Homework 1 Report

1-1 Deep vs Shallow

Simulate a Function

All models used had an input and output dimension of 1. The models I used are as follows:

- Model 1: DNN with 2 hidden layers of sizes 75 and 25. Total of 2076 parameters
- Model 2: DNN with 5 hidden layers of sizes 17, 26, 25, 23, and 12. Total of 2076 parameters
- Model 3: DNN with 1 hidden layer of size 692. Total of 2077 parameters

All models were trained for 100 epochs and I fine tuned for batch size and learning rate on all models. The optimal hyperparameters were:

- Model 1: batch_size = 8, learning_rate = 0.001
- Model 2: batch_size = 16, learning_rate = 0.001
- Model 1: batch_size = 8, learning_rate = 0.001

The loss for all models are shown in the plot below:

fn1loss

Train on Actual Task