



www.akht.or.kr

Journal of Hospitality & Tourism Studies

The Academy of Korea Hospitality & Tourism



토픽모델링을 통한 국내 대형항공사들의 서비스품질 비교: 트립어드바이저 리뷰를 중심으로

Comparison of service quality of full service carriers in Korea using topic modeling: based on reviews from TripAdvisor

조민경* · 이병주**

Cho, Min-Kyoung · Lee, Byung-Joo

* 충북대학교 대학원 경영학과 박사과정, cmr0120@naver.com, 관심분야: 서비스교육, 인적자원관리

** 경희대학교 대학원 스마트관광원 박사수료, bylee@khu.ac.kr, 관심분야: 관광마케팅, 관광빅데이터(교신저자)

요약

최근 대한항공이 아시아나항공을 인수한다는 공식 발표가 나왔다. 그러나 기대와 달리 두 대형항공사의 외형적 통합이 바로 경쟁력 확대로 이어지지 않을 수도 있기에 두 항공사의 서비스 경쟁력들을 비교, 분석하여 시너지를 창출할 수 있는 노력이 더해진다면 국내 항공 산업의 발전은 지금의 위기를 넘어 새로운 도약을 할 수 있을 것이다. 이에 본 연구에서는 고객들이 직접 생성한 콘텐츠인 온라인 리뷰를 바탕으로 고객만족에 영향을 미칠 수 있는 중요한 서비스품질 요소들을 토픽모델링을 활용하여 도출한 후 두 항공사 간 비교분석을 시행하고자 하였다. 이를 위해 트립어드바이저의 온라인 리뷰 중 영어로 작성된 리뷰만을 대상으로 파이썬 개발언어를 이용하여 자동 웹크롤링 프로그램을 개발하였다. 이렇게 수집한 자료는 총 4,763개로 대한항공이 3,147개, 아시아나항공이 1,616개이었다. 수집한 자료는 파이썬 프로그램을 이용해서 텍스트 사전처리를 수행한 후 잠재 디리클레 할당 방법론을 통해 두 항공사 모두 각 7개씩의 토픽을 산출하였고, 각 도출된 토픽 안에 분포된 단어를 바탕으로 토픽의 주제를 결정하였다. 이렇게 최종적으로 결정된 주제어들 중 두 항공사에서 공통적으로 나타난 주제어는 '고객 서비스', '기내 엔터테인먼트', '비즈니스 좌석', '승무원', '좌석 공간' 이었고, 두 항공사에서 서로 다르게 나타난 주제어는 '일등석', '환승 서비스', '일반석', '스타 얼라이언스' 가 있었다. 이러한 서비스품질 요소의 공통점과 차이점을 바탕으로 두 항공사에서 공통적으로 나타난 서비스품질 요소들은 계속적으로 품질을 유지 및 강화시켜야하며, 두 항공사에서 다르게 나타난 서비스 품질요소들은 통합된 항공사의 경쟁력 강화를 위해 잘 조화시키는 노력이 필요할 것이다.

핵심용어

토픽모델링,
대형항공사,
서비스품질,
트립어드바이저 리뷰,
잠재 디리클레 할당

ABSTRACT

Recently, an official announcement was made that Korean Air will take over Asiana Airlines. However, contrary to expectations, the external merger of the two full service carriers may not lead to increased competitiveness, so the development of the domestic aviation industry could take a new leap beyond the current crisis if efforts to compare and analyze service competitiveness of the two airlines are added. In this study, based on online reviews, which are contents directly generated by users, important service quality factors that can affect customer satisfaction were derived using topic modeling, and then a comparative analysis was conducted between the two airlines. To this end, an automated web crawling program was developed using Python, targeting only English-written reviews among TripAdvisor's online reviews. A total of 4,763, with 3,147 Korean Air and 1,616 Asiana Airlines. The collected data were followed by text preprocessing using Python, and the latent dirichlet allocation methodology produced seven topics for both airlines, and the names of each topic were determined based on the words distributed within each derived topic. Among the finally decided name of topics, the topics common to the two airlines were 'customer service', 'inflight entertainment', 'business class', 'cabin crew', and 'leg room'. The two airlines' different topics were 'first class', 'transfer service', 'economy class', and 'star alliance'. Based on the similarities and differences between these service quality factors, the factors common to both airlines will continue to maintain and enhance service quality, and the differences will need to be well harmonized to enhance the competitiveness of the airline after merger.

KEYWORDS

topic modeling,
full service carriers,
service quality,
TripAdvisor reviews,
latent dirichlet allocation

ACCEPTANCE INFO.

2021.02.02. 원고 접수
2021.02.17. 1차 수정본 접수
2021.02.25. 최종 수정본 접수
2021.02.25. 최종 게재 확정
3인 익명심사필

I. 서론

상품이나 서비스에 대한 고객들의 주관적인 평가인 고객만족(Oliver, 1980)은 고객들의 긍정적인 후기와 재 구매, 추천 등과 같은 행동을 유도하는 중요한 선행적 요인이라는 결과(Kim et al., 2009)가 다양한 연구에서 밝혀지고 있다. 특히 관광분야에서도 관광객들의 행동을 판단할 수 있는 중요한 요인 중 하나를 고객만족으로 보고 관광객들의 만족을 높일 수 있는 선행요인들에 대한 탐구가 계속되고 있다. 그리고 이러한 관심은 치열한 경쟁에 놓여 있는 항공 산업에서도 두드러지게 나타나고 있어 항공사의 서비스품질(임희진 · 임근욱, 2018), 객실승무원들의 인적 서비스(박윤미, 2017), 기내 서비스품질(정윤지 · 김미정, 2017) 등과 같이 고객만족에 영향을 미칠 수 있는 선행요소를 밝히려는 다양한 연구가 지속되고 있다.

이렇게 고객만족과 관련된 여러 선행요소들 중에서 실질적으로 항공사에 대한 고객들의 애호도 향상과 점유율 증대 등에 기여하는 요소는 항공사의 서비스품질이고 충분한 가치의 서비스품질은 고객들이 지불한 비용대비 만족을 높일 수 있는 효율적인 방안이라는 연구결과가 도출되었다(Morash & Ozment, 1994). 이에 많은 항공사 서비스품질관련 연구들에서 항공지연, 고객불만, 수화물 처리, 항공안전, 고객만족 지수 등과 같은 서비스품질 수행 척도들을 적용하여 평가하고 있다(Korfiatis et al., 2019). 이와 같은 대부분의 선행연구들은 설문방식으로 구조화된 자료를 수집(Li et al., 2017)하고, 회귀분석이나 구조방정식 모델을 이용하여 이들 요인들 간의 영향관계를 검증하고 있다. 그러나 구조화된 자료를 바탕으로 한 실증연구에서는 미리 규정된 설문지의 측정척도만을 사용하기 때문에 연구의 범위가 한정되어 있고(Korfiatis et al., 2019) 설문조사의 비용과 복잡한 실행, 그리고 설문조사에 따른 복합적인 편향이 존재(Alaei et al., 2019)한다는 단점 등이 지적되고 있다.

이에 따라 설문자료를 보완 또는 대체하기 위하여 온라인 리뷰(review)와 같은 사용자 생성 콘텐츠들인 텍스트와 이미지, 그리고 동영상 등의 빅데이터(big data)를 활용한 연구들이 항공 산업관련 연구에서 등장하고 있다. Ban & Kim (2019)의 연구에서는 세계 10여개 항공사에 대한 온라인 리뷰를 수집하여 전체적으로 리뷰 속에서 많이 나타난 단어들과 이러한 단어들 사이의 의미연결을 알아보기 위한 의미연결망분석을 실시하였다. 이를 통해 좌석에 대한 편안함, 지상 서비스, 비용대비 가치, 항공사 브랜드, 항공사 직원들, 그리고 항공기에서의 즐거움 등의 항공 서비스품질 요소들이 실질적으로 고객만족과 직접적으로 관련되어 있다는 시사점을 제시하였다. Korfiatis et al.(2019)은 트립어드바이저(TripAdvisor)의 항공사 리뷰를 모두 수집하여 토픽모델링 기법인 구조적 토픽 모델(structural topic models: STM)과 잠재 디리클레 할당(latent dirichlet allocation: LDA)를 적용하여 분석하였다. 그 결과 항공사 서비스품질 요소로 비즈니스 좌석, 수화물 정책, 좌석 공간, 항공사 직원들, 항공지연 등의 20개 토픽을 구체적으로 제시하였다. 이와 같이 빅데이터를 활용한 항공사 서비스품질에 대한 연구에서는 자유로운 형식의 비구조화된 빅데이터 자료들 속에서 고객들이 인식하는 항공사 서비스품질들의 구체적인 요인을 도출하고 그 속에 내재되어 있는 정보를 추가적으로 밝혀내어 설문방식의 문제점들을 일부 극복하는 특징을 가지고 있다(Korfiatis et al., 2019).

