

ANN

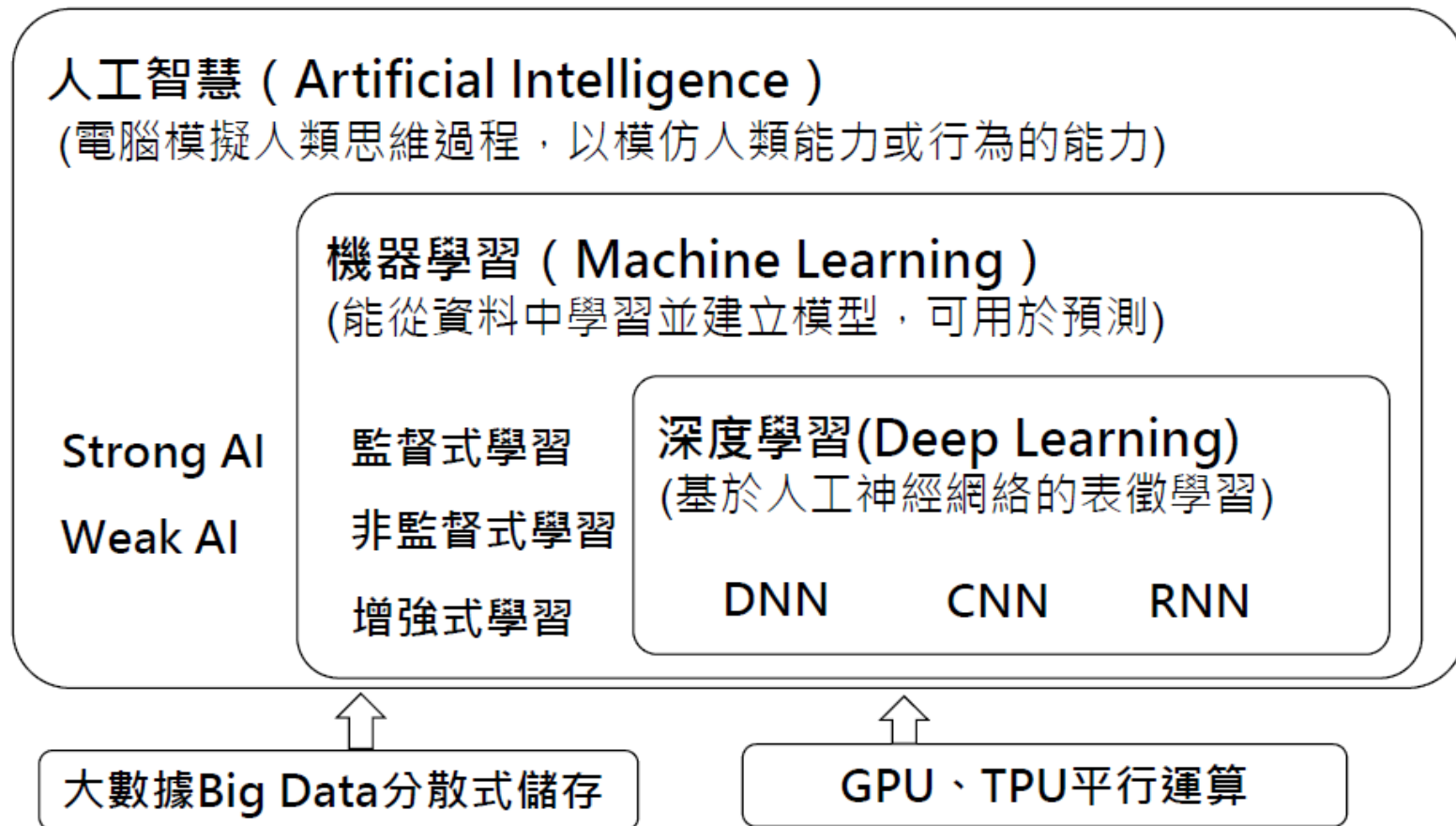
CH01

人工智慧、機器學習、
深度學習介紹

1

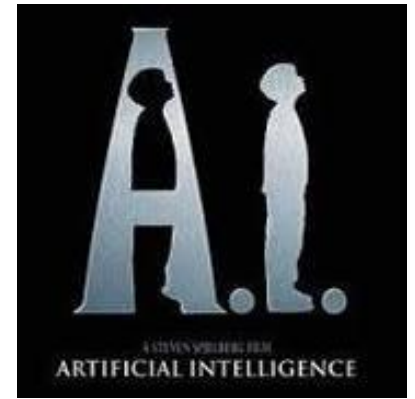
- 近年來人工智慧 AI (Artificial Intelligence) 吸引大眾與媒體的目光，尤其 AlphaGo 的成功更加吸引了大量的關注。
- 其實 AI 早已進入你我的生活中，你的手機的語音助理、人臉識別、自動篩選有興趣的新聞、影音平台的每日推薦等。
- 人工智慧才剛起步而已，未來人工智慧的應用，將會深入生活的每一個層面，也就是說未來是一個AI的時代。

1.1 人工智慧、機器學習、深度學習的關係



人工智慧 (Artificial Intelligence)

- 「人工智慧」 (**AI**) 名詞最早開始於1950年代
- 以電腦科學、生物學、心理學、語言學、數學、工程學為基礎的科學，將人類智能加以電腦化，使電腦具有思考、學習及解決問題的能力
- 目標
 - 將人類的自然智慧實現於電腦上
 - 讓電腦像人類一樣思考與學習



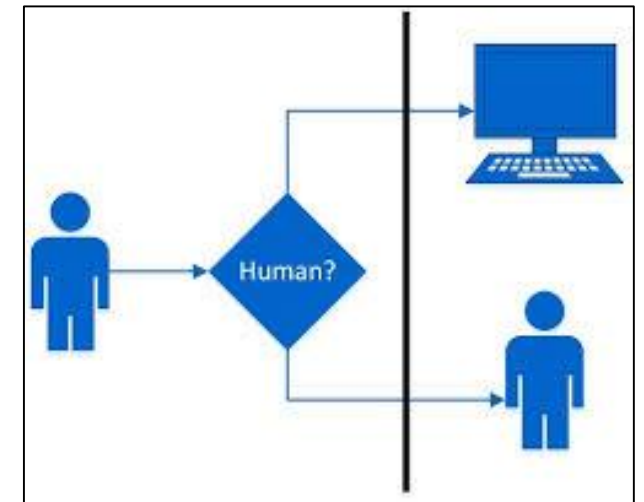
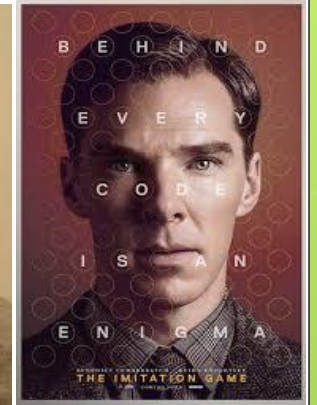
人工智慧 (Artificial Intelligence)

- 1980年代，約翰瑟爾(John Searle)提出人工智慧二分類
 - 強人工智慧(**Strong AI**)：機器具有與人類相同完整的認知與智慧能力
 - 弱人工智慧(**Weak AI**)：機器不具有與人類相同完整的認知能力，只要設計得看起來像具有智慧即可
- 目前大多數都是針對已模型化、數學化的特定問題建立的AI
 - AlphaGo(圍棋)、Deep Blue(西洋棋)
 - Watson使用自然語言來回答問題(危險邊緣)

<http://www-07.ibm.com/tw/blueview/2011apr/watson.html>

人工智慧(Artificial Intelligence)

- 圖靈(Alan M. Turing)：電腦科學與人工智慧之父
 - 「圖靈測試」(Turing Test)
 - 人類與機器透過電傳設備對話
 - 如果人類無法根據對話過程，判斷對方是機器或是人，就算通過圖靈測試
 - 可認定這台機器具有人工智慧
- Google語音助理



機器學習 (Machine Learning)

- ML是AI的分支
- ML透過演算法，使用大量資料進行訓練，藉此產生特定模型
- 未來當有新的資料，便可以使用特定模型進行預測
- ML學習策略包括「監督式」、「非監督式」與「增強式」學習
 - 監督式學習(Supervised Learning)：訓練時要給正確答案
 - 非監督式學習(Unsupervised Learning)：沒有正確答案
 - 增強式學習(Reinforcement Learning)：做對給予獎勵

