

UNIT 1

Overview of biological neurons: Structure of biological neuron, neurobiological analogy, Biological neuron equivalencies to artificial neuron model, Evolution of neural network.

Activation Functions: Threshold functions, Signum function, Sigmoid function, Tan-hyperbolic function, Stochastic function, Ramp function, , Linear function, Identity function.

ANN Architecture: Feed forward network, Feed backward network, single and multilayer network, fully recurrent network,

Course Objectives:

1. To understand the different issues involved in the design and implementation of a Neural Networks.
2. To study the basic of neural network and its activation functions.
3. To understand and use of perceptron and its application in real world
4. To develop an understanding of essential NN concepts such as: learning, feed forward and feed backward
5. To design and build a simple NN model to solve a problem

Neural Networks

- Analogy to biological neural systems,
- Attempt to understand natural biological systems through computational modeling.
- Massive parallelism allows for computational efficiency.
- Intelligent behavior as an “emergent” property of large number of simple units rather than from explicitly encoded symbolic rules and algorithms.

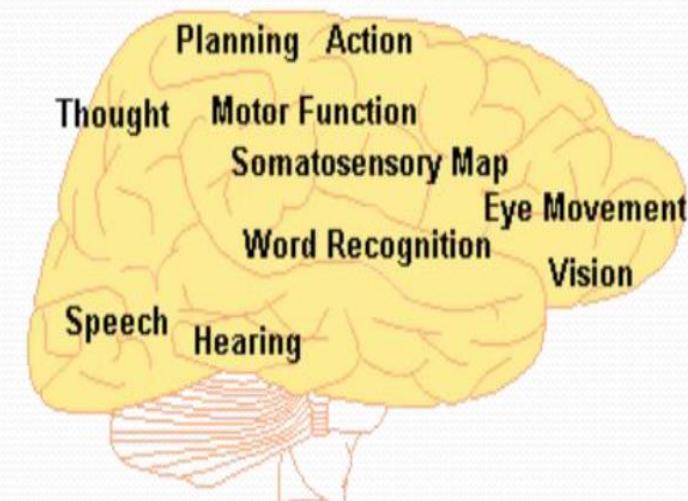
The Biological Neural Network

Characteristics of Human Brain

- Ability to learn from experience
- Ability to generalize the knowledge it possess
- Ability to perform abstraction
- To make errors

Objective

- To emulate or simulate the human brain.

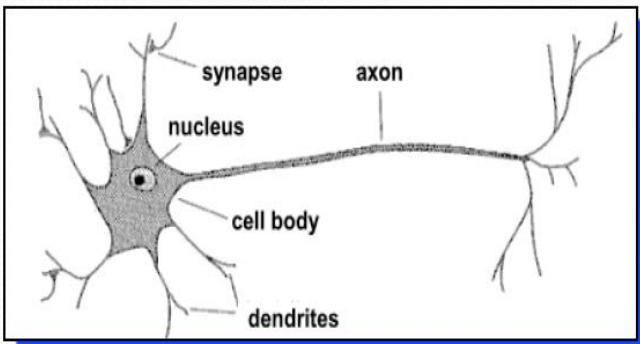


Organization of Human Brain

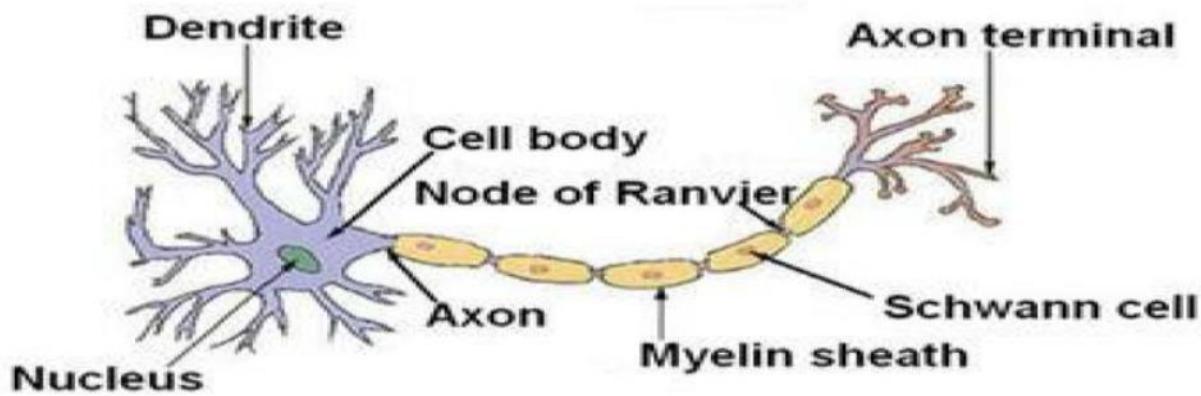


The Neuron

- Fundamental building block of the nervous system
- Performs all the computational and communication functions within the brain
- A many inputs/ one output unit



Structure of a Typical Neuron



The biological neuron has four main regions to its structure

1. The cell body, or soma
2. The axon
3. The dendrites
4. Synapse

Cell body

- It is the heart of the cell. It contains the nucleolus and maintains protein synthesis
- manufactures a wide variety of complex molecules, to keep it renewed for a life time
- manages the energy economy of the neuron
- the outer membrane of the cell body generates nerve impulses.
- Cell body is 5 to 100 microns in dia

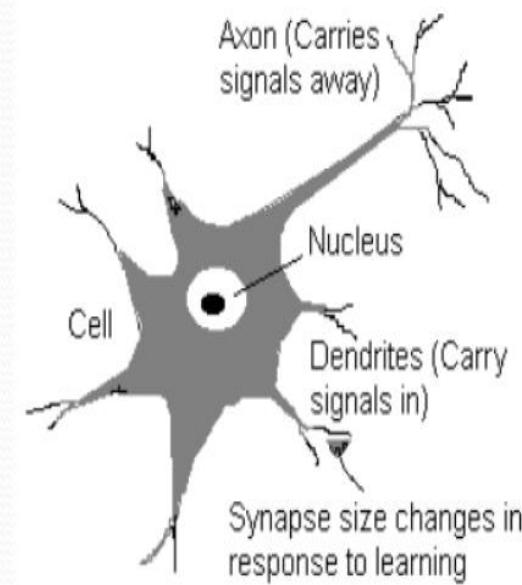
The Axon

- May be as short as 0.1 mm or it is 1 m in length.
- Has multiple branches each terminating in a *synapse*.
- The axon main purpose is to conduct electrical signals generated at the axon down its length. These signals are called *action potentials*

- The other end of the axon may split into several branches, which end in a **pre-synaptic** terminal.
- The myelin is a fatty issue that insulates the axon. The non-insulated parts of the axon area are called **Nodes of Ranvier**.
- At these nodes, the signal traveling down the axon is regenerated. This ensures that the signal travel down the axon to be fast and constant.
- The brain analyzes all patterns of signals sent, and from that information it interprets the type of information received

Dendrites

- bushy branching structure emanating from the cell body.
- Receive the signals from other cells at connection points called **synapses**.
- Usually no physical or electrical connection made at the synapse
- A neuron's dendritic tree is connected to a thousand neighbouring neurons (10,000)



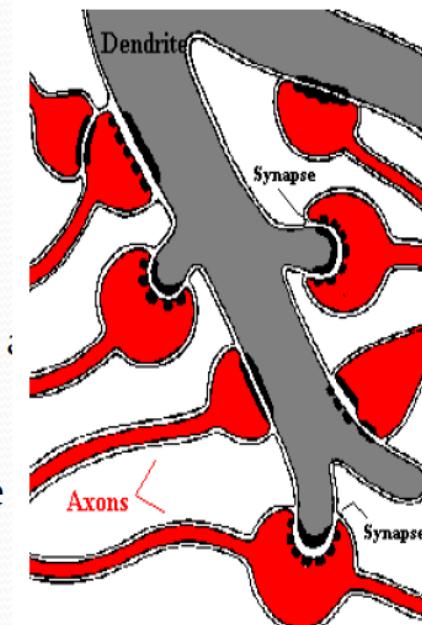
Synapse

- The synapse is the area of contact between two neurons.
- They do not physically touch because they are separated by a cleft.
- The electric signals are sent through chemical interaction.
- The neuron sending the signal is called *presynaptic cell* and the neuron receiving the electrical signal is called *postsynaptic cell*.

