In [1]:

```
import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings("ignore")
```

In [2]:

```
movie = pd.read_csv("movies.csv")
rating = pd.read_csv("ratings.csv")
```

In [3]:

```
movie.head()
```

Out[3]:

	movieId	title	genres
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	2	Jumanji (1995)	Adventure|Children|Fantasy
2	3	Grumpier Old Men (1995)	Comedy|Romance
3	4	Waiting to Exhale (1995)	Comedy|Drama|Romance
4	5	Father of the Bride Part II (1995)	Comedy

In [4]:

```
rating.head()
```

Out[4]:

	userId	movieId	rating	timestamp
0	1	31	2.5	1260759144
1	1	1029	3.0	1260759179
2	1	1061	3.0	1260759182
3	1	1129	2.0	1260759185
4	1	1172	4.0	1260759205

In [5]:

```
movie_final = pd.merge(movie,rating,on="movieId").drop(["genres","timestamp","movieId"],
axis=1)
```

In [6]:

```
movie_final["userId"] = "User"+movie_final["userId"].astype("str")
```

In [7]:

```
movie_final
```

Out[7]:

	title	userId	rating
0	Toy Story (1995)	User7	3.0
1	Toy Story (1995)	User9	4.0
2	Toy Story (1995)	User13	5.0
3	Toy Story (1995)	User15	2.0
4	Toy Story (1995)	User19	3.0
...	...	...	...
99999	The Last Brickmaker in America (2001)	User287	5.0
100000	Stranger Things	User73	4.5
100001	Rustom (2016)	User611	5.0
100002	Mohenjo Daro (2016)	User611	3.0
100003	The Beatles: Eight Days a Week - The Touring Y...	User547	5.0

100004 rows × 3 columns

In [8]:

```
from surprise import Dataset, SVD, accuracy, Reader
from surprise.model_selection import train_test_split
```

In [9]:

```
movie_final1 = movie_final
```

In [10]:

```
reader = Reader(rating_scale=(1,5))
```

In [11]:

```
data = Dataset.load_from_df(movie_final,reader)
```

In [12]:

```
trainset,testset = train_test_split(data,test_size=0.3,random_state=1)
```

In [13]:

```
svd = SVD(n_factors=100)
```

In [14]:

```
svd.fit(trainset)
```

Out[14]:

```
<surprise.prediction_algorithms.matrix_factorization.SVD at 0x1d2fedf6ac0>
```

In [15]:

```
predictions = svd.test(testset)
accuracy.rmse(predictions)
```

```
RMSE: 0.9021
```

Out[15]:

```
0.9020826498145906
```

In [16]:

```
movie_final.iloc[200]
user = "User611"
movie = "Mohenjo Daro (2016)"
```

In [17]:

```
svd.predict(user,movie)
```

Out[17]:

```
Prediction(uid='User611', iid='Mohenjo Daro (2016)', r_ui=None, est=3.540205
9941144537, details={'was_impossible': False})
```

In [18]:

```
movie_user_matrix = movie_final1.pivot_table(index="userId",columns="title",values="rating"
```

In [19]:

```
movie_user_matrix.head()
```

Out[19]:

title	"Great Performances" Cats (1998)	$9.99 (2008)	'Hellboy': The Seeds of Creation (2004)	'Neath the Arizona Skies (1934)	'Round Midnight (1986)	'Salem's Lot (2004)	'Til There Was You (1997)	'burbs, The (1989)	'night Mother (1986)
userId									
User1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
User10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
User100	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
User101	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
User102	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

5 rows × 9064 columns

In [20]:

```
movie_user_matrix.fillna(0,inplace=True)
```

In [21]:

```
movie_user_matrix.head()
```

Out[21]:

title	"Great Performances" Cats (1998)	$9.99 (2008)	'Hellboy': The Seeds of Creation (2004)	'Neath the Arizona Skies (1934)	'Round Midnight (1986)	'Salem's Lot (2004)	'Til There Was You (1997)	'burbs, The (1989)	'night Mother (1986)
userId									
User1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
User10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
User100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
User101	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
User102	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5 rows × 9064 columns

In [22]:

```
movie_watched = movie_user_matrix["Mohenjo Daro (2016)"]
movie_watched
```

Out[22]:

```
userId
User1      0.0
User10     0.0
User100    0.0
User101    0.0
User102    0.0
           ...
User95     0.0
User96     0.0
User97     0.0
User98     0.0
User99     0.0
Name: Mohenjo Daro (2016), Length: 671, dtype: float64
```

In [23]:

```
similarity_scores = movie_user_matrix.corrwith(movie_watched)
similarity_scores
```

Out[23]:

```
title
"Great Performances" Cats (1998)            -0.001718
$9.99 (2008)                                -0.002514
'Hellboy': The Seeds of Creation (2004)     -0.001493
'Neath the Arizona Skies (1934)             -0.001493
'Round Midnight (1986)                      -0.001666
                                                ...
xXx (2002)                                  -0.006556
xXx: State of the Union (2005)              -0.001493
¡Three Amigos! (1986)                       -0.008172
À nous la liberté (Freedom for Us) (1931)   -0.001493
İtirazım Var (2014)                         -0.001493
Length: 9064, dtype: float64
```

In [24]:

```
similarity_scores.sort_values(ascending=False).head(10)
```

Out[24]:

```
title
Mohenjo Daro (2016)                                              1.000000
Yeh Jawaani Hai Deewani (2013)                                   1.000000
Student of the Year (2012)                                       1.000000
Bajirao Mastani (2015)                                           1.000000
Rustom (2016)                                                    1.000000
Sherlock: The Abominable Bride (2016)                            0.552438
Hachiko: A Dog's Story (a.k.a. Hachi: A Dog's Tale) (2009)       0.488552
Pirates of the Caribbean: On Stranger Tides (2011)               0.370849
Titanic (1953)                                                   0.362719
Captain Phillips (2013)                                          0.296702
dtype: float64
```

In [ ]: