

深度學習 Pytorch手把手實作 AutoEncoder

黃志勝 (Tommy Huang)

義隆電子 人工智慧研發部

國立陽明交通大學 AI學院 合聘助理教授

國立台北科技大學 電資學院合聘助理教授



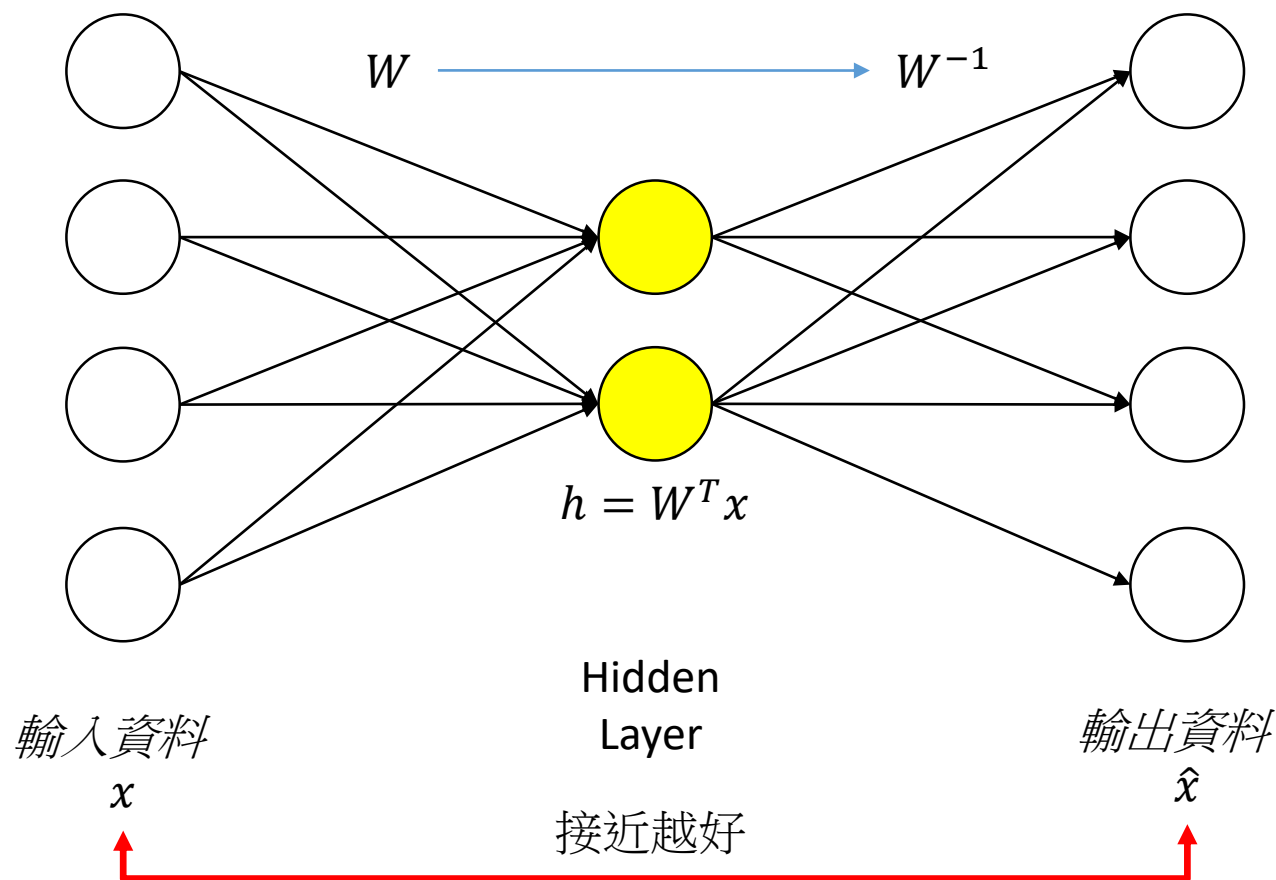
Introduction

- AutoEncoder (AE) 和 Generative Adversarial Network(GAN)都屬於unsupervised learning的領域。
- 兩種演算法看似很像，很多人會拿這兩種方法比較資料生成的效能。

