Classification of Food Reviews of Restaurant Based On Sentiment Analysis and Review Score

Md. Hasibul Hasan and Md. Abir Hasan

A Thesis in the Partial Fulfillment of the Requirements

For the Award of Bachelor of Computer Science and Engineering (BCSE)

Department of Computer Science and Engineering
College of Engineering and Technology
IUBAT – International University of Business Agriculture and Technology

Fall 2020

Classification of Food Reviews of Restaurant Based On Sentiment Analysis and Review Score

Md. Hasibul Hasan and Md. Abir Hasan

A Thesis in the Partial Fulfillment of the Requirements for the Award of Bachelor of Computer Science and Engineering (BCSE)

The thesis has been examined and approved,

_____

Prof. Dr. Md. Abdul Haque
Chairman and Professor

_____

Prof. Dr. Utpal Kanti Das
Co-supervisor, Coordinator and Professor

_____

Md. Sakibul Islam
Supervisor and Designation of the supervisor

Department of Computer Science and Engineering
College of Engineering and Technology
IUBAT – International University of Business Agriculture and Technology

Fall 2020

Abstract

Nowadays users of social networks are very much interested in expressing their opinions about different sorts of food services in social media which leads to the growth of user-generated web contents. Their reviews on social media have a significant impact on customers for making effective and optimal decisions for buying food using services. In sentiment analysis, most of the used approaches are based on machine learning techniques. In this paper, the well-known methods of machine learning are reviewed and compared against each other. Then the comparative studies on the performance of these techniques on online user reviews that come from multiple restaurant domains are performed. The experiments involve many different data sets from various domains including BD Restaurant. Well-known methods such as Support Vector Machine, Decision Tree, Bagging, Boosting, Random Forest and Maximum Entropy are implemented in the experiments. Based on the experimental results it is found that users can extract applicable information from review data sets for business intelligence and better food sales, and that Boosting and Maximum Entropy outperform the other examined machine learning algorithms for detecting sentiments in online user review scores and ratting. The Restaurant business increases day by day, also customer. For this situation customer have no enough knowledge about food quantity. For this reason, we develop a dataset about food review. We mention their quality of food it's Positive, Negative and Fake Review score by analyzing the sentiment of the review text and match with the review score to classify the reviews into various category.

Letter of Transmittal

20September 2020

The Chair

Thesis Defense Committee

Department of Computer Science and Engineering

IUBAT–International University of Business Agriculture and Technology

4 Embankment Drive Road, Sector 10, Uttara Model Town

Dhaka 1230, Bangladesh

Subject: Letter of Transmittal.

Dear Sir,

We are pleased to present to you our thesis report titled "Classification Of Food Reviews Of Restaurant Based On Sentiment Analysis And Review Score".

As required by IUBAT for the partial fulfillment of the requirements for the award of Bachelor of Computer Science and Engineering. It was indeed a great opportunity for us to work on this project to actualize our theoretical knowledge into practice. Now we are looking forward for your kind appraisal of our report.

Finally, we would like to thank you for giving us the opportunity to pursue our studies in your renowned university.

Yours sincerely,

Md. Hasibul Hasan	Md. Abir Hasan
17103121	17103272

Student's Declaration

We hereby declare that this thesis report titled 'Classification Of Food Reviews Of Restaurant Based On Sentiment Analysis And Review Score' is our original work. It has never been presented previously or concurrently for any other purpose, reward or degree at IUBAT University or any other institutions either by us or by any other student. We also declare that there is no plagiarism or data falsification and materials used in this report from various sources have been duly cited.

_____	_____
Md. Hasibul Hasan	Md. Abir Hasan
17103121	17103272

Supervisor's Certification

I certify that the student Md. Hasibul Hasan and Md. Abir Hasan carried out their thesis work 'Classification Of Food Reviews Of Restaurant Based On Sentiment Analysis And Review Score' between May 5, 2019 and December 19, 2020.During this period, they consulted me on regular basis as required by the department.

It therefore recommend that their thesis report be accepted for oral examination.

