

機器學習作業六 報告

學號	系級	姓名
B03902015	資工三	簡瑋德

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

- 模型的架構

Layer	Parameters	Name	Input
Input	shape=(1,)	userId	-
Input	shape=(1,)	movieId	-
Embedding	input_dim=6041, output_dim=8, input_length=1	uEmb	userId
Flatten	-	uVec	uEmb
Embedding	input_dim=3953, output_dim=8, input_length=1	mEmb	movieId
Flatten	-	mVec	mEmb
Embedding	input_dim=6041, output_dim=1, input_length=1	uEmb2	userId
Flatten	-	uBias	uEmb2
Embedding	input_dim=3953, output_dim=8, input_length=1	mEmb2	movieId
Flatten	-	mBias	mEmb2
dot	[uVec, mVec], axes=-1	umDot	-
add	[umDot, uBias, mBias]	output	-

- 「Normalize」的方式 - 把「rate」除以5(最大值)
- 實驗參數
 - Batch Size = 128
 - Epochs = 20
 - Validation Split = 10%
- 實驗結果

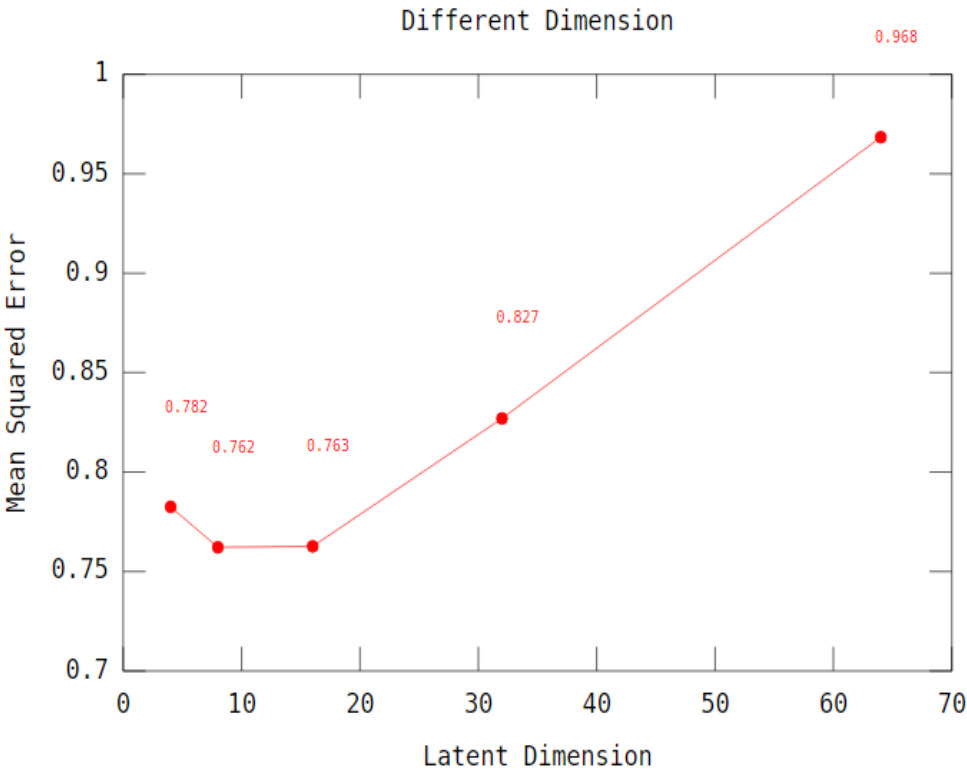
有無正規化	訓練過程	最低MSE																																																															
無	<p>Training Procedure</p> <table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>1.01</td><td>1.01</td></tr><tr><td>2</td><td>0.89</td><td>0.86</td></tr><tr><td>3</td><td>0.85</td><td>0.85</td></tr><tr><td>4</td><td>0.84</td><td>0.84</td></tr><tr><td>5</td><td>0.83</td><td>0.83</td></tr><tr><td>6</td><td>0.81</td><td>0.81</td></tr><tr><td>7</td><td>0.79</td><td>0.80</td></tr><tr><td>8</td><td>0.77</td><td>0.79</td></tr><tr><td>9</td><td>0.75</td><td>0.78</td></tr><tr><td>10</td><td>0.74</td><td>0.77</td></tr><tr><td>11</td><td>0.72</td><td>0.77</td></tr><tr><td>12</td><td>0.71</td><td>0.76</td></tr><tr><td>13</td><td>0.70</td><td>0.76</td></tr><tr><td>14</td><td>0.69</td><td>0.76</td></tr><tr><td>15</td><td>0.68</td><td>0.76</td></tr><tr><td>16</td><td>0.67</td><td>0.76</td></tr><tr><td>17</td><td>0.67</td><td>0.76</td></tr><tr><td>18</td><td>0.66</td><td>0.76</td></tr><tr><td>19</td><td>0.66</td><td>0.76</td></tr><tr><td>20</td><td>0.65</td><td>0.76</td></tr></tbody></table>	Epoch	train	validation	1	1.01	1.01	2	0.89	0.86	3	0.85	0.85	4	0.84	0.84	5	0.83	0.83	6	0.81	0.81	7	0.79	0.80	8	0.77	0.79	9	0.75	0.78	10	0.74	0.77	11	0.72	0.77	12	0.71	0.76	13	0.70	0.76	14	0.69	0.76	15	0.68	0.76	16	0.67	0.76	17	0.67	0.76	18	0.66	0.76	19	0.66	0.76	20	0.65	0.76	0.7564
Epoch	train	validation																																																															
1	1.01	1.01																																																															
2	0.89	0.86																																																															
3	0.85	0.85																																																															
4	0.84	0.84																																																															
5	0.83	0.83																																																															
6	0.81	0.81																																																															
7	0.79	0.80																																																															
8	0.77	0.79																																																															
9	0.75	0.78																																																															
10	0.74	0.77																																																															
11	0.72	0.77																																																															
12	0.71	0.76																																																															
13	0.70	0.76																																																															
14	0.69	0.76																																																															