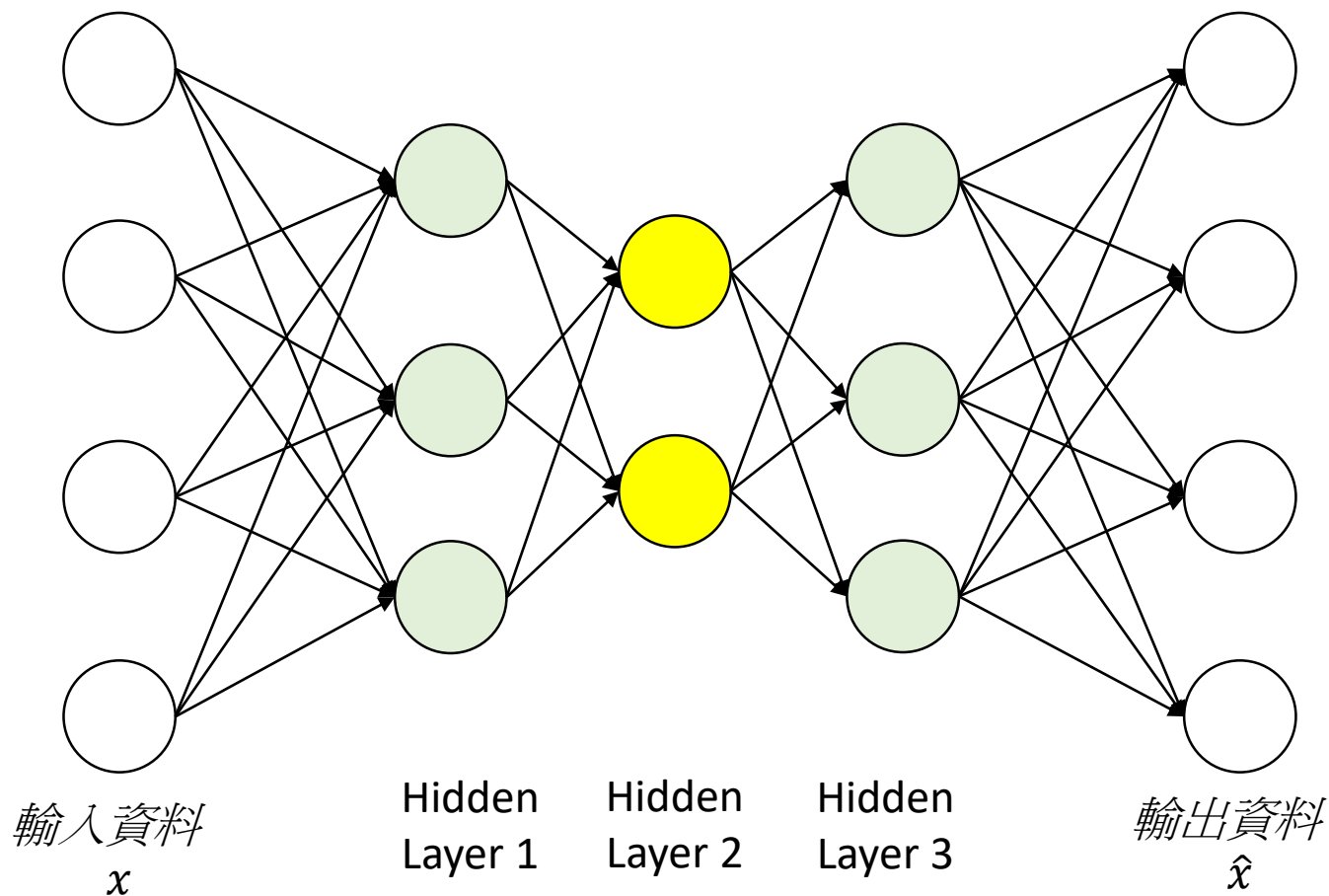


AutoEncoder (AE)

訓練AE基本上資料是不需要「標註(Labeling)」。
透過輸入盡量和輸出逼近的方式，讓神經網路架構自行學習中間的權重。



Stacked AutoEncoder (AE)