_____

Md. Sakibul Islam
Lecturer
Department of Computer Science and Engineering
IUBAT–International University of Business Agriculture and Technology

Acknowledgments

The success and final outcome of this thesis required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of our thesis. All that we have done is only due to such supervision and assistance and we would not forget to thank them. We respect and thank Md. Sakibul Islam for providing us an opportunity to do the thesis work and giving us all support and guidance, which made us complete the thesis duly. We are extremely thankful to him for providing such a nice support and guidance. We are also thankful to and fortunate enough to get constant encouragement, support and guidance from Prof. Dr. Utpal Kanti Das and all the faculties of Department of Computer Science engineering which helped us in successfully completing our thesis work. Also, we would like to extend our sincere esteems to all staff in laboratory for their timely support.

Table of Contents

List of Figures

List of Tables

1. Introduction

With the emerging use of the internet, people are showing their sentiment through different restaurant websites on social networks. From the reviews of the restaurant website, people can know and compare foods. To get knowledge from the review data, data mining techniques can be used for mining interesting, valid and significant patterns for the users. Different data mining techniques can be used for sentiment analysis. Analyzing sentiment or extracting opinion is the mathematical study of an individual's attitudes, appraisals, opinions, and emotions toward objects, issues, events, and topics. It is one of the extensively pursued fields of Natural Language Processing (NLP) and text processing. For sentiment analysis, it is very important to define the polarity of textual context which can be positive or negative. Methods of sentiment analysis can be categorized predominantly as machine learning, Lexicon-based and hybrid. In this research, the problem of determining the polarity of sentences of customer review data is studied. We have used the sentence-level classification for analysis. Given different sets of customer reviews, the task involves three subtasks. First of all, three real-world review data sets of different restaurant domains are used to determine the polarity of each sentence. Secondly, different machine learning approaches are applied to each data set and a comparative analysis is presented, and finally, the discovered information is summarized. The following example is used to illustrate the sentiment analysis process on sentence-level. Typical review sentenceslook like the following;

1.1 Food review

Positive: I bought this to use with my Kindle Fire and absolutely loved it!

Negative: This is a common food to eat, but the quality is unacceptable.

1.2 Movie review

Positive: In other words, the content level of this film is enough to easily fill a dozen other films.

Negative: A very, very slow-moving, aimless movie about a distressed, drifting young man.

For food review, in review 1, the potential customer felt positive about the food and in review 2 he gave a negative opinion about a feature. Similarly, for the movie review, in review 1, the potential customer gave a positive opinion about the movie and in review 2 his perspective was a negative opinion about it. So it is very important to get to know the manufacturer as well as the costumers about their viewpoints (whether positive or negative). For this study, structured reviews are used for training and testing purpose. Different classifiers such as Support Vector Machine (SVM), Maximum Entropy etc. are used for identifying and classifying review sentences from the data sets that are collected from Restaurant websites, Yelp and IMDb. A detailed analysis of those approaches is provided and their effects explained for understanding the significance of reviews. However, a useful summary is produced from the results of different approaches with these data sets.

As indicated above, the main contributions of this paper are:

(1) Different machine learning approaches- Support Vector Machine, Bagging, Boosting, Random Forest, Decision Tree, and Maximum Entropy are implemented to classify sentiment in sentence-level.

(2) A comparative study is performed considering various evaluation criteria that are- accuracy, recall, precision, f-score, ensemble agreement, and runtime of each algorithm.

(3) The results are analyzed with three real-world data sets from three different domains.

(4) Boosting and Maximum Entropy outperform the other examined machine learning algorithms for detecting sentiments in online user reviews.

