

<네이버 영화, 왓차피디아 간 영화 평점 경향성 차이에 관한 고찰>

김경훈, 김예진, 서상수

1. 프로젝트 목표

본 프로젝트에서는 네이버 영화와 왓차피디아 간 영화 평점 경향성 차이, 특히 영화 장르별 평점 경향성 차이에 주목하여 고찰하여, 청중들에게 네이버 영화와 왓차피디아의 영화 평점을 참고할 때의 주의점을 안내하고자 한다. 영화 리뷰 사이트 간, 같은 영화임에도 불구하고 영화 평점에 차이가 존재하는 경우가 있다는 점에 주목하여, 이러한 평점 경향성 차이에 대한 분석을 통해 청중들에게 그에 대한 유의미한 정보를 제공하고자 본 프로젝트를 실시하게 되었다.

2. 데이터 수집

- 1) 수집 방법: 네이버 영화, 왓차피디아의 영화 평점 데이터를 python으로 직접 크롤링(크롤링 방법- Appendix (1))
- 2) 수집 대상: SF, 호러, 로맨스, 애니 총 4개의 장르에 대해 장르별 30개씩 영화를 선정, 총 120개 영화에 대한 네이버 영화, 왓차피디아 영화 평점 데이터(4개 장르별 30개 영화 목록- Appendix (2), 수집데이터 형태 - Appendix (3))
- 3) 수집시 유의점
네이버 영화의 영화 평점 스케일은 '1~10', 왓차피디아의 영화 평점 스케일은 '0.5 ~ 5'였다. 데이터 간 스케일을 맞추기 위해 왓차피디아의 영화 평점 데이터에 전부 2를 곱하여 평점 스케일을 '1~10'으로 맞추어 데이터를 수집하였다.

3. 사용 기법

KL Divergence

- 1) 이유: 동일한 영화에 대해 영화 리뷰 사이트 간 영화 평점 분포의 차이를 계산하기 위하여, 확률분포 간 차이를 계산할 수 있는 'KL Divergence' 기법을 사용하였다.
- 2) 합리성: 'KL Divergence' 기법은 같은 변수에 대한 두 개의 분포를 비교하는 데에 용이한 기법이기 때문이다. 그러므로 '영화에 대한 평점'이라는 변수를 두 개의 다른 분포, 즉 네이버 영화의 평점 분포, 왓차피디아의 평점 분포에 대해서 비교하기 용이하다.
- 2) 사용 예시: '인타임' 영화에 대해 네이버 영화와 왓차피디아 간의 평점 분포의 차이를 계산한다고 가정하자. 네이버의 평점 분포를 'Pi', 왓차피디아의 평점 분포를 'Qi'라고 설정한 뒤 평점에 따른 리뷰 건수의 비율을 도출하였다.

평점	1점	2점	3점	4점	5점	6점	7점	8점	9점	10점
비율	0.071	0.023	0.024	0.032	0.072	0.096	0.103	0.116	0.094	0.364

<네이버 영화 평점 분포 'Pi'>

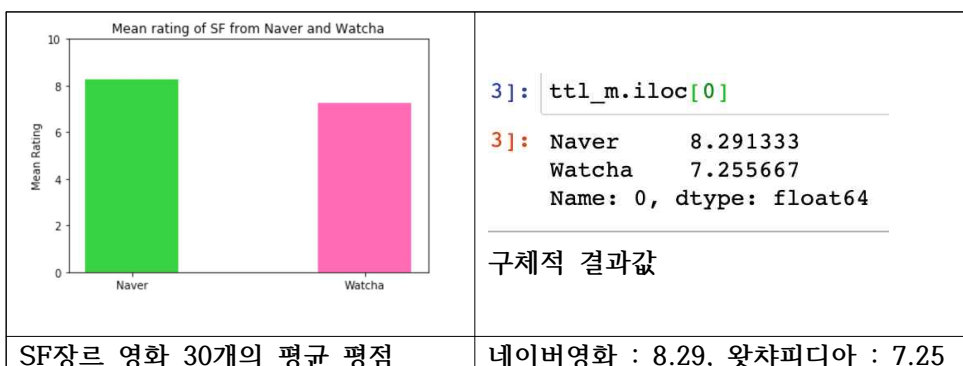
<왓차피디아 평점 분포 'Qi'>

평점	1점	2점	3점	4점	5점	6점	7점	8점	9점	10점
비율	0.018	0.031	0.044	0.142	0.151	0.229	0.150	0.132	0.039	0.059

각 평점에 따른 리뷰 건수의 전체 리뷰 건수에 대한 비율을 '확률'이라고 가정하고 계산했으며, 10개의 discrete한 평점 class에 대해 python에서 scipy.stats.entropy라는 라이브러리를 사용하여 개별 영화의 KL divergence를 계산하였다(자세한 계산 방법- Appendix (4)). 장르별 30개 영화의 KL divergence를 모두 계산한 뒤, 장르별 평균 KL divergence를 구하였다. 이를 통해 장르별로 '네이버 영화'와 '왓차 피디아'의 영화 평점의 차이를 도출해냈다.

