

# A Survey to Detect Deceptive Reviews

Viral Patel<sup>1</sup> Prof. Samina Mansuri<sup>2</sup>

<sup>1</sup>Student <sup>2</sup>Assistant Professor

<sup>1,2</sup>Department of Computer Engineering

<sup>1,2</sup>SVBIT, Gandhinagar, Gujarat, India

**Abstract**— In this Globalized world, the rapid changes in the number of e-commerce websites, has made the internet, an extensive source of product reviews. Moreover, opinions reviews are posted on the web are paramount source for both companies as well as customers. Since there is no survey regarding the quality of the review written, anyone can basically write anything which conclusively leads to Review Spams. In this paper we study about the methods used for detect the spam and non-spam reviews.

**Key words:** Opinion Mining, Opinion Spam, Deceptive Opinions, PU Learning

## I. INTRODUCTION

Online product reviews have become a requisite resource for users in order to make their decision while they purchases online products. Product reviews provide information that impacts purchasing decisions to consumers, retailers, and manufacturers. Consumers make use of the reviews for not just word of mouth information about any product, regarding product durability, quality, utility, etc. but also to give their own input regarding their experience to others. The rise in the number of E-commerce sites has led to an increase in resources for gathering reviews of consumers about their product experiences. The product is very likely buy by the users, if the opinions are mostly positives and Users choose another product, if the opinions are mostly negative. Moreover, financial gains or fames for organizations and individuals are achieved by giving Positive opinions for products. As internet has no limits anyone can write anything and get away with it, expand in the number of Review Spam.[1]These reviews are produced by the people who do not have personal experience on the subjects/product of the reviews are called fake or deceptive reviews.

Opinion spam review refers to "illegal" activities that try to misdirect readers by giving undeserving favourable opinions to some target entities with a view to promoting the entities and/or by giving false/Negative opinions to some other entities in order to damage their reputations[14]. Sometimes they just write fake reviews in order to convince users that they have to buy from particular brand/store or not [2]. Opinion spam has different forms, e.g., fake reviews, fake comments, fake blogs, fake social network postings and deceptive messages. In the last few years, Review Spam Detection has gathered a lot of attention. Over the past few years, consumer review sites like Yelp.com have been removing spurious reviews from their website using their own algorithms.

There are two distinct types of deceptive review spams: 1) Hyper spam, in which fake positive reviews are rewarded to products to promote them. 2) Defaming spam, where unreasonable negative reviews are given to the competing products to harm their reputations among the consumers [3].

The benefit of a brand or a product reviews that have been written, expressing a positive sentiment for a product are called positive deceptive reviews spam. As opposed to that, reviews that intend to defame a competing product expressing a negative sentiment towards the product are called negative deceptive review spams.

Spam refers to irrelevant posting, content that is out of place. Opinion mining refers as Human activities that try to deliberately mislead readers or automated opinion mining systems by writing positive or negative review on particular product. In general any user can give fake review to gain their benefit is called as opinion spam.[14]

## II. TYPES OF SPAM REVIEW

There are mainly three types of spam review identified:[10]

### A. Type 1 (untruthful opinions)

Those that deliberately misdirect readers. Again there are two type:

- Positive spam review: undeserving conclusion to an item to promote them.
- Negative spam review: give negative review to damage their popularity.

### B. Type 2 (reviews on brand only)

Those that comment on the brand instead of commenting on the products.

### C. Type 3 (non-reviews)

Those that are not reviews.

It has two main sub-types: (1) advertisements and (2) other irrelevant reviews containing no opinions.

## III. SPAM DETECTION METHODS

The main goal of opinion spam detection is to identify each and every fake review or reviewer. So there are mainly three methods to detect spam reviews are supervised detection method, unsupervised detection method, semi-supervised learning method.[14]

### A. Supervised detection methods

It is used to detect spam review by classification problem of separating reviews into two classes as spam review and non-spam review. The supervised learning methods depend on the existence of labelled training data.

