

감성분석을 이용한 온라인 체험 내 비정형데이터의 주관도가 고객만족에 미치는 영향 분석

Sentiment Analyses of the Impacts of Online Experience Subjectivity on Customer Satisfaction

서 예 은 (Yeeun Seo)

한양대학교 경영대학 석사

이 상 용 (Sang-Yong Tom Lee)

한양대학교 경영대학 교수, 교신저자

요 약

코로나19로 인한 팬데믹 상황에서도 여전히 여행에 대한 욕구와 수요가 시장에 존재하고 있다. 이러한 상황에서 정보기술(IT)의 발달로 인해 온라인에 대한 접근성과 유용성 및 디지털 기기의 활용도 함께 증가하였고, 비대면으로도 새로운 경험을 얻을 수 있는 ‘온라인 체험(Online Experience)’ 시장이 급격하게 성장하였다. 본 연구는 AirBnB 온라인 체험 서비스에서 서비스제공자(Provider-oriented)와 서비스이용자(User-oriented)에 의해 생성된 정형 및 비정형 데이터가 고객만족에 미치는 영향에 대해 분석하였다. 분석에 사용된 데이터는 파이썬 웹크롤러로 수집되었으며, 주요 변수인 비정형 데이터는 전처리와 감성분석을 거쳐 회귀분석에 사용되었다. 분석 결과, 주요변수인 호스트가 생성한 체험 소개글, 호스트 소개글과 같은 비정형 데이터는 텍스트의 생성 목적에 따라 주관도(Subjectivity)가 다르게 나타나며, 체험 소개글은 주관적일수록, 호스트 소개글은 객관적일수록 고객 만족에 유의한 정의 영향을 미치는 것을 확인하였다. 또한, 에어비앤비 온라인 체험에 참여한 게스트가 생성한 정형 데이터는 다른 게스트의 만족에 긍정적인 영향을 미치는 것을 확인하였다. 본 연구 결과는 온라인 공유경제 플랫폼 이해관계자 및 온라인 체험 지식경역에 관심을 갖는 연구자에게 다양한 시사점을 제공할 것으로 기대한다.

키워드 : 온라인 체험 서비스, 텍스트 마이닝, 감성분석, 고객 만족

I. 서 론

2020년부터 시작된 Covid-19 팬데믹으로 인하여, 각국은 공공장소로의 접근 제한 및 폐쇄를 통

해 국가간 이동과 여행이 제한되었다. 한국전자통신연구원(2021)은 사회적 거리두기로 인해 재택근무, 온라인 원격수업, 온라인 판매, 비대면 공연 등이 새로운 사회 흐름으로 나타나게 되었다고 보고하였다. 또한 한국관광공사에서 제시한 ‘빅데이터를 활용한 관광 트렌드 분석(2021)’의 한 조사 결과를 보면 코로나19가 확산되기 시작한 이후로

† 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A3A2A0 2093277).

‘랜선여행’, ‘방구석’이라는 단어의 월별 언급량이 급격하게 증가한 것을 발견할 수 있다. 이는 팬데믹 상황에서 직접적인 체험이 불가함에도 여행에 대한 욕구와 간접적이거나 여행에 대한 수요가 여전히 시장에 존재하는 것으로 파악할 수 있다.

팬데믹 상황부터 세계의 여행 흐름을 주도한 회사로서는 Airbnb(이하 ‘에어비앤비’)를 꼽을 수 있는데, 2008년 8월 설립 이후 190개 이상의 국가와 34,000개 이상의 도시에 진출해 있으며, 한국에는 2013년 1월부터 정식 서비스가 시작되었다. 에어비앤비는 전 세계에 있는 홈스테이 공급자와 수요자를 연결하는 온라인 숙박 중개 플랫폼으로 시작되었으며, P2P 네트워크와 IT기술을 통해 서비스를 제공하고 있다(한국문화관광연구원, 2014). 2016년부터는 ‘Experience(이하 ‘체험’)’라는 서비스를 개시하여, 단순 숙박 서비스 플랫폼에서 한 단계 더 나아가 현지 호스트가 직접 투어 프로그램을 구성함으로써, 숙박객은 보다 더 ‘현지스러운’ 음식, 문화, 관광 서비스를 제공받아 여행지를 새로운 시각에서 만나 볼 수 있게 되었다. 비록 에어비앤비의 초기 시장 진입은 느린 편이었으나 양면시장이 지닌 교차 네트워크 효과에 힘입어 빠르게 성장하며 온라인 플랫폼 시장에 자리를 잡았다.

에어비앤비의 모든 대면 활동이 코로나19로 인해 일시 중단되기도 하였으나, 빠른 사업 전환으로 2020년 4월부터 비대면 서비스인 ‘Online Experience(이하 ‘온라인체험’)’ 플랫폼 서비스를 개시하였다. 서비스가 개시된 이래로 온라인 체험 인기 호스트의 1년 최대 수입금액이 한화 약 6억 원에 달한 것과 온라인 체험 예약 수요가 2020년 4분기 대비 2021년 2분기 기준으로 약 130%가 증가한 것에서 그 성장성을 주목할 필요가 있다.

에어비앤비가 제공하는 온라인 체험 서비스는 기존 오프라인에서만 진행되던 체험을 온라인 Zoom 서비스와 제휴해 호스트와 게스트 간의 “양방향 교류”가 비대면에서도 진행될 수 있도록 하고 있다. 우리나라에서는 흔히 랜선투어라고 불리우

기도 하는데, 랜선투어는 관광지를 직접 방문하는 것이 아닌 인터넷 통신망의 콘텐츠를 통하여 온라인상에서 관광체험을 하는 것을 의미한다(김종상, 2021). 그런데, 에어비앤비의 온라인 체험 서비스는 일반적 랜선투어와 차이점이 존재한다. 에어비앤비에서는 오프라인에서 진행되던 그룹 투어 방식으로 진행되기 때문에, 호스트와 게스트 간의 양방향 뿐만 아니라, 사전 예약된 게스트간에도 소속감과 함께 상호작용이 일어날 수 있는 것이다.

가족과 함께 즐길 수 있는 ‘모로코 가족과 함께 즐기는 요리(모로코)’, ‘젤마와 함께하는 바르셀로나 여행(스페인)’ 등의 체험에서부터, ‘잠을 부르는 양과 함께하는 명상(영국)’, ‘마술의 비밀 배워보기(영국)’와 같은 개인 취미 활동 등 다양한 체험이 온라인으로 접근 가능하다. 뿐만 아니라 올림픽, 패럴림픽 출전 공식 파트너 선수들이 직접 호스트하는 ‘올림픽 리나 테일러가 들려주는 금메달을 위한 마음가짐(미국)’, ‘스콧 리어던이 알려주는 정신력 관리법(오스트레일리아)’ 등 세계 유명 선수들과 소통하며 그들의 경험과 훈련비법 등을 생생하게 들을 수 있는 서비스가 제공되고 있다. 이는 코로나19라는 절망적인 상황에서도 여행에 대한 수요와 코로나 블루(Corona Blue)를 극복하기 위한 소비자의 심리를 겨냥한 것인데, 추가적으로 경제적으로 수입이 끊겨 어려운 상황 속의 현지 호스트들에게는 새로운 수익 창출로 이어졌다.

이처럼 4차 산업혁명을 주도하고 있는 정보기술(IT)의 발달은 팬데믹 상황에서도 소비자들에게 온라인을 통해 전 세계를 여행할 수 있도록 새로운 경험 및 서비스를 제공할 수 있게 하였다. 에어비앤비가 2020년 7월, 900명의 참가자를 대상으로 자체적으로 실시한 설문조사에 따르면 온라인 체험을 경험한 응답자 절반 이상이 향후 코로나가 끝나더라도 온라인 체험 서비스를 이용할 것이며, 대면 활동이 재개된다고 하여도, 오프라인 체험을 하기에 앞서 온라인 체험을 먼저 겪어 보는 것에 긍정적인 반응을 보인 바 있다. 그 이유로는 새로운 것을 배울 수 있다는 것과 다른 문화와 라이프스타일을

만날 수 있다는 것이 꼽혔다.²⁾

위드코로나(With Corona) 시대에 오프라인 활동이 늘어나면서 온라인 체험이 다소 줄어들 가능성을 배제할 수는 없으나, 다른 한편으로는 메타버스 기술의 발전을 통하여, 온라인 체험 시장의 새로운 지평이 열릴 수도 있을 것으로 판단된다. 메타버스를 통하여 가상 온라인 체험형 콘텐츠 개발은 대한민국 정부에서도 미래 먹거리 사업의 하나로 꼽고 있다. 2020년 에 발표된 “가상융합경제 발전전략”³⁾에서는 메타버스와 융합된 온라인 가상 체험을 신 성장 동력으로 언급하고 있는데, 이는 팬데믹 이후의 시대에도 온라인 체험 시장이 성장할 수 있음을 시사하는 것이다. 따라서 본 연구를 통해 코로나19가 촉진한 디지털 전환과 비대면 트렌드 사이에서 빠르게 자리잡은 온라인 체험이 향후에도 지속적으로 고객 만족을 줄 수 있는 요인이 무엇인지를 살펴보고자 한다.

선행연구에서는 주로 체험 이용자에게 설문조사 방법을 통해 단방향적인 데이터를 통해 연구가 진행되었으나(조성호, 2019; 홍순기 등, 2022), 본 연구에서는 정보제공자가 정보이용자에게 미치는 영향(Schmitt, 1999)도 주요변수로 고려하여, 정보제공자와 정보이용자 각각에 의해 생성된 양방향 데이터가 고객만족에 미치는 영향을 보고자 한다. 따라서, 온라인 체험에서 제공되는 정형, 비정형 데이터와 요인들이 고객만족에 어떠한 영향을 미치고 있는지에 대해 실증적으로 탐구해보고, 공유경제 속 온라인 체험 서비스의 성패 요소들을 논의해보고자 한다.

