International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

LEVERAGING AI AND MACHINE LEARNING FOR IDENTIFYING FAKE REVIEWS IN E-COMMERCE

GANASREE.K[1], Dr. K.CHITRA[2]

[1] III-BSC.CS-"B" DEPARTMENT OF COMPUTER SCIENCE SRI KRISHNA ADITHYA COLLEGE OF ARTS AND SCIENCE KOVAIPUDUR,COIMBATORE

[2] HEAD OF THE DEPARTMENT, DEPARTMENT OF COMPUTER SCIENCE SRI KRISHNA ADITHYA COLLEGE OF ARTS AND SCIENCE KOVAIPUDUR,COIMBATORE

ABSTRACT:

In e-commerce, user reviews significantly impact a company's revenue, as consumers often rely on them when making purchasing decisions. These reviews provide valuable insights into the quality of products or services and help businesses understand customer sentiment and preferences. However, the presence of fake or manipulated reviews, which are designed to either artificially boost or damage a product's reputation, can mislead consumers and result in poor decision-making. Detecting such fraudulent reviews is challenging, and the issue of identifying deceptive reviews has become increasingly critical. Sentiment analysis plays a key role in identifying fake reviews.

This project focuses on developing a system to detect and eliminate fake reviews using an ensemble model capable of effectively classifying genuine and deceptive reviews. The prevention of fake reviews is addressed by ensuring that only verified individuals are allowed to submit reviews. This is done by sending a unique review ID to the registered email address of customers using a technique called Bought Auth, which relies on a Genuine Reviewer Protocol to generate review IDs for purchased products.

The unique aspect of this approach is that the features and classification labels of newly received reviews are incorporated into the initial dataset, allowing them to be treated as new samples for improved analysis. Additionally, a web-based user interface is created to provide a platform where users can input their information, and the chosen machine learning model performs fake review detection. The proposed system achieved an accuracy of 87% in detecting fake reviews written in English, surpassing the accuracy of previous methods.

KEYWORDS: Machine learning, Random Forest Classifier, Classification, Word cloud.

I.INTRODUCTION :

Fake reviews are inconsistent with real evaluations of products or services; thus, fake reviews are false, bogus, and deceptive reviews. Such reviews might be posted by different types of people, including consumers, online merchants, and review platforms. The deciding feature of fake reviews is whether they mislead consumers[1]. A fake review is a review written by someone who has not actually used the product or service. They can be written by friends, family, or employees of the company. Fake reviews are also generated by bots and companies who pay individuals to write fraudulent reviews. Companies get fake positive reviews to increase sales, or source negative reviews on other companies to bring down their competitors[2]. Such reviews can have a significant effect on product perception. Fake reviews decrease informativeness, information quality, and the effective use of online product reviews. Fake reviews also damage the credibility of reviews, and negatively affect review helpfulness. In addition, fake reviews seriously affect the development of online product reviews and stakeholders' commitment to the reduction of information asymmetry between merchants and customers. Online sellers tend to publish positive fake reviews for their products or negative fake reviews against competitors for financial gains.. In that same survey, 78% of people said Amazon product reviews play a big role in their purchase decisions. If we can't trust these online reviews—and the recent discovery that over 200,000 people were involved in a fake reviews scheme with third-party. Amazon also struggles to identify fake reviews that come from real customers who've bought and used a product. Their behavior looks legitimate, and the same customer might write some reviews that are paid and others that aren't. Another major challenge to Amazon is that the fake reviews are often coordinated on social media sites the company doesn't control. But the problem remains pervasive enough -- with many retailer's eagers to edge out their competitors -- that shopper can't really tell if the number of five-star reviews on a product is legit or artificially inflated[3]. Consumers unsure of what to believe when they're up against the prospects of dozens of copycat items in an Amazon marketplace that hosts nearly 2 million sellers globally. The effects of fake reviews have engendered serious concern and various theoretical models are employed to highlight the consequences of fake reviews[4].

