

## 온라인 리뷰의 경제적 효과, 유용성과 유용성 투표수에 영향을 주는 결정요인\*

이상재\*\* · 최준연\*\*\* · 최진호\*\*\*\*

### The Determinant Factors Affecting Economic Impact, Helpfulness, and Helpfulness Votes of Online\*

Sangjae Lee\*\* · Joon Yeon Choeh\*\*\* · Jinho Choi\*\*\*\*

#### ■ Abstract ■

More and more people are gravitating to reading products reviews prior to making purchasing decisions. As a number of reviews that vary in usefulness are posted every day, much attention is being paid to measuring their helpfulness. The goal of this paper is to investigate firstly various determinants of the helpfulness of reviews, and intends to examine the moderating effect of product type, i.e., search or experience goods on the product sales, helpfulness and helpfulness votes of online reviews. The determinants include product data, review characteristics, and textual characteristics of reviews. The results indicate that the direct effect exists for the determinants of product sales, helpfulness, and helpfulness votes. Further, the moderating effects of product type exist for these determinants on three dependent variables. The results of study will identify helpful online review and design review sites effectively.

Keyword : Sales, Helpfulness, Helpfulness Votes, Determinant Factors, Search Goods, Experience Goods

논문투고일 : 2013년 11월 07일      논문수정완료일 : 2014년 03월 13일      논문게재확정일 : 2014년 03월 17일

\* 본 논문은 2012년도 세종대학교 교내연구비 지원에 의해서 수행되었습니다.

\*\* 세종대학교 경영대학 e-비즈니스 전공 교수, 교신저자

\*\*\* 세종대학교 디지털콘텐츠학과 조교수

\*\*\*\* 세종대학교 경영학과 부교수

## 1. 서 론

온라인 소비자들의 리뷰가 구매 과정에서 담당하는 역할이 더욱 커지고 있다. 온라인 소비자들의 리뷰들은 회사 또는 제삼자 웹사이트들에 게재된 상품 평가들로써 정의될 수 있다. 대부분의 존재하는 추천 사이트들은 “별 순위” 등급과 함께 상품들의 순위만을 제공하기 때문에 사용자들은 리뷰 문장들이 내포하는 ‘실재 의미들’ 그리고 리뷰작성자의 구체적인 상품에 대한 감정들, 경험들 그리고 의견들을 이해하지 못할 수 있다. 그러므로 상품 리뷰들은 잠재적인 소비자들이 구매결정을 하는데 상품 등급보다 더 유용하다. 선행의 연구들은 순위(ordinal) 로지스틱 회귀분석[4], 최소자승회귀모형, 유용성의 로그변환[11], 유용성을 예상하기 위한 서포트 벡터 회귀 알고리즘[29] 같은 계량경제학의 방법들을 사용했다. 연구되지 않은 주제 중 하나는 소비자들에게 구매결정을 하게하는 과정에 영향을 미치는 소비자들의 리뷰들의 특성이다. 유용성 혹은 상품매출에 대한 고객의 인지된 위험에 영향을 미치는 상품 타입의 중요성[19, 24]을 고려해서, 본 연구는 상품 타입(탐색제 및 경험제)의 조절효과를 검증함으로써 유용성 예측에 대한 이전의 연구들을 확장시키고자 하였다. 즉, 상품타입이 상품매출, 리뷰의 유용성, 유용성(유용하다고 평가한) 투표수에 미치는 조절효과를 검증하였다. 이것은 리뷰들의 유용성에 대한 제조자 지향적 평가 매커니즘(즉, 경제적 영향에 대한 추정)과 소비자 지향적 매커니즘(즉, 사회적 영향 추정)에 대한 이해를 개선시킬 것이다. 유용성과 유용성 투표수를 구분하여 분석하는 이유는 총 투표수가 아주 적은 경우에는 유용성이 상품의 실제 유용성을 나타내는데 한계가 있을 수도 있기 때문이다. 본 논문은 리뷰들의 유용성에 대한 상품 데이터, 리뷰 특징들 그리고 구문의 특징들을 포함하는 결정변수의 영향을 조사하였다. 본 논문은 결정변수들, 판매량, 유용성 등에 대한 데이터를

아마존닷컴으로부터 공개적으로 추출하여 사용하였다. 이로써 본 논문은 몇몇 리뷰들은 적거나 전혀 유용성 투표를 받지 못하는 반면에 다른 리뷰들은 많은 유용성 투표를 받는지 이유를 조사하고자 한다. 본 논문의 연구 결과는 유용한 온라인 리뷰를 확인하고, 리뷰 사이트를 효과적으로 디자인하는 것을 지원할 것이다.

## 2. 연구배경

리뷰 유용성은 리뷰의 정보 분석가능성(diagnosticity)으로서 이해될 수 있다[19]. 유용성은 기존 문헌에서 제시한 정보 분석성(diagnosticity)에 기초한 의사결정과정에서 지각된 가치의 측정척도로 여겨질 수 있다[14, 22]. 구매 결정과정의 다수의 단계에 대해서 정보분석적 가치는 소비자 리뷰들을 통해서 제공될 수 있다. 소비자 리뷰들을 통해서 구매대안들에 대한 정보를 찾아내는 능력을 제고하고 소비자들이 더 나은 구매결정을 하는 것이 지원될 수 있다. 탐색제는 소비자들이 구매이전에 상품 질에 대한 정보를 얻을 수 있는 상품들이다. 반면에 경험제는 상품 질을 평가하기 위해서 샘플링이나 구매가 요구되는 상품들이다[19]. 탐색제의 예들은 컴퓨터, 카메라를 포함하며 경험제의 예들은 알코올 음료 그리고 음악들을 포함한다. 많은 상품들에 탐색제와 경험제의 속성이 동시에 존재할 수도 있으나, 탐색제와 경험제를 구분하는 것은 여전히 효과적이고 널리 적용되는 방법이다[13]. 탐색제는 상품을 사용하기 전에 상품의 품질에 대한 정보를 비교적 쉽게 얻을 수 있는 상품으로 묘사될 수 있다. 그리고 탐색제는 객관적이고 비교하기 쉬운 품질의 속성으로 평가될 수 있다. 대조적으로, 경험제는 상품을 사용하기 전에 비교적 상품 품질에 대한 정보를 찾기가 쉽지 않다. 그리고 경험제는 주관적이고 비교하기 어려운 중요한 요소로써 품질을 평가한다. 소비자들은 탐색제와 경험제에 해당하는 상품에 대한 구매결정을 돕기

위해 온라인 리뷰를 사용한다[19].

