學號：R06921066 系級：電機碩一 姓名：劉宇閎

1. (1%)請比較有無 normalize(rating)的差別。並說明如何 normalize.

	有 normalize	無 normalize
public	0.87146	0.86164
private	0.87174	0.86188

推論將 rating 做 normalize 後因分數間的比較性變低，所以有可能會導致系統判別結果不佳。

2. (1%)比較不同的 latent dimension 的結果。

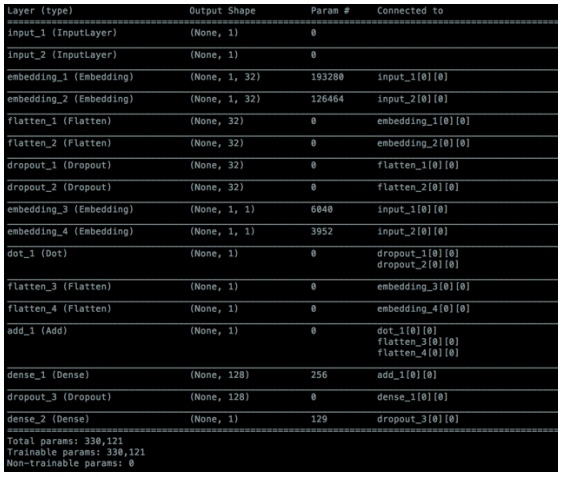
Latent dimension	4	32	64
public	0.8880	0.86164	0.86187
private	0.8853	0.86188	0.86080

3. (1%)比較有無 bias 的結果。

	有 bias	無 bias
public	0.86164	0.9756
private	0.86188	0.9766

根據此結果可以思考當 train 完後加入一個 bias 且獨立 tune 這個參數的結果，會比純粹 tune weight 的結果還要更好。

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。

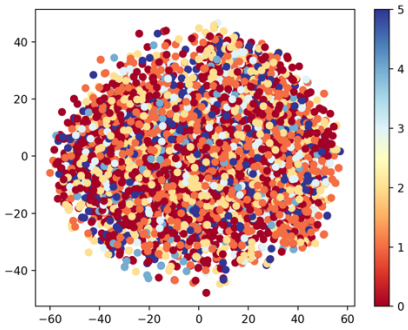


Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 1)	0	
input_2 (InputLayer)	(None, 1)	0	
embedding_1 (Embedding)	(None, 1, 32)	193280	input_1[0][0]
embedding_2 (Embedding)	(None, 1, 32)	126464	input_2[0][0]
flatten_1 (Flatten)	(None, 32)	0	embedding_1[0][0]
flatten_2 (Flatten)	(None, 32)	0	embedding_2[0][0]
dropout_1 (Dropout)	(None, 32)	0	flatten_1[0][0]
dropout_2 (Dropout)	(None, 32)	0	flatten_2[0][0]
embedding_3 (Embedding)	(None, 1, 1)	6040	input_1[0][0]
embedding_4 (Embedding)	(None, 1, 1)	3952	input_2[0][0]
dot_1 (Dot)	(None, 1)	0	dropout_1[0][0] dropout_2[0][0]
flatten_3 (Flatten)	(None, 1)	0	embedding_3[0][0]
flatten_4 (Flatten)	(None, 1)	0	embedding_4[0][0]
add_1 (Add)	(None, 1)	0	dot_1[0][0] flatten_3[0][0] flatten_4[0][0]
dense_1 (Dense)	(None, 128)	256	add_1[0][0]
dropout_3 (Dropout)	(None, 128)	0	dense_1[0][0]
dense_2 (Dense)	(None, 1)	129	dropout_3[0][0]

Total params: 330,121
Trainable params: 330,121
Non-trainable params: 0

使用在 MF 後再加入一層的 DNN 來進行比較，再經過 100 個 epochs 後可以得到比原本的 MF 更佳的結果 public = 0.855。

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



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。

   將所有的電影所對應的類別轉成 word vector 並將維度設為 4 ，user 資料也分別提取出來總共有四項，並將兩個作 dot，原本預期結果應該更好，但不知道是不是在資料處理上有誤，得到的 rmse 為 1.249。