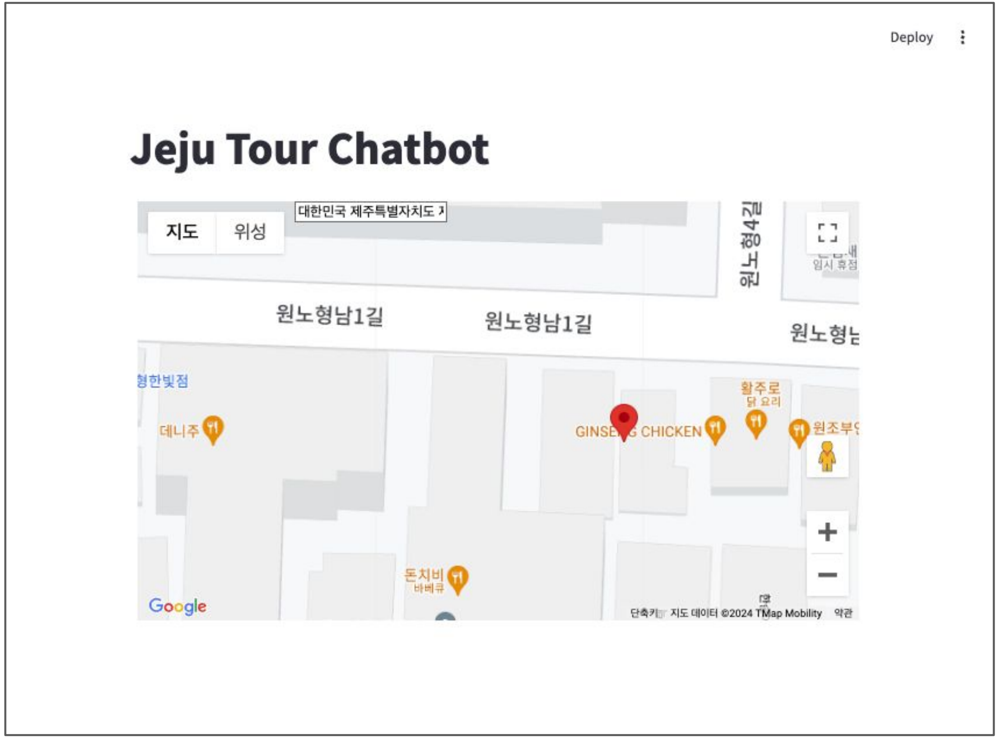
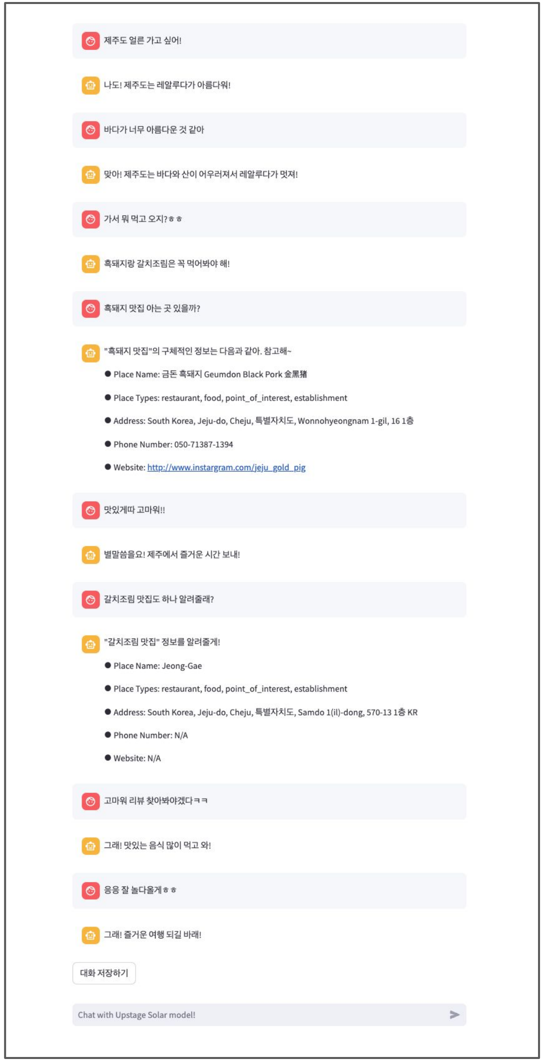
Jeju Tour Chatbot

Project Submission

Project Overview (Web page)



Jeju Tour Chatbot



제주도 얼른 가고 싶어!

나도! 제주도는 레알루다가 아름다워!

바다가 너무 아름다운 것 같아

맞아! 제주도는 바다와 산이 어우러져서 레알루다가 멋져!

가서 뭐 먹고 오지?ㅎㅎ

흑돼지랑 갈치조림은 꼭 먹어봐야 해!

흑돼지 맛집 아는 곳 있을까?

"흑돼지 맛집"의 구체적인 정보는 다음과 같아. 참고해~

- Place Name: 금돈 흑돼지 Geumdon Black Pork 金黑豬
- Place Types: restaurant, food, point_of_interest, establishment
- Address: South Korea, Jeju-do, Cheju, 특별자치도, Wonnohyeongnam 1-gil, 16 1층
- Phone Number: 050-71387-1394
- Website: http://www.instargram.com/jeju_gold_pig

맛있게따 고마워!!

