

<커뮤니케이션 텍스트 분석론>

생성형 AI 플랫폼 (ChatGPT와 Google Gemini)의 모바일
애플리케이션 리뷰 텍스트마이닝을 통한 인공지능 서비스 평가
담론의 주제 구조 및 감성 분석

과	목	명:	AI기반미디어텍스트이해
담	당	교	수: 권상희 교수
제	출	일:	2025년 8월 24일
제	출	자:	소비자학과 송종찬

요약

본 연구는 ChatGPT와 Google Gemini의 사용자 리뷰(별점 1점, 4점)를 대상으로 텍스트 마이닝 기법(TF-IDF, 이그램, LDA, 네트워크 분석, 감정분석 등)을 적용하여 비교 분석하였다. 분석 결과, ChatGPT의 부정 리뷰는 업데이트, 유료 결제, 이미지 오류 등 구체적인 기능·정책 문제에 집중된 반면 Gemini의 부정 리뷰는 Google Assistant 연동 및 기기 제어 실패와 함께 강한 부정 감정이 두드러졌다. 긍정 리뷰에서는 ChatGPT가 질문·답변 기능과 모델 성능·브랜드 신뢰에 기반한 평가를 받은 반면, Gemini는 Google 생태계와의 연계성과 정서적 호감을 중심으로 긍정 경험이 형성되었다. 네트워크와 감정분석 결과 역시 이러한 경향을 뒷받침하였으며, 종합적으로 ChatGPT는 기능 안정성 중심, Gemini는 플랫폼 연계 경험 중심의 사용자 평가 구조를 보였다. 이는 인공지능 서비스 발전에서 성능 개선뿐 아니라 정책적 수용성, 생태계 통합, 사용자 감정 관리가 중요한 과제임을 시사한다.

목 차

요약문

제1장 연구소개	1-3
----------------	-----

제2장 연구방법 및 단계	34
---------------------	----

제3장 텍스트 분석결과

3.1 이그램 빈도 분석	5-10
---------------------	------

3.1.1 ChatGPT 리뷰의 이그램 분석	5-6
--------------------------------	-----

3.1.2 Google Gemini 리뷰의 이그램 분석	7-8
--------------------------------------	-----

3.1.3 양 서비스 리뷰 간 공통 및 차별 이그램 비교	9-11
---------------------------------------	------

3.2 동시 출현어 분석	12-16
---------------------	-------

3.2.1 ChatGPT 리뷰 동시 출현어 네트워크	12-13
------------------------------------	-------

3.2.2 Google Gemini 리뷰 동시 출현어 네트워크	13-14
--	-------

3.2.3 두 서비스 리뷰 네트워크 구조 비교	15-16
---------------------------------	-------

3.3 TF-IDF 기반 핵심 키워드 비교	71
-------------------------------	----

3.4 LDA 토픽 모델링 분석	18-24
-------------------------	-------

3.5.1 LDA 토픽별 Top 5 키워드 비교	18-19
----------------------------------	-------

3.5.2 LDA 주제 분포와 주제 비율 비교	20-21
---------------------------------	-------

3.5.3 LDAvis 시각화 결과 분석과 비교	22-24
----------------------------------	-------

제4장 감성분석

4.1 감성(긍정/부정) 분포 비교	25-26
---------------------------	-------

4.2 감성별 빈도수 단어	27-28
----------------------	-------

제5장 연구 결과정리 합의	29-31
----------------------	-------

표 차례

<표 1> ChatGPT의 이그램 빈도 분석 결과	5
<표 2> Google Gemini의 이그램 빈도 분석 결과	7
<표 3> ChatGPT 리뷰 동시 출현어 상위 15개 비교	8
<표 4> Google Gemini 리뷰 동시 출현어 상위 15개 비교	9
<표 5> ChatGPT vs Google Gemini 1점 리뷰 네트워크 구조 비교	1 1
<표 6> ChatGPT vs Google Gemini 4점 리뷰 네트워크 구조 비교	2 1
<표 7> ChatGPT vs Google Gemini 1점 리뷰 TF-IDF 상위 20개 비교 4·	1
<표 8> ChatGPT vs Google Gemini 4점 리뷰 TF-IDF 상위 20개 비교 5·	1
<표 9> ChatGPT vs Google Gemini 1점 리뷰 LDA 토픽별 Top 5 키워드 비교	17
<표 10> ChatGPT vs Google Gemini 4점 리뷰 LDA 토픽별 Top 5 키워드 비교	18

그림 차례

<그림 1> google_play_scraper 라이브러리를 활용한 데이터 수집	6
<그림 1> ChatGPT vs Google Gemini 1점 리뷰 공통 이그램 비교	0 1
<그림 2> ChatGPT vs Google Gemini 4점 리뷰 공통 이그램 비교	1 1
<그림 3> ChatGPT vs Google Gemini 동시출현 네트워크 시각화	3 1
<그림 4> ChatGPT 1점 리뷰 Wordcloud 결과	6 1
<그림 5> Google Gemini 1점 리뷰 Wordcloud 결과	6 1
<그림 6> ChatGPT 4점 리뷰 Wordcloud 결과	6 1
<그림 7> Google Gemini 4점 리뷰 Wordcloud 결과	6 1
<그림 8> ChatGPT vs Google Gemini 부정적 리뷰 LDA 토픽별 Top 5 키워드 비교	19
<그림 9> ChatGPT vs Google Gemini 긍정적 리뷰 LDA 토픽별 Top 5 키워드 비교	20
<그림 10> ChatGPT 1점 리뷰 LDAvis 결과	1 2
<그림 11> Google Gemini 1점 리뷰 LDAvis 결과	2 2
<그림 12> ChatGPT 4점 리뷰 LDAvis 결과	3 2
<그림 13> Google Gemini 4점 리뷰 LDAvis 결과	4 2
<그림 14> ChatGPT vs Google Gemini 1점 리뷰 감성 분포 비교	5 2
<그림 15> ChatGPT vs Google Gemini 4점 리뷰 감성 분포 비교	6 2
<그림 16> ChatGPT vs Google Gemini 1점 리뷰 부정 감성 단어 빈도 그래프	7 2
<그림 17> ChatGPT vs Google Gemini 4점 리뷰 긍정 감성 단어 빈도 그래프	8 2

제1장 연구소개

최근 인공지능(AI) 기반 대화형 언어모델(ChatGPT, Google Gemini 등)은 전 세계적으로 주목받고 있으며, 정보 검색, 글쓰기, 학습 지원, 프로그래밍 보조 등 다양한 분야에서 활용되고 있다. 이러한 대규모 언어모델(LLM: Large Language Model)은 빠른 속도로 발전하면서도 사용자 경험에 대한 평가는 상이하게 나타난다. ChatGPT는 2022년 OpenAI에 의해 공개된 이후 폭발적인 관심을 받았으며, Google Gemini는 Google DeepMind가 2023년에 발표한 차세대 모델로써 검색 및 멀티모달 활용에 강점을 가진다. 두 모델은 각기 다른 구조적 특성과 서비스 전략을 가지고 있으므로, 사용자가 남긴 리뷰를 분석하는 것은 이들 모델의 특성과 사용자 인식을 파악하는 데 중요한 단서를 제공한다.(OpenAI, 2022)¹⁾, (Google DeepMind, 2023)²⁾

특히 ChatGPT는 2022년 출시 이후 급속도로 대중화되었고, Gemini는 2023년 말 발표되며 구글의 검색·멀티모달 기능과 결합된 차별성을 강조하였다. 두 모델 모두 기술적 완성도를 내세우고 있으나, 사용자 리뷰는 제품 성능 이상의 생생한 반응을 반영한다. 따라서 리뷰 텍스트를 분석하여 주요 토픽과 감성을 비교하는 것은 두 모델의 경쟁 구도를 이해하고, 향후 인공지능 서비스의 발전 방향을 조망하는 데 중요한 자료가 된다.

본 연구에서는 특히 모바일 애플리케이션 리뷰를 중심으로 분석을 진행하였다. 이는 ChatGPT와 Google Gemini가 모두 스마트폰 환경에서 손쉽게 접근 가능한 형태로 서비스되고 있으며, 실제 사용자들이 가장 빈번하게 경험을 남기는 공간이 앱스토어 리뷰이기 때문이다. 모바일 앱 리뷰는 사용자의 일상적 맥락 속에서 나타난 구체적 경험을 담고 있어, 단순 기술적 성능 평가를 넘어 편의성, 접근성, 기능 활용성 등 실질적인 이용 행태를 반영한

1) OpenAI. (2022). Introducing ChatGPT. OpenAI Blog. <https://openai.com/blog/chatgpt>

2) Google DeepMind. (2023). Gemini: our largest and most capable AI model. Google DeepMind Blog. <https://deepmind.google/discover/blog/google-gemini-ai>

다는 점에서 분석 대상으로서 의미가 크다.

제2장 연구방법 및 단계

본 연구의 목적은 ChatGPT와 Google Gemini의 모바일 앱 리뷰를 비교·분석하여 두 모델에 대한 사용자 인식의 차이점을 실증적으로 규명하는 것이다. 이를 위해 다음과 같은 연구 방법을 설정하였다.

본 연구는 Google 플레이스토어에 게시된 ChatGPT와 Google Gemini 애플리케이션의 사용자 리뷰를 대상으로 하였다. 긍정적 평가와 부정적 평가를 균형 있게 비교하기 위해 별점 1점 리뷰와 4점 리뷰만을 추출하여 데이터셋을 구축하였다. 특히 별점 5점 리뷰는 과도한 칭찬이나 단순 긍정 표현이 무분별하게 포함될 가능성이 높아 분석의 신뢰성을 저하시킬 수 있다고 판단하여 제외하였다. 이러한 기준에 따라 선정된 데이터셋은 두 모델에 대한 사용자 경험의 극단적인 부정적 반응과 상대적으로 긍정적인 반응을 대조적으로 분석할 수 있도록 설계되었다.

텍스트 수집 과정은 다음과 같다.

1. Google에서 제공하는 google-play-scraper 라이브러리를 활용하여 Google 플레이스토어 리뷰 데이터를 수집할 준비를 한다. 라이브러리가 프로그램을 가동할 환경에 존재하지 않다면 “pip” 명령어를 통해 설치한다. 수집 대상 애플리케이션의 데이터를 수집하기 위해선 해당 애플리케이션의 고유 식별자(id)가 필요하다. 수집 대상 앱의 Play 스토어 URL에서 “id=” 뒤에 오는 문자열이 고유 식별자이다. 본 연구에서 사용한 ChatGPT 앱의 URL은 “https://play.google.com/store/apps/details?id=com.openai.chatgpt&hl=en”이며, 이 중 “com.openai.chatgpt”이 ChatGPT의 고유 식별자이다.

