AN OPINION SPAM ANALYZER FOR PRODUCT REVIEWS USING SUPERVISED MACHINE LEARNING METHOD

[1]*DAYA L. MEVADA*, [2]*PROF.VIRAJ DAXINI*

[1]ResearchScholar, Master in Computer Engineering(M.E.-C.E.),
V.V.P. Engineering College, Rajkot-360001
[2]Assistant Professor, Computer Engineering Department,
V.V.P. Engineering College, Rajkot-360001

[1]*mevadadaya83@gmail.com*,[2]*viraj_daxini34@yahoo.com*

ABSTRACT: *While making any purchase online consumer usually checks opinions of others about the product. Manufacturer can gain insight into its products strength and weaknesses based on the reviews of the customers. So, user reviews play a crucial role in Web, since many decisions are made based on them. However usefulness of this reviewing systems motivates some people to enter their fake review to promote some products or defame some others. Opinion spam analyzer, which classifies reviews into non-spam reviews and spam reviews, and provide quality data to review mining these opinion spam should get detected and eliminated in order to prevent misleading potential customers. In this thesis, opinion spam detection approaches have been proposed and examined over a sample dataset. The proposed system will use supervised machine learning method to classify opinions into spam or non-spam. we proposed approach for performing spam detection in opinion reviews by merging methods from machine learning and text mining in one classification approach.*

KEYWORDS: Opinion mining, Review Spam, support vector machine.

1. INTRODUCTION

"What other people think" has always been an important piece of information for most of us during the decision-making process. Currently users of web 2.0 contribute content actively in product review websites, blogs and social media and web-forums. Opinions and reviews can now be found almost everywhere-blogs, social media like Facebook and Twitter, web-forums, e-commerce sites, RSS feeds etc. These opinions are helpful for both business organizations and individuals. However popularity of this reviewing systems motivates some people to enter their fake review to promote some products or defame some others.

2. PROBLEMDEFINITION

When performing any type of internet shopping, most users will spend a good amount of time reading user reviews if they are available. A survey performed by Faves.com has shown that [14] :

• More than 70% of online shoppers said they sometimes or frequently rely on online product or book reviews
• 62% rely on the popularity of information based on users' votes or ratings • 78% of have recently voted or rated something online
• 28% have recently written a product or book review

Clearly consumers value the feedback given by other users as do the companies that sell such products. Blogs, websites, discussion boards etc. are a repository of customer comments which are valuable and rich sources of textual data. Therefore individuals rely extensively on the reviews available online. It means that they make their decision of whether to buy products or not by analyzing the existing opinions on those products. In fact if a potential customer gets a positive overall impression of a product by considering the present sentiments for that product, it is highly probable that he will actually purchase the product. Normally if the percentage of positive opinions is considerable, it is likely that the overall impression will be highly positive. Likewise, if the overall impression is negative, it is less imaginable that they don't buy the product.

Now any people can write any opinion text, this can motivates the individuals, and organizations to give undeserving spam opinions to promote or to discredit some target products, services, organizations, individuals, and even ideas without disclosing their true intentions. These spammed opinion information is called opinion spam [3].

3. MOTIVATION

The motivation behind developing this system is that people are nowdays rely havily on opinions before buying anything. This motivates many peoples to write fraud opinions about other products or service. Even organizations are hiring professional to write fake reviews to promote their products or defame its competitors product. This fake opinions are misleads the customers buying experience and convince them to buy products which are based on fake opinions. So there is a need to devise a tool which can help them to find the true opinions about products, peoples and services. The proposed system will analyze the opinions and classifies them into spam and non spam. approach in both research and practice, and in information filtering and e-commerce applications [4].

4.SCOPE

Now any people can write any opinion text, this can motivates the individuals, and organizations to give undeserving spam opinions to promote or to discredit some target products, So there is a need to develop an intelligent system which automatically mine opinions and classify them into spam and non spam category. Proposed opinion spam analyzer will automatically classify user opinions into spam or non spam. This automatic system can be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing what customers are saying about products. Customers can make decision whether he/she should buy or not buy the products.