***輸入和輸出大小是一致即可。**

Stacked AE/AE 範例圖片的 Hidden 數量可為任意，數量不一定要比輸入或是輸出少，範例圖純粹是我不想畫太多。

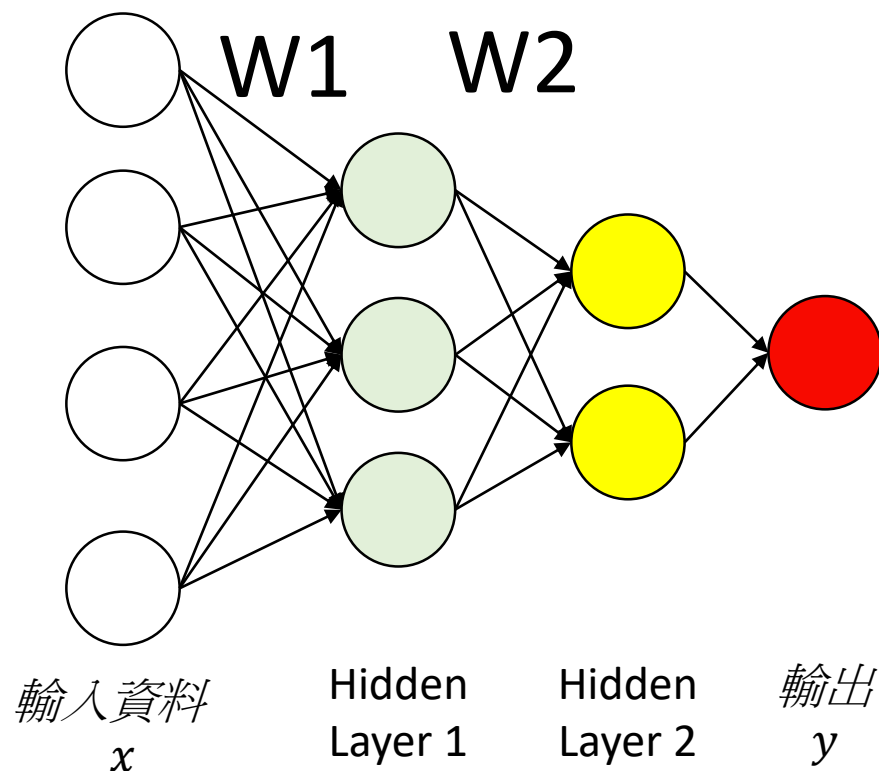


AutoEncoder (AE)

- 那AE有什麼用？ ANS: 有三個重要功能
- AE用處就是學1. pretrained weighted。

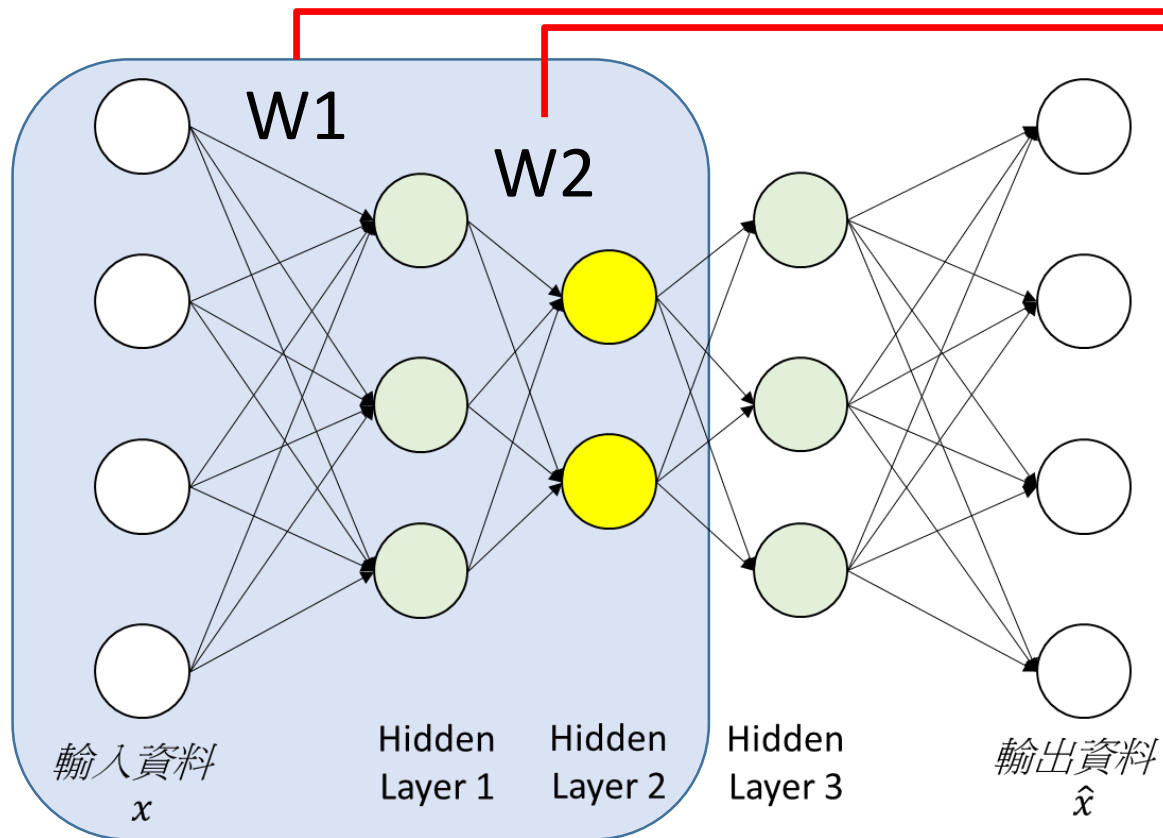
假設我們設計一個神經網路，如右圖，
然後假設 $W1$ 和 $W2$ 參數有 $1W$ 個。
我們收集的數據有 $1W$ 筆。
但標註資料只有100筆。

