Review of Machine Learning Algorithms for Fake Review Detection using Amazon Dataset

Navin Kr.Goyal¹, Bright Keswani², DineshGoyal³

¹Research Scholar, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
²Professor, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
³Professor, Poornima Institute of Engineering and Technology, Jaipur, Rajasthan, India

Abstract. In current world, fake news on social media, fake reviews of products on e-commerce platforms identification is very important and difficult area of opinion mining for sentiment analysis. The machine learning techniques can process and automate of finding fraudulent news/reviews on social media/e-commerce platforms and can differentiate into categories like ham and spam. Many systems have been developed for creating huge amount of fake reviews by the Spammers. Along this developers and researcher have developed and recommended techniques to detect fake reviews. This paper provides a detailed review of existing techniques and methods for finding spam reviews. Apart from reviewing the state-of-the-art research studies on spam review detection, a consolidated study on techniques of machine learning for spam review detection has been proposed. Moreover, its focus on existing research work gaps and future recommendations for spam review identification

Keywords: Fake Reviews, Naïve Bayes, Random Forest, Review Detection, Machine Learning.

1. Introduction

Reading product reviews before buying the product becomes a habit, especially for potential customers. If they want to buy a product, they usually read reviews from some customers about the current product. If the review is mostly positive, there is a big chance to buy the product, otherwise if it's mostly negative, they tend to buy other products. While, for a company, the positive reviews from customers can generate significant financial benefits for businesses, it can be used as input for decisions related to product design and what services are provided to customers. [15]

On E-commerce platforms, positive or negative review of a product may highly impact the consumer decision making during buying process. A meta-analysis on empirical studies found that both review quantity and review integrity is significant determinants of sellers. Despite the advantages provided by social media, the quality of news on social media is lower than traditional news organizations. However, because it is cheap to provide news online and much faster and easier to disseminate through social media, large volumes of fake news, i.e., those news articles with intentionally false information, are produced online for a variety of purposes, such as financial and political gain. [3]

2. Literature Review

In this paper author propose a feasible solution which improves the level of accuracy with good time efficiency on Tweeter data. A novel feature combination scheme which utilizes the sentiment lexicons and the extracted tweet unigrams of high information gain. We evaluate the performance of six popular machine learning classifiers among which the Naive Bayes Multinomial (NBM) classifier achieves the accuracy rate of 84.60% and takes a few minutes to complete classifying thousands of tweets. [1]

In this paper, some selected machine learning classifier algorithms were applied on crawled Twitter data after applying different types of preprocessors and encoding techniques, which ended up with satisfying accuracy. Later a comparison between the achieved accuracies was showed. Experimental evaluations show that the Neural Network Classifier' algorithm provides a remarkable accuracy of 81.33% compared with other classifiers. [2]

Social media for news consumption is a double-edged sword. On the one hand, its low cost, easy access, and rapid dissemination of information lead people to seek out and consume news from social media. On the other hand, it enables the wide spread of fake news", i.e., low quality news with intentionally false information. The extensive spread of fake news has the potential for extremely negative impacts on individuals and society. Therefore, fake news detection on social media has recently become an emerging research that is attracting tremendous attention. Fake news detection on social media presents unique characteristics and challenges that make existing detection

algorithms from traditional news media ineffective or not applicable. First, fake news is intentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content; therefore, we need to include auxiliary information, such as user social engagements on social media, to help make a determination. Second, exploiting this auxiliary information is challenging in and of itself as users' social engagements with fake news produce data that is big, incomplete, unstructured, and noisy. Because the issue of fake news detection on social media is both challenging and relevant, we conducted this survey to further facilitate research on the problem. In this survey, we present a comprehensive review of detecting fake news on social media, including fake news characterizations on psychology and social theories, existing algorithms from a data mining perspective, evaluation metrics and representative datasets. We also discuss related research areas, open problems, and future research directions for fake news detection on social media. [3]

This paper is to provide a strong and comprehensive comparative study of current research on detecting review spam using various machine learning techniques and to devise methodology for conducting further investigation. [4]

In this paper, a complete review of existing techniques and strategies for detecting spam review is discussed. Apart from reviewing the state-of-the-art research studies on spam review detection, taxonomy on techniques of machine learning for spam review detection has been proposed. Moreover, its focus on research gaps and future recommendations for spam review identification. [5]

The aim of this paper is to examine the performance of several machine learning methods used for the detection of positive and negative fake consumer reviews. Here we show that deep neural networks, including convolutional neural networks and long short term memory, significantly outperform the traditional machine learning methods in terms of accuracy while preserving desirable time performance. [6]

