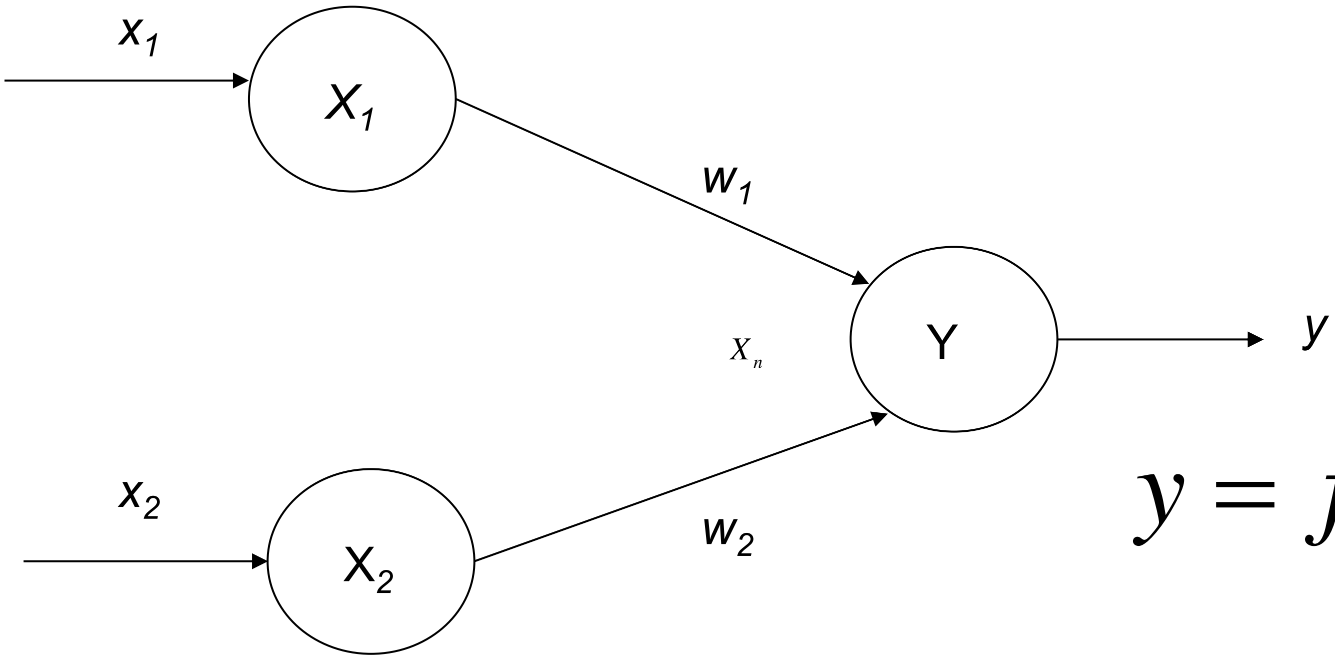
Fundamental concept

- NN are constructed and implemented to model the human brain.
- Performs various tasks such as pattern-matching, classification, optimization function, approximation, vector quantization and data clustering.
- These tasks are difficult for traditional computers

ANN

- ANN posess a large number of processing elements called nodes/neurons which operate in parallel.
- Neurons are connected with others by connection link.
- Each link is associated with weights which contain information about the input signal.
- Each neuron has an internal state of its own which is a function of the inputs that neuron receives- Activation level

