ISSN: 0976-3031

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 9, Issue, 6(C), pp. 27360-27363, June, 2018

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

DETECTING FAKE REVIEW- AN OVERVIEW

Nishi Tiku., Neha Menon and Vineeth Pillai

Department of MCA, VESIT

DOI: http://dx.doi.org/10.24327/ijrsr.2018.0906.2242

ARTICLE INFO

Article History:

Received 11th March, 2018
Received in revised form 6th April, 2018
Accepted 26th May, 2018
Published online 28th June, 2018

Key Words:

Review spam, outliner review, review Spamicity, opinion spam, and semantic review.

ABSTRACT

E-commerce is buying or selling of products using online services or over the internet. A lot of people have started preferring e-commerce over the traditional method of shopping because of the availability of variety of products all the while enjoying the comforts of one's home. Product Reviews and feedbacks have changed the game for online market since internet has become a very household thing. Sellers too have started looking at e-commerce as a way to increase their customer base. The Product Reviews are the factors which either make or break the relationship of the consumer with the store – they help build loyalty and trust and lets the potential consumer know the product much more clearly and the aspects that differentiate it from the rest of the products elsewhere [1]. This understanding among the merchants have opened a whole new Pandora's Box filled with fake reviews, baseless shaming etc. One cannot undermine the importance of product review on the success and sales of a new product and thus it is imperative that e-commerce sites identify these fake reviews and deal with them appropriately. This paper focusses on multiple fake review detection algorithms and each of their pros and cons.

Copyright © Nishi Tiku., Neha Menon and Vineeth Pillai, 2018, this is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

E-commerce has gained in popularity over the recent years. With the increasingly busy schedule of the working class, not having to waste valuable time in rummaging about for things they need is a boon. A well-placed search and a few clicks here and there will get them the products they yearn, from brands they prefer and in rates that is feasible to them. This upsurge in online buying and selling has given rise to a concept of reviews wherein customers of certain products or services leave their opinions about that product or service on the sites so that other viewers or potential customers know what to expect. These reviews can be positive or negative based on the buyers' experience. The human nature is such that, we require validation for everything that we do. In this case, these reviews have the potential to color the new buyer's opinion about a product even before buying it. This has in turn lead to the rise in fake reviews sponsored by online providers either to promote their own product or to bad mouth their competitors' products. There are a lot of fake reviews out there, ready to write untrue comments at the drop of a hat. These are fraudulent comments that influence potential buyers that even has the ability to either make or break a business. Thus, weeding out these comments is imperative to give visitors of a site true reviews of products or services they offer. There are various algorithms developed to detect fake reviews, some giving better accuracy than the others.

LITERATURE REVIEW

Analyzing and Detecting Review Spam

Jindal and Liu first categorized spam reviews into 3 categories:

Type 1: Intentionally written positive or negative reviews
Type 2: (Reviews on brand): Reviews that are on brand due to pre-conceived notification thus mostly highly biased. Intention is to promote the brand.
Type 3: (Non-reviews): Does not serve any purpose. Commenting on others reviews, advertisements, reviews on competitors etc. are some example of these category.

They have identified the type 1 spam by identifying duplicates of 3 types:

1. Duplicates from different user Id on Same Products.
2. Duplicates from Same User on Different product.
3. Duplicates from different user ids on different product.

Type 2 Spam reviews are identified by manually training the model and then using supervised learning algorithms. Logistic regression was used to find the probability that whether a review is spam or not. They considered duplicates from the

Corresponding author:* **Nishi Tiku
Department of MCA, VESIT

same user ids as case of mistake but that may not always be the case.

Issues

1. Outlier Reviews are considered based on the average rating of the product. But an expert spammer may give the average rating in alignment with others while the review content may be positive.
2. Duplicates from same user id on same product must also be taken into consideration.

Changes that can be done

Find a way to also compare the content and the rating given by the user to find deviation among them.

Review Spamicity based on rank and content of the review

They have addressed the Jindal and Liu problem of the outlier reviews. They calculated the rating based on the content of the review and compared it with the actual rating given by the user and if the difference is greater than or equal to 2 then the review is termed as suspicious. They have not considered any other feature for review classification.

Issues

1. Did not take into account that duplicate reviews for different product from same or different users may be spam.
2. Did not try and find whether the review content is related to the product or is a Type 2 or 3 review that has been categorized by Jindal & Liu., which should be fake.
3. User behavior is not taken into consideration.

