A Sentiment Analysis Approach using Effective Feature Reduction Method

International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

Volume 170
Number 4

Year of Publication: 2017

Authors:
Arash Mazidi, Elham Damghanijazi

Abstract

Sentiment analysis has attracted researchers in recent years which aims to present an automatic method for analyzing comments, assessments, opinions and sentiments of a text. In this paper, Ngram feature vector and POS (Part of Speech) are extracted from text and it is tried to find a proper combination of feature vectors so that texts can be classified as positive and negative opinions. In order to choose the most useful features, information gain ratio is used, then machine learning algorithms are used to investigate the effect of different features on sentiment analysis. In this paper, 4 groups of data are studied including film evaluation, products' evaluation (including book, DVD and electronics). Classification results are studied for three types of feature vectors: Ngram feature vector, POS feature vector and combination feature vector of Ngram with POS. results show that combinational feature vector performs better in sentiment analysis. By combining features, Boolean Multinominal Naïve Bayes (BMNB) results are improved compared to support vector machine classification algorithm.

References


Index Terms
Computer Science
Artificial Intelligence

Keywords
Sentiment analysis, Feature selection method, machine learning, support vector machine, information gain, information gain ratio.