



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 11, November 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Automated Fake Review Detection Using Machine Learning

Tamanna Maner, Prathamesh Joshi, Atharva Gadekar, Dnyaneshwar Rede, Mr.Digambar Matale

Dept. of Computer Engineering, Trinity Academy of Engineering, Pune, India

ABSTRACT : Low With the exponential growth of online reviews in e-commerce and various other domains, the presence of fake reviews has become a pressing issue. Fake reviews not only deceive consumers but also impact the reputation of businesses. This paper presents an abstract for a comprehensive approach to fake review detection. The proposed method leverages natural language processing (NLP) and machine learning techniques to detect fake reviews with high accuracy. The process involves several key steps, including data collection, pre-processing, feature engineering, and model training. We extract various features from review text, such as sentiment analysis, language patterns, and reviewer behaviour, and combine them into a unified model for fake review detection. Machine learning algorithms, specifically Logistic Regression, Naïve Bayes, Decision Tree, and Random Forest. We demonstrated that quantity, non-immediacy, affect, informality, consistency, source credibility, and deviation in behaviour are essential constructs for detecting fake review .To evaluate the effectiveness of our approach, we employ a diverse dataset comprising both genuine and fake reviews. Additionally, we explore the ethical implications of fake review detection and discuss potential countermeasures for mitigating the spread of fake reviews. We emphasize the importance of transparency in the review detection process and advocate for the responsible technologies

KEYWORDS: NLP, MI Algorithm, Reviews, Transparency,

I. INTRODUCTION

In this online world, online review platforms are widely used by people to share their opinions on products and local businesses. These reviews provide valuable impact into the experiences of others, particularly for items whose quality becomes apparent only after use. Online reviews play a pivotal role in helping consumers make informed choices, whether it's picking a restaurant or buying a product online. Reading product reviews before making a purchase has become a common practice among potential customers. They typically seek out reviews from other customers when considering a product. If the reviews are predominantly positive, it increases the likelihood of them buying the product. Conversely, if the reviews are mostly negative, they tend to look for alternative products. In the realm of online shopping, discrepancies between the information provided about products and the actual products received by consumers have made it necessary for consumers to rely on a substantial number of reviews to make informed decisions. The rise of fake reviews is a big problem on the internet. People write these fake reviews for different reasons, like making money, trying to look better than others, or beating their competitors. This report is like a journey to understand how we can find these fake reviews. It's an important task in today's connected world where we rely so much on online reviews to make decisions. As a result, product reviews not only influence consumers' purchasing intentions but also impact the interests of businesses. Favorable reviews can attract more potential customers, while unfavorable reviews can deter potential customers. To maximize profits, some unscrupulous merchants resort to hiring professional writers to create fake positive reviews for their own products, thereby increasing their products' popularity to draw in potential customers. Simultaneously, they may craft fake negative reviews about their competitors. The system we are going to design will identify the fake reviews among the reviewers using their features.

Fake reviews erode trust in online platforms and contribute to skepticism among consumers. When the lines between genuine and fabricated feedback blur, it becomes challenging for consumers to make informed decisions. Detecting fake reviews means finding out if the reviews people write online are honest or if they're pretending to be real. It's like having a detective to check if what someone says about a product or service is true or if they're just saying nice or mean things for the wrong reasons. This checking helps us trust what we read and decide better about what to buy or use.

