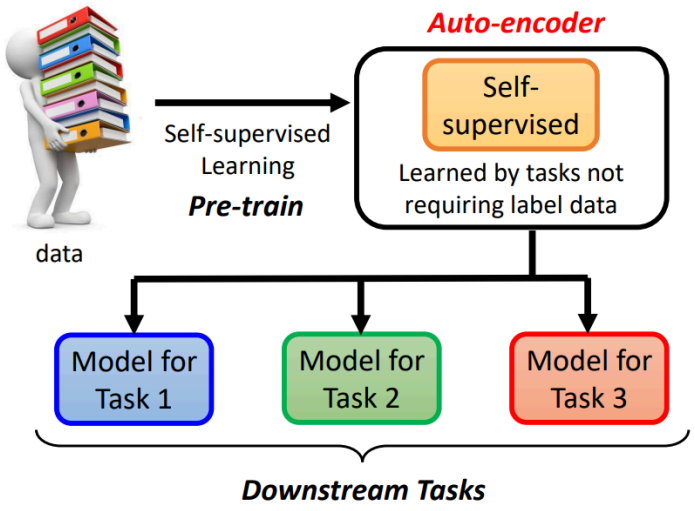
08-Auto-encoder

1. Basic Idea

1.1 基本認識

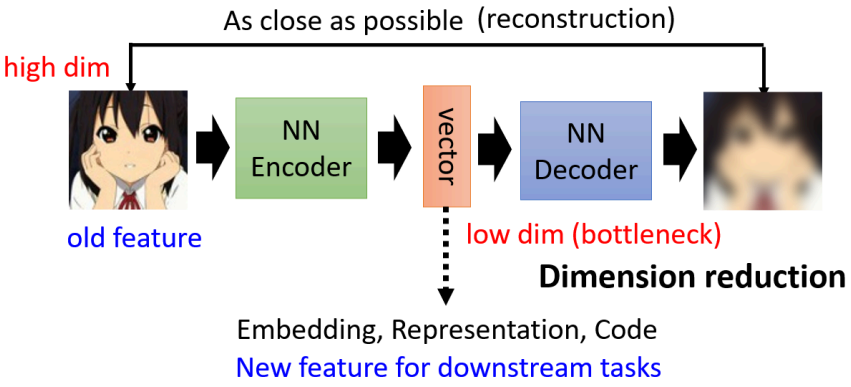
self-supervised learning 是利用不需要標註資料的任務來訓練模型，如填空題、預測下一個 token，又稱為 pre-train



auto-encoder 可以看作是 **self-supervised learning 的一種的方法**

1.2 主要架構

- encoder 讀進一張高維圖片，把這張圖片變成一個低維（bottleneck）向量（稱 **embedding**、**representation** 或 **code**）作為 decoder 的輸入。架構類似 CNN
- decoder 輸入**向量**，產生一張圖片。架構類似 GAN 的 generator



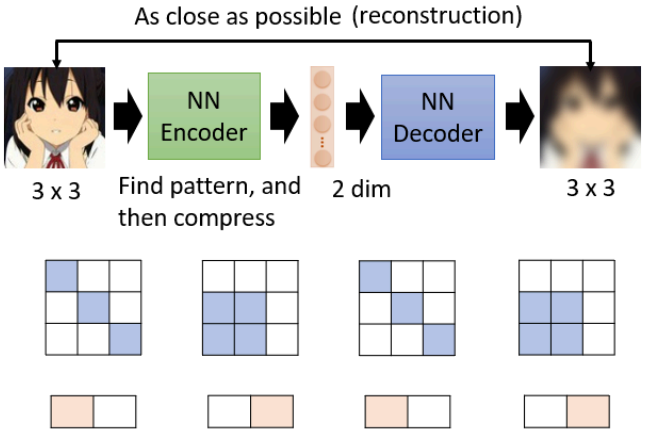
訓練的目標希望 **encoder 的輸入跟 decoder 的輸出越接近越好**（reconstruction）
與 **Cycle GAN** 做的事情其實一模一樣