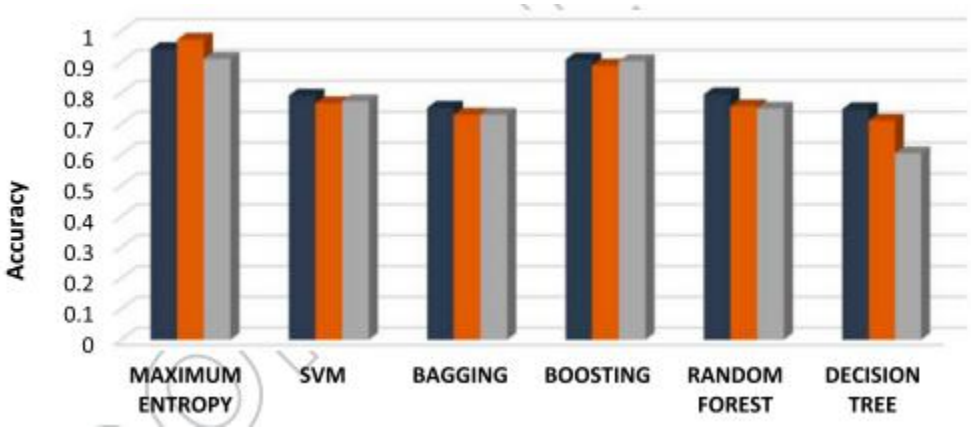

Figure 1.1: Accuracy analysis of algorithms on different review data sets based on cross-validation.

Different machine learning approaches- Support Vector Machine, Bagging, Boosting, random Forest, Decision Tree, and Maximum Entropy are implemented to classify sentiment in sentence-level.

1.3. Related work

A large number of papers focused on addressing the problem of sentiment classification. They used different techniques for classifying the sentiments of individuals at various levels. Different machine learning and lexicon-based approaches have been applied for sentiment analysis. A hybrid approach was proposed by Zhang et al. for sentiment analysis by combining both lexicon and learning based approaches and implemented on Twitter data. Not only supervised techniques but also unsupervised techniques can achieve a good accuracy which was presented by Turney. Dave, Lawrence, and Pennock designed a model of semantic orientation for positive and negative words with scoring which is used to classify the review data. Sentiment analysis is highly dependent on the topic domain and features of data. Hu and Liu presented a feature based stigmatization to determine the polarity of the review data. They used data mining and natural language processing techniques for identifying, summarizing and classifying the data as positive or negative. A lexicon-based approached was also studied by Taboada et al. in which dictionaries of words annotated with their semantic orientation. An unsupervised hierarchical Bayesian model was proposed by Lin and He. The document-level sentiment was classified by their approach to mine mixture of different topics. This is an extended version of the joint sentiment/topic (JST) and Latent Dirichlet Allocation (LDA) model for detecting the sentiment. For classifying cross-domain sentiment, Pan et al. applied Spectral Feature Alignment approach. A comparative analysis of both machine learning and lexicon-based approach for different web service was performed by Serrano-Guerrero et al. Since the existing techniques were not suitable, Liu, Hu, and Cheng proposed a novel technique to select the features of the products by analyzing the review data. They compared the opinions of the customers for different

products and demonstrated the results. Machine learning approach, rule-based approach, and lexicon based approach are used widely for sentiment polarity identification. Different machine learning algorithms such as -Naive Bayes Classifier, Max Entropy Classifier, Boosted Trees Classifier and Random Forest Classifier were implemented on text review data and after classification, an aggregate score was presented. Rule-based and lexicon-based approaches were compared with different machine learning approaches and a summary was represented in. A comparative analysis is performed on SVM and ANN in document-level sentiment analysis by Moraes et al. Tripathy et al. proposed a hybrid approach for combining SVM and ANN where SVM is used for feature selection and ANN to perform accuracy measurement. This hybrid approach provides better accuracy than single machine learning approaches. Multi-label sentiment classification is performed by Liu et al. and 11 different states are considered and compared with different approaches. Recently domain specific aspect-based sentiment analysis is widely used. Based on topic modeling and unsupervised technique an approach is proposed to sentiment classification in aspect level. This approach is performed and analyzed with different domain and languages.

2. Literature Review

2.1 Key Concepts, Theories and Studies

Early work on detection crime victim and created dataset, these approaches usually apply A person who was did crime. They are collecting the data and analysis and classify and categories the crime. They are finding the product rating which product is good they are giving accuracy. They are only excrement only Three E-Commerce website. They are using SVM Algorithm for different review data sets based on prediction. They can develop a dataset for one E-Commerce Website. They are giving rating which product are good and get satisfactory accuracy. experiments involve many different data sets from various domains including Amazon, Yelp and IMDb. Well-known methods such as Support Vector Machine, ecision Tree, Bagging, Boosting, Random Forest and Maximum Entropy are implemented in the experiments. Based on the experimental results it is found that users can extract applicable information from review data sets for business intelligence and better product salesproduction, and that Boosting and Maximum Entropy outperform the other examined machine learning algorithms for detecting sentiments in online user reviews.

