

基于协同过滤的在线教育 平台

软件项目计划书

2017-03-10 10:10

目录

1 引言	1
1.1 编写目的	1
1.2 背景	1
1.3 定义	2
1.4 参考资料	2
1.5 标准、条约和约定	3
2 项目概述	3
2.1 项目目标	3
2.2 产品目标与范围	4
2.3 假设与约束	4
2.4 项目工作范围	4
2.5 应交付成果	5
2.5.1 需完成的软件	5
2.5.2 需提交用户的文档	5
2.5.3 须提交内部的文档	5
2.5.4 应当提供的服务	5
2.6 项目开发环境	5
2.7 项目验收方式与依据	6
3 项目团队组织	6
3.1 组织结构	6
3.2 人员分工	6
3.3 协作与沟通	7
3.3.1 项目团队内部协作	7
3.3.2 项目接口人员	7
3.3.3 项目团队外部沟通与协作模式	7
4 实施计划	7
4.1 风险评估及对策	7
4.2 工作流程	8
4.3 总体进度计划	9
4.4 项目控制计划	10
4.4.1 质量保证计划	10
4.4.2 进度控制计划	14
4.4.3 预算监控计划	14
4.4.4 配置管理计划	15
5 支持条件	16
5.1 内部支持	16
5.2 客户支持	16
5.3 外包（可选）	17
6 预算	17
6.1 人员成本	17
6.2 设备成本	17
6.3 其它经费预算	17
7 关键问题	17
8 专题计划要点	18

基于协同过滤的在线教育平台项目计划书

根据《GB8567—88 计算机软件产品开发文件编制指南》中项目开发计划的要求，结合实际情况调整后的《项目计划书》内容如下：

1 引言

1.1 编写目的

本计划书的编写面向所有与本软件开发相关的开发者；

同时为了保证项目开展过程合理有序地组织，保证团队按时且高质量地完成本项目，并且便于各个成员更好的了解本项目的具体流程与实际情况，需要以文件化的形式，把对于在项目生命周期内的工作任务范围、各项工作的任务分解、项目团队组织结构、各团队成员的工作责任、团队内外沟通协作方式、开发进度、经费预算、项目内外环境条件、风险对策、关键内容等内容以书面的方式描述出来，便于达成项目团队成员以及利益相关者之间的共识与约定。

此开发计划用于指导本教育平台项目开发的顺利进行。

1.2 背景

本项目全称为基于协同过滤的在线教育平台，是使用协同过滤算法实现的教育系统。目标用户在于在校大学生群体，提供互联网的在线教育信息支持，同时也满足自学者群体的需求。

随着网络技术的进步，互联网的普及，网络阅读成本降低。在校大学生能够熟练使用计算机，是网络阅读的主要群体。调查显示，在线阅读、手机阅读、手持式阅读器阅读等数字阅读方式开始普及。网络在线阅读排第一位，手机阅读排第二位。数字媒介阅读代替书面阅读。廉价的电子书成为他们主要选择的阅读对象。惯以纸质为载体的报纸、杂志和图书所占比例较小。深受大学生依赖的网络阅读，呈现一种时效性、阶段性、冗杂性的本质，从阅读的形式上分析，这种网络阅读就是一种浅阅读。浅阅读已经成为一种流行，作为网络时代新的阅读方式，浅阅读除了与传统阅读一样获取信息外，更注重追求阅读过程中的视觉快感和心理愉悦，而难以获得实质上的深刻学习。而我们的项目希望将用户引向网络深阅读。

本项目任务由本项目组组长规划并提出；

数据显示，2014 年中国职业在线教育市场规模为 257 亿元，同比增长 18.1%，预计之后几年将保持 20% 左右的速度增长，到 2018 年达 515.5 亿元。我们分析认为，职业在线教育和用户的职业生涯息息相关，用户的付费意愿强，在线教育的形式能够 更好的被该类人群接受。用户在选择机构时，更关注品牌知名

度和机构的后续服务能力，如工作推荐等。

在线教育指的是通过应用信息科技和互联网技术进行内容传播和快速学习的方法。与传统教育机构的教育方式相比，在线教育具有效率高、方便（打破了时空限制，可碎片化学习）、低门槛、教学资源丰富的特点。基于上述特点，再加上“互联网+”推动，在线教育平台兴起，市场需求也与日俱增。

