Neural Networks Basics

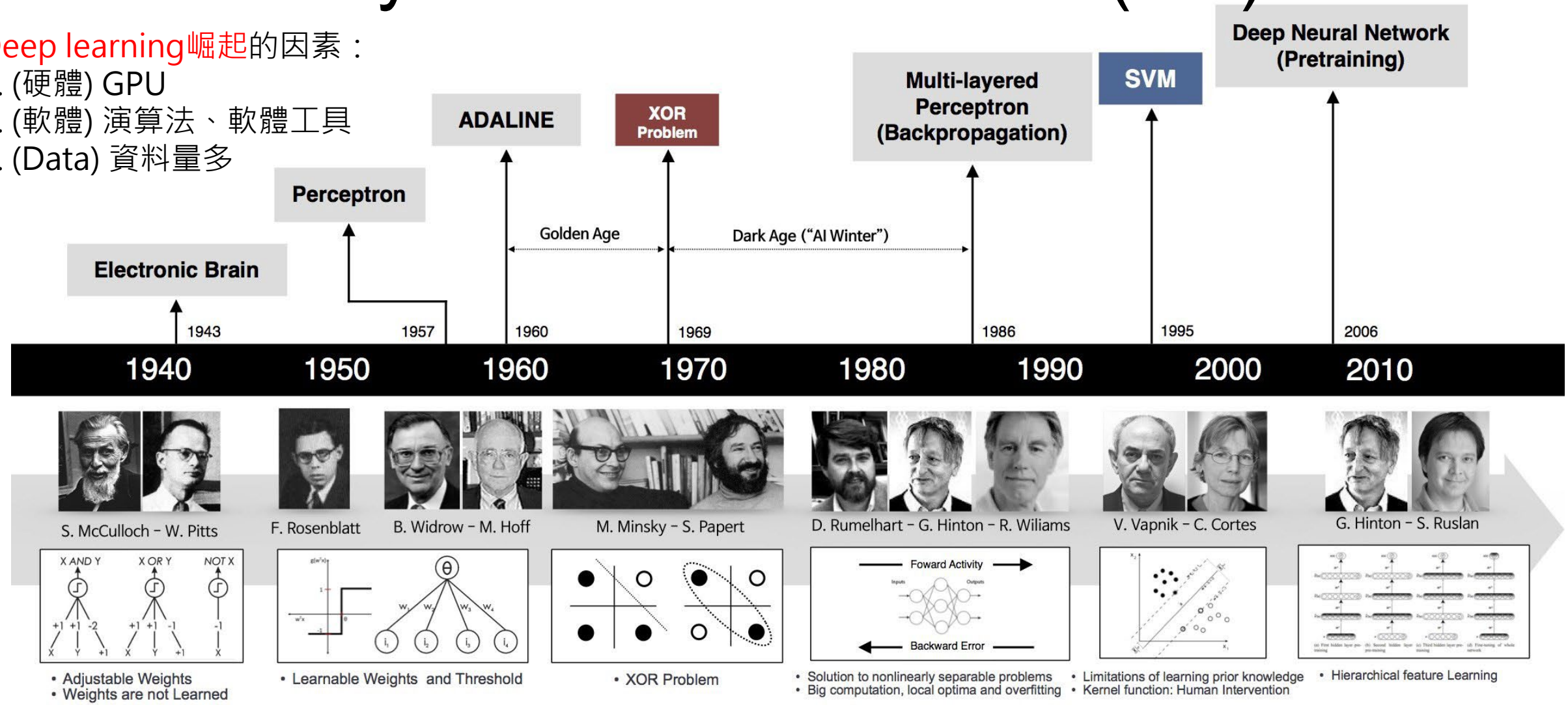
Jen-Ing Hwang
Department of Computer Science and
Information Engineering
Fu Jen Catholic University

Outline

1. Brief History of Neural Networks
2. Biological Inspiration
3. Artificial Neuron
4. Illustrative Example: Neural Representation
5. Activation Functions
6. Network Architectures
    - A Layer of Neurons
    - Multiple Layers of Neurons

Brief History of Neural Networks (1/4)

