



Instagram

@정유경 @김연지 @강창균

#이태원 맛집

#맛스타그램

#이태원

#ITEWON



실제 상권과 인스타그램 맛집의 위치 차이



Instagram



프로젝트 개요



프로젝트 개요

- 프로젝트 배경
- 프로젝트 목표

분석 계획 수립

- 데이터 구성도
- 분석 환경 및 사용기술
- 분석 절차 및 UI구성

데이터 분석 및 결과

- EDA
- 분석 방법
- 결론 및 시사점

프로젝트 마무리

- 한계점 및 개선방향
- Q & A





INSTAGRAM



Insta_matjip

...

최근 한달 내 이용 및 주 이용 소셜미디어

■ 주 이용 ■ 최근 한달 내 이용	전체						Gap 주 이용 (20-19)	성별		연령별				
	2018년		2019년		2020년			남	여	10대	20대	30대	40대	50대
	Base	(500)	(500)	(500)	(310)	(310)								
유튜브	27.6	78.8	32.6	83.8	37.6	89.2	+5.0	49.7	34.2	60.0	42.7	38.9	33.9	34.9
네이버 블로그	17.0	65.8	21.8	71.8	22.2	78.2	+0.4	15.2	21.3	1.7	12.9	19.0	29.8	27.0
인스타그램	14.8	48.4	19.2	53.0	19.0	58.6	-0.2	11.6	26.1	18.3	28.2	27.8	12.9	7.1
페이스북	15.6	60.2	9.0	55.2	7.0	53.0	-2.0	12.3	4.8	15.0	8.1	7.1	4.0	8.7
밴드	11.2	50.4	6.2	45.4	6.0	47.2	-0.2	6.1	3.5	0.0	1.6	4.8	4.8	12.7

평소 자주 보거나 공유하는 게시물

	전체		성별		연령별				
	남 (174)	여 (209)	10대 (90)	20대 (93)	30대 (89)	40대 (58)	50대 (53)		
맛집/음식	36.8	45.5	25.6	52.7	40.4	43.1	49.1	41.5	
여행	31.0	26.3	20.0	33.3	21.3	29.3	45.3		
일상생활	25.3	30.1	34.4	28.0	29.2	20.7	22.6		
패션/뷰티/잡화	16.7	33.0	31.1	24.7	27.0	27.6	13.2		
리빙/라이프스타일	14.9	23.0	4.4	19.4	27.0	24.1	26.4		
TV/연예	19.0	17.2	27.8	20.4	10.1	20.7	7.5		
뷰티	2.9	23.0	22.2	15.1	10.1	12.1	5.7		
동물	10.3	15.8	17.8	8.6	16.9	12.1	9.4		
스포츠	21.8	3.3	12.2	8.6	13.5	12.1	13.2		
유머/개그	13.2	9.1	14.4	15.1	9.0	6.9	5.7		
음악/유흥비디오	13.2	8.6	18.9	10.8	10.1	1.7	7.5		
운동/건강	12.6	8.6	6.7	8.6	13.5	12.1	13.2		
도서/공연/문화	6.3	9.1	7.8	7.5	5.6	12.1	7.5		
육아/교육	5.2	10.0	2.2	4.3	21.3	6.9	1.9		

▲ [Base : 최근 1개월 내 인스타그램 이용자, N=383, 단위 : %, 순위형 응답(1+2+3순위)] * 5% 이하인 경우 제시하지 않음 * 하늘색 음영 : 평균 대비 +5%p 이상인 데이터

20-30대가 많이 이용하는 SNS는 인스타!
인스타에서 자주 보거나 공유하는 게시물 1위가 맛집과 음식 관련 정보이고, 사진을 올리기 위해 맛집을 찾아다니는 경우가 있을 정도이다.
진짜 맛집 데이터는 인스타에 더 많이 있지 않을까? 하는 생각에서 시작하게 되었다.

