# Analysis Of Restaurant Ratings on Zomato Based on Dining And Delivery

Akshat Chourey
*Computer Science And Engineering*
*PES University*
PES2UG19CS028
akshatchourey123@gmail.com

G. Vijit Kumar Chary
*Computer Science And Engineering*
*PES University*
PES2UG19CS125
vijitkchary2001@gmail.com

Melavoy Nithin Reddy
*Computer Science And Engineering*
*PES University*
PES2UG19CS230
nithinreddy618@gmail.com

Nidhi Deshpande
*Computer Science And Engineering*
*PES University*
PES2UG19CS255
nidhidespande83@gmail.com

***Abstract*—The main purpose of this study was to examine the direct effects of restaurant location between dining and delivery service with the help of dining and delivery service review. Consumer surveys have indicated that people value the convenience and time savings of having their food delivered to them.In addition, many office-workers prefer the convenience of having food delivered to them rather than going out for lunch.**

***Index Terms*—Restaurant location, Dining services, Delivery services.**

## I. Introduction

According to Deepinder Goyal, Zomato chief executive officer and co-founder told Tech Crunch that he expects to reach 10000 restaurants in Asian country throughout a couple of months. "We have a sales team of around three hundred in Asian country and 5000-odd advertisers... these partners understand the amount we tend to bring back them thus it's quite simple for the U.S.A. to launch this new service."

The location has a significant impact on the success of the restaurant and also the sites are expensive to change, therefore location selection is a critical decision for restaurant owners.In this work, we have chosen a restaurant as a business representative and have empirically investigated to analyze how the success of a restaurant can be affected by its location and address. We have investigated the impact of restaurant type and location on the restaurant's success. To address this question, we analyzed the Zomato restaurant data-set of Bangalore - also called The Silicon Valley of India, Bangalore has a population of 10 million and is the country's third most populated city. This paper aims to analyze the types of restaurants available in a particular locality of Bangalore and how the location impacts its dining reviews as compared to its delivery reviews. The rest of the paper is organized in the following manner. First, the literature is reviewed on restaurants, various factors affecting the success or failure of restaurants like location, cuisines by comparing its dining and delivery ratings. Then research methodology is used to analyze the data and graphically represent the empirical results.

An effective restaurant recommendation is a strategy provides the opportunity to offer convenience to customers as well as to improve profits for restaurants.The main advantage of ordering food online is that everyone can order at their ease and there would not be any peer pressure while ordering the food because when going for casual dining there is also a pressure while ordering food we often ask the waiter what this food contains?And sometimes we get into situation where the dish you ordered is not that up to the mark or sometimes we need to order as per the waiter's recommendation because we don't have a brief description about that dish on the menu.

## II. Similar Approach

### A. Customer Perceptions of Restaurant Delivery[1]

Joel Collier and Sheryl Kimes studied the impact of a number of factors, including perceived convenience and control, had on satisfaction with online ordering and intention to use online ordering in the future.Sheryl Kimes studied consumer usage

of online ordering through restaurant and multi-restaurant sites and found that slightly less than half of consumers surveyed had ordered online. The key factors driving their satisfaction with online ordering were perceived convenience and control.They found that while some restaurants operate their own delivery service, most restaurants offer delivery through third-party delivery services, as noted above. Third party delivery companies typically charge restaurants a 20–25 percent commission. In exchange, the restaurant's menu is listed on the company's app or website. Consumers can then choose the food they want and,if desired.

What Makes Restaurants to choose Delivery over Dining Services?

**1.Increased Revenue** is the top reason for the outsourcing decision.Restaurants[2] aim to increase revenue through outsourcing their food delivery services to third-party online food delivery service providers.As additional stream of revenue is vital for the business to sustain as the capacity for dining-in is limited and sales turnover is then constrained.Hence, outsourcing to third-party online delivery service providers offers restaurant a lucrative income stream by having a dedicated delivery service to give their food a wider reach despite having a limited number of seats in the restaurant.

