



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

경영학석사 학위논문

리뷰가 비디오 게임의 성과에
미치는 영향:
기계학습을 적용한 트위터
데이터의 활용

2020 년 2 월

서울대학교 대학원
경영학과 경영학 전공
이 현 경

리뷰가 비디오 게임의 성과에
미치는 영향:
기계학습을 적용한 트위터
데이터의 활용

지도 교수 김 상 훈

이 논문을 경영학석사 학위논문으로 제출함

2020 년 1 월

서울대학교 대학원
경영학과 경영학 전공
이 현 경

이현경의 석사 학위논문을 인준함

2020 년 1 월

위 원 장 _____ 송 인 성 _____ (인)

부위원장 _____ 김 병 도 _____ (인)

위 원 _____ 김 상 훈 _____ (인)

초 록

오늘날 온라인을 통해 생성되는 구전은 소비자들에게 지대한 영향을 미치고 있다. 본 연구는 비디오 게임 소비에 중요한 영향을 미칠 것으로 예상되는 온라인 구전 중 트위터 리뷰에 주목하여 소비자들의 리뷰가 비디오 게임의 성과에 어떠한 영향을 미치는지를 규명하고자 하였다. 이를 위해 총 47개의 게임을 대상으로 트위터 리뷰를 수집하였다. 이후 머신러닝 기법을 적용한 감성 분석을 시행하여 리뷰의 감성을 분류하고, 이를 분석에 적용하였다. 트위터 리뷰는 텍스트로 구성되어 있어 기존 온라인 리뷰의 영향력을 검증한 문헌에서 주로 사용하고 있는 리뷰 평점과는 다른 접근 방식이 요구된다. 이에 리뷰를 분석 가능한 형식으로 가공하기 위하여 베이지안 분류 방식을 적용한 뒤, 리뷰가 긍정적인지 부정적인지를 확인하여 연구 변수로 활용하였다.

연구 결과 트위터 리뷰의 수는 판매량에 긍정적인 영향을 미치는 것으로 나타났다. 이는 게임의 리뷰가 많아질수록 소비자들의 게임 선택에 정의 영향을 미친다는 것을 뜻한다. 또한 트위터 리뷰가 긍정적일수록 비디오 게임의 성과에도 정의 영향을 미치는 것을 확인하였다. 더불어 본 연구는 정보이론에서 사용되는 엔트로피 개념을 활용하여 온라인 리뷰 감성의 분포에 따라 온라인 리뷰의 효과가 조절됨을 확인하였다. 엔트로피가 높아 소비자 리뷰의 의견이 갈릴 경우 트위터 리뷰의 양과 극성의 효과는 모두 감소하였으며, 이는 온라인 구전의 분포가 리뷰가 미치는 영향을 조절할 수 있음을 시사한다. 본 연구는 게임 산업에서의 온라인 리뷰의 영향력을 알아봄으로써 현재

온라인 매체에서의 소비자들의 활동이 어떻게 실제 산업과 매출에 영향을 주고 있는지를 확인하였다는 점에서 의의를 지닌다.

주요어 : 비디오 게임, 온라인 구전, 트위터 리뷰, 감성분석, 엔트로피, 기계학습

학 번 : 2018-24282

목 차

| | |
|-------------------------------|----|
| 제 1 장 서 론..... | 1 |
| 제 2 장 이론적 배경 및 가설 설정 | 5 |
| 제 1 절 온라인 구전에 관한 연구 | 5 |
| 제 2 절 리뷰의 감성과 엔트로피 | 9 |
| 제 3 장 연구 방법..... | 11 |
| 제 1 절 자료의 수집 및 측정..... | 11 |
| 제 2 절 감성분석..... | 15 |
| 제 3 절 모델의 설정 | 18 |
| 제 4 장 분석 결과..... | 21 |
| 제 5 장 결론 및 논의 | 24 |
| 제 1 절 연구 결과의 요약 및 시사점..... | 24 |
| 제 2 절 연구의 한계점 및 향후 연구과제 | 27 |
| 부 록 | 30 |
| 참고 문헌..... | 33 |
| Abstract..... | 38 |

표 목차

| | |
|--------------------|----|
| <표 1> 연구 변수 | 11 |
| <표 2> 기술통계표 | 17 |
| <표 3> 분석 결과표 | 19 |

그림 목차

| | |
|---------------------------|----|
| <그림 1> 트위터 리뷰 감성 분포 | 15 |
|---------------------------|----|

부록 목차

| | |
|---------------------------------|----|
| <Appendix 1> 트위터 리뷰 판별 예시 | 26 |
| <Appendix 2> 상관계수표 | 27 |

제 1 장 서 론

오늘날 발달된 온라인 환경은 소비자들의 의사결정 과정에 많은 변화를 가져왔다. 소비자들은 다양한 정보에 노출되어 있으며 넘치는 정보를 어떻게 선별하고 흡수하는지에 따라 그들의 의사결정은 달라진다. 특히 최근의 온라인 상황에서는 Facebook, Twitter, Weibo와 같은 SNS(Social Network Service)를 비롯한 다양한 매체들이 활용되고 있으며, 이러한 매체들을 통하여 수많은 상품과 서비스들에 대한 정보와 리뷰가 쏟아지고 있다. 소비자들은 그들이 경험한 재화, 서비스에 대한 정보를 적극적으로 생산하고 있으며, 이와 같은 소비자들의 정보 생산 활동은 실시간으로 소비자들의 구매 결정에 큰 영향력을 행사하고 있다. 특히 트위터(Twitter)와 같은 마이크로블로그(microblog) 서비스들은 특히 그 영향력과 파급력이 큰 것이 특징이다. 이러한 매체들은 실시간으로 대량의 정보를 교환, 확산시켜 소비자들에게 큰 영향을 미치기 때문에 마이크로 블로그를 통해 생성되는 구전은 모바일 기기의 발전과 함께 더욱 주목받고 있다(Henning-Thurau, Wiertz & Feldhaus, 2015).

트위터는 ‘트윗’이라는 140자 내외의 제한된 짧은 텍스트를 작성할 수 있는 소셜 네트워크 서비스로, 트위터를 통해 작성된 트윗은 작성자의 팔로워들에게 실시간으로 공유된다는 특징을 가지고 있다(Jin & Phua, 2014). 팔로워들은 이후 다시 그들과 네트워크를 형성하고 있는 팔로워들에게 해당 트윗을 리트윗하거나 그들이 내용을 직접

재공유함으로써 짧은 시간에 수많은 트위터 사용자들에게 정보를 퍼트린다. 이러한 특성 때문에 트위터는 실시간성과 정보의 빠른 확산력을 특징으로 한다(임준엽, 황병언, 2014). 이와 같은 파급력을 바탕으로 트위터를 통한 구전은 소비자들로 하여금 제품이나 서비스에 대한 즉각적인 감정을 불러일으켜 그들의 의사결정과 구매 과정에 결정적인 영향을 줄 수 있다(Jansen, Zhang, Sobel & Chowdury, 2009). 이와 같은 특성에 기인하여 본 연구에서는 온라인에서 생성되는 리뷰 중 트위터에서 생성되는 구전이 게임 소비자의 의사결정에 영향을 미칠 것이라고 보고 트위터를 통한 온라인 구전이 게임의 성과에 미치는 영향을 알아보고자 한다.

트위터에서 생성된 구전은 제품에 대한 평점과 같이 비교가 가능한 정형의 숫자로 이루어진 것이 아닌, 비정형의 텍스트로 이루어진 데이터라는 특징을 가지고 있다. 소비자들은 트위터를 통해 직접 구매하거나 경험한 재화에 대한 의견과 감성을 직접적으로 제시한다. 따라서 트위터를 통해 작성되는 리뷰들은 숫자 또는 점수 형식의 제품에 대한 별점, 평점의 온라인 구전과는 다른 형식을 가지고 있다고 할 수 있다. 텍스트로 이루어진 구전인 트위터 리뷰는 그 영향력을 측정하기 위해서 비정형의 텍스트를 가공하고 이를 분석하는 방식이 요구된다. 따라서 본 연구에서는 트위터 리뷰의 감성을 기계학습을 적용해 구분하고, 이를 모델에 적용하여 트위터 리뷰가 실제로 비디오 게임의 성과에 어떠한 영향을 미칠 것인지를 알아보고자 하였다. 이를 실행하기 위해 본 연구에서는 기계학습을 이용한 텍스트 분류를 시행하였다. 먼저 게임 전문 리뷰 사이트에서 평점에 따라 작성되어 있는 게임 리뷰를

긍정적인 리뷰와 부정적인 리뷰로 나누어 학습시킨 뒤 학습된 분류 기준에 따라 트위터 리뷰의 감성이 긍정적 또는 부정적인지를 판별하였다. 이후 판별된 리뷰를 활용하여 구전 효과에 대한 분석을 진행하였다.

