

온라인 상품평의 내용적 특성이 소비자의 인지된 유용성에 미치는 영향*

박윤주

서울과학기술대학교 경영학과
(yjpark@seoultech.ac.kr)

김경재

동국대학교 서울 경영대학 경영정보학과
(kjkim@dongguk.edu)

인터넷 상거래에서, 소비자들은 기존에 제품을 구매한 다른 사용자들이 작성한 상품평에 많은 영향을 받는다. 그러나, 상품평이 점차 축적되어감에 따라, 소비자들이 방대한 상품평을 일일이 확인하는데 많은 시간과 노력이 소요되고, 또한 무성의하게 작성된 상품평들은 오히려 소비자들의 불편을 초래하기도 한다. 이에, 본 연구는 온라인 상품평의 유용성에 영향을 미치는 요인들을 분석하여, 소비자들에게 실제로 도움이 될 수 있는 상품평을 선별적으로 제공하는 예측모형을 도출하는 것을 목적으로 한다. 이를 위해, 텍스트마이닝 기법을 사용하여, 상품평에 포함되어있는 다양한 언어적, 심리적, 지각적 요소들을 추출하였으며, 이러한 요소들 중에서 상품평의 유용성에 영향을 미치는 결정요인이 무엇인지 파악하였다. 특히, 경험재인 의류군과 탐색재인 전자제품군에 대한 상품평의 특성 및 유용성 결정요인이 상이할 수 있음을 고려하여, 제품군별로 상품평의 특성을 비교하고, 각각의 결정요인을 도출하였다. 본 연구에는 아마존닷컴(Amazon.com)의 의류군 상품평 7,498건과 전자제품군 상품평 106,962건이 사용되었다. 또한, 언어분석 소프트웨어인 LIWC(Linguistic Inquiry and Word Count)를 활용하여 상품평에 포함된 특징들을 추출하였고, 이후, 데이터마이닝 소프트웨어인 RapidMiner를 사용하여, 회귀분석을 통한, 결정요인 분석을 수행하였다. 본 연구결과, 제품에 대한 리뷰어의 평가가 높고, 상품평에 포함된 전체 단어 수가 많으며, 상품평의 내용에 지각적 과정이 많이 포함되어 있는 반면, 부정적 감정은 적게 포함된 상품평들이 두 제품 모두에서 유용하다고 인식되는 것을 알 수 있었다. 그 외, 의류군의 경우, 비교급 표현이 많고, 전문성 지수는 낮으며, 한 문장에 포함된 단어 수가 적은 간결한 상품평이 유용하다고 인식되고 있었으며, 전자제품의 경우, 전문성 지수가 높고, 분석적이며, 진솔한 표현이 많고, 인지적 과정과 긍정적 감정(PosEмо)이 많이 포함된 상품평이 유용하게 인식되고 있었다. 이러한 연구결과는 향후, 소비자들이 효과적으로 유용한 상품평들을 확인하는데 도움이 될 것으로 기대된다.

주제어 : 상품평, 유용성, 텍스트마이닝, LIWC, 예측모형

논문접수일 : 2017년 6월 22일 논문수정일 : 2017년 9월 14일 게재확정일 : 2017년 9월 19일
원고유형 : 일반논문 교신저자 : 김경재

1. 서론

제품을 직접 보지 못하고 구매하는 인터넷 상거래에서, 기존에 제품을 경험해 본 다른 사용자

들의 후기인 상품평(review)은 소비자들에게 유용한 정보로 인식된다. 인터넷 진흥원의 조사결과에 따르면, 소비자가 구매결정 시 다른 사용자의 상품평에 영향을 받는 비율은 무려 74%에 이

* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2015S1A5A2A03047963)

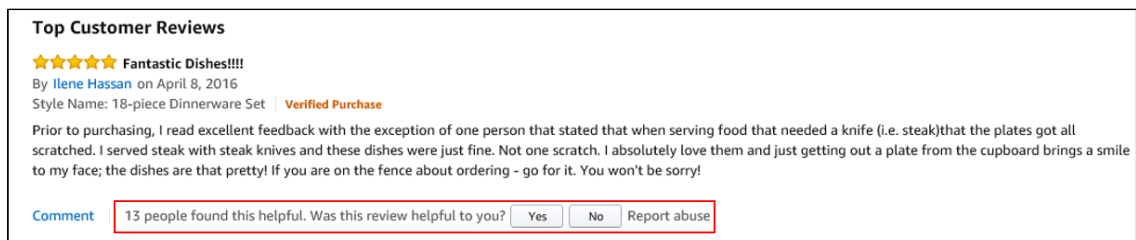
른다고 한다[Lee et al., 2010]. 이러한 상품평에는 크게 다음의 두 가지 형태가 있다. 첫째, 상품평 제공자(소위, 리뷰어)가 자신의 SNS(Social Network Services) 매체에 특정 상품에 대해서 구체적이고, 비교적 전문적인 평가를 제공하는 것이며, 둘째, 구매자가 상품을 구입한 웹사이트에 올린 댓글 형태의 구매후기이다. 인터넷 유통업체에서 제공하고 있는 대부분의 상품평은 이러한 형태로 볼 수 있다.

그러나, 이러한 두 번째 형태의 상품평의 경우, 무성의하게 작성된 상품평들로 인하여, 소비자들이 오히려 불편을 겪기도 하며, 또한 상품평이 점차 축적되어감에 따라서, 소비자들이 방대한 상품평을 일일이 확인하는 데에도 많은 시간과 노력이 소요된다. 이에, 고객들에게 유용한 상품평을 선별하여 제공하려는 노력이 학계 및 산업계에서 이뤄져 왔다. 예를 들어, 세계 최대의 인터넷 기업인 아마존닷컴(Amazon.com)은 고객들이 작성한 상품평들에 대해서, 다른 고객들이 유용성 여부를 평가하도록 하고 있다. 즉, <Figure 1>과 같이 “Was this review helpful to you”라는 질문을 통해서, 상품평을 읽은 다른 고객들이 해당 상품평의 유용성을 평가하도록 한 후, 여러 사용자들로부터 유용하다고 평가 받은 상품평 들을 화면의 상단에 노출시키는 것이다.

