**情绪分析网站——心情日记**

**一、项目信息**

我们的项目是情绪分析网站——心情日记。以日记的方式每天收集自己的情绪，用户可选择中文或者英文的文本输入，也可以选择智能语音输入。后台对文本进行分析评测，显示出当前用户此次记录时的情绪值，大体分为消极与积极两大类，同时，在分析页面也会根据用户情绪进行相应的文章推荐，帮助用户调节心情。

**二、项目详细分工情况及完成情况：**

**房友江**：英文文本情绪分析算法，文章推荐（工作完成）

**翟媛媛：**中文文本情绪分析算法，中英文语音识别算法（工作完成）

**庄玲：**整体网站框架搭建，网站页面元素设计搭配（工作完成）

**陈伟良：**中英文语音输入算法，整体网站框架搭建（工作完成）

**三、项目开发/运行环境：**JetBrains PyCharm Community Edition 2019.1.2

**四、项目系统架构/关键技术**

**整体网站框架搭建部分（包含网站主py）：**

①使用flask的Flask（app主运行）、render\_template（渲染每个页面html）、request（接收前端传送数据）、redirect（页面重定向跳转）框架模板；

②使用flask\_bootstrap 的Bootstrap前端开发框架进行页面html识别渲染；

③使用datetime模块获取系统时间，作为情绪日记和情绪分析的标签；

④使用random模块进行相关情绪的文章随机推荐

**网站页面元素设计部分**：

html，css页面搭建

**中英文语音输入部分：**

①使用pyaudio初始化pyaudio并开关数据流进行语音输入 ；

②使用wave处理音频，将录音相关内容（声道数，取样频率，量化格式等）写入wav文件；

**中英文语音识别部分：**

①使用baidu-aip调用百度语音识别API进行pcm格式的语音识别

②利用FFmpeg调用系统命令行将wav格式音频转变成pcm格式。

**英文文本情绪分析部分：**

使用nltk操作朴素贝叶斯分类器进行英文文本情绪分析

**中文文本情绪分析部分：**

①利用jieba对中文文本进行分词

②利用numpy将python的列表对象创建一个矩阵，得到情感值；np.mean，求取均值；np.std对矩阵求取标准差;np.sum可以对行，列，或整个矩阵求和

