

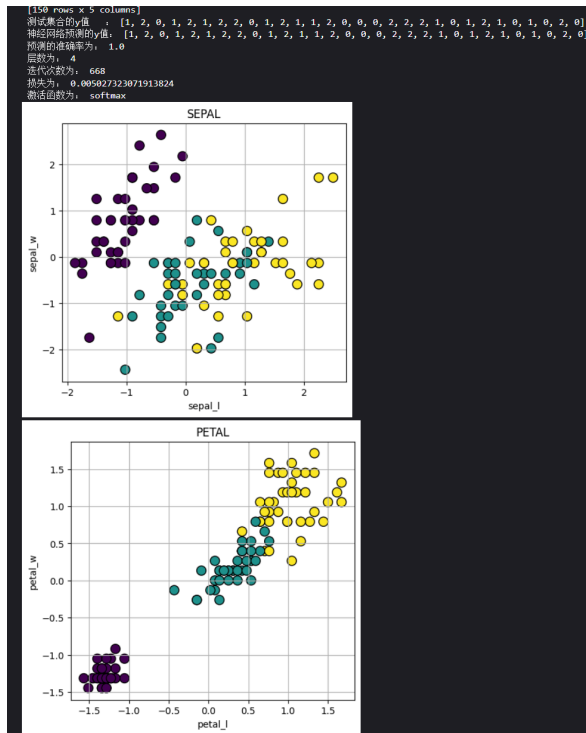


lab2_report

PB22111606_龙涵睿

1.运行结果

- 基于sklearn框架的运行结果（特征分布图是自己加的）



- 基于助教框架的运行结果

```
training.....  
Epoch 1, Loss: 0.21181034291506312  
Epoch 101, Loss: 0.04996662735177117  
Epoch 201, Loss: 0.0403792224679029  
Epoch 301, Loss: 0.035297809833111704  
Epoch 401, Loss: 0.03115171096449785  
Epoch 501, Loss: 0.027303870181452923  
Epoch 601, Loss: 0.023648389511928897  
Epoch 701, Loss: 0.020257820198469723  
Epoch 801, Loss: 0.01725040648678774  
Epoch 901, Loss: 0.014711883697791573  
测试集合的y值: [1, 2, 0, 1, 2, 1, 2, 2, 0, 1, 2, 1, 1, 2, 0, 0, 0, 2, 2, 2, 1, 0, 1, 2, 1, 0, 1, 0, 2, 0]  
神经网络预测的y值: [1, 2, 0, 1, 2, 1, 2, 2, 0, 1, 2, 1, 1, 2, 0, 0, 0, 2, 2, 2, 1, 0, 1, 2, 1, 0, 1, 0, 2, 0]  
预测的准确率为: 1.0
```

由结果可知:

TA >= sklearn is True

2.实验用时

我是在kaggle的notebook上完成的，算上在本地环境的七七八八，大概三四个小时吧

3.超参数修改

- sklearn:

```
mlp = MLPClassifier(hidden_layer_sizes=(10,10), max_iter=2000, activation='tanh', solver=
```

- 手动:

因为补充完代码后只修改了**learning_rate:0.01 -> 0.1**就已经能够跑到**acc==1**了，所以没有增加正则化等，epoch也是 default——1000，因为数据集是按比例随机划分，所以准确率会在0.9以上有所浮动，1中截图为最佳结果，即一次划分两种方法都是acc==1

4.代码相关

注释都在代码中，此外，因为readme里面有助教复现代码运行结果的有关要求，所以我用 **random_state=22111606**来实现一个可复现的运行结果实践发现就算我用了同一个 random_state也会导致结果出现细微差异，不知道为啥，，，但结果都是在0.9~0.9333...

```
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,train_size=0.8,random_state=22111606)
```

5.other

一开始不是在kaggle_nb上面跑的，而是在本地envs，但是

又 `ImportError: cannot import name '__version__' from 'numpy' (unknown location)` ,很闹麻，这个numpy版本兼容性这么差吗