

SF 영화를 통한 신기술 인지 및 관심도 분석

물리학 전공

1873026 임재랑

1. 주제 선정 과정

자율주행 자동차, 새로운 디스플레이와 센서, 로봇 등 과거 Sci-Fi 영화 속에 등장했던 기술들이 현재 많은 곳에서 상용화 된 경우를 종종 발견할 수 있다. 이처럼 SF 영화 속 특정 주제에 대한 관심은 사람들의 과학기술에 대한 관심도를 반영하며, 이에 대한 새로운 개발 및 발전이 이루어 질 것임을 암시한다. 따라서, 사람들이 관심을 가지는 SF 분야와 이에 대한 키워드를 추출하여 신기술에 대한 사람들의 인지와 관심도를 분석하고자 한다. 부가적으로 10년간 사람들의 SF에 대한 관심도 변화를 확인하여 사람들의 과학기술에 대한 인식 변화를 확인하고자 한다.

2. 데이터 수집

1) SF 영화 뉴스 크롤링

SF영화에 대해 지속적으로 언급되는 키워드에 대해 확인해보고자 줌 뉴스 페이지에서 2년 6개월 간의 기간동안 발생한 뉴스 20페이지 크롤링을 진행하였다. 검색어는 'SF 영화'로 진행하였다. 이 때 10년간 사람들의 SF에 대한 관심도 변화를 확인하고자 2010.01.01~2012.06.01 기간과 2020.01.01~2022.06.01 기간의 기사를 모두 크롤링 하였다.

2) Kaggle을 통해 인기 영화 리스트 수집

영화 산업의 추세를 분석하기 위해 Kaggle 사이트에서 2022년 업데이트 된 Highest Hollywood Grossing Movies 파일을 다운로드하였다.

3) 영화 리뷰 데이터 수집

사람들의 과학기술에 대한 선호와 인식을 확인하기 위해 크롤링과 Kaggle 데이터 분석을 통해 선정한 영화 8가지의 리뷰를 각 20개씩 수집하였다.

3. 수집 데이터 분석

1) 뉴스 데이터 워드클라우드

웹크롤링을 통해 수집한 2010.01.01~2012.06.01 기간과 2020.01.01~2022.06.01 기간의 기사의

키워드를 확인하기 위해 워드클라우드를 진행하였다. 과학기술과 관련된 키워드를 수집하기 위해 감독 이름 및 언론사에 대한 단어들은 전처리 과정을 통해 제거하였다.



그림 1 2010.01.01~2012.06.01 워드클라우드



그림 2 2020.01.01~2022.06.01 워드클라우드

결과는 다음과 같으며, 2010.01.01~2012.06.01 기간 'SF 영화'에 대한 결과 중 과학기술에 관해서 세계, 멸망, 인류 등의 단어를 찾아볼 수 있었다. 2020.01.01~2022.06.01 결과에서는 우주, 미래, 로봇 등의 키워드를 확인할 수 있었다.

2) 뉴스 잠재내용 분석

SF 영화에 대해 수집한 기사들에 대해서 어떠한 토픽으로 어떠한 내용들이 위주를 이루고 있는지 분석하고자 잠재의미 분석을 실시하였다. 토픽은 3가지로 분류하도록 하였으며, 10년 전과 최신 기사 자료 모두에 대해 실시하였다.

Topic 0	감독	리포트	개봉	관객
Topic 1	어벤져스	관객	누적	리포트
Topic 2	멸망	인류	보고서	로봇

표 1 2010.01.01~2012.06.01 잠재의미 분석 결과

<표 1>은 2010.01.01~2012.06.01 토픽별 단어 추출 결과이며 토픽 2에 집중하면 멸망, 인류, 로봇이라는 분야에 SF 분야가 집중하고 있었음을 확인할 수 있다.

Topic 0	감독	승리	우주	공개
Topic 1	감독	패틴슨	미키	봉준호
Topic 2	승리	넷플릭스	우주	로봇

표 2 2020.01.01~2022.06.01 잠재의미 분석 결과

<표 2>는 2020.01.01~2022.06.01 토픽별 단어 추출 결과이며 토픽에 집중하면 우주, 로봇, 승리 등의 키워드를 통해 10년간 과학기술과 미래에 대한 사람들의 인식이 상대적으로 긍정적으로

혹은 새로운 변화로 변화하였음을 유추할 수 있다.

