INTRODUCTION

DATE: ___/___/___

⇒ Neural Network :

→ DEFINITION OF NEURAL NETWORK :

[1] According to the DARPA NN study :
"A neural network is a system composed of many processing elements operating in parallel whose function is determined by network structure, connection strengths and the processing performed at computing elements or nodes."

[2] According to Haykin:
'A NN is massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use.

[3] According to Zurada:
'Artificial Neural systems or neural networks are physical cellular systems which can acquire, store and utilize experimental knowledge.

→ A neural network is characterized by:
(i) its pattern of connections b/w the neurons and is called its architecture.
(ii) its method ~~is~~ of determining the weights on the connections and is called its training or learning algorithm.
(iii) its activation function.

A neural network consists of a large number of simple processing elements called neurons, units, cells or nodes.

DATE: ___/___/___

→ ADVANTAGE :

(i) Inherently massively parallel.
(ii) May be fault tolerant because of parallelism
(iii) Designed to be adaptive.

→ DISADVANTAGE :-

(i) No clear rules or design guidlines for arbitrary application.
(ii) No general way to assess the internal operation of the network
(iii) Training may be difficult & sometimes impossible.
(iv) ~~Diff~~ Generalization.

→ AREAS of Applications

People:
(i) Cognitive scientists (High level brain function)
(ii) Neuro-physiologists (memory, sensory etc.)
(iii) Physicists
(iv) Biologists

Application:
- Identification and control (vehicle control, process control)
- game playing and decision making (chess, backgammon)
- pattern Recognition (radar system, face identification, object recognition
- Sequence recognition (gesture, speech, handwritten
- Medical diagnosis
- financial applications,
- data mining
- etc.

HUMAN BRAIN.

⇒ Nervous System:

The human nervous system can be broken down into three stages

Stimulus → [Receptors] ⇄ [Neural Network/ Brain] ⇄ [Effectors] → Response

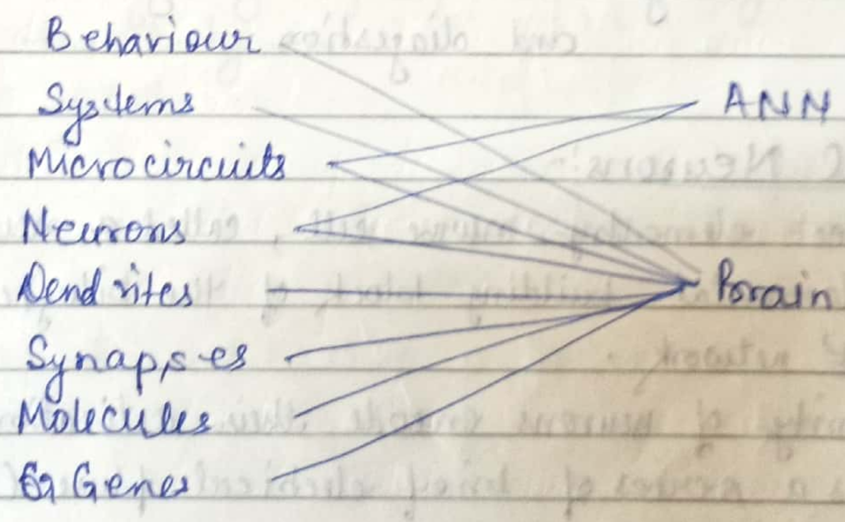
The receptors convert stimuli from the external environment into electrical impulses that convey information to the neural net (brain) – eg. photons on the retina.

The effectors convert electrical impulses ~~into~~ generated by the neural net into responses as system outputs – eg: activate muscles.

The neural net (brain) continually receives information, perceives it and makes appropriate decisions.

The flow of information is represented by arrows – feedback and feedforward.

⇒ Levels of Organization in the Brain



DATE: ___/___/___

⇒ Structure of Human Brain:-

Anatomically the brain can be subdivided into three major parts: the cerebrum, the cerebellum and brain stem.

→ Cerebrum:- It is divided in four different lobes.

Frontal lobe (self control, planning, reasoning etc.)

Parietal lobe (for body sensation)

Temporal lobe (for hearing, learning & long-term memory storage)

Occipital lobe (for reception & interpretation of vision)

→ Cerebellum: it provides precise coordination and control of the body.

→ Brain stem:- The pons & medulla oblongata constitute the brain stem.

