테스트 계획 및 결과 보고서

1. 프로젝트 개요

|  |  |
| --- | --- |
| 프로젝트 명 | 공공 데이터 및 기업 API를 활용한 반려동물 동반 여행 가이드 챗봇 – 우리 개 어디가? |
| 테스트 대상 | 프롬프트 인식 및 LLM 결과 타당성 확인 |
| 테스트 일자 | 2025.06.03. |
| 작성자 | 박슬기 |

2. 테스트 목적

- 다양한 테스트 케이스를 입력하여 LLM이 정확한 카테고리를 반환하는지 확인

- 벡터 DB와 공공 데이터 포털 API 결과의 신뢰도 확인

3. 테스트 케이스 및 결과

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | 테스트 항목 | 테스트 방법 | 결과 | 비고 |
| 1 | 프롬프트  테스트 | 프롬프트 수정 | 실패 | ‘당일치기’ 등 특정 날짜 지시 확인 못함 |
| 2 | 프롬프트  테스트 | 프롬프트 수정 | 실패 | 현재 없어진 업체의  정보가 함께 표시됨 |
| 3 | 신뢰도  테스트 | 네이버 API 검색 추가 | 성공 |  |

4. 종합 결과 요약

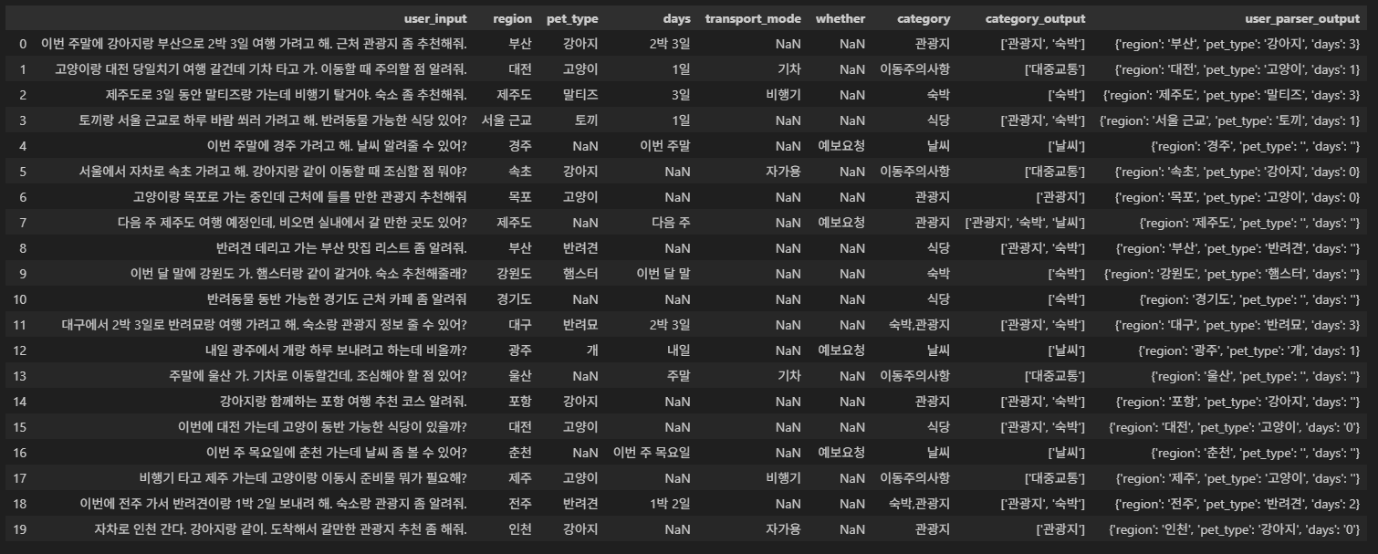
|  |  |
| --- | --- |
| 총 테스트 수 | 총 20개 프롬프트 사용, 반복 15회 |
| 성공한 테스트 수 | 전체 프롬프트 포함 최후 3회 |
| 실패한 테스트 수 | 12회 |
| 주요 이슈 요약 | 날짜 카테고리 인식 오류, 폐업 업장 정보 조회 안되도록 처리  추가적으로 환각 방지를 위해 네이버 API 호출 기능 추가 |
| 향후 조치 계획 | 테스트 케이스 추가하여 추가 예외사항 확인 |

