DEDUCTION AND REMOVAL OF UNAUTHORIZED PRODUCT REVIEW TO IMPROVE PRODUCT GENUINESS

[1]S.M.Suvetha,　[2]P.Radha,　[3]P.Mathan Kumar

[1,3] PG Scholar, Department of Computer Science, Sri Krishna arts and Science College,

[2] Assistant Professor, Department of Computer Science, Sri Krishna Arts and Science College,

ABSTRACT

Many e-commerce companies through online shopping are not able to know whether the customers are satisfied with the services provided by the company. So people find various reviews on the website, but these reviews are not recognized by the user as genuine or fake. On some websites, some good reviews are added by the product company people to generate false-positive product reviews. They provide good reviews for many different products made by their own company. Sometimes a competitor will give a false review to damage other company's products. Identifying and controlling these types of fake reviews is a very difficult task today. In this study, we propose to automatically monitor and control these fake reviews. This system will detect fake reviews made by posting fake comments about a product by identifying the IP address through review posting methods and OTP verification. And also it will remove the negative comments which are posted by the user. This system will check if the user has already purchased this product and he is eligible to provide feedback or review. Sometimes the user may be careful to give a fake. This application checks the number of review posts by user behavior and OTP verification. This system uses data mining and OTP (once password) verification. And to remove the fake comments here Naive Bayes algorithm has also been used. This system helps the user to find the right review about the product.

Keywords- *Naive Bayes algorithm, Fake reviews, IP address, Data mining.*

I. INTRODUCTION

Because most people need to review a product before they spend their money on it. So people find various reviews on the website, but these reviews are not recognized by the user as genuine or fake. On some review websites, some good reviews are added by the product company people to make the product popular, these people belong to the social media optimization team. They give good reviews for various products made by their own company. The user cannot find out if the review is genuine or fake. This "Introduction of Monitoring and Removal of Fake Product Review for Real Online Product Reviews Using Feedback Mining" has been introduced to detect fake reviews on the website. This system will detect fake reviews provided by the social media optimization team by identifying the IP address. The user logs in to the computer using his user ID and password to view various products and review the product. If the system repeatedly detects a fake review sent by the computer with the same IP address, the system will detect the user's IP address and notify the administrator to remove the review from the system. This system uses data mining. This system helps the user to find the right review about the product.

The system works as follows: -Admin will add products to the system. Admin will remove the fake review. Once the user has accessed the computer, the user can view the product and post a review about the product. The system monitors the user's IP address. If the system notices multiple duplicate reviews from the same IP address, this IP address will be monitored by the system and the administrator will be notified to remove this review from the system.

II. LITERATURE REVIEW

The concept mine has attracted a great deal of research before. However, no large-scale work has been done in this field. Review spam is very difficult to detect if not read manually. Here are some of the tasks proposed and implemented. Paper [1] Review proposes three types of new features, such as density, semantics, and emotion, and provides a model and methodology for developing each of these features. However, this is not a good metric and the reduction is not significant. Paper [2], linguistic features such as Unigram Balance, Unigram Frequency, Pigram Balance, Pigram Frequency, and Review Length have been used to create a model and detect fake reviews. However, the main problem is the lack of data and it requires both linguistic features and behavioral features. The paper [3] proposes a behavioral approach to detecting review spammers attempting to manipulate ratings in certain target products. Integrated Behavior Scoring Methods for Rating Assessors Paper [4] proposes to use online spam reviews to identify types of lexical semantic and linguistic features. We found that finding fake dummy reviews on paper [5] can be a daunting task, but finding groups is relatively easy. The paper [6] first performed a comparison using real-life filtered (fake) and unfiltered (non-filtered) reviews. The results showed that real-life data was very difficult to classify, with an accuracy of only 67.8%.

III. SYSTEM STUDY

System study is classified into two types
- Existing system and
- Proposed system

EXISTING SYSTEM

The company previously proposed to build only has a standard website. They do not have a custom online ordering system. They handle Grade I customers in the industry and deliver network products to them. The company draws customer orders through cool calls and direct appointments. They need an advanced system to receive customer orders immediately and retain them for a long time as the business grows.

