

Machine Learning-Based Fake User Identification and Sentiment Analysis of Product Reviews in E-Commerce

Mrs. N.Arockia Rosy¹, Ramanan S², Manoj Kumar P³, Gautham S H⁴

Associate Professor, Department of Information Technology, R.M.D Engineering College, Chennai,
Tamil Nadu, India¹

U.G. Student, Department of Information Technology, R.M.D Engineering College, Chennai, Tamil Nadu, India^{2,3,4}

ABSTRACT: Online e-commerce websites are heavily dependent on user-generated reviews. But with the emergence of fake reviews, it has become difficult for customers to make informed decisions. In this paper, an intelligent system has been proposed that can help e-commerce websites detect fake reviews and extract valuable insights from genuine reviews. The proposed system has used behavioural features such as user account age, review speed, verification status, and rating patterns. In addition, textual features are used, which are calculated using TF-IDF. The proposed system has used a Random Forest classifier for fake review detection. Genuine reviews are processed using a sentiment analysis model, DistilBERT. The proposed system has used a set of predefined rules for improving the performance of sentiment analysis. In addition, aspect-based keyword extraction has been used for identifying product issues. The proposed system has used visual analytics for displaying results. The proposed system has achieved high accuracy using behavioural features and textual features. The proposed system is scalable and can be used for real-world e-commerce websites.

KEYWORDS: Fake Review Detection, Natural Language Processing, Random Forest, DistilBERT, Sentiment Analysis, E-Commerce Analytics, Hybrid Models

I. INTRODUCTION

The rapid growth of e-commerce sites has profoundly affected the way in which consumers make purchase decisions. Reviews have emerged as an essential source of information, significantly influencing product perception, customer trust, and sales performance. Consumers increasingly rely on these reviews to make informed decisions about the quality and reliability of the product before making the actual purchase. Thus, the role of user-generated content is extremely crucial in determining the success of modern e-commerce platforms.

However, the rise of fake and bot-generated reviews has emerged as a significant challenge to the credibility of these e-commerce platforms. These fake reviews are created with the intention of manipulating product ratings, promoting certain products, and harming competitors. These types of reviews are created through automated means, making them extremely deceptive and difficult to detect through traditional methods of fake review detection. This has the potential to mislead customers, leading to poor purchase decisions, thereby compromising the overall trust in e-commerce platforms.

The traditional methods of fake review detection involve rule-based systems, where the process is generally simple and not flexible to adapt to new situations. These methods generally involve certain conditions, such as the frequency of the reviews, duplicate content, and ratings, to detect fake reviews. These methods, although effective in certain situations, lack the intelligence to detect more complex types of fraudulent activities.

Apart from that, sentiment analysis of reviews also has limitations. For instance, conventional sentiment analysis approaches, such as using lexicon and simple machine learning approaches, may not be able to understand the context of reviews. For example, domain-specific phrases such as “lag-free performance” or “heating issues” may be

interpreted incorrectly, which may result in inaccurate sentiment classification. This may limit the effectiveness of obtaining meaningful insights from reviews.

Recent advances in natural language processing, especially using transformer models such as DistilBERT, may be able to capture the context of reviews more accurately. Such models may be able to provide semantic understanding of reviews, which is more accurate than conventional sentiment analysis.

In order to address these challenges, this paper presents a hybrid intelligent system that uses machine learning, natural language processing, and behavioral analysis. It is proposed that the proposed system may be able to improve the accuracy of fake review detection by using user behavior data. Moreover, it may be able to provide meaningful insights by using sentiment analysis along with domain-specific corrections and aspect-based keyword extraction.

II. LITERATURE SURVEY

The rise of online shopping through e-commerce sites like Amazon, Flipkart, and Alibaba has increased the importance of online reviews in consumer purchasing decisions. However, with the rise of fake and spam reviews, concerns regarding trust and credibility have been raised in online shopping sites. Many research works have been carried out to resolve the problem of fake review detection.

Nitin Jindal and Bing Liu (2008) have proposed one of the earliest techniques for opinion spam detection based on reviewer behavior and suspicious review patterns. Myle Ott et al. (2011) have proposed a benchmark dataset for deceptive opinion spam using linguistic features with machine learning techniques for classification. Arjun Mukherjee et al. (2013) have proposed behavioral and relational features to improve fake review detection techniques.

