International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

MONITORING AND REMOVEL OF FAKE REVIEWS ON PRODUCTS

[1] Bathini Pravalika, [2] YASMEEN, [3] AMATI SANGHAVI

[1] Assistant Professor, Department of CSE, BhojReddy Engineering College for Women, Hyderabad, Telangana, India.

[1] pravalika412@gmail.com

[2,3]Students, Department of CSE, BhojReddy Engineering College for Women, Hyderabad, Telangana, India.

[2] yyaasmeen9@gmail.com , [3] sanghavisanghi22@gmail.com

Abstract

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 bythe 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 manufacturedby 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. To find out the review is fake or genuine, system will find out the IP address of the user if the system observe fake review send by the same IP Address many at times it will inform the admin to remove that review from the system. This system uses data mining methodology. This system helps the user to find out correct review of the product.

I INTRODUCTION

The Internet has vastly changed not only the customer's perspective on buying online but also the business processes. One could say, there are two worlds: one before ecommerce and one after it. Nowadays, customers prefer buying most products or services through e-commerce or online portals. These e-commerce or online portals have given rise to new techniques for marketing as well as influencing customers decision i.e. reviews. Reviews refer to any view or opinion made about a product or service by an individual usually not associated with the business. The reviews that appear on the website are specifically referred to as user generated content (UGC). Reviews present a new way to learn about customer preferences, product quality as well as product's shortcomings. A review left online is a permanent record of that customer's experience; it can be found by anyone and reach a far wider audience than ever before. Today, almost every online portal enables posting reviews, images and expressing our own views about products or services in blogs or forums or dedicated review websites like Zomato, Yelp etc. This user generated content can be used to discover customers" preferences, the strengths and weaknesses of the product, study the market conditions, identify new product launch opportunities and strategize to win from competitors. The easy possibility of monetization using the intelligence obtained from reviews has led to the problem of opinion spam or creation of fake reviews. Companies hire spammers to write undeserving positive reviews to promote their products or negative reviews to destroy the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

competitor's reputation. Unfortunately, driven by the desire for profit or publicity, fraudsters haveproduced deceptive (spam) reviews. There are various reasons that motivate people to write a review, like the desire to affect a change in the business, product or service or anger at poor product / service or delight at a great product / service or when a product / service is not as expected. Thereason could also be an inherent desire to help the public, for instance if the customer is an expertin the product and one would want to share the expertise. Before making any decision about the product, one always first checks the reviews about the product or restaurants or services etc. Positive opinions can result in significant financial gains and/or fame for organizations and individuals. This provides a good incentive for creation of review/opinion spam. Fake reviews canbe written by a shop retailer, business personnel, or individuals who maintain their online identity As the reviews have become an important decision-making factor, some business hire experts to write spam review with the objective/ intention to promote their image or damage the competitor"s reputation. There can be two types of fake review written for this purpose either forged positive review or undeserving negative review to encourage/discourage the customers from purchasing the product.

II LITERATURE SURVEY

Review spam is strenuous to detect unless read manually. Here are some of the works proposed and implemented. Paper [1] proposes behavioral approach to detect review spammers who manipulate the ratings on some target products wherein an aggregated behavior scoring methods for rank reviewers is derived. Paper [2] proposes that spotting the individual fake reviews was quite grueling unlike spotting the groups which was comparatively easier. One frequent item set mining (FIM) method is used to analyze the dataset. In paper [3], the approach was to detect the fake review by identifying the IP address of the user ID that is recorded multiple times. Paper [4] used linguisticfeatures like unigram presence, unigram frequency, bigram presence, bigramfrequency and review length to build a model and find fake reviews. Although, the main problem is data scarcity, and it requires both linguistic features and behavioral features. Paper proposes new features like review density, semantic, and emotion and givesthe model and algorithm to construct each of these features. Although, it is not a good metric, and the reduction is not substantial. In paper [6], scraping processing is used tobuild the data set from yelp and then Fake Feature Framework for organizing theextraction and characterization of features in fake detection. Their framework is composed of two main types of features: review centric and user centric. Review centric features are only related to the text of the review and User centric features show how the user behaves within the site.

III EXISTING SYSTEM

Fake reviews represent only a small proportion of the total reviews, and this causes a data imbalance problem. Addressed the problem of review spammer detection, or finding users who are the source of spam reviews. People who post intentional fake reviews have significantly different behavior than the normal user. They have identified the following deceptive rating and review behaviors.

Problems in Existing System

- Data Imbalance
- Assuring of the quality of the reviews is difficult.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

- Labeled data points to train the classifier is difficult to obtain.
- Human are poor in labeling reviews as fake or genuine

IV .PROBLEM SATEMENT

Now any people can write any opinion text or review, this can draw the individuals attention, and organizations to give undeserving spam opinions to promote or to discredit some target products. So there is a need to develop an smart system which automatically mine opinions and classify them into spam and non-spam category. Proposed opinion spam analyzer will automatically classify user opinions into spam or non-spam. This automatic system can be useful to business organizationas well as to customers. Business organization can monitor their product selling by analyzing and understand what the customers are saying about products. Customers can make decision whether he/she should buy or not buy the products. This can helpful to people to purchase valuable productand spend their money on quality product.

V PROPOSED SYSTEM

System will track the IP address of the user. For detection of fake online reviews, we start with raw text data. We have used a dataset which was already labeled by the previous researchers. Weremove unnecessary texts like article and prepositions in the data. Then these text data are converted into numeric data for making them suitable for the classifier. Important and necessary features are extracted and then classification process took place.

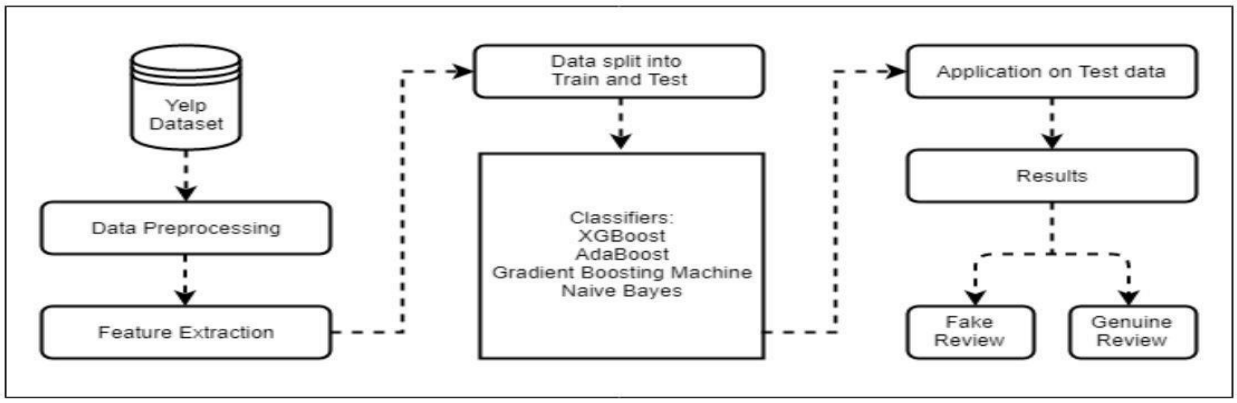
Advantages of Proposed System

- User gets genuine reviews about the product.
- Accuracy of the result.
- User can spend money on valuable products

VI ARCHITECTURE

The Model-View-Controller (MVC) framework is an architectural pattern that separates an application into three main logical components Model, View, and Controller. Hence the abbreviation MVC. Each architecture component is built to handle specific development aspect of an application. MVC separates the business logic and presentation layer from each other. It was traditionally used for desktop graphical user interfaces (GUIs). Nowadays, MVC architecture has become popular for designing web applications as well as mobile apps.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



VII MPLEMENTATION

Admin

Login

- Add Product
- View Product
- Edit Product
- Delete Product
- Find Spam Review
- View Transactions
- Update Transaction status
- Block User
- Delete Review
- logout

User

- Registration
- Login
- Search product
- View Product
- Buy Product
- View Orders
- View Order Status
- Add Review
- Add Rating

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

• Logout

(b) *Detecting by Machine Learning*

1. Dataset:

Dataset used is "amazon academic review" which contains reviews, useful votes, ratings, userid, and many other attributes. The dataset contains thousands of original and fake reviews mixed to easily assess the accuracy of the model being implemented using this dataset.

2. Pre-processing:

Pre-processing is the first step in analyzing any dataset which includes removing unnecessaryattributes, punctuations, stop words, missing words, redundant words , etc. to clean the datasetfor training purposes. This ensures proper training of the model.

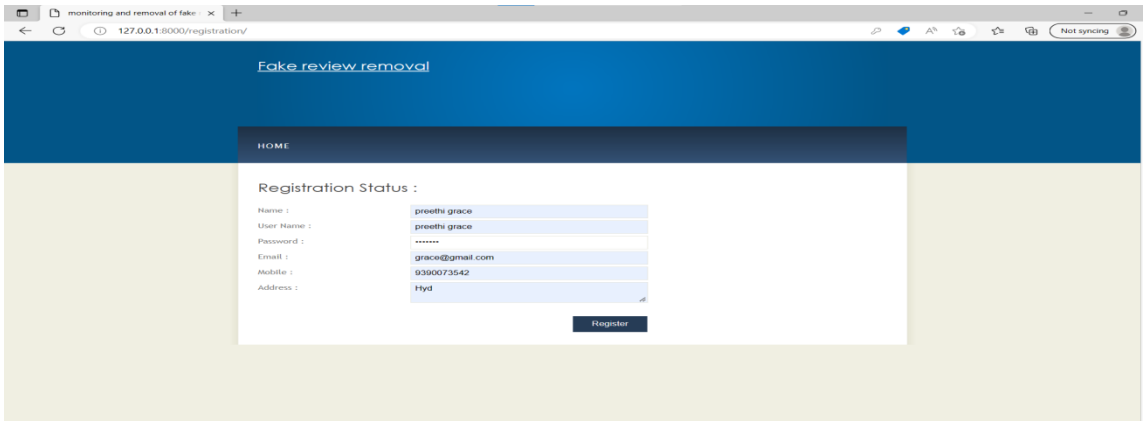
3. Feature Engineering:

This function involves all the methods to remove unwanted information from the dataset it is also called data cleaning. This step is very necessary to find the gaps and the relationship between the different attributes

4. Sampling of data:

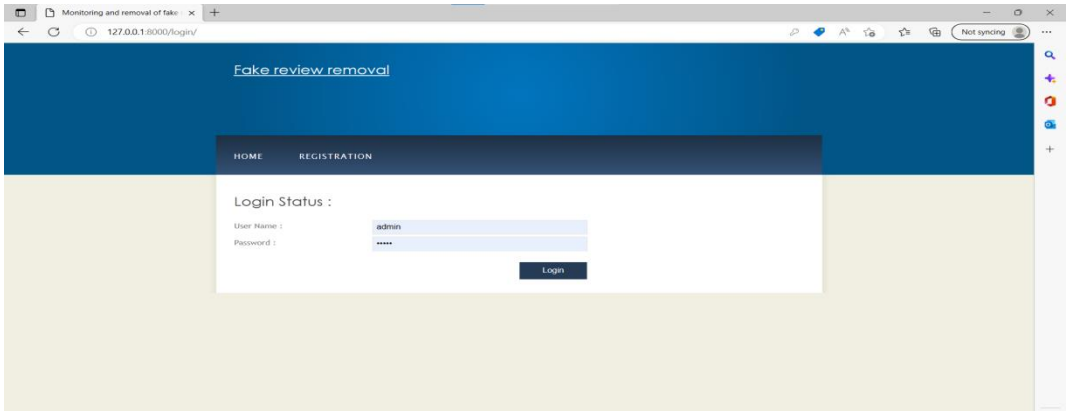
Here, different labels are used to authentic the fake reviews and then concatenate two columnsafter labelling and return the data frame.

VIII RESULTS

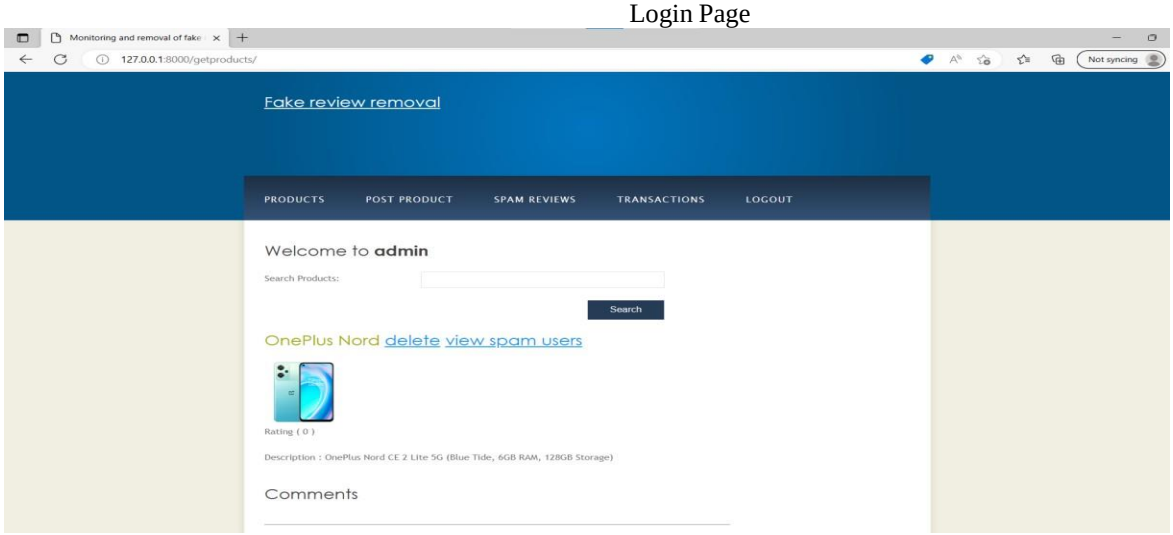


Registration Page

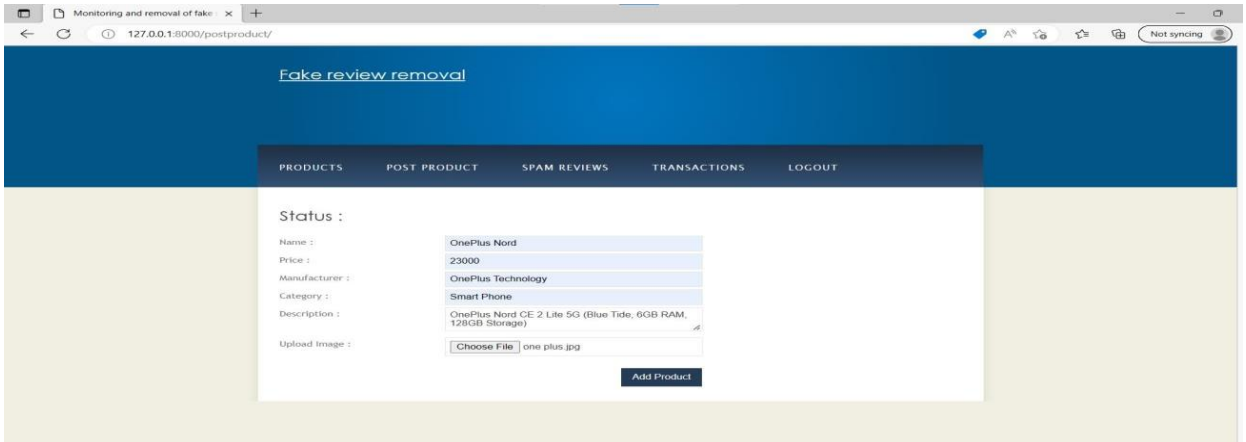
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



Login Page

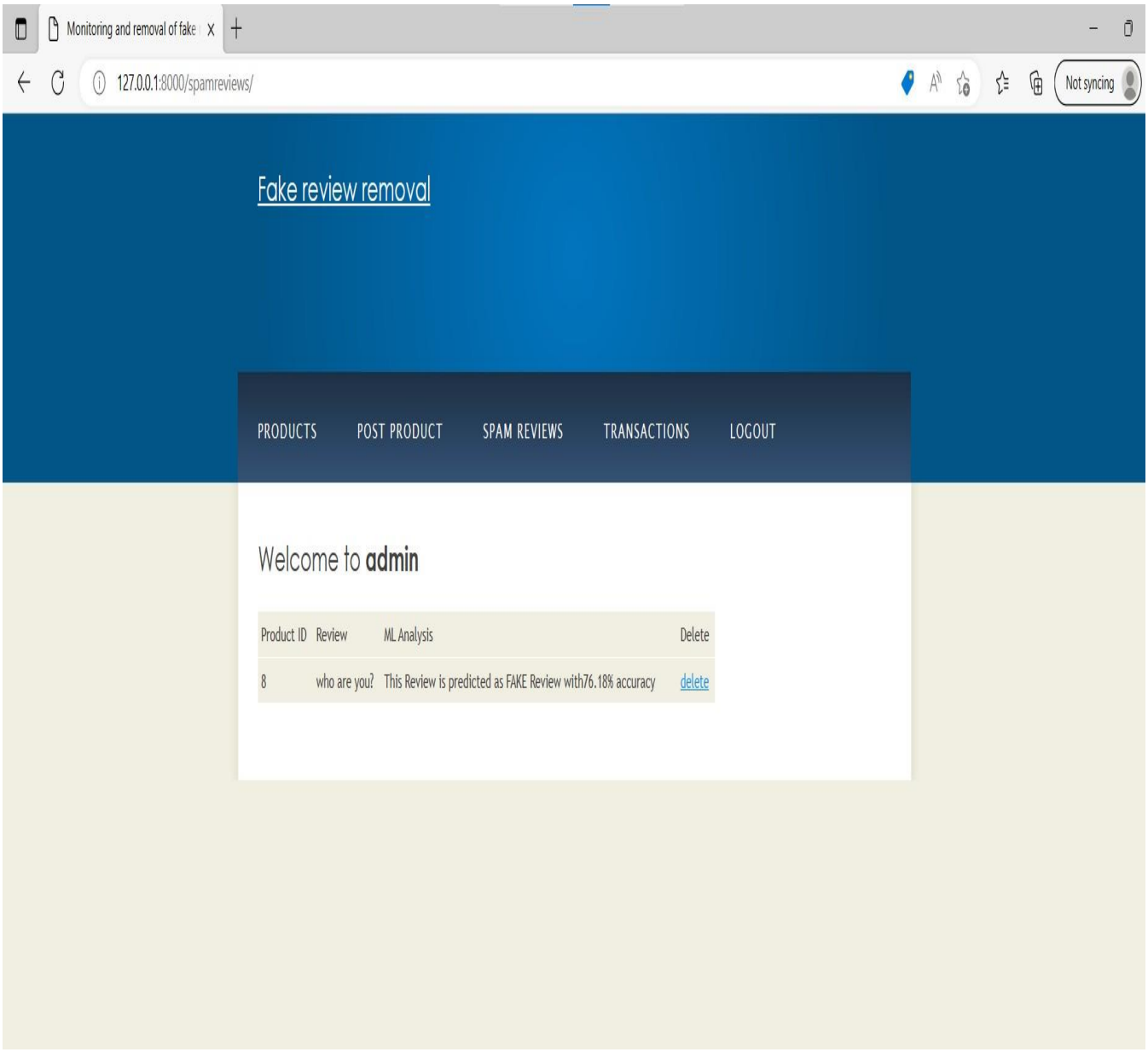


Admin Login Page

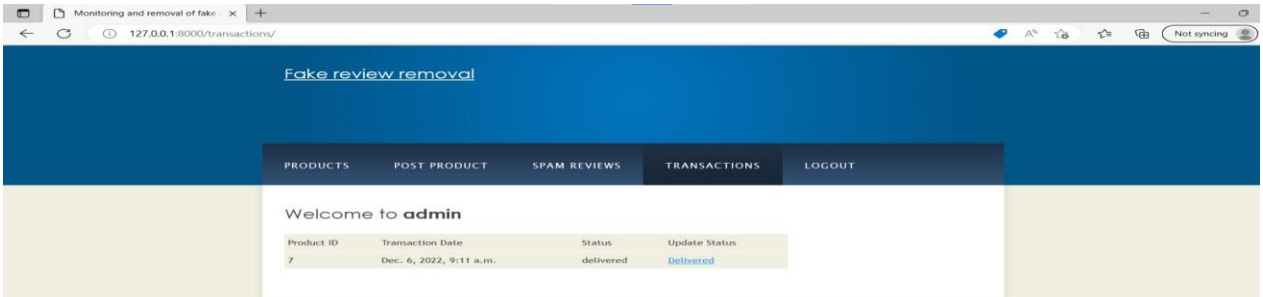


Add Product

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

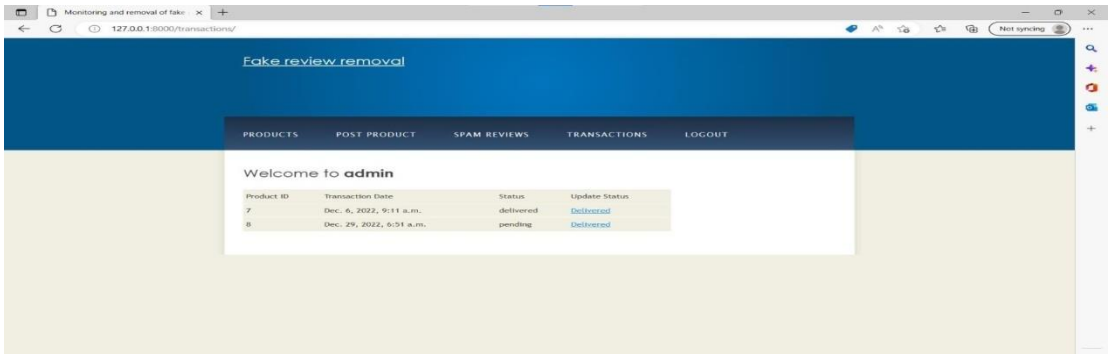


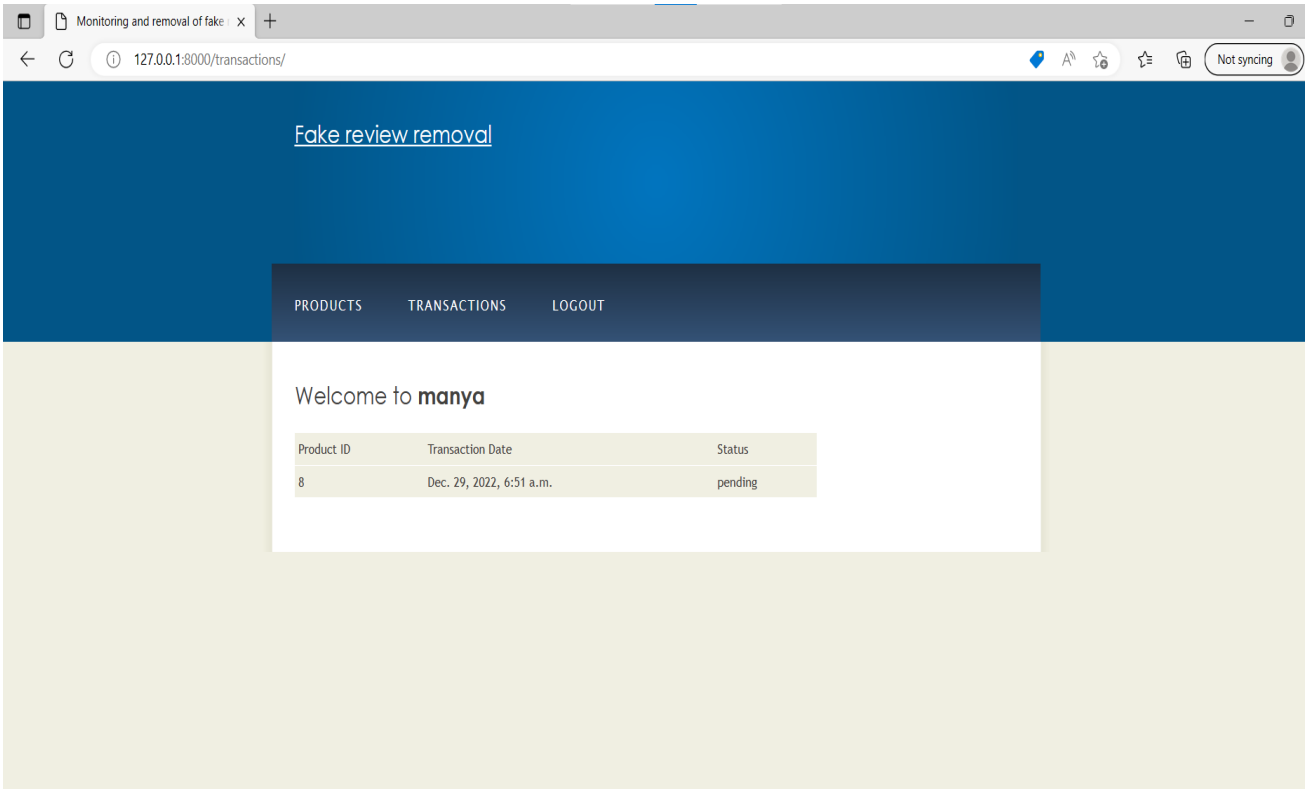
User spam Reviews



International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

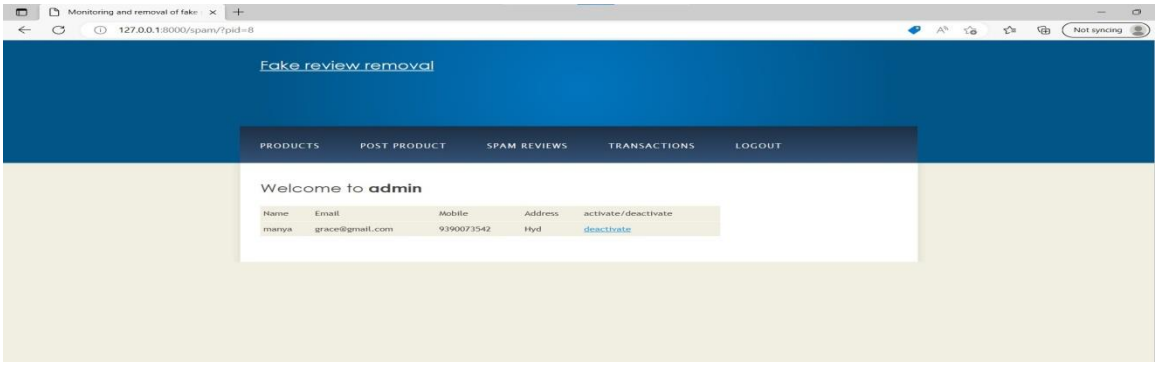
User View Status



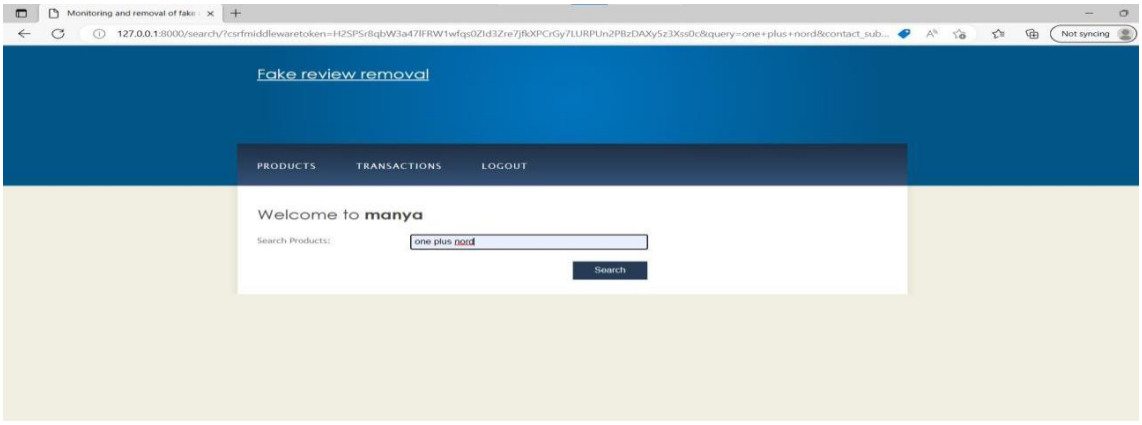


Admin Update transactions

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



Block Users



User Search Products

IX CONCLUSION

From our work we have come to a conclusion that finding the opinion spam from huge amount of unstructured data has become an important research problem. Although, some of the algorithms have been used in opinion spam