Journal of Informatics
Electrical and Electronics Engineering, 2021,
Vol. 02, Iss. 02, S. No. 004, pp. 1-4
ISSN (Online): 2582-7006

A Study on Opinion Spamming: Fake Consumer Review Detection

Aditya S. Bisht[1], Manish M. Tripathi[2]

[1]M.Tech, Scholar, Department of Computer Science & Engineering, Integral University, Lucknow, India,
[2]Associate Professor, Department of Computer Science & Engineering, Integral University, Lucknow, India,

connect2asbisht@gmail.com[1], mmt@iul.ac.in[2]

How to cite this paper: A. S. Bisht, M. M. Tripathi (2021) A Study on Opinion Spamming: Fake Consumer Review Detection. *Journal of Informatics Electrical and Electronics Engineering, Vol. 02, Iss. 02, S. No. 004, pp. 1-4, 2021.*
https://doi.org/10.54060/JIEEE/002.02.004

Received: 02/04/2021
Accepted: 23/05/2021
Published: 04/06/2021

Abstract

Online audits are the most important wellsprings of data about client feelings and are considered the columns on which the standing of an association is assembled. From a client's viewpoint, audit data is vital to settle on an appropriate choice with respect to an online buy. Surveys are for the most part thought to be a fair-minded assessment of a person's very own involvement in an item, however, the fundamental truth about these audits recounts an alternate story. Spammers abuse these audit stages unlawfully on account of impetuses engaged with composing counterfeit surveys, subsequently attempting to acquire a bit of leeway over contenders bringing about an unstable development of assessment spamming. This training is known as Opinion (Review) Spam, where spammers control and toxic substance surveys (i.e., making phony, untruthful, or misleading audits) for benefit or gain. It has become a typical practice for individuals to discover and to understand assessments/surveys on the Web for some reasons. For instance, in the event that one needs to purchase an item, one commonly goes to a vendor or audit site (e.g., amazon.com) to peruse a few surveys of existing clients of the item. In the event that one sees numerous positive audits of the item, one is probably going to purchase the item. Notwithstanding, in the event that one sees many negative surveys, he/she will in all probability pick another item. Positive suppositions can bring about huge monetary benefits and additionally popularities for associations and people. This, sadly, offers great motivating forces for input spam. Most of the momentum research has zeroed in on regulated learning strategies, which require named information, a shortage with regards to online survey spam. Examination of techniques for Big Data is of revenue, since there are a huge number of online audits, with a lot seriously being produced every day. Until now, we have not discovered any papers that review the impacts of Big Data examination for survey spam identification. The essential objective of this paper is to give a solid and far-reaching similar investigation of flow research on identifying audit spam utilizing different AI procedures and to devise a strategy for directing further examination.

Keywords

Spam, Big data, machine learning, detection

1. Introduction

In recent years, the overall Web has drastically changed the manner in which individuals convey and share their conclusions internationally. Online sentiments are currently communicated as posts [2], remarks, audits, or tweets on various online stages like internet business destinations [3], conversation gatherings, survey locales, news locales, or some other interpersonal interaction site. One of the methods of imparting an insight is to compose a survey about an item or a help reflecting the client's experience of that item or administration. [14- 25]. A client trusts in experiencing all the audits about an item prior to choosing to buy it [6], [7]. Consequently, these audits are viewed as the essential unit of business and a shocker for business associations and clients, separately [8], [9], [10]. It has become a typical practice for individuals to peruse online conclusions/surveys for various purposes.

2. Literature Survey

In a new report, a technique was proposed by **E.I Elmurngi and A. Gherbi [1]** utilizing an open-source programming apparatus called 'Weka instrument' to actualize AI calculations utilizing assessment examination to arrange reasonable and unreasonable surveys from amazon audits dependent on three unique classifications positive, negative and unbiased words. In this exploration work, the spam audits are distinguished by just including the supportiveness votes casted a ballot by the clients alongside the rating deviation are viewed as which restricts the general exhibition of the framework. Additionally, according to the analyst's perceptions and trial results, the current framework utilizes Naive Bayes classifier for spam and non-spam order where the precision is very low which may not give exact outcomes to the client.

J. C. S. Reis, A. Correia, F. Murai, A. Veloso, and F. Benevenuto [2] have proposed arrangements that relies just upon the highlights utilized in the informational index with the utilization of various AI calculations in identifying counterfeit news via web-based media. Despite the fact that distinctive AI calculations the methodology needs demonstrating how exact the outcomes are.

B. Wagh, J.V. Shinde, P.A. Kale [3] chipped away at twitter to dissect the tweets posted by clients utilizing feeling investigation to characterize twitter tweets into good and negative. They utilized K-Nearest Neighbour as a technique to assign them feeling marks via preparing and testing the set utilizing highlight vectors. In any case, the pertinence of their way to deal with other sort of information has not been approved.

B. Liu, et al [4] Although scientists have been reading spam for a long time, for example, web spam and email spam, with regards to assessment spam an unheard-of level of difficulties emerge. In contrast to different sorts of web spam (Email spam, interface spam, counterfeit news) assessment spam is hard to distinguish physically by the natural eye. This makes it practically difficult to separate important, highest quality level datasets which can be utilized to plan location calculations and Systems.