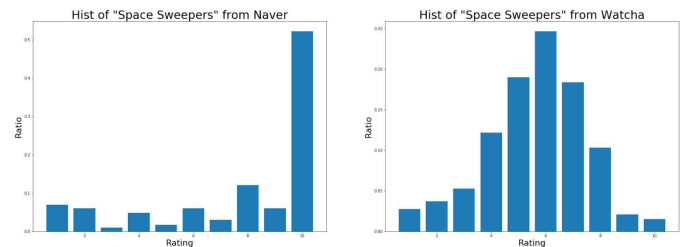
4. 데이터 요약

(I) SF 장르 영화 30개에 대한 네이버 영화, 왓차 피디아 간 영화 평점 평균 분포



우선, SF 장르 영화 30개에 대한 네이버 영화와 왓차피디아의 평균 평점을 확인한 결과, 네이버가 8.29, 왓차가 7.25로, SF 장르의 경우 네이버에서 평균적으로 사용자들이 평점을 후하게 주었다는 것을 확인할 수 있었다.(호러, 로맨스, 애니 장르의 영화 30개에 대한 두 사이트 간 영화 평점 평균 분포에 대한 정보- Appendix (5))

(2) SF 장르 개별 영화에 대한 네이버 영화, 왓차 피디아 간 영화 평점 분포



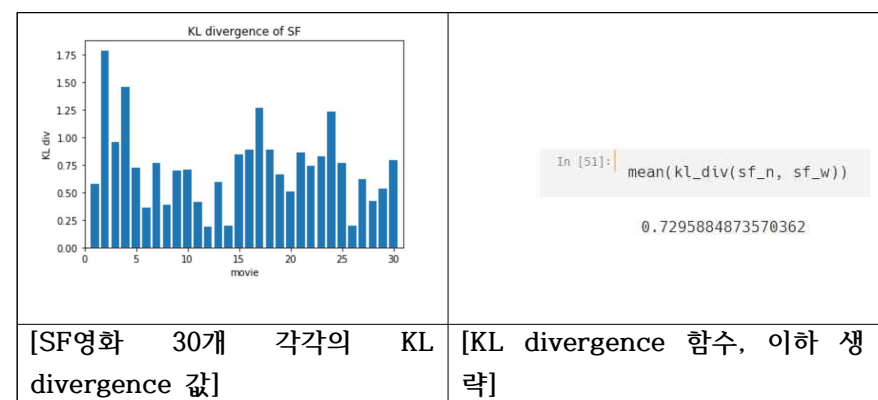
[‘승리호’ 평점 분포], 좌 네이버, 우 왓차

SF 장르 영화 30개 중 ‘승리호’ 영화에 대한 네이버 영화와 왓차 피디아 평점 분포를 확인한 결과, ‘네이버 영화’의 경우 10점으로 높은 평점을 준 비율이 상대적으로 매우 높았으며, ‘왓차 피디아’의 경우 평점 분포가 상대적으로 고르게 분포해 있으며 낮은 평점을 준 비율이 상대적으로 높다는 것을 알 수 있었다.(호러, 로맨스, 애니 장르의 개별 영화에 대한 두 사이트 간 영화 평점 분포에 대한 정보- Appendix (6))

5. KL Divergence 이용 데이터 분석 결과

SF 장르 영화 30개에 대한 네이버 영화, 왓차 피디아 간 영화 평점 차이 분포

SF 장르 영화 30개 각각의 KL divergence를 계산하고, 영화 30개 KL divergence의 평균을 구하여, SF 장르에 대한 ‘네이버 영화’와 ‘왓차 피디아’의 영화 평점의 차이 분포를 도출해냈다.



SF 장르의 평균 KL divergence 값으로 0.73 이라는 값을 도출할 수 있었다.(호러, 로맨스, 애니 장르의 영화 30개에 대한 두 사이트 간 영화 평점 차이 분포에 대한 정보- Appendix (7))

● 장르별 KL divergence 값 정리

SF	호러	로맨스	애니
0.73	0.913	0.637	1.204

[Findings & Insight]

1. 기본적으로 네이버 영화는 왓챠피디아보다 평점을 후하게 주는 경향이 있다.
2. 네이버 영화의 평점 개별 분포는 10점에 과도하게 몰려있고 나머지는 비교적 균등하게 분포되어 있는 것을 확인할 수 있다. 이로 인해 플랫폼 특성상 광고나 봇 유저의 존재 가능성을 의심해볼 수 있다.
3. 반면, 왓챠피디아에서는 아무리 유명하고 대중적 평가가 좋은 영화도 10점이 아닌 7.8점을 중심으로 분포되어 있다. 즉, 네이버는 종합적 기능의 플랫폼인데에 비해 왓챠피디아는 전문적 플랫폼이기에 네이버 영화보다 왓챠피디아의 경우, 영화를 매니악하게 감상하는 유저가 많고 때문에 평가가 까다로운 것으로 판단된다.
4. 장르별 kl divergence 분포를 보면 애니>호러>SF>로맨스 순으로 평점에 차이가 난다. 4개 장르 밖에 비교하지 못했지만 전자 2개의 장르는 후자 2개의 장르에 비해 매니아 층이 많기 때문에 위에서 언급한 매니악/헤비유저의 수가 많다는 것이 어느 정도 타당해 보인다.
5. 그러므로 청중은 네이버 영화가 기본적으로 왓챠피디아보다 평점을 후하게 주는 경향이 있다는 것, 그리고 애니, 호러 장르의 경우 네이버 영화와 왓챠피디아 간의 영화 평점 차이가 꽤 존재한다는 것을 유의하여 두 사이트의 평점을 참고하는 것이 좋을 것으로 기대된다.

[한계]

1. 본 프로젝트에서는 장르별로 30개씩 비교했지만, 이 표본 개수를 늘린다면 보다 정확한 KL divergence 평균값을 구하여 좀 더 정확한 분포 차이를 제시할 수 있을 것으로 기대된다.
2. 장르의 개수를 늘린다면 더 다양한 평점 차이를 볼 수 있어 이용자에게도 실제로 더 유의미하게 제공할만한 정보가 추가되었을 것으로 기대된다.
3. 크롤링한 데이터 중 각각 감상평에 대한 좋아요, 싫어요 수에 대한 정보가 존재하였다. 좋아요 수가 많은 감상평과 그 점수는 많은 사람들의 공감을 받았다는 뜻이기에 신빙성이 높은 감상평/점수라고 추측할 수 있다. 개별 감상평에 대한 좋아요/싫어요의 비율을 가중치로 반영하였다면 좀 더 유의미한 결과를 제시할 수 있을 것으로 기대된다.
4. KL divergence 뿐만 아닌 kolmogorov-smirnov 검정 등 다른 방법론을 도입하여 그 결과를 비교하여 만약 비슷한 경향의 결과값이 나온다면 더욱 신빙성 있는 결과가 되었을 것이다.

