HW7 CS432

Ondra Torkilson

November 30, 2020

Q1

To find the users closest to me in age, gender, and occupation, I searched the user file using 'ctrl+F' for the string '26|F|student' and found users 241 and 939. Then I searched for '25|F|student' and found user 203. See those results pictured below in Figure 1, 2, and 3.

203|25|F|student|32301

Figure 1: User 203 From u.user.txt

241 26 F student 20001

Figure 2: User 241 From u.user.txt

939|<mark>26|F|student|</mark>33319

Figure 3: User 939 From u.user.txt

Below are the results of their favorite and least favorite movies. First, you see the table with the top and bottom three for each user. Then, you can see all the ratings for each user . I had to examine further than the top and bottom three for each user, because a lot of the movies I had not seen. That will be a running theme throughout this project. Generation gaps are tough. I chose user 939 pretty much at random out of the three users.

```
C:\Users\Ondra Torkilson\Documents\CS432\HW7>python3 recommender.py

203 Favorites Emma, Swingers, Scream

203 Least Favorites A Time to Kill, Fargo, Liar, Liar

241 Favorites The English Patient, Scream, Soul Food

241 Least Favorites Seven Years in Tibet, Fallen, Alien: Resurrection

939 Favorites Courage Under Fire, Jerry Maguire, People vs. Larry Flynt

939 Least Favorites The Island of Dr. Moreau, Mortal Kombat: Annihilation, Kull the Conqueror
```

Figure 4: Top and Bottom Three For Each User

```
C:\Users\Ondra Torkilson\Documents\CS432\HW7>python3 recom
This is user 203
[('Time to Kill, A (1996)', 1.0),
    ('Fargo (1996)', 1.0),
    ('Liar Liar (1997)', 2.0),
    ('House Arrest (1996)', 2.0),
    ('Mortal Kombat: Annihilation (1997)', 2.0),
    ('Saint, The (1997)', 2.0),
    ('Mother (1996)', 3.0),
    ('Mother (1996)', 3.0),
    ('Ghost and the Darkness, The (1996)', 3.0),
    ('Playing God (1997)', 3.0),
    ('Starship Troopers (1997)', 3.0),
    ('Toy Story (1995)', 3.0),
    ('Extreme Measures (1996)', 3.0),
    ('Twelve Monkeys (1995)', 3.0),
    ('Fly Away Home (1996)', 3.0),
    ('Hercules (1997)', 3.0),
    ('Men in Black (1997)', 3.0),
    ('Contact (1997)', 3.0),
    ('Dante's Peak (1997)', 3.0),
    ('Nixon (1995)', 3.0),
    ('Nixon (1995)', 3.0),
    ('Gortact (1997)', 3.0),
    ('Rock, The (1996)', 4.0),
    ('Rock, The (1996)', 4.0),
    ('Rock, The (1996)', 4.0),
    ('Rill Wonka and the Chocolate Factory (1971)', 4.0),
    ('Fifth Element, The (1997)', 4.0),
    ('Star Trek: First Contact (1996)', 4.0),
    ('Peacemaker, The (1997)', 4.0),
    ('Peacemaker, The (1997)', 4.0),
    ('Peacemaker, The (1997)', 4.0),
    ('Star Trek: First Contact (1996)', 4.0),
    ('Star Wars (1996)', 4.0),
    ('Star Wars (1996)', 4.0),
    ('Star Wars (1996)', 5.0),
    ('Star Wars (1996)', 5.0),
    ('Star Wars (1996)', 5.0),
    ('Star Wars (1996)', 5.0),
    ('Suingers (1996)', 5.0),
    ('Suingers (1996)', 5.0),
    ('Suingers (1996)', 5.0),
    ('Emma (1996)', 5.0)]
```

