

# 온라인 리뷰 유용성과 상품매출에 영향을 주는 요인 : 중국 온라인 쇼핑 플랫폼 데이터를 기반으로

황 침\* · 권영진\*\* · 이상용\*\*\*

## Research on the Influencing Factors of the Usefulness of the Online Review and Products Sales : Based on Chinese Online Shopping Platform Data

Chim Hwang\* · Young-Jin Kwon\*\* · Sang-Yong Tom Lee\*\*\*

### Abstract

This empirical study explored characteristics that affect the usefulness of online reviews, in the China e-commerce platform, and implemented multiple regressions to find factors that significantly influence on product sales, ultimately. Till now, prior studies have continuously revealed what factor affects usefulness of online review or product sales, only in respective terms. The point of our study is that we built two-level regression models, thereby being able to comprehensively analyze these two different targets. Before plunging into running regressions, we carefully collected 192,764 online review data for 200 products extracted from the Jingdong, the second biggest e-commerce platform in China. Also, we gathered "review sentimental scores" variable from each review and used that one as a core variable in our regression model, thus we were able to implement both quantitative and qualitative research.

The evidences from the two-level regression models showed that the extent to which a product is experience good positively affects both usefulness of a review and product sales, again the usefulness of a review contributes to product sales in sequence. Also, the property of experience good has interaction effect on both for two-level regression models. Our main findings highlight the importance of role of online review to business performance of e-commerce firms.

Keywords : China e-Commerce Platform, User Review, Experience Goods and Search Goods, Crawling, Sentimental Analysis

Received : 2018. 05. 17.    Revised : 2018. 06. 15.    Final Acceptance : 2018. 06. 18.

※ This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A3A2066740).

\* School of Business, Hanyang University, e-mail : 415115178samhuang@gmail.com

\*\* School of Business, Hanyang University, e-mail : bohemian4682@gmail.com

\*\*\* Corresponding Author, School of Business, Hanyang University, 222 Wangshimri-ro, Seongdong-gu, Seoul, 04763, Korea, Tel : +82-2-2220-1814, e-mail : tomlee@hanyang.ac.kr

## 1. 서 론

### 1.1 연구 배경

최근 인터넷 쇼핑물을 이용한 전자상거래가 소비자의 소비생활에서 차지하는 비중이 높아짐에 따라, 전자상거래에 대한 관심이 커지고 있는 추세이다. 2016년 CNNIC<sup>1)</sup>에서 발표한 <2016년 중국 인터넷 시장 소비 연구보고서>에 따르면, 2016년 6월까지 중국의 인터넷 소비자는 2억 2,383만 명에 달하며 연간 증가폭은 20.2%이다. 이 연구 보고에 따르면 인터넷 쇼핑 경험에 대한 소비자 만족도는 대체로 긍정적이다. 소비자의 54%가 인터넷 쇼핑에 매우 만족한다고 답하였고, 불만족스럽거나 아주 불만족스러워 하는 소비자는 2%밖에 되지 않았다. 또한, 다른 고객들의 온라인 리뷰는 소비자들의 구매결정에 가장 큰 영향을 미치는 요소로, 지인, 친구들의 의견보다도 더 중요한 외부 정보로 평가되었다. 이에 반해 구매결정 시, 전문가 의견이나 유명 리뷰어들의 평가를 가장 중요하게 생각하는 소비자는 전체 조사대상 중 19.3%에 그쳤다. 이렇듯 온라인 리뷰는 소비자의 구매결정에 상당히 큰 영향을 미치고 있으며, 여러 정보 중에서도 비교적 영향력이 큰 정보로서 자리매김 하고 있다.

한편, 온라인 리뷰는 소비자에게 유용한 정보를 제공함과 동시에 기업이 소비자와 소통하고 질 높은 상품을 만들 수 있도록 소통의 장을 제공하였다. 점점 더 많은 기업들이 온라인 리뷰가 기업의 수익에 주는 영향을 인식하고 소비자가 상품 리뷰를 할 수 있는 공간과 플랫폼을 제공하고 있다. 그 예로 중국의 몇 대형 B2C사이트(아마존, JINGDONG, DANGDANG 등)에서는 다양한 형태로 소비자들이 자신의 구매 경험을 공유할 수

있는 서비스를 제공하고 있다. 또한 온라인 리뷰가 소비자의 소비의사결정에 미치는 영향이 증대됨에 따라 기업의 매출에 미치는 영향 또한 커지고 있다. Dellarocas[2003]는 온라인 고객 리뷰가 온라인 구전효과를 창출하는 가장 강력한 채널이라고 주장했다. 즉, 상품에 대한 부정적인 리뷰는 시장에서 상품이 설 자리를 없애고, 긍정적인 리뷰는 상품의 매출 증대에 기여한다. 온라인 리뷰 플랫폼을 처음 시도한 기업이라고 알려진 아마존은 여전히 리뷰 플랫폼을 수익 창출에 매우 핵심적인 역할을 하는 요인이라고 간주하고 있다.

한편 인터넷과 정보기술의 발전은 소비자들이 온라인상에서 상품경험을 공유할 수 있는 플랫폼의 제공을 가능하게 하고 있다. 아마존이 1995년 창립 초기에 이용자들에게 상품 리뷰기능을 제공한 이후 많은 업체와 플랫폼들이 모두 해당 기능을 사용하기 시작했는데, 이에는 Epinions.com와 같은 제3자 리뷰 사이트도 포함된다. 특히 최근 몇 년 사이에 WEB2.0 기술이 급격히 발전함에 따라 전통적인 정보전달 방식이 크게 변화하고 있다. 과거에는 소비자들이 구매결정을 내리기 전에 오프라인에서 지인을 통해 정보를 얻었다면, 현재는 온라인을 통해 이용자생성 콘텐츠(User-Generated Content, UGC)나 상품 리뷰와 같이 다른 온라인 이용자가 제공하는 정보를 더 많이 참고하는 경향이 있다. 이렇듯 온라인 쇼핑에서 이용자들의 관점 및 태도를 반영한 상품후기의 중요성은 더욱 높아지고 있다. 연구조사기업 Channel Advisor에서 2010년과 2011년에 발표한 소비자행위 보고서에 따르면 상품 리뷰가 온라인 소비자의 구매결정에 미치는 영향은 점점 더 커지고 있는데, 2011년 보고서에서는 설문조사에 참여한 이용자의 74%가 상품 리뷰를 읽어본다고 응답했다[Channel, 2010; 2011]. 그 중 43%의 이용자가 상품 리뷰 정보가 구매결정에 영향을 미쳤다고 응답했다. 그리고 댓글은 감성

1) 출처 : <2016년 중국 인터넷 시장 소비 연구보고서>, CNNIC, 2016.

에 따라 긍정적 댓글과 부정적 댓글로 분류될 수 있고, 이 여부에 따라서 다른 이용자들은 구매 의사결정에 영향을 받는 것으로 보고되고 있다 [Duan et al., 2008; Elliott, 2002; 박은아 등, 2007; 손지연과 어숙희, 2008; 손진아와 이은영, 2007].

## 1.2 연구 목적

이에 본 연구는 중국의 비교적 성공적인 온라인 쇼핑몰(Jingdong, Jingdong)의 상품 리뷰 수집을 통해 탐색재와 경험재에 따라 온라인 리뷰 유용성에 영향을 미치는 요인, 나아가 온라인 리뷰가 상품매출에 미치는 영향에 대해 연구한다. 주요 목적은 다음과 같다.

첫째, 리뷰 유용성과 상품매출에 영향을 주는 요인을 탐구하고, 그 중에서도 특히 리뷰의 감성을 수치화하여 변수로서 살펴보고자 하였다. 기존 연구는 대부분 리뷰점수라는 확실적인 수치로 리뷰 텍스트를 대체하였다. 이렇게 분석할 경우 발생하는 주된 문제점은 우선 온라인 리뷰 텍스트에 존재하는 세밀한 정보를 온라인 리뷰 점수가 대칭적으로 반영하기 어렵다는 점이다. 즉 온라인 리뷰 점수에 포함되는 정보에는 한계가 있으며, 또한 리뷰자가 특정된 리뷰점수를 부여한 구체적인 이유를 알기 어렵다. 이런 난점을 해결하기 위해, 텍스트 분석기술을 활용하여 온라인 리뷰 텍스트로부터 자동적으로 감성 단어와 감성 대상을 추출하고 감성 점수에 대해 분석할 수 있다.

