IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 3, March 2021

DOI 10.17148/IJARCCE.2021.10323

DETECTING FAKE ONLINE REVIEWS USING SUPERVISIED LEARNING

Mr. M. Ravikumar[1],Aparna R[2], Jinu T Benu[3], Jindo K Joy[4], Sandra P[5]

Department of Computer Science and Engineering,
JCT College of Engineering and Technology Coimbatore, Tamil nadu, India[1-5]

Abstract: Online reviews have great impact on today's business and commerce. Decision making for purchase of online products mostly depends on reviews given by the users. Hence, opportunistic individuals or groups try to manipulate product reviews for their own interests. This paper introduces supervised text mining models to detect fake online reviews as well as compares the efficiency of both techniques on dataset containing hotel reviews.

Keywords: Supervised learning, random forest algorithm

INTRODUCTION:

Technologies are changing rapidly. Old technologies are continuously being replaced by new and sophisticated ones.These new technologies are enabling people to have their work done efficiently. Such an evolution of technology is online marketplace. We can shop and make reservation using online websites. Almost, every one of us checks out reviews before purchasing some products or services. Hence, online reviews have become a great source of reputation for the companies. Also, they have large impact on advertisement and promotion of products and services. With the spread of online marketplace, fake online reviews are becoming great matter of concern. People can make false reviews for promotion of their own products that harms the actual users.Also, competitive companies can try to damage each others reputation by providing fake negative reviews.Researchers have been studying about many approaches for detection of these fake online reviews. Some approaches are review content based and some are based on behaviour of the user who is posting reviews. Content based study focuses on what is written on the review that is the text of the review where user behavior based method focuses on country, ip-address, number of posts of the reviewer etc. Most of the proposed approaches are supervised classification models.

Few researchers also have worked with semi-supervised models. Semi-supervised methods are being introduced for lack of reliable labelling of the reviews.

In this paper, we make some classification approaches for detecting fake online reviews, supervised learning. For supervised learning, we use Expectation-maximization algorithm. Random forest classifier used as classifiers in our research work to improve the performance of classification. We have mainly focused on the content of the review based approaches. As feature we use to detecting the hotel reviews is fake or not using machine learning algorithm.

In the following section II, we discuss about the related works. Section III describes our proposed approaches and experiment setup. Results and findings of our research are discussed in Section IV.

RELATED WORK:

Many approaches and techniques have been proposed in the field of fake review detection.The following methods have been able to detect fake online review with higher accuracy.

Sun et al. [1] divided these approaches into two categories.

a) Content Based Method: Content based methods focus on what is the content of the review. That is the text of the review or what is told in it. Heydari et al. [2] have attempted to detect spam review by analyzing the linguistic features of the review. Ott et al. [3] used three techniques to perform classification. These three techniques are- genre identificationdetection of psycholinguistic deception and text categorization [1]–[3].

1) Genre Identification: The parts-of-speech (POS) distribution of the review are explored by Ott et al. [3].They used frequency count of POS tags as the features representing the review for classification.

2) Detection of Psycholinguistic Deception: The psycholinguistic method approaches to assign psycholinguistic meanings to the important features of a review.Linguistic Inquiry and Word Count (LIWC) software was used by Pennebaker et al. [4] to build their features for the reviews.

3) Text Categorization: Ott et al. experimented n-gram that is now popularly used as an important feature in review detection.

Other linguistic features are also explored. Such as, Feng etal. [5] took lexicalized and unlexicalized syntactic features by constructing sentence parse trees for fake review detection. They shown experimenally that the deep syntactic features improve the accuracy of prediction. Li et al. [6] explored a variety of generic deceptive signals which contribute to the

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 3, March 2021

DOI 10.17148/IJARCCE.2021.10323

fake review detection. They also concluded that combined general features such as LIWC or POS with bag of words will be more robust than bag of words alone. Metadata about reviews such as reviews length, date, time and rating are also used as features by some researchers.

b) Behaviour Feature Based Methods: Behaviour feature based study focuses on the reviewer that includes characteristics of the person who is giving the review. Lim et al. [7] addressed the problem of review spammer detection, or finding users who are the source of spam reviews. People who post intentional fake reviews have significantly different behaviour than the normal user. They have identified the following deceptive rating and review behaviours.

_ Giving unfair rating too often: Professional spammers generally posts more fake reviews than the real ones.Suppose a product has average rating of 9.0 out of 10.But a reviewer has given 4.0 rating. Analyzing the other reviews of the reviewer if we find out that he often gives this type of unfair ratings than we can detect him as a spammer.

_ Giving good rating to own country's product: Sometimes people post fake reviews to promote products of own region. This type of spamming is mostly seen in case of movie reviews. Suppose, in an international movie website an Indian movie have the rating of 9.0 out of 10.0,where most of the reviewers are Indian. This kinds of spamming can be detected using address of the reviewers.

_ Giving review on a vast variety of product: Each

person has specific interests of his own. A person generally is not interested in all types of products. Suppose a person who loves gaming may not be interested in classic literature. But if we find some people giving reviews in various types of products which exceeds the general behaviour then we can intuit that their reviews are intentional fake reviews.

Deceptive online review detection is generally considered as a classification problem and one popular approach is to use supervised text classification techniques. These techniques are robust if the training is performed using large datasets of labelled instances from both classes, deceptive opinions (positive instances) and truthful opinions (negative examples).Some researchers also used semi-supervised classification techniques.