2.2 Key Debates and Controversies

We constructed embedding matrices for meaningful feature extraction using Word Embedding (Word2Vec), Bag of Words , TF-IDF, Global Vectors for Word Representation

(GloVe),fastText. We conducted several experiments to compare the performance of different machine learning algorithms, Support Vector Machine(SVM), Logistic Regression, Naive Bayes using TF-IDF and Bag of Words based on numerical feature dictionaries. Again, we conducted several experiments to compare the performance of different learning models of Artificial Neural Networks(ANN), for example Convolutional Neural Network(CNN), Recurrent Neural Network (RNN) and also their deep and shallow structured model using semantic feature based techniques, for example GloVe, FastText and Word2Vec. Finally we gave a comparison performance between machine learning algorithms as state-of- art and the Deep Neural Network algorithms in the field of Bangla Text Classification. SVM is one of the dominant classifier algorithms. The main concept of SVM is to construct a linear separator in a hyperplane that can best separate the input data into different classes. The hyperplane is learned by using an optimization procedure in training data. After that, the largest margin is selected in the hyperplane so that it can optimally separate data set into classes. For input vector x, weight vector w and bias b, the linear separator can be defined as

$$yi(w.xi + b) \geq 1 \text{ for all } 1 \leq i \leq n$$

Boosting is a good ensemble learning method for text classification Random Forest is very simple supervised ensemble learning method for classification. For prediction and decision making, a forest is created by constructing multiple trees. From the random subsets of features, the best one is selected for splitting a node in the forest. From p input variables, m numbers of variables are selected as m < f. finally, the average prediction value of its trees is used for the final prediction. Mathematically They used a subset Amazon video game user

reviews from UCSD Time-based models didn't work well as the variance in average rating between each year month, or day was relatively small.

2.3 Gaps in Existing Knowledge

We are reading some paper, for our topic we are read 1 st paper Performance Analysis of Most Prominent Machine Learning and Deep Learning Algorithms in Classifying Bangla Crime News Articles. Basically, they are developing a dataset which can show the crime category and the are develop a dataset which can show the crime information. This paper limitation is if some criminal use their system then criminal easily find out the new category for doing crime. And they are not collecting 100% accuracy so there is some lack.2 nd we are reading Sentiment Analysis of Online Food Reviews using Customer Ratings, which can develop dataset which can review only restaurant, if this system using any restaurant then many employees lost their job because rest of the working done by the system. 3 rd Restaurant Review Analysis and Classification Using SVM, they are developing a system which is review the restaurant, there is no food review. So, customer only get information about restaurant not food. 4 thAn empirical research on sentiment analysis using machine learning approaches, their experiments involve many different data sets from various domains including Amazon, Yelp and IMDb. Well-known methods such as Support Vector Machine, Decision Tree, Bagging, Boosting, Random Forest and Maximum Entropy are implemented in the experiments. Based on the experimental results it is found that users can extract applicable information from review data sets for business intelligence and better product sales production, and that Boosting and Maximum Entropy outperform the other examined machine learning algorithms for detecting sentiments in online user reviews. But

they are working a smaller number of websites, so customer cannot get proper information about other product.

2.4. Detecting food in images using CNN

CNN has been utilised for food image detection. This problem can be condensed down to a simple binary classification problem (food/nonfood). The purpose of food image detection process is to first determine iffood is present within an image or video. In regards to a food image recognition pipeline, this would be the first stage in food image recognition framework i.e. determining if the image contains food or not. In Ref. GoogleNet pretrained model was fine-tuned using Food-5K dataset. The training process in Ref. utilised a subset of Food-5K data using 1000 iterations. The learning rate was changed to 0.01 and the learning rate policy was polynomial. Results from Ref. achieved 99.2% accuracy in determining food/non-food classes. Other research also utilised CNNs for food detection [14] and used 6-fold cross validation with different hyper-parameters to determine optimal settings and experiments achieved 93.8% in food/non-food detection. It is clear from research that CNNs can be used effectively for food detection in images.