둘째, 상품이 갖는 탐색재 및 경험재 속성의 정도가, 리뷰 감성이 리뷰 유용성과 상품매출에 미치는 영향에 조절 역할을 하는지 확인하고자 하였다. 탐색재 속성이 강한 상품의 경우 상품 관련 정보의 양에 대한 소비자의 수요가 낮아지는 경향이 있다. 이 경우 리뷰 감성 점수가 리뷰 유용성 및 상품매출에 주는 영향도 함께 낮아지는지를 살펴본다. 또한, 상품의 경험재 속성이

강할 경우 소비자는 긍정적인 리뷰 텍스트에 민감하게 반응하는 반면, 상품의 위험요소에 대한 리뷰 텍스트에는 둔감하게 반응하는 경향이 있다. 따라서 소비자들이 정보 파악이 잘 안 된 상품에 대해서 구매 포기를 하는 확률은 낮다.

셋째, 상품의 가격이 차등적일 때 리뷰 감성 점수가 상품매출에 어떠한 영향을 미치는지 연구하고자 하였다. 상품의 가격이 비교적 낮을 경우, 소비자들의 상품 정보의 수와 양에 대한 요구는 모두 낮아진다. 이때, 리뷰의 감성이 상품매출에 주는 영향 또한 낮아지는지에 대해 연구한다.

## 2. 이론적 배경

Berger and Calabrese[1974]의 불확실성 감소 이론(Uncertainty Reduction Theory)에 따르면, 소비자는 상품 구매 시 해당 상품에 대한 사전 지식이 충분치 않거나, 혹은 소비에 따른 결과를 예측하기 힘들 때 구매의 불확실성을 감소시키기 위한 노력을 한다. 특히 이 노력은 구매의 효용을 극대화하기 위해 상품에 대한 관련 정보를 더 수집하는 행위로 이어진다. Nelson[1974]의 연구에 의하면, 상품이 탐색재인 경우 소비자는 해당 상품의 가치에 대해 사전 정보를 다양하게 얻을 수 있으므로 정보 수집 행위에 열성적이지 않은 경향이 있는 반면, 경험재인 경우 제품의 불확실성은 상대적으로 높기 때문에 그 불확실성을 낮추기 위해 온라인 리뷰와 같은 구매 경험이 있는 소비자의 후기를 매우 꼼꼼히 참고하는 경향이 있다. 특정 상품에 대한 온라인 리뷰는 인지된 가치(perceived value)에 대한 정보를 제공한다[Hu et al., 2008]. 이러한 리뷰들은 소비자에게 간접 경험을 제공하고, 이러한 경험은 소비자로 하여금 상품에 대한 불확실성을 낮추게 한다. 즉, 온라인 리뷰는 소비자의 구매 의사 결정에 중대한 영향을 미친다.

〈Table 1〉 Summary of Prior Studies Related

Researcher(s)	Main findings
Mudambi and Schuff[2010]	- Review stars have positive influence on usability only for experiential products. - The number of letters in the review affects the usefulness of the review. - The influence of negative on the usability of the review is whether or not the experience is good. - The number of votes in the review affects the usefulness of the review only for search good.
Chevalier and Mayzlin[2006]	- The number of reviews has a negative effect on sales. - Review stars have a negative impact on sales(-).
Chen et al.[2004]	- The number of online reviews has a positive impact on car sales.
Chatterjee [2001]	- When a consumer feels intimacy with a merchant(company), they tend to look for less WOM information about the product - Consumers tend to be more influenced by negative reviews when they search for product information of non-friendly merchandise(company) due to price factors.
Forman et al. [2009]	- Compared to online reviews with neutral content, online reviews with extreme content are more useful reviews.
Cao et al. [2011]	In a study of book sales platforms and software download platforms, it was found that extreme reviews were more useful than neutral reviews.
Park and Yu [2011]	- As a result of research on hotel products, negative reviews have a significant effect on consumers' attitudes, and positive reviews influence consumers' purchase intentions.
Kim [2006]	- Negative reviews rather than positive reviews affect consumers' attitudes toward products and purchase intentions more significantly.

Duan et al.[2008]에 의하면 구전(word-of-mouth)은 가장 효과적인 정보전달 방법 중 하나라고 알려져 있고, 특히 그 효과는 경험재에 대해 가장 강하게 작용한다고 한다[Godes and Mayzlin, 2004; Granovetter, 1973]. 또한 Reichheld[2003]는 고객이 다른 고객에게 특정 상품을 추천하려는 경향은 오늘날 성공적인 경영에 필수 불가결한 요소가 되었다고 주장한다. 그리고 그는 이 경향성이 경영 성과 예측의 측면에서 전통적인 기업 성과 측정 지표들보다 더 예측력이 뛰어나다고 주장한다.

이러한 선행 연구들의 결과를 종합하면, 기업의 입장에서 온라인 리뷰에 대한 철저한 분석은 경영 성과에 기여할 뿐만 아니라 성과에 대한 예측을 가능하게 하고 더욱 지속 가능한 경영을 할 수 있게 한다고 볼 수 있다. 하지만, 온라인 리뷰가 경영 성과에 미치는 영향을 연구한 선행 연구들의 수가 많지 않다. 또한, 소비자의 입장에서 유용하다고 판단되는 유용한 리뷰를 결정 짓는 요소들에 대해 분석한 연구는 더욱 그 수가 적고, 궁극적으로 유용한 리뷰가 경영 성과에 미치는 영향을 함께 연구한 연구는 거의 전무하다.

따라서 본 연구는 2단계의 연구 모형을 통해 이 모두를 포괄하는 분석을 했다는 점에서 기존의 선행 연구와 구분되는 차이를 갖는다.

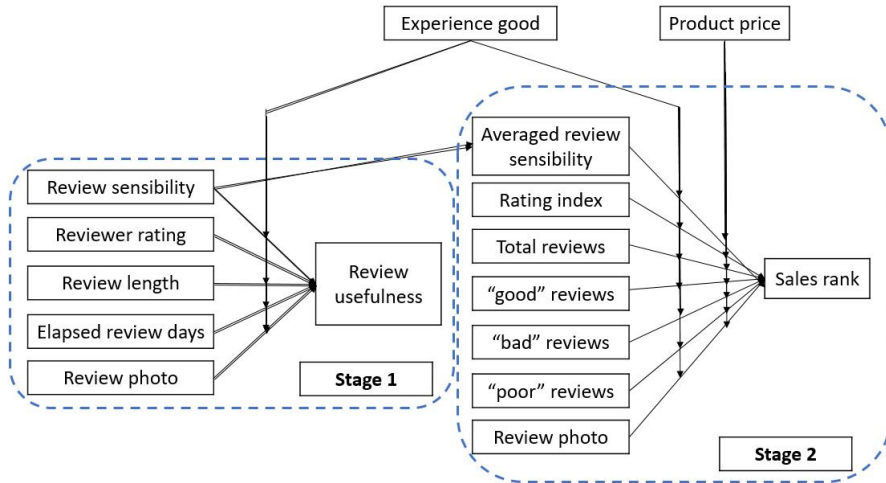
〈Table 1〉은 온라인 리뷰에 대해 연구한 관련 선행 연구들의 요약 테이블이다.

〈Table 1〉의 선행 연구 결과를 종합해보면, 온라인 리뷰가 상품매출에 유의한 영향을 미친다는 결론을 맺은 다수의 연구들이 존재하지만 그 방향성은 상이함을 알 수 있다. 따라서, 본 연구의 분석 결과는 관련 선행 연구들의 흐름에 영향을 줄 것이라고 생각한다.

### 3. 연구모형과 가설

#### 3.1 연구모형

본 연구는 온라인 리뷰 효과를 연구한 선행연구에서 활용된 모형을 기초로 두고, 발전된 연구모형을 구축하였다. 연구모형은 크게 두 가지 부분으로 구성되고, 해당 내용은 〈Figure 1〉을 통해 나타내었다.



〈Figure 1〉 Research Model

첫 번째 단계-본 연구는 온라인 리뷰의 유용성은 온라인 리뷰 감성 점수, 리뷰자의 등급, 리뷰 텍스트의 길이, 리뷰 후 경과일, 사진리뷰 등의 개별 리뷰 특성에 영향을 받았다고 가정하고 모형을 구축했다. 또한 상품의 속성을 탐색재와 경험재로 분류함으로써, 이 분류가 리뷰 유용성에 구분되는 영향을 미치는지 분석한다.

두 번째 단계-리뷰 감성 점수, 평가지수, 총 리뷰 수 등 해당 상품에 대한 리뷰 특징이 상품매출에 주는 영향을 분석함과 동시에, 경험재의 속성과 상품 가격이 조절 작용하는지 분석한다.

세 번째 단계-위 두 단계를 통해 궁극적으로 리뷰 유용성이 상품매출에 미치는 영향에 대해 연구한다. 본 연구는 특히 리뷰 감성 점수를 개별 리뷰를 묘사하는 주요 변수로 선택하였다.