Pons: (for conveying information about movement from the cerebrum to the cerebellum.

Medulla oblongata: Controlls breating, heart rate, and diagestion.

⇒ Bilogical Neurons:-

The ~~stect~~ elementry nerve cell, called a neuron, is the fundamental building block of the bilogical neural ~~net~~ network.

→ The majority of neurons encode their activations or outputs as a series of brief electrical pluse (ie. spikes or action potentials).

→ The neurons cell body (soma) processes the incoming activations and converts them into output activation.

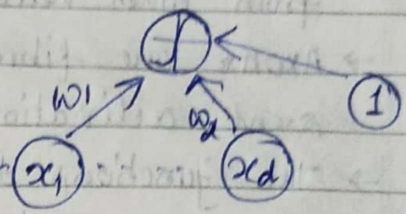
ARTIFICIAL NEURON

DATE: / /

Artificial neuron is a compilational unit which will make a particular computation based on other unit is connected.

Neuron pre-activation or i/p activation $a(x) = b + \sum_i w_i x_i$

$= b + w^T x$ → x → vector which contain scalar x_1 upto x_d



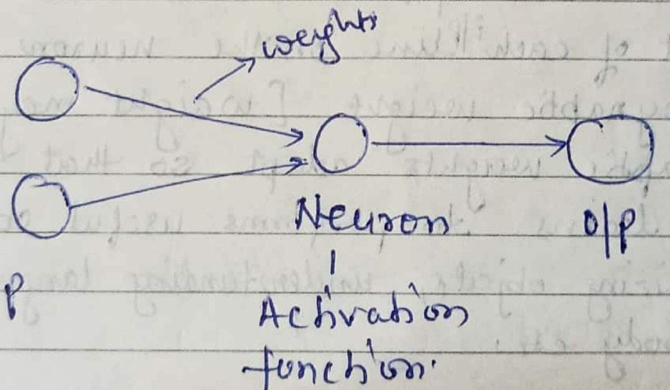
Neuron (output) activation

$h(x) = g(a(x)) = g(b + \sum_i w_i x_i) =$

w → are connection weight

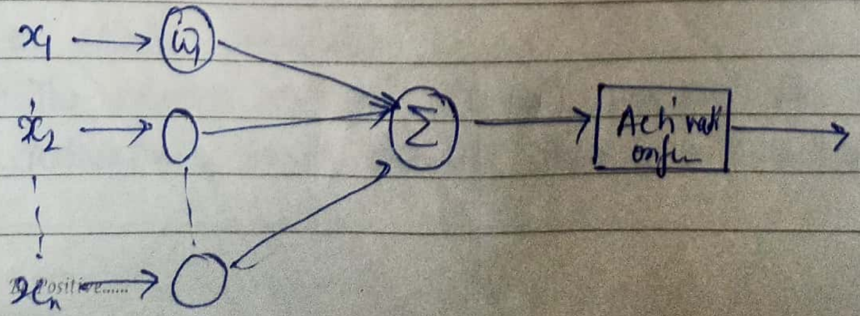
b → neuron bias

$g(\cdot)$ → activation function.



ACTIVATION FUNCTION

Activation function of a node defines the outputs of that node given input(s)

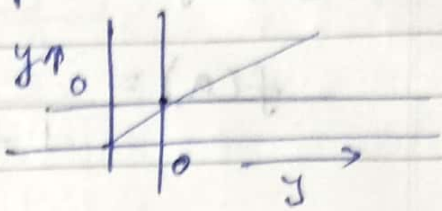


DATE: ___/___/___

(i) Linear function: these are simple but computationally limited

purelin

$f(a) = y = b + \sum_i w_i x_i$



(ii) ~~Identity function~~

$f(a) = a \quad \forall x$

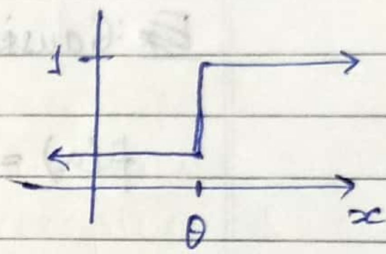
(iii) Binary step function (with threshold θ)

A threshold (hard-limiter) activation function is either a binary type or a bipolar type:

~~Sign~~

Hard limit / hardlim

$$y = f(a) = \begin{cases} 1 & \text{if } a \geq \theta \\ 0 & \text{if } a < \theta \end{cases}$$



