

OTT 앱 리뷰 분석을 통한 서비스 개선 기회 발굴 방안 연구

Exploring Service Improvement Opportunities through Analysis of OTT App Reviews

이중민¹, 송지훈^{2*}

Joongmin Lee¹, Chie Hoon Song^{2*}

〈Abstract〉

This study aims to suggest service improvement opportunities by analyzing user review data of the top three OTT service apps(Netflix, Coupang Play, and TVING) on Google Play Store. To achieve this objective, we proposed a framework for uncovering service opportunities through the analysis of negative user reviews from OTT service providers. The framework involves automating the labeling of identified topics and generating service improvement opportunities using topic modeling and prompt engineering, leveraging GPT-4, a generative AI model. Consequently, we pinpointed five dissatisfaction topics for Netflix and TVING, and nine for Coupang Play. Common issues include “video playback errors”, “app installation and update errors”, “subscription and payment” problems, and concerns regarding “content quality”. The commonly identified service enhancement opportunities include “enhancing and diversifying content quality”, “optimizing video quality and data usage”, “ensuring compatibility with external devices”, and “streamlining payment and cancellation processes”. In contrast to prior research, this study introduces a novel research framework leveraging generative AI to label topics and propose improvement strategies based on the derived topics. This is noteworthy as it identifies actionable service opportunities aimed at enhancing service competitiveness and satisfaction, instead of merely outlining topics.

Keywords : OTT, App review, Topic modeling, Prompt engineering, Service improvement

1 주저자, 석사과정, 경상국립대학교 대학원 기술경영학과
E-mail: m2live@gnu.ac.kr

2* 교신저자, 조교수, 경상국립대학교 대학원 기술경영학과
E-mail: chsong01@gnu.ac.kr

1 First Author, Graduate Student (Master's. program), Gyeongsang National University, Department of Management of Technology
E-mail: m2live@gnu.ac.kr

2* Corresponding Author, Assistant Professor, Gyeongsang National University, Department of Management of Technology
E-mail: chsong01@gnu.ac.kr

1. 서론

오늘날 우리는 디지털 콘텐츠가 범람하는 시대에 살고 있다. 한국콘텐츠진흥원이 발간하는 「2023년 상반기 콘텐츠 산업 동향분석 보고서」에 의하면 2023년 콘텐츠 산업의 매출액과 수출액은 전년 대비 각 2.5% 그리고 1.3% 증가한 것으로 나타났다[1]. 코로나19 이후 콘텐츠 산업이 회복세로 전환되며 특히 음악과 영화 분야에서 가장 큰 매출성장률을 달성했다. 2022년 콘텐츠 수출액의 경우 역대 최대인 132억 달러를 넘어서며 이는 주요 수출 품목인 가전과 이차전지를 뛰어넘는 실적을 기록했다. 한편 문화체육관광부와 한국콘텐츠진흥원이 발표한 「콘텐츠산업 2023년 결산 및 2024년 전망」에 의하면 당분간 IP 기반 산업의 강세는 지속될 것으로 예상되지만, 국내 OTT 플랫폼은 매년 점진적 영업손실 확대에 의한 오리지널 콘텐츠 제작의 축소의 위기를 맞고 있다[2]. OTT는 Over-The-Top의 약자로, OTT 서비스는 케이블이나 위성 TV 공급업체를 통하지 않고 직접 인터넷을 통해 시청자에게 제공되는 스트리밍 미디어 서비스를 지칭한다[3].

인터넷 환경과 미디어 기술 발전이 더해져 OTT 서비스는 국내·외를 막론하고 급성장하고 있으며, 코로나19 팬데믹으로 사회적 거리두기가 필연적이었던 시기에 더욱 거센 성장세를 보였다. 특히, 국내 OTT 서비스 이용률은 급증해 2022년 72.0%로 2020년 66.3%보다 5.7% 증가했다[4]. 2016년 넷플릭스가 한국에 진출하면서 국내 OTT 시장의 경쟁은 더욱 심화되었고, 각 운영 플랫폼은 경쟁의 우위를 확보하기 위해 영화와 드라마의 경계를 넘나드는 “씨네틱 드라마”와 같은 OTT 스타일의 독자적인 시리즈 콘텐츠 생산에 심혈을 기울이고 있다. 실제 <오징어게임>, <D.P>와 같은 자체 제작 콘텐츠가 출시된 2021년 9월, 넷플릭

스(Netflix) 앱의 신규 설치자가 전월 574,763명에서 1,196,987명으로 대폭 확대되었으며, 이는 자체 제작 콘텐츠가 OTT 플랫폼의 경쟁역량을 효과적으로 강화할 수 있음을 반영했다[5]. 그러나 현재 OTT 서비스 시장은 제작비 상승과 수익성 악화로 같은 요소로 경쟁이 더 치열해지고 있다[6]. 이에 더해 글로벌 OTT 플랫폼의 시장 선점과 로컬라이징 전략으로, 후발주자로 분류되는 토종 OTT 플랫폼은 국내 유통에만 의존하는 한계와 해외 제작사 대비 낮은 자본력으로 인해 경쟁력이 약화되고 있다는 지적이 나오고 있다[7]. 이를 타개하기 위해 최근 토종 OTT 플랫폼 티빙(TVING)과 웨이브(Wavve)는 합병을 위한 양해각서를 체결하였다[8]. 이러한 배경에서 OTT 서비스의 위기 개선을 위한 전략적 고찰은 시의적절하다고 볼 수 있다. 특히 서비스가 앱을 통해 소비됨에 따라, 앱 사용자의 경험을 분석하고, 신규 서비스 기회 발굴을 위한 사용자 피드백에 대한 분석이 요구되고 있다.

본 연구에서는 구글 플레이스토어에 등록된 OTT 서비스 중 국내 이용자 순위 기준 상위 3사(넷플릭스, 쿠팡플레이, 티빙)의 앱에 대한 리뷰 데이터를 수집해 분석하고자 한다. 특히, 사용자 경험을 재구성한 피드백에 대한 분석은 각 개별 OTT 서비스 앱에 대한 고객 니즈와 만족도에 영향을 주는 요인을 찾아내는 유용하게 작용할 수 있다. 본 연구에서는 수집된 리뷰 데이터에 토픽 모델링을 적용한 후, 대표적인 생성형 AI인 OpenAI의 API를 활용해 도출된 주제에 대한 레이블링과 프롬프트 엔지니어링(prompt engineering)을 통한 서비스 개선 기회를 발굴하고자 한다. 여기서 프롬프트 엔지니어링은 입력 쿼리, 즉 '프롬프트'를 설계하고 개선하여 거대 언어 모델(LLM: large language model)에서 원하는 응답을 이끌어내는 과정이다. 이를 통해 고객 관점에서 OTT

서비스 앱 사용 중 겪는 불만족 유발 요인을 이해해 서비스 개선 방향을 제시하고자 한다. 본 연구는 단순히 토픽모델링의 적용을 통한 주제 해석이 아닌, 최신 언어 모델과의 결합을 통한 구체적 서비스 기회 발굴 방안을 제시한다는 점에서 의의를 지닌다.

본 논문의 구성은 다음과 같다. 2장은 OTT 서비스 관련 선행연구를 제시한다. 3장에서는 연구에서 사용한 데이터 및 활용한 분석 프로세스에 대해 기술한다. 4장에서는 논문의 분석 결과와 서비스 개선 기회를 제시한다. 5장에서는 연구의 시사점과 한계점 및 향후 연구 방향에 대해 기술한다.

2. OTT 서비스 관련 선행연구

OTT 서비스와 관련된 선행연구는 소비자의 이용행태 및 이용자 특성 분석[9, 10], OTT 산업 현황[11], 사용자의 서비스 만족도[12], 서비스 성공요인[13] 및 사용자 경험 등으로 구분할 수 있다. 그 중 사용자 경험을 분석한 선행연구는 다음과 같이 정리할 수 있다. 조호수 외(2021)는 구글 스토어에 등록된 12개의 OTT 서비스 앱의 온라인 리뷰에 대하여 토픽모델링을 적용해 이용자 반응을 4가지의 주제(콘텐츠 품질, 서비스 품질, 비용 및 결제, 지연 및 기피)로 구분하였고, 토픽별 감성 점수를 기반으로 OTT 서비스의 강점과 약점을 평가하였다[14]. 이성준(2022)은 OTT 서비스 앱 중 웨이브의 온라인 리뷰에 대해 평점을 기준으로 긍정 또는 부정으로 분류하여 각각의 만족, 불만족 유발 요인을 토픽모델링을 활용해 추출 후 분석하였다[15]. 고민정·이상원(2021)은 넷플릭스 관련 블로그 리뷰 데이터를 통해 코로나 전·후 실용성 및 감정 관련 키워드에 대한 빈도 분석을 수행해 이용자의 관심사 변화를 포착하였

다[16]. 그 외 최재술·연승준(2021)은 넷플릭스와 웨이브 앱 리뷰 데이터에 토픽모델링을 적용해 서비스 만족도 요인을 도출 후, IS 성공 모형(Information system success model)을 이용해 분류된 품질 속성이 서비스 만족도(별점 평가)에 미치는 영향 정도를 파악하였다[17]. 기존의 연구들은 다양한 텍스트 마이닝 기법을 활용해 서비스 만족 요인에 대한 분석을 수행했지만, 주로 만족도 요인에 대한 파악에 그쳤으며 이를 서비스 기회로 발전시킨 연구는 여전히 제한적이다. 이러한 문제 인식을 바탕으로 본 논문에서는 기존 방법론을 확장해 구체적인 서비스 개선안 제시하고자 한다.

3. 연구 방법

3.1 연구 프레임워크

본 연구에서는 Fig. 1의 연구 프레임워크에 명시된 단계적 절차를 따라 분석을 수행하였다. 연구 프레임워크는 총 4단계로 구성되며, 첫째로 파이션 기반의 “google-play-scraper” 라이브러리를

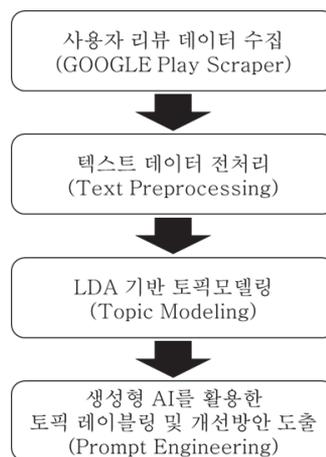


Fig. 1 Overview of the research framework

활용하여 구글 플레이 스토어에 등록된 국내 OTT 앱 중 국내 이용자 순위 기준 상위 3개에 해당하는 앱(넷플릭스, 쿠팡플레이, TVING)의 리뷰 데이터를 수집하였다. 둘째로는 수집한 리뷰 중 부정 리뷰만을 활용하여 텍스트 마이닝 기법 적용을 위한 텍스트 변환, 커스텀 단어 사전 생성, 명사 추출, 불용어 처리와 같은 텍스트 전처리 작업을 진행하였다. 세 번째 단계에서는 LDA 기반 토픽모델링을 적용해 불만족 요인에 해당하는 토픽을 도출했다. 마지막으로는 생성형 AI인 OpenAI사의 API를 활용하여 도출된 토픽의 주제 선정과 해당 토픽 내용을 토대로 앱 서비스에 대한 개선방안을 제시하였다. 아래에서는 단계별 분석 절차에 대해 보다 상세한 설명을 제시한다.

3.2 연구 데이터 수집 및 전처리

본 연구는 구글 플레이스토어에 등록된 OTT 서비스 중 국내 이용자 순위 기준 상위 3사(넷플릭스, 쿠팡플레이, 티빙)의 앱을 분석 대상으로 선정했다. 이를 위해 구글 플레이스토어에 등록된 각각의 앱 리뷰데이터를 “google-play-scraper” 라이브러리를 활용해 수집하였다. 리뷰에는 리뷰 작성자, 리뷰 코멘트, 별점 평가, 리뷰 작성 일시, 리뷰 코멘트에 대한 개발자 답변 등의 정보가 포함되는데, 본 연구에서는 그 중 한국어로 작성된 리뷰 코멘트, 별점 평가(5점 척도), 리뷰 작성 일시만을 추출해 분석에 사용하였다. 연구의 최종 목표가 리뷰 분석을 통한 서비스 개선안 제시이기 때문에, 유의미한 오피니언을 얻기 힘들다고 판단된 40자 미만의 단답형 리뷰는 원본 데이터에서 제거하였다. 나아가 리뷰 코멘트의 작성일이 국내 정식 앱 출시일 이전인 데이터들 역시 제거하였다. 이처럼 데이터에 대한 1차 필터링을 거친 결과, 넷플릭스로부터는 국내 정식 출시일인 2016년

7월 1일부터 2023년 12월 31일까지 총 8,615개의 리뷰 데이터를 수집하였고, 쿠팡플레이로부터는 2020년 12월 23일부터 2023년 12월 31일까지 총 3,410개의 데이터를 확인할 수 있었다. 마지막으로 티빙으로부터는 2014년 9월 27일부터 2023년 12월 31일까지 총 8,651개에 대한 리뷰를 획득하였다.

