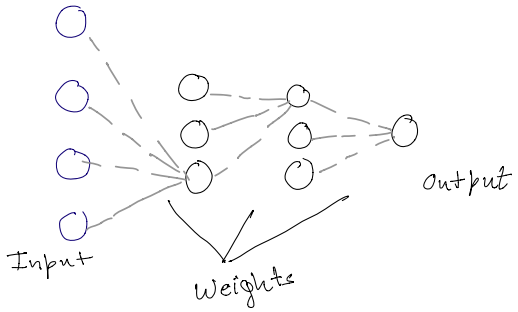
Consider a ANN below (not showing all connections for simplicity)



- Lets analyze the single cross section of the above ANN.

Given:

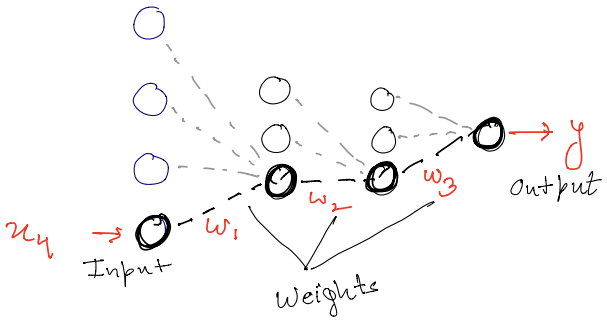
$w_1 = 0.1$

$w_2 = -0.2$

$w_3 = 0.2$

$x_4 = 1$

$y = ?$



- Since I'm not giving you all the weights of this ANN, lets assume final output of the ANN I.e. y = 3 and what we want is 2.
- Now, we shall do backward pass to change w1, w2, w3 to reduce the discrepancy between actual output (y = 3) vs desired output (y_cap = 2)
- We shall use squared error as our loss function, denoted by E.

$$E = (y - \hat{y})^2$$

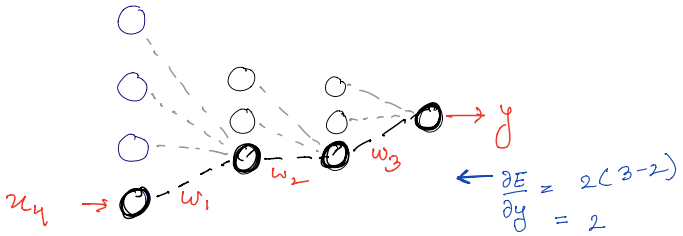
E → loss; y → actual; $\hat{y}$ → desired

- Lets perform the backward pass first we need:

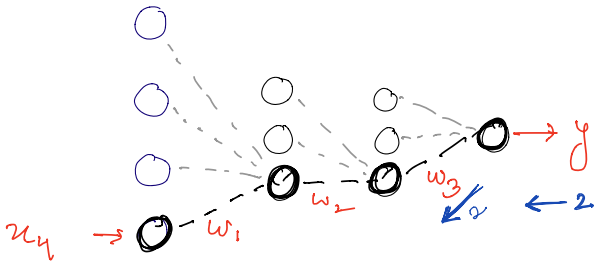
$$\frac{\partial E}{\partial y} = \frac{\partial}{\partial y}(y-\bar{y})^2 = \boxed{2(y-\bar{y})}$$

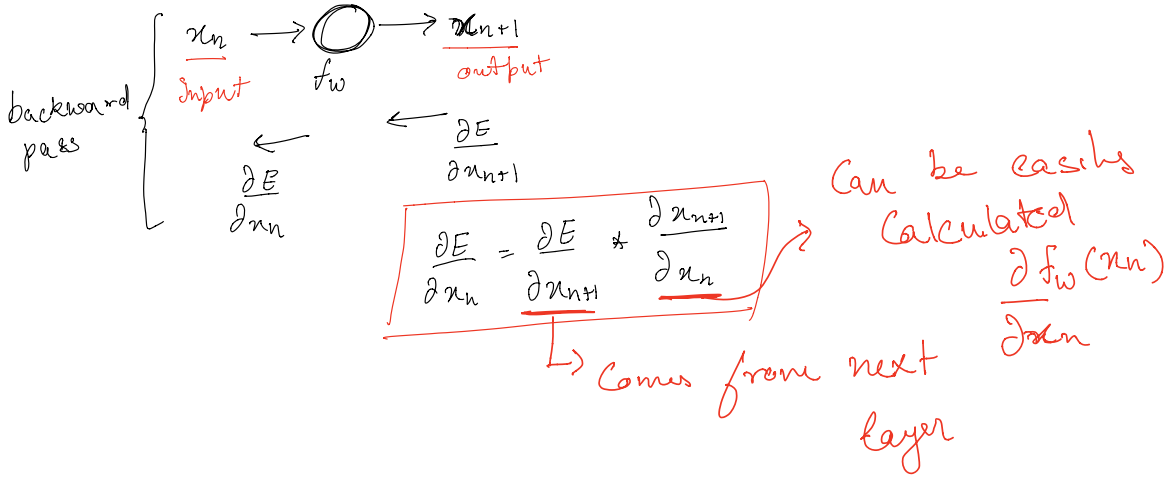
- This lets continue with the backward pass (recall slides from tutorials or lecture notes).

