第二周——项目周报和项目计划

1、项目信息：

|  |  |  |  |
| --- | --- | --- | --- |
| 项目名称 | HeyCar | | |
| 组名 | 七个老司机 | 小组成员 | 刘昱昊 罗雨辰 方潇月 张嘉麟 陈博楠 田陇宁 王珊珊 |
| 立项日期 | 2018年4月12日 | 结项日期 | 2018年5月8日 |
| 填报日期 | 2018年4月23日 | 填报人 | 王珊珊 |
| 当前所处的阶段 | 补充完善阶段 | | |
| 项目进度  4月12日立项  5月8日结项  前端后端分开工作，进行爬虫、网页设计、数据处理  前后端合并工作，处理接口，调度问题  当前进度 | | | |
| 项目进度状况 | 紧跟进度 | 项目范围状况 | 紧跟进度 |
| 项目质量状况 | 紧跟进度 | 项目总体表现 | 紧跟进度 |

2、本周任务跟踪情况

2.1本周工作任务及完成的情况

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 任务内容 | 成员分配 | 完成百分比 | 备注 |
| 1 | 文本挖掘：根据评论数据集（car comment）计算按时间顺序的一系列情感指数（如：按季度计算情感指数），以及属性标签。（如：油门硬，音响效果好等，尽量从评论中提取出来的标签而不是可以根据属性数据可以直观判断的标签） | 刘昱昊 | 50% |  |
| 2 | 数据爬取：获得一个新的数据集，包括汽车的品牌，车型，若干张图片，一系列属性（如：百公里油耗，加速度，最高速度，颜色等，选择买车的人关心并且可以通过数据直观表示的属性） | 方潇月 陈博楠 | 50% |  |
| 3 | 数据整理：   1. 整理百度指数，处理方法类似情感指数。 2. 整理销量数据，处理方法类似情感指数。 3. 把上述数据集、文本挖掘得到的结果、百度指数、销量数据整合到一起。 4. 周报/文案/ppt制作与展示。 | 王珊珊 | 50% |  |
| 4 | 网页设计：登录界面设计/服务器相关/周报/文案/ppt制作与展示 | 田陇宁 罗雨辰 张嘉麟 | 50% |  |

2.2本周未完成以及取消了的任务

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 未完成的任务 | | | | |
| 序号 | 未完成的任务 | 责任人 | 未完成的原因 | 反思 |
|  |  |  |  |  |

2.3当前项目进度与项目计划的匹配程度

项目整体实际进度与项目计划总体的进度而言，除了个别的功能之外，还算是紧跟进度，完成了对其的基本功能的要求。

2.4项目存在的问题

|  |  |  |
| --- | --- | --- |
| 序号 | 问题 | 原因分析 |
| 1 | 数据爬虫数据爬取不完整 | 平台保存数据要收费 |
| 2 | 数据集合并时表头不齐 | Pandas库的学习不够深入，正则表达式还不会应用 |

3、项目反思

|  |
| --- |
| 项目工作总体反思   1. 多看网上的爬虫、数据处理方法的教程，以免在应用时丈二和尚摸不着头脑。 2. 在今后的项目工作中要有整体的规划，使遇到的问题简单化。 |