International Journal of Emerging Science and Engineering (IJESE)
ISSN: 2319-6378 (Online), Volume-7 Issue-1, March 2021

Fake Product Monitoring and Removal for Genuine Product Feedback

Mayuri Patil, Snehal Nikumbh, Aparna Parigond, Madhavi Patil

Abstract*: *A customer's decision to purchase a product or service are primarily influenced by online reviews. Customers use online reviews, which are valuable sources of information to understand the public opinion on products and/or services. Dependability on online reviews can give rise to the potential concern that violator could give deceitful reviews in order to synthetically promote or decry products and services. This practice is known as Opinion Spam, where spammers manipulate reviews by making fake, untruthful, or deceptive reviews to get profit and boost their products, and devalue a competitor's products. In order to tackle this issue, we propose to build a fraud risk management system and removal model. This captures fraudulent transactions based on user behaviors and network, analyses them in real-time using Data Mining, and accurately predicts the suspicious users and transactions. In this system, we use two algorithms NLP and TF-IDF to differentiate between fake and genuine reviews or feedback received by the customers

Keywords*: *Genuine Reviews, Fake Reviews, Opinion Spam, Opinion Mining

I. INTRODUCTION

The internet is continuing to grow in size and importance, and hence the quantity and impact of online reviews is increasing continuously as well. Reviews can influence people of all areas but notably e-commerce is one of the most important aspect, where comments and reviews regarding products and services are often the most convenient and easy way for buyers to make a decision on whether or not to buy a product. Whenever we buy any product online on any online sites like amazon.com, flipkart.com, and many others, we first tend to check for the products reviews. If we think the product has nice and glorious reviews, we tend to immediately press the buy button, trusting these reviews. Even though most of these reviews are from genuine customers posting and sharing their feedback, there is a huge possibility that some of these reviews might be fake that are given to synthetically gain profit by attracting customers to these fake reviews.

That is not to say that online reviews are not helpful, in fact, online reviews can be helpful by helping customers understand if he should or should not buy a product, but blind trust on these reviews is treacherous for both the seller and buyer. Moreover, business owners might allure people to writes good reviews about them and then hire someone to write awful reviews about their competitor's products or services. These fake reviews are considered review spam and can have a great impact in the online marketplace.

II. LITERATURE SURVEY

Yuming Lin, et.al describes that detecting spam reviews is important for current e-commerce applications. In his paper, he and his team explore the issue on fake review reduce online opinion spam. Firstly, the characteristics of fake reviews are examined. Then, based on review contents and reviewer behaviours, six-time sensitive features are proposed to highlight the fake reviews. And then, to spot the fake reviews as soon as possible, we develop supervised solutions and a threshold-based solution. The experimental results show that their methods can identify the fake reviews orderly with high accuracy and recall. [1]

Huayi Li, et.al talked about how online reviews have become an increasingly important resource for decision making and product designing. But reviews systems are often battered by opinion spamming. Even though for years, fake review detection has been studied by researchers using supervised learning, direct observation of large datasets is still unavailable and most of existing approaches of supervised learning are not based real fake reviews. Working with Dianping, the largest Chinese review hosting site, they presented the first reported work on fake review detection in Chinese with filtered reviews from Dianping's fake review detection system. Their algorithm has a very high accuracy, but the recall is hard to know. The fake reviews detected by the system are almost surely fake but the remaining reviews may not be all genuine. It is more appropriate to treat it as an unlabelled set since the unknown set may contain many fake reviews. This calls for the model of PU learning. By controlling the complex dependencies among reviews, users and IP addresses, they first propose a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extend it to Collective Positive and Unlabelled learning (CPU). [2] Hence, the critical source of information and feedback about products are online reviews.

Manuscript received on March 19, 2021.
Revised Manuscript received on March 26, 2021.
Manuscript published on March 30, 2021.
*Correspondence Author
Mayuri Manikrao Patil*, Department of Computer Engineering, Dr. D.Y. Patil School of Engineering Academy, Pune (M.H), India. Email: mayusp2605@gmail.com
Snehal Nimba Nikumbh, Department of Computer Engineering, Dr. D.Y. Patil School of Engineering Academy, Pune (M.H), India. Email: sonusandy493@gmail.com
Aparna Parshwanath Parigond, Department of Computer Engineering, Dr. D.Y. Patil School of Engineering Academy, Pune (M.H), India. Email: aparigond98@gmail.com

[3] Deceiving or spam reviews refers to any unprompted and unneeded fats about the product and services. Spammers put together bogus audits approximately the competitor's products. [4] Deceptive or spam reviews are ones done by scammers. [5] In this way, forged reviews has been a major concern for customers to decide which products to purchase. [6] Lately, the WWW has greatly changed the way for conveying the insights. Online reviews can come from tweets, blog posts and many more online resources like survey destination, new locales, net based totally enterprise destinations. [7]

Customers explore reachable reviews and take decisions on where to shop for the products. [8]

III. PROPOSED SYSTEM

A. Problem Definition

Spammers might furnace and make fake reviews in order to artificially promote or devalue the product's quality and services. Now, customers could be deceived to make wrong decisions because of such behaviour of spammers. Therefore, detecting fake reviews is significantly important.

