VICTORY
PAGE NO.:
DATE: 9 2 23

Artificial Nueral Network - CSA4002

→ History

→ Rossemblat proposed the perceptron in 1960s

→ Werbos came up with multilayer perceptron in 1974 by the back-propogation algorithm for a three layered perceptron network.

→ 1982 - self organising map was introduced by Kohonen

→ Adaptive Resonance were termed 1987 by Stephen Grossekeberg and Gail Carpenter in 1987, which were ART - unsupervised learning technique which led to the discovery of the modified back-propating perceptron (1988-1995). Since, then various

→ Nueral Network -

→ ANN or an Artificial Nueral network is a colletively distributed network of parallely processing units with a natural prospensity (behaviour) of storing information and making available to use later.

→ Each processing unit of a ANN is a artificial nueron.

→ It resembles human brains in 2 ways -

⇒ aquiring knowledge through learning

⇒ storing information and interconnection between nuerons. This known as Synaptic Weights to store the aquired weights. (connections)

→ Knowledge is acquired by the network from its enviornment through a learning process.

→ a ANN contains a input layers of nuerons (nodes), one or two hidden layers of nuerons and a final layer of output nuerons.

→ Inspired by the functionality of human brain, to demonstrate intelligence via ANN used in various tasks such as Pattern-matching, approximation, language translation and pattern recognition, vector quantization

Computer Networks	Neural Networks
→ One CPUs	→ Multiple CPU's
→ Fast processing units	→ Slow processing units
→ Reliable units	→ Unreliable units
→ Static Infrastructure	→ Dynamic networks

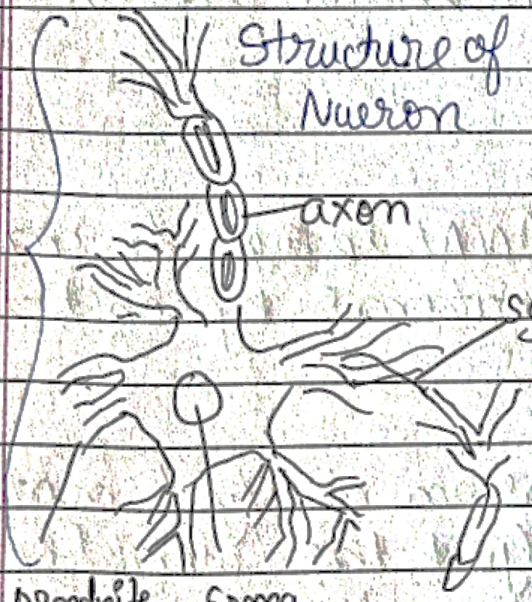
Difference between Biological Nueron and Artificial Nueron -

→ Measures	Biological Nueron (Brain)	Artificial Nueron (ANN)
	Axon	Output
	Dendrities	Input
	Cell - soma	Processing unit
	Synapse	Interconnections
Speed	Few ms.	Few nano secs.
Size & Complexity	10^{11} nuerons & 10^{15} connections	→ 1 Billion bytes of RAM, trillion of bytes of disk space (depends upon use)
Storage Capacity	Stores memory in Synapse inter-connection, so data is never lost	Stores data in contiguous memory; Data loss may occur
Tolerance	No fault tolerance	No fault tolerance if disrupted
Energy Consumption	6~10 Joules per Operation	10~16 Joules / operation
	Learning Capability	Programming Capability
Control Mechanism	Complicated involves chemicals in biological nuerons	simpler in ANN

VICTORY
PAGE NO.:
DATE:

Human Brain

→ The Brain the CPU of our body and contains huge networks of nuerons. A brain has 10 billion nuerons, each nueron has some basic elements that process the information



Structure of Nueron

Dendrite - recieves the signals from other nuerons

Synapse - point of interconnection bw. 2 nuerons and save the info.

Axon - transmits signals output of this nueron

Soma - Processor

→ Signals that electro-chemical travel through the dendrite network get propogated through the cell body.

→ A nueron requires a threshold to response in a short period of time. A response recieved is evaluated by distributing the signal throughout the network By which the senses act as input feed and certain stimulus occurs. Each nueron performs certain tasks being divided in a parallel network

→ The strength of the signal depends upon the nuero-transmitter available at that moment as the response.

VICTORY
PAGE NO.:
DATE:

Properties of ANN-

① Non-Linearity-

a) Most problems have difficulty to be solved; but ANN solves the real-world non-linear problems by learning non-linear decision making functions.

