We_R - Restaurant Recommender System

by Nishanth. S

Submission date: 06-Dec-2021 02:56PM (UTC+0530)
Submission ID: 1722053921
File name: We_R_-_Restaurant_Recommender_System.pdf (321.03K)
Word count: 1791
Character count: 9665

RESTAURANT RATING PREDICTOR MODEL

Nishanth S

PES2UG19CS265

B.E. Dept. of Computer Science & Engineering

PES University, Electronic City Campus

Bangalore, Karnataka, India

Rohith G

PES2UG19CS335

B.E. Dept. of Computer Science & Engineering

PES University, Electronic City Campus

Bangalore, Karnataka, India

P P Sujith

PES2UG19CS272

B.E. Dept. of Computer Science & Engineering

PES University, Electronic City Campus

Bangalore, Karnataka, India

Shabarinath K

PES2UG1CS373

B.E. Dept. of Computer Science & Engineering

PES University, Electronic City Campus

Bangalore, Karnataka, India

Abstract - **TripAdvisor is one of the most popular travel websites, and it includes information regarding hotel bookings, sites of attraction, restaurants reviews, and much more. The dataset deals with the details of the restaurants present in Europe. It is always a common habit to look out for the best restaurants when a tourist visits a place. The rating of a restaurant is usually done at a scale of 1 to 5 based on all the features influencing a user's experience. There are many factors that influence a user's experience when one visits a place such as a restaurant. This data modelling will help the people choose their restaurant of interest out of the best picks done based on the features that describe the restaurant.**

Keywords – **restaurants, rating, factors, user's experience, data modelling.**

I. INTRODUCTION

A recommender system is a subset of information-filtering systems that strives to pick and serve items that are believed to be of use to a user. Nowadays, recommender systems are widely used globally as businesses tend to reach out more towards understanding their customers' choices and preferences. There are many recommender system algorithms developed and used till date.

In today's world such recommender systems are used as there are so many choices available before a customer in any given enterprise. It's a tedious task to go through every single detail or choice, and hence recommender systems help solve this problem of an overabundance of choice. Recommender systems are so designed in a way to bring to the customers only those choices and options which the customer is more likely to pick. There are many ways or processes through which this can be achieved such as collaborative filtering, etc. These systems learn from the

previous user data regarding the users' choices in the past to predict what are they would choose in any given scenario.

As stated above, the problem that we chose was to analyze the correlation between different aspects of restaurants such as reviews, food served, etc. and hence build a predictor model to suggest a good restaurant to a user based on the analytics. To solve the above problem, we chose a dataset that deals with various restaurants in different European countries along with their ratings of based on the food served, service, value for money, etc.

II. RELATED WORK

After doing literature survey we found some papers which were very close to our objective. They were many different approaches used to solve the same problem of recommending restaurants.

R.M Gomathi et.al, proposes the use of NLP, i.e., Natural Language Processing algorithm to perform a collaborative filtering approach. They make use of all the features of the restaurants and reviews of the users to compute the overall percentage of comments and store them as the output. the customer chooses the features provided by different hotels according to his interests and provides them as input to the model. With the data centered on this, that relevant hotels with matching tags are fetched and user comments are verified to identify the hotel with the highest ranking. As a result, these highest ranked hotels are recommended to the user by the proposed method.[1]

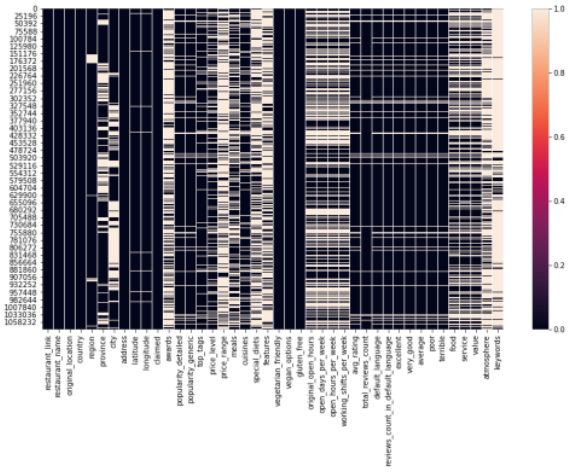
Arun Tripathi et.al, proposes the use of K nearest neighbors and multi class SVM to predict the ratings between restaurants and users. They further go on to measure the performance between the two models and conclude that the multi class SVM outperforms KNN. For rating-prediction they correlate user based and item based collaborative filtering methods. The reviews are taken by the user and based on these, the algorithm searches for the keywords in the feedback and distinguishes them into positive negative and neutral reviews. The overall rating is calculated based on these reviews and henceforth the results are estimated.[2]

