

Online Fake Review Detection based on Multiple Feature using Machine Learning Technique

Milind Mahajan¹, Vaishnavi Kulkarni², Aniket Nimbalkar³, Abhishek Thombare⁴ and Saurabh Lagad⁵

¹⁻⁵Information Technology Department, Pune Vidyarthi Griha's College of Engineering and Technology & G. K. Pate (Wani), Institute of Management, Pune-411009, India

Email: milind6787@gmail.com, vaishnavirk0411@gmail.com, aniketnimbalkar270320@gmail.com, abhisheka.thombare07@gmail.com, saurabhlagad435@gmail.com

Abstract—In today's online era while purchasing things such as clothes, laptops, home appliances, etc., customers rely on other people's reviews to know more about the product which they ought to buy. These reviews can be genuinely posted by previous purchasers or be totally deceptive i.e. spammed by bots. Therefore, the detection of fake reviews online has now become the pinnacle point of interest. Despite recent studies, the accuracy of the existing models is not up to the mark. The proposed method discusses the online fake review detection consisting of data acquisition, pre-processing, train-test splitting of data, and ensemble learning which detects these fake reviews and increases the accuracy to a higher mark. Our model uses three supervised machine learning algorithms viz. Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and trains on electronic items dataset. After training on three individual algorithms and post-calculation of the accuracy of each algorithm, the model compares and analyzes the outcomes and selects the algorithm with the highest performance accuracy.

Index Terms— Electronic Dataset, Support Vector Machine, Artificial Neural Network, Random Forest, Fake Reviews Detection, Supervised Machine Learning.

I. INTRODUCTION

Nowadays, customers can purchase various items online and also freely express their views and opinion about the purchased product in the form of reviews on the eCommerce websites as well as social media. These online reviews help future customers to decide if the product is worth purchasing or not. The reviews are generally genuine and submitted by previous customers who have used the product. But among all the existing reviews there are some reviews that are not written by customers: fake reviews.

Fake reviews are referred to as deceptive reviews which are written and spread by malicious users/bots. Unfortunately, the competition in the online market has become so cut-throat that some large enterprises are now relying on these fake reviews to sell their products or make the competitors lose out on customers. Recent studies have shown that almost 4% of all online reviews are fake [1]. Translating this into economic impact, the direct influence of fake online reviews on global online spending is \$152 billion [1]. Fake review detection is thus a significant problem for eCommerce websites as the customers rely highly on such opinions.

Filtering out fake reviews helps maintain trust among customers. Therefore, to retain the customer base it has become necessary for E-commerce businesses to identify and filter invalid reviews, protect small businesses, be proactive and prioritize action against them. To a large extent, these reviews follow certain patterns and can be detected by machine learning/deep learning algorithms. After classification, they can be removed by respective websites.

The remainder of the paper is organized as follows: In section II the related work is overviewed. Section III describes the proposed methodology. Conclusions that can be drawn are covered in section IV.

II. LITERATURE SURVEY

A. Methods

Fake reviews can be detected by supervised machine algorithms or unsupervised machine learning algorithms. But according to [2][3][4] supervised machine learning algorithms are used predominantly to identify and filter out fake reviews. Some researchers use unsupervised learning to remove fake reviews. But due to the unlabeled dataset, it is difficult to remove fake reviews in a larger dataset by using deep learning/unsupervised learning. To increase the accuracy and make stable predictions ensemble models are used by many researchers.

B. Approaches

According to [5] feature framework for detecting fake reviews that have been evaluated in the consumer electronics domain. The steps followed for detecting fake reviews in the electronic domain are (i) Acquisition of datasets (ii) feature framework definition for fake review (iii) fake review classification (iv) evaluation and analysis of the result. This framework is composed of two types of features. The features related to text review are Review centric and features that show the behavior of a user on the site are Centric features.

The survey [6] developed a supervised machine learning approach to classify the online fake reviews and filter out those reviews by using the datasets that contain hotel reviews. Also merged some features from the other research works for the development of a feature set that can perform better classification.

The paper [7] compares the reviews and performance of several experiments done on the Yelp dataset of restaurants by using different classifiers like KNN, Naïve Bayes (NB), Support Vector Machine (SVM), Logistic Regression, and Random forest. In this paper, the Yelp dataset is used to evaluate the proposed approach.

In the research paper [8], detection of fake reviews from the restaurant online reviews. This detection of fake reviews from the restaurant dataset is done by using a hybrid approach and by following data pre-processing, detection, and ensemble learning.

III. PROPOSED METHODOLOGY

Among various classification algorithms, supervised machine learning algorithms are still the most frequently used methods for fake review detection. Although deep learning methods have been proven to be effective, they require largely sufficient training data, so they might not be suitable for fake review detection due to the difficulty of labeling data manually. Given such conditions, we have chosen three machine learning algorithms as the base classifiers: Random Forest (RF), SVM, ANN. The reason for choosing ensemble modeling is that a single classifier cannot perform best for all datasets provided.

