

ChatGPT 및 한국관광공사 빅데이터 기반 사용자 맞춤 관광지 추천 애플리케이션

신지호, 권은혜, 이병정

서울시립대학교 컴퓨터과학부

sjm010529@gmail.com, enehyejjang@gmail.com, bjlee@uos.ac.kr

A User-Customized Tourist Destination Recommendation Application Utilizing ChatGPT and Big Data from the KTO

Jiho Shin, Eunhye Gwon, Byungjeong Lee

University of Seoul

요약

본 연구는 사용자가 여행 기록을 입력하면 ChatGPT 및 한국관광공사 빅데이터를 활용하여 맞춤 관광지를 추천하는 애플리케이션을 소개한다. 여행 기록은 관광지에서 할 수 있는 활동을 한 문장의 형태로 작성하며, 관광지 정보는 한국관광공사 빅데이터에서 관광지 개요를 사용한다. 추천 정확도를 높이기 위해 ChatGPT가 관광지 개요를 보고 관광지에서 할 수 있는 활동을 찾아내어 사용자가 작성한 여행 기록과 비교하여 추천한다. 사용자 데이터가 없어도 추천 가능한 방법으로 초기 사용자가 모이지 않은 관광 애플리케이션에서 활용할 것으로 기대한다.

1. 서론

전통적인 추천 방식으로는 협업 필터링이 많이 사용되지만, 사용자-아이템의 선호도 정보가 부족할 경우 성능이 저하된다[1]. 출시 초기 단계일 때 사용자 데이터가 부족하므로 사용에 어려움이 있다. 또한, 한국관광공사 기준 최소 15,985개 이상인 국내 관광지에 모든 사용자의 선호도 정보가 있을 수 없다. 따라서 많은 관광객이 다녀가지 않은 관광지들은 추천 대상에서 제외되는 경향이 있다. 이러한 배경에서 본 연구는 한국관광공사 빅데이터와 ChatGPT를 활용하여 사용자-아이템 선호도 정보가 없더라도 사용자 맞춤 관광지 추천이 가능한 방법을 제안하고, 애플리케이션으로 개발한다.

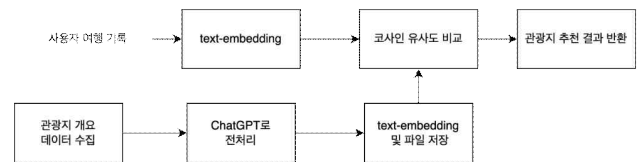
2. 관련 연구

ChatGPT를 사용하기 위해 원하는 결과를 얻으려면 프롬프트 엔지니어링이 필요하다. 프롬프트 작성 가이드라인에 따라 목표와 요구사항을 명확하게 전달하고 예시 결과를 나타내어 프롬프트를 작성한다[2]. 큰 목표인 '관광지 개요를 보고 관광지에서 할 수 있는 활동 추출하기'를 설정하고 작은 목표로 출력 형태, 문장 형식 등을 정하였다.

3. ChatGPT 및 한국관광공사 빅데이터 기반 관광지 추천

3.1 시스템 개요

사용자가 관광지에서 하였던 활동을 여행 기록으로 남기면 그 활동을 할 수 있는 다른 관광지를 찾아 추천한다. 예를 들어, 사용자가 '바다에서 산책하기'를 작성하였다면 '바다에서 산책하기' 활동을 할 수 있는 다른 적합한 관광지를 추천한다. 그림 1은 본 연구의 전체적인 개념 시나리오이다. 한국관광공사에서 관광지 개요 데이터를 수집하고 ChatGPT로 개요를 보고 관광지에서 할 수 있는 활동을 추출한다. 이후 text-embedding을 통해 벡터로 나타내어 파일로 저장한다. 사용자는 여행 기록을 입력하여 text-embedding을 한 후 관광지 데이터와 코사인 유사도를 비교하여 추천 결과를 확인한다.



(그림 1) 개념 시나리오

3.2 한국관광공사 빅데이터 전처리

관광지 정보 데이터는 한국관광공사 빅데이터에서 15,985개의 관광지의 관광지 개요를 사용한다. 그러나 관광지 개요는 배경, 볼거리 등 관광지를 설명하는 글이므로 사용자가 작성한 글과 종류가 달라서 텍스트 유사도를 비교하기 어렵다. 따라서 ChatGPT가 관광지 개요를 보고 관광지에서 할 수 있는 활동을 찾아내도록 하였다. 예를 들어, 관광지 개요가 "현재 '서울숲'은 18만 평 규모의 ... 환경친화적인 대규모 휴식공간으로 탈바꿈하였다." 라면, "한강공원에서 자연호안 산책 및 수변공간에서 휴식하기"와 같은 활동을 찾아낸다. 프롬프트 엔지니어링과 20여개의 예시 데이터를 gpt-3.5-turbo 모델에 파인 튜닝하여 적절한 결과가 출력되도록 확인한 후 15,985개의 데이터를 모두 변환하였다.

3.3 추천 시스템 구성

전처리를 마친 관광지 정보 데이터와 사용자 여행 기록과의 관계로 관광지를 추천하기 위해 텍스트 유사도를 비교한다. 관광지 정보 데이터를 ChatGPT의 text-embedding-ada-002 모델로 텍스트 임베딩하여 벡터로 나타내고 CSV 파일로 저장한다. Python의 FastAPI 패키지를 이용해 관광지 추천 서버를 구성하여 사용자 여행 기록, 추천 지역을 POST 형식으로 받는 API를 작성한다. 애플리케이션으로부터 추천 URL로 POST 요청이 오면 사용자 여행 기록을 text-embedding-ada-002 모델로 텍스트를 벡터로 나타내고 관광지 정보 데이터와 코사인 유사도를 비교한다. 서버는 추천 지역에 해당하는 관광지 중 유사도가 높은 Top-5를 반환한다. 결과적으로 사용자는 애플리케이션 메인 화면에서 서버에 API 요청하고, 관광지를 추천받아 확인한다.