2.5. Predicting food type in images using CNN

Extensive research has been carried out in utilising CNN for food item recognition. The food item recognition process would take place after the food detection phase in which the actual food item is then predicted within food image. In Ref. [15] CNNs were utilised to extract features from convolutional layers in order to classify food items and food groups, experiments achieved 70.13% for 61 class dataset and 94.01% for 7 class datasets.These

experiments used AlexNet deep features with a SVM classifier applied to PFID dataset. In Ref. the aim of the work was to compare conventional feature extraction methods with CNN extraction methods utilising UECFood-100 dataset. Results from Ref. achieved 72.6% accuracy for top-1 accuracy and 92% for top-5 accuracy. Also in Ref. as well as performing food/non-food experiments, food group classification was performed. A CNN was developed and was trained using extracted segmented patches of food items The food items used in this work were based around 7 food major types. The patches were then fed into a CNN using 4 convolution layers with different variations of filter sizes and using 5 kernels to process the patches. Results in Ref. achieved 73.70% accuracy using 6-fold cross validation. These studies confirm that CNNs provide an efficient method for food image recognition to provide for accurate food logging to promote dietary management.

3. Research Methodology

3.1 Research design

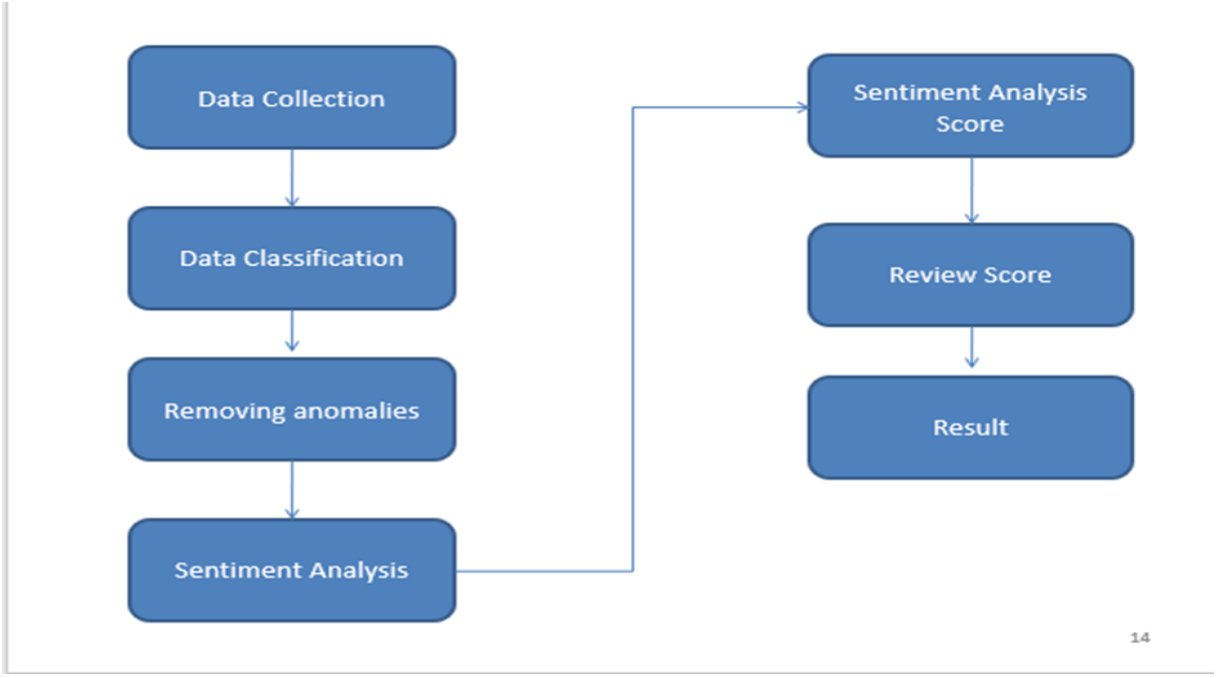

Figure 3.1: working process

For our thesis we are following this process structure. At first, we will collect data from many restaurant websites. After that we are classification those data, then we are working removing anomalies for those data. After anomalies then we are sentiment those data and given a score about those data. Again, are collecting food review score and compare those data and given a final result with accuracy score.