별말씀을요! 제주에서 즐거운 시간 보내!

갈치조림 맛집도 하나 알려줄래?

"갈치조림 맛집" 정보를 알려줄게!

- Place Name: Jeong-Gae
- Place Types: restaurant, food, point_of_interest, establishment
- Address: South Korea, Jeju-do, Cheju, 특별자치도, Samdo 1(il)-dong, 570-13 1층 KR
- Phone Number: N/A
- Website: N/A

고마워 리뷰 찾아봐야겠다ㅋㅋ

그래! 맛있는 음식 많이 먹고 와!

응응 잘 놀다올게ㅎㅎ

그래! 즐거운 여행 되길 바래!

Objectives

- A website where users can basically have multi-turn small talk conversations on the topic of Jeju Island tour and receive answers with detailed information about the place when they want.
- Utilizes a multi-turn conversational AI model created by fine-tuning a Solar model from Upstage, which is well-performing LLM model.
- Also utilizes the basic Solar model from Upstage for generating a query for searching detailed information from a Google Maps.

Process (Main Functions)

- Step 1: User inputs a utterance
- Step 2: The basic "Solar-1-mini-chat" model generates a output sentence by getting an instruction prompt as an input and indicating whether a search is necessary and the query keyword to search for.
- Step 3: Model outputs a response
    - Step 3-1: If the search is necessary, the query keyword is going to searched in Google Maps API and then the model generate a response including detailed information of the place returned from API
    - Step 3-2: The fine-tuned "Solar-1-mini-chat" model generates a small talk response about the Jeju tour.
- Step 4: The user and the model give and receive multi-turn conversations repeatedly.

Process (Fine-tuning details)

- Source Dataset: **"Korean SNS multi-turn conversation data"** (AI-Hub)
  - Use only data from conversation sessions where the conversation topic is "travel, tourism, and attractions," and among them, data from conversation sessions that include "Jeju" in the conversation session keyword.
  - pre-processed result:
    - Training
      - total 453 sessions, 7288 instances, 3640 turns
      - Each session contains an average of about 8 turns of dialogue.
    - Validation
      - total 57 sessions 916 instances, 458 turns
      - Each session contains an average of about 8 turns of dialogue.

- Fine-tuning Dataset: **train 3,640 / valid 458 pairs**
  - Input-output structure: {dialogue history up to the previous turn - model response in this turn} pair

Process (Fine-tuning details)

- Pre-trained Model: **"Solar-1-mini-chat"** (Upstage)
  - Fine-tune the model parameter-efficiently by applying a LoRA adapter.

- Prompt and Completion Formatting:
  - Prompt:
    - “**<|im_start|>user\n**제주도 여행 어땠음? 나도 가려구.ㅋㅋ**<|im_end|><|im_start|>assistant\n**혼여하는거야? 아님 같이 가는거?**<|im_end|><|im_start|>user\n**응응. 혼자가고 게하간당. 텅장이야. ㅠㅠ**<|im_end|><|im_start|>assistant\n**”
  - Completion:
    - “게하도 너무 재밌다궁. ㅋㅋ 나 이번에 야간시티투어도 타고 재밌었어.**<|im_end|>**”

Process (Generating a query details)

- Base Model: **"Solar-1-mini-chat"** (Upstage)
  - Inference the model using OpenAI API

- Instruction prompt:
    - **“You are an assistant that helps to determine if a Google Maps API query is needed and generates the query.**
      사용자 입력: "{user_input}
      이 문장이 구글 맵스 API 쿼리 검색이 필요한 문장인지 확인하고, 필요하다면 검색 쿼리를 생성해 주세요.
      검색 쿼리만 생상하고 부연설명은 함께 생성하지 마세요.
      예시:
      - 문장: "돈사돈 주소가 어떻게 돼?" -> 검색 쿼리: "돈사돈"
      - 문장: "제주토종흑돼지 월요일 휴무야?" -> 검색 쿼리: "제주토종흑돼지"
      - 문장: "카페한라산 영업시간 언제까지야?" -> 검색 쿼리: 카페한라산"
      - 문장: "오늘 날씨 어때?" -> 검색 쿼리: "N/A"
      - 문장: 애월 카페거리는 정말 멋지지! -> 검색 쿼리: "N/A"