나아가 리뷰 기반의 불만족 요인 분석을 위해 극단적 평점에(1점) 해당되는 부정적 리뷰만을 별도로 구분해 분석을 수행하였다. 이는 부정 리뷰에 긍정적인 리뷰보다 서비스 개선에 도움이 되는 정보가 더 상세하게 포함되는 경향이 있기 때문이다. 그 결과 총 20,676개의 리뷰데이터가 텍스트 전처리를 위한 다음 단계에 활용되었다(Table 1).

Table 1. Information on collected review data

구간	리뷰 수집기간	총 리뷰 수 (부정적 리뷰 수)	비고
Netflix	2016년 7월 14일- 2023년 12월 31일	8,615 (4,375)	40자 미만 리뷰 데이터 제외
쿠팡 플레이	2020년 12월 14일- 2023년 12월 31일	3,410 (1,286)	
티빙	2014년 9월 27일- 2023년 12월 31일	8,651 (3,717)	
합계		20,677 (9,378)	

3.3 텍스트 데이터 전처리

리뷰 데이터는 사용자들이 직접 겪은 경험을 토대로 작성된 텍스트 데이터로 구성된다. 텍스트 데이터는 비정형 데이터의 한 형태로, 다음과 같은 전처리 과정을 거쳐 분석에 적합한 형태로 가공된다. 먼저 정규 표현식을 활용하여 불필요한 특수문자와 기호 등을 제거하는 작업을 실시하였다. 그 후 한국어 형태소 분석을 진행하기 전

‘UI’, ‘VOD’, ‘App’과 같이 영문으로 작성된 단어들을 ‘유저인터페이스’, ‘주문형비디오’, ‘앱’과 같은 한글 단어로 변환하였다. 아울러, OTT 플랫폼 어플리케이션 리뷰의 특성상 각각의 플랫폼에서 서비스되는 다양한 프로그램의 명칭과 축약어 등의 인지를 위해, 이들을 사용자 정의 사전에 수동으로 추가하여 단어인식의 정확도를 향상시켰다. 다음으로는 파이썬 기반의 한국어 정보처리패키지인 KoNLPy의 Okt(Open Korea Text) 형태소 분석기를 활용해 명사를 추출하였다. 이는 명사의 활용이 문서의 핵심적인 특성을 도출하고 해석하는 데 유용한 역할을 하기 때문이다. 불용어 처리 단계에서는 단어의 등장 빈도가 5회 미만인 단어와 문맥상 중요도가 낮고 오히려 다른 단어의 의미를 희석시킬 수 있는 단어들을 불용어로 지정해 제거하였다. 또한, 바이그램(bi-gram) 추출을 통해 문서 내 의미적으로 연관된 어휘 조합을 찾는 것으로 전처리를 마무리하였다.

3.4 LDA 기반 토픽모델링

본 연구에서는 OTT 플랫폼 리뷰데이터 분석을 위해 토픽모델링을 적용하였다. 토픽모델링은 문서 집합에 내재 되어 있는 주제를 추론하는 데 사용이 되며, 특히 대량의 텍스트 정보를 자동으로 체계화, 이해 및 요약하는 수단으로 활용된다. 토픽모델링은 각 문서를 여러 토픽의 복잡한 조합으로, 각 토픽을 다시 여러 단어의 복잡한 조합으로 표현하며, 토픽 추론 결과를 기반으로 문서를 분류하는 텍스트 마이닝 도구로 여겨진다[18]. 본 연구는 여러 토픽모델링 기법 중 LDA(Latent Dirichlet Allocation) 알고리즘을 적용하였는데, 이는 2003년 개발된 알고리즘으로 일반적으로 LSA(Latent Semantic Analysis) 기법 대비 더 독립적인 주제의 추출이 가능하다[19]. LDA는 확률분포에 기반

하는 기법으로 문서 내 주제에 걸쳐 단어들이 어떻게 분포되어 있는지를 확률적으로 모델링 하는 방식을 따른다.

토픽모델링은 일종의 비지도 머신러닝의 한 기법이지만, 분석을 수행함에 있어 사용자는 다양한 하이퍼파라미터의 값(토픽의 수, 반복횟수 등)을 조절함으로써 결과값에 큰 영향을 미칠 수 있다. LDA에 기반한 토픽모델링은 보통 지정된 수의 토픽을 반환하며, 최적의 토픽 개수를 결정하기 위해 높은 일관성(coherence) 점수 또는 낮은 혼잡도(perplexity)를 사용할 수 있다. 본 연구에서는 토픽 내 단어들이 얼마나 서로 의미적으로 관련성이 높은지를 나타내는 일관성 점수를 활용하였다. 토픽모델링은 파이썬 기반의 gensim 라이브러리 버전 4.3.2를 활용하여 수행하였다.