최근 국내 항공 산업은 코로나19로 인해 항공사들의 어려움이 지속되는 상황 속에서 대한항공이 아시아나항공을 인수한다는 공식 발표가 나왔다. 이러한 합병은 외형적으로 자산 40조의 전 세계 7위 항공사가 탄생한다는 의미이다(매일경제신문, 2020. 11. 16). 그러나 대한항공이 갖고 있는 경쟁력과 아시아나항공이 보유한 경쟁력에는 일부 차이가 있을 수 있고 두 항공사에서 느끼던 고객들의 만족 요소에도 일부 차이가 있을 수 있기 때문에 합병으로 인한 외형적 확대가 바로 경쟁력의 확대로 이어진다는 의미로는 볼 수 없을 것이다. 두 항공사의 고유한 경쟁력 차이는 하나의 항공사로 통합되면서 교집합으로 축소되는 것이 아니라 두 항공사의 경쟁력이 더해져 시너지를 창출해야만 경쟁력이 강화될 수 있다. 이를 위해 앞에서 살펴본 온라인 리뷰의 자료 특성을 바탕으로 고객만족을 위한 두 항공사의 서비스품질 특징을 도출해낼 수 있다면 의미 있는 시사점들을 얻을 수 있을 것이라 사료된다.

이에 본 연구에서는 선행연구들을 바탕으로 사용자 생성 콘텐츠 중의 하나인 온라인 리뷰로 고객만족에 영향을 미칠 수 있는 함축적 서비스품질 요인들을 토픽모델링 방법론으로 추출하고자 한다. 이를 통해 더욱 정확하고 효과적으로 고객의 소리에 귀를 기울이고 항공사 입장에서 실질적인 도움이 될 수 있는 시사점을 도출하려고 한다. 이러한 목적을 위해 먼저 온라인 리뷰는 관광분야 리뷰 웹사이트 중에서 많은 국내외 연구에서 자주 활용되고 있는 트립어드바이저로부터 자료를 수집하였다. 연구대상은 국내 대형항공사인 대한항공과 아시아나항공으로만 한정하였다. 영어, 한글과 일본어 등의 다양한 텍스트 기반 리뷰 중에서 외국인 고객들이 작성했을 것으로 추정할 수 있는 영어 리뷰만을 추출하여 연구에 적용하였다. 이렇게 확보한 자료를 바탕으로

본 연구는 다음과 같은 구성으로 이루어지고 있다. 먼저 본 연구에서 적용한 토픽모델링 방법론에 대한 학술적 배경을 살펴본 후 자료수집 및 수집된 자료의 처리과정을 설명한다. 그리고 토픽모델링 방법을 적용하여 대한항공과 아시아나항공의 고객만족관련 토픽들을 각각 추출하고 토픽과 연관된 적합한 주제어들을 선정한다. 마지막으로 두 항공사 주제어들의 공통점과 차이점을 바탕으로 학술적 및 실무적 시사점들을 제시하고자 한다.

II. 이론적 배경

1. 토픽모델링

토픽모델링은 머신러닝 비지도 학습(unsupervised learning) 기법 중 하나로 많은 양의 텍스트를 몇 개 차원의 군집으로 축소한 후 각 차원에 속한 단어들이 나타내는 함축된 의미를 도출하는 분석기법이다(김지연 등, 2016). 토픽모델링의 구체적인 알고리즘(algorithm)으로 많이 사용되는 기법이 잠재의미분석(latent semantic analysis: LSA), 확률론적 잠재의미분석(probabilistic latent semantic analysis: PLSA)(Hofmann, 2001)과 LDA(Blei et al., 2003)라고 할 수 있다.

LSA는 행렬구조로 만든 텍스트 데이터를 선형대수학을 응용한 알고리즘을 적용하여 행렬을 축소시켜나가면서 함축된 의미를 도출해내는 기법이다(Sezgen et al., 2019). 그러나 LSA에는 확률의 개념이 적용되지는 않았고 PLSA부터 확률을 적용한 토픽모델링 기법이 등장하여 LDA 알고리즘까지 이어지게 되었다. LDA는 대량의 비구조적인 텍스트 데이터를 사람의 개입을 최소화하여 자동적으로 연관성이 높은 단어들을 기준으로 k개의 토픽으로 도출하여 많은 양의 텍스트를 요약된 정보로 축소하는 방법이다(Guo et al., 2017). 즉 LDA는 문서 내에 특정 용어가 등장하는 횟수 기준이 아닌 텍스트 내에 특정 용어가 등장할 확률을 기반으로 하고 있기 때문에 의미를 더욱 정교하게 이끌어낼 수 있다(남승주 · 이현철, 2019). 이러한 LDA는 확률을 통한 토픽 도출의 시작점인 PLSA의 알고리즘을 바탕으로 하지만 PLSA의 문제점인 새로운 문헌이 추가됐을 경우 잘 구분하지 못한다는 단점을 디리클레(dirichlet) 분포를 활용하여 보완(Titov & McDonald, 2008)하였기 때문에 텍스트 기반 토픽모델링 분석에서 가장 널리 활용(남승주 · 이현철, 2019)되고 있는 방법론이다.

이와 같은 LDA의 특성으로 항공사 온라인 리뷰관련 연구에서도 LDA를 적용한 사례가 존재하고 있다. 이시환 · 이훈영(2017)은 트립어드바이저에서 대한항공, 아시아나항공과 제주항공의 한글 고객 리뷰만을 추출하여 LDA 분석을 실시하였다. 그 결과 좌석, 가격, 승무원, 공항과 시간, 사람과 한국, 서비스, 편안함과 음식, 노선과 기종, 체크인, 기내식과 영화의 10가지로 구분하였다. 이와 함께 고객들이 평가한 항공사 평점 요소들을 바탕으로 회귀분석을 실시하여 전체적인 항공사 만족에 영향을 미치는 세부 요소들을 확인하였다. Lim & Lee(2020)의 연구에서는 항공사 및 공항을 평가하는 웹사이트로 유명한 Skytrax(airline quality.com)의 온라인 고객 리뷰를 전체적으로 추출하여 고객 리뷰를 대형항공사와 저가항공사로 구분하였다. 구분된 리뷰에서 각각 토픽을 도출하였고 각 도출된 토픽을 바탕으로 대형항공사와 저가항공사의 서비스 품질에 대한 정보를 비교하여 제시하였다.

2. 온라인 리뷰

온라인 리뷰는 상품을 사용했거나 구매한 고객들에 의해 작성되는 것으로 고객의 경험, 평가와 의견을 포함하고 있다(Litvin et al., 2008). 형태는 텍스트 기반의 리뷰와 점수 형태의 평점으로 구분할 수 있고(Liu, 2006) 관광분야에서는 호텔, 레스토랑, 관광지 등에 대한 경험을 작성하여 다른 사람들과 나누고 있는 형태이다(Ye et al., 2009). 이렇게 온라인 리뷰는 사용자들이 만드는 콘텐츠로써 많은 온라인 사용자들이 콘텐츠를 생성하고 축적함으로써 집단지성(wisdom of crowds)이 온라인에 창출되고 있는 것으로 볼 수 있다(Surowiecki, 2005). 관광분야에서도 온라인에 누적되는 고객들의 리뷰에 관광분야 연구자들이나 실무자들이 주의 깊게 관심을 갖고 있다. 실제 다양한 관광분야 연구에서 온라인 리뷰가 관광객들의 기호나 욕구를 정확히 이해하는데 중요한 자료로 활용된다는 사실이 증명되고 있다(Chau & Xu, 2012). 그 중 한 분야가 항공 산업으로 항공사에 대한 고객 리뷰의 중요성이 높아지면서 관련 연구가 국내외적으로 다양하게 진행되고 있는 상황이다.

최근 이루어지는 항공사 온라인 리뷰와 관련된 선행연구들을 살펴보면 크게 두 가지 방향으로 구분할 수 있다. 첫 번째는 온라인 리뷰에 나타난 단어들 간의 의미를 네트워크분석을 통해 연결하는 의미연결망 연구(반현정 · 김학선, 2019; 홍지숙 · 오익근, 2016; Ban & Kim, 2019)가 많이 이루어지고 있다. 이는 고객들이 리뷰에 남긴 단어들 중 높은 빈도수의 단어들을 확인하고 단어들 간의 의미를 연결해 보는 분석으로 고객의 인식, 경험, 만족을 이해하기 위한 시도로 볼 수 있다. 두 번째로는 리뷰에

나타난 단어들을 몇 개의 군집으로 구분한 후 대표적인 의미를 추출해 내는 토픽모델링을 적용한 연구(남승주 · 이현철, 2019; Sezgen et al., 2019)가 이루어지고 있다. 리뷰에서 나타난 많은 단어들을 효율적으로 몇 개의 군집으로 구분한 후 그 군집을 대표할 잠재적인 의미를 추출하여 고객만족에 영향을 미칠 수 있는 요소들을 이해하기 위한 연구방법으로 볼 수 있다. 이와 같은 두 가지 방법론 이외에도 온라인 리뷰라는 텍스트 기반 정보에서 유용한 정보를 추출하기 위한 방법론은 현재도 관광분야 연구에서 지속적으로 제시되고 있어 앞으로도 온라인 리뷰에 대한 연구는 계속될 전망이다.

