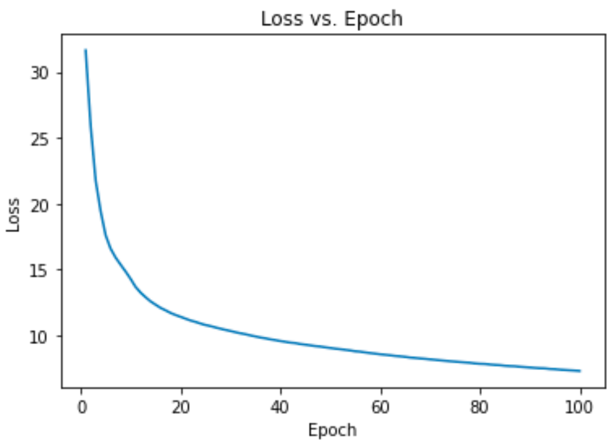
HW5 Report

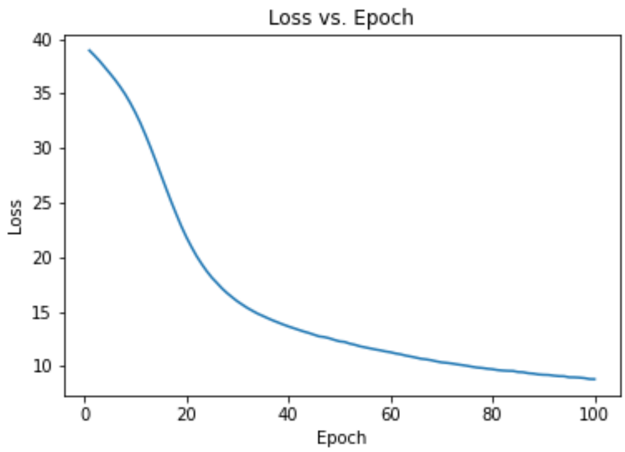
學號：B06705024　系級： 資管四　姓名：郭宇軒

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clustrering不用兩種。）

Simple_Model



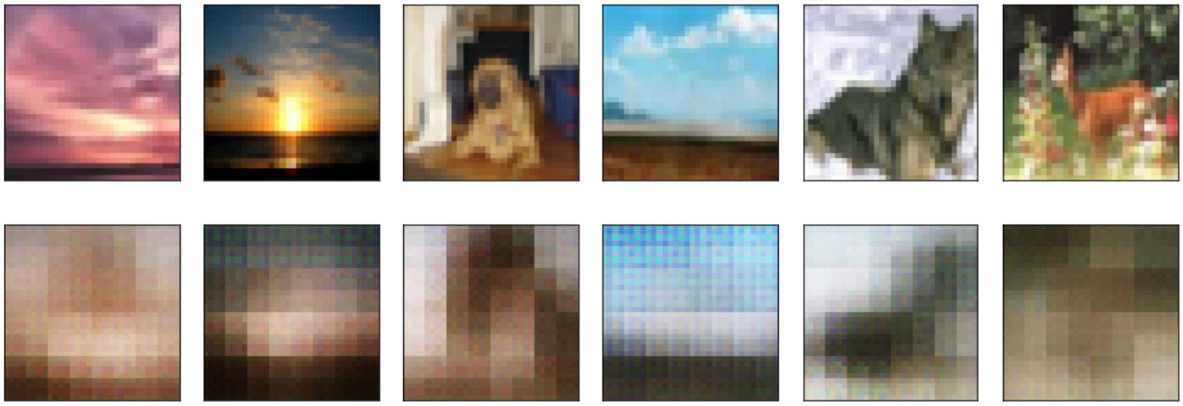
My_Model



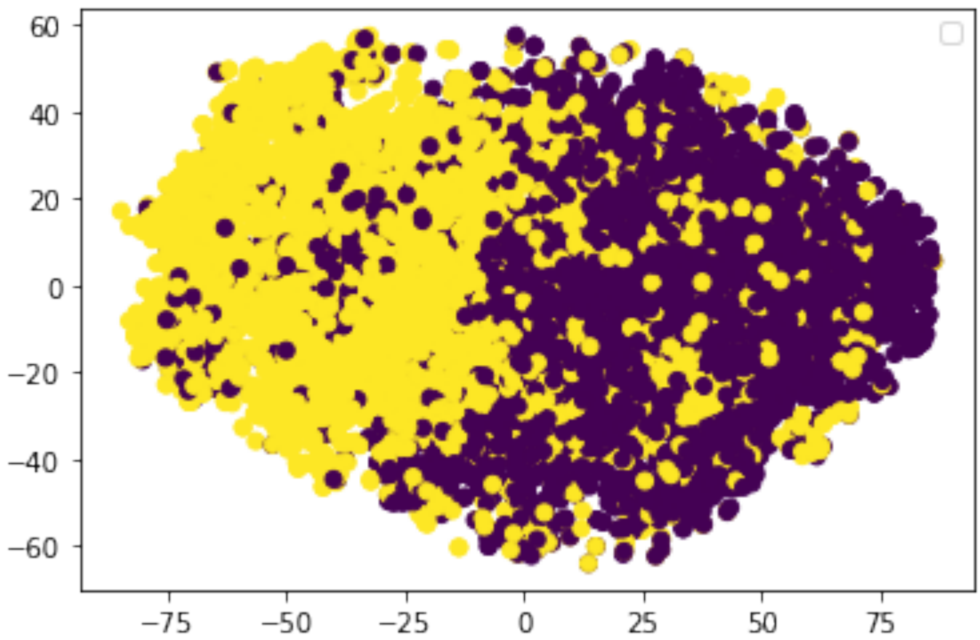
從左圖可以看到， simpel model 為我sample code中的 model，結構較簡單，因此從圖中可以看到一開始 loss 下降的比較快，約到 epoch 20時 loss 已經降到10 左右。而右圖為我自己的model，比較複雜層數也比較多，因此看到 loss 一開始下降的並不是這麼的快速，約到 epoch 20時 loss 才降到20左右。我們也可以看到 不管是哪一種分群方式，都是 My_Model的表現準確率較佳。

Kaggle Public Score	Simple_model	My_Model
KPCA	0.55365	0.69111
TSNE	0.60666	0.79761

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



3. (1%) 我們會給你dataset的label。請在二維平面上視覺化label的分佈。



4. (3%)Refer to math problem
https://drive.google.com/file/d/1-rmIFalj_6hEfJGOHLKUxlnoKMsKLHLf/view?usp=sharing

1.

when $t=1$, $z = Wx^1 + b = (0,0,0,1)\cdot(0,1,0,3)+0 = 3$

$$z_i = W_i x + b_i = (100,100,0,0)\cdot(0,1,0,3) - 10 = 90$$

$$z_f = (-100,-100,0,0)\cdot(0,1,0,3) + 110 = 10$$

$$z_o = (0,0,100,0)\cdot(0,1,0,3) - 10 = -10$$

$$c' = f(z_i)\,g(z) + c\,f(z_f) = \frac{1}{1+e^{-90}}\cdot 3 + 0\cdot\frac{1}{1+e^{-10}} \approx 3$$

$$y = f(z_o)\,h(c') = \frac{1}{1+e^{10}}\cdot 3 \approx 0.000136 \approx 0$$

$\Rightarrow$ we can get below:

T	z	z_i	z_f	z_o	c'	$f(z_i)$	$f(z_o)$	y ($= h(c')\cdot f(z_o)$)
1	3	90	10	-10	3	1	0	0
2	-2	90	10	90	1	1	1	1
3	4	90	-90	90	4	1	1	4
4	0	90	10	90	4	1	1	4
5	2	90	10	-10	6	1	0	0
6	-4	-10	110	90	6	0	1	6
7	1	190	-90	90	1	1	1	1
8	2	90	10	90	3	1	1	3

$\Rightarrow$ output $= 3$ #.

$$2.\ L = -\log \prod_{c \in C} P(W_{out}, W_{in}) = -\log \prod_{c \in C} \frac{\exp(u_c)}{\sum_{\bar{j} \in V} \exp(u_{\bar{j}})}$$

$$= -\sum \left(\log(\exp(u_c)) - \log\left(\sum \exp(u_r)\right) \right)$$

$$= -\sum u_c + \sum \log\left(\sum \exp(u_r)\right)$$

$$\frac{\partial L}{\partial W_{\bar{i}j}^{T}} = \frac{\partial L}{\partial W_{j\bar{i}}} = \sum_{k=1}^{V} \sum_{c=1}^{C} \frac{\partial L}{\partial u_{ck}} \cdot \frac{\partial}{\partial W_{\bar{i}j}} \left(\sum_{m=1}^{V} \sum_{l=1}^{V} W'_{mk} W_l x_l \right)$$

$$= \sum_{k=1}^{V} \sum_{c=1}^{C} (-\delta_{kk_c} + y_{ck}) W'_{jk} x_{\bar{i}} \ \#.$$

$$\frac{\partial L}{\partial W_{\bar{i}j}'^{T}} = \sum_{k=1}^{V} \sum_{c=1}^{C} \frac{\partial L}{\partial u_{c,k}} \frac{\partial u_{ck}}{\partial W_{\bar{i}j}'^{T}} = \sum_{c=1}^{C} \frac{\partial L}{\partial u_{c,j}} \cdot \frac{\partial u_{c,j}}{\partial W_{\bar{i}j}'^{T}}$$

$$= \sum_{c=1}^{C} (-\delta_{j j_c} + y_{cj}) \cdot \left(\sum_{k=1}^{V} W_{k\bar{i}} x_{k\bar{i}} \right) \#.$$