

소셜미디어 어널리틱스 기반 서비스품질 평가: 항공산업을 중심으로

Service Quality Evaluation based on Social Media Analytics: Focused on Airline Industry

한 명 기 (Myoung-Ki Han) 국민대학교 데이터사이언스학과 대학원 박사과정
최 병 구 (Byounggu Choi) 국민대학교 경영대학 교수, 교신저자

요 약

항공산업의 경쟁이 치열해짐에 따라 효과적인 항공사 서비스 품질 측정은 주요 과제 중 하나가 되었다. 특히 빅데이터 어널리틱스가 새로운 연구 패러다임으로 각광받게 됨에 따라 소비자가 직접 작성한 온라인 리뷰 분석을 통한 항공사 서비스 품질 측정 연구들이 새롭게 시도되고 있다. 그러나 이러한 연구들은 리뷰 제목을 분석에 활용하지 않았다는 점, 학습 데이터 셋 구축을 위한 레이블링 (labeling)에 있어 사람의 개입이 많이 요구되는 지도 학습(supervised learning)에 의존한다는 점, 서비스 품질 차원 분류에 있어 항공사 특성을 고려하지 못한다는 점 등이 문제로 지적되고 있다. 기존 연구의 한계를 극복하기 위해 본 연구에서는 제목과 본문을 포함한 온라인 리뷰 전체를 자가학습(self-training)과 감성 분석을 활용해 AIRQUAL 서비스 품질 차원으로 분류함으로써 객관적이고 정교한 서비스 품질 측정을 시도하였으며 이를 기반으로 서비스 품질 차원이 서비스 만족도에 미치는 영향을 파악하였다. 분석 결과 온라인 리뷰로부터 AIRQUAL의 다섯 가지 서비스 품질 차원을 효과적으로 추출할 수 있었으며 각 서비스 품질 차원은 모두 서비스 만족도에 유의한 영향을 미치는 것으로 나타났다. 나아가 리뷰 제목이 서비스 만족도에 미치는 영향 또한 유의한 것으로 파악되었다. 본 연구는 항공산업의 특성을 반영한 서비스 품질 차원 측정 및 이의 효과에 대한 분석이라는 측면에서 학문 및 실무적 의의가 있다.

키워드 : 항공사 서비스 품질 평가, 소셜미디어 어널리틱스, AIRQUAL 모형, 자가학습, 감성분석

I. 서 론

최근 몇십 년 동안 항공 여행은 급격한 성장을

보이고 있다. 항공 여객 수요는 코로나 19로 여행 제약이 극대화된 2020년을 제외하면 2010년부터 2019년까지 매년 지난해 대비 7%가 넘는 성장을 이루어 왔으며 2021년에는 18%, 코로나19 사태가 진정될 2022년에는 무려 51%까지 증대할 것으로 예측되고 있다(Statista, 2021). 이러한 성장세에 따

† 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A5A2A010 47070).

라 항공산업의 경쟁 역시 더욱 치열해지고 있다. 더욱이 1971년 사우스웨스트 항공에 의해 도입된 저가 항공(low-cost carriers, LCC)은 이러한 경쟁을 더욱 심화시키고 있다(Park *et al.*, 2019).

이러한 경쟁 환경에서 생존하기 위해 항공사들은 가격정책과 서비스 품질 개선을 통해 기존 고객을 유지하고 새로운 고객을 유치함으로써 경쟁우위를 확보하고자 노력하고 있다(Sukwadi *et al.*, 2021). 그러나 대다수의 항공사들이 동적 가격 모형(dynamic pricing model)을 통해 고객들에게 유사한 “최적 운임(best fares)”을 제공함으로써 가격정책은 차별적인 경쟁우위를 제공하지 못하고 있다(Park *et al.*, 2020). 나아가 저가 정책과 같은 가격정책의 경우 서비스 품질 및 고객 만족 요소를 훼손시킴으로써 장기적으로 항공사 경쟁우위 제고에 부정적인 영향을 미치고 있다(Boetsch *et al.*, 2011). 이에 따라 항공사들은 가격정책보다는 서비스 품질 향상을 통한 경쟁우위 확보에 초점을 두고 기업 전략을 수립하고 있다(Park *et al.*, 2019).

서비스 품질 향상을 위한 주요 연구과제는 서비스 품질의 효과적 측정이라 할 수 있다(Hasan *et al.*, 2019). 서비스 품질 측정의 핵심은 제공된 서비스 품질에 대한 고객의 인식과 기대수준 간의 차이를 얼마나 정확하게 평가하는가에 있다. 이와 관련하여 1980년대 초부터 북유럽과 미국을 중심으로 많은 연구가 진행되어 다양한 측정 모형이 개발되어 왔다. 특히 Parasurman *et al.*(1988)에 의해 제안된 SERVQUAL¹⁾은 서비스 품질 측정을 위해 가장 널리 활용되고 있다. 항공사의 서비스 품질 측정을 시도한 연구 또한 SERVQUAL을 기반으로 설문조사를 통해 확보된 데이터를 계량분석하거나 실증분석함으로써 서비스 품질 향상을 위

한 핵심 요인 파악, 서비스 품질과 고객 만족도 또는 고객 충성도 간의 관계 파악 등을 시도하였다(Deveci *et al.*, 2018; Tsaur *et al.*, 2002). 나아가 몇몇 연구는 항공산업의 특성을 반영한 서비스 품질 측정을 위해 제안된 AIRQUAL²⁾ 모형을 활용하여 서비스 품질을 측정함으로써 항공산업의 특성을 반영하고자 하였다(Abdel Rady, 2018; Farooq *et al.*, 2018). 최근 소셜 미디어가 급격하게 발전함에 따라 몇몇 연구들은 전통적인 설문조사 방식 대신 온라인 리뷰 데이터 분석을 통하여 서비스 품질 측정을 시도함으로써 서비스 품질과 관련된 연구의 지평을 넓혀가고 있다(Tian *et al.*, 2020).

이러한 기존 연구는 항공사의 서비스 품질 측정과 관련된 우리의 이해를 일정 정도 증진시킨 것도 사실이다. 그러나 항공산업의 서비스 품질과 관련된 기존 연구는 크게 다음과 같은 점에서 그 한계점을 보이고 있다. 첫째, 서비스 품질 측정에 있어 설문조사 기반의 자가 응답(self-reported) 기법에 지나치게 의존하고 있다. 이와 같은 전통적인 서비스 품질 측정 방식은 사회적 바람직성 편향(social desirability bias)나 묵인 편향(acquiescence bias)와 같은 응답 편향(response bias), 제한된 표본 등과 같은 문제점으로 인하여 결과의 신뢰성과 객관성 확보에 한계가 있다(Parshall *et al.*, 1995). 둘째, 설문조사 방식의 한계점을 어느 정도 극복한 것으로 평가되고 있는 온라인 리뷰 기반 연구는 불완전한 리뷰 활용과 비효과적 서비스 품질 차원 분류라는 점에서 한계가 존재한다. 리뷰 제목의 중요성에도 불구하고

1) Parasurman *et al.*(1988)은 소비자가 실제 경험한 서비스 수준에 대한 인식과 서비스에 대한 기대수준 간의 차이를 통해 서비스 품질을 측정할 수 있다고 주장하고 유형성(tangibles), 신뢰성(reliability), 대응성(responsiveness), 확신성(assurance), 공감성(empathy)의 5가지 차원으로 구성된 SERVQUAL 모형을 제안하였다.

2) AIRQUAL 모형은 항공사 유형성(airline tangibles), 터미널 유형성(terminal tangibles), 직원(personnel), 공감성(empathy), 이미지(image)의 5개 차원으로 구성되었다(Ekiz *et al.*, 2006). 이때 i) 항공사 유형성은 항공기의 유형과 항공기 내부의 시설 및 시스템을, ii) 터미널 유형성은 공항의 유형적 요소와 공항 서비스를, iii) 직원은 항공사에서 근무하는 직원을, iv) 공감성은 서비스에 대한 항공사의 약속 이행과 약속이 이행되지 않았을 때 항공사의 대처를, v) 이미지는 고객이 서비스를 받을 때 느끼는 항공사에 대한 전체적인 인상을 평가하기 위해 고안되었다.

하고 대다수 연구들은 온라인 리뷰 분석에 있어 리뷰 제목을 거의 활용하지 않고 있어 리뷰를 온전하게 활용했다고 보기 어렵다(Biswas *et al.*, 2021; Cao *et al.*, 2011). 나아가 서비스 품질 차원 분류를 위한 데이터 셋 구축에 있어 사람의 개입이 많이 요구되는 지도 학습(supervised learning)에 의존함으로써 시간과 비용 측면에서 문제점이 있다. 마지막으로, 항공산업이 갖는 특수성을 고려하지 못한다는 점에서 한계가 있다. 온라인 리뷰를 활용한 기존 연구의 경우 서비스 품질 평가를 위한 범용 모델인 SERVQUAL에 의존함으로써 항공산업의 차별적인 서비스 품질 구성 차원을 고려하지 못하고 그 결과 정교한 서비스 품질 평가가 이루어지지 않고 있다(Carman, 1990).

이러한 한계점을 보완하기 위해 본 연구에서는 i) 제목과 본문을 포함한 온라인 리뷰 전체를 ii) 자가학습(self-training)과 감성분석(sentiment analysis)을 활용해 iii) AIRQUAL의 서비스 품질 차원으로 분석함으로써 보다 객관적이고 정교하며 효율적인 서비스 품질 측정을 하고자 한다. 본 연구는 기존 연구와 달리 온라인 리뷰를 활용함으로써 전통적 설문방식의 한계점을 극복할 수 있다는 점, 리뷰 내용뿐 아니라 제목을 함께 분석함으로써 온라인 리뷰 전체를 활용할 수 있다는 점, 자가학습(self-training)을 통해 사람의 개입을 감소시킴으로써 서비스 품질 차원 측정의 효율성과 효과성을 향상할 수 있다는 점, SERVQUAL이 아닌 AIRQUAL을 사용함으로써 항공사의 특성을 반영할 수 있다는 점에서 기존 연구와 차별성이 있다.

본 연구는 다음과 같이 구성된다. 다음 장에서는 서비스 품질과 관련된 개념적 배경을 소개하고 항공산업에서의 서비스 품질 연구를 요약한다. 제 III장에서는 데이터의 수집과 평가 방법 등을 포함한 연구 방법을 설명한다. 제 IV장에서는 연구 모델을 분석하고 제 V장에서는 분석결과와 이의 함의를 논의한다. 마지막 장에서는 본 연구의 결론 및 향후 연구과제를 제안한다.

II. 개념적 배경

2.1 소셜 미디어와 온라인 리뷰

정보기술 특히 인터넷이 발전함에 따라 소비자들은 제품이나 서비스에 대한 평가를 온라인을 통해 공유할 수 있게 되었다(Avery *et al.*, 1999). 아마존에 의해 처음 제공된 온라인 리뷰는 이제 온라인에서 거래되는 거의 모든 제품과 서비스에 제공되고 있다. 온라인 소비자 리뷰(online customer review)는 기업 또는 제 3자 웹사이트에 게시된 다른 소비자의 제품이나 서비스에 대한 평가로 정의할 수 있다(Mudambi and Schuff, 2010). 온라인 소비자 리뷰는 기업이 제공하는 제품 중심의 정보에 비해 소비자의 경험이나 상황을 기반으로 사용자 관점에서 제품 정보를 제공한다. 소비자들은 온라인 리뷰로부터 해당 제품이나 서비스에 대한 정보와 평가뿐 아니라 다른 대안에 대한 정보도 얻을 수 있어 구매 의사결정에 결정적인 도움을 받을 수 있다(Chen and Xie, 2008). 반면 기업들은 온라인 리뷰를 통해 제품이나 서비스 판매를 촉진할 수 있을 뿐 아니라 소비자의 피드백(feedback)을 수집할 수 있는 도구로 활용할 수 있다(Ghose and Ipeirotis, 2010).

