

# CONTROL OF FAKE REVIEWS IN THE ECOMMERCE ENVIRONMENT

**Naveen Manikanta Sai**  
U19IT039 / Dept of IT,  
Bharath Institute Of Higher Education  
And Research,  
naveenmanikantasai@gmail.com

**M.Ashok**  
U19IT037 / Dept of IT,  
Bharath Institute Of Higher  
Education And Research,  
Ashokkumar13609@gmail.com

**Dr.R.Yogesh Rajkumar**  
Asst Professor /Dept IT  
Bharath Institute Of Higher  
Education And Research,  
Yogeshraj कुमार.it@bharathuniv.ac.in

## Abstract

With the wide spread of fake online reviews, the detection of fake reviews has become a hot research issue. Despite the efforts of existing studies on fake review detection, the issues of imbalanced data and feature pruning still lack sufficient attention. To address these gaps, the present study proposes an ensemble model for the detection of fake online reviews. The model consists of four steps, and the first three steps are proposed to optimize the base classifiers: (i) Data resampling: We propose a novel way to address the data imbalance problem by combining the resampling and the grid search technique. (ii) Feature pruning: We propose an ablation study to drop unimportant features. (iii) Parameters optimization: We apply the grid search algorithm to determine suitable values of the relevant parameters for each base classifier. (iv) Classifier ensembling: We apply majority voting and stacking strategies to integrate the optimized base classifiers. The proposed data resampling method is also applied for the meta-classifier in the stacking ensemble model. This study produces advances in terms of combining different methods or algorithms into a model and the results show that the proposed ensemble model outperforms some existing techniques, thereby providing a new way to solve the data imbalance and feature pruning issues in the field of fake review detection

comments on the website before making a purchase decision. However, the current challenge is how new individuals can distinguish truthful reviews from fake ones, which later deceives customers, inflicts losses, and tarnishes the reputation of companies. The present paper attempts to develop an intelligent system that can detect fake reviews on ecommerce platforms using n-grams of the review text and sentiment scores given by the reviewer. The proposed methodology adopted in this study used a standard fake hotel review dataset for experimenting and data preprocessing methods and a term frequency-Inverse document frequency (TF-IDF) approach for extracting features and their representation. For detection and classification, n-grams of review texts were inputted into the constructed models to be classified as fake or truthful. However, the experiments were carried out using four different supervised machine-learning techniques and were trained and tested on a dataset collected from the Trip Advisor website. The classification results of these experiments showed that naïve Bayes (NB), support vector machine (SVM), adaptive boosting (AB), and random forest (RF) received 88%, 93%, 94%, and 95%, respectively, based on testing accuracy and the F1-score. The obtained results were compared with existing works that used the same dataset, and the proposed methods outperformed the comparable methods in terms of accuracy

## 1. Introduction

Fake reviews, also known as deceptive opinions, are used to mislead people and have gained more importance recently. This is due to the rapid increase in online marketing transactions, such as selling and purchasing. E-commerce provides a facility for customers to post reviews and comment about the product or service when purchased. New customers usually go through the posted reviews or

Online reviews have become an integral part of the decision-making process for consumers. They can make or break a business, and many consumers rely on them to make informed decisions. However, fake reviews have become a significant problem for businesses and consumers. Fake reviews can mislead consumers, damage a business's reputation, and even affect their revenue. Hence, the detection of fake reviews is essential to ensure the credibility and reliability of online reviews.

## EXISTING SYSTEM:

Online reviews form a critical and unavoidable facet of e-commerce. These reviews have a significant impact on consumers purchase decisions as well as the amount of money



spent by consumers. As e-commerce grows, however, so does the prevalence of fake online reviews (from here on, “fake reviews”). The proportion of fake reviews ranges from 16% [1], 20% [2], and 25% [3] to 33.3% [4]. As early as 2012, approximately 10.3% of online products were subjected to review manipulation [5].

2. There also exist infamous cases that demonstrate the seriousness of fake reviews in e-commerce. For example, in 2012, the UK Advertising Standards Authority found TripAdvisor to be involved in creating fake reviews: approximately 50 million online reviews on its site could not be verified as trusted [6]. In 2013, Samsung was ordered to pay a fine of \$340,000 by the Taiwan Federal Trade Commission for posting negative fake reviews against its competitor HTC [7]. In 2015, Amazon sued 1114 unidentified people for posting fake reviews [8]. In 2018, Mafengwo.com, a famous tourism platform in China, was involved in review fraud, which included activities such as duplicating online reviews from competitors; the platform subsequently admitted to the issue with fake reviews [9].

