

리뷰어 평점 이력이 리뷰 조작에 대한 인식 및 리뷰 유용성에 미치는 영향: 여행플랫폼을 중심으로

장문경*

가천대학교, 경영학과 조교수

이새롬**

경북대학교, 경영학부 조교수

백현미***

고려대학교, 미디어학부 부교수

국문 요약

고객들은 조작된 온라인 리뷰가 범람하는 가운데 진정성과 가치를 지닌 리뷰를 보고자한다. 귀인 이론(Attribution theory)의 관점에서, 사람들은 리뷰어의 과거 평가 이력을 바탕으로 리뷰가 진정성 있는지를 판단하는 경향이 있다. 이러한 배경에서 본 연구의 목적은 리뷰어의 과거 평점 이력이 조작된 리뷰로 인식하는 것에 어떠한 영향을 미치며, 최종적으로 리뷰 유용성이 어떠한 영향을 미치는지 알아보는 것이다. 제안된 가설을 검증하기 위해 2차 데이터 분석(연구1)과 실험(연구2)을 수행했으며, 두 연구는 일관된 결과를 보여준다. 연구 1은 리뷰어의 과거 평가 이력이 리뷰 유용성에 미치는 영향을 분석하였다. 귀인이론에 근거하면, 사람들은 리뷰를 다른 목적을 가지고 작성되었다고 인식할 경우에 리뷰가 조작되었다고 생각하고, 그 리뷰가 물건이나 서비스의 진정한 가치를 평가하지 않았다고 간주한다. 따라서 해당 리뷰는 유용성이 낮게 평가되는 경향이 있다. 2차 데이터를 분석하기 위해 우리는 Python을 이용한 웹 스크래퍼를 개발하여 TripAdvisor(TripAdvisor.com)에서 호텔 정보, 리뷰, 리뷰 정보 등의 연구 데이터를 수집하였다. 수집한 890명 리뷰어에 대한 100,621개의 리뷰를 분석하기 위해 음이항 회귀 분석을 수행하였다. 분석 결과, 평균 평점을 낮게 주는 리뷰어의 경우에 리뷰 유용성에 유의미한 영향을 미치지 않는 것으로 나타났다. 사람들은 극단적인 평점을 거의 주지 않는 리뷰어가 작성한 리뷰가 더 도움이 된다고 평가했다. 연구 2는 리뷰어의 과거 평점 이력을 기준으로 리뷰가 조작되었다고 평가하는 사람들의 인식 프로세스를 실험하였다. 실험 결과, 사람들은 리뷰어의 과거 평점 이력이 평균적으로 평점을 낮게 주는 경우에는 리뷰가 의심스럽다고 판단하지 않는 것으로 나타났다. 그리고 사람들은 리뷰어가 대부분 극단적인 평점을 주는 이력이 있다면 해당 리뷰어가 작성한 리뷰가 의심스럽다고 판단하는 것으로 나타났다. 연구2는 사람들이 리뷰어의 과거 평점 이력을 바탕으로 리뷰가 조작되었는지 또는 리뷰가 도움이 되는지 판단하는 경향이 있음을 보여준다. 본 연구는 귀인이론을 바탕으로 리뷰어의 과거 평점 이력이 리뷰 조작성에 대한 인식과 리뷰 유용성에 미치는 영향을 분석하여, 해당 연구분야에 새로운 관점을 추가한 기여점이 있다.

핵심어: 온라인 리뷰, 리뷰 유용성, 과거 평점 이력, 리뷰 조작, 귀인 이론

1. 서론

정보통신기술(ICT)의 발전으로 소비자들은 온라인 리뷰를 활용하여 구매 결정에 대한 위험과 불확실성을 줄일 수 있다(Yang et al., 2019). 최근 온라인 리뷰가 점차 상당한 산업적 영향력에 가지면서, 온라인 리뷰는 재화나 서비스

의 구매에 이전보다 더 많은 영향을 미칠 뿐만 아니라 (Koh et al., 2010), 소비자의 구매의사 결정과정을 완전히 변화시켰다(Wan & Nakayama, 2014). 특히 관광상품이나 서비스의 경우 호텔 선정 결정 과정에서 온라인 리뷰에 의존하는 고객이 많은데, 호텔은 구매 전 평가가 불가능한 체험상품이기 때문이다(Lung et al., 2013; Xie et al., 2014).

