The Impact of Spam Reviews on Feature-based Sentiment Analysis

Nagwa M.K. Saeed
nagwa.moustafa@cis.asu.edu.eg

Nivin A. Helal
nivin.atef@cis.asu.edu.eg

Nagwa L. Badr
nagwabadr@cis.asu.edu.eg

Tarek F. Gharib
tfgharib@cis.asu.edu.eg

Information Systems Department
Faculty of Computer & Information Sciences
Ain Shams University, Cairo, Egypt

***Abstract*—With the explosive growth of online social media, people can convey their experience and share with others on a common platform. People's opinions and experience became the most important source of information in the process of sentiment analysis for making decisions. The performance of opinion mining depends on the availability of trustworthy opinions for sentiment analysis. Unfortunately, spammers write spam reviews which can be positive or negative opinions in order to promote their services or damage the reputation of their competitors' services. Spam reviews may mislead potential customers and affect their experience and influence their ideas. These spam reviews must be identified and removed to avoid possible deceitful customers. The main objective of this paper is to present an enhanced feature-based sentiment analysis algorithm that improves the performance of sentiment classification. The proposed algorithm is developed to assign accurate sentiment score to each feature in social reviews by considering spam reviews detection. The proposed algorithm also examines the effect of three different feature extraction methods on the performance of sentiment classification. Finally, the results indicate that the proposed algorithm achieves an accuracy of about 79.56% in classifying opinions.**

***Keywords*—Opinion mining; Sentiment analysis; Feature extraction; Spam reviews detection**

I. INTRODUCTION

Recently, the World Wide Web (WWW) has come to be a massive source of user-generated content and opinionative data. Several websites mainly social networks such as Twitter and Facebook encourage users to publicly express their feelings in their daily interactions and exchange their views, suggestions and opinions related to products, services, etc. The increased popularity of these sites resulted in a huge collection of people opinions and valuable information on the web in much unstructured manner for the decision making process. The extraction of the useful content from these sites and its analysis became a challenging task. This situation created the research area called opinion mining and sentiment analysis [1].

Opinion mining and sentiment analysis are considered as the extension of data mining which makes use of natural language processing techniques to extract and classify humans' opinions expressed on entities or features of entities in an automated manner where entities can be products, services, persons, organizations or events while features are attributes or components of the entities [2]. This helps in improving and addressing how many users support a sentiment whether it is against or in favor of a topic. It also helps in identifying the like-minded cluster of people and their relationship through sentiment mining [3].

Sentiment analysis can be performed at different levels; there are commonly three levels of analysis: Document, Sentence and Feature. The level of sentiment analysis is determined based on the type of data that will be processed [4].

Document level sentiment analysis is the simplest form of classification. The whole document is considered as the basic unit of information. It is assumed that the document has opinions about a single entity only. This approach is not suitable if the document includes opinions about different entities as in forums and blogs. Document level sentiment analysis is performed by classifying the whole document as positive, negative or neutral [4].

Sentence level sentiment analysis is the most fine-grained analysis of the document level. In this level, each sentence is considered as a separate unit and each sentence is classified as positive, negative or neutral [4].

Feature level sentiment analysis is an essential task for sentiment analysis where the text is analyzed and features of entities are extracted. Then these features are further analyzed one by one in terms of polarity and are given a score then based on this either the entire document or the sentence is considered as positive, negative or neutral [5].

Feature level sentiment analysis can be useful for applications in diverse fields as this level produces better sentiment information because of the fact that when a user writes an opinion, it does not mean that the user likes or dislikes everything about the service or the product. That is, users give points of view on several features that can be positive and negative. This information is essential not only for customers but also for organizations as it helps in making a decision concerning buying a product, and it can serve as the basis to enhance services and products [5].

