

Introduction to Neural Networks Homework1

N26100618 李佩萱

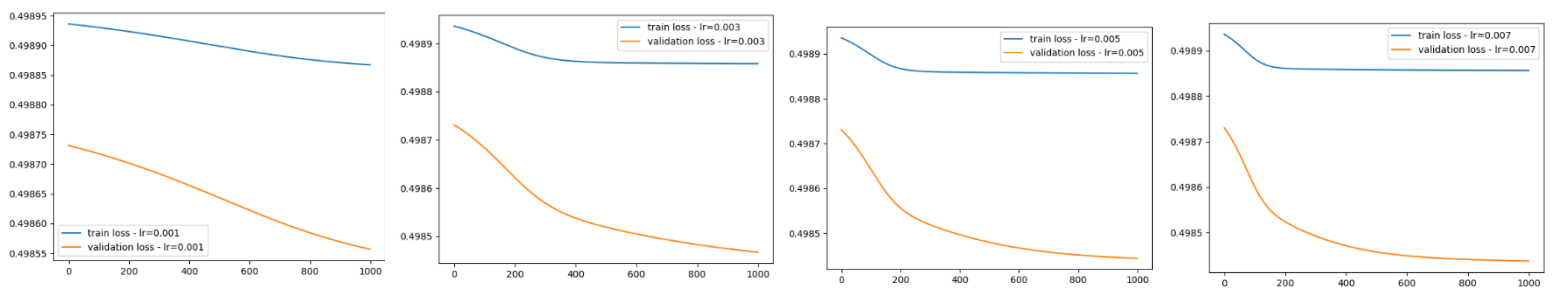
I. Data preprocessing

- 類別型資料：做 one-hot encoder。
- 數值型資料：做標準化，使平均值=0，標準差=1。
- Y：將 yes、no 轉換為 1 與 0 後，做 one-hot encoder。
- Imbalance data：答案為 0 的資料筆數過多，隨機刪減至與答案為 1 的資料筆數相同，各 4640 筆。
- Split Data：將資料分為 train 72%、validation 18%、test 10%。

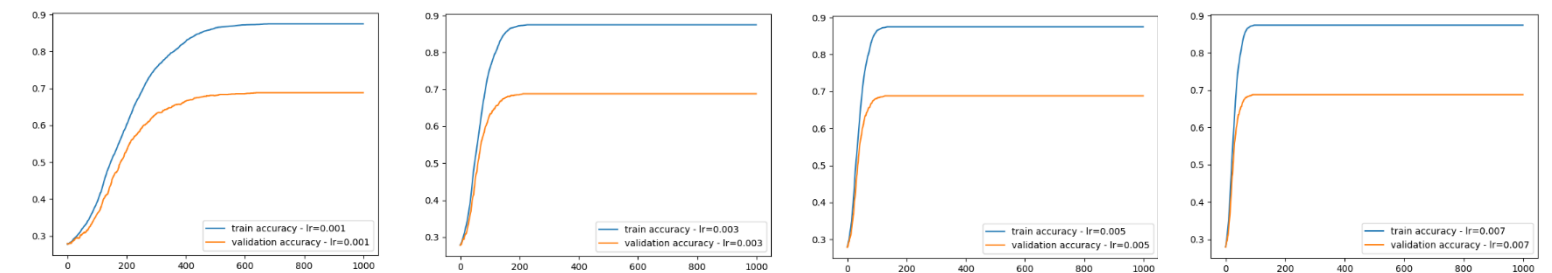
II. Neural Network (1 hidden layer)

- 設定 hidden layer 有 10 個 neuron，output size 為 2 (二分類，因此 output layer 的 activation function 固定為 softmax)，loss function 為 mean-square error。
 - 變化 learn rate、hidden layer 的 activation function，訓練 1000epochs，觀察 loss、accuracy、weight change。
- Activation function = sigmoid，learning rate = 0.001、0.003、0.005、0.007

