

기계학습 기반의 영화흥행예측 방법 비교: 인공신경망과 의사결정나무를 중심으로

권신혜¹⁾, 박경우²⁾, 장병희³⁾

A Comparison of Predicting Movie Success between Artificial Neural Network and Decision Tree

Shin-Hye Kwon¹⁾, Kyung-Woo Park²⁾, Byeng-Hee Chang³⁾

요 약

본 연구는 영화산업의 가치사슬단계에 따라 각 단계에서 고려할 수 있는 변인을 활용하여 제작/투자, 배급, 상영단계별 모형을 구성하였다. 모형의 예측력을 높이기 위해 회귀분석으로 유의미한 변인을 도출하여 모형을 추가로 설정하였다. 주어진 변인을 바탕으로 기계학습 분석방법인 인공신경망과 의사결정나무 분석방법 간의 예측력 차이를 비교하였다. 분석 결과, 제작/투자 모형과 배급 모형에서 모든 변인을 투입했을 때는 인공신경망의 정확도가 의사결정나무보다 높았으나, 회귀분석결과에 따라 선정된 변인을 투입하였을 때는 의사결정나무의 정확도가 더 높았다. 상영 모형에서는 회귀분석결과와 반영여부와 관계없이 인공신경망의 정확도가 의사결정나무의 정확도보다 높게 나타났다. 본 논문은 영화흥행 예측연구에 기계학습기법을 적용하여 예측성고가 향상됨을 확인하였다는데 의의가 있다. 선형회귀분석 결과를 기계학습기법에 반영함으로써 기존의 선형적 분석방법의 한계를 극복하고자 하였다.

핵심어 : 영화흥행예측, 선형회귀분석, 기계학습, 인공신경망, 의사결정나무

Abstract

In this paper, we constructed the model of production/investment, distribution, and screening by using variables that can be considered at each stage according to the value chain stage of the movie industry. To increase the predictive power of the model, a regression analysis was used to derive meaningful variables. Based on the given variables, we compared the difference in predictive power between the artificial neural network, which is a machine learning analysis method, and the decision tree analysis

Received (February 7, 2017), Review Result (February 21, 2017)

Accepted (February 28, 2017), Published (April 30, 2017)

¹03063 Dept. Journalism and Mass Communication, Sungkyunkwan Univ., 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul, Korea

email: shinejessi@naver.com

²(Corresponding Author) 49236. Dept. Media and Communication, Dong-A Univ., 225, Gudeok-ro, Seo-gu, Busan, Korea

email: pekawe@dau.ac.kr

³03063 Dept. Journalism and Mass Communication, Sungkyunkwan Univ., 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul, Korea

email: mediaboy@skku.edu

* 이 논문은 동아대학교의 교내연구비 지원에 의해 작성됐습니다.

method. As a result, the accuracy of artificial neural network was higher than that of decision trees when all variables were added in production/ investment model and distribution model. However, decision trees were more accurate when selected variables were applied according to regression analysis results. In the screening model, the accuracy of the artificial neural network was higher than the accuracy of the decision tree regardless of whether the regression analysis result was reflected or not. This paper has an implication which we tried to improve the performance of movie prediction model by using machine learning analysis. In addition, we tried to overcome a limitation of linear approach by reflecting the results of regression analysis to ANN and decision tree model.

Keywords : predicting movie success, linear regression analysis, machine learning, artificial neural network, decision tree

1. 서론

인공신경망(Artificial Neural Network; ANN)은 주로 공학 및 주식시장 등의 다양한 산업에 적용되어왔다[1][2]. 인간두뇌의 추론능력과 유사한 과정인 '일반화(generalization)'과정을 활용하여, 주어진 자료를 학습하여 새로운 자료에 대한 예측값을 제시하는 과정으로 구성된다[3]. 의사결정나무(Decision Tree)는 의사결정규칙(decision rule)을 나무구조로 표현하여 분석대상을 일정한 집단(classification)으로 구분하거나 예측(prediction)하는 분석방법이다[4]. 인공신경망과 의사결정나무는 비선형 문제를 해결할 수 있는 대표적인 분석방법으로서 좋은 예측력을 얻을 수 있다는 장점을 가지고 있다[5]. 영화의 성과를 예측하기 위해 활용되어온 선형회귀분석의 방법론적 한계를 보완하기 위해 인공신경망과 의사결정나무 분석방법을 도입할 수 있다.

