



유기동물을 부탁해

**Tobigs 3rd conference
TEAM ?!**

CONTENTS

1. 주제선정배경
2. 데이터 정리
3. 모델 평가 방법
4. 예측모델
5. 활용방안
6. 결론 및 제언

불법으로 운영되는 번식장
약 3000여곳





'강아지 공장' 없앤다, 반려동물·할랄·코셔 산업 육성

MBC뉴스 - 2016. 7. 7.

강아지 공장' 없앤다, 반려동물·할랄·코셔 산업 육성 ... 우선 비윤리적인 강아지 번식 공장을 없애 반려동물의 유통부터 투명화한다는 계획입니다.

[투자확신하] '강아지공장' 타촌' 등 반려동물산업 양성화 ... 온라인판매도

심리

관련



'강아지공장' 없게 허가제로

서울신문 - 2016. 7. 7.

비위생적인 관리와 동물 학대로 문제가 된 '강아지 번식 공장' 사례를 막기 위해 반려 동물 산업 관리를 강화한다. 2020년이면 5조 달러(약 5770조원) ...



연한뉴스

강제 임신·불법 제왕절개... '강아지 공장' 전수조사

조선일보 - 2016. 6. 14.

애완용 강아지를 대량 공급하는 일명 '강아지 공장'에 대해 정부가 전수조사에 착수했다. 강아지 공장의 동물 학대 논란이 확대되자 실태를 파악 ...

농식품부, '강아지공장' 실태 전수 조사 실시

연합뉴스 - 2016. 6. 13.

관련 기사 보기 (기사 29개 더보기)

정부에서도 이러한 강아지 공장을 문제로 삼고 해결하기 위해 많은 노력

동물생산업장을 현행 신고제에서 허가제로, 생산시설 기준을 마련하는 등



Knowledge • 1,343 teams

Shelter Animal Outcomes

Mon 21 Mar 2016

Sun 31 Jul 2016 (16 days to go)

Dashboard

Home

Data

Make a submission

Information

Description

Evaluation

Rules

Timeline

Forum

Kernels

New Script

New Notebook

Leaderboard

Competition Details » [Get the Data](#) » [Make a submission](#)

Help improve outcomes for shelter animals

Every year, approximately 7.6 million companion animals end up in US shelters. Many animals are given up as unwanted by their owners, while others are picked up after getting lost or taken out of cruelty situations. Many of these animals find forever families to take them home, but just as many are not so lucky. 2.7 million dogs and cats are euthanized in the US every year.





자세히 보기

공고번호 경남-고성-2016-00060
접수일 2016-07-15
품종 푸들
성별 수컷
발견장소 경남 고성군 고성읍..
특징 양호
상태 공고중



자세히 보기

공고번호 경남-거제-2016-00343
접수일 2016-07-15
품종 믹스견
성별 수컷
발견장소 능포동 애견친구 &..
특징 온순
상태 공고중

유기동물 개요

유기동물 공고

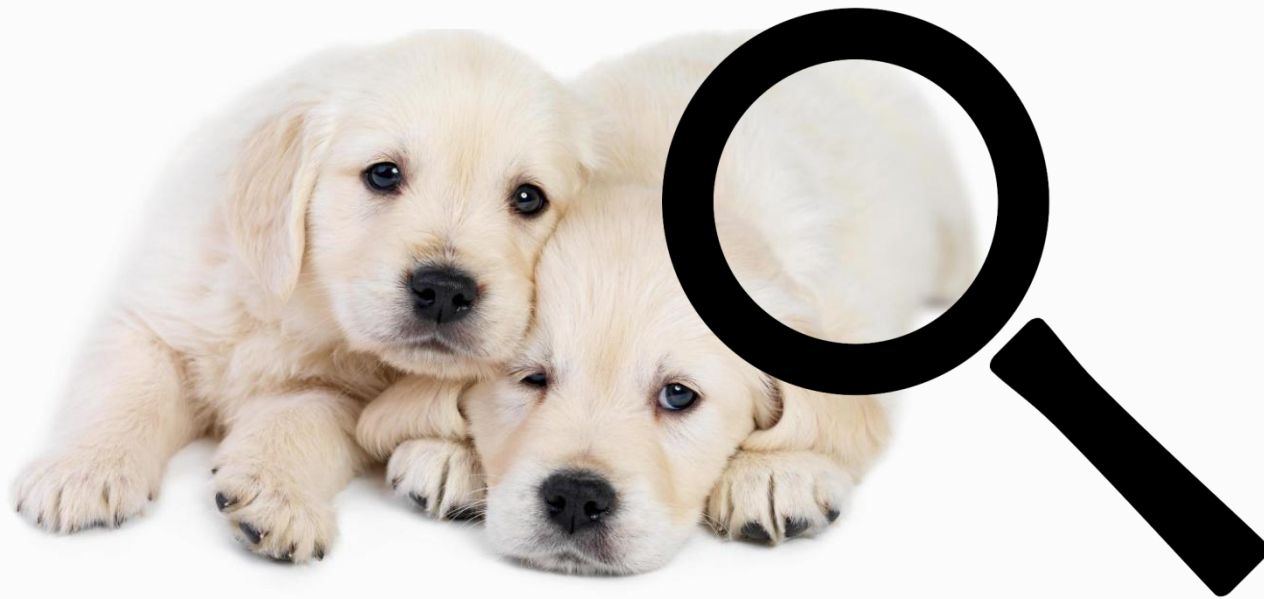
「동물보호법」 제17조, 같은 법 시행령 제7조 및 같은 법 시행규칙 제20조에 따라 구조된 동물의 보호상황을 아래와 같이 공고합니다.

