

Ch22 Deep Learning

Shun-Shii Lin

Department of Computer Science & Information Engineering
National Taiwan Normal University

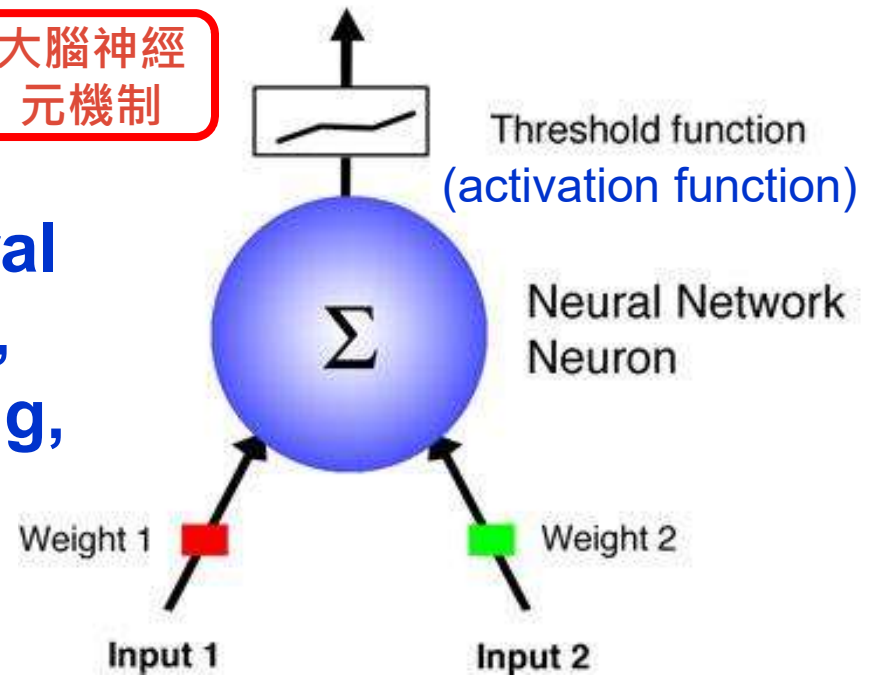


Acknowledgements: This presentation is created by Shun-Shii Lin based on the lecture slides from *The Artificial Intelligence: A Modern Approach* by Russell & Norvig, and various materials from the web.

Artificial Neural Networks

- Inspired by brain model, some of the earliest AI work aimed to create artificial neural networks(又稱connectionism, parallel distributed processing, and neural computation).

