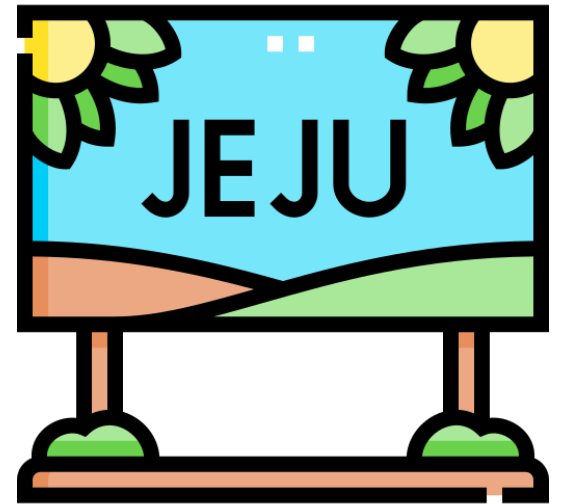


2024 빅콘테스트

# LLM 활용 제주도 맛집 추천 대화형 AI 서비스 개발



개인 참여자  
이상민



# 목차

1

대회 설명

2

데이터

- 1. 신한 카드 데이터
- 2. 웹 크롤링 데이터
- 3. 데이터 생성 디테일

3

평가

- 1. 정량 평가
- 2. 정성 평가

4

작동 방식

- 1. 도식화
- 2. 프롬프트 및 코드 리뷰

5

작동 예시

- 1. 정량 평가
- 2. 정성 평가

# 대회 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

# 데이터



## 신한 카드 데이터

- 기준 연월(YM)
- 매장명(MCT\_NM)
- 매장 오픈 일자(OP\_YMD)
- 업종(MCT\_TYPE)
- 주소(ADDR)
- 이용 건수 구간  
(6개의 구간: 상위 10%, 10~25%, 25~50%, 50~75%, 75~90%, 90%~)
- 이용 금액 구간 (위와 동일)
- 건당평균 이용 금액 구간 (위와 동일)
- 요일별 이용 건수 비중
- 시간대별 이용 건수 비중
- 현지인 이용 건수 비중
- 최근 12개월 남/녀 회원 수 비중
- 연령대별 회원 수 비중



## 크롤링 데이터

- 가게명
- 주소 (optional)
- 가게 타입
- 평점 (optional)
- 전화번호 (optional)
- 홈페이지 (optional)
- 방문자 리뷰 (optional)
- 이미지 (optional)
- 메뉴 (optional)
- 편의 (optional)

|   | YM     | 매장명    | open_date | type    | location                  | frequency | money    | avg_money_per_visited |
|---|--------|--------|-----------|---------|---------------------------|-----------|----------|-----------------------|
| 0 | 202301 | 통큰돼지   | 20110701  | 가정식     | 제주 제주시 용담이동 2682-9번지 통큰돼지 | 5_75~90%  | 4_50~75% | 3_25~50%              |
| 1 | 202301 | 해변     | 20050407  | 단품요리 전문 | 제주 제주시 애월읍 애월리 410-6번지    | 3_25~50%  | 2_10~25% | 2_10~25%              |
| 2 | 202301 | 한그릇    | 20120919  | 단품요리 전문 | 제주 서귀포시 색달동 2315-1번지 한그릇  | 3_25~50%  | 3_25~50% | 4_50~75%              |
| 3 | 202301 | 협재해녀의집 | 20130627  | 가정식     | 제주 제주시 한림읍 협재리 1459-2번지   | 3_25~50%  | 3_25~50% | 4_50~75%              |
| 4 | 202301 | 까랑고송   | 20220826  | 단품요리 전문 | 제주 제주시 노형동 3784-12번지 1층   | 5_75~90%  | 3_25~50% | 2_10~25%              |

