

论文

基于k-means和半监督机制的单类中心学习算法

李志圣 孙越恒 何丕廉 侯越先

天津大学 计算机科学与技术学院 天津大学 计算机科学与技术学院 天津大学 计算机科学与技术学院 天津大学 计算机科学与技术学院

摘要:

提出了一个基于k-means算法框架和半监督机制的single-means算法,以解决单类中心学习问题。k-means算法实质上是对一种混合高斯模型的期望最大化(EM)算法的近似,对该模型随机生成的多类混合数据集,从目标类中随机标定的初始中心出发,能确定地收敛到该类的实际中心。将single-means算法应用到对单类文本中心学习问题中,实验结果表明:在给定目标类中的小标定文本集后,新算法能够有效地改进类的初始中心,且对数据稀疏和方差较大的实际问题具有健壮性。

关键词: k-means 单类学习 半监督学习 single-means

Algorithm for learning centre of single class based on k-means and semi-supervised mechanism

Abstract:

A new algorithm named "single-means" was presented to improve the centre estimation of the object class when a hybrid data set had unknown k value and feature of accumulating to centre. Based on that k-means algorithm was equivalent to Expectation Maximum (EM) algorithm on a special hybrid Gaussian model, it was proved that given a data set generated by the above Gaussian model, the true centre of the object Gaussian distribution could be converged by a new algorithm. The new algorithm was applied in learning the centre of single text class. The experiment shows that given a small labeled text set, the new algorithm can get a better centre, and is robust on sparse data set and that with great variance.

Keywords: k-means single class learning semi-supervised learning single-means.

收稿日期 2008-04-23 修回日期 2008-05-27 网络版发布日期

DOI:

基金项目:

通讯作者: 李志圣

作者简介:

参考文献:

本刊中的类似文章

1. 白如珍 田青 徐海江.一种新的基于分水岭变换的聚类分析算法[J]. 计算机应用, 2008,28(12): 3240-3243
2. 孙秀娟 刘希玉.基于新聚类有效性函数的改进K-means算法[J]. 计算机应用, 2008,28(12): 3244-3247
3. 张磊 项学智 赵春晖.基于光流场与水平集的运动目标检测[J]. 计算机应用, 2009,29(4): 972-975,
4. 王明春 唐万生 江琪 刘鑫.基于相对距离的改进粗K-means方法[J]. 计算机应用, 2009,29(4): 1102-1105
5. 罗 冯国灿 成秋生.结合聚类和改进的C-V演化方程在医学图像分割中的应用[J]. 计算机应用, 2008,28(9): 2288-2291

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(701KB)
- ▶ [HTML全文]
- ▶ 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

- ▶ k-means
- ▶ 单类学习
- ▶ 半监督学习
- ▶ single-means

本文作者相关文章

- ▶ 李志圣
- ▶ 孙越恒
- ▶ 何丕廉
- ▶ 侯越先

PubMed

- ▶ Article by
- ▶ Article by
- ▶ Article by
- ▶ Article by

文章评论 (请注意:本站实行文责自负, 请不要发表与学术无关的内容!评论内容不代表本站观点.)

|      |                      |      |                                   |
|------|----------------------|------|-----------------------------------|
| 反馈人  | <input type="text"/> | 邮箱地址 | <input type="text"/>              |
| 反馈标题 | <input type="text"/> | 验证码  | <input type="text" value="4468"/> |