International Journal of Engineering Innovations in Advanced Technology
ISSN: 2582-1431 (Online), Volume-4 Issue-4, December 2022

Machine Learning Based Online Fake Products Review Analysis And Monitoring Using NLTP

Mallipudi Smily
Research scholar (CSE) in International School of Technology and Sciences for Womens (ISTS women's college) in NH-16, East Gonagudem, Rajanagaram, Andhra Pradesh, 533294.
E-mail: smilymallipudi39173@gmail.com

K.CHINNA NAGA RAJU
Associate professor and hod (CSE) in International School of Technology and Sciences for Womens (ISTS women's college) in NH-16, East Gonagudem, Rajanagaram, Andhra Pradesh, 533294.
E-mail: chinna.kuravati@gmail.com

Abstract: Online Shopping is increasing day by day and more people are interested in buying the products of their need from the online stores. This type of shopping takes less time and easy for customer. Customer searches the item of his/her need through online store and place the order. Only by looking at the rating and by reading the reviews related to the particular product customer places the order. Customer takes comments of other people as the source of satisfaction for the new product buyer. Here there is a possibility that the single negative review changes the angle of the customer not to buy that product. So it is possible that one review among multiple reviews is fake. This creates the difficult situation for the customer to read fake reviews and to make a decision whether to buy or not the product. In order to remove this type of fake reviews a, we proposes a Fake Product Review Finding and Reducing System to provide the users with the original reviews and rating for the products. In our proposed system, we can find the given review is genuine or fake so that User can buy a genuine product.

I. INTRODUCTION

One of the very rapid growth areas is ecommerce. Generally e-commerce provides facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services but unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. They share their thoughts on internet. Before purchasing anything, it is a normal human behaviour to do a survey on that product. Based on reviews, customers can compare different brands and can finalize a product of their interest. These online reviews can change the opinion of a customer about the product. If these reviews are true, then this can help the users to select proper product that satisfy their requirements. On the other hand, if the reviews are manipulated or not true then this can mislead user. This boosts us to develop a system which detects fake reviews for a product by using the text and rating property from a review. The honesty value and measure of a fake review will be measured by utilizing the data mining techniques. An algorithm could be used to track customer reviews, through mining topics and sentiment orientation from online customer reviews and will also blocked the fake reviews.

A) Objective of the project

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.

II. LITERATURE SURVEY

[Online product review on shopping experience in social media has promoted users to supply customer feedback. Nowadays many e-commerce sites allow customers to write down their opinion on the merchandise which they buy in the form of reviews or ratings. The reviews given by the customer can build or shatter the great name of theproduct. Thanks to these reason company personnel gets a thought of standing's of their product in the market. In order to demote or promote the merchandise, spiteful reviews or fake reviews, which are deceptive, are posted within the ecommerce site. This result will cause potential financial losses or larger amount of growth in business. We propose a system which mainly focusing on detecting fake and spam reviews by using sentiment analysis and removes out the reviews which have vulgar and curse words and make the e-commerce site fake review free online ISSN NO: 2394-2886 Page No: 16 Suraj Punj Journal For Multidisciplinary Research Volume 11, Issue 4, 2021 shopping center. In this Paper they propose a project which Focuses on detecting fake and spam reviews by using Sentiment analysis and removes out the reviews. Sentimental Analysis Algorithm Users always gets genuine reviews about any of the products. User can spend money on valuable products. [2] In this paper ,they focused on detecting fake and spam reviews by using Sentiment analysis and removes out the reviews J48 Algorithm, Naive Bayes Algorithm. So in this proposed system we will save efforts and time by helping the users and business organizations by identifying spams from different opinions quickly and also help in purchasing their valuable products from a trustworthy site. [3] In this paper ,they focuses on detecting fake review from a set of product reviews by simulating fake reviews that incorporates various types of opinion spam review features and building a training set and then classifying it using Naïve Bayes classification and ensemble classification model like random forest to test the accuracy of the model. [4] The proposed work achieved the accuracy of 87% in detecting fake reviews of written in English by using intelligent learning techniques which is greater than the accuracy of the previous systems. In this paper, they worked with the intention to remove the fake reviews from the original reviews as this is becoming the need of the hour. [5] The main objective of our work is to create a system which will detect spam and redundant reviews and to filter them so that user can collect knowledge about the product. The goal of this paper is to enhance customer satisfaction as well as to make online shopping reliable. [6] In this paper used the methodology for detecting reviews as Commentator Centric Approach-This method is predicated upon the behavior 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 [7] In this paper as an outcome of this, the amount of reviews that a product receives is growing quickly. Most of the products get thousands of reviews at some websites. Now a days any customer can write his opinion text or review, this will attract the individual's attention and organizations to offer undeserving spam opinions to market or to discredit some target products. The prevailing system doesn't restrict spam and invalid reviews and comments. So there's a requirement to develop a sensible system which automatically mine opinions and classify them into spam and non-spam category. [8] In this paper, we select a subset of highly suspicious reviewers for further scrutiny by our user evaluators with the assistance of an internet based spammer evaluation software specially developed for user evaluation experiments. We finally show that the detected spammers have more significant impact on ratings compared with the unhelpful reviewers. It uses. Sentimental Analysis Algorithm, Naive Bayes algorithm. Our results show that our proposed ranking and supervised methods are effective in discovering spammers and outperform other baseline method supported helpfulness votes alone.