3) 인기 영화 키워드 분석

Unnamed: 0	Title	Movie Info	Distributor	Release Date	Domestic Sales (in \$)	International Sales (in \$)	World Sales (in \$)	Genre	Movie Runtime	License
0	Star Wars: Episode VII - The Force Awakens (2015)	As a new threat to the galaxy rises, Rey, a de...	Walt Disney Studios Motion Pictures	December 16, 2015	936662225	1132859475	2069521700	['Action', 'Adventure', 'Sci-Fi']	2 hr 18 min	PG-13
1	Avengers: Endgame (2019)	After the devastating events of Avengers: Inf...	Walt Disney Studios Motion Pictures	April 24, 2019	858373000	1939128328	2797501328	['Action', 'Adventure', 'Drama', 'Sci-Fi']	3 hr 1 min	PG-13
2	Avatar (2009)	A paraplegic Marine dispatched to the moon Pan...	Twentieth Century Fox	December 16, 2009	760507625	2086738578	2847246203	['Action', 'Adventure', 'Fantasy', 'Sci-Fi']	2 hr 42 min	PG-13
3	Black Panther (2018)	T'Challa, heir to the hidden but advanced king...	Walt Disney Studios Motion Pictures	NaN	700426566	647171407	1347597973	['Action', 'Adventure', 'Sci-Fi']	2 hr 14 min	NaN
4	Avengers: Infinity War (2018)	The Avengers and their allies must be willing ...	Walt Disney Studios Motion Pictures	NaN	678815482	1369544272	2048359754	['Action', 'Adventure', 'Sci-Fi']	2 hr 29 min	NaN
5	Spider-Man: No Way Home (2021)	With Spider-Man's identity now revealed, Peter...	Sony Pictures Entertainment (SPE)	NaN	675813257	868642706	1544455963	['Action', 'Adventure', 'Fantasy', 'Sci-Fi']	2 hr 28 min	NaN
6	Titanic (1997)	A seventeen-year-old aristocrat falls in love ...	Paramount Pictures	December 19, 1997	659363944	1542283320	2201647264	['Drama', 'Romance']	3 hr 14 min	PG-13
7	Jurassic World (2015)	A new theme park, built on the original site o...	Universal Pictures	June 10, 2015	652385625	1018130819	1670516444	['Action', 'Adventure', 'Sci-Fi']	2 hr 4 min	PG-13
8	The Avengers (2012)	Earth's mightiest heroes must come together an...	Walt Disney Studios Motion Pictures	April 25, 2012	623357910	895457605	1518815515	['Action', 'Adventure', 'Sci-Fi']	2 hr 23 min	PG-13
9	Star Wars: Episode VIII - The Last Jedi (2017)	The Star Wars saga continues as new heroes ...	Walt Disney Studios Motion Pictures	December 13, 2017	620181382	712517448	1332698830	['Action', 'Adventure', 'Fantasy', 'Sci-Fi']	2 hr 32 min	PG-13

그림 3 Highest Hollywood Grossing Movies top 10

Kaggle에서 수집한 데이터를 dataframe 형태로 불러와 상위 10개 영화를 출력한 결과 10개 영화중 9개 영화가 Sci-Fi영화임을 확인할 수 있다. 이는 2022년 현재 많은 사람들이 SF영화와 과학기술에 높은 관심을 보이고 있음을 의미한다.

```
[ ] Counter(df_sf['Movie Info'].str.contains('space', case=False))
# SF 영화 설명에 space를 포함하는 영화의 수 (대소문자 구분 x)

Counter({False: 210, True: 18})
```



```
[ ] Counter(df_sf['Movie Info'].str.contains('future', case=False))
# SF 영화 설명에 future를 포함하는 영화의 수 (대소문자 구분 x)

Counter({False: 218, True: 10})
```



```
[ ] Counter(df_sf['Movie Info'].str.contains('robot', case=False))
# 영화 설명에 robot를 포함하는 영화의 수 (대소문자 구분 x)

Counter({False: 222, True: 6})
```

그림 4 dataframe에서 키워드의 개수 출력 결과

장르가 SF인 영화들만 dataframe에 남긴 후 영화 정보에 'space' 'future' 'robot' 키워드가 들어간 영화의 개수를 확인하였다. 그 결과 228개의 SF영화 중 space 가 18개, future가 10개, robot가 6개 포함되어 있음을 확인할 수 있었다.

Unnamed: 0	Title	Movie Info	Distributor	Release Date	Domestic Sales (in \$)	International Sales (in \$)	World Sales (in \$)	Genre	Movie Runtime	License
55	Transformers: Dark of the Moon (2011)	The Autobots learn of a Cybertronian spacecraf...	DreamWorks	June 29, 2011	352390543	771403536	1123794079	['Action', 'Adventure', 'Sci-Fi']	2 hr 34 min	PG-13
108	Gravity (2013)	Two astronauts work together to survive after ...	Warner Bros.	October 3, 2013	274092705	449100000	723192705	['Action', 'Drama', 'Sci-Fi', 'Thriller']	1 hr 31 min	PG-13
165	WALL-E (2008)	In the distant future, a small waste-collectin...	Walt Disney Studios Motion Pictures	June 26, 2008	223808164	297503726	521311890	['Adventure', 'Animation', 'Family', 'Sci-Fi']	1 hr 38 min	NaN