3. 항공사 서비스품질

일반적인 서비스품질에 대한 정의는 고객들이 평가하는 서비스의 전반적인 우수성이라고 할 수 있다(Parasuraman et al., 1988). 즉 서비스품질은 고객이 서비스에 대해서 어떻게 인지하는지가 가장 중요한 개념이며, 고객의 서비스품질 평가는 인지된 서비스의 품질과 고객 기대와의 차이에서 발생한다(Gronroos, 2001). 따라서 항공사와 같이 서비스품질이 중요한 산업에서는 다양한 고객들의 특성에 따라 항공 서비스 품질을 고객 기대수준에 맞추기 위한 노력이 필수적이라고 할 수 있다. 그러나 다양한 고객 구성으로 항공 서비스에 대한 기대도 많고 만족에 대한 평가도 다르기 때문에 일부 고객들은 불만족을 경험할 수도 있을 것이다. 이럴 경우 고객의 불만족 경험은 곧 부정적인 구전, 불평과 고객 이탈 등으로 나타날 수 있다(Lee et al., 2011).

최근 다양한 온라인 미디어의 발달로 항공사 서비스 품질에 대한 고객들의 만족과 불만족 경험은 여러 형태와 양식으로 표출되고 있다. 그 중 하나인 온라인 리뷰는 고객들이 자신의 경험을 설명하고 공유하기 위한 가치 있는 방식이라고 할 수 있다(Ban & Kim, 2019). 이러한 가치를 인정하여 온라인 리뷰를 이용한 항공 서비스 품질 연구들이 진행되었다. 한명기 · 최병구(2019)는 트립어드바이저에서 전 세계 대표적인 50여 개 항공사의 리뷰를 추출하여 감성분석(sentiment analysis)을 실시하였다. 선행연구를 통해 서비스 품질 요소들을 크게 5개로 구분한 후 각 분류에 속하는 리뷰 내용들을 대상으로 감성분석을 실시하여 고객들이 인지하는 각 서비스 품질 요소의 감성을 긍정과 부정으로 평가하였다. 이를 통해 항공사의 서비스 품질 중 항공사 유형과 개인 서비스라는 요소가 가장 중요한 지표이면서 고객들이 가장 민감하게 반응한다는 결과를 도출하였다. Korfiatis et al.(2019)은 트립어드바이저에서 수집 가능한 고객들의 항공사 리뷰를 전부 수집하여 토픽모델링 기법인 STM과 LDA로 분석하였다. 그 결과 항공사 서비스 품질 요소로 비즈니스 좌석, 수화물 정책, 좌석 공간, 항공사 직원들, 항공지연 등의 20개 토픽을 구체적으로 제시하였다. 일반적으로 고객들은 항공사에 대한 기대도 다르고 서비스 품질에 대한 만족이나 평가도 다르다고 할 수 있다(김진영, 2018).

III. 연구 설계

1. 연구문제

본 연구에서는 온라인 리뷰를 활용하여 고객만족에 영향을 미치는 중요한 서비스 품질 요인들을 밝히고자 하였다. 이를 위해 호텔분야 토픽모델링 연구인 Guo et al.(2017)과 항공사 서비스 품질 관련 토픽모델링 연구인 Korfiatis et al.(2019) 등의 선행연구를 바탕으로 토픽모델링 방법론 중 LDA를 채택하여 본 연구에 적용하였다. 이와 함께 토픽모델링의 결과를 통해 구체적으로 밝힐 수 있는 연구문제(research questions)를 다음과 같이 설정하였다.

1. 온라인 리뷰에서 나타난 국내 대형항공사의 서비스 품질 요소는 무엇인가?
2. 온라인 리뷰에서 나타난 대한항공의 서비스 품질 요소는 무엇인가?
3. 온라인 리뷰에서 나타난 아시아나항공의 서비스 품질 요소는 무엇인가?
4. 온라인 리뷰에서 나타난 대한항공과 아시아나항공의 서비스 품질 요소에 어떤 차이가 있는가?

2. 자료 수집

본 연구에서는 항공사에 대한 고객들의 리뷰 자료를 수집하기 위하여 세계적으로 널리 알려진 온라인 리뷰 웹사이트이면서 관광객들에게 무료로 온라인 여행정보 서비스를 제공하는 트립어드바이저(강성진·문혜영, 2019)를 이용하였다. 기존 항공사관련 선행연구에서 널리 사용된 설문방식의 자료와 비교할 때 온라인 리뷰 자료는 상대적으로 관광분야에서 조금 더 대표성 있는 자료(Park et al., 2020)이면서, 서비스 경쟁력을 위한 중요한 정보를 제공(조동석 등, 2020)한다고 할 수 있다.

이러한 항공사 온라인 리뷰 자료를 수집하기 위하여 파이썬(python) 개발언어를 이용하여 자동 웹크롤링 프로그램을 개발하였다. 웹크롤링 프로그램의 동작은 리뷰관련 트립어드바이저의 웹페이지 구성에 맞춰서 먼저 언어선택 기능을 수행하였다. 본 연구는 영어 리뷰만을 수집 대상으로 하였기에 ‘영어’ 버튼을 선택하여 영어 리뷰만을 추출하였다. 그 다음은 트립어드바이저의 각 리뷰가 기본적으로 일부분만 보이기 때문에 리뷰의 전체적인 내용을 확인하기 위해서 리뷰의 ‘더보기’ 버튼을 눌러 선택할 수 있도록 프로그래밍을 하였다. 트립어드바이저는 한 페이지에 5개 정도의 리뷰만 출력되기 때문에 전체 리뷰 자료를 수집하기 위해서는 리뷰의 페이지를 계속 다음 페이지로 넘기는 동적인 작업을 자동으로 수행할 수 있도록 웹크롤링 프로그램을 구성하였다. 이렇게 동적 웹페이지를 크롤링하기 위하여 구체적으로 파이썬 프로그램은 selenium을 이용하여 트립어드바이저의 내용을 가져온 후 BeautifulSoup을 통해 수집된 웹페이지의 내용을 구문 분석하여 <그림 1>과 같이 텍스트로 구성된 리뷰 자료만을 추출하였다. 대한항공과 아시아나항공에 대한 리뷰 자료는 트립어드바이저의 항공사 리뷰가 시작된 2016년 1월부터 2020년 12월 7일 사이에 작성된 리뷰 자료만을 각각 추출하였고 최종 수집된 리뷰는 총 4,763개로 대한항공이 3,147개, 아시아나항공이 1,616개였다. 트립어드바이저에는 대한항공의 고객 리뷰가 아시아나항공의 고객 리뷰보다 더 많이 누적되어 있어 수집된 리뷰 자료 개수에 큰 차이가 있었다. 그리고 이렇게 수집한 리뷰 자료는 CSV(comma-separated values) 파일 형태로 저장하여 분석을 진행하였다.



<그림 1> 트립어드바이저 웹사이트 리뷰

3. 자료 처리 및 분석

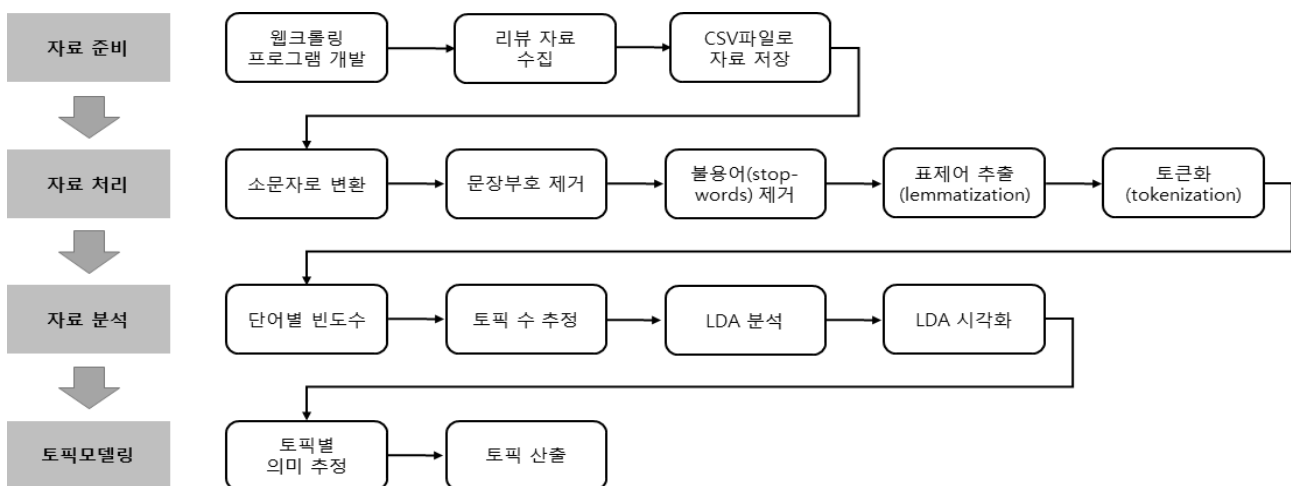
본 연구를 위해 수집한 자료는 영어 리뷰 자료로 국내 대형항공사에 대한 한글 리뷰가 영어 리뷰보다 상대적으로 적은 양이고 외국인들이 평가하는 국내 대형항공사에 대한 서비스 만족 요소를 도출하기 위하여 영어 리뷰만을 연구의 자료로 활용하였다.

