1. Obtener Usuarios similares a usuario activo : Procedimiento de NOE4J
2. Obtener las películas vistas por el usuario Activo
3. Obtener el perfil del usuario Activo
4. Obtener el perfil de los usuarios mas cercanos al usuario activo
5. Encontrar los usuarios con perfiles mas cercanos al usuario activo usando distancia Euclidiana
6. Obtener las películas vistas por los usuarios con perfiles mas cercanos al usuario activo.
7. Recomendar Top N de las películas no vistas por el usuario activo de la lista (6)

DETERMINAR PERFIL USUARIO

- Se utiliza spread activiation:
- Se utiliza el Genero Película para determinar perfil

PERFIL USUARIO ACTIVO

1. Obtener todas las películas vistas y ratings dados por el usuario activo
2. Obtener todos los géneros de las películas evaluado por el usuario activo
3. Normalizar el rating dado a cada película dividendo por el máx. ranting 5 (rta)
4. Obtener el peso (w) de cada genero en función del grado de cada genero :
    1. Se obtiene cuantos enlaces hay por cada genero de todas las película evaluadas por el usuario activo.
    2. Se suma los rta por cada enlace que salga de una película al genero.
       y se obtiene una ponderación por GENERO
        1. {ACION:9.8, DRAMA:3.77….}
    3. Se normaliza la ponderación por genero dividendo por el total de Géneros de las películas vista por el usuario activo: para obtener una medida de distribución central.
5. Finalmente so obtiene un vector que es el perfil del usuario por genero:

pefil_user_act = {ACION:0.45, DRAMA:0,034….}→PERFIL DEL USUARIO

PERFIL OTROS USUARIOS

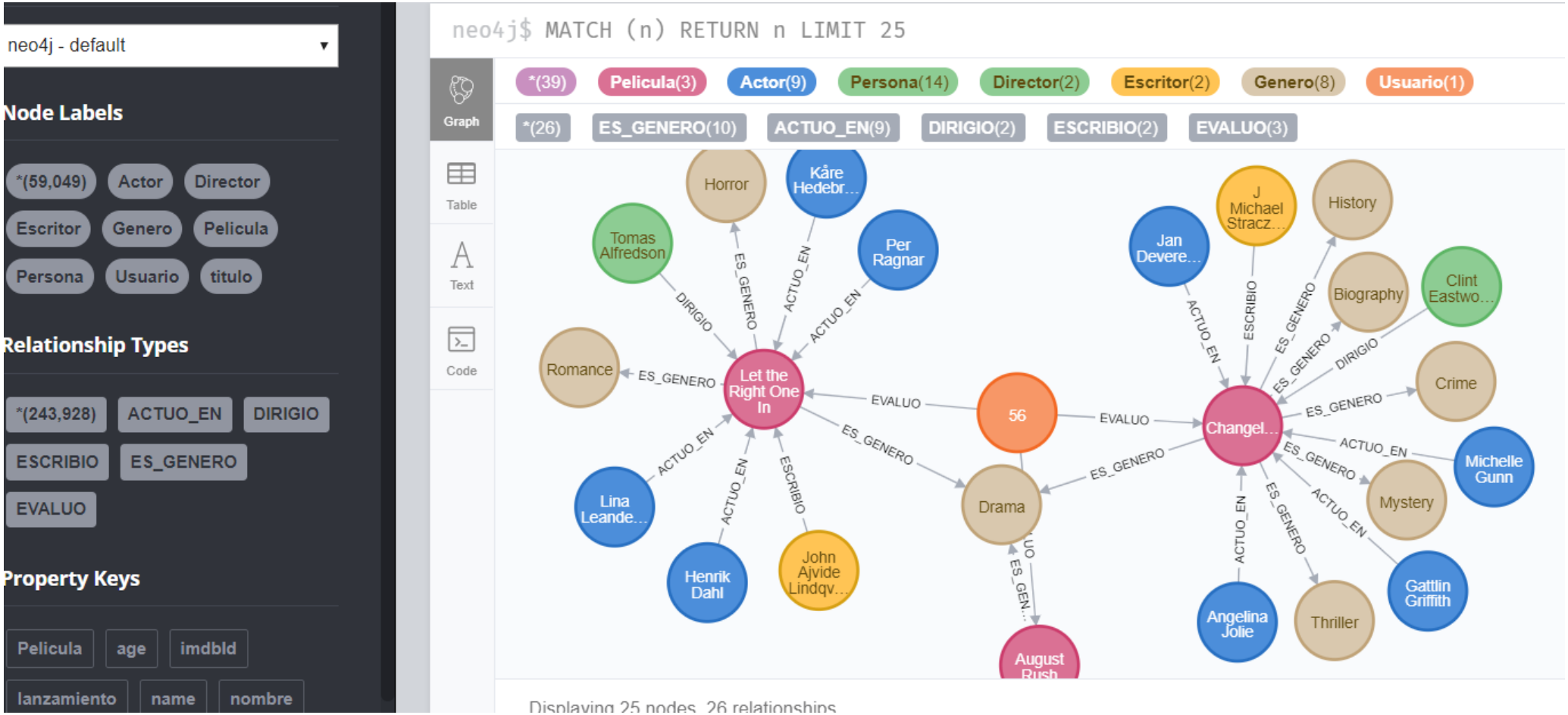
6. Se obtiene el perfil de los n usuarios mas cercanos al usuario activo.
   perfiles_user_sim =[perfil_user1, perfil_user2…….]

