DEEP LEARNING

LECTURE 3: Fundamentals of Neural Networks

Ms. Farzeen Ashfaq

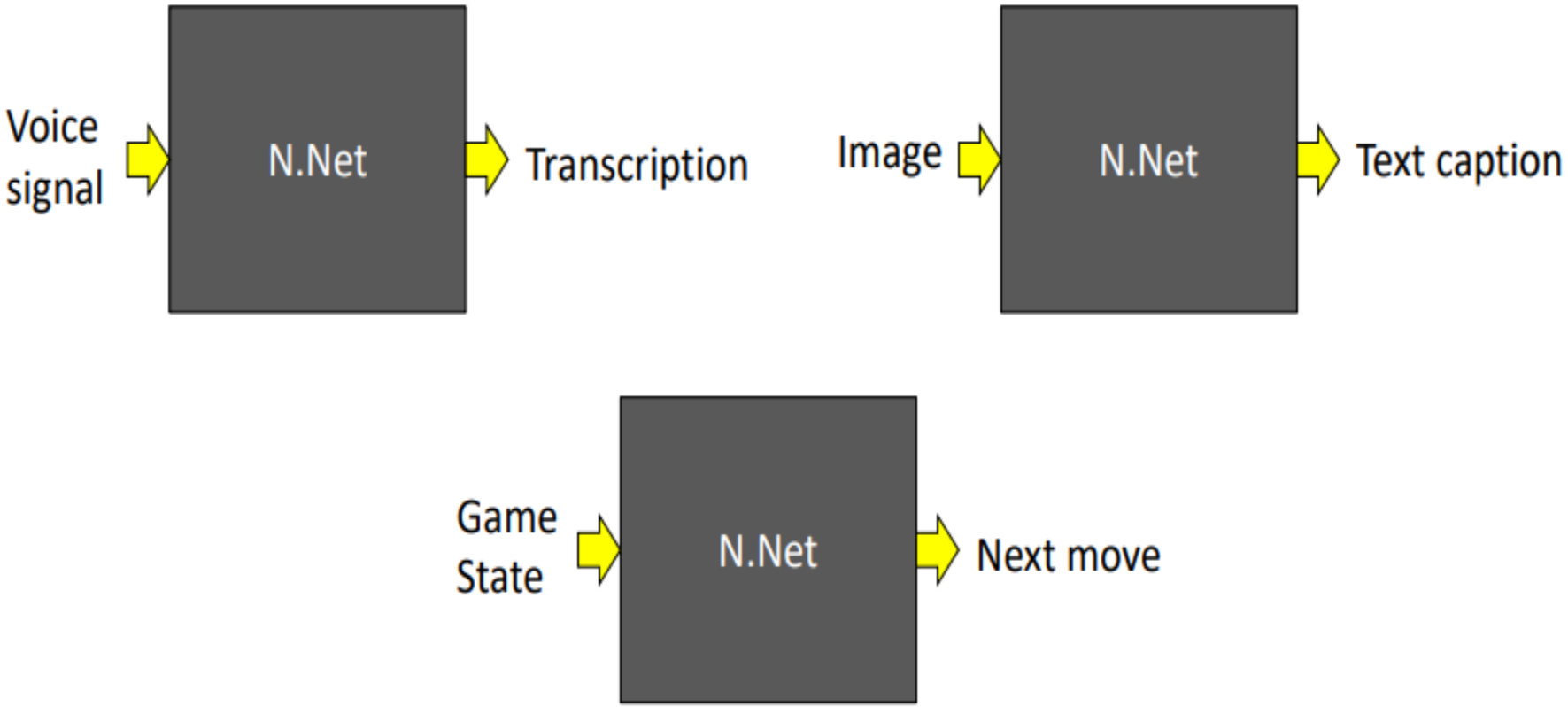
Recap: Motivation For DL

- The simple machine learning algorithms work very well on a wide variety of important problems. However, they have not succeeded in solving the central problems in AI, such as recognizing speech or recognizing objects. Deep learning was designed to o vercome these and other obstacles.

