TTIC, 2019, Vol.3, 55-58

# To dine or not to dine: Can machine learning help?

Umme Habiba Maliha, Syeda Benazir Hossain, Samsul Arefin Riffat, Sifat Momen & Shakhawat Hossain Mahi*
Department of Electrical and Computer Engineering, North South University, Bangladesh
shakhawath.mahi@northsouth.edu

***Abstract*—Finding a suitable restaurant to dine in can be considered as a complex decision problem for the residents of Dhaka, the capital of Bangladesh. This is due to the fact that people do not select a restaurant based on the food taste only. Rather, a range of parameters affect the decision making process. This has resulted diners to rely more on online restaurant reviews to decide on their choice. Restaurant ratings affect the customers' decision and consequently also affect the restaurant business. As a consequence, restaurant owners are also extremely careful in maintaining the quality of service and the feedback from the customer. In this paper, we apply machine learning approach to predict the restaurant rating. Our approach finds that we can predict the rating that a customer provide with an accuracy of 92%.**

***Index Terms*—restaurant rating, classification, machine learning, prediction**

## I. Introduction

In today's time, online restaurant reviews have become a popular culture that affects greatly to the success of a restaurant business. Many restaurant owners are now aware of the fact that positive reviews can bring profit and popularity while negative reviews can exacerbate the business loss. This makes it imperative for restaurant owners to understand how customers perceive the service and rate their restaurants.

On the flip hand, restaurant ratings are crucial to customers as it helps them to make an overall judgment on the ambience of the place they plan to visit. There are several platforms in order to reach the ratings, for instance, social media, websites, food blogging sites etc. In a recent research report, [1], approximately 61% of customers have been found to read online reviews about restaurants before visiting them. The research has further found that 34% of diners choose restaurants based only on the information that appear on these peer reviewed websites. This implies that most diners completely overlook the restaurant's website or social media pages and rely only on the data available on these sites. These statistics clearly depict the undeniable importance of restaurant ratings.

In order to predict restaurant ratings in a smarter way, machine learning techniques have been embraced in this paper. Models, in our paper, predict a rating (bad, average or good) from a number of parameters that a user feels about a restaurant. Different machine learning classifiers including decision tree, random forest, k-nearest neighbor and naïve bayes have been applied.

The rest of the paper is organized as follows: In section II, a discussion on the related work in this area has been made. In section III, we present the proposed machine learning framework. Empirical results are discussed in section IV and finally in section V we conclude the paper with remarks on our future work.

## II. Literature Review

Predicting ratings and reviews using machine learning has been an active area of research.

Wang and colleagues [2], for instance, combined online sequential extreme learning machines (OS-ELMs) and intuitionistic fuzzy sets to predict consumer sentiment and proposed a generalized ensemble learning scheme.

Zhang et al. [3] used classifiers such as naïve bayes and support vector machine (SVM) to automatically classify Cantonese-written restaurant reviews as either positive or negative. They found naïve bayes classifier to predict user reviews with better accuracy than that of SVM.

McAuley and Leskovec [4] combined latent rating dimensions with latent review topics in order to obtain highly interpretable textual labels to justify ratings with text. Their approach can predict product ratings by harnessing the information present in the review text.

Kumar and colleagues [5] proposed a novel hierarchical supervised learning method to detect online spammers who exploit consumer trust by posting fake reviews. They used several supervised learning techniques including logistic regression, support vector machine and k-nearest neighbor.

Internet review forums work on expert opinions and social network reviews to provide customers ambient rating experiences. Anderson et al. designed a regression discontinuity on the mined opinion that lead to a positive impact on restaurant ratings and increases the sale of restaurants largely [6].

Jacobsen et al. [7] found that expert opinions affects consumer ratings and this limits the possibility of actual reviews being produced.

## III. Proposed Framework

Reviews or ratings have gained trendiness amongst users in order to understand the quality of a product. The objective of this work has been three-folds: (1) to create a dataset containing opinions of users pertaining to restaurants they have visited before, (2) detect influential features that affect the restaurant ratings, and (3) predict the rating of a new restaurant with high accuracy.