5.CONTRIBUTIONS

The intention of this research is to distinguish the fake opinions posted about products to intentionally change the overall sentiment of the products. The proposed system will save the efforts and time by helping the user and business organizations to identify spam in opinions quickly.

In this paper some of the basic terminologies used in Opinion Spam Analysis and throughout this thesis are formally defined, along with a brief summary of some of the classification techniques used in this thesis along with a summary of some of the most popular related works.

6. CLASSIFIER

Support vector machine(svm)classifier

A support vector machine (SVM) is a set of related supervised learning methods used for classification and regression. In simple words, given a set of training examples, each marked as belonging to one of two categories, the SVM training algorithm builds a model that predicts whether a new example falls into one category or the other. Intuitively, SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on [15][13][18].More formally, SVM constructs a hyper plane or a set of hyper planes in a high dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance tothe nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier [15] [13].Currently, SVM is widely used in object detection and recognition, content-based image retrieval, text recognition, biometrics, speech recognition, speaker identification, benchmarking time-series prediction tests. Equation 2.7 is dot product formula and used for the output of linear SVM, where x is a feature vector of classification documents composed of words, w is the weight of corresponding x, and b is a bias parameter determined by training process.

$y = w.x - b$ ……………………….. [13]

The following summarizes SVM steps:

1. Map the data to a predetermined very high-dimensional space via a kernel function.
2. Find the hyper plane that maximizes the margin between the two classes.
3. If data are not separable find the hyper plane that maximizes the margin and minimizes the (a weighted average of the) misclassifications.

SVM can be used for both linear and nonlinear data. It uses a nonlinear mapping to transform the original training data into a higher dimension. With the new dimension, it searches for the linear optimal separating hyper plane (i.e., "decision boundary"). With an appropriate nonlinear mapping to a sufficiently high dimension, data from two classes can always be separated by a hyper plane. SVM finds this hyper plane using support vectors ("essential" training tuples) and margins (defined by the support vectors). Figure 1 shows support vectors and how margins are maximized [9] [15] [13]

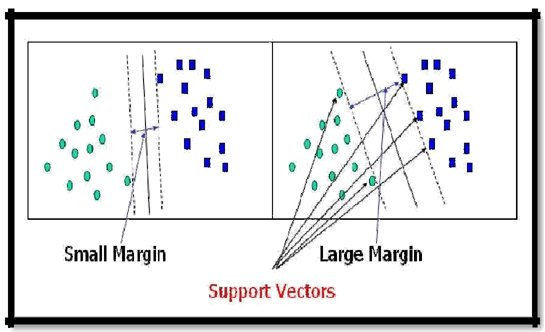

Figure1:SupportVectors

SVM is effective for high dimensional data because the complexity of trained classifier is characterized by the number of support vectors rather than the dimensionality of the data, the support vectors are the essential or critical training examples, they lie closest to the decision boundary, if all other training examples are removed and the training is repeated, the same separating hyper plane would be found. The number of support vectors found can be used to compute an (upper) bound on the expected error rate of the SVM classifier, which is independent of the data dimensionality. Thus, an SVM with a small number of support vectors can have good generalization, even when the dimensionality of the data is high [11 [15] [13] [18].

Practical real-world modelers have found that SVM have performed well when other classifiers did poorly. SVM classifiers have been widely used in text classification tasks with unbalanced training.

7. TYPEOFSPAMANDSPAMMING

Three types of spam reviews were identified in [1][8]:

Type1(untruthful opinions): these are spam reviews that are written not based on the reviewers' genuine experiences of using the products or services, but are written with hidden motives. They often contain undeserving positive opinions about some target entities (products or services) in order to promote the entities and/or unjust or false negative opinions about some other entities in order to damage their reputations.

