International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 8 ǁ August 2022 ǁ PP. 548-555

Fake Contextual Information Detection And Removal In E-Commerce Applications

A. Arthi #1
PG Student
Department of Computer Applications
UCE, Anna University, BIT campus ,Tiruchirappalli

Dr.S. Geetha #2
Assistant Professor (Sr. Gr)
Department of Computer Applications
UCE, Anna University, BIT campus, Tiruchirappalli

Abstract— *As most of the people require review a pair of product before spending their money on the merchandise. So people bump into various reviews within the web site but these reviews are genuine or fake isn't identified by the user. Positive reviews of a target object may attract more customers and increase sales; negative review of a target object may cause lesser demand and reduce in sales. These fake/fraudulent reviews are deliberately written to trick potential customers so as to promote/hype them or defame their reputations. User won't be able to learn whether the review is genuine or fake. to go looking out fake review within the web site this "Fake Contextual Information Detection and Removal in E-Commerce Applications" system is introduced. this system will discover fake reviews made by posting fake comments a pair of product by identifying the MAC address and Booking ID along with review posting patterns. User will login to the system using his user id and password and may view various products and might give review about the merchandise. to hunt out the review is fake or genuine, system will learn the MAC address of the user if the system observes fake review send by the identical MAC Address many times it'll inform the admin to induce eliminate that review from the system. This method uses data processing methodology. Recommender systems help the purchasers to hunt out accurate product from an outsized database. This method helps the user to hunt out correct review of the merchandise.*

Keywords—*Fake Review, MAC Address (Media Access Control), Text Mining ,Recommender System.*

Date of Submission: 14-08-2022 Date of acceptance: 29-08-2022

I. INTRODUCTION

The Internet has changed the method of trying to find information and, thus, has shaped our shopping behaviour. it's now not just an area of information but offers an interactive platform for any purpose of exchanging individual experience with products or services. E-commerce (electronic commerce) plays a significant role during this world, for buying and selling of products and services, over an system, primarily the web. Consequently, suppliers not retain exclusive authority over the offering of knowledge. In some review websites some good reviews are added by the merchandise company people itself so on create so on provide false positive product reviews. they provide good reviews for several different products manufactured by their own firm. Online reviews play a awfully important role in today's e-commerce for decision-making. Large an element of the population i.e. customers read reviews of products or stores before making the selection of what or from where to buy for and whether to buy for or not. As writing fake/fraudulent reviews comes with monetary gain, there has been an oversized increase in deceptive opinion spam on online review websites. This aims to style online shopping application to investigate fake contextual information posted by online users with identifying the user id and merchandise id together with review posting patterns and to recommend products for users supported sentiments. For recommendation, admin train a database which has sentiment based keywords with positivity or negativity weight. In Section II i've got to clarify Literature Survey for future circumstances, In section III I have got to clarify Research methodology for designing the web application fake review detection system. In section IV it's very useful for the superior nation and Most of the people are using the E-commerce applications.

II. LITERATURE REVIEW

In this section, we briefly review several related works on fake review detection, including classification methods, approaches.

A. Syed Mohammed Anas and Santoshi Kumari: Define Random forests model performed well compared to the Naïve Bayes algorithm by a large margin .This study conducts Naïve Bayes and random forest method.By

applying these models one can know the number of spam reviews on a website or application instantly. To counter such spammers, a sophisticated model is required in which a need to be trained on millions of reviews.
B. Jianrong Yao, Yuan Zheng and Hui Jiang: The performance of the ensemble model improves significantly on two representative datasets. In the present study, Majority voting strategy should be implemented. First select the predictive features that need. In this study,divide all the features into two categories: review-centric and reviewer-centric features. Most of the reviewer-centric features are already contained in the dataset. For textual features, we need several stages (text preprocessing) and tools (NLTK, etc.) before extracting features from the content of reviews to attain improved input data. Next, we select five supervised machine learning classifiers as the base classifiers, and three steps: data resampling , feature pruning, and parameter optimization are applied to optimize the base classifiers. Finally, we apply two ensemble strategies to integrate the base classifiers.

