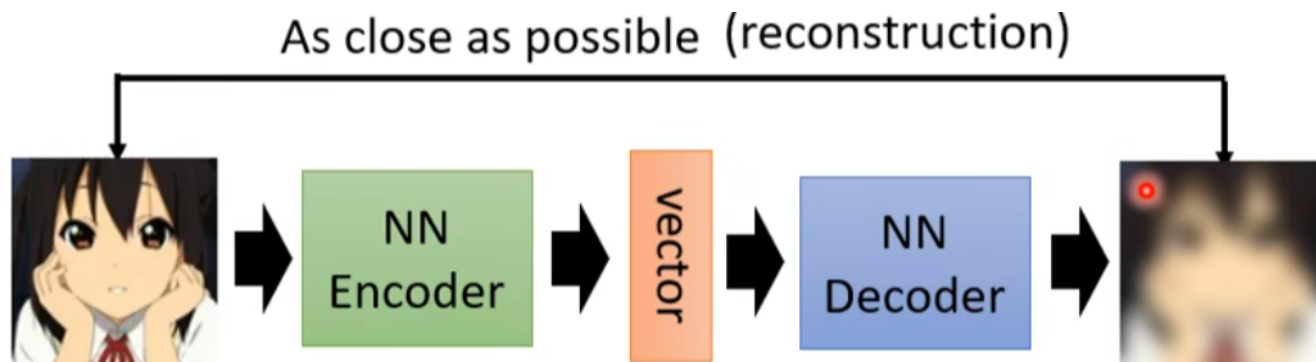


# Auto-encoder

Self-supervised Learning给没有标注的资料设计任务，放进模型学习。而自编码器Auto-encoder可以成为Pre-train的一种方法。

## 介绍

Auto-encoder希望输入的图片经过两次计算之后的输出结果和输入一样。