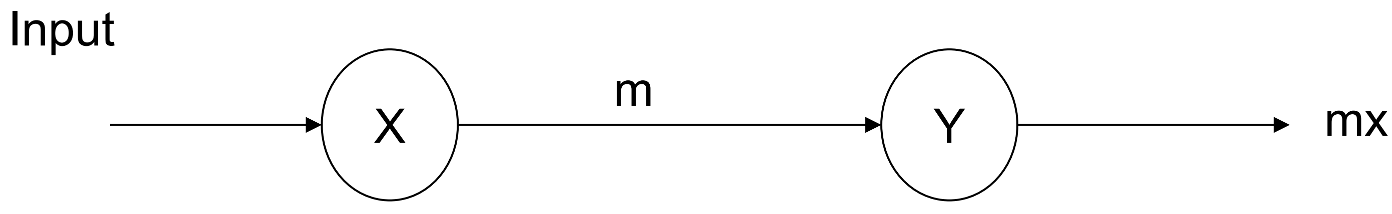
Artificial Neural Networks



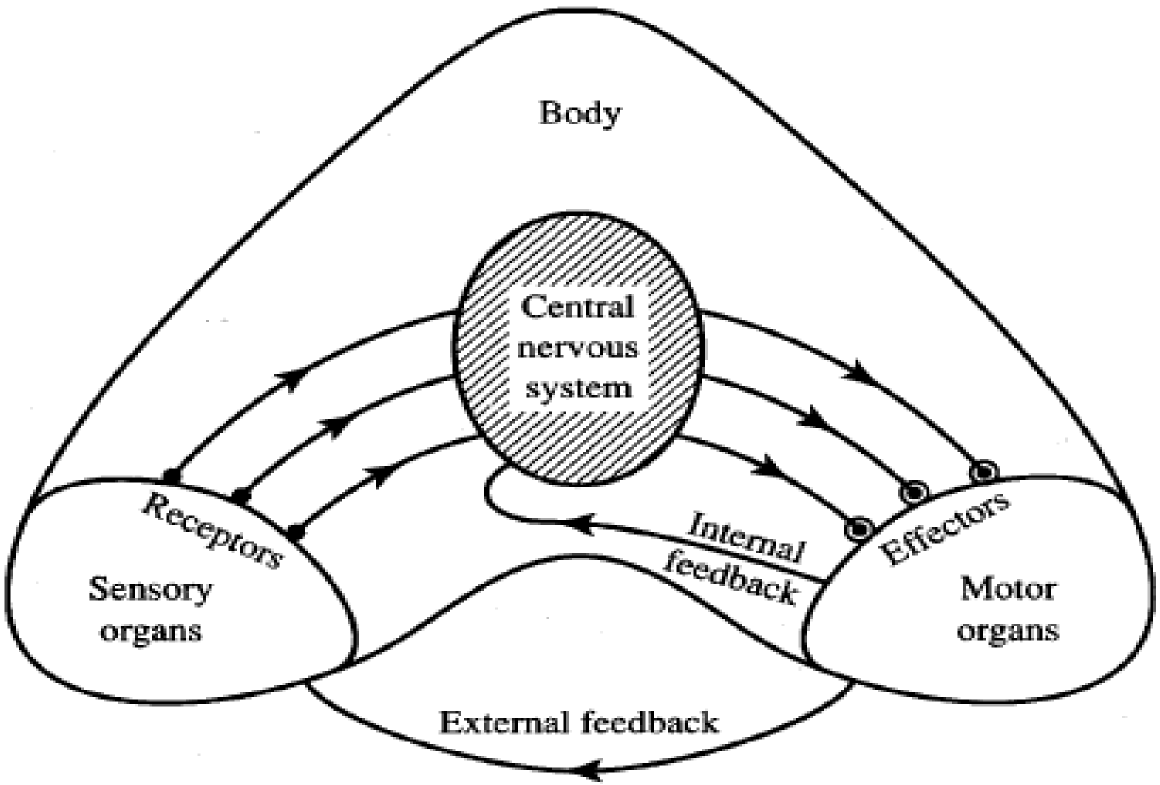
$$y = f(y_{in})$$

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

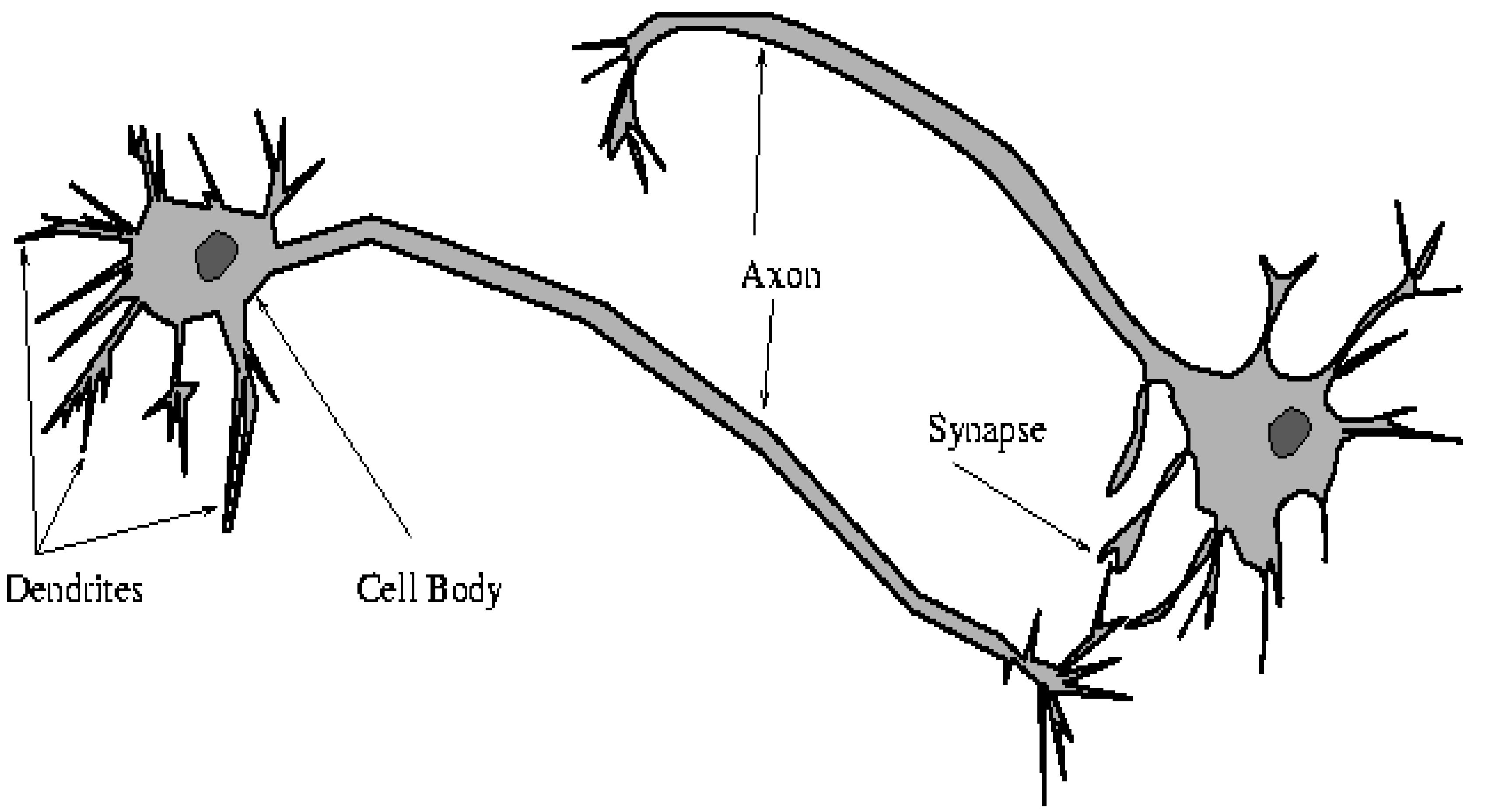
Neural net of pure linear eqn.