The auhor worked towards developing a tool which will classify the reviews as fake or genuine and provide it to the user. The proposed tool will operate on the principle of the Weighted Ensemble Classifier. The said ensemble classifier uses Support Vector Machine (SVM), Naive Bayes (NB) and KNearest Neighbor (KNN) classifiers. Experimental results show that the proposed ensemble classifier is efficient in the fake review detection task. [7]

In this paper, author introduces a hybrid method that combines: (a) basic features and feature expansion based on Term Frequency–Inverse Document Frequency (TF-IDF) and (b) basic features and feature expansion based on tweet-based features. Author trains three most common classifiers for this field, i.e., Support Vector Machine (SVM), Logistic Regression (Logit), and Naïve Bayes (NB). From those two feature expansions, we do notice a significant increase in feature expansion with tweet-based features rather than based on TF-IDF, where the highest accuracy of 98.81% is achieved in Logistic Regression Classifier. [8]

This paper evaluates the performance of ensemble learning based approaches to identify bogus online information. The application of a number of ensemble learning-based approaches to a collection of fake restaurant reviews that developed show that these ensemble learning-based approaches detect deceptive information better than conventional machine learning algorithms. [9]

This paper focuses on analysis of 2017 to 2021 papers and analysis of different fake news detection techniques. This survey gives a far-reaching review about the recent and past examinations on false news detection using different ML algorithms. [10]

3. Classification Strategies

In the wide-area of machine learning which is a sub branch of AI, bunches of classification strategies have been created. Classification strategies are utilized to group the unleveled information. These days, classification is the most helpful method to distinguish unstructured information. The primary reason for the characterization is to distinguish classes or gatherings. Grouping procedures we utilized rely upon the extremity on the grounds that each classifier works uniquely in contrast to each other.

$$Polarity = rac{P(True_Review)/P(Total_Reviews)}{P(True_Negative)/P(Total_Reviews)}$$

In this paper, we have implemented many machine learning classifiers that are given below.

- ➢ Naïve Bayes (NB)
- Support Vector Machines (SVM)
- Neural Network (NN)
- ➢ K-Nearest Neighbors (KNN)

3.1. Naïve Bayes

The Naïve Bayes algorithm is based on the Bayes theorem and consists of several popular features. Mostly, Naïve Bayes predicts the data based on various attributes. It is the most popular classification technique to solve real-world problems. The algorithm requires a small amount of training data to identify the necessary parameters and it is also a fast classifier compared to other classifiers [11]. Recently, Naïve Bayes is the most common technique for text classification into multiple classes but recently it is utilized for sentiment analysis [12].

$$P(\frac{x}{y}) = \frac{P(\frac{y}{x}) * P(x)}{P(y)}$$

Here, P(x/y) is the posterior probability of class given predictor and P(y/x) is the likelihood probability of predictor given class. On the other hand, P(x) is the prior probability class and P(y) is the prior probability of predictor.

3.2. Support Vector Machine (SVM)

Support Vector Machine is a supervised machine learning algorithm used for both classification and regression problems [13]. But it is more popular to solve the classification problems. SVM use kernel functions for categorizing data, text, images as well as vectors [14]. In the SVM algorithm, they utilize hyper-plane to separate the two classes [14]. Hyper-plane effectively separates relevant pieces of information and furthermore assists with tracking down the most extreme distance between the closest data of interest.

3.3. Neural Network

The neural network is a growing experience of the human cerebrum and comprises of a wide scope of layers of neurons. In each layer, getting inputs from past layers and ascertain the result from it again passing the results to the following layers [14]. It is a constant cycle and in the last stage gets the last results.



Fig-1 Neural Network working process

The figure-1 shows the functioning system of a neural network. Assume, I is the initial input of the interaction then it passed the neurons to the main layer (A1). From that point forward, the principal layer (A1) gets the contribution from the past and produces a result. Then, at that point, this result is again passed to the subsequent layer (B1), it is determined the result in view of the principal layer. Finally, the consolidated result of the subsequent layer is known as the last result of the model.

4. Proposed Architecture

We have previously depicted which classifiers have been utilized for our research work. Figure 2 provides us with an exceptionally essential thought of our proposed design or procedure of this work.

Here classification procedures were applied in light of the accompanying construction.

To start with, we cleaned the entire review and changed over them into CSV design. Then Tokenization was utilized to tokenize the review's statement by word. After that Stopword filter technique was utilized to eliminate the pointless words from the tokenized reviews. Then WordNet Lemmatizer was utilized to transform a word into its base word (for example eaten—eat, running \rightarrow run).

Presently comes the encoding part. Here level Encoder was utilized to encode the feeling and OneHot Encoder was utilized to encode the review. Finally, the encoded numeric structure was taken care of in the chose classification strategy to accomplish our outcome.



