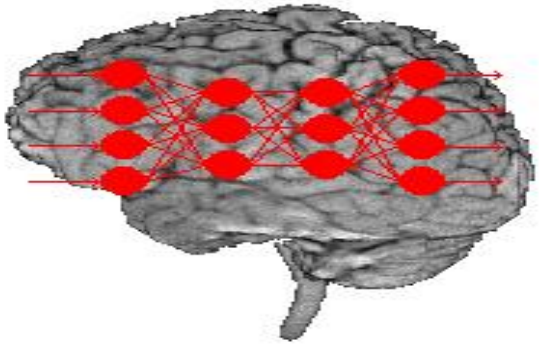
Deep Learning
Lecture-6

Artificial Neural Network



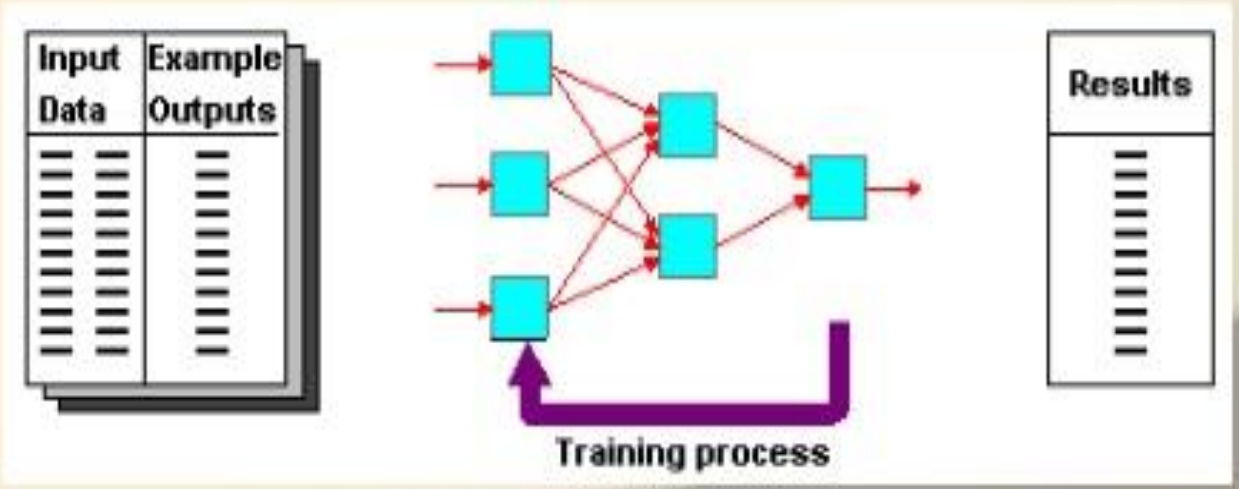
Asst. Lect. Ali Al-khawaja

2025-2026

Class Room

Definition of Supervised Learning Network

Supervised learning is used when we have a set of training data.This training data consists of some input data that is connected with some correct output values. The output values are often referred to as target values. This training data is used by learning algorithms like back propagation or genetic algorithms.



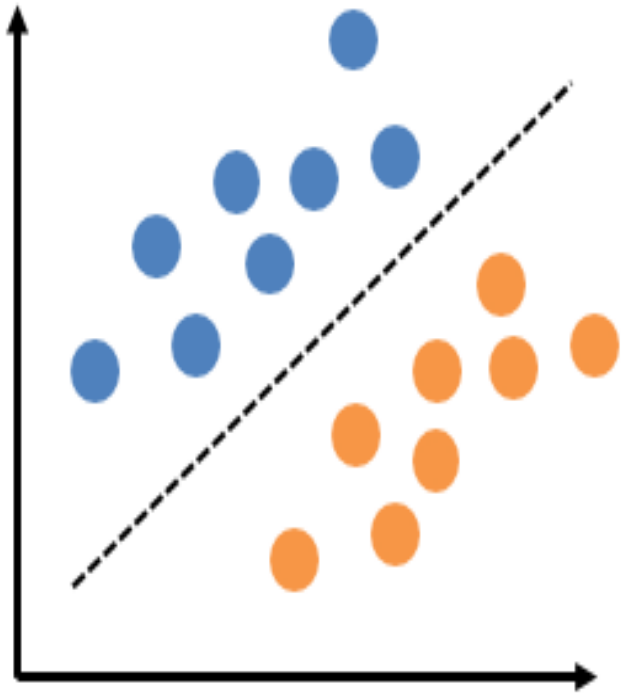
What can a Perceptron do ?