Sentiment analysis can be performed using two different approaches namely lexicon based sentiment analysis and machine learning sentiment analysis. Lexicon based sentiment analysis is an unsupervised approach based on existing sentiment lexicons. It seeks for sentiment words in a review, and then predefined lexicons are used to determine their polarities. Lexicon based sentiment analysis is domain independent which means that for all domains a single lexicon is built. It is an appealing advantage over machine learning methods, as lexicon based approach has a better performance across texts and domains. Also, this approach has the advantage that general knowledge sentiment lexicons have wider term coverage and also there is no need for labeled data and for the procedure of learning, regardless of how these words are used in a text.

Machine learning sentiment analysis is divided into supervised and unsupervised learning methods. The supervised methods are methods in which a model is learned with a training dataset for predicting sentiments (i.e., positive, neutral, negative). The unsupervised methods are used when it is difficult to find these labeled training datasets. The ability to adapt and construct trained models for particular purposes and contexts is considered as the main merit of machine learning approaches, and in practice machine learning approaches demonstrate the high accuracy of classification. However, this accuracy is achieved only with a representative collection of labeled training datasets. At the same time the limitation is that it is difficult to integrate, general knowledge which may not be acquired from training data, into a classifier.

Lexicon based sentiment analysis is more appropriate for feature level sentiment analysis and has high precision, while machine learning sentiment analysis is more appropriate for document level sentiment analysis.

There are several challenges that face opinion mining and sentiment analysis process. These challenges become obstacles in detecting the accurate sentiment polarity [6]. One of these main challenges is the detection of spam reviews. Since reviews can be freely written by anyone anonymously, there has been abuse of this anonymity feature by persons posting unethical or even fraudulent reviews to deceive consumers in order to achieve significant business advantages. Such reviews are called spam opinions and the persons indulging in these illegal activities are called spammers. Spammers may write spam reviews which can be positive opinions to promote their products or negative opinions to damage the reputation of their competitors' products. Spam opinions may prevent customers and corporations from concluding ideas about the products. Therefore, they greatly influence the online marketing and trading enterprise [7]. Therefore, Spam reviews detection is needed to identify and filter out spam opinions to provide real and trustful review service and to authenticate online opinions and gain consumer trust [8].

The contribution of this paper is divided into two main parts. First, proposing a spam detection algorithm for online reviews to enhance the accuracy of sentiment analysis process. Second, examining the impact of three different feature extraction methods on the accuracy of the feature sentiment score.

The structure of this paper is as follows: section 2 provides a brief on related works, section 3 describes the proposed algorithm, section 4 shows the conducted experimental study and the outcome results from the proposed algorithm and finally section 5 describes the conclusion and the future work.

II. RELATED WORK

There have been many researches on the subject of sentiment analysis. In this section, we present an overview on the previously proposed work which deals with sentiment analysis.

According to the sentiment analysis levels previously illustrated, some researchers applied feature-based sentiment analysis for example in [4], [5], [9] to discover sentiments on features of entities and classify the reviews based on their features.

In [4] a feature-based sentiment analysis system is proposed where the collected reviews are sent to the Part of speech tagging process to identify the features and opinion words from the reviews. All the features are extracted from the reviews and then opinion words corresponding to them are extracted. The polarity is determined for each feature using WordNet lexicon and then the polarity of the sentence is identified based on majority of opinion words, if the count of positive words is greater than the count of negative words, then the polarity of the sentence is positive otherwise the polarity of the sentence is negative. In [5] also a feature-based sentiment analysis method for the health domain based on ontologies in the diabetes domain is proposed. The sentiments of the features of each review that are close to the extracted opinion words are obtained based on lexicons related to the diabetes domain and SentiWordNet lexicon. In [9] also a feature-based sentiment analysis system is performed in which features that customer had expressed their opinions on are identified. Part of speech tagging is used and both nouns and noun phrases are extracted as product features and only words tagged as adjectives are identified as opinion words. The orientations for these opinion words are determined using SentiWordNet lexicon.

In [7], [8], [10], [11], [12], [13], [14], [15], [16] spam review detection which is an important task in opinion mining is discussed. Opinion spam detection techniques mainly rely on three information sources: review text, reviewer characteristics and sentiment information. Review text is a foremost source of information as other information is not available in most of the related datasets.

