

ZERO-BASE DS School

DL Project



위치와 사용자 취향을 반영한 데이트 경로 추천 시스템

2조 이동희, 이슬기, 장문정



목차

1. 주제

2. 사전자료

3. 프로젝트 결과 도출과정

4. 한계점, 개선방안

5. 참고자료



주제

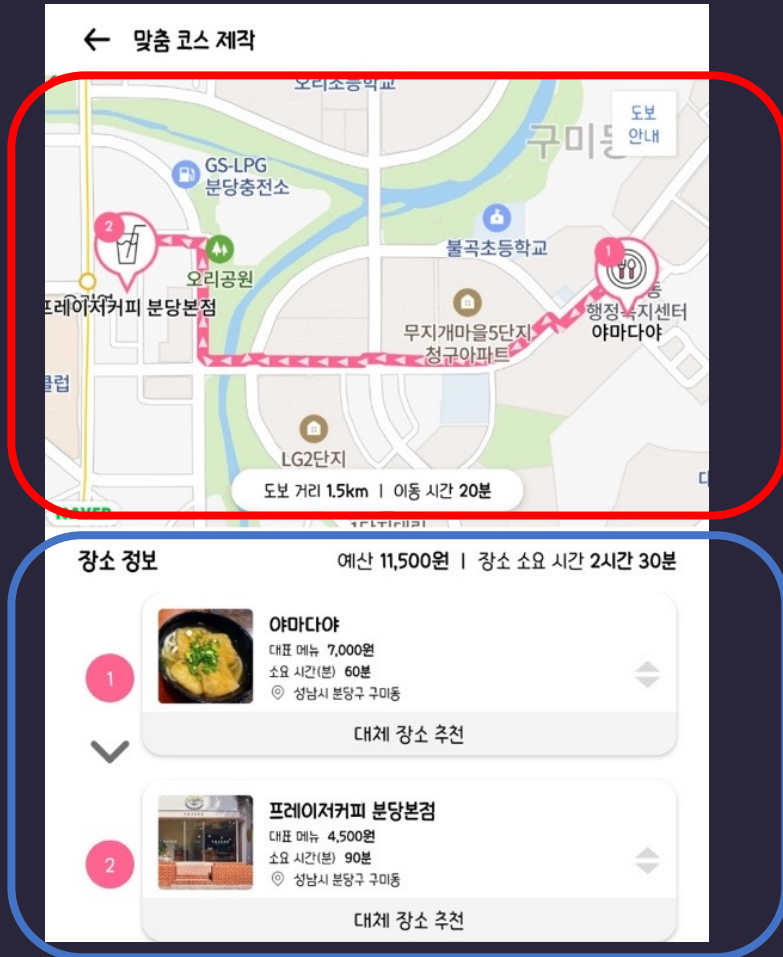
사용자 위치, 취향을 반영

맛집, 카페를 경유하는 경로 추천 시스템 개발



사전자료

모바일앱 '완벽한 하루'



