Enhancing Sentiment Analysis using Machine Learning Predictive Models to Analyze Social Media Reviews on Junk Food

Conference Paper · October 2022

DOI: 10.1109/ICCCNT54827.2022.9984355

6 authors, including:

Moshfiqur Rahman Ajmain
Daffodil International University
8 PUBLICATIONS **9** CITATIONS
SEE PROFILE

Mst. Farhana Khatun
Daffodil International University
4 PUBLICATIONS **6** CITATIONS
SEE PROFILE

Sheikh Sadi Bandan
Daffodil International University
4 PUBLICATIONS **5** CITATIONS
SEE PROFILE

Md Arifur Rahman Rejuan
Daffodil International University
3 PUBLICATIONS **5** CITATIONS
SEE PROFILE

Enhancing Sentiment Analysis using Machine Learning Predictive Models to Analyze Social Media Reviews on Junk Food

Moshfiqur Rahman Ajmain
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
moshfiqur15-14090@diu.edu.bd

Mst. Farhana Khatun
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
farhana15-14304@diu.edu.bd

Sheikh Sadi Bandan
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
sheikh15-14089@diu.edu.bd

Arifur Rahman Rejuan
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
arifur15-13810@diu.edu.bd

Nushrat Jahan Ria
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
nushratria.cse@diu.edu.bd

Sheak Rashed Haider Noori
Dept. of CSE
Daffodil International University
Dhaka, Bangladesh
drnoori@daffodilvarsity.edu.bd

***Abstract*—In the last few years, the Use of social media has increased immensely. People share different types of opinions on social media like Facebook posts, comments, tweets etc. Sentiment analysis involves the process of categorizing these opinions. The aim of this study, find out the customer's attitudes toward the restaurant. Nowadays sentiment review is gaining grip. The benefits of this sentiment analysis for restaurants is how customers like their food and as a result, the business of Bangladeshi restaurants will be more developed. The study focuses primarily on customers' behavior, tastes, preferences, conversations, reviews, and objections. For this purpose 500 data are collected. There are six attributes in the dataset and based on customer reviews they are satisfied or unsatisfied. This exploration uses different classifiers of ML to develop review analysis like SVM, Random Forest, K-nearest neighbors, Decision Tree, Logistic Regression and XGBoost Classifier. And Comparing these algorithms' performances, XGBOOST gives the greatest accuracy which is 83%.**

***Index Terms*—Food Review, Sentiment Analysis, Bangladeshi Food Review, Machine Learning Algorithms, XGBoost**

I. Introduction

Food produces energy and heat in the body making the organs healthy, strong, functional and strong. To keep the body healthy, strong and functional, a balanced diet should be taken regularly. For example - rice, pulses, fish, meat, greens, vegetables, fruits, milk, sugar, water etc. Do we know how many restaurants there are in Bangladesh? There are more than 50,000 restaurants nationally. There are 10,000 restaurants in Dhaka city. The traditional cuisine of Bangladesh is one of the most underrated cuisines in South Asia. Some of the most popular food items in Bangladesh are Shoshe Hilsa, Kachchi Biryani, Beef Kala Buna, Roasted Khichuri with Egg Baji, Shik Kebab with Naan, Dal, Bharta, Fuchka, Kolizer Shingara, Haleem etc. A part from the traditional food of Bangladesh, people are also interested in seafood, Chinese, Korean and Japanese food. Some of the food outside Bangladesh like- Noodles, Chestnuts, Sticky Rice, Sauce, Burger, Pizza, Fried Rice, Chicken Fries, Grilled Rice, Thai Soup, Thai Fried Noodles, Prawns Sizzling, Beef Tarragona, Cheesy Chicken etc. In this research paper, we have gathered food review data from several Bangladeshi restaurants. We gather 500 data from various restaurants like Bengali, Chinese, Kurian etc. With opinions of both men and women. Based on this information we can know which restaurants are providing good service and what is the quality of their food. Food data types in our review include - Burger, Naga Chicken, Pizza, Potato Wedge, Fried Rice, Soup, Pasta, Roast Khichuri, Mushroom, BBQ, Seafood, Grilled Chicken, Mutton Leg Roast, Beef and Chicken Patty, Kebab etc. Machine learning algorithms are very popular for sentiment analysis. Day by day people are interested in completing sentiment analysis employing machine learning algorithms. In our observation, we found SVM (Support Vector Machine), Logistic Regression and Naive Bayes are the most popular and used algorithms. CNN (Convolutional Neural Network) and LSTM are used for sentiment analysis. The highest accuracy using SVM is 88.38% which is very good. In This research paper, we are doing an online-based sentiment analysis of restaurants.