Figure-2 Flow diagram of proposed methodology

5. Conclusion and Future Work

Data is very critical in current scenario. A Huge amount of data in the form of reviews and opinion of users are generated on social media and other e-commerce platforms on the internet every day. By this reason, opinion mining or sentiment analysis has become very popular research area in current days. This paper, demonstrate the comparison between different machine learning techniques so different machine learning classifier algorithms were used to analyze a bulk amount of serial Twitter reviews after applying different text pre processing techniques like Removing punctuations (., ! \$() * % @), Removing URLs, Removing Stop words, converting the entire review into Lower casing, Tokenization, Stemming and Lemmatization. Presently exactnesses can be fluctuated relying upon the dataset and approach to utilizing algorithms. Among every one of the classifiers, Neural Network Classifier was the ideal one for more noteworthy exactness (81.33%) and accuracy here. The absolute most significant and entertaining thing about research is that there is no best arrangement till now. There is generally an improved

solution for the existing issues. Amazon information was utilized in this paper which can be utilized as an aiding device for various basic errands effectively like observing examples or exercises like distinguishing counterfeit surveys. There are a few restrictions and bunches of additional functioning degree on this paper. This work should be possible on a significantly greater informational collection to acquire better precision. The greater dataset will be utilized the more grounded model can be accomplished. This work can be reached out by deciding subjectivity as the machine can decide an individual's opinion whether it is unbiased or emotional. By the dataset of this paper, a module can be made utilizing an installed framework, where an individual will include his voice and in the wake of dissecting the sentence the module will answer naturally about the individual's opinion. This work can be utilized to fabricate a model that can identify sentences about numerous faculties. Likewise, an electronic utilization of opinion examination can be made from here on out.

References

- 1. A. Yang, J. Zhang, L. Pan and Y. Xiang, "Enhanced Twitter Sentiment Analysis by Using Feature Selection and Combination," 2015 International Symposium on Security and Privacy in Social Networks and Big Data (SocialSec), 2015, pp. 52-57, doi: 10.1109/SocialSec2015.9.
- Mostafa, Golam & Ahmed, Ikhtiar & Junayed, Masum Shah. (2021). Investigation of Different Machine Learning Algorithms to Determine Human Sentiment Using Twitter Data. International Journal of Information Technology and Computer Science. 2. 38-48. 10.5815/ijitcs.2021.02.04.
- 3. Shu, Kai, et al. "Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19.1 (2017): 22-36.
- 4. Crawford, M., Khoshgoftaar, T.M., Prusa, J.D. *et al.* Survey of review spam detection using machine learning techniques. *Journal of Big Data* **2**, 23 (2015). <u>https://doi.org/10.1186/s40537-015-0029-9</u>
- 5. Farooq, Muhammad Shoaib (2020): Spam Review Detection: A Systematic Literature Review. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.12951077.v1
- 6. Hájek, Petr & Barushka, Aliaksandr. (2019). A Comparative Study of Machine Learning Methods for Detection of Fake Online Consumer Reviews. 18-22. 10.1145/3383902.3383909.
- R. Furia, K. Gaikwad, K. Mandalya and A. Godbole, "Tool For Review Analysis Of Product," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-6, doi: 10.1109/ic-ETITE47903.2020.104.
- E. B. Setiawan, D. H. Widyantoro and K. Surendro, "Feature Expansion for Sentiment Analysis in Twitter," 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2018, pp. 509-513, doi: 10.1109/EECSI.2018.8752851.
- 9. Gutierrez-Espinoza, Luis, et al. "Fake reviews detection through ensemble learning." *arXiv preprint arXiv:2006.07912* (2020).
- S. Kumar, S. Kumar, P. Yadav and M. Bagri, "A Survey on Analysis of Fake News Detection Techniques," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 894-899, doi: 10.1109/ICAIS50930.2021.9395978.
- 11. Understanding Support Vector Machine (SVM) algorithm from examples https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/", Last Accessed: 09 July, 2020.
- 12. Shuo Xu "Bayesian Naïve Bayes Classifiers to text Classification", Vol.44, Issue:2018
- 13. Classification using neural Networks "https://towardsdatascience.com/classification-using-neuralnetworks-b8e98f3a904f", Last Accessed: 09 July, 2020.
- 14. Logistic Regression "https://www.statisticssolutions.com/what-is-logistic-regression/", Last Accessed: 09 July, 2020
- 15. Heydari, A., Tavakoli, M.A., Salim, N., & Heydari, Z., 2015. Detection of review spam: A survey. Expert Systems with Applications, 42(7), pp.3634–42.