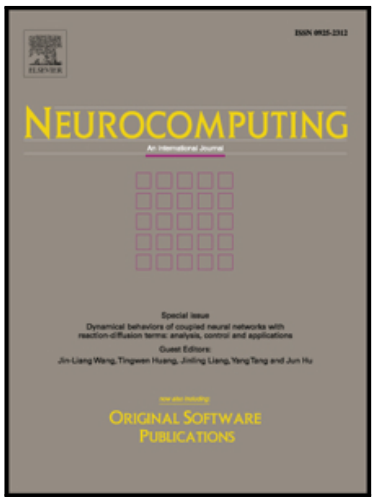

Iterative Sparsity Score for Feature Selection and Its Extension for Multimodal Data

Chen Zu, Linling Zhu, Daoqiang Zhang

PII: S0925-2312(17)30242-4
DOI: 10.1016/j.neucom.2016.08.124
Reference: NEUCOM 18038

To appear in: *Neurocomputing*

Received date: 12 February 2016
Revised date: 10 June 2016
Accepted date: 3 August 2016

Please cite this article as: Chen Zu, Linling Zhu, Daoqiang Zhang, Iterative Sparsity Score for Feature Selection and Its Extension for Multimodal Data, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2016.08.124

Iterative Sparsity Score for Feature Selection and Its Extension for Multimodal Data

Chen Zu[a,*], Linling Zhu[b,*], Daoqiang Zhang[a,**]

[a] *College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China*
[b] *Patent Examination Cooperation Jiangsu Center of the Patent Office, Suzhou, China*

Abstract

As a key dimensionality reduction technique in pattern recognition, feature selection has been widely used in information retrieval, text classification and genetic data analysis. In recent years, structural information contained in samples for guiding feature selection has become a new hot spot in machine learning field. Although tremendous feature selection methods have been developed, less important features are still used to construct the structure in those conventional structure based feature selection approaches. In this paper, we propose a new filter-type feature selection method called iterative sparsity score, which is independent of any learning algorithm. The proposed method can preserve the structural information by sparse representation, which can be efficiently solved by a ℓ_1-norm minimization problem. To exclude data noise, at one time we discard last m features and iteratively optimize the ℓ_1-norm minimization problem. We perform clustering and classification experiments on numerous bench mark datasets. Furthermore, its extension for multimodal data is also developed. We adopt the multi-modality alzheimer's disease data for classification to evaluate the extended method. The experimental results show the effectiveness of our proposed methods compared with several popular feature selection approaches.

Keywords: Feature selection, iterative sparse representation, multi-modality, alzheimer's disease, clustering, classification

2010 MSC: 00-01, 99-00

1. Introduction

With the rapid development of technology, big data have emerged with large dimensions or huge numbers. For example, high resolution face pictures include a lot of information with high dimensions, but at the same time the high dimensions limit the utilization in practical applications. One 3-D magnetic resonance image

*Chen Zu and Linling Zhu are co-first authors
**Corresponding author
Email addresses: chenzu@nuaa.edu.cn (Chen Zu), zhu_linling@126.com (Linling Zhu), dqzhang@nuaa.edu.cn (Daoqiang Zhang)