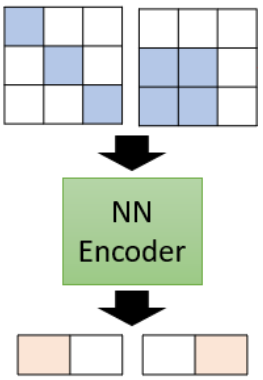
動機：

降維（dimension reduction），圖片可以看作是一個很長的向量，但這個向量太長不好處理，所以丟給 **encoder** 來壓縮輸出一個較短的向量。學習更多：PCA、t-SNE

1.3 還原為何能成功？

就算有一個高維度的向量圖片，但可能他的**變化有限**，所以只需很少的維度就能夠表示高維圖片的各種變化情況



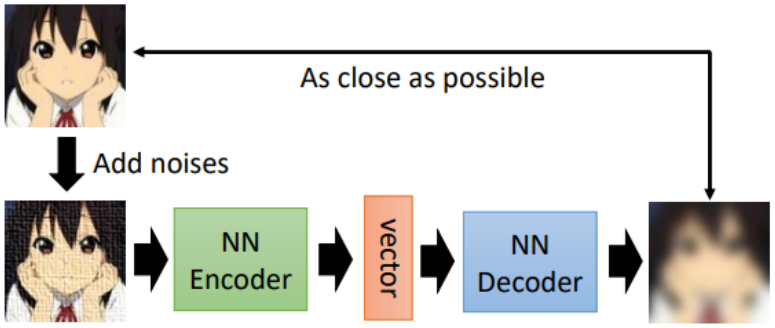


如上圖，3×3 的矩陣應當有 2^9 種變化情況，但可能只有 2 種情況會出現，因此可以只用 2 維的向量進行表示。encoder 就能夠實現這種轉換，把複雜的訊息用簡單的方法表示，實現