1.3 定义

(1) 在线教育：

或称远程教育、在线学习，现行概念中一般指的是指一种基于网络的学习行为，与网络培训概念相似。

(2) 文本分类：

文本分类用电脑对文本集(或其他实体或物件)按照一定的分类体系或标准进行自动分类标记。

(3) 网络爬虫（Reptilia）：

是一种自动获取网页内容的程序。是搜索引擎的重要组成部分，因此搜索引擎优化很大程度上就是针对爬虫而做出的优化。

(4) 协同过滤推荐（Collaborative Filtering recommendation）：

协同过滤分析用户兴趣，在用户群中找到指定用户的相似（兴趣）用户，综合这些相似用户对某一信息的评价，形成系统对该指定用户对此信息的喜好程度预测。

(5) 朴素贝叶斯算法（Naive Bayesian Model）：

朴素贝叶斯法是基于贝叶斯定理与特征条件独立假设的分类方法。

1.4 参考资料

文档格式要求按照我国 GB/T8567-1988 国家标准和 IEEE/ANSI830-1993 标准规范要求进行。

[1] 刘露，彭涛，左万利，戴耀康．一种基于聚类的 PU 主动文本分类方法[J]．软件学报, 2013, 11:2571-2583.

[2] 平源．基于支持向量机的聚类及文本分类研究[D]．北京邮电大学, 2012.

[3] 杨杰明．文本分类中文本表示模型和特征选择算法研究[D]．吉林大学, 2013.

[4] 李荣陆．文本分类及其相关技术研究[D]．复旦大学, 2005.

[5] 王煜．基于决策树和 K 最近邻算法的文本分类研究[D]．天津大学, 2006.

[6] 苏金树，张博锋，徐昕．基于机器学习的文本分类技术研究进展[J]．软件学报, 2006, 09:1848-1859.

[7] 周平红．我国高等教育信息化水平测评与发展预测研究[D]．华中师范大学, 2012.

[8] 范坤．推进我国教育信息化建设进程的对策研究[D]．华中师范大学, 2004.

[9]牛龙飞. 基于Web的我国教育信息化公共服务平台的设计与实现[D]. 华中师范大学, 2013.

[10] 艾瑞咨询2015中国在线教育行业发展报告

1.5 标准、条约和约定

本项目遵从以下标准:

GB/T 8567-2006 计算机软件文档编制规范 ;

GB/T 19003-2008 软件工程 ;

GB/T 5538-1995 软件工程标准分类法;

GB/T 9385-2008 计算机软件需求规格说明 ;

2 项目概述

在当前资源海量化, 新兴技术日益更新的时代背景下。教师的查找资料, 再学习的任务愈加艰巨, 备课压力陡增。同时, 学生缺少一款组织有序且与时俱进的学习工具。

本项目旨在开发一款基于包含爬虫、数据挖掘、个性教育的在线教育系统。用以将老师从繁重的查询资料、备课任务中解放。通过爬虫获取优质的互联网知识, 进行文本挖掘的处理后, 可以通过个性化推荐模块在既定的教学大纲内向学生推送最合适的内容, 供其挑选与学习、以达到因材施教的效果。

预期将大大减轻教师备课授课的压力, 也能极大程度的激发学生的学习兴趣, 提高学习效率。

2.1 项目目标

团队计划在本学期内完成基于协同过滤的在线教育平台, 为用户提供搜索教育资源, 个性化推荐等功能, 以满足用户的学习需要。

本目标分为以下三个阶段实现:

一、项目需求分析及软件功能设计;

二、功能模块实现及代码整合;

三、软件优化及测试。

项目需求分析方面目标做到能够针对当今线上学习者的市场痛点切入, 找到较好的解决线上轻阅读、无法高效、高利用率地使用互联网的大知识框架, 将互联网的知识信息筛选, 迎合用户的学习上的需求。功能设计方面的目标希望能够做到爬虫模块简洁、有针对性, 后端服务器资源有较好的伸缩性, 核心算法 ML 到达较理想的期望值, 前端界面用户友好。代码整合目标做到整个软件框架的高聚合、低耦合, 代码的重用性高。软件优化及测试目标是在形成产品原型的基础上针对用户反馈进行二次开发, 并且配合测试

尽可能减少代码、程序的错误，更好地满足用户的需求。

2.2 产品目标与范围

产品目标：我们的产品能有效提高学习兴趣与效率，在纷繁复杂的网络世界给同学们一张可靠又有趣学习索引。这是现在的教育所不曾具有的。这种将互联网知识浏览与教育结合起来的新思路很可能是未来教育发展的趋势。

我们团队希望不断改进和推广。在我们的 eMousika 线上教育平台产品进一步完善之后，进一步地吸纳更多的使用用户，让我们的平台为更多的人提供更方便的学习途径，旨在解决每一个寻求知识、教育途径的烦恼。为在校学生或其他使用互联网学习者学习提供尽可能全面的学习帮助，有效提高老师的备课效率和学生的学习效率。

