Abstract

In today’s world, where Internet has become a household convenience, online reviews have become a critical tool for businesses to control their online reputation. Reviewing has changed the face of marketing in this new era. Nowadays, most companies invest money in mining the reviews to gain insights into customer preferences as well as to gain competitive intelligence and are hiring individuals to write fake reviews. The fraudsters' activities mislead potential customers and organizations reshaping their businesses and prevent opinion-mining techniques from reaching accurate conclusions. Thus, it has become essential to detect fake reviews to bring to surface the true product opinion. This paper focuses on product reviews and detecting spam fake reviews among them using supervised learning techniques using synthetic fake reviews (to cover all types) as a training set. Term frequency and user review frequency are two features whose impact on classification model is studied in this paper. It classifies the reviews to test the accuracy of the model. The results have been encouraging with an accuracy of over 98%.
References


Index Terms

Computer Science

Information Sciences

Keywords

Review spam, Opinion mining, fake reviews, Naïve Bayes classification, Opinion Spamming, Random Forest Classifier, Classification Model Evaluation Measures