

## HW6

學號：b04901020 系級：電機三 姓名：解正平

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

(Collaborators:無)

答：

將全部 user 的 rating 實作 normalize(減平均除以標準差)，減少因個體差異而使評分差距大，但實際做出來的結果 normalize 沒有比較好，推測是因為 user vector 會影響 rating，做 normalize 會使得 user 影響不大，較能還原原本的 rating 分數。

| Dim = 128   | Public score | Private score |
|-------------|--------------|---------------|
| 有 normalize | 0.85122      | 0.84221       |
| 無 normalize | 0.84868      | 0.84122       |

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

(Collaborators: )

答：dimension 為 128 的效果最好，太小或太大分數都不高。

| Embedding dimesion | Public score | Private score |
|--------------------|--------------|---------------|
| 32                 | 0.85336      | 0.84646       |
| 64                 | 0.84872      | 0.84150       |
| 128                | 0.84868      | 0.84122       |
| 256                | 0.85061      | 0.84165       |
| 512                | 0.85535      | 0.84975       |

3. (1 %)比較有無 bias 的結果。

(Collaborators: )

答：有 bias 效果比較好。

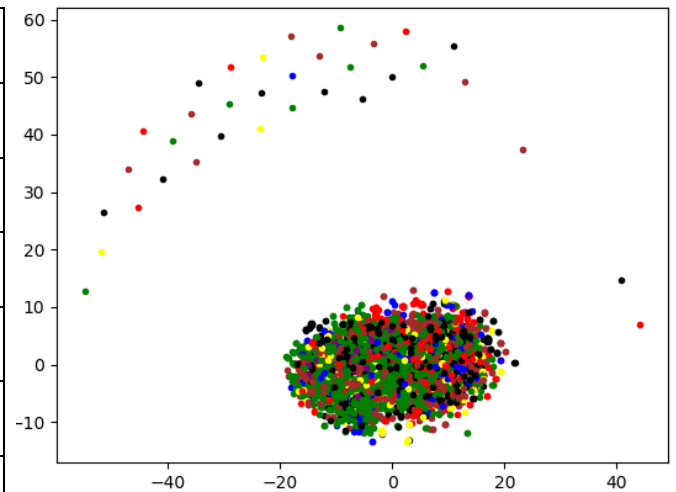
| Dim = 128 | Public score | Private score |
|-----------|--------------|---------------|
| 有 bias    | 0.84868      | 0.84122       |
| 無 bias    | 0.85148      | 0.84362       |

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

(Collaborators:無)

答：

| Color  | Movie Genre                        |
|--------|------------------------------------|
| Red    | Animation, Children's, Adventure   |
| Green  | Drama, Musical                     |
| Blue   | Fantasy, Sci-Fi, Mystery           |
| Black  | Crime, Thriller, Horror, Film-Noir |
| Yellow | Action, War, Western               |
| Brown  | Comedy, Romance                    |
| Purple | Documentary                        |



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。

(Collaborators: )

答：

除了原本 userID 和 movieID 各自做出來 128 維 vector，我還將 occupation 及 movie genre 實作 one-hot matrices，把兩個當作 vector 與 userID 及 movieID 做 inner product，看 user 及 movie 是否與 occupation 及 genre 有關聯。

另外我也將 gender 及 age 做成一維 vector，與 movieID 做 inner product，推測這兩個訊會與選擇的 movie 有關。

最後把這些 inner product 的結果 concatenate 在一起，加上 bias 項來預測 rating 分數。

|                | Public score | Private score |
|----------------|--------------|---------------|
| Normal feature | 0.84779      | 0.84039       |
| My feature     | 0.84868      | 0.84122       |