



사지마세요 입양하세요

유기동물 입양 활성화를 위한
입양 가능성 예측



CONTENTS

BACKGROUND

DATA

MODEL

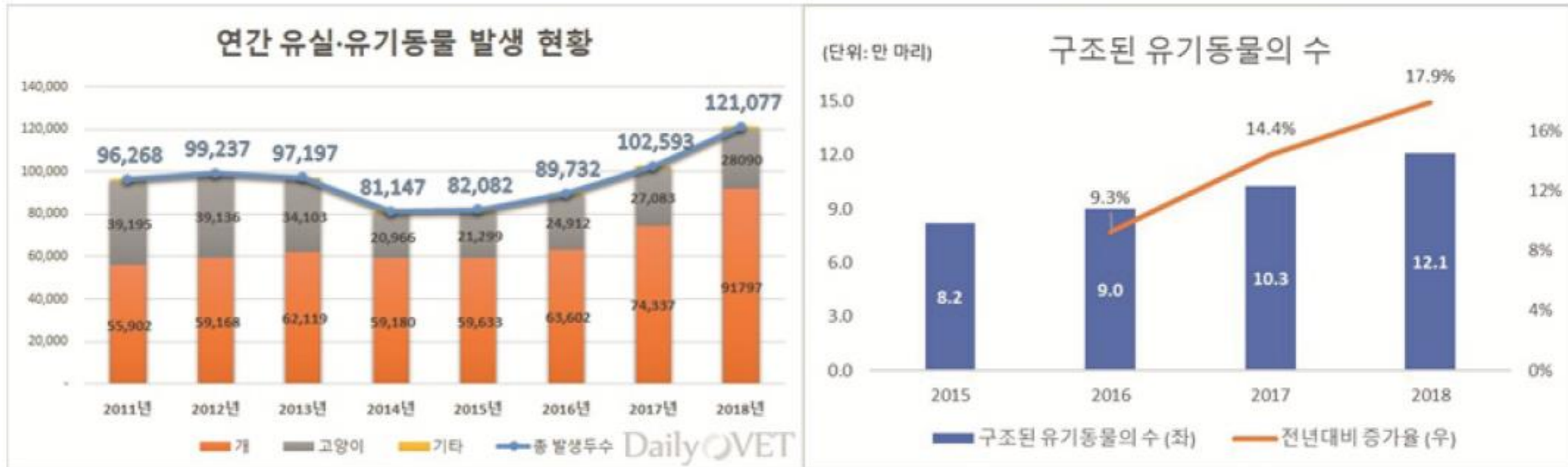
CONCLUSION



BACKGROUND

주제선정배경/분석목적

BACKGROUND 주제선정배경

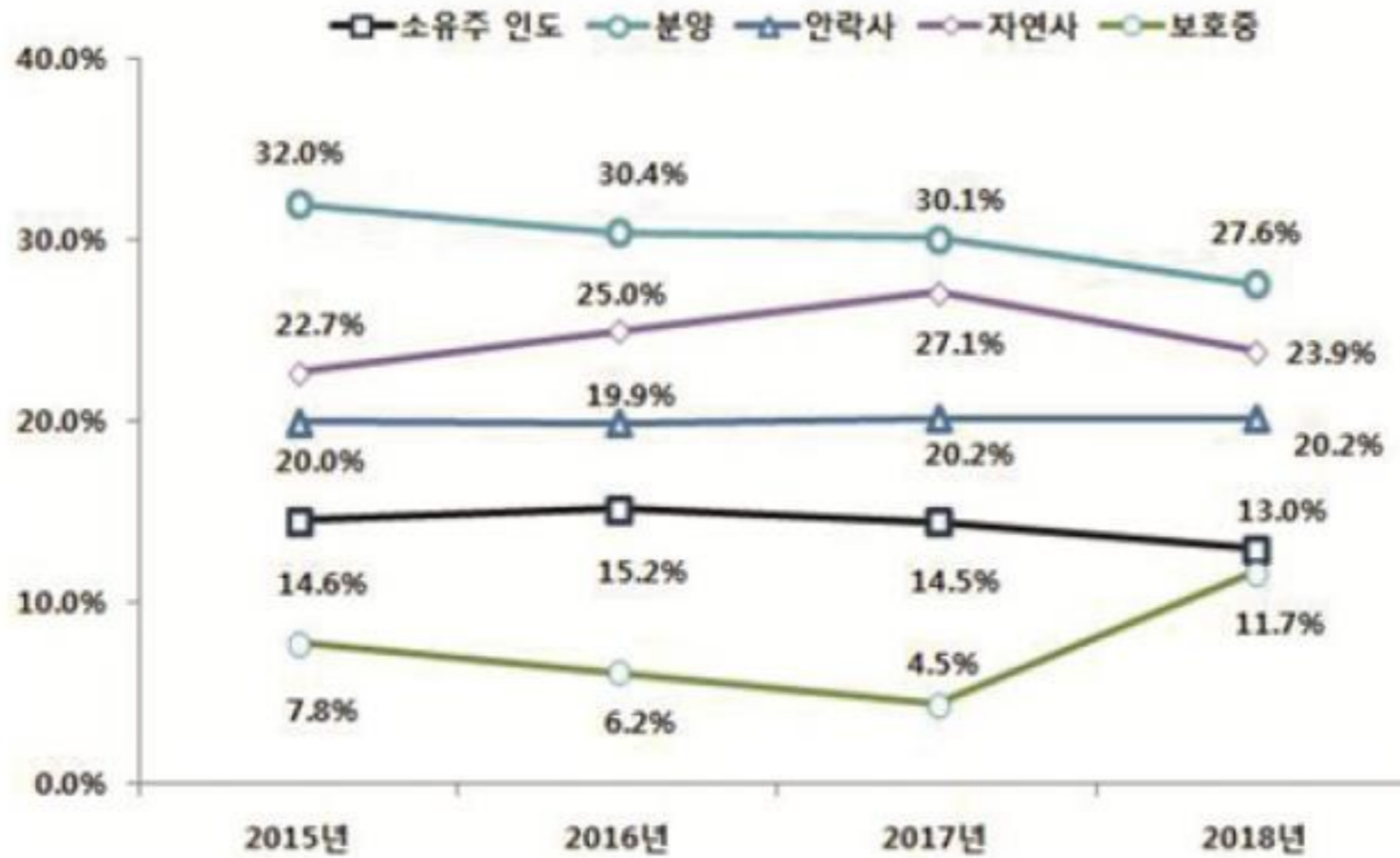


해마다 빠르게 늘어나는 유기동물의 수

2016년 89,732마리, 2017년 102,593마리, 2018년 121,077마리
 전년대비 증가율 2016년 9.3%, 2017년 14.4%, 2018년 17.9% 기록

출처: <https://mypetlife.co.kr/39524/>

BACKGROUND 주제선정배경



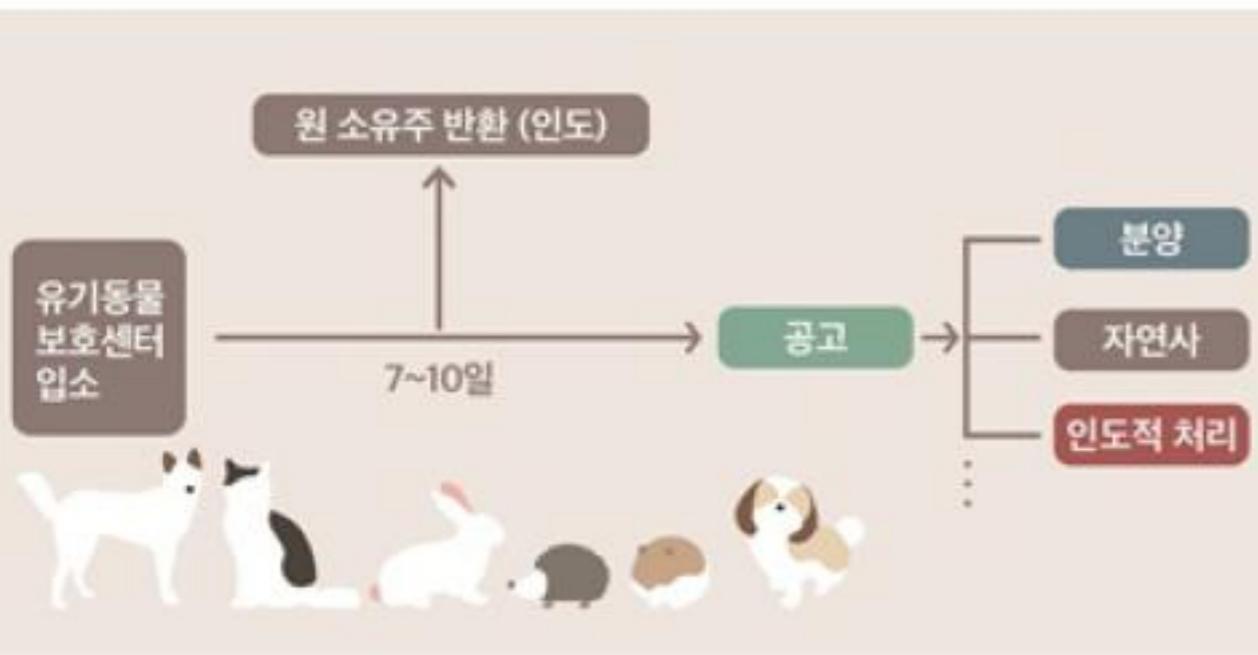
유기동물의 분양(입양)률은 매년 하락

구조된 유기동물 중 자연사하거나 안락사 되는 비율 44%, 보호자에게 다시 돌아가는 경우는 13%에 불과
특히, 유기동물 분양률은 32%를 기록한 2015년 이래 매년 하락해 작년에는 27.6%로 집계

