1. Introduction

1.1 Why this DPR Document ?

The main purpose of this DPR documentation is to add the necessary details of the project and provide the description of the machine learning model and the written code. This also provides the detailed description on how the entire project has been designed end-to-end.

Key points :

- Describes the design flow
- Implementations
- Software requirements
- Architecture of the project
- Non-functional attributes like:
  - Reusability
  - Portability
  - Resource utilization

2. General Description

2.1 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.

2.2 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.

2.3 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. Detailed Project Report 6

3. 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.

3.1 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.
- Heroku is employed for deployment.

4. Data Requirements

The Data requirements is totally supported the matter statement and also the dataset is accessible on the Kaggle within the file format of (.xlsx).

4.1 Data Collection

The data for this project is collected from the Kaggle Dataset, the URL for the dataset is https://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants?resource=download

4.2 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	Urls of the restaurants on Zomato website
address	Address of Restaurants
name	Name of the Restaurant
online_order	Whether online facility is provided or not
book_table	If advanced booking of table is allowed or not
rate	Average rate of the restaurant given by thecustomers
votes	Total number of votes
phone	Phone number of restaurant
location	Location of restaurant
rest_type	Type of restaurant
dished_liked	Most dished liked in that restaurant
cuisines	Types of cuisines served in that restaurant
approx_cost(f ortwo people	What is the approximate cost of the 2 people's meal
reviews_list	Rate and reviews given by the customers
menu_item	List of Menu Items
listed_in(type)	Type of meal
listed_in(city)	City where the restaurant is located

4.3 Importing data into Database

Created associate API for the transfer of the info into the Cassandra info, steps performed are:

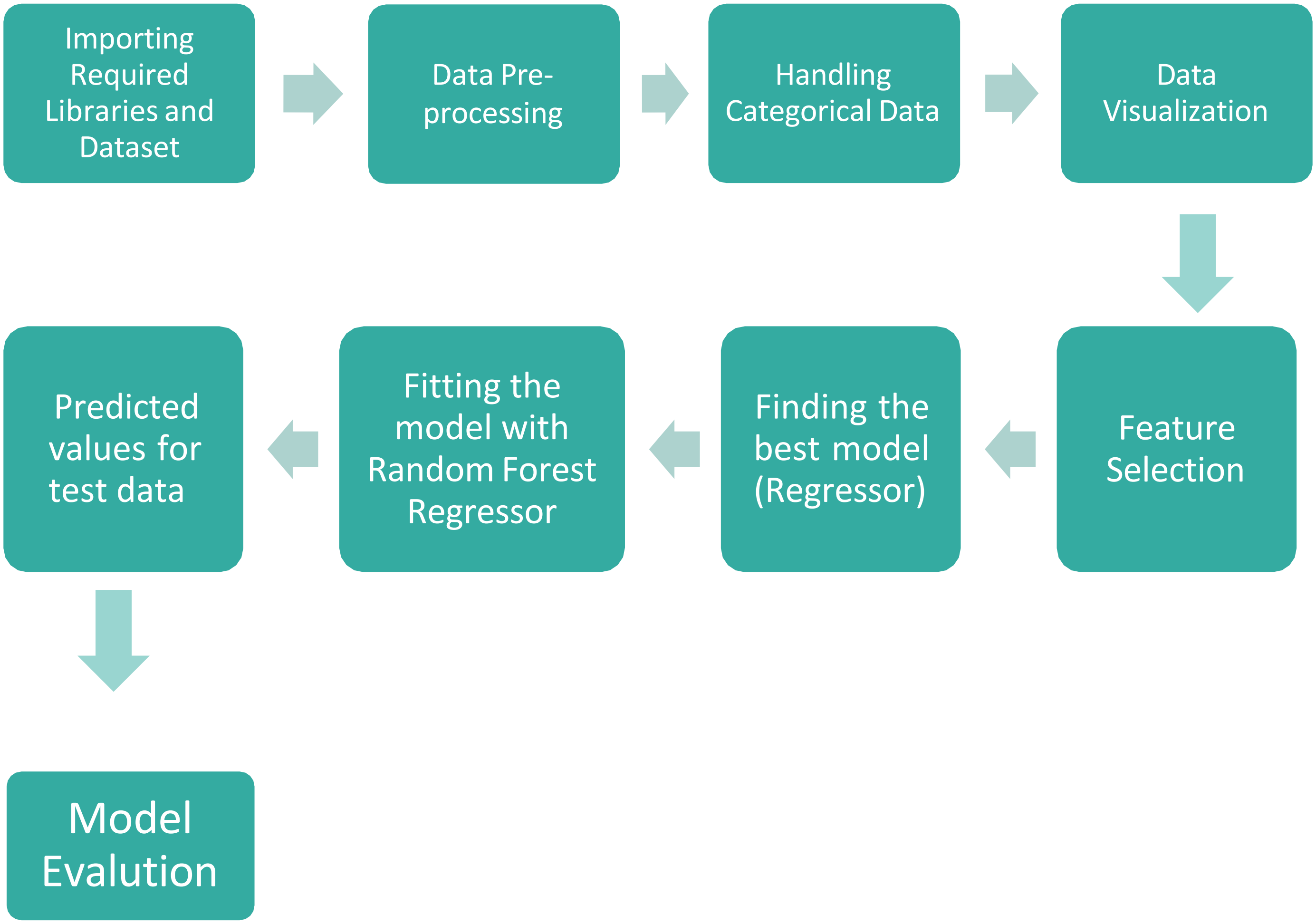
- Connection is created with the info.
- Created a info with name ZomatoInfo.
- cqlsh command is written for making the info table with needed parameters.
- And finally, a cqlsh command is written for uploading the Knowledge Set into data table by bulk insertion.

5. 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 desired columns into Datetime format.
- Performed One – Hot encoding on the desired columns.
- Checking the distribution of the columns to interpret its importance.
- Now, the info is prepared to train a Machine Learning Model.

6. Design Flow

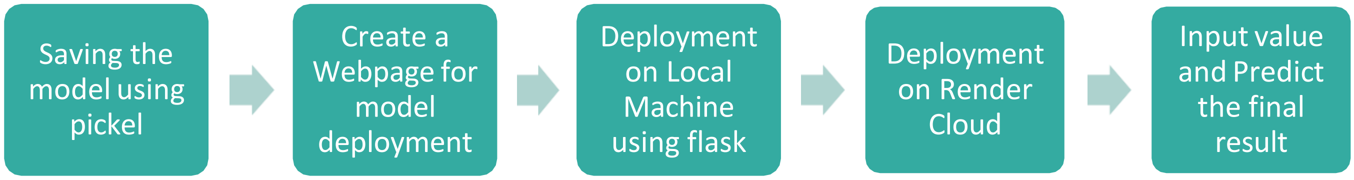
6.1 Modelling Creation and Evaluation


Importing Required Libraries and Dataset
Data Pre-processing
Handling Categorical Data
Data Visualization
Feature Selection
Finding the best model (Regressor)
Fitting the model with Random Forest Regressor
Predicted values for test data
Model Evalution

6.2 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.

6.3 Deployment Process


Saving the model using pickel
Create a Webpage for model deployment
Deployment on Local Machine using flask
Deployment on Render Cloud
Input value and Predict the final result

6.4 Logging

Logging is done for every action performed by your code use the python logging library for this.

7. Data from User

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

8. 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.

9. Rendering the Results

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

10. Deployment

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

11. 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.