产品范围：我们打造的一款在线的基于 open source learning 的教育平台，教育是万千大计，每一个渴望知识的人都理应得到最好的教育和最便捷的资源，我们的教育平台就针对对于知识渴求，但是又苦于线上互联网信息的冗杂、线下信息更新的缓慢的苦恼的学习者，他们中包括苦于备课的老师、缺乏自主学习的在校学生，渴望教育平台和二次学习资源的救治人员，互联网线上自学者，缺乏教育人力物力的偏远山区的孩子等等。同时我们的产品主要针对互联网使用者推出，应用于 PC Web 端。

2.3 假设与约束

小组成员的分工必须严格遵照计划；

小组成员要有团队意识，遵守团队规章；

预算必须遵守，背的无视计划；

系统测试人员全程跟进项目的进度，对于板块的测试和拼合起到重要作用；

小组必须一周要进行一次项目的讨论和整理报告

2.4 项目工作范围

产品范围界定：具有简介扁平化设计的教育平台前端网页。

工作范围界定：需求分析：对当今教育行业的不足提出质疑并且调查市场漏洞和需求量。

市场分析：竞品优势对比，将本项目与当今热门线上教育平台对比，分析优劣

软件设计：分析和思考设计最优化的软件结构，规定模块间的接口。

软件框架搭建：实现软件的框架结构。

代码：主要分为 Web 前端代码、Web 后段服务器、数据爬虫模块核心代码、聚类分析

核心代码

测试：对于整个系统进行测试反馈，从而进行二次优化。

2.5 应交付成果

2.5.1 需完成的软件

源程序、数据库对象创建语句、可执行程序、支撑系统的数据库数据、配置文件、第三方模块、界面文件、界面原稿文件、声音文件、安装软件、安装软件源程序文件等等。

2.5.2 需提交用户的文档

用户操作手册：详尽描述软件的实用功能、性能、用户界面，使用户对如何使用该软件得到具体了解，为操作人员提供软件各种运行情况的有关知识，特别是操作方法的具体细节。

2.5.3 须提交内部的文档

软件维护手册：主要包括软件系统说明、程序模块说明、操作环境、支持软件的说明、维护过程的说明、便于软件的维护。

2.5.4 应当提供的服务

免费咨询：客户可以在工作时间向技术人员提出问题并获得解答。

软件维护：获取软件使用中的问题，提供不定程序。

2.6 项目开发环境

硬件环境：windows PC，Android 手机，Linux 服务器

操作系统：windows，CentOS，Android

数据库系统：MySQL

开发工具：Android Studio，PyCharm，Dreamweaver

开发语言：Java（Android），Python，html+css

网络环境：校园网

2.7 项目验收方式与依据

项目验收方式：通过本学期期末课内答辩形式，由洪梅老师及助教进行验收，同时可请使用 Android 手机的同学进行运行，已进行用户验收。

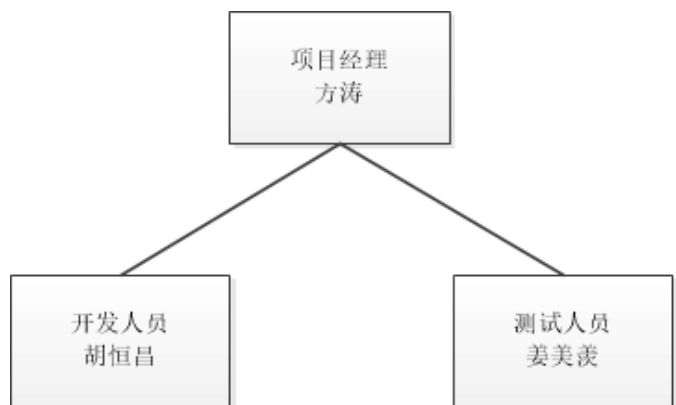
项目验收依据：老师，助教，以及体验过该软件的同学认可。

3 项目团队组织

3.1 组织结构

完成本项目所需角色：项目经理，程序开发人员，测试人员。

当前项目成员一共三位，包括一名组长、两名组员。



本项目对各个角色个人能力有一定要求：项目经理需有项目开发管理的经验、较为广泛的程序设计知识与经验；程序开发人员需在网络爬虫、机器学习等领域有较深的认识与开发经验；测试人员需具备一定的软件测试的经验。

3.2 人员分工

姓名	技术水平	角色	工作描述
方涛	中等水平，在网络爬虫编写上具有一定经验	项目经理	对项目进程管理，保证项目的高质量顺利实施，任务划分
胡恒昌	较高水平，具备机器学习、网络爬虫以及前端开发的经验	程序开发人员	开发网络爬虫脚本以及系统后端
姜美菱	较高水平，具备前端设计经验以及软件测试经验	测试人员、文档编写	开发系统前端界面以及系统测试

