IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 4, July-August 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204057

Fake Product Review Management System

Shilpirani Prdhan, Rakhshita S, Ravi Shankar Yadav

Department of MCA, CMR Institute of Technology, Bengaluru, Karnataka, India

ABSTRACT: Online product review on shopping experience in social media has promoted user to provide feedback. The majority of the e-commerce sites these days allow the customer to write his view or his rating of the product he has bought from his site. A favorable reputation of a product could be built from the review of the client, and popularize the product.

Due to this reason, in some product the customer reviews about the product are included by the item organization individuals itself so as to make so as to deliver false positive item reviews and also one can demote a product by giving the false negative review about the product.

This is a project to create an automatic fake product review monitoring system that can classify product reviews into fake and real ones by implementing machine learning algorithms.

The system will aggregate the product reviews from multiple sources such as online stores, social media, and review sites. The data will then undergo preprocessing to clean the text, eliminate stop words, and to stem or lemmatize the text. Features will be extracted from the preprocessed text data using techniques such as bag-of- words, TF-IDF, and word embeddingsThey will then be categorized into actual or spurious reviews and split into sets to train, validate, and test.Different machine learning algorithms like Naive Bayes, SVM, Random Forest, Decision Tree, KNN and logistic regression will be trained with the training data. The model will be selected.

KEYWORDS: Fake review detection, sentiment analysis, Automated review monitoring, review authenticity.

I. INTRODUCTION

With the rapid expansion of e-commerce platforms, product reviews have emerged as one of the most critical sources of information for consumers before making purchase decisions. Studies show that a majority of online buyers rely heavily on reviews to evaluate the strengths and weaknesses of products based on the experiences of previous customers. Genuine reviews play a crucial role in building trust between sellers and buyers, guiding product improvements, and shaping consumer behavior. A collection of authentic reviews not only enhances a product's credibility but also influences its popularity and sales performance.However, the increasing dependence on online reviews has also introduced the serious problem of **fake or deceptive reviews**. Many competitors and traders manipulate review systems by deliberately posting misleading opinions. Such reviews are often either excessively positive—intended to artificially boost product ratings and sales—or excessively negative—aimed at discrediting rival products. These deceptive practices mislead consumers, distort purchasing behavior, and create unfair competition in the digital marketplace. As a result, the credibility of e-commerce platforms is undermined, reducing consumer confidence in online transactions.Manual moderation is impractical due to the sheer volume of reviews generated every day. Thus, there is a pressing need for **automated review monitoring systems** that can effectively differentiate genuine feedback from fraudulent opinions. Machine Learning (ML) and Natural Language Processing (NLP) offer promising solutions by analyzing linguistic features, sentiment patterns, and reviewer behavior. These technologies can detect hidden patterns that are difficult for humans to observe at scale.This study presents a **Fake Product Review Monitoring System** that integrates NLP techniques with machine learning algorithms to classify reviews as genuine or fake. By experimenting with models such as Logistic Regression, Naive Bayes, Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN), the system compares performance metrics to identify the most effective approach. The ultimate goal is to enhance the authenticity of product reviews, restore consumer trust, and support fair practices in e-commerce platforms.

II. LITERATURE SURVEY

Several studies have explored fake product review detection using techniques such as sentiment analysis, machine learning, and deep learning. Researchers have proposed models ranging from traditional classifiers like Naive Bayes

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204057

and SVM to advanced ensemble and neural networks, highlighting both the potential and challenges in achieving accurate detection.

