Sample Solutions for Neural Networks Homework

September 8, 2002

1 Part 1

The network will adjust its weights to minimize the error function. The error is

$E = \frac{1}{2}\sum_i (T_i - Output_i)^2$

$= \frac{1}{2}[80(1 - Output_i)^2 + 20(0 - Output_i)^2]$

$= 50Output^2 - 80Output + 40$

The derivative of the error with respect to the single output $Output$ is

$\frac{dE}{dOutput} = 100Output - 80$.

Setting the derivative to zero, we find that $Output = 0.8$, the probabilistic prediction $P(Output = 1)$.

2 Part 2

The resulting networks would be different. The error used to adjust the weights of the network during training is based on the training examples given to it during that epoch. In the case of all the training examples presented in turn 1000 times, the error is based on the error over all the examples because they are all present during each epoch. In the case where each training example is presented individually for 1000 times, the first epoch error will only be based on the first training example, the second epoch error will only be based on the second training example, and so on.

The end result is the network will "forget" the original training example. As later training examples are used, the weights will move away from the weights necessary to classify the first training examples correctly.

3 Part 3

Networks of perceptrons are much more expressive than single perceptrons. For example, a single perceptron can only represent functions which are linearly seperable. Networks can represent more complex functions. However, perceptrons are guaranteed to find the global optimum because they represent a convex error surface, whereas networks may have local minima.