본 연구는, 이러한 두 번째 형태의 상품평을

대상으로, 온라인 상품평의 유용성에 영향을 미치는 요인들을 분석하여, 소비자들에게 실제로 도움이 될 수 있는 상품평을 선별하여 제공할 수 있는 예측모형을 도출하는 것을 목적으로 한다. 특히, 상품평에 포함된 언어적, 내용적 특징을 분석하여, 소비자들에게 유용하게 인식되는 상품평의 특성을 파악하였다. 즉, 상품평에 사용된 비교급 표현, 논리성, 전문성, 진정성 등을 비롯하여, 인지적 과정이나 지각적 과정, 긍/부정의 감정이 나타난 수준 등이 상품평의 유용성에 미치는 영향을 분석하였다. 이때, 제품을 직접 구매하여 사용해 보기 전에는 특성을 파악하기 어려운, 소위 ‘경험재(experience goods)’에 대한 상품평 유용성의 결정요인과, 구매 전 제품 속성을 쉽게 파악할 수 있는 ‘탐색재(search goods)’의 결정요인이 상이할 수 있음을 고려하여, 경험재의 사례인 의류제품군과 탐색재의 사례인 전자제품군을 각각 비교하여 분석하였다.

본 연구에는, 세계 최대의 인터넷 쇼핑몰인 Amazon.com에서 판매되고 있는 의류 제품군의 상품평 7,498개와 전자제품군 상품평 10,692개 데이터가 사용되었다. 이는, 앞서 <Figure 1>에서 제시한 바와 같이, Amazon의 데이터가 상품평의 유용성에 대한 소비자들의 평가 결과를 제공하고 있어, 본 연구에서 제안한 상품평 유용성 예측 모델의 성능을 평가하는데 효과적이기 때



〈Figure 1〉 An example of feedbacks of the review in Amazon.com

문이다. 이후, 언어분석 소프트웨어인 LIWC (Linguistic Inquiry and Word Count)를 사용하여, 상품평에 포함된 특징을 추출하여 설명변수로 활용하였고, 데이터마이닝 툴인 RapidMiner를 사용하여 회귀모형을 수립한 후, 모형의 설명력 및 예측 정확성을 평가하였다.

본 연구의 나머지 부분은 다음과 같이 구성된다. 우선, 2장에서는 기존의 관련연구를 조사하고, 3장에서는 연구방법을 소개하였다. 4장에서는 본 연구의 분석결과를 제시하고, 마지막으로 5장에서는 본 연구의 결론 및 한계점을 기술하였다.

2. 관련연구

기존에 상품평의 유용성에 영향을 미치는 요인을 분석한 연구들은 다수 존재한다. Mudambi와 Schuff(2010)은, 아마존닷컴(Amazon.com)의 상품평 1,587개를 분석하여, 리뷰어가 제품에 대한 부여한 평점(rating)과 상품평에 포함된 단어수가 상품평의 유용성에 미치는 영향을 분석하였다. 이들은 제품타입을 조절변인으로 설정하여, 탐색재(search goods)와 경험재(experience goods)를 비교한 연구를 수행하였다. 이러한 연구 결과, 상품평에 포함된 단어수가 증가하면, 두 제품타입 모두에서, 소비자가 느끼는 상품평의 유용성이 증가하지만, 탐색재가 더 큰 영향을 받는다고 주장하였다. 또한, 경험재의 경우, 극성이 큰(즉, 리뷰어가 제품에 대해서 극단적으로 긍정적이거나 부정적인 평가를 내린) 상품평 보다는 중도적인 상품평이 더욱 유용한 결과를 도출하였다.

Ghose와 Ikperiotis(2011)도 아마존닷컴의 상

품평을 분석하였으나, 상품평에 포함된 단어수가 증가하면, 일정수준까지는 소비자가 느끼는 상품평의 유용성이 증가하지만, 이후로는 오히려 유용성이 감소한다는 상이한 결과를 도출하였다. Forman, et al.(2008)은 소비자가 리뷰 평점이 부정적 혹은 긍정적인 방향으로 극단성을 보이는 상품평을 더욱 유용하게 느낀다고 주장하였으며, Chevalier, et al.(2006)은 아마존과 반즈 앤노블의 웹사이트에 올라와 있는 온라인 서적 리뷰를 분석하여, 부정적인 서평이 긍정적인 서평에 비해 훨씬 영향력이 크다는 것을 실증적으로 밝혀내었다.

그 외에도, Mahony와 Smyth는 TripAdvisor라는 호텔 예약 사이트의 상품평 및 관련정보를 분석하여, 유용한 상품평을 선별하여 추천하는 시스템을 제안하였다. 이들은, 리뷰어의 명성(reputation), 소셜(social), 감성(sentiment), 상품평의 내용(content) 등과 관련된 속성들을 JRip, J48, 그리고 Naive Bayes의 데이터마이닝 기법으로 분석하여 상품평 추천모형을 수립하였으며, 이 중 가장 JRip을 추천 성능이 가장 우수하다는 결과를 도출하였다.

