A Study on Restaurant Recommendation System using Feature Selection

Niketa Gupta[#1], Arvind Upadhyay[*2]

#1Research Scholar, CSE , IPS Academy ,Institute of Engineering & science Indore, Madhya Pradesh ,India

*2Associate Professor, CSE, IPS Academy ,Institute of Engineering & science Indore, Madhya Pradesh ,India

#1niketa2112.ng@gmail.com *2upadhyayarvind10@gmail.com

Abstract-In order to predict consumer demands for products, recommendation systems are frequently used. In particular, they assist the consumers to find movies, restaurants, mobiles and goods. One of the requirements of today's world is cuisine. The vast number of options for restaurants and shortage of restaurant knowledge is also a challenge to the needs of person's when selecting a restaurant. Restaurant Recommendation system is item-based reviews that suggest restaurants to users in accordance with their interests. It is also helpful in generating restaurant profit by recommending different cuisine options or suggesting restaurants to customers dependent on the type of cuisine. The current restaurant recommendation systems are analyzed in this research and the various recommendation and classification techniques are compared. In addition, we can see the benefits and challenges of each technique.

Keywords: Recommendation System, Content Filtering, Collaborative Filtering, Yelp, Zomato,   Machine Learning, Classification Techniques

1. Introduction

With the emergence of Facebook, Netflix, Amazon and several other web applications, recommendation systems have become more and more important in our lives over the last few decades. These business organizations, which are also recognized for their customized customer interactions, recommendation systems are significant and useful resources. All these businesses gathers and analyzes client statistical profile and relates it to past sales, product reviews, and user experience records. These specifics are then used to estimate how clients will rank sets of similar items, or how probably an additional item is to be purchased by a client. Recommendation systems are popular applications for machine learning which have been widely researched everywhere [1].

The goal of implementing the recommendation systems is to minimize the extensive knowledge base by retrieving the most important information and services from a large dataset and thus providing customized services. Recommendation mechanisms, however still include numerous problems, specifically in their precision and predictions. The Recommendation system is capable of predicting whether a specific user will choose an object or not based on the characteristics of the client. Recommendation systems are useful to both service providers and users [2]. In an online shopping, they help to improve the efficiency of finding and choosing items.

Recommendation systems have noticed their implementation in numerous restaurant recommendation services by using characteristics such as Facebook check-ins, user reviews and behavioral trends. The restaurant recommendation system is a really popular service whose accuracy and sophistication keeps growing every day [3]. The launch of smart phones, web 2.0 and internet services such as 3G has made this accessible to any user. In order to enhance precision, several authors have found techniques like collaborative filtering, content-based filtering and hybrid filtering [4]. We will go through different approaches of restaurant recommending systems and datasets in this paper. We will discuss how they function, identify their theoretical foundation for each of them, and analyze their strengths and limitations.

2. Background Study

2.1 Recommendation Approaches

Content-Based Filtering:

One of the most commonly used and studied approaches of recommendation is content-based filtering. The user process model, in which the preferences of consumers are interpreted from the objects interacted by users, is a core aspect of content based filtering. In content-based filtering technique, Recommendations are rendered depending on customer profiles using features derived from the content of items reviewed by the customer in the past. The customer is advised to include products mainly relevant to the strongly rated items. CBF encourages more personalized customization such that the recommendation system can independently decide the best suggestions for each user [5].

Collaborative filtering:

The recommendations are inferred in the collaborative-based recommendation framework by taking users' habits and user expectations into account. The system can also correctly suggest complicated items without considering large emphasis on the item's characteristics and contents. The collaborative filtering method works by creating a repository of user preferences for objects. To offer recommendation, it then associates users with specific personal preferences by measuring correlations between their identities. CF is domain-independent, i.e. no processing of error-prone objects is needed. Because the evaluations are performed by individuals, CF takes actual processing into consideration [6].

