



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Fake Review Detection

Ayush Tyagi

Department of Computer Science & Application,
Sharda School of Computing Science and Engineering,
Sharda University, Greater Noida, India.

Aman Raj

Department of Computer Science & Application,
Sharda School of Computing Science and Engineering,
Sharda University, Greater Noida, India.

Aditya Dayal Tyagi
Assistant Professor

Department of Computer Science & Application,
Sharda School of Computing Science and Engineering,
Sharda University, Greater Noida, India.

Abstract – Online reviews significantly influence consumer purchasing decisions. Customers come to depend increasingly on reviews when they pick a product to purchase. Yet, the effectiveness of reviews on the web is hindered by fraudulent reviews that create a misleading representation of the quality of the products. [1] Detection of such real reviews becomes imperative. Sadly, to date, automatic detection of them has been only partially successful in such a difficult task. This paper deals with the generation and detection of fake or genuine reviews. The presence of fake reviews undermines the credibility of genuine feedback. [1] Here a machine learning based solution to detect fake reviews using a linear regression model trained on text data represented with TF-IDF vectorization. A Flask -based web application interface allows users to check the authenticity of review in real -time. In the attempt to fill up the gap this survey paper discusses the fake review detection task and summarizes current datasets and their collection processes. [2] It discusses the existing feature extraction techniques. We also survey and critically traditional statistical machine learning and deep learning approaches. We also undertake a benchmark study to investigate the performance if unique previously unexplored neural network architectures and transformers in fake review detection. [2] Experimental results on two benchmark datasets look the all-model performance are boosted by 7% over existing state of the art approaches and achieve accuracy of 92% representing a future study baseline. Lastly, we explain the existing gap and future looking of study in this line of the research.[2]

Keyword – Spam review, fake review detection, linear regression, flask web app, machine learning, natural language processing, TF-IDF

1. Introduction

In the era of e-commerce, customer review has become a critical factor in product sales and brand reputation. Unfortunately, the prevalence of fake reviews undermines trust and distorts market fairness. Detecting fake reviews manually is time – consuming and impractical at scale. [2] Hence there is an increasing demand for automated fake review detection system powered by machine learning techniques. This study aims to develop a simple yet effective solution using text vectorization and predictive modelling to distinguish between genuine and fake reviews. Internet technology has revolutionized the customers buying attitude as much as it has revolutionized business processes. [2] One can say there are two world one pre-e-commerce and the other post – e-commerce. In the paper, we have addressed fake review identification as binary classification problem having two categories of class fake and real. This paper is all about identifying fake review from a collection product evaluation by synthesizing spam reviews including different kinds of opinion spam analysing characteristics creating a training set and classifying it using ensemble-based classification models such as random forest and naïve bayes classification to ensure the models accuracy. [3] Several features have been taken into consideration while classifying fake reviews. Fake reviews can be generated in two primary ways. The first is through human written content, where individuals are paid to craft realistic looking but fabricated reviews without ever having used the product. [3] The second method involves computer generated reviews, where text generated algorithms automate the creation of deceptive content. Since reviews are a strong influencer of customers buying habits, dishonest parties are lured to recruit reviewers who are specialists at or who employ algorithmic techniques for creating fabricated reviews to promote the appeal of their products and services, or to undermine competitors, reputation. [3]

2. Literature Review

Numerous studies have explored the detection of deceptive online content. Approaches range from linguistic analysis, behavioural modelling, to deep learning techniques. Traditional machine learning models like support vector machines (SVM), naïve bayes classifiers, and decision trees have demonstrated reasonable accuracy when applied to textual feature extracted from reviews. [4] More recent research highlights the advantage of deep learning models such as recurrent neural networks (RNNs) and transformers for achieving higher performance. However, simpler models like linear regression, when paired with effective feature engineering, remain relevant for lightweight and interpretable applications. [4]