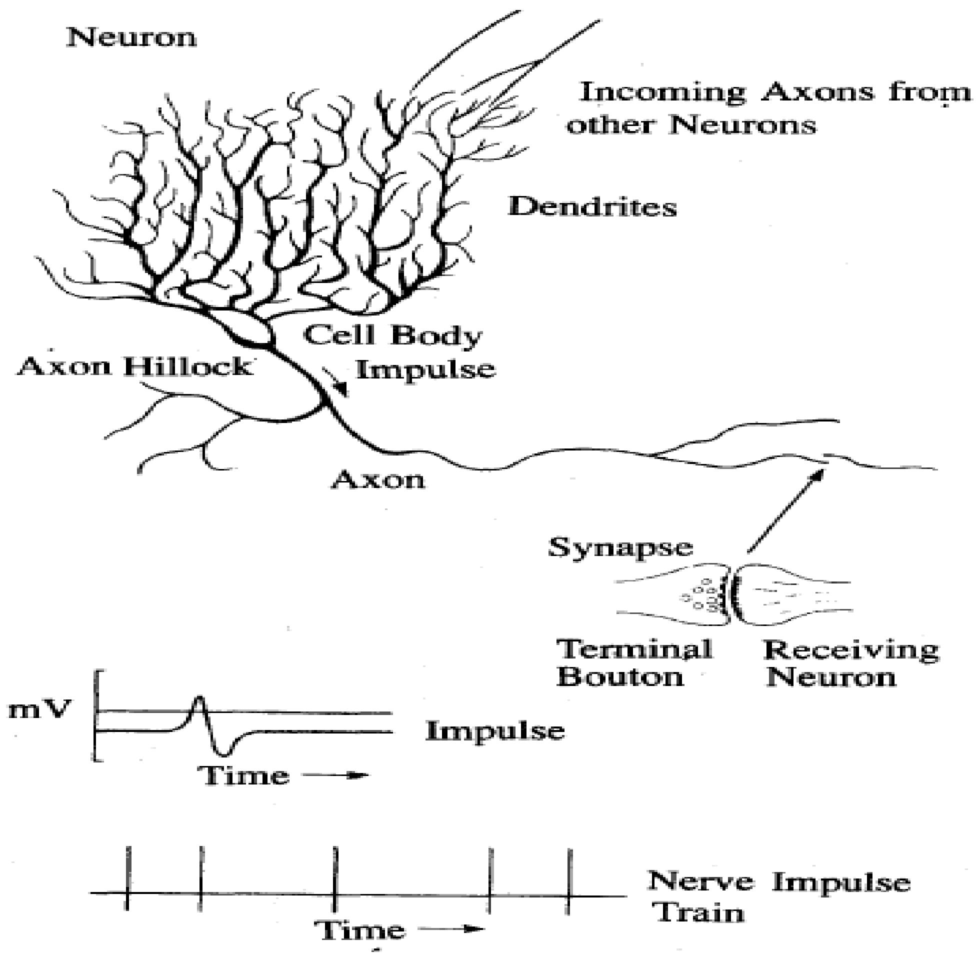
Information flow in nervous system



Biological Neural Network



Neuron and a sample of pulse train

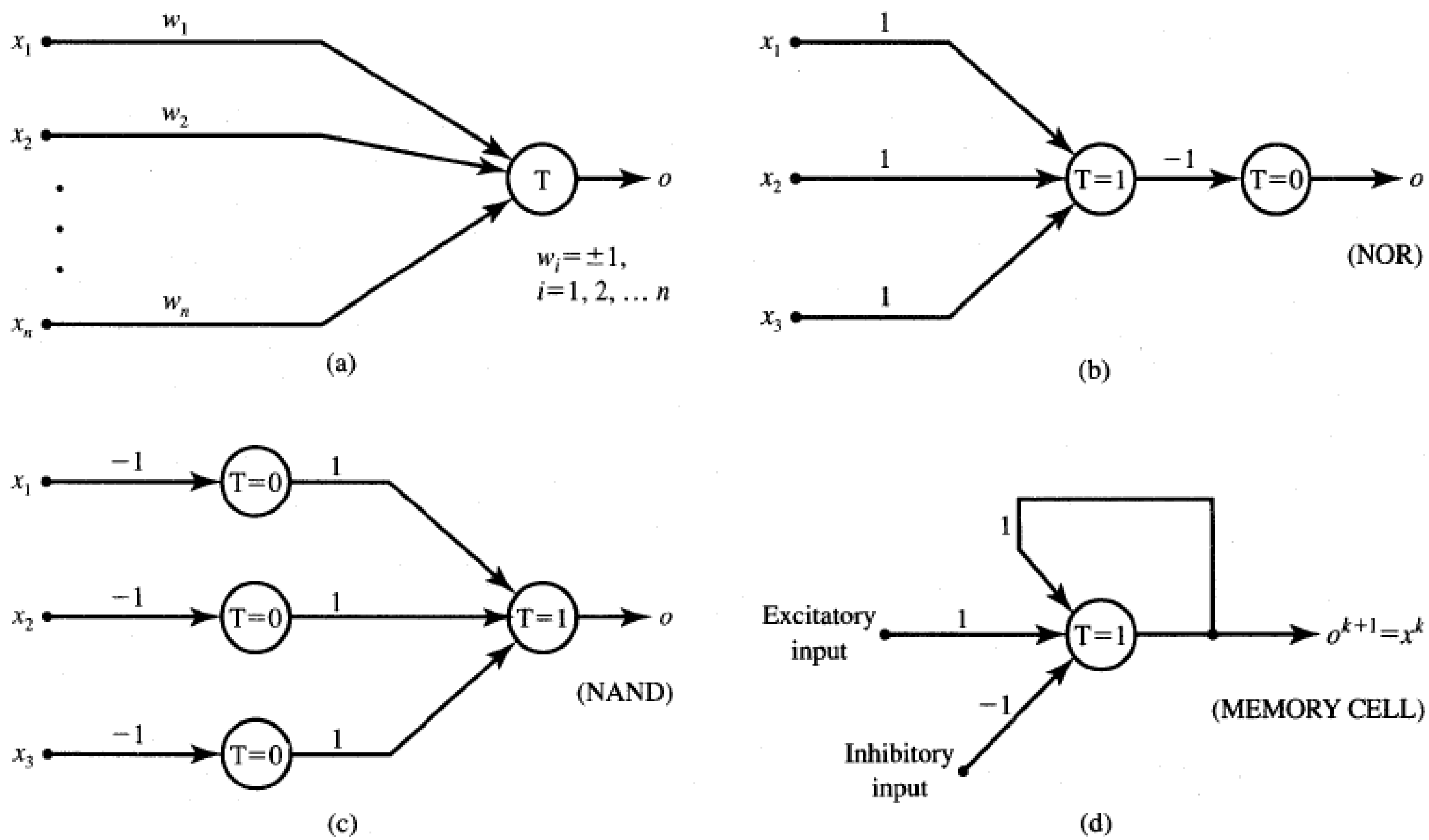


Biological Neuron

- Has 3 parts
  - Soma or cell body:- cell nucleus is located
  - Dendrites:- nerve connected to cell body
  - Axon: carries impulses of the neuron
- End of axon splits into fine strands
- Each strand terminates into a bulb-like organ called synapse
- Electric impulses are passed between the synapse and dendrites
- Synapses are of two types
  - Inhibitory:- impulses hinder the firing of the receiving cell
  - Excitatory:- impulses cause the firing of the receiving cell
- Neuron fires when the total of the weights to receive impulses exceeds the threshold value during the latent summation period
- After carrying a pulse an axon fiber is in a state of complete nonexcitability for a certain time called the refractory period.

McCulloch-Pitts Neuron Model

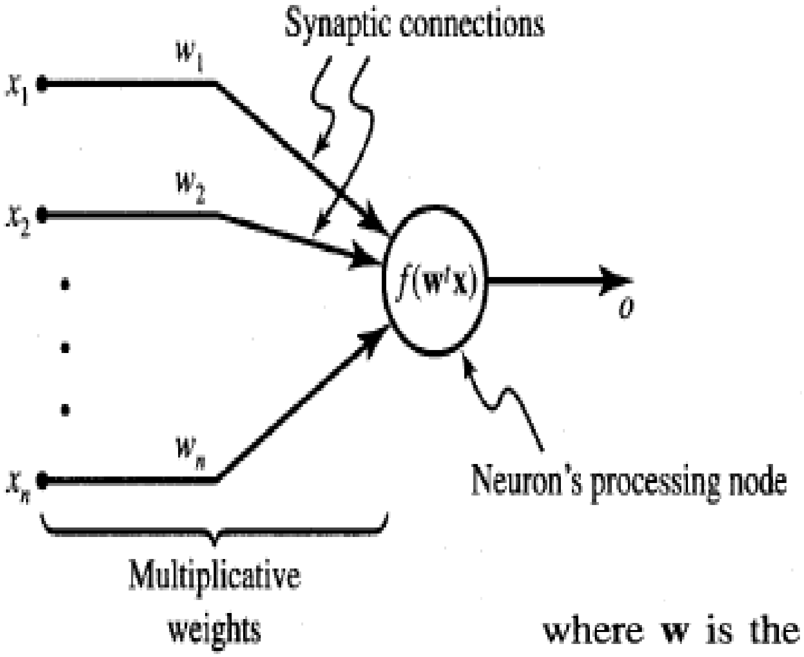
$$o^{k+1} = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w_i x_i^k \geq T \\ 0 & \text{if } \sum_{i=1}^{n} w_i x_i^k < T \end{cases}$$



Features of McCulloch-Pitts model

- Allows binary 0,1 states only
- Operates under a discrete-time assumption
- Weights and the neurons' thresholds are fixed in the model and no interaction among network neurons
- Just a primitive model

General symbol of neuron consisting of processing node and synaptic connections



$$o = f(\mathbf{w}^t\mathbf{x}), \text{ or}$$

$$o = f\left(\sum_{i=1}^{n} w_i x_i\right)$$

where $\mathbf{w}$ is the *weight vector* defined as

$$\mathbf{w} \triangleq [w_1 \quad w_2 \quad \cdots \quad w_n]^t$$

and $\mathbf{x}$ is the input vector:

$$\mathbf{x} \triangleq [x_1 \quad x_2 \quad \cdots \quad x_n]^t$$

Neuron Modeling for ANN

$$o = f(\mathbf{w}^t\mathbf{x}), \text{ or}$$

$$o = f\left(\sum_{i=1}^{n} w_i x_i\right)$$

Is referred to activation function. Domain is set of activation values *net.*

$$net \triangleq \mathbf{w}^t\mathbf{x}$$

Scalar product of weight and input vector

Neuron as a processing node performs the operation of summation of its weighted input.