II. 이론적 배경

2.1 온라인 공유경제와 에어비앤비에 대한 연구

상업경제(Commercial economy)와 대비되는 개

념으로 Lessig(2008)가 제시한 공유경제(Sharing economy)는 세계금융위기 이후 소비의 지속 불가능성과 자본주의에 대한 대안으로 등장하게 되었다. 상업경제는 화폐를 이용한 거래가 이루어지는 반면, 공유경제에서는 사람 간의 관계를 바탕으로 필요에 따라 서로의 제품과 서비스를 교환한다. 과거부터 물물교환 형식의 교환 형태는 존재해왔지만, 현재 온라인 플랫폼으로 대표되는 기술 매개(Technology-mediated) 거래는 유래 없는 방식으로 지리적 경계와 범위를 넘어 폭넓은 수준의 활동을 가능하게 하고 있다. 온라인 기술을 이용하여 일종의 협력적 소비 형태가 자리잡게 되면서, 개인은 소비자의 역할 뿐만 아니라 생산자, 혹은 공급자로서의 역할도 함께하게 되었다. 이러한 공유경제는 초기에는 비영리적인 성격을 지닌 의미로 사용되었으나, 다양한 분야에서 의미가 확장되면서 현재는 상업적 의미를 담은 ‘상업적 공유경제’의 개념으로 불리게 되었다(Belk, 2014; 정석완, 2018). 기존 상업경제에서는 행위의 주체가 소비자와 공급자 두 부류로 나뉘었다면, 상업적 공유경제는 두 주체를 연결하는 플랫폼의 역할이 중요해졌으며, 이러한 중개 서비스가 상업적 공유경제로 인식되고 있다(Kumar *et al.*, 2018).

플랫폼을 통해 형성되는 시장을 경제학에서는 양면시장(Two-sided Market)이라고 한다. 양면시장에서는 서로 다른 두 사용자 집단이 플랫폼을 통해 상호작용을 하며, 이 때 창출되는 가치는 간접적 네트워크 외부성(Indirect network externality)의 영향을 받는 시장을 일컫는다(이상규, 2014). 수요자가 많으면 공급자를 확보하기 쉽고, 반대로 공급자가 많으면 수요자 또한 확보하기 쉽다. 이처럼 양쪽의 사용자가 많아질수록 규모의 경제가 실현되며, 플랫폼의 경쟁력은 강화된다(김주락, 2021). 그러나 앞서 언급한 것처럼 온라인을 기반으로 한 공유경제에 대해 긍정적인 전망만 있는 것은 아니다. 누구나 소비자 혹은 생산자로서 시장에 자유로이 참여 가능하다는 점에서 정보 불균형에 대한 우려는 어느정도 완화될 수 있으나, 인

2) <https://www.mk.co.kr/news/culture/9831412>.

3) <https://www.korea.kr/news/policyNewsView.do?newsId=148880843>.

구학적 관점에서는 공유경제에 접근할 수 있는 문화 및 디지털 자원을 갖춘 공간의 이용자에 한해 공유경제가 이루어질 수 있으므로 오히려 공간 간 격차를 심화할 수 있다는 연구결과(Dredge and Gyimóthy, 2015)도 존재한다.

에어비엔비(Airbnb)와 같은 C2C(Customer to Customer) 공유경제 플랫폼은 사용자와 공급자가 더욱 긴밀한 관계를 맺지만 각기 다른 입장에서 플랫폼에 접근하게 된다(신진호 등, 2017). 따라서 서로 다른 이해관계를 가진 주체로 인해 생성되는 정보가 고객 만족에 미치는 영향의 정도는 다를 수 있으므로, 독립변수를 서비스 이용자가 생성한 정보(user-oriented)와 서비스 제공자가 생성한 정보(provider-oriented) 두 그룹으로 나누어 볼 수 있으며, 본 연구도 이와 같은 방식으로 분석을 진행하였다.

2.2 온라인 체험에 대한 연구

체험은 조직과 이용자 간에 감각적 자극, 감성, 정보 등을 교환하는 접점들의 집합으로 어떤 조직과의 직접적 접촉 혹은 인터넷과 같은 매개를 경유한 간접적 체험을 통해 형성된다(이정렬, 김상훈, 2011).

Schmitt(1999)에 의하면 온라인 공간에서의 체험 제공자(Experiential provider)는 체험 사용자에게 새로운 정보 및 지식의 습득과 같은 인지적 차원의 체험뿐만 아니라 다른 사용자와의 관계 맺기를 통해 감성적이고, 관계를 기반으로 한 체험 역시 느끼게 해준다고 논하고 있으며, 그는 총체적 체험을 보는 접근 방법을 감각, 인지, 감성, 행동, 관계의 5가지 유형으로 정의하였다.

앞서 모바일 관광 앱의 서비스 특성을 사용하여 소비자 만족 및 행동의도에 미치는 영향에 대해 설문조사를 통해 분석한 결과, 사용자 간의 의사소통이 용이할수록 사용자의 만족도가 높은 것으로 나타났다(조성호, 2019). 이외에도 랜선투어를 온라인 상에서 이루어지는 관광활동으로 정의하

여 설문조사법을 수행한 연구에서는 관광자의 지각된 가치는 랜선투어의 지불의도에 대해 정의의 유익한 영향을 미치는 것으로 나타났다(홍순기 외, 2022).

이처럼 최근 들어 온라인 체험에 대한 다양한 연구가 수행되고 있지만, 주로 체험 사용자에 의해 생성된 단방향적인 데이터만을 사용하여 고객 만족에 미치는 주요 변수로서 연구가 수행되었다. 그러나 체험 제공자가 체험 사용자에게 미치는 영향에 대해 고려할 때(Schmitt, 1999), 본 연구에서는 체험 제공자에 의해 생성된 데이터를 포함한 양방향적인 데이터를 주요 변수로서 활용하여 고객 만족에 미치는 영향을 분석하였다.

또한, 에어비엔비의 온라인 체험 서비스는 일반적인 랜선투어와 비교하여 크게 두가지의 차이점이 존재한다.

첫 번째로, 에어비엔비에서는 오프라인에서 진행되던 그룹 투어 방식으로 진행되기 때문에, 호스트와 게스트 간의 양방향 뿐만 아니라, 사전 예약된 게스트간에도 소속감과 함께 상호작용이 일어날 수 있는 것이다. 따라서 에어비엔비의 서비스는 다른 단순 랜선투어와 다르게, 서비스 공급자(호스트)가 제공하는 정보뿐만 아니라, 다른 사용자(게스트)들의 느낌과 상호작용을 통해 제공하는 정보가 만족도에 영향을 미치는 효과가 더 명확할 것으로 판단된다.

두 번째로, 현지 호스트 즉, 비전문가에 의해 구성되는 에어비엔비 온라인 체험 서비스는 전문가에 의해 구성되는 일반적인 랜선투어와 큰 차이점을 가진다. 일반적인 랜선투어 서비스는 전문가에 의해 정형화되고 객관적인 구조의 체험 소개를 제공받게 되나, 에어비엔비 온라인체험은 현지 호스트의 성향에 훨씬 크게 영향을 받고, 비정형화된 주관적인 구조의 체험 소개를 제공받게 된다. 따라서 본 연구를 통해 전문가 위주의 일반적인 랜선투어에서는 확인할 수 없었던, 온라인 상에서의 호스트 주관도(또는 객관도)를 주요 변수로 채택하여 연구해봄으로서 기존 연구와 차별성을 가진다.

2.3 온라인 리뷰 데이터와 고객만족에 관한 연구

온라인 리뷰 데이터는 평점과 텍스트로 된 리뷰를 포함하며, 동료가 만들어낸 웹사이트 상의 상품 평가로써 정의되며, 이러한 정보는 온라인 시장에서 소비자 구매 불확실성을 줄이고 잠재적 소비자의 구매의사결정을 돕는 지원도구로써 기능한다(Jiang and Benbasat, 2007; Metzger, 2007; Mudambi and Schuff, 2010). 더 많은 양의 정보는 소비자에게 더 많은 설득력을 가지게 되고, 특히 마케터가 아닌 사람으로부터 전달된 정보는 소비자로서 하여금 더욱 믿음을 가지도록 하기 때문에 온라인 리뷰 데이터가 고객 만족에 미치는 영향은 중요하다(Herr et al., 1991; Schwenk, 1985).

과거 기존의 연구들은 대체로 온라인 리뷰 데이터의 정량적 데이터인 리뷰의 수와 평점을 기반으로 진행되었다. Matfield(2011)는 리뷰의 양이 매출과 수익에 긍정적인 관계가 있다는 것을 데이터를 통해 확인하였고, 높은 평점은 영화의 매출 및 도서 판매량 뿐만 아니라 온라인 호텔 매출과 예약에도 긍정적인 영향을 미친다는 연구 결과가 있으며(Chintagunta et al., 2010; Kim et al., 2015), 특히 서비스 이용 후기에 대해 정량 데이터인 평점을 고객 만족으로 간주한 연구 사례가 있었으므로 본 연구에서도 주요 종속변수로 활용하였다.

평점(Rating)은 온라인 소비자로서 하여금 다른 소비자들의 만족도를 파악하고 소비를 결정할 수 있게 도와주는 역할을 한다(김홍범, 임영희; 2019). 한주형, 노은정(2019)은 온라인 호텔 리뷰 사이트인 익스피디아(Expedia)로부터 고객만족도 평점을 종속변수로 활용하여, 숙박서비스 요인이 고객만족도 평점에 유의한 결과가 있는 것을 확인하였으며, 이러한 고객만족도는 숙박시설에 따라 다르게 나타나는 것을 확인하였다. 또한 윤호민, 최규완(2021)은 고객만족을 측정하기 위해 레스토랑에서의 종합 평점(Rating)을 사용하였는데, 평점은 제한된 범위에서 점수가 매겨지므로 독립변

수의 영향을 과소 추정하지 않도록 OLS 대신 토빗 회귀분석을 활용하여 종속변수와 독립변수 간의 관계를 추정하였으므로 본 연구에서도 토빗 회귀분석을 수행하였다.

그럼에도 평점은 서비스 이용자가 실제 어떤 부분에서 만족하거나 불만족하였는지에 대한 정보를 제공하지 않기 때문에 이용자 의견에 대한 정확한 추정이 어렵고, 서비스 제공자에게 자세한 개선 방향을 전달하지 못한다는 한계점이 존재한다(임영희, 김홍범, 2019). 반면 텍스트로 된 의사 표현은 이용자에 대해 더 자세한 정보를 담고 있기 때문에 평점보다 중요할 수 있다는 연구결과가 존재하며(Schuckert et al., 2015), Hu et al.(2014)은 연구에서 평점보다 리뷰의 감성지수가 매출에 더 직접적인 영향이 있음을 확인하였다. 이에 따라 최근에는 텍스트를 수치화 할 수 있는 감성분석을 사용한 연구가 이루어지고 있으나, 대체로 아마존과 같은 온라인 상품 거래 사이트를 대상으로 하거나 혹은 호텔 서비스에 대한 연구로 아직 그 범위가 다양하지 않다.

이처럼 온라인 텍스트 리뷰는 제품 및 서비스 이용자의 만족에 대한 의사 표현을 보유하고 있으므로 평점보다 더 자세한 정보를 담고 있고, 텍스트 데이터를 기반으로 한 감성분석 결과가 고객 만족을 확인할 수 있는 유의한 변수로서 활용 가능한 것이 선행연구를 통해 확인되었다(윤호민, 최규완, 2021). 이를 바탕으로 본 연구에서는 고객 만족을 측정하기 위해 과거부터 지속적으로 쓰인 평점 데이터와 더불어 온라인 체험 서비스 이용자가 작성한 텍스트 리뷰 데이터를 감성분석을 이용해 수치 데이터로 변환하여 주요 종속변수로 활용하였다.

