

1. (1%) 請使用不同的 Autoencoder model，以及不同的降維方式(降到不同維度)，討論其 reconstruction loss & public / private accuracy。（因此模型需要兩種，降維方法也需要兩種，但 clustering 不用兩種。）

在同樣都使用 PCA 降到 50 維以及 K-means 分群的前提之下，我分別嘗試了 4 + 4 層的 autoencoder 和 3 + 3 層的 autoencoder，其結果如下：

	Reconstruct loss(MSE)	Private accuracy	Public accuracy
3 + 3 層	0.00806	0.66111	0.66407
4 + 4 層	0.01193	0.54317	0.55259

可以看到，更深的 autoencoder 效果未必比較好。

其中，單層 CNN 的深度分別為

3 + 3：3(input) >>> 32 >>> 64 >>> 128 >>> 64 >>> 32 >>> 3(output)

4 + 4：3(input) >>> 32 >>> 64 >>> 128 >>> 256 >>> 128 >>> 64 >>> 32 >>> 3(output)

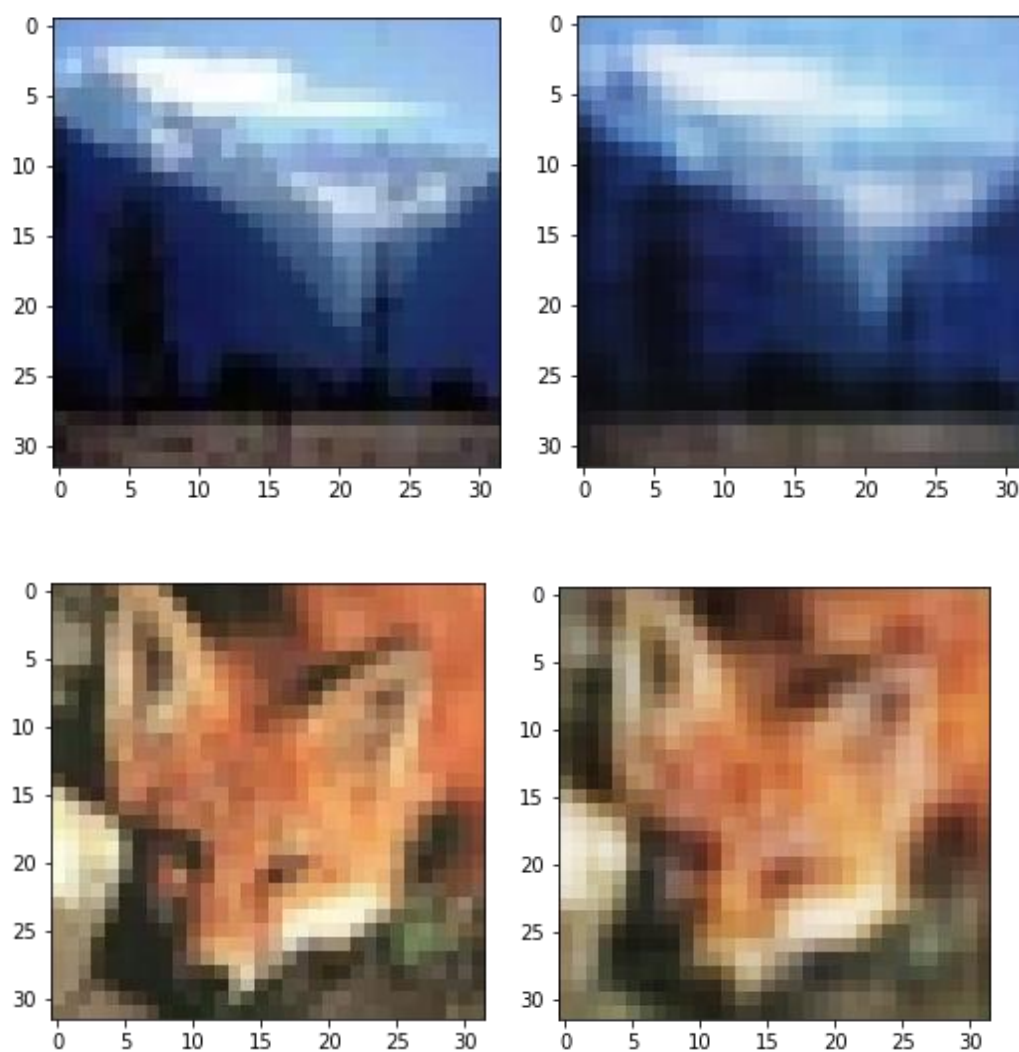
在同樣都使用 3 + 3 層的 autoencoder 以及 K-means 分群的前提之下，我分別使用一般的 PCA 降到二維，以及使用 RBF-Kernel PCA 降到二維，分數如下：

	Public accuracy	Private Accuracy
PCA	0.65222	0.65126
RBF-Kernel PCA	0.55666	0.56000