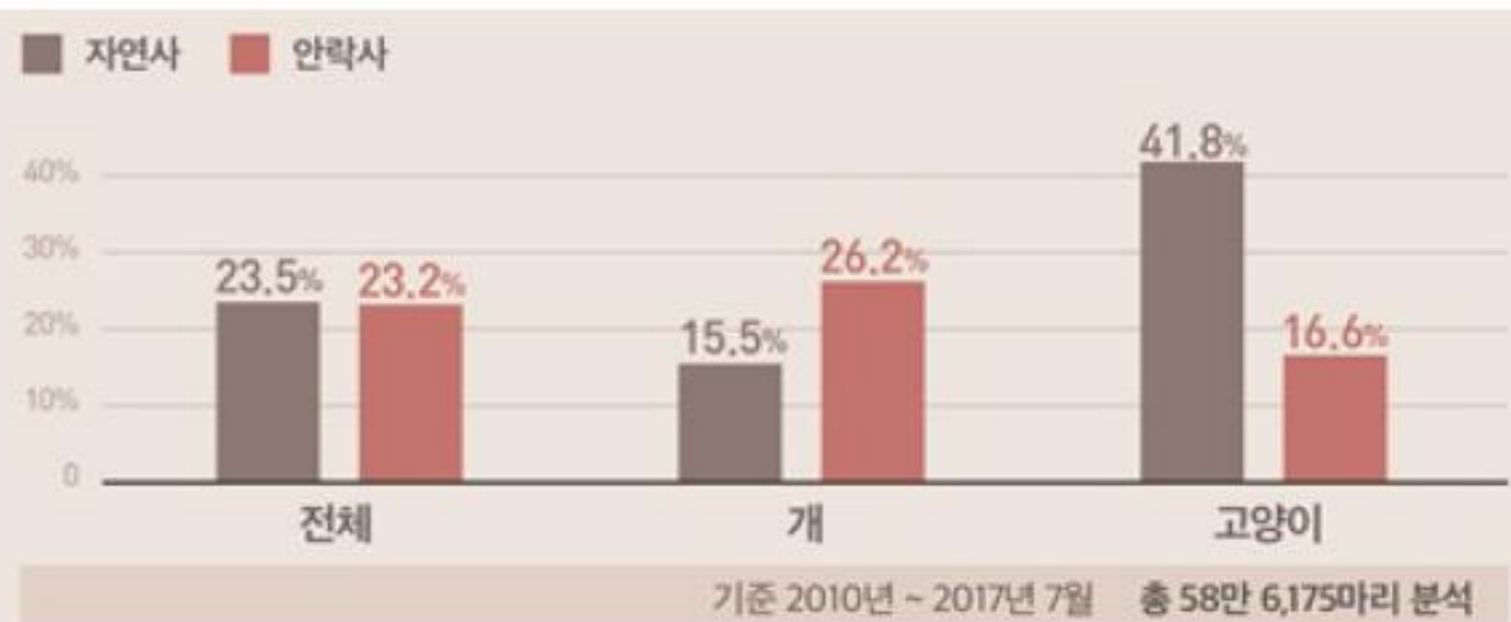
출처: <https://mypetlife.co.kr/39524/>

BACKGROUND 주제선정배경

유기동물 처리 체계도



안락사 비율 차이



운명의 10일

'인도는 12%' 나머지 유기동물은..

동물이 발견되면 최초 7일 동안 의무 공고
공고 일부터 10일이 지나면 해당 지자체가 소유권을 획득
최소 10일 동안 입양이 되지 않으면 안락사의 위험에 처하게 됨

더 살 수 있는데 죽는 동물들..

안락사 비율 '23.2%'

인도 또는 분양의 선택을 받지 못한 유기동물 대다수는
늙거나 병들어 방치되어 자연사하거나 인도적 처리 명목의
안락사에 의해 죽음이라는 안타까운 운명에 처해짐

출처: https://news.sbs.co.kr/news/endPage.do?news_id=N1004414522

BACKGROUND 분석목적

반려동물을 유기하지 않는 사회적 인식 개선과
건강한 반려동물문화형성을 위한 정책 수반이 근본적인 해결책이 될 수 있겠지만

유기된 동물들을 강제 안락사의 위험과

보호소의 열악한 환경에서 방치되어지는 것에서 보호하고자

이미 유기된 동물의 입양 활성화 방안에 포커스!



**유기동물 입양 가능성 예측 모델을 통해
유기견, 유기묘의 입양 활성화 제언**

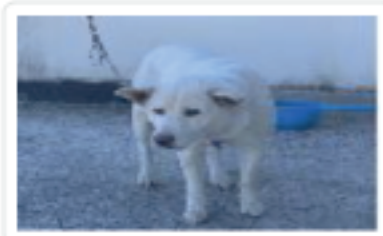




DATA

수집/전처리/정제/PCA

DATA 수집



자세히보기

공고번호	경남-하동-2020-00281
접수일자	2020-10-29
품종	믹스견
성별	암컷
발견장소	하동군 하동읍 산북..
특징	순함
상태	공고중
공고기간	2020-10-29 ~ 2020-11-08



자세히보기

공고번호	경남-거제-2020-01341
접수일자	2020-10-29
품종	한국 고양이
성별	암컷
발견장소	지세포 수협앞
특징	눈상태 안좋음
상태	공고중
공고기간	2020-10-29 ~ 2020-11-09

공고번호	경남-하동-2020-00281
품종	[개]믹스견
색상	흰색
성별	암컷
나이/체중	2017(년생) / 27 (Kg)
접수일시	2020-10-29
발생장소	하동군 하동읍 산북2길24
특징	순함
공고기한	2020-10-29 ~ 2020-11-08



** 유기동물 문의는 보호센터에 연락하시기 바랍니다.

보호센터이름	하동군유기동물보호소	전화번호	055-880-2439
보호장소	경상남도 하동군 적량면 한옥정길 90 (적량면) 유기동물보호소		
관할기관	경상남도 하동군		
담당자	하동군	연락처	055-880-2439
특이사항			

유기동물 공고 통합 시스템 데이터

각 시도군구의 동물보호업무 담당부서와 연계한 농림축산식품부 관할 동물보호관리시스템 운영 사이트

DATA 수집

instances
324,702

오픈 API
동물보호관리시스템
유기동물 조회 서비스

2017.1.1 ~ 2020.10.12
기간 데이터

features
14

유기, 유실 강아지/고양이
광고 데이터

유기번호/발견장소/품종/색상/
성별/중성화여부/특징/나이/
동물종류/몸무게/보호소이름/
관할기관/광고시작일/광고종료일

target
1

유기, 유실 강아지/고양이
관리 현황 데이터

총 6개의 클래스
보호중/종료(입양)/종료(반환)/
종료(기증)/종료(자연사)/종료(안락사)

DATA 모델링 ER 표기법



DATA 주요 데이터 탐색

변수명	수준	척도
나이	0세 ~ 30세	간격척도
성별	male / female	명목척도
중성화여부	yes / no	명목척도
타입	dog / cat	명목척도
색상	검정 / 흰색 / 갈색 / 회색 / 노랑	명목척도
몸무게	단위 kg	비율척도
품종	골든 리트리버, 닥스훈트, 믹스견, 한국 고양이 등	명목척도
공고일자	2017.01.01 ~ 2020.10.12	순서/간격척도
특징	ex. 무릎에상처가있고경계심이많음	명목척도
상태	보호중 / 종료(입양) / 종료(반환) / 종료(안락사) 등	명목척도

DATA 전처리

나이 / 몸무게

age	weight
2016(년 생)	8(Kg)
2019(년 생)	7(Kg)

— 이상치 처리 →



age	weight
5.0	8
2.0	7

성별 / 중성화여부 / 타입

sexcd	neuteryn
M	N
M	N

— 라벨 인코딩 →



sexcd	neuteryn
1	2
1	0

품종 / 색상 / 특징

kindcd	colorcd	specialmark
[고양이] 한국 고양이	치즈태비	1469
[고양이] 한국 고양이	검정색/백색	매우 사나움

— 텍스트 마이닝 →



bic_col	brown_col	white_col	grey_col	yellow_col
1	0	1	0	0
0	0	0	0	1

나이 / 성별 / 중성화여부 / 타입 / 몸무게



데이터 정제

타입 변환 (나이, 타입, 몸무게)

이상치/결측치 제거 (나이, 성별, 중성화여부, 타입, 몸무게)



수치화

1(True), 0(False)로 변환 (중성화여부)

1(dog), 0(cat)로 변환 (타입)

품종



표현의 통일

ex) 믹스, 혼종, 잡종, 믹스견 등 → 믹스견으로 통일



수치화

오픈 API의 품종 테이블 기준으로 각 품종별 숫자코드 부여



one-hot encoding

수치화된 품종코드 이용해 더미변수화

색상



표현의 통일

ex) 깜장이, 검정색, 깜냥이, 흑색, black 등
→ black으로 통일



FREQUENCY TABLE

상위 5개 색상 컬럼 생성 후 수치화
ex) 검은 배경에 흰 줄무늬

→ blc_col과 white_col에 값 1 부여 (더미변수화)

특징



텍스트 마이닝

정규화 / 토큰화 / 워드 임베딩



FREQUENCY TABLE

4가지 메인 토픽 (사회성좋음, 예민함, 질병유무, 사고이력)으로 분류한 특성 컬럼 생성 후 → 0, 1로 더미변수화
ex) 경계심있는편이고 사고로인해다리에 부상있음
→ sensitive와 accident에 값 1 부여 (더미변수화)

DATA 텍스트 마이닝

정규표현식

1차 데이터 정제

워드 토큰화

mecab

단어집합 해시맵

vocab 생성

워드 임베딩

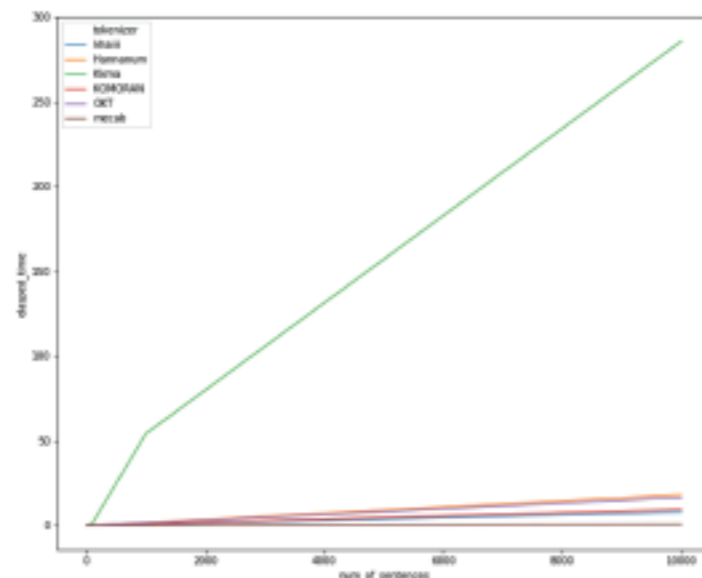
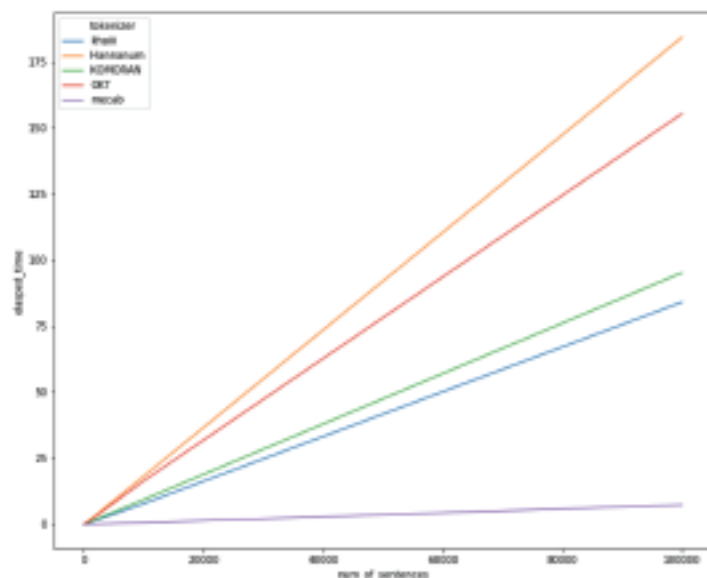
Word2Vec

토픽 워드 도출

K-Means Clustering

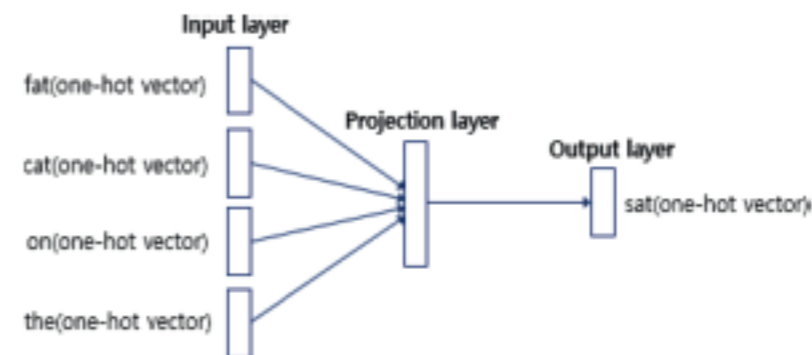
한글 형태소 분석기 mecab
이용하여 토큰화 및 불용어 제거

Skip-gram 으로
단어간 유사도 반영 위해 벡터화

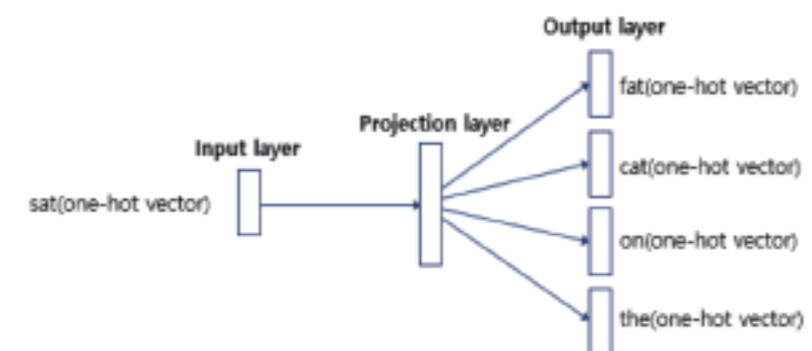


한글 형태소 분석기 성능 비교

Kharii / Hannanum / Kkma / KOMORAN / OKT / mecab
<연산 속도 및 상위권의 분석 품질을 보유한 mecab으로 결정>



<Word2Vec의 CBOW 구현 방식 인공신경망 도식화>



<Word2Vec의 Skip-gram 구현 방식 인공신경망 도식화>

good_behavior

유순, 활발, 온호기심, 좋, 순화, 밝, 순하, 착하, 착한, 조용, 소, 다르, 발랄, 영리, 애교, 이쁜, 귀여움, 귀여워, 예쁘, 착함, 깨끗, 좋아, 활달, 예쁨

sensitive

예민, 사나운, 까, 경계심, 위협, 공격, 사나움, 하, 겁, 사납, 으르렁댐, 경계, 입질, 움찔

disease

전염병, 눈병, 노출, 쇠약, 병, 증상, 모, 낭, 충, 비강, 인, 버, 시, 기, 췌, 저하, 오, 미, 치, 핑, 질, 환, 질, 병, 악, 호흡기, 탈수, 저체온증, 영양실조, 쇠약

accident

신경, 파열, 기사, 불, 능, 골, 저, 회, 아, 팽만, 뒷다리, 코, 마, 기, 립, 세, 후, 구, 탈, 전, 두, 부, 상, 쓰러져, 안면, 이상, 각, 이, 물린, 장기, 교, 통, 사, 고, 교, 통, 사, 고, 손상, 을, 무, 곤란, 보행, 아픈, 마비, 혼수

원본 데이터

	desertionno	popfile	happendt	happenplace	kindcd	colorcd	age	weight	noticeno	noticesdt	noticeedt	sexcd	neuteryn	specialmark	carenm	orgnm	lat	long	processstate	type
0	4.4854E+14	http://www	20201013	김해시 주촌면	[개]	믹스견 검흰	2016(년생)	8(Kg)	경남-김해	20201015	20201026	F	U	1469	유기동물	경상남도	35.22855	128.8894	보호중	dog
1	4.465E+14	http://www	20201013	군내면 송산길	[개]	믹스견 황색	2019(년생)	7(Kg)	전남-진도	20201015	20201022	F	N	매우 사나움	진도개메	전라남도	34.48687	126.2635	보호중	dog
2	4.465E+14	http://www	20201013	군내면 송산길	[개]	믹스견 백색	2020(년생)	10(Kg)	전남-진도	20201015	20201022	M	N	매우 사나움	진도개메	전라남도	34.48687	126.2635	보호중	dog
3	4.465E+14	http://www	20201013	고군면 가계길	[개]	믹스견 황색	2019(년생)	15(Kg)	전남-진도	20201015	20201022	F	N	사나움	진도개메	전라남도	34.48687	126.2635	보호중	dog
4	4.465E+14	http://www	20201013	금일 칠송리 미	[개]	믹스견 흰색	2017(년생)	14(Kg)	전남-완도	20201015	20201024	F	N	귀족 갈색,온	유기동물	전라남도	34.31106	126.7551	보호중	dog
5	4.1131E+14	http://www	20201013	덕릉로63가길	[개]	보스톤검흰	2017(년생)	13(Kg)	서울-도봉	20201015	20201022	F	U	유선처짐,깁	한국동물	서울특별시	37.66877	127.0471	보호중	dog
6	4.113E+14	http://www	20201013	호창공원앞 역	[개]	슈나우회색	2013(년생)	5.90(Kg)	서울-용산	20201015	20201026	M	Y	없음	열린동물	서울특별시	37.53843	126.9654	보호중	dog
7	4.5065E+14	http://www	20201013	제주시 중앙로	[개]	믹스견 흰갈	2020(년생)	11.2(Kg)	제주-제주	20201014	20201026	M	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	종료(반환)	dog
8	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2020(년생)	1.5(Kg)	제주-제주	20201014	20201026	F	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog
9	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2020(년생)	1.3(Kg)	제주-제주	20201014	20201026	F	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog
10	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2020(년생)	1.3(Kg)	제주-제주	20201014	20201026	M	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog
11	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2020(년생)	1.5(Kg)	제주-제주	20201014	20201026	M	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog
12	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2020(년생)	1.4(Kg)	제주-제주	20201014	20201026	M	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog
13	4.5065E+14	http://www	20201013	서귀포시 보목	[개]	믹스견 흰색	2018(년생)	17.7(Kg)	제주-제주	20201014	20201026	F	N	(개체관리번호)	제주 동물!	제주특별자치도	33.48901	126.4983	보호중	dog