3.4 추천 성능

이전 연구[3]에서는 관광지 추천 지역을 선택할 수 없었지만, 본 연구는 시도별 지역 설정과 시군구별 세부 지역 설정이 가능하여 원하는 지역의 관광지를 추천 받는다. 또한, 관광지 개요에서 할 수 있는 활동을 추출하는 전처리 작업을 진행하여 추천 정확도가 높아짐을 확인하였다. 표 1은 "미디어아트 관람하기"를 전국 지역으로 조회한 결과를 비교하였다. 전처리를 진행하지 않으면 미술관, 아트와 관련된 장소가 추천되는 반면 전처리를 진행한 경우 5곳 모두 미디어아트 관람이 가능한 곳으로 추천하였다.

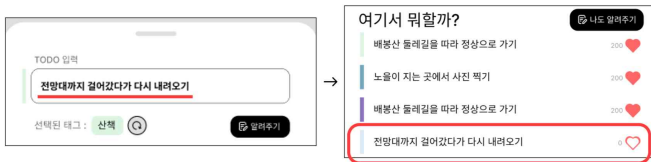
(표 1) 전처리 작업 여부에 따른 추천 결과 비교

전처리 진행 X	전처리 진행 O
벗이미술관	K-컬처 스크린 (대한민국역사박물관)
에림당아트홀	디서울라이티움
어린이회관 눈썰매장	하이커 그라운드
성남큐브미술관	가나아트센터
동작청소년문화의집	마로니에공원

4. UI

4.1 사용자의 정보 입력

사용자는 관광지에 대해 ①활동과 ②후기를 입력할 수 있고, 또 관심 있는 관광지는 ③스크랩할 수 있다.



(그림 2) 활동 입력 시나리오

그림 2와 같이, 사용자는 관광지 상세 페이지에서 해당 관광지에서 할 수 있는 활동을 입력할 수 있다. 입력된 활동은, 메인 페이지의 사용자 맞춤 추천 관광지 목록을 만들 때 반영된다.



(그림 3) 후기 입력 시나리오

그림 3과 같이, 사용자는 활동 상세 페이지에서 사진과 함께 자신의 후기를 입력할 수 있다.



(그림 4) 스크랩 시나리오

사용자는 장소 상세 페이지에서 해당 장소를 스크랩할 수 있다. 그림 4와 같이, 자신이 스크랩한 장소들은 스크랩 페이지에서 한 번에 확인할 수 있다.

4.2 사용자 맞춤 추천

사용자는 메인 페이지에서, ①지난 24시간 인기 장소, ②특정 활동을 하기 좋은 장소, 그리고 ③개인 맞춤 장소, 세 가지 목록으로 추천을 받는다.



(그림 5) 지난 24시간 인기 장소

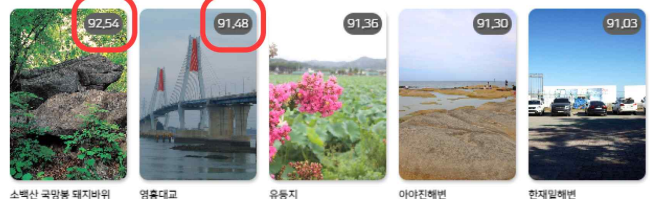
그림 5와 같이, 사용자에게 지난 24시간 동안 방문한 사람 수와 함께 인기 장소 등수 정보가 제공된다.



(그림 6) 특정 활동을 하기 좋은 장소

사용자에게, 특정 활동을 할 수 있는 장소 목록이 제공된다. 그림 6과 같이, 사용자가 활동을 선택할 수 있으며, 장소당 구체적인 활동 정보까지 함께 목록에서 볼 수 있다.

할 권은혜님이 좋아할 장소



(그림 7) 개인 맞춤 장소

사용자가 입력한 활동을 기반으로 사용자 맞춤 장소가 제공된다. 그림 7과 같이, 사용자가 입력한 활동과 관광지 정보의 유사도를 코사인 유사도로 계산한 결과를 숫자로 함께 나타낸다.

5. 결론

본 연구에서 제시한 애플리케이션을 통해, 사용자는 여행 기록을 작성하고 관광지를 추천받는다. 사용자-아이템 정보가 부족한 상황에서 ChatGPT와 빅데이터를 활용한 새로운 추천 방식을 제안하였다. 여행 분야가 아닌 데이터가 부족한 다른 분야에서도 활용할 것으로 기대한다.

참고 문헌

[1] 신경식, 강소이, "Conditional Generative Adversarial Network(CGAN) 기반 협업 필터링 추천 시스템", 지능정보연구, 27(3), pp. 157-173, 2021.
 [2] 이명숙, "ChatGPT 를 기반으로 한 프롭트 엔지니어링 기법 연구", 한국컴퓨터정보학회 학술발표논문집, 31.2, pp. 715-718, 2023.
 [3] 신지호, 권은혜, 류병욱, 이병정, "딥러닝 모델을 활용한 관광지, 활동 정보 공유 애플리케이션", 한국정보처리학회 ACK 2023 학술대회논문집, Nov. 2023.