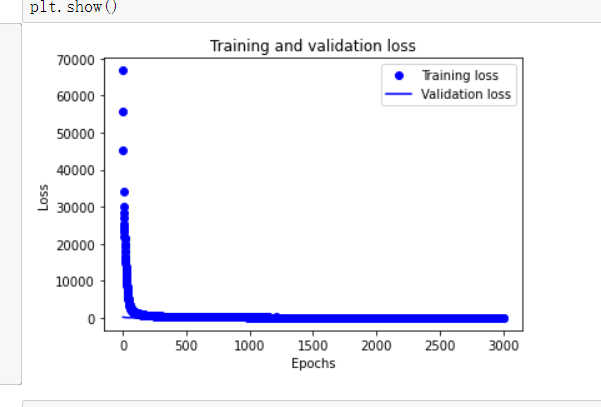
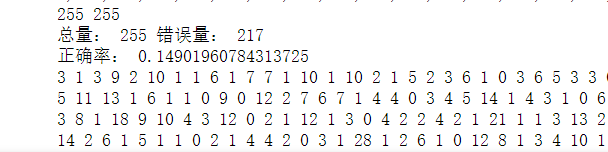
# 本周进展

## 进展1：完成时域序列机器学习的样本获取、模型搭建、模型训练和测试。样本结构如下：以16000的采样率对每个样本采样0.1s，每个样本为长度为1600的音频幅值序列，每个样本的幅值根据自身样本情况进行归一化，频率统一以最高值进行归一化或者不进行处理。模型搭建采用LSTM时间序列模型。进行了3000轮训练的结果截图如下:



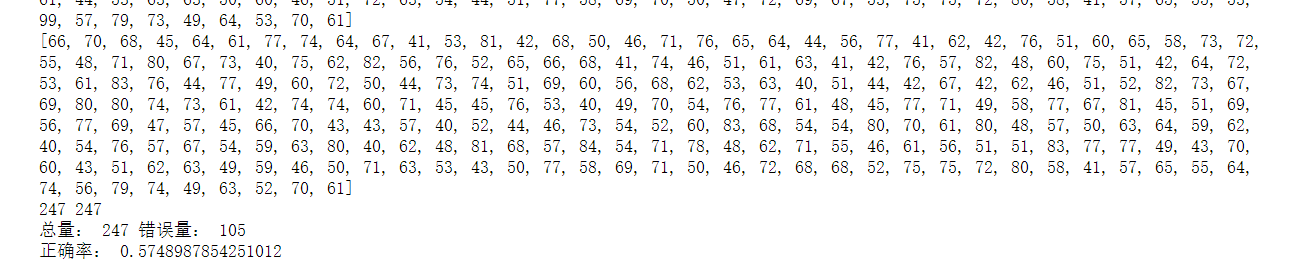


## 进展2： 重新优化修改2D、3D张量机器学习模型，找到了正确的预测率

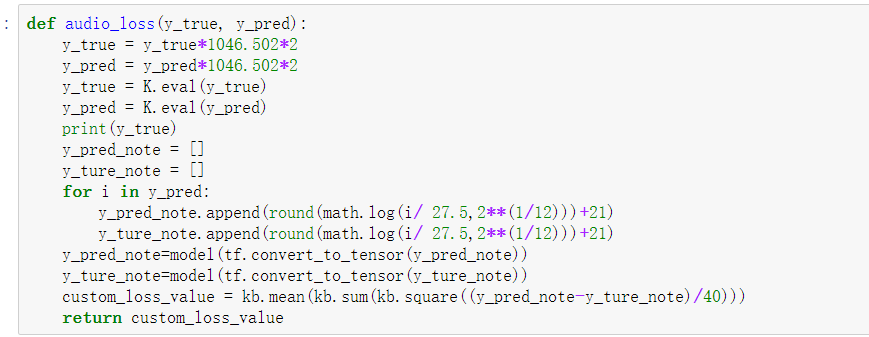
3D：

## 

2D：



## 进展3： 尝试编写对应的自定义损失函数：



# 小组分工完成情况：

冯韵菱：

1.尝试对于2D,3D的模型构建自定义损失函数

2.检查模型正确率过低原因，重新检查数据，发现组员输入的采集数据未进行裁剪，重新输入数据，得到了正确的输出结果。

许宏涛：

1. 搭建了时域的机器学习模型，输入样本为长度为1600的音频幅值序列，输出为音高类别。选择CNN网络结构，损失函数定义为交叉熵，优化器同多分类模型。
2. 《python深度学习》第六章 循环神经网络学习

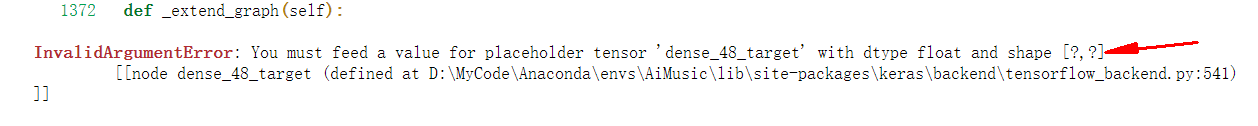
张靖鸿：

1. 完成样本数据的采集与处理，对每个样本采样1600个样值点，并进行归一化处理，保存为json格式文件，方便后续的处理。
2. 完成《python深度学习》部分的学习

# 遇到的问题及解决方法

之前的2D、3D张量模型效果不尽人意（正确率只有不到百分之20），于是我们重新从输入开始检查，发现输入数据是并未进行音频裁剪的预处理的数据，于是我们重新对所有音频进行了裁切的预处理，再归一化，送入模型，得到了正确的预测率分别为51%、57%，但还说未能达到我们的预期，于是我们思考，可能是损失函数未能对应上，因为对于回归模型，我们采用的是relu作为损失函数，其实可以看见损失最终是收敛的，也没有出现过拟合的情况，但其实我们真正的输出值并非是一个频率点，而是一个包含一定频率范围的“音高”，于是，这可能导致损失的求法有所出入，于是我们查询了损失函数的接口，其中输入分别是一个真值一个输入，返回值即当前情况下的距离（或称作损失），本身此函数是不难的，但复杂的是，传入的参数数据类型是tensor（张量），在计算的时候不同于简单的数、向量，并且出现了在运算中张量的shape是未知的情况，这导致我们在尝试自定义损失函数的时候遇到了阻碍，目前还未解决。

而同时，我们思考，可能如今的输入数据也需要调整，因为我们目前的打算是让机器学习处理时域和频域两种计算出的有出入的数据，（但如果机器学习的效果很好，我们可以考虑直接替代），所有我们可能需要调整输入数据，把训练样本就干脆换成有出入的数据。



# 下周计划

## 1测试时域的CNN模型，输入样本为1600个音频幅值采样序列，每个样本的幅值根据自身样本情况进行归一化。

## 完成时域CNN模型测试之后，我们已测试的机器学习模型有以下3个：频域LSTM(3D)，频域2D回归，时域CNN多分类，我们将观察各个模型的表现，选择准确率最高的一个进行网络的各种优化，比如调整网络层数、使用循环dropout降低过拟合

## 3 若已有的模型表现均不尽人意，考虑再搭建时域的循环RNN网络