국내연구로는 Cho & Yi(2014)가 아마존닷컴의 베스트셀러 2,600개의 상품을 대상으로 연구를 수행하여, 리뷰어의 명성, 상품평에 대한 평점의 극단성, 상품평의 길이와 객관성 등이 상품평의 유용성에 영향을 미친다는 결론을 도출하였고, 이와 유사하게, Baek et al(2012)도 아마존닷컴의 데이터를 사용하여 리뷰어의 명성, 평점, 상품평에 포함된 단어 수 및 부정어 비율 등을 결정요인으로 제시하였으나, 이들은 고가격제품과 저가격 제품의 결정요인이 다르다고 주장하였다. Hong & Yi(2015)는 상품평의 구성요소를, 상품평 자체에 포함된 정보인 ‘중심정보(central

cues)’와 외부적으로 추가된 ‘주변정보(peripheral cues)’로 구분하고, 유용하다고 평가받은 상위 50%의 상품평과 하위 50%의 상품평을 구분하여 연구를 수행하였다. 그 결과, 중심정보는 하위 50%의 상품평 집단의 유용성에 영향을 미치는 반면, 주변정보는 상위 50% 집단의 유용성의 결정요인으로 작용한다는 결과를 도출하였다.

이 외에도 상품평의 유용성에 대한 영향요인을 분석한 기존 연구들이 다수 존재한다 [Pan and Zhang(2010), Wan(2015), Baek et al.(2012), Korfiatis et al.(2012), Yin et al.(2014), Lee & Kwak(2013), Lee et al.(2014), Lee & Lee(2016), Kim & Song(2016), Choi & Lee(2011)].

3. 연구방법

3.1 데이터 및 연구변수

본 연구에는 세계 최대의 인터넷 쇼핑몰인 Amazon.com에서 판매되고 있는 의류 제품군의

상품평 7,498개와 전자제품군 상품평 10,692개 데이터가 사용되었다. 데이터는 <http://jmcauley.ucsd.edu/data/amazon/> 사이트로부터 수집되었으며, 의류군 데이터에는 의류, 신발 및 보석에 대한 상품평이 포함되어 있고, 전자제품군에는 카메라, 오디오, 메모리 스틱 등을 비롯한 다양한 제품들이 포함되어 있다. 원본 데이터는, <Figure 2>의 사례와 같이, 제품 ID(Product ID), 리뷰 작성자 ID(Reviewer ID), 상품평(Review Text), 제품에 대한 리뷰어의 평가 (Rating), 상품평 작성일(Review Time), 해당 상품평이 구매에 도움이 됐다고 표기한 사용자 수(Helpful#), 상품평을 평가한 전체 사용자 수(Total #)등으로 구성된다 [McAuley et al. (2015)]. 이때, 상품평을 평가한 전체 사용자 수인 Total #가 10미만인 데이터의 경우, 유용성 평가결과를 신뢰하기에 부족하다고 판단하여, 제거하고 사용하였다.

다음으로, 위의 원본 데이터를 변환하여, 새로운 연구변수를 구성하였다. 특히, 심리학 분야에서 널리 활용되는 언어분석 소프트웨어인 LIWC(Linguistic Inquiry and Word Count)를 사용

- Product ID: A2ENZ4FESUXXMT
- Reviewer ID: 1400501466
- Review Text: I was looking at expensive tablets that were more like mini notebook computers. I already have a high end notebook. I wanted something that was very portable and not combersome to take on trips, go to coffee shops and etc. I wanted the ability to get email, do limited surfing and read books. The Nook works flawlessly and the display is really nice. I have an N protocol router and the Nook is quick on the Net. I read some negative reviews here. They appear to be written by folks who want to take a \$200 unit and turn it into a \$500 unit with various apps and other applications. Here's a news flash for the naysayers. Go out and buy the \$500 unit and quit complaining. If you want to read books, surf and get email, you'll like this unit.
- Rating: 5.0
- Review Time: 03 12, 2012
- The number of votes on helpfulness: 9
- Total number of votes: 10

<Figure 2> An example of Amazon.com review data

하여, 상품평(ReviewText)의 내용에 포함된 단어 수(WC: Word Count), 상품평 문장의 간결성(WPS: Words Per Sentence), 상품평에 사용된 비교급 표현의 정도(Compare), 상품평 내용의 논리적 사고수준(Alytic: Analytical Thinking), 영향력(Clout), 진정성(Authentic), 인지적 과정(CogProc: Cognitive Process), 지각적 과정(Perceptual Process), 긍정적 감정(Posemo: Positive Emotion) 및 부정적 감정(Negemo: Negative Emotion)의 수준 등의 특성을 추출하여

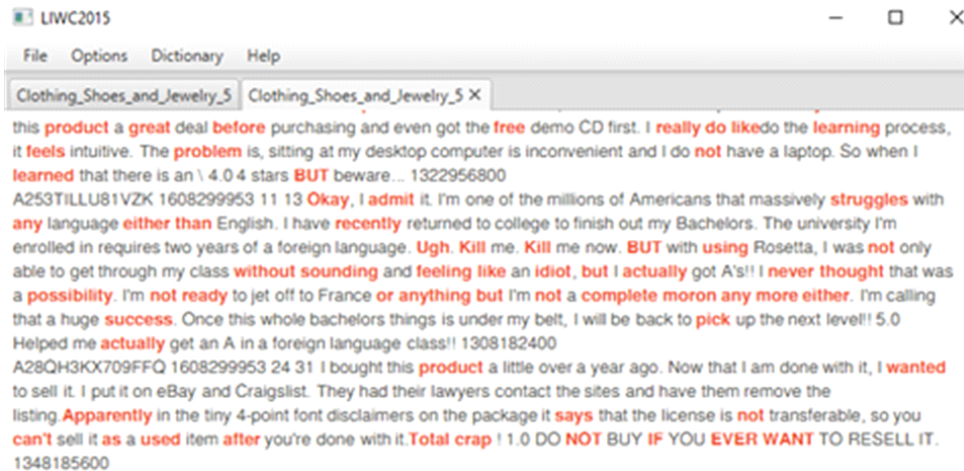
연구변수로 활용하였다. 그 외, 제품에 대한 리뷰어의 평가 값(Rating)은 원본 데이터 값을 그대로 사용하였으며, 종속변수는 상품평이 구매에 도움이 됐다고 표기한 사용자 수(Helpful#)를 상품평을 평가한 전체 사용자 수(Total #)로 나눈 비율(Helpful Ratio)을 활용하였다. <Table 1>은 본 연구에 활용된 연구변수에 대한 설명과 산출 방법을 소개하고 있다. 또한, <Table 2>는, <Table 1>의 산출방법을 통해서, <Figure 2>의 데이터가 변환된 모습을 나타내고 있다.