機器學習與傳統程式設計的比較

傳統
程式設計

Data
Rules

Traditional
Programming
傳統程式設計

Answers

- 輸入資料與規則，就能得出答案：

優點：不需要大量資料與運算

缺點：程式設計必須加入所有規則，但是規則太多很難完全列舉。

機器
學習

Data
Answers

Machine Learning
機器學習

Rules

Data

Model
模型

Prediction

- 輸入資料與答案，經過機器學習演算法訓練後，就能夠得出規則並且建立模型，後續可使用此模型輸入資料進行預測：

優點：能夠自動由資料中學習並且建立模型

缺點：需有大量資料並且需要大量運算

機器學習(ML)應用相當廣泛

- 當前熱門應用
 - 推薦引擎、搜索引擎、定向廣告、垃圾郵件過濾、詐騙偵測
 - 自然語言處理、視覺辨識、語音識別、手寫識別
 - 醫學診斷、證卷分析、需求預測

深度學習 (Deep Learning)

- ML的再分支，是目前AI應用中成長最快的領域
- 模擬人類神經網路的運作方式，常稱「類神經網路」(Artificial Neural Network, ANN)
- DL廣泛應用於自然語言處理、視覺辨識、語音識別、手寫識別等領域
- 常見的深度學習架構
 - 多層感知器(Multi-layer Perceptron, MLP)
 - 深度神經網路(Deep Neural Network, DNN)
 - 卷積神經網路(Convolutional Neural Network, CNN)
 - 迴遞神經網路(Recurrent Neural Network, RNN)

為何近年來AI發展加速？

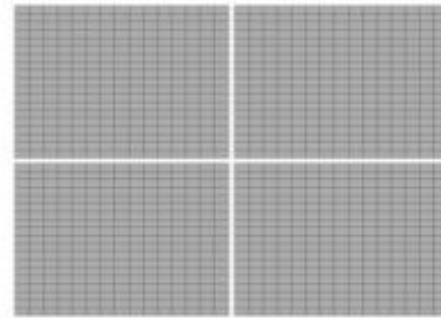
- 早在1960-70年代，科學家已提出各種機器學習演算法，但受限於電腦計算能力，以及大量資料取得不易，機器學習一直沒有很成功
- 近年來，隨著全球網路、設備與系統的連接，再加上分散式雲端儲存，例如：Hadoop、NoSQL等發展，大數據(Big Data)時代來臨
- 加上大量伺服器的平行運算功能，例如：Spark，提供了龐大的運算能量
- 大數據+高運算能力，推動機器與深度學習加速發展

GPU平行運算

- CPU (Central Processing Unit) 中央處理器：電腦的核心裝置，主要用以解讀電腦指令以及處理資料運算
- GPU (Graphics Processing Unit) 圖形處理器：原本用來處理畫面像素的運算，例如: 電玩畫面需要大量圖形運算
- CPU 與 GPU 架構上不同：CPU 只有數顆核心，為循序處理進行最佳化；GPU 有數千個小型核心，可以發揮平行運算的強大功能



CPU 數顆核心



GPU具有數千顆核心

GPU平行運算

- 深度學習以大量矩陣運算模擬人類神經元運作，矩陣運算的特性是，單一運算都很簡單，但是需要大量運算，因此適合使用 GPU 進行平行運算。
- GPU透過大量核心進行平行運算，可讓深度學習的訓練速度比 CPU 快 10~75 倍，讓訓練時間縮短。

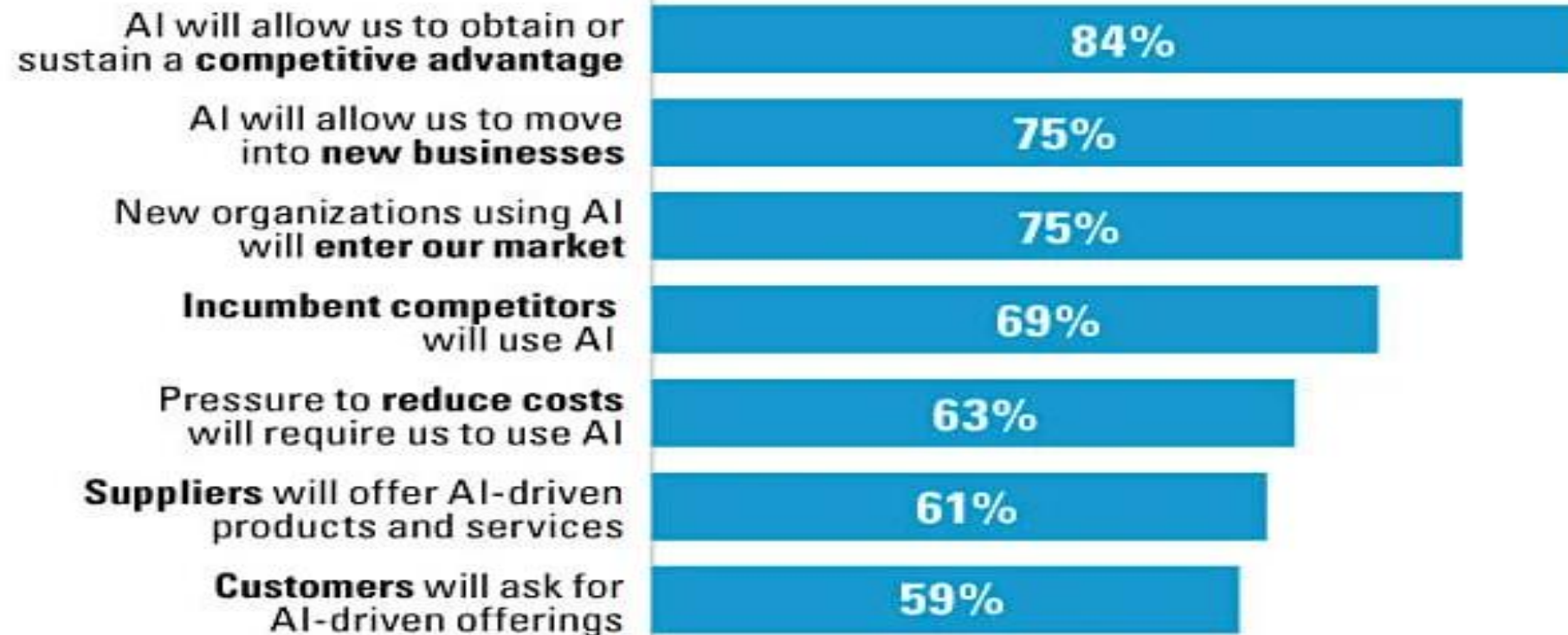


Tesla V100 16GB \$299,000
A100 Price ?



Reasons for adopting AI

Why is your organization interested in AI?



Percentage of respondents who somewhat or strongly agree with each statement

Where Is

ARTIFICIAL INTELLIGENCE

At Right Now?



2 out of 3

consumers are already using AI without even knowing they are interacting with chatbots.

AI will grow into a

\$190B

industry by 2025.



By the year 2020,

60%

of companies will be using artificial intelligence for driving digital revenue.

97%

of mobile users are already using AI-powered voice assistants.

