

영화 흥행 결정 요인과 흥행 성과 예측 연구

김연형^{1,a}, 홍정한^a

^a전주대학교 통계학과

요약

영화의 흥행 결정 요인에 대한 학문적 연구와 함께 상업적 시각에서 개별 영화의 흥행 예측에 대한 관심이 증대되고 있다. 본 연구는 2010년 한국에서 개봉된 영화를 대상으로 영화 흥행에 영향을 미치는 요인들과 영화 흥행 성과간의 관계를 분석하였다. 제작 전 투자 의사결정단계에서 영화 장르, 관람등급, 감독, 배우가 통계적으로 유의한 결과를 보였으며, 배급편성의 의사결정단계에서는 배우효과, 스크린수, 배급사파워, 소셜미디어가 통계적으로 유의한 결과를 나타내고 있다. 선택확률개념을 이용한 다항로짓모형을 통해 영화 흥행작의 성과에 영향을 미치는 요인을 검증하였으며, 인공신경망, 판별분석과 비교하여 다항로짓모형의 흥행영화 예측력을 입증하였다.

주요용어: 영화 흥행 예측, 소셜미디어, 회귀모형, 다항로짓모형.

1. 서론

영화산업에서 제작회사와 투자회사, 배급사와 극장제인은 영화 제작 이전 시점에서 개별영화의 투자 결정과 제작 후 개봉 이전 시점에서 배급 및 스크린 편성 결정을 위해 영화 흥행을 예측하려는 많은 노력을 하고 있다. 극장 영화 개봉 기간은 매우 짧아, 짧은 기간에 투자비 이상의 수익을 올리기 위해 영화시장의 소비자 반응에 민감하게 대응해야 하기 때문이다.

현재 상업적 시각에서 개별 영화의 흥행 예측은 스코어카드에 의한 평가점수합계의 주관적 방식에 의존하고 있다. 개별 영화의 흥행 예측을 위해 객관적 데이터를 수집하고 예측모형에 적용하는 과학적 접근을 쉽게 발견하기 어려운 실정이다.

이러한 맥락에서 한국 영화산업에 대한 학문적 연구는 영화의 흥행결과에 영향을 미치는 요인들을 규명하는데 주목해 왔다. 대부분의 영화흥행 연구는 기존의 선행연구에서 사용되었던 요인체계를 바탕으로 변수들이 선택되었다. 그러나, 학문적 연구가 영화의 속성 이론에 관련된 개별 요인이 흥행에 미치는 영향력을 검증하는 데 치중함으로써, 상업적 시각에서의 다양한 요인들에 의한 개별 영화의 흥행 예측에 관한 통합적 연구는 활발히 진행되고 있지 않은 실정이다.

실증적 차원에서 진행된 연구를 살펴보면 김미현 등 (2010)은 개별 한국영화를 대상으로 흥행 성공 주요 요인을 찾고자 하는 연구를 진행하였고, 김소영 등 (2010)은 상업영화와 예술영화의 흥행성과에 유의한 예측 요인을 비교 분석하였다. 정완규 (2009)는 한국에서 개봉된 애니메이션 영화의 흥행결정 요인을 고찰하였고, 장병희 등 (2009)는 경제적 변인과 함께 심리적 변인이 함께 도입되었을 경우 영화 흥행예측 연구를 진행하였다.

2000년대 연구의 가장 큰 특징은 흥행결정요인으로 온라인구전의 영향력에 주목하기 시작했다는 점이다. 오상미 (2005)는 영화관련 인터넷 게시판의 메시지 수와 영화관객수간의 유의미한 상관관계를 나타내었고, 김희경 (2007)은 온라인 구전의 조작적 정의로 네이버 영화 디렉토리의 네티즌 리뷰수

¹ 교신저자: (560-759) 전북 전주시 완산구 효자동3가 1200, 전주대학교 통계학과, 교수. E-mail: yhkim@jj.ac.kr

를 사용하였으며, 정완규 (2009), 김미현 등 (2010), 김소영 등 (2010)은 포털 사이트 영화 디렉토리의 네티즌 평점을 영화 흥행 결정요인으로 확인하였다.

2010년 이후에는 서로의 의견, 생각, 경험, 관점 등을 공유하기 위한 온라인 플랫폼인 소셜미디어를 통해 상품, 브랜드 정보를 주고 받고 있다는 점에서 소셜 미디어 마케팅의 중요성이 부각되고 있다. 그러나, 영화 흥행 요인으로서 소셜미디어의 영향력에 대한 연구는 활발하지 않으며, 실제적으로 소셜미디어 자료를 이용한 실증적 연구는 진행되고 있지 않은 형편이다.

