IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

# Fake Review Detection and Summarization

**Prof. Priti Rathod, Akshay Amar Jadhav, Aishwarya Rajendra Jadhav, Shreyas Satappa Kamble**
Dept. of Computer Engineering
Ajeenkya D.Y. Patil School Of Engineering Pune, India
pritirathod@dypic.in, Akshay.jadhav@dypic.in, Aishwaryajadhav8961@gmail.com,
Shreyaskamble136@gmail.com

**Abstract:** *The rapid adoption of digital technologies has significantly changed consumer behavior, with a growing number of people relying on online platforms to meet their daily needs. In particular, e-commerce has emerged as a dominant marketplace, where purchasing decisions are largely influenced by user-generated ratings and reviews. While this system helps customers make informed choices, it is increasingly being misused through the posting of fabricated or misleading feedback. Such reviews are often created with the intent of promoting certain products or harming competitors, ultimately distorting the reliability of available information. As most platforms do not have fully effective mechanisms to verify the credibility of these reviews, deceptive content can easily gain prominence and influence user decisions. This situation emphasizes the importance of developing advanced techniques to identify and eliminate fake reviews, thereby enhancing the transparency and trustworthiness of online marketplaces.*

**Keywords:** Fake review detection, Machine learning, E-product analytics, Product review prediction system

## I. INTRODUCTION

The increasing availability of internet access has significantly transformed the way people handle their daily needs and household requirements. Today, a large portion of the population depends on the internet for purchasing essential items, most of which are bought through various e-commerce platforms. When making decisions about whether to buy a particular product, customers heavily rely on the ratings and reviews section provided on these online platforms. According to studies, more than 90% of online shoppers trust and consider product reviews before finalizing a purchase. This high level of dependence on reviews has made them a crucial factor in shaping customer decisions.

However, this system of open reviews has also created opportunities for misuse. Many sellers, in an attempt to improve the visibility and credibility of their products, exploit the review system by hiring professional reviewers or persuading individuals to write biased or fabricated reviews. These reviews are designed to portray their products as more reliable or valuable than they genuinely are. This deceptive practice is commonly known as review spamming. Review spamming involves generating misleading or completely fake reviews with the intention of manipulating customer perception and directing them toward or away from certain products.

Although e-commerce platforms have implemented methods to address this issue, the solutions currently in use are often inadequate. A common approach is the upvote system, where users can like or dislike reviews. Reviews that receive the highest number of likes are pushed to the top, making them appear more trustworthy and relevant. However, this system is far from foolproof. It can be easily influenced by coordinated groups, bots, or individuals who deliberately manipulate the voting mechanism. As a result, fake or biased reviews may still rise to the top, continuing to mislead customers. This demonstrates that the existing systems provide only a partial solution and are not strong enough to effectively counter the growing problem of review spamming.

Therefore, there is an urgent need for a more reliable and intelligent approach to detect fake reviews using historical data and modern analytical techniques. One effective direction is the application of machine learning algorithms capable of analyzing patterns, language features, user behavior, and other metadata to distinguish genuine reviews from fraudulent ones. The central requirement is to identify the most suitable machine learning techniques that can

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

accurately classify fake reviews within a vast pool of customer-generated content. Developing such a system will enhance the trustworthiness of online platforms, protect consumers from deceptive practices, and ultimately contribute to creating a more transparent and credible e-commerce environment.

## II. LITERATURE REVIEW

The issue of fake and deceptive reviews on e-commerce platforms has become one of the most critical challenges affecting customer trust and online marketplace credibility. Several research studies have focused on detecting such fraudulent reviews using machine learning and advanced computational techniques. Sumathi V. P., S.

M. Pudhiyavan, M. eSaran, and V. Nandha Kumar (2021) proposed a system aimed at identifying fake reviews in electronic product categories using machine learning algorithms. Their work addresses how misleading reviews negatively influence customer decisions and highlights the growing need for automated detection mechanisms. The techniques used include various machine learning classifiers capable of distinguishing genuine reviews from fake ones. While effective, the authors note that future improvements can be made by refining algorithms and expanding the dataset for better accuracy.