15	0.68	0.76																																																															
16	0.67	0.76																																																															
17	0.67	0.76																																																															
18	0.66	0.76																																																															
19	0.66	0.76																																																															
20	0.65	0.76																																																															
有	<p>Training Procedure</p> <table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>0.85</td><td>0.85</td></tr><tr><td>2</td><td>0.80</td><td>0.80</td></tr><tr><td>3</td><td>0.76</td><td>0.78</td></tr><tr><td>4</td><td>0.72</td><td>0.78</td></tr><tr><td>5</td><td>0.70</td><td>0.78</td></tr><tr><td>6</td><td>0.69</td><td>0.77</td></tr><tr><td>7</td><td>0.68</td><td>0.77</td></tr><tr><td>8</td><td>0.67</td><td>0.78</td></tr><tr><td>9</td><td>0.67</td><td>0.78</td></tr><tr><td>10</td><td>0.66</td><td>0.78</td></tr><tr><td>11</td><td>0.66</td><td>0.78</td></tr><tr><td>12</td><td>0.66</td><td>0.78</td></tr><tr><td>13</td><td>0.66</td><td>0.78</td></tr><tr><td>14</td><td>0.65</td><td>0.78</td></tr><tr><td>15</td><td>0.65</td><td>0.78</td></tr><tr><td>16</td><td>0.65</td><td>0.78</td></tr><tr><td>17</td><td>0.65</td><td>0.78</td></tr><tr><td>18</td><td>0.65</td><td>0.79</td></tr><tr><td>19</td><td>0.65</td><td>0.78</td></tr><tr><td>20</td><td>0.65</td><td>0.79</td></tr></tbody></table>	Epoch	train	validation	1	0.85	0.85	2	0.80	0.80	3	0.76	0.78	4	0.72	0.78	5	0.70	0.78	6	0.69	0.77	7	0.68	0.77	8	0.67	0.78	9	0.67	0.78	10	0.66	0.78	11	0.66	0.78	12	0.66	0.78	13	0.66	0.78	14	0.65	0.78	15	0.65	0.78	16	0.65	0.78	17	0.65	0.78	18	0.65	0.79	19	0.65	0.78	20	0.65	0.79	0.7737
Epoch	train	validation																																																															
1	0.85	0.85																																																															
2	0.80	0.80																																																															
3	0.76	0.78																																																															
4	0.72	0.78																																																															
5	0.70	0.78																																																															
6	0.69	0.77																																																															
7	0.68	0.77																																																															
8	0.67	0.78																																																															
9	0.67	0.78																																																															
10	0.66	0.78																																																															
11	0.66	0.78																																																															
12	0.66	0.78																																																															
13	0.66	0.78																																																															
14	0.65	0.78																																																															
15	0.65	0.78																																																															
16	0.65	0.78																																																															
17	0.65	0.78																																																															
18	0.65	0.79																																																															
19	0.65	0.78																																																															
20	0.65	0.79																																																															

