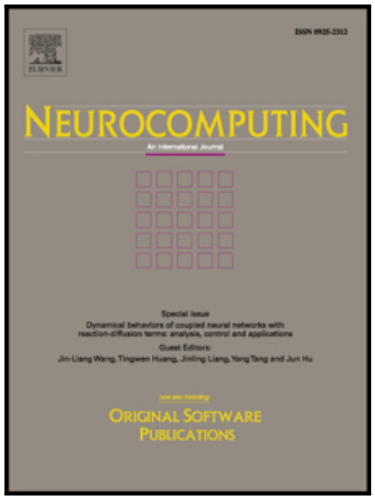
Communicated by Deng Cai

Accepted Manuscript



Semi-supervised least squares nonnegative matrix factorization and graph-based extension

Xinjun Peng, De Chen, Dong Xu

PII: S0925-2312(18)31088-9
DOI: https://doi.org/10.1016/j.neucom.2018.09.026
Reference: NEUCOM 19955

To appear in: *Neurocomputing*

Received date: 24 February 2018
Revised date: 25 June 2018
Accepted date: 4 September 2018

Please cite this article as: Xinjun Peng, De Chen, Dong Xu, Semi-supervised least squares nonnegative matrix factorization and graph-based extension, *Neurocomputing* (2018), doi: https://doi.org/10.1016/j.neucom.2018.09.026

Semi-supervised least squares nonnegative matrix factorization and graph-based extension

Xinjun Peng*, De Chen, Dong Xu

Department of Mathematics, Shanghai Normal University, Shanghai 200234, China

Abstract

Nonnegative matrix factorization (NMF) has been widely used in information retrieval, computer vision, and face recognition because it captures the underlying intrinsic structure of data by using its parts-based representations in the low dimensional space. In this paper, we first propose a novel semi-supervised least squares NMF (SLSNMF) method. This SLSNMF uses hyperplanes to separate the labelled points from different classes and minimizes the least squares loss. By considering the maximum margin principle, the discriminative abilities of clustering representations are greatly enhanced. We further present a graph-based least squares NMF (GLSNMF) method by incorporating the local manifold regularization into SLSNMF. Clustering experiments on five popular databases verify the effectiveness of our proposed SLSNMF and GLSNMF compared to the other state-of-the-art methods.

Keywords: Nonnegative matrix factorization, semi-supervised learning, maximum margin principle, least squares, graph-based manifold

1. Introduction

As an emerging technique for dimensionality reduction, nonnegative matrix factorization (NMF), proposed by Lee and Seung [1, 2], aims to find two nonnegative matrices U (basis matrix) and V (encoding matrix) whose product provides a good approximation to the original matrix X, such that $X \approx UV^T$. The nonnegative constraints lead to a parts-based representation because they allow only additive, not subtractive, combinations.

*Corresponding author. Tel.: +86-21-64324866.
Email address: xjpeng@shnu.edu.cn (Xinjun Peng)