Figure 1 shows the methodology that has been used in this work.

Supervised learning, a kind of induction learning, has been used to train the learner to predict restaurant ratings. Data is collected by asking individuals about their experiences in different restaurants. A total of 850 instances were collected which we assume to be independent and identically distributed

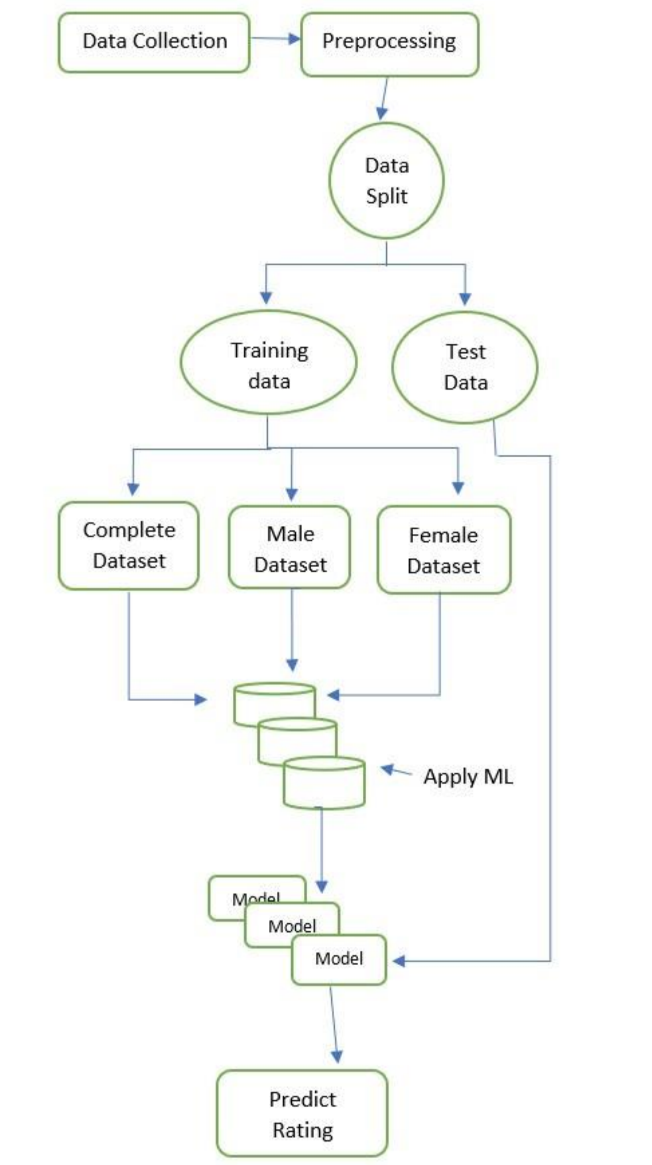


Fig. 1. Proposed framework for predicting a Restaurant's overall rating

(IID). Initially 18 attributes per instance was recorded. During the pre-processing phase, the attributes *restaurant name* and *location* were removed - thus finally 16 attributes per instance were taken into consideration. These attributes measure range of factors pertaining to the perceived quality of service including that of food quality, service, view, hygiene, facilities etc.. Details of the attributes recorded are summarised in table I.

After the collection of raw data, the data is pre-processed. Pre-processing techniques include discretizing continuous values to discrete values such as converting the numeric value of the overall rating to categories (overall rating of 1- 2 are categorized as bad, 3 as average and 4-5 as good) and also filling out obvious missing values. For instance, in multiple cases, users did not fill up information such as the location of a particular restaurant, availability of parking facility, if there

