

시켜줘 ! 명예호스트

AI 리뷰 스코어 도입을 통한 Airbnb 평점 체계 신뢰성 제고

김유찬, 김훈래, 박원우, 안병민, 채주형

Contents

서론 프로젝트 배경

문제점

프로젝트 목표

본론 모델 개요

라벨링

감성점수 추출

모델 성능 및 결과

결과 실제 구현 및 최종 목표 예시

검증 A/B TEST (효용성 검증)

결론 기대효과

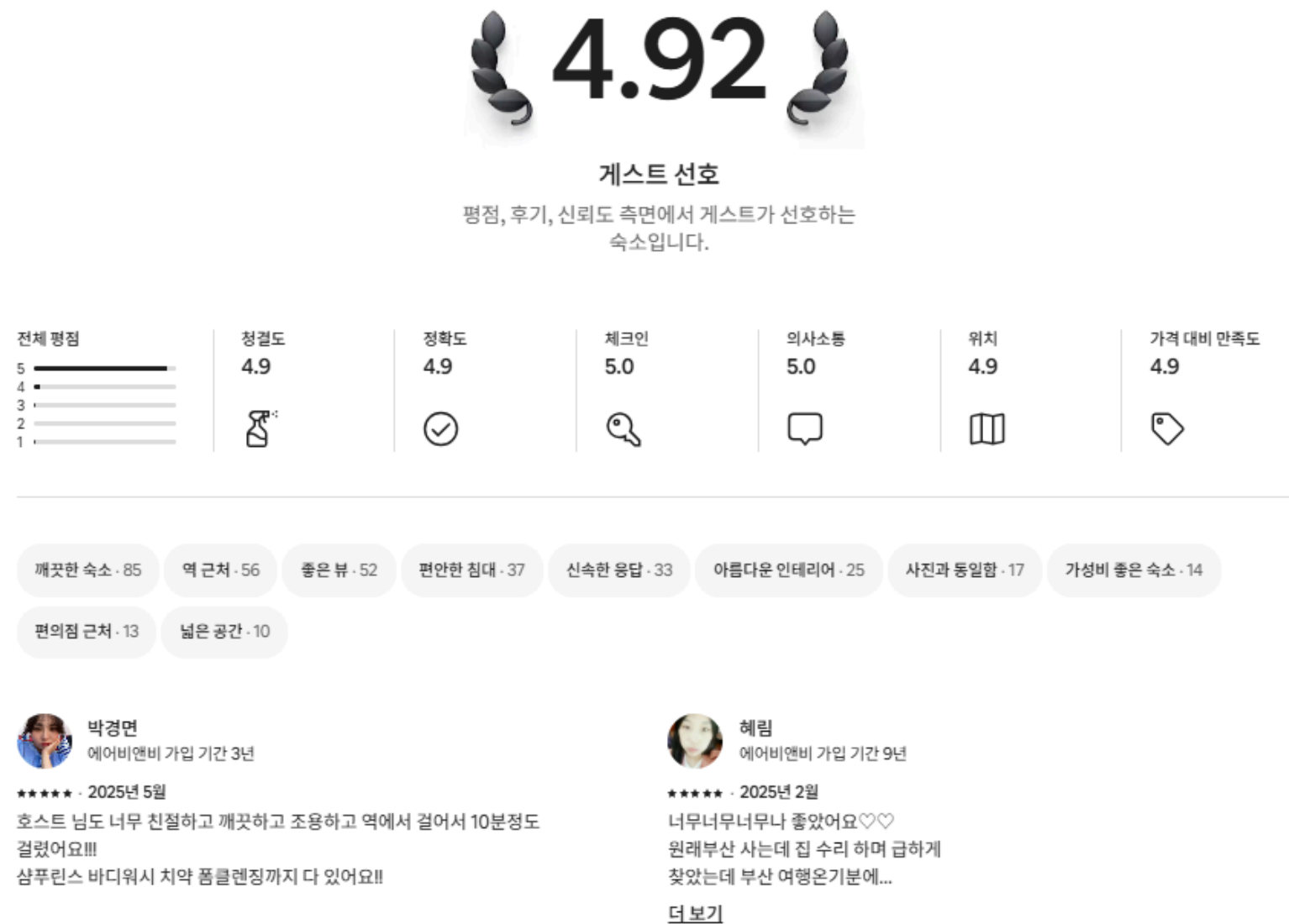
한계

문제점 인식

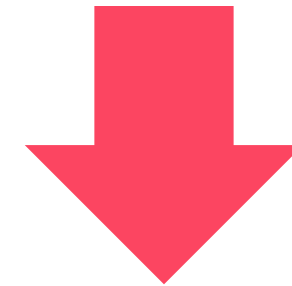


프로젝트 배경

관심변수 탐색 : 평점과 리뷰

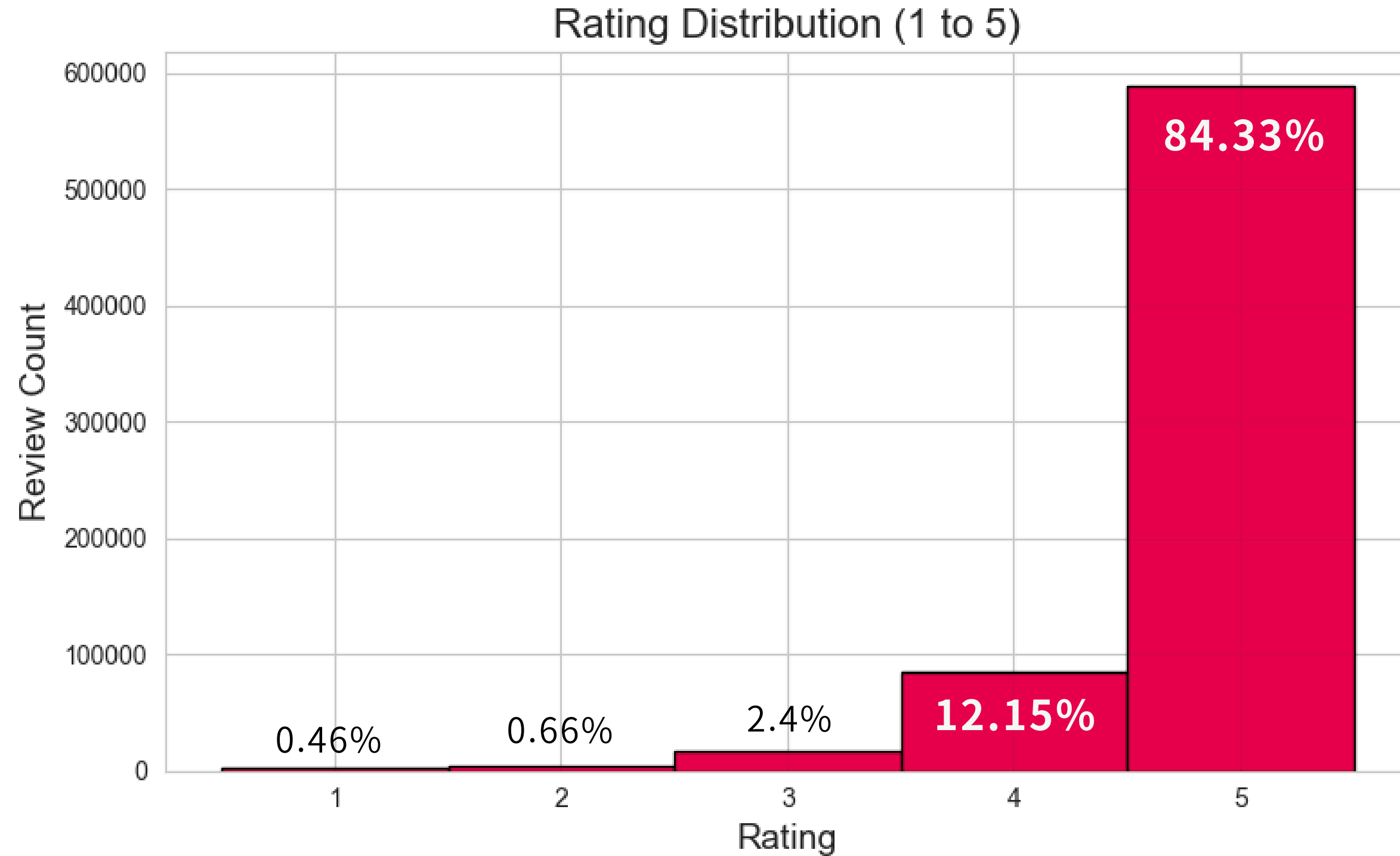


가격, 청결, 서비스 등
숙소 전반의 요소를 종합적으로 평가



실제 고객경험의 집합체

“평점”과 “리뷰”는 정확한 정보를 주는가?