본 연구는 비디오 게임을 대상으로 온라인 리뷰의 영향력을 검증한다는 점에서 기존 연구와 차별점을 가진다. 기존에 트위터 리뷰의 영향력을 검증하는 연구들은 주로 영화 산업에 집중하여 온라인 리뷰가 영화의 성과에 미치는 영향을 검증한 바 있다. 하지만 게임 산업의 경우 이와 같은 온라인 리뷰의 영향력을 확인한 연구는 비교적 적은 것이 현실이다. 하지만 게임 시장은 다양한 디바이스들의 발달과 함께 최근 그 규모가 매우 커지고 있다. 특히 게임 산업에서는 온라인 리뷰의 영향력이 매우 클 것으로 예상되는데, 게임의 경우 게임 내 스크린샷, 동영상 등을 게시하거나 스트리밍 서비스를 활용하는 등의 방식으로 소셜 미디어를 통한 소비 경험의 공유가 원활하게 이루어지고 있기 때문이다. 실제로 게임은 온라인 환경을 통해 활발한 커뮤니케이션이 이루어지고 있는데, 소셜 미디어를 통해 게임과 관련된 많은 정보가 확산되고 있으며, 일부 게임의 경우 소셜 미디어와 결합되어 온라인 네트워크 서비스 환경 내에서 게임을 즐길 수 있도록 제작되기도 하였다(Marchand, Henning-Thurau, 2013). 또한 Twitter(2018)에서 발표한 내용에 따르면 실제로 약 52%의 트위터 유저들이 게임에 관심이 있다고 밝혔으며, 또한 32%의 트위터 유저들이 새로운 게임을 접하고 난 뒤 트위터를 통해 그들의 경험을 공유하는 것으로 나타났다. 이는 게임을 즐기는 많은 소비자들이 소셜 미디어 채널의 정보에 의해 많은 영향을 받을 수

있음을 시사한다. 이러한 상황을 고려하였을 때 게임 산업에서 실시간으로 생성되는 온라인 리뷰는 매우 중요한 요인으로 자리 잡았음을 알 수 있으며, 이러한 온라인 리뷰의 영향력을 알아보는 것은 현재 온라인 매체에서의 소비자들의 활동이 어떻게 실제 게임 산업과 매출에 어떠한 영향을 주고 있는지를 이해하는 데에 도움이 될 것이다.

제 2 장 이론적 배경 및 가설 설정

제 1 절 온라인 구전에 관한 연구

구전 효과에 대한 논의는 많은 선행연구를 통해 다루어져 왔다. Borgida, Nisbett(1977)은 구전행동을 개인의 경험을 토대로 대면하여 전달되는 커뮤니케이션이라고 정의하였으며, Rosario, Ana et al.(2016)은 마케팅에서 구전이 소비자들이 재화, 서비스, 브랜드 또는 기업에 대한 정보를 다른 소비자들에게 제공하는 활동이라고 설명하고 있다. 구전은 최근 온라인 상황에서도 활발하게 생성되고 있는데, Hennig-Thurau et al.(2004)은 온라인 구전(Electronic Word of Mouth, eWOM)이 소비자가 인터넷을 통해 생성한 특정 제품이나 기업에 대한 긍정적이거나 혹은 부정적인 진술로, 누구나 자유롭게 온라인을 통해 이용 가능한 정보라고 밝혔다. 이와 같은 온라인 구전은 인터넷 환경을 통해 수많은 소비자들에게 빠르게 확산된다는 성격을 가지고 있으며, 문자를 기반으로 한다는 특징을 가지고 있다. 또한 온라인 구전은 공간과 시간적 제약이 비교적 약하여 많은 양의 정보를 시간과 장소에 관계없이 습득, 처리할 수 있어 그 파급력이 크다고 할 수 있다. 이러한 특징 때문에 온라인 구전은 소비자들이 제품에 대한 구매 의사 결정을 할 때에 지대한 영향을 미치게 된다. 이러한 온라인 구전의 효과를 게임 산업에 적용한 선행 연구의 경우, 소셜 네트워크 게임에 SNS가 미치는 영향(Hansen, Lee, 2013), 비디오게임의 판매에 인터넷 검색 활동이

미치는 영향(Ruohonen, Hyrynsalme, 2017) 등이 있다.

온라인 환경에서 생성되고 있는 구전 중 최근 그 중요성이 증대되고 있는 것은 마이크로블로그를 통해 생성되는 구전(microblog WOM, 이하 MWOM)이다. MWOM은 온라인 구전의 한 종류로서, 트위터와 같은 소셜 미디어 서비스를 통해 생성되는 정보, 메시지를 의미한다. 마이크로블로그(microblog)를 통해 개인들은 다양한 정보들을 빠르게 교환한다. 마이크로블로그 서비스 내에서 소비자들은 친구, 팔로워들과 대규모의 네트워크를 형성하고 있으며 그들이 작성하는 의견과 리뷰는 네트워크를 통해 실시간으로 빠르게 확산되게 된다. Hennig-Thurau, Wiertz & Feldhaus(2015)는 이와 관련하여 ‘트위터 효과’에 대하여 논의한 바 있다. 트위터 효과란 트위터를 비롯한 마이크로블로그 서비스를 통해 생성된 구전이 소비자들의 실제 제품 채택 행위에 영향을 주는 현상을 의미한다. 이들은 이 연구에서 마이크로블로그를 통해 생성된 WOM을 MWOM(microblogging word of mouth)라고 정의하였으며, 이러한 MWOM이 소비자들의 구매 후 제품에 대한 평가를 즉각적으로 퍼트리기 때문에 제품의 성과에 영향을 미친다고 설명하고 있다. 이와 같은 MWOM은 온라인 구전과 전통적 구전의 성격을 동시에 가지고 있다는 특징을 지니고 있다. 먼저 실시간으로 정보의 전송이 이루어진다는 점에서 MWOM은 전통적 구전의 성격을 가지고 있으며, 마이크로블로그 서비스 내에서 유저들 간의 상호 연결이 가능하다는 점 또한 전통적 구전과 유사하다. 또한 온라인 구전과도 유사한 점을 가지고 있는데, 문자로 커뮤니케이션이 이루어진다는 점과, 정보가 다수의 사람들에게 노출된다는 점에서 유사점을 갖는다. 특히

MWOM은 영화, 음악, 게임과 같은 경험재의 경우에 중요하게 작용할 수 있는데, 이는 경험재의 경우 출시 초기의 즉각적인 성공이 제품의 최종적 성과에 매우 중요하게 작용하기 때문이다. 이때에 강한 확산력을 가진 MWOM은 초기 사용자들의 경험을 빠르게 확산시키며 결과적으로 제품의 성과에 지대한 영향을 미치게 된다.

온라인 구전은 구전의 수(volume)가 많으면서, 여러 매체에 확산되어 있을수록 소비자의 의사결정에 긍정적인 영향을 미친다(Dellarocas et al., 2007). 구전은 시장에서 제품의 인지도를 상승시키고, 제품에 대한 불확실성을 줄여 제품의 성과에 영향을 주기 때문이다(Chen, Wang & Xie, 2011; Chintagunta, Gopinath & Venkataraman, 2010). 이때에 구전은 제품에 대한 정보를 담고 있기 때문에 구전의 양이 증가할수록 소비자들은 온라인 구전을 통해 해당 제품에 대한 정보를 더 많이 얻을 수 있으며 결과적으로 매출을 상승시키게 된다. 또한 구전은 리뷰의 방향성(valence)이 긍정적일수록 효과적인 것으로 알려져 있다(Dellarocas et al., 2007). 온라인 구전의 방향성은 구전의 내용이 긍정적이거나, 혹은 부정적인 정도를 의미하는데(Rosario et al., 2016), 소비자가 그들이 경험한 재화에 대하여 느낀 감성 또는 선호에 대한 표현이라고 할 수 있다. 따라서 구전의 방향성이 긍정적이라는 것은 결국 소비자들이 해당 제품을 선호하고 있음을 나타낸다고 할 수 있다.

