



웹툰을 원작으로 한 OSMU 흥행 예측 분석

DBDBDeeP
5조 디비디비딤

DBDBDeeP 팀원 소개

웹툰을 원작으로 한 OSMU 흥행 예측 분석



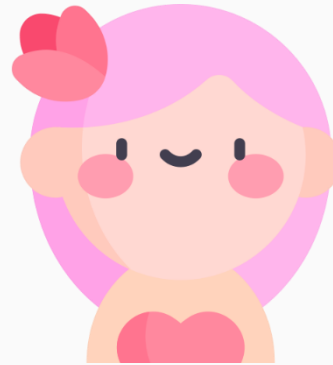
이정민



이태희



이한비



정혜리



조우현



홍사라

Contents

서론

1. 주제 선정 배경
2. 연구 목표
3. 기대 효과

본론

1. 분석 과정 개요
2. 데이터 수집
3. 데이터 전처리
4. 이미지 싱크로울
5. 모델 선정 및 평가

결론

1. 흥행 예측 분석
2. 예측 분석 결과
3. 시뮬레이션
4. 웹 구현

주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석

웹툰(Webtoon)이란?

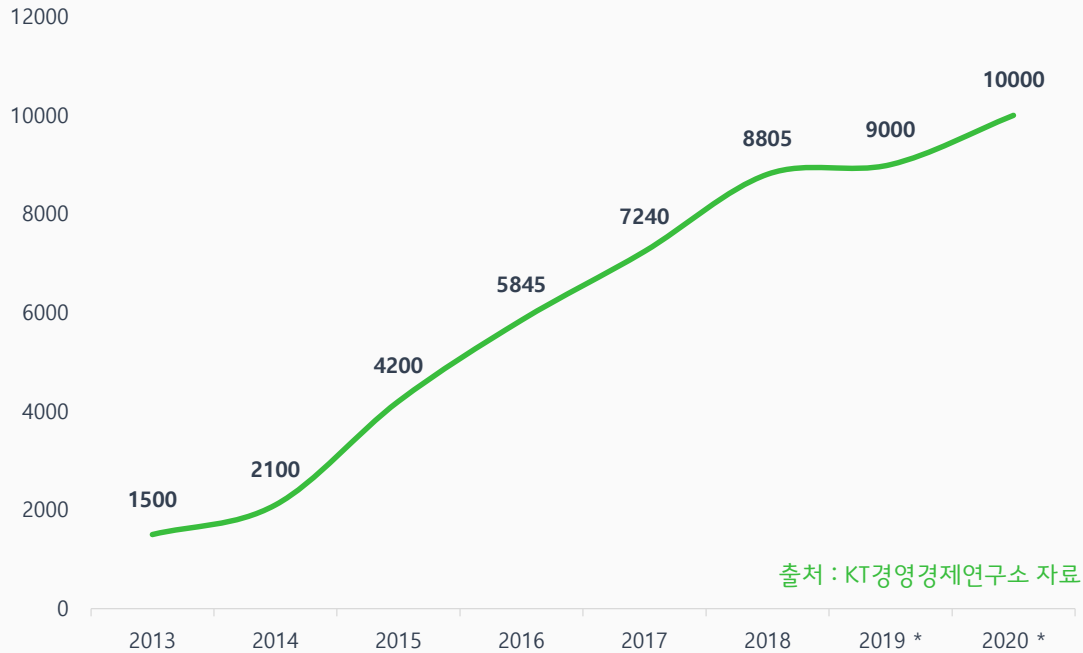
웹(Web) + 카툰(Cartoon)의 합성어,

언제 어디서나 짧은 시간에 즐길 수 있는 **스낵 컬처(Snack Culture)**



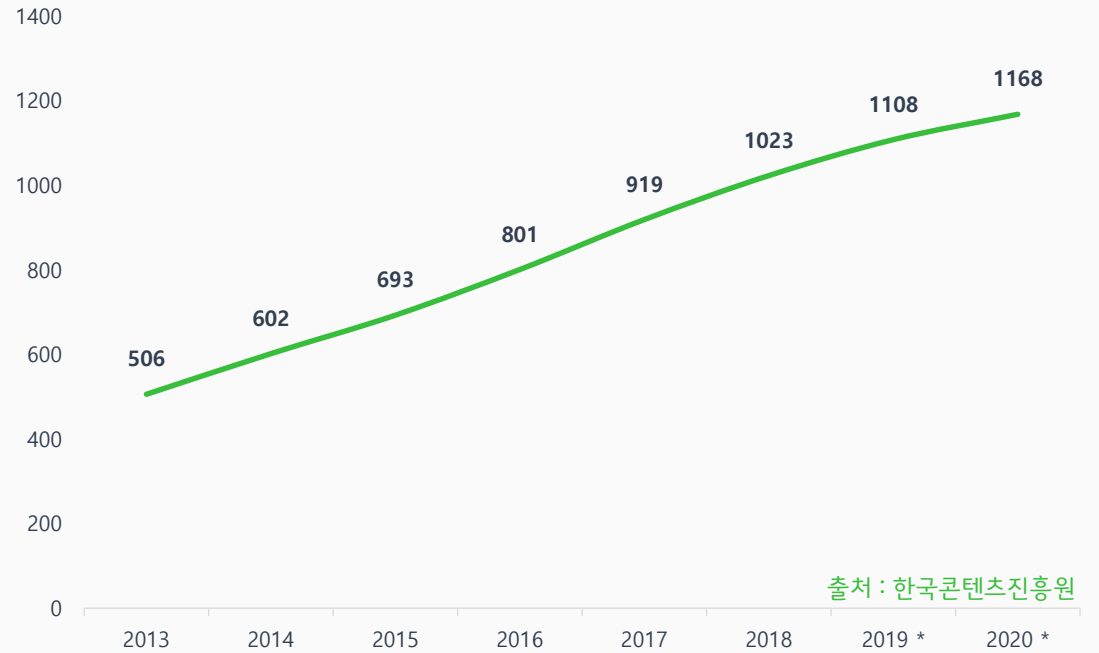
주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석



국내 웹툰 시장 규모 추이(단위: 억원)