Knowledge-based:

Knowledge-based recommendation systems are equivalent to content-based recommendation systems in that they build using a tailored model for each individual to locate appropriate things for them. However, rather than use the customer's ranking history to create the model, they ask the user, either via keywords or a set of criteria, to clearly mention their interests. This form of recommendation system aims to recommend items based on observations about the desires and needs of a consumer. Knowledge-based recommendations operate on functional knowledge recommendations: they have insight about how a specific item serves a specific user

requirement, and can thus reason about the correlation between a requirement and a potential recommendation [7].

Hybrid Recommendation System:

In the actual world, a hybrid system is far more appropriate since a mixture of parts gathered from variety methods will solve different conventional limitations. This system is a mixture of a recommendation system based on content and a recommendation system based on collaborative. The recommendation is focused on the performance of comparisons made by continuously tracking the browsing and analyzing patterns of the same types of customers in this recommendation framework [8].

2.2 Related Work

The restaurant is a public area where one meal or different foods are served. Currently, several apps include details about restaurants, such as the name of the restaurant, the restaurant menu and the price of the meals in the restaurant. The framework that is widely used in restaurant applications is the recommendation framework. The recommendation framework uses methods in all aspects of different fields including data mining, machine learning, database, tests of similarity, etc. It produces customer satisfaction estimates and/or suggests an object to a customer. Table 2.1 shows the description of the present recommendation system methods reviewed under the literature.

Table 2.1 Summary of research work done on recommender systems

S.No	Paper Title	Year of Publication	Dataset	Methods	Key Points
1	Restaurant Recommendation System for User Preference and Services Based on Rating and Amenities [9]	2019	Tripadvisor.com	Natural Language Processing	Amenities selection method yields high accuracy
2	Extraction-Based Text Summarization and Sentiment Analysis of Online Reviews Using Hybrid Classification Method[10]	2019	Zomato dataset	Random Forest	Accuracy of this machine model is 92%
3	Novel Recommendation Systems Using Personalized Sentiment Mining [11]	2018	Yelp and Movie lens dataset	Naive Bayes and Random Forest	Sentiments have been introduced to customize the recommended systems.

4	Predicting Purchase Intension According to Fan Page User's Sentiment [12]	2017	Theoretical study	Inter Transaction Association Rule Mining	Theoretical study demonstrate the functioning of the implemented model
5	Location, Time, and Preference Aware Restaurant Recommendation Method [13]	2016	Foursquare data	User preference and restaurant popularity	For improving recommendation uses customer reviews
6	Social Graph based Location Recommendation using User's behavior By locating the best route and dining in best restaurant [14]	2016	Facebook and Zomato	Sentiment Mining	Similarity with Zomato's rating 70%

3. Datasets

Zomato Dataset

Zomato is a restaurant review system beginning in 2008. It operates in 23 countries at the moment. Zomato is the right platform for restaurant assessment since it has the largest number of restaurants enrolled worldwide. The research was conducted using the Zomato API initially the raw information gathered was in JSON format subsequently stored in comma separate values ' CSV '. It has relevant data about the name of the restaurant, ratings, cuisine, location and other demographics in which an ID identifies every restaurant uniquely. The Zomato API (http:/zomato.com/api) was published with the intent of helping developers to build innovative mobile and web applications using Zomato's data. To use the API, developers can obtain the most upgraded information available on Zomato in real time specifically from inside their apps. They can access information like location coordinates, reviews, restaurant listings, discounts, photographs, menu options and relatively much anything else that we have for our consumers [15].

Yelp Dataset

Online reviews from Yelp are reliable source of information for users to select where to visit or what to eat among several choices available. Data from Las Vegas, Phoenix, Madison, Edinburgh and Waterloo is included in the dataset and gives details about 42,153 firms, 320,002 features of company, 31,617 sets of check-ins, 403,210 hints and 1,125,458 reviews of text. Specifically, the dataset is made up of five files, one for each type of object: business, review, user, check-in and tip. One json-object-per-line is for each file [16]. Therefore a company is depicted in 'business.json' as a json object file that describes the business ID, Its name, location,

stars, number of reviews, opening hours, etc. The companies listed in the Yelp dataset belong to the company various types, such as shopping, restaurants, traveling and hotels, etc. The text feedback for various types of companies can be very specific.

