

Survey on Restaurant Review Classification and Recommender

K.Soni*, M.Viren**, R.Jaikumar***, V.Naveen****

*(Student, Department of Computer Engineering,
Watumull Institute Of Electronic Engineering And Computer Technology, Ulhasnagar
Email: sonikukreja2000@gmail.com)

** (Student, Department of Computer Engineering,
Watumull Institute Of Electronic Engineering And Computer Technology, Ulhasnagar
Email: virenmakhija8901@gmail.com)

*** (Student, Department of Computer Engineering,
Watumull Institute Of Electronic Engineering And Computer Technology, Ulhasnagar
Email: jaikumarramnani122@gmail.com)

**** (Assistant Professor, Department of Computer Engineering,
Watumull Institute Of Electronic Engineering And Computer Technology, Ulhasnagar
Email: vaswani.naveen@gmail.com)

Abstract:

Recommendation system is the system in which there is huge amount of data about what are the interests and the preferences of different people about a particular product or place. Whenever we purchase a product or we want to go somewhere we always look for the ratings for that product or place. Machine learning is important in restaurant recommendation systems because it filters customers based on their previous actions. Feedbacks, ratings, reviews, and restaurant reputation, tastes provided by other customers can be used to perform restaurant recommendations. This is the survey for a restaurant recommendation system, which incorporates narrowed-down information of algorithms and methodologies that are already being used for study in this domain.

Keywords — SVM, naïve bayes, recommendation, classification, sentiment analysis.

I. INTRODUCTION

Influence of social plays an important role in our life as we depend on the opinions of the people. The recommendation system provides a wide range of options to choose the best option among the available ones. This system act as a software tool which give suggestions about the products for the users by using the machine learning techniques. The system identifies the similarities of the products and the users and recommend product or place according to their interests. The key to today's most successful online companies is a recommendation system with the best algorithms. Review services such as Yelp, Google Reviews, and others allow users and businesses to interact with one another. Reviews and

ratings are valuable sources of information, but there are significant challenges in extracting relevant information and forecasting the future using analysis and correlation of existing data. Customers review thousands of restaurants and businesses every day. A recommendation system is one that can predict a user's future preferences for a group of products and provide recommendations based on that prediction. We are presenting a survey on available approaches and their evaluation.

II. APPROACHES FOR RECOMMENDATION SYSTEM

One of the following methods can be used to build a recommendation system:

A. CONTENT-BASED FILTERING

Content-based Filtering is a Machine Learning approach that makes decisions based on feature similarity. This strategy is usually applied in recommender systems, which are algorithms that advertise or recommend items to people based on information gathered about them. This strategy is entirely based on comparing user preferences to product attributes. The products that have the most features in common with the user's interests are the ones that are recommended.

B. COLLABORATIVE FILTERING

Collaborative Filtering is a Machine Learning technique for detecting data set correlations. This method is commonly used in recommender systems to detect similarities between user data and items. This means that if Users A and B both like Product A, and User B also likes Product B, the system may recommend Product B to User A. The model keeps track of what products users like and their characteristics in order to see what products users who like similar products enjoyed. The model then makes recommendations based on this information.

C. HYBRID APPROACH

Content-based and collaborative methods are appropriate in many cases, but they have a number of limitations. The idea behind hybrid systems is to combine more than two different aspects in order to avoid flaws and achieve great results by combining the best of all methods.

D. LOCATION AND GRAPH BASED APPROACH

In location-based approach recommendation is done by considering the user's location by tracking his location through gps. The restaurant which is nearby to the user, whose ratings are high will be recommended. A group of similar type of users is formed using user's clustering. The database is being built in such a way that unwanted data is removed. Only the user's tagged locations, likes, posts, age, and gender are filtered and used for further processing.

E. SENTIMENT BASED APPROACH

A sentiment analysis model is used for classifying reviews as 0 or 1, with 0 representing negative and 1 representing positive, respectively. A review can be classified based on the average semantic orientation of the review's phrases, which include adverbs and adjectives. When semantic orientation and sentiments are combined, it is expected that there will be an efficient value. The review is only recommended if the mean is positive; otherwise, it is not.