인터넷은 블로그, 토론방, 쇼핑몰, 소셜 미디어 등과 같은 다양한 온라인 소비자 리뷰 플랫폼을 제공해 왔다. 특히 소셜 미디어는 친구나 지인 등과 같은 기존 네트워크를 활용하여 상품이나 서비스에 대한 의견 및 경험을 공유한다는 측면에서 소비자 리뷰의 신뢰성을 더욱 증진시켜 보다 강화된 상품 혹은 서비스 구매 증진 효과를 보이고 있다(Wang *et al.*, 2012). 또한 소셜 미디어는 텍스트뿐 아니라 사진, 비디오 등과 같은 시각적 효과를 증대할 수 있는 기능을 제공함으로써 소비자 리뷰를 더욱 즐겁고 매력적으로 만듦으로써 제품이나 서비스에 대한 소비자의 관심을 증대시키고 있다. 이에 따라 소비자들은 제품이나 서비스 구매를 위해, 기업은 상품이나 서비스의 판매 촉진을 위해

소셜 미디어 기반의 온라인 소비자 리뷰를 적극 활용하고 있다.

그러나 소셜 미디어를 활용한 온라인 소비자 리뷰가 소비자와 기업에게 긍정적 효과만을 가져다 준 것은 아니다. 소셜 미디어가 급격하게 발전함에 따라 소비자 리뷰의 양, 다양성, 생성 속도, 진실성(veracity) 등과 같은 다양한 측면에서 많은 도전과제가 발생하고 있다. 예를 들면, 방대한 양의 온라인 소비자 리뷰로 인해 소비자의 의사결정을 방해하는 정보 과부하(information overload)나 가짜 리뷰(fake review)를 통하여 소비자들의 의사결정을 왜곡하는 현상이 나타나고 있다. 이러한 문제를 해결하기 위해 빅데이터 어널리틱스(analytics)가 새로운 연구 패러다임으로 각광받고 있다(Xiang *et al.*, 2017). 빅데이터 어널리틱스는 기계학습(machine learning), 컴퓨터 언어학(computational linguistics), 통계 등의 통합적 활용을 통해 소셜 미디어 상의 방대한 소비자 리뷰를 효과적으로 수집, 분석, 해석할 수 있는 기반을 제공하고 있으며 이와 관련한 다양한 연구가 지속적으로 이루어지고 있다(Duan *et al.*, 2008). 수집 관련 연구들은 기업의 비즈니스 활동이나 관심사와 관련된 리뷰를 다양한 온라인 플랫폼으로부터 획득하는데 집중하고 있으며 딥웹(deep web)이나 멀티미디어 콘텐츠를 포함한 웹 크롤링(Khalsa *et al.*, 2017), 의미적 유사성(semantic similarity)에 기반한 크롤링(Lin *et al.*, 2017) 등과 같은 연구가 이루어지고 있다. 분석 관련 연구들은 수집된 온라인 소비자 리뷰의 표현을 기반으로 의미를 평가하고 의사결정을 위한 지표 생성에 집중하고 있으며(Choi *et al.*, 2021), 감성 분석, 추세 분석(trend analysis), 토픽 모델링(topic modeling), 소셜 네트워크 분석 기법 등의 개선을 위한 다양한 연구가 제시되고 있다(Fan and Gordon, 2014). 해석 관련 연구들은 분석 결과를 기반으로 비즈니스 인사이트(insights)를 획득하고 이를 의미 있는 방식으로 표현하는 연구에 집중하고 있으며 다차원척도법(multidimensional scaling, MDS), 대응 분석(corre-

spondence analysis), 포지셔닝 맵(positioning map) 등을 활용하여 도출된 결과를 소비자들에게 직관적이고 효과적으로 제시하는데 초점을 두고 있다(Tirunillai *et al.*, 2014).

소셜 미디어 상의 온라인 소비자 리뷰를 활용한 연구들은 소매, 유통, 관광, 헬스케어, 통신 등 다양한 산업에서 이루어지고 있다. O'Connor(2010)는 호텔 산업에서 소셜 미디어를 활용한 온라인 소비자 리뷰 관리의 중요성을 주장하였으며 Duan *et al.*(2016)은 비정형 온라인 소비자 리뷰에 대한 감성 분석을 기반으로 감성 점수가 호텔의 평가에 미치는 영향을 분석하였다. Yang *et al.*(2003)은 온라인 유통과 관련된 온라인 소비자 후기를 대상으로 비정형 데이터 분석을 실시하여 14개의 서비스 품질 차원을 파악하고 소비자의 만족과 불만족에 영향을 미치는 차원이 서로 상이함을 주장하였다. Ranard *et al.*(2016)은 Yelp의 리뷰를 기반으로 감성 분석을 실시한 결과 기존 설문 기반 병원 서비스 평가 방식이 발견하지 못한 추가적인 평가 차원을 발견하고 기존 방식과 온라인 소비자 리뷰 분석 간의 높은 상관관계가 있음을 실증하였다. Saputro *et al.*(2021)은 트위터에 나타난 소비자 리뷰에 대한 감성 분석을 실시하여 통신회사의 서비스 품질에 영향을 미치는 7가지 차원을 제시하였다.

항공산업 또한 온라인 소비자 리뷰를 활용한 다양한 연구가 진행되어 왔다. Lee and Yu(2018)는 온라인 리뷰 분석을 통해 공항의 서비스 품질 평가를 시행하고 이를 기반으로 온라인 서비스 품질 평가 방식이 전통적 공항 서비스 품질 평가 방식을 대체할 수 있다고 주장하였다. Sezgen *et al.*(2019)은 온라인 소비자 리뷰를 기반으로 일반 항공사와 저가 항공사의 고객 만족 및 불만족 요인을 분석하고 항공사 유형에 따라 이들 요인 간 차이가 있음을 주장하였다. Lucini *et al.*(2020)은 텍스트 마이닝(text mining) 기반 온라인 소비자 리뷰 분석을 통해 항공사 고객 만족도 측정을 위한 새로운 프레임워크를 제시하였다. Park *et al.*(2020)은 미국 20개 항공사에 대한 온라인 소비자 리뷰 분석을 통해 소비자의

만족도에 영향을 미치는 요인과 불만족에 미치는 영향이 서로 다름을 주장하였다.

2.2 온라인 소비자 리뷰 제목의 중요성

온라인 소비자 리뷰는 제목과 본문(body text)으로 구성되어 있음에도 불구하고 대부분의 빅데이터 어널리틱스 기반 연구들은 온라인 리뷰 본문만을 분석에 활용하고 있다(Biswas *et al.*, 2021). 예를 들면, Cao *et al.*(2011)은 리뷰 본문의 특성을 극단성(extremity), 주관성(subjectivity) 등을 포함한 기본적(basic) 특성, 평균 단어 및 문자 수 등을 포함한 문체적(stylistic) 특성, 단어의 의미 등을 포함한 의미론적(semantic) 특성으로 구분하고 이러한 특성이 소비자의 구매 행위에 미치는 영향을 분석하였다. 이외에도 리뷰의 문장 구조 혹은 가독성(readability) 등 주로 리뷰 본문의 특성에 초점을 두고 온라인 리뷰를 분석하였을 뿐 리뷰의 주요 구성요소인 리뷰 제목의 역할에 대한 분석은 매우 미비한 실정이다.

온라인 리뷰 분석에 있어 리뷰 제목은 매우 중요한 역할을 하고 있다. 첫째, 리뷰 제목은 강력하고 즉각적인 정보 제공을 통해 어떤 리뷰를 읽을 것인가에 대한 소비자의 의사결정에 영향을 미친다. 선택적 주의 이론(theory of selective attention)에 따르면 사람은 제한적인 정보처리 능력으로 인해 입력되는 모든 정보를 한꺼번에 처리할 수 없기 때문에 주요 정보에 선택적으로 반응하게 된다(Treisman, 1969). 이는 온라인 리뷰 상황에서도 동일하게 적용된다. 일반적으로 소비자들은 하나의 제품이나 서비스에 대한 수많은 온라인 리뷰와 작성자 정보, 유용성 등급 등과 같은 다양한 종류의 관련 정보를 접하게 된다. 이러한 모든 정보를 처리하는 것은 매우 어려울 뿐만 아니라 많은 시간이 소요된다. 따라서 소비자들은 이러한 문제 해결에 도움을 줄 수 있는 즉각적인 정보로서 온라인 리뷰 제목에 의존하게 된다(Gottschalk and Mafael, 2017). 온라인 리뷰 제목은 신문의 헤드라

인이나 광고의 태그와 같이 전체 리뷰를 함축적으로 표현함으로써 소비자들에게 짧지만 매우 강력하고 즉각적인 정보를 제공한다(Ascaniis and Gretzel, 2012). 소비자들은 온라인 리뷰 제목을 먼저 확인한 후 해당 제품이나 서비스에 대해 보다 자세한 정보가 필요한 경우 리뷰 본문을 정밀하게 파악함으로써 제한적인 정보처리 능력을 효과적으로 활용하게 된다(Salehan and Kim, 2016).

둘째, 리뷰 제목은 리뷰 작성자의 감정을 응축적으로 표현함으로써 효과적인 감정 전달의 매개체 역할을 통해 소비자의 의사결정에 영향을 줄 수 있다. 소비자는 제품이나 서비스에 대한 긍정적이거나 부정적으로 강한 정서적 경험을 하였을 때 리뷰를 남기는 경향이 있다(Nakayama and Wan, 2019). 나아가 자신이 경험한 강한 정서적 경험을 효과적으로 표현함으로써 다른 사람으로부터 감정적 지지를 획득하고자 한다. 이를 위해 소비자는 자신의 경험에 대한 특성을 종합적이고 효과적으로 표현할 수 있는 제목을 사용할 가능성이 높다. 그 결과 리뷰의 제목은 리뷰 작성자의 감정을 다른 소비자들에게 효과적으로 전달하여 리뷰 본문을 읽도록 설득할 수 있는 역할을 하게 된다(Liu, 2012).

셋째, 리뷰 제목은 제품이나 서비스를 사용하기 전 형성된 태도를 강화함으로써 소비자의 의사결정에 영향을 미친다. 선택적 주의 이론에 따르면 사람은 정보 흡수를 위해 멘탈 구조(mental structure)를 활용한다. 즉 자신의 멘탈 모형과 일치하는 정보는 받아들이는 반면 그렇지 않은 정보는 배제하는 경향이 있다(Treisman, 1969). 이는 온라인 리뷰 상황에서도 동일하다. 즉 사람들은 제품이나 서비스를 사용하기 전 이미 이에 대한 긍정이나 부정적인 태도를 형성하고 있을 가능성이 높다. 따라서 제품이나 서비스에 대해 긍정적인 태도를 가지고 있는 소비자들은 긍정적인 감정을, 부정적인 태도를 가지고 있는 소비자들은 부정적인 감정의 리뷰를 찾게 된다(Salehan and Kim, 2016). 이를 통해 자신의 태도에 대한 확신을 얻음

으로써 감정적 안정을 추구하고자 한다(Bailey, 2005). 리뷰 제목은 이러한 감정적 안정을 빠르게 확보할 수 있도록 도와줌으로써 소비자의 의사결정에 영향을 미친다.

2.3 항공사 서비스 품질 연구

기업 간 경쟁이 심화됨에 따라 서비스 품질 향상을 통한 차별화 전략의 중요성이 점점 증대해왔다(Thompson *et al.*, 1985). 높은 서비스 품질 유지를 위해서는 정확한 서비스 품질 평가가 선행되어야 하기 때문에 이와 관련한 다양한 연구가 진행되어 왔다(Parasuraman *et al.*, 1988). 이는 항공산업도 마찬가지이다. 지금까지 항공산업에서의 서비스 품질 측정과 관련한 연구는 데이터 수집 방식과 서비스 품질 차원 분류 방식을 기준으로 크게 i) 설문 기반 자체 서비스 품질 차원 활용, ii) 설문 기반 기존 서비스 품질 차원 활용, iii) 온라인 데이터 기반 자체 서비스 차원 활용, iv) 온라인 데이터 기반 기존 서비스 차원 활용의 4가지 유형으로 구분할 수 있다.