A spam review detection technique is presented and some rules related to reviewers' behavior and reviews' metadata are considered for detecting spam reviews in [7], [8], [12], [13], [14], [15]. In [10], [16] a set of discriminative rules is presented and combined with time series method in which some information is taken into consideration such as the number of reviews, posting time and ratings posted for different services in a certain time period to find out spam reviews. In [11] an automatic review spam detector is developed based on indicating the inconsistency between the sentiment score of the review and the rating given by the reviewer.

A Survey that covers the existing techniques and approaches for sentiment analysis is presented and also the challenges in this field are discussed in [1], [2], [3], [6], [17], [18], [19].

The previously mentioned studies either detected spam reviews but didn't consider their impact on the accuracy of sentiment classification or applied feature-based sentiment analysis without detecting spam reviews which lead to low accuracy in the classification of the users' sentiments.

III. THE PROPOSED ALGORITHM

The proposed algorithm is a feature-based sentiment analysis algorithm that uses a lexicon based approach and identifies accurate sentiment score for each feature in social reviews by considering spam reviews detection. In the proposed algorithm we aim to prove the impact of both using the proper feature extraction method and spam reviews detection on sentiment analysis process.

The main contribution of this proposed algorithm is applying feature-based sentiment analysis along with spam reviews detection to improve the accuracy of the feature-based sentiment analysis process. Moreover, it also improves the accuracy by comparing the impact of using three different feature extraction methods and selecting the method that yields the most accurate results.

The proposed algorithm works as follows: First, suitable reviews are collected and stored in the reviews dataset. Preprocessing is applied on the collected reviews dataset. Second, the preprocessed reviews are filtered by detecting and filtering out spam reviews. As a third phase, features and opinion words are extracted from these filtered reviews. The polarity of each one of the extracted opinion words is identified. Finally, the accumulative sentiment score of each feature is calculated, and also the reviews are categorized as positive or negative. The overall steps are illustrated in Algorithm 1.

Algorithm 1 Enhanced Feature-based Algorithm for Sentiment Analysis

Input: Reviews Dataset RD
Output: Review Polarity P

1: **for each** Review R in RD **do**
2: Preprocessed Review R_P = Apply Data preprocessing (R)
3: Calculate number of words in R_P
4: Calculate number of pronouns in R_P
5: Calculate number of unique words in R_P
6: **if** R_P has less words and R_P is diverse and R_P has less pronouns **then**
7: Truthful Review $R_T = R_P$
8: **end if**
9: Extracted Features EF = Extract features from R_T
10: Extracted Opinion Words EOW = Extract opinion words from R_T
11: Get the sentiment score for EOW from sentiment lexicon
12: Compute distance between EF and EOW
13: Assign sentiment scores of EOW to the nearest EF
14: **if** Sum of EF sentiment scores in $R_T > 0$ **then**
15: P = Positive
16: **else**
17: P = Negative
18: **end if**
19: **end for**

In the following subsections, we will explain in details the previously mentioned algorithm.

A. Preprocessing

Preprocessing process is applied to represent social reviews in a form that could be analyzed efficiently. Social reviews are filtered by applying different preprocessing tasks that include: Removing stop words which are a set of words that occur regularly without any additional information (articles, determiners and prepositions), Removing special characters and symbols (*, /, $, ?, etc.), Removing non-English words, Applying tokenization by splitting the text into tokens, Applying lemmatization by extracting the proper lemma, Excluding duplicate reviews and Applying Part of speech tagging by assigning a tag to each token based on the context.

B. Spam Reviews Detection

Spam reviews detection is a main process that helps in giving correct feedback of various customers' reviews about a given service or a product and also helps in providing real and accurate sentiment scores. In this part, a content-based method is used to detect spam reviews. The basic idea behind this method is to detect spam reviews through analyzing reviews' content.