2. ChatGPT와 Google Gemini 애플리케이션을 대상으로 리뷰를 자동 수집한다. 리뷰 텍스트, 별점, 작성일, 공감 수 등 메타데이터를 선별하여 수집한다. 본 연구에서는 별점을 통해 필터링하고 리뷰 텍스트를 중심으로 분석했다.

3. 수집된 리뷰를 CSV 파일 형식으로 저장하여 후속 분석에 활용할 수 있도록 정리한다.

4. 별점 1점 리뷰와 별점 4점 리뷰만을 선택해 데이터셋을 구성하여 별도로 저장한다.

이 과정을 구현한 파이썬 코드는 아래와 같다.

```
] from google_play_scraper import Sort, reviews_all
# API에서 제공되는 데이터는 application의 모든 리뷰를 가져오는 것은 아닙니다.
# 또한 시도마다 서버에서 제공되는 데이터 개수가 달라질 수 있습니다.
# reviews_all안의 값들은 영어 데이터를 수집하기 위해 설정되어 있습니다.
# ChatGPT
raw_data_gpt = reviews_all(
    'com.openai.chatgpt', # app
    lang='en', # defaults to 'en'
    country='us', # defaults to 'us'
    sort=Sort.MOST_RELEVANT # defaults to Sort.MOST_RELEVANT
)

# Google Gemini
raw_data_gemini= reviews_all(
    'com.google.android.apps.bard', # app
    lang='en', # defaults to 'en'
    country='us', # defaults to 'us'
    sort=Sort.MOST_RELEVANT # defaults to Sort.MOST_RELEVANT
)

# 댓글, 별점, 공감수, 시간

# --- 데이터 로드 ---
gemini_df_4 = review_df_gemini[review_df_gemini['star']==4]
gemini_df_1 = review_df_gemini[review_df_gemini['star']==1]
gpt_df_4 = review_df_gpt[review_df_gpt['star']==4]
gpt_df_1 = review_df_gpt[review_df_gpt['star']==1]
```

그림1. google_play_scraper 라이브러리를 활용한 데이터 수집

수집된 리뷰 텍스트는 분석 가능한 데이터셋으로 활용하기 위해 전처리 과정을 거쳤다. 먼저 텍스트를 소문자로 통일하고, 특수문자와 불필요한 기호를 제거하였다. 이후 불용어 제거 과정을 통해 의미 없는 단어를 걸러내었으며, 토큰화를 통해 문장을 단어 단위로 분리하였다. 마지막으로 표제어 추출과 어간추출을 실시하여 단어 형태를 정규화함으로써, 동일한 의미를 지닌 단어들이 일관성 있게 분석에 반영되도록 하였다.

분석에 적용한 기법은 다음과 같다.

1. 텍스트 마이닝 기법(빈도 분석, TF-IDF, 감성 분석)을 적용하여 두 모델 리뷰의 주요 키워드와 감성적 특징을 도출한다. 단어 출현 빈도를 계산하여 ChatGPT와 Gemini 리뷰의 핵심 키워드를 추출하고, TF-IDF 기법을 통해 각 리뷰에서 상대적으로 중요한 단어를 도출하였다.
2. LDA 기반의 토픽 모델링을 통해 ChatGPT와 Gemini 리뷰에서 공통적으로 나타나는 주제와 차별적으로 나타나는 주제를 비교한다.
3. 워드클라우드 및 시각화 기법을 활용하여 결과를 직관적으로 제시하고, 분석 결과의 시사점을 도출한다. 주요 키워드, 감성적 경향, 주제별 분포 등을 탐색한다. 이를 통해 두 모델이 사용자에게 어떻게 인식되고 있는지를 실증적으로 비교·분석하고, 향후 인공지능 서비스의 발전 방향에 시사점을 제공하고자 한다.

마지막으로, 본 연구에서는 부정적 평가(1점 리뷰)를 긍정적 평가(4점 리뷰)보다 먼저 분석한다. 이는 사용자가 경험한 구체적 불편과 불만 요인을 선행적으로 파악함으로써, 서비스의 취약 지점을 확인하고 이를 개선할 수 있는 방향성을 도출하기 위함이다. 이후 긍정적 평가 분석을 통해 강점

요인을 비교·보완함으로써, 두 모델의 전반적인 사용자 경험을 균형 있게 이해하고자 하였다.

제3장 텍스트 분석 결과

3.1 이그렘 빈도 분석 결과

3.3.1 ChatGPT 리뷰의 이그렘 분석

<표1. ChatGPT의 이그렘 빈도 분석 결과>

순위	1점 리뷰(Bigram, Freq)	4점 리뷰(Bigram, Freq)
1	chat gpt (620)	good app (787)
2	please fix (354)	chat gpt (674)
3	waste time (351)	great app (499)
4	every time (306)	really good (327)
5	use app (246)	love app (290)
6	new update (231)	nice app (276)
7	free version (216)	app good (263)
8	wrong answer (213)	best ai (259)
9	worst app (200)	best app (242)
10	many time (195)	easy use (234)
11	give wrong (193)	app help (204)
12	bring back (193)	please fix (203)
13	much time (176)	app really (184)
14	create image (170)	really helpful (180)
15	generate image (168)	give star (172)
16	even though (167)	help lot (169)
17	feel like (163)	amazing app (168)
18	went wrong (160)	ai app (163)
19	something went (156)	answer question (161)
20	try later (152)	good ai (160)

GPT의 4점 리뷰에서 가장 많이 등장한 이그렘은 “good app”, “chat gpt”, “great app” 등 긍정적 평가와 직접적인 모델 이름 언급이 주를 이루었다. “love app”, “nice app”, “best ai”와 같은 표현은 사용자가 GPT의 성능과 편리성을 높게 평가하고 있음을 보여준다. 또한 “answer question”, “really helpful” 등의 빈출은 GPT가 질문에 대한 즉각적 답변 제공 능력에서 강점을 가지고 있다는 점을 시사한다.

반면 1점 리뷰에서는 “please fix”, “waste time”, “wrong answer”, “worst app” 등이 상위에 위치하였다. 이는 사용자들이 정확하지 않은 답변이나 업데이트 후 불편함, 시간 낭비와 같은 경험을 주로 부정적으로 인식했음을 의미한다. 특히 “new update”, “free version”, “try later”와 같은 단어는 서비스 품질 관리 및 이용 정책(무료/유료 버전, 안정성 등)에 대한 불만이 빈번히 제기되었음을 보여준다.

따라서 GPT 리뷰의 전반적인 양상은 긍정적으로는 지식 탐색의 효율성, 부정적으로는 정확성과 안정성 문제로 요약할 수 있다. 이는 사용자 경험이 기능적 만족도에 크게 좌우된다는 점을 드러낸다.

3.1.2 Google Gemini 리뷰의 이그램 분석

<표2. Google Gemini의 이그램 빈도 분석 결과>

순위	1점 리뷰(Bigram, Freq)	4점 리뷰(Bigram, Freq)
1	google assistant (3488)	good app (498)
2	chat gpt (559)	google assistant (488)
3	hey google (409)	nice app (232)
4	power button (372)	good ai (223)
5	turn phone (328)	chat gpt (192)
6	use gemini (325)	great app (182)
7	set alarm (305)	best ai (176)
8	get rid (298)	far good (173)
9	generate image (290)	easy use (170)
10	every time (290)	really good (169)
11	worst ai (288)	best app (158)
12	assistant gemini (273)	pretty good (143)
13	answer question (269)	need improvement (123)
14	google gemini (268)	google gemini (121)
15	google home (262)	ai app (119)
16	gemini app (262)	app good (105)
17	like google (252)	gemini good (103)
18	play music (242)	gemini app (97)
19	want phone (239)	gemini live (90)
20	set reminder (238)	hey google (90)

Gemini의 4점 리뷰에서는 “good app”, “google assistant”, “nice app”, “good ai” 등의 표현이 두드러졌다. 특히 “google assistant”와 함께 언급된 사례가 많다는 점은 Gemini가 **기존 구글 생태계(Assistant,

Home, Android 등)**와의 연계성 속에서 긍정적 평가를 받고 있음을 보여준다. 또한 “pretty good”, “need improvement”와 같은 표현은 긍정적 평가 속에서도 일부 기능 개선에 대한 기대가 동시에 존재함을 나타낸다.

한편 1점 리뷰에서는 “google assistant”, “hey google”, “power button”, “turn phone” 등 기기 제어와 연동성 문제가 핵심 불만으로 나타났다. “use gemini”, “generate image”, “set alarm”, “play music” 등의 빈출은 사용자가 Gemini를 단순 대화형 모델이 아니라 일상적인 스마트 기기 활용 도구로 인식하고 있음을 보여준다. 그러나 동시에 “worst ai”, “get rid”, “want phone”과 같은 표현은 실제 사용 경험에서 성능 미흡, 기능 오류, 불필요성을 강하게 느낀 경우가 많았음을 의미한다.

따라서 Gemini 리뷰의 전반적인 양상은 긍정적으로는 구글 서비스와의 통합성, 부정적으로는 스마트 기기 제어 과정에서의 오류와 불편으로 정리할 수 있다. 이는 Gemini가 단순 대화 모델을 넘어 구글 생태계 플랫폼으로 자리매김하는 과정에서 겪는 과도기적 문제를 반영한다.