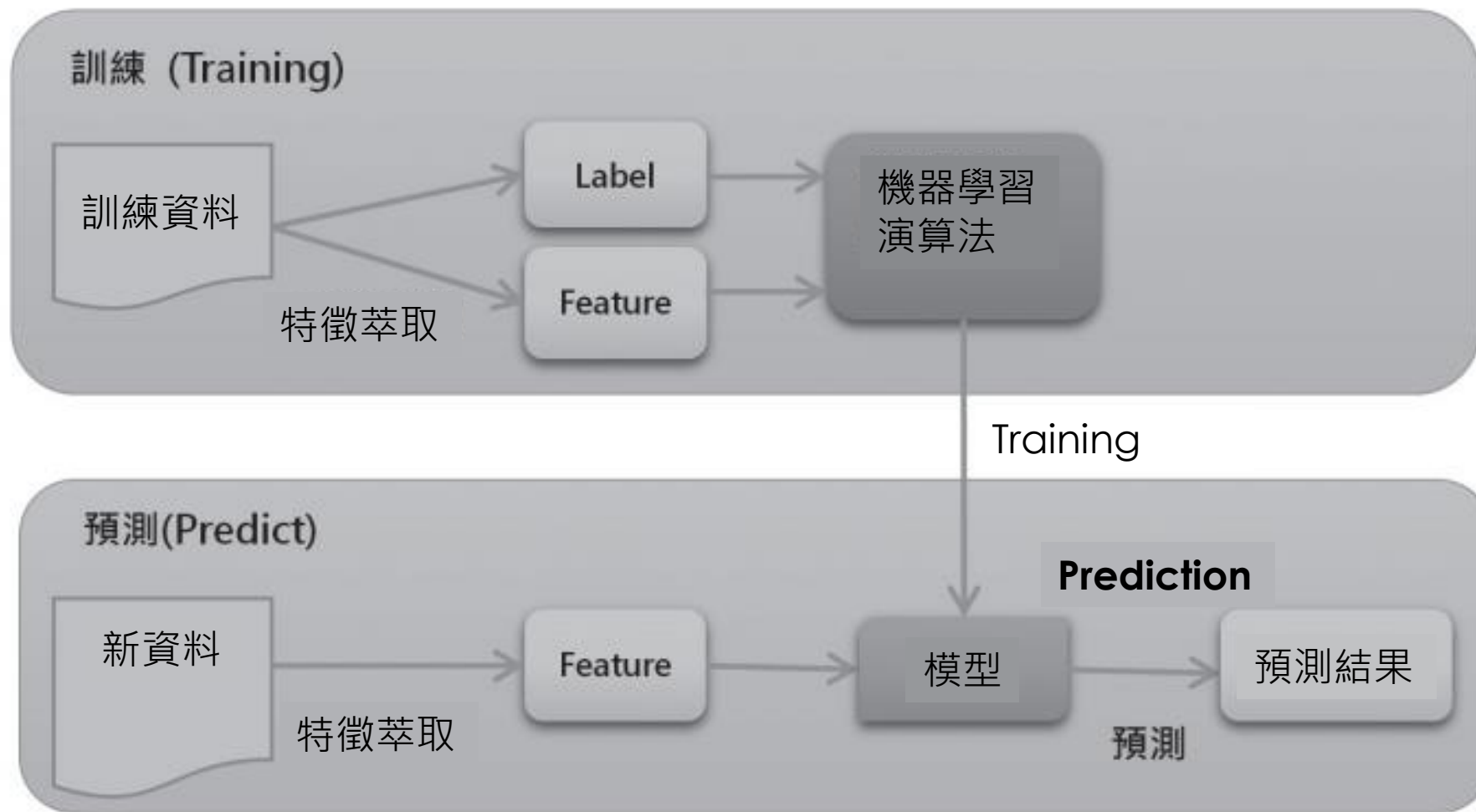
71%

B2B marketers are interested in using AI for personalization.

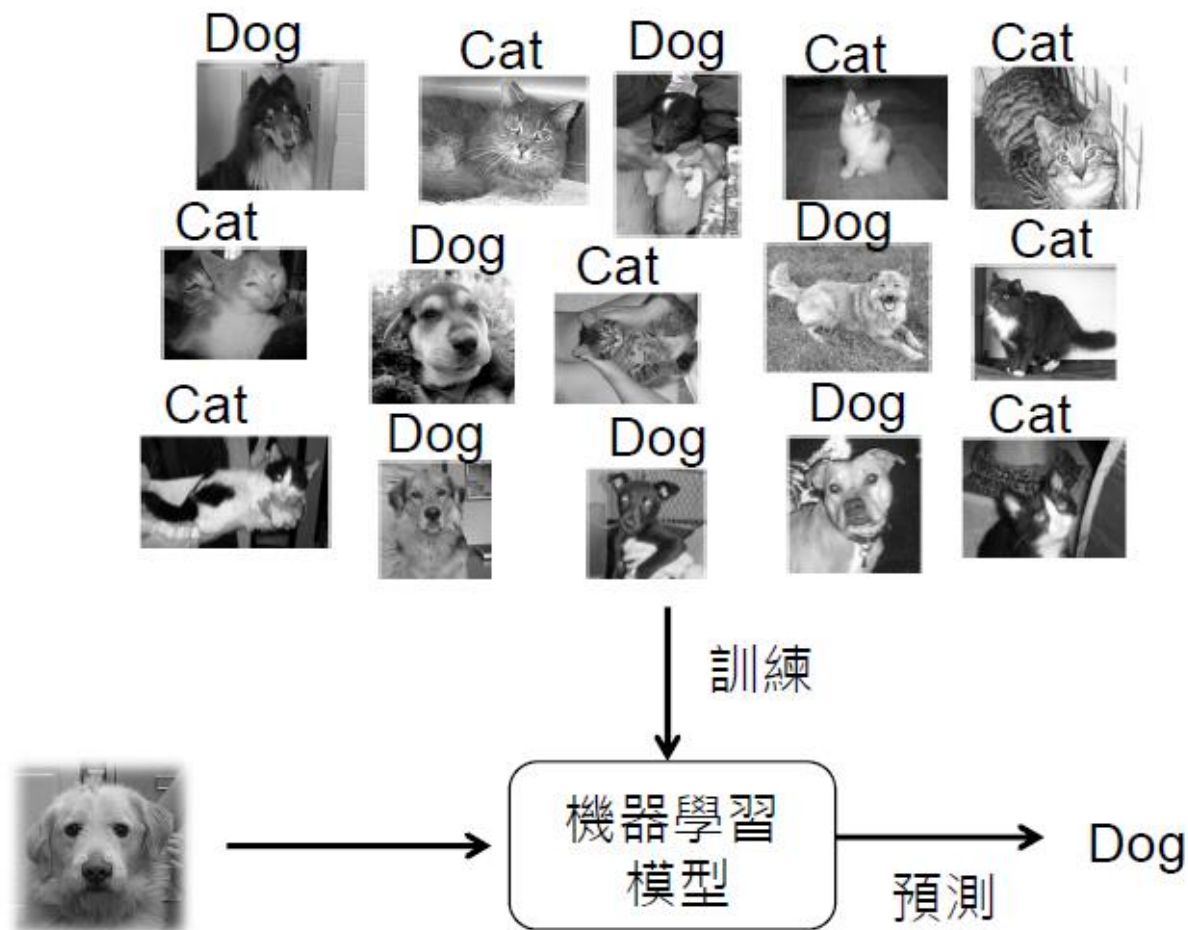
1.2 機器學習介紹

- ML透過演算法，使用歷史資料進行訓練，藉此產生特定模型
- 未來有新的資料，便可以使用特定模型進行預測
 - ML的訓練資料是由Features、Label組成
 - 資料特徵(Features)：輸入變數，例如風向、風速、氣壓、溫度
 - 資料標籤(Label)：希望預測的目標，例如天氣(1：晴天、2：雨天、3：陰天、4：下雪)

機器學習模型之建立兩階段



監督式學習 – 以分類照片為例



1.為了能進行監督式學習訓練，我們必須每一張照片加註Label(標籤)為Dog或Cat。

2.進行訓練，訓練完成後產生機器學習模型

3.訓練後，當有一張新照片，我們就可以用此模型進行預測

1.3 機器學習分類

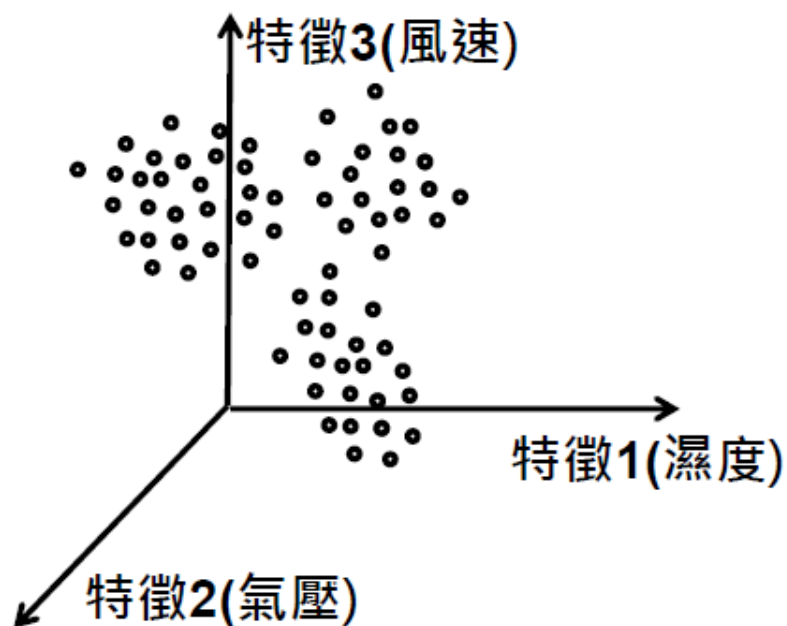
分類	細分類	Features (特徵)	Label (預測目標)
監督式學習	Binary Classification 二元分類	濕度、風向、風速、 季節、氣壓...	選項只有 0 與 1 (是非題) 0 : 不會下雨、1 : 會下雨
監督式學習	Multi-Class Classification 多元分類	濕度、風向、風速、 季節、氣壓...	有多個選項(選擇題) 1 : 晴天、2 : 雨天、3 : 陰天、 4 : 下雪
監督式學習	Regression 回歸分析	濕度、風向、風速、 季節、氣壓...	值是數值(計算題) 溫度可能是 -50 ~ 50 度的範圍