3.5 생성형 AI를 활용한 토픽 레이블링 및 개선방안 도출

프롬프트 엔지니어링은 최근 주목받고 있는 새로운 연구 분야로, LLM의 답변을 유도하고 생성된 출력값의 정확성, 관련성 및 일관성을 보장하기 위한 입력 프롬프트의 체계적인 설계 및 최적화를 의미한다[20]. 여기서 프롬프트는 생성형 AI의 출력물의 질적 수준을 보장받기 위해 LLM에 제공되는 지침(instruction)이다. 사용자는 프롬프트 엔지니어링을 통해 LLM의 출력 길이, 복잡성 및 스타일에 대한 제어가 가능하며, 특히 의학 분야에서는 프롬프트 엔지니어링을 적용함으로써 관련 연구, 교육 및 임상 진료에 새로운 영감을 얻고 있다[21]. 프롬프트 설계에는 zero-shot learning 또는 few-shot learning과 같은 기법의 적용이 가능하며, 본 연구에서는 structured prompting (또는 contextual prompting)을 사용해 토픽모델링 결과에 대한 레이블링 및 토픽 해석을 토대로

OTT 앱 서비스 개선방안에 대한 시사점을 도출하고자 한다. Structured prompting은 LLM에 추가적인 문맥정보, 요청사항, LLM 수행할 작업 방식 및 출력 형식 등에 대한 정보를 바탕으로 프롬프트를 구성하는 것을 의미한다[22].

4. 연구 결과

4.1 기술통계 분석

1차 필터링을 거치고 남은 리뷰데이터는 총 20,677개이며, 그중 9,378개의 리뷰가 극단적인 평가를 받은 부정 리뷰이다. Table 2는 OTT 서비스 앱별 별점 평가 분포를 나타낸다. 3사의 평균 평점은 2.50점으로 대체로 부정적으로 나타났으며, 이는 서비스 개선에 대한 요소가 존재함을

Table 2. Number of App Reviews by Rating

구분	별점 평가 분포					평균 평점
	매우 부정	부정	중립	긍정	매우 긍정	
	1	2	3	4	5	
Netflix	4,375	972	1,271	1,014	983	2.22
쿠팡플레이	1,286	376	582	477	689	2.68
티빙	3,717	881	1,122	891	2,040	2.61
합계	9,378	2,229	2,975	2,382	3,712	2.50

Table 3. Trend Development of Negative Reviews by Year

구분	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Netflix	-	-	22	281	439	716	1,181	983	487	266
쿠팡플레이	-	-	-	-	-	-	20	545	409	312
티빙	39	156	268	264	427	448	546	815	469	285
합계	39	156	290	545	866	1,164	1,747	2,343	1,365	863

간접적으로 시사한다. 부정 리뷰의 평균 길이는 97, 가장 긴 부정 리뷰의 길이는 648이다.

Table 3은 연도별 부정적 리뷰(1점)의 추세변화를 나타낸다. 그 결과 넷플릭스는 2020년에 직전년도 대비 부정 리뷰의 수가 약 64.9%(465개) 증가하였으며, 쿠팡플레이는 출시 직후인 2021년에 직전년도 대비 부정 리뷰의 수가 큰 폭으로 증가하였다. 다만, 쿠팡플레이의 경우 2020년 말에 서비스를 시작한 만큼, 운영 기간이 짧아 부정 리뷰의 증가율에 큰 의미를 부여하기 힘들다. 티빙의 경우 2021년도에 전년도 대비 부정 리뷰의 수가 약 49.3%(269개) 증가했는데, 이는 코로나19로 인해 OTT 서비스 이용자의 수가 급증하면서 앱 사용자의 부정적 리뷰 수 또한 증가한 것으로 보인다. 그 이후 부정적 리뷰는 감소하는 추세로 돌아섰다.

4.2 LDA 기반 토픽모델링

분석된 OTT 앱 서비스 간 차이점을 파악하기 위해 LDA 기반 토픽모델링을 적용하였다. 토픽모델링 기법 적용에 있어 중요한 과제 중 하나는 최적의 토픽 개수를 정하는 것이다. 본 연구에서는 2.4장에서 언급한 바와 같이 일관성 점수를 활용해 최적의 토픽 개수를 선택하였다. Table 4는 실험을 통해 도출된 토픽의 개수와 이에 해당하는 일관성 점수를 정리해 나타낸다. 일관성 점수는 0에서 1 사이의 값을 가지며, 일반적으로 일관성 점수가 높을수록 의미 있는 주제가 도출되었음을 시사한다. 그러나 일관성 점수가 1에 가까울수록 너

Table 4. Determination of the optimal topic number

그룹	토픽 수	Coherence score
Netflix	5	0.372
쿠팡플레이	9	0.323
티빙	5	0.393

무 일반적인 주제 또는 중복성 있는 주제가 도출이 될 가능성이 커진다. 따라서, 적절한 trade-off를 통해 유의미한 주제를 얻을 수 있도록 해야 한다. 넷플릭스와 티빙의 경우 각 5개에 해당하는 주제, 그리고 쿠팡의 경우 9개에 해당하는 주제가 도출되었다. 도출된 토픽별 키워드는 Table 5, 6 그리고 7에 정리하였다. 이렇게 도출된 토픽은 부정 리뷰를 바탕으로 하기 때문에 사용자가 느끼는 불만족 요인으로 해석할 수 있다.

Table 5. Topic modelling results for Netflix

구분	토픽 키워드
Topic 1	자막, 화면, 영상, 문제, 재생, 해결, 화질, 계속, 갑자기, 오류, 사용, 개선, 기능, 모바일, 업데이트_이후, 현상, 서비스, 영상_재생, 설정, 업데이트
Topic 2	결제, 환불, 가입, 해지, 접속, 무료, 확인, 구독_취소, 자동_결제, 구독, 계정, 로그인, 계속, 무료_이용, 오늘, 이용, 오류, 메일, 취소, 문자
Topic 3	콘텐츠, 영화, 시청, 한국, 검색, 화면, 별로, 기능, 영상, 드라마, 재생, 성인_인증, 계속, 미드, 프로그램, 영어, 미국, 시즌, 시리즈, 시작
Topic 4	자막, 서비스, 인증, 한글_자막, 성인_인증, 언제, 최악, 생각, 결제, 업데이트, 이용, 연령_인증, 무한_로딩, 이해, 유저인터페이스, 광고, 쓰레기, 카드, 설정, 해외
Topic 5	오류, 실행, 계속, 업데이트, 사용, 로딩, 로그인, 설치, 기기, 문제, 저장, 접속, 갑자기, 해결, 재생, 구글플레이, 핸드폰, 데이터, 동영상, 시도