3.1.3 양 서비스 리뷰 간 공통 및 차별 이그렘 비교

그림1의 그래프는 ChatGPT와 Google Gemini의 1점 리뷰에서 공통적으로 나타난 이그렘을 비교한 것이다. 주요 이그렘으로는 “chat gpt”, “every time”, “generate image”가 도출되었다. 두 모델 모두에서 “chat gpt”의 빈도수가 높게 나타났는데, 이는 사용자들이 불만족스러운 경험을 서술할 때도 모델명을 직접 언급하며 문제를 제기하는 경향이 강하다는 점을 보여준다. “every time”은 반복적으로 동일한 오류가 발생했음을 지적하는 표현으로, 두 모델 모두에서 유사한 수준으로 나타났다. 한편

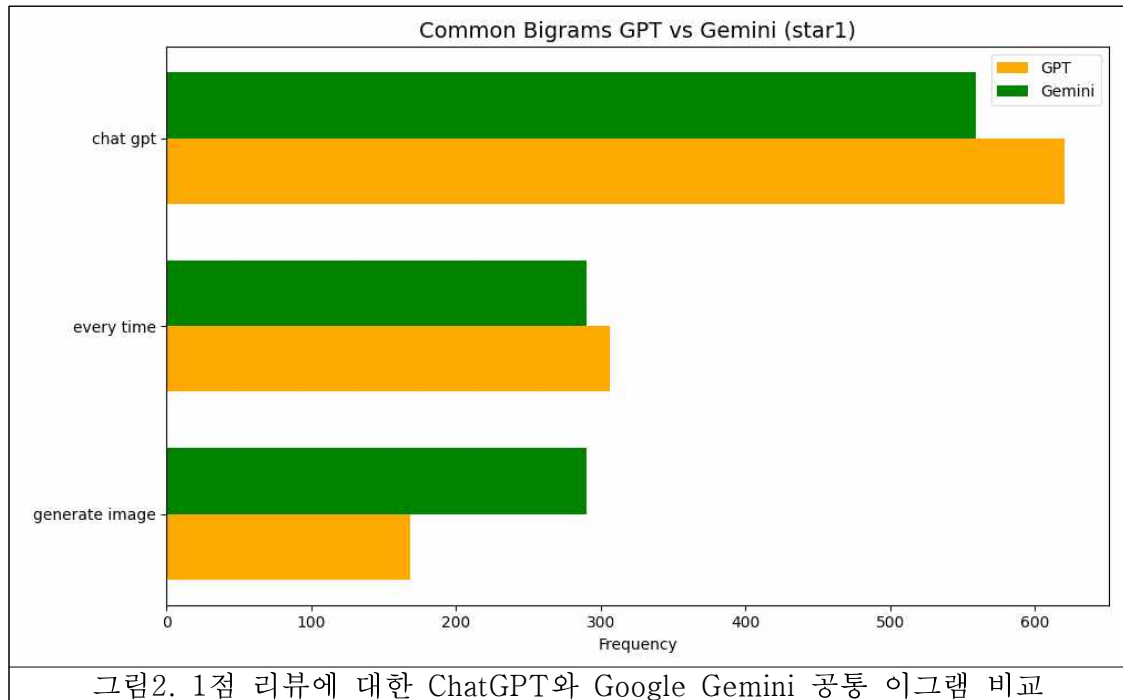


그림2. 1점 리뷰에 대한 ChatGPT와 Google Gemini 공통 이그램 비교

“generate image”는 Gemini에서 GPT보다 더 높은 빈도로 나타났는데, 이는 Gemini가 멀티모달 기능을 강조했음에도 불구하고 이미지 생성 기능이 사용자 기대에 미치지 못해 불만족으로 이어진 사례가 많았음을 시사한다. 따라서 1점 리뷰에서는 공통적으로 반복 오류와 기능적 불만이 제기되었으며, 특히 Gemini는 이미지 생성 기능에서 상대적으로 더 많은 불만을 유발한 것으로 해석된다.

그림2의 그래프는 4점 리뷰에서 공통적으로 확인된 이그램을 나타낸 것이다. 주요 표현으로는 “good app”, “chat gpt”, “great app”, “nice app”, “really good” 등이 포함되었다. 이 가운데 “good app”은 두 모델 모두에서 가장 높은 빈도를 보였으며, 특히 GPT에서 두드러지게 나타났다. 이는 GPT가 전반적으로 사용자에게 만족감을 주는 서비스로 인식되고 있음을 보여준다. 또한 “chat gpt” 역시 GPT에서 더 빈번하게 언급되었는데, 이는 사용자들이 긍정적 경험을 표현할 때 브랜드명 자체를 직접 언급하며

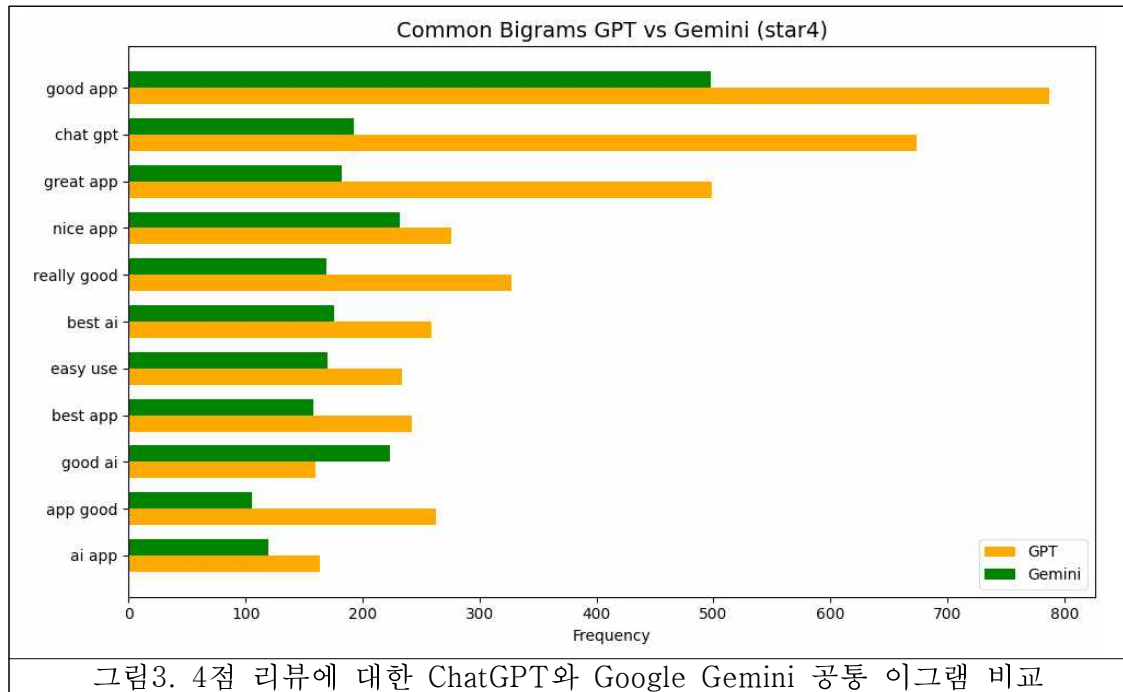


그림3. 4점 리뷰에 대한 ChatGPT와 Google Gemini 공통 이그램 비교
 신뢰와 호감을 드러낸 것으로 해석된다. 반면 Gemini는 긍정적인 표현이 다
 수 나타났으나 GPT에 비해 상대적으로 빈도가 낮아, 긍정적 반응의 강도에
 서는 차이를 보였다.

종합하면, 두 모델 모두에서 1점 리뷰는 반복 오류와 기능적 문제를 중심으로, 4점 리뷰는 전반적 만족과 긍정적 인식을 중심으로 형성되었다. 그러나 GPT는 브랜드 자체가 긍정적 평가와 밀접하게 연결되는 반면, Gemini는 이미지 생성과 같은 기능적 영역에서 불만이 집중되는 양상이 뚜렷하게 나타났다. 이는 사용자들이 GPT를 지식 탐색 중심의 대화형 도구로, Gemini를 멀티모달 및 기기 연동 중심의 플랫폼으로 인식하고 있음을 보여 준다.

3.2 동시 출현어 분석

3.2.1 ChatGPT 리뷰 동시 출현어 비교

<표3. ChatGPT 리뷰 동시 출현어 상위 15개 비교>

ChatGPT의 4점 리뷰			ChatGPT의 1점 리뷰		
word1	word2	count	word1	word2	count
app	good	1316	app	good	1316
app	great	727	app	great	727
gpt	chat	703	gpt	chat	703
really	app	461	really	app	461
time	take	426	time	take	426
app	best	419	app	best	419
app	love	416	app	love	416
app	nice	400	app	nice	400
really	good	394	really	good	394
app	helpful	393	app	helpful	393
like	app	357	like	app	357
give	answer	354	give	answer	354
app	amazing	336	app	amazing	336
app	useful	333	app	useful	333
question	answer	325	question	answer	325

표를 통해 확인할 수 있듯이, ChatGPT의 4점 리뷰와 1점 리뷰 모두에서 동일한 동시출현어 쌍이 등장하였다. “app-good”, “app-great”, “gpt-chat”이 가장 높은 빈도를 보이며, 이는 사용자가 긍정적 평가와 부정적 평가 모두에서 기본적으로 모델과 애플리케이션을 언급하는 방식이 유사함을 시사한다. 특히 “app-good”은 양쪽 리뷰 모두에서 가장 높은 빈도를 보였는데, 이는 1점 리뷰에서도 일정 부분 긍정 표현이 함께 나타날 수 있음을 보여준다. 다만 이러한 경우는 주로 ‘앱 자체는 좋지만 특정 기능이 문제다’와 같은 맥락에서 나타날 가능성이 높다.

또한 “give-answer”, “question-answer”와 같은 조합은 ChatGPT의 핵심 기능이 질문에 답변하는 역할임을 반영한다. 긍정적 리뷰에서는 이 기능이 유용하게 작동했음을 나타내지만, 부정적 리뷰에서도 같은 동시출현어가 높은 빈도를 보이는 것은, 답변 기능이 사용자 평가의 핵심 기준으로 작용했음을 보여준다.

따라서 ChatGPT의 리뷰에서 공통적으로 나타나는 동시출현어는 앱의 전반적인 효용성(앱 자체에 대한 긍정 평가)과 핵심 기능(질문-답변)의 성과가 사용자 만족 여부를 결정짓는 주요 요인임을 확인시켜 준다. 이는 긍정적 경험과 부정적 경험 모두가 같은 기능 영역을 중심으로 형성된다는 점에서 의미가 있다.

3.2.2 Google Gemini 리뷰 동시 출현어 비교

<표4. ChatGPT 리뷰 동시 출현어 상위 15개 비교>

Gemini의 4점 리뷰			Gemini의 1점 리뷰		
word1	word2	count	word1	word2	count
app	good	718	assistant	google	3760
google	assistant	522	gemini	google	859
ai	good	312	assistant	gemini	744
app	nice	281	phone	want	593
good	far	255	chat	gpt	574
app	great	248	app	google	551
gemini	google	235	answer	question	527
app	best	232	gemini	app	481
gemini	good	227	gemini	use	475
ai	best	210	assistant	use	462
chat	gpt	197	assistant	back	453
really	good	194	google	use	449
gemini	app	183	google	back	449
use	easy	182	phone	turn	438
give	answer	171	like	google	437

Gemini의 4점 리뷰에서는 “app-good”, “google-assistant”, “ai-good”, “app-nice” 등 전반적으로 긍정적인 평가가 많이 등장하였다. 특히 “google-assistant”가 상위에 위치한 것은 Gemini가 단독 애플리케이션이라기보다는 구글 어시스턴트 생태계와 결합된 서비스로 인식되고 있음을 보여준다. 또한 “use-easy”, “give-answer”와 같은 표현은 사용 편의성과 답변 기능에서 긍정적 경험을 나타낸 것으로 볼 수 있다.