But this online order booking exists in many fields, so according to the discussion with the corporate people, they have a new system that will facilitate the choice of customers and feel better with their

competitors. The new system was proposed with the association mine from the above study.

Defects of the existing system

Man High human power utilization

Satisfy does not meet customer satisfaction requirements.

As a pioneering company, they need a new system to operate in the environment.

Future Manual work does not apply as the system grows in the future.

Proposed system

The newly created system will be implemented on the company's website. Customers interested in order booking can book their orders from this module. Based on the customer's needs and preferences, the products will be categorized on the customer page. You can order their products from that customer.

This application monitors each customer in an individual session during this particular process (online orders). Their options and choices are logged into the database. The proposed mining algorithm then filters the missing orders and shows them their options instead of searching the entire database. Other activities implemented in this program will be ranked by the sales of the products and will uniquely identify their customers.

Advantages of the proposed system

- Company The forerunner of online orders for this company will increase the use of manpower
- The mining process manages customer orders and evaluates them when they go to product review
- Sales Product Sales creates an independent database and allows only ordered customers to post reviews for their products
- The email notification mechanism implemented in this program upgrades the administrator to send orders and approvals for purchases to customers.
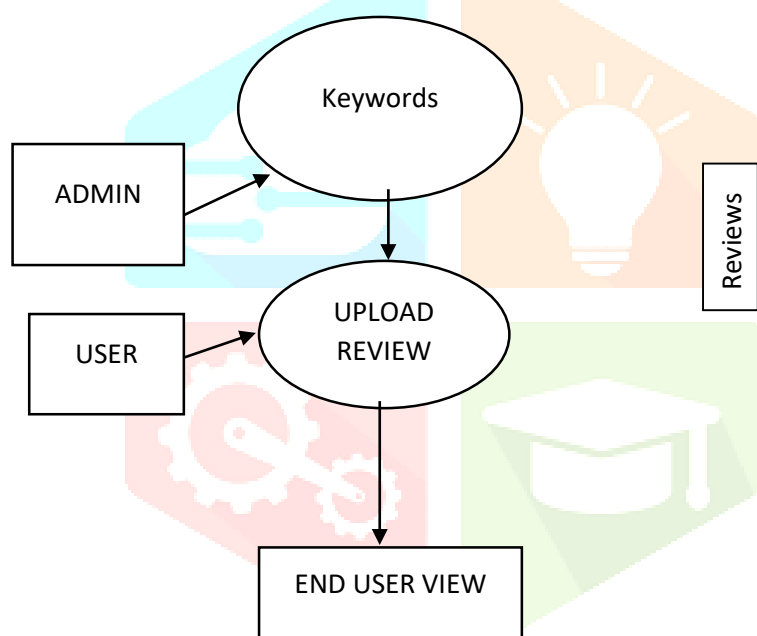
IV. SYSTEM DESIGN

In recent times the online shopping is increasing day by day. And the reviews for the online shopping was not having a genuine review. This application sends the One time Password (OTP) validation method using Data Mining Methodology to check user behavior through it. So here only registered users who brought the product can give the review. And using the user's IP address giving multiple reviews on the same product can be prevented. For finding the sentimental analysis naïve Bayes algorithm has been used. The classifier has been used and is for classifying words and labeled as negative and

positive. So that given keywords will be get trained in the dataset. If the user gave negative reviews the admin will be notified. So that genuineness can be improved with the use of this system.

V. IMPLEMENTATION METHODOLOGY

Only the registered users can use this web application. And the users only who brought the product will receive the OTP for giving the review. So if the user's given review matches means only it will allow the user to give the review. Once the user enters the review it will be start analyzed. It will analyze using the Naive Bayes algorithm whether the given review has any negative words in it. If it found any negative words then it automatically removes the words in it. And if a user had given multiple reviews on the same product it will be found out through the IP address and it will bring notice to the admin.
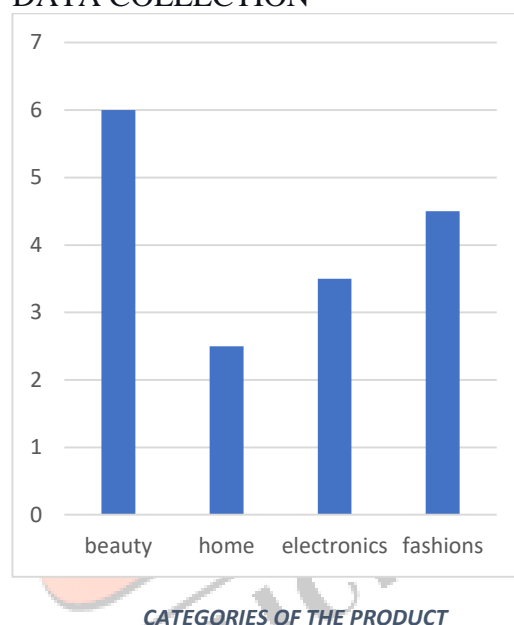