C. Ahmed M. Elmogy1,Usman Tariq2, Ammar Mohammed3 , Atef Ibrahim4: This paper applies several machine learning classifiers to identify fake reviews based on the content of the reviews as well as several extracted features from the reviewers. They apply the classifiers on real corpus of reviews taken. Besides the normal natural language processing on the corpus to extract and feed the features of the reviews to the classifiers, the paper also applies several features engineering on the corpus to extract various behaviors of the reviewers. This paper impact of extracted features of the reviewers if they are taken into consideration within the classifiers.The results in the absence and the presence of the extracted features in two different language models namely TF-IDF with bi-grams and TF-IDF with tri-grams. The results indicates that the engineered features increase the performance of fake reviews detection process.

D. Rami Mohawesh 1, Shuxiang Xu 1, Son N. Tran 1, Robert Ollington1,Mathew Springer 1, Yaser Jararweh 2, and Sumbal Maqsood1 : Developing Algorithms That Can Work Efficiently In The Real World. They use a Robustly Optimized Bert Approach (Roberta).Whether it is known as spam review detection, fake opinion detection, and spam opinion detection, the main problem associated with fake review detection is classifying the review as either fake or genuine. Machine learning plays a significant role in fake review detection . For example, supervised learning is one of the popular tasks in fake review detection, which requires labelled data to classify the fake review from genuine review based on specific features. Distinguishing a fake review from a truthful one by reading a large number of reviews is very difficult. Machine learning methods can separate fake reviews from genuine ones by revealing text hidden patterns that the human eye cannot recognize. Existing work of fake review detection can be classified according to their detecting an individual spammer, a group of spammers or fake reviews in one, mix and cross-domain . It is worth mentioning that this paper covers various techniques for fake review detection in natural language processing. As such, it is mainly focused on English language reviews, their related problems, their datasets, and their applications.

III. RESEARCH AND METHODOLOGY

The Research method begins with designing the online shopping application to analyze fake contextual information posted by online users with identifying the MAC address of the user system ,user id and Booking id along with review posting patterns and also recommend products for users based on sentiments.

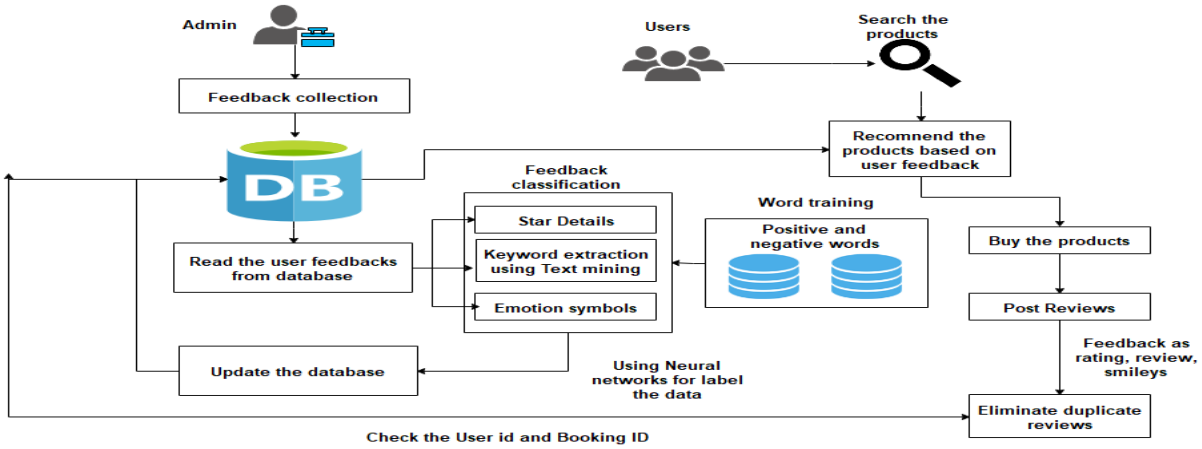

Fig 1: Architectural Diagram

To developing this online e-commerce framework, Initially the admin can login by their ID and post the products along with this price, product details. For the first time user register their personal details .User can search the products and buy it. After the shipping users will post their reviews. If the user once post their review,

couldn't be repost their reviews. So that the In this part has been checked by the Users Id, Booking Id and then MAC Address of the system. These feedbacks are collected by the admin .These feedbacks are read and classified by using the Text mining Methodology.

