Journal of Digital Convergence
Vol. 16. No. 3, pp. 145-156, 2018

ISSN 1738-1916
https://doi.org/10.14400/JDC.2018.16.3.145

감성분석 기반의 게임 소비자 온라인 구전효과 연구

정근웅[1], 김종욱[2*]
[1]성균관대학교 일반대학원 경영학과, [2]성균관대학교 경영전문대학원

A Study on the Effects of Online Word-of-Mouth on Game Consumers Based on Sentimental Analysis

Keun-Woong Jung[1], Jong Uk Kim[2*]
[1]Graduate School of Business, Sungkyunkwan University
[2]Business School, Sungkyunkwan University

요 약 배급사가 소매점을 통해 게임을 유통했던 과거와 다르게 현재는 디지털 콘텐츠인 게임을 온라인 기반의 유통채널을 활용하여 판매를 실시하고 있다. 본 연구는 온라인 디지털 콘텐츠 유통 채널인 스팀(Steam)에서 판매되는 게임의 판매량에 대해서 eWOM(전자구전효과)의 요인들이 어떤 영향을 미치는지 분석한다. 최근 빅데이터 기반의 데이터 마이닝 기법을 이용한 연구가 많이 진행되고 있는데, 본 연구에서 eWOM의 요인 중 각 리뷰의 감성을 분석할 수 있는 텍스트 마이닝 기법인 감성분석을 실시하여 eWOM의 감성지수를 도출한다. 감성분석은 나이브 베이즈(Naive Bayes)와 지지벡터기(SVM) 분류기를 활용하고, 정확도가 높은 지지벡터기(SVM) 분류기를 통해 감성지수를 산출한다. 도출한 감성지수와 eWOM의 크기인 각 게임의 리뷰의 수, eWOM의 평점인 각 게임의 유저점수를 독립변수로 하여 종속변수인 판매변화량에 대해서 회귀분석을 실시한다. 회귀분석 결과, 독립변수인 eWOM의 크기와 eWOM의 감성지수가 종속변수인 판매변화량에 영향을 미치는 것을 확인하였다. 본 연구는 연구결과를 통해 국내 게임 기업들이 스팀을 기반으로 해외진출 시 판매량에 영향을 미치는 eWOM의 요인들을 제시할 수 있는 시사점을 가진다.

주제어 : eWOM, 텍스트 마이닝, 감성분석, 오피니언 마이닝, 머신러닝

Abstract Unlike the past, when distributors distributed games through retail stores, they are now selling digital content, which is based on online distribution channels. This study analyzes the effects of eWOM (electronic Word of Mouth) on sales volume of game sold on Steam, an online digital content distribution channel. Recently, data mining techniques based on Big Data have been studied. In this study, emotion index of eWOM is derived by emotional analysis which is a text mining technique that can analyze the emotion of each review among factors of eWOM. Emotional analysis utilizes Naive Bayes and SVM classifier and calculates the emotion index through the SVM classifier with high accuracy. Regression analysis is performed on the dependent variable, sales variation, using the emotion index, the number of reviews of each game, the size of eWOM, and the user score of each game, which is a rating of eWOM. Regression analysis revealed that the size of the independent variable eWOM and the emotion index of the eWOM were influential on the dependent variable, sales variation. This study suggests the factors of eWOM that affect the sales volume when Korean game companies enter overseas markets based on steam.

Key Words : eWOM, Text Mining, Sentimental Analysis, Opinion Mining, Machine Learning

*Corresponding Author : Jong Uk Kim(jukim@skku.ac.kr)
Received December 22, 2017
Revised February 28, 2018
Accepted March 20, 2018
Published March 28, 2018

1. 서론

국내 게임 업체 블루홀이 2017년 3월 24일 스팀(Steam)을 통해 발매한 '플레이어언노운즈 : 배틀그라운드(Playerunkown's : Battleground)'는 발매 3일 만에 40만 장 이상이 판매되어 1,100만 달러의 수익을 올리면서[1] 국내 게임 업체 간 게임의 스팀 발매에 대한 관심이 높아지고 있다. 밸브 코퍼레이션(Valve Corporation)에서 개발한 디지털 콘텐츠 멀티 플레이어 플랫폼인 스팀을 통해 전 세계의 게임이 유통되고 있으며 스팀을 통해 게임 판매뿐만 아니라 게임 플랫폼으로써 구매한 게임을 클라우드 서버를 통해 언제 어디서나 실행할 수 있는 장점을 갖고 있다. 스팀 내에서 이용자들은 게임에 대한 평가뿐만 아니라 사용자간 의사소통도 가능하기 때문에 사용자는 구매자가 업로드한 제품 평가에 '유용함'이나 '유용하지 않음'을 표기할 수 있어 제품 구매에 도움이 되는 제품 평가와 각 게임에 전문가들 및 일반 사용자들이 평가한 메타크리틱(Metacritc)점수를 참고할 수 있다. 스팀의 액티브 계정은 약 1억 2천 500만 이상[2]에 달하고 있다.

국내 게임사가 쉽게 해외시장에 진출할 수 있는 스팀(Steam) 유통채널 이용에 대해서 본 연구를 통해 추후 국산 게임의 스팀 유통이 가지는 전자구전효과인 eWOM(Electronic Word-of-Mouth) 효과를 예상하고, 새로운 유통 채널에 대한 불안감을 해소하는 데 도움을 주고자 한다. 스팀은 국내 사용자뿐만 아니라 세계적으로 많은 사용자를 보유하고 있어 국내반응과 동시에 해외반응이 동시에 확인 가능한 채널이다. 따라서 본 연구를 통해 스팀의 국내·외 많은 소비자의 eWOM 효과가 매출에 미치는 영향을 분석하려 한다.

선행연구와 비교하였을 때 가지는 본 연구의 공헌점은 다음과 같다. 첫째, eWOM가 스팀을 통해 유통되는 게임의 흥행과 판매에 어떤 영향을 미치는 지에 대해서 분석 한다. 둘째, 감성 분석에서 사용하는 대표적인 두 가지 알고리즘인 나이브베이즈와 지지벡터기를 사용하여 도출한 각 감성분석 결과에 대한 비교분석을 실시한다. 셋째, 감성분석을 통해 산출한 감성지수를 독립변수로 하여 매출과 같은 종속변수에 영향을 미치는 지에 대한 회귀분석을 실시한다.

2. 이론적 배경

2.1 eWOM

구전(Word-of-Mouth)이란 소비자 간 제품에 대한 긍정적 혹은 부정적 정보를 공유하는 행위로 상업적 이익을 고려하지 않은 비공식적 의사소통으로[3], eWOM의 기반이 되는 개념이다. 전통적인 마케팅 수단인 구전은 사용자가 경험하지 못하거나 내용에 대해서 잘 모를 때 구매 결정에 있어 많은 영향을 미친다[4]. 구전이 소비자의 구매 의사 결정에 미치는 긍정적인 효과는 다음과 같다. 첫째, 구전효과는 기업의 마케팅 비용의 절감을 가능하게 한다. 둘째, 마케팅에서의 구전은 제품이나 서비스에 대한 소비자의 태도를 긍정적으로 변화시킬 수 있으며 충성 고객 확보가 가능하다. 셋째, 기업이 신제품을 출시하였을 때, 구전효과를 활용한 마케팅은 소비자의 구매 의사에 대한 불안을 감소시킬 수 있어 소비자의 구매 의사를 불러일으킬 수 있는 중요한 전략으로 사용될 수 있다[5].

eWOM은 온라인상의 소비자 간 제품 및 서비스에 관하여 긍정이나 부정적인 정보를 교환하는 비공식적인 의사소통 행위로 구전과 마찬가지로 상업적 이익을 목표로 하지 않은 공통점이 있는 반면, 전통적인 구전과 비교할 때 여러 차이점이 있기 때문에 연구자는 다음과 같은 차이점을 인지하여 연구를 진행해야한다.