전처리 완료 데이터

sexcd	neuteryn	type	blc_col	brown_col	white_col	grey_col	yellow_col	noticesdt_year	noticesdt_month	noticesdt_day	good_behavior	sensitive	disease	accident	processstate_adopt	processstate_euthanize	age	weight	kind0	kind1
1	2	1	1	0	1	0	0	2020	10	15	0	0	0	0	0	0	5	8	0	0
1	0	1	0	0	0	0	1	2020	10	15	0	1	0	0	0	0	2	7	0	0
0	0	1	0	0	1	0	0	2020	10	15	0	1	0	0	0	0	1	10	0	0
1	0	1	0	0	0	0	1	2020	10	15	0	1	0	0	0	0	2	15	0	0
1	0	1	0	0	1	0	0	2020	10	15	1	0	0	0	0	0	4	14	0	0
1	2	1	1	0	1	0	0	2020	10	15	1	0	0	1	0	0	4	13	0	0
0	1	1	0	0	0	1	0	2020	10	15	0	0	0	0	0	0	8	5.9	0	0
0	0	1	0	1	1	0	0	2020	10	14	0	0	0	0	0	0	1	11.2	0	0
1	0	1	0	0	1	0	0	2020	10	14	0	0	0	0	0	0	1	1.5	0	0
1	0	1	0	0	1	0	0	2020	10	14	0	0	0	0	0	0	1	1.3	0	0
0	0	1	0	0	1	0	0	2020	10	14	0	0	0	0	0	0	1	1.3	0	0
0	0	1	0	0	1	0	0	2020	10	14	0	0	0	0	0	0	1	1.5	0	0
0	0	1	0	0	1	0	0	2020	10	14	0	0	0	0	0	0	1	1.4	0	0

문자열 처리 / One hot encoding

카테고리형 피처를 학습시키기 위해 One hot encoder 사용

brown_col	white_col	grey_col	yellow_col	...	kind178	kind179	kind180	kind181	kind182	kind183	kind184	kind185	kind186	kind187
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	-1.155180	-0.218758	2.794427	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	-1.155180	-0.218758	2.794427	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	-1.155180	4.571260	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
1.270528	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
-0.787074	0.865666	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771
1.270528	-1.155180	-0.218758	-0.357855	...	-0.00555	-0.025379	-0.001755	-0.002482	-0.001755	-0.002482	-0.001755	-0.001755	-0.32626	-0.009771

표준화 / Standard scaling

Regressor 모델에 사용하기 위해 scaling된 데이터프레임 생성

sexcd	neuteryn	type	lat	long	blc_col	brown_col	white_col	grey_col	yellow_col	kindcd	noticesdt	noticesdt	noticesdt	good_beh	sensitive	disease	accident	processsta	processsta	age	weight	kind0
0.326745	1.450005	0.726059	-0.90491	0.15943	1.500274	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.07166	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	1.047127	-0.00168	-0.05655
0.326745	-0.72349	0.726059	-1.54425	-0.04559	-0.66654	-0.78707	-1.15518	-0.21876	2.794427	-0.47417	1.15195	1.036591	-0.07166	-0.55569	3.132218	-0.34124	-0.40459	-0.64364	-0.50591	-0.23629	-0.00171	-0.05655
-0.82463	-0.72349	0.726059	-1.54425	-0.04559	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.07166	-0.55569	3.132218	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00162	-0.05655
0.326745	-0.72349	0.726059	-1.54425	-0.04559	-0.66654	-0.78707	-1.15518	-0.21876	2.794427	-0.47417	1.15195	1.036591	-0.07166	-0.55569	3.132218	-0.34124	-0.40459	-0.64364	-0.50591	-0.23629	-0.00148	-0.05655
0.326745	-0.72349	0.726059	-1.69581	-0.00721	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.07166	1.799549	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	0.619323	-0.00151	-0.05655
0.326745	1.450005	0.726059	1.198639	0.015593	1.500274	-0.78707	0.865666	-0.21876	-0.35786	-2.2758	1.15195	1.036591	-0.07166	1.799549	-0.31926	-0.34124	2.471664	-0.64364	-0.50591	0.619323	-0.00154	-0.05655
-0.82463	0.363259	0.726059	1.086277	0.00922	-0.66654	-0.78707	-1.15518	4.57126	-0.35786	-0.29793	1.15195	1.036591	-0.07166	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	2.33054	-0.00174	-0.05655
-0.82463	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	1.270528	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00159	-0.05655
0.326745	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00186	-0.05655
0.326745	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00187	-0.05655
-0.82463	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00187	-0.05655
-0.82463	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00186	-0.05655
-0.82463	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00187	-0.05655
0.326745	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	3.132218	-0.34124	2.471664	-0.64364	-0.50591	0.191518	-0.00141	-0.05655
-0.82463	-0.72349	0.726059	-2.40444	-0.02725	1.500274	1.270528	-1.15518	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00181	-0.05655
0.326745	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00188	-0.05655
0.326745	-0.72349	0.726059	-2.40444	-0.02725	-0.66654	-0.78707	0.865666	-0.21876	-0.35786	-0.47417	1.15195	1.036591	-0.18553	-0.55569	-0.31926	-0.34124	-0.40459	-0.64364	-0.50591	-0.66409	-0.00188	-0.05655



MODEL

평가지표/모델선정/학습/평가

MODEL 평가지표

예측

		Negative 0	Positive 1
실제	Negative 0	TN True Negative 실제 입양되지 않고 입양이 되지 않을 것으로 예측	FP False Positive 실제 입양이 되지 않지만 입양이 될 것으로 예측
	Positive 1	FN False Negative 실제 입양이 되지만 입양이 되지 않을 것으로 예측	TP True Positive 실제 입양이 되고 입양이 될 것으로 예측