A. Text Mining

Data preprocessing is a critical activity , It's the major one for processing the raw data. Mining is the process of finding the relevant and appropriate data from the large dataset. A sequence of preprocessing steps have been used in this work to prepare the raw data for computational activities.

1) Tokenization: Tokenization is one of the most common natural language processing techniques. It is a basic and first step before applying any other preprocessing techniques. The text is divided into individual words called tokens. For example, if we have a sentence ("I like Travelling"), tokenization will divide it into the following tokens ("I", "like" , "travelling").

2)Token normalization: Token normalization is the process of canonicalizing tokens so that matches occur despite superficial differences in the character sequences of the tokens. It's used to cleanse the noise of the unstructured data.The most standard way to normalize is to implicitly create equivalence classes, which are normally named after one member of the set.It helps to improve the efficiency.

3) Stop Words Removal: Stop words are the words which are used the most yet they hold no value. Examples of the stop words are (the, this, that, there). In this paper, all data are cleansed from stop words before going forward in the fake reviews detection and removal process.

4) Stemming Words: Stemming is basically removing the suffix from a word and reduce it to its root word. Stemming is an important part of the data pre-processing process in Natural language processing. The input to the stemmer is tokenized words. After the stemming remove the special characters of the text.

B. Recommendetation Algorithm

To make recommendations, algorithms use a profile of the customer's preferences and an outline of an item (genre, product type, color, word length) to figure out the similarity of things using cosine and Euclidean distances. For recommendation, admin train a database which has sentiment based keywords with positivity or negativity weight. Implement Hybrid filtering technique in recommendation system with feedback analysis to enhance the recommendation system.

1)Hybrid Filtering Technique: Hybrid filters combine a variety of passive and/or active filters and their structure is also of series or parallel topology or a mix of the two.
Initialize Di, Ri, Ei .Set Pp = 0 and Pn=0 and Ki = 0 .
Read ratings R, reviews D and emotions D of every products.
If the rating Ri is beyond 5 star or 4 star, consider as positive Pp, otherwise consider as negative Pn.
Read the words from reviews datasets Di, Consider words as keywords Ki .
Match the keywords Ki with training datasets.
Labelled the review Di as "positive" and also labelled review Di as negative supported on training words .
Read the emotions Ei from datasets .If the symbol is happy symbol labelled as positive Pp, otherwise labelled as negative Pn .

IV. PROPOSED WORK

Proposed a recommendation system using hybrid filtering which is implemented by stochastic learning algorithm together with hybrid feedback. These feedbacks help the users to predict whether the item is positive or negative to buy. Eliminates the fake review posting for a product by the purchased user supported user id and merchandise id, from this genuine products are identified.It takes in consideration booking Id, user Id and MAC Address of the System while posting the reviews. This proposed model achieves accurate predictions for a product which benefits the e-commerce user to buy the net product or not.
Advantages:

Proposed System helps the user to find out correct review of the product.

- Handle large number of contextual information.
- User easily buy the genuine products.
- Recommend the positive products to trust users.
- Automatic decision making system in product recommendation.

- Helps to eliminate fake review posting using user identification.