我們用100筆資料去訓練一個 $1W$ 筆參數
的神經網路，訓練模型前的權重隨機生
成，結果一定有問題。
所以我們可以善用那 $1W$ 筆數據。

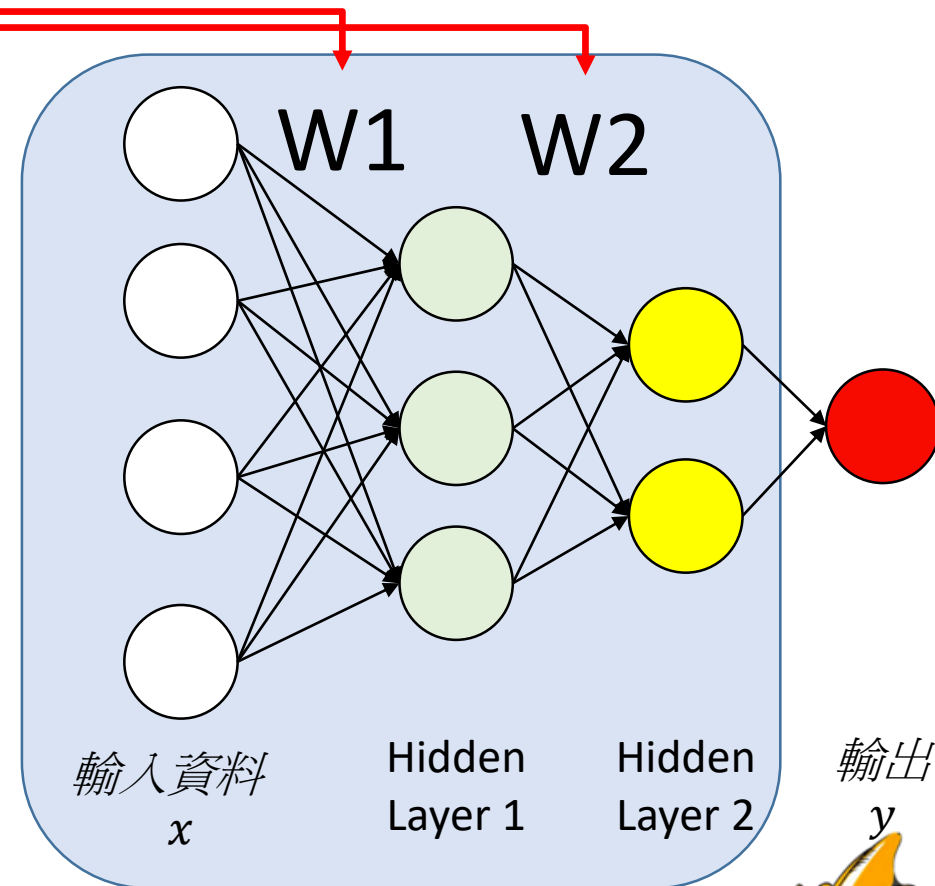


AutoEncoder (AE)

