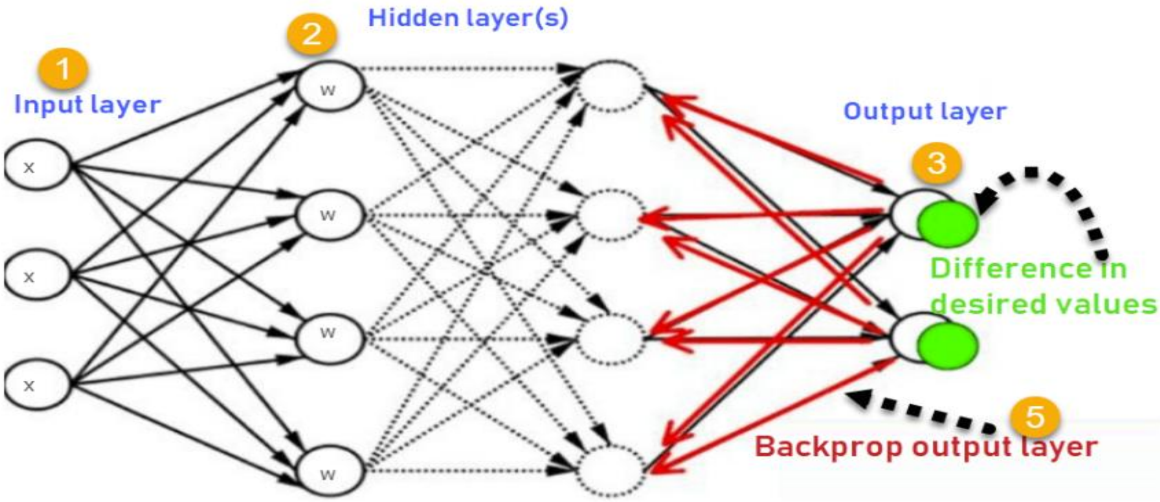
Backpropagation – Algorithm For Training A Neural Network

Backpropagation is a supervised learning algorithm, for training Multi-layer Perceptrons (Artificial Neural Networks).

Consider the following Back propagation neural network example diagram to understand:



How Backpropagation Algorithm Works

1. Inputs X, arrive through the preconnected path
2. Input is modeled using real weights W. The weights are usually randomly selected.
3. Calculate the output for every neuron from the input layer, to the hidden layers, to the output layer.
4. Calculate the error in the outputs

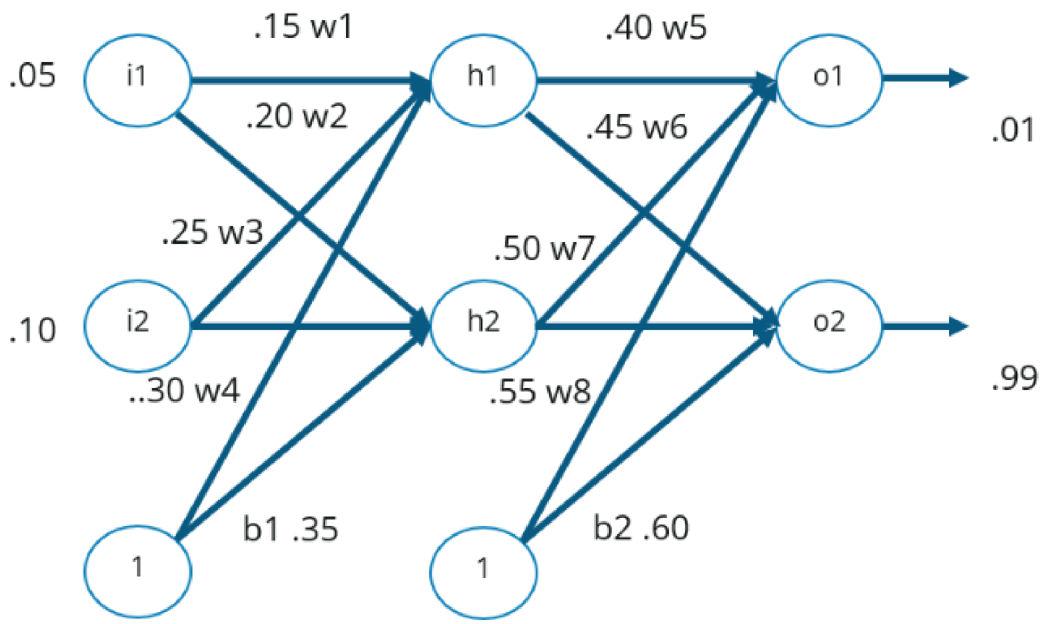
```
ErrorB= Actual Output - Desired Output
```

5. Travel back from the output layer to the hidden layer to adjust the weights such that the error is decreased.

Keep repeating the process until the desired output is achieved

How Backpropagation Works?