하지만 실제 상권과 사람들이 자주 가는 인스타 맛집 장소와 차이가 있다는 것을 알게되었다.
길 건너 가면 가게가 있는데 지리적인 문제로 상권으로 포함되지 않은 경우가 많다.
이로 인해 소비자, 공급자(상인), 배달 앱 등 지리적인 데이터를 사용하는 사람들의 불편도가 높다.
가지고 온 인스타 맛집 데이터를 이용해서 실제 상권과 비교해보면 어디가 다른지 알 수 있지 않을까?

좋아요





이태원상권



그림 1 이태원 해시태그/위치정보를 활용한 상권지도

A구역 : 현재 #이태원맛집, #이태원핫플 등의 가장 많은 해시태그가 밀집된 지역

B구역 : #한남동맛집 이라는 신규 상권으로 #이태원디저트 #이태원브런치 등의 해시태그 밀집

C구역 : #우사단길이라는 떠오르는 상권으로 이태원상권과 이어지는 트렌드 상권으로 부상

사람들이 자주가는 인스타 맛집 데이터와 나라에서 지정한 상권구역이랑 어디가 다른지 확인하여 실질적인 맛집 지도를 만들기

인스타 데이터를 이용해서 비즈니스적으로 이용할 수 있는 방안 도출해보기

다른 목표...?