**2.Increase Exposure and Wider Customer Reach** The outsourcing to third-party can help to increase their exposure as well as to achieve a wider customer reach. They mentioned that online food ordering and delivering platform is a great and quick way to get their brand name seen. There is a high possibility for restaurants to be seen by the online users that have access to the online delivery service platform when they explore for food ordering options.

**3.Convenience**[3] is one of the benefits associated with outsourcing delivery services.It is more convenient to outsource food delivery services than to establish in-house delivery service because most of the responsibilities could be delegated to the third-party online delivery service providers.

**4.Location** is the most important aspect of whether the restaurant will be a success or a failure.So to even out the odd restaurants decided to collaborate with third-party food delivery platform because the restaurant that located in areas that are surrounded by commercial lots, colleges, entertainments, and food stalls or the places where the area is still under development can have a major impact on the restaurant business. Hence, customers who visit the restaurant would encounter problem such as poor traffic conditions and limited availability of parking or under developed area.

**Plots used** to better represent the difference between the dining and delivery service **Bar Graph** as its most commonly used graph to check if anything could be inferred from the resulting graph .And in our dataset it shows which location is particularly favorable and what is the rate of dining and delivery through review counts. **Pie chart** is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart, the arc length of each slice , is proportional to the quantity it represents.And in our data set it is used to show that between delivery and dining which is most frequently used across different locations.

### *B. Collaborative Study on User ratings on Restaurant Recommendation[4]*

Recently research conduct on Online Food Delivery by Carsten Hirschberg his study indicates that introduction online food delivery penetrates the market and broke the normal restaurant dining by 30 per cent and his study also shows that it can break the normal chain restaurant by 65 per cent.

In recent years, with the popularity of internet and development of e-commerce, traditional catering industry was impacted by online activities. Thus, effective restaurant recommendation is a strategy provides the opportunity to offer convenience to customers as well as to improve profits for restaurants. They used Collaborative filtering algorithm that trains the corresponding model through the user's scoring information , and uses this model to predict the unknown data.

This paper is based on the users - item evaluation matrix of catering for the computation of similarity between users, the purpose is to predict ratings behavior of user who use the system, then according to the predicted scores to generate recommendation.The user reviews are parsed

and necessary details (like features and views) are obtained. Usually, these reviews (whether positive or negative) hold very significant for the recommender system as they are provided by the users depicting their feelings and views related to a particular restaurant. The summarized rating of a restaurant can be computed in terms of the collective reviews (be it neutral,positive or negative). The lexicon approach is applied to obtain divergence of sentiments whether Positive or Negative. It can be considered as a dictionary having collection of words/ phrases that categorize them as either positive or negative emotions.

[5]Consequently, the features are merged according to user's perspective and a score is generated for every sentence. Then the score of all the sentences are merged to generate a summarized score for a single review and thereafter using the database to store this sentimental outcome. Using the restaurant recommendation system, the basic restaurant amenities are selected by the user and based on this parameter the matching hotels are then populated.They intensified the impact of the **average score factor** on similarity and introduced **confidence level correction factor** to enhance the credibility.To show the effectiveness of the proposed AdvancedCF algorithm, they compared the performances of the algorithm against UCF. In the same condition, the smaller MEA value,the better performance. In addition, K, the number of nearest neighbors is an important factor for algorithm' performance.So they set the value of K from 5 to 50, increasing 5 each time. For some of the restaurant's the users may remark bad or poor rating which can ruin the status of the restaurant and the system too delivers minimal accuracy delivering bogus results. The machine learning algorithm is being proposed to resolve the issue of personalized Restaurant selection. In the proposed restaurant recommendation the features provided by the hotel/restaurant and users reviews are examined and tagged with the help of Natural language Processing. Using the proposed technique effective results are achieved with enhanced functionality thereby offering precise output related to user preference.

As a result, they got the MEA comparisons:

- their approach outperforms the UCF, that is to say, AdvancedCF improve recommend precision in a certain extent.
- judging from the results of RMSE,residence has the greatest effect on the prediction results,secondary effect on age and then gender. According to the above influence ratio, their weight factors were set to 0.5, 0.3,and 0.2.