## 2. Fake Review Detection

Fake review detection is a subfield of natural language processing. It aims to analyze, detect, and classify product reviews on online e-commerce domains into fake or truthful reviews. In the last two decades, fake review analysis has become a popular research topic. Many researchers have performed studies on fake/spam review identification due to its significant effect on customers and e-commerce businesses. The process of construction or extraction of important features from text data is called feature engineering. Fake review detection studies have two orientations: a behavioral approach (spammer reviewer features) and a linguistic features approach, which relies on review centric features and relates only to the single review’s content. Both features are mentioned and described below

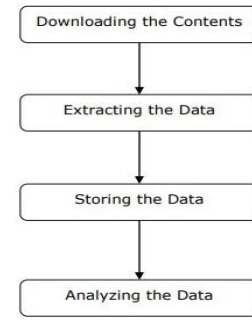
### Web Scraping

Web scraping is the process of using bots to extract content and data from a website. Unlike screen scraping, which only copies pixels displayed onscreen, web scraping extracts underlying HTML code and, with it, data stored in a database. The scraper can then replicate entire website content elsewhere

Web scraping software may directly access the World Wide Web using the Hypertext Transfer Protocol or a web browser. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot or web crawler. It is a form of copying in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis

### Methodology

In this study, we propose a hybrid model that combines text-based and metadata-based features to detect fake reviews accurately. The text-based features include bag-of-words, term frequency-inverse document frequency, and n-grams. The metadata-based features include reviewer's rating history, review length, and time of the review. We have used several classifiers, including Random Forest, Support Vector Machines, and Naive Bayes, to classify reviews as genuine or fake.



### Web Scraping

Fig. 1 presents the framework of the proposed methodology. It consists of seven steps.

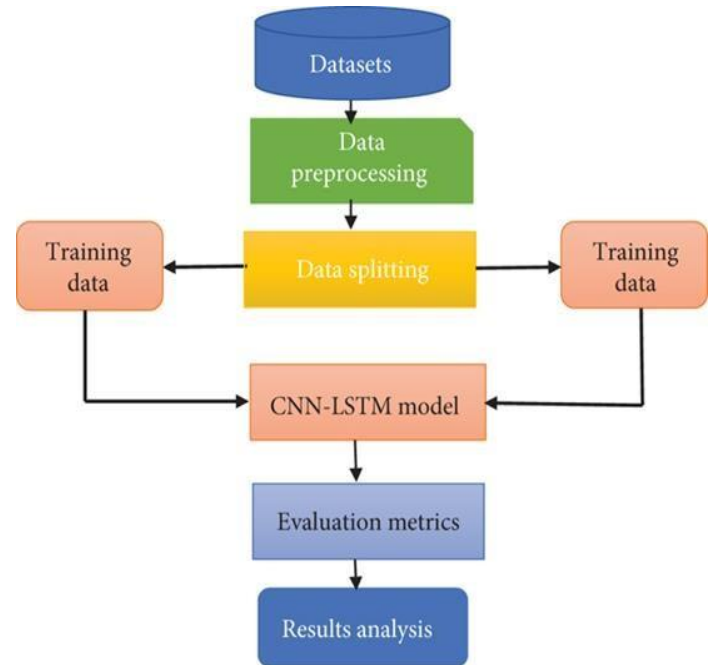
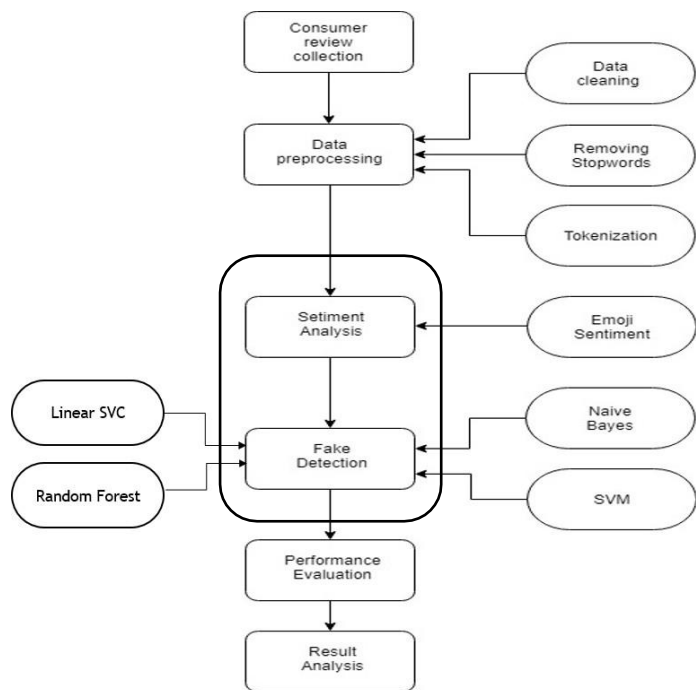


