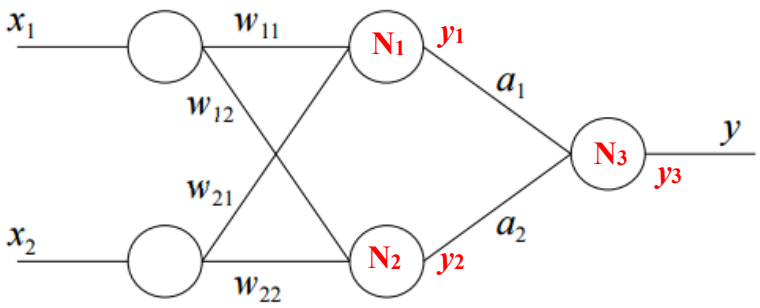
For the given dataset below, show the first iteration of training process of a MLP network which has one hidden layers with 2 neurons. Each neuron uses hyperbolic tangent activation function and the weights and parameters are w_{11}=1, w_{22}=1, w_{21}=1, w_{12}= 1, a_1= 1, a_2= -1, η = 0.5.

X_1	X_2	D
0	0	0
0	1	1
1	0	1
1	1	0



Hyperbolic tangent is $y = \frac{1-e^{-2net}}{1+e^{-2net}}$ and update is done by $\Delta w = \eta\, e\, (1- y^2)\, x$

If we say y_1, y_2, y_3 for output of N_1, N_2, N_3 respectively,

$\Delta a_1 = \eta\, e\, (1- y_3^2)\, y_1$

$\Delta a_2 = \eta\, e\, (1- y_3^2)\, y_2$

In order to update weights between input layer and hidden layer (w_{ij}), we should compute the derived error at first.

$e' = -e\,(1- y^2)$ then, $e_i' = e'.a_i$

The first sample (0,0:0)

$net_1 = x_1*w_{11}+x_2*w_{21} = 0*1+0*1 = 0$ $y_1 = 0$

$net_2 = x_1*w_{12}+x_2*w_{22} = 0*1+0*1 = 0$ $y_2 = 0$

$net_3 = y_1*a_1+y_2*a_2 = 0*1+0*(-1) = 0$ $y_3 = 0$ $e=d-y = 0-0 = 0$ no update

The second sample (0,1:1)

$net_1 = x_1*w_{11}+x_2*w_{21} = 0*1+1*1 = 1$ $y_1 = 0{,}762$

$net_2 = x_1*w_{12}+x_2*w_{22} = 0*1+1*1 = 1$ $y_2 = 0{,}762$

$net_3 = y_1*a_1+y_2*a_2 = 0{,}762*1+0{,}762*(-1) = 0$ $y_3 = 0$ $e=d-y = 1-0 = 1$

$\Delta a_1 = 0{,}5*1*(1-0^2)*0{,}762 = 0{,}381$ $a_1 = 1{,}381$

$\Delta a_2 = 0.5*1*(1-0^2)*0{,}762 = 0{,}381$ $a_2 = -0{,}619$

$e' = (-e).(1-y_3^2) = (-1).(1) = -1$

$e_i' = e'.a_i$ $e_1' = -1{,}381$ $e_2' = 0{,}619$

$\Delta w_{ij} = \eta\, e_j\, (1- y_j^2)\, x_i$

$\Delta w_{11} = 0{,}5*(-1{,}381)*(1-0{,}762^2)*0 = 0$ $w_{11} = 1$

$\Delta w_{12} = 0{,}5*0{,}619*(1-0{,}762^2)*0 = 0$ $w_{12} = 1$

$\Delta w_{21} = 0{,}5*(-1{,}381)*(1-0{,}762^2)*1 = -0{,}29$ $w_{21} = 0{,}71$

$\Delta w_{22} = 0{,}5*0{,}619*(1-0{,}762^2)*1 = 0{,}13$ $w_{22} = 1{,}13$

The third sample (1,0:1)

$net_1 = x_1*w_{11}+x_2*w_{21} = 1*1+0*0{,}71 = 1$ $y_1 = 0{,}762$

$net_2 = x_1*w_{12}+x_2*w_{22} = 1*1+0*1{,}13 = 1$ $y_2 = 0{,}762$

$net_3 = y_1*a_1+y_2*a_2 = 0{,}762*1{,}381+0{,}762*(-0{,}619) = 0{,}58$ $y_3 = 0{,}523$ $e=d-y = 1-0 = 0{,}477$

$\Delta a_1 = 0{,}5*0{,}477*(1-0{,}523^2)*0{,}762 = 0{,}132$ $a_1 = 1{,}513$

$\Delta a_2 = 0,5*0,477*(1-0,523^2)*0,762 = 0,132$ $a_2 = -0,487$

$e' = -e*(1-y_3^2) = -0,477*0,523^2 = -0,347$

$e_i' = e'.a_i$ $e_1' = -0,525$ $e_2' = 0,169$

$\Delta w_{ij} = \eta\, e_j'\,(1 - y_j^2)\, x_i$

$\Delta w_{11} = 0,5*(-0,525)*(1-0,762^2)*1 = -0,11$ $w_{11} = 0,89$

$\Delta w_{12} = 0,5*0,169*(1-0,762^2)*1 = 0,035$ $w_{12} = 1,035$

$\Delta w_{21} = 0,5*(-0,525)*(1-0,762^2)*0 = 0$ $w_{21} = 0,71$

$\Delta w_{22} = 0,5*0,169*(1-0,762^2)*0 = 0$ $w_{22} = 1,13$

The fourth sample (1,1:0)

$net_1 = x_1*w_{11}+x_2*w_{21} = 1*0,89+1*0,71 = 1.6$ $y_1 = 0,922$

$net_2 = x_1*w_{12}+x_2*w_{22} = 1*1,035+1*1,13 = 2,166$ $y_2 = 0,974$

$net_3 = y_1*a_1+y_2*a_2 = 0.922*1,513+0.974*(-0,487) = 0,92$ $y_3 = 0,726$ $e = d - y = -0,726$

$\Delta a_1 = 0.5*(-0,726)*(1-0,726^2)*0,922 = -0,158$ $a_1 = 1,355$

$\Delta a_2 = 0.5*(-0,726)*(1-0,726^2)*0,974 = -0,167$ $a_2 = -0,654$

$e' = -e*(1-y_3^2) = -1*(1-0,726^2) = 0,343$

$e_i' = e'.a_i$ $e_1' = 0,465$ $e_2' = -0,225$

$\Delta w_{ij} = \eta\, e_j'\,(1 - y_j^2)\, x_i$

$\Delta w_{11} = 0,5*0,465*(1-0,922^2)*1 = 0,035$ $w_{11} = 0,925$

$\Delta w_{12} = 0,5*(-0,225)*(1-0,974^2)*1 = -0,006$ $w_{12} = 1,03$

$\Delta w_{21} = 0,5*0,465*(1-0,922^2)*1 = 0,035$ $w_{21} = 0,745$

$\Delta w_{22} = 0,5*(-0,225)*(1-0,974^2)*1 = -0,006$ $w_{22} = 1,124$