* mkjang@gachon.ac.kr

** saeromlee@knu.ac.kr

*** lotus1225@korea.ac.kr

이에 따라 온라인 여행 플랫폼에 대한 리뷰가 증가하고 있는데, 대표적으로 트립어드바이저의 총 댓글 및 의견 수는 2020년 기준 8억8400만 건에 달한다(Statista, 2021).

이렇듯 고객들은 너무나 많은 온라인 리뷰를 접하면서 온라인 리뷰 평점이나 리뷰 콘텐츠가 제품이나 서비스를 정확하게 반영하는지에 대해 궁금해하기 시작했다. 온라인 리뷰가 제품이나 서비스의 품질을 항상 정확하게 나타내지 못하는 경우가 빈번해지고 있기 때문이다. 예를 들어, 모든 사람들이 항상 제품이나 서비스에 대한 리뷰를 매번 쓰는 것이 아니라, 그 제품이나 서비스가 특히 좋아하거나 싫어하는 상황에서만 높은 점수를 주거나 낮은 점수를 주기 위한 특정한 상황에서만 리뷰를 쓴 경향이 있다고 알려져 있다. 이를 자기선택 편향의 한 형태인 과소보고 편향이다(Bhole & Hanna, 2017; Hu et al., 2006; Koh et al., 2010). 또한, 누구나 온라인에서 리뷰를 쓸 수 있기 때문에, 리뷰를 쓰는 사람들의 실제 의도를 사용자가 정확히 알 수 없다. 리뷰가 매출에 미치는 영향이 큰 만큼 특정한 목적을 위해 온라인 리뷰를 조작하는 경우도 있다. 리뷰 조작은 일반적으로 두 가지 방향으로 이루어진다는데, 특정 제품이나 서비스의 판매를 촉진하기 위한 극단적으로 긍정적인 평점을 작성하는 경우가 있다. 이와는 반대로 경쟁 제품이나 서비스를 평가절하하기 위해 극단적으로 부정적인 평점을 작성하는 경우도 있다(Jindal & Liu, 2008). 이와 같은 상황이 빈번해짐에 따라, 악성 리뷰로 인해 피해를 보는 판매자가 많이 발생하고 있고, 많은 구매자 역시 리뷰 조작을 통한 비효율적인 구매 결정으로 어려움을 겪고 있다. 그러나 리뷰 자체의 특성(예: 리뷰의 긍정/부정성, 길이, 가독성 등)만으로 항상 편향된 리뷰를 작성하는 리뷰어인지, 나쁜 의도를 가진 리뷰어인지 판단하기는 쉽지 않다. 이는 신뢰도 문제, 즉 리뷰가 제품이나 서비스를 제대로 평가하고 있는지 확인하기 어려운 문제라고 할 수 있으며, 개별 리뷰의 특성만을 검토하는 것으로는 해결하기 어렵다. 따라서 리뷰어 자체가 리뷰 원천에 대한 신뢰도를 나타내는 지표이기 때문에 리뷰를 작성하는 리뷰어의 신뢰도를 이해하는 것이 더욱 중요해지고 있다(Metzger et al., 2003; Sundar, 2008). 이러한 연구의 중요성에도 불구하고 리뷰 조작이나 편향을 직접 검증하기 어렵기 때문에 아직까지 관련된 연구가 거의 이루어지지 않았다(Ma et al., 2019).