a. Loss 變化

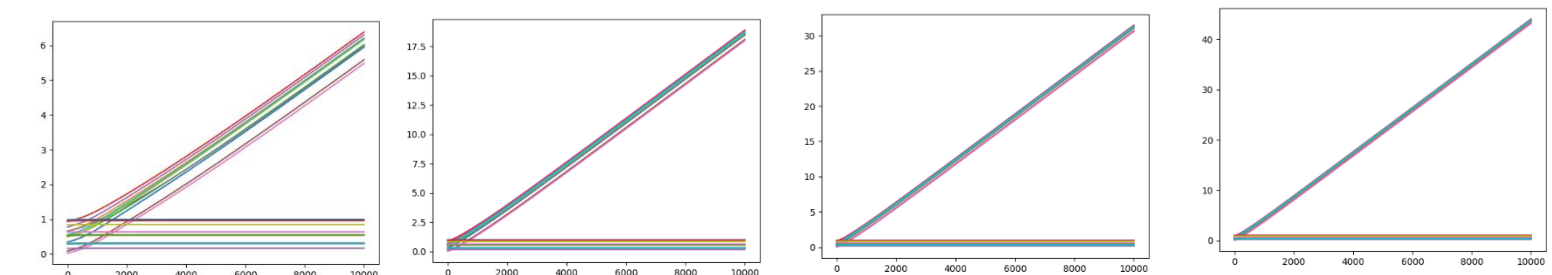


b. Accuracy 變化

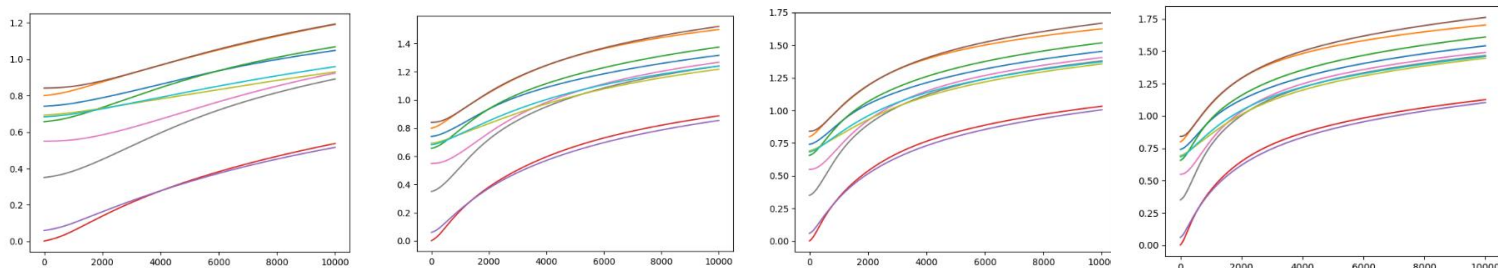


c. Weight 變化 (增加 epoch 使變化更明顯，因此拉長至 10000 epochs)