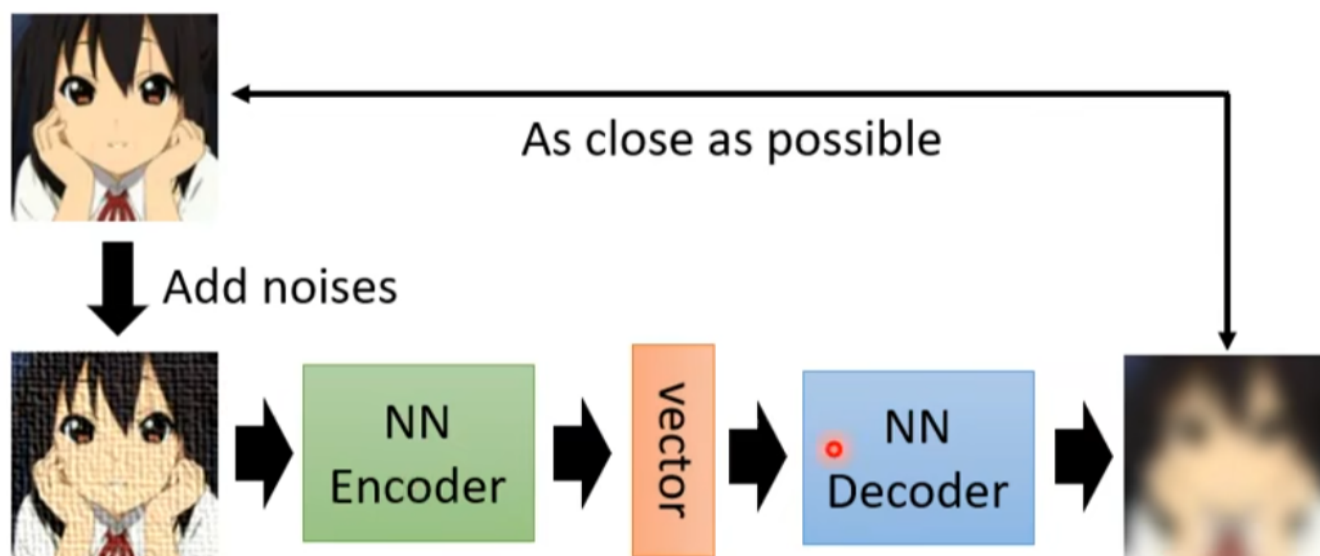


常见用法是：利用Encoder的输出结果（低维向量）运用下游任务。

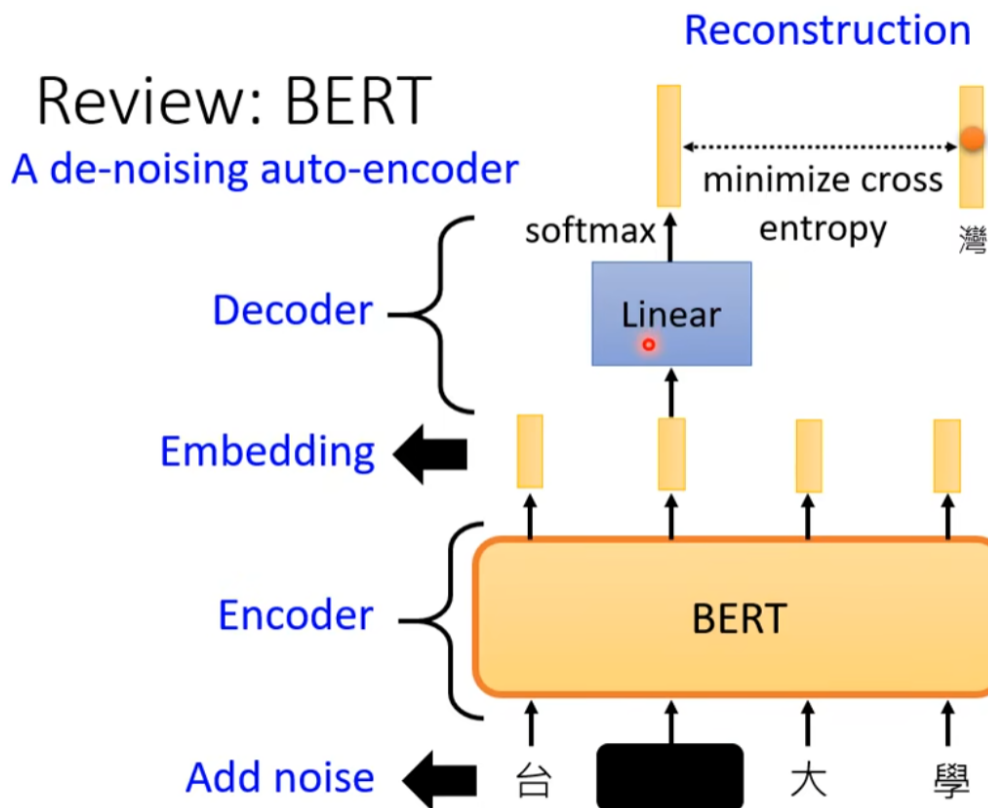
由于输入图片的变化是有限的（随便生成的噪声图像并不是输入的图像），因此对于图像的描述可以降维描述。因此，利用简单的方法描述复杂的事物，这样在下游任务中就可以利用较少的训练材料训练模型。