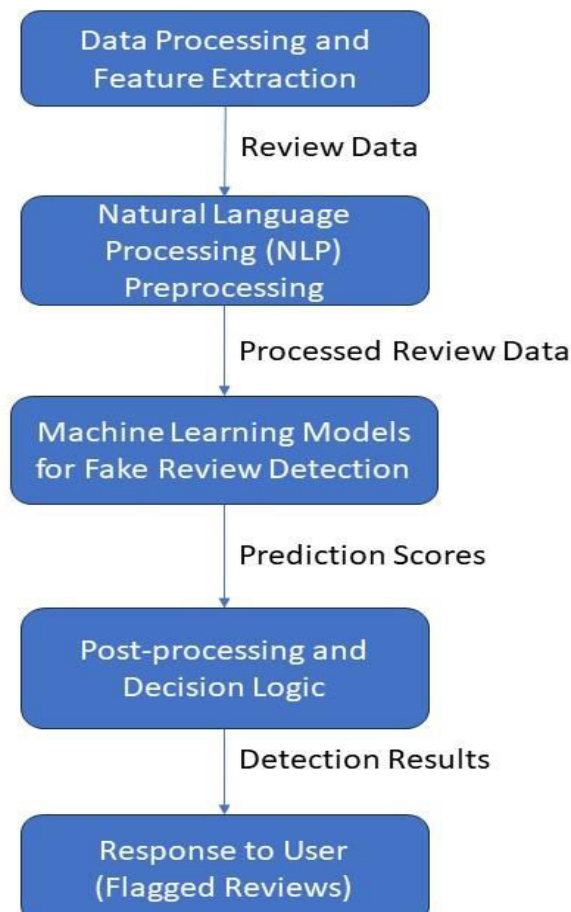
We explore the evolving landscape of online reviews, the motivations behind fake reviews, and the various techniques employed by perpetrators to create deceptive content. Moreover, we present an overview of the methodologies and technologies used to detect fake reviews, including natural language processing (NLP), machine learning algorithms, and data mining approaches. Additionally, we discuss the ethical implications of fake review detection and the need for

responsible, transparent, and fair practices in mitigating this issue. These deceptive narratives are carefully crafted to deceive potential buyers, manipulate perceptions, and bolster the reputation of a business or product. Fake reviews come in various forms, ranging from overly positive endorsements designed to inflate a product's rating to maliciously negative critiques aimed at tarnishing a competitor's image. The consequences of fake reviews are profound, as they not only mislead consumers but also erode trust in the entire review system, ultimately harming businesses and undermining the authenticity of the online marketplace.

As we delve deeper into the world of fake review detection, this report aims to shed light on the complexities of the problem and provide valuable insights into the ever-evolving battle against deceptive practices in the online review landscape. By the end of this report, readers will have a comprehensive understanding of the challenges posed by fake reviews and the strategies available for tackling this pervasive issue, thereby safeguarding the trust and reliability of online review platforms for consumers and businesses alike.

II. SYSTEM ARCHITECTURE

The system architecture for fake review detection typically involves several interconnected components and processes. Initially, the architecture begins with a data collection module, acquiring review data from various sources like e-commerce platforms or review websites. This data undergoes preprocessing and feature extraction, identifying essential elements such as sentiment analysis, text structure, language patterns, and more. These features are then fed into the machine learning models, which are trained to distinguish between genuine and fake reviews. These models, which could include decision trees, support vector machines, or neural networks, form the core of the detection system. The evaluation and tuning module continuously refines the models, adjusting hyper-parameters and optimizing feature selection to enhance accuracy and reduce misclassifications. Once trained and refined, the models are integrated into the detection system, which analyzes incoming reviews in real-time. Additionally, there might be a feedback loop to continuously improve the system based on newly identified fake review patterns. The architecture should be scalable, allowing for handling increasing review volumes, and robust enough to adapt to new forms of fraudulent behavior.



Finally, the system typically includes a reporting and analytics module to generate insights on the system's performance, helping in continual improvement and transparency.

III. DESIGN METHODOLOGY

Designing a methodology for fake review detection is a multifaceted process that involves several key steps. Below, I outline a comprehensive methodology for detecting fake reviews:

Data Collection:

Gather a dataset of online reviews from various sources, including e-commerce platforms, review websites, and social media. Collect both genuine and potentially fake reviews for a diverse set of products or services.

Pre-processing:

Clean and pre-process the text data by removing irrelevant information, special characters, and HTML tags. Tokenize the text into words or phrases. Perform stemming or lemmatization to reduce words to their base forms.