<Table 1> Explanation of the research variables

Variable	Explanation	Calculation
Rating	Rating score of a product from reviewer	Rating score scale from 1 to 5
WC	Total number of words included in a review text	Word count
WPS	Average number of words in a sentence	Words/Sentence
Compare	Comparison words where the author is comparing one entity with another, includes words like as, bigger, best, smaller, etc.	# of comparison words /total # of comparison words(317) *100
Analytic	A high number reflects formal, logical, and hierarchical thinking. lower numbers reflect more informal, personal, hereandnow, and narrative thinking.	Derived based on previously published findings from Pennebarker et al., 2014. Ranging from 0 to 100.
Clout	A high number suggests the author is speaking from the perspective of high expertise and is confident. Low Clout numbers suggest a more tentative, humble, even anxious style.	Derived based on previously published findings from Kacewicz et al., 2012.
Authentic	Higher numbers are associated with a more honest, personal, and disclosing text. Lower numbers suggest a more guarded, distanced form of discourse.	Derived based on previously published findings from Newman et al., 2003.
CogProc	Ratio of cognitive process words(cause, know, ought etc) in a review text to total 797 cognitive words in LIWC dictionary	(# of related words in a review text/total # of related words) *100
Percept	Ratio of perceptual process words(look, heard, feeling etc) in a review text to total 436 perceptual words in LIWC dictionary	(# of related words in a review text/total # of related words) *100
PosEmo	Ratio of positive emotion words(love, nice, sweet etc) in a review text to total 620 negative emotion words in LIWC dictionary	(# of related words in a review text/total # of related words) *100
NegEmo	Ratio of negative emotion words(hurt, ugly, nasty etc) in a review text to total 744 negative emotion words in LIWC dictionary	(# of related words in a review text/total # of related words) *100
Helpful_Ratio (dependent variable)	Ratio of helpful votes to total number of votes	(Helpful#/Total#) * 100

〈Table 2〉 Preprocessed review data

WC	WPS	Compare	Analytic	Clout	Authentic	CogProc	Percept	PosEmo	NegEmo
142	17.75	2.11	81.89	21.92	24.72	9.86	0	2.82	2.11



〈Figure 3〉 Key words and expression included in a review text

3.2 분석방법

3.1절에서 설명한 바와 같이, 연구변수를 추출하기 위해서 언어분석 툴인 LIWC가 활용되었다. LIWC는 자체적으로 보유한 사전을 기반으로 작동되는데, 우선 <Figure 3>와 같이, 상품평으로부터 사전에 포함된 주요 단어들이 추출된다. 추출된 단어들은, <Figure 4>과 같이, 카테고리별로 분류되며, 이를 기반으로, <Figure 5>와 같은 상품평의 특징 값들이 산출된다.

이렇게 도출된 제품군별 상품평의 특성 차이를 기술통계 값을 비교하여 분석하였으며, T검정법(t-test)으로, 이러한 차이가 통계적으로 유의한지 확인하였다.

마지막으로, 상품평의 유용성에 영향을 미치

는 결정요인을 분석하는 데에는 회귀분석이 사용되었다. 회귀분석은 SPSS등과 같은 통계 분석 도구를 활용하여 수행할 수도 있지만, 이러한 경우 도출된 회귀모형의 예측력을 평가하지 못한다. 따라서, 본 연구에서는 모형의 예측력을 평가할 수 있는 데이터마ining 툴(tool)인 Rapidminer를 사용하여 회귀모형을 수립하였다. 회귀모형은 10-fold cross validation을 사용하여, 각 제품군별로 수립되었으며, 모형의 설명력은 결정계수(R^2)를 사용하여 평가하였고, 예측력 평가에는 MAE(Mean Absolute Error)와 RMSE(Root Mean Squared Error)를 사용하였다. <Figure 6>는 회귀분석에 사용한 Rapidminer프로세스를 나타내고 있다.

Word	compare	affect	posemo	negemo	cogproc	insight	cause	percept	see	hear	feel
researched											
this											
a											
great		X	X								
deal											
before	X										
purchasing											
and											
even											
got											
free		X	X								
demo											
cd											
first											
really					X						
do like		X	X								
do											
learning					X	X					
process											
it											
feels					X	X		X			X
intuitive											
problem		X		X	X						
sitting											

