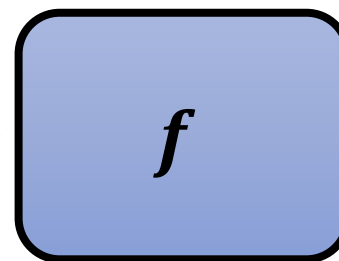


機器學習 基本概念

機器學習 \approx 機器自動找一個函式

ChatGPT

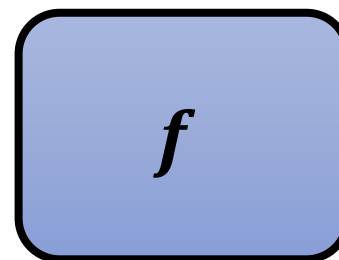
什麼是機器學習？



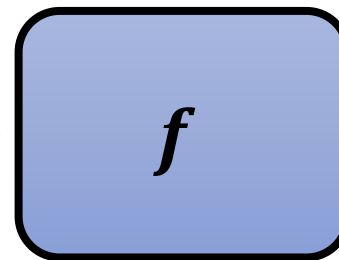
“機”

Midjourney

一隻可愛的貓



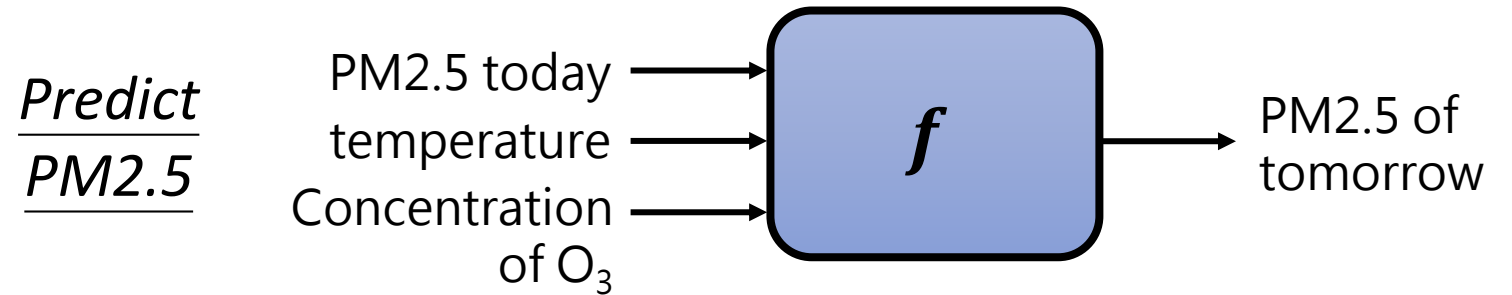
AlphaGo



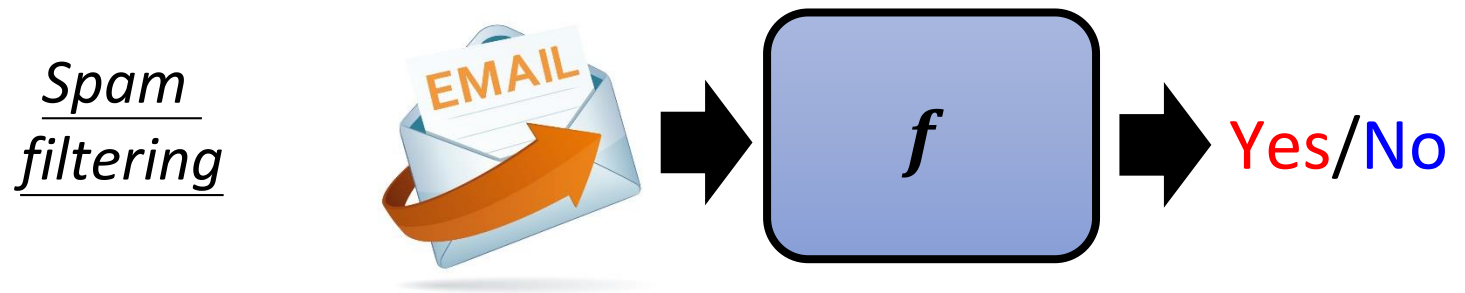
“5-5”

根據函式輸出來分類

- 迴歸 (Regression) : 函式的輸出是一個數值



- 分類 (Classification) : 函式的輸出是一個類別 (選擇題)



Structured Learning

- 生成有結構的物件
(例如：影像、文句)
- 又叫做「生成式學習」
(Generative Learning)