TABLE I
DESCRIPTION OF SELECTED ATTRIBUTES

| Variable Name | Description |
|---|---|
| Gender | Responder's gender (Male or Female) |
| Restaurant Name | Name of the Restaurant s/he visited |
| Location | Location of the restaurant |
| Restaurant Type | Type of the restaurant (Chinese, Thai, Bangla, Continental, All Cuisine, Cafe, Food Court) |
| View | View from the Restaurant (Lakeview, Roadside, Rooftop) |
| Air Conditioned | Air conditioned? (Yes or No) |
| Capacity | Number of seat capacity of the restaurant |
| Food Items | Number of Food items the restaurant provide |
| Food Quality | Quality of the Food served in the restaurant (Rate between 1 to 5) |
| Food Quantity | Quantity of the Food served in the restaurant (Rate between 1 to 5) |
| Taste | Taste of the Food Served (Rate between 1 to 5) |
| Price | Price of the Food (Low, reasonable, Expensive) |
| Service | Customer Service of the restaurant (Rate between 1 to 5) |
| Parking Space | Parking Space? (Yes or No) |
| Kids Play Zone | Have Kids Play Zone? (Yes or No) |
| Smoking Zone | Have Smoking Zone? (Yes or No) |
| Hygiene | Was the restaurant hygiene? (Rate between 1 to 5) |
| Overall Rating | Overall Rating of the restaurant (Rate between 1 to 5) In preprocessing (1,2 = bad — 3= average — 4,5 = Good) |

exists air condition in the restaurant. If these information are available to us, they have been completed.

After the pre-processing phase, the dataset is split such that 80% of the data used form the training set while the remaining 20% the test set. Four sets of data were produced from the training set - the first set comprises of the entire training set (A), the second set comprises of instances provided by the male participants (B) while the third set comprises of instances provided by the female participants (C). Following this, a Correlation Feature Set (CFS) subset evaluator is used to determine the influential parameters for predicting the rating. A fourth set of data (D) is produced that comprises of only six influential parameters identified by the CFS.

TABLE II
DESCRIPTION OF FEATURE SETS

| Dataset | Method | Selected Attributes |
|---|---|---|
| A | Complete Dataset | All Attributes |
| B | Filtered by only Male Instances | All Attributes |
| C | Filtered by only Female Instances | All Attributes |
| D | Dataset with attributes selected by CFS | Capacity, Food Quality, Food Quantity, Taste, Service, Hygiene |

## IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experiments were conducted using the WEKA 3.8.3 platform [8] and was run in a machine with 4GB RAM and 2.4 GHz processor. Models were developed using training data and validated with stratified ten-fold cross validation technique. After the model is created, it is tested using the test data.

The performance metrics used to evaluate the performance of the models include accuracy, precision and recall. These metrics can be measured using the confusion matrix [9].

Accuracy is the proportion of the total number of predictions that were correct [9].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Precision is the proportion of positive cases that were correctly identified [9].

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall is the proportion of actual positive cases which are correctly identified [9].

