International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. **11** Special Issue **03** | **2026**

Fake Product Review Monitoring System

Poonam Mahesh Walawalkar[1], Omkar Dinesh Padwale[2], Chaitrali Santosh Hadkar[3], Prof. Poonam Kadam[4]

[1,2,3] *Student, Department of Computer Engineering, Sindhudurg, Maharashtra, India*
[4]*Assistant Professor, Department of Computer Engineering, Sindhudurg, Maharashtra, India*

DOI: 10.5281/zenodo.20630170

ABSTRACT

The rapid expansion of e-commerce platforms has made online product reviews a major factor influencing consumer purchasing decisions. However, the increasing presence of fabricated or manipulated reviews has weakened the reliability of these platforms and negatively affected both consumers and genuine sellers. This project proposes a Fake Product Review Monitoring System that identifies suspicious reviews using a structured six-stage analytical pipeline combining behavioral metadata analysis with textual evaluation techniques.

The proposed pipeline detects fraudulent patterns by examining inconsistencies between review headlines and body sentiment, identifying coordinated activity across users and IP addresses, detecting review flooding behavior, identifying near-duplicate review templates, and filtering meaningless or low-quality content. The system adopts a pattern-based detection strategy, which avoids reliance on large manually labeled datasets that are typically required by complex supervised learning models. Instead, it utilizes lightweight Natural Language Processing techniques such as tokenization, TF-IDF vectorization, n-gram analysis, Cosine Similarity for identifying duplicate reviews, and Latent Semantic Analysis (LSA) to evaluate semantic coherence.The system architecture separates the major functional components, including a frontend interface with Google OAuth authentication for verified user identities, a backend database for storing review data and metadata, and a dedicated analysis engine responsible for executing the six-step detection process and updating the review status as approved or flagged.Experimental evaluation was conducted on a dataset consisting of 92,933 product reviews containing attributes such as customer ID, review ID, IP address, product title, review headline, and review body. The results demonstrate that the proposed multi-layer detection pipeline can efficiently identify suspicious review patterns while remaining computationally efficient for modest hardware environments (4–16 GB RAM) using tools such as Python and Jupyter Notebook. Additionally, the system includes an administrative monitoring dashboard and supports real-time integration through Node.js APIs that interact with the Python analysis engine.This design provides startups, educational institutions, and small online marketplaces with a cost-effective and practical solution for maintaining review authenticity and strengthening consumer trust.

Keyword : - Fake reviews, Spam detection, Sentiment analysis, Cosine similarity, Latent Semantic Analysis, Behavioral analysis

1. INTRODUCTION

Online customer reviews play a crucial role in influencing purchasing behavior on e-commerce platforms. Buyers often depend on ratings and written feedback from previous customers to determine the quality and reliability of products before making a purchase. Since reviews significantly affect product reputation and sales performance, some individuals or organizations attempt to manipulate review systems by posting misleading or fabricated feedback. These deceptive reviews may either artificially promote certain products or intentionally damage the reputation of competitors. As a result, the presence of fraudulent reviews can mislead customers and reduce overall trust in online marketplaces.

Traditional review moderation approaches such as manual monitoring, keyword filtering, or user reporting mechanisms are helpful but insufficient when dealing with the massive volume of reviews generated daily. Therefore, there is a need for an automated system capable of detecting suspicious review activity by analyzing both reviewer behavior patterns and textual characteristics.The goal of this project is to develop a monitoring system that integrates behavioral indicators such as user activity patterns and IP information with textual analysis techniques including semantic consistency and similarity detection. The system aims to detect fraudulent reviews with high accuracy while minimizing false alarms.

With the growing popularity of online shopping platforms, the number of buyers and sellers participating in e-commerce ecosystems has increased dramatically. Consequently, review sections on these platforms have become a valuable source of information for customers evaluating products before making purchasing decisions.Although these reviews can provide helpful insights into product quality and user experience, they can also be misused by individuals who deliberately post fake or misleading feedback. Sellers may attempt to

artificially increase their product ratings through fabricated reviews, while competitors may post negative reviews to damage a product's reputation. Such practices create a major challenge for maintaining the reliability and fairness of online review systems.

1.1 Problem statement

Existing review filtering mechanisms often struggle to identify advanced spam behaviors. Fraudulent reviewers may create multiple user accounts, submit reviews from various IP addresses, and craft persuasive or emotionally appealing text to imitate authentic customer feedback.Traditional filtering approaches based solely on keyword detection are insufficient to identify these sophisticated patterns. Therefore, it is necessary to develop an intelligent monitoring system capable of analyzing both reviewer behavior and textual content to accurately detect fraudulent reviews.

