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 4, November 2025
Impact Factor: 7.67

Fake Product Review Detection Using Machine Learning

Rakshitha R[1] and G Prasanna David[2]
Student, Department of Master of Computer Application[1]
Assistant Professor, Department of Master of Computer Application[2]
Vidya Vikas Institute of Engineering and Technology, Mysuru, India

Abstract: *Product reviews are crical to onlone shopping but false are deceptive reviews undermine producer confidence and undermine the legitimacy of the product . This study presents a usimg machine learning to detect bogus reviews utilizing, category, and rating data. The method includes TF-IDF for text features, One Hot Encoding for categories, and uses an XGBoost classifier for prediction. Model performance is measured using precision, recall, and F1-score. A Streamlit web app is also developed for real-time detection. Results show the system is precise and trustworthy in spotting fake reviews..*

Keywords: TF-IDF, NLP, XGBoost, , Machine Learning Fake Reviews, E-commerce, Fake Reviews and Review Detection

I. INTRODUCTION

The quick growth of online marketplaces on recent years has changed how consumers interact with products and services more so than traditional ads. Product reviews in particular have grown to be an essential information source for prospective customers. However, the increasing reliance on online reviews has also led to fraudulent activities, such as posting manipulated or misleading reviews to artificially promote or devalue products, Such actions jeopardize the legitimacy of e- commerce platforms in addition to misleading consumers.

Detecting fake reviews is a difficult process because of the complexity of natural language, the variety of writing styles, and the international efforts to disguise fraudulent content od genuine. Traditional manual approaches are inefficient and impar actional at scale, Which ha resulted in the creation of automated system based on machine learning and natural leavening and natural language processing .These methods allow for the identification of pattern in review text, metadata and user action that can be beneficial distinguish authentic reviews from deceptive ones.

This research focuses on building a machine learningbased framework that integrates textual features, categorical information, and numerical ratings to improve fake review detection.By combining term frequency-inverse document frequency for text representation with one-hot encoding or product categories and including rating data, a robust features set is created. The proposed model employs The XGBoost algorithm, renowned for its high performance in classification tasks, to achieve accurate predictions.

Additionally, a web application built with Streamlit enables real-time review analysis, making the system both practical and user-friendly. Enhancing the reliability of internet review is the main goal of this endeavour systems by providing an effective and efficient solution to detect deceptive reviews, thereby supporting consumers in making informed decisions.

II. LITERATURE REVIEW

Fake reviews mislead customers, harm product reputation and reduce confidence in the internet platforms.

[1] Researchers use machine learning to automatically detect fake reviews, as manual checking is slow and not scalable.

[2] Many models use text features like word frequency (TF- IDF), sentiment, grammar, and review length to find patterns.

[3] Some studies analyse user behaviour, such as how often a person posts reviews or gives extreme ratings.

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 4, November 2025 Impact Factor: 7.67

[4] Models that are frequently utilized such as SVM, Random Forest, XGBoost, logistic regression, and deep learning technique such as LSTM and CNN.
[5] Recent work uses BERT and graph-based models to understand deeper language meaning and reviewer relationships.
[6] Researchers often use public datasets from Amazon, Yelp, or crowdsourced fake reviews for training and testing.
[7] Models are evaluated utilizing recall, accuracy and precision, and F1-score to assess their level of detection fake reviews.
[8] Difficulties include lack of real fake review data, changing fraud tactics, and making models that compatible with various platforms.

III. METHODOLOGY

1. Dataset and Data Preparation

• The dataset contains four main attributes: Review Text, Product Category, Rating (1–5), and Label (1 = genuine, 0 = fake).
• Preprocessing includes removing duplicates and incomplete records, changing text to lowercase, eliminating numbers, punctuation, special symbols, tokenizing, and stopword removal.
• Cleaned text is stored in a separate column called clean_text for feature extraction.
• The dataset has class imbalance with more genuine than fake reviews; stratified splitting is used to keep classes balanced during train-test split.

2. Data Pre-processing Pipeline

• Text cleaning removes special symbols, numbers, extra spaces, incomplete entries, converts text to lowercase, and removes stopwords via tokenization.
• TF-IDF vectorization is applied with unigrams and bigrams, max 10,000 features, producing a sparse matrix for classifier input. This captures term frequency and importance efficiently.

3. Model Architecture Design

• Features integrated include textual (TF-IDF), categorical (One-Hot Encoding for product categories), and numerical (rating values 1–5).
• The classifier is XGBoost with hyperparameters: 200 estimators, max depth 7, learning rate 0.1, subsample and column sampling of 0.8, using logloss for evaluation.

4. Training Pipeline Implementation and Evaluation

• Data is split stratified into 80% training and 20% testing, filtering out reviews with insufficient text.
Column Transformer handles feature transformations.
• Accuracy ,precision, recall f1-score and
.confusion matrix are among the evaluation matrix Crossvalidation and hyperparameter tuning are utilized to improve and validate model robustness.

