iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** III **Month of publication:** March 2022

DOI: https://doi.org/10.22214/ijraset.2022.40743

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Survey on Comprehensive Study of Fake Reviews and Reviewers Detection using machine learning techniques

Mayuri Lahire
Post Graduate Student, Department of Computer Engineering, DYPCOE, Akurdi, Pune, Maharashtra, India- 411035

Abstract: ***Individuals and businesses are increasingly using opinionated social media, such as product evaluations, to make decisions. People, however, try to game the system for profit or fame by opinion spamming (e.g., creating bogus reviews) to promote or demote certain specific items. Such bogus reviews should be identified in order for reviews to reflect real user experiences and opinions.***
Most of the consumers are influenced by the online reviews on the product and it plays a crucial role in finalizing purchase decisions in the market. But fake reviewers or spammers misused and take advantage by writing fake reviews, positive fake reviews to promote the product, or negative fake reviews to demote the product. There has been huge research in this domain for more than a decade for detecting fake reviews or fake reviewers. Howsoever many fake reviewers work together by creating groups to target any product and writing fake reviews on product in bulk, reviewers create multiple fake IDs and write fake reviews.
Detecting false reviews and specific fraudulent reviewers was the subject of previous work on opinion spam. The primary thing of this study is to give a strong and comprehensive relative study for detecting fake reviews and reviewers using machine learning.
Keywords: Fake review, Fake reviewers Spam opinion, Opinion mining, FIM, Reviewer-centric spam, feature engineering

I. INTRODUCTION

As the Internet continues to grow in both size and importance, the quantity and impact of online reviews continually increase. People are influenced by online reviews in broad-spectrum throughout the market but mostly in the world of e-commerce where reviews and regarding products and services are often the most convenient, if not the only, way for a buyer to decide on whether to buy them.

There are various reasons for which online reviews are being generated. Often, in an effort to improve and enhance their businesses, online retailers and service allow distribution, reproduction, and unrestricted use in various mediums, that give exact credit to real authors and the sources, to indicate changes that were created, and to establish a link between licenses. Feedback regarding the experience of products or services used by customers is being asked by the service provider to ensure that customer is happy or not by buying a product.

However, customers can review as per experience that is the service is good or bad, the blind trust in these reviews is dangerous for both service providers and customers.

Before placing an online order many customers look at online reviews, but the reviews may be dangerous as can be fake, written for profit, or gain hence sometimes decisions based on online feedback can harm. Furthermore, service providers are incentives to write good reviews about their products or pay someone to write bad reviews about their competitor's products or services. Fake reviews can have a great impact on the online marketplace due to the importance of reviews.

The studies of this paper cover as follows. At the start of Section II, it highlights basic concepts regarding fake reviews and fake reviewers. Next brief details of the study in this field have been mentioned by various authors. Then Feature Engineering studies give a brief idea of feature engineering in this domain, for reviewer-centric spam detection. Next Reviewer Centric Review Spam Detection literature includes an overview of studies using reviewer-centric features. The Conclusion summarizes the overall findings.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

II. RELATED WORK

Depending upon the literature proposed by Bing Liu al. [17], there are generally three types of fake/spam reviews as identified:

Type 1 (Untruthful Reviews / fake review)	Type 2 (review on brand only)	Type 3 (non-review)
Fake or untruthful reviews are reviews that are written deliberately to mislead readers by giving undeserving positive or negative opinions to some target products in order to promote or by giving false, malicious, and unjust negative reviews to some other products to damage their reputation. This review spam is a challenging task as it is difficult, if not impossible, to distinguish between fake and real reviews by manually reading them. Ott et al. [4] has considered real and fake dataset of type 1.	These reviews only comment on the brands, but not on the specific products that they are supposed to review. Although these may be useful, they are considered spam because they are often biased. For example, to review HP printer, the fake reviewer only writes "I hate HP" or "I will never buy any of HP products".	These are not even considered reviews although they appear as reviews. For example, advertisements or other irrelevant texts like questions, answers, or any random texts.