We examine a number of properties related to the reviews' text to identify spam reviews. Spam reviews are detected by considering a combination of some of these properties such as: the number of words in a review where a spam review contains more words than that in a truthful review, the number of pronouns in a review where a spam review has more pronouns while the text of a truthful review contains more nouns, prepositions, adjectives, determiners and coordinating conjunctions and the number of unique words (diversity) where a spam review is less diverse than a truthful review. Example of a spam review: "DO NOT STAY HERE!!! My wife and I went to this place. The whole experience was very awful from its start to its end". The pseudo code of spam reviews detection process is illustrated in Algorithm 1 from step 3 to step 8.

C. Features Extraction

Feature extraction process is applied by identifying nouns and noun phrases. The performance of sentiment analysis relies on the effectiveness of the feature extraction method used.

In this part, we examine three ways to extract features from data namely extracting all nouns, extracting only the nouns that occur frequently and extracting frequent nouns by applying Apriori algorithm [20]. In extracting all nouns, we consider all nouns to be features. In extracting only the nouns that occur frequently, we count the number of times that each noun occurs and only nouns with high frequency are likely to be features. In extracting frequent nouns by applying Apriori algorithm, we apply Apriori algorithm and then frequent item sets of nouns in reviews are expected to be features while the infrequent ones are less expected to be features.

D. Opinion Words Extraction

Opinion words, the words that are used to express opinions such as adjectives, adverbs and verbs, are extracted from a review. A lexicon based approach is used to assign sentiment scores to the extracted opinion words. Example of opinion words and their sentiment scores is shown in Table I.

TABLE I. EXAMPLE OF OPINION WORDS AND THEIR SENTIMENT

Terms	Sentiment Score
Great	0.6249
Beautiful	0.5994
Perfect	0.5719
Nice	0.4215
Bad	-0.5423
Angry	-0.5106
Terrible	-0.4767
Confuse	-0.2263

E. *Polarity Classification*

In this part, the sentiment score of the extracted features is obtained according to the sentiment score of the nearest extracted opinion word and then the accumulative sentiment score of each feature is calculated, and also the reviews are categorized as positive or negative.

The review is classified as positive, if the sum of sentiment scores of the extracted features in a review is greater than zero and the review is classified as negative, if the sum of sentiment scores of the extracted features in a review is smaller than zero. The pseudo code of polarity classification process is illustrated in Algorithm 1 from step 12 to step 18.

IV. EVALUATION, RESULTS AND DISCUSSIONS

In this section, we present and assess the experimental results of the proposed algorithm. We evaluate the proposed algorithm both in sentiment score calculation at the feature level per feature and sentiment score calculation at the feature level per review using a dataset composed of real world social reviews. The proposed work is evaluated by comparing the obtained results from sentiment analysis before and after applying spam reviews detection along with using three different feature extraction methods. The comparison is done using the three types of metrics mentioned in Evaluation Metrics section.

A. *Dataset*

The Dataset used to evaluate the proposed algorithm was found in [21] and it contains 1600 reviews from TripAdvisor[1] and Yelp[2]. These reviews are for Chicago-based hotels. The Dataset is divided into positive and negative reviews. Some of these positive/negative reviews are truthful reviews while others are deceptive reviews.

[1] TripAdvisor is a website that works particularly in the field of travel and tourism. It contains travel related information and content such as location information and their reviews. (http://tripadvisor.com)

[2] Yelp is a popular review website which is crowd-sourced and reviews local stores and brands. Here, users can also interact with one another just like in social networking sites. It is like TripAdvisor, users can give their opinion about a restaurant, hotel or a location that they are visiting. (http://www.yelp.com)

B. *Evaluation Metrics*

Different evaluation metrics are used to assess the effectiveness of the proposed algorithm such as Accuracy, Precision, and Recall.

Accuracy: It is defined as "The number of correct cases divided by the total number of cases".

Precision: It is defined as "The fraction of retrieved instances that are relevant or it is the percentage of selected items that are correct".

