

K-하이테크 플랫폼 지역주민 리빙랩 지원

결과보고서

팀 명		구구구칠
아이디어명		길고양이 중성화 데이터를 이용한 대구시 길고양이 분석 데이터 서비스
성 명	팀 대표	임소영
	팀 원	이화은
		우승연
활동기간		2024.04.01. ~ 2024.06.14. (약 10주)

활용 데이터	국가동물보호정보시스템 (URL : https://www.animal.go.kr/front/awtis/roadCat/roadCatList.do?menuNo=5000000025)
--------	---

목 차


1. 활동 요약 3p
2. 추진 배경 4p
3. 리빙랩 활동 5p
4. 추진 내용 16p
5. 증빙 자료 18p
6. 소요 예산 내역 20p

K-하 이 테 크 플 랫 폼 지 역 주 민 리 빙 랩 지 원 결 과 보 고 서

< 지역주민 리빙랩 지원 활동 요약 >

문제점	개선
<ul style="list-style-type: none"> 길고양이가 주민들의 불안감과 생태계 교란을 야기 안락사가 아닌 중성화를 통해 점진적인 개체 수 조절이 필요 길고양이 개체 수와 증가폭 예측 분석을 통해 대책에 대한 논의가 필요 고양이 학대 방지를 위한 동물보호법 안내 수단에 대한 필요성 제기 	<ul style="list-style-type: none"> 각 지역별로 길고양이 관련 정보를 통계적으로 분석 알고리즘을 사용하여 분석을 통해 지역별 길고양이 및 중성화 관련 빅데이터 분석 결과 제시 인터넷 및 서적을 통해 길고양이로 인해 야기되는 각종 피해/문제에 대한 대비책 제시
활용 데이터	국가동물보호정보시스템 (URL : https://www.animal.go.kr/front/awtis/roadCat/roadCatList.do?menuNo=5000000025)

팀 명		구구구칠
주 제		길고양이 중성화 데이터를 이용한 대구시 길고양이 분석 데이터 서비스
아이디어명		길고양이 중성화 서비스
성 명	팀 대표	임소영
	팀 원	이화은
		우승연
활동기간		2024.04.01 ~ 2024.06.14.
활동지역		대구광역시 북구

A photograph showing three people in a room with wood-paneled walls. Two women and one man are gathered around a desk with two computer monitors. One woman is pointing at the screen while the others look on. The man is wearing a white cap. The room has warm lighting and a rustic feel.



1. 추진 배경

아이디어 개요	<p>길고양이 (유기묘) 관련 예측 지도:</p> <p>본 프로젝트는 신규 발생하는 길고양이, 즉 새로 태어나는 길고양이들의 위치를 예측하여 미리 포획 및 중성화 계획을 세우는 것을 목표로 한다. 이를 통해 길고양이 개체 수를 효과적으로 관리하고, 이로 인한 주민들의 피해를 줄일 수 있을 것으로 보인다.</p>
배경 및 필요성	<p>길거리에서 다양하게 만날 수 있는 동물 중 유독 사람들의 불만을 받고 있는 것이 유기묘들이다.. 특히 5 - 6월이 되면 주택가 등에서 발정기 고양이들에 의해 발생하는 메이팅 콜(교미를 위해 수컷을 부르는 암컷 고양이의 울음소리)로 인한 소음 발생과 그로 인해 조성되는 음산한 분위기로 인해 피해를 호소하는 사람들이 많다. 이러한 소음은 특히 밤 시간대에 발생하여 주민들의 수면을 방해하고, 지속적으로 반복되면서 심리적인 스트레스를 유발하기도 한다</p> <p>유기묘 문제는 단순히 소음 공해에 그치지 않는다. 이들 고양이들은 제대로 된 돌봄을 받지 못하기 때문에 질병에 걸리기 쉽고, 이는 주변의 다른 동물들에게도 전파될 수 있는 위험을 내포하고 있습니다. 또한, 고양이들이 쓰레기통을 뒤져 음식물을 찾거나, 주택의 정원 등을 파헤쳐 놓아 경제적 피해도 발생한다.</p> <p>만약 우리가 '중성화되지 않은 신규 고양이 개체의 수'를 대략적으로라도 예측할 수 있다면, 이를 통해 보다 효과적인 중성화 프로그램을 계획하고 실행할 수 있을 것이고, 중성화 수술을 통해 번식을 억제하면 길거리 고양이 수가 자연스럽게 감소하게 되어, 장기적으로 유기묘 문제를 해결하는 데 큰 도움이 될 수 있다.</p> <p>본 프로젝트는 이러한 문제의식을 바탕으로 기획되었다. 우리는 데이터 분석을 통해 주택가 내 중성화되지 않은 고양이의 수를 예측하고, 이를 기반으로 최적의 중성화 수술 시기와 대상을 선정할 것이다. 이를 통해 주민들의 피해를 최소화하고, 길고양이와 사람 간의 공존을 도모하고자 한다.</p>
목표	<p>각 지자체에서는 이러한 유기묘 문제를 해결하기 위해 길고양이들의 중성화 수술을 지원하는 사업을 진행하고 있습니다. 중성화 수술은 길고양이의 개체 수를 조절하고, 이로 인해 발생하는 다양한 문제를 줄이는 데 매우 효과적인 방법이다. 포획 위치와 방사 위치를 알 수 있다면 유독 중성화 수술을 적게 시행한 지역을 파악할 수 있다. 이러한 데이터는 특정 지역에서 중성화 수술이 얼마나 효과적으로 이루어졌는지 평가하는 데 중요한 역할을 한다.</p> <p>특히, 중성화 수술이 적게 시행된 지역에서는 신규 고양이 개체 수가 증가할 가능성이 높다. 이로 인해 해당 지역 주민들이 겪는 소음 공해와 기타 피해도 더 심각할 수 있다. 당 프로젝트는 이러한 문제를 해결하기 위해 기획되었다. 우리는 데이터 분석을 통해 중성화 수술이 적게 시행된 지역을 식별하고, 해당 지역의 신규 고양이 개체에 특별한 관심을 기울일 것이다. 이를 통해 해당 문제로 인해 야기되는 피해를 감소시키는 것이 우리의 목표이다.</p>

