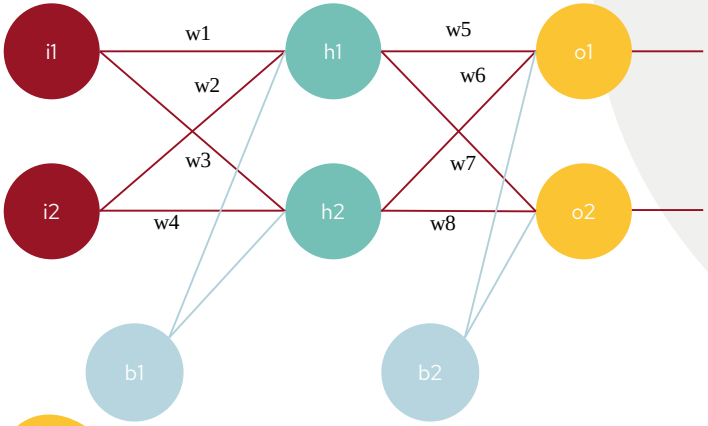
Step by step Backpropagation

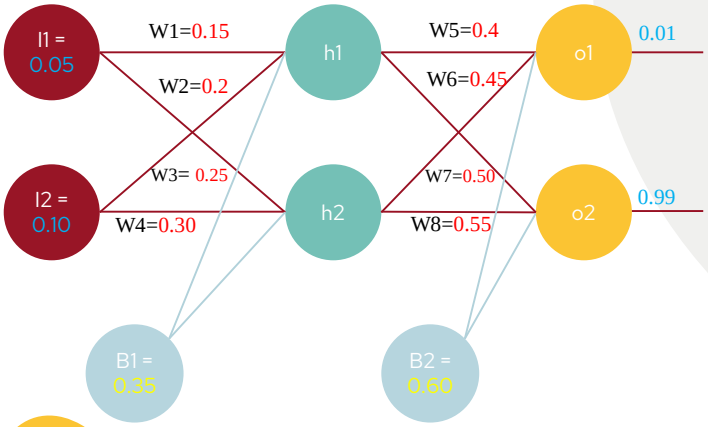
Our Neural Network



Our Goal

i1 = 0.05 ⟹ o1 = 0.01

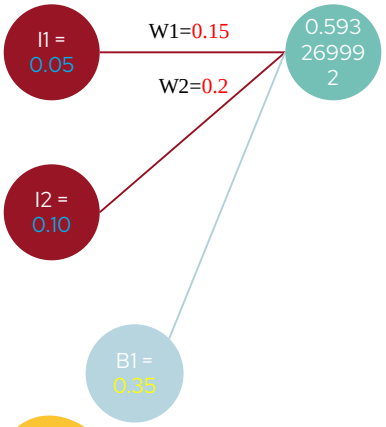
i2 = 0.10 ⟹ o2 = 0.99



Feed forward

To begin, lets see what the neural network currently predicts given the weights and biases above and inputs of 0.05 and 0.10. To do this we'll feed those inputs forward though the network.

Calculating h1



$$net_{h1} = w_1 * i_1 + w_2 * i_2 + b_1 * 1$$

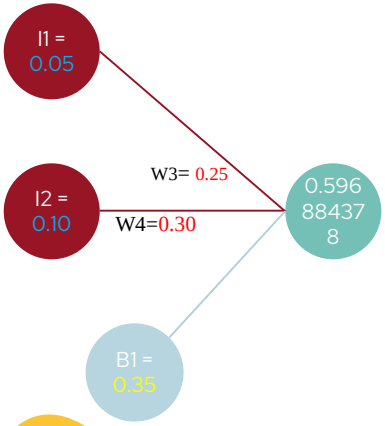
$$net_{h1} = 0.15 * 0.05 + 0.2 * 0.1 + 0.35 * 1 = 0.3775$$

$$out_{h1} = \frac{1}{1+e^{-net_{h1}}} = \frac{1}{1+e^{-0.3775}} = 0.593269992$$

Calculating h2

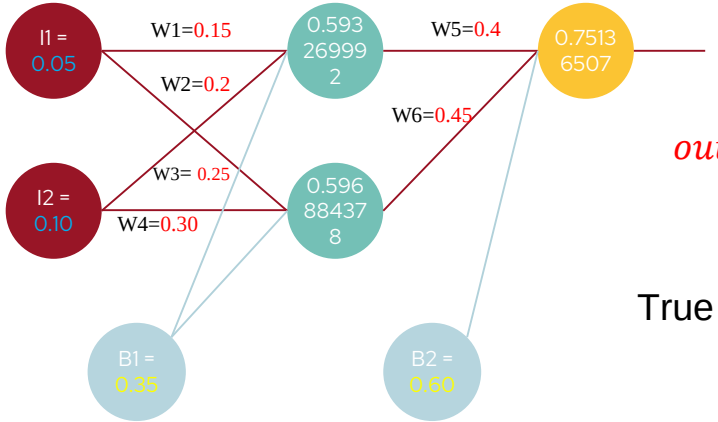
$$net_{h2} = w_3 * i_3 + w_4 * i_4 + b_1 * 1$$