However, recognizing fake reviewers who are writing fake reviews is more important in the effort to detect review spam. So, let's discuss in detail the type of spammers/fake reviewers. There can be two types as mentioned below:

Individual Fake Reviewer	Group Fake Reviewer
These spammers work individually and not with anyone else, write spam reviews. The spammer may register at a review site as a single user, or as many fake users using different user-ids. He or she can also register at multiple review sites and write spam reviews a. This kind of spammer writes fake reviews by creating multiple accounts & registering multiple times at a site using different user-ids. b. The spammer builds up reputation by reviewing other c. Spammers write either only positive reviews on his/her own products or only negative reviews on the products of his/her competitors, but not both. d. As Single user has different ID' so, their reviews or ratings do not appear as outliers.	Group fake reviewers are more damaging because they may take control of the complete sentiment on a product and completely mislead potential customers. A group of fake reviewers works together and may also register at multiple sites and spam on these sites. a. If Group reviews negatively to the same product, it can lower the rating deviation drastically. b. In this type, Group writes a review roughly at different time intervals when the product is launched c. Due to writing reviews at random intervals, it hides spikes. d. Also, Large groups are divided into sub-groups so that each sub-group can spam for different Product.

Next, the paper by Fazzolari al. [2] considered a set of effective features used for classifying fake reviews and re-engineered them by considering the Cumulative Relative Frequency Distribution of each feature. Not via providing a few changes withinside the well-tested state-of-the-art algorithms, however most effective through modifying the input used for the training phase to construct supervised classifiers. By an experimental assessment performed on actual data from Yelp.com, the paper indicates that the usage of the distributional features is capable of enhancing the performances of the classifier. The limitation observed in this paper is that it only detects individual spammers, but groups of users who, performing in a coordinated and synchronized way, the purpose is to give credit or discredit for a product.

©IJRASET: All Rights are Reserved

Further, the paper by Liu al. [3] proposes a novel approach based on the partially - supervised model, which refers to detecting opinion relations as an alignment process. Then, a graph-based co-ranking algorithm is explored to estimate the confidence of each candidate. Finally, authors have extracted candidates with higher confidence as opinion targets or opinion words. Compared to existing strategies primarily based totally on the nearest-neighbor rules, the author's model captures opinion relations more precisely, especially for long-span relations. The Limitation found in this is the presence of additional types of relations between words, such as topical relations, in Opinion, Relation Graph needs to be considered, for co-extracting opinion targets and opinion words.

The paper by Wang∗ al. [4] provides a solution to the problem of loose spammer group detection, i.e. every group members are not needed to review each target product. Paper solves this problem using bipartite graph projection. The authors proposed a set of group spam indicators to measure the spamicity of a loose spammer group and a novel algorithm to detect highly suspicious loose spammer groups is been designed in a divide and conquer manner. Experimental results show that the author's method not only can find loose spammer groups with high precision and recall but also can create more precise candidate fake reviewer groups than FIM, hence alternatively used as preprocessing tool for already existing FIM-based approaches. As a future enhancement, it needs to identify other group spamming detection methods to improve the precision and recall and look for new methods to get rid of the effect of already defined parameters such as minimum group spam score, time windows, and maximum group size. Also needed incorporation of the proposed method with existing FIM-based group spamming detection techniques.

Paper by Yang Xiao al. [5] shows that among social media users, there exist a group of users called opinion spam. They are well organized and post many purposed comments to misdirect public opinion. In this way, they significantly magnify the impact of their employers. Paper conducts quantitative analysis to study and understand the characteristics of opinion spam. Authors analyze the psycholinguistic styles of opinion spam, explore their behavior patterns and network structure. Finally, based on the analysis, context-based collective classification is proposed to detect opinion spam and the model can achieve a 91% F1 score. The issue that needs to address is that this paper focuses only on the Twitter platform, enhancement required by deploying it for other social media platforms like Facebook, LinkedIn, etc.

The authors Nitin Jindal and Bing Liu al. [6], have studied problems in the context of product reviews, which are full of opinion and are widely used by customers and product manufacturers. Several startup companies also appeared in the past two years, which aggregate opinions from product reviews. The paper shows how this is high time to study spam in reviews. To the best of mentioned details, there was no earlier published study on this topic, although Webspam and email spam have been investigated extensively. This paper analyzes such spam activities and presents some effective techniques to detect them. The limitation that needs to address is to improve the detection methods, and also look into spam in other kinds of media, e.g., forums and blogs.