1W筆數據訓練AE



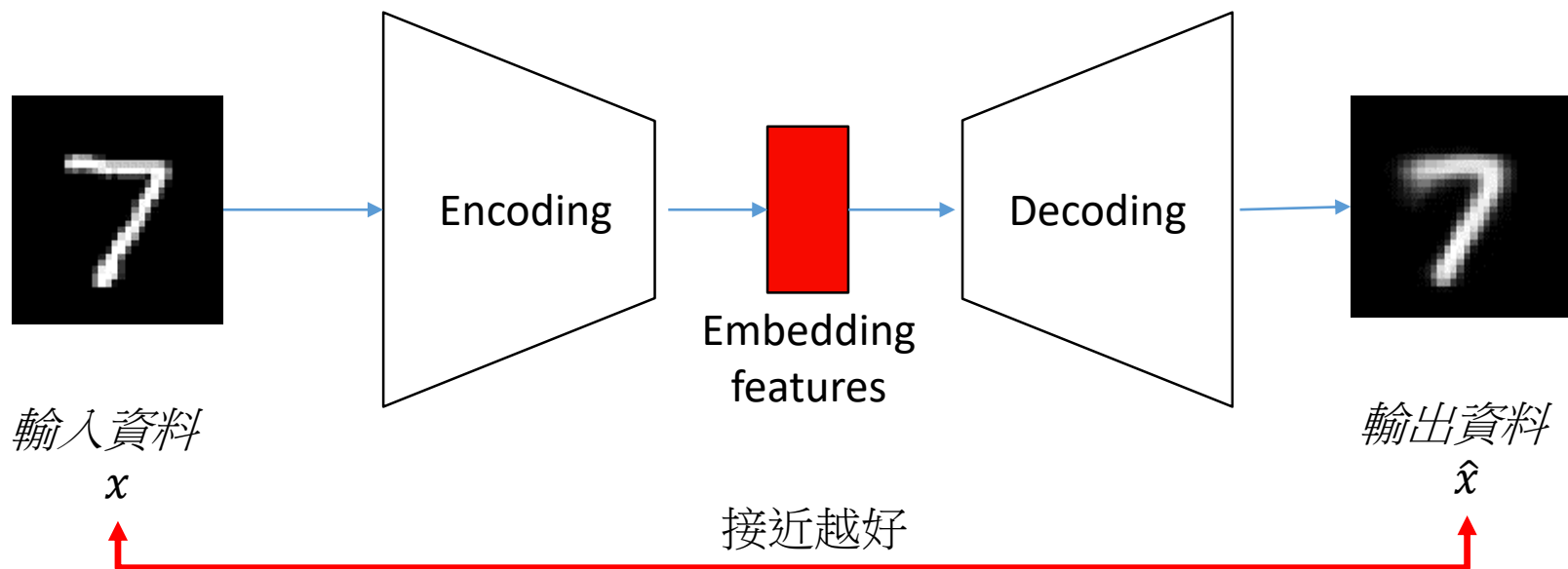
100筆數據來訓練這個網路



AutoEncoder (AE)

- 那AE有什麼用？ ANS: 有三個重要功能
- AE用處就是學2. Feature representation.

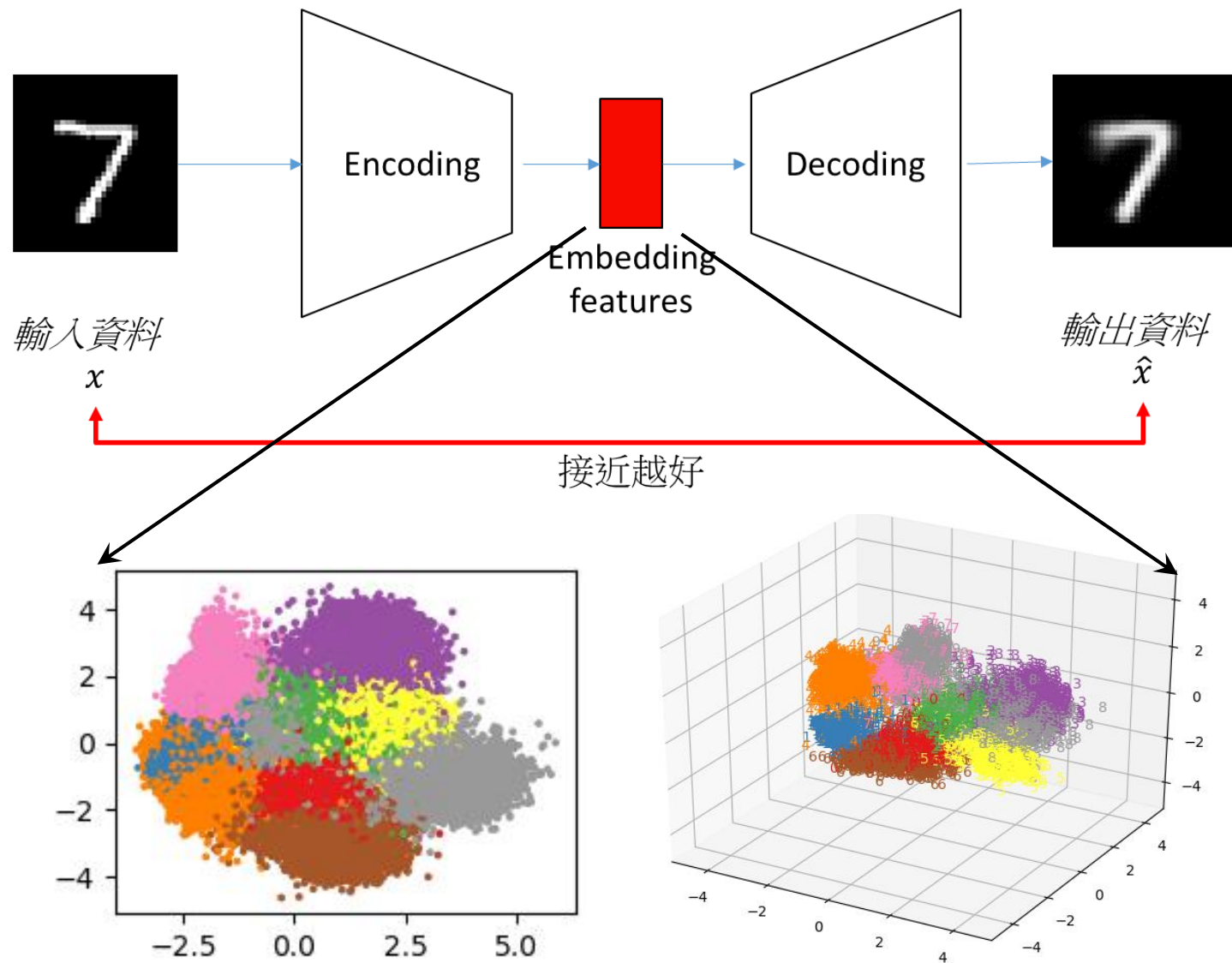
Note: 後續的Stacked AE, VAE(Variational AE), DAE(Denoise AE)都屬於此架構的變形:



紅色那塊就是「編碼後特徵」，所以資料特性(Embedding features)越像的則在此Embedding space會越接近。讓資料自行進行物以類聚的學習。

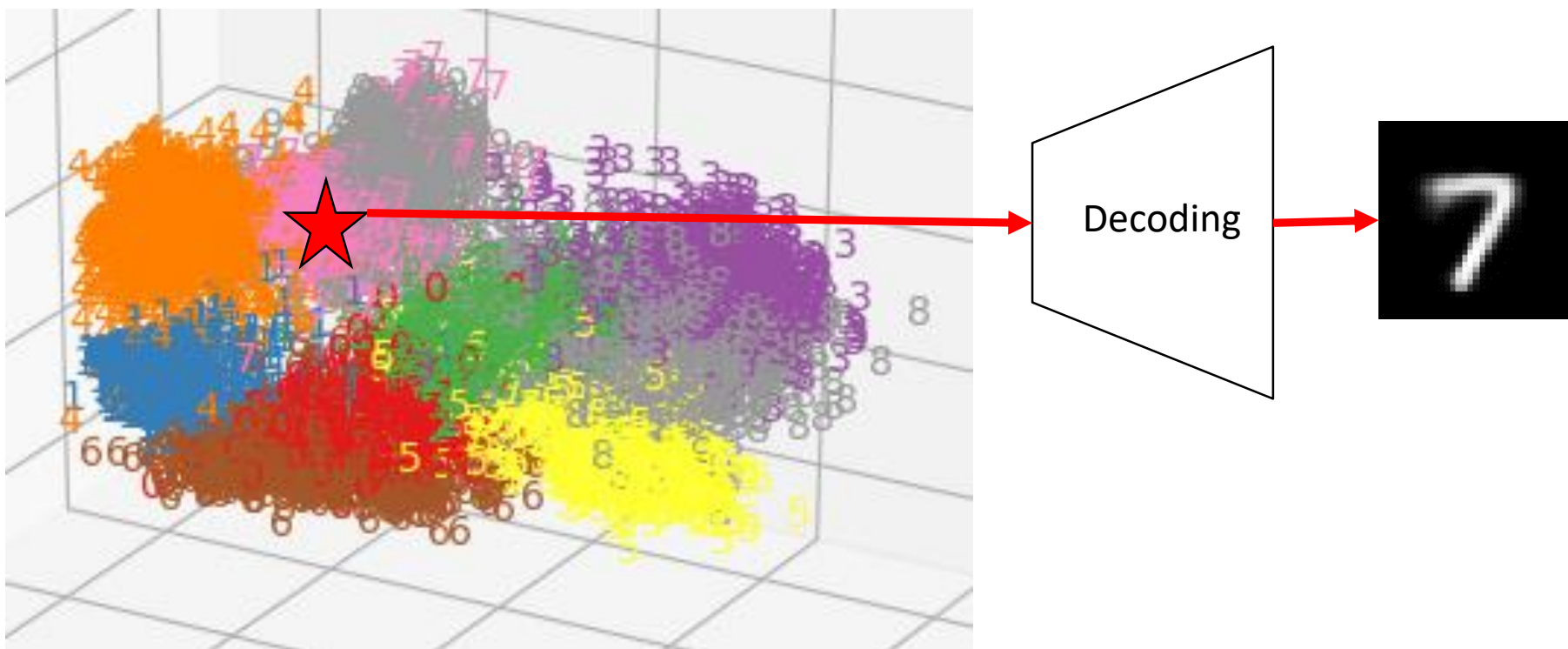


AutoEncoder (AE)