트위터 리뷰 또한 구전의 양과 방향성에 따라 효과가 달라진다는 것이 확인되었는데, Hennig-Thurau, Wiertz, and Feldhaus는 2015년의 연구에서 트위터의 구전 효과를 함께 검증하였다. 저자들은 트위터 리뷰의 양이 많을 때에 영화의 초기 성과에 긍정적인 영향을 주는 것을

확인하였으며, 부정적인 트윗의 비율이 높을수록 영화의 성과에 부정적인 영향을 미치는 것을 밝혔다. 이와 같이 트위터를 통한 온라인 구전의 경우에서도 마찬가지로 구전의 양(volume)과 극성(valence)은 소비자들의 제품 선택에 영향을 줄 수 있음을 확인할 수 있다. 이러한 논의를 토대로 본 연구는 다음과 같은 가설을 설정하였다.

가설 1: 트위터 리뷰의 양은 비디오 게임의 성과에 정(+)의 영향을 미칠 것이다.

가설 2: 긍정적인 트위터 리뷰의 비율이 증가할수록 비디오 게임의 성과에 정(+)의 영향을 미칠 것이다.

본 연구에서는 온라인 구전 중 게임의 성과에 영향을 미칠 수 있는 변수를 추가하여 게임의 성과에 구전이 미치는 영향에 대한 이해를 높이고자 하였다. 이에 온라인 구전 중 전문가 리뷰의 영향력을 추가적으로 확인하였다. 온라인 사용자 리뷰는 정보 제공자를 확실히 알 수 없다는 점에서 필연적으로 불확실성을 가지게 된다. 이러한 상황에서 전문가 리뷰는 비교적 객관성을 가지게 되며, 정보의 신뢰성을 확보할 수 있다는 점에서 일반적인 소비자들의 구전과는 다른 성격을 가진다. 전문가 리뷰가 영화의 성과에 미치는 영향은 많은 선행 연구를 통해 확인되어 왔으나, 그 효과에 대한 논의는 비교적 의견이 다양하다. 일반 소비자들에 의해 생성된 온라인 구전의 경우 영화 수요에 긍정적인 효과를 미치는 반면 전문 평론가의 평론 점수는 부정적인 효과를 미치는 것으로 나타난 연구(김일경, 전경미, 이호욱, 2012), 직접적인 소비를 하는 소비 당사자가 아닌 전문가, 미디어가 생성하는 리뷰의 경우 그

영향력이 감소함을 검증한 연구(Chen, Zhang, 2012)등은 전문가 리뷰가 부정적으로 작용할 수 있음을 증명한 연구이다. 특히 경험재의 경우 그 제품의 종류와 특성에 따라 전문가 효과는 달라질 수 있음이 선행연구를 통해 확인된 바 있다(Reinstein, Snyder, 2005). 따라서 본 연구는 비디오 게임에 온라인 전문가 리뷰가 어떠한 영향을 미치는지를 확인하고자 다음과 같은 가설을 설정하였다.

가설 3: 온라인 전문가 평점은 비디오 게임의 성과에 정(+)의 영향을 미칠 것이다.

제 2 절 리뷰의 감성과 엔트로피

앞서 논의한 것처럼 온라인 구전은 소비자가 제품에 대해 느낀 감성을 내포하게 되며 긍정적 또는 부정적인지에 따른 방향성(valence)을 가지게 된다. 이때에 제품 또는 브랜드에 따라 리뷰의 감성이 어떻게 분포되어 있는지 또한 제품에 따라 달라지게 될 것이다. 온라인 리뷰의 분포 혹은 분산 또한 소비자들의 인식에 영향을 줄 수 있음이 알려져 있는데, He, Bond(2015)는 실제로 아마존의 리뷰 데이터를 활용하여 온라인 평점의 분산에 따라 제품에 대한 소비자의 인식이 달라질 수 있음을 확인 바 있다. 본 연구에서는 이처럼 트위터 리뷰의 감성이 어떻게 분포하고 있는지에 따라 구전의 영향력이 달라질 것이라고 보고, 이를 알아보기 위해 추가적으로 엔트로피(entropy) 개념을 이용하고자 한다. 1948년 Shannon에 의해 처음 소개된 정보 엔트로피는 정보의

불확실성을 의미한다. 한 개의 단위 정보가 증가하게 되면 엔트로피는 감소하게 되며, 반대로 엔트로피가 증가한다면 정보의 불확실성은 증가한다. 따라서 만일 독립된 사건의 확률이 같을 경우 사건의 불확실성이 높아지기 때문에 엔트로피는 가장 큰 값을 가지게 된다. 0

엔트로피는 정보이론에서 제시된 이후 많은 연구에 활용되어 왔는데, 온라인 구전에 관한 선행 연구에서도 역시 활용된 바 있다. 이정은 2011년의 연구에서 정보 엔트로피의 관점으로 온라인 리뷰가 감소시키는 엔트로피의 양에 따라 제품에 대한 소비자들의 인식이 달라진다고 설명하였다. 리뷰를 통해 정보의 양이 증감함에 따라 불확실성은 변화하기 때문이다. 한편 Lee et al.(2017)은 온라인 리뷰의 극성에 엔트로피 개념을 적용하여 엔트로피가 온라인 리뷰의 효과를 조절할 수 있다는 것을 증명하였다. 이 연구에서 저자들은 온라인 리뷰가 완전히 긍정적이거나 완전히 부정적인 경우 소비자들이 해당 리뷰에 대한 정보가 없다고 판단하기 때문에 온라인 리뷰의 효과가 감소할 것이라고 보고 이를 실증적으로 확인하였다. 이러한 관점에서 본 연구는 엔트로피가 트위터 리뷰의 효과를 조절할 수 있는 것이라고 보고 다음과 같은 가설을 설정하였다.

가설4a: 트위터 리뷰 감성의 엔트로피가 증가하면 트위터 리뷰의 양이 비디오 게임의 성과에 미치는 영향이 감소할 것이다.

가설4b: 트위터 리뷰 감성의 엔트로피가 증가하면 긍정적인 트위터 리뷰의 비율이 비디오 게임의 성과에 미치는 영향이 감소할 것이다.

제 3 장 연구 방법

본 연구는 게임 소비자들이 작성한 비정형의 텍스트 리뷰가 비디오 게임의 판매에 실제로 어떠한 영향을 주는지를 규명하고자 하였다. 이를 위해 트위터를 통해 작성된 비디오 게임 소비자들의 리뷰를 수집하고 감성 분석을 활용하여 이를 분류하는 방법으로 데이터를 가공하였다. 트위터 리뷰는 텍스트로 구성되어 있어 기존 온라인 리뷰의 영향력을 검증한 문헌에서 주로 사용하고 있는 리뷰 평점과는 다른 접근 방식이 요구된다. 이에 리뷰를 분석 가능한 형식으로 가공하기 위하여 기계학습을 통한 분류를 적용하였고 분류된 트위터 리뷰를 연구 변수로 활용하였다. 이와 더불어 트위터 리뷰를 비롯해 게임 판매에 영향을 미칠 수 있는 변수들을 추가로 수집하고 모델을 설정하여 게임의 성과에 어떠한 요인들이 영향을 미치는지를 알아보았다.

제 1 절 자료의 수집 및 측정

먼저 본 연구에서는 2017년 하반기(7월~12월)부터 2018년 상반기(1월~6월)에 출시된 Playstation4 게임을 분석 대상으로 선정하였다. 이 중 출시 후 첫 주의 트위터 리뷰의 수가 50개를 넘지 않는 게임의 경우에는 분석 대상에서 제외하였다. 트위터 리뷰의 개수가 매우 작은 게임의 경우 트위터 리뷰의 영향력을 측정하기 어려우며,

연구결과를 왜곡할 수 있다고 판단하였기 때문이다. 위 과정을 거쳐 총 47편의 게임을 분석 대상으로 선정하였다. 이후 게임 각각의 출시 후 26주 동안의 주별 판매량을 수집하여 이를 종속 변수로 활용하였다. 판매 데이터의 경우 미주지역의 판매 실적을 기준으로 수집하였으며 단위는 판매된 게임의 개수(unit)로 측정하였다. 비디오 게임 시장의 경우 Sony사의 Playstation4, Microsoft의 Xbox one 등 몇 개의 대형 콘솔 게임 기기가 전체 비디오 게임 시장을 점유하고 있으며, Playstation4의 경우 현재 세대의 비디오 게임 플랫폼 시장의 50%를 넘는 점유율을 가지고 있다(VGchartz, 2019). 따라서 본 연구의 분석 대상은 Playstation4 게임에 한정하여 진행하였다.