Fig.2. Framework for the proposed methodology

### Dataset

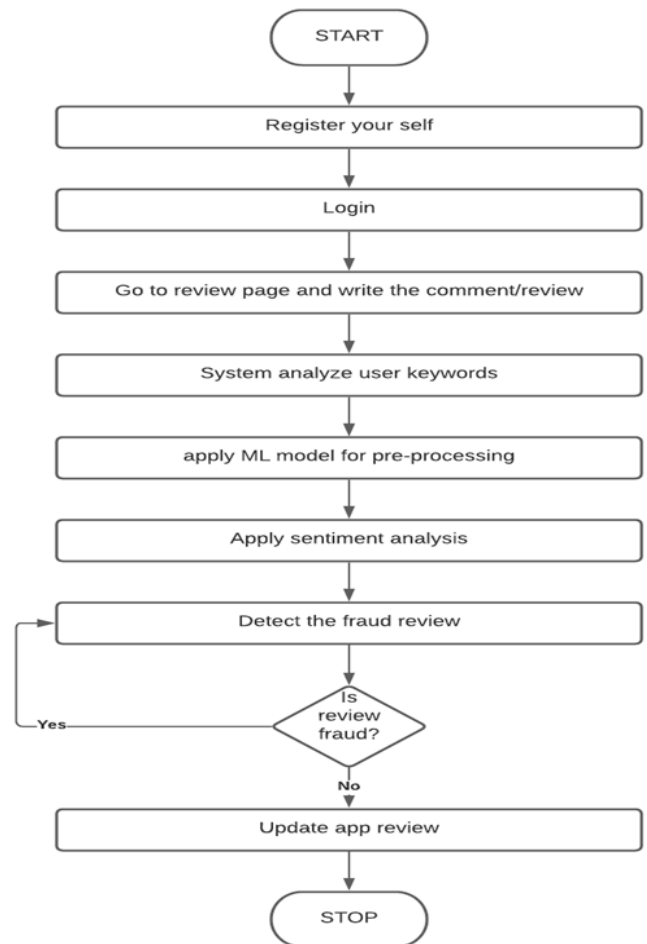
3. In this experiment, the gold standard dataset was developed by Ott et al. [34] was used. It contains 1,600 hotel reviews collected from one popular hotel booking website, Trip Advisor. The authors of this dataset refined all 5- and 3-star rated reviews from 20 hotels in Chicago. The dataset was pre-processed by adding features such as review length, four-grams, sentiment score, and POS. Every review had the following features
4. We collect dataset from amazon and flipkart



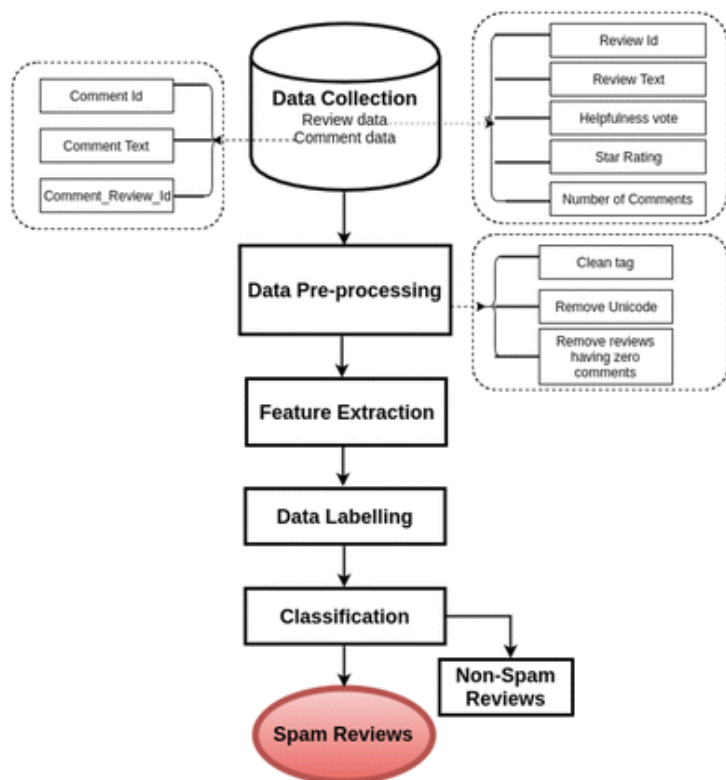
**Implementation Architecture**

### Comparative Analysis

While developing a novel system for spam/fake review detection, it is necessary to understand what techniques and datasets have already been utilized in previous studies. In this section, the results obtained by the proposed models were analyzed and compared with existing approaches based on the same dataset and accuracy metrics. shows the comparative analysis of the results of the proposed models with existing works.



**Flow diagram**



**Collect Data:** Gather a sample of reviews from the product or service in question.

**Pre-processing:** Clean the data by removing stop words, special characters, and punctuations.

**Feature Extraction:** Extract features such as sentiment, frequency of certain words, and length of the review.

**Data Split:** Divide the data into training and testing datasets.

**Train a Model:** Choose an appropriate model such as Logistic Regression, Random Forest, or Support Vector Machines (SVMs) and train it on the training dataset.

**Test the Model:** Use the testing dataset to test the accuracy of the model.

**False Positive Rate:** Check the false positive rate, which is the number of genuine reviews wrongly identified as fake.

**Adjust Threshold:** Adjust the threshold for identifying fake reviews to minimize the false positive rate.

**Predict:** Predict fake reviews using the trained model.

**Human Verification:** Verify the reviews predicted as fake through human verification to increase the accuracy of the model.

**Update Model:** Update the model regularly to improve accuracy and account for changes in the reviews over time.

Paper	Datas et used	Features used	Technique	Resul t (%)
Shojaee et.al.[11 ]	1600 hotel reviews from mechanical Turk and Trip Advisor websites.	Review's features	Naive Bayes SAGE	81 84
Fei et.al [13]	1600 hotel reviews from mechanical Turk and Trip Advisor websites.	Features of review and reviewer	Markov model	72

Ahmed et.al [21]	1600 hotel reviews from mechanical Turk and Trip Advisor websites.	Review's features	LSVM	90
Ott et.al [34]	1600 hotel reviews from mechanical Turk and Trip Advisor websites	Review's features	Naive Bayes SVM	89 88
Naryana et.al [35]	1600 hotels reviews from (AMT) Amazon.com	Review's features	Logistic regression	86 88
The Propose d work	1600 hotel reviews from Mechanical Turk and Trip Advisor websites	Review's features	Naïve Bayes SVM Ada boost RF	93 94 94

## Dataset

In this experiment, the gold standard dataset was developed by Ott et al. [34] was used. It contains 1,600 hotel reviews collected from one popular hotel booking website, Trip Advisor. The authors of this dataset refined all 5- and 3-star rated reviews

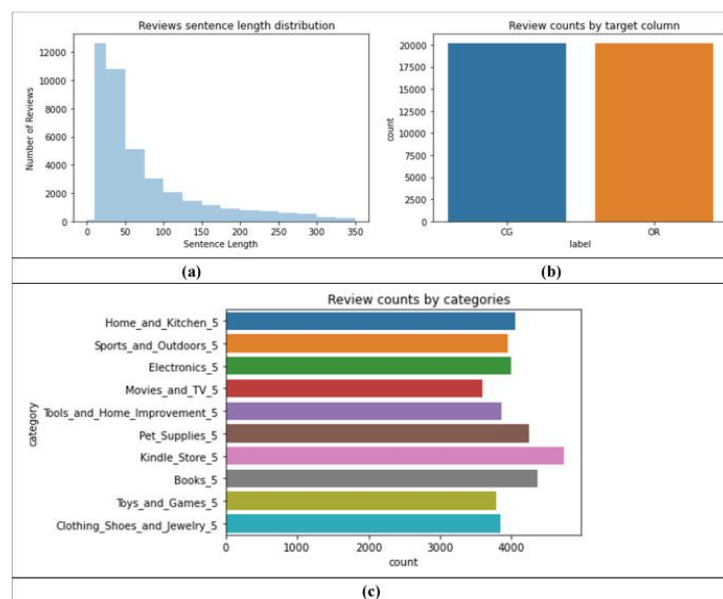
from 20 hotels in Chicago. The dataset was pre-processed by adding features such as review length, four-grams, sentiment score, and POS. Every review had the following features.

