

图书基本信息

书名：<<基于遗传算法的文本分类及聚类研究>>

13位ISBN编号：9787030227416

10位ISBN编号：7030227417

出版时间：2008-8

出版时间：科学出版社

作者：戴文华

页数：222

字数：220000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

前言

戴文华同志无论学习、研究还是工作，都非常勤奋、踏实、执著，研究能力和钻研精神都非常强。四年前，他进入华中师范大学计算机科学系攻读硕士研究生时，已经作为高校教师工作了多年，并已取得了较多的成果；然而他在自然语言处理这个研究方向上，还是一个新兵。但是，经过短短的4年，他已经在该方向做出了突出的研究成果，在文本分类和聚类的理论和方法上，做了系统深入的研究，发表了论文10多篇。现在他硕士期间的主要研究成果即将出版专著，实在是可喜可贺。文本自动分类和聚类是自然语言处理领域里重要的研究课题，有广泛的应用前景。

内容概要

文本分类和聚类技术是应信息检索和查询需要而出现的自然语言处理领域的重要研究课题。

文本分类和聚类问题中的特征选择和抽取技术、文本特征表示、聚类方法的选择和实现以及分类方法的选择和实现,都将对文本分类和聚类结果产生极大影响。

针对文本分类和聚类中的文本数据的高维性和稀疏性、同义词和近义词问题、效率与精确度之间的搭配问题以及参数优化问题,本书提出了使用遗传算法与传统分类和聚类方法相结合的思路来进行处理,充分利用了遗传算法的全局优化能力和传统分类及聚类算法的专业知识,有效地提高了文本分类和聚类的效率与精度。

本书可作为自然语言处理专业和相关专业人员自学参考书。

书籍目录

序前言第1章 绪论 1.1 研究的的意义 1.2 国内外研究现状 1.3 文本分类和聚类存在的问题 1.4 本书研究内容和目标 参考文献第2章 文本分类和聚类的基本理论及方法 2.1 文本分类和聚类的概念 2.2 文本的表示 2.3 文本预处理 2.4 文本特征选择和抽取 2.5 文本分类方法 2.6 文本聚类方法 2.7 本章小结 参考文献第3章 遗传算法基础知识 3.1 遗传算法概述 3.2 标准遗传算法 3.3 遗传算法染色体编码 3.4 适应度函数 3.5 遗传算子 3.6 遗传算法的改进 3.7 本章小结 参考文献第4章 混合并行遗传算法及其在文本聚类中的应用 4.1 k-means算法初始聚类中心的选择 4.2 混合并行遗传算法 4.3 基于并行遗传算法的文本特征词提取 4.4 基于混合并行遗传算法的文本聚类 4.5 实验设置及结果分析 4.6 本章小结 参考文献第5章 混合并行遗传算法在文本分类中的应用 5.1 基于混合并行遗传算法的潜在语义挖掘 5.2 基于混合并行遗传聚类的KNN改进算法(HPGA-KNN) 5.3 SMO-SVM算法 5.4 SMO-SVM算法核函数参数优化 5.5 改进的KNN SVM文本分类方法 5.6 实验设置及分析 5.7 本章小结 参考文献第6章 总结和研究展望 6.1 总结 6.2 后续研究工作展望附录A k-means聚类算法关键代码附录B 遗传算法聚类关键代码附录C 混合遗传聚类关键代码

章节摘录

第1章 绪论 让Internet更好地为人类服务，一直是人们努力的方向。通过不懈努力，人们逐渐掌握了多种技术对不断膨胀的信息进行收集和整理。各种海量信息的复杂性和非结构性，对信息的处理带来了极大的挑战。本书为了解决信息处理中的一些相关问题而进行了深入研究，提出了一些切实有效的措施，为海量信息的处理提供了极大帮助。

.....

编辑推荐

《基于遗传算法的文本分类及聚类研究》作者围绕文本分类和聚类研究，针对一些具体问题，提出了一系列科学可行的解决方案。

如基于并行遗传算法的特征词动态提取方法，能够降低文本对象的特征维数；采用并行遗传算法对文本分类和聚类问题中的参数进行优化，从而提高文本分类和聚类结果的精确度。

《基于遗传算法的文本分类及聚类研究》除了阐述作者自己的研究成果外，作者还结合自己的实际经验，对文本自动分类技术、文本自动聚类技术、遗传算法的相关技术都作了较详细的描述，对于准备从事相关研究的人士有很大的参考价值。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>