〈Figure 4〉 Categorization of words included in a review text

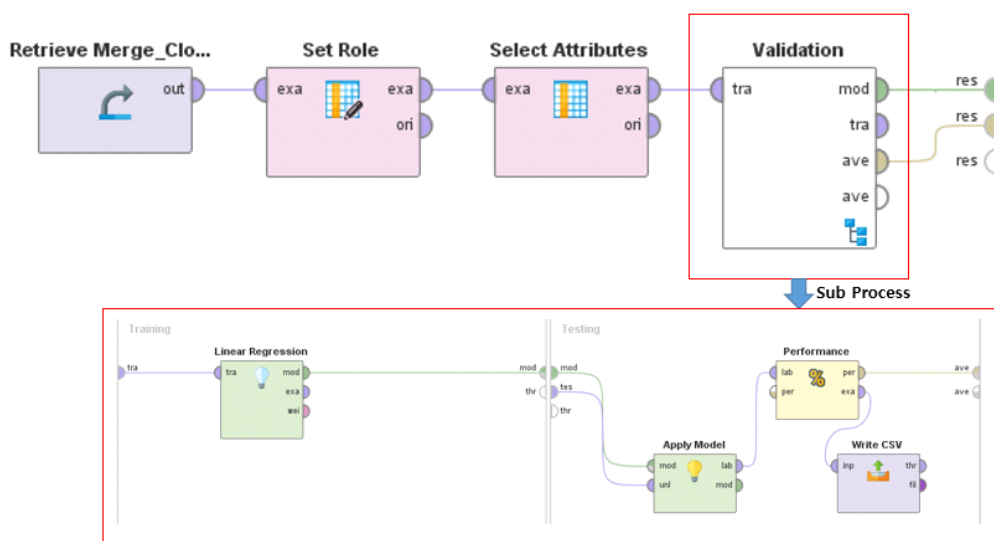
Source (A)	WC	Analytic	Clout	Authentic	Tone	WPS	compare	affect	posemo	negemo	cogproc	insight	cause	percept	see	hear	feel
I did not know that Converse stooped that low to get their products made in Vietnam or ...	28	8.99	7.67	99.00	1.00	28.00	0.00	3.57	0.00	3.57	28.57	3.57	7.14	0.00	0.00	0.00	0.00
this style does not say if it is Sandalfoot or reinforced toe...couldn't buy because of that	17	1.00	1.00	1.00	25.77	17.00	0.00	0.00	0.00	0.00	29.41	0.00	5.88	5.88	0.00	5.88	0.00
** PLEASE NOTE ** I was new to amazon at the time I posted this review, and was angry ...	272	62.14	16.03	64.08	79.41	17.00	0.74	5.88	4.41	1.47	8.46	0.37	2.21	1.84	0.00	0.00	1.84
It is a great watch. Unfortunately I bought two and I received just oneWhere is the other...	26	53.63	22.08	17.46	25.77	8.67	0.00	7.69	3.85	3.85	3.85	0.00	0.00	7.69	7.69	0.00	0.00
The title says it all. I placed my order days ago yet it sits and sits and Amazon has NO NO ...	135	53.54	22.95	92.47	15.37	19.29	0.74	0.74	0.00	0.74	15.56	3.70	2.96	1.48	0.00	1.48	0.00

〈Figure 5〉 Extracting characteristics of review text using LWC

$$MAE = \frac{\sum_{i=1}^n |e_i|}{n} \quad - \text{(equation 1)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad - \text{(equation 2)}$$

(※ n: sample size, e = prediction score - real score)



<Figure 6> RapidMiner process for regression analysis

4. 분석결과

의류군과 전자제품군 상품평의 비교결과는 <Table 3>에 제시한 바와 같다. 전자제품군 상품평에 포함된 단어 수인 WC 값은 216개로, 의류군(122개)에 비해서 매우 높아서, 리뷰어들이 전자제품군의 상품평을 작성할 때, 더 많은 단어를 사용하고 있음을 알 수 있다. 또한, 한 문장에 들어가는 평균 단어 수를 나타내는 WPS 값도 전자제품군이 22개로, 의류의 17개보다 더 높게 도출되어, 전자제품군 상품평 문장이 여러 단어들을 사용하여 길게 구성되어 있음을 알 수 있었다. 상품평의 내용상 특징을 살펴보면, 전자제품군의 상품평이 의류군보다 더 분석적(Analytic)이고, 전문적(Clout)이며, 인지적 과정(CogProc)과 관련된 단어들이 많이 포함되어 있고, 또한 부정적(NegEmo) 감정을 표현하는 단어도 많이 사용되고 있음을 알 수 있었다. 반면, 의류군 상품평

은 전자제품군에 비해서, 더 개인적이고 진솔한(Authentic) 표현들을 포함하고 있고, 긍정적 감정(PosEmo) 및 지각적 과정(Percept)을 나타내는 단어들이 상대적으로 많이 포함됨을 알 수 있었다. 이러한 차이가 통계적으로 유의한지 확인하기 위하여 T검정(t-test)을 수행한 결과, <Table 3>의 p값에서 나타난 바와 같이, 상품평에 사용된 비교급 표현을 제외한 모든 변수에서, 제품군별 상품평의 특성 차이가, 95%신뢰수준에서 통계적으로 유의하다는 결과가 도출되었다. 반면, 비교급(Compare) 표현의 경우, 두 제품군의 상품평에 유의한 차이가 없는 것으로 나타났다.

다음으로, 소비자가 느끼는 상품평의 유용성에 대한 영향요인을 선형회귀분석(Linear Regression)으로 분석한 결과를 설명하겠다. <Table 4>와 <Table 5>는 의류군과 전자제품군 상품평의 회귀분석 결과를 각각 나타내고 있다. 우선, 두 제품군 모두에 대해서 공통적으로 상

〈Table 3〉 Comparing characteristics of clothing review text with electronics

Variables	Average(Variance)		p	Significance
	Clothing	Electronics		
WC	122.846 (10770.77)	216.035 (48249.56)	0	***
Analytic	56.576 (530.74)	69.833 (363.11)	0	***
Clout	33.676 (406.22)	34.762 (323.33)	5.76E-06	***
Authentic	45.120 (837.58)	37.108 (628.70)	5.4E-117	***
WPS	16.945 (92.57)	21.873 (210.68)	0	***
Compare	2.667 (4.47)	2.635 (3.58)	0.207261	
PosEmo	4.965 (9.61)	3.536 (4.99)	0	***
NegEmo	0.978 (2.25)	1.157 (1.68)	1.48E-23	***
CogProc	9.966 (15.86)	10.381 (14.11)	2.75E-18	***
Percept	3.486 (8.16)	3.196 (6.09)	1.2E-17	***

(※ *: $p<0.05$, **: $p<0.01$, *** : $P<0.001$)

상품평의 유용성에 영향을 미치는 요인들로는, 제품에 대한 리뷰어의 평가(Rating), 상품평에 포함된 전체 단어 수(WC), 지각적 과정(Percept), 그리고 부정적 감정(NegEmo)등이 있었다. 즉, 소비자들은 제품에 대한 리뷰어의 평가(Rating)가 높을수록, 상품평이 많은 단어(WC)들로 구성되어 있을수록, 또한, 보고, 듣고, 느끼는 등의 지각적 과정(Percept)이 많이 포함되어 있을수록, 상품평을 유용하다고 느끼는 것으로 나타났다. 반면, 상품평에 부정적 감정(NegEmo)을 나타내는 단어들이 많이 포함되어 있으면, 소비자들은