2.4 감성분석과 워드 클라우드에 관한 연구

감성분석은 텍스트마이닝의 한 분야로 문서에서의 긍정 또는 부정에 대한 감성을 추측하고 분류하는 방법이다. 감성분석은 오피니언 마이닝

(Opinion mining)이라고도 불리며, 온라인 리뷰 텍스트와 같은 비정형 데이터로부터 자연어 처리를 거친 후, 감성사전을 활용하여 텍스트가 가진 감정을 식별하게 된다. 감성사전은 단어들을 긍정적인 단어와 부정적인 단어로 구분해 놓은 것으로 연구의 진행을 위해 새롭게 감성사전을 구축하는 것보다는 여러 연구기관에서 미리 구축해 놓은 감성사전을 활용하여 연구를 진행하는 경우가 많다 (최자영 등, 2020). 일반적으로 감성분석은 텍스트에 쓰인 감성을 분류하고 객관적인 수치 정보로 변환할 때 쓰이는데, 좁은 의미에서는 텍스트 상의 긍정 및 부정을 분류하는 것이다. 감성분석은 보통 세가지의 단계를 거치게 되는데, 1단계는 분석에 필요한 텍스트 데이터를 수집하고, 2단계에서는 수집된 텍스트 데이터를 기반으로 사람의 주관적 의견 또는 감성 표현 부분을 추출하는 주관성을 탐지(subjectivity detection)하고, 마지막 3단계에서는 추출한 텍스트를 긍정 또는 부정으로 분류하여 극성을 탐지(polarity detection)하는 단계이다(신수정, 2014; 장경애 등, 2015). 사공원 등(2016)은 호텔 이용자의 텍스트 리뷰를 감성분석을 통해 만족도 수치로 계량화한 뒤, 호텔 유형별로 유형성, 대응성, 신뢰성, 의사소통 등의 LQI(Lodging Quality Index) 차원으로 분류해 평가 결과를 비교하였다.

워드 클라우드(Word Cloud)는 태그 클라우드(Tag Cloud)라고도 불리며, 서적이나 온라인 상의 콘텐츠로부터 얻어진 텍스트의 단어 빈도를 시각적으로 표현한 것이다(Heimerl et al., 2014). 워드 클라우드 분석을 통해 텍스트에 주로 어떤 내용이 언급 되어있는지 쉽게 파악할 수 있다. 텍스트 분석 중에서도 가장 직관적이면서도 다양하게 활용되는 분야로, 특정 문서 집합에서 많은 빈도로 출현한 단어를 다른 단어에 비해 상대적으로 크게 표현하여 해당 단어를 강조하는 대표적인 빈도 분석 기법이다(김남규 등, 2017). 워드 클라우드를 통해 전자기록관리에 대한 국제 연구 동향을 심층 분석하거나, 항생제 오용과 관련된 증거를 도출하

기 위해 트위터 데이터를 분석하는 등 다양한 범위의 연구에서 워드 클라우드가 사용되고 있다 (Scanfeld et al., 2010; 박옥남, 박희진, 2016).

2.5 토빗 회귀분석 (Tobit Regression)

종속변수의 값이 부분적으로 절단되어 있는 특성이 있다면 최소자승법(Least Square)을 이용하는 일반적인 회귀모델로 추정하는 것은 적절하지 않다. 절단된 자료를 OLS로 추정할 경우, 도출된 예측값이 관측값의 범위를 벗어나 제한된 범위 내 독립변수의 영향을 과소 추정할 우려가 있기 때문이다(Long, 1997). 이러한 경우에는 종속 변수의 상한 값과 하한 값을 주어 제한된 범위 안에서 계수를 추정하는 토빗 회귀분석을 활용해 종속변수와 독립변수와의 관계를 추정할 수 있다.

토빗 회귀분석은 상한 또는 하한 단측만 절단된 구조를 갖도록 제약을 부여하는 단측 절단모델과 상한과 하한 모두 절단하는 양측 절단 모델이 있다.

양측 절단된 토빗 모델의 일반적인 형태는 아래의 식 (1)과 같다.

$$y_i = \beta_0 + \sum_{k=1}^K \beta_k x_{ki} + \epsilon_i \quad (1)$$

$$s.t. \ y_i = \begin{cases} U_i, & \text{if } U_i \leq y_i^* \\ y_i^*, & \text{if } L_i \leq y_i^* \leq U_i \\ L_i, & \text{if } y_i^* \leq L_i \end{cases} \quad (2)$$

U_i = 상한값, L_i = 하한값

식 (1)은 일반적인 OLS 회귀모델의 추정방식과 동일하지만, 토빗 회귀분석은 제약조건인 식(2)가 주어지는 점에서 차이가 있다. 식(2)는 종속변수가 갖는 값을 구조적으로 제한함으로써 모델의 추정을 제한 범위 내에서 MLE(최대우도법)으로 추정한다(Osgood et al., 2002).

윤호민, 최규완(2021)은 레스토랑 선택 속성과 만족도에 관한 연구를 진행하였는데 고객이 레스

토랑에 부여한 전체 종합 평점 값을 고객 만족으로 두었으며, 분석에 사용된 종합 평점 값이 10점에서 50점까지 양측이 절단된(Doubly censored) 형태였기 때문에 토빗 회귀분석을 수행하였다. 따라서 본 연구에서도 에어비앤비의 온라인 체험 서비스 특성이 고객 만족에 미치는 영향을 보기 위해 0~5점으로 절단된 종속변수인 평점(Rating)에 대해 토빗 회귀분석을 수행하였다.

III. 연구방법

3.1 데이터 수집 및 변수의 조작적 정의

본 연구를 진행하기 위해 에어비앤비 웹사이트의 온라인 체험 페이지를 대상으로 데이터를 수집하였다. 수집을 위해 기본 언어는 영어로 설정하였으며, 통화의 기준 또한 미화달러(USD)로 변경하여 수집을 진행하였다. 이는 에어비앤비 주 사용자의 사용언어, 호스트의 체험 소개글 및 게스트의 리뷰 등에 사용된 언어를 고려하였을 때 영어가 95% 이상의 비율을 차지할 뿐만 아니라, 추후 파이썬 패키지를 이용한 감성분석에서의 효율성과 타당성을 고려하여 선택하였다.

에어비앤비는 자체 검색 알고리즘을 사용하여 게스트의 관심을 끌 수 있는 검색 결과를 생성한다. 검색 알고리즘은 100개 이상의 요소를 고려하여 검색 결과에서의 체험 순위를 결정하게 되고, 검색 알고리즘은 예약으로 이어진 수백만 건의 이전 검색 결과를 바탕으로 게스트와 게스트가 원하는 체험을 연결해 주도록 설계되어 있다. 그 과정에서 게스트에게 한 번에 표출되는 체험의 개수는 최대 1,000개를 넘지 않도록 제한되어 있으므로 따라서 한 번에 수집 가능한 온라인 체험의 목록 개수는 최대 999개였다. 에어비앤비의 검색 알고리즘을 통해 반환된 수백 개의 온라인 체험은 대표 사진, 평점, 체험 제목, 가격과 함께 게스트에게 일차적으로 노출된다. 게스트에게 노출된 목록에서 하나를 선택하면 각 온라인 체험에 대한 세부

정보를 확인할 수 있으며, 이러한 데이터를 모두 수집하는 웹 크롤러(Web crawler)를 작성하여 데이터를 수집하였다. 웹 크롤러를 위해 사용된 파이썬 패키지는 Python 3.8, Selenium 3.141, 그리고 BeautifulSoup 4.6 등이다.

수집된 전체 온라인 체험의 개수는 모두 999개였으나 체험 제목을 기준으로 중복된 온라인 체험을 제거하고 난 뒤의 최종 데이터 개수는 885개였으며, 이를 바탕으로 정형데이터와 비정형 데이터의 전처리를 진행하였다. 감성분석에 앞서 텍스트 전처리가 필요한 데이터는 호스트가 작성한 체험 소개글과 호스트 소개글, 그리고 게스트가 작성한 텍스트 리뷰로써 총 세 개의 칼럼이었다. 각 영어 텍스트에 대해 먼저 모두 소문자로 변환하고 난 뒤, 띄어쓰기를 기준으로 모든 단어를 토큰화(Tokenizing)하였다. 다음으로 평가 점수에 영향을 미치지 않는 불용어(Stopwords)를 제거하고, 단·복수형의 단어를 하나로 정리해주기 위해 표제어추출(Lemmatization)을 수행하였다.

통계변수로 정의한 온라인 체험의 카테고리는 에어비앤비의 분류에 따라 Great for groups, Family friendly, Animals, Arts&writing, Baking, Cooking, Dance, Drinks, Entertainment, Fitness, History&Culture, Magic, Music, Social impact, Wellness, Olympians & Paralympians, Designed for accessibility와 같이 총 17개의 카테고리로 확인되었으며, 해당 카테고리에 속하면 1, 속하지 않으면 0인 더미변수로 변환하여 연구에 사용하였다. 이 과정에서 하나의 온라인 체험이 여러 카테고리에 중복으로 속해 있음이 확인되었다. 이러한 현상은 다중공선성 문제를 야기하여 분석에 영향을 미칠 수 있으므로, 상관관계 분석을 통해 총 9개의 카테고리로 축소(Great for groups, Cooking, Dance, Entertainment, Fitness, History & Culture, Social impact, Wellness, Olympians & Paralympians)하여 연구에 사용하였다.

본 연구를 위해 사용된 변수는 총 21개로 그 조작적 정의는 다음의 <표 1>에서 나열하였다.