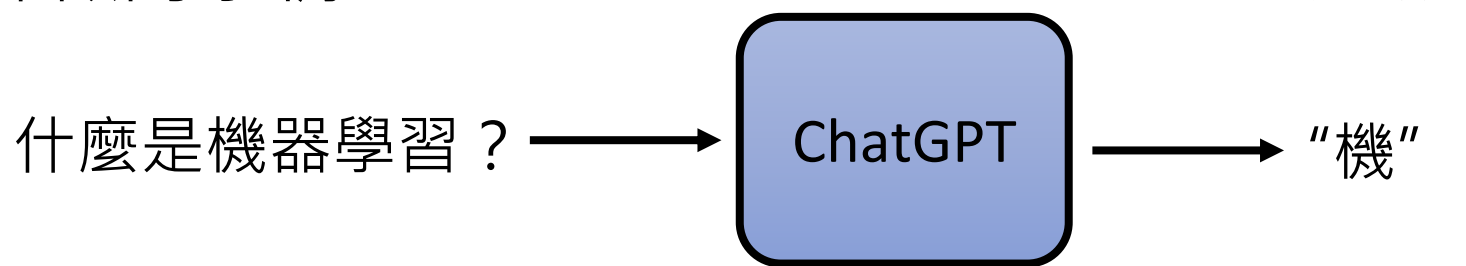


Regression,
Classification

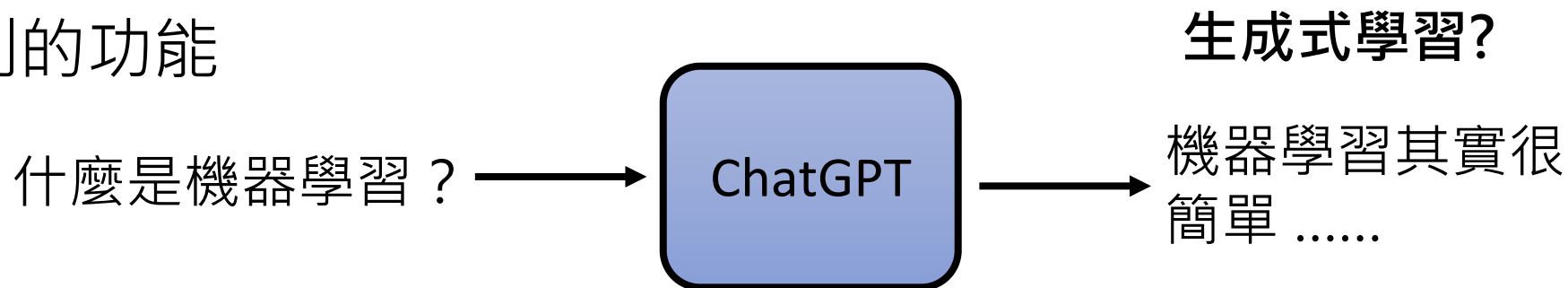
(圖片來源：漫畫獵人)

ChatGPT 是哪一類呢？

- ChatGPT 實際做的事情



- 使用者感受到的功能



把生成式學習拆解成多個分類問題

找出函式的三步驟

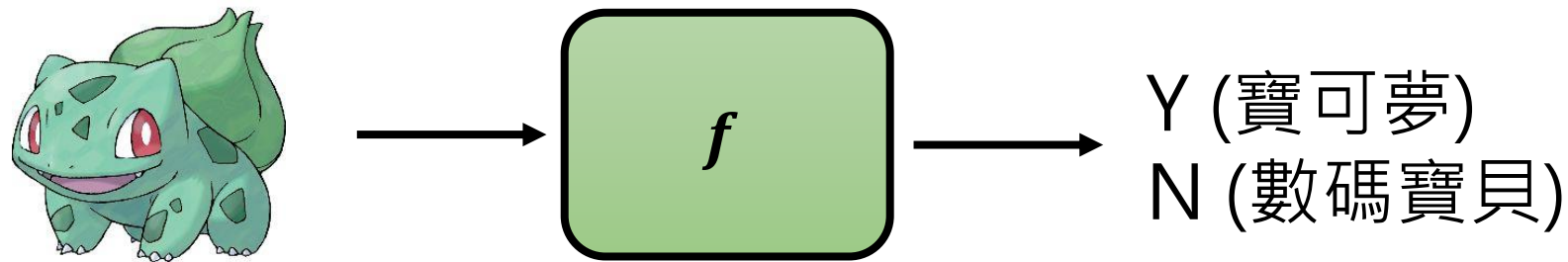
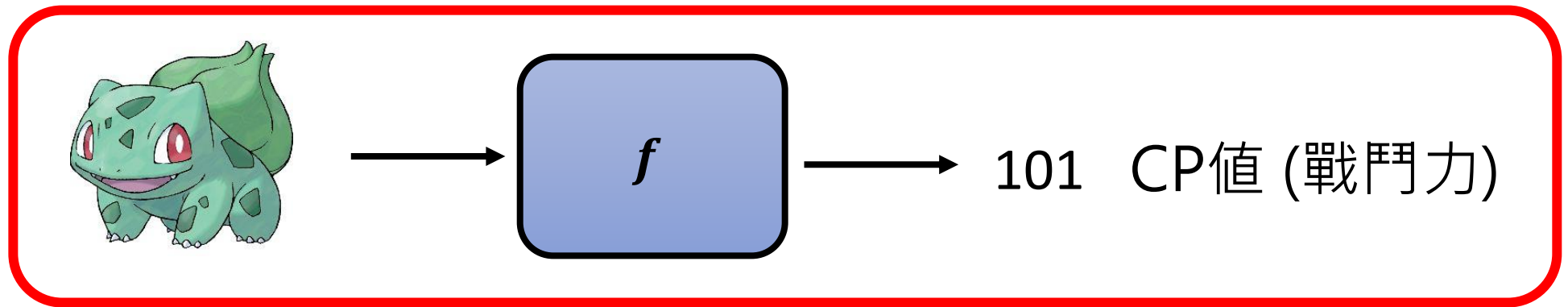
設定範圍