For supervised classification process ground truth is determined by – helpfulness vote, rating based behaviours, using seed words, human observation etc. Sun et al. [1] proposed a method that offers classification results through a bagging model which bags three classifiers including product word composition classifier (PWCC), TRIGRAMSSVM classifier, and BIGRAMSSVM classifier. They introduced a product word composition classifier to predict the polarity of the review. The model was used to map the words of a review into the continuous representation while concurrently integrating the product-review relations. To build the document model, they took the product word composition vectors as input and used Convolutional Neural Network CNN to build the representation model. After bagging the result with TRIGRAMSSVM classification, and BIGRAMSSVM classification they got F-Score value 0:77.However supervised method has some challenges to overcome.The following problems occur in case of supervised techniques.

_ Assuring of the quality of the reviews is difficult.

_ Labelled data points to train the classifier is difficult to obtain.

_ Human are poor in labelling reviews as fake or genuine.

Hence Jitendra et al. [8] proposed semi-supervised method where labelled and unlabelled data both are trained together. They proposed to use semi-supervised method in the following situations.

1) When reliable data is not available.

2) Dynamic nature of online review.

3) Designing heuristic rules are difficult.

They proposed several semi-supervised learning techniqueswhich includes Co-training, Expectation maximization, Label Propagation and Spreading and Positive Unlabelled Learning [8]. They used several classifiers which includes k-Nearest neighbor, Random Forest, Logistic Regression and Stochastic Gradient Descent. Using semi-supervised techniques they achieved highest accurace of 84%.

PROPOSED WORK:

In this paper, we make some classification approaches for detecting fake online reviews, using supervised learning. For supervised learning, Random forests classifier and used as classifiers in our research work to improve the performance of classification. We have mainly focused on the content of the review based approaches. It will give accurate results. By this method we get if the labelled reviews is fake or genuine.

PROPOSED METHODOLOGY:

1. Random Forest Algorithm

Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithm of the same type i.e. multiple decision trees, resulting in a forest of trees,hence the name "Random Forest". The random forest algorithm can be used for both

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 3, March 2021

DOI 10.17148/IJARCCE.2021.10323

regression and classification tasks.

The following are the basic steps involved in performing the random forest algorithm:
1. Pick N random records from the dataset.
2. Build a decision tree based on these N records.
3. Choose the number of trees you want in your algorithm and repeat steps 1 and 2.
4. In case of a regression problem, for a new record, each tree in the forest predicts a value for Y (output). The final value can be calculated by taking the average of all the values predicted by all the trees in forest. Or, in case of a classification problem, each tree in the forest predicts the category to which the new record belongs. Finally, the new record is assigned to the category that wins the majority vote.

2. Methodology:

Collect the dataset:
The dataset of 800 positive reviews and 800 deceptive reviews are collected in the form of text files.Reviews of 20 Chicago hotels are collected for process.Extracting features from dataset using text mining methods

Training and Testing of dataset:

Which are the basic machine learning steps to be achieved. After collecting the dataset 80% of data is taken for process and 20% of data is taken for testing. Both of the steps should be performed after the feature extraction (Abov features).Specific supervised algorithms are used for training process.Here have used randomforest algorithm.In training process algorithm needs two inputs Features and labels of each data.After specifying the parameters of random forest Training process is performed. Testing process is mainly used for of checking the accuracy and precision of the training process.Which also uses the random forest algorithm. We check the 20% of the review with the trained model using predict function in the algorithm. If specific accuracy is not achieved,we have to retune the randomforest.

Create the model:

For prediction process we have to save the model in specific format. We have used Pickle format (.pkl) for saving the model.

To fit the model we have used sklearn of Python programming language provides the needful libraries for the classifiers. We have different classification techniques in machine learning like random forest, Decision Tree and Random Forest classifiers. We have applied different predictions methods to reach the more accurate model.

Random Forest algorithm is an Ensemble model which creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting .algorithm creates decision trees on data samples and then gets the prediction from each tree and finally selects the best solution from that. This produces the highest accuracy.
Random Forest can also use for classification as well as Regression analysis. Random forest is popularly used for text categorization to predict the text with word frequencies as the features. It is typically uses bag-of-words feature from NLP to identify the fake in text categorization. SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. For SVM classifier we have gamma parameter keeping constant for perfect fit model.

RESULTS AND PERFOMANCE:

ANALYSIS

A. Experimental Environment

We have applied our experiments on a machine with Processor: Intel core i3 – 2330M and CPU- 2GHz, RAM: 4GB,S system with 64 bit OS, We have used Windows as an operating system. We have used Python as programming language with sklearn, numpy and pandas packages. Spyder 4.1.1 used as IDE.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 3, March 2021

DOI 10.17148/IJARCCE.2021.10323

B. Results

We have used random forest classifier, Decision Tree and Random Forest classifiers to classify the reviews dataset. We have divided the dataset of 1600 rows with 3 columnswith column names reviews, polarity and spamity for each classification process.

REFERENCES:

[1] Chengai Sun, Qiaolin Du and Gang Tian, "Exploiting Product Related Review Features for Fake Review Detection," Mathematical Problems in Engineering, 2016.
[2] A. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, "Detection of review spam: a survey", Expert Systems with Applications, vol. 42, no.7, pp. 3634–3642, 2015.
[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT), vol. 1, pp. 309–319, Association for Computational Linguistics, Portland, Ore, USA, June 2011.
[4] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic Inquiry and Word Count: Liwc," vol. 71, 2001.
[5] S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, Vol. 2, 2012.
[6] 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 (ACL), 2014.
[7] E. P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, "Detecting product review spammers using rating behaviors," in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), 2010.
[8] J. K. Rout, A. Dalmia, and K.-K. R. Choo, "Revisiting semi-supervised learning for online deceptive review detection," IEEE Access, Vol. 5, pp. 1319–1327, 2017.
[9] J. Karimpour, A. A. Noroozi, and S. Alizadeh, "Web spam detection by learning from small labeled samples," International Journal of Computer Applications, vol. 50, no. 21, pp. 1–5, July 2012