IV. IMPLEMENTATION

The hardware configuration includes an Intel Core 10th generation processor with 6 cores and 12 threads, 16 GB DDR4 RAM suitable for medium-to- large datasets, and a 512 GB SSD for fast read/write operations during training and testing. While GPU acceleration with NVIDIA GTX 1650 or higher can further speed up training it is not required because XGBoost operates efficacy on CPU. Python 3.12 is the programming language used in the software environment. NLTK is used for data handling and manipulation, Streamlit is employed for web-based application deployment, Scikit-Learn is employed for pipeline construction and preprocessing and XGBoost is employed for model training. To guarantee reproducibility, a virtual environment built with venc is utilized, and the operating system is windows 10.Particular hyperparameters were chosen for the experiment based on validation and experimentation .XGBoost classifier to maximize generalization and performance. These include utilizing log-loss as the evaluation metric, setting the number of estimators to 200,maximum depth to 7,learning rate to 0.1,and subsample and column

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 4, November 2025
Impact Factor: 7.67

sampling to 0.8 each. These configurations were selected to Strick a balance between generalization ability and model complexity.

Parameter	Value	Justification
Learning Rate	0.1	Ensures stable convergence without overshooting
Number of Estimators	200	Provides sufficient boosting rounds for performance
Maximum Depth	7	Balances model complexity with overfitting control
Subsample	0.8	Prevents overfitting by training on random subsets
Column Sampling	0.8	Improves generalization by using partial features
Evaluation Metric	Log-loss	Suitable for binary classification tasks
Random State	42	Guarantees reproduce-ibility of results

V. RESULTS AND DISCUSSION

Category	Precision	Recall	F1-score
Electronics	0.88	0.84	0.86
Books	0.85	0.81	0.83
Clothing	0.83	0.79	0.81
Home/Kitchen	0.86	0.82	0.84
Other	0.80	0.74	0.77

Matrix of confusion When compared to conventional machine learning baselines, the suggested XGBoost based pipeline shows notable gains. The model consistently achieves classification accuracy across training and test sets by combining textual ,categorical, and numerical features, The findings demonstrate that integrating TF-IDF representations with metadata improves generalization performance by facilitating the detection of fraudulent reviews.

The assessment of model performance across product categories reveals different levels of efficacy.Categories with more detailed review information when compared to conventional machine learning baselines,the suggested XGBoost - based pipeline shows notable gains.the model comsistently achives classification accuracy across training and test sets by combining textual categorical,and numerical features.

The findings demonstrate that integrating TF-IDF representations with metadata improves generalization by facilitating the detection of fraudulent reviews.This approach is in line with findings from related research,Where XGBoost detected phony social media accounts with high accuracy rates .

VI. CONCLUSION

Model	Accuracy	Macro F1
Logistic Regression	82.4%	0.80
Random Forest	84.1%	0.82
Proposed XGBoost	87.6%	0.85

Research offers a framework for machine learning that integrates textual, categorical, and numerical features in an XGBoost classification pipeline to detect fraudulent product reviews. The system outperformed baseline models like Random Forest and logistic Regression, achieving high accuracy and F1 score. The primary contribution of this paper are as follows: Using text

,category ,and rating features to improve classification is known as hybrid feature integration. Pipeline Development : A complete, Trained and Performance Advancement: Demonstrated improvements in accuracy and F1 compared to traditional baselines. Practical Deployment: Development of a Streamlit-based application for real-time fake review detection. While some challenges remain, particularly in handling ambiguous reviews and class imbalance, this study provides a foundation for scalable, real-time fake review detection 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 4, November 2025
Impact Factor: 7.67

REFERENCES

[1] “Opinion Spam and Analysis,” Nitin Jindal and Bing Liu and Nitin Jindal The of the 2008 International Conference on Web Search and Data Mining Proceedings

[2] Natalie Glance,Bing Liu,Vivek Venkataraman,and Arjun Mukherjee.”What Yelp Fake Review Filter could Be Doing?”The International AAAI Weblogs Conference abd Social Networks Proceedings (ICWSM), pp. 409–418, 2013.

[3] Chaochao Chen, Jun Zhou, et al. "Learning Multi-Task Topical Embeddings for phony evaluation Detection." of the 26th International Conference on World Wide Web Proceedings (WWW), pp. 1021–1030, 2017.

[4] Ahmad Alrubaian, Muhammad Al-Qurishi, Mabrook AlRakhami, and Atif Alamri. "A Reputation-Based method for Spotting fake Reviews in Online Social Networks." IEEE Access, vol. 5, pp. 16358–16375, 2017.

[5] Xuezhi Cao, Wenjing Duan, and Qiwei Gan. "Exploring Determinants of Voting for the ‘Helpfulness’ of Online User Reviews: A Text Mining Approach." Decision Support Systems, vol. 50, no. 2, pp. 511–521, 2011.

[6] Tianrui Li, Fangfang Zhang, et al. "Detecting Fake Reviews Utilizing Semi-Supervised Learning." Journal of the Association for Information Science and Technology (JASIST), vol. 66, no. 12, pp. 2461–2475, 2015.

[7] Mohammed Akbar, N. Prabhakar, and S. Ramasubbareddy. "Fake machine learning based review detection Techniques." Computer Procedia Science, vol. 171, pp. 324–331, 2020.

[8] Wei Xu, Xinyu Li, and Pengfei Yin. "Detecting Fake Online Reviews with Semi-Supervised Learning: A Survey." Information Processing & Management, vol. 59, no. 3, 102932, 2022.

[9] Steven Bird, Ewan Klein, and Edward Loper. NLP with Python: Analysing Natural Language test Toolkit. O’Reilly Media, 2009.