✧ Output layer weights

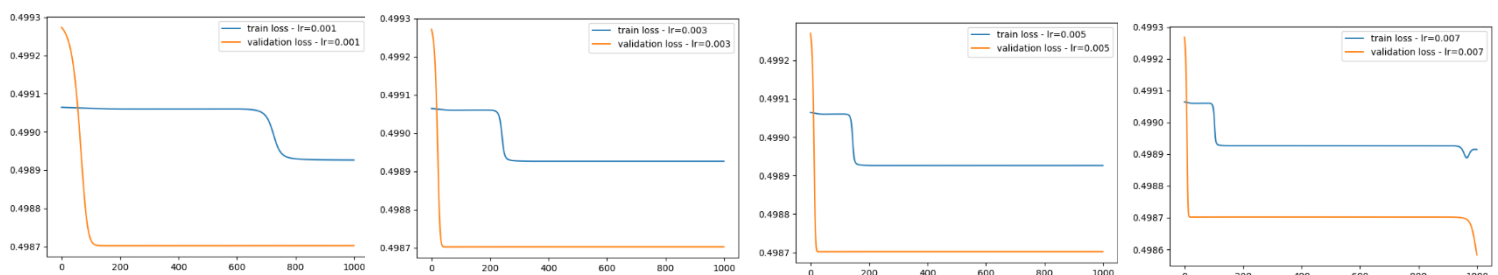


✧ Hidden layer bias

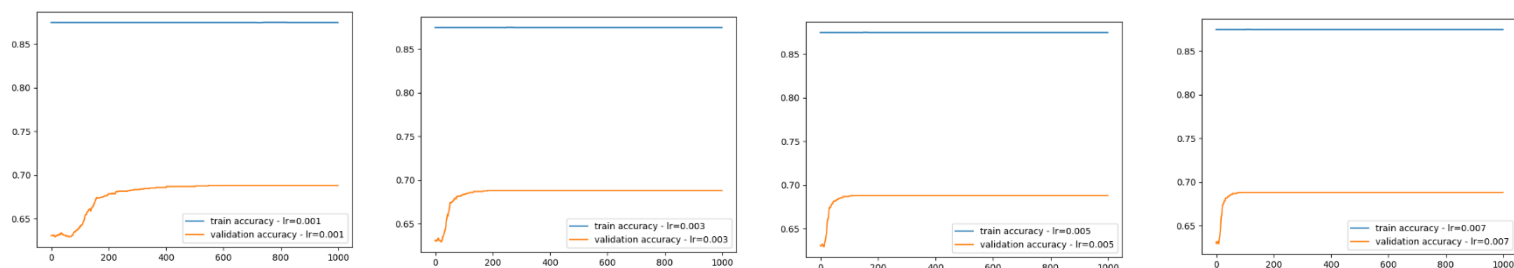


2. Activation function = relu , learning rate =0.001 、0.003 、0.005 、0.007

a. Loss 變化

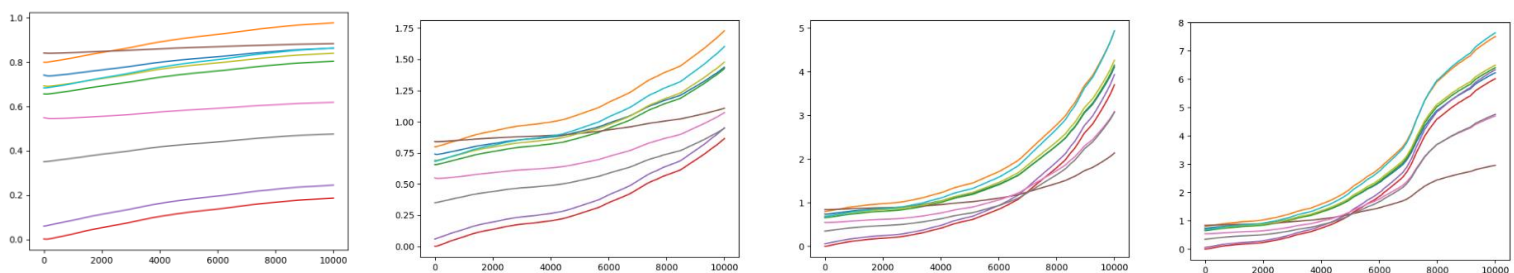


b. Accuracy 變化

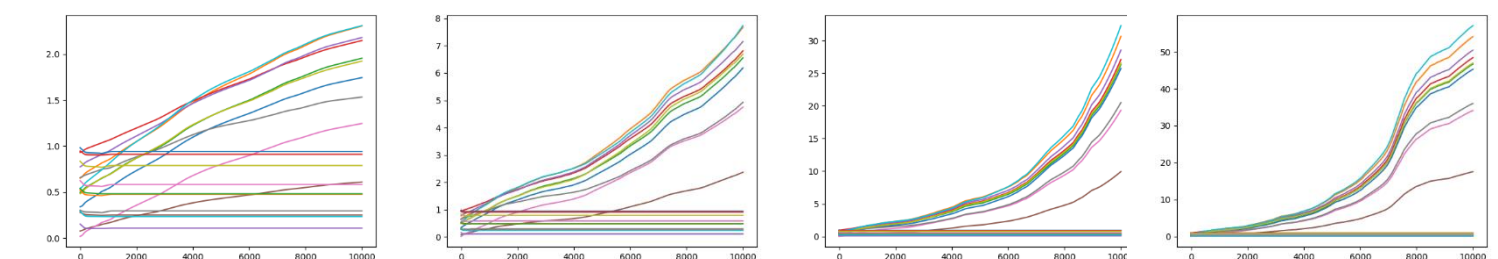


c. Weight 變化 (增加 epoch 使變化更明顯，因此拉長至 10000 epochs)

✧ Output layer weights



✧ Hidden layer bias



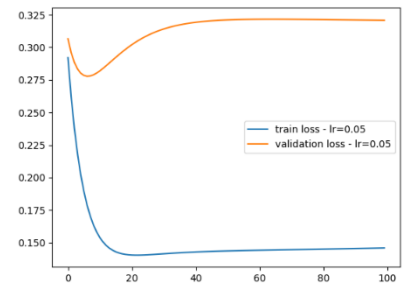
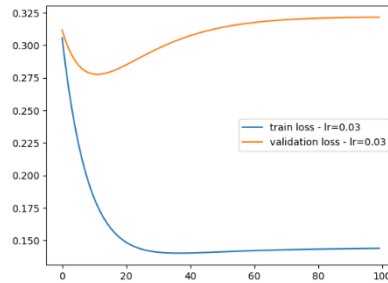
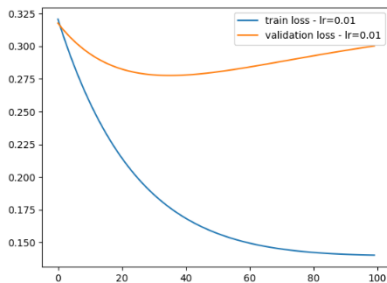
III. Radial Basis Function Network

