PREDICTING THE REVIEWS OF THE RESTAURANT USING NLP TECHNOLOGIES

SRI MV SUBBARAO
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

K. SRI SAI GAYATRI(Lead)
Department of Information Technology
S.R.K.R. Engineering
College(A) SRKR Marg,
Bhimavaram.

K.SANJANA
Department of Information Technology
S.R.K.R. Engineering
College(A) SRKR Marg,
Bhimavaram.

K. LAKSHMI SITA SRAVYA
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

K.S.S.S HARSHITHA
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Abstract—Everyone wants the best from the restaurants like food, ambiance, service, food price, etc. To know about the best restaurant, people have to spend time on it, and they have to read all the data on the internet to avoid this problem our system helps to know which restaurant is best by predicting all the positive and negative reviews from the restaurant customers. The existing solution was that for restaurant reviews most of the machine learning algorithms will take only numerical data that was for example of rating 1-5 where the people should understand one as poor two rating as worst 3 rating as average and four rating as good 5 rating as best but with this customer who are viewing the reviews have to analyse the things around that. In the proposed solution it will calculate an overall rating for the collected text reviews, by using optimal algorithms, so that the people will get an idea of which restaurants suit them. Based on their satisfaction levels the customers will give ratings for the restaurants, and this can be done by Natural Language Processing Techniques.

Keywords—Text reviews, formatting, style, stop words, text words recognization , insert (key words**)**

I. INTRODUCTION

Restaurant reviews are an important source of information for individuals who are seeking to make informed decisions about dining experiences. Ratings and reviews are given by the restaurant customers based on their comfort levels. These ratings and reviews help the other customers to decide whether going to that restaurant or not and also the reviews are very helpful to the restaurant owners to develop their business. Here the restaurant reviews are available on online review platforms, such as Yelp and Google Reviews, there is an overwhelming amount of textual data available that can be difficult to navigate and make sense of. So, data is in form of textual manner but most of the machine learning algorithms will take only numerical data. We can analyse our textual datasets through NLP methodologies. NLP provides an opportunity for data analysts to apply machine learning and deep learning algorithms to our textual datasets. Classifying reviews and recommending the best restaurant is done by using machine learning algorithms. In general, the methods are implemented in a recommended system are 3 types they are content-based, collaborative-based, hybrid-based.

II. LITERATURE SURVEY

P.SasiKala,L.M.I.Sheela[1] discussed the increase of net contributes a large amount of usercreated data such as consumer feedback, opinions, and reviews. The problem of collecting data on the net and extraction about opinions is holding by sentiment analysis on the net. It helps to determine the opinion of people about a product and how it is received in the market is done by analysing their opinions. For sentiment analysis several trading tools are available. In this paper, we propose a system based on the sentiments on the words the proposed system will classify the reviews on a scale of 1 to 5.

Tri Doan and Jugal Kalita[2] discussed the Sentiment analysis of customer reviews has a critical impact on a business's development plan. The storehouse reviews develop over time, sentiment analysis frequently depends on offline solutions where training data is collected before the model is designed despite the fact. Incremental learning becomes the best alternative solution for tasks to avoid retraining the complete model from time to time. To conduct sentiment analysis on customer reviews we present a variant of online random forests in this work. Similar to offline methods and comparable to other online models our model can attain accuracy.

Tri Doan and Jugal Kalita[3] discussed the Sentiment analysis of customer reviews has a critical impact on a business's development plan. The storehouse reviews develop over time, sentiment analysis frequently depends on offline solutions where training data is collected before the model is designed despite the fact. Incremental learning becomes the best alternative solution for tasks to avoid retraining the complete model from time to time. To conduct sentiment analysis on customer reviews we present a variant of online random forests in this work. Similar to offline methods and comparable to other online models our model can attain accuracy.

Boya Yu, Zhou, Yi Zhang, Yunlong Cao[4] proposed a machine learning-based method for various kinds of restaurants to know the features. In this paper by using a support vector machine (SVM) model to solve the sentiment tendency of each review from word frequency is the main approach. Word scores generated from the Support Vector Machine models are using for indicating the significance of each word for special types of restaurants are further processed into a polarity index. Overall tend to express more sentiment regarding service by the Customers. Results that match the common sense are obtained as for the difference between different dishes.

Kirange D and Dr. Ratnadeep R. Deshmukh [5] researched in the "Sentiment Analysis" which focused on the recognition and classification of emotions (positive, negative, 7 conflict, neutral) in the aspect of reviews. The system for recognizing and studying the sentiments using Support Vector Machines for the restaurant and laptop feedback dataset is proposed in this paper. The 4 text analysis for recognizing the aspect in the review sentences can be applied in the proposed system. We compare the performance of the system With KNN, Naïve Bayes, and Neural Network classifiers and, we can compare the performance of the system.