設定標準

達成目標

前置作業：決定要找什麼樣的函數

- 與技術無關，取決於想要做的應用



找出函式的三步驟

Model

設定範圍

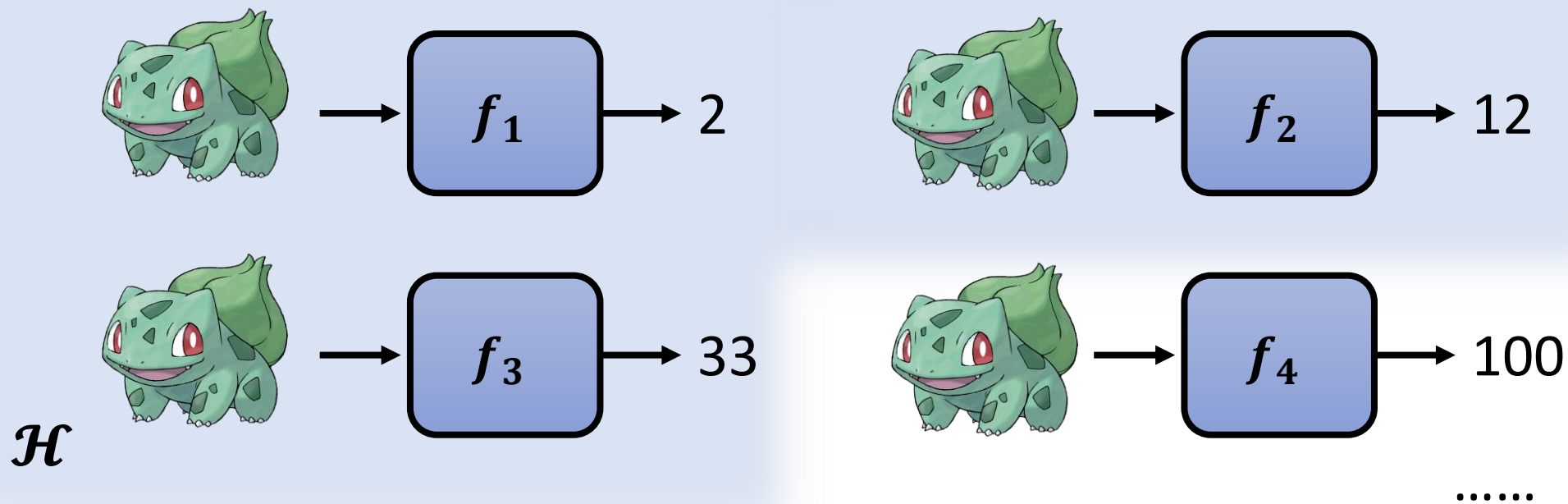
訂出候選函式的集合

設定標準

達成目標

1. 訂出候選函式的集合

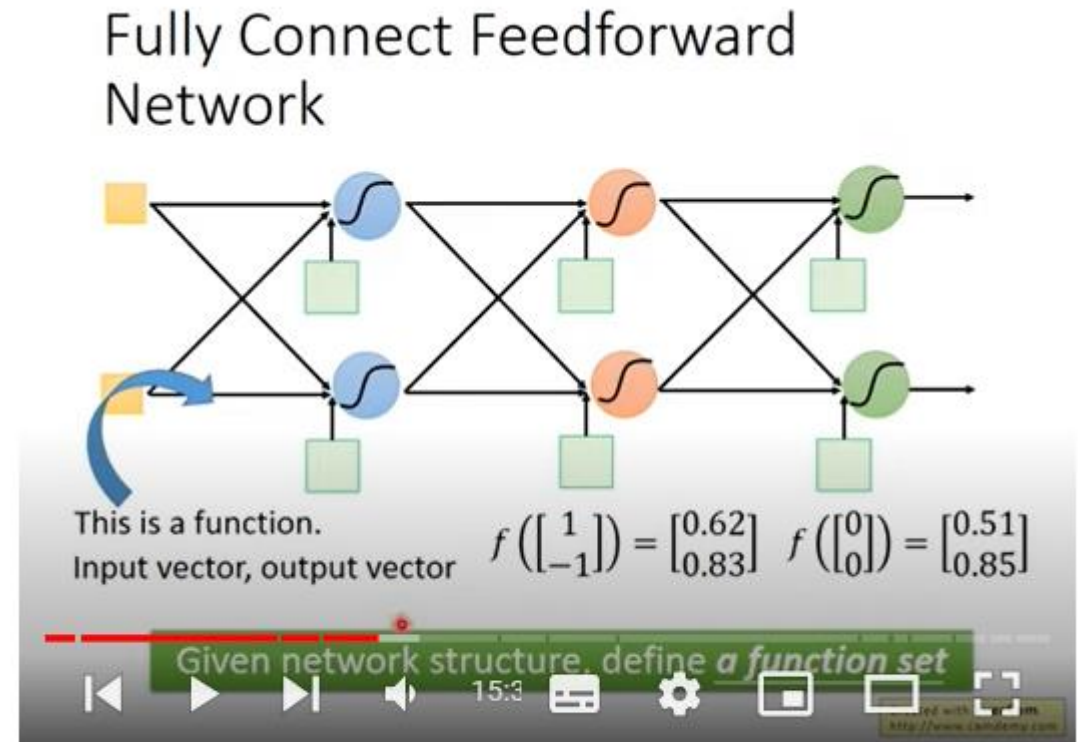
- 深度學習中類神經網路的結構 (例如：CNN, RNN, Transformer 等等) 指的就是不同的候選函式集合