大腦神經
元機制



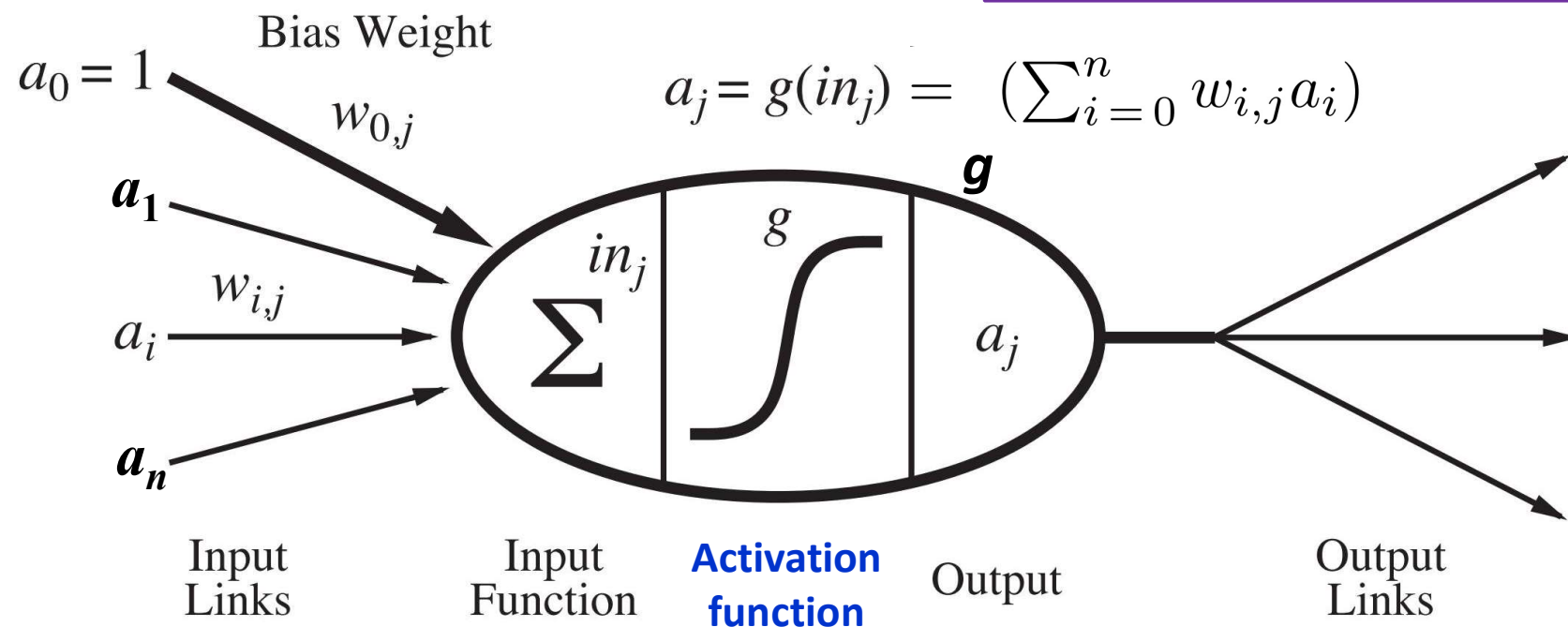
History

- 1943: McCulloch & Pitts show that **neurons** can be combined to construct a **Turing machine** (using ANDs, Ors, & NOTs).
- 1958: Rosenblatt shows that **perceptrons** will converge if what they are trying to learn can be represented. **感知器**
- 1969: Minsky & Papert showed the **limitations** of **perceptrons**, killing research for a decade. **感知器僅為單層，只能計算linear seperable function**
- 1985: **Backpropagation algorithm** revitalizes(復活) the field.

A simple mathematical model for a neuron

稱為McCulloch-Pitts neuron，
1943年 Warren McCulloch和
Walter Pitts 提出。

如果沒有**bias**，輸入**0**會永遠輸出**0.5**，不**ok**。



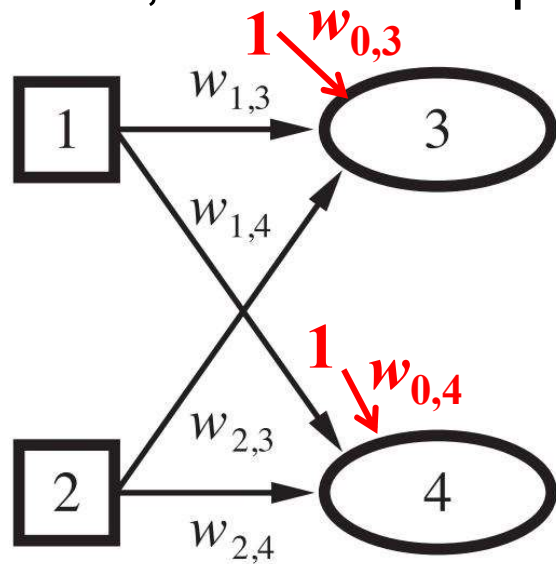
There are **2 types of components**: Linear and Non-linear.

- **Linear**: Input function
 - calculate weighted sum of all inputs.
- **Non-linear**: Activation function
 - transform sum into activation level.

22.1 Simple Feedforward Network

(a) A **perceptron** network with two inputs and two output units.

(b) A neural network with two inputs, one hidden layer of two units, and two output units.

