

AI CUP 教育部全國大專校院
人工智慧競賽
圖像辨識概念課教材

2020.03.26



What is A I





什麼是人工智慧

人工智慧三大概念：感知、認知、行動

人工智慧之父：約翰·麥卡錫





A I 三大學派

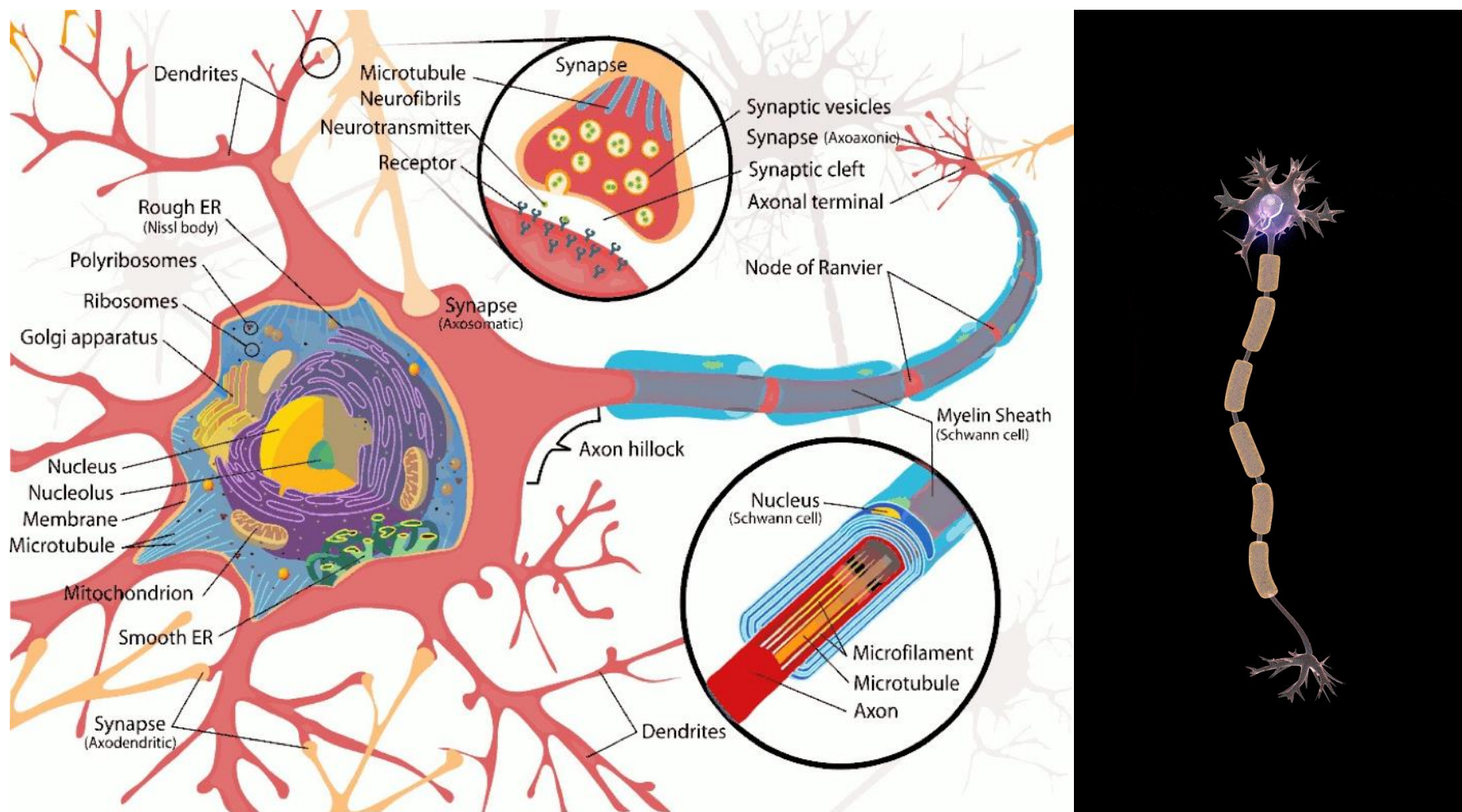
深度學習:模仿大腦皮層神經網路

知識圖譜:模仿人類邏輯與抽象推理

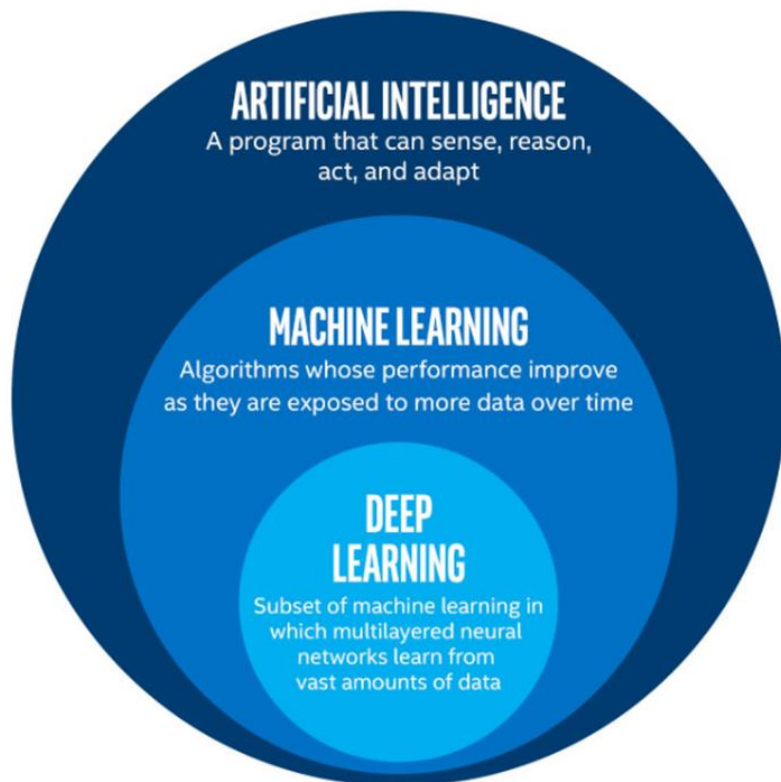
強化學習:模仿生物獎懲學習機制



A I 三大學派-模仿大腦皮層神經網路



人工智慧、機器學習和深度學習的差異



Interception from Prowesscorp website.

