

به نام خدا دانشگاه تهران دانشگده مهندسی برق و کامپیوتر



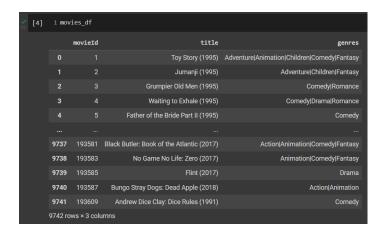
درس شبکههای عصبی و یادگیری عمیق تمرین اول

حسام اسدالهزاده – مسعود طهماسبی	نام و نام خانوادگی
810198429 - 810198346	شماره دانشجویی
14-14.19	تاریخ ارسال گزارش

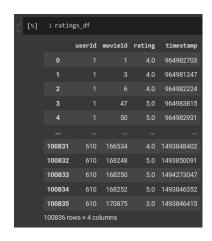
پاسخ ۳ – Restricted Boltzmann Machine

۳–۱. سیستم توصیهگر

A) پنج مورد نخست و آخر از فیلمها:



پنج مورد نخست و آخر از امتیازها:



ابعاد هر دو دیتاست:

```
[8] 1 print("Shape of movies dataset:", movies_df.shape)
2 print("Shape of ratings dataset:", ratings_df.shape)

Shape of movies dataset: (9742, 3)
Shape of ratings dataset: (100836, 4)
```

در دیتاست فیلمها، ۹۷۴۲ فیلم وجود دارد که هرکدام دارای ۳ ویژگی ۹۷۴۲ فیلم وجود دارد. میباشند. در دیتاست امتیازها نیز ۱۰۰۸۳۶ امتیاز مربوط به ۶۱۰ کاربر وجود دارد.

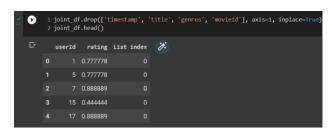
اضافه کردن ستون List index:

(10)	<pre>1 movies_df['List index'] = movies_df.index 2 movies_df</pre>									
		movieId	title	genres	List index					
	0		Toy Story (1995)	Adventure Animation Children Comedy Fantasy						
			Jumanji (1995)	Adventure Children Fantasy						
	2		Grumpier Old Men (1995)	ComedylRomance						
	3		Waiting to Exhale (1995)	Comedy Drama Romance						
	4		Father of the Bride Part II (1995)	Comedy						
	9737	193581	Black Butler: Book of the Atlantic (2017)	Action Animation Comedy Fantasy	9737					
	9738	193583	No Game No Life: Zero (2017)	Animation Comedy Fantasy	9738					
	9739	193585	Flint (2017)	Drama	9739					
	9740	193587	Bungo Stray Dogs: Dead Apple (2018)	Action Animation	9740					
	9741	193609	Andrew Dice Clay: Dice Rules (1991)	Comedy	9741					
	9742 rd	ows × 4 colu	umns							

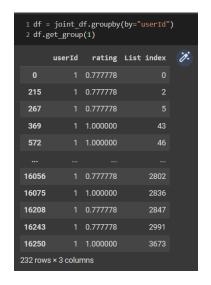
B) ادغام کردن دو دیتاست:

]	<pre>1 joint_df = ratings_df.merge(movies_df, left_on='movieId', right_on='movieId') 2 joint_df</pre>							
		userId	movieId	rating	timestamp	title	genres	List index
				0.777778	964982703	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	
				0.777778	847434962	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	C
				0.888889	1106635946	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	
	3	15		0.444444	1510577970	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	
	4	17		0.888889	1305696483	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	
	100831	610	160341	0.444444	1479545749	Bloodmoon (1997)	Action Thriller	932
	100832	610	160527	0.888889	1479544998	Sympathy for the Underdog (1971)	Action Crime Drama	9330
	100833	610	160836	0.555556	1493844794	Hazard (2005)	Action Drama Thriller	9342
	100834	610	163937	0.666667	1493848789	Blair Witch (2016)	Horror Thriller	9389
	100835	610	163981	0.666667	1493850155	31 (2016)	Horror	939
100836 rows × 7 columns								

C) ستون timestamp حاوی اطلاعات زمانی مربوط به امتیازات کاربران است که واضحا ارتباط یا correlation با خروجی سیستم پیشنهادگر ما ندارد پس قابل حذف است. ستونهای etitle و genres با خروجی سیستم پیشنهادگر ما ندارد پس قابل حذف است. ستونهای etitle و movield به عنوان حاوی اطلاعات متنی (string) مربوط به فیلمها هستند که میتوان به جای آن از string) مربوط به عنوان ویژگی عددی استفاده کرد. البته از ویژگی المovield نیز نمیتوان استفاده کرد چراکه تعداد فیلمها و درنتیجه تعداد نورونهای every etitle المستند ولی بازه movield از ۱ تا ۱۹۳۶۰۹ است درنتیجه به جای این ویژگی از ستون List index که در قسمت الف ایجاد شد استفاده میکنیم و ستون movield را نیز حذف میکنیم.



(D



Œ



H) در مرحلهی ترین کردن مدل باید به این نکته توجه کنیم که نباید اطلاعات کاربر ۷۵ را به مدل بدهیم چراکه این کار موجب data leakage می شود. نتایج زیر، فیلمهای پیشنهادی توسط مدل را نشان می دهد.

```
[41] 1 movies df['title'][ind]
                                                        Jurassic Park (1993)
      418
                                                    Sixth Sense, The (1999)
Fight Club (1999)
      2078
                                                         Pulp Fiction (1994)
Braveheart (1995)
                 Star Wars: Episode VI - Return of the Jedi (1983)
                 Raiders of the Lost Ark (Indiana Jones and the...
                                       Terminator 2: Judgment Day (1991)
                                                     American Beauty (1999)
Godfather, The (1972)
      2145
                                       Saving Private Ryan (1998)
Silence of the Lambs, The (1991)
      1503
                 Star Wars: Episode V - The Empire Strikes Back...
Star Wars: Episode IV - A New Hope (1977)
                                                          Matrix, The (1999)
      Name: title, dtype: object
```

نکته ی جالب توجه در فیلمهای پیشنهادی این است که در میان این فیلمها، برخی فیلمها توسط کاربر با کربر با مره دهی شدهاند. به عنوان مثال فیلم Star Wars: Episode IV - A New Hope توسط این کاربر با امتیاز ۵ رتبهبندی شده است؛ این یعنی مدل به خوبی توانسته است تشخیص دهد که این کاربر فیلم ذکرشده را می پسندد. همچنین مدل به خوبی تشخیص داده است که این کاربر به مجموعه فیلمهای کلات کلات و قسمتهای دیگر این مجموعه را نیز پیشنهاد داده است؛ به طوریکه کاربر ۵ قسمت از این مجموعه فیلم را تماشا کرده است و مدل، قسمت ششم این مجموعه را نیز پیشنهاد داده است.

•