DETECTION OF FAKE REVIEWS USING SENTIMENT ANALYSIS

Anantharam R U, Kirthika Gurumurthy,Ashwini Joshi
Department of Computer Science, PES University
Bangalore, India
Email: ananth99ru@gmail.com, kirthikagurumurthy@gmail.com, ashwinimjoshi@pesu.pes.edu

Abstract— **In recent times, online reviews have become become a primary factor in a consumer's decision to purchase a service or a product and are an important resource of customer opinion. Due to this heavy reliance on online reviews, there is a growing concern that some people may write false or artificial reviews to either promote their products or services or negative reviews to affect their competition. This practice of spammers manipulating and poisoning reviews (fake, untruthful or deceptive reviews) for their own gain is known as Opinion Spam. This gives rise to the need to develop techniques that can identify such deceptive reviews. Sentiment analysis has become of the most interesting topics for classification problems such as detection of fake reviews. In this paper we have employed sentiment analysis techniques, along with Machine learning models (supervised and unsupervised), lexicon based methods and hybrids of the aforementioned techniques to classify reviews as "deceptive" or "truthful".**

Key words: sentiment analysis, SVM, Naive Bayes, Decision trees, Random Forest, k-means, lexicon based, supervised, unsupervised.

I. INTRODUCTION

Online consumer reviews have become of staggering importance in any business. Research shows that 90% of customers read online reviews before visiting a business or purchasing a product. 88% of consumers trust online reviews as much as personal recommendation and 92% of users will use a local business if it has a rating of 4 stars or above. Today's consumers do research before committing to a business for a good or service. When they can't figure out on their own what to do, they take their cues from others. This is known as social proof, a term first coined by Robert Cialdini in his 1984 book, Influence: The Psychology of Persuasion. Due to this dependency there is a surge of untruthful or deceptive reviews by wrongdoers who use these tactics for their own profit or gain. Consumer reviews can be hugely influential, so it's hardly surprising there's a thriving trade in fake ones. Estimates of their prevalence vary – from 16% of all reviews on Yelp, to 33% of all TripAdvisor reviews, to more than half in certain categories on Amazon. These fake reviews are reportedly affecting the ability of almost all online shoppers to make decisions confidently regarding their purchases as ratings and reviews are their primary metrics of evaluation which creates the growing need to be able to identify such deceptive reviews.

Not all fake reviews are positive. In another study, researchers found that businesses often leave fake negative reviews to undermine their competitors and usurp potential customers. This is a common practice for restaurants and hotels who attempt to make their properties look better by ruining the star-ratings of their competitors. For example, let's say there is a major chain hotel surrounded by boutique hotels. The smaller hotels may write negative reviews about the chain in order to entice travelers to consider their properties instead. Brands and companies who have long been partaking in these activities should first begin by weeding out all fake and fraudulent content from their e-commerce websites as consumers have a right to be able to trust the reviews and ratings they encounter in the product pages.

In this study, we analyse online hotel using Sentiment analysis methods in order to detect fake reviews. One of the biggest challenges faced in the study of reviews to detect deception is obtaining labelled data. We have used the gold standard deceptive opinion spam dataset which is described in section 3. Since the dataset only contains review features and nothing about the reviewer the techniques we have employed in this study are review centric. Using both review and reviewer based methods may have a better performance, however as we could not find a labelled dataset for the same we have chose to perform the sentiment analysis on the reviews themselves. We have employed 3 main techniques to classify the reviews as "truthful" or "deceptive":

1) Supervised learning techniques -
    - Naive Bayes
    - SVM
    - Decision Trees
    - Long short term memory neural networks (LSTM)
    - Artificial Neural Networks (ANN)
2) Unsupervised learning techniques
    - k -means
3) Lexicon based approach
    - Dictionary based method

For the above methods we have identified certain features including lexical and sentiment based features which worked well in classifying these reviews (the features are described in the later section). We have also used lexicon scores such as afinn and vader for determining the degree of polarity of the sentences and as a part of the input features. For each method we have chosen the features that provide the maximum accuracy and have demonstrated the same.