Type2(reviews about brands only): these reviews do not comment on the specific products or services that they are supposed to review, but only comment on the brands or the manufacturers of the products. Although they may be genuine, they are considered as spam as they are not targeted at the specific products and are often biased.

Type3(non-reviews): these are not reviews, which have two main sub-types: (1) advertisements and (2) other irrelevant reviews containing no opinions (e.g., questions, answers, and random text).

Spam reviews may be written by many types of people, e.g., friends and family, company employees, competitors, businesses that provide spam review writing services, and even genuine customers (some businesses give discounts and even full refund to some of their customers on the condition that the customers write positive reviews for them) [3] [1] [8][18].In general, a spammer may work individually, or knowingly or unknowingly work as a member of a group (these activities are often highly secretive) [17] [8].

- **Individual spammers:** in this case, a spammer, who does not work with anyone else, writes reviews. The spammer may register at the review site as a single user, or as many fake users using different user-IDs. He/she can also register at multiple review sites and write spam reviews.
- **Group spammers:** there are two main sub-cases [16] [17].

o A group of spammers (persons) works in collusion to promote a target entity and/or to damage the reputation of another. The individual spammers in the group may or may not know each other.

O A single person registers multiple user-IDs and spam using these user-IDs. These multiple *user*-IDs behave just like a group in collusion. Group spamming is highly damaging due to the sheer number of members in a group, it can take total control of the sentiment on a product and completely mislead potential customers, especially at the beginning of a product launch [3] [5] [17] [8].

Types of Data, Features and Detection

Three main types of data have been used for review spam detection [3] [8] [18]:

- **Review Content :**the actual text content of each review. From the content, we can extract linguistic features such as word and POS n-grams and other syntactic and semantic clues for deceptions and lies. However, linguistic features may not be enough because one can fairly easily craft a spam review that is just like a genuine one.
- **Meta-data about each Reviewer:** the data such as the star rating, user-ID of the reviewer, the time when the review was posted, the host IP address, MAC address of the reviewer's computer, and the geo-location of the reviewer.
- **Product Information:** information about the entity being reviewed, e.g., the product description and sales volume/rank.

8. ARCHITECTURE

The proposed framework consists of three components. The first component is responsible for acquiring and preprocessing the data from user side or crawling the web. The second one is a text mining component which performs certain operations on data from the above component before analyzing the opinions. The third component performs spam classification of the data.

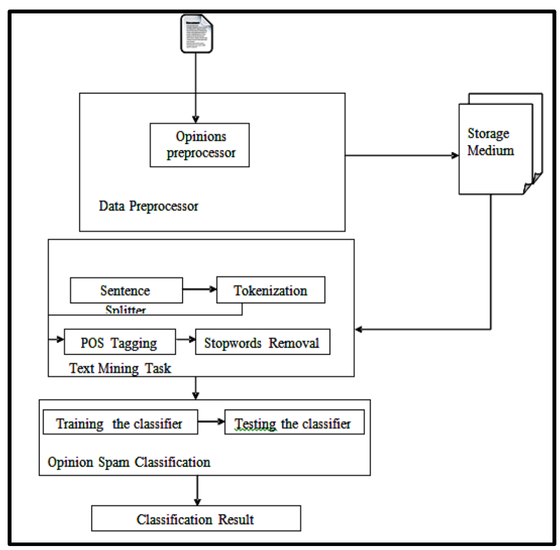

Figure 2: Architecture of SVM

COMPONENTSOFTHEPROPOSEDSYSTEM

- **Data Preprocessor Component**

Data preprocessor component aims to collecting and cleaning the data before subsequent analysis. Spidering programs will be used to collect the source data from websites as HTML pages. When data are gathered from websites it contains HTML tags and other non-textual data. opinion Spam classification task usually requires text data in the form of comments. The data preprocessor module removes such non textual contents and gives the data in structured text format for using in opinion spam classification. The preprocessor removes noisy characters from the input documents.