첫 번째 범주에 속한 연구들은 설문을 통하여 수집한 데이터를 기반으로 자체적으로 서비스 품질 차원을 도출하고 각 품질 차원이 만족도 혹은 충성도와 같은 성과에 미치는 영향을 파악하고자 하였다(Liou *et al.*, 2011; Wu and Cheng, 2013). 예를 들면, Aksoy *et al.*(2003)은 터키 승객들을 대상으로 수집된 데이터를 기반으로 요인 분석을 실시하여 음식료 서비스, 승무원, 인터넷 서비스 등을 포함한 9가지 서비스 차원을 파악하고 터키 사람들과 유럽인들이 중요하게 생각하는 서비스 차원이 통계적으로 유의한 차이가 있음을 실증하였다. Park(2007)은 한국과 호주 승객을 대상으로 수집된 데이터를 기반으로 요인분석을 실시하여 기내 서비스, 예약 관련 서비스, 공항 서비스 등의 11개 서비스 차원을 파악하고 항공사, 사용 빈도, 좌석 등급에 따라 서비스 품질 차원에 대한 인식이 유의한 차이가 있음을 주장하였다. 두 번째 범주에

속하는 연구들은 설문을 통해 수집된 데이터를 활용하여 SERVQUAL이나 AIRQUAL과 같은 기존 서비스 품질 측정 모형에서 제시된 서비스 품질 차원을 측정하고 통계나 의사결정 기법을 통해 서비스 품질 차원이 성과에 미치는 영향을 파악하고자 하였다(Deveci *et al.*, 2018; Hussain *et al.*, 2015; Perçin, 2018). 예를 들면, Tsaor *et al.*(2002)은 설문을 통하여 수집된 데이터를 AHP와 TOPSIS를 활용하여 분석함으로써 SERVQUAL의 다섯 가지 차원 가운데 유형성이 가장 중요한 반면 공감성이 가장 덜 중요한 차원이라고 주장하였다. Abdel Rady(2018)는 이집트 승객을 대상으로 수집된 데이터를 활용하여 항공산업에 특화된 AIRQUAL 모형의 다섯 가지 차원이 항공사의 이미지에 미치는 영향을 파악하였다.

세 번째 범주에 속한 연구들은 설문 기반 서비스 품질 연구의 한계를 지적하고 온라인 데이터를 기반으로 자체적인 서비스 품질 차원을 도출한 후 인공지능 기법을 활용하여 각 품질 차원이 항공사 성과에 미치는 영향을 파악하고자 하였다(Gao *et al.*, 2021; Stamolampros *et al.*, 2019). 예를 들면, Kocak and Atalik(2019)은 터키 항공사의 트위터 데이터에 대한 특성 기반 감성 분석(aspect-based sentiment analysis)을 실시하여 이미지, 안전, 가격 등을 포함한 15개의 서비스 품질 차원을 파악하고 이를 기준으로 서비스 품질 차원에 대한 지각도(perceptual map)를 작성하였다. Korfiatis *et al.*(2019)은 tripadvisor.com의 온라인 리뷰 데이터에 대한 구조적 토픽 모델링(structural topic modeling)을 실시하여 청결도, 체크인, 음식료 등을 포함한 8개의 서비스 품질 차원을 파악하고 이를 기반으로 항공사의 경쟁력을 평가하고자 하였다. 네 번째 범주에 속한 연구들은 온라인 데이터를 기반으로 SERVQUAL과 같은 기존 서비스 품질 모형에서 제시한 품질 차원을 측정하고 이를 기반으로 서비스 품질 차원이 성과에 미치는 영향을 파악하고자 하였다(Riantama *et al.*, 2021). 예를 들면, Tian *et al.*(2020)은 12개 미국 항공사의 트위터 데이터

에 대한 텍스트 마이닝을 실시하여 SERVQUAL의 서비스 품질 차원에 따른 항공사별 서비스 품질을 측정하고 이를 기존 전통 방식에 따른 서비스 품

질 측정 결과와 비교함으로써 온라인 데이터 기반 서비스 품질 측정 방식의 유효성을 검증하였다. 이를 요약하면 다음 <표 1>과 같다.

<표 1> 기존 연구 요약

데이터 수집	서비스 품질 차원	핵심 내용	분석 방법	주요 서비스 차원	주요 연구
설문	자체 분류 도출	설문을 통해 수집된 데이터를 기반으로 통계나 의사결정 기법을 실시하여 서비스 품질 차원을 자체적으로 분류하고 이를 측정하여 서비스 품질이 만족도 혹은 충성도와 같은 성과에 미치는 영향을 파악	통계 기법 (요인분석, 판별분석, 회귀분석, 구조방정식 등) 의사결정기법 (fuzzy AHP, fuzzy TOPSIS, VIKOR 등)	음식료 서비스, 승무원, 인터넷 서비스, 불만처리, 비행 스케줄 및 정보, 물리적 환경 등	Aksoy <i>et al.</i> (2003), Park(2007), Wu and Cheng(2013) Demir(2012), Li <i>et al.</i> (2017), Liou <i>et al.</i> (2011)
	기존 분류 활용	설문을 통해 수집된 데이터를 활용해 SERVQUAL이나 AIRQUAL과 같은 기존 서비스 품질 모형에서 제시된 서비스 품질 차원을 측정하고 통계나 의사결정 기법을 통해 서비스 품질 차원이 성과에 미치는 영향을 파악	통계 기법 (요인분석, 판별분석, 회귀분석, 구조방정식 등) 의사결정기법 (AHP, fuzzy ANP, fuzzy DEMATEL, Interval type-2 hesitant fuzzy sets, TOPSIS 등)	SERVQUAL 차원(유형성, 신뢰성, 대응성, 확산성, 공감성) + 이미지, 가용성 등 AIRQUAL 차원(항공사 유형성, 터미널 유형성, 직원, 공감성, 이미지) + 비행 패턴, 예약 서비스 등	Abdel Rady(2018), Ekiz <i>et al.</i> (2006), Farooq <i>et al.</i> (2018), Nadiri <i>et al.</i> (2008), Pakdil and Aydin(2007), Suki(2004) Deveci <i>et al.</i> (2018), Perçin(2018), Tsaor <i>et al.</i> (2002)
온라인 (리뷰, 트위터 등)	자체 분류 도출	온라인 데이터를 기반으로 자체적인 서비스 품질 차원을 도출한 후 인공지능 기법을 활용하여 각 품질 차원이 성과에 미치는 영향을 파악	사람 기반 데이터 셋 구축 + 지도 학습(서포트 벡터 머신, 랜덤 포레스트 등); 감성 분석, 구조토픽 모델링 등	체크인 서비스, 탑승 서비스, 음식료 서비스, 청결도, 승무원, 인터넷 서비스 등	Gao <i>et al.</i> (2021), Kocak and Atalik (2019), Korfiatis <i>et al.</i> (2019), Stamolampros <i>et al.</i> (2018)
	기존 분류 활용	온라인 데이터를 기반으로 SERVQUAL 모형에서 제시된 서비스 품질 차원을 측정하고 인공지능 기법을 활용하여 각 품질 차원이 성과에 미치는 영향을 파악	사람 기반 데이터 셋 구축 + 지도 학습 + 감성 분석, LASSO 등 준지도 학습 (semi-supervised learning) 기반 데이터 셋 구축 + 지도 학습 + 감성 분석	SERVQUAL 차원(유형성, 신뢰성, 대응성, 확산성, 공감성) + 안전성 등 AIRQUAL 차원	Riantama <i>et al.</i> (2021), Tian <i>et al.</i> (2020) This study

항공사의 서비스 품질 측정과 관련된 기존 연구들 요약하면 다음과 같은 흥미로운 사실을 발견할 수 있다. 첫째, 서비스 품질 측정을 위해 다양한 통계 기법이나 의사결정 기법을 활용하고 있으나 대다수 연구들이 설문조사 방식을 기반으로 데이터를 수집하고 있다(Aksoy *et al.*, 2003; Li *et al.*, 2017; Farooq *et al.*, 2018; Tsaor *et al.*, 2002). 그러나 설문조사에 기반한 서비스 품질 측정 방식은 다양한 응답 편익과 제한된 표본으로 인해 신뢰성과 객관성 측면에서 한계가 존재한다. 둘째, 소셜 미디어 어널리틱스가 발전함에 따라 최근 서비스 품질 측정 연구의 대부분은 트위터나 온라인 리뷰를 활용함으로써 기존 설문조사 방식의 한계점을 극복하고자 하고 있다(Kocak and Atalik, 2019; Korfiatis *et al.*, 2019; Stamolampros *et al.*, 2018; Tian *et al.*, 2020). 그러나 이러한 연구들은 리뷰 제목의 중요성에도 불구하고 리뷰 본문에만 초점을 둘 뿐 리뷰 제목에 대한 분석을 시행한 연구는 없었다(Biswas *et al.*, 2021). 셋째, 온라인 데이터를 활용한 연구의 대부분은 서비스 품질 분류 및 측정을 위한 레이블화 되어진 학습 데이터 셋의 구축에 있어 사람의 개입에 의존하고 있다(Gao *et al.*, 2021; Riantama *et al.*, 2021). 그러나 사람에 의한 레이블화 작업은 정확한 출력 데이터 구축 과정에서 많은 시간과 비용이 소요될 뿐 아니라 작은 양의 데이터로 인해 과적합(overfitting)과 같은 문제가 발생할 가능성이 높다(Dua and Chowriappa, 2013). 마지막으로 대다수 연구들이 SERVQUAL이라는 일반적 서비스 품질 평가 모형을 사용하고 있다(Tian *et al.*, 2020). 몇몇 설문지 기반 연구의 경우 항공산업의 특성을 반영한 AIRQUAL을 활용하고 있으나(Ekiz *et al.*, 2006; Farooq *et al.*, 2018) 온라인 데이터를 활용한 연구의 경우 SERVQUAL에 의존함으로써 항공산업이 갖는 서비스의 특성을 정교하게 평가하지 못하고 있다(Tian *et al.*, 2020).

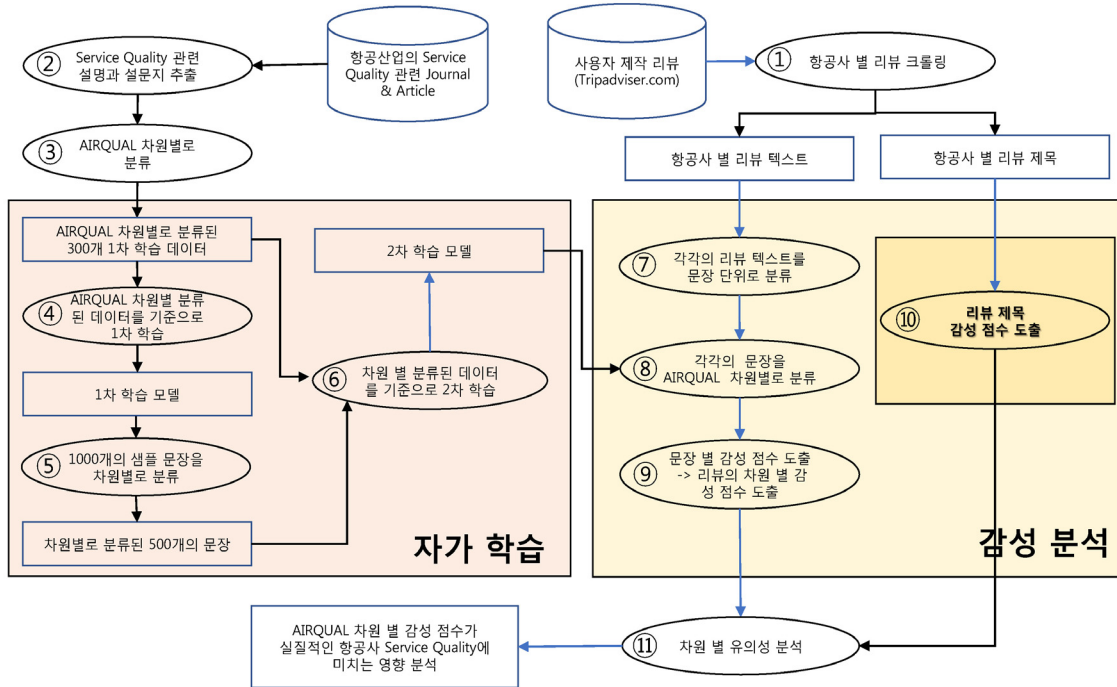
Ⅲ. 연구 방법

본 연구의 주요 목적은 온라인 리뷰 분석을 통해 항공사의 서비스 품질이 서비스 만족도에 미치는 영향을 파악하는 데 있다. 보다 구체적으로 tripadvisor.com³⁾의 온라인 리뷰에 대한 텍스트 마이닝을 통하여 고객들의 리뷰를 AIRQUAL 서비스 품질 차원으로 분류한 후, 감성 분석을 기반으로 항공사의 서비스 품질이 서비스 만족도에 미치는 영향을 파악함으로써 항공사의 서비스 품질 관리에 대한 실무적 함의(implications)를 제공하는 데 있다. 특히 본 연구에서는 기존 연구의 한계점을 극복하기 위하여 자가학습을 통한 서비스 품질 차원별 데이터 셋을 구축하고 이를 항공사 서비스에 대한 리뷰 본문만이 아니라 리뷰 제목을 포함하여 분석하고자 한다. 자가학습은 신뢰성 있는 데이터의 수를 확보하는데 필요한 시간과 비용을 획기적으로 감소시켜 적은 데이터로 인해 발생할 수 있는 과적합 문제를 해결하는데 도움을 줄 수 있을 것이다. 이를 요약하면 다음 <그림 1>과 같다.