첫째, 전통적인 구전의 경우에 특정 장소에서 소비자 간 대화를 통해 형성되는 반면, eWOM의 경우 시간과 공간에 대한 제약이 없을뿐더러 개인단위의 효과에서 그치지 않고 집단단위에서 정보공유가 이루어진다[6]. 둘째, eWOM은 텍스트로 저장이 되기 때문에 웹사이트에 기록이 남아있는 이상, 과거에 공유된 정보를 언제든지 확인할 수 있다는 장점이 있다[7]. 셋째, eWOM은 전통적인 구전과 다르게 보존가능성이 높기 때문에 오프라인정보 보다 더 많은 양으로 증가하였고, 인터넷의 발달로 많은 양의 데이터 중에서 연구자가 원하는 특징만을 분석 가능하게 되었다[8]. 넷째, 대면 접촉을 통해 상호작용을 하던 전통적인 구전과 다르게 eWOM은 온라인상에서 발생 가능한 의견을 공유하는 소비자 사이에서 상대적으로 약한 유대감을 형성시키기 때문에[9], 정보를 받아들이는 소비자가 익명의 상대로부터 받은 제품 정보에 대해 전통적인 구전에 비해 낮은 신뢰감을 갖게 된다[10].

온라인상에서 소비자 간 발생하는 구전효과인 eWOM

이 다양한 채널에서 발생되고 있는데[11], 최근에는 사용자가 쉽게 글이나 사진, 그리고 동영상을 업로드할 수 있는 유튜브(Youtube), 페이스북(Facebook), 블로그, 트위터 등과 같은 소셜 네트워크 서비스에서 활발하게 eWOM이 발생하고 있으며, 이는 소비자 간 쌍방향 의사소통 비용을 감소하게 만들어 eWOM의 파급효과는 이전보다 더 커지고 있다[12]. 또한, 온라인 환경에서 소비자들은 제품이나 서비스에 대해 자신이 겪은 경험이나 제품의 불만사항을 공유할 뿐만 아니라, 구매를 주저하고 있는 소비자와 의견공유를 통하여 구매를 일으키거나 새로운 내용을 전파하게 된다. 이러한 경향은 eWOM이 소비자의 구매의사결정에 전통적인 구전에 비해 더 많은 영향을 주고 있다는 것으로 볼 수 있다[13].

게임산업은 미래산업의 특징이라고 할 수 있는 환경오염 문제로부터 자유롭고, 지식집약적 고부가가치 산업으로[14], 한국콘텐츠진흥원에서 발표한 2013년 보고서에 따르면 국내 콘텐츠 산업 수출액 약 5조 2,017억 원 중 62%에 해당하는 3조 2,079억 원을 게임수출이 차지하고 있는 것으로 밝혔다. 최근, 스마트폰의 접속성 및 편리성에 대한 연구[15]에서 동영상, 음악, 게임 등 멀티미디어 기능이 중요 요소임을 밝혔고, 이러한 스마트폰을 사용한 소셜 네트워크 게임에 대한 연구도 지속적으로 진행되고 있다[16, 17]. 게임은 경험재에 속하기 때문에 사용자가 게임을 이용하기 전까지 게임의 내용이나 질을 평가하기 어렵다. 또한 게임의 생산비용 및 수요 규모, 기업의 이윤과 관련 없이 이용 가격이 어느 정도 고정되어 있어 게임 사용자의 추가적인 수요가 발생하지 않은 다면 게임 제작비용 회수에 어려움이 있을 수 있는 고위험 고수익 산업이다[18]. 따라서 게임을 이용하지 못한 사용자들에게 게임개발업체의 광고 및 홍보가 아닌 게임을 경험했던 사용자들의 구전정보가 매우 중요하다고 할 수 있는데, 기존 사용자들은 게임 판매 및 운영 기업이 아닌 사용자 간에 구전시키기 때문에 부정적 구전정보는 서비스 확산에 매우 부정적 영향을 줄 수 있다[19].

2.2 감성분석

2.2.1 텍스트 마이닝

텍스트 마이닝(Text-Mining)은 텍스트라는 온라인상에 등재된 게시글이나 댓글 등의 비정형 데이터에서 연구자가 원하는 정보만을 유용하게 사용하는 것으로[20], 키워드가 아닌 맥락 수준에서 의미 있는 데이터를 분석하고 가치 있는 정보를 도출한다. 이를 위해 텍스트의 자연어 처리 기술과 형태소 분석 기술을 사용하며 출현 빈도와 확률 분석을 통해 핵심 단어를 추출한다[21]. 이러한 텍스트 마이닝의 사용범위는 정보 검색, 정보 추출, 텍스트 범주화, 텍스트 군집화, 텍스트 요약뿐만 아니라 기계 학습, 통계학, 자연어 처리, 데이터 마이닝 등 다양한 분야에서 사용되고 있다[22]. 텍스트 마이닝의 한 분야인 텍스트 범주화는 단어, 구, 문장, 문단 등과 같이 구조적이지 않은 비정형 텍스트 데이터를 범주화하는 것으로[20], 감성분석은 텍스트에 내포된 감성 또는 의미를 '긍정'과 '부정'으로 극성을 분류하는 것에서 '기쁨', '슬픔', '지루함' 등 다양하게 범주화할 수 있다[23].

2.2.2 감성분석

감성분석은 오피니언 마이닝(Opinion mining)이라고도 하며 온라인상 비정형 데이터를 대상으로 자연어 처리와 텍스트 내의 '긍정'과 '부정'을 포함한 다양한 감정을 식별할 수 있는 컴퓨터 언어처리를 위해 사용된다[24]. Nasukawa & Yi은 감성 분석이 주관적인 의미를 갖고 있는 텍스트가 긍정인지 부정인지를 판단하는 것으로 정의하였다. 감성분석을 실시하기 위해 자연어 처리와 텍스트의 빈도 분석, 기존의 감성 사전이나 지도 학습 기반으로 감성 점수를 부여하는 SentiWordNet와 같은 감성사전을 사용한다[25]. 감성분석은 시장규모를 예측하거나 제품 출시에 대한 소비자의 반응정도, 그리고 구전정보 분석 등과 같이 다양하게 활용되고 있지만, 대부분 긍정과 부정을 판별하는 목적으로만 연구 및 활용되고 있어 단순한 긍정, 부정, 그리고 중립 의견 이상의 감성이나 평가 내용을 분석하기 어렵다는 문제가 있다[26].