The Bio Neuron

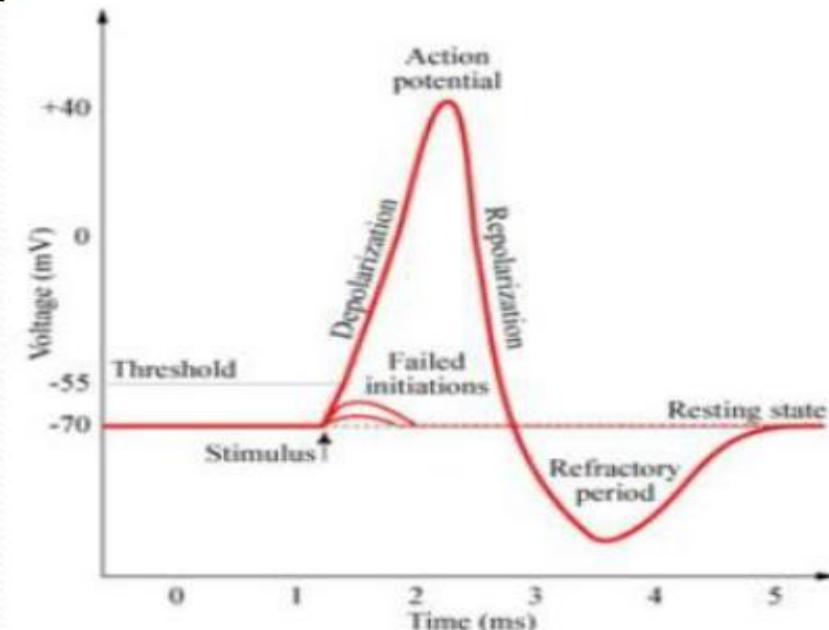
- Neurotransmitters which are specialized chemicals are released by the axon, into the synaptic cleft, diffuse across to the dendrite.
- When one of those neurons fire, a positive or negative charge is received by one of the dendrites. The strengths of all the received charges are added together through the processes of spatial and temporal summation.

- A neuron only fires if its input signal exceeds a certain amount (**threshold**) in a short time period.
- Neurotransmitters are **excitatory**, which tend to produce an output pulse.
- Some are **inhibitory**, which tend to suppress such a pulse

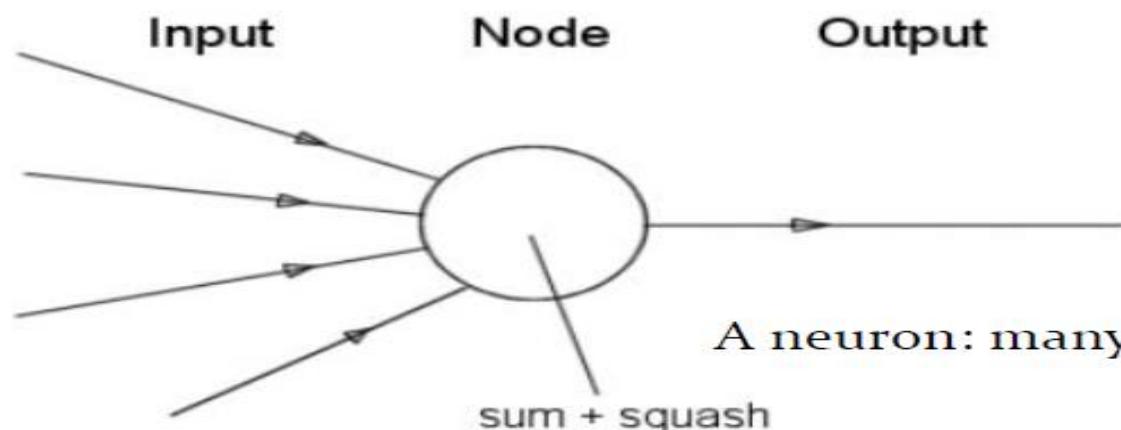
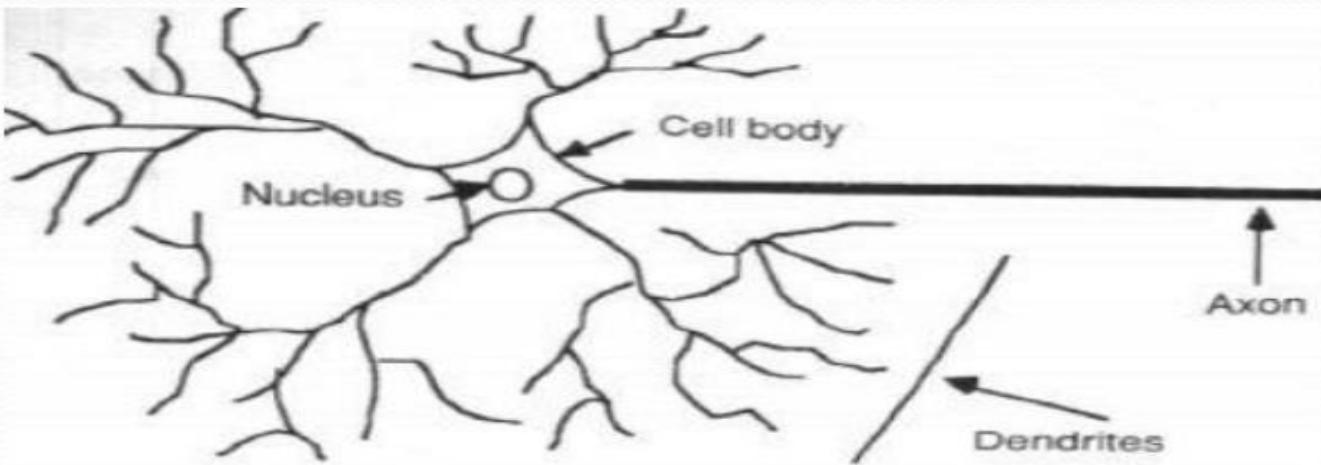


Neural Communication

- Electrical potential across cell membrane exhibits spikes called action potentials.
- Spike originates in cell body, travels down axon, and causes synaptic terminals to release neurotransmitters.
- Chemical diffuses across synapse to dendrites of other neurons.
- If net input of neurotransmitters to a neuron from other neurons is excitatory and exceeds some threshold, it fires an action potential.



