

Abstract

Sentiment in a sentence is the value of the emotion contained in the sentence can contain positive or negative emotions and usually a sentence that does not have the value of sentiment is regarded as neutral sentences. This sentiment that be in a variety of environments such as the review on a product, a person's status in social media and citations in scientific papers. Generally, in making a paper citing tend to be made with a neutral orientation, so as to avoid direct criticism on scientific papers referenced. But it is possible to make a cite with positive or negative orientation. Sentiment in the citation sentence can be used to determine the quality of a scientific work, to detect relationships between scientific work, summarisation and others. Many methods can be implemented to perform such sentiment analysis to determine the best classifier and also the best features. Due to the many methods that can be used, it appears the problem of finding the combination classifier and features used to produce high performance. In the previous studies have found the best combination in classifying sentiment on the citation sentence. But there are potential features that are not tested. In this study focuses on the analysis of the use of a combination performansi classifier with a combination of features and complementary combination of features that have not been tested in previous studies. The results of this study found that a combination of features 1-3 Grams + POSTag with Support Vector Machine classifier and 1-3Grams + Dparser combination with Logistic Regression classifier as a combination of the best features and classifier. The result of the combination is to be the best at the time of testing for 5-fold validation where each fold formed manual to mimic real events. Results from two combination of micro-f has more than 99%.

Keyword: sentiment, performance, citation, classifier, feature