the prediction. Importing Required Libraries and Dataset Data Preprocessing Handling Categorical Data Data Visualization Predicted values for test data Fitting the model with Random Forest Regressor Finding the best model (Regressor) Feature Selection Model Evalution Saving the model using pickel Create a Webpagefor model deployment.

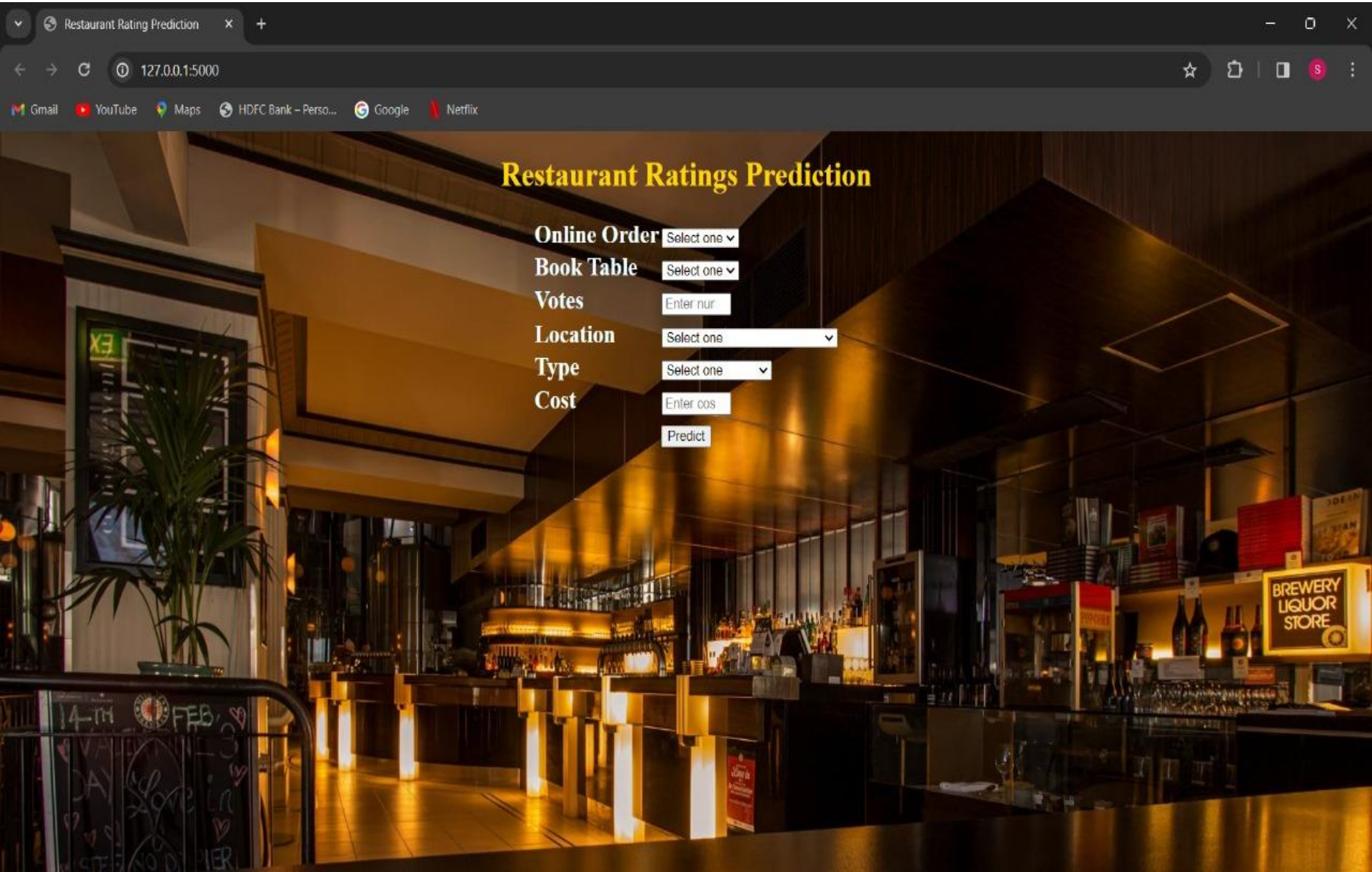
Rendering the Results

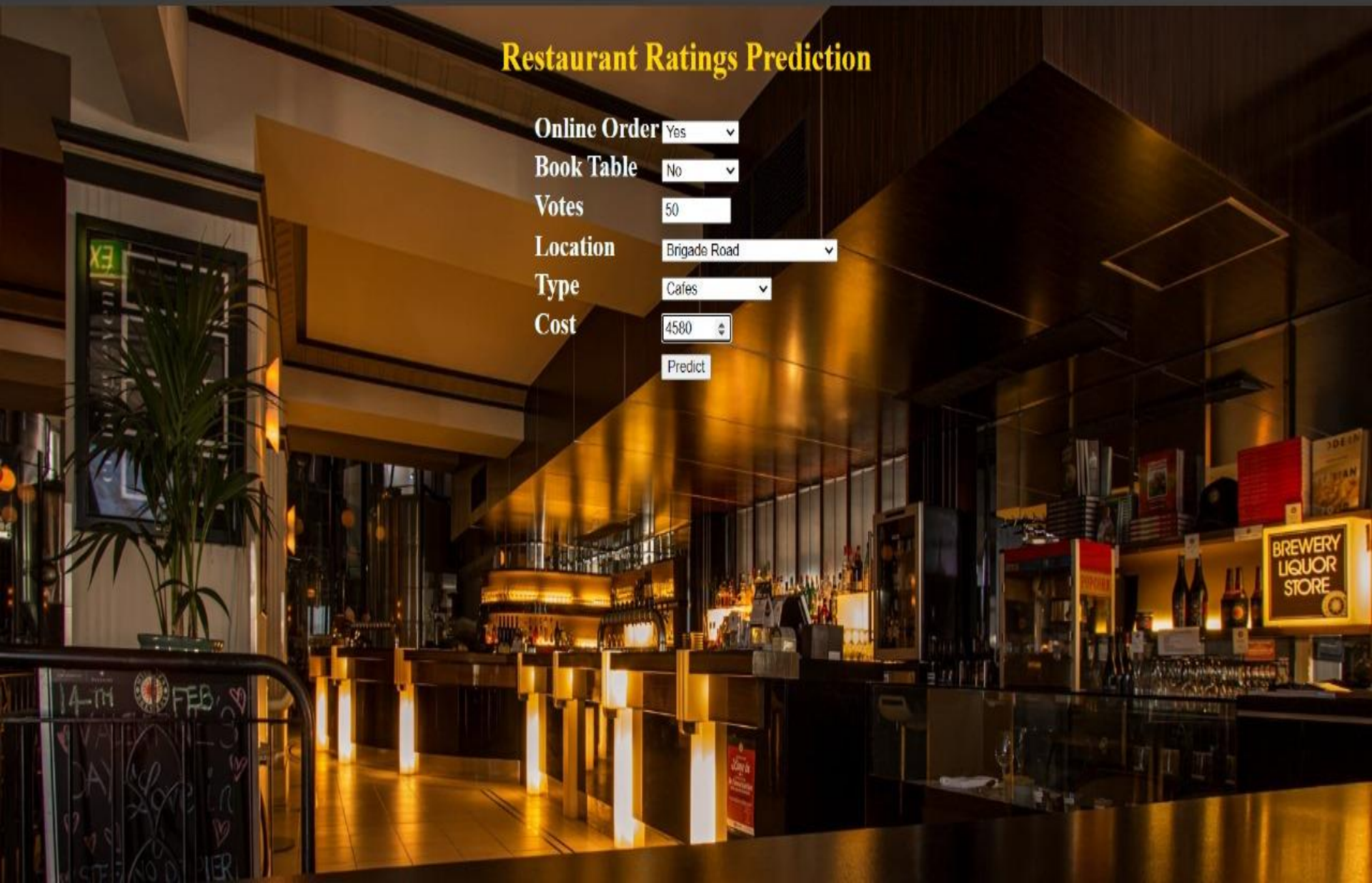
The data sent for the prediction is then rendered to the web page.