Y. Yao et al [5] proposed the possibility that despite the fact that few kinds of exploration have shown that Recurrent Neural Networks are extraordinary for producing probabilistic language models, they have missed the mark regarding genuinely imitating man composed writings. Nonetheless, this isn't the situation with regards to space explicit messages, for example, short length audits which can undoubtedly be created to copy human-composed writings. The specialists subsequently proposed that Deep neural organizations could be utilized to create assessment spam by spammers sooner rather than later and may as of now be being used for such a reason. To counter such an issue, they built up a robotized audit composing model dependent on the Recurring Neural Network (RNN), their discoveries were that normal language models have restricted execution and effectiveness when the preparation information is made out of long text based successions, though RNN settle this issue by building a memory model. Perhaps the main finishes of this examination indicated that separated from assessment spam composed by people, machine-produced audits are more earnestly to distinguish even with the most progressive and best-prepared AI calculations. To test this hypothesis, the analysts applied SVM's prepared on similitude highlights (cosine comparability of Unigrams), Semantic highlights (recurrence of positive and negative words and suppositions), syntactic highlights (recurrence of POS labels) and LIWC highlights, notwithstanding, none of the classifiers could recognize and recognize the machine-created audits from the genuine ones and passed them all as honest. This shows that spammers are getting more intelligent and there is a requirement for brilliant location frameworks to counter that spamming. Most conventional models missed the mark concerning recognizing and identifying machine produced surveys as spam and allowed them to go through the channel. Except if one approaches a machine created information corpus to additional train the models, this methodology appears to be troublesome.

M. Ott et al [6] scientists planned a few examination inquiries for the survey spam area and played out a few experimentations to do an investigation and get bits of knowledge on these issues. The investigation coordinated and put together the experimentation with respect to 4 distinct situations, for example, (disconnected learning with non-chronologically requested suppositions), and (Using surveys that are arranged on their posting time in a disconnected learning climate). Both these situations were continued utilizing surveys for online conditions. The examination utilized 2 diverse datasets, one from Yelp, which was illustrative of this present reality audits.

3. Research Gap

We can easily find plenty of research based on opinion spamming. Unfortunately, all of them lead towards mathematical and/or graphical representation of data showing either positives or negatives of the products under review. While this project is practical based, helps in effective analysis of products' reviews.

4. Problem Statement

Since we are interested in the review analysis of products, it should be done under various cases using programming utilities/libraries for data manipulation and analysis. Another problem that arises is that it is unreliable to include products with very few reviews so we will include only those products that have considerable number of reviews.

5. Conclusions

A lot of research has been done on the detection of fake and deceptive reviews and filter it from genuine truthful ones. For this study, we have surveyed most of the existing literature regarding opinion spam detection that uses machine learning and natural language processing. The objective of this study was to better understand the existing research on the methodologies and machine learning techniques used so far and to provide future insights to Researchers. The study has reviewed research work done in 3 different categories of detection methods, Review spam detection, Spam user detection, and Spammer group detection using supervised, unsupervised or semi-supervised learning. It has been noted that even though most of the literature is focused on the review centric features and that too using supervised learning, better accuracy can be attained by taking other features such as reviewer and reviewer groups centric features into account. Topological features such as social media activity of these spammer individuals can further enhance the detection results. From the reviewed literature, it is clear that the major challenge in the field of opinion spam detection is the unavailability of the labelled dataset. Although many studies have crafted their own synthetic datasets, it is noticed from the literature that these datasets do not represent the ground truth, real-world reviews as they were written not by spammers but by turkers for research.

References

[1]. E. I. Elmurngi and A. Gherbi, "Unreasonable Reviews Detection on Amazon Reviews utilizing Sentiment Analysis with Supervised Learning Techniques," Journal of Computer Science, vol. 14, no. 5, pp. 714–726, June 2018.

[2]. J. C. S. Reis, A. Correia, F. Murai, A. Veloso, and F. Benevenuto, “Supervised learning for fake news detection,” IEEE Intell. Syst., vol. 34, no. 2, pp. 76–81, May 2019.

[3]. B. Wagh, J. V. Shinde, and P. A. Kale, “A Twitter Sentiment Analysis using NLTK and machine learning techniques,” Int. j. emerg. res. manag. technol., vol. 6, no. 12, June 2018.

[4]. B. Liu, “Opinion Spam Detection,” in Sentiment Analysis and Opinion Mining, Cham: Springer International Publishing, pp. 113–125, 2012.

[5]. Y. Yao, B. Viswanath, J. Cryan, et al., “Automated crowdturfing attacks and defenses in online review systems,” in Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, Oct 2017.

[6]. M. Ott, Y. Choi, C. Cardie, et al., “Finding deceptive opinion spam by any stretch of the imagination,” pp. 309–319, July 2011.

[7]. H. Li, G. Fei, S. Wang, et al., “Bimodal distribution and co-bursting in review spam detection,” in Proceedings of the 26th International Conference on World Wide Web, pp.1063-1072, April 2017.

[8]. R. Verma, and P. Hridayalankar,"Overview PAPER ON DETECTING FAKE SELLERS USING REVIEWS "Procedures of IEEE forum International Conference, pp.5-9, Dec 2019.

[9]. D. Martens and W. Maalej, “Towards understanding and detecting fake reviews in app stores,” Empir. Softw. Eng., vol. 24, no. 6, pp. 3316–3355, Dec 2019.

[10]. X. Wang, X. Zhang, C. Jiang et al., "Distinguishing proof of phony surveys utilizing semantic and social highlights," in fourth International Conference on Information Management (ICIM), Oxford, pp. 92-97,2018.

[11]. N. A. Patel and R. Patel, “A survey on fake review detection using machine learning techniques,” in 4th International Conference on Computing Communication and Automation (ICCCA), pp. 1-6, Dec 2018.

[12]. N. S. Chowdhary and A. A. Pandit. N. S. and A. A., “Fake Review Detection using Classification,” Int. J. Comput. Appl., vol. 180, no. 50, pp. 16–21, June 2018.

[13]. W. Liu, J. He, S. Han, et al., “A method for the detection of fake reviews based on temporal features of reviews and comments,” IEEE Eng. Manag. Rev., vol. 47, no. 4, pp. 67–79, July 2019.