Machine learning techniques and sentiment analysis methods are expected to have a major positive effect especially for the detection processes of fake reviews in hotel reviews, movie reviews, e-commerce and other domains. The main goal of our study is to classify hotel reviews as real reviews or fake reviews using SA algorithms with supervised, unsupervised and lexicon based techniques. The conducted experiments have shown the accuracy, precision and recall of results through sentiment classification algorithms.

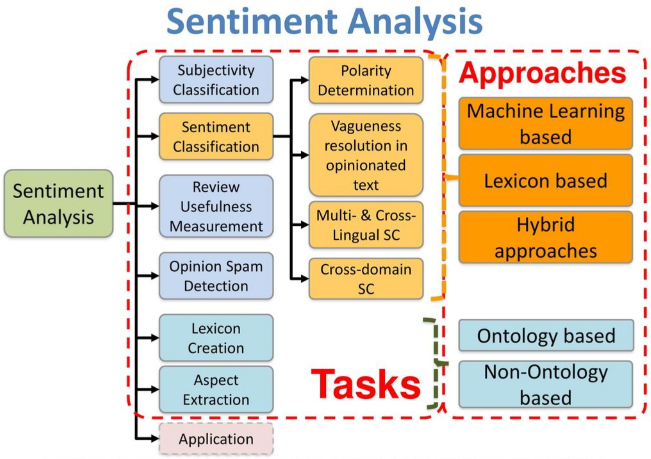
The rest of the paper is organised as follows, Section 2 presents the related work, section 3 describes the dataset used, section 4 explains the features employed and the reason for the same, section 4 shows the methodology used in all 3 techniques for each algorithm, section 5 explains the experimental results, section 6 presents the conclusion and future works, section 7 consists of the acknowledgement and finally section 8 cites the references.

II. Related Work

Many approaches and techniques have been proposed in the field of fake review detection. One strategy employed by Pennebaker et al. [1] employs Detection of Psycholinguistic Deception. The psycholinguistic technique is to assign psycholinguistic meanings to the key features of a review, use the famous Linguistic Inquiry and Word Count (LIWC) software [12] to build their features for the reviews. Another by Xinuyue et al. [2] propose two types of features and apply supervised machine learning algorithms for performing classification on Yelp's real-life data. In terms of features used, there are two new semantic feature sets: readability features and topic features. Ott et al. [3] explored the parts-of-speech (POS) distribution of the review and use the frequency of POS tags as the features representing the review to make prediction. Moraes et al. [4] proposed a technique for categorizing a single topic textual review. A sentiment classified document level is applied for stating a negative or positive sentiment. Supervised learning methods are composed of two phases, namely selection and extraction of reviews utilizing learning models such as SVM. You et al. [5] analyzed sentiment using images instead of textual contents. Nakov et al. [6] investigated sentiment analysis on Twitter. Based on a neural network method, Tang et al. [7] gave a model for predicting review ratings. Employment of linguistic features is popular in this field, Feng et al. [8] take lexicalized and unlexicalized syntactic features using sentence parse trees for deception detection.

Wenquian et al. [9] proposed a method for detection of deceptive reviews given product features wherein they first extract the review records of products to a temporal feature vector and then develop an isolation forest algorithm to detect outlier reviews by focusing on the differences between the patterns of product reviews to identify outlier reviews. They verify the effectiveness of our method and compare it to some existing temporal outlier detection methods using the crawled Amazon China dataset. While the above methods have largely been review centric as employed in our study there has been extensive research in analysing the behavioral aspects of spammers as well to detect fake reviews or the combination of both review and reviewer centric features.

Lim et al. [9] identified the anomalous rating and review behaviours such as giving unfair ratings to products and reviewing too often, so as to detect spammers. Akoglu et al. [10] proposed a novel spammer and fake reviews spotting framework which is complementary to previous works based on text and behavioural features. Fei et al. [11] exploit the burstiness nature of reviews to spot review spammer. Through a Markov Random Field model, their approach models the reviews in bursts and their co-occurrences in the same burst. Lim et al. [9] finds that spammers may write fake reviews in collusion. Based on the findings, they make composed model to integrate these features for spammer detection.