Here we have done our research paper with four processes. These are Literature review, Methodology, Results and discussion, Conclusion and Reference. As the first move, we are studying many related papers. There they discuss sentiment analysis of food, the best machine learning model & outcome. The second move is the methodology, here we discuss data collection, data preprocessing and model selection. We implement our research paper with six algorithms. Decision Tree classifier, Random Forest, Logistic Regression, SVM, K-nearest neighbors and XGBoost classifier. Third move is

Results and Discussion, XGBoost classifier is performed very well. Besides, Logistics regression, K-nearest neighbor and SVM also performed well. Finally, the worst results gave Random forest and Decision tree. Our dataset has 2 labels Satisfied and Unsatisfied which are also sentiments. Finally, we discuss the Conclusion and future research work.

II. Literature Review

The reason for writing a literature review is that many papers are read because the main purpose is to give you a summary of the paper, to help you consider what the paper is about. All display papers on other food reviews are briefly discussed. Islam et al. [1] focused on the development of analysis in product reviews and uses four Machine Learning classifiers like Random Forest, Support Vector Machine, Logistic Regression and Naive Bayes. Considering the accuracy of the four algorithms, the accuracy of Support Vector Machine was 55.40%, Random Forest 47.60%, Logistic Regression 51.60% and Naive Bayes 42.45%. Support Vector Machine and Logistic Regression gave good accuracy. Bhuiyan et al. [2] analyzed based on the reviews of the food ordered by the customers the taste, choice, objection, and behavior of the food. CNN (Convolutional Neural Network) and LSTM models are used. Analysis of the accuracy of the models showed that CNN based attention models gave 98.45% accuracy and on the contrary, CNN (Convolutional Neural Network) and LSTM models gave 98.34% and 98.23% accuracy respectively. This means that CNN-based attention models have the highest accuracy, and CNN and LSTM have the lowest accuracy. Li et al. [3] provided a method of obtaining sentence vectors by combining letter and word vectors, based on the BLSTM model. Using Att-CNN, CNN (Convolutional Neural Network), BLSTM and Att-BLSTM model. The process of concentration was introduced based on the semantics of the vocabulary in the food field to realize the extraction of distance-related sequence semantic features. The model showed that the BLSTM model performed better than CNN (Convolutional Neural Network). Rao et al. [4] analyzed Foodoholic applications by looking at the websites of different restaurants and extracting rank reviews of the recipes. Reviews were analyzed using a lexicon-based approach. Reviews were rated based on the scores of positive and negative words like delicious, bad, etc. They prepared the ordered list and displayed the correct output to the user. Anees et al. [5] analyzed product classification and response marketing through three categories namely Logistic Regression, SVM, Naive Bayes. In addition, four different weighting schemes like Term Frequency Inverse Class Frequency (TFICF), Term Frequency Inverse Document Frequency (TFIDF), X$\hat{2}$ Statistics (CHI) and Mutual Information have been used in the above three classes. The Logistic Regression classifier performed best with the MI term weighting scheme. Adnan et al. [6] categorized consumers into positive and negative comments from the restaurant website TripAdvisor with food reviews of different restaurants and its results would help them choose the best restaurant for consumers. Data were collected and processed in a programming language using WebHarvy software. In their paper, the Decision Tree-J48 method yielded an average value of Precision of 48.7%, 36.8% recall, 41.4% F-measure, and 45.6% accuracy. Ahmed et al. [7] processed sentiment analysis and strategies data from Amazon to Food Review through the Apache Spark Library. Using three algorithms in this paper namely Linear SVC, Logistic Regression, and NB we found that Linear SVC gave 88.38%, Logistic Regression gave 87.38% and NB gave 83.43% accuracy. That meant the Linear SVC classifier gave relatively good accuracy. Wu et al. [8] used a strategy to parse binary trees using a Stanford NLP parser with Fine Food Review data from Amazon. The results of Naive Bayes and RNNMS accuracy showed that the RNNMS model gave good results. That is, the learning rate of RNNMS is 0.1, while the learning rate of l2 regularization is 0.01. Alzami et al. [9] combined Word2Vector, TF-IDF, Word Bags. Word2Vector and TF-IDF extraction feature to find similarities between different machine learning models like SVM, Random Forest, Naive Bayes and KNN. tf-idf and BoW, as well as SVM's best performance, were available; The best results of random forest were found in the combination of W2V & TF-IDF; W2V & TF-IDF combination failed to give good results. TF-IDF can improve results by 87% with the help of SVM by reviewing the amazon food review dataset. Yu et al. [10] studied online-based consumers about their food feelings and experiences. Their paper analyzed data from 614 online reviews. And from there, 92% of customers discussed restaurant food, service, atmosphere, price, and location. Their research paper helped the customers to get an idea about the restaurant and to determine the quality of the food. Mohamad et al. [11] conducted sentiment analysis through product reviews in their paper. Where Four types of sentiments out of Five types of views (food, service, price, ambiance, and miscellaneous) were highlighted. Sentiment analysis was performed using naive Bayes and the best F1 score was 78.1% for the aspect-based sentiment. Dr. Wang Haoxiang [12] analyzed different data sets of bogus statistics. Information on social media is disseminated to everyone before its authenticity is verified. They did research for fake information and they used RNN(Recurrent Neural Network) and RNN classification gave 95% accuracy which gave better results than ANN (Conventional Artificial Neural Network). Dr. A. Pasumponpandian [13] Proposed an algorithm for social media E-project selection. The NSGA-II-MOIWO algorithm was developed which consists of two algorithms, MOIWO and NSGA-II. The new algorithm gave better results than these two algorithms in reducing the risk of EPPS (E-Project Portfolio Selection) problem. D. Saha et al. [14] mainly worked on hyperspectral image analysis to determine food quality and used supervised and unsupervised machine learning approaches. In his study k-means clustering achieved higher accuracy which is 83%.On the other hand logistic regression achieved lower accuracy. Analyzing many papers, no one has used XGBoost Classifier. We found this missing as well as we got better assurance using the XGBoost Classifier. Our dataset was suitable for the XGBoost classifier that's why we gain the highest accuracy from this.