Deep learning崛起的因素：
1. (硬體) GPU
2. (軟體) 演算法、軟體工具
3. (Data) 資料量多



用人工自己去調
權重
不是學習而來的

Ex:考兩科，其中一科一百分就有獎勵，考兩科更好，白色的代表沒有一百分

expression1	expression2	result
1	0	1
0	1	1
0	0	0
1	1	1

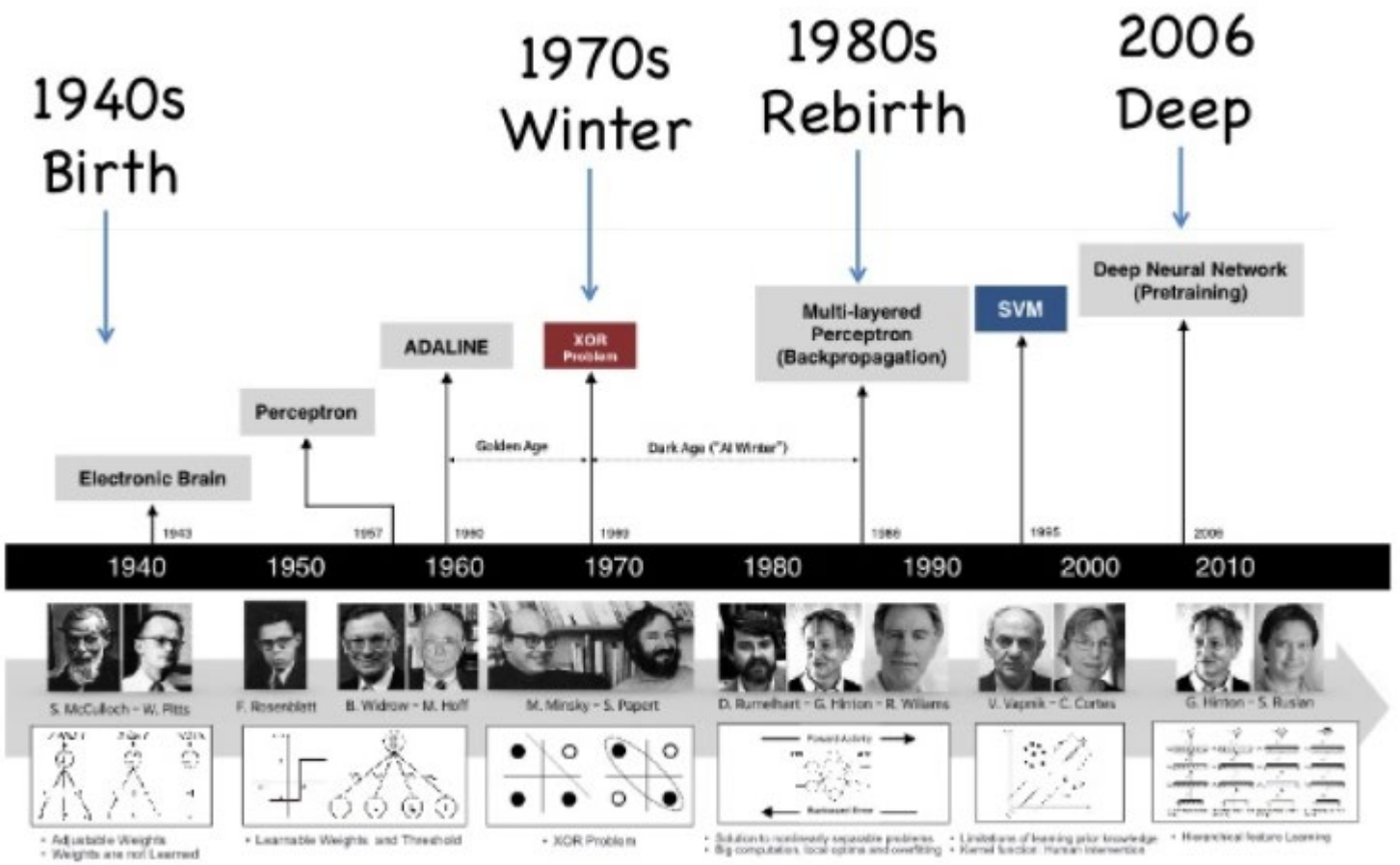
Brief History of Neural Networks (2/4)

- 1943: McCulloch & Pitts
  - Show that neurons can be combined to construct a Turing machine
- 1958: Rosenblatt
  - Shows that perceptron will converge if what they are trying to learn can be represented
- 1950s: Rosenblatt, Widrow & Hoff
  - First practical networks and learning rules
  - The Widrow-Hoff learning rule is still in use today
- 1969: Minsky & Papert
  - Show the limitations of perceptron, killing research in neural networks for a decade

Brief History of Neural Networks (3/4)

- 1980s: Rumelhart, Hinton & Williams
  - The backpropagation algorithm revitalizes the field
- 1998: Yan LeCun
  - Convolutional Neural Networks with Backpropagation for document analysis
- 2006: The Hinton lab solves the training problem for Deep Neural Networks
- 2012 - present:
  - A variety of deep learning algorithms are increasingly emerging

Brief History of Neural Networks (4/4)



Biological Inspiration (1/3)

- Biological motivation
  - Biological learning system (brain)
  - Complex network of neurons