- **Text Mining Component**

Text Mining component perform various text related operations on cleaned data obtained from previous component. It performs following operations on cleaned data.

a) Sentence Splitter: Sentence splitter splits the whole text data into individual sentences. Usually Sentences begin with a capital letter which is the most identifiable sign for sentence splits. "." mark is considered as the sentence end, if it is not preceded by words such as Pvt., Ltd., etc. Some punctuation, such as ";", "!" And "?" are also used for sentence splitting.

b) Tokenization: Tokenization process splits the text into very simple units such as numbers, punctuation and words.

c) Part-of-speech(POS)Tagging: Not all the words in source text data are useful for sentiment analysis. The part-of-speech (POS) tagging assigns each token a tag which may be adjective, verb, or adverb. Liu et al. [8] founds that nouns and noun phrases in the sentences are the features of the products and adjectives and adverbs are used to express opinions through opinion words. Part-of-speech (POS) tagging is useful for identifying adjectives and adverbs in the sentences which identify the opinion words, and nouns which identify the features of the products.

d) Stop words Removal: stop words such as 'a, an, is, the' are quite frequently seen in opinions but for classification there is no requirement of such words. So this component will remove such useless words for spam classification.

Opinion Spam Classification

In this component, two sets of documents are needed: training and a test set. A training set is used by an automatic classifier to learn the differentiating characteristics of documents, and a test set is used to check how well the classifier performs. A number of machine learning techniques have been adopted to classify the reviews. Machine Learning techniques like Naive Bayes (NB), Maximum Entropy (ME), and Support Vector Machines (SVM) have achieved great success in opinion spam classification. Machine Learning starts with collecting training dataset. The next step is to train a classifier on the training data. Once a supervised classification technique is selected, an important decision to make is feature selection. They can tell us how documents are represented. The most commonly used features in classification is Term presence and their frequency. These features include uni-grams or n-grams and their frequency or presence. These features have been widely and successfully used in classification. One of the methods of converting text into features is the bag-of-words approach, where each word is an element of the feature set, and each document is represented with a set of real numbers where each element of the set represents the frequency of that word in the document. N-gram model can be formally defined as:

Let $f1 \ldots fm$ be a predefined set of m features that can appear in a document; examples include the word "still" or the bigram "really stinks". Let $ni(d)$ be the number of times fi occurs in document d. Then, each document d is represented by the document vector

d: = (n1 (d) n2 (d)....nm (d)) [8] .

In bag of words model, a text is represented as an unordered collection of words, disregarding grammar and even word order, which causes significant data to be lost in the processing of the input. To decrease the data loss, N-gram model is used, where rather than taking each word as a feature, each sub-sequence of length n is set as a feature. This model, while increasing the complexity of the process, cannot account for long dependencies but preserves word order within N words. An N-gram of size 1 is referred to as a "unigram"; size 2 is a "bigram"; size 3 is a "trigram"; and size 4 or more is simply called an "N-gram".

An N-gram set of an example sentence is given below.

"I did not like the book."

Unigrams:
["I", "did", "not", "like", "the", "book"]

Bigrams:
["I did", "not like", "the book"]

Trigrams:
["I did not", "like the book"]

Classification techniques use this features as input and train itself using training data set. As seen from simulation study, SVM gives best accuracy among all classifiers so in this component we will use SVM as main algorithm for opinion spam classification. We also wants to check the accuracy of other supervised techniques and performs a comparative analysis about the advantages and disadvantages of each techniques in opinion spam classification.