리뷰들의 수를 기준으로 한 등급들은 긍정적으로 상품의 판매량에 영향을 준다[5, 23]. 리뷰 양은 판매량에 영향을 준다[9, 10, 17]. 더 품질 좋은 상품이 더 긍정적인 리뷰를 유도하기 때문에, 리뷰의 질은 긍정적으로 상품판매량과 소비자 구매의사와 관련된다[21].

이전의 연구들은 유용성 또는 유용성 투표에 대한 다양한 요소들을 조사했다. 예를 들어 Forman et al.[10]은 아마존닷컴의 세 가지 유형의 상품들을 조사했다(오디오와 비디오 플레이어, 디지털 카메라, DVD들). 그들은 리뷰 작성자의 정보, 리뷰 본문의 읽기 쉬움 그리고 리뷰 본문의 주관성을 조사했다. 그들은 주관적이고 객관적인 요소들의 결합을 포함한 리뷰들이 더 유용하다는 것을 보여주었다. 그들은 또한 리뷰의 읽기 쉬움이 인지된 유용성과 긍정적으로 관계된다는 것 그리고 철자 오류가 부정적으로 유용성에 영향을 미친다고 주장했다. Mudambi and Schuff[19]는 리뷰의 극단성, 리뷰 길이 그리고 제품 유형이 리뷰의 인지된 유용성에 영향을 준다는 것을 보여주었다. 리뷰에서 상품 유형은 리뷰의 유용성에 대한 리뷰 극단성이 미치는 영향을 완화시킨다. Ghose and Ipeiritos[11]는 계량경제학, 텍스트 마이닝 그리고 예측의 모델링 기술들을 사용하여 온라인 리뷰의 유용성과 경제적 영향을 평가하기 위해서 리뷰작성자의 과거 리뷰의 평균유용성 그리고 신원과 같은 리뷰작성자의 특성 그리고 리뷰 텍스트의 특징들을 조사하였다.

### 3. 판매량, 유용성, 유용성 투표수의 결정요인들

본 논문은 온라인 리뷰의 경제적 효과, 유용성, 유용성 투표수에 미치는 결정변수는 Baek et al. [3], Chevalier and Mayzlin[5], Forman et al. [10], Ghose and Ipeiritos[11], Pan and Zhang

[20]에서 제시한 변수들을 참고하여 주로 구성하였다. 유용성의 결정요인들은 상품 데이터, 리뷰 특징들, 리뷰들의 구문적 특징들을 포함한다. 상품 데이터의 변수그룹은 (1) 상품소매가격 (2) 상품 출시일로부터 경과시간 (3) 리뷰들의 수 (4) 리뷰 평가 평균점수를 포함한다.

