A literature review on text classification and sentiment analysis approaches

ABSTRACT

Sentiment analysis is an important branch task of text classification and the related system usually is applied to in perception of user emotion and public opinion monitoring. By comparison, the text classification can be applied to more fields than sentiment analysis. In the system architecture, same as text classification, the complete classification systemmainly contains data acquisition, data pre-process, feature extraction, classification algorithm and result output.TheWeb crawler usually be used in first step, the URL Link, hashtags, Non-Chinese text should be removed in second step. In feature extraction, the IG, TF-IDF, Word2vec usually be used. Then, the SVM, Naive Bayes, KNN or Neural network algorithm usually be used in classifier. Furthermore, as a system that can run automatically, the sentiment analysis system should be able to extract significant feature from corpus and make accurately analysis about emotional polarity of text corpus. At present, the system improvement direction of related system focuses on 3 aspects: data acquisition, feature extraction and classifier algorithm.