

Fake Product Review Detection and Sentiment Analysis using Machine Learning

Priti Ashok Bankar¹, Prof. Manisha Patil²

¹Student, Department of Computer Engineering, Trinity College of Engineering and Research, Pune,

²Assistant Professor, Department of Computer Engineering, Trinity College of Engineering and Research, Pune

Abstract

With the way the internet has changed shopping, people have become much more reliant on online reviews. However, this has its downsides. Even with some people trying to catch fake reviews, the methods people use to write convincing fake reviews continue to become more complex. Because of this, attempting to automate the process of distinguishing fake reviews using the older, rule-based methods has produced unsatisfactory results at a larger scale. Because of this, better and more advanced methods of machine learning must be a part of the solution to this problem. For the framework established for this study, we will be using supervised machine learning with a focus on the detection of fake product reviews. Multiple data collection methods will be combined and followed up with filtering procedures. Pre-processing and cleaning methods will be used to simplify the data. Throughout this, methods of feature extraction and feature selection will be employed to discern and obtain data relevant to the problem. To determine the most suitable approach and methods, we will be using the ANN and the Random Forest models. To assess each of these models, we will measure and evaluate them using the metrics of accuracy and the binary classifications of truth (precision), coverage (recall), and the F1 score. Out of the two-machine learning(ML) models, both were successful in detecting fake reviews. The F1 score for the ANN ML model was 98.06% which would mean that the accuracy, truth, and coverage of the model were 95.88%, 100%, and 96.19%, respectively. The Random Forest model, on the other hand, was able to achieve a 100% score for truth, coverage, F1, and an accuracy of 97.94%. Because of this, its ability to classify was better and it was overall more machine learning capable. In summary, the system detects fraudulent reviews and does this fairly and reliably. The results show that ensemble learning methods, and particularly Random Forest, outperform methods that only use one model, in this case the ANN. In this regard, the research can be expanded by adding deep learning or real-time analyses to improve the efficiency of the detection system within constantly evolving or changing environments.

Keywords: Fake reviews, Sentiment analysis, Machine Learning, Deep Learning, Opinion mining, Natural Language Processing (NLP).

1. Introduction

E-commerce means that customers consider goods or services in an entirely different way. Customer reviews provide a guide for making purchases in a digital mid-market. Therein lies a challenge. That challenge comes in the form of both positive and negative reviews being manipulated. Fake reviews mislead customers and giving products negative or positive juries. In a free market of products and services

these manipulated reviews challenge the legitimacy of a marketplace. This manipulation affects a marketplace's credibility and reconstructs relationships of trust, resulting in a financial loss for a marketplace. The time and effort needed to find fraudulent reviews in a marketplace are counterproductive because the reviewer finder ultimately relies on a human being.

Awaken your optimism, for the world has previously worked in favor of machine learning. These murky models dissect digital and inscriptive data. That being said, of utmost importance is sentiment analysis and where it pertains to a review's heart: its emotional truth or false. The fraudulent review marketplace has shifted the landscape and will continue to pose challenges to a plethora of machine learning models that utilize support vector machines or the naive Bayes, just to name a couple. The hope for the models is that they bring some semblance of control, order, and of course, ultimately and especially, organized data—specifically, what data is the most pertinent and most flexible.

This study aims to develop a system to detect fake reviews for products and to perform sentiment analysis on reviews. Roughly, the system generates fake review samples for various products for the purpose of testing. The system will then be trained to develop a sentiment analysis and classification reviews on a scale ranging from original to fake. This study aims to achieve the need for a more sophisticated sentiment analysis, review classification and system justification. The results suggest that the successful implementation of an Artificial Neural Network and Random Forest to classify review and provide sentiment analysis to the system justifies the classification of review original to fake.

2. Literature Survey

According to [1], sentiment analysis of product reviews from e-commerce platforms has been developed using an advanced deep learning framework. The primary aim is to analyze user reviews and classify the comments into positive, negative, and neutral sentiments, which improves the reliability of online reviews. The framework achieves high precision when identifying genuine comments using Convolutional Neural Networks (CNN) and Recurrent Neural Network (RNN) architectures. This approach improves customer confidence and provides businesses with vital information on the trends of customer contentment.

