International Journal of Research Publication and Reviews

FAKE PRODUCT REVIEW MONITORING SYSTEM

Prof. Archana Burujwale[1] , Sakshi Govind[2] , Nikhil Kadam[2] , Bhavesh Jadhav[2] , Prasad Patil[2]

[1,2]*Computer Departmentt , Genba Sopanrao Moze College Of Engineering , Balewadi, Pune*

A B S T R A C T

Fake review detection and its elimination from the given dataset using different Natural Language Processing (NLP) techniques is important in several aspects. In this article, the fake review dataset is trained by applying two different Machine Learning (ML) models to predict the accuracy of how genuine are the reviews in a given dataset. The rate of fake reviews in Ecommerce industry and even other platforms is increasing when depend on product reviews for the item found online on different websites and applications. The products of the company were trusted before making a purchase. So this fake review problem must be addressed so that these large Ecommerce industries such as Flipkart, Amazon, etc. can rectify this issue so that the fake reviewers and spammers are eliminated to prevent users from losing trust on online shopping platforms. This Fake review detection and its elimination from the given dataset using different Natural Language Processing (NLP) techniques is important in several aspects. In this article, the fake review dataset is trained by applying two different Machine Learning (ML) models to predict the accuracy of how genuine are the reviews in a given dataset. The rate of fake reviews in Ecommerce industry and even other platforms is increasing when depend on product reviews for the item found online on different websites and applications. The products of the company were trusted before making a purchase. So this fake review problem must be addressed so that these large Ecommerce industries such as Flipkart, Amazon, etc. can rectify this issue so that the fake reviewers and spammers are eliminated to prevent users from losing trust on online shopping platforms.

Keywords: Machine learning, Sentiment Analysis, Sentiment Classification. NLP.

1. INTRODUCTION

This is a Prototype Module and developed for Admin Panel.

The elegance with online review posting has grown at a faster rate and people buying almost everything online that gets delivered at their doorsteps. Hence, people are not subject to physically inspect the product when buying online so they drastically unwantedly/wontedly depend on reviews of other buyers this must be made truthful as much as possible so that the buyer is not cheated with fake reviewers or spammers time and again. The problem is simple yet tiring to be accomplished through/read every review to mark it as a fake or ambiguous category this must be done systematically to get to the root of the problem. This problem can be addressed by training an ML model which deals with the review section to flag a particular review as genuine or spam. The interesting thing is spammers who didn't use the product can be caught this way. A spam review or the usage of different customer id can be used to filter review of the product falsely to get a good rating of the product. This can be filtered by checking the use of words like "awesome", "so good", "fantastic" etc. can be flagged. Since they tend to hype the product or they try to emulate genuine reviews with the same words using it again and again to make an impact on the buyer. Hence the issue of spam filtering requires huge data to train and be effective with added domain knowledge such as sarcasm sentences used by users to show their dissent towards the product, sometimes the product is good but not the delivery or the packing which affects the review classification. Here, an NLP technique is used to identify such reviews instead of misclassification to a negative review as in sentiment analysis. To remove unwanted or outdated product reviews those include data preprocessing.

2. PROBLEM STATEMENT

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 using sentimental analysis, SVM and NLP.

2.1 ALGORITHMS

A) **Sentiment analysis:**

Sentiment analysis is the process of classifying whether a block of text is positive, negative, or, neutral. Sentiment analysis is contextual mining of words which indicates the social sentiment of a brand and also helps the business to determine whether the product which they are manufacturing is going to make a demand in the market or not. The goal which Sentiment analysis tries to gain is to analyze people's opinion in a way that it can help the businesses expand. It focuses not only on polarity (positive, negative & neutral) but also on emotions (happy, sad, angry, etc.). It uses various Natural Language Processing algorithms such as Rule-based, Automatic, and Hybrid.

Sentiment analysis tasks typically combine two different tasks:

1. Identifying sentiment expressions, and

2. Determining the polarity(sometimes called valence) of the expressed sentiment. These tasks are closely related as the purpose of most works is to determine whether a sentence bears a positive or a negative (implicit or explicit) opinion about the target of the sentiment.

B)   SVM (SUPPORT VECTOR MACHINE) :

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.

- Support vector machines are a set of supervised learning methods used for classification, regression, and outliers detection. All of these are common tasks in machine learning.

