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 5, Issue 1, November 2025
Impact Factor: 7.67

Fake Product Review Detection and Sentiment Analysis using Machine Learning

Prof. Shravani M Kolawale[1], Shruti Kharat[2], Sejal Kannake[3], Shruti Patil[4], Satwik Jadhav[5]
Guide, Department of Computer Science[1]
Students, Department of Computer Science[2-5]
Shrimati Kashibai Navale College of Engineering, Pune, India
sejalkannake151@gmail.com , shrutipatil0316@gmail.com

Abstract: *A fraudulent review is intentionally deceptive product feedback designed to mislead readers. A machine learning-based solution is presented to automatically identify and eliminate fraudulent reviews. The model incorporates Natural Language Processing (NLP) for feature extraction and employs hybrid methods for precise identification. The system delivers tailored and reliable review suggestions by integrating information filtering with user profiles derived via collaborative filtering. Experimental findings indicate that the suggested methodology attains superior accuracy compared to current recommendation systems.*

Keywords*:* Fake reviews, Sentiment analysis, Machine Learning, Deep Learning, Opinion mining, Natural Language Processing

I. INTRODUCTION

The widespread expansion of e-commerce platforms has motivated users to express their views and experiences online, significantly impacting purchasing choices and product credibility. However, a considerable portion of these reviews are fabricated—either written by paid individuals or generated by automated bots—to artificially alter product ratings. Identifying such misleading content demands intelligent systems that can interpret linguistic cues and behavioral patterns. Machine Learning (ML) has proven to be a powerful approach for detecting fraudulent reviews and analyzing customer sentiment. Through the use of Natural Language Processing (NLP), these models can accurately distinguish between authentic and deceptive feedback while determining the sentiment polarity as positive, negative, or neutral.

II. LITERATURE SURVEY

An sophisticated deep learning framework has been created for sentiment analysis of product evaluations from e-commerce sites, as stated in [1]. The major objective is to categorize user comments into positive, negative, and neutral attitudes, therefore enhancing the credibility of online assessments. The system attains excellent accuracy in discerning authentic opinions through the utilization of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). This method enhances customer trust and offers organizations critical insights into consumer satisfaction patterns.

A data-driven supervised learning approach is suggested in [2] for the detection of fake internet reviews. The technology examines textual patterns and behavioral indicators to detect fraudulent actions. Algorithms like Random Forest, Support Vector Machine (SVM), and Naïve Bayes were evaluated to determine performance. Research indicates that these machine learning methodologies markedly enhance the precision of identifying fraudulent reviews, hence enabling automated trust evaluations in e-commerce settings.

The SentiDeceptive model presents an innovative approach for identifying deceptive product evaluations on social media sites, as noted in [3]. It combines sentiment analysis with language feature extraction to distinguish genuine feedback from deceptive responses. Experimental findings indicate that hybrid models integrating lexicon-based and deep learning approaches surpass conventional classifiers, enhancing transparency and mitigating disinformation in online review systems.

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 5, Issue 1, November 2025
Impact Factor: 7.67

Elmogy et al. [4] investigated the efficacy of supervised learning methods in detecting fraudulent internet reviews. Their methodology utilizes textual characteristics, reviewer conduct, and review-specific elements to train categorization algorithms. Decision tree and logistic regression methods were employed to evaluate detection performance. The research shows that supervised models provide a scalable and dependable approach for detecting fake reviews on digital platforms.

As mentioned in [5], the Support Vector Machine (SVM) method was designed to detect and categorize false Amazon reviews. The study seeks to reveal misleading feedback that diminishes product trust. Preprocessing techniques, including tokenization, TF-IDF, and feature selection, are utilized to improve prediction accuracy. The experimental findings reveal that SVM delivers strong and consistent performance in differentiating legitimate from false reviews.

Alsubari et al. [6] enhanced their previous supervised learning system for detecting fraudulent reviews by integrating a diverse array of linguistic and behavioral characteristics sourced from internet data. Evaluation measures, including accuracy and recall, validate the system's enhanced detection proficiency. The study underlines the necessity of thorough data analytics for establishing scalable and effective methods to counteract false reviews.

A hybrid methodology that combines deep learning with aspect-based sentiment analysis is proposed to detect fraudulent online reviews, as stated in [7]. The program analyzes product-specific characteristics and associated sentiment expressions to identify discrepancies indicative of deceit. Deep learning models, such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), encapsulate contextual and semantic links. Findings demonstrate enhanced accuracy and clarity, associating fraudulent reviews with specific product characteristics.

Ennaouri and Zellou [8] present a comprehensive evaluation of machine learning methodologies for the identification of fraudulent reviews. Their investigation classifies detection models into supervised, unsupervised, and hybrid categories, detailing their different advantages and disadvantages. The study underscores the increasing application of deep learning and natural language processing in addressing opinion spam, while stressing the necessity of explainable models and varied benchmark datasets for forthcoming research.

