

Accepted Manuscript

Applying convolution filter to matrix of word-clustering based document representation

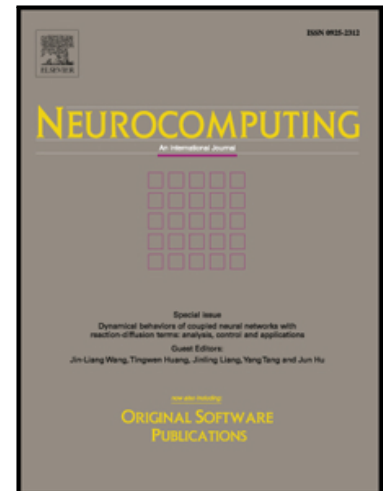
Younghoon Lee, Jinbae Im, Sungzoon Cho, Jinhae Choi

PII: S0925-2312(18)30846-4
DOI: [10.1016/j.neucom.2018.07.018](https://doi.org/10.1016/j.neucom.2018.07.018)
Reference: NEUCOM 19766

To appear in: *Neurocomputing*

Received date: 9 December 2017
Revised date: 16 June 2018
Accepted date: 2 July 2018

Please cite this article as: Younghoon Lee, Jinbae Im, Sungzoon Cho, Jinhae Choi, Applying convolution filter to matrix of word-clustering based document representation, *Neurocomputing* (2018), doi: [10.1016/j.neucom.2018.07.018](https://doi.org/10.1016/j.neucom.2018.07.018)



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 proof before it is published in its final 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.

Applying convolution filter to matrix of word-clustering based document representation

Younghoon Lee^{a,b}, Jinbae Im^a, Sungzoon Cho^{a,*}, Jinhae Choi^b

^a*Department of Industrial Engineering and Institute for Industrial Systems Innovation, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea*

^b*Data Driven User Experience team, Mobile Communication Lab, LG Electronics, 10 Magokjungang 10-ro, Gangseo-gu, Seoul 07796, Korea*

Abstract

Word-clustering based document representation approaches have been suggested recently to overcome previous limitations such as high dimensionality or loss of innate interpretation; they show higher classification performance than other recent methods. Thus, we present a novel way to combine the advantages of various word-clustering based representation approaches. Instead of previous approaches, which represent documents in vector form, we represent documents in matrix form while concatenate various representation results. And we proposed another novel way to apply convolution filter to those representation while rearranging the elements by preserving the semantic distance. In order to verify the representation performance of our proposed methods, we utilized the kinds of dataset: customer-voice data from LG Electronics, public Reuter news dataset and 20 Newsgroup dataset. The results demonstrated that the proposed method outperforms all other

*Corresponding author

Email addresses: yhoon.lee@dm.snu.ac.kr (Younghoon Lee), jinbae@dm.snu.ac.kr (Jinbae Im), zoon@snu.ac.kr (Sungzoon Cho), jin.choi@lge.com (Jinhae Choi)

Download English Version:

<https://daneshyari.com/en/article/10151171>

Download Persian Version:

<https://daneshyari.com/article/10151171>

[Daneshyari.com](https://daneshyari.com)