Intelligent Interface for Fake Product Review Monitoring and Removal

1st K.SANTHANA LAKSHMI.,M.TECH
dept of computer science and Engineering
Paavai Engineering college
Namakkal, India

2rd MANI PRABHU T
dept of computer science and Engineering
Paavai Engineering College
Namakkal, India

3nd MOHAN M
dept of computer science and Engineering
Paavai Engineering College
Namakkal, India

4th ARUN KUMAR P
dept of computer science and Engineering
Paavai Engineering College
Namakkal, India

***Abstract*— The trend of online shopping is increasing day by day and many people want to buy their required products from online stores. This type of shopping does not take much of the customer's time. A customer goes to the online store, finds what they need, and makes an order. But what people have problems with when buying products from online stores is the poor quality of the products. Customers place orders simply by looking at prices and reading reviews related to specific products. Such comments from other people are a source of satisfaction for new product buyers. Here, a single negative review can turn the corner for customers not to buy the product. In this case, this review may be fake. Therefore, we have proposed Finite Resource Finder Counterfeit Product Review and Monitoring System (FaRMS) as a smart interface to remove this type of fake reviews and provide users with genuine reviews and product-related ratings. (URL) is associated with Amazon, Flipkart and Daraz products and analyzes reviews and provides customers with real prices. A unique feature of the system is that it works with three e-commerce websites and offers reviews not only in English but also in Urdu and Roman Urdu. Previous work on fake reviews did not support the feature of analyzing reviews written in languages like Urdu and Roman Urdu and failed to capture reviews on some e-commerce websites. The proposed work achieves an accuracy of 87% in detecting fake reviews written in English using intelligent learning methods, which is higher than the previous system's accuracy.**

I.INTRODUCTION

There are several ways to shop so you can go to a store or mall and buy what you need. In this type of shopping, the seller gives a review of the product, you don't know whether it is a fake or genuine review. This is because you should review the product carefully because of the seller's own honesty, how true the seller is to his word, and the fact that you have no other choice to review the product. If you don't care to buy the product, it might be a waste for you. On the other hand, today the source of trade has changed. You can buy products from online stores of various brands. You must place an order here without viewing or reviewing the original product. You read reviews and buy products. So you depend on product reviews. This review can be genuine or fake. A customer wants to buy a unique and reliable product, which is possible when you get genuine feedback about the product. US shoppers will spend $6 billion on Black Friday 2018 sales, according to research. Americans spend 36 percent of their shopping budget online. In 2017, e-commerce generated revenue of $2.3 trillion and is expected to reach $4.5 trillion by 2021. Today, there are about 12-24 million stores in the world. Research has found that 61% of Amazon reviews in the electronics category are fake. There are several websites that are used to detect fake reviews. Akeasama is an online website that identifies fake reviews using suspicious patterns and reviewer activity. As in the process of buying a product from an online store, you should read all the reviews one by one to check the quality of the product and buy a good quality product. This is very time consuming. Reviews here can be fake or genuine.

II. LITERATURE SURVEY

Fake product monitoring and removal is an essential task for e-commerce platforms to ensure customer satisfaction and prevent fraudulent activities. An intelligent interface can aid in automating this process and provide real-time monitoring of products. In this literature survey, we will explore different research papers and articles related to intelligent interfaces for fake product monitoring and removal.

1."An Automatic System for Online Product Review Monitoring" by Y. Hu, X. Zhu, and K. YuThe paper proposes an automatic system for monitoring online product reviews to detect fake reviews. The system uses natural language processing (NLP) techniques to analyze the sentiment of reviews and identify potential fake reviews. The proposed system achieved high accuracy in detecting fake reviews, which can aid in monitoring fake products.

2."A Machine Learning Approach to Identifying Fake Products on E-Commerce Platforms" by X. Wang, Z. Wang, and W. LuThe paper presents a machine learning-based approach to identify fake products on e-commerce platforms. The authors use a combination of supervised and unsupervised learning techniques to classify products as fake or genuine.