정확도 = $TN + TP / (TN + FP + FN + TP)$

정밀도 = $TP / (FP + TP)$

실제 입양이 되지 않지만 입양이 된다고 잘못 판단하면 큰 영향 발생하는 경우 중요

재현율 = $TP / (FN + TP)$

실제 입양이 되지만 입양이 되지 않는다고 잘못 판단하면 큰 영향 발생하는 경우 중요

F1 Score = $2 * (\text{재현율} * \text{정밀도}) / (\text{재현율} + \text{정밀도})$

MODEL 모델선정

분류 모델

Light GBM

Decision Tree Classifier

Logistic Regression

Random Forest Classifier

Gradient boosting

XGboost

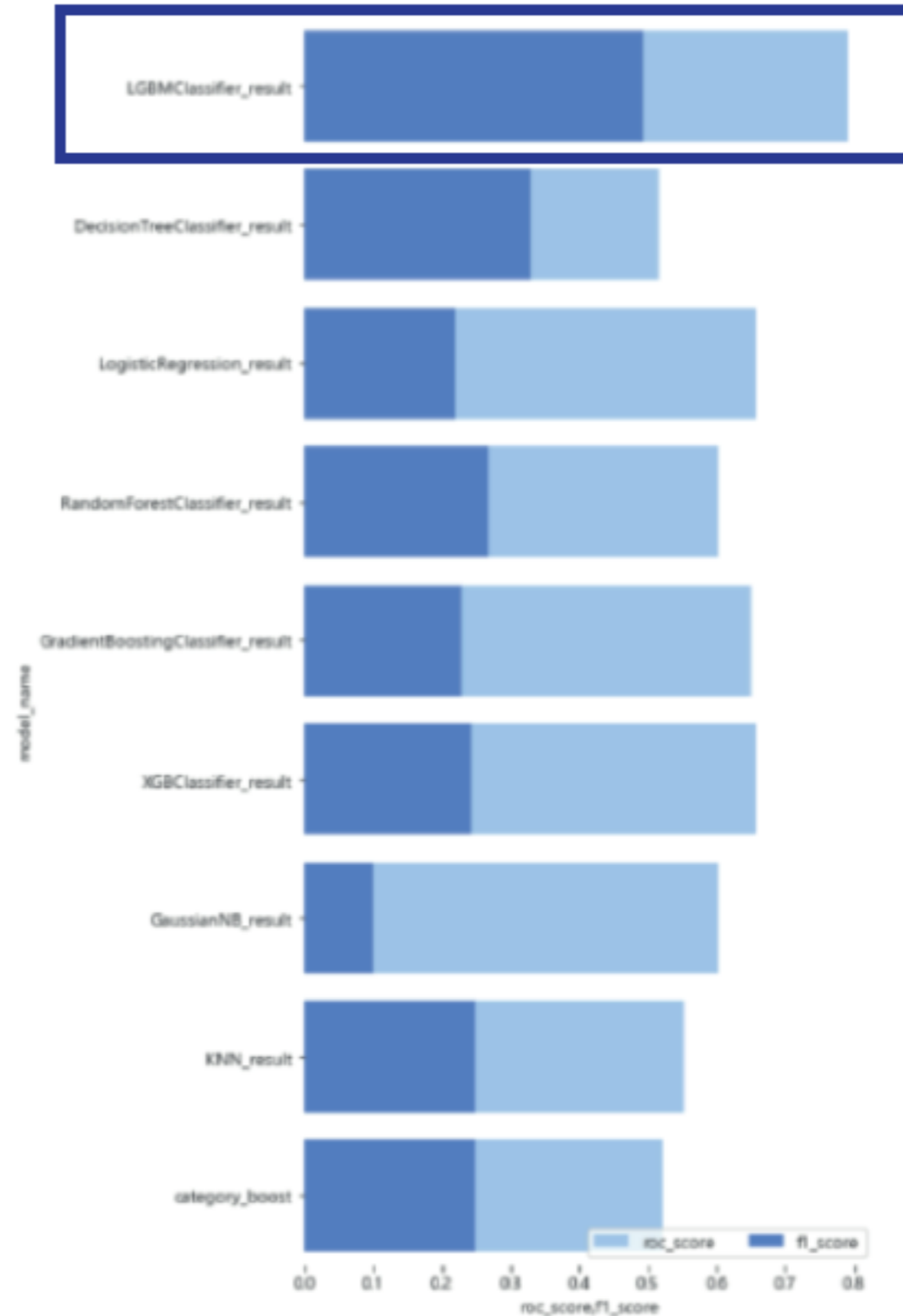
GaussianNB

K neighbors Classifier

Catboost

교차 검증

`cross_validate()`, `cv=5`



< 9개 모델별 f1_score, roc_score 비교 시각화 >

MODEL 모델 학습/하이퍼파라미터 조정



입양된 반려동물의 비율 41.43%

보다 정확한 분류 모델 학습을 위해
계층별 K-Fold 교차검증 사용

AS-IS → GridSearchCV 하이퍼파라미터 조정 → **TO-BE**

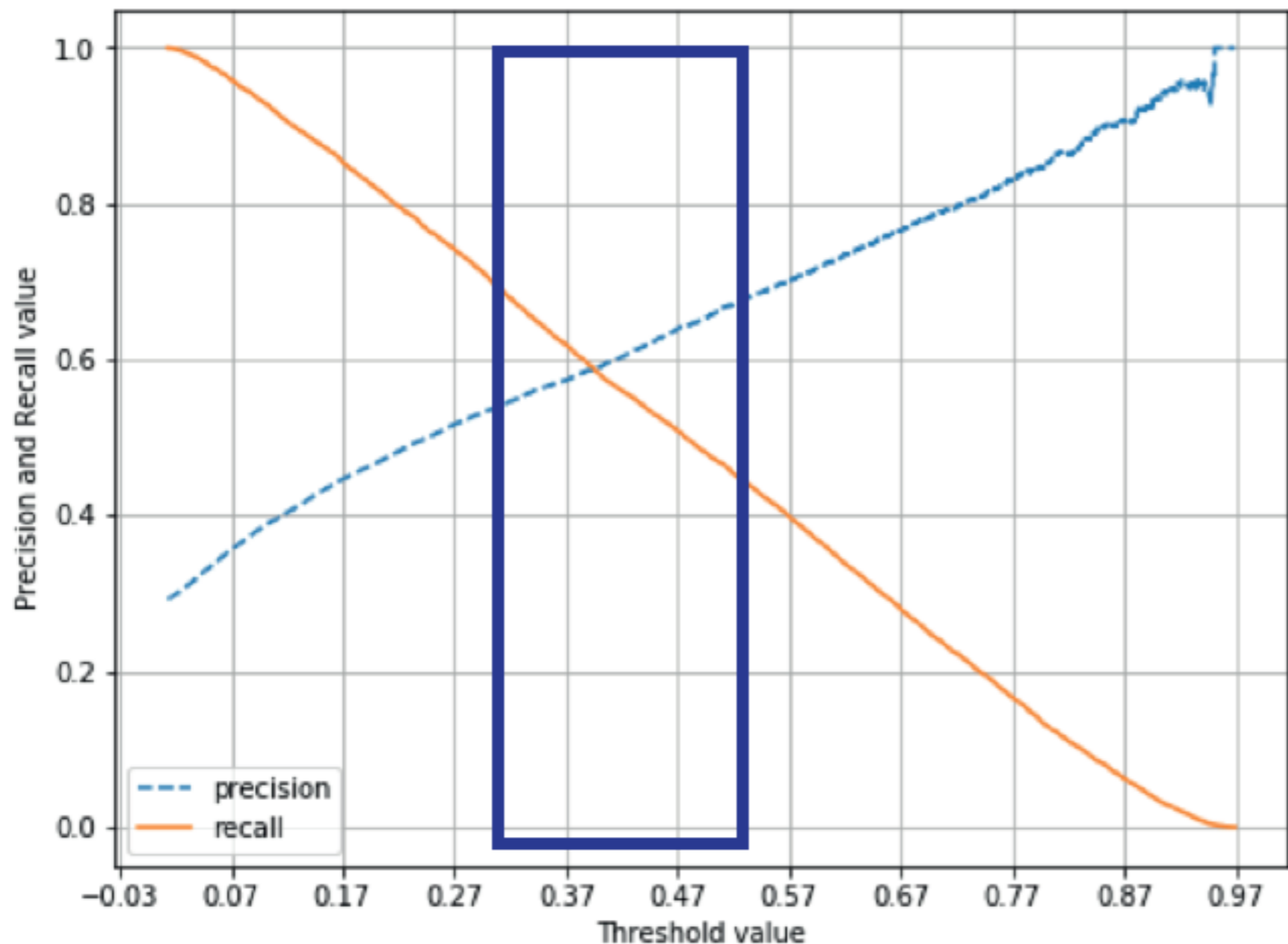
```
1 # 예측 성능 평가
2 get_clf_eval(y_test, preds, pred_proba)
```

```
[[21389  1571]
 [ 6430  3081]]
정확도: 0.7535955159988913
F1: 0.43507731412836265
정밀도: 0.6622957867583835
재현률(민감도): 0.32394070024182525
ROC: 0.6277586776470451
```

```
1 # 예측 성능 평가
2 get_clf_eval(y_test, preds, pred_proba)
```

```
[[20612  2348]
 [ 4912  4599]]
정확도: 0.7764158787841459
F1: 0.5588771418155305
정밀도: 0.6620123794443645
재현률(민감도): 0.48354536852066027
ROC: 0.6906402800791455
```

MODEL 임계값 조절



교차점을 기준으로
f1 스코어가 가장 높은
임계값 찾기

< 최종 모델 평가 점수 >

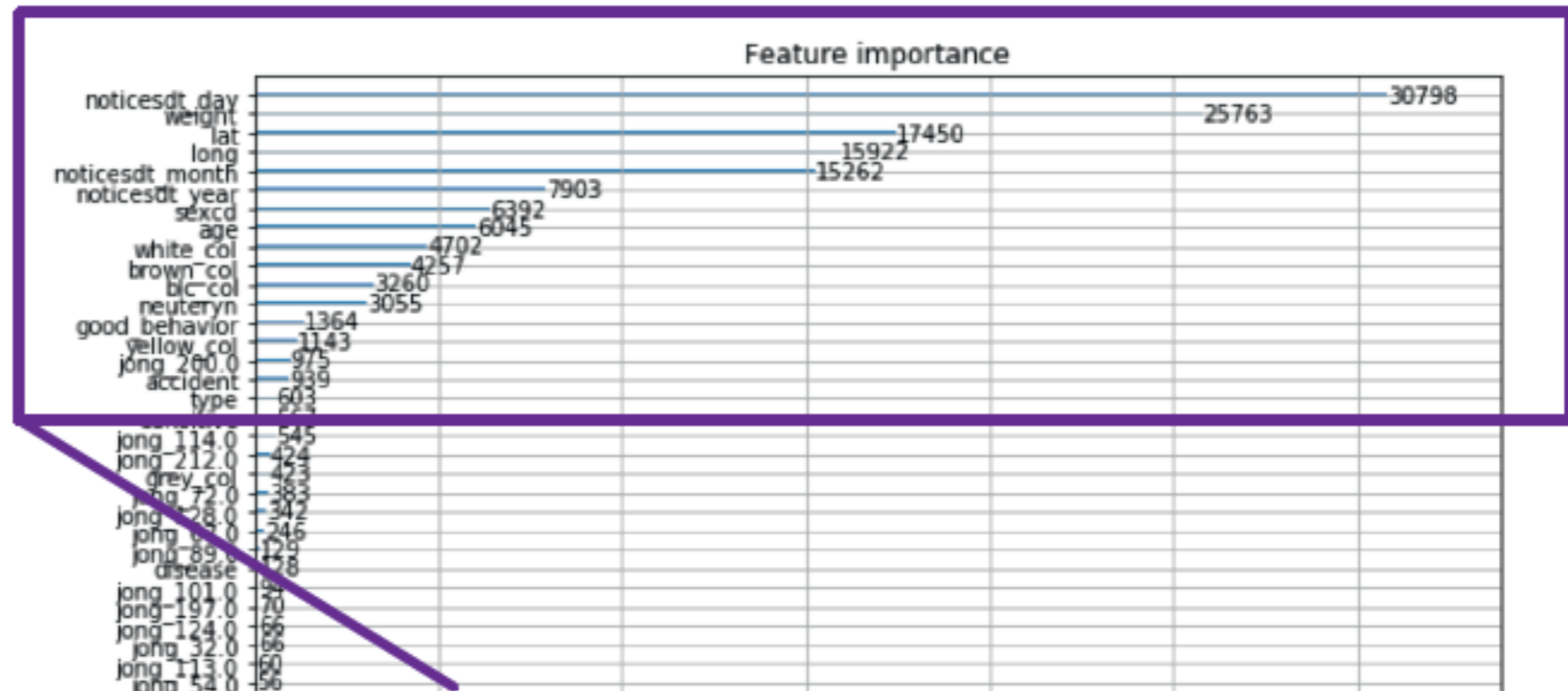
```
-----  
임계값: 0.29  
[[17263  5697]  
 [ 2637  6874]]  
정확도: 0.7433402112654368  
F1: 0.6225885336473145  
정밀도: 0.5468140959350887  
재현률(민감도): 0.7227420881085059  
ROC: 0.7373074552040788  
-----
```