機器學習分類

非監督式學習	Clustering 群集	濕度、風向、風速、 季節、氣壓...	沒有label (Cluster 集群分析)； 目的是將資料依照特徵，分成幾個相異性最大的群組，而群組內的相似程度最高
強化學習	Q-learning、 TD (Temporal Difference)		強化學習的原理，藉由定義：動作(Actions)、狀態(States)、獎勵(Rewards)的方式，不斷訓練機器循序漸進，學會執行某項任務的演算法，常用於動態系統及機器人控制等。

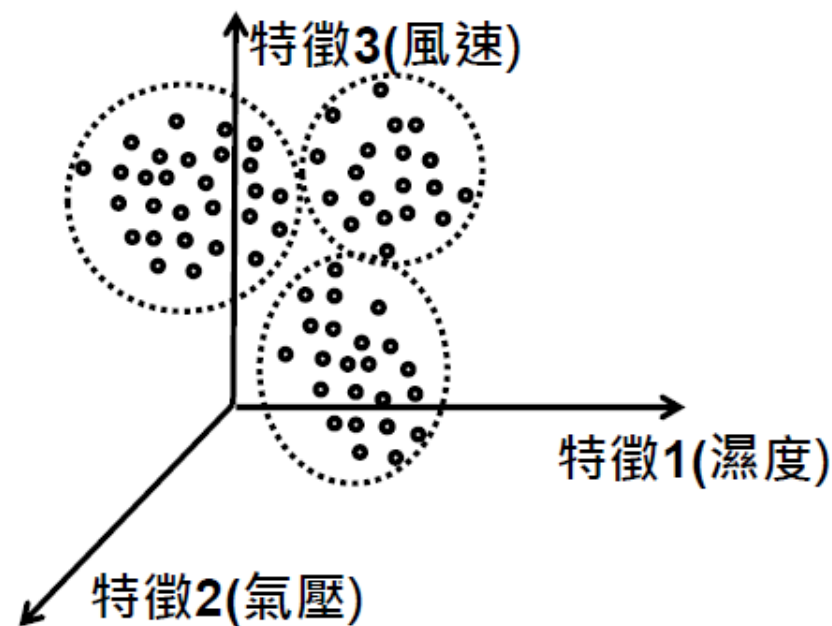
非監督式學習

例如：下列只有features特徵：
濕度、風向、風速、季節、氣
壓，無label標籤

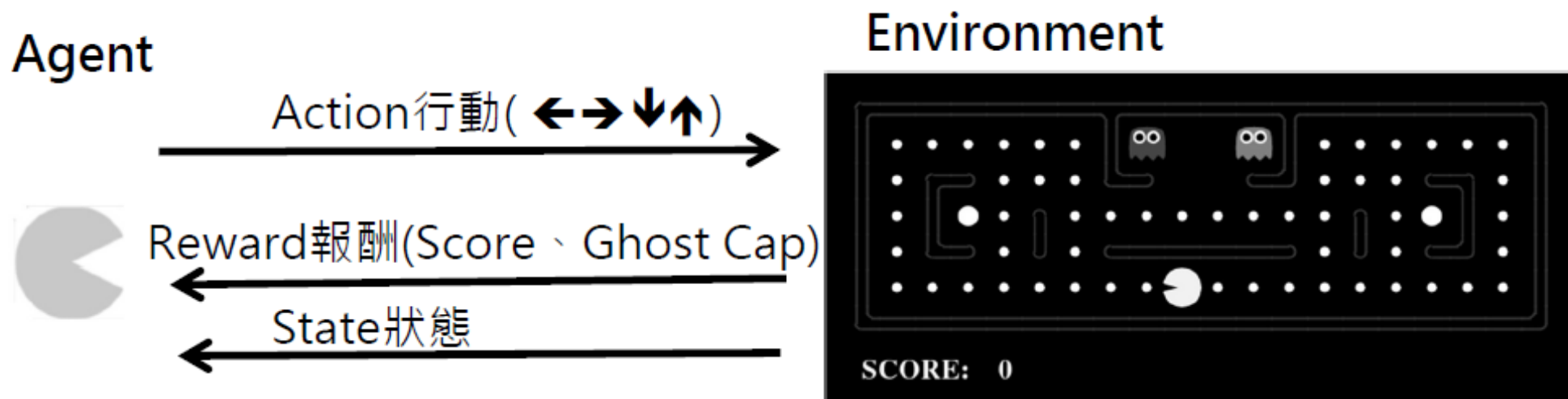


Clustering
演算法

clustering 集群演算法，將資料依照
特徵，分成幾個群組，群組內的相似
程度最高，群組之間的相異性最大。



強化學習



強化學習(reinforcement learning)的原理，藉由定義:動作(Actions)、狀態(States)獎勵(Rewards)的方式，不斷訓練機器循序漸進，學會執行某項任務的演算法。

例如訓練機器玩小精靈:

- 動作:左/右/上/下
- 狀態:目前遊戲的畫面
- 獎勵:得分/被抓到

機器學習分類

機器學習Machine Learning

監督式學習Supervised Learning

分類Classification

二元分類Binary

多元分類 Multi Class

回歸分析Regression

線性回歸Linear

DNN Regression

非監督式學習

Unsupervised Learning

Clustering
AutoEncoder

強化學習

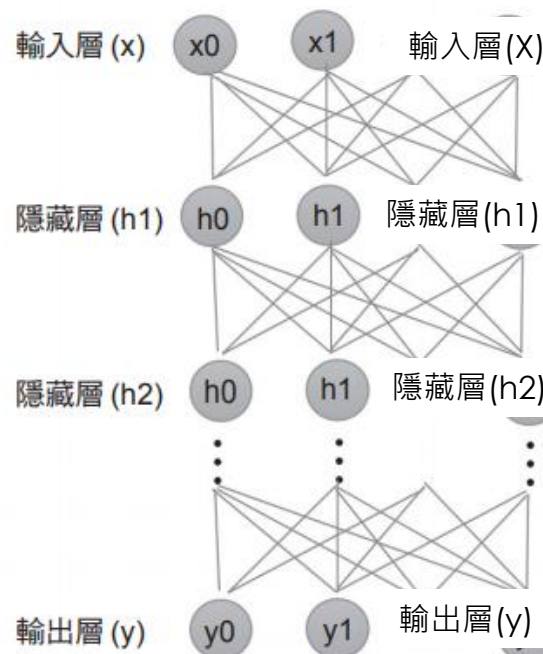
Reinforcement Learning

Q-learning
Deep Q-learning

1.4 深度學習簡介

人腦的重量大約一公斤多,結構非常複雜,預估具有860億個神經元,以及超過100兆條的神經相連,形成的網路比最先進的超級電腦還要強大。

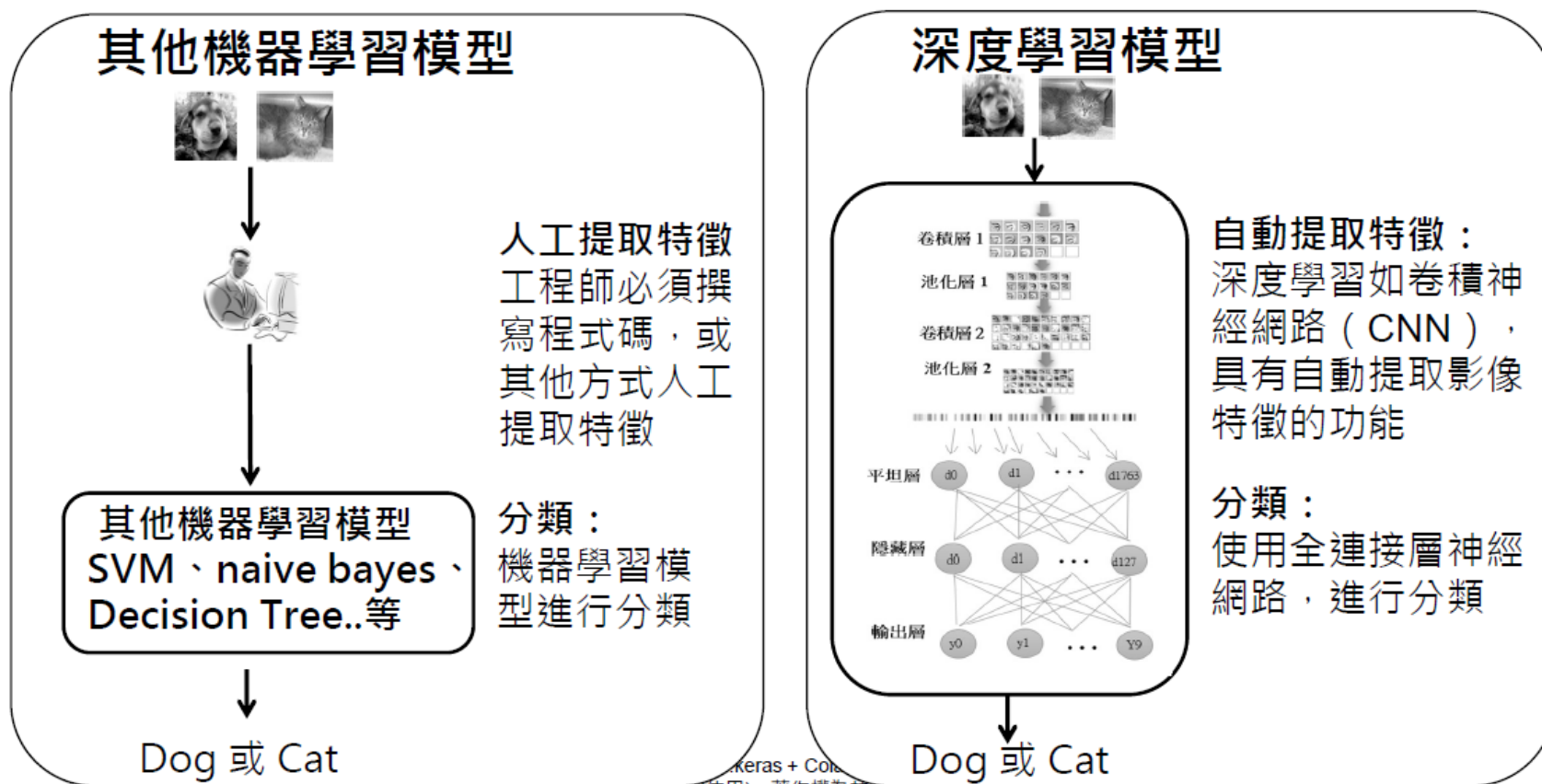
但是因為人類神經網路太過複雜,為了方便以電腦模擬,將神經元分為多層次,來模擬神經網路通常會有1個輸入層、1個輸出層、隱藏層可以非常多層,所以稱為深度學習。



隱藏層可以非常多層，
所以稱為深度學習

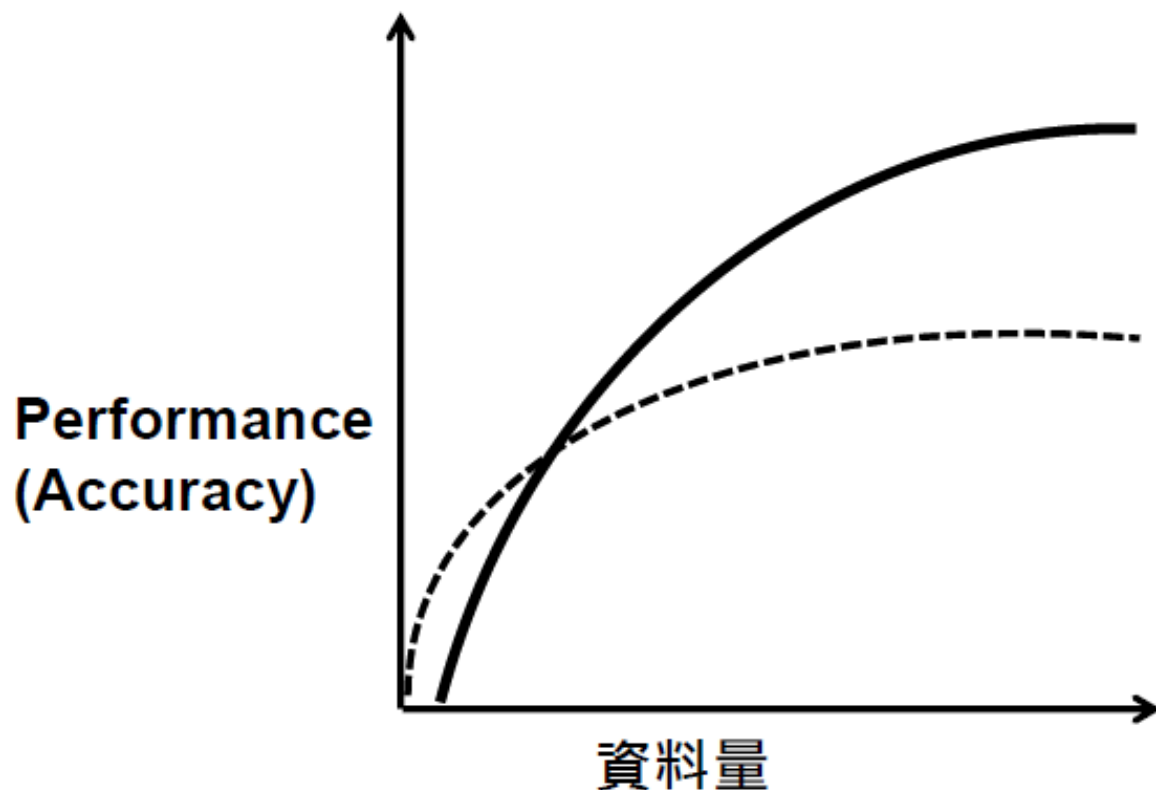
深度學習的特性

深度學習模型能自動提取特徵



深度學習的特性

- 資料量越大，深度學習表現越好



深度學習模型：

當資料量不大時，表現普通，但是資料量越大。深度學習模型表現越來越好。

其他機器學習模型：

(例如：SVM、naive bayes、Decision Tree..等) 當資料量不大時，表現還不錯，但是資料量越大時，其準確率無法繼續提昇。

為什麼使用類神經建立DL網路

- 類神經網路在處理複雜的工作時
 - 不需要針對問題定義複雜的數學模式
 - 不用去解任何微分方程、積分方程或其他數學方程式
 - 藉由學習來面對複雜的問題與不確定性的環境
- 使用ANN的原因
 - 聯想速度快（資料輸出比對或預測）
 - 網路架構容易維持
 - 具平行處理能力
 - 解決最佳化、非線性系統問題
 - 具容錯特性（減少雜訊干擾）

類神經網路的特性

- 類神經網路擁有許多與人類大腦相同功能的特性，其中最重要的三項：(1) 學習 (2) 回想 (3) 歸納推演

(1) 學習 (Learning)

- 學習方法

建立神經元間的連結模式、修正連結神經元間的權重、調整神經元活化函數中的門檻值、或是結合運用上述幾種方式

- 演算法可分成

- 即時線上學習 (on-line learning)

依據新增或即時的資訊，不斷地修正調整網路的連結權重值

- 批次學習 (batch learning or off-line learning)

先將網路訓練好才能開始使用網路，所有的訓練資料是一次一整批的用於修正網路的連結權重

類神經網路的特性

(2) 回想 (Recall)

- 當類神經網路接受到一個輸入刺激，進而依據網路架構產生一個輸出值，稱之為回想過程
- 像看到一個人的影像或聽到一個人聲音的同時，即可判斷認不認識這個人、這個人是誰、叫什麼名字等
- 回想的過程是否快速有效，影響著一個類神經網路的效能及強健性

類神經網路的特性

(2) 回想 (Recall)

- 前饋式類神經網路

只反映出目前輸入資料，直接從網路儲存的案例找到一對應的輸出，屬於瞬間表現的回想過程

- 回饋式類神經網路

輸入項與輸出項之間具互相連結的作用

時刻(t) 的輸出結果會回饋到下一時刻($t+1$)成為輸入條件

因為映射關係不斷被修正與調整，故此一系統可視為動態系統

常用於描述非線性動態系統的行為

類神經網路的特性

(3) 歸納推演 (Generalization)

