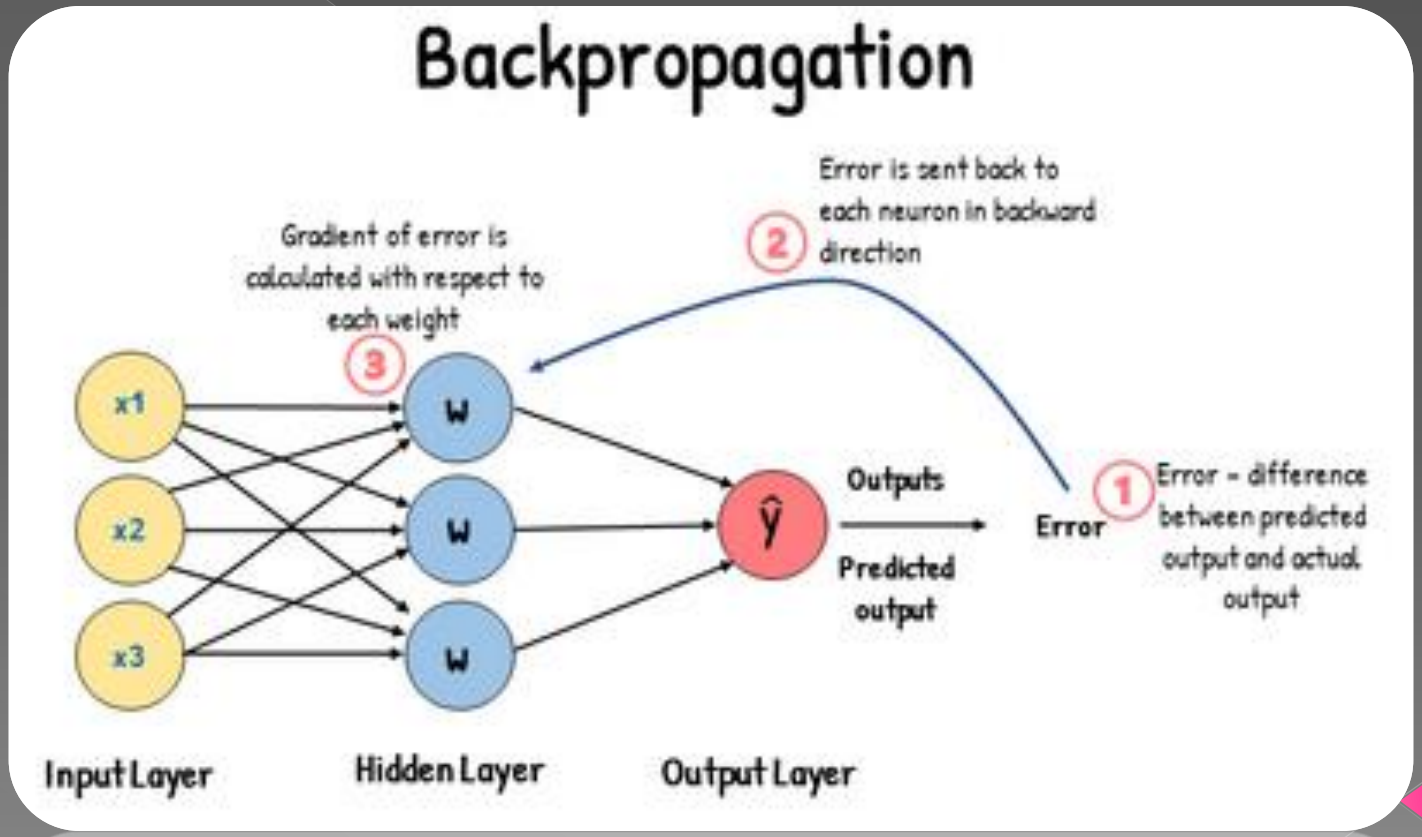
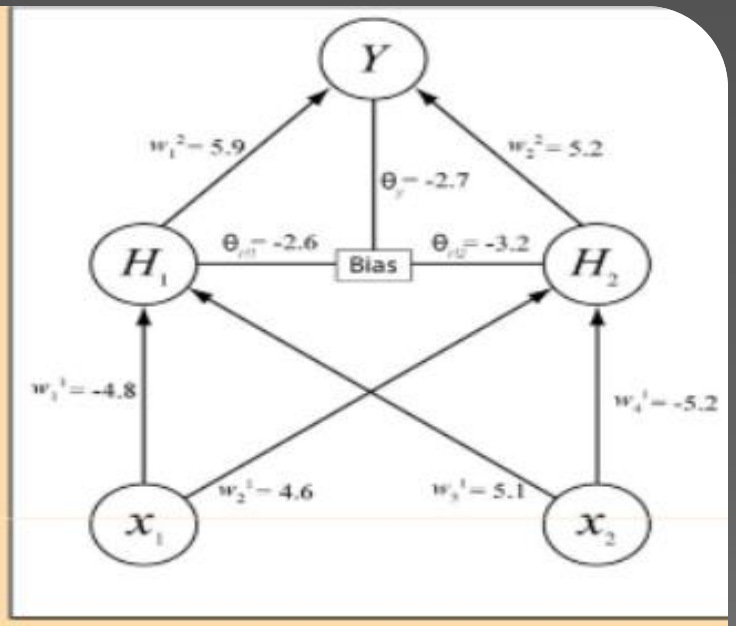