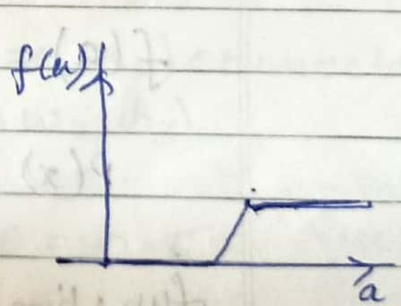
Bipolar is (1 or −1) or signum function: or sign(x)

Symmetrical Hardlim / hardlims

$$f(a) = \begin{cases} +1 & a \geq \theta \\ -1 & a < 0 \end{cases}$$

(iv) Piecewise Linear function

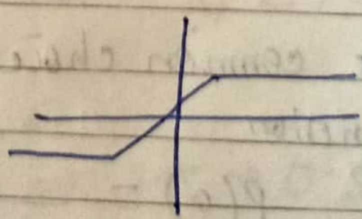
$$f(a) = \begin{cases} +1 & \text{if } a \geq 0.5 \\ a + 0.5 & \text{if } -0.5 \leq a \leq 0.5 \\ 0 & \text{if } a \leq 0.5 \end{cases}$$



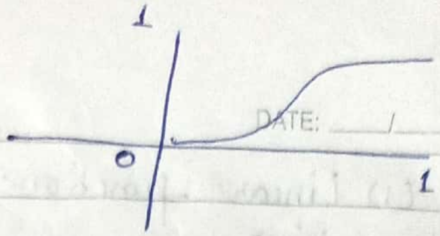
or

Bipolar

$$f(a) = \begin{cases} -1 \\ a \\ +1 \end{cases}$$

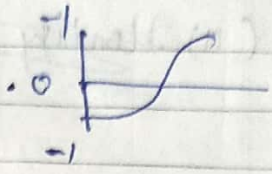


Sigmoid function :-



$$f(a) = \frac{1}{1+e^{-a}}, \quad 0 \le f(a) \le 1$$

Hyperbolic tangent (tanh) function



$$f(a) = \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}} = \frac{e^{2a} - 1}{e^{2a} + 1} \quad -1 \le f(a) \le 1$$

Gaussian function -

$$f(a) = \exp\left(\frac{(a_i - c_i)^2}{2\sigma_i^2}\right)$$

$c_i \rightarrow$ center

$\sigma_i \rightarrow$ Spread factor

~~Stocastic~~

Stochastic function :-

$$f(a) = \begin{cases} +1 & \text{with probability } P(a) \\ -1 & \text{with " } 1 - P(a) \end{cases}$$

$P(x) \rightarrow$ neuron ~~activ~~

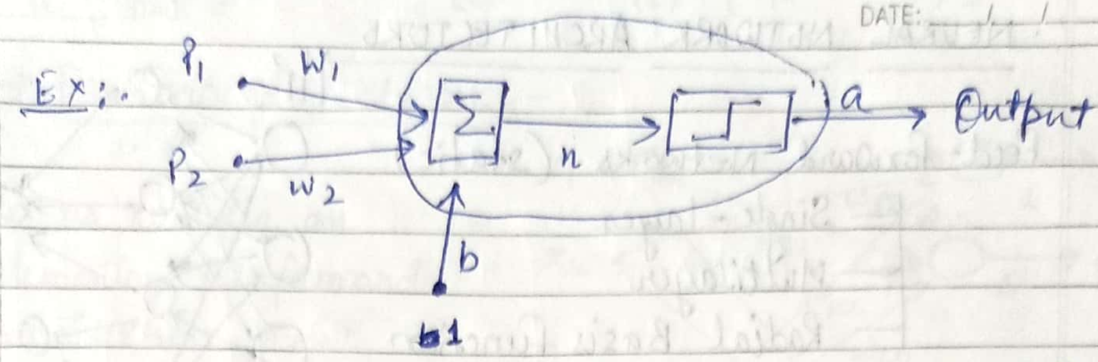
function $P(a) = \begin{cases} 0 & \text{as } a \rightarrow -\infty \\ 1 & \text{as } a \rightarrow +\infty \end{cases}$

The common choice for $P(a)$ is sigmoid-shaped function

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

$T \rightarrow$ pseudotemprature that control the noise level.

