

인공지능응용 2조

SNS 계정을 활용한 MBTI예측

大大大大

201627539 이현부
201727523 박산하
201727542 정유나
201727528 표지원
201827559 서아림





0. 조원 소개

MBTI



이현부(ENTJ)

인스타그램 크롤링
자연어 처리



박산하(ESTP)

EDA
모델링



정유나(ISFJ)

데이터 시각화
모델링



표지원(ENFP)

EDA
모델링



서아림(ENFP)

인스타그램 크롤링
PPT 제작

목차

- 1 프로젝트 배경 및 필요성
- 2 프로젝트 목적
- 3 프로젝트 방법론 및 데이터
- 4 진행된 내용 및 결과
- 5 기대효과



1. 프로젝트 배경 및 필요성

현대인들의 주 관심사인 **MBTI**

많은 마케팅에 MBTI가 활용되고 있는 상황 → 'MBTI식 마케팅'



영화 홍보 - 롯데엔터테인먼트



제품 홍보 - 공차



MBTI 성향에 따라 주스 추천 - 웅진식품



1. 프로젝트 배경 및 필요성

자주(JAJU)



개인의 생활 유형을 파악해 맞춤형 상품을 제안하는 '일상재질테스트' 진행

→ 온라인 매출이 전년 동기 대비 96% 증가

→ 고객들은 결과 분석과 상품 추천 정확도에 대해 큰 만족도를 나타냄

이렇듯 **개인별 성향**에 따라 마케팅 하는 것이 중요



2. 프로젝트 목적

□ MBTI 검사방법

MBTI 정식 검사 / 무료 성격유형검사(16Personalities)

□ 목적

- MBTI를 마케팅, 트렌드 분석에 활용하기 위해
Instagram의 여러 요소를 독립변수로 설정하여 종속변수인 MBTI 예측
팔로잉 수, 팔로워 수, 게시물 수, 게시물, 해시태그, 댓글 등
 - XGBoost를 활용한 Random Forest 모델에 대한 학습과 자연어 처리(NLP)에 대한 학습
- 현대인들의 성향별 관심사를 분석 가능
- INSTAGRAM의 게시글을 통해 MBTI를 예측할 수 있는 모델을 만들고 이를 마케팅, 트렌드 분석에 적용 및 활용할 수 있음



3. 프로젝트 방법론 및 데이터



1. 데이터 수집 방법론

20~30대 남녀가 응답한 설문조사결과

인스타그램의 어떤 요소들이 MBTI에 영향을 미칠 수 있을지를 파악

20~30대 남녀 Instagram 프로필/게시글

인스타그램 웹 크롤링 후 프로필, 게시글 등의 요소를 모델을 구축하는데 사용

XGBoost 기반의 Random Forest/SVM, 혹은 성능을 최적화할 수 있는 딥러닝 모델



3. 프로젝트 방법론 및 데이터

2. 데이터 수집 - 설문조사 진행

'SNS 활용방식에 따른 MBTI예측' 설문조사

안녕하세요!

저희는 부산대학교 산업공학과 '인공지능융합' 과목에서 'SNS 활용방식에 따른 MBTI예측'을 주제로 팀 프로젝트를 하고있는 대대대대(大大大大)입니다. 사람들의 MBTI에 따라 SNS 활용 방식이 어떻게 달라지는지, 인공지능 모델에 SNS 활용 방식 설문조사 결과를 학습시켜 정확한 모델을 만들 예정입니다.

본인이 인스타그램을 사용할 때 주로 어떤지 생각하며 답해주시면 되고, 마지막에 본인의 MBTI와 인스타그램 ID를 적어주시면 더욱 도움이 됩니다!
예상 시간은 3~5분입니다.

여러분의 정성어린 설문 결과가 많아질수록 더 좋은 모델이 나올 수 있으니 많은 참여 부탁드립니다!

[Google에 로그인](#)하여 진행상황을 저장하세요. 자세히 알아보기

* 필수항목



['SNS 활용방식에 따른 MBTI예측' 설문조사 \(google.com\)](#)

추가 설문조사 진행!