3.3 协作与沟通

3.3.1 项目团队内部协作

本节说明在项目开发过程中项目团队内部的协作模式和沟通方式、频次、沟通成果记录办法等内容。

本团队内部协作主要通过定期召开项目开发会议，会议召开频率为一周两次以上，并由方涛负责记录会议上成员提出的建议以及问题，会议召开的目的在于增加项目开发成员之间的沟通与协调，以及在会议中加深对需求的理解；

同时项目成员间通过建立 QQ 群组方式保持日常联系，通过互留电话方式满足日常开发所需的开发成员之间的沟通；

项目组组长通过邮件方式将任务分配给各个组员。

3.3.2 项目接口人员

负责本项目同用户的接口人员是姜美羨，主要负责与客户沟通并且收集整理客户对本系统的需求，并且定期更新需求，在软件开发后期负责软件的测试与调试。

3.3.3 项目团队外部沟通与协作模式

本项目开发过程中，项目组成员可通过电话、QQ 或邮件形式联系客户或项目经理与接口人员；

同时本项目组保证以每周两次以上的频率与客户保持联系，交流项目进展以及更新完善客户需求。

4 实施计划

4.1 风险评估及对策

（1）进度风险：工程存在时间限制，可能无法再规定的时间内完成计划内的所有任务；

出现概率：一般

造成的影响：不能及时地交付项目成果，同时也会影响项目答辩。

采取的对策：每周队内进行一次进度报告，项目经理确保团队的进度在计划之内或之上。

（2）技术风险：团队成员在数据库开发方面没有足够的经验，需要在项目进程中同步学习，不排除遇到技术瓶颈的问题。

出现概率：一般

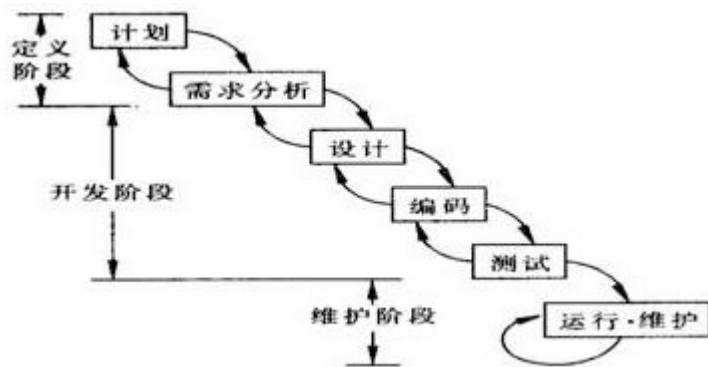
造成的影响：项目某个模块的阻塞导致整个项目的进度落下

采取的对策：在动手之前先获得足够的知识或请教老师，对于技术上进行攻克。












4.2 工作流程

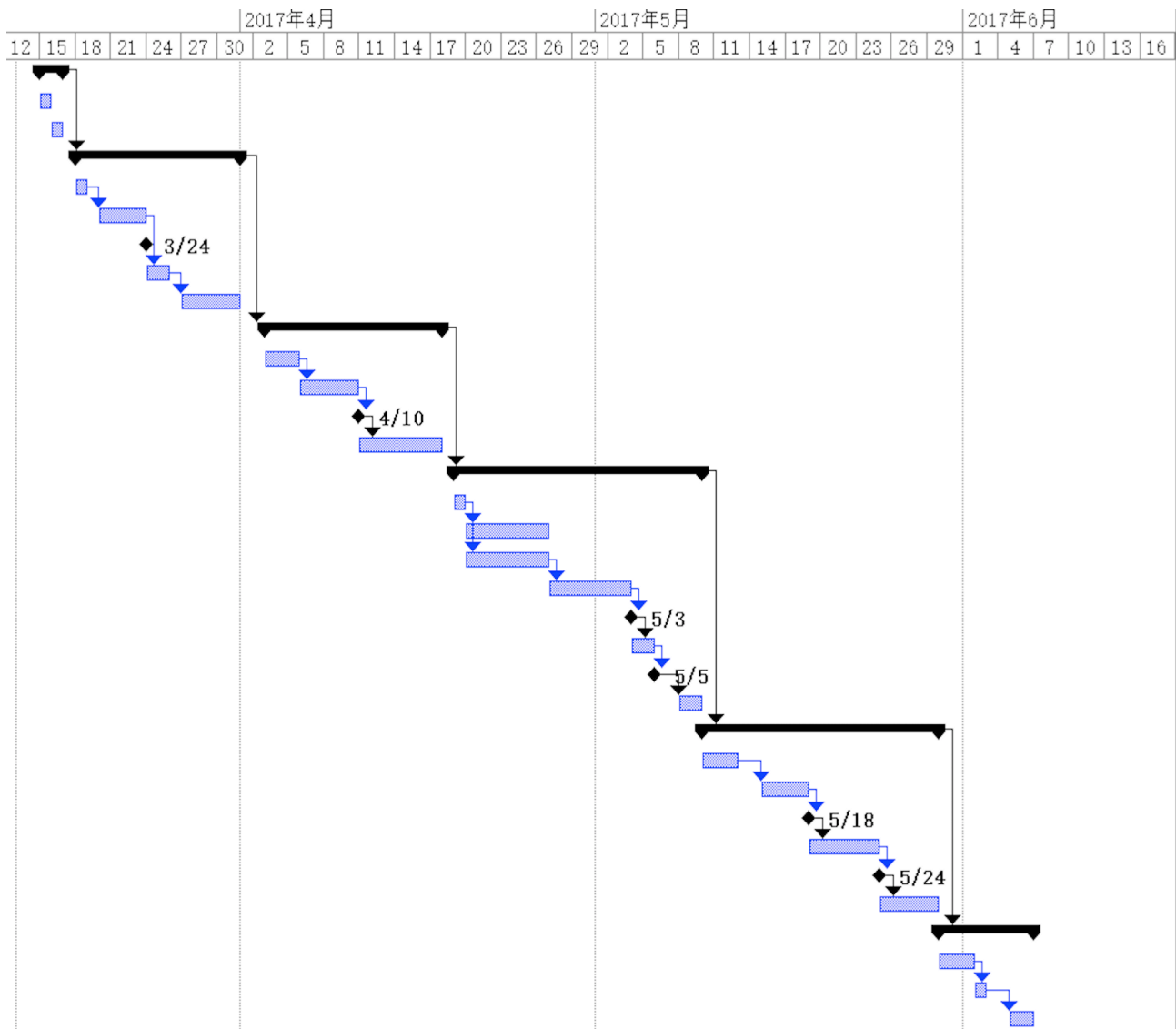
采用瀑布

开发过程是通过设计一系列阶段顺序展开的，从系统需求分析开始直到产品发布和维护，每个阶段都会产生循环反馈，因此，如果有信息未被覆盖或者发现了问题，那么最好“返回”上一个阶段并进行适当的修改，项目开发进程从一个阶段“流动”到下一个阶段。



4.3 总体进度计划

标识号	任务名称	工期	开始时间	完成时间
1	 项目启动	2 工作日	2017年3月15日	2017年3月16日
2	 小组分工	1 工作日	2017年3月15日	2017年3月15日
3	 可行性分析	1 工作日	2017年3月16日	2017年3月16日
4	需求分析	12 工作日	2017年3月18日	2017年3月31日
5	 明确需求阶段的任务并分工	1 工作日	2017年3月18日	2017年3月18日
6	 获取需求	4 工作日	2017年3月20日	2017年3月23日
7	 check point	0 工作日	2017年3月24日	2017年3月24日
8	 绘制用例图	2 工作日	2017年3月24日	2017年3月25日
9	 编写需求规格说明书	5 工作日	2017年3月27日	2017年3月31日
10	系统设计	11 工作日	2017年4月3日	2017年4月17日
11	 设计系统的功能模块	3 工作日	2017年4月3日	2017年4月5日
12	设计系统的数据库并绘制对象关系	3 工作日	2017年4月6日	2017年4月10日
13	check point	0 工作日	2017年4月10日	2017年4月10日
14	编写设计文档	5 工作日	2017年4月11日	2017年4月17日
15	编码	15 工作日	2017年4月19日	2017年5月9日
16	 明确任务并分工	1 工作日	2017年4月19日	2017年4月19日
17	 网页前端实现	5 工作日	2017年4月20日	2017年4月26日
18	朴素贝叶斯算法编码	5 工作日	2017年4月20日	2017年4月26日
19	协同过滤算法编码	5 工作日	2017年4月27日	2017年5月3日
20	check point	0 工作日	2017年5月3日	2017年5月3日
21	搭建数据库	2 工作日	2017年5月4日	2017年5月5日
22	check point	0 工作日	2017年5月5日	2017年5月5日
23	系统整合	2 工作日	2017年5月8日	2017年5月9日
24	软件测试	14 工作日	2017年5月10日	2017年5月29日
25	单元测试	3 工作日	2017年5月10日	2017年5月12日
26	集成测试	4 工作日	2017年5月15日	2017年5月18日
27	check point	0 工作日	2017年5月18日	2017年5月18日
28	系统测试	4 工作日	2017年5月19日	2017年5月24日
29	check point	0 工作日	2017年5月24日	2017年5月24日
30	编写测试分析报告	3 工作日	2017年5月25日	2017年5月29日
31	软件部署	6 工作日	2017年5月30日	2017年6月6日
32	编写用户手册	3 工作日	2017年5月30日	2017年6月1日
33	客户培训	1 工作日	2017年6月2日	2017年6月2日
34	项目结题报告	2 工作日	2017年6月5日	2017年6月6日