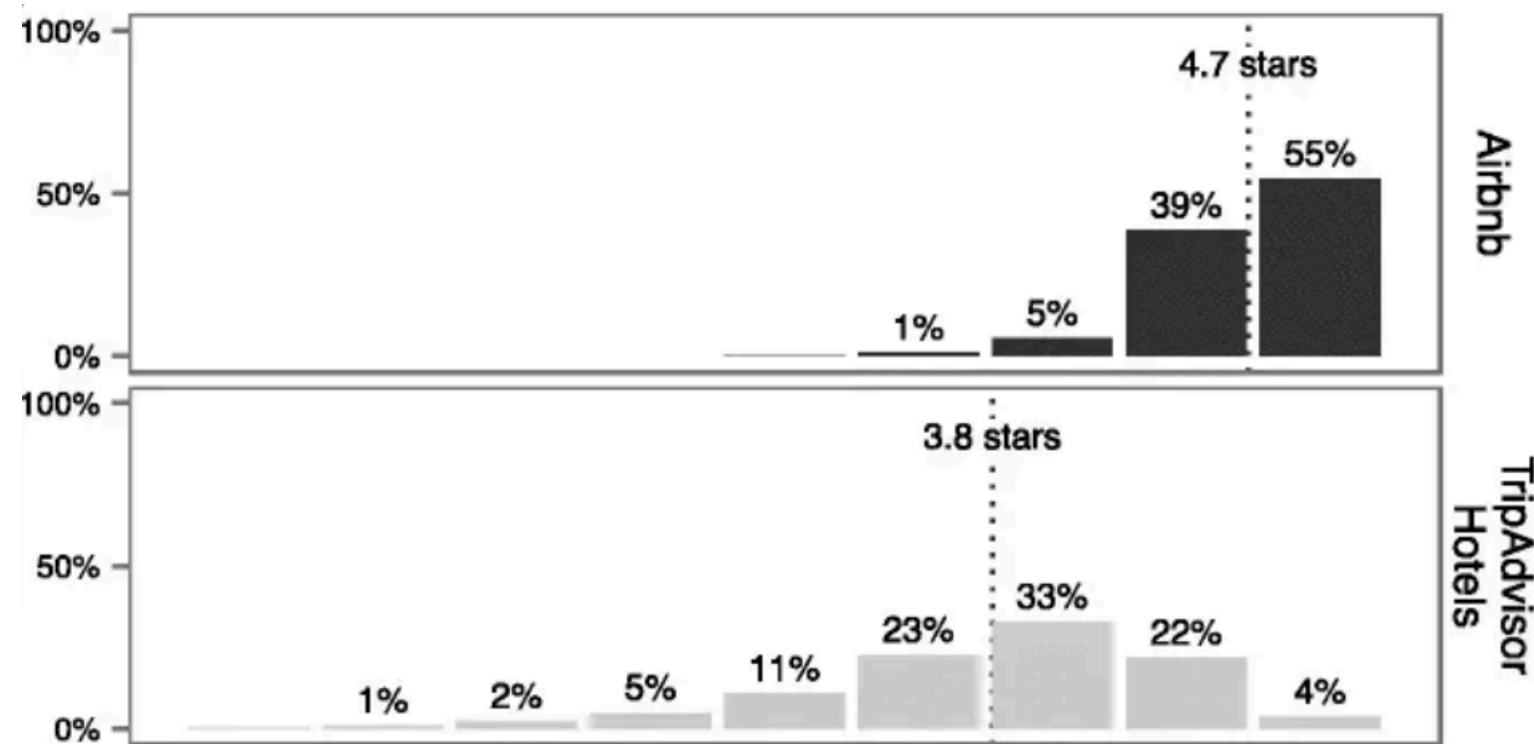


평점에 대한 EDA 결과 **긍정 편향** 확인

프로젝트 배경

에어비앤비 평점 선행연구

06



A first look at online reputation on Airbnb, where every stay is above average(Schuckert et al., 2020)

타 플랫폼에 비해 **약 1점 높음**

Star ratings on P2P accommodation platforms are highly positive. Such biases have led many users to utilise selective processing strategies to evaluate the textual content of online reviews.

From star rating to sentiment rating: using textual content of online reviews to develop more effective reputation systems for peer-to-peer accommodation platforms (Kim et al., 2022).

게스트와 호스트의 **상호 리뷰 시스템**
→ **평가가 관대**해지는 경향

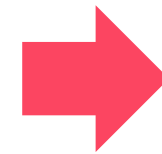
-> 에어비앤비의 평점은 **긍정편향**이 강하게 나타남

★5.0

숙소 위치는 정말 좋았어요. 다만 열쇠를 받기까지 시간이 좀 걸렸어요.

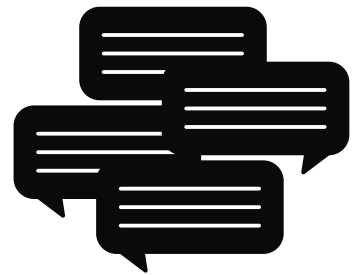
“ However, a micro-analytic comparison of **positive** reviews reveals that **less-than-positive** experiences are sometimes communicated using more **nuanced, subtle** cues. ”

If nearly all Airbnb reviews are positive, does that make them meaningless?
Current Issues in Tourism, 21(18), 2057–2075. Bridges, J., & Vásquez, C. (2016).

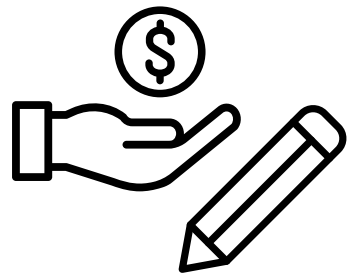


**평점보다 리뷰를
확인해야 함**

평점은 높게 주더라도 리뷰에서
불만족스러운 경험 드러냄



장문의 리뷰



리뷰이벤트

이번이 제 첫 에어비앤비
이용이었어요.

무의미한 리뷰

고객이 필요한 정보를 얻기 위해서는
많은 리뷰를 직접 읽어야 함

BUT !!

“ Those who read more than 11 reviews
constitute 15%. ”

11개 이상의 리뷰를 읽은 경우는 전체의 15%에 불과

→ 탐색비용 문제

평점의 긍정편향

- 긍정편향으로 인해 **평점이 숙소 간 비교 기준**으로서 **기능약화됨**
- 게스트의 기대와 실제 경험 간의 괴리 = **기대 불일치**
→ **만족도 저하** 및 브랜드 **충성도 감소**

평점의 신뢰성 저하

방대한 리뷰

- **평점**보다 **리뷰**에 솔직한 고객 경험이 더 잘 담겨있음
-> 게스트가 **직접 많은 리뷰**를 읽어야 함
- 리뷰가 많아 정보과부하
→ 탐색비용증가 및 의사결정 어려움

게스트의 탐색 비용 증가

문제정의

게스트가 예약결정 시 평점과 리뷰를 확인
평점의 낮은 신뢰성, 리뷰 확인 탐색비용 문제

제안 방향

“평점과 리뷰시스템 개편“
⇒ 평점의 신뢰성 회복 & 게스트의 탐색비용 감소

프로젝트 전략 수립



문제

평점보다 리뷰에서 **고객경험**
이 더 잘 담겨있음



해결방안

리뷰 감성점수를 구해
새로운 평점(AI Score) 제시



기대효과

긍정편향 완화 → **신뢰성 회복**

문제

방대한 리뷰에서 필요한
정보를 얻는 시간과 노력이 많이 듦



해결방안

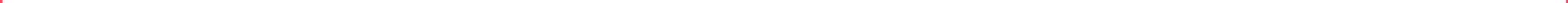
고객이 필요한 정보들
카테고리화 후 **감성분석**을 통해
필터링 기능 제공



기대효과

게스트의 리뷰
탐색비용 감소

본론



프로젝트의 NLP task

- Aspect-Based Sentiment Analysis로 **특정 항목에 대한 감성분석**을 뜻함
- 다양한 Task가 존재
- Task1: **감성 대상(Term)** 추출
- Task2: 감성 대상에 대한 **극성값(Polarity)** 추출
- Task3: **감성 대상**과 **극성값** 추출

- **term**: 감성 대상
 - ex) 방, 화장실, 지하철 역, 호스트, 가격...
- **polarity**: 극성값
 - 긍정, 부정, 중립

Ex)

input: **방**이 깨끗해요!

output: Task1 : {감성대상 : **방**}

Task2 : {극성값 : **긍정**}

Task3 : {감성대상 : **방**, 극성값 : **긍정**}

프로젝트 task 정의

term category polority

16

- 기존의 Task3 : 감성 대상과 극성값 추출
- 프로젝트 task는 감성 대상, 카테고리, 극성값 추출
 - Task3에 카테고리를 추가
- 카테고리의 의미 : 고객의 예약결정에 있어 중요한 항목들
- 제안방식은 세부 항목(카테고리)에 대해 극성값 추출
- 예시) 청결도 별 극성값, 정확도 별 극성값

프로젝트 카테고리

Airbnb 기존의 세부항목에 Amenity를 추가한 7개의 카테고리 선정

★ 5.0 · 후기 7개

전체 평점	청결도	정확도	체크인	의사소통	위치	가격 대비 만족도
5.0	5.0	5.0	4.6	4.9	4.9	5.0
						

1. 청결도 : 청소 상태, 정돈, 위생, 냄새 관련
2. 정확도 : 사진·설명과 실제 차이
3. 체크인 : 체크인/키 전달 절차 등
4. 커뮤니케이션 : 호스트의 응답, 소통, 설명 등
5. 위치 : 교통 접근성/주변환경- 이웃, 소음
6. Value : 가격 대비 만족도

[리뷰에 Amenity 관련정보 多]

7. Amenity : 편의 시설 (온수, 난방, 에어컨, 주차장 등)

프로젝트 task 예시

17

term category polarity

- 세부 항목에 대해 감성점수(AI score) 추출
→ 리뷰 문장에서 **term**-**category**-**polarity**를 뽑자
- **term**: 감성 대상
 - ex) 방, 화장실, 지하철 역, 호스트, 가격...
- **category**: 속성 분류
 - ex) 청결도, 정확도, 체크인, 의사소통, 위치, 가격대비 만족도 + 편의시설
- **polarity**: 극성
 - 긍정, 부정, 중립

“호스트가 친절해요!”



term
(감성 대상)

category
(속성 분류)

polarity
(극성)

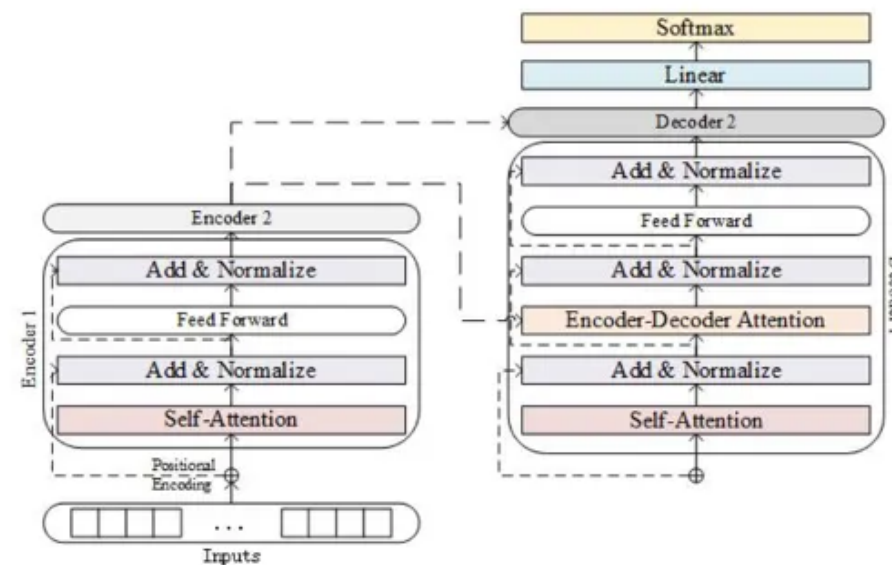


{**term**: 호스트, **category**: 의사소통, **polarity**: 긍정}

모델 선정 및 개요

InstructABSA

18



T5-Based Model for Abstractive Summarization: A Semi-Supervised Learning Approach with Consistency Loss Functions (Wang et al., Applied Sciences, 2023)