AI: 計算機模仿人類思考進而模擬人類的能力/行為。

ML: 從資料中學習模型。

DL: 利用多層的非線性學習資料表徵。

類神經網絡

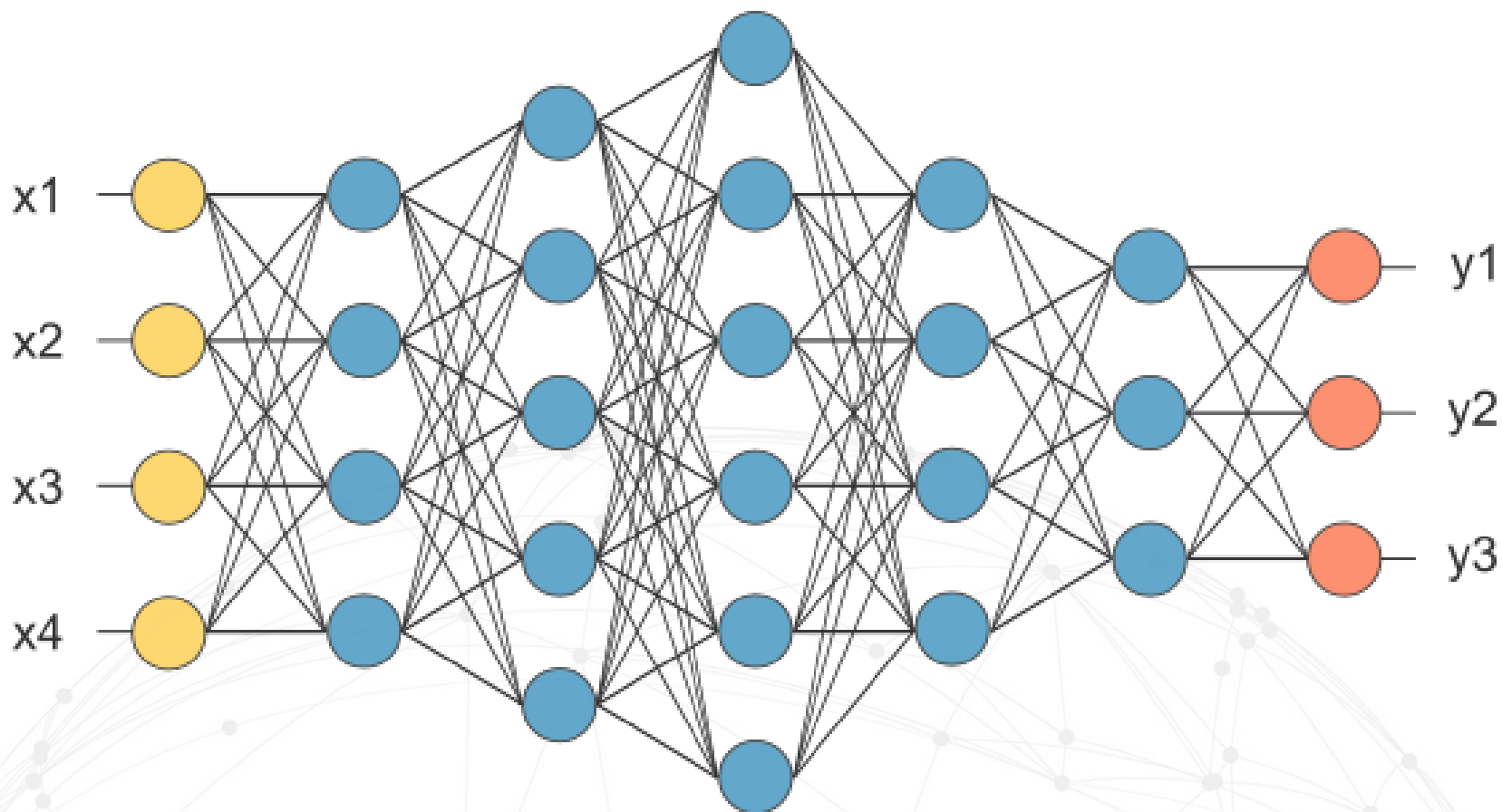


類神經網絡架構

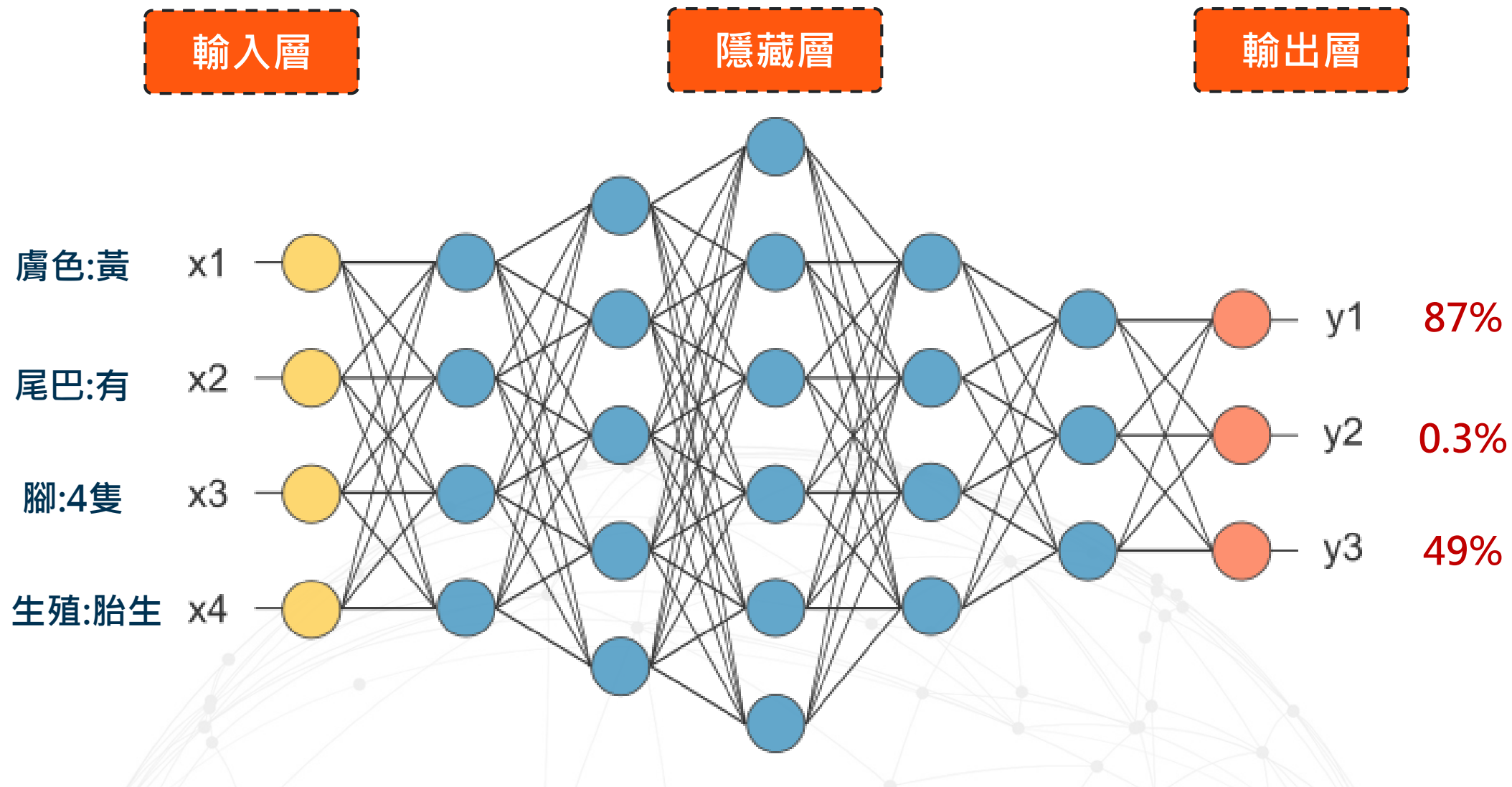
輸入層

隱藏層

輸出層



動物圖片辨識 AI



卷積神經網絡





PREMIERES MONDAY ON CNN

CUOMO
—PRIME TIME—

9^P
ET



ONLY ON CNN

A CNN ORIGINAL SERIES