3.1 데이터 수집 및 처리

본 연구에서 사용되는 데이터는 크게 tripadvisor.com으로부터 수집된 항공사별 소비자 리뷰(<그림 1>의 ① 참조)와 AIRQUAL 서비스 품질 차원 관련 논문이다(<그림 1>의 ②참조). 먼저 본 연구의 핵심 분석 대상인 온라인 리뷰 데이터는 2016년 1월부터 2018년 12월 31일까지의 기간 동안 12개 항공

3) Tripadvisor.com은 사용자 생성 콘텐츠의 얼리 어답터(early adopter)로 315,000,000이상의 회원을 보유하고 있으며 호텔, 레스토랑, 관광 명소 및 기타 여행 관련 기업 등에 대한 5억 건 이상의 리뷰를 보유하고 있다. Tripadvisor.com은 다수의 회원과 리뷰를 바탕으로 소비자 여론을 선도하고 있을 뿐 아니라 대부분의 항공사 서비스 관련 온라인 리뷰를 제공하고 있기 때문에 항공사 서비스와 관련된 리뷰 데이터로서 대표성이 있다고 할 수 있다.



〈그림 1〉 데이터 수집 및 분석 절차

사)를 대상으로 영어로 작성된 리뷰를 크롤링하여 총 40,000개를 수집하였다. 수집된 리뷰 가운데, 리뷰 본문 및 리뷰 제목이 없는 리뷰 각각 163개와 135개를 제거하였다. 또한 리뷰 본문이 3단어 미만이 짧은 리뷰 283개, 항공사 리뷰와 관련 없는 광고성 리뷰 183개, 항공사 단순 알림 67개를 각각 제거하여 총 39,169개의 리뷰를 분석에 활용하였다.

또 다른 주요 데이터는 AIRQUAL 서비스 품질 차원 관련 13개의 논문으로 AIRQUAL의 서비스 품질 차원 분류를 위한 분류기(classifier) 구축을

- 4) 2018년 Traveller's Choice가 선정한 세계 100대 항공사와 Skytrax가 선정한 세계 100대 항공사 가운데 지역과 국가를 대표할 수 있는 항공사를 선정하였다. 본 연구에 포함된 항공사는 Aeroflot, Aeromexico, Air China, American Airlines, Asiana Airlines, Avianca, EVA Air, Garuda Indonesia, Japan Airlines, Korea Air, Southwest Airlines, Vietnam Airlines이다.

위해 활용되었다. 이들 논문으로부터 AIRQUAL의 정의와 서비스 품질 차원을 측정할 설문 문구를 추출한 후 각 AIRQUAL 차원별로 분류하여 총 300개의 분류기 구축을 위한 1차 학습용 데이터 셋(training data set)을 구축하였다(〈그림 1〉의 ③ 참조). 추출된 차원별 핵심 문구 리스트와 관련 논문의 예는 다음 <표 2>와 같다.

3.2 자가학습(Self-Training)

AIRQUAL 관련 주요 논문을 기반으로 구축된 학습 데이터 셋의 경우 차원 별 단어의 다양성이 부족하기 때문에 분류의 정확도가 저조할 가능성이 높다. 이러한 한계를 극복하기 위해 자가학습을 수행하였다(〈그림 1〉 참조). 자가학습은 학습용으로 사용된 레이블화 된 데이터(labeled data)가 부족할 때 학습 효과를 높이기 위해 사용된다 (Zhu and Goldberg, 2009). 자가학

〈표 2〉 주요 논문 기반 AIRQUAL 차원 별 문구 추출 결과 예

차원	추출 문구(예)	주요 관련 논문
Airline tangibles	<ul style="list-style-type: none"> • Overall condition of aircraft which includes its interior and exterior, catering services, seating comfort and cleanliness • I found the interiors of the aircraft fresh and visually appealing • The quality of food served is good 	Ali <i>et al.</i> (2015) Farooq <i>et al.</i> (2018) Hasan <i>et al.</i> (2019) Li <i>et al.</i> (2017) Park(2007)
Terminal tangibles	<ul style="list-style-type: none"> • Sign-board system at terminal, security and control system, an air conditioning system at the airport, toilet conditions and information counters for passengers • Clear terminal announcements at the airline • Number of shops in airport is sufficient 	Ali <i>et al.</i> (2015) Ekiz <i>et al.</i> (2006) Farooq <i>et al.</i> (2018) Suki(2004)
Personnel	<ul style="list-style-type: none"> • The attitude and behaviors of the airline’s staff towards its customers is the subject matter of personnel services. It includes services of the staff at terminal and attendants in flight • Crew members take immediate response of my needs • Employees will never be too busy to respond to customers’ requests 	Abdel Rady(2018) Ekiz <i>et al.</i> (2006) Hasan <i>et al.</i> (2019) Nadiri <i>et al.</i> (2008) Tasur <i>et al.</i> (2002) Park(2007)
Empathy	<ul style="list-style-type: none"> • In the airline industry, empathy results from hassle-free service; that involves carefully handling the luggage, well-developed compensation plan in case of any loss and courteous ticketing service • The airline handles complaints of passengers very well • Transportation between city and airport was not a problem 	Ali <i>et al.</i> (2015) Ekiz <i>et al.</i> (2006) Hasan <i>et al.</i> (2019) Suki(2014)
Image	<ul style="list-style-type: none"> • Overall image of the company, value for money, and promotional offers and goodwill • The airlines announced when services will be performed • Have customer’s best interests at heart 	Abdel Rady(2018) Farooq <i>et al.</i> (2018) Hasan <i>et al.</i> (2019) Nadiri <i>et al.</i> (2008) Park(2007)

습은 먼저 적은 양의 분류 데이터로 학습을 수행하여 분류기를 구축한 후 이를 활용하여 레이블화되지 않은 데이터 (unlabeled data)를 분류하고 이를 원래의 분류 데이터에 추가하여 다시 학습하는 절차를 반복함으로써 분류의 정확도를 높이게 된다.

본 연구에서는 AIRQUAL 관련 논문을 기반으로 구축된 학습 데이터를 다차원 분류 모델 (multi-class classification model)을 기반으로 학습을 진행하여 1차 분류기를 구축하였다(<그림 1>의 ④참조). 이렇게 구축된 1차 분류기를 활용하여 1000개의 미분류 샘플 문장을 분류하였다(<그림 1>의 ⑤참조). 1차 분류기를 통해 분류된 1,000개

의 문장에 대해 항공 산업 분야 전문가 3명이 각각 문장의 차원별 분류가 정확한지 아닌지를 평가하고 2명 이상의 전문가로부터 적절하게 분류되었다고 판정된 문장 가운데 분류 정확도가 높은 500개 문장을 선정하였다. 평가자 간 신뢰도 분석 결과(Fleiss’ Kappa =0.74) 분류가 적절히 이루어졌음을 확인하였다. 이러한 과정을 거쳐 선정된 500개의 문장을 1차 학습용 데이터 셋에 추가하여 총 800개의 2차 데이터 셋을 구성한 후 이를 다시 다차원 분류 모델로 학습하여 2차 분류기를 구축하였다(<그림 1>의 ⑥참조). 자가학습 후 추출된 AIRQUAL 차원 별 단어의 예를 나타내면 다음 <표 3>과 같다.

〈표 3〉 자기학습 후 AIRQUAL 차원 별 단어 추출 예

차원	주제1	주제2	주제3	주제4	주제5
Airline Tangibles	Seats: Comfortable, Leg, Room, Economy, New, Aircraft, Clean	Entertainment: Inflight, TV, Movies, Old, Screen, Like	Food: Meal, Water, Served, Drinks, Choice, Fine, Good, Service	Boarding: Extra, Early, Front, Tip, Travel, Row	Others: Worth, Assigned, Time, Selection, Aircraft
Terminal tangibles	Check-in: Check, Bags, Quick, Easy, Efficient, Line, Smooth	Baggage: Bags Everything, Fast, Line, Free Arrived, Small	Boarding: Gate, Security, Terminal, Check, Early, Late, Time, Hours	Airport: International, Lounge, Flights, Travel, Food, Fly, Hotel	Others: Airport, Trip, Airlines, One, Really
Personnel	Staff: Crew, Attendants, Cabin, They, Peoples	Response: Always, Delays, Helpful, Efficient, Attentive, Smooth, Know, See, Said	Attitude: Experience, Service, Friendly, Professional, Rude, Accommodating,	Customers: Business, Class, Bad, First, Need, Gate,	Others: Food, Boarding, Check-in, Travel
Empathy	Luggage: Baggage, Get, Check, Free, Pay, Said, Asked, Extra, Charge	Flying: Fly, Back, Booked, Due, Last, City, Problems, Always	Time: Hour, Return, Connecting, Change, Arrived, Departure, Left, Late, Delayed, Early	Service: First, This, Experience, Next, Business, Ever, Really, Class	Others: Passenger, People, Others, Customer, Another
Image	Booking: Easy, Early-Bird, Possible, Priority, Pay	Service: Recommend, Class, Premium, Money, Price, Better, Extra	Flying: Best, Love, Experience, Always, First, Next, Definitely	Legroom: Short, Long, Time, Make, Sure, Right, Return, Take	Others: Travel, Economy, Business, Airline, Trip, Food, Many

〈표 3〉에 나타난 단어들은 각 차원으로 분류된 문장들의 말뭉치를 대상으로 5개의 주제를 기준으로 잠재 토픽 모델링(latent dirichlet allocation, LDA)을 수차례 진행하여 가장 선명하게 주제 분류가 가능한 시점에서 나타나는 단어들을 주제 별로 추출하여 나타낸 것이다.