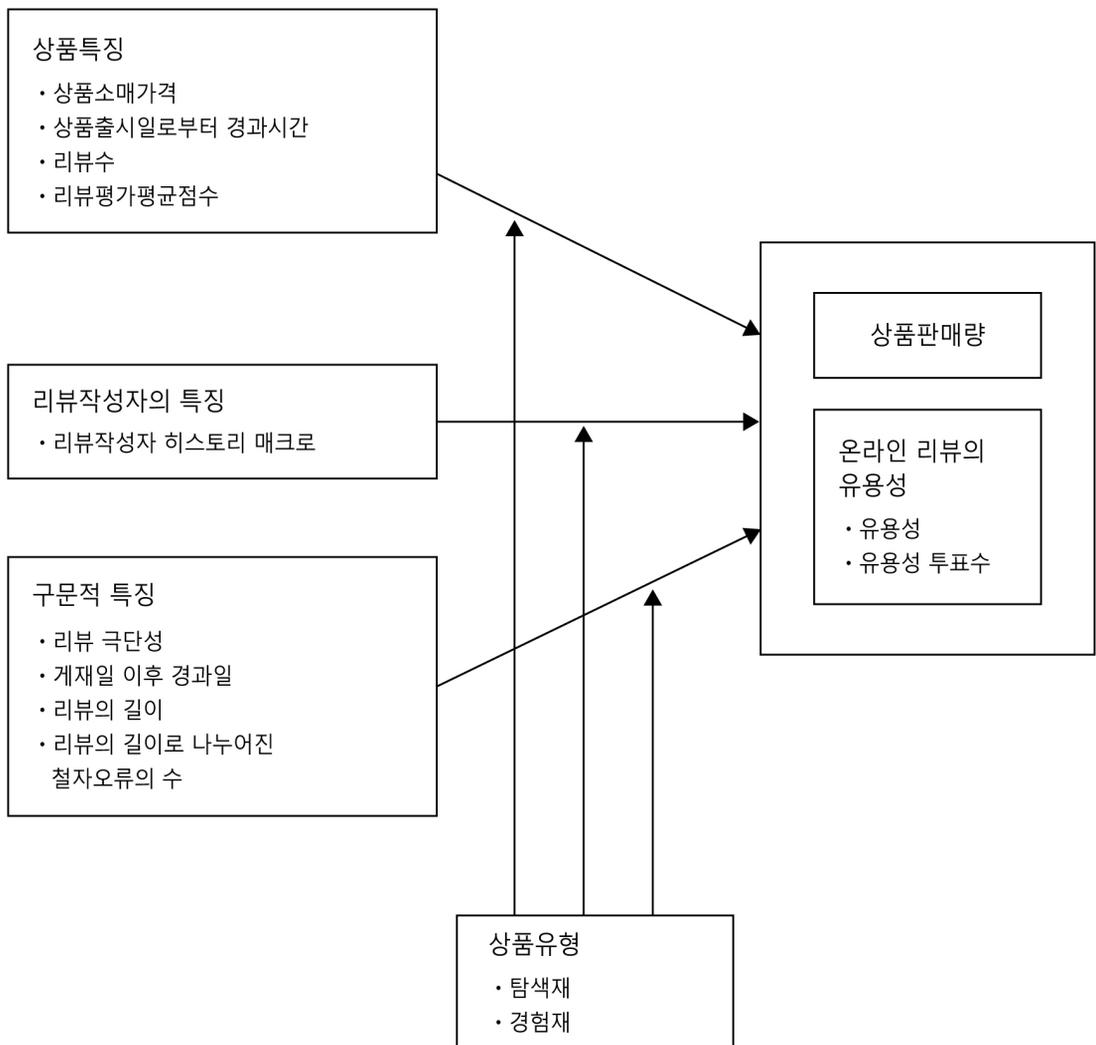
유용성을 측정하기 위해서, 사용자들은 “이 리뷰가 당신에게 유용합니까?”라는 질문에 대답할 것이 요구되었다. 유용성은 질문에 대해 “네”라고 말하는 응답자들의 비율이라고 정의된다[10]. 유용성 투표수는 해당 리뷰가 받은 총 유용성 (유용하다고 표시한) 투표수이다. 아마존 사용자평가 순위들은 [투표한 멤버들의 수]중에서 “[유용한 투표의 수] 형식으로 게재된 리뷰 옆에 실시간으로 나타내진다. 이 전체 투표수와 유용하게 평가한 투표수는 유용한 리뷰로 평가된 투표의 비율을 계산하는 것을 가능하게 한다. 구전(word-of-mouth)과 관련된 연구들의 한 흐름은 리뷰작성자들의 인지된 속성이 소비자들로 하여금 리뷰들에 응답하도록 만든다는 것을 나타낸다[10]. 본 논문은 각각의 리뷰작성자들의 과거 리뷰들과 과거리뷰들 각각의 유용성의 전체 투표수를 모았다. 리뷰작성자의 특징들은 (5) 리뷰작성자가 받은 전체 투표수를 리뷰작성자가 받은 전체 유용성 투표수로 나눈 비율로 정의되는 리뷰작성자 히스토리 매크로(reviewer history macro), (6) 리뷰작성자 랭크(“랭크”는 높아질수록 순위는 떨어지는 것으로 본 논문에서는 표기함)를 포함한다.

각각의 리뷰에서, 본 논문은 리뷰의 실제 구문적 내용과 리뷰작성자들에 의해 받은 리뷰 순위들을 발췌한다. 리뷰작성자들이 리뷰된 상품에 부여하는 순위는 별의 숫자 1개에서 5개까지로 나타내진다. 리뷰들의 구문적 특징들은 대부분 문체의 기초적인 특징들이고, 이에 대한 근거는 문체의 특징들이 얼마나 많은 유용성 투표 리뷰수를 받는데 영향을 끼치는지 연구한 Cao et al.[4], 및 Ghose and Ipeiritos[11]의 연구에 기초할 수 있다. 그러

므로, 본 논문은 (7) 리뷰작성자가 부여한 평가점수와 모든 사용자가 부여한 평가의 평균값과의 차이의 절대값으로 정의되는 리뷰 극단성 (8) 게재일 이후의 경과시간 (9) 리뷰의 길이 (10) 각각의 리뷰 안에서의 철자 오류들의 수를 리뷰의 길이(문자수를 단위)로 나눈값을 포함한다. [그림 1]은 본 논문에서 결정요인들과 종속변수 사이의 관계들을 보여준다.

결정요인들의 영향을 검증하기 위해서, 본 논문

은 Baek et al.[3], Chevalier and Mayzlin [5], Forman et al.[10], Ghose and Ipeiritis[11], Pan and, Zhang[20]에서 사용된 것과 비슷한 모델을 사용하였다. Chevalier and Mayzlin[5]는 가격, 리뷰들의 수 그리고 리뷰 valence가 아마존 그리고 반즈 · 노블 서점의 상품판매량에 어떻게 영향을 미치는지를 조사했다. 이러한 연구들에 기초하여 본 논문의 회귀분석에 사용되어질 결정요인들이 선택되어졌다.



[그림 1] 본 논문에서 결정요인과 종속변수 사이의 관계들

판매량의 랭크로 측정된 수요는 파레토 분포를 따른다(즉, a power law)[6]. 따라서 아래에 제시된 파레토 관계를 이용하여 판매량 랭크를 수요로 전환시키는 것이 가능하다.

$$\ln(D) = a + b \cdot \ln(S)$$

여기에서 D는 관찰되지 않은 상품 수요를 나타내고, S는 관찰된 판매량 랭크들을 나타낸다. 그리고  $a > 0$ ,  $b < 0$ 는 특정 산업용 변수이다. 그러므로, 아마존닷컴의 상품판매량의 랭크의 로그는 상품 수요의 로그의 대체변수(proxy)로 사용될 수 있다. 본 논문은 어떤 유형의 리뷰가 판매량에 중요한 영향을 끼치는지를 조사하려고 한다. 아마존의 판매량 랭크는 수요의 대리변수로서 제공될 수 있다. 아마존 소매가격은 소비자들의 구매 행동에 영향을 끼칠지도 모른다. 시장에 상품이 출시된 날짜는 상품의 출시일로부터 경과시간을 계산하는데 사용된다. 오래전에 출시된 상품인 경우 시간이 경과할수록 판매량이 감소하기 쉽다.

