A study on detecting Opinion Spam Concerning the Issues, Challenges and Opportunities.

[1]Dr.R.Mahalakshmi

Assistant Professor
Department of Computer Science and Engineering
Presidency University,Bengalure, India

***Abstract**:* With the impulsive growth of social media on the Web, individuals and organizations are more and more using public opinions for their decision making. The opinionated text retrieved from protracted forum postings, discussion rooms, chats and blogs. It is difficult for the average human reader identifying relevant posts and summarizing the information accurately. Moreover, it is human pschycology that people often pay superior attention to opinions that are harmonizing with their own preferences. As there are ample of sites available, extracting the factual information is a challenging task, rather than, finding relevant information on the Web. Keeping in the mind, the easy accessibility of the reviews and the significant impacts in the digital market, there is an increasing support to manipulate the reviews, mostly profit driven. In near future, spam reviews will damage the entire online review systems and finally cause a gradual loss of credibility. Hence, the first step towards securing the online review system is detecting the spam reviews.There are no clear hints or signs from the text of the reviews that it is real or fake. Web Mining and Machine learning techniques both put forward an inspiring contribution to detecting fraudulent reviews The key objective of this paper is to provide a strong and comprehensive analysis on current research and detecting review spam machine learning techniques.

The organization of the paper is as follows, section 1 gives the introduction of Opinion mining and review spam detection, section 2 briefs the related work presented in the literature, section 3 explains the representative techniques in identifying, summarizing the various opinionated texts and categorizing the same as true or spam review. Section 4 details the issues of detecting spam or fake online reviews, and section 5 gives the conclusion.

***Key words* - Web mining; Review spam; Opinion mining; Machine learning; Clustering**

I. Introduction

Now a day's most of the customers do online shopping, naturally rely on online reviews, gives rise for fake reviews. Wrongdoers may create false reviews either to artificially promote or devalue products and services. This practice is known as Opinion/Review Spam. Since not all online reviews are truthful and trustworthy, it is important to develop techniques for detecting review spam. WEB 2.0 and WEB 3.0 has provided many open ended platforms to share the views and feeling on anything and everything in the world. Actually every individual enjoy sharing their views on the open platform with hidden identity.

In general, opinions can be expressed about anything, e.g., a product, a service, an individual, an organization, an event, or a topic, by any person or organization. The public opinions are extracted to obtain subjective and factual information. Thus, it is the process of categorising hidden information about user's intensions, likeliness and taste. This will in turn help the traders and customers to make their decisions rightly. Review spam is intended to damage the reputations of a product or the manufacturer either directly or indirectly.

In today's world, Fake reviews are appealing as common problem on the Web. It is very important for the Traders and customers to distinguish truthful reviews from fake ones. It is really a challenging task as there are no distinctive words or indications available in the text corpus.

Web Mining and Machine learning techniques offer an amazing contribution to detecting falsified reviews. Web mining is to discover and extract intelligent information from Web documents and services automatically. Web mining could be classified into three types, such as structure, content and usage mining. Content mining is concerned with knowledge and information extraction, and categorizing entities using data mining and machine learning approaches. A direct example of content mining is opinion mining. Opinion mining consists of attempting to ascertain the sentiment of a text passage by analysing the features of that passage. Sentiment is polarity of two types Positive and Negative

It is a very vital information that the existing machine learning techniques are not adequate for efficient review spam detection, remains more trustworthy than manual approach. Crawford et al.[1] had done an extensive survey on machine learning techniques that have been proposed in the literature for the detection of online review spam. The author also highlighted the importance of on feature engineering and the impact of those features on the enactment of the spam detectors. Further, the authors had investigated the merits of supervised, unsupervised and semi-supervised learning methods and a comparative analysis results were presented in the paper. A widespread investigation was done on various aspects of the review spam detection, given further direction for future research.

A classifier can be trained to classify new instances by analysing the text features associated with multitude of opinions with respect to their sentiments. Review spam detection lies in the category of content mining, but also utilizes features not directly linked to the content [6]. Constructing features to describe the text of the review involves text mining and Natural Language Processing (NLP). Additionally, there may be features associated with the review's writer, its post date/time and how the review deviates from other reviews for the same product or service.

