

## 模型架構

- 輸入層：x (1 維)
- 隱藏層：2 層，每層包含 64 個神經元，激活函數採用 tanh
- 輸出層：1 層

## 訓練架構

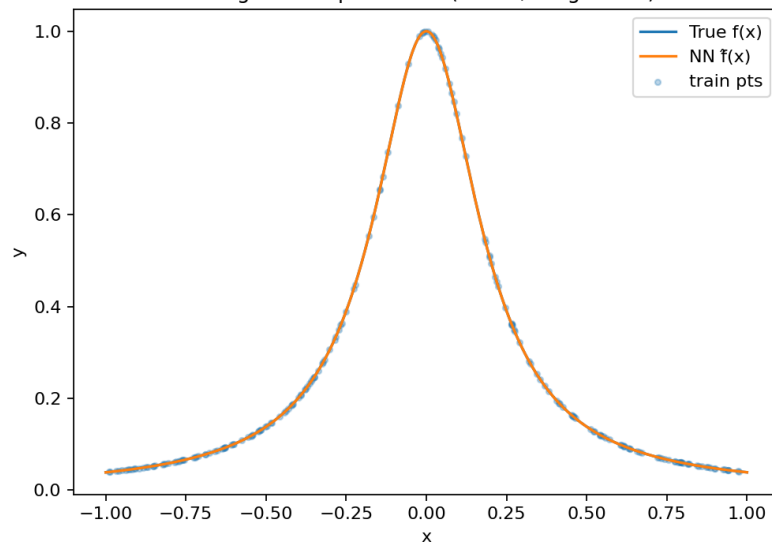
- **Loss function**：MSELoss，包含 function loss 與 derivative loss
- **Optimizer**：Adam，學習率  $10^{-3}$
- **學習率調整**：ReduceLROnPlateau (factor = 0.5, patience = 50)
- **Early stopping**：若驗證誤差連續 150 epoch 無改善則提前停止

## 資料集

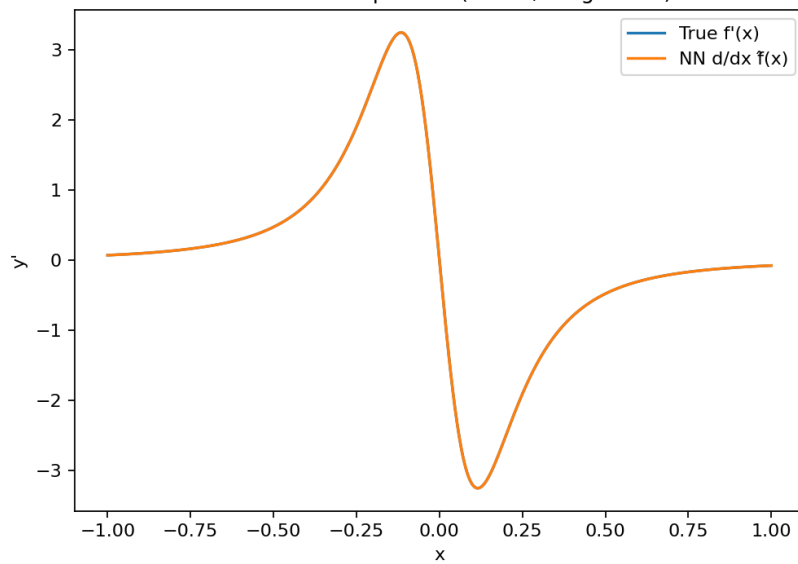
- 訓練集：200 點
- 驗證集：50 點
- 測試集：1000 點，取自區間  $[-1,1]$

