

제12회 2024 빅콘테스트 생성형 AI 분야

LLM 활용 제주도 맛집 추천 대화형 AI 서비스 개발

배경

데이터로부터 인사이트를 얻어내는 것은 쉽지 않습니다. 그 양이 많다면 더욱 그렇죠.
복잡한 데이터 처리나 계산 없이도 데이터로부터 필요한 정보를 바로 얻어낼 수 있다면 얼마나 좋을까요?
그것도 바로 우리에게 익숙한 '대화'라는 수단을 통해서 말이죠.

여기 방대한 신한카드의 제주 가맹점 이용 데이터와 제주관광공사의 관광 데이터가 있습니다.
이를 활용해 핫한 제주도 맛집을 추천해주는 대화형 AI 서비스를 개발해주세요.
복잡하고 번거로운 검색을 하지 않아도, 몇 마디 질문 만으로 마음에 쏙 드는 제주도 맛집을 찾을 수 있도록요!

주제

LLM 활용 제주도 맛집 추천 대화형 AI 서비스 개발

설명

- 제주도내 핫플레이스 맛집을 추천하는 대화형 AI 모형 개발
- LLM을 활용해 개발한 AI 모형에 기반해야 하며, 사용자가 입력하는 자연어(한국어)에 알맞은 맛집을 추천
- 추천 대상 가맹점은 신한카드에 등록된 가맹점 중 매출 상위 9,252개 요식업종(음식점, 카페 등) 가맹점으로 제한

평가

1. 검색형 질문 응답 정확도 - 정량 평가 (30%)

- 검색형 질문 : JEJU_MCT_DATA.CSV 파일에 포함된 정보(지역, 연령대, 업종 등)를 조합한 단순 정보 검색성 질문
 - 예시 1. *제주시 한림읍에 있는 카페 중 30대 이용 비중이 가장 높은 곳은?*
 - 예시 2. *제주시 노형동에 있는 단품요리 전문점 중 이용건수가 상위 10%에 속하고 현지인 이용 비중이 가장 높은 곳은?*
- 모든 팀에 동일한 5개의 질문을 입력해 응답 정확도를 평가 (각 6점 - 정확히 일치 : 6점, 부분 일치 : 3점, 불일치 : 0점)
 - 대회 종료 후 질문 세트 공개 예정

2. 추천형 질문 응답 구현도 - 정성 평가 (40%)

- 팀별로 한 가지 이상의 질문 테마를 정해 해당 테마에 특화된 대답을 하도록 구성
 - 예시 1. 지역별 요새 뜨는 맛집 추천 (*애월에서 요새 뜨는 카페가 어디야? / 제주도에서 요새 제일 핫한 메뉴가 뭐야?*)
 - 예시 2. 유사한 컨셉의 맛집 추천 (*중문 숙성도처럼 숙성 고기 파는데 웨이팅은 적은 식당 있을까?*)
 - 예시 3. 동선, 시간대를 고려한 맛집 추천 (*아침에 우진 해장국 오픈런 할 건데, 근처에 후식으로 디저트 먹으러 갈 카페 알려줘.*)
- 테마의 참신성, 흥미성, 구현 완성도를 종합 평가
- 데이터에 직접 언급되는 정보만을 활용하는 것이 아닌, 질문의 맥락을 파악한 추론이 필수적인 테마일 수록 가산점
 - 예시. *바다 보이는 횡집 추천해줘. 제주 신화월드 근처에 부모님 모시고 가기 좋은 집 추천해줘.* 등

3. 기타 (30%)

- 실제 제주도 맛집을 검색하는 유저 관점에서 서비스를 사용해보며 느껴지는 점을 토대로 서비스 완성도를 종합 평가
- 규칙(4~5페이지) 미 충족 시 감점 및 실격으로 처리될 수 있음

규칙

분류	번호	규칙
LLM		LLM은 Gemini 1.5 Flash로 사용 한정
	1	<ul style="list-style-type: none"> - 유료 결제 및 Fine-Tuning 제한 (무료로 사용가능 한 리소스 내에서 활용) - 코드 제출 시 API 키는 삭제하고 제출 (실수로 키가 공개되지 않도록 유의)
		Inference시 신한카드 데이터에 기반한 응답이 이루어져야 함
	2	<ul style="list-style-type: none"> - 단순 Gemini LLM에 기반한 응답 인정 안됨 - 제주관광공사 데이터 활용은 선택
	3	싱글턴(Single-Turn) / 멀티턴(Multi-Turn) 방식 모두 허용
LLM		한 번 질의에 LLM 호출은 최대 3회 제한
	4	<ul style="list-style-type: none"> - 각 LLM 호출 시 최대 5,000 토큰 제한 (Input)
LLM		임베딩 모델은 아래 리스트에 있는 것으로 사용 한정
	5	<ul style="list-style-type: none"> - Hugging Face에 업로드 되어있으며 추가적인 Fine-Tuning 제한 [List] - intfloat/multilingual-e5-large-instruct - jhgan/ko-sroberta-multitask - upskyy/bge-m3-Korean