Architecture

Transformer 기반 T5모델을,
ABSA task에 맞도록
Instruction tuning한
text-to-text 모델

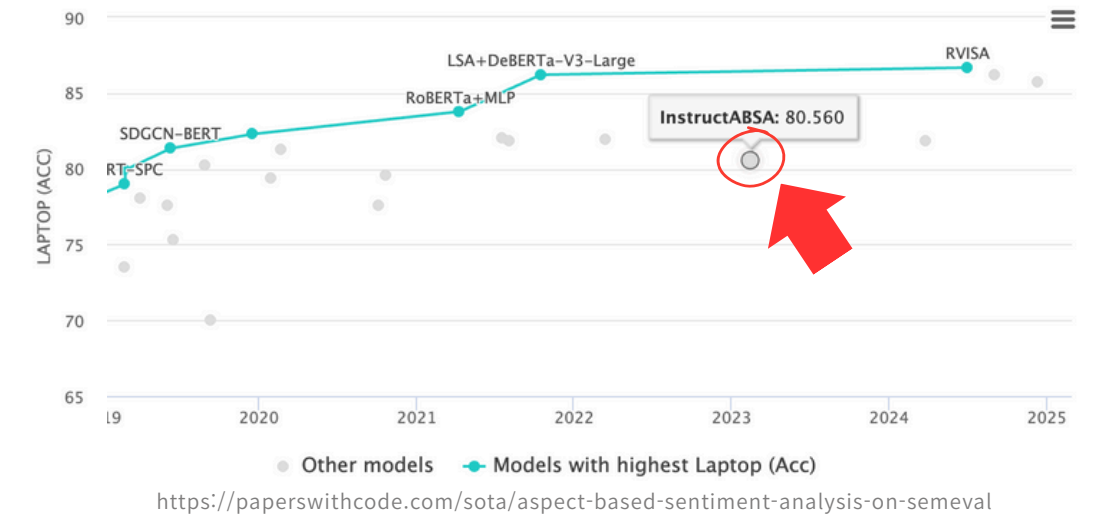
“Definition: The output will be the aspect terms,
their predefined categories, and sentiment polarity.

Positive example 1
input: The room was clean
output: room-Cleanliness-positive

...

Instruction 예시

Instruction으로 task 정의
다양한 ABSA task 수행 가능
→ 범용성/유연성이 특징
(instruction구성: def - example)



SemEval Dataset benchmark

SemEval benchmark
ABSA task에서
13등으로 준수한 성능

→ InstructABSA를 감성 분석 모델로 선정

모델 적용 과정 🔍

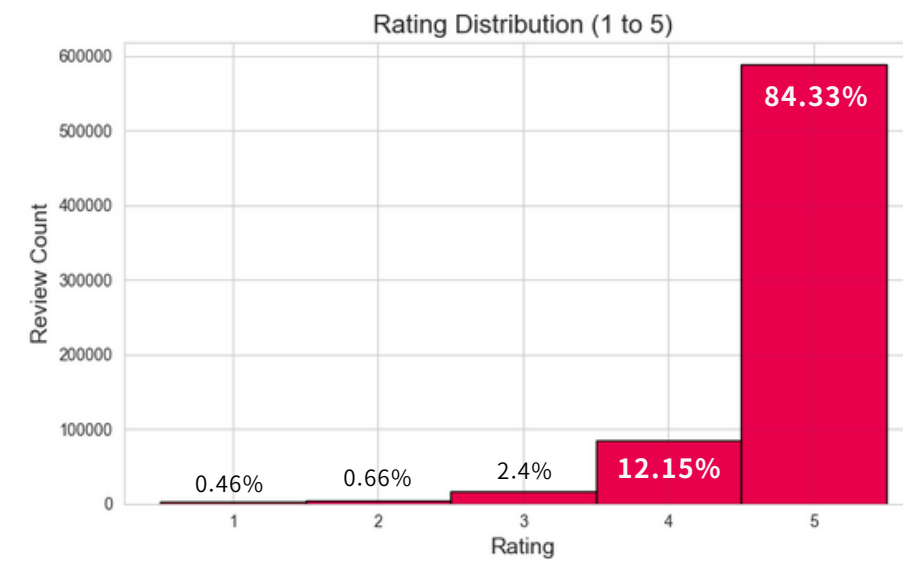
숙박 리뷰 도메인으로 학습되지 않았음
&
프로젝트 task로 학습된 모델 없음
(term-category-polarity 추출)



에어비앤비 데이터를 라벨링하여
학습데이터 생성 후 fine tuning

라벨링한 데이터의 양

- 총 1,000개의 리뷰 수작업 라벨링 진행



샘플링 방법

1. 4~5점의 96퍼여서 1~2점대 리뷰의 고객의 pain point를 반영위해 비율을 높여 포함



리뷰에서 항목별 극성값 을 추출할 때

극성값이 긍정/중립/부정일 확률을 계산

각 확률의 가중합(긍정:5, 중립:2.5, 부정:1) → 항목별 감성 점수(AI Score 산출)

감성점수 출력 예시

극성 확률 및 감성점수 예시

"The **location** is terrible but
host is responsive"



location(Location) → 2.62점
host(Communication) → 3.65점

"The **location** is terrible but
host is “**really**” responsive"



location(Location) → 2.63점 ✨
host(Communication) → 4.03점 ✨

```
input_sentence = "The location is terrible but host is responsive"
result = analyze_aspects(input_sentence)

pprint(result)

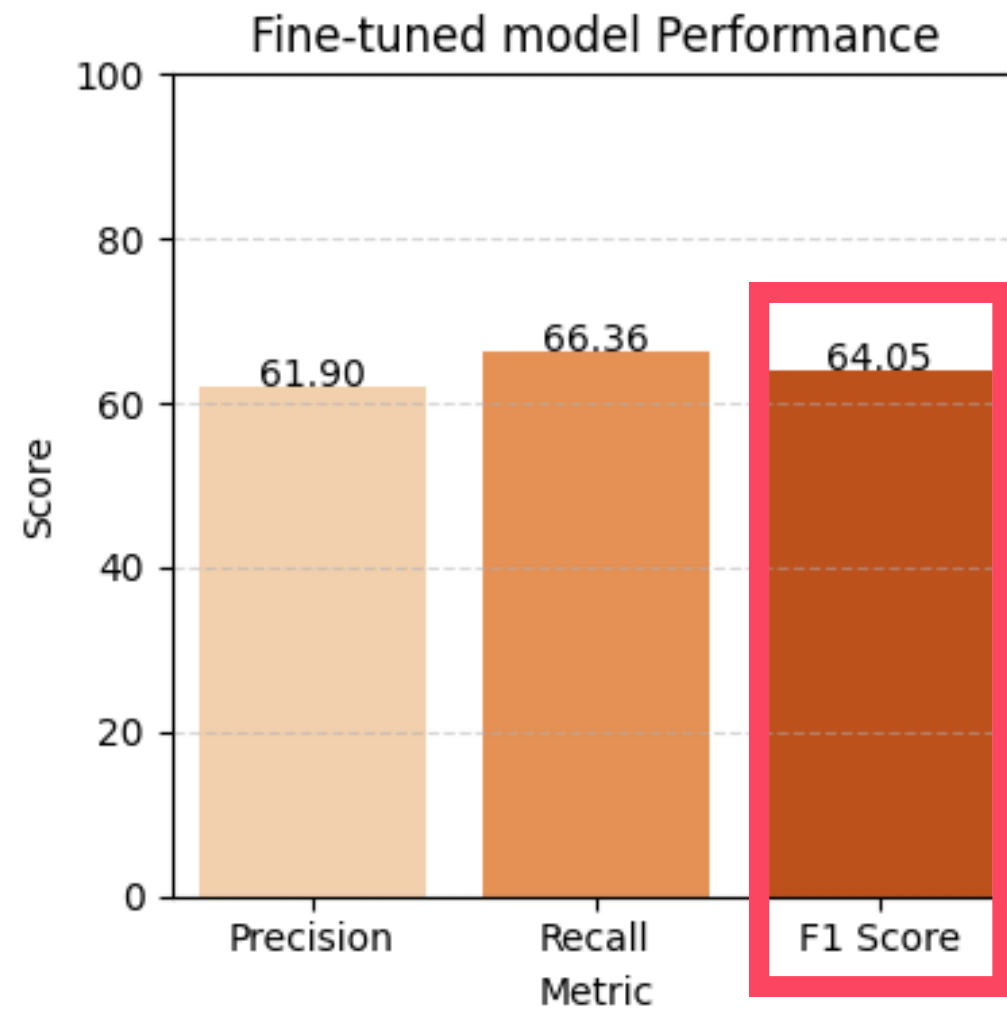
{'host.Communication': {'probs': {'negative': 0.3049,
                                  'neutral': 0.2767,
                                  'positive': 0.4184},
                        'score': 3.65},
 'location.Location': {'probs': {'negative': 0.5227,
                                  'neutral': 0.2569,
                                  'positive': 0.2204},
                        'score': 2.62}}
```

```
input_sentence = "The location is terrible but host is really responsive"
result = analyze_aspects(input_sentence)

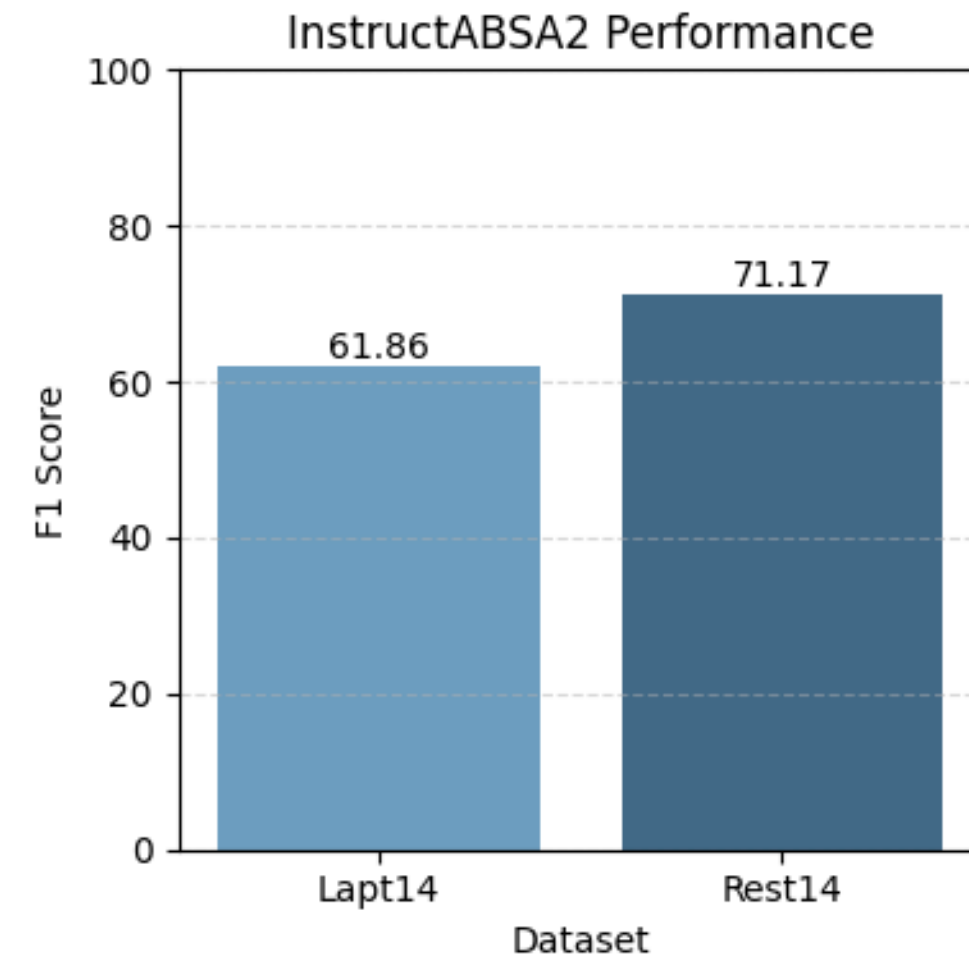
pprint(result)

{'host.Communication': {'probs': {'negative': 0.2351,
                                  'neutral': 0.2647,
                                  'positive': 0.5002},
                        'score': 4.03},
 'location.Location': {'probs': {'negative': 0.5177,
                                  'neutral': 0.2618,
                                  'positive': 0.2205},
                        'score': 2.63}}
```