Changes that can be done

1. Train Data for identifying Type 2 and Type 3 spams.
2. Incorporate methods to identify duplicate reviews on different products
3. Context sensitive detection for reviews (like sarcasm)

Detecting Fake Reviews Utilizing Semantic and Emotion Mode

This paper focusing on detecting spam reviews by taking into account a varied number of features associated with the review, reviewer and user behavior. They used a labelled dataset with fake and non-fake reviews from professional review website. They have taken into considerations user related behavior features to identify expert spammers. The authors have identified the emotion expressed in reviews, similarity between reviews, and category, time and store density for finding fake reviews. It is assumed that mostly spammers spam the same category product reviews. Similar Reviews are identified using cosine similarity. Using word vector, they have identified semantically similar reviews. Reviews expressing extreme emotions are considered as spam.

Table 1 Comparison of Literature Review

Factors	Jindal & Liu Spam Detection Strategy	Review Spam Based on Rank and Content of the Review	Opinion Spam Detection Using Feature Selection	Detecting Fake Reviews using Semantic and Emotion Model
Approach	Identified 3 categories of spam and based on the type strategy for detecting them were used for Type 2 and Type 3 spams Training data was manually labelled and logistic regression model was used to predict whether the review was spam or not. Type 1 spam was identified by detecting duplicates and near duplicates.	Data is pre-processed using tokenization and stop word removal, positive and negative words were identified and stored in a different database with their count. Opinion matrix and rank were created. Rating based on the contents of the review was computed, compared and difference was calculated for computed rating and the rating given by reviewer. As the rating scale will not go > 5, threshold of 2 was chosen so if the difference was >=2, reviews were suspected for Spamicity.	Pre-processing of data was done by converting words to lowercase and stop words removal and the data was represented in 3 different vector representations namely: Boolean, bag-of-words, TFIDF along with unigram, bigram or bigram+ approaches for sequence of words. Then manually training half dataset 2 classification techniques: Naïve Bayes and LS-SVM method was used to classify the results.	In this Model, different features related to user behavior, review information characteristics, user characteristics, content related features, review density features, semantic similarity between different elements as well as the emotion expressed in the review was taken into consideration. Using this data, a training set was prepared and then 3 classification techniques were used namely Naïve Bayes, SVM and decision tree.
Assumption	For Type 2 & 3, they trained the data so no assumption was made. But for type 1 spam, the assumption is that all outlier reviews are spam.	Assumption is that a review cannot have extreme positive or negative emotions, a genuine review will have a balanced emotion ratio.	That Context of the opinion words will not matter.	Product related features have no impact on the spam detection techniques.
Storage requirements	Review Data along with 36 features are stored in a database.	3 databases: Raw review database, Opinion analyzed database and Opinion words database.	3 different vector representations of reviews are required	Reviews along with 18 features are stored in 1 database.
Join computation	No Joins	Need joining of 3 tables	No Joins	No Joins
Processing efficiency	Simple arithmetic calculations and Logistic Regression Modeling is used.	Only Simple Arithmetic Calculations are done	Requires implementation of complex Naïve Bayes and LS-SVM Techniques.	Involves complex Arithmetic calculations along with implementation of classification techniques.
Factors considered	Considers the following features: Reviewer Centric Features Product Centric Features Review Centric Features	Only Content of the Review and rating is taken into consideration	Requires implementation of complex Naïve Bayes and LS-SVM Techniques	User Behavior Diversity Features Reviewer characteristics & Information Feature Content related feature Review Density Features Semantics & Emotions
Training set required	Yes	No	Yes	No
Accuracy	90% - 99%	19.65% -40%	70% - 89%	75% - 93%

Issues

1. Does not identifies non-reviews or reviews on brand only rather than product
2. Does not takes into consideration that similar reviews for similar product with just one or two different features may be genuine rather than fake.

Changes that can be done

1. Incorporate Jindal & Liu Technique of training the dataset to identify Type 2 and Type 3 spams.
2. Identify if reviews are similar then are the products similar too.

Opinion Spam Detection Using Feature Selection

The authors detect spam via 3 approaches: by Boolean representation, bag-of-words and TFIDF. They have considered spam detection as a binary classification problem. They have undertaken different features and then have used Naive Bayes classifier and LS-SVM to train the classifiers and calculated the accuracy of both the approaches.

Issues

Independence among the features are assumed which may not be the case.

Comparison

Working Structure

***Opinion Spamming**:* [3] It refers to "illegal" activities (e.g., writing fake reviews, also called shilling) that try to mislead readers or automated opinion mining and sentiment analysis systems by giving undeserving positive opinions to some target entities in order to promote the entities and/or by giving false negative opinions to some other entities in order to damage their reputations. Opinion spam has many forms, e.g., fake reviews (also called bogus reviews), fake comments, fake blogs, fake social network postings, deceptions, and deceptive messages.