지도시각화

현재위치 - 식당 - 카페

driving 경로

거리(km), 택시비

Place 추천

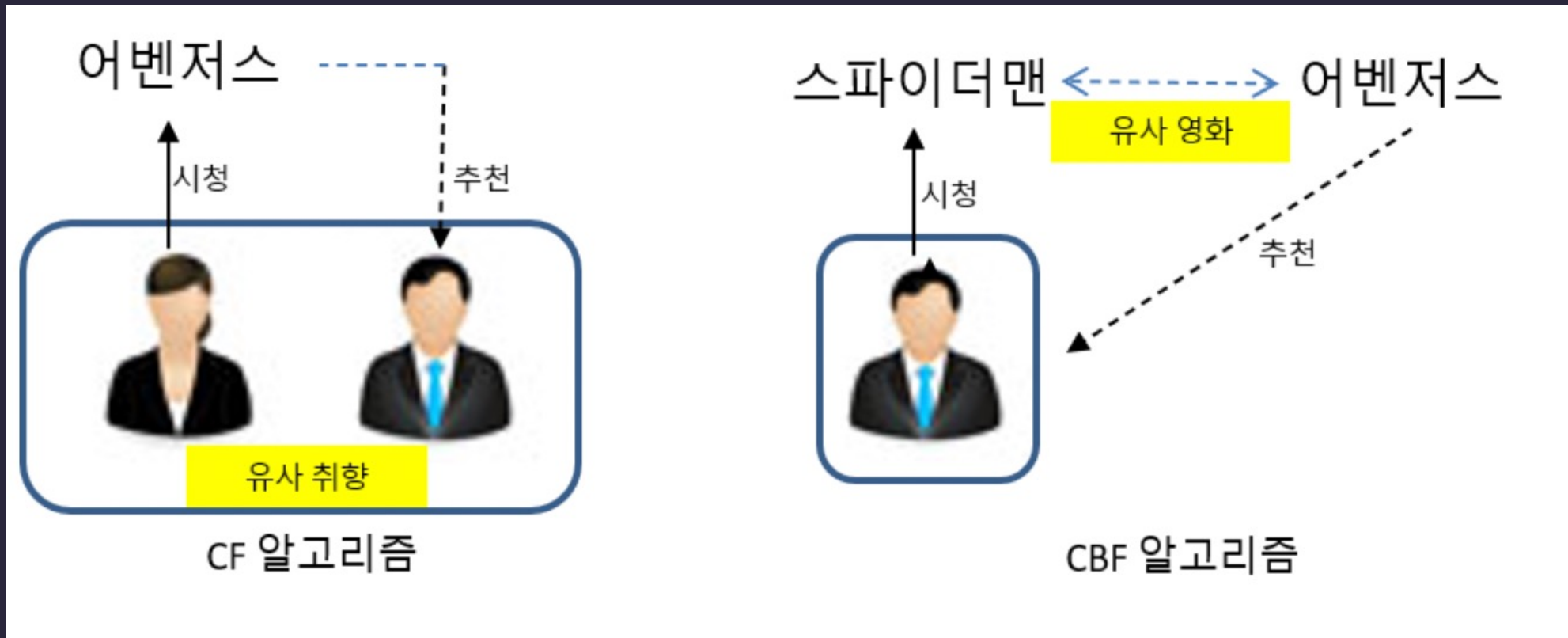
사용자의 위치와 취향을 반영

맛집 / 카페 / 장소



구현모델

Hybrid model = Contents-based Filtering + Collaborative Filtering





NCF 모델 학습(강남구)

분야별 데이터수(유저수, 아이템 수, 평가 수)와 학습 방법

	Restaurant	Cafe
Total users	4181	2588
Total Items	773	874
Total Ratings	17135	11294

모델 공통 파라미터

MLP layers = [64, 32, 16, 8]

Activation function = tanh

Optimizer = adam

Call back = 5 patience



Random search hyper parameter

Random search 시행 후 최적의 파라미터 결과

```
# RandomSearch
rs_neumf = RandomSearch(
    model = neumf,
    space=[
        Discrete("num_epochs", [50, 100, 150, 200]),
        Discrete("num_factors", [4, 8]),
        Discrete("batch_size", [128, 256, 512]),
        Continuous("lr", 0.001, 0.01)
    ],
    metric = fm,
    eval_method = ratio_split
)
```

Restaurant

```
print('Random search: ', rs_neumf.best_params)
```

Random search: {'batch_size': 512, 'lr': 0.0057864482838717955, 'num_epochs': 50, 'num_factors': 8}

Cafe

```
print('Random search: ', rs_neumf.best_params)
```

Random search: {'batch_size': 128, 'lr': 0.0075614473664563754, 'num_epochs': 200, 'num_factors': 4}



NCF (Neural Collaborative Filtering) model

Restaurant 결과

...		F1@5		NDCG@5		Precision@5		Recall@5		Train (s)		Test (s)
-----	+	-----	+	-----	+	-----	+	-----	+	-----	+	-----
NeuMF		0.3284		0.7045		0.2745		0.7057		783.9693		6.2789

Café 결과

	F1@5	NDCG@5	Precision@5	Recall@5	Train (s)	Test (s)
-----	+	-----	+	-----	+	-----
NeuMF	0.4784	0.9779	0.4306	0.9024	1899.4704	3.5955

	u_id	p_id	ncf_score
0	0	296.0	9.624095e-01
1	0	459.0	9.935983e-01
2	0	54.0	1.490116e-06
3	0	84.0	9.967474e-01
4	0	110.0	7.807910e-04
...
867	2587	563.0	2.980232e-08
868	2587	562.0	0.000000e+00
869	2587	560.0	0.000000e+00
870	2587	558.0	0.000000e+00
871	2587	0.0	0.000000e+00