47개 게임에 대한 트위터 리뷰는 다음과 같은 방법으로 수집하였다. 먼저 각 게임 i 출시 후 26주 동안 게임 이름 전체가 포함되어 작성된 트윗을 모두 수집하였다. 트윗 수집을 위해서는 Python을 사용한 웹 크롤링 기법을 적용하였다. 수집된 트위터 리뷰의 전처리에는 다음과 같은 방식으로 진행하였다. 먼저 리뷰 중에서 텍스트 내에 알파벳, 숫자, 문장 부호를 제외한 기타 언어가 포함된 트윗을 제외하였다. 이는 미주 지역의 판매에 영향을 줄 수 있는 트윗만을 선별하기 위함이다. 이 과정을 통해 한국어, 중국어 등 다른 문자로 작성된 트윗을 제거하였다. 이후 URL이 포함된 트윗을 분석 대상에서 제외하였다. 이는 수집된 트윗 중 광고를 필터링하는 가장 효율적인 방법으로 Rui et al.(2013)의 연구에서 사용한 방법을 적용하였다. 실제로 트위터 데이터의 경우 Youtube 광고, 웹사이트 광고등이 URL을 직접 링크하여 게시되고 있었으며, 이와 같은 방법으로 많은 수의 광고를 제거할 수 있었다.

위와 같은 과정을 거쳐 47개 게임에 대한 총 224,578개의 리뷰 텍스트를 수집하였다. 이 외에 추가적으로 사전 구전의 양을 알아보기 위하여 게임 i 의 출시 전 일주일 동안의 트위터 리뷰의 개수를 함께 수집하였다. 이 때에 수집된 트윗은 텍스트 형식이 아닌 단위 기간 동안 작성된 트윗 수로 수집하였다. 자료 수집은 2019년 11월 20일부터 25일까지의 6일간 시행하였다.

트위터 리뷰 외에 분석에 사용하기 위하여 수집한 변수와 데이터의 출처는 다음 <표1>에 정리된 내용과 같다. 먼저 비디오 게임 성과에 영향을 주는 주요 변수인 전문가 리뷰 변수는 *professional*으로, Metacritic 사이트에서 대상 i 게임에 대하여 미디어, 또는 전문가들이 매긴 평점의 평균을 수집하여 이를 10점 만점으로 변환한 값을 변수로 사용하였다. 이외에 게임의 성과에 대한 영향을 통제하고자 수집한 변수는 다음과 같다. 먼저 *priorsales*는 전기($t-1$)의 판매 개수이며, 관측된 주의 한주 전의 판매 개수(*unit*)를 의미한다. 게임의 특성과 관련된 변수로는 해당 게임이 Playstation4외에 다른 플랫폼에 얼마나 많이 배포가 되어 있는지, 해당 게임이 유명한 배급사의 게임인지 등을 변수에 포함하였다. *Entropy* 변수는 트위터 리뷰 감성의 엔트로피를 계산한 값이다. 이는 긍정적 또는 부정적 감성의 리뷰 비율에 따른 불순율을 의미하며, 트위터의 극성이 얼마나 일관적인지를 판별하기 위한 변수로 사용하였다. 앞서 논의한 엔트로피는 정보의 불확실성을 의미한다. 따라서 리뷰가 긍정적이거나 부정적일 확률이 동일하다면 정보의 불확실성은 상승하며, 엔트로피 또한 증가할 것이다. 예를 들어, 만일 게임 i 의 트위터 리뷰가 모두 긍정적이거나 혹은 부정적인 경우에

<표1> 연구 변수

| 변수 | 설명 | 출처 |
|---------------------------------|---|------------|
| <i>sales_{it}</i> | 게임 i의 출시 이후 t 번째 되는 주의 판매 unit 수 | Vgchartz |
| <i>priorsales_{it}</i> | 게임 i의 출시 이후 t-1 번째 되는 주의 판매 unit 수* | Vgchartz |
| <i>count_{it}</i> | 게임 i의 출시 이후 t 번째 되는 주의 트윗 개수의 합 (단위: 100개) | Twitter |
| <i>PNratio_{it}</i> | 게임 i의 출시 이후 t 번째 되는 주의 긍정적 트윗 개수를 부정적 트윗 개수로 나눈 비율 | Twitter |
| <i>Entropy_{it}</i> | 게임 i의 출시 이후 t 번째 되는 주의 긍정/부정 트윗의 비율에 따른 엔트로피(불순률)** | Twitter |
| <i>precount_t</i> | 게임 i의 출시 전 일주일 동안의 트위터 개수의 합 (단위: 100개) | Twitter |
| <i>price_t</i> | 게임 i의 출시 이후 t 번째 되는 주의 월별 평균 가격 | Metacritic |
| <i>numplatform_{it}</i> | 게임 i가 출시 된 플랫폼의 개수 | Metacritic |
| <i>professional_t</i> | 게임 i의 전문가 리뷰 평점 평균 (10점 만점 기준) | Metacritic |
| <i>pub_t</i> | 게임 i의 배급사가 2018년 기준 매출액 상위 25위에 속할 경우 1, 아닌 경우 0 | |
| <i>adapt_t</i> | 게임 i가 다른 콘텐츠(영화, 소설)를 적용하여 개발된 경우 1, 아닌 경우 0 | |
| <i>sequel_t</i> | 게임 i가 기존 게임의 후속작인 경우 1, 아닌 경우 0 | Metacritic |

* $\ln(\text{priorsales})$ 변환을 위해 값에 0.0001을 더한 값을 사용

** $entropy = - \sum_{k=1}^m p_k \log_2(p_k)$ 공식을 적용

엔트로피는 0에 가까울 것이다. 반면에, 긍정적인 리뷰와 부정적인 리뷰가 동등하게 혼재되어 있을 경우 엔트로피가 1에 가까워지게 되어 트위터의 극성이 균일하게 분포되어 있는 것으로 판단할 수 있다. 엔트로피를 계산하기 위해 사용한 공식은 다음 (1)과 같으며, 여기서 엔트로피는 긍정과 부정의 두 개의 클래스를 가진다.

$$entropy = - \sum_{k=1}^m p_k \log_2(p_k) \quad (1)$$

제 2 절 감성분석

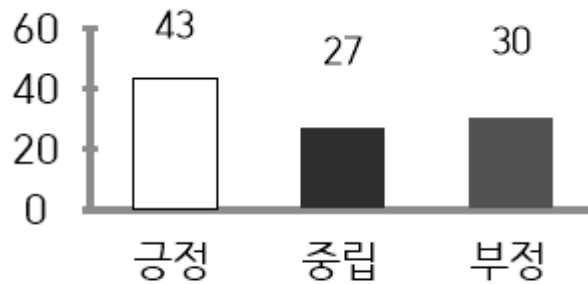
감성분석은 트위터에서 수집한 게임 관련 리뷰로부터 주관적인 감정, 의견 등을 텍스트로부터 추출해내어 분석에 적용하기 위한 방법으로, 다양한 기계학습 기법을 적용하여 이루어지고 있다(김유영, 송민, 2016). Pang et al.(2002)은 영화 리뷰 데이터를 학습 데이터로 활용하여 다양한 기법의 기계학습 분류를 실행하고 이를 비교하였는데, 나이브 베이지스, 최대 엔트로피, 서포트 벡터 머신(SVM)을 사용한 모델을 설정하고 이를 소개하였다. 이 중 본 연구에서는 나이브 베이시안 분류(Naïve Bayesian classification)방법을 활용하여 트위터 리뷰의 감성이 긍정적인지, 부정적인지를 분류하였다. 나이브 베이시안 분류 방법은 독립적인 속성들을 통해 결과를 분류하는 방법으로, 특히 분류의 축이 두개일 때 그 성능이 우월한 것으로 알려져 있다(Ferguson et al., 2009). 본 연구에서는 트위터 리뷰를 긍정과 부정의 두 축으로 나누어

분석을 시행할 것이기 때문에 베이지안 분류 방식의 기계학습 방법이 적합하다고 판단하였다. 베이지안 분류를 통한 기계학습을 위해서 먼저 학습에 사용되는 텍스트 데이터를 단어를 기준으로 토큰화하여 텍스트 데이터에 어떠한 단어가 속해 있는지를 구분하였다. 이후 구분된 단어들이 긍정적인 리뷰에 속할 것인지, 부정적인 리뷰에 속할 것인지 그 확률을 계산하여 학습시키는 모델을 생성하였고, 해당 모델을 수집된 트위터 리뷰에 적용하였다. 이 방식을 통해 수집된 트위터 리뷰가 각각 긍정적인지 부정적인지 판별 확률이 계산되며 결과적으로 리뷰가 내포하고 있는 텍스트의 감성이 부정과 긍정 중 어느 클래스에 속하게 될지를 결정하였다. 이 때에 판별 확률이란 그 값이 1에 가까울수록 해당 트윗이 긍정적이라고 판별될 확률이 높아지는 것을 의미한다.

