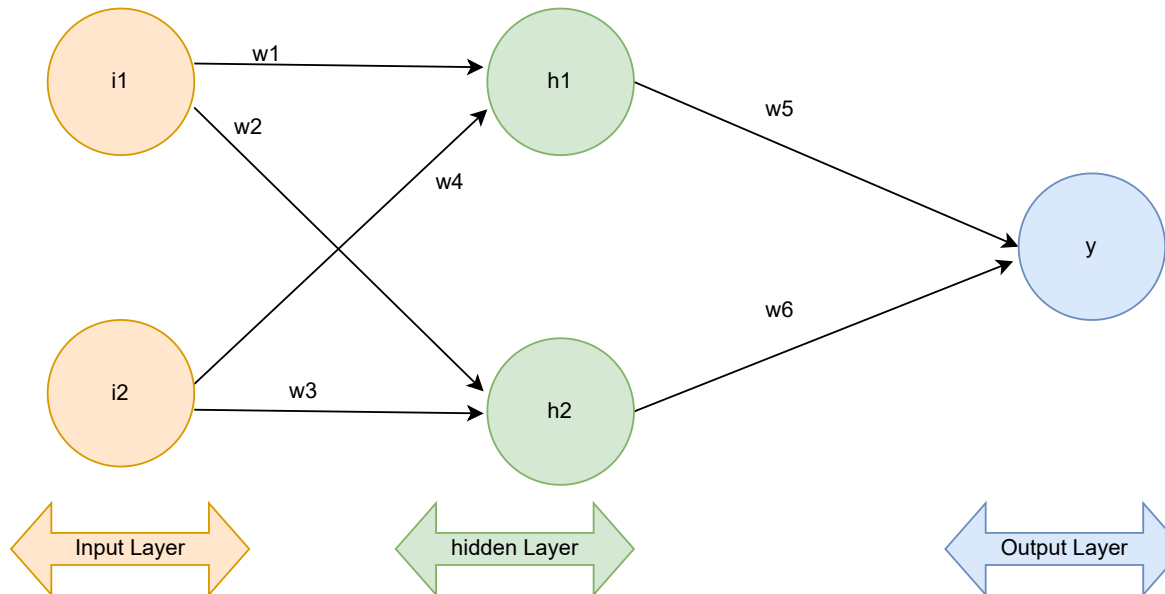
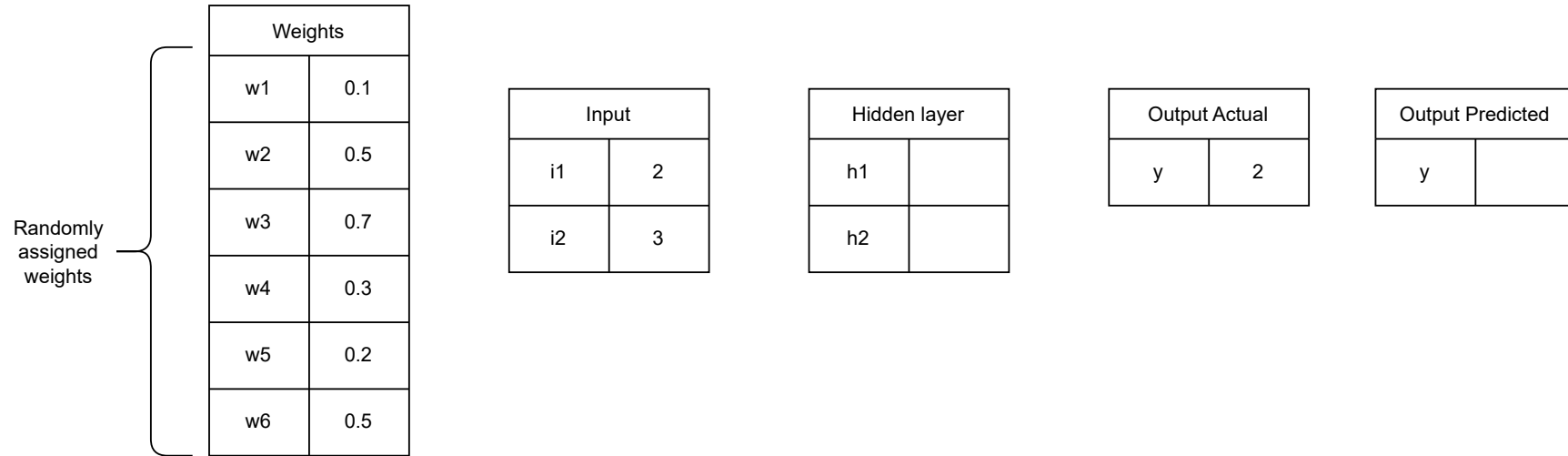