|   | monday   | tuesday  | wednesday | thursday | friday   | saturday | sunday   | 남성 이용 | 여성 이용 | 현지인 비율   |
|---|----------|----------|-----------|----------|----------|----------|----------|-------|-------|----------|
| 0 | 0.161290 | 0.032258 | 0.129032  | 0.096774 | 0.161290 | 0.161290 | 0.258065 | 0.610 | 0.390 | 0.707763 |
| 1 | 0.090909 | 0.121212 | 0.045455  | 0.136364 | 0.181818 | 0.242424 | 0.181818 | 0.542 | 0.458 | 0.230928 |
| 2 | 0.224719 | 0.112360 | 0.000000  | 0.089888 | 0.146067 | 0.179775 | 0.247191 | 0.529 | 0.471 | 0.347059 |
| 3 | 0.000000 | 0.271739 | 0.195652  | 0.086957 | 0.130435 | 0.119565 | 0.195652 | 0.484 | 0.516 | 0.016096 |
| 4 | 0.272727 | 0.121212 | 0.060606  | 0.151515 | 0.151515 | 0.151515 | 0.090909 | 0.513 | 0.487 | 0.264463 |

|   | 5시부터 11시 | 12시부터 13시 | 14시부터 17시 | 18시부터 22시 | 23시부터 4시 | 20대 이하 | 30대   | 40대   | 50대   | 60대 이상 |
|---|----------|-----------|-----------|-----------|----------|--------|-------|-------|-------|--------|
| 0 | 0.000000 | 0.000000  | 0.161290  | 0.838710  | 0.000000 | 0.103  | 0.124 | 0.245 | 0.387 | 0.142  |
| 1 | 0.015152 | 0.181818  | 0.242424  | 0.560606  | 0.000000 | 0.221  | 0.201 | 0.195 | 0.244 | 0.139  |
| 2 | 0.101124 | 0.651685  | 0.247191  | 0.000000  | 0.000000 | 0.138  | 0.309 | 0.342 | 0.164 | 0.047  |
| 3 | 0.250000 | 0.239130  | 0.358696  | 0.152174  | 0.000000 | 0.323  | 0.321 | 0.209 | 0.109 | 0.039  |
| 4 | 0.000000 | 0.000000  | 0.181818  | 0.727273  | 0.090909 | 0.412  | 0.311 | 0.176 | 0.076 | 0.025  |



**통큰돼지** 육류,고기요리

★4.59 · 방문자 리뷰 95 블로그 리뷰 7

출발 도착

저장 거리뷰 공유

홈 메뉴 리뷰 사진 정보

제주 제주시 용문로 134 1층

공황에서 3~5분 거리에 sk현대주유소 맞편에 위치해 있습니다. 버스 이용시 용문마을에서 도보 1분 거리에 있습니다.

영업 종료 · 16:00에 영업 시작

064-713-9255 복사

<https://www.instagram.com/p/CK8i6wOhw0a/?i...>

포장, 단체 이용 가능, 유아의자, 예약

#### 방문자 리뷰 95

|               |    |
|---------------|----|
| 😊 "음식이 맛있어요"  | 69 |
| 🥩 "고기 질이 좋아요" | 22 |
| 🍲 "양이 많아요"    | 21 |
| 💰 "가성비가 좋아요"  | 17 |
| 💖 "친절해요"      | 16 |

#### 메뉴 6

메뉴판 이미지로 보기

|                    |    |         |
|--------------------|----|---------|
| 제주산 양념갈비(1인분 350g) | 대표 | 16,000원 |
| 흑돼지 김치찌개           | 대표 | 8,000원  |
| 흑돼지 오겹살(1인분 200g)  |    | 20,000원 |
| 통큰 양념구이(1대 350g)   |    | 13,000원 |

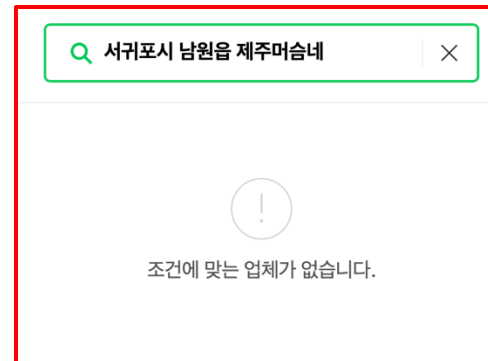
출처: 네이버 맵

# 데이터 (생성 디테일)

## Crawling

### Library

- Selenium (동적 크롤링)
- 사용한 이유: Naver map에 없는 가게도 존재하기 때문



### 방식

Naver map 페이지에서 “00시 00읍 가게명”으로 입력해서 관련도가 가장 높은 가게의 정보 수집

### 추출 정보

Naver map에 있는 정보 중  
- 기본 정보(주소, 찾아오는 길, 영업시간, 편의 등), 가게 종류, 평점, 메뉴, 방문자 리뷰, 이미지

# 데이터 (생성 디테일)

## Naver API

### Geocoding

입력: 위치 주소를 받음(도로명, 지번 등)

출력: 해당 주소의 위도, 경도, 지번 주소, 도로명 주소 → 주소를 찾을 수 없는 경우도 있음!