본 연구의 목적은 기존의 영화흥행요인들로 밝혀진 변수들을 통합적으로 분석하고 상업적 시각에서 영화 제작 전 시점과 개봉 전 시점에서 입수 가능한 정보를 기반으로 영화 수요를 어떻게 예측할 수 있는 지 실증 분석하는 것이다. 또한, 기존 연구에서 다루지 못했던 소셜미디어를 고려하여, 영화 흥행 결정 요인으로서의 중요도와 흥행 성과 예측에 있어서의 영향력을 검증하고자 하였다.

2. 연구 방법

2.1. 연구 문제

Eliashberg 등 (2006)은 영화산업의 가치사슬(value chain)에 있어 핵심단계를 제작 (production), 배급(distribution), 상영(exhibition)의 3단계로 구분하였다.

상업적 시각에서 제작, 배급, 상영 각 단계별로 영화흥행예측과 관련된 연구문제를 다음과 같이 정리하여 볼 수 있다.

- 제작 단계: 영화흥행예측모형이 초기지표(대본, 캐스팅, 예상상영등급)를 토대로 얼마나 정확하게 개발될 수 있을 것인가
- 배급 단계: 영화마케팅을 위한 예산을 다양한 미디어-전통적 미디어, 온라인 미디어-에 어떻게 배분하는 게 최적인가? 영화 흥행에 온라인 리뷰, 온라인 구전이 얼마나 영향을 미치는가?
- 상영 단계: 최적의 스크린 수를 결정하기 위해 어떤 요인을 고려해야 하는가?

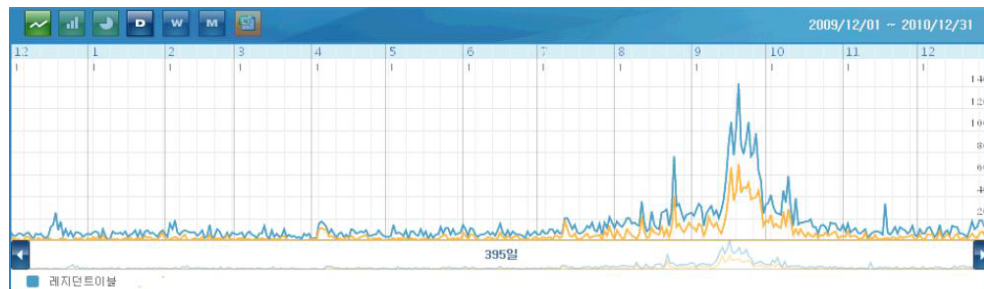
스크린 수는 할리우드 영화를 중심으로 한 연구에서 흥행에 중요한 변수로 제시되었다. 한국은 2000년대 이후 멀티플렉스 확산으로 전국 동시 개봉이 일반화되어 마케팅비용과 개봉 스크린 수가 관객 수에 영향을 미치는 요소로 간주되어 왔다. 인터넷이 중요한 커뮤니케이션 매체로 등장하면서 온라인 구전은 어떤 미디어보다도 강력한 구전커뮤니케이션 채널로서 소비자에게 위협을 회피할 수 있는 정보의 중요한 원천으로 작용하고 있다.

블로그가 국내에 소개된 것은 2003년이었지만 본격적으로 블로그가 활성화되기 시작한 때는 2005년 후반부터이다. 2010년부터는 블로그가 소셜미디어의 중요한 축으로 강력한 위력을 과시하고 있는 가운데, 페이스북, 트위터, 미투데이와 같은 다양한 형태의 소셜미디어가 빠른 속도로 확산되고 있다.

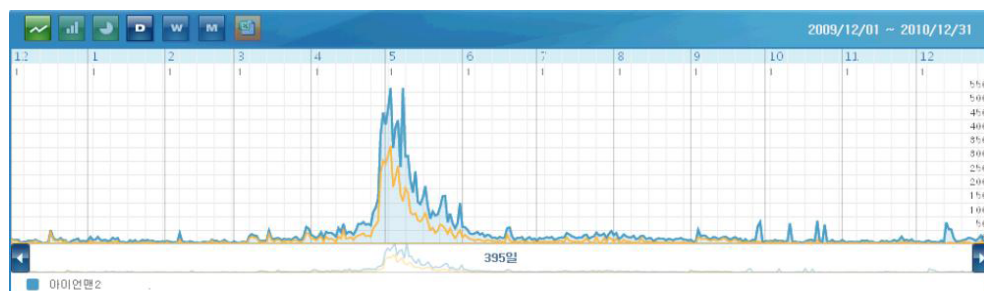
그 동안의 영화흥행연구에서는 네티즌이 포털에서 평가한 영화기대평점을 온라인 구전의 Proxy 변수로 사용하여 왔다. 영화 흥행 요인으로서 소셜미디어의 영향력에 대한 실증 연구가 아직 없었기 때문에, 블로그 문서 중 해당 영화에 대한 코멘트가 포함된 문서 수를 소셜미디어에 의한 온라인 구전의 Proxy 변수로 정의하여 모형에 포함시켰다.