3."Detecting Fake Online Product Reviews Using Machine Learning Techniques" by R. Anand and V. A. NarayananThe paper proposes a machine learning-based approach to detect fake online product reviews. The authors use NLP techniques to extract features from reviews and train a machine learning model to classify reviews as genuine or fake. The proposed approach achieved high accuracy in detecting fake reviews, which can aid in monitoring fake products

4."A Deep Learning Based Framework for Fake Product Detection in E Commerce" by H. Xu, H. Cheng, and Y. LiuThe paper proposes a deep learning-based framework for detecting fake products in e-commerce. The authors use convolutional neural networks (CNNs) to extract features from product images and classify products as fake or genuine. The proposed framework achieved high accuracy in identifying fake products, which can aid in their removal
5."Fake Review Detection using Machine Learning Techniques: A Review" by S. S. Raman, S. T. Jothi, and S. SivakumarThe article provides an overview of different machine learning techniques used for fake review detection. The authors compare the performance of different techniques, including supervised and unsupervised learning, and suggest that a combination of techniques can provide better results. The review can aid in choosing the appropriate technique for fake product monitoring.
The literature survey highlights different approaches and techniques used for fake product monitoring and removal. NLP techniques and machine learning-based approaches are commonly used to detect fake reviews and products. Deep learning techniques, such as CNNs, can aid in identifying fake products based on product images. A combination of different techniques can provide better results in fake product monitoring.

Fake products with opinion mine:
On e-commerce websites or Flipkart, Amazon, Snap deal etc. In this paper, we propose a framework using fake product reviews or spam reviews. Mine thought. Opinion Mining is also called Sentiment Analysis. In sentiment analysis, we try to find out what the customer thinks about a piece of text. First, we review and verify that the review is related to a specific product using the Decision Tree. We use the Spam Dictionary to define the word spam in reviews. In Text Mining we use several algorithms and based on these algorithms we get specific results.

Black Friday Purchases:
Laptops are top sellers on Black Friday, with many people taking advantage of the savings to buy new devices. Other top sellers are games such as God of War and Go Go Pikachu, and children's toys such as Fingerlings. It's a reminder that while Black Friday is more often seen as a day for adults to buy big-ticket items like electronics and gadgets, the kids' toy market still makes up a big part of sales.
Research has found that 61 percent of electronic reviews on Amazon are fake:
The trend of online shopping is increasing day by day and many people want to buy their required products from online stores. This type of shopping does not take much of the customer's time. A customer goes to the online store, finds what they need, and places an order. But the thing that people find difficult when buying products from online stores is the poor quality of the products. Customers place orders simply by looking at ratings and reading reviews related to specific products. Such comments from other people are a source of satisfaction for new product buyers. Here, a single negative review can turn the corner of the customer not buying the product. In this case, this review may be fake. Therefore, to remove this type of fake reviews and provide users with genuine reviews and product-related ratings, we have proposed a Fake Product Monitoring and Removal System (FaRMS) that takes an End-to-end Resource Locator as an interface. URL) related to Amazon, Flipkart products and analyze reviews and provide actual prices to customers. The proposed work achieves an accuracy of 87% in detecting fake reviews written in English using intelligent learning methods, which is higher than the previous system's accuracy.

State Sentiment Analysis:
The whole world is changing rapidly in today's innovation. The Internet has become a basic requirement for everyone and the Web is used in every field. With the rapid development of social networking applications, people are using these platforms to express their views on everyday issues. Public goods, public services, etc. it is important to collect and analyze their reaction to buy. Sentiment analysis (or sentiment mining) is a general discourse analysis task that aims to identify the sentiment of ideas in the text on various topics. In recent years, sentiment analysis researchers have become interested in analyzing sentiment in a variety of topics, including movies, commercial products, and everyday social issues.

Fake Reviewer Groups in Customer Reviews:
Social media such as product reviews are now widely used by individuals and organizations to make decisions. However, for profit or fame, people try to game the system by spamming (eg writing fake reviews) to promote or demote certain targeted products. Such spam reviews should be found for reviews reflecting real user experiences and opinions. Previous work on opinion spam focused on detecting fake reviews and individual fake reviewers. However, fake reviewer groups (groups of reviewers who work together to write fake reviews) are even more dangerous because they can completely control the sentiment on the target product due to their size. This paper explores spam detection in a collaborative environment to detect fake watchdog groups. The proposed method first uses frequent feature extraction to find a set of candidate clusters. Then using some behavioral models derived from the association between pseudo-observer and relational model based on relationships.

