Abstract

It is taxing to understand the current trends in the online market and then abridge the general opinions about the products due to the existence of diversified social media data. This has created a need for real time opinion mining which is analysis of the sentiments that classifies the text into positive and negative emotion polarities. In this paper, the author explores the most important step in sentiment analysis that is data pre-processing and analyses the different techniques used for pre-processing in R. The results show that using library packages provides better results with respect to the method where direct functions are used.

References


14. Peng Taoxin. A framework for data cleanings in data warehouses. School of computing, napier University, 10 Colinton Road, Edinburgh, EH10 5DT, UK.


Index Terms

Computer Science Information Sciences

Keywords

Sentiment analysis, data pre-processing, Tm Library, Natural language processing.