Appendix (1) 크롤링 방법

코멘트 10000+

더보기

이동진 평론가

★ 4.0

하나를 위한 전체. 세상에서 가장 낙천적인 (SF)재난영화.

2395 35

좋아요

신준섭

★ 4.0

스포일러가 있어요!! 보기

213 86

좋아요

	id	rating	text	num
36049	이동진 평론가	4.0	하나를 위한 전체. 세상에서 가장 낙천적인 (SF)재난영화.	11
36050	신준섭	4.0	NaN	11
36051	Arcturus	4.0	정말 간절히 원하면 천 우주가 나서서 다 같이 도와준다.	11
36052	외계이	5.0	한단어로 말해보자.wn콘택트 : Meetwn그래비티 : Livewn인터스텔라 : S...	11
36053	루르나	4.0	살면서 본 가장 멋진 귀농	11
...
42215	여름SUMMER	4.0	마크 워트니는 마블의 워트를 뛰어 넘는다	11
42216	송광	4.5	CG 자연스럽고 결말도 좋음	11
42217	주은	4.5	한 사람을 살리려는 여러 사람의 노력과wn그 한 사람의 준비로 이뤄진 화성구출기	11
42218	재은	4.0	저런 마인드로 살아가면 뭐든 못할까	11
42219	김경수	4.5	무렵고진지할줄알았는데유쾌했음흥사형에서살아남기wn감자먹고싶다	11

6171 rows × 4 columns

마션, 왓차피디아 리뷰

마션, 크롤링 결과 dataframe

[왓차 크롤링 예시]

관람객 평점 14,729건

내 평점 등록

공감순 최신순 평점 높은 순 평점 낮은 순 스포일러 보기 관람객 평점만 보기

★★★★★ 10

관람객

보면 감자먹고 싶어지는 영화

e2es**** | 2015.10.08 11:20 | 신고

4299 448

★★★★★ 9

사람은말단독특해야된다

왕눈이(sooc****) | 2015.10.08 22:23 | 신고

2927 120

★★★★★ 8

화성판 삼시세끼...

드러그엔젤(drug****) | 2015.10.08 11:32 | 신고

3243 447

	id	date	rating	text	watched	like	dislike	num
64146	e2es****	2015.10.08 11:20	10.0	부면 감자먹고 싶어지는 영화	1	4299	448	11
64147	왕눈이(sooc****)	2015.10.08 22:23	9.0	사람은말단독특해야된다	0	2927	120	11
64148	드러그엔젤(drug****)	2015.10.08 11:32	8.0	화성판 삼시세끼...	0	3243	447	11
64149	박공95(joon****)	2015.10.08 01:12	10.0	최고의 영화. 다만 이 영화를 인터스텔라같은 영화라고 기대하고 보지 마시고 테스트 ...	0	2497	359	11
64150	d34d****	2015.10.08 08:51	10.0	내가더낫조? 낄 암스트롱? 이말한마디로 지켰다.	0	2537	453	11
...
78871	STEELHEART(mool****)	2018.08.10 12:24	2.0	인터스텔라 직후리로 만든느낌. 넷딩이는 조연급으로 나와야 영화가상당. 라이언왕벌부터...	0	2	13	11
78872	*ㅈㅇ(yje****)	2015.10.08 10:43	6.0	결말이 별하게 보여서 긴장감이 좀 떨어지고, 감동면에서나 쿼트면에서나 다소 황당...	1	14	26	11
78873	니제다(niet****)	2019.06.29 16:40	1.0	말도 안되고 개지루한 영화. 리플리스로감독이라 기대했지만 인간은 농오면 정신적으로로...	0	5	19	11
78874	havehappymind(tele****)	2019.07.07 14:49	1.0	별 한개도 아까워 저발 보지마	0	1	19	11
78875	자유의 남신상(ehdt****)	2015.10.08 00:03	1.0	레브라이브 만세! 레브라이브드스콜아이들루비 빼고는 다 쓰레기 영화!ㅜㅜ	0	5	34	11

14730 rows × 8 columns

마션, 네이버 리뷰

마션, 크롤링 결과 dataframe

[네이버 크롤링 예시]

Appendix (2) 4개 장르별 30개 영화 목록

* 장르마다, 순서대로 1~30으로 넘버링

1) SF

[인 타임/승리호/로보캅/터미네이터:미래 전쟁의 시작/아바타/매트릭스/아이,로봇/컨택트/E. T./테넷/마션/인터스텔라/엣지 오브 투모로우/그래비티/아마겔돈/지구가 멈추는 날/공각기동대/아일랜드/마이네리티 리포트/맨 인 블랙/흑성탈출: 종의 전쟁/프로메테우스/트론:새로운 시작/패신저스/퍼시픽 림/인셉션/더 문/레디 플레이어 원/핀치/토탈 리콜]

2) 호러

[랑종/곤지암/더 년/장산범/그것/직쏘/장화, 홍련/겟 아웃/더 퍼지/오편: 천사의 비밀/컨저링/검은 사제들/콘스탄틴/링/미드소마/숨바꼭질/곡성/맨 인 더 다크/쏘우/쥬스/어스/해피 데스데이/파라노말 액티비티/알포인트/여고괴담 3:여우계단/데드 사일런스/스크림/제인 도/박쥐/고사: 피의 중간고사]

