开题报告

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 题目 | 基于协同过滤的电影推荐系统开发 | | | |
| 姓名 | 蔡晓聪 | 学号 | | 19720152203406 |
| 研究目标 | 通过实现基于协同过滤的推荐算法，完成一个电影推荐系统，用户可以在系统上为电影评分，后台再根据用户对过往看过的电影评分来建立一个推荐模型。并依据该推荐模型向用户推荐用户未看过并可能感兴趣的电影。通过建立该系统来帮助解决用户选看电影时的信息过载问题，同时帮助更好地发掘一些冷门电影推荐给对其可能感兴趣的用户，挖掘电影的长尾价值。 | | | |
| 研究思路 | 首先需要收集相关的电影数据。可以使用一些类似MovieLens这样的开源数据来作为推荐算法设计时的训练和测试集。之后收集大量推荐系统相关的文章，结合后续预计开发的电影推荐应用系统确定推荐算法的大体设计方案。然后开始编码，编写具体的推荐算法，再使用该推荐算法训练出相应的模型。用测试集对模型进行测试直至达到满意的准确度为止。  等到算法编写完成测试通过。再使用node.js或python等服务器端编程语言融合推荐算法编写出服务端API接口供前端调用。然后编写前端页面完成整个系统的开发。  在前期不具有大量用户行为数据的时候，还需要考虑推荐系统的冷启动问题。拟定使用基于内容过滤算法来完成系统的冷启动问题以积累用户的行为数据。 | | | |
| 研究方法 | 算法研究：主要是编写推荐算法，然后离线对一些开源数据集例如Movielens测试算法的性能。  实践推广：将开发的电影推荐系统应用在小范围推广开，根据实际产生的用户行为数据对应用系统的效果进行在线的测试和评估。  Log Server  (日志记录用户行为)  Frontend  (微信小程序)  行为数据  Algorithm Model  (算法模型)  Web Server  (Node.js) | | | |
| 具体进度安排 | 起讫时间 | | 计划完成内容  (一般可分为资料文献搜索、拟定方案（提纲）、试验或初稿、定稿等阶段) | |
| 2018年12月1日-2019年1月1日 | | 搜集相关文献资料、电影数据 | |
| 2019年 1月2日-2019年1月20日 | | 初步拟定算法设计思路 | |
| 2019年 1月21日-2019年3月1日 | | 编码实现系统 | |
| 2019年 3月 1日-2019年3月31日 | | 论文初稿完成 | |
| 2019年 4月 1日-2019年4月30日 | | 论文定稿 | |
| 签名 | 学生签名：蔡晓聪 2018年 12月 4日 | | | |
| 指导教师签名： 年 月 日 | | | |

注：不足部分可加页。