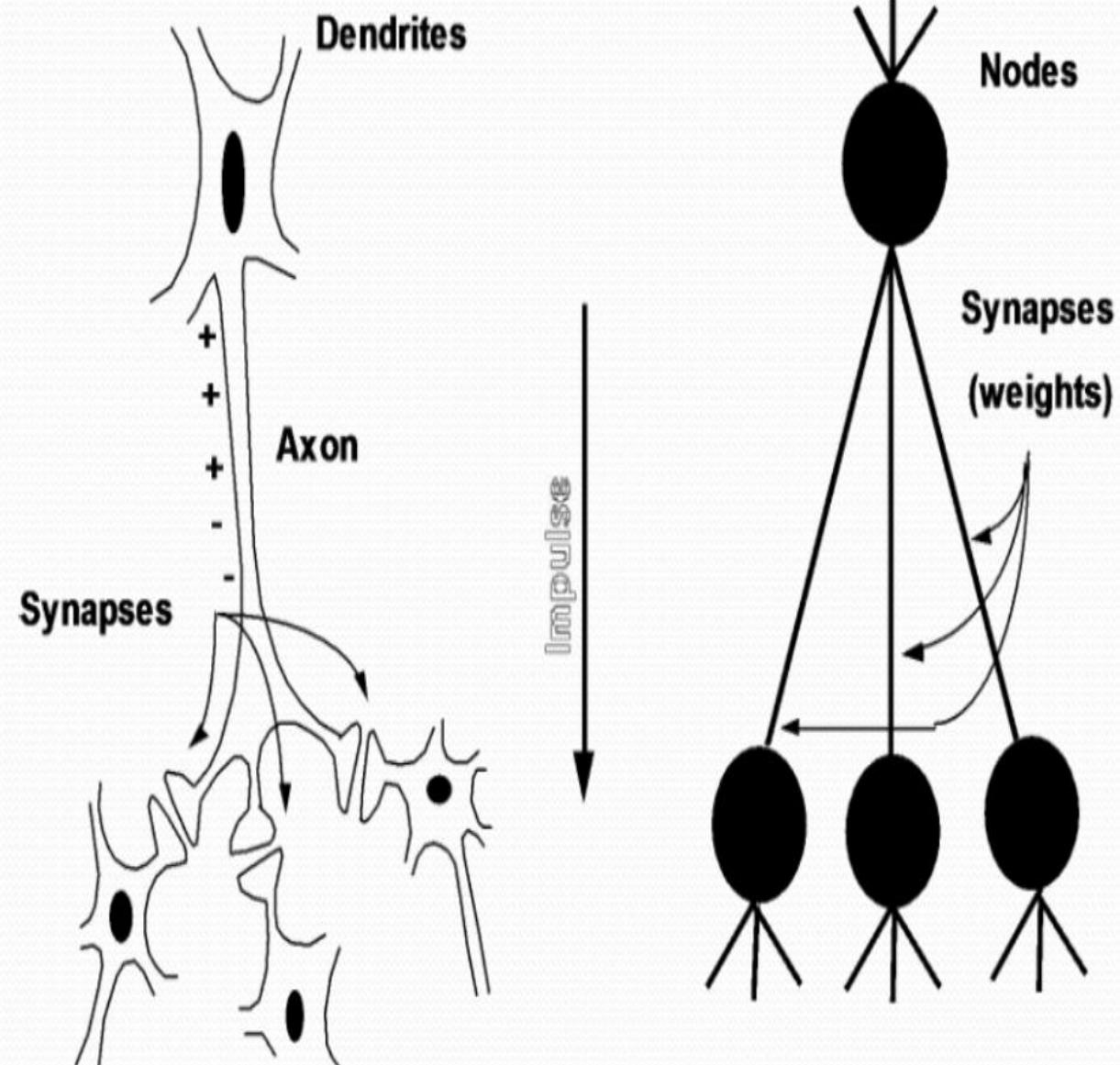
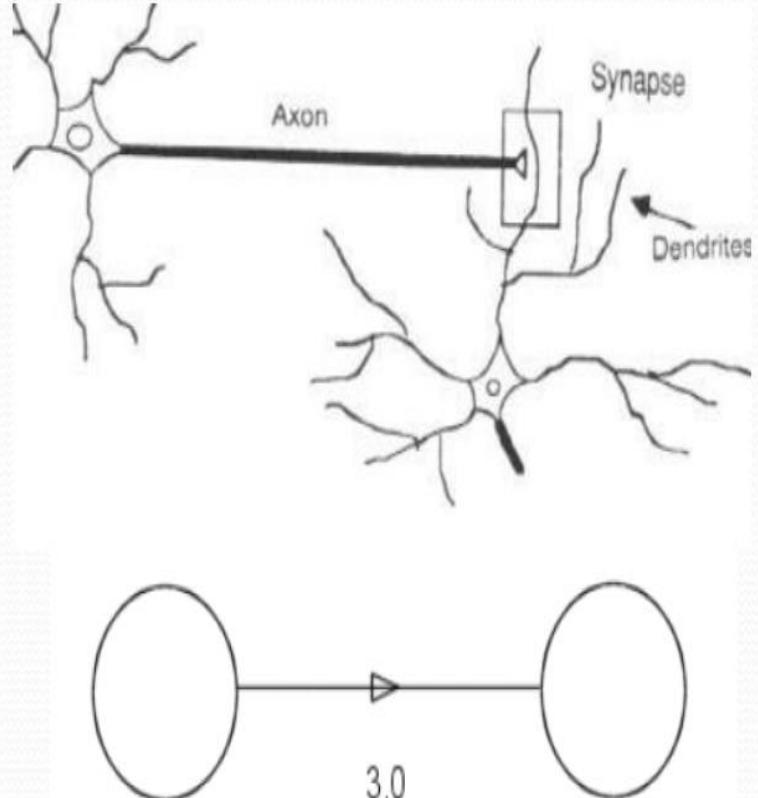
Neuron vs. Node



A neuron: many-inputs / one-output unit

Synapse vs weight

--Axon turn the processed inputs to outputs. --- Synapses are the electrochemical contact between neurons.



- Each neuron receives inputs from other neurons

- A few neurons also connect to receptors.

- Cortical neurons use spikes to communicate.

- The effect of each input line on the neuron is controlled

- by a synaptic weight

- The weights can be positive or negative.

- The synaptic weights adapt so that the whole network learns to perform useful computations

- Recognizing objects, understanding language, making plans, controlling the body.

- We have about 10^{11} neurons each with about 10^4 weights.

- A huge number of weights can affect the computation in a very short time.

