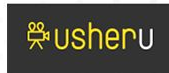
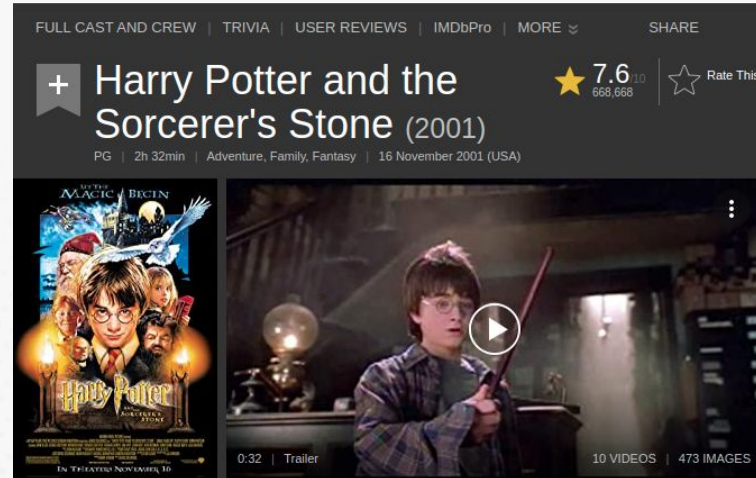
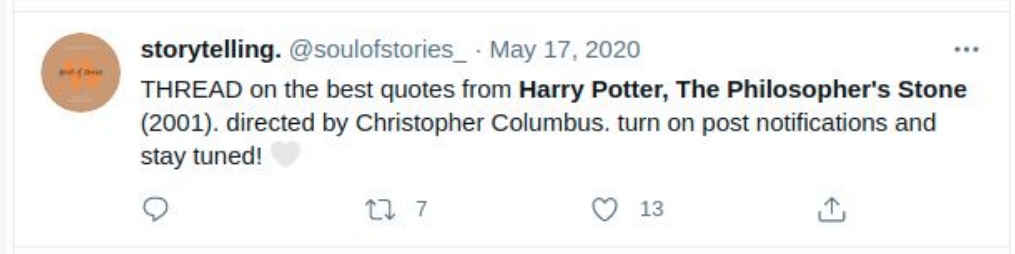


b33mo

b33mo



Problem



Problem

- Do the titles X and Y represent the same movie?
- Bunch of movie titles -> Group them depending on the movie they represent!