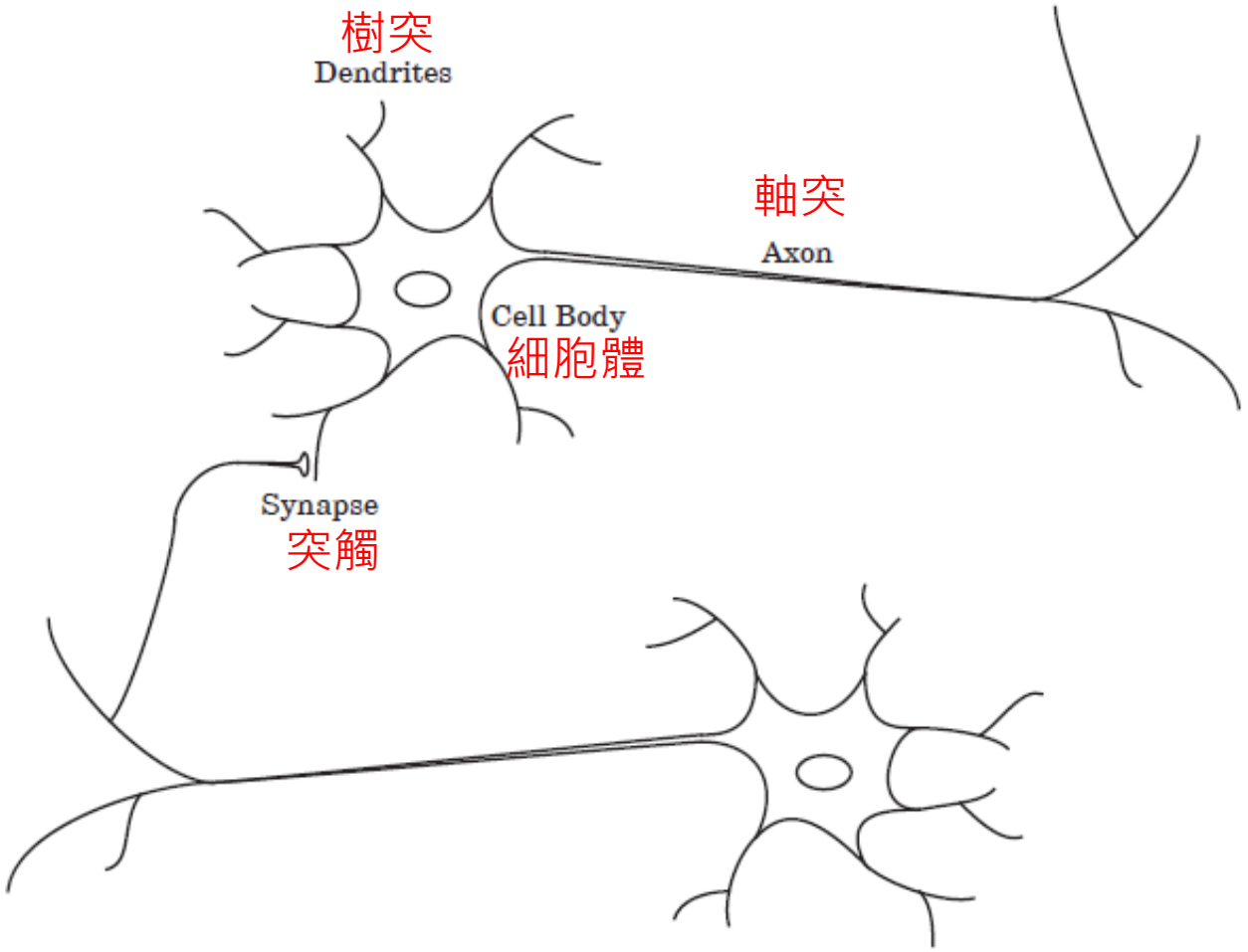


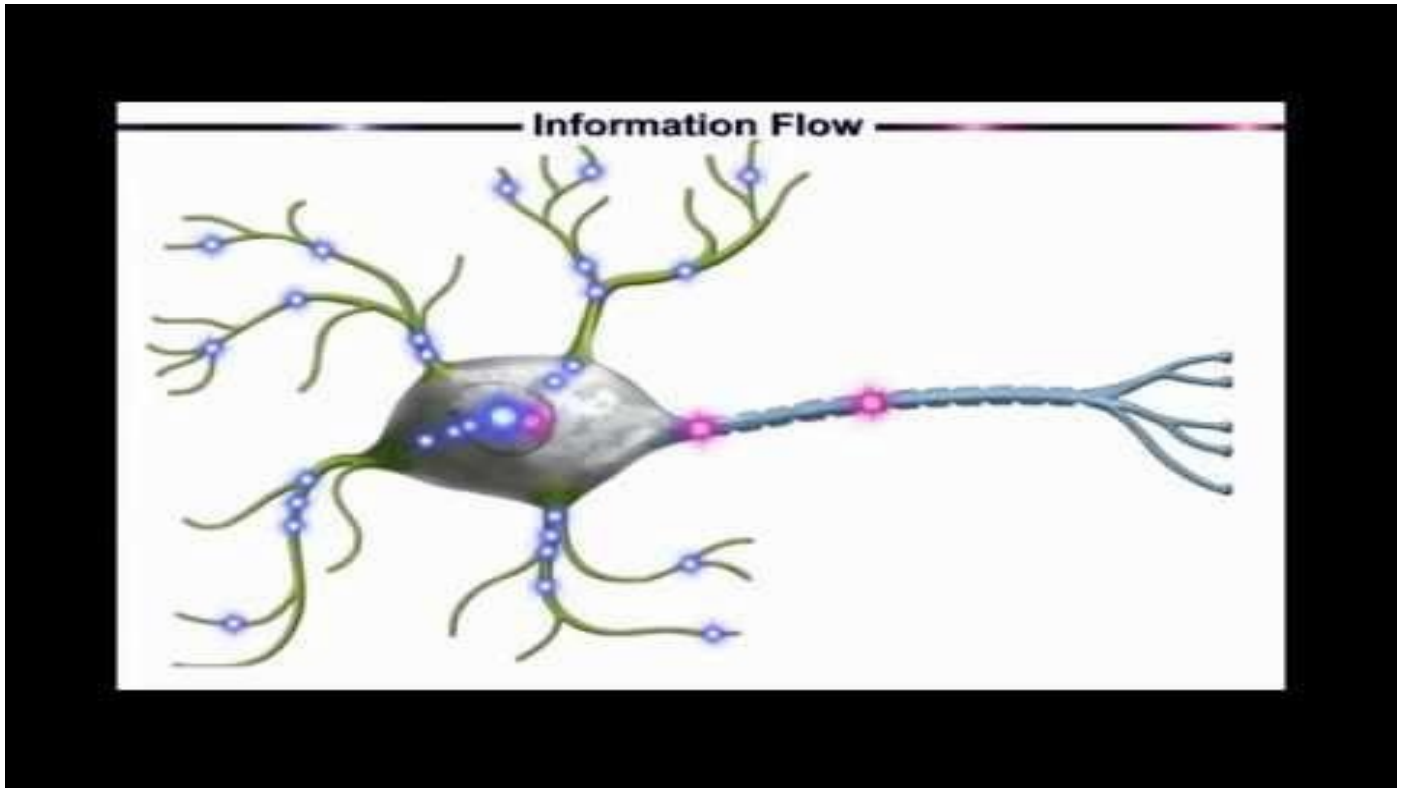
Figure adapted from Fig 1.1 in Ref. [2]

