

AIATC 素養評測 考前閱讀資料 (AI Overview)



Taiwan AI Academy



「著作權聲明」

本備考教材由台灣人工智慧學校基金會精心編寫，僅供本會學生於學習與 AIATC 素養認證考試準備之用。

教材中所有內容，包括但不限於文字、圖像、表格等，均受著作權法保護，權利歸台灣人工智慧學校基金會所有。任何單位或個人不得以任何形式轉載、複製、分發、出版或作其他用途。違者必依法追究其法律責任。

目錄

1. AI 的歷程演進
2. 機器學習
3. 深度學習
4. 遷移學習
5. 雲端運算與邊緣運算
6. 生成式 AI
7. AI 輔助數位學生
8. 人工智慧安全
9. AIATC 素養認證測驗說明
10. GenAI 補充資料及自學資源



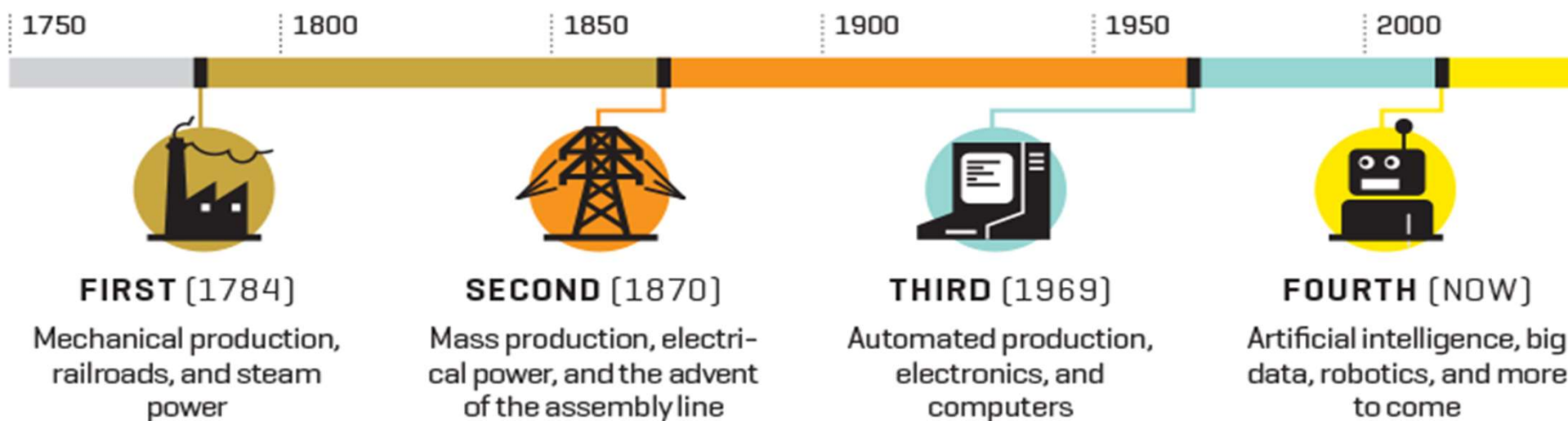


01 AI 的歷程演進

人類科技的歷史演進



THE FOUR INDUSTRIAL REVOLUTIONS



工業革命

人類發明了機器，
機器是能**規律運作**的設備，
用來取代人類(重複性)的勞力

科技革命

人類發明了電腦，電腦是照著
固定規則的軟體來運行，
用來取代人類(重複性)的腦力

AI時代

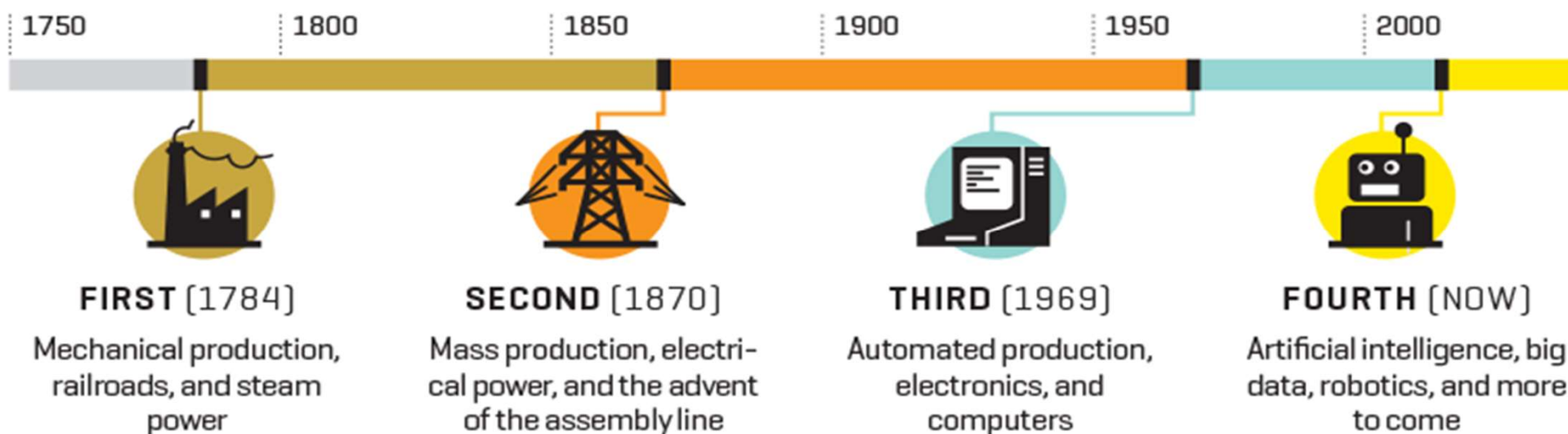
機器或電腦
運作規則是**透過數據**，
自我學習而得

source : <https://finance.yahoo.com/news/ceos-revolution-coming-140021597.html> 5

從規則驅動到資料驅動



THE FOUR INDUSTRIAL REVOLUTIONS



規則驅動

Rule-based

資料驅動

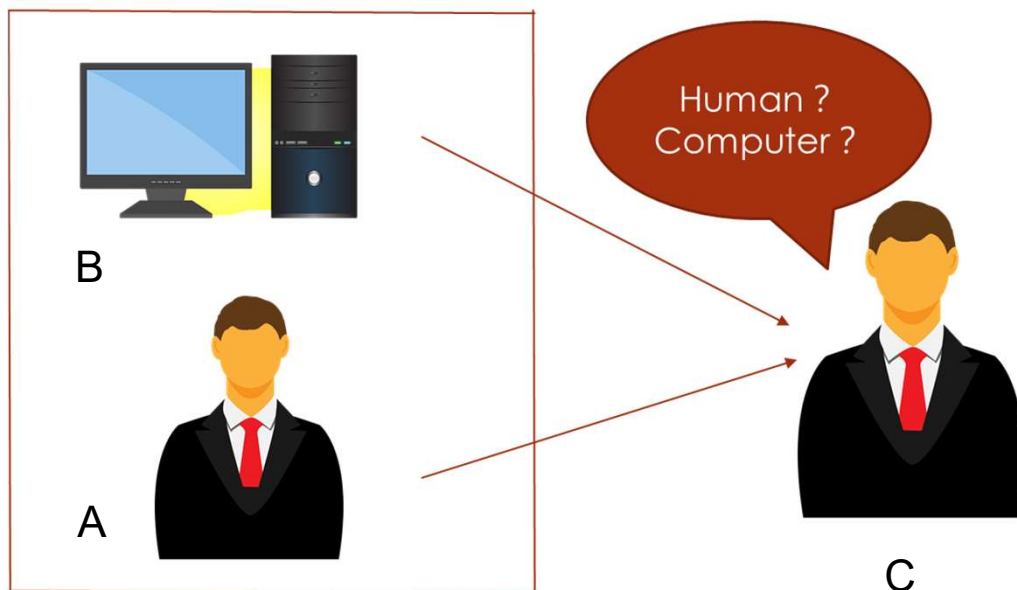
Data-driven

source : <https://finance.yahoo.com/news/ceos-revolution-coming-140021597.html>

機器是否會思考-圖靈測試(1950)



- 艾倫.圖靈 (A. M. Turing) 提出了一套評估「機器能否思考」
- 圖靈測試 (Turing Test)



圖靈測試的定義：

在C看不見A、B的狀況下，C詢問A、B一串問題，經若干問答後，C無法區別出A、B間機器或人的身份，則稱 B 通過「圖靈測

人工智慧的演進



資料來源：《人工智慧在台灣》，劉奕西整理

人工智慧的強弱之分



- 弱人工智慧 (Weak AI) :

- 也被稱為狹隘人工智慧 (Artificial Narrow Intelligence , ANI) , 意即在某些條件限制下可以表現得很有智慧。它們或許能在一件小事上表現得不錯, 但卻不能解決其他對人類而言相對容易的問題。

- 強人工智慧 (Strong AI) :

- 也稱為通用人工智慧 (Artificial General Intelligence , AGI) 。可以跟人一樣有很廣泛的智慧, 具備執行智慧行為的能力, 但缺乏心靈或自我意識。其中一種強人工智慧的判斷指標為「咖啡測試」, 現今強人工智慧仍是人工智慧的發展目標, 尚未能實現。

強人工智慧	弱人工智慧
一般的工作 (AGI) => Really can think	特定的工作 => Act as it can think



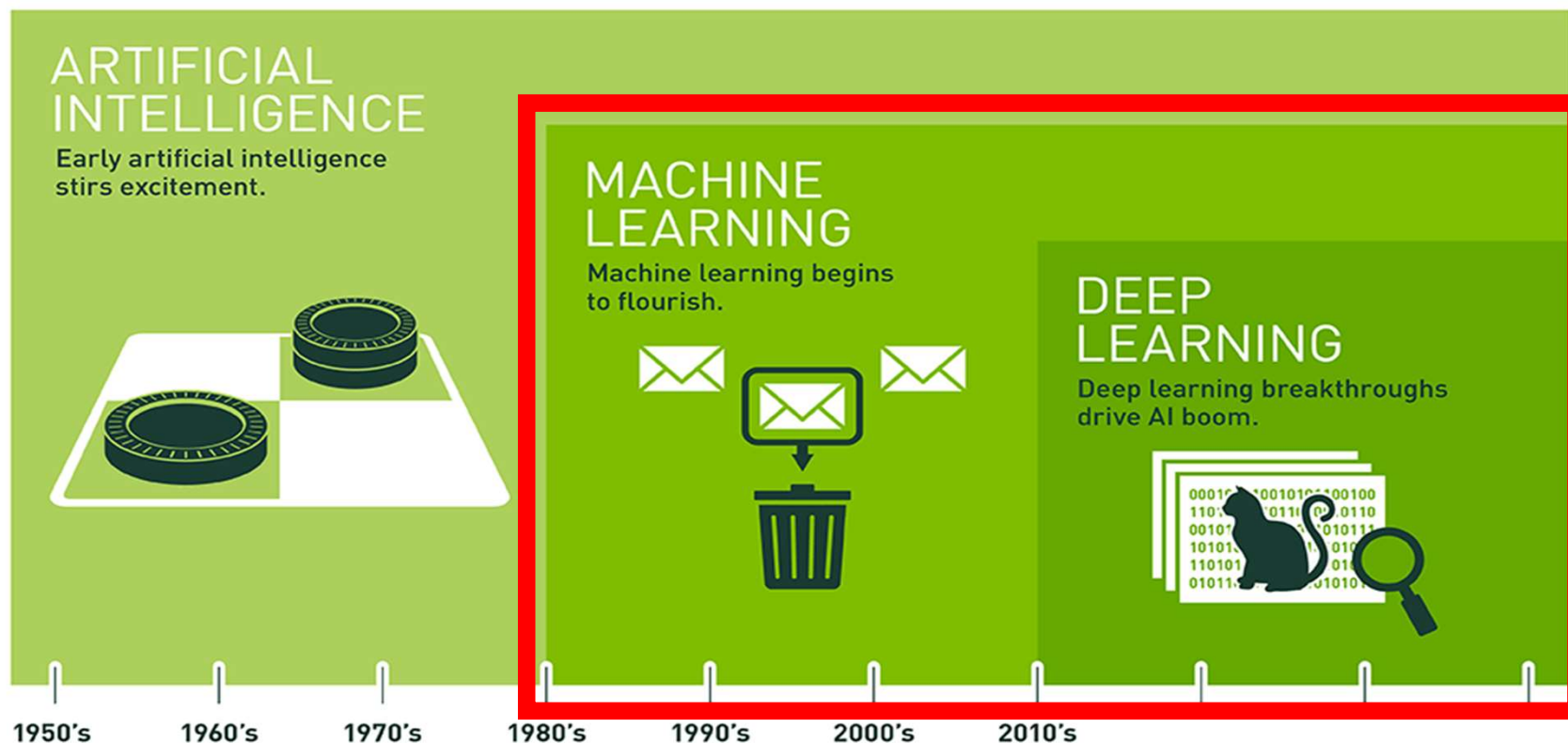
02 機器學習



Part 1

何謂機器學習

機器學習 – 人工智慧的一種技術



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

source: <https://blogs.nvidia.com.tw/2016/07/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

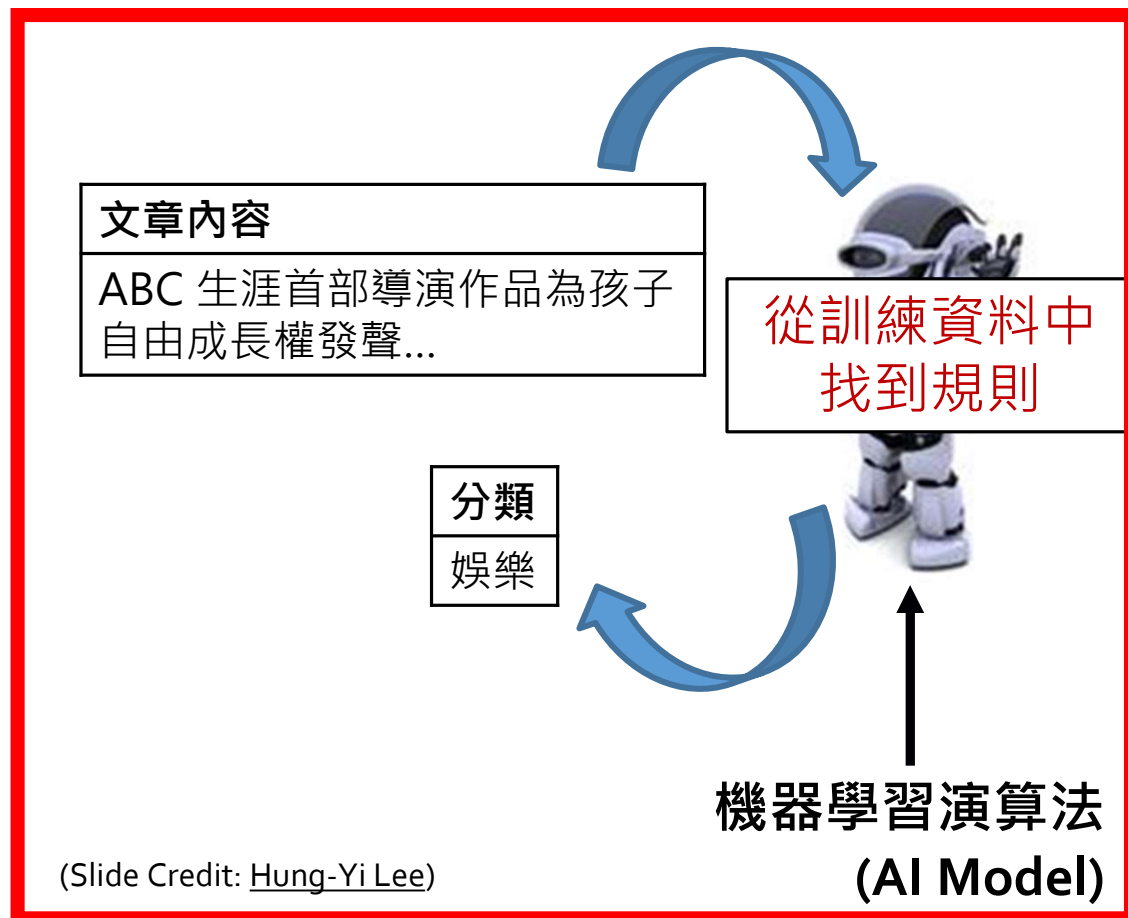
機器學習的定義



機器學習是一種
**「從過去資料學習規律，
進而預測未知資料」**
的方法與技術



讓機器自己從資料中學習規則



文章內容	分類
李維拉來了！美國職棒名人堂球星 傳奇守護神今抵台...	運動
美國聯準會主席鮑爾在國會暗示利率可能升得更高、維持更久...	財經
中央流行疫情指揮中心今日公布國內上週平均每日新增...	健康
...	...
摺疊手機去年出貨倍增 三星領軍群雄角逐...	科技