공고번호	충남-부여-2016-00053
품종	[개]믹스견
색상	황색
성별	수컷
중성화 여부	아니오
나이/체중	2014(년생) / 30.00(Kg)
접수일시	2016-07-15
발생장소	충청남도 부여군 규암면 진변리 213-6
특징	
공고기한	2016-07-15 ~ 2016-07-25



** 유기동물 문의는 보호센터에 연락하시기 바랍니다.

보호센터이름	녹십자동물병원	전화번호	041-835-2879
보호장소	충청남도 부여군 부여읍 성왕로 202-3 녹십자동물병원		
관할기관	충청남도 부여군		
담당자	부여군	연락처	041-830-2254
특이사항			



이미 유기된 동물에 포커스

TOBIGS

유기동물을 부탁해

project

유기동물을 부탁해

170,000

동물보호관리시스템에서
총 십 칠만건의
데이터를 크롤링

8

종, 품종, 나이, 색,
들어 온 날짜, 나간 경로,
중성화 여부, 발견장소와 특징 파악

170,000

동물보호관리시스템에서
총 십 칠천건의
데이터를 크롤링

8

종, 품종, 나이, 색,
들어 온 날짜, 나간 경로,
중성화 여부, 발견장소의 특징 파악



이를 통해 특징에 따른 입양 가능성을 파악

변수명	수준
보호소	강원, 경기, 경남, 경북, 광주, 대구, 대전, 부산, 서울, 세종, 울산, 인천, 전남, 전북, 제주, 충남, 충북
개/고양이	dog , cat
성별	male ,female
중성화 여부	0 (no), 1(yes)
몸무게	몸무게
입양여부	안락사, 자연사 , 입양됨, 반환 등
연도	2014, 2015, 2016년
달	1-12월

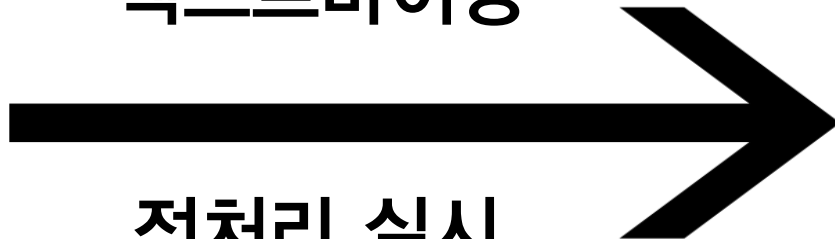
색상

품종

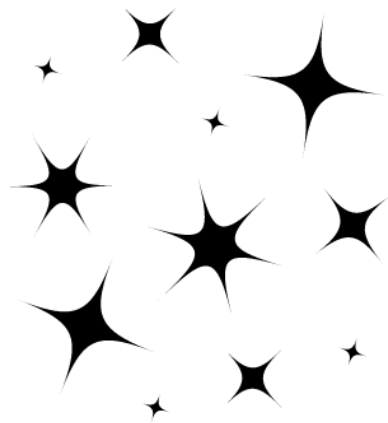
나이

특징

“텍스트마이닝”