〈표 1〉 변수의 조작적 정의

구분	Variable(ENG)	Variable(KOR)	Definition
User-oriented	Review Polarity	리뷰 감성	각 온라인 체험에 쓰인 리뷰의 전반적인 감성점수 [-1, 1]
	Rating	평점	각 온라인 체험이 가진 종합 평점
	Review Number	리뷰 수	각 온라인 체험에 달린 리뷰의 수
	Review Depth	리뷰 깊이	게스트가 온라인 체험 이후 작성한 리뷰 길이의 평균
Provider-oriented	Experience Subjectivity	체험소개글 주관도	각 온라인 체험에 대한 소개글의 주관도 [0, 1]
	Host Subjectivity	호스트소개글 주관도	각 온라인 체험의 호스트에 대한 소개글의 주관도 [0, 1]
	Price	가격	각 온라인 체험의 가격 (단위: USD)
	Runtime	진행시간	각 온라인 체험이 진행되는 시간 (단위: 분)
	Language	사용언어 수	각 온라인 체험의 호스트가 사용 가능한 언어의 수
	Photo	사진 수	각 온라인 체험을 소개하기 위해 포스팅 되어있는 사진의 수
	Hosting Years	호스팅 근속연수	각 온라인 체험의 호스트가 호스팅을 개시한 이후로의 근속 연수
	Co-host	공동호스트 수	각 온라인 체험에서의 공동 호스트 인원 수
Category	카테고리	온라인 체험이 속한 각각의 카테고리 변수	

주요 독립변수는 정보 생성의 주체가 누구인지를 기준으로 서비스 이용자(User-oriented)와 서비스 제공자(Provider-oriented) 두 가지 관점에서 구분하여 정의하였다. 정보 생성의 주체가 서비스 이용자인 경우로 정의된 변수는 리뷰 감성, 평점, 리뷰 수, 리뷰 깊이가 있다. 여기서 리뷰 깊이는 리뷰길이를 리뷰수로 나눈 평균값을 사용하였으며, Mudambi and Schuff(2010)가 온라인 리뷰의 지각된 유용성에 미치는 영향을 연구하기 위해 리뷰 깊이를 단어 수로 측정된 방법에서 차용하였다.

정보 생성의 주체가 온라인 체험 서비스 제공자인 경우의 독립변수에는 체험소개글 주관도, 호스트소개글 주관도, 가격, 진행시간, 사용언어 수, 사진 수, 호스팅 근속연수, 공동 호스트 수가 있다. 통제변수는 각 온라인 체험이 속한 카테고리이며 각각 Great for groups, Cooking, Dance, Entertainment, Fitness, History & Culture, Social impact, Wellness, Olympians & Paralympians로 구성되었다.

3.2 감성분석과 워드클라우드

본 연구의 감성분석(Sentiment analysis)을 위해 MIT 라이선스 하에 운영되는 파이썬 라이브러리 중 하나인 TextBlob을 사용하였다. TextBlob은 어휘 사전 기반 방식으로 수많은 어휘 자원(Lexical resources)에 쉽게 접근할 수 있도록 해주는 NLTK를 사용하며, 이는 사용자가 자연어의 분류 및 기타 여러 작업을 수행할 수 있도록 지원한다. 일반적으로 감성분석을 수행할 때는 부정 및 긍정 단어들로 분류된 사전의 선(先)정의가 필요하다. TextBlob은 NLTK 라이브러리에 구축되어 있는 감성사전을 바탕으로 극성(Polarity)과 주관도(Subjectivity)의 값을 반환하는데, 극성은 [-1, 1] 사이의 값을 반환하며, -1에 가까울수록 부정적인 감성을 나타내고 1에 가까울수록 긍정적인 감성을 나타낸다. 주관도는 [0, 1] 사이의 값을 반환하며, 0에 가까울수록 텍스트의 내용이 객관적임을 나타내고, 1에 가까울수록 텍스트의 내용이 주관적임을 나타낸다.

수집된 호스트 및 체험 소개글, 그리고 온라인

체험 리뷰는 텍스트 형태의 데이터로 저장되어 있으므로 회귀분석에 앞서 감성분석을 통해 수치 값으로 변환해주는 과정이 필수적이다. 온라인 체험 소개글과 호스트의 자기 소개글의 경우에는 일반적인 사회통념 상, 홍보를 위한 소개글에 부정적인 내용의 글을 쓸 이유가 없으므로 TextBlob을 이용한 감성분석 결과에서 주관도 점수만 차용하여 독립변수로 사용하였으며, 온라인 이용자의 텍스트 리뷰는 전처리를 끝낸 뒤 전체 텍스트를 병합하여 각 체험에 대한 전반적인 감성을 나타내는 극성 점수만 차용하여 일반 선형 회귀분석에서의 종속변수로 사용하였다.

워드 클라우드(Word Cloud)는 텍스트 분석 중에서도 가장 직관적이면서도 다양하게 활용되는 분야로, 특정 문서 집합에서 많은 빈도로 출현한 단어를 다른 단어에 비해 상대적으로 크게 표현하여 해당 단어를 강조하는 대표적인 빈도 분석 기법이다. 따라서 본 연구에서도 에어비앤비 온라인 체험 사용자가 작성한 텍스트 리뷰 데이터의 전반적인 방향과 단어 목록의 시각화를 위해 워드 클라우드를 통한 빈도분석을 수행하였다.

3.3 일반선형회귀분석과 토빗회귀분석

본 연구는 감성분석을 이용하여 온라인 체험에서 제공되는 서비스 특성이 고객 만족에 미치는 영향을 분석한 것이다. 고객 만족을 보기 위한 변수로는 선행연구를 근거로 평점과 리뷰감성 두가지로 설정하였으며, 종속변수 데이터의 특성을 고려하여 각 변수에 대해 일반선형 회귀분석 및 토빗 회귀분석을 수행하였다.

먼저 식 (3)은 종속변수를 리뷰감성(Review Polarity)으로 두어 일반선형 회귀모형을 설계한 것이다. 여기서 리뷰감성과 리뷰의 수, 그리고 리뷰 깊이는 서비스 이용자(User-oriented)에 의해 생성된 변수로 분류하였다. 앞의 세 가지를 제외한 온라인 체험 소개글 주관도, 호스트 소개글 주관도, 가격, 진행시간, 사용언어 수, 사진 수, 호스팅

근속연수, 공동 호스트 수, 카테고리는 서비스 제공자(Provider-oriented)에 의해 생성된 변수로 분류하였다.

$$\begin{aligned} Review \text{ Polarity} = & \beta_0 + \beta_1 Rating & (3) \\ & + \beta_2 Review \text{ Number} + \beta_3 Review \text{ Depth} \\ & + \beta_4 Experience \text{ Subjectivity} \\ & + \beta_5 Host \text{ Subjectivity} + \beta_6 Price \\ & + \beta_7 Runtime + \beta_8 Language + \beta_9 Photo \\ & + \beta_{10} Hosting \text{ Years} + \beta_{11} Co_host + \beta_{12} \end{aligned}$$

다음으로 아래의 식 (4)는 종속변수를 평점(Rating)으로 두어 토빗 회귀분석(Tobit Regression)을 적용하였다. 종속변수에 쓰인 평점은 값의 범위가 0점에서 5점 사이로 그 범위가 절단되어 있기 때문에 일반선형 회귀모형에서 쓰이는 최소승법으로 계수를 추정하게 되면 독립변수의 영향을 과소 추정할 우려가 있다(Long, 1997). 따라서 양측 절단 모형을 적용한 토빗 회귀모형으로 분석을 수행하였다.

$$\begin{aligned} Rating = & \beta_0 + \beta_1 Review \text{ Polarity} & (4) \\ & + \beta_2 Review \text{ Number} + \beta_3 Review \text{ Depth} \\ & + \beta_4 Experience \text{ Subjectivity} \\ & + \beta_5 Host \text{ Subjectivity} \\ & + \beta_6 Price + \beta_7 Runtime \\ & + \beta_8 Language + \beta_9 Photo \\ & + \beta_{10} Hosting \text{ Years} + \beta_{11} Co_host \\ & + \beta_{12} Category + \epsilon \end{aligned}$$

앞서 만든 식 (3)과 마찬가지로 평점, 리뷰 수, 리뷰 깊이는 서비스 이용자(User-oriented)에 의해 생성된 변수로 분류되고, 온라인 체험 소개글 주관도, 호스트소개글 주관도, 가격, 진행시간, 사용언어 수, 사진 수, 호스팅 근속연수, 공동 호스트 수, 카테고리는 서비스 제공자(Provider-oriented)로 분류하였다.

IV. 분석 결과

4.1 기초통계량

본 연구를 위해 에어비앤비 웹사이트의 온라인 체험 페이지에 접근할 수 있도록 Python을 기반으로 한 웹 크롤러를 제작하여 수치 데이터 및 텍스트 데이터를 수집하였다.

<표 2>는 각 변수에 대한 기초통계량으로 각각의 관측치, 평균, 표준편차, 최소값, 최대값을 나열하였다. 전처리를 거쳐 최종적으로 분석에 쓰인 데이터는 총 885개이며 누락된 값은 없었다. 에어비앤비의 온라인 텍스트 리뷰 감성은 평균이 0.31로 대체로 긍정적인 감성을 보이고 있으며, 평균 평점 또한 4.21로 텍스트 리뷰 감성과 마찬가지로 높은 값을 보여주었다. 각 온라인 체험에 등록된 리뷰의 평균 개수는 약 136개였으며 최소값은 0

개, 최대값은 3,661개로 그 격차로 인해 리뷰 수의 표준편차는 330.93으로 큰 값을 보였다. 리뷰의 깊이는 각 온라인 체험에서 쓰인 리뷰 내의 전체 단어 개수를 구한 뒤, 이를 리뷰 개수로 나눈 평균값으로 변환하여 분석에 사용하였다.

호스트가 작성한 온라인 체험 소개글의 주관도와 호스트에 대한 소개글 주관도에 대한 평균과 표준편차는 각각 0.50(0.12)과 0.41(0.13)을 보여주며 고른 분포를 보였다. 또한 온라인 체험에 대한 가격은 평균 USD \$ 22.81이며, 체험의 진행시간은 평균 69분으로 한 시간 내외의 수준에서 진행되는 것을 확인할 수 있었다. 호스트가 사용하는 언어의 개수는 평균 1.64개 정도이며, 온라인 체험의 소개를 위한 사진 게시는 약 9장이고, 오프라인 체험에서 시작하여 온라인 체험으로 넘어온 호스트를 고려했을 때 평균 호스팅 경력은 5.6년 정도인 것으로 확인되었다.

<표 2> 기초 통계량

Variable	Obs.	Mean	Std. Dev.	Min	Max
Review Polarity	885	0.31	0.18	-0.01	0.80
Rating	885	4.21	1.78	0	5
Review Number	885	136.19	330.93	0	3661
Review Depth	885	26.24	19.54	0	599
Experience Subjectivity	885	0.50	0.12	0	1
Host Subjectivity	885	0.41	0.13	0.03	0.77
Price	885	22.81	21.81	0	300
Runtime	885	69.93	24.90	30	390
Language	885	1.64	1.34	1	18
Photo	885	8.96	3.65	4	47
Hosting Years	885	5.59	2.74	1	12
Co-host	885	0.33	0.73	0	3
Great for groups	885	0.25	0.44	0	1
Social impact	885	0.05	0.21	0	1
History and culture	885	0.46	0.50	0	1
Cooking	885	0.19	0.39	0	1
Dance	885	0.02	0.16	0	1
Entertainment	885	0.09	0.29	0	1
Fitness	885	0.01	0.09	0	1
Wellness	885	0.16	0.37	0	1
Olympians and Paralympians	885	0.01	0.03	0	1

4.2 상관관계 분석

본 연구에 사용된 변수들 간의 상관관계는 다음의 <표 3>과 같다. 이를 통해 종속변수인 리뷰감성과 독립변수인 평점, 리뷰 수, 리뷰 깊이, 온라인 체험 소개글 주관도, 진행시간, 호스트가 사용하는 언어의 수, 체험 소개 사진의 수, 호스팅 근속연수와 동료 호스트의 수가 양의 상관관계가 있음이

확인되었다. 그러나 리뷰감성과 가격 사이에는 음의 상관관계가 있는 것으로 확인되었다.

