

Assignment #5, 26 Sept 2014

Read Haykin Chapter #4, Sections 4.1-4.0, 4.13-4.17, 4.19  
Due: Tuesday 14 Oct 2014

Computer Problem Back Propagation Learning  
Due: Tuesday 21 Oct 2014 in PDF format by email.

a) Implement a simulation of the Backprop Learning Rule in a program and conduct the experiments described below. Write a concise report on the results of the experiments using figures and tables. For plots, label all axis, provide a short caption with figure number, and make sure all labels and numbers on the axis are readable. Provide title and table numbers for all tables as well. Include code in the same document as an appendix. Use PDF format. Make sure your name and assignment number are embedded in the file name.(i.e. SmithHW5.pdf )

b) Included in the email are files containing sets of training and testing data drawn from the same parent probability distributions. This is a four-class problem as can be seen from the class labels. Plot the training and testing data as two separate 2D scatter plots, clearly annotating the four classes of points with colors. Are these classes linearly separable? What form do the parent distributions appear to have from looking at the plots?

c) Set up a two input (plus bias node), four output network with multiple hidden layer network topologies, including appropriate bias nodes. Train and test a set of BP network architectures with the included data.

Specifically, run numerical experiments with one and two hidden layer network topologies. For each topology, vary the initial random weights and the number of hidden nodes in the layers at least three times, ranging from 2 to 50 or more per layer. It's your choice what numbers go in what layers. Summarize the training and testing performance in tables. In addition, for each network, produce several generalization plots calculated with "interesting" parameter settings. For these cases, produce plots of the decision regions (and therefore the boundaries) using a 2D sample grid, for multiple parameter setting for each type of algorithm.

Explain your stopping criterion and discuss the results.

d) Repeat part (c) with at least two (2) other training data orderings. I suggest you use a random shuffle algorithm to create different orders. Discuss the effects on convergence rates and the number of epochs required to reach a common terminal error. Include tabular performance results and an example generalization plot for each ordering

For the Training and Testing files, the column headings are: "c" the class label, "i" the sample index within the class, and "x1" and "x2" the feature vector components for input. Map the class labels to the desired output values (d1, d2, d3, d4) for the four output neurons as follows:

	d1	d2	d3	d4
Class 1:	1	0	0	0
Class 2:	0	1	0	0
Class 3:	0	0	1	0
Class 4:	0	0	0	1

The Training and Testing files each contain 800 samples, 200 per class.