Artificial neural networks

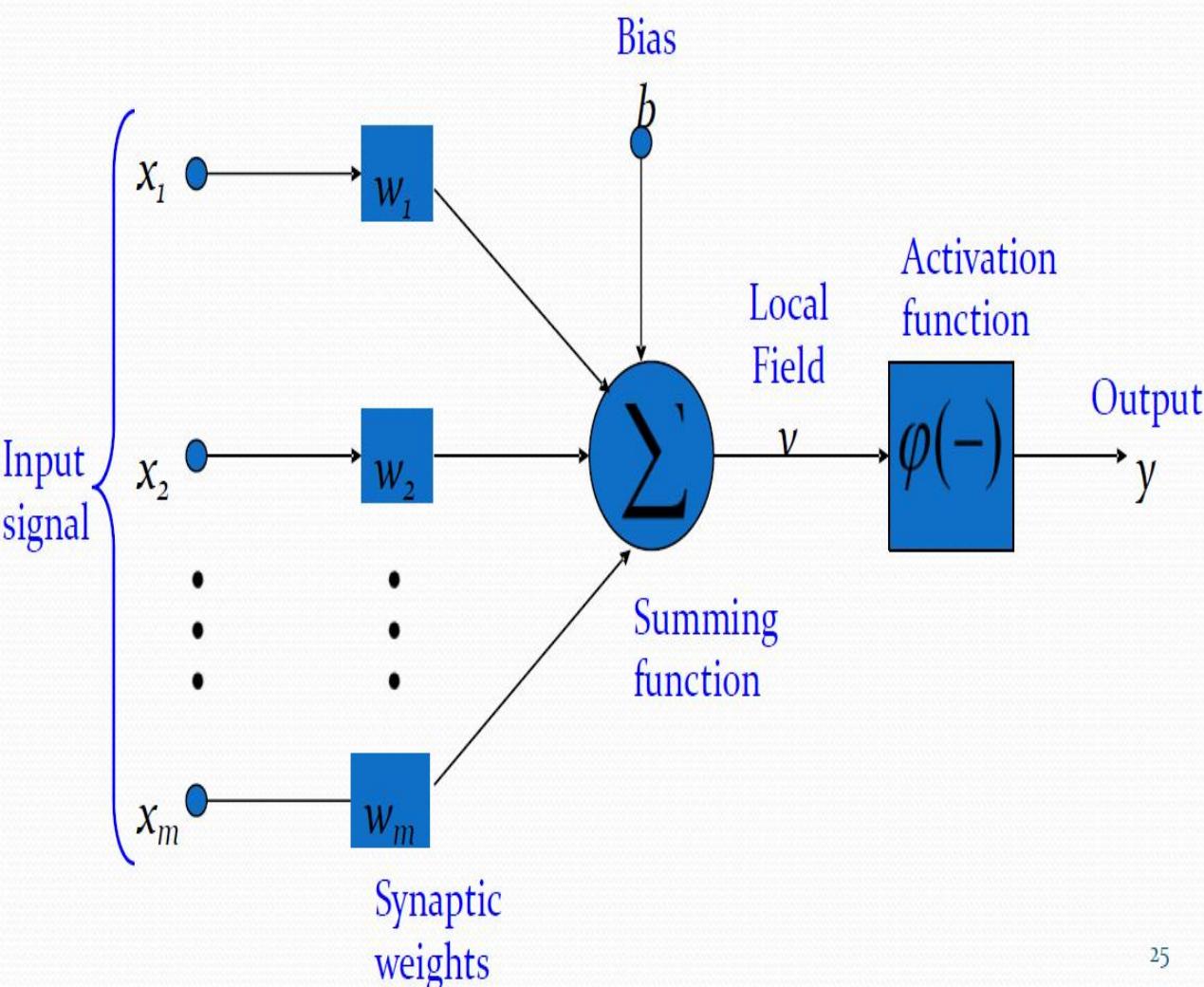
- An artificial neural network is composed of many artificial neurons that are linked together according to a specific network architecture.

- Signals (action potentials) appear at the node's inputs (synapses).

- Each input is multiplied by a certain weight, before being added together at the node (neuron) to produce an overall activation.

- If this exceeds a threshold, the node fires, sending signals to other nodes.

The Artificial Neuron



The Neuron

- The neuron is the basic information processing unit of a NN.
- It consists of:
 - 1 A set of **synapses** or **connecting links**, each link characterized by a **weight**: (W_{kj})
 W_1, W_2, \dots, W_m
$$u = \sum_{j=1}^m w_j x_j$$
 - 2 An **adder** function (combiner) which computes the weighted sum of the inputs:

3 Activation function (squashing function) for limiting the amplitude of the output of the neuron

$Y = \text{activation potential/ induced local field}$

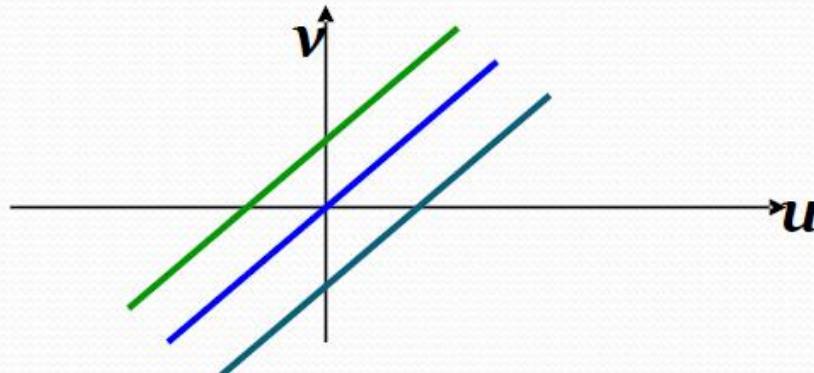
$$y = \varphi(u + b)$$

Bias of a Neuron

- Bias b has the effect of applying an affine transformation to u

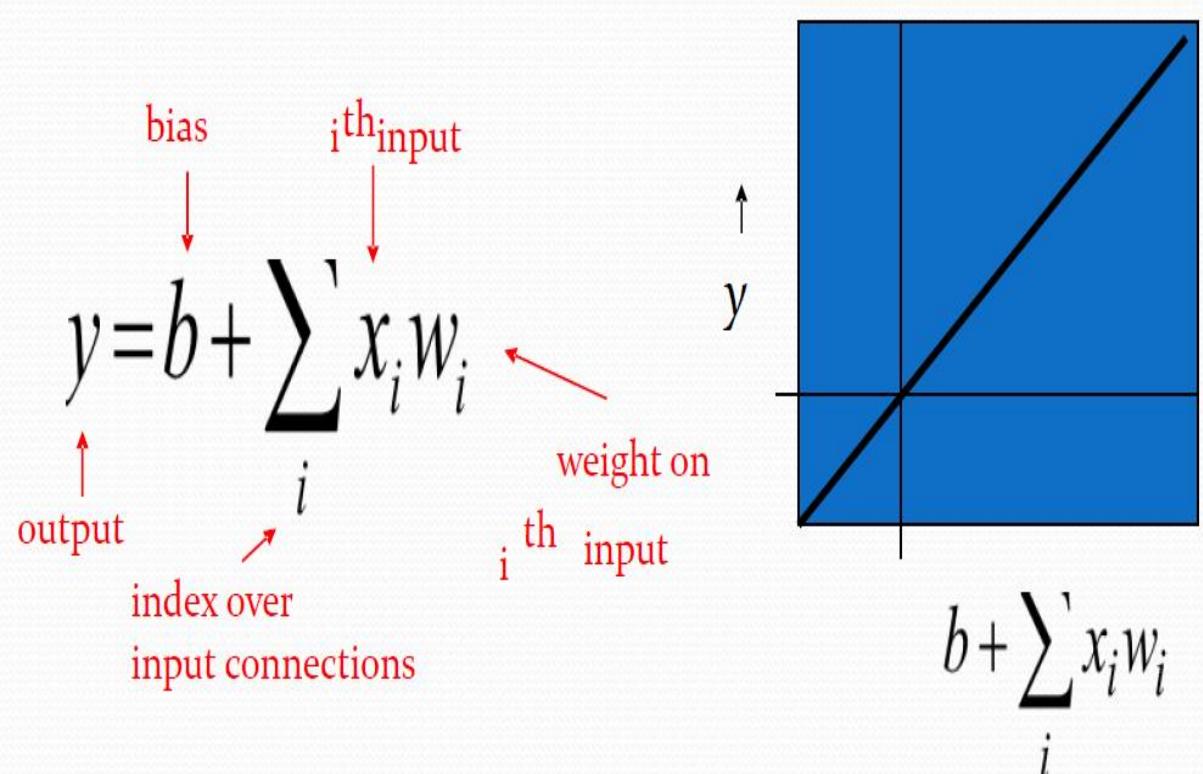
$$v = u + b$$

- v is the **induced field** of the neuron



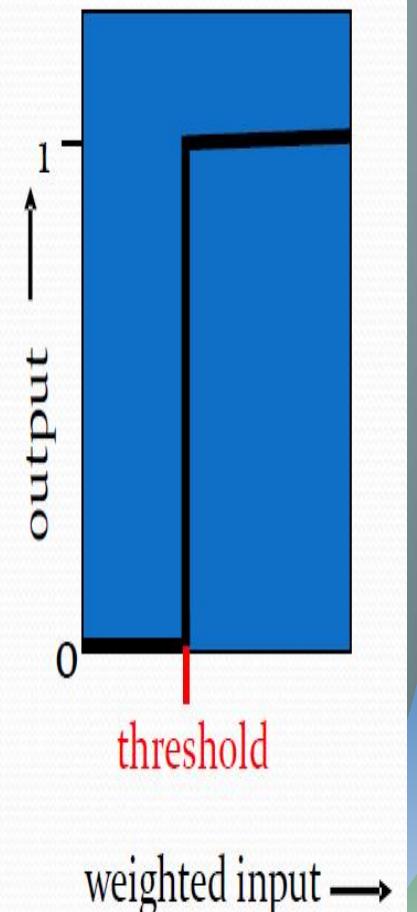
Linear neurons

- These are simple but computationally limited
 - If we can make them learn we **may** get insight into more complicated neurons.