ANTHONY BOURDAIN
PARTS UNKNOWN

SUNDAY

9^P
ET/PT



11:43 AM PT

NEWSROOM

資料匯入流程

讓圖片及標籤可以順利讀取進程式中，為讓後續可餵到模型中進行學習。



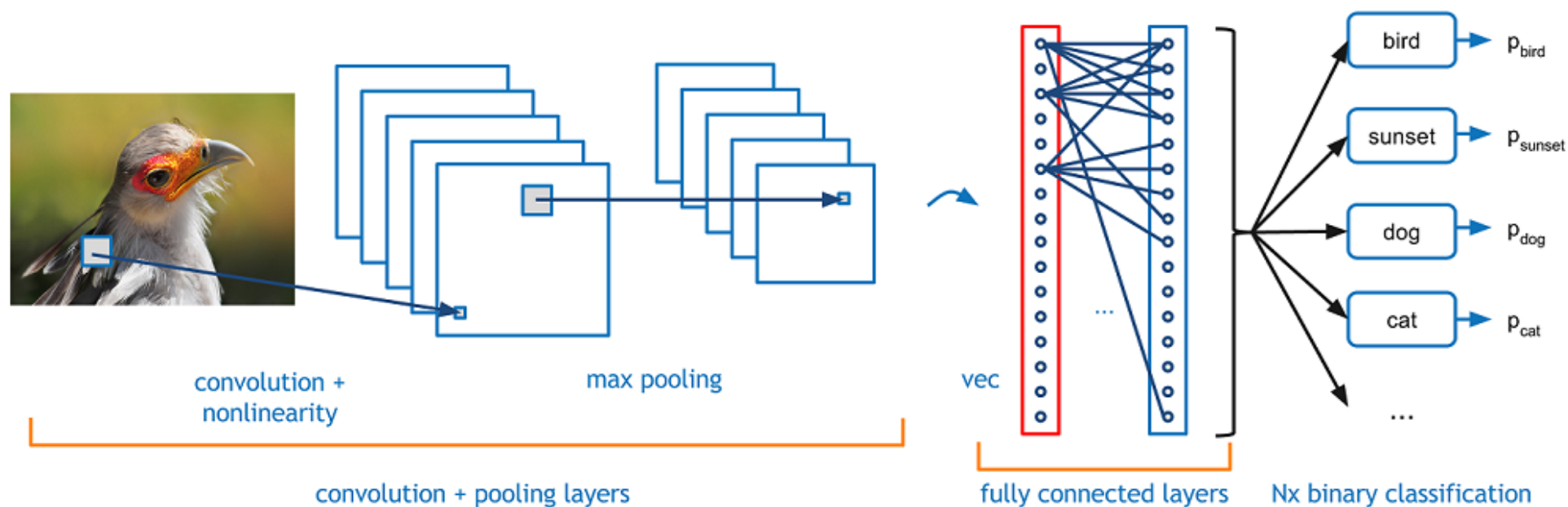
資料前處理流程

讓模型可以更有效果的學習。



卷積神經網路

Convolutional Neural Network (CNN) 中文稱作卷積神經網路，是一種深度神經網路的架構，它主要被應用於影像辨識、視訊分析、自然語言處理以及圍棋等，尤其它在影像處理的應用特別出色。