A machine learning-based system has been developed to autonomously identify and eliminate false reviews, ultimately enhancing customer trust. The methodology employs text mining and classification algorithms to eliminate spam content. Data preparation techniques, like stop-word elimination and stemming, improve data quality. The results indicate that machine learning-based detection techniques significantly diminish fraudulent reviews and improve the reliability of e-commerce platforms.

Qazi et al. [10] performed an extensive assessment of machine learning methodologies for identifying opinion spam. The research evaluates algorithms like SVM, decision trees, and ensemble approaches across many datasets to assess performance variations. It underscores the significance of deep learning and hybrid systems in enhancing detection precision and flexibility. The article finishes by highlighting significant research gaps, such as class imbalance and the evolution of spam strategies, and advocates for robust and adaptable algorithms for detecting bogus reviews.

TABLE I

Sr. No	Methodology	Algorithm Used	Limitations
[11]	Developed a user-centered fake news detection framework integrating user engagement features and textual attributes for fake content identification.	Random Forest, Decision Tree, SVM	Model performance depends heavily on feature selection; lacks adaptability across platforms and languages.
[12]	Proposed an online detection system for fake reviews using classical ML models with review text preprocessing and sentiment classification.	Logistic Regression, Random Forest, Naïve Bayes, SVM	Limited evaluation on large-scale, multilingual datasets; relies only on text features without behavioral analysis.
[13]	Designed a real-time fake product review monitoring system integrating sentiment polarity detection with ML classification.	Decision Tree, KNN	Accuracy affected by noisy data; lacks integration with deep learning or contextual embeddings.

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 5, Issue 1, November 2025
Impact Factor: 7.67

[14]	Conducted a systematic review of ML-based fraud detection methods applied in e-commerce systems, including fake review identification.	Ensemble Learning, SVM, Gradient Boosting	Focused more on fraud detection in general; limited exploration of text-based fake review detection techniques.
[15]	Presented a detailed analysis of opinion spam detection methods using ML, covering datasets, features, and performance metrics.	Deep Neural Networks, SVM, Naïve Bayes	Lack of benchmark standardization; limited real-time and cross-domain validation.

A. Gap Analysis

Despite substantial research in fake review identification and sentiment analysis, several enduring obstacles impede the scalability, interpretability, and practical implementation of existing algorithms. The principal research deficiencies mentioned in the literature are encapsulated as follows:

Restricted Cross-Domain Adaptability: Current models frequently demonstrate robust performance on certain datasets (e.g., Amazon or Yelp) yet falter in generalizing across other platforms and product categories due to discrepancies in language patterns and user behavior [1], [8], [10].

Reliance on Annotated Data: Supervised learning methods, like SVM, Random Forest, and Logistic Regression, are significantly dependent on extensive, labeled datasets, which are expensive and labor-intensive to produce [2], [4], [5].

Inadequate Contextual and Semantic Comprehension: Conventional machine learning models often struggle to grasp contextual subtleties, sarcasm, or nuanced language signals, leading to diminished precision in detecting fraudulent reviews [3], [7], [9].

Absence of Real-Time Detection Systems: Most existing frameworks function in offline environments and are unable to dynamically identify fraudulent reviews on live e-commerce platforms [12], [13].

Restricted Application of Behavioral Indicators: Numerous research prioritize textual content, overlooking user behavioral indicators like posting frequency, rating consistency, or activity history, which might enhance detection reliability [6], [8], [14].

Limited Explainability and Transparency: Deep learning models such as CNN, LSTM, and BERT demonstrate superior performance yet function as "black boxes," providing scant interpretability and rationale for their predictions [1], [7], [11].

Dataset Imbalance and Bias: Fake review datasets frequently demonstrate an unequal distribution of authentic and fraudulent reviews, resulting in biased models that perform inadequately for the minority class [5], [8], [15].

Multilingual and Cultural Limitations: The majority of methodologies primarily concentrate on English datasets, resulting in insufficient examination of multilingual and cross-cultural situations [9], [10].

The absence of established assessment protocols complicates model comparison, since conflicting measures like accuracy, F1-score, and precision-recall underscore the necessity for uniform benchmarking criteria [10], [14], [15].

III. CHALLENGES

Publicly accessible datasets often have a much higher number of authentic reviews compared to fraudulent ones, resulting in biased model training and inadequate identification of minority (fraudulent) samples.

Identifying the most informative linguistic, behavioral, and semantic characteristics is a challenging endeavor. Inadequate or superfluous feature selection might diminish interpretability and model accuracy.

Traditional machine learning algorithms fail to identify semantic nuances like sarcasm, irony, or implicit attitude, which frequently appear in misleading evaluations.