The supervised learning approach that is based on data in [2] proposes a strategy to identify fake reviews on the web. The said approach utilizes textual recordings and behavioral patterns to identify fraudulent activities. The performance of the methodologies (Random Forest, Support Vector Machine (SVM), and Naïve Bayes) were individually studied. The studies prove that the aforementioned machine learning techniques significantly improve the ability to identify fake reviews which, in turn, facilitates the automation of trust assessments in online business transactions.

SentiDeceptive [3] is the first model of its kind. Its originality lies in its ability to identify social media product reviews that are deceptive. The model employs sentiment analysis and feature extraction. Both of which assist in genuine feedback versus false feedback. The model's experimental results show that hybrid models that involve both the lexicon and the deep learning sides of things. More than traditional classifiers, it improves transparency and decreases misinformation in reviews that are online.

In [4], Elmogy et al. tried to assess the results of the various techniques arising from the field of supervised learning and the specific problem of detecting fraudulent reviews of the web. In addition to the reviews, the review writers, and specific pieces that the review is composed of, they way in which traits are used to formulate the classification algorithm. The evaluation of the classification algorithm's predictive capability was performed using both the decision tree and logistic regression techniques. The

results of the technique show that the supervised learning paradigm is a promising model for the automation of the review detection mechanism in online review systems.

As described in [5], the Support Vector Machine (SVM) methodology was implemented for detection and classification of false Amazon reviews. The aim of the study is to uncover unsubstantiated reviews that erode the trust in the products. In order to improve the predictive accuracy, SVM is combined with a certain set of Preprocessing techniques (like tokenization, TF-IDF, and feature selection). The results from the experiments show that SVM is one of the best and reliable techniques to identify genuine reviews from the fraudulent ones.

Alsubari et al. [6] built on their previous work and further streamlined their fraud review detection Supervised Learning system by incorporating a system with a more varied set of the internet data, including linguistic and behavioral attributes. The improved review detection system is supported by the evaluation metrics, namely accuracy and recall. The study stresses the importance of data analytics in developing efficient and comprehensive strategies for the review fraud problem.

The use of a hybrid method involving Deep Learning and Aspect-based Sentiment Analysis for identifying fraudulent online reviews is detailed in [7]. The software examines product-related features and corresponding sentiment expressions in order to spot inconsistencies that signify fraudulent activity. Deep Learning techniques like Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) capture contextual and semantic relationships. Enhanced accuracy and better association of fraudulent reviews with particular product features is reported.

Ennaouri and Zellou [8] extensively survey review fraud detection with the use of machine learning. It categorizes detection models into three categories: supervised, unsupervised, and hybrid; and explains each of their respective strengths and weaknesses. The survey notes the increasing application of Deep Learning and Natural Language Processing to fight opinion spam but calls for better explainable models and varied benchmark datasets in future studies.

A [9] machine learning-based system has been created to improve customer confidence by automatically detecting and removing fraudulent reviews. The approach uses text mining and classification algorithms to identify and remove spam. Several data preparation strategies like removing stop-words and stemming were also used to improve the outcome of the data. The findings reveal that machine learning-based review fraud detection techniques are useful in reducing fraudulent reviews and increasing the trustworthiness of e-commerce websites.