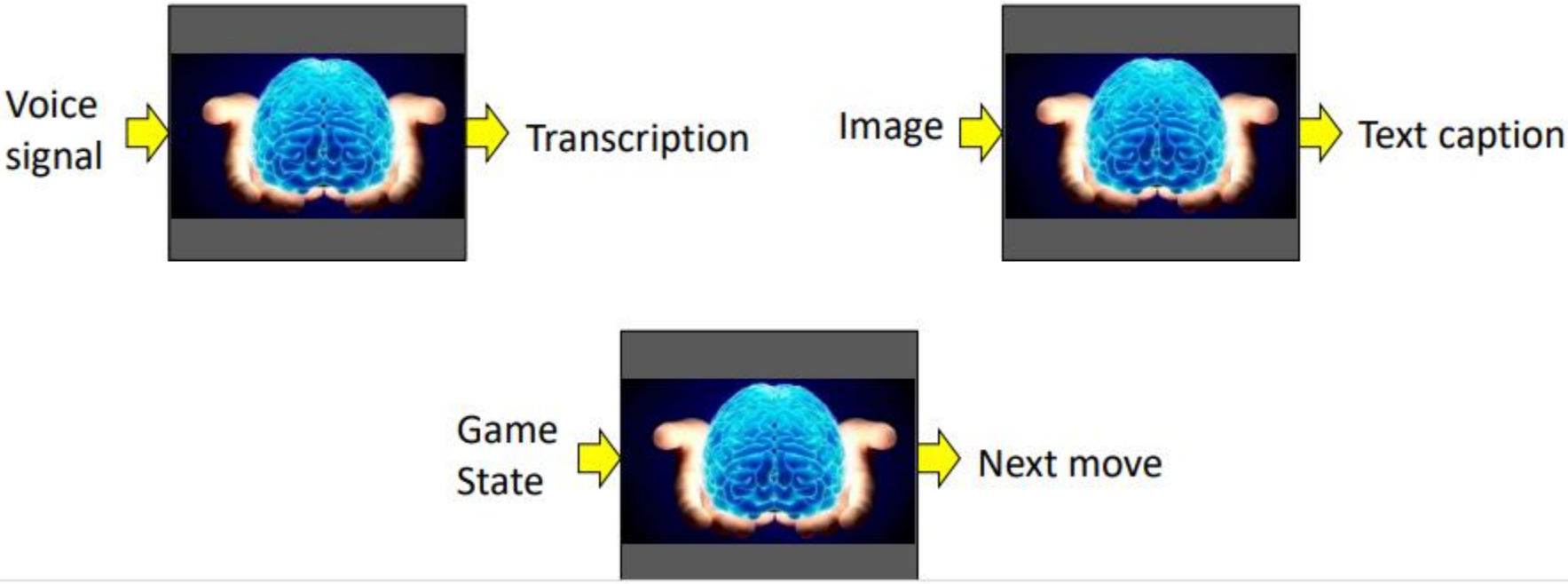
Recap: Biological Inspiration

- Deep learning is an aspect of artificial intelligence (AI) that is to simulate the activity of the **human brain** specifically, pattern recognition by passing input through various layers of the neural network.

What are Neural Networks



The Human Perspective



Fundamental Concept

- ***Neural networks*** are those information processing systems, which are constructed and implemented to model the human brain.
- The main objective of the neural network research is to develop a computational device for modeling the brain to perform various computational tasks at a faster rate .than the traditional systems

The Human Brain

(According to a computer scientist)

- Send electro-chemical signals
- Network of ~100 Billion Neurons
- Each ~1,000 – 10,000 connections
- Activation time ~10 ms second
- ~100 Neuron chain in 1 second

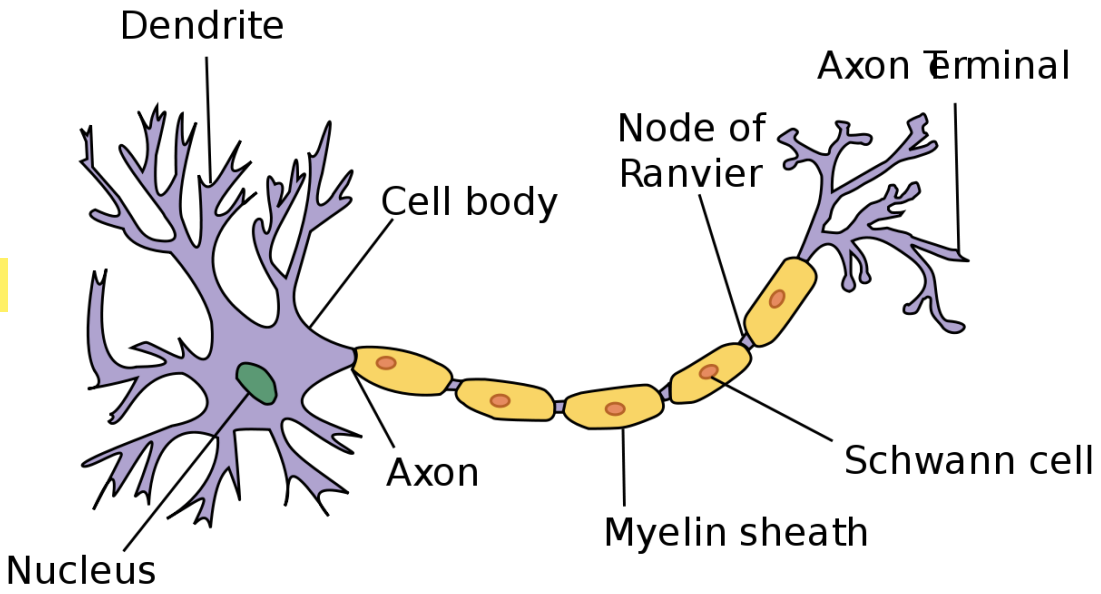
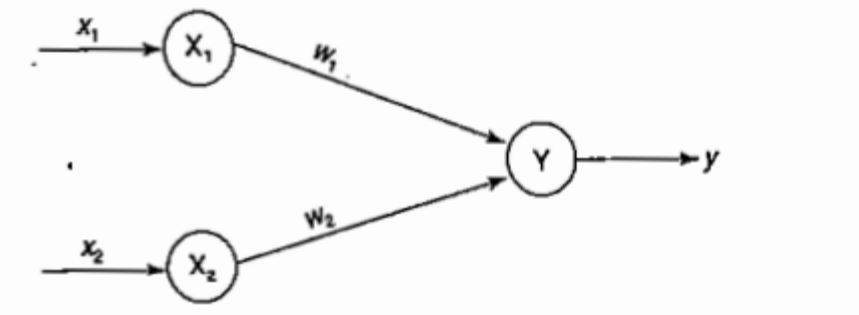