Models developed on particular platforms such as Amazon or Yelp sometimes struggle to generalize to other domains owing to variations in terminology, review structure, and user intent.

Supervised algorithms rely significantly on annotated datasets; yet, the manual labeling of fraudulent reviews is resource-intensive, subjective, and prone to errors.

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 5, Issue 1, November 2025
Impact Factor: 7.67

Fraudsters consistently alter their writing patterns and tactics, making static or rule-based detection algorithms more ineffective over time.

Despite the remarkable accuracy attained by deep learning models like CNNs and BERT, their "black box" characteristics hinder transparency and diminish trust in decision-making.

Most research concentrate on English-language evaluations, overlooking cross-linguistic and cultural differences that affect the expression of ideas.

The deployment of detection systems in live e-commerce settings necessitates scalable, low-latency models—attributes that most existing methodologies currently fail to provide.

The examination of user-generated material for the identification of fraudulent reviews presents ethical and privacy dilemmas, especially with the utilization of sensitive information such as user identity or location.

IV. APPLICATIONS OF MACHINE LEARNING TECHNIQUES ON DIFFERENT DATASETS

Machine Learning (ML) and Deep Learning (DL) methodologies are widely employed to identify fraudulent product evaluations and do sentiment analysis on established datasets such as Amazon, Yelp, and TripAdvisor. These datasets encompass a variety of language styles and rating behaviors, enabling researchers to evaluate model efficacy across different circumstances. The subsequent table encapsulates main uses of machine learning techniques for the identification of fraudulent reviews and sentiment analysis, emphasizing the datasets and algorithms employed in current research.

Table 3: Overview of Machine Learning Applications on Different Datasets

Sr. No.	Dataset Description	ML and DL Technique / Algorithm Used
[1]	E-commerce product review dataset collected from multiple online platforms to analyze sentiment polarity.	Deep Learning algorithms (CNN, LSTM) for sentiment classification and opinion mining.
[2]	Review dataset containing both genuine and deceptive comments for fake review identification.	Supervised learning algorithms such as Logistic Regression, Random Forest, and SVM.
[3]	Social media product review data with deceptive rating information for sentiment-based deception detection.	SentiDeceptive framework using sentiment lexicons and hybrid ML techniques.
[4]	Publicly available e-commerce datasets consisting of textual and behavioral review data.	Supervised ML algorithms including Naïve Bayes, SVM, and Decision Tree.
[5]	Amazon review dataset focusing on classification of fake and genuine reviews using polarity-based features.	Support Vector Machine (SVM) model with supervised learning.
[6]	Large-scale dataset of consumer reviews collected from multiple sources for fake review detection.	Data analytics with supervised ML classifiers such as Random Forest and Logistic Regression.
[7]	Aspect-based review dataset for deep learning-driven fake review detection.	CNN and LSTM models integrating aspect extraction and attention mechanisms.
[8]	Comprehensive multi-domain datasets (Amazon, Yelp, IMDb) analyzed in a systematic review.	Machine Learning algorithms including Decision Tree, SVM, and Gradient Boosting.
[9]	Mixed product review dataset for identifying and removing fake opinions using ML.	Naïve Bayes and Random Forest for text classification and spam detection.
[10]	Multi-platform datasets (Amazon, Yelp, TripAdvisor) used for opinion spam detection.	Deep Neural Networks and SVM with sentiment and linguistic feature fusion.
[11]	User-centered fake news and review dataset integrating behavioral and textual attributes.	Random Forest, Decision Tree, and SVM classification algorithms.
[12]	Online fake review dataset developed for real-time detection of deceptive reviews.	Logistic Regression, Random Forest, Naïve Bayes, and SVM.
[13]	Real-time e-commerce dataset designed for fake product review monitoring.	Decision Tree and K-Nearest Neighbors (KNN) algorithms.

ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

[14]	E-commerce fraud and fake review datasets analyzed for ML-based fraud detection.	Ensemble Learning, SVM, and Gradient Boosting techniques.
[15]	Multi-domain datasets summarized in a systematic review of opinion spam detection studies.	Deep Learning models and traditional ML algorithms like SVM and Naïve Bayes.

V. FUTURE RESEARCH DIRECTIONS

Despite substantial breakthroughs in fake review identification and sentiment analysis using Machine Learning (ML) and Deep Learning (DL) approaches, several issues persist that require further exploration to enhance model robustness, transparency, and scalability. Subsequent research should focus on creating intelligent, adaptable, and explicable systems proficient in handling varied, extensive, and perpetually developing datasets. The principal directives for forthcoming research are indicated below:

• Future systems must incorporate Explainable AI (XAI) methodologies to render model choices interpretable, enabling users and organizations to comprehend the rationale behind a review being classified as legitimate or fraudulent.

• Future research should concentrate on developing models capable of efficiently adapting across diverse domains—such as electronics, hospitality, and fashion—and managing multilingual data to guarantee inclusion and global significance.

