**SJTU公司**

**立项建议书**

项目名称：“智读”有声书

√ 多人项目 □单人项目

联系人学号和姓名： 517021910589 姬浩迪

联系电话：15201975051

电子邮箱：575650866@qq.com

**2019 年 7 月**

1. 项目的必要性

有声阅读的起源很早，其历史可以追溯到清末民初时的说书这一传统文化·活动。进入移动互联时代以来，电子阅读已经被人们广泛接受。与此同时诞生的，还有另一种阅读形式——有声阅读，即我们常说的听书。当今市场上的有声读物一般分为两种，一种是基于语音合成技术的机器朗读，另一种是人工录制的有声书。前者广泛嵌入于各大电子书 app 中，可以低成本而灵活地向用户提供听书服务。但是，目前语音合成技术合成的声音往往缺乏感情，语调平淡，机器本身无法理解文本内容，朗读时照本宣科、毫无感情，给人的体验并不好。后者是将文字经说书人之口演绎出来，往往是包含了说书人的简介与情感，而且多有后期的音效和背景音乐添加以渲染气氛，是真正的在单纯的文字之上增加了一个声音的维度，给人以良好的听觉体验。不过，这种听书方式不够灵活，听众不能随心所欲地选择要听的章节，而且有声读物的录制也需要大量的人力物力。如何在提高用户的听书体验的同时有效降低制作成本，成为限制有声读物软件发展的一大问题。因此，有必要开发一款可以充分发挥语音合成技术和人工录制音频长处的软件。

1. 项目目标和特性

### 1、本项目的主要建设目标

本项目基于深度学习的语音合成识别与语言情感分析，通过运用文本情感识别算法判断文本包含的情绪，并赋予与之对应的背景音乐，完成一款基于人工智能的有声书产品。

### 2、本项目的特性(功能需求与非功能需求)：

功能需求： 基本需求：

1. 用户注册登录与用户个人信息管理(优先级高)
2. 上传有声书文本与有声书制作和阅读(优先级高)
3. 有声书文本感情赋予与背景音乐配置(优先级高)
4. 提供用户分享有声书的平台，可以收听自己和他人的作品(优先级高)

进阶需求：

1. 通过音频生成有声书(优先级低)
2. 基于人工智能的推送，根据用户偏好进行推荐(优先级低)

非功能需求：

1. 安全性：保证用户信息不会错误，被盗取或恶意利用(优先级中)
2. 可维护性：当软件出现问题时，可以快速找到 bug 的位置，及时修复(优先级中)
3. 系统响应时间小于 5s(优先级中)