- In machine learning, the **perceptron** is an algorithm for supervised classification of an input into one of several possible non-binary outputs.
- Perceptron can be defined as a single artificial neuron that computes its weighted input with the help of the threshold activation function or step function.
- The Perceptron is used for binary Classification.
- The Perceptron can only model linearly separable classes.
- First train a perceptron for a classification task.
  - Find suitable weights in such a way that the training examples are correctly classified.
  - Geometrically try to find a hyper-plane that separates the examples of the two classes.

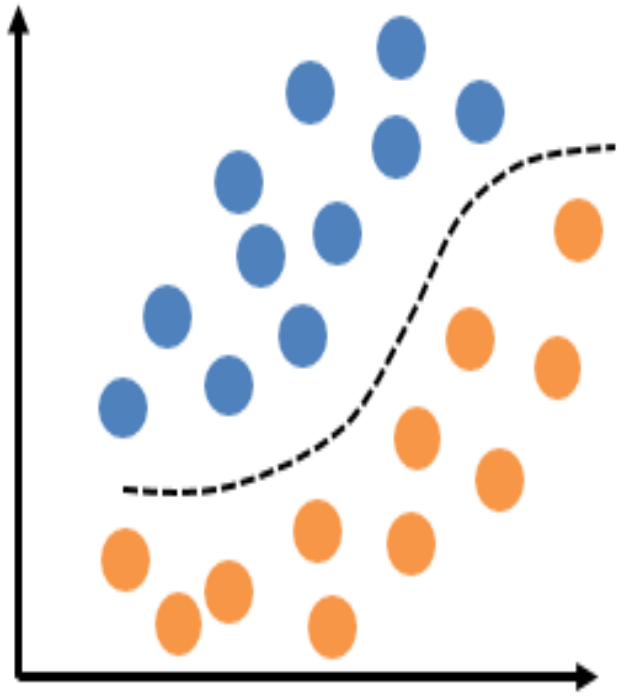
Linear Separability:

- Linear separability is the concept wherein the separation of the input space into regions is based on whether the network response is positive or negative.
- When the two classes are not linearly separable, it may be desirable to obtain a linear separator that minimizes the mean squared error.
- Definition : Sets of points in 2-D space are linearly separable if the sets can be separated by a straight line.
- Generalizing, a set of points in n-dimensional space are linearly separable if there is a hyper plane of (n-1) dimensions separates the sets.