3.3 감성 분석(Sentiment Analysis)

감성 분석은 연속된 단어 이면에 있는 감정적 톤(tone)을 결정하는 과정으로 온라인 텍스트에 표현된 사람들의 태도, 의견, 감정 등을 이해하는 데 활용된다(Tian et al., 2020). 이러한 감성 분석은 정성적인(qualitative) 온라인 리뷰를 정량적인

(quantitative) 감성 점수로 변경하여 다양한 계량적 분석이 가능하도록 해준다. 감성 분석을 위해 먼저 수집된 각 온라인 리뷰 텍스트를 문장별로 분해한 후(〈그림 1〉의 ⑦참조), 구축된 2차 분류기를 활용하여 분해된 문장을 AIRQUAL 차원별로 분류하였다(〈그림 1〉의 ⑧참조). 이렇게 분류된 각 문장별 감성 점수를 구하기 위해 먼저 각 문장에 포함된 단어에 대한 감성 점수를 구할 필요가 있다. 이를 위해 AFINN 영어 감성 단어 사전을 활용하였다. AFINN은 -5와 +5 사이의 정수로 평가된 영어 단어 목록으로 본 연구에서는 AFINN-96을 활용하였다. 각 문장별 감성 점수는 각 문장에 포함된 단어에 대한 감성 점수를 합산하는 방식으로 도출하였다. 이를 도식화하면 다음 식 (1)과 같다.

$$S_i = \sum_{i=1}^n (\text{문장내 단어별 감성점수}) \quad (1)$$

이때 S_i : 문장별 감성 점수
 n : 문장 내 단어의 수

이를 바탕으로 각 리뷰의 AIRQUAL 서비스 품질 차원별 감성 점수(R_i)를 도출하기 위해 문장별 감성 점수(S_i)를 서비스 품질 차원별로 합산한 후 각 차원에 포함되어 있는 문장의 수로 나누어 도출하였다(<그림 1>의 ⑨참조). 이를 도식화하면 다음 식 (2)와 같다. 이상의 각 리뷰 본문에 대한 감성 분석 절차를 요약하면 다음 <그림 2>와 같다.

$$R_i = \frac{\sum_1^N S_i}{N_i} \quad (2)$$

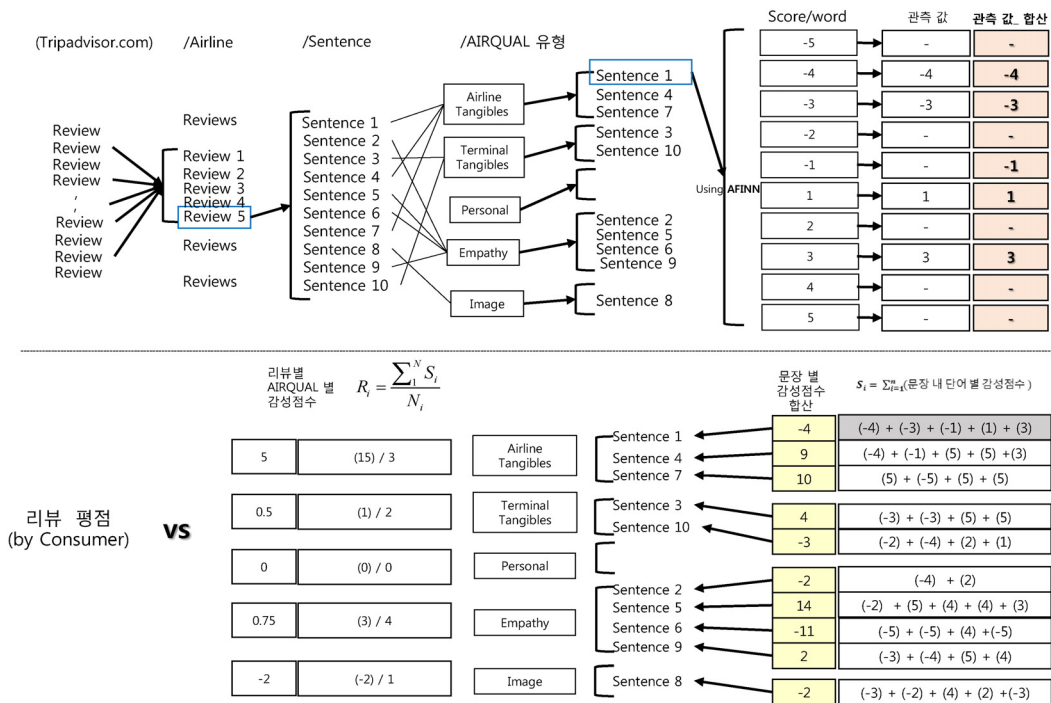
이때 S_i : 문장별 감성 점수
 N : 각 리뷰의 AIRQUAL 차원별 문장 수

리뷰 제목은 서비스 품질에 대한 소비자의 정서적 감정을 집약적으로 나타내는 감성 분석을 위한 정보의 중요한 원천이다(Zappavigna, 2015). 리뷰 제목에 대한 감성 점수(T_i)는 제목에 포함된 단어에 대한 감성 점수를 합산하는 방식으로 도출하였다. 이를 도식화하면 다음 식 (3)과 같다.

$$T_i = \sum_{i=1}^n (\text{제목내 단어별 감성점수}) \quad (3)$$

이때 T_i : 제목별 감성 점수
 n : 제목 내 단어의 수

리뷰 본문과 제목에 대한 감성 분석 결과를 요약하면 다음 <표 4>와 같다. AIRQUA의 각 차원별 감성 점수의 평균은 각각 1.234(항공사 유형), 0.151(터미널 유형), 1.907(직원), -0.104(공감), 1.064(이미지)로 나타났으며 리뷰 제목의 감성 점수 평균값 약 0.863으로 조사되었다. 이를 통해 공



<그림 2> 감성 분석 절차

〈표 4〉 감성 분석 통계량

차원	관측치	평균	표준편차	최소	최대
리뷰 별점	39,169	3.633	1.385	1	5
AIRQUAL 차원별 감성 점수					
항공사 유형	39,169	1.234	3.618	-19	35
터미널 유형	39,169	0.151	1.812	-17	21
직원	39,169	1.907	3.988	-23	31
공감	39,169	-0.104	3.693	-36	31
이미지	39,169	1.064	2.709	-18	26
리뷰 제목 감성 점수	39,169	0.863	2.309	-11	16

감성에 대한 평균 감성 점수가 가장 낮은 반면 직원에 대한 평균 감성 점수가 가장 높은 것을 알 수 있다. 또한, 본 연구에서는 항공사 서비스에 대한 만족도를 측정하기 위한 대리 변수(surrogate variable)로 tripadvisor.com 항공사 리뷰 별점(review star)을 사용하였다(최지은 등, 2018). 고객의 리뷰 별점은 항공사 서비스를 종합 평가한 후 고객이 해당 서비스에 대하여 부여하는 점수이기 때문에 서비스에 대한 종합적인 만족도를 평가한 것으로 간주할 수 있다. 리뷰 별점은 최저 1점에서 최고 5점까지의 값을 가지고 있으며 평균값은 약 3.633로 나타났다.

IV. 연구 모델

본 연구의 주요 목적은 온라인 리뷰로부터 항공사 서비스 품질을 AIRQUAL 서비스 차원으로 분류할 수 있는가와 이렇게 분류된 항공사의 서비스 품질이 만족도에 미치는 영향을 파악하는 데 있다. 이를 위해 서비스 만족도의 대리 변수인 리뷰 별점을 종속변수로 하고 AIRQUAL 차원별 감성 점수를 독립변수로 하여 회귀분석을 실시하였다(<그림 1>의 ①참조). 먼저 AIRQUAL 차원별 감성 점수가 리뷰 별점에 미치는 영향을 파악하기 위한 다음과 같은 모델 1을 설정하였다.

모델 1:

$$Rating_{i,j} = \beta_0 + \beta_1(Terminal\ Tangibles_{i,j}) + \beta_2(Airlines\ Tangibles_{i,j}) + \beta_3(Personnel_{i,j}) + \beta_4(Empathy_{i,j}) + \beta_5(Image_{i,j}) + \epsilon_i$$

Rating_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i*의 고객 별점

Terminal Tangibles_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i*의 터미널 유형성(terminal tangibles) 감성 점수

Airlines Tangibles_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i*의 항공기 유형성(airline tangibles) 감성 점수

Personnel_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i*의 직원(personnel) 감성 점수

Empathy_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i*의 공감(empathy) 감성 점수

Image_{i,j}: 항공사 *j*에 대한 온라인 리뷰 *i* 이미지(image)의 감성 점수

ϵ_i : 오차항

모델 1을 구성하는 변수 가운데 β_0 는 상수항을 의미한다. $\beta_1 \sim \beta_5$ 는 AIRQUAL의 다섯 가지 차원인 터미널 유형성, 항공기 유형성, 직원, 공감성,

이미지에 대한 리뷰 i 의 감성 점수가 항공사 j 에 대한 평가에 미치는 영향을 측정할 변수이며 마지막으로 ϵ_i 는 오차항으로 정규분포를 따르는 것으로 가정한다.

본 연구의 주요 특징 가운데 하나인 리뷰 제목의 감성 점수가 항공사 서비스 품질 평가에 미치는 영향을 파악하기 위하여 리뷰 제목을 변수로 추가하여 모델 2를 구축하였다.

모델 2:

$$\begin{aligned} \text{Rating}_{i,j} = & \beta_0 + \beta_1(\text{Terminal Tangibles}_{i,j}) \\ & + \beta_2(\text{Airlines Tangibles}_{i,j}) \\ & + \beta_3(\text{Personnel}_{i,j}) + \beta_4(\text{Empathy}_{i,j}) \\ & + \beta_5(\text{Image}_{i,j}) + \beta_6(\text{Review Title}_{i,j}) \\ & + \epsilon_i \end{aligned}$$

Review Title_{i,j}: 항공사 j 에 대한 온라인 리뷰 i 의 제목(*Review title*) 감성 점수

마지막으로 리뷰 제목이 AIRQUAL 서비스 품질 차원의 서비스 만족도에 미치는 영향에 대한 조절 효과를 파악하기 위해 각 서비스 품질 차원과 리뷰 제목의 곱을 추가하여 다음 모델 3을 구축하였다.

모델 3:

$$\begin{aligned} \text{Rating}_{i,j} = & \beta_0 + \beta_1(\text{Terminal Tangibles}_{i,j}) \\ & + \beta_2(\text{Airlines Tangibles}_{i,j}) \\ & + \beta_3(\text{Personnel}_{i,j}) + \beta_4(\text{Empathy}_{i,j}) \\ & + \beta_5(\text{Image}_{i,j}) + \beta_6(\text{Review Title}_{i,j}) \\ & + \beta_7(\text{Terminal Tangibles}_{i,j} \\ & \times \text{Review Title}_{i,j}) \\ & + \beta_8(\text{Airlines Tangibles}_{i,j} \\ & \times \text{Review Title}_{i,j}) \\ & + \beta_9(\text{Personnel}_{i,j} \times \text{Review Title}_{i,j}) \\ & + \beta_{10}(\text{Empathy}_{i,j} \times \text{Review Title}_{i,j}) \\ & + \beta_{11}(\text{Image}_{i,j} \times \text{Review Title}_{i,j}) \\ & + \epsilon_i \end{aligned}$$

리뷰 제목은 소비자의 주의를 끌 수 있는 매우 간결하며 강력한 장치이다(Chua and Banerjee, 2016). 리뷰 제목은 리뷰 본문에 관한 함축된 정보를 제공하기 때문에 소비자들은 리뷰 제목을 기반으로 어떤 리뷰를 읽을 것인가를 결정한다. 이러한 중요성으로 인해 몇몇 연구들이 리뷰 제목 역할에 대한 검증을 실시하였다(Biswas *et al.*, 2021; Chua and Banerjee, 2016). 그러나 이러한 노력에도 불구하고 아직까지 리뷰 제목의 역할에 대한 일치된 견해는 존재하지 않는다. Hu *et al.*(2014)은 리뷰 제목의 감성이 제품 판매 순위에 부정적인 영향을 미친다고 주장한 반면 Salehan and Kim(2016)은 리뷰 제목의 감성이 독자층(readership) 증대에 긍정적인 영향을 미친다고 주장하였다. 나아가 Zhou *et al.*(2020)은 리뷰 제목과 본문의 텍스트 유사성이 리뷰 유용성에 미치는 영향이 제목의 감성에 따라 달라진다고 주장하였다. 따라서 본 연구에서는 이러한 논의를 바탕으로 모델 3을 통해 서비스 서비스 품질 차원과 서비스 만족도 간의 관계에 있어 리뷰 제목의 조절 효과를 분석하고자 한다.

V. 분석 및 논의

5.1 AIRQUAL 서비스 차원과 만족도

온라인 리뷰를 AIRQUAL 서비스 품질 차원으로 분류하기 위한 학습을 위해서는 데이터에 대한 레이블링 과정이 필수적이다. 그러나 이러한 과정을 통해 레이블화 된 데이터는 데이터의 본질적인 모호성 또는 레이블링 오류 등과 같은 다양한 소스에서 비롯된 다양한 수준의 불확실성으로 인해 등분산성(homoscedasticity)의 가정을 위배할 가능성이 존재한다(Cao *et al.*, 2021). 따라서 브로이슈-파간 검정(Breusch-Pagan test)를 실시하여 이분산성(heteroscedasticity)이 존재하는지 분석하였다(Breusch and Pagan, 1979). 분석 결과 등분산성을 위배하는 것으로 나타났다(statistic = 123.506, $p < 0.01$). 이를 해결하기 위해 이분산성-일치 표준오

차(heteroskedastic-consistent standard errors)를 사용하여 분석을 실시하였다(White, 1980).