II. Related Work

These user generated contents to express personal opinions about objects are known as reviews, which includes a rating with or without narrative comments. Reviews about goods, books, movies, news, services, etc. are deliberated as product reviews [3 ,5], and reviews that express overall attitudes of organizations or stores are classified as store reviews [6,8,10]. Reviews usually come with ratings. Respectively, there exists two types of review spams, inserting dishonest ratings or inserting unreliable comments on product or stores. Detecting unfair or fraudulent ratings has been studied in several works including [5, 17]. Different behaviours of the spammers have been observed when inserting fraudulent reviews into two types of review systems[11].spamming by identifying singleton reviews on the reviewed items[17]. Singleton reviews are the reviews written by users who contribute only one review each.

Opinion holders are usually the authors of the postings, [5, 13]. Opinion holders are also called opinion sources [4]. Opinions could be classified in to two main types, such as regular opinions and comparative opinions. The former is often referred to the general opinions reflect some sensitivity analysis. The later expresses few resemblances or differences between two or more entities, with or without a preference of the opinion holder. [4]. A comparative opinion is typically written using the adjectives or adverbs either in comparative or superlative form. The terms regular opinion and opinion is used interchangeably in the literature. It [5] is simply a positive or negative sentiment, attitude, excitement or assessment about an entity or an feature of the entity from an opinion holder point of view. Positive, negative and neutral are called opinion orientations also known as sentiment orientations, semantic orientations, or polarities. There are two more significant concepts related to opinion mining, such as, subjectivity and objectivity. Expressing some factual information about the world is objectivity, while expressing some personal feelings, views or beliefs is subjectivity. The score [15] function and [20], feature weighting schemes are used to enhance classification accuracy.

As stated earlier detecting review spam is a challenging task, since no one knows the amount of spam in existence accurately. Due to the openness of the social media, spammers can post spammed reviews by hiding their original identification. Which makes the task harder to eliminate completely. Spam reviews typically look as normal as original. In order to identify the fake ones the users have to make the additional comparisons, results in tedious and non-trivial. One approach taken by review site such as Amazon.com is to allow users to label or vote the reviews as helpful or not.

Naive Bayesian classification, and support vector machines (SVM) are most commonly used machine learning techniques. Pang et al. [13] and many others took this approach to classify reviews into two classes, positive and negative. Most of the authors used unigrams, n grams or a bag of individual words as features in classification and stated that it performed well with either naive Bayesian or SVM. The TF-IDF weighting scheme from information retrieval, POS, Opinion words and phrases may also be applied too. Thus, adjectives and opinion phrases are important indicators, have been commonly used to express positive or negative sentiments. For example, attractive, delightful, superb, Hats off, wonderful, and marvellous are positive opinion words, and bad, poor, awful rubbish, junk, and crap, and hate are negative opinion words[].

Many researchers have worked towards the product review spam detection, content based filtering or behaviour based filtering. In [10]. Similarity based methods were adopted to identify near-duplicate content, which was considered as spam reviews and used as labelled training data to build the discriminator for 2-class classification. In [14] the reviews on only brand and bon-reviews [reviews with no features to classify further] have been considered as spam in [14] and automatically labelled out to construct the classifier for supervised learning.

The unexpected pattern in rating distribution [16,4,3] , strange behaviours , an unexpected rapid boosting or downgrading change in ratings and a suspicious time interval were studied by the authors and used rule mining method to detect the anomaly. Also discussed [16,12,15]the inter-relationships between product, groups, and group members, described the architecture for blogs to get the summary of the text, highlighted the Appraisal Theory, whereas classified the opinions into three categories as Affect, Appreciation or Judgment. Cambria et al., 2011[21] had done a research based on a common sense and emotion representation, concluded to have more research in this zone. It have been used for short texts to infer emotional states over the web [21]. Yi et.al. [20], developed the concepts called e-sentiment analyzer for online text documents.