3) 로맨스

[500일의썸머/건축학개론/그여자작사그남자작곡/나는내일어제의너와만난다/남자가사랑할때/내머리속의지우개/너는내운명/노트북/노팅힐/라라랜드/러브,로지/러브레터/러브앳유얼리/로맨스/만추/맘마미아/먼훗날우리/미비포유/반창꼬/비긴어게인/비포선라이즈/어바웃타임/연애의온도/엽기적인그녀/오직그대만/월스/이터널선샤인/조제,호랑이그리고물고기들/타이타닉/트와일라잇/플립]

[겨울왕국/라푼젤/시간을 달리는 소녀/엽/하울의움직이는성/라따뚜이/토이스토리/너의이름은/몬스터주식회사/니모를 찾아서/인사이드아웃/라이온킹/인크레더블/쿵푸팬더/슈렉/몬스터대학교/물란/슈퍼배드/미니언즈/마다가스카/꿀벌대소동/볼트/해피피트/보스베이비/초속5센티미터/이웃집토토로/모노노케히메/벼랑위의포뇨/알라딘/아이스에이지]

1) SF

[illegible]

	id	date	rating	text	watched	like	dislike	num
0	헛살(hhs0****)	2021.07.14 13:41	6.0	공포가 아니라 그 날, 기억할, 나중에 가길 그 기억할까 게도 아니니들, 헤어리 다행히.	1	3645	536	1
1	원희보(wonhi****)	2021.07.14 16:35	1.0	이런 영화보고 하 려 시간과 돈을 남 기까지 왜로 세 상에는 아직 못하 를 다른 좋은 ...	1	3647	1102	1
2	jeongleej****	2021.07.14 13:01	7.0	로그인하고 기본이 자랑스럽게 되어서 진짜 좋으니 좋았 다.	1	1862	214	1
3	윤선영(yoon****)	2021.07.14 14:20	1.0	종간 19일은 딱히 좋은 영화. 티비 vod 중에 무조건 좋은 10년 만 영화 는 없음...	1	2242	754	1
4	박동권(dong****)	2021.07.15 02:14	1.0	버이탈 마케팅에 나름대로 남비한 것 에 대해서 좋구나 는 영화 그렇게 우 습지도 않고 재미...	1	1475	340	1
...
216250	해리캐인(paul****)	2017.03.03 10:32	3.0	소재는 꽤나 괜찮 았으나 역시나 날 수확 거당성이 많 아서는 전혀 우성 지도 않은 영화	0	0	3	30

	id	rating	text
0	박재민	3.5	황동호 나재석의 옷차 크렌트 보트wn웃자의 영입이와 적자책이 매년wn커지고있는 이...
1	CineVet	1.5	귀신에게 책임을 전가한 스티프 필름. <황동>에서 카미라는 귀신의 공범이다.
2	박종민	3.5	NaN
3	석미민	3.0	관동: 카미라를 잡으려 안돼?
4	이동진 권문가	3.5	전자는 어질지 않아서 만들어 줄요로 만든 거저한 여긴다. (노자 '도덕경')
...
34605	이쁜미래	4.0	불만였다
34606	May	1.0	장문은 있는데 스토리가 없다...
34607	박희성	2.0	인우서운 스티리 드드해...
34608	안녕장구물	1.0	ㅎㅎ 후속이 나올까와 두려워
34609	군파지보	1.5	정신 초라함... 학교에서 일어나는 이야기라 좀더 흥미이 잘 되었던... 하지만...

34610 rows x 4 columns

호러 네이버 리뷰 전체 데이터	호러 왓차 리뷰 전체 데이터
------------------	-----------------

3) 로맨스

id date rating text watched like dislike num									
0	kunw****	2013.07.01 23:47	10.0	법마의 심리 버전으로 영화 나왔으면 좋겠다.	0	3389	88	1	
1	국태지(jpsyc****)	2013.06.17 16:47	7.0	이렇게 여름(summer)이 끝나고 가을(autumn)이 온다.	0	2395	83	1	
2	유시은	2015.07.24 01:59	8.0	좋은 파티에 가져간 선물이 간혹적이 아니라 핑크스타 열병이었던 어렸을때, 핑크는 ...	0	2161	48	1	
3	o o (jd_dd****)	2015.05.28 17:10	10.0	영화가 훌륭한 남자는 자신이 알고있는 책의 내용을 묻는, 자신이 좋아하는것에 관심.	0	2151	44	1	
4	찰리슈슈(doub****)	2014.05.17 09:41	9.0	인간을 발라도 너를 잘 이해할까. 그게 이 영화가 볼수록 마음 아를 수 있을 것 같...	0	2031	139	1	
...	
233913	Stechkin(ykuh****)	2017.07.12 10:02	1.0	평점조절용 1점 두려함	0	7	41	30	
233914	비비씨(trgu****)	2018.01.04 18:14	5.0	너무기다를 않아해서 너무너무 좋았다.	1	11	46	30	

id rating text num			
0	이동진 평론가	9.0	사형은 꼭 그 사람일 필요가 없는 운명을 반드시 그 사람이어야만 하는 운명으로 바꾸... 1
1	RHY	9.0	NaN 1
2	최영훈	10.0	운명론자인 주인공, 운명을 부정하는 심마, 둘 다 정답은 아니었다.wns님없이 스쳐... 1
3	유시은	10.0	범마가 나쁜년인 이유는 이 영화가 순전히 좋은 이야기이기 때문이다. 1
4	sue	10.0	날 사랑하고있던 너의 마음을 사랑하고 있는건 아닌지 1
... 1
44121	우비니니니	10.0	사랑말네 30
44122	영화부족민	6.0	뜻밖한 소품. 30
44123	tokibo	5.0	난 잘 모르겠어요. 그저 작하기만해 30
44124	림스지킨	9.0	이 영화를 좋아하는 사람을 좋아할 것 같다 30
44125	주은지	10.0	진짜 이런 영화 또 없나. 30
44126 rows × 4 columns			