- 变体De-noising Auto-encoder

原本的输入加入了噪声，训练中加入去噪的功能。



- BERT也是一种特殊的De-noising Auto-encoder

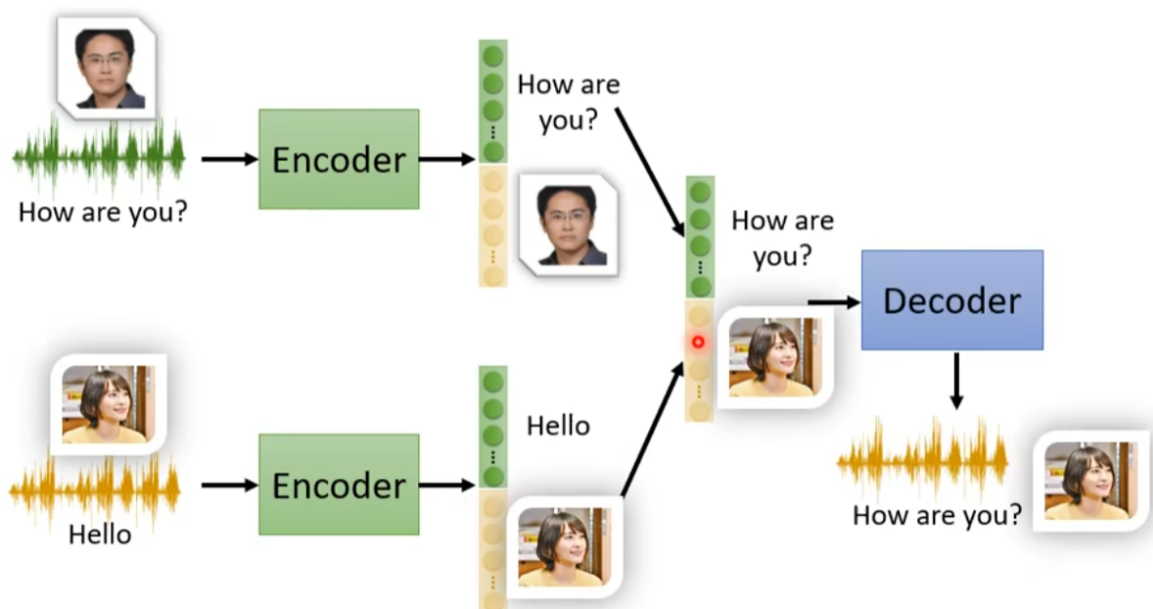


## Auto-encoder的其他运用

- Feature Disentangle: 可以知道每一维代表什么资讯

Disentanglement: 把原本纠缠的东西解开。

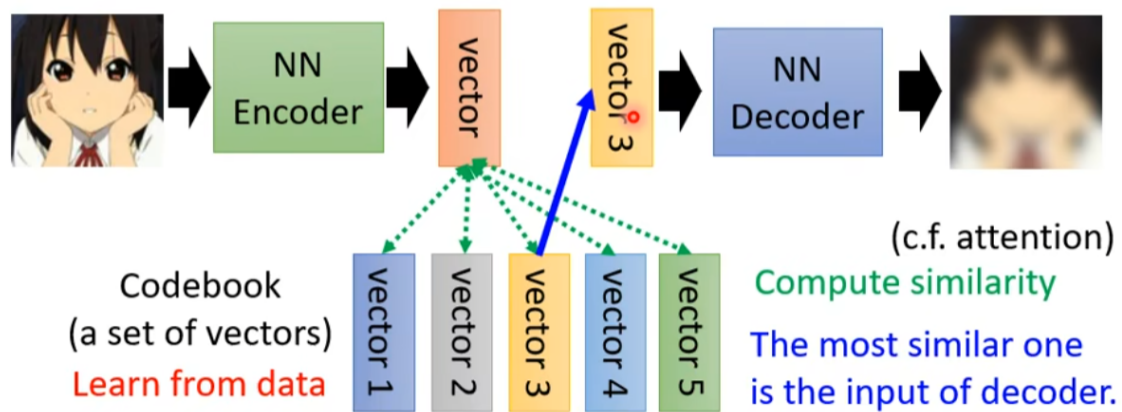
**Apply:** 语音转换 (Voice Conversion) 模仿声音的输入不用和输出配对。通过把文字信息的维度和身份信息的维度进行替换,



- 离散表示 (Discrete Representation)

**Apply:** VQVAE。把Encoder的输出结果和Codebook进行相似度匹配选定输入，这样的输入Decoder的所有可能就都是固定的。这样的话就可以学习到最基础的发音单位。

- Vector Quantized Variational Auto-encoder (VQVAE)



