

EncyCatalogRec: 针对百科文章补全的目录推荐

鲁伟明, 刘佳卉, 徐玮, 王鹏, 魏宝刚

浙江大学计算机科学与技术学院, 中国杭州市, 310027

摘要: 目前, 在线百科(如维基百科等)已提供海量且主题多样的文章。然而, 部分文章内容仍不够完善。本文提出EncyCatalogRec, 一种能为百科文章推荐相关目录, 从而帮助用户更好完善百科内容的系统。首先, 将百科文章和目录项表达为内嵌向量, 基于局部敏感哈希方法检索得到相关文章, 并以这些文章的目录项为候选项; 然后, 基于检索得到的文章及其目录项构建关系图, 进一步转为乘积图; 在乘积图上, 将目录推荐问题转为直推式学习问题; 最后, 基于学习排序算法对推荐得到的目录项排序。热启动和冷启动场景实验均证实, 本文所提方法性能优于已有方法。最后通过示例验证了所提方法性能。

关键词: 目录推荐; 百科文章补全; 乘积图; 直推式学习

<https://doi.org/10.1631/FITEE.1800363>