Figure 5: User 203 Preferences

```
This is user 241

[('Seven Years in Tibet (1997)', 2.0),
    ('Fallen (1998)', 2.0),
    ('Alien: Resurrection (1997)', 2.0),
    ('I Know What You Did Last Summer (1997)', 2.0),
    ('Scream 2 (1997)', 2.0),
    ('Liar Liar (1997)', 3.0),
    ('Kiss the Girls (1997)', 3.0),
    ('La. Confidential (1997)', 3.0),
    ('Jackie Brown (1997)', 3.0),
    ('Jackie Brown (1997)', 3.0),
    ('Jackie Brown (1997)', 3.0),
    ('Gattaca (1997)', 3.0),
    ('Gattaca (1997)', 3.0),
    ('Rosewood (1997)', 4.0),
    ('Titanic (1997)', 4.0),
    ('Eve's Bayou (1997)', 4.0),
    ('Rainmaker, The (1997)', 4.0),
    ('Air Force One (1997)', 4.0),
    ('Air Force One (1997)', 4.0),
    ('Amistad (1997)', 5.0),
    ('Scream (1996)', 5.0),
    ('Scream (1996)', 5.0),
    ('English Patient, The (1996)', 5.0)]
    This is user 939

[('Island of Dr. Moreau, The (1996)', 2.0).
              his is user 241
(Stream (1996), 5.0),
('English Patient, The (1996)', 5.0)]
This is user 939
[('Island of Dr. Moreau, The (1996)', 2.0),
('Mortal Kombat: Annihilation (1997)', 2.0),
('Bull the Conqueror (1997)', 2.0),
('Diabolique (1996)', 3.0),
('Girl 6 (1996)', 3.0),
('Greacher's Wife, The (1996)", 3.0),
('Children of the Corn: The Gathering (1996)', 3.0),
('Lost World: Jurassic Park, The (1997)', 3.0),
('Contact (1997)', 4.0),
('Mr. Wrong (1996)', 4.0),
('Jack (1996)', 4.0),
('Fathers' Day (1997)", 4.0),
('Eraser (1996)', 4.0),
('Nutty Professor, The (1996)', 4.0),
('Glimmer Man, The (1996)', 4.0),
('Sense and Sensibility (1995)', 4.0),
('Sense and Sensibility (1995)', 4.0),
('Jackal, The (1997)', 5.0),
('First Wives Club, The (1996)', 5.0),
('That Old Feeling (1997)', 5.0),
('Star Trek: First Contact (1996)', 5.0),
('G.I. Jane (1997)', 5.0),
('Associate, The (1996)', 5.0),
('Associate, The (1996)', 5.0),
('Mirror Has Two Faces, The (1996)', 5.0),
('Godfather, The (1972)', 5.0),
('Thister (1996)', 5.0),
     This is user 939
```

Figure 6: User 241 and 939 Preferences

$\mathbf{Q2}$

For questions 2, 3, and 4, I directly edited the 'recommender.py' script. To find the correlations of the other users to substitute me for Q2, I used the built-in function 'sim_pearson' within a for loop to calculate the correlation of each user in the preferences dictionary to user 939. Then, I sorted the list of correlations and examined the top and bottom of the list.

```
#q2
correlationsToMe = []
for i in prefs:
    correlation = sim_pearson(prefs, '939', i)
    userAndCorrelation = {'i':i, 'correlation':correlation}
    correlationsToMe.append(userAndCorrelation)

sorted_correlations = sorted(correlationsToMe, key = lambda item: item['correlation'])
pprint.pprint(sorted_correlations)
#q2
```

```
{'correlation': 1.0, 'i': '801'},
    {'correlation': 1.0, 'i': '826'},
    {'correlation': 1.0, 'i': '816'},
    {'correlation': 1.0, 'i': '846'},
    {'correlation': 1.0, 'i': '873'},
    {'correlation': 1.0, 'i': '939'},
    {'correlation': 1.000000000000004, 'i': '603'}]
```

Figure 7: Positively Correlated Users to 939

```
[{'correlation': -1.0, 'i': '154'}, {'correlation': -1.0, 'i': '111'}, {'correlation': -1.0, 'i': '281'}, {'correlation': -1.0, 'i': '118'}, {'correlation': -1.0, 'i': '171'}, {'correlation': -1.0, 'i': '153'},
```