很多訓練資料



Part 2

機器學習的資料準備

機器學習的資料型態 - 結構化資料



- 結構化資料的資訊內容有精確定義的模式。若要簡單定義，意即所有可以透過表單系統（如 Google 試算表、Microsoft Excel）呈現出來的資料都是結構化資料。
- 資料可以透過行列式表格呈現出來。每一行都代表一種特殊的屬性，而每一列會個別列出與該屬性相關的資料。行與列組成了表格，因而可以輕鬆引用。
- 不同的表格可以互相連結，兩個表格之間同一列的資料可以互相關聯。

一個結構化資料的例子



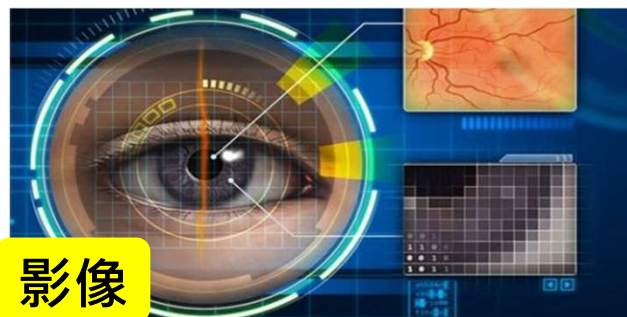
acqic	bacno	cano	conam	contp	csmcu	ecfg	etymd	flbm	flg_3dsmk	hcefg	insfg	iterm	locdt	loctm	mcc	mchno	ovrlt	scity	stocn	stscd	txkey	fraud_ind
6881	113261	38038	513.8	5	0	N	0	N	N	5	N	0	33	172652	457	59333	N	0	102	0	516056	0
0	134508	45775	465.62	5	0	N	2	N	N	0	N	0	9	105114	451	0	N	5817	102	0	4376	0
6881	15408	188528	513.8	5	0	N	0	N	N	5	N	0	6	152458	457	59333	N	0	102	0	483434	0
6716	157159	29967	1016.11	5	62	N	5	N	N	5	N	0	5	172946	247	50436	N	3281	102	0	1407164	0
5975	105985	81305	713.66	5	62	N	4	N	N	5	N	0	6	182129	263	93775	N	5817	102	0	1051004	0
0	78377	2295	465.62	5	0	N	2	N	N	0	N	0	6	104918	451	0	N	5817	102	0	2943	0
6411	94435	49219	1806.49	3	62	N	4	N	N	5	N	0	6	172624	339	0	N	5865	102	0	1622153	0
6769	112032	177989	526.88	6	62	N	2	N	N	5	N	0	7	34933	373	79200	N	5817	102	0	57795	0
6092	92294	85535	201.39	2	62	N	2	N	N	5	N	0	6	64652	264	8335	N	3585	102	0	836165	0
3288	88538	81033	1.38	5	62	Y	8	N	N	5	N	0	7	45457	337	20984	N	621	93	2	651056	1
0	16279	110755	465.62	5	0	N	2	N	N	0	N	0	10	104447	451	0	N	5817	102	0	4507	0
6882	49387	57083	0	5	0	N	0	N	N	0	N	0	5	210254	459	2467	N	0	102	0	611092	0
0	24677	4485	465.62	5	0	N	2	N	N	0	N	0	9	104728	451	0	N	5817	102	0	4143	0
6883	40691	47685	513.8	5	0	N	0	N	N	5	N	0	5	172346	457	10045	N	0	102	0	482472	0
6881	12981	178029	513.8	5	0	N	0	N	N	5	N	0	6	191233	457	86584	N	0	102	0	483717	0
6882	43457	161279	0	5	0	N	0	N	N	0	N	0	6	164857	459	2459	N	0	102	0	611116	0

結構化資料 (Table)

機器學習的資料型態 - 非結構化資料



- 所有不是結構化資料的資料都可以被歸類為非結構化資料，形式包括了文字、聲音、圖片、影像等



機器學習的資料型態 - 半結構化資料



- 半結構化資料是指既有某種結構但又不完全像傳統結構化資料那樣嚴格的資料。它介於結構化資料和非結構化資料之間，例如電子郵件包含結構化的標題部分（如發件人、收件人、時間等），但正文部分可以是非結構化的文本
- 其他如XML 文件、JSON 格式、Web 檔案等也是半結構化資料

Garbage in, garbage out



準備許多資料(Data)來教AI成長，同時資料的品質對機器學習的影響至關重要！

我們要準備許多數據(Data)來讓AI成長...，數據品質與AI表現習習相關！



數據決定了整件事情的上限，而AI演算法決定了下限。

資料收集、運用的倫理



- **資料去識別化**：將個人資料中可辨識個人身份的信息移除或修改的過程，以保護個人隱私。有助於減少資料被濫用的風險。
- **資料同意和透明度**：獲取使用者的明確同意來收集和使用他們的資料，並對如何使用這些資料進行全面的透明披露。
- **公平性和無偏見**：確保AI系統不會放大現有的社會偏見，並努力使算法決策公平，避免對特定群體產生不公平的影響。
- **資料安全**：採取必要的措施來保護存儲和傳輸中的資料，防止未授權的訪問和資料洩露。
- **負責任的資料分享**：分享資料時考慮到倫理和隱私問題，只有在有明確目的且合法律規範下進行。



Part 3

機器學習的型態與對應任務

機器學習的型態



資料**是否有標籤**(標準答案) →

- 監督式學習 (資料有標籤, Labelled Data)
- 非監督式學習 (資料無標籤, Unlabelled Data)
- 半監督式學習 (資料部分有標籤, 部分無標籤)

透過**獎懲機制**讓模型學習 →

- 強化式學習 (增強學習)
每次學習後給予正回饋或負回饋(獎勵和懲罰的機制), 讓模型進行試誤學習, 直到學會預期行為(如下圍棋、走迷宮)

機器學習的運作兩階段



訓練階段

(模型讀書階段)

大量的訓練資料 X



圖片來源：<https://twitter.com/teenybiscuit>

反覆迭代運算

從訓練資料中找出規則



$f(X)$

推論階段

(模型考試階段，對新資料預測)

訓練集外的未知資料



“Dog”

一個監督式學習的例子



1. 模型訓練階段:

- 先讓AI模型看過大量的圖片
- 每張圖片都有**標籤(Label)**:
狗或是瑪芬，兩個類別

2. 模型推論階段:

- 給一張新的圖片，讓AI模型預測
這張圖是狗還是瑪芬



Karen Zack/@teenybiscuit

監督式學習可以處理的任務



- **分類任務 (Classification)** (讓模型預測新資料屬於哪一個類別) :
 - 個人資料：男/女、已婚/未婚
 - 醫療影像診斷：癌症第0期、第1期、第2期、第3期、第4期
 - 信用卡交易判斷：正常交易、被盜刷
- **迴歸任務 (Regression)** (讓模型預測一個連續型的數值) :
 - 個人資料：身高、體重
 - 明日氣溫預測：30度、35度、40度
 - 氣體濃度預測：二氧化碳濃度 5%、10%、15%

一個非監督式學習的例子



1. 模型訓練階段:

- 讓AI模型看過大量的圖片
- 但是每張圖片**沒有**任何的標籤
- AI模型將**特徵接近**的圖片分在同一群

2. 模型推論階段:

- 給一張新的圖片，讓AI模型預測這張圖屬於哪一群



非監督式學習常見任務



- 分群任務 (Clustering) (讓模型預測新資料屬於哪一個分群) :
 - 根據資料的**特徵**來分群
 - 資料庫裡面找尋相似的照片/影片
- 降維任務 (Dimension Reduction) :
 - 因為特徵太龐大，需要整合以方便電腦計算
 - 常用方法: **PCA主成分分析**、**T-SNE**

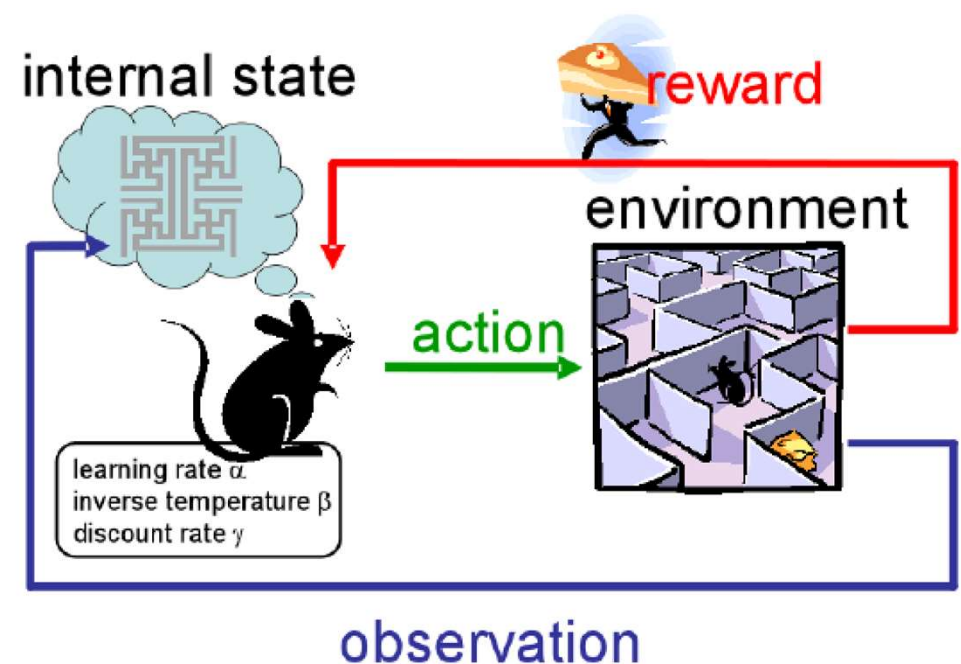
表格化資料	特徵1	特徵2	...	特徵M	答案
第1筆資料					標記1
第2筆資料					標記2
...					...
第N筆資料					標記N



非監督式學習，
資料**不需要**標籤

強化學習處理的任務

- 讓AI模型學習**長期的行為目標** (非即時的答案)，例如下圍棋、玩遊戲
- 與環境互動：觀察然後行動，最後達到最好的結果
- 適合的問題：環境能有**明確的反饋**且能進行大量試驗 / 模擬



source: <https://becominghuman.ai/the-very-basics-of-reinforcement-learning-154f28a79071>

監督式學習 v.s. 非監督式學習



機器學習 - 從數學的角度來看



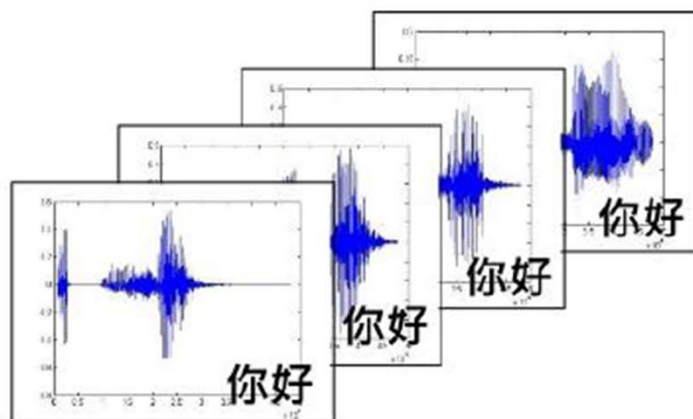
給定輸入 X

讓機器透過學習(訓練)得到函式

$f(X)$

即為 AI 模型

使它能够產生我們要的結果 Y



讓機器能從訓練資料
萃取出規則的演算法

問題與機器學習任務的對應



X

$f(X)$

Y

X光照片	→	$f_1(X)$	→	是否得癌症	→	分類 (Classification)	(預測離散的類別)
金融新聞	→	$f_2(X)$	→	股票指數	→	迴歸 (Regression)	(預測連續的數值)
各式照片	→	$f_3(X)$	→	風格相似的分到同一群	→	分群 (Clustering)	(依照特徵相似度分群)

常見的機器學習演算法 (AI 模型)



迴歸任務 (Regression)

線性迴歸

多項式迴歸

決策樹迴歸

KNN 迴歸

SVM 迴歸

分類任務 (Classification)

邏輯迴歸 (二分類)

KNN

SVM (支持向量機)

決策樹

分群任務 (Cluster)

K-means

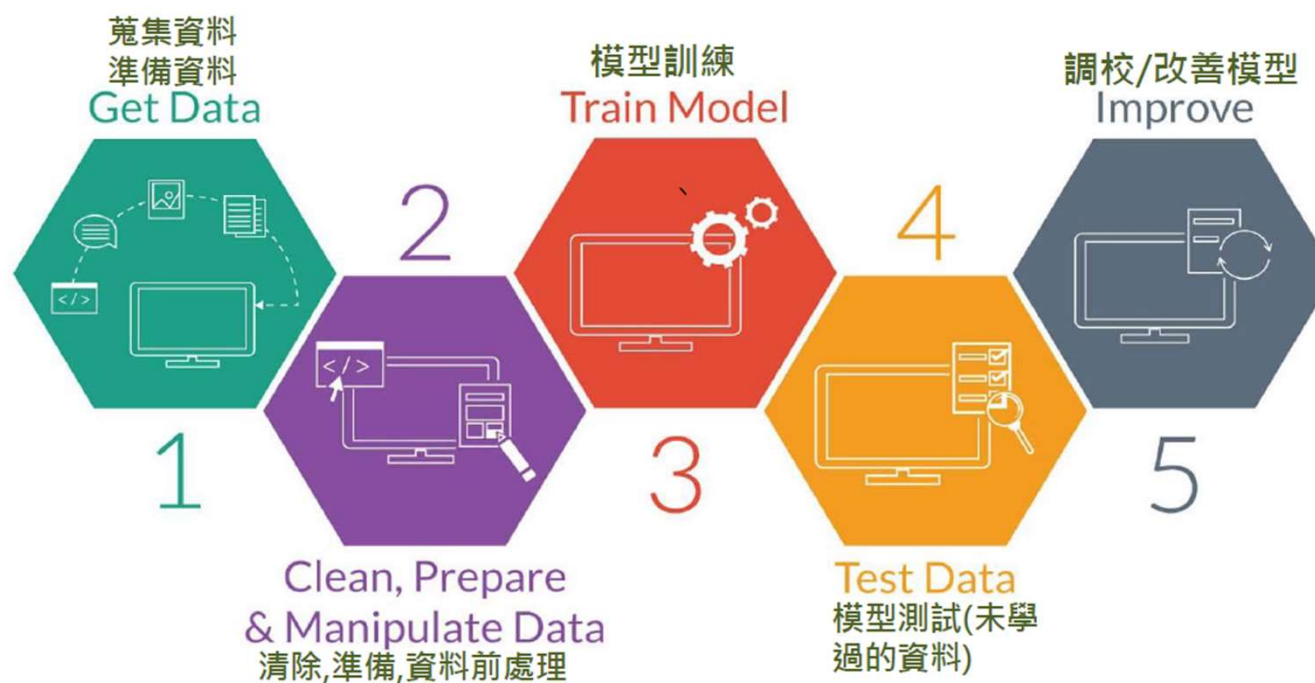
DBSCAN

機器學習的步驟



8 整個機器學習的步驟，第1~2步資料收集、前整理通常是最耗時的
收集到資料後，要依據選用的AI模型特性進行適當的資料處理

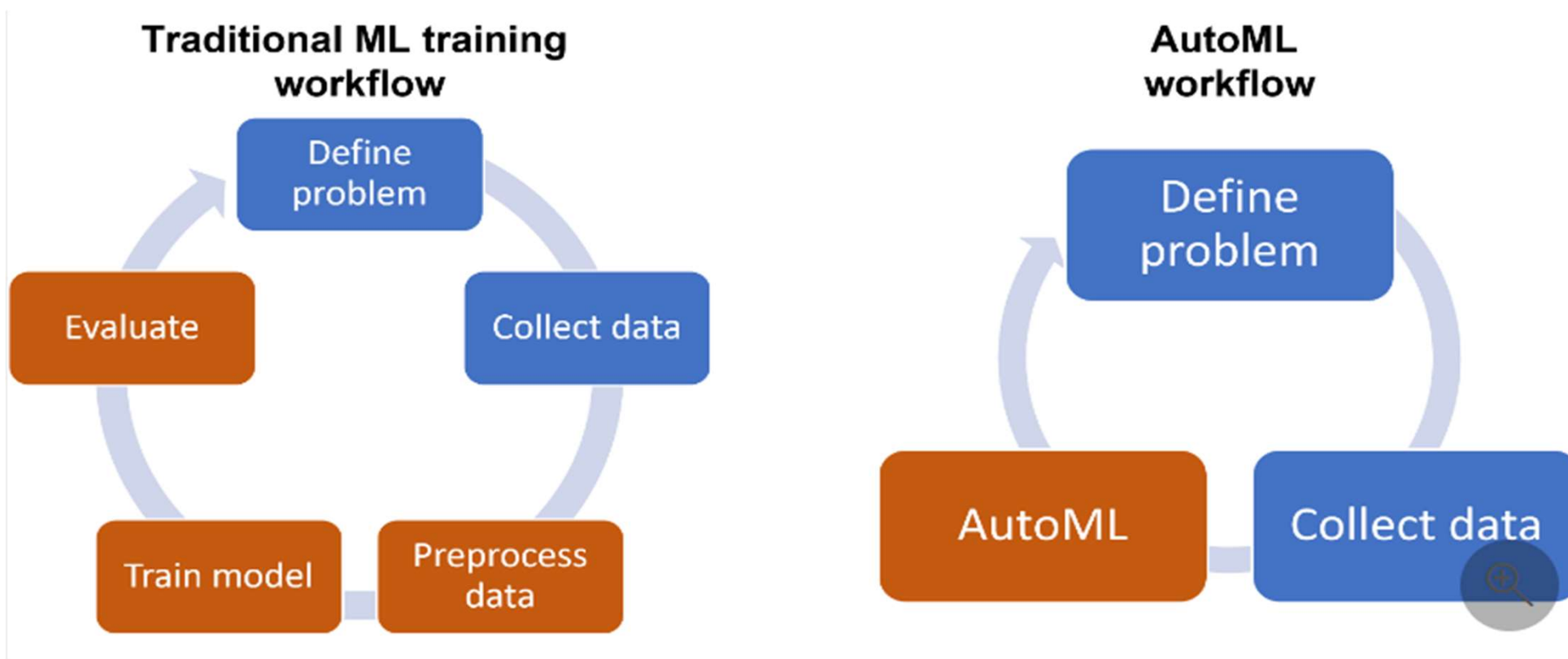
取得資料 → 資料整理/清除/編碼 → 訓練模型 → 測試模型 → 改善模型



自動化機器學習 (Auto ML)



