

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

	Normalization	Non- normalization
kaggle RMSE	0.8686	0.8505

首先要說明一下，我的Normalization 是  $(rating - mean) / (max - min)$ ，Non-normalization 則是  $(rating - mean)$ ，在這樣子情況下發現，有 normalization 並沒有比較好。

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

Latent Dimension	8	16	32	64	128	256
Validation RMSE	0.8555	0.8557	0.8521	0.8486	0.8462	0.8466

可以大致看得出，latent dimension 是8 的時候效果就蠻不錯的了，但隨著 dimension 越來越高，訓練效果的確會比較好，只是沒有大幅度的進步。

3. (1%)比較有無bias的結果。  
(collaborator:)

	Bias	No Bias
Validation RMSE	0.8466	0.8539

沒有 Bias的情況明顯比較差，由此的確可以說明每個user對不同的電影有不同的評分傾向。

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:)

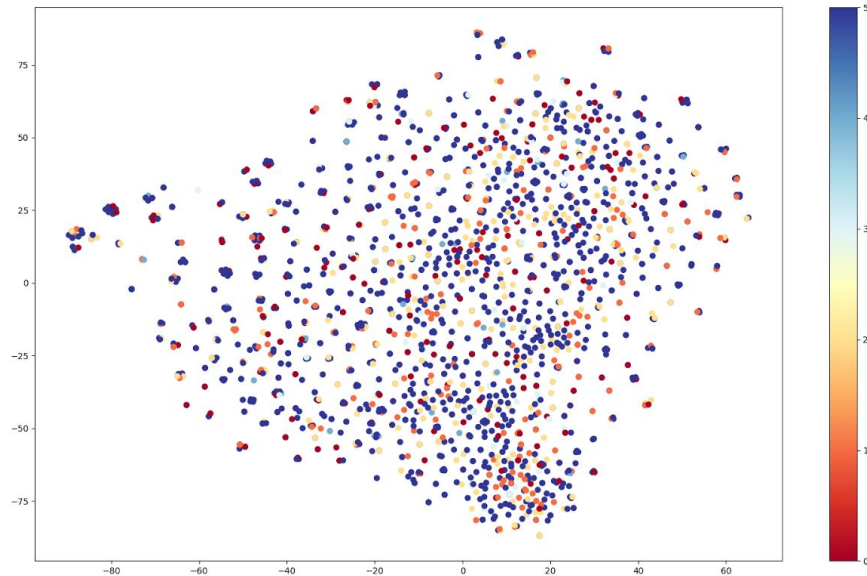
	MF	DNN
Validation RMSE	0.8466	0.8633

將user embedding以及movie embedding concatenate 起來，再將其丟入三層的 hidden layer，每層都有加入dropout 0.5，但這樣子效果不如普通的MF好。

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

圖。

(collaborator:)



我將所有電影類別分成如下五類，將我覺得類似的放在一起，例如西部、動作、戰爭的這種高張累得分成一類，在這之前嘗試了很多種，但都分不太開來，不過喜劇類的。

0:['Adventure','Children's','Animation']

1:['Film-Noir', 'Thriller','Crime','Horror'],

2:['Action', 'Western','War']

3:['Documentary']

4:['Sci-Fi', 'Fantasy', 'Mystery'],

5:[ 'Romance', 'Musical', 'Comedy','Drama']

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:)