Recall: It is defined as "The fraction of relevant instances that are retrieved or it is the percentage of correct items that are selected".

The above mentioned metrics can be calculated by using the equations discussed below using the confusion matrix presented in Figure 1.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

Fig. 1 Confusion Matrix

True Positive (TP) – "the number of correct predictions that an instance is positive".

True Negative (TN) – "the number of correct predictions that an instance is negative".

False Positive (FP) – "the number of incorrect predictions that an instance is positive".

False Negative (FN) – "the number of incorrect predictions that an instance is negative".

C. *Experimental Results*

In this section, three experiments are conducted to evaluate the efficiency of the proposed algorithm. There are three objectives for these experiments. First, to investigate the effectiveness of the methods used to extract features from reviews. Second, to evaluate the effectiveness of the different spam detection methods. Third, to evaluate the influence of detecting spam reviews on determining the overall sentiment for each feature in social reviews.

1) *Feature Extraction*

The first experiment aims to compare the results obtained by sentiment analysis using different feature extraction methods namely extracting all nouns, extracting only the nouns that occur frequently and extracting frequent nouns by applying Apriori algorithm. The results obtained from the method “extracting frequent nouns by applying Apriori algorithm” shown in Table II give accuracy of 69.03% which is fairly better than the results obtained from the methods “extracting all nouns” and “extracting nouns that occur frequently” which obtain accuracy of 67.41% and 68.84% respectively. The results also show an improvement in precision value when using the method “extracting frequent nouns by applying Apriori algorithm” where it is 62.11%. The results obtained from extracting features by the methods “extracting all nouns” and “extracting nouns that occur frequently” give precision of 60.71% and 61.87% respectively. This increase in precision is due to the decrease in the number of false positives which means that there is an increase in the number of reviews that are correctly classified as negative reviews. On the other hand, the value of recall decreases slightly with the use of the method “extracting frequent nouns by applying Apriori algorithm” and this due to the increase in the number of false negatives.

TABLE II. PERFORMANCE OF SENTIMENT ANALYSIS USING DIFFERENT FEATURE EXTRACTION METHODS

	Accuracy	Precision	Recall
Extract All Nouns	67.41%	60.71%	98.51%
Extract Frequent Occurrence Nouns	68.84%	61.87%	98.01%
Extract Apriori Frequent Nouns	69.03%	62.11%	97.38%

2) Spam Reviews Detection

The second experiment aims to study the effect of three properties related to reviews’ text used for the detection of spam reviews namely calculating number of words in a review, calculating number of unique words in a review and calculating number of pronouns in a review. This experiment is divided into two parts. The first part of the experiment, presented in Figure 2 shows a comparison between using each of the three above mentioned properties, related to reviews’ text, individually and combining the three properties together for the detection of spam reviews.

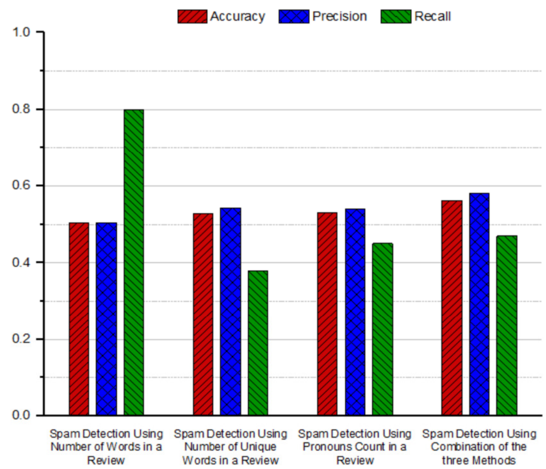

Fig. 2 Comparing different spam detection methods

In Table III, removing spam reviews with the three different properties related to reviews’ text which are calculating number of words in a review, calculating number of unique words in a review and calculating number of pronouns in a review gives accuracy of 50.37 %, 52.80%, and 53.05% respectively and the accuracy is increased to 56.28% on using a combination of the three properties. The results also show an improvement in precision value on using a combination of the three properties where it increased to 58.02% due to the decrease in the number of false positives. This decrease in the number of false positives shows that there is an increase in the number of reviews that are correctly classified as spam reviews. On the other hand, the number of false negatives increases and therefore the value of recall decreases.

