

# 평점 예측 모델 개발을 위한 관광지 만족도 정량 지수 구축: 제주도 관광지 리뷰를 중심으로\* \*\*

윤동규

충북대학교 빅데이터 협동과정  
(dongkyu.yun@cbnu.ac.kr)

박기태

충북대학교 빅데이터 협동과정  
(rlxogustn@naver.com)

최상현

충북대학교 경영정보학과/빅데이터 협동과정  
(chois@cbnu.ac.kr)

코로나19 팬데믹 이후 관광 산업이 회복되면서 많은 관광객들이 다양한 플랫폼을 활용하고 리뷰를 남기고 있지만, 대량의 데이터 속에서 유용한 정보를 찾기 어려워 아직도 여행지 선정 과정에서 많은 시간과 비용이 낭비되고 있다. 이에 따라 많은 연구들이 진행되고 있지만, 평점이 없거나 플랫폼별로 다른 형태의 평점 제공으로 인해 연구에 한계를 가지고 있으며, 평점과 리뷰 내용이 일치하지 않는 경우도 있어 추천 모델 구축에 어려움을 주고 있다. 본 연구에서는 이러한 문제를 해결하기 위해 7,104개의 제주도 지역 관광지 리뷰를 활용하여 제주도에 특화된 관광지 만족도 정량 지수를 개발하고 이를 활용하여 ‘평점 예측 모델’을 구축하였다. 모델의 성능을 확인하기 위해 실험 데이터 700건의 평점을 본 연구에서 개발된 모델과 LSTM을 활용하여 예측 하였으며, 제안된 모델이 LSTM 보다 약 4.67% 높은 73.87%의 가중 정확도로 성능이 더 우수한 것을 확인하였다. 본 연구의 결과를 통해 평점과 리뷰 내용 사이의 불일치 문제를 해결하고, 평점이 없는 리뷰나 다양한 형태의 평점을 정형할 수 있으며, 다른 도메인에 적용하여 여행의 모든 분야에서 신뢰할 수 있는 평점 지표를 제공할 수 있을 것으로 기대된다.

**주제어** : 관광지, 제주도, 만족도, 리뷰, 평점, 예측, 정량지수

논문접수일 : 2023년 11월 28일    논문수정일 : 2023년 12월 8일    게재확정일 : 2023년 12월 11일  
원고유형 : Fast Track    교신저자 : 최상현

## 1. 서론

최근 COVID-19(이하 코로나19) 팬데믹(pandemic) 해제 및 새로운 여행지의 지속적인 개발로 인하여 여행객 수가 급격히 증가하고 있다. 최근 정부에서도 지역 활성화, 상권 및 관광 산업 활성화를 명분으로 숙박 할인대전, 교통 지원비 등 다양한 지원을 함으로써 국내 관광 산업에 많은 지원을 하고 있는 추세이다(서운암 등, 2022). 한국 관광

데이터 랩 통계를 정리한 <표 1>을 보면, 한국에 방문한 외국 관광객은 2,922,178,516명이며, 그중에서도 제주도는 123,394,908명으로 많은 방문객들이 방문하고 있음을 알 수 있다. <표 2>와 같이 제주특별자치도 관광협회에서 제공하는 제주관광 통계의 2012년부터 2023년 9월까지 관광객 입도 현황을 보면, 실제로 내국인 통계에서는 2015년부터 꾸준히 천만 명 이상의 관광객이 입도하였다. 외국인의 경우 코로나19의 여파로 2020년 급감

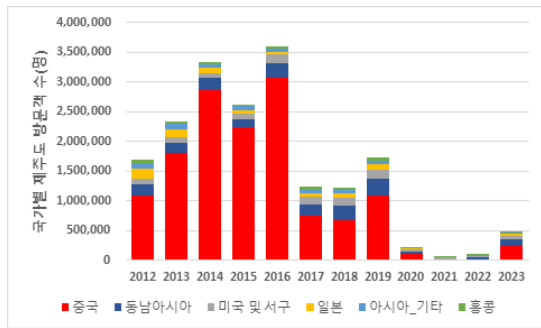
\* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2023년 문화체육관광 연구개발사업의 연구결과로 수행되었음 (RS-2023-00227532)

\*\* 이 논문은 한국전자통신연구원에서 공개한 한국어 언어모델(KorBERT)를 사용함(No.2013-2-00131, 휴먼 지식증강 서비스를 위한 지능진화형 WiseQA 플랫폼 기술 개발)

하였지만, 2023년 9월 통계를 기준으로 전년대비 443%나 증가한 443,885명이 방문한 것을 확인하였다. <그림 1>을 보면 국가별 주요 현황으로는 중국이 전체의 75%, 태국과 베트남을 포함하는 동남아시아 국가 10%, 미국 및 서구 6% 순 등으로 많은 것을 확인할 수 있다.

<표 1> 전국 및 제주지역 관광 현황(2023)

	전국	제주
외부인 방문자(b+c)	2,922,178,516 명	123,394,908 명
현지인 방문자(a)	9,518,273,730 명	485,879,250 명
외지인 방문자(b)	2,900,762,893 명	122,693,737 명
외국인 방문자(c)	21,415,623 명	701,171 명
목적지 검색건수	468,495,730 건	19,988,089 건



<그림 1> 국가/연도별 제주도 방문객 수

<표 2> 연도별 내/외국인 제주도 입도 현황

연도	내국인 관광객 수(명)	내국인 증감률(%)	외국인 관광객 수(명)	외국인 증감률(%)
2012	8,010,304	-	1,681,399	-
2013	8,517,417	6.331	2,333,848	38.804
2014	8,945,601	5.027	3,328,316	42.611
2015	11,040,135	23.414	2,624,260	-21.154
2016	12,249,959	10.958	3,603,021	37.297
2017	13,522,632	10.389	1,230,604	-65.845
2018	13,089,129	-3.206	1,224,832	-0.469
2019	13,560,004	3.597	1,726,132	40.928
2020	10,023,678	-26.079	212,767	-87.674
2021	11,960,159	19.319	48,278	-77.309
2022	13,803,058	15.409	86,444	79.055
2023.9	9,610,307	-30.376	<b>470,156</b>	<b>443.885</b>

이와 같이 제주도는 독특한 자연환경과 문화적 매력으로 국내외 수많은 여행객들의 주요 관광지 이면서 다양한 경로 선택지가 존재한다. 하지만 다양한 경로 선택지가 존재하는 만큼 수많은 여행객들은 제주도 여행지 선정 과정에서 많은 시간과 비용을 소모하고 있는 상황이다(이종수, 채동규, 2021). 따라서 수많은 관광객들은 시간과 비용을 줄이기 위해 주로 방대한 양의 정보가 있는 인터넷, 특히 여행 전문 사이트를 방문하여 기존 사용자들이 남긴 정보를 습득하여 계획에 반영한다.

실제로, 대다수의 검색엔진 사이트에서는 원하는 지역에 대해 고객들에게 추천 순, 평점 좋은 순 등 방문자들의 리뷰를 기반으로 한 관광지를 추천해 주고 있으며, 여행객들 대부분은 방문지 선정 과정에서 온라인 플랫폼을 주로 이용한다. 이러한 이유는 직접 방문하여 경험하지 않으면 파악하기 어려운 정보를 해결하기 위해 과거에 방문한 여행객들의 리뷰를 찾아 질적 속성 및 양적 속성 간의 상호작용 가능성을 고려해 보며 최종적으로 자신에게 맞는 관광지를 선정하곤 한다(노민정, 2021).

이처럼 온라인 리뷰의 품질은 리뷰의 유용성 부분에서 상당히 중요한 역할을 하며 마찬가지로 평점으로 파악하기 어려운 구체적인 정보를 제공하며 관광객의 긍정적 혹은 부정적 감성을 파악할 수 있어(Kavita et al., 2009), 잠재 여행객들의 관광지 선정을 위한 의사결정에 유용한 정보를 제공한다.

뿐만 아니라, 서비스 및 제품에 대한 만족도를 나타내는 평점 데이터는 고객들이 실제로 경험한 서비스 또는 제품에 대한 만족도를 수치화한 척도와 함께 그에 맞는 구체적인 평가와 개선점이 서술되어 있어(조호수, 류민호, 2020), 이와 같은 데이터를 통해 해당 여행지를 방문할 예정 고객들과 잠재 여행객들에게 다양한 선택지와

정보를 제공해 줄 수 있다. 또한, 평점 데이터는 업체 자체의 비즈니스 성능을 평가하고 경쟁 업체와 비교할 수 있는 중요한 지표로 활용되기에, 긍정적인 평가는 브랜드 이미지 강화와 신규 고객 유치에 긍정적인 영향을 미치며, 부정적인 평가는 서비스 품질 개선을 위한 피드백 소스로 활용되며, 지속적인 품질 향상을 이끌어낸다(강희재 등, 2019).