9. CONCLUSION AND FUTURE WORK

Finding the opinion spam from huge amount of unstructured data has become an important research problem. Now business organizations and academics are putting forward their efforts to find the best system for opinion spam analysis. Although, some of the algorithms have been used in opinion spam analysis gives good results, but still no algorithm can resolve all the challenges.More future work is needed on further improving the performance of the opinion spam analysis. There is a huge need in the industry for such applications because every company wants to know how consumers really feel about their products and services and those of their competitors by analyzing true reviews not spam reviews.This research proposes an opinion spam analyzer which automatically classifies input text data into either spam or non spam category. The proposed system will use machine learning supervised technique. The chosen algorithm based on simulation work is Support Vector Machine (SVM).

A direction for future research is to implement the system and check performance by applying proposed approach to various benchmark datasets. Comparing performance of different classification methods to find the best one for our proposed opinion spam classification method could be another future research direction.

10. REFERENCES:

[1]Nitin Jindal, Bing Liu, Review spam detection, Proceedings of the16th international conference on World Wide Web, May08-12, 2007, Banff, Alberta, Canada.

[2]Nitin Jindal, Bing Liu, Analyzing and DetectingReviewSpam,ICDM2007.

[3]Nitin Jindal, BingLiu, Opinion spam and analysis, Proceedings of the international conference onWebsearchandwebdatamining,February11-12,2008,PaloAlto, California, USA.

[4]K.Sharma and K. Lin, “Review Spam Detector with RatingConsistencyCheck,”2013.

[5]Y. Lin, T.Zhu, X. Wang, J. Zhang, A. Zhou,“ Towards Online Review Spam Detection,”WWW'14 Companion,, April 7-11, 2014pp. 341-342, Apr. 2014.

[6]Y. Ma and F.Li, “Detecting Review Spam: Challenges and Opportunities,” Proc. 8thIEEEInt.Conf. Collab .Comput. Networking, Appl. Work., 2012.

[7]MyleOtt, Yejin Choi, ClaireCardie, JeffreyT.Hancock ,Finding deceptive opinion spam by anystretch ofthe imagination, Proceedings ofthe 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, June 19-24, 2011, Portland, Oregon.

[8]BingLiu, Sentiment analysis and subjectivity, Handbook of Natural Language Processing, second edition, 2010.

[9]Arjun, M. and Liu,B.(2012)Modeling Review Comments. In Proceedings of the50th Annual Meeting of Association for ComputationalLinguisticsACL-2012.

[10]Y. Ma and F.Li, “Detecting Review Spam: Challenges andOpportunities,”Proceedingsof8th IEEEInternational Conference Collaborate com, Pittsburgh, U.S. pp. 640-643, 2012.

[11]Bramer, M. (2007)Principles of Data Mining. Springer-Verlag,London.(ISBN 1-84628-765 0).

[12]Dunham, M. (2003)Data Mining :Introductory and Advanced Topics. 1st Edition, Pearson Education.

[13]Han,J., and Kamber, M.(2006)Data Mining: Concepts and Techniques. 2nd Edition. Morgan Kaufmann Publishers, San Francisco, USA.(ISBN1-55860-901-6).

[14]Cios, K.J., Pedrycz, W., and Swiniarski, R.W. (1998)DataMiningMethodsforKnowledgeDiscovery. KluwerAcademicPublishers, Norwell, MA, USA.(ISBN0-7923-8252-8).

[15]Frank, E., and Witten,I.(2005)DataMining:PracticalMachineLearningToolsandTechniques. 2nd Edition, Morgan Kaufmann, San Francisco.

[16] Arjun, M. and Liu,B. (2012)Modeling Review Comments. In Proceedings of the50th Annual Meeting of Association for Computational Linguistics(ACL-2012)(Accepted for Publication).

[17]Arjun, M., Glance, N. and Liu, B.(2012)Spotting Fake Reviewer Group sin Consumer Reviews. In Proceedings of International World Web Conference (WWW-2012).

[18] Ahmad S. J. Abu Hammad, “An approach for Detecting Spam in Arabic Opinion Reviews,” Master’s Thesis, Islamic University of Gaza, January 2013.