

사용자 경험과 서비스 평가의 변화에 관한 연구

- 넷플릭스 앱 리뷰 토픽 모델링을 통해

2024.04.11

연어 황서진

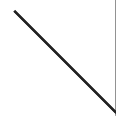
목차

- 1 연구 배경
- 2 데이터 수집 및 전처리
- 3 토픽모델링(LDA)
- 4 결론 및 한계
- 5 Q&A

01. 연구 배경

코로나19의 발생으로 OTT 서비스 이용 증가 . 특히, Netflix 1,600만 명의 가입자가 증가

⇒ OTT 이용이 증가하면서 사용자 경험도 증가 & OTT 서비스 간 경쟁도 치열해지면서 OTT 사용자 경험을 연구하는것이 중요



코로나19 대유행 전·후 넷플릭스 앱 사용자가 남긴 리뷰의 텍스트를 토픽 모델링으로 분석하여 사용자 경험 증가에 따른 서비스 평가 변화를 알아보고자 함

이를 통해 사용자들이 직접 표현한 의견을 분석하여 코로나19 대유행 전·후 넷플릭스 서비스의 사용자 경험 변화를 이해하고, 서비스 평가에 대한 통찰을 얻고자 함

02. 데이터 수집 및 전처리

LDA 방식을 사용하여 OTT 서비스에 대한 사용자들의 평가를 분석

1 데이터 수집

- 구글 플레이 스토어에서 '넷플릭스' 앱 리뷰 수집
- 영어권의 이용자가 가장 많은 플랫폼이므로 영어 리뷰만 수집
- 2018년 이후 전체적인 OTT 시청 시간 증가 → 2018년 1월부터 2022년 9월까지 558,592개의 데이터 수집

표 1. 코로나19 대유행 전후 데이터 수집 현황

	코로나19 대유행 이전	코로나19 대유행 이후
리뷰 수	195,515개	363,077개
리뷰 전체 평균 글자 수	790자	942자

*대유행 기준 : 2020년 03월
대유행 이전 : 2018/01 ~ 2020/02
대유행 이후 : 2020/03 ~ 2022/09

코로나19 대유행 이후
리뷰 수와 리뷰 전체 평균 글자 수 모두 늘어남

02. 데이터 수집 및 전처리

2 데이터 전처리

NLTK 라이브러리 사용

- 토큰화
 - 수집된 데이터들을 리스트 형태로 변환
 - 텍스트를 문장 단위로 나눔
 - 각 문장을 단어로 나누는 토큰화
 - 특수문자나 숫자 등을 제거, 소문자로 변환
- 불용어 제거
 - 'a', 'is', 'that' 등의 불용어 제거
 - 해당 서비스 이름인 'netflix'는 분석에 영향을 미치지 않는다고 판단되어 불용어 처리 목록에 추가하고 삭제
- 표제어 추출
 - 단어들의 명사, 형용사, 부사, 동사와 같이 품사를 지정한 후 원형으로 변환하는 작업

03. 토픽 모델링

토픽 모델링

- 다수의 텍스트로부터 토픽들을 추출하는 텍스트 마이닝 기법

텍스트 마이닝

- 비정형의 텍스트 데이터를 분석하여 드러나지 않는 의미를 찾고 다른 정보와의 관계성까지 파악할 수 있게 해주는 분석 기법

03. 토픽 모델링

LDA (Latent Dirichlet Allocation)

- 문서의 집합으로부터 어떤 토픽이 존재하는지를 알아내기 위한 토픽 모델링의 대표적인 알고리즘
- 3개의 베이지안 계층으로 구성된 주제 추출 모델이며 키워드가 집단으로 표현될 확률을 추정하여 나타냄
- 일반적으로 복잡도와 일관성 등의 지표를 사용하여 토픽의 개수를 정함
- LDA는 문서들은 토픽들의 혼합으로 구성되어져 있으며, 토픽들은 확률 분포에 기반하여 단어들을 생성한다고 가정
- 데이터가 주어지면 , LDA는 문서가 생성되던 과정을 역추적
- 빈도수 기반의 표현 방법인 BoW의 행렬 DTM 또는 TF-IDF 행렬을 입력으로 함 → 단어의 순서는 신경쓰지 않음

03. 토픽 모델링

LDA 수행하기

1. 사용자는 알고리즘에게 토픽의 개수 k 를 알려줌 → LDA는 토픽의 개수 k 를 입력받으면, k 개의 토픽이 M 개의 전체 문서에 걸쳐 분포되어 있다고 가정하게 됨
2. 모든 단어를 k 개 중 하나의 토픽에 랜덤하게 할당
3. 모든 문서의 모든 단어에 대해 단어 w 가 두 가지 기준인 $p(t|d)$, $p(w|t)$ 에 따라 토픽을 재할당, 이를 반복
 - a. 이 때 자신만이 잘못된 토픽에 할당되어 있고 다른 모든 단어는 올바른 토픽에 할당되어 있다고 가정
 - b. $p(t|d)$: 문서 d 의 단어들 중 토픽 t 에 해당하는 단어들의 비율
 - c. $p(w|t)$: 각 토픽들 t 에서 해당 단어 w 의 분포

