

软件技术与数据库

基于世界知识的深网数据源增强分类模型

黄 黎^{1,2}, 赵朋朋¹, 方 巍¹, 崔志明¹, 孙振强³

(1. 苏州大学智能信息处理及应用研究所, 苏州 215006; 2. 江苏广播电视大学, 南京 210017; 3. 南大苏富特科技有限公司, 苏州 215006)

收稿日期 修回日期 网络版发布日期 接受日期

摘要 针对传统词袋方法在深网(Deep Web)数据源分类应用中的局限性, 提出一种基于世界知识的Deep Web数据源增强分类模型, 通过对外部知识库的主题分析, 建立特征映射, 构造基于领域概念的辅助分类器, 丰富Deep Web查询表单的特征集合。基于Wikipedia百科知识库对真实Web数据进行分类。实验结果证明该模型有效。

关键词 [深网](#); [数据源分类](#); [主题分析](#); [特征映射](#); [世界知识](#)

分类号 [TP393](#)

DOI:

通讯作者:

作者个人主页: [黄 黎^{1;2}](#); [赵朋朋¹](#); [方 巍¹](#); [崔志明¹](#); [孙振强³](#)

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF \(239KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“深网; 数据源分类; 主题分析; 特征映射; 世界知识” 的相关文章](#)

▶ [本文作者相关文章](#)