2. 리빙랩 활동

① 문제점

길고양이는 다른 동물에 비해 귀여운 외모로 야생 동물임에도 불구하고 먹이를 주거나 돌보는 사람이 많다. 또한 번식률이 높고 도심의 길고양이들은 천적이 없는 등의 이유로 개체 수가 증가하게 되었다

그러나 이로 인해 여러 가지 문제를 초래하고 있다.

길고양이 개체 수 증가와 관련된 문제점

- ① 주거 지역 내 갈등 발생
 - 길고양이의 소음과 배설물로 인한 주민 불편
 - 쓰레기 봉투 및 차량 훼손 등의 피해 발생
 - 길고양이를 선호하지 않는 주민들의 불안감 유발
- ② 생태계 교란
 - 고양이에 의한 조류 사냥으로 생태계 균형 파괴
 - 적이 없는 도심 지역에서 개체 수 급증
- ③ 위생 및 미관 문제
 - 고양이 보호자들에 의해 방치된 사료 그릇과 사료로 인한 악취 발생
 - 주변 환경 미관 저해

길고양이 문제의 근본 원인은 길고양이의 귀여운 외모와 높은 번식률, 그리고 도심 지역 내 천적 부재 등으로 인한 개체 수 급증에 있다. 이로 인해 주거 지역 내 갈등, 생태계 교란, 위생 및 미관 문제 등 다양한 부작용이 발생하고 있다.

② 아이디어 변화

길고양이는 사람들에 따라 다르게 인식되고 있다. 일부 사람들에게는 귀여움과 애정의 대상이지만, 지역 사회에서는 소음, 악취, 건물 파손 등의 문제를 일으키는 존재로 여겨지기도 한다. 이러한 갈등을 해결하기 위해, 많은 지방자치단체에서는 'TNR(Trap-Neuter-Return)' 사업을 시행하고 있다. 본 활동에서는 TNR 사업에 대해 두 가지 관점에서 분석을 진행하였다. 먼저 TNR 사업을 통해 수집된 길고양이 데이터를 바탕으로 머신러닝 알고리즘을 사용한 지역별 길고양이 특성 분석 및 분류 모델을 만들었다.

그리고 같은 데이터를 바탕으로 TNR 사업이 활발한 지역과 그렇지 않은 지역의 개체 수를 비교하는 등의 길고양이 개체 수 관련 경향 분석을 진행하였다.

③ 아이디어 상세 설명

[길고양이 특성들을 활용하여 대구 권역을 유추하는 분류 모델 생성]

각 권역별 길고양이 특성에 대한 분류 작업을 실시하면 약간의 단서가 되는 길고양이의 정보만으로 권역 유추가 가능하지 않을까? 하는 아이디어에서 시작되었다.

예를 들어, 고양이의 성별, 몸무게, 털 색을 입력하면 어느 지역에서 가장 많이 발생할

길고양이의 특성을 입력해서 각 권역별에 해당할 가능성이 높은 지역을 도출해내고자 한다.

도출된 결과를 분석하여 추후 문제를 더 확인하고 이에 대한 대비책 및 보완점을 구축하고자 한다.

[길고양이 개체 수 관련 경향 분석]

먼저 TNR 사업의 개요 및 목적 대상에 대해 조사하고 시행 현황에 대해 알아봄으로써 각 지자체에서 TNR 사업에 어떤 노력을 쏟고 있는지 확인한다.