2.2   WORKING:

Admin:

- **Admin Login**: - Admin login to the system using his admin ID and password.
- **Add product:-** Admin will add product to the system.
- **Delete product**:- Admin will be able to delete the product from the system
- **View Transaction:** Admin can view transaction done by the user.
- **Statistics:**-Admin can determine overall statistics of reviews.
- **Delete Review**:- Admin will remove the review which tracked by the system as fake.

User:

- User once access the system, user can view product and can post review about the product.
- User can purchase the product using Dummy payment process.
- System will check whether the user review is fake or genuine using algorithms.

Result:

The Main objective of this project is to find the Fake Reviews given to the product. User can add the review to product by purchasing or without purchasing the product. User can view the product in Product section and if he likes the product he can add that product to the cart .In the cart the user can directly add the feedback/review without purchasing the product ie. Fake review that can be positive or Negative, the User can buy the product using dummy payment process and after buying the product he/she can state the feedback for the product according to their convenience. i.e. Original and it can be classified as positive or negative by specifying   it's feedback. Using sentiment.

When the user has given the feedback/review, then the admin will check that the feedback is original or fake. And can also be further classified into positive or negative .Admin can also be able to remove the fake feedback/review from the Database.

Analyzing the reviews:

- **Case 1:** User Purchases the Product and gives Positive Review-Original Review (Positive).
- **Case 2:** User Purchases the Product and gives Negative Review-Original Review (Negative).
- **Case 3:** Without Purchasing the Product if the user gives Positive Review-Fake Review (Positive).
- **Case 4:** Without Purchasing the Product if the user gives Negative Review-Fake Review (Negative)

STATISTICAL ANALYSIS OF REVIEW
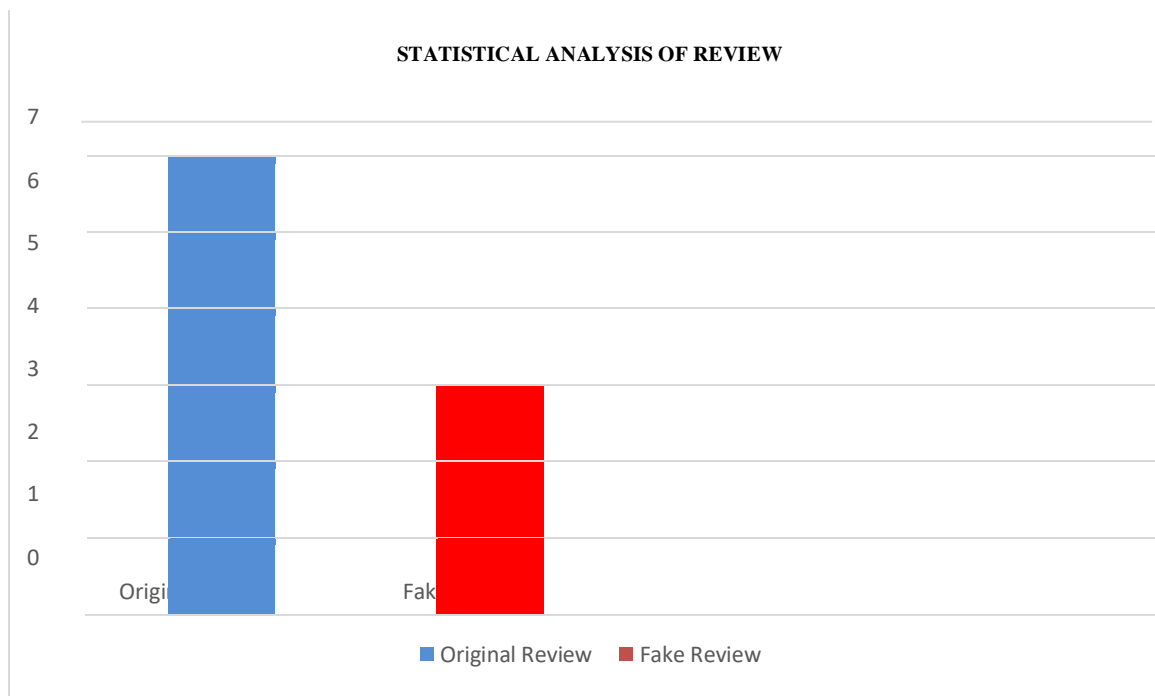
FUTURE SCOPE:

1) As this is a prototype module it can be used in future by companies to design their E-Commerce site.
2) In future we can show the reviews to the User
3) The user can be able to view only the Original Reviews and the Fake reviews would get removed

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 .