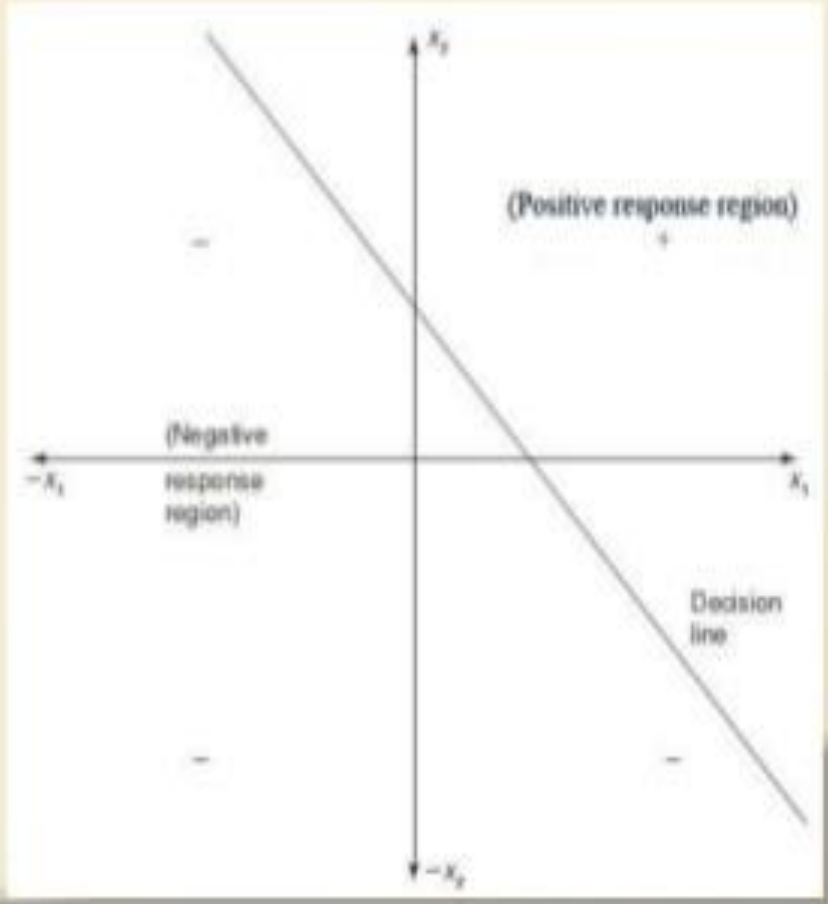
Linear



Nonlinear



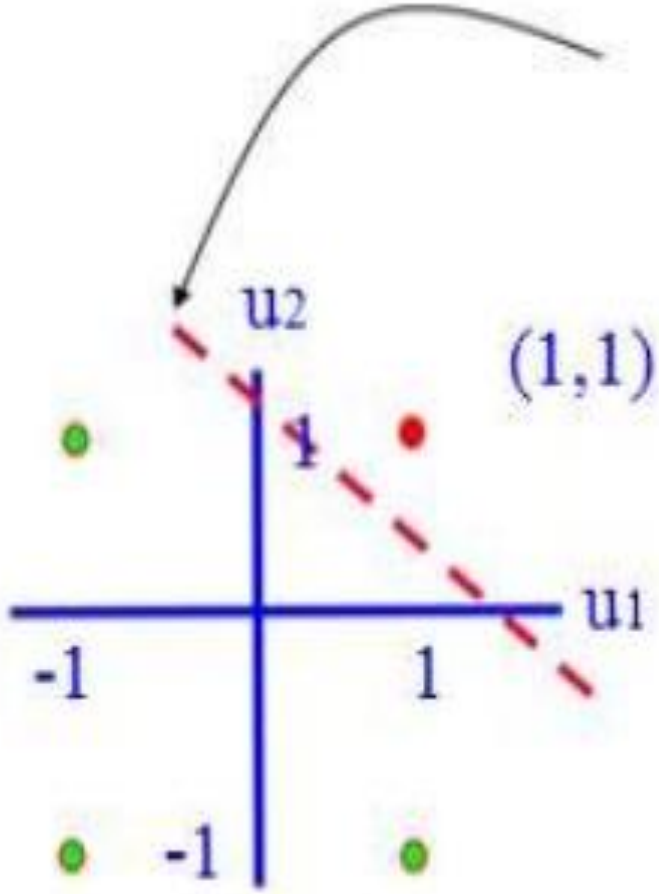
- Consider a network having positive response in the first quadrant and negative response in all other quadrants (AND function) with either binary or bipolar data, then the decision line is drawn separating the positive response region from the negative response region.