본 연구의 분석대상인 영화의 성과예측은 산업 전반에 걸친 의사결정단계와 밀접한 연관이 있다. 일반적으로 영화산업의 가치사슬은 제작, 배급, 상영 단계로 구분되어 있으며[6], 각 단계에서는 영화의 성과를 예측하기 위해 고려할 수 있는 요인이 다르다. 예를 들면, 제작 단계에서는 스토리와 감독 및 배우, 배급 단계에서는 상영관 수와 스크린 수, 상영 단계에서는 전문가 평가, WOM(Word Of Mouth) 등의 요소들에 대한 의사결정이 이루어진다. 모형의 개발 시에 이를 반영하는 것은 당연하다. 영화홍행예측연구의 결과를 실무적으로 활용할 여지가 많음에도 불구하고, 의사결정자들은 정성적이고 경험적인 평가에 의존하고 있는 실정이다. 학술적 연구의 결과가 실무적으로 활용도가 낮은 원인은 영화홍행예측연구가 가치사슬단계를 고려하지 않고 포괄적으로 각 변인의 영향력 검증 위주로 연구되었기 때문이다[7].

본 연구의 목적은 첫째, 영화홍행예측 분야에 기계학습 방법론의 하나인 인공신경망과 의사결정나무 분석방식을 새롭게 적용하여 방법론적 확장을 시도하는 것이다. 둘째, 영화산업단계에 따라 영화홍행요인을 달리 적용하여 연구결과의 활용도를 높이는 것이다.

2. 이론적 논의

2.1. 방법론적 측면에서의 영화홍행 예측연구

영화홍행 예측연구는 경제학적 접근방식에서부터 수용자 조사, 새로운 방법론의 적용에 이르기까지 다양한 관점에서 연구되어 왔다. 영화홍행예측과 관련된 기존 연구들은 산업단계별 모형 연구의 부재, 통계학적 한계 등의 단점을 가지고 있다.

첫째, 전통적인 영화홍행 예측연구는 선형회귀분석을 적용한 경제학적 접근방식(economic approach)과 수용자를 분석대상으로 하는 심리적 접근(psychological approach)으로 분류된다. 경제학적 접근방식은 산업 데이터를 이용하며, 영화상품의 제작과 관련된 속성과 배급과 관련된 속성과 같은 영화산업의 여러 영역에 속하는 영향변인을 제시해 왔다. 변인영역은 창작영역, 배급영역, 마케팅과 홍보영역으로 나누어진다. 창작영역에는 장르, 스타파워, 감독파워, 제작비, 속편 등이, 배급영역에는 배급사 파워, 개봉시기 등이, 마케팅영역에는 홍보, 광고, 비평가 평가, 수용자 평가 등이 속한다[8]. 이러한 접근방식은 개별영화 상품에 대한 데이터를 활용하기 때문에 개별 영화가 가진 특성을 반영한다는 의미를 가진다. 각 영역에 속하는 변인들은 영화가 제작되고 배급되는 과정에 따라 중요도가 다를 수 있다. 예를 들어, 상영단계에서 수용자들이 영화를 선택할 때, 영화가 어떤 장르인가보다 영화에 대한 타인의 평가가 더 중요하게 작용할 수 있다. 따라서 변인의 적용시기에 따라 영향력을 분석할 필요성이 제기된다.

