IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124187

A Survey on Fake Product Review Monitoring

Gowrishankar B S[1], Bhoomika H Y[2], Purushottam B N[3], Shalini N[4], Anusha M[5]

Dept. of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysore[1-5]

Abstract: As we can see, the trend of online shopping has evolved over the past two decades into a practical choice for everyone of us. Due to the expanding demand and desire for online shopping and online enterprises, Businesspeople must constantly rely on computer science and related technology to understand what the customer wants. Yet, because the number of consumers is expanding quickly and on a larger scale online, it is challenging for interested parties to fully comprehend the reviews they require to evaluate a product. Moreover, some product reviews are fake. Consumers and product providers that seek to meet the demands of the client encounter obstacles as a result of this**.**

The quantity of client reviews for the goods increases quickly as e-commerce expands and gains popularity every day. A well-liked product may have hundreds of reviews, perhaps thousands. Because of this, it is challenging for a potential customer to read them and decide whether or not to acquire the items. In this research, a method is implemented that makes use of data mining to analyze real product evaluations submitted by real consumers, notifies the creators and other consumers of the positive or negative review, and blacklists bogus accounts.

Moreover, there are a few bogus product evaluations from time to time. Consumers and product manufacturers that make an effort to understand the demands of the customer both encounter difficulties as a result. Finally, after they have reviewed the product themselves, the comment will be automatically checked to see if it is positive or negative. Fake accounts will also be blocked and prevented from accessing further. Based on experimental research and surveys, this review monitoring technique was found to be successful and efficient.

Keywords: online shopping, client needs, product reviews, false reviews, e-commerce, data mining, authentication, summary of reviews, positive/negative comments, fake accounts, review monitoring, successful and efficient.

I. INTRODUCTION

Purchasing desired products and commodities has always been a part of humanity's modernization process. Based on our own experiences, we frequently advise friends, family members, and acquaintances which things to purchase and which not to. Similar to this, we chat with people who are knowledgeable in that sector when we want to buy something we have never done before.

Manufacturers have also used this method of gathering customer feedback to determine the items or product features that will pleasure customers the most. This tendency has continued since that time, but in the era of digital technology, it has transformed into online reviews. As a result of the development of e-commerce in our contemporary times, focusing on such online business components has become essential.

Thankfully, practically all online retailers have integrated the use of review platforms with their merchandise. Yet, keeping track of evaluations and organizing them becomes difficult due to the vast number of online connected people that live all over the world. To address these difficulties, via this document, we have developed a new Review monitoring technique. The organizing of user reviews will be aided by this technique, enabling both potential buyers and producers to decide quickly whether to buy or sell various goods. With our approach, customers will have the ability to access reviews of certain products as well as reviews of particular product characteristics.

according to This enables users to base judgements on both the benefits and the shortcomings experienced by past clients. For the producers, it will make it easier for them to get user input on their product, which will help them determine whether to keep selling it or make the necessary modifications. Also, due to the nature of internet platforms, some reviews may be false or incorrect. This problem is also addressed by our solution, as there will be adequate user authentication and subsequent banning of irrelevant or suspect accounts.

DOI: 10.17148/IJARCCE.2023.124187

II. LITERATURE REVIEW

1] An idea for Review Graph Based Online Shop Review Spammer Detection was put up in 2011. Researchers in this study created a unique idea of a heterogeneous review graph in order to describe the connections between reviewers, reviews, and the stores that the reviewers had evaluated. They looked at how interactions between nodes in this graph may help identify spam's source, in addition to offering an iterative process for detecting dubious reviewers. For the purpose of detecting review spam, these complex relationships have never before been found. Additionally, they create a powerful computing technique to quantify the honesty, dependability, and trustworthiness of stores, reviewers, and reviews themselves. Unlike previous methods, they did not rely on review text data. Thus, the model is compatible with current methods and is capable of identifying.

[2] In 2014, there was a paper distributed on Spotting Counterfeit Surveys by means of Aggregate Positive - Unlabeled Learning. In this research, experts focused on Dianping, the largest Chinese audit enabling site, It filtered surveys from the architecture of Dianping's bogus survey locations. While Dianping's estimate has a very high degree of accuracy, understanding the review is challenging. This paper concentrated on the issue of phoney survey recognition in the Aggregate PU learning system. The names of audits have very high accuracy but opaque review, For the varied structure of surveys, consumers, and IPs, this study first recommends controlled learning calculation (MHCC). Afterwards, it is expanded to include a computer chip model, which is more pertinent to the PU learning problem. With the marked information given by the audit facilitating site Dianping, they have led a few examinations to show that joining aggregate order and PU learning, proposed computer processor model enjoys significant upper hands over existing cutting edge gauge calculation. This model has language autonomous highlights due to which model is summed up to numerous different dialects.