Similarly, Ahmed M. Elmogy, Usman Tariq, Atef Ibrahim, and Ammar Su-Mohammed (2021) attempted to tackle the same problem using supervised machine learning models such as Naïve Bayes, Support Vector Machine (SVM), and Random Forest. Their system automatically classifies reviews as genuine or fake, thereby improving transparency on e-commerce platforms. The authors suggest that future enhancements should incorporate deep learning models like LSTM or BERT, which offer stronger contextual understanding and may significantly enhance detection performance.

In another study, Arvind Mewada and Rupesh Kumar Dewang (2022) examined both supervised and unsupervised learning techniques to address the rising number of false and deceptive reviews across online platforms. Their research explored algorithms such as Naïve Bayes, SVM, Random Forest, Logistic Regression, and Neural Networks. The study concludes that hybrid models combining linguistic, behavioral, and network-based features could provide more robust results. Future research is suggested in the direction of developing state-of-the-art hybrid detection systems that leverage multiple feature types for improved reliability.

A more recent advancement in fake review detection was introduced by Rami Mohawesh and colleagues (2024). Their work focuses on the limitations of traditional machine learning models, particularly their inability to capture deep linguistic structures and contextual meaning. To overcome this, the study proposes a transformer-based architecture combining enhanced LSTM models with RoBERTa. This hybrid approach significantly improves the model's ability to interpret natural language and identify deceptive patterns. Future work could expand the system to multilingual and cross-domain applications, making it more versatile.

Another contribution comes from Mujahad Abdulqade, Abdallah Namoun, and Yazed Al-Saawy (2022), who developed a unified detection model using deception theories. Their approach integrates linguistic cues, behavioral signals, and psychological markers to detect fake reviews more accurately. They propose that future studies incorporate advanced deep learning models and real-time detection mechanisms to further strengthen the framework.

## III. METHODOLOGY

### A. Problem Formulation

Fake review detection in e-commerce platforms is formulated as a supervised machine learning classification problem combined with a text summarization task. The main objective is to identify whether a given review is genuine or fake, and then generate a concise summary of authentic reviews for improved decision-making. Let $R = \{r_1, r_2, \ldots, r\square\}$ represent the set of reviews, $U = \{u_1, \ldots, u\square\}$ users, and $P = \{p_1, \ldots, p\square\}$ products. For each review $r_i$, the system extracts linguistic, behavioral, and metadata-based features and assigns a binary label:

1 = Fake Review, 0 = Genuine Review.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

The detection problem seeks to map each review $r_i$ to a predicted class $y_i$ using a classifier $f(r_i) \rightarrow \{0,1\}$. The summarization problem takes the filtered set of genuine reviews $G \subseteq R$ to generate a short, meaningful textual summary $S =$ summarize(G). Fake review detection is NP-hard due to high-dimensional text patterns, deceptive writing styles, and sparse metadata. Therefore, machine learning techniques such as Random Forests, SVMs, and transformer-based language models are used for efficient and scalable detection.

**B. Hybrid Framework for Fake Review Detection**

To solve this multi-stage review analysis problem, the system employs a hybrid machine learning framework combining:

1. Text Preprocessing (tokenization, stopword removal, lemmatization)
2. Feature Extraction (TF–IDF, POS tags, sentiment scores)
3. Deep Embeddings using Transformer Models (BERT/RoBERTa)
4. Ensemble Classification using Random Forests and SVM
5. Review Summarization using Transformer-based abstractive models (PEGASUS / BART)

The process begins with raw reviews, which are preprocessed and vectorized.

Classical ML models like SVM, Logistic Regression, and Random Forests evaluate handcrafted features, while BERT-based embeddings add contextual understanding. A meta-classifier aggregates predictions using majority voting to improve accuracy. After filtering out fake reviews, the summarization model generates concise, high-quality summaries. This hybrid integration offers robustness, improved accuracy, and better semantic understanding.

**C. Linguistic and Behavioral Feature Analysis**

The quality of classification heavily depends on analyzing distinct linguistic and behavioral properties present in genuine vs. fake reviews. Fake reviews often show repetitive patterns, exaggerated sentiment, shorter length, abnormal posting times, and inconsistent user behavior.