본 연구는 영화 흥행 성과 예측 요인을 규명하고자 다음과 같이 연구 문제를 설정하였다.

- 연구 문제 1: 개봉 영화의 스크린 수는 영화 흥행성공에 영향이 있는가?
- 연구 문제 2: 포털 평점은 영화 흥행성공에 영향이 있는가?
- 연구 문제 3: 소셜미디어는 영화 흥행성공에 영향이 있는가?



(a) 레지던트아일랜드(개봉일: 2009년9월16일, 총관객수: 1,194,085 (37위), 총Buzz: 5,233 (35위))



(b) 아이언맨2(개봉일: 2009년4월29일, 총관객수: 4,452,072 (11위), 총Buzz: 15,043 (11위))

그림 1: 영화개봉과 소셜미디어(Online Buzz) 발생 추이

2.2. 변수의 정의

영화 흥행에 관한 변수들은 크게 영화의 내적 요인과 외적 요인으로 구분되고 있다. 영화의 외적 요인은 다시 구전커뮤니케이션 영역과 배급유통경쟁영역으로 나누어 진다.

영화 흥행에 대한 영화의 내적 요인의 영향을 분석하기 위해 해당 영화의 장르는 더미변수를, 관람 등급은 순서형변수를 이용하였다. 또한 영화 흥행에 기여하는 감독의 효과를 측정하기 위하여 감독이 해당 작품이 개봉되기 직전 3년 동안 감독한 영화가 동원한 관객 수의 자연로그 값을 이용하였다. 주연 배우의 효과를 보다 구체적으로 측정하기 위하여 주연배우가 해당 작품이 개봉되기 직전 3년 동안 주연으로 출연한 영화가 동원한 관객 수의 자연로그 값을 이용하였다.

영화 흥행에 영향을 미치는 영화 외적 요인으로는 배급사효과, 스크린 수, 포털 평점 그리고 소셜미디어 등을 이용하였다. 최근 영화 관객 동원에 있어 배급사의 영향력에 대한 관심이 높아지고 있는데, 영화 관람에 구조적으로 영향을 미치는 대표적인 요인은 스크린 수이며, 스크린 수 확보는 배급사의 파워와 어느 정도 연관이 있다. 구전효과를 영화 기대치를 평가한 네티즌 평점과 아울러, 블로그, 트위터 등 소셜미디어로부터 생성되는 온라인 버즈(Online Buzz)의 크기를 함께 이용하였다.

이론적으로 볼 때, 네티즌 평점, 소셜미디어와 같은 구전커뮤니케이션은 개봉 전 흥행예측의 사전 지표로서보다는 개봉이후 흥행의 지속 여부를 가늠하는 요인에 좀 더 가깝다고 할 수 있다. 현실적으로는 매우 짧은 개봉 기간 내에 투자비 이상의 수익을 올리기 위해서 개봉 전 소비자의 호기심을 유발 하면서도 흥미를 떨어뜨리지 않도록 전력투구하는 배급사의 입장에서, 영화 개봉 전 사전홍보수단으로서 온라인구전마케팅, 소셜미디어마케팅의 중요성이 점점 높아지고 있다.

소셜미디어의 예시는 그림 1과 같다. 그림 1의 첫 번째 그래프는 손익분기점 수준의 영화[레지던트아일랜드]의 경우로서, Buzz가 개봉일 3주 전후 기간 동안 좌우대칭의 분포(peak 412)를 보여주고 있다.

표 1: 변수의 정의

구분	변수	추정방법	변수형태	제작전 투자결정	개봉전 최적편성
독립 변수	영화 속성	감독	감독 작품 개봉 이전 3년 감독 영화 평균 동원관객 수	O	O
		배우	주연배우 작품 개봉 이전 3년 감독 영화 평균 동원관객 수	O	O
		국적	한국 미국 기타	더미변수	O
		장르	코미디 액션 스릴러 멜로 드라마 공포	더미변수	O
		관람등급	전체 12+ 15+ 18+	O	O
	구전 효과	포털평점	포털사이트 일반인의 영화기대평점 (10점 만점)		O
		소셜미디어	블로그 문서 중 검색어 문서 건수		O
	경쟁 요소	스크린 수	전국 개봉 스크린 수		O
		배급사	배급사 작품 개봉 이전 3년 배급 영화 평균 동원관객 수		O
	종속변수	관객수	영화진흥위원회 Box Office 누적관객 수		O

두 번째 그래프는 흥행성공영화[아이언맨2]의 경우로서, 개봉일에 다가올수록 급격히 상승한 Buzz (peak 1,762)가 개봉 후 1개월이 경과할 때까지 높은 수준으로 지속하였음을 볼 수 있다.