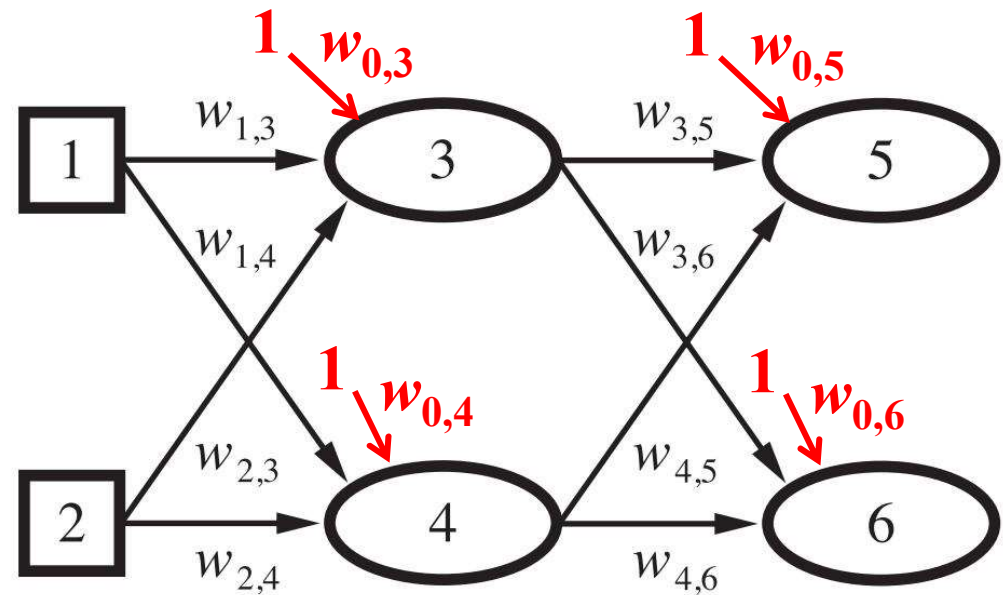


感知器

(a)

- Perceptron Network, single-layer network (proposed by **Frank Rosenblatt** in 1958, 1962) has no hidden layer.

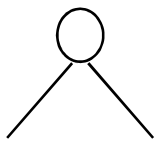
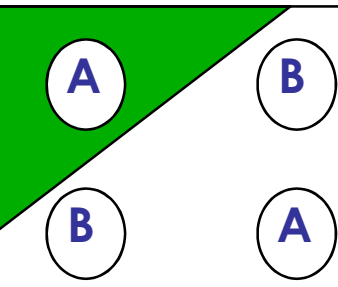
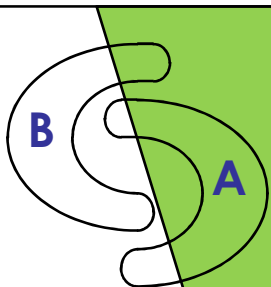
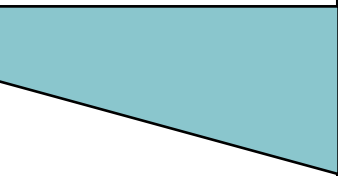
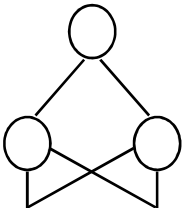
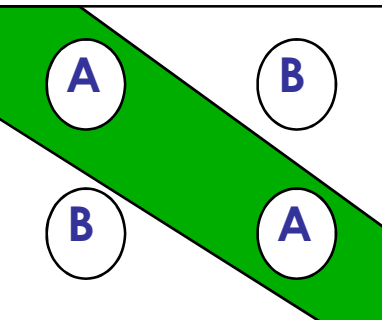
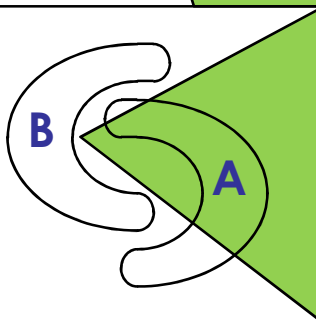
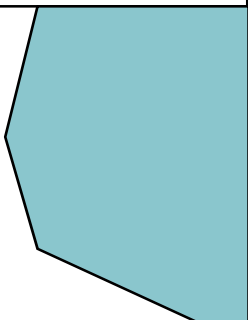
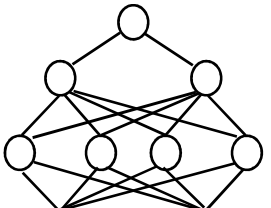
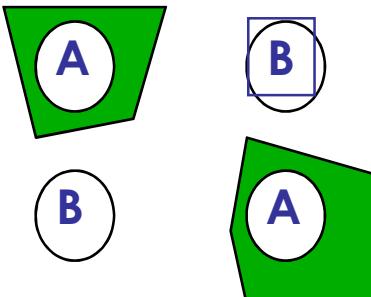
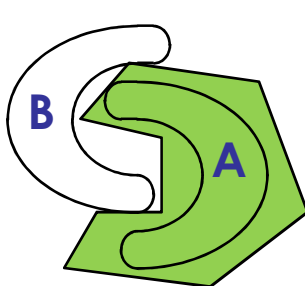
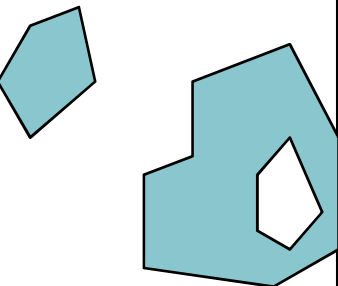
感知器只能計算linear seperable function