로맨스 네이버 리뷰 전체 데이터

로맨스 왓차 리뷰 전체 데이터

4) 애니

id date rating text watched like dislike num									
0	윙표(grea****)	2014.01.16 02:05	10.0	수많은 세월동안 축적된 디즈니의 장수를 보여준 영화. 겨울왕국의 성공은 디즈니와 함께...	0	6316	677	1	
1	보그잇나(kung****)	2014.01.16 01:35	10.0	프로튼 할 장 이요...	0	5467	666	1	
2	추차용(jifq****)	2014.01.16 01:24	10.0	아...너무 감동적이라 눈물이 땀(바다)도 올라...	0	5177	715	1	
3	시나리요(hong****)	2014.01.16 01:24	9.0	환상미 10/10재미 10/10노래 10/10재치덕 9/10 스토리 7/10스토리가 ...	0	4359	838	1	
4	책구인(dif****)	2014.01.16 14:32	10.0	그래...전론 상우를쓰리니까, 아이들이나 유행어 낱말하는거그면 어용지않는네...	0	1941	184	1	
...	
197605	김원성(phl****)	2015.03.01 17:20	1.0	뭔 애니가 이렇게 쏘나 지루하고 오글거림	0	1	3	30	
197606	ssh0****	2019.06.02 00:21	1.0	hglkytgjkrthr	0	0	3	30	
197607	shasha(jgh****)	2019.03.10 13:26	5.0	기대여 못미친다 스토리전개가 지루	0	0	3	30	
197608	hang****	2014.12.02 20:46	1.0	알도 안되는 스토리네. 발판들이 좋았는데 보고 사람들이 생긴 장면을 알은다는데...	0	2	5	30	
197609	죽목종종종종종종 (jji****)	2018.12.04 06:08	2.0	송추니 너무으로 좋아하는 막장전개. 둘 아떡하라는 거지???	0	0	4	30	
197610 rows × 8 columns									

id rating text num			
0	TUDDI	5.0	스토리는 라몬질이 한수위지만, 캐릭터와 영상미, let it go 부르는 황... 1
1	신정훈	4.0	디즈니가 보여주고 싶은 것은 어떤 작품적인 요소가 아니라, 사랑 그 자체이다. 1
2	Park Sho Yeon	5.0	진짜 let it go 부를때환상이었음 스토리는그중부족한데 캐릭터들이 환상이었다 ... 1
3	송잔국	3.0	사운드 트랙은 어론 눈놀이wns스토리는 어린이 눈놀이 1
4	이동진 평론가	3.5	종목적인 노래들과 낭만적인 겨울왕국의 위력. 1
... 1
49578	Rina	4.0	안어플라는 조합이 나중에는 하나로 융된다~ 우린 친구 30
49579	나탈	4.0	아이스 어이자는 먼 첫 연이 제일 재밌다 30
49580	Yeezus	3.0	풍광 얼어붙은 빙하기 시대에도 따뜻한 정은 살아있다. 30
49581	snoopy3	2.0	보다가 좋았음 ...ㅋㅋㅋ 30
49582	크리다를	3.0	온 가족 애니메이션 30
49583 rows × 4 columns			

애니 네이버 리뷰 전체 데이터

애니 왓차 리뷰 전체 데이터

Appendix (4) 개별 영화의 KL divergence 계산 방법

$$KL(P||Q) = \sum p_i(x) \log\left(\frac{p_i(x)}{q_i(x)}\right)$$

scipy.stats.entropy

`scipy.stats.entropy(pk, qk=None, base=None, axis=0)` [\[source\]](#)

Calculate the entropy of a distribution for given probability values.

If only probabilities *pk* are given, the entropy is calculated as *S* = -sum(*pk* * log(*pk*), axis=*axis*).

If *qk* is not None, then compute the Kullback-Leibler divergence *S* = sum(*pk* * log(*pk* / *qk*), axis=*axis*).

This routine will normalize *pk* and *qk* if they don't sum to 1.

Parameters: *pk* : *array_like*

Defines the (discrete) distribution. Along each axis-slice of *pk*, element *i* is the (possibly unnormalized) probability of event *i*.

qk : *array_like, optional*

Sequence against which the relative entropy is computed. Should be in the same format as *pk*.

base : *float, optional*

The logarithmic base to use, defaults to *e* (natural logarithm).

axis : *int, optional*

The axis along which the entropy is calculated. Default is 0.

Returns: *S* : (*float, array_like*)

The calculated entropy.

(출처: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.entropy.html>)

