Lecture Resources



**ABSTRACT**

Zomato is one of the best online food delivery apps which gives users ratings and reviews on restaurants all over India. These ratings and reviews are considered as one of the most important deciding factors which determine how good a restaurant is.

We will therefore use the real-time Data set with various features a user would look into regarding a restaurant. We will be considering Banglore City in this analysis.

**Content**

The basic idea of analyzing the Zomato dataset is to get a fair idea about the factors affecting the establishment of different types of the restaurant at different places in Bengaluru, aggregate rating of each restaurant, Bengaluru being one such city has more than 12,000 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. In spite of increasing demand, it however has become difficult for new restaurants to compete with established restaurants. Most of them serving the same food. Bengaluru is the 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 restaurants, it has therefore become important to study the demography of a location. What kind of food is more popular in a locality. Do the entire locality loves vegetarian food. If yes then is that locality populated by a particular set of people for eg. Jain, Marwaris, Gujaratis who are mostly vegetarian. This kind of analysis can be done using the data, by studying factors such as

1. • Location of the restaurant
2. • Approx Price of food
3. • Theme based restaurant or not
4. • Which locality of that city serves that cuisines with maximum number of restaurants
5. • The needs of people who are striving to get the best cuisine of the neighborhood
6. • Is a particular neighborhood famous for its own kind of food.

“Just so that you have a good meal the next time you step out”

The data is accurate to that available on the zomato website until 15 March 2019. The data was scraped from Zomato in two-phase. After going through the structure of the website I found that for each neighborhood there are 6-7 categories of restaurants viz. Buffet, Cafes, Delivery, Desserts, Dine-out, Drinks & nightlife, Pubs, and bars.

Phase I,

In Phase I of extraction only the URL, name, and address of the restaurant were extracted which were visible on the front page. The URl's for each of the restaurants on the zomato were recorded in the CSV file so that later the data can be extracted individually for each restaurant. This made the extraction process easier and reduced the extra load on my machine. The data for each neighborhood and each category can be found here

Phase II,

In Phase II the recorded data for each restaurant and each category was read and data for each restaurant was scraped individually. 15 variables were scraped in this phase. For each of the neighborhoods and for each category their onlineorder, booktable, rate, votes, phone, location, resttype, dishliked, cuisines, approxcost(for two people), reviewslist, menu\_item was extracted. See section 5 for more details about the variables.

Acknowledgements The data scraped was entirely for educational purposes only. Note that I don’t claim any copyright for the data. All copyrights for the data is owned by Zomato Media Pvt. Ltd..

**Main Objective:**

The main agenda of this project is:

Perform extensive **Exploratory Data Analysis(EDA)** on the Zomato Dataset.

Build an appropriate **Machine Learning Model** that will help various Zomato Restaurants to predict their respective Ratings based on certain features

**DEPLOY** the Machine learning model via **Flask** that can be used to make live predictions of restaurants ratings

1. Source: Kaggle

[Click Here To Download Files](https://bit.ly/2I0Ypdy)