



영화산업에서 빅데이터의 활용방안 연구 - 넷플릭스(Netflix) 분석기술 중심으로

A study on the use of Big Data in Film Industry - Focused on Netflix Analytical Tools

저자 (Authors)	김진욱 Kim Jinwook
출처 (Source)	한국예술연구 (25) , 2019.9, 51-64(14 pages) The Korean Journal of Arts Studies (25) , 2019.9, 51-64(14 pages)
발행처 (Publisher)	한국예술종합학교 한국예술연구소 Korean Research Center For The Arts
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE08770219
APA Style	김진욱 (2019). 영화산업에서 빅데이터의 활용방안 연구 - 넷플릭스(Netflix) 분석기술 중심으로. 한국예술연구(25), 51-64
이용정보 (Accessed)	연세대학교 165.132.5.*** 2020/04/27 10:44 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

영화산업에서 빅데이터의 활용방안 연구

- ‘넷플릭스(Netflix)’ 분석기술 중심으로*

김진욱**

「차례」

1. 서론
2. 영화/영상 콘텐츠 추천 서비스 ‘넷플릭스’
3. 영화 온라인 추천서비스 및 빅데이터 기술
4. 결론

<국문초록>

최근 문화예술콘텐츠 전반에 빅데이터의 활용은 다양하게 적용되고 있다. 본고에서는 영화산업 분야에서 급속도로 발전되고 있는 거대 공룡기업인 온라인 추천서비스 넷플릭스의 빅데이터 분석기술에 대해 살펴보고, 비슷한 형태의 온라인 추천서비스 기업인 ‘아마존’의 프라임 비디오(Prime Video), 지니(Jinni)와 국내기업 왓챠(Watcha), 애플(Apple)TV+까지 어떠한 분석기술로 서비스를 하는지에 대한 형태도 살펴본다. 앞으로의 빅데이터는 데이터 자체가 스마트화 되며 IoT***를 통해 생산량이 급격히 증가될 것으로 기술적으로는 알고리즘과 인공지능이 발달하면서 대용량 데이터 처리에 가속도가 붙을 것으로 전망되며, 개인정보 보호 문

* 이 논문은 한국예술연구소에서 2019년 5월에 개최한 춘계학술대회 《빅데이터와 예술》에서 발표한 내용을 보완한 것이다.

** 평택대학교 공연영상콘텐츠학과 부교수.

*** ‘IoT’: ‘사물인터넷’으로 여러 사물에 정보통신기술이 융합되어 실시간으로 데이터를 인터넷으로 주고받는 기술.

제와 함께 미디어 업계와 IT 기술 업계의 긴밀한 협력을 통한 빅데이터 관리전문 인력 확보가 선행되어야 할 것으로 보인다.

주제어

영화산업, 빅데이터, 넷플릭스, 자연어 처리분석, 기계학습 기반

1. 서론

최근 문화예술콘텐츠 전반에 빅데이터의 활용은 다양하게 적용되고 있다. 장르별 대표적인 활용사례로는 방송산업에서는 DB 분석 기반으로 하는 시청자 맞춤형 추천정보가 있고, 음악분야에서는 DB 비교분석(패턴매칭) 기반으로 하는 음향 정보 검색, 게임에서는 이용자들의 플레이정보 DB분석 기반으로 하는 유지보수 및 이벤트 릴리즈 등이 있다.

영화산업에서는 영화의 인지도와 호감도 및 버즈량 분석에 기반한 소셜 데이터와 연동하여 영화마케팅 및 흥행예측 등이 있다. 김진옥의 박사논문 「빅데이터를 활용한 영화마케팅 연구」¹⁾에서는 SNS의 입소문으로 성공한 <설국열차> 등의 영화를 구글(google)에서 제공한 빅쿼리²⁾ 분석을 통해 영화 흥행 여부를 어떻게 예측할 수 있는지에 대한 연구와 <변호인> 등의 영화들에 대한 트위터에서의 버즈량을 분석하는 기술을 통해 흥행 요인 분석을 살펴본 사례가 있었다.

본고에서는 영화 온라인 추천서비스의 최대 기업인 ‘넷플릭스(Netflix)’의 빅데이터 분석 기술을 보다 구체적으로 살펴보고, 이에 따른 영화산업에서의 빅데이터 분석기술을 통한 활용을 전망해 보고자 한다.

