

ZERO-BASE DS School

DL Project



위치와 사용자 취향을 반영한 데이트 경로 추천 시스템

2조 이동희, 이슬기, 장문정



목차

1. 주제
2. 배경/기획의도
3. 사전 자료
4. 구현 모델
5. 향후 진행 방안



주제

사용자 위치 기반

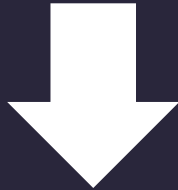
사용자 취향을 반영

주변의 명소, 맛집, 카페를 잇는 데이트 코스 추천 시스템 개발



배경/기획의도

- SNS상 맛집 추천 BUT, 위치를 직접 찾아야하는 번거로움
- 데이트 코스를 정하는 데에 어려움을 겪는 커플에게 도움



'사용자 위치 + 사용자 취향'을 반영하여 데이트 코스 추천



사전자료

모바일앱 '완벽한 하루'



지도시각화

출발지 - 경유지 - 도착지

driving 경로

거리(km), 이동시간(분), 택시비

Place 추천

사용자의 위치와 취향을 반영

맛집 / 카페 / 장소

대체장소 추천



지도시각화

출발지 - 경유지 - 도착지

driving 경로

거리(km), 이동시간(분)

- NAVER-api(Direction 5)
- Folium



Place 추천

사용자의 위치와 취향을 반영

맛집 / 카페 / 장소

대체장소 추천

- Crawling
- KR-WordRank
- 코사인유사도
- CNN



구현모델

Crawling

검색 필터

평점순 인기순

가격/1인당

만원미만 1만원대 2만원대 3만원대

중복 선택 가능



지역

서울-강남 서울-강북 경기도 인천 대구 더보기 >

가로수길 강남역 강동구 개포/수서/일원

관악구 교대/서초 구로구 금천구

중복 선택 가능

음식종류

한식 일식 중식 양식

세계음식 뷔페 카페 주점

중복 선택 가능



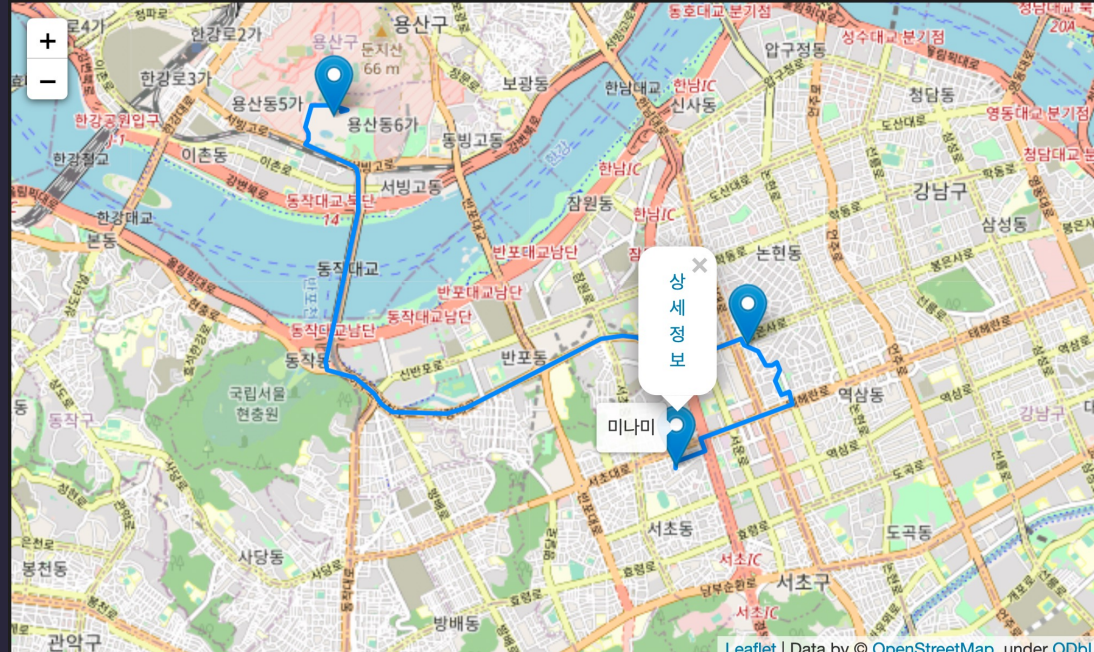
구현모델

거리, 이동시간(driving), 택시비 추출

NAVER-api
(Direction 5)

```
distance : 10334 meters  
time : 1856841 milisecond(1/1000초)  
fare : 12700 원
```

Folium





구현모델

WordRank

- graph-ranking
- 일본어, 중국어
- ranking이 높은 마디 -> 단어 , KeyWord
- 가정 : 단어의 좌/우에는 단어
substring의 좌/우에는 substring
단어는 질 좋은 links가 많이 연결



KR-WordRank

- 토큰나이저(X)
- 띄어쓰기 정보를 이용
- $L + [R]$
- L part 사전화
- 상위 rank R은 조사나 어미
- 'unsupervised' Korean 'keyword' extraction



구현모델

KR-WordRank

```
keyWORDS('신토불이떡볶이')
```

✓ 0.5s

Output exceeds the [size limit](#)

scan vocabs ...

num vocabs = 886

done = 10

떡볶이: 14.5042

핫도그: 6.9355

너무: 5.8616

맛있: 4.4467

맛이: 4.4172

매운: 3.2561

진짜: 2.7016

생각: 2.3637

그냥: 2.1400

```
keyWORDS('우마이도')
```

✓ 0.1s

Output exceeds the [size limit](#)

scan vocabs ...