해당 상품평이 유용하지 못하다고 느끼는 것으로 나타났다. 이러한 결과는, 리뷰어가 제품에 만족하여 높은 평점(Rating)을 부여했을 때, 상품평도 성의 있게 작성하기 때문인 것으로 보이며, 또한 여러 단어(WC)들을 사용하여 작성된 상품평이 더 많은 정보를 담고 있기 때문인 것으로 판단된다. 그 외, 기존에 제품을 구매한 다른 소비자들이 보고, 듣고, 느낀 경험을 기술한 지각적 과정(Percept)은, 제품을 직접 보지 못한 온라인 소비자들에게 매우 중요한 정보로 인식되기 때문에, 유용성의 결정요인으로 도출된 것

〈Table 4〉 Results of regression analysis of review text (Clothing)

Attribute	coefficient	std.error	std.coef.	tolerance	t-test	p-val	
Rating	3.998	0.141	0.343	0.843	28.351	0.000	****
WC	0.008	0.002	0.056	0.998	4.792	0.000	****
Clout	-0.041	0.009	-0.054	0.979	-4.671	0.000	****
Compare	0.350	0.079	0.049	1.000	4.404	0.000	****
NegEmo	-0.421	0.114	-0.042	0.935	-3.706	0.000	****
WPS	-0.041	0.018	-0.026	1.000	-2.276	0.023	**
percept	0.110	0.057	0.021	0.995	1.922	0.055	*
Analytic	0.009	0.008	0.014	1.000	1.190	0.234	
posemo	0.047	0.060	0.010	0.903	0.781	0.435	
cogproc	-0.058	0.046	-0.015	0.995	-1.242	0.214	

〈Table 5〉 Results of regression analysis of review text (Electronics)

Attribute	coefficient	std.error	std.coef.	tolerance	t-test	p-val	
Rating	7.414	0.049	0.453	0.840	152.428	0.000	****
WC	0.011	0.000	0.097	0.976	34.178	0.000	****
Analytic	0.136	0.004	0.107	0.983	37.160	0.000	****
Clout	0.022	0.004	0.016	0.990	5.477	0.000	****
Authentic	0.041	0.003	0.043	0.996	14.213	0.000	****
posemo	0.108	0.031	0.010	0.953	3.486	0.000	****
cogproc	0.093	0.019	0.014	0.984	4.934	0.000	****
negemo	-0.151	0.053	-0.008	0.881	-2.844	0.004	***
percept	0.058	0.026	0.006	0.990	2.198	0.028	**
WPS	0.007	0.005	0.004	0.990	1.639	0.101	

으로 생각된다.

제품군별 결정요인의 차이를 살펴보면, 의류 제품군의 경우, 위의 공통 변인들과 더불어, 비교급(Compare) 표현이 많이 포함된 상품평이 유용하다고 평가받고 있었다. 반면, 전문성(Clout)

지수가 높고, 상품평 한 문장에 포함된 단어 수(WPC)가 많은, 즉 간결하지 못한 문장의, 상품평은 유용하지 못하다고 평가 받고 있음을 알 수 있었다.

전자제품군의 경우, 전문성(Clout) 지수가 높

은 상품평의 유용성이 높게 나타났고, 또한, 분석적(Analytic)이고, 진솔하며(Authentic), 인지적 과정(CogProc)에 대한 표현이 많고, 긍정적 감정(PosEmo)을 나타내는 단어들이 많이 포함된 상품평이 유용하다는 결과가 도출되었다. 이는, 일반적으로 전자제품들이 의류에 비해 이해하기 어렵고 복잡하기 때문에, 소비자들이 보다 전문적이고, 분석적인 상품평들을 선호하기 때문인 것으로 생각된다.

마지막으로, 이러한 회귀모형들에 대한 성능 평가를 수행하였다. 의류의 경우, 회귀모형에서 독립변수들이 종속변수를 어느 정도 설명하는지를 나타내는 결정계수(R^2) 값이 13.1로 도출되어, <Table 4>의 독립변인들이 상품평의 유용성을 13%정도 설명하고 있음을 알 수 있었다. 또한, 이러한 회귀모형을 사용하여 새로운 상품평의 유용성을 예측할 때의 예측오류는, MAE기준 9.216, RMSE기준 14.108을 보여, 본 회귀모형으로 임의의 새로운 상품평에 대한 유용성을 예측할 경우, MAE기준 평균적으로 9.2(100대비)정도 틀릴 수 있음을 알 수 있었다. 유사하게, 전자제품 상품평 회귀모형의 결정계수(R^2)값은 0.252로 도출되어, <Table 5>의 독립변인들이 상품평의 유용성을 25%정도 설명하고 있음을 알 수 있었으며, 이러한 회귀모형을 사용하여 새로운 상품평의 유용성을 예측할 경우, 예측오류 MAE는 14.907, RMSE는 20.903으로 도출되었다.

5. 결론 및 향후과제

본 연구는 온라인 상품평의 유용성에 영향을 미치는 언어적, 내용적 요인들을 분석하여, 소비자들에게 실제로 도움이 될 수 있는 상품평을 선

별하여 제공할 수 있는 예측모형을 도출하였다. 또한, 경험재인 ‘의류제품군’에 대한 상품평 유용성의 영향요인과, 탐색재인 ‘전자제품군’의 영향요인이 상이함을 실증적으로 분석하고, 어떠한 요소들이 각 제품군의 유용성을 결정하는지 파악하였다.