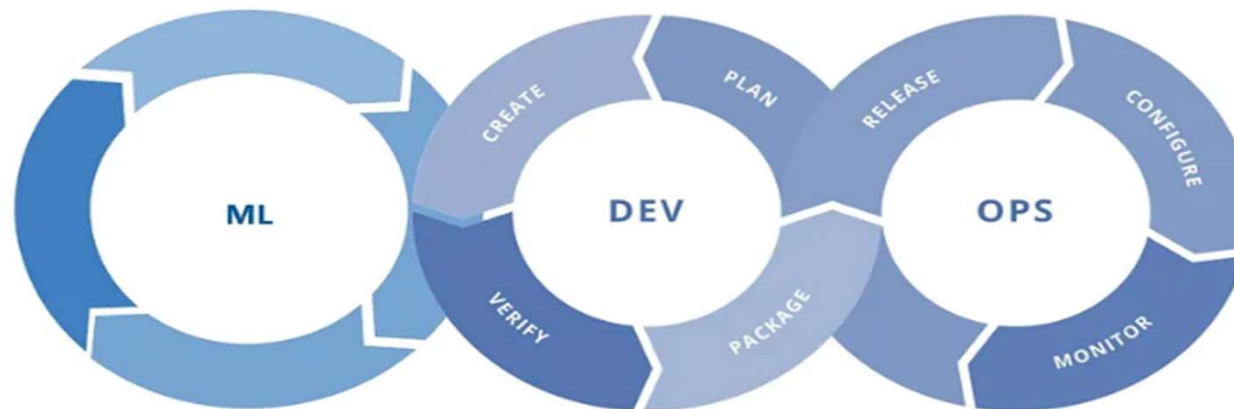
- 資料前處理、訓練模型和評估模型是實驗性和反覆性的程序，AutoML 可協助自動化這些步驟



機器學習營運 (Mlops)



Machine Learning + DEV+OPS 三個部分的縮寫合併
也就是把軟體專案中機器學習模型的開發、軟體開發、與系統持續維運的
整個生命週期串接在一起



Experiment

獲取資料
理解業務
模型初始設計

Develop

模型 + 測試
持續整合
持續部屬

Operate

持續交付
資料品質回饋
系統與模型監控

(綠字：機器學習 藍字：DevOps)

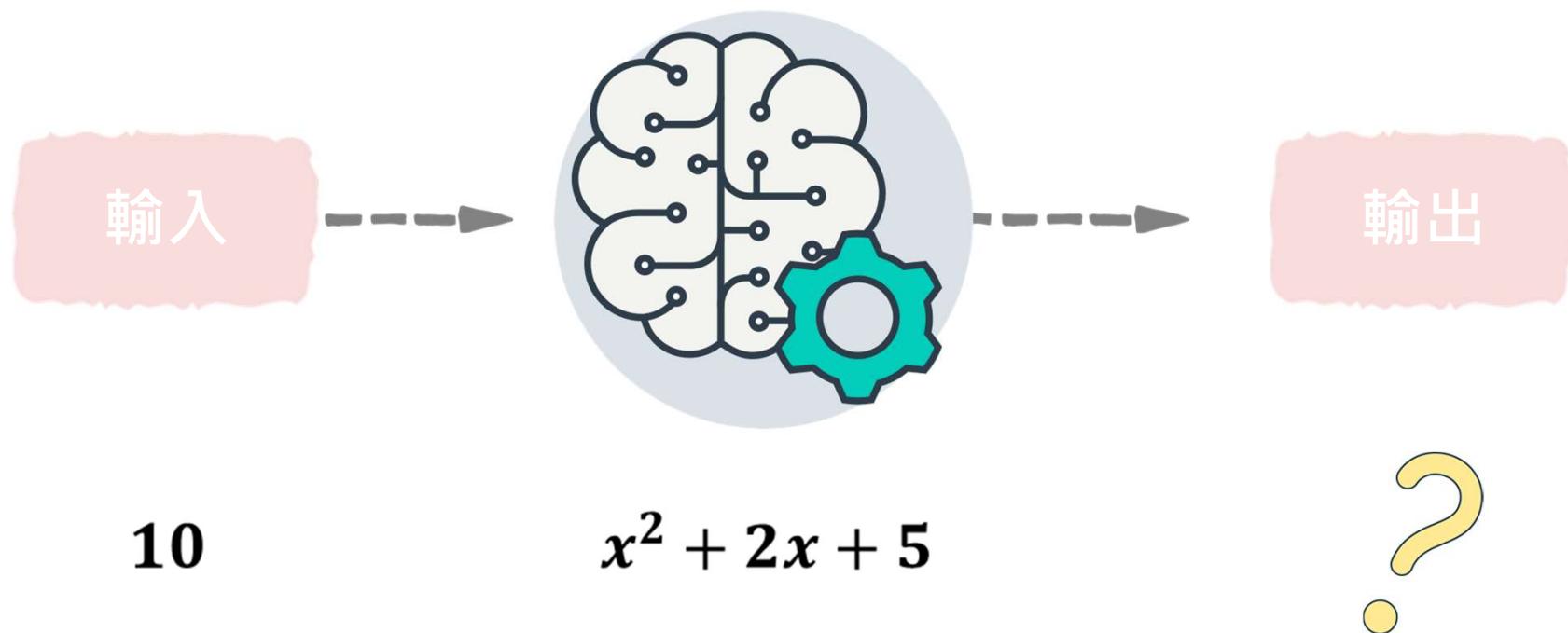
Source



Part 4

AI模型的訓練

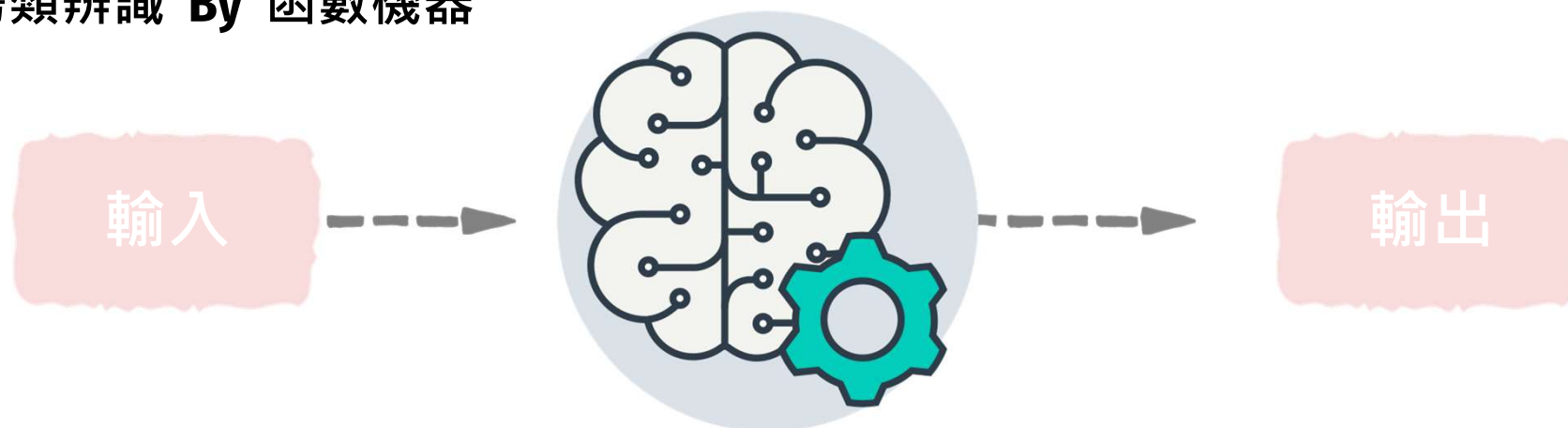
將觀察到的特徵數值輸入函數，輸出預測結果



建立函數進行分類預測



❖ 鳥類辨識 By 函數機器



$$f(\text{Bird Pic}) = \text{Bird's Name}$$

麻雀

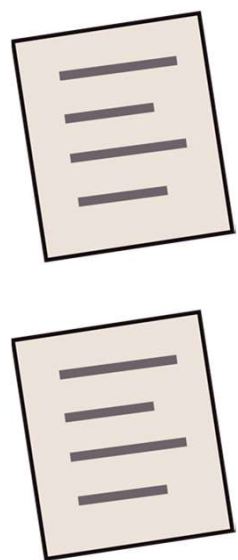
如何訓練一個好的函數 (AI模型)



● AI 模型的學習比照人類讀書考試：



測試函數機器的能力



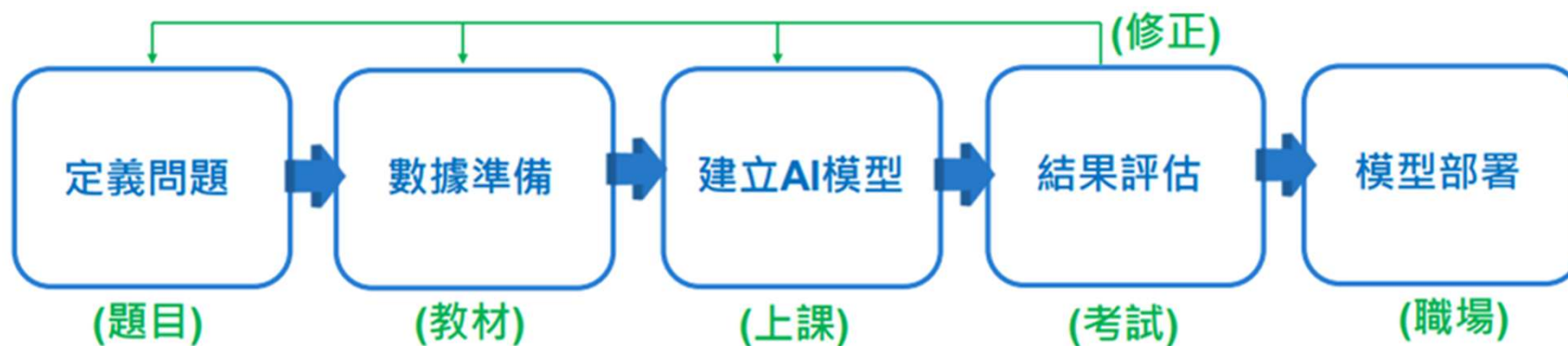
92

86

複習結束！

拿新的考題測試

訓練AI模型就像是教小朋友



機器學習的運作兩階段



訓練階段

(模型讀書階段)

大量的訓練資料 X



圖片來源：<https://twitter.com/teenybiscuit>

從訓練資料中找出規則

反覆迭代運算



$f(X)$

推論階段

(模型考試階段)

訓練集外的未知資料



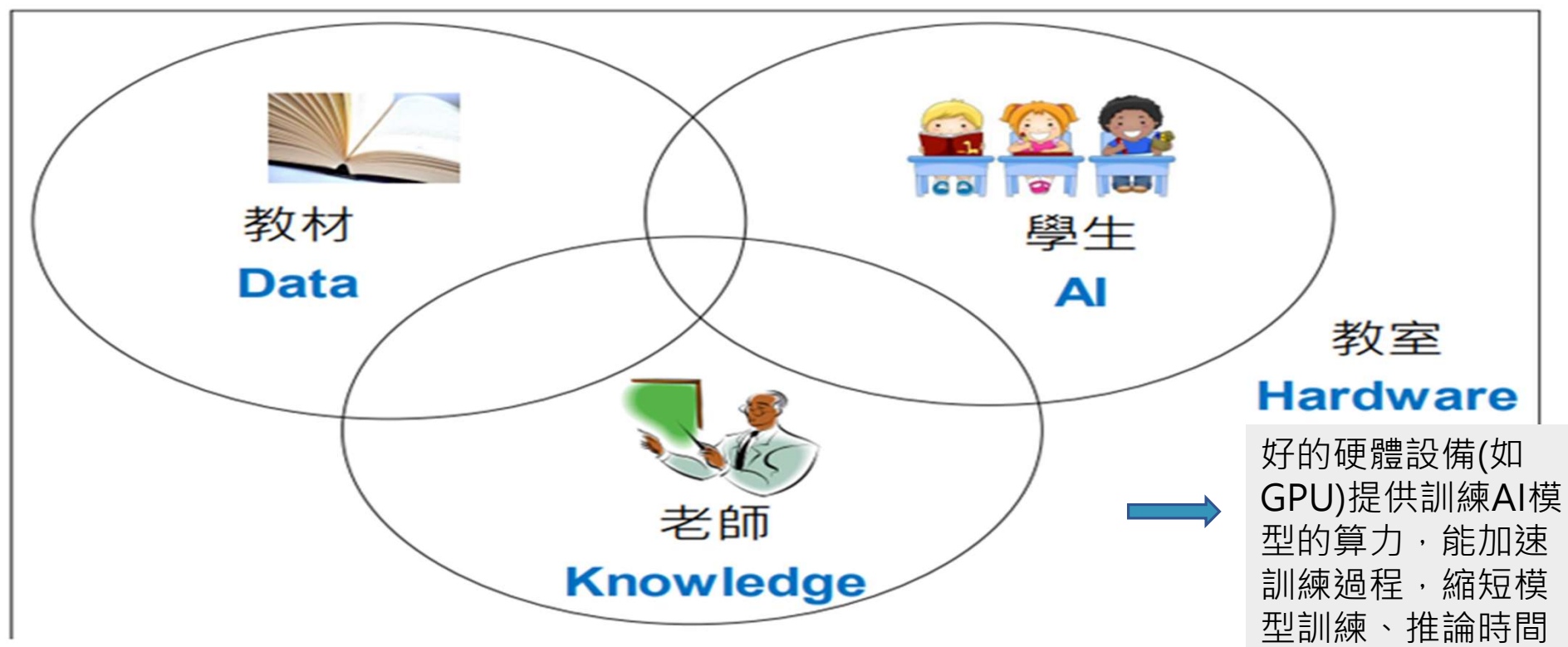
“Dog”