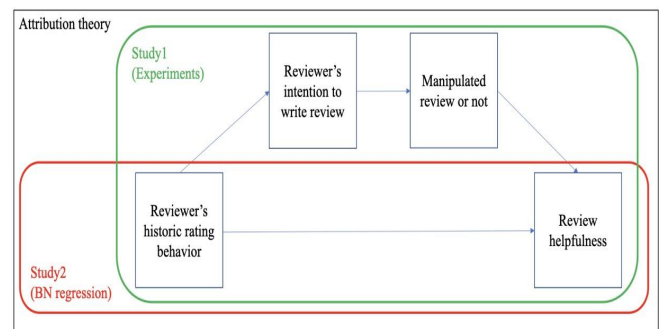
이와 같은 배경에서 본 연구는 리뷰어를 리뷰 원천에 대한 신뢰도를 나타내는 지표로 보고, 리뷰어의 과거 리뷰 작성 행동이 리뷰의 유용성 및 리뷰 조작 인식에 미치는 영향을 알아보는 것을 목적으로 한다. 구체적으로, 본 연구는 귀속 이론(Kelley, 1973)을 이론적 근거로 하여 사람

들이 리뷰어의 과거 리뷰 행동(부정성과 극단성)을 바탕으로 리뷰 작성 동기를 평가함으로써 리뷰의 신뢰성을 평가한다고 상정하였다. 귀속 이론에 따르면, 리뷰어가 다른 목적을 가지고 리뷰를 작성하여 리뷰가 조작되었다고 판단할 경우, 리뷰가 해당 제품이나 서비스의 진정한 가치를 반영하고 있다고 간주되지 않기 때문에 리뷰 유용성이 낮은 경향이 있다. 예를 들어, 사람들은 과거 리뷰에서 다양한 평점으로 평가를 하는 리뷰어보다 항상 최고 평점을 주는 리뷰어를 더 신뢰할지, 또한 리뷰가 유용한지 여부를 결정할 수 있다. 리뷰어의 평점 기록을 바탕으로 리뷰어가 편향되어 있는지 또는 제품 등급을 조작하려는 시도에 관여하고 있는지 여부를 판단할 수 있다고 알려져 있다(Jindal & Liu, 2008).

II. 연구방법

2.1. 연구개요

연구목적 달성을 위해 <그림 1>과 같이 실험(연구1)과 2차 데이터 분석(연구2)을 수행하였다.

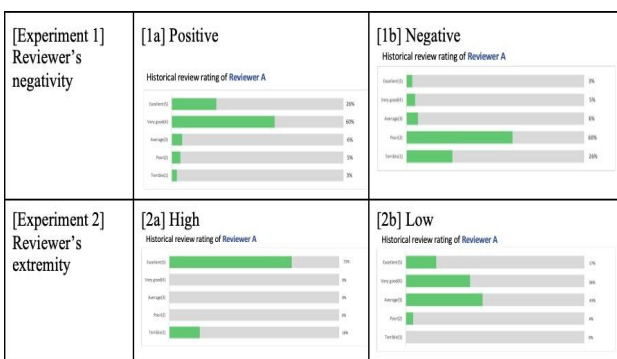


<그림 1> 연구개요

2.2. 연구1: 리뷰어의 과거 평점 이력이 리뷰 조작에 대한 인식과 리뷰 유용성에 대한 영향에 대한 실험

연구 1의 주요 목적은 사람들이 리뷰어의 리뷰 작성 의도를 어떻게 인지하고, 이를 바탕으로 리뷰 조작 여부에 대한 인식을 어떻게 하는지를 검증하는 것이다. 본 연구는 리뷰어의 과거 평점 이력의 부정성과 극단성에 중점을 두기 때문에 각각 두 가지 측면을 보기 위해 두 가지 실험을 수행한다. 실험을 위해 2022년 8월 2일부터 8월 29일까지 엠브레인(embbrain.com)에서 328명의 참가자를 모집하고,