CONCLUSION

결과및제언/한계및보완점/활용방안

CONCLUSION 중요 FEATURES



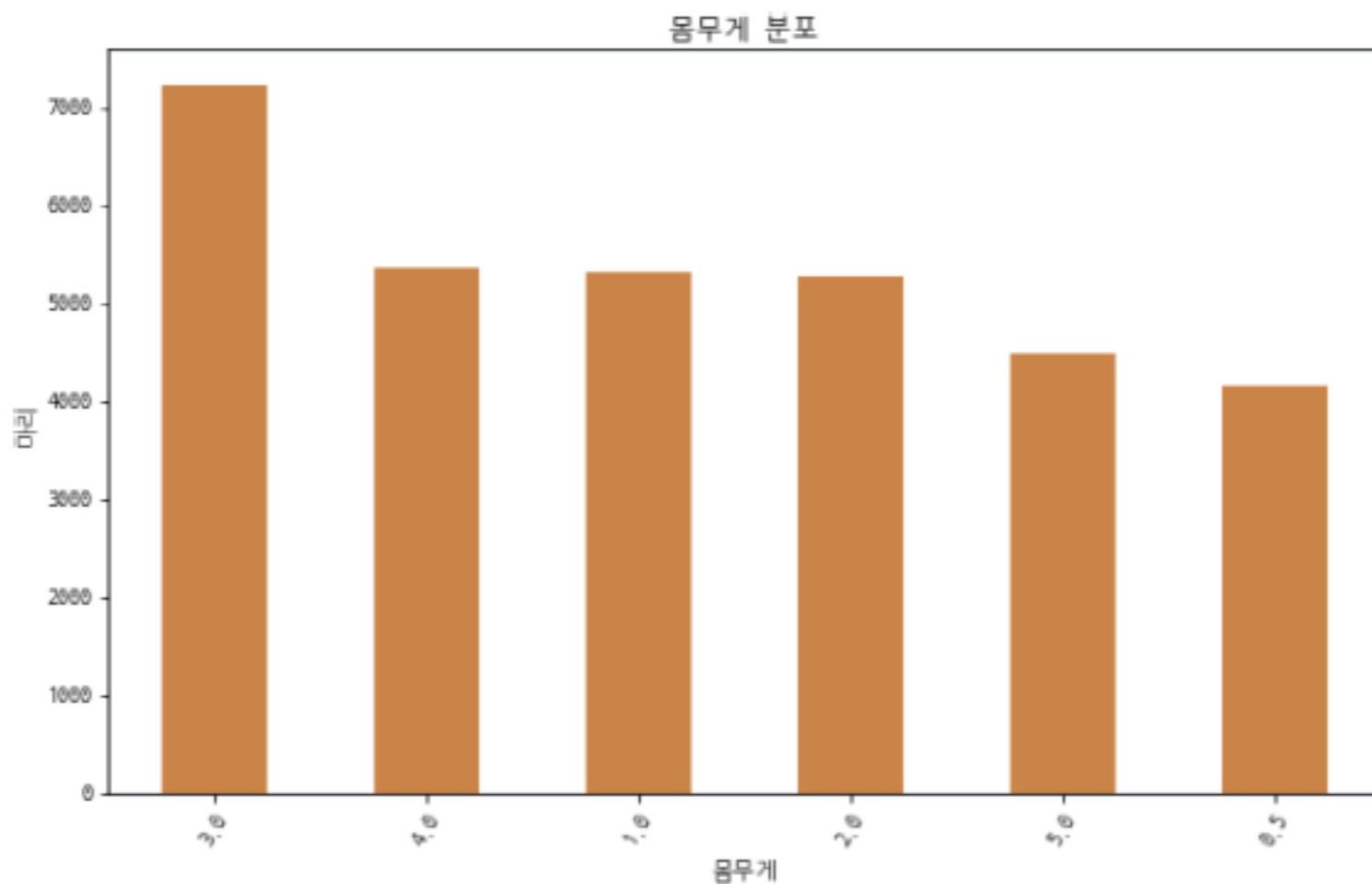
TOP 6



CONCLUSION

중요 FEATURES

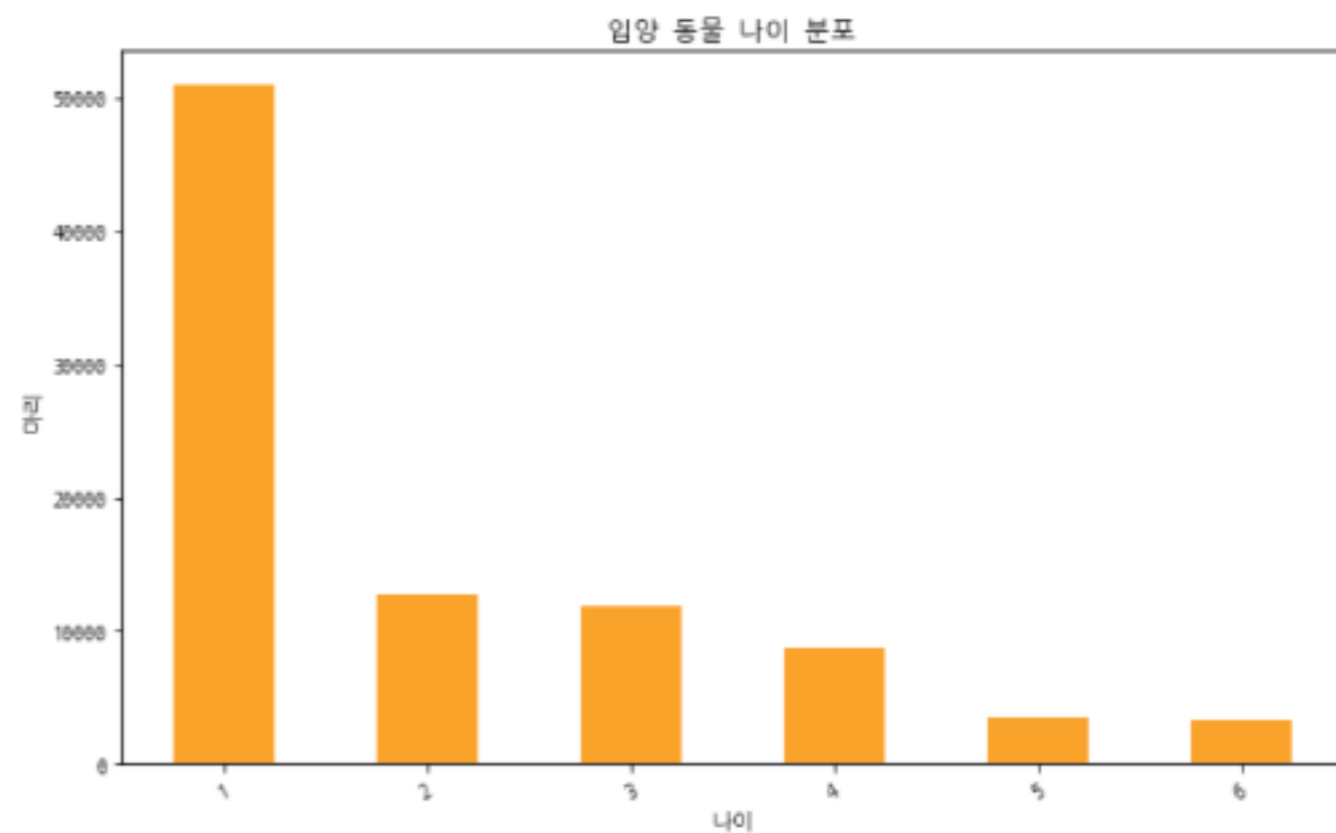
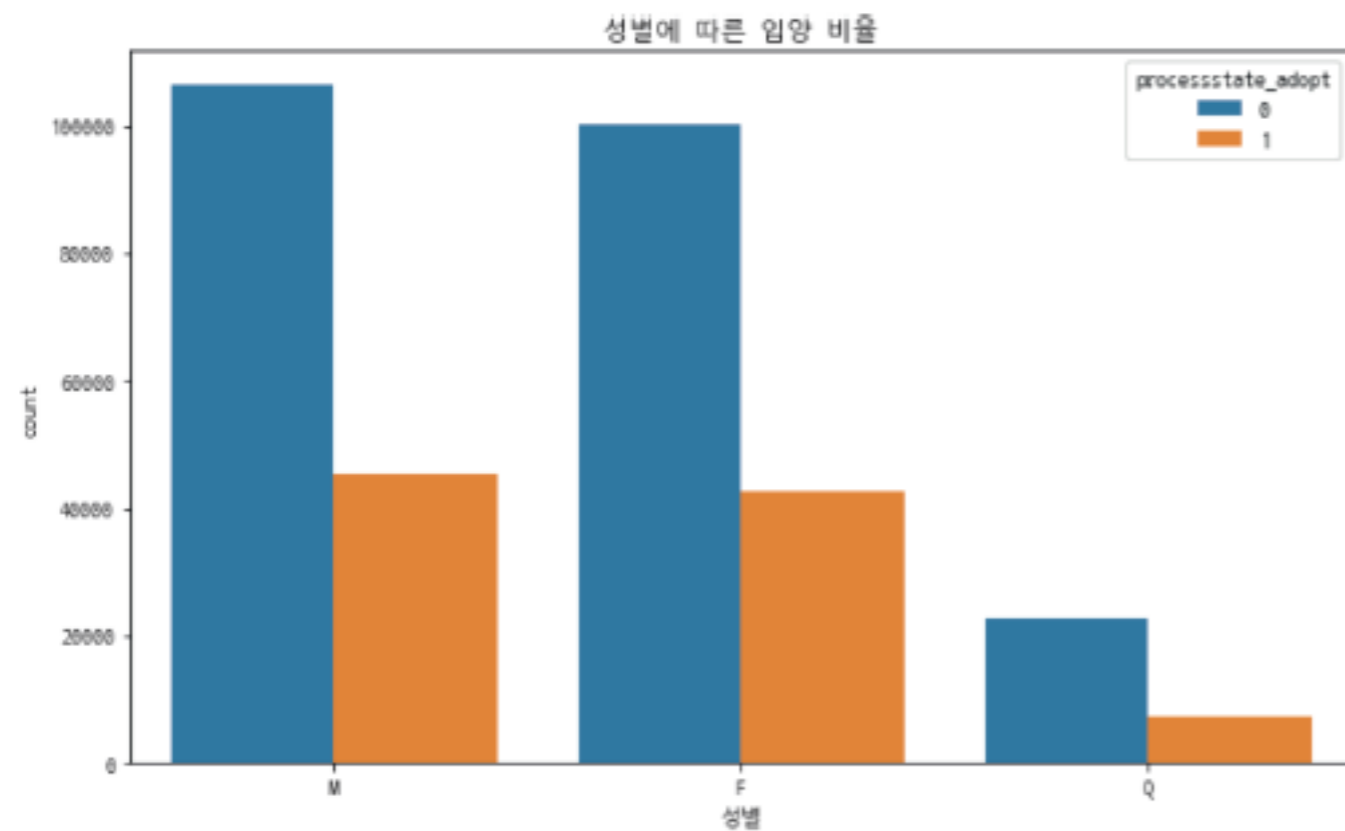
입양된 유기견/유기묘 몸무게