III. THE EXISTING SYSTEM

A system where you can find genuine reviews and ratings related to specific products, then it's a source of satisfaction and reliability for you. In the proposed method, reviews related to a given product URL are retrieved. The system then detects fake reviews and by analyzing this review system, finds genuine reviews of the product. Previous research has detected fake reviews using several approaches, including deterministic addressing, opinion mining and sentiment analysis, and machine learning approaches. That is why we have proposed a Fake Product Monitoring and Removal System (FaRMS) where customers can get the best products from online stores in no time with genuine reviews related to those products. This system provides genuine testimonials with real reviews from people related to the product. Some popular products may have hundreds of reviews on various major shopping sites, and FaRMS filters out fake reviews and provides promising reviews, which you can then decide to buy or not.
Fake review detection for Yelp was developed to filter out fake reviews from genuine reviews as it is the need of the hour. The proposed classification system takes the text and

other information from the reviews and concludes whether the reviews are trustworthy or not. The dataset used in this project was originally for Yelp.com by Ryan and Akoglu. They used 16282 observations and divided these into 0.7 training sets, 0.2 dev and 0.1 test sets. Extracting predictive features from observations is the most difficult part of the project. They mainly release two types: review-centric and viewer-centric features. First, they calculated the percentage of each unigram and bigram number for false and non-false observations. Then they came up with the top 100 unigrams and bigrams with the most different percentages of fake and non-fake reviews. The second approach leads to better performance because it processes all unigrams and bigrams. They tested several machine learning algorithms, but using the Neural System achieved the highest accuracy of 81.92%. This system is good at detecting fake reviews, but there is a need to improve the accuracy of filtering reviews.

.

IV. THE PROPOSED SYSTEM

The product rating technique is developed to represent a product rating model that uses weights on product rating factors to calculate a product rating score. In this proposed system, there are sentences that are not related to product quality, such as customer service or a sentence related to In this paper, pre-processing is done using Support Vector Machine (SVM). First, it removes comments that are not related or related to the quality of the product. The second phase describes the review weights based on votes. In the last stage, the overall rating of the product is calculated. The review score is calculated based on the relevance of the review to the quality of the product, the content of the review and the date the review was published. They use 10-fold cross-validation on the training set. In the ranking process, they use two measures to quantify the effectiveness of the ranking model, which are as follows: the correlation between the ranking method and the Amazon ranking, and the second is the mean average accuracy (MAP), which is a very commonly used ranking ranking technique. accuracy. Because this system detects fake reviews using only two features of reviews, but according to the future work described in the paper, more features can be used to detect fake reviews more accurately.

Spam review detection using temporal pattern detection is designed to track reviews related to the normal reviewer arrival pattern and fake reviewer arrival pattern and find that the normal reviewer arrival pattern is stable and temporally uncorrelated with their rating pattern. Spam attacks, on the other hand, are usually bursty and either positively or negatively correlated with the rating pattern. The dataset they took is a snapshot of a review website dated October 6, 2010. It contains 408,469 reviews written by 343,629 reviewers, which are written for 25,034 website stores. For each review, they collect the following information such as rating, postdate and whether it is a Spammer Review (SR) or not. In the evaluation process, they select 53 stores, each with more than 1,000 reviews. Human evaluators decide whether trades are SR spam or not, if two or more evaluators declare the trade as SR spam, before the system considers the trade to be fraudulent in selling. Out of 53 stores, 34 are suspicious and the rest are normal. Out of 34 stores, 22 stores have at least two suspicion votes. The recall related to the system is 75.86%, which shows that the system detects most of the trades with the SR spam attack. The accuracy associated with the proposed approach is 61.11%. This proposed system is good in terms of training their model to find the relationship because the model is trained using a large number of reviews included in their dataset.

Previous works detect fake reviews using IP address, opinion mining, sentiment analysis and some of them use machine learning approaches. In some approaches, the dataset is very small, and others use few features related to the reviews to find fake reviews. In the proposed system, a large dataset of English reviews is used to train the model. In this way, the system can more effectively find hidden patterns in reviews. The accuracy achieved by using the proposed technique is greater than the accuracy of previous systems in terms of English overviews. This system also works to detect fake reviews of Urdu and Roman Urdu reviews.

