

Fake Product Review Monitoring and Removal System.

Mr. Autade Sanket Balasaheb
UG Student,
Department of Computer
Engineering
SNDCOE & Rc Yeola

Mr. Bhawar Shubham Balu.
UG Student,
Department of Computer
Engineering
SNDCOE & Rc Yeola

Mr. Shaikh Zuber Harun
UG Student,
Department of Computer
Engineering
SNDCOE & Rc Yeola

Prof. Dhakane Vikas Nivrutti.
Assistant Professor,
Department of Computer
Engineering
SNDCOE & Rc Yeola

ABSTRACT

In the current scenario, the data on the web is growing exponentially. Social media is generating a large amount of data such as reviews, comments, and customer's opinions on a daily basis. This huge amount of user generated data is worthless unless some mining operations are applied to it. As there are a number of fake reviews so opinion mining technique should incorporate Spam detection to produce a genuine opinion. Nowadays, there are a number of people using social media opinions to create their call on shopping for product or service. Opinion Spam detection is an exhausting and hard problem as there are many faux or fake reviews that have been created by organizations or by the people for various purposes. They write fake reviews to mislead readers or automated detection system by promoting or demoting target products to promote them or to degrade their reputations. The proposed technique includes Ontology, Geo location and IP address tracking, Spam words Dictionary using Naïve Bayes, Brand only review detection and tracking account used.

Keywords

Fake Reviews Detection, Text Classification, Natural Language Processing, Machine Learning, Bigrams, Term Frequency and Inverse Document Frequency..

1. INTRODUCTION

There are different ways to shop like you can buy a specific thing of your need by going to a store or mall. In this style of shopping the seller gives you the feedback of the product, you do not know whether he/she is giving a fake feedback or original. Because, it is upon seller honesty, how much the seller is true in his/her words and you have to carefully examine the product because you do not have any other option in examining the product. If you don't pay attention in buying that product then it may be proved a waste for you. On the other hand, nowadays source of shopping has been changed. You can buy the products from the online stores of different brands. Here, you have to place the order without seeing and examining the original product. You read the reviews and buy the product. Therefore, you are dependent on the reviews

about the product. These reviews may be the original or fake. The customer wants to buy an original and reliable product, which is possible only when you get the original feedback related to that product.

1.1 Objective

The identified challenges motivate to bring up a solution to all the problems stated in the above problem statement section. Following are the objectives of the proposed approach and this thesis work: To implement different algorithm to get better Spam Detection Graphical representation of work. To deals with 6 different types of Spam Reviews. To presents Opinion Mining on Spam Filtered Data. To implement Ontology in Spam Detection. Our application which will help the user to pay for the right product without any getting into any scams.

2. ANALYSIS OF THE SYSTEM

2.1 Existing System

When performing any type of internet shopping, many of the users will spend their quality time into reading other user reviews if they are available. Clearly consumers value the feedback given by other users as do the companies that sell such products. Blogs, websites, discussion boards etc. are a repository of customer suggestions which are valuable and important sources of textual data. Therefore, today's individuals and older ones extensively rely on reviews available on line. It means that people make their decisions of whether to purchase the products or not by analyzing and reflecting the existing opinions on those products.

2.2 Proposed System

As most of the people require review about a product before spending their money on the product. So people come across various reviews in the website but these reviews are genuine or fake is not identified by the user. In some review websites some good reviews are added by the product company people itself in order to make product famous this people belong to Social Media Optimization team. They give good reviews for many different products manufactured by their own firm. User will not be able to find out whether the review is genuine or fake. To find out fake review in the website this "Fake

Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining” system is introduced. This system will find out fake reviews made by the social media optimization team by identifying the IP address. User will login to the system using his user id and password and will view various products and will give review about the product. And the user will get genuine reviews about product. And while reviewing he needs to enter the email id from which he is reviewing and it would be verified. If he writes a fake review then his id will be blocked but allowing him to share his opinions again.