* : 추정치, 2018년도 기준

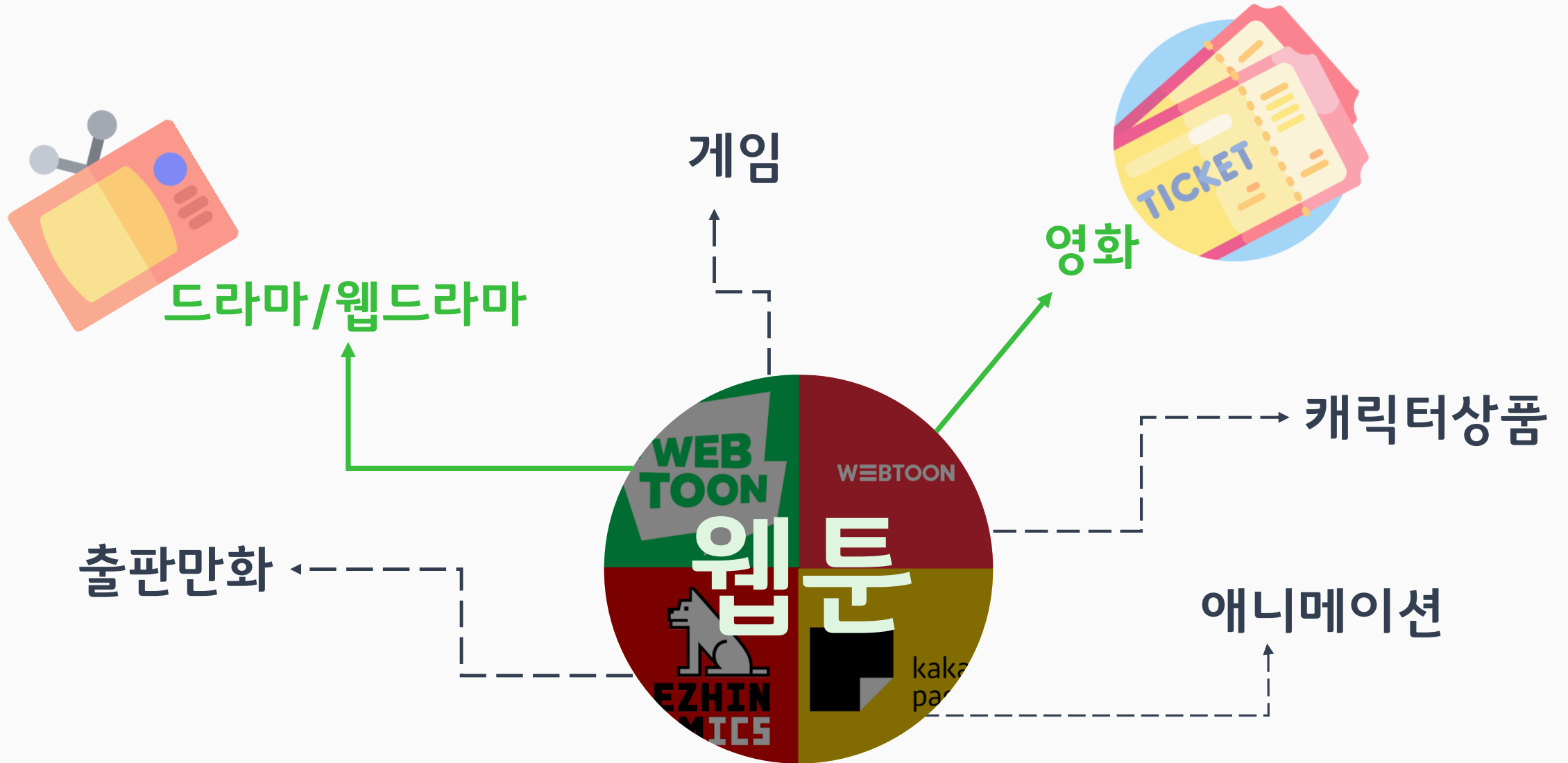


글로벌 웹툰 시장 규모 전망(단위: 만달러)

* : 추정치, 2018년도 기준

주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석



주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석

OSMU(one source multi use)란?

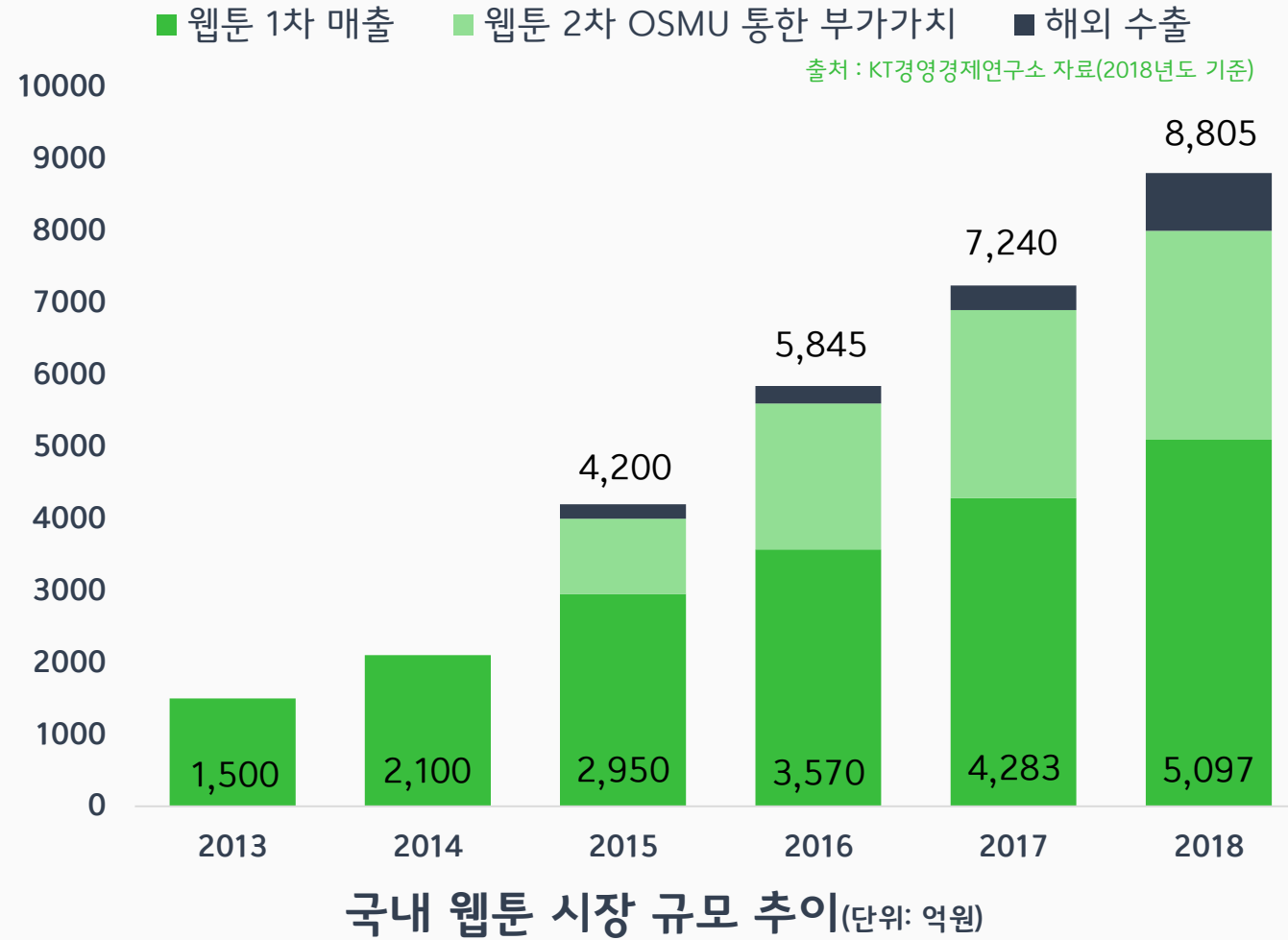
하나의 콘텐츠를 영화, 게임, 책 등 **다양한 방식**으로 개발하여 판매하는 전략

최소의 투자 비용으로 높은 부가가치를 얻을 수 있는 장점.

웹툰 서비스만으로는 산업을 성장시키는데 **한계**가 있기 때문에,
OSMU는 웹툰 사업 성장을 위한 **필수 전략**.

주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석



전체 웹툰 시장에서 OSMU 시장이 차지하는 비율이 빠른 속도로 증가

주제 선정 배경

웹툰을 원작으로 한 OSMU 흥행 예측 분석

"안 어울려" 웹툰 원작 드라마 싱크로율, 선택 아닌 필수? [ST이슈]

[스포츠투데이 김나연 기자] 웹툰을 원작으로 한 드라마의 강세가 지속되고 있는 가운데, 주인공의 캐스팅은 가장 큰 관심사다. 캐스팅이 알려지면 곧바로 웹툰 원작 역할과 배우와의 싱크로율이 예비 시청자들의 큰 관심을 받는다. 인물 간 '싱크로율'을 따지는 일은 선택이 아닌 필수가 된 셈이다.

'미생'을 시작으로 '치즈인더트랩', '김비서가 왜 그럴까', '내 아이는 강남미인', '타인은 지옥이다', '이태원 클라쓰' 등 많은 웹툰 원작 드라마들이 시청자들의 사랑을 받았다.

제작자들이 웹툰이라는 콘텐츠에 주목하는 이유는 인기 웹툰은 댓글이나 독자 반응을 통해 작품성과 대중성에 대한 검증이 이미 끝났고, 탄탄한 마니아층을 갖췄다는 점이다.

인기 웹툰이 원작인 드라마는 원작의 인기를 이어받을 수 있을 뿐만 아니라 해당 작품의 독자들을 시청자로 끌어모을 가능성이 있다는 점, 초기에 화제성을 잡고 시작할 수 있다는 것 또한 장점이다.