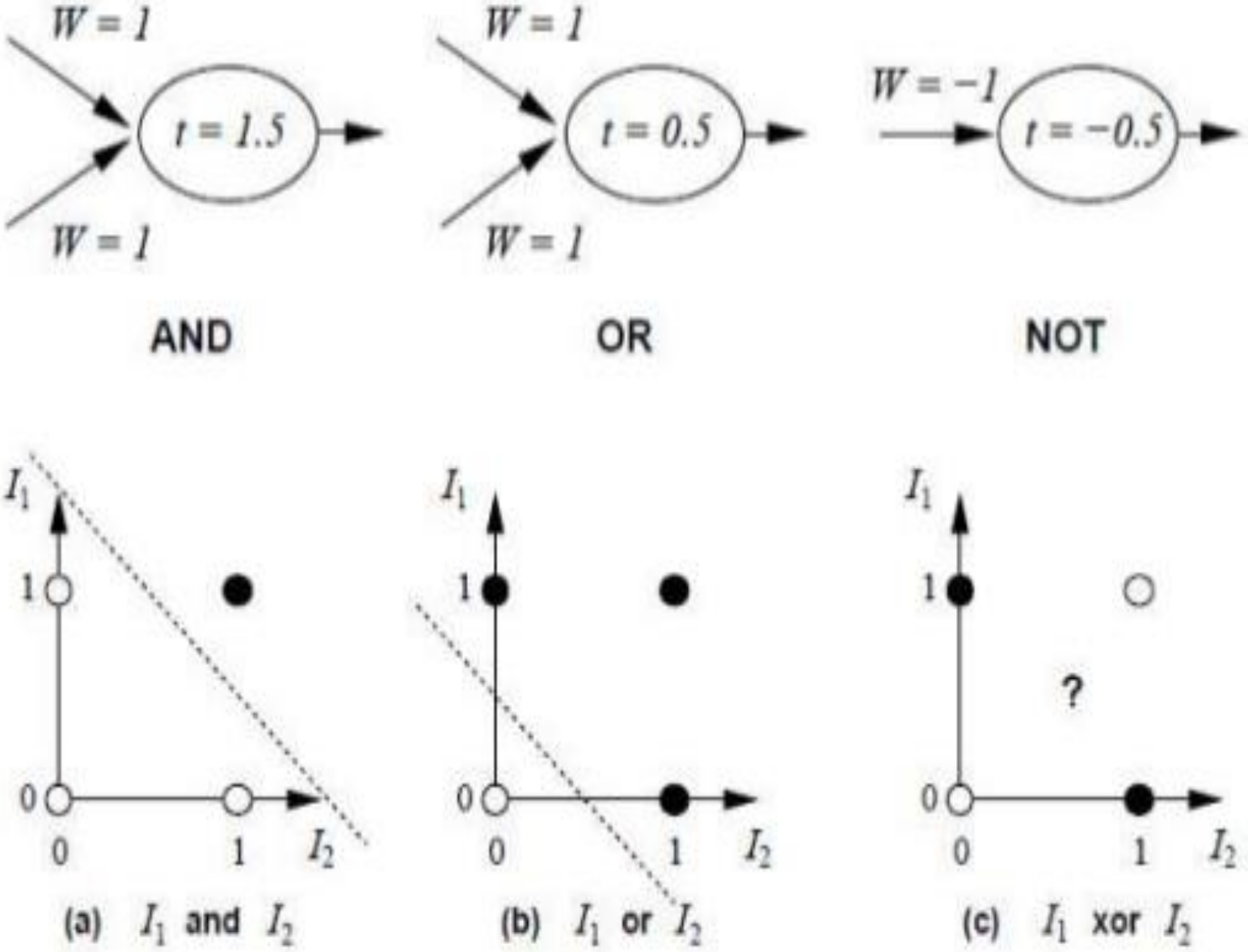


AND Gate is linearly Separable

:-

u1	u2	AND
-1	-1	-1
1	-1	-1
-1	1	-1
1	1	1




W = 1
t = 1.5
W = 1
AND
W = 1
t = 0.5
W = 1
OR
W = −1
t = −0.5
NOT
I1
1
0
0
1
I2
(a) I1 and I2
I1
1
0
0
1
I2
(b) I1 or I2
I1
1
?
0
0
1
I2
(c) I1 xor I2

- The net input to the output Neuron is:

$$Y_{in} = w_0 + \Sigma_i x_i w_i$$

Where Y_{in} = The net inputs to the ouput neurons.

i = any integer

w_0 = initial weight

- The following relation gives the boundary region of net input.

$$b + \Sigma_i x_i w_i = 0$$

- The equation can be used to determine the decision boundary between the region where $Y_{in} > 0$ and $Y_{in} < 0$.
- Depending on the number of input neurons in the network. this equation represents a line, a plane or a hyper-plane.
- If it is possible to find the weights so that all of the training input vectors for which the correct response is 1. lie on the either side of the boundary, then the problem is called linearly separable.
- Otherwise. If the above criteria is not met, the problem is called linearly non-separable.