2.2 Proposed System

This system provides solution to improve the speed of purchasing of products and faster payment option. In this solution we are using Raspberry-Pi, barcode scanner, Raspberry-Pi touch screen display and a button were placed appropriately in the shopping cart. Each and every product has barcode tag on it. Barcode scanner reads the product information before put into the trolley. When the customer wants to remove the product from the cart, a push button is placed in the cart. Then customer presses the button and scans the product then automatically reduces the cost of the product in the total amount and as well as it removes from cart.

3. LITERATURE REVIEW

As of late, the World Wide Web has greatly changed the technique for imparting the insights. Online audits are feedback, tweets, posts, conclusions on numerous on line tiers like survey destinations, news locales, net based totally enterprise destinations or some different lengthy range interpersonal communication destinations. Sharing audits is one of the strategies to compose a survey approximately administrations or items. Surveys are considered as a person's near domestic concept or experience about items or administrations. Client dissects reachable audits and takes preference whether to shop for the object or no longer. In this way on-line audits are critical wellspring of information approximately patron conclusions. Phony or spam audit alludes to any spontaneous and superfluous facts about the item or administration. Spammer composes counterfeit audits approximately the contenders' item and advances possess items. The surveys composed by spammers are called phony audits or unsolicited mail audits. In this manner counterfeit surveys discovery has grown to be fundamental trouble for customers to decide higher preference on items dependable simply because the sellers to make their purchase.

4. PROBLEM DEFINATION

In recent years, online reviews have been playing an important role in making purchase decisions. This is because, these reviews can provide customers with large amounts of useful information about the goods or service. However, to promote factitiously or lower the quality of the products or services, spammers may forge and produce fake reviews. Due to such behaviour of the spammers, customers would be mislead and make wrong decisions. Thus detecting fake (spam) reviews is a significant problem. Opinion spamming refers to the use of excessive and illicit methods, such as creating a large volume of fake reviews, in order to generate biased positive or negative opinions for a target product or service with the intention of promoting or demoting it, respectively. The reviews created for this purpose are known

as fake, spam or bogus reviews, and the authors responsible for composing such deceptive content are known as fake or spam reviewers.

5. ARCHITECTURE

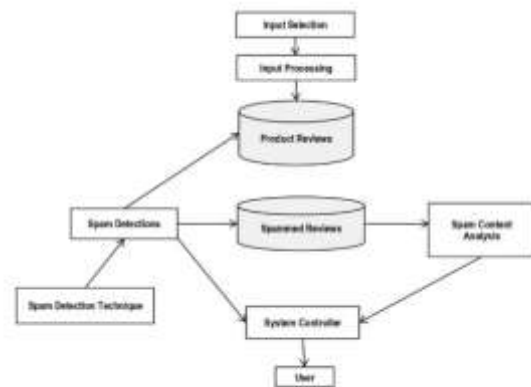


Fig No 1 Architecture

5.1. Implementation Details (Modules)

1. Administrator Login:

Admin login to the framework utilizing his administrator ID and mystery key.

- Add item: Admin will upload object to the framework.

- Delete Review: Admin will evacuate the survey which observed by means of the framework as phony.

2. Client Login:

User will login to the framework making use of his patron ID and mystery phrase.

- View object: User will see object.

- Post Review: User can submit audit approximately the item.

- Tracks IP Address: If the framework reveals an audit is phony it will light up the administrator to expel the phony survey.

Table No 1.Review Assessment

QF	90	80	70	60	50
Electronics	0.983	0.959	0.91	0.83	0.77
Home Appliances	0.987	0.962	0.92	0.85	0.76
Groceries	0.989	0.957	0.92	0.82	0.77

Table 1 explains that the quality factor analysis with respect to different types of product reviews which are discussed in above table.1. Here the Groceries achieves more review factor rather than other to digital review techniques.