**五、项目方法/资源**

**英文文本情绪分析部分：**

朴素贝叶斯步骤如下：

(1)收集数据：可以使用任何方法。

(2)准备数据：需要数值型或者布尔型数据。

(3)分析数据：有大量特征时，绘制特征作用不大，此时使用直方图效果更好。

(4)训练算法：计算不同的独立特征的条件概率。

(5)测试算法：计算错误率。

(6)使用算法：一个常见的朴素贝叶斯应用是文档分类。可以在任意的分类场景中使用朴素贝叶斯分类器，不一定非要是文本。

**中文文本情绪分析部分：**

(1)读取文本信息，对文本进行分句

(2)查找句子的情感词，记录积极或者消极，以及位置

(3)往情感词前查找程度词，找到就停止搜索。为程度词设置权值，乘以情感词

(4)往情感词前查找否定词，找全部否定词，数量为奇数则乘-1，偶数乘1

(5)判断分句结尾是否有感叹号，有则往前找情感词，并且情感词 \*2

(6)计算文本所有分句的情感值，有list记录

(7)计算并记录所有分句的情感值

(8)计算文本的句子积极情感均值，消极情感均值，及方差

(9)算出文本的积极情感值占积极与消极情感值总和的百分比

**中英文语音识别部分：**

(1)使用FFmpeg调用系统命令行将wav格式音频转变成pcm格式。

(2)调用百度API并设置相关参数进行pcm格式音频的语音识别

**整体网站框架搭建部分：**

(1)使用flask的Flask作为网站主程序框架

(2)使用flask的render\_template模板渲染每个页面html文件，并传送页面中相关参数

(3)使用flask的request方法接收前端传送数据

(4)使用flask的redirect模板进行页面重定向跳转

(5）使用flask\_bootstrap 的Bootstrap前端开发框架进行页面html识别渲染；

