**语音质检项目产品构思**

1. **问题描述**

随着网络服务的兴起，越来越多的网络技术公司会使用人工坐席来进行语音服务，这样就不可避免的会为了提高质量而进行服务质量检查。而用人工来进行质量检查的问题有：  
a)耗费时间太多  
b)耗费太多劳动力成本  
c)很有可能会因为质检员分神等原因使检测效率变低

1. **产品愿景和商业机会**

**定位：**

为设有语音客服服务的企业提供针对客服的语音质量检测评估，使企业的语音客服服务更加完善。

**商业机会：**

* 设有语音客服的企业很多对语音质检有需求，客户来源确定
* 当前语音质检技术还不够成熟，企业应用率不高

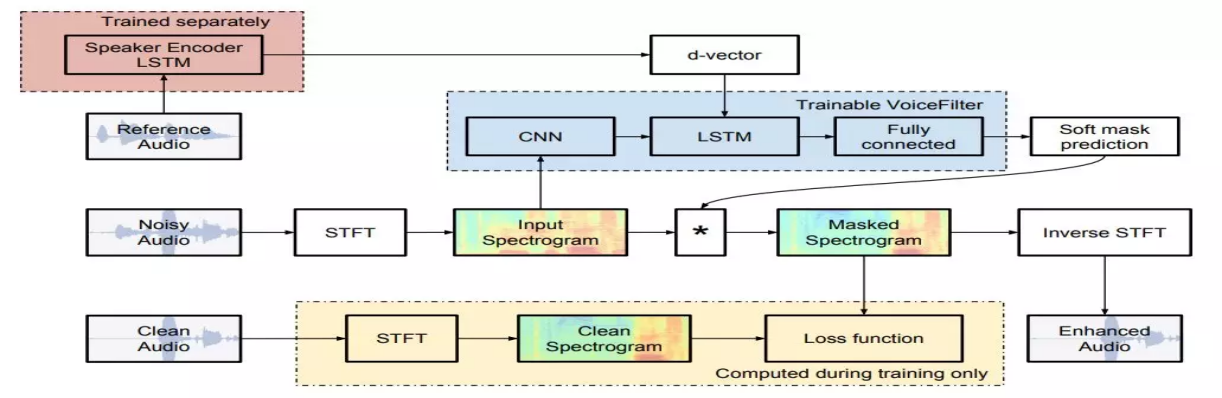
**商业模式：**

* 技术销售

1. **用户分析**

语音质检项目主要服务两类人：  
**1、使用人工坐席的企业经理（简称商家）：**  
 愿望：希望给用户更好的人工座席服务，并且节省提高服务质量的成本  
 痛点：使用人工进行通话质量检测服务耗时大，效率低，成本高  
**2、使用人工坐席的客户：**  
 愿望：能通过人工服务让问题得到解决，人工服务的态度良好，让自己感到舒服  
 客户范围：所有有能力使用手机并懂得拨打电话的人  
**3、打分观念：**只要大致满足自己的需求并且服务舒服，大部分会愿意给高分

1. **技术分析**
2. **话者分离**



主要由两部分组成:声纹识别模块(红色) 声音过滤模块(蓝色)

声纹识别模块：它使用的网络是三层的递归神经网络加上GEE损失，输入1600ms的mel话谱，最后将三个d-vectrorL2正则化后取平均值得出一个具有说话者声纹特征的向量

声音过滤模块:输入d-vector向量和混合有噪声的音频，由八个卷积神经网络和一层递归神经网络组成，这其中除了最后的输出层的激活函数使用sigmoid函数之外，其余的隐藏层全部使用ReLu函数。将d-vector和时间轴重复之后与卷积层的结果合并输入到递归神经网络中最后得出单纯的参考人的说话音频

1. **语音转写**

**语音识别技术采用开源ASRT**

**仓库地址**：https://github.com/nl8590687/ASRT\_SpeechRecognition

ASRT 是一套基于深度学习实现的[语音识别](https://www.jiqizhixin.com/articles/021102)系统，全称为 Auto Speech Recognition Tool，由 AI 柠檬博主开发并在 GitHub 上开源（GPL 3.0 协议）。本项目声学模型通过采用[卷积神经网络](https://www.jiqizhixin.com/articles/021102)（CNN）和连接性时序分类（CTC）方法，使用大量中文语音数据集进行训练，将声音转录为中文拼音，并通过[语言模型](https://www.jiqizhixin.com/articles/021102)，将拼音序列转换为中文文本。基于该模型，作者在 Windows 平台上实现了一个基于 ASRT 的语音识别应用软件它同样也在 GitHub 上开源了。

**系统流程：**

特征提取：将普通的 wav 语音信号通过分帧加窗等操作转换为神经网络需要的二维频谱图像信号，即语谱图。

声学模型：基于 Keras 和 TensorFlow 框架，使用这种参考了 VGG 的深层的卷积神经网络作为网络模型，并训练。

CTC 解码：在语音识别系统的声学模型输出中，往往包含了大量连续重复的符号，因此，我们需要将连续相同的符号合并为同一个符号，然后再去除静音分隔标记符，得到最终实际的语音拼音符号序列。

语言模型：使用统计语言模型，将拼音转换为最终的识别文本并输出。拼音转文本本质被建模为一条隐含马尔可夫链，这种模型有着很高的准确率。

**模型使用：**

test\_mspeech.py

1. **关键字检索**

* 关键词检索根据应用场景不同分为两种
* 模式固定，文本不定

场景：事先定义一些模式(比如说黄色关键词)。对每个给定的文本，确定是否含有这些关键词

* 文本固定，模式不定

场景：已知网络上的网页库，给定一个字符串，寻找含有这些字符串的网页。

* 我们的需求是在客服与用户的对话中检索一些业务用语，所以我们的检索是“模式固定，文本不定”类型。
* “模式固定，文本不定”检索的算法研究：
* 单模式匹配(仅仅有单一模式文本)

KMP算法

* 多模式匹配

AC(Aho-Corasick, KMP)的升级版本

1. **资源需求估计**
2. **人员**

产品经理：依据本产品的商业背景和定位，吸取已有产品的成熟经验，结合项目特点和用户特征，设计符合用户应用场景的产品。

技术人员：架构和实现产品，同时确保产品质量效率和正确率，以及改变应用场景后产品的适应性。

用户的客户代表：有较多投诉经历或与客服沟通经历，帮助分析客服的语音态度用词问题等；

用户代表：主要负责监管客服的人员，帮助分析客服的话术及用词；

1. **资金**

产品验证阶段前暂无需要。完成产品验证后，需要资金集中快速完成商家扩充和宣传推广；

1. **设备**

一台云端PC服务器；

1. **设施**

10平米以内的固定工作场地；

1. **风险分析**

|  |  |  |  |
| --- | --- | --- | --- |
| 编号 | 事件描述 | 根本原因 | 风险类型 |
| R1 | 客户有担忧 | 需要进行通话录音，可能会泄露客户私人信息 | 信息风险 |
| R2 | 语音分析不到位 | 技术不够成熟 | 技术风险 |
| R3 | 企业应用率不高 | 销售推广不到位 | 商业风险 |
| R4 | 产品完成周期过长 | 技术人员有可能不能及时到位 | 人员风险 |

1. **收益分析**

根据与企业的合作情况进行定位