Biological Inspiration (2/3)

- A neuron consists of a cell body (細胞體)(中心體) , dendrites (樹突)(輸入) and an axon (軸突)(輸出).
- Neurons are massively interconnected, where an interconnection is between the axon of one neuron and a dendrite of another neuron.
- This connection is referred to as a synapse (突觸)->用來決定每個輸入值的重要程度

Biological Inspiration (3/3)

Neural Networks: A Simple Explanation



https://youtu.be/gcK_5x2KsLA

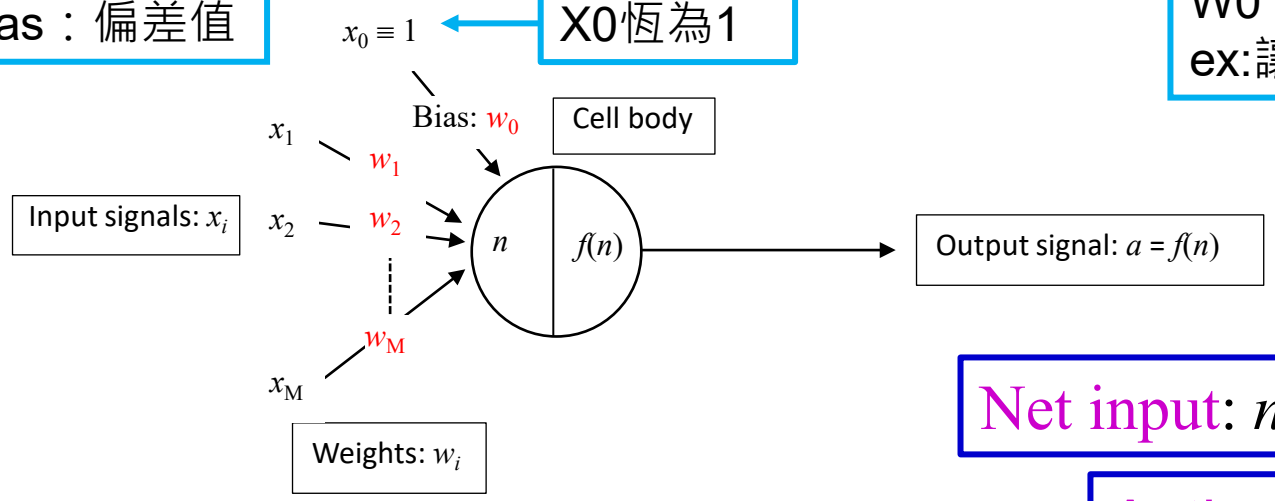
Artificial Neuron (1/2)

- An artificial neuron is a model of a biological neuron.
  - Based on model of brain
  - Network of simple units
  - Real-valued inputs & outputs