(b)

- Multi-layer Network has one or more **hidden** layers.
- 2 layer networks : can compute **any continuous** function
- 3 layer networks : can compute **any** function

Behavior of multilayer neural networks

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

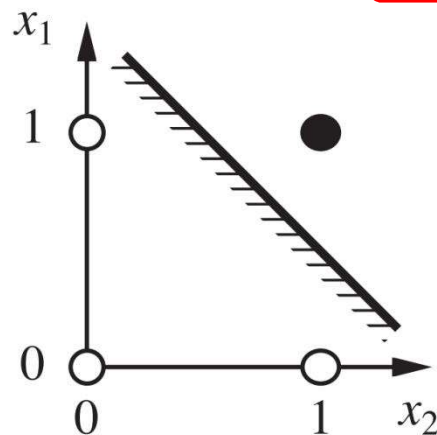
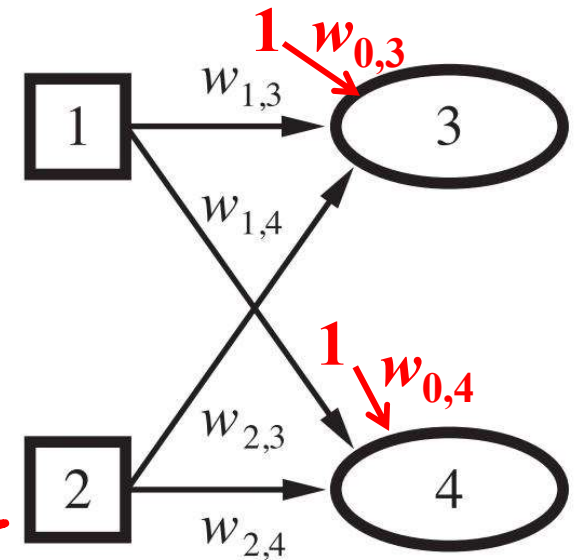
Single-layer network can not solve the two-bit adder function. ∴ XOR無解

x_1	x_2	y_3 (carry)	y_4 (sum)
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

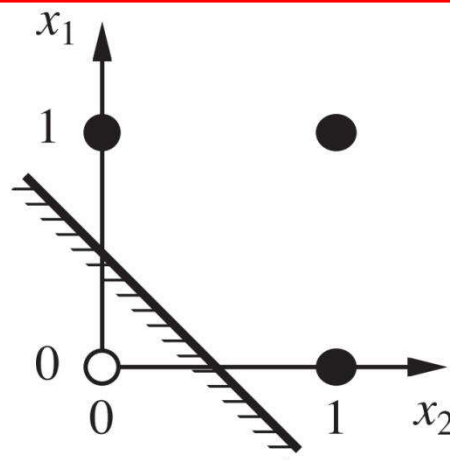
AND

XOR

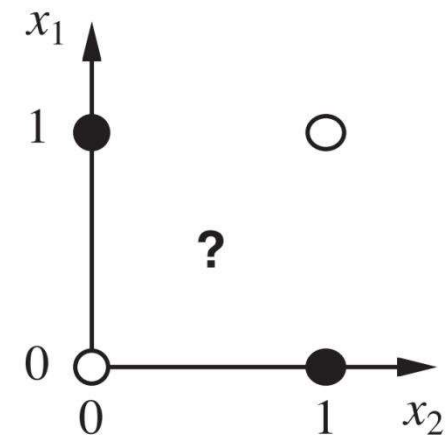
感知器只能計算linear seperable function



(a) x_1 and x_2



(b) x_1 or x_2

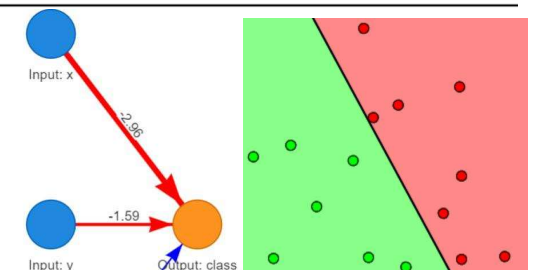


(c) x_1 xor x_2

Figure 18.21 Linear separability in threshold perceptrons.