영화의 흥행을 나타내는 변수로는 영화진흥위원회 Box Office 전국 관객 수 통계를 이용하였으며, 다음 표 1에 각각의 변수의 정의를 설명하였다. 제작 전 투자결정 단계에서 고려할 수 있는 요인과 개봉 전 최적편성 단계에서 고려할 수 있는 요인들을 O로 분류하였다.

영화제작 단계에서 배급사나 창투사가 투자자의 형태로 영화제작에 관여하게 되는데, 투자단계에서 고려할 수 있는 요인은 국적, 대본의 완결성, 장르와 예상관람등급, 감독과 배우의 캐스팅 등 영화의 내적요인이다. 또한, 완성된 영화의 배급상영 단계에서 극장체인은 영화의 내적요인 이외에 구전커뮤니케이션, 배급사의 마케팅능력, 개봉 스크린 수 등과 같은 영화외적요인의 경쟁력을 함께 고려하여 최적 편성을 시도하기 때문이다.

2.3. 분석 방법

영화의 내적 요인 중 관람객 수에 영향을 미치는 독립변수는 앞 절에서 제시한 바와 같이 국적, 장르, 등급, 감독효과, 배우효과로 설정하여 <모형1>로 제시하였다. <모형1>은 영화 제작 전 투자 의사결정단계에서 영화 내적 속성만을 가지고 관객수를 예측하는 모형이라고 할 수 있다.

<모형2>는 <모형1>에 배급사효과, 스크린 수, 포털 평점 등을 변수로 추가하였다. 모형2>는 개봉 전 영화의 최적 편성을 위한 의사결정 단계에서 관객 수를 예측하는 모형이라 할 수 있다.

또한, <모형3>은 <모형2>에 소셜미디어(Online Buzz)를 변수로 추가하였다. 기존 모형에 최근 각광받고 있는 소셜미디어에 의한 온라인구전효과를 포함시킨 모형이다.

<모형1>: 누적관객수 = 국적 + 장르 + 등급 + 감독 + 배우

<모형2>: 누적관객수 = 국적 + 장르 + 등급 + 감독 + 배우 + 배급사 + 스크린수 + 포털평점

<모형3>: 누적관객수 = 국적 + 장르 + 등급 + 감독 + 배우 + 배급사 + 스크린수 + 포털평점 + 소셜
미디어

영화의 흥행 성공 결정요인을 도출하는 또 다른 접근으로 연속형자료인 종속변수를 이산형자료로 범주화하여 분석하는 방법이 있다. 종속변수가 범주형인 경우 인공신경망, 다항로지토형, 판별분석을

표 2: 주요 변수의 기술통계량

변수	N	Minimum	Maximum	Mean	Standard Deviation
관객수	316	3,434	8,173,375	470,409	994,691
감독	316	-	8,426,973	275,409	1,010,208
배우	316	-	11,707,650	623,555	1,609,059
배급사	316	-	9,438,537	329,883	1,210,954
스크린수	316	1	1,621	409	427
포털평점	308	5	10	8	1.088
소셜미디어	315	0	155,574	2863.54	10334.575

표 3: 2010년 한국영화 관람객 수 분포

그룹	기준	영화수	Percent	Cumulative Percent
1	50만 미만	245	77.5	77.5
2	50만~100만	26	8.2	85.8
3	100만~300만	35	11.1	96.8
4	300만~500만	6	1.9	98.7
5	500만 이상	4	1.3	100.0
Total		316	100.0	

이용하여 쉽게 모형화 할 수 있다. 인공신경망은 복잡한 문제에 대하여 좀 더 정확한 예측모형을 제공하나 모형의 직관적인 해석력이 떨어지고, 판별분석은 독립변수들의 정규성과 종속변수의 공분산행렬이 동일하다는 가정에 의존하는데 비해, 다항로짓모형은 회귀분석의 연속선 상에서 자료를 분석하면 서도 오차항에 대한 정규성이나 등분산성 가정을 두지 않는 장점이 있다. <모형3>을 다항로짓모형으로 바꾸면 다음과 같이 쓸 수 있다.

<모형4>:

$$\Pr(y_i = j) = \frac{\exp(X_i\beta_j)}{1 + \sum_j \exp(X_i\beta_j)}.$$

3. 실증분석

3.1. 기술 통계

본 절에서는 회귀분석을 수행하기에 앞서, 설명변수와 종속변수들에 대한 탐색적 기초통계분석 및 가공을 수행하였다. 통계적 모형은 자료에 대한 정규분포 가정에 기반한 추론이 이루어짐으로 로그변환을 시도하는 등 안정적인 모형이 도출될 수 있도록 하였다.

본 연구에서는 2010년 1월부터 2010년 12월까지 한국에서 개봉된 영화 316편의 영화를 분석에 이용하였다. 표 2는 본 연구에서 사용한 주요 변수에 대한 기술통계량이며, 분석대상에 이용된 316편의 평균 관객 수는 약 47만명으로 나타났다.