이와 같이 리뷰와 평점은 관광객들이 여행지를 선정하고 경로를 선택하는데 매우 중요한 정보이기 때문에 이를 활용하여 여행지 및 여행 경로를 추천해 주는 텍스트 분석 연구들이 꾸준히 진행되고 있다(박영욱, 정규엽, 2021). 하지만 관련 연구를 위해서는 리뷰와 평점을 활용해야 하는데, 소수의 여행 예약 플랫폼에서는 음식점이나 숙박업소가 아닌 관광지의 특성상 리뷰를 남기지 않아, 리뷰 데이터를 구하기가 쉽지 않은 실정이다(박소형, 조성준, 2022). 또한 우리나라의 대표적인 포털 사이트인 네이버에서 운영하는 네이버 플레이스에서는 소위 ‘별점 테러’ 등의 이유로 2021년 10월 26일을 기준으로 별점 평가에 대한 서비스가 종료되어 더 이상 평점 데이터를 확보할 수 없으며, 평점을 제공하는 타 플랫폼의 경우에도 의도적인 ‘악성 평가’ 또는 사용자가 직접 설정하기 때문에 리뷰와 평점이 일치하지 않는 등의 문제가 있어 정확한 평점 데이터를 확보하는데 어려움이 있다.

본 연구에서는 이러한 문제를 해결하기 위해 국내 관광 특화 도시인 제주도 지역 관광지의 리뷰와 평점 데이터를 활용하여 제주도 지역 관광지에 특화된 만족도 정량 지수를 구축하고, 이를 기반으로 리뷰 평점 예측 모델을 개발하고자 한다. 이러한 연구를 토대로 서비스를 이용하는 고객들에게 올바른 기존 고객 만족도 정보를 제공할 수 있으며, 다양한 관광지 추천 연구를 위한

정형화된 평점을 제공할 수 있다. 이를 통해 잠재적인 여행객들이 보다 신속하고 효율적으로 관광지를 선택할 수 있도록 도울 수 있으며, 여행지 관리자 및 정책 입안자들에게 중요한 피드백을 제공하여 관광지의 서비스 개선과 마케팅 전략 수립에 기여할 것으로 기대된다.

본 연구의 2장에서는 리뷰 데이터 분석과 관광객 만족도 분석을 위한 기존 연구들을 검토하고, 이 분야에 대한 선행연구를 심층적으로 살펴본다. 3장에서는 본 연구를 위한 데이터 수집 방법과 ‘평점 예측 모델’의 구축 과정에 대해 상세히 설명한다. 이어지는 4장에서는 개발된 모델을 활용하여 실제 리뷰 데이터의 평점을 예측하고, 모델의 성능을 다양한 측면에서 평가한다. 마지막으로, 5장에서는 연구 결과를 요약하고, 연구의 의의 및 한계점을 논의하며, 향후 연구 방향에 대해 제안한다.

## 2. 선행 연구

### 2.1 리뷰 데이터 분석 방법론

관광지의 리뷰 데이터 분석은 관광객 만족도 분석 및 관광 상품 추천 모델을 개발하는 데에 핵심적인 요소이다. 리뷰 데이터는 소비자의 솔직한 의견, 선호도 및 기대를 포함하고 있으며, 이는 관광지의 강점과 약점을 파악하는 데에 매우 중요하다(심영석, 김홍범, 2018).

리뷰 데이터는 주로 비정형 데이터 즉, 텍스트 형태로 존재하며 이를 분석하기 위해서는 감성 분석 또는 TF-IDF(Term Frequency-Inverse Document Frequency) 등 가중치 모델을 이용한다(홍태호, 2022).

### 2.1.1 리뷰 데이터의 감성 분석

감성 분석은 텍스트 데이터에서 작성자의 의도에 맞게 감정적 색채를 분석하고 분류하는 자연어 처리 기술을 의미하며, 분석 방법에 따라 크게 지식 기반 기술, 통계적 방법 그리고 하이브리드 접근 방법으로 구분할 수 있다(Devika et al., 2016).

지식 기반의 감성 분석 방법은 사전에 정의된 감정 어휘와 규칙에 기반을 둔 방법이다. 이 방법은 ‘즐거움’, ‘지루함’, ‘화남’ 과 같은 감정을 표현하는 단어를 구분하는 감정어 사전을 만들어 텍스트를 분석한다(박상민 등, 2018). 감정어 사전을 만들 때에는 객관적으로 특정 감정 상태와 명확하게 연관되도록 단어를 매칭하여야 분석할 때 단어의 존재와 맥락을 파악하여 전체적인 감정을 결정할 수 있다(Jo, 2019). 지식 기반 기술은 특히 명확하고 구조화된 데이터를 분석하는 것에 적합하며 복잡한 통계적 모델링 없이도 간단하고 효과적으로 감정을 분석할 수 있다는 장점이 있다. 다만, 사전에 정의된 규칙에만 의존하기 때문에 새로운 단어에 대해 전혀 반영하지 못하고 한국어의 미묘한 의미나 은유적, 비유적 표현이 자주 쓰이는 도메인에는 성능이 낮아질 수 있다(조은경, 남길임, 2022). 이러한 경우 성능을 고려하여 각 문서의 도메인에 맞는 맞춤형 감정사전 구축을 고려할 필요가 있다(이상훈, 최정, 2016).

통계적 감성 분석 방법은 기계 학습 알고리즘을 활용한 감성 분석 방법이다. 대표적으로 잠재 의미 분석, SVM(Support Vector Machines), Bag of Words와 같은 모델을 활용한다. 이 방법은 복잡한 연산을 할 수 있는 고급 알고리즘을 이용하여 대규모 텍스트 데이터에서 패턴을 식별하고

학습함으로써 단어의 사용 빈도뿐만 아니라 문맥을 고려하여 감정을 분석할 수 있다(Nasir et al., 2020). 특히 BoW(Bag of Words) 모델은 임베딩 기법으로 단어를  $n$  차원의 벡터로 변환하여 단어들이 어떻게 표현되는지를 분석하고, 단어 간 관계 통계적으로 정의한다. 그렇기에 복잡한 텍스트 데이터에서 구조와 의미를 파악하고 이를 바탕으로 보다 정확한 감정 분석 결과를 보여준다(어균선, 이진창, 2019). 다만, 통계적 방법도 지식 기반의 감성 분석 방법의 감정어 사전과 같이 대량의 라벨링 된 데이터를 준비하여야 한다. 만약 라벨링의 품질이 낮거나 데이터의 수가 부족하다면 모델의 신뢰성이 저하된다.

하이브리드 감성 분석 접근법은 지식 기반 기술의 감정어 사전과 통계적 방법의 기계 학습 알고리즘을 결합하여 두 방식의 장점을 모두 이용하는 방법이다(채영길, 이종혁, 2022). 이 방법은 온톨로지와 의미 네트워크를 사용하여 텍스트의 의미 구조를 분석한다. 온톨로지란 단어들의 개념과 속성, 관계를 포함하는 체계적인 언어 표현 방식으로 문장 안에서 개념이 어떻게 연결되어 있는지를 파악하는 데 사용하며, 의미 네트워크는 단어와 개념 간의 연결을 시각적으로 표현하여 감정적 의미와 문맥적 관계를 이해하는 데 사용한다(Shi et al., 2017).

하이브리드 감성 분석의 경우, 많은 양의 텍스트 데이터가 필요하고 보다 심층적인 학습이 필요하기 때문에 RNN, LSTM 등 딥러닝 알고리즘을 적용하기도 한다(김백기, 장경배, 2023). 이러한 방법은 문맥상의 복잡성을 보다 잘 해석할 수 있어 감성 분석의 정확도와 신뢰성을 크게 높일 수 있다. 다만, 각기 다른 접근법을 통합하여야 하기 때문에 모델을 최적화하는 데에 시간적·기술적 노력이 필요하다(채영길, 이종혁, 2022).

### 2.1.2 단어 가중치 모델 : TF-IDF

단어 가중치 모델이란 문서에서 사용된 단어에 가중치를 할당하여 단어마다의 중요도를 정량화하는 기술이다. 문서 내에서 단어의 순서나 문맥은 배제하고 단순히 단어가 사용된 횟수를 계산하는 단어 빈도 모델(Term Frequency Model)이 가장 기본적인 형태의 단어 가중치 모델이라고 볼 수 있다(Liu et al., 2018).

TF-IDF(Term Frequency-Inverse Document Frequency)는 대표적인 단어 가중치 모델이다. TF-IDF의 TF(Term Frequency)는 식 (1)과 같이 단어의 출현 횟수로 문서 내에서 단어가 얼마나 자주 등장하는지를 의미하며, IDF(Inverse Document Frequency)는 식 (2)와 특정 단어가 전체 문서에 등장한 횟수로 전체 문서에서 특정 단어가 얼마나 드물게 나타나는지를 측정한 값을 의미한다(Eisenstein, 2019). 즉, TF-IDF는 식 (3)과 같이 단어 빈도(TF)와 역 문서 빈도(IDF)로 구성되며 TF와 IDF의 값을 곱하여 계산한다.

$$TF(t, d) = \frac{\text{number of times } t \text{ appears in } d}{\text{total number of terms in } d} \quad (1)$$

$$IDF(t) = \log \frac{N}{1+df} \quad (2)$$

$$TF-IDF(t, d) = TF(t, d) * IDF(t) \quad (3)$$

단어 가중치 모델은 단어의 중요도를 분석하기 때문에 불용어를 처리하는 작업에서 이용할 수 있으며, 일반적으로 문서를 분류하거나 텍스트 기반의 추천 시스템 개발 또는 감성 분석 등 텍스트 마이닝을 진행할 때 데이터의 특징을 추출하는 전처리 단계에서 활용된다(노성화, 김정, 2022).