1) 김진옥, 「빅데이터를 활용한 영화마케팅 연구 - SNS 이용자(특성) 중심으로」, 박사학위논문, 서강대학교, 2013.

2) ‘빅쿼리: 페타 바이트급의 데이터 저장 및 분석용 클라우드 서비스이다. <https://bcho.tistory.com/1116>, 접속일 2019.8.25.

2. 영화/영상 콘텐츠 추천 서비스 ‘넷플릭스’

‘넷플릭스’는 빅데이터 분석을 통해 가입자의 콘텐츠 기호를 파악하고 여기에 맞는 영화나 TV 프로그램을 추천하는 서비스를 제공한다. 콘텐츠 추천을 위해 활용되는 빅데이터에 대해 구체적인 내용을 공개하지 않고 있으나, 넷플릭스의 데이터 담당자가 거론했던 분석 내용을 참조해보면 매우 다양한 데이터를 활용하고 있는 것으로 드러났다. 이 데이터에는 2,500만 명 이상의 이용자들의 일시 정지, 되감기 등의 이용 행태를 포함한 하루 평균 3,000만 건의 동영상 재생기록, 최근 3개월 기간에 해당하는 20억 시간 이상의 동영상 시청 기록이 포함되었다. 이렇듯 넷플릭스의 빅데이터 분석은 가입자의 콘텐츠 기호 파악을 가능케 해 다른 콘텐츠 제작 시에도 상당한 기여를 하는 것으로 풀이된다. 현재 넷플릭스는 총 1억 5,000만 명 이상의 가입자를 확보하고 있으며, 이들 가입자들을 겨냥해 빅데이터 기법을 이용해 자사에서 자체 제작한 드라마인 ‘넷플릭스 오리지널(Netflix Original)’시리즈를 추천하고 있다.

보통 방송사들이 신규 프로그램 방영 시 시청자 확보를 위해 초기에 막대한 광고비를 투자하는 데 비해, 넷플릭스는 빅데이터 기법을 통해 잠재 시청자들에게 자사 콘텐츠를 추천함으로써 엄청난 광고 효과를 얻고 있는 것이다. 넷플릭스는 <하우스 오브 카드(House of Cards)> 이후, <못 말리는 패밀리(Arrested Development)>, <오렌지 이즈더뉴 블랙(Orange is the New Black)> 등 자체 제작 드라마 시리즈인 ‘넷플릭스 오리지널’을 통해 계속해서 돌풍을 일으키고 있다.

1) 넷플릭스의 추천 알고리즘

넷플릭스를 성공으로 이끄는 알고리즘 중에는 ‘추천’ 알고리즘이 있는데, 영화를 보고 나면 넷플릭스는 고객에게 ‘당신이 좋아할 것으로 생각되는 영화(Movies we think you will love)’를 추천한다. 이 추천 알고리즘은 전체 매출의 60%를 만들어내는 핵심 기술이다. 이렇듯 넷플릭스는 끊임없이 추천 알고리즘 개선에 투자하며 대표적으로 ‘넷플릭스 프라이즈(Netflix Prize)’를 통해 데이터를 공개하고 외부 전문가들 중 자사의 알고리즘 10%의 개선 조건으로 100만 달러의 상금을

수여한다. 넷플릭스 알고리즘은 ‘씨네매치(Gene Match)’→‘다이노소어 플레닛(Dinosaur Planet)’→‘그래비티(Gravity)’→‘프래그매틱 카오스(Pragmatic Chaos)’ 등 다양한 방식으로 추천 정확성을 향상시키며 발전하고 있다.

넷플릭스 추천 알고리즘의 정확도가 높은 이유 중 하나는 자체적으로 운영하고 있는 마이크로 장르 시스템인데 넷플릭스는 각 영화별로 수많은 메타데이터를 태깅³⁾하여 입력하고 이를 기반으로 다양한 장르를 만들어 낸다. 실제로 넷플릭스가 관리하는 장르는 약 77,000개에 이르는 것으로 알려져 있다.

2) 빅데이터를 활용한 투자와 관리

넷플릭스는 콘텐츠 수급 일환으로 자체 제작, 투자한 <하우스 오브 카드(House of Cards)>는 2013년 미국 상반기를 휩쓸었던 드라마로 1억 달러의 제작비가 투입되었으며 최고의 드라마를 대상으로 시상하는 제 65회 에미상에서 최우수 감독상을 비롯한 3관왕을 차지함으로써 작품성을 인정받았다. 넷플릭스는 지상파나 케이블 방송사가 아닌 만큼 따로 방영시간을 정해두지 않고, 온라인상에서 스트리밍으로 이용할 수 있도록 에피소드를 한편씩 차례대로 공개했다. <하우스 오브 카드>는 본래 영상 스트리밍 서비스만 담당하던 넷플릭스가 드라마 제작에 뛰어들어 자체적으로 제작한 ‘넷플릭스 오리지널(Netflix Original)’시리즈의 일환으로 기획되었다. 제작 과정에서 전문적인 데이터 마이닝⁴⁾ 과정을 통해 유명 감독 데이빗 핀치(David Fincher)를 제작에 참여시키고, 케빈 스페이시(Kevin Spacey) 등의 배우들을 캐스팅함으로써 큰 성공을 거두었고 이를 통해 할리우드로부터 주목받기 시작하였다.

3) ‘태깅: 사이트의 관리자가 사이트의 이미지나 텍스트를 관련된 주제나 카테고리의 형태로 키워드 처리를 해주는 것으로, 쉽게 말하면 태그를 다는 것. 코리아인터넷닷컴, <http://www.korea.com/>, 접속일 2019.8.25.

4) ‘데이터 마이닝: 대규모로 저장된 데이터 안에서 체계적이고 자동적으로 통계적 규칙이나 패턴을 찾아내는 것. <https://bit.ly/2P74pnF>, 접속일 2019.8.25.

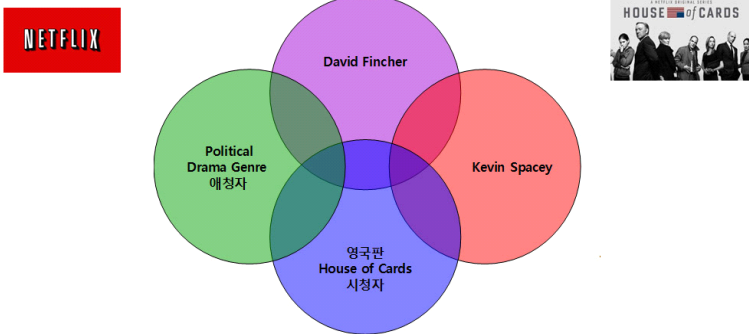


그림 1. House of card(하우스 오브 카드) 흥행요인⁵⁾

3. 영화 온라인 추천서비스 및 빅데이터 기술

다양하고 방대한 영화콘텐츠를 온라인으로 자유롭게 스트리밍 방식으로 즐기는 시대가 도래하면서 온라인 기업들은 영화 추천 서비스를 앞 다투어 개발하고 있다. 앞서 살펴본 넷플릭스는 3만여 개의 영화와 TV 프로그램을 제공하고 있으며, 후발 주자인 ‘아마존의 프라임 비디오(Prime Video)’는 약 8만 여개의 콘텐츠를 제공하고 있다. 이는 가정뿐만 아니라 야외, 스마트폰, 웨어러블 등의 스마트 기기로 편하게 영화를 볼 수 있는 서비스가 필요함에 따라 생겨나는 자연스런 현상이 되었다.

넷플릭스의 영화 추천사이트 ‘시네매치(Cine-match)’뿐 아니라, ‘지니(Jinni)’ 그리고 최근에는 국내의 벤처기업인 영화 추천 서비스인 ‘왓챠(Watcha)’가 사용자를 점차 늘려가고 있는 추세이다.

1) 빅데이터 분석기술

영화 추천 서비스에 사용되는 분석기술로는 대표적으로 빅데이터 분석(Big Data)분석, 자연어 처리 분석(Natural Language Processing), 기계학습 기법(Machine

5) Netflix, www.netflix.com, 접속일 2019.8.25.

Learning)이 있다.

