

数据挖掘

一种面向专利文献数据的文本自动分类方法

蒋健安¹;陆介平²;倪巍伟¹;孙志挥¹

东南大学¹

江苏省镇江市科技局²

收稿日期 2007-7-11 修回日期 2007-9-10 网络版发布日期 2008-1-1 接受日期

摘要 中文专利文献自动分类目前尚无成熟适用的方法。分析了文本自动分类的关键技术,并结合专利数据的特点对无词典分词和权重计算进行了改进,提出了一种适用于专利数据分类的层次分类方法,给出了面向专利文献数据的文本自动分类系统的框架模型。实验表明,该系统具有较好的分类精度与效率。

关键词 [文本分类](#) [专利文献](#) [IPC](#) [K-近邻](#)

分类号

DOI:

对应的英文版文章: [A7073726](#)

通讯作者:

蒋健安 jja65@163.com; jja1982@gmail.com

作者个人主页: [蒋健安](#) [陆介平](#) [倪巍伟](#) [孙志挥](#)

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF \(628KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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](#)

· [蒋健安](#)

· [陆介平](#)

· [倪巍伟](#)

· [孙志挥](#)