第二次实验

一、实验目的

多次实验求acc平均值，构造confusion matrix

二、实验记录

从官网下载下来sleep-edf，通过matlab将原始文件转为matlab存储数据的.mat文件。并同python将.mat文件中的数据选取Pz-Cz通道，随后reshape成3000\*n的矩阵。其中3000为30s乘以100Hz，3000\*n=总的采集到的数据数量。并将睡眠健康的受试者数据合并。对于标签，则是5\*n的矩阵，5是五分类标签，对应target\_class = ['W', 'N1', 'N2', 'N3', 'REM']。随导入之前训练好的模型，使用源文件中的utils调用confusion matrix的函数。得到如下结果：

Confusion matrix, without normalization

[[83758 743 364 18 224]

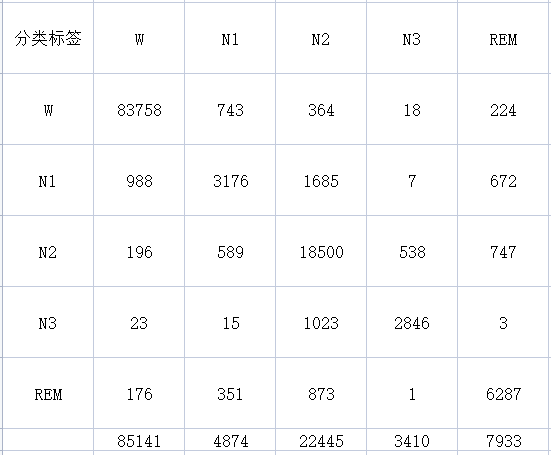
[ 988 3176 1685 7 672]

[ 196 589 18500 538 747]

[ 23 15 1023 2846 3]

[ 176 351 873 1 6287]]

为了方便理解我绘制了下图：



混淆矩阵

随后，计算每一个类的召回率、准确率和F1 score：

W：98.41 98.38 98.39

N1：48.65 65.16 55.71

N2：89.94 82.42 86.02

N3：72.79 83.46 77.76

REM：81.78 79.25 80.5

现在我们调整random\_state，也就是随机种子，来对实验结果求平均，使其具有普适性，我们重复了三次，这三次的结果如下：

Random\_state=0：89.62%~90.76%

Random\_state=1025：87.62%~90.91%

Random\_state=2022：88.84%~90.72%

三、实验结果

构建了实验的混淆矩阵，对比了不同类的分类精确度，除了N1阶段以外，其他四个类的准确度都在80%以上。而整体精确度都在90%以上应该是被Wake的准确度拉上去的，因为W阶段数据样本量大，且容易区分。此外反倒是数据量倒数第二多的N1阶段最难区分。

四、实验计划

我想可以开始搭建代码了。但是对于数据的预处理、设计、大小我还在忧虑中...