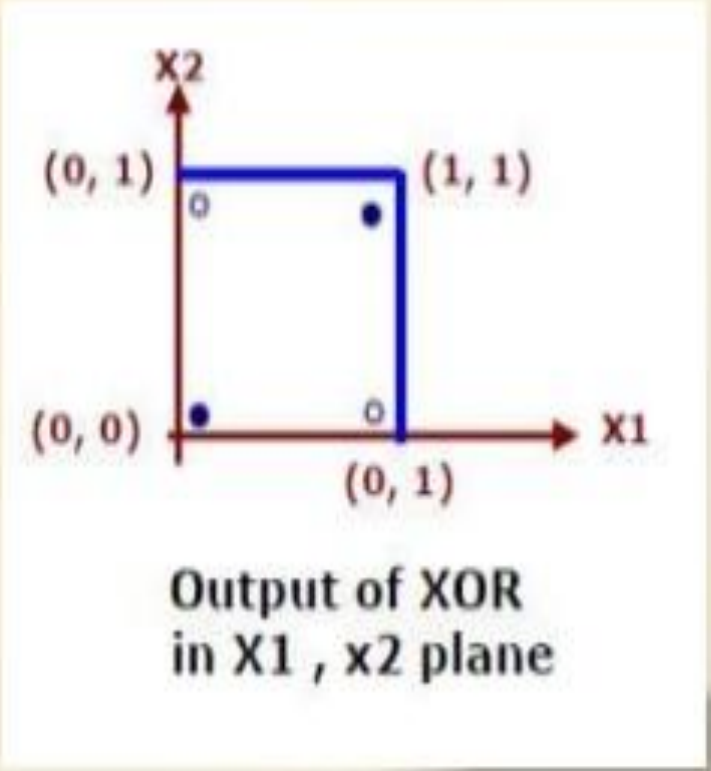
Exclusive OR (XOR) Problem :-

Input x1	Input x2	Output	
0	0	0	} Even parity •
1	1	0	
0	1	1	} Odd parity °
1	0	1	

XOR truth table

- Even parity means even number of 1 bits in the input
- Odd parity means odd number of 1 bits in the input

- There is no way to draw a single straight line so that the circles are on one side of the line and the dots on the other side.
- Perceptron is unable to find a line separating even parity input patterns from odd parity input patterns.