다만, 이 장점은 단점으로 작용하기도 한다. 기존 웹툰 팬들이 웹툰 원작에 대한 애정을 갖고 있는 만큼 원작 캐릭터와 싱크로율이 맞는 배우와 웹툰과 비슷한 스토리 전개, 배경 등을 원하기 때문에 그렇지 않을 경우 불만을 안고 갈 수밖에 없다.

제작자들도 이러한 반응을 알기 때문에 인기 웹툰이 영상화하는 과정에서 원작과 유사한 이미지의 배우들을 캐스팅하는 데 주력하고 있다. 인기 웹툰인 '치즈인더트랩'과 '내 아이는

강남미인'이 가상 캐스팅에서 1순위로 언급됐던 배우 박해진, 임수향을 캐스팅한 것만 봐도 그렇다.

웹툰 원작의 역할과 배우의 싱크로율을 높이는 것은 제작자가 고려해야 할 중요한 요소이지만, 흥행이나 작품성의 필수 요소는 아니라는 의견도 많다. 웹툰을 원작으로 하지만, 드라마는 또 다른 콘텐츠로 생각해야 한다는 것이다.

캐스팅이 알려지면 곧바로 원작 역할과 배우와의 싱크로율이 예비 시청자들의 큰 관심을 받는다. 인물 간 '싱크로율'을 따지는 일은 선택이 아닌 필수가 된 셈이다.

웹툰 원작의 역할과 싱크로율을 높이는 것은 제작자가 고려해야 할 중요한 요소이지만, 흥행이나 작품성의 필수 요소는 아니라는 의견도 많다. 웹툰을 원작으로 하지만, 드라마는 또 다른 콘텐츠로 생각해야 한다는 것이다.

- 기사 중 일부 발췌 -

싱크로율이 중요하다는 것은 알지만
그 반대 사례도 존재

싱크로율과 여러 변수간 관계를 통해
복합적으로 흥행에 미치는 영향과 그 관계를 파악하고자 함



목표

웹툰을 원작으로 한 OSMU 흥행 예측 분석



여신강림

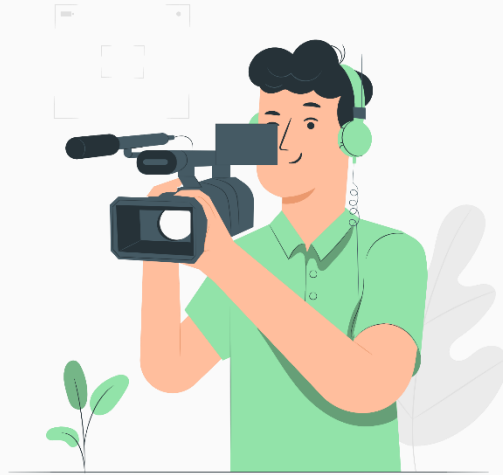


경이로운 소문

2020년 방영 예정작 '여신강림', '경이로운 소문' **흥행 여부 예측**

기대효과

웹툰을 원작으로 한 OSMU 흥행 예측 분석



영화/드라마 제작사

OSMU의 수익성
창출에 필요한
요소들 미리 파악



투자자

흥행 예측 분석 시스템을
통해 투자자본수익률
높일 수 있음

분석 과정

웹툰을 원작으로 한 OSMU 흥행 예측 분석

데이터 수집

데이터 전처리

모델 구축(학습)

이미지분석

시각화



WEB CRAWLING



Flask



Bootstrap

데이터 수집 기준

웹툰을 원작으로 한 OSMU 흥행 예측 분석

드라마 흥행 예측에 대한 선행 연구		
연구자	흥행 영향 요인	흥행 지표
배진아 (2005)	배우 파워 초기 시청률 채널	시청률
김경묵 (2009)	배우 파워 작가 파워 감독 파워 장르	
이관열 (2015)	배우 파워 장르 채널 드라마 유형 (일일/주말 등)	

웹드라마 흥행 예측에 대한 선행 연구		
연구자	흥행 영향 요인	흥행 지표
강여경 외 2명 (2018)	배우 파워 감독 파워 장르 댓글 수	누적 재생 수
홍성현 외 1명 (2017)	배우 파워 감독 파워 장르 제작사 파워 원작 여부	
김선기 외 1명 (2018)	배우 파워 감독 파워 제작사 파워 원작 유무	
오경석 (2014)	배우 파워 작가 파워 감독 파워	

영화 흥행 예측에 대한 선행 연구		
연구자	흥행 영향 요인	흥행 지표
조아름 (2019)	배우 파워 장르 제작사 파워 개봉 시기 국적	매출액
이찬복 (2013)	배우 파워 감독 파워 제작사 파워 스크린 수	손익분기점
박성복 (2019)	배우 파워 감독 파워 장르	매출액

배우 파워, 작가 파워, 감독 파워, 장르 4가지 변수는 흥행 예측에 관한 선행 연구들에서 **공통적으로 사용된 변수**.

추가적으로 웹툰 원작 2차 창작물 흥행 예측에 유의하다고 판단되는 변수들을 조사

데이터 수집

웹툰을 원작으로 한 OSMU 흥행 예측 분석

2020년 8월 18일 기준 OSMU 콘텐츠화 된 작품 목록 100개에 대한 데이터 수집

원작	웹툰	원작 웹툰 명
		웹툰 시작일
		원작플랫폼
		방영 당시 완결여부
		좋아요 수
		웹툰 회차
		웹툰 댓글
	웹툰 작가	작가 명
		작가 전작 수
		작가수상횟수
		작가 데뷔 년도

OSMU 작품	드라마	드라마, 웹드라마, 영화
		방송 기간
		제목
		방송사
		방송 회차
		시청률
		동시간대 방송사 평균 시청률
		웹드라마 평균 재생 수
	영화	장르
		방송 기간
		제목
		장르
		손익분기점
	감독	감독 명
		감독 데뷔 년도
		감독 전작 수
		감독 수상횟수
	배우(주연배우1)	배우 전작 수(드라마)
		배우 전작 수(영화)
		배우 수상횟수
		배우 사진
	배우(주연배우2)	배우 전작 수(드라마)
		배우 전작 수(영화)
		배우 수상횟수
		배우 사진

데이터 수집

웹툰을 원작으로 한 OSMU 흥행 예측 분석

웹 크롤링을 통해 OSMU 콘텐츠화 된 작품 100개에 대한 총 8,046개의 회차에
해당하는 **좋아요**와 **댓글**을 가져와 **JSON파일로 수집 및 저장**



JSON파일 저장

```
[{"CommentList": [{"webtoon_name": "저녁 같이 드실래요?",  
"comment_cnt": ["929",  
"623",  
"510",  
"611",  
"496",  
"318",  
"409",  
"430",  
"511",  
"682",  
"676",  
"538",  
"644",  
"621",  
"586",  
"614",  
"397",  
"784",  
"660",  
"721",  
"919",  
"578",  
"688",  
"737",  
"594",  
"716",  
"709",  
"467",  
"520",
```

```
["WebtoonList": [{"webtoon_name": "저녁 같이 드실래요?",  
"webtoon_num": "32",  
"like": ["20216",  
"13693",  
"11784",  
"12488",  
"11469",  
"10387",  
"11456",  
"10827",  
"11875",  
"10986",  
"11691",  
"11817",  
"11258",  
"12991",  
"12365",  
"12287",  
"11821",  
"13378",  
"11587",  
"12959",  
"14805",  
"13497",  
"12130",  
"13945",  
"13131",  
"14302",  
"14391",  
"12331",
```