입력한 주소: 제주 랜디스 도넛 제주애월점

주소 정보를 찾을 수 없습니다. 반환하는 정보: (0, 0, 0, 0)

입력한 주소: 제주특별자치도 제주시 애월읍 애월리 2471

위도: 33.4612048, 경도: 126.3114628, 지번 주소: 제주특별자치도 제주시 애월읍 애월리 2471, 도로명 주소: 제주특별자치도 제주시 애월읍 애월로 27-1

### Direction 5

입력: 출발 주소와 도착 주소의 '위도 & 경도'

출력: 거리(m), 시간(milli second)

시작 위치: 제주특별자치도 제주시 애월읍 광지리 1566, 도착 위치: 제주특별자치도 제주시 애월읍 애월리 2471

거리와 소요 시간 함수 결과값: (2.2, 7). 즉, 거리: 2.2km, 소요 시간: 7분

# 데이터 (생성 디테일)

## 거리 정보

### 1차 필터링

1. 주소 데이터를 {'시', '동', '읍', '면', '리'} 만을 가지는 주소로 분류  
ex) 제주 제주시 애월읍 곽지리 1566 → 제주시 애월읍 곽지리
2. 해당 주소들을 모아, 주소 간 거리 및 시간을 계산 (Naver API 사용)
3. 해당 주소 간의 거리, 시간 정보를 distance.json 파일로 미리 저장

(예외 사항, '우도'는 다른 지역과 교류하기 힘들다고 판단  
→ 'start: 우도, end: 제주도' 또는 'start: 제주도, end: 우도' 이동 불가능하게 처리)

### 실시간 정보

1. 1차 필터링 정보로 특정 거리, 시간 내의 지역에 위치한 가게를 필터링
2. 이후, 다른 키워드 정보들을 필터링
3. 추천할 가게와 현재 위치 간의 '거리'와 '현재 걸리는 시간'을 계산 (Naver API 사용)
4. 해당 정보를 사용 → 가게 필터링 or 거리 정보 반환



# 평가

## 정량 평가

---

- 검색형 질문 🔍
- 신한 카드 데이터 파일에 포함된 정보를 조합한 단순 정보 검색형 질문
- 평가 방법: 모든 팀에 동일한 5개의 질문  
→ 응답 정확도 평가  
(정확히 일치: 6점, 부분 일치: 3점, 불일치: 0점)
- 질문 예시)
  - 제주도 한림읍에 있는 카페 중 30대 이용 비중이 가장 높은 곳은?
  - 제주도 노형동에 있는 단품요리 전문점 중 이용 건수가 상위 10%에 속하고 현지인 비율이 가장 높은 곳은?

## 정성 평가

---

- 추천형 질문 🗣️
- 신한 카드 데이터 내 가맹점의 naver map 정보를 가져와, 수치정보가 아닌 키워드 정보를 활용한 추천형 질문
- 평가 방법: 테마를 정해 해당 테마의 참신성, 흥미성, 구현 완성도 종합 평가
- 테마)
  - 자신의 현재 위치 기반, 특정 거리 또는 특정 시간 내의 가게 추천
  - 편의 시설 키워드 이용 가능 (반려동물 동반, 장애인 편의시설, 단체 이용 가능 등)
  - 해당 가게의 상위 5개 방문자 리뷰 이용 (친절해요, 대화하기 좋아요, 혼잡하기 좋아요 등)

# 평가 (정성 평가 테마)

정성 평가  
사용 요소

위치 기반 거리 서비스 (거리, 소요 시간)