Fuzzy artificial bee colony–based CNN–LSTM and semantic features for fake product review classification. (2021). [Journal/Conference].This work couples semantic feature extraction with a CNN–LSTM pipeline and optimizes training using a fuzzy artificial bee colony approach.The aim is to reduce overfitting and improve generalization on deceptive-review corpora. Reported results indicate higher accuracy than traditional ML baselines. The study highlights deep models' strength on contextual cues but notes sensitivity to training data quality and size.Jain, [First Initial.], [Co-authors]. (Year). Fake Product Review Monitoring System.Proposes a hybrid framework combining ontology-based text analysis with geolocation/IP tracking and a spam-word dictionary (Naive Bayes classifier). The method targets linguistic ambiguity by enforcing consistency checks between content and metadata. Experiments suggest improved detection of biased/inorganic opinions. Challenges include ontology maintenance and potential privacy concerns with IP/geodata.Desai, S., [Co-authors]. (Year). Fake Product Review Monitoring and Removal for Proper Ratings. [Journal/Conference]. Introduces rating-deviation analysis alongside sentiment analysis and IP tracking to flag anomalies. An admin workflow adjudicates suspicious items before removal. The system enhances moderation precision through multi-signal corroboration. Limitations involve manual oversight load and threshold selection for deviation metricsBurujwale, A., Govind, S., Kadam, N., Jadhav, B., & Patil, P. (Year). Fake Product Review Monitoring System. [Journal/Conference]. Uses SVM for text classification plus sentiment analysis within an admin–user platform (add/view/delete products, manage reviews). The pipeline supports direct removal of flagged content by admins. Results show practical deployability with clear roles. Generalization across domains and evolving spam tactics remain open issues.Priya, M., & Praba, S. R. (Year). Fake Product Review Monitoring and Removal for Genuine Online Product Review Using IP Address Tracking. [Journal/Conference].Focuses on IP-based heuristics to detect spam clusters (multiple reviews from same source/IP). Effective at surfacing coordinated campaigns and sockpuppetry. While simple and inexpensive, it risks false positives (shared networks/VPNs) and lacks robustness against distributed attacks.Fayaz, M., Khan, A., Rahman, J. U., Alharbi, A., Uddin, M. I., & Alouffi, B. (Year). Ensemble Machine Learning Model for Classification of Spam Product Reviews. [Journal/Conference]. Evaluates ensemble learning with comprehensive feature sets to boost accuracy and stability. Findings show ensembles outperform single models across metrics. Emphasis on feature selection improves efficiency without sacrificing performance. Future work suggests stacking and model interpretability (XAI).Ata-Ur-Rehman, [Co-authors]. (2019, Sept. 11–13). Intelligent Interface for Fake Product Review Monitoring and Removal. In Proc. 16th Int'l Conf. on Electrical Engineering, Computing Science and Automatic Control (CCE), Mexico City, Mexico. Splits pipeline for verified vs. non-verified reviews: sentiment–rating consistency checks for verified, SVM/TF-IDF/Count features for others. Reports SVM ≈ 87%, Naive Bayes ≈ 85%, Logistic Regression ≈ 81% on Yelp-like data. Strengths include multilingual support (e.g., Urdu/roman Urdu). Scope limited to major e-commerce domains.Kumar, B. S. (2022). Fake Product Review Monitoring System. International Journal of Engineering Innovations in Advanced Technology, 4(4). Presents a TF-IDF–driven pipeline with Naive Bayes and Passive-Aggressive classifiers for rapid text filtering. Emphasizes straightforward data collection, feature extraction, and binary classification. Argues for distinguishing opinion spam from traditional email/web spam. Notes the need for richer behavioral signals to reduce false positives.

III. PROPOSED METHODOLOGY

The proposed Fake Product Review Monitoring System employs a systematic pipeline that integrates **Natural Language Processing (NLP)** techniques with machine learning algorithms to accurately identify and filter deceptive reviews. The methodology consists of several sequential stages, each playing a crucial role in ensuring robust detection.The process begins with **data collection**, where product reviews are gathered from multiple sources such as Amazon, Flipkart, Yelp, or benchmark datasets like the Amazon Fine Food Reviews and the Ott Deceptive Opinion Spam dataset. This ensures diversity in data and a balanced representation of both genuine and fake reviews.Once the dataset is obtained, **data preprocessing** is performed to eliminate noise and standardize the textual content. This step involves removing HTML tags, punctuation marks, numerical digits, and stop words. Additionally, the text is normalized using techniques such as stemming and lemmatization, which reduce words to their base or root forms. This preprocessing step significantly improves consistency and enhances the quality of the dataset for subsequent analysis.Following preprocessing, **feature extraction** is carried out to transform unstructured text into numerical representations suitable for machine learning models. Techniques such as **Bag-of-Words (BoW)**, **Term Frequency–Inverse Document Frequency (TF–IDF)**, and **word embeddings** (e.g., Word2Vec or GloVe) are used to capture linguistic and semantic patterns from the text. These features allow models to learn discriminative characteristics of genuine versus fake reviews.The dataset is then **labeled** into genuine and fake categories and split into **training (80%)**, **validation (10%)**, and **testing (10%)** subsets. In the **model building stage**, multiple machine learning classifiers are

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204057

implemented, including **Logistic Regression, Naive Bayes, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, and Support Vector Machine (SVM)**. Each model is trained to identify deceptive patterns by leveraging both linguistic cues and contextual signals.

