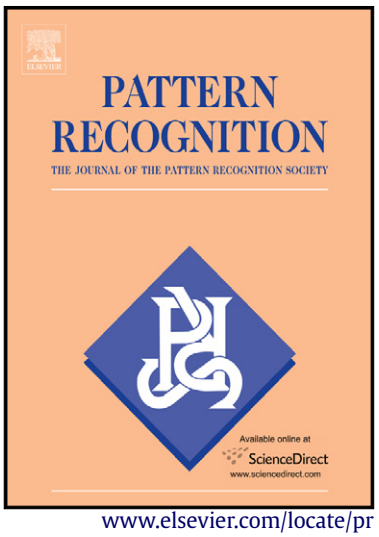
Author's Accepted Manuscript



www.elsevier.com/locate/pr

Incremental partial least squares analysis of big streaming data

Xue-Qiang Zeng, Guo-Zheng Li

PII: S0031-3203(14)00224-6
DOI: http://dx.doi.org/10.1016/j.patcog.2014.05.022
Reference: PR5130

To appear in: *Pattern Recognition*

Received date: 22 January 2014
Revised date: 18 May 2014
Accepted date: 31 May 2014

Cite this article as: Xue-Qiang Zeng, Guo-Zheng Li, Incremental partial least squares analysis of big streaming data, *Pattern Recognition,* http://dx.doi.org/10.1016/j.patcog.2014.05.022

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

Incremental Partial Least Squares Analysis of Big Streaming Data

Xue-Qiang Zeng[a,b], Guo-Zheng Li[c,*]

[a]*Computer Center, Nanchang University, Nanchang 330031, China*
[b]*Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 201804, China*
[c]*Department of Control Science & Engineering, Tongji University, Shanghai 201804, China*

Abstract

Incremental feature extraction is effective for facilitating the analysis of large-scale streaming data. However, most current incremental feature extraction methods are not suitable for processing streaming data with high feature dimensions because only a few methods have low time complexity, which is linear with both the number of samples and features. In addition, feature extraction methods need to improve the performance of further classification. Therefore, incremental feature extraction methods need to be more efficient and effective. Partial least squares (PLS) is known to be an effective dimension reduction technique for classification. However, the application of PLS to streaming data is still an open problem. In this study, we propose a highly efficient and powerful dimension reduction algorithm called incremental PLS (IPLS), which comprises a two-stage extraction process. In the first stage, the PLS target function is adapted so it is incremental by updating the historical mean to extract the leading projection direction. In the second stage, the other projection directions are calculated based on the equivalence between the PLS vectors and the Krylov sequence. We compared the performance of IPLS with other state-of-the-art incremental feature extraction methods such as incremental principal components analysis, incremental maximum margin criterion, and incremental inter-class scatter using real streaming datasets. Our empirical results showed that IPLS performed better than other methods in terms

*Corresponding author at: Room 605, Dianxin College Building, Caoan Road 4800, China. Tel.: +86 21 6958 3706; fax: +86 21 6958 9241.
Email address: gzli@tongji.edu.cn (Guo-Zheng Li)

Accepted manuscript

Download English Version:

https://daneshyari.com/en/article/10360342

Download Persian Version:

https://daneshyari.com/article/10360342

Daneshyari.com