평점 정보

가게 타입

고기, 한식, 국수, 해산물, 중식 등

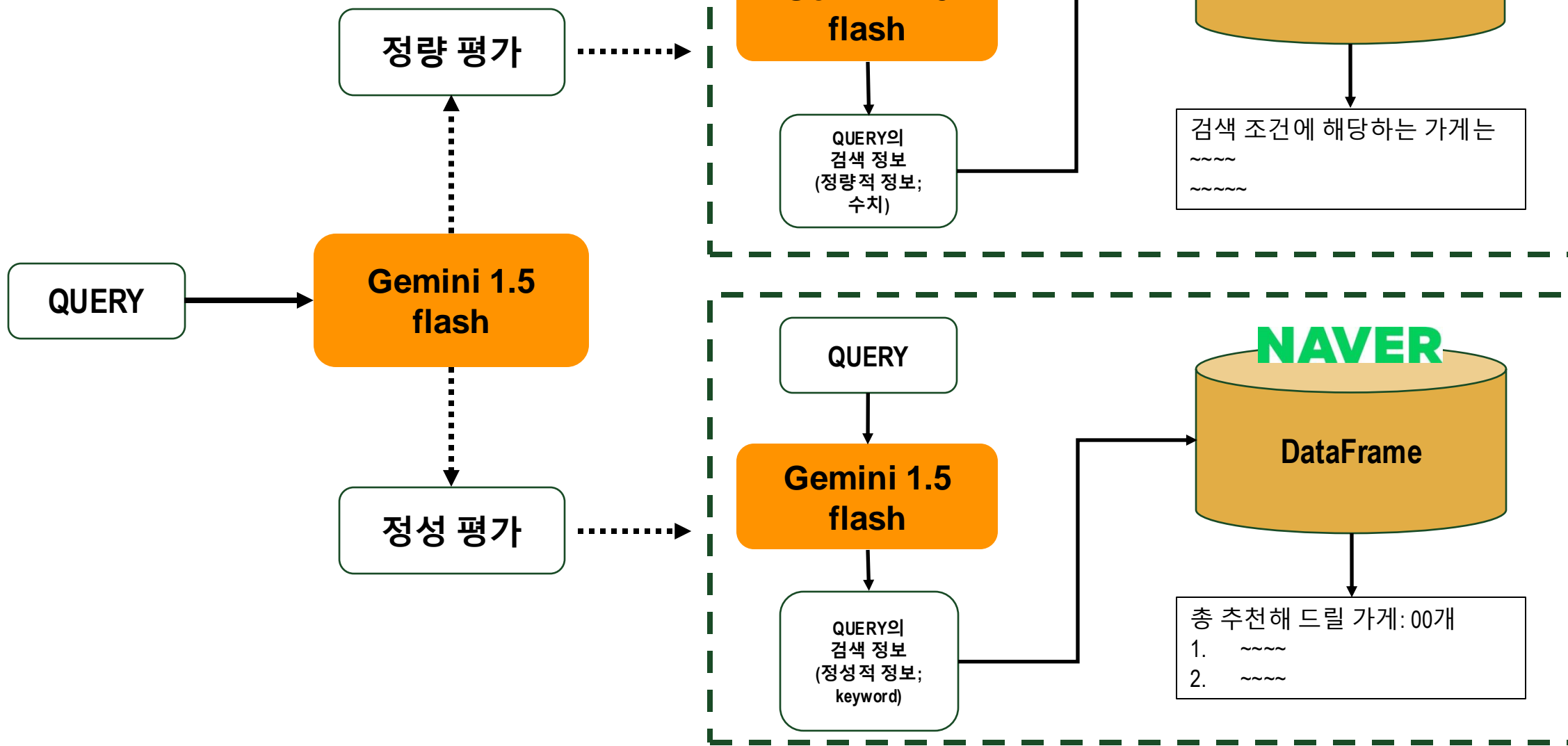
편의 정보

주차 유/무, 반려동물 동반, 무선 인터넷 등

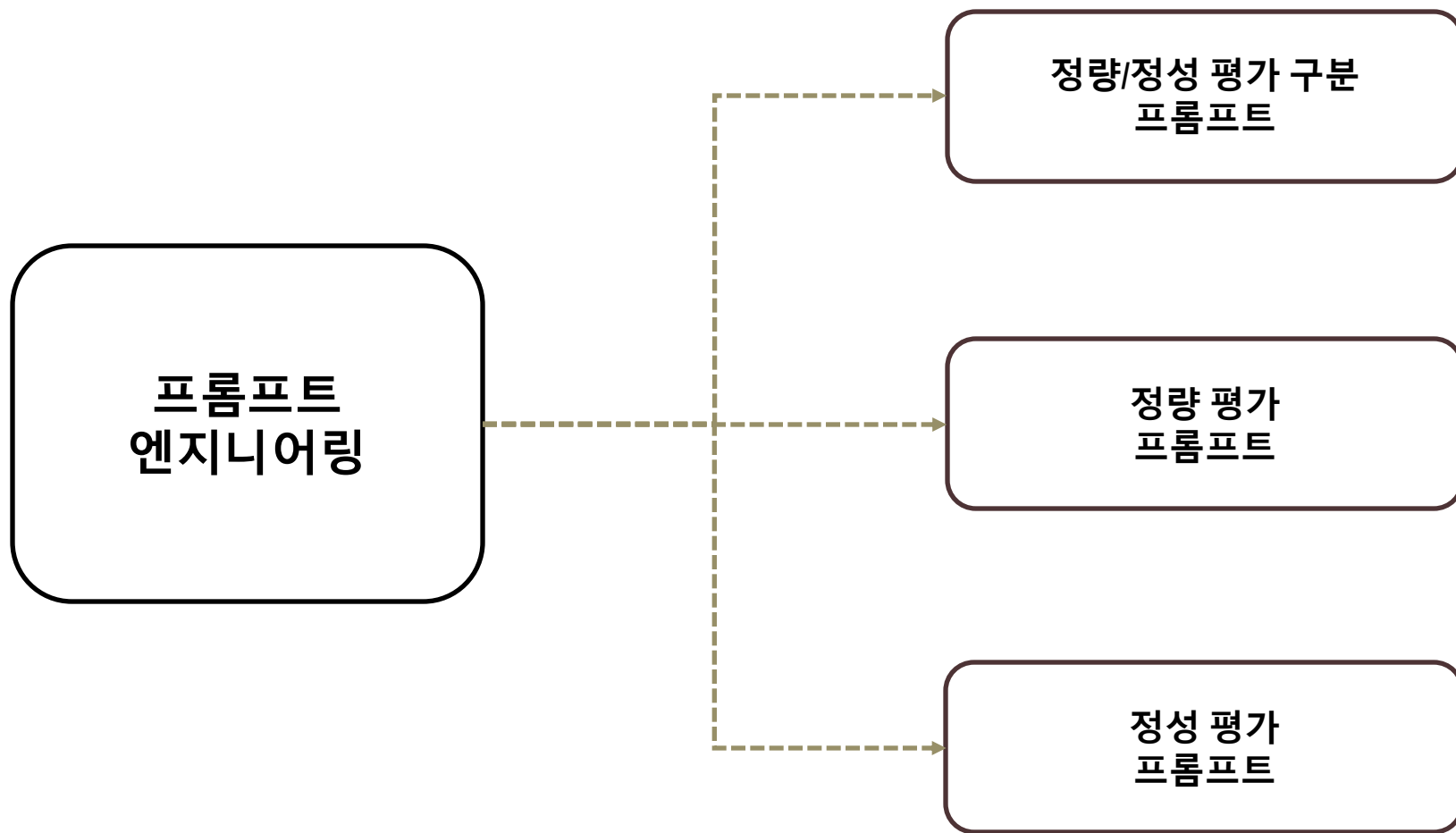
방문자 리뷰

친절해요, 가성비가 좋아요, 뷰가 좋아요 등

# 작동 방식 (도식화)



# 작동 방식 (코드 리뷰)



# 작동 방식 (코드 리뷰)

## 정량/정성 평가 구분 프롬프트

```
DISTINGUISH_PROMPT = '''
```

너는 질문이 정성 평가에 대한 질문인지, 정량 평가에 대한 질문인지 구분해야해.

각 평가에 대한 특징을 알려줄게.

정량 평가) 대체로 질문의 형태가 의문형. 조건들에 맞는 특정 가게를 찾아내야 하는 목적. 사용하는 데이터는 오직 신한 카드 데이터만 사용해.

정성 평가) 대체로 질문의 형태가 요청형. 현재 상황 또는 제한 사항에 맞는 가게들을 추천해줘야 하는 목적. 사용하는 데이터는 네이버 맵 데이터와 신한 카드 데이터를 사용할 수 있어.

평점 얘기가 나오면 '정성 평가'야.

만약 너가 질문이 정량 평가라고 판단한다면 '정량' 만 return 해줘.

만약 정성 평가라고 판단했다면 '정성' 만 return 해줘.