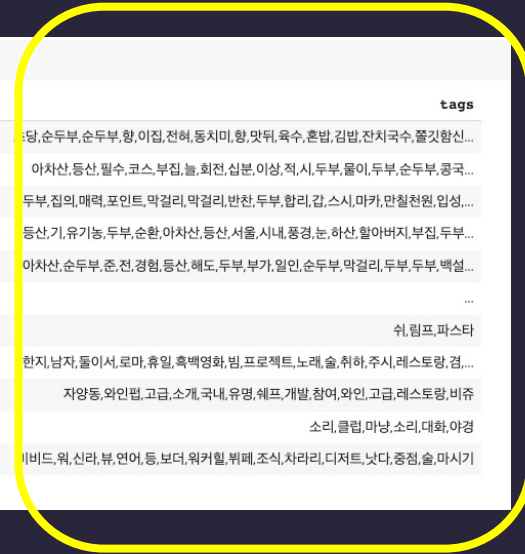
ncf score data



Konlpy를 통해 tag 추출

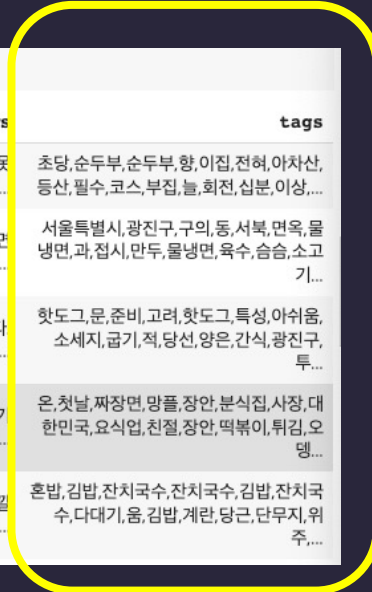
user
matrix

[] eat_review_by_user	
	reviews
715	예전 초당할머니 손두부는 살짝 변해서여길 못따라옴원손두부는 맛과 향이 살짝 비릴수 ...
453	아차산 등산 후 내려오는 길에 필수코스처럼 되어버린 두부집. 늘 대기가 조금씩 있는...
1080	<다른메뉴 없고 손두부만 파는 집>이 집의 매력포인트 중 하나는 막걸리가 정말 싸다...
1479	등산 후 맛보는 기가막힌 #원조할아버지손두부- 직접 만든 유기농 100% 손두부- ...
954	드디어 먹어본 아차산 손두부!! 같이 가 준 친구들이 이미 전에도 맛본 경험이 있어...
...	...
537	오픈매장인거 보고 왔다갔어요!개인적으로 쉬림프파스타 정말 맛있었어요!!또 들들래요 ...
1597	오픈한지 별로 안된 가게라서 가게에 사람들이 많이 없더라고요...사실 들어가기 두려...
603	자양동에 이런 분위기 와인샵이 들어올줄은 몰랐네요 엄청 고급집니다. 소개해보면 국내...
655	소리 너무너무너무 시끄러워요. 클럽마냥 소리질러야 대화가 됩니다~ 아경은 예뻐요.
58	비비드아워로 방문 종류는 아주 다양하다거나 신라 정도의 퀄리티는 아니지만 뷰가 예쁘...
1730 rows x 2 columns	



