English and Malay cross-lingual sentiment lexicon acquisition and analysis

ABSTRACT

Sentiment analysis finds opinions, sentiments or emotions in user-generated contents. Most efforts are focusing on the English language, for which a large amount of sources and tools for sentiment analysis are available. The objective of this paper is to introduce a cross-lingual sentiment lexicon acquisition method for the Malay and English languages and further being test on a set of news test collections. Several part of speech tags are being experimented using the Word Score Summation technique in order to classify the sentiment of the news articles. This method records up to 50% as experimental accuracy result and works better for verbs and negations in both the English and Malay news articles.