2024 鐵人賽 – 我數學就爛要怎麼來學 DNN 模型安全
Day 01 – 機器學習是幹啥的?

大綱

- 機器學習
  - 定義
  - 運作方式及流程
  - 類型
- 結論

機器學習 - 定義

https://aws.amazon.com/tw/what-is/machine-learning/

- 機器學習是一門開發演算法和統計模型的科學，這些算法和模型**不需要由人類輸入精確的指令**，而是可以讓電腦系統根據模式和推理來執行任務

- 電腦系統使用機器學習演算法**處理大量的歷史資料**，並從中**找出資料的模式**，這讓電腦能更精準地根據輸入的資料集預測結果

機器學習 – 舉例

- 假設今天有個商店統計了數據，想知道來客數跟他們獲利的關係

來客數	5	3	2	1	7	10	4
獲利	10	6	4	2	14	20	8

- 透過觀察後會得出以下程式碼

```
def count_profit(num_customer) :
    return num_customer*2
```

機器學習 – 舉例

- 但現實的問題背後的關聯是很複雜的，比方說如下：

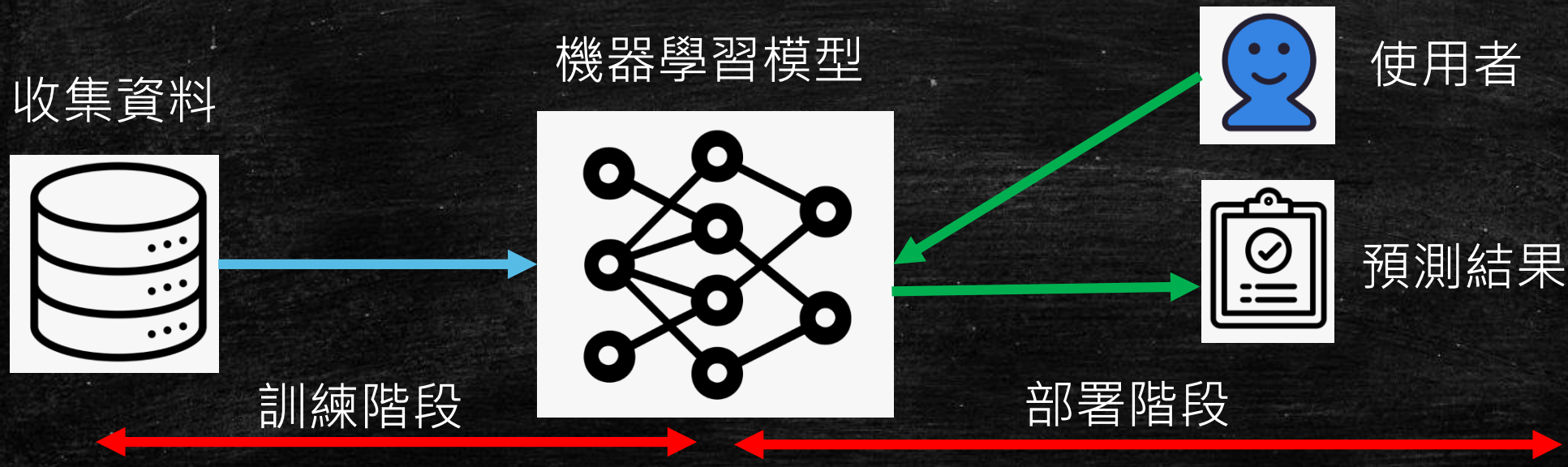
來客數	5	3	2	1	7	10	4
獲利	211	43	13	3	603	1821	105

- 這邊的公式是 $y=x^3 - 2x^2 + 2x + 1$

- 這時候以人工方式要有效率的得到背後的關聯是困難的，所以才會有機器學習的使用的需求

機器學習 - 運作方式及流程

- 機器學習的核心概念是找到**輸入與輸出資料對之間的數學關聯**，但隨著給予其足夠的資料，模型的預測會越來越準確
- 通常會分為兩個階段，一個是**訓練學習階段**，另一個則是學習完後的**部署階段**



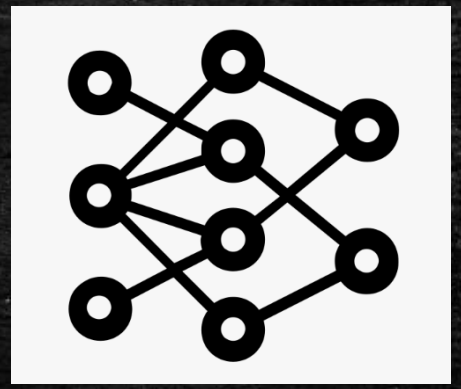
機器學習 – 類型

- **監督式機器學習**
  - 監督學習需要訓練資料有相對應的標記才能進行
- 非監督機器學習
  - 無監督學習演算法使用未標記的資料進行訓練，多半以分群的方式呈現結果

演算法 [編輯]

具體的機器學習演算法有：

- 構造間隔理論分布：聚類分析和圖型識別
  - 類神經網路
  - 決策樹
  - 感知器
  - 支援向量機
  - 整合學習AdaBoost
  - 降維與度量學習
  - 聚類
  - 貝氏分類器



結論

- 類神經網路 Deep Neural Network (深度神經網路) 是機器學習的一種，因此必須先理解其運作方式及開發流程，後續才有辦法判斷有可能產生安全風險的位置

- 在切入類神經網路之前，明天會先來點簡單的數學熱身，內容包含簡單的線性代數、微積分