1차: 2021.11.11~16

2차: 2021.11.28~12.10

추가 200여개의 설문조사 결과를 얻음

설문조사 내용 시각화

① 질문별로 설정된 MBTI 요소 시각화

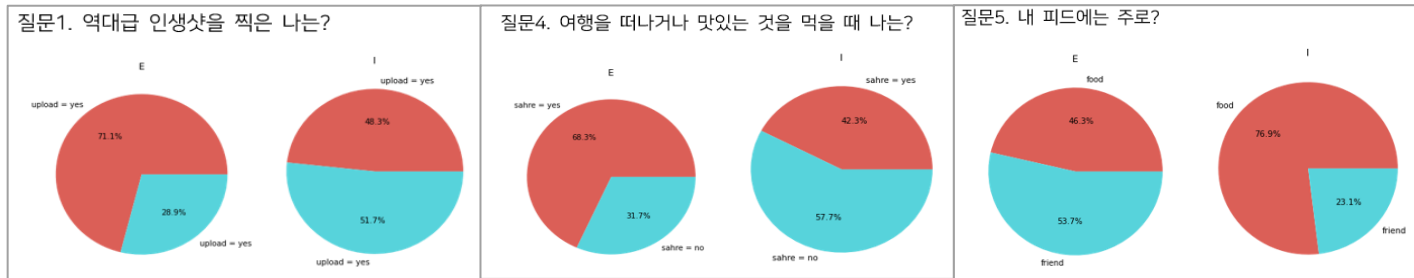
:질문 작성시 생각해둔 MBTI요소로 시각화 진행

MBTI 요소	질문 문항
I/E	1,2,4,5,12번
S/N	7,13,15,16,17,18번
F/T	3,6,8,9,10,11번
J/P	14,19,20번



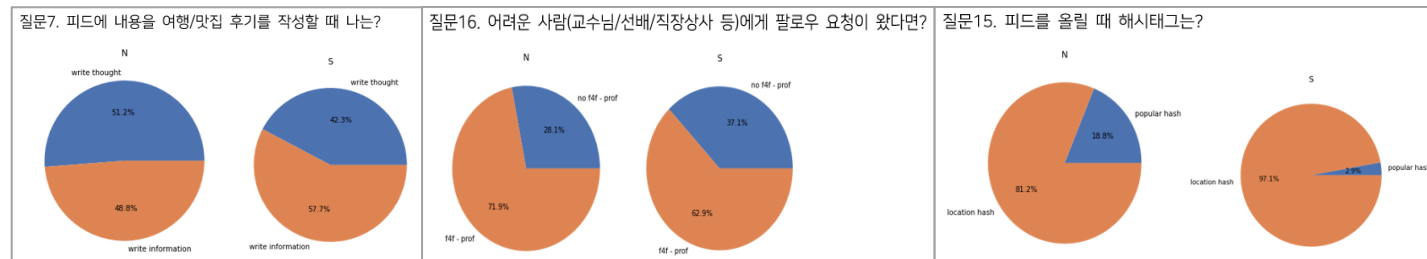
3. 프로젝트 방법론 및 데이터

① 질문별로 설정된 MBTI요소 시각화



<I/E>

- 인스타그램에 게시된 글 수에 따라 I/E 판단가능
 - 피드 사진이 주로 사람/사물·음식인가에 따라 I/E 판단가능
- 게시글 수나 피드 사진을 모델에 활용가능



<S/N>

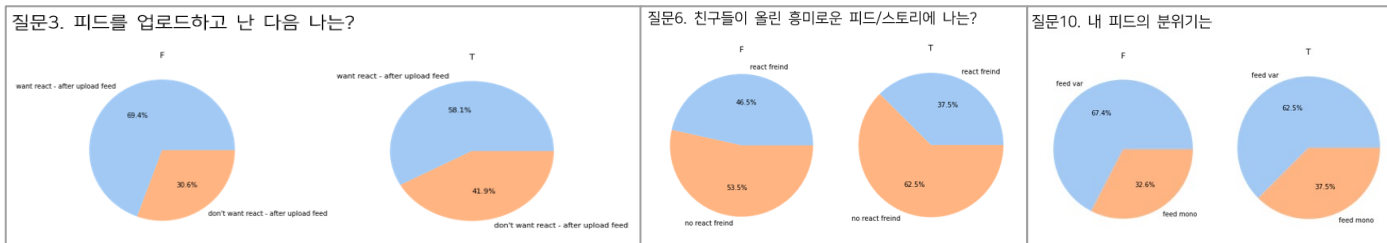
- 피드 내용이 정보/감상인가에 따라 N/S 판단가능
- 해시태그나 팔로우상태는 다른 요소도 고려해볼 필요