Perceptron demo

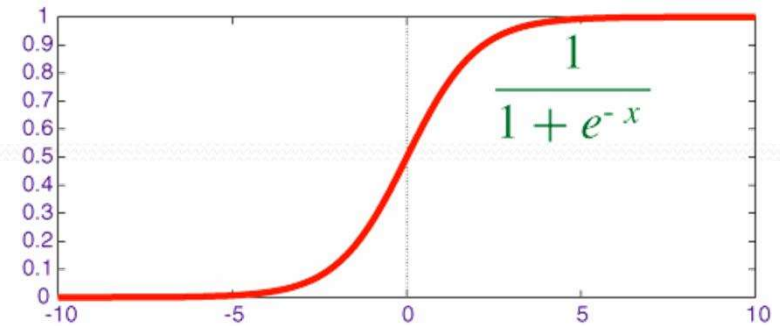
<https://lecture-demo.ira.uka.de/neural-network-demo/?preset=Rosenblatt+Perceptron>



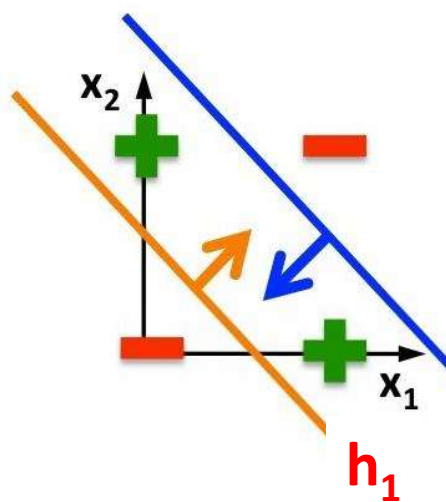
Solving XOR with a Multi-layer Network

2 layer networks : can compute any **continuous** function.