[3] 2015 saw the publication of the study BILD Testing for Picking out Suspect Reviews, Suspicious Reviewers, and Group Spammers. Using the review data in this paper, researchers expect to offer parameterized algorithms for spotting suspicions. Internet audits provide buyers with useful information on goods and services. On the other side, spammers are also a part of the group and are creating fake polls to get people to follow them. Previous attempts to identify spammers relied on the behavior of reviewers, text similarities, linguistic factors, and rating patterns. These studies have the ability to identify specific types of spammers, such as those that post several identical surveys regarding a single target material. Private information in review metadata includes MAC addresses, IP addresses, locations, dates, and browsers.

[4] Representation Learning Based Opinion Spam Detection (RLOSD) research was published in 2017. To distinguish between honest and dishonest evaluations, the researchers in this study developed a decision tree-based technique. to obtain the appropriate attributes, evaluate them with a decision tree, they have applied unsupervised representation learning in addition to conventional feature selection techniques. The model considers data correlation for choosing the appropriate characteristics. To lower the dimensionality of the feature space and get rid of duplicate and pointless features, a combinational model made up of feature engineering and feature reduction stages is constructed. The efficacy of review spam detection is increased as a result. There is an initial pre-processing stage. Stop words are eliminated throughout this process, along with stemming and POS tagging. In order to distinguish between spam and non-spam reviews, the next stage employs feature engineering approaches, such as eventually using a classifier that leverages information obtained to rank features and detect fraudulent reviews. As compared to the most popular methods in this field, the outcome demonstrates the greater performance in detecting opinion spam.

[5] Opinion Mining utilizing Ontological Spam Detection was the working title of a study proposal that was posted in 2017. In order to obtain more accurate results from opinion mining, the suggested technique in this study includes Ontology, Geo location and IP address tracking, Spam terms Dictionary utilizing Nave Bayes, Brand only review identification and tracking account. Nowadays, there is a significant increase of online data. In order to obtain a valid opinion given the prevalence of false product evaluations, opinion mining techniques involve spam detection.

Reviews play a major role in influencing online purchases, but not all of them are reliable or honest. Some ratings are manipulated by suppliers, sellers, or publications to boost their profits. Review manipulation is the practice of creating reviews that are fictitious and inaccurate. A new Dataset is established after the previous Dataset's Spam reviews have been identified and removed. Finally, a novel algorithm is put out that does opinion mining on the Spam Filtered Dataset and more accurately detects spam reviews. This study shows that the suggested algorithm is effective and scalable regardless of the kind of dataset employed.

DOI: 10.17148/IJARCCE.2023.124187

[6] 2017 saw the development of a technique employing opinion mining for monitoring and deleting false product reviews from genuine online product reviews. This essay examines consumer reviews of products and contrasts them based on reviews collected in one location. This research differs from others since it uses a technique that only gathers data about a product about which customers have provided feedback and determines whether those opinions are favorable or unfavorable. It takes five steps to get the desired outcome.

First, the customer's email address is used to verify that they are who they say they are; next, product features that customers have reported; finding opinion statements in each review; recognizing if each comment is good or negative; and ultimately. The email address is banned if the provided opinion is false, but the results are summarized in the last step. The issue with this article is therefore how to identify the spam viewpoint among the enormous amount of unstructured data.

[7] Using opinion mining, a monitoring tool for bogus product reviews was developed in 2018. This study proposes an open mining-based solution for identifying fake or spam product reviews. Researchers are seeking to determine a customer's perspective through some written content in this essay. The approach involves gathering reviews and using a decision tree to determine whether they are relevant to a particular product. To find the spam words in the reviews, a spam dictionary is employed. Many algorithms are used in text mining to provide the required results.