III. Methodology

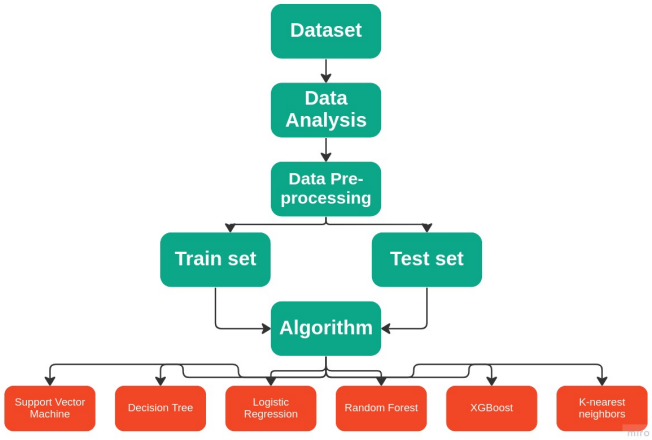

Fig. 1. Method Process

A. Data Collection Overview

This dataset is based on online data collection using Facebook, Twitter, and Google Maps where food quality is collected based on online food reviews. The dataset is divided into 6 columns, namely: Restaurant Name, Restaurant Food Type, Overall Review, Food Review Type, Customer Opinion, and Gender. In these columns we cover in detail - which restaurants the customers ate at, what was the food type of the restaurant, what kind of reviews they gave overall, food reviews in a specific way, customer ratings based on food quality, and their gender. Thus, data collection has been done from the said social media platform based online.

B. Pre-Processing

A plain text review dataset is basically unstructured data. So after data collection, the data needs to be processed. Because the data taken from online social media contain duplicate data, there is also data that is not understandable. Data processing is done to convert unique and understandable data. That is, to apply the data pre-processing method to our dataset, we follow the steps below-

- Remove URL, screen name, and hashtag. (We work with 500 data and some data has URL, so remove URL manually by unlinking)
- Remove symbols, emojis, punctuation's & numbers.
- Remove all retweets.