그림 5 영화 정보에 space 단어를 포함하는 영화 상위 3개

Unnamed: 0	Title	Movie Info	Distributor	Release Date	Domestic Sales (in \$)	International Sales (in \$)	World Sales (in \$)	Genre	Movie Runtime	License
3	3 Black Panther (2018)	T'Challa, heir to the hidden but advanced king...	Walt Disney Studios Motion Pictures	NaN	700426566	647171407	1347597973	['Action', 'Adventure', 'Sci-Fi']	2 hr 14 min	NaN
122	122 Star Trek (2009)	The brash James T. Kirk tries to live up to his...	Paramount Pictures	May 6, 2009	257730019	127950427	385680446	['Action', 'Adventure', 'Sci-Fi']	2 hr 7 min	PG-13
148	148 War of the Worlds (2005)	An alien invasion threatens the future of huma...	Paramount Pictures	June 29, 2005	234280354	369592765	603873119	['Action', 'Adventure', 'Sci-Fi']	1 hr 56 min	PG-13

그림 6 영화 정보에 future 단어를 포함하는 영화 상위 3개

Unnamed: 0	Title	Movie Info	Distributor	Release Date	Domestic Sales (in \$)	International Sales (in \$)	World Sales (in \$)	Genre	Movie Runtime	License
165	165 WALL-E (2008)	In the distant future, a small waste-collecting...	Walt Disney Studios Motion Pictures	June 26, 2008	223808164	297503726	521311890	['Adventure', 'Animation', 'Family', 'Sci-Fi']	1 hr 38 min	NaN
166	166 Big Hero 6 (2014)	A special bond develops between plus-sized infant...	Walt Disney Studios Motion Pictures	October 25, 2014	222527828	435341858	657869686	['Action', 'Adventure', 'Animation', 'Comedy', ...]	1 hr 42 min	PG
367	367 Terminator 3: Rise of the Machines (2003)	A machine from a post-apocalyptic future travels...	Warner Bros.	July 2, 2003	150371112	283000000	433371112	['Action', 'Sci-Fi']	1 hr 49 min	R

그림 7 영화 정보에 future 단어를 포함하는 영화 상위 3개

이후 영화에 대한 사람들의 선호도 파악을 위해 전체 SF 영화 dataframe 중 해당 3개의 키워드에 대해서 상위 3개 영화를 출력하였다.

4) LTSM기반 영화 리뷰 분석

3)에서 출력한 영화는 트랜스포머(2011), 그래비티, 월-e, 블랙팬서, 스타트랙(2009), 워 오브 더 월즈, 월-e(중복), 빅히어로(2014), 터미네이터(2003)으로, 중복을 제거하여 8개의 영화에 대해 네이버 리뷰를 수집하였다.

영화 리뷰에 대해 훈련된 모델을 통해 수집한 네이버 영화 리뷰에 대해 테스트하였으며, 테스트 정확도는 0.8313으로 높은 정확도를 확인하였다.

```
sentiment_predict('미래')
```

© 84.19% 확률로 긍정 😊 리뷰입니다.

```
sentiment_predict('우주')
```

© 53.04% 확률로 부정 🙄 리뷰입니다.

```
sentiment_predict('로봇')
```

© 76.23% 확률로 부정 🙄 리뷰입니다.

그림 8 3가지 영화 키워드에 대한 리뷰 감정 테스트

각 키워드에 대해 리뷰 감정 분석을 진행한 결과 '미래' 키워드에 대해서는 긍정적인 평가를 주었고 '우주' 키워드에 대해서는 53% 부정으로 거의 긍정 혹은 부정적 선호가 없음을 확인할 수 있다. 반면 '로봇'에 대해서는 대체로 부정적인 경향을 확인할 수 있다.

4. 결론

영화를 통해 확인한 결과 많은 사람들이 Sci-Fi 영화와 이에 활용된 다양한 과학기술에 관심을 가지고 있음을 확인할 수 있었다. 워드클라우드와 잠재의미 분석을 한 결과, 최근에는 '우주', '미래', '로봇'에 대해 많은 관심을 가지고 있었다. 또한 10년 전과 비교하였을 때, 과학기술에 대한 인식이 긍정적인 방향 혹은 새로운 발전으로 변화하였음을 확인할 수 있었다.

2022년 기준 해당 키워드가 들어간 영화의 리뷰를 분석한 결과, 미래에 대한 키워드에는 긍정적인 선호도를 갖고 우주에 대해선 확실한 선호도를 보이지 않았다. 반면 로봇 키워드에 대해서는 다양한 디스토피아적인 SF영화의 영향으로 부정적인 인식을 보이고 있다고 추측할 수 있었다.