Image from Wikipedia

Architecture of a simple artificial Neuron Net

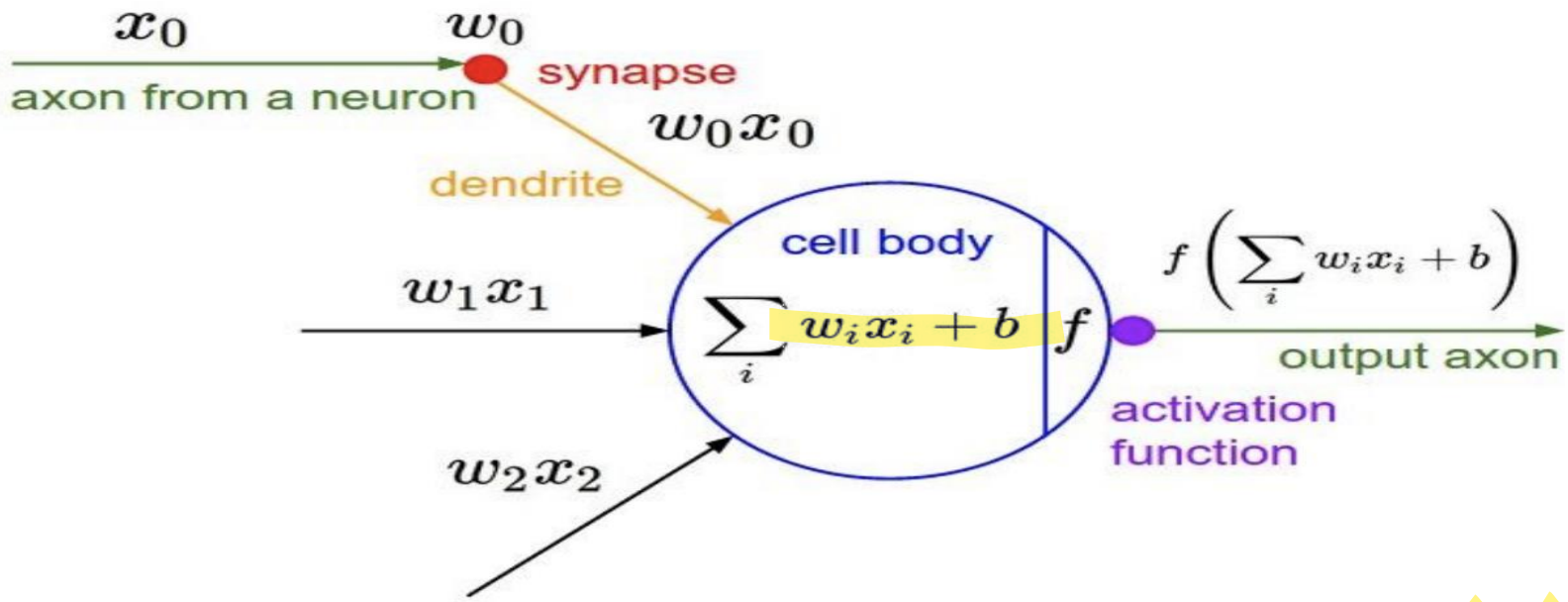


$$y_{in} = + x_1 w_1 + x_2 w_2$$

- To depict the basic operation of a neural net,
- consider a set of neurons, say x_1 and x_2, transmitting signals to another neuron, Y.

- Here x_1 and x_2 are input neurons, which transmit signals, and Y is the output neuron, which receives signals.
- Input neurons x_1 and x_2 are connected to the output neuron Y, over a weighted interconnection links (w_1 and w_2, as shown in Figure

A More In Depth Mathematical Model of Artificial Neuron



- We can draw a diagram that makes the analogy between the neuron structure and the artificial neurons in a neural network.

Table 2-1 Terminology relationships between biological and artificial neurons

Biological neuron	*Artificial neuron*
Cell	Neuron
Dendrites	Weights or interconnections
Soma	Net input
Axon	Output

ANN

- An artificial neural network (ANN) is an efficient information processing system which resembles in characteristics with a biological neural network.
- ANNs possess large number of highly interconnected processing elements called nodes or units or neuron, which usually operate in parallel and are configured in regular architectures

