A Survey on Restaurant Recommendation System Based on Feature Selection and Classification Method

Shreya Joshi[1], Jigyasu Dubey[2]
[1]*Shri Vaishnav Institute of Technology and Science*
[2]*Shri Vaishnav* Vidyapeeth Vishwavidyalaya
(E-mail: Shreya.270892@gmail.com)

***Abstract*— In the recent years, opinion of person has become the most useful metric for recommendation systems. We look for ratings before purchasing any product. Social media growth has given the opportunity to know what likeminded people are interested in. In this paper, we are presenting a survey for a restaurant recommendation system by incorporating the narrowed down information of methods and data which are already being used for research in this domain. To make precise predictions and provide proficient recommendations the data and method are most important factors. By thoroughly analyzing the literature we came across the fact that the Facebook and Yelp are most successfully used datasets. We are also presenting a survey on various techniques used and their advantages.**

Keywords— **Recommendation System, Classification, Content Based method, Collaborative Approach, Accuracy**

I. INTRODUCTION

This social influence plays an important role in selection now days as we rely on people's opinions. It was not the case of traditional systems. The recommendation system provides a range of options to choose best among available options. We, find many recommendation systems which are based on rating. These systems are based on filtering the mammoth information. One domain, which is receiving a relatively little more attention, relating to leisure and entertainment, is the recommendation of food items [1].

When you want to eat out, it's difficult to choose which restaurant to go to. It gets even tough when you are finding something suitable for a larger group, because there are huge choices available proportionally the selection time increases [2]. We are presenting a survey on available approaches and there evaluation. This motivates us to create a recommendation application based on social data or dataset to recommend best restaurant based on required attributes.

II. BACKGROUND STUDY

The dataset based recommendation systems is being used for many years so apparently it has lot of research content availability, on the contrary the alternate sources like social media based recommendation system is not very old approach, still a decent body of literature exists in this domain.

A. Approaches for Recommendation Engine:

Content-Based Recommendation

The content-based system requires the content regarding both user and item. On the basis of shared attributes among users we can construct user-profile and item-profile. For example, for a restaurant we can present our opinion in form of rating by stars and the quality on basis of attributes chosen. For user profile, you can do the same thing based on the users likes some Restaurant category such as Chinese or continental etc. To calculate how good a restaurant is to a user, we use cosine similarity.

Collaborative Recommendation Engine:

Collaborative methods, based prediction works on a pair of likeable items for a specific user, based on similar items rating done by other users in past. This approach is also known as the wisdom of the crowd and assumes that users, who had similar tastes in the past, will have similar tastes in the future [7].

Hybrid Approach

Content-based and collaborative methods are suitable in many cases but also lot of limitations. The idea behind hybrid systems [8] is to combine more than two different aspects in a way to avoid the shortcomings and achieve great results by using benefits of goodness of all methods.

The authors proposed a hybrid model that combines simultaneously user-based collaborative filtering and item-based collaborative filtering by adding ratings from each technique and normalize it by a multiplication factor [9]. This proposed approach benefit simultaneously from user and item both. The accuracy of recommender systems and sparsity of data is addressed and improved by simultaneously incorporating users' correlations and items ones.

Location and Graph based Approach

In location based approach only user profile is considered. The demographic information is collected from social network like Zomato or Facebook using tools. The most important social network Facebook data is being analyzed by tools using the Facebook API and packages and libraries. The information on Facebook in form of comments and likes is most useful along with user ID, places visited, gender and age.Users clustering are used to form group of similar type of users. The database is being created in such a manner than remove unwanted data. Only the tagged locations, likes, posts post, age and gender of the user is filtered and used for further processing.

III. LITERATURE REVIEW

Other than various approaches the current work done by authors in different sub-domains useful in our approach is being presented in table below.