TABLE III. COMPARING DIFFERENT SPAM DETECTION METHODS

	Accuracy	Precision	Recall
Spam Detection Using Number of Words in a Review	50.37%	50.39%	79.95%
Spam Detection Using Number of Unique Words in a Review	52.80%	54.35%	37.87%
Spam Detection Using Pronouns Count in a Review	53.05%	53.93%	45.05%
Spam Detection Using Combination of the three Methods	56.28%	58.02%	47.03%

From this part of the experiment we recognized that the accuracy obtained from detecting spam reviews using a combination of the three properties is fairly better than the accuracy obtained from detecting spam reviews using a single property.

The second part of this experiment which is presented in Figure 3 is based on the results obtained from the first part of this experiment. This part of this experiment compares the results obtained from sentiment analysis before and after removing spam reviews using a combination of the three properties related to reviews’ text. These results indicate that the detection of spam reviews allows getting better results for feature-based sentiment analysis. The comparison is also done using the three feature extraction methods mentioned before in the first experiment.

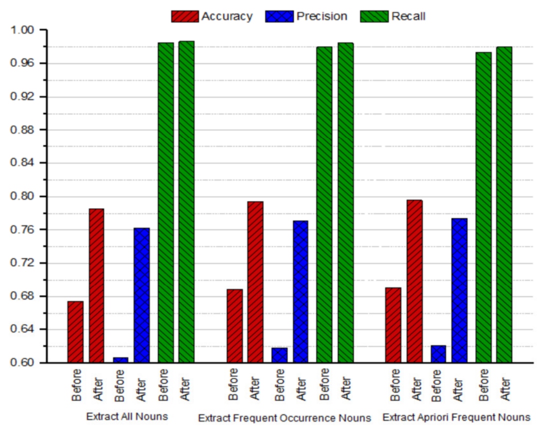

Fig. 3 Performance of sentiment analysis on using the three feature extraction methods before and after removing spam reviews

In Table IV, the accuracy of sentiment analysis on using the feature extraction method “extracting all nouns” before removing spam reviews is 67.41% while after removing spam reviews is 78.52% which means that it is improved by 16.48%. The results also show an improvement in precision value when using this feature extraction method after removing spam reviews where it increased to 76.20%.

TABLE IV. PERFORMANCE OF SENTIMENT ANALYSIS ON USING THE FEATURE EXTRACTION METHOD “EXTRACTING ALL NOUNS” BEFORE AND AFTER REMOVING SPAM REVIEWS

	Accuracy	Precision	Recall
Extract All Nouns Before Removing Spam Reviews	67.41%	60.71%	98.51%
Extract All Nouns After Removing Spam Reviews	78.52%	76.20%	98.67%

In Table V, the accuracy of sentiment analysis on using the feature extraction method “extracting nouns that occur frequently” before removing spam reviews is 68.84% while after removing spam reviews is 79.41% which means that it is improved by 15.35%. The results also show an improvement in precision value when using this feature extraction method after removing spam reviews where it increased to 77.08%.

TABLE V. PERFORMANCE OF SENTIMENT ANALYSIS ON USING THE FEATURE EXTRACTION METHOD “EXTRACTING NOUNS THAT OCCUR FREQUENTLY” BEFORE AND AFTER REMOVING SPAM REVIEWS

	Accuracy	Precision	Recall
Extract Frequent Occurrence Nouns Before Removing Spam Reviews	68.84%	61.87%	98.01%
Extract Frequent Occurrence Nouns After Removing Spam Reviews	79.41%	77.08%	98.45%

In Table VI, the accuracy of sentiment analysis on using the feature extraction method “extracting frequent nouns by applying Apriori algorithm” before removing spam reviews is 69.03% while after removing spam reviews is 79.56% which means that it is improved by 15.25%. The results also show an improvement in precision value when using this feature extraction method after removing spam reviews where it increased to 77.41%.