본 연구에서는 TF-IDF 알고리즘을 변형하여 관광객이 직접 작성한 리뷰와 평점을 이용해서

평점별로 특정 단어의 중요도를 파악하고, 관광지에서 특화된 정량 지수를 구축하여 텍스트(리뷰)의 평점 정보를 자동으로 생성하는 모델을 설계하고자 한다.

## 2.2 관광객 만족도 분석 연구

인터넷과 소셜 미디어의 확산 및 코로나19 팬데믹의 해제로 관광객들은 다시금 관광지에 대한 방대한 양의 정보에 노출되고 있다. 이러한 배경은 관광객에게 많은 정보를 얻을 수 있는 기회이기도 하지만, 많은 정보 중에서 관광객이 원하는 정보를 찾아내기까지 시간이 많이 소요되고 또한 거짓 정보로 인해 오히려 만족스러운 정보를 찾기 어려운 실정이다(조재현, 2023). 때문에 많은 연구자들은 방대한 데이터에서 관광객에게 개인화된 여행 상품을 추천해 주기 위한 관광 상품 추천 모델을 연구하고 있다.

관광 상품 추천 모델을 개발하는 데에 가장 필요한 정보는 관광객의 만족도이다. 만족도는 관광지에 대한 강점과 약점을 파악하는 데에 효과적이며, 분석을 통해 관광 상품 서비스를 개선하고 개인화된 추천을 가능하게 한다. 즉, 관광객의 만족도를 분석하여야 고객 경험을 향상시키고 장기적인 고객 충성도를 증진시키는 데에 중요한 역할을 한다(박지수, 2021). 그렇기에 리뷰 데이터를 가지고 소비자의 만족도를 분석하려는 연구가 많이 진행되었다.

박지원 등(2022)은 관광객이 만족도를 평가할 때 어떠한 요소를 고려하는지 분석하기 위해 한국의 트립어드바이저 웹사이트에서 수집한 온라인 리뷰 데이터를 가지고 관광지 선택 속성과 관광객의 만족도 간 관계 분석을 진행하였다. 해당 연구는 리뷰 데이터에서 관광지 선택 속성을 도출

하기 위해 TF-IDF 가중치가 부여된 LDA(Latent Dirichlet Allocation) 토픽 모델링을 사용하였으며, 토픽 모델링의 결과를 FsQCA(Fuzzy-set Qualitative Comparative Analysis) 분석으로 확장하여 리뷰 별로 할당되는 토픽의 비율 값과 관광지 만족도 점수 간 관계를 확인할 수 있었다.

현지연 등(2019)는 온라인 리뷰의 감성분석과 평점을 결합하여 추천 시스템의 성능을 향상시키고자 하였다. 해당 연구는 감성 분석을 위해 도메인 맞춤형 감성 사전을 구축하였으며 Lasso regression, Ridge regression, ElasticNet 등 알고리즘을 이용하여 감성 점수를 산출하였다. 그리고 평점을 가중치로 하여 감성 점수와 결합하여 사용자의 의견을 반영한 조정된 감성 점수 값을 도출할 수 있었다. 고안한 방법을 paired t-test 검정으로 비교하여 전통적인 협업 필터링 방식보다 우수함을 확인하였다.

김은미(2021)는 관광지에 대한 정확한 수요예측을 위하여 관광객의 리뷰 데이터 기반의 감성 분석 연구를 진행하고, 감성분석 결과 기반의 관광지 수요예측 모델을 개발하였다. 연구를 위해 서울 관광지 10곳의 온라인 리뷰 29,467개에 대한 평점, 리뷰 데이터를 활용하여 감성 분석을 진행하고 긍·부정에 대한 감성 값을 도출하였다. 그 다음, ARIMA와 RNN 등 기계학습 알고리즘으로 온라인 리뷰가 반영된 관광지 수요 예측 모델을 개발하였으며 MAPE 0.13으로 우수한 성능을 보였다.

다만 이러한 연구들은 리뷰뿐만 아니라 평점도 같이 표기되어 있어야 하는 상황에서만 분석이 가능하다는 한계가 있다. 대다수의 관광지 온라인 리뷰 웹사이트는 평점에 대한 일관된 기준이 없기 때문에 리뷰 데이터를 분석할 때에 각기 다른 리뷰 정보를 결합하는 데에 어려움이 있으며 또한, 평점이 있더라도 리뷰와 평점이 불일치

하는 문제가 많아 연구 모델의 신뢰성을 저하시키는 요인이 된다(윤민지 등, 2021).

홍태호(2022)는 관광객의 평점과 리뷰 내용 사이의 불일치 문제를 해결하기 위해 온라인 리뷰 평점 분류 모형을 연구하였다. 해당 연구는 불일치 정도를 정형화하기 위하여 리뷰에 대한 감성 분석으로 감성 값을 뽑고 평점과의 차이를 비교하였다. 구체적으로 TF-IDF 기반의 감성분석 결과를 변수로 선정하고 로지스틱 회귀분석, 인공신경망, SVM 등 기계학습 모델에 학습하여 평점 분류 모델을 설계하였다. 이를 통해 평점과 불일치하는 리뷰를 제거할 수 있는 방안을 제안하였다.

Yang et al. (2020)은 평점이 없는 리뷰에 대한 계산된 평점 생성과 리뷰에 작성된 소비자의 만족도를 반영한 조정된 평점을 생성하기 위하여 평점 추천 모델을 제안하였다. 해당 연구는 사용자와 텍스트에서 추출된 잠재 요인을 연결하여 공간적 컨볼루션을 수행할 수 있는 TextOG 모델을 설계하여 리뷰 내용 기반의 평점을 생성하였다. 이 모델은 리뷰에서 사용자의 행동과 상품 속성을 추출하여 계산하기 때문에 평점에 대한 신뢰성을 높일 수 있었다. 다만, 명확한 불만족과 만족에 대한 리뷰의 평점 정확도는 우수하였으나 애매한 리뷰에 대해서는 상대적으로 낮은 성능을 보였다. 이는 일반적으로 리뷰에 사용된 단어의 많은 수가 특정 클래스로 분류하기 어려운 애매한 의미를 가지고 있기 때문이다.

이에 따라 본 연구에서는 리뷰에 포함되어 있는 단어들의 독립적인 가중치를 산출 및 적용하여 리뷰에 사용된 애매한 의미를 가진 단어의 영향력을 낮춰 평점 예측에 대한 정확도와 신뢰성을 높일 수 있는 모델을 설계하고자 한다.

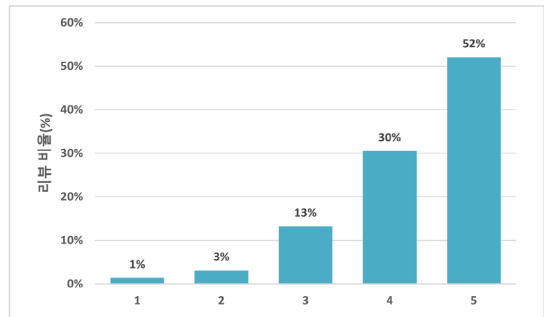
### 3. 연구 방법

#### 3.1 데이터 수집

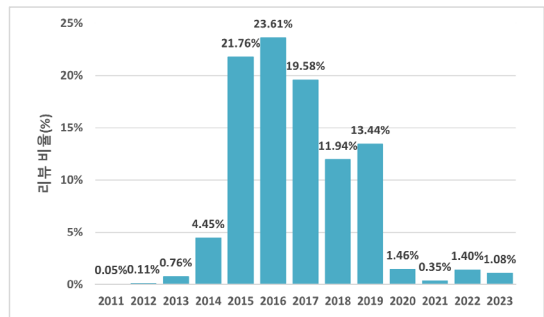
본 연구에서는 관광 명소에 특화된 만족도 정량 지수 개발을 통해 평점을 예측하여 자동으로 부여하는 모델을 개발하고자 한다. 모델 개발을 위해 세계 최대의 여행 플랫폼 중 하나인 ‘TripAdvisor’의 데이터를 웹 크롤링하여 분석에 활용하였다. 2011년 8월부터 2023년 10월까지 작성된 제주특별자치도 소재 200개 관광지에 대한 7,104개의 리뷰 데이터를 수집하였으며, 수집 항목으로는 작성 연월, 리뷰, 평점, type이다. 또한 이 플랫폼에서의 관광지 만족도 평가 방법은 사용자가 만족도에 대해 평점을 직접 설정하게 하며, 만족도가 가장 낮은 1부터 가장 높은 5까지 선택할 수 있도록 구성되어 있다. 이렇게 수집된 데이터를 기반으로 제주지역 여행자들의 만족도별 댓글 수의 분포를 확인하였다. 분석 결과 <그림 2>와 같이 평점 1 93개(1.31%), 평점 2 217개(3.06%), 평점 3 936개(13.18%), 평점 4 2,163개(30.45%), 평점 5 3,694개(52.01%)로 각 평점별로 상이하였으며, 특히 평점 4와 5에 집중되어 있는 것을 확인하였다. 따라서 본 연구자는 평점별로 상이한 분포율을 해결하기 위해 만족도를 정량 지수화하고, 정규화 함으로써 평점별로 비중 차이를 보완할 수 있었다.