The fake review tracking system focuses on detecting spam and fake reviews using sentiment analysis and removes reviews that contain profanity and profanity. In the proposed system, a web crawler is used to scrap data on the Web. In pre-processing, the data is converted into the required format and then fake reviews are removed from the mix of genuine and spam reviews. Fake reviews are detected by the fake review detector. Each review must be passed through a classifier that calculates the sentiment score of the review. Cosine similarity is used to measure this type of similarity. If the cosine value to be calculated is greater than 0.5, the review is considered false. The system developed found 111 reviews out of 300 to be fake. However, the dataset used to train the model is very small, so suspicious patterns cannot be found more accurately.

V. METHODOLOGY

- Data collection
- Data preprocessing
- Feature Extraction
- Machine Learning Model Training
- Model Evaluation
- Integration with Interface

The methodology used for an intelligent interface for fake product monitoring and removal typically involves the following steps:

Data Collection: The first step is to collect data from various sources, such as e-commerce platforms, social media platforms, and online review sites. The data can include product information, user reviews, and ratings.

Data Preprocessing: Once the data is collected, it needs to be preprocessed to remove any noise or irrelevant information. This step involves data cleaning, data integration, and data transformation.

Feature Extraction: In this step, relevant features are extracted from the preprocessed data, such as product images, product descriptions, user reviews, and ratings. Natural Language Processing (NLP) techniques can be used to extract features from text data.

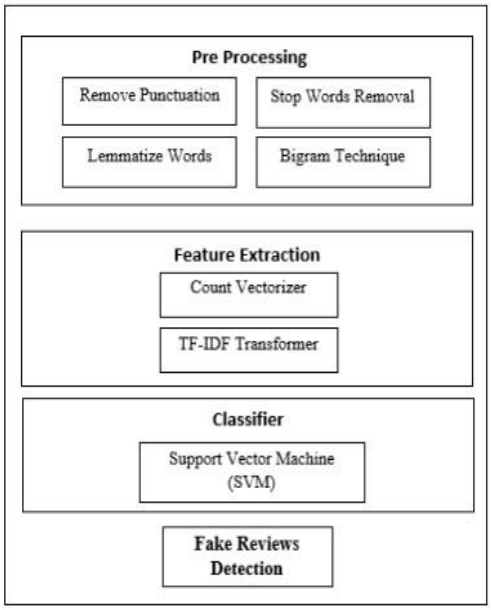
Machine Learning Model Training: The extracted features are then used to train a machine learning model, such as a classification model or a regression model. The model can be trained using supervised or unsupervised learning techniques.

Model Evaluation: Once the model is trained, it needs to be evaluated to determine its accuracy and performance. This step involves testing the model on a separate dataset and comparing the predicted outputs with the actual outputs.

Integration with Interface: The final step is to integrate the machine learning model with the intelligent interface for fake product monitoring and removal. The interface can be designed to provide real-time monitoring of products and automatically remove fake products.

Overall, the methodology used for an intelligent interface for fake product monitoring and removal involves data collection, preprocessing, feature extraction, machine learning model training, model evaluation, and integration with the interface. The choice of specific techniques and algorithms may vary depending on the type and amount of data, as well as the requirements of the interface.

VI . SYSTEM ARCHITECTURE



VII . HARDWARE REQUIREMENTS

Equipment requirements:
The maximum requirements specified by an operating system or software program are physical computer resources, also known as hardware. The list of hardware requirements is often accompanied by a list of hardware compatibility (HCL) in the case of the operating system. A list of tested, compatible, and sometimes incompatible HCLs for specific operating systems or applications. The following sections discuss various aspects of hardware requirements.

Development:
All computer operating systems are designed for a specific computer architecture. Most software is limited to a specific operating system that runs on a specific architecture. Although architecture-independent operating systems and applications exist, many must be redesigned to work with the new architecture. See also the list of common operating systems and supported architectures.