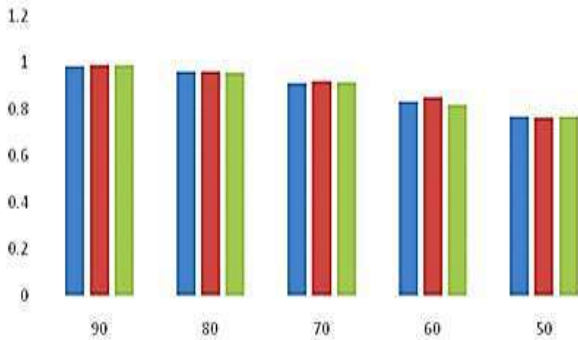


Figure No 2. Review Analysis System

Fig.2 explains that digital review rating system with respect to 2D mechanism. In this situation 3 level transformation mechanism has been used to extract the reviews with windowing techniques by coefficients. At this stage efficient review is obtained but robustness is required to improve.

$$NC = \frac{\sum_i \sum_j w(i,j) \sum_i \sum_j w^1(i,j)}{\sum_i \sum_j w(i,j)^2}$$

$$\text{Review} = 20 \log (255 \sqrt{NR})$$

Table No 2 Review with respect to falser

parameter	review_100	review_100	review_100
electronics	0.99	0.994	0.98
Home appliances	0.995	0.995	0.992
Groceries	0.94	9 0.94	0.97
electronics	0.96	0.967	0.98
Home appliances	0.88	0.89	0.89

The above table explains that different types of reviews on online trading. In this all corrections and elements has attain with efficient manner but extraction is complex procedure compared to traditional machine learning models.

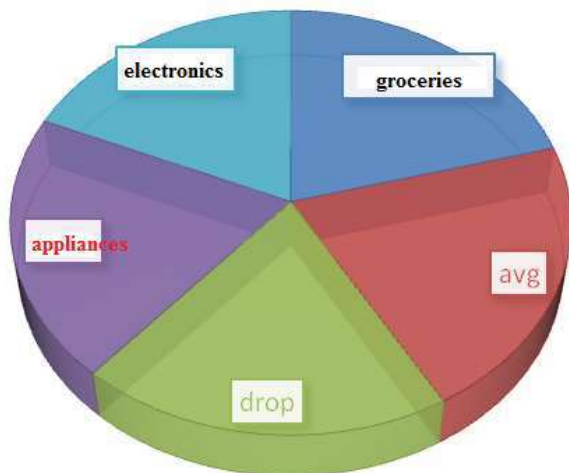


Fig.No 3 Graphical Representations of Reviews

Fig 3 explains that different reviews on online systems, at this stage some of techniques has been failed because of conventional insecure methods. These are limitations.

6. METHODOLOGY

Commentator Centric Approach-This method is predicated upon the behaviour of analysts. This methodology considers data about clients and all surveys that are composed with the aid of them. Highlights utilized proper now account age, profile image, URL duration, IP address, variety of composed audits through one commentator, maximum severe rating every day and so on. **Item Centric Approach-**This technique for the most component facilities around the object related information. Right now, rank of object, value of object and so on are taken into consideration as highlights. At first phony audit identification turned into supplied via Jinal et al. There are special methods to differentiate counterfeit surveys. AI system is one of the procedures to distinguish counterfeit surveys. AI model learns and make forecast. The essential advances associated with AI are records making ready, highlight extraction, include determination, characterization model age. This technique is appeared in Fig. 1:

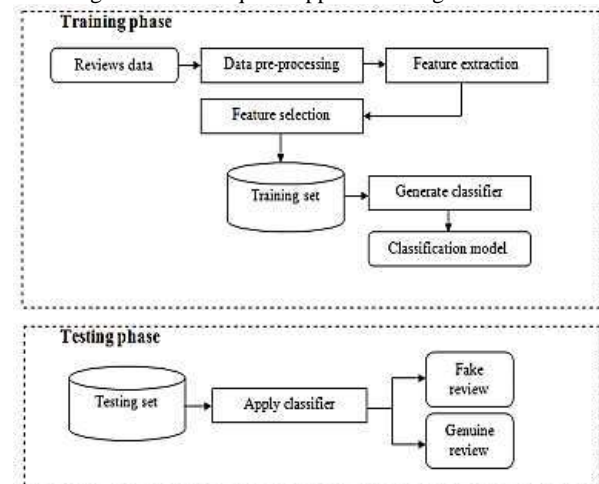


Fig No 4 Training Phase

System works as follows: -

- Admin will add gadgets to the framework.
- Admin will erase the audit that is phony.
- User as soon as get to the framework, patron can see object and can post audit about the item.
- System will comply with the IP deal with of the consumer.
- If the framework watches counterfeit audit originating from equal IP address several a couple of times this IP cope with might be followed with the aid of the framework and will advise the administrator to expel this survey from the framework.