흥행에 성공한 영화의 요인별 특징을 제시하기 위해 관객 수를 기준으로 분석 구간을 설정하였다. 본 연구에서는 이경재와 장우진 (2006), 김미현 등 (2010)의 연구를 참고하여 100만 이상관객 수를 흥행에 성공한 한국영화의 최소 기준으로 보았다. 표 3을 보면 영화 관객 수의 편차가 매우 크다는 것을 알 수 있는데, 300만 이상의 관객을 동원하는 영화는 전체 개봉영화의 3.2%에 불과할 정도로 그 숫자가 매우 적은 것을 볼 수 있다.

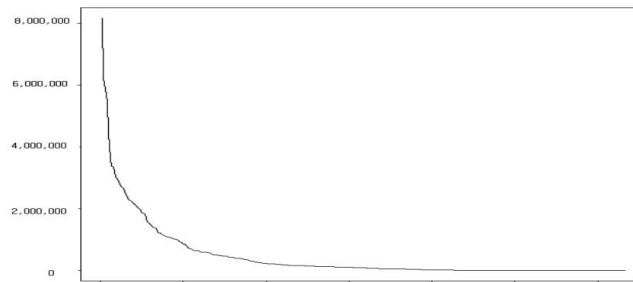


그림 2: 2010년 한국영화 관람객 수 분포(단위: 명)

표 4: 주요 변수 간 상관관계

		관객수	관람 등급	감독	배우	배급사	스크린수	포털평점	소셜미디어
관객수	상관계수	1	.104	.199	.399	.608	.954	-.023	.405
	유의확률(양측)	.	.084	.001	.000	.000	.000	.691	.000
	N	316	316	316	316	316	316	309	316
관람 등급	상관계수	.104	1	.122	.235	.091	.184	-.393	.021
	유의확률(양측)	.084	.	.029	.000	.149	.003	.000	.712
	N	316	316	316	316	316	316	309	316
감독	상관계수	.199	.122	1	.292	.258	.202	-.050	.191
	유의확률(양측)	.001	.029	.	.000	.000	.000	.379	.001
	N	316	316	316	316	316	316	309	316
배우	상관계수	.399	.235	.292	1	.348	.417	-.109	.179
	유의확률(양측)	.000	.000	.000	.	.000	.000	.057	.002
	N	316	316	316	316	316	316	309	316
배급사	상관계수	.608	.091	.258	.348	1	.630	-.110	.359
	유의확률(양측)	.000	.149	.000	.000	.	.000	.054	.000
	N	316	316	316	316	316	316	309	316
스크린 수	상관계수	.954	.184	.202	.417	.630	1	-.123	.355
	유의확률(양측)	.000	.003	.000	.000	.000	.	.031	.000
	N	316	316	316	316	316	316	309	316
포털 평점	상관계수	-.023	-.393	-.050	-.109	-.110	-.123	1	.052
	유의확률(양측)	.691	.000	.379	.057	.054	.031	.	.387
	N	309	309	309	309	309	309	309	309
소셜 미디어	상관계수	.405	.021	.191	.179	.359	.355	.052	1
	유의확률(양측)	.000	.712	.001	.002	.000	.000	.387	.
	N	316	316	316	316	316	316	309	316

3.2. 상관성의 검토

표 4는 더미변수를 제외한 주요 변수 간, 즉 종속변수와 모형에 포함된 독립변수들과의 스피어만 상관분석을 실시한 결과이다. 관객 수와 배우효과, 배급사파워, 스크린 수, 소셜미디어는 유의미한 상관관계를 보인 반면, 관객 수와 관람등급, 감독효과, 포털평점 간에 상관관계가 없는 것으로 나타났다.

주연배우가 해당 작품의 개봉 이전 3년 동안 주연한 영화의 평균 동원 관객 수를 나타내는 배우효과는 스크린 수와 상관관계가 있음을 알 수 있다. 스타배우의 캐스팅은 상대적으로 고 제작비 영화에서 이루어지는데, 제작비가 많이 든 블록버스터 영화일수록 파워가 큰 배급사를 통해 스크린 수를 많이 확보함으로써 투자금액을 조기 회수하려 하기 때문으로 해석할 수 있다. 그러나 배우효과는 네티즌 평점, 소셜미디어와는 유의한 상관관계를 보이지 않았다.