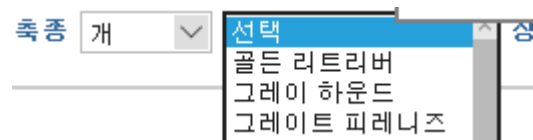
전처리 실시



- 표현의 통일

- ① 단위(나이)의 통일
(ex) 년, 월, 일 등 → 개월
- ② 색상은 동의어로 정리
(ex) 흰색, 하얀, 백구 등 → white

- 품종은 페이지 내에 있는 기준 변수를 이용하여 범주 정리



- 특징 ⇒ FREQUENCY TABLE

사람좋아함, 공격성, 겁있음, 활발, 순함, 건강함, 건강미흡, 사고여부, 병, 주인있음 → 0,1로 더미변수화

X
사나움, 백내장



accident	active	charm	disease	
0	0	0	1	
fear	like	mild	non	aggressive
0	0	0	0	1

최종 데이터 12만개



입양 : 34.6%
입양 × : 65.4%

상태 변수를 TARGET 변수

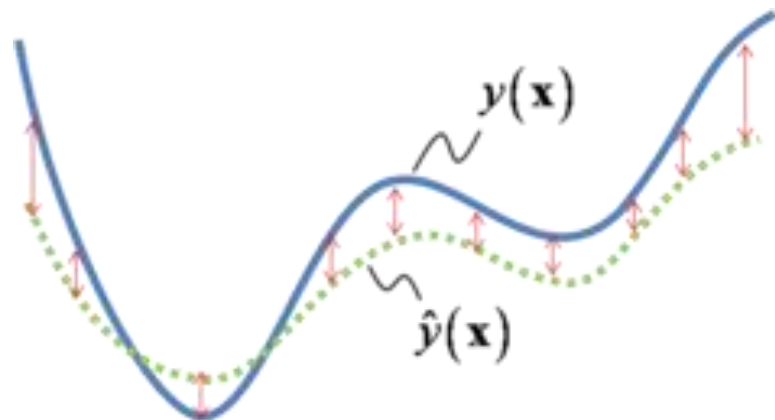


입양(YES) : 1
입양(NO) : 0

TRAIN (50%) : TEST (50%)



분석



$$RMSE = \sqrt{\frac{1}{n_{test}} \sum_{i=1}^{n_{test}} [y(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)]^2}$$

$y(\mathbf{x}_i)$: \mathbf{x}_i 에서의 실제 함수값

$\hat{y}(\mathbf{x}_i)$: \mathbf{x}_i 에서의 근사 함수값

n_{test} : 근사모델 평가(validation)를 위한 실험점의 수

TOBIGS 데이터 예측 모델링

시행횟수 ↓

MODEL →

1	Id	knn1	knn3	logistic	tree.3	bagging	boosting2	svm1	lasso	naive
2	0.33551	0.358449	0.33868	0.329321	0.329211	0.347425	0.32351	0.346029	0.343514	0.469216
3	0.334254	0.359469	0.338146	0.33089	0.327343	0.336425	0.334306	0.345982	0.343514	0.483608
4	0.335667	0.355373	0.340908	0.331173	0.32816	0.33925	0.328798	0.346029	0.346025	0.46942
5	0.336263	0.359642	0.341568	0.330482	0.328834	0.343487	0.331199	0.346029	0.344989	0.46942
6	0.336953	0.359814	0.34094	0.330859	0.328254	0.339815	0.329881	0.344946	0.344204	0.468447
7	0.337205	0.358653	0.339998	0.329839	0.333511	0.338873	0.327574	0.346029	0.345334	0.469672
8	0.335384	0.360238	0.341662	0.331377	0.327139	0.343251	0.329975	0.346029	0.34411	0.470974
9	0.334442	0.35735	0.341568	0.331314	0.32918	0.345417	0.33021	0.346076	0.344392	0.469687
10	0.336577	0.358731	0.339245	0.329855	0.325523	0.342357	0.337084	0.346029	0.344236	0.470189