Source: Kumar Ravi and Vadlamani Ravi (2015), "A survey on opinion mining and sentiment analysis: tasks, approaches and applications." Knowledge-Based Systems, 89, pp.14-46.

Fig. 1. Various approaches for sentiment analysis.

III. Dataset

The corpus chosen consists of truthful and deceptive hotel reviews of 20 Chicago hotels contains:

- 400 truthful positive reviews from TripAdvisor
- 400 deceptive positive reviews from Mechanical Turk
- 400 truthful negative reviews from Expedia, Hotels.com, Orbitz, Priceline, TripAdvisor and Yelp
- 400 deceptive negative reviews from Mechanical Turk Each of the above datasets consist of 20 reviews for each of the 20 most popular Chicago hotels . The files are named according to the following conventions:
  - Directories prefixed with fold correspond to a single fold from the cross-validation experiments reported in [1] and [2].
  - Files are named according to the format $\%c\%h\%i.txt, where :$
    - %c denotes the class: (t)ruthful or (d)eceptive
    - %h denotes the hotel:
      - ∗ affinia: Affinia Chicago (now MileNorth, A Chicago Hotel)
      - ∗ allegro: Hotel Allegro Chicago - a Kimpton Hotel
      - ∗ amalfi: Amalfi Hotel Chicago
      - ∗ ambassador: Ambassador East Hotel (now PUBLIC Chicago)
      - ∗ conrad: Conrad Chicago

- ∗ fairmont: Fairmont Chicago Millennium Park
- ∗ hardrock: Hard Rock Hotel Chicago
- ∗ hilton: Hilton Chicago
- ∗ homewood: Homewood Suites by Hilton Chicago Downtown
- ∗ hyatt: Hyatt Regency Chicago
- ∗ intercontinental: InterContinental Chicago
- ∗ james: James Chicago
- ∗ knickerbocker: Millennium Knickerbocker Hotel Chicago
- ∗ monaco: Hotel Monaco Chicago - a Kimpton Hotel
- ∗ omni: Omni Chicago Hotel
- ∗ palmer: The Palmer House Hilton
- ∗ sheraton: Sheraton Chicago Hotel and Towers
- ∗ sofitel: Sofitel Chicago Water Tower
- ∗ swissotel: Swissotel Chicago
- ∗ talbott: The Talbott Hotel

We have combined the above files into one dataset having the columns - text(review), hotel, polarity and deceptive(indicating if the review is truthful or deceptive). This dataset is considered as a gold standard for labelled corpus in the field of fake review identification. The procedure of the dataset creation can be found in references [3] and [14].

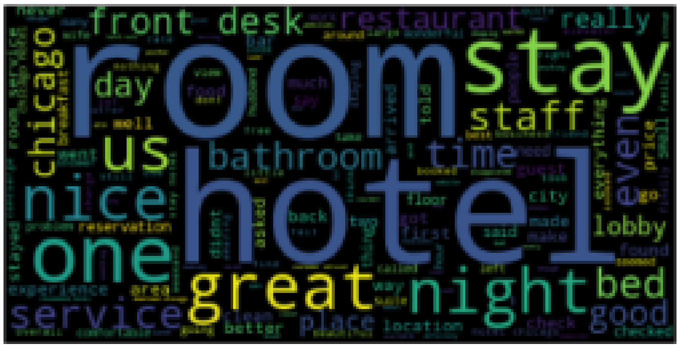


Fig. 2. Word Cloud of hotel review dataset

IV. Features

We have used a combination of the features given below to get the best accuracy from each algorithm employed.