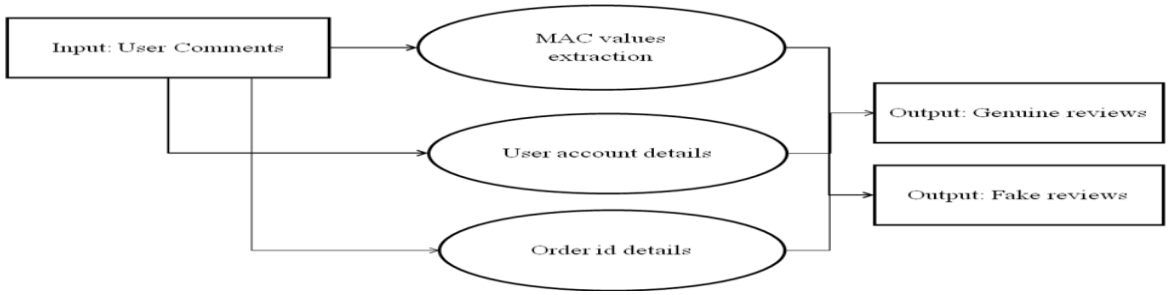

Fig 2: Detecting and removal the fake reviews

In this diagram briefly describes the users reviews are analyzed that is Genuine or Fraudulent review. Reviews detected by the MAC address of the users system , User Id ,Account Details, Booking Id Details.

V. IMPLEMENTATION

Customers are going to be ready to make selection of whether or not they will purchase the products. This will beneficial to the human beings to urge precious product and also spend their money on excellent products.

1. User are allowed to review providing he's logged into our online portal.
2. After logging in user are going to be allowed to review for the merchandise.
3. Once the user post their review, then the reviews are processed and analyzed for spam supported the subsequent
conditions:
a) Analyzing whether multiple reviews have come from the identical user.
b) Analyze whether identical Product Booking Id or identical MAC address are used for multiple reviews on the identical product.
c) Analyze the ratings or reviews to detect whether the reviews are spam or not.
4. If the review uploaded by the user satisfies any of the above specified conditions then it'll be considered as spam or fake reviews.
5. Once the review is detected as spam review or fake review, then it will be removed and review are reported to the administrator.
6. Recommend the products supported the analyze of the user review.

A. Online E-Commerce framework

This module is used to create web site and buy or post products for users. There are two accounts such as admin and user account. Admin login to the framework utilizing their Administrator ID and mystery key. Admin can login to the system and post products with features. User can login to the system to view product details. Admin can categorize the products based on type, gender and so on. User can submit audit approximately the item.

B. Review Collection

Admin collect reviews and have various types of reviews.
Reviews may be rating reviews, text reviews and smiley reviews. All reviews are stored in database for future evaluation. Ratings are in the form of star values. Reviews may be unigrams, bigrams or n-grams. Smiley specifies the symbols of happy and sad.

C. Sentimental Analysis

Sentiment analysis gives significant information for decision making in various domains .Admin can analyze whether the product is positive or negative. In star rating, we can calculate star count values. In text reviews, extract keywords and matched with database. Then smiley reviews are calculated based happy and sad symbols.

D. Recommendation System

User can search the product in search bar. And view the list of products based on features and reviews. Using support vector machine, recommend the products based on product categorization.If the product has negative review means, automatically the positive products in recommendation panel.

E. Eliminate the fake Reviews

In this module, fake reviews are analyzed by admin. Admin can get user account details, MAC address and Order id details. So one user can post one reviews that will be genuine reviews. If the framework reveals an audit is phony it will light up the administrator to expel the phony survey.A novel implementation of a mobile product recommendation based on hybrid filtering with feedback analysis to improve the recommendation system. This feedback includes Ratings, reviews and emoticons are analysed for a product and categorized the product such as positive or negative for the customers to purchase the product. MAC based filtering approach can be used to avoid fake reviews to be posted.

VI. RESULTS AND DISCUSSIONS

In the testing of Application has running well. The users can post their reviews by their Booking Id.The reviews are in the form of star Ratings, Text feedback, and then Emoticons.These values are extracted and calculated for the Recommender system. If the user give the fake Booking Id,it will display the text push up into the back page .By analyzing these user reviews it will recommend the products based on their reviews .