Feature Extraction:

Extract relevant features from the review data, which can include:

Textual Features: Analyse the content of the review, including sentiment, sentiment polarity, and linguistic patterns.

Metadata Features: Utilize metadata such as review date, user profile information, and product details.

User Behaviour Features: Consider user-specific information like the number of reviews posted, review frequency, and review history.

Social Network Features: Analyse connections between users and the social network structure.

Content and Sentiment Analysis Features: Use NLP techniques for sentiment analysis, topic modelling, and emotion detection.

Labelling and Annotation:

Annotate the dataset by labelling reviews as genuine or fake. Ensure a balanced dataset with a mix of both types of reviews.

Machine Learning Model Selection:

Choose appropriate machine learning algorithms for classification. Common choices include:

Supervised Learning: Utilize algorithms like Logistic Regression, Random Forest, Support Vector Machines, or deep learning techniques such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs).

Model Training and Validation:

Split the dataset into training, validation, and test sets. Train the machine learning model on the training data and fine-tune hyper parameters using the validation set. Evaluate model performance using appropriate metrics like accuracy, precision, recall, and F1-score.

Documentation and Reporting:

Document the methodology and report the results of the fake review detection system's performance.

Designing an effective fake review detection system requires a combination of data pre-processing, feature engineering, machine learning, and ongoing monitoring. The choice of algorithms and features may vary depending on the specific application and the nature of the reviews being analysed. Additionally, ethical considerations should be paramount throughout the design and implementation process.

IV. CONCLUSIONS

In conclusion, We introduce increasing the accuracy of fake review detection involves a combination of techniques of strategies and consideration that encompasses the human aspect problem. Some points are consider for expanding the scope of fake review detection and improving accuracy. Gather the information of amazon dataset of fake reviews includes various features of reviews characteristics and ensuring the dataset is accurately labelled with clear distinctions between real and fake reviews. Utilize the techniques like NLP (Part of speech, semantic analysis). Incorporate contextual information such as the reviewer's profile, their purchase history, and the timing of their reviews relative to their interaction with the product and services.



V. ACKNOWLEDGEMENTS

The authors thank Trinity Academy of Engineering for their support and providing the access of dataset and published paper.

REFERENCES

1. Received September 8, 2020, accepted September 27, 2020, date of publication October 5, 2020, date of current version October 16, 2020. {IEEE}
2. Received January 17, 2022, accepted February 10, 2022, date of publication February 18, 2022, date of current version March 10, 2022. {IEEE}
3. Received 30 November 2022, accepted 4 December 2022, date of publication 8 December 2022, date of current version 14 December 2022. Digital Object Identifier 10.1109/ACCESS.2022.3227631 Fake Online Reviews: {IEEE}
4. Aug 10th, 12:00 AM Fake Review Detection - The Value of Domain-Specificity René Theuerkauf MartinLutherUniversity Halle-Wittenberg, rene.theuerkauf@wiwi.uni-halle.de Ralf Peters Institute for Information Systems and Operations Research, ralf.peters@wiwi.uni-halle.de. {AISel}
5. Creating and detecting fake reviews of online products Joni Salminen a,b,, Chandrashekhar Kandpal c, Ahmed Mohamed Kamel d,Soon-gyo Jung a Bernard J. Jansen a Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar b Turku School of Economics at the University of Turku.
6. Fake Reviews Detection through Ensemble Learning Luis Gutierrez-Espinoza¹, Faranak Abril, Akbar Siami Namin¹, Keith S. Jones², and David R. W. david.sears@ttu.edu.
7. Fake Reviews Detection through Analysis of Linguistic Features Faranak Abri¹, Luis Felipe Gutierrez ¹, Akbar Siami Namin¹, Keith S. Jones², and David R. W. Sears³1Department of Computer Sciencettu.edu.
8. Received April 1, 2021, accepted April 21, 2021, date of publication April 26, 2021, date of current version May 6, 2021. Digital Object Identifier 10.1109/ACCESS.2021.3075573 Fake Reviews Detection(rami.mohawesh@utas.edu.au)



Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details