AIRQUAL의 서비스 품질 차원이 항공사 서비스에 대한 만족도에 미치는 영향을 파악하기 위하여 이분산성-일치 표준오차를 활용하여 모델 1을 분석하였다. 분석 결과 tripadvisor.com에서 평가되는 항공사의 사용자 리뷰 평점에 대한 AIRQUAL의 5가지 차원의 감성 점수가 모두 유의한 영향을 미치는 것으로 파악되었다(<표 5 >참조). 특히 직원(personnel)이 가장 강한 양의 관계를 나타내고 있는(비표준화 $\beta_3 = 0.104, p < 0.001$) 반면 터미널 유형성(terminal tangibles)의 경우 상대적으로 약한 양의 관계를 나타내고 있다(비표준화 $\beta_1 = 0.096, p < 0.001$). 모델 1에 전반적인 모형은 유의한 것으로 나타났다(Likelihood $\chi^2 = 15577.728, p < 0.001$). 또한, 공선성 분석 결과(VIF=1.024~1.060) 다중공선성(multicollinearity) 문제는 없는 것으로 나타났다. 모델 2를 통해 리뷰 제목의 감성 점수가 항공사

서비스 만족도에 미치는 영향을 분석하였다(<표 5 >참조). 분석 결과는 AIRQUAL의 서비스 차원과 추가 변수인 리뷰 제목 모두가 유의한 영향을 미치는 것으로 파악되었다. 특히 서비스 차원에 대한 감성 점수에 비해 리뷰 제목이 가장 강한 양의 관계를 미치고 있는 것으로 나타났다(비표준화 $\beta_6 = 0.215, p < 0.001$). 또한, 모형의 Likelihood χ^2 값(= 22189.286)이 통계적으로 유의한 값을 보임에 따라($p < 0.001$) 모형은 전반적으로 유의하다는 것을 확인하였다.

마지막으로 AIRQUAL의 다섯 가지 서비스 차원이 서비스 만족도에 미치는 영향이 리뷰 제목에 의해 어떻게 변화하는가를 파악하기 위해 모델 3에 대한 분석을 실시하였다(<표 5 > 참조). 분석 결과 AIRQUAL의 다섯 가지 차원 모두가 (터미널 유형성: 비표준화 $\beta_7 = -0.008$, 항공기 유형성: 비표준화 $\beta_8 = -0.012$, 직원: 비표준화 $\beta_9 = -0.011$, 공감성: 비표준화 $\beta_{10} = -0.002$, 이미지: 비표준화

<표 5> 이분산성(Heteroscedasticity)을 고려한 분석 결과

변수	모델1		모델2		모델3	
	비표준화 β	왈드 카이제곱 (Wald χ^2)	비표준화 β	왈드 카이제곱 (Wald χ^2)	비표준화 β	왈드 카이제곱 (Wald χ^2)
Constant	3.254***	191859.339	3.178***	214742.908	3.208***	202049.632
Terminal Tangibles	0.096***	702.424	0.071***	477.313	0.073***	422.873
Airline Tangibles	0.070***	1742.356	0.045***	845.318	0.061***	951.216
Personnel	0.104***	4499.830	0.078***	2871.266	0.087***	2587.182
Empathy	0.094***	2688.220	0.071***	1854.540	0.064***	1336.660
Image	0.084***	1479.074	0.058***	816.052	0.071***	784.533
Review title			0.215***	6704.662	0.262***	7204.473
Terminal Tangibles × Review title					-0.008***	40.334
Airline Tangibles × Review title					-0.012***	306.058
Personnel × Review title					-0.011***	380.007
Empathy × Review title					-0.002**	4.301
Image × Review title					-0.011***	184.722
Likelihood χ^2	15577.728***		22189.286***		23752.830***	

주) ** $p < 0.05$, *** $p < 0.001$.

$\beta_{11} = -0.011$) 리뷰 제목과 유의한 음의 상호 작용이 나타나는 것으로 파악되었다. Likelihood χ^2 값 (= 23751.830)을 통해 모형이 전반적으로 유의하다는 것을 파악하였다($p < 0.001$).

5.2 논의

본 논문은 항공사 서비스 품질 측정 지표인 AIRQUAL의 다섯 가지 차원(항공사 유형성, 터미널 유형성, 직원, 공감성, 이미지)의 감성 수준이 서비스 만족도에 미치는 영향에 대해 살펴보았다. 전통적 설문 방식의 한계점을 극복하기 위해 tripadvisor.com의 소비자 리뷰를 텍스트 마이닝 한 후 감성 분석을 통해 다섯 차원에 대한 감성 점수를 도출하였다. 회귀분석을 통해 AIRQUAL의 다섯 차원에 대한 감성 점수가 리뷰 별점으로 측정된 서비스 만족도에 모두 유의미한 양의 영향을 미치는 것을 파악하였다. 본 연구의 결과는 전통적 설문 방식을 활용한 기존 연구 결과와도 일치한다(Farooq *et al.*, 2018). 예를 들어, 본 연구 결과 다섯 가지 차원 가운데 직원이 서비스 만족도에 대한 가장 높은 예측력을(비표준화 $\beta_3 = 0.104$) 보이는 것으로 나타났다. 이러한 결과는 Ali *et al.* (2015)나 Nadiri *et al.*(2008) 등의 연구에서도 동일하게 확인할 수 있다. 또한, Farooq *et al.*(2018)의 연구와 유사하게 터미널 유형성의 경우 상대적으로 덜 중요하게 나타났다(비표준화 $\beta_1 = 0.096$).

본 연구에서는 또한 그 중요성에도 불구하고 기존 연구에서 도외시 되었던 리뷰 제목이 리뷰 별점에 미치는 영향을 분석하였다. 리뷰 제목은 사용자가 경험한 서비스에 대한 전체적인 감정을 응축적으로 드러내기 위해 매우 감정적인 언어로 작성되는 경향이 있다. 온라인 환경에서 이러한 감정적인 언어는 중립적인 언어와 비교할 때 보다 많은 주의(attention)와 피드백(feedback)을 받을 가능성이 있다(Huffaker, 2010). 따라서 감정적인 단어가 많이 포함되어 있는 리뷰 제목은 메시지의 효과적 전달을 가능하게 하여 결과적으로 다른 소

비자들의 인식에 매우 유의한 영향력을 미칠 가능성이 높다. 본 연구의 분석 결과는 리뷰 제목의 이러한 특성이 잘 나타나 있다. 즉, AIRQUAL의 다섯 가지 서비스 차원에 비해 서비스 만족도를 예측하는데 리뷰 제목이 오히려 더 중요한 변수로 나타났다(비표준화 $\beta_6 = 0.215$). 이러한 결과는 리뷰 제목이 리뷰의 유용성(helpfulness)에 미치는 영향을 파악한 기존 연구의 결과와도 어느 정도 일맥상통한다고 할 수 있다(Salehan and Kim, 2016).

본 연구의 분석 결과 AIRQUAL의 서비스 차원의 서비스 만족도에 미치는 영향에 대한 리뷰 제목의 조절 효과를 확인할 수 있었다. 항공기 유형성, 터미널 유형성, 직원, 공감성, 이미지 모두 리뷰 제목과 유의한 음의 상호 작용이 있는 것으로 파악되었다(<표 5> 참조). 이러한 흥미로운 결과는 선택적 주의 이론으로 설명 가능하다. 온라인 리뷰의 경우 그 양이 방대하기 때문에 모든 리뷰 본문을 읽는 것은 매우 어려울 뿐 아니라 많은 시간을 요구한다. 따라서 소비자들은 리뷰 제목을 기준으로 리뷰 본문을 계속 탐색할 것인지 아닌지를 결정하게 된다(Salehan and Kim, 2016). 따라서 이미 리뷰 제목을 읽은 소비자는 리뷰 본문에 대한 일정한 기대를 갖게 되며 그 결과 AIRQUAL 서비스 품질 차원이 서비스 만족도에 미치는 영향은 리뷰 제목으로 인해 감소하게 된다.

5.3 시사점

본 연구는 다음과 같은 점에서 학문적, 실무적 효과가 있을 것으로 기대한다. 먼저 학문적 측면에서 보면 본 연구는 첫째, 온라인 리뷰를 활용한 서비스 품질 평가 방안을 제시함으로써 전통적 서비스 품질 평가 방식의 단점을 보완할 수 있을 것으로 기대된다. 기존 서비스 품질 평가와 관련한 연구들은 설문이나 인터뷰와 같은 전통적 방식에 의하여 진행되어 왔다(Deveci *et al.*, 2018; Tsaur *et al.*, 2002). 이러한 방식은 서비스 차원별로 뚜렷하게 구분되는 질문을 통해 서비스 차원에 따른

서비스 품질을 측정할 수 있었다는 점에서 서비스 품질 연구에 일정 정도 기여한 것도 사실이다. 그러나 전통적 서비스 품질 평가 방식은 설문 응답자가 동료, 친구, 부모 등과 같은 타인에게 긍정적으로 평가받기를 원하는 경향 때문에 그 응답이 과장될 수 있어 신뢰성과 객관성 측면에서 문제가 있는 것도 사실이다. 또한, 소수의 응답자를 표본으로 조사를 진행하기 때문에 주관적인 편이가 발생할 수 있어 전체 고객의 의도를 반영하는 데 한계가 있다. 본 연구는 이러한 문제점을 해결할 수 있는 방안을 제시함으로써 서비스 품질 평가 연구의 지평을 넓혔다고 할 수 있다. 둘째, 온라인 리뷰 제목이 갖는 함의를 포괄할 수 있는 방법론을 제안함으로써 기존 온라인 리뷰 기반 서비스 품질 연구를 확대 발전시킬 수 있을 것이다. 최근 전통적 방식의 설문 기반 서비스 품질 평가 연구의 한계점을 극복하기 위하여 온라인 리뷰를 활용한 다양한 연구가 제시되고 있다(Tian *et al.*, 2020). 그러나 이러한 연구들은 온라인 리뷰 제목이 중요함에도 불구하고 온라인 리뷰 본문만을 분석에 활용하고 있다(Biswas *et al.*, 2021). 온라인 리뷰의 경우 제목과 본문으로 구성되어 있을 뿐 아니라 소비자들은 리뷰 제목을 확인한 후 추가적인 정보를 필요로 할 경우 본문을 읽는 경향이 있다는 점에 비추어 볼 때 온라인 리뷰 제목에 대한 분석은 매우 중요하다(Salehan and Kim, 2016). 본 연구에서는 온라인 리뷰 제목을 활용하여 이러한 기존 연구의 단점을 어느 정도 극복하였다. 셋째, 본 연구는 서비스 품질 차원 분류에 있어 신뢰성 있는 데이터의 수를 확보하는데 필요한 시간과 비용을 획기적으로 줄이고 이를 통해 적은 데이터로 인해 발생할 수 있는 과적합의 문제를 해결할 수 있는 자가 학습 방법을 도입함으로써 방법론적으로 기존 연구의 한계를 극복하였다(Dua and Chowriappa, 2013). 온라인 리뷰에 대한 소셜 어널리틱스 분석에 있어 기존 연구들은 사람의 개입에 의해 학습 데이터 셋이 구축되는 지도 학습에 의존하고 있다(Gao *et al.*, 2021; Riantama *et al.*, 2021). 그러나 사람의 개

입에 의한 학습 데이터 셋 구축에는 많은 시간과 비용이 소요된다. 이러한 시간과 비용을 줄이기 위해 사람의 개입을 감소시키게 되면 학습에 필요한 적절한 수준의 데이터를 확보할 수 없게 되며 그 결과 과적합과 같은 문제가 발생할 가능성이 높다. 본 연구는 자가학습과 감성 분석을 결합함으로써 이러한 문제점을 효율적이고 효과적으로 개선하였다. 마지막으로 항공산업에 특화된 모형을 활용하여 보다 정확한 서비스 품질 평가를 가능하게 함으로써 기존 연구를 확대 발전시키려고 노력하였다. 산업에 따라 그 특징이 다름에도 불구하고 기존 연구들은 주로 SERVQUAL 모형을 적용하여 서비스 품질을 평가하고자 하였다. 특히, 소셜 어널리틱스를 활용한 항공사 서비스 품질 측정 연구의 경우 모든 연구들이 SERVQUAL 모형에 의존하고 있다(Riantama *et al.*, 2021; Tian *et al.*, 2020). 이로 인해 항공산업의 특성을 정확하게 반영하고 있지 못하고 있는 것도 사실이다. 본 연구는 항공산업의 특성을 정확하게 반영할 수 있는 AIRQUAL 모형을 활용함으로써 이러한 문제점을 어느 정도 해결하려고 노력하였다.

