

本期目录 | 下期目录 | 过刊浏览 | 高级检索  
页] [关闭]

[打印本

机器学习与数据挖掘

谱聚类中特征向量的Bagging选取方法

王兴良,王立宏\*,李海军

烟台大学计算机学院, 山东 烟台 264005

摘要:

谱聚类算法中用亲和矩阵特征值最大的k个特征向量并不总是能有效地发现数据集的结构。为了选取较好特征向量,提出了一种特征向量的Bagging选取算法。以对约束计分方法为评价标准,对特征向量进行评价并选出较好的特征向量,将多次选择的特征向量进行Bagging集成(Bootstrap aggregating),得出k个特征向量的组合。该算法能够较好地选取出特征向量,根据UCI实验数据集的测试,证实该算法对测试数据集可以得出较好的预测结果。

关键词: 特征向量选择 谱聚类 Bagging方法 约束计分 拉普拉斯矩阵

Eigenvector selection in spectral clustering based on Bagging

WANG Xing-liang, WANG Li-hong\*, LI Hai-jun

School of Computer Science & Technology, Yantai University, Yantai 264005, China

Abstract:

For the spectral clustering algorithm, the largest k eigenvectors of the affinity matrix derived from the dataset were not always able to find the structure of dataset effectively. An eigenvector selection algorithm in spectral clustering based on Bagging was proposed in order to select better eigenvectors. The eigenvectors were evaluated by pairwise constraints score. First, some eigenvectors were ranked according to their constraint scores, and then the suitable eigenvectors were selected from the ranking list, finally the optimal combination of k eigenvectors was obtained by Bagging based ensemble algorithm. The better eigenvectors could be achieved. Experimental results on UCI benchmark datasets showed that this algorithm could gain satisfactory prediction results.

Keywords: eigenvector selection spectral clustering Bagging method constraint score Laplacian matrix

收稿日期 2012-10-12 修回日期 网络版发布日期

DOI:

基金项目:

国家自然科学基金资助项目(61170224);山东省计算机网络重点实验室开放课题基金资助项目(SDKLCN-2012-03)

通讯作者: 王立宏(1970-),女,吉林镇赉人,教授,博士,硕士生导师,主要研究方向为数据挖掘.E-mail:wanglh-000@163.com

作者简介: 王兴良(1988-),男,山东临朐人,硕士研究生,主要研究方向为数据挖掘.E-mail:wangxingliang0911@163.com

作者Email:

PDF Preview

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(1485KB)
- ▶ 参考文献[PDF]
- ▶ 参考文献

服务与反馈

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ 引用本文
- ▶ Email Alert
- ▶ 文章反馈
- ▶ 浏览反馈信息

本文关键词相关文章

- ▶ 特征向量选择
- ▶ 谱聚类
- ▶ Bagging方法
- ▶ 约束计分
- ▶ 拉普拉斯矩阵

本文作者相关文章

PubMed