Bias：偏差值

X0恆為1

W0：用來移動直線
ex:讓線不通過原點



Net input: $n = w_0 x_0 + w_1 x_1 + \ldots + w_M x_M$

Activation Function: $f(n)$

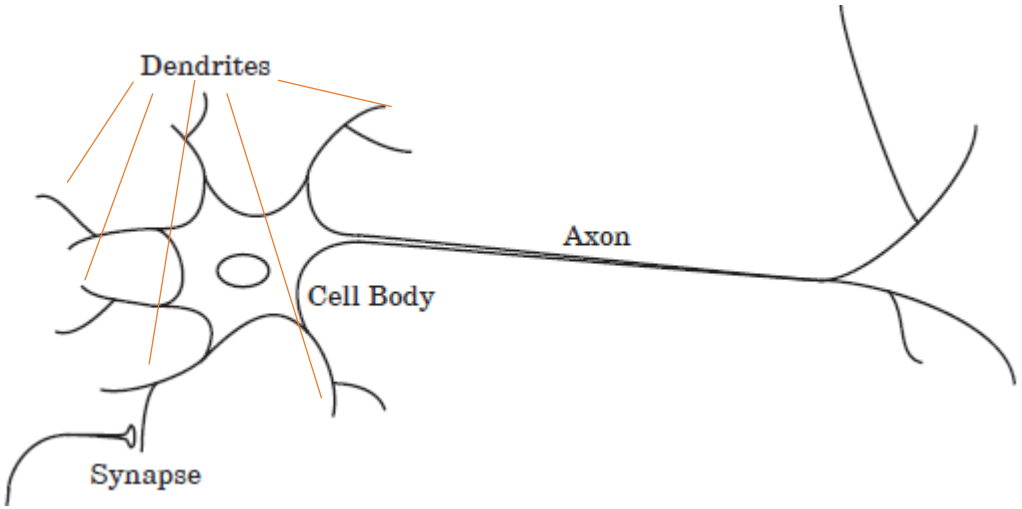
Artificial Neuron (2/2)

Bias and Weight

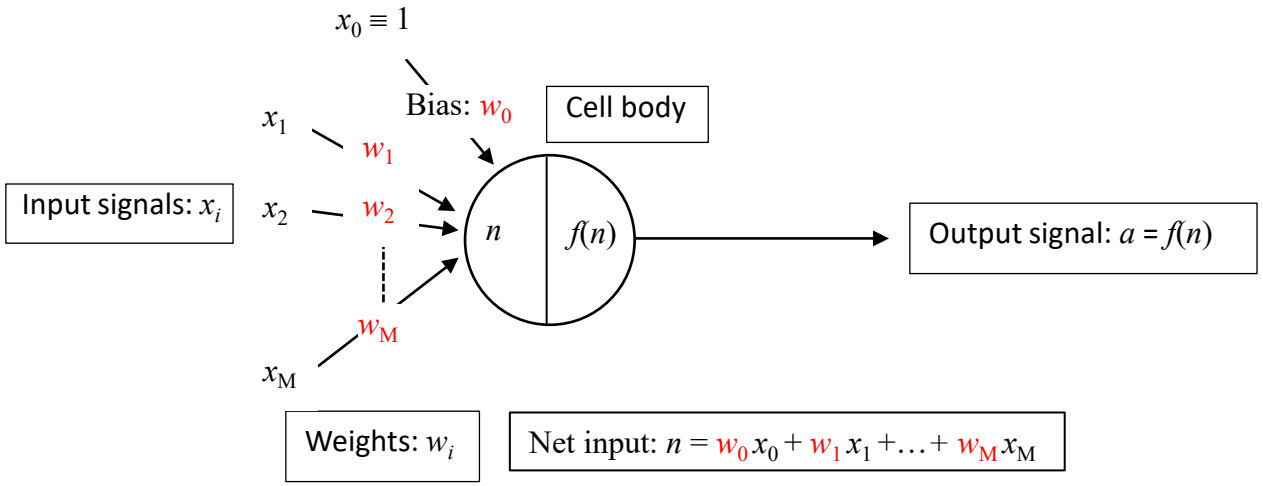
- The **bias** *b* is much like a **weight** *w*, except that it has a constant input of 1. It can be **omitted** for some situation.
- **The bias** *b* (= w_0) and **weights** w_i, I = 1, 2,…, M, are adjustable 可調整的 scalar parameters of the neuron. They can be adjusted by some learning algorithm so that the neuron input/output relationship meets some particular goal.

Biological Neuron Vs. Artificial Neuron

Biological Neuron



Artificial Neuron



Dendrites (樹突)	⟷	Input signals (x_i)
Synapses (突觸)	⟷	Weights (w_i)
Cell Body (細胞體)	⟷	Cell Body (n, $f(n)$)
Axon (軸突)	⟷	Output signal (a)