$$net_{h2} = 0.25 * 0.05 + 0.3 * 0.1 + 0.35 * 1 = 0.3925$$



$$out_{h2} = \frac{1}{1+e^{-net_{h2}}} = \frac{1}{1+e^{-0.3925}} = 0.596884378$$

Calculating o1



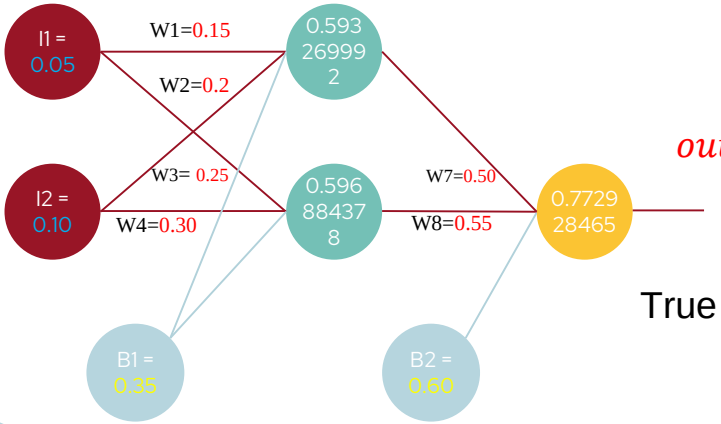
$$neto1 = w5 * outh1 + w6 * outh2 + b2 * 1$$

$neto1$ = 0.4 * 0.593269992 + 0.45 * 0.596884378 + 0.6 * 1
= 1.105905967

$$out_{o1} = \frac{1}{1+e^{-net_{o1}}} = \frac{1}{1+e^{-1.105905967}} = 0.75136507$$

True output = 0.01 != Our output = 0.75136507

Calculating o2



$$neto2 = w7 * outh1 + w8 * outh2 + b2 * 1$$

$neto2$ = 0.5 $*$ 0.59326992 + 0.55 * 0.596884378 + 0.6 * 1
= 1.22492136

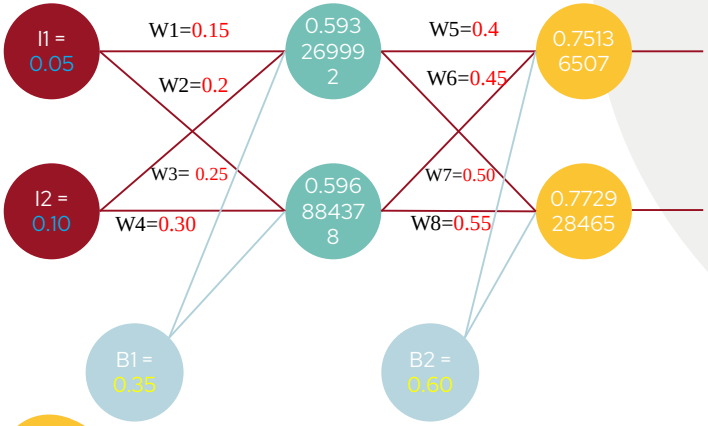
$$out_{o2} = \frac{1}{1+e^{-net_{o2}}} = \frac{1}{1+e^{-1.22492136}} = 0.772928465$$

True output = 0.99 $!=$ Our output = 0.772928465

Our neural network in first train

i1 = 0.05 ⇒ o1 = 0.75136507

i2 = 0.10 ⇒ o2 = 0.772928465



Calculating the total error1

$$E_{total} = \sum \frac{1}{2} (target - output)^2$$

True output = 0.01 != Our output = 0.75136507

$$E_{o_1} = \frac{1}{2} \left(target_{o_1} - output_{o_1} \right)^2 = \frac{1}{2} (0.01 - 0.75136507)^2 = 0.274811083$$

Calculating the total error2

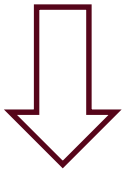
$$E_{total} = \sum \frac{1}{2} (target - output)^2$$

True output = 0.99 != Our output = 0.772928465

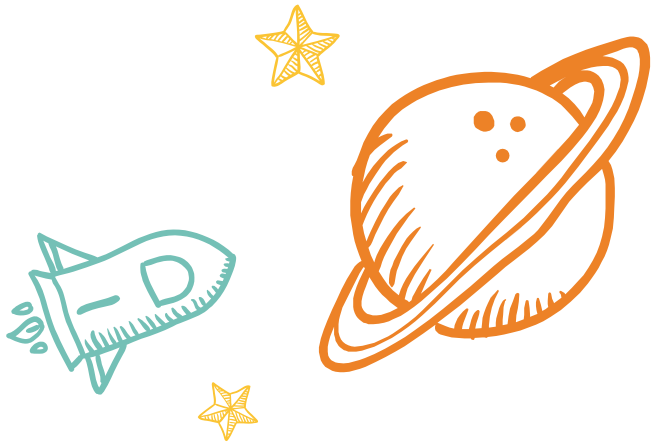
$$E_{o_2} = \frac{1}{2} (target_{o_2} - output_{o_2})^2 = \frac{1}{2} (0.99 - 0.772928465)^2 = 0.023560026$$

Calculating the total error3

$$E_{total} = E_{o_1} + E_{o_2}$$



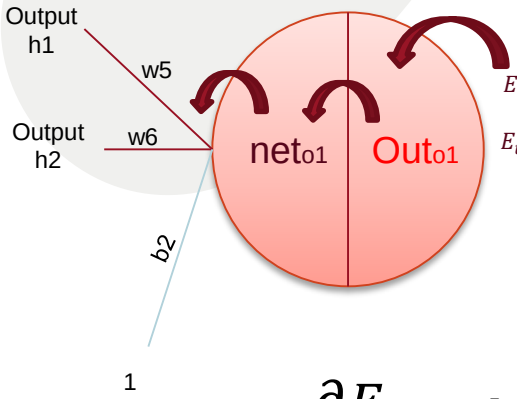
$$E_{total} = 0.274811083 + 0.023560026 = 0.298371109$$



The backward pass

Our goal with backpropagation is to update each of the weights in the network so that they cause the actual output to be closer the target output, thereby minimizing the error for each output neuron and the network as a whole.