‘빅데이터 분석(Big Data)’에는 빅쿼리, 하둡처리⁶⁾ 등이 있고, 이 분석 방법은 소셜미디어와 웹 데이터 수집으로 유저들과 평론가들에게 얻어지는 영화 평가나 리뷰, 메타데이터로 방대하게 확장된다.

‘자연어 처리분석(Natural Language Processing)’은 인간의 언어를 컴퓨터가 이해할 수 있는 데이터로 변환해 분석하는 기술로, 영화의 내용이나 분위기와 같은 특징이 인간의 주관적인 감정을 반영하는 단어로 입력되기 때문에 이를 컴퓨터로 처리하기 위해 필요한 기술이다. 이를 위해 마이닝분석과 전처리과정⁷⁾을 거쳐야 하는데 마이닝 분석은 데이터마이닝, 텍스트마이닝, 오피니언마이닝 이렇게 세분류 분석을 통해 얻어낼 수 있다.⁸⁾ 전처리 과정은 트윗, URL, 문장부호, 불가능용어 제거를 통해 정확도 검사⁹⁾를 마치고 얻어낼 수 있다.

‘기계학습 기법(Machine Learning)’은 ‘딥 러닝(Deep Learning)’이라고도 불리는 일종의 인공지능 기술로, 컴퓨터가 스스로 각종 데이터를 통해 학습하고 판단할 수 있는 알고리즘과 처리 방식으로 개발하는 기법으로 2016년 3월 세계를 살짝 놀라게 하며 이슈가 됐던 바둑경기 ‘이세돌과 알파고’를 연상하면 쉽게 이해할 수 있다.

2) 영화 추천 서비스 기술의 원리

영화 데이터 수집은 분석 시스템의 설계에 따라 분류 기준이 정해지고, 컴퓨터가 자동적으로 인터넷이나 클라우드에 수집하거나, 사람이 수작업으로 판단하여 입력할 수도 있다. 이렇게 수집된 데이터를 위에 언급한대로 전처리 과정, 마이

6) ‘하둡 처리’: 대용량 데이터를 적은비용으로 더 빠르게 분석할 수 있는 소프트웨어이며, 빅데이터 처리와 분석을 위한 플랫폼. <http://bitly.kr/6WHy81y>, 접속일 2019.8.25.

7) ‘전처리 과정’: 데이터들의 목적파일을 만들기 전에 처리하는 특별한 과정. <https://writeren.tistory.com/99>, 접속일 2019.8.25.

8) 마이닝의 세 분류: 마이닝에는 가장 큰 규모의 데이터 마이닝이 있고, 다음에 텍스트 마이닝, 그다음에 감정과 상태 등을 추출하는 오피니언 마이닝, 이렇게 세 분류가 있음. <http://bitly.kr/gMGvBT0>, 접속일 2019.8.25.

9) ‘정확도 검사’: 데이터들의 목적파일을 만들기 위한 오타자 등을 제거하고 정확한 값을 나타내기 위한 검사. <http://bitly.kr/tfyAmiW>, 접속일 2019.8.25.

닝 분석을 거쳐 감상자의 선호에 맞게 속성, 추출한 값을 부여한 후 데이터베이스로 저장해 컴퓨터의 알고리즘으로 처리한다. 이러한 기계학습 기법은 영화 추천 서비스뿐만 아니라 검색 엔진에서부터 음악이나 게임, 데이트 상대 매칭서비스까지 다양한 분야에 활용되고 있다.

영화 추천 기술 구현은 영화 관련 및 감상자 데이터 수집, 영화의 속성 분류 및 평가, 알고리즘에 따른 영화 선호도와 영화 추천으로 이루어지고, 이런 정량적 데이터는 영화의 장르, 상영 시간, 개봉 일시, 흥행 성적과 같은 객관적인 데이터를 포함하며 영화의 내용적 데이터는 영화의 분위기(공포스릴거나 또는 사랑스러운), 스토리의 복잡함(비)선형구조, 액자구조, 반전구조)과 같이 영화를 본 감상자의 주관적인 판단과 평가로 분류한다.

