Machine Learning Based Class Level Prediction of Restaurant Reviews

F.M Takbir Hossain
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
takbir3038@diu.edu.bd

Md. Ismail Hossain
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
ismailhossainroyal@diu.edu.bd

Ms. Samia Nawshin
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
samia@daffodilvarsity.edu.bd

***Abstract*- Nowadays with the proliferation of location aware technologies and smart phones people tend to give reviews for all types of products services and place them online. It is very important to extract knowledge or information occupies in the vast amount of available text reviews. For these, user's sentiment is also monumental. If any business owners want to take decision on future planning, they must consider their clients sentiment. In this research, we proposed a noble strategy to predict user sentiment from their online reviews given for a particular business by using supervised machine learning techniques. Our proposed machine learning model will give a hand to restaurant owners to identify their customer's feedback and market positions.**

Keywords— Website, Social Media, Sentiment, Prediction, Machine Learning

I. INTRODUCTION

In recent years it is noticeable that sharing text reviews on various businesses specially restaurants through website and social media is a very common phenomenon. Online reviews reflect customer's opinion. By expressing own sentiment, users actually rate the restaurants and their services. That's why these reviews can be the source for the sentiment analysis of a user about a restaurant. In recent years, the number of internet and social media users are growing dramatically in Bangladesh. The restaurant owners offer the users to share their valuable feedback on social media or website and start to care about their customer's key point of interest of their services. This huge collection of user data in terms of text reviews can be analyzed to identify user's sentiment and their demand also. Here users are the primary sources. Text reviews are the complete reflection of user's sentiment and also owned by them. Restaurant owners can get very useful information from the user's sentiment analysis. Measuring user's sentiment will also be able to find out the market position of a restaurant. By making the machine learned about the total reviews and their class levels as positive or negative, it will be able to categorize new user reviews [2]. To identify the class of a review text we collected reviews from a Bangladeshi website named Priyo review (Beta).

In this paper, we took a step forward by combining user review texts which were collected from that website to build a model that can predict a review asserting positive or negative[1, 2, 7]. Key benefit of our approach is that, by using our proposed model restaurant owners can identify the main focused term from the review of customers and also can take future step to work on that. We are also able to publish the position of a restaurant by counting that how many reviews are positive comparative to negative. If number of positive reviews is greater than negative one then it can be said that status of the restaurant is in a good position otherwise just opposite to it. As this model is based on text document, it will be very perfect work in all terms and condition because text document shows almost the best predicting result of user's sentiment than that of star rating does [4].

In our study we have collected review data from local businesses. Then we selected valid review texts which are given by the users in English. Then we employed four machine learning based classification algorithms for finding the best classification model for predicting the class level of new review texts.

In summary, we have the following contributions.

- At first we used data from a website that is owned by Bangladesh.
- In our experimental study, we collected 827 reviews of different restaurants and measure the efficiency of our approach.
- We have described the techniques of finding positive negative keywords as identifier from the text documents.
- By integrating all the reviews we can ensure the market position of a restaurant.

II. RELATED WORK

Online reviews are often accessed by the users to buy products, see a movie or going to a restaurant. However online reviews are a useful resource for sentiment analysis therefore it is easier to work with it for classification problem. We can

978-1-5386-2175-2/17/$31.00 ©2017 IEEE

identify the work in two ways. Some works have been done with just for answering the positive and negative class level, some work with the stars. Many of the work have gone through both.

Text reviews in social media and website are progressive. On the basis of text documents lots of work has already been done for sentiment analysis [1, 6, 11]. It is always important to identify user's sentiment for predicting any review finding from online. Collecting human sentiment is also crucial because it gives us more insight into how the sentiment of the entire excerpt is formed from its constituents [8]. One way described in different papers that working with textual document results better than the star rating which is in numeric form [4]. Score rating from text document analysis has also mentioned for improving the quality of a classifier [5]. Using the user's opinion is always appreciating to all because users are the most serious person who actually deals with review text though online. By using those review texts giving only the answer of positive and negative is also a text mining for understanding the user's sentiment [1, 2, 7]. Some others say that classifying a review document as positive sentiment or negative sentiment using supervised learning algorithm there may be tendency of positive classification accuracy is approximately 10% higher than the negative one. In order to mitigate they proposed an improved Naïve Bayes algorithm which can express the average value of the two accuracies [9]. Again another option came with neutral value which was also a good one because there were three class levels at the end [10]

In recent time Proposing two ranking mechanisms for ranking product reviews: one is consumer-oriented ranking mechanism that ranks the reviews according to their expected helpfulness and another is manufacturer oriented ranking mechanism which ranks the reviews according to their expected effect on sales is also a newer thing [3]. A paper published in 2014, they attempt to bridge the gap between phrase-level and review/document-level sentiment analysis by leveraging the results given by review [11].

Our paper presents the first approach that asserting or predicting the class level of a review is not just enough and so we elaborate this by giving the answer of total ratio of positive and negative reviews of a specific restaurant. By these a restaurant owners will be able to know that what the position of their restaurant due to the class level prediction of total reviews of a restaurant. Positioning of a restaurant is totally new thinking and implementation apart from all those had before. That is the main identical part of our research.

