

Learning Parameters via *Back-Propagation*

not(XOR)

x_1	x_2	y
0	0	1
0	1	0
1	0	0
1	1	1

1. Initialize all the parameters ($w_1, w_2, w_3, w_4, w_5, w_6, b_1, b_2, b_3$) to random values.
2. Forward pass: input data samples (e.g., $x_1 = 0, x_2 = 0$) to the neural network, and then get the output \hat{y} .
3. Compare \hat{y} with the desired output y , and the difference between them is regarded as the error: $\text{error} = \hat{y} - y$.
4. **Back-propagate the error to the beginning of the network. During back-propagation, we adjust the parameters (e.g., if we find error is caused by that w_5 is too big, then we decrease w_5).**

