EasyChair Preprint

№ 7686

Recognition of Fabricated Online Reviews Using Semi-Supervised Learning

Satyanarayana Botsa and Dinesh Chandra Pancharia

April 2, 2022

Recognition of fabricated online reviews using Semi-Supervised learning

B Satyanarayana[1] Dinesh Chandra Pancharia[2]

[1]*Dept.ofComputerScience&Engineering,AVANTHI InstituteofEngineering&Technology*,Vizianagaram,A.P., INDIA.

[2]*Dept.ofComputerScience&Engineering,AVANTHI InstituteofEngineering&Technology*,Vizianagaram,A.P., INDIA.

[1]Satyanarayana.botsa@gmail.com,

[2] *dinesh.pancharia@gmail.com*

Abstract— In the World Online reviews have major impact on daily business and e-commerce . Purchase of products in online mostly depends on reviews given by the users, reviews are the major parameter of decision making . Thus, opportunistic individuals or groups attempt to use product reviews to advance their own interests. We propose some semi-supervised text mining models to detect false online reviews in this paper.

IndexTerms—Fakereviews,semi-supervisedlearning,super-visedlearning,NaiveBayesclassifier,SupportVectorMachineclassifier,Expectation-maximizationalgorithm.

I. INTRODUCTION

Sophisticated and new technologies continually replace the old ones . These the new technologies are enabling people to have their work done efficiently. Such an evolution of technology is the online marketplace. We can shop and make a reservation using online websites. Almost, Before purchasing a product or service, we all check reviews. so the online reviews have become a great source of a reputation for companies. Also, they have a large impact on the advertisement and promotion of products and services. With the spread of the online marketplace, fake online reviews are becoming a great matter of concern. People can make false reviews for the promotion of their products that harms the actual users. some of , competitive companies can try to damage each other's reputation by providing fake negative reviews.

Researchers have been studying many approaches for the detection of these fake online reviews. Some approaches are review content-based and some are based on the behavior of the user who is posting reviews. The content-based study focuses on what is written on the review that is the text of the review whereas the user behavior-based method focuses on country, IP address, the number of posts of the reviewer. There are a large number of models based on supervised classification. There is a need for semi-supervised methods because reviews cannot be reliably labelled.

In this paper , we use classification approaches for detecting fabricated online reviews , some of which are semi-Supervised. For semi-supervised learning ,we use Expectation-maximization algorithm. Naïve Bayes classifier and Support Vector Machines(SVM) are used as classifiers in our research work, to get best performance of classification. Review-based approaches have mainly been examined in terms of content . As feature we have used word frequency count, sentiment polarity and length of review.

In the following section-II, we discuss about the related works. Section-III describes our proposed approaches and experiment setup . Results and finding so four research are discussed in Section IV. Section-V includes conclusions and future work.

II. RELATEDWORK

In the field of fabricated review detection, a number of approaches and techniques have been proposed . The following methods have been able to Recognition fabricated online review with higher accuracy.

Content Based Method: Methods that focus on content examine what is in the review. These three techniques are 1. genre identification 2. detection of psycholinguistic deception 3. text-categorization

1) Genre Identification :It is explored in Ottetal how the review's parts-of-speech distribution is distributed .In order to classify reviews, the features used were the frequency count of POS tags.
2) Detection of Psycholinguistic Deception :Assigning psycholinguistic meaning to the important feature so far is the goal of the psycholinguistic method . Linguistic Inquiry and Word Count(LIWC)software was used by Pennebakeretal. To build their features for the reviews.
3) Text Categorization : Ottetal. A popular feature of fake review detection is the n-gram, which was initially a research project. Other linguistic features are also explored. Such as ,Fengetal. Took lexicalized and un-lexicalized syntactic features by

Constructing sentence parse trees for fabricated review detection . They shown experimentally that the deep syntactic features improve the accuracy of prediction. explored a different of generic deceptive signals which contribute to the fabricated review detection and They also concluded that merged general features such as LIWC or POS with bag of words will be more robust than bag of words alone. Meta data about reviews such as reviews length ,date time and rating are also used as features by some researchers.

a) Behavior Feature Based Methods: it focuses on the reviewer that includes character-istics of the person who is giving the review .addressed the problem of reviews detection, or finding users who are the source of spam reviews. People who post intentional fabricated reviews have different behavior than the normal user reviews . They have notified the below ambiguous rating and review actions.