TOBIGS

예측모델

LDA

NAIVE

KNN

BAGG
ING

예측 모델

LASSO

BOOST
ING

LOGIS
TIC

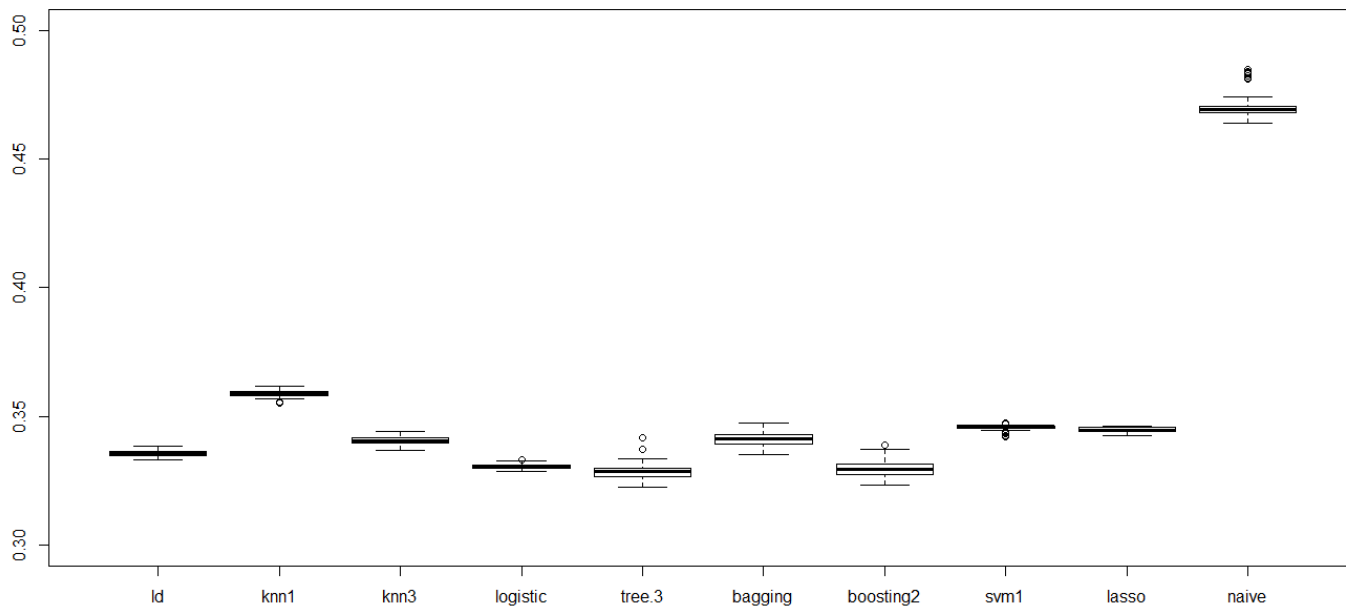
SVM

TREE

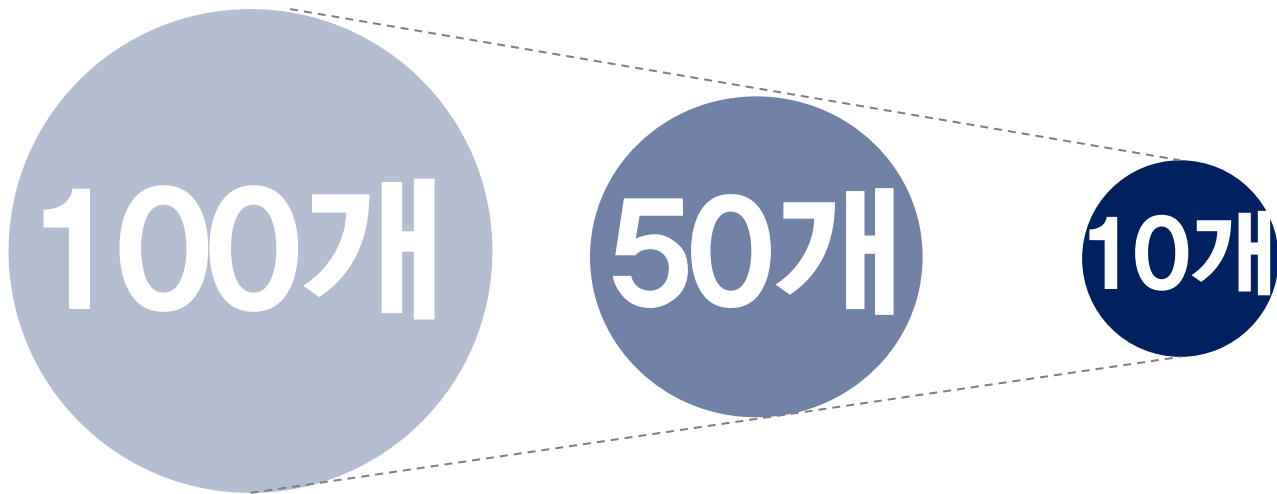
전체 데이터를 사용하면?