III. ALGORITHMS

A. SUPPORT VECTOR MACHINE

The powerful machine learning algorithm known as support vector machine can be used for regression and classification techniques. SVM is primarily concerned with classifying data points using the hyperplane. In an n-dimensional space, a hyperplane is simply the decision boundary. The hyperplane can be identified by plotting the classes as datapoints in an n-dimensional space, and the distance between the datapoints can be used to predict the hyperplane. SVM's main goal is to divide datasets into classes so that new upcoming data can be plotted using the hyperplane. This algorithm will improve the accuracy of classifying datapoints based on specific factors. SVM is classified into two types: linear SVM and non-linear SVM. Linear SVM is based on a straight line to classify a dataset into two classes, whereas non-linear SVM does not use a straight line to classify a dataset into two classes.[1][3][4][5][9][10][13]

B. NAÏVE BAYES

One of the most effective supervised machine learning algorithms for classification problems is the Naive Bayes algorithm. This algorithm is based primarily on the Bayes theorem. The Bayes theorem is based on the concept of conditional probabilities and prior experience. It is primarily used for text classification based on results probabilities. It will predict the results quickly by employing the conditional probability concept. Conditional

probability finds the results based on prior knowledge and predicts the final results with the greatest degree of certainty. [1][8][9][10][11] Bayes' theorem can be expressed as follows:

$$P(A|B) = P(B|A) P(A) / P(B).$$

C. K-NEAREST NEIGHBOR

A supervised ML model is the k-nearest neighbours (KNN) machine learning algorithm. It is a straightforward and efficient model that can be applied to classification tasks. The K-NN model is based on the assumption that similar items can be found nearby. That is, it operates on the simple principle that "similar things are closer to each other." The distance can be calculated using a method known as "Euclidean distance." [5][6]

D. LOGISTIC REGRESSION

Although the name of this algorithm includes the word "regression," it is only used for classification. When the predicted variable is categorical (or labelled), it is used. This model is based on a

probability model. We start with linear regression, and then we need to set a threshold based on which classification can be done. If that value is less than 0.5, we classify the sample. If that value is greater than 0.5, we move the sample to a different class. [5]

E. APRIORI

Apriori is a mining algorithm for frequent item sets and generating association rules. This algorithm has traditionally been known for its utility in market basket analysis. It employs a single minimum support as a criterion to identify a set of frequently occurring itemset. It narrows the search space and reduces the number of frequent patterns that are generated. Using a single minimum support, on the other hand, assumes that all items in the dataset are of the same nature and have the same frequencies, which is not the case in real-world applications. [7]

IV. LITERATURE SURVEY

The tables below show the work done by the others authors which are useful and related to our work.

TABLE I
COMPARISON OF DIFFERENT ALGORITHMS FOR RESTAURANT REVIEW CLASSIFICATION AND RECOMMENDATION

S.No	Title	Year of Publication	Dataset	Algorithm	Accuracy	Improvement
1	Restaurant Recommendation System Using Support Vector Machine and Naive Bayes Classifier Machine Learning Algorithms	2020	Real time data	support vector machine, naïve bayes, Content-Based Recommendation System, Collaborative Filtering Technique.	Collaborative Filtering gives 87% accuracy, Content Based Filtering gives 93% accuracy and Hybrid Filtering gives 96% accuracy.	Deep learning and neural networks will be used in conjunction to find non-linear patterns.

2	Restaurant Recommendation System Using Customer's Data Analysis	2018	Zomato	---	---	---
3	Aspect Based Sentiment Classification for Restaurant Reviews Using Deep Learning	2021	---	Support vector machine, Content-Based Recommendation System	---	To deal with the issue of cold beginnings from certain criticism.
4	Cuisine Recommendation, Classification and Review Analysis Using Supervised Learning	2020	Real time data	Support vector machine	Proposed system gives an average performance of 91 %.n.	---
5	Restaurant Reviews Classification Using NLP Technique	2019	Real time data	K-NN, Logistic Regression, Support Vector Machine	K-NN gives 75 accuracy, Logistic Regression gives 80 accuracy, Support Vector Machine gives 70 accuracy.	---
6	A Survey on Restaurant Recommendation System Based on Feature Selection and Classification Method	2018	Yelp.com	K-NN	---	Personalized recommendation and product category profit can be used for further improvement in recommendation system and restaurant brand image
7	Food Recommendation Using	2019	Real time data	Modified classifier algorithm	---	The algorithms can work fine with the larger