Activation function

- Bipolar binary and unipolar binary are called as hard limiting activation functions used in discrete neuron model
- Unipolar continuous and bipolar continuous are called soft limiting activation functions are called sigmoidal characteristics.

Activation functions

Bipolar continuous

$$f(net) \stackrel{\Delta}{=} \frac{2}{1+\exp(-\lambda net)} - 1$$

$$\lambda > 0$$

$$f(net) \stackrel{\Delta}{=} \operatorname{sgn}(net) = \begin{cases} +1, & net > 0 \\ -1, & net < 0 \end{cases}$$

Bipolar binary functions

Activation functions

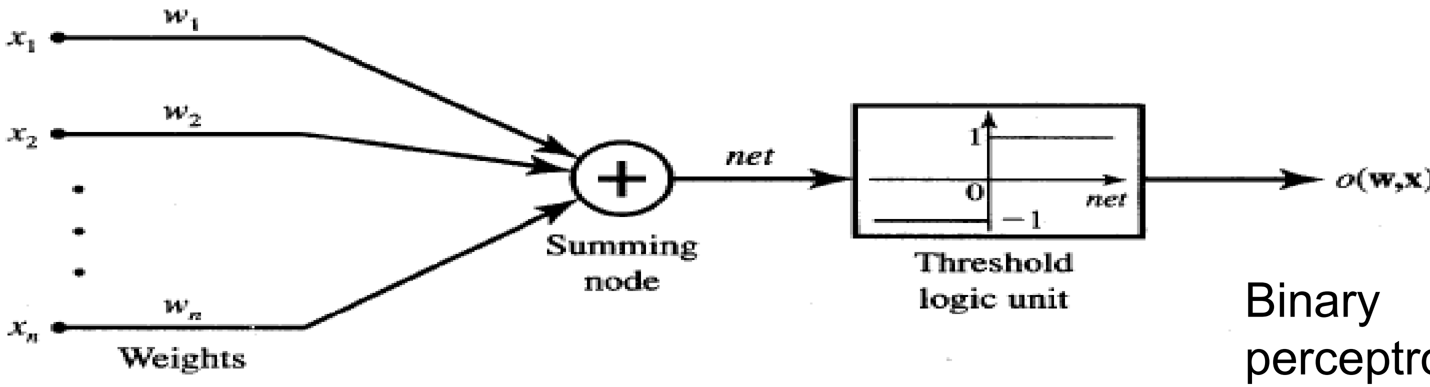
Unipolar continuous

$$f(net) \triangleq \frac{1}{1 + \exp(-\lambda net)}$$

Unipolar Binary

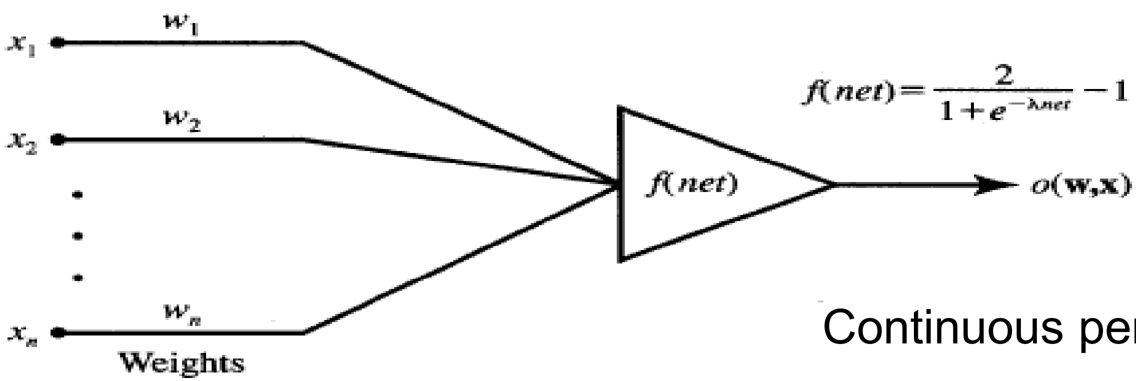
$$f(net) \triangleq \begin{cases} 1, & net > 0 \\ 0, & net < 0 \end{cases}$$

Common models of neurons



(a)

Binary perceptrons



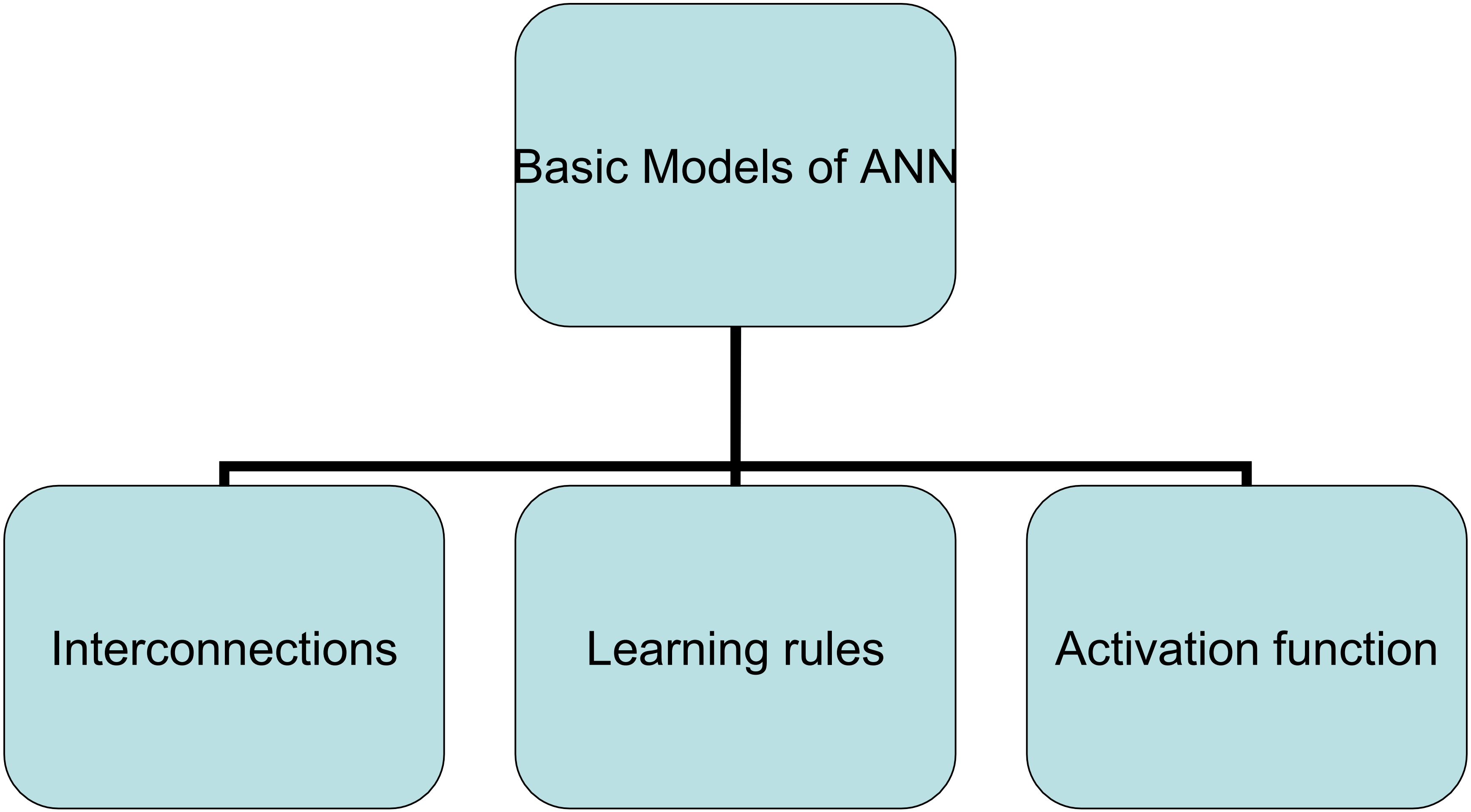
(b)