데이터 전처리

웹툰을 원작으로 한 OSMU 흥행 예측 분석

흥행 여부 변수 생성

불필요한 데이터 삭제

드라마

방영한 모든 드라마의
평균 시청률을 기준으로
흥행여부 판단

웹드라마

방영한 플랫폼의 회차 별 평균
재생 수 100만 건을 기준으로
흥행여부 판단

영화

손익분기점을 기준으로
흥행여부 판단

변수
삭제

웹툰 작가 명
웹툰 시작일
웹툰 작가 데뷔 년도
웹툰 회차
OSMU 작품 제목
방송 년도
방송사
OSMU 회차
동시간 방송사 평균 시청률
OSMU 작품 시청률
웹드라마 재생 수
웹드라마 재생 평균
영화 손익분기점
감독 명
감독 데뷔 년도

데이터
삭제

네이버, 다음 이외의
기타 플랫폼에 해당하는
18개의 데이터

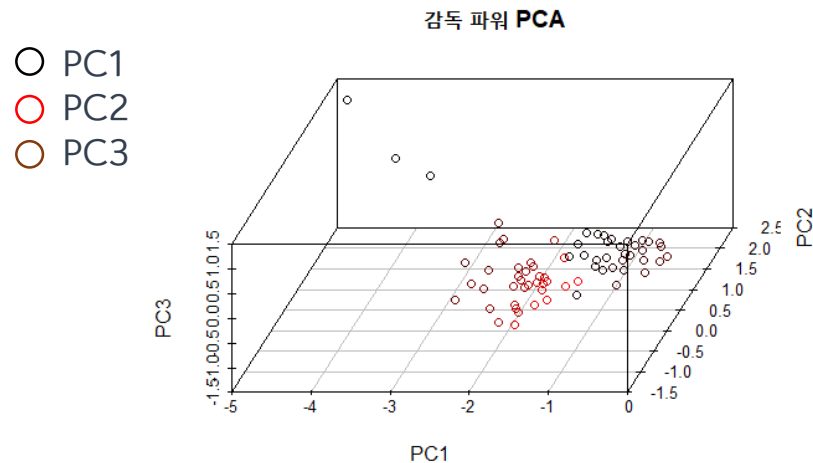
데이터 전처리

웹툰을 원작으로 한 OSMU 흥행 예측 분석

변수 축소 및 파워변수 생성

PCA 분석 사용하여
작가 파워, 배우 파워, 감독 파워 변수 생성

차원축소(Dimensionality Reduction)와 변수추출(Feature Extraction) 기법



감독 파워 = 감독 활동 연수 + 감독 수상 횟수 + 감독 전작 수

```
> summary(web_pca1) #높음
```

Importance of components:

	PC1	PC2	PC3
Standard deviation	1.5312	0.6701	0.45425
Proportion of Variance	0.7815	0.1497	0.06878
Cumulative Proportion	0.7815	0.9312	1.00000

PC1만 사용해도
변수 3개의 **78% 설명 가능**

데이터 전처리

웹툰을 원작으로 한 OSMU 흥행 예측 분석

변수 축소 및 파워변수 생성

웹툰 작가 전작 수

+

웹툰 작가 수상 횟수

+

웹툰 작가 활동 연수



작가 파워

감독 활동 연수

+

감독 전작 수

+

감독 수상 횟수



감독 파워

배우1 전작 수(방송)

+

배우1 전작 수(영화)

+

배우1 수상 횟수



배우1 파워

배우2 전작 수(방송)

+

배우2 전작 수(영화)