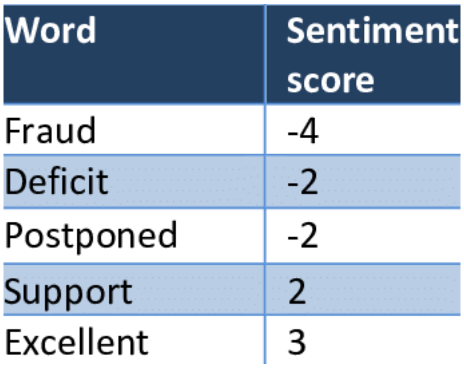
- n-grams : Unigram and bigram features have been taken as input for the classification.For models requiring numeric input tfidf was used which is a way to to convert textual data to numeric form, and is short for Term Frequency-Inverse Document Frequency. The vector value it yields is the product of these two terms; TF and IDF.TF stands for term frequency while Inverse Document Frequency (IDF) measures how important a word is. It works better than the simple bag of words representation which just keeps track of the term frequency.
- Total length of the review - the reason behind choosing this feature was the assumption that very short reviews are probably fake as they are vague and don't have much details on the products or services. Generalities are used with no specifics. Even if a spammer trued to use words similar to real reviews, he probably didn't spend as much time in writing the review.Lack of domain knowledge also increases the reason for the same.But it could also so happen that the spammer tried to overdo his job and wrote a very long review which marks this as an important feature.
- Ratio of capital letters - Fake reviews try to grab the attention of viewers and this could be one of the ways to do so.
- Ratio of capital words - For the same reason as mentioned above.
- Count of first pronouns - Reviewers often try to establish themselves as genuine by using a lot of personal pronouns - I,we,my. Using too many personal pronouns sounds stiff and not like something a real person would write.
- Count of sentiment words - Fake reviews tend to consist of more sentiment words than true reviews. Bing Liu and Minquing Hu sentiment lexicon was used to for sentiment words consisting of 6787 sentiment words which were compared against the words review to obtain a sentiment word count.
- Ratio of verbs to nouns - Researches have found that deceivers tend to use more verbs than nouns.POS tagging was employed to find the count of the number of verbs and nouns and the ratio was then obtained by dividing the same.
- AFINN score - AFINN lexicon is one of the simplest and most popular lexicons that can be used extensively for sentiment analysis. It associates a polarity score with each word indicating the degree of polarity. A normalized afinn score was found for each review by dividing by the word count. This uses the logic that fake reviews tend to be either extremely positive or extremely negative. An example is shown in Figure 2.
- VADER score - VADER is a lexicon and rule based sentiment analysis tool that is specifically attuned to sentiments expressed in social media.It scores the polarities primarily based on these features - punctuation,capitalization,degree modifiers,conjunctions,preceding trigram. The logic for using this is same as that for AFINN.
- Fake score - A fake score was found by us by constructing a vocabulary of words that occur most often in deceptive reviews.

V. Methodology

A. *Pre-processing and Feature Extraction*

The preprocessing steps include :

1) Tokenization - Each review was tokenized into words by using the nltk.tokenize module of nltk.



Word	Sentiment score
Fraud	-4
Deficit	-2
Postponed	-2
Support	2
Excellent	3

Fig. 3. Example of afinn score.

2) Removal of stopwords - The stopwords are the words we want to filter out, eliminate, before training the classifier. Some of these words are commonly used (“a,”the”,”it”) etc. that do not provide any substantial information to our labelling scheme but only confuse our classifiers.However the features such as count of personal pronouns are extracted from the original text.To compare with the stop words the text was converted to lower-case.

Feature extraction - A list of first personal pronouns was defined and the count for the same in each review was found. In a similar manner, the total length, capital-letters ratio, capital-words ratio was found. The sentiment words text file from Bing Liu and Minquing Hu sentiment lexicon was opened and the words were inserted into a sentiment list and a comparison was done with the words in the review column to get the count of the sentiment words. POS tagging was done to find out the count of the number of verbs and number of nouns and the ratio for the same was found. (Verbs can be ‘VERB,’ VB’, ‘VBN’, ’VBD’, ’VBZ’, ’VBG’ ,’ VBP’ and nouns can be ‘NOUN’, ’NNP’, NN’, ’NUM’, ’NNS’, NP’, ’NNPS’). The afinn lexicon and VADER packages were installed and the respective scores were calculated. A fake words vocabulary dictionary was constructed by creating BOW for both the deceptive and truthful reviews in the training dataset. More words were added into these vocabularies by using the dictionary based lexicon method [13].

One simple technique using only afinn scores uses the logic that reviews with afinn scores close to the mean score would be classified as true and reviews with afinn scores lying more than 1.5 standard deviations away is labelled fake as spammers give an extreme opinion of either positive or negative polarity. Just using this simple technique did not yield a very high accuracy, however the recall was really high.