### B. Unsupervised detection methods

It is difficult to accurately producing labeled datasets, so supervised detection method is not every time works and not always applicable. So one of solution is to used unsupervised detection method. It doesn't require labeled data.

### C. Semi-supervised detection method

It has been found that using unlabeled data in conjunction with a small amount of labeled data can considerably improve learner accuracy compare to other method.

#### IV. RELATED RESEARCH WORK

Nitin Jindal and Bing Liu find study on categorization of spam reviews and basic of opinion spam detection. Duplicates and near-duplicates can be detected using the shingle method. Describe main three types of spam reviews are false opinion, Review on brands only, Non-reviews and also describe spam detection strategy. Duplicates and near-duplicates can be detected using the shingle method. [10]

M.N. Istiaq Ahsan , Tamzid Nahian , Abdullah All Kafi, Md. Ismail Hossain, Faisal Muhammad Sha Introduce an ensemble learning approach which combines two different types of learning methods active and supervised.[11]

Donato Hern, andez Fusilier , Manuel Montes-y-Gomez, Paolo Rosso, Rafael Guzm an Cabrera focused on Detection of deceptive opinion spam using PU-learning. It is partially Supervised classification technique.[7]

Donato Hern, andez Fusilier , Manuel Montes-y-Gomez, Paolo Rosso, Rafael Guzm an Cabrera, Detecting positive and negative deceptive opinion using PU-learning Propose a novel method of PU learning that with respect to its original version is much more conservative at the moment of selecting the negative example.[12]

Fang Lu, Qingyuan Bai proposed Semi-supervised Text Categorization with Only a Few Positive and Unlabeled Documents in which not only extracts more reliable negative examples from U but also extracts most reliable positive examples from U to increase the size of P. Two-step algorithm, in Step 1, propose a more effective extraction method of purifying the initial negative set N and expanding the original small positive set P. Then in step 2, adopt Naïve Bayes which is a popular text classifier. [9]

Huayi Li, Zhiyuan Chen , Bing Liu, Xiaokai Wei Spotting Fake Reviews via Collective Positive-Unlabeled Learning In which Proposed a supervised algorithm MHCC (Multi-typed Heterogeneous Collective Classification) for the heterogeneous network of review, users and IPs and then extended it to CPU model which is more appropriate for PU learning problem. [6]

Yixing Xu, Chang Xu, Chao Xu, Dacheng Tao Multi-Positive and Unlabeled learning Used Multi-Positive and Unlabeled algorithm for solving the multi-positive and unlabeled learning problem. Proposed method that directly enables multi-class model to be trained using the given input multi-class data and that predicts the label based on the model decision. Construct different convex loss function.[13]

#### V. RELATED WORK

Here in this table displays the survey undertaken for the spam review detection. As shown below the table gives brief description about the different approaches as well as the dataset used.

Paper Title/ Publication	Method Used	Dataset Used
Opinion spam and analysis ACM[10]	Duplicates and near-duplicates can be detected using the shingle method.	Use reviews from amazon.com. in June 2006.extract 5.8 million reviews, 2.14 reviewers and 6.7 million products
An Ensemble approach to detect Review Spam using hybrid Machine Learning Technique. IEEE[2016] [11]	Introduce an ensemble learning approach which combines two different types of learning methods active and supervised.	creating a hybrid dataset of both real-life and pseudo reviews Labeled reviews taken from ott dataset and the unlabeled dataset reviews extracted from
Using PU-Learning to Detect Deceptive Opinion Spam, Association for Computational Linguistics [7]	The PU-learning approach which learns only from a few positive examples and a set of unlabeled data	Dataset of reviews assembled by Ott et al. (2011).
Detecting positive and negative deceptive opinion using PU learning. Science direct[2014] [12].	Propose a novel method of PU learning that with respect to its original version is much more conservative at the moment of selecting the negative example.	Dataset of reviews assembled by Ott et al (2013).
Semi-supervised text categorization with only a few positive and unlabelled documents. IEEE [9]	Propose a refined method to do the PU-Learning with the known technique combining Racchio and k mean algorithm.	Use TanCorp-12 dataset of Tan-Corp corpus which is a popular data set for Chinese text classification experiment
Spotting Fake Reviews via Collective Positive-Unlabeled Learning. Dianping[6]	Propose a collective classification algorithm called multi-typed Heterogeneous collective classification(MHC C)and then extend it to collective	Reviews from 500 restaurant in shanghai, china