The need for pre-processing is-

- Remove noise
- Remove the missing values
- Null values
- Random error
- Irrelevant data

C. Analysis of Statistical

- The dataset includes 500 food reviews and data collected from online social media based.
- The dataset keeps 6 columns.
- The data type is string.
- The training portion contains 80% data (The data with which we will train the computer is called training data).
- The test portion contains 20% data. (And the data with which we will analyze the performance of the computer after the completion of training, how well it has done its work, we will call that data test data.)
- Reviews are classified into 2 steps (Satisfied defined as 1 and Unsatisfied defined as 0).

D. Confusion Matrix

Confusion matrixes are used to visualize important predictive analyzes such as recall, precision, accuracy and precision. The outcome of the prediction falls into two categories, namely the positive class and the negative class. Confusion matrixes are useful because they directly compare values such as true positive, false positive, true negative, and false negative. In the assessment procedure using the illusion matrix, the values of accuracy, retraction value and accurateness are obtained from the following equations.

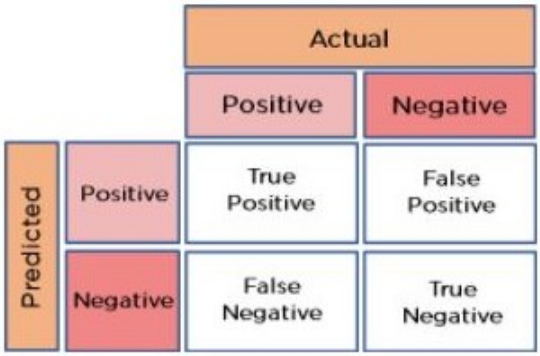

Fig. 2. Represent of confusion matrix terminology

- True Positive means actually positive and predicted value positive.
- True Negative means actually negative and predicted value negative.
- False Positive means actually negative and predicted value positive.
- False Negative means actually positive and predicted value negative.

IV. Practical Implications

A. Logistic Regression

Logistic regression is a supervised machine learning algorithm. which predicts classification problems through binary outcomes, such as yes or no, 0 or 1. It provides results based on probabilities of continuous variables versus categorical variables and predictive analytics algorithms.

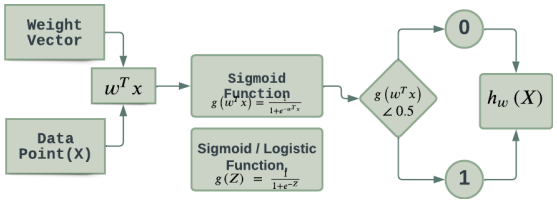

Fig. 3. Logistic Regression model

1) Summary of Logistic Regression Report: The Logistic Regression gives train accuracy of 81.06% and test accuracy of 82.00%. True positive - 9 data correctly predicted and the true value is true. True negative - 73 data points are not correctly predicted and the true value is false. False positives - 17 data correctly predicted and the true value false. False negative - 1 data not correctly predicted and the actual value is true.

TABLE I
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR LOGISTIC REGRESSION

	0	1	Accuracy	Macro-avg	WeightedAvg
Precision	0.90	0.81	0.82	0.86	0.83
Recall	0.35	0.99	0.82	0.67	0.82
F1-score	0.50	0.89	0.82	0.70	0.79
Support	26.00	74.00	0.82	100.00	100.00

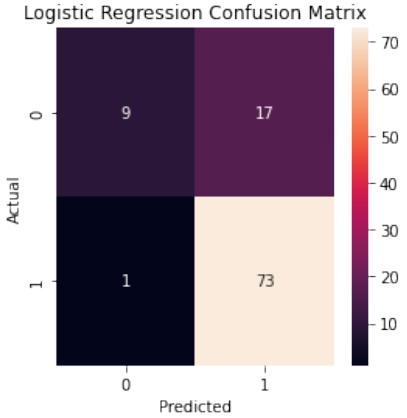

Fig. 4. Confusion Matrix of Logistic Regression

B. K-Nearest Neighbors

K-nearest neighbor or k-nn is a popular and very comfortable classification algorithm. Its practical application is very fruitful.

