******[***www.ijwer.com***](http://www.ijwer.com/)

***International Journal of Worldwide Engineering Research***

## (Peer-Reviewed, Open Access, Fully Referred International Journal) Volume:02/Issue:01/January-2024

***Impact Factor- xxxx e-ISSN : xxxx-xxxx pp : 1-2***

Fake Review Detection System

**Charushila Patil \*1,Isha Zambare\*2, Bhumika Waghulde\*3, Bhavika Sadhwani\*4**

Department of Artificial Intelligence & Data Science, Guru Gobind Singh College of Engineering and Research Centre, Nashik, Maharashtra, India

# ABSTRACT

With the exponential growth of online reviews influencing consumer decisions, the rise of fake reviews has become a critical issue for e-commerce platforms. Fake reviews, often generated to manipulate product or service reputations, undermine trust and distort market dynamics. This research paper examines the various techniques used to detect fake reviews, including text-based analysis, behavioral modeling, and machine learning approaches. It also discusses the challenges associated with detection, such as the scarcity of labeled data and the increasing sophistication of deceptive content. By analyzing existing methodologies and identifying gaps in current systems, this study aims to contribute to the development of more reliable and ethical detection mechanisms for maintaining the integrity of online platforms.

**Keywords:** Fake review detection, E-commerce, Machine learning, Text analysis, Behavioral modeling, Deceptive content, NLP, Consumer trust.

**INTRODUCTION**

In the digital age, online reviews significantly influence consumer behavior and purchasing decisions. However, the growing presence of fake reviews has raised serious concerns about the reliability of online platforms. These deceptive reviews can either unfairly promote a product or tarnish a competitor’ s reputation Detecting fake reviews has become crucial to maintaining trust in e- commerce and service platforms. This paper explores various techniques used to identify fake reviews, evaluates their effectiveness, and highlights the challenges in developing robust detection systems. Web product reviews have emerged as the bedrock of online consumption choice -making. Yet, with the advent of opinion spam fake or deceptive reviews there has been growing concern about the validity and authenticity of these user reviews This research explains how these fake reviews can be detected using natural language processing (NLP) and machine learning (ML) methods. They are typically constructed either manually by interested individuals, or automatically by the use of generative facilities. They aim to mislead consumers by manipulating the perceived quality of a product artificially. Successful detection mechanisms thus play a significant role in safeguarding the integrity of review-based recommendation systems.

# METHODOLOGY

# Data and Preprocessing: The study utilized both real and simulated labelled reviews. Text preprocessing steps included converting text to lowercase, removing special characters, tokenizing, eliminating stop words, and converting text to numeric format using TF-IDF. This prepared the text for use in machine learning models.

# Feature Engineering: Key features were extracted from the TF-IDF matrix. Additionally, keyword-based heuristics highlighted domain-specific terms like "buy," "recommend," "quality," and "amazing," which were strongly associated with consumer sentiment.

# Model Training: A Logistic Regression (LR) model was selected for its interpretability an effectiveness in binary classification.

# The model was trained on TF-IDF vectors and evaluated using accuracy, precision, recall, and F1-score achieving strong performance on both training and validation datasets

#

# MODELLING AND ANALYSIS

# A deceptive review detection system is intended to detect deceptive or spam reviews on websites such as Yelp or Amazon by looking at both review content and corresponding user behavior. The system starts with gathering labeled datasets, and then cleaning and preprocessing the text data. Extracted features include key aspects from review text (e.g., sentiment and word patterns), user activity (e.g., frequency and timing of reviews), and metadata (e.g., reviewer account age or if the purchase was verified). Then, machine learning models such as Logistic Regression or Random Forest, and deep learning models such as BERT or LSTM, are trained to identify reviews as real or fake. The model is assessed based on precision, recall, F1-score, and ROC-AUC. For deployment in real-world scenarios, explanations such as SHAP or LIME are utilized to explain model decisions, and the system needs to be updated regularly and defended against changing manipulation methods.

.

# RESULTS and DISCUSSION

# The system for detecting fake reviews performed robustly, especially with the integration of textual, behavioral, and metadata features. Among the models evaluated, the highest accuracy of 92.3% and F1-score of 92.2% was achieved by a fine-tuned BERT classifier compared to conventional models such as Random Forest and Logistic Regression. Patterns of bursty review behavior and limited reviewer history were also useful to identify spurious content, whereas linguistic signals indicated that such deceptive reviews tended to contain generic or extremely positive language. False positives in some instances consisted of legitimate but brief or effusive reviews, and false negatives included well-written spurious reviews imitating real ones. These findings illustrate the power of deep learning in this application, yet also underscore the necessity of continuous model revision and extra signals of data to remain in front of changing spam strategies.

1. **CONCLUSION**

Fake reviews continue to be a persistent issue in online commerce and content sites. This research demonstrates that the integration of classical NLP methods with machine learning models can provide solid detection performance. Although existing systems offer strong accuracy, future development such as transformer- based models and tracking reviewer behavior will be required to keep pace with changing s pam strategies. Sustaining cooperation among platform sand researchers will be essential to building reliable online environments.

1. **REFERENCES**
2. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108–121
3. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. Springer- Verlag, Berlin Heidel berg New York (1997) 415–438
4. van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidel berg New York (1995)
5. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rdedn. Springer- Verlag, Berlin Heidel berg New York (1996)