Table 6. Topic modeling results for Coupang Play

구분	토픽 키워드
Topic 1	다운로드, 오류, 재생, 다운, 삭제, 콘텐츠, 드라마, 사용, 페이지, 영화, 개선, 전혀, 최신, 쿠팡, 계속, 설정, 알림, 시도, 화면, 며칠
Topic 2	문제, 실행, 계속, 설치, 다운, 연결, 티비, 안드로이드, 기기, 재생, 스트리밍, 다운로드, 삭제, 사용, 넷플릭스, 오류, 업뎃, 업데이트, 오늘, 쿠팡플레이
Topic 3	기능, 자막, 화면, 영화, 설정, 자체, 기본, 넷플릭스, 핸드폰, 영상, 화면_회전, 업데이트, 버튼, 드라마, 서비스, 데이터, 사용, 이용, 와이파이, 방향

Table 6. (Continued)

구분	토픽 키워드
Topic 4	재생, 티브이, 넷플릭스, 루팅, 쿠팡플레이, 답변, 결제, 회차, 쿠팡, 확인, 개선, 중계, 콘텐츠, 미러링, 검색, 버튼, 이용, 축구_중계, 쿠팡, 회
Topic 5	업데이트, 문제, 영화, 자막, 화면, 사용, 쿠팡플레이, 언제, 오류, 기기_호환, 노트, 시즌, 가격, 갤럭시, 소리, 해결, 설치, 버그, 쿠팡, 화면_소리
Topic 6	버퍼링, 가입, 계속, 삭제, 개선, 로그인_자체, 중계, 쿠팡, 플레이, 짜증, 생중계, 로그인_로그인, 멤버십, 시리즈, 시청, 방송, 손흥민_경기, 광고, 콘텐츠, 무료_체험
Topic 7	로그인, 축구, 접속, 결제, 경기, 시간, 서버, 가입, 콘텐츠, 중계, 쿠팡, 계속, 유료, 쿠팡플레이, 환불, 로딩, 최악, 무한_반복, 서비스, 토트넘
Topic 8	화면, 로그인, 계속, 설치, 쿠팡, 사용, 회원, 갑자기, 쿠팡플레이, 인증, 해결, 에러, 삭제_설치, 아이디, 성인_인증, 뭔가, 플레이, 새로고침, 정작, 쿠팡_로그인
Topic 9	영상, 재생, 화면, 시청, 문제, 다운로드, 다운, 자막, 오티티, 개선, 쿠팡플레이, 기능, 태블릿, 영화, 렉, 콘텐츠, 소리, 티빙, 갑자기, 현상

Table 7. Topic modeling results for TVING

구분	토픽 키워드
Topic 1	결제, 환불, 답변, 이용권, 문의, 해지, 티빙, 서비스, 계속, 보고, 정기_결제, 고객센터, 취소, 확인, 오류, 문제, 시간, 재생, 자동_결제, 해결
Topic 2	업데이트, 로그인, 오류, 설치, 사용, 버전, 화면, 조치, 계속, 접속, 티빙, 시청, 가입, 실행, 갑자기, 어제, 해도, 태블릿, 창, 점검
Topic 3	영상, 화면, 티빙, 재생, 사용, 계속, 문제, 개선, 와이파이, 실행, 로딩, 영상_재생, 종료, 연결, 정도, 데이터, 갑자기, 이용, 기능, 해결
Topic 4	무료, 광고, 채널, 방송, 티빙, 유료, 계속, 로그인, 라이브, 실시간, 다시보기, 드라마, 회원_가입, 티브이, 업뎃, 삭제, 시간, 구매, 결제, 갑자기
Topic 5	티빙, 버퍼링, 콘텐츠, 영화, 이용, 결제, 최악, 크롬캐스트, 넷플릭스, 티비, 생각, 사용, 오티티, 프로그램, 화질, 시청, 드라마, 개선, 서비스, 보고

4.3 생성형 AI를 활용한 토픽 레이블링

토픽 레이블링은 연구자가 주관적인 판단을 통해 도출된 토픽이 어떤 주제나 내용을 대표하는지

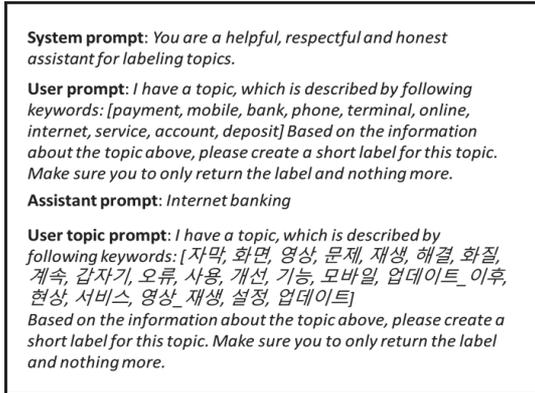


Fig. 2 Example prompt for labeling of topics

Table 8. Topic labeling results using generative AI

구분	Topic	주제
Netflix	1	영상 재생 및 화질 문제
	2	결제 및 구독 서비스 이슈
	3	콘텐츠 선택과 관련한 불만
	4	자막 및 인증 서비스 문제
	5	앱 실행 및 로딩 이슈
쿠팡플레이	1	다운로드 및 콘텐츠 이용
	2	앱 설치 및 연결 문제
	3	영상 및 자막 설정 기능
	4	콘텐츠 재생 및 검색
	5	TV 및 다른 기기와의 연동 문제
	6	로그인 및 멤버십
	7	스포츠 중계 및 결제 이슈
	8	로그인 및 앱 설치
	9	영상 재생 및 다운로드
티빙	1	결제 및 환불 관련 문제
	2	로그인 및 업데이트 오류
	3	영상 재생 및 연결 문제
	4	광고 및 콘텐츠 이용권 관련 문제
	5	콘텐츠 품질 및 버퍼링 문제

이해하기 쉽도록 지정하는 절차를 의미한다. 본 연구에서는 연구자의 주관적인 판단 대신 생성형 AI 모델(GPT-4)을 활용한 자동화된 토픽 레이블링을 수행하였다. 이를 위해 시스템 프롬프트를 통해 AI에 역할을 부여하고 토픽별 도출된 키워드를 컨텍스트(context)로 제시하여 토픽 레이블링을 수행하도록 프롬프트를 구성하였다. 프롬프트의 예시는 Fig. 2와 같다.