item
matrix

eat_info									
	name	address	category	price	rvw_cnt	rating	p_id	reviews	tags
0	원조할아버지손두부	서울특별시 광진구 자양로 324	한정식 / 백반 / 정통 한식	5000	72	4.5	553	예전 초당할머니 손두부는 살짝 변해서여길 못따라옴원손두부는 맛과 향이 살짝 비릴수 ...	초당,손두부,손두부,향,이집,전혀,아차산,등산,필수,코스,부집,늘,회전,심분,이상,...
1	서북면옥	서울특별시 광진구 자양로 199-1	국수 / 면 요리	5000	94	4.2	383	서울특별시 광진구 구의2동에 위치한 서북면옥에서 물냉면(9천원)과 찹시만두 6조각(...	서울특별시,광진구,구의,동,서북,면옥,물냉면,과,찹시,만두,물냉면,육수,슴슴,소고기,...
2	케도에오르다	서울특별시 광진구 광나루로 416 건희빌딩 1F	기타 한식	5000	40	4.0	78	기본에 충실해서 더 맛있었던 핫도그였습니다 들어가서 음식이 나오기까지는 꽤 시간이...	핫도그,문,준비,고려,핫도그,특성,아쉬움,소세지,곱기,적,당선,양은,간식,광진구,투...
3	콩이네	서울특별시 광진구 긴고랑로9길 28 중곡제일시장 103호	기타 한식	5000	13	4.0	715	이사는 첫날 점심은 짜장면인데 집에서 먹기 싫어서밖으로 나와 망플을 겁니다..1위가...	은,첫날,짜장면,망플,장안,분식집,사장,대한민국,요식업,친절,장안,떡볶이,튀김,오뎅,...
4	건대우동집	서울특별시 광진구 아차산로 201 동성빌딩 1F	기타 한식	5000	9	4.0	45	혼밥하기 좋은곳.김밥도 괜찮고 잔치국수도 깔끔하니 괜찮다.. 잔치국수와 건대김밥 주...	혼밥,김밥,잔치국수,잔치국수,김밥,잔치국수,다대기,음,김밥,계란,당근,단무지,위주,...





리뷰 Word Clustering – word2vec

	열쇠	지상	제시	육질	무제한	서원	밀집	자본주의
0	0.041388	-0.664884	0.873832	0.378289	-0.096717	1.283177	-1.333698	-0.984570
1	0.533276	-0.878393	-0.709686	0.046813	0.674410	1.579470	0.628633	-0.848251
2	-0.710033	-0.303712	0.455485	0.073390	-0.556135	-0.893533	-1.205475	1.239331
3	-0.099551	0.465191	0.312420	0.014859	-0.276005	0.041519	0.494988	-0.533101
4	0.320853	0.119387	-0.933840	-0.227507	0.576302	0.890856	-2.604765	0.441561
...
195	-0.458551	0.640496	-0.457630	0.637339	0.045735	3.372293	1.659957	-1.478075
196	-0.325321	0.046069	0.029208	0.389370	0.277043	-1.130511	-0.656631	1.071058
197	-0.032340	1.024354	0.261130	0.083895	0.888093	-1.475212	0.719932	0.877240
198	0.247157	-0.195966	-0.045529	0.130099	-0.602561	1.656126	1.588254	1.373969
199	0.177988	-0.872825	-0.394103	-0.059823	0.154985	-0.755136	-0.911105	0.486669



	group_21	group_0	group_1	group_2	group_3	group_4	group_5	group_6	group_7	group_8
0	무제한	밀집	기능	망	스톡	구름	조치	요정	오스	쿠
1	단골	수도원	만족	족	옴	핵	수단	도둑	코스	0
2	아이덴티티	사원	최적	과테말라	보로	가열	목적	주인	피아	0
3	바코드	울타리	도록	굴	콜리	농도	조항	바보	칸	0
4	불꽃놀이	탑	초점	파파	부스	함유	사항	천사	폴리	0
5	방탄	잔디	차원	주도	펠	용액	취지	변신	파	0
6	도통	창문	동작	캉	쿨	금속	명분	아빠	아르	0
7	호명	우물	있	몽	한스	분해	명목	외계인	나스	0
8	비바람	벽면	적응	사스	소시	전지	방침	선생님	라스	0
9	여운	산책로	탐색	낭	콥	반응	특권	동료	트리	0
10	웬	안치	균형	상	홀딩	얼음	제도	악마	카스	0

word2vec 으로 임베딩

K-means clustering 으로 word 그룹화
최종 클러스터수 :120



user/item 클러스터별 점수 부여 및 정규화

user
matrix

	group_73	group_0	group_1	group_2	group_3	group_4	group_5
u_id							
559	0	0	0	0	0	0	0
3735	0	0	0	0	0	0	0
1961	2	0	0	0	3	0	0
4100	0	0	0	0	0	1	0
3990	12	0	0	0	0	0	2
...



	0	1	2	3	4	5
u_id						
559	0.000000	0.0	0.0	0.0	0.000000	0.000000
3735	0.000000	0.0	0.0	0.0	0.000000	0.000000
1961	0.004821	0.0	0.0	0.0	0.252345	0.000000
4100	0.000000	0.0	0.0	0.0	0.000000	0.116959
3990	0.005080	0.0	0.0	0.0	0.000000	0.000000
...

item
matrix

	group_73	group_0	group_1	group_2	group_3	group_4	group_5
p_id							
0	3	0	8	0	1	1	0
1	10	0	10	3	1	0	1
2	65	0	28	6	0	5	4
3	31	0	14	0	0	5	1
4	32	0	15	1	2	3	2
...