Ex:-



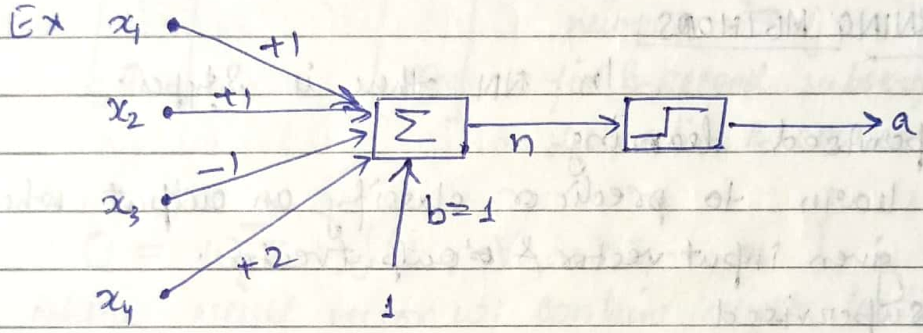
$p_1 = 2.5$; $p_2 = 3$; $w_1 = 0.5$; $w_2 = -0.7$, $b = 0.3$

$$a = w^T x + b = b + \sum_i w_i x_i$$

$$= 0.3 + (w_1 x_1 + w_2 x_2) = 0.3 + (0.5 \times 2.5 + (-0.7 \times 3))$$

$$= ~~0.37~~ \; -0.55$$

Ex



$$a = b + w^T x = b + \sum w_i x_i = b \cdot 1 + \sum_{i=1} w_i x_i$$

Ex → Suppose we have 3-D i/p $x = (x_1 x_2 x_3)$ connected to neuron with weights $w = (w_1 w_2 w_3)$

$x_1 = 2$ $w_1 = 1$

$x_2 = -1$ $w_2 = -0.5$ & bias = 0.5

$x_3 = 1$ $w_3 = 0$

logistic sigmoid = 0.95, linear = 3, biT = 1, Rectified Lin = 3

Ex → target of for x is $t = 1$

$y = w^T x + b \Rightarrow 2.5 + 0.5 = ③$

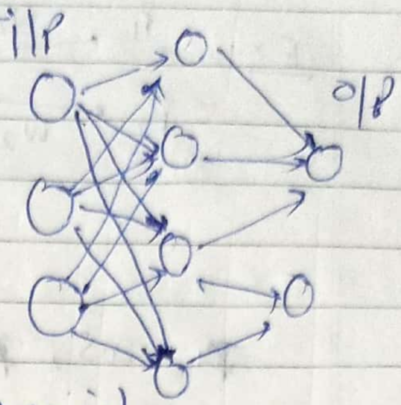
$[2 \; -1 \; 1] \begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix}$ ~~[2 -1 1]~~ $= [2 + 0.5 + 0] = 2.5$

$\frac{1}{2}(2.5 - 1)^2 = \frac{1}{2} \times 1.5 \times 1.5 = 1.125$

2.t

NEURAL NETWORK ARCHITECTURE

Feed-forward Networks (static)
- Single-layer
- Multilayer
- Radial Basis function



Recurrent/feedback Networks (Dynamic)
- Competitive Networks
- Kohonen's SOM
- Hopfield Network
- ART Models

LEARNING METHODS:

In NN there is 3 types.

Supervised learning
learn to predict or classify an output when an given input vector & output vector.

Unsupervised
Discover good internal representation of input

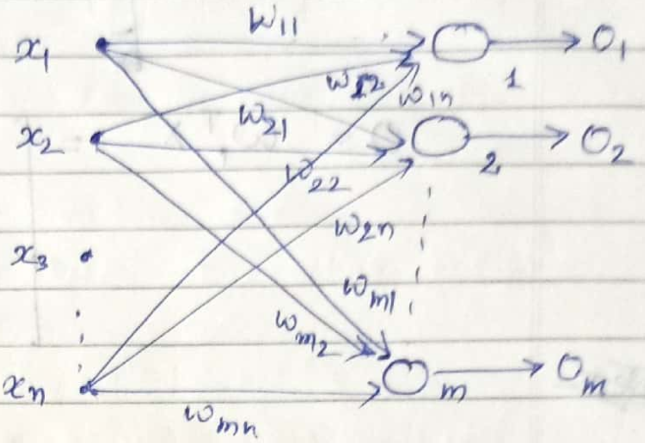
Reinforcement

	Supervised	Unsupervised
Feed forward	Perceptron, Adaline, Madaline, backpropagation, Cauchy M/c, etc.	Learning Matrix, Linear Associative Memory, Fuzzy Associative Memory etc.
Feed Back	Boltzmann M/C, Fuzzy Cognitive map, Learning Vector Quantization etc.	Additive Grossberg, Binary Adaptive Resonance Theory, SOM, etc.

