Quick Guide to Build a Restaurants Recommendation System

Kunal Bhashkar

kbhashkar@rediffmail.com

Introduction

The dataset is obtained from a recommender system prototype from UCI Machine Learning Repository of Restaurant and Consumer data Data Set. The task is to generate a top-n list of restaurants according to the consumer preferences.

Be it a fresher or an experienced professional in data science my main purpose to write this article is to get your started with recommendation systems so that you can build one.

We can simply say that Recommendation engines are nothing but just an automated form of a person who is sitting at shop counter and predict about customer on the basis of their previous experience or customer choice. If customer ask him for a product then he not only to show that product, but also the related ones which you could buy. They are well trained in cross selling and up selling. So, does our recommendation engines.

The ability of these engines to recommend personalized content, based on past behavior is incredible. It brings customer delight and gives them a reason to keep returning to the website. In this post, I will cover the fundamentals of creating a recommendation system using in Python. I will get some intuition into how recommendation work and create basic popularity model and a collaborative filtering model.

Topics Covered

1. Type of Recommendation Engines
   i. Recommendation by most popular Item
   ii. Recommendation Algorithm
2. Data Set and Attribute Information
3. Analysis,Visualization Data Set by using Matplotlib and Plotly in python and build a and user-user collaborative filtering recommandation system
4. Result
5. Conclusion and Observations

1. Type of Recommendation System

i. Recommendation on the basis of most popular Items

In this type of recommendation system we recommend only those items which is most popular in the users i.e which is liked by most number of users. As for example, In Restaurant & consumer data Data Set, we have rating_final.csv file which contain UserID, PlaceID, rating, food_rating, service_rating attributes. In this data set, we take only those list of restaurant which having high rating value in their respective categories.

ii. Recommendation Algorithm

In this type of algorithm in which there are two kind of algorithm takes place as below,

Context Based Algorithm

As the name suggest, these algorithms are strongly based on driving the context of the item. Once you have gathered this context level information on items, you try to find *look alike* items and recommend them.

Collaborative filtering algorithms

This kind of filtering technique is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. This type of filtering technique has three different type of category which is as below,

a. User-User Collaborative filtering

In this type of filtering algorithm we look alike customer to every customer and offer products which first customer’s look alike has chosen in past. This algorithm is very effective but takes a lot of time and resources since it requires to compute every customer pair information. Therefore, for big base platforms, this algorithm is hard to implement without a very strong parallelizable system.

b. Item-Item Collaborative filtering

It is quite similar to previous algorithm, but instead of finding customer look alike, we try finding item look alike. Once we have item look alike matrix, we can easily recommend alike items to customer who have purchased any item from the store. This kind of algorithm is very less resource consuming as compare to user-user collaborative filtering. Hence, for a new customer the algorithm takes far lesser time than user-user collaborate as we don’t need all similarity scores between customers.As we can say that with fixed number of products, product-product look alike matrix is fixed over time.

c. Market basket analysis

Market Basket Analysis is a modelling technique based upon the theory that if you buy a certain group of items, you are more (or less) likely to buy another group of items. Market basket is a widely used analytical tool in retail industry. However, retail industry use it extensively, this is no way an indication that the usage is limited to retail industry. This algorithm don't have high predictive power.

2. Data Set and Attribute Information

Here i am using the data set of Reataurant and consumer data Data set which has been collected from UCI Machine Learning Repository (Link: https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data). The basic information about this data set is it has total nine csv file in which five of them have Restaurant category, three are from customer category and one is from rating category.

You can download this file from **https://archive.ics.uci.edu/ml/machine-learning databases/00232/Rcdata.zip**

3. Analysis, Visualization of Data Set by using Matplotlib and Plotly and Build a user-user collaborative filtering recommandation system

Now i start some hands-on experience to Analyse and visualize the data for Reataurant and consumer data Data Set.

Code Example:

You can find code on my github link : <githublink>

4. Result

Our main purpose is "to generate a top-n list of restaurants according to the consumer preferences" by applying user-user collaborative filtering algorithm. Hence finally we got the recommandation of restaurant according to their rating with their category wise as below

#Recommandation on the basis of overall rating

```
Rcuisine_American              0.001860

Rcuisine_Armenian              0.014780
Rcuisine_Bakery                0.058773
Rcuisine_Bar                   0.019786
Rcuisine_Bar_Pub_Brewery       0.116406
Rcuisine_Breakfast-Brunch     -0.058542
Rcuisine_Burgers              -0.083035
Rcuisine_Cafe-Coffee_Shop      0.112542
Rcuisine_Cafeteria             0.010247
Rcuisine_Chinese               0.013291
Rcuisine_Contemporary          0.110369
Rcuisine_Family                0.146112
Rcuisine_Fast_Food            -0.098107
Rcuisine_Game                  0.067152
Rcuisine_International         0.197299
Rcuisine_Italian              -0.028509
Rcuisine_Japanese              0.086276
Rcuisine_Mediterranean         0.161423
Rcuisine_Mexican              -0.101195
Rcuisine_Pizzeria             -0.054471
Rcuisine_Regional             -0.205184
```