+

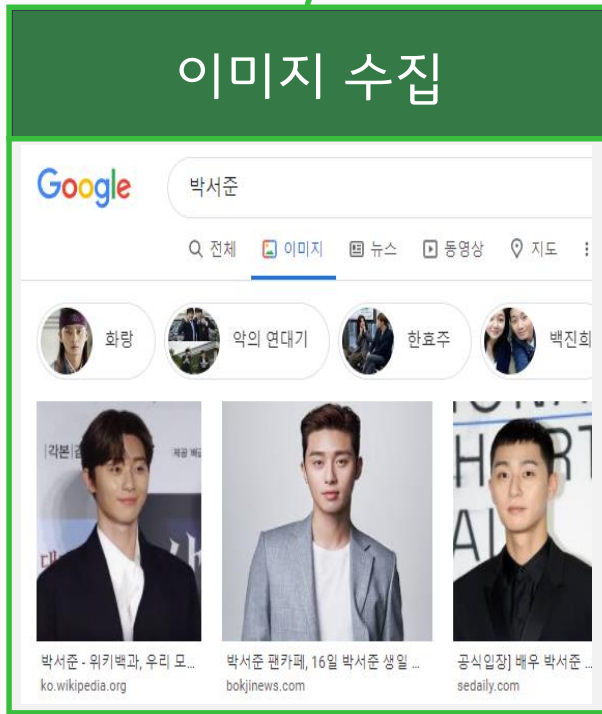
배우2 수상 횟수



배우2 파워

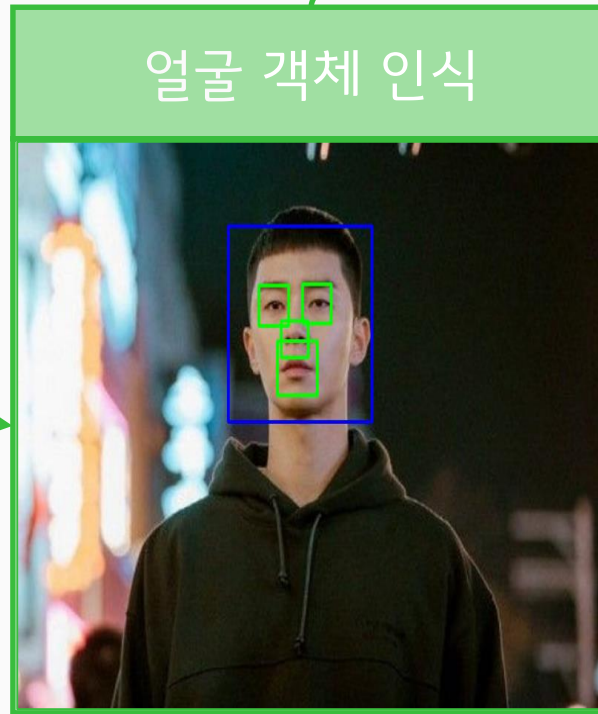
이미지 싱크로율 분석

웹툰을 원작으로 한 OSMU 흥행 예측 분석



배우 별 이미지 수집
각 110명의 남녀 배우

YES



cv2.CascadeClassifier 사용
배우 얼굴을 추출 및 저장

YES

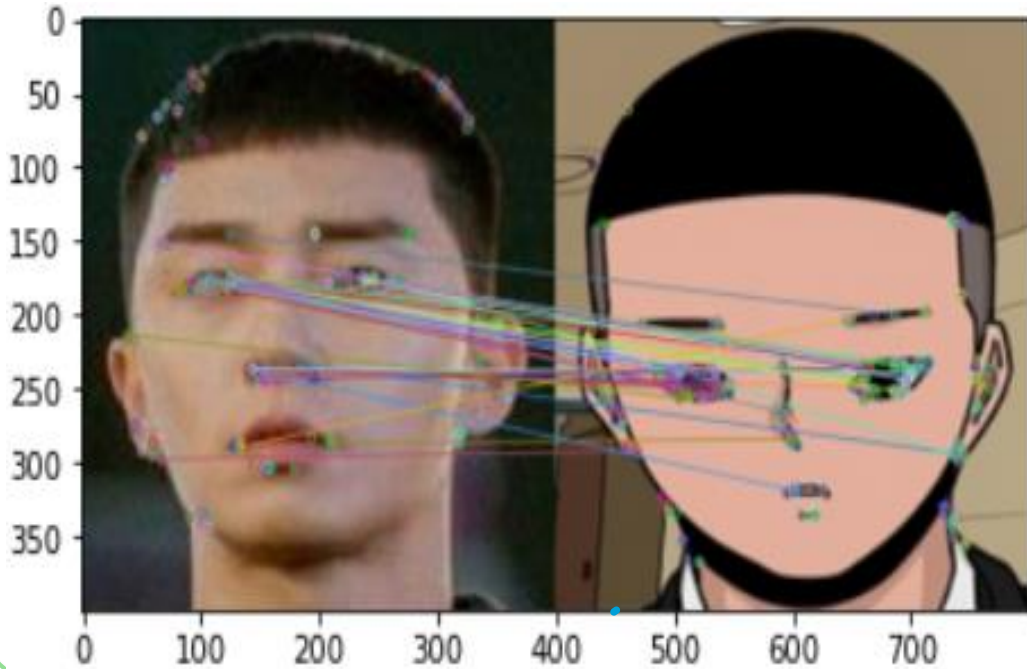


추출한 얼굴 사진
400x400형태의 jpg파일로 변경

이미지 싱크로율 분석

웹툰을 원작으로 한 OSMU 흥행 예측 분석

추출된 Feature들의 매핑 과정 **시각화**



이미지 유사도 측정과정

orb.detectAndCompute 사용

사진당 **100개의 피쳐**를 추출

scipy.spatial의 **distance**를 통해

각 사진들의 매핑된 Feature간의 **거리를 계산**하여 저장

이미지 싱크로율 분석

웹툰을 원작으로 한 OSMU 흥행 예측 분석



Original

박새로이



1st similar one

박서준

0.0157



2nd similar one

조병규

0.0181



3rd similar one

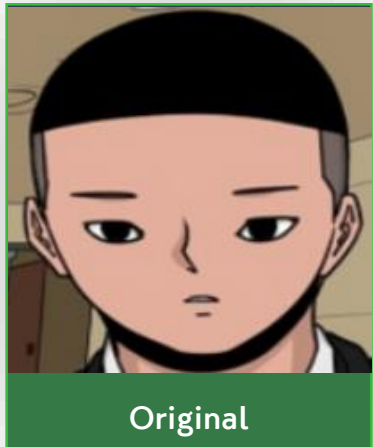
김래원

0.0327

111명의 남자배우들의 Feature distance를 캐릭터사진의 Feature와 가장 가까운 순서대로 정렬

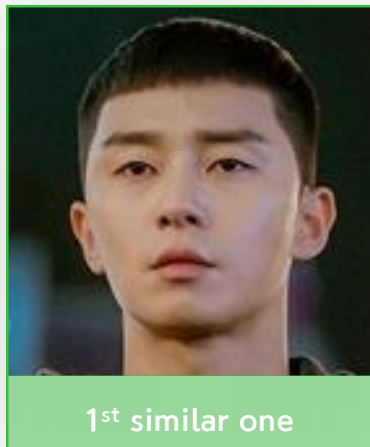
이미지 싱크로율 분석

웹툰을 원작으로 한 OSMU 흥행 예측 분석



Original

박새로이



1st similar one

박서준

0.0157



2nd similar one

조병규

0.0181



3rd similar one

김래원

0.0327

원작_웹툰	배우1_이미지_싱크로율	배우2_이미지_싱크로율	총행여부
저녁 같이 드실래요?	75	5	0
계약우정	24	56	0
어서와	64	81	0
쌍갑포차	20	98	1
이태원 클라쓰	1	87	1
메모리스트	102	65	1
어쩌다 발견한 7월	98	28	0
녹두전	92	89	1
동네변호사 조들호	25	46	1
쌈니다 천리마마트	81	60	0
타인은 지옥이다	8	109	0
빙의	39	92	0
좋아하면 울리는	101	96	1
죽어도 좋아	85	28	0
참치와 돌고래	33	5	0
내 ID는 강남미인!	74	59	1
탁구공	26	4	1
계룡선녀전	14	90	1
은주의 방	28	36	0
파라다이스	49	51	1
마음의 소리	84	58	1
한번 더 해요	26	17	1
세상 밖으로	104	6	0
멜로홀릭	34	72	0
부암동 복수자 소셜클럽	12	11	1

EX)

박서준 = 1 이므로,
112 - 1 = 111 사용

캐릭터를 연기했던 배우의 사진의 순위를 추출
이미지 변수 값 = 전체 목록 수 - 순위 값

데이터 전처리

웹툰을 원작으로 한 OSMU 흥행 예측 분석

수치형 데이터 처리

MinMaxScaling
을 통한 데이터 scaler 조정

범주형 데이터 처리

Dummy Variable
더미 변수 처리

변수	데이터형
원작 웹툰	수치형
웹툰 평균 좋아요 수	수치형
웹툰 평균 댓글 수	수치형
작가 파워	수치형
배우1 파워	수치형
배우2 파워	수치형
감독파워	수치형
배우1 이미지 싱크로율	수치형(순서)
배우2 이미지 싱크로율	수치형(순서)

변수	데이터형	변수 설명
웹툰 플랫폼	범주형	1. 네이버 2. 다음
완결여부	범주형	1. 완결 2. 연재 중(드라마 방영 당시 기준)
종류	범주형	1. 드라마 2. 영화 3. 웹드라마
범주화 된 장르	범주형	1. 코미디 2. 스릴러 3. 판타지 4. 로맨스 5. 드라마 6. 액션
OSMU 사전제작여부	범주형	1. 사전제작 2. 방영 동시 촬영

데이터 전처리

웹툰을 원작으로 한 OSMU 흥행 예측 분석

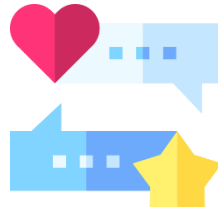
● 최종 사용 변수 개수 : 16개



웹툰 플랫폼



웹툰 평균 좋아요 수



웹툰 댓글



완결 여부



범주화장르



작가 파워



감독 파워



배우 파워 1, 2



이미지 싱크로율 1, 2



흥행 여부

분석 과정

웹툰을 원작으로 한 OSMU 흥행 예측 분석

K-Fold(교차검증)

$n = 8$  Test  Train

Model 1 

만든 데이터셋의 행개수가 적어 정확한 검정을 위해서 **K-Fold** 사용

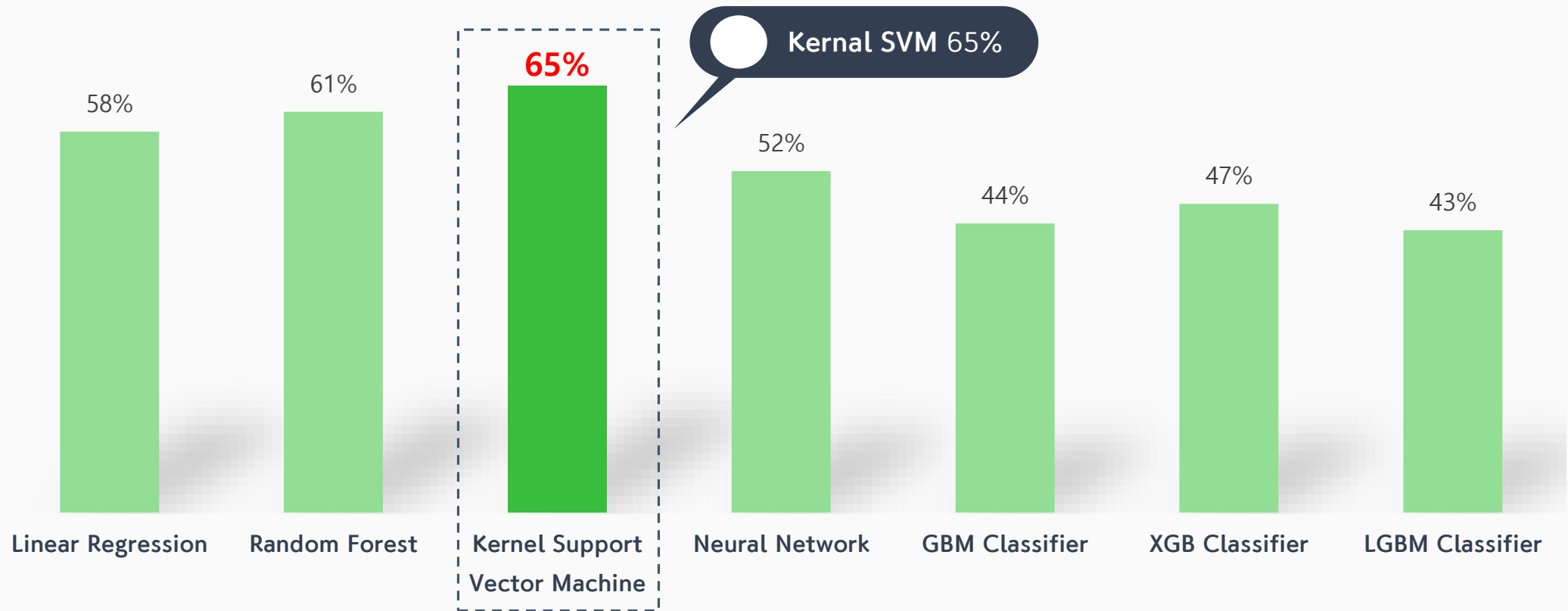
훈련, 검정 데이터를 **8:2**로 설정하였고 이에 맞춰 모든 데이터를 검정용으로 사용해보기 위해서 **n_fold=5**로 적용