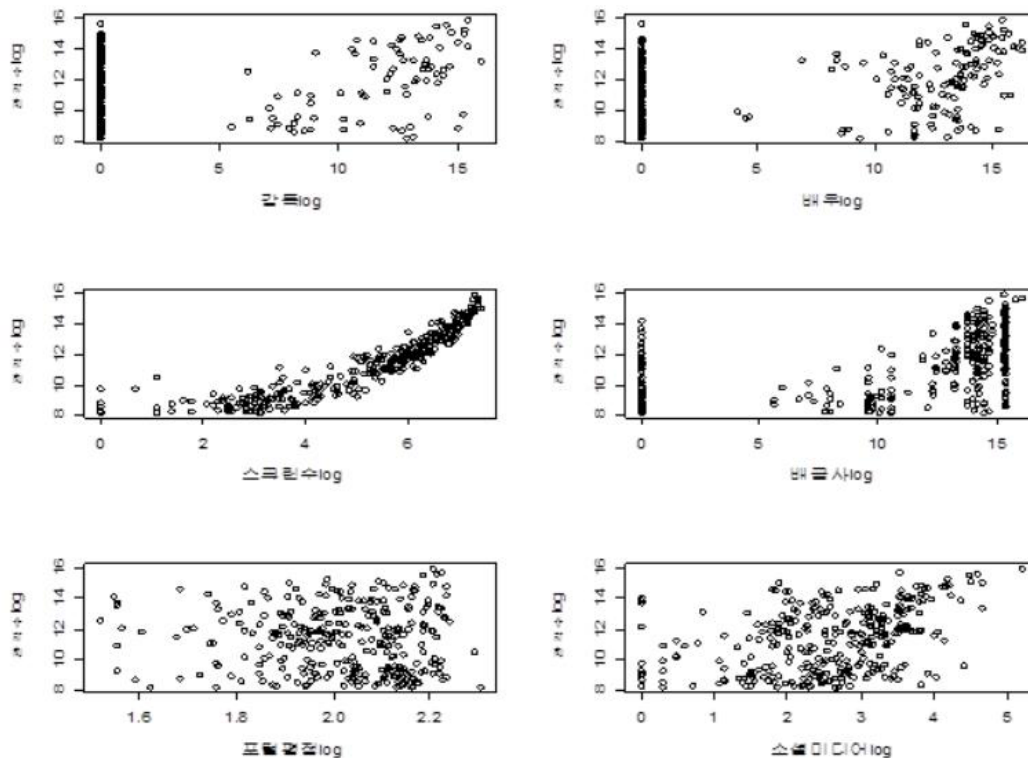


그림 3: 주요 변수 간 분포도

스크린 수는 배우효과, 배급사파워, 소셜미디어와 상관성이 있으며 소셜미디어는 배급사파워, 스크린수와 상관관계가 나타난다. 포털평점은 관람등급과 음(-)의 상관관계를 보이고 있는 반면, 나머지 변수와는 뚜렷한 상관관계를 보이지 않았다.

2차적으로 다중공선성을 판단하기 위해 VIF를 이용하였다. 변수들의 VIF가 1에서 10 미만의 값으로 다중공선성이 문제가 되지 않는 것으로 판단되었다.

그림 3은 각 변수와 관객수 간의 산점도이다. 감독효과, 배우효과, 배급사파워의 그래프에서 X축 좌표 0의 값에 몰려있는 것을 발견할 수 있다. 이는 영화 개봉 이전 3년 기간 평균 동원관객수가 많은 흥행작이 있는 감독과 배우라고 해서 항상 관객동원에 성공하는 것이 아니듯, 전체 개봉작 중 상당한 비중을 차지하는 무명 감독이나 신인 배우의 작품이라도 작품 여하에 따라 흥행에 성공할 수도 실패할 수도 있음을 보여주는 것이다. 또한, 이전 3년 기간 평균 동원관객수가 많은 배급사라고 해서 반드시 흥행에 성공하는 것이 아니고, 군소회사가 배급하더라도 작품 여하에 따라 관객동원에 성공할 수도 실패할 수도 있음을 보여준다.

3.3. 모형 간의 성과 비교

먼저 영화의 국적, 장르, 관람등급, 감독효과, 배우효과만을 변수로 분석한 <모형1>의 결과, 59.8%의 설명력을 보였으며, 국적을 제외하고 감독, 배우, 장르, 관람등급이 모두 유의미했다.

배급사파워, 스크린수, 포털평점을 포함한 <모형2>에서는 94.9%의 설명력이 있는 가운데 장르 중 액션, 스릴러, 드라마, 스크린수, 배급사파워는 흥행에 양(+)의 영향을 주는 것으로 나타났고 국적, 관

표 5: 2003~2008년 한국영화 흥행에 영향을 미치는 요인(회귀분석)