AI 模型學習的理解方式

以監督式學習為例，每次模型預測的結果會跟真實答案比對計算誤差，模型透過誤差來學習，讓下一次的預測能更接近正確答案



訓練AI模型的關鍵要素





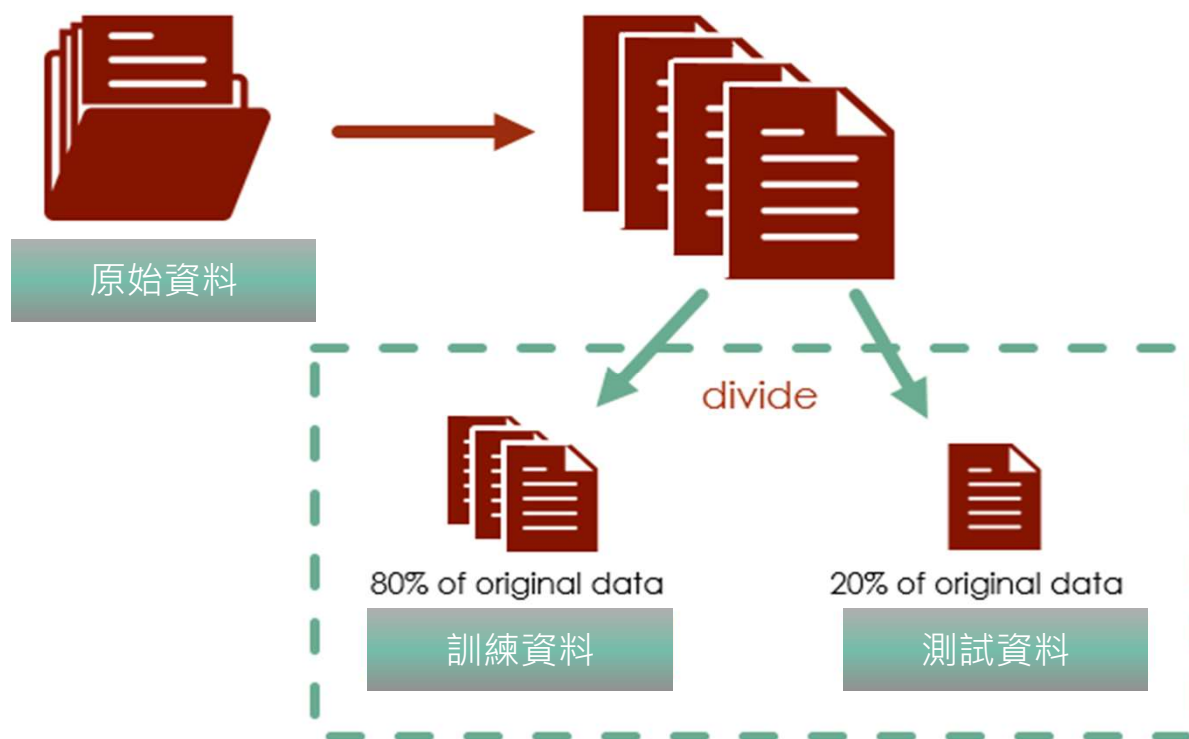
Part 5

訓練模型的資料切分

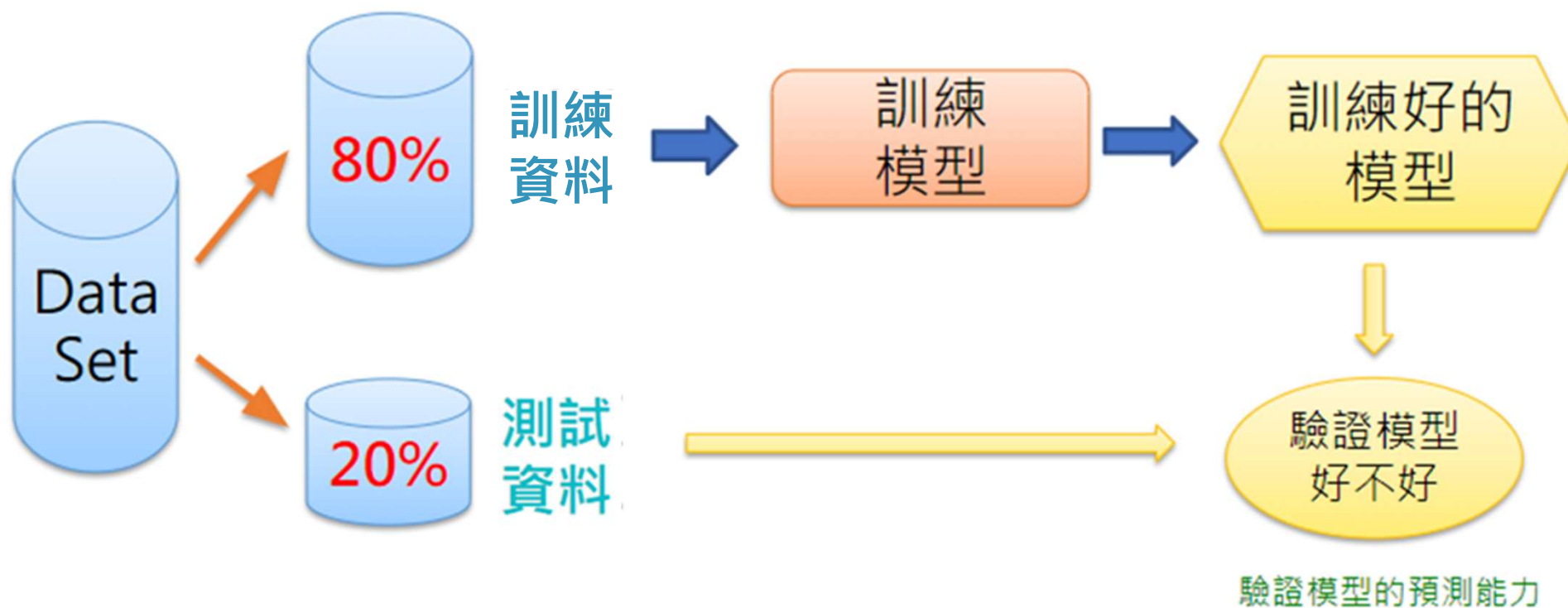
資料集分割（訓練集、測試集）



- 通常會將原始資料切割成「**訓練資料**」與「**測試資料**」，訓練資料，是用來教AI模型的，待訓練完後，再以模型沒看過的測試資料，測試模型的預測準確度，如下圖：



資料集分割（訓練集、測試集）

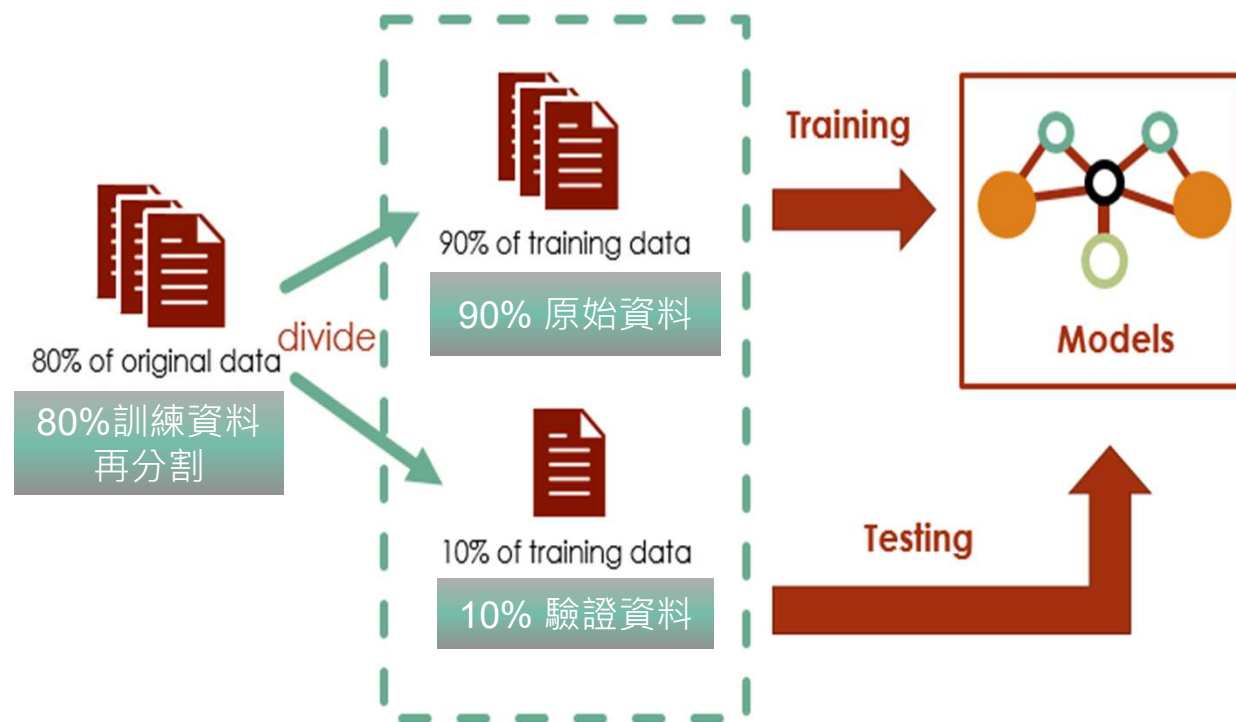


資料集分割（訓練集、驗證集、測試集）

- 在訓練過程中仍需不斷去驗證模型，有時也會再從訓練資料中切割少數資料作為「**驗證資料**」，如下圖所示：

訓練完模型之後先使用**驗證集**資料來自我驗證，確定模型的有效性之後再拿**測試集**測試。

有點像是讓模型先參加**模擬考**，再去考**大考**的概念



常見模型評估指標



迴歸任務

- 均方誤差(MSE)
- 平均絕對誤差(MAE)
- 判定係數 (R平方值)

分類任務

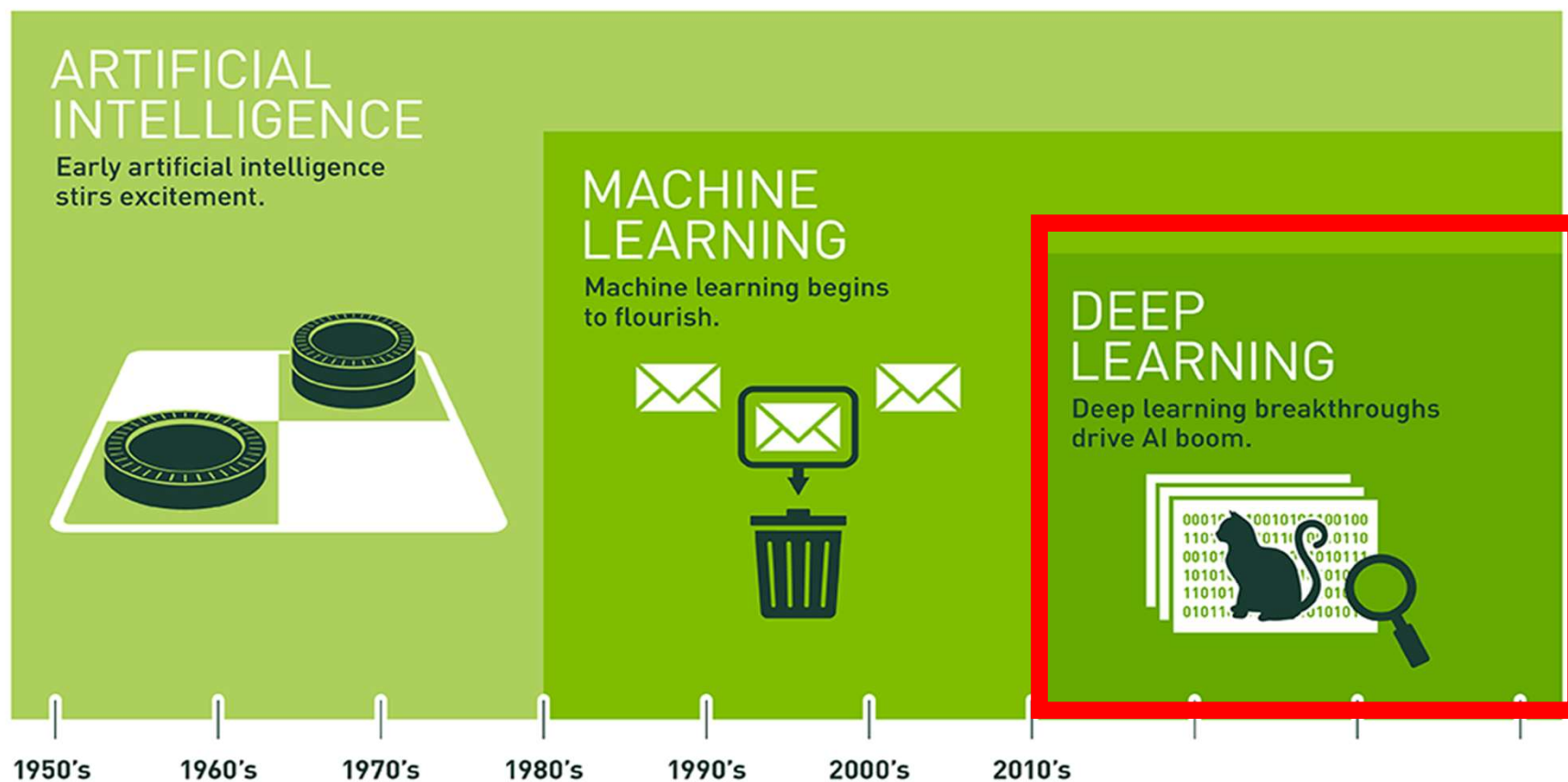
- 分類準確率(Accuracy)
- 混淆矩陣(Confusion Matrix)
- 精確率 (Precision Rate)
- 召回率 (Recall Rate)
- AUC曲線

p.s. 能將任務與指標對應即可



03 深度學習

深度學習 - 機器學習的子集



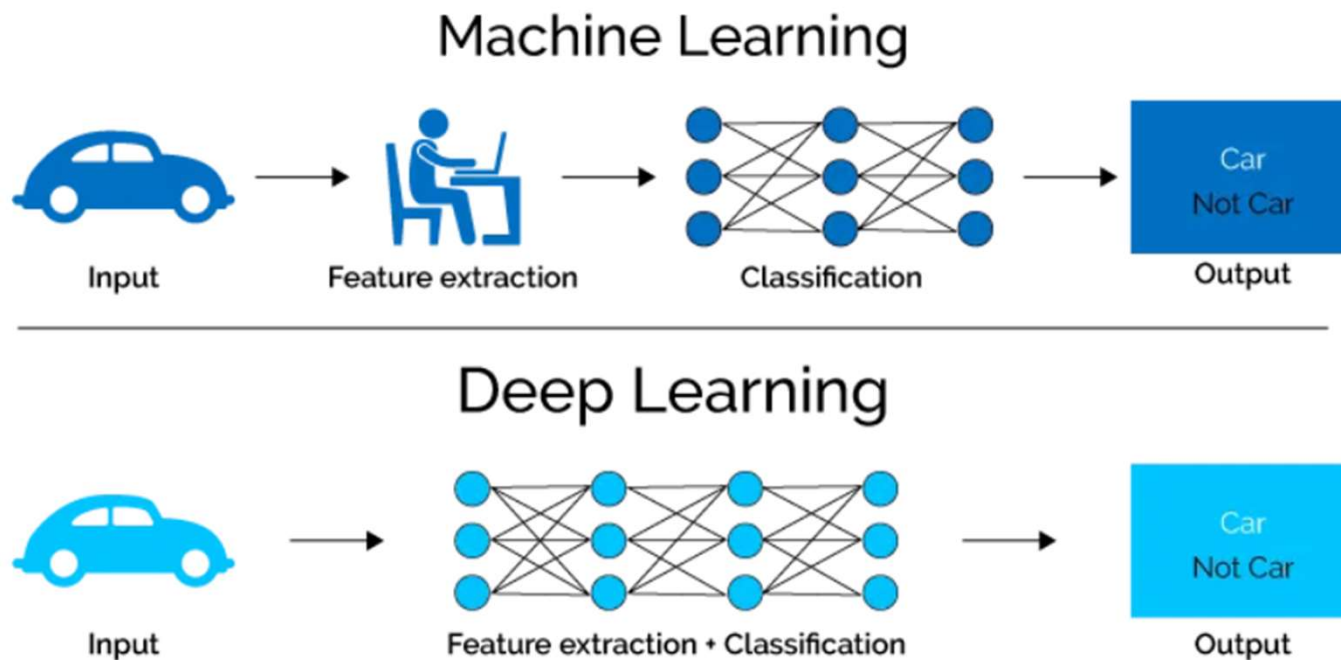
source : <https://blogs.nvidia.com.tw/2016/07/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/> 52

深度學習 V.S.機器學習



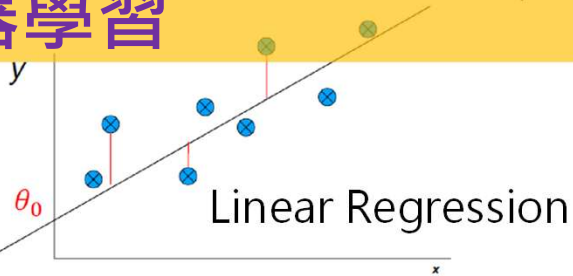
跟傳統機器學習的最主要差異在於：

深度學習可自動進行資料特徵萃取，不像機器學習需要人為制定特徵



(傳統)機器學習

Slope = θ_1

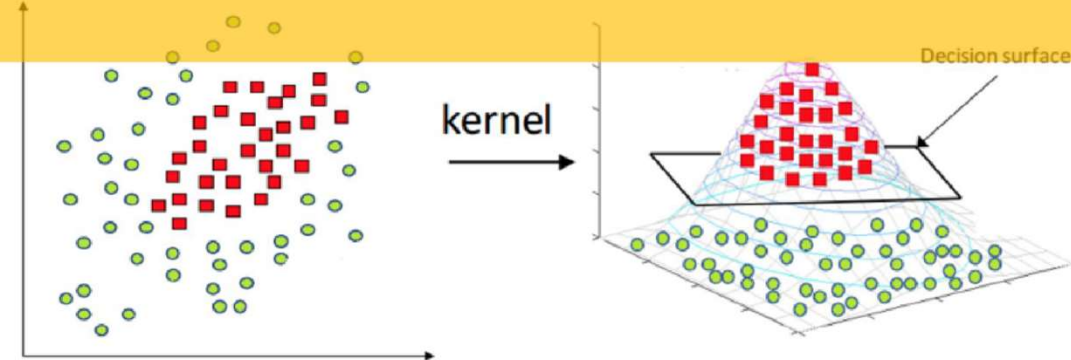


Logistic Regression

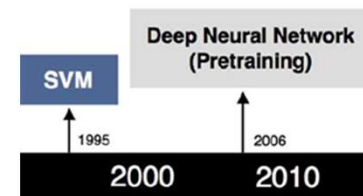
Linear function: $y = \theta_0 + \theta_1 x$