卷積神經網路簡史

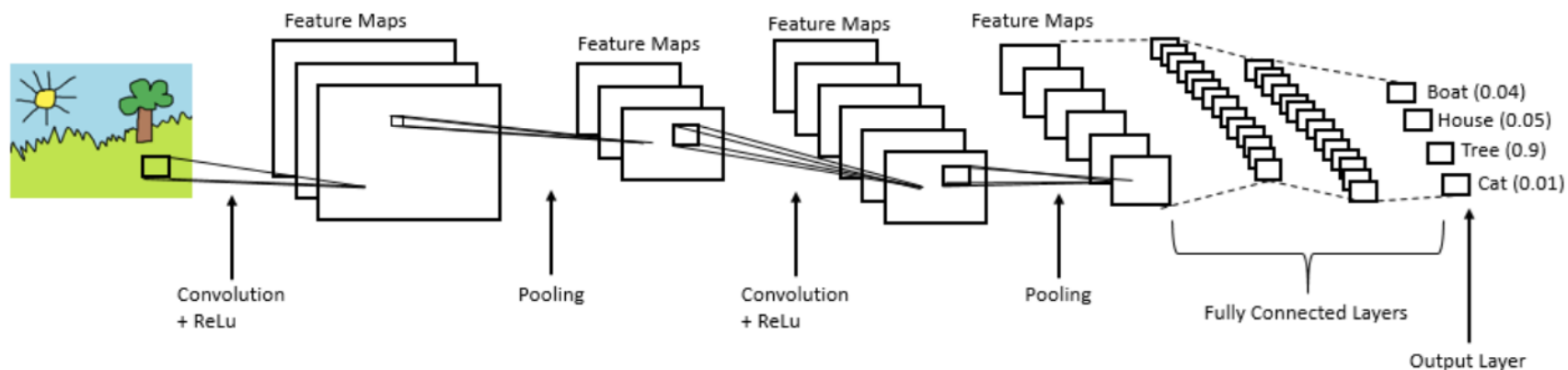


楊立昆(Yann LeCun)，法國裔美國計算機科學家，他在機器學習、計算機視覺、移動機器人和計算神經科學等領域都有很多貢獻。他最著名的工作是在光學字符識別和計算機視覺上使用卷積神經網絡 (CNN)，他也被稱為卷積網絡之父。

卷積神經網路架構

CNN架構基本包含：

- 卷積層 (Convolutional layer)
- 線性整流層 (Rectified Linear Units layer, ReLU layer)
- 池化層 (Pooling layer)
- 全連接層 (Fully Connected layer)



認識圖片的像素

以下圖為例(左圖 $18 \times 18 = 324$ ，右圖 $50 \times 70 = 3500$)