TABLE I. LITERATURE SUVEY

Title	Year of Publi-cation	Data-set	Methods	Improvement Scope
A Restaurant Recommendation System for Turkish Based on User Conversations [10]	2018	Chat System	Rule-based approach	Rule based approach can be made dynamic along with machine learning approach.
Social Graph based Location Recommendation using User's behavior By locating the best route and dining in best restaurant [11]	2016	Facebook and Zomato	Sentiment Mining	Accuracy is 70 % which can be improved. Parameters of filter can be changed.
Recommending Nearby Strangers Instantly Based on Similar Check-In Behaviors [12]	2015	Gowalla dataset	Kernel Density Estimation (KDE)-based user check-in probability	In spite 24* 7 cycles of check-in data weekly data can improve recommendation.
Location, Time, and Preference Aware Restaurant Recommendation Method [13]	2016	Foursquare data	User preference and restaurant popularity	Customer reviews can be used in matric for better recommendation.
You are where you eat: Foursquare check ins as indicators of human mobility and behavior [14]	2012	Flickr4 data	Location and check in based filter	Mobility behavior of individuals with that of their interest preferences combination can give realistic results.
A New Framework of a Personalized Location-based Restaurant Recommendation System in Mobile Application [15]	2013	Foursquare data	Rule Based appro-ach	Preference probability and user profile can be added with new attributes to get more effective outcomes.
Domain Knowledge Based Personalized Recommendation Model and Its Application in Cross-selling [16]	2012	Real time super market data	Customer domain knowledge-based clustering preprocessing	Personalized recommendation and product category profit can be enhanced

Review of related works in this field.

IV DATASETS

A. Restaurant & Consumer Data Dataset

The dataset was incurred from a recommender system prototype. The task was to create a top-n list of restaurants according to the consumer preferences. This dataset is from UCI Machine Learning Repository. The data comprise is of multivariate characteristic and the number of instances are 138. It is having 47 significant attributes and number of hits on web are 101442. This dataset also have missing values which can affect the classification accuracy [3].

B. Restaurant Review Dataset

This data has been collected by Adwait University of Chicago (in a project with NoemieElhadad) from http://newyork.citysearch.com/ in August 2006. Out of 17843 Restaurants, only 5531 had reviews which give us a total of 52077 reviews. Maximum number of reviews is 242 (to give better idea for distribution: 25 restaurants >=100 reviews, 103 restaurants >=10 reviews). Here is disribution of ratings (Columns = 1: Rating, 2: Review counts, 5: Percent) and cuisines (Columns = 1: Cuisine, 2: Restaurant Count, 4: Review Count - note than one restaurant can have multiple cuisines) [4].

C. Yelp Dataset

Yelp Dataset Challenge data (http://www.yelp.com/dataset challenge/) that contains actual business, user, and users' review data from the greater Phoenix, AZ metropolitan area. After filtering the data for food-related businesses and reviews, there remain > 6, 000 businesses, > 184, 000 reviews, and > 44, 000 users who gave food reviews. By using and combining various data fields, we identify both similar businesses and users to aggregate the likely sparse ratings per average user.

D. External Resource: Facebook Data

The data from above mentioned sources suffer with the problem of data sparsity, different methods to overcome or reduce data sparseness are being proposed by researchers. One method to overcome sparsity and cold start is to retrieve data from unconventional external sources that can reduce or balance missing rating data in order to get more accurate endorsements. Due to the strong recognition of social networks and the enormous amount of personal data they hold, social networks could be used as valuable external sources. One of the biggest giant in this category is Facebook. It can be a valuable alternative the only problem is vastness of data and can be handled by various information filtering approaches [6].

IV. CHALLENGES OF RECOMMENDER SYSTEM

For any particular system we must look at the problem and issues faced which can be considered as scope of improvement and can be resolved with better approach [19].

A. Cold Start

In case of user profile it's difficult to provide recommendation to naïve user as there no past experiences or ratings available. Also if we work on item profile and its mapping can't be used. This issue makes both possibilities of using user and item information a later objective. This slow nature is called Cold Start. In spite using collaborative or context based method the issue can be resolved by using Hybrid Approach.

