

# 第四周编程作业

## 4.3 练习Softmax回归

Softmax的预测公式如下：

$$f(\vec{z}) = \frac{e^{wx_i}}{\sum_j e^{wx_j}}$$

(1) Softmax Regression是非线性分类器。因为Softmax的输出不满足基本的线性性质：齐次性和叠加性。

(2)

- 随机抽取两个人的图像：

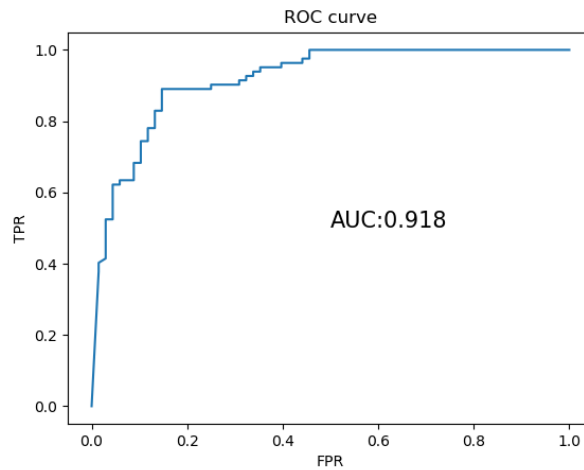
题目中需要用到的计算公式如下：

$$\begin{aligned} SN &= TPR = \frac{TP}{TP + FN} \\ FPR &= 1 - TNR \\ SP &= TNR = \frac{TN}{TN + FP} \\ FNR &= 1 - TPR \\ FDR &= \frac{FP}{FP + TP} \end{aligned}$$

正确率、TPR、FPR、TNR、FNR、sensitivity、specificity、FDR等数据展示如下：

```
测试集准确率为0.853
TPR为0.854
FPR为0.146
TNR为0.853
FNR为0.147
sensitivity为0.854
specificity为0.853
FDR为0.125
```

ROC曲线和对应的AUC值（在表格中）：



- 随机抽取五个人的图像，正确率数据如下：

测试集准确率为0.725

- 抽取所有人的图像，正确率数据如下：

测试集准确率为0.673

### (3) 结论

由运行结果可以看到随着分类数量的增加，测试集的准确率逐渐下降。这是因为分类数量的增加导致了训练需要的模型复杂度逐渐增加，此时Softmax回归不能很好地拟合测试数据，需要更加复杂的模型。

## 4.4 练习多层感知机

本题修改上一题的模型初始化方法即可。

### (1) 完成实验

本次实验用到的模型参数如下：

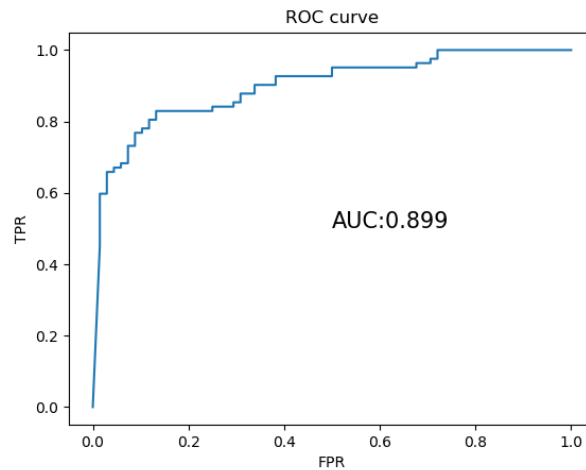
```
model = MLPClassifier(  
    hidden_layer_sizes=(3000, 1000),  
    activation='relu',  
    solver='adam',  
    batch_size='auto',  
    learning_rate='adaptive',  
    learning_rate_init=0.00001,  
    max_iter=5000,  
    shuffle=True,  
    random_state=1  
)
```

- 随机抽取两人图像

正确率、TPR、FPR、TNR、FNR、sensitivity、specificity、FDR等数据展示如下：

测试集准确率为0.840  
TPR为0.805  
FPR为0.195  
TNR为0.882  
FNR为0.118  
sensitivity为0.805  
specificity为0.882  
FDR为0.108

ROC曲线和对应的AUC值（在表格中）：



- 随机抽取五人图像，测试集准确率如下：

测试集准确率为0.704

- 抽取十人的图像，测试接准确率如下：

测试集准确率为0.589

## (2) 讨论网络参数影响（以2分类为例）

- 隐层节点数

第一层节点数	测试集准确率
(3000,1000)	0.840
(2000,1000)	0.907
(1000,1000)	0.673

可以看出，隐层节点数目和最终训练结果的效果并不是线性的关系，上图中当第一层节点数为2000时测试集准确率更高，这说明在实际训练中往往需要我们手动去调试网络结构，但注意节点数目不应过大或过小，容易导致欠拟合或过拟合

- 隐层层数

隐层层数	测试集准确率
(3000,1000)	0.840
(3000,1000, 200)	0.9
(3000)	0.813

由上图可以看到随着网络层数增加，复杂度逐渐升高，测试集准确率也相应提高，但注意可能会出现过拟合的情况，因此有些时候不能一味增加网络复杂度，而是应该根据数据特点调试。

● 学习率

学习率	测试集准确率
0.00001	0.840
0.0005	0.847
0.00005	0.927

学习率过小会导致最终可能收敛到局部最优解，且学习速度很慢。而学习率过大，可能会导致结果难以收敛，因此实际调参时学习率对与最终结果的影响很重要。

(3) 对比本次实验与上次实验

两次实验数据对比表格如下：

	Softmax准确率	MLP准确率
2分类	0.853	0.840
5分类	0.725	0.704
10分类	0.673	0.589

由表格数据可以看出，两类网络都是随着类别增加准确率降低。MLP参数更多，训练时占用的内存更多，训练时间更慢。尽管MLP的网络复杂度更高，但在测试集上的准确率似乎并没有比Softmax高，这可能是因为MLP训练涉及的参数更多，需要我们手动去调试参数，以达到更优的训练效果。