بسم الله الرحمن الرحیم

Backpropagation Algorithm (BP)



What happen if the error is large?

What Happen if the Network output is NOT approaching to the target values and/or the error is Large?



Input		Target Output	Network Output (Y)
0	1	1	0.6401
1	0	1	0.5138
0	0	0	0.3102
1	1	0	0.2142

BACKPROPAGATION ALGORITHM (BP)

Backpropagation is a supervised learning Algorithm used by multilayered neural network for learning purposes.

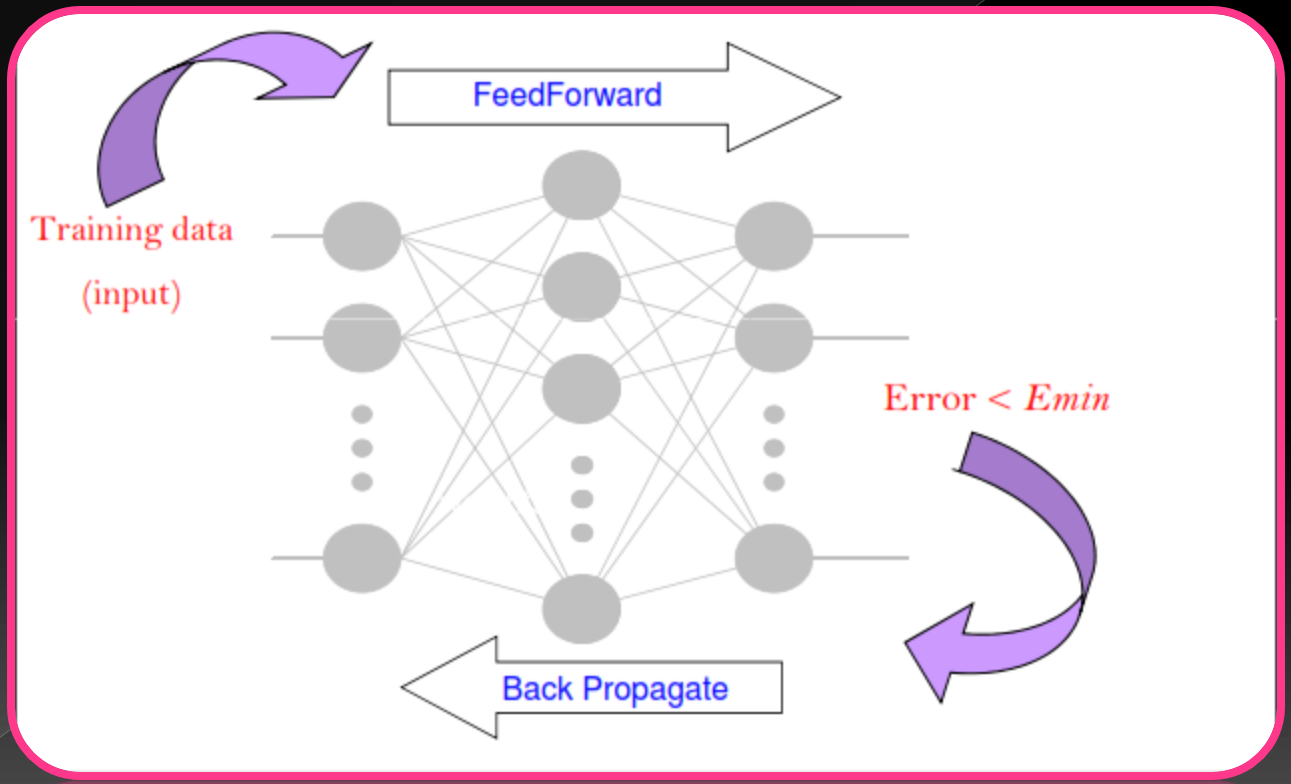
Now, the BP algorithm is the most widely used supervised learning technique to train feedforward ANN despite its existence for almost 5 decades

How does back propagation algorithm work?