Though good results have been achieved with these techniques, most of the techniques were based solely on text features like TF-IDF and did not effectively utilize behavioral features of reviewers. In recent times, metadata features like account age, review frequency, and verification status have been used to detect suspicious activity in review accounts. Ensemble techniques like Random Forest have been used to improve robustness in fake review detection techniques.

Moreover, models such as DistilBERT have also shown promising results for sentiment analysis because of their ability to understand context while reading review text. However, it has been observed that most of the existing studies on fake review detection and sentiment analysis have been carried out separately. Therefore, a unified hybrid framework has been proposed for fake review detection, sentiment analysis, domain correction, and aspect-based keyword extraction.

III. PROPOSED SYSTEM

The proposed system is a Machine Learning-Based Fake User Identification and Sentiment Analysis system designed to improve the reliability of e-commerce review platforms.

- The system integrates two major components:
- Fake User Detection Module
- Sentiment Analysis Module

In the fake user detection module, user behavioral features and review-related characteristics are extracted and used to train a supervised machine learning model. The trained model classifies users as genuine or fake based on learned patterns.

In the sentiment analysis module, review text is processed using Natural Language Processing techniques. Text preprocessing steps such as lowercasing, tokenization, stop word removal, and vectorization (TF-IDF) are applied. A classification model is then used to predict whether the review sentiment is positive or negative.

The system is implemented using Python and machine learning libraries. A web-based interface is provided using Fast API, where users can input review details and receive prediction results in real time.

The proposed system offers an automated, scalable, and intelligent solution compared to traditional review filtering systems.

IV. SYSTEM DESIGN AND WORKING PRINCIPLES

System Design

The proposed system is designed as a modular framework that integrates machine learning and natural language processing techniques to analyze e-commerce reviews. The system processes both textual and behavioral data to identify fake reviews and extract meaningful insights from genuine customer feedback.

Initially, the dataset containing review text and user-related features such as account age, review frequency, verification status, and ratings is collected. The data is then preprocessed by removing noise, converting text to a standardized format, and transforming review content into numerical features using the TF-IDF technique.

A Random Forest classifier is employed to detect fake reviews by analyzing both behavioral and textual features. Reviews identified as fake are filtered out, while genuine reviews are forwarded to the sentiment analysis stage. Sentiment classification is performed using the transformer-based model DistilBERT, which categorizes reviews into positive, neutral, or negative sentiments.

The system also performs aspect-based keyword extraction to identify commonly mentioned product features such as battery, camera, or performance. Finally, the results are presented using visual charts, including sentiment distribution and frequently reported product issues.

Working Principles

The system operates through a sequential pipeline. First, reviews are collected and preprocessed to generate feature vectors. The fake review detection model then classifies reviews as fake or genuine. Genuine reviews are passed to the sentiment analysis module for contextual sentiment classification.

Next, aspect extraction identifies key product features mentioned in the reviews. The extracted information is aggregated and displayed through graphical visualizations. This workflow enables automated detection of fake reviews and provides useful insights into customer feedback for improving product quality and decision-making.

V. METHODOLOGY

Random Forest

Reasons to choose the Random Forest Classifier:

- Robust against Overfitting
- Handles mixed feature types
- Effective on medium-sized datasets
- Easy to Interpret
- Model Description

The model uses an 80/20 train/test split and cross-validation to tune the hyperparameters.

The model will output a binary label as well as a prediction probability.

DistilBERT

DistilBERT is a lighter version of the BERT model, which has less computation overhead without losing the major part of its semantic capabilities. BERT has the ability to understand the context, which the traditional lexicon-based tools like TextBlob.

Example:

- “lag-free performance” → Positive
- “lightweight but fragile” → Mixed

The model categorizes the reviews as:

- Positive
- Neutral
- Negative

Domain-Aware Rule-Based Sentiment Correction

• Even though the transformer-based models, such as DistilBERT, have shown promising contextual understanding, domain-specific words used in electronics reviews can result in misinterpretations of the polarity of the sentences. Pre-trained models, such as general-purpose models, have been trained on large datasets, but the domain-specific words used in the product reviews can be misinterpreted by the model, as the pre-trained models have not been exposed to such words during their pre training period.

