

基于互信息的推荐系统方法研究

郑诚^{1, 2}, 徐启南^{1, 2}, 章金平^{1, 2}

(1 安徽大学 计算智能与信号处理实验室, 安徽 合肥 230039; 2 安徽大学 计算机学院, 安徽 合肥 230601)

摘要: 随着网络信息的爆炸增长, 越来越多的信息, 让人目不暇接, 推荐系统这个时候应运而生, 本文提出了一种基于互信息的推荐系统, 从三方面分析数据集, 包括基于用户行为、商品标题以及商品标签, 用互信息来计算物品之间的相似度, 并且验证不同的权重, 观察对推荐效果的影响, 最终得到最好的推荐列表. 另外, 在本文数据集下, 本文方法和协同过滤推荐做对比, 证明本文的方法推荐效果优于协同过滤算法.

关键词: 推荐系统; 互信息; 权重; 协同过滤

Research on Recommender System Based on Mutual Information

ZHENG Cheng^{1, 2}, XU Qi-nan^{1, 2}, ZHANG Jin-ping^{1, 2}

(1 Key Laboratory of ICSP Ministry of Education, Anhui University, Hefei 230039, China; 2 School of Computer Science and Technology, Anhui University, Hefei 230601, China)

Abstract: With the explosion of network information, more and more information is dizzying and the recommendation system came into being. This paper presents a recommendation system based on mutual information, which analyzes the data set from three aspects, including user behavior, product titles and product tags, using mutual information to calculate the similarity between items, verify the different weights, observe the effect on the recommended effect, and finally get the best recommended list. In addition, under the data set of this paper, the proposed method and the collaborative filtering recommendation are compared to prove that the proposed method is better than collaborative filtering.

Key words: recommendation system; mutual information; weights; collaborative filtering

作者简介:

郑诚男, (1964-), 博士, 副教授. 研究方向为数据挖掘、自然语言处理、人工智能、信息检索.

徐启南 (通讯作者) 男, (1991-), 硕士研究生. 研究方向为数据挖掘、自然语言处理、信息检索. E-mail: 1850889412@qq.com.

章金平女, (1991-), 硕士研究生. 研究方向为数据挖掘、自然语言处理、信息检索.