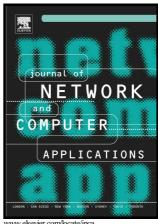
Author's Accepted Manuscript

SentiRelated: a Cross-Domain Sentiment Classification Algorithm for Short Texts through Sentiment Related Index

Lei Wang, Jianwei Niu, Houbing Song, Mohammed Atiquzzaman



www.elsevier.com/locate/jnca

PII: S1084-8045(17)30358-2

DOI: https://doi.org/10.1016/j.jnca.2017.11.001

Reference: YJNCA2004

To appear in: Journal of Network and Computer Applications

Received date: 28 June 2017

Revised date: 10 September 2017 Accepted date: 3 November 2017

Cite this article as: Lei Wang, Jianwei Niu, Houbing Song and Mohammed Atiquzzaman, SentiRelated: a Cross-Domain Sentiment Classification Algorithm for Short Texts through Sentiment Related Index, *Journal of Network and Computer Applications*, https://doi.org/10.1016/j.jnca.2017.11.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

SentiRelated: a Cross-Domain Sentiment Classification Algorithm for Short Texts through Sentiment Related Index

Lei Wang^a, Jianwei Niu^{a,*}, Houbing Song^b, Mohammed Atiquzzaman^c

^aState Key Laboratory of Software Development Environment, School of Computer Science and Engineering, Beihang University, Beijing 100191, China ^bDepartment of Electrical and Computer Engineering, West Virginia University, WV 25136, USA ^cSchool of Computer Science, University of Oklahoma, Norman, OK 73019, USA

Abstract

Sentiment classification for short texts, aiming at predicting sentiment polarity of short texts automatically, has attracted more and more attentions due to its wide applications. Traditional supervised classification approaches perform well in predicting the sentiment polarity for a given domain, but the performance decreases drastically when a classifier trained on a specific domain is directly applied to predict the sentiment polarity of another domain because the words used in the trained domain may not appear in the test domain. Moreover, the same word may indicate different sentiment polarities in different domains. In this paper, to bridge the gap between different domains, we create a Sentiment Related Index (SRI) to measure the association between different lexical elements in a specific domain with the help of domain-independent features as a bridge. Then we propose a novel cross-domain sentiment classification algorithm based on SRI, which is termed SentiRelated, to analyze the sentiment polarity for short texts. SentiRelated utilizes SRI to expand feature vectors based on unlabeled data from the target domain. In this way, some important sentiment indicators for the target domain are appended to feature vectors. At last, we validate our SentiRelated algorithm on two typical datasets. The experimental results demonstrate that, compared with state-of-the-art algorithms, our SentiRelated algorithm can improve the per-

Email addresses: lei@buaa.edu.cn (Lei Wang), niujianwei@buaa.edu.cn (Jianwei Niu), Houbing.Song@mail.wvu.edu (Houbing Song), atiq@ou.edu (Mohammed Atiquzzaman)

^{*}Corresponding author

Download English Version:

https://daneshyari.com/en/article/6884902

Download Persian Version:

https://daneshyari.com/article/6884902

<u>Daneshyari.com</u>