Continuous perceptrons

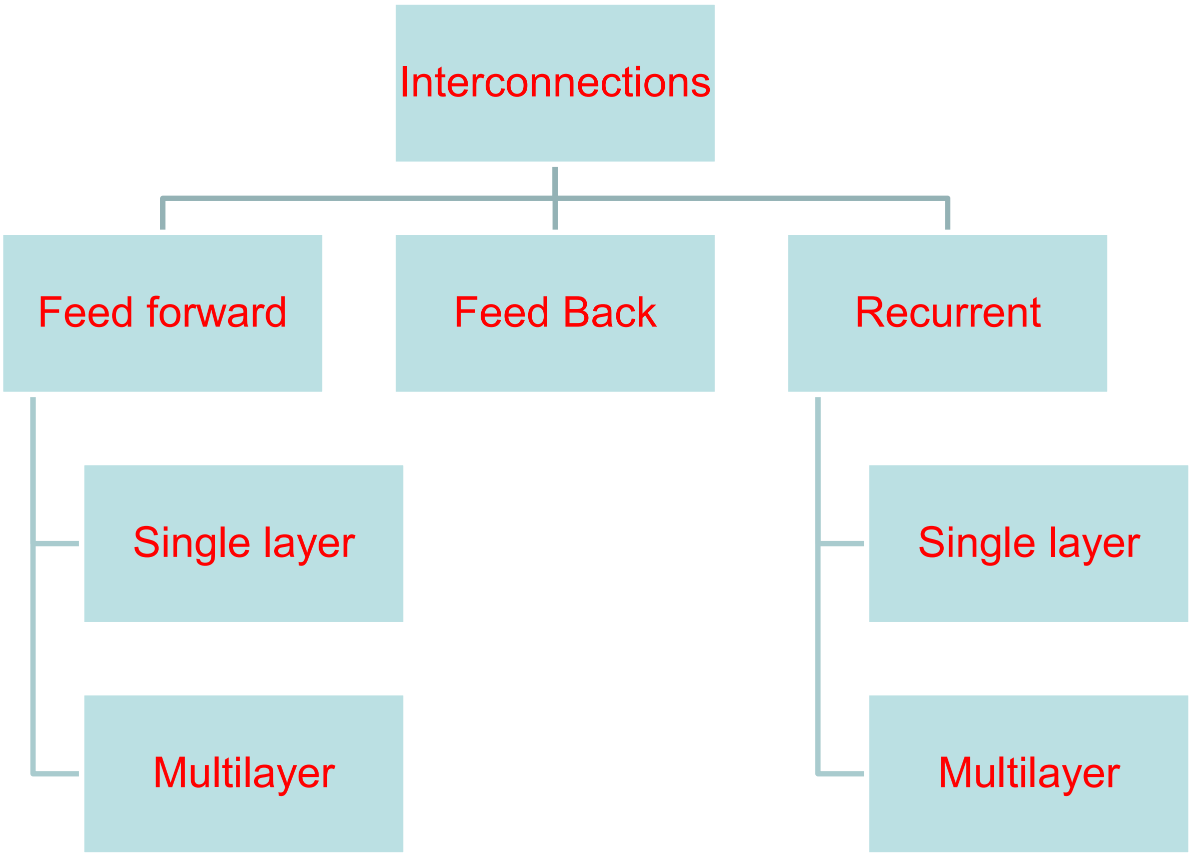
Comparison between brain verses computer

	Brain	ANN
Speed	Few ms.	Few nano sec. massive ||el processing
Size and complexity	10^{11} neurons & 10^{15} interconnections	Depends on designer
Storage capacity	Stores information in its interconnection or in synapse. No Loss of memory	Contiguous memory locations loss of memory may happen sometimes.
Tolerance	Has fault tolerance	No fault tolerance Inf gets disrupted when interconnections are disconnected
Control mechanism	Complicated involves chemicals in biological neuron	Simpler in ANN