유기동물의 몸무게가 입양에 두번째로 큰 **영향력**을 가진 **특성**임을 확인
8kg 미만 소형견으로 분류
체중이 적게 나가는 유기동물의 입양률이 상대적으로 높음

CONCLUSION 중요 FEATURES

입양된 유기견/유기묘 성별 / 나이



유기견과 유기묘의 **성별**이
입양의 선호도와 상관성이 적음

1살인 유기동물의 **입양**이 절대적으로 많음
어린 연령대의 유기견과 유기묘의 **입양**을 선호

CONCLUSION

중요 FEATURES

공고일자 월별 유기 및 입양

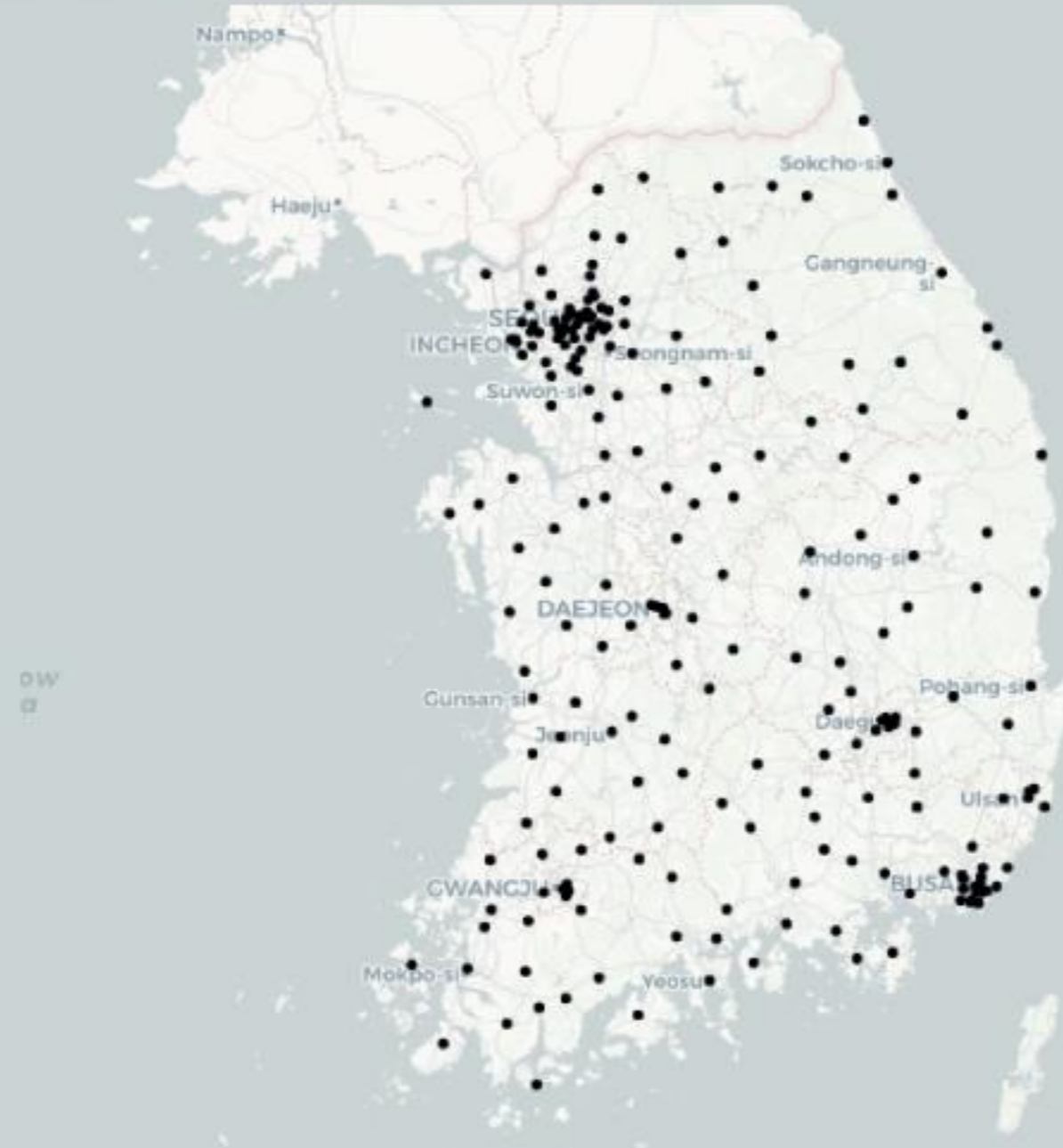


공고일자가 입양에 가장 큰 영향
7~10월 에 유기되는 동물의 수가 현저히 많음

공익 광고 및 캠페인 진행에 참고 가능

CONCLUSION 중요 FEATURES

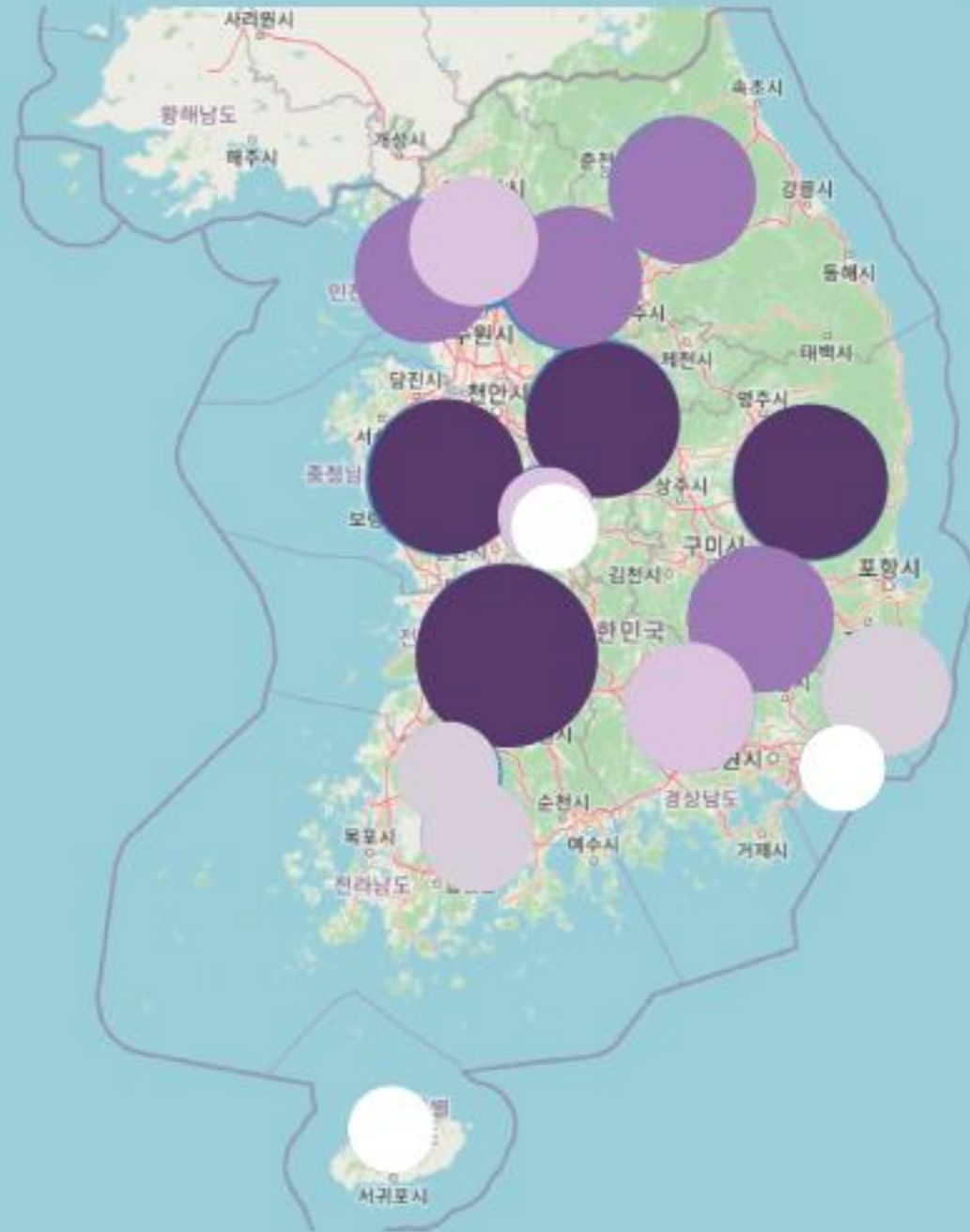
유기된 장소 현황 파악 보호소 위치/입양 장소와 동일



유기동물보호소는 서울, 부산, 대구, 대전, 광주 등 광역시 위주로 위치하고 있음을 확인

CONCLUSION 중요 FEATURES

보호소 지역별 입양률



입양률 매우 높음

전북 32%
경북 28%
충남 28%
충북 27%



입양률 높음

강원 26%
대구 25%
경기 25%
인천 25%



입양률 보통

세종 24%
서울 24%
경남 23%



입양률 낮음

울산 22%
전남 19%
광주 18%



입양률 매우 낮음

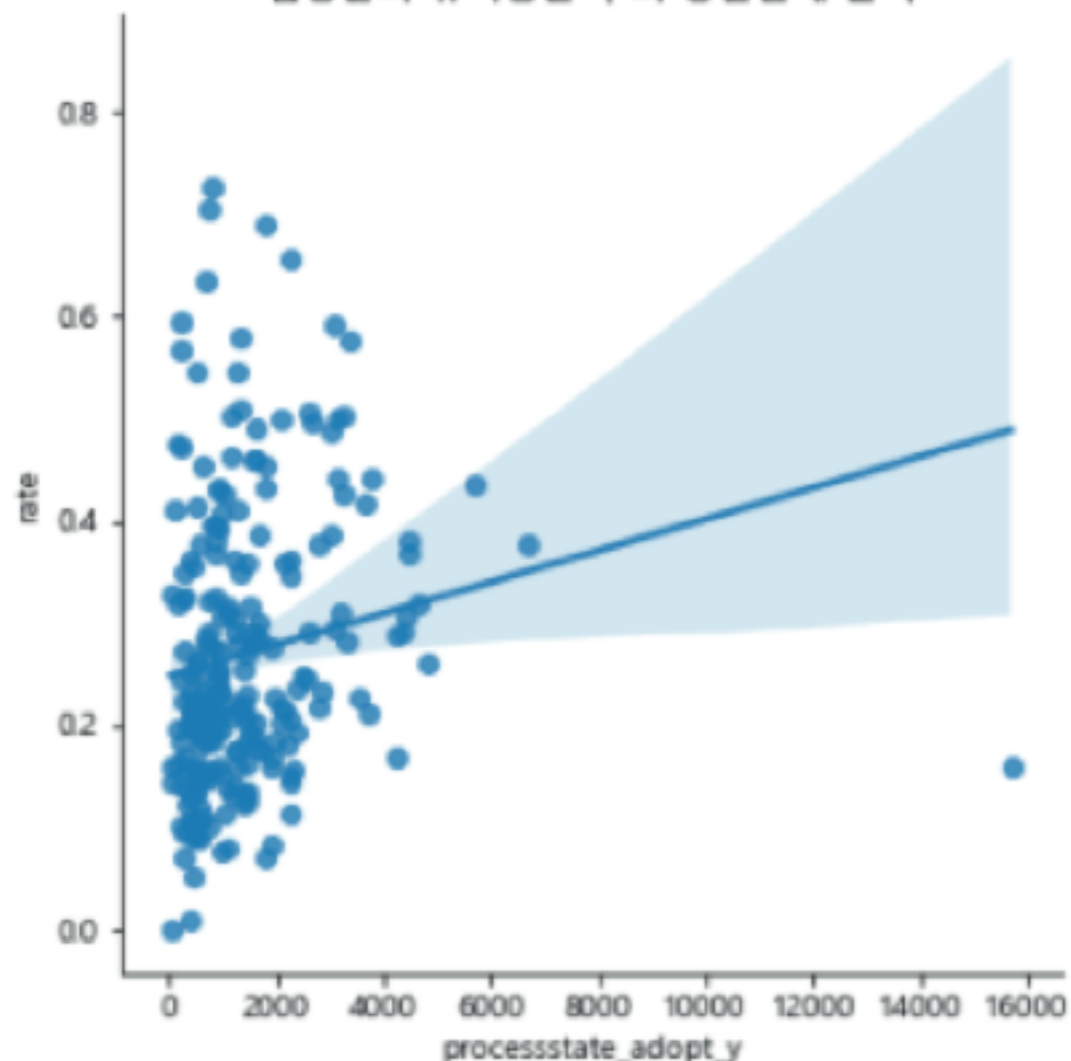
대전 17%
부산 15%
제주 14%

CONCLUSION

중요 FEATURES

입양률과 유기동물 수의 상관관계

입양률과 유기동물 수의 상관관계 분석



	processstate_adopt_x	processstate_adopt_y	rate
processstate_adopt_x	1.000000	0.836537	0.534415
processstate_adopt_y	0.836537	1.000000	0.163918
rate	0.534415	0.163918	1.000000