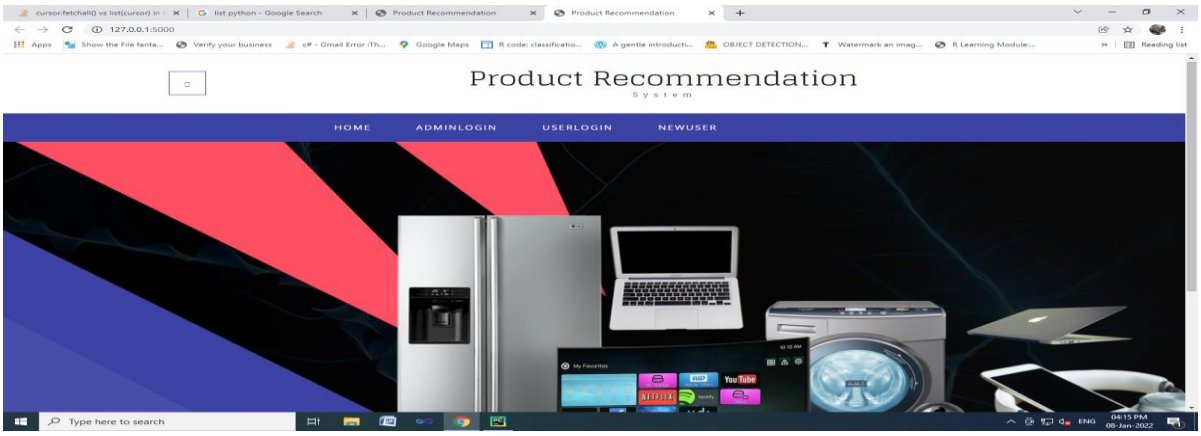

Fig 3: Home page

The Figure 3 displays the homepage of the Fake contextual Information detection and removal website.

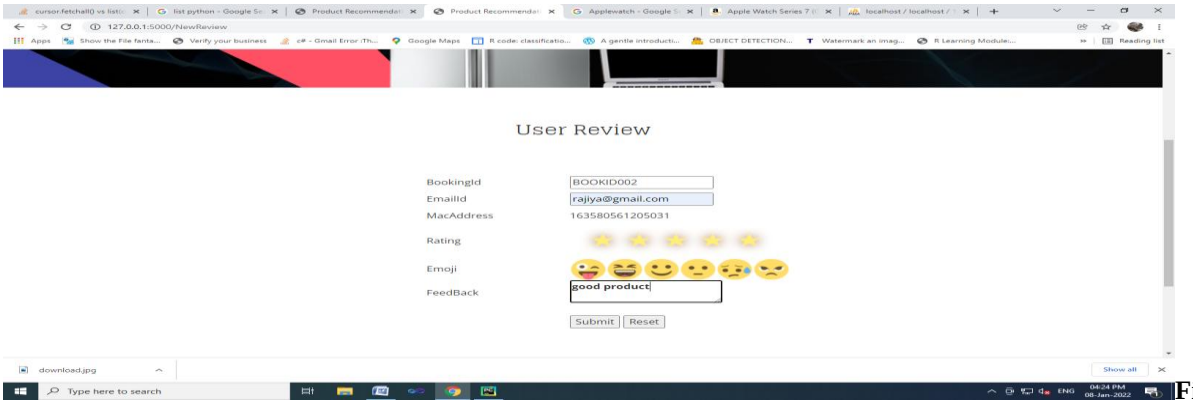

Fig 4 :Users Review submission

Figure 4 describes, the Users should posting their reviews by their user Id, Id and the reviews are in the form of star ratings, Emotions ,and then the feedback text box. After enter the review users can click the submit button comment review enter successfully will be display.