BEST MODEL “TREE”

boxplot 100 simulation

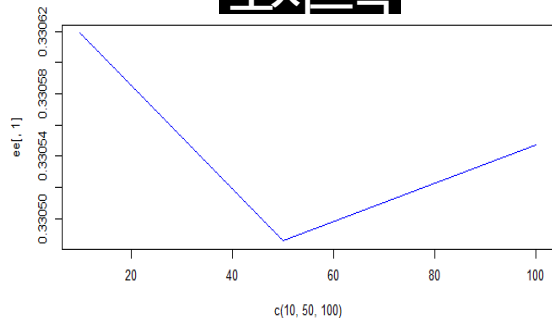


시뮬레이션을 줄이면?

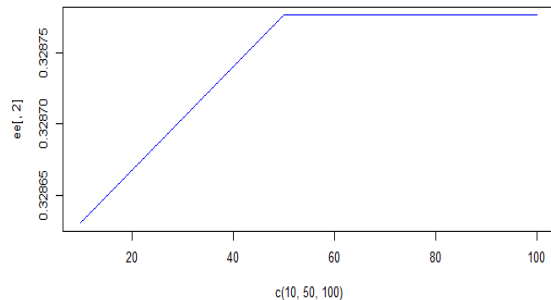


SIMULATION NUMBER를 줄여도 BEST MODEL이 동일한지,
그리고 **오류률**이 비슷하게 나오는지 확인