Illustrative Example for Linear Classification

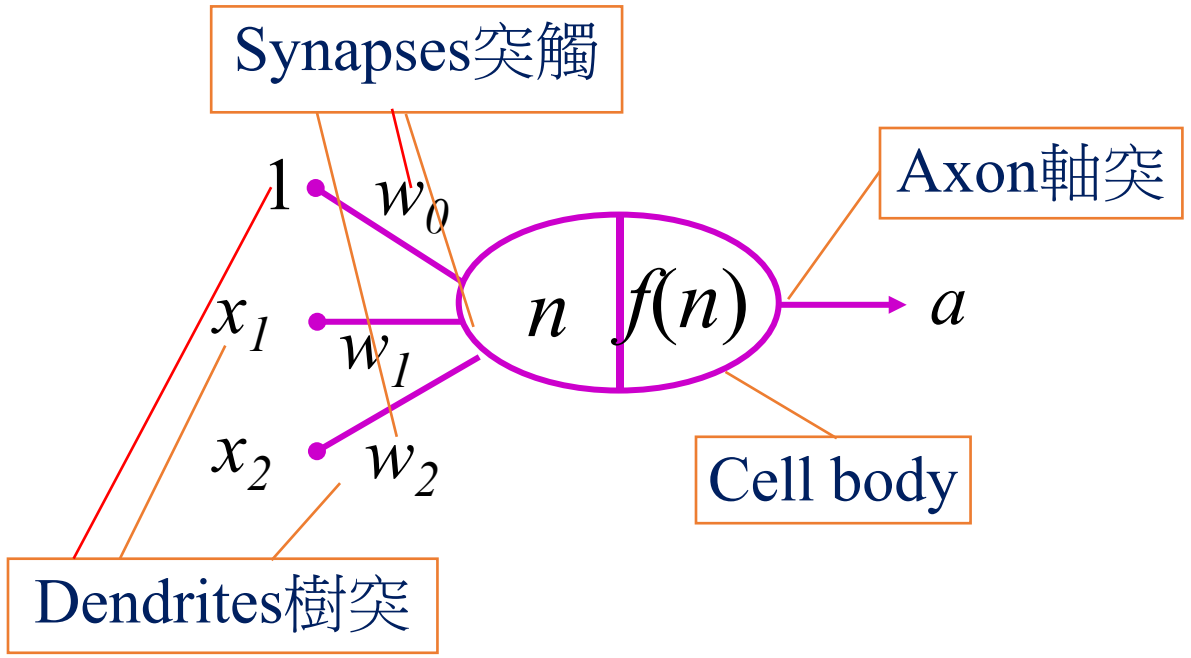
- Dataset *D*
  - Students take a class on a **Pass/Fail** basis

Students	x_1 (Midterm)	x_2 (Final)	y (Pass/Fail)
A	80	60	+1 (Pass)
B	50	50	–1 (Fail)
C	90	80	+1 (Pass)
D	30	60	–1 (Fail)
E	40	90	+1 (Pass)
F	90	50	+1 (Pass)

Illustrative Example: Neural Representation

Students' Performance Dataset

Students	x_1 (Midterm)	x_2 (Final)	y (Pass/Fail)
A	80	60	+1
B	50	50	–1
C	90	80	+1
D	30	60	–1
E	40	90	+1
F	90	50	+1



$$f(x_1, x_2) = \text{sign}\,((w_1 x_1 + w_2 x_2) - \text{threshold})$$
$$= \text{sign}\,(w_0 x_0 + w_1 x_1 + w_2 x_2)$$

– threshold(門檻值) 1

Ex:
0.4 X1 + 0.6 X2 > 50 及格
則0.4 = W1, 0.6 = W2

Activation Functions (1/3) (= transfer function)

- Activation function $f(n)$:
  - Define the output of a node when given the inputs of the node
    - The output of a neuron can either inhibit (抑制、不觸發) or excite (激發、觸發) the net input *n*

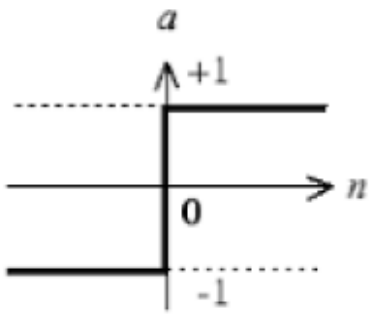
Ex: 及格 -> 觸發
不及格 -> 不觸發

Activation Functions (2/3)

函式考試必考!!

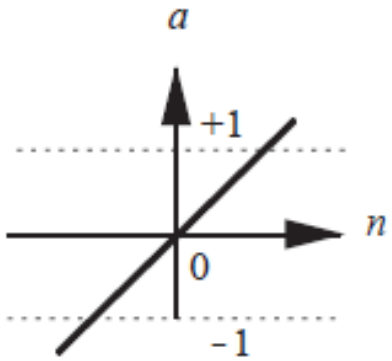
一個東西有很多名稱，表示它很重要，表示有個領域的人來做研究，並用不同的名稱來稱呼它

- The activation function *f* may be a ***linear*** or ***nonlinear*** function of *n*



Sign function:

$f(n) = \text{sign}\ (n)$

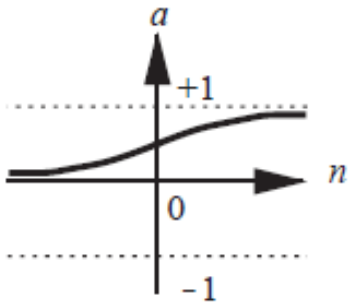


Linear function:

$f(n) = n$

0.4 $X1$ + 0.6 $X2$ = 0

Ex: 老師完全不調分，也沒有門檻值



Logistic sigmoid function:

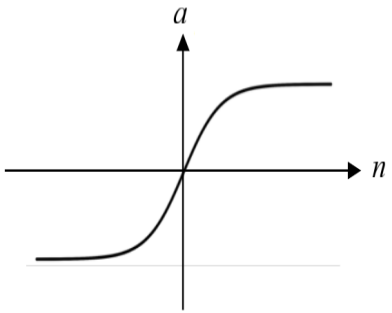
$$f(n) = \sigma(n) = \frac{1}{1+e^{-n}}$$

n = 0 -> $\sigma(n) = ½$

n = ∞ -> $\sigma(n) = 1$

Activation Functions (3/3)

