

Restaurant Reviews 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 systems are intended to give the user with a personalized product or service. Its goal is to determine users' future activities based on their previous behavior and give recommendations accordingly. As a restaurant, it is necessary to pay greater attention to customer service by continuously improving it. The scenario has an impact on the restaurant's brand image, which is influenced by customer satisfaction. Restaurants also need recommendation systems in order to attract more customers to experience popular, well-known dishes in the restaurant. In this paper, we present the recommendation system for restaurants which will first perform sentiment analysis on the restaurant reviews using machine learning algorithm and then recommend restaurants according to their positive percentage.

Keywords — recommendation systems, sentiment analysis, logistic regression, machine learning.

I. INTRODUCTION

Recommendation systems have recently gained popularity due to their application in a variety of domains. A recommendation system is a collection of algorithms that learn from input and, after processing, can recommend products or places to customers. Customers rate and write reviews about restaurants based on their level of satisfaction. These ratings and reviews help other customers in deciding whether or not to visit those restaurants. These ratings are also useful for restaurant owners who want to improve their business by making changes based on customer feedback.

Reviews and ratings are useful sources of information, but there are significant challenges in

extracting relevant information and predicting the future using analysis and correlation of existing data. Customers review thousands of restaurants and businesses every day.

The major goal of the work proposed in this paper is to improve the user experience by analysing restaurant reviews, labelling them as positive or negative, and recommending restaurants based on the percentage of positive reviews calculated, so that a user can quickly discover about the restaurant.

II. LITERATURE SURVEY

The tables below show the work done by the others authors which are useful and related to our work. [16]

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

S.No	Title	Year of Publication	Dataset	Algorithm	Accuracy	Improvement
1	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.
2	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
3	Restaurant Recommendation System Using Machine Learning Algorithms	2021	kaggle dataset	Collaborative filtering technique	---	---
4	Developing Restaurant Recommendation System With Neural Collaborative Filtering Method	2021	Kaggle dataset	Neural Collaborative Filtering technique	---	---
5	Restaurant Recommender System	2021	TripAdvisor.com	NLP techniques	Proposed system can provide users	group-based restaurant recommendation

	Based on Sentiment Analysis				with 92.8% accuracy in the Top5 mode	, the similarity between users' preferences and their favorite foods in each restaurant should be specified.
6	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.
7	Cuisine Recommendation, Classification and Review Analysis Using Supervised Learning	2020	Real time data	Support vector machine	Proposed system gives an average performance of 91 %.	---
8	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
9	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.
10	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.	---
11	Food Recommendation Using Classifier and Modified Apriori Algorithm	2019	Real time data	Modified classifier algorithm	---	The algorithms can work fine with the larger datasets
12	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.	---
13	Sentiment Analysis of	2019	yelp.com	Naive Bayes classifier	---	---

	Customer Feedback on Restaurant Reviews					
14	Restaurant Recommendation System Using Customer's Data Analysis	2018	Zomato	---	---	---
15	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

III. PROPOSED SYSTEM

A. Algorithms

Below table shows Various algorithms which we used for comparison, their accuracy and confusion matrix.

TABLE II
 COMPARISON OF DIFFERENT ALGORITHMS BASED ON THE ACCURACY

S.No	Name	Category	Accuracy
1	Decision Tree	Supervised	77.33%
2	Random Forest	Supervised	61.66%
3	KNN	Supervised	65%
4	Gaussian Naïve Bayes	Supervised	72.33%
5	Logistic Regression	Supervised	84%

6	Gradient Boosting	Supervised	67.66%
7	Bagging Classifier	Supervised	71.66%
8	Support Vector Machine	Supervised	84%
9	XGB Classifier	Supervised	78.66%
10	Multilayer Perceptron	Supervised	80.33%
11	Bernoulli NB	Supervised	84%

B. Datasets

The following are the datasets we used for our project.

1) **Zomato Bangalore Restaurant Dataset:** Zomato Bangalore restaurant dataset uploaded in 2019. The data was scraped from Zomato for educational purposes only. Dataset consists of 17 columns and 8792 Unique restaurants. From which we used 61 restaurants information and 5 columns labelled as name, locality, address, phone number and reviews. This dataset is available at Kaggle website.

2) **Google Reviews Ulhasnagar Dataset:** This dataset was made by us using Google reviews of Ulhasnagar restaurants in different localities. This dataset contains 48 unique restaurants in Ulhasnagar and 3 columns named as name, locality and reviews.

3) **Kaggle Labelled Dataset:** The dataset contains 1000 rows and 2 columns. The first column represents the reviews for the restaurant and second column represents the sentiment of the review. Further, we have expanded this dataset till 1500 labelled reviews. This dataset is used for training our model.

C. Methodology

Restaurant Reviews Classification and Recommender is basically a system in which we classify the reviews of the restaurant in to positive or negative review using the ML algorithm and then calculate the positive percentage of each restaurant and display the list of the restaurants in descending positive percentage.