- 觀察和比較
 - 有做正規化的模型，不論在「Train」或「Validation」損失都下降得很快，約在四個「epochs」時「Validation」就收斂
 - 沒做正規化的模型，雖然下降得較慢，但反而有比較好的結果
 - 兩個模型的參數、架構相同，只差在「Loss」的計算，進而影響「Gradient」的大小，結果卻相差甚大

2. 比較不同的「latent dimension」的結果。

- 模型架構同上一題
- 實驗參數
 - Batch Size = 128
 - Epochs = 12
 - Validation Split = 10%

- 實驗結果



- 觀察和比較
 - 「Latent Dimension」是8或16的時候結果較好
 - 維度太低的话，損失不容易降低；維度超過16，則很容易「overfit」，可能需要搭配「early-stopping」或「regularization」來維護「Validation」和「Test」的表現

3. 比較有無「bias」的結果。

- 模型架構同第一題
- 實驗參數
 - Batch Size = 128
 - Epochs = 20
 - Validation Split = 10%
- 實驗結果

有無偏置	訓練過程	最低MSE																																																															
有	<p>Training Procedure</p> <table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>1.01</td><td>1.01</td></tr><tr><td>2</td><td>0.88</td><td>0.86</td></tr><tr><td>3</td><td>0.85</td><td>0.85</td></tr><tr><td>4</td><td>0.84</td><td>0.84</td></tr><tr><td>5</td><td>0.83</td><td>0.83</td></tr><tr><td>6</td><td>0.81</td><td>0.81</td></tr><tr><td>7</td><td>0.79</td><td>0.80</td></tr><tr><td>8</td><td>0.77</td><td>0.79</td></tr><tr><td>9</td><td>0.75</td><td>0.78</td></tr><tr><td>10</td><td>0.74</td><td>0.77</td></tr><tr><td>11</td><td>0.72</td><td>0.77</td></tr><tr><td>12</td><td>0.71</td><td>0.76</td></tr><tr><td>13</td><td>0.70</td><td>0.76</td></tr><tr><td>14</td><td>0.69</td><td>0.76</td></tr><tr><td>15</td><td>0.68</td><td>0.76</td></tr><tr><td>16</td><td>0.67</td><td>0.76</td></tr><tr><td>17</td><td>0.67</td><td>0.76</td></tr><tr><td>18</td><td>0.66</td><td>0.76</td></tr><tr><td>19</td><td>0.66</td><td>0.76</td></tr><tr><td>20</td><td>0.65</td><td>0.76</td></tr></tbody></table>	Epoch	train	validation	1	1.01	1.01	2	0.88	0.86	3	0.85	0.85	4	0.84	0.84	5	0.83	0.83	6	0.81	0.81	7	0.79	0.80	8	0.77	0.79	9	0.75	0.78	10	0.74	0.77	11	0.72	0.77	12	0.71	0.76	13	0.70	0.76	14	0.69	0.76	15	0.68	0.76	16	0.67	0.76	17	0.67	0.76	18	0.66	0.76	19	0.66	0.76	20	0.65	0.76	0.7564
Epoch	train	validation																																																															
1	1.01	1.01																																																															
2	0.88	0.86																																																															
3	0.85	0.85																																																															
4	0.84	0.84																																																															
5	0.83	0.83																																																															
6	0.81	0.81																																																															
7	0.79	0.80																																																															
8	0.77	0.79																																																															
9	0.75	0.78																																																															
10	0.74	0.77																																																															
11	0.72	0.77																																																															
12	0.71	0.76																																																															
13	0.70	0.76																																																															
14	0.69	0.76																																																															
15	0.68	0.76																																																															
16	0.67	0.76																																																															
17	0.67	0.76																																																															
18	0.66	0.76																																																															
19	0.66	0.76																																																															
20	0.65	0.76																																																															
無	<p>Training Procedure</p> <table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>0.97</td><td>0.88</td></tr><tr><td>2</td><td>0.86</td><td>0.86</td></tr><tr><td>3</td><td>0.85</td><td>0.85</td></tr><tr><td>4</td><td>0.84</td><td>0.85</td></tr><tr><td>5</td><td>0.84</td><td>0.84</td></tr><tr><td>6</td><td>0.83</td><td>0.83</td></tr><tr><td>7</td><td>0.81</td><td>0.82</td></tr><tr><td>8</td><td>0.79</td><td>0.80</td></tr><tr><td>9</td><td>0.77</td><td>0.79</td></tr><tr><td>10</td><td>0.76</td><td>0.78</td></tr><tr><td>11</td><td>0.75</td><td>0.77</td></tr><tr><td>12</td><td>0.73</td><td>0.77</td></tr><tr><td>13</td><td>0.72</td><td>0.76</td></tr><tr><td>14</td><td>0.71</td><td>0.76</td></tr><tr><td>15</td><td>0.70</td><td>0.76</td></tr><tr><td>16</td><td>0.69</td><td>0.76</td></tr><tr><td>17</td><td>0.69</td><td>0.76</td></tr><tr><td>18</td><td>0.68</td><td>0.76</td></tr><tr><td>19</td><td>0.68</td><td>0.76</td></tr><tr><td>20</td><td>0.67</td><td>0.76</td></tr></tbody></table>	Epoch	train	validation	1	0.97	0.88	2	0.86	0.86	3	0.85	0.85	4	0.84	0.85	5	0.84	0.84	6	0.83	0.83	7	0.81	0.82	8	0.79	0.80	9	0.77	0.79	10	0.76	0.78	11	0.75	0.77	12	0.73	0.77	13	0.72	0.76	14	0.71	0.76	15	0.70	0.76	16	0.69	0.76	17	0.69	0.76	18	0.68	0.76	19	0.68	0.76	20	0.67	0.76	0.7589
Epoch	train	validation																																																															
1	0.97	0.88																																																															
2	0.86	0.86																																																															
3	0.85	0.85																																																															
4	0.84	0.85																																																															
5	0.84	0.84																																																															
6	0.83	0.83																																																															
7	0.81	0.82																																																															
8	0.79	0.80																																																															
9	0.77	0.79																																																															
10	0.76	0.78																																																															
11	0.75	0.77																																																															
12	0.73	0.77																																																															
13	0.72	0.76																																																															
14	0.71	0.76																																																															
15	0.70	0.76																																																															
16	0.69	0.76																																																															
17	0.69	0.76																																																															
18	0.68	0.76																																																															
19	0.68	0.76																																																															
20	0.67	0.76																																																															