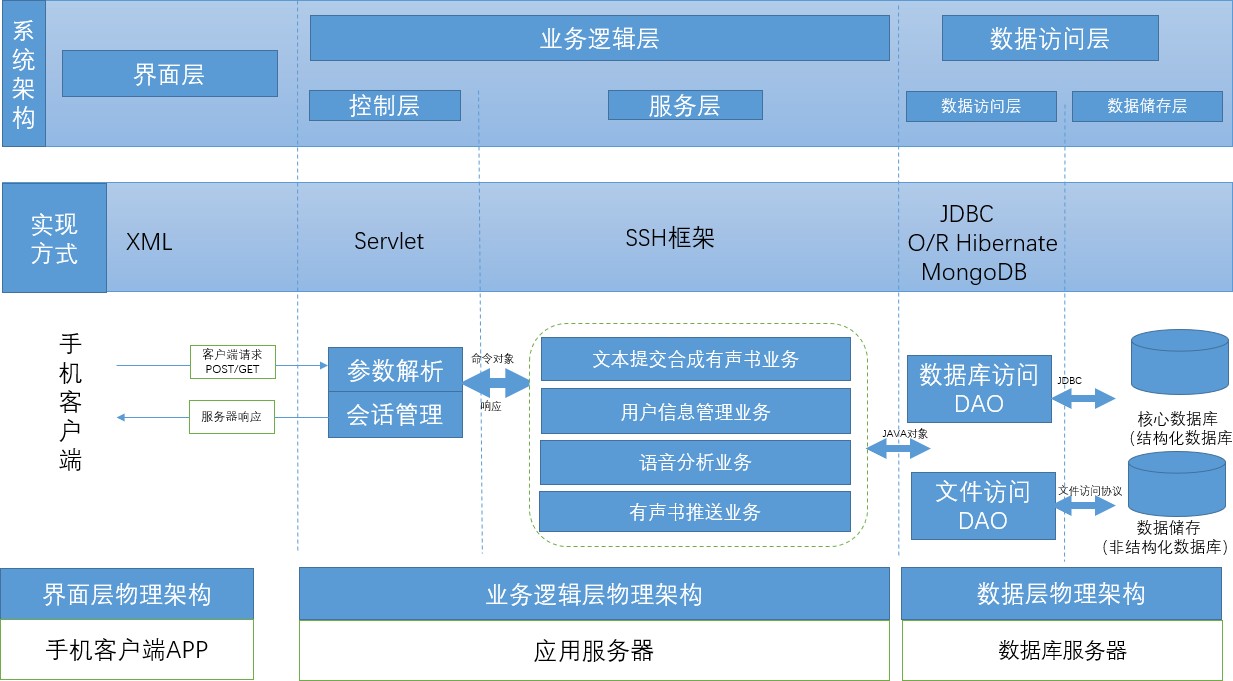
### 3、项目的创新点与优势

**１*.*本项目赋予机器带有情感的阅读。**通过文本情感分析为问本提供背景音乐和上下文音效。相较于无情感阅读的有声书能够满足用户对有声书更好的听书体验的需求。

***2.*本项目可以根据用户生成的文档自动产生对应的有声读物，生产效率高，成本低可以快速占领市场。**由于只需要用户提交文本之后 app 就可以根据用户提交的文本就可以生成有声书。可以在当前该读物人工录制的有声书尚未产生时凭借其超高的生产效率快速占领市场。由于有声书属于单次消费物，用户在收听一次之后很少会再重复收听一次。因此凭借其超高的生产效率可以快速占领市场，在有限的" 蛋糕"里，分得最大的一块。

1. 项目技术方案

**本项目的软件架构示意图如下**：



本项目软件架构图

采用“三层”架构模式

系统主要分为界面层，业务逻辑层，数据访问层三层。其中界面层主要用于呈现软件界面，传递客户请求，接收逻辑层响应；业务逻辑层主要用于处理客户请求等业务逻辑；数据访问层用于提供访问数据的接口并储存用于合成有声书的音频文件和用户的个人数据。

本项目使用 powerdesinger 工具进行架构图类图的建模设计，使用 java 编程语言进行开发，通过 android studio 工具进行编程，采用 SSH 框架，开发方法为基于UML 的面向对象方法

1. 项目风险分析和里程碑计划

风险分析：

1）进度风险

由于小组内同学对于人工智能了解并不是十分深入，需要花大量时间进行相关知识的学习，加上开发时间紧张，可能会存在无法按照预计进度完成相应阶段的开发任务的风险。

## 2）技术风险

小组成员对本软件功能实现所需要的算法之前并没有过多了解，存在技术能力难以 完成软件开发的风险。由于需求中有制造符合根据上下文环境的声音音效，可能由 于训练不充分导致软件对于某写字段识别错误。如：当文本中出现"此时风和日丽"， 软件却会配以狂风的声音。

里程碑计划：

1. 第一次迭代(7.1 ~~7.7)

主要任务：

1、分析需求。

2、撰写《立项建议书》

3、架构设计

4、完成前端 app 界面的设计。

5、完成用户注册功能

6、完成用户登录功能

7、完成对用户信息的管理与维护。

8、系统测试

提交成果:

1、第一次迭代计划

2、第一次迭代评估报告

3、具有前端和注册登录功能和完整界面的 app

4、《立项建议书》

5、架构文档和分析设计模型

主要风险：

**技术风险**：注册用户激活可能存在一定的问题。

1. 第二次迭代(7.8~~7.19)

主要任务：

1、完成通过文本制作人工合成的无感情语音内容。

2、复现 hanLP 算法，实现根据文本内容选择合适的背景音乐和音效。

3、建立常用音乐库。

4、系统测试

提交成果：

1、具有用户上传文本和制作有声书功能的 app

2、常用音乐库文件

3、第二次迭代计划

4、第二次迭代评估报告

1. 第三次迭代（7.20~~8.2）：

主要任务：

1、实现用户有声书的收听与分享

2、允许用户搜索、浏览、收藏、收听其他用户制作的有声书

3、实现通过音频生成有声书功能

4、复现基于CNN神经网络的文本分类算法，完成用户偏好推荐。

5、通过部分用户进行软件测试，收集反馈意见，根据反馈意见进行调试，修改。

主要风险：

**技术风险**：可能存在算法复现困难的问题

提交成果：

1、基本完善的 app

2、第三次迭代计划

3、第三次迭代评估报告

1. 第四次迭代（8.3-9.8）：

主要任务：

1、新功能开发：夜间模式、定时退出

2、软件优化

3、验收准备

提交成果：

1. 最终的 app
2. 第四次迭代计划
3. 第四次迭代评估报告
4. 最终提交内容中未提交的部分
5. 项目预期成果

项目验收时交付以下成果：

《立项建议书》

《迭代计划》

《迭代评估报告》

《SRS 文档》和用例模型

《软件架构文档》和分析设计模型

《测试用例》和《测试报告》

《项目总结报告》 源代码和可执行代码

演示视频文件（包括安装、运行、功能等） 演示 PPT