3.2 Methods and Sources

For our thesis we are using TF-IDF algorithm with Machine Learning and Deep learning, NLP, SVM algorithm. Also, we develop a Dataset based on customer review about food.

1. TF-IDF is a statistical measure that evaluates how relevant a word is to a document in a collection of documents. It has many uses, most importantly in automated text analysis, and is very useful for scoring words in machine learning algorithms for Natural Language Processing (NLP).
2. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.
3. Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabelled. Also known as deep neural learning or deep neural network.
4. SVM is a supervised machine learning algorithm which can be used for classification or regression problems. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs.
5. We are using those tools for develop our Dataset.

3.3 Practical Considerations

Practically we are working on Python and collecting data from different kind of restaurant. When we are collecting data then we were observed many restaurants removing their bad review which is collecting from the customer. So, we cannot find data easily. But we are hardworking and collect many data from many types of restaurant. There is some limitation. Many restaurant owners think if we are giving review negative about their restaurant then customer cannot buy their food. But we are said them, we are just collecting the customer review on food, nothing else. After that, we are analysis those data on sentiment analysis. Before sentiment analysis we are categories our data then collect sentiment analysis score. After getting sentiment analysis score then we are comparing those scores with review score which we were collecting from the restaurant website.

1.Procedure Scoreorientation (score t ,summary s_i)

2. begin

3.gor each opinion word op in summary s_i

4. if(t>3)

5. s_i orientation = positive

6. else

7. s_i orientation = negative

8.endfor

9.end

Text preprocessing Text preprocessing is an important part of opinion mining. The unstructured and unorganized text data needs further processing to extract predominant

knowledge from it. The key steps of preprocessing are syntax and semantic preprocessing. Stop words elimination Stop words are extremely common words but contain little significance for any analysis. The elimination of stop words makes the text less heavy and reduces dimensionality. Determiners, coordinating conjunctions, prepositions etc. are some types of stop words that are not meaningful for the documents. Some customized keywords can also be used based on the content of the testing data. For text mining applications, stop words are eliminated from documents as those words are not appraised as keywords.

Decision Tree classifier can be used for text classification with a small variety of different Food. A successor of Decision Tree algorithm C5 was used for document classification Lewis and Ringuette made a comparative study of Decision Tree and Naïve Bayes algorithm for text categorization. Here, to perform split in a single attribute, the existence of phrases or words in a tree at a specific node was inspected.

Data sets collection In order to evaluate and compare the different approaches of sentiment analysis, customer review data from the UCI machine learning repository1 is used. This review data contains three different real-world data sets which are collected from Amazon, Yelp and IMDb. Each data set contains Sentiment Labeled Sentences with positive or negative sentiment. The characteristic of the data set is text type containing 3000 number of instance. It contains positive or negative sentiment labeled sentences with corresponding sentence scores. The score is either 1 (for positive) or 0 (for negative). The first data set is of review and accessory category that is collected from Amazon.com. It contains reviews about cell phones that are sold on Amazon. In this review data set, the 1000 labeled sentence consists of total 11,248 words. The next data set is about restaurant reviews that are extracted from Yelp which consists of 11,894 words. The third and last sentiment data set regards movie review that is referred from

IMDb. This is the largest data set with 15,366 words for 1000 sentences. For each of the data sets, 1000 labeled sentences are split off for testing and training. Among them, 50% represent positive polarity and the rest 50% represent negative polarity. These data sets are used to evaluate sentence-level classifiers and model training.

Steps of sentiment analysis (1) Data collection: Data collection is the very first step to sentiment analysis. Data can be collected from different Restaurant, social networks, blogs, review sites and so on. This unstructured and unorganized text data of individual domains are used for the next steps. (2) Text preprocessing: Text preprocessing helps to make data useable for further analysis. Removal of digits, punctuation, stop words, case conversion, and tokenization make the unprocessed data organized and efficient. (3) Sentiment classification: In this step, the polarity of each sentence is classified into positive or negative. (4) Presentation of output: At the last step, the unstructured text data is transformed into significant and useable information. The resultant output is represented graphically using charts and purposeful.