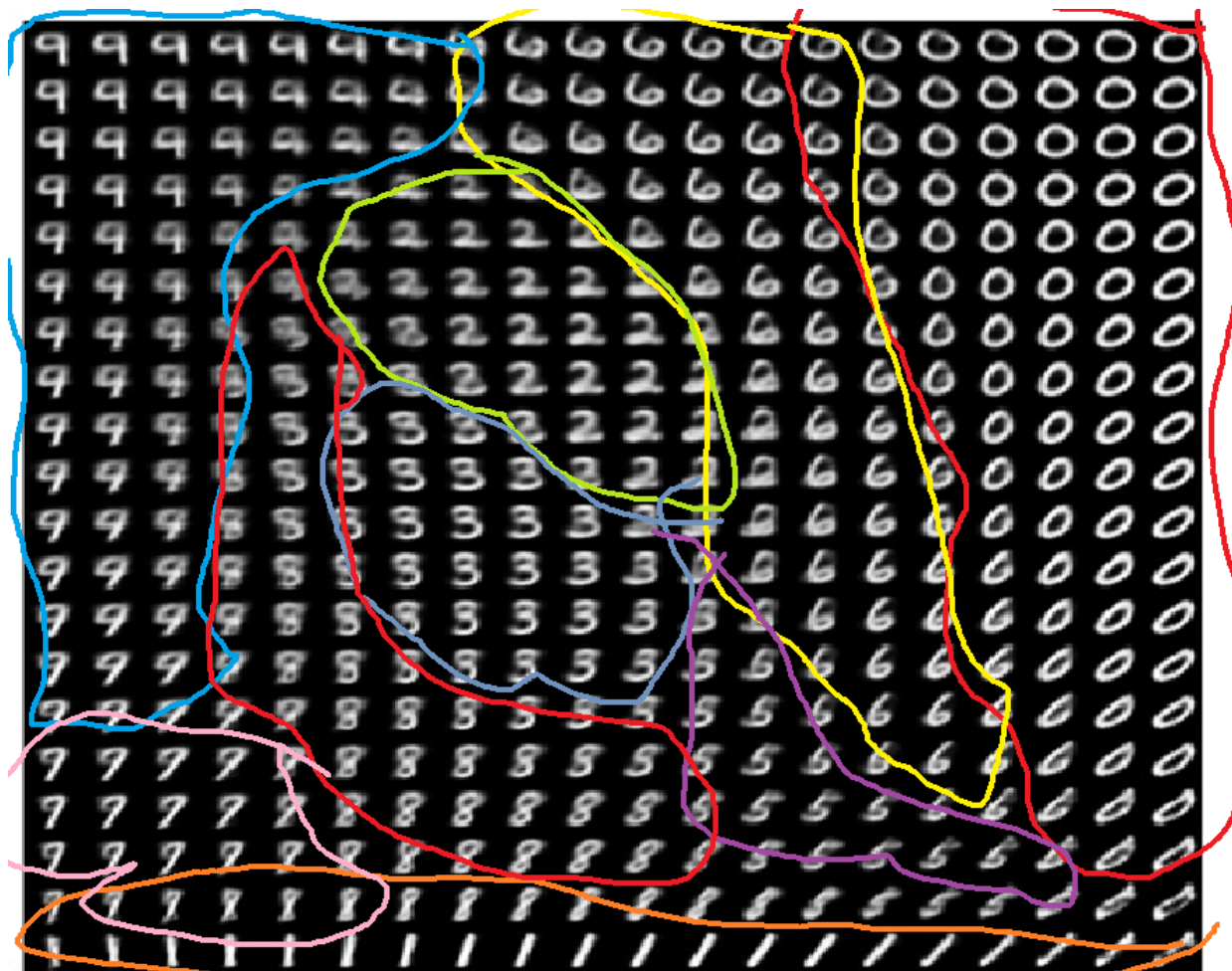
AutoEncoder (AE)

- 那AE有什麼用? ANS: 有三個重要功能
- AE用處就是學 3. Data Generation.



AutoEncoder (AE)

