

Week 2 Program Assignment Report

Design of Network

activation function: $\sigma(x) = \tanh(x)$

Input layer: 1 input, x

second layer: 7 neurons, and σ

third layer: 7 neurons, and σ

output layer: 1 output

Training Process

loss function: MSE loss

optimizer: Adam

learning rate: 0.01

train times: 2000

Training Result

The results show that this model can approximate the original function very well:

- 1) The curve predicted by the model is obviously very similar to the true function. (see Figure 1)
- 2) The convergence speed of model training is very fast. In 2000 rounds of training, it basically converges to the minimum value in about 200 rounds. (see Figure 3)
- 3) The MSE loss of test data and true function value < 0.00001
- 4) The network structure is simple, takes less time. Even if only use cpu, it only takes about 10 seconds to complete the training.

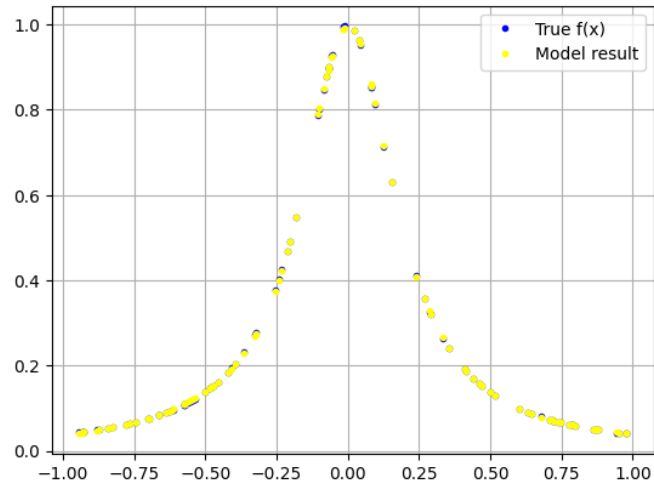


Figure 1: Comparison of test set

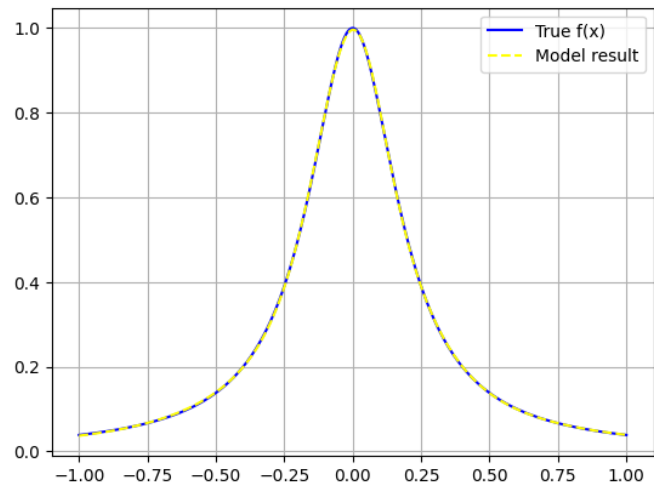


Figure 2: Comparison of prediction and function

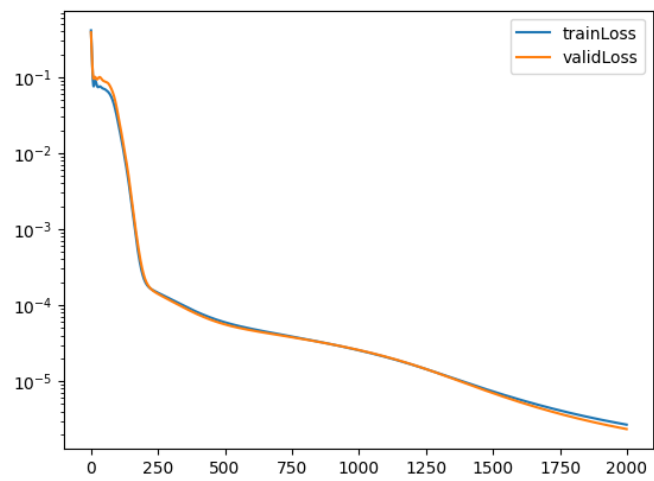


Figure 3: Train and valid error curves