TABLE VI. PERFORMANCE OF SENTIMENT ANALYSIS ON USING THE FEATURE EXTRACTION METHOD “EXTRACTING FREQUENT NOUNS BY APPLYING APRIORI ALGORITHM” BEFORE AND AFTER REMOVING SPAM REVIEWS

	Accuracy	Precision	Recall
Extract Apriori Frequent Nouns Before Removing Spam Reviews	69.03%	62.11%	97.38%
Extract Apriori Frequent Nouns After Removing Spam Reviews	79.56%	77.41%	98.00%

As noticed on using the three different feature extraction methods after removing spam reviews, there is a remarkable increase in precision value but there is a very slight increase in recall value. This increase in precision means a lower false positive rate and the increase in recall means a lower false negative rate. The increase in both shows that the proposed algorithm is returning accurate results in classifying truthful reviews.

3) Sentiment Score Calculation

The aim of this experiment is to evaluate the influence of detecting spam reviews on determining the overall sentiment for each feature in social reviews.The feature sentiment scores shown in Table VII indicate that our algorithm is able to classify and assign accurate sentiment score to each feature in social reviews after removing spam reviews where the accumulated sentiment score for features is equal or nearly close to the sentiment score calculated from ground truth dataset.

TABLE VII. FEATURE SENTIMENT SCORE BEFORE AND AFTER REMOVING SPAM REVIEWS

	Odor	Steak	laundry	Room	Service	Staff	Food
Feature Score Before Removing Spam Reviews	-2.57	2.76	-0.29	177.88	51.26	136.25	35.06
Feature Score After Removing Spam Reviews (Proposed Algorithm)	0.03	0.64	0.2	86.53	30.63	77.24	16.88
Feature Score After Removing Spam Reviews (Ground Truth Dataset)	0.03	0.64	0.2	86.65	30.11	80.66	12.07

As shown from the above table, the sentiment score of the features: “Odor” , “Steak”and “Laundry” is equal to their sentiment score in ground truth dataset while the sentiment score of the features “Room”, “Service”, “Staff” and “Food” is close to their sentiment score in ground truth dataset.

V. CONCLUSION

The process of spam reviews detection in sentiment analysis and opinion mining is a major research issue. In this paper, we proposed an enhanced feature-based algorithm which detects spam reviews based on the reviews’ text properties. We presented some experiments aiming to endorse the proposed algorithm for enhancing the accuracy of feature level sentiment analysis.

As an Overall, the results lead us to conclude that applying the feature extraction method “extracting frequent nouns by applying Apriori algorithm” gives better accuracy based on the investigation applied in the first experiment in this paper to compare the effectiveness of the different methods used to extract features from reviews.

It is also concluded that applying a combination of three properties related to the reviews’ text to detect spam reviews gives better accuracy than using a single property based on the investigation applied in the first part of the third experiment in this paper.

It is also concluded that detecting spam reviews using a combination of three properties related to the reviews’ text after applying the feature extraction method “extracting frequent nouns by applying Apriori algorithm” gives better accuracy based on the investigation applied in the second part of the second experiment in this paper.

In addition, we evaluate the effectiveness of the proposed algorithm in classifying and assigning an accurate sentiment score to each feature in social reviews after removing spam reviews. Finally, the accumulated sentiment score for features is equal or nearly close to the sentiment score calculated previously from ground truth dataset. In terms of future work, we suggest to consider other challenges that face opinion mining and sentiment analysis process to enhance the proposed feature-based sentiment analysis algorithm.

REFERENCES

[1] H. P. Patil and M. Atique, "Sentiment Analysis for Social Media: A Survey," 2015 2nd International Conference on Information Science and Security (ICISS), Seoul, Korea (South), 2015, pp. 1-4.

