CS491 Quiz 7 Name Solution

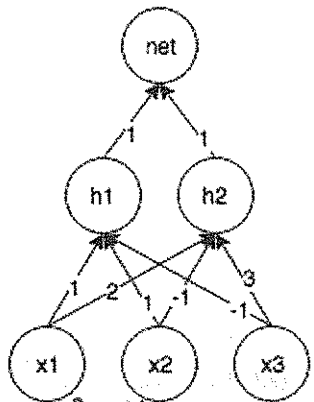
1. What would the neural network below predict for the following test data? That is, what is net for each test input? (3 points)
$t_1 = (1, 1, 1)$
$t_2 = (-5, -2, 1)$
$t_3 = (3, 1, 4)$

$net = h_1 + h_2$
$h_1 = x_1 + x_2 - x_3$
$h_2 = 2x_1 - x_2 + 3x_3$

① $h_1 = \cancel{2}$ $h_2 = 4$
$= 1$ $= 1$
$net = 2$

② $h_1 = -1$ $h_2 = -1$
$net = -2$

③ $h_1 = -1$ $h_2 = 1$
$net = 0$. or 2 if $sign(0) = +1$



h1=sign(<w1,x>)
h2=sign(<w2,x>)
net=<v,h>

<a,b> is a dot product b

2. Give an equation for the output of the network above, net, in terms of the inputs x_1, x_2, x_3. Do not use dot product notation. (2 points)

$net = \text{sign}(x_1 + x_2 - x_3) + \text{sign}(2x_1 - x_2 + 3x_3)$

3. Let's say I have a two layer neural network, with w_{ij} being the weight between the j^{th} input and the i^{th} hidden unit, and v_i being the weight between the i^{th} hidden unit and the final output unit. Each hidden unit applies the same non-linear function to its input: $h_i = f(w_i \bullet x)$. The loss function we are trying to minimize is $L = e^{-ynet}$ where net is the output of the network: $net = v \bullet h$. What are $\nabla_w L$ and $\nabla_v L$? Show your work. (5 points)

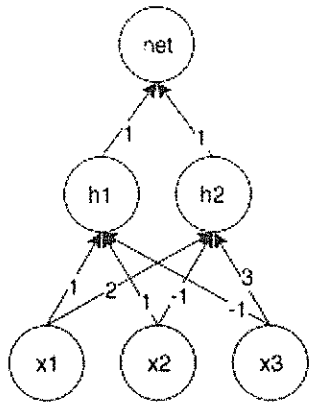
$L = e^{-y \cdot net}$ $net = v \cdot h = v \bullet f(w \bullet x)$

$\nabla_v L = -y f(w \bullet x) e^{-ynet} = -y h e^{-ynet}$

$\nabla_w L = \frac{\partial L}{\partial f} \cdot \frac{\partial f}{\partial w} = -y v f'(w \bullet x) \bullet x\, e^{-ynet}$

CS691 Quiz 7 Name Solution

For the following problems, use the network given below.



h1=sign(<w1,x>)
h2=sign(<w2,x>)
net=<v,h>

<a,b> is a dot product b

1. What would the neural network below predict for the following test data? That is, what is net for each test input? (2 points)
   $t_1 = (1, 1, 1)$ 2
   $t_2 = (-5, -2, 1)$ -2 Same as other side

2. Give an equation for the output of the network above, net, in terms of the inputs x_1, x_2, x_3. Do not use dot product notation. (2 points)

   Same as other side

3. We are trying to train the network above. Ignore the weights on the figure. We start with all $w_{ij} = 0$ and $v_i = 0$. Remember the w are the weights for the first layer and the v are the weights for the second layer. Assume the loss function is $L = \frac{1}{2}(y - net)^2$.

   (a) How do h_1 and h_2 behave in this setting? Why? (2 points)

   h_1 & h_2 will do the same thing. Their inputs have same values & backprop same gradient.

   (b) What is the value of net for the first input sample? (1 points)

   $net = h_1 + h_2 = 0$ if sign(0) = 0 if sign(0) = -1 net = -2

   (c) Give the equation for the weight update for v. (3 points) $net = v_1 h_1 + v_2 h_2 = v \cdot h$

   $\nabla_v L = (y - net) \cdot h$

   $v = v - \eta (y - net) h$ gradient descent.