Another technique using the fake scores uses the logic that for every word in reviews that is in the fake vocabulary we subtract the fake score by 1 else we add it by 1 if it belongs to the not-fake vocabulary. If the fake score ends up being negative the review is labelled as fake and if it ends up positive it is labelled as not fake. Just using this technique gives an accuracy of 70%. More information on this method is given in Section 6.4.

B. Supervised Learning

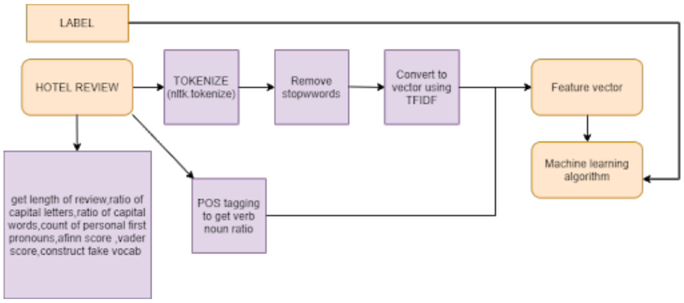

LABEL
HOTEL REVIEW
TOKENIZE (nltk.tokenize)
Remove stopwwords
Convert to vector using TFIDF
Feature vector
Machine learning algorithm
get length of review,ratio of capital letters,ratio of capital words,count of personal first pronouns,afinn score ,vader score,construct fake vocab
POS tagging to get verb noun ratio

Fig. 4. Flowchart for Supervised Learning algorithms.

1) NAIVE BAYES - Multinomial Naive Bayes classification algorithm tends to be a baseline solution for sentiment analysis task. The basic idea of Naive Bayes technique is to find the probabilities of classes assigned to texts by using the joint probabilities of words and classes..We built the naive bayes classifier from scratch and found that the features other than the TFIDF vector did not have much of an impact only about 2-3 Algorithms that implemented this theorem can run much faster due to its simplicity.Credulous Bayes classifier accepts that the impact of a specific element in a class is free of different highlights. For instance, an advance candidate is attractive or not relying upon his/her pay, past credit and exchange history, age, and area. Regardless of whether these highlights are associated, these highlights are as yet considered freely. This supposition improves calculation, and that is the reason it is considered as credulous. This supposition that is called class contingent autonomy. P(A/B)=P(B/A)*P(A)/P(B) A classifier is an ML model that is utilized to separate various articles dependent on specific highlights. A Naive Bayes classifier is a probabilistic ML model that is utilized for order task. The core of the classifier depends on the Bayes hypothesis. Utilizing Bayes hypothesis, we can discover the likelihood of An event, given that B has happened. Here, B is the proof and An is the theory. The suspicion made here is that the indicators/highlights are free. That is nearness of one specific component doesn’t influence the other. Henceforth it is called innocent.

2) SVM - SVM is considered a state-of-the-art machine learning technique and is used in many real-life classification problems. It is a robust model that does well in sentiment analysis. In short, what a SVM does is finding the best hyperplane that separates the data points of two different classes (deceptive and truthful) .Both sides of the plane represent a different class and this plane can thus be seen as a decision boundary for any new point, because we can easily classify this new point based on which side of the plane it lies.

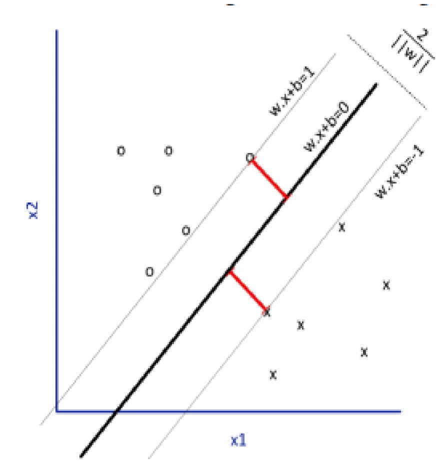


Fig. 5. SVM classifier.

Therefore, after a decision boundary is obtained, a SVM is very useful for making accurate predictions of new data points.

