

- >> 首页
- >> 被收录信息
- >> 投稿须知
- >> 模板下载
- >> 信息发布
- >> 常见问题及解答
- >> 合作单位
- >> 产品介绍
- >> 编委会/董事会
- >> 关于我们
- >> 网上订阅
- >> 友情链接

#### 友情链接

- >> 中国期刊网
- >> 万方数据资源库
- >> 台湾中文电子期刊
- >> 四川省计算应用研究中心
- >> 维普资讯网

## Web文本特征选择算法的研究

### Research about Algorithm of Web Text Feather Selection

摘要点击: 103 全文下载: 156

[查看全文](#) [查看/发表评论](#) [下载PDF阅读器](#)

中文关键词: 文本表示; 向量空间模型; 特征选择; 熵

英文关键词: Text Represents; Vector Space Model; Feather Selection; Entropy

基金项目: 国家教育部科研基金重点项目(200305); 河南省自然科学基金资助项目(0311012600)

作者	单位
冯长远, 普杰信	(河南科技大学 电子信息工程学院, 河南 洛阳 471003)

中文摘要:

以向量空间模型作为Web文本的表示方法, 结合Web文本的结构特征对向量空间模型中的特征选择算法进行了分析并加以改进。在改进的算法中, 体现出了特征词在Web文档结构中的位置信息; 引入了信息论中熵的概念, 用词的熵函数对权值进行调整, 从而更加准确地选取有效的特征词。实验验证了改进算法的可行性和有效性。

英文摘要:

This paper uses vector space model as the description of the Web text, analyses the feather selection algorithm and brings forward an improved algorithm in view of the construct character of the Web text. The new algorithm describes the situation information of the feather terms in Web text, introduces the concept of entropy and adjusts the weighting by the entropy function of the words, thus it can select feather terms more effectively. The experiment shows the feasibility and the validity of this method in feather selection.

[关闭](#)

您是第938069位访问者

主办单位: 四川省电子计算机应用研究中心 单位地址: 成都市武侯区成科西路3号

服务热线: 028-85249567 传真: 028-85210177 邮编: 610041 Email: arocmag@163.com; srcea@sichuan.net.cn

蜀ICP备05005319号 本系统由北京勤云科技发展有限公司设计