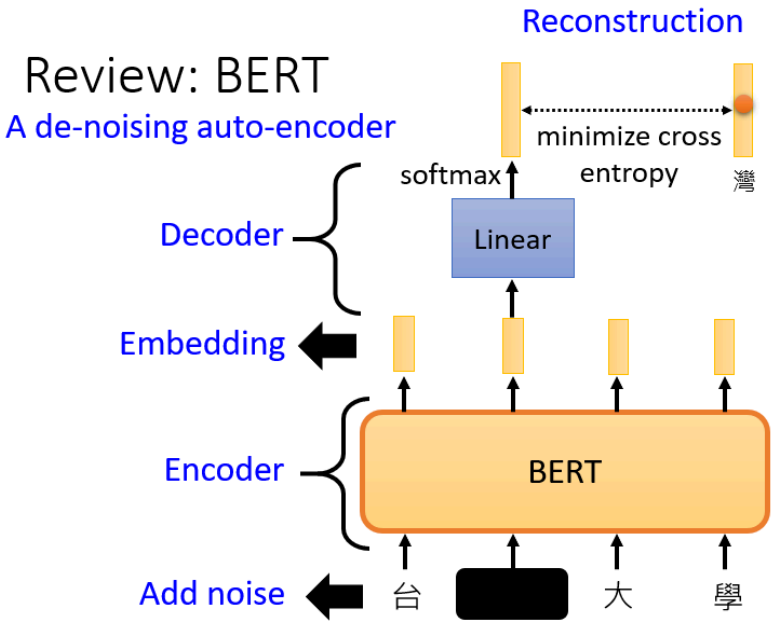
dimension reduction

1.4 De-noising Auto-encoder

De-noising auto-encoder 是將圖片送入 encoder 之前加一些雜訊，要 decoder 把向量**還原成加入雜訊前的結果**

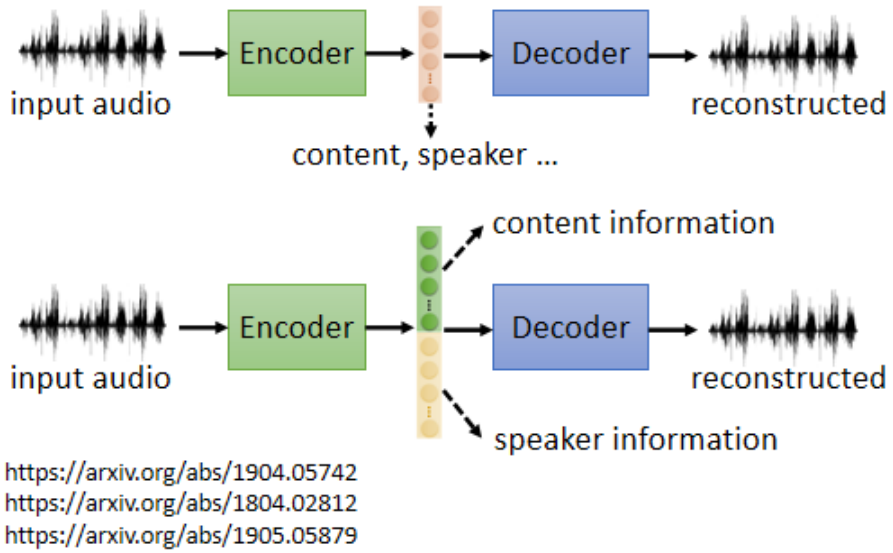


與 BERT 做的事很像，可以說 **BERT 就是一個 De-noising Auto-encoder**



2. Feature Disentanglement

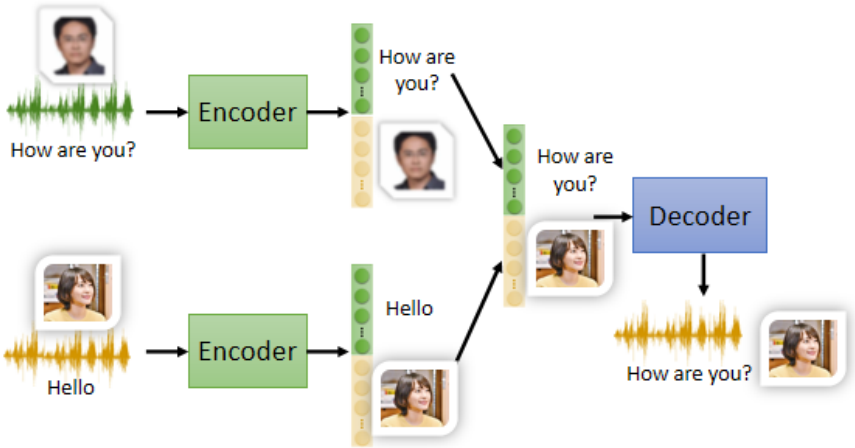
由於 embedding 向量能夠還原回原來的數據，這說明 **auto-encoder 能夠讓 embedding 中包含原數據中的所有訊息**。例如把一段聲音丟到 encoder 變成向量，這個向量包含了語音裡所有重要的資訊，包括這句話的內容是什麼、這句話是誰說的等等



Feature Disentangle 就是希望在訓練一個 auto-encoder 時，同時有辦法知道這個 embedding 的**哪些維度代表了哪些資訊**，詳細可參考：

1. *One-shot Voice Conversion by Separating Speaker and Content Representations with Instance Normalization*
2. *Multi-target Voice Conversion without Parallel Data by Adversarially Learning Disentangled Audio Representations*
3. *AUTOVC: Zero-Shot Voice Style Transfer with Only Autoencoder Loss*