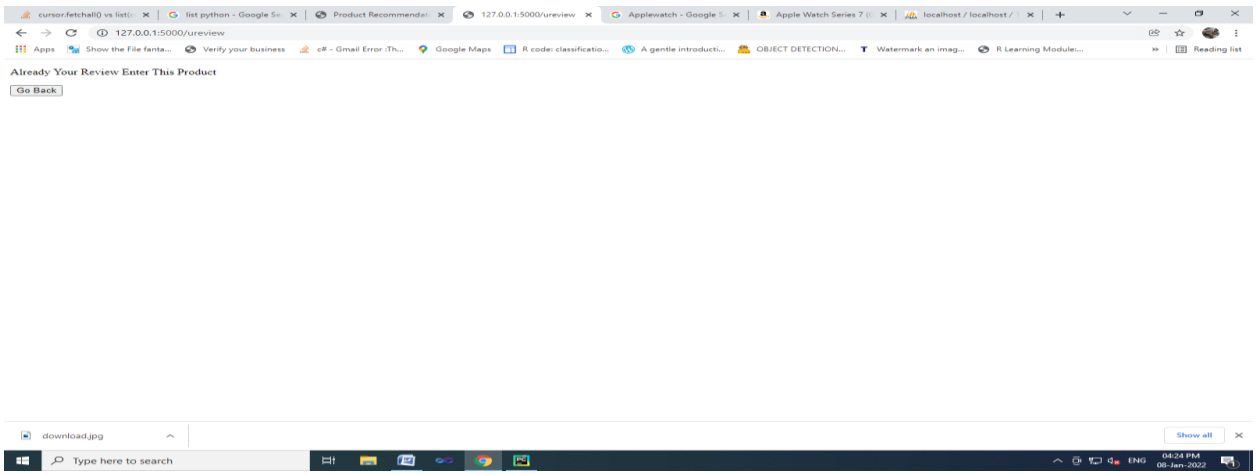

Fig 5: Detecting the fake review

In this Figure 5 says to detect the fake review reviews. Fake reviews should be considered as user should only post the review at one time .

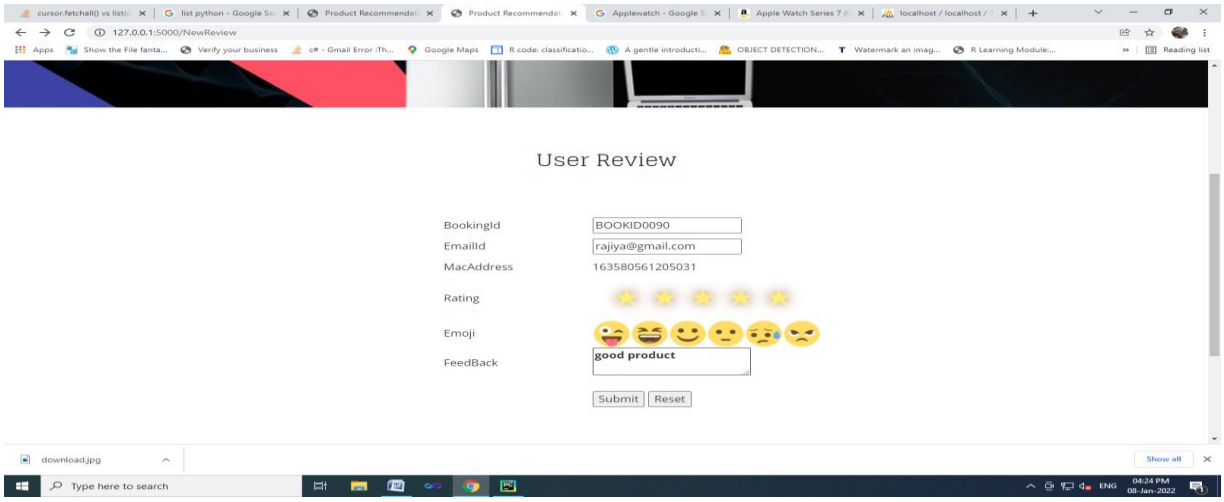

Fig 6: Detecting fake Reviews

In Figure 6 describes If the user could give another fake Booking id ,for posting review that cannot allow to post review.