The growth of dependence on online reviews for customer decision-making has heightened attempts to identify spurious reviews, a common problem for online marketplaces. Fake review identification techniques have been widely researched, with conventional as well as innovative computation methods employed. [4] NLP is a core player in such attempts, with machine learning often being combined with it for the identification of linguistic as well as structural inconsistencies. [4]

Fake review identification is a special case of the problem of deception identification, for which both verbal as well as non-verbal indicators may be employed. Research for fake review identification has typically employed text as well as behavioural indicators, whereas other methods have also incorporated social or time – based criteria. [5] Text – based indicators have been introduced for several works. Have utilized psycholinguistic indicators derived from LIWC with regular word as well as part of speech (POS) features. Build upon that work incorporating style as well as POS based feature for instance deep syntax as well as sequence patterns. [5]

Research Objective

- **To develop a system that detects fake reviews** – To develop a system that automatically detects fake reviews from online platforms, ensuring the reliability and authenticity of consumer feedback. [5]
- **To utilize TF-IDF vectorization for feature extraction from textual data** – Identifying important words in the reviews by considering how frequently they appear in a document relative to the entire corpus. [6] Converting unstructured textual data into structured numerical vectors that machine learning models can understand. By applying TF-IDF, the system ensures that the machine learning models captures meaningful patterns from the review text, which are essential for accurate classification. [6]

- **To build and evaluate a linear regression – based classification model** – Training a linear regression model on the extracted TF-IDF features to predict whether a given review is genuine or fake. Although linear regression is a technique, it is adapted here by applying a threshold to treat the output as a binary classification. [6] Analysing the model's strengths and weaknesses to understand its reliability in real world applications.
- **To deploy a user-friendly web application for real time fake review detection** – Integrating the trained machine learning model into flask-based web framework, ensuring the backend can handle review inputs and provide quick predictions. [6] Designing a simple, clean, and intuitive user interface where users can submit reviews for analysis without needing technical expertise. Providing instant feedback to users by displaying whether the review is detected as genuine or fake along with a confidence score. [7] Ensuring system reliability and responsiveness to make the tool suitable for practical use in e-commerce platforms, business, or academic demonstrations. [7]

3. Methodology

The methodology for this project encompasses several critical stages, each contributing to the development of an effective fake review detection system. The process includes data preparation, feature extraction, model training, evaluation, and deployment. [8] The methodology for developing the fake review detection system involves several systematic steps, starting from data collection to model deployment. the following stages were executed. [8]

The process involves six produce sample reviews through two language models. Quantitatively and qualitatively evaluate the produced sample reviews. select the most effective language model for the development of a dataset of fake and original, human-written reviews. [8] Train classifiers on distinguishing between artificially written and authentic reviews. Hire crowd workers to label a sample of original and fake reviews. Test crowd workers and the classifiers via statistical tests on accuracy. Our methodology relies on synthetic review production, where the fake reviews are created on the basis of existing authentic reviews. [8]

3.1 Data Preparation

A small dataset was manually created consisting of sample reviews. Each review was labelled either as genuine (1) or fake (0) based on manual assessment. The dataset served as a proof-of-concept for the initial model training and evaluation. Thus, clean and pre-processing was an important step. [8]

The data was cleaned to obtained the rating and the influence number format from text. The date created field was properly formatted to transform it from text to date type. Likewise, other categories like verified purchase were also formatted as desired. All white space and empty values were replaced with nulls for textual columns and 0for quantity fields. Following the preliminary cleansing the succeeding activity was to replace abbreviation such as eliminate punctuations convert the number to words. [8]

Stop word removal in natural language processing is the process of extracting words that would not add classification that is the removal of most commonly used words in the text that even search engines cannot utilize for indexing such as stemming is the process shortening the word to its root or base or stem word. After that the, the data is ready for next step. [8]

3.2 Data Preprocessing & Analysis

The TF-IDF (Term frequency-inverse document frequency) method was employed to convert text data into numeric vectors. [9]