- 文字的Auto-encoder

本质是一个seq2seq2seq模型。但编码器的输出并不是人类可以理解的摘要。

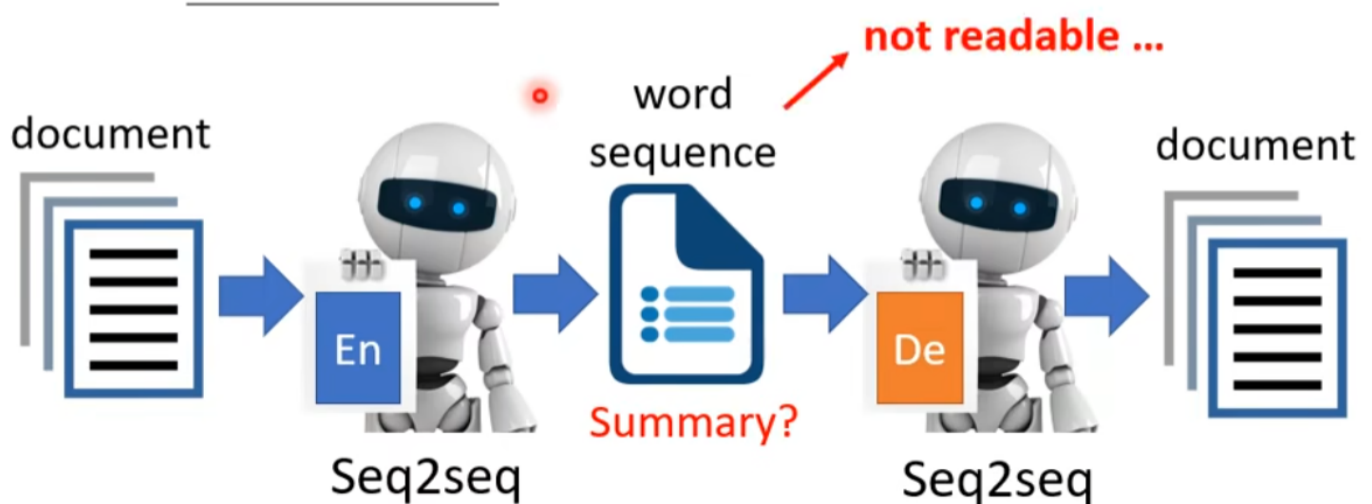
## Text as Representation

Only need a lot of documents to train the model

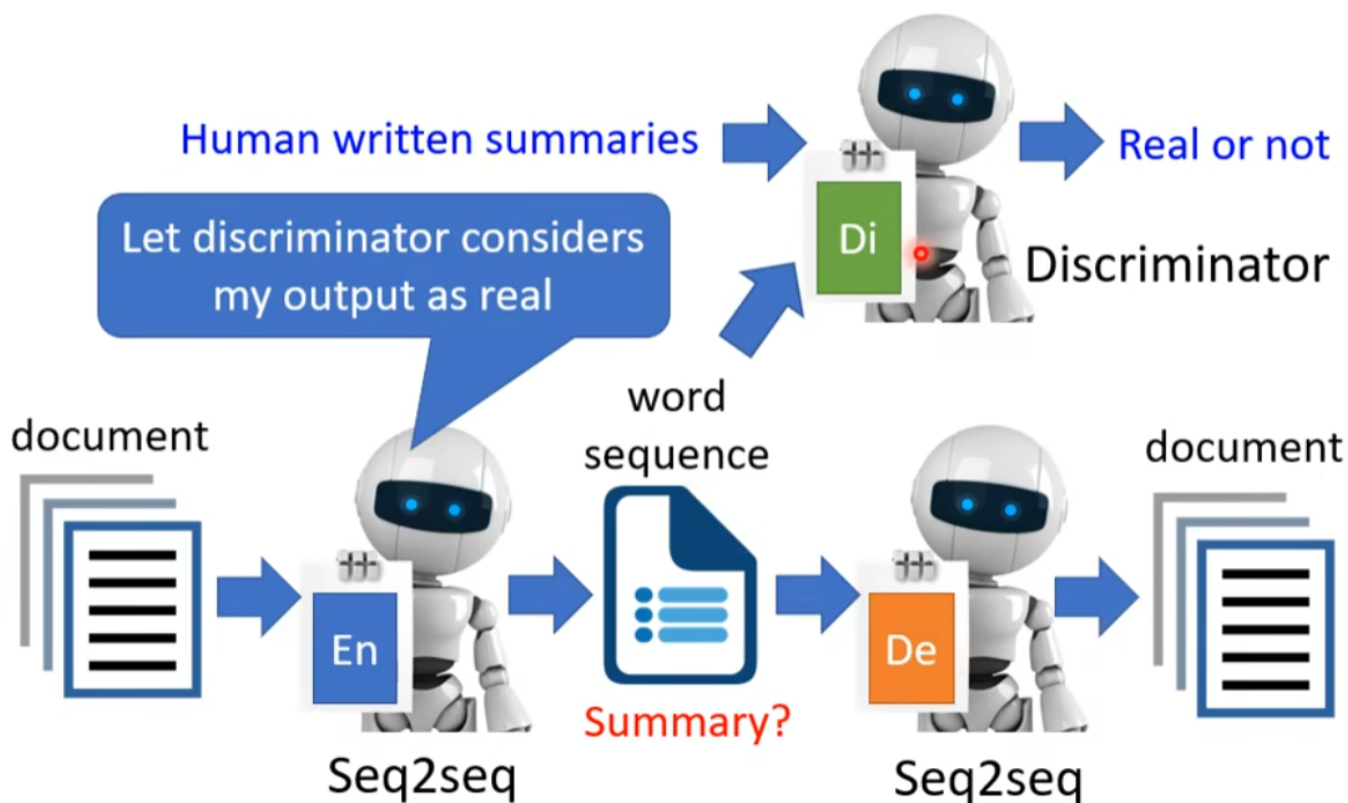


seq2seq2seq  
auto-encoder

**Unsupervised Summarization**



因此，我们可以利用GAN的概念增加一个辨别器。从而强迫产生一段人类可读的密码给解码器破解。



- 解码器也是一个生成器e.g.:variational auto-encoder (VAE)
- Encoder的输出可以成压缩，Decoder输出就是解压缩。
- 异常检测（Anomaly Detection）：难点在于搜集异常数据信息

如果是异常，无法重建（输入和Decoder差距很大），说明检测出异常。