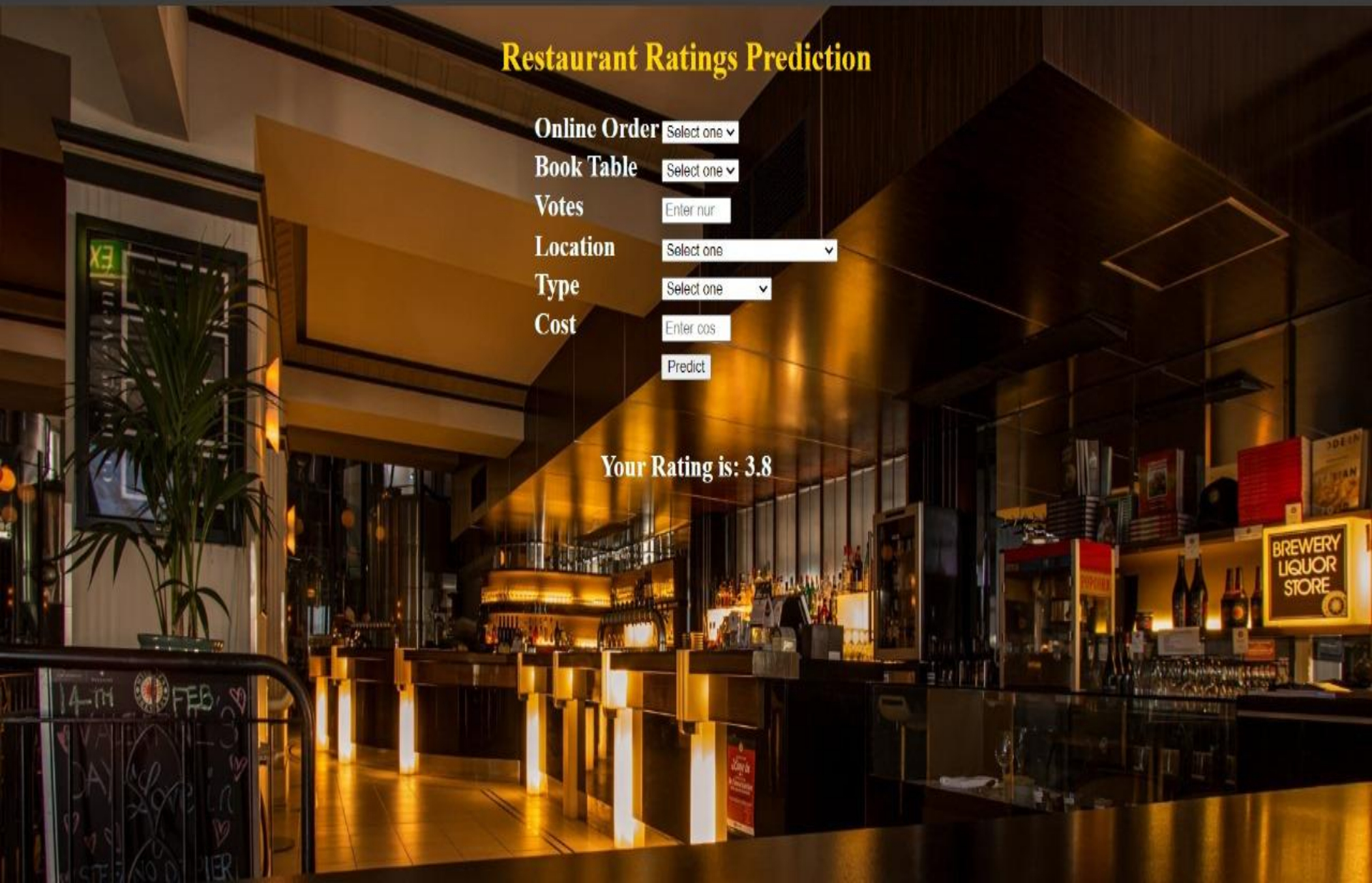
Deployment

The tested model is then deployed to Heroku. So, users can access the project from any internet devices.

DEMONSTRATION:




Restaurant Rating Prediction
127.0.0.1:5000
Gmail
YouTube
Maps
HDFC Bank – Perso...
Google
Netflix
Restaurant Ratings Prediction
Online Order
Yes
Book Table
No
Votes
50
Location
Brigade Road
Type
Cafes
Cost
4580
Predict
BREWERY LIQUOR STORE


Restaurant Rating Prediction
127.0.0.1:5000/predict
Gmail
YouTube
Maps
HDFC Bank – Perso...
Google
Netflix
Restaurant Ratings Prediction
Online Order
Select one
Book Table
Select one
Votes
Enter nur
Location
Select one
Type
Select one
Cost
Enter cos
Predict
Your Rating is: 3.8
BREWERY LIQUOR STORE

ACKNOWLEDGMENT

I would like to extend my sincere appreciation to all those who played a crucial role in the development and completion of the Restaurant Rating Machine Learning Project.

Special thanks to the members of my project team for their hard work, dedication, and collaborative spirit. Each member brought a unique set of skills and perspectives, contributing significantly to the success of this endeavor.

Last but not least, I extend my heartfelt thanks to my friends and family for their unwavering support and encouragement throughout this challenging yet rewarding process.

This project has been a truly collaborative effort, and I am fortunate to have had such a talented and supportive team. Thank you to everyone involved for your contributions and commitment to excellence.

Conclusion:

The Restaurant Rating Prediction system will predict the rating for helping the customers with the trained knowledge with set of rules. The user can use this system to recognize the approximate rating of the restaurant.