此圖在說明，若圖片的像素越高，影像則越清晰。



<http://blog.logo123.net/wp-content/uploads/2015/01/250-250-1.jpg>

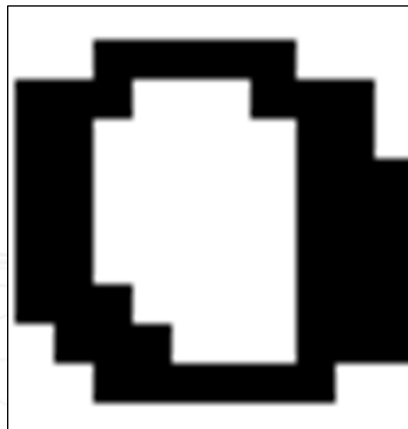


<http://tech.sina.com.cn/it/2015-07-04/doc-ixfesftm9556659.shtml>

以下圖為例，是一個數字0(像素值為 $10 \times 10 = 100$)，

人的眼睛看到的是左圖，而電腦看到的話是右圖。

原始影像

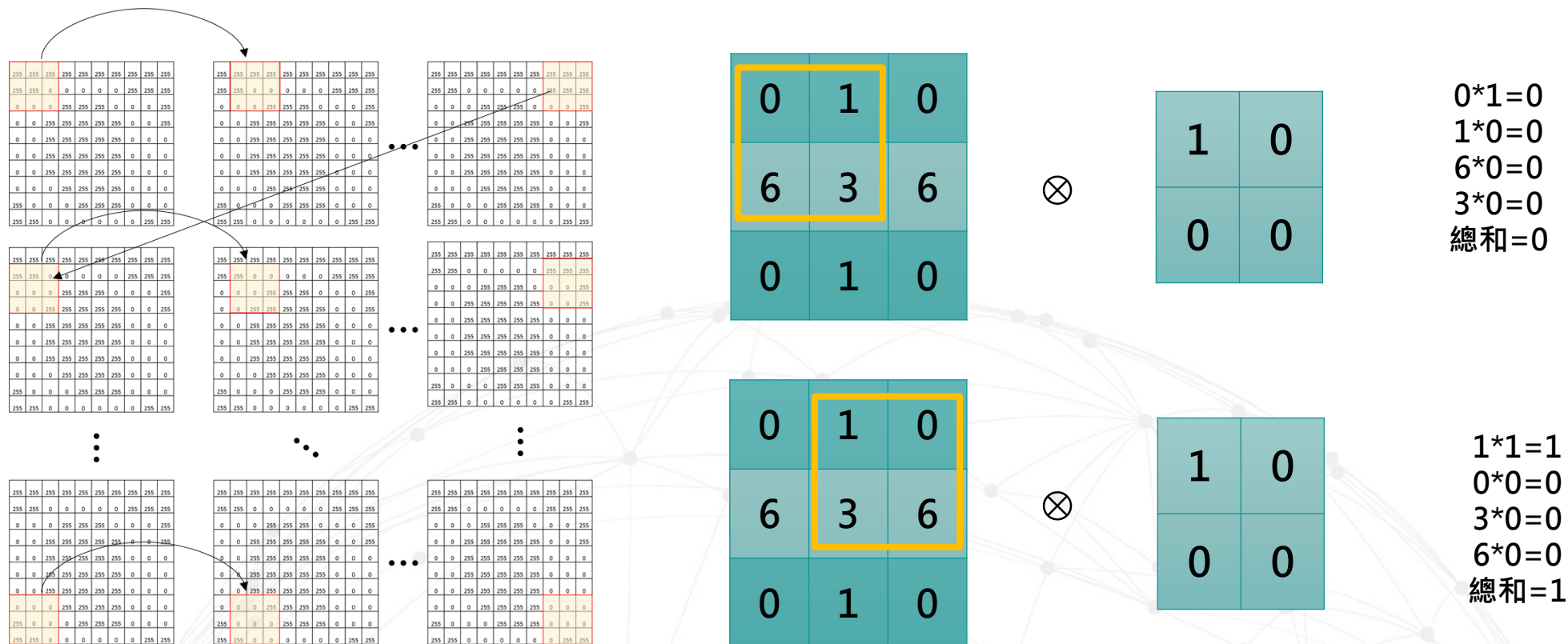


數位化的值

255	255	255	255	255	255	255	255	255	255
255	255	0	0	0	0	0	255	255	255