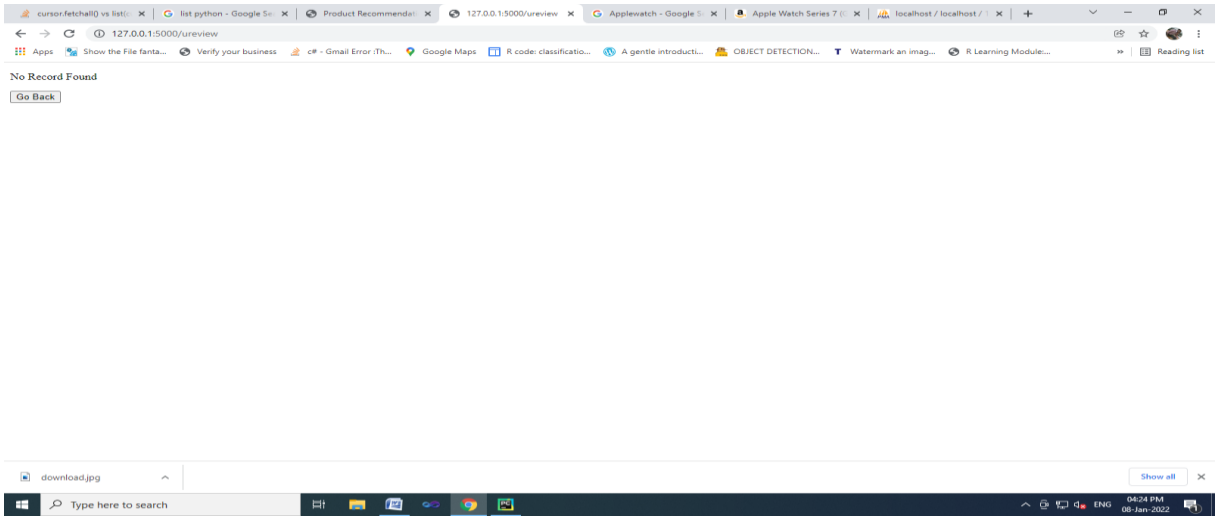

Fig 7: Detecting fake Reviews

In this Figure 7 describes By posting the review using fraud booking id ,it will displays the no data found message.

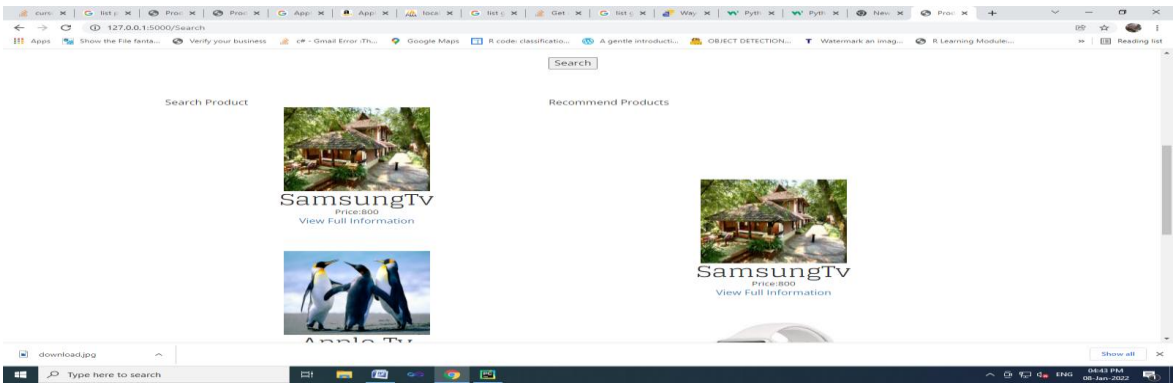

Fig 8: Recommend the products

In this Figure 8 Displays the product recommendation .After that the fake review detection and removal process the system can recommend the genuine products for the users. So that users will shop only the genuine products or goods. By analyzing the review what kind of products the users needed.The user review are analyzed by the admin, these data are collected for the recommendation.