WORKING OF THE SYSTEM

User Upload Review: To ensure authentic reviews the customer whom ordered from the website alone can give the reviews for the particular ordered product and OTP verification which confirms the user identity before they send review about products. And any illegal posts are uploaded by the uploader, and when the user gave negative feedback, the system will automatically control the uploader ID and remove the uploader from the database. The system will verify through their IP configuration.

Add keywords: Adding filtered illegal words and some passwords for security purposes. The system will analyze the keyword with the administrator's security keywords and provide alerts for loading regarding the security issue.

End-User View: User can see without illegal words, and with the genuine reviews. So this module is helpful for both uploader and end-user.

DATA COLLECTION

CATEGORIES OF THE PRODUCT

The product reviews represented here are belongs to different categories such as beauty, home, electronics, fashions. Those online reviews were posted by the users for 1546 products. Each reviews in this contains Id, product Id, IP address, time of the review, date of the review.

ALGORITHM

Data mining is the discovery of hidden relationships in data, which is part of a broader process called "knowledge discovery". Describes the steps that need to be taken to ensure that knowledge discovery achieves meaningful results. The use of data mining tools does not eliminate the need for business knowledge, data understanding, or statistical methods. It also lacks clear forms of knowledge. Current methods of data analysis, which operate based on reviewing the origin or annual statistical maps, have several limitations to predict the performance and availability of the

areas produced. So today, researchers are coming to use superior forms to maximize availability. In recent years, data processing has become one of the most common methods for processing and detecting hidden patterns. Provides new fields for equipment manufacturing and industry such as data warehouses, data processing. So by using these tools, companies can achieve competitive advantages. In particular, through data processing - the extraction of hidden information from large databases - companies can predict future behaviour and make informed decisions.

VI. CONCLUSION

The System is easy to use and implement. It can also be used to track the Ip addresses thus preventing false information regarding the product. And it is used to prevent the negative words using the Naive Bayes Algorithm. The OTP validation method was also used. So in this, the project can determine genuineness in the review. The major advantage of the project is that it is easy to use and implement. The analyzing process of the reviews in the system can be tracked easily by the admin.

VII. FUTURE SCOPE

Will be able to use advanced algorithm ARM for Distributed databases, Hardware enhanced ARM is added in this application for working in a Distributed Environment. And the result would yield output for multiple branches. So that in the future we will do further different investigate of features to make accurate predictions.

VIII. REFERENCE

[1] Yuejun Li, Xiao Feng, ShuwuZhang,"Detecting Fake Reviews Utilizing Semantic and Emotion Model" IEEE2016

[2] KolliShivagangadhar, Sagar H, SohanSathyan, Vanipriya C.H ,"Fraud Detection in Online Reviews using Machine Learning Techniques." IJCER-2015.

[3] SushantKokate, Bharat Tidke, "Fake Review and Brand Spam Detection using J48 Classifier." IJCSIT, 2015.

[4] Amir Karam, Bin Zhou, "Online Review Spam Detection by New Linguistic Features." UNIVERSITY OF MARYLAND BALTIMORE COUNTY"15

[5] Kolhe N.M., Joshi M.M.Jadhav A.B., Abhang P.D., "Fake Reviewer Groups" Detection System." IOSR-JCE 2014.

[6] Arjun Mukherjee, VivekVenkataraman, Bing Liu, NatalieGlance,"Fake Review Detection:Classification and Analysis of Real and Pseudo-Reviews."\-CS"13