2.2.3 감성사전

감성사전은 긍정적인 의미의 단어와 부정적인 의미의 단어를 구분해 놓은 단어사전으로, 가장 많이 이용되는 사전은 Hu & Liu의 영문 감성어휘사전인 Opinion Lexicon과 Esuli & Sebastiani의 'SentiWordNet' (https://sentiwordnet.isti.cnr.it/)이 있다[27, 28]. Opinion Lexicon의 경우 긍정단어 2,006개와 부정단어 4,783개로 구성되어 있으며 단어를 단순히 긍정과 부정으로 분류한다. 반면 SentiWordNet의 경우 단어를 명사, 동사, 형용사, 부사로 구분하여 품사마다 긍정, 부정뿐만 아니라 중립으로 나누어 각 단어의 극성에 대한 점수를 부여한다.

SentiWordNet의 품사 별 단어의 개수는 명사 82,115개, 동사 13,767개, 형용사 18,576개, 부사 3,621개이며 각 품사별 감성 값을 통해 문장의 극성을 계산할 수 있다. 각 단어에 대한 극성 점수는 감성이 긍정일 경우, 0에서 1사이의 값으로 계산되며 감성이 부정일 경우 -1에서 0사이의 값으로 계산된다. 1에 가까울수록 긍정적 성향이 강하며 -1에 가까울수록 부정적 성향이 강하다고 볼 수 있다. 0의 경우 단어의 감성이 중립이란 것을 의미한다. 각 단어 집단에는 동의어나 유의어가 하나의 그룹으로 지정되어 있으며 기계 학습에 의해 구축된 SentiWordNet은 기존의 WordNet의 데이터를 기반으로 현재 가장 큰 감성사전이라고 볼 수 있다[29]. 이러한 감성사전 기반의 감성값 합계와 평균 그리고 빈도수뿐만 아니라 문서의 특성에 맞는 가중치를 부여하여 분석하기도 한다[30]. 본 연구에서 사용자 리뷰 텍스트 데이터에 담긴 감성이 긍정인지 부정인지에 대한 극성을 분류하기 위해서는 SentiWordNet 감성사전 기반으로 하는 감성 점수를 이용한다.

3. 연구 모형 및 가설 설정

3.1 연구모형

본 연구에서는 Fig. 1, Fig. 2와 같은 연구 모형을 제시한다. Fig. 1은 본 연구에서 진행할 감성분석 모형으로 스팀 웹사이트로부터 각 게임에 해당하는 온라인 리뷰 데이터를 수집하고 전처리 과정을 거쳐 감성사전 데이터와 알고리즘을 이용하여 샘플마다 감성분석 데이터를 산출한다. Fig. 2는 본 연구의 회귀분석을 진행하기 위한 모형으로, 모형을 기반으로 eWOM요인들과 감성분석을 통해 도출한 각 샘플의 감성수치들이 종속변수인 게임 판매량(판매 변화량)에 미치는 영향을 알아보기 위해 회귀분석을 실시한다.

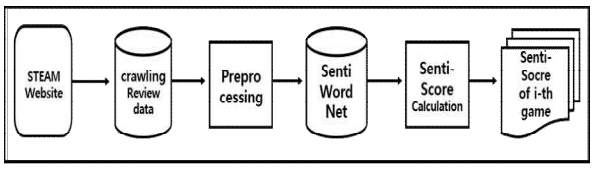

Fig. 1. Sensitivity Analysis Process

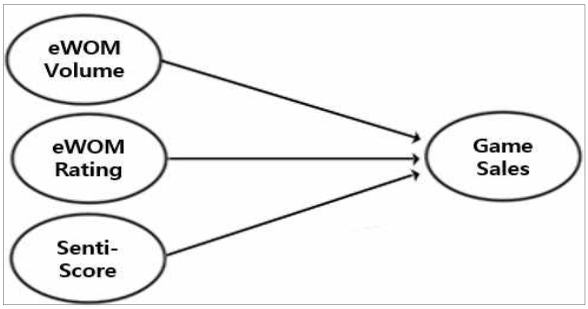

Fig. 2. Research Model

3.2 가설설정

온라인상에서 사용자가 제품을 구매 및 사용 후의 경험이나 느낌을 기록한 소비자 제품 리뷰에 대한 연구들[31, 32]에서 공통적으로 소비자들은 구매 결정전에 온라인상에 존재하는 사용자의 리뷰로부터 유의미한 정보를 얻을 수 있으며 이를 효과적으로 반영한다는 것을 밝혔다[33]. 스팀에서 각 게임에 대한 리뷰를 남기기 위해선 사용자가 게임을 직접 구매해야 하기 때문에 구매 시 각 상품 평가를 참고하여 구매에 대한 의사결정을 진행하게 된다. 또한 게임이라는 디지털 매체 거래의 경우 상품의 가치를 직접 평가하기 어려워 다른 구매자의 상품 평가에 의존하는 경향이 크다[34]. 따라서 본 연구에서 다루는 디지털 매체로서 게임이 구전효과에 대해 잘 반영된다고 보았다. eWOM는 온라인상의 구전효과로 시간 및 공간의 제약에서 벗어나 소비자의 구매 결정에 매우 큰 영향을 차지하는 채널로 관심을 받고 있다[34]. 눈에 보이는 재화뿐만 아니라 서비스 그리고 디지털 매체의 거래에서 구매자가 eWOM의 간접 정보만을 통해 구매자가 제품에 대한 직접 경험 없이 제품 구매에 대한 의향을 높일 수 있다[35]. 또한, 구매자는 사용자가 남긴 제품리뷰가 유용한지 아닌지에 대한 평가가 가능해지면서 제품 리뷰의 유용성도 높아지고 있다.

최근 온라인상의 구전효과 eWOM을 주제로 하는 연구가 다양한 분야에서 진행되고 있으며, 드라마나 영화와 같은 미디어 중심의 연구가 주로 진행되고 있다[36, 37]. 기술의 발달로 게임은 음악, 이미지, 영상, 텍스트 등이 공존하는 동시에 기술이 의존하는 비율이 높은 뉴 미디어로써 몰입도가 높아 역동적인 이미지와 감각적 상호작용이 가능하다[38]. 따라서 본 연구는 기존의 미디어 중심의 eWOM 연구를 기반으로 뉴미디어인 게임에 적용하여 가설 설정 및 분석을 실시한다. 과거의 미디어 기반 eWOM 선행연구는 주로 매출액, 이익, 매출액 순위와

같은 정량화 지수를 종속변수로 사용하였고 독립변수로 댓글(리뷰)의 개수와 평점(Rating), 그리고 최근에는 텍스트 마이닝 기법을 사용하여 댓글 또는 블로그 게시글과 같은 비정형데이터를 대상으로 한 감성분석을 사용하여 연구가 진행되고 있다[39].

3.2.1 eWOM 크기

eWOM의 크기는 온라인 구전의 발생 정도로 상기의 선행연구에서 공통으로 드라마 리뷰의 수나 댓글 혹은 리뷰의 개수로 측정하고 각 연구에 있어 수익이라고 할 수 있는 종속변수에 긍정적인 영향을 미치는 것으로 밝혔다. Liu의 연구에서도 eWOM의 크기가 수익에 긍정적인 상호작용이 작용한다는 것을 밝혔다[40]. 따라서 본 연구는 각 게임에 해당하는 리뷰의 숫자가 종속변수인 게임소지자 수에 미치는 영향을 연구한다.

