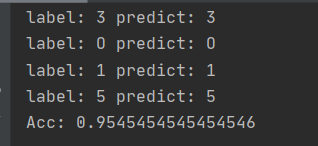
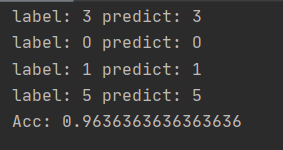
作业1：

Q：完善DNN代码，并基于该DNN实现11个数字识别

dnn = DNN(429, 11, 128, 1)



dnn = DNN(429, 11, 200, 1)



作业2：基于kaldi下该数据集的标注脚本，梳理基于DNN-HMM的语音识别系统的流程，其有哪些步骤，每一步的输入、输出，步骤间的相互关系等

1. 先进行GMM-HMM模型训练，viterbi算法得到每个特征帧被识别到每个状态的概率分布，将特征帧作为DNN的输入，将概率分布作为DNN要预测的标签，送入DNN网络进行有监督学习。
2. 生成Fbanks特征

“steps/make\_fbank.sh”：

输入：“data/fbank/”，

输出：fbank特征；

“steps/compute\_cmvn\_stats.sh“：

进行倒谱均值归一化

3.帧交叉熵训练——把帧分到对应的三因素状态

“steps/nnet/train.sh”：

这里用的是mini-batch梯度下降法,更新速度快。参数选择Mini-batch=256, 使用Sigmoid隐层单元，Softmax输出单元和全连接层。学习率为0.008。

“steps/nnet/decode.sh”：

对word和phone进行解码

4.训练dnn的序列辨别MPE/sMBR

“steps/nnet/align.sh”：

将‘data’对齐到transition-ids序列上

“steps/nnet/make\_denlats.sh“：

为MMI/MPE/sMBR 训练创建词图

5.重复进行3轮MPE训练

“steps/nnet/train\_mpe.sh”：

输入对齐文件和词图，使用SGD梯度下降法，最大化从参考的对齐中得到的状态标签的期望正确率。