某個類神經網路的結構

1. 訂出候選函式的集合

- 深度學習中類神經網路的結構 (例如：CNN, RNN, Transformer 等等) 指的就是不同的候選函式集合



<https://youtu.be/Dr-WRIEFefw>

找出函式的三步驟

設定範圍

Model

訂出 候選函式的集合

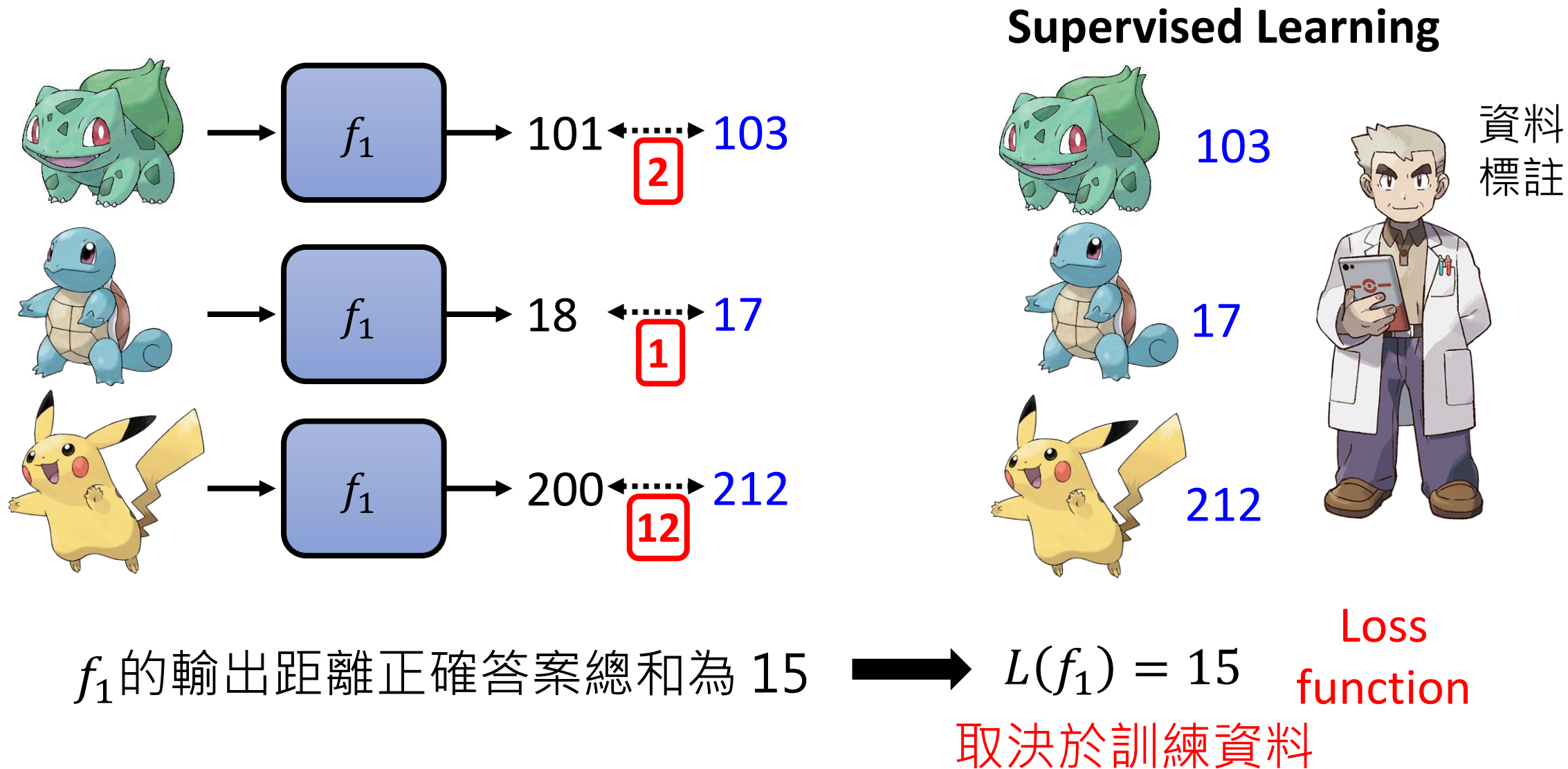
設定標準

Loss

訂出 「評量函式好壞」的標準

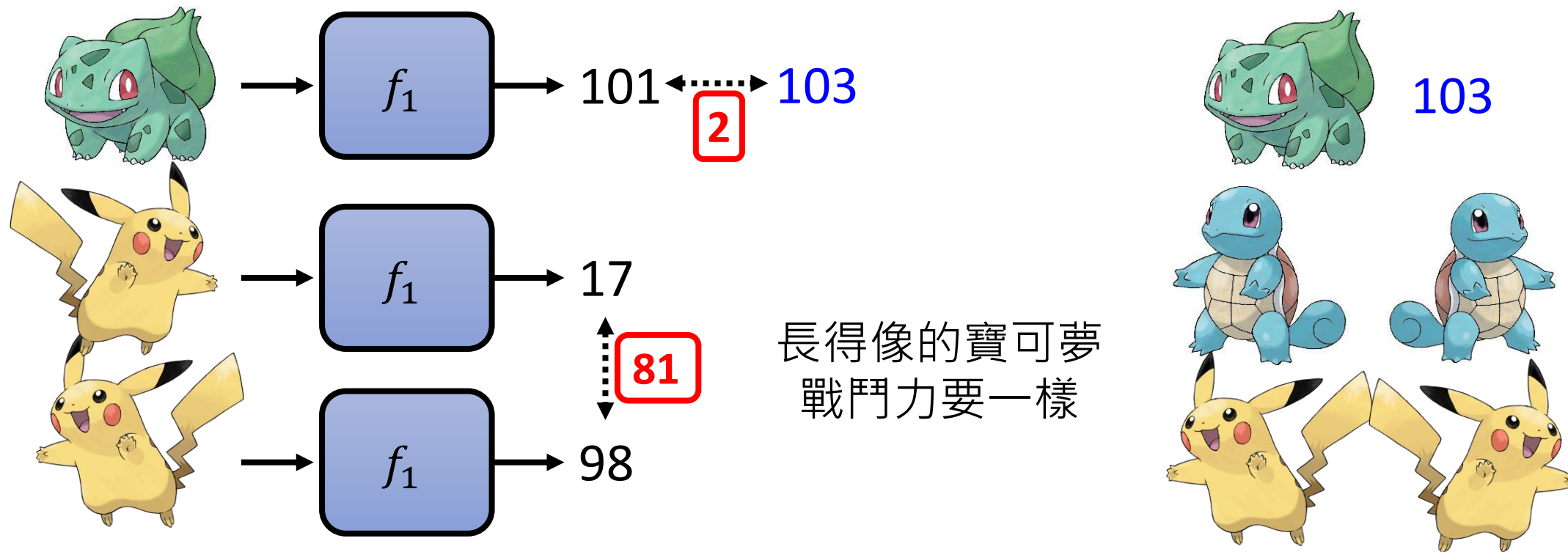
達成目標