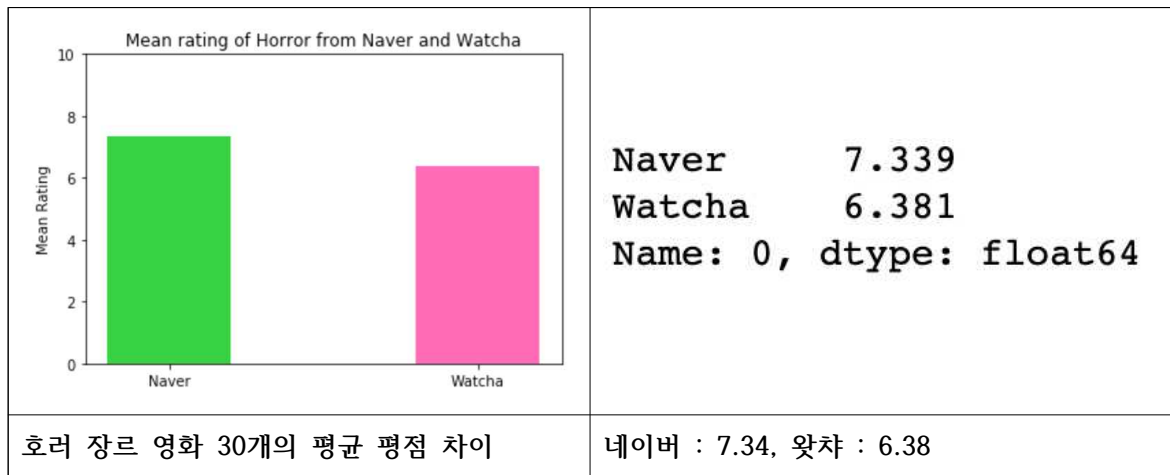
Python에서 scipy.stats.entropy라는 라이브러리를 사용

```
In [42]: import pandas as pd
         from scipy.stats import entropy
         from statistics import mean

In [48]: def kl_div(df1, df2):
         kl_div_result = []
         for i in range(30):
             n_list = list(df1.iloc[i, 1:11])
             w_list = list(df2.iloc[i, 1:11])
             kl_val = entropy(pk = n_list, qk = w_list)
             kl_div_result.append(kl_val)
         return kl_div_result
```

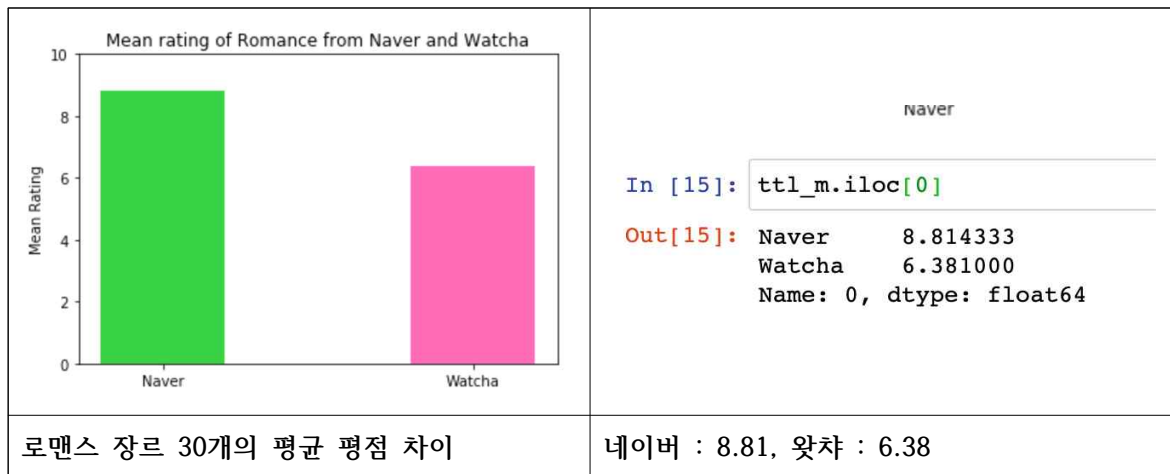

Appendix (5) - 호러, 로맨스, 애니 장르의 영화 30개에 대한 두 사이트 간 영화 평점 평균 분포에 대한 정보

1) 호러



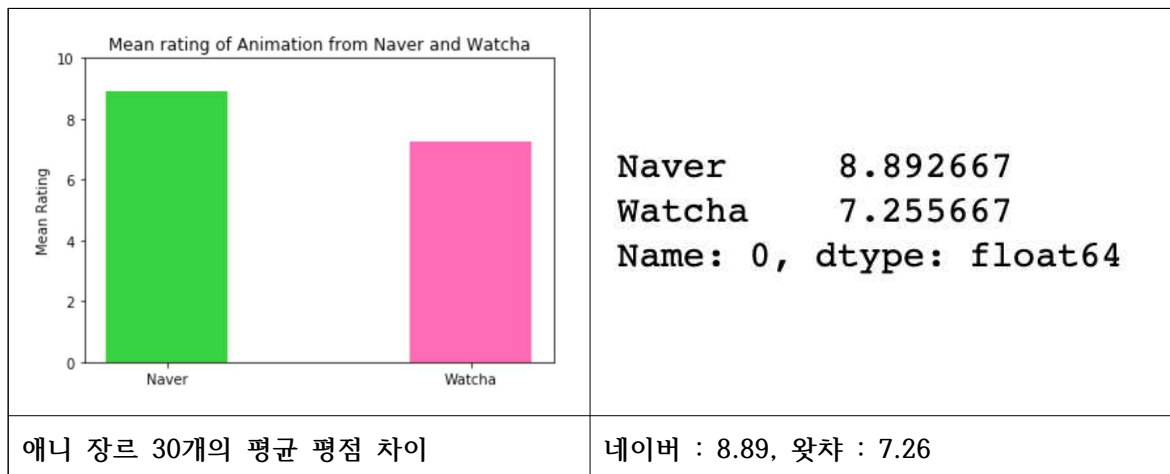
호러 장르 영화 30개에 대한 네이버 영화와 왓챠피디아의 평균 평점을 확인한 결과, 네이버가 7.34 왓챠가 6.38로, 호러 장르의 경우도 네이버에서 평균적으로 사용자들이 평점을 후하게 주었다는 것을 확인할 수 있었다.

2) 로맨스



로맨스 장르 영화 30개에 대한 네이버 영화와 왓챠피디아의 평균 평점을 확인한 결과, 네이버가 8.81, 왓챠가 6.38로 로맨스 장르에서도 네이버에서 평균적으로 사용자들이 평점을 후하게 주었다는 것을 확인할 수 있었다.