Embedding
Features 2

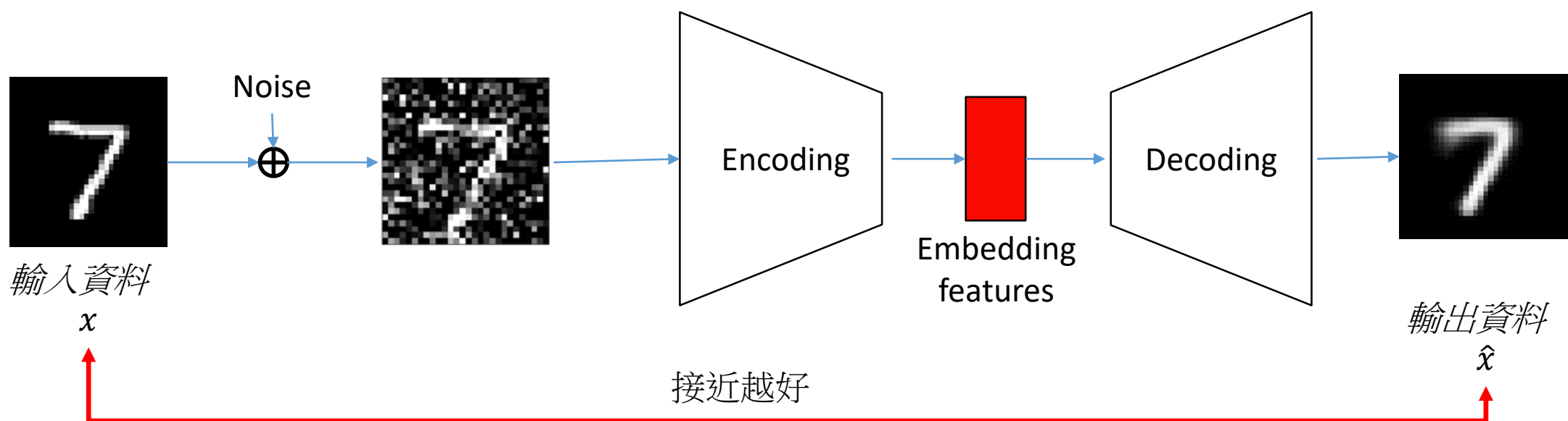


假設Embedding feature只有兩個。右圖為在Embedding Space的數字Decoder回圖片的結果。

Embedding
Features 1



AutoEncoder (AE)衍生

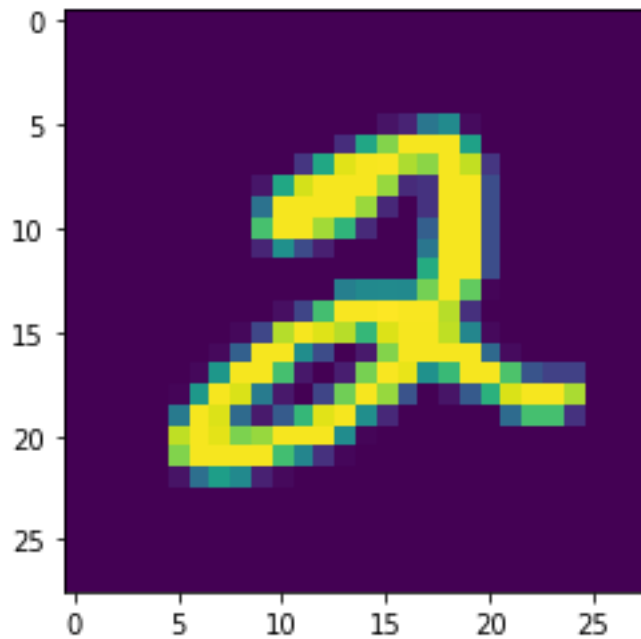


只要在輸入圖片的部分加上雜訊，經由AE的處理後，輸出的 \hat{x} 跟原本的 x 越接近，代表這個AE可以做到去雜訊的功能，這也就是Denoise Auto-encoder

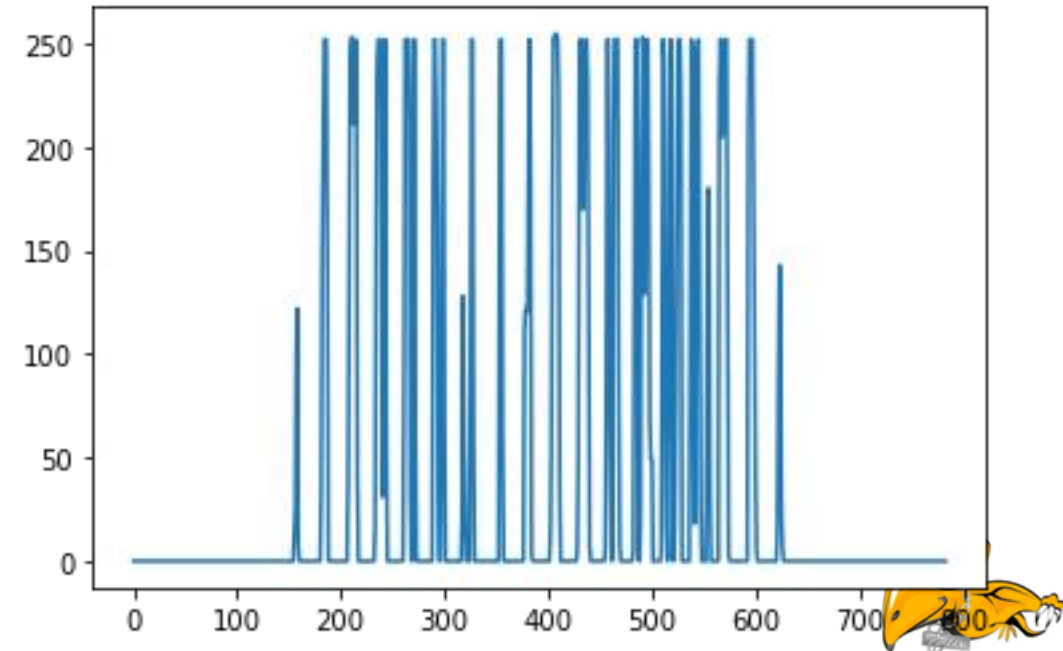


Hand-by-hand pytorch implementation.

- Here I will build a MLP-based AutoEncoder for MNIST.
- A single digit in MNIST is 28×28 8bits Gray image.
- If we flatten 2D image with 1D array, which contains with 784 elements.



flatten

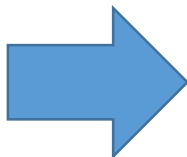


2D image flatten to 1D array

1ch 2D image

1	2	3
4	5	6
7	8	9

flatten



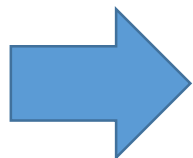
1D array

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

2ch 2D image

1	2	3	10	11	12
4	5	6	13	14	15
7	8	9	16	17	18

flatten



1D array

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----



AE model