한편, 근 3년간 코로나19로 인해 국내 여행객이 많이 감소했을 것을 감안하여 전체 데이터 리뷰 개수 대비 연도별 리뷰 개수 비율을 확인함으로써, 국내 관광 산업의 동향을 확인하였다. <그림 3>을 보면, 데이터의 연도별 리뷰의 수는 2011년부터 2016년까지 전체 데이터 기준 누적 46.7%로 꾸준히 증가하였지만, 중국의 고고도 미사일 방어체계(THAAD, 사드) 배치 보복에 따른 한국 관광 전면 금지 조치가

있었던 2017년부터 감소세를 보였다. 또한, 예상했던 것과 마찬가지로 코로나19의 영향을 받은 2020년에는 전체 댓글 수의 1.58%, 2021년에는 0.41%까지 감소한 것을 볼 수 있으며, 2022년부터 다시 약간의 증가세를 보이는 것을 확인할 수 있었다.



<그림 2> 평점별 리뷰 비율(전체 데이터 대비)



<그림 3> 연도별 리뷰 비율(전체 데이터 대비)

리뷰와 평점 외에 추가적으로 수집한 type은 어떤 인원 구성으로 여행을 왔는지 확인할 수 있는 데이터이며 가족, 개인, 비즈니스, 친구, 커플 총 5개의 종류로 분류되어 있다. 전체 데이터 대비 가장 많은 비율을 차지한 type은 가족이 30.24%로 가장 많았고 커플(20.02%), 친구(22.86%), 미기업(13.11%), 개인(11.98%), 비즈니스(1.80%) 순으로 분포되어 있었다. type은 전체 데이터의 13.11%인 약 931개의 데이터가 사용자가 기입하지 않은

결측치임을 확인하였으며, 이를 대체할 수 있는 데이터를 확보하기 어려워 본 연구에서는 분석 대상에서 제외하였다.

### 3.2 모델 구축

#### 3.2.1 형태소 분석

제주도 관광지 리뷰 데이터는 분석에 사용하기 위해 먼저 한 글자 단어와 특수문자를 제거하였고, 불필요한 띄어쓰기가 되어 있는 부분을 제거하여 기본적인 전처리를 진행하였다.

이후 ETRI(Electronics and Telecommunications Research Institute, 한국전자통신연구원)에서 API 형태로 제공하는 KorBERT를 활용하여 문장 형태의 리뷰 데이터를 한국어에서 의미를 가지는 최소 단위인 형태소 단위로 분절하였다. 본 연구에서 활용한 Korean\_BERT\_Morphology는 KorBERT의 형태소 분석 기반의 세부 언어 모델로 신문기사와 백과사전 등 23GB의 원시 말뭉치(47억 개 형태소)를 학습한 한국어 언어학습 모델이다. <그림 4>는 ETRI에서 개발한 형태소 기반 언어 모델의 분석 단위를 시각화한 것이다.



<그림 4> ETRI 형태소 기반 언어 모델 분석 단위

연구에 사용된 형태소는 API 레퍼런스에 따라 TTA 표준 형태소 태그셋(TTAK.KO-11.0010/R1)에 기반하여 대분류, 소분류, 세분류, 세종 태그 중 세종 태그를 사용하였으며, 사용자의 관광지 만족도를 파악할 수 있는 일반명사(NNG), 형용사(VA), 일반부사(MAG) 품사만 사용하였다. KorBERT를 활용한 형태소 분석 결과인 <표 3>에서 평점 1의 리뷰에서 추출한 단어들은 아쉽, 실망, 혼잡, 비하 등 부정적인 반면, 평점 5에서 추출한 단어에서는 매력, 하이라이트, 추천 등 긍정적인 단어들이 상대적으로 많이 출현하는 것을 확인할 수 있다.

<표 3> KorBERT를 활용한 형태소 분석 결과(예시)

Name	Score	Morp
우도	1	관광:NNG, 객:XSN, 이:JKS, 너무:MAG, 많:VA, 어서:EC, 조용:NNG, 하:XSA, 게:EC, 경치:NNG, 를:JKO, 즐기:VV, ㄷ:ETM, 수:NNB, 없:VA, 었:EP, 고:EC, ,:SP, 일부:NNG, 해변:NNG, 의:JKG, 청결:NNG, 상태:NNG, 도:JX, 아쉽:VA, 었:EP, 어요:EF, ,:SF, 또한:MAG, ,:SP, 맛:NNG, 집:NNG, 과:JC, 카페:NNG, 가:JKS, 혼잡:NNG, 하:XSA, 고:EC, 서비스:NNG, 도:JX, 기대:NNG, 예:JKB, 못:MAG, 미치:VV, 었:EP, 습니다:EF, ,:SF, 제주도:NNP, 의:JKG, 다른:MM, 명소:NNG, 들:XSN, 예:JKB, 비하:VV, 어:EC, 다소:MAG, 실망:NNG, 스럽:XSA, 었:EP, 어요:EF, ,:SF
우도	5	우도봉:NNP, 에서:JKB, 바라보:VV, ㄴ:ETM, 탁:MAG, 트이:VV, ㄴ:ETM, 바다:NNG, 전망:NNG, 과:JC, 울창하:VA, ㄴ:ETM, 자연:NNG, 경관:NNG, 이:JKS, 인상:NNG, 적:XSN, 이:VCP, 었:EP, 습니다:EF, ,:SF, 특히:MAG, 하:VV, 고:EC, 수동:NNG, 해변:NNG, 의:JKG, 에메랄드:NNG, 빛:NNG, 바다:NNG, 는:JX, 매우:MAG, 매력:NNG, 적:XSN, 이:VCP, 었:EP, 어요:EF, ,:SF, 우도:NNP, 는:JX, 여행:NNG, 의:JKG, 하이라이트:NNG, 중:NNB, 하나:NNG, 로:JKB, ,:SP, 다시:MAG, 방문:NNG, 하:XSV, 고:EC, 실:VX, 은:ETM, 못:NNG, 이:VCP, 예요:EF, ,:SF, 강력:NNG, 추천:NNG, 하:XSV, ㅂ니다:EF, !:SF



### 3.2.2 관광지 리뷰에 특화된 만족도 정량 지수

#### 3.2.2.1 만족도 정량 지수 도출 방법

앞서 형태소 분석을 통해 필요한 품사를 추출한 후 문서 내 단어의 중요도를 나타내는 가중치인 TF-IDF와 대상 형태소로 분절된 평점별 리뷰 데이터 셋을 활용하여 개별 평점/리뷰별로 각 단어가 문서에서 얼마나 중요한지를 나타내는 TF-IDF 가중치 값을 추출하였다.

연구에서는 개별 리뷰 간 중요 단어를 추출하고자 하는 것이 아닌 전체 문서 내에서 중요한 단어를 추출하는 것이 핵심이므로 식 (4)와 같이 평점 별 문서 내에서 추출한 각 단어의 개별 TF-IDF 가중치 값은 모두 더하였다.

$$Sum-TF-IDF(w,s) = \sum_{des} TF-IDF(w,s,D) \quad (4)$$

하지만 이렇게 단순히 더한 가중치 값을 그냥 사용할 경우 리뷰의 수가 압도적으로 많은 평점 4와 5의 모든 단어에서 가중치가 높게 나오게 된다. 따라서 <표 4>의 예시와 같이 식 (5)를 적용

하여 각 평점별 가중치가 전체 리뷰의 수에 영향을 받지 않고 각 평점 그룹에서 추출된 단어의 상대적 중요도를 반영할 수 있도록 가중치를 문서의 수로 나누고 100을 곱하여 정규화 된 새로운 가중치 값을 부여하였다.

$$Norm-TF-IDF(w,s) = \left( \frac{Sum-TF-IDF(w,s)}{\text{number of review in } s} \right) \times 100 \quad (5)$$

정규화 된 가중치를 부여한 각 평점별 데이터 셋은 weight를 기준으로 내림차순 정렬하여 가중치가 높은 순으로 'rank'를 부여하였으며, 이후 평점별 5개의 데이터 셋을 병합하여 총 14,959개의 단어를 가진 하나의 통합 데이터 셋으로 구축하였다.

#### 3.2.2.2 중복 단어 제거

본 연구에서는 리뷰에 나타나는 모든 단어를 사용하는 것이 아닌 평점별로 핵심 단어의 가중치만 활용하여 새로운 리뷰의 평점을 부여하기 위해 중복되는 단어를 제거하였다. 먼저 각 별점

<표 4> 관광지 리뷰 만족도 정량 지표 데이터 셋 예시

평점 2 데이터 셋 단어 TF-IDF 가중치 추출 및 합계						완성된 데이터 셋(예시 : 평점 2)				
	별로	실망	솔직히	불거리	그냥	Score	Rank	Word	Weight	Scaled weight
review 1	0.117	0.180	0.166	0.231	0.148	2	1	그냥	11.909	2.381
review 2	0.132	0.164	0.259	0.231	0.174		2	불거리	5.096	1.019
review 3	0.105	0.125	0.211	3.455	0.164		2	솔직히	2.457	0.491
review 4	0.252	0.151	0.198	0.187	5.812		2	별로	1.971	0.394
review 5	0.399	0.254	0.270	0.148	0.482		2	실망	1.874	0.374
review 6	0.190	0.110	0.252	0.224	0.363					
review 7	0.162	0.118	0.264	0.150	0.845					
review 8	0.148	0.274	0.206	0.144	0.231					
...	...	...	...	...	...					
review n	0.214	0.207	0.428	0.149	3.455					
<b>Summary</b>	<b>1.971</b>	<b>1.874</b>	<b>2.455</b>	<b>5.096</b>	<b>11.909</b>					

데이터 셋에서 rank가 높은 단어는 해당 별점 데이터에서 중요도가 높은 단어라고 판단하여 rank를 기준으로 가장 높은 단어만 남기고 중복 제거를 하였다. 또한 평점별로 rank가 같은 경우가 있는데 정규화 된 weight 값을 활용하여 가중치 값이 가장 높은 단어만 남기고 중복되는 단어는 데이터 셋에서 제거하였다.