(6)使用datetime模块获取系统时间，作为情绪日记和情绪分析的标签

(7)使用random模块获取文章文本列表随机内容，进行相关情绪的文章随机推荐

**网站页面元素设计部分：**

(1)html文件设计登录、注册、首页、情绪分析页面的排版设计

(2)css使html页面多彩化

**中英文语音输入部分：**

(1)实例初始化pyaudio，设置相关参数（声道、数据流块、取样频率、量化格式等）并打开数据流，开始语音输入

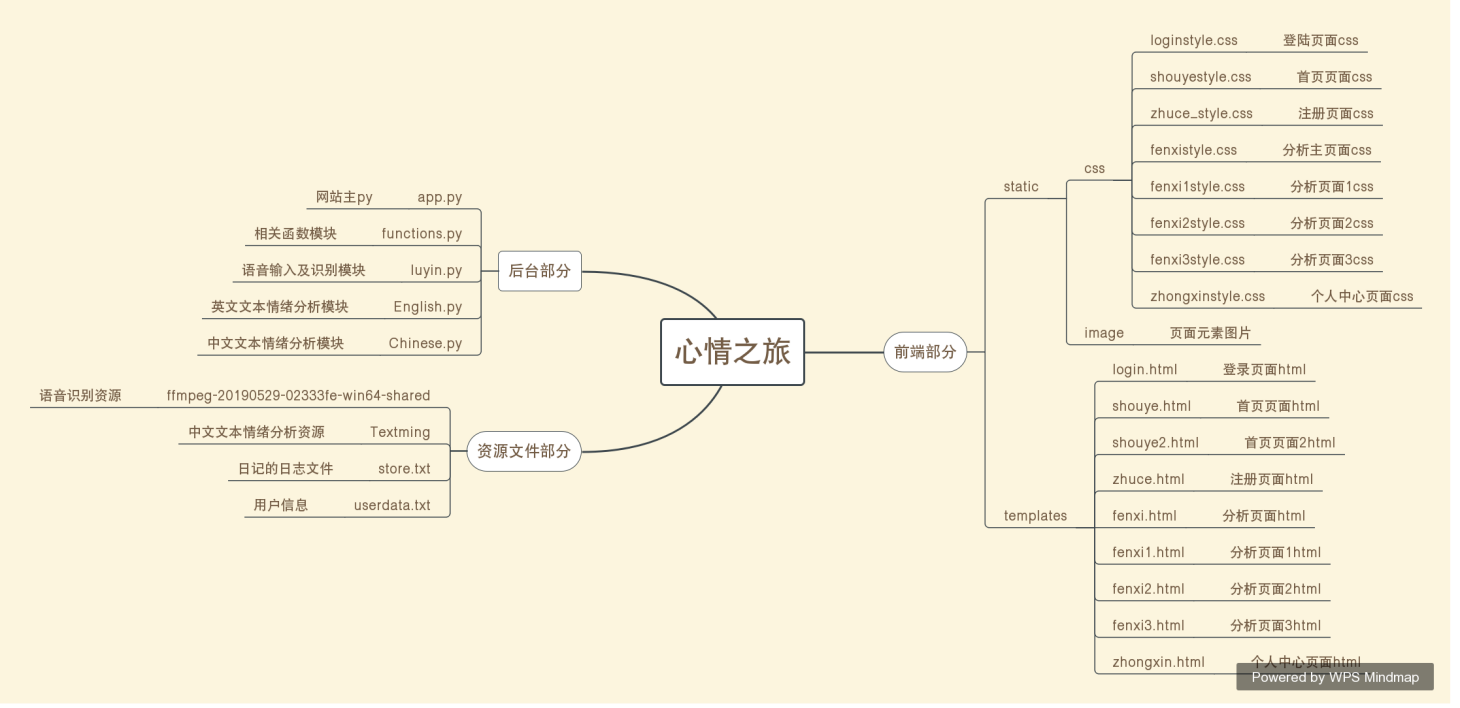
(2)关闭数据流，关闭pyaudio，停止语音输入

(3)使用wave.open新建音频文件，将录制的音频内容写入音频文件

(4)设置音频文件中音频内容相关参数（声道、取样频率、量化格式等）

**六、项目备注**

**（1）项目文件结构**

****

**（2）项目截图**

（采用pycharm运行app.py）

**登录界面：**

****

**注册页面（未联网）**

****

**首页**

****

****

**未输入文本时的分析页面**

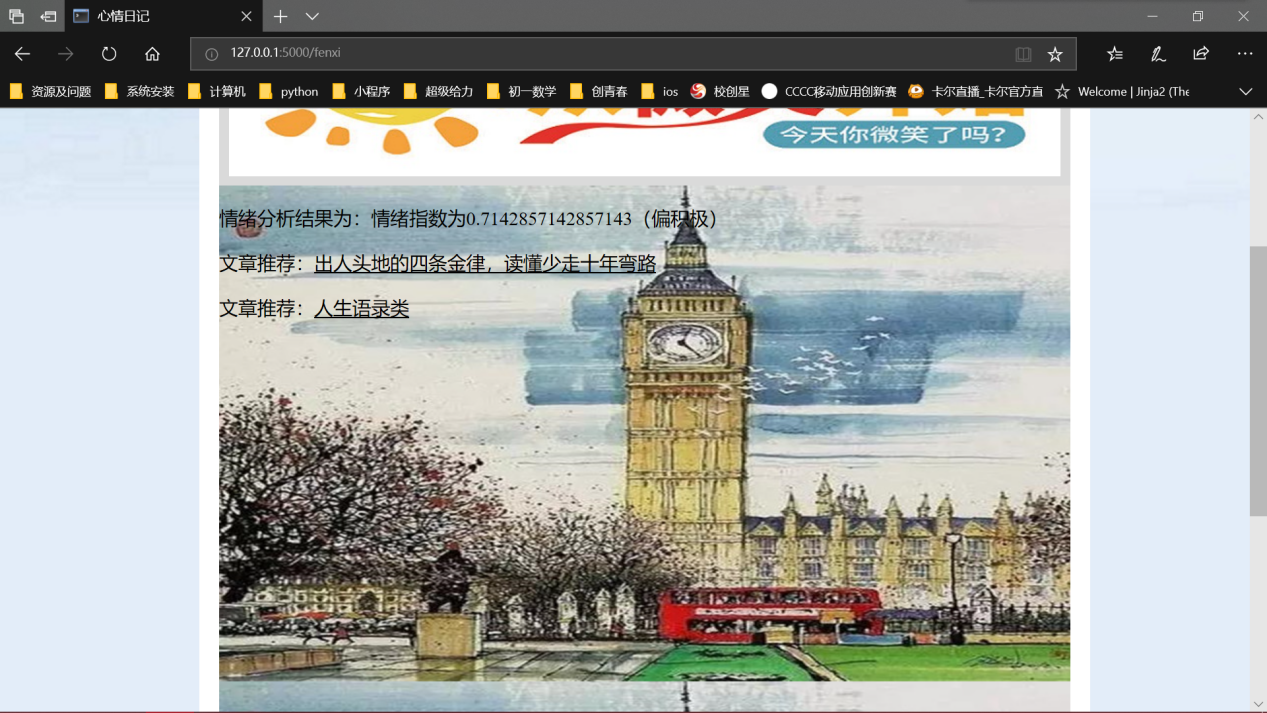
****

**个人页面（未联网）**

****

**输入“我今天超级开心”，点击保存分析，下面是分析结果页面**

****



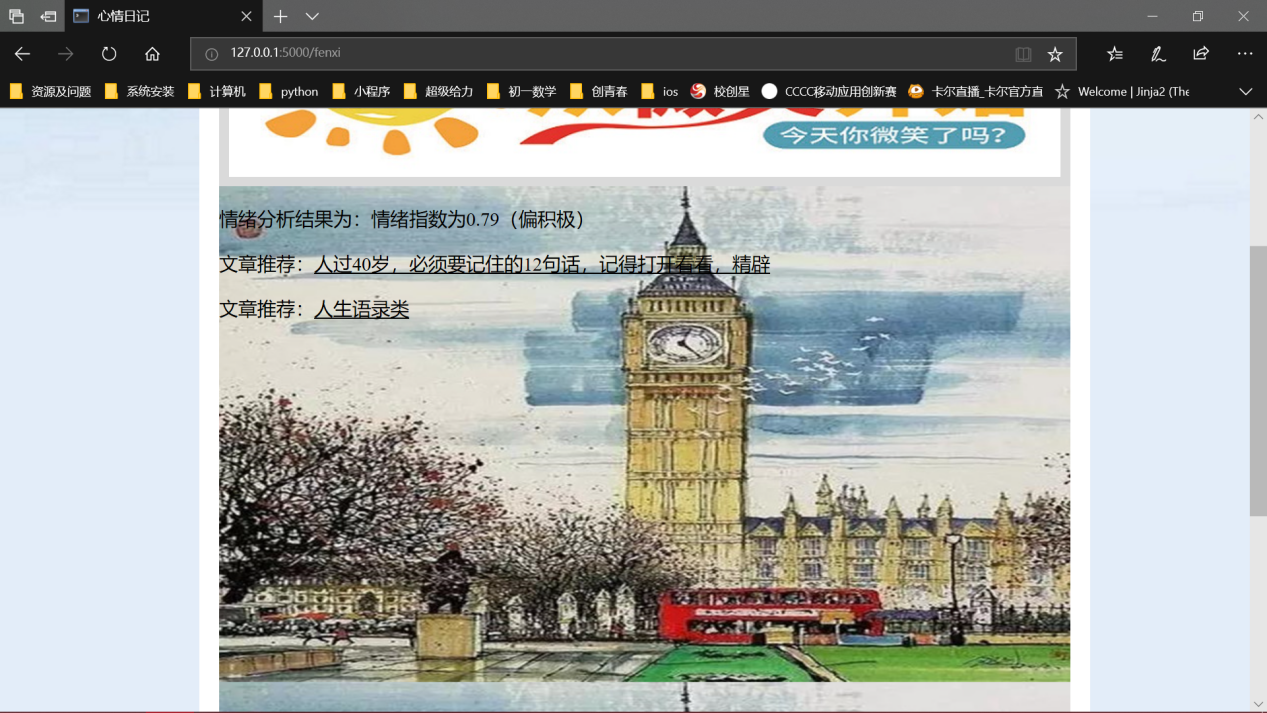
**输入“我今天好难过”，点击保存分析，下面是分析结果页面**

****



**输入“I am very happy”,点击保存分析，下面是分析结果页面**

****



**点击语音输入，开始说话，时间为7秒（可设置）**

**时间截止则文字识别到文字框**



**点击保存分析**