Fine-tuning VS InstructABSA



Fine-tuning 모델
(term-category-polarity)추출

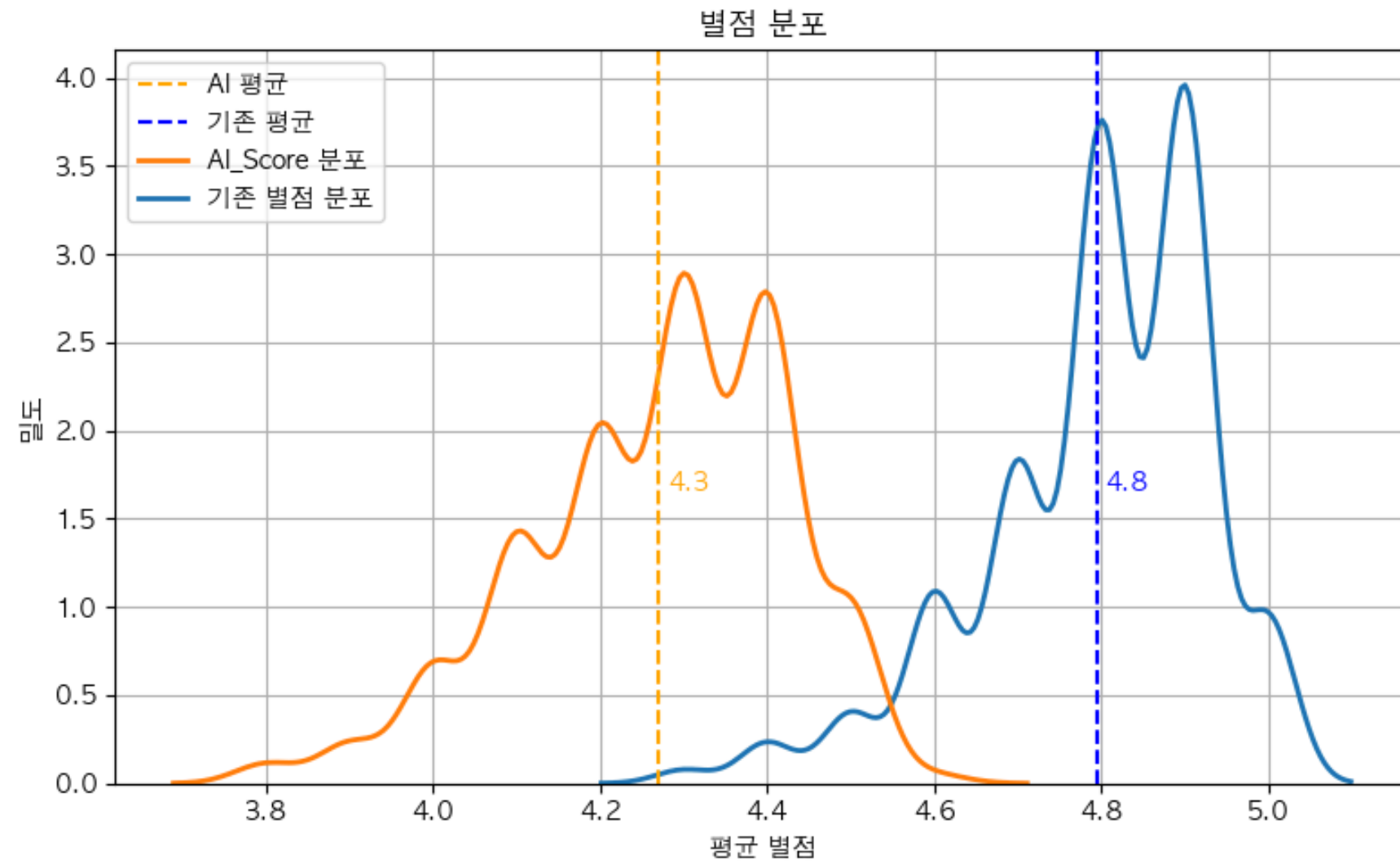


InstructABSA
(term-opinion-polarity)추출
(프로젝트처럼 3개의 output을 출력하는 가장 유사한Task)

AI score 결과 분석

기존 별점 VS AI score

24



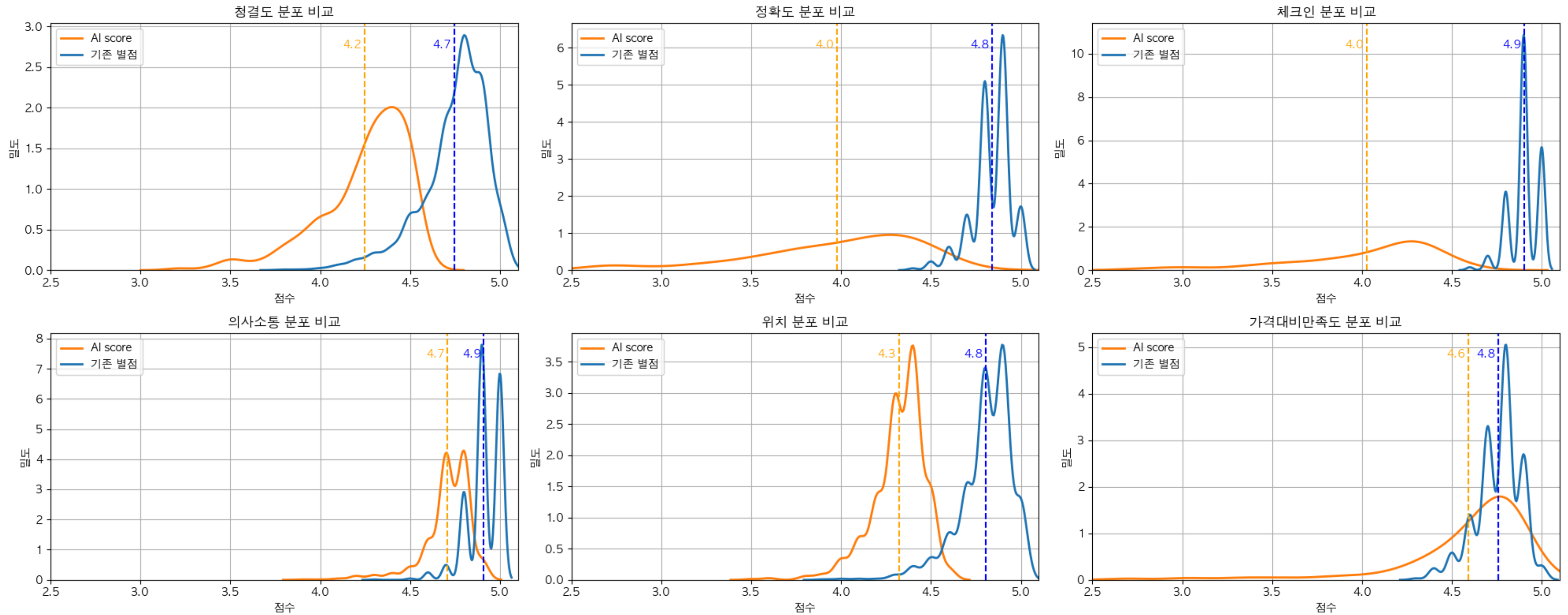
기존 별점 VS AI Score 밀도 그래프

- 기존 별점: 전체 평점, 세부 항목(청결, 위치...)을 사용자가 입력
 - AI SCORE : 리뷰 세부항목별 AI SCORE 평균
 - 기존 별점 평균: 4.8
 - AI score 평균: 4.3
- AI score 분포는 기존 별점 분포보다 **넓게 분산**된 것을 확인