Review content, emotional diversity, user behaviour analysis, along with training models for classifying has been utilized by various researchers to identify fake reviews. [9] A majority of the past research on fake review detection is grouped under 3 principal detection mechanism. Review content-based detection, detections in the rating-based detection and review content along review content along with user behaviour-based detection. [9]



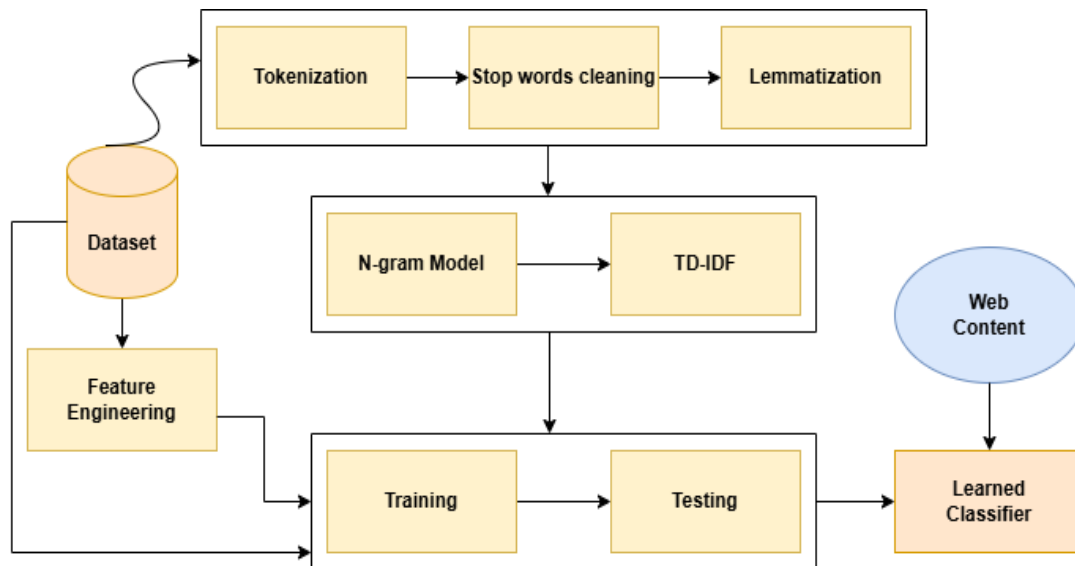


Fig. The proposed framework.

3.3 Model Training

Model choice: A linear regression model was selected. Although traditionally a regression algorithm, linear regression was adapted for binary classification by applying a threshold on the predicted output. [9]

Training process: The dataset was split into training and testing sets. The model was trained on the TF-IDF features vectors and the corresponding labels.[9]

```

app.py > ...
10 with open("model/fake_review_model.pkl", "rb") as f:
11     model, vectorizer = pickle.load(f)
12
13 @app.route("/", methods=["GET", "POST"])
14 def index():
15     prediction = ""
16     if request.method == "POST":
17         review = request.form["review"]
18         transformed = vectorizer.transform([review])
19         result = model.predict(transformed)[0]
20         label = "Genuine Review" if result > 0.5 else "Fake Review"
21         prediction = f"Prediction: {label} (Confidence: {result:.2f})"
22     return render_template("index.html", prediction=prediction)
23
24 if __name__ == "__main__":
25     app.run(debug=True)
  
