

2019-03~19~2019-03~26 week4进度总结

苏昭帆

1. 调研：常见的语音信号特征（MFCC、共振峰频率、基音周期、过零率、能量等）；部分特征集（ComParE features set, 6373d）（GeMAPS, 62d）（GeMAPS, 88d）；语音情感识别算法（常见的机器学习分类器，如SVM等）（声谱图+CRNN）（手工特征+CRNN）；离散情感数据库（Belfast英语：anger、sadness、happiness、fear、neutral）（EMO-DB德语，7种情感；FAU AIBO儿童德语）（EMO-DB德语，7种情感；FAU AIBO儿童德语）（ACCorpus汉语）等。更多调研结果参见（#url address）
2. 实验：获取中文情感数据集（#name），包括六个情感维度，共计1100条训练样本和100条测试样本；利用sklearn提供的api提取wav音频文件的MFCC特征，利用SVM算法建立模型，准确率约为30%；数据集及代码地址如下：（#url address）

刘娅璇

1. 分析并绘制基于“升金风暴”的业务流程图，流程图地址如下：（#url address）

邹笑寒

1. 文字转语音：PocketSphinx支持离线和中文文，但识别正确率较低；在Mac系统上能正常运行，但在Windows系统上出现了配置问题，目前还未解决；要提高正确率的话可以考虑自建语言包；之后将继续用在线API进行行行试验。
2. 语音转文字：pyttsx3支持离线和中文文，能做到棒读给定文文字。
3. 代码整合至项目代码中，地址如下：<https://github.com/XLab-Tongji/SemanticAnalysis/tree/Speech-and-Text>

刘岳涵

1. 代码工作：在python代码中实现了基本的对话流逻辑，建好了基础分类器类（基于模版匹配）。代码已经整合至项目代码中。

郭辉

1. 调研：查找可以将声音转换为特征向量的工具，选定pythonspeechfeatures包和librosa音频处理库，对二者的使用方法进行了总结，并进行了简单的尝试。代码地址如下：（#url address）；可视化语音分析开源程序，https://github.com/fedden/umaptsneembedding_visualiser；
2. 阅读论文：《语音情感识别分类算法研究综述》和《语音情感识别综述》，简单了解语音情感的方法。

论文上传至github(#url)