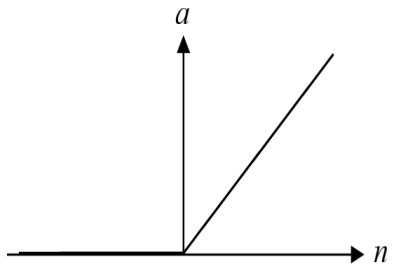
- The activation function *f* may be a ***linear*** or ***nonlinear*** function of *n*



Hyperbolic Tangent Tanh function（雙曲函數）:

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

是**Logistic sigmoid function** 的變形
Ex: 把同學的分數壓縮在1和-1之間



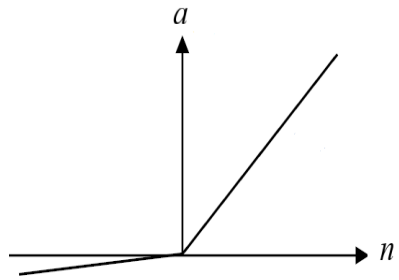
Rectified Linear Unit（ReLU） function（線性整流函數）:

$$f(n) = \max(0, n)$$

-> 變形的線性函數

訊號很強 就希望它的輸出也很強

如果訊號很弱 甚至是負的 就當作它們不重要



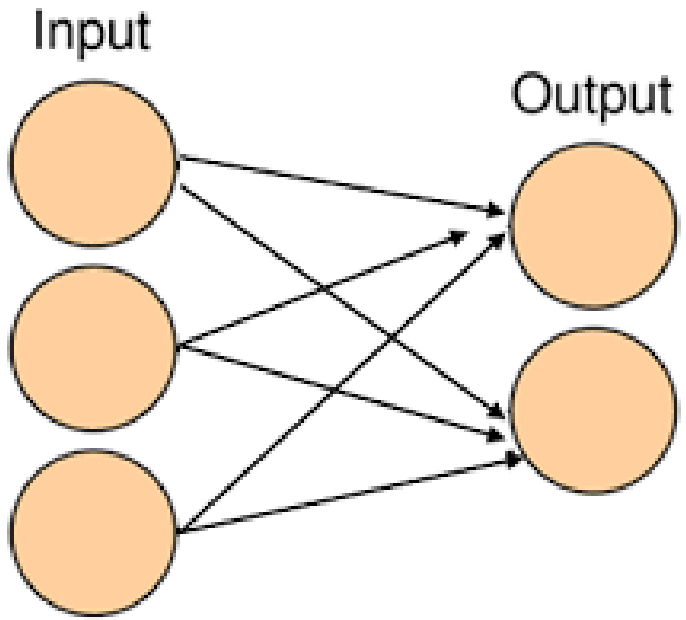
Leaky ReLU function:

$$f(n) = \max(0.01n, n)$$

實務上，負的有跟沒有差不多，所以比較 prefer **Rectified Linear Unit ReLU**

Network Architectures (1/3)