[2] H. R. P, “Opinion mining and sentiment analysis - challenges and applications,” International Journal of Application or Innovation in Engineering and Management (IJAIEM), vol. 3, 2014.

[3] Saqib Iqbal, Ali Zulqurnain, Yaqoob Wani and Khalid Hussain , “The survey of sentiment and opinion mining for behavior analysis of social media”, International Journal of Computer Science & Engineering Survey (IJCSES) Vol.6, No.5, October 2015.

[4] Richa Sharma, Shweta Nigam and Rekha Jai , “Mining Of Product Reviews At Aspect Level”, IJFCST, Vol.4, No.3, pp 87-95, May 2014

[5] M. del Pilar Salas-Zárate et al., "Sentiment analysis on tweets about diabetes: An aspect-level approach", Comput. Math. Methods Med., vol. 2017, Feb. 2017.

[6] D.M.E.D.M Hussein, "A survey on sentiment analysis challenges" in Journal of king saud university-engineering science, Press, 2016.

[7] F. Li, M. Huang, Y. Yang, X. Zhu. “Learning to identify review Spam” In IJCAI, Vol. 22, p. 2488, 2011.

[8] S. Shojaee, M. A. A. Murad, A. B. Azman, N. M. Sharef, and S. Nadali, “A framework for fake review annotation” In Proceedings of the 2015 17th uksim-amss international conference on modelling and simulation , pp. 153–158, 2015.

[9] I R Jayasekara , W M J I Wijayanayake, "Opinion mining of customer reviews: Feature and Simile based approach", International journal of Data mining & Knowledge Management process (IJKDP), Vol.6, No.1, pp.1-11, January 2016.

[10] Qingxi Peng, Ming Zhong, "Detecting Spam Review through Sentiment Analysis", Journal of Software, vol. 9, no. 8, 2014.

[11] Sharma Kuldeep, King-Ip Lin, "Review spam detector with rating consistency check", Proceedings of the 51st ACM Southeast Conference on - ACMSE 13, 2013, p.34.

[12] M. Tavakoli, A. Heydari, Z. Ismail, N. Salim, "A Framework for Review Spam Detection Research", World Academy of Science Engineering and Technology International Journal of Computer Electrical Automation Control and Information Engineering, vol. 10, no. 1, pp. 67-71, 2015.

[13] K. H. Yoo, and U. Gretzel, "Comparison of deceptive and truthful travel reviews" ICT in Tourism, pp. 37-47. 2009.

[14] S. Ahmad, A. Pathak, S. Jaiswal, “A survey about spam detection and analysis using users’ reviews”, Malaya Journal of Matematik (MJM), pp. 1-4. 2018.

[15] Neha S Chowdhary and Anala A Pandit. “Fake Review Detection using Classification”, International Journal of Computer Applications vol. 180, no.50, pp. 16-21, June 2018.

[16] M. Singh, L. Kumar, S. Sinha, “Model for detecting fake or spam reviews” In Ict based innovations, Springer, 2018, pp. 213–217.

[17] B. Verma, RS. Thakur, “Sentiment analysis using lexicon and machine learning-based approaches: A survey”, In Proceedings of international conference on recent advancement on computer and communication, 2018, pp. 441–447.

[18] A. Shaikh, M. Rao, “Survey on sentiment analysis”, presented at International Conference On Emanations in Modern Technology and Engineering (ICEMTE- 2017), 5(3), 69–73.

[19] Charu C. Aggarwal, “Opinion Mining and Sentiment Analysis”. In: Machine Learning for Text. Springer, Cham, 2018, pp 413-434

[20] R. Agrawal, R. Srikant, “Fast algorithms for mining generalized association rules”, in: Proceedings of the 20th International Conference on Very Large Databases (VLDB94), Santiago, Chile, 1994, pp. 487–499.

[21] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination", in HLT '11 Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 309-319, Jun. 2011.