2.訂出「評量函式好壞」的標準



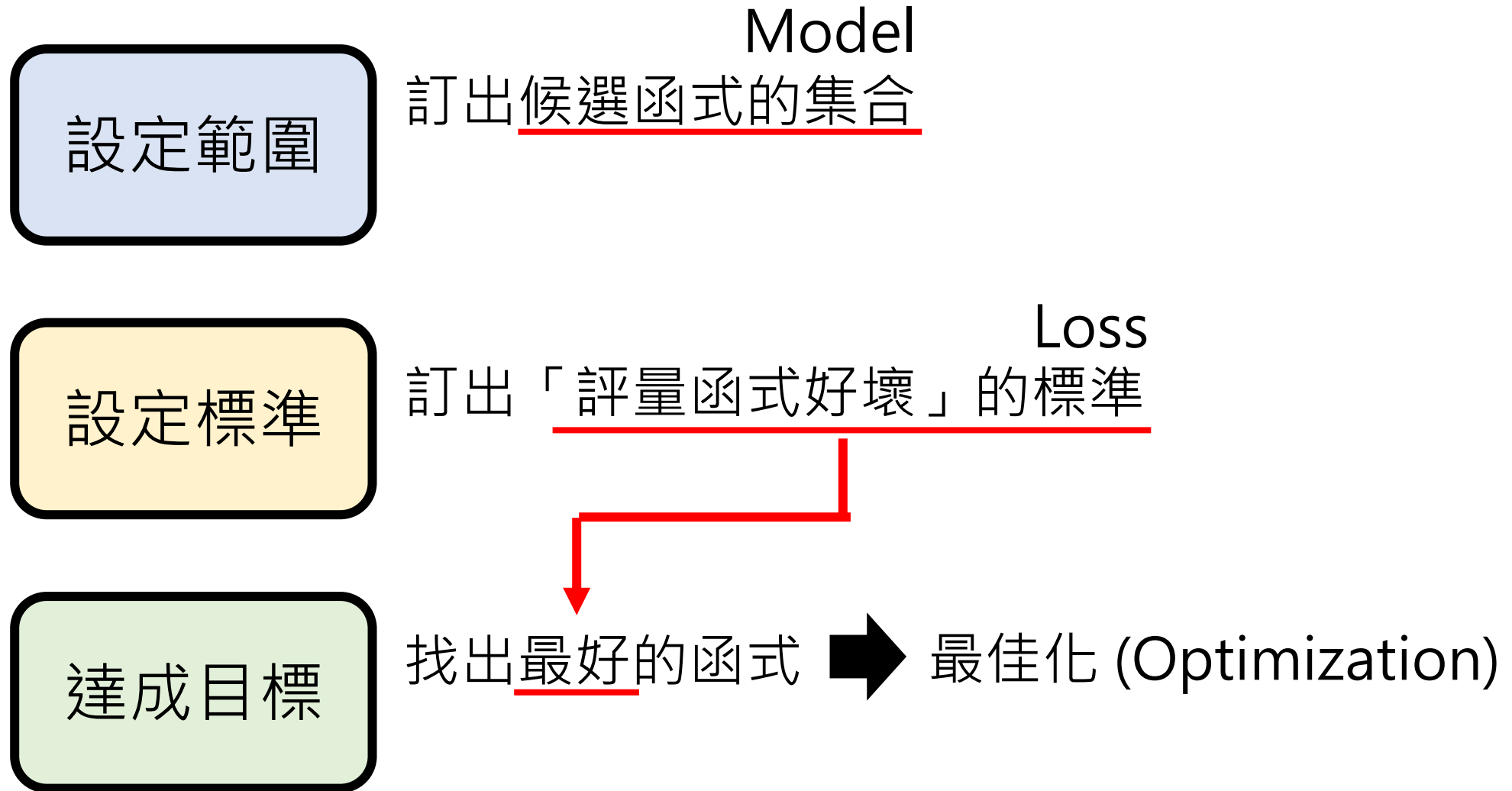
2.訂出「評量函式好壞」的標準

Semi-supervised Learning

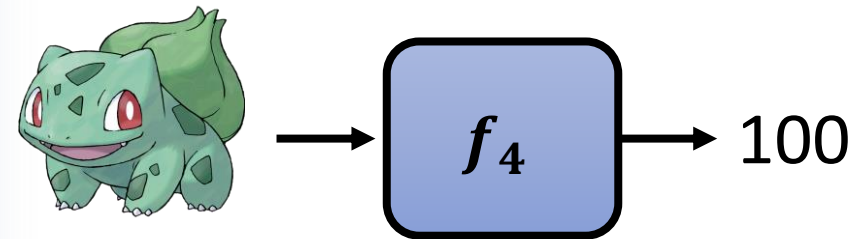
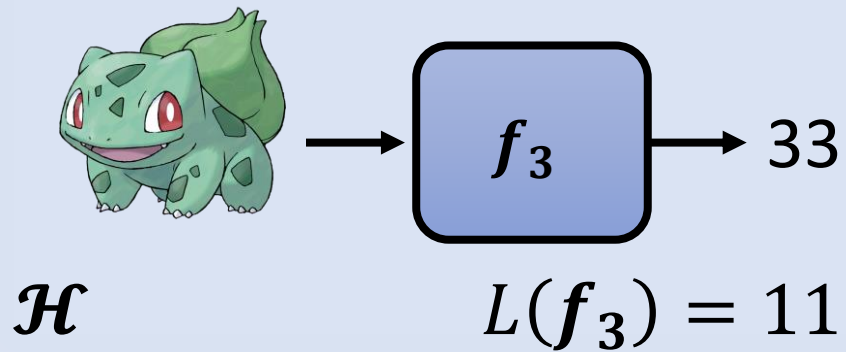
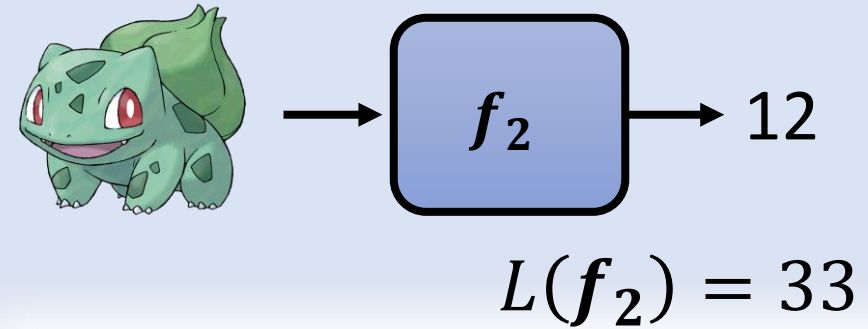
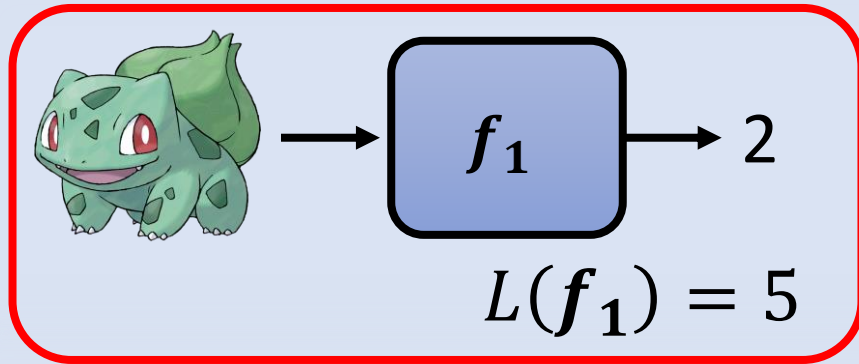