실험 참가자는 표본 편향을 피하기 위해 성별과 연령별로 고르게 할당되었다. 실험 참가자는 참가 동의서를 제출한 후, 먼저 웹사이트에서 소비자들이 다양한 호텔에 대한 리뷰어의 평점 이력을 읽을 수 있다는 설명을 들었다. 이어 참가자들에게 웹사이트에서 목을 적당한 호텔을 찾고 있다고 상상하고 리뷰어의 과거 평점 이력을 꼼꼼히 읽어보라고 했다. 다음으로, 참가자들은 네 가지 조건 중 하나에 무작위로 노출되었다. 각각의 조건에서 참가자들은 같은 호텔에 대한 리뷰와 리뷰어의 과거 평점 이력 분포를 보았다. 네 가지 조건에서 리뷰의 텍스트와 평점은 동일하고, 유일한 차이점은 리뷰어의 과거 평점 이력 분포이다. 그림 2은 리뷰어의 과거 평점 이력 분포와 관련한 자극물을 보여준다. 실험 1에서 참가자는 두 조건 중 하나에 무작위로 할당되었다(과거 평점 이력 분포의 부정성: 높음 대 낮음). 과거 평점 이력의 높은 부정성 하에서, 리뷰어의 과거 평점 이력의 평균은 2이다. 과거 평점 이력의 낮은 부정성 하에서, 리뷰어의 과거 평점 이력의 평균은 4이다. 자극물의 t-테스트 결과는 [1a] 자극의 평균 등급(M=5.44, SD=1.06)이 [1b] 자극의 평균 등급(M=3.53, SD=1.88, $t(160)=7.96, p=0.00$)보다 높다는 것을 보여준다. 실험 2에서 참가자는 두 조건 중 하나에 무작위로 할당되었다(과거 평점 이력의 극단성: 높음 대 낮음). 리뷰어의 과거 평점 이력 평균은 과거 평점 이력의 높은 극단성과 낮은 극단성 모두에 대해 3.9이지만, 과거 평점 이력분포의 분산은 다르다. 자극물의 t-테스트 결과는 [2a] 자극물의 높은 극단성(M=5.02, SD=1.20)이 [2b] 자극물의 평균 극단성(M=4.04, SD=1.17, $t(164)=5.36, p=0.00$)보다 높다는 것을 보여준다.

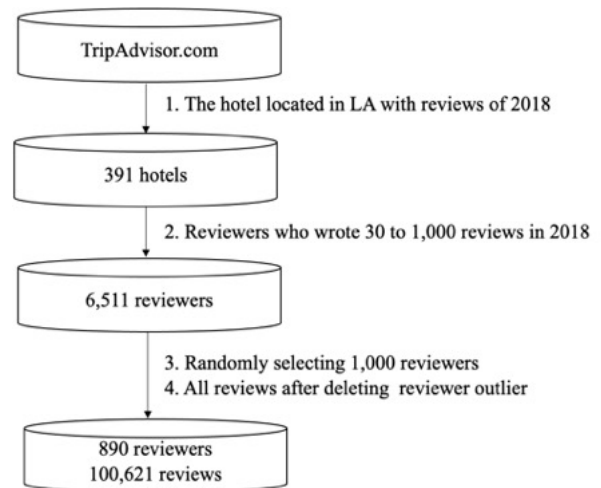


<그림 2> 연구1의 자극물

실험 결과, 사람들은 리뷰어의 과거 리뷰 행동이 대부분 긍정적일 경우에는 리뷰가 의심스러웠다고 판단하는 경향을 보이지 않았다. 그리고 사람들은 리뷰어의 과거 평점 분포가 대부분 극단적일 경우에 리뷰가 의심스러웠다고 판단하는 경향을 보였다.