Binary threshold neurons

- McCulloch-Pitts (1943): influenced Von Neumann.
 - First compute a weighted sum of the inputs.
 - Then send out a fixed size spike of activity if the weighted sum exceeds a threshold.
 - McCulloch and Pitts thought that each spike is like the truth value of a proposition and each neuron combines truth values to compute the truth value of another proposition!



Binary threshold neurons

- There are two equivalent ways to write the equations for a binary threshold neuron:

$$v = \sum_i x_i w_i$$

$$\phi(v) = \begin{cases} 1 & \text{if } v \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

$$\boxed{\theta = -b}$$

$$v = b + \sum_i x_i w_i$$

$$y = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

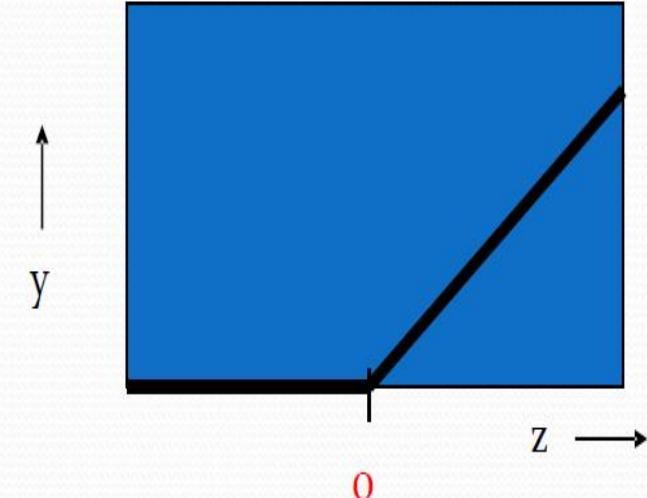
Rectified Linear Neurons

(sometimes called linear threshold neurons)

They compute a **linear** weighted sum of their inputs.
The output is a **non-linear** function of the total input.

$$v = b + \sum_i x_i w_i$$

$$y = \begin{cases} v & \text{if } v > 0 \\ 0 & \text{otherwise} \end{cases}$$



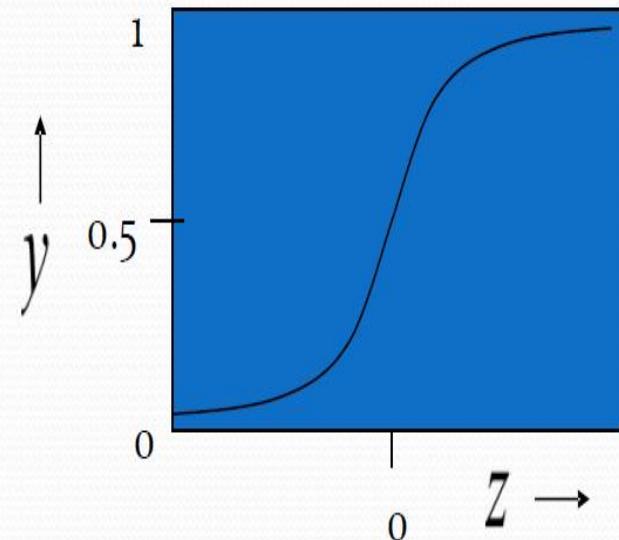
Sigmoid neurons

These give a real-valued output that is a smooth and bounded function of their total input.

- Typically they use the logistic function

- They have nice derivatives which make learning easy

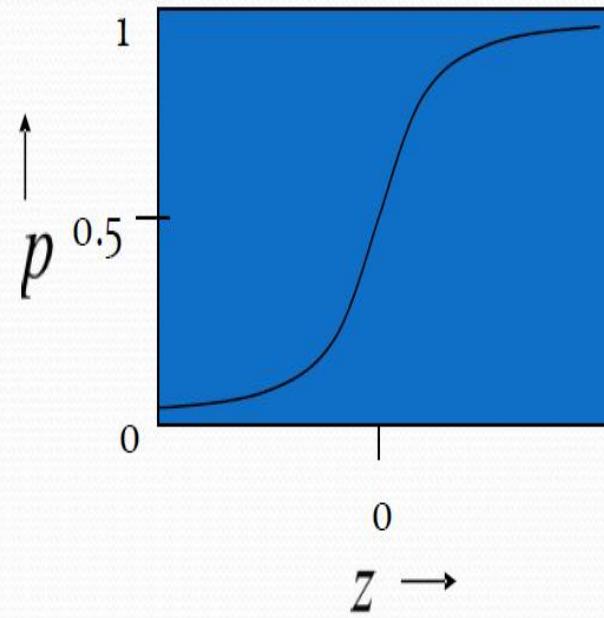
$$v = b + \sum_i x_i w_i \quad y = \frac{1}{1 + e^{-av}}$$



Stochastic binary neurons

- These use the same equations as logistic units.
 - But they treat the output of the logistic as the **probability** of producing a spike in a short time window.
- We can do a similar trick for rectified linear units:
 - The output is treated as the Poisson rate for spikes.

$$z = b + \sum_i x_i w_i \quad p(s=1) = \frac{1}{1 + e^{-z}}$$



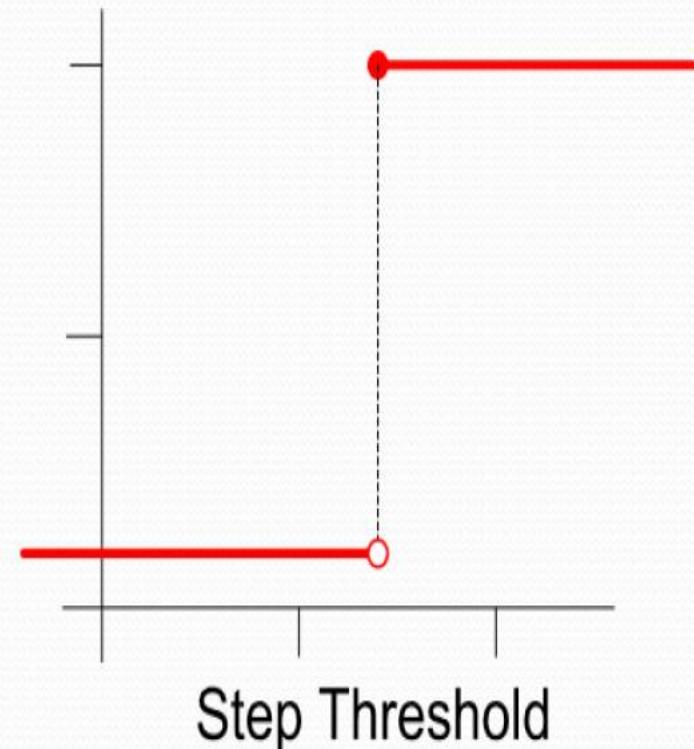
ACTIVATION FUNCTIONS

- To calculate the output response of a neuron
- Transforms neuron's input into output.
- Features of activation functions:
 - A squashing effect is required
 - Prevents accelerating growth of activation levels through the network.
 - Simple and easy to calculate