The paper by Asghar al. [7] enriches the feature set of a baseline Spam detection method with Spam detection features (Opinion Spam, Opinion Spammer, Item Spam). Using a dataset of reviews from the Amazon site and sentences labeled for Spam detection, authors evaluate the role of spamicity-related features in detecting and classifying spam (fake) clues and distinguishing them from genuine reviews. For this purpose, the author introduced a rule-based feature weighting scheme and propose a method for tagging the review sentence as spam and non-spam. The addition of a revised feature weighting scheme provides better accuracy from 93% to 96%. Also, a hybrid set of features improves the performance of Fake review detection in terms of better precision, recall, and F-measure values. This work shows that combining spam-related features with a rule-based weighting scheme can improve the performance of even baseline Spam detection methods. But limited feature set is used, which if extended in diverse domains, can produce better results. Secondly, feature selection is performed manually, however, automated feature selection may yield improved results using deep learning models, and the proposed feature weighting scheme operates on a limited set of spamicity features, which can be extended for obtaining more robust results.

Also, authors Lau RY al. [8] present semantic language modeling, and a text mining-based computational models are efficient for the detection of fake reviews, even if fake reviewer exercise complicated strategies. Particularly, the proposed supervised model SVM outperforms than other well-known baseline models in analyzing the Amazon review dataset with true positive rate of over 95% in fake review detection. In addition, authors need to examine more sophisticated language modeling approaches, such as n-gram language models, to improve the efficiency of the fake review detection method.

Finally, the authors Dixit al. [9] used the integration of Naïve Bayesian classification with conceptual and semantic similarity technique is proposed spam detection. For best analysis experiments were conducted on benchmark data sets such as PU1, Linkspam, Spambase, and , Enron corpus. Experimental results achieved highest accuracy of 98.89% than existing. But model trained with fewer e-mail so the documents may decrease classifier accuracy through overfitting.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

III. FEATURE ENGINEERING FOR FAKE REVIEW DETECTION

Feature engineering is used to extract or construct features from data. Crawford et. Al [20], highlighted previous studies, that have used several different types of features that can be extracted from reviews, in the review's text most common words found is mostly used feature referred as bag of words approach, where features for each review made from individual or small groups of words.
Also, researchers have used other characteristics of the reviews, reviewers, and products, such as features describing reviewers behavior or syntactical with lexical features used by Shojaee et al. [11] for classification model SVM with 84% accuracy and naïve bayed with 74% accuracy with AMT dataset of 400 deceptive reviews. The features are categorized as review and reviewer-centric features [20]. Review-centric features are made up with information in a single review whereas reviewer-centric features use all reviews written by particular author, along with information about the author. Here identifying fake reviewer results better as reviewer identification helps to prevent many fake reviews, so in this studies reviewer centric features are focused more. Whereas using a combination of features has generally resulted better performance than any single type of feature to train a classifier, as demonstrated in Jindal et al. [12] & [13] with 5.8 million reviews crawled from amazon site to Logistic Regression classifier. Then Li et al. [10], Fei. et al. Mukherjee et al. [14] and Hammad [15]. Li et al. [16] used general features (e.g., Linguistic Inquiry Word Count and Part Of Speech) in combination with bag-of-words, is a more robust approach than bag-of-words alone. A study by Mukherjee et al. [14] used YELP filtered dataset for training with SVM as the best classifier and got 86% accuracy. This paper focuses on reviewer-centric features as discussed below:

IV. REVIEWER-CENTRIC FEATURES

As highlighted earlier, figuring out spammers can enhance the detection of fake reviews, due to the fact many spammers proportion profile traits and pastime styles. Various combo of capabilities engineered from reviewer profile traits and behavioral styles had been studied mentioned by Jindal et al. [12], Jindal et al. [13], Li et al. [10], Fei et al. [22], and Mukherjee et al. [14]. Details of reviewer-centric features are described in Mukherjee et al. [14] as mentioned in the below observations follows:

A. Maximum Number Of Reviews

It turned into located that approximately 75 % of spammers write greater than five reviews on any given day. Therefore, contemplating the quantity of reviews consumer writes consistent with day can assist discover spammers due to the fact 90 % of valid reviewers in no way create multiple review on any given day.

B. Percentage Of The Positive Reviews

Around 85 % of fake reviewer wrote greater than 80 % in their reviews as positive review, for this reason a more percent of fine review is probably a demonstration of an untrustworthy reviewer.

C. Review Length

The common review length can be an crucial indication of reviewers with questionable intentions due to the fact approximately 80 % of spammers don't have any reviews longer than one hundred thirty five word even as greater than 92 % of dependable reviewers have a median reviews length of more than two hundred phrases.

