**附录 2**

**工程实践选题申请表**

**班级：ZF20214 学期：2021年秋季学期**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **分组情况** | | | | |
| **学号** | **姓名** | **组内分工** | **联系方式（手机）** | **Email** |
| ZF2021428 | 张韬 | 主体代码和前端代码实现，相关文档写作 | 15601132130 | hexwar@buaa.edu.cn |
| ZF2021440 | 赵庆国 | 模拟环境代码实现，相关文档写作 | 13998496887 | 308540851@qq.com |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| **选题说明** | | | | |
| **选题名称** | 基于强化学习的文本推荐系统的设计与实现 | | | |
| **研究思路简述** | 本项目旨在实现一个可以对文本网页周期性爬取，对文本数据进行清洗分类后，使用NLP技术生成摘要，再根据用户的浏览偏好，基于主流的推荐算法和强化学习，进行个性化推荐的推荐系统。  **本项目实现内容包括以下推荐系统的核心模块**：  1 数据收集模块，构建推荐模型需要收集很多数据，包括用户行为数据，用户相关数据及推荐文本的相关数据。  2 ETL模块，通过后台管理模块将数据存储到SQLite数据库中。  3 特征工程模块，对原始数据进行降维和特征提取，将数据转换为推荐算法可以学习的特征矩阵。  4 推荐算法模块，推荐系统尝试采用多种机器学习算法、以及强化学习算法来学习用户偏好，在比较不同算法的效果后，从中选取最优者，并基于用户偏好来为用户推荐文本。  5 推荐结果存储模块，周期性的更新用户推荐结果，保存在Redis等易于横向扩展的数据库。  6 Web服务模块，主要作用是当用户在UI上触达推荐系统时，触发推荐接口，为用户提供个性化推荐。  7 模拟环境模块，一个用于训练强化学习的环境。  **推荐系统架技术选型：**  主要使用BeautifulSoup库以及正则表达式来爬取文本信息并进行初步的过滤，将纯文本结果存入SQLite数据库。  使用抽取式（textrank）和生成式（seq2seq）方法分别进行摘要的自动生成，并进行效果比较。  基于K-means等机器学习方法对文本进行分类，根据pagerank方法进行评分，用于在推荐系统冷启动或者探索时，在评分最高的类里选取topk。  使用DQN，A3C和IRecGAN的等强化学习方法对推荐Agent在模块环境中进行训练，以测试不同方法的效果。  使用Flask搭建web服务模块，与用户交互，留存日志做为下次推荐的依据。  **采用项目：**  IRecGAN参考由Xueying Bai, Jian Guan, Hongning Wang于2019年发表的论文Model-Based Reinforcement Learning with Adversarial Training for Online Recommendation  该项目提出了一种基于模型的强化学习解决方案，该方案通过生成性对抗网络对离线策略学习的用户-代理交互进行建模。为了减少学习策略中的偏差，使用鉴别器评估生成序列的质量并重新缩放生成的奖励。 | | | |
| **初步计划** | **按月列出计划进度**  10月 完成技术积累，并实现基本的框架代码。  11月 完成主体和模拟环境的编码，单元测试和系统集成。  12月 完成集成测试和技术文档写作。 | | | |
| **备注** |  | | | |
| **指导教师意见** |  | | | |
| **任课教师意见** |  | | | |

指导教师（签字）： 任课教师（签字）：