피드 내용을 모델에 활용가능



3. 프로젝트 방법론 및 데이터

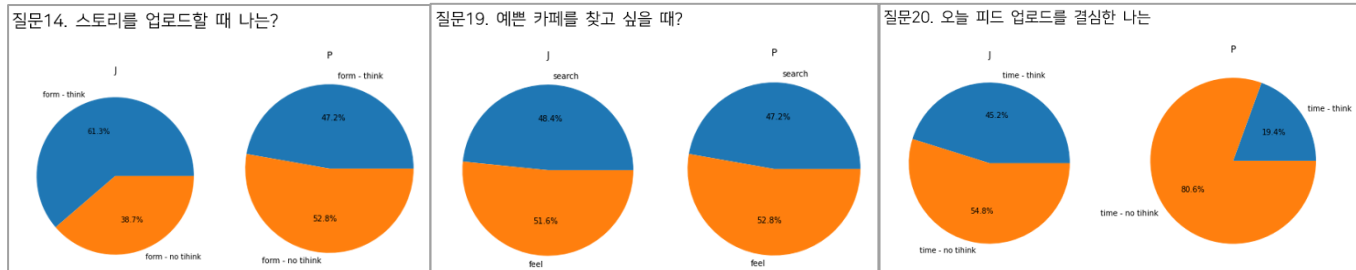
① 질문별로 설정된 MBTI요소 시각화



<F/T>

- 피드 분위기가 정돈/다양한가는 다른 요소도 고려할 필요

F/T를 구분할만한 다른 요소를 추가적으로 고민



<J/P>

- 스토리 업로드 시 구성을 고민하는가에 따라 J/P판단 가능(크롤링 가능한 영역인지 고민)

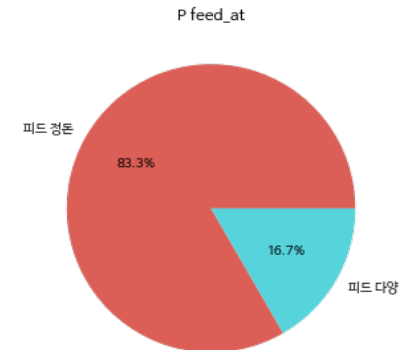
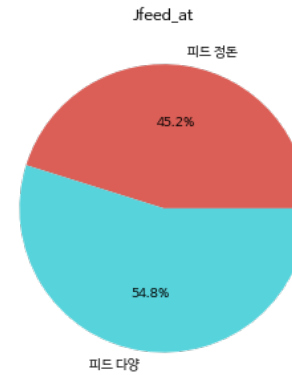
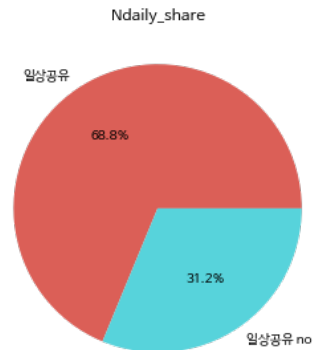
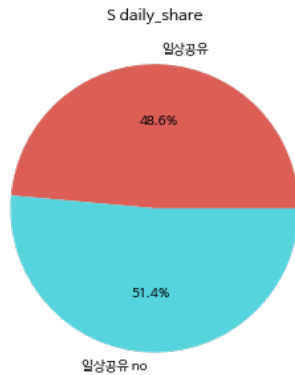
J/P를 구분할 만한 요소 추가적으로 고민



3. 프로젝트 방법론 및 데이터

② 질문별로 전체 MBTI요소 시각화

: 이전과정을 통해 다른 요소를 고려해봐야 하는 질문들이 있음을 파악



4번. 피드 사진이 사람/사물인가
I/E외에도 S/N도 영향을 미칠 수 있음을 파악

10번. 피드가 다양/정돈인가
F/T로는 파악 불가 - J/P로 판단가능성 확인

MBTI의 여러 요소에 의한 교호작용이 있을 수 있음을 고려하고 모델구성 !!