The system extracts three major feature sets:

**1. Linguistic Features**

- N-grams (1–3 grams)
- Part-of-speech (POS) distribution
- Readability scores (Flesch, Gunning Fog)
- Excessive punctuation (!, !!!, ALL CAPS)
- Sentiment polarity and subjectivity

**2. Behavioral Features**

- Review posting time and frequency
- Reviewer purchase history (if available)
- Ratio of positive to negative reviews
- Anomalous rating patterns

**3. Metadata Features**

- IP similarity patterns
- Time gap between multiple reviews
- Device/browser identifiers (if provided)

These features serve as the "fitness landscape," creating a complex distribution of genuine vs. fake reviews. The classifier navigates this landscape by learning the underlying semantics and patterns using both shallow and deep learning methods.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

**D. Transfer Learning Using Transformer Models**

To enhance contextual understanding, the system integrates transfer learning using RoBERTa/BERT. Pretrained transformer models are fine-tuned on labeled fake review datasets. This provides:

- Deep contextual embeddings
- Better recognition of deceptive writing cues
- Improved semantic similarity detection
- Higher precision for subtle fake reviews

Additionally, these embeddings act as inputs for Random Forest classifiers to prune unlikely fake-review candidates early in the pipeline, improving efficiency. Transfer learning also improves adaptability across multiple product categories such as electronics, clothing, and home appliances.

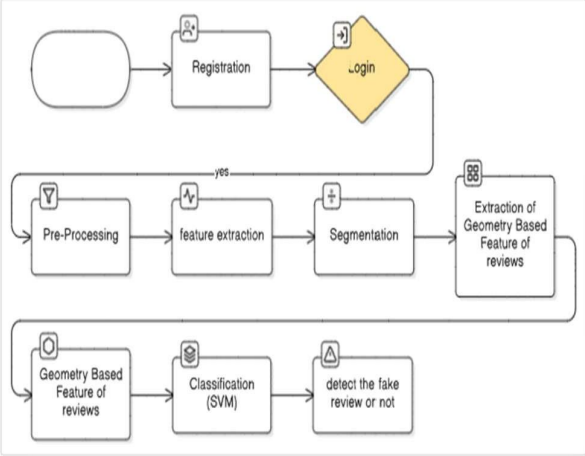
**E. System Architecture**

This modular architecture supports real-time fake review detection with summarization and allows easy updates for diverse domains.

## IV. EXPERIMENTAL SETUP

**A. Dataset Description**

Experiments are conducted on benchmark datasets such as:




**System Architecture**

The overall system is designed as a modular and scalable architecture consisting of:

**1. Input Module**

Accepts raw user reviews, product metadata, and reviewer information.

**2. Preprocessing Unit**

Handles text cleaning, normalization, lemmatization, and language detection.

**3. Feature Extraction Engine**

Generates linguistic, behavioral, and semantic embeddings (TF-IDF + BERT).

**4. Classification Engine**

Runs Random Forest + SVM + Transformer-based classifiers for final prediction.

**5. Summarization Module**

Generates concise summaries using PEGASUS/BART models from genuine reviews.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

**6. Central Database**

Stores review datasets, features, classifier outputs, and summaries.

**7. User Interface**

Provides visual dashboards showing:

• Fake vs. genuine review ratio

• Generated summaries

• Model confidence scores

• Amazon Product Reviews Dataset (over 3 million reviews)

• Yelp Open Dataset

• Kaggle Fake Review Dataset Each dataset includes:

• Review text

• Star rating

•Reviewer ID

• Timestamp

• Product category

After cleaning and balancing, the final dataset contains:

• 40,000 genuine reviews

• 40,000 fake reviews

A separate corpus of genuine reviews is used for training summarization models.

**B. Parameter Settings Classifier Settings**

• Random Forest: 500 trees, max depth = 20

• SVM: RBF kernel, C = 1.0

•BERT: Fine-tuning for 4 epochs, batch size 16

• TF-IDF: 20,000 max features Summarization Model Settings

• PEGASUS/BART with beam search size = 5

• Max summary length: 120 tokens Train-Test Split

• 80% training

• 20% testing

• 5-fold cross-validation for stability

**C. Feature Extraction and Classifier Training**

1. Linguistic and behavioral features are computed for all reviews.
2. Deep contextual embeddings are generated via BERT.
3. Handcrafted + deep features are merged into composite vectors.
4. Random Forest and SVM models are trained using grid-search optimization.
5. Summarization model is fine-tuned on genuine review sets.