감상자 데이터는 유저의 별점 평가, 평론가들의 영화 평가 점수와 같이 숫자로 정량화될 수 있는 평가 수치를 포함하고, 다양한 데이터를 수집, 조합해 영화의 속성을 분류하고 평가하는 과정을 거친다. 이 같은 데이터를 알고리즘에 따라 처리하여 개인이 어떤 속성의 영화를 선호하는지 분석하고, 이와 비슷한 영화를 추천한다. 예를 들어, 어떤 사용자의 영화 선호 경향을 분석해 코미디 장르, 공포영화의 분위기, 단순한 스토리 구조를 선호하는 것으로 판단될 경우, 이와 유사한 속성을 지닌 영화들을 추천한다.

영화 추천 서비스 기술은 데이터의 양이 많아질수록 더욱 정교하고 정확해지며, 영화의 속성을 분류하고 가중치를 부여하는 방식에 따라 같은 알고리즘을 사용하더라도 다른 결과가 나올 수 있다. 이는 빅데이터의 분석기술과 인공지능(딥러닝)의 결합된 산물이라고도 볼 수 있다.

기본적으로 빅데이터 분석 기법을 활용하기 때문에 영화와 관련된 각종 데이터, 감상자의 평가 데이터가 많아질수록 추천 결과가 개인의 선호도를 더욱 정확하게 반영하게 되고, 각 영화 추천 서비스들은 저마다 서로 다른 영화 속성 분류 기준을 적용하고 있으며, 어떤 기준으로 분류하느냐에 따라 동일한 알고리즘을 사용하더라도 영화 추천 서비스의 결과가 달라질 수 있기 때문이다.

3) 주요 영화 추천 서비스 분석

넷플릭스(Netflix)는 자사 서비스 가입자들로부터 얻어진 방대한 데이터를 기반

으로 ‘넷플릭스 퀀텀 이론(Netflix Quantum Theory)’이라고 불리는 기법을 활용한 다. 넷플릭스 내부적으로 콘텐츠 평가 및 분류 기법에 의해 각 영화별로 태그를 만들고 데이터베이스를 만들고, 이를 그대로 콘텐츠를 구성하는 ‘양자(Quantum)’가 무엇인지를 파악한다.

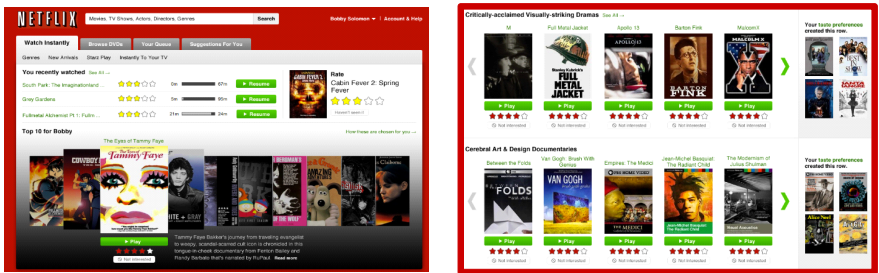


그림 2. 넷플릭스의 영화 추천 화면¹⁰⁾

‘지니(Jinni)’는 2008년 1월 설립 후 260만 달러, 2011년 5월 500만 달러의 벤처 투자금을 유치했고, 2010년에는 구글 TV(Google TV)와의 전략적인 제휴를 맺고, 2012년에는 스위스컴(Swisscom)과 자사의 시멘틱 검색¹¹⁾ 기술을 제공하는 계약을 맺었다. 2013년 11월에는 웹사이트 전면 개편과 함께 아이패드용 앱을 출시했고, 영화나 TV와 같은 콘텐츠를 개인 취향에 따라 분석, 시멘틱 검색과 소셜 기반 추천 서비스를 제공한다.

지니는 영화의 일명 ‘판도라(Pandora)¹²⁾’ 서비스라고 지칭되고 있으며, 영화 종류, 장르, 분위기, 배우 등의 분류에 따라 영화 추천을 제공한다. 추천 결과를 개봉일자, 장르 등의 기준으로 다시 정렬하거나 좁히는 것도 가능하며, 제정을 보유하고 있을 경우 추천 결과와 영화 선호도를 저장할 수 있다. 즉, 지니의 영화

10) Netflix, 접속일 2019.8.25.