중복되는 단어를 제거하는 과정을 설명하는 예시 데이터 셋인 <표 5>를 보면 ‘시간’과 ‘그냥’이라는 단어는 평점 1, 평점 2, 평점 3에 모두 출현하는 단어이다. 먼저 초록색으로 표시된 ‘그냥’이라는 단어는 단순히 rank 순으로만 비교하였을 때 평점 2가 1위로 가장 높기 때문에 평점 2를 제외한 다른 별점 리스트에서 해당 단어를 제거하였다. 다음으로 노란색으로 표시된 ‘시간’이라는 단어는 rank를 기준으로 봤을 때 평점 1과 평점 3에서 같은 순위를 가진다. 이러한 문제를 해결하기 위해서는 weight 값을 비교하여 더 높은 weight 값을 갖는 평점 1의 단어 리스트에만 포함하고, 마찬가지로 다른 별점 리스트에서는 제거하였다.

<표 5> 중복 단어 제거 과정 예시

Rank	Score1 word	Score1 weight	Score2 word	Score2 weight	Score3 word	Score3 weight
1	시간	3.875	그냥	3.153	사관	2.621
2	크냥	3.739	사관	2.951	관광	2.402
3	여행	3.179	방문	2.867	사람	2.330
4	정말	3.048	관광	2.736	방문	2.281
5	관광	2.899	생각	2.696	생각	2.121
6	사람	2.841	너무	2.528	너무	2.013
7	방문	2.554	추천	2.180	사진	1.967
8	추천	2.471	여행	2.092	크냥	1.740
9	생각	2.339	도로	1.926	많이	1.708
10	아갑	2.044	입장	1.818	아이	1.624

위 방법과 같이 중복되는 단어를 제거한 후 8,115개의 평점과 중요도에 따른 가중치가 부여된 단어 데이터 셋을 활용하여 ‘관광지에 특화된 만족도 정량 지수’를 개발하였으며, 평점별 ‘만족도 정량 지수’ 데이터 셋의 단어 특성을 확인하기 위해 워드 클라우드(Word cloud) 기법을 활용하여 시각화하였다.

3.2.2.3 평점별 만족도 정량 지수 시각화 분석



<그림 5> 평점 1 단어와 weight를 활용한 시각화

평점 1의 단어들을 시각화 한 <그림 5>를 보면 ‘아갑’, ‘낭비’, ‘기분’, ‘비싸’, ‘최악’과 같이 전반적으로 단어들이 매우 부정적이며, 평점 1의 실제 리뷰를 살펴보면 주로 가격이 비싸거나 가이드 또는 여행 상품 패키지로 인해 기분이 좋지 않은 리뷰들이 포함된 것을 확인할 수 있었다.



<그림 6> 평점 2 단어와 weight를 활용한 시각화



‘NEG’ 그룹으로 구성하였으며, 평점 3은 중립적인 단어 위주의 리뷰로 구성되어 있어 ‘NEU’ 그룹으로 구성하였다. 평점 4와 평점 5는 관광지에 만족하여 긍정적인 단어 위주의 리뷰가 많았기 때문에 ‘POS’ 그룹으로 구성하였다. 이 그룹을 활용하여 모델 성능 평가 시 정확도 산식에서 가중치를 부여하였다.

### 3.2.3 만족도 정량 지수 기반의 평점 예측 모델

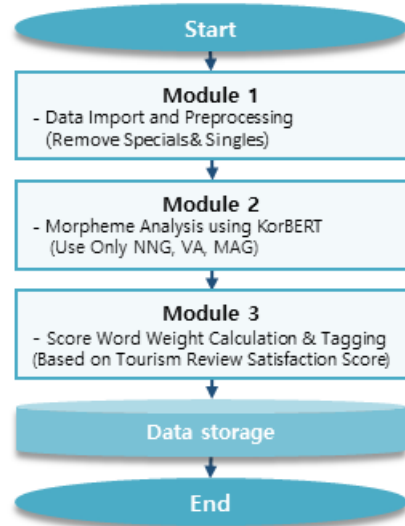
본 연구에서 제안하는 모델은 제주도 관광지 리뷰의 평점을 예측하는 모델로서 제주도 관광지 리뷰 데이터를 기반으로 개발한 ‘관광지 리뷰 만족도 정량 지표’를 적용하여, 새로운 리뷰 데이터에 대해 1에서 5까지의 평점을 예측한다.

본 모델의 모델 구축 과정은 세 가지 주요 모듈로 설계하여 진행되었다. 첫 번째 단계는 입력된 리뷰 데이터를 정제하는 ‘데이터 전처리 모듈’로, 데이터 정제 과정에서 한 글자 단어와 특수문자를 제거하여 기본적으로 데이터의 품질을 향상시키고, 분석의 효율성을 높였다.

두 번째 단계는 Korean\_BERT\_Morphology를 활용하여 리뷰 데이터에서 의미 있는 형태소를 추출하는 ‘형태소 분석 모듈’로, 특히 일반명사(NNG), 형용사(VA), 일반부사(MAG) 등 필요한 품사에 초점을 맞추어 한국어의 복잡한 문법적 특성을 고려한 분석을 수행하는 모듈이다.

마지막 단계는 ‘평점 예측 및 데이터베이스 저장’ 모듈이다. 해당 모듈은 먼저 리뷰 내 각 단어에 대한 평점과 weight를 ‘관광지 리뷰 만족도 정량 지수’에서 찾아, 리뷰 문장 전체에 걸쳐 각 평점의 weight를 합산한 후 가장 높은 weight를 가진 평점을 해당 리뷰에 매칭(Matching)하고, 최종 결과를 데이터베이스에 저장하는 모듈이다.

제안하는 모델의 상세 구축 Process는 <그림 10>과 같다.



<그림 10> 평점 예측 모델 구축 Process

위에서 설명한 과정을 통해 통합된 제주도 관광지 버전의 평점 예측 모델은 기존에 널리 사용되는 딥러닝 방법과 몇 가지 중요한 측면에서 차별화된다. 평점 예측 모델은 제주도 관광지 리뷰 데이터에 특화된 접근 방식을 사용하며, 각 리뷰 내에서 평점별 단어의 중요도를 정밀하게 파악하는 데 초점을 맞추고 있다. 이 모델에서는 각 리뷰가 가진 평점을 기반으로 TF-IDF 및 변형된 Sum-TF-IDF, Norm-TF-IDF 방법을 통해 단어별 weight를 계산한다. 이러한 방식은 리뷰 내 개별 단어의 중요성을 더 세밀하게 분석하고, 리뷰의 본질적인 의미를 효과적으로 포착하는 데 유리한 방법이며, 이는 제주도 관광지에 대한 리뷰의 특성을 잘 반영하여 높은 정확도와 신뢰성을 제공하는 데 기여한다. 반면, 딥러닝 기반의 모델

들은 전체 텍스트의 맥락을 학습하는 데 강점을 가지지만, 개별 단어의 중요도와 그 의미를 구체적으로 분석하는 데는 한계를 가질 수 있다. 따라서 본 연구에서 제안하는 평점 예측 모델은 개별 단어 및 그 조합에 대한 이해를 바탕으로 보다 세분화된 분석 결과를 제공할 수 있다.

연구 결과에서는 이러한 차이점을 바탕으로, 본 연구에서 개발된 평점 예측 모델과 딥러닝 기반의 별점 예측 모델 간의 성능을 비교 분석하였으며, 비교를 통해 각 모델의 장단점을 보다 명확하게 이해하고 더 나은 리뷰 분석 방법을 제시한다.

## 4. 연구 결과

### 4.1 관광지 만족도 정량 지수를 활용한 평점 예측 모델의 성능 평가 방법

본 연구에서는 제주도 관광지 리뷰 데이터를 활용하여 평점이 없는 관광지 리뷰에 과거 리뷰 데이터로 구축한 단어별 가중치를 활용하여 평점을 생성하는 ‘평점 예측 모델’을 설계하였다.

평점 예측 모델의 성능을 확인하기 위해 관광지 만족도 정량 지수 개발에 사용하였던 전체 리뷰 데이터 7,104건의 약 10%에 해당하는 700건의 관광지 리뷰 데이터를 추가로 수집하여 시험 데이터로 사용하였다. 시험 데이터는 350건씩 2가지 종류로 확보하였으며, 본 연구에서 제안하는 모델이 제주도 지역 관광지에 특화된 모델임을 확인하기 위해 우리나라 주요 관광지인 서울, 부산, 여수, 제주, 경주의 리뷰를 50개씩 포함하는 Dataset A와 제주도 소재 관광지 리뷰로만 구성되어 있는 Dataset B로 구성하여 비교에 사용하였다.