0	0	0	255	255	255	0	0	0	255
0	0	255	255	255	255	255	0	0	255
0	0	255	255	255	255	255	0	0	0
0	0	255	255	255	255	255	0	0	0
0	0	255	255	255	255	255	0	0	0
0	0	0	255	255	255	255	0	0	0
255	0	0	0	255	255	255	0	0	0
255	255	0	0	0	0	0	0	255	255

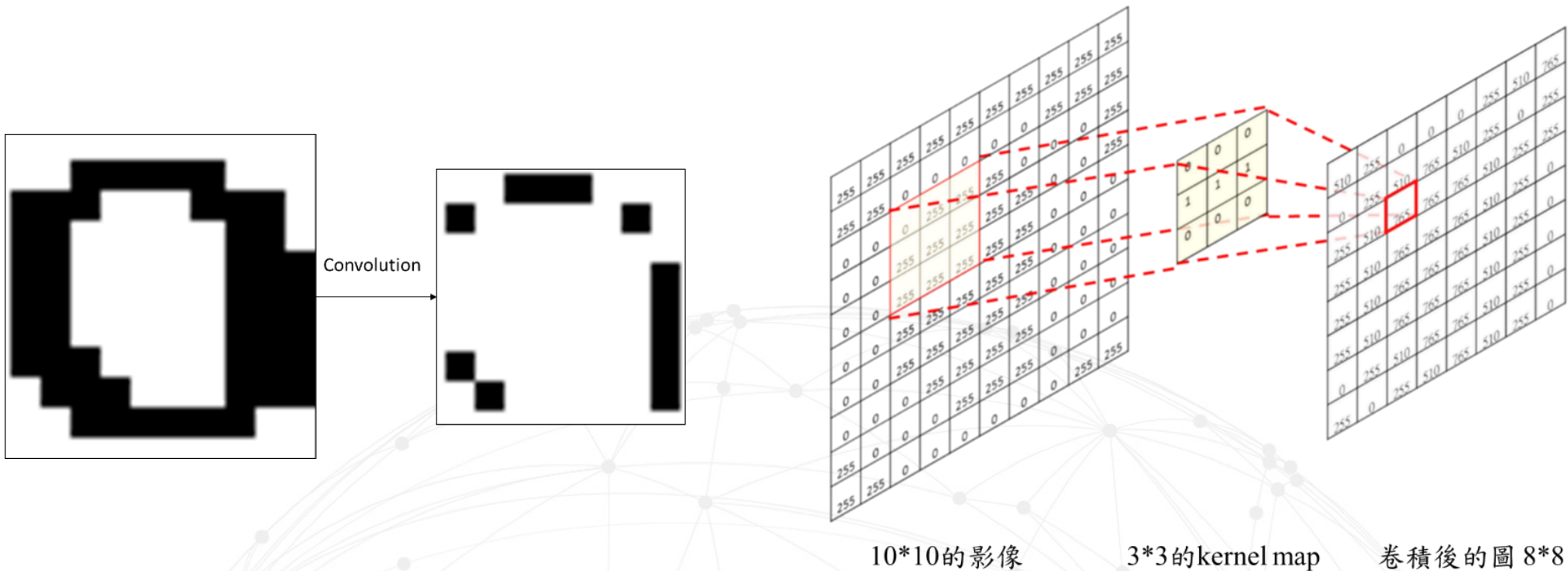
卷積層 Convolutional layer

以圖片辨識為例，CNN的input為圖片的像素(pixel)組成的矩陣，而filter是特徵矩陣，透過convolution，也就是feature mapping，可以將圖片的特徵利用filter捕捉起來。