Finally, the system undergoes **evaluation and comparative analysis** using metrics such as **Accuracy, Precision, Recall, F1-score, and Confusion Matrix**. The performance comparison identifies the most effective model, which is then deployed into the review monitoring system to flag suspicious reviews in real time, thereby enhancing the credibility and trustworthiness of e-commerce platforms.

IV. ALGORITHMS USED

A. Decision Tree Classifier Algorithm – [10] A Decision Tree Classifier is a supervised machine learning algorithm that predicts outcomes by splitting data into branches based on feature values. Each internal node represents a decision rule, and each leaf node represents a class label. It is easy to interpret and works with both numerical and categorical data, but it may overfit if not properly pruned.

.

B. Logistic Regression - explained that Logistic Regression is a supervised learning algorithm used for binary classification tasks. It estimates the probability that a given input point belongs to a particular class using the logistic (sigmoid) function. The model is trained by minimizing a loss function (typically cross-entropy) using gradient descent. It assumes a linear relationship between the independent variables and the log-odds of the dependent variable. Logistic Regression is particularly advantageous due to its simplicity, computational efficiency, and interpretability, which makes it suitable for initial fraud detection models. Implemented as a baseline model to understand feature impact and establish a performance benchmark.

C. Naive Bayes - also analyzed Naive Bayes, a family of probabilistic classifiers based on applying Bayes' Theorem with strong (naive) independence assumptions between features. Despite the simplicity of this assumption, Naive Bayes classifiers perform well in many complex domains. They are particularly efficient for high-dimensional data and work best when the features are conditionally independent given the class label. In fraud detection, this model offers high speed and scalability, although its performance may drop if the independence assumption is heavily violated. Applied for quick classification due to its speed, helping in initial detection of fraudulent transactions.

D. K-Nearest Neighbors (KNN) – It maintains all of the training instances and assigns a new instance a class from the majority of the k-nearest neighbours (utilizing distance measures like the Euclidean distance or Manhattan distance). It is also straightforward and intuitive, and it deals well with non-stationary distributions of the data. Nonetheless, it is noisy-data-sensitive and computationally costly when dealing with very large datasets. For detection of fraud, it is extremely dependent on appropriate scaling and feature selection. It is employed to categorize transactions, depending on similarity to well-known fraudulent and genuine cases.

E. Random Forest – Declares that Random Forest is a technique of ensemble learning which creates many decision trees in the process of training and produces the majority vote for problems of classification. The individual trees are produced by bootstrapping a small subset of the data and by random feature selection, which prevents overfitting and diversifies the trees. The Random Forests are also very resistant to noise and do a good job when dealing with imbalanced-class datasets like fraud detection because they minimize the variance without increasing the bias substantially. It is trained to capture complex, non-linear patterns in the dataset, improving fraud identification.

F. Ada Boost - proposed AdaBoost, a boosting algorithm that builds a strong classifier by iteratively combining several weak learners (typically decision stumps). AdaBoost, in each step, gives higher weights to the wrongly classified instances so that the subsequent classifier pays specific attention to the problematic instances. It continues to do this until the model reaches a pre-determined number of iterations or reaches minimum error. AdaBoost performs excellently when handling sophisticated classification issues and is less likely to overfit than individual learners. Its dynamic learning methodology also suits fraud detection, where the detection of rare occurrences of fraud is of importance. It focuses on misclassified transactions, enhancing detection of rare fraud cases.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204057

V. DATASET USED

For this study, datasets containing product reviews were collected from publicly available sources such as **Amazon Fine Food Reviews** (Kaggle) and the **Ott Deceptive Opinion Spam dataset**. The Amazon dataset consists of over **500,000 reviews**, including textual feedback, product ratings, and metadata such as helpfulness scores. The Ott dataset contains manually labeled hotel reviews categorized as truthful or deceptive, which is widely used as a benchmark for fake review detection.Before model training, the datasets underwent preprocessing steps, including removal of special characters, stop words, and HTML tags, as well as stemming and lemmatization for text normalization. Feature extraction methods such as **Bag-of-Words (BoW)** and **TF–IDF** were applied to represent reviews numerically.To build and evaluate the models, the data was split into **training (80%)** and **testing (20%)** subsets. Additionally, **SMOTE (Synthetic Minority Oversampling Technique)** was used to handle class imbalance between genuine and fake reviews. This ensured that the classifiers were trained on a balanced dataset and could effectively distinguish between authentic and deceptive opinions.

