

软件技术与数据库

一个对不带类别标记文本进行分类的方法

蒋志方<sup>1</sup>, 祝翠玲<sup>2</sup>, 吴 强<sup>1</sup>

(1. 山东大学计算机科学与技术学院, 济南 250061; 2. 山东经济学院信息管理学院, 济南 250014)

收稿日期 修回日期 网络版发布日期 2007-6-15 接受日期

**摘要** 利用无监督聚类方法和朴素贝叶斯分类的特点, 把UC获得的预分类结果作为朴素贝叶斯分类器的训练样本, 将处在聚类结果中类属模糊区域的文本交给训练好的朴素贝叶斯分类器再进行分类, 实现了对不带任何类别标记文本的准确分类, 可得到较准确的分类结果。

**关键词** [文本分类](#) [无监督文本聚类](#) [朴素贝叶斯分类](#) [欧氏距离](#)

分类号 [TP181](#)

**DOI:**

通讯作者:

作者个人主页: [蒋志方<sup>1</sup>](#); [祝翠玲<sup>2</sup>](#); [吴 强<sup>1</sup>](#)

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(139KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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](#)

相关信息

▶ [本刊中 包含“文本分类”的 相关文章](#)

▶ [本文作者相关文章](#)

▶ [蒋志方<sup>1</sup>, 祝翠玲<sup>2</sup>, 吴 强<sup>1</sup>](#)