- 觀察和比較
 - 有偏置的模型參數稍微多了一些，「Train」的損失降得比較低
 - 在「Validation」上的表現，兩者相差不大，「MSE」都能降到0.75左右
4. 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。
- 模型架構

Layer	Parameters	Name	Input
Input	shape=(1,)	userId	-
Input	shape=(1,)	movieId	-
Embedding	input_dim=6041, output_dim=8, input_length=1	uEmb	userId
Flatten	-	uVec	uEmb
Embedding	input_dim=3953, output_dim=8, input_length=1	mEmb	movieId
Flatten	-	mVec	mEmb
concatenate	[uVec, mVec]	concat	-
Dense	units=1	output	concat

- 實驗參數
 - Batch Size = 128
 - Epochs = 20
 - Validation Split = 10%
- 實驗結果

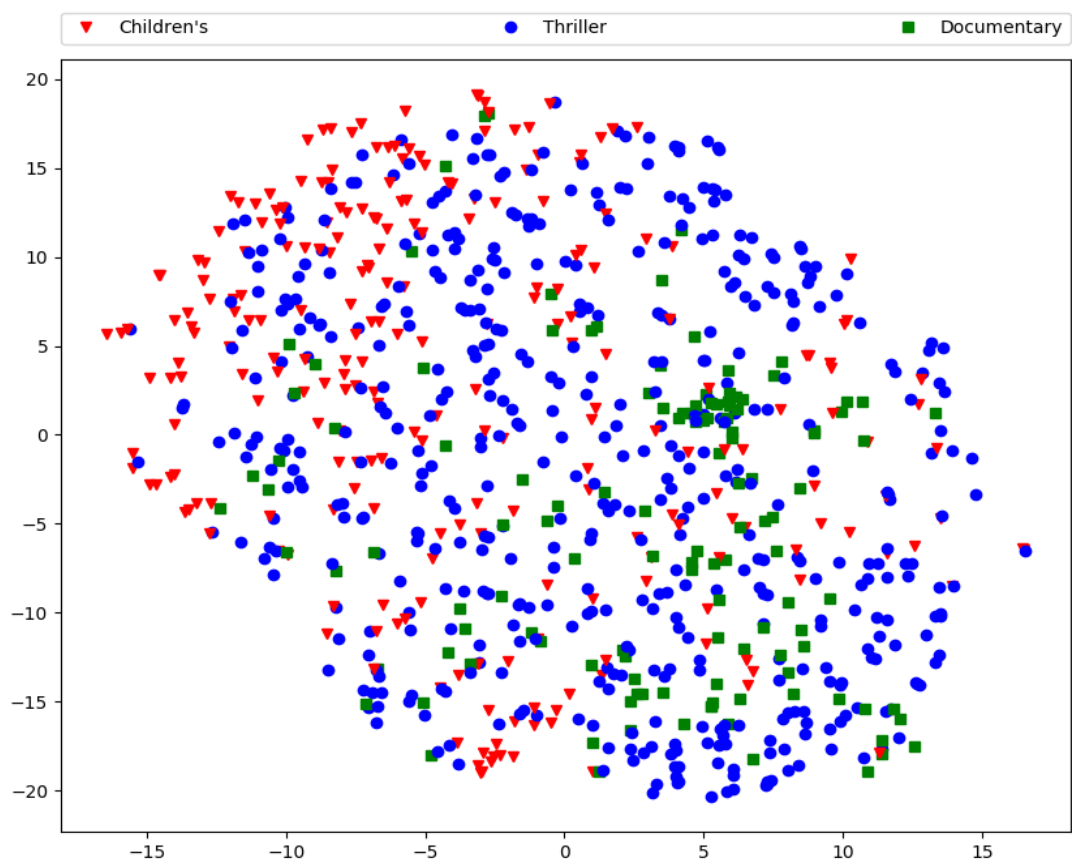
架構	訓練過程	最低MSE																																																															
FM	<div><p>Training Procedure</p><table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>1.01</td><td>1.01</td></tr><tr><td>2</td><td>0.88</td><td>0.86</td></tr><tr><td>3</td><td>0.85</td><td>0.85</td></tr><tr><td>4</td><td>0.84</td><td>0.84</td></tr><tr><td>5</td><td>0.83</td><td>0.83</td></tr><tr><td>6</td><td>0.81</td><td>0.81</td></tr><tr><td>7</td><td>0.79</td><td>0.80</td></tr><tr><td>8</td><td>0.77</td><td>0.79</td></tr><tr><td>9</td><td>0.75</td><td>0.78</td></tr><tr><td>10</td><td>0.74</td><td>0.77</td></tr><tr><td>11</td><td>0.72</td><td>0.77</td></tr><tr><td>12</td><td>0.71</td><td>0.76</td></tr><tr><td>13</td><td>0.70</td><td>0.76</td></tr><tr><td>14</td><td>0.69</td><td>0.76</td></tr><tr><td>15</td><td>0.68</td><td>0.76</td></tr><tr><td>16</td><td>0.67</td><td>0.76</td></tr><tr><td>17</td><td>0.67</td><td>0.76</td></tr><tr><td>18</td><td>0.66</td><td>0.76</td></tr><tr><td>19</td><td>0.66</td><td>0.76</td></tr><tr><td>20</td><td>0.65</td><td>0.76</td></tr></tbody></table></div>	Epoch	train	validation	1	1.01	1.01	2	0.88	0.86	3	0.85	0.85	4	0.84	0.84	5	0.83	0.83	6	0.81	0.81	7	0.79	0.80	8	0.77	0.79	9	0.75	0.78	10	0.74	0.77	11	0.72	0.77	12	0.71	0.76	13	0.70	0.76	14	0.69	0.76	15	0.68	0.76	16	0.67	0.76	17	0.67	0.76	18	0.66	0.76	19	0.66	0.76	20	0.65	0.76	0.7564
Epoch	train	validation																																																															
1	1.01	1.01																																																															
2	0.88	0.86																																																															
3	0.85	0.85																																																															
4	0.84	0.84																																																															
5	0.83	0.83																																																															
6	0.81	0.81																																																															
7	0.79	0.80																																																															
8	0.77	0.79																																																															
9	0.75	0.78																																																															
10	0.74	0.77																																																															
11	0.72	0.77																																																															
12	0.71	0.76																																																															
13	0.70	0.76																																																															
14	0.69	0.76																																																															
15	0.68	0.76																																																															
16	0.67	0.76																																																															
