Machine Learning Assignment 5

Dhruv Chauhan

December 29, 2016

1 Neural Networks

1.1 Neural network implementation

1.2 Neural network training

The plots were attempted to be combined, but due to the number of lines they became difficult to understand, so I instead include the graphs on the next page, for 2 and 20 hidden neurons, and with learning rates of $[0.01, 0.1, 0.5, 1, 3]$.

Looking at the graphs, for small learning rates the error takes a long time to converge, (around 1000 epochs for 0.01). With a bigger learning rate the rate converges much faster, but at a local minimum that is higher than with lower rates (see learning rate of 3 for example).

Below are the graphs for the input values [-10...10], (excluding 0 due to the division error in the sinc function) with the trained neural network and the sinc function, for both 2 and 20 hidden neurons.
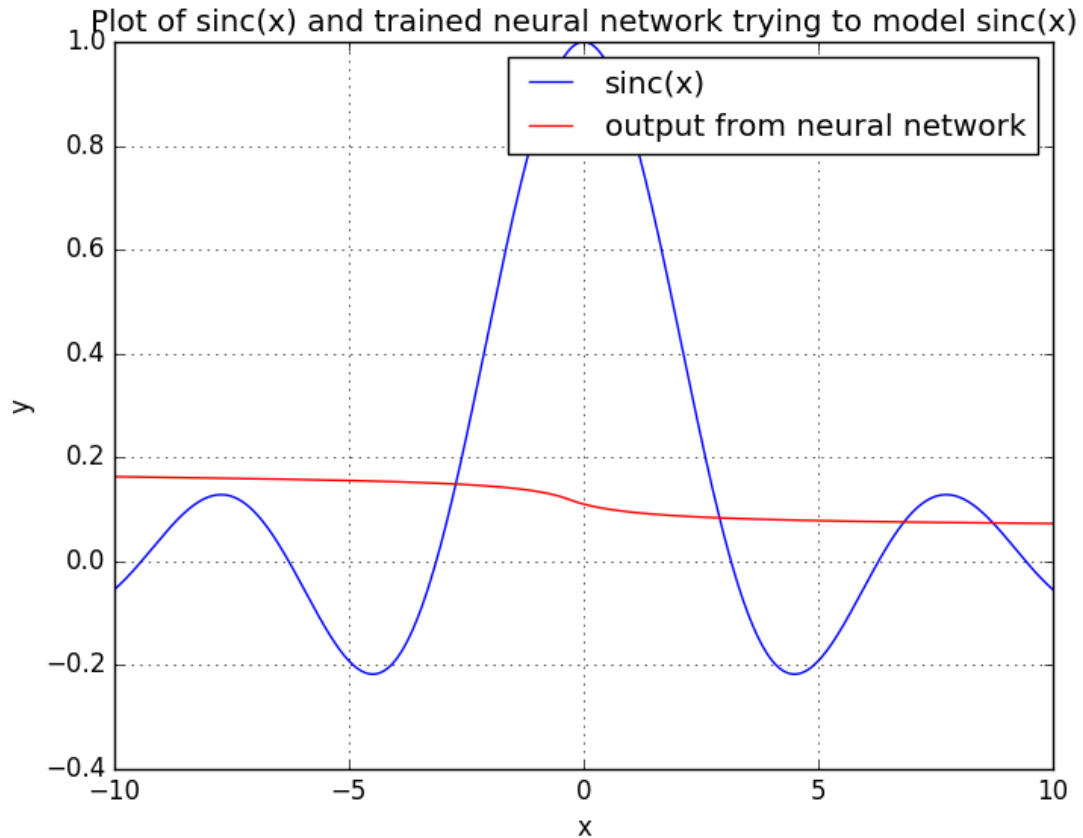

Figure 1: Neural Network trained with 2 Hidden Neurons

Both models seemed to have missed the actual result by a bit, but the model with 2 hidden neurons roughly averages out the function, except around 0. The one with 20 hidden neurons oddly doesn't seem more accurate, although in total it crosses the true graph more times. Overfitting seemed to occur with a very low learning rate and a high number of epochs. We know that over fitting is occuring when the error rate on the validation set starts to increase again. By using early-stopping we prevent the over-training / this increase in error rate, and as a result we get a more suitable network for the function we are trying to model. We attempt to minimise the mean-squared error (until a noticeable convergence occurs), and before the error starts to increase again.
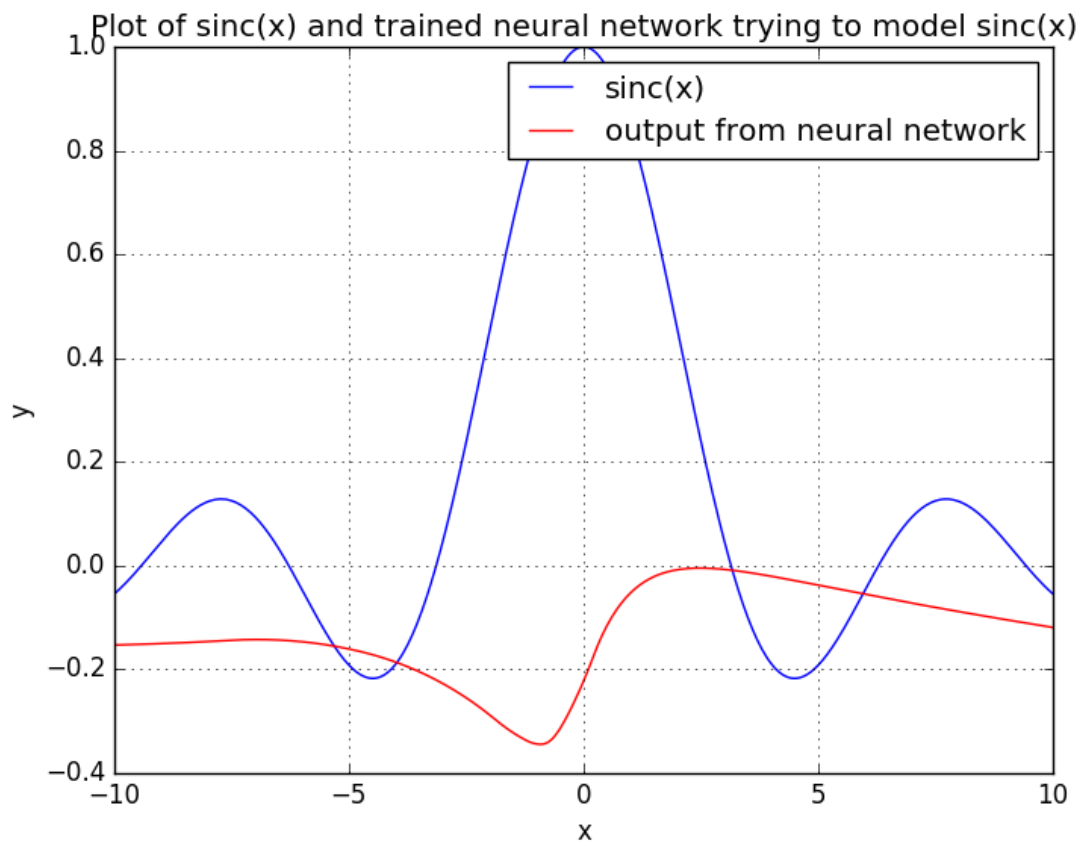
Figure 2: Neural Network trained with 20 Hidden Neurons

2   The growth function

2.1

2.2

2.3

2.4

2.5

3   VC-Dimension

3.1

3.2