Programming assignment Week 2

Question

Use a neural network to approximate the Runge function:
$f(x) = \frac{1}{1+25x^2}, \quad x \in [-1, 1]$.

The way to approximate the function

(This report used chatgpt for assistants to complete the code)
Function hypothesis: $h(x) = W_2 tanh(W_1 x + b_1) + b_2$, with 1 neuron input layer, 1 neuron output layer, and 30 neurons hidden layer. That is, $W_1 \in M^{1\times 30}$, $W_2 \in M^{30\times 1}$. We randomly take 400 data; 280 of them are used to train, and 60 of them are used to check validation loss, and the last 60 of them are used to evaluate final MSE. During training, we first compute the prediction and MSE. Then by back propogation, compute $\frac{\partial L}{\partial W_2}, \frac{\partial L}{\partial W_1}, \frac{\partial L}{\partial b_2}, \frac{\partial L}{\partial b_1}$ for updating the parameters. The learning rate is 0.03 with 300 epochs.

Result

```
Test MSE = 9.5773e-03
Test max abs error = 2.7207e-01
```

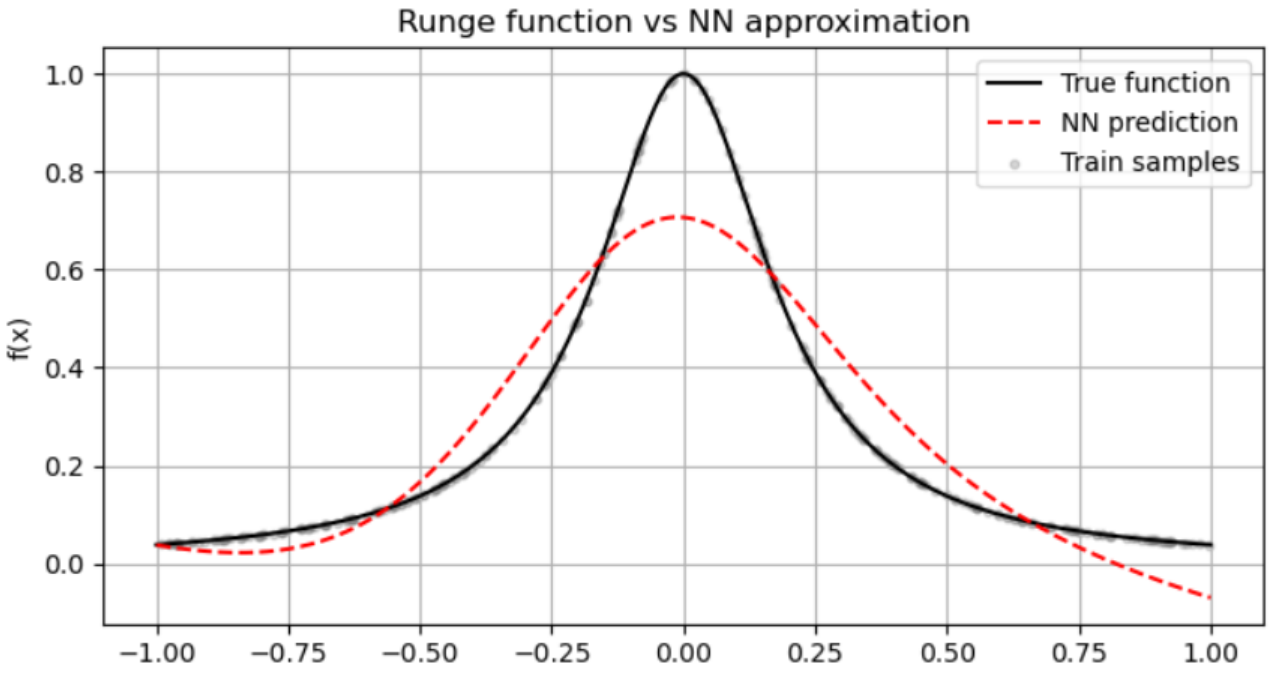

Figure 1: The true function and The neural network prediction

At first, both train loss curve and validation loss curve are high. After about 10 Epochs, they drop sharply. Besides, the two curves are almost overlap. So I think the result is not bad. As for the final NN prediction, the MSE is 0.0095773, I think it is low enough, while max error is 0.27207, which still can be improved.

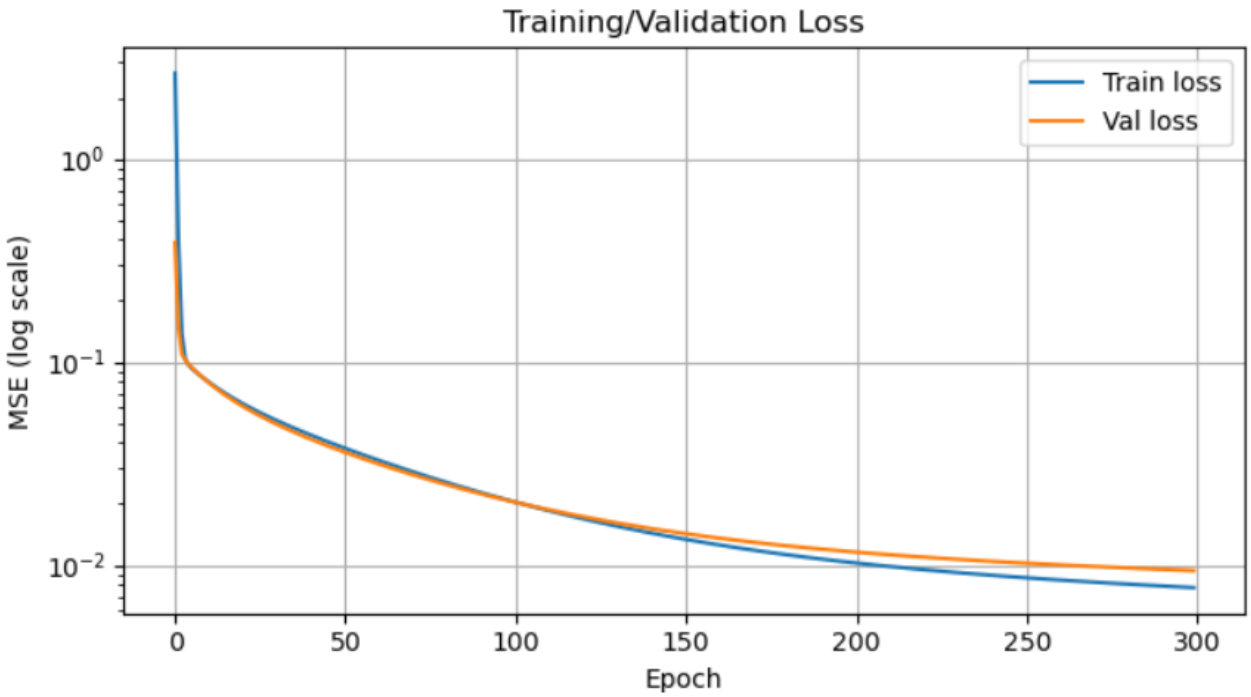

Figure 2: The training/validation loss curves