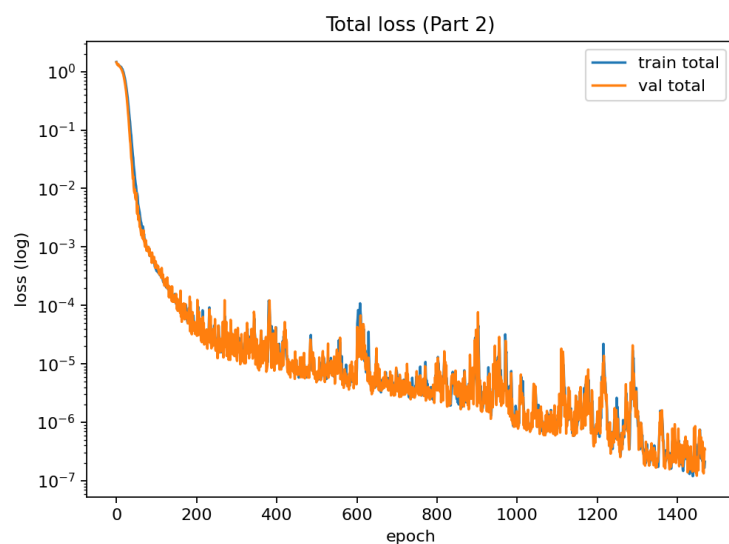
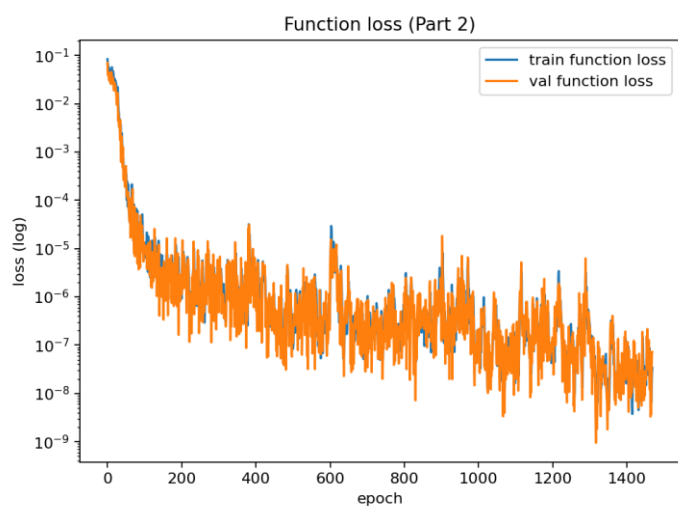
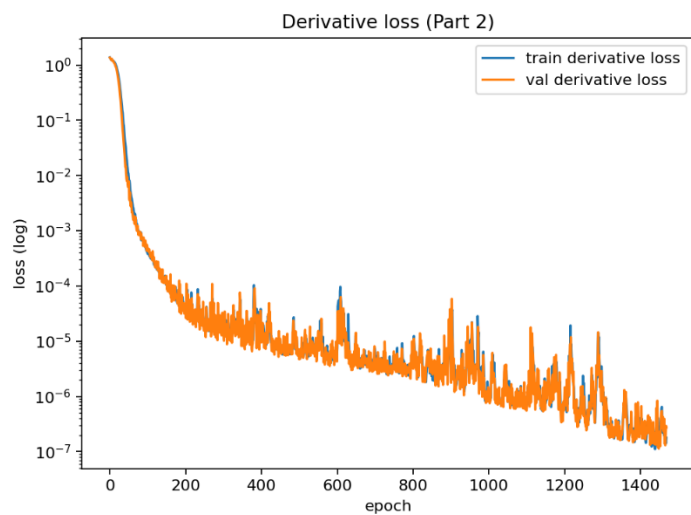
## 結果 2

Runge vs NN prediction (Part 2, no-grad fix)



Derivative comparison (Part 2, no-grad fix)





