



Detective System of Fake Online Reviews using Semi-Supervised and Supervised Learning

M.Bhavya Chamundeswari¹ | B.N.S Gupta²

¹PG Scholar, Department of Computer science, SVKP &Dr K S Raju Arts & Science College(A), Penugonda, W.G.Dt., A.P, India

²Associate Professor in Computer science, SVKP &Dr K S Raju Arts & Science College(A), Penugonda, W.G.Dt., A.P, India

To Cite this Article

M.Bhavya Chamundeswari and B.N.S Gupta. Detective System of Fake Online Reviews using Semi-Supervised and Supervised Learning. International Journal for Modern Trends in Science and Technology 2022, 8(09), pp. 131-133.

<https://doi.org/10.46501/IJMTST0809026>

Article Info

Received: 22 August 2022; Accepted: 12 September 2022; Published: 17 September 2022.

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 some semi-supervised and supervised text mining models to detect fake online reviews as well as compares the efficiency of both techniques on dataset containing hotel reviews.

KEYWORDS : *Decision Making, online reviews, online products, great impact, reviews, online, well, users, try, text, techniques, supervised*

1. INTRODUCTION

Technology is evolving quickly. New, advanced technology are constantly replacing outdated ones. With the aid of these modern technology, people may do their task quickly. Online markets are a result of this technological advancement. Using internet, we may purchase and make reservations. Nearly all of us read reviews before deciding to buy a product or service. Online reviews are now a valuable source of reputation for businesses. Additionally, they play a significant role in product and service promotion and advertising. Fake internet reviews are a growing concern as online marketplaces grow in popularity. For the purpose of promoting their own items, someone may write phoney evaluations that are harmful to the real users.

Additionally, rival businesses may attempt to harm one another's reputations by posting phoney bad evaluations.

Numerous methods for identifying these fraudulent online reviews have been studied by researchers. Some strategies are focused on the reviews' content, while others are based on how the users who write the reviews behave. The text of the review is the emphasis of the content-based study, whereas the nation, IP address, and number of postings of the reviewer are the focus of the user behavior-based method. The majority of the methods that are suggested use supervised classification models. Only a few number of researchers have also used semi-supervised models. Because the reviews

cannot be labelled with sufficient accuracy, semi-supervised methods are being presented.

In this article, we provide a variety of classification methods for identifying false internet reviews, some of which are semisupervised and others are supervised. We employ the Expectation-maximization technique for semi-supervised learning. In this research, we use Support Vector Machines (SVM) and Statistical Naive Bayes classifiers as classifiers to enhance classification performance. We have mainly concentrated on the review-based techniques' content. Word frequency, emotion polarity, and review length were employed as features.

2. LITERATURE SURVEY

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 behavior 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 labeling of the reviews.

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 identification, detection 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 fake review detection.

Other linguistic features are also explored. Such as, Feng et al. [5] took lexicalized and unlexicalized syntactic features by constructing sentence parse trees for fake review detection. They shown experimentally 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 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.

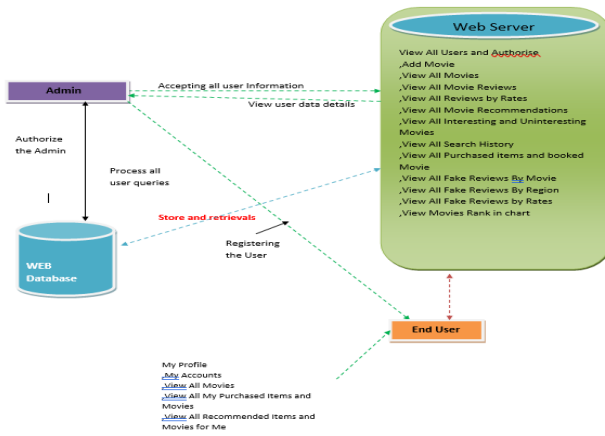
3. PROBLEM STATEMENT

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 identification, detection of psycholinguistic deception and text categorization.

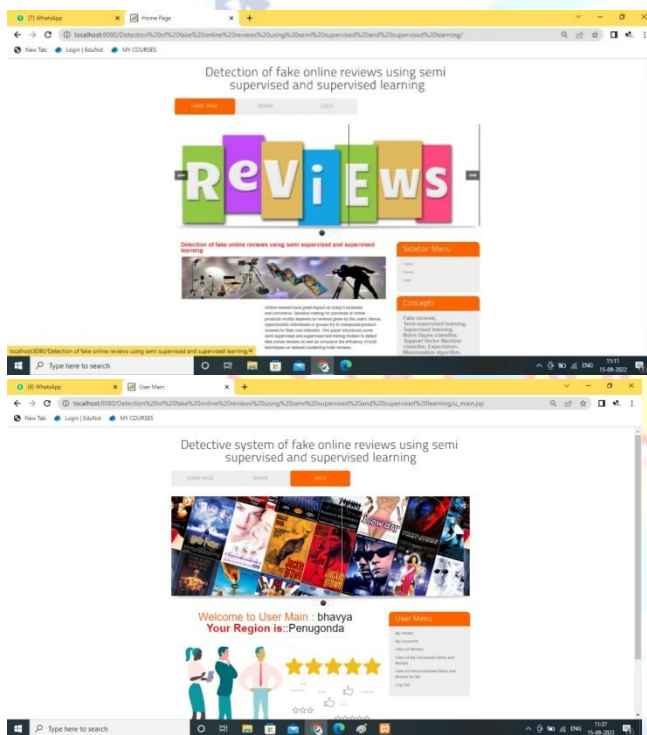
Behavior feature based study focuses on the reviewer that includes characteristics of the person who is giving the review. Lim et al. [7] addressed the problem of review spammer detection, or finding users who are the source of spam reviews. People who post intentional fake reviews have significantly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

Deceptive online review detection is generally considered as a classification problem and one popular approach is to use supervised text classification techniques [5]. These techniques are robust if the training is performed using large datasets of labeled instances from both classes, deceptive opinions (positive instances) and truthful opinions (negative examples) [8]. Some researchers also used semi-supervised classification techniques.

4. ARCHITECTURE



5. RESULTS



6. CONCLUSION

In this study, we have demonstrated a number of semi-supervised and supervised text mining approaches for identifying false internet reviews. To improve the feature set, we incorporated features from other research studies. Additionally, we attempted a few different classifiers that weren't used in the earlier study. As a result, Jiten et al.earlier .'s semisupervised approaches have improved in accuracy [8]. Additionally, we discovered that the supervised Naive Bayes classifier provides the best accuracy. This guarantees that our dataset is accurately labelled, as we are aware that semi-supervised models perform well in the absence of accurate labelling.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

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

Authors Biography



M. Bhavya Chamundeswari currently pursuing MCA in SVKP & Dr KS Raju Arts and Science College(A), Affiliated to Adikavi Nannaya Univesity ,Rajahendravaram. Her research interests include Machine learning ,Data Mining



B.N. Srinivasa Gupta is working as Associate Professor in SVKP & Dr K S Raju Arts & Science College, Penugonda, A.P. He received Masters Degree in Computer Applications from Andhra University and Computer Science & Engineering from Jawaharlal Nehru Technological University Kakinada (JNTUK), Kakinada, India. His research interests include Data Mining, Cyber Security, Artificial Intelligence