4. Result and Discussion

This section gives an overview of accuracy rates of the trained classifiers. All the calculations are done in Weak tool which runs on java virtual machine. Sentiment Analyzer Results The above sections discussed the method followed to train the classifier used for sentiment analysis of tweets. The classifier with features as Word2vec representations of human annotated tweets trained on Random Forest algorithm with a split percentage of 90 for training the model and remaining for testing the model showed an accuracy of 70.2%. With N-gram representations, the classifier model with same algorithm and with same dataset showed an accuracy of 70.5%. Though the results are very close, model trained with word2vec representations is picked to classify the nonhuman annotated tweetsbecause of its promising accuracy for large datasets and the sustainability in word meaning. Numerous studies have been conducted on people and they concluded that the rate of human concordance, that is the degree of agreement among human son the sentiment of a text, is between 70% and 79%. They have also synthesized that sentiment analyzers above 70% are very accurate in most of the cases. Provided this information, the results we obtained from the sentiment classification can be observed as very good figures while predicting the sentiments in short texts, tweets, less than 140 characters in length. Table-2 depicts the results of sentiment classification including accuracy, precision, F-measure and recall when trained with different machine learning algorithms. ROC curves are plotted for detailed analysis. B. Stock Price and Sentiment Correlation Results A classifier is presented in the previous sections that is trained with aggregate sentiment values for 3-day period as features

and the increase/decrease in stock price represented by 1/0 as the output. Total data is split into two parts, 80 percent to train the model and remaining for testing operations. The classifier results show an accuracy value of 69.01% when trained using Logistic regression algorithm and the accuracy rate varied with the training set. When the model with LibSVM is trained with 90 percent of data, it gave a result of 71.82%. These results give a significant edge to the investors and they show good correlation between stock market movements and the sentiments of public expressed in twitter. This trend shows that with increasing dataset the models are performing well. We would like to incorporate more data in our future work. we demonstrate our experiment results and assess performance based on various machine learning classifiers. We run the classifiers on each data set and compare the performance based on cross-validation, recall, precision, f-score, ensemble agreement, and runtime. We got Bag of Words and TF-IDF vectors from our dataset, which were used in training state of the art machine learning algorithms (Naive Bayes, Logistic Regressing and Support Vector Machine). We also got the three separated embedded matrices from these different approaches that are Word2Vec, fastText and GloVe. After that these results had been used both for CNN and RNN. The dimensions which were set and the total number of words in our feature dictionary are given in table II. These embedded matrices had been used for training our deep neural network (CNN and RNN) models. We calculated accuracy, precision, recall and F1 score by validating with our test data for all machine learning algorithms used, which are displayed in table III and IV. We had done experiments in Google Colab environment. It gave us an environment with Intel(R) Xeon(R) CPU @ 2.30GHz, 12 GB of RAM and a GPU. The execution time for the deep learning algorithms depends on number of hidden layers. So deep CNN would take much time to train than shallow CNN.

And memory required for these deep learning algorithms depends on the dimensions and number of hidden layers. In order to generate an optimal neural network that classifies instances in question to good or not good categories, the intelligent problem solver (IPS) has been repeatedly launched, generating neural linear networks, MLP and RBF with diverse architecture in subsequent cycles; out of the generated set of best networks in any given cycle it retained 10 containing the least learning, validation and testing inaccuracies. For every given network a determination of the classification quality through the designation of classification coefficients of learning, validation and test sets has been conducted. The classification coefficient Q is the ratio of correctly classified cases to the total case number in a set. The Q coefficient is also expressed by a percentage (Q%) of correct classified cases in an entire set. The IPS used procedures that allowed it to avoid network 'over-learning' during network generationi.e. excessive matching of the neural model to learning data. Network over learning is characterized by a high degree of a hyper spectral curve modeling the input signal space relation, owing to the fact that the function represented by the network usually curves in a way that allows it to pass through all the learning points. A model defined by the MLP with an s-shaped (logistic or hyperbolic) activation function is characterized by a curve increasingly higher the greater the weight values are. One of the methods of preferring models that relate to a smaller hyper spectrum is the tendency towards learning procedures that guarantee smaller weights. In the Weygand network reduction method it is achieved by attaching an additional segment to the error function. The segment 'punishes' the network whenever it tries to set higher weight values in the learning process. Because of that the network tolerates only higher weights that are necessary to create a model of the problem in question; the remaining weights are zeroed.

