

Fake Online Review Detection using Machine Learning

A Machine Learning Approach to Identify Deceptive and Misleading Customer Reviews

Lakhan Mehra

Department of Computer Science
Dr. D. Y. Patil Arts, Commerce & Science College, Pimpri
Pune, Maharashtra, India

Sharayu Dube

Department of Computer Science
Dr. D. Y. Patil Arts, Commerce & Science College, Pimpri
Pune, Maharashtra, India

ABSTRACT : Online reviews play an important role in shaping customer opinions on e-commerce and service-based platforms. People often depend on these reviews to decide whether a product or service is reliable. However, the increasing number of fake online reviews has become a serious issue. Fake reviews are usually written to mislead users by promoting products dishonestly or damaging the reputation of competitors. This creates confusion for customers and reduces trust in online platforms.

This research focuses on detecting fake online reviews using machine learning techniques. Manual identification of fake reviews is difficult due to the large volume of online data. To solve this problem, a machine learning-based system is proposed that automatically analyzes review text to identify whether a review is genuine or fake. The approach includes collecting review data, cleaning the text, and extracting useful features such as word usage and writing patterns. Machine learning algorithms like Naive Bayes, Logistic Regression, and Support Vector Machine are applied for classification.

The outcome of the study shows that machine learning techniques can successfully detect fake online reviews with good accuracy. This approach helps improve transparency, increases customer trust, and supports fair practices on digital platforms.

Keywords: Fake Online Reviews, Machine Learning, Text Classification, Online Platforms, Artificial Intelligence

INTRODUCTION

Nowadays, almost everyone reads online reviews before buying a product, booking a hotel, or choosing any service. Reviews help people understand the quality, performance, and reliability of what they plan to purchase. Because of this, online reviews have become very powerful. A few positive reviews can increase sales, while negative ones can reduce customer interest.

However, not all reviews available online are genuine. Many companies or individuals create fake reviews to promote their own products or to damage the image of competitors. These reviews are written in a way that looks real, so normal users often cannot tell the difference. As a result, people may make wrong decisions, waste money, or lose trust in online platforms.

Since thousands of reviews are posted every day, checking each review manually is not possible. This situation creates

the need for an automatic system that can identify fake reviews quickly and accurately. Machine Learning (ML) offers a practical solution to this problem. By studying patterns in review text, ML models can learn how fake reviews differ from genuine ones.

This research focuses on using machine learning techniques to detect fake online reviews and improve the reliability of online review systems.

OBJECTIVES OF THE STUDY :

The main goals of this research are:

- To understand the issue of fake online reviews
- To study how machine learning can help in review classification
- To analyze features present in genuine and fake reviews
- To build a system that can automatically detect fake reviews
- To increase trust in online review platforms

PROBLEM STATEMENT

Fake reviews are increasing on online platforms such as shopping websites, hotel booking portals, and food delivery apps. These reviews are created to mislead users and influence their decisions unfairly. Customers find it difficult to identify whether a review is real or fake. Due to the large number of reviews posted daily, manual checking is not practical. Therefore, there is a need for an intelligent system that can automatically detect fake reviews using machine learning techniques.

LITERATURE REVIEW

Researchers have shown interest in the problem of fake review detection for several years. Earlier methods depended mainly on manual analysis and simple rule-based systems. Later, with the growth of data science, machine learning became a popular approach. Studies suggest that fake reviews often contain exaggerated language, repeated phrases, and unusual patterns in writing style.

Many researchers have used techniques like text mining and sentiment analysis to understand review content. Classification algorithms such as Naive Bayes and Support Vector Machine have shown promising results in text-based prediction tasks. Even though progress has been made, the problem still exists because fake review writers continuously change their methods. This research contributes by applying simple and effective ML techniques for practical fake review detection.

METHODOLOGY

5.1 Data Collection

1. A dataset of online reviews is collected from publicly available sources where reviews are labeled as genuine or fake.