기계학습을 적용한 분류 모델을 만들기 위하여 본 연구에서는 먼저 Metacritic에서 제공하고 있는 게임 리뷰를 수집하여 학습을 위한 재료 데이터로 활용하였다. 2019년 11월 기준 TOP 랭킹에 올라 있는 게임 107개의 리뷰를 대상으로 수집하였고, 긍정적인 리뷰와 부정적인 리뷰를 각각 5,585개 수집하여 총 11,170개의 리뷰를 수집하였다. 리뷰 개수의 단위는 개행 문자로 나뉘어진 문단이다. 이후 Metacritic에서 선별된 리뷰 데이터를 기반으로 베이지안 분류 모델을 생성하였다. 위에서 언급한 바와 같이, 기계학습을 위해 수집한 Metacritic 리뷰 데이터는 긍정적인 리뷰와 부정적인 리뷰로 나뉘어져 있다. 따라서 리뷰 텍스트를 토큰화한 뒤 해당 감성에 따라 어떠한 단어가 사용되었을 때 두 클래스인 긍정과 부정 중 어떤 축에 속할 확률이 높은지를 결정하게 된다.

기계학습을 위해 수집한 Metacritic 리뷰 데이터의 약 80%에 해당하는 텍스트는 학습을 위한 트레이닝 자료로 사용하였으며, 20%에 해당하는 텍스트는 모델의 성능을 확인하기 위한 테스트 셋으로 나누어 구성하였다. 테스트 셋을 활용하여 본 연구에서 사용한 분류 모델의 성능을 측정한 결과 정확도는 약 87%로 감성분석을 시행한 선행 연구와 유사하거나 비교적 높게 나타나(Hodeghatta, 2013; Das, Chen, 2007) 모델 활용에 무리가 없다고 판단하였다. 다음 단계는 최종적으로 생성된 모델을 앞서 수집한 트위터 리뷰에 적용하여 비디오 게임에 대한 트위터 리뷰가 각각 어떠한 극성을 가지고 있는지를 판별하는 것이다. 이 결과 수집한 224,578개의 트위터 리뷰가 각각 긍정적인 리뷰로 판별될 확률을 계산할 수 있었다.

위에서 판별한 결과값을 기준으로 본 연구에서는 판별 확률이 0.7 이상일 경우에 긍정적인 리뷰로, 0.3 이하인 경우 부정적인 리뷰, 그리고 0.7 초과 0.3 미만인 경우에는 중립적 리뷰라고 판단하였다. 이는 리뷰의 극성을 분류하기 위해 판단 기준을 두어 긍정 또는 부정이 판별되지 않는 중립(또는 관련 없는) 리뷰들을 제외하기 위함이다. 수집된 224,578개의 트위터 리뷰의 판별 결과는 다음 <그림 1>과 같다. 약 43%의 리뷰가 긍정적인 것으로 판별되었으며, 부정적인 리뷰는 30%로 비교적 적은 것을 알 수 있다. 중립 또는 감성과 관련이 없는 리뷰들은 전체 리뷰의 27%를 차지하였다.



단위: 백분율

<그림 1> 트위터 리뷰 감성 분포

제 3 절 모델의 설정

아래의 수식 (2)은 위에서 논의한 변수를 종합한 최종 모델이다. 종속변수로서 판매 성과를 나타내는 *sales*의 경우 게임에 따라 편차가 매우 커 그 값이 크게 편중되어 있기 때문에 로그변환을 취하여 모델을 설정하였다. 로그변환을 통해 정규 분포로 근사할 수 있도록 하는 방법은 영화에 대한 구전 효과를 검증한 선행 연구를 참조하였다(Chevalier, Mayzlin 2006; Hennig-Thurau, Wiertz & Feldhaus, 2015). 아래의 식에서 *count*변수는 리뷰의 양을, 그리고 *PNratio*는 리뷰의 극성을 나타내며 각각이 게임의 판매 성과에 어떠한 영향을 주는지를 알아보하고자한 변수이다. *PNratio*는 게임 *i*의 출시 이후 *t* 번째 되는 주의 긍정적 트윗 개수를 부정적 트윗 개수로 나눈 비율로

측정하였다.

$$\begin{aligned}
 LN(sales) = & \beta_0 + \beta_1 LN(priorsales) + \beta_2 count + \beta_3 PNratio \\
 & + \beta_4 Entropy + \beta_5 count * Entropy + \beta_6 PNratio * Entropy \\
 & + \beta_7 professional_i + \beta_8 pub_i + \beta_9 numplatform_i + \beta_{10} adapt_i + \beta_{11} sequel_i \\
 & + \beta_{12} price_i + \beta_{13} count_i + \epsilon
 \end{aligned} \tag{2}$$

리뷰의 긍정적 트윗 개수와 부정적 트윗 개수는 앞절에서 논의한 감성분석을 통해 판별한 것으로 리뷰의 극성이 게임 성과에 미치는 영향을 알아보기 위한 변수이다. *Entropy*는 감성분석을 통해 비디오 게임의 리뷰의 감성이 얼마나 고르게 분포되어 있는지, 혹은 긍정, 부정으로 치우쳐져 있는지를 알아보기 위한 변수로 사용하였으며, 긍정과 부정 두 개의 클래스로 나누어 감성이 얼마나 고르게 분포되어 있는지에 따라 게임의 성과에 어떠한 영향을 미치는지를 알아보고자 하였다. 더불어 본 연구에서는 엔트로피가 게임의 성과에 주로 영향을 줄 것으로 예상되는 리뷰의 양과 극성을 조절할 것이라고 예상하였기 때문에 *entropy*와 *count*(리뷰의 양), *PNratio*(리뷰의 극성)의 상호작용항을 각각 추가하였다. 이 외에 비디오 게임의 판매에 영향을 미칠 것으로 예상되어 이를 통제하고자 사용한 변수들을 추가하였는데, 먼저 해당 게임 *i*가 Playstatin4 외에 PC 또는 모바일과 같은 다른 플랫폼에 출시되었을 경우에 출시된 플랫폼의 수를 측정한 *numplatform*과, 해당 게임이 책, 또는 영화와 같은 기존에 있던 콘텐츠를 적용하여 개발된 것인지에 대한 변수 *adapt*가 있다. 이는 플랫폼에 게임이 많이 배포되어 있을수록, 게임의 소재가 이미 다른 콘텐츠를 통해 알려져 있을수록 소비자들의 게임에 대한 인식이 상승할 것이라고 예상하였기 때문이다. 이 외에 게임 *i*의 출시 가격과 출시 전

일주일 동안의 사전 구전의 양을 추가하였다. 이는 출시 전 구전의 양이 제품에 대한 소비자들의 기대를 반영할 것으로 예상하였기 때문이다. 이와 같은 변수들은 트위터 리뷰가 영화의 초기 성과에 영향을 주는 것을 검증한 Hennig-Thurau, Wiertz & Feldhaus(2015)의 연구를 참고하였다. 모델에 사용한 변수에 대한 기술통계값은 <표2>에 수록하였다.