5. Conclusion

In this paper, we present a vast experimental analysis of sentiment classification using different well-known machine learning approaches. Different classification techniques such as Maximum Entropy, SVM, Bagging, Boosting, Random Forest and Decision Tree are applied to conduct the experiments. For these experiments, three well-known review data sets collected from different restaurant websites domains such as food, restaurants, and movies are used. To conduct the experiments, these review data sets are used to evaluate these models and classify sentiment polarity. For every different data sets accuracy, performance factors (recall, precision, f-score, and ensemble agreement) and runtime are considered to evaluate the result. Finally, the results are analyzed, compared and presented through a statistical approach. For further research, the following directions could be followed to extend the work. Firstly, mining not only positive or negative polarity but also neutral opinion from review data. Secondly, making a classifier which will be topic and domain independent. Thirdly, combining the existing approaches to make a new model that can perform better with a large amount of data. At last, collaborating with various organizations to observe the influence on products with the analyzed results of review data and making advance decisions for the businesses. Sentiment analysis is the process of identifying the feeling expressed in the text or document. We proposed a methodology for mining the food reviews based on score combined with existing text analyzing packages. The proposed system has produced a very good result using the score ratings. The limitation of this system is, it works better only for the open sentiments like rating or scores. The results were not

promising for hidden sentiments. In Future work, prediction-based methods will be implemented with existing approach. More features will be extracted to handle the implicit sentiment analysis.

References

Dębska, B. and Guzowska-Świder, B., 2011. Application of artificial neural network in food classification. Analytica Chimica Acta, 705(1-2), pp.283-291.

De Albornoz, J.C., Plaza, L., Gervás, P. and Díaz, A., 2011, April. A joint model of feature mining and sentiment analysis for product review rating. In European conference on information retrieval (pp. 55-66). Springer, Berlin, Heidelberg.

Eshan, S.C. and Hasan, M.S., 2017, December. An application of machine learning to detect abusive bengali text. In 2017 20th International Conference of Computer and Information Technology (ICCIT) (pp. 1-6). IEEE.

Haque, T.U., Saber, N.N. and Shah, F.M., 2018, May. Sentiment analysis on large scale Amazon product reviews. In 2018 IEEE International Conference on Innovative Research and Development (ICIRD) (pp. 1-6). IEEE.

Kabir, M., Kabir, M.M.J., Xu, S. and Badhon, B., 2019. An empirical research on sentiment analysis using machine learning approaches. International Journal of Computers and Applications, pp.1-9.

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.

McAllister, P., Zheng, H., Bond, R. and Moorhead, A., 2018. Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. Computers in biology and medicine, 95, pp.217-233.

Ortigosa, A., Martín, J.M. and Carro, R.M., 2014. Sentiment analysis in Facebook and its application to e-learning. Computers in human behavior, 31, pp.527-541.

Sasikala, P. and Sheela, L.M.I., 2018. Sentiment Analysis of Online Food Reviews using Customer Ratings. International Journal of Pure and Applied Mathematics, 119(15), pp.3509-3514.

Sheela, L.M.I., SENTIMENT ANALYSIS OF ONLINE FOOD REVIEWS USING CUSTOMER RATINGS.

Pagolu, V.S., Reddy, K.N., Panda, G. and Majhi, B., 2016, October. Sentiment analysis of Twitter data for predicting stock market movements. In 2016 international conference on signal processing, communication, power and embedded system (SCOPES) (pp. 1345-1350). IEEE.

Waikul, V., Ravgan, O. and Pavate, A., Restaurant Review Analysis and Classification Using SVM. optimization, 2, p.7.