- Giving unfair rating too often: the Professional spammers generally posts more fake reviews than there alones . let Suppose a product has average rating of 9.0 out of 10. But a viewer has given rating is 4.0 . The other reviews of the reviewer can reveal whether he is a spammer by analyzing his other reviews if he consistently gives misleading and unjust ratings.
- Giving good rating to own country's product :False reviews are sometimes posted by people to promote the products of their own language . Reviews of movies are usually the target of this type of spamming . For example, suppose an Indian movie receives a 9.0 out of 10 rating on an international website that most of the reviewers are Indian . This kinds of spamming can easly identified using reviewers
- address .
- Review on a vast variety of product : Every Indiidual person has specific interests of his own. In general, people aren't interested in every product . Suppose a person who loves eletrical devices may not be interested in gaming devices . We can tell if some people are deliberately giving fake reviews when their behavior exceeds the general behavior when they give reviews

in various types of products.

As the detection of fraudulent online reviews is a classification problem, supervised text classification techniques are being used as one popular approach As long as large datasets of labeled instances from both classes, deceptive opinions (positive examples) and truthful opinions (negative examples), are used, the techniques are likely to be robust.Some researchers also used semi-supervised classification techniques.

For sem-supervised classification process ground truth is deter-mined by–helpfulness vote, rating based behaviors ,using seed words, human observation etc. An improved method is proposed in which a bagging model bags three classifiers: product word composition classifier (PWCC), *trigrams SVM classifier (*$TRIGRAMS_{SVM}$*), and bigrams SVM* classifier (*BAGRAMS*$_{SVM}$).To predict the polarity of reviews, the authors developed a product word composition classifier .The model was used to map the wordsof a re

view.

Including the relationships between products and reviews within the continuous representation .For the formulation of the document model, they used the product word composite as input and created a representation model with Convolutional Neural Network (CNN).Following the $TRIGRAMS_{SVM}$ and $BIGRAMS_{SVM}$ classification, the F-score was 0.77 after bagging the result.

How ever semi-supervised method has some challenges to over-come.Supervised techniques present the following problems .

- Assurance of the quality of reviews is challenging.
- It is difficult
- to obtain labeled data points for training classifiers.
- Humans aren't very good at identifying fake reviews

.Therefore, a semisupervised method is proposed where labeled and unlabeled data are trained together.

Following are some examples of semi-supervised method s they proposed.

1) When reliable data is not available.
2) Dynamic nature of online review.
3) Designing heuristic rules are difficult.

These semi-supervised learning techniques include co-training, expectation maximization, label propagation, and positive unlabelled learning . A number of classifiers were utilized including K-Nearest Neighbor, Random Forest, Logistic Regression, and Stochastic Gradient Descent . They achieved a high accuracy of 84% using semi-supervised techniques.

III. PROPOSEDWORK

A. DatasetDescription

In this paper ,the 'gold standard' data set developed by Ottetal. Is used in our evaluations.The data set contains 1,600 reviews in text formaton 21 restrunts in United States America .Here we have 800 fake reviews and 800 true reviews. For the evaluations ,a value of '0' is denotes false reviews , where as '1' is denotes Trusted reviews .In the dataset, more than 400 genuine reviews had a negative sentimental polarity, while 400 demonstrated a positive sentimental polarity.Additionally, 400 fake reviews contain positive sentiments while the other 400 are negative in nature. We collected these reviews from a variety of sources.In addition to the deceptive reviews, other reviews were obtained from Yelp,TripAdvisor,Expedia, and Hotels, among others.cometc.

