

# Esercitazione 8 MNI

Studente: Giuseppe Cardaropoli

Matricola: 052251310

## Esercizio n°1

### Punto a

```
Misura di Similarità: Correlazione di Pearson

Lisa Rose Similarity: 0.99
Film: Lady in the Water - Voto: 2.5 - Film Similarity: 2.48
Film: Just My Luck - Voto: 3.0 - Film Similarity: 2.97
Film: The Night Listener - Voto: 3.0 - Film Similarity: 2.97

Gene Seymour Similarity: 0.38
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 1.14
Film: Just My Luck - Voto: 1.5 - Film Similarity: 0.57
Film: The Night Listener - Voto: 3.0 - Film Similarity: 1.14

Claudia Puig Similarity: 0.89
Film: Just My Luck - Voto: 3.0 - Film Similarity: 2.68
Film: The Night Listener - Voto: 4.5 - Film Similarity: 4.02

Mick LaSalle Similarity: 0.92
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 2.77
Film: Just My Luck - Voto: 2.0 - Film Similarity: 1.85
Film: The Night Listener - Voto: 3.0 - Film Similarity: 2.77

Jack Matthews Similarity: 0.66
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 1.99
Film: The Night Listener - Voto: 3.0 - Film Similarity: 1.99

Film consigliati per Toby: ['The Night Listener', 'Lady in the Water', 'Just My Luck']
```

A lato sono riportati i risultati del programma `collaborative_filtering.py`

Tali risultati combaciano con quelli attesi:

Critic	Similarity	Night	S.xNight	Lady	S.xLady	Luck	S.xLuck
Rose	0.99	3.0	2.97	2.5	2.48	3.0	2.97
Seymour	0.38	3.0	1.14	3.0	1.14	1.5	0.57
Puig	0.89	4.5	4.02			3.0	2.68
LaSalle	0.92	3.0	2.77	3.0	2.77	2.0	1.85
Matthews	0.66	3.0	1.99	3.0	1.99		
Total			12.89		8.38		8.07
Sim. Sum			3.84		2.95		3.18
Total/Sim. Sum			3.35		2.83		2.53

L'utente più simile a Toby risulta essere Lisa Rose con similarity 0.99.

### Punto b

```
Misura di Similarità: Distanza Euclidea

Lisa Rose Similarity: 0.35
Film: Lady in the Water - Voto: 2.5 - Film Similarity: 0.87
Film: Just My Luck - Voto: 3.0 - Film Similarity: 1.04
Film: The Night Listener - Voto: 3.0 - Film Similarity: 1.04

Gene Seymour Similarity: 0.26
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 0.77
Film: Just My Luck - Voto: 1.5 - Film Similarity: 0.39
Film: The Night Listener - Voto: 3.0 - Film Similarity: 0.77

Michael Phillips Similarity: 0.39
Film: Lady in the Water - Voto: 2.5 - Film Similarity: 0.97
Film: The Night Listener - Voto: 4.0 - Film Similarity: 1.55

Claudia Puig Similarity: 0.36
Film: Just My Luck - Voto: 3.0 - Film Similarity: 1.07
Film: The Night Listener - Voto: 4.5 - Film Similarity: 1.61

Mick LaSalle Similarity: 0.4
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 1.2
Film: Just My Luck - Voto: 2.0 - Film Similarity: 0.8
Film: The Night Listener - Voto: 3.0 - Film Similarity: 1.2

Jack Matthews Similarity: 0.27
Film: Lady in the Water - Voto: 3.0 - Film Similarity: 0.8
Film: The Night Listener - Voto: 3.0 - Film Similarity: 0.8

Film consigliati per Toby: ['The Night Listener', 'Lady in the Water', 'Just My Luck']
```

A lato sono riportati i risultati del programma `collaborative_filtering.py` utilizzando la **distanza euclidea** come misura di similarità.

Possiamo notare che vengono considerati tutti gli utenti del dataset a differenza dell'esecuzione con la correlazione di Pearson in cui consideriamo solo gli utenti che hanno valutato TUTTI i film valutati da Toby. Ad esempio, l'utente Micheal Phillips non è stato considerato nella prima esecuzione perché non ha valutato il film "You, Me and Dupree".

Adesso, con la distanza euclidea, l'utente più simile a Toby è Claudia Puig con similarity 0.4.