III. Methodology and Techniques

Let us assume that the review document is represented by D, which has an opinion on a single entity[e].opinion authors are represented by [r]. all document-level sentiment classification are based on supervised learning, although there are also some unsupervised methods. Each review has a reviewer assigned rating [1-5 stars].we can group the reviews based on their ratings 4-5 one group Positive, 3 as a separate group Neutral and 1-2 Negative.

To achieve this objective, one needs to perform the following tasks:

Task 1: Review Extraction and Grouping:

Extract all review expressions in the set D, and group identical expressions into review clusters. Each expression cluster indicates a unique entity r_1.

Task 2: Feature Extraction and Grouping:

Extract all features of the reviews, and group them into clusters. Each cluster of entity r_i indicates a unique feature F_{ij} .

Task 3: Opinion Holder and Time extraction:

Excerpt the source of information from the text or unstructured data.

Task 4: Feature sentiment classification: D
Classify the each opinion based on features f_{ij} - as positive, negative or neutral.

The following are the popular approaches used for sentiment analysis are:

• Subjective lexicon – is a list of words where each word is assigned a score that indicates nature of word in terms of positive, negative or objective and tried to identify subjectivity of adjectives using Word Net.

• Using N-Gram modelling- for a given set of training data, make a N-Gram model (uni-gram, bigram, tri-gram or combination for classification.

• Machine learning – the supervised or semi- supervised learning is used for extracting the features from the text and learn the model.

Some algorithms like Support Vector Machine (SVM), Conditional Random Field (CRF) have been used for clustering of opinions of same type [6].Subjectivity means the text contains opinion and objectivity means text contains no opinion but contains some fact. In precise form, Subjectivity can be explained as the Topical Relevant. Sentiment Analysis explained till now is not sufficient to satisfy the needs of end user, because the latter is not interested in binary output in terms of positive or negative but interested in aspectual sentiment classification.

Most of the current research has focused on supervised learning methods, which require labelled data for training the classifiers. With millions of online reviews, being generated more and more every day, it is very difficult to label them manually. The unsupervised techniques may yield better results for the above problem by cluster and label the reviews efficiently. To date, not found much papers that study the effects of unsupervised learning for review spam detection.

.

IV. ISSUES ,CHALLENGES AND OPPORTUNITIES,:

The acquisition of unlabelled raw reviews from online review systems is not a challenging job, making it labelled data is an exorbitant affair. In addition to that the quality of training data plays a critical role in developing accurate classifiers. A common method used in most of the existing approaches is engaging one or multiple skilled human agents.

Another challenging task is to identify a list of spamming signs and indicators. Strategies are designed with the facts extracted from the consumer studies and explicit domain knowledge for distinguishing genuine and spamming reviews. There are few assumptions in the literature and researches:

- Genuine author will not write multiple review texts with ratings and
- Smaller reviews are more genuine than larger texts.

Both need to be revisited. Sometimes the vice versa may also be true so as to enable the content rich in clarity .For example extreme ratings with empty review will give an indication for a fake review. But if that is improperly abstracted then that will add a bias to the labelled data [16]. It has to be studied further.
Amazon Mechanical Turk was first used in [5] to collect arbitrary reviews for a set of target stores.

The training time and accuracy of the algorithm can sometimes be quite sensitive to getting just the right settings. Typically, algorithms with large numbers parameters require the most trial and error to find a good combination. Having many parameters typically indicates that an algorithm has greater flexibility. It can often achieve very good accuracy. Provided with the right combination of parameter settings.

Few cases the number of features are very large compared to the number of data points. This is often the case with genetics or textual data. This may take training time unfeasibly long. Support Vector Machines are particularly well suited to this case.Some learning algorithms make particular assumptions about the structure of the data ,features or the desired results, If one that fits the needs, it can provide more useful results, accurate predictions, or faster training times.

