

HW #5

① Bishop 5.26

② Carry out regression on the Bessel function ($J_0(x)$) dataset from previous HW's using an ANN with a single hidden layer. Either implement the ANN or use an existing package.

Train the model using $M = 2, 4, 6, 8, 10$.
For each value of M , ~~fit~~ ^{fit} the model 15 times starting from random weights and biases. Plot an equivalent of Fig. 5.10 in Bishop: test-set sum-of-squares error vs. M .
↑ set some data aside, describe it briefly

Repeat the process for two hidden-layer activation functions of your choice, comment on the differences (if any). Use early stopping if necessary.

Extra credit (+5): explore two SGD flavors (i.e., produce 4 plots instead of 2)