VII. CONCLUSIONS

In this paper, we showed the importance of reviews and how they affect almost everything related to web based data. The proposed framework presented a product recommendation system based on hybrid filtering recommendation algorithm. The main advantages of this framework is to provide a genuine products for the user. This framework also provide a simple method to search the products anywhere and anytime. Ratings, reviews and emoticons are analyzed and categorized as positive and negative comments. Hybrid Recommendations is one of the main modules of the system which helps overcome the drawbacks of the traditional Collaborative and Content Based Recommendations.

VIII. FUTURE SCOPE

In the proposed framework implemented a detecting the reviews by the Star ratings,smileys,text feedback.For future research,we can add a image review processing. So that users can easily identify the genuine products.

ACKNOWLEDGMENT

We would like to thank our guide, Dr.S.Geetha Mam Assistant Professor, University College of Engineering ,BIT Campus, Tiruchirapalli , who guided us throughout this project and helped in developing the research paper.

REFERENCES

[1]. Anas, Syed Mohammed, and Santoshi Kumari. "Opinion Mining based Fake Product review Monitoring and Removal System." 2021 6th International Conference on Inventive Computation Technologies (ICICT). IEEE, 2021.
[2]. Yao, Jianrong, Yuan Zheng, and Hui Jiang. "An Ensemble Model for Fake Online Review Detection Based on Data Resampling, Feature Pruning, and Parameter Optimization." IEEE Access 9 (2021): 16914-16927.
[3]. Elmogy, Ahmed M., et al. "Fake Reviews Detection using Supervised Machine Learning." Int. J. Adv. Comput. Sci. Appl 12 (2021).
[4]. Mohawesh, Rami, et al. "Fake Reviews Detection: A Survey." IEEE Access 9 (2021): 65771-65802.
[5]. Hassan, Rakibul, and Md Rabiul Islam. "A Supervised Machine Learning Approach to Detect Fake Online Reviews." 2020 23rd International Conference on Computer and Information Technology (ICCIT). IEEE, 2020.
[6]. Mupparam Sowjanya, K.Shnati latha,Ch.hyma,K.Naresh "Implementation of fake product review monitoring system and real review generation by using data mining mechanism" Journal of Xi'an University of Architecture & Technology, 2020.
[7]. Joni Salminen, Chandrashekhar Kandpal, Ahmed Mohamed Kamel, Soon-gyo Jung, Bernard J. Jansen," Creating and detecting fake reviews of online products", Journal of Retailing and Consumer Services(2022).
[8]. Piyush Jain, Karan Chheda, Mihir Jain, Prachiti Lade, "Fake Product Review Monitoring System", International Journal of Trend in Scientific Research and Development (IJTSRD), Volume 3, Issue 3, pp. 105-107, Mar-Apr 2019.
[9]. Guangxia Xu, Mengxiao Hu, Chuang Ma, Mahmoud Daneshmand, "GSCPM: CPM-based group spamming detection in online product reviews", 2019.
[10]. Ata-Ur-Rehman, Nazir M. Danish, Sarfraz M. Tanzeel, Nasir Usama, Aslam Muhammad, Martinez-Enriquez A. M., Adrees Muhammad, "Intelligent Interface for Fake Product Review Monitoring and Removal", 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) Mexico City, Mexico, September 11-13, 2019.
[11]. Nour Jnoub, Wolfgang Klas, "Declarative Programming Approach for Fake Review Detection", 2020.
[12]. M.Abi Priya, S.Hema, R.Dhivya Praba," Fake Product Review Monitoring and Removal for Genuine Online Product Review Using IP Address Tracking", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,2020

[13]. Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., & Chau, M. (2004). Crime data mining: a general framework and some examples. *computer*, *37*(4), 50-56.
[14]. Sinha, A., Arora, N., Singh, S., Cheema, M., & Nazir, A. (2018). Fake Product Review Monitoring Using Opinion Mining. International Journal of Pure and Applied Mathematics, *119*(12), 13203-13209.