D. Reviewer Deviation

It turned into located that spammers' rankings generally tend to deviate from the common reviews score at a miles better charge than valid reviewers, for this reason figuring out consumer score deviations can also additionally assist in the detection of cheating reviewers.

E. Maximum Content Similarity

The presence of comparable opinions for unique merchandise with the aid of using the identical reviewer has been proven to be a sturdy indication of a spammer. Mukherjee et al. [14] used cosine similarity; however, different greater superior similarity capabilities primarily based totally upon phrase meanings as opposed to the phrases themselves have proven promise [8].
Similarly, reviewer-centric features for groups are included in the paper by Gupta et al. [1], in which they have collected lakhs of reviews on brands from the Amazon product review site and manually labeled a set of 923 candidate reviewers groups.

©IJRASET: All Rights are Reserved

The groups are extracted using FIM over brand similarities such that users are clustered together if they have mutually reviewed (products of) a lot of brands. Surprisingly, the authors observed that there are a lot of verified reviewers showing extreme sentiment, which, on further investigation, leads to ways to avoid the current mechanisms in place to prevent unofficial encouragements on Amazon. Various features, in brief, are mentioned for detection of a group of fake reviewers[1]:

F. Average Rating

It is used to captures the average rating given by group G to a certain brand B. It does the average of the reviews given by group members to products of the given brand and takes the mean of these ratings. Also predicts that an extreme group may give an average rating value at the extremes, i.e., closer to five stars or one star.

G. Average Upvotes

It captures how many upvotes the given group receives with respect to the given brand. This is the mean/ average of upvotes taken across reviews posted by the group members, for products belonging to the given brand.

H. Average Sentiment

This analyzes the review text that what kind of views are there is it positive, negative or neutral, for the given group and brand pair and finds out the average sentiment of these reviews.

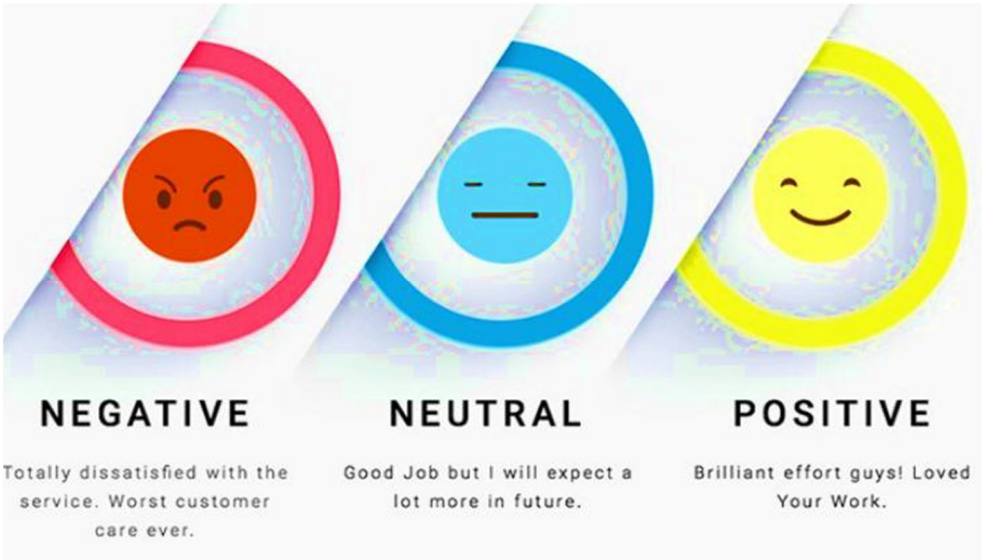


Fig. Types of Sentiment

I. Verified Purchase

A review where the product was bought by the reviewer holds more credibility than the opposing case. This feature determines the fraction of reviews posted by the group for the brand.

J. Review Count

This feature counts the number of reviews and how many views are written for the brand by the group. The fake reviewers are writing more reviews as compared to other users.

K. Early Time Window (ET)

It measures the time gap since the product was produced on the marketplace, and the last review posted on it by the group. The mean value is taken across all the products for the brand.