이전의 연구들은 리뷰 사이트에서 수적인 순위와 차후의 책의 판매량 사이[5, 23] 또는 리뷰의 수와 판매량 사이[9, 10, 17]의 관계들을 제안하였다. 리뷰들의 수적인 순위는 더 좋은 품질의 상품이 더 많은 긍정적인 리뷰들을 얻는다는 점에서 판매량과 긍정적으로 관계된다. 소비자들에게 유용하게 인지된 리뷰들은 판매량을 증가시킨다는 점에서 리뷰들의 수적인 순위는 회사의 이익과 연관성이 있다[5, 7].

Log(판매량 순위)<sub>kr</sub>

$$\begin{aligned} &= \alpha + \beta_1 \cdot \log(\text{소매가격})_{kr} \\ &+ \beta_2 \cdot \log(\text{상품 출시일로부터 경과시간})_{kr} \\ &+ \beta_3 \cdot \log(\text{리뷰수})_{kr} + \beta_4 \cdot \text{리뷰평가 평균점수}_{kr} \\ &+ \beta_5 \cdot \text{리뷰극단성}_{kr} + \beta_6 \cdot \log(\text{계제일 이후 경과일})_{kr} \\ &+ \beta_7 \cdot \log(\text{리뷰의 길이})_{kr} + \beta_8 \cdot \log(\text{철자오류})_{kr} \end{aligned}$$

$$\begin{aligned} &+ \beta_9 \cdot \log(\text{소매가격})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{10} \cdot \log(\text{상품 출시일로부터 경과시간})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{11} \cdot \log(\text{리뷰수})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{12} \cdot \text{리뷰평가 평균점수}_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{13} \cdot \text{리뷰극단성}_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{14} \cdot \log(\text{계제일 이후 경과일})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{15} \cdot \log(\text{리뷰의 길이})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{16} \cdot \log(\text{철자오류})_{kr} \cdot \text{상품유형}_{kr} \end{aligned}$$

여기에서  $\mu_k$ 는 상품들에 고정된 속성값으로 관찰되지 않은 이질성(heterogeneity)을 설명하는 값이고  $\varepsilon_k$ 는 에러값이다.

본 논문은 어떤 리뷰의 유형이 사용자들에게 가장 유용한지 확인하려고 한다. 다른 조건이 동일할 때 더 유용한 리뷰들은 합리적인 길이를 가지고, 읽기 쉬우며, 철자와 문법 오류들이 없는 리뷰이고 그리고 읽기 어렵고 오류들이 있는 다른 리뷰들과 비교하였을 때 영향력이 있다. 본 논문은 각각의 리뷰 안의 철자 오류 수를 리뷰의 길이(문자수를 기준으로)로 나눈 정규화된 수를 사용했다. 본 논문은 대문자로 표시된 단어나 숫자가 들어간 단어는 무시하였고 맞춤법 검사 프로그램을 활용하여 측정했다. 리뷰를 읽기 위해 사용자들이 필요한 인지 노력을 측정하기 위해서 리뷰의 글자 길이 수가 측정 되었다.

Log(유용성)<sub>kr</sub>

$$\begin{aligned} &= \alpha + \beta_1 \cdot \log(\text{리뷰수})_{kr} \\ &+ \beta_2 \cdot \text{리뷰작성자 히스토리 매크로}_{kr} \\ &+ \beta_3 \cdot \log(\text{리뷰랭크})_{kr} + \beta_4 \cdot \text{리뷰극단성}_{kr} \\ &+ \beta_5 \cdot \log(\text{리뷰길이})_{kr} + \beta_6 \cdot \log(\text{철자오류})_{kr} \\ &+ \beta_7 \cdot \log(\text{리뷰수})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_8 \cdot \text{리뷰작성자 히스토리 매크로}_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_9 \cdot \log(\text{리뷰작성자랭크})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{10} \cdot \text{리뷰 극단성}_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{11} \cdot \log(\text{리뷰 길이})_{kr} \cdot \text{상품유형}_{kr} \\ &+ \beta_{12} \cdot \log(\text{철자오류})_{kr} \cdot \text{상품유형}_{kr} \end{aligned}$$

$$\begin{aligned}
 & \text{Log}(\text{유용성 투표수})_{kr} \\
 &= \alpha + \beta_1 \cdot \log(\text{리뷰수})_{kr} \\
 &+ \beta_2 \cdot \text{리뷰작성자 히스토리 매크로}_{kr} \\
 &+ \beta_3 \cdot \log(\text{리뷰랭크})_{kr} + \beta_4 \cdot \text{리뷰극단성}_{kr} \\
 &+ \beta_5 \cdot \log(\text{계재일로부터 경과일})_{kr} \\
 &+ \beta_6 \cdot \log(\text{리뷰의 길이})_{kr} + \beta_7 \cdot \log(\text{철자오류})_{kr} \\
 &+ \beta_8 \cdot \log(\text{리뷰의 수})_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_9 \cdot \text{리뷰작성자 히스토리 매크로}_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_{10} \cdot \log(\text{리뷰작성자랭크})_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_{11} \cdot \text{리뷰극단성}_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_{12} \cdot \log(\text{계재일이후 경과일})_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_{13} \cdot \log(\text{리뷰의 길이})_{kr} \cdot \text{상품유형}_{kr} \\
 &+ \beta_{14} \cdot \log(\text{철자오류})_{kr} \cdot \text{상품유형}_{kr}
 \end{aligned}$$

판매량 랭크값의 증가는 낮은 판매량을 나타내기 때문에, 판매량 랭크변수의 음의 계수는 판매량랭크변수가 판매량을 증가시킨다는 것을 나타낸다. 본 논문은 큰 값을 유연하게 하고 잠재적 비선형성을 설명하기 위해서 종속변수와 몇몇의 결정요인들의 로그를 사용하였다.