반면 1점 리뷰에서는 “assistant-google”, “gemini-google”, “assistant-gemini”가 가장 높은 빈도를 기록하며, Gemini가 구글 서비스 및 기기 제어와 밀접하게 연동되어 있음에도 불구하고 불만족스러운 경험을 제공했다는 점을 시사한다. “phone-want”, “phone-turn”, “assistant-back”, “google-back”과 같은 조합은 기기 제어 기능에서 오류나 기대 미달이 발생했음을 의미한다. 또한 “answer-question”, “chat-gpt” 등의 조합도 등장하여, Gemini가 대화형 답변 기능에서 기대만큼의 성능을 제공하지 못했다는 사용자 인식이 반영되었다.

종합하면, Gemini의 4점 리뷰는 편리성과 긍정적 경험을 강조하는 반면, 1점 리뷰는 구글 생태계 연계성의 문제와 기기 제어 기능에서의 오류를 집중적으로 지적하고 있다. 이는 Gemini가 단순 언어모델을 넘어 구글 어시스턴트의 확장형으로 자리매김하고 있지만, 동시에 그 연동성이 사용자 불만족의 주요 원인이 되고 있음을 보여준다.

3.2.3 두 서비스 리뷰 네트워크 구조 비교

<표5. ChatGPT vs Google Gemini 1점 리뷰에 대한 네트워크 구조 비교 분석 결과>

구분	노드 수	엣지 수	밀도	평균 차수	평균 클러스터링 계수
ChatGPT	3356	300	0.000053	0.178784	0.002070
Gemini	3952	300	0.000038	0.151822	0.000705

1점 리뷰 네트워크 분석 결과, Gemini의 1점 리뷰 네트워크는 ChatGPT보다 노드 수가 더 많으나 밀도와 평균 차수, 그리고 평균 클러스터링 계수는 낮게 나타났다. 이는 Gemini의 1점 리뷰에서 더 다양한 단어들이 분포하지만, 단어 간의 연결 관계는 상대적으로 느슨하게 형성되어 있음을 의미한다. 반면 ChatGPT의 1점 리뷰 네트워크는 노드 수는 적지만 밀도와 결합도가 상대적으로 높아, 불만이 특정 키워드와 주제(예: 오류, 업데이트 문제, 답변 정확성)에 집중적으로 연결되어 있음을 보여준다. 따라서 Gemini

의 부정적 평가가 여러 기능적 문제로 분산되는 경향이 있는 반면, ChatGPT의 부정적 평가는 비교적 소수의 핵심 불만 요인에 집중된다고 해석할 수 있다.

<표6. ChatGPT vs Google Gemini 4점 리뷰에 대한 네트워크 구조 비교 분석 결과>

구분	노드 수	엣지 수	밀도	평균 차수	평균 클러스터링 계수
ChatGPT	2861	300	0.000073	0.209717	0.001457
Gemini	1953	300	0.000157	0.307220	0.001979

4점 리뷰 네트워크 분석 결과, ChatGPT는 Gemini보다 더 많은 노드를 포함하고 있으나 밀도와 평균 차수, 평균 클러스터링 계수는 낮게 나타났다. 이는 ChatGPT의 긍정적 리뷰가 다양한 키워드들로 구성되어 있지만 단어 간 연결성은 상대적으로 약하다는 점을 보여준다. 반면 Gemini의 경우 노드 수는 적지만 네트워크 밀도와 평균 차수, 클러스터링 계수가 모두 높게 나타나, 사용자의 긍정적 리뷰가 특정 주제와 키워드에 집중적으로 형성되어 있음을 알 수 있다. 즉, Gemini의 긍정적 평가가 소수의 핵심적 경험(예: 구글 생태계 연계성, 편리성)에 강하게 묶여 있음을 시사한다.

1점 리뷰 네트워크 분석에서는 ChatGPT가 불만족 요인이 몇몇 핵심 키워드에 집중된 반면, Gemini는 다양한 기능과 연계성 문제에 불만이 분산되어 나타났다. 반대로 4점 리뷰 분석에서는 ChatGPT가 폭넓은 키워드를 포함하나 연결성은 약한 반면, Gemini는 적은 수의 노드 내에서 밀도 높은 네트워크를 형성하여 긍정 평가가 특정 주제에 응집되어 있음을 보여주었다. 종합하면, ChatGPT는 부정적 경험이 특정 문제에 집중되고 긍정적 경험은 다양하게 분산되는 경향을 보인다. 반면 Gemini는 부정적 경험이 여러 문제로 흩어지지만, 긍정적 경험은 소수의 핵심 요인에 강하게 결집하는 특징을 가진다. 이러한 차이는 두 모델이 사용자에게 제공하는 경험의 성격이 다르다는 점을 드러내며, ChatGPT는 기능적 성능에 대한 평가 중심, Gemini는 플랫폼 연계성과 기기 활용 경험에 대한 평가 중심으로 구분된다고 할 수

있다.

3.3 TF-IDF 기반 핵심 키워드 비교

<표7. ChatGPT vs Google Gemini 1점 리뷰에 대한 TF-IDF
상위 20개 단어 비교 분석 결과>

ChatGPT	TF-IDF	Google Gemini	TF-IDF
app	0.034538	good	0.029976
time	0.023737	google	0.024667
image	0.020912	gemini	0.021891
even	0.016567	assistant	0.020889
working	0.015547	phone	0.019578
chatgpt	0.015060	app	0.018685
gpt	0.014159	want	0.018462
chat	0.014007	bad	0.017493
work	0.013909	like	0.014442
give	0.013163	ai	0.013858
use	0.013090	google assistant	0.013390
slow	0.012765	work	0.013355
photo	0.012600	use	0.010770
like	0.012457	nice	0.010559
bad	0.012301	useless	0.010380
update	0.011966	even	0.010167
wrong	0.011792	hate	0.010068
answer	0.011728	answer	0.009991
error	0.011635	language	0.009328
get	0.011519	suck	0.009271

<표8. ChatGPT vs Google Gemini 4점 리뷰에 대한 TF-IDF
상위 20개 단어 비교 분석 결과>

ChatGPT	TF-IDF	Google Gemini	TF-IDF
app	0.042635	good	0.176285
good	0.037183	nice	0.061192
helpful	0.020292	app	0.026179
like	0.018662	great	0.022178
great	0.018012	best	0.022075
image	0.017236	gemini	0.017447
really	0.016696	helpful	0.016811

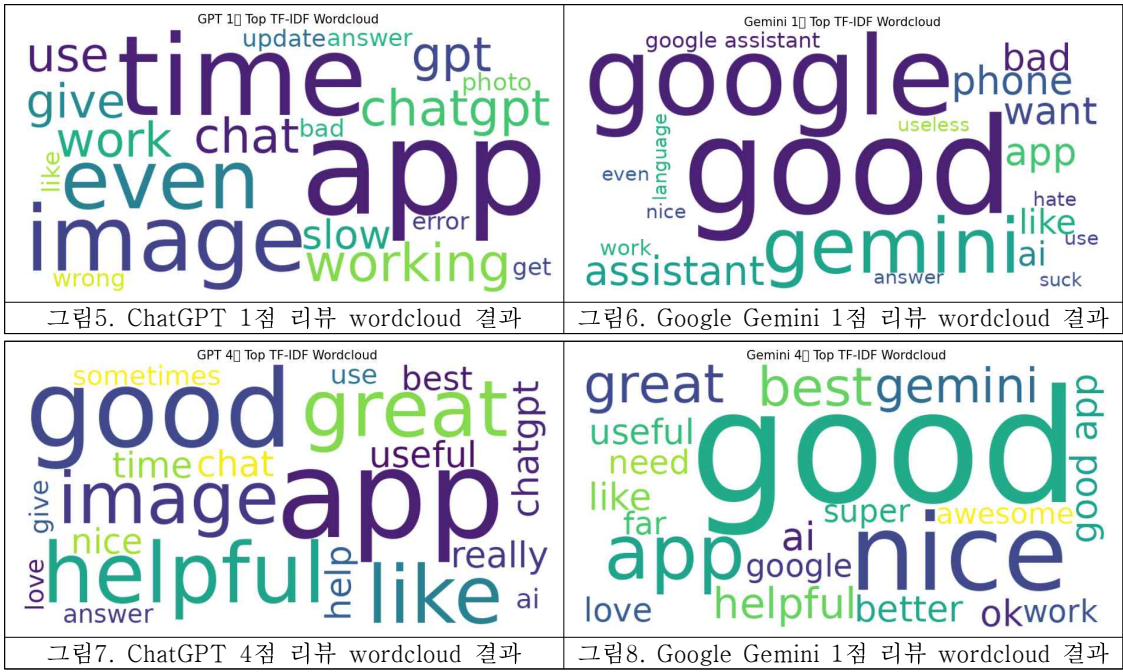
chatgpt	0.016407	ai	0.014354
time	0.016336	like	0.012823
useful	0.016167	ok	0.012444
chat	0.015999	good app	0.012378
help	0.015336	better	0.012093
nice	0.015255	useful	0.011688
best	0.014768	need	0.010434
sometimes	0.014430	love	0.010286
use	0.014418	google	0.009486
ai	0.014395	far	0.009468
give	0.013128	super	0.008882
love	0.012778	work	0.008720
answer	0.012491	awesome	0.008691

1점 리뷰의 분석 결과를 보면, ChatGPT에서 “time”, “image”, “working”, “update”, “error” 등 기능적 문제와 성능 저하에 관한 키워드가 상위에 집중되었다. 이는 주로 답변 속도, 이미지 생성 기능, 오류 발생, 업데이트 이후 문제 등 구체적인 불만사항이 드러난 결과이다. 반면 Gemini의 경우 “google”, “gemini”, “assistant”, “phone” 등이 상위에 위치해, 구글 생태계 및 기기 연동성과 관련한 불만이 핵심으로 나타났다. 또한 “bad”, “useless”, “hate”, “suck”과 같은 강한 부정 감정 표현이 포함되어 Gemini 사용자 불만이 감성적으로도 더 직설적으로 드러나는 특징을 보였다.

4점 리뷰의 TF-IDF 결과에서는, ChatGPT는 “app”, “good”, “helpful”, “like”, “great”과 같은 앱 자체와 유용성 관련 단어가 상위에 위치하였다. 특히 “chatgpt” 자체가 중요한 키워드로 등장한 것은 사용자들이 브랜드 자체를 긍정적으로 언급하는 경향이 강함을 보여준다. 반면 Gemini는 “good”, “nice”, “great” 등 기본적인 긍정 표현과 함께 “gemini”, “google” 등 브랜드명 중심 단어가 상위에 포함되었다. 또한 “ok”, “better”, “need” 같은 표현은 Gemini의 긍정적 평가에도 불구하고 기능 개선 요구가 함께 언급된 점을 반영한다. 즉, ChatGPT는 앱 효용성

과 직접적인 경험, Gemini는 브랜드 인식과 개선 기대가 긍정적 평가의 핵심으로 나타났다.

종합적으로 볼 때, ChatGPT는 긍정 리뷰에서 앱 유용성과 브랜드 자체의 호감도가 강조되고 부정 리뷰에서는 성능과 기능적 오류에 대한 불만이 나타났다. 반면 Gemini는 긍정 리뷰에서 브랜드·생태계 결합의 긍정적 인식이 두드러졌으나 동시에 개선 요구가 나타났으며, 부정 리뷰에서는 구글 생태계 연동의 실패 경험과 더불어 감정적으로 강한 부정 표현이 집중적으로 드러났다.



추가적으로, 각 단어의 중요도와 출현 빈도를 직관적으로 확인하기 위해 워드클라우드를 생성하였다. 본 연구에서는 ChatGPT와 Gemini 각각의 1점(부정) 및 4점(긍정) 리뷰에 대한 워드클라우드를 제시하여, 어떤 단어가 사용자 평가에서 중심적인 의미를 갖는지를 시각적으로 비교하였다. 이를 통해 표와 수치로만 확인하기 어려운 사용자 반응의 직관적 흐름과 핵심 키워드의 집중도를 효과적으로 이해할 수 있다.

3.4 LDA 토픽 모델링 분석

3.4.1 LDA 토픽별 Top 5 키워드 비교

<표9. 1점 리뷰의 LDA 토픽별 Top 5 키워드 비교>

ChatGPT			Google Gemini		
topic	토픽 라벨	keywords	topic	토픽 라벨	keywords
Topic 1	업데이트 및 대화 기능 불만	chat, gpt, update, like, voice, new, chatgpt, back, conversation, text	Topic 1	앱 품질 불만 및 낮은 평가	bad, ai, nice, app, worst, useless, better, poor, chatgpt, chat
Topic 2	유료 결제와 제한 불만	free, limit, plus, subscription, use, version, money, pay, paid, get	Topic 2	답변 정확도 문제	answer, time, question, give, gemini, wrong, asked, ask, information, say
Topic 3	오류 및 로그인 문제	app, working, work, error, phone, please, issue, problem, use, login	Topic 3	구글 어시스턴트 관련 불편	google, assistant, work, gemini, like, ok, use, even, voice, app
Topic 4	이미지 생성 및 속도 불만	image, time, app, photo, slow, bad, picture, working, take, create	Topic 4	기기 연동 및 사용 거부	want, phone, suck, app, hate, language, gemini, change, get, uninstall
Topic 5	답변 부정확성	give, answer, app, question, ai, wrong, time, information, ask, good	Topic 5	이미지 생성 성능 문제	good, working, hai, image, slow, generate, properly, people, picture, nahin

<표10. 4점 리뷰의 LDA 토픽별 Top 5 키워드 비교>

ChatGPT			Googel Gemini		
topic	토픽 라벨	keywords	topic	토픽 라벨	keywords
Topic 0	앱 유용성과 답변 만족	app, good, answer, helpful, help, give, question, useful,	Topic 0	구글 및 기능 활용	gemini, google, like, voice, assistant, need, feature, thing, would,

		best, really			make
Topic 1	브랜드 및 기능 호감	chatgpt, user, ai, voice, like, response, would, feature, need, language	Topic 1	기본적 긍정 평가	good, nice, app, love, cool, far, bad, work, really, working
Topic 2	앱 개선 요구	chat, gpt, app, like, please, would, one, update, message, feature	Topic 2	답변 품질 및 정보 제공	great, ok, excellent, answer, app, help, question, information, give, better
Topic 3	앱 전반 경험	app, good, like, photo, problem, time, best, really, work, ai	Topic 3	사용 편리성과 도움	helpful, useful, awesome, use, super, app, better, google, assistant, easy
Topic 4	이미지 및 속도 관련 긍정 경험	good, image, app, free, nice, time, use, take, great, slow	Topic 4	AI 및 브랜드 긍정 인식	best, ai, app, gemini, like, amazing, need, still, response, chatgpt

먼저 1점 리뷰의 토픽 키워드를 살펴보면, ChatGPT는 “update”, “subscription”, “error”, “image”, “login” 등 구체적인 기능적 문제와 서비스 정책과 관련된 불만이 중심을 이루었다. 사용자는 앱의 업데이트 이후 불편, 유료 결제나 사용 제한에 대한 불만, 로그인 오류와 이미지 생성 기능의 문제를 반복적으로 지적하였다. 이는 ChatGPT의 부정적 인식이 주로 서비스 이용 과정에서 발생하는 기능적 결함과 정책적 제약에 집중되어 있음을 보여준다. 반면 Gemini는 “google”, “assistant”, “phone”, “uninstall”, “suck”, “hate” 등 구글 생태계 및 기기 연동성 문제와 더불어 강한 부정적 감정을 담은 표현이 두드러졌다. 이는 Gemini가 단순 언어 모델이 아니라 구글 생태계와 통합된 서비스라는 특성이 사용자 불만의 핵심 요인으로 작용하고 있음을 보여주며, 동시에 ChatGPT보다 더 감정적으로 격한 반응을 유발했음을 시사한다.

4점 리뷰의 토픽을 보면, ChatGPT는 “helpful”, “useful”, “answer”, “chatgpt” 등에서 알 수 있듯이 앱의 실질적 유용성과 답변 기능의 만족도가 강조되었다. 사용자는 질문-답변 기능, 텍스트 생성 능력, 앱의 전반적 편리성을 긍정적으로 평가하였다. 또한 “update”,

“feature”, “voice” 등의 키워드에서 나타나듯, 긍정적 평가 속에서도 추가 기능 개선 기대가 함께 드러났다. 반면 Gemini의 4점 리뷰에서는 “google”, “assistant”, “gemini” 등 브랜드명과 연계된 키워드가 다수 등장하였다. 이는 Gemini가 독립된 서비스라기보다는 구글 생태계의 일부로서 긍정적으로 평가되었음을 의미한다. 또한 “helpful”, “useful”, “awesome”, “easy” 등의 키워드는 Gemini가 편리성과 사용성 측면에서 긍정적인 평가를 얻었음을 보여준다. 다만 “need”, “still” 같은 단어는 여전히 기능적 개선 요구가 병행됨을 보여준다.

종합하면, ChatGPT는 부정적 리뷰에서 정책적 제약과 기능 오류에 집중된 반면, 긍정적 리뷰에서는 실질적 유용성과 답변 만족도가 강조되었다. Gemini는 부정적 리뷰에서 구글 생태계 연동 문제와 강한 부정 감정이 중심을 이루었으며, 긍정적 리뷰에서는 브랜드와 플랫폼 결합에 대한 긍정적 인식과 편리성이 두드러졌다. 이러한 차이는 ChatGPT가 언어모델 자체의 성능과 기능 안정성을 평가받는 반면, Gemini는 플랫폼 연계성과 사용자 경험을 중심으로 평가받는다라는 점을 보여준다.

3.4.2 LDA 주제 분포 비교

그림8은 ChatGPT와 Google Gemini의 부정적 리뷰에 대해 동일 키워드에 따른 LDA 토픽의 분포를 비교한 그래프이다. 분석 목표는 ChatGPT와 Google Gemini의 부정적 리뷰에서 사용자 불만이 어떤 주제에 집중되는지를 비교하는 것이다.

분석 결과, 두 모델 모두에서 공통적으로 “app”, “work”, “image”, “answer”와 같은 키워드가 상위에 포함되었다. 이는 사용자 불만이 주로 앱 사용성, 동작 오류, 이미지 생성 문제, 답변의 정확성에 집중되

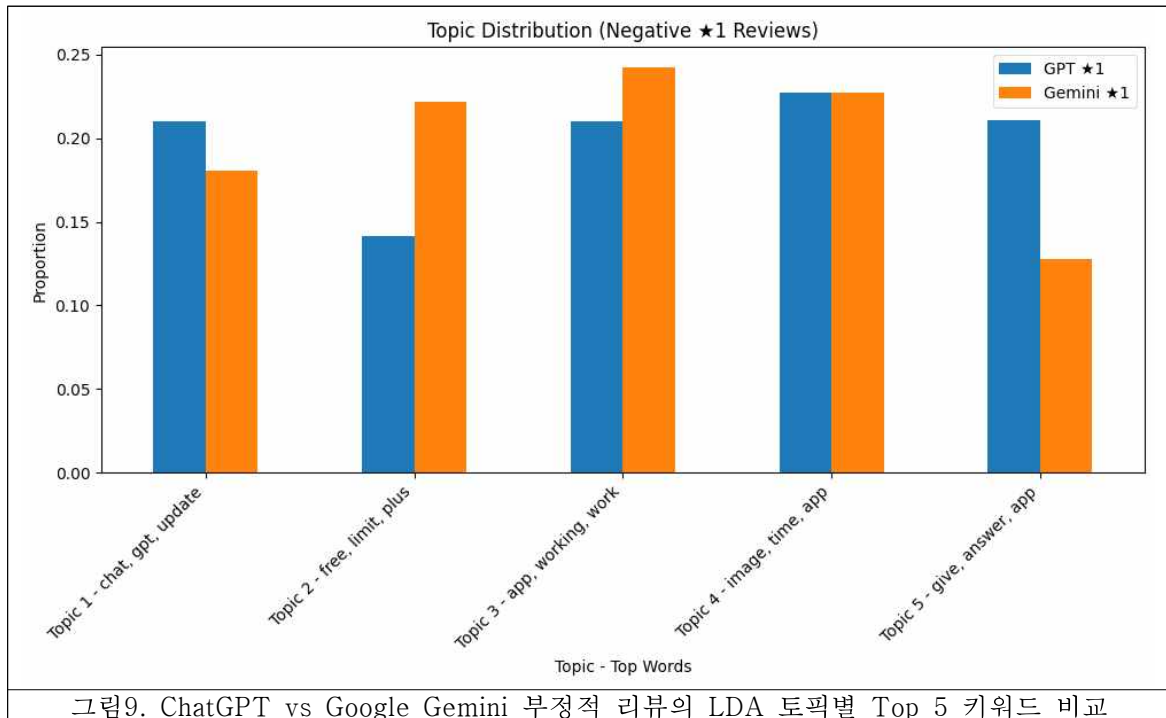


그림9. ChatGPT vs Google Gemini 부정적 리뷰의 LDA 토픽별 Top 5 키워드 비교

어 있음을 의미한다. 그러나 세부적으로 살펴보면, ChatGPT는 “update”, “subscription”, “limit” 등 정책적 제약(업데이트 불편, 유료 결제, 사용 제한)에 대한 불만이 뚜렷하게 나타났고, Gemini는 “google”, “assistant”, “phone”, “uninstall” 등 구글 생태계 연동 및 기기 활용 문제가 핵심 불만으로 드러났다.