This approach will be similar in many ways to the first paper mentioned above since we are also going to recommend restaurants based on the tags associated with them. However, we are going to ignore comments by users in our approach and take input directly from the user regarding his choices for a particular restaurant. Also, we will be calculating the distance between the user's location and the restaurants' location which will be helpful in recommending the nearest restaurants available to the customer.

III. PROPOSED WORK

Our approach will be to take in input from the user based on their interests along with their location coordinates and then perform a collaborative filtering approach along with natural language processing to predict the best restaurants available to the customer.

A) Preprocessing: This process is performed to ensure that the data set doesn't contain redundant information such as null values, missing values, etc.



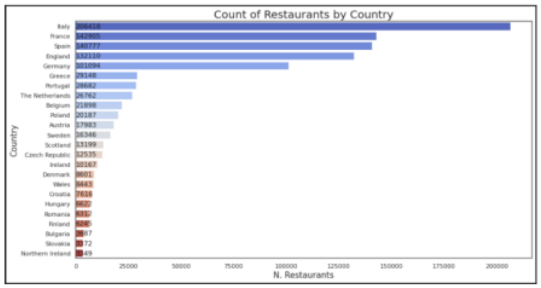
The above plot shows the number of null values present in each column. This provides an insight into which of the columns are we going to keep for our future calculations.

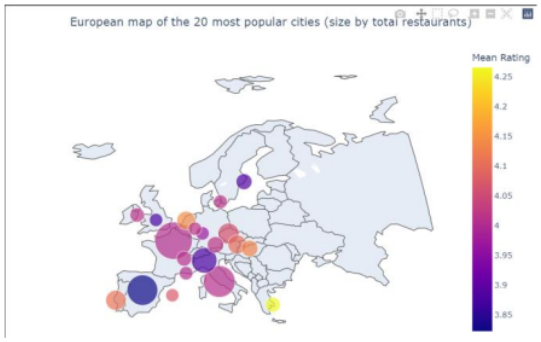
We also carry out several other preprocessing techniques such as finding correlation between the features and extracting only those that are

necessary. We also drop all those rows containing null values.

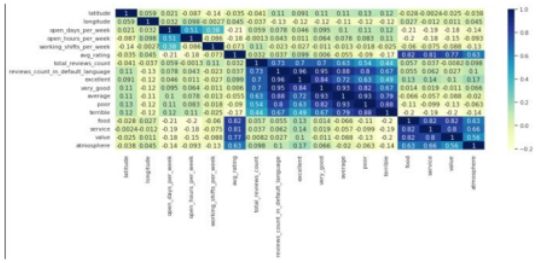
Some columns are left untouched even with null values as they are used later for model building.

Exploratory analysis helps us visualize country-wise restaurant distribution.





Italy seems to have the highest number of restaurants and the most popular cuisine seems to be Italian. Athens seems to be an outlier given both its geographical location as well as the abnormally high mean rating. Most of the popular cities with respect to the restaurants seem to be distributed towards the west of Europe. A few of the columns such as food, service, value, and atmosphere seem to be highly correlated indicating that they have quite a strong relationship.



B) Modelling: To build the model, we first perform correlation analysis on all the features of the dataset to find out their interdependencies along with the objective of locating the target variable.

Three models are built.

1) To predict the average rating of a restaurant given its food rating, service rating, atmosphere rating, and value rating. This approach is done using the method of linear regression.

As we have seen in the above plot, it's more convenient to consider only the features present in the above correlation as they show high correlation towards average rating of the restaurants.
For cross validation purpose, the data set was divided into training set and testing set at a split ratio of 70:30. The average rating being the target value, was predicted after training the model with the training data set.

