

人工智能

基于向量空间模型的中文文本层次分类方法研究

肖雪, 何中市

重庆大学计算机学院

收稿日期 2005-11-3 修回日期 2006-1-13 网络版发布日期 接受日期

**摘要** 在文本分类的类别数量庞大的情况下, 层次分类是一种有效的分类途径。针对层次分类的结构特点, 考虑到不同的层次对特征选择和分类方法有不同的要求, 提出了新的基于向量空间模型的二重特征选择方法FDS以及层次分类算法HTC。二重特征选择方法对每一层均进行一次特征选择, 并逐层改变特征数量和权重计算方法; HTC算法把分别对粗分和细分更有效的类中心向量法与SVM方法相结合。实验表明, 该方法相对于平面分类和一般的层次分类方法, 有较高的准确率。

**关键词** [层次分类, 向量空间模型, 二重特征选择, 权重计算](#)

分类号

**DOI:**

对应的英文版文章: [5115760](#)

通讯作者:

肖雪 [zaigjdh@hotmai.com](mailto:zaigjdh@hotmai.com)

作者个人主页: 肖雪; 何中市

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF \(640KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“层次分类, 向量空间模型, 二重特征选择, 权重计算”的 相关文章](#)

▶ 本文作者相关文章

· [肖雪](#)

· [何中市](#)