測試集表現

根據程式輸出結果：

```

PS C:\Users\user> python -u "c:\Users\user\Desktop\大四上\機器學習\HW3\assignment3_programing_2.py"
Epoch 100 | tr_tot 4.183e-04 | tr_f 1.347e-05 | tr_fp 4.048e-04 | va_tot 5.006e-04 | va_f 1.133e-05 | va_fp 4.892e-04
Epoch 200 | tr_tot 3.631e-05 | tr_f 1.243e-06 | tr_fp 3.507e-05 | va_tot 3.509e-05 | va_f 1.877e-06 | va_fp 3.321e-05
Epoch 300 | tr_tot 1.863e-05 | tr_f 2.540e-06 | tr_fp 1.609e-05 | va_tot 1.636e-05 | va_f 1.303e-06 | va_fp 1.505e-05
Epoch 400 | tr_tot 1.756e-05 | tr_f 1.686e-06 | tr_fp 1.588e-05 | va_tot 1.432e-05 | va_f 1.110e-06 | va_fp 1.321e-05
Epoch 500 | tr_tot 5.561e-06 | tr_f 1.795e-07 | tr_fp 5.382e-06 | va_tot 4.845e-06 | va_f 1.182e-07 | va_fp 4.727e-06
Epoch 600 | tr_tot 5.751e-06 | tr_f 2.897e-07 | tr_fp 5.462e-06 | va_tot 4.479e-06 | va_f 2.094e-07 | va_fp 4.269e-06
Epoch 700 | tr_tot 7.746e-06 | tr_f 3.999e-07 | tr_fp 7.346e-06 | va_tot 3.788e-06 | va_f 5.417e-07 | va_fp 3.247e-06
Epoch 800 | tr_tot 3.831e-06 | tr_f 4.520e-07 | tr_fp 3.379e-06 | va_tot 8.245e-06 | va_f 1.199e-06 | va_fp 7.046e-06
Epoch 900 | tr_tot 1.064e-05 | tr_f 2.800e-06 | tr_fp 7.835e-06 | va_tot 8.516e-06 | va_f 1.457e-06 | va_fp 7.059e-06
Epoch 1000 | tr_tot 1.168e-06 | tr_f 9.663e-08 | tr_fp 1.071e-06 | va_tot 1.495e-06 | va_f 1.910e-07 | va_fp 1.304e-06
Epoch 1100 | tr_tot 8.472e-07 | tr_f 5.170e-08 | tr_fp 7.955e-07 | va_tot 6.287e-07 | va_f 7.820e-08 | va_fp 5.505e-07
Epoch 1200 | tr_tot 8.201e-07 | tr_f 1.053e-07 | tr_fp 7.148e-07 | va_tot 1.134e-06 | va_f 1.440e-07 | va_fp 9.903e-07
Epoch 1300 | tr_tot 1.394e-06 | tr_f 2.898e-07 | tr_fp 1.104e-06 | va_tot 1.346e-06 | va_f 2.411e-07 | va_fp 1.105e-06
Epoch 1400 | tr_tot 4.424e-07 | tr_f 3.689e-08 | tr_fp 4.055e-07 | va_tot 2.514e-07 | va_f 3.937e-08 | va_fp 2.121e-07
Early stopping @ 1469 | best val=1.911149e-07
[Test] f(x): MSE=1.325996e-08 | Max|err|=2.086870e-04
[Test] f'(x): MSE=2.281066e-07 | Max|err|=1.449585e-03

```

- $f(x)$  : MSE =  $1.325996 \times 10^{-8}$  , 最大誤差  $\approx 2.09 \times 10^{-4}$
- $f'(x)$  : MSE =  $2.281066 \times 10^{-7}$  , 最大誤差  $\approx 1.45 \times 10^{-3}$

## 說明

- 本次程式同時將函數與導數誤差納入訓練，結果顯示模型能夠有效學習 Runge 函數的整體形狀，並能在高曲率區域（例如靠近  $x=0$  附近）提供精確的導數近似。
- 從 loss 曲線可以看到，function loss 與 derivative loss 都能隨著 epoch 持續下降，並且在驗證集上保持穩定，顯示模型沒有發生過度推論。
- 與僅訓練函數值的情況相比，加入 loss 訓練方式能顯著改善導數近似能力，尤其在 Runge 函數變化劇烈的區域。