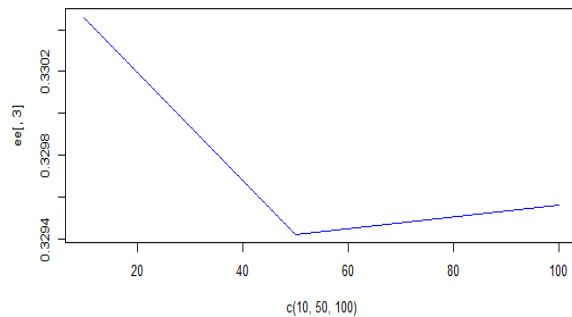
로지스틱



트리

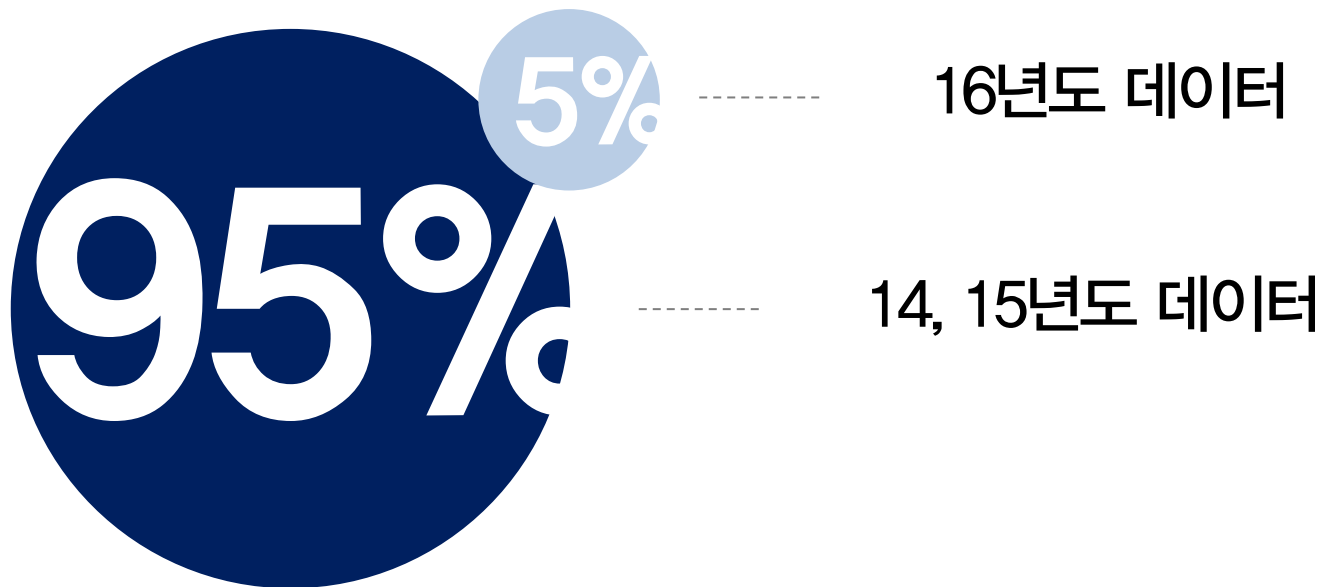


부스팅

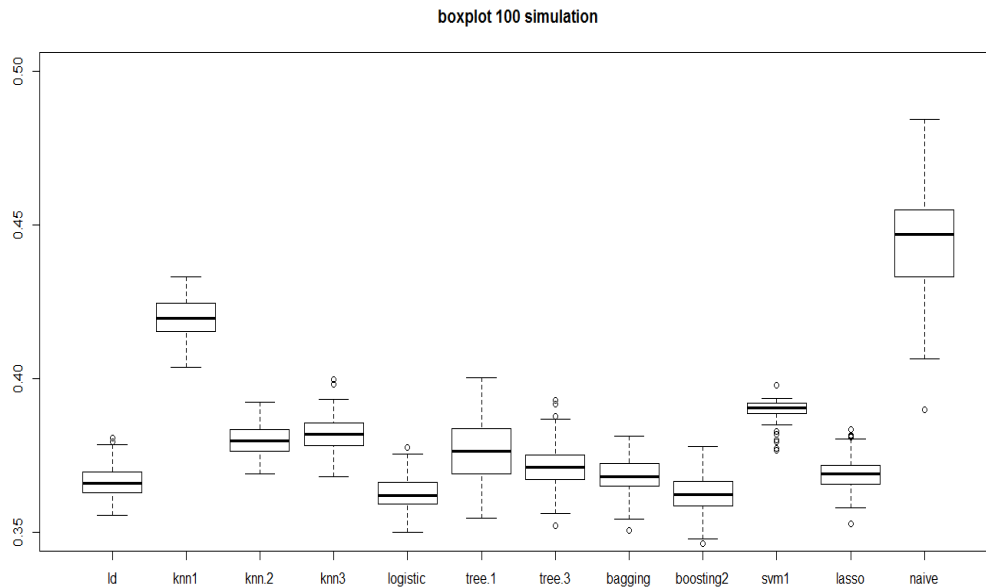


그 결과, SIMULATION NUMBER를 변경해도
best 모델은 **TREE MODEL** 로 동일

2016 data만을 사용할 경우엔?