	positive and unlabeled learning(CPU).	
Multi-Positive and Unlabeled learning. IJCAI[2017][13]	Used Multi-Positive and Unlabeled algorithm for solving the multi-positive and unlabeled learning problem.	Use four different dataset (Image Segment, Letter, USPS, MNIST) and the relevant metadata for each dataset.

Table 1: Related Work

## VI. CONCLUSION

Opinion spam detection is very broad field and it is still growing. Research on which we are working is present in the market but we can achieve optimize and flexible solution. In this paper, some introduction about opinion mining and opinion spam. Supervised spam detection method and unsupervised spam detection method are used to detect spam review. Related work is described and also table 1 gives brief description about the different approaches as well as the dataset used. Basically some basic information about opinion spam and spam detection are given in this paper.

## REFERENCES

- [1] Kyung-Hyan Yoo and Ulrike Gretzel. Comparison of deceptive and truthful travel reviews. Information and communication technologies in tourism 2009, pages 37{47, 2009.
- [2] Qingxi Peng and Ming Zhong. Detecting spam review through sentiment analysis. Journal of Software, 9(8):2065{2072, 2014.
- [3] C Harris. Detecting deceptive opinion spam using human computation. In Workshops at AAAI on Artificial Intelligence, 2012.
- [4] Myle Ott, Claire Cardie, and Jeffrey T Hancock. Negative deceptive opinion spam. In HLT-NAACL, pages 497{501, 2013.
- [5] Siddu P. Algur, Amit P. Patil, P.S Hiremath, S. Shivashankar, Conceptual level Similarity Measure based Review Spam Detection, karnatak, India, copy right 2010 IEEE.
- [6] Huayi Li, Zhiyuan Chen, Bing Liu, Xiaokai Wei and Jidong Shao, Spotting Fake Reviews via Collective Positive-Unlabeled Learning, IL, Dianping.
- [7] Donato Hernandez Fusilier, Manuel Monstes-y-Gomez, Paolo Rosso ,Using PU-Learning to Detect Deceptive Opinion Spam, Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, page 38-45, Atlanta, Georgie, 14 June 2013, @2013. Association for Computational Linguistics.
- [8] Rinki Patel , Priyank Thakkar, Opinion Spam Detection Using Feature Selection, Copy right 2014 IEEE.
- [9] Fang Lu, Qingyuan Bai, Semi-supervised Text Categorization with only a Few Positive and Unlabeled Document , Fuzhou, China, copy right 2010 IEEE.
- [10] Nitin Jindal and Bing Liu, Opinion Spam and Analysis, palo Alto, USA Copy right 2008 ACM.
- [11] M.N. Istiaq Ahsan, Tamzid Nahian, Abdullah All Kafi, Md. Ismail Hossain, Faisal Muhammad Shah, An Ensemble approach to detect review Spam using hybrid Machine learning technique, Bangladesh ,copy right 2016 IEEE.
- [12] Donato Hernandez Fusilier, Manuel Monstes-y-Gomez, Paolo Rosso, Detecting positive and negative deceptive opinions using PU-learning, Information Processing and Management, 2014.
- [13] Yixing Xu, Chang Xu, Chao Xu, Dacheng Tao, Multi-Positive and Unlabeled learning, proceeding of the twenty-sixth International Joint Conference on Artificial Intelligence (IJCAI), 2017.
- [14] Pooja H. Ghelani, Tosal M. Bhalodia, Opinion Mining and Opinion Spam Detection, International Research Journal of Engineering and Technology (IRJET), 2017.
- [15] <https://www.cs.uic.edu/~liub/FBS/fake-reviews.html>