Processing Power:
Central processing unit (CPU) power is a basic system requirement for any software. Most software running on the x86 architecture defines processing power as a CPU model and clock speed. Many other CPU features that affect speed and power, such as bus speed, cache, and MIPS, are often ignored. This definition of power is often misleading because AMD Athlon and Intel Pentium CPUs with similar clock speeds often have different clock speeds. Intel Pentium CPUs are very popular and often mentioned in this category.

minutes:
All software resides in the computer's random access memory (RAM) when it is running. Requirements are determined after reviewing applications, operating systems, software and files, and other operational processes. This requirement also applies to the optimal performance of other software that is not related to running in a multitasking computer system.

Secondary storage:
Storage device requirements vary depending on the size of the software installation, temporary files created and saved when the software is installed or running, and the use of swap space (if RAM is not enough).

Display adapter:
Software that requires a better-than-average PC graphics display, such as high-end graphics editors and games, often specifies a high-end display adapter in the system requirements.

Environment:
Some software applications require extensive and/or specialized use of certain peripherals that require high performance or peripheral functionality. Such peripheral devices include CD-ROM drives, keyboards, pointing devices, network devices, etc. enter

platforms:
A computing platform describes any framework in hardware or software that allows software to run. Common platforms include computer architecture, operating systems or programming languages, and runtime libraries.

The operating system is one of the requirements that are mentioned when determining the system requirements (software). Software may not be compatible with different versions of the same operating system, but some compatibility criteria are generally maintained. For example, most software designed for Microsoft Windows XP does not work on Microsoft Windows 98, but the opposite is not always true. Likewise, software developed using the newer features of Linux Kernel v2.6 usually does not run correctly (or at all) on Linux distributions using Kernel v2.2 or v2.4.

API and Driver:
Software that uses more specialized hardware, such as high-end display adapters, will need special APIs or new device drivers. DirectX is a collection of APIs for solving problems related to multimedia, especially game applications on Microsoft platforms.

Web Browser:
Most web applications and software depend heavily on web technologies to use the default browser installed on the system. Despite its weaknesses, Microsoft Internet Explorer is the most popular type of software that uses ActiveX controls in Microsoft Windows.

Other requirements:

Some software have other requirements for proper operation. Internet connection (browsing and speed) and screen resolution are notable examples.

Examples:

Some examples of the increasing trend of defining system requirements and resource requirements for popular PC games are:

For example, StarCraft (1998):

System requirements [hide]

	Requirements
Windows	
Operating system	Windows 95 or NT or superior
CPU	Pentium processor at 90 MHz or higher
Memory	16 MB RAM
Free space	80 MB available
Media	CD-ROM, 2x or higher
Graphics hardware	DirectX 3.0 or higher

System requirements [hide]

	Requirements
Windows	
Operating system	Windows 2000/XP
CPU	Pentium 4 1.5 GHz or Athlon XP 1500+ processor or higher
Memory	384 MB RAM
Free space	2.2 GB free space
Media	8x Speed CD-ROM
Graphics hardware	3D Hardware Accelerator - 64MB of memory minimum DirectX 9.0b
Sound hardware	DirectX 9.0b compatible 16-bit sound card

System requirements [hide]

	Requirements
Windows	
Operating system	Windows Windows XP SP3, Windows Vista SP2, Windows 7
CPU	Core 2 Duo or Athlon X2 at 2.4 GHz
Memory	2 GB RAM
Free space	8 GB of free space, 23.8 GB + 1 GB Swap File space
Graphics hardware	DirectX 9.0c compatible video card. 3D Hardware Accelerator - 256MB of memory minimum
Sound hardware	DirectX 9.0c compatible sound card

System requirements [hide]

	Requirements
Windows	
Operating system	Windows 8.1 64 Bit, Windows 8 64 Bit, Windows 7 64 Bit Service Pack 1, Windows Vista 64 Bit Service Pack 2
CPU	Core 2 Quad Q6600 at 2.4 GHz or AMD Phenom 9850 at 2.5 GHz
Memory	4 GB RAM
Free space	65 GB of free space
Graphics hardware	DirectX 10-compatible GPU: GeForce 9800GT 1GB or ATI Radeon HD 4870 1GB
Sound hardware	DirectX 10 compatible sound card

VIII . SOFTWARE REQUIREMENTS

The software requirements for an intelligent interface for fake product monitoring and removal would depend on various factors such as the scale of the operation, the complexity of the algorithms used, and the data processing requirements. Here are some general software requirements that may be needed:

Programming Language: The choice of programming language would depend on the specific requirements of the application. Popular languages used for machine learning and deep learning applications include Python, R, and Java.