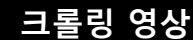
3. 프로젝트 방법론 및 데이터

2. 데이터 수집 - 인스타그램 크롤링

수집 데이터

- 팔로잉, 팔로워 수
- 게시글 수
- 해시태그
- 프로필 바이오
- 게시글 내용 및 날짜
- 좋아요 수
- 비공개 여부

: 팀원들 모두가 크롤링한 후, 차이가 있는 계정을 비공개 계정으로 판단



	name	cnt	follower_ing	bio	content	date	like	place	tags			
0	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic몇년 만의 증명사진인지 모르겠지만...347조는 기망의사 없으니까^-^	2021-09-27	g_ju_님 외 74명이 좋아하니		0		
1	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic2021년 생일 정말 행복했어요! 시간차 공격으로 어마무시하게 선물을 받	2021-06-11	dong_2님 외 111명이 좋아하니		0		
2	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_	2021-01-27	dong_2님 외 69명이 좋아하니		0		
3	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_🚗🚗	2021-01-16	4min님 외 130명이 좋아하니		0		
4	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_🚗🚗	2021-01-08	ey_song님 외 94명이 좋아하니	부산대학교 Pusan National U	0		
5	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_Kairos for me, by me, with me, to me.~ 2020.11.29.사랑합니다. 카이로	2020-11-29	dong_2님 외 85명이 좋아하니		0		
6	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_기대 오로라 레지스레이션 하러 오는 이들도 크르르르르	2020-08-27	dong_2님 외 103명이 좋아하니		0		
7	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_평범하게 자랐습니다.역변해서 그럴지	2020-04-23	dong_2님 외 69명이 좋아하니		0		
8	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_내 전완근은 쓰레기였어-	2020-02-18	dong_2님 외 81명이 좋아하니		0		
9	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_너를 만나 감사했어 짹짹🥰Gratitude to befriend with you guy Jack:-)	2020-02-09	ved_chan님 외 103명이 좋아하니		0		
10	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_And ye shall know the truth, and the truth shall make you free.John8	2020-02-07	ved_chan님 외 101명이 좋아하니		0		
11	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_부시시한 모습으로 내게 지약을 짜 주던 그 기분이 이런 기분이었을까. 2019-04-29	2019-04-29	ved_chan님 외 72명이 좋아하니		0		
12	wiseb0_0	게시물	28팔로워	473팔로우	592	gefährlic_그대도 비비빅을 좋아하고 싶지 않았겠습니까 마음이 곱지 아니한 이를	2019-04-01	alf_moon님 외 58명이 좋아하니		0		
116	s_arimmie	게시물	86팔로워	222팔로우	219	PNU INDL_아부지랑랑랑 똑 닮아쪄...#부산 #동래 #동래읍성 #겨울 #겨울인데 #1	2021-02-13	.zzi_님 외 51명이 좋아하니	동래읍성북문	0		
117	s_arimmie	게시물	86팔로워	222팔로우	219	PNU INDL_집에 트리가 없어 집앞에 구경 나와보았습니다...#멜크 #크리스마스 #2020-12-25	2020-12-25	dong_2님 외 52명이 좋아하니	집앞길바닥	0		
118	s_arimmie	게시물	86팔로워	222팔로우	219	PNU INDL_우우우우우...인생은 다시 내려갈길이 있는 오르막#20201106 #부산	2020-11-07	dong_2님 외 59명이 좋아하니	Gathercoffeelab	0		
119	s_arimmie	게시물	86팔로워	222팔로우	219	PNU INDL_👉1002	2020-10-04	5	미술관	0		
120	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_오늘 배터져ㅋㅋㅋㅋ	2020-07-31	_aelle_님 외 9명이 좋아하니	Busan, South Korea	0		
121	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_전공문화체험(서울, 세종나들이)👉	2019-07-06	_aelle_님 외 14명이 좋아하니	Seoul, Korea	0		
122	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_경주여행👉	2019-05-05	_aelle_님 외 16명이 좋아하니	경주 황리단길	0		
123	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_경희랑 부대에서👉#부산대축제 #로이킴	2018-11-01	8	부산대학교 Pusan National U	0		
124	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_맑은하늘👉	2018-07-11	_smile__님 외 7명이 좋아하니	부산 화명동	0		
125	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_👉충강기념으로 놀러가기!! 태경이가 발탈출하다가 자기만 살겠다고 니	2018-06-22	_smile__님 외 9명이 좋아하니	Yangsang	0		
126	jynhm	게시물	7팔로워	38팔로우	166	PNU IE_오랜만에 엄청 놀았다👉카페에서 공부한다고 하고선 놀고ㅋㅋㅋ급창전	2018-06-06	_smile__님 외 8명이 좋아하니	Busan, South Korea	0		



4. 진행된 내용 및 결과

2. 데이터 전처리

인당 30개의 게시글(row)를 하나의 row로 바꾸는 과정

- 게시글 -> 글자 수 평균
- Like -> 좋아요 수 평균
- Tag_cnt -> 태그 수의 평균
- Timediff -> 시간 간격의 평균
- Mbti -> 0~15로 범주화

