Tutorial session 6. Neural Networks

Question 1:

x_1	x_2	y
1	2	1
3	1	0
1	1	1
2	0	0

dataset.

$$\underbrace{g(x) = \begin{cases} 1 & \text{if } s > 0 \\ 0, & \text{otherwise} \end{cases}}_{\text{step function.}}$$

So, in this exercise, they are asking us to use the correction method to update the weights. There are several methodologies to update our weights. The correction method is one of them. Just be careful, in other book the name of this method might change. The correction method is as follows:

$w_i \leftarrow w_i + \alpha(t - g(s))x_i$ → so here we use the output of the activation function to update our weights

Now we are going update our weights. We update instance by instance.

Tutorial session 06 : Neural Networks.

Correction method.

So the step function is $g(s) \begin{cases} 1 & \text{if } s > 0 \\ 0 & \text{otherwise.} \end{cases}$

We need to calculates s.

$s = \sum w_i x_i$, where x_0 is 1 because we are using a bias.

So for the first instance:

$s = w_0 \cdot x_0 + w_1 \cdot x_1 + w_2 \cdot x_2$

$s = 1 \cdot 1 + 1 \times 1 + 2 \times 2 = 6$

now it is time to update the weights

$\rightarrow w_0 \leftarrow w_0 + \alpha(t - g(s)) x_0$

$1 + 0.5(1 - \underbrace{g(6)}_{1}) 1 = 1$

$\rightarrow w_1 \leftarrow w_1 + \alpha(t - g(s)) x_1$

$1 + 0.5(1 - 1) 1 = 1$

$w_2 = 2$

1st iteration.

we use the weights we got in iteration 1

2nd Iteration.

$g(6) = 1$ and in this case it will be an erroneous prediction.

$s = 1 \times 1 + 1 \times 3 + 2 \times 1 = 6$ x_0

$w_0 = 1 + 0.5(0 - g(6)) \times 1 \rightarrow 0.5$

$w_1 = 1 + 0.5(0 - g(6)) \times \underset{x_1}{3} \rightarrow -0.5$

$w_2 = 2 + 0.5(0 - g(6)) \times 1 = 1.5$

3 iteration → we use the weights we got from previous iteration

$$s = \overset{w_0}{0.5} \times 1 + \overset{w_1}{-0.5} \times 1 + \overset{w_2}{1.5} \times 1 = 1.5$$

$$g(s) = 1$$

$$w_0 = 0.5 \times 0.5(1-1) \times 1 = 0.5$$
$$w_1 = -0.5 \times 0.5(1-1) \times 1 = -0.5$$
$$w_2 = 1.5 \times 0.5(1-1) \times 1 = 1.5$$

third ~~fourth~~ iteration

4 iteration, we use the weights we got from the previous iteration.

$$s = \overset{w_0}{0.5} \times \overset{x_0}{1} + \overset{w_1}{-0.5} \times \overset{\cancel{x_1}}{2} + \overset{w_2}{1.5} \times \overset{x_2}{0} = -0.5$$

$$g(-0.5) = 0$$

$t = 0$ ~~[illegible]~~ and $g(s)$ is 0. There won't be any update.

Tutorial session 6. Neural Networks.

Now we are using the delta rule to update our models. The delta rule is as follows:

$w_i \leftarrow w_i + \alpha\,(t - s)\,x_i$ → So here we will not use an activation function when updating

And just to make it clear, we use s for the update. For classification we use $g(s)$

1st Iteration.