1) The main four steps of K-Nearest Neighbors are::

- Find the distance from the unknown data point to all remaining data points.
- Data points should be sorted from smallest to largest size (or, Ascending Order) according to the distance value.
- The first K-number of points needs to be taken from the sorted data points.

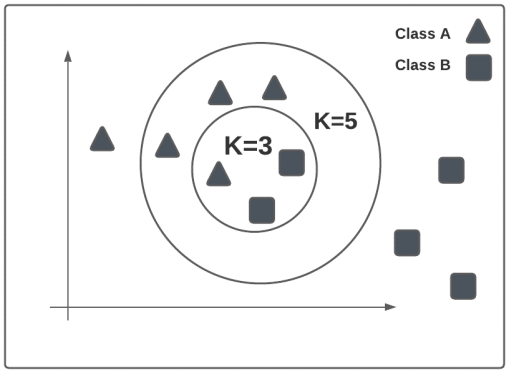

Fig. 5. Represent KNN algorithm terminology

- Among these K-number of data points, the class that has the most number of points, the unknown data point should be identified as being in that class.

2) Summary of K-Nearest Neighbors Report: The K-Nearest Neighbors gives train accuracy of 85.86% and test accuracy of 80.00%. True positive - 11 data correctly predicted and the true value is true. True negative - 69 data points are not correctly predicted and the true value is false. False positives - 15 data correctly predicted and the true value false. False negative - 5 data not correctly predicted and the actual value is true.

TABLE II
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR K-NEAREST NEIGHBORS

	0	1	Accuracy	Macro-Avg	Weighted Avg
Precision	0.69	0.82	0.80	0.75	0.79
Recall	0.42	0.93	0.80	0.68	0.80
F1-score	0.52	0.87	0.80	0.70	0.78
Support	26.00	74.00	0.80	100.00	100.00

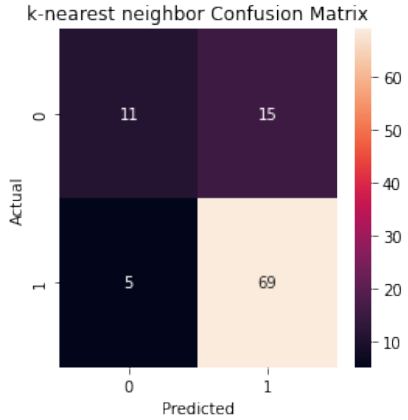

Fig. 6. Confusion Matrix of K-Nearest Neighbors

C. Support Vector Machine

Support vector machine which short form is SVM. It is the most broadly used and popular algorithm in machine learning. This algorithm has been used in many places, on many problem sets. There are numerous optimization methods for this, as well as various applications.

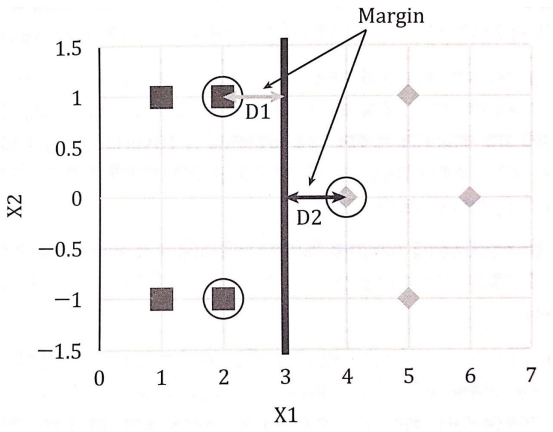

Fig. 7. Linearly Separable Data points in SVM

1) Summary of Support Vector Machine(SVM) Report: The SVM (Support Vector Machine) gives train accuracy of 90.91% and test accuracy of 80.00%. True positive - 8 data correctly predicted and the true value is true. True negative - 72 data points are not correctly predicted and the true value is false. False positives - 18 data correctly predicted and the true value false. False negative - 2 data not correctly predicted and the actual value is true.