② Input-Output Mapping

→ Just like the brain, the ANN learns to map the possible input-output outcomes by constructing them at hand

③ Adaptibility - The ANN has an ability to not only learn but to re-adjust the weight of its synapse according to the changes in the environment, meaning the ANN can be re-trained for minor changes using enviorment conditions

④ Evidential Response - ANN can not only predict patterns and recognize them but also create a confidence about its outcomes and decisions

⑤ Fault tolerance - The ANN has a great inherity fault tolerant hardware implementation. Thus it provides robust output in danger/ faliure situations as desired.

⑥ Uniformity of Analysis and Design

⑦ VLSI compatibility - The ANN has a great massive collection of parallel processing nuerons at a high speeds which makes them suitable for the implementation of large scale integrated technology for certain tasks.

⑧ Contextual Information - The information that deals with information naturally.

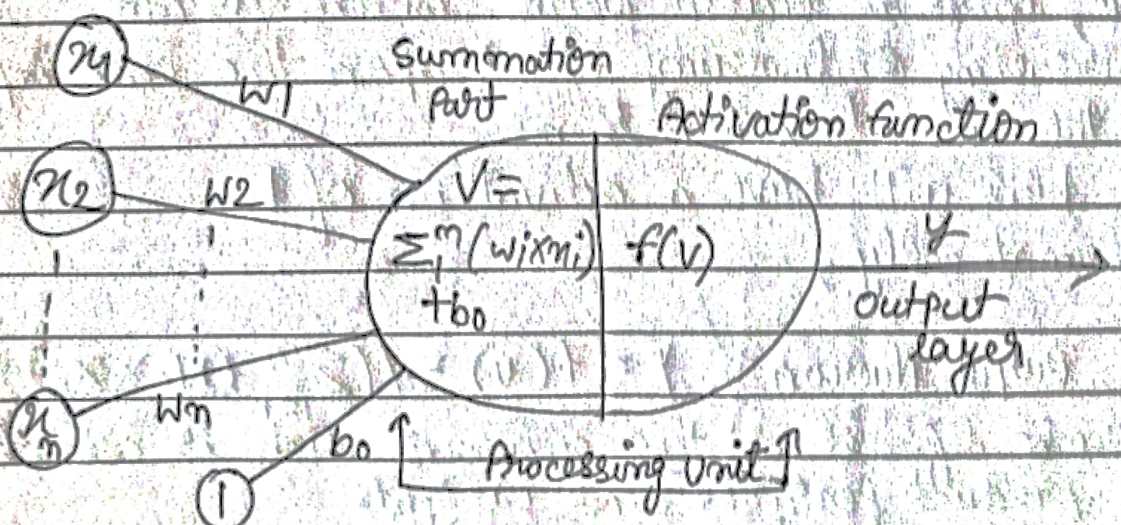
⑨ Nuero-Biology - The ANN analogy helped to iterate that the fault tolerant parallel processing can not just be possible but can be more powerful and fast.

Models of Artificial Nuerons

→ The models of ANN are specified by following entities

1) Model's Synaptic Interconnection
2) Synaptic Interconnection weights
3) Summation
4) Activation Function
5) Training / learning rules for updating and adjusting the connection weights.

System Architecture of Single Nuron



→ Information aquired by nueron may certainly come from a single or multiple source
→ Each information synapse with nueron is associated with a weight.
→ The processing unit of nueron has 2 parts majorly →

1) Information Accumulation part
2) Activation function

Input layer

1) → During the information accumulation the multiplication of input to the corresponding weights are formed.
2) Activation function → Here n is the size of input
A nueron applies an activation function on the applied upon the summation
→ It compares the input value to the threshold value, so if it's greater then the nueron is activated. It's disabled if the input value is not enough

$$V = W_1 x x_1 + W_2 x x_2 + \cdots + W_n x x_n + b_0 \quad \text{—①}$$

$$\boxed{V = \sum_{i=1}^{n} (W_i x x_i) + b_0} \rightarrow \text{Activation function}$$

Here, b_0 is a biased weight, which has an increasing or lowering effect upon the net input of the activation function.
→ The bias will have a activation of 1
Mathematically,

$$\text{Output Nueron} \Rightarrow y = f(v) = f\left(\sum_{i=1}^{n} (W_i x x_i) + b_0\right)$$