BP for MLP Network

1. A BP network is an algorithm that can propagate and back propagate the network from one layer to other layers.
2. The network consists of an input layer of The network consists of an input layer of source neurons, at least one middle or hidden layer of computational neurons, and an output layer of computational neurons.
3. The input signals are propagated in a forward direction on a layer-by-layer basis and backward direction on a layer-by layer basis.

Multilayer Network with BP algorithm



How BP Learns ?

Before the learning process starts, all the weights (synapses) in the network are initialized with pseudo random numbers.

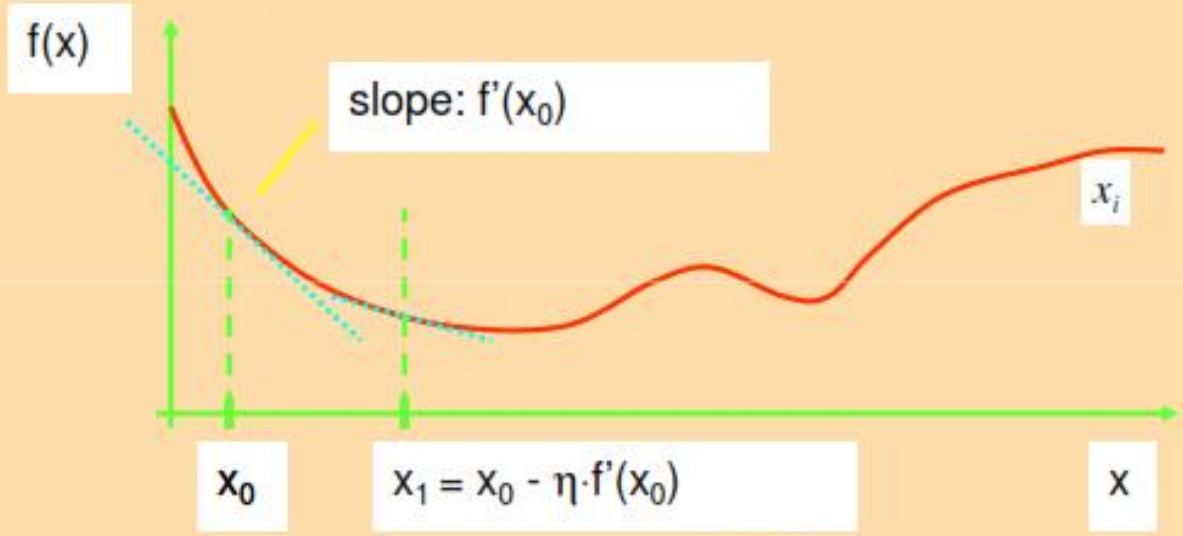
1. We have to provide a set of training patterns (exemplars). They can be described as a set of ordered vector pairs $\{(x_1, y_1), (x_2, y_2), \ldots, (x_P, y_P)\}$.
2. First, a training input pattern is presented to the network input layer. The network propagates the input pattern from layer to layer until the output pattern is generated (MLP Procedure)
3. Second, if this pattern is different from the desired output, an error is calculated and then propagated backwards through the network from the output layer to the hidden layer and input layer. The weights are modified as the error is propagated.
4. Then we can start the backpropagation learning algorithm.This algorithm iteratively minimizes the network's error by finding the gradient of the error surface in weight-space and adjusting the weights in the opposite direction (Gradient-Descent method).

WHAT IS GRADIENT DESCENT METHOD?

Gradient Descent is an optimization algorithm.

GRADIENT DESCENT METHOD

Gradient-descent example: Finding the absolute minimum of a 1-dimensional error function f(x):



Repeat this iteratively until for some x_i , $f'(x_i)$ is sufficiently close to 0.

BP Algorithm

1. Randomly select a vector pair (x_p, y_p) from the training set and call it (x, y).
2. Use x as input to the BPN and successively compute the outputs of all neurons in the network (bottom-up) until you get the network output (O).
3. Compute the error signal δ_k, for the pattern p across all k output layer units by using the formula:

$$\delta_k = (t_k - o_k) f'(net_k)$$

$$\delta_k = e_k . f'(net_k)$$

BP Algorithm

4. Compute the error δ_j, for all j hidden layer units by using the formula:

$$\delta_j = f'(net_k)\sum_{k=1}^{K}\delta_k w_{kj}$$

$$\delta_j = \sum_k w_{kj}\delta_k f'(net_j)$$

5. Update the connection-weight values to the hidden layer by using the following equation:

$$w_{ji}(t+1) = w_{ji}(t) + \eta\delta_j x_i$$