성능 평가 방법은 먼저 본 연구에서 개발한 평점 예측 모델을 활용하여 시험 데이터의 평점을 예측한다. 이후 비교 대상으로 평점을 예측할 때 보편적으로 사용하는 LSTM 모델로 전체 데이터 7,104개를 학습하고 시험 데이터의 평점을 예측한다. LSTM에서 사용하기 위해 원본 데이터에서 한글자 단어와 특수문자를 제거하고 KorBERT로 형태소 분석하는 과정까지는 본 연구의 모델 구축 방법과 같이 진행하고, 추가적으로 토큰화와 패딩 과정을 거쳐 모델의 Input으로 사용하였다. 비교 대상으로 구축한 LSTM 예측 모델의 활성화 함수로는 ‘linear’를 사용하였다. 또한 ‘mean\_squared\_error’ 손실 함수와 ‘Adam’ 옵티마이저를 사용하여 컴파일하였으며, Adam 옵티마이저의 학습률은 0.001로 설정하였다. 주요 하이퍼 파라미터로 batch\_size는 32, epochs는 10으로 설정하였다. 이렇게 구성된 두 모델의 시험데이터 평점 예측 결과를 비교하여 성능을 평가하였다.

성능 평가 지표로는 데이터의 특성을 고려하여 가중치를 적용한 ‘Weighted Accuracy’를 사용하였다. 앞 장에서 설명한 바와 같이 만족도를 나타내는 평점의 특성상 평점 1과 2는 부정, 평점 3은 중립, 평점 4와 5는 긍정의 의미를 갖기 때문에 ‘NEG’, ‘NEU’, ‘POS’ 3가지 그룹으로 구성하였다. 이에 따라 각 모델의 성능을 평가할 때, 비슷한 감성 값을 가지는 긍정 그룹과 부정 그룹 내에서는 예측이 틀린 경우 오답으로 처리하는 것이 아니라 0.8의 가중치를 주었다. 예를 들어 정답이 평점 1인데 평점 2로 예측한 경우 같은 그룹의 평점이기 때문에 0.8의 가중치를 주지만, 평점 3이 정답인데 평점 4로 예측한 경우 다른 그룹에 해당하기 때문에 오답으로 처리한다. 가중치를 활용한 평가 지표의 계산 과정을 보여주는 소스코드는 <그림 11>과 같다.

평점 예측 모델은 사전 기반의 방법 특성상 같은 모델과 데이터를 사용할 경우 항상 정확도가 같기 때문에 한 번의 Test만 진행하여 정확도를 산출하였으며, 딥러닝 기반의 예측 방법인 LSTM의 경우 예측할 때마다 정확도의 차이가 있기 때문에 신뢰성 있는 결론은 도출하기 위해 3번의 Test의 평균 가중 정확도를 성능 비교에 활용하였다.

```
def calculate_weighted_accuracy(df):
    total_weight = 0
    for _, row in df.iterrows():
        actual, predicted = row['y'], row['pred']

        if actual == predicted:
            weight = 1.0
        elif (actual in [1, 2] and predicted in [1, 2]) or \
             (actual in [4, 5] and predicted in [4, 5]):
            weight = 0.8
        else:
            weight = 0.0

        total_weight += weight

    return total_weight / len(df)

weighted_accuracy = calculate_weighted_accuracy(df)
```

〈그림 11〉 Weighted accuracy 계산식

평점 예측 모델은 사전 기반의 방법 특성상 같은 모델과 데이터를 사용할 경우 항상 정확도가 같기 때문에 한 번의 Test만 진행하여 정확도를 산출하였으며, 딥러닝 기반의 예측 방법인 LSTM의 경우 예측할 때마다 정확도의 차이가 있기 때문에 신뢰성 있는 결론은 도출하기 위해 3번의 Test의 평균 가중 정확도를 성능 비교에 활용하였다.

## 4.2 연구 결과

본 연구에서 제안하는 평점 예측 모델의 성능 확인을 위해 앞서 설명한 바와 같이 두 가지 모델과 두 가지 시험 데이터 셋을 활용하여 가중 정확도(Weighted Accuracy)를 비교하였다.

먼저 모델별로 가중 정확도를 비교한 결과를 확인해 보면, 본 연구에서 제안하는 평점 예측 모델이 LSTM보다 가중 정확도가 약 4.67% 높아 더 우수한 모델임을 확인하였다. 다음으로 데이터 셋 구성을 기준으로 성능을 비교하면, 평점 예측 모델은 제주도를 포함한 우리나라 주요 관광도시 5곳의 관광지 리뷰 데이터를 포함하는 Dataset A보다 제주도의 관광지 리뷰로만 구성된 Dataset B의 가중 정확도가 4.69% 더 높았다. 이를 통해 평점 예측 모델은 제주도 관광지에 특화된 모델임을 확인할 수 있었다. 반면에 LSTM의 경우 Dataset A와 Dataset B의 가중 정확도의 차이가 약 0.2%로 두 데이터 셋 간의 성능 차이가 없음을 확인하였다.

본 연구의 결과를 정리한 <표 6>을 보면, 평점 예측 모델(Proposed Model)을 활용하여 제주도 지역 관광지 데이터로 구성된 Dataset B의 평점을 예측했을 때, 가중 정확도가 76.22%로 가장 성능이 좋은 것을 확인할 수 있다.

〈표 6〉 모델 및 데이터셋별 성능 비교

(Weighted accuracy, %)

	Proposed Model	LSTM	Average	
<b>Dataset A</b> (Korea's 5 Major Tourist Cities)	71.53	Test 1	69.1	70.31
		69.31		
		Test 2		
		68.93		
Test 3	69.05			
<b>Dataset B</b> (Jeju Island)	76.22	Test 1	69.3	72.76
		69.35		
		Test 2		
		68.21		
Test 3	70.35			
Average	73.87	69.2		

다음으로 예측 결과의 일부를 추출하여 두 모델의 특징을 확인하였다. 먼저 평점 예측 모델의 결과를 보여주는 <표 7>의 case 1을 보면, 제주도의 관광지 리뷰에서 추출한 단어들을 평점별로 가중치를 주었기 때문에 전반적으로 각 평점마다의 특징을 잘 파악하여 예측한 것을 볼 수 있다. 평점 4의 morp을 보면, 대부분 긍정의 의미와 약간의 아쉽다는 느낌의 단어들을 잘 반영하여 예측했으며, 특히 긍정과 부정의 의미가 없는 중립적인 단어들을 많이 포함하고 있는 평점 3을 예측하였을 때도 성능이 좋은 것을 확인하였다.

하지만 잘못 예측한 경우를 보여주는 case 2를 보면 짧은 리뷰의 경우 대체적으로 정확한 평점

을 계산할 수 있는 단어의 수가 부족하기 때문에 특정 평점 단어의 가중치가 높을 경우 성능이 좋지 않은 경우를 확인하였다.

다음으로 LSTM의 평점 예측 결과에 대한 분석을 진행하였다. <표 8>의 case 1을 살펴보면, 본 연구에서 구축한 예측 모델과 다르게 상대적으로 짧은 리뷰를 잘 예측하는 것을 볼 수 있다. 그러나 case 2에서는 LSTM 모델이 대부분의 경우에 평점 3을 잘못 예측하는 경향을 보였으며, 이는 주로 3으로 예측되어야 할 리뷰를 다른 평점으로 잘못 분류하는 문제점을 드러낸다. 이러한 결과는 중립적인 평점, 즉 평점 3을 나타내는 리뷰의 단어들의 특성을 LSTM 모델이 충분히

<표 7> 제안된 모델을 활용한 평점 예측 결과

Case	Morp	Y	Pred
1	없 가이드 설명 제품 구매 설명 대놓고 제품 구매 곳 안내 토속 모습 구경 기분 그냥 지자체 관리 제대로 안 좋 기억 곳 같 굉장히 불쾌하 기억 사진 없 추천	1	1
	입장 좋 관광 호객 행위 좀 골치 아프 차 라리 입구 부스 필요 사람 안내 요청 좋	2	2
	종료 모두 노란색 갈색 많이 있 권장 있 월 경치 그다지 좋 여전히 재미	3	3
	아이 좋 바다 작 아이 오히려 파도 약 하 정도 주차 조금 더 여유 있 좋 해수 욕 영역 옆 물론 바다 좋 공식 해수 욕 화장실 안전 시설 없 아쉽..	4	4
	에메랄드 빛 바다 도 아름답 해변 변화하 이유 충분하 때 새롭 가게 해변 서 해변 가장 색 아름답	5	5
2	차 좋 곳 나이 등산 어렵 정도 같	4	2
	위 아름답 도착 약간 시간 즐겁 시간	4	1
	협제 여름 수영 켈 좋 여름 왜 가격 비싸 이유 수영..	5	1
	비 날 아무 것 없 서운하 커피 한잔 좋 날 다시 해수 욕 좋 같	5	1
	물 에메랄드 색 아름답 해변 너무 사람 많 너무 많 파라솔 상점 실망 표준 서로 다르 간단하 음식 사람 그늘 있 곳	2	5