4.4 项目控制计划

4.4.1 质量保证计划

4.4.1.1 引言

4.4.1.1.1 目的

本计划的目的在于对所开发的软件规定各种必要的质量保证措施，以保证所交付的软件能够满足项目预定需求，能够满足本项目总体制定的且经老师批准的该软件系统需求规格说明书中规定的各项具体要求。

4.4.1.1.2 定义

本计划用到一些术语的定义按 GB/T 11457 和 GB/T 12505

4.4.1.1.3 参考资料

GB/T 11457 软件工程术语

GB 8566 计算机软件开发规范

GB 8567 计算机软件产品开发文件编制指南

GB/T 12504 计算机软件质量保证计划规范

GB/T 12505 计算机软件配置管理计划规范

4.4.1.2. 管理

4.4.1.2.1 机构

在本软件系统整个开发期间，必须成立软件质量保证小组负责质量保证工作。软件质量保证小组由软件开发小组成员组成，由项目组长担任质量保证小组组长。软件质量保证小组必须检查和督促本计划的实施。

4.4.1.2.2 任务

软件质量保证工作涉及软件生存同期各阶段的活动，应该贯彻到日常的软件开发活动中，而且应该特别注意软件质量的早期评审工作。因此，对新开发的各个子系统，要按照 GB 8566 与本计划的各项规定进行各项评审工作。评审与检查的目的是为了确保在软件开发工作的各个阶段和各个方面都认真采取各项措施来保证与提高软件的质量。在软件开发过程中，在软件开发过程中，应该进行以下三次评审：第一次评审软件需求、概要设计、验证与确认方法；第二次评审详细设计、功能测试与演示，并对第一次评审结果复核；第三次是功能检查、物理检查和综合检查。

4.4.1.2.3 职责

在软件质量保证小组中，各方面人员的职责如下：

组长全面负责有关软件质量保证的各项工作和有关阶段评审、项目进展报表检查以及软件验收准备等第三方中的质量保证工作；

组员负责测试复查和文档的规范化检查工作，协助组长开展各项软件质量保证活动。

4.4.1.3 文档

为确保软件的实现满足需求，应该编写以下文档：

软件需求规格说明书；

软件设计说明书；

软件测试计划；

软件测试报告；

用户手册；

源程序清单；

项目实施计划；

项目开发总结。

4.4.1.4. 标准、条例和约定

文档标准：符合《质量体系文件》中定义的模块格式和内容；

软件产品标准：必须通过单元测试、集成测试和验收测试；

阶段结束的条件和标准：通过阶段评审，并产生评审报告；

缺陷等级：严重、一般、建议；

优先级：高、中、低。

4.4.1.5. 评审和检查

就整个软件开发过程而言，至少要进行软件需求评审、概要设计评审、详细设计评审、软件验证和确认评审、功能检查、物理检查、综合检查以及管理评审等八个方面的评审和检查工作。把评审分成三次进行。在每次评审之后，要对评审结果作出明确的管理决策。下面给出每次评审应该进行的工作。

第一次评审：对软件需求、概要设计以及验证和确认方法进行评审；

第二次评审：对详细设计、功能设计与演示进行评审，并对第一次评审结果进行复核。

第三次评审：进行功能检查、物理检查和综合检查。这些评审应在集成测试阶段结束后进行。

4.4.1.6. 工具、技术和方法

在软件的开发过程中，都应该合理地使用软件质量支持工具、技术和方法。这些工具主要有两种：软件配置管理工具和文档辅助生成工具与图形编辑工具。

4.4.1.7. 媒体控制

为了保护计算机程序的物理媒体，以免非法存取，意外损坏或自然老化，软件系统都必须设立软件配置管理人员，妥善管理和存放各个子系统及其专用支持软件的媒体。

4.4.1.8. 记录的收集、维护和保存

在于协同过滤的在线教育平台开发期间，要进行各种软件质量保证活动，准确记录、及时分析并妥善保存有关这些活动的记录，是确保软件质量的重要条件。由组长负责收集、汇总与保存有关软件质量保证活动的记录。