3) Decision trees - Decision tree classifiers are utilized as a well known classification technique in different pattern recognition issues, for example, image classification.Decision tree classifiers perform more successfully, specifically for complex classification problems, due to their high adaptability and computationally effective features.The parameter were adjusted for maximum accuracy by trial and error and the attribute selection method chose was ‘entropy’.
4) Random forest - Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithm of the same type i.e. multiple decision trees, resulting in a forest of trees, hence the name "Random Forest". The following are the basic steps involved in performing the random forest algorithm:
   - Pick N random records from the dataset.
   - Build a decision tree based on these N records.
   - Choose the number of trees you want in your algorithm and repeat steps 1 and 2.
   - For the classification problem, each tree in the forest predicts the category to which the new record belongs. Finally, the new record is assigned to the category that wins the majority vote.
5) LSTM - Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. LSTMs are a complex area of deep learning. The methodology for classification using LSTM is illustrated in Figure 4. LSTMs have an edge over conventional feed-forward neural networks and RNN in many ways. This is because of their property of selectively remembering patterns for long durations of time.

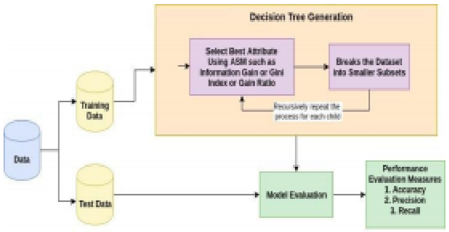

Decision Tree Generation
Select Best Attribute Using ASM such as Information Gain or Gini Index or Gain Ratio
Breaks the Dataset into Smaller Subsets
Recursively repeat the process for each child
Data
Training Data
Test Data
Model Evaluation
Performance Evaluation Measures
1. Accuracy
2. Precision
3. Recall

Fig. 6. Decision Tree classifier.

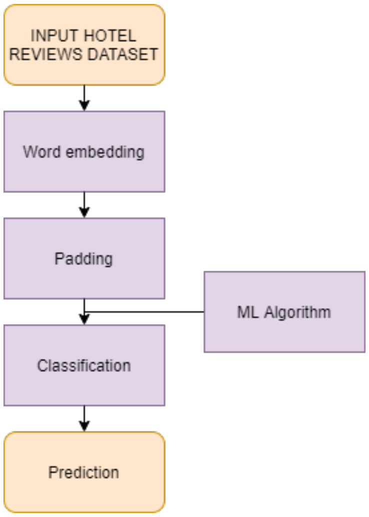


Fig. 7. LSTM Flowchart.

6) ANN - We implemented an Artificial Neural Network from scratch. Artificial neural networks or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains. The NN consists of an input layer, hidden layer and an output layer. The number of units in the input layer is equal to the size of the BOW consisting of words that occur frequently in the training dataset. More words were added to the datasets using the dictionary based lexicon method [13]. The number of hidden units is taken to be twice the number of input units. Output layer has two units corresponding to each label. Initially all the units are assigned values between -0.05 and 0.05. Word vectors of each review are passed as input.

C. *Unsupervised Learning*

K-Means - The K-means clustering algorithm collects the extracted terms according to their feature values into K number clusters, and K is any positive number that is used to determine the number of clusters. The K-means clustering algorithm performs the following steps:

1) Pick a number (K) of cluster centers (at random)
2) Assign every item to its nearest cluster center (e.g., using Euclidean distance)

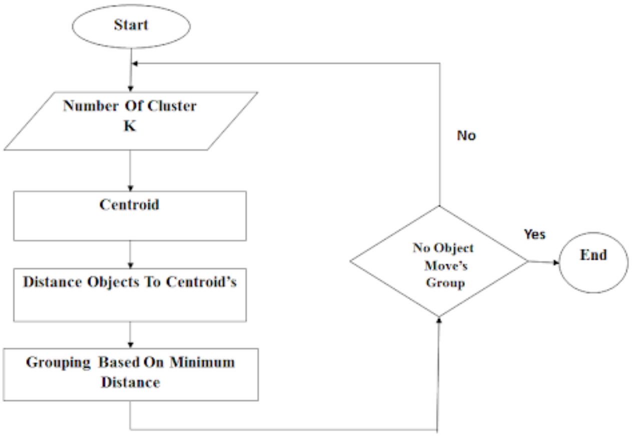


Fig. 8. K-Means Flowchart.