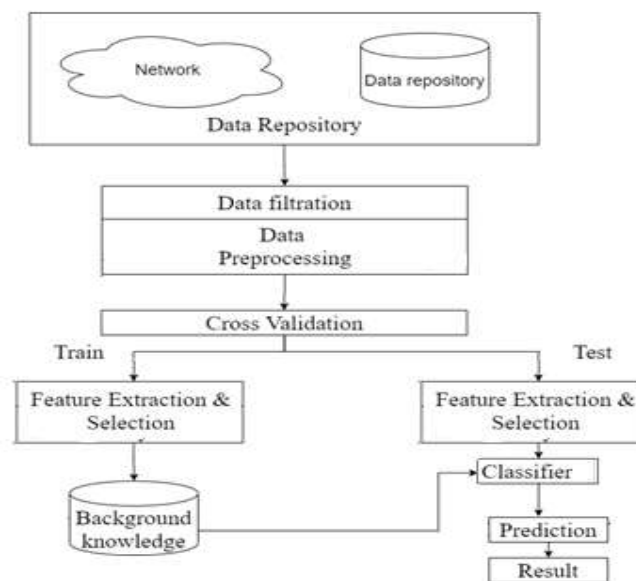
Qazi et al. [10] conducted an opinion spam detection research by examining various machine learning methods. The study explores various datasets, presenting performance differences of SVM, decision trees, and ensemble methods. The study emphasizes the importance of integrated systems and deep learning for detection accuracy and adaptability. The study identifies the class imbalance, dynamic spam detection, and flexible, comprehensive algorithms to detect false reviews as key research areas

3. Methodology

The system has the ability to predict online reviews by analyzing and processing text data. The system and real time text data that any application can work with can be viewed as an integrated ecosystem. The system has the modules for both the training and testing phases of the classification. It can classify text data using symbolic representation and can predict and differentiate Review data forgery. The system has a collection of algorithms for the classification of Review data and differentiates based on the potentially harmful Review data. Review the system for its accuracy and false prediction ratio.

The two classification machine learning models are Artificial Neural Network (ANN) and Random Forest (RF). ANN has a higher ability to learn and recognize nonlinear relationships and uses hidden layers. These hidden layers of ANN are trained using a technique called backpropagation. The Random Forest classifier is a bit different. It builds a classifier based on an ensemble of decision trees. During the training of this ensemble, each decision tree is trained on different random subsets of the data and features. The final prediction is made based on the votes of the decision tree ensemble. Due to this voting mechanism, Random Forests are less prone to overfitting to the training data and captures the data pattern more reliably..

Figure 1: Proposed System Architecture



Model Process

For our fake review detection system, we’ve laid out some simple steps to help us organize and communicate our progress. Our upcoming engineering collaboration with Natural Language Processing and Machine Learning has made us excited and optimistic for our upcoming releases.

Data Collection: The goal for step one is to gather product reviews online. We want to capture reviews from both e-commerce sites and social media to collect an assortment of reviews. Because reviews are purchased, we need to capture locations with known reviews and random unreviewed locations to capture real and fake reviews.

Data Preprocessing: Text data is incredibly messy, we will need to clean this up. We remove words that aren't as important, we stem and lemmatize, and we will need to tokenize as well. After this we will be left with much less noise with much higher quality data.

Feature Extraction and Feature Selection: Machine Learning requires us to create a list of features that we will extract, and we will do this by implementing Bag of Words. However, we won’t be able to include every feature and will need to make a significant judgment to help the system learn to spot which reviews are fake or real!

Training Models: After we create our data and feature set, we are going to run a few Machine Learning type models. We will be using Artificial Neural Networks and Random Forest models among other types.

Classification and Prediction: The system can classify whether a review is fake or genuine in real-time or during batch processing. It may also conduct a sentiment review of the study. A new method according to the proposal will be shown in a report or in a graphic. The model is tested with data that hasn't been seen before. Accuracy, precision, recall, and F1-score give an indication of how effective the system is

Algorithm Process

Algorithm 1: RF

Input: Selected feature of all test instances $D[i\dots n]$, Training database policies $\{T[1] \dots T[n]\}$

Output: No. of probable classified trees with weight and label.

Step 1: Read (D into $D[i]$)

$V \leftarrow$ Extract features (D)

Step 2: $N \leftarrow$ Count Features (D)

Step 3: for each(c into TrainDB)

Step 4: $Nc[i] \rightarrow$ Ext Features(c)

Step 5: select relevant features of $w = \{Nc[i], N\}$

Step 6: Statement ($w > t$)

Step 7: Return Tree Instance $\{ Nc[i], N, w, \text{label} \}$

Algorithm 2: ANN

Training Process

Input: Training dataset Train-Data [], Many activation functions [], Threshold Th

Output: Extracted Features Feature set[] for a trained module that has been finished.