	Classifier and Modified Apriori Algorithm					datasets
8	Restaurant Review Classification Using Naives Bayes Model	2021	---	Naives Bayes Model	Naive Bayes technique gives an accuracy score of 77.5 percent	The accuracy of sentiment analysis might be enhanced by combining the use of semantic resources such as WordNet and SentiWordNet with a statistical method
9	Restaurant Review Classification and Analysis	2020	Zomato.com	Support Vector Machine (SVM), Naive Bayes Model	MultinomialNB model performs with an accuracy of 86.6%, SGD Classifier performs with an accuracy of 85.8% and Random Forest performs with 82.5%.	Neural network architectures can be used to make the summarization of text more accurate and readable and to the point
10	Predicting The Helpfulness of Online Restaurant Reviews Using Different Machine Learning Algorithms	2019	Yelp.com	Naïve Bayes, Support Vector Machine, Fuzzy Domain Ontology	Naïve Bayes gives an accuracy rate from 67.68 to 71.20, Support Vector Machine gives an accuracy rate from 69.52 to 71.20, Fuzzy Domain Ontology gives an accuracy from 79.59 to 81.62.	---

11	Sentiment Analysis of Customer Feedback on Restaurant Reviews	2019	yelp.com	Naive Bayes classifier	---	---
12	Restaurant Recommendation System Using Machine Learning Algorithms	2021	kaggle dataset	Collaborative filtering technique	---	---
13	Opinion Recognition System for Restaurants	2020	Zomato dataset	Support Vector Machine	Support Vector Machine gives an accuracy of 88.2%	It's exciting to see how far this work may be extended to better the scientific classification of restaurants, when fresh views are discovered while doing surveys.
14	Developing Restaurant Recommendation System With Neural Collaborative Filtering Method	2021	Kaggle dataset	Neural Collaborative Filtering technique	---	---
15	Restaurant Recommender System Based on Sentiment Analysis	2021	TripAdvisor.com	NLP techniques	Proposed system can provide users with 92.8% accuracy in the Top5 mode	group-based restaurant recommendation, the similarity between users' preferences and their favorite foods in each restaurant should be specified.

V. DATASETS

A. *YELP*

Yelp is well recognized for its online business reviews. It began as an internet website and has since evolved into an app for user-generated ratings of local businesses and communities. Customers can use Yelp for a variety of things other than submitting reviews, such as locating events, learning about new businesses, and communicating with other Yelpers. Users who submit evaluations and engage with one another through the website and application are referred to as "Yelpers." On average, Yelp receives around 25 million unique visitors via its app and 72 million unique visitors via its website.[6][11][10]

B. *KAGGLE*

Kaggle is a resource platform for data scientists, both experienced and aspiring. For anyone wishing to get started or contribute to projects in order to enhance their skills or build up their data science portfolios, they provide 19,000 public datasets and 200,000 public notebooks. They provide a Jupyter Notebooks environment that requires no setup and is fully customizable, as well as free GPUs and a massive archive of community-published data and code.[12][14]

C. *TRIP ADVISOR*

TripAdvisor is a travel information and booking website with reviews, photos, and discussions about hotels and resorts all around the world. The most popular site for hotel and restaurant reviews is still TripAdvisor. According to an Uberall survey, more than three out of four (76%) UK consumers have examined online reviews on TripAdvisor, Google, Facebook, or Yelp, but TripAdvisor is the clear victor in terms of popularity, with 68 percent of those surveyed stating the site is their "go-to" for reviews.

Customers' reliance on internet reviews for the hospitality industry is also progressively increasing, according to the report. Consumers have checked online reviews for hotels 79 percent of the time and restaurants 77 percent of the time.[15]

D. *ZOMATO*

Zomato is a search and discovery engine for restaurants that was created in 2008. India, Australia,

and the United States are among the 23 nations where it now operates. It includes user reviews and ratings, as well as restaurant information such as scanned menus and images obtained by local street teams.[2][9][13]

VI. CONCLUSION

Based on other users' reviews and preferences, the restaurant suggestion system can present consumers with reliable and useful restaurant information. The many methods and challenges of existing methodologies are examined in this research. When used alone, the collaborative and context-based method has issues with cold start, reliability, and sparsity. Since a result, the hybrid strategy is a preferable option, as we can combine the benefits of both methods while avoiding their shortcomings. After a thorough examination of the papers reviewed in the literature review section, we discovered that rule-based approaches and machine learning can be combined to make the system more dynamic.