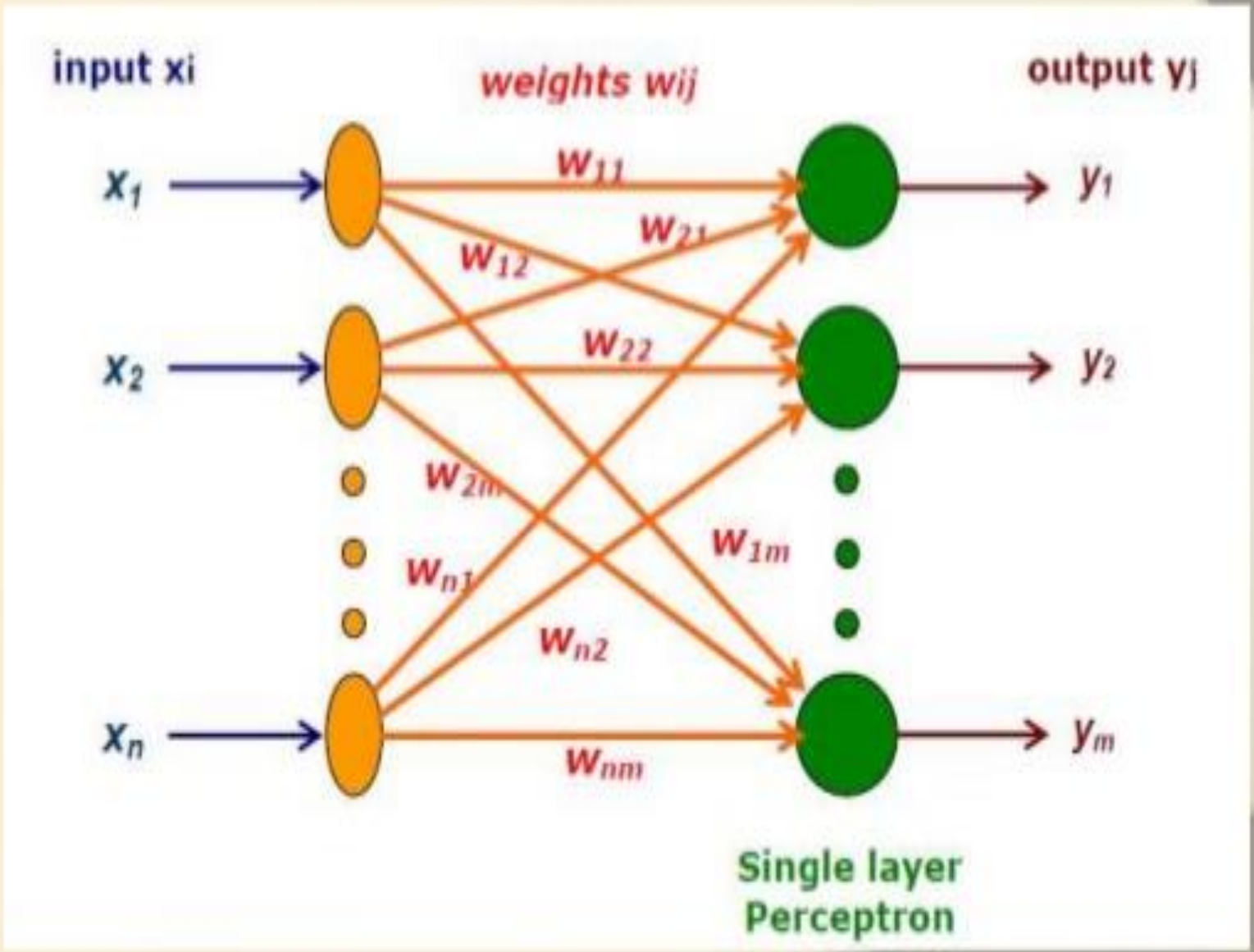


Limitation of Perceptron :-

- The perceptron can only model linearly separable functions,
  - those functions which can be drawn in 2-dim graph and single straight line separates values in two part.

  Boolean functions given below are linearly separable:
  - AND
  - OR
  - COMPLEMENT

  It cannot model XOR function as it is non linearly separable.
  - When the two classes are not linearly separable, it may be desirable to obtain a linear separator that minimizes the mean squared error.


input xi
weights wij
output yj
x1
x2
xn
w11
w21
w12
w22
w2m
wn1
w1m
wn2
wnm
y1
y2
ym
Single layer
Perceptron

Multi Layer Perceptron :-

- Multilayer perceptrons (MLP) are the most popular type of neural networks in use today. They belong to a general class of structures called feedforward neural networks, a basic type of neural network capable of approximating generic classes of functions, including continuous and integrable functions.
- A multilayer perceptron:

  has one or more hidden layers with any number of units.

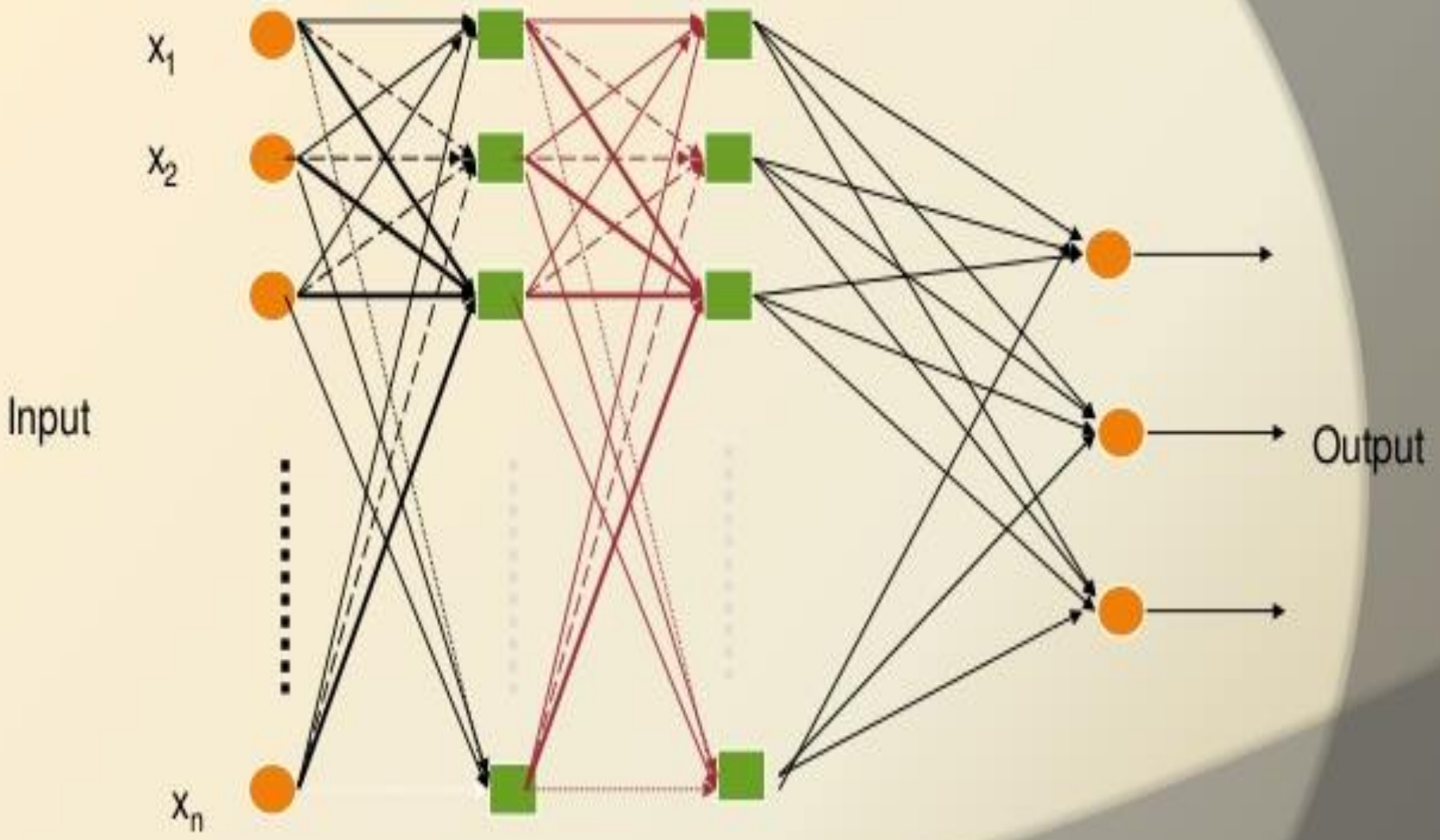
  uses linear combination functions in the input layers.