The data

LANGUAGE PROBLEM

Harry Potter & the Sorcerer's Stone
Harry Potter and the Sorcerer's Stone

Harry Potter One
Harry Potter 1
First Harry Potter Movie

Harry Potter & the Philosopher's Stone
Harry Potter and the Philosopher's Stone

The data

LANGUAGE PROBLEM

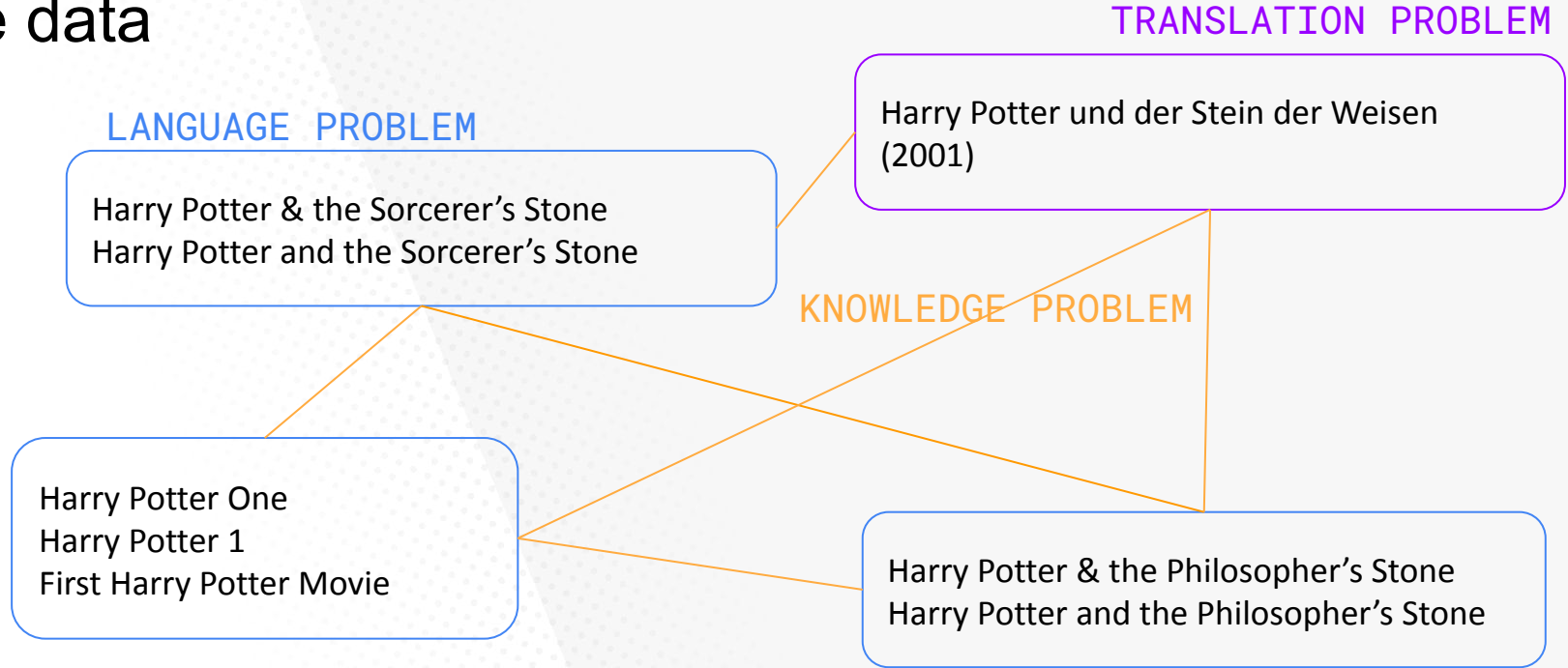
Harry Potter & the Sorcerer's Stone
Harry Potter and the Sorcerer's Stone

KNOWLEDGE PROBLEM

Harry Potter One
Harry Potter 1
First Harry Potter Movie

Harry Potter & the Philosopher's Stone
Harry Potter and the Philosopher's Stone

The data



Solution

LANGUAGE PROBLEM - NLP (tokenization/ lemmatization/ etc)

Harry Potter & the Sorcerer's Stone -> ['harry', 'potter', 'sorcerer', 'stone', '']

Harry Potter and the Sorcerer's Stone -> ['harry', 'potter', 'sorcerer', 'stone', '']

Harry Potter One -> ['harry', 'one', 'potter']

Harry Potter 1 -> ['harry', 'one', 'potter']

Solution

KNOWLEDGE PROBLEM - GOOGLE KNOWLEDGE GRAPH API



Harry Potter & the Philosopher's Stone

Harry Potter and the Philosopher's Stone
(113216.796875)

Harry Potter i la Pedra Filosofal
(440.1052856445312)

Harry Potter and the Chamber of Secrets
(201.3001251220703)

Harry Potter (107.7324752807617)

Harry Potter & The Philosopher's Stone
(100)

Harry Potter and the Philosopher's Stone
(79.19971466064453)

Harry Potter and the Philosopher's Stone:
Harry's Wondrous World (48)

Harry Potter und der Stein der Weisen (2001)

Harry Potter (92.34211730957031)

Harry Potter and the Philosopher's Stone
(63.07793045043945)

Harry Potter and the Philosopher's Stone
(8.806763648986816)

Harry Potter and the Chamber of Secrets
(3.844590425491333)

Harry Potter and the Philosopher's Stone
(2.563870906829834)