PROCESO DE RECOMENDACIÓN

1. Se obtiene distancia euclidiana entre el perfil de usuario activo y los perfiles de los n usuarios mas cercanos

    1. Por cada “GENERO” del usuario activo en su perfil (“pefil_user_act “) , se busca el genero en el perfil del usuario similar, en la lista “perfiles_user_sim”
    2. Se obtiene el valor del correspondiente “GENERO” en el perfil del usuario similar
    3. Si no se encuentra este “GENERO” en el usuario similar , este valor es de cero (0).
    4. Se obtienen dos vectores el del usuario activo y de el usuario similar por “GENERO”
    5. Se calcula distancia Euclidiana entre estos dos vectores.
    6. El proceso se repite por cada n usuario similar en “perfiles_user_sim”
2. Se Obtienen los k usuario mas cercanos (distancia Euclidiana cerca a cero), del proceso anterior.
3. Por cada k usuario mas cercano se obtiene las películas vistas
4. Se filtra las películas del paso anterior con las vistas por el usuario activo
5. Se hace recomendación las n top películas del paso anterior, no vistas por el usuario activo.

Nota:
El vector perfil del usuario se puede aumentar con otras características del enriquecimiento ontológico con ACTORES, DIRECTORES, ESCRITORES. Y el proceso básicamente seria el mismo


neo4j - default
Node Labels
*(59,049)
Actor
Director
Escritor
Genero
Pelicula
Persona
Usuario
titulo
Relationship Types
*(243,928)
ACTUO_EN
DIRIGIO
ESCRIBIO
ES_GENERO
EVALUO
Property Keys
Pelicula
age
imdbId
lanzamiento
name
nombre
neo4j$ MATCH (n) RETURN n LIMIT 25
Graph
Table
Text
Code
*(39)
Pelicula(3)
Actor(9)
Persona(14)
Director(2)
Escritor(2)
Genero(8)
Usuario(1)
*(26)
ES_GENERO(10)
ACTUO_EN(9)
DIRIGIO(2)
ESCRIBIO(2)
EVALUO(3)
Horror
Kåre Hedebr...
Tomas Alfredson
Per Ragnar
Romance
Let the Right One In
Lina Leande...
Henrik Dahl
John Ajvide Lindqv...
56
Drama
August Rush
Jan Devere...
J Michael Stracz...
History
Biography
Clint Eastwo...
Crime
Changel...
Michelle Gunn
Mystery
Gattlin Griffith
Angelina Jolie
Thriller
DIRIGIO
ES_GENERO
ACTUO_EN
ESCRIBIO
EVALUO
Displaying 25 nodes, 26 relationships

DISTANCIA EUCLIDIANA CON K CERCANOS

```
get_recomendacion_user(6,10,10,5)
```

```
[('7384', 0.0593503160770430 94),
 ('7345', 0.0942595701375552 6),
 ('5985', 0.0977459077704163 5),
 ('7657', 0.1052474186587717 7),
 ('4958', 0.1420167500475796)]
```

```
: get_recomendacion_user(6,10,100,5)
```

```
: [('7384', 0.0593503160770430 94),
   ('4637', 0.0698836037521386),
   ('5626', 0.0755366267124562),
   ('214', 0.0763228725253330 5),
   ('430', 0.0766778677179055 8)]
```

RECOMENDACIONES

```
 #u=user activo, n= recomendaciones, n= vecinos mas similares, k=mas cercanos
reco,existe_ =get_recomendacion_user(7384,10,10,10)
```

```
reco
```

```
['Desperado',
 'Mary Shelleys Frankenstein',
 'Little Women',
 'Drop Zone',
 'Don Juan DeMarco',
 'Waterworld',
 'Under Siege 2 Dark Territory',
 'Judge Dredd',
 'Hackers',
 'The Birdcage']
```

Links

Ratings

```
ratings.head()
```

	userId	movieId	rating	imdbId
0	14	1	4.5	0114709
112	14	87876	3.5	1216475
113	14	88405	3.0	1632708
114	14	89745	4.0	0848228
115	14	90647	3.5	0448694