  uses generally sigmoid activation functions in the hidden layers.

  has any number of outputs with any activation function.

  has connections between the input layer and the first hidden layer, between the hidden layers, and between the last hidden layer and the output layer.

Three Layer Network :-



Layers in Neural Network :-

- **The input layer:**
  - Introduces input values into the network.
  - No activation function or other processing.

The hidden layer(s):

- Performs classification of features.
- Two hidden layers are sufficient to solve any problem.
- Features imply more layers may be better.

The output layer:

- Functionally is just like the hidden layers.
- Outputs are passed on to the world outside the neural network.

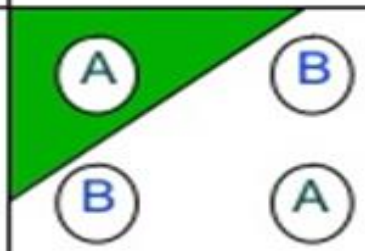
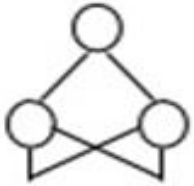
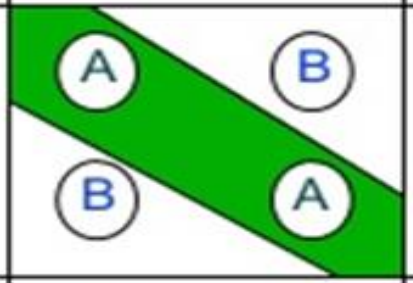
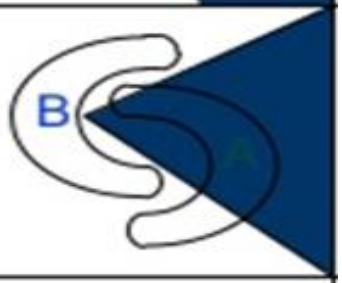
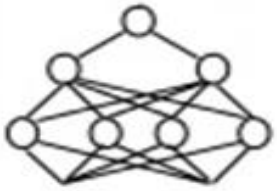
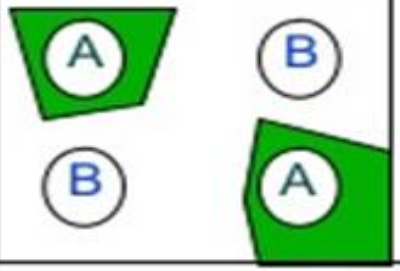
Why the MLP?

- The single-layer perceptron classifiers discussed previously can only deal with linearly separable sets of patterns.

- The multilayer networks to be introduced here are the most widespread neural network architecture
  - Made useful until the 1980s, because of lack of efficient training algorithms (McClelland and Rumelhart 1986)
  - The introduction of the backpropagation training algorithm.