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

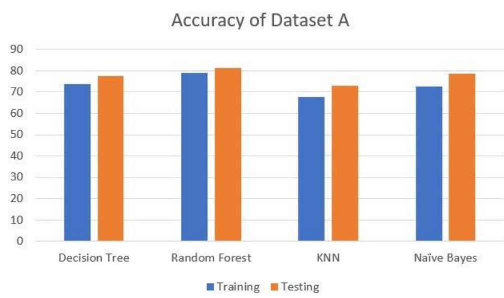


Fig. 2. Training and Testing Accuracy on different Algorithms on Dataset A

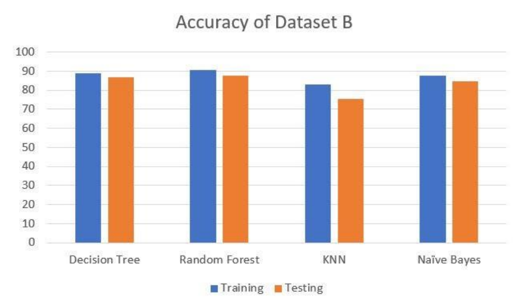


Fig. 3. Training and Testing Accuracy on different Algorithms on Dataset B

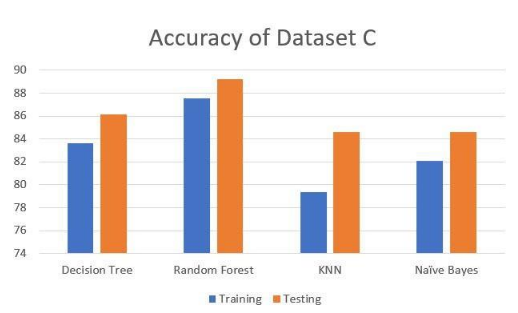


Fig. 4. Training and Testing Accuracy on different Algorithms on Dataset C

The results produced are summarised in table III. The results indicate that the highest test accuracy is achieved using the K-nearest neighbor classifier on Dataset D.

The precision-recall curve shows the tradeoff between precision and recall for different thresholds. It is a way to assess the effect of false positive rate on false negative rate. A good PR curve has the characteristic of having greater AUC (area under curve). A high area under the curve represents both high recall and high precision, where high precision relates to a low false positive rate, and high recall relates to a low false negative rate.

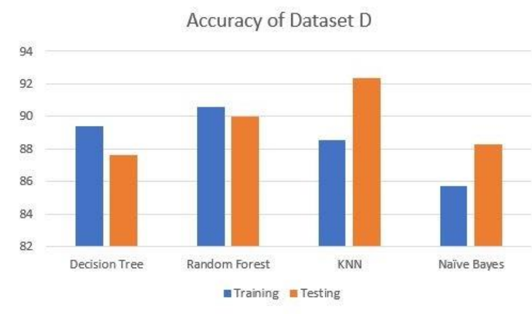


Fig. 5. Training and Testing Accuracy on different Algorithms on Dataset D

TABLE III
THE PREDICTION ACCURACY MATRIX

| Dataset | Algorithm | Training Accuracy | Test Accuracy | Precision | Recall |
|---|---|---|---|---|---|
| A | Decision Tree | 0.85 | 0.89 | 0.90 | 0.89 |
| | Random Forest | 0.89 | 0.91 | 0.91 | 0.91 |
| | K-NN | 0.83 | 0.80 | 0.79 | 0.80 |
| | Na¨ıve Bayes | 0.86 | 0.86 | 0.87 | 0.86 |
| B | Decision Tree | 0.89 | 0.86 | 0.87 | 0.86 |
| | Random Forest | 0.91 | 0.88 | 0.87 | 0.87 |
| | K-NN | 0.83 | 0.87 | 0.87 | 0.87 |
| | Na¨ıve Bayes | 0.87 | 0.85 | 0.85 | 0.85 |
| C | Decision Tree | 0.83 | 0.86 | 0.87 | 0.86 |
| | Random Forest | 0.87 | 0.89 | 0.89 | 0.89 |
| | K-NN | 0.79 | 0.84 | 0.84 | 0.84 |
| | Na¨ıve Bayes | 0.82 | 0.84 | 0.87 | 0.84 |
| D | Decision Tree | 0.89 | 0.87 | 0.88 | 0.87 |
| | Random Forest | 0.91 | 0.90 | 0.91 | 0.90 |
| | K-NN | 0.89 | 0.92 | 0.93 | 0.92 |
| | Na¨ıve Bayes | 0.86 | 0.88 | 0.90 | 0.88 |

We have found that the test accuracy of KNN on dataset D is the highest (92%). We also found that the variance between train and test accuracy is lowest with random forest algorithm.

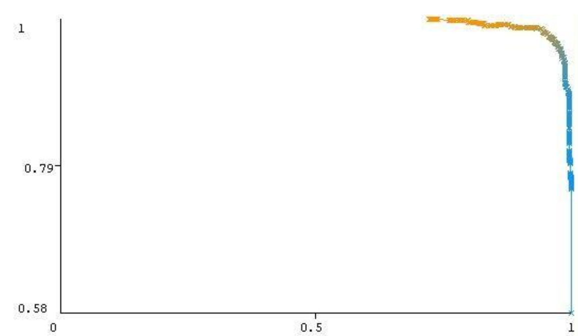


Fig. 6. PR curve for Random Forest Algorithm in Dataset D for class value ”Good”

The value of precision and recall is very high when random forest algorithm is applied on dataset D (figures 6,7,8). This results in a high area under the curve.