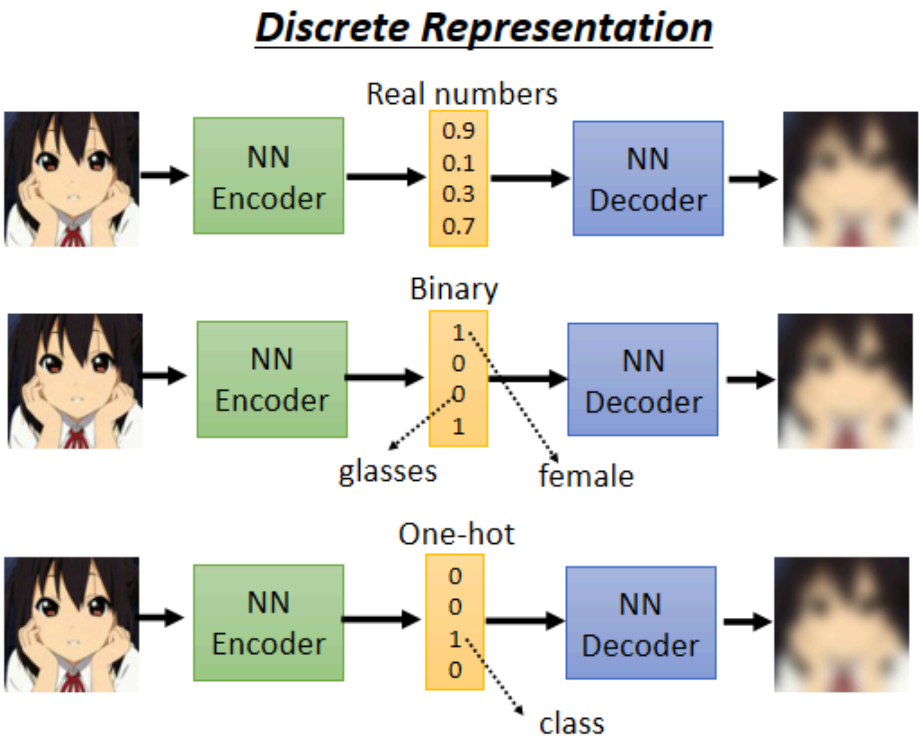
2.1 應用：Voice Conversion



利用 **Feature Disentangle** 可以知道向量中哪些維度代表語音的內容、哪些維度代表語音的聲音，只要把其中一人說話的內容的部分取出來，把另一人說話的聲音特徵的部分取出來，將二者併起來丟到 decoder 裡面就可以實現變聲

3. Discrete Latent Representation

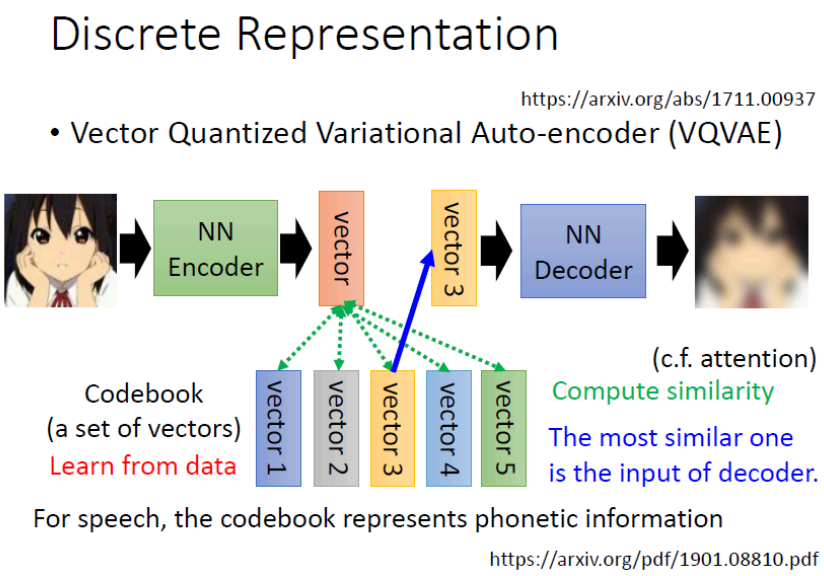
embedding 的一些不同形式



- embedding 是一連串的**實數數字**
- embedding 只有 **0 跟 1**，每一個維度它就代表了**某種特徵的有或者是沒有**
  如第一維 0 代表男生、1 代表女生；第三維 0 代表有戴眼鏡、1 代表沒戴眼鏡
- embedding 是 **one-hot vector**，可以在完全沒有 label data 的情況下讓**機器自動學會分類**
  如手寫數字辨識，embedding 就設十維