Column별 데이터

계정, bio내용 길이, 게시물 수, 팔로워 수, 팔로우 수, 게시글 길이, 좋아요 수, 태그 개수, 장소입력 수, 게시물 날짜 간격, MBTI, 공개/비공개 여부

	name	bio_len	cnt	followers	follows	content_len	like	tag_cnt	place	timediff	mbti	i/e	s/n	f/t	j/p	open
0	1ooo_jaeni	14	9	214	192	73.88888889	41.88888889	6.444444444	5	115.8888889	ISFJ	0	0	0	0	1
1	2021_happiest_man	1	3	183	262	54.33333333	49.66666667	1	3	137	ESTP	1	0	1	1	1
2	7h_34min	9	2	312	512	54.5	43	4.5	1	37.5	ENFJ	1	1	0	0	0
3	acoustic_string_cow	50	45	329	327	118.3793103	38.06896552	10.79310345	0	18.24137931	ENTP	1	1	1	1	1
4	annn_aelle_	1	138	1420	645	89.02941176	185.0294118	9.058823529	19	7.147058824	ENFP	1	1	0	1	1
5	azu_c_e_re	1	45	121	85	128.9333333	73.4	16.4	25	21.96666667	ESFJ	1	0	0	0	0
6	dal_sxxu	10	85	138	160	0	100	1	0	167	ENFP	1	1	0	1	1
7	deogun04	48	21	567	443	29.06666667	145.0333333	1	27	26.36666667	ESFP	1	0	0	1	1
8	dreamer_cha	82	1	238	227	53	54	8	1	37	ENTJ	1	1	1	0	1
9	eastshine96	1	0	158	194	0	0	1	0	1000000	ESFJ	1	0	0	0	0
10	f_ishfo_od	33	2	272	280	13.5	78	1	2	194	ESTJ	1	0	1	0	0
11	garden_0908	1	0	99	112	0	0	1	0	1000000	INFP	0	1	0	1	1
12	geon_tb	1	10	94	107	24.6	33.4	1	1	72.8	INFJ	0	1	0	0	1
13	ggggg_geo	1	33	207	110	107.2580645	46.09677419	11.22580645	9	24.90322581	INFP	0	1	0	1	1
14	gimume	1	3	154	151	0	0	1	0	1000000	ENFP	1	1	0	1	0
15	gwak.nayeon	10	21	320	325	36.28571429	42.19047619	1.666666667	1	45.33333333	ESFP	1	0	0	1	1

●

●

●

85	j_h0ung	3	0	60	115	0	0	0	0	1000000	ISFJ	0	0	0	0	0
86	kji_917	6	0	48	116	0	0	0	0	1000000	ISFJ	0	0	0	0	0
87	1_2x11	9	7	245	94	7	28	0	5	69	ESTP	1	0	1	1	1
88	ji_e_s_	17	15	328	35	7	102	0	4	6	ESTP	1	0	1	1	1
89	dear__ssu	11	13	513	278	9	123	0.2	2	8	ESTP	1	0	1	1	1
90	liyanwxx	19	31	836	288	6	237	0.1	20	4	ESTP	1	0	1	1	1
91	suk_is_	1	0	585	364	0	0	0	0	1000000	INTP	0	1	1	1	1
92	jinsol__p	16	140	171	174	10	37	0.8	0	11	INTP	0	1	1	1	1
93	sxxkvly	10	360	87	27	8	13	0.1	22	1	INTP	0	1	1	1	1
94	d20_ny	9	0	90	163	0	0	0	0	1000000	INTP	0	1	1	1	1
95	yeaeaan	0	0	33	40	0	0	0	0	1000000	ISFP	0	0	0	1	0
96	jeewonyee	0	7	80	167	222.71	30.42857143	4	5	37	INTP	0	1	1	1	1
97	je_...hui	0	19	96	123	41.11	31.57	2.8	18	95	ISFP	0	0	0	1	1



4. 진행된 내용 및 결과

3. 자연어 처리

- 대중적 태그(#f4f, #ootd, #선팔하면맛팔 등)를 자주 사용하는지의 여부 확인 (I/E, S/N)

→ 대중적 태그 자주 쓰는 사람 : E, N

- 복수 명사 및 접미사('들')와 부사(우리, 함께, 같이) 등의 명사 사용빈도(I/E)