2.3. 연구2

연구 2에서는 Python을 이용한 웹 스크레이퍼를 개발하여 트립어드바이저(TripAdvisor.com)에서 호텔 정보, 리뷰, 리뷰 정보 등의 연구 데이터를 수집하였다. <그림 3>은 데이터를 수집하기 위한 구체적인 과정을 나타낸다. 본 연구는 캘리포니아 주 로스앤젤레스의 391개 호텔에서 데이터를 수집했으며 리뷰를 위한 데이터(예: 리뷰 등급, 리뷰 텍스트, 리뷰 발행일 및 리뷰 유용성)와 리뷰를 위한 데이터(예: 리뷰어의 ID, 이름 등)를 수집했다. 특이치를 제거하기 위해 매우 적거나 매우 많은 리뷰를 가진 리뷰어를 포함하지 않고 2018년에 30개 이상의 리뷰와 1,000개 미만의 리뷰를 작성한 6,511명의 리뷰어로부터 데이터를 수집했다. 다음 단계로 1,000명의 검토자를 무작위로 선정하여 과거 평점 이력을 조사하였다. 이어 선정된 1,000명의 리뷰어가 트립어드바이저에 가입 이후부터 2018년까지 작성한 리뷰를 모두 수집했다. 데이터 특이치가 발생한 경우를 제외하고 890명의 리뷰어가 작성한 100,621개의 리뷰를 수집하였다. 이 100,621건의 리뷰는 리뷰어의 과거 평점 이력에 대한 다른 변수를 만들기 위해 890개의 리뷰 수준으로 통합되었다. 리뷰어가 부여한 평점 평균, 평점 표준편차, 평점 극단비 등 각 검토자의 과거 평점 이력에 대한 변수를 얻었다. 또한 리뷰어의 프로필 페이지를 통해 리뷰어가 여행한 국가의 수에 대한 데이터를 얻었다.



<그림 3> 연구2 데이터 수집 과정

890명의 리뷰어에 대한 100,621개의 리뷰를 분석하기 위해 음이항 회귀 분석을 수행하였다. <표 1>은 음이항 회귀 분석의 결과이다. 분석 결과는 과거 평점의 평균이 리뷰 유용성(계수=-0.0954, $p=0.669$)에 유의미한 영향을 미치지

않는 것으로 나타났다. 이는 리뷰어의 과거 평점 이력 부정성이 리뷰 유용성에 유의미한 영향을 미치지 않음을 의미한다. 과거 평점 이력 중 평점 극단성의 비율은 리뷰 유용성에 유의미하게 부정적인 영향을 미치는 것으로 나타났다(계수=-3.8548, $p<0.001$). 이는 주로 극단적인 평점을 주는 리뷰어가 작성한 리뷰가 도움이 되지 않는 경향이 있음을 나타낸다. 통제변수인 ExtremeRating, WordCount, ExcidedDay 및 Country는 모두 유용성에 긍정적인 영향을 미치는 것으로 나타났다. 이 결과는 평점이 극단적일수록 리뷰에 도움이 된다고 생각했고, 리뷰에 단어가 많을수록 리뷰에 도움이 된다고 생각했으며, 리뷰가 발표된 지 오래될수록 리뷰에 도움이 된다고 생각했으며, 리뷰가 방문한 국가가 많을수록 리뷰에 도움이 된다고 생각한다고 해석할 수 있다. 설명력은 0.0798로 분석 신뢰도가 높은 것으로 나타났다.

<표 1> 음이항 회귀 분석 결과

Variable	Coefficient	Std. error	Z	p
RatingAverage	-0.0954	0.2233	-0.43	0.669
ExtremeRatio	-3.8548***	1.0150	-3.80	0.000
Country	0.0820***	0.0126	6.53	0.000
ReviewNum	-0.0012	0.0009	-1.24	0.215
ExtremeRating	0.4192**	0.1733	2.42	0.016
WordCount	0.0026***	0.0006	4.56	0.000
Picture	0.3257	0.2015	1.62	0.106
ElapsedDay	0.4025***	0.1026	3.92	0.000
Constant	-4.4407***	1.0648	-4.17	0.000
Likelihood-ratio test of alpha=0: chibar2(01)=417.42 p>= chibar2=.000 Number of obs=1,780 Pseudo R ² =0.0798				