둘째, 방법론적 관점에서 기존 영화홍행연구를 살펴보면, 선형회귀분석은 종속변인과 독립변인 간의 비선형 관계가 있을 경우 모델의 성능이 떨어질 수 있다. 이를 해결하기 위해 로그-선형 형태의 연립방정식, 유한 혼합 모형(Finite Mixture Model)을 활용한 군집분석과 판별분석, 인공신경망(ANN; Artificial Neural Network), 의사결정나무(Decision Tree) 등의 방법론을 활용하는 시도들이 있다[5][9]. 이러한 시도들은 전통적인 통계모형과는 달리 변수들의 분포 가정에 대해 자유롭다. 다양한 정보들을 반영할 수 있기 때문에 예측력을 높일 수 있으며, 연구모형의 강건성(robustness) 측면에서 장점을 가진다. 인공신경망은 독립변수와 종속변수 간의 관계를 정의하기 까다로운 데이터에 대해 상대적으로 좋은 결과를 낼 수 있다는 장점이 있다. 의사결정나무는 데이터를 계층적으로 분류하며, 각 계층마다 하나 이상의 분류 기준을 적용한다. 가장 큰 영향을 미치는 요소부터 적용하여 분류되며, 각 중간 분류 결과를 다음의 적합한 속성으로 다시 분류해 나간다. 본 연구에서는 새로운 방법론으로서 인공신경망과 의사결정나무를 영화 성과의 예측에 적용함으로써 예측력을 높여 영화홍행예측연구 분야의 방법론적 확장에 기여하고자 한다.

2.2. 영화산업

영화산업은 영화상품이 경험재이기 때문에 성공에 대한 불확실성이 높다[10]. 각 산업단계에서 의사결정자들이 객관적으로 참고할 수 있는 정보가 적다는 것을 의미하기도 한다. 예를 들어, 투자자 입장에서는 투자여부를 결정하는 시기가 제작 전 단계이기 때문에 제작이 완결되기 전 시놉시스와 트리트먼트 수준의 스토리라인, 기획한 감독, 출연예정인 배우 등과 같이 한정된 정보를 바탕으로 결정하게 된다. 이러한 한계점을 반영하여 주차별로 개봉 전과 개봉 후 주차별 관객 수의 변화와 그에 대한 흥행요소를 밝혀내는 연구도 있으나[11], 여전히 투자 단계에서 고려할 수 있는 부분은 변인에서 제외되었거나 우회적 측정이라는 한계를 가진다. 따라서 영화산업의 가치사슬의 각 단계에 따라 변인의 영향력이 다를 수 있으며, 영화흥행 예측 모형에 반영할 필요가 있다.

3. 연구문제

영화산업의 가치사슬단계는 제작 및 투자, 배급, 상영의 세 단계로 나눌 수 있으며, 각 단계에 따라 다른 변인으로 모형을 구성할 수 있다. [표 1]은 제작/투자, 배급, 상영 모형별로 고려할 수 있는 변인을 보여준다. 상영모형은 영화가 개봉된 후를 의미하며, 개별 영화의 첫 주 전국 주말관객 수를 독립변인으로 투입한다. 영화의 성과를 예측하는 시기에 따라 변인의 영향력이 달라질 수 있다는 가정을 바탕으로 본 연구에서는 모형별 선형회귀분석을 실시하여 영향력을 미치는 변인을 살펴본다. 모형별로 영향을 미치는 변인에 차이가 발생한다면, 이를 반영한 모형은 선형회귀분석 결과의 반영여부와 분석기법에 따라 예측력에서 차이가 발생할 수 있다. 이를 검증하기 위해 다음과 같이 연구문제 1과 2를 제시한다.

연구문제 1: 영화산업단계별 모형의 선형회귀분석 결과, 영향을 미치는 변인에 차이가 있는가?

연구문제 2: 선형회귀분석결과를 반영한 인공지능망모형과 의사결정나무모형의 예측력은 어떠한가?

[표 1] 영화산업단계별 영화흥행 영향요인

[Table. 1] The factors affecting movie success in each stage of movie industry value-chain

변인군	세부변인	조작적 정의	제작/투자 (모형 1)	배급 (모형 2)	상영 (모형 3)
종속변인 (DV)	전국 총 관객 수(Y_0)	영화진흥위원회 Boxoffice	0	0	0
독립변인 (IV)	창작	감독과위(X_1)	0	0	0
		배우과위(X_2)	0	0	0
		제작국가(X_3)	0	0	0
		장르(X_4)	0	0	0

		관람등급(X ₅)(더미)	전체, 12+, 15+, 18+(더미)	0	0	0
		러닝타임(X ₆)	러닝타임(분)		0	0
	배급	스크린 수(X ₇)	전국 개봉 스크린 수		0	0
	구전효과	전문가 평가(X ₈)	포털사이트 전문가 평점			0
		관객 평가(X ₉)	포털사이트 네티즌 평점			0
	첫주 흥행	첫 주 전국 주말 관객 수 (X ₁₀)	영화진흥위원회 Boxoffice			0