Figure 8: Negatively Correlated Users to 939

$\mathbf{Q3}$

For question 3, I used the built-in function get Recommendations on the user 939. Then, I was able to see the top and bottom recommended movies from the returned list which I called 'recommendations.' You can see these pictured below, and I included the top and bottom 10.

```
#q3
prefs = loadMovieLens()
recommendations = getRecommendations(prefs, '939', sim_pearson)
pprint.pprint(recommendations)
#q3
```

```
\Users\Ondra Torkilson\Documents\CS432\HW7>python3 recommender.py
[5.0, 'Two or Three Things I Know About Her (1966)'),
[5.0, 'They Made Me a Criminal (1939)'),
[5.0, 'Santa with Muscles (1996)'),
[5.0, 'Saint of Fort Washington, The (1993)'),
[5.0, 'Prefontaine (1997)'),
[5.0, 'Marlene Dietrich: Shadow and Light (1996) '),
[5.0, 'Marlene Dietrich: Shadow and Light (1996) '),
[5.0, 'Horse Whisperer, The (1998)'),
[5.0, 'Great Day in Harlem, A (1994)'),
[5.0, 'Aparajito (1956)'),
[5.0, 'Aiqing wansui (1994)'),
```

Figure 9: Top Ten Recommendations For Substitute Me, User 939

```
Venice/Venice (1992)),
'Turbo: A Power Rangers Movie (1997)'),
                    (1.0, 'Turbo: A Power Rangers Movie (1997)'),
(1.0, 'Tough and Deadly (1995)'),
(1.0, 'Time Tracers (1995)'),
(1.0, 'Time Tracers (1995)'),
(1.0, 'The Deadly Cure (1996)'),
(1.0, 'Stefano Quantestorie (1993)'),
(1.0, 'Somebody to Love (1994)'),
(1.0, 'Shadow of Angels (Schatten der Engel) (1976)'),
(1.0, 'Relative Fear (1994)'),
(1.0, 'Power 98 (1995)'),
(1.0, 'Power 98 (1995)'),
(1.0, 'Open Season (1996)'),
(1.0, 'New York Cop (1996)'),
(1.0, 'New Age, The (1994)'),
(1.0, 'New Age, The (1994)'),
(1.0, 'New Age, The (1994)'),
(1.0, 'Dem Season (1996)'),
(1.0, 'New York Cop (1996)'),
(1.0, 'New Age, The (1994)'),
(1.0, 'Mr. Mago (1997)'),
(1.0, 'Mr. Magoo (1997)'),
(1.0, 'Mr. Magoo (1997)'),
(1.0, 'Mille bolle blu (1993)'),
(1.0, 'Mille bolle blu (1993)'),
(1.0, 'Low Life, The (1994)'),
(1.0, 'Loch Ness (1995)'),
(1.0, 'Loch Ness (1995)'),
(1.0, 'Loch Ness (1995)'),
(1.0, 'Loch Ness (1995)'),
(1.0, 'Hugo Pool (1997)'),
(1.0, 'Hugo Pool (1997)'),
(1.0, 'Homage (1995)'),
(1.0, 'Good Morning (1971)'),
(1.0, 'Good Morning (1971)'),
(1.0, 'Gotdigan's Island: The Movie (1998)"),
(1.0, 'Getting Away With Murder (1996)'),
(1.0, 'Falling in Love Again (1980)'),
(1.0, 'Every Other Weekend (1990)'),
(1.0, 'Crude Oasis, The (1995)'),
(1.0, 'Crude Oasis, The (1995)'),
(1.0, 'Butterfly Kiss (1995)'),
(1.0, 'Boys in Venice (1996)'),
(1.0, 'Bloody Child, The (1996)'),
(1.0, 'Brid of Prey (1996)'),
(1.0, 'Brid of Prey (1996)'),
(1.0, 'Brid of Prey (1996)'),
(1.0, 'Ann Rand: A Sense of Life (1997)'),
(1.0, 'Angust (1996)'),
(1.0, 'Amityville: Dollhouse (1996)'),
(1.0, 'Amityville: A New Generation (1993)'),
(1.0, 'Amityville Curse, The (1990)'),
(1.0, 'Amityville Supposition of the Mega Mountain (1998)')]
```

Figure 10: Bottom Ten Recommendations For Substitute Me, User 939

$\mathbf{Q4}$

It was slightly difficult to choose a least favorite and favorite film from a list of 1682 films, most of which I had never seen before. Luckily, I found Groundhog Day, which is definitely one of my favorites. As for the least favorite, I chose one of the lowest rated films from substitute me, 'Mortal Kombat: Annihilation'. Figures 11, 12, 13, and 14 show the highest and lowest correlated movies to 'Groundhog Day' and the highest and lowest correlated movies to 'Mortal Kombat: Annihilation'.