가설1. eWOM 크기는 게임 판매변화량에 정의 영향을 미친다.

가설 1-1. 긍정적 eWOM의 크기는 게임 판매변화량에 정의 영향을 미친다.

가설 1-2. 부정적 eWOM의 크기는 게임 판매변화량의 부의 영향을 미친다.

3.2.2 eWOM 평점

평점은 사용자가 사용한 제품이나 경험한 서비스에 대하여 정량적인 수치로 평가되는 것으로, 소비자가 제품 리뷰를 접함으로써 상품에 대한 정보를 획득하고 이는 상품의 가치를 만들며 평점은 소비자의 상품 선택에 있어 선택의 기준이 된다[41]. Sawhney & Eliashberg은 영화 리뷰 평점이 영화 매출에 있어 긍정적인 효과를 가진다는 것[42]을 주장한 반면 Liu은 실증연구를 통하여 평점이 매출에 아무런 영향을 미치지 못한다는 것을 밝혔다[39]. 하지만 Duan et al.은 평점과 수익 간 관계에 대해 연구를 하였는데, 상품이 좋아 평점이 높은 것인지, 평점이 높아서 상품이 좋은 것인지 구분하기 어렵다는 것을 밝혔다[43]. 본 연구에서 게임 구매 시 소비자가 고려할 수 있는 각 게임의 평점이 게임 판매에 어떠한 영향을 미치는지 밝히고자 한다.

가설 2. eWOM 평점은 게임 판매변화량에 정의 영향을 미친다.

3.2.3 지도학습을 통한 감성분석 연구

Pang et al.의 연구는 지도학습 방법을 이용한 대표적인 감성분석 연구로 영화 리뷰 자료를 이용하였다. 나이브 베이즈, 최대 엔트로피, 지지벡터기 세 가지 학습분류기를 이용하여 리뷰 평점이 5점 만점일 때, 4 - 5점일 때를 긍정, 1 - 3점일 때를 부정으로 나누어 학습하여 결과를 비교하였고 가장 높은 성능은 지지벡터기 분류기임을 밝혔다[44]. 나이브 베이즈(Naive Bayes)는 각 데이터의 요소들이 상호 독립이라는 가정을 기반으로 통계적 확률을 추정하여 학습 한 뒤 실험 데이터를 확률이 높은 방향으로 분류한다. 나이브 베이즈 알고리즘의 경우 학습시간이 높은 단점에 비해 통상적으로 높은 정확도를 갖고 있으며, 학습 데이터의 양과 정확도가 비례한다[45]. 지지벡터기(Support Vector Machine, SVM)는 1979년 Vapnik에 의해 기계학습 분야 중 하나의 분류 모델로 제안된 기법이다. 지지벡터기는 각 데이터를 두 집단으로 분류시키는 초평면(Hyperplane)을 찾는 알고리즘으로, 다른 부류 사이에 존재하는 margin을 최대화하여 오차를 최소화하도록 설계되었다[46]. 지지벡터기는 기계학습에 필요한 조정 변수가 적어 비교적 간단하게 준비를 할 수 있을 뿐만 아니라 수학적 식에 근거하여 데이터를 분류하기 때문에 지속적으로 텍스트 마이닝 분야에서 연구되고 있다[47].

Pak & Paroubek의 연구에서는 확률 언어 모델인 n-gram과 품사 태깅(Part-of-speech tagging, POS tagging)으로 키워드를 추출하였을 때 더 높은 정확도를 낼 수 있다는 것을 밝혔으며[48], Kumar et al.은 트위터 데이터를 통해 감성분석을 진행하였을 때 이모티콘과 같은 감성에 가중치를 부여하여 더 정확한 감성분석을 실시하였다[49]. 황보현우 & 김종혁은 댓글의 수, 평점 그리고 텍스트 마이닝을 동시에 수행한 연구를 진행하였고 eWOM의 사용자 리뷰의 감성분석이 한국 드라마 수출에 긍정적인 작용을 한다는 것을 밝혔다[38]. 본 연구에서는 지지벡터기(SVM)를 활용한 지도학습 방법으로 스팀 내 각 게임 상점페이지마다 게임 소지자가 작성 가능한 리뷰의 감성적인 측면, 즉 텍스트로 나타낸 리뷰가 실제 게임 판매변화량에 어떠한 영향을 미치는지를 분석하고자 한다.

가설3. 사용자 제품 리뷰의 감성 지수는 게임 판매변화량에 영향을 미친다.

가설 3-1. 사용자 제품 리뷰의 긍정 감성 지수는 게임 판매변화량에 정의 영향을 미친다.

가설 3-2. 사용자 제품 리뷰의 부정 감성 지수는 게임 판매변화량에 부의 영향을 미친다.

4. 연구 방법

4.1 자료의 수집 및 측정

본 연구는 eWOM이 게임의 판매변화량에 미치는 영향을 분석하기 위한 것으로 상기의 3개의 가설과 감성 분석 과정을 바탕으로 설계한 연구 모형에 따라 진행한다. 게임에 대한 스팀 사용자의 리뷰는 웹 사이트 스팀에서 파이썬(Python)을 이용한 웹크롤러로 수집하고[50], 품사 태깅과 토큰화 작업 등 자연어처리를 거치고, SentiWordNet감성 사전을 기반으로 도출한 각 감성 값을 사용하여 독립변수인 감성지수를 산출한다. 이후 각 i번째 게임의 eWOM크기와 평점, 그리고 스팀 사용자 리뷰 감성지수가 종속변수인 게임 판매변화량에 미치는 영향을 분석한다. 각 독립변수에 대한 측정은 Table 1을 따른다.

Table 1. Independent Variable Measurement

Independent Variable Measurement	
eWOM Volume	User reviews for i-th game
eWOM User Rating	OpenCritic score of i-th game
eWOM User Sentiment Score	User review emotion index of i-th game

4.2 데이터 처리 과정

본 연구에서는 텍스트 마이닝 기법 중 감성분석을 이용하기 위해 수집한 비정형 데이터를 기반으로 감성분석을 하고자 한다. 감성분석 모델 구축을 위해 스팀에서 제공하는 게임 리뷰 9,291개의 데이터를 학습용 데이터로 이용한다. 감성분석 모델은 전처리, 사전 및 어휘 리스트의 구축, 키워드 추출, 그리고 기계학습을 통한 분류작업을 거쳐 구축된다[51]. 학습용 데이터의 전처리 과정(Preprocessing)에서는 문장 내 공백을 구분자로써 문장을 나누는 토큰화(Tokenization)와 각 단어의 품사를 결정하는 품사 태깅 단계로 나눠진다. 사전과 어휘 리스트의 구축에서 사전의 경우 단어의 단순한 나열인 말뭉치(Bag of words)로 구성되며, 단어 리스트의 경우 단어의 출현 횟수를 내림차순으로 정렬된 것으로 구성된다. 전처리 과정 후 품사 태깅 작업이 끝난 텍스트 데이터에 대한 명사와 동사 리스트를 구축하고, 긍정 형용사와 부사, 부정 형용사와 부사 각각의 단어리스트를 구축한다. 품사별로 분류된 텍스트 데이터를 감성사전인 SentiWordNet을 이용하여 리뷰에 대한 감성 값을 계산한다. 이후 지지벡터기 분류기(SVM)를 통하여 분류과정을 거쳐 감정 분류에 대한 기계학습을 실시한다. 지지벡터기 분류기(SVM)으로 학습된 데이터를 토대로 하나의 게임당 최대 4,000개의 리뷰 데이터에 대한 감성지수를 산출한다.