B. Reliability

The opinion of recent user with smaller span of joining the system have very less past references compared to the users who joined the system way to earlier have rich reference and can be definitely more trustable. The user can be assigned a priority which can be used as authorization based discrimination.

C. Sparsity

In case of restaurant recommendation system the user and item base is huge. The user who is only interested in few items will provide review of specific items of his/her interest only. In such cases many items do not have users recommendations associated with it. These recommendation systems usually work on collaborative methods uses neighborhood properties and due to lack of information for user's taste they can easily relate to wrong neighborhood.

V. SELECTING FEATURE MATRIX

To refine our feature selection process, we examine the most and least significant features for users, here we are presenting few most useful feature in restaurant selection as per our review is location and similarity in user profile.

- Location (state, city) of the place matches the current location of the user
- Type of the place (e.g. Club, Restaurant/Cafe, Bar, Food/Beverages)
- The number of users who ever liked this place on Facebook

- Average rating of the place from social data/ Yelp
- Number of reviews in last month from Yelp [17]

VI. USEFUL CLASSIFIER

After reading available literature and looking at the performance of classifiers, we found that these classifiers are most capable and widely used in recommendation systems for classification purposes.

A. Support Vector Machine (SVM):

A support vector machine constructs a hyperplnein a high dimensional space, which is used for classification or regression. Hence, how happy and satisfied the user is about the restaurant corresponds to the distance of the restaurant to their hyperplane. We can evaluate model on the test set by comparing the user's actual rating of the restaurant to the result obtained by SVM [18].

B. K-Nearest Neighbor

Among collaborative filtering machine-learning algorithms, the kNearest Neighbors algorithm (k-NN) is a non-parametric method used for classification and regression. For the k-NN method, user, restaurant, user to restaurant rating and relationship is used. It's an item-based k-NN method. The possible value for rating is limited as 1, 2, 3, 4 and 5, and k-NN treats these five values as five classes [18].

VII. CONCLUSION

The restaurant recommendation system can provide users with accurate and effective restaurant information based on user`s profile information and preference. This paper analyses the various approaches and challenges of existing techniques. The collaborative and context based method when used alone have problem of cold start, reliability and sparsity. This makes the hybrid approach a better choice where we can use advantages of both methods and eliminate its drawbacks. We have also considered the feature selection method and possible metrics are presented. Various data sources are being presented. We also have uncovered areas that are open to many further improvements, and where still a scope expansion of existing research. After a detailed analysis we got following insights from the papers considered in literature review section, to make the system more dynamic rule based approach and machine learning can be used together. Any other combination of approaches makes hybrid approach looks more useful. Time(Weekend data), Filters, Customer Reviews, Mobility Behavior, User Profile and preference combination, Preference probability and user profile can be added with new attributes to get more effective outcomes. Personalized recommendation and product category profit can be used for further improvement in recommendation system and restaurant brand image. These all are the parameter if not all some from these we can incorporate in our work.

ACKNOWLEDGEMENT

This research is result of support provided by Mr. Jigyasu Dubey who mentored me and granted me insight for this research area. He assisted me through the related work, its pros and cons and scope of improvement that created a foundation for future dissertation. This research is an opportunity provided by Rajeev Gandhi technical university, Bhopal(M.P.) including the research as a part of post graduation curriculum.

REFERNCES

[1]ChristophTrattner, David Elsweiler, "Food Recommender Systems Important Contributions, Challenges and Future Research Directions" of literature exists. Table 1 provides a list of important rearXiv:1711.02760v2 [cs.IR] 10 Nov 2017

[2] A Survey on Food Recommendation Systems, Shashank S U et. al. International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 4, April 2016

[3] Blanca Vargas-Govea, Juan Gabriel GonzÃ¡lez-Serna, Rafael Ponce-MedellÃ-n. Effects of relevant contextual features in the performance of a restaurant recommender system. In RecSysâ€™11: Workshop on Context Aware Recommender Systems (CARS-2011), Chicago, IL, USA, October 23, 2011.