II. PROBLEM STATEMENT :

The framework proposed for this project is illustrated in Figures 1 and 2. The initial step involves selecting a relevant fake news dataset from kaggle.com and performing data preprocessing. Following this, the TF-IDF technique is applied to extract word features, and the dataset is split using 10-fold cross-validation. The next phase involves classifying the data using a Random Forest classifier, with model performance evaluated through various metrics such as accuracy, recall, and precision【5】.

III. SYSTEM ANALYSIS L

E-Commerce Website

This module involves the development of an e-commerce platform that facilitates the buying and selling of physical goods, digital products, and services online. It integrates the F2RSpot API, which is used to analyze reviews and determine whether they are genuine or fake.

F2RSpot API

In this module, a web application is designed to perform analytics on fake reviews and reviewer behavior. It leverages Amazon's consumer demographic and product review data【6】 to identify and analyze fraudulent reviews, providing insights into the authenticity of the feedback.

Fake Review Classification

1. **Amazon Review Dataset**:The product review data used in this project is sourced from Kaggle. The dataset contains reviews labeled as positive, neutral, or negative.
2. **Preprocessing**:This step focuses on modifying the text data to prepare it for analysis. The preprocessing includes converting text to lowercase, removing unnecessary characters such as hashtags, multiple spaces, tabs, and stop words. Additionally, the start and end of sentences were not marked, and the text was converted into numerical vectors as required by the classifier.
3. **Feature Extraction**:We used the Tfidf (Term Frequency and Inverse Document Frequency) vectorizer from the sklearn module to convert the text data into a tf-idf representation. This method is widely used in document classification and information retrieval, aiding in sentiment score classification【7】.
4. **Random Forest Classification**:The Random Forest algorithm was utilized to classify reviews. It works by fitting multiple decision tree classifiers on various sub-samples of the dataset, and then averaging their results to improve prediction accuracy and reduce overfitting. After training the classifier on the training subset, the model was evaluated using a test subset, and performance was visualized through a confusion matrix generated using Seaborn's heatmap function.

Fake Review Prediction

In this module, users log into the e-commerce website, make purchases or choose not to, and leave reviews on products. The reviews are then analyzed using the trained Amazon review dataset to identify fake content【8】.

Fake Review Elimination

To eliminate fake reviews, the system sends an email to the account holder who has posted a review for a product they did not purchase. This ensures that only verified buyers can leave product reviews on the site.

Fake Review Prevention

* *Review ID Generation System* Each customer is assigned a unique review ID upon purchasing a product, using a method known as the BoughtAuth Technique. This ensures that only verified buyers can leave reviews for the product they've purchased, making the feedback more trustworthy.

* *Review ID Distribution* The review ID is sent to the customer's email after a purchase. By requiring the review ID to post feedback, administrators can minimize the risk of receiving inaccurate or inappropriate reviews from individuals who did not purchase the product. This helps ensure that future customers can rely on authentic, detailed reviews on e-commerce platforms like Amazon, Flipkart, Snapdeal, etc.

Performance Evaluation

* *Precision* is defined as the ratio of true positive (TP) instances to the total of true positive (TP) and false positive (FP) instances, measuring the accuracy of positive predictions in a system.

$$Precision = \frac{TP}{TP + FP}$$

**Recall* is the proportion of true positive (TP) cases out of the total of true positive (TP) and false negative (FN) cases, indicating the system's ability to correctly identify positive instances.

$$Recall = \frac{TP}{TP + FN}$$

*The F1-score is a comprehensive metric that combines both precision and recall into a single value. It serves as a weighted average, where "1" indicates the best possible score and "0" represents the worst. The F1-score is calculated using the following formula:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

*Accuracy is the proportion of correctly identified samples compared to the total number of samples.

$$Accuracy = \frac{TP + tn}{(TP + FP) + (TN + FN)}$$

*AUC (Area Under the Curve) is a performance metric used for binary classification models. It reflects the ability of the model to distinguish between positive and negative samples when selected at random.