5. 연구 결과 및 분석

5.1 감성분석 정확도 분석

본 연구에서 기계 학습을 과정을 통해 대량의 데이터에 대한 감성 지수를 산출하기 위해 파이썬을 사용하였다. 긍정리뷰와 부정리뷰가 비슷한 비율로 있는 게임을 선정하여 총 9,291개의 리뷰를 학습용 데이터(training set)로 설정하였고, 평가용 데이터(test set)의 경우에는 2016년에 발매한 게임 90여 개에 대한 각각의 리뷰 데이터로 설정하였다. 각 게임의 리뷰 데이터의 경우 최대 4,000개로 제한하였다. 학습용 데이터를 사용하여 지지벡터기(SVM)로 기계 학습을 통해 수행하고 2016년에 발매한 게임 각각에 지지벡터기(SVM) 분류기를 통해 감성지수를 산출하였다. 지지벡터기(SVM)의 분류 기법은 데이터가 선형으로 분류되는 경우와 분류되지 않은 경우가 다른데, 본 연구에서는 비선형 데이터를 사용하기 때문에 데이터 일부를 고차원으로 바꾼 뒤 두 그룹으로 분류 가능한 평면을 찾아야 한다. 이러한 과정에서 비선형 분류 알고리즘은 커널 함수를 사용하여 분류 가능한 초평면을 찾는다. 정확도 비교를 위해 학습용 데이터에 나이브 베이즈를 통해 기계 학습을 수행하고 평가용 데이터에 적용하여 감성지수를 산출하였고 RMSE(Root Mean Square Error)를 사용하여 Fig. 3에서 90개의 게임 중 15개의 게임에 대한 분류기 간 정확도 값의 비교를 나타내었다[52]. RMSE값이 작을수록 높은 정확도를 나타낸다. 지지벡터기(SVM)가 전반적으로 높은 정확도를 보여주었다.

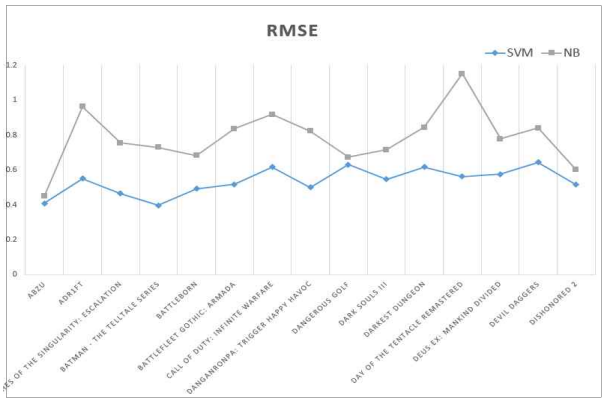

Fig. 3. Sentimental Analysis Accuracy

5.2 회귀 분석

산출된 각 90개의 게임에 대한 감성지수는 긍정 리뷰 점수(*P.sen*)와 부정 리뷰 점수(*N.sen*)로 나눠 독립변수로 설정한다. eWOM의 크기(*Vol*)는 각 샘플에 대한 긍정 리뷰 수(*P.Vol*)와 부정 리뷰 수(*N.Vol*), 그리고 eWOM의 평점(*U.Rat*)인 각 샘플에 대한 유저 스코어(User Score)를 추가적으로 독립변수로 설정하여 종속변수인 게임 소지자의 수(*Owner*)의 변화량에 영향을 미치는 지에 대한 회귀 분석을 진행한다. 이 때, eWOM의 크기와 감성지수인 부정 리뷰 점수와 긍정 리뷰 점수, 그리고 게임 소지자의 수는 게임 별 편차가 크고, 정규 분포를 이루지 못하기 때문에 로그 처리를 한다[39]. 본 연구에서의 가설 검정을 위해 아래와 같은 회귀 식을 도출하였다.

$$\ln_e(Owner) = b_0 + b_1\ln_e(P.Vol) + b_2\ln_e(N.Vol) + b_3(U.Rat) + b_4\ln_e(P.sen) + b_5\ln_e(N.sen) + b_6(E.Rat) + \varepsilon$$

분석을 시행하기 위해 IBM의 SPSS v25를 사용하여 Table 2와 같이 기본 통계량을 구하였다. 또한, 각 변수 간 다중공선성 문제를 방지하기 위해 Table 3과 같은 상관분석을 실시하였다. 마지막으로 각 독립변수가 종속변수인 게임 소지자의 수에 영향을 미치는 지에 대해 영향을 알아보기 위해 다중 회귀 분석을 실시한 결과, Table 4와 같은 결론을 도출하였다.

Table 2. Descriptive Statistics Analysis

Variables	Sample	Average	Standard	Minimum Value	Maximum Value
ln(P.Vol)	90	7.75	1.38	3.29	11.10
ln(N.Vol)	90	6.09	1.46	2.30	9.87
U.Rat	90	68.49	14.30	26.00	86.00
ln(P.sen)	90	8.83	.87	6.02	10.10
ln(N.sen)	90	6.20	1.63	-.47	8.98
ln(Owner)	90	11.16	1.37	4.26	13.46

Table 3. Correlation Analysis

	P. Vol	N. Vol	U. Rat	P. sen	N. sen	Owner
P.Vol	1	.698	.115	.704	.473	.629
N.Vol		1	-.392	.406	.736	.391
U.Rat			1	.354	.248	.180
P.sen				1	.403	.453
N.sen					1	.139
Owner						1

Table 4. Regression Analysis

Dependent	Independent	Standard Error	Beta	t-value	P	VIF
ln (Owner)	Constant	1.226	-	4.596	.000	-
	ln(P.Vol)	.161	.537	3.309	.001	3.984
	ln(N.Vol)	.177	.278	1.473	.145	5.376
	U.Rat	.011	.143	1.209	.230	2.114
	ln(P.sen)	.204	.031	.240	.811	2.545
	ln(N.sen)	.106	-.297	-2.354	.021	2.404
	$R = .667 \quad R^2 = .445$ $Adjusted\,R^2 = .412$ $F = 13.467 \quad P^F = .000$ $Dubin - Watson = 2.253$					

게임 소지자 수에 미치는 영향에 대한 회귀모형의 분산분석에 대한 F분포는 유의확률 .000에서 13.497의 수치를 보이며, 회귀 식에 대한 수정된 R제곱은 .416로 41%의 설명력을 보인다. Dubin -Watson 값은 2.253로 잔차 간 상관관계가 없어 회귀모형이 적합한 것으로 나타나고 있다. 5개의 독립변수 중 2개의 독립변수가 종속변수에 유의한 영향을 주는 것으로 보이며, 로그 처리된 긍정적 eWOM의 크기인 P.Vol 변수(t=3.309, p<.001)와 로그 처리된 부정리뷰의 감성지수인 N.sen 변수(t=-2.354, p<.021)가 종속변수 게임 판매변화량인 Owner에 유의한 영향을 미친다. 또한 종속변수에 영향을 미치는 독립변수의 중요도를 판단하기 위해 유의한 독립변수들의 회귀계수 베타(β)값을 비교한 결과 긍정적 eWOM 크기의 베타 값은 .537(β = .537), 부정리뷰 감성지수의 베타 값은 -.297(β = -.297)으로 긍정적 eWOM의 크기는 게임 소지자 수에 정의 영향을 미치는 반면, 부정리뷰의 감성지수는 상대적으로 음의 영향을 미치고 있다. 본 연구의 결과를 요약하자면 다음의 Table 5와 같다.