#Recommandation on the basis of service rating

```
Rcuisine_American            0.046926

Rcuisine_Armenian            0.047020
Rcuisine_Bakery              0.092859
Rcuisine_Bar                -0.017405
Rcuisine_Bar_Pub_Brewery     0.104559
Rcuisine_Breakfast-Brunch   -0.063335
Rcuisine_Burgers            -0.174886
Rcuisine_Cafe-Coffee_Shop    0.097953
Rcuisine_Cafeteria          -0.006442
Rcuisine_Chinese            -0.035177
Rcuisine_Contemporary        0.136682
Rcuisine_Family              0.208064
Rcuisine_Fast_Food          -0.208497
Rcuisine_Game               -0.029380
Rcuisine_International       0.221540
Rcuisine_Italian            -0.043081
Rcuisine_Japanese            0.106286
Rcuisine_Mediterranean      -0.029380
Rcuisine_Mexican            -0.041290
Rcuisine_Pizzeria           -0.035325
Rcuisine_Regional           -0.258578
```

#Recommandation on the basis of food rating

```
Rcuisine_American           -0.092815

Rcuisine_Armenian            0.235444
Rcuisine_Bakery              0.048637
Rcuisine_Bar                -0.105803
Rcuisine_Bar_Pub_Brewery    -0.005632
Rcuisine_Breakfast-Brunch   -0.145089
Rcuisine_Burgers             0.014586
Rcuisine_Cafe-Coffee_Shop    0.027880
Rcuisine_Cafeteria          -0.102661
Rcuisine_Chinese            -0.020661
Rcuisine_Contemporary        0.073894
Rcuisine_Family              0.102352
Rcuisine_Fast_Food          -0.165883
Rcuisine_Game               -0.075901
Rcuisine_International       0.234228
Rcuisine_Italian            -0.013019
Rcuisine_Japanese            0.143163
Rcuisine_Mediterranean       0.157608
Rcuisine_Mexican             0.042431
Rcuisine_Pizzeria           -0.067518
Rcuisine_Regional           -0.075901
```

Hence we can say that Rcuisine_International, Rcuisine_Armenian, Rcuisine_International is in top priority list of overall, food and servive category respectively. Here the main disadvantage is items with lots of history get recommended a lot, while those without never make it into the recommendation engine, resulting in a positive feedback loop. At the same time, new users have no history and thus the system doesn’t have any good recommendations.

5. Conclusion and Observations

Here we observe that, correlations are slightly different all three categorical rating. Rcuisine_International gets top rating in combined and service rating both but Rcuisine_Armenian gets the top rating in food category. Whereas Rcuisine_Regional has lowest rating (-0.205184) in combined category, Rcuisine_Breakfast-Brunch has lowest rating in food category and Rcuisine_Regional also has lowest rating in service category.

Now i'll discuss the features which affect the rating most. We see that a list of features like parking, alcohol, Smoking Area, Rambience, Other Service, Price, Dress Code, Accessibility, Area will affect the ratings.

i. Parking

As we can see that parking affect the combined rating but not food rating. The parking has many options like: 'parking_lot_none', 'parking_lot_public', 'parking_lot_valet parking', 'parking_lot_yes' different values. In this, parking_lot_public affect the restaurant and service rating both because it only matter for the car owner.

ii. Alcohol

There are basically three kind of service No_Alcohol_Served, Wine-Beer, Full_Bar, so all different types of drinker have different choice.

iii. Smoking Area

We see that smoking area are most important for smokers. The restaurant having smoking area are highly rated as compare to no smoking area of Restaurant.

iv. Ambience

We see that rambience will always affect the rating, because consumer always prefer those restaurant which have family type or someone wants friends type.

v. Price

The price is totally depends on the budget of the people. Mostly medium and high-price restaurants have better ratings.

vi. Dress Code

The dress code will also affect the rating. We see that the restaurant have formal dress code have always good rating.

vii. Other Services

We see that the restaurant having variety of other services having good rating as compare to those having internet only. So, other services increase the rating.

viii. Accessibility

Accessibility doesn’t affect the rating more, because sometimes the area where the people prefer close space as compare to open space, for cuisine, what type of food people prefer is depends on which rating we are predict. As per example, fast food restaurant have not good rating.