學號：R06942082　系級： 電信碩一　姓名：黃釋平

1. (1%)請比較有無 normalize(rating)的差別。並說明如何 normalize.
(collaborator:)
normalize 方法:將 rating/5，即限制於 0~1 之間

	normalized	**No-normalize**
Training loss	0.136	0.699
Kaggle public	0.873	0.904

可見 normalize 有相當好的效果

2. (1%)比較不同的 latent dimension 的結果。
(collaborator:)

	128	**64**	**32**
Training loss (使用 mse)	0.029	0.030	0.030
Kaggle public	0.850	0.868	0.872

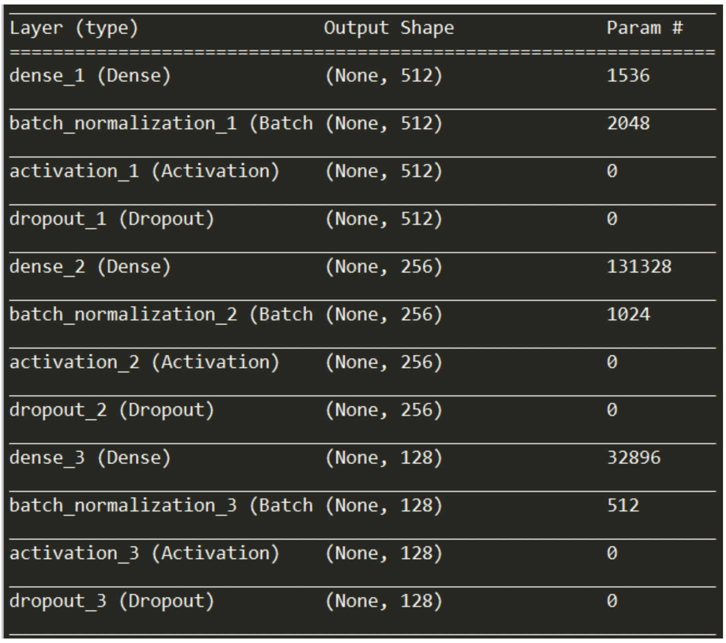
3. (1%)比較有無 bias 的結果。
(collaborator:)

	有 bias(normalized)	**無 bias(normalized)**
Training loss	0.136	0.699
Kaggle public	0.873	0.904

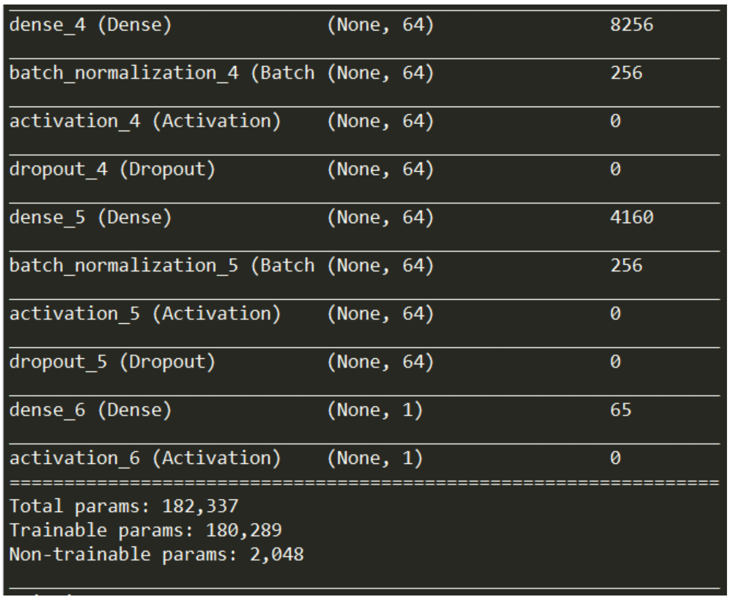
4. (1%)請試著用 DNN 來解決這個問題，並且說明實做的方法(方法不限)。並比較 MF 和 NN 的結果，討論結果的差異。
(collaborator:)