Based on the correlations, I agree with some of them and not others. For example, none of the movies correlated with Groundhog Day seemed interesting to me. I had to watch trailers, because I had not seen any of them.

Infinity did not seem like a movie I would like, even though I really enjoy Patricia Arquette.

Bliss was a German movie, subtitled in English about some woman in her midlife crisis cheating on her husband with an author. Not my style.

Some of the other top correlations with Groundhog Day were foreign films with no translated trailers, such as Nelly & Monsuier Armaud, Le Colonel Chabert, and La Flor De Mi Secreto.

```
'Colonel Chabert, Le (1994)'}
                                        'Perfect Candidate, A (1996)'},
correlation'
                                       'Outlaw, The (1943)'},
correlation
                            Secret, The (Flor de M1 sec.),
, 'i': 'Murder, My Sweet (1944)'},
, 'i': 'Nelly & Monsieur Arnaud (1995)'},
, 'i': 'Nelly & Book: Mowgli & Balo
      'Flower of My
                                        'Second Jungle Book: Mowgli & Baloo, The (1997)'}
                                        'Pompatus of Love, The (1996)'},
'Gabbeh (1996)'},
'Wonderland (1997)'},
correlation
correlation
correlation
                                        'Infinity (1996)'},
'Bliss (1997)'},
'Wild Reeds (1994)'
correlation
correlation
                                        "Ed's Next Move (1996)"},
'Across the Sea of Time (1995)'},
'Two or Three Things I Know About Her (1966)'},
correlation
correlation
correlation
                                       'Boys Life (1995)'},
'Savage Nights (Nuits fauves, Les) (1992)'},
'Paradise Road (1997)'},
 correlation
                                        'Mr. Jones (1993)'},
'Wedding Gift, The (1994)'},
'New Jersey Drive (1995)'},
correlation
correlation
                                        'Safe Passage (1994)'},
'Romper Stomper (1992)'},
 orrelation
correlation
                                        Guantanamera (1994)
                                        'Lamerica (1994)'}]
correlation':
```

Figure 11: Positive Correlation: Groundhog Day

Some of the films negatively associated with 'Groundhog Day', I think I would have liked such as 'I Like It Like That', 'The Scarlet Letter,' and 'The Joy Luck Club.' The first was a comedic drama based on the life of a strong-willed mother in the Bronx. The second two are based off of books that I read and enjoyed, although this does not mean they would be good movies.

I do agree that I probably would not like 'Bad Moon.' though, which seemed to be a werewolf thriller.

```
00000007, 'i': 'Dingo (1992)'},
00000007, 'i': 'I Like It Like That (1994)'},
"C'est arrivé près de chez vous (1992)"},
                                 -1.000000000000000007,
correlation'
 correlation'
 correlation':
                                                          "C'est arrivé près de chez vous (1992)"},
'Scarlet Letter, The (1995)'},
'Bad Moon (1996)'},
'Talking About Sex (1994)'},
'Search for One-eye Jimmy, The (1996)'},
'Broken English (1996)'},
'Neon Bible, The (1995)'},
'The Innocent (1994)'},
'Love and Death on Long Island (1997)'},
'Nil By Mouth (1997)'},
'Vermin (1998)'},
'Caro Diario (Dear Diary) (1994)'},
'Palmetto (1998)'}.
correlation':
                                  -1.0,
correlation':
                                  -1.0,
correlation'
correlation'
                                  -1.0,
correlation'
correlation':
correlation':
correlation'
correlation'
correlation'
correlation'
                                                           'Cano Diario (Dear Diary) (1994)')
'Palmetto (1998)'},
'Pushing Hands (1992)'},
'Joy Luck Club, The (1993)'},
'Four Days in September (1997)'},
"Enfer, L' (1994)"},
'Collectionneuse, La (1967)'},
'Schizopolis (1996)'},
'Show, The (1995)'},
'Dream Man (1995)'},
'Aparajito (1956)'},
'American Dream (1990)'}.
correlation'
correlation
correlation'
                                  -1.0,
correlation'
correlation'
correlation':
correlation'
correlation'
correlation
correlation
                                                            'Aparajito (1956) },
'American Dream (1990)'},
'Clean Slate (Coup de Torchon) (1981)'},
'Glass Shield, The (1994)'},
'Condition Red (1995)'},
 correlation
correlation
correlation
correlation
                                                              'It Takes Two (1995)'},
'Slingshot, The (1993)'
correlation
 correlation
                                                              Angel Baby
```