Step 1: Set the data input block $d[]$, the activation function, and the epoch size.

Step 2: Features-pkl \leftarrow Feature-Extraction ($d[]$)

Step 3: Feature-set [] \leftarrow optimized (Features-pkl)

Step 4: Return Feature-set []

Testing Process

Input: Extracted features of testing instances set Data $[i\dots n]$, Train data policies PSet[41].....T[n]

Output: Predication Real and Fake.

Steps:

1. For each (Data [i] into Data) choose n attributes from Data [i] using below formula,

$$\text{Testset}(k) = \sum_{k=1}^n \text{attribute } [D[i]k \dots D[n]n]$$

2. For each (PSet [i] from PSet),

$$\text{Train}[m] = \sum_{m=1}^n \text{attribute } [T[i]k \dots T[n]n]$$

3. Evaluate train and test instances using below formula,

$$\text{weight}[k] = \text{similarity}(\text{Testset}[k] \sum_{m=1}^n \text{Trainset } [m])$$

4. If ($\text{weight}[k] > \text{Th}$),

Testset [k].class \leftarrow Train[m]: class

Break;

5. Return class label

4. Results

The results from the experiments show the proposed detection system for fake product-based reviews using ANN and Random Forest (RF) classifiers is promising. For ANN models the results are accuracy = 95.88%, precision = 100%, recall = 96.19%, and F1 score = 98.06%. These results show the strength of the model in classifying product reviews and indicates a relatively low false positive rate. For Random Forest the results showed a higher score than ANN at 97.94% for accuracy, and a perfect score of 100% for precision, recall, and F1 score. These results show a better system in identifying negative and positive reviews of products. The graphical analysis shows that the Random Forest classifiers provided the best and most consistent results in all performance measures.

Figure 2: ANN Performance Evaluation

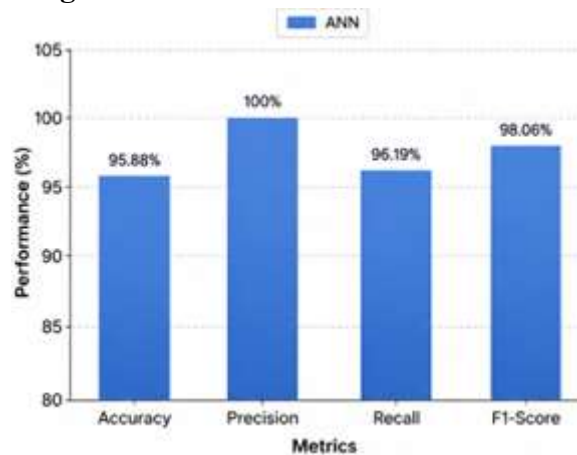


Figure 2 shows the performance metrics (accuracy, precision, recall, and F1-score) of the Artificial Neural Network (ANN) model. The ANN model exhibits 95.88% accuracy, 100% precision, 96.19% recall, and 98.06% F1-score. The ANN model shows that it can detect fake reviews use easily with very few false positives and very good general classification capability.

Figure 3: RF Performance Evaluation

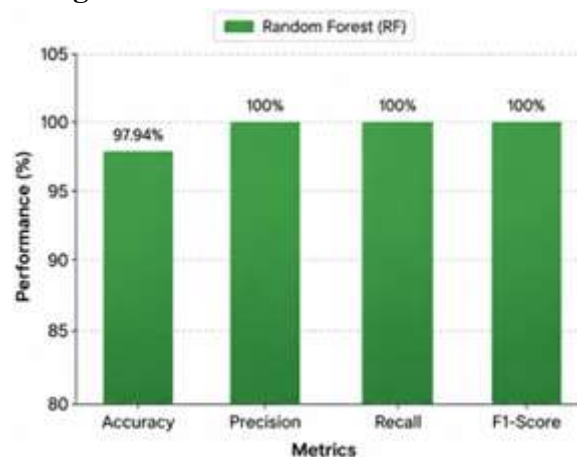


Figure 3 shows how the Random Forest (RF) model performs using the same evaluation criteria. The RF model reports 97.94% accuracy, and 100% for precision, recall, and F1-score. The Random Forest model

performs significantly better than the ANN, as the Random Forest model offers a more accurate, and more robust classification with 0 false positives. Thus, the Random Forest model serves more effectively to identify fake reviews.

