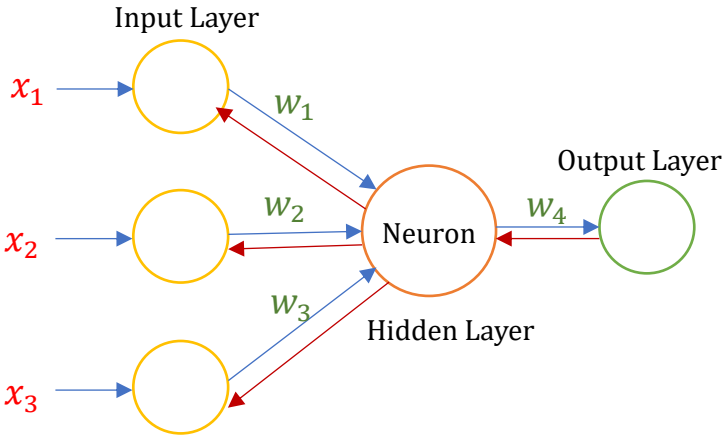
Forward and Backward Propagation in NN Training

1. Forward propagation
2. Backward propagation



Let's consider an example data as shown in below to understand these concepts better. It is a time management data of different students.

study	play	sleep	Good time maintenance
2hr	4hr	9hr	1
1hr	2hr	8hr	0
5hr	1hr	9hr	1
2hr	2hr	5hr	0

Here x_1 = Study

x_2 = Play

x_3 = Sleep

Output = Good time maintenance

In the about diagram blue arrows indicates the forward propagation and red arrows indicates the backward propagation

1. Forward Propagation:

Now we need to pass this data to above neural network for training purpose. This process is called forward propagation because the data and the training flow is in forward direction in neural network.

Let's understand forward propagation very clearly with below steps:

Step - 1:

Providing input data (x_1 , x_2 , x_3) to neural network.

Step - 2:

Neuron will calculate Y value with below mathematical equation.

$$z = w_1x_1 + w_2x_2 + w_3x_3 + bias$$

Step - 3:

Neuron will calculate the activation function value by using above y value, and pass the generated z value to output layer.

$$z=act(y)$$

2. Backward Propagation:

- The backward propagation starts immediately after received the output in forward propagation. The following steps involved in backward propagation.
- By using this backward propagation, we will decrease the loss and optimize the training.

Step - 1:

First it will calculate the loss value by using predicted and actual Y values with below equation.

$$loss = \sum_{i=1}^{n} (y_i - \hat{y}_i)$$

Now the neural network task is reducing this loss by update the weights (w_1, w_2, w_3, w_4), because this loss is nothing but the total generated error in the neural network.

Step - 2:

In order to reduce the loss, it will update the weights (w_1, w_2, w_3, w_4) by using below equation. This equation provides the new weight by using old weight (w_{old}), learning rate (α) and derivative of loss function.

$$\omega_{new} = \omega_{old} - \alpha \frac{\partial L}{\partial \omega_{old}}$$

There is a huge subject behind the above equation we can learn all those things in gradient decent topic.

Step - 3:

After update the all the weights in the neural network with this backward propagation it will again start the forward propagation operation. This cycle continues until to get the best accuracy with less loss value.