11) ‘시멘틱 검색(Semantic Search)’: 문서의 의미(Semantic)를 분석해 검색하는 기술을 말함. 즉, 문장이나 단락에 기술된 주제를 파악하고 이를 대상으로 검색. <http://bitly.kr/Gud1pa8>, 접속일 2019.8.25.

12) ‘판도라: 세계 최대의 인터넷 음악 라디오 서비스로 정교한 음악 데이터 분석을 통한 추천 기능을 갖추고 있음. <http://www.pandora.tv/>, 접속일 2019.8.25.

데이터 분석은 '무비 게놈 프로젝트(Movie Genome Project)'에 의해 자체적인 기준으로 수집되고 분류된다.

지니는 시멘틱 검색을 통해 '갱스터들이 출연하는 영화', '미래를 배경으로 한 액션 영화', '데이트용 영화'와 같은 문구를 검색하면 해당 키워드에 맞는 영화들을 추천해 주는 기능을 제공한다. 지니의 영화 추천은 무비 게놈 프로젝트를 통한 자체적인 데이터분석과 '머신 러닝(Machine-Learning)' 및 '자연어 분석(Natural Language Processing)' 기술에 기반을 두고 이루어진다.

영화 게놈 분석은 영화의 분위기(mood, tone)와 플롯(전쟁 이야기, 선악 대립), 구조(비)전형적이야기, 액자 구조, 기준(폭력, 선정성, 멜로)과 같은 이야기적 요소로 구분되고, 이러한 영화 게놈의 기준은 영화 전문가들에 의해 지니 내부적으로 사람의 수작업으로 선정되는데, 수 천 개의 게놈 중에 영화 한 편을 특징짓는 게놈은 50개 정도 결정된다. 또한 영화 평론가 평가, 유저 평가, 메타데이터를 기반으로 머신 러닝 기술과 자연어 분석 기술을 통해 지니의 시스템이 영화 작품들을 자동적으로 분류하게 되고, 사람의 수작업과 시스템의 자동화된 처리를 병행함으로써 영화 추천의 질을 향상시킨다고 밝혔다.

'프로그램스(Programs)'라는 한국의 벤처 회사에서 개발해 2013년 5월부터 서비스를 개시한 영화 추천 서비스인 '왓챠(Warcha)'는 최근 인기를 얻으면서 사용자가 증가하고 있다. 그 당시, 왓챠는 25억의 벤처 투자 자금을 유치할 정도로 기술력과 서비스 역량을 인정받았으며, 페이스북(facebook) 아이디어에 기반을 둔 유저들의 참여를 바탕으로 현재까지 2억만 개 이상의 영화평가 DB를 확보하고, 좋아하는 영화 20여 편에 대한 별점을 매기면 이와 유사한 영화를 영화감독, 배우, 장르 등의 조건에 맞춰 보여주는 방식으로 페이스북 친구의 별점 평가와 리뷰도 확인할 수 있다. 넷플릭스와 마찬가지로 자신이 원하는 취향에 맞는 영화를 제법 잘 찾아준다는 입소문과 함께 사용자들의 높은 평가를 받고 있는 것으로 알려졌다.

왓챠는 최근 새로운 버전을 내놓으면서 기능을 추가했으며, 언어적인 제약이 크지 않은 서비스 특성 상 해외 진출도 가능할 것으로 전망된다. 왓챠 2.0 버전에서는 사용자가 추천받은 영화를 바로 스마트폰이나 IPTV에서 볼 수 있도록 연결하는 기능이 추가되었다. 왓챠의 유저 평점 수는 2억만 건을 돌파할 정도로 단기간에 성장했으며, 영화 평점 데이터가 늘어날수록 유저 취향 분석이 더욱 정확

해지는 강점을 보유하고 있다.

4. 결론

현재 문화예술콘텐츠 분야에서 빅데이터의 활용 사례는 다양하게 적용되고 있다. 본고에서는 영화산업 분야에서 급속도로 발전되고 있는 거대 공룡기업인 온라인 추천서비스 넷플릭스(Netflix)의 빅데이터 분석기술에 대해 살펴보고, 비슷한 형태의 온라인 추천서비스 기업인 아마존의 프라임 비디오(Prime Video), 지니(Jinni)와 국내기업인 왓챠(Watcha)까지 어떠한 분석기술로 서비스를 하는지에 대한 형태도 살펴보았다.