→ 사용빈도 높은 사람 : E

- 감정 형용사(정말, 진짜, 완전, 진심)를 수식하는 부사 빈도 (T/F)

→ 사용빈도 높은 사람 : F

- 음식 게시물 태그의 빈도(I/E)

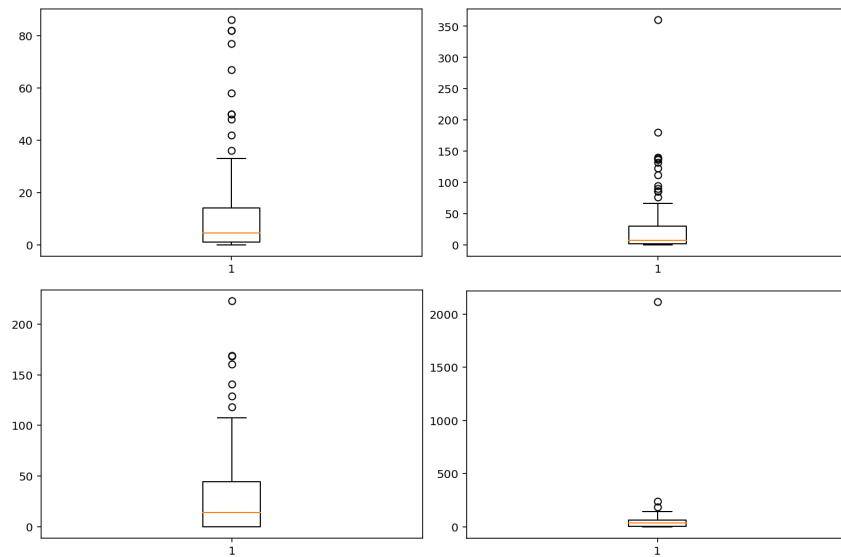
→ 사용빈도 높은 사람 : E

index	name	tag_cat	avg_special	lenN	lenVV	lenVA	lenMAG	OftenTags
0	0 z0_heon	1	1	57	13	4	3	[[('카즈탄탄화명정', 1), ('복구맛집', 1)]]
1	1 yun_ss_	1	3.4	72	45	24	27	[[('말레이시아', 2), ('pulispringsresort', 2), ('해동공사', 1), ('오션뷰부동산투기', 1), ('바다는연제나얌은뒤르림백만년만에', 1), ('happybirthday', 1), ('malaysia', 1), ('금산', 1), ('조호
2	2 yu_oni	0	0.666667	94	60	12	33	[]
3	3 yeon_smile	0	0	0	0	0	0	[]
4	4 xxco_verm	2	9.9	194	57	60	27	[[('중반', 13), ('좋아요반사', 11), ('오오티다', 11), ('고양국제고', 10), ('Hf', 10), ('교복스타그램', 9), ('첫돌반사', 9), ('likeforfollow', 9), ('likeforlikes', 9), ('교복', 8)]]
5	5 x_sit0	1	2.566667	238	125	39	51	[[('서영In제주', 5), ('서영In경주', 3), ('제주', 2), ('제주도', 2), ('경주월드', 1), ('', 1), ('도두동무지개해안도로', 1), ('도두봉', 1), ('무지개해안도로', 1), ('제주노형수퍼마켓', 1)]]
6	6 wjdvdgns	0	0	4	0	0	0	[]
7	7 wiseb0_0	2	6.862069	314	438	129	141	[[('일본', 9), ('여행스타그램', 8), ('여행', 8), ('japan', 8), ('소통', 7), ('교토', 7), ('kyoto', 7), ('선팔하면맛팔', 7), ('tenryuji', 6), ('은각사', 6)]]
8	8 tjddnr_	0	0	0	0	0	0	[]
9	9 ssung_ho	0	0	0	0	0	0	[]
10	10 serom_dal	1	3.233333	569	921	291	516	[[('폰카메라', 5), ('갤럭시s105g', 4), ('캠퍼스', 3), ('나홀로출사', 3), ('출사', 3), ('폰카', 2), ('갤럭시노트20울트라', 2), ('달', 2), ('보름달', 2), ('광안리해수욕장', 2)]]
11	11 seongyeog	0	0.5	0	0	0	0	[]
12	12 s_arimmie	1	3.1	188	78	9	51	[[('부산', 5), ('카페', 3), ('광안리', 2), ('동래읍성', 2), ('', 1), ('테제', 1), ('수플레팬케이크', 1), ('온천천', 1), ('온천천카페', 1), ('카페맛집', 1)]]
13	13 ppaeng_gv	1	4.5	174	144	39	96	[[('tr', 4), ('일본', 4), ('sonya6000', 3), ('turkey', 3), ('日本', 3), ('横浜', 3), ('happy_jaeuns_day', 2), ('720', 2), ('터키', 2), ('', 1), ('', 1), ('', 2)]]
14	14 perfect.kj	0	0.1	31	21	6	15	[]
15	15 park_9000	0	0	0	0	0	0	[]
16	16 opti_min_	0	0.285714	105	99	39	45	[]
17	17 my_0_o	0	0.423077	40	30	0	3	[]
18	18 mj_stagra	1	3.333333	26	12	0	6	[[('부산대학교산업공학과', 1), ('청년', 1), ('끝났다', 1), ('산공으로끝까지', 1), ('엽사그만', 1), ('산공이면', 1), ('필드하자', 1), ('금주', 1)]]
19	19 mina0_011	1	3.428571	34	51	15	36	[[('nan', 1), ('20190209', 1), ('행복해', 1)]]
20	20 kosori71	1	0.6	190	99	6	33	[[('saladeo10days', 1), ('세종시맛집샐러드', 1), ('부산대카페', 1), ('묘한다방', 1), ('고양이뿔뿔', 1), ('복덕와플세트', 1), ('치즈팔빙수', 1), ('묘한알바ㅋㅋㅋㅋ', 1), ('미나리', 1), ('표
21	21 klmiuhan	0	0.290323	74	90	12	33	[]