III. PROPOSED FRAMEWORK

In our proposed system we find the review is fake or not. If the review is fake the user account will be noticed and give warning to the user instruct the user can create maximum two accounts. From this account only they can purchase and review the products. That reviews is also a fake the account will be blocked. So we can reduce number of fake reviews. So the user can buy a valuable product.

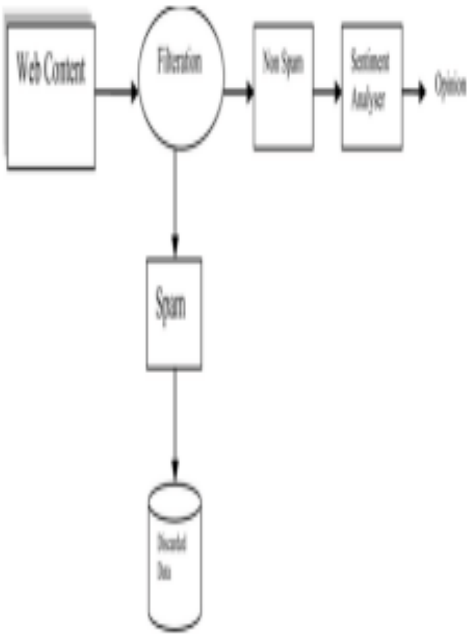

Fig.1 Block Diagram

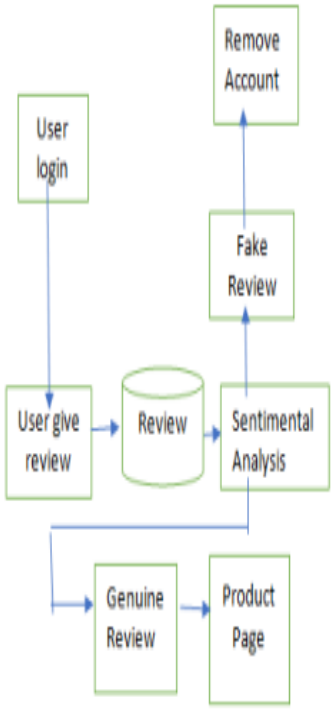

Fig 2 MODULES AND DESCRIPTION

IV. SYSTEM ARCHITECTURE

A system architecture which describes conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal representation of a system, which are organized in a way that supports reasoning about the structures and behaviors of the system. System architecture comprises about system components, the externally visible properties of these components, the relationships (e.g. the behavior) between them. It can provide an idea from which products are often procured, and systems developed, which will work together to implement the general system. There are efforts to formalize languages to explain system.collectively these are called architecture description languages.

V. RELATED WORK

Duhan & Mittal proposed a paper "Opinion Mining Using Ontological Spam Detection" which will help us to find out fake reviews by using as algorithm. To find out fake review in the website this "Fake Product Review Monitoring System" system is introduced. This system will find out fake reviews made by the customers and it will block the users.

$$P(y|x_1,\ldots,x_n) = \frac{P(x_1|y)P(x_2|y)\ldots P(x_n|y)P(y)}{P(x_1)P(x_2)\ldots P(x_n)}$$

$$P(y|x_1,\ldots,x_n) \propto P(y)\prod_{i=1}^{n} P(x_i|y)$$

To find out the review is fake or genuine, we will use some classification such as Tracking IP address of the user to detect if the reviews are from a Spammer. If multiple reviews are from the same IP address then the Reviews are considered Spam. Using Account Used to check whether the reviews are done using the same account. Brand only Review detection i.e. whether the reviews are on only Brand not the product. It's not helpful to consider only the Brand value to judge a product. Using Negative Dictionary i.e. the negative words are identified in the review. If there are more than five Negative Words then the review is a Spam. For instance, a user has posted a Review: "This product

is not good; the design is bad, quality is worst and it is worthless to buy." Here, this sentence consists of 4-5 negative words. So, the system will check the count of negative words, if the count exceeds, then it will be considered as spam review. Therefore Negative Word Dictionary will be used with customized Senti strength algorithm. According to this approach, probability of given review to be Spam is more so it will be considered a Spam. Using Ontology: For instance, if the review posted on a product is not about that product but talking about something else then ontology is used to identify and classify such reviews as spam. If Class: Toshiba Context: Laptop Review: Dell is not so good. Here User is Posting Reviews about Laptop that comes under the class Toshiba. But his Review contains Dell Keyword. In order to identify this Review as Spam we are going to use Ontology. This system uses data mining methodology and Opinion mining technology. This system helps the user to find out correct review of the product, will also help the user to detect fake review and makes them to block the fake reviews automatically.

Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining

Kohli, Mishra & Gupta proposed a paper "Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining" which help us in detecting the fake reviews and track down the user. 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 bot allowing him to share his opinions again. System works as follows: Admin will add products to the system. User need to enter their email id and OTP no to enter the system User once access the system, user can view product and can post review about the product. For posting reviews, the user's id will be verified. And admin will also block the email id of the user if reviews are spammed. Admin will delete the review which is fake. Admin Login: Admin login to the system using his admin ID and password. Add product: Admin will add product to the system. Delete Review: Admin will remove the review which tracked by the system as fake. User Login: User will login to the system using his user ID and password. View product: - User will view product. Post Review: User can post review about the product.

A New Approach for Identifying Manipulated Online Reviews using Decision Tree

Now-a-days an internet has become an essential thing, as it provides more facilities to its users. There are many social networking sites which offer users to share their views. People share their thoughts about politics, social issues as well as about different products. It is a common practice today that before purchasing anything user checks the reviews of that product online. There are multiple sites which deal with these reviews. They provide ratings for the products as well as show comparison between different products. Some enterprises attempt to create fake reviews to affect customer behaviours and increase their sales. But, how to identify those fake reviews is a difficult task for customers. In today's world of competition it is necessary for any enterprise to maintain its reputation in a market. So it is necessary for both, i.e. enterprise and customer to identify manipulated reviews. This paper studies different approaches for identifying manipulated reviews and proposes a new approach to identify those manipulated reviews using Decision Tree (DT).

A study on Review Manipulation Classification using Decision Tree

Identifying review manipulation has become one of hot research issues in e-commerce because more and more customers make their purchase decisions based on some personal comments from virtual communities and e-business websites. Customers consider these personal reviews are more reliable than the existing internet advertisements. Consequently, some enterprises attempt to create fake personal comments to affect customer behaviours and increase their sales. But, how to identify those manipulated reviews is a difficult task for customers. Therefore, this study employs Decision Tree (DT) to improve the classification performance of review manipulation by introducing eight potential review manipulation attributes. In addition, we attempted to discover the important factors of identifying manipulated reviews using correlation analysis and extracted knowledge rules. Finally, a real case of online users' comments

regarding smart phones has been employed to testify the effectiveness of the proposed method.

Multiple Aspect ranking using the Good Grief Algorithm

We address the problem of analyzing multiple related opinions in a text. For instance, in a restaurant review such opinions may include food, ambience and service. We formulate this task as a multiple aspect ranking problem, where the goal is to produce a set of numerical scores, one for each aspect. We present an algorithm that jointly learns ranking models for individual aspects by modelling the dependencies between assigned ranks. This algorithm guides the prediction of individual rankers by analyzing meta-relations between opinions, such as agreement and contrast. We prove that our agreement based joint model is more expressive than individual ranking models. Our empirical results further confirm the strength of the model: the algorithm provides significant improvement over both individual rankers and a state-of-the-art joint ranking model.

VI. CONCLUSION

From our work we've come to a conclusion that finding the fake review from huge amount of unstructured data has become a crucial research problem. Although, a number of the algorithms are utilized in review analysis gives good results, but still no algorithm can resolve all the challenges and difficulties faced by today's generation. it's vital to think about certain quality measures like helpfulness, usefulness and utility while analyzing each review. From the literature survey there are many sophisticated methods given intimately which uses the sentiment analysis with reference to different aspects. Our application which can help the user to buy the proper product with none stepping into any scams. Our application will do analysis then post the real reviews on genuine product. In future we might attempt to improve the tactic of calculating the sentiment score of the reviews. Aim of our project is to reinforce customer satisfaction also on make online shopping reliable

VII REFERENCES

[1] Rajashree S. Jadhav, Prof. Deipali V. Gore, "A New Approach for Identifying Manipulated Online Reviews using Decision Tree ". (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2), pp 1447-1450, 2014

[2] Long- Sheng Chen, Jui-Yu Lin, "A study on Review Manipulation Classification using Decision Tree", Kuala Lumpur, Malaysia, pp 3-5, IEEE conference publication, 2013.

[3] Benjamin Snyder and Regina Brazil, "Multiple Aspect ranking using the Good Grief Algorithm "Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology2007.

[4] Ivan Tetovo, "A Joint Model of Text and Aspect Ratings for Sentiment Summarization "Ivan Department of Computer Science University of Illinois at Urbana, 2011

[5] N. Jindal and B. Liu, "Analyzing and detecting review spam," International Conference on Web Search and Data Mining, 2007, pp. 547-552.

[6] N. Jindal and B. Liu, "Opinion spam and analysis," International Conference on Web Search and Data Mining, 2008, pp. 219-230.

AUTHORS PROFILE

Mallipudi Smily completed her B.Tech in 2018 in Computer Science and Engineering and has interest in Software Testing, Data Structures, App Development, DBMS, Techniques For Fake Product Review Monitoring and removal for genuine online products review using opinion mining as a part of research.

K.Chinna Naga raju (M.tech) he is working as an associate professor in ISTS women's engineering college.