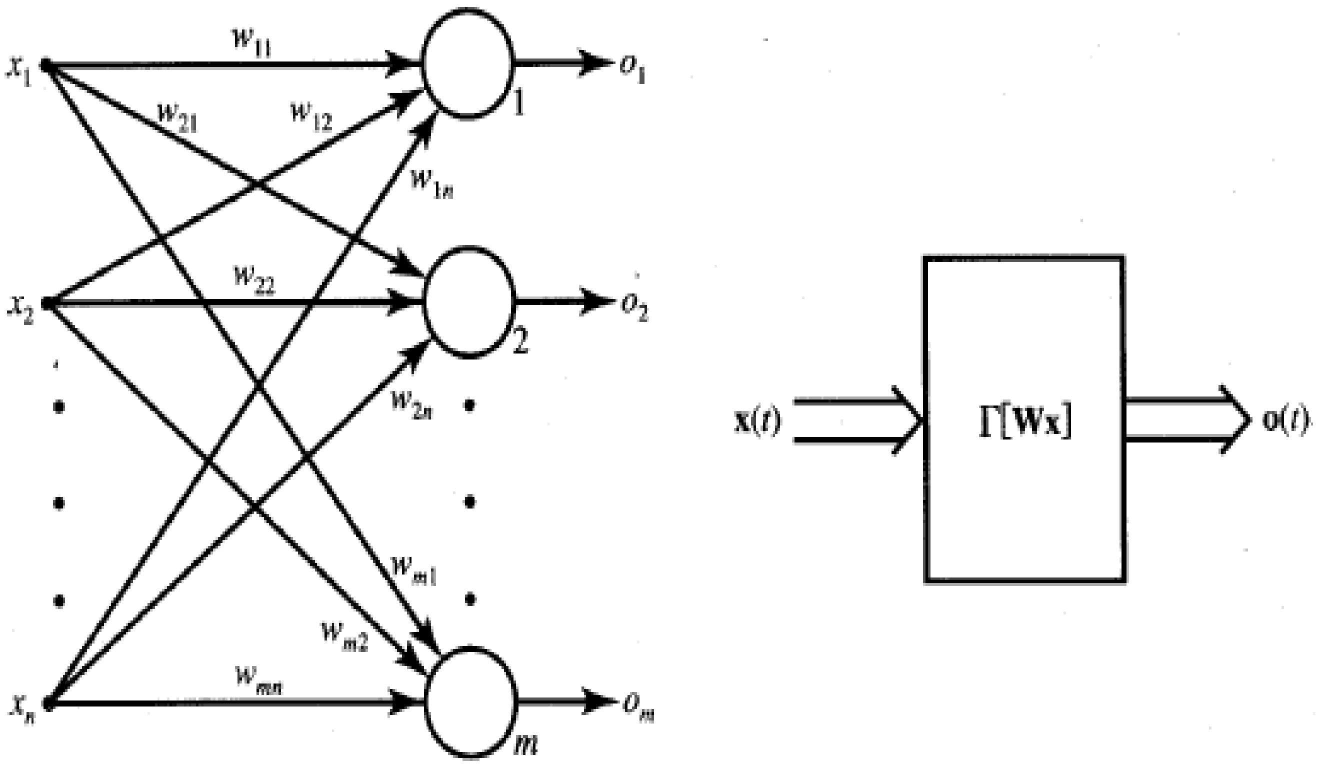
Basic models of ANN



Classification based on interconnections



Single layer Feedforward Network



Feedforward Network

- Its output and input vectors are respectively

$$\mathbf{o} = [o_1 \quad o_2 \quad \cdots \quad o_m]^t$$
$$\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^t$$

- Weight w_{ij} connects the *i'th* neuron with *j'th* input. Activation rule of ith neuron is

$$net_i = \sum_{j=1}^{n} w_{ij}x_j, \quad \text{for } i = 1, 2, \ldots, m$$

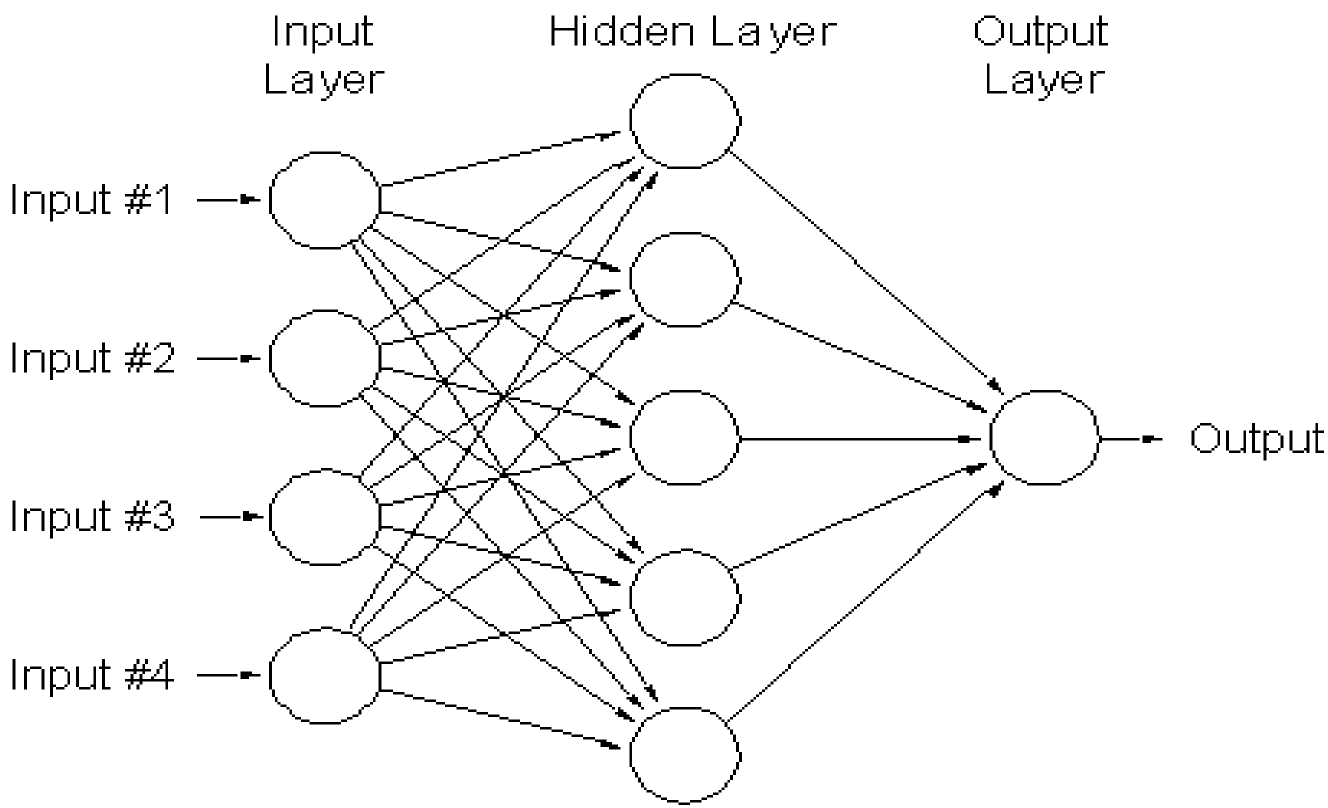
$$o_i = f(\mathbf{w}_i^t\mathbf{x}), \quad \text{for } i = 1, 2, \ldots, m$$