좋아요

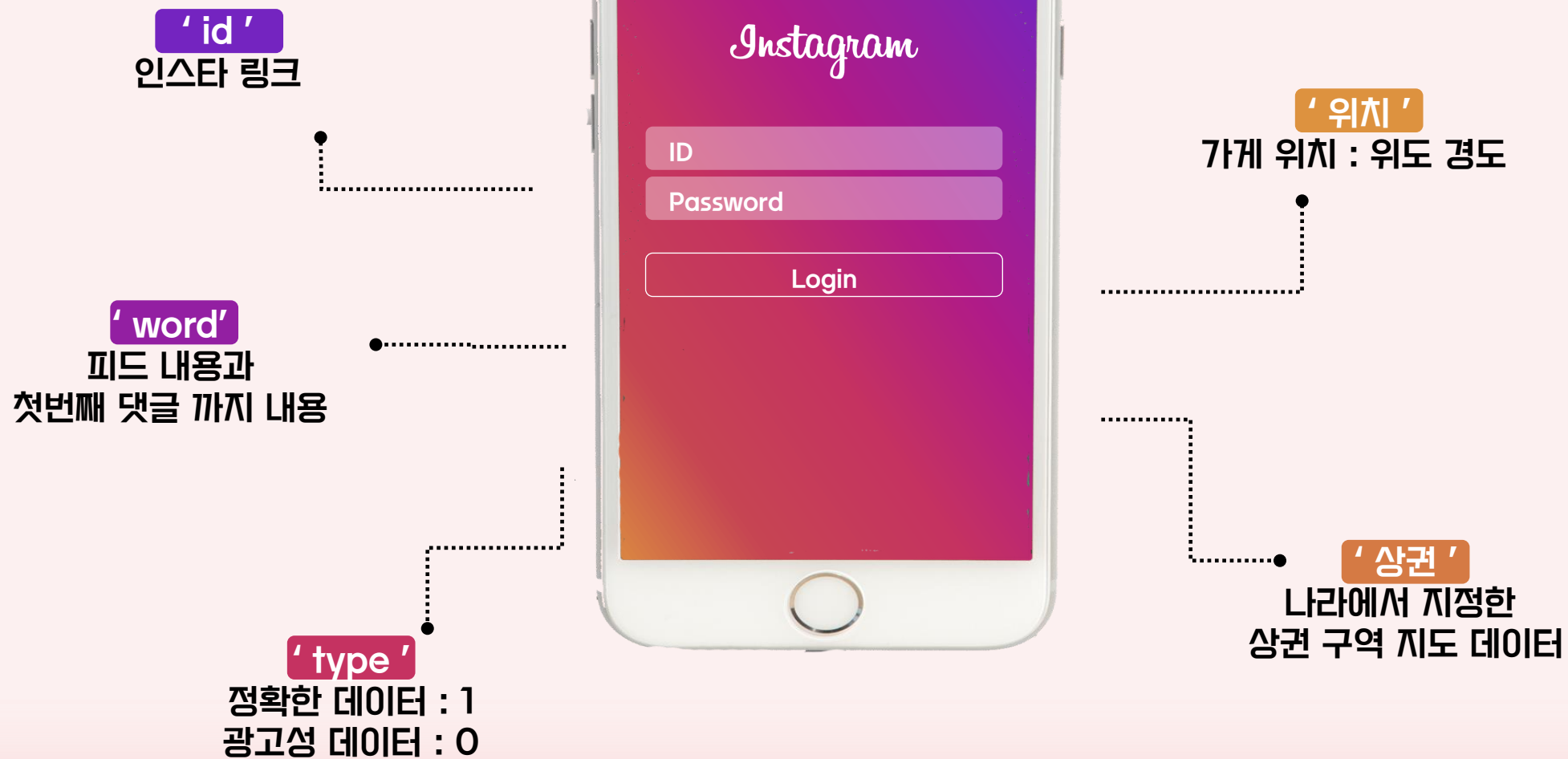




Instagram



분석 계획 수립



좋아요





분석툴



시각화툴



Folium



좋아요





데이터 수집

전처리

EDA

모델 학습

실제 상권
비교

위치 데이터

좋아요





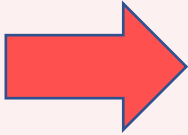
Instagram



데이터 분석 및 결과

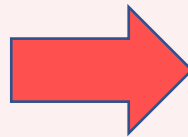


멘토님



Instagram

DATA 공공데이터포털
.GO .KR

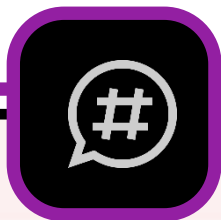


실제 상권 데이터





“#” 해시태그
빼주기



이모티콘
빼주기
EX)
<U+653C>



특수기호
들어간 것들
빼주기







total



전체 데이터
이태원
이태원맛집
먹스타그램
우사단길
이태원카페
일상



type1



정확한 데이터
이태원
이태원맛집
이태원카페
먹스타그램
한남동맛집
경리단길카페



type0



광고 데이터
이태원
우사단길
이태원맛집
일상
맛팔
힐링

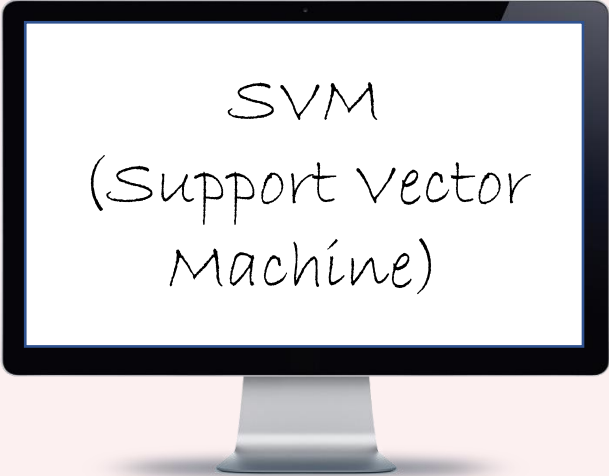




노션 6/2 :
전체 데이터
라벨1값 영향
력이 큰 단어
/ 라벨0값 영
향력이 큰 단
어

좋아요





SVM
(Support Vector
Machine)