실무적 관점에서 보면 본 연구는 첫째, 국내 항공산업 관계자로 하여금 서비스 만족도 개선을 위해 어떤 차원이 중요한가를 규명함으로써 효과적인 고객 만족 프로그램을 수행할 수 있는 가이드 라인을 제시하였다. 예를 들면, 본 연구 결과 직원이 서비스 만족도에 가장 높은 영향을 미치고 있는 것으로 나타났다. 따라서 양질의 직원을 고용하고 훈련시켜 소비자와 함께 공감할 수 있는 개인화된 서비스를 제공함으로써 서비스 만족도를 개선할 수 있을 것이다(Farooq *et al.*, 2018). 나아가 항공기 유형성을 높이기 위해 제공되는 음식료의 질을 개선하거나 비행기 내부의 위생 상태를 청결하게 함으로써 서비스 만족도를 높일 수 있을 것이다. 둘째, 본 연구에서 제시한 방법론을 기반으로 항공사 관계자는 자사의 서비스 품질 관리를 계량화하여 보다 정교하게 수행할 수 있을 것이다. 본 연구에서 제안한 감성 분석의 경우 AIRQUAL

차원별로 실제 해당 항공사를 경험한 고객의 감성을 지표화하여 제시하였다. 이를 바탕으로 다른 항공사와 자신의 항공사를 계량적으로 비교 분석할 수 있을 뿐 아니라 객관적 지표를 바탕으로 보다 효율적인 서비스 품질 관리가 가능할 것이다. 마지막으로 본 연구에서 제안한 방법론을 기반으로 효율적인 서비스 품질 측정이 가능할 것이다. 항공사 마케팅 담당자는 본 연구에서 제안한 방법론을 활용함으로써 전통적 설문조사 방식에 기반한 서비스 품질 측정과 동일하거나 혹은 유사한 효과를 거두면서 시간과 비용은 획기적으로 절감할 수 있을 것이다. 나아가 온라인 리뷰에 대한 정확한 분석을 통해 소비자의 피드백을 정확하게 분석하여 소비자 만족도를 증대시키고 궁극적으로 판매를 촉진할 수 있을 것이다(Ghose and Ipeirotis, 2010).

VI. 결 론

항공사의 서비스 품질 향상을 위한 효과적인 서비스 품질 측정은 오랜 연구과제이다. 특히 빅데이터 어널리틱스가 새로운 연구 패러다임으로 각광받게 됨에 따라 소비자가 직접 작성한 온라인 리뷰 분석을 통한 서비스 품질 측정이 점점 중요해지고 있다. 본 연구는 제목과 본문을 포함한 온라인 리뷰 전체를 자기학습과 감성 분석을 활용해 AIRQUAL 서비스 품질 차원으로 분류하고 이를 분석함으로써 보다 객관적이고 정교한 서비스 품질 측정을 시도하였으며 이를 기반으로 서비스 품질 차원이 서비스 만족도에 미치는 영향을 파악하였다. 분석 결과 AIRQUAL의 다섯 가지 서비스 품질 차원은 모두 서비스 만족도에 유의한 영향을 미치는 것으로 나타났다. 나아가 리뷰 제목이 서비스 만족도에 미치는 영향 또한 유의한 것으로 파악되었다. 본 연구는 항공산업의 특성을 반영한 서비스 품질 차원 측정 및 이의 효과에 대한 분석이라는 관점에서 학문 및 실무적 의의가 있다고 할 수 있다.

본 연구는 다음과 같은 한계점이 있으며 이를 해결하기 위한 향후 연구가 필요하다. 첫째, 비록 기존 연구들에서 리뷰 별점이 서비스 만족도를 측정하는 대리 변수로 활용되고 있으나(Korfiatis et al., 2019) 몇몇 연구의 경우 별점과 리뷰 내용 간의 불일치를 주장하고 있다(Shan et al., 2018). 따라서 서비스 만족도를 리뷰 별점이 아닌 유용성(Salehan and Kim, 2016) 또는 항공사 수준 평가 점수(airline quality rating)(Tian et al., 2020)와 같은 형태로 측정함으로써 보다 강건한(robust) 분석이 가능할 것이다. 둘째, 본 연구는 영어로 작성된 온라인 리뷰만을 대상으로 분석을 실시하였다. 따라서 본 연구 결과를 비영어권 국가에 그대로 적용하는 데 무리가 있다. 따라서 향후 연구에서는 다양한 언어를 포함함으로써 소비자 만족도에 영향을 미치는 보다 적절한 요인과 이유를 파악할 필요가 있다. 마지막으로 본 연구의 경우 국적, 성별, 나이 등과 같은 다양한 조절 변수를 고려하지 못했다. 이러한 조절 변수들은 연구 결과를 정확하게 이해하는 데 도움을 줄 수 있을 것이다. 예를 들면, 국적을 기반으로 문화적 배경에 따라 서비스 품질 차원이 만족도에 미치는 영향이 어떻게 달라지는 가를 파악하는 것은 매우 흥미로운 연구가 될 것이다.

참 고 문 헌

- [1] 최지은, 김소담, 김희웅, “의료 서비스 리뷰의 감성 수준이 병원 평가에 미치는 영향 분석”, *Information Systems Review*, 제20권, 제2호, 2018, pp. 111-137.
- [2] Abdel Rady, H., “Measuring airline service quality using AIRQUAL model: A study applied to Egyptair”, *International Journal of Heritage, Tourism and Hospitality*, Vol.12, No.1, 2018, pp. 271-290.
- [3] Aksoy, S., E. Atilgan, and S. Akinci, “Airline services marketing by domestic and foreign firms: Differences from the customers’ viewpoint”,

- Journal of Air Transport Management*, Vol.9, No.6, 2003, pp. 343-351.
- [4] Ali, F., B. L., Dey, and R. Filieri, “An assessment of service quality and resulting customer satisfaction in Pakistan international airlines: Findings from foreigners and overseas Pakistani customers”, *International Journal of Quality & Reliability Management*, Vol.32, No.5, 2015, pp. 486-502.
- [5] Ascaniis, S. and U. Gretzel, “What’s in a travel review title?”, in: Fuchs, M., F. Ricci, and L. Cantoni (eds.), *Information and Communication Technologies in Tourism*, Springer, Vienna, 2012, pp. 494-505.
- [6] Avery, C., P. Resnich, and R. Zeckhauser, “The market for evaluations”, *American Economic Review*, Vol.89, No.3, 1999, pp. 564-584.
- [7] Bailey, A. A., “Consumer awareness and use of product review websites”, *Journal of Interactive Advertising*, Vol.6, No.1, 2005, pp. 68-81.
- [8] Biswas, B., P. Sengupta, and B. Ganguly, “Your reviews or mine? Exploring the determinants of, “perceived helpfulness” of online reviews: A cross-cultural study”, *Electronic Markets*, 2021, in Press.
- [9] Boetsch, T., T. Bieger, and A. Wittmer, “A customer-value framework for analyzing airline services”, *Transportation Journal*, Vol.50, No.3, 2011, pp. 251-270.
- [10] Breusch, T. S. and A. R. Pagan, “A simple test for heteroskedasticity and random coefficient variation”, *Econometrica*, Vol.47, No.1979, pp. 1287-1294.
- [11] Cao, K., Y. Chen, J. Lu, N. Arechiga, A. Gaidon, and T. Ma, “Heteroskedastic and imbalanced deep learning with adaptive regularization”, *Proceedings of the International Conference on Learning Representations*, 2021, <https://openreview.net/forum?id=mEdwVCRJuX4>.
- [12] 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.
- [13] Carman, J. M., “Consumer perceptions of service quality: An assessment of the SERVQUAL dimensions”, *Journal of Retailing*, Vol.66, No.1, 1990, pp. 33-55.
- [14] Chen, Y. and J. Xie, “Online consumer review: Word-of-mouth as a new element of marketing communication mix”, *Management Science*, Vol.54, No.3, 2008, pp. 477-491.
- [15] Choi, J., A. Menon, and H. Tabakovic, “Using machine learning to revisit the diversification-performance relationship”, *Strategic Management Journal*, Vol.42, No.9, 2021, pp. 1632-1661.
- [16] Chua, A. Y. and S. Banerjee, “Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality”, *Computers in Human Behavior*, Vol.54, 2016, pp. 547-554.
- [17] Demir, H. H., “Evaluation of service quality of airway companies giving domestic services in Turkey with fuzzy set approach”, *International Journal of Electronics, Mechanical and Mechatronics Engineering*, Vol. 2, No. 3, 2012, pp. 233-239.
- [18] Deveci, M., E. Özcan, R. John, and S. C. Öner, “Interval type-2 hesitant fuzzy set method for improving the service quality of domestic airlines in Turkey”, *Journal of Air Transport Management*, Vol.69, 2018, pp. 83-98.
- [19] Dua, S. and P. Chowriappa, *Data Mining for Bioinformatics*, CRC Press, Boca Raton, FL, 2013.
- [20] Duan, W., B. Gu, and A. B. Whinston, “Do online reviews matter?: An empirical investigation of

- panel data”, *Decision Support Systems*, Vol.45, No.4, 2008, pp. 1007-1016.
- [21] Duan, W., Y. Yu, Q. Cao, and S. Levy, “Exploring the impact of social media on hotel service performance: A sentimental analysis approach”, *Cornell Hospitality Quarterly*, Vol.57, No.3, 2016, pp. 282-296.
- [22] Ekiz, H. E., K. Hussain, and A. Bavik, “Perceptions of service quality in North Cyprus national airline”, *Proceedings of the International Conference on Tourism and Hospitality Industry*, 2006, pp. 778-790.
- [23] Fan, W. and M. D. Gordon, “The power of social media analytics”, *Communications of the ACM*, Vol.57, No.6, 2014, pp. 74-81.
- [24] Farooq, M. S., M. Salam, A. Fayolle, N. Jaafar, and K. Ayupp, “Impact of service quality on customer satisfaction in Malaysia airlines: A PLS-SEM approach”, *Journal of Air Transport Management*, Vol.67, 2018, pp. 169-180.
- [25] Gao, K., Y. Yang, and X. Qu, “Examining non-linear and interaction effects of multiple determinants on airline travel satisfaction”, *Transportation Research Part D: Transport and Environment*, Vol.97, 2021, Article 102957.
- [26] Ghose, A. and P. G. Ipeiritos, “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, 2010, pp. 1498-1512.
- [27] Gottschalk, S. A. and A. Mafael, “Cutting through the online review jungle: Investigating selective eWOM processing”, *Journal of Interactive Marketing*, Vol. 37, 2017, pp. 89-104.
- [28] Hasan, M., M. N. Khan, and R. Farooqi, “Service quality measurement models: Comparative analysis and application in airlines industry”, *Global Journal of Enterprise Information System*, Vol.11, No.2, 2019, pp. 29-41.
- [29] 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.
- [30] Huffaker, D., “Dimensions of leadership and social influence in online communities”, *Human Communication Research*, Vol. 36, No. 4, 2010, pp. 593-617.
- [31] Hussain, R., A. Al Nasser, and Y. K. Hussain, “Service quality and customer satisfaction of a UAE-based airline: An empirical investigation”, *Journal of Air Transport Management*, Vol.42, 2015, pp. 167-175.
- [32] Khalsa, S. J. S., C. A. Mattmann, and R. Duerr, “Deep web crawling for insights from polar data”, *Proceedings of the International Geoscience and Remote Sensing Symposium*, 2017, pp. 376-379.
- [33] Kocak, B. B. and Ö. Atalik, “Perceptual maps of Turkish airline services for different periods using supervised machine learning approach and multidimensional scaling”, *International Journal of Sustainable Aviation*, Vol.5, No.3, 2019, pp. 205-229.
- [34] Korfiatis, N., P. Stamolampros, P. Kourouthanassis, and V. Sagiainos, “Measuring service quality from unstructured data: A topic modeling application on airline passengers’ online reviews”, *Expert Systems with Applications*, Vol. 116, 2019, pp. 472-486.
- [35] Lee, K. and C. Yu, “Assessment of airport service quality: A complementary approach to measure perceived service quality based on Google reviews”, *Journal of Air Transport Management*, Vol.71, 2018, pp. 28-44.
- [36] Li, W., S. Yu, H. Pei, C. Zhao, and B. Tian, “A hybrid approach based on fuzzy AHP and