The classification of sentiment polarity, a key problem in sentiment analysis, is covered in this article. The software that the researchers have developed and that is described in this paper will assist the user in making the proper purchase. This programme is used for analysis, and if a particular IP address repeatedly submits phoney reviews, the ability to blacklist that IP address is available to admin users. It notifies the user through email of the blacklisted IP address. As a result, the consumer will be certain that the product is available on that application and that any fraudulent reviews posted on any product are being monitored. In 2018, a method for identifying fake reviews that poses problems and challenges was presented.

[8]. In this essay, a general overview of the concerns, difficulties in identifying false reviews, and discussions of fake reviewers have been covered. This research proposes a framework that can handle both labelled and unlabeled data. The accuracy rates provided by the framework employed in this study are 83.70% for semi-supervised learning and 90.19% for supervised learning. In this research, topics and issues from the literature are discussed. In this study, every potential feature was looked into and discussed. This paper conducts the experiment and compiles the findings using the given dataset. This report addressed every prospective direction that could be taken.

[9] A study research paper named Fake Product Review Monitoring System was released in 2019. The goal of this study was to find a solution to every issue surrounding spotting phoney reviews. This essay explores a variety of techniques for spotting false reviews in order to do opinion mining. How to recognise spam reviews in the existing system has been hotly contested. In order to get more reliable results from the opinion mining, other techniques are included, including IP address tracking and Ontologies to identify spam reviews. When it is determined that the original dataset contained spam reviews, a new dataset is produced that is devoid of these reviews, and opinion mining is then carried out on the recently spam-filtered dataset. Opinion mining is then performed on the newly spam-filtered dataset when it is discovered that the original dataset contains spam reviews. Brand-only review detection and tracking accounts, a Naive Bayes spam terms dictionary, geolocation and IP address tracking, and ontology are some of the techniques used in this study.

[10] 2019 saw the publication of a study titled GSCPM: CPM based group spamming detection in online product reviews. Researchers in this work developed the GSCPM, a novel, three-step, totally unsupervised method to identify spammer groups based on the Clique Percolation Method (CPM). It initially constructs a suspicious reviewer graph based on behavioral data signals (timestamp, rating), relational data (network), and behavioral data (timestamp, rating). Depending on CPM, the whole network of suspect reviewers is subsequently separated into k-clique clusters. These K-clique clusters fall under the category of very suspicious probable group spammers. In the end, prospective groups are ranked based on both group-based and individual-based spam indications. Using three authentic Yelp.com review datasets, they evaluated the efficiency of the suggested strategy. The recommended method outperforms four other tactics, according to experimental findings, in terms of prediction accuracy. It is recommended to just label a limited amount of data to increase detection precision. The adoption of partially supervised techniques is the way of the future.

[11] In 2019, the Intelligent Interface for Monitoring and Removing False Product Reviews was made available. The suggested study uses the n-gram approach to recognize bogus reviews across various languages, including Urdu and Roman Urdu, as it is challenging to spot phoney evaluations on your own. During the investigation, it was discovered that the best method for identifying bogus reviews is text categorization using an SVM classifier. Hence, fraudulent reviews are distinguished.

DOI: 10.17148/IJARCCE.2023.124187

[12] A paper on Declarative Programming Method for Review Detection was released in 2020. The method used in this paper presentation to identify talk reviews combined attributes derived from the reviews' text with author behavior. Having a white box strategy for the industry is more important than ever because of the growing social unease around judgements made using personal data.

This research presents a rule-based phoney review detection system based on Answer Set Programming, a strong way for detecting hostile behavior patterns suggested by a range of limits. This approach combines information on the total number of reviews, the number of dislikes, the number of reviews, analysis of the submission times, the number of dislikes, the number of reviews, the quality of the content based on similarity metrics, and the classification of reviews and the product. The creation of a system for identifying spam, spammers, and fraudulent reviews is advised in this article.

III. OUR FINDINGS FROM STUDIED LITERATURE

The difficulties found inspire efforts to offer solutions for each of the issues listed in the problem description section above. The goals of the literature under review are as follows:

• To use several methods, including IP Address, Account Used, Negative Word Dictionary Using Senti-strength, and Ontology, in order to improve Spam Detection.

• Work that is represented graphically.

• Identify and address the six main types of spam reviews.

• Opinion Mining on Spam Filtered Data will be the subject of the presentation.

• Spam detection will include ontologies.

• To show how a method that combines opinion mining and spam detection may be used.

• Literature was used to outline problems and difficulties.