모델 성능향상 목적이 아닌 성능평가의 일반화를 위한 과정: **StratifiedKFold** 사용하여 **Label**을 고르게 분포 시킴

모델 선정

웹툰을 원작으로 한 OSMU 흥행 예측 분석

분류 모델 별 정확도



모델 별 정확도

동일 데이터로 분석한 7개 분류 모델 중 정확도가 가장 높은 Kernal SVM 선정

모델 설명

웹툰을 원작으로 한 OSMU 흥행 예측 분석

SVM(Support Vector Machine)

두 그룹에서 각각의 데이터 간 거리를 측정하여 두 데이터 사이의 중심을 구한 뒤
그 가운데에서 최적의 초평면(Optimal Hyper Plane)을 구함으로써 그룹을 나누는 방법

장점

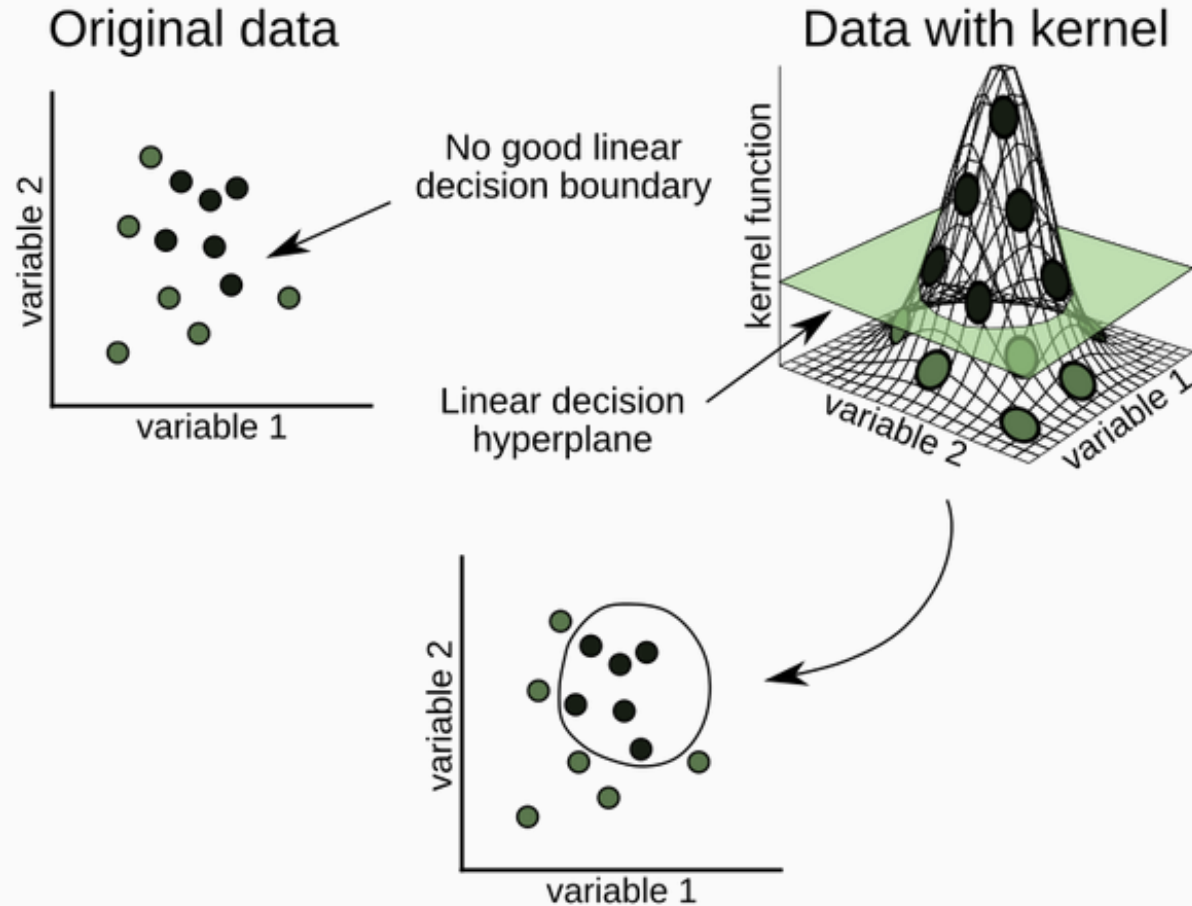
1. 저차원, 고차원 공간의 적은 데이터에 대해서 일반화 능력이 우수
2. 잡음에 강함
3. 데이터 특성이 적어도 높은 성능을 냄
4. 커널 함수 커스터마이징 가능
5. 신경망 기법에 비해 과적합 정도가 덜함

단점

1. 최적화된 모형을 위해서 커널과 모델 파라미터를 조절하기 위한 많은 테스트 필요
2. 데이터 수가 많을 수록 속도가 느려 지고 메모리적으로 힘들
3. 데이터 특성의 스케일링에 민감
4. 결과에 대한 설명력이 떨어짐

모델 설명

웹툰을 원작으로 한 OSMU 흥행 예측 분석



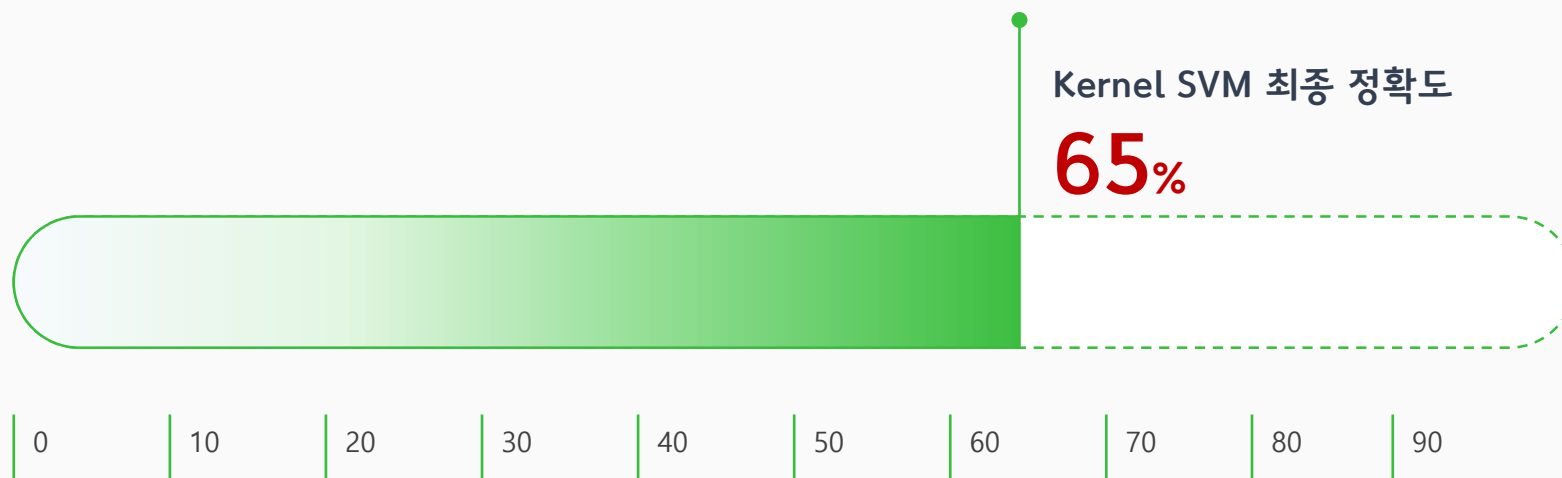
Kernel SVM

- 기존의 선형모델로 분류가 불가능 한 경우 데이터를 고차원특징으로 사상하여 해결하는 Kernel SVM 사용
- 두 매개변수 Cost와 Gamma에 따라 과소적합과 과대적합이 일어 날 수 있기에 적정 값을 찾아야함
- 웹툰 데이터의 행이 적다는 특성 때문