일반적으로 리뷰 자료는 긴 문자열로 이루어져 비구조적인 형태로 되어 있어 컴퓨터가 처리할 수 있도록 작은 의미 단위인 토큰(token)으로 나누어야 한다(Kwon et al., 2020). 그러나 그 전에 제대로 된 결과를 산출할 수 있는 토큰을 생성하기 위해서는 파이썬 프로그램을 이용해서 몇 가지 사전 자료 처리과정을 시행하였다. 가장 먼저 리뷰 자료에 있는 대문자를 모두 소문자로 바꾸어 대·소문자에 따른 단어의 차이를 모두 통일시켰다. 그 후 문장부호, 숫자와 문자 이외의 심볼 등을 제거하는 과정을 거쳤고 ‘a’, ‘the’, ‘that’ 등의 불용어라고 불리는 영어 단어들도 제거하였다. 이런 불용어들은 일반적인 문장을 구성할 때는

필요한 요소이나 토픽모델링 분석에서는 의미를 제공하지 못하기 때문에 포함시키지 않았다. 영어 단어들 중 ‘goes’와 같은 3인칭 단수, ‘went’와 같은 과거와 ‘gone’과 같은 과거 분사형 등을 단어의 원형인 ‘go’의 형태로 바꾸는 표제어 추출작업을 통해 분석의 효율을 높이려고 하였고 이런 과정을 모두 거친 후 마지막으로 컴퓨터가 처리할 수 있도록 단어들을 토큰화 하였다.

이러한 자료 처리과정을 통해서 대한항공은 3,147개의 리뷰로부터 총 1,866개의 단어를, 아시아나항공은 1,616개의 리뷰로부터 총 1,297개의 단어를 추출하여 토픽모델링 분석에 사용하였다. 먼저 추출된 단어들의 일반적 빈도를 확인하기 위하여 빈도분석을 실시하였다. 항공사 서비스관련 고객만족 요인의 토픽을 토큰화된 단어들로부터 산출하기 위하여 LDA를 실시하였다. 그러나 LDA는 자체적으로 가장 적합한 토픽의 수를 결정하지 못하고 분석자가 직접 토픽 수를 지정해 줘야하기 때문에 적절한 토픽의 수를 결정하기 위하여 토픽 일관성(topic coherence) 분석을 실시하였다. 그 결과 토픽의 수가 7개에서 높은 일관성을 보인다는 결과가 산출되어 대한항공과 아시아나항공 모두 LDA 분석 시 토픽 수를 7개로 지정하여 분석하였다. 또한 LDA는 머신러닝의 비지도 학습 방식이므로 매번 LDA를 실시할 때마다 다른 결과가 도출된다. 이에 같은 결과를 산출하기 위하여 파이썬 프로그램에서 난수의 초기 값을 지정(random_state=1)하여 고정하였다. 분석의 마지막으로 파이썬 라이브러리인 LDAvis를 이용하여 토픽들을 시각화하여 토픽 주제어 결정 시 시각화된 자료를 활용하였다.

이렇게 도출된 토픽들은 각 토픽에 분포된 단어를 바탕으로 토픽의 주제어를 결정하였다. 이러한 작업은 본 연구의 저자들이 먼저 단어들의 의미를 통해 주제어를 정한 후 추가로 항공사에 10년 이상 근무한 경험을 갖고 있는 객실승무원 출신 실무 전문가 1명과 항공서비스학과 교수 1명의 검증을 거쳤다. 이러한 주제어 선정과정의 반복을 통해 최종적으로 토픽 주제어를 결정하였다. 전체적인 리뷰 자료의 수집, 처리와 분석 등에 관한 체계는 <그림 2>와 같다.



<그림 2> 연구방법의 체계

IV. 실증 분석

1. 자료의 특성

1) 리뷰 작성자의 특성

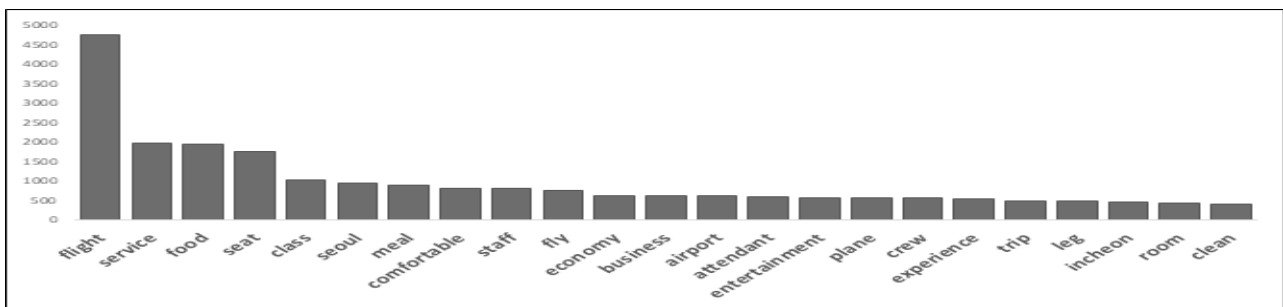
트립어드바이저는 리뷰 작성자의 위치를 알 수 있는 정보를 표시해 준다. 모든 리뷰 작성자들의 위치가 표시되지는 않지만, 위치가 표시된 리뷰 작성자들의 자료는 웹크롤링으로 수집이 가능하다. 위치가 리뷰 작성자의 국적을 의미하지는 않지만 거주지의 근거가 될 수 있기 때문에 본 연구에서는 <표 1>과 같이 대한항공 2,659명, 아시아나항공 1,362명의 리뷰 작성자들의 위치정보를 수집하였다. 대한항공 리뷰 작성자의 위치 중 미국이 37.2%로 가장 높았으며, 그 다음이 호주 10.4%, 영국 6.3%, 캐나다 6.3% 순이었다. 아시아지역에서는 싱가포르가 4.1%로 가장 높게 나타났다. 아시아나항공은 미국이 44.6%, 영국 7.3%, 필리핀 6.7%, 호주 5.4%의 순이었다. 특히 아시아나항공에서는 필리핀과 일본의 위치가 대한항공보다 높게 나타난 특징을 보이고 있다.

〈표 1〉 리뷰 작성자의 위치

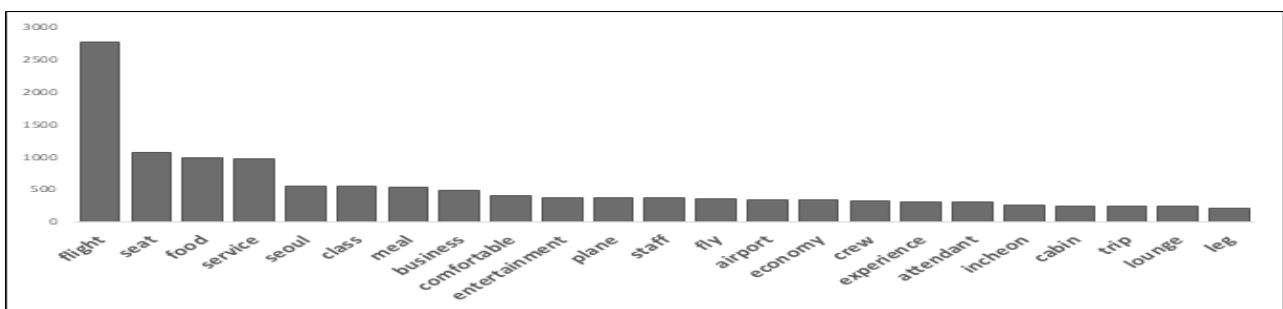
항목	빈도(명)	비율(%)	항목	빈도(명)	비율(%)
대한항공(N=2,659)			아시아나항공(N=1,362)		
미국	990	37.2	미국	607	44.6
호주	276	10.4	영국	100	7.3
영국	168	6.3	필리핀	91	6.7
캐나다	168	6.3	호주	73	5.4
싱가포르	108	4.1	싱가포르	72	5.3
한국	101	3.8	캐나다	55	4.0
태국	97	3.6	한국	42	3.1
필리핀	91	3.4	일본	35	2.6
뉴질랜드	73	2.7	태국	35	2.6
기타	587	22.2	기타	252	18.4

2) 리뷰 단어 빈도분석

대한항공과 아시아나항공의 리뷰에서 많이 나타난 단어들의 빈도를 <그림 3>, <그림 4>와 같이 분석하였다. 대한항공을 이용한 고객들의 리뷰에서는 ‘flight’이라는 단어가 가장 많은 4,747회 나타났고 그 다음이 ‘service’ 1,977회, ‘food’ 1,945회, ‘seat’ 1,764회, ‘class’ 1,032회, ‘seoul’ 939회 등의 순으로 나타났다. 아시아나항공의 리뷰에서는 ‘flight’ 2,761회, ‘seat’ 1,073회, ‘food’ 995회, ‘service’ 975회, ‘seoul’ 549회, ‘class’ 543회 등의 순으로 대한항공과 거의 유사하게 나타났다. 즉 두 항공사의 리뷰에서 산출한 높은 빈도의 단어들만 서로 비교할 경우 순서에서 일부 차이는 있었으나 거의 비슷한 단어들이 나타나고 있었다.