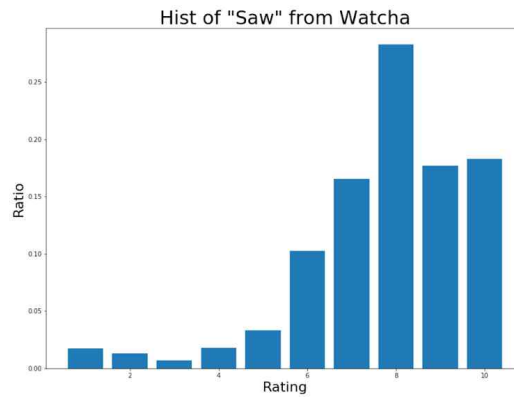
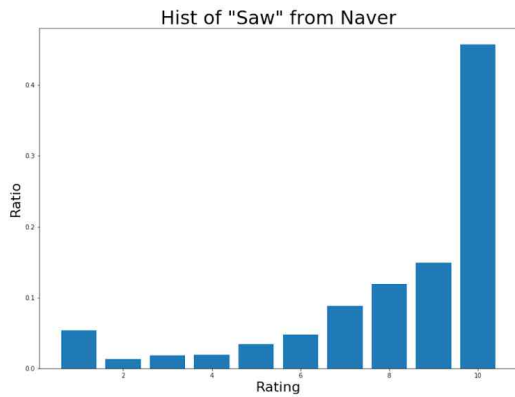
3) 애니



애니 장르 영화 30개에 대한 네이버 영화와 왓챠피디아의 평균 평점을 확인한 결과, 네이버가 8.89, 왓챠가 7.26으로 애니 장르에서도 네이버에서 평균적으로 사용자들이 평점을 후하게 주었다는 것을 확인할 수 있었다.

Appendix (6) - 호러, 로맨스, 애니 장르의 개별 영화에 대한 두 사이트 간 영화 평점 분포에 대한 정보

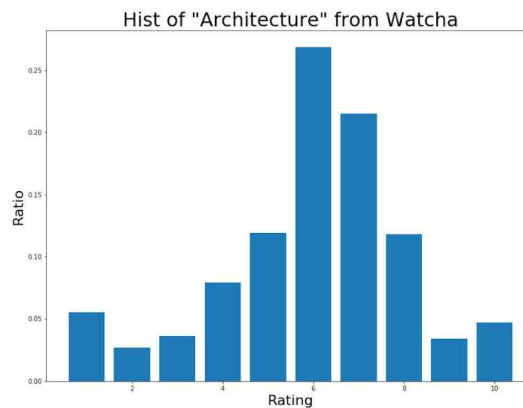
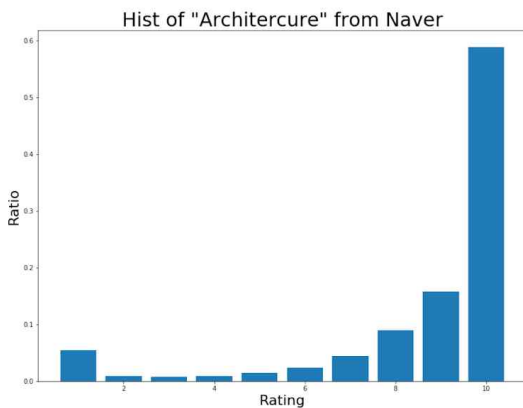
1) 호러



['쏘우' 평점 분포], 좌 네이버, 우 왓차

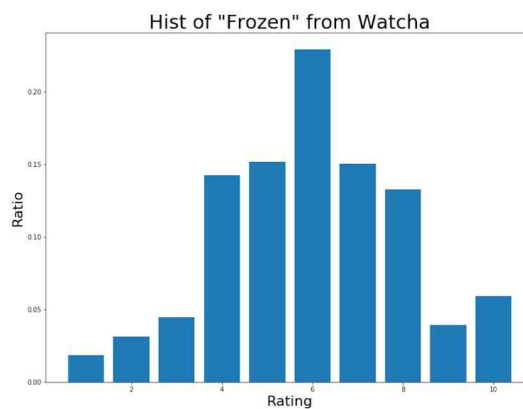
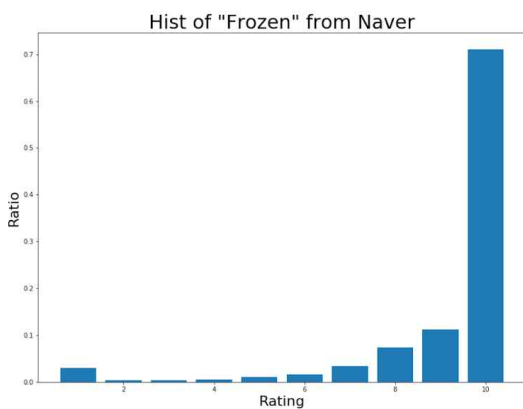
위 그림은 '쏘우'의 평점 분포를 나타낸 그래프이다. 마찬가지로 네이버는 높은 평점의 비율이, 왓차는 상대적으로 낮은 평점의 비율이 높았다는 것을 확인할 수 있었다.

2) 로맨스



위 그림은 '건축학개론'의 평점 분포를 나타낸 그래프이다. 마찬가지로 네이버는 높은 평점의 비율이, 왓차는 상대적으로 낮은 평점의 비율이 높았다는 것을 확인할 수 있었다.