3.1 Vector Quantized Variational Auto-Encoder（VQVAE）

encoder 輸出一個向量，與 codebook 的每個向量計算相似度，挑選相似度最高的向量再輸入進 decoder

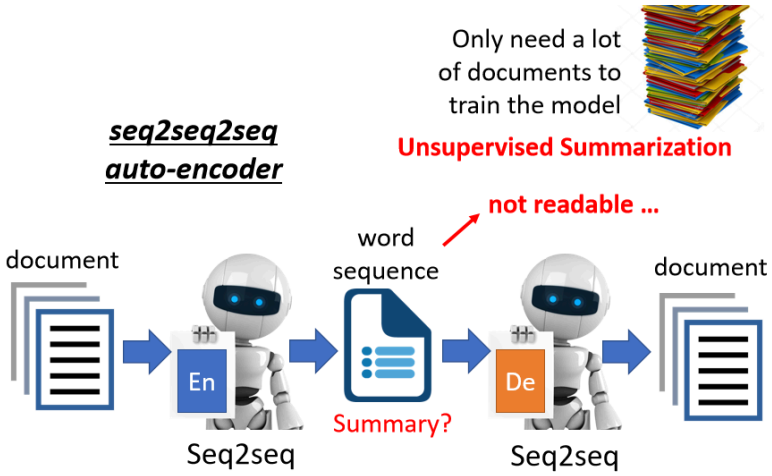


好處：

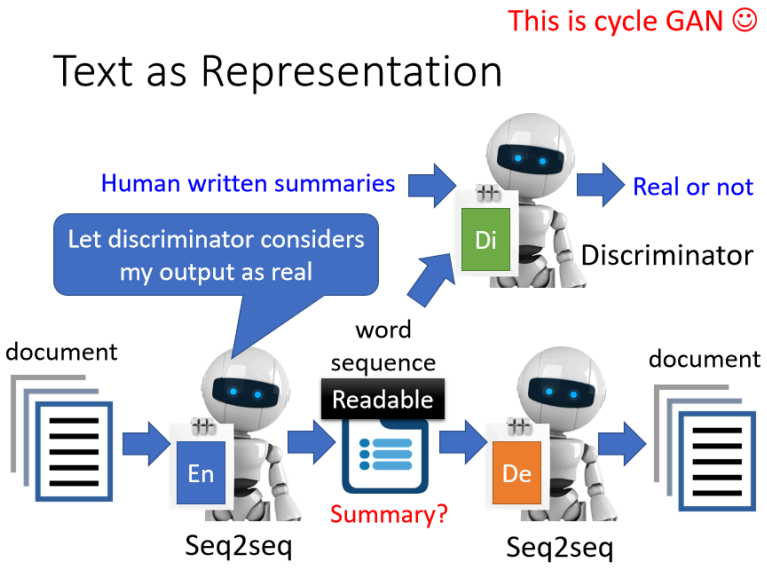
Discrete Latent Representation，假設 codebook 裡面有 32 個向量，那 decoder 的輸入就只有 32 種可能，等於是讓 **embedding 是離散的，沒有無窮無盡的可能**

3.2 Text as Representation

讓 **representation（embedding）是文字**，比如做摘要，給一段文章輸入到 encoder 輸出摘要，再輸入到 decoder 還原，但會發現單純這樣**訓練不起來**



再用上 **GAN 的 discriminator**，discriminator 看過人寫的句子，所以知道人寫的句子長什麼樣子

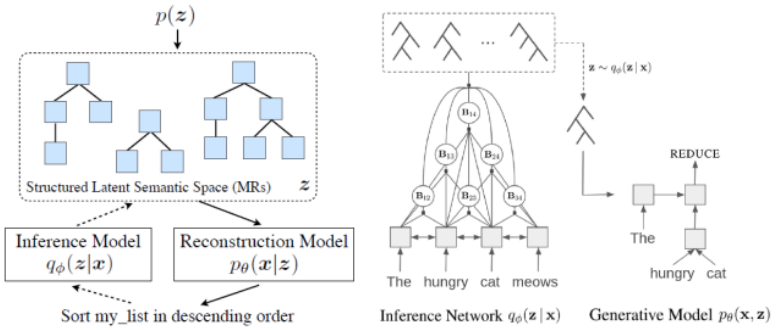


另一角度看 CycleGAN

encoder 要想辦法產生一段句子，這段句子不只可以**透過 decoder 還原回原來的文章**，還要是**discriminator 覺得像是人寫的句子**