Threshold Activation Function

- Binary classifier functions

Heaveside function



Binary Threshold Signal Function

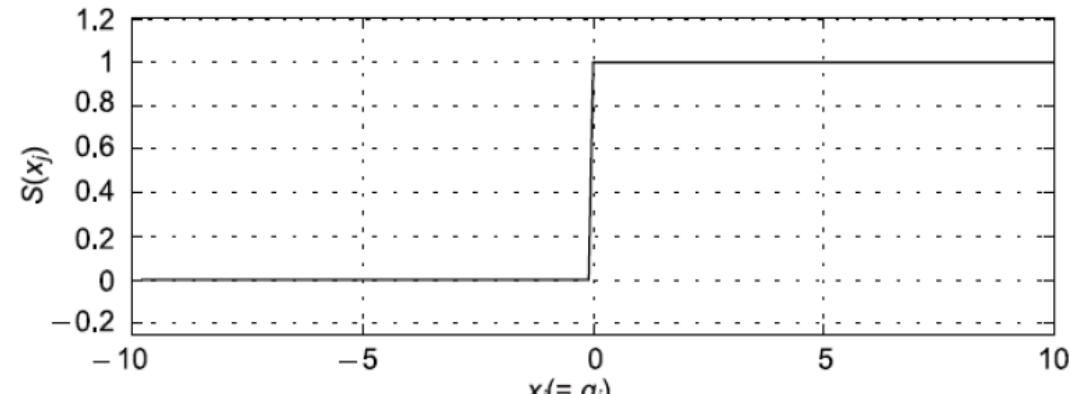
- The threshold logic neuron is a two state machine
 - $s \in S(x) \in \{0, 1\}$

- Net positive activations translate to a +1 signal value
- Net negative activations translate to a 0 signal value.

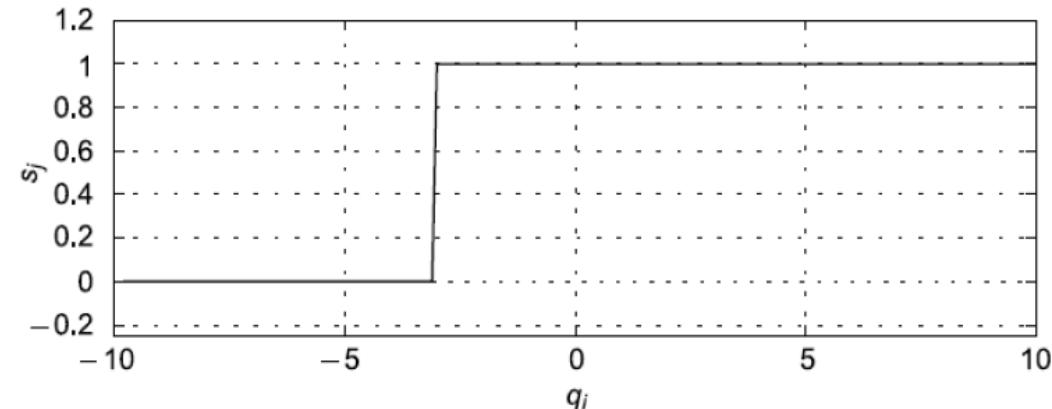
$$\Phi(v) = \begin{cases} 1 & v \geq 0 \\ 0 & v < 0 \end{cases}$$

$$Y_k = \begin{cases} 1 & v \geq 0 \\ 0 & v < 0 \end{cases}$$

Neuron Signal Functions: Binary Threshold Signal Function



(a) Binary threshold function: $\theta_j = 0$

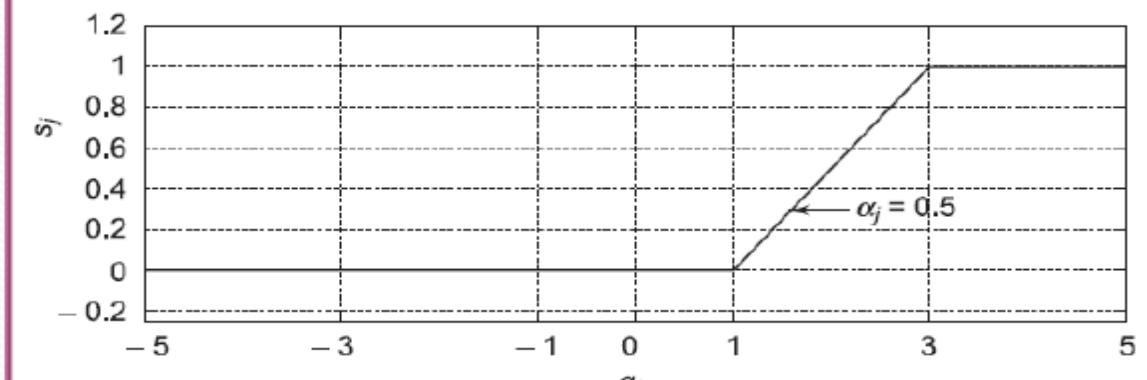
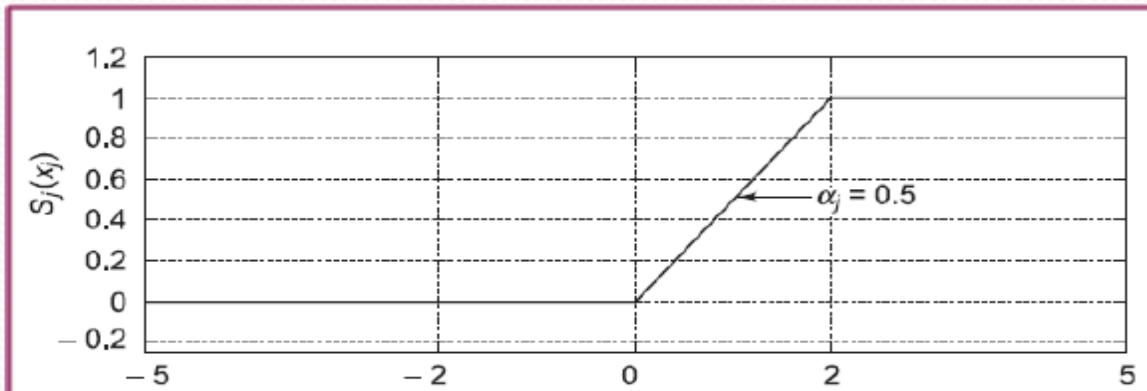