• Integrating behavioral markers such as reviewer engagement, posting frequency, and temporal trends with textual content analysis may significantly enhance detection accuracy.

• Integrating the advantages of conventional machine learning classifiers with deep learning feature extraction can improve contextual awareness and mitigate the likelihood of overfitting.

• Advanced transformer models, such as BERT, RoBERTa, and GPT, are adept at understanding intricate contextual and semantic links, hence enhancing the detection of fraudulent reviews and sentiment categorization.

• Future systems must emphasize the creation of scalable structures that provide real-time monitoring and detection of fraudulent reviews, rendering them appropriate for integration with extensive e-commerce and social media platforms.

• Researchers ought to investigate sophisticated methodologies, including data augmentation, active learning, and semi-supervised learning, to mitigate class imbalance and diminish reliance on labor-intensive human labeling.

• Integrating adversarial training techniques might enhance models' resilience to manipulated or AI-generated fraudulent evaluations, hence augmenting their flexibility to new deceptive schemes.

• Implementing privacy-preserving techniques like federated learning and differential privacy would provide secure model training while safeguarding user anonymity.

• Creating standardized benchmark datasets and uniform assessment standards would enhance fairness, repeatability, and transparency in forthcoming studies on fake review identification.

VI. CONCLUSION

The suggested system presents a sentiment classification methodology focused on product aspects, consisting of three primary modules: preprocessing, product aspect identification, and sentiment classification. The preprocessing module executes tokenization, eliminates stop words, and applies stemming to enhance the incoming data. This approach improves accuracy, retrieval efficiency, and recognition performance by removing duplicate functions and using previously unidentified characteristics. It further integrates tailored detection of fraudulent reviews from social media platforms. Preliminary assessments indicate encouraging accuracy, and further enhancements may use hybrid models alongside sophisticated feature selection methods. By employing unlabeled data, examining sentiment-score consistency, and implementing text representation through feature fusion, the system attains enhanced classification accuracy and more dependable sentiment analysis results.

REFERENCES

[1] Alzahrani, Mohammad Eid, et al. "Developing an Intelligent System with Deep Learning Algorithms for Sentiment Analysis of E-Commerce Product Reviews." Computational Intelligence and Neuroscience 2022.1 (2022): 3840071.

[2] Alsubari, S. Nagi, et al. "Data analytics for the identification of fake reviews using supervised learning." Computers, Materials & Continua 70.2 (2022): 3189-3204.
[3] Marwat, M. Irfan, et al. "Sentiment Analysis of Product Reviews to Identify Deceptive Rating Information in Social Media: A SentiDeceptive Approach." KSII Transactions on Internet & Information Systems 16.3 (2022).
[4] Elmogy, Ahmed M., et al. "Fake reviews detection using supervised machine learning." International Journal of Advanced Computer Science and Applications 12.1 (2021).
[5] Tabany, Myasar, and Meriem Gueffal. "Sentiment analysis and fake amazon reviews classification using SVM supervised machine learning model." Journal of Advances in Information Technology 15.1 (2024): 49-58.
[6] Alsubari, S. Nagi, et al. "Data analytics for the identification of fake reviews using supervised learning." Computers, Materials & Continua 70.2 (2022): 3189-3204.
[7] Bathla, Gourav, et al. "Intelligent fake reviews detection based on aspect extraction and analysis using deep learning." Neural Computing and Applications 34.22 (2022): 20213-20229.
[8] Ennaouri, Mohammed, and Ahmed Zellou. "Machine learning approaches for fake reviews detection: A systematic literature review." Journal of Web Engineering 22.5 (2023): 821-848.
[9] Bhattacharya, Swarnajit. "Monitoring and Removal of Fake Product Review Using Machine Learning (Ml)." (2023).
[10] Qazi, Atika, et al. "Machine learning-based opinion spam detection: A systematic literature review." IEEE Access 12 (2024): 143485-143499.
[11] Park, Minjung, and Sangmi Chai. "Constructing a user-centered fake news detection model by using classification algorithms in machine learning techniques." IEEE Access 11 (2023): 71517-71527.
[12] Alshehri, Asma Hassan. "An Online Fake Review Detection Approach Using Famous Machine Learning Algorithms." Computers, Materials & Continua 78.2 (2024).
[13] Jaiswal, Mohini, and Deepali Javale. "Fake Product Review Monitoring System." 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). Vol. 2. IEEE, 2024.
[14] Mutemi, Abed, and Fernando Bacao. "E-commerce fraud detection based on machine learning techniques: Systematic literature review." Big Data Mining and Analytics 7.2 (2024): 419-444.
[15] Qazi, Atika, et al. "Machine learning-based opinion spam detection: A systematic literature review." IEEE Access 12 (2024): 143485-143499