where

$$\mathbf{w}_i \triangleq [w_{i1} \quad w_{i2} \quad \cdots \quad w_{in}]^t$$

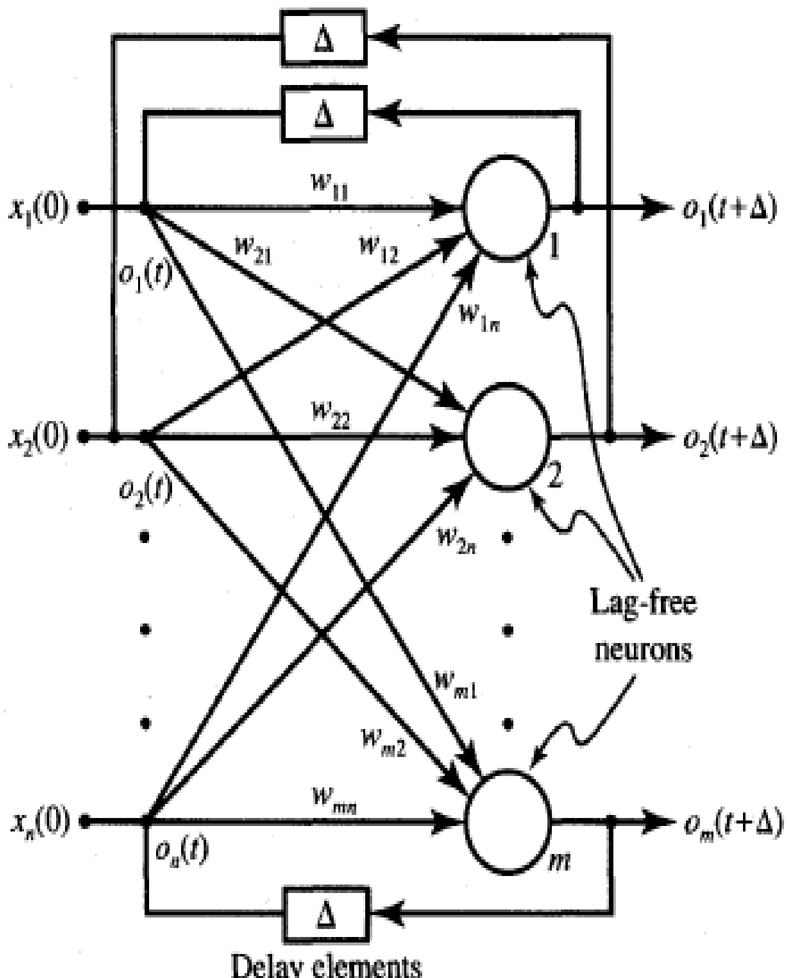
EXAMPLE

Multilayer feed forward network

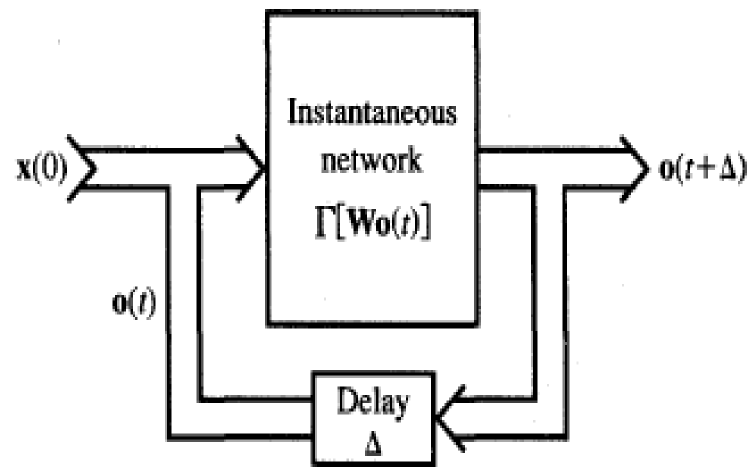


Can be used to solve complicated problems

Feedback network

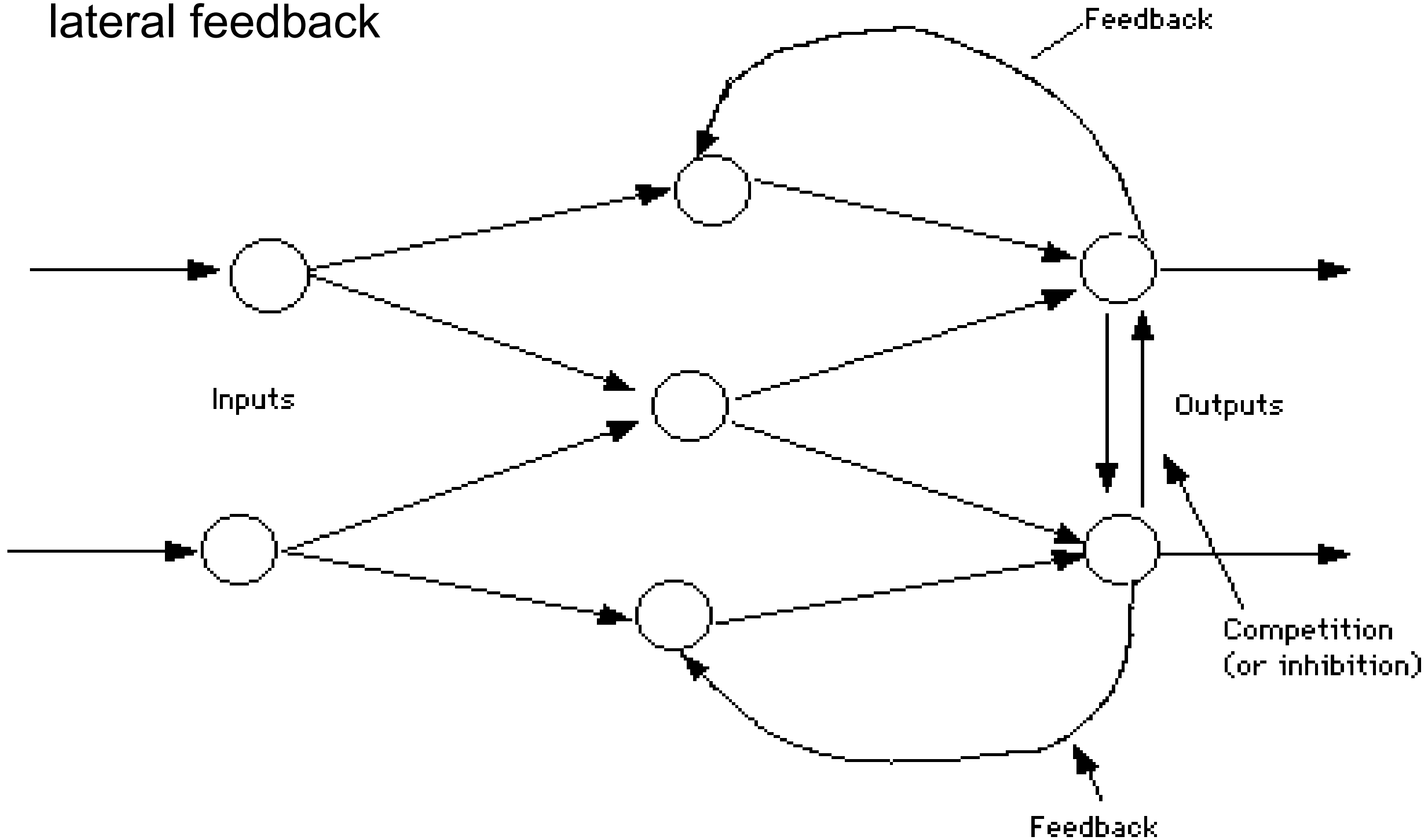


When outputs are directed back as inputs to same or preceding layer nodes it results in the formation of feedback networks



Lateral feedback

If the feedback of the output of the processing elements is directed back as input to the processing elements in the same layer then it is called lateral feedback



Recurrent n/ws

Feedback networks with closed loop are called **Recurrent Networks**. The response at the k+1'th instant depends on the entire history of the network starting at k=0.