4. Classifiers

Support Vector Machine Classifier:

Support Vector Machine (SVM) is regarded as the significant classifier that generates the reliable results in concerns with face recognition. For the nearest qualified instances, they accomplished by constructing a hyper plane with highest possible Euclidean distance. Vapnik is the founder and developer of the Support Vector Machines (SVM). The SVM is well known among researchers because of the characteristics and reliability of the empirical results. SVM is a robust machine learning system that is more capable of analyzing the hidden consistency of different data sets. Its main feature is to boost the ability of learning machines to generalize.[17]

Naive Bayes Classifier

One of the popular supervised learning methods that are often used for classification purposes is a Naive Bayesian classifier. A classifier is called naive because it recognizes the possibilities that are really connected are not based on the further. Naive Bayes is a system of classification invented by Thomas Bayes. This system, which learns from information and predicts the likelihood of each class [18].Naive Bayesian classifier is a probabilistic classifier and claims significantly that the attributes are isolated from each other depending on the theorem of Bayes. Traditional applications of this classifier involve spam filtering, face detection, prediction of sentiments, etc.

Random forest

Tin Kam Ho first presented the random forest algorithm. Random forest is a machine learning ensemble classifiers tend to create many or more decision trees and trees combines them in order to obtain a more reliable and precise prediction. N random forest for training data choose N arbitrarily data with the permissible replacement data in training. After building numerous or several trees, it conducts the majority voting [19].

5. Identifying Evaluation Characteristics

Consumers are now becoming more refined and challenging, along with changing industry patterns. Consumer satisfaction is an important business challenge, as merchants. They also recognized that favorable user experience is necessary for a successful long-term activity. If clients have an outstanding dining experience, then they will suggest the restaurant to many others and gives positive feedback or being a regular consumer. To optimize our method of characteristics identification, we investigate the most and least important attributes of restaurants, which we are discussing here some of the most valuable restaurant selection features are ambience, price range and food quality.

Ambience: Every individual wants to visit a restaurant where they experience a sense of class, pleasure, beauty, convenience and, most of all, protection. Ambience include parking, Wi-Fi, T.V., dining space and credit card payment facility.

Price Range: The price of an item will influence the consumer satisfaction level since it has a related sense of value. Interpretations of unfair pricing by a consumer relate to negative outcomes, For example, a lower level of purpose to revisit, frustration and bad publicity.

Food Quality: The quality of food is an essential element and has repeatedly been shown to be a primary value that a client considers while selecting restaurant.The quality of food is considered a very important element for consumer loyalty and the desire of clients to revisit.

6. Conclusion

A few observations can be reached on the basis of the results of the systematic literature on studies published in recent years using all recommendations system approaches. We have evaluated the current systems and identified that they are not sufficiently flexible. Data science and collaborative filtering can therefore guarantee the recommendation system's high scalability and power. We analyze that Zomato and Yelp dataset are two most widely dataset used for research in restaurant recommendation system. Each methods has its advantages and disadvantages among different methods, and we recognize ensemble model to be the most efficient as we can customize data-based operations and integrate advantages of more than one technique.

References

[1] Satya Prakash Sahu, Anand Nautiyal, Mahendra Prasad*,"Machine Learning Algorithms for Recommender System - a comparative analysis",* International Journal of Computer Applications Technology and Research Volume 6–Issue 2, 97-100, 2017, ISSN:-2319–8656.