### *C. Restaurant Analysis and Prediction[6]*

The focus of the article 'Zomato Bangalore Restaurant Analysis and Rating Prediction'[] is to make a model to predict the ratings of restaurants based on the information contained in the data set. They made use of their data set to also learn various other information from the data set such as :

- The number of restaurants giving online delivery option.
- The most ordered dish.
- The cuisines which are ordered by most people.
- The type of restaurant liked by most people.
- Does the restaurant offer to book tables in advance?

[7]They proceeded to first plot various charts and graphs to determine relationship between the variables present in the data set. The graphs and charts used are the variables present in the data set. The graphs and charts used are:

- Bar charts used to show the relative numbers for the attributes.
- Pie charts to show which attribute appear more number of times and which appear fewer number of times.

These methods of plotting help to answer most of the secondary question asked about the information in the data set and visualize the data. Python was used for model building because of the wide number of libraries available for data processing and analytic available in Python. Libraries such as Pandas,matplotlib ,Seaborn ,Numpy ,Sklearn are used. The authors then proceed to building a model to determine the ratings of restaurants.

They began by cleaning the data such as dropping columns which are not required, handling missing values by replacing either with mean or mode or deleting the entire row. The data was further processed by applying certain encoding on the data

because categorical variables were present in the data but machine learning can be applied only to numerical data so the data was encoded like 0/1 for No/Yes and so on.One-hot encoding is applied to the categorical data.

[8]The correlation between variables is calculated and independent variables are singled out. Feature selection is used to select only the necessary features for models and this is done by making sure that there is no reduction in classification accuracy. This prevents over-fitting. The data set was then split into training data set and testing data set using the 70-30 rule, 70% of data for the training data set and 30% of data for testing data set chosen randomly for fairness of the model.The data was standardized to allow for easier processing of data as variables have different ranges. For prediction of values several Machine learning methods were used such Ridge regression ,k-Nearest neighbours,Logistic regression,lasso regression and Random Forest regression. The accuracy of each algorithm was calculated for the testing data set.Out of these methods Random Forest Regression resulted in the best accuracy.

### *D. PREDICTING CONSUMER PURCHASING DECISIONS IN THE ONLINE FOOD DELIVERY INDUSTRY[9]*

In this paper, they used several prediction models to determine whether a customer would purchase again from the online food delivery platforms. The ability to do so provides a strong predictive tool for online food delivery providers to have a better understanding of their customers, and to improve their services accordingly Three prediction models: decision tree, random forest, and rule-based classifier, will be compared based on their performances. In all three prediction models, the attributes of marital status, occupation, educational qualifications, family size, ordering medium preferences, meal preferences, and cuisine preferences are eliminated. In all three models, the training data covers 75% of the data, while the remaining 25% is assigned to test data. The method for data training is cross-validation with 10 folds. Finally, they are compared based on prediction accuracy. The prediction accuracy is derived from the confusion matrix which summarizes a classifier's classification performance in relation to some test data.

- Decision tree
  - The decision tree is used for regression predictive modeling problems. It is a binary recursive partitioning tree, where each parent node in the tree is split into two child nodes. Further,decision tree is known for its simple interpretation and inherent logic. Here,decision tree is used to predict the purchasing decisions of online food consumers.
- Random forest
  - A random forest is made up of many separate decision trees that work together to form an ensemble. Each tree in the random forest produces a class prediction, and the class performing the best becomes the prediction of the model.
- Rule-based classifier
  - The rule-based classifier is employed in the class prediction method to give the rules a rating, which is then used to predict the class of future cases. When compared with the other prediction models used, it performs less than the random forest model

## III. Problem Statement

Using Exploratory Data Analysis and Predictive statistics on the various restaurants available in Bengaluru, we would like to enhance the feasibility for the people to choose the restaurant in their locality or anywhere in the city on the basis of their preference of dining ,i.e, whether they would like to dine in the restaurant or prefer delivery to their homes.

For this we pre-process the dataset and make the observations:

## IV. Inference

- Different restaurants have different types of cuisines available.People we free to choose from an enormous "menu" and had varieties of food available for them.