Automaton: A system with discrete time inputs and a discrete data representation is called an automaton

- Single node with own feedback
- Competitive nets
- Single-layer recurrent nts
- Multilayer recurrent networks

Activation Function

1. Identity Function
   f(x)=x for all x
2. Binary Step function

$$f(x) = \begin{cases} 1 \; if \; x \geq \theta \\ 0 \; if \; x < \theta \end{cases}$$

3. Bipolar Step function

$$f(x) = \begin{cases} 1 \; if \; x \geq \theta \\ -1 \; if \; x < \theta \end{cases}$$

4. Sigmoidal Functions:- Continuous functions
5. Ramp functions:-

$$f(x) = \begin{cases} 1 \; if \; x > 1 \\ x \; if \; 0 \leq x \leq 1 \\ 0 \; if \; x < 0 \end{cases}$$

Important terminologies of ANNs

- Weights
- Bias
- Threshold
- Learning rate
- Momentum factor
- Vigilance parameter
- Notations used in ANN

Weights

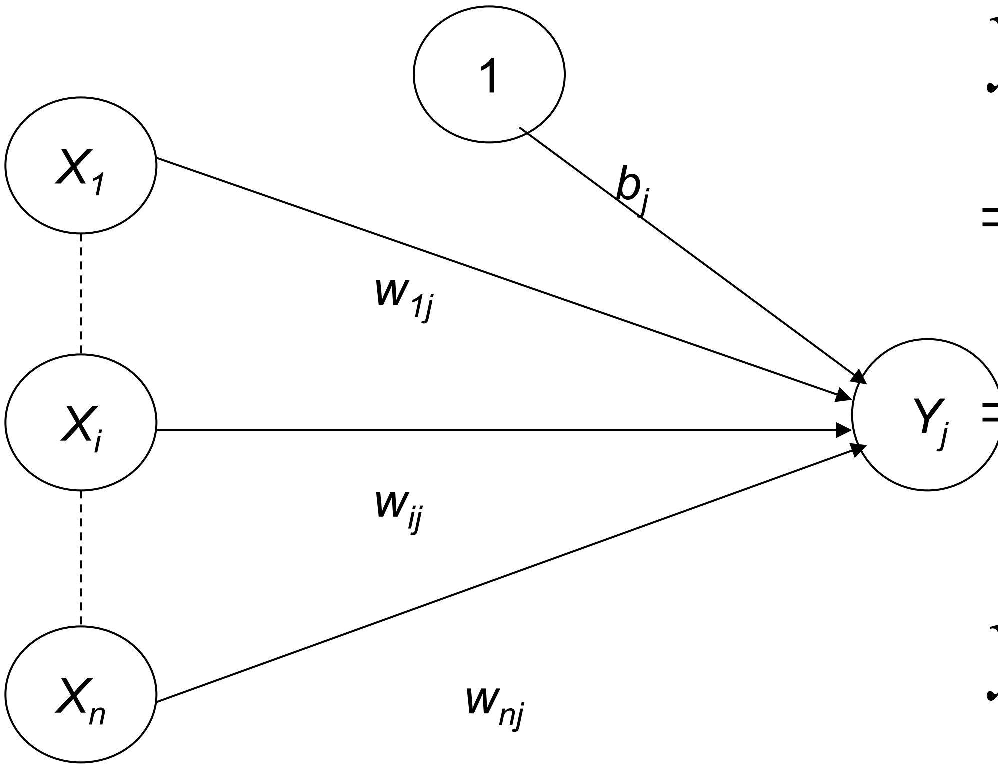
- Each neuron is connected to every other neuron by means of directed links
- Links are associated with weights
- Weights contain information about the input signal and is represented as a matrix
- Weight matrix also called connection matrix

Weight matrix

$$W = \begin{bmatrix} w_1^T \\ w_2^T \\ w_3^T \\ . \\ . \\ . \\ . \\ . \\ w_n^T \end{bmatrix} = \begin{bmatrix} w_{11} w_{12} w_{13} \cdots w_{1m} \\ w_{21} w_{22} w_{23} \cdots w_{2m} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ w_{n1} w_{n2} w_{n3} \cdots w_{nm} \end{bmatrix}$$

Weights contd…

- w_{ij} _is the weight from processing element *"i"* (source node) to processing element *"j"* (destination node)



$$y_{inj} = \sum_{i=0}^{n} x_i w_{ij}$$

$$= x_0 w_{0j} + x_1 w_{1j} + x_2 w_{2j} + ... + x_n w_{nj}$$

$$= w_{0j} + \sum_{i=1}^{n} x_i w_{ij}$$

$$y_{inj} = b_j + \sum_{i=1}^{n} x_i w_{ij}$$

Activation Functions

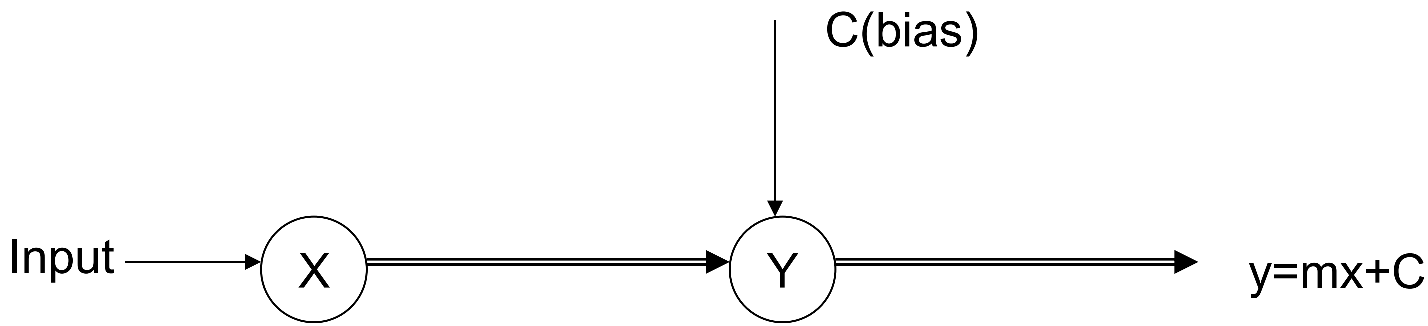
- Used to calculate the output response of a neuron.
- Sum of the weighted input signal is applied with an activation to obtain the response.
- Activation functions can be linear or non linear
- Already dealt
  - Identity function
  - Single/binary step function
  - Discrete/continuous sigmoidal function.

Bias

- Bias is like another weight. Its included by adding a component $x_0=1$ to the input vector X.
- $X=(1,X_1,X_2\ldots X_i,\ldots X_n)$
- Bias is of two types
  - Positive bias: increase the net input
  - Negative bias: decrease the net input

Why Bias is required?

- The relationship between input and output given by the equation of straight line y=mx+c



Threshold

- Set value based upon which the final output of the network may be calculated
- Used in activation function
- The activation function using threshold can be defined as

$$f(net) = \begin{cases} 1 if net \geq \theta \\ -1 if net < \theta \end{cases}$$

Learning rate

- Denoted by α.
- Used to control the amount of weight adjustment at each step of training
- Learning rate ranging from 0 to 1 determines the rate of learning in each time step

Other terminologies

- Momentum factor:
  - used for convergence when momentum factor is added to weight updation process.
- Vigilance parameter:
  - Denoted by ρ
  - Used to control the degree of similarity required for patterns to be assigned to the same cluster