accuracy : 0.8905
precision : 0.8942
recall : 0.9395
f1 score : 0.9163



Naïve Bayes
Classifier

accuracy : 0.8770
precision : 0.8820
recall : 0.9320
f1 score : 0.9063

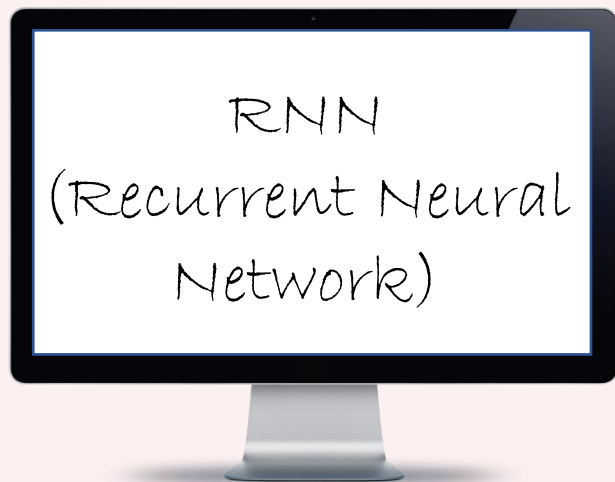


Logistic
Regression

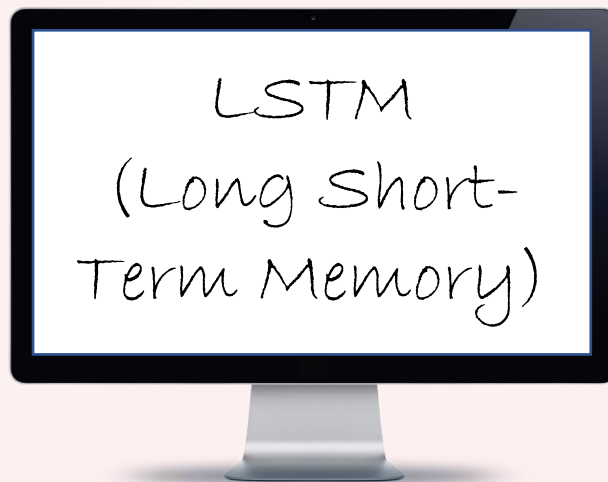
accuracy : 0.8848
precision : 0.8911
recall : 0.9335
f1 score : 0.9118