6.1 Algorithm

- Step 1: Enter login Details of Admin, Buyer, Seller.
- Step 2: Seller uploads the product and give certain description.
- Step 3: Buyer selects the product and give reviews.
- Step 4: Admin verifies the given review.
- Step 5: If “Brand Spam” is displayed, go to Step 8.
- Step 6: If “Fake review” is displayed, go to Step 8.

Step 7: If nothing is displayed, go to Step 9

Step 8: The Review is fake and Admin removes it from the database.

Step 9: The Review is genuine.

Step 10: Exit.

6.2 Application

This application will be used by the people who like to spend their money on valuable products

CONCLUSION

Determining and classifying a review into fake or truthful one is an important and challenging problem. As part of future work, we can incorporate review spammer detection into the review detection and vice versa. Exploring ways to learn behavior patterns related to spamming so as to improve the accuracy of the current regression model. So as to evaluate our proposed methods, that conducts user evaluation on an Amazon dataset containing reviews of different manufactured products. We derive an aggregated behavior scoring methods for rank reviewers according to the degree that they demonstrate spamming behaviors. We found that here proposed methods generally outperform the baseline method based votes. We further learn a regression model from the user labeled ground truth spammers.

Acknowledgments

A very firstly we gladly thanks to my project guide Prof. V.N. Dhakane, for his valuable guidance for implementation of proposed system. We will forever remain a thankful for their excellent as well as polite guidance for preparation of this report. Also we would sincerely like to thank to HOD A.S. Chandgude and other staff for their helpful coordination and support in project work.

REFERENCES

- [1] N. Jindal and B. Liu, "Review spam detection", Proceedings of the 16th international conference on World Wide Web - WWW 07 (2007), ACM, pp. 1189-1190, 2007
- [2] N. Jindal and B. Liu., "Opinion spam and analysis", Proceedings of the international conference on Web search and web data mining - WSDM 08 (2008), ACM, pp. 219-230, 2008.
- [3] 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.
- [4] M. Ott, Y. Choi, C. Cardie and J.T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination", ACM, pp.309-319, 2011.
- [5] P. Rosso, D. Cabrera, M. Gomez, "Using PU Learning to Detect Deceptive Opinion Spam", pp.38-45, 2013.
- [6] W. Zhang, R. Y. K. Lau and Li. Chunping, "Adaptive Big Data Analytics for Deceptive Review Detection in Online Social Media", Thirty Fifth International Conference on Information Systems, Auckland 2014, pp.1-19, 2014.
- [7] 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.
- [8] J. Koven, H. Siadati, and C. Y. Lin, "Finding Valuable Yelp Comments by Personality, Content, Geo, and Anomaly Analysis," 2014 IEEE International Conference on Data Mining Workshop, pp. 1215-1218, 2014.
- [9] S. Rayana and L. Akoglu, "Collective Opinion Spam Detection: Bridging Review Networks and Metadata," SIGKDD, pp. 985-994, 2015
- [10] 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.
- [11] S. Shojaee, A. Azman, M. Murad, N. Sharef and N. Sulaiman, "A Framework for Fake Review Annotation", 2015 17th UKSIM-AMSS International Conference on Modelling and Simulation, IEEE, pp. 153-158, 2015
- [12] 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.
- [13] M. I. Ahsan, T. Nahian, A. A. Kafi, M. I. Hossain, and F. M. Shah, "Review spam detection using active learning," 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016.
- [14] A. Rastogi, M. Mehrotra, "Opinion spam Detection in Online Reviews", Journal of information and Knowledge Management, vol. 16, no. 04, pp. 1-38, 2017.
- [15] W. Etaiwi, G. Naymat, "The impact of applying preprocessing steps on review spam detection", The 8th international conference on emerging ubiquitous system and pervasion networks, Elsevier, pp. 273-279, 2017.