A fixed partition of the dataset is used for the evaluations.There are two sets of examples created from the 1600 examples in the corpus, such as : the training set and the test set .In ratios of 75:25 and 80:20, the corpus is divided between arteries and veins respectively.

B. Proposedmethodology

Recognizing fabricated online reviews using raw text data .Data that had already been labeled by previous researchers has been used by us .We remove unnecessary text such as prepositions and articles from the data

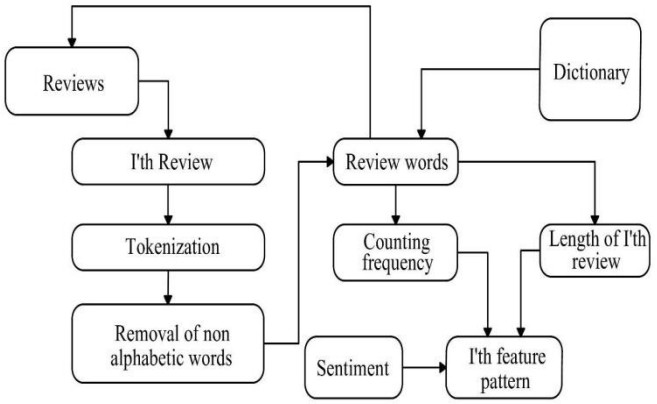

Fig.1. Stages of proposed feature extraction process

Algorithm 2 EM Algorithm

INPUT: Labeled instance set L, and unlabeled instance set U.

OUTPUT: Deployable classifier, C.

```
C ← train(L);
PU = ∅;
while true do
    PU = predict(C, U);
    if PU same as in previous iteration then
        return C;
    end if
    C ← train(L ∪ PU);
end while
```

Fig.2. Expectation-MaximizationAlgorithm

In order to make them suitable for the classifier, these text data are converted into numeric ones .A classification process took place after extracting important and necessary features

.

Based on the 'gold standard' dataset prepared by Ottetal.It wasn't necessary to deal with missing values, remove inconsistencies, remove duplicates, etc. In order to preprocess the text, we had to merge it, make a dictionary, and turn it into numeric value . Among the features that we analyzed were word frequency count, sentiment polarity, and

the length of the review . We taken 2000 words as a features words .Therefore, we need a feature vector of 160×2002. n-grams and parts of speech were not taken into account because they are derived features from bag of words and may cause over-fitting . Figure 1 summarizes the feature extraction process.

From the figure1 , we can see that, when we are working with *i'th* review ,In the following procedure, its corresponding features are generated.

1) Tokenization is the first step in every review. A set of candidate feature words is then generated
2) after removing unnecessary words.
3) In the feature vector, the frequency of each candidate feature word is calculated and added to the column that corresponds to the numeric map of the word
4) whose entry is found in the dictionary.
5) We also measure the length of a review along with counting frequency.
6) To finish, the feature vector
7) is enhanced with sentiment score, which is available in the dataset . The feature vector has zero values for negative sentiment and some positive values for positive sentiment.

We have implemented semi-supervised classifications.For semi-supervised classification of the dataset, we used Expectation-Maximization(EM) algorithm. it is first proposed by Karimpour, is designed to label unlabeled data to be used for training.

The algorithm operates as follows :The labeled dataset is first used to create a classifier . To label unlabeled data, the classifier is then applied. We'll call this predicted set of labels PU. The unlabeled data set is then reclassified using another classifier that is derived from the combined sets of both labelled and unlabelled data sets. Once the set PU stabilizes, the process is repeated . After as table *PU* set is produced , we have trained the classification algorithm with the combined training set of both labeled and unlabeled data sets and deploy it for predicting test dataset .The algorithm is given below.