STEP 01



STEP 02 | Forward Pass

Weights	
w1	0.1
w2	0.5
w3	0.7
w4	0.3
w5	0.2
w6	0.5

Input	
i1	2
i2	3

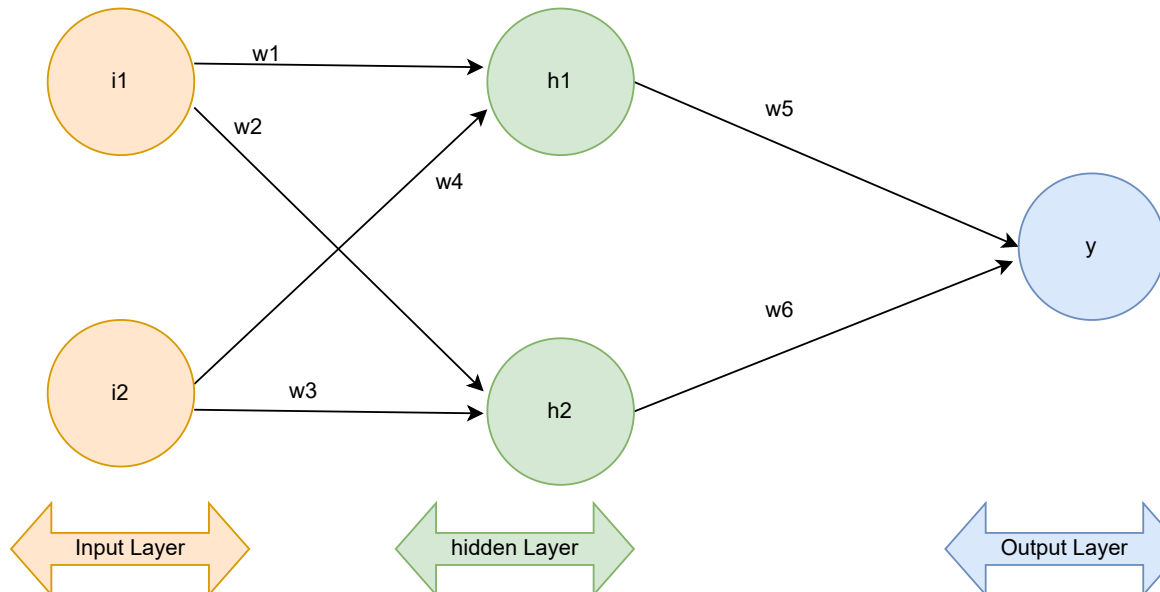
Hidden layer	
h1	1.1
h2	3.1

Output Actual	
y	2

Output Predicted	
y	1.77

$$\begin{aligned}h1 &= x1*w1 + x2*w4 \\h2 &= x1*w2 + x2*w3 \\y &= h1*w5 + h2*w6\end{aligned}$$

$$\begin{aligned}h1 &= 2*0.1 + 3*0.3 = 1.1 \\h2 &= 2*0.5 + 3*0.7 = 3.1 \\y &= h1*0.2 + h2*0.5 = 1.77\end{aligned}$$



