5.项目研究调查阶段，我们团队查阅并研读了国外相关研究方向的学术论文，深入了解了技术细节，并为下一阶段的开展提供思路。

6．小组利用python爬虫技术进行原始数据的收集工作。这是代码实现和收集数据的格式。

7.人声分离阶段，为了取得更好的分声效果，团队采用卷积神经网络模型，并实现了CNN损失函数。同时，团队对卷积神经网络进行网络结构优化来实现模型结构调整，并对调整后的模型进行训练。

8.数据预处理阶段，团队将人声分离后得到的音频文件转化为非压缩波形文件。之后，对音频进行分帧操作。

9.预处理结束后，需要利用HMM模型进行处理。其中，利用到Baum-Welch和Viterbi算法具体实现。同时，团队对Viterbi算法进行了完善和优化。调整后得到DNN-HMM模型，并对其进行训练。