〈그림 3〉 대한항공 리뷰의 단어 빈도



〈그림 4〉 아시아나항공 리뷰의 단어 빈도

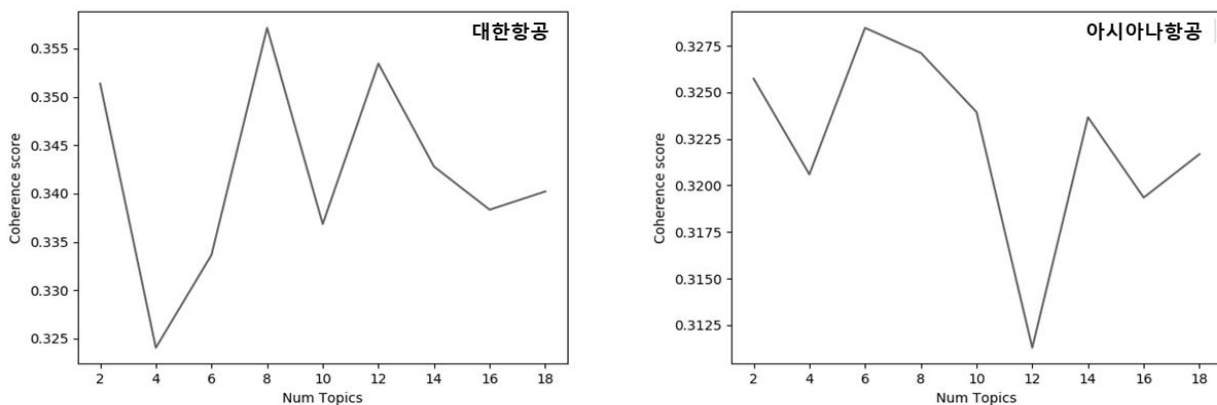
2. 토픽 분석

1) 토픽 일관성 검증

토픽모델링을 평가하는 척도로는 혼란도(perplexity)와 토픽 일관성이 있다(Newman et al, 2010). 이러한 모델링 평가를 통해 적절한 토픽의 개수를 결정하기도 한다. 토픽의 개수는 연구자가 임의로 결정할 수도 있지만 일반적으로 평가 척도인 혼란도나

토픽 일관성 검증을 통해 토픽의 수를 결정하기도 한다. 선행연구들에서는 두 가지 방법 중 어느 방법이 더 우수하거나 검증효과가 좋다는 비교 연구나 주장을 찾기는 어려웠다. 이에 본 연구에서는 혼란도보다 더 후에 제시된 방법론인 토픽 일관성을 사용하여 토픽의 개수를 결정하고자 하였다.

토픽 일관성은 각 토픽들 안에서 높은 빈도로 도출된 단어들의 의미적인 일관성을 측정하는 척도이며 이를 통해 토픽의 개수에 따라 토픽들의 일관성 변화가 나타난다(Stevens et al., 2012). 즉 토픽 일관성이 가장 높은 지점이 토픽의 의미적 구분이 가장 잘 구현된 지점으로 볼 수 있고 그 지점을 최적의 토픽 개수로 판단하기도 한다. 이러한 기준을 바탕으로 <그림 5>에서와 같이 대한항공과 아시아나항공에 대한 리뷰 자료의 토픽 일관성 검증을 실시한 결과 대한항공의 일관성 검증 그래프는 토픽의 수가 8개일 때 가장 높은 일관성 수치를 보이고 있다. 이는 대한항공 리뷰 자료를 바탕으로 토픽을 구분할 때 8개의 토픽을 구성하는 것이 일관성에서 가장 높은 결과를 보인다는 의미로 해석할 수 있다. 그러나 아시아나항공의 토픽 일관성 검증 결과는 대한항공과 조금 다르게 토픽이 6개에서 가장 높은 일관성 수치를 나타내었다. 결과적으로 대한항공은 8개의 토픽, 아시아나항공은 6개의 토픽이 적합한 것으로 보였다. 그러나 본 연구의 목적에 따라 두 항공사의 고객만족 요인에 대한 공통점과 차이점을 비교해 보기 위해서는 두 항공사의 토픽 개수를 일치시켜 비교의 정확성을 높이는 것이 필요하다고 판단되었다. 토픽 일관성 검증을 통한 결과가 항상 더 좋은 정보를 나타내지 않을 수도 있고(Stevens et al., 2012) 두 항공사 모두 일관성 검증에서 7개의 토픽일 때 공통적으로 높은 일관성 수치를 보이고 있다고 할 수 있기에 본 연구에서는 대한항공과 아시아나항공의 토픽 개수를 동일하게 7개로 결정하였다.



<그림 5> 토픽 일관성 검증

2) LDA 분석

고객들의 리뷰 중 단순히 일정이나 여행을 소개한 경우도 있지만, 대부분의 리뷰에서는 항공사를 이용하면서 느꼈던 만족 사항이나 불만족 사항들에 대한 고객들의 언급이 주를 이룬다. 즉 고객 리뷰에서 많이 언급되는 단어의 의미는 결국 고객만족과 연결된 항공사의 중요한 유·무형의 서비스품질에 대한 고객 평가로 볼 수 있으며, 항공사의 경쟁력 강화를 위한 핵심적인 요소로도 볼 수 있다.

이러한 서비스품질과 관련된 토픽들을 리뷰로부터 도출하기 위하여 LDA 분석을 실시하였고, 토픽 일관성 검증을 참조하여 대한항공과 아시아나항공 모두 7개씩의 토픽을 산출하였다. 그리고 각 토픽에는 약 30개씩의 단어들이 구성되었으나 토픽 주제어 선정을 위해 각 토픽별 단어의 비중이 높으면서 주제어 선정을 위한 중심적 의미를 가진 핵심적인 단어들 5개씩만 <표 2>와 <표 3>에 명시하였다. 각 토픽을 구성하고 있는 이 단어들을 바탕으로 본 연구의 저자들이 먼저 핵심 단어들 사이의 논리적인 연결을 통해 주제어 후보를 선정하였다. 예를 들어 대한항공의 토픽을 나타내는 <표 2>에서 첫 번째 토픽 주제어인 '기내 엔터테인먼트'는 'inflight entertainment'와 'movie'라는 단어들을 바탕으로 주제어를 유추하였다. 함께 나타난 단어들 중 'economy class'와 'flight attendant' 단어들도 '기내 엔터테인먼트'와의 연관성을 확인하기 위하여 관련 내용을 실제 고객들의 리뷰 자료에서 검토하였다. 확인 결과 'inflight entertainment'를 언급할 때 많은 경우 'economy class' 좌석 모니터에 대한 언급과 'flight attendant'의 모니터 기기 사용 시 도움 제공 등과 같은 내용이 많아 토픽 주제어로 '기내 엔터테인먼트'가 적합한 것으로

판단하였다. 이와 같은 과정으로 각 토픽별 비중이 높은 단어들을 바탕으로 대표적인 주제어를 7개씩 선정하였으며 항공사별로 선정된 7개의 주제어들이 토픽의 단어들을 잘 반영하도록 조정하였다. 그런 후 최종 후보 주제어들을 항공사에서 10년 이상 근무한 객실승무원 출신 실무 전문가 1명과 항공서비스학과 교수 1명으로부터 의미의 타당성 검증을 받았다. 일부 토픽의 주제어가 해당 단어들을 대표하는 주제어로 적합하지 않거나 의미가 불분명할 경우 다시 다른 주제어를 선정한 후 확인 작업을 반복하였다.

이러한 과정을 거쳐 최종적으로 결정된 대한항공의 주제어들은 ‘기내 엔터테인먼트’, ‘고객 서비스’, ‘일등석’, ‘비즈니스 좌석’, ‘좌석 공간’, ‘환승 서비스’, ‘승무원’이었다. 아시아나항공의 주제어는 ‘비즈니스 좌석’, ‘고객 서비스’, ‘좌석 공간’, ‘일반석’, ‘스타얼라이언스’, ‘기내 엔터테인먼트’, ‘승무원’이었다. 항공사의 주요 상품이 좌석의 품질이기 때문에 좌석에 대한 주제어가 많이 보이고 있고 기내 서비스의 핵심인 승무원에 대한 주제어도 두 항공사에서 모두 나타나고 있다. 이러한 결과를 요약해보면 두 항공사의 공통적인 주제어는 ‘비즈니스 좌석’, ‘고객 서비스’, ‘좌석 공간’, ‘기내 엔터테인먼트’, ‘승무원’ 이고 두 항공사에서 서로 다르게 나타난 주제어는 ‘일등석’, ‘환승 서비스’, ‘일반석’, ‘스타얼라이언스’이었다.

〈표 2〉 대한항공의 각 토픽별 주제어 선정

토픽	Topic 1: 기내 엔터테인먼트	Topic 2: 고객 서비스	Topic 3: 일등석	Topic 4: 비즈니스 좌석
구성 단어	economy class (.027)	flight attendant (.031)	first class (.026)	business class (.085)
	inflight entertainment (.025)	customer service (.022)	cabin crew (.020)	long haul (.022)
	flight attendant (.024)	check (.011)	staff friendly (.013)	lounge (.013)
	movie (.010)	ticket (.008)	serve (.009)	prestige class (.010)
	best (.009)	airport (.007)	drink (.006)	flat bed (.008)
토픽	Topic 5: 좌석 공간	Topic 6: 환승 서비스	Topic 7: 승무원	
구성 단어	leg room (.051)	hotel (.018)	flight attendant (.031)	
	plane (.011)	incheon (.013)	best (.011)	
	economy (.010)	grand hyatt (.011)	helpful (.011)	
	plane clean (.009)	airport (.010)	friendly (.010)	
	seat comfortable (.008)	breakfast (.010)	crew (.008)	

주: 괄호안의 숫자는 단어의 비중을 의미함.