Logistic function:
 $y = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x)}}$

Regression-based

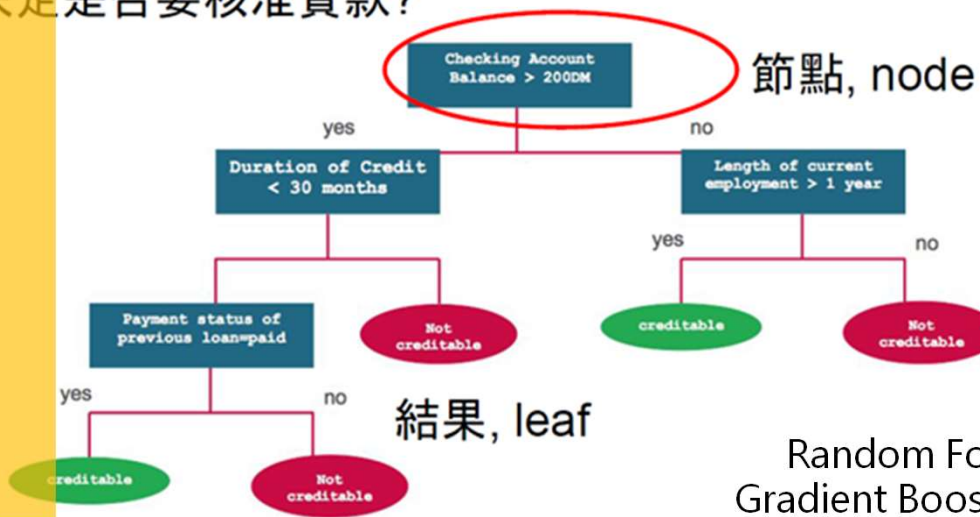


Support Vector Machine



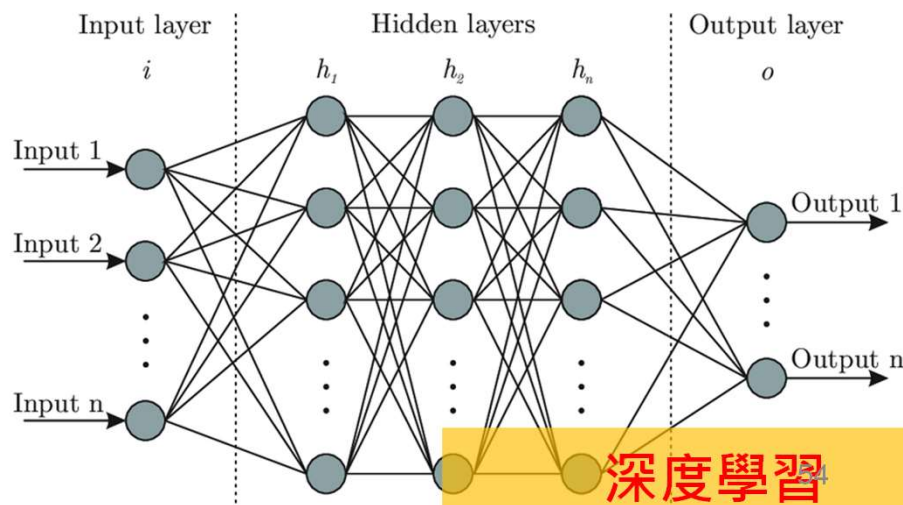
Decision Tree-

決定是否要核准貸款?



Random Forest
Gradient Boosting

Neural Network Family



https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051

深度學習大爆發的背後

- 理論齊備 : DNN, CNN, RNN ...
 - DNN: Gradient Descent and Backpropagation
 - CNN: Convolution and NN-ized
- 硬體能力提升 : Nvidia GPU + CUDA
- 巨量資料 : ImageNet and ILSVRC



Geoffrey Hinton



Yann LeCun



黃仁勳
Nvidia



李飛飛
Stanford



Part 1

神經網路基本架構

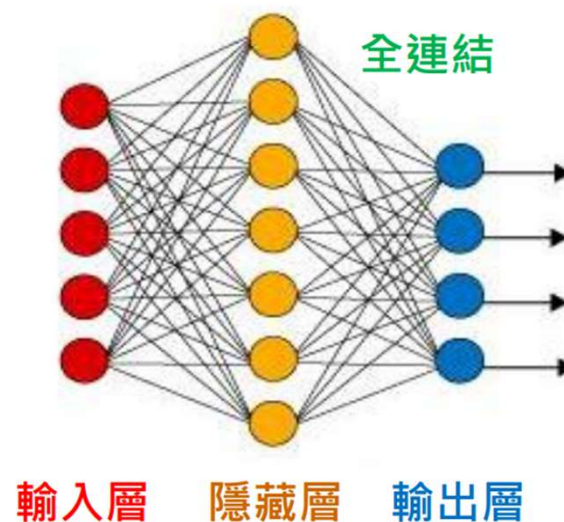
神經網路架構



概念源自生物神經網路，希望透過模仿生物神經網路的運作方式，讓電腦具備學習及記憶的能力，對新舊事物產生連結，進而做出推理判斷並解決問題



人腦有多個如上圖的神經元，
每個神經元都有受器和動器



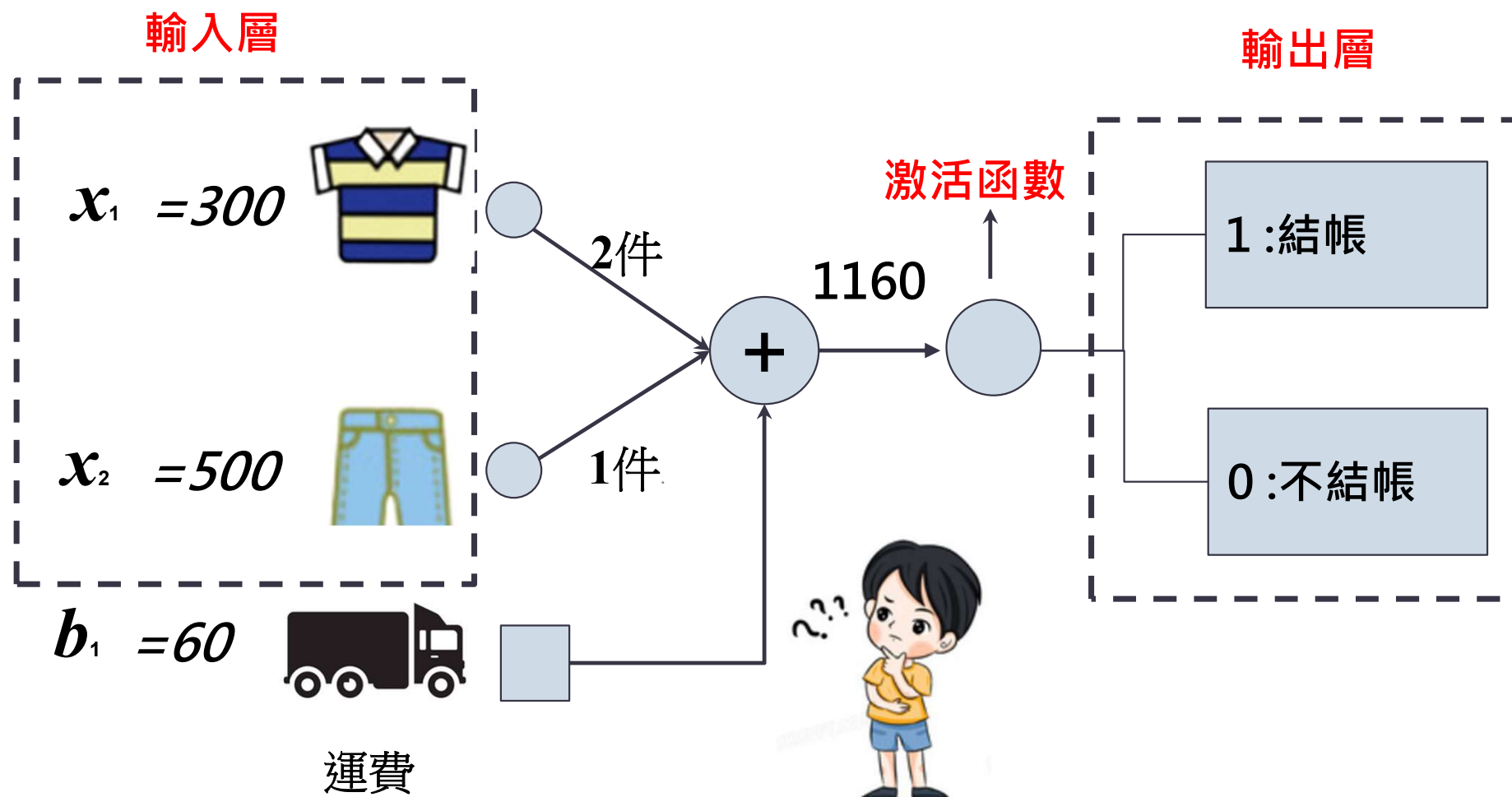
深度學習的神經網路，
每個圈圈都在模擬一個神經元

神經網路基本要件

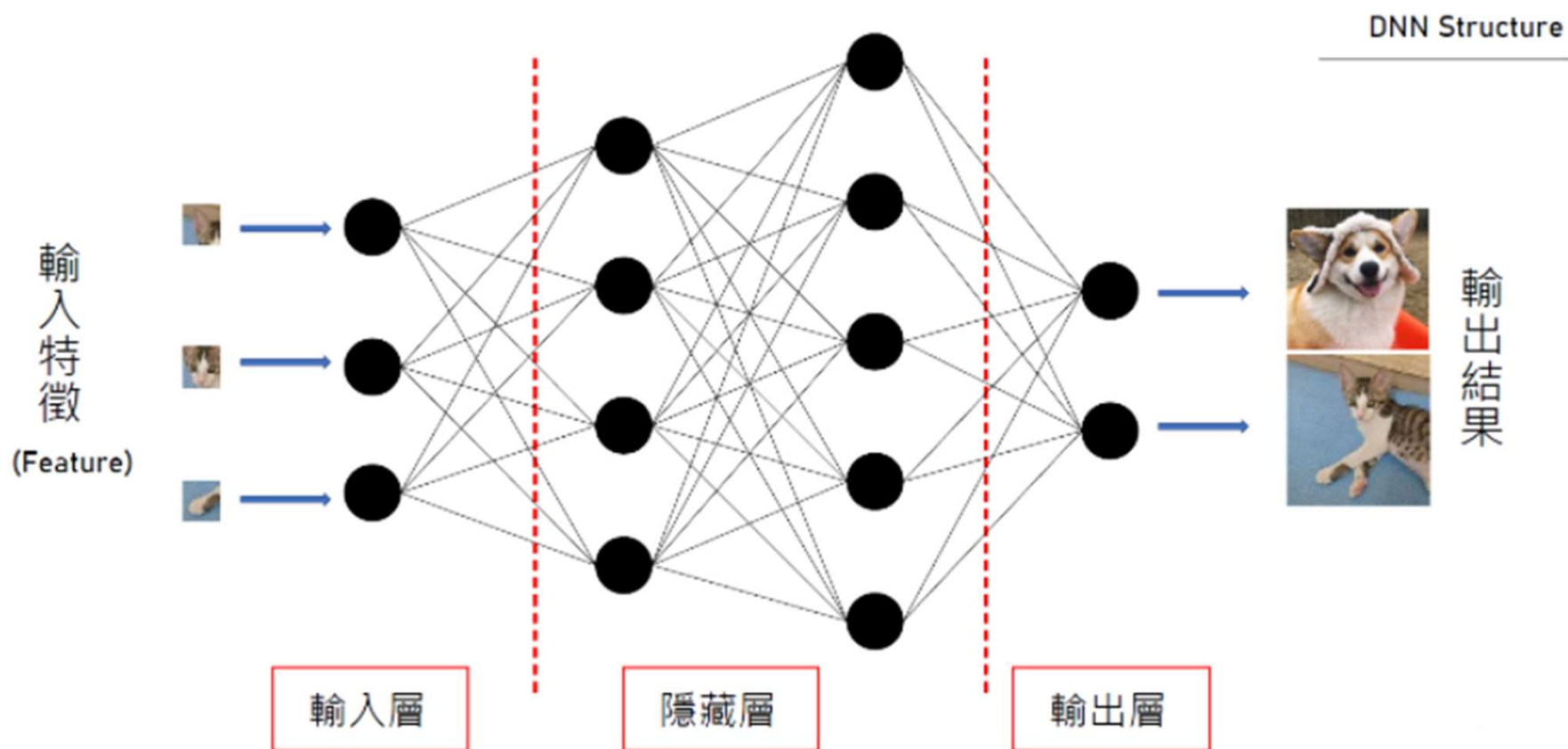


- **輸入層** -接受刺激的神經元，如同神經系統中的受器一樣，不同的輸入會觸發不同的神經元，受到刺激的神經元會將訊息往後傳遞下去。
- **隱藏層** -夾在輸入層與輸出層中間，不跟外部有直接的接觸，就像生物的中樞神經系統一般，主要的功能是對所接收到的資料進行處理。這個階段會對資料進行某些形式的轉換，並整理得到的資料訊息，再將得到的結果往後傳遞。
- **輸出層** -像是神經系統的動器(Acuator)，在接收到傳遞的訊息後特定的神經元會做出反應，其中反應訊號最強的神經元代表的項目就是這些資料辨識的結果。

簡易神經網路運作示例



深度神經網路-三層以上 (Deep Neural Network)

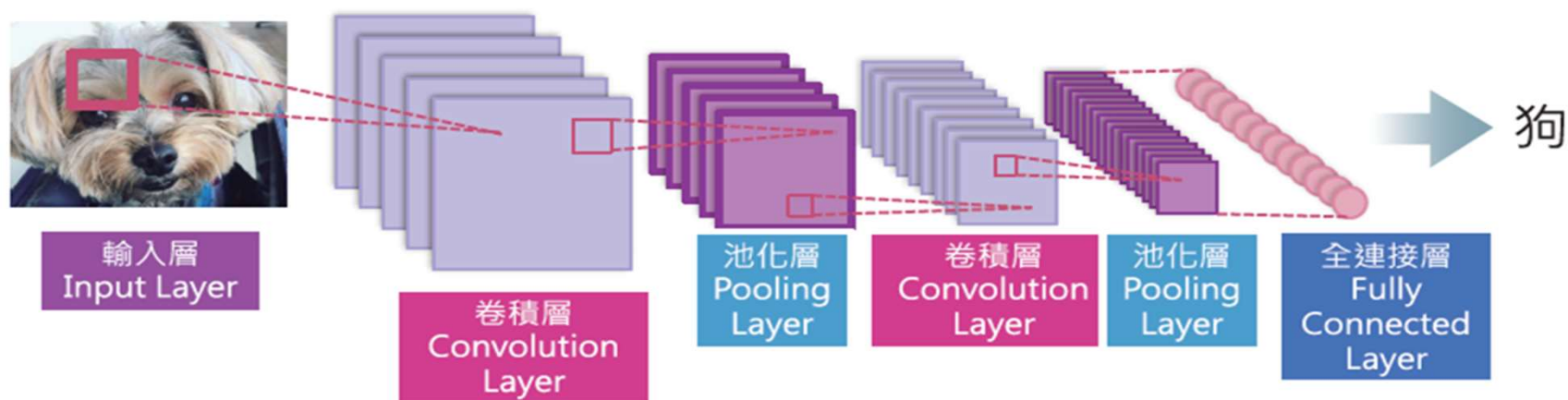


『深』代表有許多隱藏層！

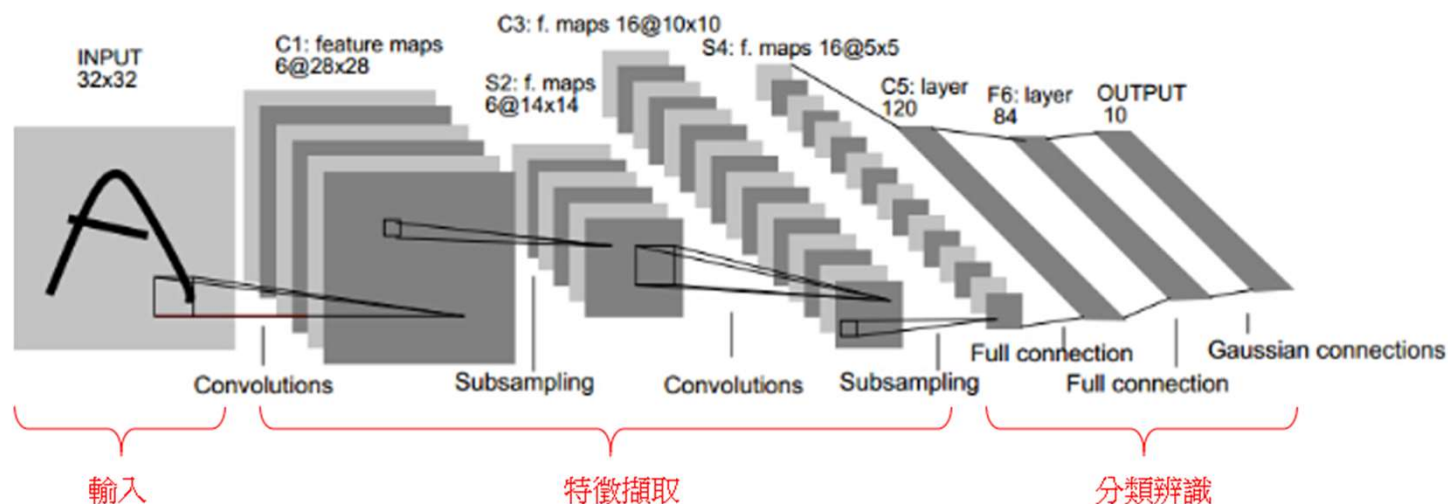
CNN : 卷積神經網路



- 卷積神經網路 (Convolutional Neural Network, CNN)
- 輸入資料為影像，適用 CNN 架構的神經網路
- 人類視覺: 由“圖片特徵”來辨別物體
- 電腦視覺: 透過卷積運算萃取圖片特徵後，再來辨識物體