Different Non-Linearly Separable Problems

http://www.zsolutions.com/light.htm

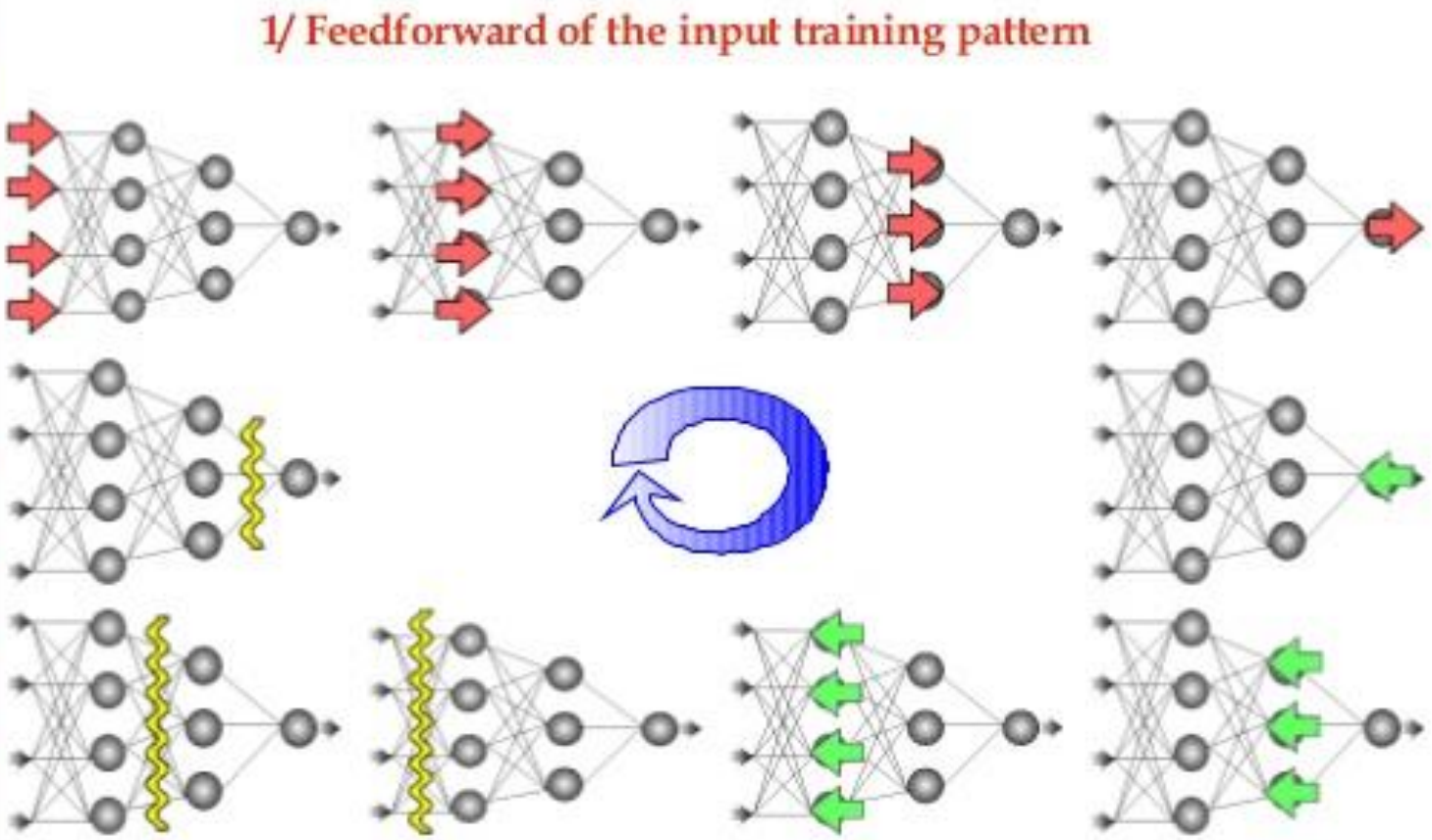
Structure	Types of Decision Regions	Exclusive-OR Problem	Classes with Meshed regions
Single-Layer 	Half Plane Bounded By Hyperplane		
Two-Layer 	Convex Open Or Closed Regions		
Three-Layer 	Arbitrary (Complexity Limited by No. of Nodes)		

What is Backpropagation?

- Supervised Error Back-propagation Training
  - The mechanism of backward error transmission (delta learning rule) is used to modify the synaptic weights of the internal (hidden) and output layers
    - The mapping error can be propagated into hidden layers
  - Can implement arbitrary complex/output mappings or decision surfaces to separate pattern classes
    - For which, the explicit derivation of mappings and discovery of relationships is almost impossible
  - Produce surprising results and generalizations

Backpropagation training cycle

Reference Eric Plummer



Thank you…

Any questions??

يرجى مسح رمز الاستجابة السريعة QR Code لتعبئة نموذج التغذية الراجعة حول المحاضرة. ملاحظاتكم مهمة لتحسين المحاضرات القادمة.

My google site