본 연구는 이와 같이 리뷰 유용성이 상품매출에 미치는 영향을 분석하는 모형을 구축했고, 이때 경험재의 속성이 조절 효과를 갖는지에 대한 분석에 초점을 맞추고 연구를 진행했다. 탐색재 및 경험재의 속성은 리뷰에 영향을 받는 이용자가 감지하는 상품의 품질, 구매 실패 확률 및 데이터 처리 원가에 영향을 준다. 즉 경험재의 속성은 리뷰 감성 점수, 리뷰 유용성 및 상품매출

에 주는 영향을 증가 또는 감소시킬 수 있다. 모형 2단계의 “리뷰 감성 점수 평균치”는 모형 1단계의 “리뷰 감성 점수”를 “리뷰 유용성(투표수)”으로 곱한 값을 나타낸다. 모형 1단계에서 사용된 “리뷰 감성 점수”를 사용하지 않고 해당 값에 리뷰 투표수를 곱한 값(“리뷰 감성 점수 평균치”)를 사용한 이유는, 리뷰 감성 점수의 평균치가 2단계 모형에서 궁극적으로 알아보고자 한 “제품 판매 순위”에 더 가시적인 영향을 미칠 것이라고 가정한 뒤 모형을 설정했기 때문이다. 즉, 1단계 모형과 2단계 모형의 연계성을 높일 수 있다고 판단했다.

### 3.2 연구가설

본 연구는 리뷰 유용성에 영향을 미치는 요인들을 분석한 다수의 선행 연구들을 비교 및 취합했고, 궁극적으로 이 요인들이 상품 판매 순위에 미치는 영향을 분석했다. 우리는 개별 리뷰의 특성, 리뷰 감성 점수, 리뷰자의 등급, 리뷰 길이, 리뷰한 후 경과일, 그리고 리뷰 사진 여부가 리뷰 유용성에 영향을 미칠 것이라고 가정했고, 해당 내용은 <Table 2>의 연구 가설에 반영되어 있다. 가설 1

부터 가설 5까지는 여러 요인들이 리뷰 유용성에 어떠한 영향을 미치는지에 알아보기 위한 가설들에 해당하고, 이때 경험재 속성이 조절 작용하는지에 대한 하위 가설 또한 포함한다. 다음으로, 가설 6부터 가설 12까지는 본 연구에서 궁극적으로 알아보자 하는 상품 판매 순위에 영향을

미치는 요인들에 대한 가설이고, 이때 경험재의 속성과 상품의 가격이 조절 효과를 갖는지에 대한 하위 가설을 포함한다. 가설 검증에 대한 결과는 결과 섹션의 결과표에 나타내어져 있다. 본 연구는 개념 간에 명확한 인과성을 규명하기보다는, 탐색적인 물음을 가정한다.

〈Table 2〉 Research Hypotheses

Numb.		Hypothesis	
Review usefulness	H1	1	The review sensitivity score will positively affect the overall usefulness of review.
		1-1	The attribute of experience good will positively affect the relation between review sensitivity score and review usefulness.
Review usefulness	H2	2	The Reviewer ratings will positively affect the overall usefulness of review.
		2-1	The attribute of experience good will positively affect the relation between Reviewer ratings and review usefulness.
Review usefulness	H3	3	The review length will positively affect the overall usefulness of review.
		3-1	The attribute of experience good will positively affect the relation between review length and review usefulness.
Review usefulness	H4	4	The Elapsed review days will negatively affect the overall usefulness of review.
		4-1	The attribute of experience good will positively affect the relation between Elapsed review days and review usefulness.
Review usefulness	H5	5	The Reviews photos will positively affect the overall usefulness of review.
		5-1	The attribute of experience good will positively affect the relation between Reviews photos and review usefulness.
Product sales ranking	H6	6	The average reviews sensibility will positively affect the product sales rank.
		6-1	The attribute of experience good will positively affect the relation between average reviews sensibility and product sales rank.
		6-2	The product price will positively affect the relation between average reviews sensibility and product sales rank.
Product sales ranking	H7	7	The rating index will positively affect the product sales rank.
		7-1	The attribute of experience good will positively affect the relation between rating index and product sales rank.
		7-2	The product price will positively affect the relation between rating index and product sales rank.
Product sales ranking	H8	8	The total reviews will positively affect the product sales rank.
		8-1	The attribute of experience good will positively affect the relation between total reviews and product sales rank.
		8-2	The product price will positively affect the relation between total reviews and product sales rank.
Product sales ranking	H9	9	The "Good" reviews will positively affect the product sales rank.
		9-1	The attribute of experience good will positively affect the relation between "Good" reviews and product sales rank.
		9-2	The product price will positively affect the relation between "Good" reviews and product sales rank.
Product sales ranking	H10	10	The "Poor" reviews will negatively affect the product sales rank.
		10-1	The attribute of experience good will positively affect the relation between "Poor" reviews and product sales rank.
		10-2	The product price will positively affect the relation between "Poor" reviews and product sales rank.
Product sales ranking	H11	11	The "Neutral" reviews will positively affect the product sales rank.
		11-1	The attribute of experience good will positively affect the relation between "Neutral" reviews and product sales rank.
		11-2	The product price will positively affect the relation between "Neutral" reviews and product sales rank.
Product sales ranking	H12	12	The reviews photos will positively affect the product sales rank.
		12-1	The attribute of experience good will positively affect the relation between reviews photos and product sales rank.
		12-2	The product price will positively affect the relation between reviews photos and product sales rank.

### 4. 연구 방법

#### 4.1 연구 대상

본 연구는 Jingdong(www.jd.com) 전자상거래 사이트를 연구 대상으로 선택하였다. 해당 사이트는 중국 전자상거래 사이트 중 알리바바(Alibaba) 다음으로 규모가 큰 사이트이다. jd.com을 연구 대상으로 선택한 이유는 해당 사이트상의 상품들은 품목 종류가 다양할 뿐만 아니라 온라인 리뷰도 상당히 많기 때문이다. 탐색재와 경험재의 분류도 명확하여, 본 연구에서 알아보고자 한 독립변수와 종속 변수들에 대한 데이터를 포괄적으로 수집할 수 있었다.

jd.com 사이트는 비교적 복잡한 리뷰 시스템을 가지고 있다. 소비자는 자신의 실제 사용 현황에 따라 리뷰를 작성할 수 있는데, 한 가지는 이미

구매한 상품 또는 서비스에 대한 리뷰이고 다른 가지는 기타 소비자의 리뷰에 투표를 진행하는 방식이다. 이러한 리뷰는 판매량에 영향을 주는데, 예를 들어 판매자들은 소비자들이 제기한 의견을 상품에 반영함으로써 상품과 유통의 품질을 높일 수 있다. 반대로, 리뷰는 소비자들에게도 매우 도움이 되는데, 예를 들어 잠재적인 소비자는 온라인 리뷰 내용을 근거로 상품의 품질, 기능 등에 대해 자세하게 알 수 있다. 또 한 가지 매우 중요한 점은 jd.com 사이트는 리뷰의 유용성에 따라 상품 투표 내용을 열거한다. 위의 <Figure 2>는 온라인 리뷰 작성할 때의 화면인데, 여기에서 알 수 있듯이 상품에 대한 1에서 5까지의 점수를 부여할 수 있다. 이러한 이용자들의 평가가 반영된 지수가 <Figure 3>에 있는 “평가 지수”이다. 이 점수에 따라 해당 상품은 “좋음”(5점, 4점), “보통”(3점, 2점), 또는 “나쁨”(1점) 중 하나로 분류된다.



<Figure 2> Comment Captured on Jindong Website



<Figure 3> Review System on Jindong Website

### 4.2 데이터

본 연구는 쭈쭈미 크롤링 프로그램을 통하여 온라인 쇼핑몰 사이트의 리뷰 데이터를 수집했다. 수집 과정은 기본 정보 입력, 작업 흐름 설계, 실행 계획 입력 및 수집의 일련의 4가지 단계를 거친다. 본 연구는 jd.com에서 판매중인 텔레비전, 디지털 카메라, 외장하드, 메모리 카드 각각에 대해 판매량이 상위 50위에 속하는 상품을 대상으로 데이터를 수집했고 총 192764개의 리뷰를 수집할 수 있었다. 특정 상품에 대해 수집할 정보는 다음과 같다 : 상품명, 판매순위, 상품 가격, 총 평가 지수, 총 리뷰 수, 좋은 리뷰 수, 보통 리뷰 수, 나쁜 리뷰 수, 사진 있는 리뷰 수, 리뷰자 등급, 리뷰 내용, 리뷰한 시각. <Figure 4>는 쭈쭈미 프로그램의 화면을 나타낸다.