TABLE III
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR SUPPORT VECTOR MACHINE

	0	1	Accuracy	Macro-Avg	Weighted Avg
Precision	0.80	0.80	0.80	0.80	0.80
Recall	0.31	0.97	0.80	0.64	0.80
F1-score	0.44	0.88	0.80	0.66	0.77
Support	26.00	74.00	0.80	100.00	100.00

D. Decision Tree Classifier

A decision tree is a supervised learning algorithm in machine learning that is used to solve problems like regression and classification but is more commonly used in classification problems. It is a tree-structured classification, where internal nodes represent features of a dataset and branches represent decision rules that help provide results. This decision matrix tree helps you decide on any type of process. The decision tree process starts with the data at the root node, which then selects the attributes given by the logic test of that attribute. The data concepts in Decision Tree 48 include:

- Data can be expressed in the form of a table that contains features and records.

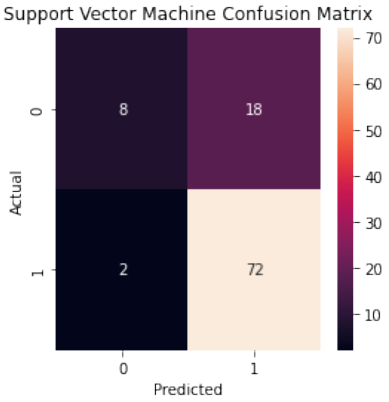

Fig. 8. Confusion Matrix of Support Vector Machine

- The properties show the parameters that are created as the criteria for tree formation A feature called data per-item data solutions is called Target Attribute.
- Features values that are called instances.

1) Summary of Decision Tree Classifier Report: The Decision Tree classifier gives train accuracy of 100.00% and test accuracy of 78.00%. True positive - 12 data correctly predicted and the true value is true. True negative - 66 data points are not correctly predicted and the true value is false. False positives - 14 data correctly predicted and the true value false. False negative - 8 data not correctly predicted and the actual value is true.

TABLE IV
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR DECISION TREE CLASSIFIER

	0	1	Accuracy	Macro-Avg	Weighted Avg
Precision	0.60	0.82	0.78	0.71	0.77
Recall	0.46	0.89	0.78	0.68	0.78
F1-score	0.52	0.86	0.78	0.69	0.77
Support	26.00	74.00	0.78	100.00	100.00

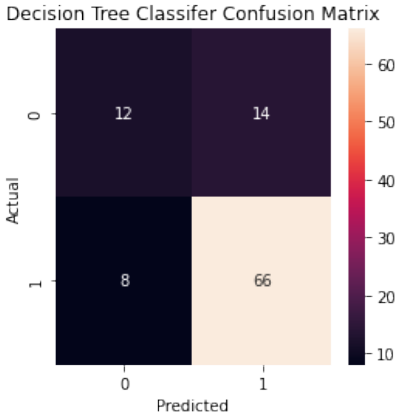

Fig. 9. Confusion Matrix of Decision Tree Classifier

E. Random Forest Classifier

Random forest is a widely used and popular algorithm in machine learning. Through this, all problems like classification and regression of machine learning can be solved. It averages to improve predictive elimination by adopting different machine learning classifiers like decision trees on various subsets of the dataset. Then solves the fitting problem leading to high accuracy by predicting the final result.

1) Summary of Random Forest Report: The Random Forest classifier gives train accuracy of 100.00% and test accuracy of 79.00%. True positive - 10 data correctly predicted and the true value is true. True negative - 69 data points are not correctly predicted and the true value is false. False positives - 16 data correctly predicted and the true value false. False negative - 5 data not correctly predicted and the actual value is true.

TABLE V
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR RANDOM FOREST CLASSIFIER

	0	1	Accuracy	Macro-Avg	Weighted Avg
Precision	0.67	0.81	0.79	0.74	0.77
Recall	0.38	0.93	0.79	0.66	0.79
F1-score	0.49	0.87	0.79	0.68	0.77
Support	26.00	74.00	0.79	100.00	100.00

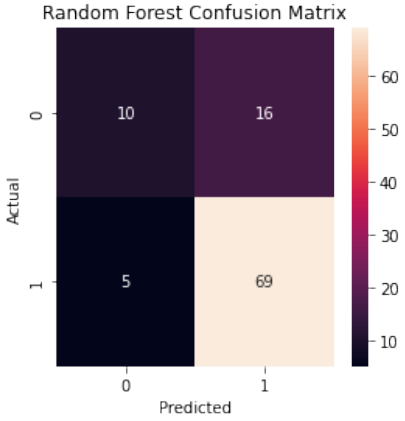

Fig. 10. Confusion Matrix of Decision Random Forest Classifier

F. XGBoost Classifier

XGBoost Classifier is an open-source execution of gradient boost decision trees. It implements ML algorithm under supervised learning algorithm and optimized distributed gradient boosting library. It provides a parallel tree boosting and distributed computing to solve various problems accurately and quickly.