Table 5. Summary of Results

Summary of Results	
H1	-
H1-1	significant
H1-2	insignificant
H2	insignificant
H3	-
H3-1	significant
H3-1	insignificant

5.3 분석결과

5.3.1 가설1 검정 논의

본 연구에서는 eWOM의 크기가 커질수록 게임 판매변화량에 정의 영향을 미치는 가설1이 채택되었으며, 리뷰의 방향성을 고려하여 긍정적인 eWOM의 크기와 부정적인 eWOM의 크기를 각각 독립변수로 종속변수인 게임 판매변화량에 미치는 영향에 대해 분석한 결과, 긍정적인 리뷰의 수가 판매변화량에 긍정적인 영향을 미치는 것으로 나타났다. 이는 상기의 선행연구와 같이 eWOM의 크기가 커지면 매출에 정의 영향을 미치는 결과보다 더욱 구체적인 결과로, 이용자들은 부정적인 리뷰나 댓글보다는 긍정적인 리뷰나 댓글의 수를 보고 구매를 결정하는 것으로 보인다.

5.3.2 가설2 검정 논의

가설2는 기각되어 eWOM의 평점이 게임 판매변화량에 유의한 영향을 미치지 못하는 것으로 나타났다. 본 연구에서 eWOM의 평점이 판매 현황만을 고려한 것이 아니고 게임의 작품성, 예술성, 그리고 호환성 등을 모두 고려한 것이기 때문에 평점이 높다고 하여 해당 작품의 흥행과 직접적인 관련이 없는 것으로 보인다.

5.3.3 가설3 검정 논의

가설3의 경우에는 사용자 리뷰의 감성지수가 게임 판매변화량에 영향을 미치는 것으로 나타났다. 하지만 가설 3-1의 긍정적인 리뷰의 경우 게임 판매변화량에 유의한 영향을 미치지 못 하였고, 가설 3-2의 부정적인 리뷰만이 게임 판매변화량에 유의한 음의 영향을 미치는 것으로 나타났다. 이는 구매자가 부정적인 리뷰를 통해 제품의 단점을 고려하여 구매를 하지 않은 것으로 볼 수 있다. Coovert와 Reeder는 구전정보에 있어 긍정적인 구정보보다 부정적인 정보가 구매의도에 더 큰 영향을 미친다고 밝혔다[53]. 본 연구에서는 구매자가 리뷰를 통해 제품의 긍정적인 면을 주의 깊게 보는 것보다 부정적이고 비판적인 글과 함께 제품의 단점을 더 주의 깊게 보며 이러한 리뷰가 구매에 직접적인 영향을 미치는 것으로 추측한다.

6. 결론

6.1 연구의 결론

본 논문에서는 감성분석을 실시하기 위해 약 1만건의 사용자 리뷰 데이터를 기반으로 기계학습을 실시하여 90개의 게임의 각 리뷰에 대한 감성분석을 실시하였다. 감성분석은 긍정적인 리뷰와 부정적인 리뷰로 나누어 각각 따로 감성지수를 산출하였고 지지벡터기(SVM) 분류기와 나이브 베이즈(Naive Bayes) 분류기를 사용하여 정확도를 비교하였다. 이후 감성분석을 실시하여 도출된 긍정리뷰와 부정리뷰에 대한 감성지수와 리뷰의 방향성을 적용한 eWOM의 크기와 평점을 독립변수로 하여 종속변수인 게임 판매변화량에 미치는 영향을 알아보기 위해 회귀분석을 실시하였다. 감성분석 정확도 비교는 RMSE(Root Mean Square Error) 수치를 사용하였고 지지벡터기(SVM) 분류기가 나이브 베이즈(NB) 분류기보다 전반적으로 높은 정확도를 보였다. 회귀분석을 실시한 결과로는 크게 3개의 가설 중 2개의 가설인 eWOM의 크기와 eWOM의 감성지수가 판매변화량에 유의한 영향을 미치는 것으로 나타났다. 그 중 긍정적인 eWOM의 크기가 게임 판매변화량에 정의 영향을, 부정 리뷰 감성지수가 게임 판매변화량에 부의 영향을 미치는 것을 확인하였고 긍정적인 댓글의 리뷰 숫자가 게임 판매량에 가장 큰 영향을 미치는 것을 확인하였다.

6.2 연구의 시사점

본 연구의 시사점은 다음과 같다.

첫째, 본 연구에서는 eWOM가 스팀을 통해 유통되는 게임의 판매변화량에 미치는 영향을 분석하였다. 본 연구에서 eWOM를 기존에 사용된 척도인 크기와 평점뿐만 아니라 감성분석을 통해 산출된 리뷰의 감성지수를 포함하여 분류하였다. 이 과정에서 리뷰의 감성점수만을 산출한 것이 아닌 긍정 리뷰와 부정 리뷰로 분류하여 각각이 갖는 감성지수가 게임 판매변화량에 어떤 영향을 미치는지 분석하여 텍스트 마이닝 기법을 활용한 방법이

eWOM 연구를 확장시킬 수 있다는 학술적 시사점을 가진다.

둘째, 최근 자연어처리기술이 발전하여 뉴스, 서적, 온라인 리뷰 등에 대한 감성분석을 활용한 연구가 진행되고 있다. 다양한 알고리즘을 활용하여 더 정확한 감성분석을 실시할 수 있음을 시사하였고 영어뿐만 아닌 한국어로 되어있는 데이터도 정확한 감성분석이 가능하다는 것을 밝혔다[51]. 이와 같이 더 높은 정확도를 가진 분류기를 소개하는 연구와 달리 도출된 결과값을 변수로 활용하여 종속변수에 영향을 미치는 지에 대한 연구는 잘 이루어지지 않고 있어, 본 연구에서는 기계학습을 통한 감성분석에서 정확도 분석 기준이 되는 지지벡터기(SVM)와 나이브 베이즈(Naive Bayes)를 사용하여 감성지수를 도출하였고 이를 독립변수로 설정하여 다른 독립변수와 함께 종속변수에 대한 다중회귀분석을 진행하였다. 본 연구는 감성분석으로 도출된 결과 값을 독립변수로 하여 회귀분석이 가능하다는 시사점을 갖는다.