$$AUC = \frac{\int_0^1 TPdFP}{(TP + FN)(TN + FP)}$$

IV. RESULTS L

The proposed system was evaluated using the Amazon Electronics Product Review Dataset, which consists of 4,915,063 reviewers. The reviews are categorized into two groups: 709 labeled as genuine and 1,144 labeled as fake. Amazon has classified these reviews as either real or fake. Each review in the dataset contains the review date, review ID, reviewer ID, product ID, review label, and star rating. A summary of the dataset statistics is provided in Table I. The longest review in the dataset consists of 875 words, while the shortest contains only 4 words. The average review length is 439.5 words. The total number of words across all reviews is 103,052, with 102,739 unique words.

No	Model	Accuracy	Precision	Recall	F1 Score
1	Logistic Regression	0.8201	0.82911	0.8201	0.82319
2	Naïve Bayes Classifier	0.73631	0.73331	0.73631	0.74834
3	K-NN	0.71012	0.8288	0.79241	0.79839
4	SVM	0.83148	0.87527	0.84742	0.85236
5	RF	0.95859	0.97684	0.96865	0.98457

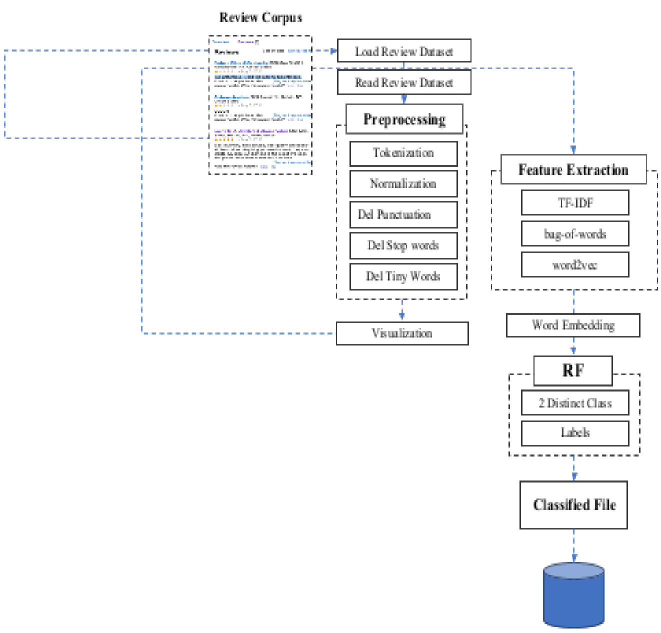
TABLE-1

V. DIAGRAM :

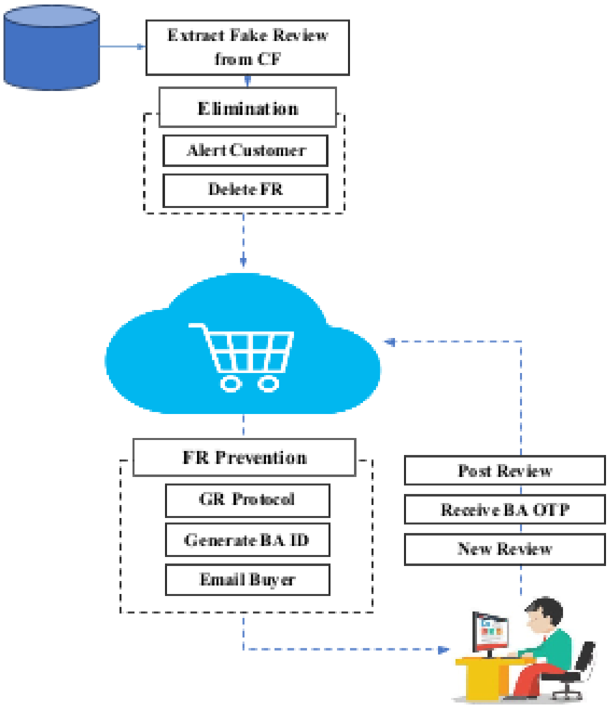
SYSTEM ARCHITECTURE

System architecture is a conceptual framework that outlines the structure, functions, and various perspectives of a system. It provides a formal representation and description of the system, organized to facilitate analysis and understanding of its components and behaviors. This helps in reasoning about how the system operates and interacts.