$$\begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix} \begin{bmatrix} i/p1 \\ i/p2 \end{bmatrix} = \begin{bmatrix} (i/p1 * w_{11}) + (i/p2 * w_{21}) \\ (i/p1 * w_{12}) + (i/p2 * w_{22}) \end{bmatrix}$$

DATE:

⇒ Feedforward Network :-



Let us consider an elementary feedforward architecture of m neurons receiving n inputs. Its output & input vectors are respectively : (we are using vectors as coloumn vectors)

$$O = \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_m \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix};$$

weights w_{ij} connects the i'th neuron with j'th input. first & second subscript denote the destination & source nodes, respectively.

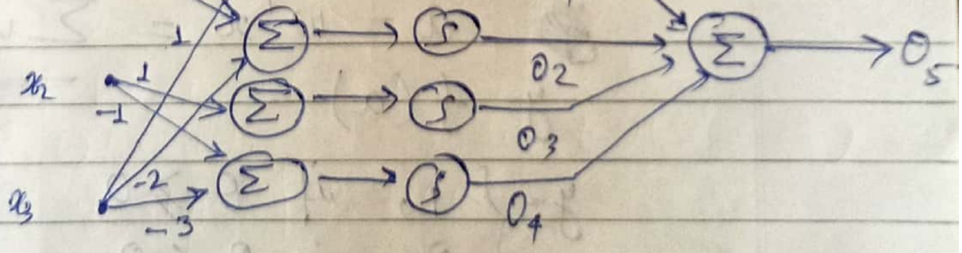
$$O = w_i^T x = f(\cdot \quad \forall\ i = 1, 2, \ldots m \}$$

where weight vector w_i contains weights leading towards the ith output node.

$$w_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{in} \end{bmatrix}; \quad w = \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{m1} \\ w_{12} & w_{22} & \cdots & w_{m2} \\ \vdots & \vdots & & \vdots \\ w_{1n} & w_{2n} & \cdots & w_{mn} \end{bmatrix}$$

Ex →

$$O = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ O_4 \end{bmatrix}; \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



DATE: / /

for first layer. $W_1 = \begin{bmatrix} +1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & -2 & 0 & -3 \end{bmatrix}$

$$O = W_1^T X = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ O_4 \end{bmatrix}$$

Note:-

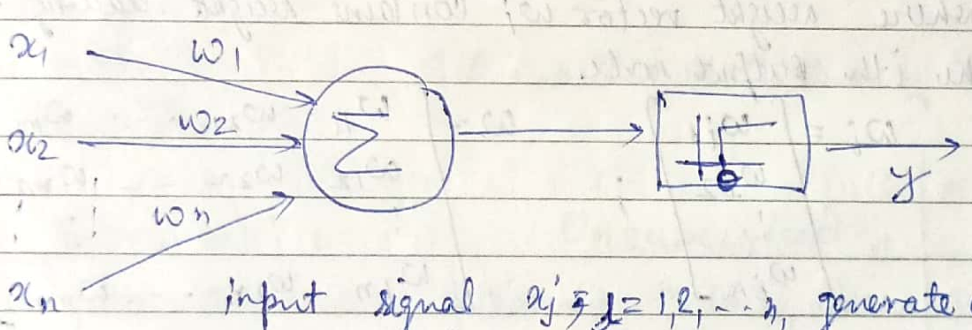
⇒ Mc.Culloch & Pitt Model: (1943)

This is first ANN model.

→ Activation of a McCulloch - Pitts neuron is binary.

→ A connection path is excitatory if the weight on the path is positive; otherwise it is inhibitory.

→ net input to the neuron is greater than threshold the neuron ~~fire~~ fires.



input signal x_j; $j = 1, 2, \ldots, n$, generate output of 1 if this sum is above a certain threshold θ, Otherwise output of 0

$$a = w^T x = \sum w_i x_i - \theta$$

$$y = f(a)$$

$$y = \begin{cases} 1, & \text{if } \theta \geq 0 \\ 0 & \text{if } \theta < 0 \end{cases}$$

Be Positive......