	모형1		모형2		모형3	
	Coefficient	Significance	Coefficient	Significance	Coefficient	Significance
(Constant)	4.311	.038	4.618	.015	3.855	.031
국적_한국	-.127	.900				
국적_미국	-.224	.810	-.083	.673	.015	.938
장르_코미디	4.283	.000	.625	.189	.580	.194
장르_액션	4.657	.000	1.037	.043	.755	.118
장르_스릴러	5.570	.000	.965	.088	.912	.086
장르_멜로	3.315	.001	.539	.184	.596	.120
장르_드라마	3.912	.000	.850	.086	.750	.114
관람등급	-.771	.023	-.082	.543	-.034	.797
감독효과	.260	.015	-.051	.394	-.066	.241
배우효과	.263	.013	.109	.065	.105	.057
스크린수			.897	.000	.909	.000
배급사파워			.091	.001	.078	.003
네이버영화평점			.499	.451	.508	.410
소셜미디어					.120	.028
N	316		316		316	
adj R ²	0.598		0.949		.958	
F	8.291		53.979		58.477	

표 6: 2010년 흥행 영화의 성공 요인 (다항로짓모형)

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	14.417	1	.000
국적	4.645	1	.031
장르	7.315	5	.198
관람등급	20.585	3	.000
감독	2.090	1	.148
배우	14.290	1	.000
배급사	31.350	1	.000
스크린수	7.761	1	.005
포털평점	2.049	1	.152
소셜미디어	6.286	1	.012

람등급, 감독효과, 포털평점은 영향이 없는 것으로 나타났다.

다음으로 소셜미디어를 포함한 <모형3>을 분석한 결과, 설명력은 95.8%로 영화 흥행에 있어서 배우효과, 스크린수, 배급사파워, 소셜미디어가 주요 요인임을 확인할 수 있다.

<모형3>의 설명력은 <모형2>에 비해 1% 정도 개선된 가운데 포털평점이 유의한 영향이 없는 반면, 소셜미디어가 영화 흥행에 유의미한 영향을 가지고 있는 것으로 나타났다.

포털평점은 영화를 본 네티즌들이 포털 영화사이트에 들어가 평점을 매기는 방식인데, 네이버영화를 비롯한 영화포털사이트의 네티즌 평점은 어느 정도 한계가 있다. 우선 평점 인플레이션이 지나치고, 영화 평점 참여 인원이 몇 백 명에 그치는 영화도 많기에 평점을 통해 영화를 판단하기에 힘든 면이 있다. 그 동안 여러 영화흥행 결정요인 연구에서 포털영화평점을 영화흥행 영향요인으로 설정하여 가설 검증하였으나, 소셜미디어가 본격 등장한 2010년 이후부터는 온라인구전효과와 proxy변수로서 포털영화평점 대신 소셜미디어를 사용하는 것이 바람직할 것으로 기대된다.

<모형4>는 회귀분석 모형인 <모형3>의 연속형 종속변수를 이산형으로 범주화하여 다항로짓모형

표 7: 인공신경망, 다항로짓, 판별분석 간 예측률 비교

Group	기준	TRUE	인공신경망	다항로짓	판별분석
1	50만 미만	237	98.7%	97.5%	97.0%
2	50만~100만	26	88.2%	65.4%	7.7%
3	100만~300만	35	84.2%	100.0%	57.1%
4	300만~500만	6	25.0%	100.0%	0.0%
5	500만 이상	4	33.3%	100.0%	75.0%
Overall		308	94.1%	95.1%	82.8%

으로 분석한 것이다.

표 6과 같이 다항로짓모형분석 결과, 국적, 관람등급, 배우효과, 배급사파워, 스크린수, 소셜미디어가 유의미한 요인인 것으로 나타났다. 이는 배우효과, 스크린수, 배급사파워, 소셜미디어가 주요 흥행 요인으로 나타난 <모형 3>의 결과와 대부분 일치한다.

다음으로 신규영화의 흥행예측을 위해 다항로짓모형을 어떻게 활용할 수 있는 지 살펴보았다. 상업적으로 영화 흥행을 예측한다는 것은 흥행에 성공할 영화를 예측하는 것이지, 흥행에 실패할 영화를 예측할 이유가 없으므로, 영화 흥행 성과 예측은 전반적인 예측률(overall hit ratio)보다는 흥행 성공 영화(4, 5그룹)의 예측력을 모형간에 비교하는 편이 합목적적이다.

표 7과 같이 다항로짓모형, 인공신경망모형, 판별분석 간 예측률을 비교하여 볼 때, 인공신경망과 다항로짓모형 모두 전반적 예측률이 우수한 것으로 나타났다. 그러나, 흥행 성공 영화(4, 5 그룹)의 예측력을 비교하여 볼 때, 다항로짓모형의 흥행영화 예측력이 더 우수하게 나타났다.

4. 결론

본 연구에서는 2010년 상영된 영화를 대상으로 영화 흥행 결정 요인을 파악하고 흥행 성과를 예측하였다. 선행 연구 결과와 비교하기 위해 회귀분석모형을 통해 분석하였으며, 기존 연구에서 고려하지 못한 소셜미디어를 활용하여 온라인구전의 영향력을 검증하였다.

