Abstract

In the current years, social media became one of the most important sources of data for different data analytic purposes. One of the most important issues is how to map different trends in social media and define the relation between different groups based on their sentiment or interests. In this paper, a two-phase approach is used for clustering a set of blogs. At the first phase, the approach builds a lexicon that provides the polarity of each word. In the second phase, the approach clusters the blogs bases on the polarity features and the Power Link features. The output of the second phase is used as the input of the first phase to get an improved lexicon. This process will continue in a loop between phase one and phase 2 till a stable set of clusters is gotten. The approach aims to develop a non-supervised cluster agent that can correctly cluster micro blogs and define different interests of different groups of people. The results of the approach are expressed in terms of precision, recall and F-measure.

References


Index Terms

Computer Science Artificial Intelligence

Keywords

Sentiment Analysis, Evolution Calculation, Genetic Algorithms, Power Links, Social Media, Micro Blogs.