4.4.2 进度控制计划

进度控制依旧项目实施计划表，定期对项目各任务完成情况进行评估，小组讨论后给出参考意见，督促项目各部分的完成。

每周进行一次组内讨论，汇报项目进度，项目经理保证项目进度在计划之内或者计划之上。如果进度落后于计划，则需要加紧项目工作。

4.4.3 预算监控计划

小组成员各部分开销由项目经理统一管理并根据需要分配限额，本着节约的原则尽量减少个本分的开支。

4.4.4 配置管理计划

4.4.4.1. 引言

4.4.4.1.1 目的

本文档目的在于对基于协同过滤的在线教育平台进行软件配置管理，提高软件质量，降低软件开发成本。

本文档内容主要参考研发中心相关的 ISO 程序和制度文档，并在此基础上整理成适合本项目的软件配置管理，为项目经理、配置管理员及相关人员提供日常的配置管理操作步骤。

4.4.4.1.2 定义

软件配置管理：是在项目开发中，标识、控制和管理软件变更的一种管理。配置管理的使用取决于项目规模和复杂性以及风险水平。软件的规模越大，配置管理就显得越重要。

配置管理员：项目组中负责配置管理工作的角色，该角色可以兼职。在某一开发阶段通过评审或某一质量检查点通过审核后，配置管理员负责统一添加或修改相关文档的最新有效版本以及审批人签字。

配置标识：（Configuration Identification）对软件项目在开发过程中的资源进行标识，以便识别。

配置检查：（Configuration Audit）对软件配置管理过程中的行动进行检查。

4.4.4.1.3 参考资料

《研发中心配置管理制度》 《产品的标识与可追溯性程序》 《开发手册》

4.4.4.2. 管理

SCM 管理信息描述了组织和个人在项目的 SCM 活动中的责任和权限。

4.4.4.2.1 组织

组织结构包括技术和管理两方面，计划中的并将要被实施的配置管理活动必须被描述。计划必须说明以下问题：

- a) 在项目中，参与或对所有配置管理活动负责的组织单位； b) 在项目结构中，组织单位的功能角色；
- c) 各组织单位之间的关系。

组织单位可能包括供应商和客户，主承包商和分承包商，或在组织中的其他团队。组织图表、功能和关系的状态的补充可以成为表现信息的有效途径。

4.4.4.2.2 配置管理责任

本项目配置管理计划对组织单位 SCM 活动的分配必须被详细说明。对每个在 SCM 中列出的活动，必须提供执行活动的组织单位或工作标题的名字。可以使用矩阵清楚地表示和说明以上定义的组织和 SCM 功能、活动和任务之间的关系。

对于那些在项目中为执行 SCM 活动而建立的任何审查委员会或特定组织，计划中必须描述有关它们的以下方面：

- 目的和目标； b) 成员和联系； c) 有效期； d) 权限范围； e) 操作程序。

4.4.4.3 软件配置管理活动

4.4.4.3.1 配置标识管理

i. 文档

根据配置管理计划和配置库中的文档清单，配置管理员要检查需要提交的文档是否都按时提交，文档

数目是否符合，文档的标识、命名以及版本等是否符合程序规定。

ii. 程序

所有属于该项目的程序、分程序、模块和程序单元，都要按照由项目组和配置管理员制订的软件系统的命名约定的规定来标识。

要求所有模块的源代码都需记录模块编号，且模块编号在整个系统中是唯一的。模块编号在系统设计完成之后，由项目组和配置管理员共同根据系统设计进行编制。

iii. 基线

所有属于本项目及其各子系统的各类基线，首先要按照计划书、软件需求规格说明书、软件项目详细分析设计说明书的规定确定其技术内容，

配置管理员负责在项目开发的每一个里程碑处、每一个阶段性的版本发布时负责为整个配置库设立书签，划定配置管理基线，并以文档的方式记录下这些书签的定义。

(1) 配置库控制

在项目开发和实施的整个过程中，配置管理员应根据配置管理计划及管理规则对配置库应进行管理和控制。配置管理员负责检查项目组成员使用配置库是否正确。包括是否及时检入最新版本、是否添加了注释、是否及时更改配置状态，是否存在项目组成员修改了不属于自己负责的配置项，项目组成员是否完成了自己负责的配置项的检入，测试版本的构造是否从配置库中取出等。

在项目的各个开发阶段，应建立起各阶段各子系统的软件开发库，同时建立起想对应的有关该系统及其子系统的软件受控库。在每个阶段结束或里程碑，需让各子系统提交相关的产品并送入软件受控库，由配置管理员统一管理，以后再有对产品的变更需求，应按照正常的变更程序来控制并检查相关的变更文档。当全部开发工作结束，需建立起软件产品库，将所有可交付的产品都送入软件产品库。