〈표 3〉 아시아나항공의 각 토픽별 주제어 선정

토픽	Topic 1: 비즈니스 좌석	Topic 2: 고객 서비스	Topic 3: 좌석 공간	Topic 4: 일반석
구성 단어	business class (.069)	passenger (.010)	leg room (.014)	economy class (.012)
	flight attendant (.042)	airport (.010)	window seat (.009)	seat comfortable (.010)
	excellent (.007)	customer service (.009)	well (.008)	experience (.008)
	upper deck (.006)	staff (.009)	economy (.007)	flight attendant (.007)
	a380 (.006)	ticket (.008)	screen (.006)	recommend (.007)
토픽	Topic 5: 스타얼라이언스	Topic 6: 기내 엔터테인먼트	Topic 7: 승무원	
구성 단어	star alliance (.026)	inflight entertainment (.031)	cabin crew (.035)	
	connect flight (.011)	economy class (.012)	Korean air (.024)	
	recommend (.009)	movie (.011)	friendly (.008)	
	customer (.009)	experience (.010)	staff (.006)	
	incheon airport (.008)	selection (.009)	well (.006)	

주: 괄호안의 숫자는 단어의 비중을 의미함.

3. 토픽 주제어 비교

Guo et al. (2017)은 호텔들의 리뷰를 바탕으로 도출한 서비스품질 토픽들을 선행연구에서 제시한 토픽들과 비교하기 위하여 자카드 계수(Jaccard coefficient)라는 유사성 지표를 제시하였다. 유사한 연구들의 결과를 비교하기 위하여 적용할 수 있는 수치로 계수의 값이 0에 가까울수록 유사성이 낮다는 의미이며 1에 가까울수록 유사성이 매우 높아진다는 의미로 해석할 수 있다. 이러한 자카드 계수의 계산공식은 두 집합의 교집합을 두 집합의 합집합으로 나누면 되고 그 값은 음의 값이 없는 0~1 사이의 절대값이다.

〈표 4〉 LDA 방법론을 적용한 항공서비스 토픽 주제어 비교

구분		본 연구		선행연구	
				이시환 · 이훈영(2017)	Lim & Lee(2020)
연구 방법	연구대상 항공사	대한항공	아시아나항공	대한항공, 아시아나항공, 제주항공	세계10대 대형항공사
	리뷰 언어 및 출처	영어, TripAdvisor	영어, TripAdvisor	한글, TripAdvisor	영어, Skytrax
	표본의 수	3,147개	1,616개	900개	7,462개
토픽 주제어	고객 서비스	○	○	○	○
	기내 엔터테인먼트	○	○	○	○
	비즈니스 좌석	○	○	×	×
	스타얼라이언스	×	○	×	×
	승무원	○	○	○	○
	일등석	○	×	×	×
	일반석	×	○	×	×
	좌석 공간	○	○	○	○
	환승 서비스	○	×	×	○

주: ○: 포함, ×: 미포함; 자카드 계수: 대한항공과 아시아나항공 .56, 대한항공과 이시환 · 이훈영(2017) .31, 대한항공과 Lim & Lee(2020) .16.

〈표 4〉는 대한항공과 아시아나항공의 주제어 비교 및 추가적으로 대형항공사의 리뷰를 바탕으로 LDA를 적용한 선행연구들과의 주제어 비교 결과이다. 먼저 이 결과를 바탕으로 자카드 계수를 산출해 보면 대한항공과 아시아나항공의 자카드 계수는 .56이다. 즉 두 항공사에서 도출된 토픽들의 유사성은 약 56%라고 해석할 수 있다. 이와 함께 선행연구들과의 비교에서는 대표적 국내 대형항공사인 대한항공을 기준으로 선행연구 이시환 · 이훈영(2017)과의 자카드 계수는 .31로 나타났고 선행연구 Lim & Lee(2020)와의 자카드 계수는 .16으로 나타났다. 이는 대한항공이나 아시아나항공에서 산출된 토픽의 수가 7개인 반면 선행연구 이시환 · 이훈영(2017)은 10개, Lim & Lee(2020)는 30개의 토픽을 도출하여 서로 다른 토픽의 수를 바탕으로 공통적인 요소들을 비교하였기에 자카드 계수에 영향을 미친 것으로 볼 수 있다. 그러나 Guo et al. (2017)의 연구에서는 자카드 계수가 낮을 경우에도 두 연구가 서로 다른 새로운 결과들을 도출하였다는 의미를 제시하였다.

이러한 유사성 지표를 참고로 조금 더 구체적인 토픽 주제어들의 비교 결과를 살펴보면 대한항공과 아시아나항공에서 공통적으로 나타난 주제어는 ‘고객 서비스’, ‘기내 엔터테인먼트’, ‘비즈니스 좌석’, ‘승무원’, ‘좌석 공간’이었다. 이는 본 연구에서 제시한 첫 번째 연구문제인 온라인 리뷰에서 나타난 국내 대형항공사의 서비스품질 요소는 무엇인가라는 질문에 대한 결과 제시로 볼 수 있다. 이러한 공통적인 요소들은 국내 대형항공사들이 고객만족을 위해 집중해야 할 중요한 서비스품질 요소들이라고 할 수 있다. 특히 이 중에서 ‘비즈니스 좌석’을 제외한 다른 4가지 요소들은 선행연구들의 결과와도 일치한다. 제주항공을 포함한 이시환 · 이훈영(2017)의 연구나 전 세계 많은 항공사들을 대상으로 한 Lim & Lee(2020)의 연구에서도 같은 결과를 보인다는 의미는 ‘고객 서비스’, ‘기내 엔터테인먼트’, ‘승무원’, ‘좌석 공간’의 요소는 어떤 항공사이든 고객들이 가장 기본적으로 고려하는 서비스품질이라고 해석할 수 있다. 결국 항공사 경쟁력 향상을 위해서는 이러한 기본적인 요소들에 대한 서비스품질 유지 및

강화가 선행되어야 할 것이다.

앞에서 살펴본 바와 같이 두 항공사에서 공통적으로 나타난 주제어와는 별도로, 아시아나항공에서 나타나지 않고 대한항공에만 존재하는 주제어는 ‘일등석’, ‘환승 서비스’이다. 반대로 아시아나항공에서만 보이는 주제어는 ‘스타얼라이언스’와 ‘일반석’으로 나타났다. 이는 본 연구의 두 번째, 세 번째 연구문제인 온라인 리뷰에서 나타난 대한항공의 서비스품질 요소와 아시아나항공의 서비스품질 요소를 밝혀낸 것으로 공통적인 요소 5가지와 각 항공사별 고유한 주제어 2개씩으로 구성되어 대한항공의 서비스품질 요소 7개, 아시아나항공의 서비스품질 요소 7개를 연구문제의 결과로 제시할 수 있다. 마지막 연구문제인 두 항공사의 서비스품질 요소에는 어떠한 차이가 있는가에 대해서도 두 항공사의 고유한 주제어가 2개씩 도출되어 이 요소들을 네 번째 연구문제에 대한 결과로 제시할 수 있다. 그러므로 본 연구에서 제시한 4개의 연구문제는 토픽모델링 LDA 분석을 통해 결과를 모두 확인할 수 있었다.

V. 결 론

본 연구는 국내 최대 항공사인 대한항공이 아시아나항공을 인수하면서 합병이 단순히 항공사의 규모만 확장시키는 것이 아니라 세계적으로 치열한 항공 산업 경쟁에서 국내 항공사의 서비스 경쟁우위 확보를 위한 시사점을 제시하고자 하였다. 이를 위해 고객들이 직접 생성한 콘텐츠인 리뷰 자료를 활용하여 대한항공과 아시아나항공을 이용한 경험을 갖고 있는 해외 고객들이 중요하게 언급한 서비스품질 요소들을 도출하여 비교 분석하였다. 이러한 분석 결과들을 바탕으로 다음과 같은 학술적 및 실무적 시사점들을 제시할 수 있다.

먼저 학술적 시사점으로 첫 번째는 고객들의 리뷰 자료를 토픽모델링의 방법론으로 분석하여 항공사 서비스 연구에 적용하였다는 사실이다. 최근 고객들의 리뷰 자료를 이용한 연구들을 어렵지 않게 접할 수 있고 관광분야에서도 관광객들이 작성한 리뷰 분석을 위해 다양한 방법론이 소개되고 있다. 특히 그 중에서 토픽모델링은 언어기반의 자료를 처리하기 위한 방법론의 하나로 관광분야 연구에서는 이미 국내의 연구에서 소개된 상황이다. 그러나 상대적으로 지금까지의 연구 주류인 설문방식에 의한 요인들 간의 영향관계 분석 연구들과 비교하면 리뷰 자료를 활용한 연구는 아직 많이 부족한 상황이며 더욱이 항공사관련 토픽모델링 연구는 상대적으로 많이 부족한 것이 사실이다. 이에 본 연구는 리뷰 자료를 바탕으로 토픽모델링을 적용한 국내 대형항공사의 서비스품질관련 연구 부재를 보완하고자 하였다.