AI score 결과 분석

세부항목 분포 비교

25

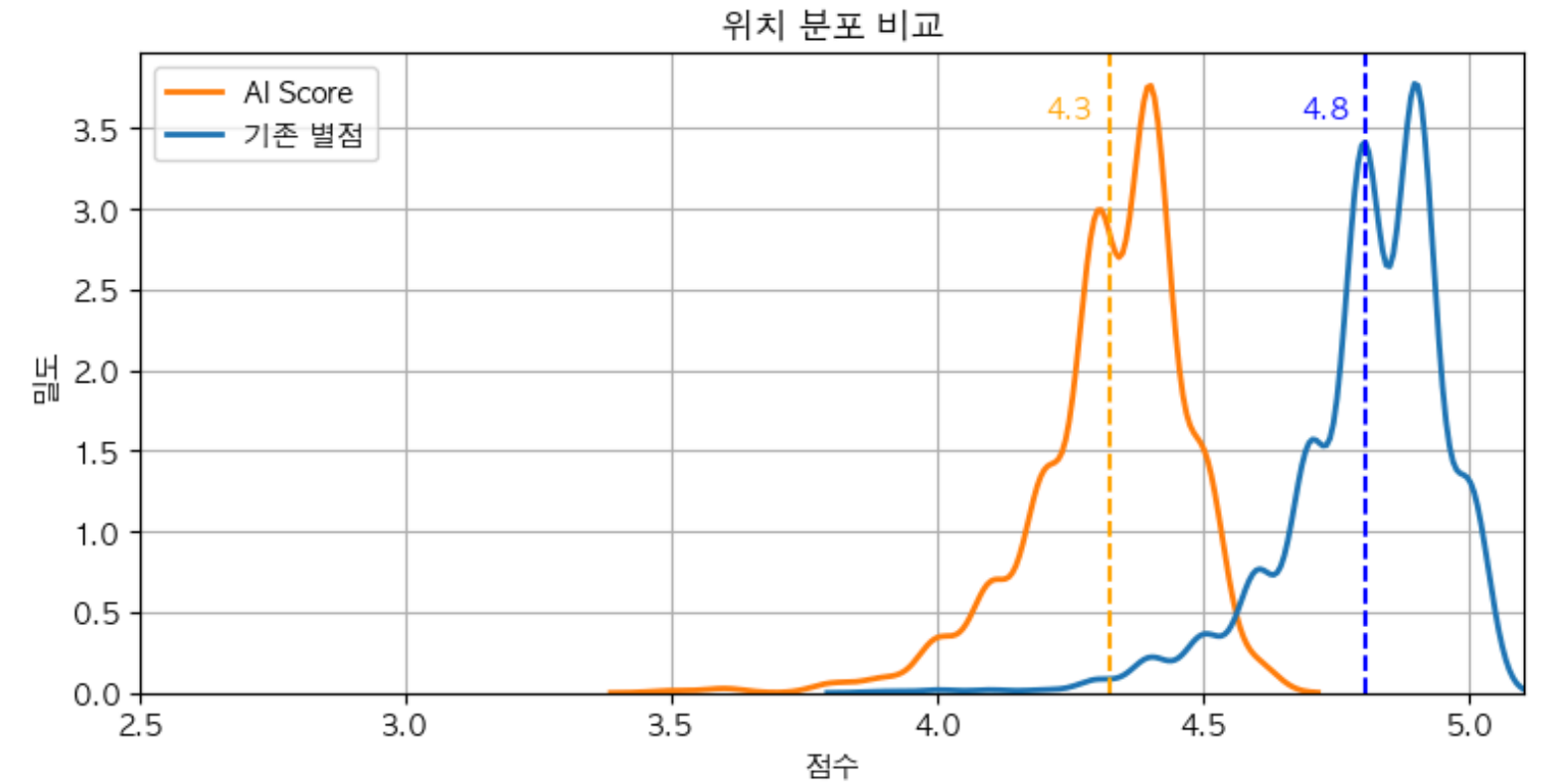
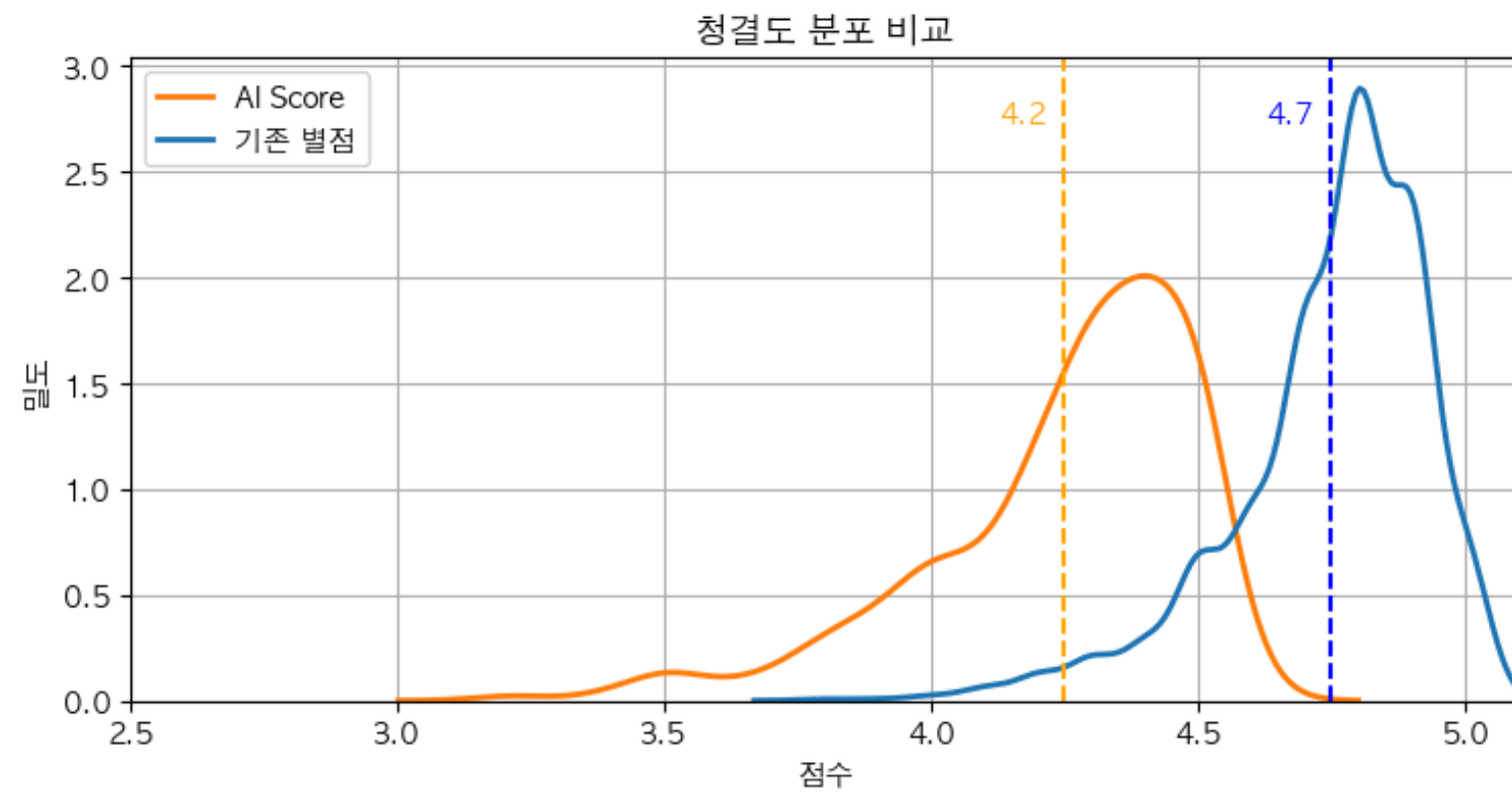


6가지 세부항목에 대한 AI score vs 기존 별점 분포 비교 그래프

AI score 결과 분석

세부항목 분포 비교

세부항목에 대해 AI score와 기존 별점의 분포가 유사한 경우
“청결도 & 위치”



특징:

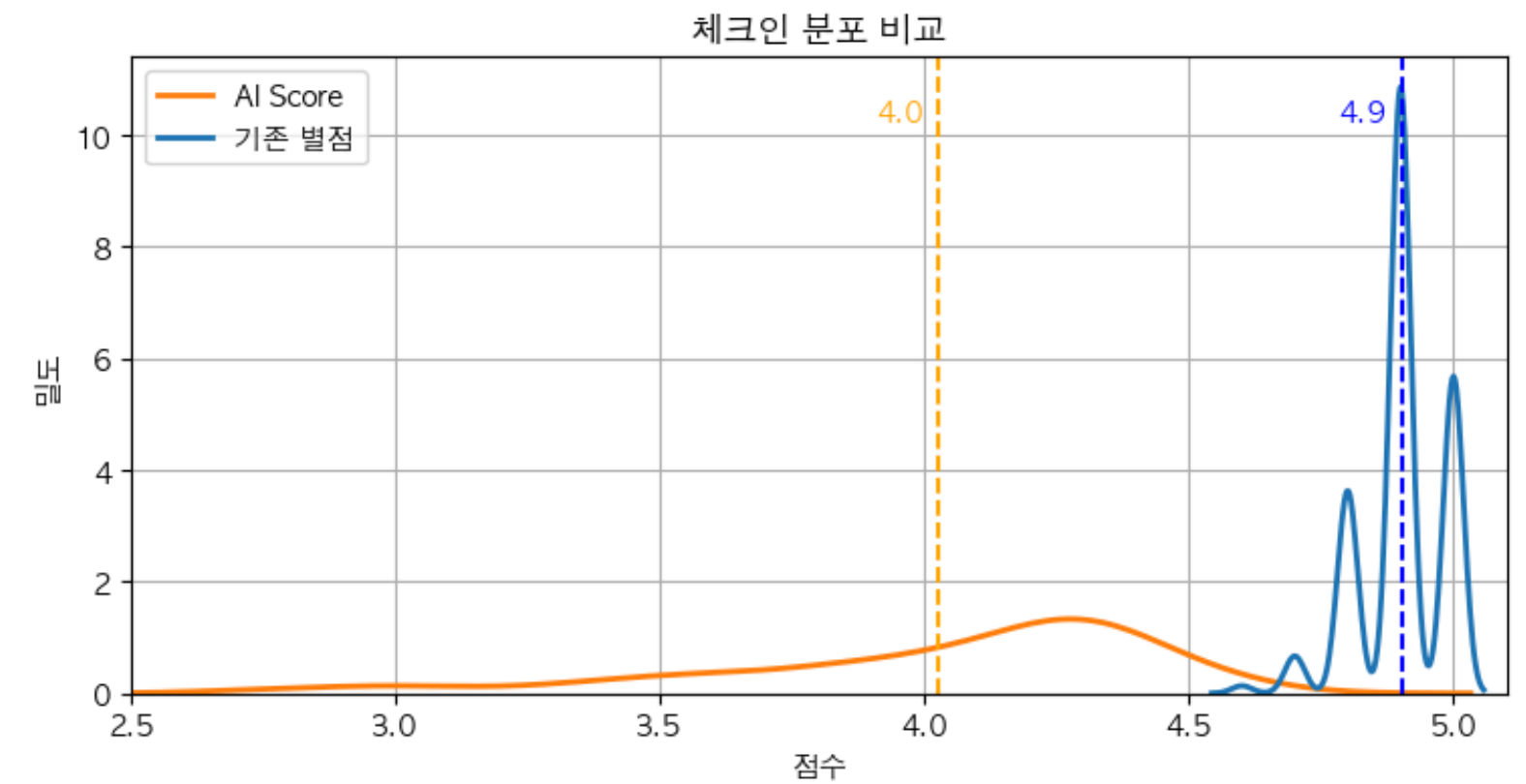
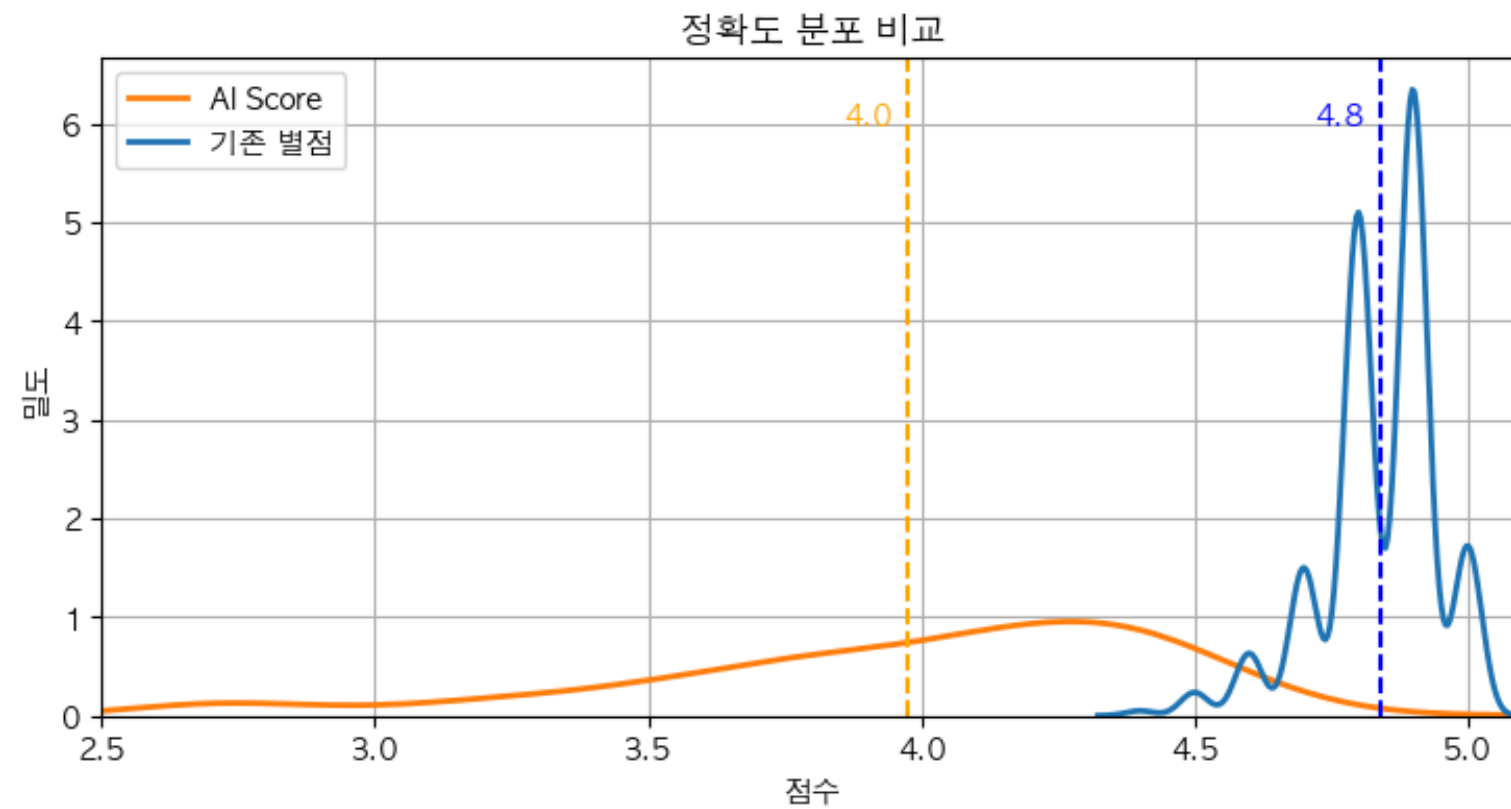
- 리뷰에서 해당 항목에 대한 자주 언급
- 리뷰에 긍정적인 반응이 많음(리뷰와 별점이 일치하는 경우)

AI score 결과 분석

27

세부항목 분포 비교

세부항목에 대해 AI score와 기존 별점의 분포가 **상이한** 경우
“정확도 & 체크인”



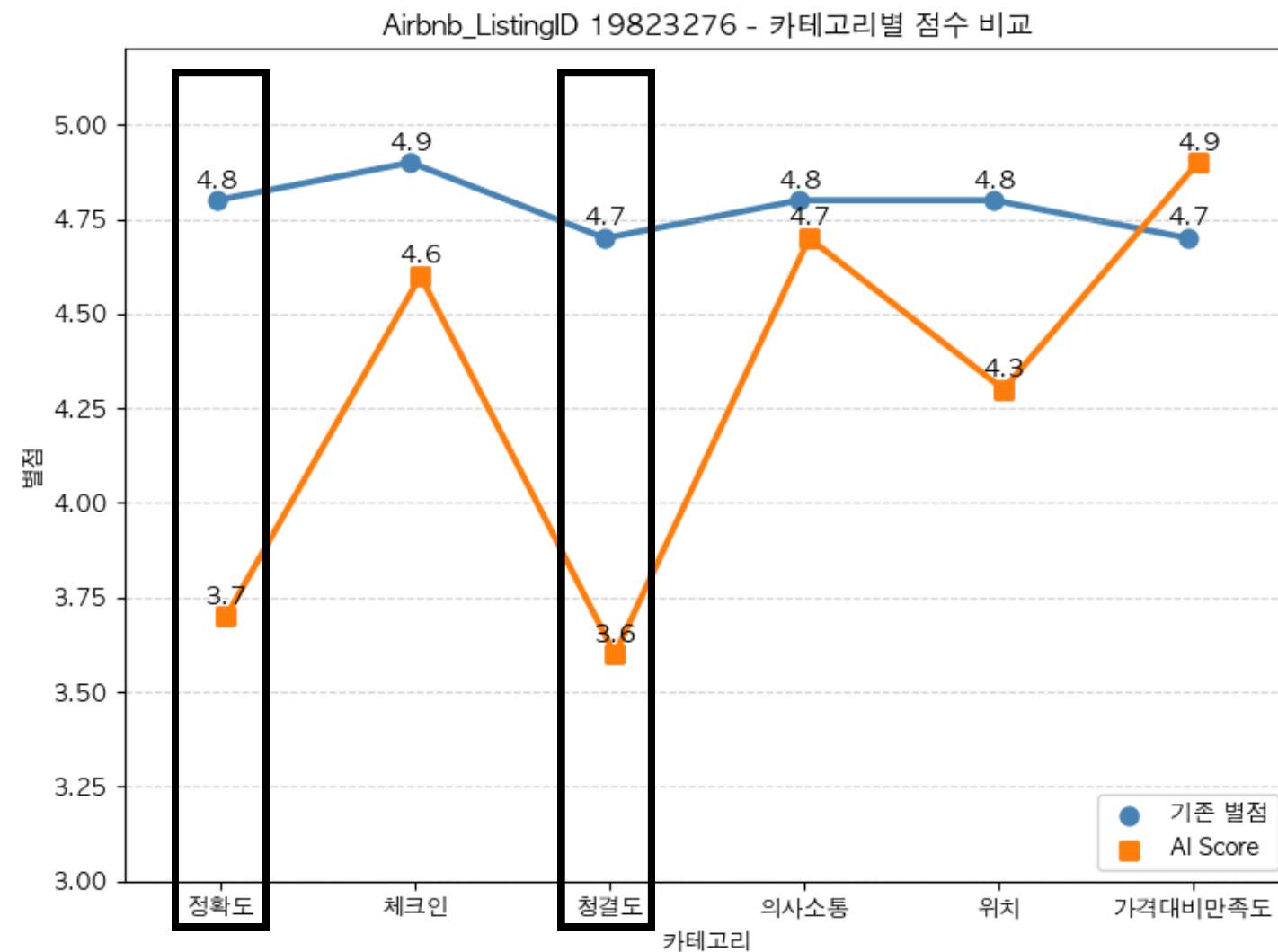
특징

- 리뷰에서 해당 항목에 대한 내용이 적음
- 분산이 크고 극단값이 자주 등장
- 실제경험에서는 차이가 존재하지만 평점에는 잘 반영되지 않음

AI score 결과 분석

숙소 기존 별점 vs AI score 비교

28



청결도 리뷰

1. There was a smell that was not so pleasant.
2. Wall was dirty, esp the one behind the desk has food stains.
3. Glasses were not cleaned ... floor had lots of crumbs...

...

정확도 리뷰

1. It didn't align with the photos
2. The photos match the listing but probably a few years back.
3. ~but a bit less nice compared to the pictures.

...