Calculating the w5 change1

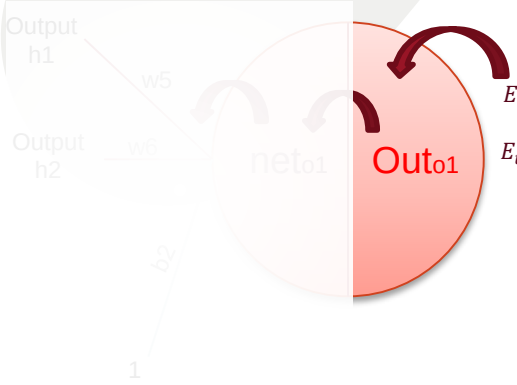


$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial w_5}$$

Calculating the w5 change2



$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

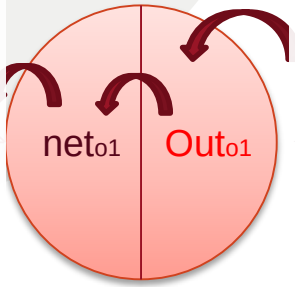
$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial out_{o_1}} = 2 * \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^{2-1} * -1 + 0$$

$$\frac{\partial E_{total}}{\partial out_{o_1}} = -(target_{o_1} - output_{o_1}) = -(0.01 - 0.75136507)$$

$$= 0.74136507$$

Calculating the w5 change2

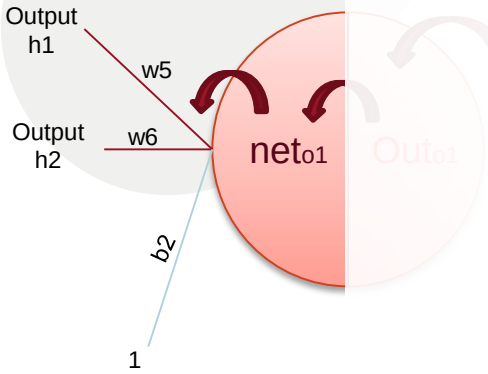


$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o_1}(1- out_{o_1})$$

$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o_1}(1- out_{o_1}) = 0.75136507\ (\ 1 - 0.75136507\ )$$

$$= 0.186815602$$

Calculating the w5 change2



$$\frac{\partial net_{o_1}}{\partial w_5} = 1 * out_{h1} * w_5^{(1-1)} + 0 + 0$$

$$\frac{\partial net_{o_1}}{\partial w_5} = out_{h1} = 0.593269992$$

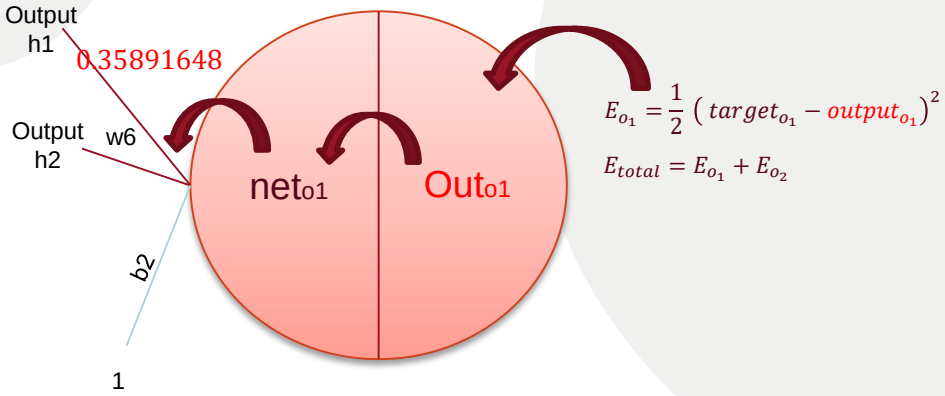
Calculating the w5 change2

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial w_5}$$

$$\frac{\partial E_{total}}{\partial w_5} = 0.74136507 * 0.186815602 * 0.593269992 = 0.082167041$$

$$w_5^+ = w_5 - \eta * \frac{\partial E_{total}}{\partial w_5} = 0.4 - 0.5 * 0.082167041 = 0.35891648$$

Calculating the w5 change1



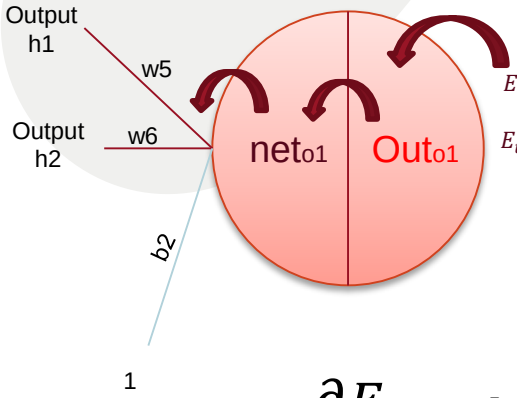
$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