5.2 Data Preprocessing

1. Before analysis, the text data is cleaned:
2. Removing unnecessary words
3. Converting text to lowercase
4. Eliminating symbols and numbers

5.3 Feature Extraction

Important characteristics of reviews are identified, such as:

1. Word frequency
2. Length of review
3. Emotional tone (sentiment)
4. Writing patterns

5.4 Machine Learning Models Used

Three algorithms are applied:

1. Naive Bayes
2. Logistic Regression
3. Support Vector Machine (SVM)

5.5 Model Training and Testing

1. The models are trained using labeled data and tested on new data to measure accuracy.

RESULTS

The machine learning models were able to classify reviews as fake or genuine with good performance. Naive Bayes gave fast results, while SVM showed better precision in some cases. The system successfully detected suspicious review patterns, showing that ML can be a useful tool in solving this problem. The experimental evaluation shows that the machine learning models were able to classify online reviews as fake or genuine with satisfactory performance. Different algorithms behaved

differently during testing. Naive Bayes produced faster results and handled text data efficiently, making it useful for large datasets. Logistic Regression showed balanced performance, while Support Vector Machine (SVM) provided better precision in identifying fake reviews in several test cases. This indicates that SVM was more effective in reducing false classifications when review patterns were complex.

DISCUSSION

The findings of this research highlight the effectiveness of machine learning in addressing the problem of fake online reviews. Traditional manual verification methods are slow and impractical, especially when thousands of reviews are generated daily. The proposed automated system reduces human effort and speeds up the detection process. By analyzing writing patterns and textual characteristics, machine learning models can identify hidden signs of deception that are difficult for humans to notice.

However, the study also shows that no detection system can be completely perfect. Some fake reviews are written very carefully to appear genuine, which may reduce detection accuracy. Continuous training of models with updated datasets is necessary to handle changing review patterns. Despite these challenges, the system significantly improves reliability compared to manual methods and contributes to building more trustworthy digital platforms.

CONCLUSION

Fake online reviews have become a serious issue in today's digital environment. They influence customer decisions, damage fair competition, and reduce trust in online services. This research demonstrates that machine learning techniques offer a practical solution for detecting fake reviews automatically. By analyzing text features and using classification algorithms, the system can effectively distinguish genuine reviews from misleading ones.

The study proves that ML-based systems can process large amounts of review data quickly and accurately. Such systems can be integrated into e-commerce platforms and service websites to filter suspicious reviews before they affect customers. Implementing automated detection improves transparency, strengthens user confidence, and supports ethical business practices. Overall, the research shows that machine learning plays an important role in solving real-world problems related to online information reliability.

FUTURE SCOPE

Future improvements can make the system even more effective. Deep learning models such as neural networks may provide higher accuracy by learning more complex patterns in review text. Another possible extension is detecting fake reviews in multiple languages to support global platforms. Combining text analysis with user behavior data, such as posting patterns and account history, can further improve detection accuracy. Real-time fake review detection systems can also be developed to prevent misleading content from appearing online.

FLOWCHART :

[8]

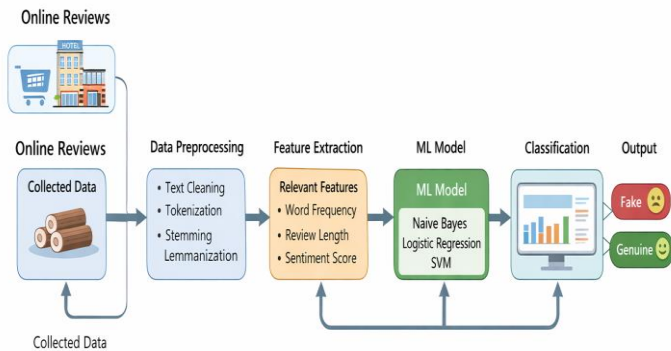


Fig. 1: System Architecture of Fake Review Detection Using Machine Learning

Keywords: Fake Online Reviews, Machine Learning, Text Classification

REFERENCES

- [1] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis.
- [2] Jurafsky, D., & Martin, J. (2020). Speech and language processing.
- [3] Aggarwal, C. (2018). Machine learning for text mining.
- [4] Jindal, N., & Liu, B. (2008). Opinion spam and analysis. Proceedings of the International Conference on Web Search and Data Mining.
- [5] Ott, M., Choi, Y., Cardie, C., & Hancock, J. (2011). Finding deceptive opinion spam by any stretch of the imagination. Proceedings of the Annual Meeting of the Association for Computational Linguistics.
- [6] Mukherjee, A., Liu, B., & Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. Proceedings of the International Conference on World Wide Web.
- [7] Li, F., Huang, M., Yang, Y., & Zhu, X. (2011). Learning to identify review spam. Proceedings of the International Joint Conference on Artificial Intelligence.