19Z601- MACHINE LEARNING

UNIT- 3

NEURAL NETWORKS AND DECISION TREES : **Feed-forward Networks - Network Training - Delta Rule- Gradient Descent – Error Backpropagation - Regularization in Neural Networks - Generalisation - Decision Tree Learning- Representation - Inductive Bias- Issues** (9)

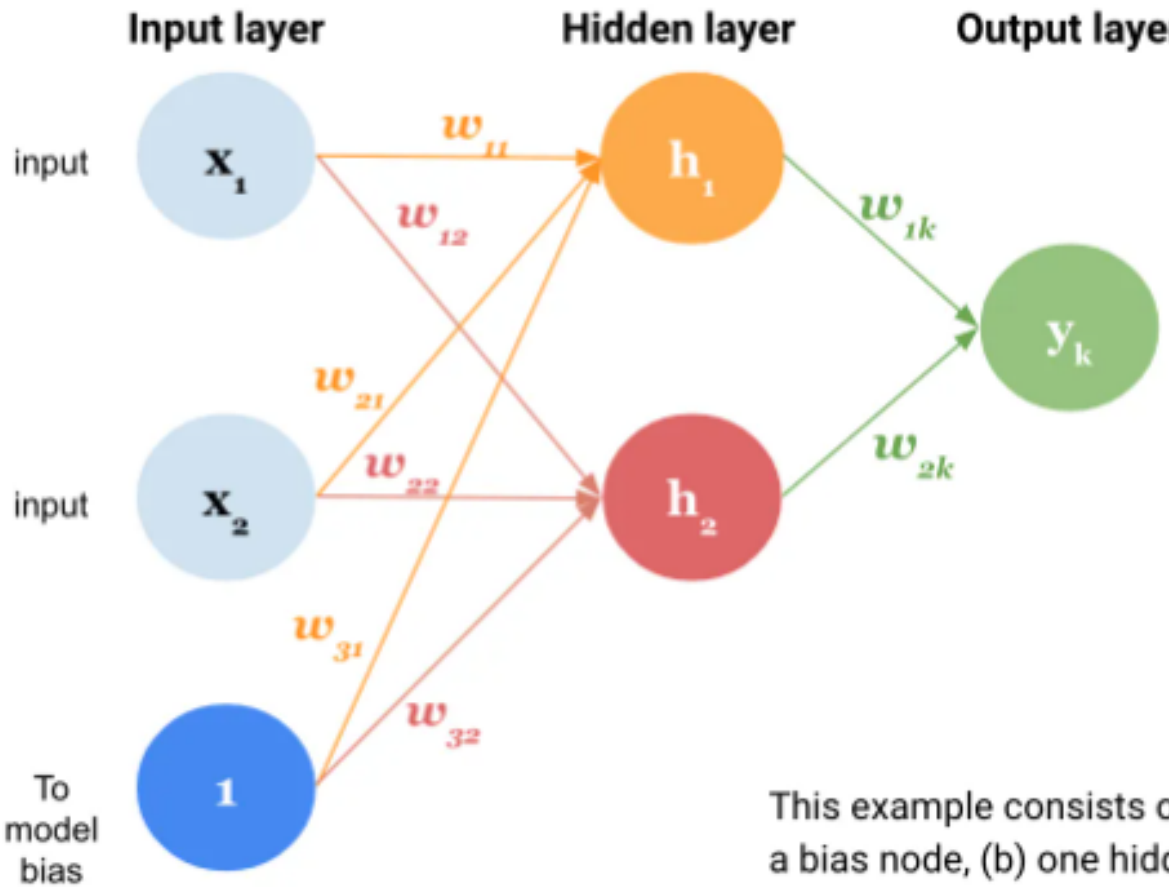