(b) Binary threshold function: $\theta_j = +3$

Linear Threshold Signal Function

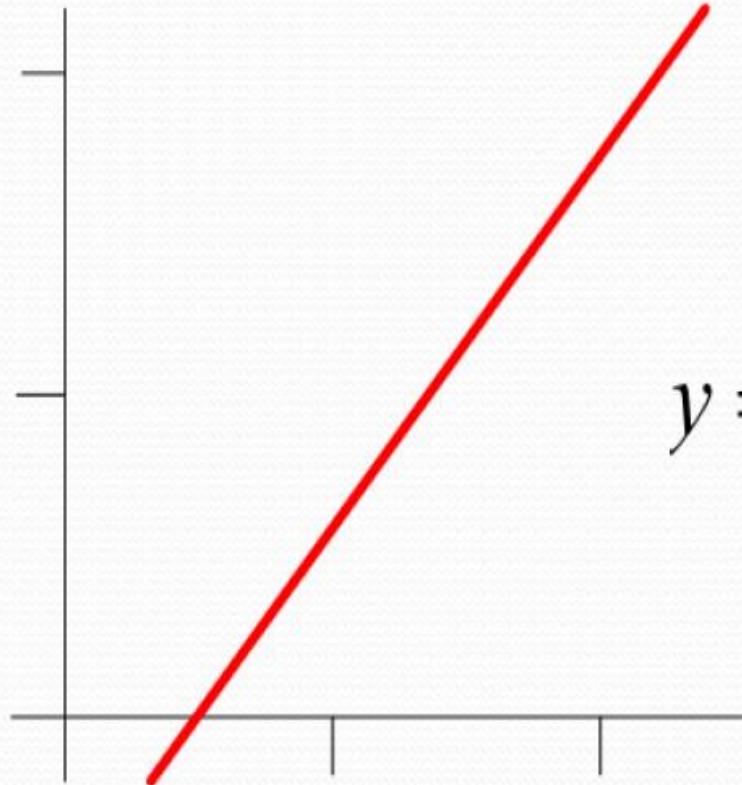
$$S_j(x_j) = \begin{cases} 0 & x_j \leq 0 \\ \alpha_j x_j & 0 < x_j < x_m \\ 1 & x_j \geq x_m \end{cases}$$

- $\alpha_j = 1/x_m$ is the **slope parameter** of the function
- Figure plotted for $x_m = 2$ and $\alpha_j = 0.5$.



Linear Activation functions

- Output is scaled sum of inputs



$$y = u = \sum_{n=1}^N w_n x_n$$

Linear

Threshold Logic Neuron (TLN) in Discrete Time

- The updated signal value $S(x_j^{k+1})$ at time instant $k + 1$ is generated from the neuron activation x_j^{k+1} , sampled at time instant $k + 1$.
- The response of the threshold logic neuron as a two-state machine can be extended to the *bipolar* case where the signals are
 - $s \in \{-1, 1\}$

$$S(x_j^{k+1}) = \begin{cases} 1 & x_j^{k+1} > 0 \\ S(x_j^k) & x_j^{k+1} = 0 \\ 0 & x_j^{k+1} < 0 \end{cases}$$

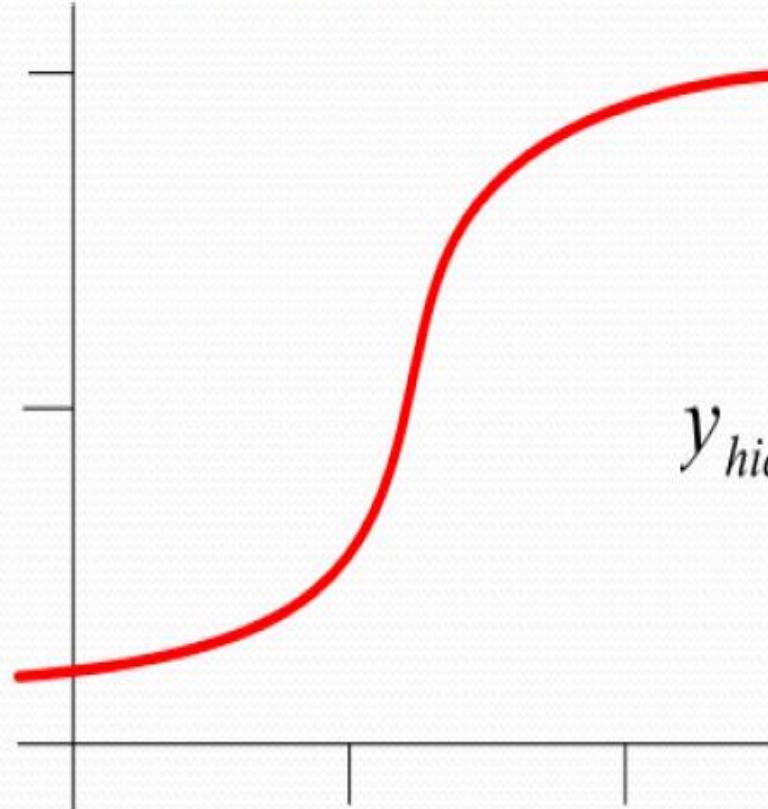
$$S(x_j) = \begin{cases} +1 & x_j > 0 \\ -1 & x_j < 0 \end{cases}$$

Threshold Logic Neuron (TLN) in Discrete Time

- The resulting signal function is then none other than the *signum function*, $\text{sign}(x)$ commonly encountered in communication theory.

Nonlinear Activation Functions

- Sigmoid Neuron unit function

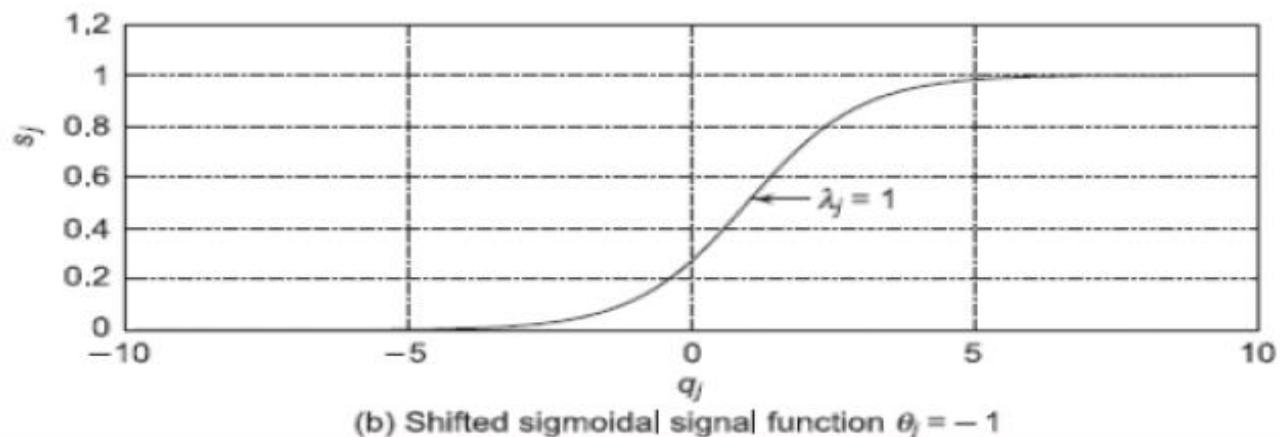
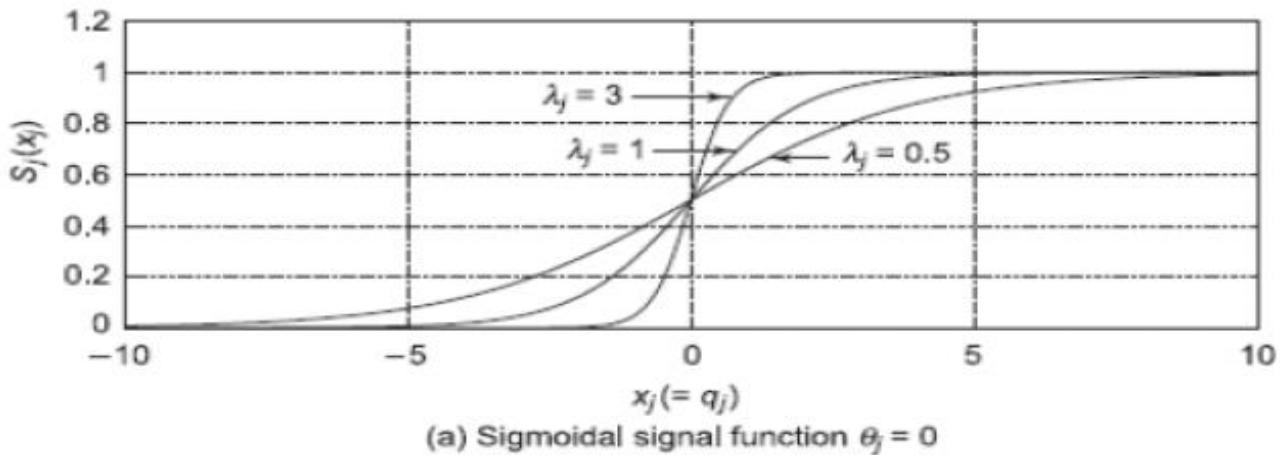