셋째, 게임을 판매하기 위해서는 '게임물 등급위원회'와 '청소년 보호위원회'의 각 심의를 받아야하지만 두 기관의 이중 등급분류제도는 중복 심의문제를 낳았고, 「게임산업진흥에 관한 법률」 제21조(등급분류) 제1항 개인이 사업 창업 절차를 받아 사업자로 등록해야 함을 내포하여 개인적인 한계점을 가지고, 제11조의 3항으로 게임 버전 업그레이드나 패치제공으로 게임을 보완 시 재등급 분류를 받아야하여 개인 게임 제작자나 중소 게임기업에게 있어 큰 무리가 있는 상황이다[54]. 따라서 게임 판매에 있어 제한을 받는 국내 시장이 아닌 스팀을 활용한 세계시장으로 진출을 고려해보는 것에 있어, 본 연구는 스팀을 통한 감성분석 결과를 제시하여 스팀에서의 판매량에 영향을 줄 수 있는 eWOM 요인을 제시한다는 실무적 시사점을 갖는다.

6.3 연구의 한계점 및 향후 연구

본 연구는 다음과 같은 한계점을 갖고 있다.

첫째, 2016년에 발매된 게임 중 유료로 판매되며 10,000개 이상이 판매된 90개의 샘플을 선정하였지만, 이는 안정적인 회귀분석을 진행하기 위해 샘플 수가 부족하였다는 것이다. 추가로 각 게임의 경우 2016년에 발매했지만, 2016년 상반기에 발매된 게임도 있지만, 하반기에 발매된 게임도 있어 판매변화량의 상대적인 차이점이 있음에도 불구하고 이를 고려하지 않고 샘플을 90개로 한정하였다. 이후 연구에서 발매 연도를 고려하여 더 많은 샘플을 확보한 뒤 연구가 진행되어야 할 것이다.

둘째, 선정된 게임의 경우 단순히 발매 연도를 기준으로 선정되었기 때문에 장르나, 연령 제한 등을 고려하지 않고 감성분석에 있어 단순히 SentiWordNet을 기준으로 감성지수를 산출하였기 때문에 실제 각 게임마다 상이한 감성 단어에 차별점을 주지 못했다는 것이다. 이후 각 게임에 해당하는 장르, 연령제한을 고려하여 세분화된 감성사전으로 연구가 진행되어야 할 것이다.

셋째, 본 연구의 결론 중 긍정적인 리뷰의 감성이 판매량에 정의 영향을 미치지 않은 것과 부정적인 리뷰의 크기가 판매량에 부의 영향을 미치지 않은 것과 반대로 부정적인 리뷰가 판매량에 부의 영향을 주는 것과 긍정적인 리뷰의 크기가 판매량에 정의 영향을 주는 것은 현업에서 종사하고 있는 실무자의 입장에서 추측이 가능할 것으로 사료된다.

마지막으로, 종속변수인 게임 판매변화량은 스팀 스파이(https://steamspy.com/)에서 제공하는 게임 소지자의 수로 측정할 경우, 스팀 내에서 프로필이 공개 상태의 사용자만을 집계한 수로 일부 비공개 프로필 사용자의 소지 숫자는 파악이 어렵다는 한계점을 갖고 있다. 일부 게임 배급사에서 정확한 게임의 판매변화량을 제공하거나 스팀에서 비공개 프로필 사용자에 대한 정보를 공개할 경우 더 정확한 연구가 가능할 것으로 사료된다.

REFERENCES

[1] H. J. Jang. (2017. 3. 31). *Worldwide attention, Steam 1st battle ground, Game performance is the most important.* THIS IS GAME. http://www.thisisgame.com/webzine/news/nboard/5/?n=70753

[2] S. Y. Park. (2015. 2. 27). *Steam, active account exceeded 125 million*, ZD NET Korea. http://www.zdnet.co.kr/news/news_view.asp?artice_id=20150227091302.

[3] Christiansen. T & S. S. Tax. (2000). Measuring Word of Mouth: The Questions of Who and When. *Journal of Marketing Communications, 6(3),* 185-199.

[4] Goh, J. M., G. G. Gao. & R. Agarwal. (2016). The creation of social value: Can an online health community reduce rural-urban health disparities?. *Management Information Systems Quarterly, 40(1),* 247-263.

[5] S. R. Back. (2005). An exploratory study of motives toward word of mouth activities on the Internet. *The Korean Journal of Advertising and Public Relations, 7(1),* 108-144.

[6] Zhang. W. & S. Watts. (2003). Knowlege Adoption in Online Communities of Practice. *International Conference on Information Systems, Atlanta, AIS, 35(3),* 96-109.

[7] Henning-Thurau. T., K. P. Gwinner., G. Walsh. & D. D. Gremler. (2004). Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motives Consumers to Articulate Themselves on the Internet?. *Journal of Interactive Marketing, 18(1),* 32-52.

[8] Chatterjee. P. (2001). Online Review - Do Consumers Use Them?. *Advances in Consumer Research, 28,* 129-133.

[9] Mcknight. H. D., V. Choudhury., & C. Kacmar. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research, 13(3),* 334-359.

[10] Brown. J. J., A. J. Broderick., & N. Lee. (2007). Word of Mouth Communication with Online Communication : Conceptualizing the Online Social Network. *Journal of Interactive Marketing, 21(3),* 2-20.

[11] Schindler. R. M. & B. Bickart.(2005, January). Published Word of Mouth: Referable, Consumer-Generated Information on the Internet. *In C. P. Haugtvedt, K. A. Machleit and R. F. Yalch(eds.)* .(pp. 35-61). Online Consumer Psychology, NJ: Lawrence Erlbaum Associates.

[12] I. K. Kim. (2016). *The dynamics of online word-of-mouth and marketing performance : exploring mobile game application reviews using text-mining and machine-learning.* ph.D. dissertation. Korea University, Seoul.

[13] B. Y. Choi. (2017). *Understanding and application of consumer behavior.* Seoul : Parkyongsa.

[14] H. S. Byeon & M. S. Yim. (2014). The Impact of Users' Congruity and Emotion on Intention to Game Use. *Journal of Digital Convergence, 12(11),* 89-98.

[15] Y. J. Jo. (2015). A Study on the Influence of Connectivity and Convenience of Smartphones of Word-of mouth Intentions in the Convergence Era : Focused on the Mediating Effects of Application. *Journal of Digital Convergence, 13(5),* 69-78.

[16] D. S. Yorm. (2016). Factors Affecting User Satisfaction of Mobile Social Network Games : Focusing on the Quality and Self-determination. *Journal of Digital Convergence, 14(11),* 459-467.

[17] D. S. Youm. (2017). The Effect of Perceived Enjoyment and User Characteristics on Intention of Continuous Use of Mobile Social Network Games : Focusing on Mediating Effect of Flow Experience. *Journal of Digital Convergence, 15(9),* 415-425.

[18] J. W. Kang. (2008). *Game and Culture Research.* Seoul : Communication Books.

[19] S. T. Park., H. C. Lee., T. U. Kim & S. M. Choi. (2012). A Study on Factors Influencing Attachment of Gamers to MMORPG On-line Games. *Journal of Digital Convergence, 10(2),* 109-119.