### 4.3 분석 방법

본 연구는 머신 러닝을 기반으로 분석을 진행했고 다음의 3단계를 거쳤다.

- 1) 현재 가장 높은 빈도로 사용되는 중국어 분류 모듈인 jieba 모듈을 도입하여, 스페이스가 없는 중국어 문장을 독립적인 단어들로 나뉘었으며 단어 빈도를 분석하였다.
- 2) 중국어 텍스트 내용을 처리하는 전문적 파이썬 라이브러린 SnowNLP 라이브러리를 도입하여, 문장 감성을 분석하였다. 이 파이썬 라이브러리는 기존에 설계된 머신 러닝 알고리즘을 통해 리뷰 감성 분석 시 자주 사용된다.
- 3) 2단계를 통해 감성 분석된 특정 리뷰에 대해 점수를 부여했다. 예를 들어서, 특정 리뷰가 0부터 1까지의 점수를 받았다면, 해당 리뷰는



<Figure 4> Crawling Program

리뷰 내용	리뷰 감성 극단성	리뷰 길이	리뷰 작성 시간
用工具看了下，硬盘全新！	0.982211829965778	42	2017-05-24 09:23
第一次在京东买东西，下面一波三折，终于收到硬盘。	-0.999994496785612	420	2017-06-12 22:13
物流很快，第二天就到了，	-0.99999998964929	247	2017-06-03 21:28
西数品牌不错的，本来想买	0.938443702355066	108	2017-06-10 21:29
物流速度非常的快。实物包	0.998156697186332	94	2017-06-03 12:44
送货速度快，包装也很好，	0.996577639996822	23	2017-05-20 16:21
很久之前就想买一个移动硬	0.993167766648774	68	2017-06-29 16:14
1t的，比我以前买的小多了	0.256870773470378	51	2017-06-06 11:03
不错，第一次买，价格合适	0.874673284526943	37	2017-06-03 22:13
一直我用西数 固态 机械卷	0.981591499461899	25	2017-06-19 11:33
非常完美昂打6900k 小超3.8	-0.996780657948785	144	2017-06-14 20:59
客观说一下：1.今天刚到手	0.960158726047492	151	2017-06-16 22:35

<Figure 5> Example of Measuring Review Sensibility



긍정적인 감성을 나타낸다고 판단된다. 그 다음, 이 점수에서 0.5를 차감한 뒤 2를 곱하기 하여 리뷰의 최종 감성점수 점수가 부여된다. 즉, 감성 점수가 1에 가까울수록 해당 리뷰는 긍정적이고, 감성 점수가 -1에 가까울수록 해당 리뷰가 부정적임을 나타낸다. 0인 경우는 임계점이라 방향성이 없다고 간주된다.

## 5. 연구 결과

### 5.1 기초 통계

본 연구는 총 200개 상품, 192,764개의 리뷰 데이터를 바탕으로 연구를 진행하였다. 데이터는 리뷰 유용성(투표수), 리뷰 감성, 리뷰 길이, 리뷰자 등급, 리뷰한 경과일, 사진 여부, 판매 순위, 리뷰 감성 점수 평균치, 평가 지수, 총 리뷰 수, 사진 있는 리뷰 수, “좋음” 리뷰 수, “보통임” 리뷰 수, “나쁨” 리뷰 수, 가격 정보를 포함한다. 본 연구는 연구 모델 중의 두 가지 단계에 대하여

모두 기초 통계량 분석을 진행 하였으며, <Table 3>과 <Table 4>를 통해 해당 데이터 수치의 범위, 최대값, 평균, 표준편차를 나타냈다.

<Table 4>에서 “총 리뷰 수”는 “좋음 리뷰 수”와 “보통 리뷰 수”, 그리고 “나쁨 리뷰 수”에 더하여 평점이 누락된 일반 리뷰 수도 포함한다. 즉, “총 리뷰 수 = 좋음 리뷰 수+보통 리뷰 수+나쁨 리뷰 수”일 때 선형중속의 이유로 발생 가능한 추정 불가능성(회귀분석 시)도 배제할 수 있었다.

### 5.2 상관 분석

본 연구는 먼저 리뷰 감성 점수, 리뷰자 등급, 리뷰 텍스트 길이, 리뷰한 후 경과일, 리뷰 사진 여부, 리뷰 유용성 간의 상관 분석을 진행하였고 <Table 5>를 통해 확인할 수 있다. 다음으로 리뷰 감성 점수의 평균치, 평가지수, 총 리뷰 수, “좋음” 리뷰 수, “보통” 리뷰 수, “나쁨” 리뷰 수, 사진을 포함한 리뷰 수, 판매량 순위에 대하여 상관을 진행하였고 <Table 6>을 통해 확인할 수 있다.

<Table 3> Summary Statistics for Model 1

	N	range	Min.	Max	Average	S.E.	Variance
Review usefulness	192764	1164	0	1164	3.62	6.742	45.448
Reviews sensibility	192764	2.000	-1.000	1.000	.339	.733	.537
Review length	192764	566.0	.0	566.0	33.612	38.883	1511.875
Reviewer ratings	192764	4.0	1.0	5.0	3.995	.935	.875
Elapsed review days	192764	1929	3	1932	70.48	103.602	10733.420
Reviews photos	192764	1	0	1	0.392	0.654	0.461

<Table 4> Summary Statistics for Model 2

	N	range	Min.	Max	Average	S.E.	Variance
Sales rank	200	49	1	50	25.50	14.467	209.296
Reviews average sensibility	200	.684	-.065	.619	.366	.127	.016
Rating index	200	7	93	100	97.28	1.122	1.258
Total reviews	200	789950	50	790000	25689.70	76458.628	5845921775.789
Reviews photos	200	451	49	500	458.82	102.448	10495.549
“Good” reviews	200	769950	50	770000	25120.91	75106.030	5640915715.921
“Neutral” reviews	200	19000	0	19000	399.97	1610.086	2592376.778
“Poor” reviews	200	6500	0	6500	228.295	630.2769	397249.013
Price	200	36467	33	36500	3484.89	4989.758	24897688.815

&lt;Table 5&gt; Model 1 Correlation Matrix

		Reviews sensibility	Review length	Elapsed review days	Reviewer ratings	Reviews photos	Review usefulness
Reviews sensibility	Correlation significance N	1 .000 192764					
Review length	Correlation significance N	-.212** .000 192764	1 192764				
Elapsed review days	Correlation significance N	-.214** .000 192764	.049** .000 192764	1 192764			
Reviewer ratings	Correlation significance N	.092** .000 192764	-.033** .000 192764	.016** .000 192764	1 192764		
Reviews photos	Correlation significance N	.167** .000 192764	.361** .000 192764	.027** .000 192764	.083** .000 192764	1	
Review usefulness	Correlation significance N	-.087** .000 192764	.170** .000 192764	.198** .000 192764	0.285** .000 192764	0.285** .000 192764	1 192764

\*\* p < 0.01, \* p < 0.05, + p < 0.1.

&lt;Table 6&gt; Model 2 Correlation Matrix

		Reviews average sensibility	Rating index	Total reviews	"Good" reviews	"Neutral" reviews	"Poor" reviews	Reviews photos	Sales rank
Reviews average sensibility	Correlation significance N	1 .000 200							
Rating index	Correlation significance N	.103** .000 200	1 200						
Total reviews	Correlation significance N	.058** .000 200	.098** .000 200	1 200					
"Good" reviews	Correlation significance N	.069** .000 200	-.043** .000 200	.346** .000 200	1 200				
"Neutral" reviews	Correlation significance N	.017** .000 200	-.013** .000 200	.316** .000 200	.102** .000 200	1 200			
"Poor" reviews	Correlation significance N	-.083** .000 200	-.033** .000 200	.216** .000 200	.043** .000 200	.158** .000 200	1 200		
Reviews photos	Correlation significance N	-.085 .000 200	.110 .000 200	.198** .000 200	.093** .000 200	.058** .000 200	.165** .000 200	1 200	
Sales rank	Correlation significance N	0.395** .000 200	.185** .000 200	.078** .000 200	.145** .000 200	-.042** .000 200	-.088** .000 200	.059 .000 200	1 200

\*\* p < 0.01, \* p < 0.05, + p < 0.1.

<Table 5>를 통해 5개의 독립변수와 리뷰 유용성 간에 통계적으로 매우 유의한 상관관계가 있다는 것을 확인할 수 있다. 특히 리뷰 감성과 리뷰 유용성 간에는 부정적 상관관계가 존재한다는 사실을 알 수 있었고, 나머지 변수들은 리뷰 유용성과 긍정적 상관관계가 있음을 알 수 있었다. 또한, 다수의 독립변수들 간에도 상당히 유의한 상관관계가 있다는 결과를 도출할 수 있었는데, 후술하게 될 회귀분석 결과 섹션에서 변수들 간의 다중공선성에 대한 확인을 시행했다.