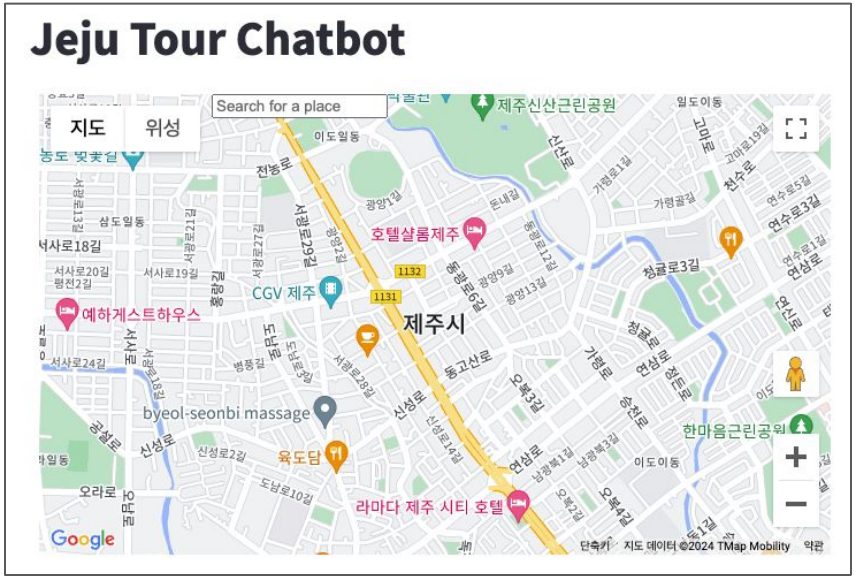
Process (Searching the information details)

- API: **"Google Maps API"** (Google)
    - Searching the query keyword into Google Maps API and returned detailed information of the place from API

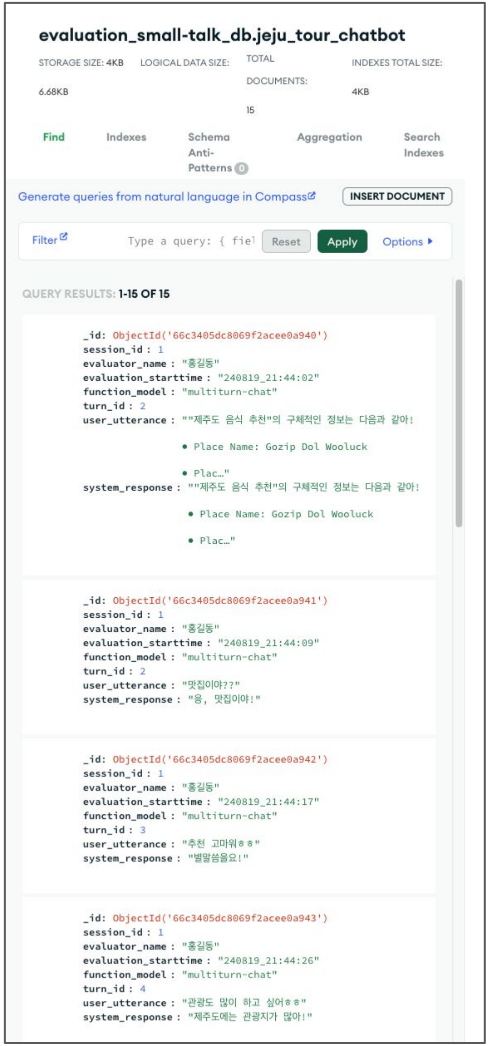
- Detailed information
    - Place Name
    - Place Types
    - Address
    - Phone Number
    - Website

Process (Implementation tools)

- Web: Streamlit
- APIs: Google Maps API
- Database: MongoDB



Jeju Tour Chatbot



evaluation_small-talk_db.jeju_tour_chatbot

STORAGE SIZE: 4KB
LOGICAL DATA SIZE: 6.68KB
TOTAL DOCUMENTS: 15
INDEXES TOTAL SIZE: 4KB

QUERY RESULTS: 1-15 OF 15

```
_id: ObjectId('66c3405dc8069f2acee0a940')
session_id : 1
evaluator_name : "홍길동"
evaluation_starttime : "240819_21:44:02"
function_model : "multiturn-chat"
turn_id : 2
user_utterance : ""제주도 음식 추천"의 구체적인 정보는 다음과 같아!

                • Place Name: Gozip Dol Wooluck

                • Plac…"
system_response : ""제주도 음식 추천"의 구체적인 정보는 다음과 같아!

                 • Place Name: Gozip Dol Wooluck

                 • Plac…"

_id: ObjectId('66c3405dc8069f2acee0a941')
session_id : 1
evaluator_name : "홍길동"
evaluation_starttime : "240819_21:44:09"
function_model : "multiturn-chat"
turn_id : 2
user_utterance : "맛집이야??"
system_response : "응, 맛집이야!"

_id: ObjectId('66c3405dc8069f2acee0a942')
session_id : 1
evaluator_name : "홍길동"
evaluation_starttime : "240819_21:44:17"
function_model : "multiturn-chat"
turn_id : 3
user_utterance : "추천 고마워ㅎㅎ"
system_response : "별말씀을요!"

_id: ObjectId('66c3405dc8069f2acee0a943')
session_id : 1
evaluator_name : "홍길동"
evaluation_starttime : "240819_21:44:26"
function_model : "multiturn-chat"
turn_id : 4
user_utterance : "관광도 많이 하고 싶어ㅎㅎ"
system_response : "제주도에는 관광지가 많아!"
```

Outcomes

- Fine-tuned model: the fine-tuned adapter by LoRA method
- Web page: the website developed based Streamlit
- Github: https://github.com/sallyy1/Jeju-Tour-Chatbot