num vocabs = 726

done = 10

라멘: 7.4586

맛있: 4.9690

돈코츠: 4.1469

국물: 3.7293

건대: 3.0668

너무: 2.9868

차슈: 2.8138

있는: 2.4945

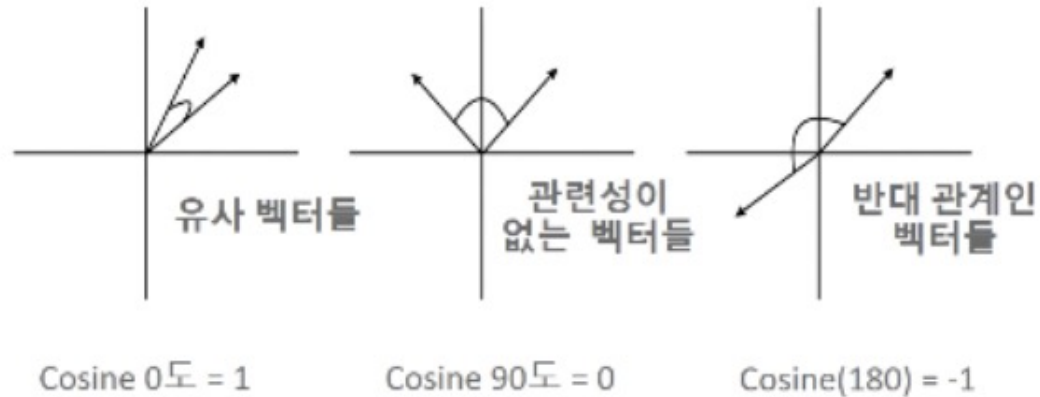
조금: 2.3931



구현모델

- 중복되는 단어가 많을수록 높은 점수
- 문장 계량화 = 피쳐 벡터화
- 유사한 단어들은 1, 관련성이 없는 단어들은 0

코사인 유사도





구현모델

코사인 유사도

```
place_simi_co = (  
    place_simi_cate * 0.3 # 카테고리 유사도  
    + place_simi_review * 1 # 리뷰 텍스트 유사도  
    + np.repeat([review['point']], len(review['point']), axis=0) * 0.005  
    # 평가가 얼마나 좋은지  
    + np.repeat([review['qty'].values], len(review['qty']), axis=0) * 0.001  
    # 리뷰가 얼마나 많은지  
)
```



구현모델

코사인 유사도

```
find_simi_place(review, place_simi_co_sorted_ind, '아찌떡볶이', 3)
```

✓ 0.4s

Python

7	골목에골목	깔끔하고 맛있고 양도 푸짐하다사장님 고생하시는거 같아 마...	한식
23	뉴욕떡볶이	년째 자주가는 동네떡볶이집가격도 착하고 맛있고 최고예요단골집 잘되길 바라는마...	한식
4	콩이네	이사온 첫날 점심은 짜장면인데 집에서 먹기 싫어서밖으로 나와 망플을 컵니다워...	한식

```
find_simi_place(review, place_simi_co_sorted_ind, '서북면옥', 3)
```

✓ 0.1s

Python

31	태천면옥	지금 능라도강남점삼성동이 역삼동에 있을때 그곳에서 주방 면장을 하셨던 분이 ...	한식
40	오징어상미수산	연어회가 꽃처럼 피었습니다신선하고 윤기좌르르 흐르는 도톰한 연어를가성비 좋은...	일식
85	중경삼림	무나리뷰 둘 다 쥐긴 쉽지않을겉건대 커먼그라운드 층에 위치한 중경삼림은 오픈...	중식



CLIP

(by open api)

- CLIP은 OpenAi에서 발표한 최초의 비전 텍스트 multimodal
- 이미지와 텍스트를 모두 인풋으로 사용 가능
- 텍스트는 한 단어 이상도 사용 가능
- zero-shot learning
- domain/distribution shift에 강점



구현모델

텍스트 쿼리를 이용해서 이미지 매칭 가능

CLIP

(by open api)

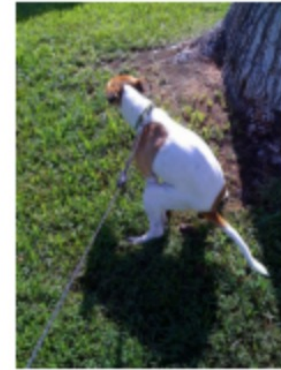
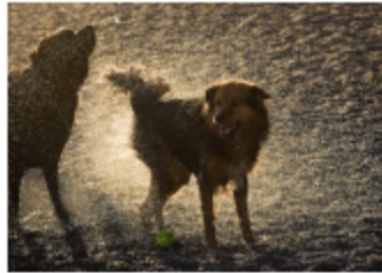
```
find_matches(model,  
             image_embeddings,  
             query="one dog sitting on the grass",  
             image_filenames=valid_df['image'].values,  
             n=9)
```



구현모델

텍스트 쿼리를 이용한 이미지 매칭 결과

CLIP
(by open api)





문제점

CLIP
(by open api)

- Image Dataset
(“Naver 가볼만한 곳” 에서 명소 목록 크롤링)
- 한글 텍스트 적용 불가
(KoClip 한국어 가능; 아직 사용 가능 확인 못함)



Place 추천





향후 진행 방안

- 크롤링
- 알고리즘 성능
- 알고리즘을 연결 시켜서 구현
- 사용자 인터페이스

감사합니다