최근 애플(Apple)은 애플TV 발표에 이어 ‘애플TV+’를 새롭게 출시하면서 방송 뿐만아니라 영화, 게임, 음악까지도 클라우딩을 통해 서비스 하고 있다. 이렇듯 영화 추천 기술과 서비스의 중요성은 계속해서 증가하며 스마트폰과 웨어러블 기기 발달과 함께 이어나갈 전망이다.

물론 콘텐츠의 양과 가짓수가 너무 많은 경우에 사용자가 자신이 선호하고 좋아하는 것을 찾아내는 것이 매우 힘들어진다는 문제도 발생하기도 할 것이다. 앞서 살펴봤듯이, 넷플릭스는 경쟁사인 아마존에 비해 적은 콘텐츠 라이브러리¹³⁾로도 더 높은 매출과 더 큰 가입자 만족도를 이끌어내는 이유는 넷플릭스만의 특유의 추천 분석 시스템이기 때문이다. 따라서 영화 콘텐츠 제공 업체의 경쟁력은 얼마나 ‘많은’ 콘텐츠를 제공하느냐가 아닌, 얼마나 ‘양질’의 작품들을 소비자에게 ‘효율적으로’ 추천하느냐에 의해 결정될 것이다.

현재 영화 추천 서비스에 이용되는 데이터는 사용자가 과거 어떤 특징의 영화를 선호했는가와 영화의 객관적 데이터 및 내용적 정보와 같이 과거에 존재했던 불변하는 데이터만을 사용하고 있지만 앞으로는 날씨가 어떤지, 누구와 함께 있는지, 어디서 영화를 볼 것인지, 영화를 보고 싶은 시간이 언제인지와 같은 상황

13) Netlix가 약 3만여 개의 콘텐츠 라이브러리를 확보한 데 비해 후발 주자인 Amazon의 Prime Video는 약 8만 6천여 개의 콘텐츠를 제공 중. <http://bitly.kr/FL0Wx5b>, 접속일 2019.8.25.

인지 정보(Context-aware Information)를 활용하는 방향으로 진화할 전망이다.

이러한 상황 인지 정보는 스마트폰과 착용형 단말(Wearable Device)의 발달¹⁴⁾에 따라 과거보다 훨씬 간편하고 정확하게 수집 및 분석될 수 있다. 다만, 이러한 상황 인지 정보가 개인 프라이버시에 밀접하게 연관되어 있기 때문에, 이를 영화 추천에 반영하기 위해서는 자신의 상황 인지 정보를 활용하는 데 대한 유저의 거부감을 줄이고, 개인정보 유출에 대한 우려를 해소할 필요도 있다.

앞으로의 빅데이터는 데이터 자체가 스마트화 되며 IoT를 통해 생산량이 급격히 증가될 것으로 보인다. 따라서 영화산업에서도 넷플릭스 뿐만 아니라 다양한 온라인 서비스 스트리밍 업체들은 발빠르게 개발하고, 뛰어들 것이다. 기술적으로는 알고리즘과 인공지능이 발달하면서 대용량 데이터 처리에 가속도가 붙을 것으로 전망되며, 개인정보 보호 문제와 함께 미디어 업계와 IT 기술 업계의 긴밀한 협력을 통한 빅데이터 관리전문 인력 확보가 선행되어야 할 것으로 보인다. 인력 측면에서는 데이터 사이언티스트(data scientist)나 알고리즘미스트(algorithmist) 같은 전문직이 부상할 것으로 보이고, 이에 영화/영상 콘텐츠 장르에 전문지식이 덧 붙여 진 전문가라면 국내 뿐 아니라 해외에서도 적극 필요로 할 것이다. 따라서 외부로부터의 우수한 빅데이터 관련 인력 확보와 전문성 부여를 통한 인력 육성이 매우 필요할 전망이다.

14) 스마트폰을 통해 위치 정보와 시간 정보가 파악할 수 있고, 착용형 단말을 통해 전달되는 심박수와 같은 생체신호를 통해 개인의 신체적이고 감정적인 상태 또한 간접적으로 파악 가능함.