결과물 적용

추천 숙소

검색어를 입력하세요



숙소 호스팅하기

메뉴

당신의 완벽한 휴식을 찾아보세요

AI가 추천하는 최고의 숙소에서 특별한 경험을 만나보세요

전체

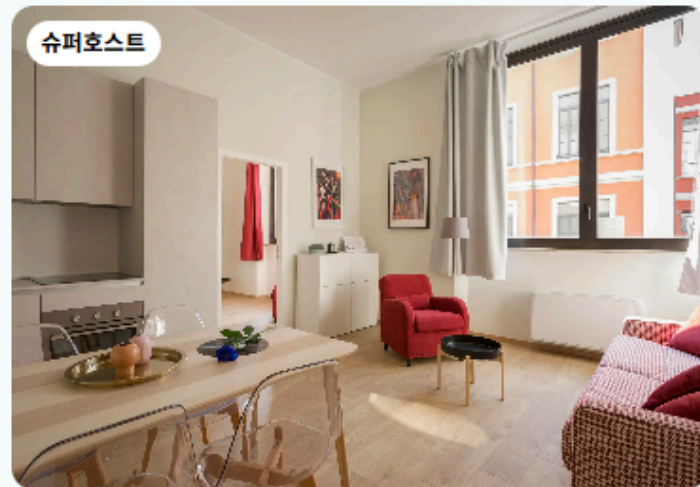
아파트

원룸

펜션

호텔

추천 숙소



서울 시청 근처 아파트

★ 4.7 AI 4.3

서울시 중구

₩120,000 / 박

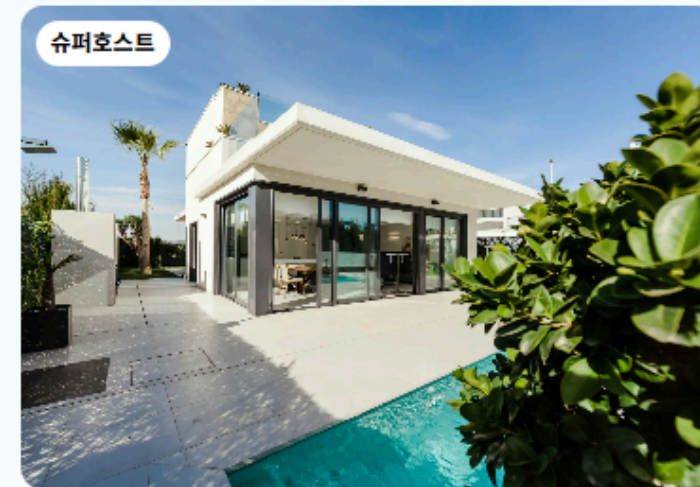


강남 스타일 원룸

★ 4.83 AI 4.15

서울시 강남구

₩150,000 / 박



홍대 근처 로프트

★ 4.6 AI 3.7

서울시 마포구

₩100,000 / 박

4.92

(4.5)

게스트 선호

평점, 후기, 신뢰도 측면에서 게스트가 선호하는 숙소입니다.

전체 평점

5

4

3

2

1

청결도

4.9 (4.7)

정확도

4.9 (4.8)

체크인

5.0 (4.9)

의사소통

5.0 (4.9)

위치

4.9 (3.3)

가격 대비 만족도

4.9 (4.4)

어메니티

4.9 (4.5)

* ()의 숫자는 AI가 리뷰를 분석하여 산출된 점수입니다.

깨끗한 숙소 · 85

역 근처 · 56

좋은 뷰 · 51

편안한 침대 · 36

신속한 응답 · 33

아름다운 인테리어 · 24

사진과 동일함 · 17

가성비 좋은 숙소 · 14

편의점 근처 · 13

넓은 공간 · 10

Karine

Granby, 캐나다

★★★★★ · 2일 전

모든 것이 완벽하고 깨끗했으며, 지침을 따라 쉽게 이동할 수 있었고 전망이 훌륭했습니다. 정말 마음에 들었습니다.

Julie-Éléonore

에어비엔비 가입 기간 10년

★★★★★ · 1주 전

훌륭한 숙소! 훌륭한 위치, 조용함, 아름다운 인테리어, 매우 편안한 침대, 훌륭한 전망! 정말 마음에 들었어요! 100% 추천합니다

4.92

(4.5)

게스트 선호

평점, 후기, 신뢰도 측면에서 게스트가 선호하는 숙소입니다.

전체 평점

5

4

3

2

1

청결도

4.9 (4.7)

정확도

4.9 (4.8)

체크인

5.0 (4.9)

의사소통

5.0 (4.9)

위치

4.9 (3.3)

가격 대비 만족도

4.9 (4.4)

어메니티

4.9 (4.5)

후기 233개

후기 운영 방식 알아보기

최신순

후기 검색

버스 (4.2)

자하철 (4.1)

주변 (3.8)

관광지 (3.1)

오르막길 (1.3)

* ()의 숫자는 AI가 리뷰를 분석하여 산출된 점수입니다.

Étienne

에어비엔비 가입 기간 4년

★★★★★ · 2주 전

전반적으로 만족스러운 숙소였습니다. 특히 위치가 좋았는데, 시내 중심가에 인접하여 이동이 편리했습니다. 다만, 숙소로 향하는 길이 약간의 오르막길에 위치하여, 짐이 많은 경우 다소 불편함이 있을 수 있습니다. 침구류는 쾌적했습니다. 이로 인해 편안한 휴식을 취할 수 있었습니다. 가성비 역시 훌륭하여, 다음 기회에 재방문 의사가 있습니다.

프랑스어에서 번역됨 원문 보기

Oliver

에어비엔비 가입 기간 8개월

★★★★★ · 3주 전

주요 관광 명소와 인접해 있어 도보로 이동하기에 용이했습니다. 그러나 숙소 입구까지의 오르막길은 다소 힘든 부분이었습니다. 정확도 측면에서는 게시된 사진과 실제 숙소가 거의 일치하여 실망하는 일은 없었습니다. 밤새도록 등을 지지는 온돌의 따스함 덕분에, 여행의 피로가 완벽히 사라져 편안한 휴식을 취할 수 있었습니다. 이 숙소를 매우 추천합니다!

영어에서 번역됨 원문 보기

李伟

에어비엔비 가입 기간 2년

★★★★★ · 2025년 2월

친구와 함께 방문한 숙소인데 만족스러웠습니다. 위치는 역과 인접하여 접근성이 양호했으나, 숙소 바로 앞까지는 짧은 오르막길이 있어 도보 이동이 필요했습니다. 청결도는 매우 우수하여 안심하고 이용할 수 있었어요.

중국어(간체)에서 번역됨 원문 보기

주요 기능

카테고리별 AI Score 표기, 리뷰 분석 기반 대표 키워드 추출, 키워드별 AI Score 추출, 키워드별 리뷰 필터링

프로젝트 효용성 검증



효용성 검증 (A/B Test)