A. Classifier Ensemble Approach:

Ensemble learning methods are popular and widely used as they reduce bias and achieve better performance. An Ensemble strategy combines the prediction of two or more other models. An ensemble can make better predictions than any single contributing model. An ensemble reduces the spread or dispersion of the predictions and model performance. A classifier is built based on the identified features. And those features are assigned to the algorithms depending on the classified training sets. In this section, we briefly discuss the several proposed algorithms on fake review detection, including classification, stop word removal, and stemming.

Support Vector Machine

Support vector machine is a supervised machine learning algorithm that can be used for both types of problems i.e. regression and classification. SVM is used to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The distance between a data point and hyperplane is margin and a decision plane that separates between a set of data points having different classes is called a hyperplane. Support vectors are the

hyperplane used to maximize the margin between two classes. A hyperplane having the maximum distance to the nearest training data item of both the classes is considered to be the most appropriate hyperplane.

SVM is performed with the help of kernels. Kernels are very helpful in classification and are used to analyze the patterns in a given dataset. For the fake review detection, the linear kernel will give better accuracy as it falls in the category of text classification problems. The 2nd most commonly used kernel is the gaussian Radial Basis Function (RBF). It has two hyper-parameters cost (C) and gamma that decide the performance of the SVM model.

Random Forest

Another algorithm that is used for filtering out fake reviews is Random Forest (RF) which is also a supervised machine learning algorithm used for classification and regression. Selected features are given as input to the RF algorithm and it creates several probable classified decision trees with weight and label. It is a collection of multiple decision trees and is also less sensitive to the training data. The main reason to use the Random Forest algorithm over decision trees is that decision trees are highly sensitive to the training data which could result in high variance.

Two techniques are present in Random Forest, bootstrapping and random feature selection. The idea behind bootstrapping is to create several subsets of data from a training set chosen randomly with replacement. On each bootstrapped dataset a decision tree is trained independently. Subsets of features are randomly chosen and used for training. The process in which the results are combined from multiple models is called aggregation. The bootstrapping and aggregation together are called bagging. Bootstrapping ensures that the same data is not used for every tree so in a way it helps our model to be less sensitive to the original training dataset. The random feature selection helps to reduce the correlation between trees.

Artificial Neural Network

An Artificial Neural Network is simply called a neural network that mimics the behavior of the human brain to solve complex data-driven problems. ANN consists of 3 layers viz. input layer, hidden layer, and output layer as shown in Fig1. The feature vectors will be provided as an input to the input layer. The hidden layer connects both the input layer and output layer that consists of neurons and also in this layer all calculations are performed to find the hidden features and patterns. And the output layer will give a neural network response.

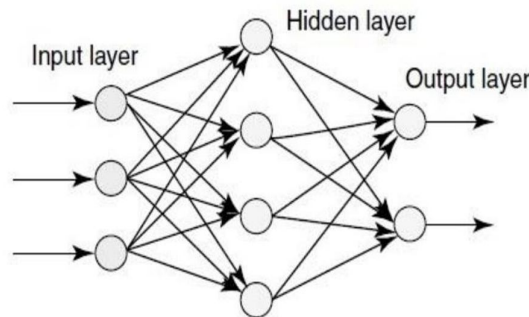


Fig. 1. Multilayered artificial neuron network [9]

From Fig.1, the arrows drawn from the input to the hidden layer as well as from the hidden to the output layer indicate the pathway for the flow of information. Each arrow represents a connection between two neurons. The connections have the weight, an integer value that controls the signal between two neurons. If the model is giving poor/undesired output then the model alters the weights to improve the desired output. However, if the model is generating good/desired output then there is no need to adjust the weights between connections.

B. Proposed Architecture

First of all, data acquisition will take place i.e. the reviews submitted by the users to the website will get stored in the data set. Then that data set will go through pre-processing which consists of lexical analysis, stop word removal, and stemming. In lexical analysis, the review gets separated into tokens concerning delimiters such as space, newline, tab. In the stop word removal process, the words that occur commonly and frequently like a, an, the, of, etc. will be ignored. Now, in the Stemming process, reducing a word to its word stem takes place. In