Consider the below Neural Network:



The above network contains the following:

- two inputs
- two hidden neurons
- two output neurons
- two biases

Below are the steps involved in Backpropagation:

- Step – 1: Forward Propagation
- Step – 2: Backward Propagation
- Step – 3: Putting all the values together and calculating the updated weight value

Step – 1: Forward Propagation

We will start by propagating forward.

Step – 1: Forward Propagation

We will start by propagating forward.

Net Input For h1:

net h1 = w1*i1 + w2*i2 + b1*1 → net h1 = 0.15*0.05 + 0.2*0.1 + 0.35*1 = 0.3775

Output Of h1:

out h1 = $1/1+e^{-net\ h1}$ → $1/1+e^{.3775}$ = 0.593269992

Output Of h2:

out h2 = 0.596884378

We will repeat this process for the output layer neurons, using the output from the hidden layer neurons as inputs.

Output For o1:

net o1 = w5*out h1 + w6*out h2 + b2*1 → 0.4*0.593269992 + 0.45*0.596884378 + 0.6*1 = 1.105905967

Out o1 = $1/1+e^{-net\ o1}$ → $1/1+e^{-1.105905967}$ = 0.75136507

Output For o2:

Out o2 = 0.772928465

Now, let's see what is the value of the error:

Error For o1:

E o1 = $\Sigma 1/2(target - output)^2$ → $½ (0.01 – 0.75136507)^2$ = 0.274811083

Error For o2:

E o2 = 0.023560026

Total Error:

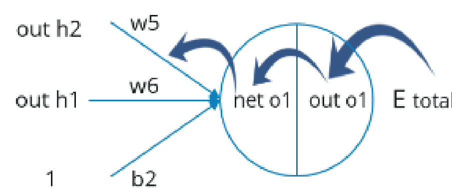
$E_{total} = E\ o1 + E\ o2$ → $0.274811083 + 0.023560026 = 0.298371109$

Step – 2: Backward Propagation

Now, we will propagate backwards. This way we will try to reduce the error by changing the values of weights and biases.

Consider W5, we will calculate the rate of change of error w.r.t change in weight W5.

$$\frac{\delta Etotal}{\delta w5} = \frac{\delta Etotal}{\delta out\ o1} * \frac{\delta out\ o1}{\delta net\ o1} * \frac{\delta net\ o1}{\delta w5}$$



Since we are propagating backwards, first thing we need to do is, calculate the change in total errors w.r.t the output O1 and O2.

Part: 2: Calculating backward propagation error
(output layer → hidden layer)

w_5, w_6, w_7 and w_8.

first let us adjust w_5

$$w_5^* = w_5 - \eta \frac{\partial E_{total}}{\partial w_5}$$

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5}$$

$$\frac{\partial total}{\partial out_{o1}} = out_{o1} - target_{o1} = 0.751365 - 0.01$$

$$= 0.7413565$$

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

$$= 0.751365(1 - 0.751365) = 0.186815602$$

$$\frac{\partial net_{o1}}{\partial w_5} = out_{h1} = 0.593269992$$

$$\frac{\partial E_{total}}{\partial w_5} = 0.7413565 \times 0.186815602 \times 0.593269992$$

$$= 0.08216704$$

$$w_5^* = w_5 - \eta \frac{\partial E_{total}}{\partial w_5}$$

$$= 0.4 - 0.6 \times 0.08216704 = 0.350699776.$$

Part:3: calculating backward propagation of error (hidden → input layer)

(w_1, w_2, w_3, w_4)

first lets adjust w_1

$$w_1^* = w_1 - \eta \frac{\partial E_{total}}{\partial w_1}.$$

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h1}} \times \frac{\partial out(h1)}{\partial net_{h1}} \times \frac{\partial net_{h1}}{\partial w_1}$$

$$\frac{\partial E_{total}}{\partial out(h_1)} = \frac{\partial E_{o1}}{\partial outh_1} + \frac{\partial E_{o2}}{\partial outh_1}$$

$$\frac{\partial E_{o1}}{\partial neto_1} \times \frac{\partial neto_1}{\partial outh_1}$$

$$\frac{\partial E_{o1}}{\partial outo_1} \times \frac{\partial outo_1}{\partial neto_1}$$

$$\frac{\partial E_{o2}}{\partial neto_2} \times \frac{\partial neto_2}{\partial outh_1}$$

$$\frac{\partial E_{o2}}{\partial outo_2} \times \frac{\partial outo_2}{\partial neto_2}$$

$$\frac{\partial E_{o2}}{\partial outo_2} = (outo_2 - targeto_2)$$

$$= 0.77292846 5 - 0.99$$

$$= -0.217071535$$

$$\frac{\partial outo_2}{\partial neto_2} = outo_2(1 - outo_2)$$

$$= 0.7729(1 - 0.7729)$$

$$= 0.17551 0052$$