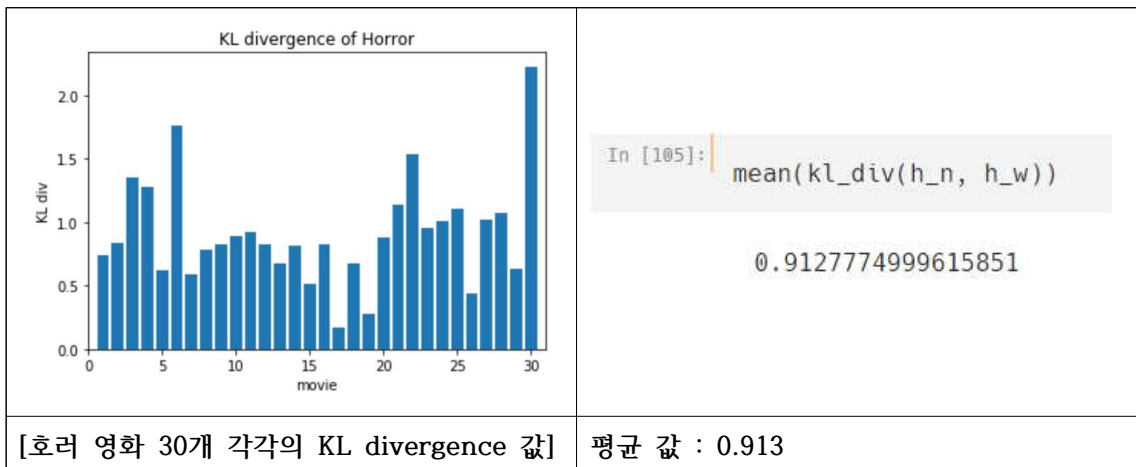
3) 애니



위 그림은 '겨울왕국'의 평점 분포를 나타낸 그래프이다. 마찬가지로 네이버는 높은 평점의 비율이, 왓차는 상대적으로 낮은 평점의 비율이 높았다는 것을 확인할 수 있었다.

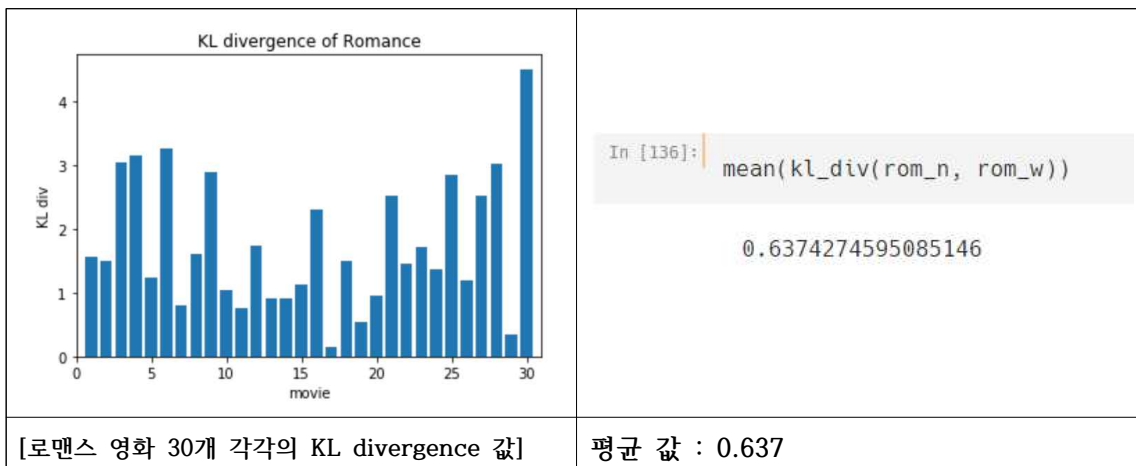
Appendix (7) -호러, 로맨스, 애니 장르의 영화 30개에 대한 두 사이트 간 영화 평점 차이 분포에 대한 정보

1) 호러



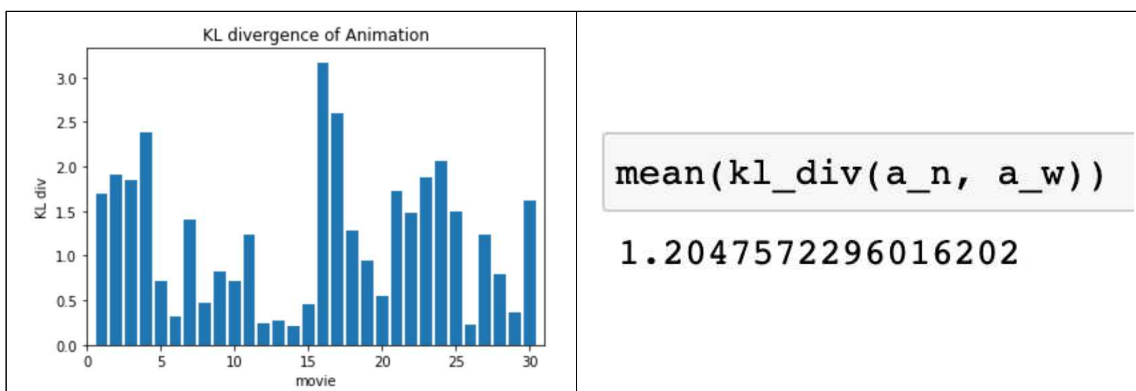
마찬가지로 KL div를 계산하고 평균을 낸 결과 0.913이라는 값을 얻었으며, SF 장르에 비해 더 큰 차이가 난다는 사실을 알 수 있었다.

2) 로맨스



KL div의 평균 값은 0.637이라는 값을 얻었으며, 타 장르 대비 두 사이트 간 분포 차이가 평균적으로 덜하다는 것을 알 수 있었다.

3) 애니



[애니 영화 30개 각각의 KL divergence 값]	평균 값 : 1.204
---------------------------------	--------------

KL div의 평균 값으로 1.204라는 값을 얻었으며, 타 장르 대비 두 사이트 간 분포 차이가 가장 두드러진다는 것을 알 수 있었다.