STEP 03 | Error Calculation

Weights	
w1	0.1
w2	0.5
w3	0.7
w4	0.3
w5	0.2
w6	0.5

Input	
i1	2
i2	3

Hidden layer	
h1	1.1
h2	3.1

Output Actual	
y	2

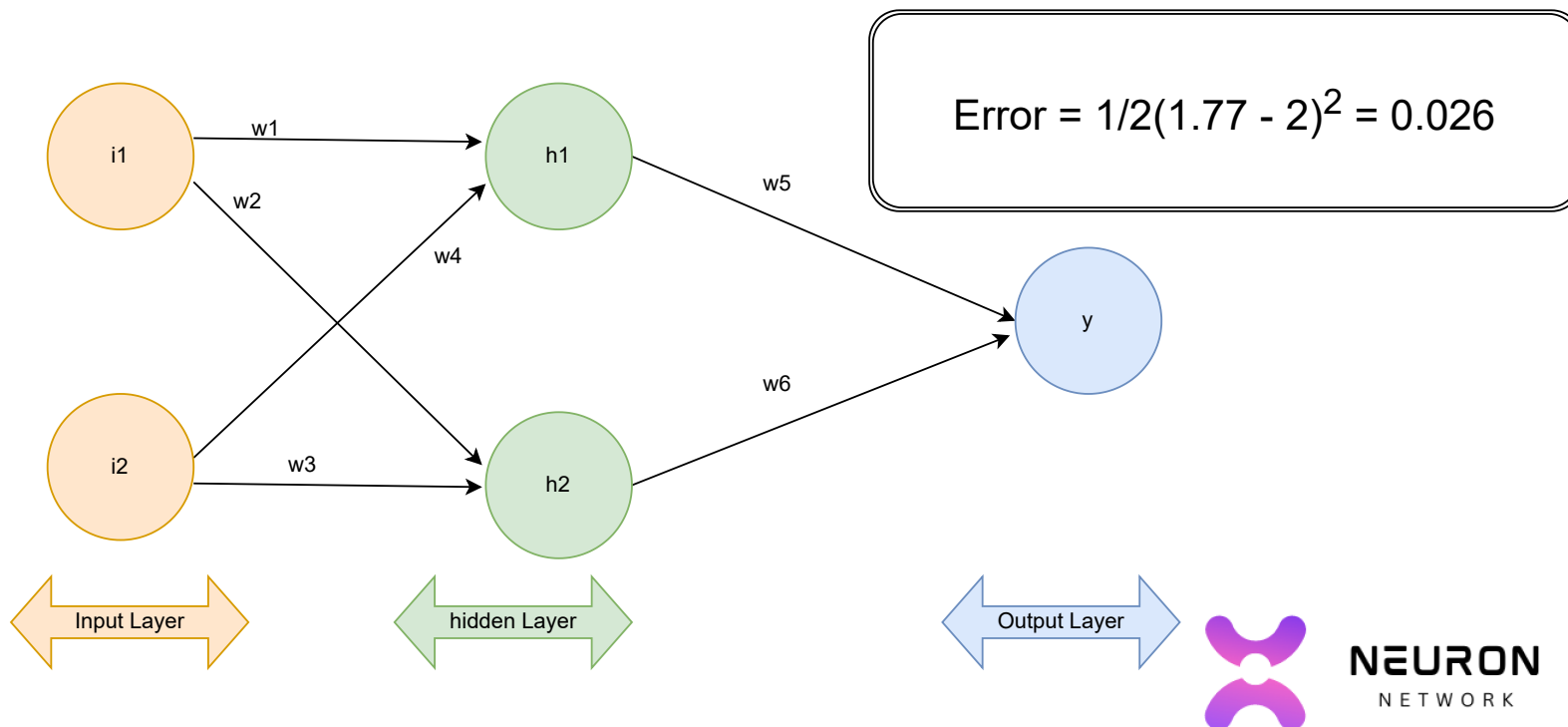
Output Predicted	
y	1.77

Error = 0, if prediction = actual

$$\text{Error} = \frac{1}{2}(\text{prediction} - \text{actual})^2$$

Error is always positive because of the square

$\frac{1}{2}$ is added to ease the calculation of the derivative



STEP 04 | BackProp | Weight Adjustment

Weights	
w1	0.09999428
w2	0.49400858
w3	0.69999428
w4	0.29997582
w5	0.199714
w6	0.499194