Presented by

Ms.Anisha.C.D

Assistant Professor

CSE

Multi-Layer Perceptron



x_1, x_2 : *input data features*

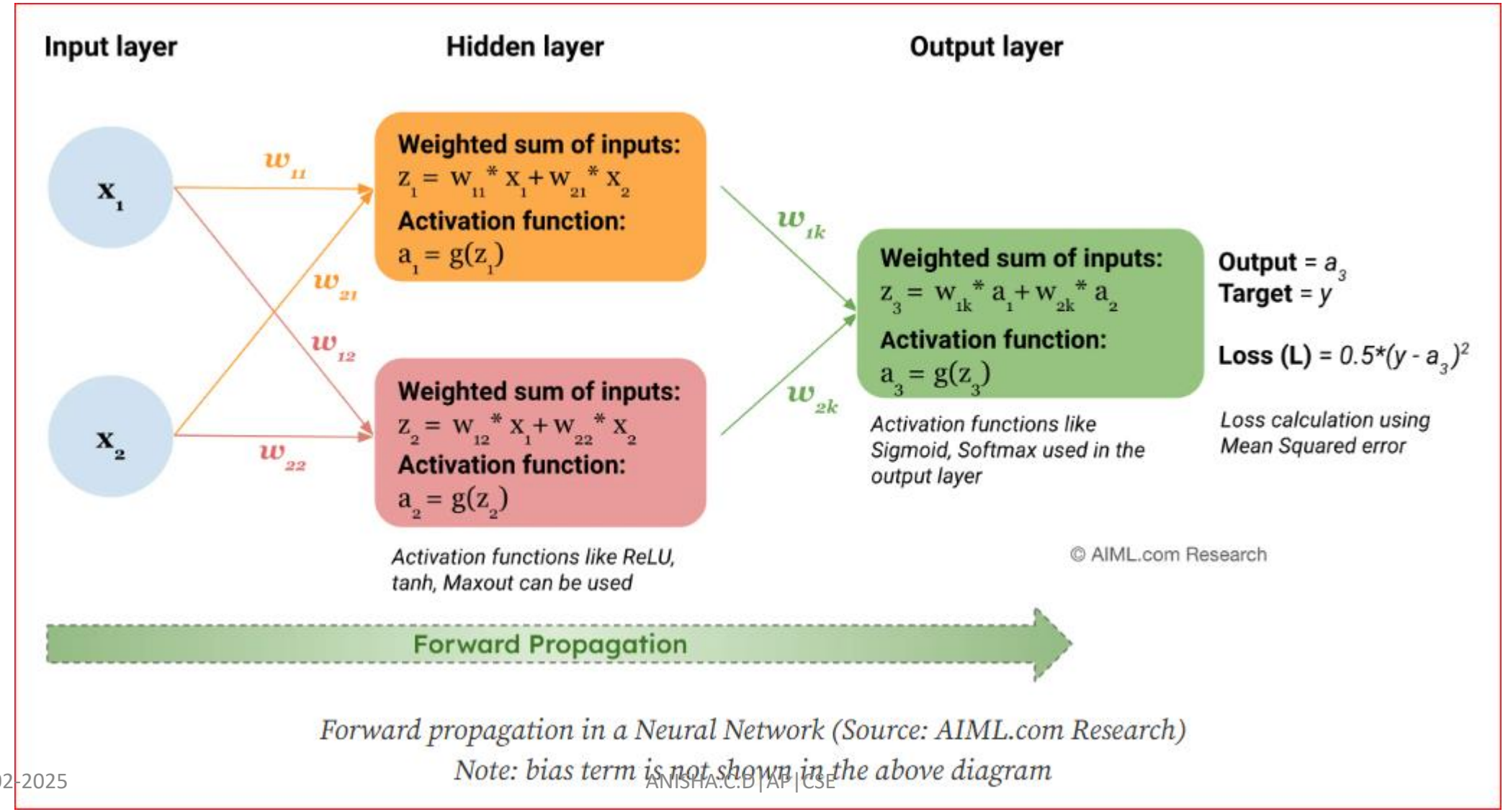
w_{ij} : *weights of the network*

h_1, h_2 : *nodes in the hidden layer*

y_k : *output variable*

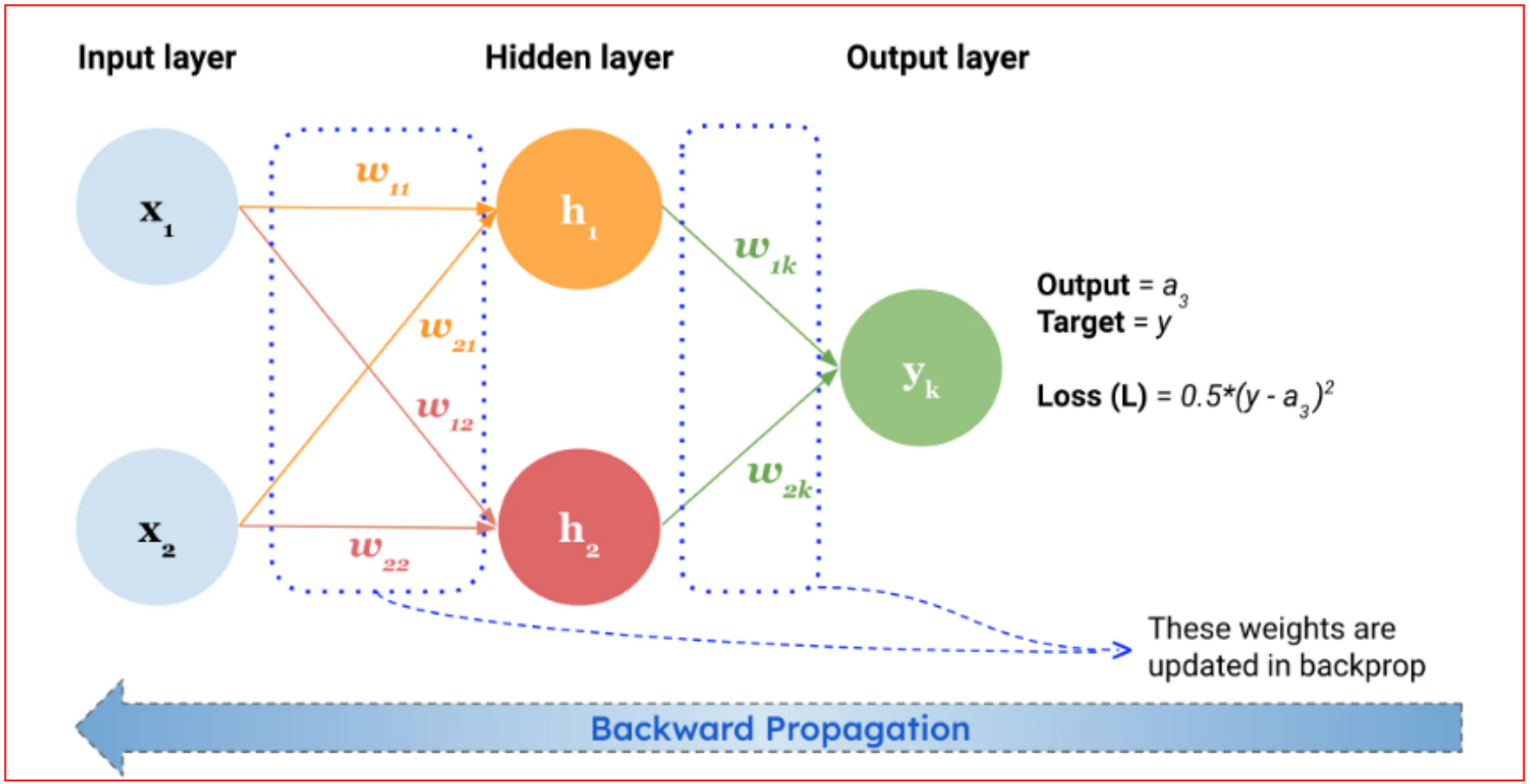
This example consists of: (a) an input layer with two input nodes and a bias node, (b) one hidden layer with two neurons, and (c) an output layer with one neuron

Working of Multi-Layer Perceptron



Forward propagation in a Neural Network (Source: AIML.com Research)
Note: bias term is not shown in the above diagram