3) Move each cluster center to the mean of its assigned items
4) Repeat steps 2 and 3 until convergence is achieved (change in cluster assignments less than a threshold)

The fakescore is used as a feature here to assign clusters. Once the final clusters are formed, the cluster whose centroid is closer to 0 is labelled as fake.

D. Lexicon Based Approach

Application of a lexicon is one of the two main approaches to sentiment analysis and it involves calculating the sentiment from the semantic orientation of word or phrases that occur in a text.We have used the following approach described for our sentiment analysis:

Dictionary based approach - In the dictionary based method targeted opinion words with an identified orientation are collected and are then searched from the WordNet dictionary for their antonyms and synonyms. The newly found words are added to the seed list. This iterative process is continued until no new words are found. The limitation of this method is that it is usually difficult to find different opinion words for a specific domain [13].

Figure 7 depicts the difference between ML approach and Lexicon based approach.

VI. Results

It was found that for Naive Bayes and SVM the features other than the TFIDF vectors of the word had a small increase in accuracy but for Decision tree and Random forest the other features showed a significant increase in accuracy. The lexical approach gives a lower accuracy than the ML based approaches.The accuracy, precision and recall of the different classifiers are given in the Table 1 and Figure 8.

VII. Conclusion and Future Work

Our dataset is taken from Non-spam hotel reviews by TripAdvisor and Spam reviews by Amazon Mechanical Turk. Dataset consists 20 hotels in Chicago area and size is 1.6mb. Total number of reviews collected is 1600 and

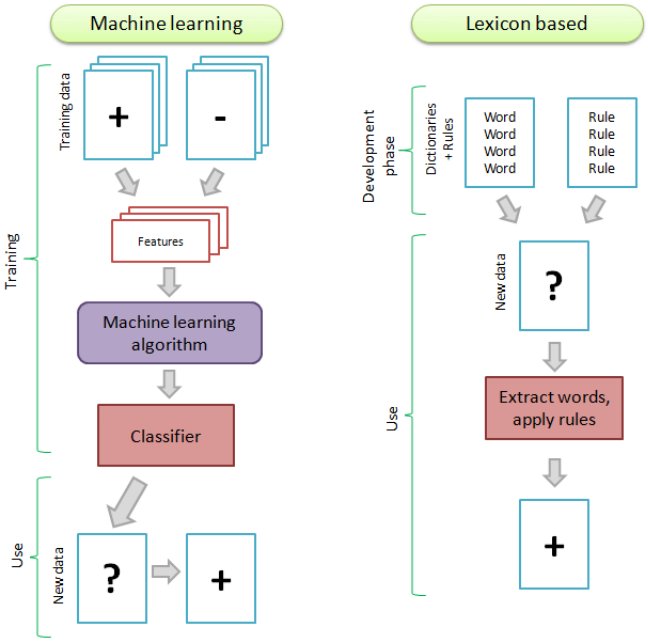


Fig. 9. ML vs Lexical approach.

Technique	Classifier	Accuracy	Precision	Recall
Supervised learning	Naive Bayes	85.416 %	84 %	84.75 %
	SVM	88.75 %	90 %	88 %
	Decision Trees	71.25 %	75.00 %	66 %
	Random Forest	84.375 %	86 %	84 %
	LSTM	86.25 %	85 %	90 %
	ANN	87.58 %	87.72 %	88.75 %
Unsupervised Learning	K-Means	64.79 %	62.72 %	72.91 %
Lexicon Approach	Dictionary method	70.825 %	67.68 %	75.46 %

TABLE I
Performance of different Models.

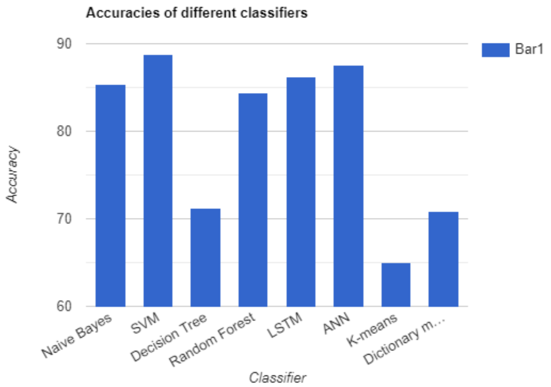


Fig. 10. Bar-graph showing the accuracies of each classifier.