#### 4. 분석데이터

본 연구목적을 위해서 분석데이터는 Amazon사의 리뷰시스템에서 거래데이터 및 리뷰데이터를 대상으로 2013년 7월에 수집되었고 리뷰데이터는 1998년부터 2013년까지 작성된 리뷰로 구성되었다. <표 1>에 샘플데이터의 연도, 제품군, 제품유형분포가 제시되어 있다. 2000년부터는 데이터의 분포가 거의 균일하다. 샘플에서 서로 다른 제품유형의 수는 1,834개이다. Amazon 사이트에서 각 제품카테고리에서 상위 10개의 베스트셀러 제품을 중심으로 제품을 선정했다. 최소한 두 개 이상의 리뷰 및 상품평가를 한 리뷰작성자의 수는 13320명이다. 최종 샘플데이터는 총 28699개의 리뷰를 포함한다.

<표 1> 본 논문의 샘플

(a) 연도분포

연도	리뷰의수
1998	182
1999	583
2000	2414
2001	1817
2002	1838
2003	1850
2004	2381
2005	2737
2006	2242
2007	2190
2008	1953
2009	2499
2010	2083
2011	1701
2012	1122
2013	1106

(b) 제품군 및 유형분류(\* : 해당사항없음)

제품군	리뷰의수	제품유형
유아용품	4	*
서적	418	탐색재
소비자전자제품	3	*
DVD	9186	경험재
가정용품	7	탐색재
기타 가정용품	7	탐색재
주방용품	2	탐색재
음반	17423	탐색재
PC 악세서리	77	탐색재
애완용품	6	*
메모리카드	19	*
소프트웨어	3	*
스포츠	135	*
장난감용품	388	경험재
비디오	935	경험재
비디오게임	43	경험재
무선제품	43	탐색재
합계	28699	

<표 2>는 이 연구에서 사용되어진 기술통계학을 나타낸다. <표 3>은 탐색재에 대한 리뷰가 28700개의 전체 리뷰 샘플중에서 62.7%(경험재, 36.8%)을 차지한다는 것을 보여준다.

〈표 2〉 결정요인과 종속변수의 기술통계학

	최소	최대	평균	표준편차
상품소매가격	0.01	265	11.97	9.98
상품출시일로부터 경과시간	0	61905	2029.54	1682.39
리뷰의 수	0	1830	44.48	65.00
리뷰평가 평균점수	1	155	4.65	4.92
리뷰작성자 히스토리 매크로	0	19	0.68	0.53
리뷰작성자 랭크	0	15891617	4333650.80	5837966.86
리뷰 극단성	0	15846951	20219.53	427707.53
게재일로부터 경과일	1	11794153	5003.19	142871.41
리뷰 길이	1	4565	115.29	117.67
리뷰길이를 나눈 각 리뷰의 철자오류의 수	0	76	2.56	2.35
판매량 랭크	2.36	2199944	29245.84	53783.51
유용성	0	1	0.52	0.44
유용성투표	0	234576	2372.31	6506.94

〈표 3〉 상품유형의 배분

	빈도수	퍼센트비율
탐색재	17,996	62.7
경험재	10,552	36.8
무소속	152	0.5
총	28,700	100.0

〈표 4〉 Log(판매량 랭크)를 종속 변수로 한 회귀분석 결과

변수	표준화계수	표준오차	t-값
Log(소매가격)	.031	.019	4.880***
Log(상품출시로부터 경과시간)	.057	.015	6.390***
Log(리뷰수)	-.599	.016	-70.955***
리뷰평가 평균점수	-.135	.013	-17.814***
리뷰극단성	.006	.006	1.093
Log(게재일 이후 경과일)	.200	.017	20.342***
Log(리뷰길이)	.044	.017	5.329***
Log(철자오류)	-.015	.021	-1.861**
Log(소매가격)×상품유형	.582	.029	29.867***
Log(상품출시로부터 경과시간)×상품유형	-.069	.020	-1.658**
Log(리뷰 수)×상품유형	-.076	.019	-3.869***
리뷰평가 평균점수×상품유형	.188	.018	3.609***
리뷰극단성×상품유형	-.532	.071	-5.851***
Log(게재일로부터 경과일)×상품유형	-.396	.022	-8.459***
Log(리뷰길이)×상품유형	-.031	.022	-1.077
Log(철자오류)×상품유형	.001	.027	.097
F 값 = 1594.19***			
결정계수 = 0.588			

\* p &lt; 0.1, \*\* p &lt; 0.05, \*\*\* p &lt; 0.01.

## 5. 결 과

<표 4>의 회귀분석 결과는 리뷰들의 수, 리뷰평가가 평균점수 그리고 철자오류들이 상품 판매량에 긍정적으로 영향을 끼친다는 것을 보여준다(부정적으로 판매량 랭크값에 영향을 끼친다). 이것은 리뷰의 수가 많아지고 상품의 평균 리뷰 평가가 높아질수록 상품판매량도 많아진다는 것을 보여준다. 리뷰의 수가 많아질수록 철자오류도 많아지므로 철자오류는 리뷰들의 수와 연관성이 있다고 가정할 수 있고 따라서 판매량과 긍정적으로 관계가 있을 수 있다. 소매가격, 상품출시로부터 경과시간, 게재일 이후의 경과일, 리뷰길이는 부정적으로 상품 판매량과 관계되었고 이는 낮은 가격, 최근 출시된 상품 또는 최근 게재된 리뷰, 짧은 리뷰가 잘 팔리는 상품들의 특징임을 나타낸다. 탐색재의 상품출시로부터 경과시간, 리뷰 수, 리뷰극단성 그리고 게재일 이후의 경과일은 경험재보다 상품 판매량에 더 많은 영향을 미치는 것으로 나타났다. 소매가격 그리고 리뷰평가평균점수는 탐색재보다

경험재의 상품 판매량에 영향을 더 미치는 것으로 나타났다. 이것은 상품판매량에서 탐색재가 상품 출시로부터 더 많이 경과된 시간, 더 많은 리뷰들 그리고 더 극단적인 리뷰 그리고 게재일로부터 더 긴 시간을 요구한다는 것을 보여준다. 경험재의 경우 더 비싼 가격과 더 높은 리뷰평가가 평균점수의 상품이 많이 팔린다.

