_____

# DETECTION OF FAKE REVIEWS ON SOCIAL MEDIA USING MACHINE LEARNING ALGORITHMS

*Huy Le, Seattle University, lehuy@seattleu.edu*
*Ben Kim, Seattle University, bkim@seattleu.edu*

## ABSTRACT

*With the development of the Internet, technology and e-commerce, online-purchasing is easier and more convenient these days. Online reviews become the main source of information that customers usually refer to for their purchasing decision. However, many of reviews given by the online users are not considered truthful. Because of commercial benefits, fake reviews were generated to mislead customers. Therefore, it is necessary to detect fake reviews effectively. This paper aims to improve the performance of fake review classifiers by integrating different techniques into classifying models. More specifically, we analyzed similarity between reviews and utilized the EM (Expectation Maximization) clustering algorithm to recognize the review patterns. We also applied the sentiment analysis to analyze the reviews. Using the results from clustering models, sentiment analysis, and non-textual features of reviews and reviewers, we built machine learning models to classify fake reviews. We compare three supervised machine learning algorithms: Support Vector Machine, Artificial Neural Network and Random Forest. The empirical results from our experiments showed that the Random Forest algorithm outperforms against other algorithms. It also proved our assumption about text clustering and non-textual features in fake review detections.*

**Keywords:** Machine Learning, Text Mining, Fake Reviews, Random Forest, EM Clustering, Text Clustering, Text Classification

## INTRODUCTION

In the era of the Internet and e-commerce, when online businesses are becoming continuously developed and dominant, writing online reviews of products is now a common practice for consumers. This is one of the most convenient ways for consumers to express their opinion about the services or products they purchased. The reviews have become valuable sources of information for potential customers by helping them increase their insights into the products or services that they are going to purchase. These user-generated contents are also useful sources for the online-business entities. Merchants can use this information to improve their products, services, marketing strategies or analyzing their competitors.

A new issue has arisen when businesses or reviewers create fake reviews for spreading deceptive information. These counterfeit contents can be used to promote or demote specific businesses/products. This activity is known as fake reviews, review spams or opinion spams. The main problem of review spams is that reviewers can easily create a hype for products or services by writing positive reviews in bulk. These spam reviews now play as key factors that can easily sway customers' perceptions. Positive reviews can bring significant financial benefits or fame for organizations while negative reviews can dramatically ruin their reputation**.** Reviews can be generated by an automated system or paid reviewers. Companies and merchants can hire individuals or third-party organizations to write fake positive reviews for their products or services. Furthermore, the trend of spamming fake reviews on e-commerce websites has increased since everyone can easily write and post a review on the internet. Taylor (2019, April) has reported that Amazon was flooded with fake five-star reviews. Liu, a data mining expert at the University of Illinois, Chicago estimated that one-third of the reviews on the Internet are fake reviews (Streitfeld, 2012). Fake reviews are becoming more sophisticated as reviewers tried to mimic genuine reviews or work in groups. Thus, it has become more difficult for customers to retrieve helpful information without being deceived by those fake reviews.

Because of these concerns, the fake review problem has gained a higher level of interest from both academics and industry. It is also drawing attention from legal regulations. To counter this issue, scientists have done a great deal of research on opinion spams. Commercial hosting sites, such as yelp.com and amazon.com have also integrated their classifiers to prevent deceptive reviews. However, as the problem is becoming complicated, we need to continue improving the techniques for fake review detection.

_____

**Definition of Fake Reviews**

What is a fake review? According to Jindal & Liu (2007, May), a fake review is considered a type of opinion spamming. This is referred to as an illegal activity that reviewers try to mislead readers or automated opinion mining and sentiment analysis systems rather than express their genuine opinions or experience. These reviewers post either positive reviews to a product, service or business entities to promote them or negative opinions about some other entities in order to destroy their reputations. In other words, a fake review is one that deliberately delivers wholly or partially untruthful information, opinion, experience or irrelevant information about the review objects. Fake reviews are also called bogus reviews, shilling reviews, deceptive reviews, or spam reviews.