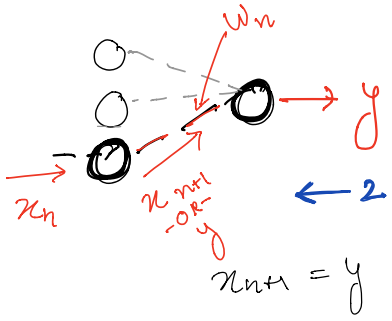
Step 1



- Updating weight w3







Now, y can be written as:

$$y = A(w_n * x_n)$$

A: Activation function

Lets say activation function is ReLu

$$A = \max(0, x)$$

Now y becomes

$$y = \max(0, w_n x_n)$$

↳ Assuming it is +ve.

$$\frac{\partial y}{\partial w_n} = x_n$$

if $(w_n x)$ was −ve then $\frac{\partial y}{\partial w}$ would have been zero.

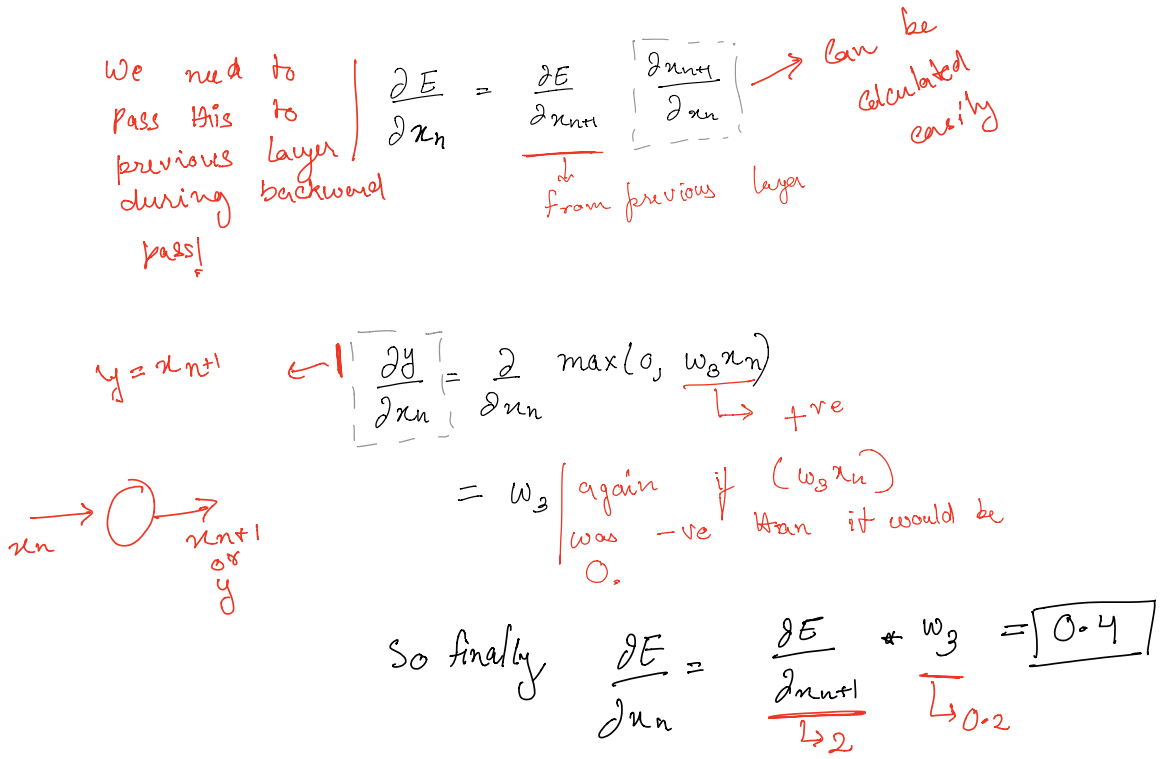
- Finally updating the w3

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial w_3} = 2 * x_n$$

$\frac{\partial E}{\partial y}$: from previous layer = 2

x_n: from forward pass

- Passing the partial derivate to the previous layer.



- 0.4 is passed onto the previous layer and used to update the w2. It's important to remember that we need values from forward pass in order to do the actual update of the weights. This document assumes you already know how to do the forward pass in the ANN, thus it should be taken as a simple guide to get insight about backward pass. It is left on you you practice a forward and backward pass in a simple ANN.