3.3 Tree as Embedding

給一段文字轉為 tree structure，再把 tree structure 轉回為原文字

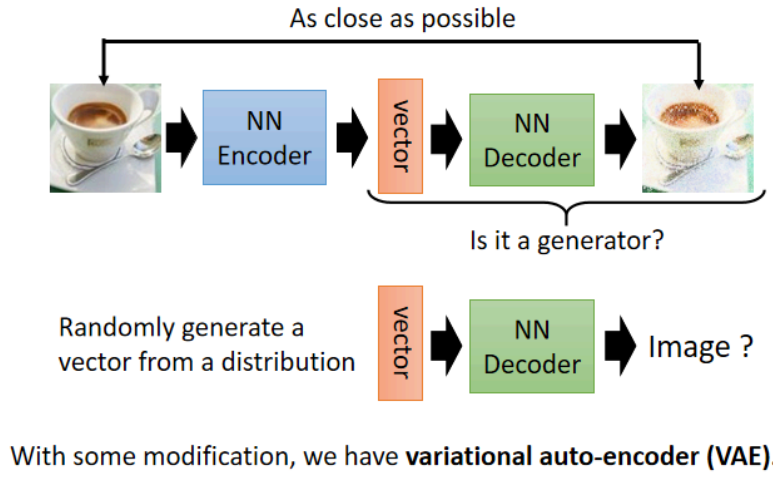


https://arxiv.org/abs/1806.07832 https://arxiv.org/abs/1904.03746

https://arxiv.org/abs/1806.07832、https://arxiv.org/abs/1904.03746

4. More Applications

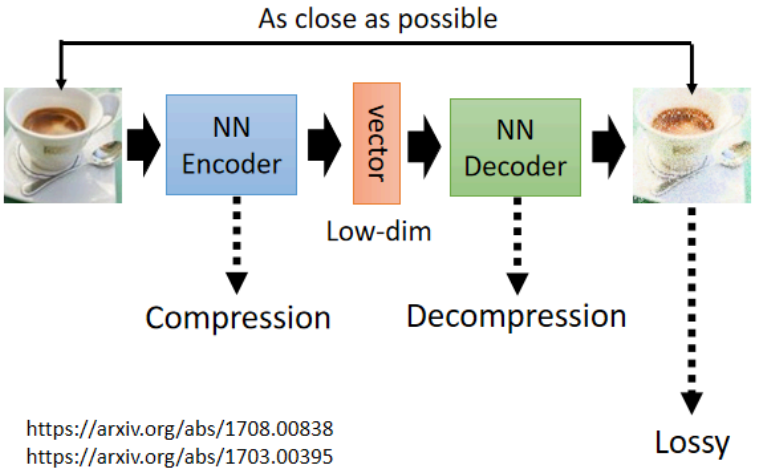
4.1 Generator



With some modification, we have **variational auto-encoder (VAE).**

decoder 正好是輸入一個向量，產生一張圖片，所以可以把它當做一個 generator 來使用

4.2 Compression

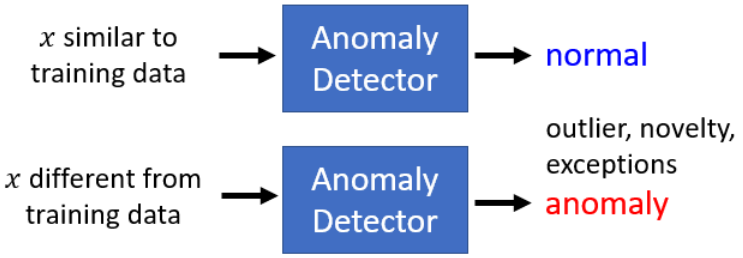


encoder 的輸出會把高維向量變為低維向量，encoder 做壓縮，而 decoder 做解壓縮

4.3 Anomaly Detection（異常檢測）

判斷一筆新的資料跟之前在訓練資料裡面**看過的資料相不相似**

- Given a set of training data $\{x^1, x^2, \cdots, x^N\}$
- Detecting input x is *similar* to training data or not.