또한 Gemini의 1점 리뷰에서는 “suck”, “hate”, “useless”와 같은 감정적으로 강한 부정 표현이 포함된 반면, ChatGPT는 상대적으로 기능적 문제 지적이 중심이었다. 이는 ChatGPT의 부정적 평가는 주로 기능적 불만에, Gemini의 부정적 평가는 생태계 연동 실패와 감정적 거부감에 초점이 맞춰져 있음을 보여준다. 그 외 키워드들은 비교적 급격하게 비중이 감소하는 경향을 보인다. 이는 블로그 담론이 특정 경험 장면이나 주제를 중심으로 단일 축으로 전개되며 주제 집중도가 높다는 점을 시사한다.

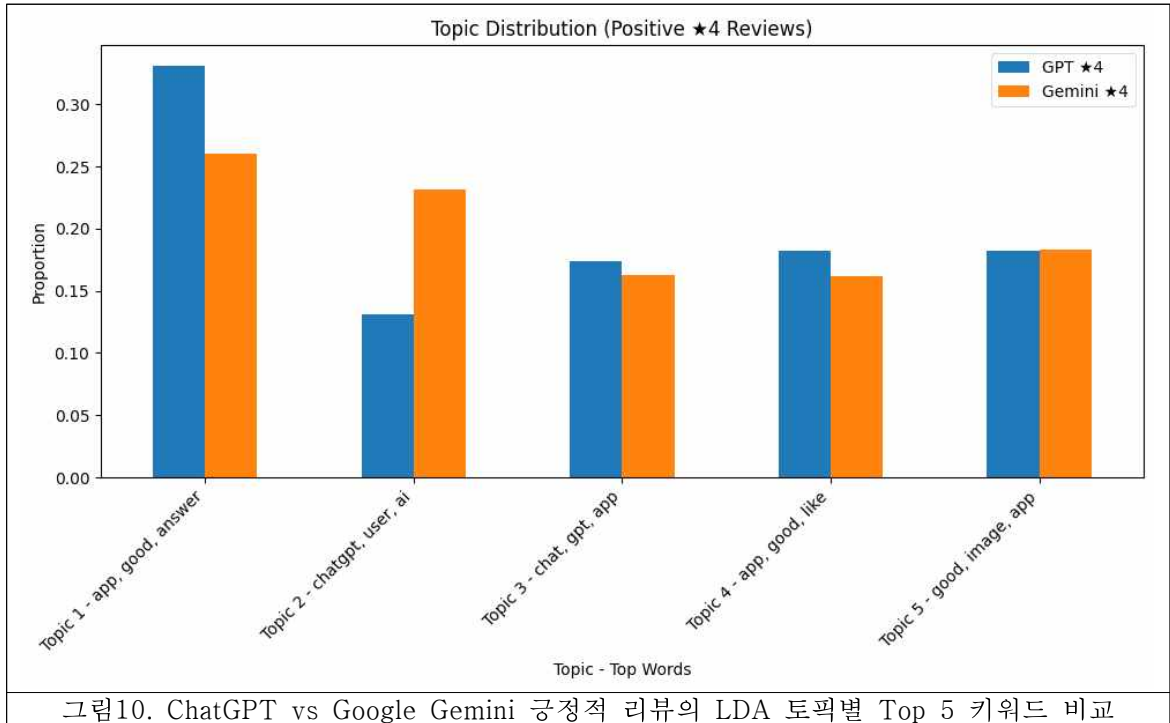


그림10. ChatGPT vs Google Gemini 긍정적 리뷰의 LDA 토픽별 Top 5 키워드 비교

그림9은 ChatGPT와 Google Gemini의 긍정적 리뷰에 대해 동일 키워드에 따른 LDA 토픽의 분포를 비교한 그래프이다. 분석 목표는 두 모델에 대한 긍정적 경험이 어떤 요인에서 비롯되는지를 규명하는 것이다.

분석 결과, ChatGPT와 Gemini 모두 “app”, “good”, “answer”, “helpful”, “useful”과 같은 긍정 키워드가 상위에 위치하였다. 이는 공통적으로 앱의 편리성과 답변 기능의 유용성이 긍정적 평가의 주요 요인임을 보여준다. 그러나 ChatGPT의 경우 “chatgpt”, “ai”, “voice”와 같은 브랜드 및 기능적 특징이 직접적으로 언급되어, 긍정적 인식이 모델 자체 성능과 신뢰성에 기반함을 확인할 수 있다. 반면 Gemini는 “google”, “assistant”, “gemini” 등 브랜드와 플랫폼 연계성이 반복적으로 등장하여, 긍정적 경험이 구글 생태계와의 결합 및 편리성에서 비롯된다는 점이 특징적이다. 또한 Gemini는 “cool”, “awesome”, “super” 등 감정적으로 호의적인 단어가 포함되어, 사용자의 정서적 만족이 긍정 평가에 중요한 역할을 하고 있음을 알 수 있다.

정리하면, 1점 리뷰에서는 ChatGPT가 정책 및 기능적 제약에 대한 불만이 두드러지고, Gemini는 생태계 연동 실패와 강한 감정적 불만이 집중적으로 드러났다. 4점 리뷰에서는 두 모델 모두 앱 유용성과 답변 기능이 긍정적 평가를 이끌었으나, ChatGPT는 모델 자체의 성능과 브랜드 신뢰가 강조되었고, Gemini는 구글 생태계와의 연계성 및 정서적 만족이 주요 요인으로 나타났다. 이는 두 모델이 제공하는 경험의 성격 차이를 잘 보여주며, ChatGPT는 기능 안정성 중심의 평가, Gemini는 생태계 활용 경험 중심의 평가로 구분할 수 있다.

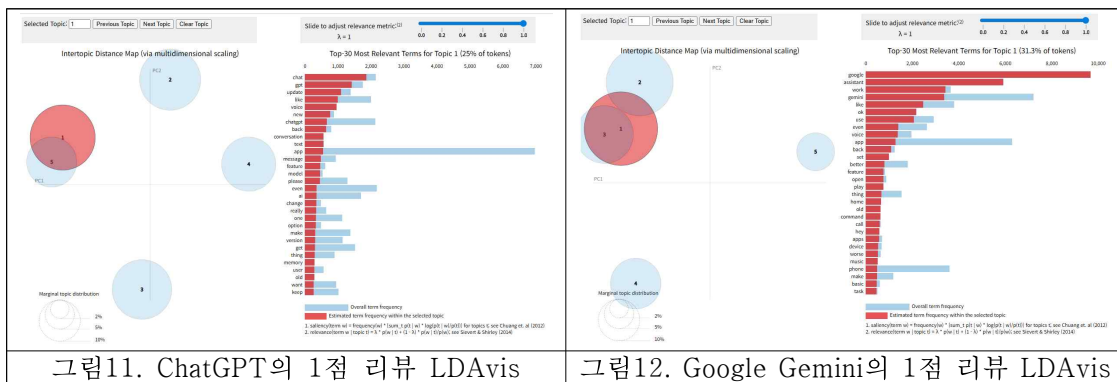
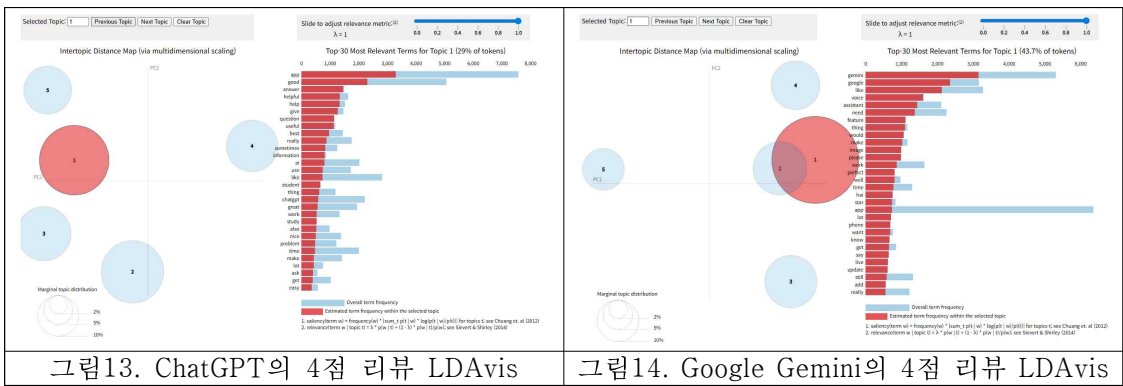


그림12. Google Gemini의 1점 리뷰 LDAvis

ChatGPT의 1점 리뷰에서는 토픽 원의 크기와 위치를 통해 확인할 수 있듯이, 사용자 불만이 정책적 제약과 기능 오류에 집중되어 있었다. 특히 “update”, “subscription”, “limit”, “error”, “image” 등이 상위 키워드로 나타나며, 업데이트 이후 발생하는 불편, 유료 결제 및 사용 제한 문제, 이미지 생성 기능의 오류가 주요 불만 요인으로 확인되었다. LDAvis 시각화에서도 이들 토픽이 상대적으로 큰 원으로 형성되어 있어, 전체 불만 중 높은 비중을 차지함을 보여준다. 또한 토픽 간 거리가 비교적 가까워, ChatGPT의 부정적 평가가 몇 가지 기능적 문제를 중심으로 응집되어 있음을 알 수 있다.

반면 Gemini의 1점 리뷰는 ChatGPT와 달리 구글 생태계 연동 문제가 핵심 불만으로 드러났다. LDAvis 결과에서 “google”, “assistant”, “phone”, “uninstall” 등 구글 어시스턴트 및 기기 제어 관련 키워드가 뚜렷하게 나타났고, 이들 토픽의 비중이 크고 서로 밀접하게 연결되어 있었다. 또한 Gemini의 1점 리뷰에서는 “suck”, “hate”, “useless” 등 감성적으로 강한 부정 표현이 두드러졌는데, 이는 LDAvis에서 별도로 분리된 토픽으로 나타나 단순 기능적 문제를 넘어 강한 거부감을 반영한다는 점에서 ChatGPT와 차이를 보였다. 정리하면, ChatGPT는 구체적인 기능·정책 문제에 대한 불만이 중심이었고, Gemini는 구글 생태계 연계 실패와 감성적 반응이 복합적으로 나타났음을 확인할 수 있다.