FAKE REVIEW DETECTION



FAKE REVIEW ELIMINATION AND PREVENTION



VI. CONCLUSION :

In today's digital world, reviews are a crucial source of information for customers when making decisions about products or services. However, fake reviews can mislead consumers, leading to substantial financial losses and public trust issues. Given the significant role reviews play in influencing decisions, detecting fake reviews has become a prominent and ongoing area of research. This paper presents F2RSpot API, a fake review detection, elimination, and prevention method based on the ensemble learning algorithm, Random Forest. The proposed approach takes into account both the content of the reviews and the behavioral patterns of the reviewers. The evaluation of the approach is conducted using the Amazon dataset, with experimental results showing that this method outperforms traditional algorithms. By utilizing unlabeled data, it enhances the performance of the classification system and achieves higher classification accuracy. Furthermore, the approach analyzes the consistency between sentiment and ratings, performs feature extraction through a text representation model, and integrates these features with external information. This combination effectively improves the classification results of the model.

VII. FUTURE ENHANCEMENTS :

In the future, we aim to develop more efficient and accurate detection methods that can identify false information across various domains, such as fake reviews, fake news, and rumors. With the growing prevalence of deepfake technology, there is an increasing need for algorithms capable of detecting reviews generated using AI-powered text or voice synthesis. This will help identify artificially created content that is often challenging to differentiate from authentic reviews.

REFERENCES :

1. 1.Y. Fang, H. Wang, L. Zhao, F. Yu, and C. Wang, ``Dynamic knowledge graph based fake-review detection,'' Int. J. Speech Technol., vol. 15, pp. 1-15, Jul. 2020.
2. J. Wang, H. Kan, F. Meng, Q. Mu, G. Shi, and X. Xiao, ``Fake review detection based on multiple feature fusion and rolling collaborative training,'' IEEE Access, vol. 8, pp. 182625-182639, 2020.
3. R. de Castro Oliveira, E. C. Gonçálves Dá Rós Baldam, F. Reis da Costa, and A. S.Pelissari, ``The effect of perceived usefulness of online reviews on hotel booking intentions,'' Revista Brasileira Pesquisa Turismo, vol. 14, no. 2, pp. 1-5, 2020.
4. R. Barbado, O. Araque, and C. A. Iglesias, “A framework for fake review detection in online consumer electronics retailers,” Information Processing & Management, vol. 56, no. 4, pp. 1234 – 1244, 2019.
5. R. Barbado, O. Araque, and C. A. Iglesias, ``A framework for fake review detection In online consumer electronic of sretailers,'' Inf. Process. Manage., vol. 56, no. 4, pp.1234-1244, Jul. 2019.

6. N. Jain, A. Kumar, S. Singh, C. Singh, and S. Tripathi, ``Deceptive reviews detection using deep learning techniques,'' in Proc. Int. Conf. Appl. Natural Lang., 2019, pp. 79-91.
7. D. Radovanovic and B. Krstajic, ``Review spam detection using machine learning,'' in Proc. 23rd Int. Sci.-Prof. Conf. Inf. Technol., Feb. 2018, pp. 1-4.
8. 8. H. Ahmed, I. Traore, and S. Saad, ``Detecting opinion spams and fake news using text classification,'' Secur. Privacy, vol. 1, no. 1, p. e9, 2018.
9. W. Zhang, Y. Du, T. Yoshida, and Q. Wang, ``DRI-RCNN: An approach to deceptive review identication using recurrent convolutional neural network,'' Inf. Process. Manage., vol. 54, no. 4, pp. 576-592, Jul. 2018.
10. Z.Wang, S. Gu, X. Zhao, and X. Xu, ``Graph-based review spammer group detection,'' Knowl. Inf. Syst., vol. 55, no. 3, pp. 571-597, Jun. 2018.