B. System Architecture

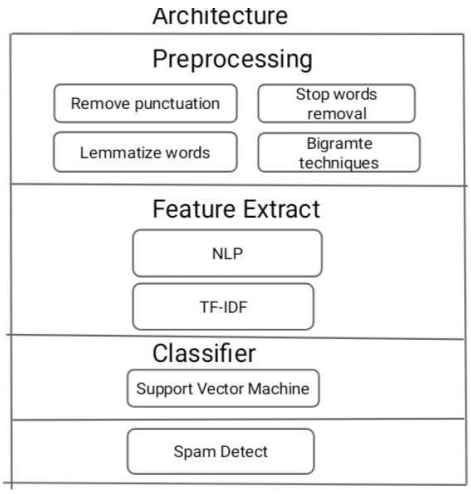

Fig. 1. System Architecture

C. Explanation of System Architecture

There are three stages in our system:

1) Preprocessing: Preprocessing consist of removing any punctuation made on the reviews. Along with it any stop words, for example, "the", "a", "an", "in","" and" are removed. Any lemmatized words are also removed like say, words "am", "are", "is" will become "be". After all this we move on to the next stage Feature Extract

2) Feature Extract: In feature extraction, we apply two algorithms, NLP and TF-IDF. We use NLP to predict if reviews are positive or negative. This is done by bag of words. TF-IDF is used to analysis the importance of a word by analyzing and checking the keyword's relevance throughout internet.

3) Classifier: We use Support Vector Machine for classifying fake reviews from the genuine ones.

IV. ALGORITHM OF PROPOSED SYTEM

1) Sentence Tokenization
2) Create the Frequency matrix
3) Term Frequency calculation and generating a matrix
4) Creating a table for documents per words
5) Calculating IDF and generating a matrix
6) TF-IDF calculation and generation of a matrix
7) Scoring sentences
8) Finding threshold
9) Summary Generation

V. UML DIAGRAM OF PROPOSED SYSTEM

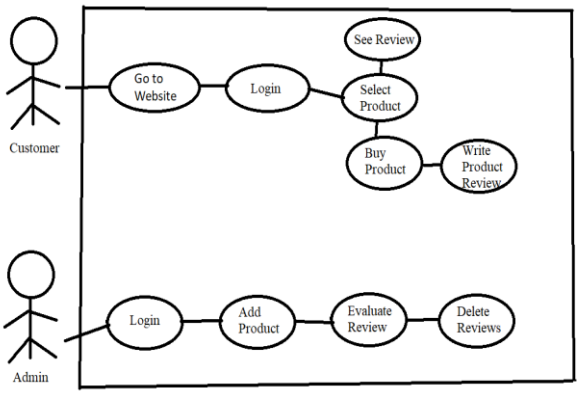

Fig. 2. UML Diagram of System

VI. CONCLUSION

To conclude, we know that now-a-days technology is growing day by day, and there are so many websites and applications available online that sell their products, and every single product have millions if not gazillions of reviews, And since most people decide to buy the products mostly judging by these convenient reviews, there is a need to make sure that there are some genuine reviews out there. Here is where out project will help users buy the right product by doing analysis and then if fake reviews are found from any IP address consistently as well as if the user ID are similar then the admin can take action against such users. In this way, we can monitor the fake reviews made on any product and hence saving the customer form any huge loss.

REFERENCES

1. Yuming Lin, Tao Zhu, Hao Wu, Jingwei Zhang, Xiaoling Wang and Aoying Zhou, "Towards Online Anti-Opinion Spam: Spotting Fake Reviews from the Review Sequence", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2014, pp.261-264.
2. Huayi Li, Zhiyuan Chen, Bing Liu, Xiaokai Wei and Jidong Shao,"Spotting Fake Reviews via Collective Positive-Unlabeled Learning".
3. R.Narayan,J. Rout and S. Jena, "Review Spam Detection Using Semisupervised Technique", Progress in Intelligent Computing Techniques: Theory, Practice, and Applications, pp. 281-286, 2018.
4. C. Lai, K. Xu, R. Y. Lau, Y. Li, and L. Jing, "Toward a Language Modeling Approach for Consumer Review Spam Detection," 2010 IEEE 7th International Conference on E-Business Engineering, pp. 1–8, 2010.
5. J. Rout,S. Singh, S. Jena, and S. Bakshi, "Deceptive review detection using labeled and unlabelled data", Multimedia Tools and Applications,vol.76, no. 3, pp. 3187-3211, 2016.
6. S. Banerjee and A.Y.K. Chua. 2014. "Applauses in hotel reviews: Genuine or deceptive ?", 2014 Science and Information Conference (2014), pp. 938–942,2014.

7. A. Rastogi, M. Mehrotra, "Opinion spam Detection in Online Reviews", Journal of information and Knowledge Management, vol. 16, no. 04, pp. 1-38, 2017.
8. S. Banerjee, A. Chua, J. Kim, "Using Supervised Learning to Classify Authentic and Fake Online Reviews ", Proceeding of the 9th International Conference on Ubiquitous Information Management and Communication", ACM, 2015.

AUTHORS PROFILE

Mayuri Manikrao Patil, Dr. D.Y. Patil School of Engineering Academy, Ambi, Pune, India studying in the field of Computer Engineering, Email: mayusp2605@gmail.com

Snehal Nimba Nikumbh, Dr. D.Y. Patil School of Engineering Academy, Ambi, Pune, India studying in the field of Computer Engineering, Email: sonusandy493@gmail.com

Aparna Parshwanath Parigond, Dr. D.Y. Patil School of Engineering Academy, Ambi, Pune, India studying in the field of Computer Engineering, Email: aparigond98@gmail.com