$s = \overset{w_0}{1} \times \overset{x_0}{1} + \overset{w_1}{1} \times \overset{x_1}{1} + \overset{w_2}{2} \times \overset{x_2}{2} = 6 \rightarrow g(s) = 1$ ← this is the classification

$w_0 \leftarrow \overset{w_0}{1} + \overset{\alpha}{0.5}\,(\overset{t}{1} - \overset{s}{6}) \times \overset{x_0}{1} = -1.5$

$w_1 \leftarrow \overset{w_1}{1} + \overset{\alpha}{0.5}\,(\overset{t}{1} - \overset{s}{6}) \times \overset{x_1}{1} = 1.5$

$w_2 \leftarrow \overset{w_2}{2} + \overset{\alpha}{0.5}\,(\overset{t}{1} - \overset{s}{6}) \times \overset{x_2}{2} = -3$

2 iteration, we use the weights we got in the first iteration

$s = \overset{w_0}{-1.5} \times \overset{x_0}{1} + \overset{w_1}{-1.5} \times \overset{x_1}{3} + \overset{w_2}{-3} \times \overset{x_2}{1} = -9 \rightarrow g(s) = 0$

$w_0 = \underset{w_0}{-1.5} \times \underset{\alpha}{0.5}\,(\underset{t}{0} - \underset{s}{(-9)}) \times \underset{x_0}{1} = 3$

Tutorial session 06. Neural Networks.

→ 2 Iteration continuation

$$w_1 = \overset{w_1}{-1.5} + \overset{\alpha}{0.5} \times (\overset{t}{0} - (\overset{s}{-9})) \times \overset{x_1}{3} = 12$$

$$w_2 = \overset{w_2}{-3} + \overset{\alpha}{0.5} \times (\overset{t}{0} - (\overset{s}{-9})) \times \overset{x_2}{1} = 1.5$$

3 Iteration, we use the results from the second Iteration

$$s = \overset{w_0}{3} \times \overset{x_0}{1} + \overset{w_1}{12} \times \overset{x_1}{1} + \overset{w_2}{1.5} \times \overset{x_2}{1} = 16.5 \rightarrow g(s) = 1$$

classification

$$w_0 \leftarrow \overset{w_0}{3} + \overset{\alpha}{0.5}(\overset{t}{1} - \overset{s}{16.5}) \times \overset{x_0}{1} = -4.75$$

$$w_1 \leftarrow \overset{w_1}{12} + \overset{\alpha}{0.5}(\overset{t}{1} - \overset{s}{16.5}) \times \overset{x_1}{1} = 4.25$$

$$w_2 \leftarrow \overset{w_2}{1.5} + \overset{\alpha}{0.5}(\overset{t}{1} - \overset{s}{16.5}) \times \overset{x_2}{1} = -6.25$$

4 Iteration, we use the results we got from the third iteration

$$s = \overset{w_0}{-4.75} \times \overset{x_0}{1} + \overset{w_1}{4.25} \times \overset{x_1}{2} + (\overset{w_2}{-6.25}) \times \overset{x_2}{0} = 3.75$$

$$g(s) = g(3.75) = 1 \rightarrow \text{classification}$$

$$w_0 \leftarrow \overset{w_0}{-4.75} + \overset{\alpha}{0.5}(\overset{s}{0} - \overset{t}{3.75}) \times \overset{x_0}{1} = -6.625$$

$$w_1 \leftarrow \overset{w_1}{4.25} + \overset{\alpha}{0.5}(\overset{t}{0} - 3.75) \times \overset{x_1}{2} = 0.5$$

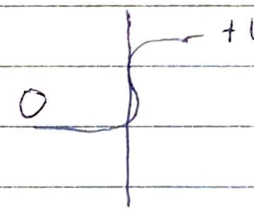
$$w_3 \leftarrow \overset{w_2}{-6.25} + \overset{\alpha}{0.5}(0 - \overset{t}{3.75}) \times \overset{x_2}{0} = -6.25$$

Tutorial session 06. Neural Networks

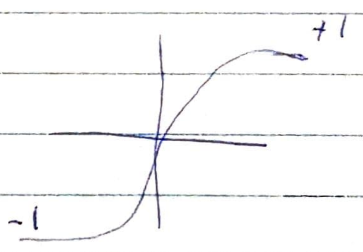
Now, we have a different update rule. the generalised delta rule. This is one of The most common update rules. Just be carefull because the name changes across The literature.