본 연구결과, 소비자들은 리뷰어가 제품에 대해서 높은 평점을 부여하고, 상품평에 포함된 단어 수가 많으며, 상품평의 내용에 지각적인 과정이 많이 포함되어 있는 반면, 부정적 감정은 적게 포함된 상품평을 유용하게 인식함을 알 수 있었다. 그 외, 제품군별 특징으로는, 의류군의 경우, 비교급 표현이 많고, 전문성 지수는 낮으며, 한 문장에 포함된 단어 수가 적은 간결한 상품평이 유용하다고 인식되고 있었으며, 전자제품의 경우, 전문성 지수가 높고, 분석적이며, 진솔한 표현이 많고, 인지적 과정과 긍정적 감정(PosEmo)이 많이 포함된 상품평을 유용하게 인식하고 있었다. 이러한 연구결과는, 상품평의 언어적 특징들뿐만 아니라, 지각적, 인지적, 감정적 요소들이 상품평의 유용성을 결정하는데 영향을 미친다는 사실을 실증적으로 확인하였다는 점에서 의의가 있다.

본 연구는 다음의 한계점을 갖는다. 첫째, 본 연구는 상품평 특징을 추출하는데 LIWC에서 도출된 변수를 그대로 사용하였으며, 이러한 특징들이 정말 해당 항목을 정확히 측정한 것인지에 대한 검증은 수행하지 않았다. 예를 들어, 상품평의 전문성에 대한 특징은 LIWC의 clout변수를 통해서 측정된 값을 활용하였으나, 해당 값이 정말 전문성을 정확히 반영한 것인지에 대한 검증이 추가적으로 수행된다면, 더욱 정확한 연구가 될 수 있을 것으로 생각된다. 둘째, 본 연구결과, 제품에 대한 리뷰어의 평점이 높고, 부정적인 감

정은 적게 표현된 상품평이 소비자들에게 더욱 유용하다고 인식되는 경향이 있다고 도출되었으나, 이는 기존의 일부 연구[Chavalier, et al.(2006); Forman, et al.(2008)] 결과와는 차이가 있다. 이는 기존 연구들이 사용한 데이터와 본 연구의 데이터가 다르기 때문에 나타난 결과로 생각되지만, 보다 보편적인 결과를 도출하기 위해서는, 이러한 상이한 결과에 대한 이유를 면밀히 검토해 볼 필요가 있겠다. 마지막으로, 본 연구에서는 경험재와 탐색재 사례로 각 하나의 제품군에 대해서 실험하였으나, 보다 다양한 제품군으로 실험을 확장한다면, 보다 일반적인 결과를 도출할 수 있을 것으로 생각된다.

참고문헌(References)

- Baek, H., J. Ahn, and Y. Choi (2012), "Helpfulness of Online Consumer Revers: Reader's Objectives and Review Cues", *International Journal of Electronic Commerce*, vol.17(2), pp.99-126.
- Cao, Q., D. W. J., and Gan, Q. W (2011). "Exploring Determinants of Voting for the 'Helpfulness' of Online User Reviews: A Text Mining Approach," *Decision Support Systems*, vol.50(2), pp.511-521.
- Chevalier, J. A. & D. Mayzlin, (2003), "The Effect of Word of Mouth on Sales: Online Book Reviews", *Journal of Marketing Research*, vol.43(3), pp.345-354.
- Cho, S.-H., M.Y. Yi (2014), "Business Implications of the Factors that Determine Online Review Helpfulness", *Entrue Journal of Information Technology*, vol.13(1), pp.29-40.
- Choi, J.-W., and H.-J. Lee, "The Effects of Customer Product Review on Social Presence in Personalized Recommender Systems", *Journal of Intelligence and Information Systems*, Vol. 17, No. 3 (2011), 115~130.
- Forman, C., A. Ghose, B. Wiesenfeld (2008), "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets", *Information Systems Research*, vol.19(3), pp.291-313.
- Ghose, A. and P.G. Ipeirotis (2011), "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics", *IEEE Transactions on Knowledge and Data Engineering*, vol.23 (10), pp.1498-1512.
- Hong, Won Eui, M.Y. Yi (2015), "Helpfulness in Unhelpfulness: A Study of the Flip Side of the Coin on Review Helpfulness", KAIST MS. thesis
- Hyunmi Baek, J. H. Ahn, S. W. Ha (2011), "Identifying Factors Affecting Helpfulness of Online Reviews: The Moderating Role of Product Price", *The Journal of Society for e-Business Studies*, vol.16(3), pp.93-112.
- Kim, Y., and M. Song, "A Study on Analyzing Sentiments on Movie Reviews by Multi-Level Sentiment Classifier", *Journal of Intelligence and Information Systems*, Vol. 22, No. 3 (2016), 71~89.
- Korfiatis, N., E. Garcia-Bariocanal, and S. Sanchez-Alonso (2012), "Evaluating Content Quality and Helpfulness of Online Product Reviews: The Interplay of Review Helpfulness vs. Review Content", *Electronic Commerce Research and Applications*, vol.11(3), pp.205-217.