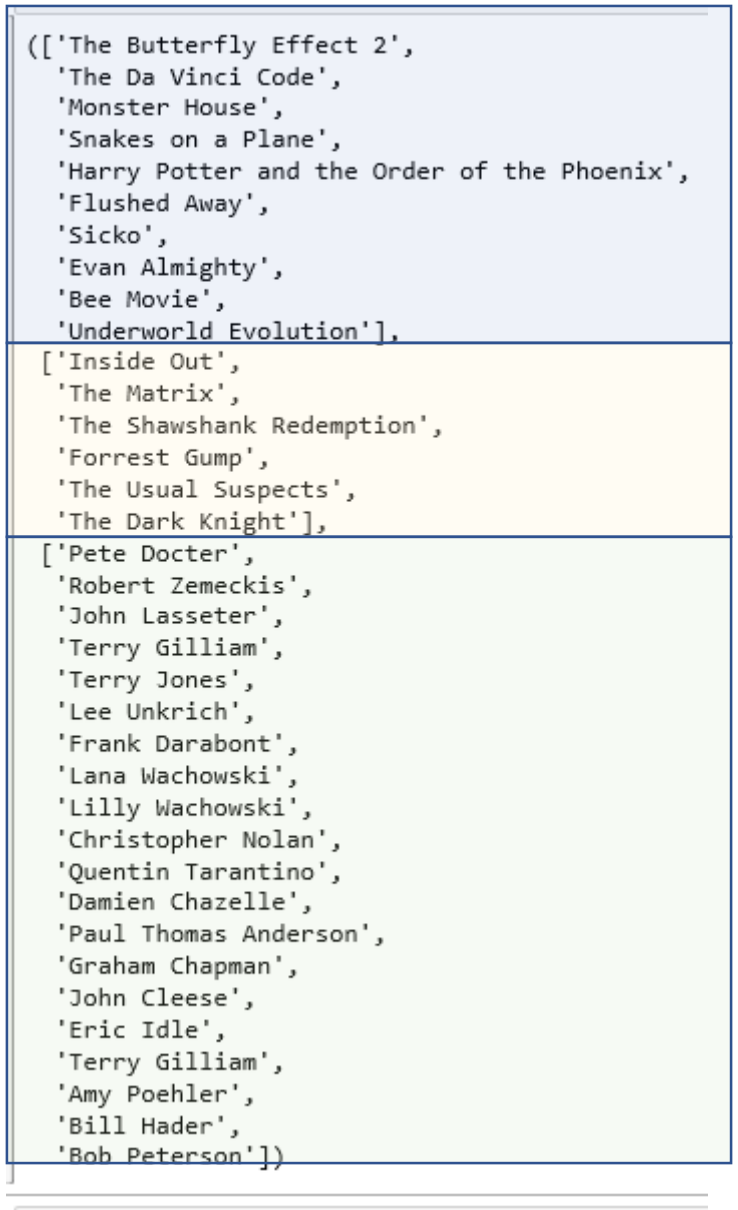
OMDb API

```
for item in resp:
    print(item," -- " ,resp[item])
```

```
Title  --  Forrest Gump
Year  --  1994
Rated  --  PG-13
Released  --  06 Jul 1994
Runtime  --  142 min
Genre  --  Drama, Romance
Director  --  Robert Zemeckis
Writer  --  Winston Groom (novel), Eric Roth (screenplay)
Actors  --  Tom Hanks, Rebecca Williams, Sally Field, Michael Conner Humphreys
Plot  --  The presidencies of Kennedy and Johnson, the events of Vietnam, Watergate and other historical events unfold throug
the perspective of an Alabama man with an IQ of 75, whose only desire is to be reunited with his childhood sweetheart.
Language  --  English
Country  --  USA
Awards  --  Won 6 Oscars. Another 44 wins & 74 nominations.
Poster  --  https://m.media-amazon.com/images/M/MV5BNWIwODRlZTUtY2U3ZS00YzgiLWJhNzYtMmZiYmEyNmU1NjMzXkEyXkFqcGdeQXVyMTQxMzMzN
@._V1_SX300.jpg
Ratings  --  [{'Source': 'Internet Movie Database', 'Value': '8.8/10'}, {'Source': 'Rotten Tomatoes', 'Value': '70%'}, {'Sour
e': 'Metacritic', 'Value': '82/100'}]
Metascore  --  82
imdbRating  --  8.8
imdbVotes  --  1,722,338
imdbID  --  tt0109830
Type  --  movie
DVD  --  N/A
BoxOffice  --  N/A
Production  --  N/A
Website  --  N/A
Response  --  True
```

Python Component

NEO4J DB



N Recomendaciones

Recomendaciones vistas por el Usuario Activo

Características que se repiten en el vector perfil