$$L(f_1) = \text{輸出距離正確答案} + \text{長得像的寶可夢差距}$$

找出函式的三步驟



3. 找出最好的函式

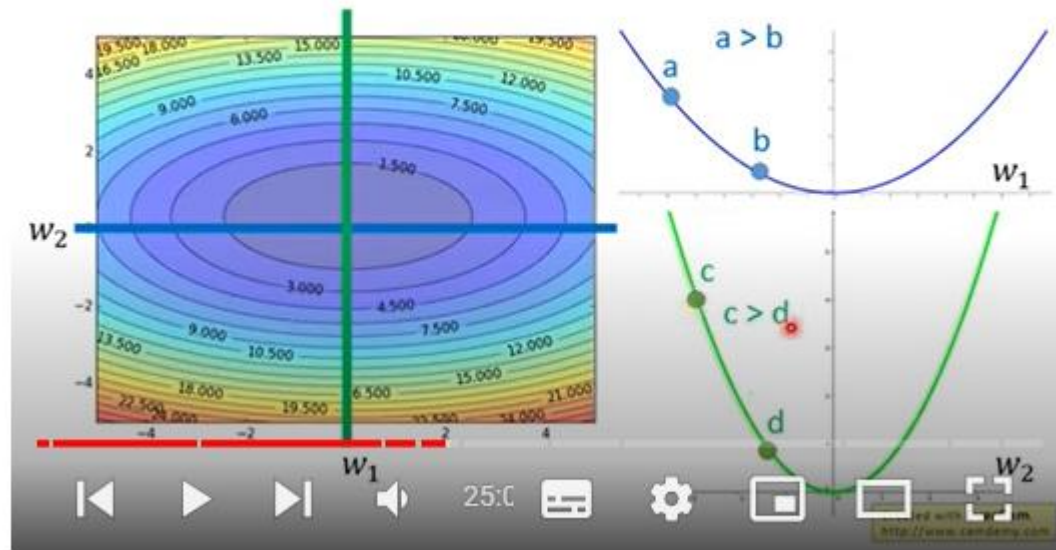


$$f^* = \arg \min_{f \in \mathcal{H}} L(f)$$

3. 找出最好的函式

Comparison between different parameters

Larger 1st order derivative means far from the minima



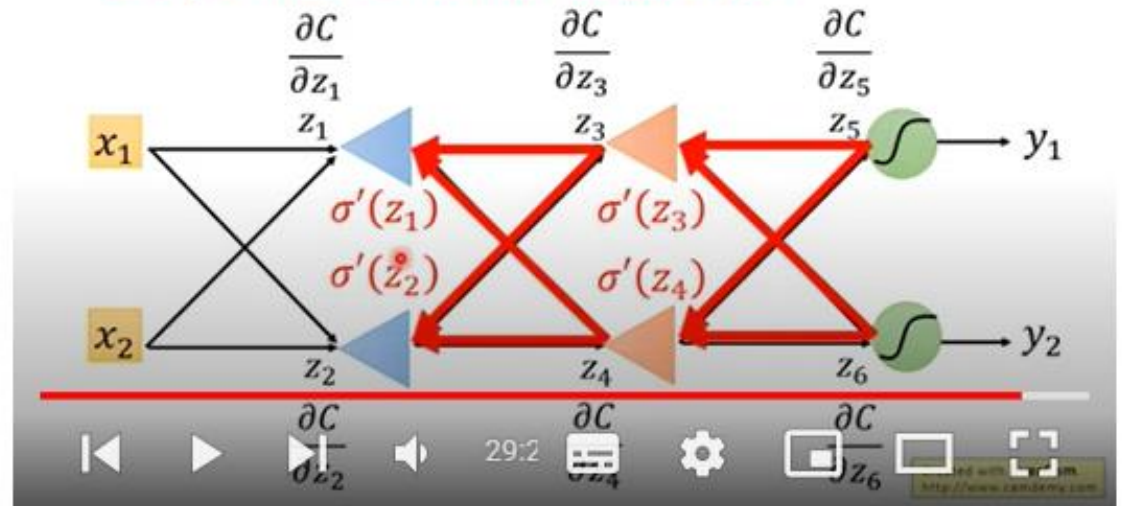
ML Lecture 3-1: Gradient Descent

<https://youtu.be/yKKNr-QKz2Q>

Backpropagation – Backward Pass

Compute $\frac{\partial C}{\partial z}$ for all activation function inputs z

Compute $\frac{\partial C}{\partial z}$ from the output layer



ML Lecture 7: Backpropagation

<https://youtu.be/ibJpTrp5mcE>

找出函式的三步驟

“RL 要取代 Deep Learning 了”

設定範圍

訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.

設定標準

訂出「評量函式好壞」的標準

Supervised Learning, Semi-supervised Learning, RL, etc.

達成目標

找出最好的函式

Gradient Descent (Adam, AdamW ...), Genetic Algorithm, etc.

達成目標

找出最好的函式

Gradient Descent (Adam, AdamW ...), Genetic Algorithm, etc.

\mathcal{H}

L

最佳化演算法
(Optimization Algorithm)

更快

f^*

$L(f^*)$ 越低越好

需要先設定

Learning
Rate

Batch
Size

How to
Init

超參數 (Hyperparameter)

對超參數不敏感

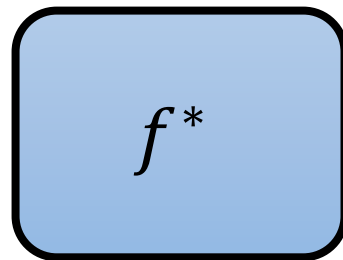
設定標準

訂出「評量函式好壞」的標準

Supervised Learning, Semi-supervised Learning, RL, etc.



訓練資料



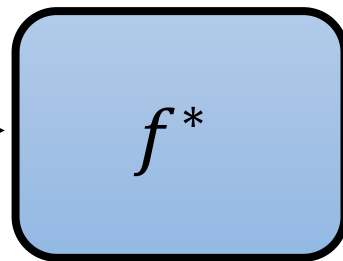
$$f^* = \arg \min_{f \in \mathcal{H}} L(f)$$

$L(f)$ 小

Training



測試資料



測試不一定好

Testing

???

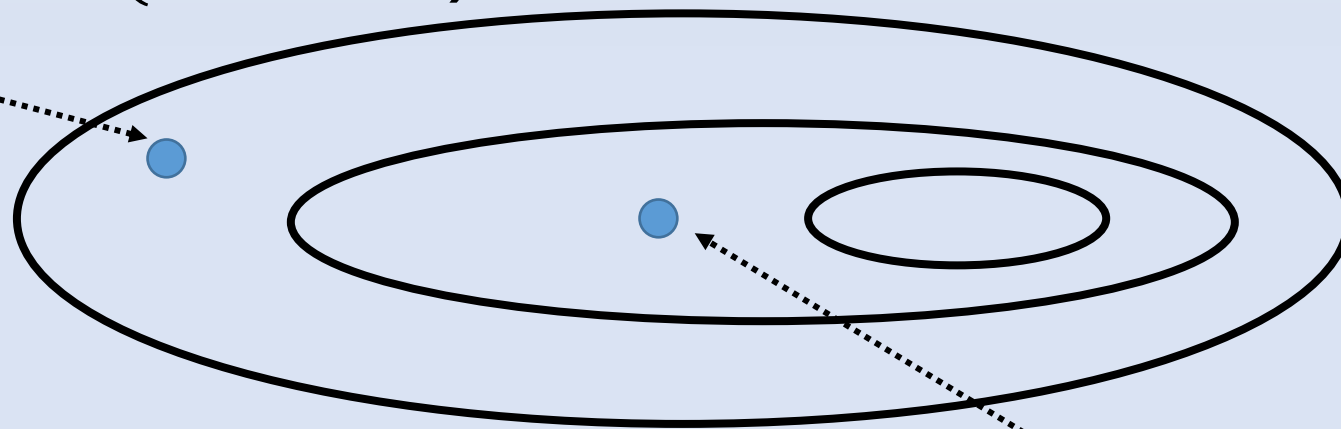
在 Loss 上做額外考量 (如 : Regularization)

設定範圍

考量任務特性，訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.

$L(f)$ 小，測試不好 (硬記答案)



$L(f)$ 小，測試好

設定範圍

考量任務特性，訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.

寶可夢、數碼寶貝分類器

淺談機器學習原理



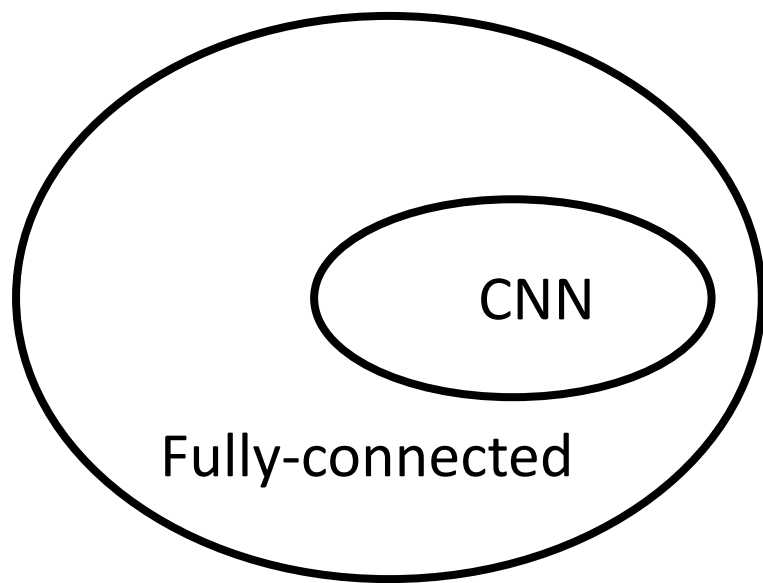
https://youtu.be/_j9MVVcvyZI

【機器學習 2022】再探寶可夢、數碼寶貝分類器 – 淺談機器學習原理

設定範圍

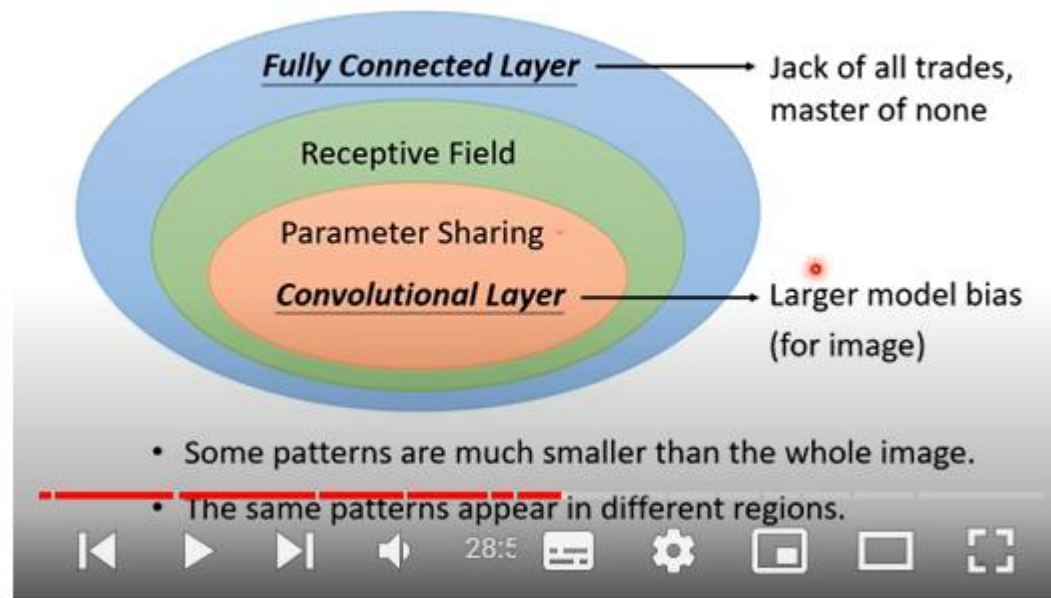
考量任務特性，訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.