## Conclusions

The fake review detection is designed for filtering the fake reviews. In this research work SVM classification provided a better accuracy of classifying than the Naïve Bayes classifier for testing dataset. On

data. Revealing that it can generalize better and predict the fake reviews efficiently. This method can be applied over other sampled instances of the dataset. The data visualization helped in exploring the dataset and the features identified contributed to the accuracy of the classification. The various algorithms used, and their accuracies show how each of them have performed based on their accuracy factors.

Also, the approach provides the user with a functionality to recommend the most truthful reviews to enable the purchaser to make decisions about the product. Various factors such as adding new vectors like ratings, emojis, verified purchase have affected the accuracy of classifying the data better.



## References

- [1] J. Leskovec, "WebData Amazon reviews," [Online]. Available: <http://snap.stanford.edu/data/web-Amazon/links.html> [Accessed: October 2018].
- [2] J. Li, M. Ott, C. Cardie and E. Hovy, "Towards a General Rule for Identifying Deceptive Opinion Spam," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, vol. 1, no. 11, pp. 1566-1576, November 2014.).
- [3] M. Luca and G. Zervas (2016) challenged the organisations 'Fake it till you make it: Reputation, competition, and yelp review fraud'.
- [4] B. Liu and M. Hu, "Opinion Mining, Sentiment Analysis and Opinion Spam Detection,"[Online]. Available: [https://www.cs.uic.edu/~liub/FBS/sentiment\\_analysis.html#lexicon](https://www.cs.uic.edu/~liub/FBS/sentiment_analysis.html#lexicon) [Accessed: January 2019]. the other hand, the Naïve Bayes classifier has performed better than other algorithms on the training

- [5] Suresh Kumar, K., Nassa, V. K., Uike, D., Sahu, A. K., Athavale, V. A., & Saravanan, V. (2022). A Comparative Analysis of Blockchain in Enhancing the Drug Traceability in Edible Foods Using Multiple Regression Analysis. *Journal of Food Quality*, 2022.
- [6] E. I. Elmurngi and A.Gherbi, "Unfair Reviews Detection on Amazon Reviews using Sentiment Analysis with Supervised Learning Techniques," *Journal of Computer Science*, vol. 14, no. 5, pp. 714–726, June 2018
- [7] N. O'Brien, "Machine Learning for Detection of Fake News," [Online]. Available: <https://dspace.mit.edu/bitstream/handle/1721.1/119727/1078649610-MIT.pdf> [Accessed: November 2018]
- [8] J. C. S. Reis, A. Correia, F. Murai, A. Veloso, and F. Benevenuto, "Supervised Learning for Fake News Detection," *IEEE Intelligent Systems*, vol. 34, no. 2, pp. 76-81, May 2019.
- [9] B. Wagh, J. V. Shinde and P. A. Kale, "A Twitter Sentiment Analysis Using NLTK and Machine Learning Techniques," *International Journal of Emerging Research in Management and Technology*, vol. 6, no. 12, pp. 37-44, December 2017.
- [10] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," in *Proceedings of AAAI-98 Workshop on Learning for Text Categorization*, Pittsburgh, PA, USA, vol. 752, no. 1, pp. 41-48, July 1998.
- [11] B. Liu and M. Hu, "Opinion Mining, Sentiment Analysis and Opinion Spam Detection," [Online]. Available: <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon> [Accessed: January 2019].
- [12] C. Hill, "10 Secrets to Uncovering which Online Reviews are Fake," [Online]. Available: <https://www.marketwatch.com/story/10-secrets-to-uncovering-which-online-reviews-are-fake-2018-09-21> [Accessed: March 2019].
- [13] .J. Novak, "List archive Emojis," [Online]. Available: <https://li.st/jesseno/positive-negative-and-neutral-emojis-6EGfnd2QhBsa3t6Gp0FRP9> [Accessed: June 2019].