독립변수 간의 상관관계도 확인되었는데, 평점과 리뷰 깊이 사이에 0.57의 다소 높은 상관관계가 존재하였으나, 다중공선성의 문제가 발생할 정도는 아니므로 회귀분석 후에 VIF를 확인하여 연구 결과의 타당성을 확보하였다.

또한 통제변수로 쓰인 카테고리 변수 간의 상관

<표 3> 주요변수간 상관관계 분석

Variable	1	2	3	4	5	6	7	8	9	10	11	12
Review Polarity	1											
Rating	0.75***	1										
Review Number	0.19***	0.18***	1									
Review Depth	0.53***	0.57***	-0.24***	1								
Experience Subjectivity	0.17***	0.12***	0.04	0.03	1							
Host Subjectivity	0.03	0.05	0.01	0.05	0.17***	1						
Price	-0.09**	-0.09***	-0.01	-0.06*	0.03	-0.01	1					
Runtime	0.00*	-0.05	0.06*	-0.02	-0.02	0.01	0.13***	1				
Language	0.12**	0.11***	0.09***	0.03	0.05	0.07*	-0.03	-0.05	1			
Photo	0.28***	0.28***	0.22***	0.14***	0.09***	0.10***	-0.06*	0.06*	0.04	1		
Hosting Years	0.07**	0.08**	0.11***	0.06*	0.03	0.01	0.08**	0.03	0.12***	0.13***	1	
Co_host	0.13***	0.13***	0.32***	-0.04	0.02	0.09**	-0.06*	0.01	0.12***	0.14***	-0.01	1

<표 4> 카테고리 간 상관관계 분석

Variable	1	2	3	4	5	6	7	8	9
Great for groups	1								
Social Impacts	0.11***	1							
History and Culture	-0.07**	-0.03	1						
Cooking	-0.15***	-0.11**	-0.36***	1					
Dance	0.06*	0.07**	0.03	-0.06*	1				
Entertainment	0.06*	-0.05*	-0.17***	-0.11***	0.08**	1			
Fitness	0.01	0.04	-0.08**	-0.04	0.07**	0.06*	1		
Wellness	-0.05	-0.04	-0.28***	-0.18***	-0.03	-0.07**	0.03	1	
Olympians and Paralympians	-0.02	-0.01	-0.03	-0.02	-0.01	-0.01	0.00	-0.01	1

관계 분석결과는 <표 4>와 같으며, 각 테고리 간의 상관계수를 확인하였을 때 유의하게 높은 수치는 존재하지 않았고, 주요변수와 카테고리 간의 상관관계 분석에서도 특이점은 존재하지 않았다.

4.3 워드 클라우드

워드 클라우드는 단어의 빈도분석 및 시각화를 통해 데이터에 대한 초기 인사이트를 얻기 위해 사용된다. 따라서 에어비앤비의 온라인 체험을 경험한 서비스 이용자들이 느낀 전반적인 느낌을 표현하기 위해 주로 사용한 단어와 그 빈도를 파악해보고자 분석을 진행하였다. 워드 클라우드는 소문자화, 불용어의 제거 및 표제어 추출의 전처리 과정을 마친 텍스트 데이터를 사용하였다. 그러나 텍스트 데이터의 전처리 과정에서 표제어 추출을 거쳤음에도 불구하고 워드 클라우드를 통해 확인된 단어를 살펴보면 ‘enjoyed’, ‘amazing’, ‘learned’ 등과 같이 그 단어의 원형으로 변환되지 않은 단어들 존재하였다. 그러나 이러한 결과는 정해진 규칙에 의해 단어의 어미를 제거하는 어간추출(Stemming)과는 달리 문맥을 고려하는 표제어 추출(Lemmatization)을 전처리 과정으로 채택한 것이므로 그 결과값이 감성사전에 존재하는 형태이자, 단어의 의미를 파악하기 용이한 형태로 보존된 것이므로 텍스트 전처리가 적절히 잘 이루어진 것으로 볼 수 있다.

워드 클라우드는 높은 빈도를 가진 단어들이 상대적으로 크게 표현되기 때문에 연구자가 직관적으로 이해하기에 용이하다. 따라서 텍스트의 크기가 가장 크나 중립적인 의미를 가진 ‘experience’ 등의 단어를 제외하고 눈에 띄는 단어들을 살펴봤을 때 ‘great’, ‘fun’, ‘amazing’, ‘enjoyed’, ‘thank’와 같이 긍정적인 의미를 가진 단어가 주로 확인되었다. 이는 앞서 기초 통계량에서 살펴본 텍스트 리뷰 감성의 평균값이 양수로써 긍정적인 것을 고려하였을 때 감성분석을 통한 수치화가 적절히 이루어졌음으로 볼 수 있다.



<그림 1> 온라인 텍스트 리뷰의 워드 클라우드

4.4 리뷰 감성에 대한 일반회귀분석

텍스트 리뷰는 평점과 비교하였을 때, 제품 혹은 서비스에 대해 더 자세한 의사표현과 정보를 담고 있기 때문에 평점보다 더 중요하다고 할 수 있다(Schuckert et al., 2015; 임영희, 김홍범, 2019). 따라서 텍스트 리뷰 데이터를 이용한 감성분석 결과가 단순히 평점을 이용한 고객의 반응 및 만족을 연구하는 것과 다른 의미로 주요 변수가 될 수 있으므로 첫 번째 분석에서 종속변수를 리뷰감성으로 하여 일반선형 회귀분석을 시행하였다.

<표 5>에서 주요 변수로 보고자 하였던 정보 생성 주체가 온라인 서비스 제공자(Provider-oriented)인 경우의 감성분석 결과가 유의하게 나타났는데, 온라인 체험 소개글에 대한 주관도는 유의한 정의 관계($B=0.095^{***}, p<.01$)를 가지며, 호스트의 자기 소개글에 대한 주관도는 리뷰감성과 유의한 부의 관계($B=-0.058^{**}, p<.05$)를 가지는 것으로 확인되었다. 이 같은 결과는 온라인 체험에 대한 호스트의 소개글이 주관적이고 친근하게 쓰일수록 고객 만족에 긍정적인 영향을 가지는 반면, 호스트 자신에 대한 소개글은 객관적이고 공정하게 쓰일수록 고객 만족에 긍정적인 영향을 가지는 것이라고 볼 수 있다.

정보 생성의 주체가 온라인 체험 이용자(User-oriented)인 경우는 앞의 분석결과와 마찬가지로 모두 유의한 정의 관계를 가지고 있음이 확인되었다. 리뷰 평점($B=0.054^{***}, p<.01$)을 비롯하여 리뷰

수(B=0.001^{***}, $p<.01$)와 리뷰 깊이(B=0.002^{***}, $p<.01$) 모두 각각 유의한 정의 관계인 것으로 나타났다. 이는 공유경제 플랫폼에서 같은 이용자들의 참여로 인해 생성된 정보가 다른 이용자의 만족감에 매우 긍정적인 영향을 가지고 있는 것을 실증적으로 확인하였다.

다만 가격과 관련해서는 분석결과가 유의하지 않았으며, 진행시간, 호스팅 근속연수, 동료 호스트의 수 또한 유의하지 않았다. 반대로 호스트가 사용할 가능한 언어의 수(B=0.005^{*}, $p<.1$)와 온라인 체험에 업로드 된 사진의 수(B=0.002^{**}, $p<.05$)에서는 각각 정의 유의한 관계가 있음이 확인되었다. 카테고리 변수에서는 Great for groups(B=0.030^{***}, $p<.01$), Social impacts(B=0.038^{**}, $p<.05$), Cooking(B=0.033^{***},

$p<.01$), Entertainment(B=0.033^{**}, $p<.05$)에서 각각 유의한 정의 관계가 확인되었다.

4.5 평점에 대한 토빗회귀분석

서비스 이용 후기에 대해 온라인 상의 정보 공유가 활발해지면서 정량적 데이터인 평점을 고객 만족으로 간주한 연구도 여럿 수행되었다. 윤호민, 최규완(2021)은 고객만족을 측정하기 위한 종속변수로 레스토랑에서의 종합 평점(Rating)을 사용하였는데, 평점은 제한된 범위에서 점수가 매겨지므로 독립변수의 영향을 과소 추정하지 않도록 OLS 대신 토빗 회귀분석을 활용해 종속변수와 독립변수 간의 관계를 추정하였다.

<표 5> 리뷰 감성에 대한 일반선형 회귀모형 추정 결과

DV I: Review Polarity		Coef.	Std. Err.	t	P> t
User-oriented	Rating	0.054^{***}	0.003	19.01	0.000
	Review Number	0.001^{***}	0.000	3.38	0.001
	Review Depth	0.002^{***}	0.000	8.14	0.000
Provider-oriented	Experience Subjectivity	0.095^{***}	0.032	3.00	0.003
	Host Subjectivity	-0.058^{**}	0.028	-2.03	0.043
	Price	0.001	0.001	-0.55	0.579
	Runtime	0.001	0.001	0.78	0.437
	Language	0.005[*]	0.003	1.67	0.095
	Photo	0.002^{**}	0.001	2.00	0.046
	Hosting Years	-0.001	0.001	-0.96	0.338
	Co-host	0.001	0.005	0.15	0.879
	Great for groups	0.030^{***}	0.010	2.92	0.004
	Social Impacts	0.038^{**}	0.018	2.11	0.035
	History and Culture	-0.009	0.009	-0.99	0.321
	Cooking	0.033^{***}	0.012	2.83	0.005
	Dance	-0.002	0.024	-0.07	0.941
	Entertainment	0.033^{**}	0.014	2.33	0.020
	Fitness	0.043	0.042	1.02	0.309
	Wellness	-0.018	0.012	-1.49	0.138
	Olympians and Paralympians	-0.014	0.110	-0.13	0.897
Constant	-0.038	0.026	-1.46	0.144	

* $p<.1$, ** $p<.05$, *** $p<.01$ Observation = 885, Adjusted $R^2 = 0.614$, $F = 71.43$ ***, $R^2 = 0.623$

따라서 고객 만족을 평점(rating)으로 측정할 때, 정보 생성의 주체를 서비스 이용자(User-oriented)와 서비스 제공자(Provider-oriented) 두 가지 관점으로 나누어 정의하였고, 분석 결과는 <표 6>과 같다. 전체 모델이 가지는 설명력은 0.674이고, 정보 생성의 주체자가 서비스 이용자인 독립변수의 경우에는 모두 유의하게 나타났다. 이는 리뷰 감성이 긍정적일수록 평점에도 유의한 정의 관계가 있으며(B=1.233^{***}, p<.01), 동일하게 리뷰 수(B=0.001^{***}, p<.01)와 리뷰 깊이(B=0.006^{***}, p<.01)에 대해서도 각각 유의한 정의 관계가 있음이 확인되었다. 이는 리뷰 수가 많아질수록 그리고 그 리뷰가 길게 쓰일수록 이러한 정보를 이용하는 고객의 만족도에 통계적으로 유의한 정의 관계가 있는 것으로 확인되었다. 반면에 정보 제공의 주체

가 서비스 제공자인 호스트의 경우에는 온라인 체험 소개글과 호스트 소개글에 대한 감성분석의 결과가 유의하게 나타나지 않았다. 다만, 온라인 체험의 진행시간이 유의한 부의 관계로 나타났으며(-0.001^{**}, p<.05), 호스트가 온라인 체험 소개를 위해 업로드한 사진의 수에서는 유의한 정의 관계가 확인되었다(0.005^{**}, p<.1).