	0	1	2	3	4	5
p_id						
0	0.001151	0.0	0.010374	0.000000	0.013388	0.013760
1	0.002574	0.0	0.008699	0.022153	0.008981	0.000000
2	0.002991	0.0	0.004354	0.007920	0.000000	0.008250
3	0.002100	0.0	0.003205	0.000000	0.000000	0.012146
4	0.003296	0.0	0.005221	0.002955	0.007187	0.011080
...



CBF 점수 산출

user matrix

	0	1	2	3	4	5	6	7	8	9
0	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000
1	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000
2	0.004821	0.0	0.0	0.0	0.252345	0.000000	0.000000	0.0	0.061629	0.070733
3	0.000000	0.0	0.0	0.0	0.000000	0.116959	0.000000	0.0	0.041689	0.095694
4	0.005080	0.0	0.0	0.0	0.000000	0.000000	0.020243	0.0	0.027057	0.008281
...
4176	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000
4177	0.018337	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.358685
4178	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.222418
4179	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000
4180	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000

두 벡터간의 cosine similarity 측정

4 0.005080 0.0 0.0 0.0 0.000000 0.000000 0.020243 0.0 0.027057

place matrix

	0	1	2	3	4	5
0	0.001151	0.0	0.010374	0.000000	0.013388	0.013760
1	0.002574	0.0	0.008699	0.022153	0.008981	0.000000
2	0.002991	0.0	0.004354	0.007920	0.000000	0.008250
3	0.002100	0.0	0.003205	0.000000	0.000000	0.012146
4	0.003296	0.0	0.005221	0.002955	0.007187	0.011080
...
768	0.000662	0.0	0.006717	0.000000	0.000000	0.000000
769	0.003246	0.0	0.000000	0.000000	0.000000	0.000000
770	0.013393	0.0	0.000000	0.192109	0.000000	0.000000
771	0.006584	0.0	0.000000	0.000000	0.000000	0.000000
772	0.013287	0.0	0.000000	0.000000	0.000000	0.000000

preference

0	0.016090
1	0.016106
2	0.011974
3	0.004561
4	0.005723
...	...
768	0.005259
769	0.001718
770	0.000000
771	0.015681
772	0.010548

