Title : **A review of natural language processing techniques for opinion mining systems**
Writer : **Shiliang Sun, Chen Luo, Junyu Chen**
Citations : **166**
Date Submitted : **2016**
Date Published : **2017**
Journal : **Information Fusion 36 (2017)**
Source :**Department of Computer Science and Technology, East China Normal University**

Mohammad Sabik Irbaz (160041004)

Starting Date : January 28, 2020

Overall Summary :

First, they introduced general NLP techniques which are required for text preprocessing. Second, they investigated the approaches of opinion mining for different levels and situations. Then they introduce comparative opinion mining and deep learning approaches for opinion mining. Opinion summarization and advanced topics are introduced later. Finally, they discuss some challenges and open problems related to opinion mining.
Document and sentence level opinion mining can be formulated as classification problems which determine whether a positive or negative sentiment is expressed. Classifiers are trained to determine the polarities of forthcoming texts. Naïve Bayes classifier, maximum entropy classifier and Support Vector Machine (SVM) are the most commonly used models. However, the requirement of annotated corpora (**language science**) is not easily satisfied, especially for cross-domain and cross-lingual situations. Semi-supervised methods, which train classifiers on both annotated and unannotated corpora in different domains and languages, are developed to deal with the lack of annotated corpora. Several unsupervised methods based on Latent Dirichlet Allocation (LDA) have been proposed to release the dependence of annotated corpora. (**In natural language processing, the latent Dirichlet allocation (LDA) is a generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar.**)
Lexicon approaches determine the sentiment score of text according to sentiment lexicons in an unsupervised manner. **A lexicon is a dictionary of sentiment words and phrases with their polarities and strengths.**