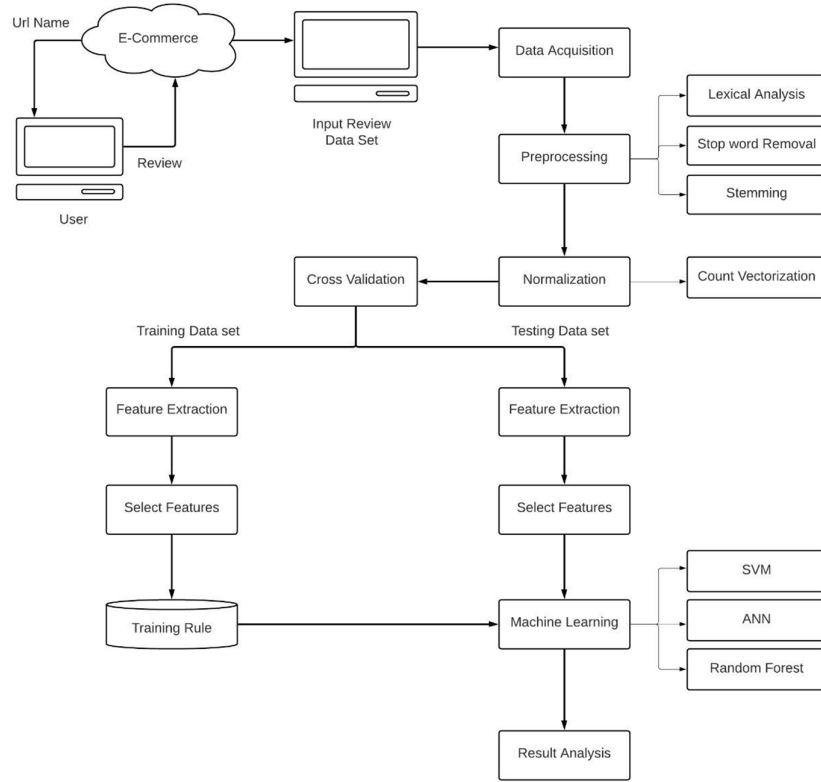


Fig. 2. Architecture Diagram of Online Fake Review Detection System

short, the system will replace all the variants with a single stem word. These variants include plurals, gerund forms i.e. ing form, third-person suffixes like 's', 'es', and past tense suffixes like 'ed'.

The next implementation step is normalization where a count vectorizer will be applied on pre-processed data followed by cross-validation. Subsequently, the data is split into training and testing data sets. And afterward, feature extraction, the process in which the numbers of features are reduced by creating new features from existing features, is used on both training as well as testing data sets. These extracted features have both positive and negative values. With the help of a hybrid approach and by using supervised machine learning algorithms viz. Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest (RF), all the reviews will be classified as genuine or deceptive. Lastly, we analyze the accuracy of the proposed system and compare the highest acquired accuracy with other existing systems.

IV. CONCLUSION

The major drawback of today's existing systems is the low accuracy obtained by the models. To increase the accuracy of the classification of fake reviews, we proposed a sentiment classification approach based on product aspects. The modules for this include Preprocessing Module, Product Aspect Identification, Sentiment Classification. The proposed system describes a personalized-based fake reviews detection from social media. In the initial research, the system evaluates the system performance with this dataset and measure the accuracy. By using this proposed system, a more accurate result can be generated which will directly help make the process easier and less straining for the eCommerce websites. Also, customers can trust the authenticity of the reviews on that particular website without any hesitation.

REFERENCES

- [1] <https://theprint.in/opinion/almost-4-of-all-online-reviews-are-fake-their-impact-is-costing-us-152-billion/715689/>

- [2] Martinez-Torres, M., Toral, S.: "A machine learning approach for the identification of the deceptive reviews in the hospitality sector using unique attributes and sentiment orientation," *Tourism Manage.*, vol. 75, pp. 393–403, Dec. 2019.
- [3] Cardoso, E., Silva, R., Almeida, T., "Towards automatic filtering of fake reviews," *Neurocomputing*, vol. 309, pp. 106–116, Oct. 2018.
- [4] Ott, M., Cardie, C., Hancock, J., "Negative deceptive opinion spam," in *Proc. Conf. North Amer. Chapter Assoc. Compute. Linguistics, Hum. Lang. Technologies (HLT- NAACL)*, Atlanta, GA, USA, Jun. 2013, pp. 497–501.
- [5] Barbado, Rodrigo, Araque, O., Iglesias, C.: "A framework for fake review detection in online consumer electronics retailers." *Information Processing & Management* 56.4 (2019): 1234-1244.
- [6] Hassan, Rakibul, Islam, Mr.: "A Supervised Machine Learning Approach to Detect Fake Online Reviews." *2020 23rd International Conference on Computer and Information Technology (ICCIT)*. IEEE, (2020).
- [7] Elmogy, Ahmed M., et al. "Fake Reviews Detection using Supervised Machine Learning." *Int. J. Adv. Comput. Sci. Appl* 12 (2021).
- [8] Hegde, S., Rai, R., Hiremath S., Gangisetty, S.: "Fake Review Detection Using Hybrid Ensemble Learning". In: *Advances in Computing and Network Communications*, Springer (2021).
- [9] Sharma, V., Rai, S., Dev, A.: "A comprehensive study of artificial neural networks" (2012). (ISSN: 2277 128X).