user의 특성에 맞는 item을
탐색 후 ranking



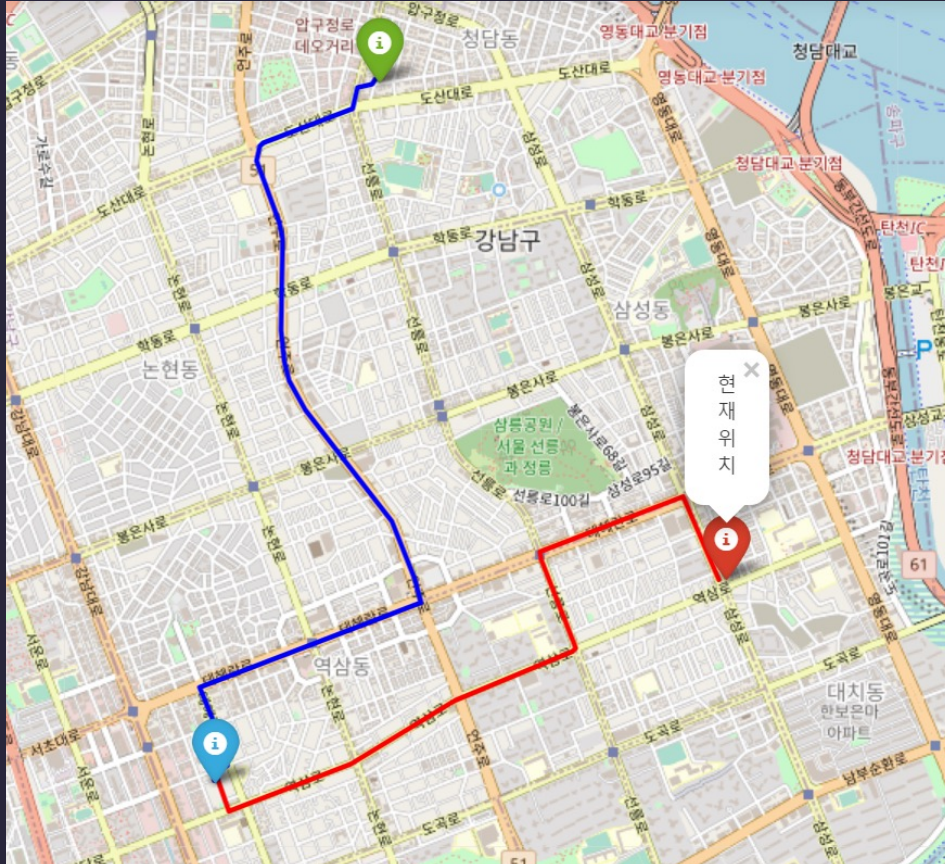
hybrid model 결과

	p_id	name	address	category	price	cbf_score	ncf_score	result
494	15	가가솔밥	서울특별시 강남구 역삼로5길 19 1F	한정식 / 백반 / 정통 한식	15000	0.001246	0.062523	0.000078
15	37	강승월한식당	서울특별시 강남구 논현로175길 61 2F	한정식 / 백반 / 정통 한식	5000	0.000688	0.071217	0.000049
129	10	Lu307	서울특별시 강남구 압구정로10길 30-7 1F	이자카야 / 오뎅 / 꼬치	15000	0.001880	0.025679	0.000048
663	68	구구당	서울특별시 강남구 강남대로102길 35	퓨전 양식	15000	0.001384	0.033634	0.000047
144	16	가로수길 소소	서울특별시 강남구 강남대로154길 41	이자카야 / 오뎅 / 꼬치	25000	0.000146	0.135507	0.000020
368	75	그 우동집	서울특별시 강남구 역삼로7길 7	기타 한식	5000	0.000468	0.027335	0.000013
117	153	덴푸라키이로	서울특별시 강남구 논현로153길 51 102호	돈부리 / 일본 카레 / 벤토	15000	0.000142	0.076490	0.000011
192	425	신사완쯔	서울특별시 강남구 강남대로154길 11 애경빌딩	퓨전 중식	15000	0.000541	0.011556	0.000006
767	2	29펍	서울특별시 강남구 강남대로110길 14	치킨 / 호프 / 펍	25000	0.000087	0.066681	0.000006
646	435	써브웨이	서울특별시 강남구 강남대로96길 12 지천빌딩 1F	브런치 / 버거 / 샌드위치	5000	0.000579	0.009117	0.000005





Folium 지도시각화, Naver api 경로탐색, 최종 추천



107번님, 강남구에서 데이트 중이신가요? 혹시 '가가솔밥'에서 식사 어떠세요?
이후에 '버터핑거팬케이크'에서 커피 한잔 어떠세요?

'가가솔밥'

주소 : 서울특별시 강남구 역삼로5길 19 1F

종류 : 한정식 / 백반 / 정통 한식

예상가격 : 15000

'버터핑거팬케이크'

주소 : 서울특별시 강남구 선릉로152길 11

종류 : 카페 / 디저트

예상가격 : 5000

'현재위치' 에서 '가가솔밥' 까지

예상 거리 : 1089m

예상 택시비 : 3800원

'가가솔밥' 에서 '버터핑거팬케이크' 까지

예상 거리 : 4613m

예상 택시비 : 6000원



한계점, 향후 개선방안

- 전체데이터의 성능 확인
- 가상의 107번 user 지정
- 새로운 user 추가 및 만족도 설문



한계점, 향후 개선방안

데이터 확보를 위한 설문지 작성

- CBF : 선택한 명사가 들어있는 클러스터에 점수를 부여
 - 분야별 cluster를 대표하는 단어를 도출하여 1-5점의 점수 설문
- NCF : 사용자가 소비해본 아이템 조사
 - 맛집, 카페 각각 상위 10개의 장소를 골라 방문 유/무와 만족도 설문
- 추천 받을 행정구역 선택



참고자료

- **NCF, CBF**

[__https://github.com/DAC-KHUPID/seoul-date-course-recommendation](https://github.com/DAC-KHUPID/seoul-date-course-recommendation)

- **Naver API Direction5**

https://velog.io/@choonsik_mom/Naver-API%EB%A1%9C-%EA%B8%B8%EC%B0%BE%EA%B8%B0with-python

감사합니다