ML HW6

b05902127 資工二 劉俊緯

請比較有無**normalize**的差別。並說明如何**normalize.**

- Normalize: $\frac{score-1}{4}$, scale$[1,5]$to$[0,1]$
  - Normalization is better I think because of kernel initializer are Xavier.

	Private Score	Public Score
With Normalize	0.84728	0.85456
Without Normalize	0.85355	0.86438

比較不同的**embedding dimension** 的結果。

- In the result, I think larger embedding size with dropout will easily extract more features.

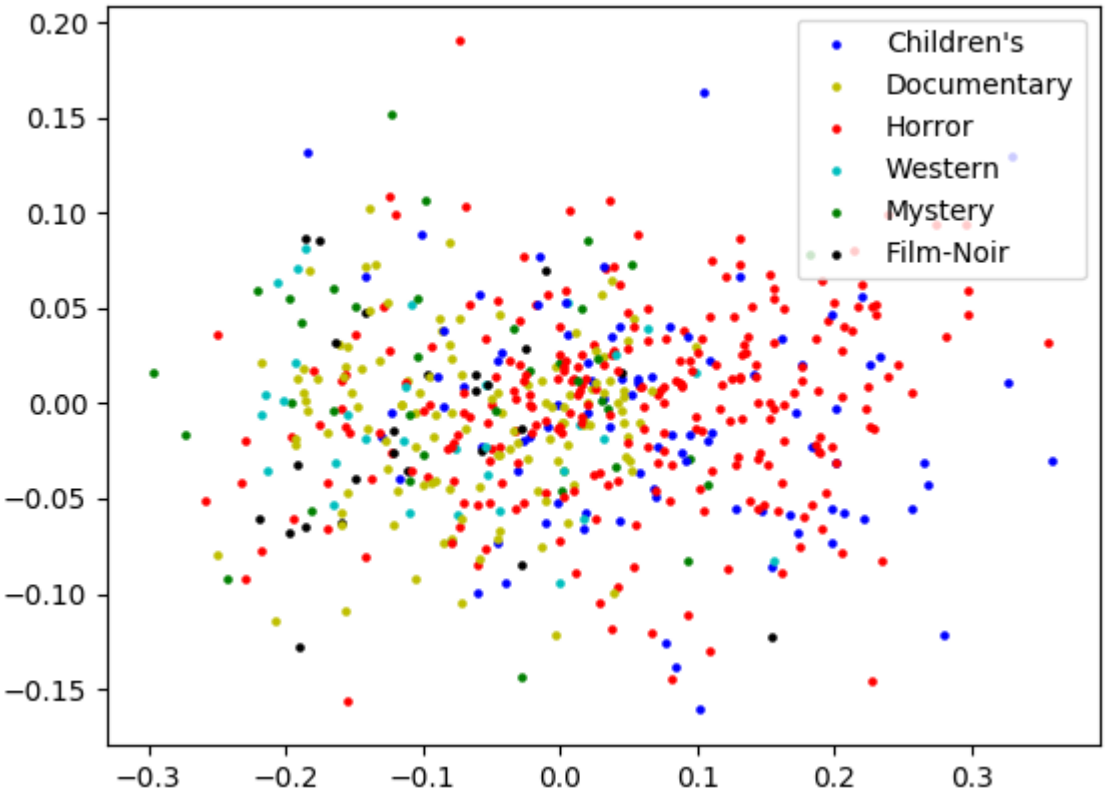
Embedding Size	Private Score	Public Score
300	0.84728	0.85456
100	0.86923	0.87387

比較有無**bias**的結果。

	Private Score	Public Score
With bias	0.86900	0.87477
Without bias	0.86923	0.87387

- In the result, adding bias will imporve a little bit both in private and public score.

請試著將**movie**的**embedding**用**tsne**降維後，將**movie category**當作**label**來作圖。



- 在這張圖我們可以看到Documentay和Children's是被明顯分開的。這與事實挺相符的，因為Children's與Documentay的性質相差甚遠。
- 此外，我們也觀察到Film-Noir(黑色電影)與Children's也被隔開來了。

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。

	Private Score	Public Score
More Feature	0.84728	0.85456
Less Feature	0.87117	0.87963

- concat user_embedding (dim 300 + bias) and Dense(300,user information)
- concat movie_embedding(dim 300+bias) and Dense(300,movie information)
- both two Dense(1024), and then Dot together.
- More detail code is is below.

```
user_output = Add()([Embedding(6041,300, input_length=[1])(user_idx) ,
Embedding(6041,1, input_length=[1])(user_idx)])
movie_output = Add()([Embedding(3953,300, input_length=[1])(movie_idx),
Embedding(3953,1, input_length=[1])(movie_idx)])
user_output = Dropout(0.5)(Flatten()(user_output))
movie_output= Dropout(0.5)(Flatten()(movie_output))
user_info_output = Dropout(0.5)(Dense(300)(user_info))

movie_info_output = Dropout(0.5)(Dense(300)(movie_info))
```

```
user = concatenate([user_output,user_info_output])
movie = concatenate([movie_output,movie_info_output])
user_output = Dense(1024,activation='tanh')(user)
movie_output= Dense(1024,activation='tanh')(movie)
user_output = Dropout(0.5)(user_output)
movie_output = Dropout(0.5)(movie_output)
y = Dot(1)([user_output , movie_output])
y = Dense(1)(y)
```