CNN 之父 - Yann Le Cun



楊立昆（法語：Yann Le Cun，英語：Yann LeCun，原中文譯名揚·勒丘恩，1960年7月8日 - ），法國籍計算機科學家，2018 年圖靈獎得主，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。他同 Léon Bottou 和 Patrick Haffner 等人創建了 **DjVu** 圖像壓縮技術。他同 Léon Bottou 開發了 Lush 語言。2019 年他同 Yoshua Bengio 以及 Geoffrey Hinton 共同獲得計算機學界最高獎項圖靈獎 **Prix Turing 2019**

RNN：循環神經網路



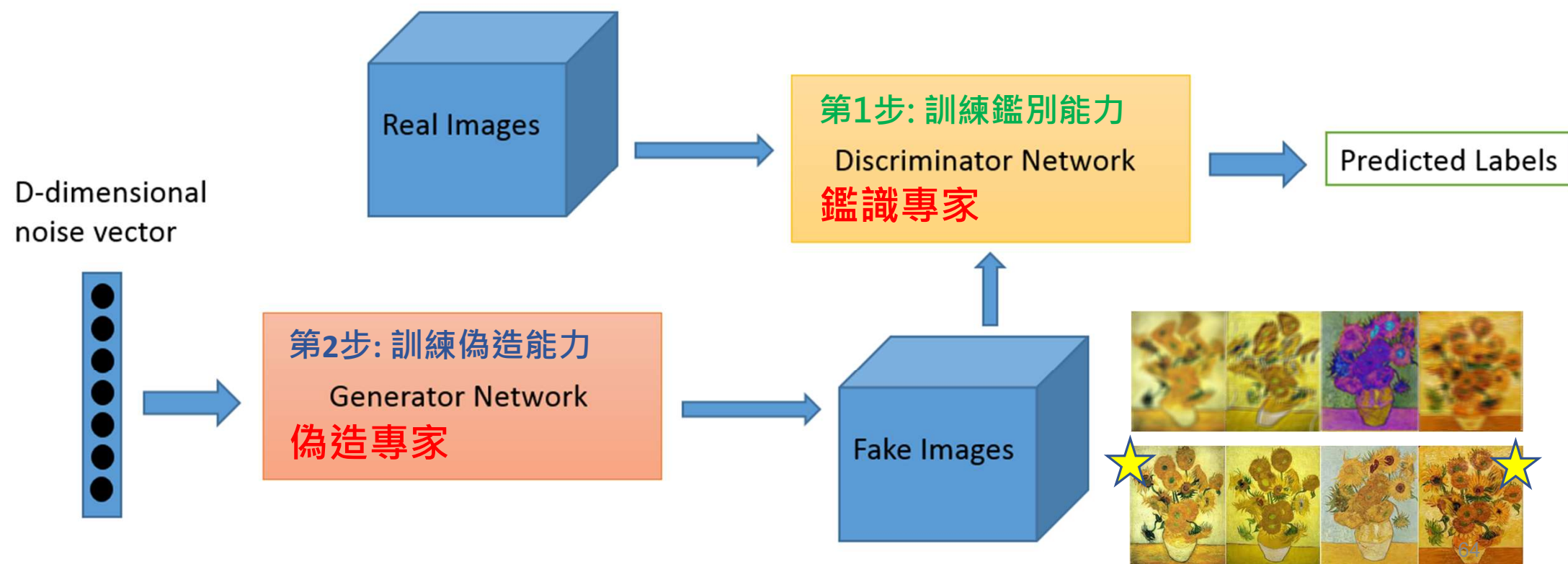
- 遞迴神經網路 **RNN** (Recurrent Neural Network, RNN)
- 如果輸入資料是**序列型資料**，例如一段聲音片段，適合使用 RNN 架構的神經網路。
- 由RNN往後延伸出的神經網路: **LSTM**, GRU

GAN : 生成對抗網路



Generative Adversarial Network

透過生成器與判別器的遞迴對抗來生成圖片，常應用於生成擬真的圖片



應用領域 / 資料型態 / 代表技術



結構化資料 (Table)

表格 (Excel)

ML : Decision Tree

電腦視覺 (CV)

影像 (照片、X光)

DL : CNN

自然語言處理 (NLP)

文字 (文章、留言)

DL : Transformer

時間序列 (Time Series)

訊號 (聲音、股市)

DL : RNN

生成 (Generation)

影像, 文字, 訊號

DL : GAN,
Diffusion

P.S.以上舉例和技術僅列出常見代表，並不是絕對的



Part 2

電腦視覺任務

影像分類

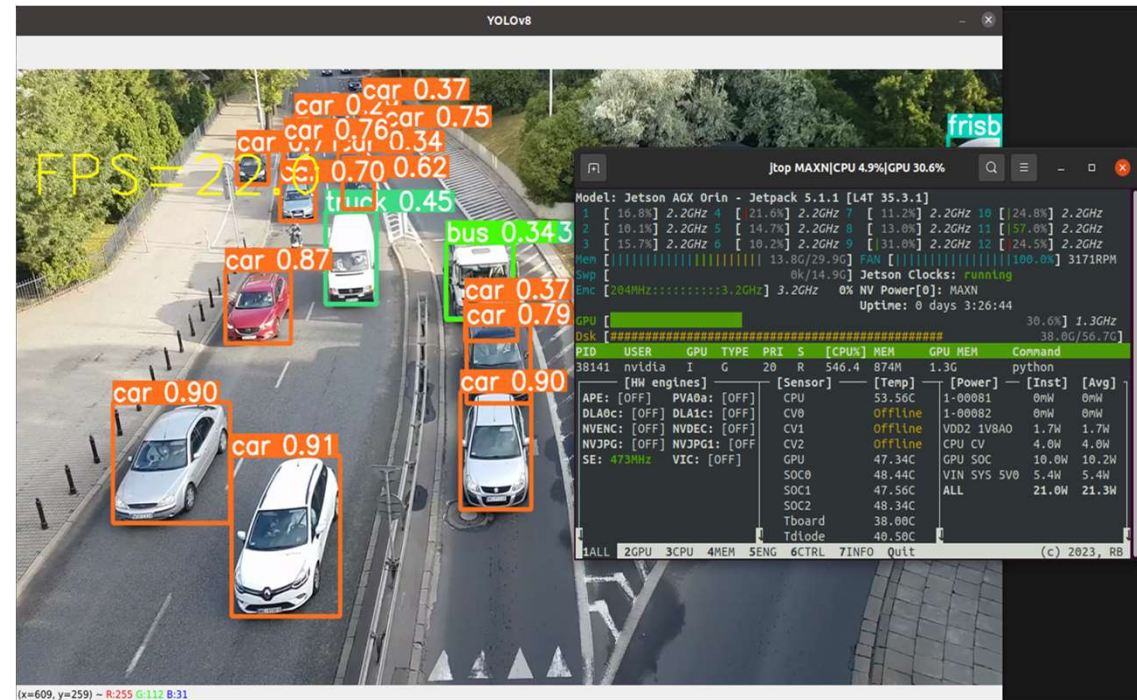


- 影像辨識，或稱為影像分類 (Image Classification)
- 根據訓練好的深度神經網路模型將指定畫面分類到多個類別(class)其中之一。
- 可根據神經網路推論的信心指數來判斷其分類成效。



物件偵測

- 物件偵測 (Object Detection)
- 訓練模型將欲辨識的物體框出



影像分割 - 語義分割



- 語義分割 (semantic segmentation)
- 對圖像中的每個像素打上類別標籤，如下圖，把圖像分為人（紅色）、樹木（深綠）、草地（淺綠）、天空（藍色）標籤



語義分割特別適用於**環境感知**

(可對每張畫面中多個可能出現的物體進行密集的像素級別分類推論，包含畫面的前景及後景)

影像分割 - 實例分割



- 實例分割 (Instance segmentation)
- 目標檢測和語義分割的結合，先在圖像中將目標檢測出來（目標檢測），然後對每個像素打上標籤（語義分割）



實例分割只對圖像中的目標（如上圖中的人）進行檢測和按像素分割，區分不同實例（使用不同顏色）

影像分割 - 全景分割



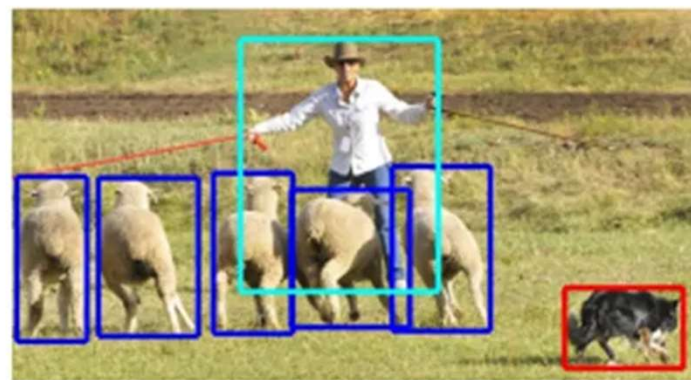
- 對所有目標都檢測出來，又要區分出同個類別中的不同實例



全景分割是對圖中的所有物體包括背景都要進行檢測和分割，區分不同實例（使用不同顏色）



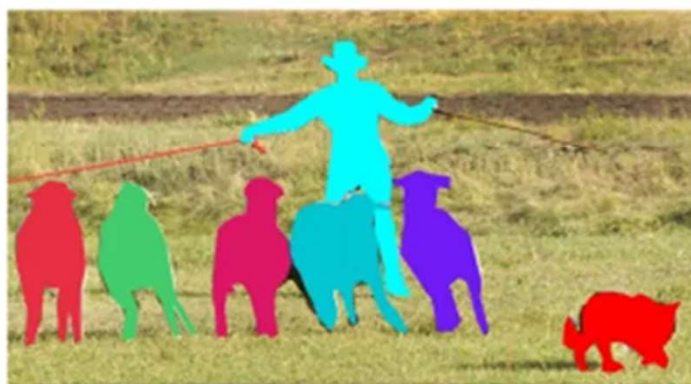
(a) Image classification



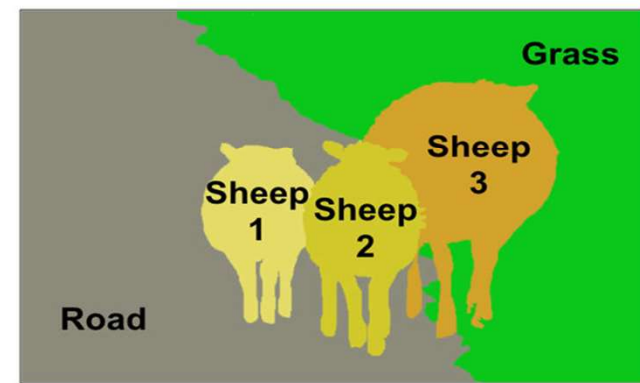
(b) Object localization



(c) Semantic segmentation



(d) Instance Segmentation



(e) Panoptic Segmentation



04 遷移學習

預訓練模型 (Pretrained Model)



- 預訓練模型是指在大量資料上提前訓練好的模型，它們學習了很多通用的知識和特徵。這些模型可以在進行特定任務時直接使用，只需使用少量資料額外訓練，就能達到很好的效果。
- 簡單來說，預訓練模型就像是已經學好了基本技能的學生，當他們需要解決特定問題時，只需進行一些針對性的學習就可以快速上手並表現優異。

遷移學習



- 以影像辨識為例：你可以下載別人訓練好的人臉辨識模型(預訓練模型，pretrained model)，接著拿自己公司員工的照片對模型微調(再訓練)，即可應用於公司門禁人臉辨識系統
- 將一個任務上訓練好的預訓練模型應用於另一個相關任務，使用自己的資料，就可以對預訓練模型進行微調，達到因地制宜的效果。
- 透過遷移學習，可以大幅提高機器學習的效率和準確性



05 雲端運算與邊緣運算

雲端運算 v.s. 邊緣運算



特性	雲端運算	邊緣運算
定義	使用遠端服務器網絡來儲存、處理和管理資料。	通常是在本地機器或單板電腦上進行運算
延遲	較高，因為數據需在遠端處理	較低，因為數據就近處理
網路頻寬要求	較高，需要將大量數據傳輸到雲端	較低，大多數數據在本地處理
適用場景	需要大量計算和儲存空間的應用	需要快速反應和處理即時數據的應用
成本	通常較高，取決於使用的服務和儲存量	可能較低，因為降低了數據傳輸和存儲需求
安全性	受到雲服務提供商的安全措施保護	資料可以在本地處理，減少資料外洩風險
擴展性	高，可以輕鬆增加更多的存儲和處理能力	受限於本地硬件和資源

聯邦學習



聯邦學習 (Federated Learning) 是一種機器學習方法，其主要特點如下：

- 1. 分散式資料訓練**：在聯邦學習中，資料保持在本地設備上，不需要將其集中或傳輸到中心伺服器。 **模型訓練分布在多個設備**（例如智能手機或個人電腦）上。
- 2. 隱私保護**：由於資料不離開本地設備，聯邦學習可以**提高用戶的隱私保護**。敏感數據如個人資料不需要共享給模型訓練者。
- 3. 模型更新共享**：雖然資料留在本地，但**訓練過程中學習到的模型更新**（例如**權重更新**）**會被發送回中央服務器共享**。在那裡，這些更新被綜合以改進全局模型。



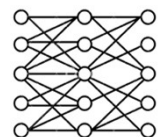
06 生成式AI

判別式 AI v.s. 生成式 AI

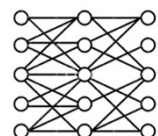
輸入資料類型

AI模型

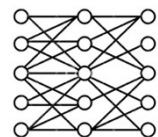
預測結果



盜刷
or
非盜刷



狗
貓
老虎
獅子
·
·
·



正常
or
異常

2012
AlexNet

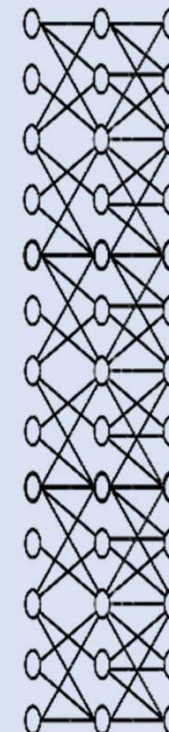
2016
 AlphaGo

輸入 prompt

AI 模型

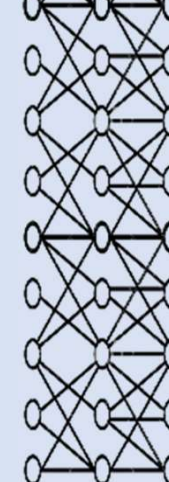
生成結果

台灣在哪裡？



台灣是一個位於
東亞的島嶼，位
於西太平洋，東
鄰太平洋...