규칙

분류	번호	규칙
데이터		신한카드 데이터 활용 필수
	6	<ul style="list-style-type: none"> - 제주관광공사 데이터 활용은 선택 (일부만 사용하는 것도 가능) - 신한카드와 제주관광공사 데이터 간 결합키는 따로 제공되지 않으며 가맹점 이름, 주소 등을 활용한 자체 결합 필요
	7	크롤링 등 외부 데이터 사용 가능 <ul style="list-style-type: none"> - 출처 및 수집 방법 표기
	8	제공 데이터(신한카드, 제주관광공사)의 전처리 및 가공 가능
	9	Python Streamlit 패키지를 활용해 웹 환경에서 동작하도록 구현 <ul style="list-style-type: none"> - 샘플 코드 제공 예정
기타		평가 시 Colab Pro 환경에서 테스트 예정
	10	<ul style="list-style-type: none"> - Colab을 반드시 사용해서 개발해야 하는 것은 아니나 평가가 Colab Pro 환경에서 이루어지므로 Colab Pro 환경을 초과하는 리소스를 사용하지 않도록 유의 - 참가자에게 Colab Pro 월 이용 요금(\$9.99) 지원
	11	Colab Pro 모든 유료 서비스 사용 제한 <ul style="list-style-type: none"> - Streamlit으로 개발된 웹 서버를 띄우기 위한 단순 용도의 서버(AWS 등) 활용은 가능

결과물 제출

1. 서비스 설명서

- 팀별로 정한 추천형 질문 테마를 포함하여 서비스에 대한 상세 설명 작성 (ppt로 작성 후 pdf로 변환하여 제출)
- 코드, 데이터 등 서비스 개발 과정 전반에 대한 설명 포함
- 2차 발표 평가 시 제출한 파일 그대로 발표 예정

2. Streamlit으로 개발된 웹 서비스 시연 동영상

- 팀별로 정한 테마에 맞는 추천형 질문을 하고 응답이 나오는 모습을 포함한 서비스 시연 동영상
- 동영상의 총 길이는 3분 이내로 하며, 재생 속도나 컷 편집 등 어떠한 편집도 하지 않은 원본 그대로 제출
- 장면에 대한 설명이 필요한 부분은 서비스 설명서에 포함
- 2차 발표 평가 시 Streamlit 시연 예정 (제출한 영상 활용 또는 당일 현장에서 실시간 시연)

3. 코드

- 서비스 동작에 필요한 코드 전체
- 패키지 설치(requirements.txt 등), 환경 설정, 데이터 로드 및 전처리 코드 등 포함
- Streamlit 구동을 위한 코드 포함

4. 데이터

- 서비스 동작에 필요한 데이터 전체
- Read/Write 시 경로 에러가 나지 않도록 유의

Streamlit 구현 예시

🍊 참신한! 제주 맛집

언드레 가신디가?

저녁

어드레가 맘에 드신디가?

- ☒ 제주도민 맛집
☐ 관광객 맛집

Clear Chat History

흔저 읍서예! 🙌

군맛난 제주 밥집 🍱 추천해드릴게예

#흑돼지 #갈치조림 #옥동구이 #고사리해장국 #전복뚝배기 #한치물회 #빙떡 #오메기떡.. 🍽



🍽 어드런 식당 찾으시쿠과?

🍽 애월 쪽에 흑돼지 맛집 있어?

🍽 애월 쪽 흑돼지 맛집을 찾으시는군요! 특히 제주도민(현지인) 맛집을 선호하신다니, 딱 맞는 정보를 드릴게요.

제가 제공받은 정보를 바탕으로, 현지인 이용 비중이 가장 높은 곳은 "엄지흑돼지" 입니다.

- **엄지흑돼지:** 현지인 이용 비중이 0.3867로, 세 곳 중 가장 높습니다. 단품요리 전문점으로, 맛도 좋고 현지인들에게 인기가 많다는 뜻이죠.

다만, 현지인 비중만으로 판단하기에는 부족한 면이 있습니다.

추가적으로 고려해볼 만한 점:

- **이용건수:** 엄지흑돼지와 흑돼지가노는집은 동일 업종 내 상위 10~25%로 이용건수가 높습니다. 많은 사람들이 찾는다는 의미죠.
- **요일/시간대:** 흑돼지가노는집은 월요일에 이용건수가 가장 높고, 엄지흑돼지는 금요일에 높습니다. 방문 계획에 맞춰 고려해보세요.

Your message



샘플 코드 제공 예정
(추후 공지사항 확인)