With E-M algorithm, SVM and NB classifiers were used as classifiers. A Python library for these classifiers is provided by Scikit- Learn, a Python package. The SVM parameters have been tuned for better results. For semi-supervised classification,we have used NaïveBayes and SVM classifiers. There is a Naive Bayes classifier that can be implemented in a way that maintains the conditional independence property. Because the text comes from user's mind, we cannot predict what line and word will follow. It is a probabilistic method therefore it can be used for both classification and regression. It is also very fast in calculating results. Therefore, Naive Bayes is widely used in text mining.

IV. RESULTSANDPERFORMANCEANALYSIS

A. ExperimentalEnvironmentandTools

We have applied our experiments on amachine with Processor: Intel(R)Core(TM)i5-4200U and CPU-1.6GHz, RAM:6GB, Systemtype:64bitOS,x64-basedprocessor, HardDisk:1TB.We have used Linux(Ubuntu) operating system . Programming was done in Python using Scikit-Learn and Numpy packages.

B. Results

Our semi-supervised classification system has been based on the Expectation Maximization (EM) algorithm . As classifier we have used

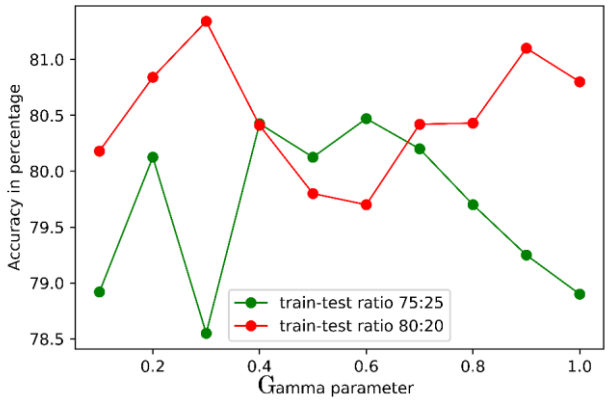

Fig.3. Graph showing Gamma parameter vs Accuracy for Supervised

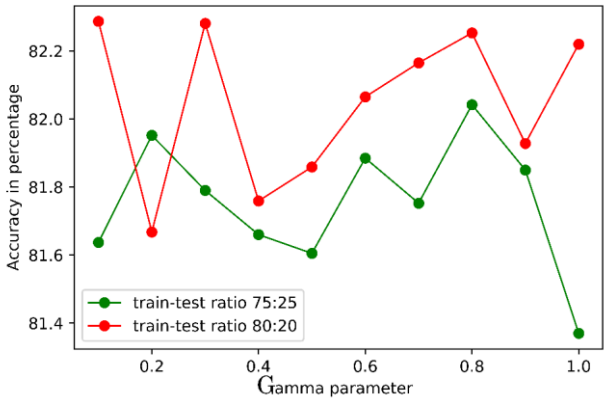

Fig.4.Graph showing Gamma parameter vs Accuracy for Semi-supervised SVM classier

Support Vector machines(SVM) and NaiveBayes classifier.For each classification process, we have divided our data into a ratio of 75:25 and 80:20

.

We have tuned different gamma parameters for semi-supervised classification with SVMs while keeping parameter constant .The percentage accuracy graphis shown in the figure3. In the above graph we can find that the accuracy of supervised classification with KNN classifier is 81.34% for 80:20 split ratio and 80.47% for 75:25 split ratio with gamma equal to 0.3 and 0.6 respectively. With a split ratio of 80:20 and 75:25 we have achieved an accuracy of 85.21 and 84.87 percent respectively for semi-supervised classification with SVM and Naive Bayes classifier.

C. PerformanceAnalysis

The above figure 5 shows a histogram of our implemented techniques as well as previous results on the dataset.

To reduce overfitting, we have carefully selected our features in our research .We have not taken derived features like

The accuracy of semi-supervised classification to 83.21% .We achieved a high accuracy rate of 85.75 % .Furthermore, supervised classification with Naive Bayes classier yielded the highest accuracy of 85.75 %. findings are summarized in the table I.