도출된 토픽 레이블명은 Table 8에 나열하였다. 그 결과 분석한 모든 OTT 앱에서 “영상 재생 오류”, “앱 설치 및 업데이트 오류”, “구독 및 결제”, 그리고 “양질의 콘텐츠 제공”과 관련 문제점들이 공통된 불만 요인으로 나타났다. 특히, 넷플릭스와 티빙의 경우 콘텐츠의 선택과 품질에 관한 이슈가 불만 요인으로 작용했다면, 쿠팡플레이의 경우 다른 기기와의 호환성 및 스포츠 콘텐츠 관련 불만족 사항이 존재했다. 실제로 콘텐츠 선택의 다양성과 관련된 이슈는 OTT 서비스 제공업체들이 공통으로 직면한 문제로, 이러한 상황을 타개하기 위해 티빙은 웨이브와의 합병을 논의 중으로 알려졌다. 이들은 합병을 통해 더욱 원활한 콘텐츠의 공급과 규모의 경제 달성을 통해 수익성을 개선할 수 있을 것으로 기대한다[23]. 종합적으로 생성형 AI를 활용한 토픽 레이블링 결과는 연구자의 주관적인 판단에 의한 결과와 크게 다르지 않음을 전문가와의 협의를 통해 확인할 수 있었다. 전문가로는 토픽모델링 관련 연구 경험을 지닌 연구자와 콘텐츠 관련 분야 종사자 2명을 선정하였고, 생성형 AI를 통해 얻은 토픽 레이블링의 해석에 대해 5점 척도로 동의 또는 비동의 여부를 조사하였다. 그 결과 넷플릭스의 사례를 제외하고는 평균이 4.4로 전반적으로 획득한 레이블링에 대한 동의도가 높음을 확인할 수 있었다. 넷플릭스의 평균은 3.8로 중간 이상으로 높게 나타났다.

4.4 생성형 AI를 활용한 개선방안 도출

본 장에서는 앞서 도출된 토픽을 기반으로 생성형 AI를 활용하여 토픽별로 3가지 서비스 개선 방안을 도출토록 하였다. 적용한 프롬프트의 예시는 Fig. 3과 같다.

도출된 서비스 개선방안은 Table 9, 10, 11에 정리하였다.

System prompt: You are a helpful, respectful and honest assistant for providing service improvement suggestions.

User prompt: I identified following customer pain point from the analysis of user review: 'Subscription service support'. Based on the given dissatisfaction factor, please suggest service improvement plans. Provide only three improvement suggestions without any additional explanations.

Fig. 3 Example prompt for deriving service improvement suggestions

Table 9. Suggestions for service improvement in case of Netflix

주제	서비스 개선방안
영상 재생 및 화질 문제	1. 화질 및 재생 안정성 최적화 2. 영상 로딩 속도 개선 3. 사용자 화질 설정 기능 강화
결제 및 구독 서비스 이슈	1. 결제 및 환불 프로세스 간소화 2. 구독 및 해지 관리 기능 개선 3. 자동결제 알림 기능 강화
콘텐츠 선택과 관련한 불만	1. 다양한 콘텐츠 확보 및 업데이트 2. 검색 기능 및 결과 개선 3. 사용자 맞춤 콘텐츠 추천 기능 강화
자막 및 인증 서비스 문제	1. 자막 품질 및 다양성 증대 2. 인증 절차 간소화 및 안정성 강화 3. 사용자 친화적인 UI/UX 디자인 개선
앱 실행 및 로딩 이슈	1. 앱 로딩속도 및 안정성 개선 2. 로그인 및 계정 연동 문제 해결 3. 앱 사용 중 발생하는 오류 수정

Table 10. Suggestions for service improvement in case of Coupang Play

주제	서비스 개선방안
다운로드 및 콘텐츠 이용	1. 콘텐츠 다운로드 속도 개선 2. 콘텐츠 업데이트 빈도 증가 3. 콘텐츠 제공 다양화

Table 10. (Continued)

주제	서비스 개선방안
앱 설치 및 연결 문제	1. 앱 설치 및 실행 안정성 향상 2. 스트리밍 연결 안정성 개선 3. 기기 호환 범위 확장
기능 및 화면 설정	1. 사용자 친화적인 화면 및 기능 설정 2. 자막 설정 개선 3. 화면 회전 기능 개선
콘텐츠 재생 및 검색	1. 콘텐츠 재생 안정성 개선 2. 콘텐츠 검색 기능 개선 3. 빠른 콘텐츠 업데이트
TV 및 다른 기기와의 연동 문제	1. TV 및 다른 기기와의 연동 개선 2. 루팅, 미러링 문제 해결 3. 중계 콘텐츠 품질 개선
로그인 및 멤버십	1. 로그인 프로세스 간소화 2. 멤버십 혜택 강화 3. 로그인 문제 해결에 대한 솔루션 제공
스포츠 중계 및 결제 이슈	1. 스포츠 중계 안정성 개선 2. 결제 절차 간소화 및 안정성 강화 3. 경기 스케줄 안내 개선
로그인 및 앱 설치	1. 로그인 오류 수정 2. 앱 설치 문제 해결 3. 사용자 인증 프로세스 개선
영상 재생 및 다운로드	1. 영상 재생 품질 개선 2. 다운로드 속도 개선 3. 자막 및 더빙 서비스 개선

Table 11. Suggestions for service improvement in case of TVING

주제	서비스 개선방안
결제 및 환불 관련 문제	1. 결제 시스템에 대한 안정성 개선 2. 환불 관련 고객 응대 개선 3. 문의 및 이용권 해지 절차 단순화
로그인 및 업데이트 오류	1. 앱 버전 호환성 개선 2. 로그인 절차 간소화 3. 앱 오류 수정 및 업데이트 안정화
영상 재생 및 연결 문제	1. 영상 재생 품질 개선 2. 와이파이 및 데이터 연결 안정성 강화 3. 로딩 속도 및 버퍼링 개선
광고 및 콘텐츠 이용권 관련 문제	1. 광고 시간 및 빈도 조정 2. 유료/무료 콘텐츠 구분 및 공지 개선 3. 로그인 및 회원 가입 절차 단순화
콘텐츠 품질 및 버퍼링 문제	1. 영상 및 사운드 품질 개선 2. 콘텐츠 다양성 확보 3. 크롬캐스트 및 기타 디바이스와의 호환성 개선