참고문헌

1. 단행본

Mayer-Schönberger, Viktor & Cukier, Kenneth. 이지연 옮김. 『빅데이터가 만드는 세상』. 서울: 21세기북스, 2013.

2. 논문

김진욱. 「빅데이터를 활용한 영화마케팅 연구 - SNS 이용자(특성) 중심으로」. 박사 학위논문, 서강대학교, 2013.

송명빈 · 이상호. 「소셜미디어와 영화산업 관계의 정책적 함의 연구」, 『디지털융복합연구』Vol.12 No.1. 천안: 한국디지털정책학회, 2014.

이오준 외. 「소셜 빅데이터를 이용한 영화 흥행 요인 분석」. 『한국콘텐츠학회논문지』Vol.14 No.10. 대전: 한국콘텐츠학회, 2014.

최화열 · 황규일 · 이문규. 「디지털 미디어 융합 환경에 따른 영화산업의 내외적 변화」. 『서비스마케팅저널』Vol.4 No.1. 세종: 서비스마케팅학회, 2011.

한국콘텐츠진흥원. 「빅데이터 분석을 통한 한류지도 구축방안」. 『KOCCA포커스 2014-04호』 나주: 한국콘텐츠진흥원, 2014.

_____. 「영화 추천 기술과 서비스」. 『KOCCA포커스』2014-09호 나주: 한국콘텐츠진흥원, 2014.

_____. 「콘텐츠 분야에서의 빅데이터 기법 활용 사례, 빅데이터 시장 현황과 콘텐츠산업 분야에 대한 시사점」. 『KOCCA포커스』2013-11호 나주: 한국콘텐츠진흥원, 2013.

한국정보화진흥원. 「빅데이터의 진화 : 스마트 데이터」. 대구: 한국정보화진흥원 2013.

_____. 「새로운 미래를 여는 Big Data 시대」. 대구: 한국정보화진흥원, 2013.

3. 인터넷 및 기타자료

- Economist. "Data, data everywhere". 2010. <http://goo.gl/oumB>. 접속일 2019.8.25.
- firmex.com. "7 Big Data Techniques That Create Business Value". 2013.1.18.
<https://www.firmex.com>. 접속일 2019.8.25.
- McKinsey. "Big Data: The next frontier for innovation, competition, and productivity". 2011. <https://www.mckinsey.com>. 접속일 2019.8.25.
- Marr, B.. "Why Big Data Will Change the World". 2013. <http://goo.gl/S30FxE>. 접속일 2019.8.25.
- Scientific American. <http://bitly.kr/pl8bFQH>. 접속일 2019.8.25.
- Singel, R.. "Netflix Spilled your Brokeback Mountain Secret, Lawsuit Claims, Wired". 2009. <http://goo.gl/oO1rwh>. 접속일 2019.8.25.
- The Atlantic. "How Netflix Reverse Engineered Hollywood". 2014.1.2.
<http://bitly.kr/uXKigQ>. 접속일 2019.8.25.
- Wikipedia. "Big Data". http://en.wikipedia.org/wiki/Big_data. 접속일 2019.8.25.

A study on the use of Big Data in Film Industry

– Focused on ‘Netflix’ Analytical Tools

Kim Jinwook*

Big data is widely applied to cultural and arts content. In this article, we will look at how Netflix, a rapidly developing media company, uses big data by examining their analytical tools as well as those of Amazon’s ‘Prime Video’, ‘Jinni’ the domestic company ‘Watcha’, and ‘Apple TV+’ as points of comparison. In the future, big data will become smarter, and rapidly increase through the popularization of IoT devices. Technically speaking, algorithms and artificial intelligence are expected to improve the processing power of large amounts of data. Finally, media companies and I.T companies must work together to protect data privacy and experts in big data analysis must be recruited as a matter of priority.

key words

Film industry, Big data, Netflix, Natural language processing analysis, Machine learning

접 수 일 : 2019년 7월 4일

심사기간 : 2019년 7월 19일~2019년 9월 6일

게재결정 : 2019년 9월 6일

* Associate Professor, Department of Media contents and Performing Art, Pyeongtaek University.