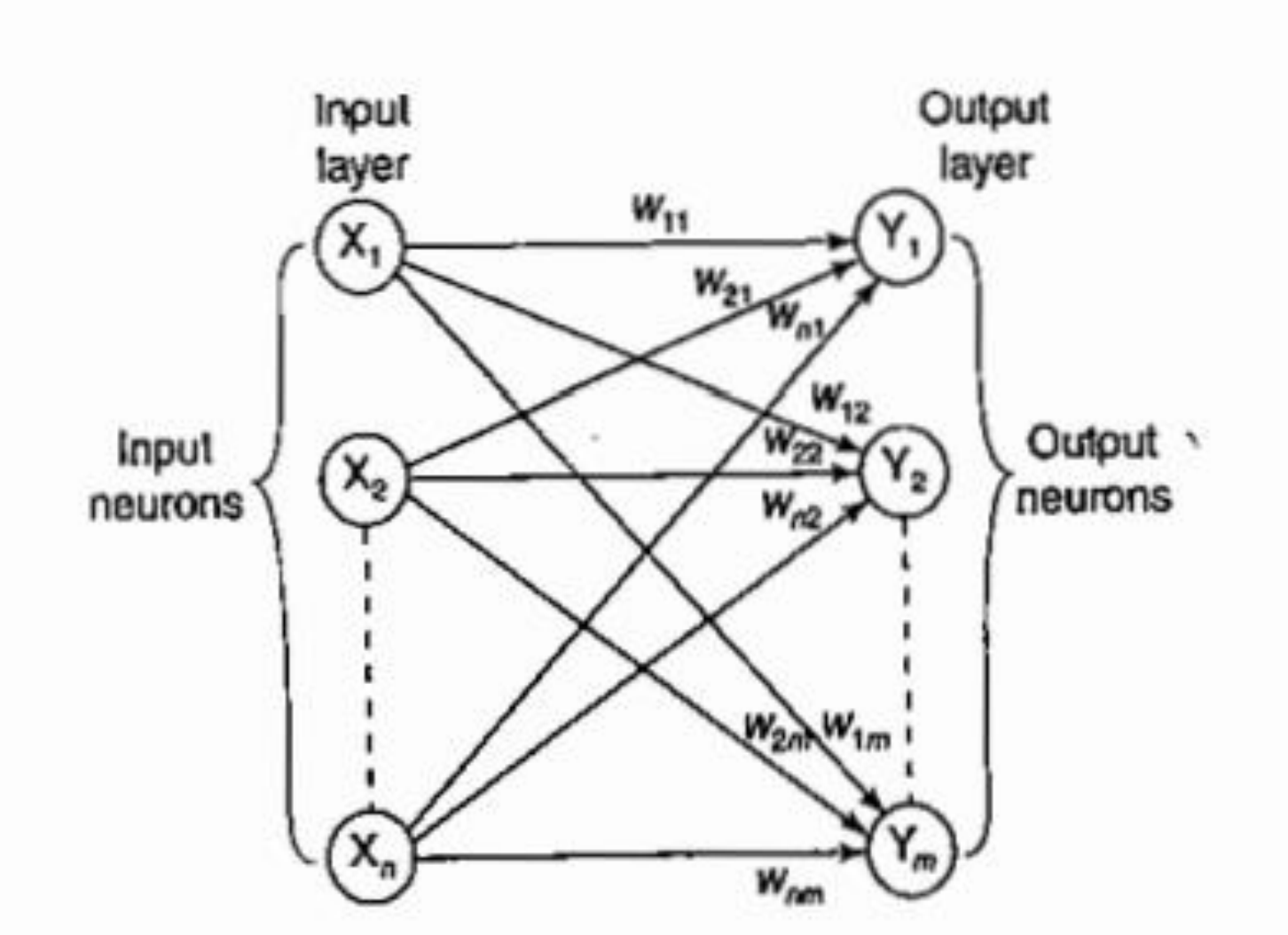
Network Architecture

- The arrangement of neuron to form layers and the connection pattern formed and between