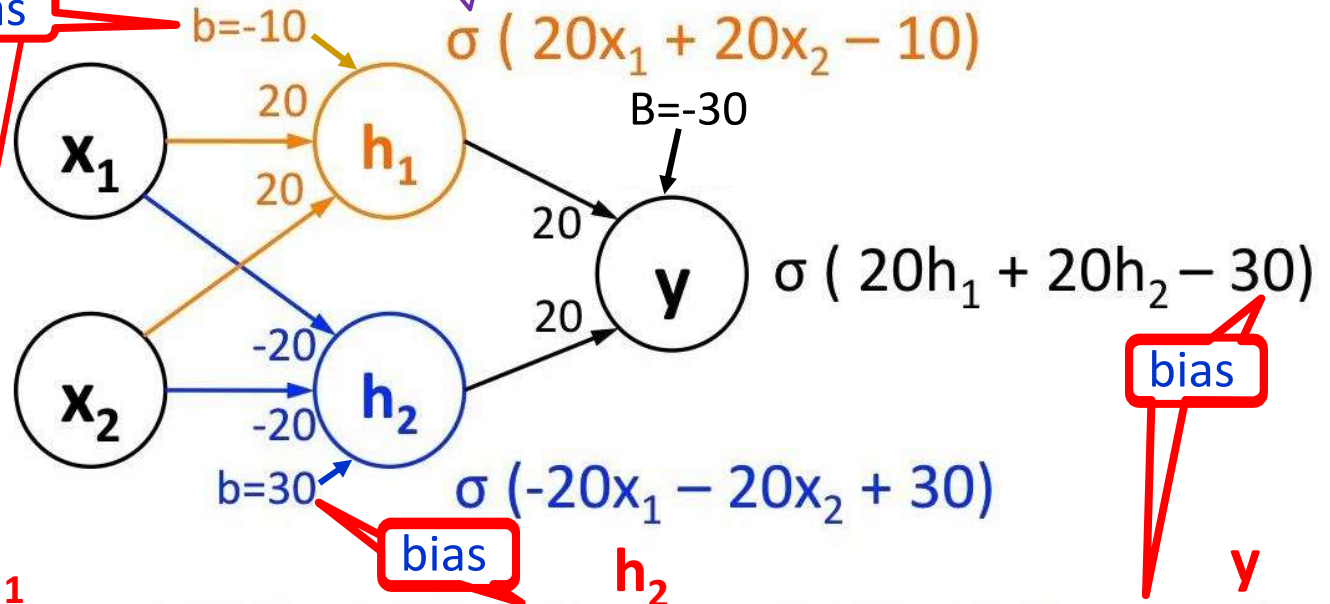
activation function σ : sigmoid



Linear classifiers cannot solve **XOR**



bias



$$\sigma(20 \cdot 0 + 20 \cdot 0 - 10) \approx 0$$

$$\sigma(20 \cdot 1 + 20 \cdot 1 - 10) \approx 1$$

$$\sigma(20 \cdot 0 + 20 \cdot 1 - 10) \approx 1$$

$$\sigma(20 \cdot 1 + 20 \cdot 0 - 10) \approx 1$$

$$\sigma(-20 \cdot 0 - 20 \cdot 0 + 30) \approx 1$$

$$\sigma(-20 \cdot 1 - 20 \cdot 1 + 30) \approx 0$$

$$\sigma(-20 \cdot 0 - 20 \cdot 1 + 30) \approx 1$$

$$\sigma(-20 \cdot 1 - 20 \cdot 0 + 30) \approx 1$$

$$\sigma(20 \cdot 0 + 20 \cdot 1 - 30) \approx 0$$

$$\sigma(20 \cdot 1 + 20 \cdot 0 - 30) \approx 0$$

$$\sigma(20 \cdot 1 + 20 \cdot 1 - 30) \approx 1$$

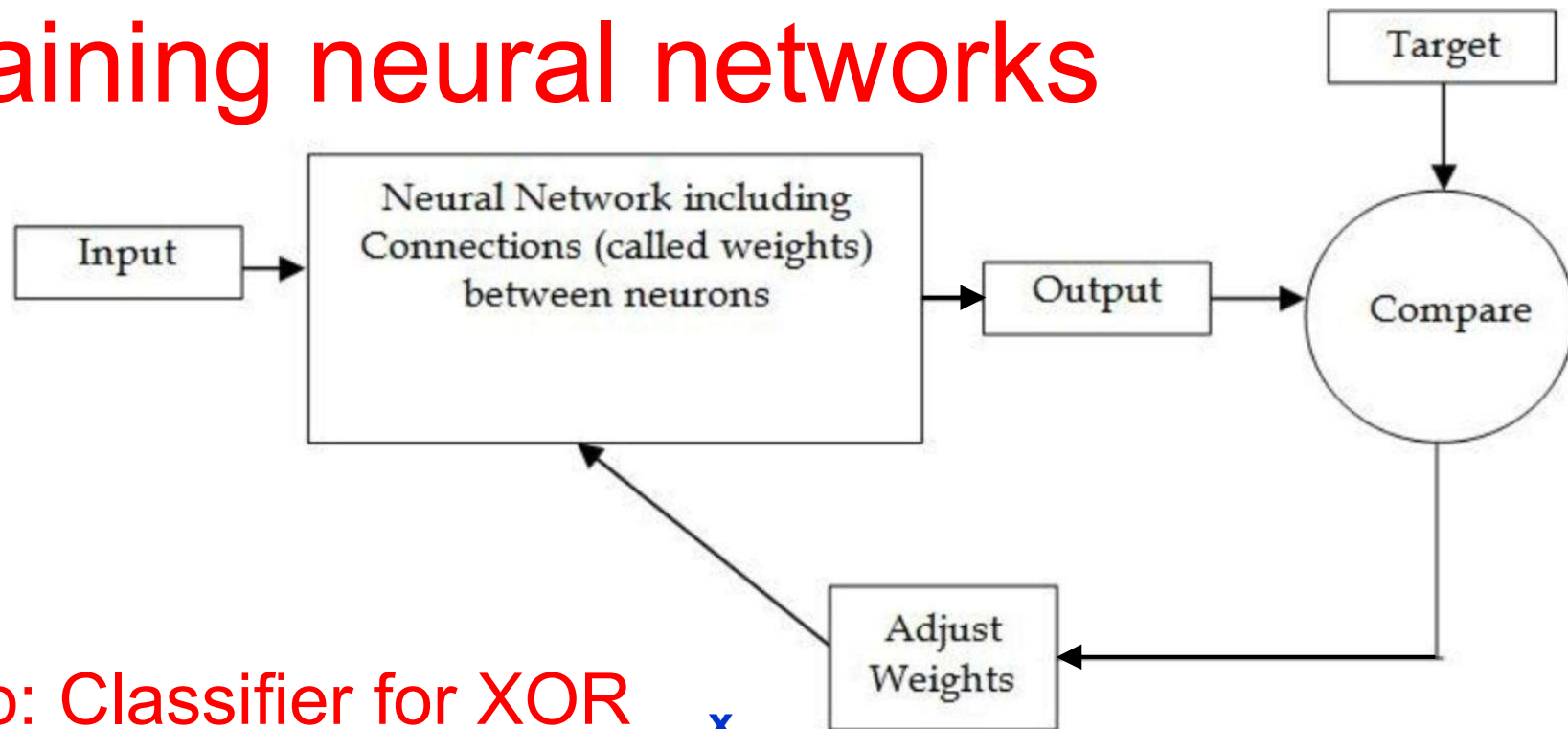
$$\sigma(20 \cdot 1 + 20 \cdot 1 - 30) \approx 1$$