(첨부 1) 테스트 케이스

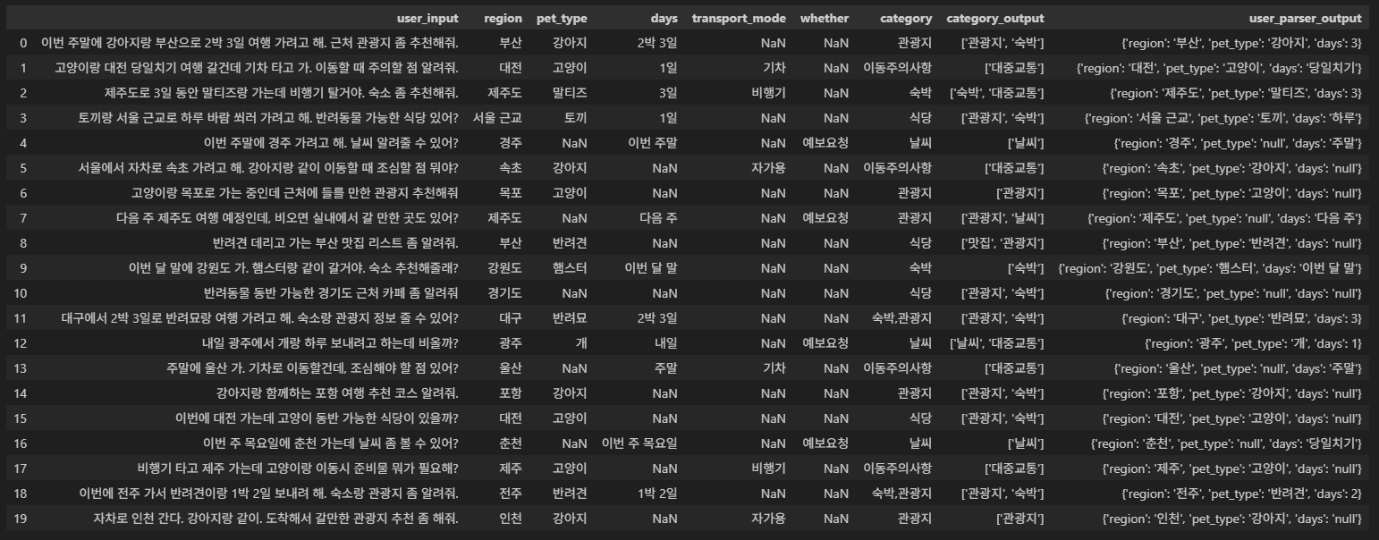


(첨부 2) 테스트 결과 사진

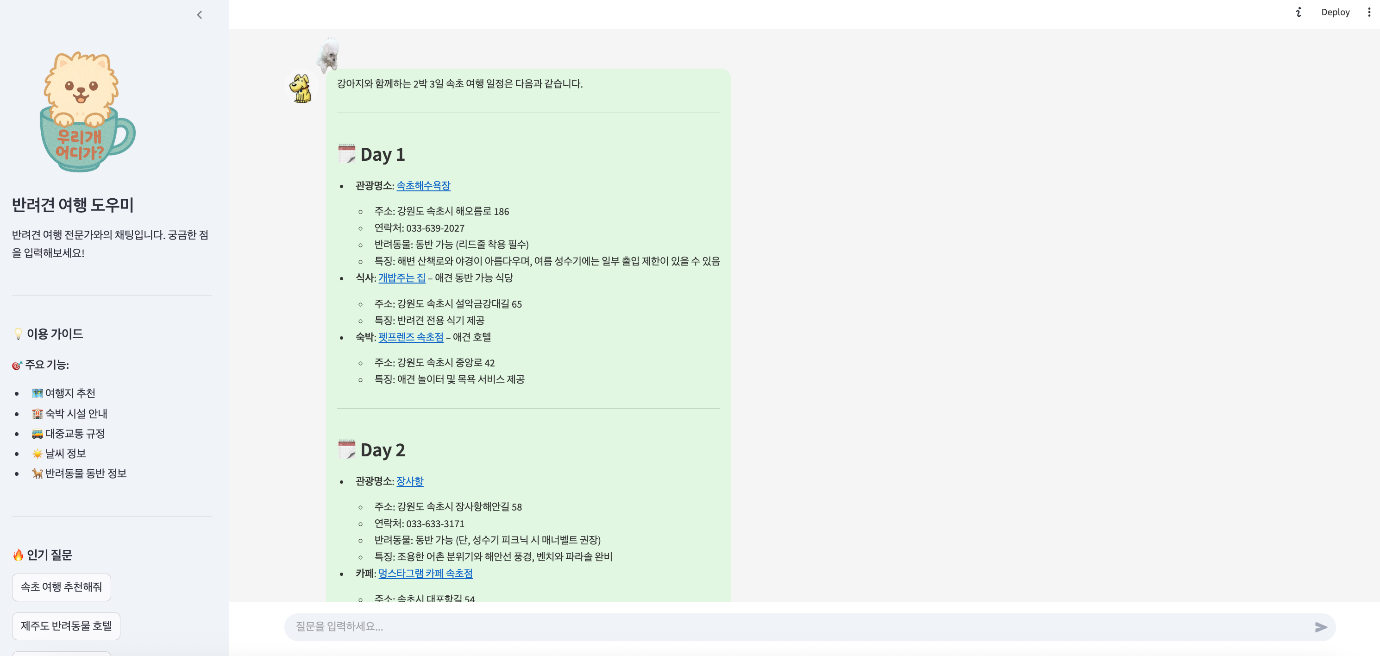
1차 테스트 결과