Old w5 = 0.4

New w5 = 0.35891648

Calculating the w6 change1

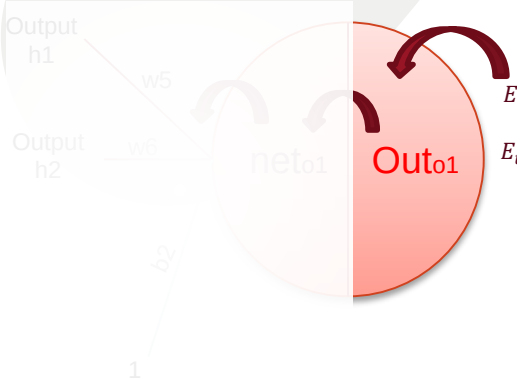


$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial w_6} = \frac{\partial E_{total}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial w_6}$$

Calculating the w6 change2



$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

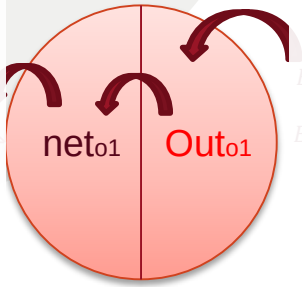
$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial out_{o_1}} = 2 * \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^{2-1} * -1 + 0$$

$$\frac{\partial E_{total}}{\partial out_{o_1}} = -(target_{o_1} - output_{o_1}) = -(0.01 - 0.75136507)$$

$$= 0.74136507$$

Calculating the w6 change3

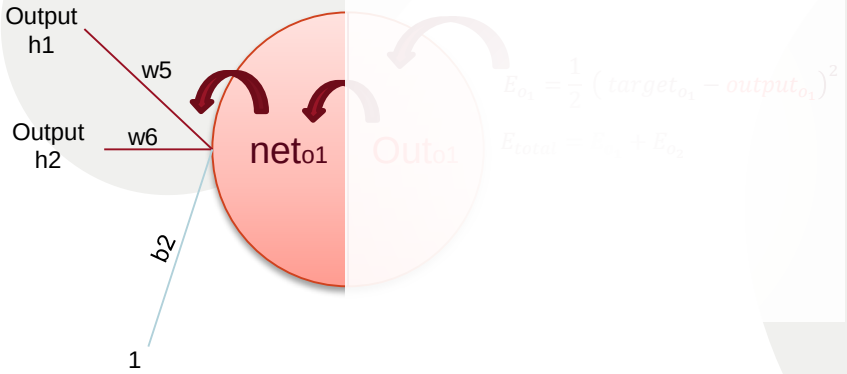


$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o_1}(1- out_{o_1})$$

$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o_1}(1- out_{o_1}) = 0.75136507\ (\ 1 - 0.75136507\ )$$

$$= 0.186815602$$

Calculating the w6 change4

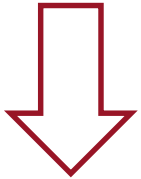


$$\frac{\partial net_{o_1}}{\partial w_6} = 1 * out_{h2} * w_6^{(1-1)} + 0 + 0$$

$$\frac{\partial net_{o_1}}{\partial w_6} = out_{h2} = 0.596884378$$

Calculating the w6 change5

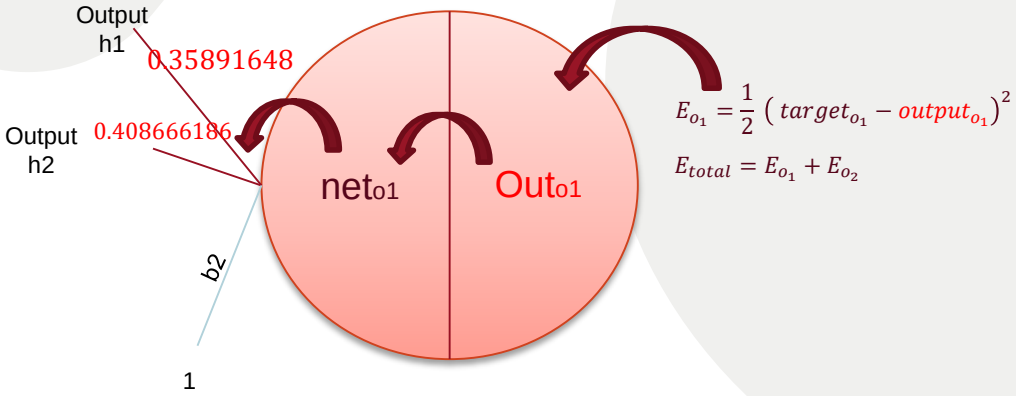
$$\frac{\partial E_{total}}{\partial w_6} = \frac{\partial E_{total}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial w_6}$$



$$\frac{\partial E_{total}}{\partial w_6} = 0.74136507 * 0.186815602 * 0.596884378 = 0.08266762$$

$$w_6^+ = w_6 - \eta * \frac{\partial E_{total}}{\partial w_6} = 0.45 - 0.5 * 0.08266762$$
$$= 0.408666186$$