1.2 Objectives

The primary objective of the proposed Fake Product Review Monitoring System is to design an intelligent framework capable of automatically identifying fraudulent product reviews on e-commerce platforms.The system analyzes both behavioral and textual aspects of reviews by examining features such as review submission timing, rating distribution patterns, reviewer activity frequency, and inconsistencies in textual content. Natural Language Processing techniques including tokenization and sentiment analysis are applied to extract meaningful features from review text.Ultimately, the goal of the system is to improve the credibility, transparency, and reliability of online review platforms by helping administrators identify suspicious reviews more effectively.

1.3 Scope

This research focuses on analyzing English-language product reviews that include metadata attributes such as user ID, IP address, timestamp, and product identifier. The system does not attempt to support multilingual reviews or advanced deep learning models such as BERT.Instead, the study emphasizes a pattern-based and explainable detection approach that can operate efficiently with limited computational resources. The system is designed primarily for educational use and small-scale production environments where access to large labeled datasets and expensive computing infrastructure may be limited.

2. SYSTEM ARCHITECTURE AND IMPLEMENTATION

The proposed system follows a modular architecture consisting of several interconnected components:

- Data Input and Authentication Layer - A frontend interface that uses Google OAuth for secure user authentication.
- Data Storage Layer - A database that stores reviews along with associated metadata.
- Preprocessing and Feature Extraction Module - Responsible for cleaning and preparing data for analysis.
- Six-Step Analysis Pipeline - Performs behavioral and textual detection checks.
- Classification and Flagging Module - Assigns review status such as approved or suspicious.
- Monitoring Dashboard and Reporting Interface - Allows administrators to review flagged cases.

The analysis engine is implemented using Python, while the system integration layer utilizes Node.js APIs that communicate with the Python modules. This architecture supports modularity, explainability, and near real-time review processing.

Initially, review data is collected and stored in a structured database. Each record contains information such as reviewer ID, IP address, timestamp, rating value, and review text. During preprocessing, the system removes duplicate records, cleans textual data, standardizes formats, and handles missing values. After preprocessing, both behavioral and linguistic features are extracted for further analysis.

2.1 Metadata-Driven Behavioral Analysis

The detection framework begins with behavioral analysis based on review metadata. Behavioral features are important because fraudulent reviewers often display activity patterns that differ from normal users.

The system analyzes several behavioral indicators, including:

- Number of reviews submitted from a single IP address
- Time gap between consecutive reviews
- Rating distribution patterns for individual users
- Repetitive reviews posted across multiple products
- Account creation timing and posting frequency

Statistical thresholds are used to detect unusual patterns. For instance, if a single IP address posts more than five reviews within a short period, those reviews are marked as suspicious. Similarly, if a user consistently posts only extremely positive or extremely negative ratings across many products, this behavior may indicate review manipulation.

Threshold values were determined through percentile analysis on the dataset of 92,933 reviews. For example,

posting intervals shorter than 60 seconds or more than five reviews per day from the same IP address may trigger a warning flag.

Applying these behavioral filters helps reduce the amount of data that must be processed by the computationally heavier text analysis modules.

Table -1 : Metadata Features Used for Behavioral Analysis

Feature	Description	Detection Method	Risk Level
IP Activity Rate	Reviews per IP per day	Threshold-based	High
Rating Variance	Extreme rating patterns	Statistical	Medium
Review Interval	Time gap between reviews	Time analysis	High
Account Similarity	Similar usernames	Pattern matching	Medium

2.2 NLP-Based Content Classification Model

The second detection layer analyzes the textual content of reviews using Natural Language Processing techniques. While behavioral analysis detects suspicious activity patterns, content analysis identifies deceptive writing styles and artificial text structures.

The text processing pipeline includes several steps:

a) Text Preprocessing: Stop words, punctuation, and special characters are removed, and text is converted to lowercase to standardize the dataset.

b) Feature Extraction:Unigrams, bigrams, and trigrams are extracted, and TF-IDF weighting is applied to determine the importance of words within the review corpus.

c) Similarity Detection:Cosine Similarity is calculated between TF-IDF vectors to identify near-duplicate reviews or templated text patterns. Reviews with similarity values above a defined threshold (e.g., cosine similarity > 0.9) are considered duplicates.

d) Sentiment and Polarity Analysis:Sentiment scores are calculated to detect exaggerated emotional expressions that may indicate manipulated reviews.

Combining behavioral analysis with textual analysis significantly improves detection performance compared to relying on either method individually.

3. PERFORMANCE EVALUATION AND EXPERIMENTAL ANALYSIS

The performance of the Fake Product Review Monitoring System was evaluated using the structured six-step analytical pipeline. The evaluation aimed to assess the effectiveness of metadata-based behavioral analysis, similarity detection, semantic validation, and coordinated activity identification.

