

學號：r08246009 系級：應數所 碩二 姓名：許哲維

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。)

<ans>

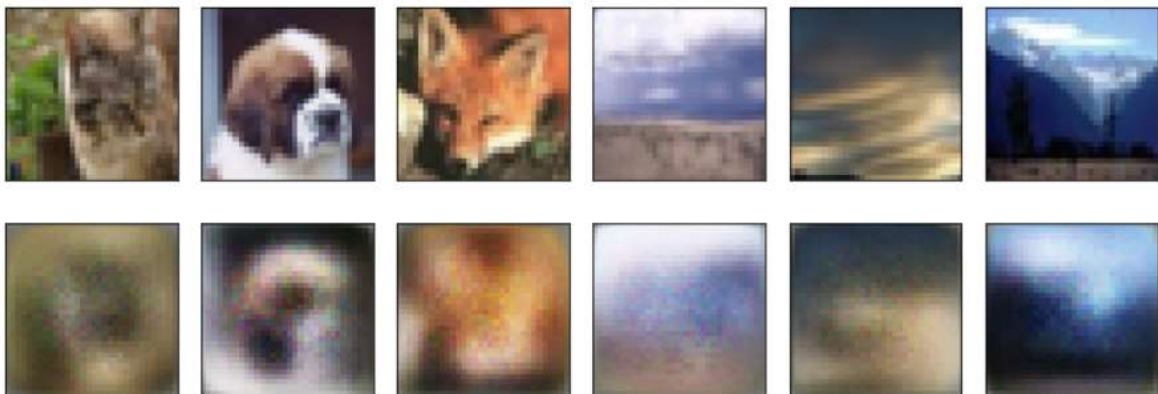
皆使用minibatch Kmeans的方式做分群，得到的kaggle prediction error如下

	Baseline	Deeper
Dimension=200	0.53063	0.79617(the best model)
Dimension=400	0.52904	0.76793

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

<ans>

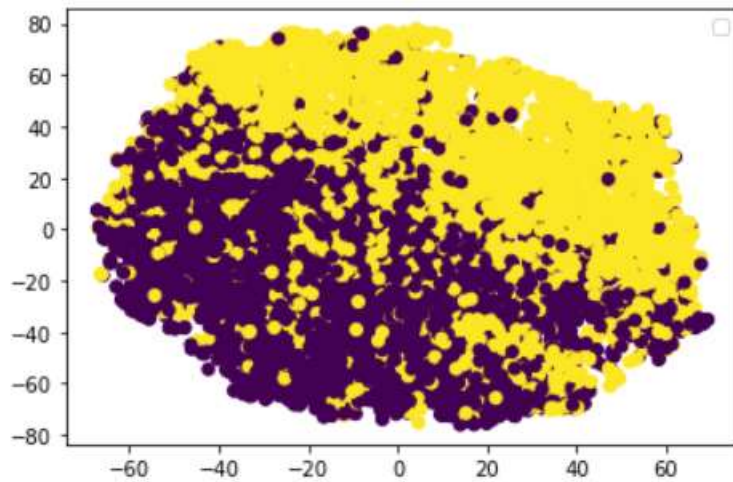
以下呈現第1,2,3,6,7,9張的原圖，以及reconstruct後的圖片



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

<ans>

Accuracy 有 0.611



4. (3%)Refer to math problem

LSTM cell

以下是  $y_1$  到  $y_8$  的值

```

1  import numpy as np
2  x = [[0,1,0,3],[1,0,1,-2],[1,1,1,4],[0,1,1,0],[0,1,0,2],[0,0,1,-4],[1,1,1,1],[1,0,1,2]]
3
4  x = np.array(x).T
5
6  w = np.array([[0,0,0,1]])
7  b = np.array([[0]])
8
9  wi = np.array([[100,100,0,0]])
10 bi = np.array([[10]])
11
12 wf = np.array([[-100,-100,0,0]])
13 bf = np.array([[110]])
14
15 wo = np.array([[0,0,100,0]])
16 bo = np.array([[10]])
17
18 z = w.dot(x) + b
19 zo = wo.dot(x) + bo
20 zi = wi.dot(x) + bi
21 zf = wf.dot(x) + bf
22
23 def f(x):
24     return 1/(1+np.exp(-x))
25 def g(x):
26     return x
27 def h(x):
28     return x
29 c = 0
30 c_tmp = f(zi)*g(z) + c*f(zf)
31 y = f(zo)*h(c_tmp)
32 y

```

```

array([[ 1.36193606e-04, -2.00000000e+00,  4.00000000e+00,
         0.00000000e+00,  9.07957374e-05, -1.81591475e-04,
         1.00000000e+00,  2.00000000e+00]])

```



Problem 2:

$$h = W^T x$$

$$L = -\log \prod_{c \in C} \frac{\exp(\mu_c)}{\sum_{i \in V} \exp(\mu_i)}$$

$$\mu = W^T h$$

$$= -\sum_{c \in C} [\mu_c - \log \sum_{i \in V} \exp(\mu_i)]$$

$$= [ |C| \log \sum_{i \in V} \exp(\mu_i) ] - \sum_{c \in C} \mu_c$$

①

$$\frac{\partial L}{\partial w_{ij}^T} = \frac{\partial L}{\partial \mu_{c_{ij}}} \times \frac{\partial \mu_{c_{ij}}}{\partial w_{ij}^T}$$

$$= h_i \times \left( \sum_{c \in C} \frac{\exp(\mu_{c_{ij}})}{\sum_{i \in V} \exp(\mu_i)} - \delta_{c_{ij}} \right)$$

②

$$\frac{\partial L}{\partial w_{ij}^T} = \left( \sum_{k \in V} \frac{\partial L}{\partial \mu_k} \frac{\partial \mu_k}{\partial h_i} \right) \frac{\partial h_i}{\partial w_{ij}^T}$$

$$= \left[ \sum_{k \in V} \sum_{c \in C} \left( \frac{\exp(\mu_{c,k})}{\sum_{i \in V} \exp(\mu_i)} - \delta_{c,k} \right) w_{ik}^T \right] x_j$$