03. 토픽 모델링

doc1					
word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2					
word	cute	book	king	apple	apple
topic	B	B	B	B	B

- $p(t|d)$: 문서 d의 단어들 중 토픽 t에 해당하는 단어들의 비율

doc1					
word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2					
word	cute	book	king	apple	apple
topic	B	B	B	B	B

1. doc1의 단어들이 어떤 토픽에 해당하는지 봄
 2. doc1의 모든 단어들을 토픽A와 토픽 B에 대해 50대 50의 비율로 할당되어 있음
- => 단어 apple은 토픽 A 또는 토픽 B 둘 중 어디에도 속할 가능성 있음

- $p(w|t)$: 각 토픽들 t에서 해당 단어 w의 분포

doc1					
word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2					
word	cute	book	king	apple	apple
topic	B	B	B	B	B

1. 단어 apple이 전체 문서에서 어떤 토픽에 할당되어 있는지 봄
=> 단어 apple은 토픽 B에 할당될 가능성이 높음

두 가지 기준을 참고하여 LDA는 doc1의 apple을 어떤 토픽에 할당할지 결정

03. 토픽 모델링

1

토픽 수 구하기

- 토픽 일관성 분석을 진행
- 토픽 일관성은 토픽 내의 단어들이 함께 등장하는 정도를 측정하여 토픽의 의미를 평가하는 지표를 말함
- 일관성 분석을 통해 각 토픽의 일관성 점수를 계산하고, 이를 기반으로 최적의 토픽 수를 결정
- 토픽 수가 4개 일 때 가장 높은 일관성

2

LDA

- 토픽 모델링 방법
- 코로나 대유행 전후 모두 토픽의 수를 4개로 선정하여 분석

3

토픽 추출

- 토픽은 넷플릭스 앱 기능 부분, 콘텐츠에 대한 평가, 서비스 이용, 총평의 4가지로 나타남
- 각 토픽을 워드클라우드로 시각화

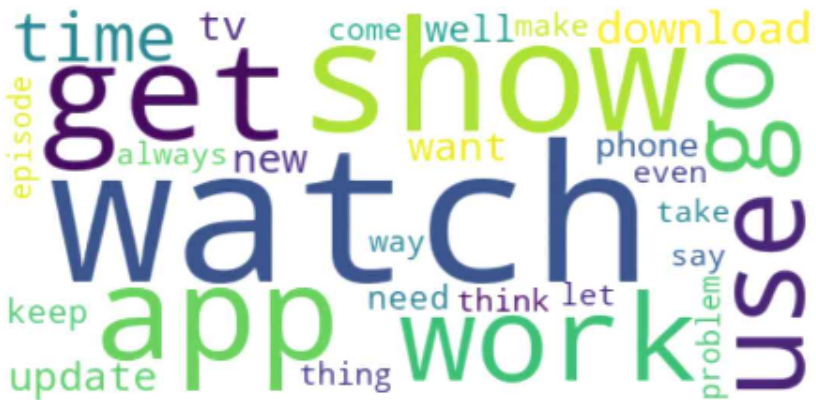

표 2. 각 토픽별 리뷰 개수

	코로나19 대유행 이전	코로나19 대유행 이후
Topic 1	90,914개	363,077개
Topic 2	48,488개	95,126개
Topic 3	31,869개	94,763개
Topic 4	24,244개	59,908개

03. 토픽 모델링

- Topic1
- 앱 자체 기능
 - 코로나19 대유행 전후 넷플릭스의 기능적인 면을 나타내는 키워드 모임

표 3. Topic 1의 코로나19 대유행 전후

넷플릭스 앱 기능	코로나19 대유행 전
	
	watch(0.050), show(0.032), get(0.023)
	코로나19 대유행 후
	
	excellent(0.146), super(0.043), watch(0.011)

- 대유행 이전
- watch, show, work, download 등 넷플릭스의 기능적인 면을 설명하는 단어들의 빈도가 높게 나옴
- 대유행 이후
- excellent, super, watch, dope과 같은 기능에 대한 긍정적인 감정을 나타내는 단어가 등장
 - quality, phone, streaming 등의 넷플릭스 기능의 이용에 대한 좀 더 세부적인 표현들이 나타남
 - 그 중 phone은 이용자의 휴대전화를 나타내는 단어였고 사용자와 이용량이 늘어나면서 구체적인 상황을 언급하며 전보다 자세한 평가를 남긴다고 해석할 수 있음