그 결과 “콘텐츠 품질 개선 및 다양화”, “영상 화질 품질과 데이터 사용량 최적화”, “외부기기 호환성”, “결제 및 해지”와 같은 개선방안이 분석 대상 앱 모두에서 공통으로 도출되었다. 이는 OTT 서비스 이용자에게 양질의 콘텐츠 제공이 서비스의 질을 결정짓는 중요한 요인으로 기인한다는 추정이 가능하다. 특히, 콘텐츠의 다양화뿐만 아니라 업데이트 주기에 대한 개선과 추천 알고리즘의 개선을 통해 사용자 경험을 향상시킬 수 있을 것으로 기대한다. 나아가 사용자가 스마트폰, 태블릿PC, 스마트TV 등과 같은 다양한 기기를 통해 서비스를 이용하기 때문에, 개별적인 사용환경에 따른 영상 품질 등이 중요한 역할을 함을 파악하였고, 이를 위해서는 타 기기와의 호환성 및 연결 안전성 강화를 위한 개발 노력이 요구됨을 인지할 수 있었다. 그 외 구독 결제 및 해지 과정에서의 절차 단순화 등에 대한 방안도 새로운 서비스 개선 기회로 도출되었다.

각 서비스를 개별적으로 살펴보면, 넷플릭스는 여러 국가에서 제작한 다양한 콘텐츠를 제공하는 점을 특징으로 볼 수 있는데, 이와 연계하여 “자막 품질 및 다양성 증대”, “사용자 맞춤 콘텐츠 추천 기능 강화” 등 글로벌 서비스 콘텐츠에 대한 서비스 개선방안이 도출되었다. 쿠팡플레이의 경우 스포츠 중계권 확보를 통한 이용자 확보 전략^[24]으로 인해 “경기 스케줄 안내 개선”, “스포츠 중계 안정성 개선” 등과 같은 서비스가 개선 방안으로 도출되었으며, 이는 정확한 중계 일정 안내에 대한 요구와 중계의 품질에 관한 이슈가 중요함을 시사한다. 또한, 다른 기기와의 호환성에 대하여 “루팅, 미러링 문제 해결” 등과 같은 개선방안이 도출된 것으로 볼 때 스포츠 중계 콘텐츠의 경우 스마트폰과 같은 환경보다는 TV 등과 연결하여 큰 화면으로 시청이 선호됨을 알 수 있다. 티빙의 경우 “유료/무료 콘텐츠 구분 및 공지

개선”, “콘텐츠 다양성 확보”와 같은 콘텐츠 관련 서비스 개선방안이 제시되었는데, 만약 티빙과 웨이브의 합병이 완료되어 통합적인 서비스가 이루어진다면 국내 이용자를 위한 양질의 콘텐츠가 확보되어 서비스의 품질이 향상될 것으로 기대된다.

5. 결론

본 연구에서는 OTT 앱 사용자 리뷰 중 부정 리뷰의 분석을 중심으로 서비스 개선안을 제시하고자 했다. 그 결과 넷플릭스와 티빙으로부터 총 5개의 불만족 요인이 도출되었고, 쿠팡플레이로부터는 총 9개에 달하는 불만족 요인을 파악하였다. 공통된 불만족 요인으로는 “영상 재생 오류”, “앱 설치 및 업데이트 오류”, “구독 및 결제”, “양질의 콘텐츠 제공”이라 할 수 있다. 이는 OTT 앱 서비스가 갖추어야 할 기본적인 서비스 경쟁력 강화방안으로 해석될 수 있다. 또한, 앱의 특성에 따라 도출된 개별 서비스 개선방안 역시 해당 OTT 앱의 이용자 확보전략과 대체로 부합한다. 넷플릭스의 글로벌 서비스 콘텐츠에 대한 자막 제공 서비스 강화, 쿠팡플레이의 실시간 국내·외 스포츠 중계에 대한 스케줄 안내 및 중계 안정성 강화, 티빙의 유/무료 콘텐츠 구분 및 공지 개선 등과 같은 개선 기회는 유저들이 해당 서비스를 이용하는 강력한 동기 중 하나로 사용자의 이탈을 막기 위해 개별적으로 개선해야 할 부분이다.

본 연구의 주요 기여는 다음과 같다. 기존 토픽 모델링의 경우 연구자가 키워드를 해석해 주관적인 관점에서 주제를 도출하는 방식을 따랐다면, 본 연구에서는 이를 생성형 AI를 활용해 자동화시킴으로써 대용량 문서에 대한 레이블링을 더 용이하게 해주는 연구 프레임워크를 제시하는 데 그 의의를 갖는다. 나아가 대다수의 토픽모델링 관련

연구들이 도출된 주제를 기반으로 제품이나 서비스의 강·약점에 대한 함의만을 파악하고자 하였다면, 본 연구의 결과는 구체적인 action plan으로 이어질 수 있는 구체적 개선방안을 제공함으로써 실무적으로도 실제 서비스 개선에 활용될 수 있음을 시사한다.

본 연구의 한계점으로는 다음과 같은 사항들이 있다. 먼저 수집된 사용자 리뷰 중 1점의 극단적 평점을 부여받은 리뷰를 부정 리뷰로 가정하였고, 40글자 미만의 리뷰에서는 유의미한 의미 도출이 어려울 것으로 여겨 분석에 포함하지 않았다. 그러나 별점 점수와 무관하게 서비스 개선에 대한 아이디어는 긍정적 리뷰 또는 중립적인 리뷰에서도 획득이 가능할 수 있기에, 후속 연구에서는 수집한 모든 리뷰 데이터에 대한 감성분석을 사전에 적용해 부정 리뷰를 분류하고, 이후 본 연구에서 제시한 분석 기법을 적용한다면 더 정확한 서비스 개선 기회를 발굴할 수 있을 것으로 기대된다.