Machine Learning Libraries: The system would require machine learning libraries such as Scikit-learn, Tensorflow, Keras, and PyTorch to perform the required analysis and training of the models.

Natural Language Processing (NLP) Libraries: NLP libraries such as NLTK, SpaCy, and Gensim may be required if the system needs to process text data.

Image Processing Libraries: Image processing libraries such as OpenCV, Pillow, and Scikit-image may be required if the system needs to process image data.

Data Storage: The system would require a database management system (DBMS) to store and manage the data. Popular DBMS used for machine learning applications include MySQL, MongoDB, and PostgreSQL.

Web Development Frameworks: If the system needs to be deployed as a web application, web development frameworks such as Django, Flask, and Ruby on Rails may be required.

Cloud Services: Depending on the scale of the operation,

cloud services such as Amazon Web Services (AWS), Google Cloud Platform (GCP), or Microsoft Azure may be required for data storage, computing, and deployment.

Overall, the software requirements for an intelligent interface for fake product monitoring and removal would be significant and would depend on the specific requirements of the application. The system would require a combination of machine learning libraries, NLP libraries, image processing libraries, web development frameworks, and cloud services to operate effectively.

IX . CONCLUSION

The results In the proposed work, a dataset containing Urdu and Roman Urdu reviews has been developed. Fake reviews are hard to spot on your own. Therefore, the n-gram approach is used to detect false positives in several languages. Text category and SVM classifier was observed to be the best approach to detect fake reviews. Today, technology is developing day by day and online marketers can sell their products and have millions of reviews about those products. There are some organizations that post fake reviews for seller products to raise or lower the price of the product. Therefore, it is proposed to classify reviews into authentic ones, which detect fake reviews in different languages like English, Urdu and Roman Urdu. It helps users to buy products from Daraz, Flipkart and Amazon with satisfaction and pay for good quality products. Similarly, e-commerce stores like AliExpress and Alibaba have many reviews in different languages. It would be better if the proposed system found a way to process and filter comments for other languages.

X . REFERENCES

[1] A. Sinha, N. Arora, S. Singh, M. Cheema, and A. Nazir, “Fake Product Review Monitoring Using Opinion Mining,” vol. 119, no. 12, pp. 13203–13209, 2018.

[2] Torbet, Georgina. “U.S. Customers Spent over $6 Billion on Black Friday Purchases.” Digital Trends, Digital Trends, 25 Nov. 2018, www.digitaltrends.com/web/shopping-totals-black-friday/

[3] Sterling, Greg. “Study Finds 61 Percent of Electronics Reviews on Amazon Are 'Fake'.” Marketing Land, 19 Dec. 2018, marketingland.com/study-finds-61-percent-of electronics-reviews-on-amazon-are-fake 254055.

[4] K. Khan, W. Khan,A. Rehman, A. Khan, Asfandyar. Khan, A. Ullah Khan, B. Saqia, "Urdu Sentiment Analysis," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 9, 2018.

[5] A. Mukherjee, B. Liu, and N. Glance, "Spotting Fake Reviewer Groups in Consumer Reviews," 2012.

[6] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, “What Yelp Fake Review Filter Might Be Doing?,” Aaai, pp. 409–418, 2013

[7] Z. Wang, Y. Zhang, and T. Qian, “Fake Review Detection on Yelp Dataset and features,” pp.

[8] S. Xie, G. Wang, S. Lin, and P. S. Yu, “Review spam detection via temporal pattern discovery,” p. 823, 2012.

[9] C. Paper, “Mining millions of reviews: A technique to rank products based on importance of reviews Mining Millions of Reviews: A Technique to Rank Products Based on Importance of Reviews,” no. November, 2015.

[10] V. K. Madhura N Hegde, Sanjeetha K Shetty, Sheikh Mohammed Anas, “Fake product review monitoring,” Int. Res. J. Eng. Technol., vol. 05, no. 06, p. 4, 2018.

[11] S. Rayana and L. Akoglu, “Collective Opinion Spam Detection: Bridging Review Networks and Metadata,” Sigkdd, pp. 985–994, 2015.