<표 8> LSTM을 활용한 평점 예측 결과

Case	Morp	Y	Pred
1	굉장하 자부 애정 민속 마을 정말 신 분 목 마치 공식 안내 접근 멀리 바가지 접근 장사 도 민속 마을 호객 행위 바가지 대책 조속히 마련 이미지 전부	1	1
	경로 꽤 실망 겨울 어찌면 가을 예쁘	2	2
	민속 마을 한번 곳 가족 단위 커플 간단하 음식 상당 비싸 맛 별로 강 옛날 모습 생각 구경 예요	3	3
	항구 도착 전 온 라인 예약 투어 민속 마을 마을 방문 포함 마을 있 입 장 마을 주변 오래 건물 농 ...	4	4
	매혹 마을 아름답 집 매우 흥미 곳 방문	5	5
2	제품 구매 곳 안내 토속 모습 구경 기분 그냥 지자체 관리 제대로 안 좋 기억 곳 같 굉장히 불쾌하 기억..	1	3
	시간 낭비 황량하 곳 절대 실망 더 좋 곳 있	1	3
	곳 광고 이해 그렇 일 요일 시간 있 많 귀엽 작 집 얼마나 내구 흥미 경 우 기본 스트 진흥 고려	1	3
	꽤 실망 곳 많 있 가이드 설명 기록 영업 이야기 예상 로컬 제품 계속 빠 오일 호스 특화 정보 차	1	3
	차 귀도 섬 바다 일몰 유명 최근 모래 모래 사장 비닐 아쉽 여전히 예쁘 바다	3	5

과약하지 못했음을 의미한다. 이러한 결과는 LSTM 모델이 리뷰 텍스트의 복잡성과 다양성을 완전히 이해하고 처리하는 데에 있어 일정한 한계를 가지고 있음을 시사한다. 특히, 중립적인 의미를 담은 리뷰의 경우, 특정한 감정이나 평가가 명확하게 표현되지 않기 때문에 모델이 이러한 리뷰의 내용을 정확히 이해하고 평점을 예측하기 어렵다. 또한, LSTM은 주로 시퀀스의 맥락을 파악하는데 중점을 두기 때문에, 각 단어의 미묘한 의미나 문맥상의 뉘앙스를 정확히 포착하는 데에는 한계가 있다.

이와 더불어 LSTM 모델이 장기 의존성을 다루는 데 탁월한 성능을 보이긴 하지만, 실제로는 매우 긴 시퀀스 데이터에서는 이러한 장점이 충분히 발휘되지 않을 수도 있다. 이는 텍스트의 길이가 길어질수록 모델이 중요한 정보를 유지하고 기억하는 데 어려움을 겪을 수 있기 때문이다. 따라서, 리뷰의 길이가 길거나 내용이 복잡한 경우, LSTM 모델은 중요한 특징을 놓치거나 오해할 가능성이 있다.

결론적으로, LSTM을 사용한 리뷰 텍스트 기반의 평점 예측 모델은 간결하고 명확한 표현을 담은 리뷰에서는 비교적 높은 정확도를 보이지만, 중립적인 의미나 복잡한 내용을 담은 리뷰에 대해서는 예측 정확도가 상대적으로 낮을 수 있다는 한계가 존재한다는 것을 확인하였으며, 본 연구에서 제한한 예측 모델의 우수성을 확인할 수 있었다.

<표 9>는 이를 검증하기 위해 제주도 지역 관광지 리뷰 데이터로만 구성된 Dataset B에서 각 평점 그룹(NEG, NEU, POS)별로 두 모델이 잘 예측한 비율을 비교하는 표이다. 상대적으로 명확한 극성 값을 가지는 그룹인 NEG와 POS에서는 두 모델 다 높은 예측 성능을 보이지만 NEU

그룹에서는 본 연구에서 제안한 모델이 약 1.8배 정확도가 높은 것을 확인할 수 있다.

<표 9> 모델별 각 평점 그룹을 정확히 예측한 비율

Dataset / Group		Proposed Model	LSTM	Average
Dataset B (Jeju Island)	NEG (평점 1, 2)	83.06	84.62	83.84
	NEU (평점 3)	56.31	31.53	43.92
	POS (평점 4, 5)	89.28	91.76	90.52
Average		76.22	69.30	

## 5. 결론

코로나19의 팬데믹이 종료됨에 따라 관광 산업은 다시금 호황을 누리고 있으며, 수많은 관광객들이 다양한 플랫폼을 통해 자신의 경험을 리뷰로 남기고 있다. 하지만 그럼에도 불구하고 대다수의 관광객들은 방대한 데이터 속 만족스러운 정보를 찾지 못해 관광지를 찾는데 많은 시간과 비용을 낭비하고 있다. 이에 따라 최근 관광지를 추천할 수 있는 다양한 연구가 진행되고 있지만, 사용자의 만족도를 알 수 있는 지표인 평점이 없거나 플랫폼마다 제공하는 형태가 달라서 추천 모델을 구축하는데 한계점이 있다. 또한 평점과 리뷰 내용이 불일치한 경우가 있는 데이터의 경우 그대로 추천 모델 개발용으로 사용하기에 어려움이 있다. 본 연구에서는 이러한 문제를 해결하기 위해 관광지에 방문한 사용자 리뷰와 평점을 활용하여 관광지에 특화된 만족도 정량 지수를 개발하고 평점 예측 모델 설계하였다.



만족도 정량 지수를 개발하고 평점 예측 모델을 설계하기 위해 온라인 여행 플랫폼 TripAdvisor에서 약 12년 동안 수집된 국내 관광 특화 도시인 제주도 관광지에 대한 7,104개의 온라인 리뷰 데이터를 수집 및 활용하였다. 먼저 리뷰 데이터셋을 평점별로 나누고 TF-IDF를 활용하여 출현하는 단어의 TF-IDF weight 값을 도출한 후 모두 더하여 각 평점에서의 단어 중요도를 확인하였다. 마지막으로 전체 데이터 대비 각 리뷰 데이터 수의 비율을 고려하여 각 weight 값을 리뷰의 수로 나누어 정규화 된 weight 값을 도출하였다. 마지막으로 rank와 weight에 따라서 서로 중복되는 단어를 제거하여 제주지역 관광지에 특화된 ‘만족도 정량 지수’를 개발하였다. 이후 ‘만족도 정량 지수’를 활용하여 새로운 리뷰 데이터에 활용할 수 있는 평점 예측 모델을 개발하였다.

모델의 성능을 시험하기 위해 전체 데이터의 10%에 해당하는 700건의 시험 데이터를 서울, 부산, 여수, 제주, 경주 등 5가지 우리나라 주요 관광지 리뷰가 섞여있는 Dataset A 350건과 제주도 관광지 리뷰만으로 구성된 Dataset B 350건으로 추가 수집하였다. 이후 평점 예측 모델과 비교 모델인 LSTM으로 해당 데이터들의 평점을 예측하여 성능을 확인하였다. 예측 결과, 본 연구에서 개발한 평점 예측 모델로 제주도 데이터만 사용한 Dataset B의 평점을 예측했을 때, 가중 정확도 76.22%로 가장 성능이 좋은 것을 확인하였다. 평점 예측 모델은 각 평점별로 중요한 단어를 계산하여 추출하고 가중치를 사용하였기 때문에 긍정적인 단어들과 부정적인 단어들이 섞여 있는 애매한 리뷰에서도 성능이 좋았지만, LSTM은 대부분의 잘못 예측한 값들이 평점 3점으로 예측하거나 평점 3점을 다른 값으로 예측하는 등 단어의 극성이 애매한 경우 예측 성능이

매우 낮은 것을 확인하였다.

이러한 모델을 활용하면 관광객의 평점과 리뷰 내용 사이의 불일치 문제를 해결할 수 있을 뿐 아니라 평점이 없는 리뷰나 각기 다른 형태의 평점을 정형화하여 사용할 수 있다는 장점이 있다. 또한 제주도뿐만 아니라 국내 다른 관광지 또는 해외 관광지에 대한 특화된 평점 예측 모델을 개발하여 사용할 수 있다. 나아가 이러한 정량 지표 방법론을 확장하여 식당, 숙소, 이동 수단 등 다양한 도메인의 텍스트 리뷰 데이터를 만족도 정량 지수로 개발하여 적용한다면, 여행에 대한 모든 분야에서 정형화된 평점을 보고 관광지를 선택할 수 있는 지표가 될 수 있을 것으로 예상된다.

다만 본 연구의 평점 예측 모델은 평점별로 심한 데이터의 불균형을 정규화로 완전히 해결하지 못하였으며, 짧은 리뷰에서는 성능이 좋지 못한 것을 확인하였다. 향후 연구에서는 부족한 평점에 대한 데이터를 추가 수집 또는 증강하여 확보하고, 여행 타입, 리뷰 제목 등 리뷰에 포함된 다른 데이터를 사용하는 컨텍스트 확장 등 다양한 방법을 통해 짧은 댓글에 대한 성능을 보완할 수 있는 연구를 진행하여 모델을 개선하고자 한다.

## 참고문헌(References)

### [국내 문헌]

- 강희재, 김창희, 성원우. (2019). 텍스트 분석을 통한 서비스 온라인 리뷰의 부정적 편향성에 관한 실증 연구. *한국생산관리학회 학술대회 논문집*, 2019(1), 6-6.
- 김백기, 장경배. (2023). 딥러닝과 감성사전을 결합한 하이브리드 감성분석 시스템 개발. *한국사물인터넷학회 논문지*, 9(4), 21-31.