在哲學和邏輯學中

- 演繹推理(deductive reasoning)

「從已知為事實的前提，『必然的』得出結論的推理」
如果前提為真，則結論必然為真

- 歸納推理(inductive reasoning)

「論證的前提支持結論，但『不確保』結論的推理」
基於對特殊代表的有限觀察，把性質或關係歸結到類型
或基於對反覆再現的現象的模式(pattern)的有限觀察，以公式表達規律

類神經網路的特性

(3) 歸納推演 (Generalization)

- 從一個系統中局部觀察描述出其整體特性的過程
- 不論是從特殊例子推演到整體的事件，或是從範例的認知去定義出物件的種類，稱為歸納推演過程
- 讓我們找出所有可以適用於全部類別物件的重要因子
- 而非僅將記憶中的特殊事件運用於個別物件
- 提供一個有效率的記憶與儲存模式
- 若是沒有歸納推演過程，我們便可能需要以記憶與回想組合出無限個特殊事件、因子與關係

類神經網路的特性

(3) 歸納推演 (Generalization)

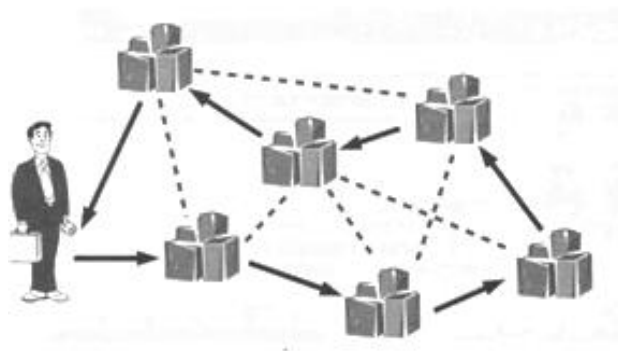
- 類神經網路對輸入資料，具有特徵萃取的能力
- 例如，可以訓練ANN辨識一組扭曲的字母 “A”，然後網路能夠辨識出另一組扭曲的字母 “A”
- 也就是可以對過去訓練範例中沒見過的輸入作出正確的回應
- 經過訓練的ANN，可以對曾見過但有些微差異的輸入訊號作出正確的輸出
- 是網路模式中不可或缺的部分

適合使用類神經網路解決的問題

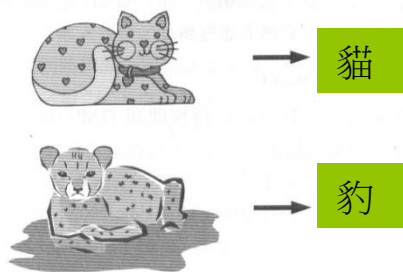
- 適合使用類神經網路解決的問題通常有以下的特徵
 - 問題及相關條件難以完整定義
 - 需要快速得到問題解答且解答不用完全精確
 - 問題非常複雜或是「非線性」的問題，無法由一連串已知的數學方程式來描述並求得解答者

- 這些問題可歸納為下面幾類：

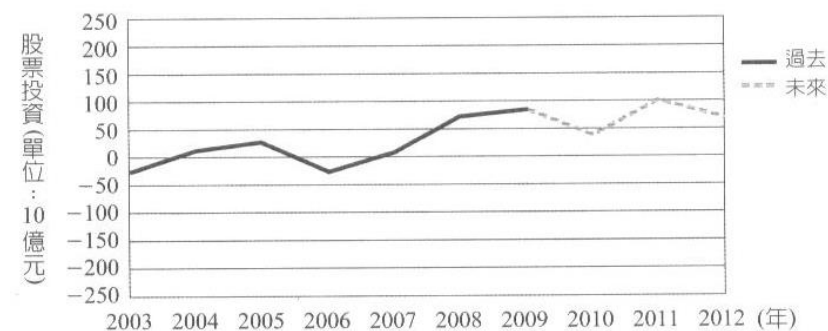
1. 最佳化 (optimization)



2. 辨識/分類 (recognition /classification)



3. 預測 (prediction)

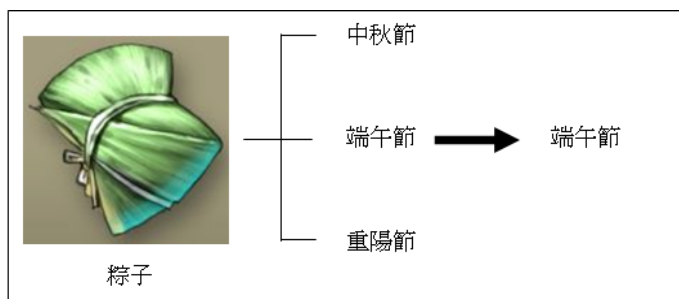


4. 評估、診斷/決策 (diagnosis /decision)

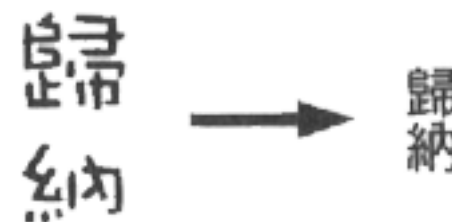
發燒 2-3 天，全身酸痛、頭痛、喉嚨痛，眼睛痛與疲倦	➡ 感冒
發燒、咳嗽、喉嚨痛、全身酸痛、頭痛、寒顫與極度倦怠，有些病例出現腹瀉、嘔吐及流鼻涕症狀	➡ H1N1

- 這些問題可歸納為下面幾類：

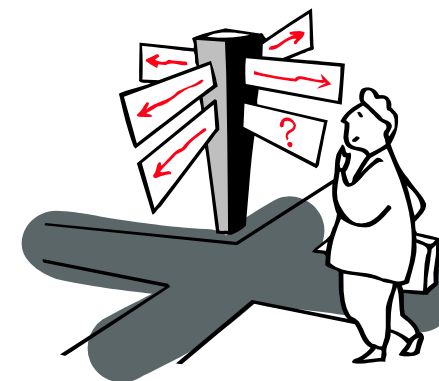
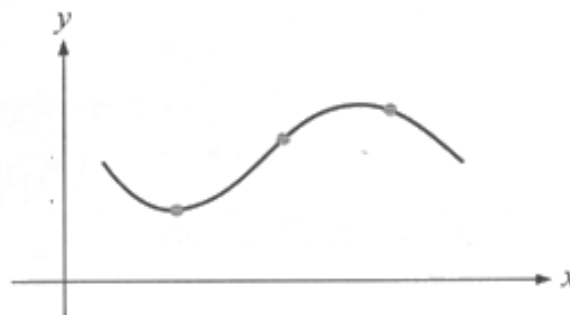
5. 聯想 (association)