4. 연구방법

4.1. 데이터 수집과 분석방법

본 연구의 분석대상은 2004년 1월부터 2014년 12월까지 국내에서 상영된 상업영화 중, 매해 국내 총 관객 수를 기준으로 상위 100위를 기록한 1,100편의 영화이다. 기계학습 모형을 위해 1,100편의 영화중에서 990편의 데이터를 학습데이터로 활용하고, 110편의 데이터를 테스트 데이터로 활용한다. 독립변인인 장르, 감독, 배우, 영화등급, 국적, 스크린 수는 영화진흥위원회의 영화통합전산망에서 제공하는 2차 산업데이터를 활용한다. 관람객평가와 전문가평가는 포털사이트 네이버 영화에서 제공하는 네티즌평가와 전문가평가 점수를 활용한다.

본 연구는 두 단계로 구성된다. 첫째, 선형회귀분석을 통해 모형별로 예측요인으로서 유의미한 결과를 나타내는 변인을 살펴본다. 둘째, 인공신경망과 의사결정나무에 선형회귀분석 결과를 반영하여 각 분석결과의 예측력을 비교해본다. 선형회귀분석을 위해 통계 패키지인 SPSS ver. 21을 활용하였으며, 인공신경망과 의사결정나무 분석을 위해 기계학습 툴 키트(Tool-kit)인 웨카(Weka ver.3.8)를 활용한다.

4.2.1. 선형회귀분석

선형회귀분석은 선형적 관계를 가정하여 X변수를 통해 Y변수를 예측하기 위한 것으로서, 본 연구에서는 모형에 따라 식 (1), 식 (2), 식 (3)으로 다중선형회귀분석모형을 적용하였다. Y₀는 종속 변수의 값, β₀, β₁ ...β₁₀은 상수와 회귀계수다.

$$\ln(Y_0) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon \quad (1)$$

$$\ln(Y_0) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \epsilon \quad (2)$$

$$\ln(Y_0) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \epsilon \quad (3)$$

4.2.2. 인공신경망

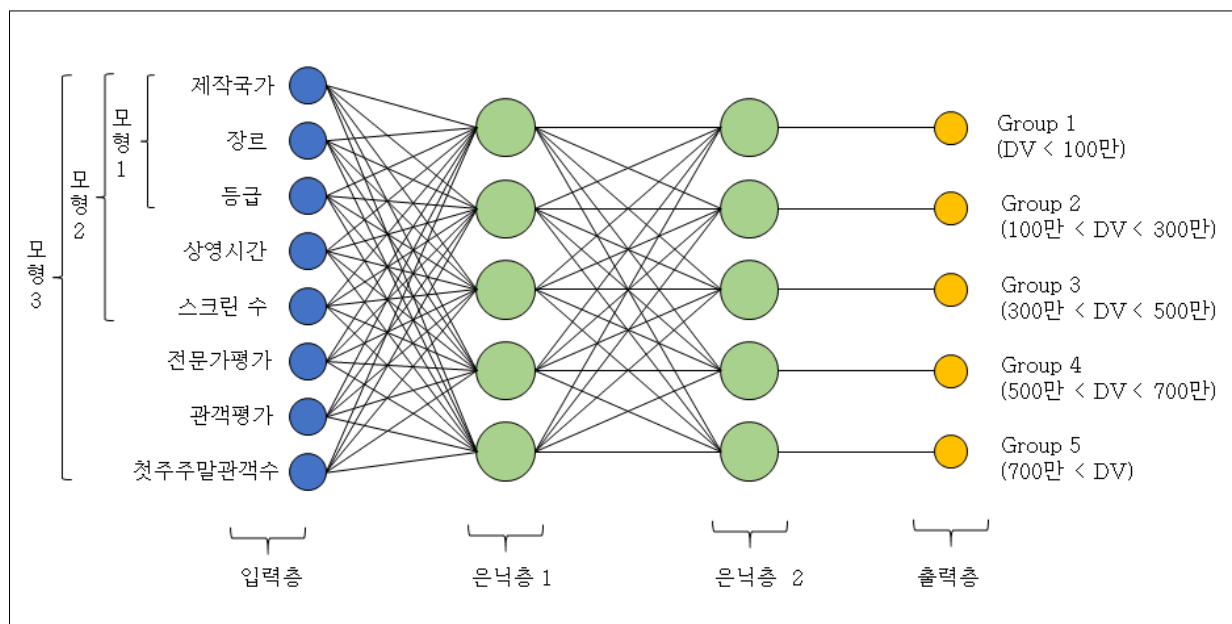
인공신경망은 입력층(Input Layer), 은닉층(Hidden Layer), 출력층(Output Layer)으로 이루어진