幫我畫一隻
在太空游泳
的狗



幫我作一首有微
風感覺的鋼琴曲

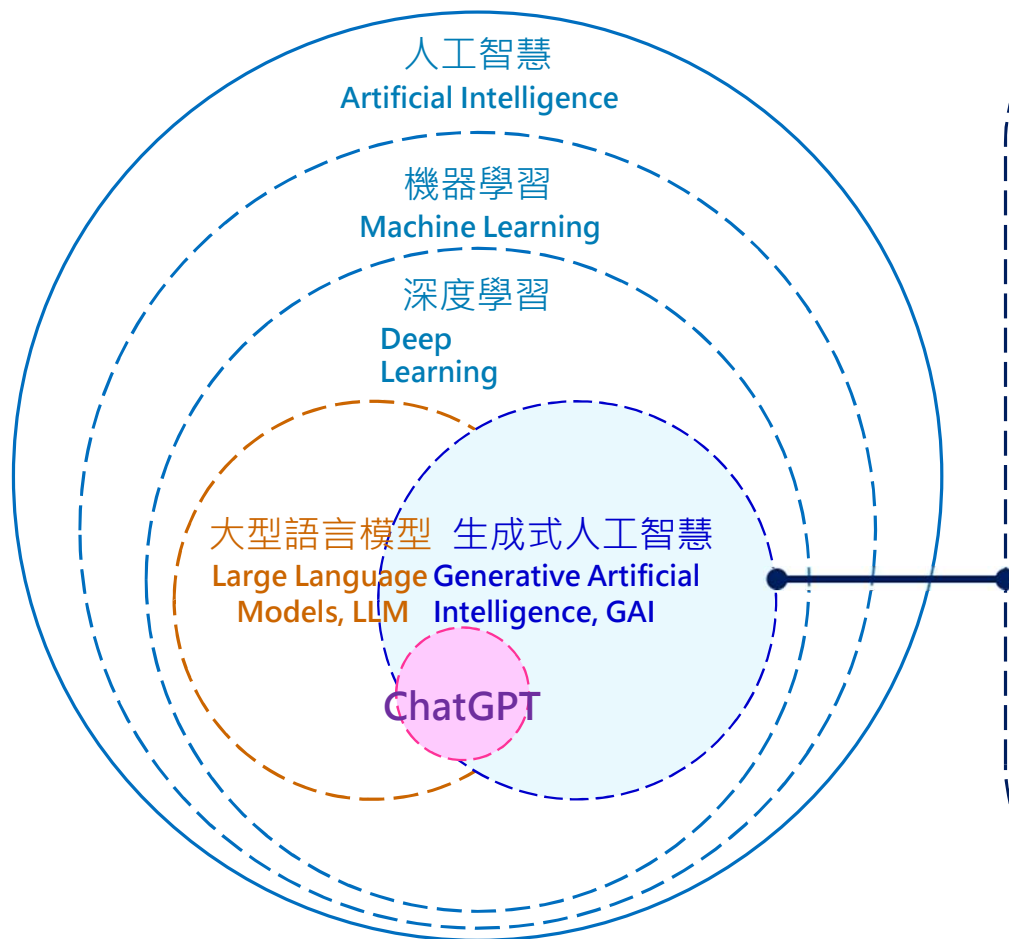


2023
 ChatGPT

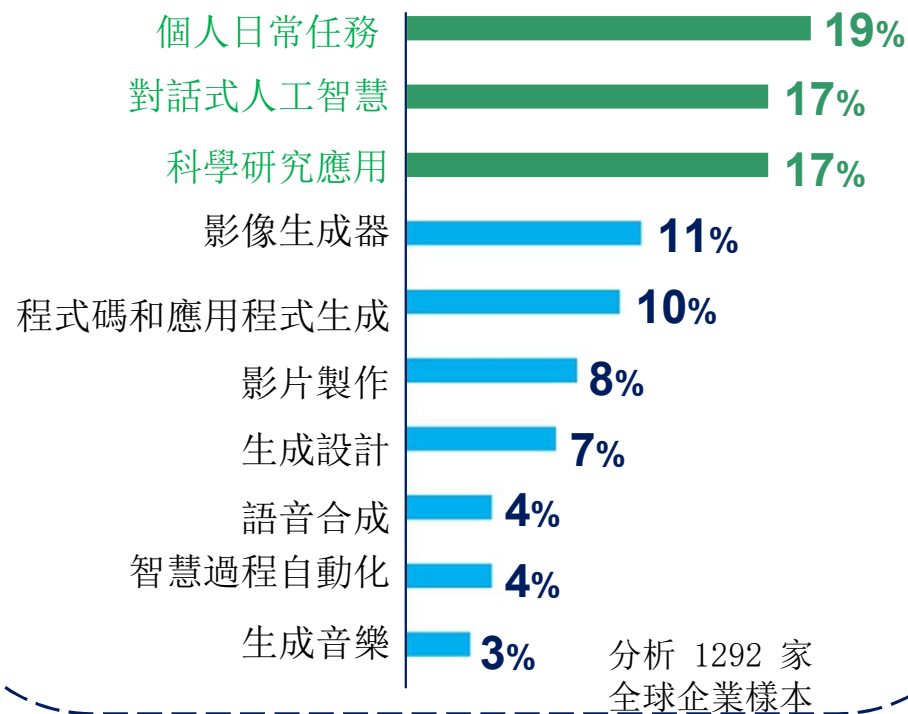
DALL·E 3

The iPhone
moment of A.I.

生成式 AI 的主要應用



2023年十大生成式AI應用趨勢



Generative AI (GenAI)



- 主要用於**創造性的工作**，例如文章生成、影像生成、音樂生成等。
- 機器根據已有的資料庫自主學習、創造、生成包含圖像、文本、聲音等新的資料內容，或是完成特定任務的人工智慧模型，
- 應用範圍包括(但不限於)：語言模型、自然語言生成、圖像生成、影片生成、音樂生成、藝術創作...等。
- GenAI 根據**機率**回答的答案**可能不完全是事實**，輸出內容可能帶有偏見、生成的圖像可能缺乏構圖完整性，這與模型當初訓練時的**資料品質**有關。

Generative AI 的應用場景



AIGC是什麼呢？

AIGC全名為
AI Generated Content,
即人工智慧生成內容, 所以AI繪圖, AI作曲, AI回
答問題都算是AIGC。



應用場景	功能
數據分析	自動化數據收集、整理、分析和報告
文字生成	自動生成新聞、文章、網頁內容等
語音生成	能夠以自然的聲音調式和音色生成語音內容，如新聞廣播、有聲書、自動語音助手等
影像生成	能夠自動生成圖片、視頻等多媒體內容，如藝術品、照片、影片等
電子商務	幫助企業自動生成產品描述、產品評論、客戶評價等內容
遊戲和娛樂產業	自動生成遊戲和虛擬現實（VR）的遊戲內容
廣告和營銷	幫助企業自動生成廣告文案和視頻廣告內容
安全領域	自動檢測和評估系統漏洞和網絡安全風險
醫療領域	幫助醫生自動生成病例報告、診斷和治療方案
智能客服和客戶服務	幫助企業實現自動化客戶服務，如自動回答客戶問題、提供虛擬客服等

來源

文字
影像
訊號
聲音
影片

.....



生成

產出

文字
影像
訊號
聲音
影片

.....

ChatGPT ≠ 一切問題的答案



Newtalk新聞

ChatGPT成洩密管道！資安公司監測160萬人 7天內4.96萬人外洩機敏資料

ChatGPT爆紅，雖然吸引許多企業開始嘗試導入在工作當中，但也開始傳出企業機密資料外洩災情，一家資安公司日前統計了旗下企業用戶的使用。

ENN台灣電報

ChatGPT闖禍！三星遭洩密 義大利開禁用第1槍

商傳媒 | 綜合報導南韓三星工程師因使用ChatGPT為輔助工具，快速修復原始程式碼錯誤，卻不慎洩露了會議紀錄、工廠性能、產量等機密資訊。

1 month ago

iThome

員工外洩內部機密！三星開放ChatGPT

南韓媒體《Economist》等，而分別在不同

大紀元

AI洩密及虛假訊息事件頻傳 多國考慮禁用

隨著ChatGPT帶起人工智能 (AI) 的熱潮，與AI相關的洩密和假訊息等問題也逐漸浮現。這讓不少科技先驅者呼籲監管機構管制AI。意大利政府因ChatGPT洩密和...



經濟日報

日企憂洩密 禁用ChatGPT

日經新聞報導，ChatGPT等AI工具下，...



AI) 3
務...



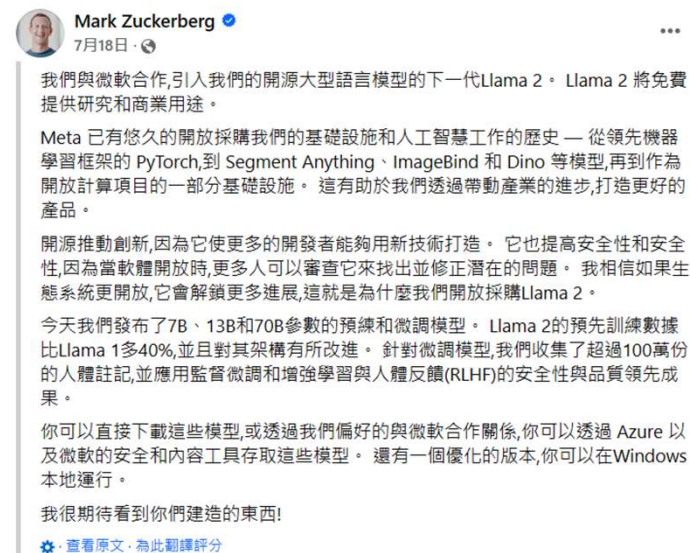
開源大語言模型（ open-source LLM ）



- 分享模型參數、模型架構、訓練方法以及訓練程式碼，有些模型會分享訓練資料
- 開發者可以在現有模型的基礎上進行擴展和改進，無需從頭開始訓練模型

Meta 執行長 Mark Zuckerberg：

**『開源（ open-source ）推動創新，讓更多開發人員
能使用新技術來進行開發』**



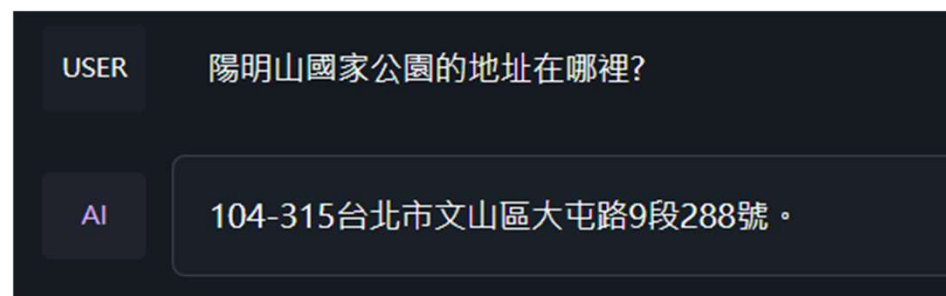
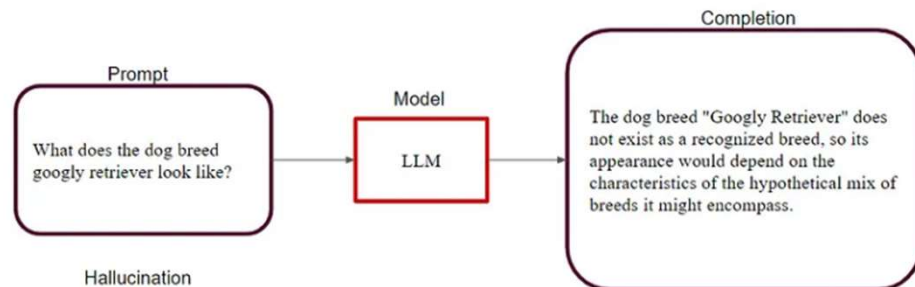
Reference: [Mark Zuckerberg's Facebook](#)

大型語言模型的限制(1)-幻覺



幻覺：一本正經的胡說八道

幻覺包含邏輯錯誤、無中生有等，溫度參數值設定太高可能導致幻覺



大型語言模型的限制(2)-偏見



- 模型根據訓練資料學習，如果訓練資料中存在偏見，那麼模型也會學到這些偏見，生成不準確或誤導性的訊息。
- 某些主題缺乏足夠的訓練資料，此情況會特別明顯
- 生成式 AI 可能含有種族、性別偏見內容

大型語言模型的限制(3)-知識斷點



- 例如GPT-3.5 模型的知識目前只更新到2022年1月，向模型詢問最近的NBA總冠軍得主，無法得到正確資訊。
- 同樣的問題問GPT-4 (知識更新到2023年4月)，可得到正確答案

GPT3.5

ET You

2022年NBA總冠軍是哪一隊？

ChatGPT

很抱歉，我的知識截止日期是2022年1月，我無法提供2022年NBA總冠軍的信息。要獲取最新的NBA冠軍資訊，建議您查詢網上新聞或訪問NBA官方網站。

GPT4

ET You

2022年NBA總冠軍是哪一隊？

ChatGPT

2022年NBA總冠軍是金州勇士隊 (Golden State Warriors)。

免費、可商用的開源模型出現



- Meta 於 2023 年 7月發佈免費商用模型 **Llama2**，有 7B、13B、70B 三種版本
- 70B 參數的版本，在一些任務上的表現不輸gpt-3.5-turbo
- 與 Microsoft Azure 平台合作，加速企業開發企業大腦

MODEL SIZE (PARAMETERS)	PRETRAINED	FINE-TUNED FOR CHAT USE CASES
7B	Model architecture: Pretraining Tokens: 2 Trillion Context Length: 4096	Data collection for helpfulness and safety:
13B		Supervised fine-tuning: Over 100,000
70B		Human Preferences: Over 1,000,000



Reference: <https://www.jonpeddie.com/news/llama-2-to-run-locally-for-you-with-snapdrag>

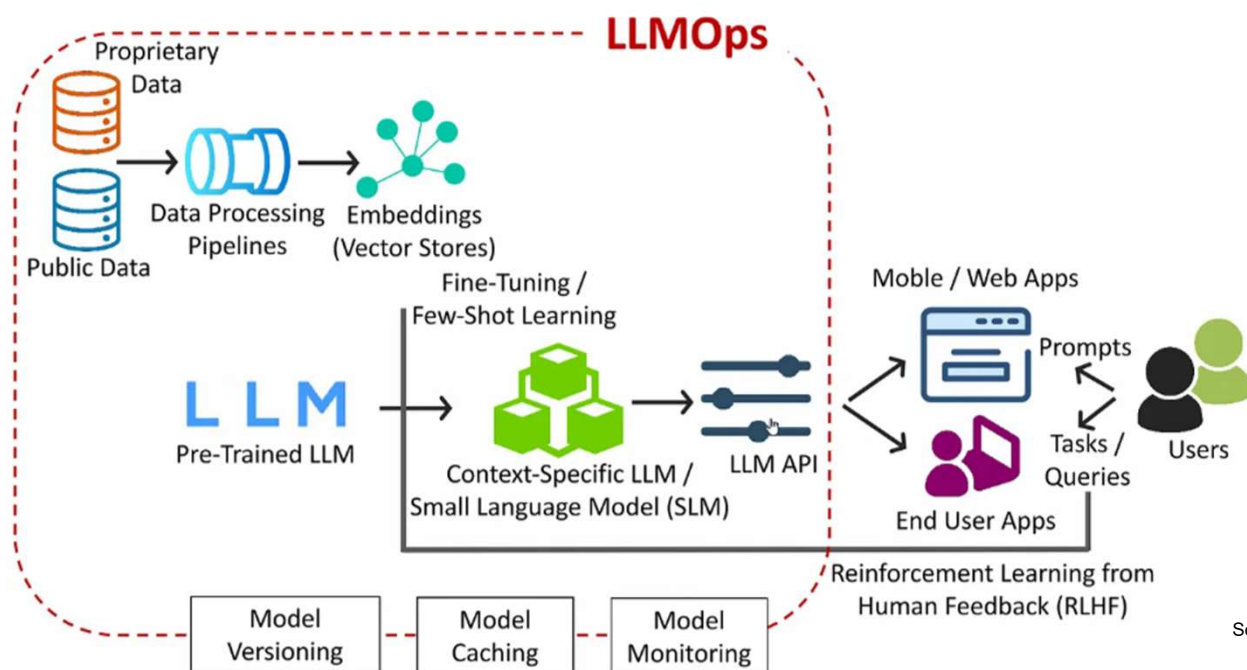
運用 LLM 大型語言模型的兩種方式



	(1) 利用API 串接未開源模型 (例如 ChatGPT)	(2) 使用開源模型再加上自己的資料訓練(例如 Llama 3)
適合用戶	用量較少的個人或中小企業	擔心資料外洩的用戶 想要建立私有 LLM 的企業
優點	能透過 API 將自家產品與模型連結 產生的回應品質領先市面上其他模型	使用者能用自己的資料從零開始訓練 模型改造為企業獨有的 LLM
缺點	可能洩漏企業機敏資料	需要龐大的基礎設備、金錢成本、專業 AI 人才

LLMOps

- LLMOps，與 MLOps 類似，是一個將 LLM 模型部署到實際環境中的一個流程
- LLMOps 的流程包含了模型的訓練、模型的壓縮、模型的部署、模型的應用等



Source : <https://medium.com/@bakingai/llmops-the-future-of-mlops-for-generative-ai-aed95decf21e>

模型開源共享平台 - Hugging Face



- 2016 年在紐約成立，原本是一個開發青少年聊天機器人應用為主的公司
- 逐漸轉型成允許使用者共享機器學習模型和資料集的平台
- 保持 AI 研究空間的開放，是一股強而有力的民間力量
- 共享超過10萬個預訓練模型，上萬資料庫，並提供模型託管、測試的服務