Figure 12: Negative Correlation: Groundhog Day

The films positively and negatively correlated with my least favorite movie seemed more accurate.

Positive associations with 'Mortal Kombat: Annihilation' all seemed pretty bad My Own Private Idaho had a really bizzare trailer, and I do not think I would have liked that movie. Same goes for Trigger Effect, Jeffrey, and The Last Supper.

Films negatively associated with 'Mortal Kombat: Annihilation' such as 'Little Women,' 'Les Miserables,' and 'Anne Frank Remembered' all seemed interesting, as I love period pieces and Holocaust remembrance pieces. Also, 'The Truman Show' was good.

```
'Stuart Saves His Family (1995)'},
'My Own Private Idaho (1991)'},
'Young Guns II (1990)'},
'Now and Then (1995)'},
'Last Supper, The (1995)'},
'Jeffrey (1995)'},
'Mortal Kombat: Annihilation (1997)'},
'Rome Is Blooding (1993)'
correlation'
correlation':
                            1.0,
correlation':
                            1.0,
correlation':
                            1.0,
correlation'
                             1.0.
correlation'
                                                  'Romeo Is Bleeding (1993)'},
correlation'
                             1.0,
                                                 Komeo Is Breeding (1993) },
'Judgment Night (1993)'},
'Trigger Effect, The (1996)'},
'Mute Witness (1994)'},
'Village of the Damned (1995)'},
'National Lampoon's Senior Trip (1995)"},
'Charade (1963)'},
'Paris Is Burning (1990)'},
correlation'
correlation'
                             1.0,
                            1.0,
correlation'
correlation'
                             1.0,
                             1.0,
correlation'
                             1.0,
correlation'
                            1.0,
correlation': 1.0,
i': 'Until the End of
                                            the World (Bis ans Ende der Welt) (1991)'},
': 'Beverly Hillbillies, The (1993)'},
': 'Children of the Corn: The Gathering (1996)'},
': 'Jury Duty (1995)'},
': 'Jane Eyre (1996)'},
': 'Bloodsport 2 (1995)'},
': 'Wonderland (1997)'},
': 'Addiction, The (1995)'},
': 'Sunset Park (1996)'},
': 'Mary Reilly (1996)'},
': 'Poetic Justice (1993)'},
correlation': 1.0,
correlation':
correlation':
                            1.0,
correlation'
                            1.0,
correlation'
                            1.0,
correlation'
                                .ǿ,
 correlation'
correlation'
correlation'
                             1.0,
                                                 Mary Reilly (1996)',

Poetic Justice (1993)',

'Night Flier (1997)',

'Double Happiness (1994)',

'That Old Feeling (1997)',

'April Fool's Day (1986)",

'Time Tracers (1995)',

'Frisk (1995)',

'Tom and Huck (1995)',

'Head Above Water (1996)',

'North (1994)';
correlation'
                             1.0,
correlation'
                             1.0,
correlation'
                             1.0,
correlation'
correlation'
                             1.0,
                            1.0,
correlation'
correlation'
                             1.0,
correlation'
                                .0,
                                                  'North (1994)'},
'Bushwhacked (1995)'},
'Killer (Bulletproof Heart) (1994)'},
correlation'
                             1.0,
correlation'
                             1.0,
                                                  'Killer (Burre.,
'Stalker (1979)'},
'Graph' 'i': 'Angus (1995)'},
correlation'
                            1.0,
correlation'
                            1.0, 'i': 'Stalker (
1.00000000000000007,
                            1.00000000000000001,
        correlation': 1.00000000000000022,
```