- Price for dining had a causal relationship with the cuisines being served there.

- No. of Reviews vs Rating: Little-to-no relationship.

- Prices for dining were related to the location where the restaurant was present.But they didn't change even when there were a considerable amount of restaurants in the particular location.

- Some restaurants had large number of dining reviews as well as delivery reviews indicating that particular restaurants were particularly famous in that location.

- Even though a restaurant provided variety of food, the rating for that restaurant was not affected by it.Rather than it depended on the quality of food and service provided by the restaurant.

## V. Work Done

Our Study shows that a restaurant is not defined by the varieties of food it provides but rather than the different amenities it has, like the taste,the ambience, the service and so on.To study the trends and correlations among these factors we performed an extensive EDA and predictive analytics on our dataset.

We critiqued different research papers that had the basis of research related to our work.We came to find that many of the research papers had taken up data that was inaccurate and prone to minimal data analysis.Some of the research papers had data that was either underfitting or overfitting the model they were incorporating in the research paper. One of the research paper we had studied focused more on the qualitative analysis rather than the more accurate,quantitative analysis.But again another research paper put more of the focus on the modelling and analysis of the dataset that they deviated from the original cause of the research paper.

Hence we would like to work on the parts that the other research papers missed out on and improve the basis of understanding of the work done.So the steps taken by us include:

- Data Collection. The dataset we will be using in our project is the Zomato Resaturants in Metropolitan Areas of India(Bangalore) from Kaggle(https://www.kaggle.com/aestheteaman01/zomato-restaurant-caf-dataset-pune-bangalore). This dataset contains 13 features and 5110 rows of which there are 4882 unique values.

- Data Cleaning. Before getting on to the usage of the dataset we need to clean it.The datatset has a total of 1436 missing values. We impute these missing values with the mean. We choose to drop the Website,Address, Phone_No,Latitude and Longitude columns from the dataset as they were not giving us any meaningful data(redundant) that could have been worked upon.

- Some features had to be modified so as to get a better understanding about them.Names of some columns were fixed for easier readability.The duplicate rows were removed from the dataset.

- Dataset Inspection. This phase includes inspection of the data,and visualisation by plotting various graphs between different attributes of the dataset and drawing conclusions from them.This is where the EDA on the dataset is performed.

## VI. References

[1]Dr. Anupam Bhatia and Ms.Sneha* "Big Data Analytics and Artificial intelligence in business and marketing ,LOCATION BASED RESTAURANT PREFERENCES IN BANGALORE"

[2]Joel Collier and Sheryl E.Kimes "Customer Perceptions of Restaurant Delivery"

[3]GohSee-Kong, NgSoo-Ryue, WongShiun-Yi, Chong-Lily "OUTSOURCING TO ONLINE FOOD DELIVERY SERVICES: PERSPECTIVE OF FI&B BUSINESS OWNERS"

[4]L. Li, Y. Zhou, H. Xiong, C. Hu and X. Wei, "Collaborative filtering based on user attributes and user ratings for restaurant recommendation," 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation

Control Conference (IAEAC), 2017, pp. 2592-2597, doi: 10.1109/IAEAC.2017.8054493.

[5]R. M. Gomathi, P. Ajitha, G. H. S. Krishna and I. H. Pranay, ”Restaurant Recommendation System for User Preference and Services Based on Rating and Amenities,” 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862048.

[6]Zomato Prediction and Classification - https://www.kaggle.com/nrdnandan/zomato-prediction-classification

[7]Predicting Restaurants' Rating And Popularity Based On Yelp Dataset- http://cs229.stanford.edu/proj2017/final-reports/5244334.pdf

[8]Zomato Bangalore Restaurant Analysis and Rating Prediction- https://purnasaigudikandula.medium.com/zomato-bangalore-restaurant-analysis-and-rating-prediction-df277321c7cd

[9]Predicting consumer purchasing decisions in the online food delivery industry - https://arxiv.org/ftp/arxiv/papers/2110/2110.00502.pdf