2) A progressive filtering model was built to suggest restaurants based on user's preferences. The conditions that were set as filters include restaurants having vegan options, vegetarian-friendliness, and gluten-free cuisine. The user had to provide input based on the above criteria to meet their interests of choice. After the data was inputted, the algorithm would perform a filtering approach and classify the restaurants and provided the user the same after sorting them on the basis of the restaurants distance from the user's location.

3) A recommender system was created using the relevant key words of each restaurant. Each restaurant was first ordered based on distance to the current location of the user. Using other features input by the user, the

restaurant's priority value was calculated based on similarity between the expected features. Assuming that at least some of the highest priority restaurants fit the users taste, those restaurants that were most similar to these were given a higher priority. The restaurants with the highest priority were recommended to the system.

Algorithm:

```
Inp_loc -> user location
Inp_features -> user preferred features
Sort(restaurants) #on distance to user

for restaurant in restaurants:
    restaurant[priority] = no of words in
    both inp_features and restaurant[features]
sort(restaurants) #on priority
max_priority = restaurants[0][priority]

for restaurant in restaurants:
    if restaurant[priority] < max_priority:
        restaurant[priority] +=
            max(similairity between restaurant
            and other restaurants with
            max_priority)

sort(restaurants) #based on priority
Recommended -> highest(restaurants)
```

C) Evaluation: When the linear regression approach was followed, it was possible to predict the average ratings of the restaurants based on the cross-validation tests but was insufficient to recommend a restaurant to a customer based on linear regression model only.

When the progressive filtering model was used, the model was able to classify the restaurants according to user's given criteria which was received as user-input. However, this method was limited in its reach, as the user was restricted to only certain attributes which was prompted by the model, and hence quality of the recommendations still wasn't up to par.

The NLP recommender system above built models work based on the explicit information provided by the user. If the user happens to give invalid or no input such that a non-zero priority value cannot be calculated, the model fails. Evaluation of such a model is also quite hard as it requires explicit feedback from the user or information regarding if a restaurant was chosen, which can be collected after the model is deployed (earliest at alpha testing).

IV. EXPERIMENTAL RESULTS

The algorithm proposed was used to rank restaurants based on what the user is most likely looking for. It hopes to solve the problem of making a better-informed decision, and hence, a better recommendation even when the information given by the user is slightly inadequate. Testing is to be deferred to the alpha testing phase.

V. CONCLUSION

As shown in the above methods, the Natural Language Processing strategy was more efficient compared to the other two models. The attributes of the dataset were considered for calculation based on customer's sentiments which made the model to yield results with a high precision.

VI. CONTRIBUTION

1. **Nishanth S, PES2UG19CS265**
   Literature survey, Preprocessing and modeling.
2. **P P Sujith, PES2UG19CS272**
   Literature survey, Preprocessing and modeling.
3. **Rohith G , PES2UG19CS335**
   Literature survey, Preprocessing and modeling.
4. **Shabarinath K, PES2UG19CS373**
   Literature survey.

VII. REFERENCES

[1] R.M.Gomathi, P.Ajitha, G. Hari Satya Krishna, I. Harsha Pranay, "Restaurant Recommendation System for User Preference and Services Based on Rating and Amenities", Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019).

[2] Arun Tripathi, A. K. Sharma, "Recommending Restaurants: A Collaborative Filtering Approach", 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Amity University, Noida, India. June 4-5, 2020.

We_R - Restaurant Recommender System

ORIGINALITY REPORT

6%	1%	5%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Raunak Bhupal, Kalidindi Lakshmi Sanjana, Nikhil Kishan Khaneja, Prakhar Bhartiya, Achaladi Spoorthi Bhat, Sandesh B J. "Finger Vein Authentication System", 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021 Publication	3%
2	Arun Tripathi, A.K. Sharma. "Recommending Restaurants: A Collaborative Filtering Approach", 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020 Publication	1%
3	R.M. Gomathi, P. Ajitha, G. Hari Satya Krishna, I. Harsha 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 Publication	1%

4	www.digitalmarketingchandigarh.com Internet Source	1%
5	www.mdpi.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		