A. 設定 hidden layer 有 10 個 neuron，output size 為 1 (regression)，loss function 為 mean-square error。

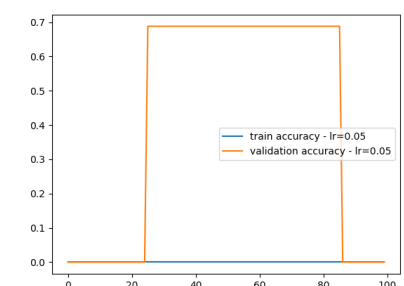
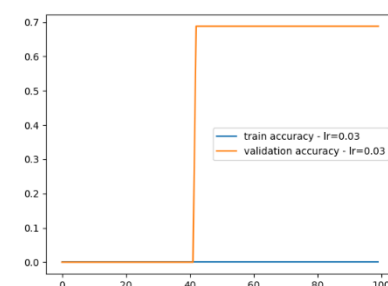
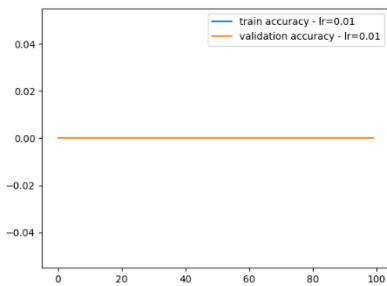
B. 變化 learn rate、radial basis function，訓練 100epochs，觀察 loss、accuracy、weight change。

1. Radial basis function $\varphi(r) = e^{-(\varepsilon r)^2}$ ，learning rate = 0.01、0.03、0.05。

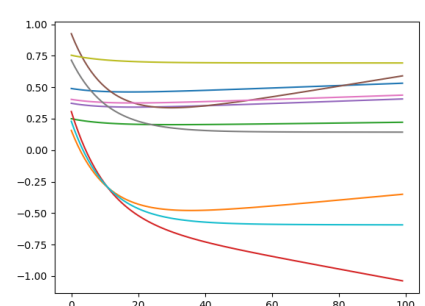
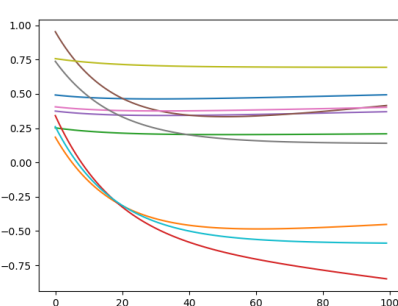
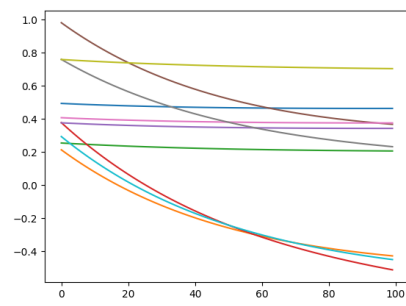
a. Loss 變化



b. Accuracy 變化

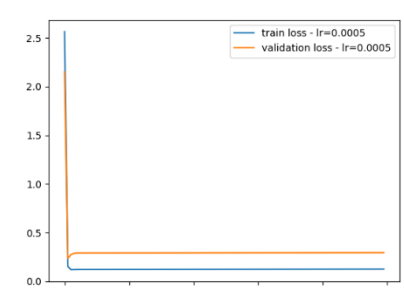
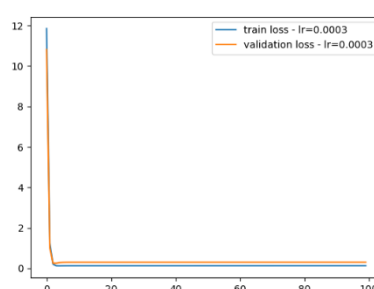
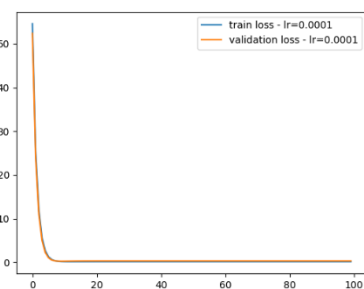


c. Weight 變化 (Output layer)