17	0.67	0.76																																																															
18	0.66	0.76																																																															
19	0.66	0.76																																																															
20	0.65	0.76																																																															
DNN	<div><p>Training Procedure</p><table border="1"><thead><tr><th>Epoch</th><th>train</th><th>validation</th></tr></thead><tbody><tr><td>1</td><td>0.85</td><td>0.85</td></tr><tr><td>2</td><td>0.84</td><td>0.84</td></tr><tr><td>3</td><td>0.84</td><td>0.84</td></tr><tr><td>4</td><td>0.83</td><td>0.83</td></tr><tr><td>5</td><td>0.83</td><td>0.83</td></tr><tr><td>6</td><td>0.83</td><td>0.83</td></tr><tr><td>7</td><td>0.83</td><td>0.83</td></tr><tr><td>8</td><td>0.82</td><td>0.83</td></tr><tr><td>9</td><td>0.82</td><td>0.83</td></tr><tr><td>10</td><td>0.82</td><td>0.83</td></tr><tr><td>11</td><td>0.82</td><td>0.83</td></tr><tr><td>12</td><td>0.82</td><td>0.83</td></tr><tr><td>13</td><td>0.82</td><td>0.83</td></tr><tr><td>14</td><td>0.82</td><td>0.82</td></tr><tr><td>15</td><td>0.82</td><td>0.82</td></tr><tr><td>16</td><td>0.82</td><td>0.82</td></tr><tr><td>17</td><td>0.82</td><td>0.82</td></tr><tr><td>18</td><td>0.82</td><td>0.82</td></tr><tr><td>19</td><td>0.81</td><td>0.82</td></tr><tr><td>20</td><td>0.81</td><td>0.83</td></tr></tbody></table></div>	Epoch	train	validation	1	0.85	0.85	2	0.84	0.84	3	0.84	0.84	4	0.83	0.83	5	0.83	0.83	6	0.83	0.83	7	0.83	0.83	8	0.82	0.83	9	0.82	0.83	10	0.82	0.83	11	0.82	0.83	12	0.82	0.83	13	0.82	0.83	14	0.82	0.82	15	0.82	0.82	16	0.82	0.82	17	0.82	0.82	18	0.82	0.82	19	0.81	0.82	20	0.81	0.83	0.8233
Epoch	train	validation																																																															
1	0.85	0.85																																																															
2	0.84	0.84																																																															
3	0.84	0.84																																																															
4	0.83	0.83																																																															
5	0.83	0.83																																																															
6	0.83	0.83																																																															
7	0.83	0.83																																																															
8	0.82	0.83																																																															
9	0.82	0.83																																																															
10	0.82	0.83																																																															
11	0.82	0.83																																																															
12	0.82	0.83																																																															
13	0.82	0.83																																																															
14	0.82	0.82																																																															
15	0.82	0.82																																																															
16	0.82	0.82																																																															
17	0.82	0.82																																																															
18	0.82	0.82																																																															
19	0.81	0.82																																																															
20	0.81	0.83																																																															

