Accepted Manuscript

Random Multi-Graphs: A Semi-Supervised Learning Framework for Classification of High Dimensional Data

Qin Zhang, Jianyuan Sun, Guoqiang Zhong, Junyu Dong

PII: S0262-8856(16)30130-5
DOI: doi: 10.1016/j.imavis.2016.08.006
Reference: IMAVIS 3541

To appear in: *Image and Vision Computing*

Received date: 14 May 2016
Revised date: 22 July 2016
Accepted date: 16 August 2016

Please cite this article as: Qin Zhang, Jianyuan Sun, Guoqiang Zhong, Junyu Dong, Random Multi-Graphs: A Semi-Supervised Learning Framework for Classification of High Dimensional Data, *Image and Vision Computing* (2016), doi: 10.1016/j.imavis.2016.08.006

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.

ACCEPTED MANUSCRIPT

Random Multi-Graphs: A Semi-Supervised Learning Framework for Classification of High Dimensional Data

Qin Zhang[a,b], Jianyuan Sun[a], Guoqiang Zhong[a], Junyu Dong[a,*]

[a] *Department of Computer Science and Technology, Ocean University of China, No.238 Songling Road, Qingdao, 266100, China*
[b] *Department of Science and Information, Agriculture University of Qingdao, No.700 Changcheng Road, Qingdao, 266109, China*

Abstract

Currently, high dimensional data processing confronts two main difficulties: inefficient similarity measure and high computational complexity in both time and memory space. Common methods to deal with these two difficulties are based on dimensionality reduction and feature selection. In this paper, we present a different way to solve high dimensional data problems by combining the ideas of Random Forests and Anchor Graph semi-supervised learning. We randomly select a subset of features and use the Anchor Graph method to construct a graph. This process is repeated many times to obtain multiple graphs, a process which can be implemented in parallel to ensure runtime efficiency. Then the multiple graphs vote to determine the labels for the unlabeled data. We argue that the randomness can be viewed as a kind of regularization. We evaluate the proposed method on eight real-world data sets by comparing it with two traditional graph-based methods and one state-of-the-art semi-supervised learning method based on Anchor Graph to show its effectiveness. We also apply the proposed method to the subject of face recognition.

Keywords: Semi-supervised learning, Graph, Regularization, Randomness, Anchors

2010 MSC: 00-01, 99-00

*Corresponding author

Email addresses: zhang_qin1005@163.com (Qin Zhang), sunjianyuan11@163.com (Jianyuan Sun), gqzhong@ouc.edu.cn (Guoqiang Zhong), dongjunyu@ouc.edu.ch (Junyu Dong)

Download English Version:

https://daneshyari.com/en/article/4968917

Download Persian Version:

https://daneshyari.com/article/4968917

Daneshyari.com