4.6 추가 분석: 로지스틱 회귀분석

앞서 진행한 일반선형 회귀분석에 추가하여, 리뷰 감성(Review Polarity)에 대해 추가적인 로지스틱 회귀분석을 진행하였다. 리뷰 감성의 값은 사전을 기반으로 측정이 되는데, 아주 작은 양적 점수 차이가 반드시 만족도의 차이를 표현하고 있

<표 6> 평점에 대한 토빗 회귀모형 추정 결과

DV II : Rating		Coef.	Std. Err.	t	P> t
User-oriented	Review Polarity	1.233^{***}	0.070	17.59	0.000
	Review Number	0.001^{***}	0.001	4.02	0.000
	Review Depth	0.006^{***}	0.001	10.27	0.000
Provider-oriented	Experience Subjectivity	0.053	0.079	0.67	0.501
	Host Subjectivity	0.048	0.070	0.69	0.491
	Price	0.001	0.001	-0.60	0.549
	Runtime	-0.001^{**}	0.000	-2.28	0.023
	Language	0.001	0.007	0.15	0.882
	Photo	0.005[*]	0.003	1.96	0.050
	Hosting Years	0.001	0.003	0.31	0.755
	Co-host	0.013	0.013	1.02	0.307
	Great for groups	0.034	0.024	1.38	0.169
	Social Impacts	0.003	0.043	0.06	0.954
	History and Culture	0.015	0.023	0.66	0.511
	Cooking	0.004	0.029	0.13	0.893
	Dance	0.024	0.058	0.42	0.674
	Entertainment	-0.002	0.034	-0.05	0.958
	Fitness	-0.013	0.101	-0.12	0.901
	Wellness	-0.002	0.029	-0.08	0.940
	Olympians and Paralympians	-1.413	36.044	-0.04	0.969
	Constant	4.179	0.064	65.33	0.000

* p<.1, ** p<.05, *** p<.01, Observation = 885, Pseudo R² = 0.674, Log likelihood = -194.84

는지에 의문을 가질 수 있다. 그러나 하위 그룹의 값과 상위 그룹의 값의 큰 차이는 만족도의 질적 차이를 일정 정도 표현한다고 볼 수 있을 것이다. 이에, 전체 데이터에 대해 종속변수인 review polarity 값을 기준으로 재정렬한 뒤, 세 구간으로 나누어 가운데 구간을 제거한 데이터를 사용하였다. 좌측 구간은 상대적으로 감성점수가 낮은 부분으로 더미변수 0으로, 우측 구간은 상대적으로 감성점수가 높은 부분으로 더미변수 1로 설정하였다.

분석결과는 다음의 <표 7>과 같으며, 전반적인 결과는 앞서 수행한 회귀분석과 유사하였다. 서비스 이용자(User-oriented)에 속하는 평점, 리뷰 수, 리뷰의 깊이는 높은 값을 가질수록 리뷰 감성이 긍정적으로 높은 값을 가질 확률이 높았다. 또한 서비스 제공자(Provider-oriented)에서도 마찬가지로 호스트가 작성한 체험 소개글에 대해서는 정의

유의함($B=3.400^{***}, p<.01$)을, 그리고 서비스 제공자의 자기 소개글에 대해서는 부의 유의함($B=-2.119^{**}, p<.05$)을 확인할 수 있었다. 이는 앞선 결과와 같이 체험 소개글은 주관적으로 기술될수록 고객 만족에 긍정적인 영향을, 그리고 자기 소개글은 객관적으로 기술될수록 고객 만족에 긍정적인 영향을 미치는 것으로 일관적인 결과를 확인할 수 있었다.

V. 결 론

5.1 연구결과 요약

본 연구는 감성분석을 이용하여 온라인 체험에서 제공되는 서비스 특성이 고객 만족을 평점과 리뷰감성으로 두었을 때 각각에 어떠한 영향이 있

<표 7> 리뷰 감성에 대한 로지스틱 회귀모형 추정 결과

DV I: Review Polarity		Odds Ratio	Coef.	Std. Err.	z
User-oriented	Rating	3.65	1.295^{**}	0.502	2.58
	Review Number	1.00	0.001^{**}	0.001	2.07
	Review Depth	1.03	0.027^{***}	0.007	3.63
Provider-oriented	Experience Subjectivity	29.96	3.400^{***}	1.033	3.29
	Host Subjectivity	0.12	-2.119^{**}	0.943	-2.25
	Price	0.99	-0.009	0.007	-1.15
	Runtime	1.01	0.007	0.006	1.22
	Language	1.24	0.219[*]	0.117	1.88
	Photo	1.03	0.030	0.036	0.85
	Hosting Years	0.96	-0.046	0.044	-1.04
	Co-host	0.84	-0.172	0.175	-0.98
	Great for groups	1.81	0.591[*]	0.331	1.79
	Social Impacts	2.67	0.983[*]	0.565	1.74
	History and Culture	0.53	-0.627^{**}	0.293	-2.14
	Cooking	2.07	0.726[*]	0.381	1.90
	Dance	0.93	-0.068	0.810	-0.08
	Entertainment	2.67	0.982^{**}	0.489	2.01
	Wellness	0.44	-0.818^{**}	0.405	-2.02
Constant	0.00	-8.236^{***}	2.624	-3.14	

* $p<.1$, ** $p<.05$, *** $p<.01$, Observation = 585, Pseudo $R^2 = 0.419$, Log likelihood = -235.65

는지에 대해 연구하고자 에어비앤비 웹페이지로부터 데이터를 수집하였다. 텍스트로 수집된 비정형 데이터의 경우에는 텍스트 마이닝에서 감성분석을 할 때 쓰이는 3단계 방법을 적용하여 전처리, 주관성 탐지, 극성 탐지(최자영 등, 2020)를 거쳐 정량적인 수치 데이터로 변환하여 분석에 사용하였다.

종속변수 중 하나인 리뷰감성에 나타나는 단어들의 빈도분석과 직관적인 해석을 위해 워드클라우드를 활용하였다. 중립을 나타내는 단어들을 제외하면 대부분이 ‘great’, ‘fun’, ‘amazing’, ‘enjoyed’, ‘thank’와 같은 긍정 어휘들이었음을 확인할 수 있었다. 리뷰감성에서 대부분이 긍정 어휘로 이루어진 이유는 에어비앤비의 알고리즘 특성 상 1점에서 3점 사이의 온라인 체험들은 우선적으로 노출시키지 않기 때문에 웹 크롤링으로 모아진 데이터들은 높은 평점을 가진 온라인 체험이었던 것으로 여겨진다.

온라인 구매 시 이용자는 제한적으로 노출된 정보를 바탕으로 의사결정을 해야 하므로, 상품 및 서비스를 경험한 기존 소비자들이 작성한 텍스트 리뷰와 평점에 대한 의존도가 매우 높고(Li and Hitt, 2008), 텍스트 리뷰는 소비자의 긍정과 부정을 보다 정확하게 추정 가능하다는 연구결과가 있다(Walter, 2012). 따라서 텍스트 리뷰와 평점은 이용자의 구매 의사결정 과정에서 주요한 역할을 함과 동시에 고객 만족을 나타내는 척도로써 자리매김하여 많은 연구에서 주요변수로 활용되었으므로, 본 연구에서도 주요 종속 변수로 채택하였다.

첫 번째 분석에서는 종속변수를 텍스트 기반 리뷰감성(Review Polarity)으로 두고 일반선형 회귀분석을 시행하였다. 분석결과, 정보 제공의 주체가 같은 온라인 체험 이용자(User-oriented)일 때 고객 만족에 유의한 정의 영향을 가지는 것은 일관되게 나타났다. 이는 공유경제 플랫폼에서 이용자들의 참여로 인해 생성된 정보가 다른 이용자의 만족감에 매우 긍정적인 영향을 가지고 있으며, 그렇기 때문에 추후 비즈니스 전략 수립 시 다수

의 서비스 이용자로부터 생성된 정보들이 다른 이용자에게 더 효과적으로 노출될 수 있도록 해야 될 것임을 실증적으로 확인하였다. 다만, 정보 제공의 주체가 온라인 체험 서비스 제공자(Provider-oriented)일 때는 변수 그룹에서 부분적으로 유의성이 확인되었다. 먼저, 온라인 체험 서비스 제공자가 작성한 온라인 체험 소개글의 주관도에서 유의한 정의 관계가 나타났다. 이는 온라인 체험에 대해 설명한 글은 객관적인 것 보다 주관적일수록 고객만족에 긍정적인 영향을 주었다고 해석할 수 있는데, 이 같은 결과는 이용자들이 온라인 체험에 참여하는 이유가 즐거움을 얻기 위해서라고 한다면, 온라인 체험 소개글이 보다 주관적이고 친근하게 쓰일수록 고객 만족에 긍정적인 영향을 전달하는 것이라고 볼 수 있다. 반면에 서비스 제공자의 본인 소개글에 대해서는 유의한 정의 관계가 확인되었는데, 이는 자기 자신에 대한 소개글은 객관적으로 작성될수록 고객만족에 긍정적인 영향을 준다는 것으로 볼 수 있다. 따라서 온라인 체험을 주관하는 책임자의 입장이라면, 친근하게 여겨지는 단어의 사용보다는 객관적인 단어를 사용함으로써 보다 전문적인 모습을 보여주었을 때 서비스 이용자들로 하여금 신뢰감 형성에 도움을 줄 수 있기 때문인 것으로 여겨진다.