```

PROBLEMS DEBUG CONSOLE OUTPUT TERMINAL PORTS

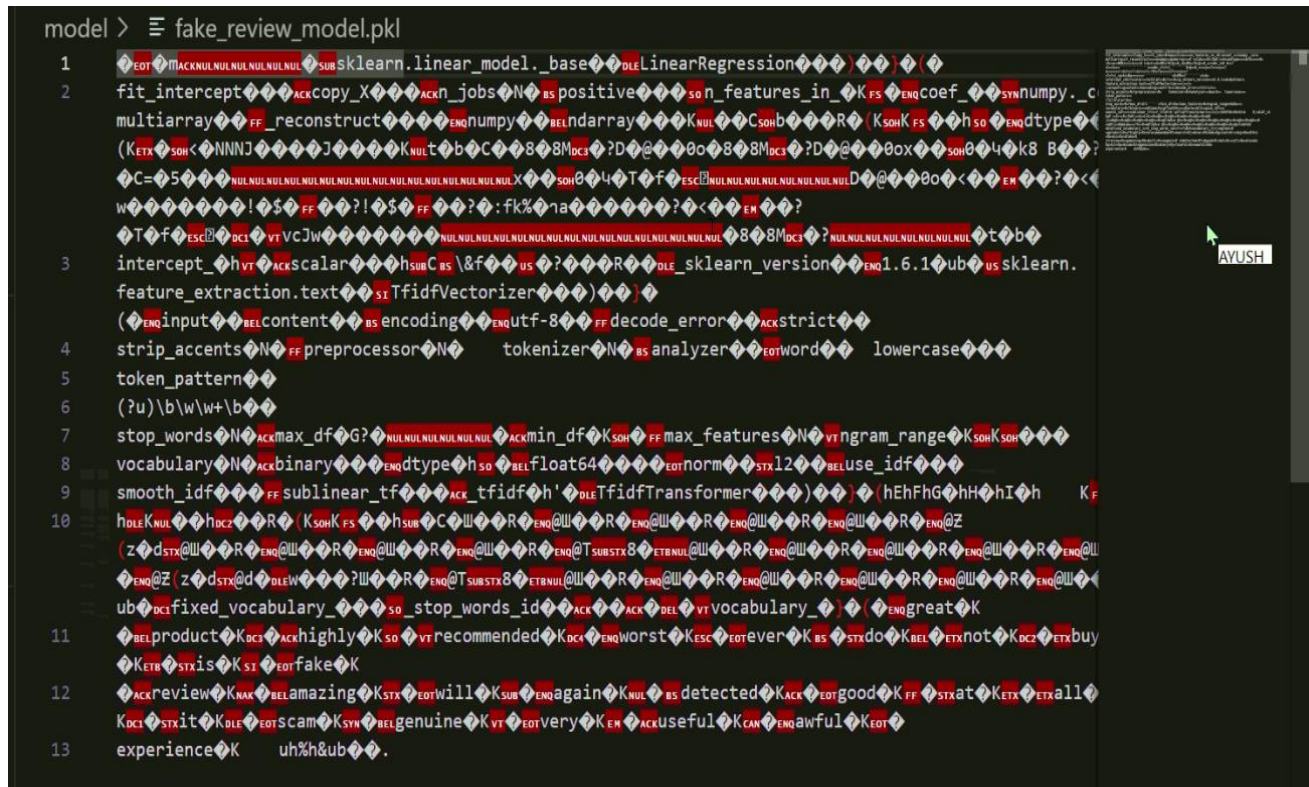
```

WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on http://127.0.0.1:5000
Press CTRL+C to quit
* Restarting with stat
* Debugger is active!
* Debugger PIN: 100-073-533
127.0.0.1 - - [13/May/2025 21:01:17] "GET / HTTP/1.1" 200 -
  
```

3.4 Model Evaluation

The performance of the model was evaluated using standard classification metrics: [10]

- . Accuracy: Percentage of total correct predictions.
- . Precision: Proportion of relevant instance among retrieved instance.
- . Recall: Proportion of relevant instance that were retrieved.
- . Evaluation was conducted utilizing scikit learn library classification report. [10]



4. Web Application Development

Backend:

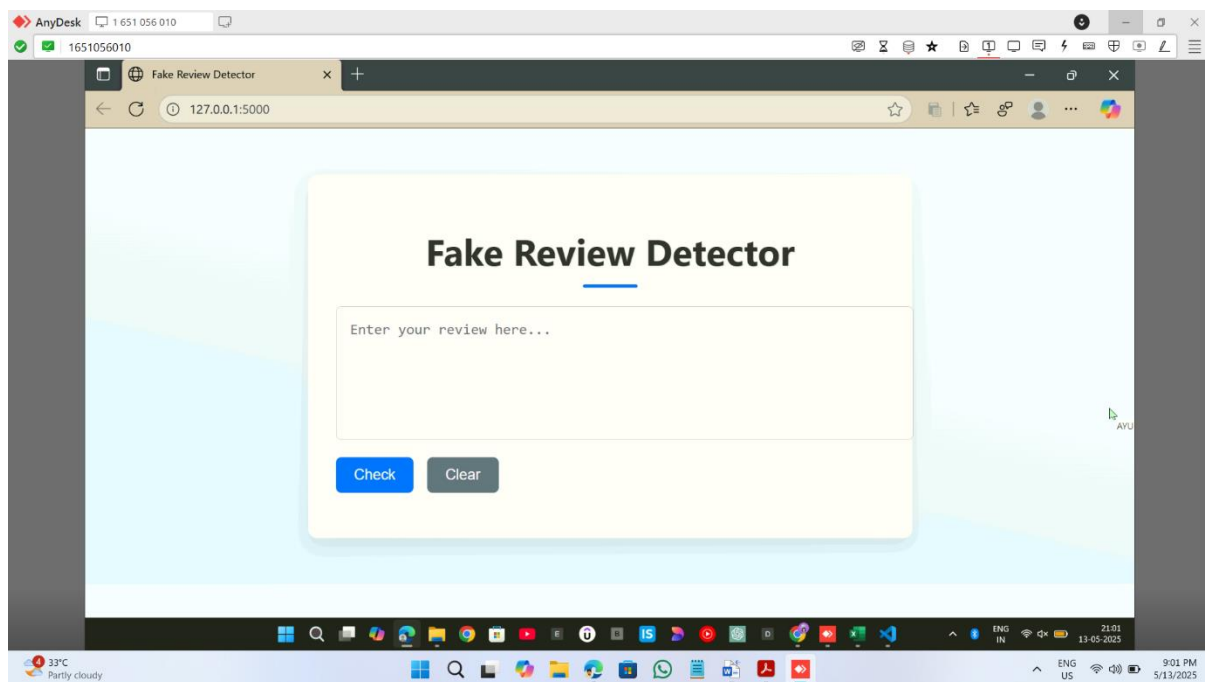
- A flask application was created to load the trained model and vectorizer.
- The app received review text from users transformed it using the saved TF-IDF vectorizer and predicted the reviews authenticity. [10]

Frontend:

- . A simple HTML form was designed to collect the review text from users.
- . The prediction result (Genuine/Fake with confidence) was displayed on the same page. [10]

User Experience:

- . Clean and minimalistic interface.
- . Real time predictions and immediate feedback. [10]



5. Conclusion

This study demonstrates that simple machine learning models when combined with text features engineering can serve as an effective solution for fake review detection. [11]

Although further improvement is needed the result are promising for real world application. Fake review detection is a challenge for researcher's online merchants' website and companies conducting e-business. Our finding show that current text generation techniques produce false reviews that seem so realistic that it is difficult for a human being to recognize them. Thankfully machine learning classifiers fare much better on this count and nearly perfect accuracy for detecting reviews written by other machines. [11] This suggests that machines can battle machines when combating false reviews. More research is required to experiment using additional data sets and platforms. The reviews have a significant part to play regarding how end users perceive a product item. It is natural for person to gauge a metaphorical book by its review. Regardless of it is that will tend to accept the world of mouth. [11] Therefore, false reviews present an insidious threat to e-commerce companies. It is unethical but common. The advance in technology have come hand-in-hand with solution for dealing with the said fake reviews and the aftermath. Utilizing these tools which can help prevent this malpractice and bring integrity to the online business. [12]