試過幾次 DNN，發現精準度無法提高，覺得像這種類型對應輸出的問題，也許比較適合 MF，也就是資料能畫出關係表格的。

model 如下:



```
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 512)               1536
_________________________________________________________________
batch_normalization_1 (Batch (None, 512)               2048
_________________________________________________________________
activation_1 (Activation)    (None, 512)               0
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_2 (Dense)              (None, 256)               131328
_________________________________________________________________
batch_normalization_2 (Batch (None, 256)               1024
_________________________________________________________________
activation_2 (Activation)    (None, 256)               0
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_3 (Dense)              (None, 128)               32896
_________________________________________________________________
batch_normalization_3 (Batch (None, 128)               512
_________________________________________________________________
activation_3 (Activation)    (None, 128)               0
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
```



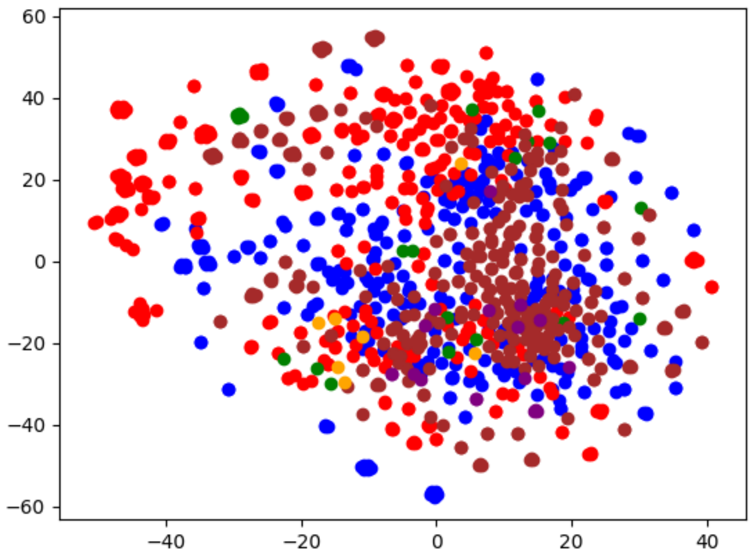
```
_________________________________________________________________
dense_4 (Dense)              (None, 64)                8256
_________________________________________________________________
batch_normalization_4 (Batch (None, 64)                256
_________________________________________________________________
activation_4 (Activation)    (None, 64)                0
_________________________________________________________________
dropout_4 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_5 (Dense)              (None, 64)                4160
_________________________________________________________________
batch_normalization_5 (Batch (None, 64)                256
_________________________________________________________________
activation_5 (Activation)    (None, 64)                0
_________________________________________________________________
dropout_5 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_6 (Dense)              (None, 1)                 65
_________________________________________________________________
activation_6 (Activation)    (None, 1)                 0
=================================================================
Total params: 182,337
Trainable params: 180,289
Non-trainable params: 2,048
_________________________________________________________________
```

最高 validation:**0.922**，kaggle:**1.144**

5. (1%)請試著將 movie 的 embedding 用 tsne 降維後，將 movie category 當作 label 來作圖。
(collaborator:)

分類(有多種性質的 movie 會取第一個作為代表)

1	2	3	4	5	6
Animation Children's Comedy Fantasy	Adventure Action Crime War	Thriller Horror	Drama Romance Musical	Documentary Sci-Fi Mystery Film-Noir	Western



目前還找不到更好的分類方式。

6. (BONUS)(1%)試著使用除了 rating 以外的 feature, 並說明你的作法和結果，結果好壞不會影響評分。

(collaborator:)

有嘗試過下列幾種作法:

(1)將每一行 user 的資訊(Gender,Age,Occupation,Zip-code)和 movie 的資訊(Genres*6，未達 6 種的補 0)取出，並合成一條 11 個 element 的向量，輸入 DNN，輸出為 1 個值，該值當作 bias 與原本的 MF 相加，輸出最後結果。但不論 Dense、batchnormalize、activation、drpoout、不使用 zip-code、減少 Genres 數量，結果 loss 都無法降低，上傳結果亦很慘(>1)。

(2)與上種類似，但是不輸入 DNN，而是將 user 資訊、資訊分別做 word embedding 後內積，也就是 MF，此值當作 bias 再與原本的 MF 相加，輸出最後結果。

以上經過長時間的搶救，最終宣告不治