Calculating the w6 change6



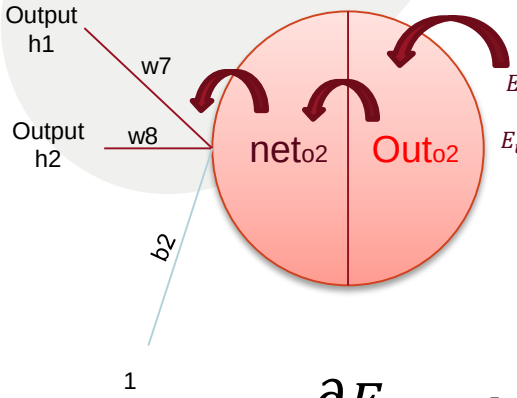
$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

Old w6 = 0.45

New w6 = 0.408666186

Calculating the w7 change1

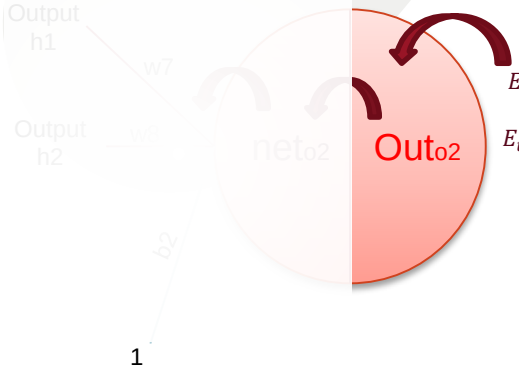


$$E_{o_1} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial w_7} = \frac{\partial E_{total}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}} * \frac{\partial net_{o_2}}{\partial w_7}$$

Calculating the w7 change2



$$E_{o_2} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

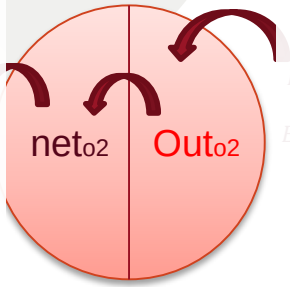
$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial out_{o_2}} = 2 * \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^{2-1} * -1 + 0$$

$$\frac{\partial E_{total}}{\partial out_{o_2}} = -(target_{o_2} - output_{o_2}) = -(0.99 - 0.772928465)$$

$$= -0.21707153$$

Calculating the w7 change3

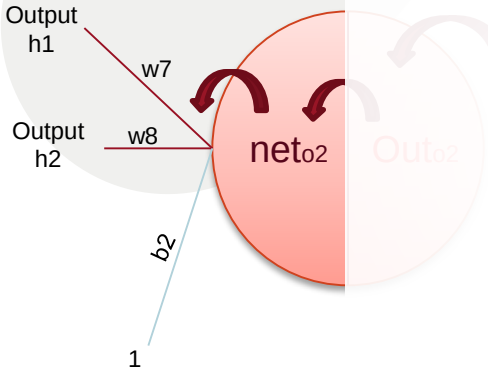


$$\frac{\partial out_{o_2}}{\partial net_{o_2}} = out_{o_2}(1- out_{o_2})$$

$$\frac{\partial out_{o_2}}{\partial net_{o_2}} = out_{o_2}(1- out_{o_2}) = 0.772928465(1 – 0.772928465)$$

$$= 0.17551005$$

Calculating the w7 change4



$$\frac{\partial net_{o_2}}{\partial w_7} = 1 * out_{h1} * w_7^{(1-1)} + 0 + 0$$

$$\frac{\partial net_{o_2}}{\partial w_7} = out_{h1} = 0.593269992$$

Calculating the w7 change5

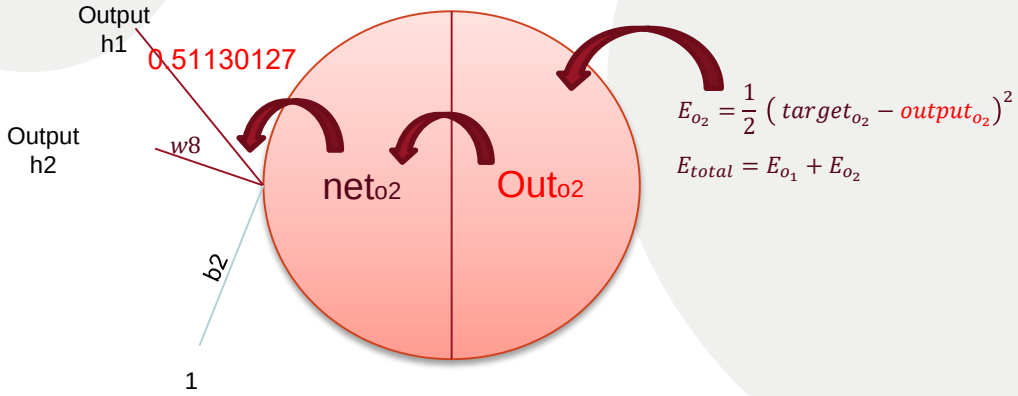
$$\frac{\partial E_{total}}{\partial w_7} = \frac{\partial E_{total}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}} * \frac{\partial net_{o_2}}{\partial w_7}$$

$$\frac{\partial E_{total}}{\partial w_7} = -0.21707153 * 0.17551005 * 0.593269992 = -0.02260253$$

$$w_7^+ = w_7 - \eta * \frac{\partial E_{total}}{\partial w_7}$$

$$= 0.5 - 0.5 * (-0.02260253) = 0.51130127$$

Calculating the w7 change6



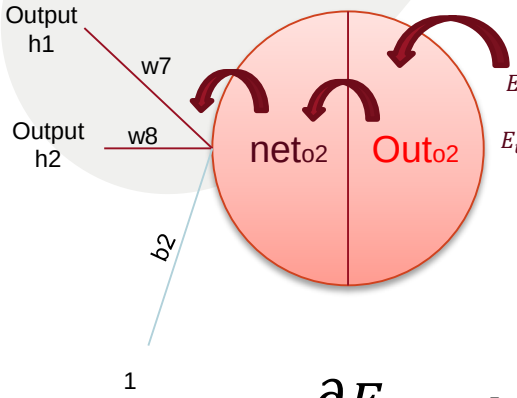
$$E_{o_2} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