L. Group Time Window (GT)

The gap between the latest and the earliest review posted is given by Group Time Window. A lower value of GT is the indication of the group is closely linked together and spoils in review spamming.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

M. Rating Deviation

The deviation from the mean rating is calculated with this feature for brands by the group. The fake reviewers show an extremist nature group tends to deviate from 1 to 5.

In addition, these features can be used in a hybrid approach in combination with spam feature scanning to improve classification results. The features listed below are hybrid features used in earlier studies [20].

N. The Ratio of Amazon verified Purchase (RAVP)

This feature is the number of Amazon confirmed purchases divided by the total rating of this user. Reviewers with a high RAVP are considered more reliable because validated purchase ratings are likely to reflect actual ratings.

O. Rating Deviation (RD)

This feature measures the average deviation of a reviewer's evaluation. Spammers can experience significant differences in rating behavior, as the expected behavior of reviewers is to give similar ratings to other users of the same product.

P. Burst Review Ratio (BRR)

This score is calculated as the ratio of the reviewers' ratings displayed in bursts to the total number of reviews written by the reviewers.

Q. Review Content Similarity (RCS)

Average pairwise cosine similarity for all reviews from reviewers. Higher values may indicate that you are likely to be a spammer.

R. Reviewer Burst (RB)

This measures the number of reviews displayed in both Reviewer Burst and Product Burst. In case this feature has higher value the person is more likely to be fake reviewer.

V. CONCLUSION

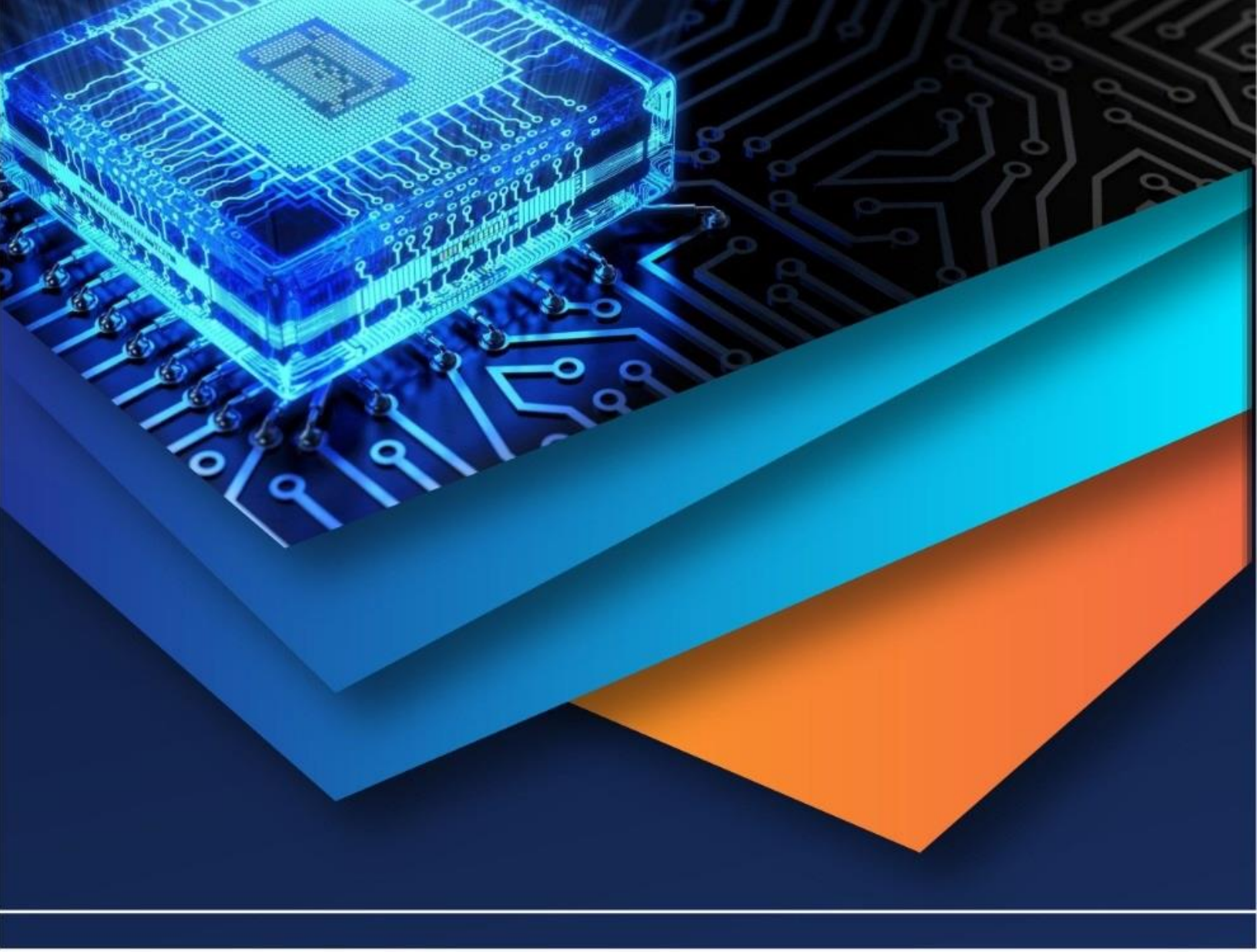
Over the past few years, online reviews have become very important, since reviews influence but fake reviews by the group can drastically influence the purchase decision of consumers and the reputation of businesses. Therefore, the practice of writing fake reviews can have severe consequences on customers and service providers, it highlights basic concepts regarding fake reviews and fake reviewers. This paper covers brief details in this field that have been mentioned by various authors. The literature described in this paper focuses primarily on feature engineering, Feature Engineering studies give a brief idea of features mainly used for reviewer-centric spam detection. The main goal of this paper to focus on Reviewer-centric features is because detection of one fake reviewer can potentially help to identify a number of fake reviews hence it has proven to be important for accurate detection of review spam, as shown in [1, 10, 12, 13, 14, 15]. Despite many studies focused on feature engineering, experiments use different datasets, so it is not possible to identify the optimal type of feature.