둘째, 국내 항공 산업의 시대적 변화에 발맞춘 연구로 의미를 부여할 수 있다. 코로나19 이전에 이미 아시아나항공의 매각은 결정된 사실이었다. 그러나 예상치 못한 코로나19로 항공 산업의 위기가 높아지면서 아시아나항공이 우여곡절 끝에 대한항공에 인수되는 것으로 최근 결정되었다. 이러한 통합은 항공 산업의 위기를 기회로 바꾸기 위한 업계와 정부의 노력으로 볼 수 있다. 이에 학계에서도 이러한 노력에 빠르게 동참하면서 국내 항공 산업의 발전을 위한 다양한 연구가 제시되어야 할 것이다. 특히 코로나19로 인해 항공 산업이 거의 멈춰 있는 현 시점이 항공사들의 변화와 혁신을 위한 최적의 시간이라고 할 수 있다. 그러므로 본 연구도 현 시점에서 하나의 항공사로 통합되는 대한항공과 아시아나항공의 서비스 경쟁력들이 통합에 의해 시너지가 창출될 있도록 기여하고자 하였다.

셋째, 고객들이 생성한 콘텐츠를 바탕으로 국내 두 대형항공사의 실질적인 서비스품질 요소를 제시하였다는 의미를 갖는다. 고객들이 리뷰에서 많이 언급하는 사항은 항공사 이용 시 고객들이 서비스에 만족했거나 또는 서비스에 아쉬웠던 내용들이 대부분이다. 즉 항공사별로 고객들의 언급이 많은 요소가 고객만족과 직접적으로 연결된 실질적인 서비스품질 요소로 볼 수 있으며 이러한 요소들은 항공사 간의 공통적인 요소들과 차별적인 요소들로 구성될 수 있다. 공통적인 요소는 치열한 시장경쟁에서 기본적으로 갖추어야 할 항공사의 서비스 경쟁력일 수 있으며 차별화된 요소는 경쟁우위를 위한 해당 항공사만의 강점이 될 것이다. 그러므로 본 연구는 하나의 항공사로 통합되는 국내 두 대형항공사 간의 공통적인 서비스품질 요소와 차별화된 서비스품질 요소에 대한 고객들의 소리를 제공하고자 하였다.

이와 같은 살펴본 학술적 시사점들과 함께 통합을 앞둔 두 항공사의 서비스품질 요소에 대한 실무적인 시사점들은 다음과 같다. 첫째, 국내 두 항공사에서 공통적으로 도출된 주제어는 ‘고객 서비스’, ‘기내 엔터테인먼트’, ‘비즈니스 좌석’, ‘승무원’, ‘좌석 공간’이었다. 공통적으로 도출되었다는 의미는 항공사 운영에 있어서 가장 중요하고 기본이 되는 서비스품질 요소라고 할 수 있다. 특히 이러한 요소들은 단지 국내 두 대형항공사만의 결과가 아니라 세계 10대 대형항공사의 리뷰를 대상으로 연구한

Lim & Lee(2020)의 결과에서도 거의 유사한 결과를 보였다. 결국 이와 같은 서비스 요소들은 세계적인 대형항공사들이 고객만족을 위해 기본적으로 갖추어야 할 서비스품질이라고 할 수 있다. 이에 조금 더 정확한 시사점 도출을 위해 5개의 공통적인 주제어들에 대한 실제 대한항공과 아시아나항공의 고객 리뷰들을 자세히 들여다보면 두 항공사의 5가지 서비스 요소에 고객들은 대부분 만족했다는 의견이 많이 나왔다. 그러나 부분적으로 불만족했다는 의견도 다수 보이고 있었다. 이러한 만족과 불만족에 대한 의견과 함께 두 항공사를 모두 이용해봤던 고객들이 서비스를 비교하는 의견을 남기기도 하였다. 구체적으로 비교 의견에서는 기내 엔터테인먼트 요소에서 대한항공이 아시아나항공보다 영어로 관람 가능한 기내 영화가 많다는 일부 의견과 비즈니스 좌석에서는 아시아나항공의 비즈니스 좌석이 최신의 A380 기종 등을 제외하면 전체적으로 조금 낡았다는 의견이 일부 나타나고 있었다. 또한 승무원 요소에서도 대한항공 승무원들이 아시아나항공 승무원들보다 더욱 친절하다는 의견이 부분적으로 나타나고 있었다. 이는 아시아나항공의 경영위기 이후 상대적으로 항공기 좌석에 대한 투자 부족과 승무원들에 대한 정신적, 물질적 사기진작 시행 등이 부족했을 것이라고 유추해 볼 수 있다. 그러므로 대한항공은 이러한 연구결과를 바탕으로 중요한 서비스품질 요소들에 대한 지속적인 품질 유지와 함께 통합되는 아시아나항공의 노후 좌석 등에 대한 투자 및 승무원들에 대한 사기진작 정책을 우선적으로 고려해야 할 것이다.

둘째, 대한항공에서만 나타난 서비스관련 토픽 주제어는 ‘일등석’과 ‘환승 서비스’이었다. 먼저 일등석 요소에서는 아시아나항공이 판매가 저조한 일등석을 2017년부터 점차적으로 줄였기(매일경제신문, 2019. 05. 07) 때문에 대한항공에서 일등석이라는 주제어가 유일하게 도출된 것으로 볼 수 있다. 그러나 일등석이라는 주제어를 바탕으로 고객들의 실제적인 리뷰를 검토한 결과 단지 아시아나항공의 일등석 미판매가 주요한 원인이 아니라 대한항공의 일등석 자체가 좌석의 편리함, 훌륭한 식음료 서비스, 그리고 일등석 고객에 대한 승무원들의 높은 서비스 수준 등으로 대한항공의 일등석이 좋은 평가를 받고 있었다. 이는 대한항공이 프리미엄 항공사라는 이미지를 고객들에게 심어주는 원동력이며 대한항공만의 차별화된 경쟁력이라고 할 수 있다. 또 하나의 차별화된 주제어는 환승 서비스이다. 환승 서비스 주제어에 포함된 단어들을 살펴보면, 대한항공 직영호텔인 ‘그랜드 하얏트 인천’과 함께 ‘호텔’, ‘인천국제공항’ 등이 언급되고 있다. 이를 토대로 리뷰를 확인하면 대한항공 이용 시 인천공항에서 가까운 그랜드 하얏트 호텔을 이용하기 편리하며 호텔의 서비스도 인상적이라는 의견들이 많이 보이고 있었다. 이러한 고객들의 평가는 대한항공이 일등석과 함께 특급 호텔을 연계한 호텔과 공항 간 위탁수하물 연계 서비스, 호텔과 공항 간 무료 셔틀버스 운행, 국제선 탑승권 발급 등과 같은 편리한 환승 서비스를 통하여 프리미엄 항공사로서의 이미지를 굳히는데 원동력이 되고 있다. 그러므로 대한항공은 아시아나항공과의 통합으로 규모면에서 세계적인 항공사로 거듭날 뿐만 아니라 항공사의 이미지에서도 프리미엄 항공사의 대명사가 될 수 있도록 일등석, 환승 서비스에 대한 경쟁력을 유지, 발전시켜야 할 것이다.

셋째, 아시아나항공에서만 나타난 서비스관련 토픽 주제어는 ‘스타얼라이언스’와 ‘일반석’이었다. 잘 알려진 대로 아시아나항공은 1997년 첫 번째로 만들어져 역사가 가장 오래되었고 규모면에서도 제일 큰 스타얼라이언스에 2003년에 가입하였다. 동맹 항공사로는 미국 유나이티드항공, 싱가포르항공, 독일 루프트한자항공, 일본 ANA항공 등이 속해있다. 이에 아시아나항공을 이용하면서 마일리지 적립과 마일리지 항공권 이용 그리고 공동운항 항공기 이용을 통하여 부족한 노선을 보완하여 많은 고객들이 만족을 보이고 있었다. 그러나 다른 항공 동맹체에 가입되어 있는 대한항공과의 통합과정에서 스타얼라이언스의 탈퇴가 자연스러운 수순일 것이다. 그렇기에 아시아나항공의 스타얼라이언스 주제어는 통합되는 두 항공사를 위한 서비스 경쟁력의 시너지 창출 요소로 고려하기는 힘든 상황이다. 그러나 두 항공사의 합병으로 인하여 고객들은 두 항공사의 노선을 모두 이용할 수 있기 때문에 이점은 보완될 것이다. 아시아나항공에서만 나타난 또 하나의 주제어인 일반석은 중요한 정보를 제공하고 있다. 일반석이라는 주제어를 바탕으로 아시아나항공의 고객 리뷰들을 구체적으로 살펴보면 많은 고객들이 아시아나항공의 일반석 식사나 스크린 모니터와 좌석 공간 등에 대해 좋은 평가를 내리고 있었다. 그래서 실제 대한항공과 아시아나항공의 홈페이지를 통해 좌석 사이의 공간이나 좌석의 모니터 사이즈 등을 비교한 결과 A380, B747, B777, A330 등의 기종에서 좌석 공간이나 모니터 사이즈에서 두 항공사 간 차이가 거의 없었다. 하지만 두 항공사의 다른 점은 아시아나항공의 일반석 좌석은 이코노미 레그룸, 이코노미 스마티움, 이코노미로 세분화되어 있어서 이코노미 한 가지 좌석으로 운행이 되는 대한항공에 비해 고객들에게 다양한 선택의 폭을 제공하였고 이와 함께 일반석에 대한 아시아나항공이 쌓아온 긍정적인 이미지의 결과라고도 해석할 수 있다. 결국 아시아나항공은 일반석의 서비스 경쟁력으로 대중적인 이미지를 가진 항공사로 자리매김했다고 할 수 있다. 이런 결과는 두 항공사의 통합과정에서 언론을 통해 보도된 아시아나 브랜드를 저비용항공사의 브랜드로 활용하는 방안(한국경제신문, 2020. 11. 17)과 일맥상통한다고 할 수 있다. 대한항공의 프리미엄 항공사 이미지와 아시아나항공의 대중적인 이미지를 고려하여 아시아나라는 브랜드를

소멸시키지 않고, 두 항공사가 기존에 축적한 이미지를 활용하여 대형항공사와 저비용항공사의 브랜드로 양분화 시킨다면 두 항공사의 통합에서 창출할 수 있는 또 하나의 시너지 효과가 될 수 있을 것이다.