• In the context of electronics-based e-commerce platforms, it has often been seen that customers use complex sentences and technical words to express their sentiments. For instance, sentences like “overheating issue,” “battery drain problem,” “screen flickering” express negative sentiment. However, the use of neutral words like “issue” and “problem” might not always assure the maximum level of negative sentiment with the use of deep learning-based inference methods. Similarly, sentences like “value for money,” “lag-free performance,” “fast charging support” express positive sentiment but might not always get the maximum confidence with the use of general sentiment based inference methods

• In the case of the electronics ecommerce site, it has been seen that the customers use compound words that show implicit sentiments. For instance, words such as “over heating issue,” “battery drain problem,” “screen flickering” show negative sentiments. However, the presence of neutral words such as “issue” or “problem” does not guarantee the maximum confidence in negative sentiments using deep learning inference techniques. Similarly, words such as “value for money,” “lag-free performance,” “fast charging support” show positive sentiments specific to the electronics category, though they may not get the maximum confidence using generic sentiment inference techniques.

• Overall, the use of domain-aware rule-based correction is beneficial for improving interpretability, reducing false neutral predictions, and improving sentiment reliability for electronics product reviews. The use of a hybrid approach is an advancement over the use of deep learning based sentiment analysis systems. This approach is effective in improving the precision of sentiment labeling.

Aspect-Based Keyword Extraction

To enable the provision of useful insights to the seller, the system also offers aspect-level analysis. Instead of the overall sentiment of the review, the system identifies the following product aspects:

- Battery
- Camera
- Display
- Price

This process is achieved using a keyword-matching technique along with a limited amount of linguistic pre-processing. During this process, each review is checked to see if it contains any specific keyword related to a particular aspect. For example, the keyword “battery life” or “charging speed” will be used to represent the “Battery” aspect. If a keyword related to a particular aspect is found in a review, the review will be marked as such. When the system encounters a negative review, it aggregates the frequency of the occurrence of the specific keyword related to the particular aspect. This allows the system to generate a list of issues related to the product, indicating the problems that customers face. For example, if the keyword “battery drain” or “heating problem” is encountered a high number of times, it can be determined that the battery life is a major problem. Using the statistics related to the polarity and the occurrence of the specific aspect in the reviews, the system can generate a structured representation such as “Battery-related complaints constitute 35% of negative feedback.”

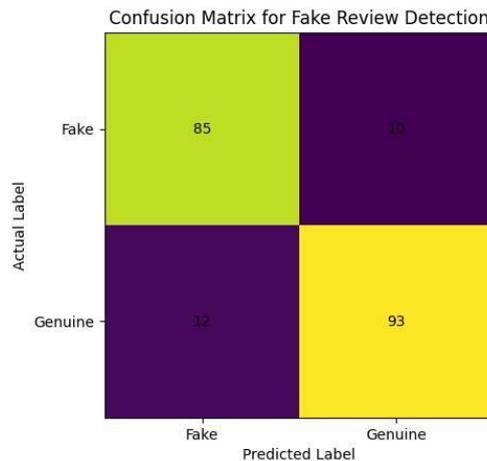
Evaluation Metrics

The model's performance is measured using standard classification metrics.

1. Confusion Matrix Analysis:

Confusion matrix for fake review detection using the Random Forest classifier.

The model correctly classified 85 fake reviews and 93 genuine reviews, while misclassifying 10 fake reviews as genuine and 12 genuine reviews as fake.



2. Model Performance Comparison

This comparative performance, as depicted in the results in the form of a table, clearly highlights the effectiveness of the different sentiment classification techniques when used on actual product reviews.

It can be seen that the Lexicon based TextBlob model had a moderate performance in terms of accuracy.

However, it was also evident that the model was unable to handle domain-specific words and nuances in the text. This is clearly depicted in the lower recall value that the model was unable to achieve.

The performance of the DistilBERT model was significantly high in all the performance parameters.

This can be attributed to the fact that the model uses a transformer architecture to learn the representation of the text in the context in which it is used.

This has resulted in a high precision and recall performance in comparison to the Lexicon-based model.

The proposed model, which is a hybrid model that combines the performance of the DistilBERT model with the domain specific rules in the form of a correction module, has the best performance in all the parameters.

The addition of the rules to handle the specific sentiments related to the electronics domain has resulted in a high F1 score performance.