$$y_{hid}(u) = \frac{1}{1 + e^{-u}}$$

Sigmoidal Signal Function

$$s_j(x_j) = \frac{1}{1 + e^{-\lambda_j x_j}}$$

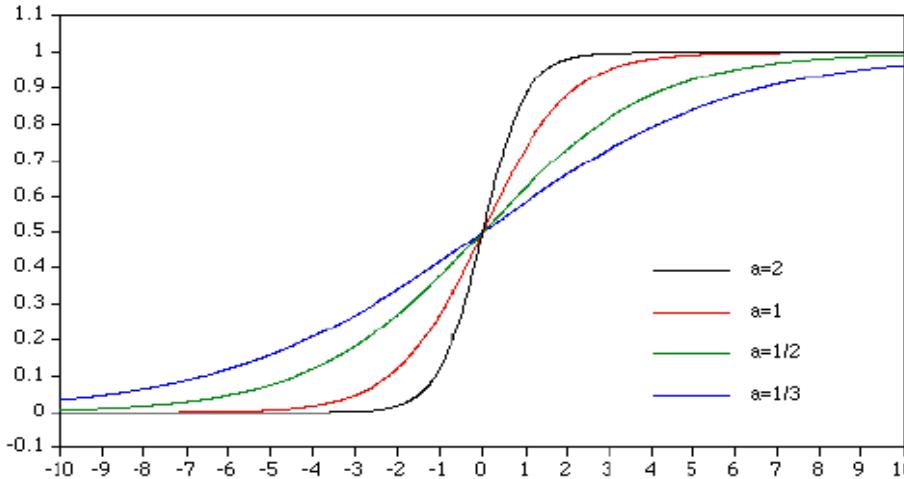
- λ is a gain scale factor
- In the limit, as $\lambda_j \rightarrow \infty$ the smooth logistic function approaches the non-smooth binary threshold function.



Activation Function

- Squashing Function or Logistic Function or Sigmoid Function.

-



$$Y = \frac{1}{1+e^f}$$

$$f = 0$$

$$Y = 0.5$$

$$f > 0$$

$$Y = 1$$

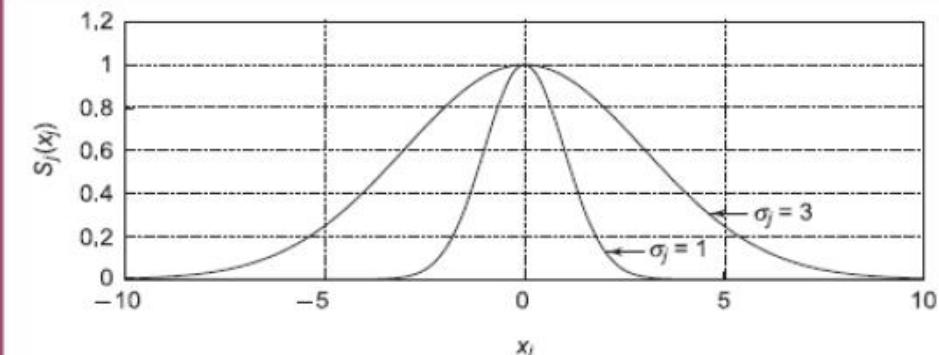
$$f < 0$$

$$Y = 0$$

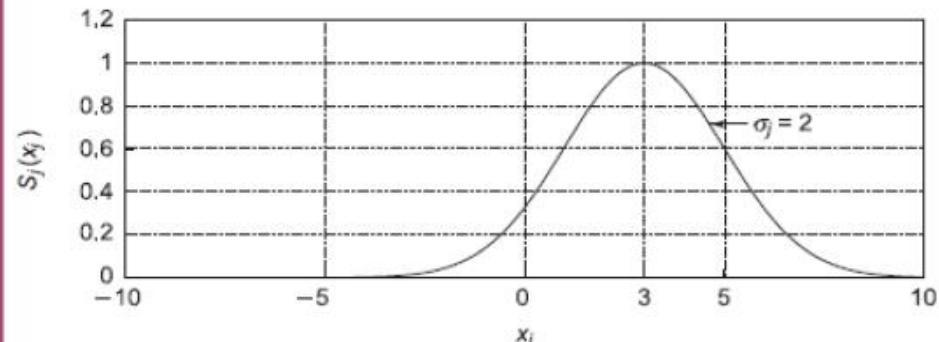
Gaussian Signal Function

$$S_j(x_j) = \exp\left(-\frac{(x_j - c_j)^2}{2\sigma_j^2}\right)$$

- σ is the Gaussian spread factor and c is the center.
- Varying the spread makes the function sharper or more diffuse.



(a) Gaussian signal function: center = 0



(b) Gaussian signal function: center = 3

Stochastic Neurons

- The signal is assumed to be two state
 - $s \in \{0, 1\}$ or $\{-1, 1\}$
- Neuron switches into these states depending upon a *probabilistic function of its activation*, $P(x_j)$.

$$P(x_j) = \frac{1}{1 + e^{-x_j/T}}$$

Summary of Signal Functions

Name	Function	Characteristics
Binary threshold	$s(x_j) = \begin{cases} 1 & x_j \geq 0 \\ 0 & x_j < 0 \end{cases}$	Non-differentiable, step-like, $s_j \in \{0, 1\}$
Bipolar threshold	$s(x_j) = \begin{cases} 1 & x_j \geq 0 \\ -1 & x_j < 0 \end{cases}$	Non-differentiable, step-like, $s_j \in \{-1, 1\}$
Linear	$s_j(x_j) = \alpha_j x_j$	Differentiable, unbounded, $s_j \in (-\infty, \infty)$
Linear threshold	$s_j(x_j) = \begin{cases} 0 & x_j \leq 0 \\ \alpha_j x_j & 0 < x_j < x_m \\ 1 & x_j \geq x_m \end{cases}$	Differentiable, piece-wise linear, $s_j \in [0, 1]$
Sigmoid	$s_j(x_j) = \frac{1}{1 + e^{-\lambda_j x_j}}$	Differentiable, monotonic, smooth, $s_j \in (0, 1)$
Hyperbolic tangent	$s_j(x_j) = \tanh(\lambda_j x_j)$	Differentiable, monotonic, smooth, $s_j \in (-1, 1)$
Gaussian	$e^{-(x_j - c_j)^2 / 2\sigma_j^2}$	Differentiable, non-monotonic, smooth, $s_j \in (0, 1)$
Stochastic	$s_j(x_j) = \begin{cases} +1 & \text{with probability } P(x_j) \\ -1 & \text{with probability } 1 - P(x_j) \end{cases}$	Non-deterministic step-like, $s_j \in \{0, 1\}$ or $\{-1, 1\}$

SECTION – A

- 2. Differentiate between artificial neural networks and biological neurons in detail. Explain various applications of Neural Networks.** 20
- 3. (a) Describe the following learning Rules : 10**
- (i) Hebbian learning rule**
 - (ii) Perception learning rule**
- (b) What is activation function ? Explain the models of Artificial Neural Networks.** 10





THANKS