這些權重有沒有辦法找出來？只有唯一解嗎？

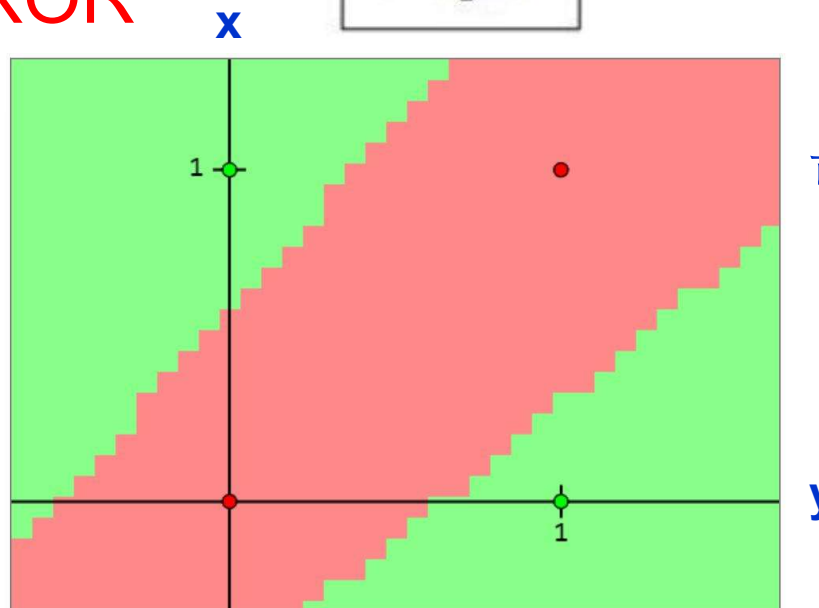
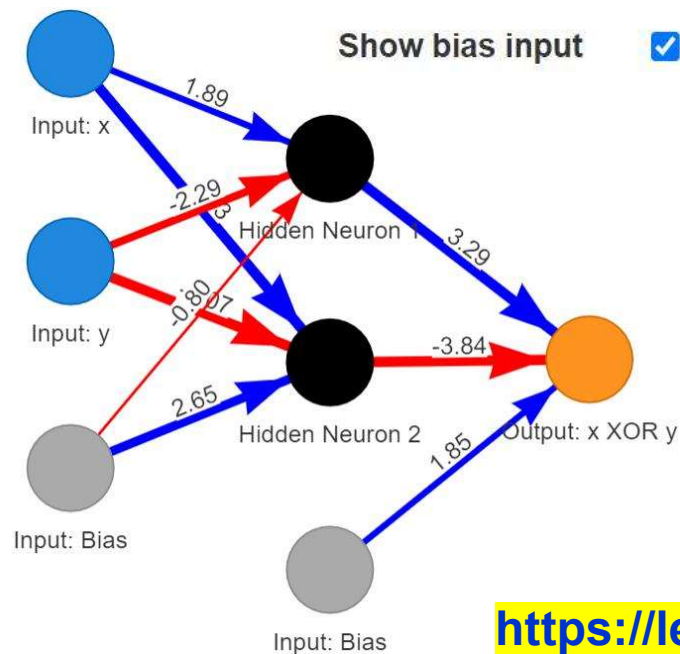
Copyright © 2014 Victor Lavrenko

Ch22-7

Training neural networks



demo: Classifier for XOR



可增加紅點、綠點

Add Red Add Green