The review spam affects the sales performance of the victim products, spread rumours especially in review sites that will affect others buying and selling the products. Further consequences are the existing customers could be disappointed by purchased products or misinterpreting the good products. It is thus an important task to detect review spam and remove them so as to safeguard the genuine interests of the consumers and product vendors.

The general issue is what to extract from the review. The holder of the review is not an expert writer, since there are several opinion in a single review. Hence each sentence carries different meaning on the same issue. In the first Sentence express some positive opinions on overall view, while other two sentences may express negative opinions or emotions with respect to different aspects. It may also be noticed that all opinions have some targets. The target of the Opinion also vary as a whole or part of the entity in a sentence, In all the cases the author of the blog or review has some personal association. It may be from his past experience with the product or some sentiment inclination. Some people will give like because of his favourite star is recommending the product in the advertisement. Hence the reviewer's personal association will not be taken into consideration. It may not be always true that all products of the particular company is always good. They may be good in one product, it does not hold good for other products.

It is really a challenging task to identify the psychological reason for authors to give positive or negative opinion. We need to follow them in their social media, and find out whether he/ she is a die heart fan of the actor come in the product advertisement. As far as my knowledge is concerned no such research is carried out in this aspects. Many papers had been published with respect to their behavioural pattern. Why do they like or dislike the particular product and giving abnormal comments or ratings? This needs some psychological analysis along with usual machine learning and mining techniques.

A practical challenge along this direction is that the dataset could be extremely sparse since each reviewer only reviews limited number of products. The limitation of the dataset could be addressed by actively integrating the review social networks with common social networks

Opinion words and phrases are instrumental to sentiment analysis for obvious reasons.
Negations: Clearly negation words are important because their appearances often change the opinion orientation. For example, the sentence \I don't like this camera" is negative. However, negation words must be handled with care because not all occurrences of
such words mean negation. For example, \not" in \not only but also" does not change the orientation direction Word dependency based features generated from parsing or dependency trees are also tried by several researchers.Instead of using a standard machine learning method, researchers have also proposed several custom techniques specifically for sentiment classification.

As the nature of the review are subjective ,boundary between personalized review, poor-qualitied review and fake review is unclear, It is a principal challenge for all patterns based on supervised learning. From the overview of the existing review spam detection approaches, that most schemes highly depend on the modelling of suspicious spammer behaviours. Therefore, the behaviour of suspicious spammers need to be studied carefully to construct more accurate indicators, which is again an another area of research needs some attention.

The association between a review and a purchase making will increase the integrity of the product store, in other words it may shake the credibility by fake review. Hence the major online shopping portals started working towards the same. Amazon provides verified purchase review service to authenticate the review which was made after the purchase. By attaching Certified review tag on all reviews by verifying their recent visit to the store and the actions taken by the users. This will help the user to evaluate the credibility of the reviews. Hence this could be incorporated into the learning model. It doesn't mean that reviews without tags are fake reviews. Sometimes the opponents give multiple negative comments, do not mind spending a reasonable amount to make it verified review, but will affect the system. This is again a challenge to the So careful study on this is required.

Another source for verification is location. Though no major research has been observed from this aspect, it is believed that the location information should also be utilized to assess the credibility of reviews submitted on-site, within geographical proximity of the store being reviewed. Since the volume of mobile submitted reviews is considerably smaller than Web submitted reviews.this could be tackled by adopting semi-supervised learning model.

V CONCLUSION

This paper discusses the issues, challenges, and research opportunities of online review spam detection. The overview of existing detection methods based on supervised machine learning and data mining is described. Most of the researchers stated that SVM gives good accuracy compare to other classification techniques, but still it has some limitations. More and more future work needed. The prime focus in this research, is to address few most challenging issues in review spam detection to be useful for researchers and facilitate them later.

A few challenges on reliable data acquisition and behavior correlation have been analysed along with a discussion on the new opportunities from integrating current review systems.Many different types of techniques are combined to overcome individual's limitations and benefit from each other's merit and measure the performance of classification technique. large number of studies and research have helped monitor the trending new research increasing year by year.