layers is the network architecture. There exist five basic types of neuron connection architectures.

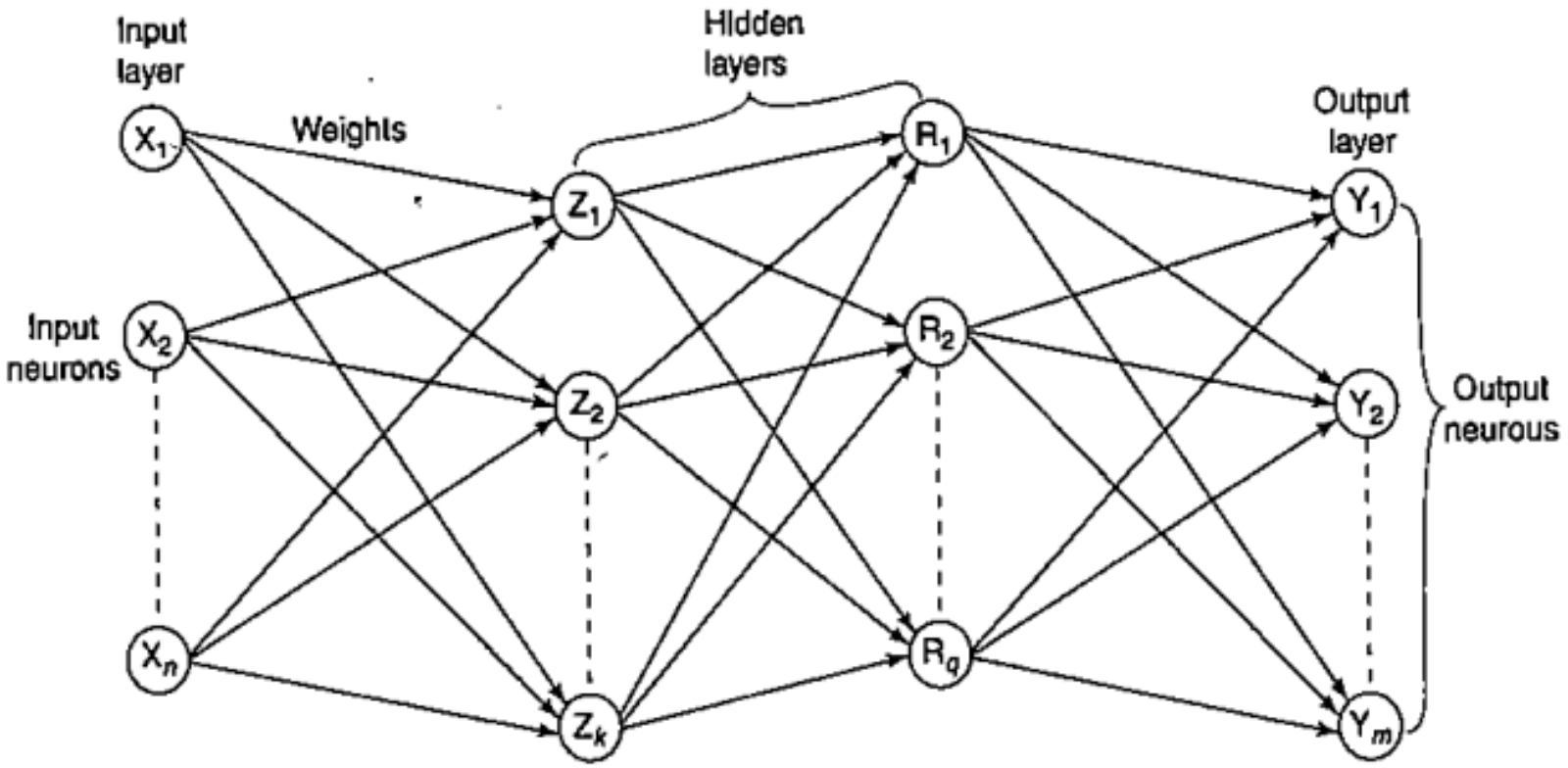
They are:

1. Single-layer feed-forward network;
2. Multilayer feed-forward network;
3. Single node with it's own feedback;
4. Single-layer recurrent network;
5. Multilayer recurrent network.

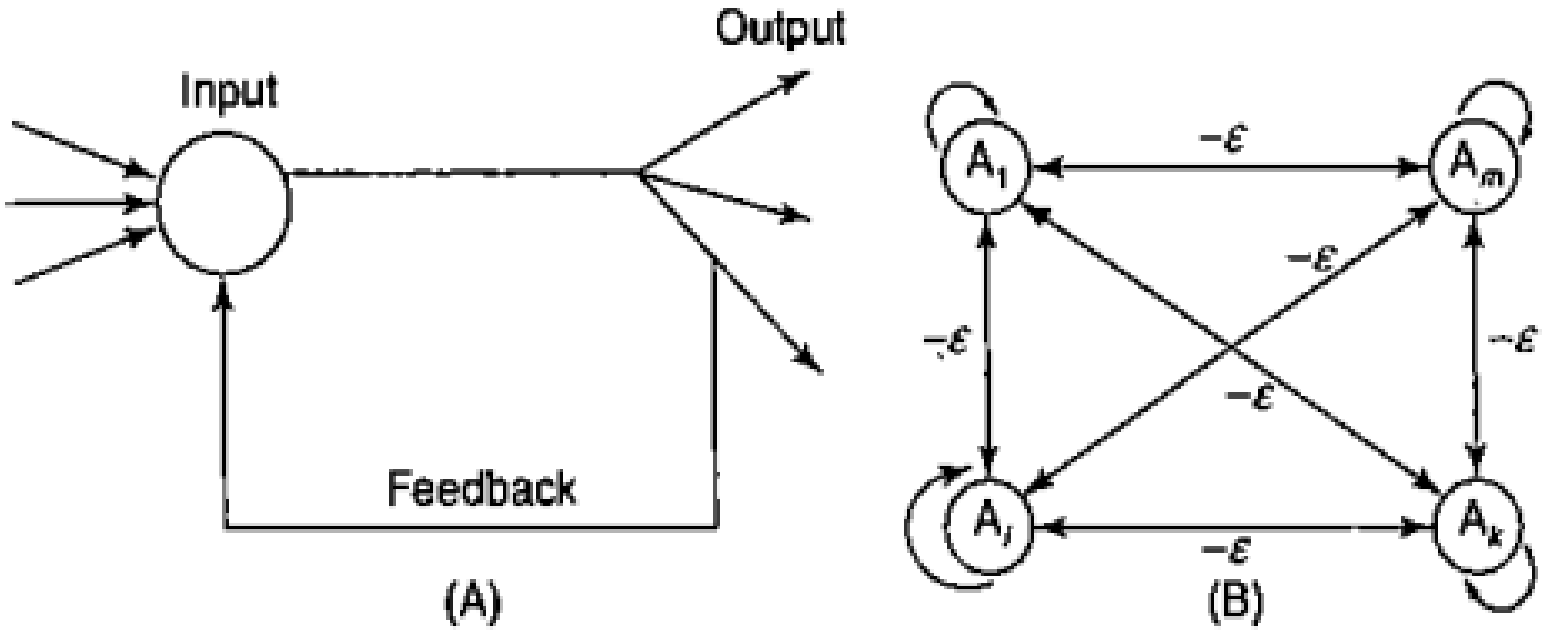
Single layer Feed Forward Network



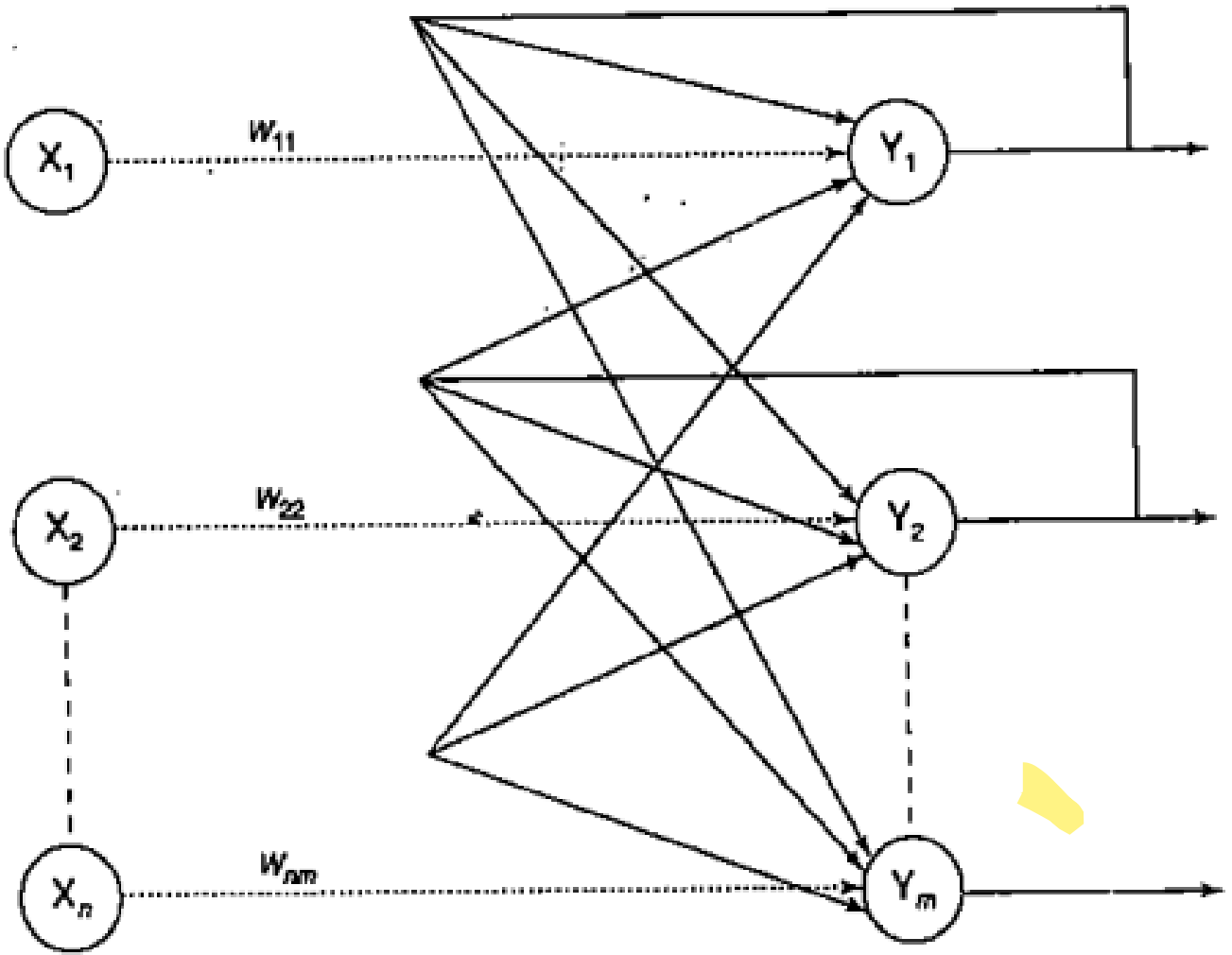
Multi Layer Feed Forward Network



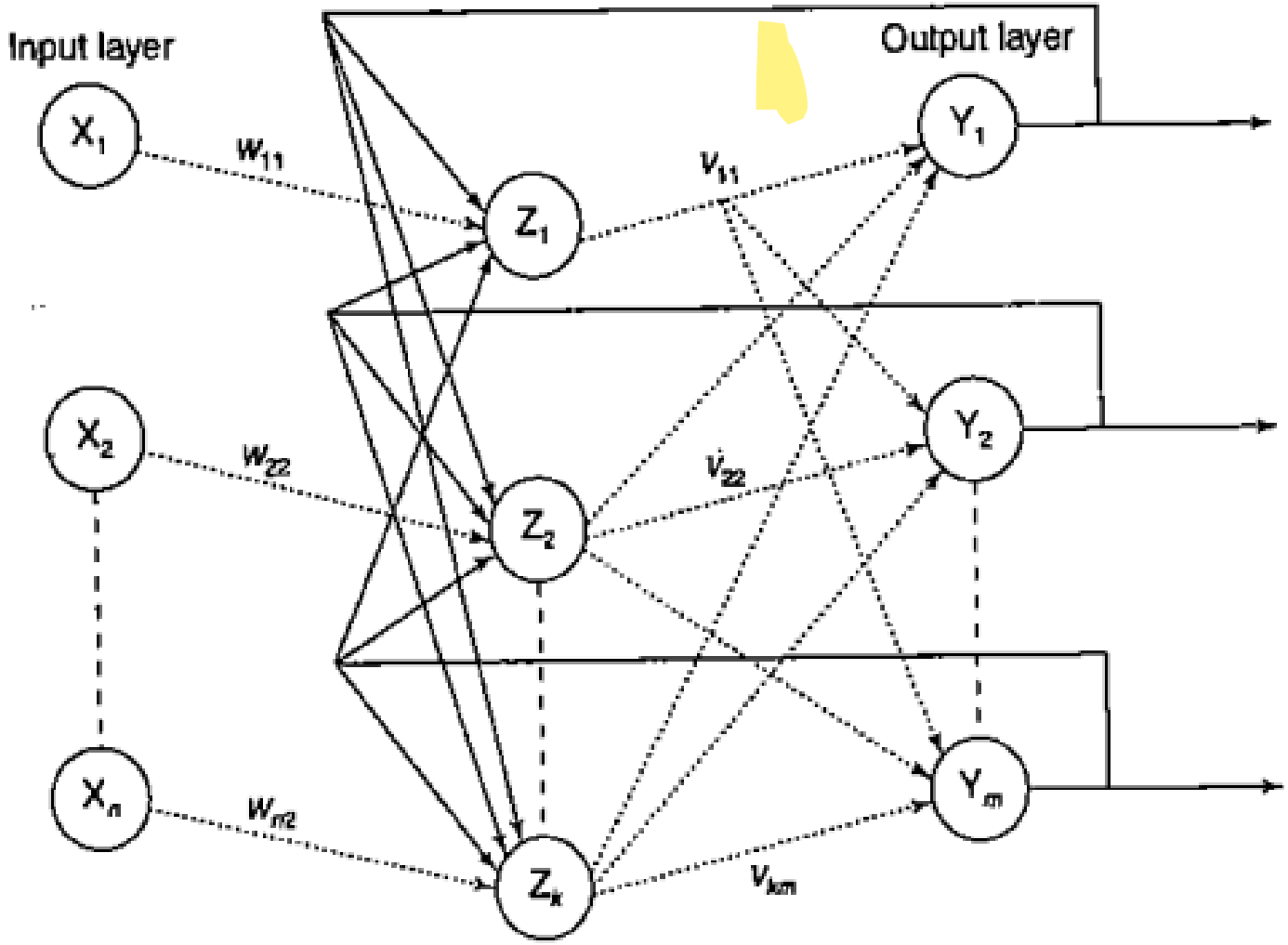
Single Layer with own Feedback



Single Layer Recurrent Network



Multi Layer Recurrent Network



Activation Functions

Sign function:

is one of the nonlinearity used in a neuron, which has value of 1 for positive x and value of -1 for negative x. Mathematically it can be expressed as

$$\sigma(x) = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{otherwise} \end{cases}$$

Activation Functions

Sigmoid function Sigmoid function:

Sigmoid function, also known as logistic function, has as ”S” shape curve. For $x \to \infty$, sigmoid function approaches 1, while for $x \to -\infty$, it approaches to zero. Mathematically sigmoid function can be expressed as,