[4] OpenNLP (http://opennlp.sourceforge.net/) which is create on top of Maximum Entropy tools (http://maxent.sourceforge.net/) inspired from thesis work of AdwaitRatnaparkhi (UPenn 1998). The models are trained on Penn Treebank.

[5] Yelp Food Recommendation System, SumedhSawant and Gina Pai, http://cs229.stanford.edu/proj2013/ SawantPai-YelpFoodRecommendationSystem.pdf

[6] Utilizing Facebook Single and Cross Domain Data for Recommendation Systems BrachaShapira1 ,Lior Rokach1 , Shirley Freilichman2, http://citeseerx.ist.psu.edu/ viewdoc/download? doi=10.1.1.408.763 &rep=rep1&type=pdf

[7] Personalized Food Recommendations Exploring Content-Based Methods Jorge Almeida, jorge. https://fenix.tecnico.ulisboa.pt/downloadFile/1126295043834284/resumo.pdf

[8] Robin Burke. Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction, 12(4):331–370, 2012.

[9] G. Badaro, H. Hajj, W. El-Hajj and L. Nachman, "A hybrid approach with collaborative filtering for recommender

systems," 2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, 2013, pp. 349-354.
[10]M. Elıfoğlu and T. Güngör, "A restaurant recommendation system for Turkish based on user conversations," *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir, 2018, pp. 1-4.
[11] AnuTaneja, Prashant Gupta and AyushGarg, "Social Graph based Location Recommendation using User's behavior By locating the best route and dining in best restaurant", 2016, Fourth International on Parallel, Distributed and Grid Computing.
[12] X. Qiao et al., "Recommending Nearby Strangers Instantly Based on Similar Check-In Behaviors," in IEEE Transactions on Automation Science and Engineering, vol. 12, no. 3, pp. 1114-1124, July 2015.
[13] M. A. Habib, M. A. Rakib and M. A. Hasan, "Location, time, and preference aware restaurant recommendation method," 2016 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, 2016, pp. 315-320.
[14] G.B. Colombo, M.J. Chorley, M.J. Williams, S.M. Allen, R.M. Whitaker, "You are where you eat:Foursquare checkins as indicators of human mobility and behaviour", International Workshop on the Impact of Human Mobility in Pervasive Systems and Applications 2012, Lugano
[15] Z. Wen-ying and Q. Guo-ming, "A new framework of a personalized location-based restaurant recommendation system in mobile application," 2013 International Conference on Management Science and Engineering 20th Annual Conference Proceedings, Harbin, 2013, pp. 166-172.
[16]LinglingZhanga,b,*, CaifengHua , QuanChena,b, YibingChena,b,YongShib, "Domain Knowledge Based Personalized Recommendation Model and Its Application in Cross-selling" International Conference on Computational Science, ICCS 2012
[17] Qiaosha Han, Vivian Lin and Wenqing Dai, Restaurant Recommendation for Facebook Users, s229.stanford.edu/.../HanLinDai RestaurantRecommendationForFacebookUsers.pdf
[18] Rui Jiang, A Customized Real Time Restaurant Recommendation System, https://etda.libraries.psu.edu/files/final_submissions/8189, 2015
[19] Khusro, Shah & Ali, Zafar&Ullah, Irfan. (2016). Recommender Systems: Issues, Challenges, and Research Opportunities. 1179-1189. 10.1007/978-981-10-0557-2_112.

Shreya Joshi, Post graduate Scholar specialisation in Artiicial Intelligence in Shri Vaishnav Institute of of Science & technology, Indore(M.P.) affiliated to R.G.T.U. Bhopal(M.P.).
3.E (C.S.E.) from RGTU

Jigyasu Dubey, Associate professor, Departmentof Information Technology, Shri Vaishnav Vidyapeeth Vishwavidyalaya Indore(M.P.)
Ph. D. Computer Engineering