This work suggests two novel kind of features user review frequency within the same product and term feature. Experiment indicate that the model and method established are precise at distinguishing between spurious and real reviews the experimental outcomes suggest outperforms naïve bayes and can be employed to identify the real and the fraudulent reviews. [12] One thing to note here is that if one interest is to bring to surface the real opinion regarding the since the outcomes reflects its bias to detect wrote them well. [12]

6. Reference

1. Creating and detecting fake reviews of online products (Joni salminen , Chandrashekhhar Kandpal, Ahmed Mohamed Kamel, Soon-gyo jung, Bernard j. jansen www.elsevier.com/locate/jretconser)
 2. Research on false review detection methods Astate of the art review (Arvind mewada , Rupesh Kumar Dewang)
Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211004, India www.sciencedirect.com
 3. Fake review Detection using Classification (Neha Chowdhary, Anala Pandit)
International Journal of Computer Applications (0975 – 8887)
Volume 180 – No.50, June 2018
 4. Fake reviews detection: A survey (RAMI MOHAWESH 1, SHUXIANG XU 1, SON N. TRAN 1, ROBERT OLLINGTON1, MATTHEW SPRINGER 1, YASER JARARWEH 2, AND SUMBAL MAQSOOD1)
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
 5. Fake Reviews detection using supervised machine learning (Ahmed M. Elmogy¹, Usman Tariq², Atef Ibrahim, Ammar Mohammed) (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 1, 2021
 6. Fake review detection system using supervised machine learning (1Prof. Shekhar Patle, 2Gaurav Pawar, 3Shrinivas Pawar, 4Onkar Davkare, 5Rutuja Javalekar) © 2023 IJCRT | Volume 11, Issue 5 May 2023 | ISSN: 2320-2882
 7. Fake review detection using hybrid model based on CNN and LSTM (1Jitendra Bhaware and 2Vivek Sharma)
International Journal of Trend in Research and Development, Volume 12(1), ISSN: 2394-9333 www.ijtrd.com
 8. Fake review detection using machine learning techniques (Shilpa yadav¹, Dr.Gulbakshee Dharmela² , Khushali Mistry³)© 2021 JETIR April 2021, Volume 8, Issue 4
 9. Fake review detection from a product review using modified of iteration computation framework (Eka Dyar Wahyuni 1 and Arif Djunaidy) DOI: 10.1051/mateconf/20165803003 MATEC Web of Conferences 58 03003 (2016) mateconf/20165803003 BISSTECH 2015
 10. Recent state of art of fake review detection a comprehensive review (Richa Gupta^{1,2} , Vinita Jindal¹ , and Indu Kashyap²)The Knowledge Engineering Review (2024), 39, e8, pp. 1–47
[doi:10.1017/S0269888924000067](https://doi.org/10.1017/S0269888924000067)
 11. Intelligent fake review detection based on aspect extraction and analysis using learning (Gourav Bathla¹, Pardeep Singh¹, Rahul Kumar Singh¹ , Erik Cambria² , Rajeev Tiwari¹)Neural Computing and Applications (2022) 34:20213–20229
[https://doi.org/10.1007/s00521-022-07531-8\(0123456789\(\),-volV\)\(0123456789,-\(\).volV\)](https://doi.org/10.1007/s00521-022-07531-8(0123456789(),-volV)(0123456789,-().volV))
- Received: 12 August 2021 / Accepted: 9 June 2022 / Published online: 20 July 2022
 _ The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022
12. Fake review detection : Classification and analysis of real and pseudo reviews (Arjun Mukherjee[†], Vivek Venkataraman[‡], Bing Liu[†], Natalie Glance) [†]University of Illinois at Chicago, [‡]Google Inc
arjun4787@gmail.com, vivek1186@gmail.com, liub@cs.uic.edu, n glance@google.com