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

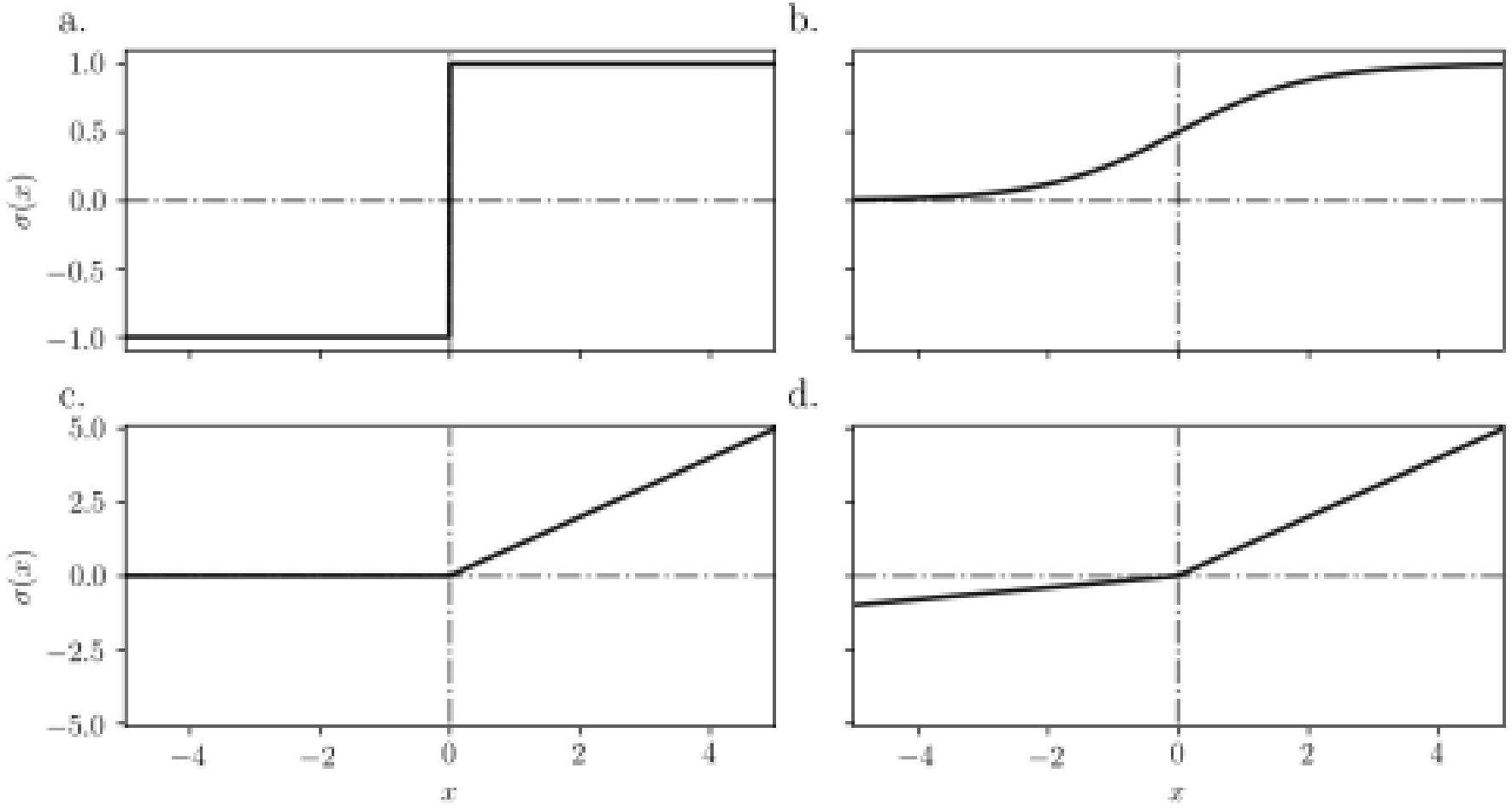
Activation Functions

ReLU function: The rectifier linear unit (ReLU) is an activation function defined as the positive part of its argument

$$\sigma(x) = \max(0, x) = [x]^+$$

Leaky ReLU: Leaky ReLUs a modification of ReLU allow a small, positive gradient when the unit is not active

$$\sigma(x) = \max(\alpha x, x) \text{ where } 0 < \alpha < 1$$


a.
b.
c.
d.
1.0
0.5
0.0
−0.5
−1.0
5.0
2.5
0.0
−2.5
−5.0
σ(x)
−4
−2
0
2
4
x

Important Terminologies of ANNs

- **Weights :** In the architecture of an ANN, each neuron is connected to other neurons by means of directed communication links, and each communication link is associated with weights. The weights contain information about input signal.

Important Terminologies of ANNs

- **Bias:** The bias included in the network has its impact in calculating the net input. The bias is included by adding a component x_o 1 to the input vector us, the input vector becomes:

$$X = (1, X_1, \ldots, X_i, \ldots, X_n)$$

- The bias can be of two types: positive bias and negative bias.
- The positive bias helps in increasing input of the network and the negative bias helps in decreasing the net input of the network, Thus the result of the bias effect, the output of the network.

McCulloch-Pitts Neuron

- Earliest neural network discovered.
- MP Neuron is binary, either the neuron may fire or not.
- Most widely used in logic functions.
- The M-P neuron has no particular training algorithm.
- An analysis has to be performed to determine the values of the weights and the threshold. H
- Here the weights of the neuron are set along with the threshold to make the neuron "perform a simple logic function

Limitation of MP-neurons and Solution

- Weights and thresholds are analytically determined. Cannot learn
- Very difficult to minimize size of a network
- What about non-discrete and/or non-binary tasks?

Perceptron solution [Rosenblatt, 1958]

- Weights and thresholds can be determined analytically or by a learning algorithm
- Continuous, bipolar and multiple-valued versions
- Efficient minimization heuristics exist

Theory of Perceptron

- Perceptron networks come under single-layer feed-forward networks and are also called simple perceptron
- The perceptron network consists of three units, namely, sensory unit (input unit), associator unit (hidden unit), response unit (output unit).
- The perceptron learning rule is used in the weight updation between the associator unit and the response unit. For each training input, the net will calculate the response and it will determine whether or not an error has occurred.

Learning Rule

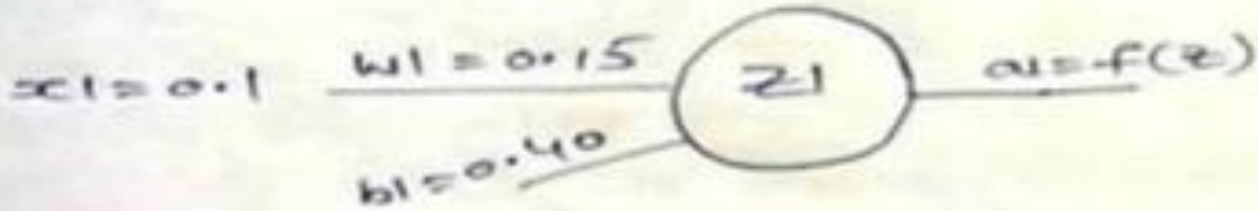
- The weights will be adjusted on the basis of the learning rule if an error has occurred for a particular training pattern

$$w_i(\text{new}) = w_i(\text{old}) + \alpha t x_i$$
$$b(\text{new}) = b(\text{old}) + \alpha t$$

- If no error occurs, there is no weight updation and hence the training process may be stopped

Example

Eg: i) $x1 = 0.1$ & want to predict o/p for this. The nlw has optimized wt. & bias where $w1 = 0.15$ & $b1 = 0.4$.



$$z1 = x1 \cdot w1 + b1$$
$$= (0.1)(0.15) + 0.40$$
$$= 0.415.$$

$$a1 = f(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(0.415)}} = 0.6023$$