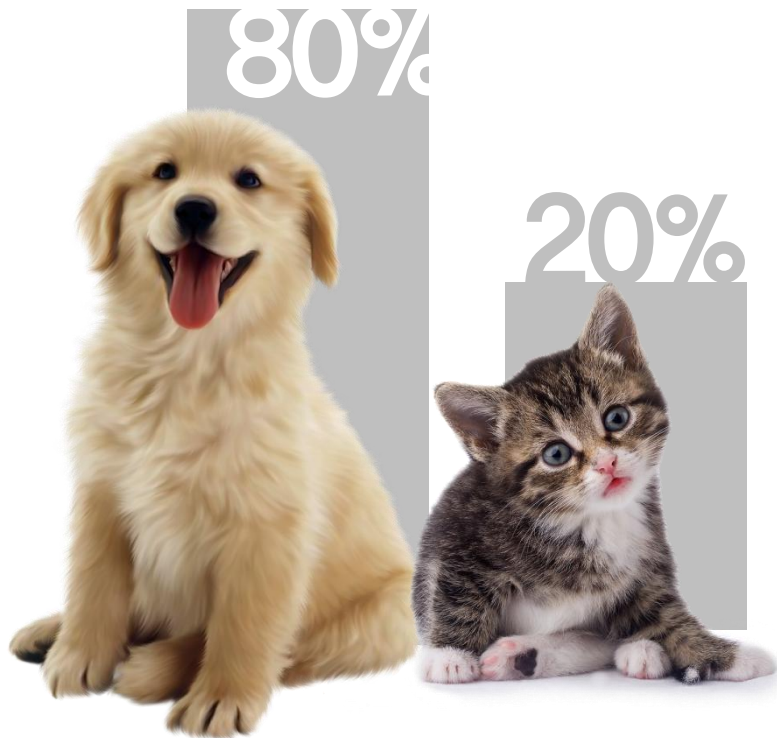


SO 2016만으로도 같은 결과가 나오는지 확인

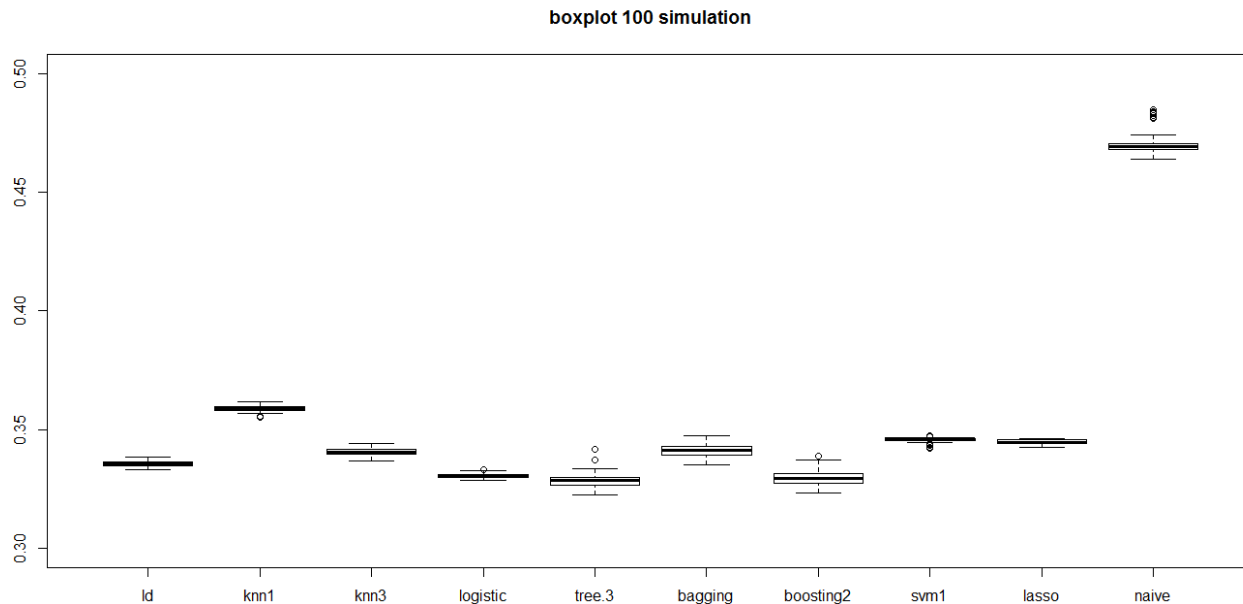


2016 데이터만을 이용했을 때는
Best model로 **로지스틱**이 뽑혔다

고양이 데이터만을 사용할 경우엔?



전체 데이터의 **20%**만을
차지하는 고양이 데이터만으로
분석이 가능할까?



고양이 데이터만을 이용했을 때는
Best model로 **Tree Model**이 뽑혔다



이러한 예측결과를 어떻게 활용할수 있을까?

미국



한국



공고기간보호기간

공고기간보호기간

10

day

10

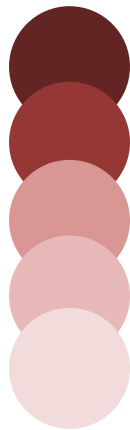
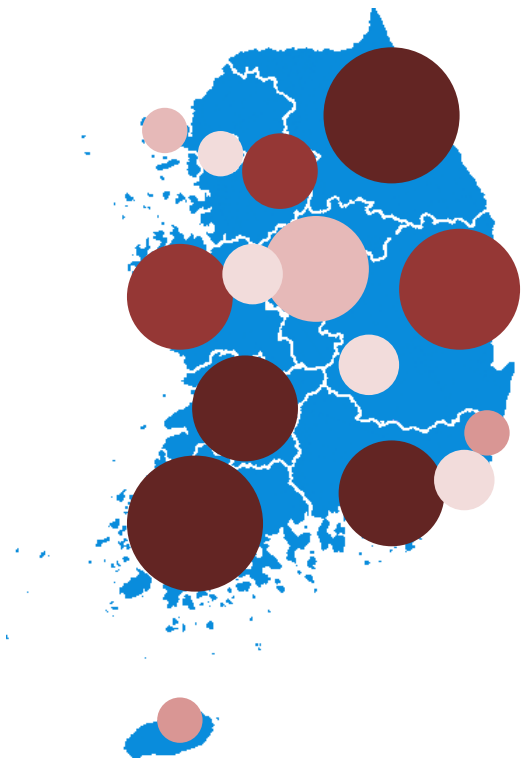
day



연관성 분석으로 보는 "입양" Insight !