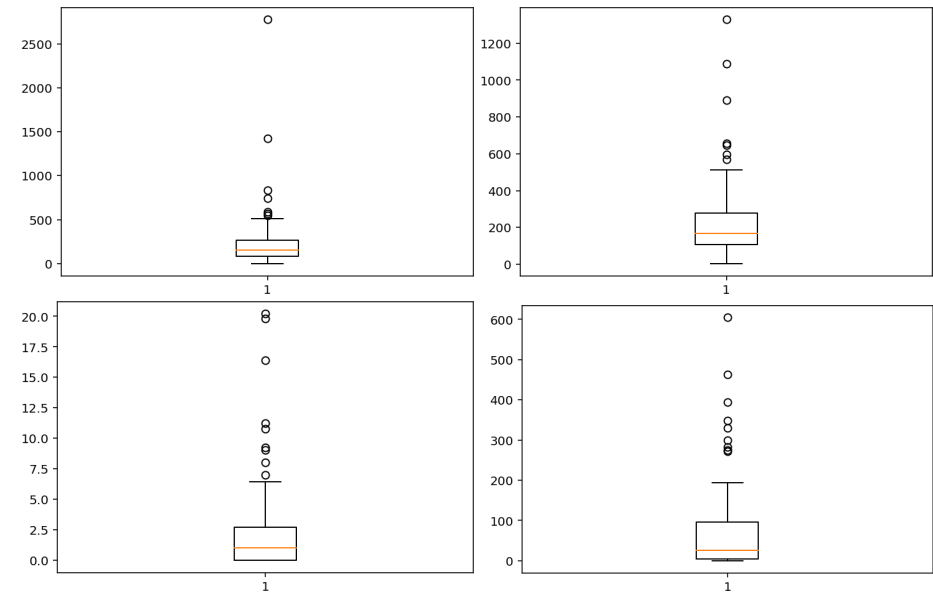


4. 진행된 내용 및 결과

4. EDA - boxplot



1. Bio의 길이는 보통 10자 이내임을 확인, 그러나 꽤 많은 수의 예외 또한 분포
2. 게시글 수는 평균 20~30개 내에 분포
3. 1 개의 게시글 당 30 자 내외임을 확인
4. 게시글 1개 당 평균 좋아요 수는 약 62개
5. 공개 계정/ 전체 계정 = 75.5%