The dataset we used in this research was obtained from Mukherjee, Venkataraman, Liu, and Glance (2013, June). This is the Yelp review dataset, which already has fake and non-fake labels based on Yelp's fake review filter. Mukherjee and others performed an investigation of the nature of fake reviews in Yelp.com. Their experiment demonstrated that Yelp's filtering system is reliable. In other words, we can confidently use this labeled data set for training classifying models for fake reviews. We do not know how Yelp's filter system classifies fake reviews because it is their trade secret. Mukherjee and his team had figured out that there was not much difference in words used in fake (filtered) and non-fake (unfiltered) reviews in Yelp dataset. Hence, this is much harder to recognize a fake review by using linguistic features or human intuition. Mukherjee et al. (2013, June) had strong evidence to believe that Yelp had used behavioral data of reviewers and other internal data to support their filter system.

Reviewing existing works showed that most of them mainly focused on supervised machine learning approaches. They either focused on analyzing review contents or the reviewer's behavioral attributes. Both methods have shown the pros and cons in detecting fake reviews. There is no single representation method that can adapt to every part of the problem.

Our research seeks to improve fake review classifiers by examining the effectiveness of a Part of Speech (PoS) in review representations and integrating the clustering model into detecting the fake reviews. Our experiments verify the potential ability of PoS tagging in detecting fake reviews by comparing performance of model using PoS and N-grams. In addition to contributing to fake review classifications, we demonstrate the ability of text clustering technique in finding the hidden structures of text data. We believe clustering algorithms such as k-means and Expectation-Maximization (EM) can read and organize those hidden structures and their outcomes to support our classification models. Finally, our analysis supports the idea of incorporating unsupervised and supervised learning into fake review detection. We built supervised learning models with Support Vector Machine (SVM), Neural Network (NN) and Random Forest (RF) using behavioral features and clustering results as inputs.

We believe that this study will make contributions to the e-commerce industry in training and developing a fake review detection system. It will be beneficial for both consumers and business owners when fake review classifiers are more accurate. Business owners can protect their business by detecting fake negative reviews and consumers can retrieve helpful information without worrying about misleading information. This research will also serve as a reference for future research on the subject of detection of fake reviews and text classification.

## RESEARCH METHODOLOGY

**Detection Techniques for Fake Reviews**

As we already mentioned above, the goal of fake review detection is to develop a method that can fully incorporate all valid information regarding the reviews and reviewers to accurately identify fake reviews. Our research was an attempt to solve the fake review problems based on the machine learning approach. The classification models take the reviews and reviewer's attributes as input values and returned a label for each review indicating whether the review is a truthful or untruthful one as an output. The following sections explain each step in our experiment process: feature extraction, review clustering, and classifying fake reviews. Figure 1 describes the process of our experiment.
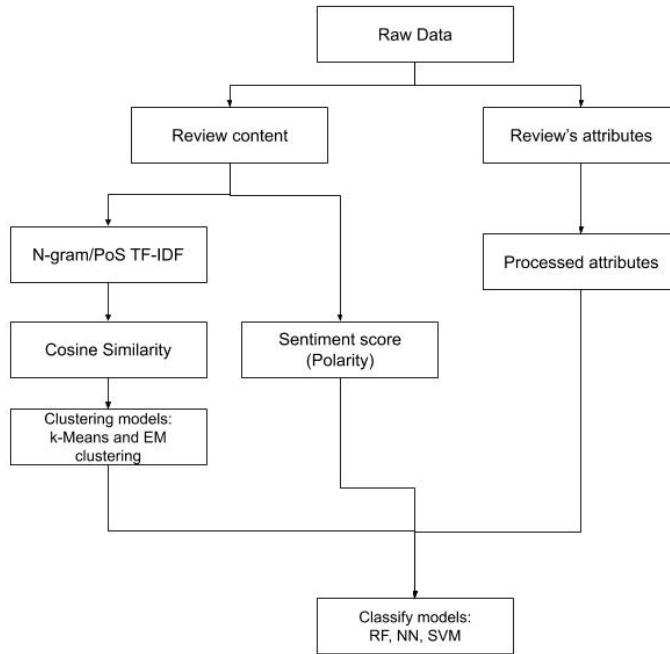
_____