Input	
i1	2
i2	3

Hidden layer	
h1	1.1
h2	3.1

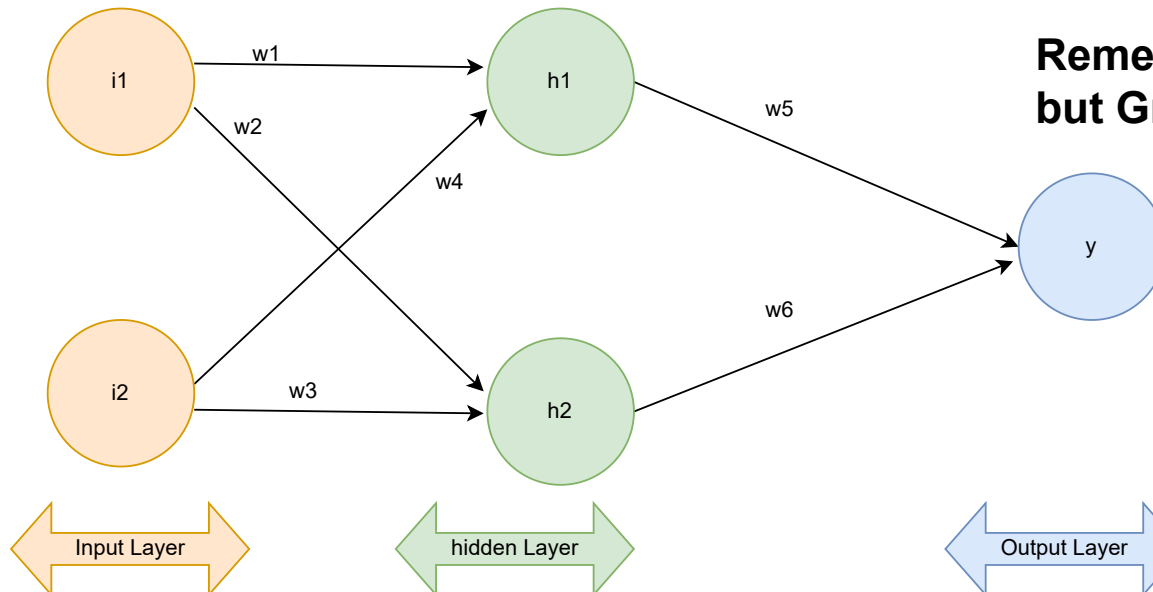
Output Actual	
y	2

Output Predicted	
y	1.77

$$= \text{Error} = \frac{1}{2}(1.77 - 2)^2 = 0.026$$

updated weights

$$\begin{aligned} *w_6 &= w_6 - a (h_2 \cdot \Delta) \\ *w_5 &= w_5 - a (h_1 \cdot \Delta) \\ *w_4 &= w_4 - a (i_2 \cdot \Delta w_6) \\ *w_3 &= w_3 - a (i_1 \cdot \Delta w_6) \\ *w_2 &= w_2 - a (i_2 \cdot \Delta w_5) \\ *w_1 &= w_1 - a (i_1 \cdot \Delta w_5) \end{aligned}$$



Remember BackProp is nothing but Gradient Descent with Chain Rule

a = learning rate = 0.01



NEURON
NETWORK

STEP 05 | Forward Pass Again

Weights	
w1	0.09999428
w2	0.49400858
w3	0.69999428
w4	0.29997582
w5	0.199714
w6	0.499194

Input	
i1	2
i2	3

Hidden layer	
h1	1.09991602
h2	3.088

Output Actual	
y	2

Output Predicted	
y	1.85

$$\Delta = \text{Error} = 1/2(1.85 - 2)^2 = 0.01125$$

Notice that the error has reduced in this iteration.
Perform several such steps to converge