Ekaterina Pronoza, Y.Elena, Svetlana Volskaya, Andrey Lyashin[6] Suggested that the information removal method for the restaurant recommendation system, which is intended to be a module of the recommendation system and aims for the development of an information extraction system? Collecting the information about various features of restaurants from online reviews, structure it, and feed the recommendation module with the collected data is done by the information extraction system. The Service, food quality, and cuisine, price level, noise level are including in the studied frames. Service quality, cuisine type, and food quality are regarded in this paper. A method for Russian reviews corpus analysis (as part of information extraction) is proposed apart from the corpus pre-processing phase. It is shown that for improving machine learning models' performance the ideas obtained at the corpus preprocessing stage can help.

III. SYSTEM IMPLEMENTATION (Methodology)

There are various methodologies for implementing sentiment analysis such as using Machine learning algorithms like Natural Language Processing or standard libraries, Naive Bayes, that are already available. But we are implementing a Machine learning algorithm for sentiment analysis which is Natural Language Processing (NLP).

Work flow:

Phase 1: Data Collection

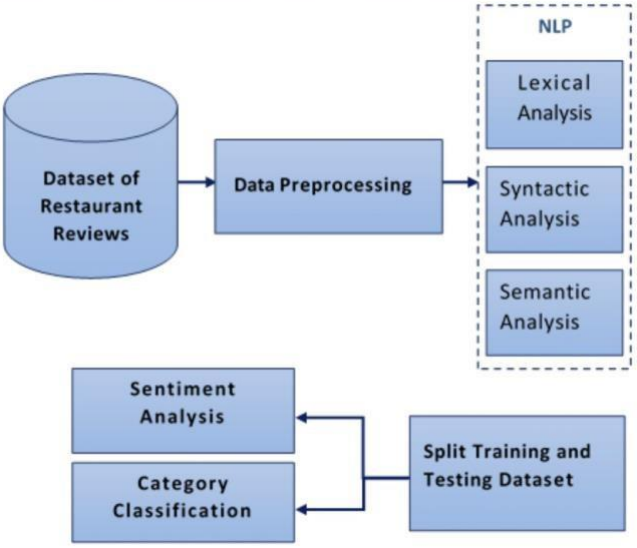
Data collection deals with the collection of the right dataset for classification and sentiment analysis. The data is collected in form of textual reviews from the reviews given by customers. The labeled dataset for sentiment analysis is taken from Kaggle.

Phase 2:

This phase deals with the aspect classification and preprocessing part for sentiment analysis. Initially, we pass the data and data is tokenized based on white spaces between and treat each token as one unit. Data will be pre-processed and also includes removal of stop words, POS tagging, lemmatization, stemming, etc. Feature generation is performed using TF-IDF (Term Frequency) it normalizes the document term matrix i.e., Confusion matrix. It is the product of TF and IDF. Finally, the dataset is prepared into vectors using the count vectorization for sentiment prediction.

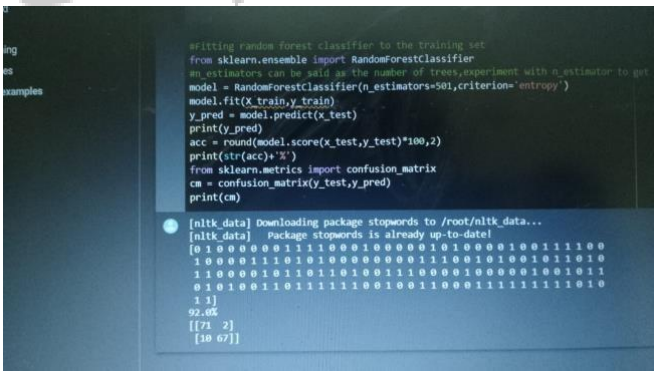
Phase 3:

This is the final phase where we are going to do the sentiment analysis prediction on the preprocessed data which is obtained from phase 2. The sentiments are predicted for each aspect and analyzed. The analyzed results are then displayed.



IV. EXPERIMENTS & RESULTS

For classifying restaurant reviews using NLP techniques, we have taken a dataset from kaggle and trained the model. After going through various experiments with different models, the random forest algorithm has taken into consideration as it gives the best accuracy among models.



Random Forest classifier has given the best accuracy with 92% for classifying restaurant reviews given by customers in form of textual reviews to positive or negative.