A/B Test 진행 방식

33

제작한 웹사이트의 기술적 한계로

구글폼 설문조사를 통해 간접적으로 A/B Test 진행



STEP1. 설문지 제작

- 기존방식 : 현재 에어비앤비 리뷰 및 평점 화면 제공
- 제안방식 : 최종 구현 예시 화면 제공

STEP2. 설문지 배포

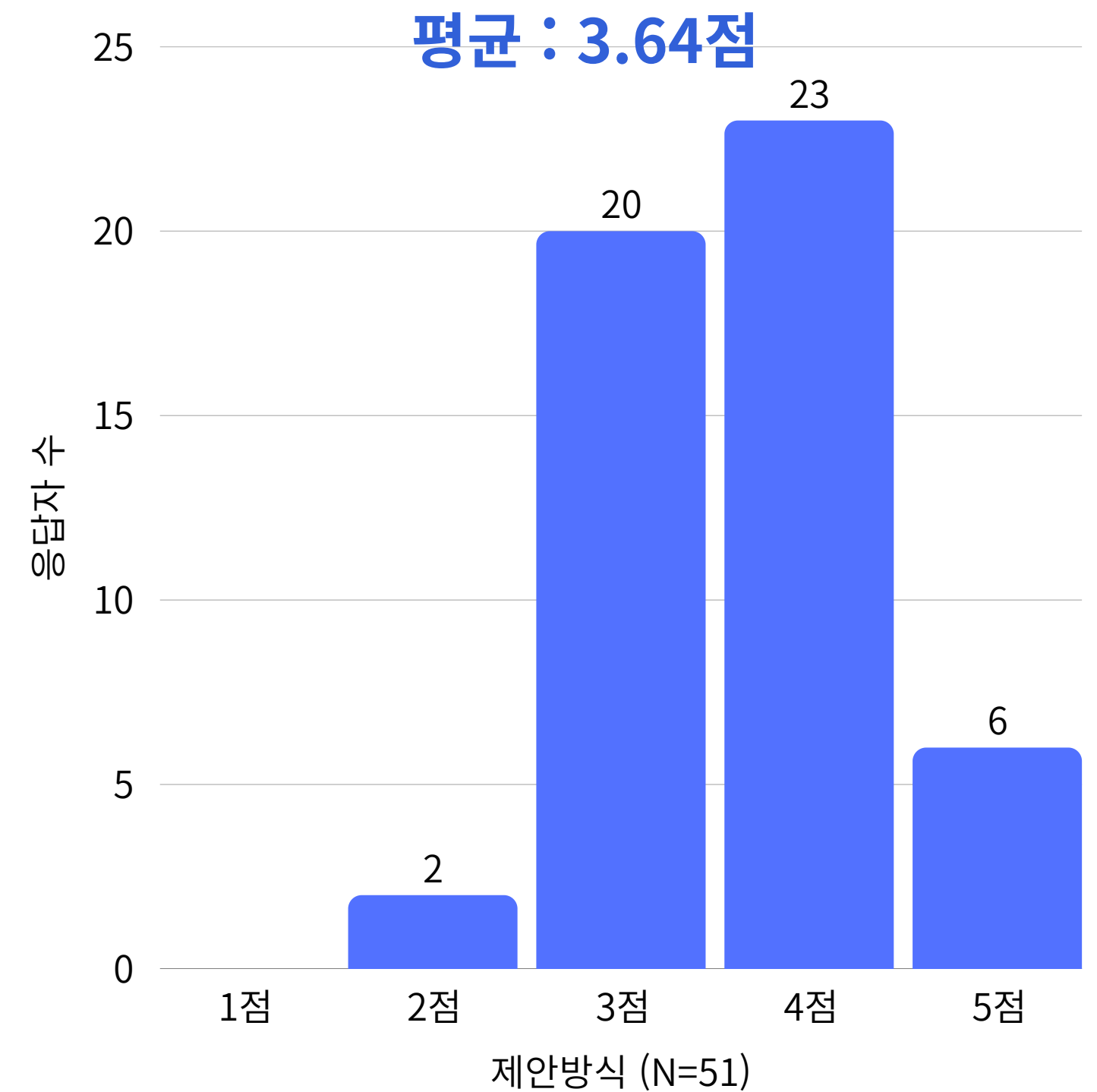
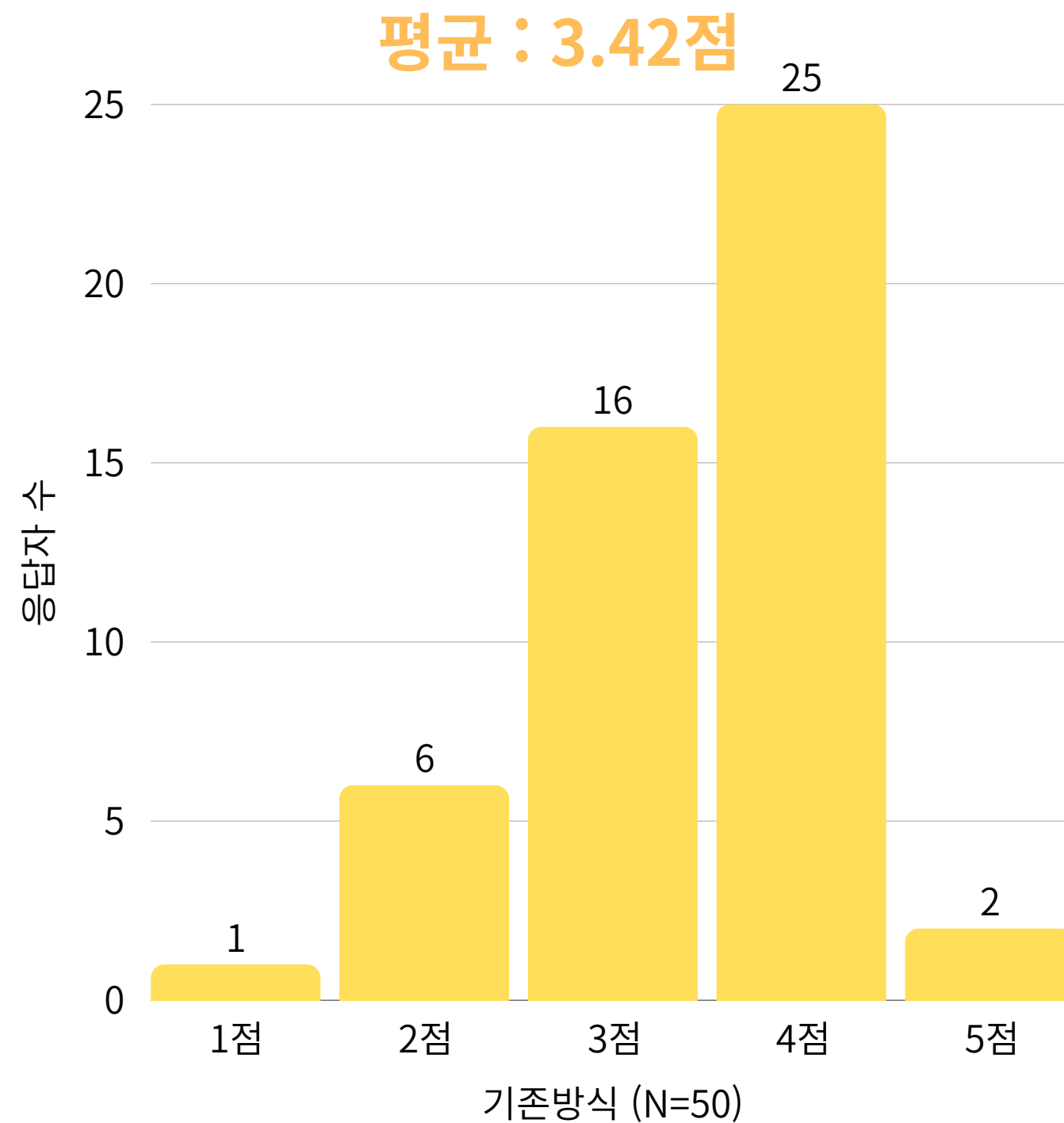
- 2개의 설문지를 분리하여 설문지 배포
- 기존방식 : 50명 응답
- 제안방식 : 51명 응답

효용성 검증 (A/B Test)

신뢰성 관련

34

해당 정보를 신뢰하시나요?



정보의 신뢰도를 다음과 같이 평가하신 이유는 무엇인가요?

기존방식

- 사람들의 귀찮음을 믿을 수 없음
- 너무 높은 점수는 광고같은게 들어갔을까 의심이되요
- 원하는 점수가 다 나와있음
- 점수가 과하게 높다
- 너무 최악이 아닌 이상 웬만한 사람들은 보통 4~5점정도 점수를 주기 때문에
- 구체적이지 않고 형식적인 답변이어서
- 믿을게 평점이랑 리뷰밖에 없다

제안방식

- ai의 추가적인 점수 제공이 있으니 생각 없이 준 별점이 포함되어있는 일반적인 리뷰보다 신뢰도가 높을 것 같습니다
- 자동으로 많이 나온 단어를 묶어줘서 문제점들을 파악하기가 쉬움
- 실제 이용자의 리뷰에 AI기능을 접목하여 편의성을 제공하며 신뢰도를 높임
- 사용자 점수와 AI 점수를 구분하여 비교 가능했다. 키워드를 필터링 해 볼 수 있어 편한 것 같다. 그러나 AI평점이 어떤 기준으로 도출된 것인지 명확하게 알 수 없다. 어떤 키워드로 인해 평점을 높이고 낮추는지 불투명한것같다.

효용성 검증 (A/B Test)

제안 방식의 효용성 기대 확인

36

AI Score의 도입이 평점의 신뢰성을
올리는데 영향을 주었나요? (N=51)



신뢰도 개선 질문에서
평균 3.9점으로 긍정적인 반응

AI Score가 숙소 예약 과정에서 시간을
단축시키는데 도움을 줄 것 같나요? (N=51)



숙소 탐색 비용 절감 관련 질문에서
평균 4.0점으로 긍정적인 반응

기존 문제점

- 평점의 **긍정편향**으로 인한 **신뢰성 저하**
- 다량의 **리뷰확인**으로 인한 **탐색 비용 증가**

제안방식의 효용성

- A/B Test 결과 AI Score 도입 시 **평점의 신뢰성 향상** 확인
- 리뷰 필터링 시스템-> **숙소 탐색 비용 감소**
긍정적 기대 확인

기대효과 및 한계 🔍



신뢰도 높은 정보 제공

- Alscore로 통해 객관적인 평점 비교
- 평점의 긍정편향 완화

숙소 탐색 비용 감소

- 리뷰를 일일이 읽지 않고
필요한 정보만 빠르게 파악
→ 의사 결정 피로도 감소

만족도 높은 숙소 선택

세부항목별 필터링 신뢰도 높은 평점으로 자신에게 중요한 기준에 맞는 숙소 선택 가능



플랫폼 신뢰도 및 경쟁력 강화

- 구조화된 리뷰-평점 시스템 제공
- 사용자 신뢰 향상 → 플랫폼 차별화



객관적인 피드백 수집

- 객관적인 리뷰 요약으로 개선점 파악

→ 빠른 피드백 반영,
서비스 품질 개선 및 마케팅 활용

[학습 데이터로 제한 모델 성능의 제약] [웹페이지 구현 기술적인 제약]



- 제한된 시간·자원으로 학습용 라벨링 데이터가 부족
- 일정 수준 이상의 성능은 확보했지만 데이터가 충분했다면 더 높은 정밀도와 일반화 가능성

- 웹 구현 미완성으로 A/B Test를 간접 방식으로만 진행
→ **A/B TEST 검증의 의미가 약화됨**
- 실제 모델 output 데이터와 연동하여 서비스 구현을 하지 못한점

P4.
https://www.airbnb.co.kr/rooms/682756408556877714?check_in=2025-08-01&check_out=2025-08-03&photo_id=1938686872&source_impression_id=p3_1749266293_P3OJJfuordThwVBK&previous_page_section_name=1000

P6.
Zervas, Georgios and Proserpio, Davide and Byers, John, A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average (December 28, 2020). Available at SSRN: A first look at online reputation on Airbnb, where every stay is above average(Schuckert et al., 2020

Zolbanin, H. M., & Wynn, D. (2022). From star rating to sentiment rating: using textual content of online reviews to develop more effective reputation systems for peer-to-peer accommodation platforms. Journal of Business Analytics, 6(2), 127–139. <https://doi.org/10.1080/2573234X.2022.2122880>

P7.
Bridges, J., & Vásquez, C. (2016). If nearly all Airbnb reviews are positive, does that make them meaningless? Current Issues in Tourism, 21(18), 2057–2075. <https://doi.org/10.1080/13683500.2016.1267113>

P8.
Rudolph, S. (2015). The Impact of Online Reviews on Customers’ Buying Decisions [Infographic]. Business 2 Community. <https://www.business2community.com/infographics/impact-online-reviews-customers-buying-decisions-infographic-01280945>

P13.
Wang, Mingye, Pan Xie, Yao Du, and Xiaohui Hu. 2023. "T5-Based Model for Abstractive Summarization: A Semi-Supervised Learning Approach with Consistency Loss Functions" Applied Sciences 13, no. 12: 7111. <https://doi.org/10.3390/app13127111>
<https://paperswithcode.com/sota/aspect-based-sentiment-analysis-on-semeval>

P18.
T5-Based Model for Abstractive Summarization: A Semi-Supervised Learning Approach with Consistency Loss Functions (Wang et al., Applied Sciences, 2023)

P32.
https://www.airbnb.co.kr/rooms/682756408556877714?check_in=2025-08-01&check_out=2025-08-03&photo_id=1938686872&source_impression_id=p3_1749266293_P3OJJfuordThwVBK&previous_page_section_name=1000

감사합니다.