*. $p<0.05$, **. $p<0.01$, ***. $p<0.001$
Dependent Variable: Helpfulness

III. 결론

이 연구는 크게 두 가지의 연구의의를 가진다. 첫째, 많은 선행연구들이 인지된 온라인 리뷰 유용성에 영향을 미치는 다양한 요인을 조사하였지만, 리뷰어의 평가 이력을 바탕으로 판단된 리뷰 신뢰도가 인지된 리뷰 유용성에 어떤 영향을 미치는지 조사한 연구는 거의 없다. 본 연구는 리뷰어의 과거 리뷰 행동 중 부정성과 극단성에 초점을 맞추어, 리뷰어의 과거 리뷰 행동이 인식된 리뷰 유용성에 미치는 영향을 종합적으로 검토하였다. 둘째, 본 연구는 귀속이론에 기초하여 편향과 조작이 리뷰 신뢰성에 어떠한 부정적인 영향을 미치는지 살펴보았다. 또한, 이 연구는 리뷰어의 과거 평점 기록이 리뷰어가 편향성을 가지고 있는지 또는 리뷰 조작에 관여하고 있는지 여부를 결정하

는 데 도움이 될 수 있음을 보여준다. 따라서 본 연구의 결과는 리뷰어의 과거 평점 이력에 대한 정보를 이론적으로 그리고 실질적으로 관리할 필요성을 다루고 있다. 이러한 연구 의의에도 불구하고 이 연구 역시 한계가 있다. 먼저, 본 연구는 간접적으로 리뷰의 과소 보고 편향을 측정했다. 과소 보고 편향을 직접 측정할 수 있으려면 실제 구매 건수 대비 실제 리뷰 작성 건수에 관한 데이터가 필요하다. 이 데이터를 이용할 수 있다면, 리뷰를 구매 건수보다 적게 작성하는 성향을 가진 검토자가 작성한 리뷰의 유용성을 연구할 수 있을 것이다. 또한 본 연구에서는 2차 데이터와 실험을 통해 리뷰어의 과거 평점 이력이 리뷰 유용성이나 리뷰 조작에 미치는 영향에 대해 분석하였으나, 다른 방법론을 추가적으로 고려하여 연구 결과를 더욱 풍부하게 만들 수 있을 것이다. 예를 들어, 질적 분석을 통해 리뷰 텍스트의 내용을 통제 변수로 더 고려할 수 있을 것이다. 또한, 극단적 평가를 주로 하는 리뷰어의 리뷰나 과거 평점 분포의 표준편차가 낮은 리뷰어의 글을 신뢰하지 않는 이유에 대한 설문조사를 추가 연구로 진행할 수 있을 것이다.

참고문헌

Bhole, B., & Hanna, B.(2017). The effectiveness of online reviews in the presence of self-selection bias. *Simulation Modelling Practice and Theory*, 77, 108-123.

Hu, N., Pavlou, P. A., & Zhang, J.(2006). Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication. *Proceedings of the 7th ACM Conference on Electronic Commerce*, 324-330.

Jindal, N., & Liu, B.(2008). Opinion spam and analysis. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 219-230.

Jindal, N., Liu, B., & Lim, E. P.(2010). Finding unusual review patterns using unexpected rules. *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, 1549-1552.

Koh, N. S., Hu, N., & Clemons, E. K.(2010). Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electronic Commerce Research and Applications*, 9(5), 374-385.

Ma, H., Kim, J. M., & Lee, E.(2019). Analyzing dynamic review manipulation and its impact on movie box office revenue. *Electronic Commerce Research and Applications*, 35, 100840, 1-13.

- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & Mccann, R. M.(2003). Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment. *Communication Yearbook*, 27, 293-335.
- Statista (2021). Tripadvisor-statistics & facts, available at: (Retrieved 2022.01.03.), <https://www.statista.com/topics/3443/tripadvisor/#dossierKeyfigures/>.
- Sundar, S. S.(2008). *The MAIN model: A heuristic approach to understanding technology effects on credibility*", Metzger M. J. and Flanagin, A. J. (Ed.). Cambridge, MA: Digital media, youth, and credibility, MIT Press.
- Wan, Y., & Nakayama, M.(2014). The reliability of online review helpfulness. *Journal of Electronic Commerce Research*, 15(3), 179-189.
- Yang, S., Zhou, Y., Yao, J., Chen, Y., & Wei, J.(2019). Understanding online review helpfulness in omnichannel retailing. *Industrial Management & Data Systems*, 119(8), 1565-1580.