다른 표현은 생성하지마.

```
질문: {query}
```

```
'''
```

”정량”, “정성”을 return 받기 위한 프롬프트

# 작동 방식 (코드 리뷰)

## 정량평가 프롬프트

```
SEARCH_PROMPT = '''
# 너는 질문을 받아서 맛집 검색을 위한 조건을 찾아줘야 해.
# 조건은 총 11개의 타입을 가지고 있고, 각각 필요한 값은 아래에 정의되어 있어.

# 만약 해당 정보에 대한 조건이 질문에 없다면, 해당 값을 null으로 표기해줘.

1. location
- string. 어떤 지역의 맛집인지에 대한 지역값
2. type
- 다음 값들 중 하나의 string: '가정식', '구내식당/푸드코트', '기사식당', '기타세계요리', '꼬치구이', '단품요리 전문', '도너츠', '도시락', '동남아/인도음식'
3. weekday
- dictionary 요일에 대한 조건값
- dictionary의 형태: {"value": "weekday_value", "operator": "weekday_operator_value"}
  - "weekday_value"는 다음 7개의 값 중 하나의 string: "monday", "tuesday", "wednesday", "thursday", "friday", "saturday", "sunday"
  - "weekday_operator"는 다음 2개의 값 중 하나의 string: "max", "min"
4. age
- dictionary 나이에 대한 조건값
- dictionary의 형태: {"value": "age_value", "operator": "age_operator_value"}
  - "age_value"는 다음 5개의 값 중 하나의 string: "20대 이하", "30대", "40대", "50대", "60대 이상"
  - "age_operator"는 다음 2개의 값 중 하나의 string: "max", "min" # 가장 많이 방문하는, 가장 적게 방문하는 -> 20대 애들은 시끄러워서 20대가 가장 적게
5. local
- dictionary 현지인 이용 비율에 대한 조건값
- dictionary의 형태: {"value": "local_value", "operator": "local_operator_value"}
  - "local_bool"은 현지인 이용에 대한 내용이 있으면 '현지인 비율'
  - "local_operator"는 다음 2개의 값 중 하나의 string: "max", "min"
6. time
- dictionary 시간대 별 이용 비율에 대한 조건값
- dictionary의 형태: {"value": "time_value", "operator": "time_operator_value"}
  - "time_value"는 다음 10개의 값 중 하나의 string: "5시부터 11시", "12시부터 13시", "14시부터 17시", "18시부터 22시", "23시부터 4시"
  - "time_operator"는 다음 2개의 값 중 하나의 string: "max", "min"
7. gender
- dictionary 남녀 이용 비율에 대한 조건값
- dictionary의 형태: {"value": "gender_value", "operator": "gender_operator_value"}
  - "gender_value"는 다음 2개의 값 중 하나의 string: "남성 이용", "여성 이용"
  - "gender_operator"는 다음 2개의 값 중 하나의 string: "max", "min"
8. open_date
```

1. Location
2. Type
3. Weekday
4. Age
5. Local
6. Time
7. Gender
8. Open\_date
9. Frequency
10. Money
11. Avg\_money\_per\_visited

정보를 json 형식으로 반환하기  
위한 프롬프트

# 작동 방식 (코드 리뷰)

## 정성평가 프롬프트

```
RECOMMEND_PROMPT = '''
# 너는 조건을 받아서 맛집 추천을 위한 조건을 찾아줘야 해
# 조건은 총 6개가 있고, 각각 필요한 타입은 아래에 정의되어 있어.
# 카페는 커피를 마시는 곳으로 밥을 먹는 곳이 아니다.

# 만약 해당 정보에 대한 조건이 질문에 없다면, 해당 값을 null로 표기해줘.

1. rating
- float, 몇 점 이상 별점을 가진 가게를 찾기 위한 조건값

2. type
- string, 원하는 식당의 타입, 타입은 아래의 리스트 중 하나의 값을 가져야 해
  - 한식
  - 고기
  - 해산물
  - 카페
  - 치킨, 닭강정
  - 중식당
  - 술집
  - 양식
  - 피자
  - 돈가스
  - 국수

3. concept
- List[string],
- 설명: 유저가 원하는 해당 식당의 컨셉, 여러 컨셉의 값이 나올 수 있고, 컨셉 값은 아래의 정해진 값들만 가지고 구성해야 해! '직접적인 언급'이 있어야만 반환해.
- 추가 설명: 혼밥이란, 혼자 밥먹는 것을 말해. 가성비란, 가격대비 성능이 좋은 것을 말해. 혼술이란, 혼자 술먹는 것을 말해.
  - '친절해요'
  - '가성비가 좋아요'
  - '양이 많아요'
  - '뷰가 좋아요'
  - '혼밥하기 좋아요'
  - '매장이 청결해요'
  - '대화하기 좋아요'
  - '혼술하기 좋아요'
  - '직접 잘 구워줘요'

4. distance
- data type: int
- 설명: 현재 내위치에서 추천하는 가게까지 몇 km 안에 도착했으면 좋겠다는 정보야. 이에 대한 정보가 있으면 'km' 기준으로 정수형으로 반환, 그렇지 않으면 null 반환해줘
```

