Zomato Restaurant Rating Prediction

About

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 theirrespective Ratings based on certain features

Life cycle of Machine learning Project

- Understanding the Problem Statement
- Data Collection
- Exploratory data analysis
- Data Cleaning
- Data Pre-Processing
- Model Training
- Choose best model

Project Overview

1. Understanding the Problem Statement
2. Understanding the Data set
3. Understanding the solution
4. Code understanding and walkthrough
5. Understanding the Deployment

zomato

Gaurav Sahu

1. Problem statement.

The underlying problem here is it has become difficult for non-established restaurants to compete with already-established restaurants. You are required to predict the rating for their better future.

In this project we are going to use the data given to build a Classification model:

Data is formed by taking 12,000 restaurants, serving dishes from all over the world. The data include the Location of the restaurant, Theme based on the restaurant or not, and many more thing

Objective

The primary goal is to gain insights from the Zomato dataset, understanding the dynamics influencing the establishment of diverse restaurant types in different areas of Bengaluru. With over 12,000 restaurants serving a global array of dishes, Bengaluru, as a city, continues to witness new restaurant openings regularly. Despite the growing demand, new establishments face challenges competing with well-established counterparts, often offering similar cuisines.

2. Understanding the dataset

1. url: The URL of the restaurant on the Zomato website.
2. address: The physical address of the restaurant.
3. name: The name of the restaurant.
4. online_order: Indicates whether the restaurant accepts online orders (Yes/No).
5. book_table: Indicates whether the restaurant allows table booking (Yes/No).
6. rate: The average rating of the restaurant (in the format "X.X / 5.0").
7. votes: The number of votes received by the restaurant.
8. phone: Contact phone number(s) of the restaurant.
9. location: The area or locality where the restaurant is situated.
10. rest_type: The type or category of the restaurant (e.g., Casual Dining, Cafe).
11. dish_liked: Some dishes that are popular or liked at the restaurant.
12. cuisines: The types of cuisines offered by the restaurant.
13. approx_cost(for two people): Approximate cost for two people to dine at the restaurant.
14. reviews_list: List of reviews and ratings given by customers.
15. menu_item: Menu items offered by the restaurant.
16. listed_in(type): The type of service the restaurant is listed under (e.g., Dine-out).
17. listed_in(city): The city or area where the restaurant is listed.

3. Solution

Zomato Rating Prediction

You have to build a solution that should be able to predict the ratings of the restaurants listed in the dataset.

4. Solution approach

1. Machine learning: ML classification Algorithms
2. Deep Learning: Custom ANN with sigmoid activation Function

Here we are going to take the first approach using the ML Algorithm

5. Solution Proposed

1. Load the data from MongoDB
2. Perform EDA and feature engineering to select the desirable features.
3. Fit the ML classification Algorithm and find out which one performs better.
4. Select the top few and tune hyperparameters.
5. Select the best model based on desired metrics

6. Deployment:

1. Docker
2. Cloud services
3. Adding self-hosted runner
4. Workflows