Old w7 = 0.5

New w7 = 0.51130127

Calculating the w8 change1

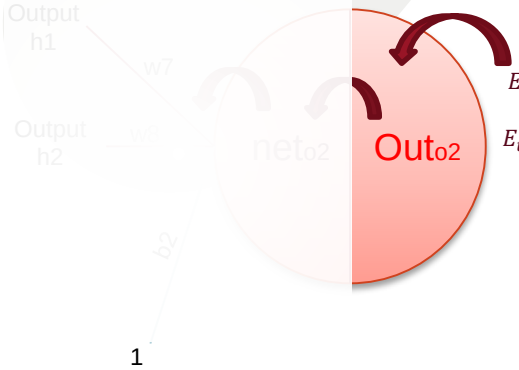


$$E_{o_1} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial w_8} = \frac{\partial E_{total}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}} * \frac{\partial net_{o_2}}{\partial w_8}$$

Calculating the w8 change2



$$E_{o_2} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

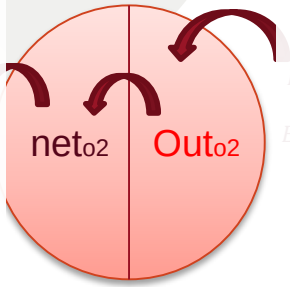
$$E_{total} = E_{o_1} + E_{o_2}$$

$$\frac{\partial E_{total}}{\partial out_{o_2}} = 2 * \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^{2-1} * -1 + 0$$

$$\frac{\partial E_{total}}{\partial out_{o_2}} = -(target_{o_2} - output_{o_2}) = -(0.99 - 0.772928465)$$

$$= -0.21707153$$

Calculating the w8 change3

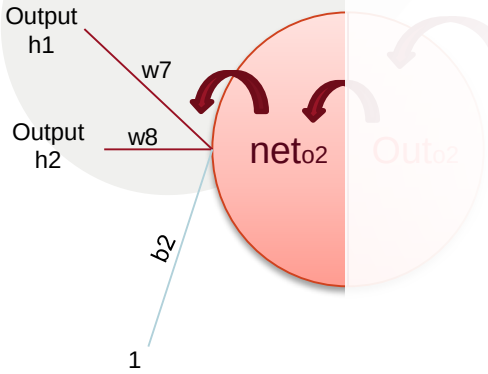


$$\frac{\partial out_{o_2}}{\partial net_{o_2}} = out_{o_2}(1- out_{o_2})$$

$$\frac{\partial out_{o_2}}{\partial net_{o_2}} = out_{o_2}(1- out_{o_2}) = 0.772928465(1 – 0.772928465)$$

$$= 0.17551005$$

Calculating the w8 change4

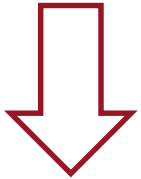


$$\frac{\partial net_{o_2}}{\partial w_8} = 1 * out_{h2} * w_8^{(1-1)} + 0 + 0$$

$$\frac{\partial net_{o_2}}{\partial w_8} = out_{h2} = 0.596884378$$

Calculating the w8 change5

$$\frac{\partial E_{total}}{\partial w_8} = \frac{\partial E_{total}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}} * \frac{\partial net_{o_2}}{\partial w_8}$$

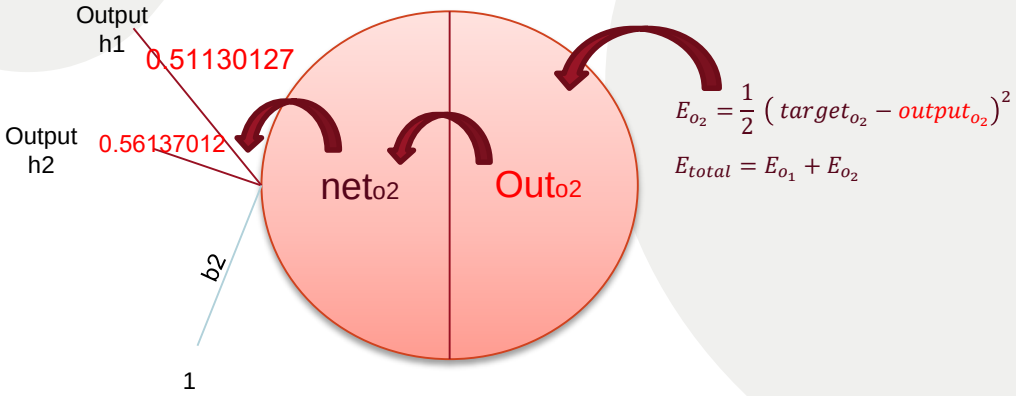


$$\frac{\partial E_{total}}{\partial w_8} = -0.21707153 * 0.17551005 * 0.596884378 = -0.02274024$$