모델 평가

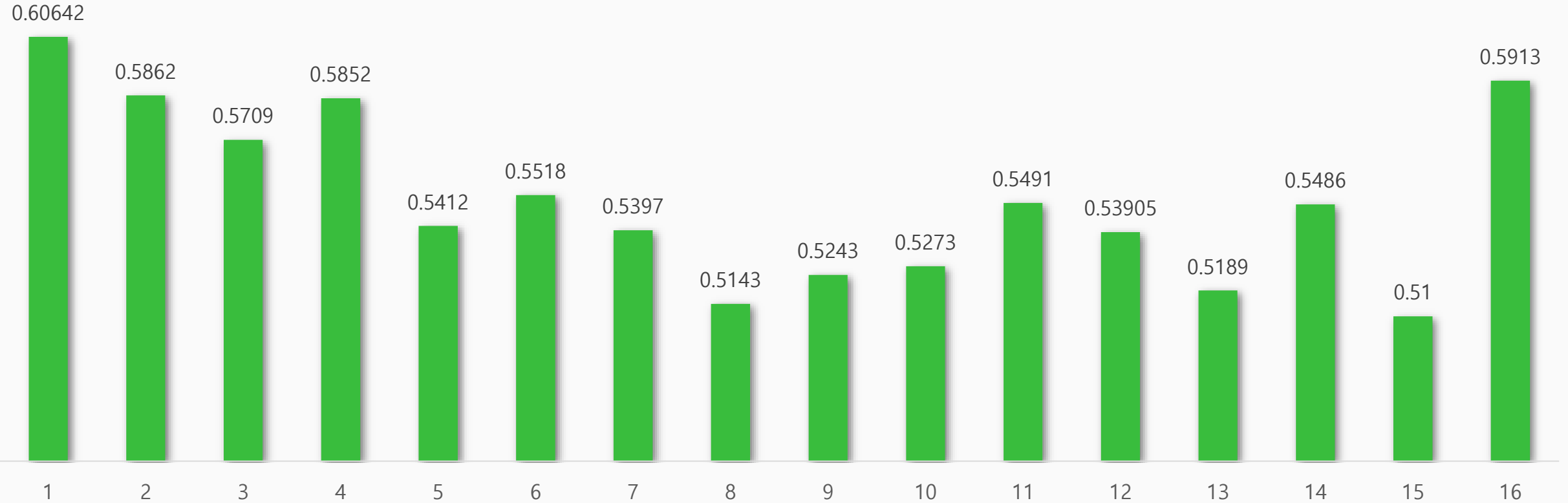
웹툰을 원작으로 한 OSMU 흥행 예측 분석

모델을 돌려보며 **최적의 파라미터 값**을 찾음으로써 최종 정확도 도출



모델 변수 중요도

웹툰을 원작으로 한 OSMU 흥행 예측 분석



완결 여부	웹툰 평균 좋아요 수	웹툰 플랫폼	배우1 파워	배우2 파워	배우1 싱크로율 점수	배우2 싱크로율 점수	장르 (코미디)	장르 (스릴러)	장르 (판타지)	장르 (로맨스)	장르 (드라마)	장르 (액션)	작가 파워	감독 파워	웹툰 댓글
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

흥행 예측 분석

웹툰을 원작으로 한 OSMU 흥행 예측 분석



선택된 모델에 2020년 제작 예정작인 ‘여신강림’, ‘경이로운 소문’ 데이터로 흥행 예측

웹툰을 원작으로 한 OSMU 흥행 예측 분석

[illegible]

Train Data

원작_웹툰	웹툰_플랫폼	웹툰_평균_좋아요수	완결여부	작가_파워	범주화된_장르	배우1_파워	배우2_파워	배1_싱점	배2_싱점	흥행여부
경이로운 소문	0	4928	0	-1.6	2	-0.64	-1.37	71	25	0
여신강림	1	48926	0	-0.62	4	-0.64	1.37	78	38	1

Validation Data

“경이로운 소문”, “여신 강림”

