

<b>Topic:</b>	Restaurant Rating Prediction
<b>Document Type:</b>	Low-Level-Design (LLD)
<b>Author:</b>	Lakshita Joshi
<b>Date:</b> <b>30/8/2024</b>	PW Skills: Full Stack Data Science Pro

## 1. Introduction:

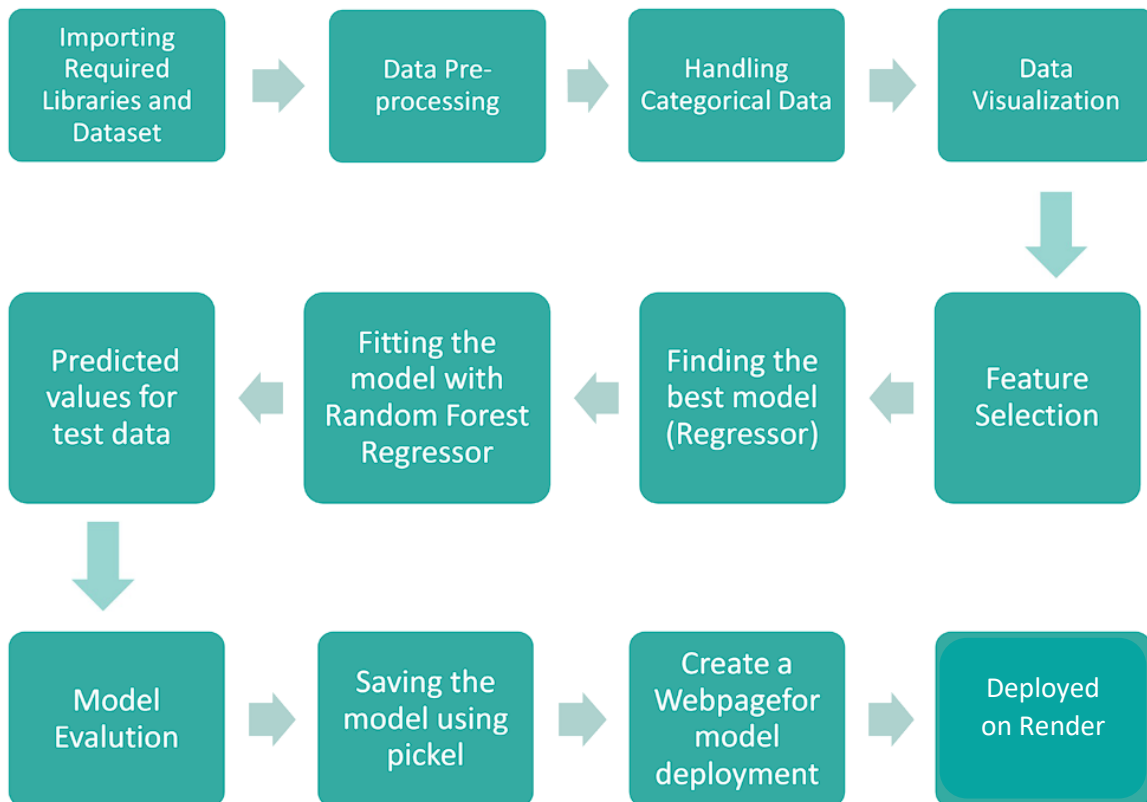
The Restaurant Rating Prediction Tool is an advanced application designed to accurately predict restaurant ratings by analyzing various key factors. It allows users to select from a wide range of cuisines, including Chinese, Thai, and street food, as well as different dining formats like dine-in, online delivery, and quick bites. This diversity ensures that the predictions are tailored to various dining preferences and operational styles.

In addition to these options, the tool considers other critical elements such as online ordering availability, table booking facilities, location, and cost. By incorporating these features, the app delivers reliable ratings that reflect the true essence of the dining experience. Whether you're choosing a quick meal or planning a special outing, the Restaurant Rating Prediction Tool helps you make informed decisions with confidence.

## 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.

## 3. Architecture



### 3. Architecture Description:

#### 3.1 Data Description:

The dataset was taken from Kaggle (URL: <https://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants>)

#### 3.2 Data Description:

There are 17 variables in the dataset which contains information about the restaurant.

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

### 3.2 Data Preprocessing:

This included importing important libraries such as Seaborn, Matplotlib, pandas etc. We imported the same dataset mentioned above from Kaggle.

- 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
- Performed Labelled Encoding and One hot Encoding on the desired columns
- Checking the distribution of the columns to interpret its importance
- Now, the information is prepared to train a machine learning model

### 3.3 Model Creation:

The pre-processed info is now envisioned and drawn insights helps us to select the feature that improves the accuracy of the model. The info is randomly used for modelling with different ML algorithms to create a model to predict the Rating of the Restaurants. After performing on different algorithms, we use Random Forest Regression to create a model and then also perform Hyperparameter Tuning to improve the accuracy of the model.

### **3.4 Data from user:**

It is retrieved from the created HTML web page

### **3.5 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.

### **3.9 Deployment:**

The deployed model through Render.