5. Conclusion

This study is a good example because it offers a fake product review detection system along with a sentiment analysis system, both powered by machine learning. The product reviews and relevant user content speak for themselves. The study focuses on building a machine learning system with a well-defined pipeline comprising data carving, feature designing, and sorting. This pipeline looks to solve the problems posed by the data which is mostly unstructured, and in this case, mostly noise. Some preliminary strategies such as tokenization and the removal of both stop-words and morphology lead to a considerable amount of decimation of data for the reason that the irrelevant elements are eliminated. This classification framework detects fake online reviews with the help of machine learning. Both the ANN and Random Forest models had good results. However, the Random Forest model was superior due to the “ensemble learning” methodology. The study showed that the framework was able to detect fraudulent reviews. Further study can improve the framework with the fusion of deep learning and the real time processing of reviews/sentiments.

6. References

1. Alzahrani, Mohammad Eid, et al. "Developing an Intelligent System with Deep Learning Algorithms for Sentiment Analysis of E - Commerce Product Reviews." *Computational Intelligence and Neuroscience* 2022.1 (2022): 3840071.
2. Alsubari, S. Nagi, et al. "Data analytics for the identification of fake reviews using supervised learning." *Computers, Materials & Continua* 70.2 (2022): 3189-3204.
3. Marwat, M. Irfan, et al. "Sentiment Analysis of Product Reviews to Identify Deceptive Rating Information in Social Media: A SentiDeceptive Approach." *KSII Transactions on Internet & Information Systems* 16.3 (2022).
4. Elmogy, Ahmed M., et al. "Fake reviews detection using supervised machine learning." *International Journal of Advanced Computer Science and Applications* 12.1 (2021).
5. Tabany, Myasar, and Meriem Gueffal. "Sentiment analysis and fake amazon reviews classification using SVM supervised machine learning model." *Journal of Advances in Information Technology* 15.1 (2024): 49-58.
6. Alsubari, S. Nagi, et al. "Data analytics for the identification of fake reviews using supervised learning." *Computers, Materials & Continua* 70.2 (2022): 3189-3204.
7. Bathla, Gourav, et al. "Intelligent fake reviews detection based on aspect extraction and analysis using deep learning." *Neural Computing and Applications* 34.22 (2022): 20213-20229.
8. Ennaouri, Mohammed, and Ahmed Zellou. "Machine learning approaches for fake reviews detection: A systematic literature review." *Journal of Web Engineering* 22.5 (2023): 821-848.
9. Bhattacharya, Swarnajit. "Monitoring and Removal of Fake Product Review Using Machine Learning (MI)." (2023).
10. Qazi, Atika, et al. "Machine learning-based opinion spam detection: A systematic literature review." *IEEE Access* 12 (2024): 143485-143499.

11. Park, Minjung, and Sangmi Chai. "Constructing a user-centered fake news detection model by using classification algorithms in machine learning techniques." *IEEE Access* 11 (2023): 71517-71527.
12. Alshehri, Asma Hassan. "An Online Fake Review Detection Approach Using Famous Machine Learning Algorithms." *Computers, Materials & Continua* 78.2 (2024).
13. Jaiswal, Mohini, and Deepali Javale. "Fake Product Review Monitoring System." 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). Vol. 2. IEEE, 2024.
14. Mutemi, Abed, and Fernando Bacao. "E-commerce fraud detection based on machine learning techniques: Systematic literature review." *Big Data Mining and Analytics* 7.2 (2024): 419-444.
15. Qazi, Atika, et al. "Machine learning-based opinion spam detection: A systematic literature review." *IEEE Access* 12 (2024): 143485-143499.

Licensed under [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)