processstate_adopt_x = 입양된 유기동물 수
processstate_adopt_y = 총 유기동물 수
rate = 입양률 (입양된 유기동물 수/총 유기동물 수)

pearson 상관계수를 구함
rate와 총 유기동물 수 사이의 상관계수가 0.163918로 낮은 것을 확인 할 수 있음

유기동물의 입양 활성화를 위해 보호소 설치 증가 정책보다는 다른 방안이 필요함을 인식
입양률이 높은 지역의 유기동물 입양 추진 캠페인 시행 여부 확인

입양률이 높은 지역의 유기동물 입양 추진 캠페인 사례



입양률이 높은 전라북도

유기동물 사랑 나눔 캠페인

군산 유기동물 보호소와 (주)핏펫 연계 캠페인

-> 지자체의 적극적인 캠페인 추진

<http://www.koreadognews.co.kr/m/page/view.php?no=3082>

입양률이 낮은 제주도

여름 휴가철 동물 보호 캠페인

홍보물 배포, 현수막 게시 등 인식 제고 위주 캠페인

-> 상대적으로 소극적인 캠페인 추진

<http://www.headlinejeju.co.kr/news/articleView.html?idxno=424432>

입양 된 유기동물 특성 워드클라우드



CONCLUSION 중요 FEATURES

입양이 되지 않은 유기동물 특성 워드클라우드



CONCLUSION 분석결과/제언



7~10 월 반려동물 유기 증가 기간 집중 캠페인 시행
보호소 신규 개설 보다 직접적인 영향을 미치는 제도 마련
질병 및 사고로 입양 가능성이 낮은 유기동물의 질병 치료 우선



CONCLUSION 한계/보완점



데이터의 한계

공고 기간이나 유기 동물 입양 여부까지만 확인 가능
입양된 시점에 대한 데이터 부재
구조적인 데이터 수집 필요



분석 기법의 한계

입양에 중요한 영향을 미칠 것이라고 예상되는
유기동물의 이미지를 사용한 딥러닝 모델링 미적용
딥러닝을 통한 정확한 분류 예측 모델 필요



물류 시스템 분류 예측 모델 적용 가능



제발

사지마세요 입양하세요

Thank you !