- Lee, G.N, Y.J. Won, C.H. Cho, G.H Lee, E.A. Na, H.S. Hwang and M.J. Shin, 「2010 Survey on the Internet Usage, Seoul: Korea Internet & Security Agency, 2010.
- Lee, H. G., H. Kwak (2013), "Investigation of Factors Affecting the Effects of Online Consumer Reviews", *Informatization Policy*, vol.20(3), pp.3-17.
- Lee, S.J. Lee, J. Y. Choeh, J. H. Choi (2014), "The Determinant Factors Affecting Economic Impact, Helpfulness, and Helpfulness Votes of Online", *Journal of IT Service*, vol.13(1), pp.43-55.
- Lee, M., and H. J. Lee, "Increasing Accuracy of Classifying Useful Reviews by Removing Neutral Terms", *Journal of Intelligence and Information Systems*, Vol. 22, No. 3 (2016), 129~142.
- Mahony, M. P., & B. B. Smyth (2010), A Classification-Based Review Recommender", *Knowledge-Based Systems*, vol.23(4), pp.323-329.
- McAuley and Leskovec2013] Julian McAuley and Jure Leskovec. 2013. Hidden factors and hidden topics: Understanding rating dimensions with review text. In *Proceedings of the 7th ACM Conference on Recommender Systems, RecSys' 13*, pp.165-172.
- McAuley, J., R. Pandey, and J. Leskovec (2015), "Inferring Networks of Substitutable and Complementary Products", *In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785-794.
- Mudambi, S.M. and D. Schuff (2010), "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com", *MIS Quarterly*, vol.34(1), pp.185-200.
- Pan, Y., and J. Q. Zhang (2011) "Born unequal: a study of the helpfulness of user-generated product reviews", *Journal of Retailing*, vol.87(4), pp.598-612.
- Wan, Y. (2015), The Matthew Effect in Social Commerce, *Electronic Markets*, pp.1-12.
- Yin, G., L. Wei, W. Xu, and M. Chen (2014), "Exploring Heuristic Cues for Consumer Perceptions of Online Reviews Helpfulness: The Case of Yelp. com.", *PACIS 2014 Proceedings*, 52.

Abstract

Impact of Semantic Characteristics on Perceived Helpfulness of Online Reviews*

Yoon-Joo Park** · Kyoung-jae Kim***

In Internet commerce, consumers are heavily influenced by product reviews written by other users who have already purchased the product. However, as the product reviews accumulate, it takes a lot of time and effort for consumers to individually check the massive number of product reviews. Moreover, product reviews that are written carelessly actually inconvenience consumers. Thus many online vendors provide mechanisms to identify reviews that customers perceive as most helpful (Cao et al. 2011; Mudambi and Schuff 2010). For example, some online retailers, such as Amazon.com and TripAdvisor, allow users to rate the helpfulness of each review, and use this feedback information to rank and re-order them. However, many reviews have only a few feedbacks or no feedback at all, thus making it hard to identify their helpfulness. Also, it takes time to accumulate feedbacks, thus the newly authored reviews do not have enough ones. For example, only 20% of the reviews in Amazon Review Dataset (Mcauley and Leskovec, 2013) have more than 5 reviews (Yan et al, 2014).

The purpose of this study is to analyze the factors affecting the usefulness of online product reviews and to derive a forecasting model that selectively provides product reviews that can be helpful to consumers. In order to do this, we extracted the various linguistic, psychological, and perceptual elements included in product reviews by using text-mining techniques and identifying the determinants among these elements that affect the usability of product reviews. In particular, considering that the characteristics of the product reviews and determinants of usability for apparel products (which are experiential products) and electronic products (which are search goods) can differ, the characteristics of the product reviews were compared within each product group and the determinants were established for each.

This study used 7,498 apparel product reviews and 106,962 electronic product reviews from

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A5A2A03047963)

** Department of Business Administration, Seoul National University of Science and Technology

*** Corresponding Author: Kyoung-jae Kim

Department of MIS, Dongguk University, Seoul,

30, Pildong-ro 1-gil, Chung-gu, Seoul 04620, Korea

Tel: 82-2-2260-3324, Fax: 82-2-2260-3684, E-mail: kjkim@dongguk.edu

Amazon.com. In order to understand a review text, we first extract linguistic and psychological characteristics from review texts such as a word count, the level of emotional tone and analytical thinking embedded in review text using widely adopted text analysis software LIWC (Linguistic Inquiry and Word Count). After then, we explore the descriptive statistics of review text for each category and statistically compare their differences using t-test. Lastly, we regression analysis using the data mining software RapidMiner to find out determinant factors.

As a result of comparing and analyzing product review characteristics of electronic products and apparel products, it was found that reviewers used more words as well as longer sentences when writing product reviews for electronic products. As for the content characteristics of the product reviews, it was found that these reviews included many analytic words, carried more clout, and related to the cognitive processes (CogProc) more so than the apparel product reviews, in addition to including many words expressing negative emotions (NegEmo). On the other hand, the apparel product reviews included more personal, authentic, positive emotions (PosEmo) and perceptual processes (Percept) compared to the electronic product reviews.

Next, we analyzed the determinants toward the usefulness of the product reviews between the two product groups. As a result, it was found that product reviews with high product ratings from reviewers in both product groups that were perceived as being useful contained a larger number of total words, many expressions involving perceptual processes, and fewer negative emotions. In addition, apparel product reviews with a large number of comparative expressions, a low expertise index, and concise content with fewer words in each sentence were perceived to be useful. In the case of electronic product reviews, those that were analytical with a high expertise index, along with containing many authentic expressions, cognitive processes, and positive emotions (PosEmo) were perceived to be useful. These findings are expected to help consumers effectively identify useful product reviews in the future.

Key Words : review text, helpfulness, text mining, LIWC, prediction model

Received : June 22, 2017 Revised : September 14, 2017 Accepted : September 19, 2017

Publication Type : Regular Paper Corresponding Author : Kyoung-jae Kim

저 자 소 개



박 윤 주

고려대학교 컴퓨터학과에서 학부 및 석사학위를 취득하였으며, 2006년 한국과학기술원에서 경영공학 박사학위를 취득하였다. 이 후, New York University의 Stern Business School에서 초빙연구원으로 근무하였으며, 삼성생명 정보기획부서에서 과장으로 근무하였다. 2010년부터 현재까지 서울과학기술대학교 경영학과에서 부교수로 재직 중이다. 주요 연구분야는 데이터마이닝을 활용한 경영문제 해결이다.



김 경 재

현재 동국대학교 서울 경영대학 경영학부 교수로 재직 중이다. KAIST에서 경영정보시스템을 전공으로 박사학위를 취득하였으며, 연구관심분야는 비즈니스 애널리틱스, 고객관계관리, 빅데이터 등이다.