卷積運算 Convolution

左邊的圖是原始影像，在經由中間的mask去做完卷積後，
可以得到右邊那張圖，圖的大小會變小，從原本的10x10變成8x8。



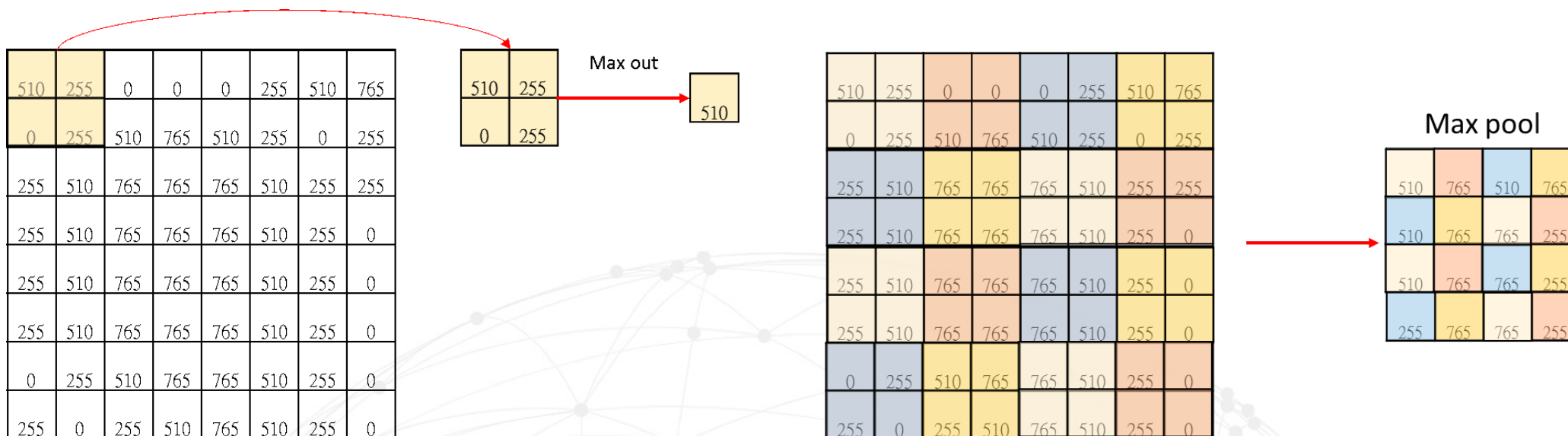
卷積運算的體驗

0~255的數值，呈現該點的數值越大，顏色越深



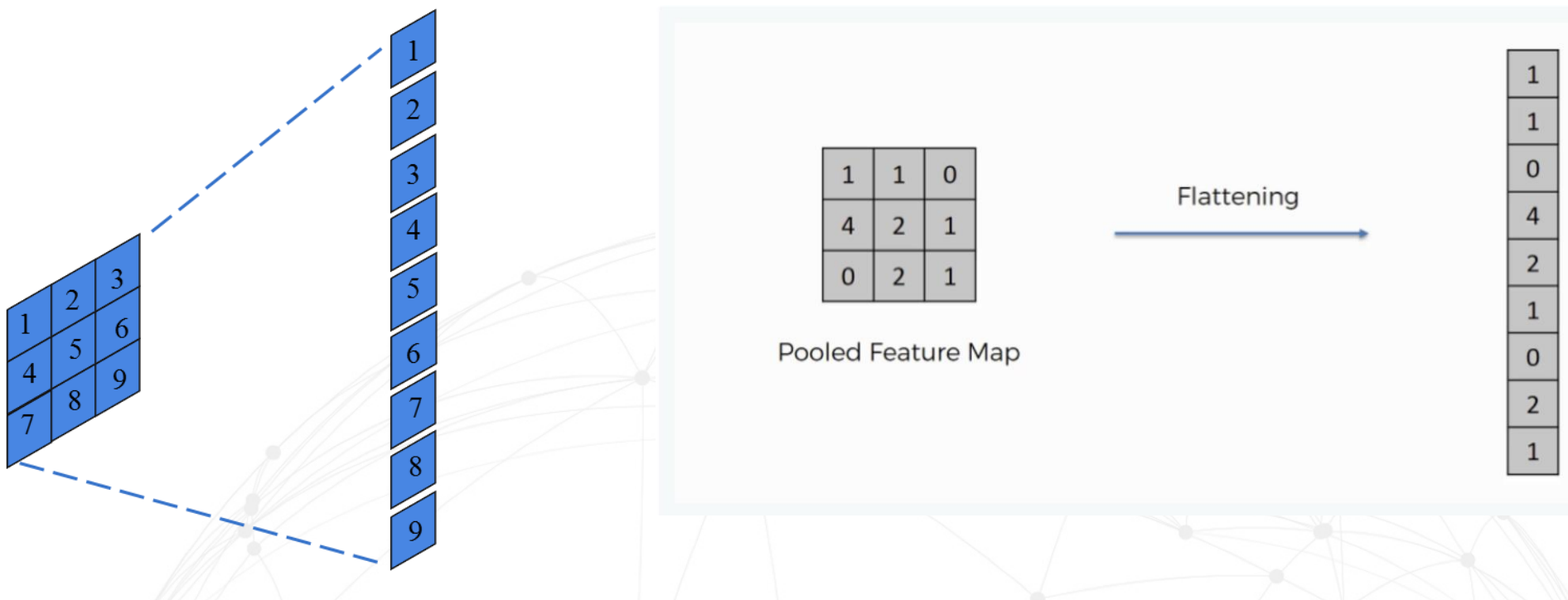
池化層 Pooling

池化目的只是在將圖片資料量減少並保留重要資訊的方法，把原本的資料做一個最大化或是平均化的降維計算。本文取一個2x2的最大池化法(max pooling)當作例子。



全連接層 Fully Connected layer

因為做完卷積運算和池化法後得到的特徵圖還是一個2-D的圖片，到全連接層前要先轉成1-D的陣列。(如果做完卷積或是池化後結構是1x1的feature map，此步驟可以省略)



基礎介紹-損失函數(loss function)

「實際值和預測值的殘差」

是用來評估模型的預測值與真實值不一致的程度，也是神經網絡中優化的目標函數，神經網絡訓練或者優化的過程就是最小化損失函數的過程，損失函數越小，說明模型的預測值就越接近真是值，模型的健壯性也就越好。

$$\text{loss/residual} = y - \hat{y}$$



模型評估-驗證指標(validation index)