흥행 예측 분석 결과

웹툰을 원작으로 한 OSMU 흥행 예측 분석

흥행 예측 결과



흥행 성공



女神降臨
여신강림



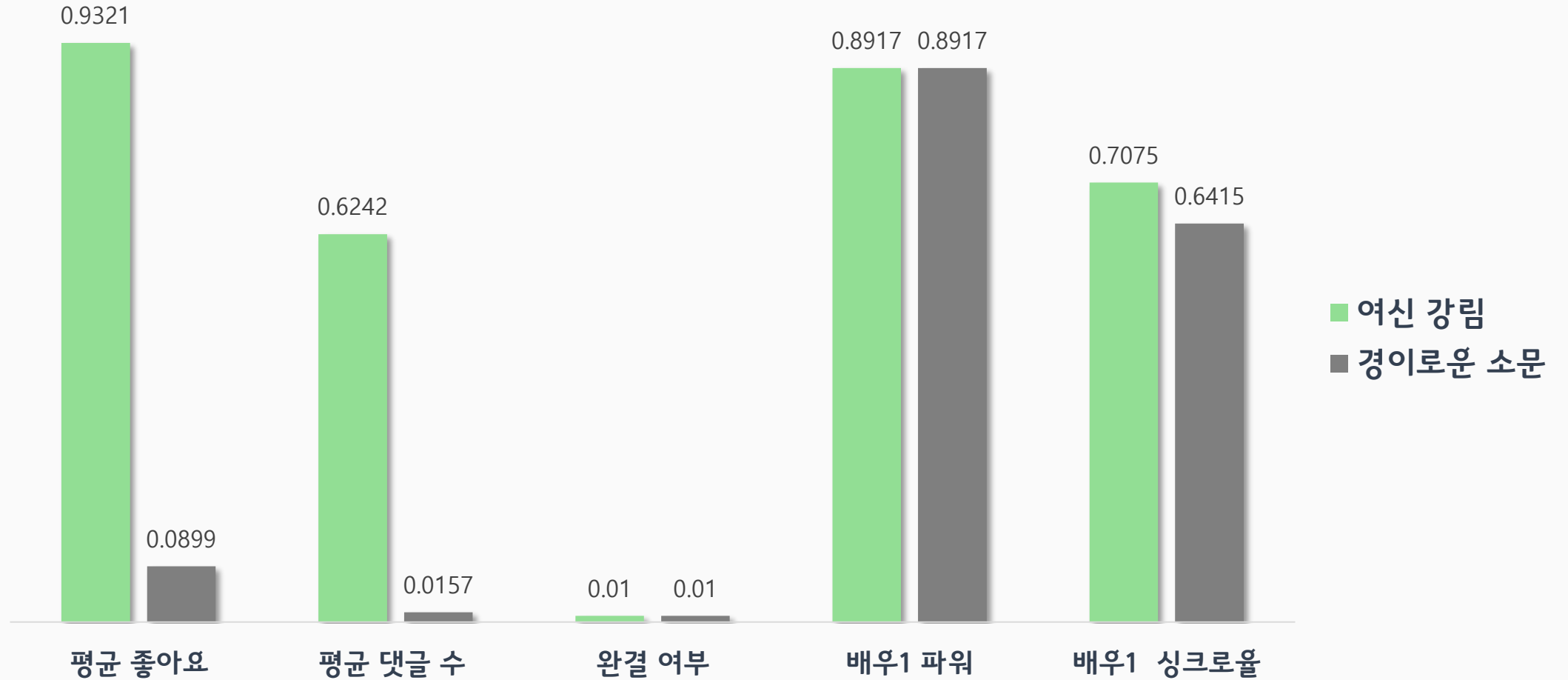
흥행 실패



경이로운
소문

결론

웹툰을 원작으로 한 OSMU 흥행 예측 분석



“여신강림”과 “경이로운 소문”의 변수중요도가 높은 변수 값들의 차이

시뮬레이션

웹툰을 원작으로 한 OSMU 흥행 예측 분석

웹툰	웹툰 좋아요 수	흥행여부
경이로운 소문	4958	0



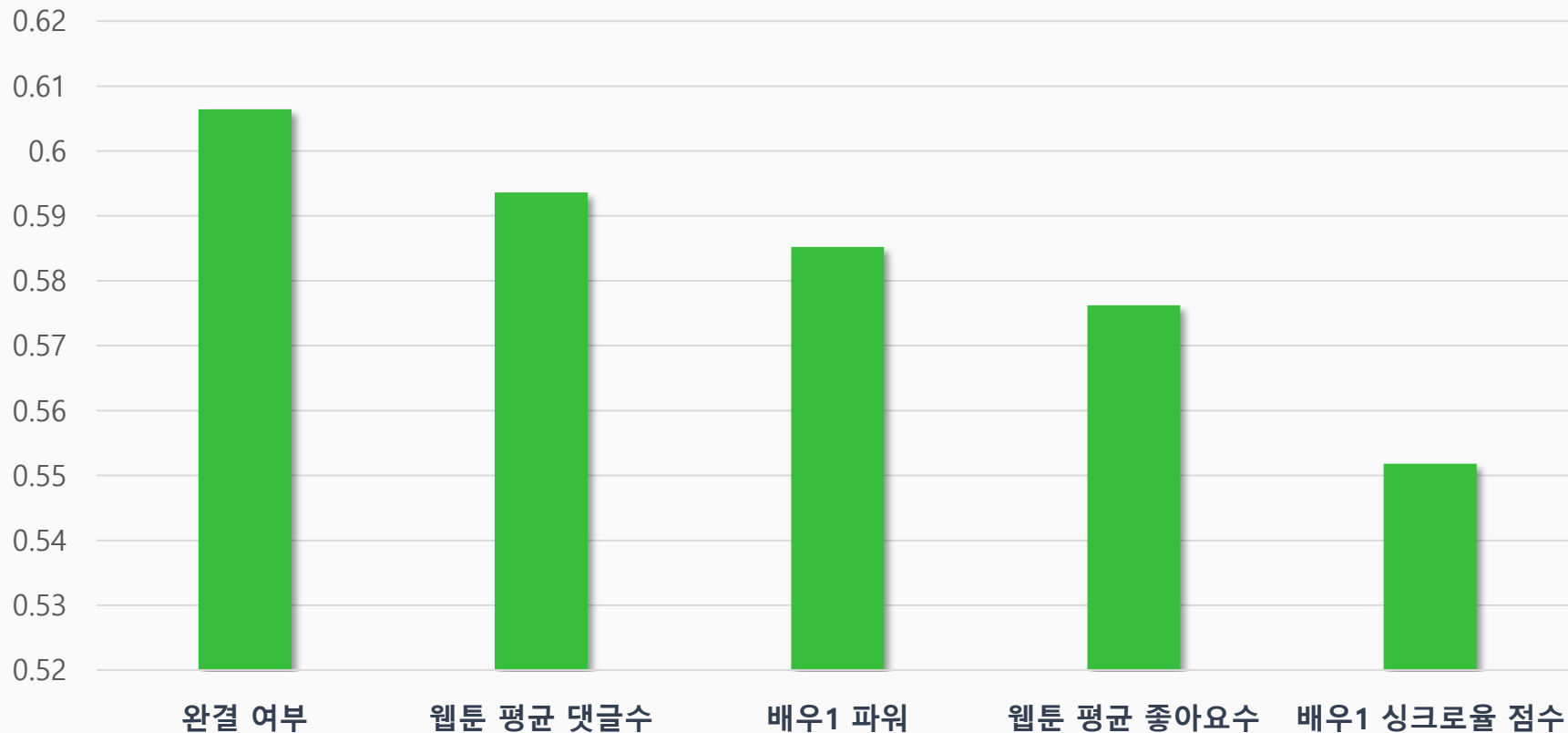
웹툰	웹툰 좋아요 수	흥행여부
경이로운 소문	30000	1

평균 좋아요수, 웹툰 댓글과 같은 **웹툰의 관심도** 및 **인지도**가 흥행여부에 중요한 요소임

제작 전 웹툰 관심도에 대한 조사가 중요할 것으로 보임(제작사, 투자자 등)

기대 효과

웹툰을 원작으로 한 OSMU 흥행 예측 분석



중요한 변수들은 웹툰 자체나 배우에 대한 변수가 높게 나옴

투자를 결정할 때 감독보다는 **웹툰** 자체나 **배우**들에 대한 고려를 더 하는 것이 투자에 좋은 영향을 끼칠 것이라 기대

웹 구현

웹툰을 원작으로 한 OSMU 흥행 예측 분석



Reference

웹툰을 원작으로 한 OSMU 흥행 예측 분석

참고 논문

1. 강여경, “웹 드라마의 흥행성과 결정요인에 관한 통합적 연구”, 중앙대학교, 2018.
2. 김경목, “대 흥행(blockbuster) 드라마의 조건”, 한국경영학회 통합학술발표논문집, 2008.
3. 김선기 “원작의 유무와 아이돌 출연이 웹드라마 흥행에 미치는 영향”. 한국콘텐츠학회논문지, 2018.
4. 박성복 “스타 배우들의 영화 출연 구조가 영화 흥행에 미치는 영향”, 동국대학교, 2019.
5. 이관열, “스타의 역량이 방송프로그램 성공에 미치는 영향”, 한국콘텐츠진흥원, 2015.
6. 이찬복, “한국 영화 흥행 데이터 마이닝에 관한 연구”, 상명대학교, 2013.
7. 조아름, “국내 개봉 흥행 영화의 속성에 따른 배우 캐스팅의 특성과 성과 연구”. 한양대학교, 2018.
8. 최은선, “지상파 TV 주간 드라마의 시청률 예측 요인에 관한 연구 : 인접효과, 경쟁구조, 스타파워, 드라마 장르를 중심으로”, 연세대학교, 2015.

참고문헌 및 기사

1. 고찬수, “웹드라마, 성공의 조건”, KBS N스크린기획팀, 2015.
2. 강여경 외 2명, “한국 웹 드라마의 흥행성과 결정요인에 관한 연구”, 전략경영연구, 2018.
3. 권호영, “드라마에서 스타 참여의 효과와 지속성 분석”, 한국콘텐츠진흥원, 2009.
4. 김나연, “'안 어울려' 웹툰 원작 드라마 싱크로율, 선택 아닌 필수?”, 스포츠투데이, 2020년 7월 11자.
5. 배진아, “드라마 시청률 영향요인 분석: 드라마 속성 및 수용자 요인을 중심으로,” 한국방송학보, 2005.
6. 홍성현 외 1명, “웹 드라마의 흥행에 영향을 미치는 요인에 대한 연구.”, 방송통신연구, 2017.
7. 김태영, “2019 만화 이용자 실태조사”, 한국콘텐츠진흥원, 2019.



감사합니다.

