Accepted Manuscript

Title: Automatic Text Classification Algorithm based on Gauss Improved Convolutional Neural Network

Author: Jian-hai Du

Please cite this article as: Jian-hai Du, Automatic Text Classification Algorithm based on Gauss Improved Convolutional Neural Network, Journal of Computational Sciencehttp://dx.doi.org/10.1016/j.jocs.2017.06.010

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.

Automatic Text Classification Algorithm based on Gauss Improved Convolutional Neural Network

Jian-hai Du[1].

1. State Key Laboratory of Software Development Environment, Beihang University,Beijing China (100191)

Highlights

- Firstly, the effective dimensionality reduction of similarity calculation amount of the algorithm is realized based on the clustering center, and the original large-scale document samples are replaced with a relatively small number of clustering sample centers to realize improvement of the KNN query process.

- Secondly, MapReduce parallel framework is used to meet real-time demand of large-scale text classification and calculation combined with features of text classification, and to effectively overcome slow speed of the KNN query process and ensure accuracy of text classification as higher as possible.

- Finally, the classification speed of proposed algorithm can be effectively improved under the premise of ensuring sufficient accuracy through comparison in experiment of text classification accuracy and algorithmic efficiency with the similar single-threaded algorithm.