다계층 구조로 구성된다[12]. 본 논문에서는 역전파 학습방법(Backpropagation, BPN)을 이용하여 신경망의 연결정도를 최적화하였다. 입력층은 선형회귀분석 결과를 바탕으로 구성한다. 출력층은 흥행 성적인 총 관객 수 분포를 기준으로 5개의 그룹으로 분류한다([표 2] 참조). [그림 1]은 본 연구의 인공신경망 모형을 나타낸 그림이다. 모형에 따라 입력층에 투입되는 변인은 달라지며, 모형 간 비교를 위해 출력층은 동일하게 구성하였다.

[표 2] 2004-2014년 한국영화 총 관객 수 분포

[Table. 2] The proportion of 2004-2014 Korean movies' total box-office

그룹	기준	영화 수 (편)	백분율(%)
1	100만 미만	2670	84
2	100만 이상 ~ 300만 미만	372	12
3	300만 이상 ~ 500만 미만	87	3
4	500만 이상 ~ 700만 미만	26	1
5	700만 이상	26	1
합계		3181	100



[그림 1] 인공신경망 다중퍼셉트론 모델

[Fig. 1] Graphical representation of MLP neural network model

4.2.3. 의사결정나무

인공신경망 예측력과의 비교분석을 위해 의사결정나무 분석방법을 적용한다. 본 연구에 적용한 M5P 알고리즘은 의사결정나무 기법의 대표적인 알고리즘이다. M5P 알고리즘은 데이터의 엔트로피 값을 바탕으로 의사결정 구조를 구성한다. 과적합(overfitting)을 피하기 위해 10배 교차검증

(10-fold cross validation)을 적용한다.

5. 결과

본 연구에서는 2004년부터 2014년까지 국내에서 상영된 영화를 대상으로 영화 흥행 영향요인을 살펴보고 흥행성과를 예측하였다. 먼저, 영화 산업단계별로 고려할 수 있는 변인을 다르게 적용하여 세 가지 모형을 대상으로 선형회귀분석을 실시하였으며, 모형별로 유의한 변인에서 차이가 나타났다([표 3] 참조).

세부적으로 살펴보면, 모형 1($R^2=0.12$, $F=11.09$)의 경우, 제작국가($\beta=0.50$, $p<0.00$), 코미디 장르($\beta=0.56$, $p<0.00$), 드라마 장르($\beta=-0.34$, $p<0.00$), 공포 장르($\beta=-0.84$, $p<0.00$), 12세 이상 관람가($\beta=0.28$, $p<0.00$)가 유의한 변인으로 나타났다.

[표 3] 선형회귀분석결과

[Table. 3] The results of multiple regression analysis for total admission

변인	제작/투자 (모형 1)	배급 (모형 2)	상영 (모형 3)
(상수)	13.51**	9.30**	4.45**
감독과위	0.01	-0.01	-0.01
배우과위	0.01	0.00	0.00
제작국가	0.47**	0.48**	0.35**
장르	액션	-0.05	-0.04
	코미디	-0.55**	-0.33**
	드라마	-0.34**	-0.27**
	어드벤처	-0.06	-0.15
	공포	-0.79**	-0.60**
	스릴러	-0.32*	-0.28*
등급	전체	-0.16	0.06
	12세+	0.27**	0.30**
	15세+	0.07	0.12
ln_상영 시간	-	0.61**	0.42**
ln_스크린수	-	0.30**	0.10**
ln_전문가평가	-	-	0.02*
ln_관객평가	-	-	0.21**
ln_첫주전국주말관객수	-	-	0.50**
R^2	0.12	0.24	0.47
Adjusted R^2	0.11	0.23	0.46
F	12.49**	24.16**	55.50**