(2) 配置的检查 and 评审

配置的检查 and 评审可通过研发中心配置管理制度的审核内容来进行检查。配置管理员应配合研发中心产品管理部定期对项目进行配置管理的审核。在审核过程中，提供所需要的配置管理计划及相关资料，在项目开发结束后，需提交所有关于项目的软件配置库。

4.4.4.4. 工具、技术和方法

4.4.4.4.1 软件环境

(1) Python 2.7

(2) MySQL 数据库

(3) Anacoda

(4) Microsoft Project

4.4.4.4.2 硬件环境

(1) 服务端：

1GB 内存，1MB/s 带宽，500GB 内存

(2) 普通个人计算机

v. 对供货单位的控制

本项目无供货单位。

5 支持条件

5.1 内部支持

- i. 项目成员拥有丰富的项目经历，对于时间管理有较好的把控度。
- ii. 课堂中有经验丰富的老师和可靠的同学能够共同相互讨论沟通，帮助我们解决问题。
- iii. 腾讯云稳定、可靠
- iv. 队员统一使用最稳定的 Python 2.7 版本

5.2 客户支持

客户主要承担部分软件功能测试及体验工作，并提出改进意见。完成时间在项目各功能模块实现测试之后的一周内获取用户反馈。

且该项目面向用户广泛，容易获得客户的反馈。每一位同学都可以做我们的支持客户。

5.3 外包（可选）

无。

6 预算

6.1 人员成本

本项目组成员工作量分配（预计月数）：

小组成员	工作量（月）
方涛	3
胡恒昌	3
姜美羨	2

本项目主要工作任务：

任务名称	所需人员	时间（月）
需求确认	3	0.5
需求分析	3	0.5
软件建模	2	0.5
编码实现	2	1
软件测试	1	0.5
软件部署、维护	2	0.5

无劳务费。

6.2 设备成本

本项目使用现有设备：

Windows 个人电脑 1 台，安卓手机（Android4.4）设备 1 台

拟购置的设备：

腾讯云 CentOS 服务器（1M/s 带宽，1GHz 处理器），价格为 65 元/月。

6.3 其它经费预算

无其余经费。

7 关键问题

关键问题	等级	难点
文本分类准确率	高	本项目组使用朴素贝叶斯算法对文本进行分类，由此可能带来算法分类准确率不高问题； 另一方面由于爬虫从网络获取的文本信息可能有大量冗余信息，导致分类不准
教学资源的获取	高	使用爬虫从网络获取的信息可能包含大量无用字符，需进行额外的工作进行过滤； 此外，高频率地使用爬虫可能导致账号被某些网站识别禁用导致无法继续工作
数据传输时间可能过长	中	因服务器性能（带宽）有限，可能

		导致用户使用客户端访问服务器、应用程序或网页端调取数据库数据时存在较长延迟
推荐算法地准确性	中	在系统运行之初,由于缺乏大量用户的使用与评分数据,推荐算法会因缺乏足够样本集导致无法推荐或推荐不准问题
数据更新	低	对系统中文本内容的更新依赖于爬虫目标网站的内容更新

8 专题计划要点

专题计划也就是因为项目的需要在本文档之外独立建立的计划,本节说明本项目开发中需要制定的各个专题计划的要点。专题计划可能包括分合同计划、分项目计划、项目团队成员培训计划、测试计划、安全保密计划、质量保证计划、配置管理计划、用户培训计划、系统安装部署计划。

8.1 项目团队成员培训计划

为小组成员高效地完成本项目,现拟自 2017 年 3 月 25 日开始,至 2017 年 4 月 2 日止,在为期一周时间内,对小组成员进行项目所涉及相关开发知识的培训,由方涛进行知识讲解。

培训内容	培训时间
Python 基本用法	3 月 25 日
爬虫示例	3 月 27 日
朴素贝叶斯算法	3 月 29 日
推荐算法	4 月 2 日

8.2 测试计划

拟在 2017 年 5 月下旬对本系统进行各项测试,测试人员由小组成员姜美菱担任,主要测试内容有:
 安卓客户端的稳定性;
 教育平台网页端;
 服务端数据处理与反馈。

8.3 质量保证计划

本项目开发过程中由开发小组组长负责项目进程的监管,保证严格按照软件开发过程中的各项步骤具体实施,包括可行性分析,需求分析,项目开发计划等,保证本软件质量。

8.4 用户培训计划

本项目组计划在软件具体部署前一周,由胡恒昌对客户进行软件使用培训。

8.5 配置管理计划

项目开发小组人数: 3 人;
 组长: 方涛;

组员：胡恒昌，姜美羨

8.6 安全保密计划

无。

8.7 合同计划

无。