<표 5>의 회귀분석 결과는 리뷰의 수와 리뷰극단성이 유용성에 부정적으로 영향을 끼치는 반면에 리뷰작성자 히스토리 매크로와 리뷰작성자의 랭크가 긍정적으로 유용성에 영향을 끼친다는 것을 나타낸다. 리뷰작성자 랭크의 긍정적인 영향은 유용성의 정의에서 그 이유를 이해할 수 있다. 즉, 유용성은 구체적인 리뷰를 위한 전체 투표수 중에 유용한 투표의 비율이므로 만약 리뷰들의 전체 투표수가 많지 않다면 랭크값과 유용성 사이의 관계는 항상 부정적일 필요가 없는 것이다. 특정 리뷰 작성자에 대해서 전체 투표수를 모든 과거 유용하게 평가한 투표수로 나눈 리뷰작성자 히스토리 매크로값은 유용성에 긍정의 영향을 가진다. 전체

〈표 5〉 Log(유용성)를 종속변수로 한 회귀분석 결과

변수	표준화계수	표준오차	t-값
Log(리뷰수)	-.026	.007	-2.124**
리뷰작성자 히스토리 매크로	.306	.016	22.173***
Log(리뷰작성자 랭크)	.041	.002	2.918***
리뷰 극단성	-.464	.002	-63.695***
Log(리뷰)길이	-.006	.008	-.516
Log(철자오류)	-.013	.009	-1.150
Log(리뷰 수)×상품유형	-.014	.008	-.515
리뷰작성자 히스토리매크로×상품유형	-.029	.019	-.848
Log(리뷰작성자 랭크)×상품유형	-.052	.002	-1.852**
리뷰극단성×상품유형	-.104	.016	-1.496*
Log(리뷰길이)×상품유형	.181	.010	3.889***
Log(철자오류)×상품유형	.008	.011	.551
F 값 = 679.18*** 결정계수 = 0.385			

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

〈표 6〉 Log(유용성투표)를 종속 변수로한 회귀분석 결과

변수	표준화계수	표준오차	t-값
Log(리뷰수)	-.099	.015	-7.156***
리뷰작성자 히스토리 매크로	.154	.037	9.894***
Log(리뷰작성자 랭크)	-.073	.004	-4.515***
리뷰 극단성	.142	.005	17.372***
Log(계재일로부터 경과일)	.319	.020	24.003***
Log(리뷰길이)	.202	.018	14.403***
Log(철자오류)	.001	.020	.040
Log(리뷰수)×상품유형	-.098	.018	-3.236***
리뷰작성자 히스토리매크로×상품유형	-.103	.044	-2.677***
Log(리뷰작성자 랭크)×상품유형	-.149	.005	-4.599***
리뷰극단성×상품유형	.265	.052	2.348**
Log(상품출시로부터 경과시간)×상품유형	-.008	.025	-.092
Log(리뷰길이)×상품유형	.072	.023	1.386
Log(철자오류)×상품유형	-.007	.025	-.458
F값 = 270.2			
결정계수 = 0.225			

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

리뷰수 그리고 리뷰극단성이 커질수록 리뷰의 유용성은 작아진다. 이것은 극도의 평가 경향과 리뷰의 수와 같은 상품 특징이 유용성에 실제로 부정적인 영향을 가진다는 것을 나타낸다.

리뷰의 길이는 경험제보다는 탐색제에 더 큰 영향을 끼친다. 경험제에는 리뷰작성자의 랭크값과 리뷰 극단성이 탐색제보다 유용성에 더 큰 영향을 끼친다. 이것은 극단적인 리뷰가 경험제의 유용성에 결정적인 반면에 더 긴 리뷰와 리뷰순위(랭크가 아님)가 탐색제의 유용성에 중요한 요소인 것을 보여준다.

〈표 6〉의 결과는 리뷰작성자 히스토리 매크로, 리뷰 극단성, 계재일 이후 경과일 그리고 리뷰의 길이가 리뷰의 유용성 투표수에 긍정적인 영향을 끼친다는 것을 보여준다. 보다 유용한 리뷰작성자, 보다 극도의 리뷰평가, 계재일이 보다 오래된 리뷰들 그리고 보다 긴 리뷰들은 전체 유용한 투표수를 증가시킴을 나타낸다. 상품 리뷰의 수와 작성자 랭크는 유용성 투표수에 부정적으로 관계된

다. 예상된 것처럼, 리뷰 작성자들의 유용성을 부정적으로 나타내는 리뷰작성자의 더 높은 랭크값은 더 유용성이 낮은 리뷰를 결과로 초래한다. 전체 리뷰의 수의 부정적인 영향은 상품의 전체 리뷰 수에 비해 유용성 투표수가 적음을 나타낸다. 즉 유용성 투표수가 전체 투표수에 비해서 적은 부분을 차지함을 나타낸다.

리뷰 극단성은 경험제보다는 탐색제의 유용성 투표수에 더 크게 관련되어 있다. 리뷰의 수, 리뷰작성자 히스토리 매크로, 그리고 리뷰작성자랭크는 탐색제보다 경험제의 유용성 투표수에 더 큰 영향을 끼친다.