<표2> 기술통계표

| Variables | Mean | SD | Min | Max |
|---------------------|-------|-------|-------|-------|
| <i>lnsales</i> | 7.55 | 1.71 | 0.69 | 13.98 |
| <i>lnprior</i> | 6.93 | 3.64 | -9.21 | 13.98 |
| <i>count</i> | 1.83 | 3.83 | 0.01 | 66.53 |
| <i>PNratio</i> | 3.66 | 5.73 | 0 | 59.00 |
| <i>professional</i> | 7.28 | 0.93 | 4.30 | 9.30 |
| <i>pub</i> | 0.55 | 0.50 | 0 | 1 |
| <i>numplatform</i> | 3.28 | 1.24 | 1 | 7 |
| <i>adapt</i> | 0.30 | 0.46 | 0 | 1 |
| <i>sequel</i> | 0.60 | 0.49 | 0 | 1 |
| <i>price</i> | 33.69 | 12.63 | 13.97 | 96.50 |
| <i>precount</i> | 3.68 | 4.84 | 0.17 | 25.67 |

제 4 장 분석 결과

본 연구는 패널회귀분석을 시행하여 게임의 성과에 리뷰가 미치는 영향을 알아보았다. 패널회귀분석을 통해 게임의 개별적 특성과 출시 후 시간에 따른 변화를 통제하고자 하였기 때문이다. 이를 분석한 결과는 다음 <표3>과 같다. 먼저 트위터 리뷰의 양(volume)은 비디오 게임의 성과에 정(+)의 영향을 미치는 것으로 나타났다. 이는 가설1을 지지하는 결과이다. 또한 긍정적인 트위터 리뷰의 비율이 증가할수록 비디오 게임의 성과에 유의한 영향을 미치는 것으로 나타나 가설2 또한 지지되었다. 이는 온라인 리뷰에 관련한 선행 연구의 방향과 유사하다. 많은 사전 연구에서 온라인 구전의 양과 긍정적인 극성은 성과에 정의 영향을 주는 것을 확인할 수 있다. 본 연구의 결과를 통해 게임 산업에서 또한 마찬가지로 온라인 리뷰가 동일한 영향을 미치는 것을 알 수 있다. 한편 전문가의 평점 변수는 기존 문헌과 차별되는 연구 결과를 보여주었는데, 전문가의 평점은 비디오 게임의 성과에 정(+)의 영향을 미치는 것으로 나타났다. 전문가 평점의 경우 그 영향이 재화의 특성에 따라 매우 다르게 나타나는 것으로 알려져 있으며, 특히 경험재일 때 그 영향이 달라짐을 선행 연구를 통해 확인한 바 있다. 분석 결과 비디오 게임에 전문가 평점은 매출에 큰 영향을 미치는 것을 확인할 수 있다. 엔트로피는 트위터 리뷰의 감성이 혼재되어 있는 경우 게임의 성과에 비교적 부정적인 영향을 주는 것으로 나타났으나 통계적인 타당성은 부족하였다. 다만 엔트로피의 조절 효과는 본 연구를 통해 확인할 수 있다. 엔트로피가 높아 소비자 리뷰의 의견이 갈려, 감성의 불확실성이

<표3> 분석 결과표

| <i>DV = Ln(sales)</i> | | | |
|------------------------|-----------|-----------------------|----------|
| Model | | Without interactions | |
| Variables | Coef. | Variables | Coef. |
| (intercept) | 1.998 ** | (intercept) | .892 |
| <i>Ln(priorsales)</i> | .055 *** | <i>Ln(priorsales)</i> | .063 *** |
| <i>count</i> | .408 *** | <i>count</i> | .063 *** |
| <i>PNratio</i> | .033 * | <i>PNratio</i> | -.007 |
| <i>entropy</i> | -.782 . | <i>entropy</i> | -.259 |
| <i>count*entropy</i> | -.282 *** | | |
| <i>PNratio*entropy</i> | -.291 *** | | |
| <i>professional</i> | .600 *** | <i>professional</i> | .601 *** |
| <i>pub</i> | .702 *** | <i>publisher</i> | .690 *** |
| <i>numplatform</i> | .215 *** | <i>numplatform</i> | .232 *** |
| <i>adapt</i> | -.034 | <i>adapt</i> | -.026 |
| <i>sequel</i> | -.022 | <i>sequel</i> | .022 |
| <i>price</i> | .011 ** | <i>price</i> | .012 *** |
| <i>pre_count</i> | .050 *** | <i>pre_count</i> | .049 *** |
| R-Squared: 0.418 | | R-Squared: 0.394 | |
| Adj. R-Squared: 0.411 | | Adj. R-Squared: 0.389 | |

높을 때 트위터 리뷰의 양과 극성의 효과는 모두 감소하는 것으로 나타났다. 이는 비디오 게임의 평가가 ‘호불호’가 갈리게 된다면 트위터 리뷰의 효과에는 부정적인 영향을 미치는 것으로 이해할 수 있다. 이는 가설 4a와 4b를 모두 지지하는 결과이다.

다만 본 분석에서 price 변수의 경우 일반적인 선행 연구를 통해 예상되는 값과 다른 결과가 도출되었다. 일반적으로 가격이 높을수록 수요는 감소한다. 하지만 분석 결과 게임 가격이 증가할수록 게임 판매에 긍정적인 영향을 주는 것으로 나타났다. 이러한 결과가 도출된 원인은 다음과 같이 추정할 수 있다. 먼저 게임의 가격이 제대로 통제되지 못하였을 수 있다. 특히 게임 가격의 경우 그 분포가 매우 치우쳐 있는 것으로 나타났다. 실제로 분석에 사용된 47개 게임 중 26개의 출시 가격이 같으며 게임 가격이 게임 배급사에 의해 일률적으로 책정되어 있다. 이러한 연구 결과가 도출된 다른 이유로는 가격이 제품의 품질을 반영한다는 점을 들 수 있다. 본 연구에서 제시한 모델에서는 제품의 게임성, 또는 품질에 관한 변수를 추가하지 않았기 때문에 가격이 이를 반영하였을 수 있다.

제 5 장 결론 및 논의

제 1 절 연구 결과의 요약 및 시사점

본 연구는 비디오 게임 소비에 중요한 영향을 미칠 것으로 예상되는 온라인 구전 중 트위터 리뷰에 주목하여 소비자들의 트위터 리뷰가 비디오 게임의 성과에 어떠한 영향을 미치는지를 규명하고자 하였다. 총 47개의 게임을 대상으로 트위터 리뷰를 수집하여 이를 패널 회귀 분석에 적용한 결과는 다음과 같다. 첫째, 트위터 리뷰의 수가 비디오 게임의 성과에 미치는 영향을 확인하였다. 분석 결과에 따르면 트위터 리뷰의 수는 판매량에 긍정적인 영향을 미치는 것으로 나타났다. 이는 게임의 리뷰가 많아질수록 소비자들의 게임 선택에 도움이 된다는 것을 의미한다. 둘째, 트위터 리뷰가 긍정적일수록 비디오 게임의 성과에 긍정적인 영향을 미치는 것을 확인하였다. 선행 연구에 따르면 리뷰의 극성은 제품의 특성이나 성격에 따라 그 효과가 달라지는 것으로 알려져 있다. Hennig-Thurau, Wiertz & Feldhaus(2015)의 연구에서는 긍정적인 트위터 리뷰는 영화의 초기 성과에 영향을 주지 못하였지만, 부정적인 트위터 리뷰의 경우 성과에 부정적인 영향을 주는 것으로 나타났다. 본 연구에서는 리뷰의 극성이 미치는 효과를 측정하기 위해 제품에 대한 트위터 리뷰의 감성을 판별하고 긍정적인 리뷰와 부정적인 리뷰의 비율을 측정하여 온라인 구전의 극성을 파악하였다. 그 결과 비디오 게임에 대한 긍정적인 평가가 많을수록 판매량에 긍정적인 영향을

미치는 것을 확인할 수 있었다. 셋째 비디오 게임에 온라인 전문가 리뷰가 영향을 미치는 것을 확인하였다. 전문가 리뷰는 제품의 불확실성을 줄이고 신뢰감 있는 정보를 준다는 점에서 소비자의 의사결정의 긍정적인 영향을 준다. 하지만 영화, 음악과 같은 경험재에서는 전문가의 영향이 제품군에 따라 달라진다. 본 연구 결과를 통해 비디오 게임의 경우 전문가 리뷰가 효과적으로 작용하는 콘텐츠임을 확인할 수 있었다. 마지막으로 온라인 리뷰 감성의 엔트로피가 온라인 리뷰의 효과를 조절하는 것을 본 연구를 통해 확인할 수 있다. 엔트로피가 높아 소비자 리뷰의 의견이 갈릴 때 트위터 리뷰의 양과 극성의 효과는 모두 감소하였다. 이를 통해 온라인 리뷰에서 감성이 어떻게 분포되어 있는지에 따라 리뷰의 효과가 조절됨을 확인할 수 있다. ‘호불호’가 강하여 특정 제품에 대한 전반적인 감성을 파악하기 어려운 경우 리뷰 효과는 감소하게 되는 것이다.