- 2-tuple fuzzy linguistic method for evaluation in-flight service quality”, *Journal of Air Transport Management*, Vol.60, 2017, pp. 49-64.
- [37] Lin, J. W., F. Wang, and P. Chu, “Using semantic similarity in crawling-based web application testing”, *Proceedings of the International Conference on Software Testing, Verification and Validation*, 2017, pp. 138-148.
- [38] Liou, J. J., C. Y. Tsai, R. H. Lin, and G. H. Tzeng, “A modified VIKOR multiple-criteria decision method for improving domestic airlines service quality”, *Journal of Air Transport Management*, Vol. 17, No. 2, 2011, pp. 57-61.
- [39] Liu, B., “Sentiment analysis and opinion mining,” *Synthesis Lectures on Human Language Technologies*, Vol.5, No.1, 2012, pp. 1-167.
- [40] Lucini, F. R., L. M. Tonetto, F. S. Fogliatto, and M. J. Anzanello, “Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews”, *Journal of Air Transport Management*, Vol.83, 2020, pp. 101760.
- [41] Mudambi, S. M. and D. Schuff, “What makes a helpful review? A study of customer reviews on amazon.com”, *MIS Quarterly*, Vol.34, No.1, 2010, pp. 185-200.
- [42] Nadiri, H., K. Hussain, E. H. Ekiz, and S. Erdogan, “An investigation on the factors influencing passengers’ loyalty in the North Cyprus national airline”, *The TQM Journal*, Vol.20, No.3, 2008, pp. 265-280.
- [43] Nakayama, M. and Y. Wan, “The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews”, *Information & Management*, Vol.56, No.2, 2019, pp. 271-279.
- [44] O’Connor, P., “Managing a hotel’s image on TripAdvisor”, *Journal of Hospitality Marketing & Management*, Vol.19, No.7, 2010, pp. 754-772.
- [45] Pakdil, F. and Ö. Aydin, “Expectations and perceptions in airline services: An analysis using weighted SERVQUAL scores”, *Journal of Air Transport Management*, Vol.13, No.4, 2007, pp. 229-237.
- [46] Parasuraman, A., V. A. Zeithaml, and L. L. Berry, “SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality”, *Journal of Retailing*, Vol.64, No.1, 1988, pp. 12-40.
- [47] Park, E., Y. Jang, J. Kim, N. J. Jeong, K. Bae, and A. P. del Pobil, “Determinants of customer satisfaction with airline services: An analysis of customer feedback big data”, *Journal of Retailing and Consumer Services*, Vol.51, 2019, pp. 186-190.
- [48] Park, J. W., “Passenger perceptions of service quality: Korean and Australian case studies”, *Journal of Air Transport Management*, Vol.13, No.4, 2007, pp. 238-242.
- [49] Park, S., J. S. Lee, and J. L. Nicolau, “Understanding the dynamics of the quality of airline service attributes: Satisfiers and dissatisfiers”, *Tourism Management*, Vol.81, 2020, Article 104163.
- [50] Park, S., J. S. Lee, and J. L. Nicolau, “Understanding the dynamics of the quality of airline service attributes: Satisfiers and dissatisfiers”, *Tourism Management*, Vol.81, 2020, pp. 104-163.
- [51] Parshall, C. G., P. D. B. Houghton, and J. D. Kromrey, “Equating error and statistical bias in small sample linear equating”, *Journal of Educational Measurement*, Vol.32, No.1, 1995, pp. 37-54.
- [52] Perçin, S., “Evaluating airline service quality using a combined fuzzy decision-making approach”, *Journal of Air Transport Management*, Vol.68, 2018, pp. 48-60.
- [53] Ranard, B. L., R. M. Werner, T. Antanavicius,

- H. A. Schwartz, R. J. Smith, F. Meisel, D. A. Asch, H. Ungar, and R. M. Merchant”, Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care”, *Health Affairs*, Vol.35, No.4, 2016, pp. 697-705.
- [54] Riantama, D., A. Tsai, S. Lin, and H. Lin, “Evaluating airline passengers’ satisfaction during the COVID-19 pandemic: A text mining approach”, *Proceeding of the International Conference of Information Communication Technologies Enhanced Social Sciences and Humanities*, 2021, <https://doi.org/10.21428/7a45813f.d5497ab0>.
- [55] Salehan, M. and D. J. Kim, “Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics”, *Decision Support Systems*, Vol.81, 2016, pp. 30-40.
- [56] Saputro, B., A. N. Hidayanto, Z. Abidin, and N. Paoprasert, “Measuring service quality in the telecommunications industry from customer reviews using sentiment analysis: A case study in PT XL Axiata”, *International Journal of Innovation and Learning*, Vol.30, No.2, 2021, pp. 188-200.
- [57] Sezgen, E., K. J. Mason, and R. Mayer, “Voice of airline passenger: A text mining approach to understand customer satisfaction”, *Journal of Air Transport Management*, Vol.77, 2019, pp.65-74.
- [58] Shan, G., D., Zhang, L. Zhou, L. Suo, J. Lim, and C. Shi, “Inconsistency investigation between online review content and ratings”, *Proceedings of the Americas Conference on Information Systems*, 2018, <https://aisel.aisnet.org/amcis2018/eBusiness/Presentations/23/>.
- [59] Stamolampros, P., N. Korfiatis, P. Kourouthanassis, and E. Symitsi, “Flying to quality: Cultural influences on online reviews”, *Journal of Travel Research*, Vol.58, No.3, 2019, pp. 496-511.
- [60] Statista, Annual growth in global air traffic passenger demand from 2006 to 2022, 2021, <https://www.statista.com/statistics/193533/growth-of-global-air-traffic-passenger-demand/>.
- [61] Suki, N. M., “Passenger satisfaction with airline service quality in Malaysia: A structural equation modeling approach”, *Research in Transportation Business & Management*, Vol.10, 2014, pp. 26-32.
- [62] Sukwadi, R., A. Susanto, and Y.C. Liang, “Airline service quality evaluation for Indonesian low-cost carriers based on Extenics innovation theory”, *Archives of Transport*, Vol.58, No.2, 2021, pp. 7-20.
- [63] Thompson, P., G. DeSouza, and B. T. Gale, “The strategic management of service quality”, *Quality Progress*, Vol.18, No.6, 1985, pp. 20-25.
- [64] Tian, X., W. He, C. Tang, L. Li, H. Xu, and D. Selover, “A new approach of social media analytics to predict service quality: Evidence from the airline industry”, *Journal of Enterprise Information Management*, Vol.33, No.1, 2020, pp. 51-70.
- [65] Tirunillai, S. and G. J. Tellis, “Mining marketing meaning from online chatter: Strategic brand analysis of big data using Latent Dirichlet Allocation”, *Journal of Marketing Research*, Vol.51, No.4, 2014, pp. 463-479.
- [66] Treisman, A.M., “Strategies and models of selective attention”, *Psychological Review*, Vol.76, No.3, 1969, pp. 282-299.
- [67] Tsaur, S. H., T. Y. Chang, and C. H. Yen, “The evaluation of airline service quality by fuzzy MCDM”, *Tourism Management*, Vol.23, No.2, 2002, pp. 107-115.
- [68] Wang, C. Y. and Y. Wei, “Social media peer communication and impacts on Purchase intentions: A consumer socialization framework”,

- Journal of Interactive Marketing*, Vol.26, pp. 198-208.
- [69] White, H., “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”, *Econometrica*, Vol.48, No.4, 1980, pp. 817-838.
- [70] Wu, H. C. and C. C. Cheng, “A hierarchical model of service quality in the airline industry”, *Journal of Hospitality and Tourism Management*, Vol.20, 2013, pp. 13-22.
- [71] Xiang, Z., Q. Du, Y. Ma, and W. Fan, “A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism”, *Tourism Management*, Vol. 58, 2017, pp. 51-65.
- [72] Yang, Z., R. T. Peterson, and S. Cai, “Services quality dimensions of Internet retailing: An exploratory analysis”, *Journal of Services Marketing*, Vol.17, No.7, 2003, pp. 685-700.
- [73] Zappavigna, M., “Searchable talk: The linguistic functions of hashtags”, *Social Semiotics*, Vol.25, No.3, 2015, pp. 274-291.
- [74] Zhu, X. and A. B. Goldberg, “Introduction to semi-supervised learning”, *Synthesis Lectures on Artificial Intelligence and Machine Learning*, Vol.3, No.1, 2009, pp. 1-130.
- [75] Zhou, Y., S. Yang, Y. Li, Y. Chen, J. Yao, and A. Qazi, “Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining”, *Information Processing & Management*, Vol.57, No.2, 2020, Article 102179.

Service Quality Evaluation based on Social Media Analytics: Focused on Airline Industry

Myoung-Ki Han* · Byounggu Choi**

Abstract

As competition in the airline industry intensifies, effective airline service quality evaluation has become one of the main challenges. In particular, as big data analytics has been touted as a new research paradigm, new research on service quality measurement using online review analysis has been attempted. However, these studies do not use review titles for analysis, rely on supervised learning that requires a lot of human intervention in learning, and do not consider airline characteristics in classifying service quality dimensions. To overcome the limitations of existing studies, this study attempts to measure airlines service quality and to classify it into the AIRQUAL service quality dimension using online review text as well as title based on self-training and sentiment analysis. The results show the way of effectively extracting service quality dimensions of AIRQUAL from online reviews, and find that each service quality dimension has a significant effect on service satisfaction. Furthermore, the effect of review title on service satisfaction is also found to be significant. This study sheds new light on service quality measurement in the airline industry by using an advanced analytical approach to analyze the effects of service quality on customer satisfaction. This study also helps managers who want to improve customer satisfaction by providing high quality service in the airline industry.

Keywords: *Airline Service Quality Evaluation, Social Media Analytics, AIRQUAL Model, Self-Training, Sentiment Analysis*

* Ph.D. Candidate, Department of Data Science, Kookmin University

** Corresponding Author, Professor, College of Business Administration, Kookmin University

◎ 저 자 소 개 ◎



한 명 기 (hmkcmk@gmail.com)

국민대학교 데이터사이언스학과에서 석사학위를 취득하였으며, 현재 동대학원 박사과정 중이다. 주요 관심분야는 데이터사이언스, 소셜 어널리틱스 등이며 경영정보학회 학술대회에서 다수의 논문을 발표하였다.



최 병 구 (h2choi@kookmin.ac.kr)

현재 국민대학교 경영대학 경영학부 교수로 재직 중이다. KAIST경영공학 석사 및 박사학위를 취득하였다. 국민대학교에 부임하기 전에는 University of Sydney, School of Information Technologies에서 조교수로 재직하였다. 주요 연구분야는 지식경영, 소셜미디어 어널리틱스, 데이터사이언스 등이며 지금까지 이와 관련하여 Journal of Association for the Information Systems, Journal of MIS, IEEE Transactions on Engineering Management, I&M, APJIS, 지식경영연구 등을 포함한 다수의 국내 외 학술지에 논문을 게재하였다.

논문접수일 : 2021년 12월 28일

게재확정일 : 2022년 02월 21일

1차 수정일 : 2022년 02월 17일