It is computationally very expensive and possibly impractical with large corpus of reviews to train and training the classifiers.Apart from this eature selection techniques have received little attention.. Further work needs to be conducted to establish how many features are required and what types of features are the most beneficial. Feature selection should not be considered optional when training a classifier in a big data domain with potential for high feature dimensionality.derive an aggregated behavior scoring methods to rank reviewers according to the degree they demonstrate spamming behaviors

REFERENCES

[1]. Crawford, M., Khoshgoftaar, T.M., Prusa, J.D. et al. Journal of Big Data (2015) 2: 23. https://doi.org/10.1186/s40537-015-0029-9

[2]. Read J, carol J.," Weakly supervised techniques for domain independent sentiment classification", In: Proceeding of the 1st international CIKM workshop on topic-sentiment analysis for mass opinion; P.45-52, 2009.

[3]. Zhang Wenhao, HuaXu, Wan Wei.," Weakness finder: find product weakness from Chinese reviews by using aspects based sentiment analysis", Expert SystAppl;39:10283-91,2012.

[4]. Y. Yang, X. Liu, "A re-examination of text categorization methods", in: Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), ACM New York, NY, USA, pp.42-49,1999.

[5]. M. Annett, G. Kondrak, "A comparison of sentiment analysis techniques: Polarizing movie Blogs", In Canadian Conference on AI, pp. 25-35, 2008.

[6]. Maks Isa, VossenPiek. “A lexicon model for deep sentiment analysis and opinion mining applications.”,Decis Support Syst;53:680-8,2012.
[7]. N. Jindal and B. Liu. Review spam detection. In WWW (poster), 2007. [10] N. Jindal and B. Liu. Opinion spam and analysis. In WSDM, 2008.
[8]. N. Jindal, B. Liu, and E.-P. Lim. Finding unusual review patterns using unexpected rules. In CIKM, 2010.
[9]. Qian T, Liu B (2013) Identifying Multiple User ids of the Same Author. In: EMNLP., pp 1124–1135.
[9]. Ellen Riloff and JanyceWiebe, “Learning extraction patterns for subjective expressions”, Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2003.
[10]. 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.
[11]. A. Mukherjee, B. Liu, and N. Glance, "Spotting fake reviewer groups in consumer reviews," in Proceedings of the 21st international conference on World Wide Web.
[12]. Alexander Pak and Patrick Paroubek, “Twitter as a corpus for Sentiment Analysis and Opinion Mining”, Proceedings of the Seventh conference on International Language Resources and Evaluation, pp. 1320-1326, 2010. [2]
[13]. Bo Pang, Lillian Lee and Vaithyananthan, “Thumbs up? Sentiment Classification using Machine Learning Techniques”, Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, Vol. 10, pp. 79-86, 2002.
[14]. Thorsten Joachims. “Text categorization with support vector machines: Learning with many relevant features”, Proceedings of the European Conference on Machine Learning, pp. 137–142, 1998.
[15]. A Hogenboom et al., Lexicon-based sentient analysis by mapping conveyed sentiment to intended sentiment. Int. J. Web Eng. Technol. 9, 125–147 (2014)
[16]. Michael Gamon, “Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis”, Proceedings of the 20th International Conference on Computational Linguistics pp. 841-847, 2004.
[17]. Hsinchun Chen and David Zimbra, “AI and Opinion Mining”, IEEE Intelligent Systems, Vol. 25, No. 3, pp. 7480, 2010.
[18]. Hsinchun Chen, “AI and Opinion Mining, Part 2”, IEEE Intelligent Systems, pp. 72-79, 2010.
[19].Arti Buche, Dr. M. B. Chandak, Akshay Zadgaonkar, “Opinion Mining and Analysis: A survey” arXiv:1307.3336.
[20]. Yang et al. “A web sentiment analysis method on fuzzyclustering for mobile social media users” EURASIP Journal on Wireless Communications and Networking DOI 10.1186/s13638-016-0626-0.
[21]. Erik Cambria Amir Hussain “A Common-Sense-Based Framework for Concept-Level Sentiment Analysis” http://www.springer.com/series/13199.