<Table 6>을 통해 7개의 독립변수와 판매 순위 간에 통계적으로 매우 유의한 상관관계가 있다는 것을 확인할 수 있다. 특히 판매 순위와 “보통” 리뷰 수, “나쁨” 리뷰 수 간에는 각각 부정적 상관관계가 존재한다는 사실을 알 수 있었고, 나머지 변수들은 대체적으로 판매 순위와 긍정적 상관관계가 있음을 알 수 있었다. 또한, 모형 2단계의 다수의 독립변수들 간에도 상당히 유의한 상관관계가 있다는 결과를 도출할 수 있었는데, 후술하게 될 회귀분석 결과 섹션에서 변수들 간의 다중공선성에 대한 확인을 시행했다.

### 5.3 회귀 분석 결과

앞에서 언급한 2단계의 모형에 대해 회귀분석을 진행하였다. 먼저 <Table 7>은 1단계 모형에 대한 회귀 분석 결과를 나타낸다. 분석 결과, 해당 상품이 경험재 속성이 많을수록, 리뷰의 길이가 길수록, 리뷰한 후 경과일이 길수록, 리뷰자의 등급이 높을수록, 사진이 있는 리뷰일수록 리뷰 유용성에 긍정적인 영향을 미쳤음을 알 수 있었다. 우리는 투표수가 특정 리뷰의 유용성을 대표한다고 판단했다는 점에서, 리뷰한 후 경과일이 길거나 리뷰자의 등급이 높을수록 투표수가 높다는 사실은 매우 상식에 부합하는 결과라고 생각한다. 그리고 사진이 있는 리뷰의 경우 리뷰

<Table 7> Regression Analyses for Model 1

Dep. variable : Review usefulness	Model 1	Model 2	Model 3
(constant)	-1.231**	.616**	.615**
Experience good		.160**	.265**
Reviews sensibility	-.223**	-.220**	-.212**
Review length	.025**	.025**	.024**
Elapsed review days	.011**	.010**	.010**
Reviewer ratings	.105**	.072**	.069**
Reviews photos	0.349**	0.245**	0.243**
Experience good× Reviews sensibility			-.156**
Experience good× Review length			.002**
Experience good× Elapsed review days			.009**
Experience good× Reviewer ratings			.028**
Experience good× Reviews photos			0.210**
N	192763	192763	192763
Adjusted R squared	.165	.166	.184
F	3372.199**	2731.295**	1955.116**

\*\*p < 0.01, \*p < 0.05, +p < 0.1.

이용자에게 보다 더 신뢰를 줄 수 있기 때문에, 투표수에 정(+)의 영향을 미친다는 결과 또한 상식적이다. 분석 결과 중 매우 흥미로운 사실은, 상품이 경험재 속성이 높을수록, 리뷰 감성이 낮을수록 리뷰 유용성에 양(+)의 영향을 미쳤다는 것이다. 이것을 종합해서 해석하면, 특정 리뷰어가 경험재 속성이 높은 상품에 부정적인 내용을 담은 리뷰 게시물을 남겼을 때 그 리뷰를 읽은 사람들은 더욱 민감하게 반응한다고 볼 수 있다. 즉, 구매 예정자들의 위험 회피 성향이 반영된 결과라고 해석할 수 있다.

우리는 기본 모형에 대한 회귀 분석에 그치지 않고, 상품의 경험재 속성이 각 독립 변수에 어떠한 조절 효과를 갖는지에 대한분석을 진행했다. <Table 7>에서 알 수 있듯, 모델이 발전됨에 따라 수정된 R제곱의 값이 점점 증가하고 있다.

또한, 모델 1과 모델 3간의 수정된 R제곱 차이에 대한 유의성 검정(F-test) 결과, 그 차이가 매우 유의함을 알 수 있었다(F value : 0.0426, p-value < 0.01). 이는 상품의 경험재 속성이 뚜렷한 조절 효과를 갖는다는 것을 나타낸다.

<Table 7>에서 도출된 회귀 분석 결과에서, 독립 변수 간에 다중 공선성이 존재하였는지 확인하기 위해 분산 팽창 인수(VIF, Variance Inflation Factor) 확인을 통한 다중공선성 검정을 시행했다. 본 연구는  $0 < VIF < 10$ 일 때 변수들은 공선성이 존재하지 않음을 나타내고,  $10 < VIF < 100$ 일 때 변수들의 공선성은 강함을 나타내며,  $VIF > 100$ 일 때는 변수들의 공선성은 매우 강함을 나타낸다고 가정했다. 다중 공선성 검정에 대한 결과는 <Table 8>을 통해 알 수 있다. 분석 결과, 어떠한 독립변수 간에도 다중공선성이 없음을 확인하였고, 따라서 <Table 7>에서 도출된 회귀분석 결과에 대한 신뢰성을 제고하였다.

우리는 <Table 7>에서 도출된 회귀 분석 결과를 더 세분화해서 분석해 본다면 더 의미 있는 시사점을 발견해낼 수 있다고 생각했고, 리뷰 이용자들이 리뷰에 부여한 점수를 바탕으로 “좋은” 리뷰, “보통” 리뷰, “나쁨” 리뷰의 3 그룹으로 분류했다. 분류 결과 총 192,763개의 리뷰가 각각 순서대로 163,225개, 14,516개, 15,022개로 분류되었다. <Table 9>~<Table 11>은 세분화된 분류 결과를 나타낸다.

<Table 9>~<Table 11>에서 도출된 결과 중 <Table 7>의 전체 결과 구분되는 결과는, “좋은” 리뷰의 경우 다른 두 그룹과는 반대로 리뷰 감성이 리뷰 유용성에 통계적으로 유의한 양(+)의 영향을 미친다는 것이다. 즉, 긍정적인 내용을 내포하는 리뷰의 경우 리뷰 이용자들은 그 내용이 긍정적일수록 해당 리뷰가 유용하다고 판단한다는 것을 알 수 있었다. <Table 9>, <Table 10>의 결과와 종합해서 해석하면, 리뷰 사용자

<Table 8> Multicollinearity Test for Model 1

Independent	Tolerance	VIF
model 1		
Reviews sensibility	0.561	1.537
Review length	0.234	3.413
Elapsed review days	0.453	2.945
Reviewer ratings	0.391	3.682
Reviews photos	0.212	3.246
Model 2		
Experience good	0.312	2.692
Reviews sensibility	0.618	1.532
Review length	0.233	3.691
Elapsed review days	0.410	2.941
Reviewer ratings	0.359	2.322
Reviews photos	0.214	4.912
Model 3		
Experience good	0.325	3.124
Reviews sensibility	0.289	4.568
Review length	0.512	1.574
Elapsed review days	0.321	2.815
Reviewer ratings	0.284	4.945
Experience good× Reviews sensibility	0.256	5.162
Experience good× Review length	0.381	2.487
Experience good× Elapsed review days	0.531	1.729
Experience good× Reviewer ratings	0.236	4.479
Experience good× Reviews photos	0.414	1.595

들은 매우 긍정적인 리뷰이거나, 매우 부정적인 리뷰를 더욱 유용하다고 생각한다. 만약 특정 리뷰가 중립적인 내용을 담을 경우 상대적으로 덜 유용하다고 판단하는 것으로 보인다. 결국 리뷰는 다소 자극적인 내용을 가질 때 비로소 그 유용성을 극대화할 수 있다는 것이다. 이는 매우 상식에 부합하고 논리적으로 이해 가능한 분석 결과라고 보인다. 일반적으로 누군가 특정 상품을 적극적으로 권유하거나 부정할 때 우리는 대체적으로 더욱 영향을 받게 마련이기 때문이다.

