

$$y_{in}(2) = 1 = t = -1 \quad \{ -1 y_{in}(2) \}$$

$$1. -1 + 1 = 0$$

$$y_{in}(0) = 1 = t = 1$$

1) write the purposes of weights in the artificial neural network models.

A) weights play a crucial role in determining the network's ability to learn and make accurate predictions.

i) control the strength of connections.

ii) facilitate learning.

iii) enable generalization.

iv) control Activation of neurons.

v) help with Nonlinearity.

vi) contribute to model flexibility.

2) illustrate the uses of different activation functions for artificial neural networks.

1) sigmoid \rightarrow Binary classification output layer.

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

- smooth output probabilistic output
- vanishing gradient near zero centered

2) tanh \rightarrow Hidden layers for data between -1 and 1

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

- zero-centered better gradients.
- vanishing gradient for large values.

3) ReLU \rightarrow Hidden layers for deep networks.

$\text{ReLU}(x) = \max(0, x)$

- Efficient; simple sparse activations.
- Dying ReLU problem (inactive neurons).

4) Leaky ReLU \rightarrow Deep networks to prevent dying ReLU

$\text{Leaky ReLU}(x) = \max(ax, x)$

- mitigates dying ReLU problem.
- slower convergence in some cases.

5) softmax \rightarrow multi-class classification output.

$$\text{softmax}(Z_i) = \frac{e^{Z_i}}{\sum_j e^{Z_j}}$$

- converts to probability distribution
- computationally expensive.

3) write the algorithm used in a single layer perceptron model for learning.
 The LP model is a type of artificial neural network used for binary classification tasks.

$$\hat{y}_{in} = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b$$

$$w_{i, \text{new}} = w_{i, \text{old}} + \alpha t \cdot o_i$$

$$w_{j, \text{new}} = w_{j, \text{old}} + \alpha t \cdot x_j$$

$$b_{\text{new}} = b_{\text{old}} + \alpha t$$

$$\begin{cases} 1 & y_{in} \geq 0 \\ -1 & y_{in} < 0 \end{cases}$$

4) Distinguish the following.

1) Threshold:

- def: A value used to determine when a neuron activates.
- role: determine the decision boundary for neuron activation.
- usage: used in step functions to decide if the neuron should activate.
- input: directly impacts when a neuron fires, making the model binary in simpler cases.

2) Learning rate: A hyperparameter that controls the step size in weight ^{updates}.
 role: control how much the weights are updated during training.
 usage: used during training to adjust weights according to the current gradient.
 input: Impact the speed and stability of training.

3) Activation: A function applied to the weighted sum of inputs to determine neuron output.

- role: introduces non-linearity, allowing the network to learn complex functions.
- input: the ability of the model to learn complex patterns & representations.