Figure 13: Positive Correlation: Mortal Kombat: Annihilation

```
i': 'Made in America (1993)'},

i': 'Bogus (1996)'},

i': 'M. Butterfly (1993)'},

i': 'Vampire in Brooklyn (1995)'},

i': 'Cemetery Man (Dellamorte Dellamore) (1994)'},

i': 'Specialist, The (1994)'},

i': 'Two if by Sea (1996)',

i': 'Milk Money (1994)'},

i': 'Jade (1995)'},

i': 'Bad Company (1995)'},

i': 'Bride of Frankenstein (1935)'},

i': 'Old Man and the Sea, The (1958)'},
   correlation': -1.0,
  correlation: -1.0, correlation: -1.0,
  correlation': -1.0,
correlation': -1.0,
    correlation': -1.0,
  'correlation': -1.0,
'correlation': -1.0,
'correlation': -1.0,
'correlation': -1.0,
'correlation': -1.0,
   correlation': -1.0,
'correlation': -1.0, 'i': 'Old Man and the Seā, The (1958) },
'correlation': -1.0,
'i': 'Umbrellas of Cherbourg, The (Parapluies de Cherbourg, Les) (1964)'},
'correlation': -1.0, 'i': 'Secret Agent, The (1996)'},
'correlation': -1.0, 'i': 'Beautiful Thing (1996)'},
'correlation': -1.0, 'i': 'Primary Colors (1998)'},
'correlation': -1.0, 'i': 'NeverEnding Story III, The (1994)'},
'correlation': -1.0, 'i': 'Anne Frank Remembered (1995)'},
'correlation': -1.0, 'i': 'Heaven's Prisoners (1996)"},
'correlation': -1.0, 'i': 'House of Yes, The (1997)'},
'correlation': -1.0,
    correlation': -1.0,
 correlation: -1.0,

'i': 'Old Lady Who Walked in the Sea, The (Vieille qui marchait dans la mer,

'La) (1991)'},

'correlation': -1.0, 'i': 'Heaven & Earth (1993)'},

'correlation': -1.0, 'i': 'Something to Talk About (1995)'},
                                                                 'i': 'Heaven & Earth (1993)'},
'i': 'Something to Talk About (1995)'},
'i': 'Man of the House (1995)'},
'i': 'Above the Rim (1994)'},
'i': 'Bread and Chocolate (Pane e cioccolata) (1973)'},
'i': 'Truman Show The (1998)'}.
  correlation : 1.0, 'i':
correlation': -1.0, 'i':
correlation': -1.0, 'i':
                                                                  'i': 'Truman Show, The (1998)'},
'i': 'Beans of Egypt, Maine, The (1994)'}
    correlation':
  'correlation': -1.0,
```

Figure 14: Negative Correlation: Mortal Kombat: Annihilation

In order to get correlations on the movies, I used the built-in function transformPrefs which took the dictionary of users and their ratings and transformed that into a dictionary of movies with user ratings. Then, I just iterated through that dictionary and ran the sim_pearson function on each movie plus my favorite and least favorite movie, sorted those lists, and examined the top and bottom of those lists.

```
#q4
transformPrefs = transformPrefs(prefs)
correlationsToMovie = []
for i in transformPrefs:
    correlation = sim_pearson(transformPrefs, 'Mortal Kombat: Annihilation (1997)', i)
    movieAndCorrelation = {'i':i, 'correlation':correlation}
    correlationsToMovie.append(movieAndCorrelation)
sorted_correlations = sorted(correlationsToMovie, key = lambda item: item['correlation'])
pprint.pprint('Correlations to "Mortal Kombat: Annihilation (1997)"')
pprint.pprint(sorted_correlations)
#pprint.pprint(transformPrefs)
#a04
```

Overall, the recommender was not awe some. I can be kind of picky when it comes to movies, and I do not think I share interests with people my age. I love period pieces such as 'Anna Karenina' and 'Pride and Prejudice'. I also love Holocaust remembrance pieces and historical documentaries. I dislike action and horror. I think it can be really difficult to make recommendations on artistic prefrences. Recommendations for shopping may be easier, as customers who buy one item may certainly be interested in buying similar items or items from a similar brand. Art such as movies and music are totally subjective, and although genres may be similar, people may prefer certain actors, aestetics, and settings.