또한, 도출된 토픽에 대해 생성형 AI인 GPT-4 모델을 적용해 레이블링 및 개선안 도출을 시도하였는데, 생성형 AI의 특성상 활용한 모델에 따라 답변의 질이 달라질 수 있다. 따라서 향후 연구에서는 더 다양한 방식의 프롬프트 개선에 대한 실험과 더불어, 여러 모델의 결과값 비교를 통해 더 폭넓은 시선에서 서비스 개선방안을 도출하고자 하는 노력이 요구된다. 이를 위해 최근 좋은 성능을 보이는 LLama 2 또는 Mistral 모델 등을 후속 연구에 적용해볼 수 있다. 특히, 생성형 AI는 방대한 양의 데이터를 기반으로 학습한 후, 새로운 답변을 생성한다는 특징이 있다. 따라서, 상대적으로 한국어 데이터에 특화된 LLM을 적용한다면, 보다 리뷰가 내포하고 있는 문맥적 특성을 효과적으로 이해해 한국에서 서비스되는 OTT 앱에 대한 개선전략을 보다 정확히 도출할 수 있을 것으로 기대한다.

사 사

본 논문은 산업통상자원부의 ‘융합기술사업화 확산형 전문인력 양성사업’의 지원을 받아 수행된 논문임.

참고문헌

- [1] 한국콘텐츠진흥원, 2023, “2023년 상반기 콘텐츠산업 동향분석 보고서”, <<https://welcon.kocca.kr/ko/info/trend/1953540>>, (electronic article), viewed 15 Feb (2024).
- [2] 문화체육관광부, 한국콘텐츠진흥원, 2023, “콘텐츠산업 2023결산 2024 전망 세미나 자료집”, <<https://www.kocca.kr/shortUrl/vh8p>>, (electronic article), viewed 15 Feb (2024).
- [3] 한국수출입은행, 2021, “OTT산업과 K콘텐츠 수출 : K드라마, K무비를 중심으로”, <<https://keri.koreaexim.go.kr/HPHFOE052M01/63856?currentPage=2>>, (electronic article), viewed 16 Feb (2024).
- [4] 정용찬, 김윤화, “2022년도 방송매체 이용형태 조사”, 방송통신위원회, p.126-131, (2022).
- [5] 한국콘텐츠진흥원, 2022, “콘텐츠 중심의 OTT 산업 육성방향”, <<https://www.kocca.kr/shortUrl/8v7p>>, (electronic article), viewed 15 Feb (2024).
- [6] 한국수출입은행, 2023, “영화·OTT산업 위기론과 투자” <<https://keri.koreaexim.go.kr/HPHFOE052M01/105984?currentPage=1>>, (electronic article), viewed 16 Feb (2024).
- [7] 한국콘텐츠진흥원, 2023, 글로벌 OTT 동향분석(Vol.3), “영화에 비춰보는 OTT 투자”, <<https://www.kocca.kr/globalOTT/vol03/html/main.html>>, (electronic article), viewed 16 Feb (2024).
- [8] 한국경제, “[단독] 티빙·웨이브 합병... 단숨에 토종 1위 OTT”, <<https://www.hankyung.com/article/2023112924491>>, viewed 7 Feb (2024).
- [9] 유지훈, 박주연, 글로벌 OTT 서비스 이용자의 지속적 이용 의도에 미치는 요인 연구: 넷플

- 릭스 사례를 중심으로, 방송통신연구, 102, p.46-79, (2018).
- [10] 변은지, 백현미, “OTT 서비스 복수 구독 현황 및 이용자 특성에 관한 탐색적 연구”, 방송통신연구, 117, p.43-69, (2021).
- [11] 곽승준, 최은철, 국내 OTT 서비스 산업 현황 및 수요 분석, *Journal of The Korean Data Analysis Society*, 22(6), p.2561-2572, (2020).
- [12] 이준영, 전범수, OTT 서비스 만족도와 지속이용의도 결정 요인 연구, *한국방송학보*, 34(4), p.116-144, (2020).
- [13] 이소율, 박현준, 국내 OTT (Over-The-Tops) 서비스 성공 요인에 관한 비교 연구, *한국산업정보학회논문지*, 26(6), p.135-154, (2021).
- [14] 조호수, 강성안, 류민호, 텍스트마이닝을 활용한 OTT 서비스 리뷰 분석 : 로컬 서비스 경쟁우위 확보 방안을 중심으로, *정보통신학회 논문지*, 46(4), p.722-733, (2021).
- [15] 이성준, 온라인 리뷰 분석을 통한 국내 OTT 서비스 이용자 만족 및 불만족 유발 요인 연구 : 어휘 분석 및 LDA 토픽 모델링 분석 방법을 중심으로, *커뮤니케이션학 연구*, 30(2), p.14-74, (2022).
- [16] 고민정, 이상원, 텍스트 마이닝을 활용한 코로나 19 전후 온라인 동영상 서비스 (OTT) 리뷰 비교분석 연구-정서 중심 대처와 노스텔지어를 중심으로, *한국콘텐츠학회논문지*, 21(11), p.375-388, (2021).
- [17] 최새술, 연승준, 앱 리뷰 분석을 통한 국내 OTT 서비스 만족 요인 분석 : 넷플릭스와 웨이브 사례 비교, *한국콘텐츠학회 종합학술대회 논문집*, p.373-374, (2021).
- [18] 강범일, 송민, 조화순, 토픽 모델링을 이용한 신문 자료의 오피니언 마이닝에 대한 연구, *한국 문헌정보학회지*, 47(4), p.315-334, (2013).
- [19] 김현구, 황종국, 황재승, LSA와 LDA를 이용한 풍공학회지 토픽모델링, *한국풍공학회 논문집*, 24(3), p.113-120, (2020).
- [20] 정천수, LLM 애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현: RAG모델과 LangChain 프레임워크 기반, *한국지능정보시스템학회*, 29(4), p.129-164, (2023).
- [21] Meskó, B., Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial, *Journal of Medical Internet Research*, 25(e50638), p.1-6, (2023).
- [22] Heston, T.F., Khun, C., Prompt Engineering in Medical Education, *International Medical Education*, 2(3), p.198-205, (2023).
- [23] 연합뉴스, “웹이브-티빙 합병 예고...불가피한 생존전략, 문제는 이후”, <<https://www.yna.co.kr/view/AKR20231130147600017>>, viewed 12 Mar(2024).
- [24] 매일경제, “쿠팡플레이, 스포츠 콘텐츠로 입지 다진다”, <<https://www.mk.co.kr/news/it/10937974#none>>, viewed 12 Mar (2024).