VI. EXPERIMENTAL RESULTS

The proposed Fake Product Review Monitoring System was implemented using Python libraries such as **Scikit-learn, Pandas, and NLTK**. After preprocessing and feature extraction using **Bag-of-Words (BoW)** and **TF–IDF**, the dataset was split into **80% training** and **20% testing** subsets. To address class imbalance, the **SMOTE technique** was applied, ensuring equal representation of genuine and fake reviews.Multiple machine learning algorithms were trained and evaluated on the dataset. The performance of each classifier was measured using **Accuracy, Precision, Recall, F1-score, and Confusion Matrix**. Results showed that **Support Vector Machine (SVM)** achieved the highest accuracy of **87.89%**, closely followed by **Logistic Regression** with **86.31%**. **Multinomial Naive Bayes** also performed well, reaching **84.77% accuracy**, while **Random Forest** achieved **83.81%**. In contrast, **Decision Tree** recorded **73.95%**, and **K-Nearest Neighbors (KNN)** had the lowest accuracy at **57.74%**.These findings demonstrate that **SVM and Logistic Regression are the most effective classifiers** for fake review detection, particularly with high-dimensional text data. Although ensemble methods like Random Forest provided moderate results, further optimization and the use of advanced deep learning models (e.g., LSTMs, BERT) may yield even higher accuracy in future research.

VII. CONCLUSION AND FUTURE SCOPE

In this study, we proposed a **Fake Product Review Monitoring System** using machine learning algorithms to detect and classify deceptive reviews in e-commerce platforms. The work highlights the increasing problem of opinion spam, where businesses attempt to manipulate product reputations through fake positive or negative reviews. By applying Natural Language Processing (NLP) techniques such as Bag-of-Words and TF–IDF for feature extraction, and testing multiple classifiers including Logistic Regression, Naive Bayes, Decision Tree, Random Forest, K-Nearest Neighbors, and Support Vector Machines (SVM), the system demonstrated the ability to effectively distinguish between genuine and fake reviews. Experimental results revealed that SVM achieved the highest accuracy (87.89%), followed by Logistic Regression (86.31%) and Naive Bayes (84.77%), while KNN performed poorly. These findings reinforce the importance of choosing models well-suited for high-dimensional textual data.The proposed system not only improves the reliability of online reviews but also contributes to maintaining consumer trust and fair competition among sellers. It shows that integrating machine learning with linguistic and behavioral cues can significantly enhance review authenticity monitoring.

Future Scope involves several directions for improvement. First, integrating **deep learning models** such as LSTMs and transformer-based architectures (e.g., BERT, RoBERTa) can capture deeper semantic and contextual relationships within reviews, potentially improving classification accuracy. Second, the system can be extended to include **reviewer behavioral analysis**—such as IP tracking, account activity, and temporal posting patterns—to strengthen detection capabilities. Third, the development of **real-time monitoring systems** that can be integrated directly into e-commerce platforms would make detection immediate and actionable. Finally, adopting **Explainable AI (XAI)** approaches can improve the transparency of the models, enabling users and administrators to understand why a review is flagged as fake.

By addressing these enhancements, the Fake Product Review Monitoring System can evolve into a robust, scalable, and industry-ready solution for combating fraudulent online reviews.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204057

REFERENCES

1. McAuley, J., & Leskovec, J. (2013). From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews. Proceedings of the 22nd International Conference on World Wide Web (WWW '13), 897–908. https://doi.org/10.1145/2488388.2488466
2. Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT), 309–319.
3. Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81–106. https://doi.org/10.1007/BF00116251
4. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018
5. McCallum, A., & Nigam, K. (1998). A comparison of event models for Naive Bayes text classification. AAAI-98 Workshop on Learning for Text Categorization, 752, 41–48.
6. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964
7. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
8. Jindal, N., & Liu, B. (2008). Opinion spam and analysis. Proceedings of the 2008 International Conference on Web Search and Data Mining (WSDM), 219–230. https://doi.org/10.1145/1341531.1341560
9. Mukherjee, A., Liu, B., & Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. Proceedings of the 21st International Conference on World Wide Web (WWW '12), 191–200. https://doi.org/10.1145/2187836.2187863
10. Crawford, M., Khoshgoftaar, T. M., Prusa, J. D., Richter, A. N., & Al Najada, H. (2015). Survey of review spam detection using machine learning techniques. Journal of Big Data, 2(23). https://doi.org/10.1186/s40537-015-0029-9
11. Banerjee, S., Naskar, S., & Sen, A. (2020). Ensemble learning for opinion spam detection. Information Processing & Management, 57(6), 102360. https://doi.org/10.1016/j.ipm.2020.102360

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com