이것은 리뷰의 수와 작성자의 평균유용성이 경험제의 유용성투표에 결정적인 영향을 미치는 반면에 극단의 리뷰와 리뷰작성자의 순위가 탐색제의 유용성 투표수에 중요한 요소라는 것을 보여준다. 〈표 7〉은 상품판매량, 유용성 그리고 유용성 투표수에 영향을 미치는 중요한 결정요소를 요약해서 나타낸다.

〈표 7〉 서비스 유형의 중요한 결정요인

종속변수	결정요인	탐색제	경험제
판매량 <sup>a</sup>	소매가격* 상품출시이후 경과시간* 리뷰 수 리뷰평가 평균점수 계재일이후 경과일* 리뷰길이* 철자오류	상품출시이후 경과시간 리뷰 수 리뷰극단성 계재일이후 경과일	소매가격 리뷰평가 평균점수
유용성	리뷰 수* 리뷰작성자 히스토리 매크로 리뷰작성자 랭크 리뷰극단성*	리뷰길이	리뷰극단성 리뷰작성자 랭크
유용성투표수	리뷰 수* 리뷰작성자 히스토리 매크로 리뷰작성자 랭크* 리뷰극단성 계재일 이후 경과일 리뷰길이	리뷰극단성	리뷰 수 리뷰작성자 히스토리 매크로 리뷰작성자 랭크

<sup>a</sup>상품판매량 랭크값과 부정적으로 연관.

\*종속변수와 부정적으로 연관.

## 6. 시사점과 결론

본 논문에서 사용되어진 결정변수들의 유용성에 미치는 영향은 Baek et al.[3] 및 Pan and Zhang [20] 등의 유용성의 결정변수에 관한 기존의 연구 결과를 지지하거나 확장했다고 볼 수 있다. 본 논문은 각 결정변수에 대한 이론적 근거를 검증하는 가설검증논문이라기 보다는 각 결정변수의 직접적인 영향 및 상품타입을 도입함으로써 각 결정변수가 어떻게 종속변수에 미치는 영향이 달라지는지를 제시하는 귀납적 분석을 위한 연구이다. 따라서 판매량, 유용성, 유용성 투표수의 개별 결정요인에 대한 이론적 근거를 토대로 한 개별변수들의 가설검증에 대한 부분은 추후의 연구에서 다루어질 수 있을 것이다.

본 논문의 결과는 리뷰들을 상품제조자들에게 보여주기 위한 순위 제도를 제안하기 위해 사용될 수 있다. 판매량과 매우 관계있는 리뷰들은 제조자들에게 처음으로 보여지는 리뷰들로 구성된다. 그 리뷰들은 상품에 관심 있는 소비자들의 주의를 끄는 것이 수월하다. 제조자는 마케팅 전략을 세

우고 상품의 미래 버전들을 발달시키는데 리뷰들을 참고할 수 있다. 온라인쇼핑에 대한 태도와 긍정적으로 관계가 있기 때문에 유용성과 유용성 투표수를 정확하게 측정하는 것은 중요하다.

본 연구의 결과들은 실무자들이 온라인 리뷰 사이트의 중요한 특성들을 정의하는데 기초들을 제공한다. 마케터매니저들은 유용한 리뷰들을 제공하고 이 연구에서 중대하게 제안되는 특징들을 가진 리뷰 사이트들에 집중해야만 한다. 매니저들은 탐색제와 경험제를 위한 온라인 사이트의 설계에서 소비자들의 상품판매량, 리뷰의 유용성과 유용성투표수의 결정요인에 집중할 수 있다. 사용자 리뷰 시스템의 설계는 어떠한 결정요인이 유용성 투표를 하는 것을 유도하는지에 대한 이해에 기초해서 유용성과 유용성 투표를 증가시키는 방향으로 진행될 수 있다.

## 참고 문헌

- [1] Aciar, S., D. Zhang, S. Simoff, and J. Debenham, "Informed recommender : Ba-

- sing recommendations on consumer product reviews”, *IEEE Intelligent Systems*, Vol.22(2007), pp.39-47.
- [2] Aladag, C. H., E. Egrioglu, and S. Gunay, “A new architecture selection strategy in solving seasonal autoregressive time series by artificial neural networks”, *Hacettepe Journal of Mathematics and Statistics*, Vol. 37, No.2(2008), pp.185-200.
- [3] Baek, H., J. Ahn, and Y. Choi, “Helpfulness of online consumer reviews : readers’ objectives and review cues”, *International Journal of Electronic Commerce*, Vol.17, No.2(2012~2013), pp.99-126.
- [4] Cao, Q., W. Duan, and Q. Gan, “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.
- [5] Chevalier, J. A. and D. Mayzlin, “The effect of word of mouth on sales : online book reviews”, *Journal of Marketing Research*, Vol.43, No.3(2006), pp.345-354.
- [6] Chevalier, J. A. and A. Goolsbee, “Measuring prices and price competition online : Amazon.com and BarnesandNoble.com”, *Quantitative Marketing and Economics*, Vol.1, No.2(2003), pp.203-222.
- [7] Clemons, E., G. Gao, and L. Hitt, “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.
- [8] Chung, W. and T.-L. Tseng, “Discovering business intelligence from online product reviews : A rule-induction framework”, *Expert Systems with Applications*, Vol.39, No.15(1, 2012), pp.11870-11879.
- [9] Duan, W., B. Gu, and A. B. Whinston, “The dynamics of online word-of-mouth and product sales : an empirical investigation of the movie industry”, *Journal of Retailing*, Vol.84, No.2(2008), pp.233-242.
- [10] Forman, C., A. Ghose, and B. Wiesenfeld, “Examining the relationship between reviews and sales : the role of reviewer identity disclosure in electronic markets”, *Information Systems Research*, Vol.19, No. 3(2008), pp.291-313.
- [11] Ghose, A. and P. G. Ipeirotis, “Estimating the helpfulness and economic impact of product reviews : Mining text and reviewer characteristics”, *IEEE Transactions on Knowledge and Data Engineering*, Vol.23, No.10(2011), pp.1498-1512.
- [12] Ghose A. and A. Sundararajan, “Evaluating pricing strategy using e-commerce data : Evidence and estimation challenges,” *Statistical Science*, Vol.21, No.2(2006), pp. 131-142.
- [13] Huang, P., N. H. Lurie, and S. Mitra, “Searching for experience on the Web : An empirical examination of consumer behavior for search and experience goods”, *Journal of Marketing*, Vol.73, No.2(2009), pp.55-69.
- [14] Jiang, Z. and I. Benbasat, “Investigating the influence of the functional mechanisms of online product presentations”, *Information Systems Research*, Vol.18, No.4(2007), pp.221-244.
- [15] Khashei, M. and M. Bijari, “An artificial neural network(p, d, q) model for time series forecasting”, *Expert Systems with Applications*, Vol.37, No.1(2010), pp.479-489.