하지만 이러한 시사점에도 불구하고 본 연구는 다음과 같은 한계를 가지고 있다. 우선 ‘영어’ 리뷰만을 대상으로 연구를 진행하였다. ‘영어’ 리뷰가 다른 언어 리뷰들 보다 수량에서 많아 표본 확보에 유리하였다. 그러나 ‘한글’ 등을 사용하는 고객들의 의견까지는 반영하지 못했기에 결과의 일반화에는 한계가 있다고 할 수 있다. 본 연구에서 도출한 주제어들은 확률기반의 LDA 방법론을 통해 도출한 단어들을 기반으로 연구자들이 주제어들을 산정하였기에 실제 두 항공사의 서비스 요소들에 대한 고객들의 긍정, 부정 평가까지는 명확히 확인하기 힘든 한계가 있다. 이에 이러한 한계점을 극복하기 위하여 향후 연구에서는 ‘한글’ 등 분석 가능한 언어를 대상으로 한 연구들이 수행된다면 연구결과에 대한 일반화가 가능할 것이다. 또한 고객들의 리뷰에서 서비스 요소들에 대한 긍정, 부정의 감성까지 도출할 수 있는 감성분석을 시행한다면 항공사 고객만족 서비스 요소에 대한 이해의 폭을 더욱 넓힐 수 있을 것이다.

참 고 문 헌

- 강성진·문혜영(2019). 북한 평양방문 인바운드 관광객의 호텔서비스 평가: 트립어드바이저 리뷰를 중심으로. *호텔관광연구*, 21(2), 75-89.
- 김지연·조우용·최정혜·정예림(2016). 온라인상의 기업 및 소비자 텍스트 분석과 이를 활용한 온라인 매출 증진 전략. *한국경영과학회지*, 41(2), 81-100.
- 김진영(2018). 항공사 상용고객우대제도의 서비스품질이 고객만족과 애호도에 미치는 영향. *호텔관광연구*, 20(1), 190-206.
- 남승주·이현철(2019). LDA 토픽 모델링을 활용한 항공승객 유형별 특성 분석. *경영과학*, 36(3), 67-85.
- 매일경제신문(2019. 05. 07). 아시아나항공 일등석 없앤다. 9월부터 비즈니스 스위트로. Retrieved December 15, 2020, from <https://www.mk.co.kr/news/business/view/2019/05/298766/>
- 매일경제신문(2020. 11. 16). 대한항공-아시아나 합병하면 자산 40조 전 세계 7위 항공사 탄생해. Retrieved December 10, 2020, from <https://www.mk.co.kr/news/business/view/2020/11/1177432/>
- 박윤미(2017). 항공사 객실승무원의 인적서비스가 고객배대성과 고객만족에 미치는 영향. *관광연구저널*, 31(9), 123-135.
- 반현정·김학선(2019). 온라인 리뷰의 의미연결망 분석을 통한 국내 대형항공사 인식에 관한 연구. *Culinary Science & Hospitality Research*, 25(7), 205-215.
- 이시환·이훈영(2017). 한국어 자연어 처리와 LDA, 감성분석을 이용한 항공사 서비스 만족요인 연구. *한국경영정보학회 추계 학술대회는 문집*, 12월, 323-330.
- 임희진·임근욱(2018). 항공사 서비스품질이 고객만족 및 재구매의사에 미치는 영향. *관광연구저널*, 32(5), 131-144.
- 정윤지·김미정(2017). 항공사 기내서비스 품질이 고객만족과 충성도에 미치는 영향. *관광연구저널*, 31(1), 215-228.
- 조동석·리드열·이해영(2020). 호텔 서비스 평가에 대한 온라인 리뷰 빅데이터 분석: 솔로 여행자와 동반 여행자의 비교. *호텔관광연구*, 22(3), 123-136.
- 한국경제신문(2020. 11. 17). 내년 말 사라질 ‘아시아나’ 통합 LCC 브랜드로 재탄생한다. Retrieved December 28, 2020, from <https://www.hankyung.com/economy/article/2020111798481>
- 한명기·최병구(2019). 데이터 마이닝 기반 항공사 서비스 품질 분석. *한국경영정보학회 춘계학술대회 논문집*, 5월, 692-705.
- 홍지숙·오익근(2016). 소셜미디어 빅데이터를 활용한 항공사 이미지 변화 분석. *관광연구저널*, 30(6), 119-133.
- Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: capitalizing on big data. *Journal of Travel Research*, 58(2), 175-191.
- Ban, H. J., & Kim, H. S. (2019). Understanding customer experience and satisfaction through airline passengers' online review. *Sustainability*, DOI 10.3390/su11154066.
- Blei, D. M., Andrew, Y. N. G., & Michael, I. J. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(4-5), 993-1022.
- Chau, M., & Xu, J. (2012). Business intelligence in blogs: understanding consumer interactions and communities. *MIS Quarterly*, 36(4), 1189-1216.
- Gronroos, C. (2001). The perceived service quality concept: a mistake? *Managing Service Quality*, 11(3), 150-152.
- Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: tourist satisfaction analysis using

- latent dirichlet allocation. *Tourism Management*, 59(April), 467-483.
- Hofmann, T.(2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1-2), 177-196.
- Kim, W. G., Ng, C. Y. N., & Kim, Y. S.(2009). Influence of institutional DINESERV on customer satisfaction, return intention, and word-of-mouth. *International Journal of Hospitality Management*, 28(1), 10-17.
- Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V.(2019). Measuring service quality from unstructured data: a topic modeling application on airline passengers' online reviews. *Expert Systems with Applications*, 116(February), 472-486.
- Kwon, W., Lee, M., & Back, K. J.(2020). Exploring the underlying factors of customer value in restaurants: a machine learning approach. *International Journal of Hospitality Management*, 91(October), 102643.
- Lee, M. J., Singh, N., & Chan, E. S.(2011). Service failures and recovery actions in the hotel industry: a text-mining approach. *Journal of Vacation Marketing*, 17(3), 197-207.
- Li, W., Yu, S., Pei, H., Zhao, C., & Tian, B.(2017). A hybrid approach based on fuzzy AHP and 2-tuple fuzzy linguistic method for evaluation in-flight service quality. *Journal of Air Transport Management*, 60(May), 49-64.
- Lim, J., & Lee, H. C.(2020). Comparisons of service quality perceptions between full service carriers and low cost carriers in airline travel. *Current Issues in Tourism*, 23(10), 1261-1276.
- Litvin, S. W., Goldsmith, R. E., & Pan, B.(2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), 458-468.
- Liu, Y.(2006). Word of mouth for movies: its dynamics and impact on box office revenue. *Journal of Marketing*, 70(3), 74-89.
- Morash, E. A., & Ozment, J.(1994). Toward management of transportation service quality. *Logistics and Transportation Review*, 30(2), 115-140.
- Newman, D., Jey, H. L., Karl, G., & Timothy, B.(2010). Automatic evaluation of topic coherence. Paper presented at the 2010 annual conference of the North American Chapter of the Association for Computational Linguistics, June, 100-108.
- Oliver, R. L.(1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4), 460-469.
- Parasuraman, A., Zeithaml, V. A., & Berry L. L.(1988). Servqual a multi-item scale for measuring customer perceptions of service quality. *Journal of Retailing*, 64(1), 1-31.
- Park, S., Lee, J. S., & Nicolau, J. L.(2020). Understanding the dynamics of the quality of airline service attributes: satisfiers and dissatisfiers tourism management. *Tourism Management*, 81(December), 104163.
- Sezgen, E., Mason, K. J., & Mayer, R.(2019). Voice of airline passenger: a text mining approach to understand customer satisfaction. *Journal of Air Transport Management*, 77(June), 65-74.
- Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttler, D.(2012). Exploring topic coherence over many models and many topics. Paper presented at the 2012 joint conference on empirical methods in natural language processing and computational natural language learning of the Association for Computational Linguistics, July, 952-961.
- Surowiecki, J.(2005). *The Wisdom of Crowds*, New York, NY: Anchor.
- Titov, I., & McDonald, R.(2008). Modeling online reviews with multi-grain topic models. Paper presented at the 17th international conference on World Wide Web, April, 111-120.
- Ye, Q., Zhang, Z., & Law, R.(2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36(3), 6527-6535.