

志愿服务推荐系统设计与实现

赵丰

2018 年 6 月 14 日

概要

主题背景

系统设计

算法实现

“WA” ,IBM 企业社会责任聚焦 “公益大数据”



参与者针对“志愿者管理”、“筹款募捐”、“社区参与”场景下的提升目标，用 WA 对样本数据进行分析，依据分析结果制定改善方案。

《社会与公益》2016 年 7 月刊

合作伙伴“灵析”是目前国内快速崛起、影响力最大的公益机构数据服务提供商，其捐赠人、志愿者 SaaS 产品服务超过 3000 家机构。

搜狐新闻，公益中国 2016 年 7 月 19 号

“中国公益基金会数据骇客马拉松竞赛”圆满落幕



当下社会对数据科学的热情主要集中在金融，互联网，医疗等热点，却忽视了公益领域的需求，这种局面需要得到改变。

搜狐新闻，基金会中心网 2017 年 9 月 6 日
引用 DAL 数据应用学院联合创始人陈晓理

竞赛选题

1. 基金会中心网推荐系统：从公司与基金会的需求出发
2. 舆情分析：社会大众对基金会的认知及公益基金会社会影响力

设计目标

1. 对志愿者：实现与志愿组织的合理匹配
2. 对志愿组织：帮助志愿组织找到合适的志愿者

我们利用志愿北京共计 66 个和清华、北大有关的志愿组织以及约 13500 名志愿者的服务数据，希望能构建一个匹配志愿组织和志愿者的推荐系统。

数据描述

类似 MovieLen 数据, 一条有效的记录为

volunteer_id	organization_id	joined_times
31240041	3474144	3

上表的含义是甘文灼同学共参加清华大学机械系紫荆志愿者支队 3 次公开招募的活动。

共采集到 17436 条有效记录。

推荐算法描述

我们采用协同滤波的算法。

协同滤波

协同滤波是一种基于收集多个用户对商品的喜好对已知用户未作评价的物品进行量化的一种推荐算法。

我们用 Agnes Johannsdottir 在个人博客中介绍的以下三种方法进行无监督学习：

- ▶ 基于用户的协同滤波
- ▶ 基于物品的协同滤波
- ▶ SVD

基于物品的协同滤波

1. 计算志愿组织的相似度
2. 利用志愿者已参加的志愿组织发布的活动计算对未参加的志愿组织的意愿