<https://youtu.be/OP5HcXJg2Aw>

Benefit of Convolutional Layer

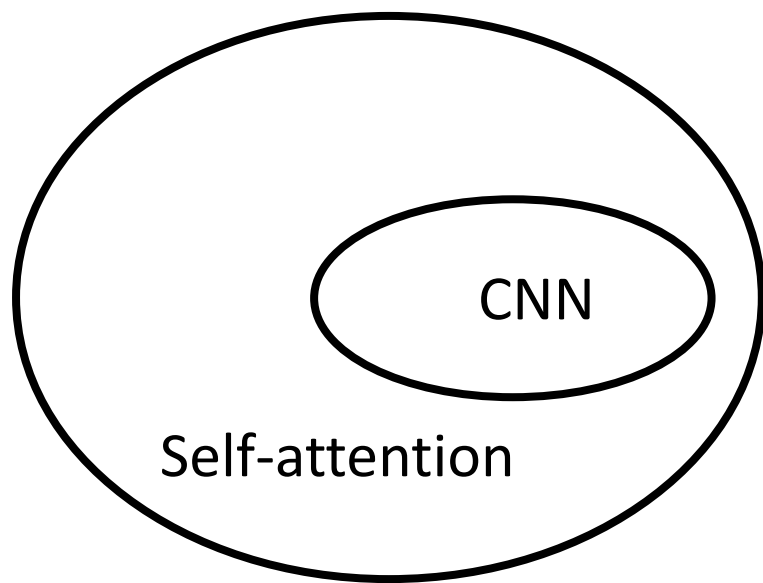


【機器學習2021】卷積神經網路 (Convolutional Neural Networks, CNN)

設定範圍

考量任務特性，訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.



<https://youtu.be/gmsMY5kc-zw>

Self-attention v.s. CNN

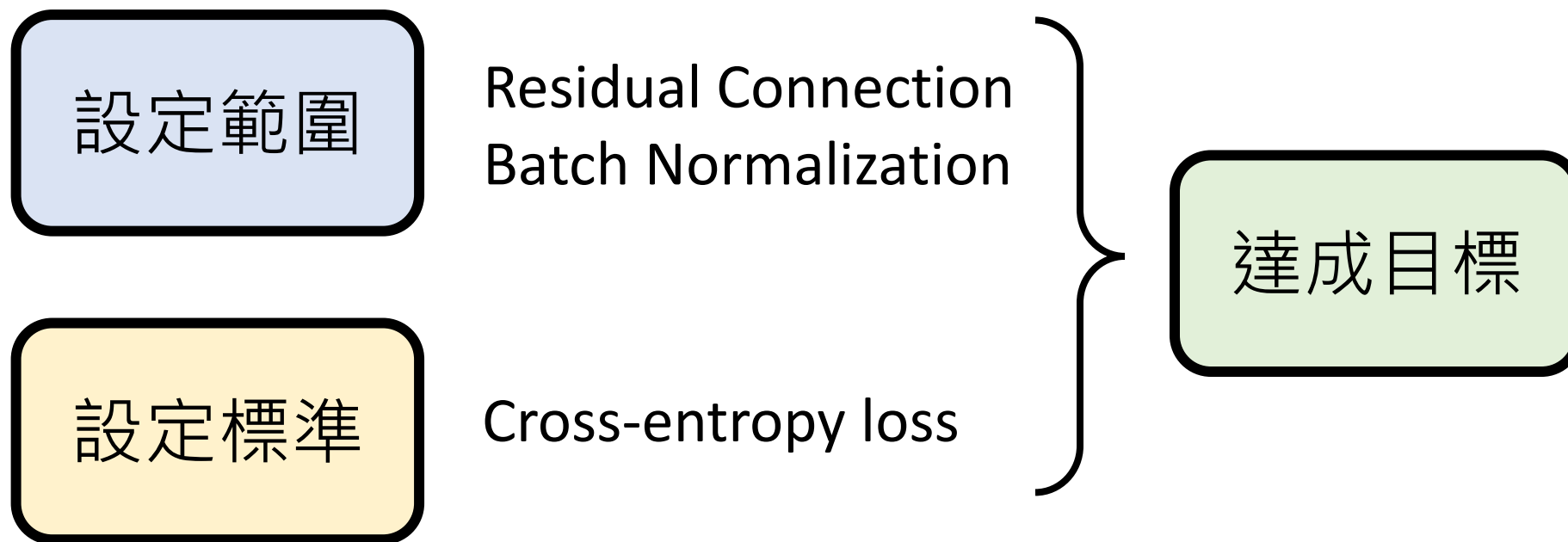
Multi-Head Self-Attention Layer

On the Relationship between Self-Attention and Convolutional Layers

32:1 <http://arxiv.org/abs/1803.02629>

【機器學習2021】自注意力機制 (Self-attention) (下)

有一些方法的好是可以支援其他步驟



總結

設定範圍

考量任務特性，訂出候選函式的集合

Deep Learning (CNN, Transformer ...), Decision Tree, etc.

設定標準

訂出「評量函式好壞」的標準

Supervised Learning, Semi-supervised Learning, RL, etc.

達成目標

找出最好的函式

Gradient Descent (Adam, AdamW ...), Genetic Algorithm, etc.