Steps involved in our project are explained below:

- 1) **Data collection:** This phase is concerned with gathering the appropriate dataset for classification and sentiment analysis. The datasets mentioned above were collected and used for this project.
- 2) **Data Pre-processing:** This step is responsible for pre-processing for sentiment analysis on the reviews. The pre-processing part includes dropping special characters, converting all words in the reviews in to small case, dropping stop-words and stemming from reviews and then finally convert the clean dataset in to Bag of Words (BOW) representation.
- 3) **Data Training:** After pre-processing of the data, the data is trained using the machine learning algorithm, from the above mention algorithm table, we have selected logistic regression algorithm because of accuracy and performance results.
- 4) **Data Classification:** After training the model, we have passed unlabelled dataset to our model, which gives the reviews with the predicted label i.e., 0 or 1. The label 1 represents the positive review and the label 0 represents the negative reviews.
- 5) **Recommendation:** After classification of the reviews, the percentage for each restaurant is calculated based on the number of the positive reviews and the list of the restaurants based on the percentage is shown on our front-end technology.

IV. IMPLEMENTATION AND RESULTS

A. Accuracy Calculation

As, we have used logistic regression for our project, the following formula is used for calculating the accuracy using the confusion matrix for our project.

Confusion matrix:

True Positive (TP)	False Positive (FP)
False Negative (FN)	True Negative (TN)

Confusion matrix values of our model:

[[110 28]
[20 142]]

Formula for Accuracy:

$(TP+TN)/(TP+TN+FP+FN)$.

Accuracy for our project:

$(110+142)/(110+142+28+20) = 0.84 = 84\%$

We have trained our model using Logistic Regression and after training, we have predicted the sentiment for 4000 reviews. We have observed accuracy of 84% for this model.

	A	B	C	D	E
1	" Locality"	Name	Reviews	predicted_label	
2	banashankari	Jalsa	[('Rated 4.0', 'RATED	1	
3	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
4	banashankari	Jalsa	" ('Rated 2.0', 'RATED	0	
5	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
6	banashankari	Jalsa	" ('Rated 5.0', 'RATED	1	
7	banashankari	Jalsa	" ('Rated 5.0', 'RATED	1	
8	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
9	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
10	banashankari	Jalsa	" ('Rated 5.0', 'RATED	1	
11	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
12	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
13	banashankari	Jalsa	" ('Rated 4.0', 'RATED	1	
14	banashankari	Spice Elephant	[('Rated 4.0', 'RATED	1	

Fig. 2 Bangalore Restaurant reviews and their predicted label

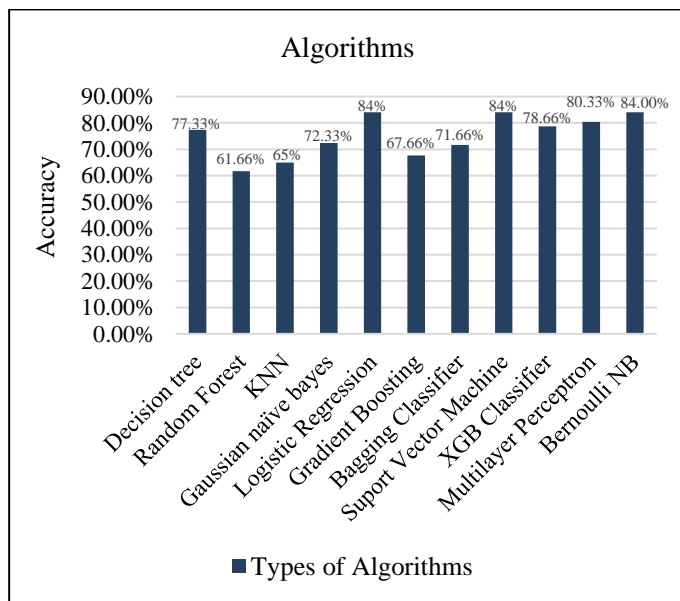


Fig. 1 A bar Graph showing types of algorithms we used and their accuracy

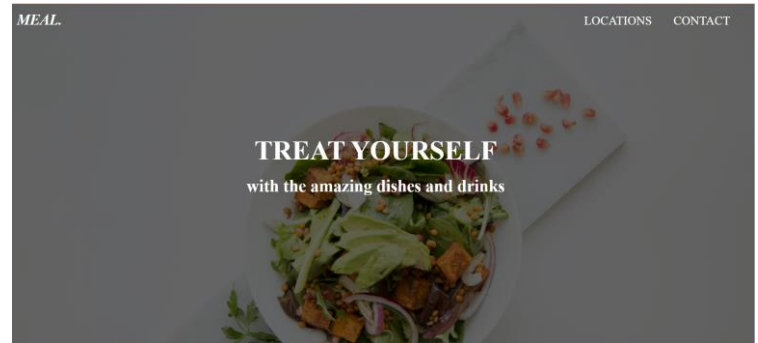


Fig. 3 Homepage of our website



Fig. 4 Localities section



Fig. 5 Categories section



Fig. 6 restaurants in particular locality

V. CONCLUSIONS

In this paper, we have proposed the restaurant reviews classification and recommender which is successfully predicting and recommending the restaurants based on the percentage calculated based on the positive reviews with the accuracy of 84%. Whereas the website is showing only static data is calculated in the backend. With more work, we can add post review section in the restaurant page and can convert this website to show results dynamically.

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 Hamitovali "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.
- [16] K.Soni 1, M.Viren 2, R.Jaikumar 3, V.Naveen 4 "Survey on Restaurant Review Classification and Recommender" International Journal of Scientific Research and Engineering Development— Volume 4 Issue 5, Year 2021, ISSN : 2581-7175.