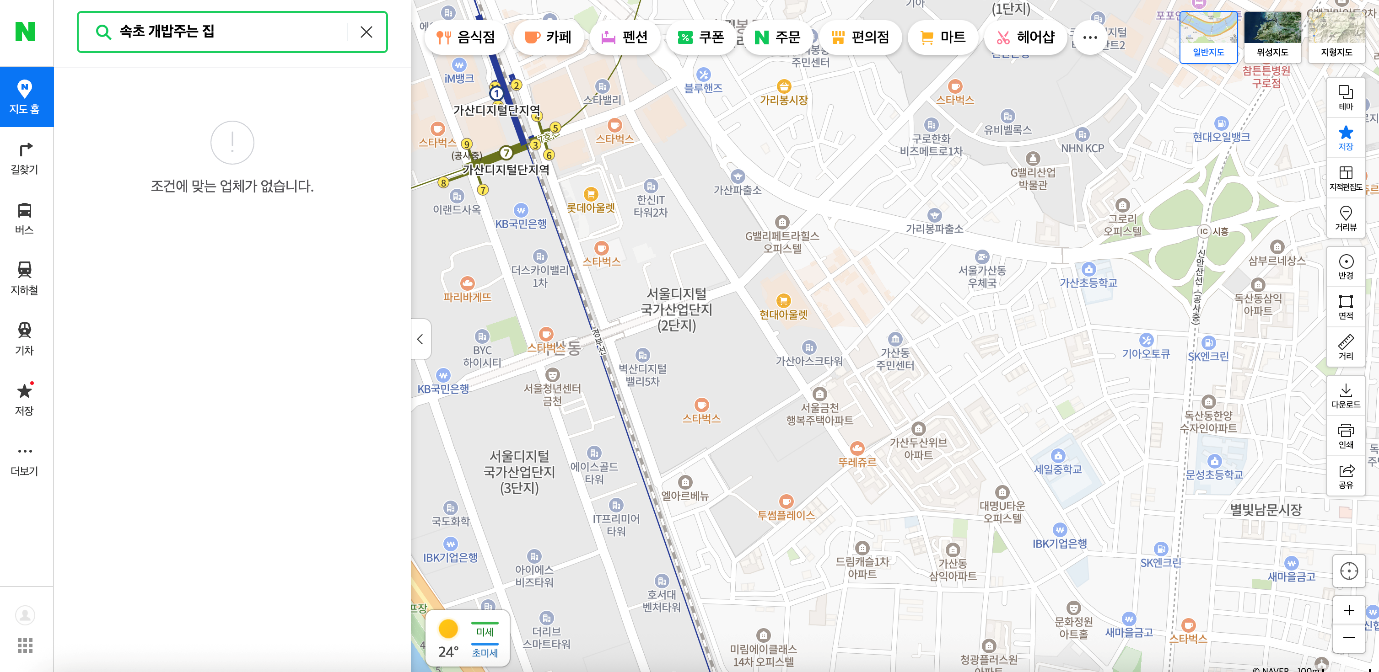


2차 테스트 결과



3차 테스트 결과 (존재하지 않는 업체 검색)





결과 네이버 API를 이용하여 현재 영업중인 업체만 선별