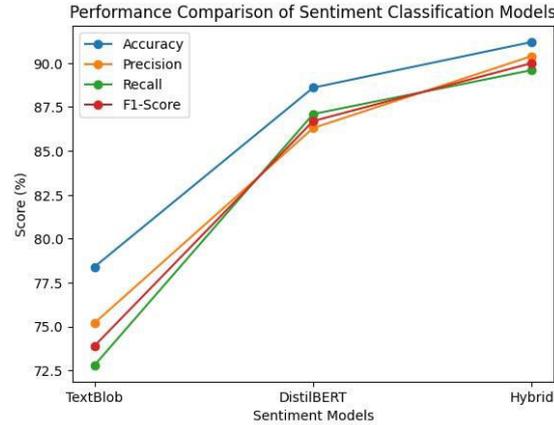


Fig illustrates the comparative performance of TextBlob, DistilBERT, and the proposed Hybrid model across four evaluation metrics. The proposed hybrid approach demonstrates superior performance in all metrics, achieving the highest accuracy and F1-score.

VI. RESULTS AND DISCUSSION

The proposed system was evaluated to measure its ability to detect fake reviews and analyze genuine customer feedback. The dataset was divided into training and testing sets, and both behavioral features (account age, review frequency, verification status) and textual features extracted using TF-IDF were used to train the Random Forest classifier. The results show that combining behavioral and textual features improves the accuracy of fake review detection compared to using textual features alone.

The Random Forest model effectively identified suspicious reviews and filtered them from the dataset. The remaining genuine reviews were analyzed using the transformer-based model DistilBERT to classify sentiments into positive, neutral, and negative categories. This model provided better contextual understanding compared to traditional sentiment analysis methods.

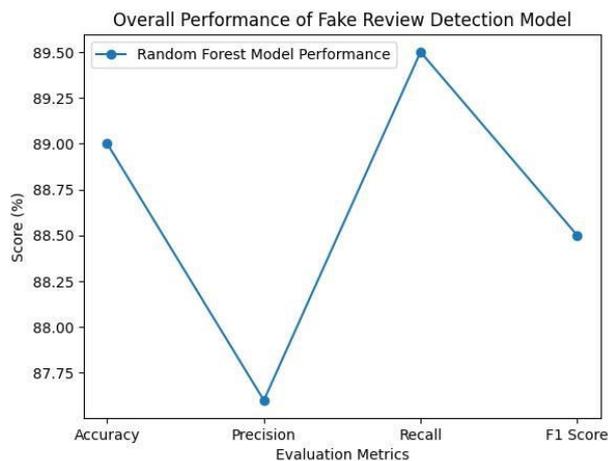


Figure1 The Random Forest model effectively detected fake reviews (89.0% accuracy, 88.5% F1) by leveraging both TF-IDF textual and behavioural features

The system also performed aspect-based keyword extraction to identify frequently mentioned product features such as battery, camera, and performance. The results were visualized using charts to present sentiment distribution and

common product issues. Overall, the system successfully detects fake reviews and provides useful insights from genuine customer feedback.

VII. CONCLUSION

This research has proposed an intelligent system for the analysis of customer reviews using behavioral machine learning and deep learning techniques in natural language processing to improve the accuracy of detecting fake reviews and sentiments in e-commerce sites.

The proposed system has used a hybrid model that combines the power of TF-IDF feature extraction and behavioral analysis using a Random Forest classifier to accurately identify genuine and fake customer reviews. In addition, the system has used a transformer model based on the DistilBERT algorithm to improve the accuracy of the sentiment analysis of genuine customer reviews. To overcome the misclassification problem in the case of electronics product reviews, a correction layer has been added to the system using a set of rules.

The system has used a keyword extraction technique to identify the frequently used words in the customer reviews related to the aspects of the product. This is further supported by the experimental results, which show that the proposed hybrid model significantly outperforms the traditional lexicon-based model and the transformer model with respect to the following parameters: Accuracy, Precision, Recall, and F1-Score. Finally, the proposed model was implemented as a FastAPI based microservice integrated with the MERN stack-based e-commerce platform to validate the practicality of the proposed model. The proposed model would be beneficial to improve the trust, transparency, and reliability of the online review ecosystems.

REFERENCES

- [1] N. Jindal and B. Liu, "Opinion spam and analysis," in Proc. Int. Conf.
- [2] Web Search and Data Mining, 2008, pp. 219–230.
- [3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in Proc. ACL, 2011,
- [4] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "Fake review detection: Classification and analysis of real and pseudo reviews," IEEE



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details