영화 내적 요인으로는 감독, 배우, 관람등급, 영화 외적 요인으로는 스크린 수, 배급사파워, 소셜미디어 등이 영화 관객을 유인하는 요인으로 분석되었다. 영화 제작 전 투자의사결정단계에서는 영화 내적 요인인 장르, 관람등급, 감독, 배우가 일정 수준 영화흥행 영향요인으로 기여할 것으로 볼 수 있다.

영화 제작 후 상영 배급 편성의 의사결정단계에서는 영화 내적 요인 중 감독의 영향력은 떨어지는 대신, 배우효과와 함께 영화 외적 요인인 스크린수, 배급사파워, 소셜미디어가 영화흥행 결정요인으로 작용하고 있다.

선택확률개념을 이용한 다항로짓모형을 통해 영화 흥행작의 성과에 영향을 미치는 요인을 검증하였으며, 인공신경망모형, 판별분석과 비교하여 다항로짓모형의 흥행영화 예측력이 더 우수하게 나타났다.

국내 영화 산업에서 상업적인 영화흥행예측은 스코어카드에 의한 평가점수합계로 상영예정 영화의 흥행성적을 예측하고 있다. 스코어카드를 구성하는 항목들의 가중치가 주관적 경험에 근거한 배분으로 이루어지고 있는데, 다항로짓모형을 통해 도출한 흥행결정요인들의 상대적 영향력을 가중치로 하여 각 요인에 대한 상영 예정 영화의 평가점수를 가중 평균한다면 보다 합리적이고 객관적으로 흥행 성과를 예측할 수 있으리라 기대된다.

참고 문헌

김미현, 김수은, 최영준 (2010). 한국영화의 흥행 결정요인과 정책적 시사점, <영화연구>, 46, 31-55.

- 김소영, 임승희, 정예슬 (2010). 영화 유형별 영화 흥행 성과 예측 요인의 비교 연구, <한국콘텐츠학회 논문지>, **10**, 381-393.
- 김연형 (2010). 지역문화 축제의 서비스 품질과 고객충성도에 관한 연구, <한국데이터정보과학지>, **21**, 437-446.
- 김희경 (2007). <한국 영화 흥행 결정요인에 대한 연구>, 고려대학교 박사학위논문.
- 다음소프트 (www.daumsoft.com) 소셜메트릭스DB.
- 영화진흥위원회 Box Office (www.kobis.or.kr) 영화관입장권통합전산망DB.
- 오상미 (2005). <영화관련 구전커뮤니케이션 공간으로서 인터넷 게시판의 메시지 분석>, 연세대학교 석사학위논문.
- 이경재, 장우진 (2006). 베이지안 선택모형을 이용한 영화흥행 예측, <대한산업공학회/한국경영과학회 2006 춘계공동학술대회 논문집>, 1851-1856
- 인터넷 사이트 네이버 (www.naver.com) 영화정보수퍼DB.
- 장병희, 이양환, 김병선, 남상현 (2009). 심리적 변인 활용을 통한 영화흥행 예측의 정교화, <한국언론학보>, **53**, 346-371.
- 정완규 (2009). 애니메이션 영화의 흥행결정 요인에 관한 연구, <한국만화애니메이션학회>, 21-32.
- Eliashberg, J., Elberse, A. and Leenders, M. A. A. M. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions, *Marketing Science*, **25**, 638-661.
- Sharda, R. and Delen, D. (2006). Predicting box-office success of motion pictures with neural networks, *Expert Systems with Applications*, **30**, 243-254.

2011년 8월 접수; 2011년 10월 채택

A Study for the Development of Motion Picture Box-office Prediction Model

Yon Hyong Kim^{1,a}, Jeong Han Hong^a

^aDepartment of Statistics, Jeonju University

Abstract

Interest has increased in academic research regarding key factors that drive box-office success as well as the ability to predict the box-office success of a movie from a commercial perspective. This study analyzed the relationship between key success factors of a movie and box office records based on movies released in 2010 in Korea. At the pre-production investment decision-making stage, the movie genre, motion picture rating, director power, and actor power were statistically significant. At the stage of distribution decision-making process after movie production, among other factors, the influence of star actors, number of screens, power of distributors, and social media turned out to be statistically significant. We verified movie success factors through the application of a Multinomial Logit Model that used the concept of choice probabilities. The Multinomial Logit Model resulted in a higher level of accuracy in predicting box-office success compared to the Artificial Neural Network and Discriminant Analysis.

Keywords: Forecasting movie success, social media, regression model, Multinomial Logit Model.

¹ Corresponding author: Professor, Department of Statistics, Jeonju University, Jeonju-Si, Jeollabuk-Do 560-759, Korea.
E-mail: yhkim@jj.ac.kr