**Figure 1.** The Proposed Model Workflow

**Feature Extraction**

In text-mining, the textual content is one of the essential characteristics of a document. In this problem, it is the review content that expresses the experience or opinion of a reviewer regarding a product or service. To use the textual content as the inputs for the machine learning model, the textual content needs to be transformed to machine-readable values. Previous studies used the N-gram based features on one or multiple word levels and they yielded a satisfactory result with a high accuracy (Mukherjee et al, 2013; Ott et al, 2011, June). We became curious about whether Part of Speech can be used as N-gram alternatives. Usually, Part of Speech of a document is generated in its form as arrays where each tuple represents a word and its tag (see Figure 2). Thus, we need to convert the tuples in PoS arrays into a single string before we compute the Term Frequency-Inverse Document Frequency (TF-IDF). The tuple in PoS arrays have a form as "word tag" (Figure 3). We then generate a TF-IDF matrix based on the PoS matrix. Simultaneously, the TF-IDF matrix based on the N-gram features is still generated. We are intent to compare the impacts of N-gram and PoS tags in fake review classifications.

_____

```
pos = word_tok.apply(pos_tag)
print(pos[:10])

0    [(great, JJ), (food, NN), (great, JJ), (locat,...
1    [(beef, NN), (flemmington, NN), (drool, NN), (...
2    [(went, VBD), (drink, NN), (dessert, NN), (hum...
3    [(went, VBD), (birthday, JJ), (dinner, NN), (f...
4    [(pizza, RB), (good, JJ), (fri, NN), (quick, J...
5    [(know, NNS), (could, MD), (pass, VB), (place,...
6    [(place, NN), (aw, NN), (overpr, IN), (thank, ...
7    [(ador, JJ), (motel, NN), (reason, NN), (other...
8    [(look, NN), (juici, NN), (juici, NN), (luci, ...
9    [(eaten, RB), (two, CD), (visit, NN), (minneap...
Name: reviewcontent, dtype: object
```

**Figure 2.** Part of Speech Tagging

```
print(pos[:10])

0    place_NN amaz_NN partner_NN came_VBD first_JJ ...
1    visit_NN number_NN number_NN yum_JJ visit_NN h...
2    great_JJ atmospher_RB great_JJ servic_NN aroun...
3    ok_JJ cant_JJ recal_JJ mani_NN time_NN driven_...
4    hotel_NN staff_NN incred_VBN help_NN room_NN p...
5    place_NN definit_NN fresh_JJ fresh_JJ fresh_JJ...
6    boyfriend_NN treat_NN mini_JJ weekend_NN getaw...
7    number_NN year_NN patron_RB origin_JJ uncommon...
8    prepar_NN treat_NN poorli_NN promis_IN number_...
9    enjoy_VB anoth_DT fabul_NN dinner_NN last_JJ n...
Name: reviewcontent, dtype: object
```

**Figure 3.** Processed Part of Speech Tags

The next step is computing the similarity scores and sentiment scores. Probably one of the most useful techniques to recognize spamming activity in online reviews is examining the duplicates of reviews (Jindal & Liu, 2007 October). For example, if we see many reviews in one or many products that are similar, there is a high probability that they were written by one person although their user-names are different, and they are likely to be spam reviews.

One of the most common metrics used to measure how similar the documents are is Cosine Similarity (CS). CS measures the cosine of the angle between two vectors projected in a multi-dimensional space. The reason we chose CS to measure the similarity between reviews is that CS can measure the document similarity regardless of document's size. It is more advantageous than the distance-based method. The higher cosine we get, the smaller angle between the two vectors is the more similar between two documents and vice versa. We applied the cosine-similarity function from the scikit-learn library.