〈Table 9〉 “Good” Reviews Regression Results

Dep. variable : Review usefulness	Model 1	Model 2	Model 3
(constant)	-1.321**	.532**	.530**
Experience good		.145**	.312**
Reviews sensibility	.125**	.125**	.192**
Review length	.023**	.023**	.023**
Elapsed review days	.011**	.011**	.011**
Reviewer ratings	.107**	.105**	.058**
Reviews photos	0.369**	0.249**	0.249**
Experience good× Reviews sensibility			.136**
Experience good× Review length			.003**
Experience good× Elapsed review days			.009**
Experience good× Reviewer ratings			.036**
Experience good× Reviews photos			0.225**
N	163225	163225	163225
Adjusted R squared	.164	.164	.182
F	3122.124**	2746.153**	1512.466**

〈Table 11〉 Regression Results(“Poor” Reviews)

Dep. variable : Review usefulness	Model 1	Model 2	Model 3
(constant)	-1.512**	.682**	.676**
Experience good		.162**	.282**
Reviews sensibility	-.421**	-.421**	-.485**
Review length	.253**	.262**	.261**
Elapsed review days	.028**	.028**	.028**
Reviewer ratings	.212**	.275**	.274**
Reviews photos	0.336**	0.242**	0.242**
Experience good× Reviews sensibility			-.446**
Experience good× Review length			.002**
Experience good× Elapsed review days			.009**
Experience good× Reviewer ratings			0.207**
Experience good× Reviews photos			
N	15022	15022	15022
Adjusted R squared	.168	.168	.191
F	1976.642**	1942.665**	1175.965**

〈Table 10〉 Regression Results(“Poor” Reviews)

Dep. variable : Review usefulness	Model 1	Model 2	Model 3
(constant)	-1.238**	.521**	.518**
Experience good		.163**	.285**
Reviews sensibility	-.120**	-.120**	-.231**
Review length	.051**	.052**	.052**
Elapsed review days	.015**	.015**	.015**
Reviewer ratings	.103**	.092**	.065**
Reviews photos	0.312**	0.223**	0.223**
Experience good× Reviews sensibility			-.172**
Experience good× Review length			.002**
Experience good× Elapsed review days			.008**
Experience good× Reviewer ratings			.037**
Experience good× Reviews photos			0.205**
N	14516	14516	14516
Adjusted R squared	.168	.168	.174
F	1535.592**	1125.579**	858.129**

다음으로 <Table 12>는 모형 2단계에 대한 회귀 분석 결과를 나타낸다. 즉, 모형 1단계에서 종속 변수로서 사용 됐던 리뷰 유용성(리뷰 투표 수)이 상품 판매 순위에 어떠한 영향을 미치는지 중점적으로 분석했고, 앞으로는 해당 내용에 대한 결과를 자세히 서술할 것이다. 본 연구는 상품 판매 순위가 상품매출을 대표한다고 가정했고, 결국 리뷰의 유용성에 따른 차이가 상품의 판매에 영향을 미치는지에 대해 관찰하였다. 변수의 쓰임을 기준으로 모형 1단계에서 진행되었던 회귀 분석과의 다른 점은, 상품 가격, 평가 지수, 총 리뷰 수의 변수가 새로 추가 반영되었다는 것이고, 리뷰 감성은 리뷰 감성 평균치로 대체되었다. 이렇게 새로운 변수들을 추가한 이유는, 모형 2단계의 회귀 분석은 그 대상이 리뷰가 아니라 상품 자체이기 때문이다. 따라서 분석 시 적용 대상은 200개의 상품이다. 또한, 모형 1단계와의 차이점은 “좋은”, “보통”, 그리고 “나쁨”에 대한 리뷰 수를 개별적인 독립 변수로 사용했다는 점이다.

〈Table 12〉 Regression Analyses for Model 2

Dep. Variable : Sales ranking	M1	M2	M3	M4	M5
(Constant)	4.652**	25.453**	17.409**	23.844**	17.409**
Experience good		.079**		.115**	
Product price			-.084**		-.063*
Reviews average sensibility	7.945*	-7.521	8.124	-5.945	5.412*
Rating index	.142**	.138**	.112**	.173**	.110**
Total reviews	.241**	.239**	.236**	.221**	.233**
Reviews photos	-.014	-.013	-.022	-.013	-.022
“Good” reviews	.175**	.173**	.185**	.171**	.177**
“Neutral” reviews	-.008*	-.008*	-.012	-.008*	-.013
“Poor” reviews	-.075**	-.074**	-.063*	-.070**	-.062*
Experience good×Reviews average sensibility				-1.463	
Experience good×Rating index				.021**	
Experience good×Total reviews				.009**	
Experience good×Reviews photos				-.003	
Experience good×“Good” reviews				.016**	
Experience good×“Neutral” reviews				.002*	
Experience good×“Poor” reviews				-.003**	
Product price×Reviews average sensibility					2.124
Product price×Rating index					.046*
Product price×Total reviews					.082
Product price×Reviews photos					-.012
Product price×“Good” reviews					.047
Product price×“Neutral” reviews					.023
Product price×“Poor” reviews					-.075
N	200	200	200	200	200
Adjusted R squared	.216	.222	.233	.248	.227
F	9.209**	8.017**	8.466**	5.637**	5.714

2단계에 대한 회귀 분석 결과를 모델 1을 기준으로 설명하면 가장 먼저, 모형 1단계에서는 리뷰 유용성에 부정적인 영향을 미쳤던 리뷰 감성 변수가 상품 판매 순위에는 긍정적인 영향을 미쳤다. 즉, 특정 상품에 대한 리뷰가 대체적으로 긍정적이라면 상품매출에 기여한다고 볼 수 있다. 이 결과와 모형 1단계에서 도출된 결과를 종합하면, 리뷰의 내용이 자극적일수록 리뷰 이용자들은 민감하게 반응하지만, 결국 긍정적인 내용을 담은 리뷰가 대체적으로 상품매출에 기여한다. 그리고 평가 지수, 총 리뷰 수, “좋음” 리뷰 수가 상품 판매 순위에 양(+)의 영향을 미치는 것으로 드러났고,

상식에 부합하는 결과라고 볼 수 있다. 반대로, 상품 가격과 “나쁨” 리뷰 수가 상품 판매 순위에 음(-)의 영향을 미친다는 결과 또한 충분히 합리적으로 추론 가능한 결과이다.

본 연구에서 중점적으로 보고자 했던 독립 변수들과 상품 경험재 속성, 상품 가격 간의 상호작용 효과에 대한 분석 결과는 다음과 같다.

모형 1과에서 모형 4로 발전됨에 따라 수정 R제곱의 값이 커지기 때문에 상품의 경험재 속성은 조절 효과가 있다고 볼 수 있다(F-value : 0.2419, p-value < 0.01). 또한 모형 1에서 모형 5로 발전됨에 따라 수정 R제곱의 값이 커지고 그 차이에 대한

〈Table 13〉 Multicollinearity Test for Model 2

Independent	Tolerance	VIF
model 1		
Reviews average sensibility	0.467	1.894
Rating index	0.346	2.691
Rating index	0.264	2.913
Reviews photos	0.246	3.124
“Good” reviews	0.365	2.245
“Neutral” reviews	0.168	5.815
“Poor” reviews	0.282	2.532
model 5		
Reviews average sensibility	0.327	2.248
Rating index	0.356	2.137
Total reviews	0.373	2.459
Reviews photos	0.267	1.595
“Good” reviews	0.265	2.873
“Neutral” reviews	0.179	4.435
“Poor” reviews	0.310	1.762
Product price× Reviews average sensibility	0.421	1.687
Product price×Rating index	0.573	1.518
Product price×Total reviews	0.387	2.172
Product price×Reviews photos	0.297	2.473
Product price×“Good” reviews	0.265	2.234
Product price×“Neutral” reviews	0.378	1.923
Product price×“Poor” reviews	0.281	2.387

유의성도 확인됐기 때문에(F-value : 0.0266, p-value < 0.01) 상품 가격 또한 조절 효과가 있는 것으로 드러났다. 분석 결과 상품 경험재 속성은 평가지수, 총 리뷰 수, “좋은” 리뷰 수, “보통” 리뷰 수, “나쁨” 리뷰 수가 상품 판매 순위에 미치는 영향력을 낮추는 조절 작용을 한 것으로 드러났다. 즉, 양(+)과 음(-)의 영향력 여부에 상관없이 그 정도와 크기를 낮추었다.

모형 1단계의 회귀 분석 결과에 적용했던 분산 팽창 인수(VIF, Variance Inflation Factor) 검정 기준을 동일하게 모형 2단계의 회귀 분석 결과에도 적용했다. 검정 결과 모든 개별 변수의 VIF는 0과 10사이의 값을 가졌고, 따라서 모형 2단계의 회귀 분석 시 독립 변수들 간에는 다중공선성이 없다고 판단했다.

