Bryan Faryadi (801178567) - Homework 6

Github: <https://github.com/bfaryadi/ECGR4105_hw6_bf>

Problem 1

My final loss with one hidden layer (of 8 nodes) and 3000 epochs was 3.550840e+12. This is worse than the results from the linear model in homework 5, and increasing the number of epochs did not yield more than marginal decreases. This training took 6.46 seconds.

After adding two more hidden layers (both with 8 nodes), the final loss ended slightly lower at 3.550506e+12, but still seems to be converging on the same value as in part A. The difference is that it reached this value far, far before the model with one hidden layer would. More epochs yield no results, as it had converged on this value by epoch 500. Surprisingly, this training only took slightly longer than part A at 6.62 seconds.

These results similar to that of SVR with RBF or polynomial models. However, the linear model far outperformed the neural networks at around 2e+12.

Problem 2

For part A, I had one hidden layer with 32 nodes. My final loss was 8.012608e-02. My accuracy was 0.9649, which is 110/114 correctly classified. I did this training with 3000 epochs, and it took 4.73 seconds. These results are very slightly worse than homework 4 using SVC with a linear kernel.

For part B, I added two hidden layers with 32 nodes each. My final validation loss increased to 8.196887e-02, but the accuracy went up to 0.9824, which is 112/114 correctly classified. There may be some overfitting here, as the pool of 114 includes the training set as well as the validation set, which could explain why the accuracy increased even though the validation loss also did.

Problem 3

After 100 epochs with a model with just one hidden layer of 512 nodes my final loss was 0.0884. Validation accuracy was 0.4563, and the training took 1746.70 seconds.

For the second part, I did three layers of 1024, 512, and 256 nodes respectively. For this model, the final loss was 0.0006, but the validation accuracy was only slightly better at 0.4718. This indicates overfitting, as the loss is only calculated for the training set. This training took 2507.21 seconds.