We also included sentiment ratios into our feature sets. The sentiment ratio of a review was calculated based on the Textblob library (https://textblob.readthedocs.io/). Textblob is a Python library that offers a simple API for performing NLP tasks. Textblob provides two metrics for sentiment analysis: Polarity and Subjectivity. Polarity simply means emotions expressed in the reviews. The Polarity ratio obtains the float value in the range of [-1.0,1.0] with -1 is extremely negative, 1 is extremely positive, and 0 is neutral. Subjectivity is a subjective ration of reviews; it presents either a review is subjective or objective. The subjectivity is a float number within the range [0.0, 1.0] where 0.0 is very objective and1.0 is very subjective. The sentiment properties were generated by taking the processed review contents as inputs and returns the sentiment score. By default, "Textblob.sentiments" module implements an analysis by applying Pattern Analyzer based on the pattern library (https://www.clips.uantwerpen.be/pattern). We can override the analyzer by Naïve Bayes Analyzer, which is from the Natural Language toolkit (NLTK) library (Bird, Klein, & Loper, 2009). In this study, we used Pattern Analyzer.

_____

**Behavioral Features**

Nonverbal behavioral features were selected based on our assumptions about their possible influence on fake review classifications and the findings from existing works (Mukherjee et al., 2013; Zhang et al., 2016). Most of the behavioral features already appeared in the dataset. Those others were computed based on some criteria. The detail non-verbal feature sets and their description are presented in Table 1.

**Table 1**. Behavioral Features

| Feature name | Description |
|---|---|
| monmembership | A number of month membership of reviewer at the time given the review. |
| reviewrating | Posted rating of this review |
| reusefulcount | Number of useful votes from other users for this review |
| recoolcount | Number of cool votes from other users for this review |
| reviewDate | Number of funny votes from other users for this review |
| friendcount | Number of friends of a reviewer |
| fancount | Number of fans of a reviewer |
| tipcount | Number of tips of a reviewer |
| reviewcount | Number of written reviews of a reviewer |
| firstcount | Number of times being the first review of a reviewer |
| usefulcount | Number of useful votes for this reviewer |
| coolcount | Number of cool votes for this reviewer |
| complimentcount | Number of compliments for this reviewer |
| funnycount | Number of funny votes for this reviewer |
| busrating | The rating of this business |
| pricerange | Level of the Price range, from 1 to 4 |
| firstreview | 1: this is the first review, 0: this is not the first review on this business page. |
| maxReviewDay | Maximum number of reviews were written within a day of a reviewer |
| avgReviewDay | The average number of reviews those observations within a day of reviewer |

_____

| Feature name | Description |
|---|---|
| avgpostedrating | Average posted ratting of a reviewer |
| avgreviewlen | Length of review content, count by character |

## Clustering Methods

In this section, we used clustering as a data preprocessing step. Cluster labels that are generated from the clustering algorithms are considered as independent nominal data. After that, cluster labels are integrated into dataset for training classifying models.

The purpose of this step is to reveal the hidden structure of fake and non-fake reviews, which would support our review classification models. We used popular clustering methods such as Gaussian EM Clustering. To build a cluster model, the clustering input is the cosine similarity matrix. This matrix was generated by applying cosine similarity on both Unigram and Unigram-PoS (one work) to compare the effects of these text features (see Table 2).

## Classification Methods

The final section of this experiment is to build models for classifying fake reviews. We selected three classification algorithms, including Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest (RF).

SVM in machine learning is a classification method for both linear and non-linear data. The operation of the SVM algorithm is based on finding the hyperplane that segregates multi-dimensional data into classes. SVM is one of the most commonly used classification algorithms for fake review detection (Mukherjee et al., 2013; Mukherjee et at, 2013, June; Zhang et al, 2016).

Artificial Neural Network (ANN) consists of a input layer, one or more hidden layers, and one output layer. ANN is generally applied in computer vision; however, they are recently applied to various text mining problems, especially text classifications (Luo et al, 2017 July).

Random Forest is an ensemble classifier that operates as a combination of multiple decision trees. Each tree in the forest is generated using a random selection of attributes at each node to determine the split. Random forests operate on a set of randomly selected features. Thus, high dimensionality of data can be less of a problem with RF. Although in our research we do not build the RF model with high dimensional text features, we still apply RF because of its outstanding performance in the previous studies.