This update rule use to come together with flexible ~~transfer~~ activation functions such as The sigmoid and the Tanh activation function

sigmoid: $g(s) \quad \dfrac{1}{1+e^{s}}$



Tanh: $g(s) \quad 2\left(\dfrac{1}{1+e^{-2s}}\right)-1$



~~[illegible]~~

The updating rule is as follows:

$$w_i \leftarrow w_i + \alpha\,(t - g(s))\,x_i\,g(s)\,(1-g(s))$$

Tutorial session 06. Neural Networks

Question 3 continuation

1st Iteration

$$s = \overset{w_0}{1} \times \overset{x_0}{1} + \overset{w_1}{1} \times \overset{x_0}{1} + \overset{w_2}{2} \times \overset{x_2}{2}$$

$$g(s) = \underbrace{0.998}_{\text{The class would be 1}} \longrightarrow \frac{1}{1+e^{-6}}$$

The update would be like:

$$\overset{w_0}{w_0} \leftarrow 1 + \overset{\alpha}{0.5}(\overset{t\, -\, g(s)}{1 - 0.998}) \times \overset{x_0}{1} \times \overset{g(s)}{0.998} \times (\overset{1}{1} - \overset{g(s)}{0.988})$$

$$\leftarrow 1.002$$

$$w_1 \leftarrow 1 + 0.5(1 - 0.998) \times 1 \times (0.998) \times (1 - 0.988)$$

$$\leftarrow 1.002$$

$$w_2 \leftarrow 2 + 0.5(1 - 0.998) \times 2 (0.998) \times (1 - 0.988)$$

$$\leftarrow 2.004$$

Second iteration. We use the weights we got before

$$s = 1.002 \times 1 + 1.002 \times 3 + 2.004 \times 1$$

$$s = 6.0012$$

$$g(s) = 0.998$$

$$w_0 \leftarrow \overbrace{1.002 + 0.5(0 - 0.998) \times 1 \times (0.998)(1 - 0.998)}^{0.999}$$

$$w_1 \leftarrow 1.002 + 0.5(0 - 0.998) \times 3 \times (0.998)(1 - 0.998)$$

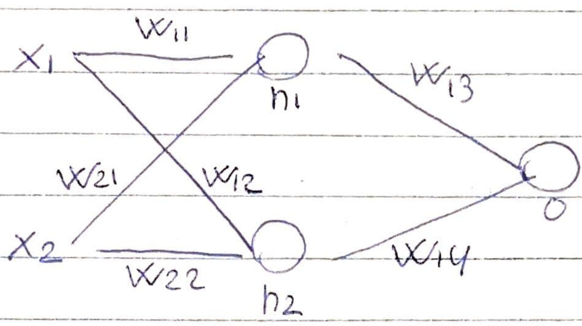
$$\leftarrow 0.997$$

$$w_2 \leftarrow 2.004 + 0.5(0 - 0.998) \times 1 \times (0.998)(1 - 0.998)$$

$$\leftarrow 1.999$$

Tutorial session 06. Neural Networks

Question 4



The mayor challenge here is that we have several layers. So we have to use the back propogation algorithm to update the weights w_{11}, w_{21}, w_{22}, and w_{12},

So we will get a prediction: $g(o)$
or $g(\text{bias} \times \text{bias weigh} + h_1 \times w_{13} + h_2 \times w_{14})$

The error is: $\underbrace{\overset{\text{actual label}}{y} - g(o)}$

assuming that we are using the sigmoid function, the derivative of this is

$$\Delta = (y - g(o)) \times (g(o) \times (1 - g(o)))$$

So the update of the middle layer neurons will be

$$w_{13} \leftarrow w_{13} + \alpha \overset{\text{output neuron}}{\Delta} g(h_1)$$
$$w_{13} \leftarrow w_{14} + \alpha \Delta g(h_2)$$

Tutorial session 06. Neural Networks.

So, now, we have to move further the error. The derivative of the error in h_1 is

$$\Delta h_1 = g(h_1) \times (1 - g(h_1)) \cdot W_{13} \cdot \Delta$$

So the updates will be

$$W_{11} \leftarrow W_{11} + \alpha \Delta h_1 \cdot x_1$$
$$W_{12} \leftarrow W_{12} + \alpha \Delta h_1 \cdot x_2$$