1) Summary of XGBoost Classifier Report: The XGBoost classifier gives train accuracy of 87.37% and test accuracy of 83.00%. True positive - 11 data correctly predicted and the true value is true. True negative - 72 data points are not correctly predicted and the true value is false. False positives - 15 data correctly predicted and the true value false. False negative - 2 data not correctly predicted and the actual value is true.

TABLE VI
SCORE OF ACCURACY, RECALL, F1SCORE, ACCURACY FOR XGBOOST CLASSIFIER

	0	1	Accuracy	Macro-Avg	Weighted Avg
Precision	0.85	0.83	0.83	0.84	0.83
Recall	0.42	0.97	0.83	0.70	0.83
F1-score	0.56	0.89	0.83	0.73	0.81
Support	26.00	74.00	0.83	100.00	100.00

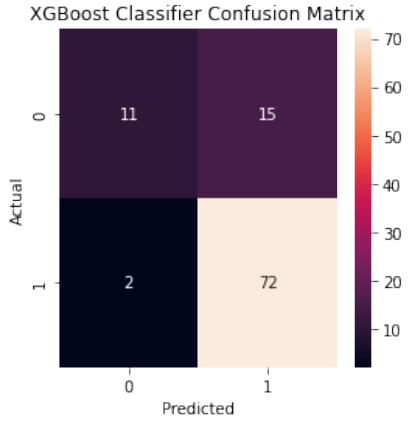

Fig. 11. Confusion Matrix of XGBoost Classifier

G. Feature Extraction:

Feature extraction is another necessary step before model training. Feature extraction techniques are operated to reduce the number of features in a dataset by creating a new feature dataset from the properties provided in the dataset. This technique is used when the model becomes difficult to fit with the dataset, the dataset has a large number of features. Feature extraction works to create models with fewer features each time. It basically considers the most varied features of the data set and then brings out the most important features for the model after re-evaluating the models by all possible features. There are many feature extraction techniques including TF-IDF, Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and many more. The TF-IDF technique is used to feature extraction. TF counts the words in the text and IDF estimates how common and rare words are in all documents or text.

V. EXPERIMENT AND ANALYSIS

This portion gives a consideration of the analysis and assessment result and also finding of the order report of utilized calculations. We split our dataset into two parts, one is training and another is test data. The number is 80% for trained data and 20% for test data. Our total data is 500. SVM, K-nearest Neighbors, Logistic Regression, Random Forest, XGBoost Classifier and Decision Tree were applied. Python is best for using Machine Learning to find accuracy. Sklearn is doing a good job on this. It helps to complete the model train.

TABLE VII
MULTICLASS CLASSIFICATION ACCURACY TABLE

Classifier	F1-score	Accuracy
SVM	80.00	80.00%
Decision Tree	78.00	78.00%
Logistic Regression	82.00	82.00%
Random Forest	79.00	79.00%
XGBoost Classifier	83.00	83.00%
K-nearest neighbors	80.00	80.00

A. *Analysis Diagram:*

Here we show six algorithms and their accuracy performance in the table. We find multiclass classification algorithms in the table and The XGBoost classifier gives very good accuracy. Logistic regression, Support Vector Machine and K-nearest neighbors give good accuracy. Decision trees and Random Forests give poor accuracy.