$$w_8^+ = w_8 - \eta * \frac{\partial E_{total}}{\partial w_8}$$

$$= 0.55 - 0.5 * (-0.02274024) = 0.56137012$$

Calculating the w8 change6



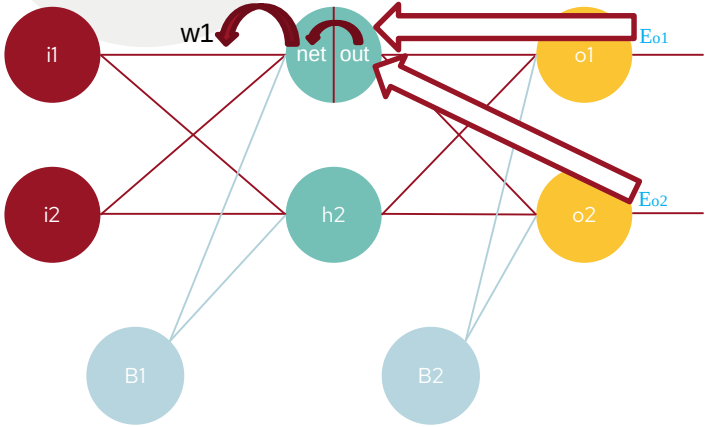
$$E_{o_2} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2$$

$$E_{total} = E_{o_1} + E_{o_2}$$

Old w8 = 0.55

New w8 =0.56137012

The backward pass for w1



$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$$

$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\boldsymbol{\partial E_{o_1}}}{\boldsymbol{\partial out_{h_1}}} + \frac{\partial E_{o_2}}{\partial out_{h_1}}$$

Calculating the w1 change1

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial out_{h_1}}$$

$$\frac{\partial E_{o_1}}{\partial net_{o_1}}$$

$$\frac{\partial E_{o_1}}{\partial net_{o_1}} = \frac{\partial E_{o_1}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}}$$

$$E_{o_1} = \frac{1}{2}\left(target_{o_1} - output_{o_1}\right)^2 \longrightarrow \frac{\partial E_{o_1}}{\partial out_{o_1}} = 2 * \frac{1}{2} - \left(target_{o_1} - output_{o_1}\right)^{2-1} = 0.74136507$$

Calculating the w1 change3

$$\frac{\partial E_{o_1}}{\partial net_{o_1}}$$

$$\frac{\partial E_{o_1}}{\partial net_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}}$$

$out_{o1} = \frac{1}{1+e^{-net_{o1}}}$

$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o1}(1 - out_{o1})$$

$= 0.75136507 * (1 - 0.75136507)$

$= 0.186815602$

$$\frac{\partial E_{o_1}}{\partial net_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{o_1}} * \frac{\partial out_{o_1}}{\partial net_{o_1}} = 0.74136507 * 0.186815602 = 0.138498562$$

Calculating the w1 change4

$$\frac{\partial net_{o1}}{\partial out_{h1}}$$

$$net_{o1} = w_5 * out_{h1} + w_6 * out_{h2} + b_2 * 1$$

$$\frac{\partial net_{o1}}{\partial out_{h1}} = w_5 * out_{h1} + 0 + 0 = w_5 = 0.4$$

Calculating the w1 change5

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial net_{o_1}} * \frac{\partial net_{o_1}}{\partial out_{h_1}}$$

$\frac{\partial E_{o_1}}{\partial out_{h_1}}$

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = 0.138498562 * 0.4 = 0.055399425$$

Calculating the w1 change6

$$\frac{\partial E_{o_2}}{\partial net_{o_2}} = \frac{\partial E_{o_2}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}}$$

$$\frac{\partial E_{o_2}}{\partial net_{o_2}}$$

$$out_{o2} = \frac{1}{1+e^{-net_{o2}}} = out_{o2}(1- out_{o2}) = 0.772928465\ (1-0.772928465)$$

$$= 0.17551005$$

$$E_{o_2} = \frac{1}{2}\left(target_{o_2} - output_{o_2}\right)^2 \rightarrow \frac{\partial E_{o_2}}{\partial out_{o_1}} = 2 * \frac{1}{2} -\left(target_{o_2} - output_{o_2}\right)^{2-1} = -0.217071535$$

$$\frac{\partial E_{o_2}}{\partial net_{o_2}} = \frac{\partial E_{o_2}}{\partial out_{o_2}} * \frac{\partial out_{o_2}}{\partial net_{o_2}}$$

$$= -0.217071535 * 0.17551005 = -0.03809823$$

Calculating the w1 change7

$$net_{o2} = w_7 * out_{h1} + w_8 * out_{h2} + b_2 * 1$$

$$\frac{\partial net_{o_2}}{\partial out_{h_1}} = w_7 * out_{h1} + 0 + 0 = w_8 = 0.5$$

$$\frac{\partial net_{o_2}}{\partial out_{h_1}}$$

Calculating the w1 change8

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial out_{h_1}}$$

$$\frac{\partial E_{o_1}}{\partial out_{h_1}}$$

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = -0.03809823 * 0.5 = -0.01904911$$