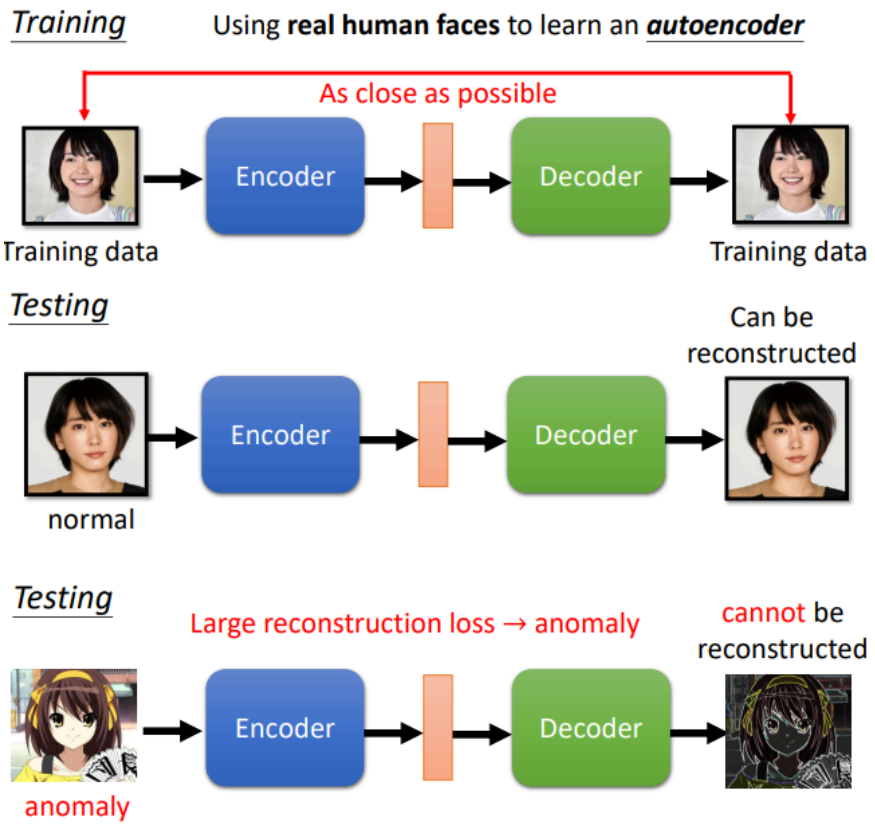
4.3.1 應用

- Fraud Detection
  - Training data: credit card transactions, x: fraud or not
  - Ref: https://www.kaggle.com/ntnu-testimon/paysim1/home
  - Ref: https://www.kaggle.com/mlg-ulb/creditcardfraud/home
- Network Intrusion Detection
  - Training data: connection, x: attack or not
  - Ref: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
- Cancer Detection
  - Training data: normal cells, x: cancer or not?
  - Ref: https://www.kaggle.com/uciml/breast-cancer-wisconsin-data/home

- **詐欺偵測**
  假設訓練資料有許多信用卡的交易紀錄，訓練一個異常檢測的模型，有一筆新的交易紀錄進來，可以讓機器判斷這筆紀錄算是正常的還是異常的
- **網路侵入偵測**
  收集許多正常的連線的紀錄，訓練出一個異常檢測的模型，看看新的連線是正常的連線還是異常的連線
- **癌細胞檢測**
  收集許多正常細胞的資料，訓練一個異常檢測的模型，看到一個新的細胞可以知道這個細胞有沒有突變，是不是一個癌細胞

難點：

與分類問題很相像，但**並沒有分類問題簡單**，通常收集到的**只有某個類別的資料**，另外一類別的資料極少或根本沒有，這種分類的問題又叫做 **one class** 分類問題



根據 **reconstruction 的好壞來判斷是否異常**

4.3.2 More about Anomaly Detection

- Part 1: https://youtu.be/gDp2LXGnVLQ
- Part 2: https://youtu.be/cYrNjLxkoXs
- Part 3: https://youtu.be/ueDlm2FkCnw
- Part 4: https://youtu.be/XwkHOUPbc0Q
- Part 5: https://youtu.be/Fh1xFBktRLQ
- Part 6: https://youtu.be/LmFWzmn2rFY
- Part 7: https://youtu.be/6W8FqUGYyDo