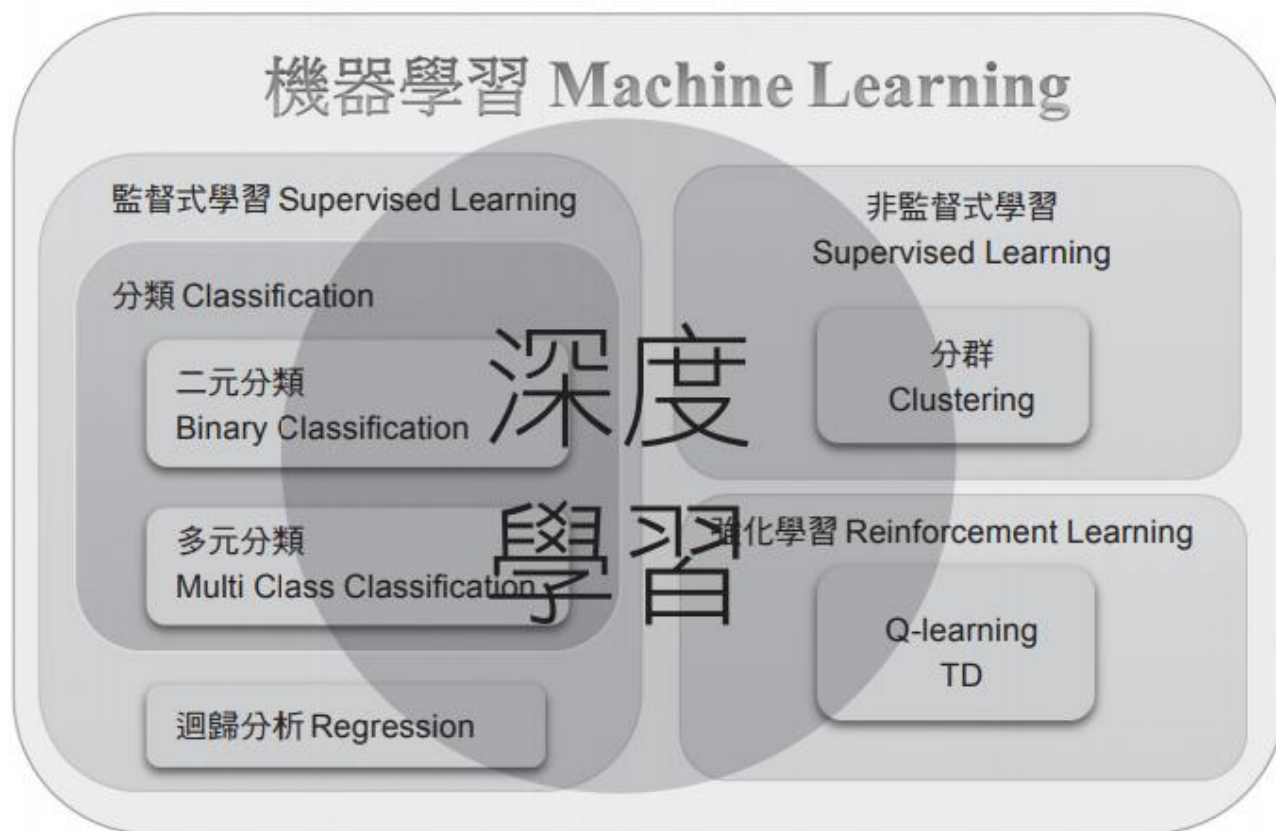
7. 歸納推演 (generalization)



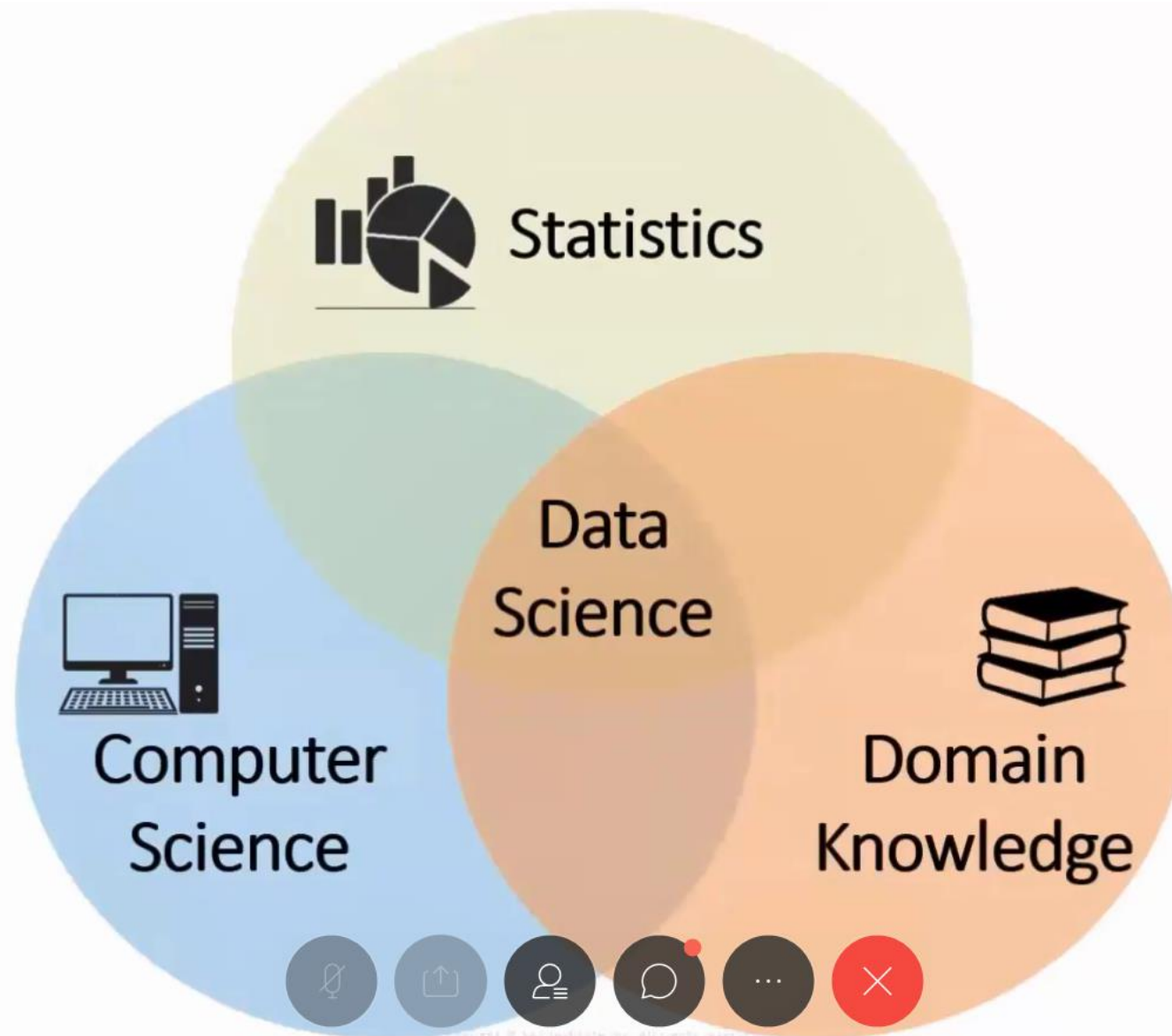
6. 近似 (approximation)