본 논문이 가지는 이론적 의의는 다음과 같다. 먼저 본 연구는 마이크로블로그에서 생성되는 비정형의 온라인 리뷰를 가공하여 텍스트 리뷰가 실제로 어떠한 영향을 미치는지를 확인한 연구라는 점에서 의의를 가진다. 온라인 리뷰의 영향에 대해서는 기존의 많은 선행 연구를 통해 밝혀진 바 있다. 하지만 텍스트로 이루어진 비정형의 트위터 리뷰의 영향력을 마케팅 관점에서 다뤄진 연구는 매우 적은 것이 현실이다. 따라서 본 연구는 비정형의 온라인 리뷰를 분석하는 방법을 제시하였다는 점에서 의의가 있다. 둘째, 본 연구는 트위터의 감성 분석 뿐 만 아니라 감성의 엔트로피를 적용한 연구라는 점에서 의의를 지닌다. 온라인 리뷰의 감성이 어떻게 분포되어 있는지에 대한 연구는 실제

산업에서의 데이터를 토해 다루어진 경우가 비교적 적다. 온라인 구전의 분산이나 분포의 경우 실험을 활용한 연구들이 진행되어 왔는데, 소비자의 제품에 대한 평가가 수렴하는지 또는 발산하는지에 따라 소비자의 반응을 확인한 연구(주우진, 노민정, 2009; 최지은, 여민선, 2017)들이 진행되어 구전의 분산에 대한 이해를 높인 바 있지만, 이를 실제 데이터를 활용하여 분석한 연구는 적은 것이 실정이다. 본 연구는 이같은 점을 보완하고 비디오 게임 데이터를 활용하여 리뷰의 분산을 확인하고자 한 연구라는 점에서 의의를 가진다.

한편 본 연구가 가진 실무적 시사점은 다음과 같다, 첫째, 게임 산업에서의 온라인 구전이 실제로 영향을 미칠 수 있음을 데이터를 통해 확인하였다. 따라서 기업의 소셜 네트워크 서비스들을 이용한 커뮤니케이션 전략에 타당성을 부여할 수 있다. 본 연구는 마이크로블로그를 비롯한 온라인 SNS들이 비디오 게임 구매를 위한 마케팅 도구로 활용될 수 있음을 증명한다. 현재에도 많은 기업들은 소셜 네트워크 서비스를 활용하여 제품과 브랜드에 대한 홍보를 지속적으로 펼치고 있다. 본 연구는 이러한 소셜 미디어 커뮤니케이션 전략이 게임 산업에서도 여전히 유효함을 시사한다. 따라서 게임 개발사와 배급사 또한 SNS를 적극적으로 활용하여 온라인에서 구전을 과생시킬 수 있는 마케팅 전략을 구축하여야 한다. 둘째로 구전의 양을 늘리고, 긍정적인 구전을 발생시키기 위한 환경을 조성해 나가야 한다. 그 방안으로 트위터와 같은 소셜 네트워크 매체와의 연결을 원활하게 하는 게임 시스템을 구축하는 것을 제안할 수 있다. 특히 최근 많은 게임들은 온라인 미디어의 발달과 더불어 게임 내에서 바로 SNS

서비스를 이용할 수 있게 하거나, 게임 플레이 영상 또는 스크린 샷을 바로 공유할 수 있는 서비스들을 제공하고 있다. 이러한 게임 내 시스템은 트위터와 같은 온라인 매체에서의 게임의 구전을 확산시키는 방법이 될 수 있다. 셋째로, 본 연구는 전문가 평점이 긍정적인 영향을 미친다는 점을 통해 게임 산업에서 전문가 또는 미디어의 영향력을 활용할 수 있다는 점을 시사한다. 연구 결과에 따르면 게임 산업의 경우 영화 콘텐츠와는 다르게 전문가와 미디어의 영향력이 유효한 것으로 나타났다. 따라서 게임 배급사는 전문가나 미디어가 제품을 어떻게 평가하는지 또는 이를 어떻게 홍보하는지를 파악하고 이를 관리하기 위한 실무적인 전략을 구축하여야 할 것이다.

제 2 절 연구의 한계점 및 향후 연구과제

본 연구는 트위터 리뷰를 활용하여 비정형의 온라인 리뷰가 실제 제품에 미치는 영향에 대한 이해를 높이고자 하였다. 이와 같은 연구는 다음과 같은 한계점을 가진다. 첫째, 본 연구에서는 게임을 콘솔 게임에 한정하여 연구를 진행하였다. 최근 비디오 게임 산업의 경우 다양한 모바일 디바이스의 발전과 게임 시스템의 진화를 통해 다양한 기기를 통한 게임 플레이가 가능해졌다. 특히 모바일 게임의 경우 그 성장세가 매우 두드러지는 것으로 보인다. 하지만 PC와 모바일 기기를 제외한 콘솔 게임에 한정하여 분석을 진행하였기 때문에 전체 게임 산업에서의 온라인 리뷰의 영향력을 고려하지 못하였을 수 있다. 이를 보완하기

위하여 향후 연구에서는 비디오 게임에 대상을 한정하지 않고 모바일, PC등의 다른 플랫폼 게임들에 트위터 리뷰가 미치는 영향을 검증한다면 게임 산업에서 온라인 구전이 미치는 영향에 대한 보다 구체적인 이해를 도울 수 있을 것이다. 둘째, 모형을 설계하는 과정에서 게임 자체의 품질에 대한 변수를 제시하지 못하였다. 이에 따라 가격이 품질을 반영하는 변수로 작용하여 가격이 상승할수록 게임 판매가 증가하는 결과가 도출되었다. 가격 외에 게임의 성과에 큰 영향을 미칠 것으로 예상되는 마케팅 변수 또는 게임 제작비 등의 변수를 포함하지 못한 점 또한 모형 설계의 한계점으로 뽑을 수 있다. 이와 같은 마케팅 변수들은 실제로 게임의 성과에는 주요하게 작용할 수 있기 때문에 본 연구에서 설정한 모델에 대한 결과값은 한계가 있을 수 있다. 향후 연구에서는 게임의 특성이나 특히 게임의 품질을 반영하는 변수를 추가하고 온라인 구전에 영향을 미치는 기타 다른 마케팅 변수들에 대한 추가적인 설정이 필요할 것으로 보인다. 특히 온라인 구전의 영향변수는 선행 연구를 통해 많이 다루어져 왔기 때문에 게임 산업에서 적용할 수 있는 변수를 확인하여 적용함으로써 보다 심도있는 논의가 가능해질 것으로 기대한다. 마지막으로 트위터 데이터의 수집 과정에 한계점을 들 수 있다. 트위터 데이터는 작성 시점과 수집 시점이 달라 데이터가 완전하게 수집되지 않았을 것으로 판단된다. 수집 시점과 게임 발매 시점이 크게는 2년 정도 차이가 있기 때문에 그 기간 동안 게임 초기에 작성된 트위터 리뷰가 삭제되어 증발하였을 수 있기 때문이다. 이 경우 리뷰의 실제 영향력에 차이를 발생시킬 수 있다. 또한 게임과 관련이 없는 트위터 리뷰를 모두 제거할 수 없었다는 점 또한 데이터의 한계로 들 수 있다. 본 연구는 트위터에서 발생하는 대량의 리뷰를 대상으로 분석을

진행하였기 때문에 수집한 모든 트위터 리뷰를 검토하는 과정이 부족한 것이 사실이다. 따라서 수집된 데이터가 실제 게임의 리뷰와 거리가 있을 수 있기 때문에, 이를 더욱 면밀하게 검토해 볼 필요가 있다. 이와 같은 한계점은 향후 연구를 통해 추가적으로 논의를 진행하여야 할 것이다.