좋아요

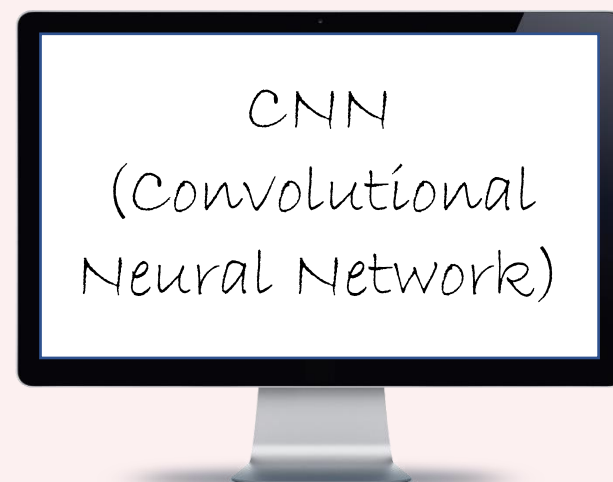




accuracy: 0.8682
precision: 0.8938
recall: 0.9008
f1 score: 0.8948



accuracy: 0.8875
precision: 0.8872
recall: 0.9414
f1 score: 0.9117



accuracy: 0.8907
precision: 0.9007
recall: 0.9302
f1 score: 0.9132

좋아요





노션 6/13 :
모델 돌려본
후 test데이터
라벨1값 영향
력이 큰 단어 /
라벨0값 영향
력이 큰 단어

노션 6/10:
모델 돌려본
후 틀린 10%
이유

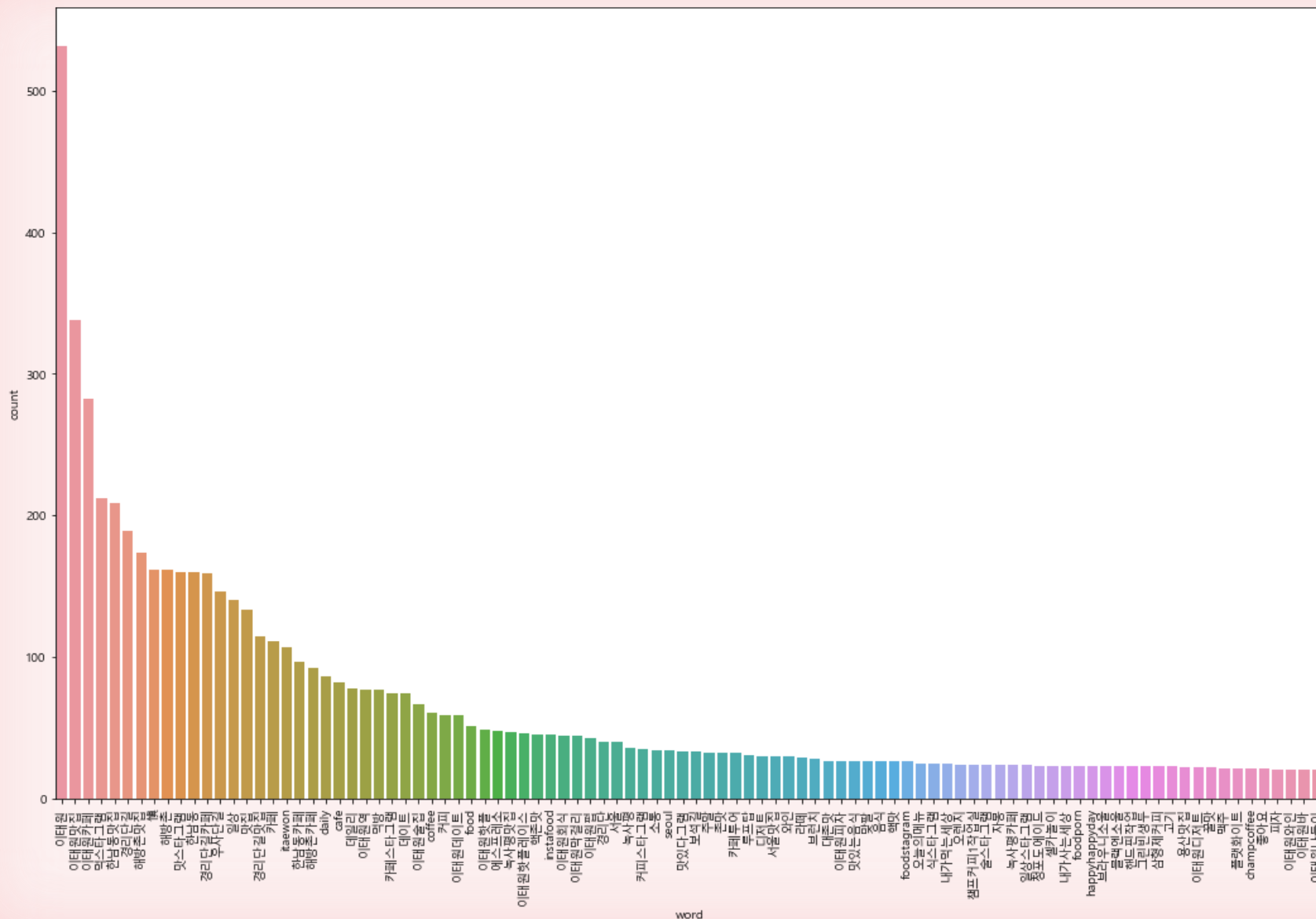
노션 6/3 :
임계값??