두 번째로 평점을 종속변수로 한 토빗 회귀분석(Tobit Regression)을 시행했을 때, 선행연구와 마찬가지로 온라인 체험 서비스 이용자(User-oriented)로부터 이루어진 변수들에 대해 모두 정의 유의함이 발견되었다. 즉, 리뷰 감성이 긍정적이며, 리뷰 수가 많고 길게 작성되어 있을수록 온라인 체험 서비스 이용자가 해당 온라인 체험에 대해 가지는 만족감이 높은 것으로 나타났다. 이는 공유경제 플랫폼에서 이용자 간의 의견 공유가 고객 만족에 긍정적인 영향을 미치고 있음으로 볼 수 있다. 또한, 온라인 체험의 진행시간은 짧을수록, 그리고 온라인 체험을 소개하는 사진의 수는 많을수록 고객 만족에 긍정적인 영향을 주는 것으로 확인되었다. 다만 온라인 체험 소개 주관도 및

호스트 소개 주관도는 평점과 유의미한 관계가 확인되지 않음으로써, 리뷰감성과는 다른 결과가 나타났다. 리뷰감성과 평점에 미치는 주관도의 차별적 효과에 대해서는 추가적 데이터 확보 및 분석을 통하여 보다 구체화된 설명이 필요할 것으로 판단된다.

그리고 두 분석 모두에서 가격, 호스팅 근속연수, 동료 호스트의 수에 대해서는 유의성이 나타나지 않았는데, 특히 가격에 대해서는 사용자가 실제 오프라인 체험에 참여할 때에 드는 시간과 비용에 비해 온라인 체험을 선택함으로써 이미 상당한 비용을 절감하였으므로 상대적으로 고려의 대상이 되지 않은 것으로 보인다. 이러한 결과를 바탕으로 차후 온라인 체험과 같은 비대면 서비스를 제공하는 입장에서는 객관적인 지표 뿐만 아니라, 전달하고자 하는 정보의 성격에 따라 전략적으로 문맥의 주관도를 고려하여 작성한다면 고객만족에 긍정적인 영향을 끼칠 수 있을 것으로 생각된다. 특히나 본 연구에서 사용된 감성사전은 이미 유명한 파이썬 라이브러리인 NLTK에 구현되어 있는 것을 활용한 것이므로, 차후 알고리즘을 이용한 문장 자동완성 서비스 등의 개선에 쉽게 적용가능한 한 방안이 될 것으로 기대한다.

이외에도 고객만족에 유의한 정의의 관계를 가진 변수로는 서비스 제공자가 사용가능한 언어의 수와 온라인 체험을 소개하기 위해 업로드 한 사진의 수가 있었다. 또한 고객만족에 긍정적인 영향을 끼친 카테고리는 **Great for groups, Social impacts, Cooking, Entertainment**가 있었고, 이러한 카테고리의 전반적인 공통점은 비대면 상황 속에서도 혼자보다는 여럿이 함께 즐길 수 있는 분야라는 것이다. 특히 **Social impacts**와 같은 온라인 체험은 다른 나라의 문화, 전통의상, 음식에 대해 설명 혹은 체험을 제공해주는데, 이는 코로나19와 같은 상황에서도 다른 나라의 문화를 접해보고자 하는 온라인체험을 통한 투어의 욕구가 반영된 것으로 보인다.

이와 같은 결론에서 말미암아 온라인 체험에서

의 서비스 소개가 정량적으로 수치화 된 정보도 고객 만족에 긍정적인 영향을 주지만 텍스트로 된 정성적 서비스 정보 또한 고객 만족에 긍정적인 영향을 미치고 있음을 실증적으로 확인하였다. 또한, 정보의 성격에 따라 텍스트가 지닌 주관도가 다르며 이 또한 고객 만족에 주요한 영향을 가진 것이 분석을 통해 확인되었다.

5.2 기여도

본 연구의 기여도는 이론적 기여와 실무적 기여 두 가지 관점에서 설명할 수 있다. 이론적 관점의 기여도 첫 번째로, 온라인리뷰는 eWOM와 같이 온라인 상의 소비자 간 제품 및 서비스에 관해 의견을 교환하는 행위로 정의될 수 있으며, 이는 잠재적 소비자가 구매의사결정을 내릴 때 중요한 지원 도구로써 기능한다. 그렇기 때문에 많은 선행 연구에서 온라인 리뷰 정보를 사용한 분석을 시도해왔고, 대표적으로 쓰이는 정량적 데이터는 평점이며, 최근 감성분석을 이용한 정성적 데이터인 텍스트 정보 또한 연구에 많이 활용되고 있다. 그러나 각각의 변수가 고객 만족에 대해 어떤 식으로 더 유의한 영향이 있는지 그 차이에 대해서는 아직 연구가 미진하였다. 따라서 본 연구에서는 코로나19 이후 두각을 나타내게 된 온라인 체험이라는 분야에서 평점과 리뷰감성 두가지의 종속변수를 가지고 각각 회귀분석을 시도하여 두 분석결과에 대한 차이점과 공통점을 살펴보았다.

이론적 관점의 기여도 두 번째로, 에어비앤비와 같은 C2C 공유경제 플랫폼에서 사용자와 공급자는 더욱 긴밀한 관계를 맺지만 각기 다른 입장에서 플랫폼에 접근하게 된다(신건호 등, 2017). 이로 인해 생성되는 정보의 목적과 그 정보에 대한 의존도는 다를 수 있으므로, 본 연구에서는 이를 서비스 이용자가 생성한 정보(User-oriented)와 서비스 제공자가 생성한 정보(Provider-oriented), 두 그룹으로 나누어 분석을 진행하였고, 각 그룹에 속한 변수들과 고객 만족에 대한 관계를 알아

보았다. 결과적으로 실질적인 서비스 이용자가 더 많은 정보를 생성하여 해당 플랫폼에서 공유될 수 있도록 하는 등의 비즈니스 전략 수립이 필요한 실무적 인사이트를 제시하였다.

특히, 에어비엔비의 온라인 체험 서비스는 일반적 랜선투어와 차이점이 존재한다. 에어비엔비에서는 오프라인에서 진행되던 그룹 투어 방식으로 진행되기 때문에, 호스트와 게스트 간의 양방향 뿐만 아니라, 사전 예약된 게스트간에도 소속감과 함께 상호작용이 일어날 수 있는 것이다. 따라서 에어비엔비의 서비스는 다른 단순 랜선투어와 다르게, 서비스 공급자(호스트)가 제공하는 정보뿐만 아니라, 다른 사용자(게스트)들의 느낌과 상호작용을 통해 제공하는 정보가 만족도에 영향을 미치는 효과가 더 명확할 것으로 판단된다. 에어비엔비 온라인 체험 서비스를 대상으로 하는 본 연구는 그러한 관점에서 기존의 온라인 서비스에 관한 연구와 차별성이 있다고 하겠다.

실무적 관점에서 본 연구의 기여도는 다음과 같다. 메타버스(Metaverse)가 각광받고 있는 현 시점에서 온라인에 기반한 서비스 접근성을 높이기 위해 소비자가 적극적으로 편견 없이 수용할 수 있는 환경을 조성하는 것은 아주 중요하다. 따라서 온라인에서의 경험을 접하는 소비자들이 어떠한 부분에서 심리적으로 만족감을 얻는가에 대해 그 요인을 탐구하고 어떠한 영향을 미치는가에 대해 실증적으로 분석해본 것은 앞으로의 관련 정책 방향과 사업 전략 수립에 기여할 수 있을 것으로 기대한다.

5.3 한계점 및 추후 연구 방향

본 연구를 진행함에 있어, 에어비엔비로부터 온라인체험과 관련된 전체 데이터 및 콘텐츠를 수집할 수 없었던 점이 한계로 남았다. 특히, 코로나 19 확산과 맞물려 온라인체험이 등록된 시점을 평점과 리뷰에 영향을 미치는 주요 변수로 활용하고자 하였으나, 등록 시점에 대한 데이터는 웹사이트

트에 공개되어 있지 않아 시점에 대한 고정효과 분석은 시행할 수 없었다. 또한 에어비엔비 웹사이트의 보안 설정으로 인해 서비스 이용자가 한번에 접근 가능한 최대 온라인 콘텐츠의 수는 1,000개 미만일 뿐더러, 또한 검색 알고리즘이 매번 서비스 이용자의 검색 기록(log)에 접근하여 가장 알맞은 검색결과를 일관적으로 노출시키므로 새로 고침에서 갱신되는 온라인 체험의 수가 많지 않아 더 다양한 데이터 수집이 어려웠던 점이 한계로 남았다.

또한 본 연구 결과의 해석에 있어 주의할 필요가 있는 것도 본 연구의 한계점이다. 본 연구를 통하여 평점과 리뷰 감성의 두 변수간의 상관관계를 일반회귀분석과 토빗 회귀분석을 통하여 살펴 보았다. 그런데, 두 변수의 상관관계는 서로 영향을 주고 받았다기보다는 만족도에 영향을 받는, 일종의 만족도의 부분집합과 같은 변수이기에 나타난 현상으로 해석할 수도 있다. 즉, 만족도가 올라감에 따라 두 변수는 동시에 올라가는 것일 수도 있는 것이다. 따라서, 본 연구의 상관관계 분석을 인과관계로 해석하기 어렵다는 것도 본 연구의 한계점이며, 향후에는 시계열 데이터의 확보를 통하여 인과관계 분석으로 발전시킬 필요가 있을 것이다.

마찬가지로 게스트는 어떤 체험을 예약하든 일정 수준 이상의 체험을 즐길 수 있기를 기대하므로 에어비엔비의 검색 알고리즘이 이용자에게 노출하는 온라인 체험의 평균 평점은 4.21로 상당히 높은 비율로 인기있는 온라인 체험만을 노출시켰다. 그러나 선행연구에 따르면 소비자들은 제품에 대한 단편적인 정보만 받을 때 보다 다양한 정보를 제공받았을 때 제공된 정보를 수용할 가능성이 더 높은 것으로 나타났기 때문에(Maheswaran, et al., 1990) 시스템 상 접근의 한계로 인해 더 많은 온라인 체험 데이터의 수집이 이뤄지지 않아 추후 좀 더 폭 넓은 범위에서 일반화할 수 있는 데이터를 수집한 뒤 새로이 분석이 진행되어야 할 것이다. 마지막으로 에어비엔비의 온라인 체험 서비

스는 사실상 코로나19로 인해 대면 서비스가 불가능해지자 새로이 출시된 서비스이기 때문에 코로나19가 종식되고 난 뒤에 추가적이고 지속적인 분석을 통해 사회 트렌드의 변화와 그에 상응하는 연구 결과의 방향성을 파악해볼 필요가 있을 것이다.

