

عنوان مقاله:

An Experimental Study on Performance of Text Representation Models for Sentiment Analysis

محل انتشار:

فصلنامه سیستم های اطلاعاتی و مخابرات، دوره 8، شماره 1 (سال: 1399)

تعداد صفحات اصل مقاله: 8

نویسندگان:

Sajjad Jahanbakhsh Gudakahriz - *Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

Amir Masoud Eftekhari Moghadam - *Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

Fariborz Mahmoudi - *Advanced Analytics Dept., General Motors Company, Warren, Mi, USA*

خلاصه مقاله:

Sentiment analysis in social networks has been an active research field since ۲۰۰۰ and it is highly useful in the decision-making process of various domains and applications. In sentiment analysis, the goal is to analyze the opinion texts posted in social networks and other web-based resources to extract the necessary information from them. The data collected from various social networks and web sites do not possess a structured format, and this unstructured format is the main challenge for facing such data. It is necessary to represent the texts in the form of a text representation model to be able to analyze the content to overcome this challenge. Afterward, the required analysis can be done. The research on text modeling started a few decades ago, and so far, various models have been proposed for performing this modeling process. The main purpose of this paper is to evaluate the efficiency and effectiveness of a number of commons and famous text representation models for sentiment analysis. This evaluation is carried out by using these models for sentiment classification by ensemble methods. An ensemble classifier is used for sentiment classification and after preprocessing, the texts is represented by selected models. The selected models for this study are TF-IDF, LSA, Word2Vec, and Doc2Vec and the used evaluation measures are Accuracy, Precision, Recall, and F-Measure. The results of the study show that in general, the Doc2Vec model provides better performance compared to other models in sentiment analysis and at best, accuracy is ۰.۷۲.

کلمات کلیدی:

Text Representation Models; Sentiment Analysis; Sentiment Classification; Ensemble Classifiers

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1352394>