- A Layer of Neurons

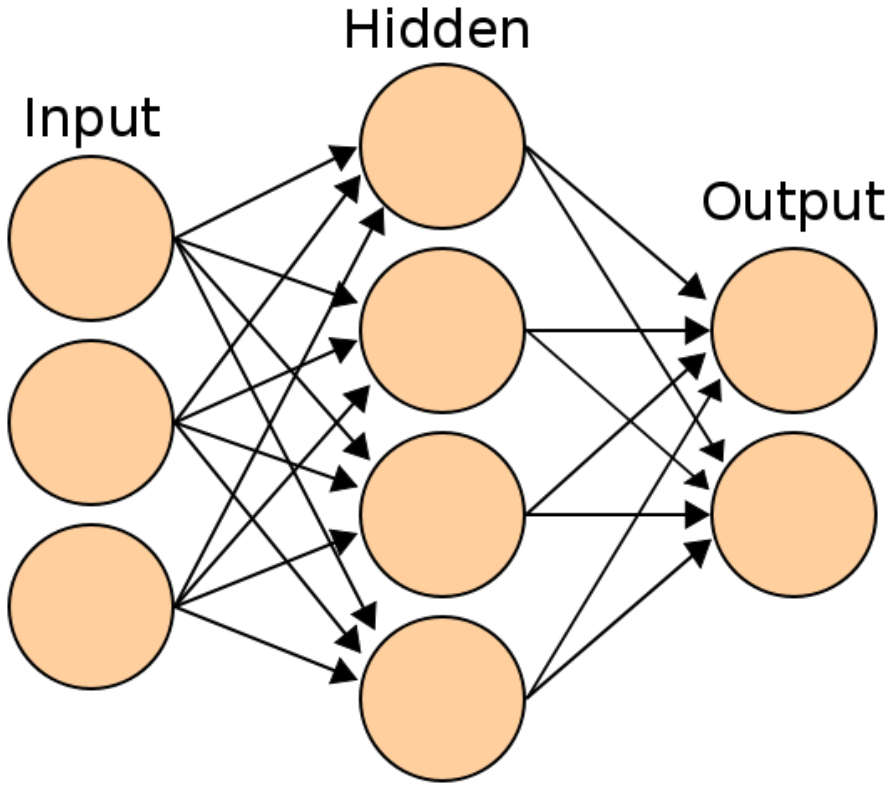


Ex:這個神經元做pass or fail

Ex:這個神經元做數值

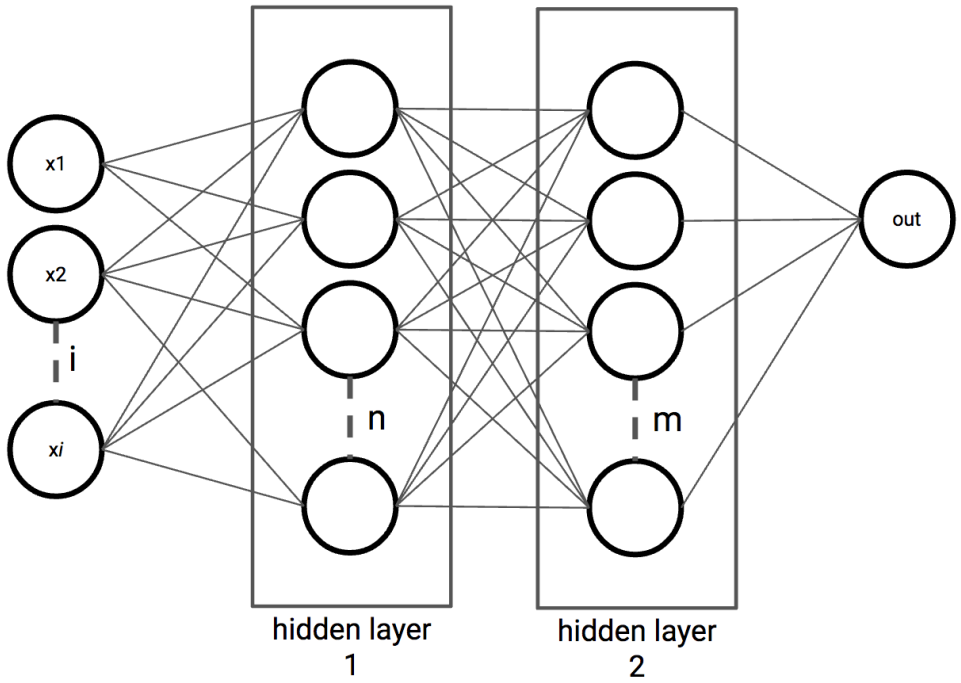
Network Architectures (2/3)

- Multiple Layers of Neurons



Network Architectures (3/3)

- Multiple Layers of Neurons
  - The layer that receives inputs is called the input layer.
  - The outputs of the network are generated from the output layer.
  - Any layer between the input and the output layers is called a hidden layer.



-> 如果問題很簡單 不需要隱藏層
輸入層和輸出層永遠只有一個 但是隱藏層可以有很多個
Ex: 如果有兩個隱藏層 一個輸出層 我們稱它為三層的神經元
因為真正有神經元的是隱藏層一 隱藏層二和輸出層 (但也有人說它是四層 ex:有人說 tree 的 root 是在 level 1也有人說在level 0)

Representational Power of Feedforward（前饋）Networks

- Boolean functions
  - Every boolean function can be represented exactly by some network with two layers (*i.e.*, one single hidden layer) of units
  - But might require exponential (in number of inputs) hidden units
- Continuous functions
  - Every bounded continuous function can be approximated with arbitrarily small error by a network with two layers
- Arbitrary functions
  - Any function can be approximated to arbitrary accuracy by a network with three layers (*i.e.*, two hidden layers)

Note: Although a limited depth of feedforward networks can provide a very expressive hypothesis space for Backpropagation, the number of neurons in a hidden layer can be exponential in the input dimension.

Summary

- An Artificial Neural Network (ANN) is a computational model inspired by biological network systems.
- A biological neuron:
  - Dendrites, a cell body, and an axon.
  - Synapse is between the axon of one neuron with the dendrite of another neuron.
- An artificial neuron:
  - A set of input nodes, a set of weights, an activation function that relates the total synaptic input to the output (activation) of the neuron.
- Activation function
  - Define the output of a neuron
- The essential elements of an artificial network are an input layer, an output layer, and may contain hidden layers.

References

[1] Networks are like onions: Practical Deep Learning with TensorFlow, Barbara Fusinska, London Tensorflow Meetup, Jun 21, 2017.

[2] Neural Network Design (2nd Edition), Martin T Hagan , Howard B Demuth , Mark H Beale , and Orlando De Jesús, Martin Hagan, 2014.

[3] Pattern Recognition and Machine Learning, Christopher Bishop, Springer-Verlag New York, 2006.

[4] Machine Learning, Tom M. Mitchell, McGraw-Hill, 1997.