모형 2($R^2=0.24$, $F=22.66$)의 경우, 제작국가($\beta=0.47$, $p<0.00$), 12세 관람가($\beta=0.31$, $p<0.00$), 상영시간($\beta=0.61$, $p<0.00$), 스크린 수($\beta=0.29$, $p<0.00$)가 정적영향을 미치는 것으로 나타났으며, 코미디 장르($\beta=-0.33$, $p<0.00$), 드라마 장르($\beta=-0.27$, $p<0.00$), 공포 장르($\beta=-0.59$, $p<0.00$), 스릴러 장르($\beta=-0.28$, $p<0.00$)의 경우 부적 영향력을 보였다. 마지막으로 모형 3($R^2=0.47$, $F=53.04$)의 경우, 제작국가

($\beta=0.34$, $p<0.00$), 상영시간($\beta=0.43$, $p<0.00$), 스크린 수($\beta=0.10$, $p<0.00$), 전문가 평가($\beta=0.03$, $p<0.00$), 관객 평가($\beta=0.21$, $p<0.00$), 첫 주 전국 주말 관객 수($\beta=0.51$, $p<0.00$)가 정적 영향을 미치는 것으로 나타났으며, 배우 파워($\beta=-0.01$, $p<0.00$), 공포 장르($\beta=-0.36$, $p<0.00$)가 부적 영향력을 보였다.

다음으로 선형회귀분석에서 유의미한 변인을 바탕으로 모형 1,2,3을 재구성하여, 인공신경망과 의사결정나무 분석에 활용하였다. 선형회귀분석 결과의 반영여부의 효과성을 검증하기 위해 전체 변인을 대상으로 분석한 결과(미반영)와 결과를 반영하여 유의미한 변인만을 투입하여 인공신경망과 의사결정나무 결과값을 도출하였다([표 4] 참조). 인공신경망과 의사결정나무 예측력 비교기준으로 RMSE(Root Mean Square Error; 최소자승평균오차) 값을 활용하였다. RMSE 값은 모형이 예측한 값과 실제값 간의 차이를 평가하기 위한 척도로서, 오차정도를 파악하는데 주로 활용된다. 연구 결과, 회귀분석결과를 반영하지 않은 경우에는 세 모형에서 의사결정나무 모형보다 인공신경망 분석방법의 정확도가 우수한 것으로 나타났다. 한편, 인공신경망 분석의 경우, 선형회귀분석 결과의 반영여부가 정확도에 크게 영향을 미치지 않았다. 의사결정나무의 경우, 정확도가 평균 3.54% 개선되며, RMSE 수치가 개선되는 것으로 나타났다. 구체적으로, 제작/투자 모형에서는 미반영의 경우 인공신경망의 정확도가 50.50%이고 의사결정나무의 결과는 47.89%였으나, 결과를 반영한 경우에는 인공신경망의 정확도가 50.51%로 변동이 미미했으나 의사결정나무의 경우 54.26%로 인공신경망보다 높은 정확도를 보였다. 배급 모형에서도 미반영의 경우, 의사결정나무가 59.79%로 인공신경망의 결과(61.35%)보다 낮았으나, 결과를 반영한 후에는 의사결정나무가 62.33%로 인공신경망 결과(61.59%)보다 높았다. 상영 모형에서는 선형회귀분석결과 반영여부와 관계없이 인공신경망의 성능이 좋은 것으로 나타났다(미반영=73.53%, 결과반영=71.31%).

[표 4] 인공신경망과 의사결정나무 결과비교

[Table. 4] The Results of Artificial Neural Network Analysis and Decision Tree

		제작/투자 (모형 1)		배급 (모형 2)		상영 (모형 3)	
		미반영	결과반영	미반영	결과반영	미반영	결과반영
ANN	정확도(%)	50.50	50.51	61.35	61.59	73.53	71.31
	RMSE	0.34	0.34	0.32	0.32	0.28	0.29
DT	정확도(%)	47.89	54.26	59.79	62.33	68.40	70.11
	RMSE	0.38	0.34	0.36	0.34	0.33	0.31

