1.Data	- 輸入： $x \in [-1,1]$ - 訓練集：200 點 - 驗證集：100 點 - 測試集：500 點 - 輸出：Runge 函數 $f(x)=\frac{1}{1+25x^2}$ 及其導數 $f'(x)=\frac{-50x}{(1+25x^2)^2}$
2.Hypothesis	Feedforward Neural Network, MLP，近似 $f(x)$，導數由自動微分 (autograd) 計算
3.Hidden Layers	2 層
4.Activation Function	Tanh
5.Neurons in Each Hidden Layer	第一層：64 個 第二層：64 個
6.Loss Function	1. 函數損失 (Function Loss)：$L_f = MSE(\hat{f}(x), f(x))$ 2. 導數損失 (Derivative Loss)：$L_{f'} = MSE(\frac{d}{dx}\hat{f}(x), f'(x))$ 總損失：$L = L_f + L_{f'}$
7.Stopping Criteria	固定訓練 2000 個 epoch

Epoch 0, Train Total Loss: 2.003874, Func Loss: 0.082507, Deriv Loss: 1.921367, Val Loss: 2.633613

Epoch 200, Train Total Loss: 0.000009, Func Loss: 0.000000, Deriv Loss: 0.000009, Val Loss: 0.000009

Epoch 400, Train Total Loss: 0.000007, Func Loss: 0.000000, Deriv Loss: 0.000007, Val Loss: 0.000007

Epoch 600, Train Total Loss: 0.000009, Func Loss: 0.000000, Deriv Loss: 0.000009, Val Loss: 0.000009

Epoch 800, Train Total Loss: 0.000006, Func Loss: 0.000000, Deriv Loss: 0.000006, Val Loss: 0.000006

Epoch 1000, Train Total Loss: 0.000003, Func Loss: 0.000000, Deriv Loss: 0.000003, Val Loss: 0.000003

Epoch 1200, Train Total Loss: 0.000005, Func Loss: 0.000000, Deriv Loss: 0.000005, Val Loss: 0.000005

Epoch 1400, Train Total Loss: 0.000004, Func Loss: 0.000000, Deriv Loss:

0.000003, Val Loss: 0.000002

Epoch 1600, Train Total Loss: 0.000000, Func Loss: 0.000000, Deriv Loss: 0.000000, Val Loss: 0.000000

Epoch 1800, Train Total Loss: 0.000068, Func Loss: 0.000006, Deriv Loss: 0.000062, Val Loss: 0.000126

Function Approximation -> MSE: 5.926243e-10, Max Error: 4.903061e-05

Derivative Approximation -> MSE: 4.962465e-07, Max Error: 1.934788e-03