## DATA MINING MODELS AND EVALUATIONS

### Dataset Description

All our experiments were implemented on datasets of 10,000 reviews that are randomly selected from original datasets. The training and testing data are divided by an 80:20 ratio in which we have 50% for both fake and non-fake reviews. The original dataset includes yelp review data from 2004 to 2012. Because the size of the dataset is too big and too far from now, we only chose those observations from 2010 to 2012. We also limited the review dataset in two business categories: restaurant and hotel. The other thing noticed by the author who crawled the datasets was that the label column named 'flagged' had 4 categories 'Y', 'YR', 'N', 'NR'. Y/N reviews were obtained from the business page, YR/NR reviews were obtained from the reviewer profile page. Y means the review was filtered by Yelp's filtering

_____

system or fake review and N means non-fake review. The author only used reviews with labels Y and N. Therefore, to make our results are comparable and avoids duplication in dataset we only user Y and N labels.

**Experimental Setup**

As has been described in the previous section, two different sets of experiments have been conducted. We used the result from two different clustering models as independent features for our classifiers. In the first one, we used the result from the clustering model, which was trained by cosine similarity based on PoS while the second one trained by cosine similarity based on Unigram. The reason for these setups is to determine the effect of PoS and clustering on text classification in latter step. We also conducted the experiments with various settings, which are full sample datasets with and without cluster labels, and within each cluster.

The experiments showed that clustering by using Unigram based and Unigram PoS-based slightly improved the classifier performance. It is important to notice that increasing the N-gram in the TF-IDF generating step did not help in clustering. When we increased n by more than 1 a major part of datasets belong to one cluster since most values in the cosine similarity matrix are close to 1. Due to the limitation in computing power, we do not present the clusters' characteristics in this research.

**Empirical Results**

We used the confusion matrix and related measurements to evaluate and compare the performance of each model on a standardized level. The table below presents the accuracy, recall, precision and F1 score from each model in different settings (Table 2).

At the first glance, it appears that there is a small margin of differences between using PoS clusters and the Unigram cluster. Between all three models, it appears that Random Forest provided the highest accuracy and recall with 92.55% and 95.27%, respectively against SVM and Neural Network. Furthermore, we observe that cluster 1 in PoS-based cluster and cluster 3 in the Unigram-based cluster produce the highest performance when compared to the result from different settings with the same algorithms. Another interesting point is that in the results from Random Forest, the accuracy and other measures slightly decrease when we remove the cluster labels from the independent feature sets.

To verify the effect of clustering and other features in classification, we conducted a further investigation on finding the most important features. We applied three different methods to find out the most important features: logistic regression with stepwise selection, random forest selection, and decision tree. As can be seen in Figure 4, our results showed that behavior-related features play more important roles in fake review classifications than text-related features because most features that yielded from selection method are behavior features. Only polarity appeared in the stepwise selection, there are no textual features in a random forest selection, and finally cluster labels appeared in level 6 of the decision tree. The graph below presents the top 10 important features in RF models (Figure 4). The 'usefulcount' is the most important feature with 0.16 while the two textual features, 'subjective' and 'polarity', appear in 10th and 11th respectively. These importance scores measure the ability reducing the information impurity of features in the decision tree as measured in calculating Gini-indices. The total value of the important scores of all features in the tree is equal to 1.

_____

**Table 2**. Results of classifying model integrated with clustering based on Unigram

| Clustering | Model | Clusters | Accuracy | Error | Recall | Precision | F1 |
|---|---|---|---|---|---|---|---|
| Uni-gram similarity | SVM | Full clusters | 87.30% | 12.70% | 91.84% | 84.06% | 87.78% |
| | | 0 | 85.91% | 14.09% | 92.75% | 82.40% | 87.27% |
| | | 1 | 84.60% | 15.40% | 92.31% | 82.56% | 87.16% |
| | | 2 | 88.84% | 11.16% | 88.89% | 85.11% | 86.96% |
| | | 3 | 88.25% | 11.75% | 90.57% | 84.71% | 87.54% |
| | | Without Cluster label | 87.40% | 12.60% | 91.84% | 84.21% | 87.86% |
| | Random Forest | Full clusters | 92.25% | 7.75% | 94.86% | 90.06% | 92.40% |
| | | 0 | 90.82% | 9.18% | 94.44% | 88.66% | 91.46% |
| | | 1 | 90.36% | 9.64% | 94.78% | 88.92% | 91.76% |
| | | 2 | 90.23% | 9.77% | 87.78% | 88.76% | 88.27% |
| | | 3 | 93.12% | 6.88% | 95.60% | 89.94% | 92.68% |
| | | Without Cluster label | 92.40% | 7.60% | 94.96% | 90.24% | 92.54% |
| | Neural Network | Full clusters | 88.70% | 11.30% | 95.87% | 83.73% | 89.39% |
| | | 0 | 87.42% | 12.58% | 93.24% | 84.28% | 88.53% |
| | | 1 | 87.87% | 12.13% | 93.96% | 85.93% | 89.76% |
| | | 2 | 84.65% | 15.35% | 75.56% | 86.08% | 80.47% |
| | | 3 | 86.53% | 13.47% | 88.68% | 82.94% | 85.71% |
| | | Without Cluster label | 87.90% | 12.10% | 94.16% | 83.56% | 88.54% |
| | Decision Tree | Full clusters | 88.30% | 11.70% | 88.62% | 87.91% | 88.26% |
| | | 0 | 85.79% | 14.21% | 87.20% | 85.75% | 86.47% |
| | | 1 | 86.00% | 14.00% | 86.26% | 88.70% | 87.47% |
| | | 2 | 88.37% | 11.63% | 80.00% | 91.14% | 85.21% |
| | | 3 | 85.10% | 14.90% | 81.76% | 84.97% | 83.33% |
| | | Without Cluster label | 88.40% | 11.60% | 88.32% | 88.32% | 88.32% |

_____
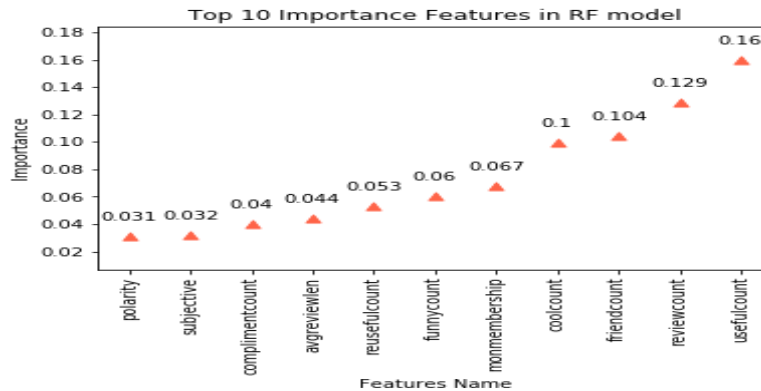


**Figure 4**. Important Features selected by Random Forest

## DISCUSSION

From the results in the previous section, the Random Forest model gave us the highest accurate results in both three different settings. Our findings concern the role of clustering in shaping the fake review classifier. The effect of clustering is not significant in this research. We found that by increasing the number of clusters we can increase the performance of classifying models. Our experiment shows that when we clustered with k = 8, the accuracy of RF can reach 94%. However, because of the difficulty in visualizing the characteristics of the clusters, we used 4 clusters in this experiment. In addition, if we consider separate classification tasks within each separate cluster, we can see that the total number of true positives is higher than when classifying the whole datasets.

These experimental results also prove that utilizing behavioral features is more effective than textual features in fake review classifying problems. The top five important features are i. useful count, ii. review count, iii. friend count, iv. cool count, and v. length of membership. This finding indicates that the credibility of reviewers is an effective factor in evaluating the trustworthiness of a review. In other words, we believe that instead of focusing on analyzing the reviewer's writing styles and word choices, we can develop a framework for analyzing the reviewer's behavior and credibility to improve the performance of fake review detection system.