- 觀察和比較
 - 同樣的參數量，「DNN」明顯「underfit」，不論在「Train」或是「Validation」，「MSE」都超過0.8
 - 從實驗結果，可以得知，比起「concat」，「dot」本身有「consine-similarity」的意義，足以代表使用者和電影的相容性

5. 請試著將「movie」的「embedding」用「tsne」降維後，將「movie category」當作「label」來作圖

- 我選了三個比較特別的類別：「Thriller」、「Chrildren's」和「Documentary」

- 作圖結果



- 觀察和比較
 - 「Documentary」特別集中，我想它應該是最特別的一類電影，「Embedding」自然特別突出
 - 「Thriller」幾乎到處都有散布，可能是因為，不論甚麼樣的電影，多少都能夠加入一些恐怖、嚇人的元素，只是程度的差別，所以沒有很明顯的被區分出來
 - 「Children's」則落在偏左上角的部分。因為兒童電影的種類可能還能再細分，所以比起「Documentary」，它的分布比較沒那麼集中

6. 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。

- 額外使用的「feature」：年齡、性別、職業、電影年份、電影分類
- 實驗作法
 - 和「MF」相同，也有「User/Movie Embedding」和「User/Movie Bias」
 - 讓「職業」和「電影分類」，也各自有一個「Embedding」(職業使用「Embedding Layer」，而電影分類使用「Dense」)
 - 每筆資料現在都有4個「Embedding」，任兩個作內積，共可得到 $C_2^4 = 6$ 個內積結果
 - 把6個內積結果，以及「年齡」、「性別(0/1)」和「電影年份」，「concat」成一個9維的向量
 - 過一層「Dense Layer」，拿到一個實數，加上「User/Movie」的「Bias」，就是最後預測的評分
- 實驗結果
 - 「Validation MSE」最低可到0.74左右，相較於原始的「MF」，進步了一些