REFERENCES

[1] Viresh Gupta , Aayush Aggarwal , and Tanmoy Chakraborty. “Detecting and Characterizing Extremist Reviewer Groups in Online Product Reviews”. IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, 2020
[2] Michela Fazzolari, Francesco Buccafurri, Gianluca Lax, Marinella Petrocchi, “Experience: Improving Opinion Spam Detection by Cumulative Relative Frequency Distribution”. ACM Journal of Data and Information Quality, Vol. 13, No. 1, Article 4. Publication date: January 2021
[3] Kang Liu, Liheng Xu and Jun Zhao, “Co-extracting Opinion Targets and Opinion Words from Online Reviews Based on the Word Alignment Model”. 2013 IEEE
[4] Zhuo Wang∗, Tingting Hou, Dawei Song, Zhun Li And Tianqi Kong , “Detecting Review Spammer Groups via Bipartite Graph Projection”. The Computer Journal, 2015
[5] Yang Xiao , JieFan Qiu, “Exploring and Detecting Opinion Spam on Social Media” IEEE 2020
[6] Nitin Jindal and Bing Liu , “Opinion Spam and Analysis”. Copyright 2008 ACM 978-1-59593-927-9/08/0002
[7] Muhammad Zubair Asghar, Asmat Ullah, Shakeel Ahmad, Aurangzeb Khan , “Opinion spam detection framework using hybrid classification scheme”. June 2019 Springer
[8] Lau RY, Liao SY, Kwok RCW, Xu K, Xia Y, Li Y (2011) Text mining and probabilistic language modeling for online review spam detecting. ACM Trans Manage Inf Syst 2(4):1–30
[9] Dixit S, Agrawal AJ (2013) Survey on review spam detection. Int J Comput Commun Technol ISSN (PRINT) 4:0975–7449
[10] Ott M, Choi Y, Cardie C, Hancock JT (2011) Finding deceptive opinion spam by any stretch of the imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1 (pp. 309–319). Association for Computational Linguistics