〈Table 14〉 Regression Result for Model 2

Dep. variable :	Model 1	Model 2	Model 3
Review usefulness			
(Constant)	17.294**	25.453**	23.844**
Experience good		.079**	.115**
Reviews average sensibility	19.652**	19.578**	15.834**
Rating index	.112**	.110**	.089**
Total reviews	.233**	.232**	.198**
Reviews photos	-.020	-.013	-.013
“Good” reviews	.185**	.176**	.141**
“Neutral” reviews	-.012	-.011*	-.008*
“Poor” reviews	-.062*	-.061**	-.048**
Product price× Reviews average sensibility			6.873**
Product price ×Rating index			.028**
Product price× Total reviews			.009**
Product price× Reviews photos			-.003
Product price× “Good” reviews			.018**
Product price× “Neutral” reviews			-.002*
Product price× “Poor” reviews			-.003**
N	200	200	200
Adjusted R squared	.245	.252	.281
F	8.466**	8.231**	5.851**

〈Table 14〉의 모형 2단계 회귀 분석은 〈Table 11〉의 회귀 분석을 수정 및 발전시킨 회귀 분석 결과표이다. 그리고 리뷰 감성 점수 값과 모형 1단계에서 도출된 리뷰 유용성 값을 곱한 뒤 생성된 값의 평균치를 새로운 변수로 적용시켰다. 〈Table 12〉의 회귀분석 결과와 〈Table 14〉의 결과를 비교한 내용은 다음과 같다. 먼저 상품 경험재 속성은 뚜렷하게 상품 판매 순위에 양(+)의 영향을 미치는 것으로 드러났고, 반대로 “보통” 리뷰 수와 “나쁨” 리뷰 수는 보다 더 현저히 상품 판매 순위에 통계적으로 유의한 음(-)의 영향을 미치는 것을 확인할 수 있었다. 하지만, 여

전히 사진 있는 리뷰 수는 종속 변수에 유의한 영향을 미치지 못했다. 그리고 주목해야 할 점은 “(리뷰 감성×리뷰 유용성)의 평균치” 변수가 상품 매출에 통계적으로 매우 유의한 양(+)의 영향을 미친다는 사실이다(회귀계수 : 15.834, p-value < 0.01).

또한 <Table 14>의 모델이 발전함에 따라 수정 R제곱의 값이 뚜렷하게 증가했고(모델 1과 모델 3 간의 수정된 R제곱의 차이 : F value : 0.0938, p-value < 0.01), 이는 상품 경험재 속성이 조절 효과를 갖는다는 사실을 나타낸다. 즉, 특정 상품의 경험재 속성에 따라 해당 상품에 대한 리뷰가 상품 판매 순위에 미치는 영향이 다를 수 있다는 사실을 발견했다. 이는 상품 리뷰가 실질적으로 상품매출에 영향을 미친다는 것을 의미하고, 특히 해당 상품의 경험재 여부가 그 영향의 크기에 유의한 영향을 준다는 것을 나타낸다.

본 연구는 2단계의 연구 모형을 수립했고 데이터를 수집한 뒤 각각의 모형에 대한 회귀 분석을 시행했다. 회귀 분석 결과를 바탕으로, 분석 전에 설정했던 다수의 가설들에 대한 검증을 해볼 수 있었고 그 결과는 <Table 15>, <Table 16>에 나타나어 있다. <Table 15>는 연구 모형 1단계에 대한 가설 검증, 즉 종속 변수를 리뷰 유용성으로 설정한 연구 모형에 대한 가설 검증 결과를 나타낸다. <Table 16>은 연구 모형 2단계에 대한 가설 검증, 즉 종속 변수를 상품 판매 순위로 설정한 연구 모형에 대한 가설 검증 결과를 나타낸다.

## 6. 결 론

본 연구는 각각 리뷰 유용성과 상품 판매 순위를 종속변수로 설정한 2단계의 회귀 모형을 수립했고, 다수의 변수를 기준으로 횡단면의 데이터를 수집한 후 회귀 분석을 진행했다. 아래의 <Table 15>와 <Table 16>은 연구자들이 연구

<Table 15> Hypothesis Test Results for Model 1

Hypotheses	Expected	Revealed
H1	Positive	Accepted
H1-1	Positive	Accepted
H2	Positive	Accepted
H2-1	Positive	Accepted
H3	Positive	Accepted
H3-1	Positive	Accepted
H4	Negative	<b>Rejected</b>
H4-1	Positive	Accepted
H5	Positive	Accepted
H5-1	Positive	Accepted

<Table 16> Hypothesis Test Results for Model 2

Hypotheses	Expected	Revealed
H6	Positive	Partially Accepted
H6-1	Positive	Partially Accepted
H6-2	Positive	<b>Rejected</b>
H7	Positive	Accepted
H7-1	Positive	Accepted
H7-2	Positive	<b>Rejected</b>
H8	Positive	Accepted
H8-1	Positive	Accepted
H8-2	Positive	<b>Rejected</b>
H9	Positive	Accepted
H9-1	Positive	Accepted
H9-2	Positive	<b>Rejected</b>
H10	Negative	Accepted
H10-1	Positive	Accepted
H10-2	Positive	<b>Rejected</b>
H11	Positive	<b>Rejected</b>
H11-1	Positive	Accepted
H11-2	Positive	<b>Rejected</b>
H12	Positive	<b>Rejected</b>
H12-1	Positive	<b>Rejected</b>
H12-2	Positive	<b>Rejected</b>

초기에 설정했던 가설 내용과 예상 관계에 대한 검증 결과를 간략히 나타낸 표이다. <Table 15>는 1단계 모형에 대한 가설 검증 결과를 나타내고, <Table 16>은 2단계 모형에 대한 가설 검증 결과를 나타낸다.



## 6.1 연구 결과 요약

모형 1단계에 대한 분석 결과, 리뷰 감성, 리뷰자 등급, 리뷰의 길이, 리뷰한 후 경과일, 리뷰 사진 여부는 모두 리뷰 유용성에 통계적으로 유의한 양(+)<sup>1)</sup>의 영향을 미쳤다. 따라서 4번째 가설을 제외한 모든 가설은 기각되지 않고 채택되었다. 우리는 리뷰한 경과일이 길수록 해당 리뷰에 대해 리뷰 이용자들은 유용하지 못한 정보라고 판단할 것이라 예상했지만, 분석 결과 오래된 리뷰일수록 유용성이 높다고 판단되었다.

또한, 경험재 속성이 강할수록 분석에 사용된 독립변수가 종속변수에 양(+)<sup>2)</sup>의 영향을 미칠 것이라는 본 연구의 가설은 모두 채택되었다. 즉, 상품의 경험재 속성이 강할수록 리뷰 감성, 리뷰자 등급, 리뷰의 길이, 리뷰한 후 경과일, 리뷰 사진 여부는 리뷰 유용성을 제고시키는 것으로 드러났다.

모형 2단계에 대한 분석 결과, 리뷰 감성, 평가 지수, 총 리뷰 수, “좋음” 리뷰 수는 상품 판매 순위에 통계적으로 유의한 양(+)<sup>3)</sup>의 영향을 미치는 것으로 드러났다. 반대로 “나쁨” 리뷰 수, “보통” 리뷰 수는 모두 상품 판매 순위에 음(-)<sup>4)</sup>의 영향을 미친다는 사실을 알 수 있었다. 하지만 사진을 포함한 리뷰 수가 상품 판매에 유의한 영향을 미친다는 사실을 밝힐 수 없었다. 따라서 해당 가설은 기각되었다. 또한 조절 변수로서의 ‘상품의 가격’ 변수는 그 유의성이 입증되지 않아 해당 가설 또한 기각되었다. 하지만 상품의 경험재 속성은 조절 효과를 갖는 것으로 드러났고 관련 가설들은 채택되었다.

## 6.2 연구 의의

실증적으로 검증된 본 연구의 연구 결과는 학문적 관점에서 의의를 지니고 산업적 관점에서는 시사하는 바가 있다.

먼저 학문적 관점에서의 의의는 다음과 같다.

첫째, 일련의 2단계 연구 모형을 수립함으로써 기존의 단일적인 연구 모형 수립에 더 나아가, 포괄적이고 단계적인 연구 모형을 수립하였다. 리뷰 유용성을 연구한 기존의 선행 연구들은 대부분 유용한 리뷰의 특징에 대해 연구하였거나, 단일 리뷰 시스템이 상품의 판매량에 미치는 영향을 분석하는 것에 그쳤다. 하지만 본 연구는 유용한 리뷰를 결정짓는 요소를 연구했음에 더하여 궁극적으로 상품 판매에 영향을 미치는 요인들을 탐색하고 발견했다.

둘째, 기존의 많은 연구들은 온라인 리뷰의 수치적 특징에 기반해서 상품 판매량에 영향을 주는 요인들을 연구했지만, 본 연구는 리뷰의 질적 내용에 대한 감성 분석 모형을 이용하여 리뷰 텍스트를 분석하고, 수치화 한 후 계량 분석을 진행했다. 이때 파이썬 분석 도구를 적극 활용함으로써 보다 더 실증적인 연구 접근이 가능했다.

셋째, 본 연구는 데이터 수집의 대상으로 중국에서 두 번째로 큰 전자상거래 사이트인 Jingdong을 선택했고, 이 선택은 연구 대상 범위를 중국 전 지역으로 확대했다는 점에서 연구 결과 설득력을 제고할 수 있었다.

