

RESEARCH ARTICLE | OCTOBER 04 2023

Recognition of fabricated online reviews using semi-supervised learning

B. Satyanarayana ; Dinesh Chandra Pancharia

 Check for updates

AIP Conf. Proc. 2794, 020008 (2023)

<https://doi.org/10.1063/5.0165688>


View
Online


Export
Citation

CrossMark

Articles You May Be Interested In

A functional integral formalism for quantum spin systems

J. Math. Phys. (July 2008)

Modes selection in polymer mixtures undergoing phase separation by photochemical reactions

Chaos (June 1999)

Spreading of a surfactant monolayer on a thin liquid film: Onset and evolution of digitated structures

Chaos (March 1999)

500 kHz or 8.5 GHz?
And all the ranges in between.

Lock-in Amplifiers for your periodic signal measurements



Find out more

 Zurich
Instruments

Recognition of Fabricated Online Reviews Using Semi-Supervised Learning

B Satyanarayana^{1,a)}, Dinesh Chandra Pancharia^{1,b)}

¹Dept. of Computer Science & Engineering, AVANTHI Institute of Engineering & Technology, Vizianagaram, A.P, India.

^aCorresponding author: satyanarayana.botsa@gmail.com, ^bdinesh.pancharia@gmail.com

Abstract— In the World Online reviews have major impact on daily business and e-commerce. Purchase of products in online mostly depends on reviews given by the users, reviews are the major parameter of decision making. Thus, opportunistic individuals or groups attempt to use product reviews to advance their own interests. We propose some semi-supervised text mining models to detect false online reviews in this paper.

Index Terms: Fake reviews, semi supervised learning, supervised learning, Naïve Bayes classifier, Support Vector Machine classifier, Expectation-maximization algorithm.

I. INTRODUCTION

Sophisticated and new technologies continually replace the old ones. These the new technologies are enabling people to have their work done efficiently. Such an evolution of technology is the online marketplace. We can shop and make a reservation using online websites. Almost, before purchasing a product or service, we all check reviews. so the online reviews have become a great source of a reputation for companies. Also, they have a large impact on the advertisement and promotion of products and services. With the spread of the online marketplace, fake online reviews are becoming a great matter of concern. People can make false reviews for the promotion of their products that harms the actual users. some of , competitive companies can try to damage each other's reputation by providing fake negative reviews. Researchers have been studying many approaches for the detection of these fake online reviews. Some approaches are review content-based and some are based on the behavior of the user who is posting reviews. The content-based study focuses on what is written on the review that is the text of the review whereas the user behavior-based method focuses on country, IP address, the number of posts of the reviewer. There are a large number of models based on supervised classification. There is a need for semi-supervised methods because reviews cannot be reliably labelled.

In this paper, we use classification approaches for detecting fabricated online reviews, some of which are semi Supervised. For semi-supervised learning, we use Expectation-maximization algorithm. Naïve Bayes classifier and Support Vector Machines (SVM) are used as classifiers in our research work, to get best performance of classification. Review-based approaches have mainly been examined in terms of content. As feature we have used word frequency count, sentiment polarity and length of review.

In the following section-II, we discuss about the related works. Section-III describes our proposed approaches and experiment setup. Results and finding so four research are discussed in Section IV. Section-V includes conclusions and future work.

II. RELATEDWORK

In the field of fabricated review detection, a number of approaches and techniques have been proposed. The following methods have been able to Recognition fabricated online review with higher accuracy.

Content Based Method: Methods that focus on content examine what is in the review. These three techniques are 1. genre identification 2. detection of psycholinguistic deception 3. text-categorization

1)Genre Identification: It is explored in [3] how the review's parts-of-speech distribution is distributed. In order to classify reviews, the features used were the frequency count of POS tags.

2)Detection of Psycholinguistic Deception: Assigning psycholinguistic meaning to the important feature so far is the goal of the psycholinguistic method. Linguistic Inquiry and Word Count (LIWC)software was used by [4]. To build their features for the reviews.

3)Text Categorization: [3] A popular feature of fake review detection is the n-gram, which was initially a research project. Other linguistic features are also explored. Such as [5]. Took lexicalized and un-lexicalized syntactic features by Constructing sentence parse trees for fabricated review detection. They shown experimentally that the deep syntactic features improve the accuracy of prediction. explored a different of