참고 문헌

- [1] 김남규, 이동훈, 최호창, “텍스트 분석 기술 및 활용 동향”, *한국통신학회논문지*, 제42권, 제2호, 2017, pp. 471-492.
- [2] 김주락, “포스트코로나 관광에서 온라인 여행 플랫폼 주민 호스트의 역할에 대한 탐색적 연구”, *한국지역지리학회지*, 제27권, 제3호, 2021, pp. 372-386.
- [3] 김홍범, 임영희, “데이터마이닝을 활용한 외국인 호텔 고객의 온라인 평점 분석에 관한 연구”, *호텔경영학연구*, 제28권, 제4호, 2019, pp. 159-176.
- [4] 박옥남, 박희진, “전자기록관리에 대한 국제 연구 동향 분석”, *한국기록관리학회지*, 제16권, 제1호, 2016, pp. 89-120.
- [5] 사공원, 하성호, 박경배, “온라인 후기에 내재된 고객의 감성분석과 LQI 차원별 호텔서비스 품질 평가”, *정보시스템연구*, 제25권, 제3호, 2016, pp. 217-245.
- [6] 신건호, 박규홍, 박용진, 안재현, “C2C 공유경제 서비스 참여자 간의 비대칭적 플랫폼 참여 의도”, *Information Systems Review*, 제19권, 제3호, 2017, pp. 47-67.
- [7] 신수정, “글에서 감정을 읽다! 감성분석의 이해”, *IT World IDG Tech Report*, 2014, pp. 5-7.
- [8] 윤호민, 최규완, “감성분석을 활용한 레스토랑 선택속성과 만족도에 관한 연구: 외래관광객 온라인 리뷰 중심으로”, *호텔경영학연구*, 제30권, 제6호, 2021, pp. 99-117.
- [9] 이경렬, 김상훈, “온라인 상의 체험의 유형화와 체험이 웹사이트에 미치는 영향에 관한 연구: 슈미트(Schmitt, 1999)의 총체적 체험이론을 중심으로”, *광고학연구*, 제22권, 제1호, 2011, pp. 57-79.
- [10] 이상규, “양면시장의 정의 및 조건”, *정보통신정책연구*, 제21권, 제4호, 2014, pp. 73-105.
- [11] 임영희, 김홍범, “호텔 온라인 리뷰 빅데이터를 활용한 감성분석에 관한 연구”, *호텔경영학연구*, 제28권, 제7호, 2019, pp. 105-123.
- [12] 장경애, 박상현, 김우제, “인터넷 감정기호를 이용한 긍정/부정 말뭉치 구축 및 감정분류 자동화”, *정보과학회논문지*, 제42권, 제4호, 2015, pp. 512-521.
- [13] 정석완, “공유경제 개념의 변화와 한국의 공유경제”, *산은조사월보*, 제756호, 2018, pp. 71-85.
- [14] 조성호, “모바일 관광 앱 서비스특성이 지각된 가치, 만족 및 행동의도에 미치는 영향”, *산경연구논집*, 제10권, 제9호, 2019, pp. 45-52.
- [15] 최자영, 김현아, 김용범, “온라인 리뷰가 매출에 미치는 영향력 분석: 텍스트기반 감성지수를 중심으로”, *유통연구*, 제25권, 제3호, 2020, pp. 1-21.
- [16] 한국관광공사, 『빅데이터를 활용한 2021 관광 트렌드 분석』, 실태보고서, 2020.
- [17] 한국문화관광연구원, 『창조경제시대 산업간 융합을 통한 관광산업 발전방안』, 실태보고서, 2014.
- [18] 한국전자통신연구원, 『코로나19 이후 일반인의 IT 활용변화』, 실태보고서, 2021.
- [19] 한주형, 노은정, “가격이 중저가호텔과 게스트하우스의 고객만족도 평점에 미치는 영향: 온라인리뷰 빅데이터 분석을 중심으로”, *관광경영연구*, 제88권, 2019, pp. 331-352.
- [20] 홍순기, 이진희, 김남조, “랜선투어를 통한 간접체험이 지각된 가치와 지불의도에 미치는 영향: 체험경제이론의 적용”, *관광레저연구*, 제34권, 제6호, 2022, pp. 61-76.
- [21] Belk, R., “You are what you can access: Sharing and collaborative consumption online”, *Journal*

- of *Business Research*, Vol.67, No.8, 2014, pp. 1595-1600.
- [22] Chintagunta, P. K., S. Gopinath, and S. Venkataraman, "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets", *Marketing Science*, Vol.29, No.5, 2010, pp. 944-957.
- [23] Dredge, D. and S. Gyimóthy, "The collaborative economy and tourism: Critical perspectives, questionable claims and silenced voices", *Tourism Recreation Research*, Vol.40, No.3, 2015, pp. 286-302.
- [24] Heimerl, F., S. Lohmann, S. Lange, and T. Ertl, "Word cloud explorer: Text analytics based on word clouds", In *2014 47th Hawaii International Conference on System Sciences* (pp. 1833-1842), IEEE, 2014.
- [25] Herr, P. M., F. R. Kardes, and J. Kim, "Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnostics perspective", *Journal of Consumer Research*, Vol.17, No.4, 1991, pp. 454-462.
- [26] Hu, N., N. S. Koh, and S. K. Reddy, "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales", *Decision Support Systems*, Vol.57, 2014, pp. 42-53.
- [27] Jiang, Z. and I. Benbasat, "The effects of presentation formats and task complexity on online consumers' product understanding", *MIS Quarterly*, Vol.31, No.3, 2007, pp. 475-500.
- [28] Kim, W. G., H. Lim, and R. A. Brymer, "The effectiveness of managing social media on hotel performance", *International Journal of Hospitality Management*, Vol.44, 2015, pp. 165-171.
- [29] Kumar, V., A. Lahiri, and O. B. Dogan, "A strategic framework for a profitable business model in the sharing economy", *Industrial Marketing Management*, Vol.69, 2018, pp. 147-160.
- [30] Lessig, L., *Remix: Making Art and Commerce Thrive in the Hybrid Economy*, The Penguin Press, New York, 2008.
- [31] Li, X. and L. M. Hitt, "Self-selection and information role of online product reviews", *Information Systems Research*, Vol.19, No.4, 2008, pp. 456-474.
- [32] Long, J. S. and J. S. Long, *Regression models for categorical and limited dependent variables* (Vol. 7). Sage, 1997.
- [33] Maheswaran, D. and B. Sternthal, "The effects of knowledge, motivation, and type of message on ad processing and product judgments", *Journal of consumer Research*, Vol.17, No.1, 1990, pp. 66-73.
- [34] Matfield, "The Impact of Customer Reviews and Ratings on Conversion Rates", 2011, Available at <https://www.smartinsights.com/conversion-optimisation/product-page-optimisation/reviews-conversion-rate-impact/>.
- [35] Metzger, M. J., "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research", *Journal of the American Society for Information Science and Technology*, Vol.58, No.13, 2007, pp. 2078-2091.
- [36] Mudambi, S. M. and D. Schuff, "Research note: What makes a helpful online review? A study of customer reviews on Amazon.com", *MIS Quarterly*, Vol.34, No.1, 2010, pp. 185-200.
- [37] Osgood, D. W., L. L. Finken, and B. J. McMorris, "Analyzing multiple-item measures of crime and deviance II: Tobit regression analysis of transformed scores", *Journal of Quantitative Criminology*, Vol.18, No.4, 2002, pp. 319-347.
- [38] Scanfeld, D., V. Scanfeld, and E. L. Larson,

- “Dissemination of health information through social networks: Twitter and antibiotics”, *American Journal of Infection Control*, Vol.38, No.3, 2010, pp. 182-188.
- [39] Schmitt, B. H., “Experimental marketing: How to get customers to sense, feel, think, act and relate to your company and brands”, *European Management Journal*, Vol.18, No.6, 1999, p. 695.
- [40] Schuckert, M., X. Liu, and R. Law, “Hospitality and tourism online reviews: Recent trends and future directions”, *Journal of Travel & Tourism Marketing*, Vol.32, No.5, 2015, pp. 608-621.
- [41] Schwenk, C. R., “The use of participant recollection in the modeling of organizational decision process”, *Academy of Management Review*, Vol.10, No.3, 1985, pp. 496-503.
- [42] Walter, K., “Sentiment analysis for hotel reviews”, *Speech Technologies*, Vol.2, 2012, pp. 96-109.

Sentiment Analyses of the Impacts of Online Experience Subjectivity on Customer Satisfaction

Yeeun Seo* · Sang-Yong Tom Lee**

Abstract

The development of information technology(IT) has brought so-called “online experience” to satisfy our daily needs. The market for online experiences grew more during the COVID-19 pandemic. Therefore, this study attempted to analyze how the features of online experience services affect customer satisfaction by crawling structured and unstructured data from the online experience web site newly launched by Airbnb after COVID-19. As a result of the analysis, it was found that the structured data generated by service users on a C2C online sharing platform had a positive effect on the satisfaction of other users. In addition, unstructured text data such as experience introductions and host introductions generated by service providers turned out to have different subjectivity scores depending on the purpose of its text. It was confirmed that the subjective host introduction and the objective experience introduction affect customer satisfaction positively. The results of this study are to provide various implications to stakeholders of the online sharing economy platform and researchers interested in online experience knowledge management.

Keywords: *Online Experience Service, Text Mining, Sentiment Analysis, Customer Satisfaction*

* Graduate Student, School of Business, Hanyang University

** Corresponding Author, Professor, School of Business, Hanyang University

● 저 자 소 개 ●



서 예 은 (dpdms1044@hanyang.ac.kr)

한양대학교 대학원에서 경영정보시스템전공 석사학위를 취득하였다. 주요 관심 분야는 소셜미디어, 빅데이터애널리틱스, 메타버스 등이다.

.



이 상 용 (tomlee@hanyang.ac.kr)

현재 한양대학교 경영대학 교수로 재직 중이다. 주요 관심 분야는 정보경제, 개인정보 보호(privacy) 및 보안, 소셜미디어, 빅데이터애널리틱스 등이다. 관련 연구들을 MIS Quarterly, Management Science, Journal of Management Information Systems 등을 비롯한 다수의 저널에 관련 논문을 게재하고 있다.

.

논문접수일 : 2022년 11월 18일

1차 수정일 : 2023년 01월 17일

게재확정일 : 2023년 02월 14일

2차 수정일 : 2023년 02월 13일