III. Implementation

The process of text mining is different from mining numeric data because in algorithm we cannot use the text document directly. Unlike from numeric data finding features from text data is not a straightforward process. As our goal is to classify text reviews either positive or negative, so the task of selecting features from the reviews for classification is our main challenge for completing this work.

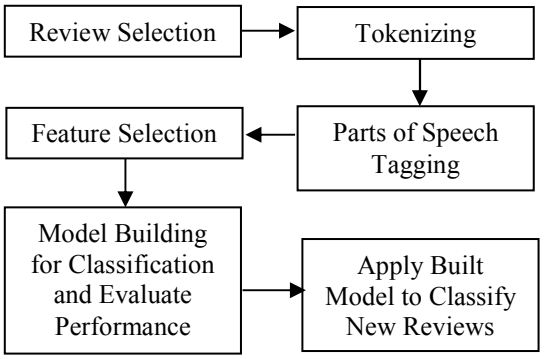

Figure 1: Stepwise working flow of the whole process.

At first we have pre-processed raw review texts. Figure 1 shows our whole process stepwise. For preprocessing we tokenize each review and tag the parts of speech and select features. Then we applied different classification algorithm on the processed data. We have used classification algorithm to assert a review text whether it is positive or negative. Here we have made the machine learned about a specific class or target. After the machine learned by the training data, it will categorize the new user reviews. Finally from the ability of categorizing a review, our model will find the ratio of positive and negative reviews of a restaurant which is the reflection of a restaurants market position.

A. Data set

We collected our dataset from one Bangladeshi review website popularly known as Priyo Review (Beta). This dataset includes user review ratings on business, business id, unique user id, location, image etc. We are mainly dealing with restaurant's review text. From that total dataset, we only examined restaurant review text and then we studied about those data to extract usable feature for building classification model. Here we found almost 50 restaurants with more than 1000 reviews. As the mentioned website is new so currently the number of reviews is small but this number is getting bigger every day.

B. Tools

We used Natural language Toolkit (NLTK) for data pre-processing which is widely used for processing text data. We preeminently used Sci-kit learn library for building classification model. Both tools are developed by Python programming language.

C. Creating feature

a) Review Selection: Users often express their positive or negative opinion by using some common sentence like "The food was good and tasty", "The food quality was bad". For those reviews the feature can be selected easily but most of the time the users use very complicated words, critical for identifying positive or negative sentiment. Using more unreliable words as feature can generate low performance. On

the other hand choosing some specific words will cause the loss of information. So for creating features we selected the maximum possible words containing most information.

Users often share their opinion on various websites which mostly indicate the quality of food or services on a specific restaurant. Sometimes they did not say anything about the quality of food or service in their review and say about only the occasion of going the restaurants but those are not indicating any sentiment of the users about the restaurants. So using those reviews in training set are not necessary. Selecting appropriate reviews for experiment is important for building a good model. As we are working with Priyo review (Beta) data set and it collects reviews for local business of Bangladesh we also found some reviews which was written in Bengali, for example " *Pasta khete valoi lagse, khub e valo poribesh*" or "*Food quality khub e kharap, onek deri kore serve korse, faltu .... dhurrrr"* . We discarded those reviews from our training dataset. Finally we selected 503 review texts for our training dataset.

b) Tokenizing and parts of speech tagging: Tokenizer breaks the full documents by identifying insight of a document, that's why it makes easy to learn about the sentiment of a user review. For separating each and every sentences of a review text in our data set, we have used sentence tokenizer. To identify the word quantity we have used word Tokenizer of NLTK. Then we have used parts of speech tagging. Here we are showing an example of review "*The place is really very large. The foods are delicious all the time. I always enjoy that .The tea and leg roast are the best but I can never find chicken jhal fry whenever go there after 7:30.but all foods are delicious*"

[('the', 'NNP'), ('foods', 'NNS'), ('are', 'VBP'), ('delicious', 'JJ'), ('all', 'PDT'), ('the', 'DT'), ('time', 'NN'), ('.', '.')]
[('I', 'NN'), ('always', 'RB'), ('enjoy', 'VBP'), ('that', 'IN'), ('.the', 'NNP'), ('tea', 'NN'), ('and', 'CC'), ('leg roast', 'NN'), ('are', 'VBP'), ('the', 'DT'), ('best', 'JJS'), ('but', 'CC'), ('i', 'NN'), ('can', 'MD'), ('never', 'RB'), ('find', 'VB'), ('chicken', 'JJ'), ('jhal', 'NN'), ('fry', 'NN'), ('whenever', 'WRB'), ('go', 'VBP'), ('there', 'RB'), ('after', 'IN'), ('7:30.but', 'CD'), ('all', 'DT'), ('foods', 'NNS'), ('are', 'VBP'), ('delicious', 'JJ')]

We focused mostly on adjective (tagged as 'JJ') and noun (tagged as 'NN') for making feature because other words generally not contain any logical information or mean any sentiment of a user.

c) Selecting keywords: After identifying and removing the unnecessary words, we have found very clean and specific words which can identify the positive or negative sentiment of a review. In the training data set, we have found 2018 words those are only adjective and noun. The adjectives are mostly used to identify whether an item is positive or negative. For example, if we find five words which are 'good' , 'bad' , 'nice' , 'chicken' , 'rice' . We consider 'good' and 'nice' as positive word and 'bad' as negative word. After detecting the total positive and negative keywords of a review, we use this knowledge to decide whether a review is positive or negative. As we have used sci-kit learn library for building model, it considers each and every word as feature.

D. Building Classification Model

We have run our pre-processed dataset into Sci-kit learn for building the classification model. Here every review is separate from each other and the response vector of these reviews is represented as 1 for positive review and 0 for negative review. We have used countvectorizer in Sci-kit learn [12] by four steps model building pattern: *import*, *instantiate, fit and predict*. As we have preprocessed and cleaned the training set before, our built model responds very well for new review's class level (positive or negative) prediction.

For this experiment we have used four proper algorithms for building proper model by using training dataset. Those are Multinomial Naïve Bayes (MNB), Support Vector Machine (SVM) K nearest neighbor (KNN) and Linguistic Regression (LR). For measuring the performance of our built model we have used k-fold cross validation which is very effective way for finding the accuracy of a predictive model. It divides the data set into training and testing set and computes the performance. We have used 10 fold in our experiment that means k=10. Linguistic Regression (LR) has given us the best result with more than 77% accuracy. Table 1 presents the prediction accuracy, true positive rate (TPR), true negative rate (TNR) and area under the ROC curve (AUC) of different classification algorithms for our built model.

Table 1: Performance of different classifiers.

Algorithm	Accuracy	TPR	TNR	AUC
MNB	0.7619	0.87	0.57	0.77
SVM	0.7142	0.74	0.65	0.73
KNN	0.7460	0.89	0.48	0.75
LR	0.7779	0.84	0.62	0.79

Finally, we calculated the number of positive and negative reviews predicted by our model for all the reviews of a given restaurant and measure the ratio between positive and negative reviews.

IV. CONCLUSION

In this paper, we represented the experiments we have done on sentiment classification of user review texts using a large dataset by Sci-kit learn library. Our built model can predict the class level of user review as positive or negative. More importantly, this paper shows that sentiment classification is possible to be learned from online review text given by the users. In addition, we have shown performances of our built model using different classification algorithms. We are the first to work with a Bangladeshi website containing

reviews of local businesses. To the best of our knowledge, this is the first work of counting the ratio of positive and negative reviews of a restaurant. This ratio is an indicator of a restaurant's popularity or market position. Determining the position of a restaurant by using users' review is totally a matchless work relative to others.

We used supervised machine learning for predicting class levels. In future, it will be absolutely very exciting to find the hidden points from review text with unsupervised learning. Now our model can predict between two class levels positive and negative, but there are reviews with mixed feelings. In future we plan to build multi-level classifier model which will pursuit to predict the best class label from possible options. Multi-level classifier will reflect market position of a business more accurately.

REFERENCES

[1] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. "Thumbs up?: sentiment classification using machine learning techniques." *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*. Association for Computational Linguistics, 2002.

[2] Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004.

[3] Ghose, Anindya, and Panagiotis G. Ipeirotis. "Designing novel review ranking systems: predicting the usefulness and impact of reviews." *Proceedings of the ninth international conference on Electronic commerce*. ACM, 2007.

[4] Ganu, Gayatree, Noemie Elhadad, and Amélie Marian. "Beyond the Stars: Improving Rating Predictions using Review Text Content." *WebDB*. Vol. 9. 2009.

[5] Lee, Moontae, and R. Grafe. "Multiclass sentiment analysis with restaurant reviews." *Final Projects from CS N* 224 (2010).

[6] Socher, Richard, et al. "Semi-supervised recursive autoencoders for predicting sentiment distributions." *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, 2011.

[7] Zhang, Ziqiong, et al. "Sentiment classification of Internet restaurant reviews written in Cantonese." *Expert Systems with Applications* 38.6 (2011): 7674-7682.

[8] Wu, Jean Y., and Yuanyuan Pao. "Predicting Sentiment from Rotten Tomatoes Movie Reviews."

[9] Kang, Hanhoon, Seong Joon Yoo, and Dongil Han. "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews." *Expert Systems with Applications* 39.5 (2012): 6000-6010.

[10] Khairnar, Jayashri, and Mayura Kinikar. "Machine learning algorithms for opinion mining and sentiment classification." *International Journal of Scientific and Research Publications* 3.6 (2013): 1-6.

[11] Zhang, Yongfeng, et al. "Do users rate or review?: Boost phrase-level sentiment labeling with review-level sentiment classification." *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. ACM, 2014.

[12] Learn about scikit-learn, available at << http://scikit-learn.org/stable/tutorial/statistical_inference/supervised_learning.html>>, last accessed on 30-08-2017 at 10:00pm.