[11] Shojaee S, Murad MAA, Bin Azman A, Sharef NM, Nadali S (2013) Detecting deceptive reviews using lexical and syntactic features. In: Intelligent Systems Design and Applications (ISDA), 2013 13th International Conference on (pp. 53–58). IEEE, Serdang, Malaysia
[12] Jindal N, Liu B (2007) Review spam detection. In: Proceedings of the 16th international conference on World Wide Web (pp. 1189–1190). ACM, Lyon, France
[13] Jindal N, Liu B (2008) Opinion spam and analysis. In: Proceedings of the 2008 International Conference on Web Search and Data Mining (pp. 219–230). ACM, Stanford, CA
[14] Mukherjee A, Venkataraman V, Liu B, Glance NS (2013) What yelp fake review filter might be doing? Boston, In ICWSM.
[15] Hammad ASA (2013) An Approach for Detecting Spam in Arabic Opinion Reviews. Doctoral dissertation, Islamic University of Gaza
[16] Li J, Ott M, Cardie C, Hovy E (2014) Towards a general rule for identifying deceptive opinion spam. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, pages 1566–1576, Baltimore, Maryland, USA, June 23-25 2014. ACL
[17] Bing L (2008) Web Data Mining. Book. Springer, Berlin Heidelberg New York
[18] Li F, Huang M, Yang Y, Zhu X (2011) Learning to identify review spam. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence, vol 22, No. 3., p 2488
[19] A. Kim. (2017). that review you wrote on Amazon? Priceless. [Online]. Available: https://www.usatoday.com/story/tech/news/ 2017/03/20/review-you-wrote-amazon- pricess/99332602/
[20] Michael Crawford, Taghi M. Khoshgoftaar, Joseph D. Prusa, Aaron N. Richter & Hamzah Al Najada Survey of review spam detection using machine learning techniques, Crawford et al. Journal of Big Data
[21] E. Gilbert and K. Karahalios, "Understanding deja reviewers," in Proc. ACM Conf. Comput. supported Cooperat. Work (CSCW), 2010, pp. 225–228, doi: 10.1145/1718918.1718961.
[22] Amazon.in. (2018). Review Community Guidelines. [Online]. Available: https://www.amazon.in/gp/help/customer/display.html?nodeId= 201929730
[23] Geli Fei1 Arjun Mukherjee1 Bing Liu1 Meichun Hsu2 Malu Castellanos2 Riddhiman Ghosh2, "Exploiting Burstiness in Reviews for Review Spammer Detection", Association for the Advancement of Artificial Intelligence 2013.
[24] A. Mukherjee, B. Liu, and N. Glance, "Spotting fake reviewer groups in consumer reviews," in Proc. 21st Int. Conf. World Wide Web (WWW), 2012, pp. 191–200.
[25] Y. Lu, L. Zhang, Y. Xiao, and Y. Li, "Simultaneously detecting fake reviews and review spammers using factor graph model," in Proc. 5th Annu. ACM Web Sci. Conf. (WebSci), 2013, pp. 225–233.
[26] S. Rayana and L. Akoglu, "Collective opinion spam detection: Bridging review networks and metadata," in Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD), 2015, pp. 985–994.
[27] S. Dhawan, S. C. R. Gangireddy, S. Kumar, and T. Chakraborty, "Spotting collective behaviour of online frauds in customer reviews," 2019, arXiv:1905.13649. [Online]. Available: http://arxiv.org/abs/1905. 13649
[28] 27. Mayzlin D, Dover Y, Chevalier JA (2012) Promotional reviews: An empirical investigation of online review manipulation (No. w18340). National Bureau of Economic Research, Nashville, TN
[29] Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182
[30] Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ (2007) The development and psychometric properties of LIWC2007
[31] Eisenstein J, Ahmed A, Xing EP (2011) Sparse additive generative models of text. In: Proceedings of the 28th International Conference on Machine Learning (ICML-1)., pp 1041–1048
[32] Abbasi A, Chen H, Nunamaker JF (2008) Stylometric identification in electronic markets: Scalability and robustness. J Manage Inf Syst 25(1):49–78
[33] Mukherjee A, Venkataraman V, Liu B, Glance N (2013) Fake review detection: Classification and analysis of real and pseudoreviews. Technical Report UIC-CS-2013-03, University of Illinois, Chicago
[34] Chapelle O, Schölkopf B, Zien A (2006) Semi-supervised learning. Vol. 2. Cambridge: MIT press.
[35] Blum A, Mitchell T (1998) Combining labeled and unlabeled data with co-training. In: Proceedings of the eleventh annual conference on Computational learning theory (pp. 92–100). ACM, Madison, WI
[36] Liu B, Dai Y, Li X, Lee WS, Yu PS (2003) Building text classifiers using positive and unlabeled examples. In: Data Mining, 2003. ICDM 2003. Third IEEE International Conference on (pp. 179–186). Melbourne, Florida, IEEE
[37] Hernández D, Guzmán R, Móntes y Gomez M, Rosso P (2013) Using PU-learning to detect deceptive opinion spam. In: Proc. of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)