ChatGPT의 4점 리뷰에서는 긍정적 평가가 앱의 실질적 유용성과 답변 기능의 만족도에 집중되었다. LDAvis 결과에서 큰 원으로 나타난 토픽들은 “helpful”, “useful”, “answer”, “chatgpt”, “ai”와 같은 키워드 중심으로 형성되었으며, 이는 사용자가 GPT의 답변 능력과 전반적 활용성에 높은 만족을 보였음을 시사한다. 동시에 “update”, “feature”, “voice”와 같은 단어도 토픽으로 나타나, 긍정적 평가 속에서도 기능 개선 요구가 병행되고 있음을 보여주었다. 토픽 간 거리는 다소 분산되어 있어, 다양한 긍정 경험이 골고루 언급된 특징이 있었다.

Gemini의 4점 리뷰에서는 브랜드와 생태계 연계성이 핵심 긍정 요인으로 나타났다. LDAvis 시각화에서 “gemini”, “google”, “assistant”가 포함된 토픽이 크게 형성되어, 사용자 만족이 구글 서비스와의 통합적 사용 경험에서 비롯

되었음을 확인할 수 있었다. 또한 “helpful”, “useful”, “awesome”, “super”와 같은 감정적으로 호의적인 단어가 토픽별 상위 키워드로 등장하여, Gemini 사용자들이 기능적 만족뿐 아니라 정서적 긍정 경험도 강조하는 경향을 보였다. 다만 “need”, “still” 같은 단어가 토픽에 포함된 것은 여전히 일부 개선 요구가 존재함을 보여준다.

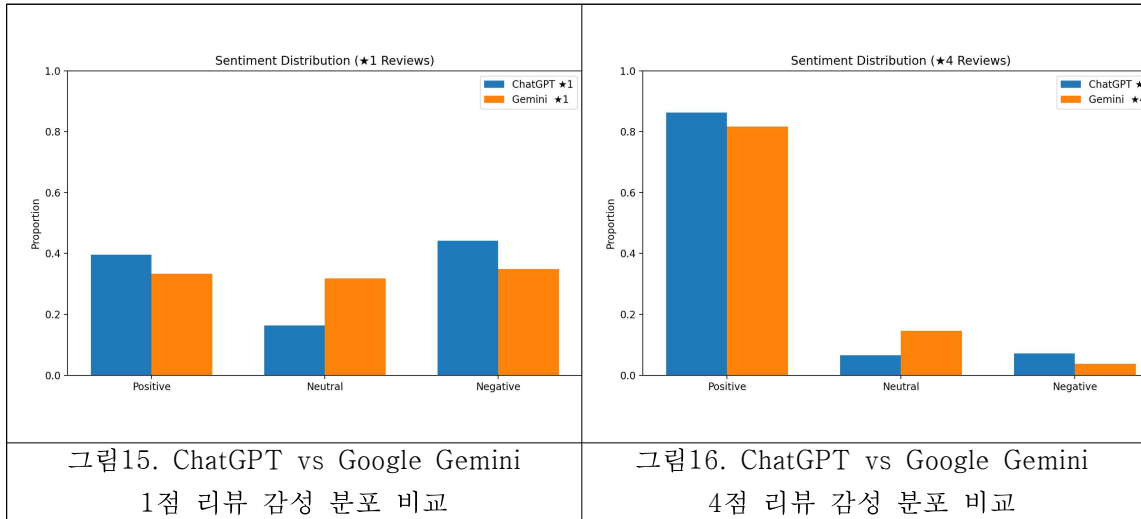
LDAvis 시각화 결과를 종합하면, 1점 리뷰에서는 ChatGPT가 업데이트, 결제, 기능 오류와 같은 구체적 문제에 불만이 집중된 반면, Gemini는 구글 어시스턴트 및 기기 연동 실패와 더불어 강한 감정적 불만이 중심을 이루었다. 4점 리뷰에서는 ChatGPT가 답변 정확성과 앱의 유용성에 대한 긍정 평가가 두드러진 반면, Gemini는 구글 생태계와의 통합 경험과 정서적 만족감이 주요 긍정 요인으로 나타났다. 이는 ChatGPT가 기능 안정성과 모델 성능 중심으로 평가되는 반면, Gemini는 플랫폼 연계성과 사용자 체감 경험을 중심으로 평가된다는 차이를 보여준다.

본 연구에서는 ChatGPT와 Google Gemini의 사용자 리뷰를 대상으로 LDA 토픽 모델링을 수행하여 긍정적(4점) 및 부정적(1점) 평가에 내재된 주제들을 비교하였다. 비교 결과를 종합하면, ChatGPT는 부정적 경험이 구체적인 기능·정책 문제에 집중되고, 긍정적 경험은 모델 자체의 성능과 신뢰성에서 비롯되는 특성이 있다. 반대로 Gemini는 부정적 경험이 생태계 연동 실패와 감정적 불만으로 확산되는 반면, 긍정적 경험은 구글 생태계 통합성과 감정적 호감에 기반하는 특징을 보였다. 이는 두 모델이 사용자에게 제공하는 경험의 성격이 상이하다는 점을 보여주며, ChatGPT는 기능 안정성과 텍스트 처리 능력을, Gemini는 플랫폼 연계성과 사용자 체감 경험을 핵심 가치로 인식시키고 있음을 확인할 수 있다.

4. 감성 분석

4.1 감성(긍정/부정) 분포 비교

1점 리뷰의 감성 분포는 ChatGPT와 Gemini 모두 부정적 감정이 높은 비중을 차지했다. ChatGPT는 약 44%가 부정으로 분류되었으며, Gemini는 약

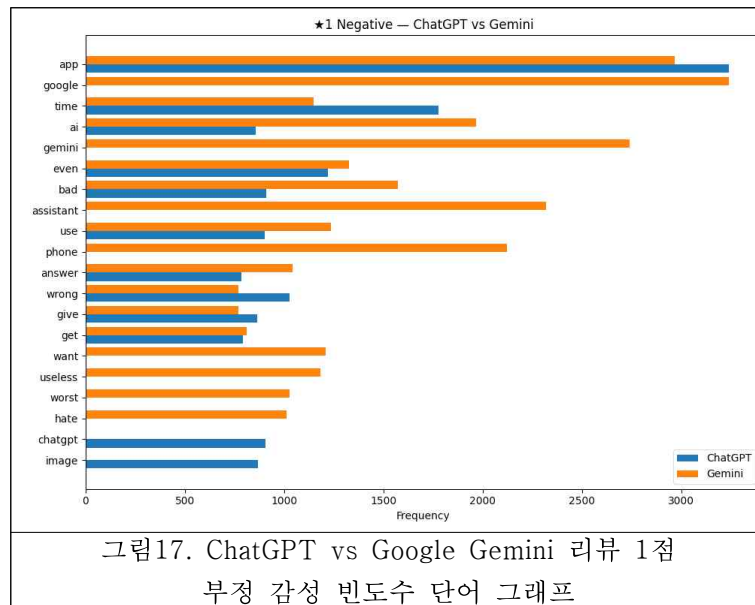


35%로 상대적으로 낮았다. 또한 ChatGPT는 긍정 감정(약 40%)과 부정 감정이 비슷한 수준으로 공존했는데, 이는 일부 사용자들이 낮은 별점을 주면서도 특정 기능에 대해서는 긍정적으로 평가했음을 보여준다. 반면 Gemini는 긍정과 부정이 모두 30%대 중반으로 분포했으며, 중립 감정의 비율(약 32%)이 ChatGPT보다 높았다. 이는 Gemini 사용자들이 불만을 직접적으로 표출하기보다 애매하거나 복합적인 평가를 남기는 경향이 상대적으로 강했음을 의미한다.

반면에 4점 리뷰에서는 ChatGPT와 Gemini 모두 압도적으로 긍정 감성 비율이 높게 나타났다. ChatGPT는 약 86%가 긍정으로 분류되었으며, Gemini 역시 약 82% 수준으로 높게 나타났다. 이는 두 모델이 모두 사용자에게 전반적으로 만족감을 제공했음을 보여준다.

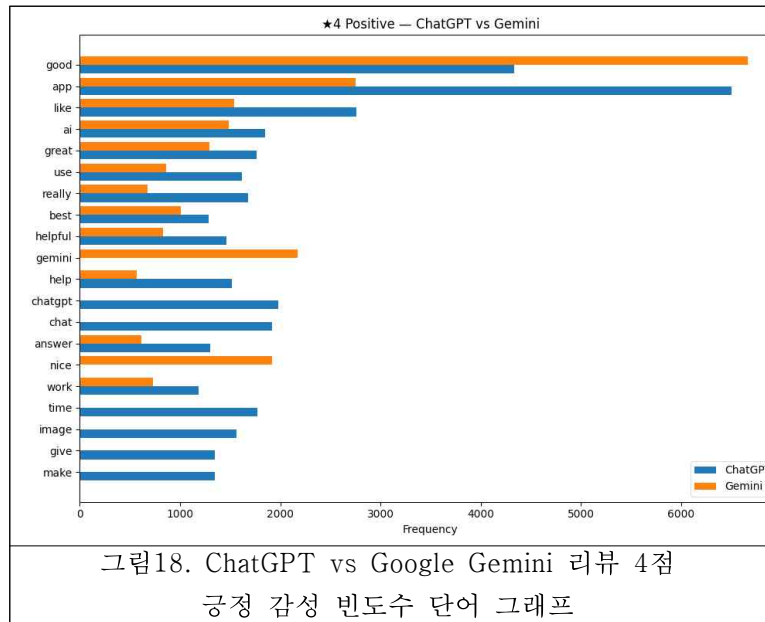
다만 세부적으로 보면, Gemini는 중립 감정(약 15%)의 비중이 ChatGPT(약 6%)보다 높게 나타났다. 이는 Gemini 사용자들이 긍정적인 평가를 남기면서도 기능 개선이나 추가 요구사항을 언급하는 경우가 많았음을 의미한다. 반대로 부정적 감성 비율은 ChatGPT에서 약간 더 높아, 일부 사용자가 성능 문제나 기능 오류를 여전히 지적했음을 시사한다.