1. 보호소 지역별
2. 중성화의 여부
3. 질병의 여부

보호소 지역별 입양



입양률 매우 높음 | 강원, 전남, 광주, 전북, 경남

입양률 높음 | 경북, 충남, 경기

입양률 보통 | 제주, 울산

입양률 낮음 | 인천, 충북

입양률 매우 낮음 | 서울, 부산, 대전, 대구

중성화에 따른 입양



연관성 분석을 통해서 유기동물의
중성화 여부와 입양여부는 **서로 독립**



중성화보다는 다른 요인을
우선적으로 고려해야할 것

질병에 따른 입양

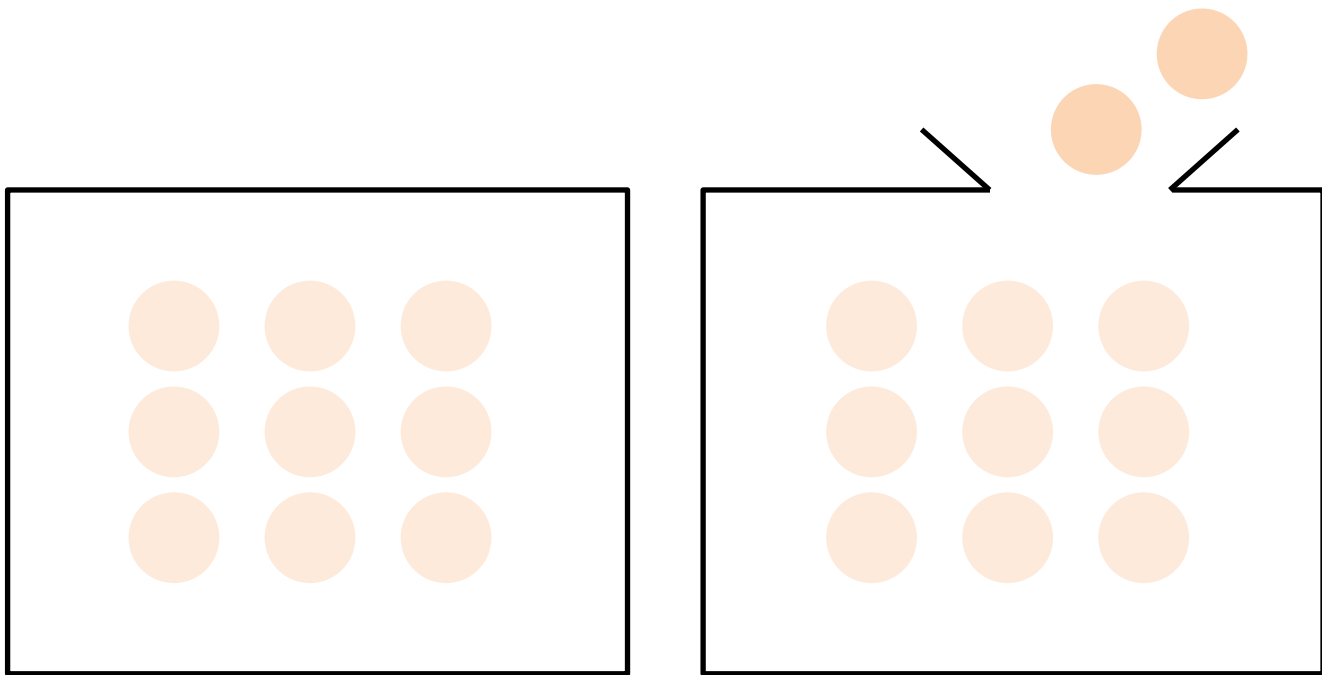
질병이 없는 경우, 입양에 큰 영향 미치지 않지만 있는 경우 부정적인 영향



입양 가능성 낮은 유기동물들의
질병치료가 가장 우선시 되어야



제안 I



지역별 선례를 다른 지역으로 확대하고,
각 동물보호 시민단체나 정부부처에서
효율적인 계획을 수립할 수 있도록 제도적 지원 마련!

제안 |



중성화보다는 입양가능성이 낮은
동물들의 "질병치료"를 우선 시



보다 많은 유기동물들에게
입양의 기회 마련



Thanks a lot!

**Tobigs 3rd conference
TEAM ?!**