

HW6 Report

姓名：施力維 系級：電機二 學號：b04502031

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

model 使用有 embedding dimension 為 10 且有 bias 的 MF 來進行測試，其中 Normalize 的方法是將 label 從(1, 5) linear mapping 到(-1, 1)，測試結果如下：

	Normalize	No normalize
Public	0.874	0.869
Private	0.866	0.861

在這個測試中 normalize 對於結果的影響比第三題中 bias 的影響還來得大。Normalize 後會有相當的進步。

2. (1 %)比較不同的 embedding dimension 的結果。

model 使用有 bias 的 MF 來進行測試，其他參數都固定

dim	1	2	5	10	20	50	100
Private	0.882	0.881	0.869	0.868	0.873	0.876	0.880

在這邊的測試中，5 維以及 10 維的 RMSE 最小，embedding dimension 而往左右變大變小則逐漸增加。

3. (1 %)比較有無 bias 的結果。

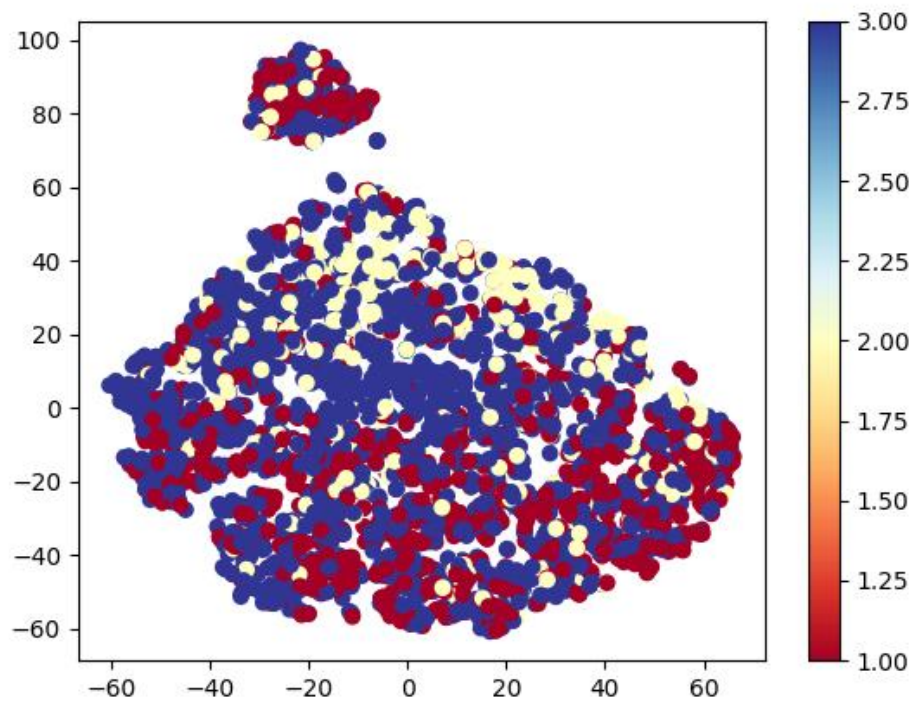
model 使用有 embedding dimension 為 10 的 MF 來進行測試，測試結果如下：

	bias	No bias
Public	0.874	0.875
Private	0.866	0.867

可以看到在這邊其實 bias 對於結果的影響沒有很多。

4. (1 %)請試著將 movie 的 embedding 用 tsne 降維後，將 movie category 當作 label 來作圖。

將 Drama, Musical 分成一類，Thriller, Horror, Crime 分為一類，其餘的分成一類，作圖如下，可以看出黃紅藍在下面那一大塊有根據層分開的傾向，黃在上，藍在中，紅在下，因此可以得出 movie 的 category 與使用者對於電影的評價相關的結論。



5. (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。

Model 嘗試使用有 category 來作為分類標準, 將 category 作為 BOW 轉成 19 維的向量, 在丟進 Dense 成 1 維, 最後再跟 MF 的結果做 Dense, 測試結果如下:

	MF	Using Category
Public	0.874	0.866
Private	0.866	0.859

由此可知加入 category 對於準確率確實有幫助, 驗證上議題的結論。