V. EVALUATION METRICS

Evaluation of the proposed model can be calculated using various metrics such as, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). And these can be calculated using the following formulae:

Precision=TP/(TP+F

P) Recall=TP/

(TP+FN)

Accuracy = (TP+TN)/ (TP+TN+FP+ FN)

Classifier Model	Validation Accuracy (%)	Recall (%)	Precision (%)
Model 1(Support vector Machine)	79	81	76
Model 2(Random Forest)	92	97	87
Model 3(Naïve Bayes)	73	56	82

VI. CONCLUSION

Normally, a lot of businesses are remained as failures due to lack of profit, lack of proper improvement measures. Mostly, restaurant owners face a lot of difficulties to improve their productivity. This project really helps to those who want to increase their productivity, which in turn increases their business profits.Classifying restaurant reviews using NLP techniques can provide valuable insights for businesses to improve their products and services, as well as for consumers to make informed decisions about where to eat. This is the main objective of the project. What we are doing is that the restaurant owner gets to know about drawbacks of his restaurant such as most disliked food items of his restaurant by customer's text review which is processed with ML classification algorithms. Predicting the reviews of the restaurant is done by the customer reviews. In the project for efficiency, we applied the word stemming by removing the stop words in the dataset. For converting the text data to numeric data we used the vectorization techniques (count vectorizer).

VII. FUTURE WORK

The work includes several potential areas for future work in classifying restaurant reviews using NLP techniques. In the future, a real-time application can be made so that people can use it to analyze the reviews more effectively and grow their businesses. Now all the works are related to binary sentiment classification i.e., positive or negative, in future we can extend it to degree of positivity and negativity with reviews. Restaurant reviews often contain not only text but also visual elements, such as photos and videos. Future work could extend on how to integrate these multimodal features into sentiment analysis to improve the accuracy and depth of analysis. Neural network architectures can be used for summarizing of texts more accurate and readable and to the point. Summarization of text reviews and the power to get insights from them can open an all-new world into the field of analytics and how we use data in businesses. There is a vast array of techniques that are left to explore which can make this an even more exciting area to study and improve. This study will have a vast room for future research and this can also be performed for various other industries like transportation, hospitality, healthcare as well as education.

ACKNOWLEDG MENT

We are very thankful for Govindarajan sir, Sasikala, Tri Doan, Waikul, Veda and many more for their generous works towards classifying restaurant reviews, as they helped us as a reference for research on this project. Also, grateful for everyone who has shown their immense support in completion of the project.

REFERENCES

[1] Govindarajan, M. "Sentiment analysis of restaurant reviews using hybrid classification method." International Journal of Soft Computing and Artificial Intelligence 2.1 (2014): 17-23.

[2] Sasikala, P., and L. Mary Immaculate Sheela. "Sentiment analysis of online food reviews using customer ratings." International Journal of Pure and Applied Mathematics 119.15 (2018): 3509-3514.

[3] Yu, Boya, et al. "Identifying restaurant features via sentiment analysis on yelp reviews." arXiv preprint arXiv:1709.08698 (2017).

[4] Kirange, D. K., and Ratnadeep R. Deshmukh. "Emotion classification of restaurant and laptop review dataset: Semeval 2014 task 4." International Journal of Computer Applications 113.6 (2015).

[5] Pronoza, Ekaterina, et al. ``Restaurant Information Extraction (Including Opinion Mining Elements) for the Recommendation System." Mexican International Conference on Artificial Intelligence. Springer, Cham, 2014.

[6] Tri Doan and Jugal Kalita discussed the Sentiment analysis of customer reviews has a critical impact on a business's development plan.

[7] Waikul, Veda, Onkar Ravgan, and Aruna Pavate. "Restaurant Review Analysis and Classification Using SVM." optimization

[8] Lee, Pei-Ju, Ya-Han Hu, and KuanTing Lu. "Assessing the helpfulness of online hotel reviews: A classification-basedapproach." Telematics and Informatics 35, no. 2 (2018): 436-445.

[9] Barbado, Rodrigo, Oscar Araque, and Carlos A. Iglesias. "A framework for fake review detection in online consumer electronics retailers." Information Processing & Management 56, no. 4 (2019): 1234- 1244.

[10] Thumbs Up Thumbs Down by D.Tumey, Turney, (2020).

21. Priyantina, Reza Amalia, and Riyanarto Sarno. "Sentiment Analysis of Hotel Reviews " International Journal of Intelligent Engineering and Systems 12, no. 4 (2019): 142- 155.

[11]M. Al-Smadi, O. Qawasmeh, B. Talafha, M. Quwaider, Human annotated Arabic dataset of book reviews for aspect based sentiment analysis, in: 3rd International Conference on Future Internet of Things and Cloud (FiCloud), IEEE, pp. 726–730 (2015).

[12] Mohammad Al-Smadi, Omar Qawasmeh, Mahmoud Al- Ayyoub, Yaser Jararweh, Brij Gupta "Deep Recurrent neural network vs. support vector machine for aspectbased sentiment analysis of Arabic hotels' reviews Journal of Computational Science 27 (2018) 386–393".

[13] SemEval-2014 Task-4 dataset for aspect classification. Available at: http://alt.qcri.org/semeval2014/task 4/ (2014).

25. Zomato Bangalore Restaurant Dataset Available at: https://www.kaggle.com/himanshup oddar/zo mato- bangalore-restaurants (2019).