[20] Dang, Shilpa & Peerzada Hamid Ahmad. (2014). Text Mining: Techniques and its Application. *International Journal of Engineering & Technology Innovations, ISSN (Online) : 2348-0866, 1(4),* 22-25.

[21] C. N. Jun. & I. O. Seo. (2013). Analyzing the Bigdata for Practical Using into Technology Marketing : Focusing on the Potential Buyer Extraction. *Korean Strategic Marketing Association, 21(2),* 181-203.

[22] Dang, Dr. Shilpa & Peerzada Hamid Ahmad. (2015). A Review of Text Mining Techniques Associated with Various Application Areas. *International Journal of Science and Research (IJSR), 4(2),* 2461-2466.

[23] S. H. Seo & J. T. Kim. (2016). Deep Learning Based Emotion Analysis Research Trend. *Korea Multimedia Society, 20(3),* 8-22.

[24] Chen, H. & Zimbra. D. (2010). AI and opinion mining. Intelligent Systems. *IEEE, 25(3),* 74-80.

[25] Nasukawa. T. & Yi. J. (2003). Sentimentanalysis: Capturing favorability using natural language processing. *In Proceedings of the 2nd international conference on Knowledge capture.* (pp. 70-77). ACM.

[26] O'Connor. B., Balasubramanyan. R., Routledge B. R., & Smith. N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *ICWSM, 11,* 122-129.

[27] Hu. M. & Liu. B. (2004). Mining and summarizing customer reviews. *KDD'04 Proceedings of the tenth international conference on knowledge discovery and data mining.* (pp. 168-177). ACM SIGKDD.

[28] Esulim. A. & Sebastiani. F. (2006). SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. *Proceedings of LREC.* (pp. 417-422). ITALY

[29] Baccianella. S., Esuli. A. & Sebastiani. F. (2010). SentiWordNet 3.0: An Enganced Lexical Resource for Sentiment Analysis and Opinion Mining. *LREC, 10,* 2200-2204.

[30] Liu. S. M & Chen. J. H. (2015). A multi-label classification based approach for sentiment classification. *Expert Systems with Applications, 42(3),* 1083-1093.

[31] Arcjak. N., A. Ghose & P. G. Ipeirotis. (2007). Show me the money!. *Proceedings on the 13th International Conference.* (pp. 56-65). ACM SIGKDD.

[32] Archak. N., A. Ghose, & P. G. Ipeirotis. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science, 57(8),* 1485-1509.

[33] Berger. J., A. T. Sorensen. & S. J. Rasmussen. (2010). Positive Effects of Negative Publicity: When Negative Reviews Increase Sales. *Marketing Science, 29(5),* 815-827.

[34] W. J. Chu. & M. J. Roh. (2014). Exploring the Role of Preference Heterogeneity and Causal Attribution in Online Ratings Dynamics. *Asia Marketing Journal, 15(4),* 61-101.

[35] J. H. Lee, S. Hong & D. Kang. (2014). The Marketing Success Factors of Hyundai Card Company: Business Model, Development of Goods and BTL Marketing. *Korea Business Review, 18(3),* 147-170.

[36] Chintagunta. P. K., S. Gopinath & S. Venkataraman. (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.

[37] Dellarocas. C., G. Gao. & R. Narayan. Are Consumers More Likely to Contribute Online Reviews for Hit or Niche Products?. *Journal of Management Information Systems. 27(2),* 127-157.

[38] H. K. Lee & H. Kwak. (2013). Investigation of Factors Affecting the Effects of Online Consumer Reviews. *Informatization policy, 20(3),* 3-17.

[39] H. W. Hwangbo & J. H. Kim. (2016). A Study on the Factors Affecting to the Export Performance for Korean Drama Using Sentimental Analysis. *The e-Business Studies, 17(6),* 87-99.

[40] Y. Liu. (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing, 70(3),* 74-89.

[41] Y. K. Kim. (2013). A Study on Relationship between the Relationship Benefit, Customer Satisfaction and Loyalty of Internet Shopping Malls. *Daehan Academy of Management Information Systems, 32(4),* 155-187.

[42] Sawhney. M. S. & J. Eliashberg. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science, 15(2),* 113-131.

[43] Duan. W., B. Gu. & B. Whinston. (2008). Do online reviews matter?: An empirical investigation of panel data. *Decision Support Systems, 45(4),* 1007-1016.

[44] Pang. B., Lee. L. & Vaithyanathan. S. (2002). Thumbs up?: sentiment classification using machine learning techniques. *In Proceedings of the ACL-02 conference on Empirical methods in natural language processing.* (pp. 79-86). ACM.

[45] Manning. C. D., Raghavan. P. & Schutze. H. (2008). Introduction to Information Retrival. *Cambrige: Cambridge university press, 1(1).*

[46] Tay. F. E. & Cao. L. (2001). Application of support vector machines in financial time series forecasting. *Omega, 29(4),* 309-317.

[47] Tong. S. & Koller. D. (2002). Support Vector Machine Active Learning with Applications to Text Classification. *The Journal of Machine Learning Research, 2,* 45-66.

[48] Pak. A. & P. Paroubek. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *Proceedings of the Seventh International Conference on Language Resources and Evaluation*(pp. 1320-1326). Valletta.

[49] Kumar. A. & Sebastian. T. M. (2012). Sentiment Analysis on Twitter Issue. *IJCSI, 9(3), 372-378*

[50] H. J. Kim, K. H. Han & S. S. Shin. (2017). Crepe Search System Design using Web Crawling, *Journal of Digital Convergence, 15(11),* 261-269.

[51] Y. Y. Kim & M. Song. (2016). A Study on Analyzing Sentiments on Movie Reviews by Multi-Level Sentiment Classifier, *Journal of intelligence and information systems, 22(3),* 71-89.

[52] P. G. Preethi., V. Uma & Ajit kumar. (2015). Temporal Sentiment Analysis and Causal Rules Extraction from Tweets for Event Prediction. *Procedia Computer Science, 48,* 84-89.

[53] Coovert. M. D. & G. D. Reeder. (1990). Negativity Effects in Impression Formation: The Role of Unit Formation and Schematic Expectations. *Journal of Experimental Social Psychology, 26(1),* 49-62.

[54] J. S. Kim, T. Y. Lee, T. G. Kim & H. W. Jung. (2015). Studies on the development scheme and the current state of Korea Game Industry. *Journal of Digital Convergence, 13(1),* 439-447.

정 근 웅(Jung, Keun Woong) [학생회원]

- 2016년 2월 : 협성대학교 경영정보학과(학사)
- 2018년 2월 : 성균관대학교 경영정보전공(경영학석사)
- 관심분야 : 빅데이터, 감성분석, 경영정보
- E-Mail : wowhbo@skku.edu

김 종 욱(Kim, Jong Uk) [정회원]

- 1982년 2월 : 서울대학교 산업공학과 (공학사)
- 1990년 12월 : (미)조지아주립대학교 경영정보학전공(경영정보학 석사 수료)
- 1994년 8월 : (미)조지아주립대학교 경영정보학전공(경영정보학박사)
- 2002년 3월 ~ 현재 : 성균관대학교 경영전문대학원 교수
- 관심분야 : 빅데이터, 감성분석, 데이터 인텔리전스
- E-Mail : jukim@skku.ac.kr