1. 팔로워의 수는 평균 200 명 내외
2. 팔로우 수는 평균 200명 이하
3. 평균 태그의 수는 1~2개
4. 평균 게시글 간격은 ~~일



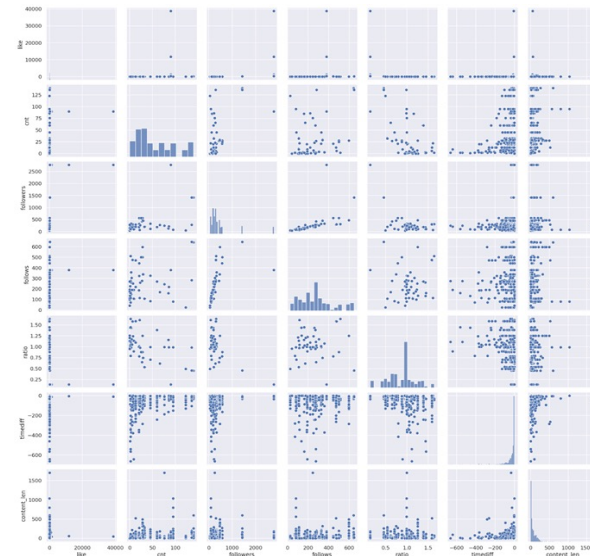
4. 진행된 내용 및 결과

4. EDA - correlation



<변수간의 상관관계>

대체적으로 변수끼리의 상관관계가 낮음
-> 전체 데이터를 활용하여 학습에 사용



<scatter plot>



4. 진행된 내용 및 결과

5. 모델 구축

```
from sklearn.metrics import confusion_matrix
from sklearn.ensemble import RandomForestClassifier

import sklearn.preprocessing
scaler = sklearn.preprocessing.StandardScaler()
X = scaler.fit_transform(df_mbti_1row_1211_cate_X1)

X_train, X_test, y_train, y_test = train_test_split(X, df_mbti_1row_1211_cate_y, test_size=0.15, random_state=41, shuffle=True)

res = []
for i in range(30):

    model = RandomForestClassifier(n_estimators = i+20,
                                   max_depth = 6,
                                   min_samples_leaf = 2,
                                   min_samples_split = 6,
                                   random_state = 42,
                                   n_jobs = -1)

    model = model.fit(X_train, y_train)
    pred_test = model.predict(X_test)
    m = confusion_matrix(y_test, pred_test)
    acc = np.sum(np.diag(m))/len(y_test)
    res.append(acc)
    print('n_estimators {} 결과의 정확도는 {:.4f}% 입니다'.format(i+21, acc*100))
```

-> Standard Scaler 사용하여 스케일링

-> Date set 분할(Train/test)

-> Random Forest Classifier 이용
Accuray 기반 최적 파라미터 찾기



4. 진행된 내용 및 결과

5. 모델 구축

n_estimators	29	결과의	정확도는	60.000000%	입니다
n_estimators	30	결과의	정확도는	60.000000%	입니다
n_estimators	31	결과의	정확도는	60.000000%	입니다
n_estimators	32	결과의	정확도는	60.000000%	입니다
n_estimators	33	결과의	정확도는	60.000000%	입니다
n_estimators	34	결과의	정확도는	60.000000%	입니다
n_estimators	35	결과의	정확도는	53.333333%	입니다
n_estimators	36	결과의	정확도는	60.000000%	입니다
n_estimators	37	결과의	정확도는	66.666667%	입니다
n_estimators	38	결과의	정확도는	60.000000%	입니다
n_estimators	39	결과의	정확도는	53.333333%	입니다
n_estimators	40	결과의	정확도는	53.333333%	입니다
n_estimators	41	결과의	정확도는	53.333333%	입니다
n_estimators	42	결과의	정확도는	53.333333%	입니다
n_estimators	43	결과의	정확도는	60.000000%	입니다
n_estimators	44	결과의	정확도는	66.666667%	입니다
n_estimators	45	결과의	정확도는	66.666667%	입니다
n_estimators	46	결과의	정확도는	66.666667%	입니다
n_estimators	47	결과의	정확도는	66.666667%	입니다
n_estimators	48	결과의	정확도는	66.666667%	입니다
n_estimators	49	결과의	정확도는	66.666667%	입니다
n_estimators	50	결과의	정확도는	66.666667%	입니다

- **Accuracy 66.7일 때의 Hyper Parameter**

- n_estimators = 37
- max_depth = 6
- min_samples_leaf = 2
- min_samples_split = 6
- random_state = 42
- n_jobs = -1



5. 기대효과

사회적 기대효과



- 개인 성향 분석에 대한 새로운 관점에서의 접근효과
- 사회에 편재한 성격 별 구매 성향 정보, 데이터를 활용해 개인에게 맞춤형 된 마케팅, 광고효과

학문적 기대효과



- 개인성향을 분석할 수 있는 자연어 처리 노하우 정립
- Random Forest 모델에 대한 심화학습



THANK YOU