REFERENCES

- [1] Dr.J.Jeyabharathi 1, Dr.K.Loheswaran 2, V.Subba Ramaiah 3, T.Kumaravel 4 "Restaurant Recommendation System Using Support Vector Machine and Naive Bayes Classifier Machine Learning Algorithms" International Journal of Future Generation Communication and Networking Vol. 13, No. 4, (2020), pp.3710-3714
- [2] Bhagyashree Basudkar 1, Shruti Bagayatkar 2, Meghana Chopade 3, Sachin Darekar 4 "Restaurant Recommendation System Using Customer's Data Analysis" Volume 8 Issue No.4,2018 IJESC
- [3] Mr. Arvind Gautam 1, Prof. Vandana Navale 2 "Aspect Based Sentiment Classification for Restaurant Reviews Using Deep Learning" International Journal Of Advance Scientific Research And Engineering Trends Volume 6 || Issue 6 || June 2021 || ISSN (Online) 2456-0774
- [4] Aruna Pavate, Priti Mishra, Anita Chaudhary, Mansi Shah, Pranav Nerurkar "Cuisine Recommendation, Classification and Review Analysis Using Supervised Learning" 2020 IEEE International Conference on Convergence to Digital World-Quo Vadis (ICCDW 2020)
- [5] Nuradha Tutika, M Y V Nagesh "Restaurant Reviews Classification Using NLP Techniques" Journal of Information and Computational Science Volume 9 Issue 11 – 2019 ISSN: 1548-7741
- [6] Shreya Joshi,Jigyasu Dubey "A Survey on Restaurant Recommendation System Based on Feature Selection and Classification Method" Volume VII, Issue X October - 2018 ISSN (Online) 2278-8034
- [7] Sreyam Dasgupta, Suparna Karmakar "Food Recommendation using Classifier and Modified Apriori Algorithm" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-12, October 2019
- [8] Kothapally Nithesh Reddy, Dr. B. Indira Reddy "Restaurant Review Classification Using Naives Bayes Model" Journal of University of Shanghai for Science and Technology ISSN: 1007-6735

[9] Yi Luo 1 and Xiaowei Xu 2 "Predicting the Helpfulness of Online Restaurant Reviews Using Different Machine Learning Algorithms: A Case Study of Yelp" Published: 25 September 2019

[10] Dhiraj Kumar, Gopesh2, Avinash Choubey, Ms.Pratibha Singh "Restaurant Review Classification and Analysis" Vol 11, Issue 8, August/2020 ISSN NO:0377-9254

[11] Spoorthi C, Dr. Pushpa Ravikumar, Mr. Adarsh M.J "Sentiment Analysis of Customer Feedback on Restaurant Reviews" International Journal of Research in Advent Technology, Vol.7, No.6S, June 2019 E-ISSN: 2321-9637

[12] Asst. Prof. Sathya Seelan K, Asst. Prof. K. Prabhu "Restaurant Recommendation System Using Machine Learning Algorithms" Sathyaseelan K,2021, 9:3, ISSN (Online): 2348-4098, ISSN (Print): 2395-4752

[13] Dr. Kanika Gupta 1, Sagar Soni 2, Saumya Mathur 3, Vaishali Chauhan 4 "Opinion Recognition System for Restaurants" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume: 07 Issue: 04 | Apr 2020, p-ISSN: 2395-0072

[14] Ilham Huseyinov and Tolga Hamitovoli "Developing Restaurant Recommendation System with Neural Collaborative Filtering Method" 2021 JETIR August 2021, Volume 8, Issue 8

[15] Elham Asani, Hamed Vahdat-Nejad, Javad Sadri "Restaurant recommender system based on sentiment analysis" Machine Learning with Applications 6 (2021) 100114