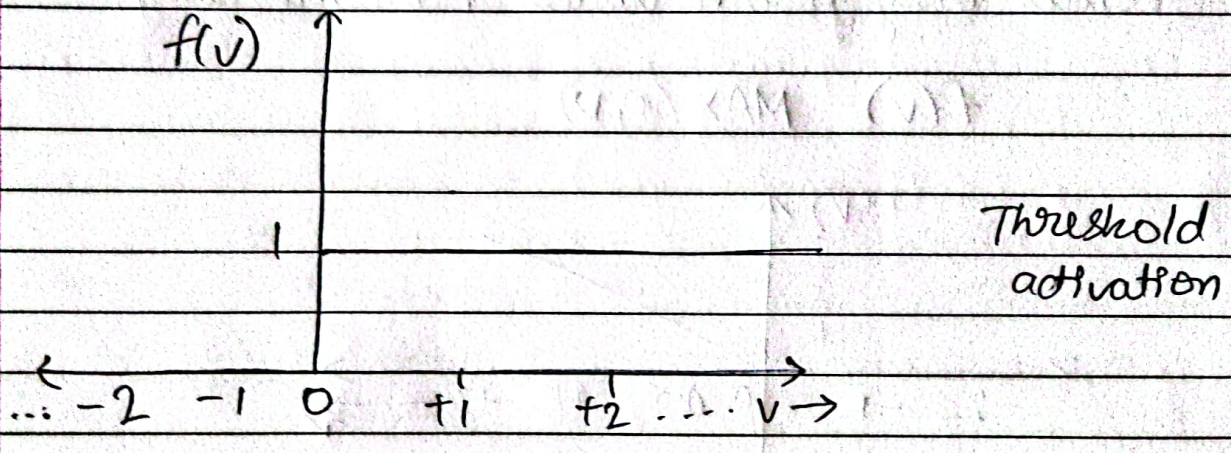
→ It can be defined as the constant which is added to the product features and weights

Types of Activation functions-

① Threshold function

→ This activation function will result to 1 if the input summation to the nueron is greater than 0 or equal otherwise it will result 0.

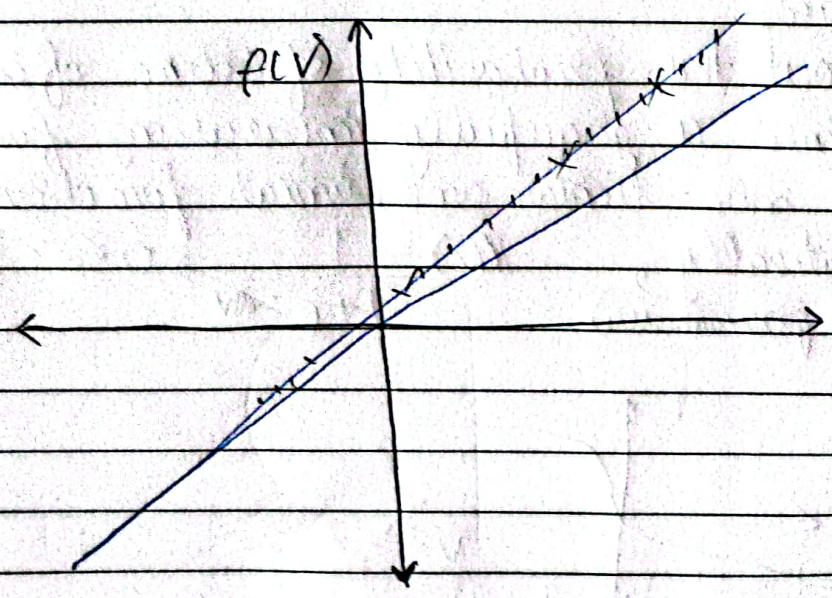
$$f(v) = \begin{cases} 0 & \text{if } v < 0 \\ 1 & \text{if } v \geq 0 \end{cases}$$



② Linear function-

It will provide the same output as the summation of the input to nueron.