- [16] Leahy, P., G. Kiely, and G. Corcoran, "Structural optimisation and input selection of an artificial neural network for river level prediction", *Journal of Hydrology*, Vol.355 (2008), pp.192-201.
- [17] Liu, Y., "Word of mouth for movies : Its dynamics and impact on box office revenue", *Journal of Marketing*, Vol.70, No. 3(2006), pp.74-89.
- [18] Min, H.-J. and J. C. Park, "Identifying helpful reviews based on customer's mentions about experiences", *Expert Systems with Applications*, Vol.39(2012), pp.11830-11838.
- [19] Mudambi, S. M. and D. Schuff, "What makes a helpful online review? A study of customer reviews on Amazon.com", *MIS Quarterly*, Vol.34, No.1(2010), pp.185-200.
- [20] Pan, Y. and, J. Q. Zhang, "Born Unequal : A Study of the Helpfulness of User-Generated Product Reviews", *Journal of Retailing*, Vol.87, No.4(2011), pp.598-612.
- [21] Park, D. H., J. Lee, and I. Han, "The effect of on-line consumer reviews on consumer purchasing intention : The moderating role of involvement", *International Journal of Electronic Commerce*, Vol.11, No.4(2007), pp.125-148.
- [22] Pavlou, P., H. Liang, and Y. Xue, "Uncertainty and mitigating uncertainty in on-line exchange relationships : A principal-agent perspective", *MIS Quarterly*, Vol. 31, No.1(2007), pp.105-131.
- [23] Reinstein, D. and C. M. Snyder, "The influence of expert reviews on consumer demand for experience goods : a case study of movie critics", *Journal of Industrial Economics*, Vol.53, No.1(2005), pp. 27-51.
- [24] Weathers, D., S. Sharma, and S. L. Wood, "Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods", *Journal of Retailing*, Vol.83, No.4(2007), pp.393-401.
- [25] West, D., S. Dellana, and J. Qian, "Neural network ensemble strategies for financial decision applications", *Computers and Operations Research*, Vol.32(2005), pp.2543-2559.
- [26] Yoon, Y., T. Guimaraes, and G. Swales, "Integrating artificial neural networks with rule-based expert systems", *Decision Support Systems*, Vol.11(1994), pp.497-507.
- [27] Yuen, C. W. M., W. K. Wong, S. Q. Qian, L. K. Chan, and E. H. K. Fung, "A hybrid model using genetic algorithm and neural network for classifying garment defects", *Expert Systems with Applications*, Vol.36, No.2(2009), pp.2037-204.
- [28] Zhang, R. and T. Tran, "Helpful or unhelpful : a linear approach for ranking product reviews", *Journal of Electronic Commerce Research*, Vol.11, No.3(2010), pp. 220-230.
- [29] Zhang, Z., "Weighing stars : Aggregating online product reviews for intelligent e-commerce applications", *IEEE Intelligent Systems*, Vol.23(2008), pp.42-49.

## ◆ 저 자 소 개 ◆

**이 상 재 (sangjae@sejong.ac.kr)**

KAIST에서 박사학위(경영정보전공)를 받고 현재는 세종대학교 경영대학 경영정보 전공 교수로 재직 중이다. 주요 연구분야는 전자상거래전략, 데이터마케팅, 빅데이터이다. Decision Support Systems, European Journal of Information Systems, Telecommunication Systems, Information Processing and Management, Technological Forecasting and Social Change, Information and Management 등 다수의 국외 유명저널에 논문을 게재하고 있다. AIS(Association for Information Systems)가 2009년도에 아시아 태평양 지역에서 한국 교수 중 Top Journals 게재기준 연구 최우수교수로 선정하였다.

**최 준 연 (zoon@sejong.ac.kr)**

KAIST 전기 및 전자공학과에서 학사, 테크노경영대학원 경영공학과에서 석사 및 박사학위를 취득하였으며, 현재 세종대학교 디지털콘텐츠학과 조교수로 재직 중이다. 주요 관심분야는 지능형시스템, 소셜 데이터 마이닝, 추천시스템 등이다. Expert Systems, AI Communications, International Journal of Computational Intelligence Systems, Expert Systems with Applications 등의 학술지에 논문을 게재하였다.

**최 진 호 (jhchoi@sejong.ac.kr)**

KAIST 산업경영학과에서 학사, 테크노경영대학원 경영공학과에서 석사 및 박사학위를 취득하였으며, 현재 세종대학교 경영학과 부교수로 재직하고 있다. 주요 관심분야는 플랫폼비즈니스, 지식네트워크, 경제성분석 등이다.