Abstract

Most of the e-commerce sites ask their customers to provide relevant reviews on their products which could help other customers to decide their choice. A slew of reviews is being generated on a daily basis due to an increase in the usage of e-commerce sites. A potential customer may need to go through thousands of reviews before arriving at a firm decision, which is time-consuming. The project elaborated below aims at reducing this time constraint, by providing an effective summarization of reviews in a manner suitable for users. Usage of MapReduce technique provided by Apache Hadoop is highly emphasized for processing reviews. The summarization of reviews is limited to attributes that the potential customers might be interested while looking for the particular product. In this paper, the technique used for the same is described which substantially reduces time complexity when implemented.

References
2. Yun Niu, MSc, Xiaodan Zhu, MSc, Jianhua Li, MSc and Graeme Hirst, PHD. “Analysis of Polarity Information in Medical Text” AMIA 2005 Symposium.

Index Terms

Computer Science
Information Sciences

Keywords

Sentiment Analysis, Opinion Mining, Product Reviews, Hadoop, MapReduce, OpenNLP, SentiWordNet.