1. Rating
2. Type
3. Concept
4. Distance
5. Time
6. Convenience

정보를 json 형식으로 반환하기  
위한 프롬프트

# 작동 방식 (코드 리뷰)

## 전체 프롬프트

```
response = chat(RECOMMEND_PROMPT, query)
result = parse_json(response)
rating = result.get('rating', None)
concept = result.get('concept', None)
restaurant_type = result.get('type', None)
distance = result.get('distance', None)
time = result.get('time', None)
convenience = result.get("convenience", None)
sub_df = crawling_data

if rating is not None:
    sub_df = sub_df[sub_df.reputation >= rating]
if restaurant_type is not None:
    sub_df = sub_df[sub_df.naver_type == restaurant_type]

if concept is not None:
    sub_df = sub_df[sub_df.visitor_review.apply(
        lambda x: (x is not None) and (not pd.isna(x)) and all([c in x for c in concept])
    )]]

# 시간에 대한 언급이 없으면 20분 이내의 가게를 찾음 -> 거리는 조절 X. 거리는 그냥 있으면 필터링 하고, 아니면 그냥 낮둑기
if distance is not None:
    sub_df = sub_df[sub_df['name'].str.contains(distance_time_restriction(start_address, distance_dict, distance, 'distance'))]
    # distance = 15

# 시간도 이에 맞게 조절
if time is None and time != -20:
    time = 20
    sub_df = sub_df[sub_df['name'].str.contains(distance_time_restriction(start_address, distance_dict, time, 'time'))]

# convenience filtering!! <-- 이거 해야함!
if convenience is not None:
    sub_df = sub_df[sub_df['편의'].apply(
        lambda x: all([c in x for c in convenience]) if isinstance(x, list) else False
    )]]
```

1. 해당 조건들을 필터링 한  
Sub\_df 추출

2. 해당 sub\_df의 가게 정보를  
streamlit에 반환할 수 있게  
string 형태로 가게 설명



# 작동 예시

## 정량 평가

안녕하세요, 현재 위치 또는 시작 위치 기반으로 제주도 맛집을 추천해드립니다 🙌  
편하게 질문해주세요.



제주시 한림읍에 있는 카페 중 30대 이용 비중이 가장 높은 곳은?

나이스데이밀

## 정성 평가

강아지랑 갈 수 있는 카페 추천해줘

총 추천해 드릴 가게: 3개

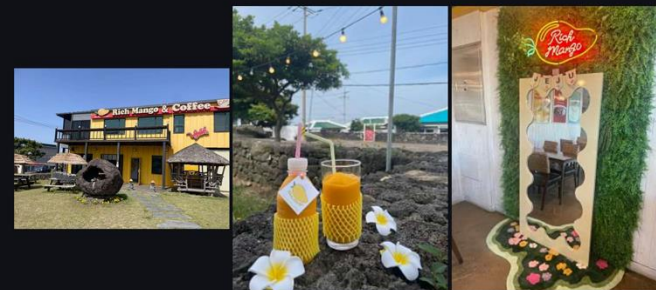
### 1. 리치망고 협재점

- ★ 별점: 4.47
- 📍 주소: 제주 제주시 한림읍 한림로 482
- 🚗 거리: 11.0km (약 17 분 소요)

[T 편의]

- 예약
- 단체 이용 가능
- 포장
- 남/녀 화장실 구분
- 반려동물 동반
- 유아의자
- 무선 인터넷
- 간편결제
- 주차

[가게 관련 사진]



출처: 네이버 맵