分類指標(Classification metrics)

二元相關(二元混淆矩陣和相對應驗證指標、ROC曲線、AUC)和多元相關(多元混淆矩陣和相對應驗證指標)。

回歸指標(Regression metrics)

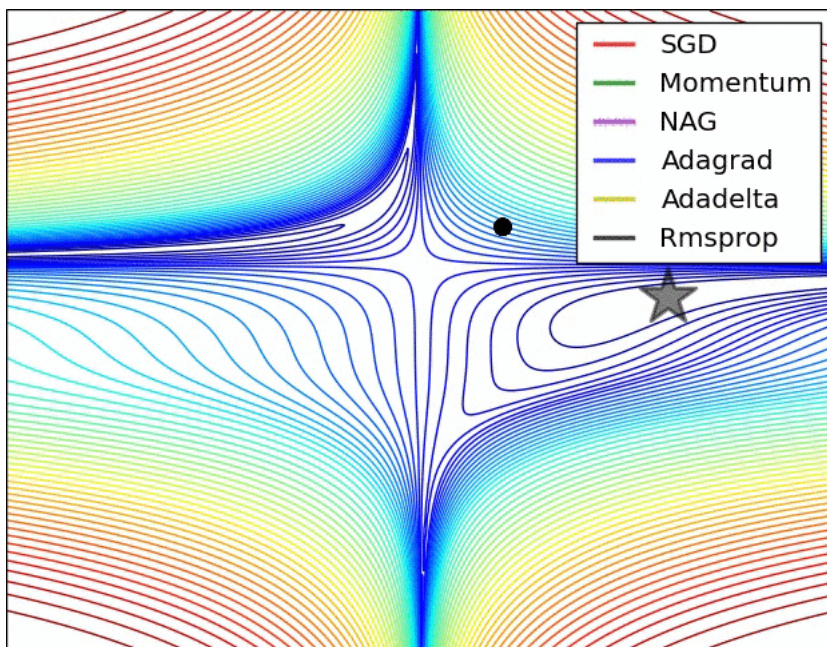
平均均方誤差(Mean Squared Error, MSE)、平均絕對誤差(Mean Absolute Error, MAE)和平均均方對數誤差(Mean Squared Logarithmic Error, MSLE)

		True Condition	
	Total Population (I)	Positive	Negative
Predicted outcome	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

混淆矩陣

深度學習優化器 optimizer 的選擇

機器學習算法中，大部分算法的本質就是建立優化模型，通過最優化方法，對目標函數進行優化從而訓練出最好的模型，而優化算法的功能，是通過改善訓練方式，來最小化 (或最大化) 損失函數。



在 keras 中也有 SGD , RMSprop ,
Adagrad , Adadelta , Adam 等....

<https://keras.io/optimizers/>