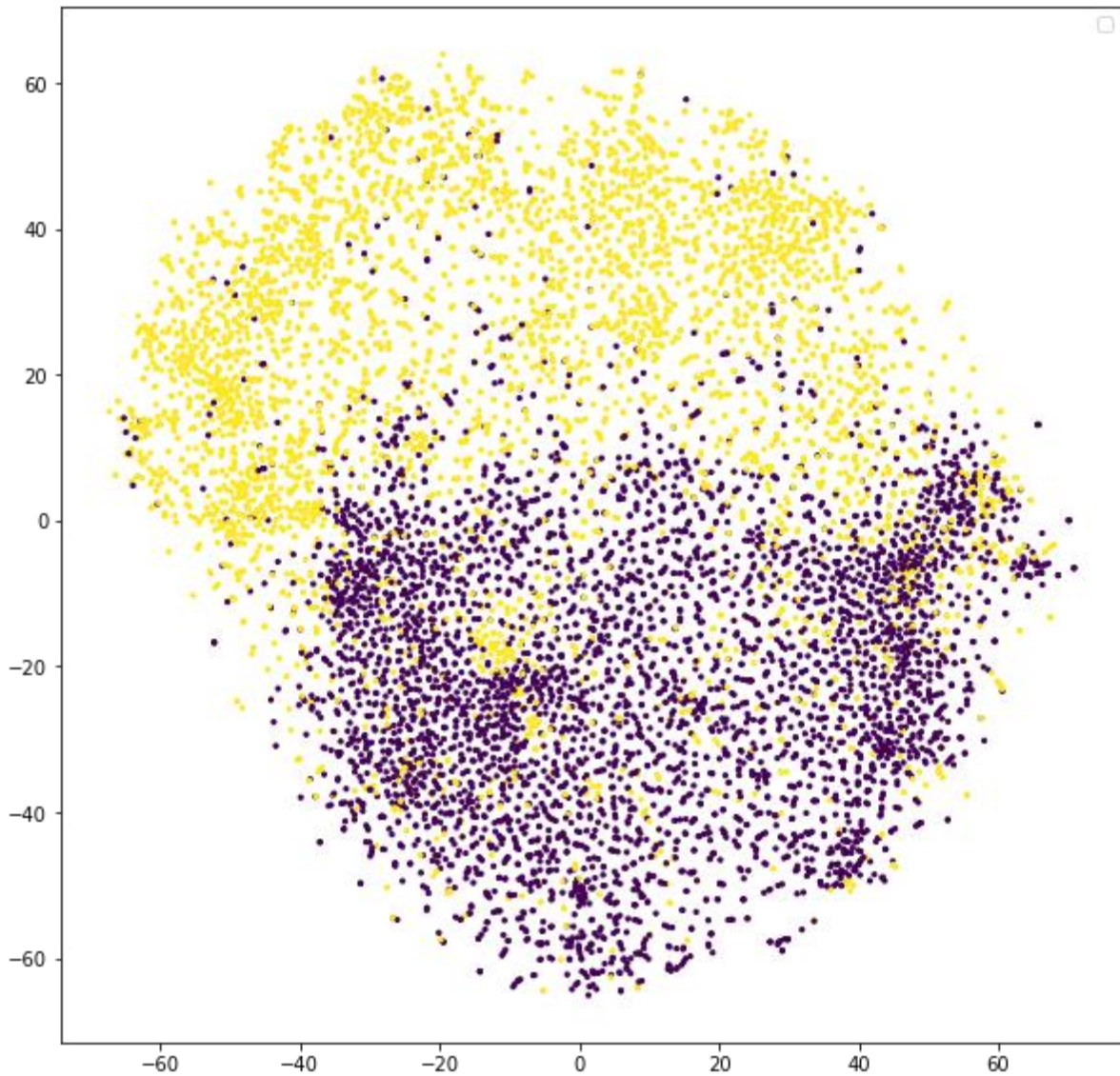
可以發現，直接降到二維時，PCA 的效果比使用 RBF-Kernel 的 PCA 好，使用 Kernel Function 並沒有使模型的表現比較好。

2. (1%) 從 dataset 選出 2 張圖，並貼上原圖以及經過 autoencoder 後 reconstruct 的圖片。

以下左邊為原圖，右邊為重建之後的圖



3. (1%) 在之後我們會給你 dataset 的 label。請在二維平面上視覺化 label 的分佈。  
下圖中，黃色的為 label 為 1 的點；紫色的為 label 為 2 的點



4. (3%) Refer to math problem  
[https://drive.google.com/file/d/1e\\_IDAV2yv0YEhIuVWpDdaH4Pzz5s1p2P/view?fbclid=IwAR0tO9NRxK9JZeUDNdawNuSbGTvqI7niuMX3Kkk9arauC8O6p6iJc7oMz84](https://drive.google.com/file/d/1e_IDAV2yv0YEhIuVWpDdaH4Pzz5s1p2P/view?fbclid=IwAR0tO9NRxK9JZeUDNdawNuSbGTvqI7niuMX3Kkk9arauC8O6p6iJc7oMz84)