HW7 CS440

Steven Tang stang17

May 2016

Problem 1:

a) To get AND, set $w_1, w_2, w_3 = -1$ and $b = 3$

b) To get OR, set $w_1, w_2, w_3 = 1$ and $b = -1$

c) To get at-most-2, set $w_1, w_2, w_3 = -1$ and $b = 2$

d) 2-out-of-3 is not possible because there is no combination of w_1, w_2, w_3 and b that would make 0, 1, and 3 on get a value less than 0, and make 2 on get a value greater than or equal to 0.

Problem 2:

(a) Circles of radius r that are centered at the origin have a VC dimension of 1 because it can shatter one point that is either inside or outside the circle.

(b) Circles of radius r that may be arbitrarily centered at a point x_0 in 2D have a VC dimension of 3. This is because if we tried to shatter with 4 points that alternated between + and -, it would be impossible to shatter with a circle.

Problem 3:

(a) 5 hidden nodes

(b) Based on the graphs below, I tried running the neural net with 2,5, and 30 hidden nodes. For 2 nodes, the data converged to around the same value as 5 nodes, but contained more noise in the data. Trying with a large number of hidden nodes like 30 took a very long time to run, and the error decreased very slowly. Based on multiple trials, I decided to have 5 hidden nodes.

(c) Using 5 hidden neutrons, the error rate on the training set after 5,000 runs is close to 0.0%.

(d) Using 5 hidden neutrons, the error rate on the testing set after 5,000 runs is 1.52%.

(e) I did not divide the training set and used all of it as a learning set.

(f) I tested for convergence by checking if the past 500 runs produced an error difference that is less than 0.001.

(g) The value for the learning rate was 0.2.

(h) Using my chosen number of hidden neurons and learning rate, the neural net would converge at an average of 8,500 runs.

Graphs:

Average Error vs Iteration Run with 2 Hidden Nodes for 5,000 Runs

Error(%) vs Number of Runs

Average Error vs Iteration Run with 5 Hidden Nodes for 5,000 Runs

Average Error vs Iteration Run with 30 Hidden Nodes for 5,000 Runs

3