Calculating the w1 change9

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$$

$$\frac{\partial E_{total}}{\partial out_{h_1}}$$

$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{o_2}}{\partial out_{h_1}}$$

$$\frac{\partial E_{total}}{\partial out_{h_1}} = 0.055399425 + (-0.01904911)$$

$$= 0.036350306$$

Calculating the w1 change10

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$$

$$\frac{\partial out_{h_1}}{\partial net_{h_1}}$$

$$\frac{\partial out_{h_1}}{\partial net_{h_1}} = out_{h_1}\left(1 - out_{h_1}\right) = 0.59327(1 - 0.59327)$$

$$= 0.241300709$$

Calculating the w1 change11

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$$

$$\frac{\partial net_{h_1}}{\partial w_1}$$

$$net_{h1} = w_1 * i_1 + w_2 * i_2 + b_1 * 1$$

$$\frac{\partial net_{h_2}}{\partial w_1} = i_1 = 0.05$$

Calculating the w1 change12

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$$

$\frac{\partial E_{total}}{\partial w_1}$

$$\frac{\partial E_{total}}{\partial w_1} = 0.036350306 * 0.2413 * 0.05 = 0.000438568$$

$$w_1^+ = w_1 - \eta * \frac{\partial E_{total}}{\partial w_1} = 0.15 - 0.5 * 0.000438568 = 0.149780716$$

Calculating the w2 change1

$$\frac{\partial E_{total}}{\partial w_2} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_2}$$

$\frac{\partial E_{total}}{\partial w_2}$

$$\frac{\partial E_{total}}{\partial w_2} = 0.03635013 + 0.2413007086 + 0.1 = 0.000877131$$

$$w_2^+ = w_2 - \eta * \frac{\partial E_{total}}{\partial w_2} = 0.20 - 0.5 * 0.000877131 = 0.19956143$$

Calculating the w3 change1

$$\frac{\partial E_{total}}{\partial w_3} = \frac{\partial E_{total}}{\partial out_{h_2}} * \frac{\partial out_{h_2}}{\partial net_{h_2}} * \frac{\partial net_{h_2}}{\partial w_3}$$

$$\frac{\partial E_{total}}{\partial w_3}$$

$$\frac{\partial E_{total}}{\partial w_3} = 0.00043856$$

$$w_3^+ = w_3 - \eta * \frac{\partial E_{total}}{\partial w_3} = 0.25 - 0.5 * 0.00043856 = 0.24978$$

Calculating the w4 change

$$\frac{\partial E_{total}}{\partial w_4} = \frac{\partial E_{total}}{\partial out_{h_2}} * \frac{\partial out_{h_2}}{\partial net_{h_2}} * \frac{\partial net_{h_2}}{\partial w_4}$$

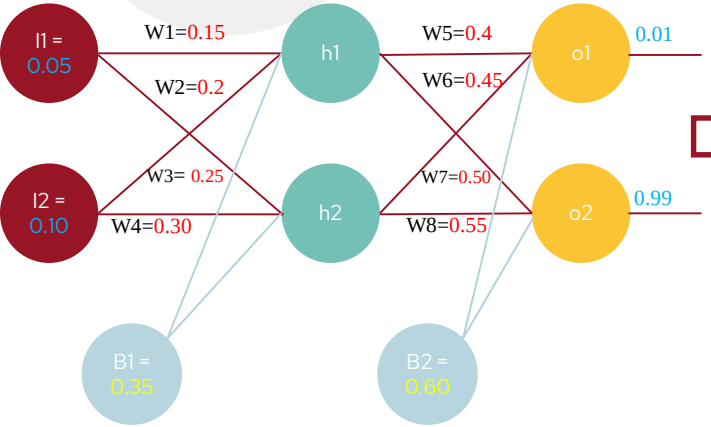
$$\frac{\partial E_{total}}{\partial w_4}$$

$$\frac{\partial E_{total}}{\partial w_4} = 0.041369954 + 0.240613417 + 0.1 = 0.00099541$$

$$w_4^+ = w_4 - \eta * \frac{\partial E_{total}}{\partial w_4} = 0.3 - 0.5 * 0.00099541 = 0.29950229$$

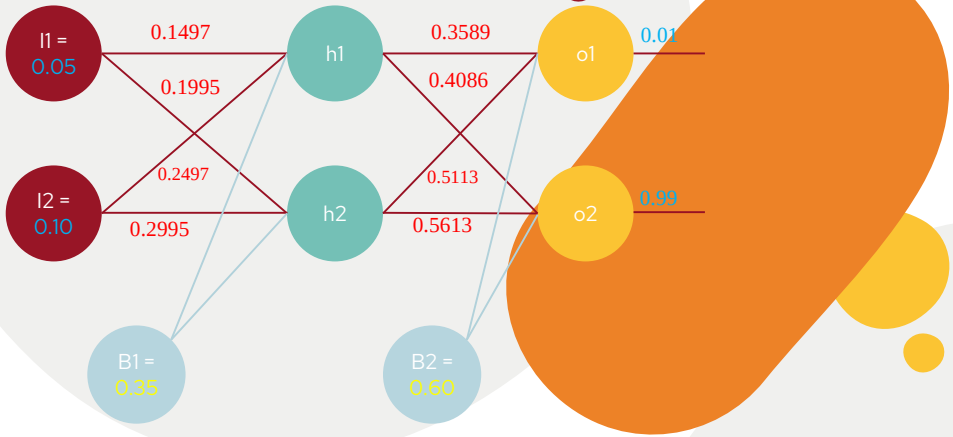
Diffrence in one run

old



Error = 0.298371109

new



Error = 0.291027924

After 10000 times reapiting

input	New output	output
i = 0.05	o1 = 0.015912196	o1 = 0.01
i = 0.10	o2 = 0.984065734	o2 = 0.99

Thanks!

Any questions?

You can find me at
mahdi.2000musa@gmail.com