4.2 감성별 빈도수 단어 그래프



1점 리뷰에서 ChatGPT는 이미지 생성 오류, 답변의 부정확성, 업데이트와 같은 기능적 문제에 대한 불만이 핵심으로 나타났다. 사용자는 앱이 기본적으로 제공해야 할 기능이 제대로 작동하지 않는 점을 주로 지적하며, 성능 안정성에 대한 요구를 강하게 드러냈다. 반면 Gemini의 부정 리뷰는 구글 어시스턴트와의 연동 문제, 휴대폰 제어 기능의 오류 등 생태계 차원의 불편이 중심에 있었고, 여기에 “useless”, “hate”, “worst”와 같은 감정적으로 강한 표현이 더해졌다. 이는 Gemini 사용자가 단순히 기능적 불만에 그치지 않고, 서비스 자체를 불필요하거나 부정적으로 인식하는 경향이 있음을 보여준다.

4점 리뷰에서는 두 모델 모두 “good”과 “app”을 중심으로 전반적인 만족감이 공통적으로 나타났지만, 강조되는 긍정 경험의 성격은 달랐다. ChatGPT는 “helpful”, “answer”, “useful”과 같은 키워드에서 드러나듯이 질문·답변 기능과 앱의 실질적인 효용성이 긍정 평가의 핵심으로 작용했다. 특히 “chatgpt”라는 브랜드명 자체가 반복적으로 언급된 점은 사용자들이 모델의 성능과 신뢰도를 높게 평가했음을 의미한다. 반대로 Gemini는 “google”, “gemini”, “assistant”와 같은 단어가 두드러지며, 독립적인 앱이라기보다



구글 생태계와의 연계성 속에서 긍정적인 경험이 형성되었음을 보여주었다. 또한 “nice”, “cool”, “awesome”과 같은 감정적으로 호의적인 표현이 다수 등장해, Gemini는 기능적 만족뿐 아니라 정서적 호감 역시 긍정 평가를 강화하는 요인으로 작용했음을 알 수 있다.

제5장 연구 결과정리 합의

본 연구에서는 인공지능 기반 대화형 언어모델(ChatGPT와 Google Gemini)의 사용자 리뷰를 대상으로 다양한 텍스트 마이닝 기법을 적용하여 비교 분석을 수행하였다. 분석 대상은 Google 플레이스토어에 게시된 두 모델의 리뷰 중 별점 1점과 4점에 해당하는 데이터를 선별하였다. 이는 극단적인 긍정·부정 경험을 통해 사용자 인식의 핵심 요인을 파악하기 위함이며, 특히 별점 5점은 무분별한 칭찬과 단순 긍정 반응이 다수 포함될 가능성이 높아 제외하였다. 따라서 본 연구는 기능적 불만과 부분적 만족을 드러내는 1점·4점 리뷰에 집중하여, 실제 사용자가 경험하는 문제와 긍정 요인을 균형 있게 규명하는 것을 목표로 하였다.

분석 결과, ChatGPT와 Gemini는 공통적으로 “app”, “good”, “answer”와 같은 핵심 키워드가 긍정과 부정 양쪽에서 반복적으로 나타나며, 두 모델 모두 사용자 평가가 앱 자체의 유용성과 답변 기능을 중심으로 형성되고 있음을 확인할 수 있었다. 그러나 세부적으로 살펴보면 양자의 특성은 뚜렷하게 구분된다.

먼저 부정적 리뷰(1점 리뷰)의 분석에서 ChatGPT는 업데이트 불편, 유료 결제 및 사용 제한, 이미지 생성 오류, 로그인 문제와 같은 구체적인 기능적 문제가 핵심 불만으로 나타났다. TF-IDF와 이그램 분석에서 “update”, “subscription”, “error”, “image” 등의 단어가 상위에 위치했으며, LDA 토픽 모델링에서도 이와 유사한 패턴이 확인되었다. 네트워크 분석 결과 ChatGPT의 부정 토픽들은 몇몇 키워드에 응집되어 나타났는데, 이는 사용자 불만이 다소 집중적이고 구체적인 문제에 국한되어 있음을 시사한다. 다시 말해, ChatGPT 사용자들은 기능적 오류와 정책적 제약이 불만의 주된 원인이었으며, 서비스 전반에 대한 부정적 감정보다는 특정 기능이 제대로 작동하지 않는 점을 지적하는 경향이 강했다.

반면 Gemini의 1점 리뷰에서는 구글 어시스턴트와의 연동, 휴대폰 제어 기능, 기기 호환성과 관련된 불편이 반복적으로 언급되었다. 특히 “google”, “assistant”, “phone”, “gemini” 등이 상위 키워드로 나타났으며, 여기에 “useless”, “hate”, “worst”, “suck”과 같은 감정적으로 강한 부정 표현이 함께 등장하였다. 이는 Gemini 사용자들이 단순히 특정 기능의 오류를 지적하는 차원을 넘어, 서비스 전반에 대한 실망감과 거부감을 강하게 표현하고 있음을 보여준다. LDAvis 시각화에서도 Gemini의 부정 토픽은 기기 연동 문제와 감정적 불만이 결합된 형태로 나타났으며, 이는 ChatGPT의 기능 중심 불만과 대조적이다. 감정분석 결과 또한 이를 뒷받침

했는데, ChatGPT는 부정 리뷰에서 부정 감정 비율이 상대적으로 높고 명확하게 드러난 반면, Gemini는 부정·긍정이 혼재하고 중립 감정의 비율이 더 높아, 불만의 감정 표현이 다층적이고 복합적인 양상을 띠고 있음을 알 수 있었다.

긍정적 리뷰(4점 리뷰)의 분석에서는 두 모델 모두 전반적으로 높은 만족도가 확인되었다. ChatGPT의 경우 “helpful”, “useful”, “answer”, “chatgpt”, “ai” 등의 단어가 상위에 등장하여, 사용자들이 질문·답변 기능과 앱의 전반적 유용성에 만족하고 있음을 보여주었다. 특히 “chatgpt”라는 브랜드명 자체가 긍정 키워드로 반복적으로 나타난 것은 ChatGPT의 성능과 신뢰도가 사용자 인식에 깊이 자리잡고 있음을 시사한다. LDA 결과에서도 ChatGPT의 긍정 토픽은 “answer”, “helpful”, “update”, “voice” 등을 포함하여, 기능적 만족과 더불어 추가 기능 개선에 대한 기대가 동시에 존재하는 양상을 보였다.

Gemini의 4점 리뷰에서는 “gemini”, “google”, “assistant”가 핵심 키워드로 도출되었으며, 이는 Gemini가 독립적인 대화형 모델이라기보다는 구글 생태계 속에서 통합적으로 활용되는 서비스로 인식되고 있음을 보여준다. 또한 “nice”, “cool”, “awesome”, “super” 등 감정적으로 긍정적인 단어가 두드러지게 나타나, Gemini 사용자들은 기능적 만족뿐 아니라 정서적 호감과 즐거움을 긍정 평가에 반영하고 있음을 확인할 수 있었다. 감정분석 결과에서도 ChatGPT는 명확하게 긍정 감정이 높게 나타났고, Gemini는 긍정 비율이 다소 낮지만 중립 감정이 높게 나타나 긍정 속에서도 개선 요구가 함께 내포된 평가가 이루어지고 있음을 알 수 있었다.

네트워크 분석 또한 이러한 결과를 뒷받침하였다. ChatGPT의 긍정 네트워크는 노드 수가 많지만 연결 밀도가 낮아 다양한 키워드가 분산되어

있었으며, 이는 긍정적 경험이 다채롭게 분포하고 있음을 의미한다. 반면 Gemini의 긍정 네트워크는 노드 수가 적지만 밀도와 클러스터링 계수가 높아, 사용자들이 소수의 핵심 주제(구글 연계성, 사용 편리성)에 긍정적 경험을 집중하고 있음을 보여주었다.

이상의 결과를 종합하면, ChatGPT는 부정 경험이 기능적 문제에 집중되고 긍정 경험은 모델 성능과 실질적 유용성을 중심으로 형성되는 경향이 뚜렷하다. 반대로 Gemini는 부정 경험이 구글 생태계 연동의 실패와 감정적 불만에 의해 강화되며, 긍정 경험은 브랜드와 플랫폼 결합 및 정서적 만족에 의해 형성되는 특징을 보였다. 이는 두 모델이 사용자에게 제공하는 경험의 성격이 본질적으로 다르다는 점을 의미한다. ChatGPT는 기능 안정성과 텍스트 처리 성능을 통해 신뢰를 얻고 있으며, Gemini는 구글 플랫폼과의 통합성, 그리고 감정적으로 긍정적인 사용자 경험을 통해 호감을 형성한다.

이러한 결과는 학문적·실무적 함의를 동시에 지닌다. 학문적으로 본 연구는 대규모 언어모델의 사용자 경험을 텍스트 마이닝 기법으로 실증 분석하였다는 점에서 의미가 있다. 기존 연구들이 모델의 성능 지표에 집중하는데 비해, 본 연구는 실제 사용자 리뷰 데이터를 활용하여 서비스 사용성, 정책 수용성, 감정적 반응을 포괄적으로 분석하였다. 실무적으로는 두 모델의 서비스 전략에 중요한 시사점을 제공한다. ChatGPT는 기능 안정성과 정확성 개선을 통해 부정적 경험을 줄이고, 기존의 높은 신뢰도를 강화하는 방향으로 발전할 필요가 있다. 반면 Gemini는 구글 생태계와의 연동 안정성을 높이고, 기기 제어 문제를 해결함으로써 부정적 인식을 완화할 수 있을 것이다. 또한 Gemini가 보유한 감정적 호감 요인을 유지·확장하는 전략은 긍정적 사용자 경험을 더욱 강화할 수 있는 기반이 된다.

종합적으로 본 연구는 ChatGPT와 Gemini라는 대표적 대화형 언어모

텔을 비교함으로써, 향후 인공지능 서비스가 나아가야 할 방향을 제시하였다. 사용자 경험은 단순히 모델 성능의 결과가 아니라, 정책, 생태계 연계성, 기능 안정성, 감정적 반응이 종합적으로 작용한 결과라는 점을 확인하였다. 이는 대규모 언어모델 연구와 서비스 전략 수립 모두에 중요한 함의를 제공하며, 향후 다양한 사용자 집단과 플랫폼 환경에서 추가적인 비교 연구가 필요함을 시사한다.

참 고 자 료

OpenAI. (2022). Introducing ChatGPT. OpenAI Blog.

<https://openai.com/blog/chatgpt>

Google DeepMind. (2023). Gemini: our largest and most capable AI model. Google DeepMind Blog.

<https://deepmind.google/discover/blog/google-gemini-ai>

학생 소개

송종찬 (SONG JONGCHAN)

AI미디어텍스트이해 수강생

성균관대학교 소비자학과 학부생

<관심분야> : 인공지능, 웹/앱개발, 운영체제