EDA

Page. 01



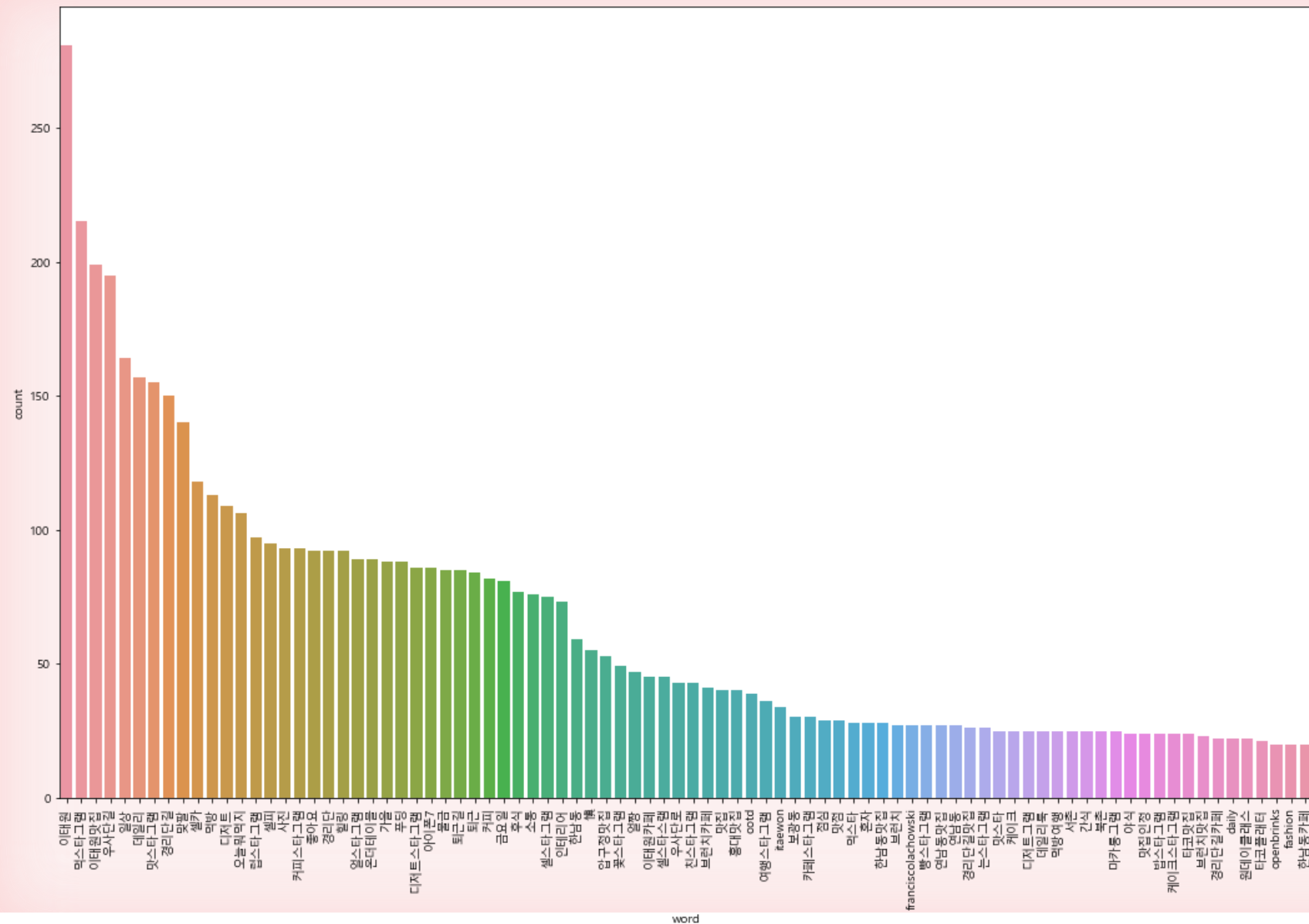
Instagram

Test type 1
예측 맞춘
해시태그
키워드



좋아요





Instagram

Test type 0
예측 맞춘
해시태그
키워드



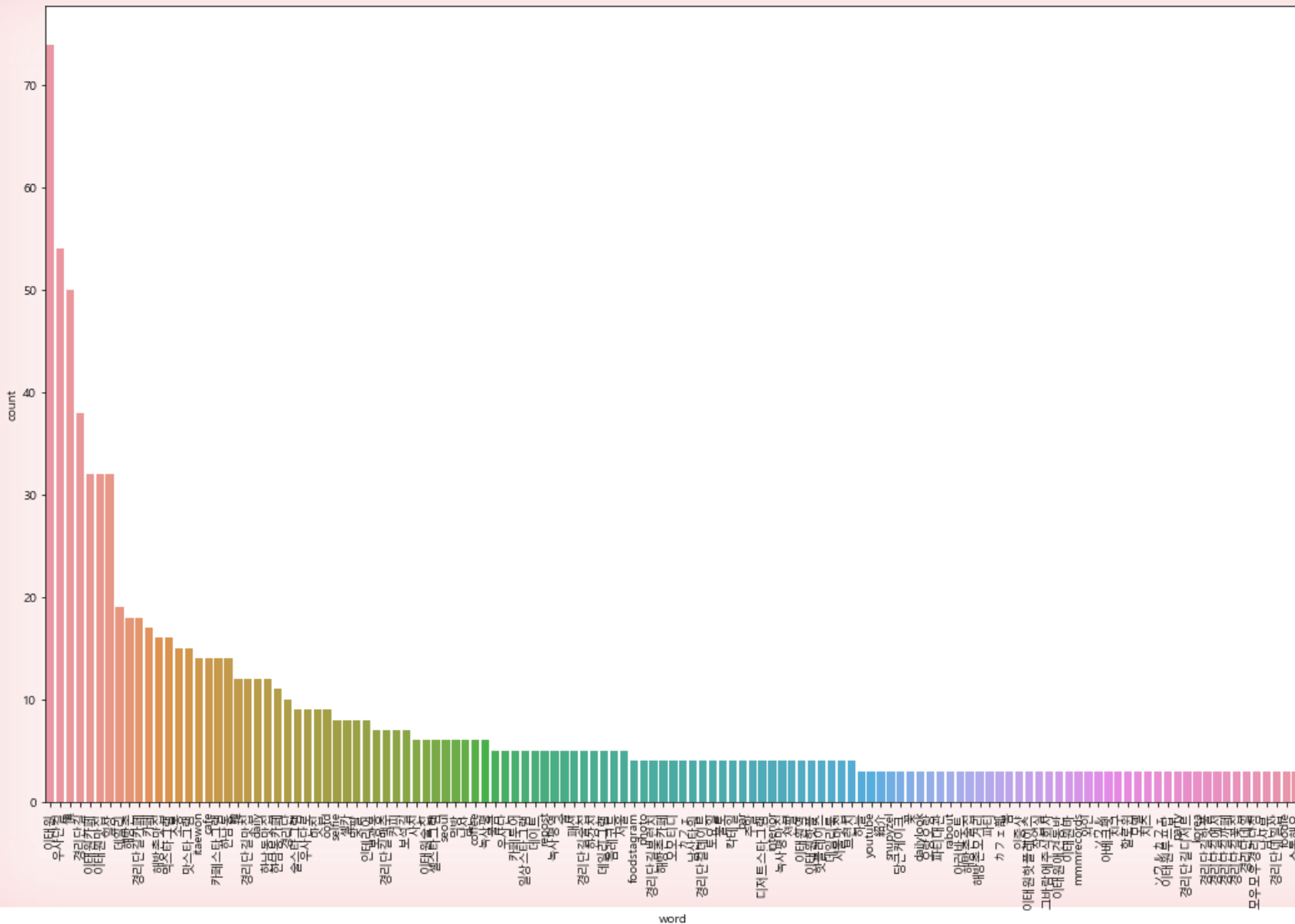
좋아요





↳ 해시 태그 내용과 실제 내용이 다르다





Instagram

Test type 1
예측 0
해시태그
키워드



↳ 해시 태그 내
용과 실제 내용이
다르다

좋아요





가장 잘 나온 모델
설명 / 토큰화/ 벡터
화 설명 / 단어유사
도 사용하지 않은 이
유 설명 -> 멘토님
한테 여쭙보기 / 평
가지표이유

좋아요





결론(실제 상권과 인스타그램 맛집 비교)