In entrambi i casi la lista dei film consigliati è la stessa

## Punto c

La seguente stima di valutazione  $\hat{r}_{ki} = \mu_k + \frac{\sum_{j \in N(i)} \text{sim}(k, j)(r_{ji} - \mu_j)}{\sum_{j \in N(i)} \text{sim}(k, j)}$ , è implementata in questa funzione:

```
def recommendations_approximation():  
    # Per ogni utente k  
    for user in dataset:  
        average_vote = 0.0  
        average_local = 0.0  
        sum_sims = 0.0  
        sum_sims_weighted = 0.0  
  
        print(f"\nUtente: {user}")  
  
        # Calcolo il suo voto medio  
        for item in dataset[user]:  
            average_vote += dataset[user][item]  
        average_vote /= len(dataset[user])  
  
        print(f"Voto medio: {round(average_vote,2)}")  
  
        # Per ogni film i che ha votato l'utente k  
        for item in dataset[user]:  
            for other in dataset:  
                # saltiamo l'utente stesso  
                if other == user:  
                    continue  
  
                # consideriamo solo gli altri utenti che hanno votato lo stesso film  
                if item in dataset[other]:  
                    # ne calcoliamo la similarità  
                    sim = pearson_correlation(user, other)  
  
                    # calcolo la media di questo utente  
                    for other_item in dataset[other]:  
                        average_local += dataset[other][other_item]  
                    average_local /= len(dataset[other])  
  
                    # aggiorniamo le somme delle similarità  
                    sum_sims_weighted += sim * (dataset[other][other_item] - average_local)  
                    sum_sims += sim  
  
                # aggiorniamo la raccomandazione dell'utente k sul prodotto i  
                dataset[user][item] = average_vote + (sum_sims_weighted / sum_sims)  
  
        print(f"Raccomandazione per il prodotto {item}: {round(dataset[user][item],2)}")
```

```
Lisa Rose Similarity: 0.19  
Film: Lady in the Water - Voto: 2.3 - Film Similarity: 0.43  
Film: Just My Luck - Voto: 2.13 - Film Similarity: 0.39  
Film: The Night Listener - Voto: 2.03 - Film Similarity: 0.38
```

```
Gene Seymour Similarity: 0.21  
Film: Lady in the Water - Voto: 2.37 - Film Similarity: 0.49  
Film: Just My Luck - Voto: 2.28 - Film Similarity: 0.48  
Film: The Night Listener - Voto: 2.62 - Film Similarity: 0.55
```

```
Michael Phillips Similarity: 0.22  
Film: Lady in the Water - Voto: 1.36 - Film Similarity: 0.31  
Film: The Night Listener - Voto: 0.4 - Film Similarity: 0.09
```

```
Claudia Puig Similarity: 0.16  
Film: Just My Luck - Voto: 2.28 - Film Similarity: 0.38  
Film: The Night Listener - Voto: 2.16 - Film Similarity: 0.36
```

```
Mick LaSalle Similarity: 0.15  
Film: Lady in the Water - Voto: -0.21 - Film Similarity: -0.03  
Film: Just My Luck - Voto: 1.22 - Film Similarity: 0.19  
Film: The Night Listener - Voto: 3.26 - Film Similarity: 0.5
```

```
Jack Matthews Similarity: 0.29  
Film: Lady in the Water - Voto: -0.49 - Film Similarity: -0.14  
Film: The Night Listener - Voto: 3.36 - Film Similarity: 0.97
```

```
Film consigliati a Toby: ['The Night Listener', 'Just My Luck', 'Lady in the Water']
```

I risultati ottenuti applicando tale funzione di approssimazione sono mostrati a lato.

Possiamo notare che l'utente Micheal Philips è stato considerato perché abbiamo stimato le sue valutazioni per film che non ha visto.

In questo caso l'utente più a Toby è Jack Matthews.

Anche in questo caso la lista di film consigliati a Toby è la medesima.

## Esercizio n°2

### Punto a

```
coseno dell'angolo tra Luke e Ryne: 0.99  
coseno dell'angolo tra Luke e Erin: 0.78  
coseno dell'angolo tra Luke e Nathan: 0.23  
coseno dell'angolo tra Luke e Pete: 0.96
```

```
Utenti più simile a Luke: [('Ryne', 0.99), ('Pete', 0.96), ('Nathan', 0.23), ('Erin', 0.78)]
```

Gli utenti che risultano più simili a Luke sono Ryne e Pete.

### Punto b

Per  $k = 3$

```
coseno dell'angolo tra Luke e Ryne: 0.92  
coseno dell'angolo tra Luke e Erin: -0.49  
coseno dell'angolo tra Luke e Nathan: -0.18  
coseno dell'angolo tra Luke e Pete: 0.35
```

```
Utenti più simile a Luke: [('Ryne', 0.92), ('Pete', 0.35), ('Nathan', -0.18), ('Erin', -0.49)]
```

Gli utenti più simili a Luke sono Ryne e Pete.

Per  $k = 4$

```
coseno dell'angolo tra Luke e Ryne: 0.64  
coseno dell'angolo tra Luke e Erin: -0.54  
coseno dell'angolo tra Luke e Nathan: -0.27  
coseno dell'angolo tra Luke e Pete: 0.48
```

```
Utenti più simile a Luke: [('Ryne', 0.64), ('Pete', 0.48), ('Nathan', -0.27), ('Erin', -0.54)]
```

Gli utenti più simili a Luke sono Ryne e Pete

## Esercizio n°3

La porzione di energia conservata con  $k=5, 20, 100$  e riportata direttamente sulle immagini generate:

$r = 5$  - energy = 0.4



$r = 20$  - energy = 0.53



$r = 100$  - energy = 0.71



### **Punto Facoltativo**

Il valore minimo di  $k$  per cui si conserva l'80% dell'energia totale è 193. Tale valore è stato determinato all'interno del programma.

$r = 193 - \text{energy} = 0.8$