80 reviews per hotel. It consists 40 spam reviews and 40 non spam reviews per hotel.This set consists both positive and negative reviews.This data corpus contains 400 truthful positive reviews from tripadvisor, 400 deceptive positive reviews from mechanical turk, 400 truthful negative reviews from expedia,hotels.com,tripadvisor,yelp and 400 deceptive negative reviews from mechanical turk.

We have thus performed three methods of sentiment classification on this dataset - supervised,unsupervised and lexicon based to detect fake reviews.We observe that the machine learning methods perform far better using the feature vectors chosen than the other methods. Scope for future work would be to improve the lexicon based approach and to form more hybrids for better accuracy. Another aspect we would like to explore is to scrape reviewer centric features and use them in addition to the existing features for classification.There is also more scope in improving the accuracies for unsupervised learning algorithms such as k-means.Finding a dataset with labelled data with respect to decptive reviews has been difficult,but to improve out models we would like to scrape more data and obtain the labels with the right permissions from organisations such as Yelp.

Acknowledgement

We would like to thank the Department of Computer Science and Engineering for giving us this opportunity to do a research project on the classification of fake reviews.

We would like to thank the chairperson, Dr. Shylaja S for giving us this opportunity and for her continuous support. We would finally like to thank all the staff members of the computer science department for their round the clock guidance.

References

[1] J. W. Pennebaker, M. E. Francis, and R. J. Booth, Linguistic Inquiry and Word Count: Liwc 2001, vol. 71, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2001.

[2] Jia, Shaohua Zhang, Xianguo Wang, Xinyue Liu, Yang. (2018). Fake reviews detection based on LDA. 280-283. 10.1109/INFOMAN.2018.8392850.

[3] Ott, M., Choi, Y., Cardie, C., Hancock, J.T. (2011). Finding Deceptive Opinion Spam by Any Stretch of the Imagination. ArXiv, abs/1107.4557.

[4] R. Moraes, J. F. Valiati, and W. P. G. Neto, "Document-level sentiment classification: An empirical comparison between svm and ann," ExpertSystems with Applications, vol. 40, no. 2, 2013, pp. 621–633.

[5] Q. You, J. Luo, H. Jin, J. Yang, "Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks," AAAI, 2015.

[6] Preslav Nakov, Alan Ritter, Sara Rosenthal, Veselin Stoyanov, and Fabrizio Sebastiani. 2016a. SemEval2016 task 4: Sentiment analysis in Twitter. In Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval '16, San Diego, California

[7] Tang, B. Qin, T. Liu, Y. Yang, "User Modeling with Neural Network for Review Rating Prediction," IJCAI, 2015

[8] S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers— Volume 2, pp. 171–175, Association for Computational Linguistics, 2012.

[9] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, "Detecting product review spammers using rating behaviors," in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM '10), pp. 939–948, ACM, Toronto, Canada, 2010.

[10] L. Akoglu, R. Chandy, and C. Faloutsos, "Opinion fraud detection in online reviews by network effects," ICWSM, vol. 13, pp. 2–11, 2013.

[11] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Exploiting burstiness in reviews for review spammer detection," ICWSM, vol. 13, pp. 175–184, 2013.

[12] B. Pang and L. Lee, "A sentimental education: Sentiment analysis Using subjectivity summarization based on minimum cuts," in Proceedings ofthe 42nd annual meeting on Association for Computational Linguistics.Association for Computational Linguistics, 2004, p. 271. [Online]. Available from: http://www.cs.cornell.edu/People/pabo/movie%2Dreview%2Ddata

[13] Naveed Hussain , Hamid Turab Mirza , Ghulam Rasool , Ibrar Hussain2 and Mohammad Kaleem. Spam Review Detection Techniques: A Systematic Literature Review.

[14] Myle Ott, Claire Cardie, Jeffrey T. Hancock. Negative Deceptive Opinion Spam.