[2]F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, *"Recommendation systems: Principles, methods and evaluation"*, Egyptian Informatics Journal,Volume 16, Issue 3, 2015, Pages 261-273, ISSN 1110-8665

[3]Pu P, Chen L, Hu R*, "A User-Centric Evaluation Framework for Recommender Systems",* In Proceedings of the fifth ACM conference on Recommender Systems (RecSys'11), ACM, New York, NY, USA; 2011. p. 57–164.

[4] R. Chen, Q. Hua, Y. Chang, B. Wang, L. Zhang and X. Kong, *"A Survey of Collaborative Filtering-Based Recommender Systems: From Traditional Methods to Hybrid Methods Based on Social Networks,"* in IEEE Access, vol. 6, pp. 64301-64320, 2018, doi: 10.1109/ACCESS.2018.2877208.

[5] R.Manjula and A. Chilambuchelvan, *"Content Based Filtering Techniques in Recommendation System using user preferences,"* Int. J. Innov. Eng. Technol., vol. 7, no. 4, 2016.

[6] U. Farooque, B. Khan, A. B. Junaid and A. Gupta, *"Collaborative Filtering based simple restaurant recommender,"* International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2014, pp. 495-499, doi: 10.1109/ IndiaCom. 2014. 6828187.

[7] S. Subbotin, O. Gladkova and A. Parkhomenko, *"Knowledge-Based Recommendation System for Embedded Systems Platform-Oriented Design,"* IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, 2018, pp. 368-373, doi: 10.1109/STC-CSIT.2018.8526659.

[8] G. Badaro, H. Hajj, W. El-Hajj and L. Nachman, *"A hybrid approach with collaborative filtering for recommender systems,"* 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, 2013, pp. 349-354.

[9] 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)*, C*hennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862048.

[10] N. Yadav, R. Kumar, B. Gour and A. U. Khan, *"Extraction-Based Text Summarization and Sentiment Analysis of Online Reviews Using Hybrid Classification Method,"* 2019 Sixteenth International Conference on Wireless and Optical Communication Networks (WOCN), Bhopal, India, 2019, pp. 1-6, doi: 10.1109/WOCN45266.2019.8995164.

[11] B.S. Govind, R. Tene and K. L. Saideep, *"Novel Recommender Systems Using Personalized Sentiment Mining,"* 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, 2018, pp. 1-5, doi: 10.1109/CONECCT.2018.8482394.

[12] L. Kao and Y. Huang, *"Predicting purchase intention according to fan page users' sentiment,"* IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, 2017, pp. 831-835, doi: 10.1109/SMC.2017.8122712.

[13] M. A. Habib, M. A. Rakib and M. A. Hasan, *"Location, time, and preference aware restaurant recommendation method,"* 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, 2016, pp. 315-320.

[14] Anu Taneja, Prashant Gupta and Ayush Garg, *"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.

[15] Shina, S. Sharma and A. Singla, *"A Study of Tree Based Machine Learning Techniques for Restaurant Reviews,"* 4th International Conference on Computing Communication and Automation (ICCCA)*, Gre*ater Noida, India, 2018, pp. 1-4, doi: 10.1109/CCAA.2018.8777649.

[16] Nabiha Asghar, *"Yelp Dataset Challenge: Review Rating Prediction",* CoRRabs/1605.05362 (2016).

[17] Yu Boya, Zhou Jiaxu, Zhang Yi, Cao Yunong, "Identifying Restaurant Features via Sentiment Analysis on Yelp Reviews", ARXIV, 2017.

[18] R. A. Laksono, K. R. Sungkono, R. Sarno and C. S. Wahyuni, *"Sentiment Analysis of Restaurant Customer Reviews on TripAdvisor using Naïve Bayes,"* 2019 12th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 2019, pp. 49-54, doi: 10.1109/ICTS.2019.8850982.

[19] H. Valecha, A. Varma, I. Khare, A. Sachdeva and M. Goyal, *"Prediction of Consumer Behaviour using Random Forest Algorithm,"* IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gorakhpur, 2018, pp. 1-6,doi:10.1109/UPCON.2018.8597070