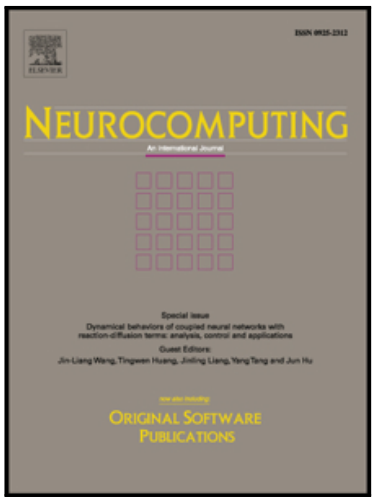

Semi-supervised Metric Learning in Stratified Spaces via Intergrating Local Constraints and Information-theoretic Non-Local Constraints

Zohre Karimi , Saeed Shiry Ghidary

PII: S0925-2312(18)30674-X
DOI: 10.1016/j.neucom.2018.05.089
Reference: NEUCOM 19650

To appear in: *Neurocomputing*

Received date: 23 March 2017
Revised date: 18 March 2018
Accepted date: 21 May 2018

Please cite this article as: Zohre Karimi , Saeed Shiry Ghidary , Semi-supervised Metric Learning in Stratified Spaces via Intergrating Local Constraints and Information-theoretic Non-Local Constraints, *Neurocomputing* (2018), doi: 10.1016/j.neucom.2018.05.089

Semi-supervised Metric Learning in Stratified Spaces via Intergrating Local Constraints and Information-theoretic Non-Local Constraints

Zohre Karimi
Department of Computer Engineering & IT
Amirkabir University of Technology
Tehran, Iran
z_karimi@aut.ac.ir

Saeed Shiry Ghidary
Department of Computer Engineering & IT
Amirkabir University of Technology
Tehran, Iran
shiry@aut.ac.ir

Abstract— Considerable research efforts have been done in learning semi-supervised distance metric learning based on both manifold and cluster assumptions in the past few years. However, there is a major problem with them once they are applied to data lying on stratified space. The problem is that label smoothness assumption on manifold and cluster may be violated in the intersecting regions of manifolds. This problem is caused by overlearning of locality that misleads the metric learning process in the absence of enough labeled data.

In this paper, we will propose a novel semi-supervised metric learning for stratified spaces (S2MLS2) which removes unsuitable local constraints in the manifold based methods for adapting to the smoothness assumption on multi manifolds. We will also impose some non-local constraints to detect the shared structures at different positions in the absence of enough supervised information. Besides, a novel bootstrapping method based on smoothness assumption on multi manifolds will be proposed to enlarge the labeled data.

The proposed algorithm is based on different behavior of Laplacian of piecewise-smooth function on multi manifolds in the neighborhood of non-interior points of the manifolds as compared with interior points of the manifolds. Experiments on artificial and real benchmark data sets demonstrate that the proposed metric learning method outperforms many state-of-the-art metric learning methods.

Keywords-Manifold, Semi-supervised, Laplacian, Stratified Space, Kullback-Leibler, Metric learning.

1. INTRODUCTION

Recent studies show that metric learning has vital role in machine learning tasks [26, 29, 37, 41, 42, 56]. Indeed, appropriate choice of a metric has substantial impact on performance of both clustering and classification tasks. The goal of metric learning is to learn a pairwise real-valued distance function which preserves similar data close together and keeps dissimilar data apart. Distance metric learning algorithms are categorized into three types: supervised, unsupervised and semi-supervised.

Unsupervised metric learning methods attempt to preserve local topology and/or geometric structure of data as much as possible. It is formulated as a low-dimensionality reduction or/and