Working of Multi-Layer Perceptron

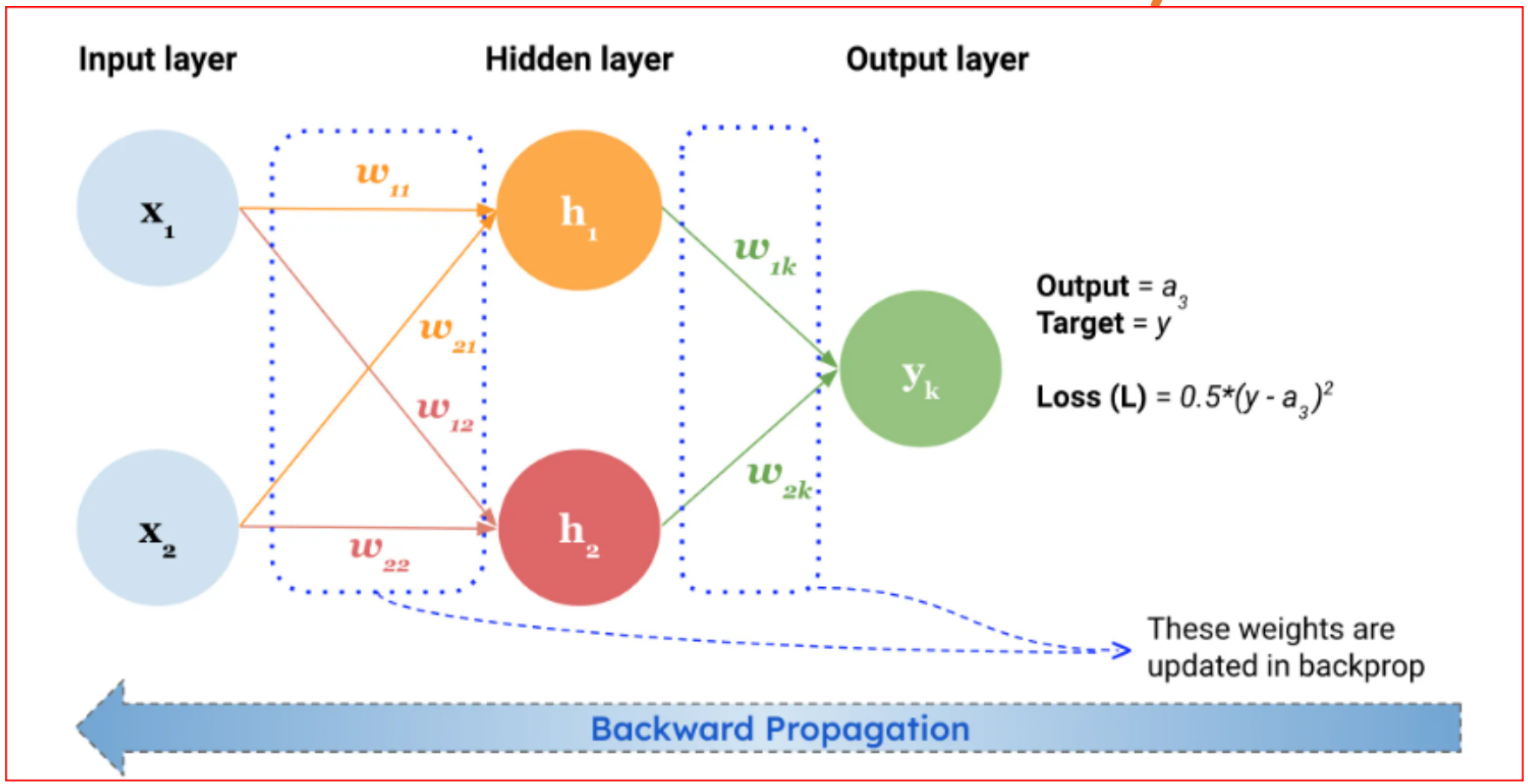


$$\frac{\delta Ed}{\delta w_{ji}}$$

$$\frac{\delta Ed}{\delta net_j} \qquad \frac{\delta net_j}{\delta w_{ji}}$$

$$net_j = \sum_i w_{ji}\, xj_i$$

Working of Multi-Layer Perceptron – Case 1 : Output Layer (Unit j is the Output Unit)



$$\frac{\delta Ed}{\delta net_j}$$

$$\frac{\delta Ed}{\delta o_j} \qquad \frac{\delta o_j}{\delta net_j}$$

Working of Multi-Layer Perceptron – Case 2 : Hidden Layer (unit j is the internal unit)