機器學習與深度學習的關係



如上圖,深度學習的應用很廣泛,你可以將深度學習技術,應用在監督式學習、非監督式學習、強化學習等領域。

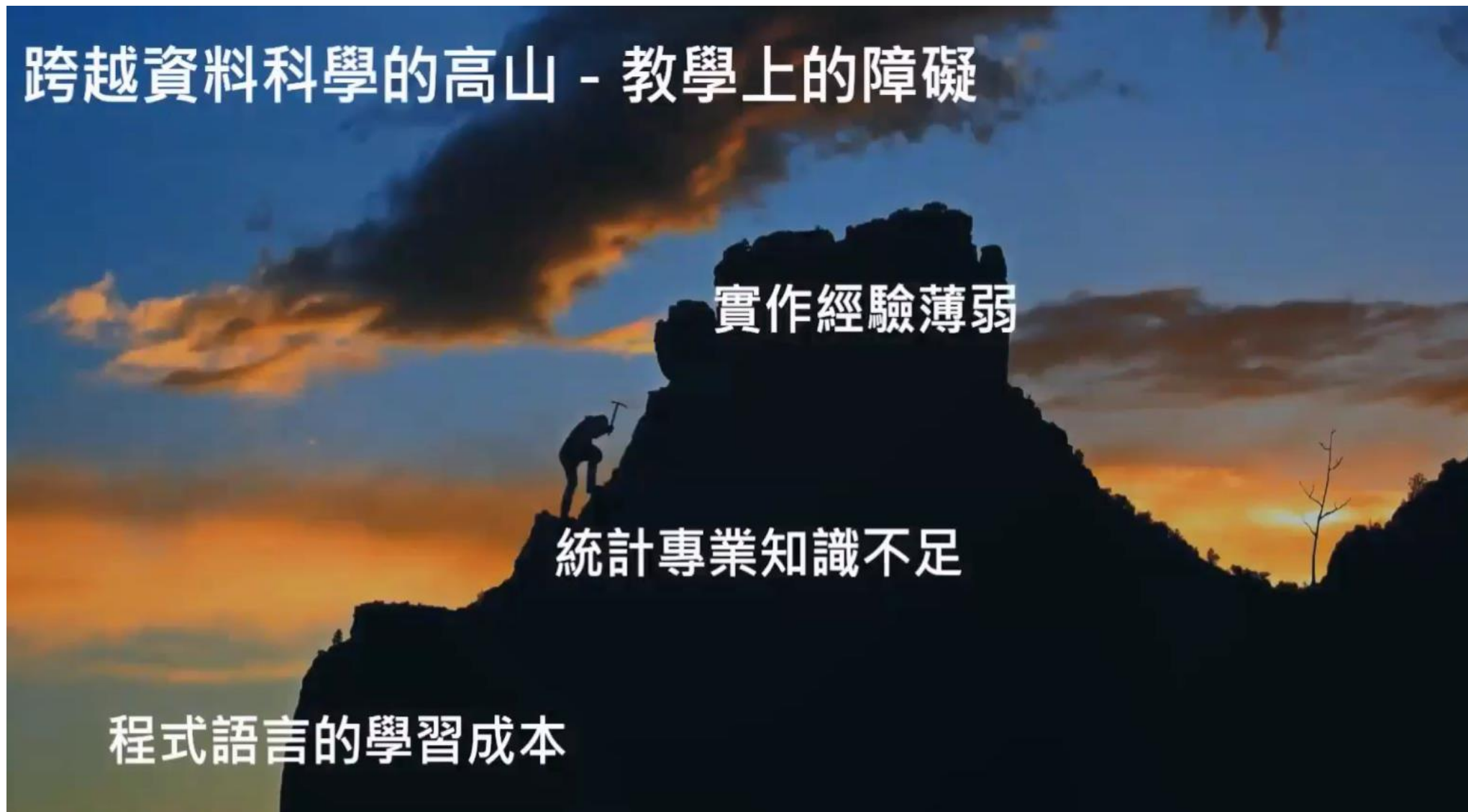


跨越資料科學的高山 - 教學上的障礙

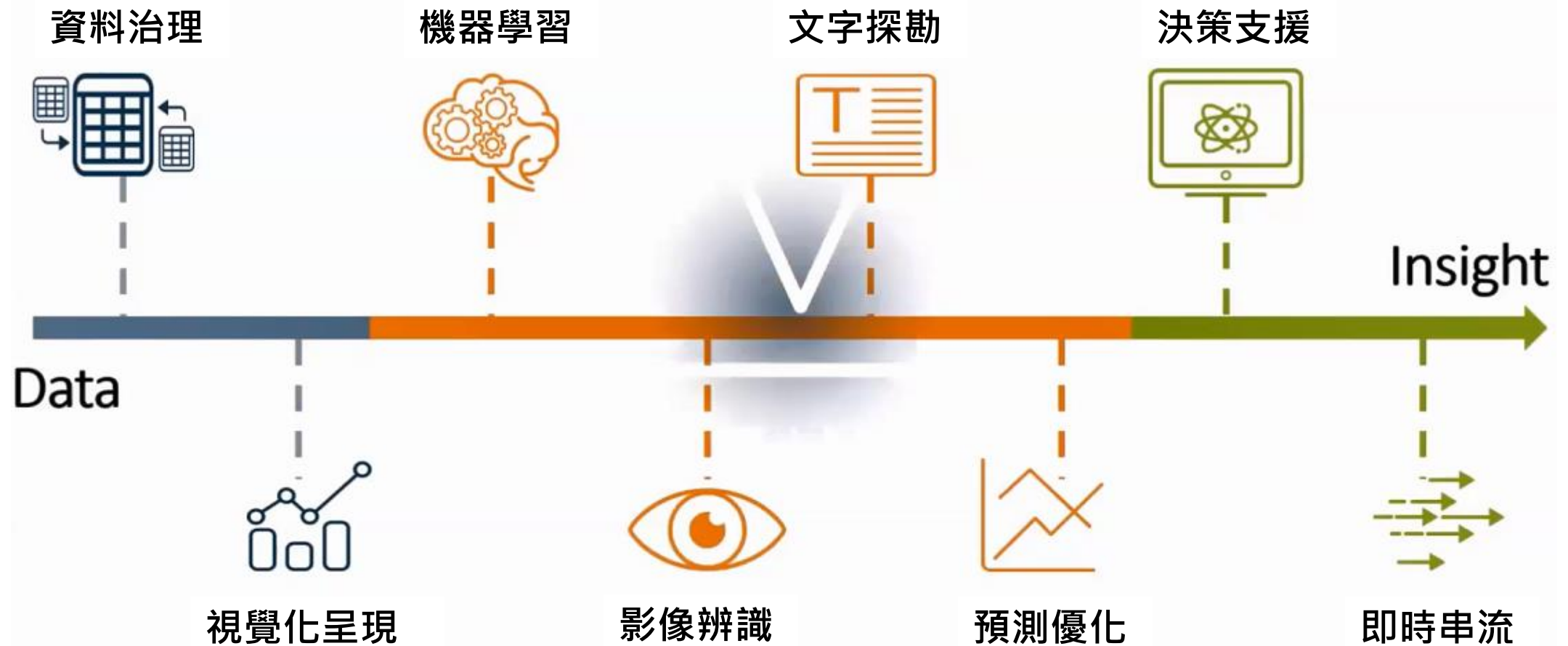
實作經驗薄弱

統計專業知識不足

程式語言的學習成本



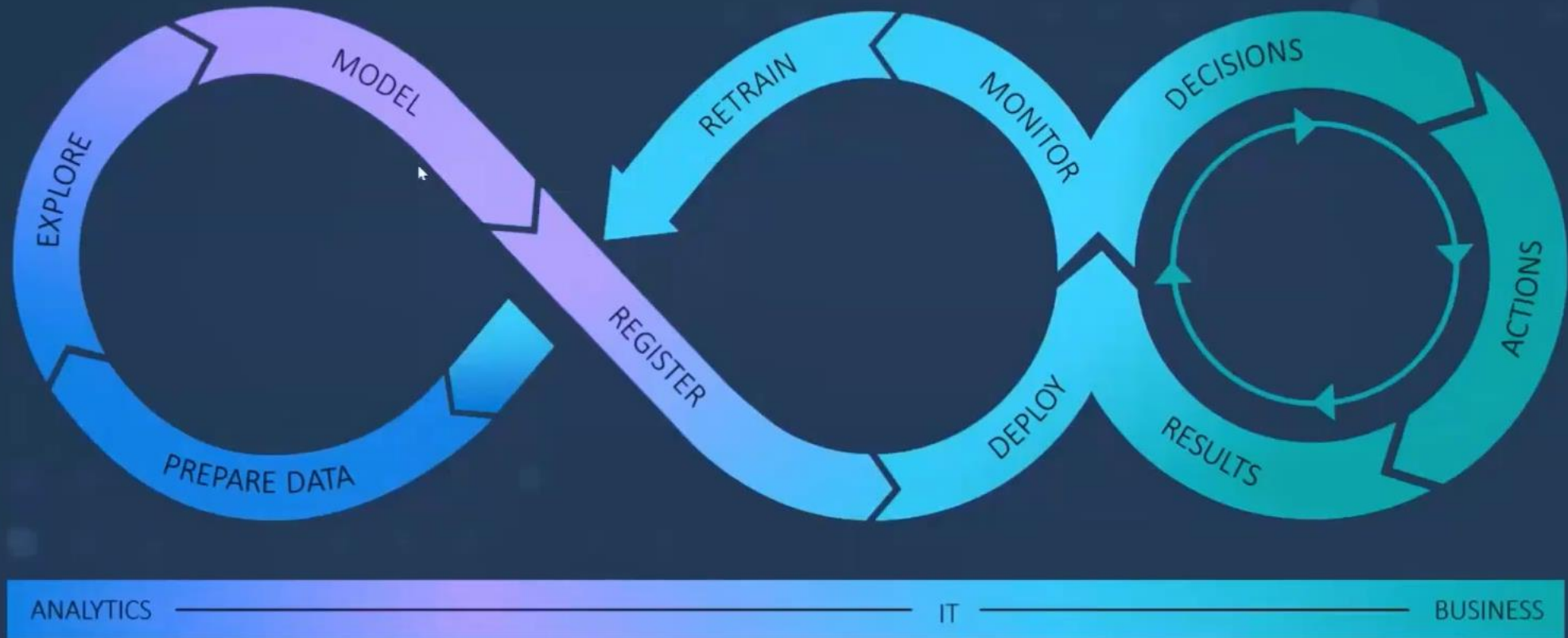
資料科學旅程



企業級AI特色



學生畢業後需面對的資料分析生命週期



結論

- 本章介紹AI、ML、DL、ANN的關係以及基本概念
- DL不同於傳統數學難以建立函式關係的困境，可藉由學習與記憶的方式來處理複雜且具不確定性的問題
- 以大量平行分布的計算單位(神經元)所連結組成多層級的網路架構
- 具有從輸入的環境訊息中獲取並累積經驗、儲存知識，進而能加以利用的智慧型演算程序
- ANN在處理樣本比對、分類、函數近似、最佳化(optimization)及資料聚類等，都有很好的效果