analysis gives good results, but still no algorithm can resolve all the challenges and difficulties faced by today's generation. It is veryimportant to consider certain quality measures like helpfulness, usefulness and utility while analysing each review. In the literature survey there are many sophisticated methods explainedwhich defines the sentiment analysis with respect to different aspects. Our application which will help the user to pay for the right product without any getting into any scams. Our application will do analysis and then post the genuine reviews on genuine product. And user can be sure about the products availability on that application and reviews too. In future we would try to improve the method of calculating the sentiment score of the reviews. We wouldalso try to update our dictionary containing sentiment word. We would try to add more wordsin our dictionary and update the weights given to those words to get more accurate calculated score of the reviews. Sentiment analysis or opinion mining can be applied for any new applications which follow data mining rules. A direction for future research is to implement the system and check

performance by applying proposed approach to various benchmark datasets. The main objective of our work is to create a system which will detect spam and redundantreviews and to filter them so that user correct knowledge about the product. Aim of our projectis to enhance customer satisfaction as well as to make online shopping reliable

REFERENCES

[1] Chengai Sun, Qiaolin Du, and Gang Tian "Exploiting Product Related Review Features forFake Review Detection" College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

[2] Eka Dyar Wahyuni and Arif Djunaidy," Fake Review Detection from a Product Review using Modified Method of Iterative Computation Framework" MATEC Web of Conferences BISSTECH 2015

[3] Kaushik Varadha Rajan, Nivasse Ajagane, Shubham Srivastav, "Evaluating Performance of Semi-Supervised Self Training in Identifying Fake Reviews," North Carolina State University, Raleigh NC 27606, USA

[4] Shashank Kumar Chauhan, Anupam Goel, Prafull Goel, Avishkar Chauhan and Mahendra K Gurve "Research on Product Review Analysis and Spam Review Detection" 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)

[5] Shivaprasad TK Jyothi Shetty "Sentiment Analysis of Product Reviews: A Review" International Conference on Inventive Communication (ICICCT 2017)