V. CONCLUSIONSANDFUTUREWORK

In this study, we used semi-supervised text mining methods for the detection of fake online reviews .A better set of features has been developed by combining features from several research works. As a result, we have tried some other classification methods in addition to those used previously . Thus,we have been able to increase the accuracy of previous supervised techniques done by Jitenetal.We have al so found out that Semi- supervised Naïve Bayes classifier gives the highest accuracy .This ensures that our data set is labeled well

TABLEI
COMPARATIVESUMMARYOFSEMI-SUPERVISEDANDSUPERVISEDLEARNINGTECHNIQUES

	Features	Algorithm Type	Classifier Used	Train-Test Ratio	Accuracy
Jitendra et al[8]	Bigrams, sentiment,Score, POS, LIWC	Supervised	K-NN	75:25 80:20	0.8300 0.8313
			Logistic Regression	75:25 80:20	0.8300 0.8375
Proposed Work	Word Frequency count, Sentiment score , review, size	Semi-Supervised	Naïve Bayes	75:25 80:20	0.8487 0.8521
			SVM	75: 25 80:20	0.8047 0.8134

When reliable labeling is not available, the semi-supervised model works well .

In our research work we completly foucs on reviews given by user . A better classification model can be built in the future by combining user behavior with texts . To make the dataset more accurate, preprocessing tools can be used for tokenization. A larger dataset can be used to evaluate the proposed methodology . English reviews are the only ones being looked at in this research .

REFERENCES

[1]ChengaiSun,QiaolinDuandGangTian,*"ExploitingProductRelatedReviewFeaturesforFakeReviewDetection,"*MathematicalProblemsinEngineering,2016.

[2]A.Heydari,M.A.Tavakoli,N.Salim,andZ.Heydari,"*Detectionofreviewspam: asurvey"*,ExpertSystemswithApplications,vol.42,no.7,pp.3634–3642,2015.

[3]M.Ott,Y.Choi,C.Cardie,andJ.T.Hancock,*"Findingdeceptiveopinionspambyanystretchoftheimagination,"*inProceedingsofthe49thAnnualMeetingoftheAssociationforComputationalLinguistics:HumanLanguageTechnologies(ACL-HLT),vol.1,pp.309–319,AssociationforComputationalLinguistics,Portland,Ore,USA,June2011.

[4]J.W.Pennebaker,M.E.Francis,andR.J.Booth,"*LinguisticInquiryandWordCount:Liwc*,"vol.71,2001.

[5]S.Feng,R.Banerjee,andY.Choi,*"Syntacticstylometryfordeceptiondetection,"*inProceedingsofthe50thAnnualMeetingoftheAssociationforComputationalLinguistics:ShortPapers,Vol.2,2012.

[6]J.Li,M.Ott,C.Cardie,andE.Hovy,*"Towardsageneralruleforidentifyingdeceptiveopinionspam,"*inProceedingsofthe52ndAnnualMeetingoftheAssociationforComputationalLinguistics(ACL),2014.

[7]E.P.Lim,V.-A.Nguyen,N.Jindal,B.Liu,andH. W.Lauw,*"Detectingproductreviewspammersusingratingbehaviors,"*inProceedingsofthe19thACMInternationalConferenceonInformationandKnowledgeManagement(CIKM),2017.

[8]J.K.Rout,A.Dalmia,andK.-K.R.Choo,*"Revisitingsemi-supervisedlearningforonlinedeceptivereviewdetection,"*IEEEAccess,Vol. 5,pp.1319–1327,2019.

[9]J.Karimpour,A.A.Noroozi,andS.Alizadeh,*"Webspamdetectionbylearningfromsmalllabeledsamples,"*InternationalJournalofComputerApplications, vol.50,no.21,pp.1–5,July2020.

[10] https://www.kaggle.com/rtatman/deceptive-opinion-spam-corpus