Although our research produces a satisfying result, which supports our assumption of the effect of clustering in text classifications, many constraints are identified. Further research needs to be conducted to obtain a better solution in fake review problems. The most critical limitation is computing power. As we mentioned above that classification models can increase their performance by increasing the number of clusters. Initially, we thought that clustering would significantly improve the classifier accuracy even with a small number of clusters. However, because of high dimensionality of cosine similarity matrix, the maximum number of clusters that we can perform is 8 and it took several hours to produce a result. This problem also limited our capability in analyzing cluster characteristics, optimizing clustering, training models with larger datasets and applying other algorithms for clustering such as deep neural networks. This also calls into another question for us that whether there is no difference in textual structure among fake and non-fake reviews and for that reason textual features were not as significant in our research.

This study is the first step towards enhancing our understanding of utilizing the clustering method as preprocessing steps in text classification problems. We hope that our research will serve as a base for future studies, which will investigate more on clustering text data and developing a framework for evaluating the reviewer's credibility and online behavior. Further studies on this topic should concentrate on applying deep neural networks in clustering text data and combining verbal and non-verbal data in classification. It is also important to merge review data from different websites in training data sets. Furthermore, we may use different sentiment analysis algorithms for building polarity and subjective scores. Finally, one thing we would like to mention is a possible approach for future works. We have observed the different attributes and preprocessing techniques between text and non-text data. There exist a great deal of research addressing the problem by combining these types of data together. We believe that the problem can be solved by separating these data into two parts, then we apply suitable machine learning techniques.

_____

Consequently, we incorporate these models together as ensemble learning methods to obtain better predictive performance.

## CONCLUSION

The main idea of our research problem is to recognize the hidden patterns of fake reviews by using a clustering model based on cosine similarity among the reviews. We would like to emphasize that the objective is not using unsupervised learning to address the text classification problems but incorporating the result from clustering into the set of predictor attributes as an input to build the fake review classifiers. Our research underlined the importance of integrating the clustering step into data preprocessing. Although it was not significant, clustering can improve text classifying performance. By conducting separate reviews for each cluster, machine learning models can perform better. Non-text features are truly significant in solving fake review problems. In this study, what we are concerned about is the trustworthiness of the reviews, thus we need a metric that can evaluate the credibility of reviewers. Yelp has done an excellent job in evaluating the reviewers by allowing their customers to assess the reviews and reviewers. The length of a reviewer's membership at the time of the reviews also demonstrated a significant impact on classifying fake reviews. We explained that fake reviewers usually create new accounts for their activities. Hence, we believe the future research needs to include the features to be obtained by tracking reviewers' activities and utilizes those features to measure the reviewers' credibility.

## REFERENCES

Jindal, N., & Liu, B. (2007, May). Review spam detection. *Proceedings of the 16th international conference on World Wide Web,* 1189-1190.

Jindal, N., & Liu, B. (2007, October). Analyzing and detecting review spam. *Seventh IEEE International Conference on Data Mining (ICDM 2007),* 547-552. IEEE.

Luo, N., Deng, H., Zhao, L., Liu, Y., Wang, X., & Tan, Z. (2017, July). Multi-aspect Feature based Neural Network Model in Detecting Fake Reviews. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE),* 475-479. IEEE.

Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013, June). What yelp fake review filter might be doing? *Seventh international AAAI conference on weblogs and social media.*

Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). Fake review detection: Classification and analysis of real and pseudo reviews. *Technical Report UIC-CS-2013–03*, University of Illinois at Chicago, Tech. Rep.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011, June). Finding deceptive opinion spam by any stretch of the imagination. In Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1 (pp. 309-319). Association for Computational Linguistics.

Streitfeld, D. (2012) The best book reviews money can buy. *New York Times* 25. 2012. Retrieved from https://www.nytimes.com/2012/08/26/business/book-reviewers-for-hire-meet-a-demand-for-online-raves.html

Taylor, C. (2019, April 16). Amazon flooded with thousands of fake reviews, report claims. Retrieved from https://www.cnbc.com/2019/04/16/amazon-flooded-with-thousands-of-fake-reviews-report-claims.html

Zhang, D., Zhou, L., Kehoe, J., & Kilic, I. (2016). What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews. Journal of Management *Information Systems*, 33(2), 456-481.