• To address the problems, a framework has been suggested.

• Potential features were investigated and spoken about.

• Potential directions for the future were considered.

IV. CONCLUSION AND FUTURE WORK

As a research topic, locating opinion spam in enormous volumes of unstructured data has gained significance. Currently, professionals, academics, and corporate organizations are competing to develop the best system for analyzing opinion spam. Internet evaluations that are false have an increasing negative impact on customers and businesses. Therefore, it is crucial to identify and remove such bogus reviews from online sources. This study offers many techniques for review spam detection and performance indicators. In order to be more precise and helpful, many strategies for detecting spam reviews have been examined in this research. A thorough examination of the methods used to determine if a review is spam is offered.

One direction for further research is to put the system in place and evaluate performance using the suggested approach on various benchmark data sets. Future research may also evaluate the effectiveness of several classification algorithms to select the best one for our suggested categorization method for opinion spam. Other types of review or reviewer-related features do exist, though, and they are likely to contribute to the prediction task. In order to make forecasts that are more accurate, we will do additional research into various feature types in the future.

V. DISCUSSION

Based on the information in the pertinent study publications, the existing literature is summarized. The detection of false reviews is a challenging problem. This is because phoney reviews have at least two significant root causes: First comes the review's substance, followed by the reviewer's actions. The results from publications based on the same study are acknowledged last. The main points of the researched research articles are briefly discussed. In particular, this research offered advice on how to address the challenge of identifying fake/spam reviews. Beginners will benefit much from the material offered.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.124187

REFERENCES

[1] Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu, "Review Graph based Online Store Review Spammer Detection", 2011 11th IEEE International Conference on Data Mining, pp. 1242-1247, 2011.
[2] Huayi Li∗, Zhiyuan Chen∗, Bing Liu∗, Xiaokai Wei∗ and Jidong Shao†, "Spotting Fake Reviews via Collective PositiveUnlabeled Learning", 2014 IEEE International Conference on Data Mining, pp. 899-904, 2014.
[3] Tanya Gera, Deepak Thakur, Jaiteg Singh, "BILD Testing for Spotting Out Suspicious Reviews, Suspicious Reviewers and Group Spammers", 2015 Fifth International Conference on Communication Systems and Network Technologies, pp. 976-981, 2015.
[4] Zeinab Sedighi, Hossein Ebrahimpour-Komleh, Ayoub Bagheri, "RLOSD: Representation Learning based Opinion Spam Detection", 2017 3rd Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), pp. 74-80, 2017.
[5] Neelam Duhan, Divya, Mamta Mittal, "Opinion Mining using Ontological Spam Detection", 2017 International Conference on Infocom Technologies and Unmanned Systems (ICTUS'2017), pp. 557-562, 2017.
[6] Manleen Kaur Kohli, Shaheen Jamil Khan, Tanvi Mirashi, Suraj Gupta, "Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 1, pp. 106-109, January 2017.
[7] Anusha Sinha, Nishant Arora, Shipra Singh, Mohita Cheema, Akthar Nazir, "Fake Product Review Monitoring Using Opinion Mining", International Journal of Pure and Applied Mathematics, Volume 119, No. 12, pp. 13203-13209, 2018.
[8] Jitendra Kumar Rout, Amiya Kumar Dash, Niranjan Kumar Ray, "A Framework for Fake Review Detection: Issues and Challenges", 2018 International Conference On Information Technology (ICIT), pp. 7-10, 2018.
[9] Piyush Jain, Karan Chheda, Mihir Jain, Prachiti Lade, "Fake Product Review Monitoring System", International Journal of Trend in Scientific Research and Development (IJTSRD), Volume 3, Issue 3, pp. 105-107, Mar-Apr 2019.
[10] Guangxia Xu, Mengxiao Hu, Chuang Ma, Mahmoud Daneshmand, "GSCPM: CPM-based group spamming detection in online product reviews", 2019.
[11] Ata-Ur-Rehman, Nazir M. Danish, Sarfraz M. Tanzeel, Nasir Usama, Aslam Muhammad, Martinez-Enriquez A. M., Adrees Muhammad, "Intelligent Interface for Fake Product Review Monitoring and Removal", 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) Mexico City, Mexico, September 11-13, 2019.
[12] Nour Jnoub, Wolfgang Klas, "Declarative Programming Approach for Fake Review Detection", 2020.