부 록

<Appendix 1> 트위터 리뷰 판별 예시

| Game | Positive probability | text |
|----------------------------------|----------------------|--|
| Uncharted_ The Lost Legacy | 0 | Now i know even before Uncharted: The Lost Legacy came out people already said the game just looked like a tomb raider knock off when really its just bc they were both women, no one called nathan drake a lara croft copy even tho.....same storyline |
| Project Cars 2 | 0 | Yeah it sucks wish I could have the choice to Uninstall the patch for now atleast it's like the titanfall 2, project cars 2 "4k" patches they screwed up the games so bad until they patched there patch hopfully they can fix this... I guess there's a learning curve lol |
| Assassin's Creed Origins | 0 | Is there any reason for me to continue playing Assassin's Creed: Origins? I'm 90 min into it and there's basically no story, slow pace, little action, constant and confusing feedback, a poorly designed UI, and a clunky climbing system. How did they screw this up so badly? |
| WWE 2K18 | 0 | i feel like i wasted 63 on wwe 2k18! im pissed it wont let me do anything til i beat this first match and i cant do it for some reason this is why i dont buy wrestling games they suck! |
| WWE 2K18 | 0.49 | That's fine. I can pause in-between while playing WWE 2k18 lol. Designated coffee break included. |
| Cities_ Skylines | 0.48 | This will be me most of the weekend. Hiding in bed playing cities skylines. Wish they could sort out the traffic AI pic.twitter.com/y1Lu3fFsAh |

| | | |
|-------------------|------|---|
| Cities_Skylines | 0.49 | So I checked my most recent Cities Skylines map, and I own 5 hospitals. I'm over his level! |
| WWE 2K18 | 0.48 | Dear @ 2K – I won't be buying WWE 2K18 on Switch, because you're forcing me to download 24GB as opposed to not being greedy. Just so ya know. |
| Cities_Skylines | 0.49 | just watched someone accidentally gentrify the working class neighbourhood in his city in cities: skylines lol |
| The Evil Within 2 | 0.99 | From what I understand about The Evil Within 2, it's a game about a white dude who uses the knowledge, friendship, and strength of WoC that help him till they die to rescue his daughter that he refers to as The Girl. Top notch gaming. Well done game makers. |
| Yakuza Kiwami | 0.99 | Yakuza Kiwami has put me in a very joyful mood, this is great video gaming, a feel-good experience for sure |
| Call of Duty_WWII | 0.99 | "Call of Duty: WWII returns to its roots with a breathtaking experience that redefines World War II for a new gaming generation." redefines historical inaccuracy too |

<Appendix 2> 상관계수표

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-----------------------|--------|---------|---------|---------|---------|---------|---------|---------|--------|-------|----|
| 1 <i>lnsales</i> | 1 | | | | | | | | | | |
| 2 <i>lnprior</i> | .138** | 1 | | | | | | | | | |
| 3 <i>count</i> | .452** | -.219** | 1 | | | | | | | | |
| 4 <i>PNratio</i> | 0.053 | 0.032 | -.074** | 1 | | | | | | | |
| 5 <i>professional</i> | .320** | .148** | 0.028 | .156** | 1 | | | | | | |
| 6 <i>pub</i> | .136** | .062* | -0.04 | .176** | -0.025 | 1 | | | | | |
| 7 <i>numplatform</i> | .069* | 0.031 | .071* | -.082** | -0.027 | -.425** | 1 | | | | |
| 8 <i>adapt</i> | -.065* | -0.03 | -.065* | .132** | -.198** | .305** | -.184** | 1 | | | |
| 9 <i>sequel</i> | .166** | .076** | .058* | -.072* | .164** | .216** | 0.008 | -.128** | 1 | | |
| 10 <i>price</i> | .110** | -.063* | .176** | -0.035 | -.182** | .170** | -.099** | .294** | .102** | 1 | |
| 11 <i>precount</i> | .232** | .102** | .293** | -.114** | -.141** | .104** | -.201** | -.068* | .159** | 0.035 | 1 |

* 상관계수는 0.05 수준(양쪽)에서 유의

** 상관계수는 0.01 수준(양쪽)에서 유의

참고 문헌

- Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, 53(3), 297-318.
- Borgida, E. and R. E. Nisbett (1977), "The Differential Impact of Abstract vs. Concrete Information on Decisions," *Journal of Applied Social Psychology*, Vol. 7, No. 3, pp. 8-271.
- Chen, Y., Liu, Y., & Zhang, J. (2012). When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews. *Journal of Marketing*, 76(2), 116-134.
- Chen, Y., Wang, Q., & Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research*, 48(2), 238-254.
- Chevalier, J., & Mayzlin, D. (2006). The effect of word of mouth on sales: online book reviews. *Journal of Marketing Research*, 43(August), 345-354.
- Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The

effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, 29(5), 944–957.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing*, 21(4), 23–45.

Ferguson, P., O'Hare, N., Davy, M., Bermingham, A., Sheridan, P., Gurrin, C., & Smeaton, A. F. (2009). Exploring the use of paragraph-level annotations for sentiment analysis of financial blogs.

Hansen, S. S., & Lee, J. K. (2013). What drives consumers to pass along marketer-generated eWOM in social network games? social and game factors in play. *Journal of theoretical and applied electronic commerce research*, 8(1), 53–68.

He, S. X., & Bond, S. D. (2015). Why is the crowd divided? Attribution for dispersion in online word of mouth. *Journal of Consumer Research*, 41(6), 1509–1527.

Hennig-Thurau, Thorsten, Kevin P. Gwinner, Gianfranco Walsh, and Dwayne Gremler (2004), “Electronic Word-of-Mouth Via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?” *Journal of Interactive Marketing*, 18(1), 38–52.

- Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science*, 43(3), 375-394.
- Hodeghatta, U. R. (2013, August). Sentiment analysis of Hollywood movies on Twitter. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining pp. 1401-1404. ACM.
- Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, 60(11), 2169-2188.
- Jin, S. A. A., & Phua, J. (2014). Following celebrities' tweets about brands: The impact of twitter-based electronic word-of-mouth on consumers' source credibility perception, buying intention, and social identification with celebrities. *Journal of Advertising*, 43(2), 181-195.
- Lee, J. H., Jung, S. H., & Park, J. (2017). The role of entropy of review text sentiments on online WOM and movie box office sales. *Electronic Commerce Research and Applications*, 22, 42-52.

- Marchand, A., & Hennig-Thurau, T. (2013). Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities. *Journal of Interactive Marketing*, 27(3), 141-157.
- Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics* (p. 271). Association for Computational Linguistics.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing—Volume 10* (pp. 79-86). Association for Computational Linguistics.
- Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The journal of industrial economics*, 53(1), 27-51.
- Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems*, 55(4), 863-870.
- Ruohonen, J., & Hyrynsalmi, S. (2017). Evaluating the use of internet

- search volumes for time series modeling of sales in the video game industry. *Electronic Markets*, 27(4), 351-370.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3), 379-423.
- Twitter (2018), How Twitter can connect brands with gamers, https://blog.twitter.com/marketing/en_gb/topics/research/2018/Twitter-brands-Gamers.html
- 김유영, & 송민. (2016). 영화 리뷰 감성분석을 위한 텍스트 마이닝 기반 감성 분류기 구축. *지능정보연구*, 22(3), 71-89.
- 김일경(Il Kyoung Kim), 전경미(Kyung Mi Chun) , 이호욱(Houk Lee). (2012). 전문가 리뷰가 인터넷 사용자 리뷰와 박스 오피스(Box Office)에 미치는 영향. *인사조직연구*, 20(3): 1-27
- 임준엽, & 황병연. (2014). 트위터를 이용한 기계학습 기반의 영화 흥행 예측. *정보처리학회논문지. 소프트웨어 및 데이터 공학*, 3(7), 263-270.
- 이정. (2011). 정보이론의 엔트로피 관점에서의 바라본 온라인 소비자 리뷰의 소비자 의사결정에 있어 불확실성 감소 효과. *한국전자거래학회지*, 16(4), 241-256.

Abstract

The Impact of Reviews on Performance of Video Game Sales: Analyzing Twitter Data with Machine Learning

Lee HyunGyung

College of Business Administration

The Graduate School

Seoul National University

Today's online word of mouth has a profound impact on consumers decision processes. This study focuses on Twitter reviews which is expected to have a significant impact on video game consumption. We collected Twitter reviews for 47 games, and then classified them through sentiment analysis using machine learning method. Twitter reviews are text-based, requiring a different approach than the review ratings typically used in previous literature that validates the impact of e-WOM. In order to process the reviews into an analytical format, we applied Bayesian classification and classified the twitter reviews

as either positive or negative to use as research variables.

The result showed that the volume of Twitter reviews has a positive effect on sales. In addition, the more positive the Twitter review is, the more positively it affects the performance of video games. Furthermore, it was confirmed that the effect of online review was moderated by distribution of online review sentiment To find out this result, the study borrowed the concept of entropy used in information theory. High consumer entropy or high uncertainty of consumer review sentiment reduced the effects of both volume and valence on Twitter reviews, suggesting that the distribution of e-WOM can moderate the impact of reviews. This result is meaningful in that it examines the influence of online reviews in the game industry and identifies how consumers' activities in the online media are affecting the actual game industry.

Keywords : video game, online WOM, Twitter review, sentiment analysis, entropy, machine learning

Student Number : 2018-24282