企業導入生成式 AI 的步驟



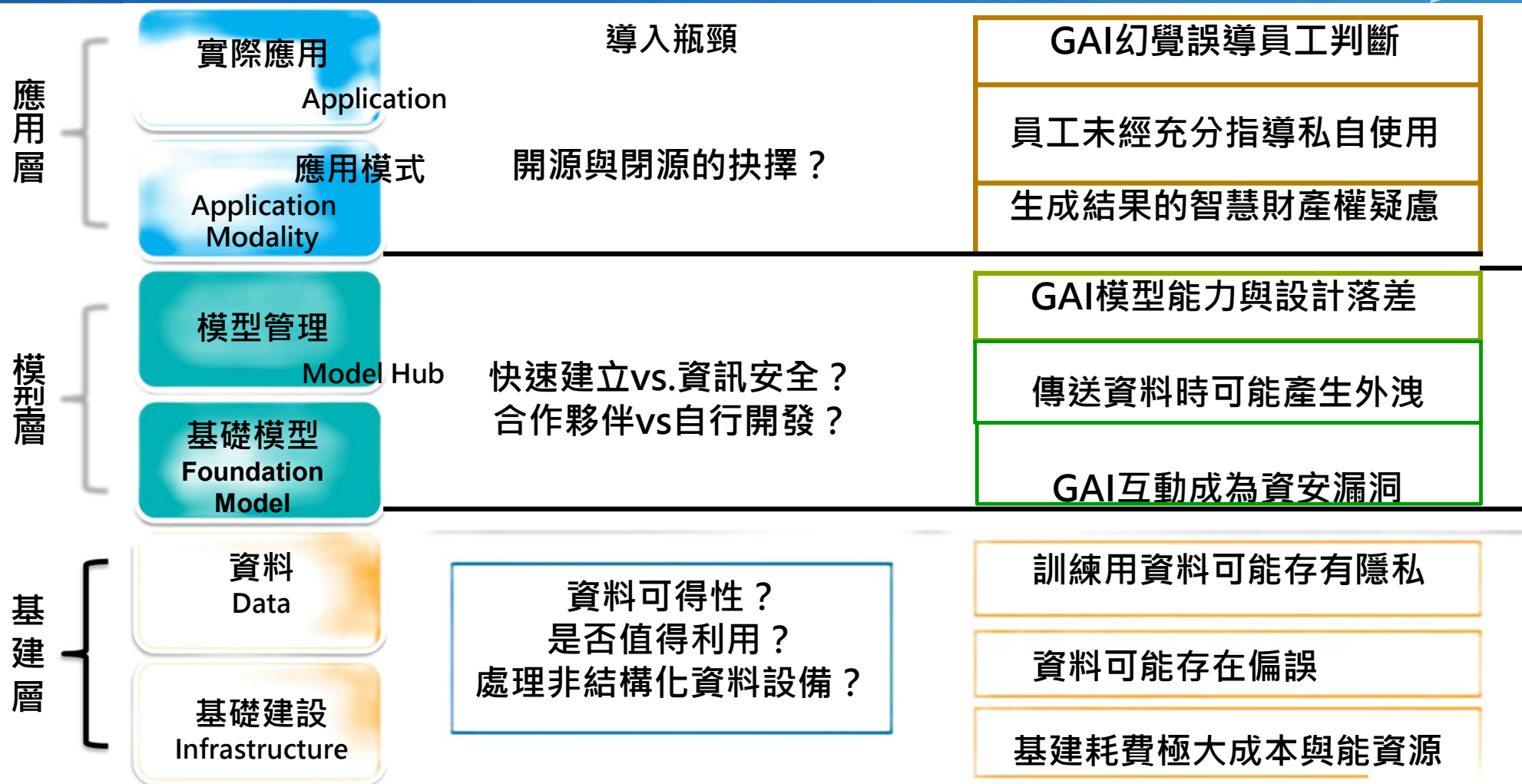
吳恩達 (Andrew Ng) 在 “ AI Transformation Playbook ” 中，對於想要利用人工智慧來轉型的公司，提出了5個步驟：

- Step1: 執行試點的專案從而獲得前進的動力
- Step2: 建立公司內部的 AI 團隊
- Step3: 對員工進行廣泛的 AI 培訓
- Step4: 制定 AI 的戰略
- Step5: 在公司內外建立良好的溝通管道



生成式AI應用瓶頸與風險

9大應用風險



Source: 工研院產業科技國際策略發展所

檢索資料回答特定領域知識 (檢索增強生成)

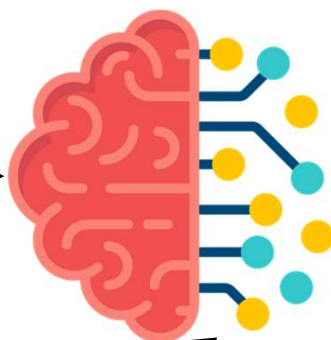


模型生成回應

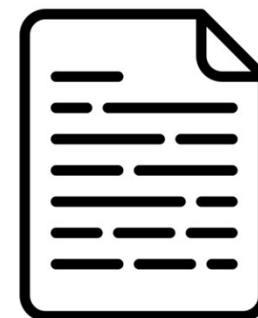


問題

多模態/大語言模型



回答



輸入資料給模型



檢索增強生成 (Retrieval-Augmented Generation, RAG)

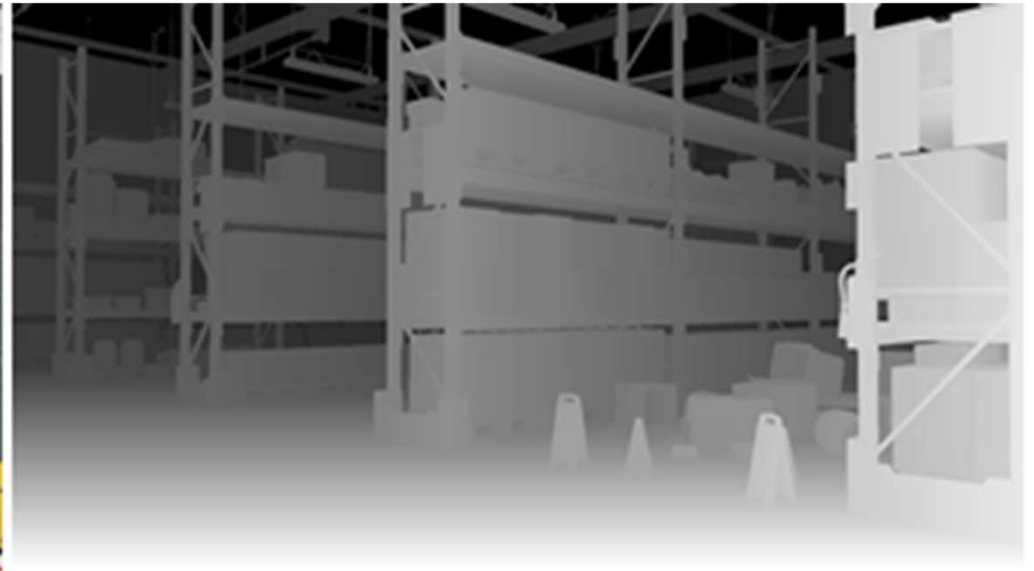
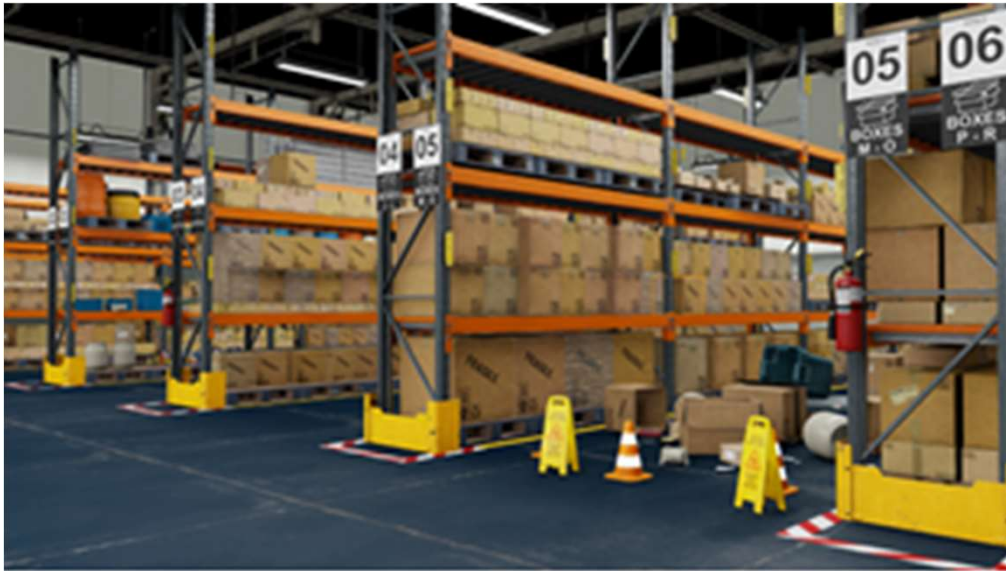


07 AI 輔助數位學生

數位孿生 (Digital Twin)



- 數位孿生，或稱為數位映射、數位對映、數位分身、數位雙生
- 數位孿生是現實世界的物理實體（例如人、商品、流程或整套系統）在數位平台上的虛擬「雙胞胎」，可用來觀察實體、或是模擬實體所發生的變化。
- 可用來監督和管理實體，或是測試影響實體的行為和決策。





08 人工智慧安全

資料隱私的疑慮



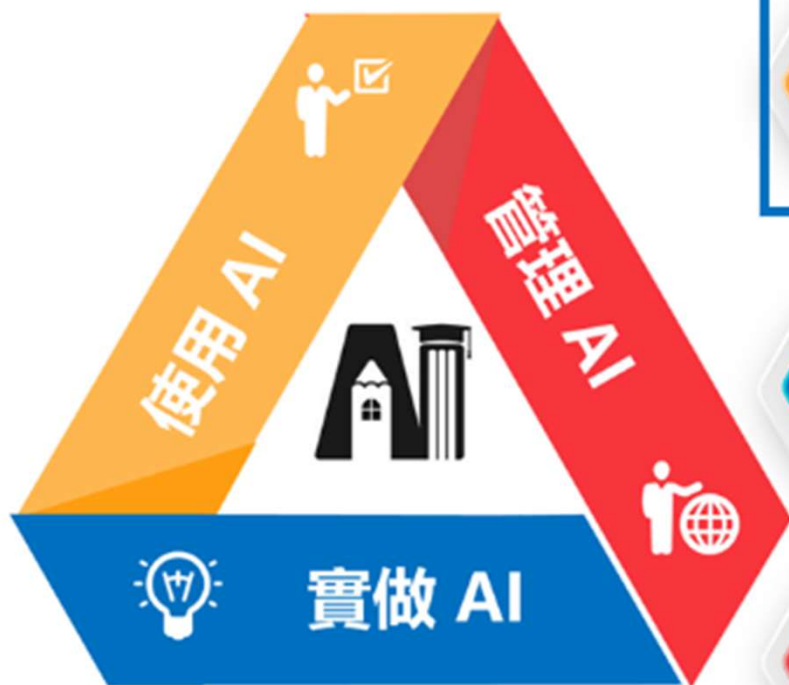
- 資料難以追蹤、合法保存與使用或移除機制，當機敏資料不小心進入模型是否會產生不預期的外流
- AI 系統儲存大量資料，容易成為駭客或惡意軟體攻擊目標

三星開放 內部使用	機密資料 接連外洩	不得使用生成 式 AI	自行研發
<ul style="list-style-type: none">• 20 天內三起機密資訊外流事件• 緊急啟動資安指施，限制使用流量	<ul style="list-style-type: none">• 再次有工程師將內部原始程式碼上傳到 GhatGPT	<ul style="list-style-type: none">• 規定公司內不得使用• 不得上傳個人裝置，違規者恐被開除	<ul style="list-style-type: none">• 自行研發 LLM，協助員工完成翻譯、知識搜索以及摘要等任務• 避免過度依賴海外科技公司技術
2023.3	2023.4	2023.5	2023.6



09 AITC 素養認證考試

AITC 認證



01

素養級: 用AI TCLiteracy

評估具備基本的AI素養，如：AI的基礎知識與原理、生成式AI的使用、場域應用與實務、對倫理與社會影響、技術限制與前景、AI的規範與重要性等必須素養能力。

02

工程級: 做AI TCEngineering

有程式設計基礎，具資料處理分析、機器學習和深度學習的應用能力，如：SVM、決策樹、神經網路、機器學習、深度學習的算法等，並要求在真實資料集上進行模型的訓練、部屬與評估等工程技術能力。

03

管理級: 管AI TCManagement

AI資安、架構設計和AI倫理等，此等級要有能力在實際情境中建立安全、效能優良的AI系統，並了解AI的倫理使用原則等管理治理能力。

AIATCL 素養試卷題型 & 計分方式

Taiwan AI Academy 人才認證 (AIATC™) 是 2024 年由台灣人工智慧學校 (AIA) 開始創建主辦的 AI 人才素養與專業能力認證，包含素養級 (AIATCL)、工程級 (AIATCE) 與管理級認證 (AIATCM) 等三種認證，認證內容對應到台灣人工智慧學校以「用 AI、做 AI 與管 AI」為主軸所規劃的各種 AI 課程。



AIATCL 素養試卷題型 & 計分方式

素養級認證測驗 (TCL) 總測驗時間為 **50分鐘**，共分為以下三種題型：

- 一、單選題：共 15 題，每題 2.5 分，共 37.5 分。
- 二、多選題：共 5 題，每題 2.5 分，共 12.5 分。
- 三、閱讀題組：共 20 題，每題 2.5 分，共 50 分。

通過標準為 **80分**，最多可錯 8 題

考試期間將有監考人員在場。無論是採用電子系統還是紙筆測驗，本次考試均為閉卷考試(禁止攜帶任何參考資料)。



10 GenAI 補充資料 & 自學資源

GenAI 素養參考資源

- 人工智慧安全 (AI Safety) :
 - 控制邊緣：未來科技與全球秩序的抉擇 - DeepMind 共同創辦人蘇萊曼從歷史角度分析科技演變，評估新興技術利弊。他指出 AI 帶來進步也帶來威脅，人類須在矛盾中尋求平衡，主動參與塑造未來。
 - Introduction to Generative AI: An Ethical, Societal, and Legal Overview - 介紹生成式 AI 技術，探討其工作原理、應用方法、社會影響等。內容涵蓋大型語言模型、AI 整合策略、創新與責任平衡、法律政策等議題，適合想了解生成式 AI 的讀者。
 - AI 安全：技術與實戰 - 關於人工智慧安全的全面著作，內容涵蓋 AI 發展歷史、各國戰略、安全技術框架、攻防案例分析、未來展望等。

GenAI 素養參考資源

- 機器學習 & 深度學習：

- [AI科學家李飛飛的視界之旅](#) - 講述李飛飛從移民孩童到人工智能領域頂尖科學家的傳奇故事。她分享個人成長經歷，探討人工智能的發展與影響，呼籲以人為本、符合倫理地運用這項技術。
- [深度學習 \(Deep Learning\)\(繁體中文版\)](#) - 深度學習聖經，由 Ian Goodfellow, Yoshua Bengio, Aaron Courville 撰寫。
- [人工智慧在台灣](#) - 適合剛接觸人工智慧的一般大眾，是初學者快速了解人工智慧的入門指南。
- [台大 李宏毅 機器學習](#) - 從底層的理論技術到應用都有全面性的介紹，適合想要從技術面完整理解機器學習的人。

GenAI 素養參考資源

- 生成式 AI :

- 資策會 [企業應具備的生成式AI素養](#) - 生成式 AI 導入指引
- [台大 李宏毅 生成式AI導論](#) - 《生成式人工智慧導論》將重點放在解釋生成式人工智慧的基本原理上，適合已具備機器學習概念的人延伸學習。
- [Learnprompting](#) - 提示工程 (Prompt Engineering) 是與 AI 進行有效溝通來實現預期結果的過程。隨著 AI 技術持續快速的發展，掌握提示工程的技能變得尤為重要。提示工程技術可以應用於各種各樣的任務，使其成為任何尋求提高日常和創新活動效率的人的有用工具。

GenAI 素養參考資源

- 資料科學與資料準備：

- 製造數據科學：邁向智慧製造與數位決策 - 台大李家岩老師整合數據科學與決策科學，教導從問題發現到前瞻性決策的過程。書中涉及數據科學的基礎與進階知識，以製造業為例探討特徵工程、故障預測等技術，也介紹了自適性演算法設計來優化系統決策。
- 資料科學入門完全指南：資料分析的觀念處理實作 - 這本書適合對資料科學有興趣但無程式基礎的讀者，涵蓋了數值、影像、音訊及文字資料的處理與分析，並詳述資料前處理方法和實作案例。
- 資料科學的統計實務：探索資料本質、扎實解讀數據，才是機器學習成功建模的第一步 - 這本書探討了資料科學的核心概念，從統計基礎到分析技術，並指出常見的錯誤，目的是提升讀者的資料科學素養，幫助他們成為卓越的資料科學家。