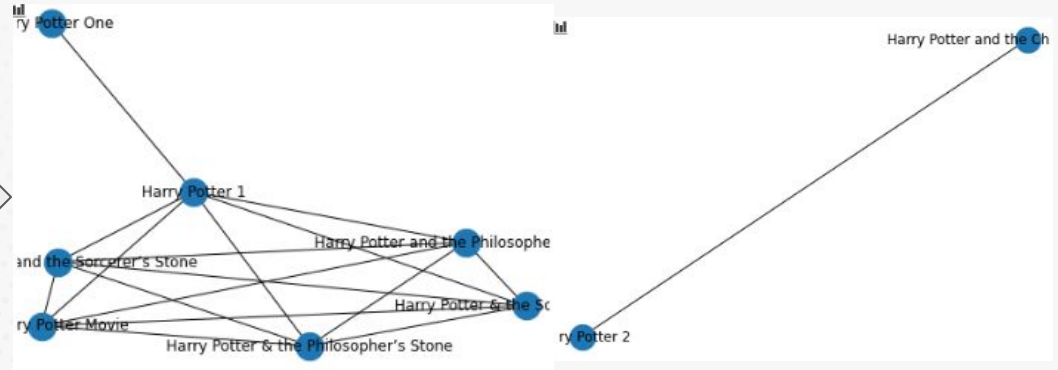
Harry Potter and the Philosopher's Stone
(2.234956979751587)

Harry Potter and the Goblet of Fire
(1.793914198875427)

Solution

Are X and Y the same movie? -> if so, connect them in the graph

```
hp = [  
    "Harry Potter & the Philosopher's Stone",  
    "Harry Potter and the Philosopher's Stone",  
    "Harry Potter & the Sorcerer's Stone",  
    "Harry Potter and the Sorcerer's Stone",  
    "Harry Potter One",  
    "Harry Potter 1",  
    "First Harry Potter Movie",  
    "Harry Potter and the Deathly Hallows – Part 1",  
    "Harry Potter and the Chamber of Secrets",  
    "Harry Potter 2"  
]
```



```
0 ['Harry Potter & the Philosopher's Stone', 'Harry Potter and the Philosopher's  
Stone', 'Harry Potter & the Sorcerer's Stone', 'Harry Potter and the Sorcerer's  
Stone', 'Harry Potter One', 'Harry Potter 1', 'First Harry Potter Movie']  
1 ['Harry Potter and the Deathly Hallows – Part 1']  
2 ['Harry Potter and the Chamber of Secrets', 'Harry Potter 2']
```

Results

```
1 Demon Slayer The Movie: Mugen Train
2 Demon Slayer - Kimetsu No Yaiba - The Movie: Mugen Train
3 DEMON SLAYER THE MOVIE: MUGEN TRAIN
4 Tom & Jerry
5 Tom y Jerry
6 Tom & Jerry The Movie
7 Maya the Bee 3: The Golden Orb
8 MAIJA MEHILÄINEN JA KULTAINEN AARRE
9 Peter Rabbit 2: The Runaway
10 Peter Rabbit 2: A la fuga
11 Pelle Kanin 2: På rymmen
12 Peter Rabbit: Coelho à Solta
13 Peter Hase 2 - Ein Hase macht sich vom Acker
14 Minions: El origen de Gru
15 KÄTYRIT: GRUN TARINA
16 Minions: The Rise of Gru
17 BLACK WIDOW
18 Black Widow
19 Viuda Negra
20 Marvel Black Widow
21 FAST & FURIOUS 9
22 Fast and Furious 9
23 Fast & Furious 9
24 F9 - Fast & Furious 9
25 Fast & Furious Presents: Hobbs & Shaw
26 Space Jam 2
27 Space Jam: A New Legacy
28 Doctor Strange: In the Multiverse of Madness
29 Doctor Strange in the Multiverse of Madness
30 The Matrix 4
31 The Matrix Resurrections
32 No Time To Die
33 Sin tiempo para morir
34 007 No Time to Die
35 The Boss Baby: Family Business
36 Baby Boss 2 : une affaire de famille
37 Ainbo: Hrdinka pralesa
38 Ainbo
39 Charlie's Angels (2020)
40 Charlie's Angels
41 Godzilla: King of the Monsters
42 Godzilla II - King of the Monsters
43 POKÉMON Detective Pikachu
44 Pokémon: Detective Pikachu
45 Roger Waters - Concert
46 Roger Waters: Us + Them
47 Roger Waters: Us + Them (Digital 5.1)
48 The Lion King
49 The Lion King (2019)
50 Zombieland 2
51 Zombieland: Double Tap
```