## V. CONCLUSION

There has always been search on how to attract customers coming to one's restaurant. In this paper, a machine learning framework has been developed to predict the rating of a restaurant. KNN classifier has been found to produce the best accuracy while random forest algorithms yielded lowest

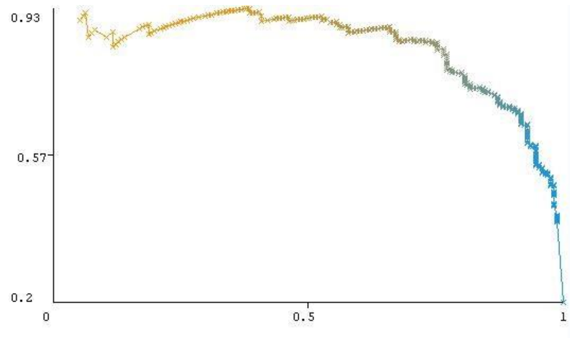


Fig. 7. PR curve for Random Forest Algorithm in Dataset D for class value "Average"

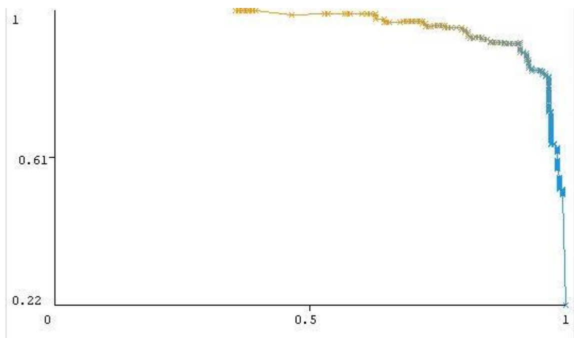


Fig. 8. PR curve for Random Forest Algorithm in Dataset D for class value "Bad"

variance in terms of train and test accuracy. In our future work, we plan to understand how males and females rate restaurants.

## REFERENCES


[1] M. R. M. . (https://www.modernrestaurantmanagement.com), "The impact of reviews on the restaurant market (infographic): Modern restaurant management: The business of eating restaurant management news," Apr 2017. [Online]. Available: https://www.modernrestaurantmanagement.com/the-impact-of-reviews-on-the-restaurant-market-infographic/

[2] H. Wang, G. Qian, and X.-Q. Feng, "Predicting consumer sentiments using online sequential extreme learning machine and intuitionistic fuzzy sets," *Neural Computing and Applications*, vol. 22, no. 3-4, pp. 479–489, 2013.

[3] Z. Zhang, Q. Ye, Z. Zhang, and Y. Li, "Sentiment classification of internet restaurant reviews written in cantonese," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7674–7682, 2011.

[4] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: understanding rating dimensions with review text," in *Proceedings of the 7th ACM conference on Recommender systems*. ACM, 2013, pp. 165–172.

[5] N. Kumar, D. Venugopal, L. Qiu, and S. Kumar, "Detecting review manipulation on online platforms with hierarchical supervised learning," *Journal of Management Information Systems*, vol. 35, no. 1, pp. 350–380, 2018.

[6] M. Anderson and J. Magruder, "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database," *The Economic Journal*, vol. 122, no. 563, pp. 957–989, 2012.

[7] G. D. Jacobsen, "Consumers, experts, and online product evaluations: Evidence from the brewing industry," *Journal of Public Economics*, vol. 126, pp. 114–123, 2015.

[8] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[9] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation," in *Australasian joint conference on artificial intelligence*. Springer, 2006, pp. 1015–1021.