- 김은미. (2021). 온라인 리뷰의 감성분석과 순환 신경망을 적용한 국내 인바운드 관광수요 예측 모형. *관광연구저널*, 35(3), 69-79. DOI : 10.21298/IJTHR.2021.3.35.3.69
- 노민정. (2021). 리뷰의 품질이 리뷰의 유용성에 미치는 영향: 리뷰 별점의 조절효과를 중심으로. *한국디지털콘텐츠학회 논문지*, 22(6), 999-1007. DOI : 10.9728/dcs.2021.22.6.999
- 노성화, 김정. (2022). 텍스트마이닝을 위한 한국어 범용 불용어 목록 연구. *한말연구*, 63(13), 1-15. DOI : 10.16876/klrc.2022.63.13.1
- 박상민, 나철원, 최민성, 이다희, 온병원. (2018). Bi-LSTM 기반의 한국어 감성사전 구축 방안. *지능정보연구*, 24(4), 219-240.
- 박소형, 조성준. (2022). 이모지 임베딩을 활용한 한글 소셜 미디어 텍스트 임베딩 방법론 연구. *한국경영과학회 학술대회논문집*, 3853-3853.
- 박영욱, 정규엽. (2021). DMR(Dirichlet Multinomial Regression) 토픽모델링을 이용한 온라인 리뷰 빅데이터 기반 고객감성 분석에 관한 연구: 국내 5성급 호텔의 외국인 이용객 리뷰를 중심으로. *호텔경영학연구*, 30(2), 1-20. DOI : 10.24992/KJHT.2021.2.30.02.1
- 박지수. (2021). 관광이벤트산업에 관한 서비스 품질이 고객만족도와 고객충성도에 미치는 영향에 관한 연구. *동북아관광연구*, 17(2), 23-44.
- 박지원, 이형룡. (2022). 관광지 만족도를 평가하는 관광지 선택속성에 관한 연구: 텍스트 마이닝과 퍼지셋 질적비교분석의 활용. *관광학연구*, 46(6), 105-126.
- 서윤암, 김희수, 윤상후. (2022). 날씨와 인기도를 고려한 경북 관광지 추천 알고리즘 개발에 관한 연구. *한국데이터정보과학회지*, 33(5), 845-856. DOI : 10.7465/jkdi.2022.33.5.845
- 심영석, 김홍범. (2018). 온라인 리뷰 빅데이터 기반의 Word2Vec 기법을 활용한 관광지 성과와 여행객 평점 간 구조적 관계 분석. *관광학연구*, 42(8), 165-189. DOI : 10.17086/JTS.2018.42.8.165.189
- 어균선, 이진창. (2019). 속성선택방법과 워드임베딩 및 BOW (Bag-of-Words)를 결합한 오피니언 마이닝 성과에 관한 연구. *디지털융복합연구*, 17(2), 163-170. DOI : 10.14400/JDC.2019.17.2.163
- 윤민지, 조창규, 서용원. (2021). 온라인 리뷰의 텍스트 분석을 통한 고객 평점 예측에 관한 연구. *한국생산관리학회지*, 32(1), 89-103.
- 이상훈, 최정, & 김종우. (2016). 영역별 맞춤형 감성사전 구축을 통한 영화리뷰 감성분석. *지능정보연구*, 22(2), 97-113. DOI : <https://doi.org/10.13088/jiis.2016.22.2.097>
- 이종수, 채동규. (2021). 제주도 여행지 추천 시스템: 딥러닝 기반의 여행객 평점 및 다양성 분석 기법. *한국정보과학회 학술발표논문집*, 1470-1471.
- 정효선, 이선이, 윤혜현. (2021). O2O 외식 배달 앱 서비스 품질이 고객만족도 및 행동의도에 미치는 영향: 식생활 라이프 스타일의 조절효과를 중심으로. *관광연구저널*, 35(2), 193-208. DOI : 10.21298/IJTHR.2021.2.35.2.193
- 조은경, 남길임. (2022). 한국어 텍스트 감성 분석. 서울: 커뮤니케이션북스.
- 조재현. (2023). 정보과다의 위험성과 대응방안에 관한 연구. *東亞法學*, -(98), 1-30. DOI : 10.31839/DALR.2023.02.98.1
- 조호수, 류민호. (2020). 사업자의 네트워크 중심성이 이용자 만족도에 미치는 영향 연구: 지능정보서비스를 중심으로. *정보사회와 미디어*, 21(3), 65-87. DOI : 10.52558/ISM.2020.12.21.3.65
- 채영길, 이종혁. (2022). 국제 뉴스에서 국가별 보도 빈도와 보도 태도에 대한 차이 탐색: 감성

- 사전과 인공지능경망을 연계한 ‘하이브리드 감성분석’을 통해. *정치커뮤니케이션 연구*, -(65), 5-64. DOI : 10.35731/kpca.2022..65.001
- 현지연, 유상이, 이상용. (2019). 평점과 리뷰 텍스트 감성분석을 결합한 추천시스템 향상 방안 연구. *지능정보연구*, 25(1), 219-239.
- 홍태호. (2022). 방한 관광객의 온라인 리뷰에 대한 빅데이터 분석 기반의 감성분석 및 평점 예측모형, *지식경영연구*, 23(1), 187-201. DOI : 10.15813/kmr.2022.23.1.010
- [국외 문헌]**
- Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment analysis: a comparative study on different approaches. *Procedia Computer Science*, 87, 44-49. DOI : 10.1016/j.procs.2016.05.124
- Eisenstein, J. (2019). Introduction to natural language processing. MIT press.
- Ganesan, K., Zhai, C., & Viegas, E. (2012). Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions. *In Proceedings of the 21st international conference on World Wide Web*, 869-878. DOI : 10.1145/2187836.2187954
- Jo, T. (2019). Text mining (Vol. 45). Cham, Switzerland: Springer International Publishing.
- Liu, C. Z., Sheng, Y. X., Wei, Z. Q., & Yang, Y. Q. (2018). Research of text classification based on improved TF-IDF algorithm. *In 2018 IEEE International Conference of Intelligent Robotic and Control Engineering(IRCE)*, 218-222. DOI: 10.1109/IRCE.2018.8492945
- Nasir, A. F. A., Nee, E. S., Choong, C. S., Ghani, A. S. A., Majeed, A. P. A., Adam, A., & Furqan, M. (2020). Text-based emotion prediction system using machine learning approach. *In IOP Conference Series: Materials Science and Engineering*, 769(1), 012022. DOI:10.1088/1757-899X/769/1/012022
- Shi, F., Chen, L., Han, J., & Childs, P. (2017). A data-driven text mining and semantic network analysis for design information retrieval. *Journal of Mechanical Design*, 139(11), 111402. DOI : 10.1115/1.4037649
- Yang, Z., & Zhang, M. (2020). TextOG: A recommendation model for rating prediction based on heterogeneous fusion of review data. *IEEE Access*, 8, 159566-159573. DOI: 10.1109/ACCESS.2020.3020942

Abstract

## Development of a Tourist Satisfaction Quantitative Index for Building a Rating Prediction Model: Focusing on Jeju Island Tourist Spot Reviews

Dong-kyu Yun\* · Ki-tae Park\* · Sang-hyun Choi\*\*

As the tourism industry recovers post the COVID-19 pandemic, an increasing number of tourists are utilizing various platforms to leave reviews. However, amidst the vast amount of data, finding useful information remains challenging, often leading to time and cost inefficiencies in selecting travel destinations. Despite ongoing research, there are limitations due to the absence of ratings or the presence of different rating formats across platforms. Moreover, inconsistencies between ratings and the content of reviews pose challenges in developing recommendation models.

To address these issues, this study utilized 7,104 reviews of tourist spots in Jeju Island to develop a specialized satisfaction index for Jeju tourist attractions and employed this index to construct a ‘Rating Prediction Model.’ To validate the model’s performance, we predicted the ratings of 700 experimental data points using both the developed model and an LSTM approach. The proposed model demonstrated superior performance with a weighted accuracy of 73.87%, which is approximately 4.67% higher than that of the LSTM. The results of this study are expected to resolve the discrepancies between ratings and review contents, standardize ratings in reviews without ratings or in various formats, and provide reliable rating indicators applicable across all areas of travel in different domains.

**Key Words** : Tourism, Jeju Island, Satisfaction, Review, Rating, Prediction, Quantitative Index

Received : November 28, 2023   Revised : December 8, 2023   Accepted : December 11, 2023

Corresponding Author : Sang-Hyun Choi

---

\* Dept. of Bigdata, Chungbuk National University  
\*\* Corresponding Author: Sang-Hyun Choi  
Dept. of MIS/Bigdata, Chungbuk National University  
1, Chungdae-ro, Seowon-gu, Cheongju-si, Chungcheongbuk-do, Korea  
Tel: +82-43-261-3742, E-mail: chois@cbnu.ac.kr

## 저자 소개



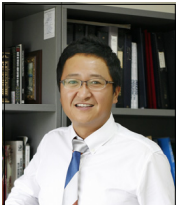
**윤동규**

현재 충북대학교 빅데이터 협동과정 박사과정에 재학 중이다. 주요 연구 분야는 Data Science, Deep Learning, Machine Learning 등이다.



**박기태**

현재 충북대학교 빅데이터 협동과정 석사과정에 재학 중이다. 주요 연구 분야는 Data Science, Image Analysis, Deep Learning 등이다.



**최상현**

현재 충북대학교 경영정보학과와 빅데이터 협동과정 교수로 재직 중이다. KAIST에서 경영정보공학을 전공으로 박사학위를 취득하였으며, 주요 연구 분야는 Business Process, Data Analysis, Machine Learning, Deep Learning 등이다.