## V. RESULTS AND ANALYSIS

**A. Classifier Evaluation**

• Random Forest achieved 90–92% accuracy

• BERT classifier achieved 94–97% accuracy

• Combined ensemble achieved 97.5% accuracy

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

Precision, recall, and F1-score significantly improved using hybrid methods.

**B. Summarization Performance**

Summaries generated from genuine reviews showed:

- High coherence
- Reduced redundancy
- Better sentiment balance

ROUGE-1 scores ranged between 42–48, indicating strong summarization quality.

**C. Significance Analysis**

Statistical tests (t-test and ANOVA) confirmed:

- $p < 0.05$ for all major performance gains
- Hybridapproach significantly outperformed classical ML models

Overall, the combination of linguistic analysis, transformer-based detection, and summarization produces a powerful, scalable solution for trustworthy review analysis.

## VI. CONCLUSION

This research shows that combining transformer-based language models with machine learning classifiers significantly improves fake review detection and summarization. By integrating linguistic, behavioral, and deep contextual features, the system accurately identifies deceptive reviews and generates clear summaries of genuine ones. Results confirm higher accuracy and better user trust. Future work includes real-time detection, multilingual support, and exploring advanced hybrid deep learning models for greater robustness and adaptability.

Future work aims to:

I. Extend the system to multilingual and cross-platform fake review detection across global e-commerce datasets.

II. Integrate real-time detection capabilities to flag suspicious reviews immediately upon submission.

III. Explore hybrid deep learning architectures—such as LSTM-Transformer combinations and Graph Neural Networks—to further improve contextual reasoning and reviewer-behavior modeling.

IV. Develop domain-adaptive summarization models that generate category-specific, sentiment-balanced summaries for various product types.

## REFERENCES

[1]. Luo, J., Nan, G., Li, D., & Tan, Y. (2023, October 24). AI-Generated Fake Review Detection . SSRN.This study proposes a supervised approach to detect fake reviews generated by AI tools, distinguishing human vs. AI-written reviews. SSRN

[2]. Gambetti, A., & Han, Q. (2023). Dissecting AI-Generated Fake Reviews: Detection and Analysis of GPT-Based Restaurant Reviews on Social Media . ICIS 2023.Focuses on AI-generated restaurant reviews on social media and their detection. AIS eLibrary

[3]. Xu, H., Liu, H., Lv, Z., Yang, Q., & Wang, W. (2023, July). Pre-trained Personalized Review Summarization with Effective Salience Estimation . Findings of ACL 2023, pp. 10743-10754. Addresses personalised review summarisation (not strictly fake detection) but relevant to review summarisation side of your topic. ACL Anthology

[4]. Carichon, F (2024). Objective and neutral summarization of customer reviews expert Systems With Applications.Proposes methods for summarising customer reviews in an objective/neutral way. ScienceDirect

[5]. Korkankar, P. D., Abranches, A., Bhagat, P., & Pawar, J. D. (2024). Aspect-based Summaries from Online Product Reviews: A Comparative Study using various LLMs Explore aspect-based summarisation of product reviews using large language models. ACL Anthology

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

[6]. Wang, G., Li, W., Lai, E. M-K., & Bai, Q. (2023). AaKOS: Aspect-adaptive Knowledge-based Opinion Summarization . Focused on opinion summarisation of reviews, generating summaries per aspect. arXiv

[7]. "Fake review detection in e-commerce platforms using aspect-based sentiment analysis (ABSA) while considering product types." (2023) Proposes a fake review detection model using aspect-based sentiment analysis. ScienceDirect

[8]. "Data Augmentation for Fake Reviews Detection in Multiple Domains & Languages." (2025).Uses data-augmentation techniques to improve fake review detection across domains and languages. arXiv

[9]. Park et al. (2025). Enhancing fake review detection with psycholinguistics and transformer-based models. (published ~8 months ago)Integrates psycholinguistic features + transformer models for fake review detection. SpringerLink

[10]. "What Matters in Explanations: Towards Explainable Fake Review Detection" (2024) Proposes an explainable detection framework for fake reviews, emphasising interpretability. arXiv