Abstract

Sentiment analysis, also known as opinion mining, is the analysis of the feelings (i.e., attitudes, emotions and opinions) behind the words. Sentiment analysis involves classifying the opinions as positive, negative, or neutral. Classification of textual objects in accordance with sentiment is considered to be a more difficult task than classification of textual objects in accordance with the content because opinions in natural language can be expressed in subtle and complex ways containing slang, ambiguity, sarcasm, irony, and idiom. This paper investigates the problem of sentiment analysis of online review. A Jaccard index based clustering algorithm (JIBCA) is proposed to support mining online reviews and predicting sales performance. The information gain is the change in information by considering number of datasets. The performance of information gain varies depending on the dataset. It is observed that the information gain performed better in JIBCA than existing methods for the movie review dataset. It is therefore recommended that JIBCA can be a good feature selection method for sentiment classification tasks. This paper also proposes a new approach for movie reviews classification based on extraction and analysis of appraisal groups such as action, thrill, comedy, and romantic.
- Xiaohui Yu, Yang Liu, Xiangji Huang, Aijun, "Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain", A Knowledge and Data Engineering, IEEE Transactions on 24, 2012.
- Deng Bin, Peiji, Shao, Zhao Dan, "E-Commerce Reviews Management System Based on Online Customer Reviews Mining Innovative Computing & Communication", 2010 Intl Conf on and Information Technology & Ocean Engineering, 2010 Asia-Pacific Conf on (CICC-ITOEO) 2010
- Peng Jiang, Chunxia Zhang, Hongping Fu, Zhen dong Niu, Qing Yang, "An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews, Data Mining (ICDM), 2010 IEEE 10th International Conference on 2010
- Lai C. L., Xu K. Q., Lau R. Y. K. Yue feng Li, Dawei Song, High-Order Concept Associations Mining and Inferential Language Modeling for Online Review Spam Detection, Data Mining Workshops (ICDMW), 2010 IEEE International Conference, on 2010
- Weishu Hu, Zhiguo Gong, Jingzhi Guoe, "Mining Product Features from Online Reviews", Business Engineering (ICEBE), 2010 IEEE 7th International Conference, on 2010
- Samsudin N., Puteh M., Hamdan A. R., Bess or xbest: Mining the Malaysian online reviews&quo, Data Mining and Optimization (DMO), 2011 3rd Conference, on 2011
- XuXueke, Cheng Xueqi, Tan Songbo, Liu Yue, Shen Huawei, Aspect-level opinion mining of online customer reviews Communications, China102013
- Lin E., Shiaofen Fang, Jie Wang, "Mining Online Book Reviews for Sentimental Clustering", Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on 2013
- Anwer, N. Rashid, A. Hassan, Feature based opinion mining of online free format customer reviews using frequency distribution and Bayesian statistics; Networked Computing and Advanced Information Management (NCM), 2010 Sixth International Conference on 2010
- Algur S. P., Patil A. P., Hiremath P. S., Shivashankar, Conceptual level similarity measure based review spam detection; Signal and Image Processing (ICSIP), 2010 International Conference on 2010
- Rahayu, D. A., Krishnaswamy S., Alahakoon O., Labbe C., RnR: Extracting Rationale from Online Reviews and Ratings; Data Mining Workshops (ICDMW), 2010
IEEE International Conference on 2010
  - Minqing Hu, Bing Liu, "Mining and summarizing customer reviews", Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004

**Index Terms**

Computer Science  
Algorithms

**Keywords**

Sentiment Analysis  
Review Classification  
Opinion Mining  
Review Mining  
Prediction.