The experimental dataset consisted of 92,933 product reviews stored in a flat structured file with attributes such as customer ID, review ID, IP address, product title, review headline, and review body.

The evaluation results indicate that integrating behavioral analysis with NLP-based content analysis improves detection reliability compared to using standalone text-based techniques.

3.1 Validation of Six-Step Detection Pipeline

The experimental evaluation followed the same six-stage workflow:

1. Detection of contradictory sentiment between review headline and body text
2. Identification of duplicate users or coordinated IP activity
3. Detection of review flooding behavior
4. Identification of duplicate templates using Cosine Similarity
5. Filtering of meaningless or low-quality content

6. Semantic coherence validation using Latent Semantic Analysis (LSA) 7.

The IP grouping mechanism successfully identified clusters of reviews originating from the same IP address. The flooding detection mechanism detected repeated review submissions by individual users. Cosine Similarity successfully identified near-duplicate review templates with minor variations in wording. Additionally, the LSA-based validation process helped identify reviews that lacked semantic coherence.

The layered detection strategy ensures that reviews are not flagged based on a single indicator but rather through the accumulation of multiple suspicious signals.

3.2 System Workflow Verification & Practical Performance

The overall workflow was evaluated based on several operational aspects:

- Sequential Processing Flow:The six detection stages operate sequentially. Behavioral filtering is applied first to reduce the dataset size before performing computationally intensive NLP operations such as TF-IDF vectorization and similarity calculations.
- Modular Integration:The analysis module can integrate with a Node.js backend through APIs, enabling deployment beyond the experimental notebook environment.
- Explainable Detection Mechanism:Each flagged review is associated with a specific rule or condition that triggered the detection, allowing administrators to understand the reasoning behind the decision.

- Scalability:The system successfully processed the full dataset of 92,933 reviews, demonstrating its capability to handle real-world data volumes.

4. CONCLUSIONS

The proposed Fake Product Review Monitoring System provides an effective and practical solution for detecting fraudulent reviews using a structured six-step analytical pipeline. By combining behavioral analysis with lightweight Natural Language Processing techniques, the system can identify various types of review spam without requiring large labeled datasets or computationally expensive deep learning models.Experimental evaluation on a dataset of 92,933 reviews demonstrates that the system can efficiently detect suspicious review patterns while operating on modest hardware resources. As a result, the proposed system offers a reliable and cost-effective solution for improving trust and transparency in online review platforms.

5. REFERENCES

[1]. Chuhao Yao, Jiahong Wang, Eiichiro Kodama, "A spam Review detection method by verfying Consistency among multiple Review states"IEEE 21st International Conference on High Performance Computing and Communications, 2019.
[2]. Cennet Merve Yılmaz,Ahmet Onur Durahim."A Semi-Supervised Spam Review Detection Framework" IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2018.
[3]. Lekshmi M B, Deepthi V R, "Spam Detection Framework for Online ReviewsUsing Hadoop's Computational Capability"2018 International CET Conference on Control, Communication, and Computing 2018.
[4]. Sruthi Sathyanandan, Dhanya Sreedharan, "An e-commerce feedback review mining for a trusted seller's profile by classifying fake and authentic feedback comments" International Conference on circuits Power and Computing Technologies 2017.
[5]. Nandakishor Prabhu RAsha Ashok "Effect of Feature Reduction using BigramTechnique for detection of Forged Reviews" International Conference on Advances in Computing, Communications and Informatics 2018.
[6]. Ning Luo, Huaxun Deng, Linfeng Zhao, Yuan Liu*, Xingwei Wang, Zhenhua Tan,"Multiaspect Feature based Neural Network Model in Detecting Fake Reviews" 4th International Conference on Information Science and Control Engineering, 2017.
[7]. M.N. Istiaq Ahsan, Tamzid Nahian, Abdullah All Kafi, Md. Ismail Hossain, Faisal Muhammad Shah,"An Ensemble approach to detect Review Spam using hybrid Machine Learning Technique" 19th International Conference on Computer and Information Technology 2016.
[8]. Lu Zhang, Yang Yuan, Zhiang Wu, Jie Cao, "Semi-supervised Learning based Spammer Group Detectionin Product Reviews" Fifth International Conference on Advanced Cloud and Big Data 2017.
[9]. Jitendra kumar Rout, Amiya Kuar Dash, Niranjan Kumar Ray, "A framework for Fake Review Detection: Issues and Challenges" International Conference on Information Technology 2018.
[10]. Sanjay K.S Dr. Ajit Danti, "Detection of fake opinions on online products using Decision Tree and Information Gain" 3rd International Conference on Computing Methodologies and Communication 2019.