



北京航空航天大学
B E I H A N G U N I V E R S I T Y

聊天机器人 ChatterBot 软件项目计划书

E 组

郭维泽 王元玮 张延钊 吕江枫 王云杰 李书缘

2020 年 3 月 15 日

目 录

1. 项目概述.....	3
1.1 背景.....	3
1.2 简介.....	3
2. 项目内容.....	4
2.1 项目目标.....	4
2.2 具体内容.....	4
3. 团队组织.....	5
3.1 团队成员.....	5
3.2 人员分工.....	5
3.3 协作方式.....	5
4. 项目计划.....	6
4.1 预计计划和任务量.....	6
4.2 甘特图.....	6
5. 参考资料.....	6

1. 项目概述

1.1 背景

随着机器学习尤其是自然语言处理技术的不断发展和突破，语音助手和聊天机器人已经走进了人们的生活。为了更好的研究聊天机器人的相关技术，并让大家更方便的与聊天机器人开展对话。我们选取了 ChatterBot 这个项目，进行相应的改进，并尝试部署到微信上。

1.2 简介



ChatterBot 是一个 Python 库，在 GitHub 上开源。它。ChatterBot 可以使用一系列机器学习算法，轻松生成对不同类型的用户输入的合理响应。ChatterBot 运行简单，可以轻松创建聊天机器人并自动与用户进行对话。其在 GitHub 上拥有 9600 个 star，是最受欢迎的聊天机器人项目之一。

其主要模块有以下五个部分：

- (1) **Trainer:** 负责根据不同的训练器执行不同的训练方法。
- (2) **Input/output adapter:** 负责预处理输入或输出文本。
- (3) **Logit adapter:** 负责根据不同的策略搜索正确的答案。
- (4) **Storage adapter:** 负责提供连接到各种存储系统的接口。

(5) ChatterBot: 对所有模块的整合。

大致的执行流程是：用户给定一个输入、对输入进行预处理、利用不同的策略寻找可能的答案、对候选答案进行排序得到最终答案、将最终答案返回给用户。

2. 项目内容

2.1 项目目标

改进 ChatterBot 的内部算法，实现更好的聊天功能。采用更丰富的训练样本，同时增加对中文的支持。改进 ChatterBot 的数据接入方式，增加更多类型的数据库接入服务。接入微信个人号，部署在阿里云平台上。最终目标是完成一个可以对外提供聊天机器人服务的微信个人号。

2.2 具体内容

(1) 改进 logic adaptor，使用深度学习模型。主要考虑目前效果较好的 LSTM 模型、BERT 模型等。

(2) 查找中文语料库，并将语料库导入到现在的模型中。同时增加新训练样本，尝试提高模型表现。

(3) 优化 storage 模块，目前仅支持 MongoDB，要增加对更多数据库接入的支持。

(4) 接入微信前端，创建注册微信个人号，将项目部署到阿里云，并成功运行。

3. 团队组织

3.1 团队成员

本团队成员分为组员和组长。组长与组员沟通协调，安排具体事宜、分配相应的任务。而组员则需要根据组长的安排要求，完成任务。

- 组长：郭维泽
- 组员：李书缘、吕江枫、王元玮、王云杰、张延钊

3.2 人员分工

郭维泽：负责项目管理和软件功能测试。

李书缘：负责微信个人号前端的需求分析和开发。

吕江枫：负责软件功能测试。

王元玮：负责整体设计和 logic adaptor 模块的需求分析与改进。

王云杰：负责 storage 模块的需求分析与改进

张延钊：负责 logic adaptor 模块的需求分析与改进

3.3 协作方式

简单问题在微信群里及时讨论，重大或复杂问题需要商定时间通过腾讯会议开会讨论决定。

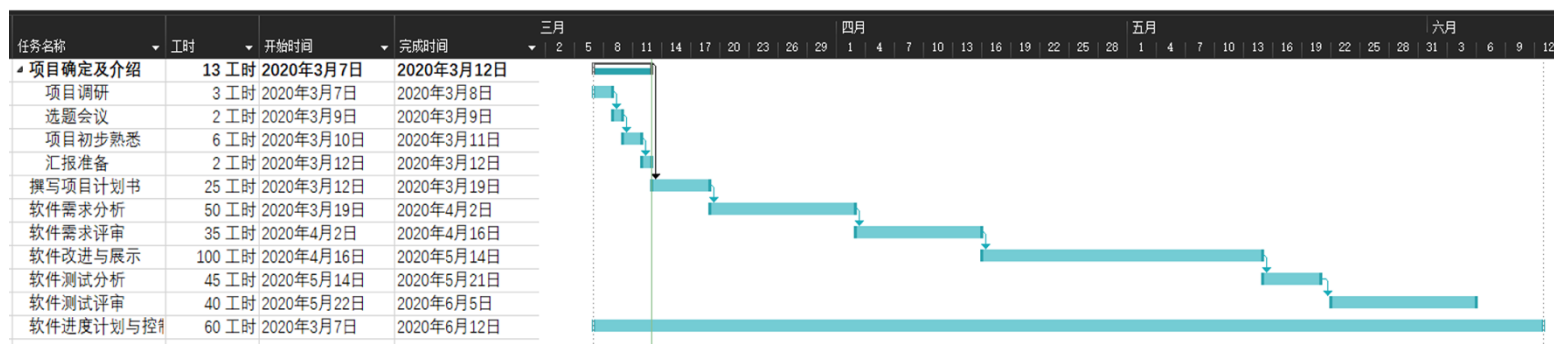
代码同步在 GitHub 上进行，文档同步在腾讯为文档上进行，项目进度管理使用 Microsoft Project 软件。

4. 项目计划

4.1 预计计划及任务量

任务名称	任务量	开始日期	结束日期
项目确定及介绍	13小时	3月7日	3月11日
项目调研	3小时	3月7日	3月7日
选题会议	2小时	3月9日	3月9日
项目初步熟悉	6小时	3月10日	3月11日
汇报准备	2小时	3月12日	3月12日
撰写项目计划书	20--30小时	3月13日	3月19日
软件需求分析	40--50小时	3月20日	4月2日
软件需求评审	30--40小时	4月3日	4月16日
软件改进与展示	80--100小时	4月17日	5月14日
软件测试分析	40--50小时	5月15日	5月21日
软件测试评审	30--40小时	5月22日	6月5日
软件进度计划与控制	60--70小时	3月7日	6月12日

4.2 甘特图



5. 参考资料

(1) ChatterBot 参考文档: <https://chatterbot.readthedocs.io/en/stable/>