0 ['Demon Slayer - Kimetsu No Yaiba - The Movie: Mugen Train', 'DEMON SLAYER THE MOVIE: MUGEN TRAIN']

1 ['Tom & Jerry', 'Tom y Jerry', 'Tom & Jerry The Movie']

2 ['Maya the Bee 3: The Golden Orb']

3 ['MAIJA MEHILÄINEN JA KULTAINEN AARRE']

4 ['Peter Rabbit 2: The Runaway', 'Peter Rabbit 2: A la fuga', 'Pelle Kanin 2: På rymmen', 'Peter Rabbit: Coelho à Solta', 'Peter Hase 2 - Ein Hase macht sich vom Acker']

5 ['Minions: El origen de Gru', 'KÄTYRIT: GRUN TARINA', 'Minions: The Rise of Gru']

6 ['BLACK WIDOW', 'Black Widow', 'Viuda Negra', 'Marvel Black Widow']

7 ['FAST & FURIOUS 9', 'Fast and Furious 9', 'Fast & Furious 9', 'F9 - Fast & Furious 9']

8 ['Fast & Furious Presents: Hobbs & Shaw']

9 ['Space Jam 2', 'Space Jam: A New Legacy']

10 ['Doctor Strange: In the Multiverse of Madness', 'Doctor Strange in the Multiverse of Madness']

11 ['The Matrix 4']

12 ['The Matrix Resurrections']

13 ['No Time To Die', 'Sin tiempo para morir', '007 No Time to Die']

14 ['The Boss Baby: Family Business']

15 ['Baby Boss 2 : une affaire de famille']

16 ['Ainbo: Hrdinka pralesa', 'Ainbo']

17 ['Charlie's Angels (2020)', 'Charlie's Angels']

18 ['Godzilla: King of the Monsters', 'Godzilla II - King of the Monsters']

19 ['POKÉMON Detective Pikachu', 'Pokémon: Detective Pikachu']

20 ['Roger Waters - Concert', 'Roger Waters: Us + Them']

21 ['Roger Waters: Us + Them (Digital 5.1)']

22 ['The Lion King', 'The Lion King (2019)']

23 ['Zombieland 2', 'Zombieland: Double Tap']

Results

96% correct association - out of 51 movie titles

Peter Rabbit 2: The Runaway
Peter Rabbit 2: A la fuga
Pelle Kanin 2: På rymmen
Peter Rabbit: Coelho à Solta
Peter Hase 2 - Ein Hase macht
sich vom Acker

Minions: El origen de Gru
KÄTYRIT: GRUN TARINA
Minions: The Rise of Gru

Limitations and perspectives for the future

- A bit dependent on the Google Knowledge API:
2 ['Maya the Bee 3: The Golden Orb']
3 ['MAIJA MEHILÄINEN JA KULTAINEN AARRE']
- Funny example:
7 ['FAST & FURIOUS 9', 'Fast and Furious 9', 'Fast & Furious 9', 'F9 - Fast & Furious 9']
8 ['Fast & Furious Presents: Hobbs & Shaw']
- Future: What countries have the most different titles?
The Nightmare Before Christmas (EN) - Jack's Odd World (PT)
101 Dalmatians (EN) - The Night of The Cold Noses (ES)

b33mo

- Joana Ferreira
- Cristina Meng