넷째, 본 연구는 상품의 경험재 속성과 상품 가격을 조절 변수로서 회귀 모형에 적용시켰다. 회귀 분석 결과 상품의 경험재 속성은 다른 주요 변수들 간의 상호작용 관계가 있음을 알 수 있었고, 따라서 리뷰 이용자들에게 특정 상품의 경험재 여부는 관련 리뷰들을 수용함에 있어 매우 중요한 변수라는 것을 발견했다.

다음으로 산업적 관점에서의 시사점은 다음과 같다.

첫째, 본 연구의 결과는 온라인 리뷰 이용자, 특히 중국의 전자상거래 기업들에게 유용한 참고 문헌이 될 수 있다. 중국의 인터넷 발전 현황 및 향후 몇 년간의 발전방향을 봤을 때 중국의

인터넷 이용자 규모는 계속해서 증가하는 추세를 보이고 있으며, 전자상거래의 거래액도 계속하여 증가하고 있다. 그 중 농촌 지역의 인터넷 이용자 집단이 향후 인터넷쇼핑의 주요 소비자가 될 가능성이 높으며, 관련 지식에 대한 부족이나 검색의 한계로 인하여 온라인 리뷰에 대한 수요가 많을 것으로 예상된다. 따라서 리뷰 이용자들이 특정 상품에 대한 리뷰를 어떻게 인식하는지를 파악한다면 잠재적 소비자들의 요구를 만족시킬 수 있을 것이다.

둘째, 본 연구의 결과는 전자상거래 기업들의 관점에서 상품매출에 효과적인 리뷰 프레임워크를 구축할 수 있는 지침 자료로 활용될 수 있다. 우리는 상품의 경험재 속성에 따라 상품매출에 차등적인 영향을 미친다는 사실을 발견했다. 이것을 활용하면, 예를 들어서 경험재로 분류가 되는 상품과 탐색재로 분류되는 상품에 대한 리뷰 프레임을 달리 구축할 수 있을 것이다.

### 6.3 한계점

본 연구의 한계점은 먼저, 우리는 표본을 제한적으로 수집하였고 선택한 상품의 수량도 상대적으로 적었다. jd.com에서 판매중인 텔레비전, 디지털 카메라, 외장하드, 메모리 카드만을 기준으로 관련 리뷰를 수집했기 때문에 연구 결과를 모든 상품에 적용시키는 것은 한계가 있다. 향후의 연구는 더 많은 종류의 상품을 대상으로 표본을 수집한 후 연구를 진행한다면 더욱 신뢰성을 제고할 수 있을 것이다. 다음으로, 탐색재와 경험재의 구분에 대한 모호함이다. 일부 소비자들에게 탐색재 상품과 경험재 상품에 대한 구분은 명확하지 않을 수 있다. 따라서, 향후의 연구를 진행할 때는 사전에 간단한 설문조사를 통해 소비자 관점에서의 상품 분류 체계를 모형에 반영하면 더 타당한 연구 결과를 얻을 수 있을 것이다.

## References

- [1] Basuroy, S., Chatterjee, S., and Ravid, S. A., "How critical are critical reviews? The box office effects of film critics, star power, and budgets", *Journal of Marketing*, Vol. 67, No. 4, 2003, pp. 103-117.
- [2] Berger, C. R. and Calabrese, R. J., "Some explorations in initial interaction and beyond : Toward a developmental theory of interpersonal communication", *Human Communication Research*, Vol. 1, No. 2, 1974, pp. 99-112.
- [3] Cao, Q., Duan, W., and Gan, Q., "Exploring determinants of voting for the 'helpfulness' of online user reviews : A text mining approach", *Decision Support Systems*, Vol. 50, No. 2, 2011, pp. 511-521.
- [4] Channel Advisor, 2010 Consumer Shopping Habits Survey, 2010.
- [5] Channel Advisor, 2011 Global Consumer Shopping Habits Survey, 2011.
- [6] Chatterjee, P., "Online review : do consumers use them?", *Advances in Consumer Research*, Vol. 28, 2001, pp. 129-133.
- [7] Chen, P. Y., Wu, S. Y., and Yoon, J., "The impact of online recommendations and consumer feedback on sales", *Proceedings of the International Conference on Information Systems*, 2004, pp. 711-724.
- [8] Chen, Y., Fay, S., and Wang, Q., "Marketing implications of online consumer product reviews", *Business Week*, Vol. 7150, 2003, pp. 1-36.
- [9] Chevalier, J. A. and Mayzlin, D., "The effect of word of mouth on sales : Online book

- reviews”, *Journal of Marketing Research*, Vol. 43, No. 3, 2006 pp. 345-354.
- [10] Clemons, E. K., Gao, G. G., and Hitt, L. M., “When online reviews meet hyperdifferentiation : A study of the craft beer industry”, *Journal of Management Information Systems*, Vol. 23, No. 2, 2006, pp. 149-171.
- [11] Dellarocas, C., “The digitization of word of mouth : Promise and challenges of online feedback mechanisms”, *Management science*, Vol. 49, No. 10, 2003, pp. 1407-1424.
- [12] Duan, W., Gu, B., and Whinston, A. B., “Do online reviews matter? An empirical investigation of panel data”, *Decision support systems*, Vol. 45, No. 4, 2008, pp. 1007-1016.
- [13] Elliott, K. M., *Understanding consumer-to-consumer influence on the web*(Doctoral dissertation), 2004.
- [14] Forman, C., Ghose, A., and Goldfarb, A., “Competition Between Local and Electronic Markets : How the Benefit of Buying Online Depends on Where You Live”, *Management Science*, Vol. 55, No. 1, 2009, pp. 47-57.
- [15] Godes, D. and Mayzlin, D., “Using online conversations to study word-of-mouth communication”, *Marketing Science*, Vol. 23, No. 4, 2004, pp. 545-560.
- [16] Granovetter, M., “The strength of weak ties”, *American Journal of Sociology*, Vol. 78, No. 6, 1973, pp. 1360-1380.
- [17] Hu, N., Liu, L., and Zhang, J. J., “Do online reviews affect product sales? The role of reviewer characteristics and temporal effects”, *Information Technology and Management*, Vol. 9, No. 3, 2008, pp. 201-214.
- [18] Kim, C. H., “The Influences of Internet WOM (word Of Mouth) Information Valence and Contents to WOM Effects-Focus on Cell-Phone Purchasing for Student via Online WOM”, *Korea Association for International Commerce and Information*, Vol. 8, No. 4, 2006, pp. 23-41.
- [19] Mudambi, S. M. and Schuff, D., “What makes a helpful online review? A study of customer reviews on Amazon.com”, *MIS Quarterly*, Vol. 34, 2010, pp. 185-200.
- [20] Nelson, P., “Advertising as Information”, *Journal of Political Economy*, Vol. 82, No. 4, 1974, pp. 729-754.
- [21] Park, J. Y. and Yu, Y. S., “A Study on the Influence of Online Word-Of-Mouth Information’s Direction on Word-Of-Mouth Effects : Focusing on the Moderating Effects of Brand Recognition”, *Korea Information Technology Service Industry Association*, Vol. 8, No. 3, 2011, pp. 15-28.
- [22] Reichheld, F. F., “The one number you need to grow”, *Harvard Business Review*, Vol. 81, No. 12, 2003, pp. 46-55.
- [23] Son, J. A. and Rlee, E. Y., “The effects on Information Characteristics(Direction, Consensus) on Word-of-Mouth Performance in Online Apparel Shopping”, *Journal of the Korean Society of Clothing and Textiles*, Vol. 31, No. 8, 2007, pp. 1157-1167.
- [24] Son, J. Y. and Eo, S. H., “An Effect of Directiveness of Online Transmissive Information When Purchasing Cosmetics Online on Purchase Intention of Consumer”, *Journal of Korea Social Cosmentics*, Vol. 14, No. 3, 2008, pp. 929-945.

## ■ 저자소개



Chim Hwang

Chim Hwang received his Master's degree in Information Systems from School of Business, Hanyang University in 2018. He is currently working

on a Korea-China trading company. His research interests are electronic commerce and social media.



Young-Jin Kwon

Young-Jin Kwon received his Bachelor's degree from School of Business, Hanyang University and he is currently a Master's student in Information

Systems of the same school. His major research interests are information security and privacy, IT investment, and social media.



Sang-Yong Tom Lee

Sang-Yong Tom Lee is a Professor in School of Business, Hanyang University, Seoul, Korea. He was on the faculty of Department of Information

Systems, National University of Singapore before joining Hanyang University. He received his Ph.D degree from Texas A&M University (1999). His research interests include economics of information systems, online information privacy, and value of IT. His papers have been published in Management Science, MIS Quarterly, Journal of Management Information Systems, Communications of the ACM, IEEE Transactions on Engineering Management, Information and Management, and others.