Fake Review Detection: We have used supervised learning, pattern discovery, graph-based methods, and relational modeling to solve the problem. Below are some main signals that we have used:

1. ***Review content:*** Lexical features such as word n-grams, part-of-speech n-grams, and other lexical attributes. Content and style similarity of reviews from different reviewers.
2. ***Reviewer abnormal behaviors:*** Public data available from Web sites, e.g., reviewer id, time of posting, frequency of posting, first reviewers of products, and many more. For example, do you see anything wrong with the reviews from this user, Big John? What about after you see the reviews of these two users, Cletus and Jake? In fact, if you browse the reviews of their reviewed products, you will find another suspicious user/reviewer. This is just one example of atypical behaviors that our algorithms are able to discover.
3. ***Product related features:*** E.g., product description, sales volume, and sales rank.
4. ***Relationships:*** Complex relationships among reviewers, reviews, and entities such as products and stores.

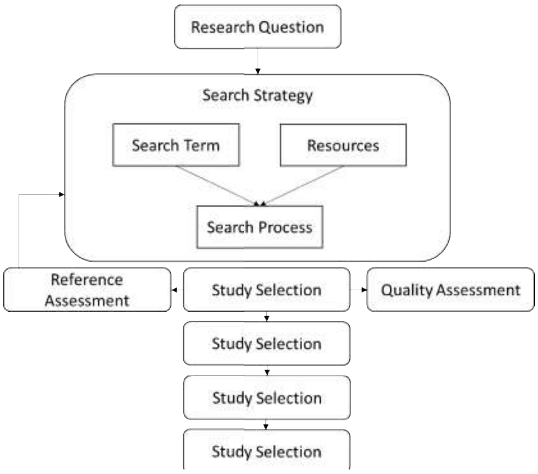

Figure 1 Review flow graph

CONCLUSION

The conclusion is that the three main types of spam were identified. Detection of such spam is done first by detecting duplicate reviews. We then detected type 2 and type 3 spam reviews by using supervised learning with manually labeled. Future work includes collecting abundant review data from other review web sites, computer assisted labeling of reviews to reduce the workload of human experts, more efficient model of detecting the relationship of reviews, reviewers and stores as for detected gaps in literature, our future work will be extracting the most effective and efficient features reported as best in the literature and create a collection of efficient features to be used in future proposed techniques.

References

1. R. Baeza-Yates, C. Castillo & V. Lopez. PageRank increase under different collusion topologies. AIRWeb'05, 2005.
2. Heydari, A., *et al*. Detection of review spam: A survey. Expert Systems with Applications (2014), http://dx.doi.org/10.1016/j.eswa.2014.12.029
3. https://www.cs.uic.edu/~liub/FBS/fake-reviews.html
4. A. Z. Broder. On the resemblance and containment of documents. In Proceedings of Compression and Complexity of Sequences 1997.
5. K. Dave, S. Lawrence & D. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. WWW'2003.
6. I. Fette, N. Sadeh-Koniecpol, A. Tomasic. Learning to Detect Phishing Emails. WWW'2007.
7. Z. Gyongyi and H. Garcia-Molina. Web Spam Taxonomy. Tech. Report, Stanford University, 2004.
8. M. Hu & B. Liu. Mining and summarizing customer reviews. KDD'2004.
9. N. Jindal and B. Liu. Review Spam Detection.WWW '2007. (poster paper)
10. B. Mobasher, R. Burke & J. J Sandvig. Model based collaborative filtering as a defense against profile injection attacks. AAAI'2006.
11. A. Ntoulas, M. Najork, M. Manasse & D. Fetterly. Detecting Spam Web Pages through Content Analysis. WWW'2006.

12. N. Jindal, B. Liu. "Opinion spam and analysis." International Conference on Web Search and Data Mining ACM, 2008, pp. 219--230.
13. A. Mukherjee, A. Kumar, B. Liu, *et al*. “Spotting opinion spammers using behavioral footprints”, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2013:632-640.
14. S. Feng, R. Banerjee, Y. Choi, “Syntactic Stylometry for Deception Detection”, ACL (2011), pp. 171-175.
15. M. Ott, Y. Choi, C. Cardie, J.T. Hancock, “Finding Deceptive Opinion Spam by Any Stretch of the Imagination”, ACL (2011), 309-319.
16. RYK. Lau, SY. Liao, RCW. Kwok, K. Xu, Y. Xia, Y. Li, “Text mining and probabilistic language modeling for online review spam detection”, ACM Trans. on Management Information Systems(TMIS), 2011,2(4):25

How to cite this article:

Nishi Tiku., Neha Menon and Vineeth Pillai.2018, Detecting Fake Review- an Overview. *Int J Recent Sci Res.* 9(6), pp. 27360-27363. DOI: http://dx.doi.org/10.24327/ijrsr.2018.0906.2242

****