HW3·Part1

Xiang Pan 潘翔
1900941520
hover@hust.edu.cn

Q1 PLA Result

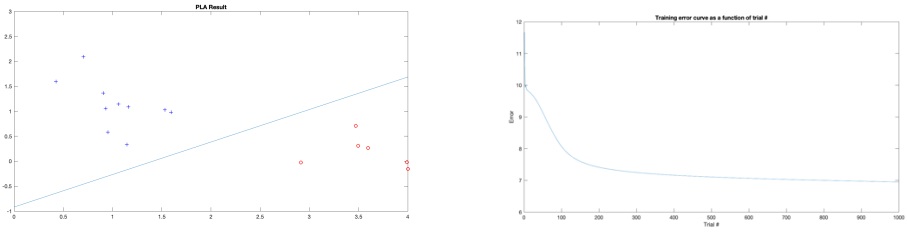


Image 1.1 The Result and Train Loss of PLA

Q2 MLP Result

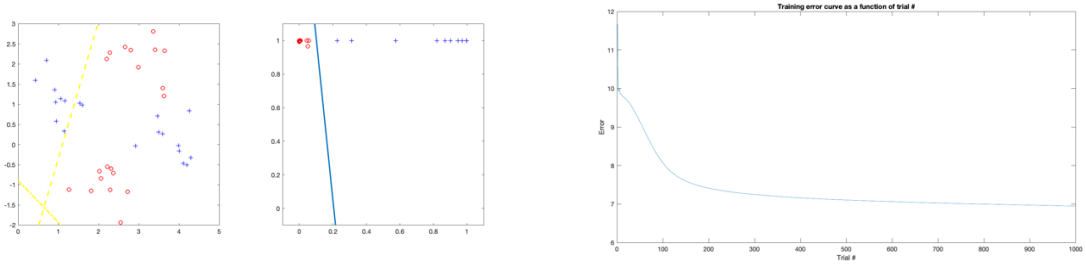
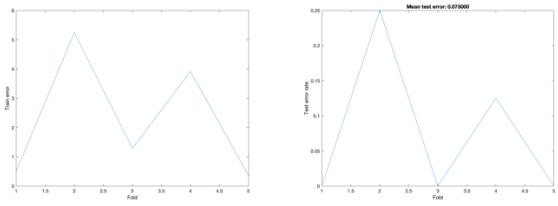


Image 2.1 The Result and Train Loss of MLP



Q3 5-Fold

Image 3.1 Train Loss of 5 Folds(Classication)

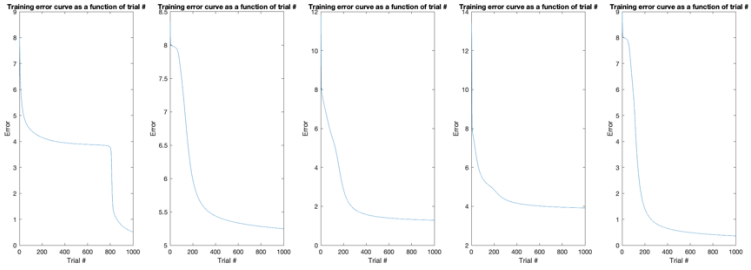


Image 3.2 Train Error of 5 Folds(XOR)

HW3·Part2

QC: MLP for Classification

Net1

Layer (type)	Output Shape	Param #
dense_73 (Dense)	(None, 4)	12
dense_74 (Dense)	(None, 4)	20
dense_75 (Dense)	(None, 1)	5

Total params: 37
Trainable params: 37
Non-trainable params: 0

Image C.1 Network Summary1

```
MLP_Split(0.8,0.2):0.875
MLP_overfit:1.0
MLP_5_Fold:
```

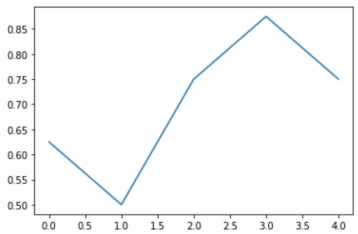


Image C.2 Train Error of 5 Folds(Apple & Orange)

Net2

Layer (type)	Output Shape	Param #
dense_76 (Dense)	(None, 10)	30
dense_77 (Dense)	(None, 10)	110
dense_78 (Dense)	(None, 1)	11

Total params: 151
Trainable params: 151
Non-trainable params: 0

Image C.3 Network Summary2

```
MLP_Split(0.8,0.2):1.0
MLP_5_Fold:
```

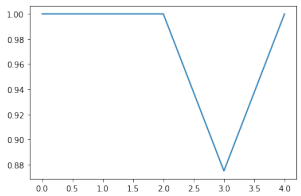


Image C.4 Train Error of 5 Folds(Apple & Orange)

MLP for Spirals:

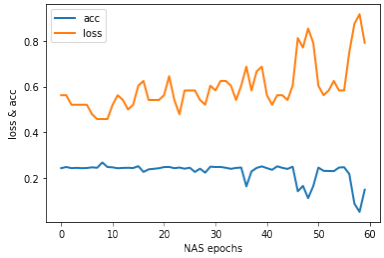


Image C.5 MLP of Spirals

Because of the highly nonlinear of spirals' data, the MLP cannot balance the simple network architecture and generalization. In fact, the highest acc is about 0.92, but the model will overfit in someway.

NAS for MLP

MLP is the simplest NN, so it is very useful in understanding how NN works. In part.1 we try to adjust the hyperparameter to minimize the NN scale and maximize the performance of task. Here, I try to use NAS Method for more smart network design.

Here, I use the simplest violent search for MLP Network Architecture.

```
for neuron_indx in range(neuron_max // 10):
    for layer_indx in range(layer_max):
    model = basic_model(layers, neurons)
```

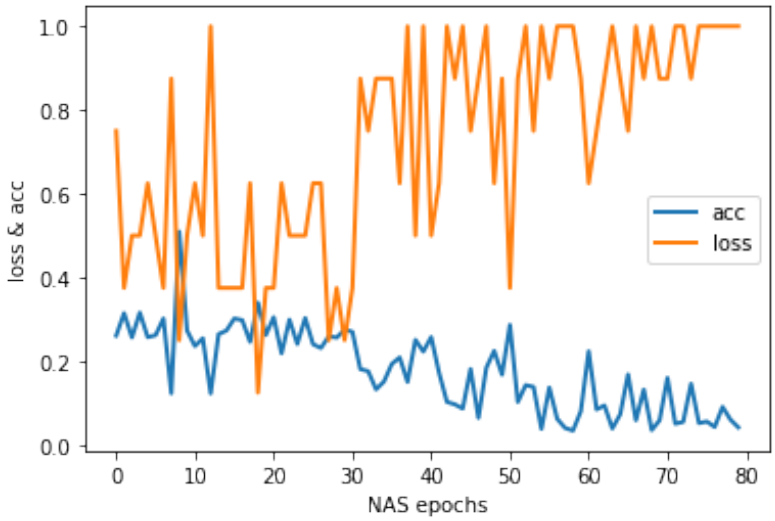


Image C.3 NAS for MLP