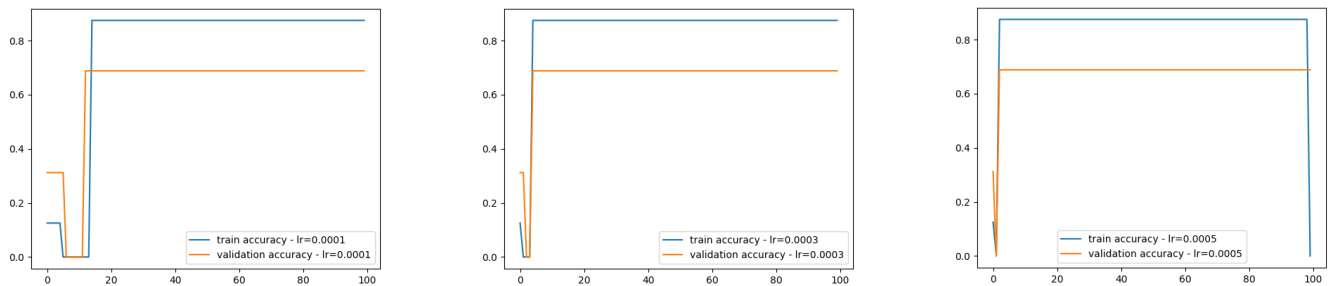


2. Radial basis function $\varphi(r) = \sqrt{1 + (\varepsilon r)^2}$ ，learning rate = 0.0001、0.0003、0.0005。

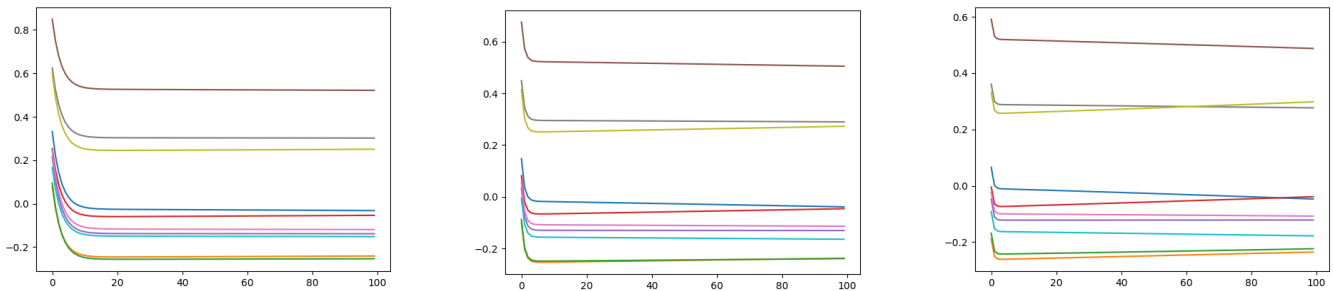
a. Loss 變化



b. Accuracy 變化



c. Weight 變化 (Output layer)



IV. 比較與討論

A. Neural network

1. 比較相同 activation function，不同 learning rate
 - a. learning rate 越大，loss、accuracy、weight 收斂越快。
2. 比較 relu 與 sigmoid 兩個 activation function
 - a. relu 一開始即有較高 accuracy、較低 loss，收斂也比 sigmoid 快
 - b. weight 變化的部分，sigmoid 的變化較 relu 平滑。

B. Radial Basis Function Network

1. 比較相同 radial basis function，不同 learning rate
 - a. learning rate 越大，loss、accuracy、weight 收斂越快。
 - b. learning rate 較大則較容易 overfitting (epoch 增加反而 validation 的 loss 上升、accuracy 下降)。
2. 比較 2 個 radial basis function
 - a. $\varphi(r) = e^{-(\varepsilon r)^2}$ 的收斂速度較慢，需將 learning rate 調大或訓練非常多 epoch，參數才可收斂，但 train 的 accuracy 仍訓練不起來。
 - b. $\varphi(r) = \sqrt{1 + (\varepsilon r)^2}$ 收斂快，在 accuracy 的表現相對較佳也較穩定。

C. NN 與 RBFN 比較

1. NN 可以訓練的參數較多，搭配足夠多的數據，較容易訓練起來。
2. RBFN 較不容易訓練成功，需多試幾組 hyperparameters。且訓練起來後(loss 明顯下降)非常容易一下子就 overfitting。