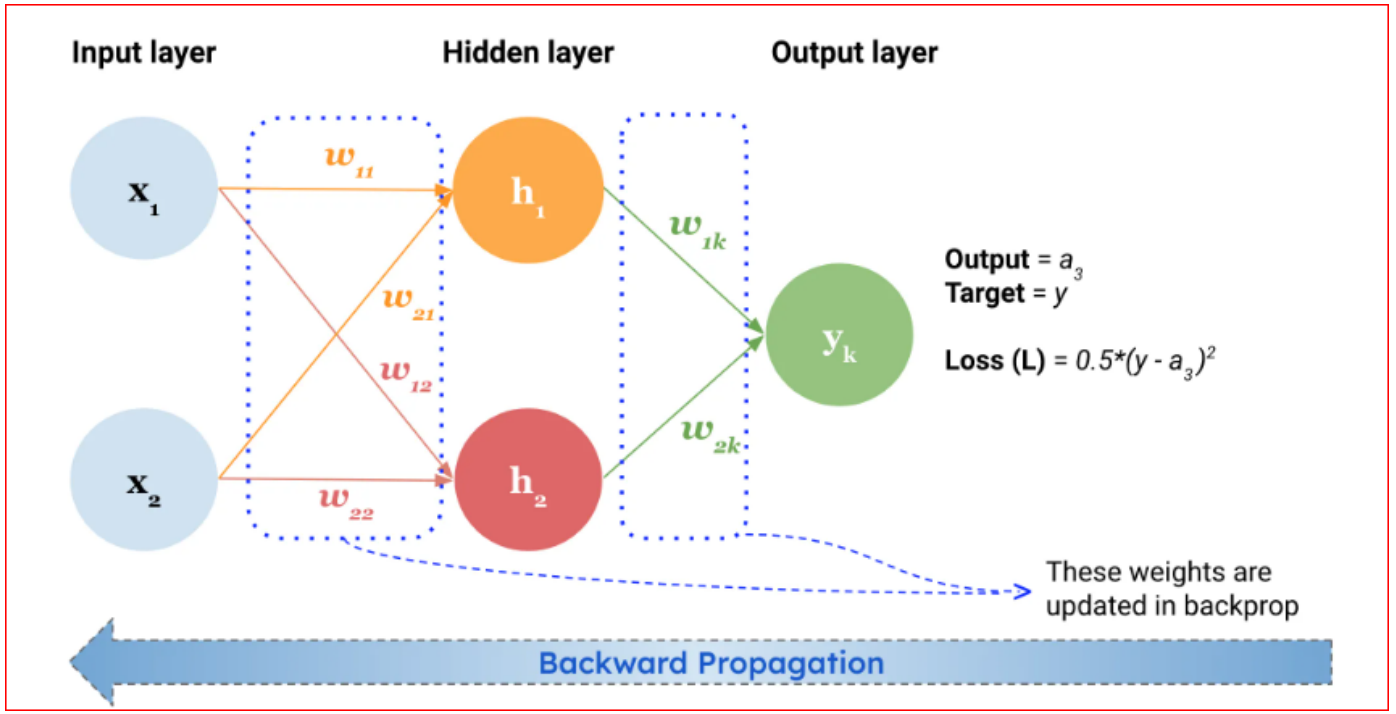
- Terminology : Downstream Neurons

Hidden Layer:

- Neurons **h_1 and h_2** receive inputs from **x_1 and x_2**
- So, **h_1 and h_2 are the downstream neurons of x_1 and x_2**

Output Layer:

- Neuron **y_k** receives inputs from **h1 and h2**
- So, **h_1 and h_2 are the downstream neurons of x_1 and x_2**



Working of Multi-Layer Perceptron – Case 2 : Hidden Layer (unit j is the internal unit)

$$\frac{\partial E_d}{\partial net_j} = \sum_{k \in Downstream(j)} \frac{\partial E_d}{\partial net_k} \frac{\partial net_k}{\partial net_j}$$

$$= \sum_{k \in Downstream(j)} -\delta_k \frac{\partial net_k}{\partial net_j}$$

$$= \sum_{k \in Downstream(j)} -\delta_k \frac{\partial net_k}{\partial o_j} \frac{\partial o_j}{\partial net_j}$$

$$= \sum_{k \in Downstream(j)} -\delta_k \, w_{kj} \frac{\partial o_j}{\partial net_j}$$

$$= \sum_{k \in Downstream(j)} -\delta_k \, w_{kj} \, o_j(1 - o_j)$$

$$\delta_j = o_j(1 - o_j) \sum_{k \in Downstream(j)} \delta_k \, w_{kj}$$

$$\Delta w_{ji} = \eta \, \delta_j \, x_{ji}$$