03. 토픽 모델링

표 4. Topic 2의 코로나19 대유행 전후

콘텐츠에 대한 평가	코로나19 대유행 전
	
	love(0.131), good(0.118), movie(0.066)
	코로나19 대유행 후
	
	good(0.241), awesome(0.064), amazing(0.043)

Topic 2

- 넷플릭스의 콘텐츠에 대한 키워드 집합

대유행 이전

- love, good, movie, amazing, favorite등의 단어들로 넷플릭스 콘텐츠에 대한 소감들을 나타내는 단어들이 등장

대유행 이후


- good, nice, awesome 등이 여전히 나타남
- original과 series는 넷플릭스에서 자체제작, 소수의 콘텐츠들을 칭하는 것인데 대유행 이전에는 인기가 많은 특정한 콘텐츠에 대해 언급했다면, 대유행 이후에는 콘텐츠 양이 늘어나면서 콘텐츠 자체에 대한 의견을 나타내는 양상으로 변화

03. 토픽 모델링

Topic 3

- 넷플릭스의 서비스 이용 측면에 관한 토픽

표 5. Topic 3의 코로나19 대유행 전후

서비스 이용	코로나19 대유행 전
	
	awesome(0.123), add(0.030), thank(0.020)
	코로나19 대유행 후
	
	superb(0.035), experience(0.033), gud(0.022)

대유행 이전

- awesome, add, thank, easy와 같이 서비스의 이용과 관련된 내용을 담고 있음

대유행 이후



- superb, experience, gud, service 등의 단어가 나타남
- 서비스를 나타내는 service라는 직접적인 단어의 언급이 증가
- bad, problem 등 서비스 이용에서의 오류 및 불편함을 나타내는 단어가 나타남
- 이를 통해 사용자와 이용자의 증가로 서비스에 오류 발생이 증가했고 사용자의 오류 경험도 증가하면서 서비스에 대해 직접적으로 언급하는 등 이용과 관련된 추가적인 내용을 기재한다고 판단할 수 있음

03. 토픽 모델링

Topic 4

- 넷플릭스에 대한 사용자들의 총평을 나타내는 키워드 모임

표 6. Topic 4의 코로나19 대유행 전후

총평	코로나19 대유행 전
	
	great(0.065), app(0.065), nice(0.043)
	코로나19 대유행 후
	
	great(0.119), love(0.102), app(0.070)

대유행 이전

- great, nice, excellent 등의 긍정적인 표현들의 빈도가 높음

대유행 이후

- great, love, nice 등 넷플릭스를 칭찬하는 뉘앙스의 단어들이 등장
- 사용자가 증가하여도 넷플릭스 전체에 대한 평가는 꾸준히 긍정적임

04. 결론 및 한계

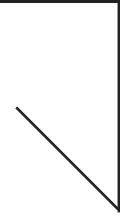
결론

- 토픽별 키워드 분석 결과, 코로나19 대유행 이후 넷플릭스 이용자들은 좀 더 구체적인 상황 언급과 세부적인 단어들을 선택하여 서비스와 앱을 평가
 - 이를 통해 사용자들의 경험증가는 사용자들이 서비스에 대해 더 자세한 이용 경험을 나누고자 한다는 점을 알 수 있음
 - 경험을 나누는 과정에서 서비스 제공자와 이용자 간의 의사소통 필요성이 높아짐
 - 다양한 콘텐츠 양이 늘어나면서 특정 콘텐츠에 대한 평가보다는 전반적인 콘텐츠의 질에 대한 의견이 증가했음
 - 이용량이 증가한만큼 불편함과 부정적인 의견도 증가했고 서비스의 안정성과 신뢰성을 높이기 위해 지속적인 업그레이드와 개선을 추진해야 할 필요성을 가지게 됨
 - 이는 새로운 콘텐츠 개발과 OTT 시장에서의 경쟁력을 강화하는데 도움이 됨
-

04. 결론 및 한계

한계

- 데이터 수집에서 코로나19 대유행 이전은 2년 1개월, 대유행 이후는 2년 6개월간의 데이터를 수집하여 두 기간에 차이가 있기에 코로나19 대유행을 기준으로 정확하게 비교하기에는 아쉬움
- 영어로 된 리뷰만을 선정하여 분석하였기 때문에 영어권 이용자들의 반응 위주로 파악할 수밖에 없음



향후 연구

- 코로나19 대유행을 기점으로 발생 이전과 이후의 데이터 수집 기간을 동일하게 설정한다면 더 정확한 비교가 가능할 것
- 6개월에서 1년 정도에 해당하는 가까운 과거의 데이터를 사용하면 이용자들의 최근 반응 흐름과 새로운 토픽의 등장을 살펴볼 수 있을 것
- 비영어권 이용자들의 리뷰까지 살펴본다면 이용자들에 대해 다양하고 더욱 폭넓은 이해를 할 수 있을 것

Q&A