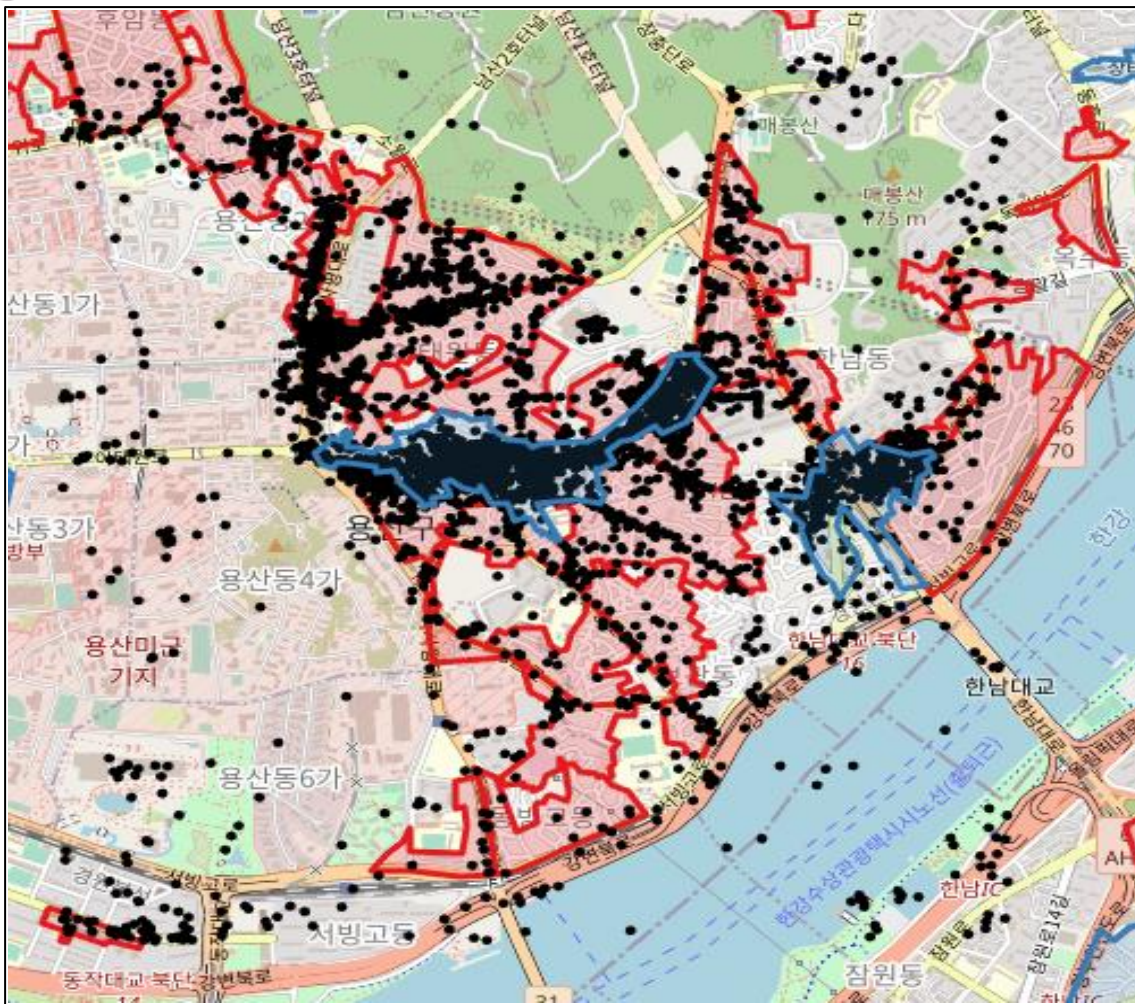


INSTAGRAM



Insta_matjip

...



***공데 상권 지도와 인스타 지도 데이터의 차이점**

1.공데 상권데이터에서 아파트 단지, 병원, 학교, 박물관, 공군기지, 공원 내 맛집들의 정보가 빠짐(인스타는 포함되어 있다)

2.한강 공원과 한강에서 배달하고, 해당 맛집을 태그 했기때문에 한강과 그 건너편 데이터까지 찍힌 것으로 판단된다

3.호텔에 맛집 찍힌 이유는 호텔에서 룸서비스, 호텔 내 맛집을 간 후 태그해서 데이터에 나타난 것으로 판단된다

4.서빙고역 쪽 데이터 - 기간 차이가 있음(현재 공사 중인 맛집들)

좋아요





비즈니스적으로
활용시 가정
재현율 이용
-> 멘토링





Instagram



프로젝트 마무리



한계점 및 개선방향

Page. 01





팔로잉

내 소식



2_조님이 회원님을 팔로우 하기 시작했습니다.

팔로잉



2_조님이 댓글에서 회원님을 언급했습니다 :

@insta_matjip **Q&A 시간입니다**



2_조님이 댓글을 남겼습니다 :

질문이 있으시다면 질문해주세요



MEDICI님이 회원님을 팔로우 하기 시작했습니다.

팔로우



Instagram

**THANK
YOU**



#메디치

#수고하셨습니다:D