$$f(v) = \{v, \forall v$$



3) Piecewise Linear Function

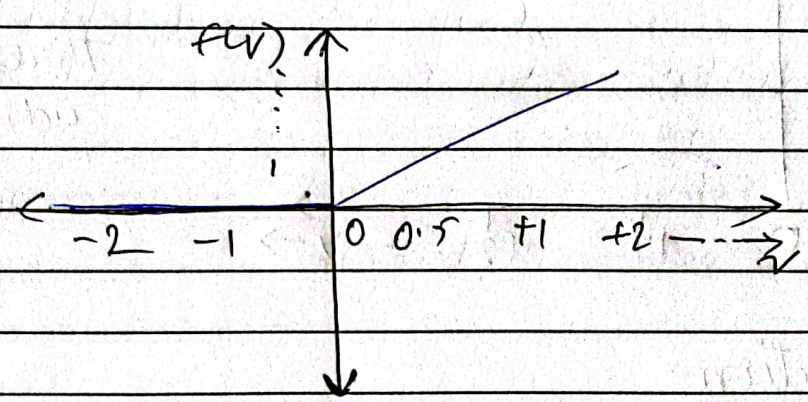
→ Its the variant of linear activation function which can be represented as-

$$f(V) = \begin{cases} 1, & V \geq 0.5 \\ V, & -0.5 < V < 0.5 \\ 0, & V \leq 0.5 \end{cases}$$

4) Rectified Linear Unit (ReLU)

→ It is a variant of linear activation function which results only positive values. Mathematically,

$$f(V) = MAX(0, V)$$



5) Sigmoid Function

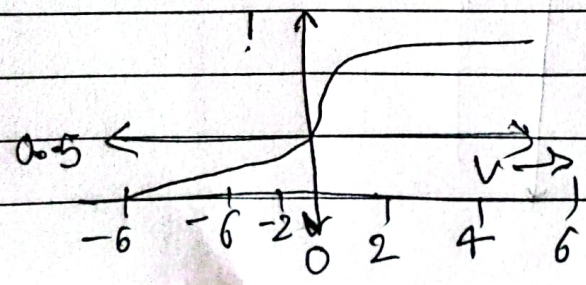
→ It is a logistic function where a S-shaped graph is obtained.

→ It maps out the probability of score of a neuron

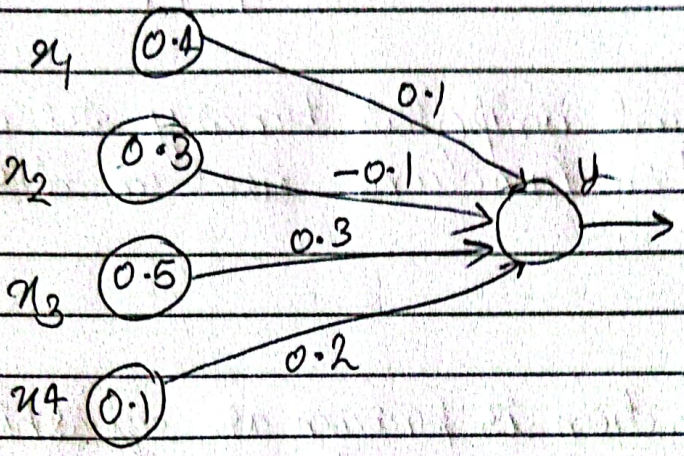
→ It is defined a gracefully increasing function about the non-linear and linear functions.

Mathematically, $f(V) = \frac{1}{1+e^{-\alpha V}}$, $0 < \alpha \leq 1$

(α is slope parameter)



Ex-) Find output of neural architecture, considering the activation function as linear, ReLU and sigmoid. Consider the slope parameter for sigmoid as 1.



⇒ Summation function

⇒ $V = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4$

$= 0.4 \times 0.1 + 0.3 \times (-0.1) + 0.5 \times 0.3 + 0.1 \times 0.2$

⇒ $V = 0.18$, output of the neuron $y = f(V)$

for different activation functions,

a) Linear ⇒ $y = f(V) = 0.18$

b) ReLU ⇒ $y = f(V) = MAX(0, V) = MAX(0, 0.18) = 0.18$

c) SIGMOID

⇒ $y = f(V) = \frac{1}{1 + e^{-\alpha V}}$

$= \frac{1}{1 + e^{-0.18}}$

$= 0.54$

Bipolar Sigmoid

$$f(x) = \frac{-1 + 2}{1 + e^{-x}}$$

Binary Sigmoid

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

⇒ Bipolar binary and unipolar binary are called limiting activation functions used in discrete nueron model

Squashing Function

The sigmoid function is known as squashing function, as it maps the whole real range of z into [0,1] in the g(z). It can be used as conditional probability distribution model as sigmoid function proposes the probability of the nueron source and probability always ranges from 0 to 1.