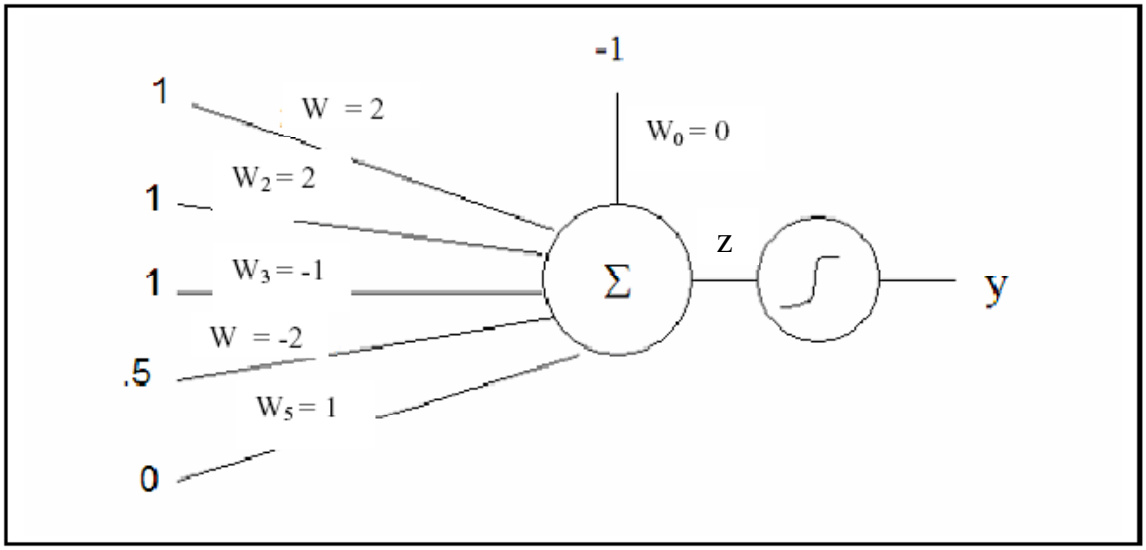
6.034 Recitation 7: Neural Nets (11/5/04)

LOrtiz (Orig. by KKoile)

A. Consider the following neural net (taken from the 2001 final exam).



The desired value, y*, is 1. Run one step of training to adjust the weights so that the net gets closer to that value. Assume a learning rate of 100, initial weights shown above, sample input shown above, and the standard error function, $E = 1/2(y^* - y)^2$. (Useful data point: sigmoid(2) = .9.) To run one step of training do the following:

1. Calculate the final (and in this case, only) node's contribution to the error. This quantity is δ, which Jake likes to call the "blame". It is the derivative of E with respect to the total input, and is equal to $(y - y^*)y(1 - y)$. (The formula's derivation via the chain rule: $dE/dz = (dE/dy)(dy/dz) = (y - y^*)(d(s(z))/dz) = (y - y^*)y(1 - y)$.)

z =

y = s(z) =

δ =

2. Figure out how to distribute the "blame" over contributors to the error. In this case, the weights on the input links contribute to the error. (If instead of input values, another layer of nodes contributed values, the error would be distributed to them first. See problem B.) Adjust the weights using the formula: $w_i' = w_i - r\delta x_i$. (This equation tells us to distribute the error in proportion to an input value, amount of blame, and r, a learning rate.) Assume r = 100.

w_0' =

w_1' =

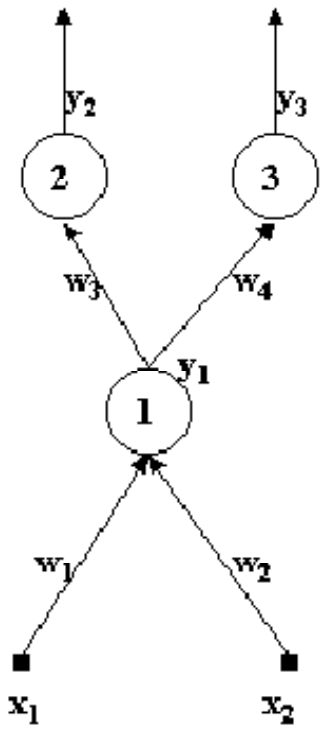
w_2' =

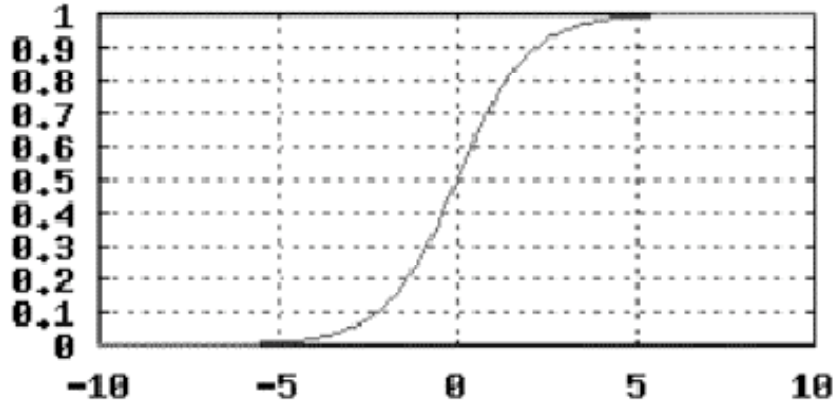
w_3' =

w_4' =

w_5' =

B. Now consider a more complex neural net (taken from exam 2, 2000):





Sigmoid values: s(-5) = 0; s (-4) = .02; s(-3) = .04; s(-2) = .12; s(-1) = .27; s(0) = .5; s(5) = 1; s(1) = .98
s(3) = .96; s(2) = .88; s(1) = .73

In order to distribute the error for this net to all the weights, we must not only distribute the blame δ directly from the final nodes 2 and 3 to w3 and w4, we must also distribute the blame from node 1 to w1 and w2. The blame for node 1 is calculated from the blame for nodes 2 and 3.

Assume a standard error function and values: $w_1 = 2$, $w_2 = -2$, $w_3 = 4$, $w_4 = 0$; $x_1 = 2$, $x_2 = 2$; $r = 1$; $y_2^* = 0$, $y_3^* = 1$, $r = 1$.

1. Calculate the y values and the δ values.

$y_1 = s(w_1x_1 + w_2x_2) =$

$y_2 = s(w_3y_1) =$

$y_3 = s(w_4y_1) =$

$\delta_2 = y_2 (1 - y_2) (y_2 - y_2^*) =$

$\delta_3 = y_3 (1 - y_3) (y_3 - y_3^*) =$

$\delta_1 = y_1(1 - y_1)(\Sigma\delta_k w_{1k}) = y_1(1 - y_1)(\delta_2 w_3 + \delta_3 w_4) =$

2. Calculate the new weights.

$w_1' = w_1 - r\delta_1 x_1 =$

$w_2' = w_2 - r\delta_1 x_2 =$

$w_3' = w_3 - r\delta_2 y_1 =$

$w_4' = w_4 - r\delta_3 y_1 =$