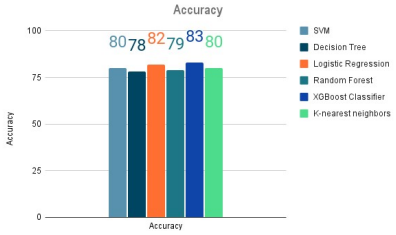

Fig. 12. Accuracy graph for food review analysis

This figure shows, the six algorithms we are use on our dataset. Here we see that XGBoost algorithm gives the best accuracy of all algorithms is 83%. Logistic regression after XGBoost gives a good accuracy of 82%, followed by both SVM and k-nearest neighbors at 80%. And Random Forest algorithm gives 79% accuracy. while the decision tree gives the lowest accuracy is 78%.

VI. CONCLUSION

Sentiment Analysis is important for better business, food quality & service improvements. For this, we need analysis with text data. We analyze 33 Bangladeshi restaurant food reviews using machine learning-based algorithms. This analysis consists of two different types of sentiments, satisfied and unsatisfied. Here we implement six prevalent machine learning algorithms to analyze our model. We start with SVM for better performance and gradually Decision Tree, Logistic Regression, Random Forest, XGBoost Classifier and K-nearest neighbors. But only the XGBoost classifier performs well and gives a better result. The best F1-Measure of XGBoost is 83%. Although our models don't reach a hundred percent accuracy but give us an excellent prediction for our model. In this research, our results and discussion will help a person to know about one restaurant to go to and also restaurant owners improve their business strategies and the current status of restaurants. The limitation of our food review analysis is only for Bangladeshi restaurants. In the future, Our research paper will improve. Better work can be done by increasing our dataset. And also use two popular deep learning architectures CNN and LSTM.

REFERENCES

[1] Islam, N., Akter, N. and Sattar, A., 2021, March. Sentiment analysis on food review using machine learning approach. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 157-164). IEEE.

[2] Bhuiyan, M.R., Mahedi, M.H., Hossain, N., Tumpa, Z.N. and Hossain, S.A., 2020, July. An Attention Based Approach for Sentiment Analysis of Food Review Dataset. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.

[3] Li, Y., Jin, Q., Zuo, M., Li, H., Yang, X., Zhang, Q. and Liu, X., 2020. RETRACTED: Multi-neural network-based sentiment analysis of food reviews based on character and word embeddings. The International Journal of Electrical Engineering & Education, p.0020720920928492.

[4] Rao, S. and Kakkar, M., 2017, January. A rating approach based on sentiment analysis. In 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence (pp. 557-562). IEEE.

[5] Anees, A.A., Gupta, H.P., Dalvi, A.P., Gopinath, S. and Mohan, B.R., 2019, May. Performance analysis of multiple classifiers using different term weighting schemes for sentiment analysis. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 637-641). IEEE.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] Ahmed, H.M., Javed Awan, M., Khan, N.S., Yasin, A. and Faisal Shehzad, H.M., 2021. Sentiment analysis of online food reviews using big data analytics. Hafiz Muhammad Ahmed, Mazhar Javed Awan, Nabeel Sabir Khan, Awais Yasin, Hafiz Muhammad Faisal Shehzad (2021) Sentiment Analysis of Online Food Reviews using Big Data Analytics. Elementary Education Online, 20(2), pp.827-836.

[8] Wu, J. and Ji, T., 2016. Deep learning for amazon food review sentiment analysis.

[9] Alzami, F., Udayanti, E.D., Prabowo, D.P. and Megantara, R.A., 2020. Document preprocessing with TF-IDF to improve the polarity classification performance of unstructured sentiment analysis. Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, pp.235-242.

[10] Yu, C.E. and Zhang, X., 2020. The embedded feelings in local gastronomy: a sentiment analysis of online reviews. Journal of Hospitality and Tourism Technology.

[11] Mubarok, M.S., Adiwijaya and Aldhi, M.D., 2017, August. Aspect-based sentiment analysis to review products using Naïve Bayes. In AIP Conference Proceedings (Vol. 1867, No. 1, p. 020060). AIP Publishing LLC.

[12] Haoxiang, W., 2020. Emotional analysis of bogus statistics in social media. Journal of Ubiquitous Computing and Communication Technologies (UCCT), 2(03), pp.178-186.

[13] Pasumponpandian, A., 2020. A Hybrid-Algorithm for E-project Selection on Social Media. Journal of Information Technology, 2(02), pp.116-122.

[14] Saha, D. and Manickavasagan, A., 2021. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review. Current Research in Food Science, 4, pp.28-44.