본 연구는 의의는 다음과 같다. 첫째, 영화 흥행 예측연구에 기계학습 기법을 적용하여 성과를 향상시킬 수 있었다. 현재 국내영화시장을 분석대상으로 하는 흥행예측분야에서는 기계학습 기법의 활용이 미비하다. 본 연구는 초기연구로서 대표적인 기계학습 기법인 인공신경망과 의사결정나무 기법을 활용하였으며, 영화 데이터에 활용하기에 적합한 분석기법을 도출하기 위해 분석방법 간 성능을 비교하였다. 시기별 모형에 따라 분석기법의 정확도 결과에서 차이가 있었다. 제작/투자 모형과 배급 모형에서는 모든 변인을 투입(미반영)하였을 때 인공신경망의 정확도가 의사결정나무보다 높았으나, 회귀분석결과를 반영한 후에는 의사결정나무의 정확도가 인공신경망보다 높았

다. 반면, 상영 모형에서는 회귀분석결과의 반영여부와 관계없이 인공지능망의 정확도가 의사결정 나무의 정확도 보다 높게 나타났다. 이러한 결과는 각 모형별로 적합한 분석기법이 달라질 수 있다는 것을 의미한다.

둘째, 영화흥행 예측의 정확도 개선을 위해 선형회귀분석과 기계학습 분석을 접목하였다. 선형 회귀분석을 통해 변인의 유의미성을 확인한 후, 그 결과를 기계학습 분석에 적용하였다.

본 연구에서는 영화흥행예측분야에 기계학습 기법을 도입한 초기연구로서 대표적인 기법인 인공신경망과 의사결정나무 기법에 초점을 맞추었다. 후속연구에서는 서포트 벡터 머신(SVM), 랜덤 포레스트 등 다른 기계학습 기법도 추가적으로 분석할 수 있기를 기대한다. 또한, 본 연구에서는 선형회귀분석 결과를 바탕으로 모형을 추가하여 기계학습 분석에 활용하였다. 후속연구에서는 기계학습 기법의 결과와 선형회귀분석의 결과를 직접적으로 비교할 수 있기를 기대한다.

References

- [1] R. Lippmann, An introduction to computing with neural nets., IEEE. (1987), Vol.4 No.2, pp.4-22.
- [2] B. K. Wong, T. A. Bondnovich, and S. Yakup, Neural Network applications in business: A review and analysis of the literature, Decision Support Systems. (1997), Vol.19, pp.301-320.
- [3] R. P. Gorman and T. J. Sejnowski, Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets, Neural Networks. (1988), vol.1, pp.75-89.
- [4] J. H. Choi and D. S. Seo. Journal of The Korean Official Statistics. (1999), vol.4, No.1, pp.61-83.
- [5] T. W. Kim, Monthly precipitation forecasting using rescaling errors, Journal of Civil Engineering, KSCE. (2006), Vol.10, No.2, pp.137-143.
- [6] J. Eliashberg, A. Elberse, and M. A. A. M. Leenders, The motion picture industry: Critical issues in practice, Neural Networks. (1988), Vol.1, pp.75-89.
- [7] Y. H. Kim and J. H. Hong, A Study for the Development of Motion Picture Box-office Prediction Model., Communications for Statistical Applications and Methods. (2011), Vol.18, No.6, pp.859-869.
- [8] B. R. Litman, Predicting success of theatrical movies: An empirical study. The Journal of Popular Culture. (1983), Vol. 16, No.4, pp.159-175.
- [9] R. Sharda and D. Delen, Predicting Box-Office Success of Motion Pictures with Neural Networks, Expert Systems with Applications. (2006). Vol.30, pp.243-254
- [10] J. W. Chung, The Journal of Asia Movie. (2012), Vol.5, No.1, pp.173-203
- [11] S. J. Kwon, A Study on the Structure of the Secondary Market in the Korean Film Industry The Journal of Asia Movie. Review of Culture & Economy. (2014), Vol.17, No.1, pp. 35-55.
- [12] J. Y. Yim and B. Y. Hwang, Data Engineering : Predicting Movie Success based on Machine Learning Using Twitter, Korea Information Processing Society Journal of Information Processing Systems Tr. Software and Data Eng. (2014), Vol.3, No.7 pp.263-270