대구광역시 내에서 가장 많은 개체가 포획된 지역을 도출하여 해당 지역은 다른 지역과 비교하고 어떤 특징이 있는지 분석한다. 또한 TNR이 해당 지역에 어떤 영향을 끼쳤는지 알아본다.

④ 과정 및 활동 내용

[주제 선정을 위한 검색어 관련 wordcloud 생성]



[길고양이 특성들을 활용하여 대구 권역을 유추하는 분류 모델 생성]

참고 사이트에서 필요한 정보를 선정하고 이를 활용할 수 있는 데이터를 수집하는 단계가 선행되어야 했다. 먼저 BeautifulSoup4 모듈을 이용하여 Chrome에서 크롤링을 진행하였다. 지역, 수집장소, 성별, 중성화 여부, 종에 대한 내용을 크롤링하였다. 약 5000개의 내용이 크롤링한 후 csv 파일 형식으로 저장하였다.

그 후 사용할 머신러닝 모델에 대한 탐구를 하였다. scikit-learn의 document에 제시된 모델들 중 수집된 자료를 가장 적절하게 활용할 방향을 고민하여 3가지를 결정하였다.

그 3가지로는 LogisticRegression, KNeighborsClassifier, RandomForestClassifier으로 도출되었다. 도출된 결과를 바탕으로 분석을 실시하고 아이디어를 구체화하였다.

[길고양이 개체 수 관련 경향 분석]

1. 세부 주제 :

- 1) 대구광역시 내 TNR 개체 최다 포획 지역 특징 분석
- 2) 시군구별 TNR 건수와 해당 시군구의 인구 수 비교

2. 데이터 정보 :

- 1) 파일 형식 : csv
- 2) 내용
 - (1) 포획 기간 : 2023년 4월 27일 ~ 2024년 05월 01일
 - (2) 포획 지역 : 대구광역시 전역

.....
(3) 포획 건수 : 총 3,719건

① 지역별 포획 건수

- 달서구 604 건
- 수성구 526 건
- 달성군 506 건
- 중구 466 건
- 북구 391 건
- 서구 387 건
- 동구 374 건
- 군위군 241 건
- 남구 216 건

② 몸무게

- 평균 몸무게 : 3.576955kg
- 최소 몸무게 : 1.900000kg
- 최대 몸무게 : 7.500000kg

3. 분석 방법 :

1) 데이터 수집 :

- 공공데이터포털 길고양이 중성화사업(TNR)의 공공데이터포털 LINK 서비스에서 길고양이 중성화 사업(TNR) 대상 개체의 정보를 csv파일로 다운로드

2) 데이터 전처리 :

(1) csv 파일을 Pandas의 데이터프레임 형태로 변환

(2) 관리번호, 포획 사진 및 중성화 사진, 수술 이후 사진, 방사 사진 컬럼 삭제

(3) 추가 컬럼 삭제

① 사업지역(시도) : 대구 지역 자료이기 때문에 삭제

② 털색, 품종, 특이사항, 건강상태 : 분석 내용과 상관 관계가 없어 삭제

③ 연령 : 추정치가 많아 삭제

④ 종결 정보 : '상태' 컬럼과 일치

⑤ 상태가 '방사'가 아닌 경우를 찾아 종결 정보 확인 (총 8건)

⇒ 수술 후 사망한 경우 해당 개체는 수술로 인한 개체 수 변화에 영향을 끼치지 않을 것으로 판단하여 해당 행 삭제 후 '상태' 열 삭제.

(4) 포획 일자, 방사 일자 datetime으로 형변환

(5) 포획 장소 및 방사 장소 일치 여부 확인

⇒ 지역별 개체 수 변동 가능성 최소화

※ 이슈 발생 : 주소 작성 방식 차이 및 오타로 인해 포획 장소와 방사 장소 불일치 건수 총 850건

⇒ 방사포획일치 컬럼을 추가하고 'preprocess.csv'로 파일 저장 후 데이터 확인

① 1차 확인 (850건→ 43건)

- 띄어쓰기 차이 : (예 : 가창리1037 - 가창리 1037)

• 오기 : (예 : 녹원맨션 - 녹원맨션, 국채보상로 - 구개보상로)

• 누락 : (예 : 월배로 83안길18 - 월배로 83안길18 2층, 성동로37길39-3 - 성동로37길39)

.....

② 2차 확인 (43건 → 24)

- 인근인 경우 : (예: 동덕로30길 45 - 동덕로 122, 구지면 과학마을로 3길 - 구지면 과학마을로 2길, 옥수천로7 - 달구벌대로650길78)
- 주소 표기 차이 : (화원읍 화원로 36 - 화원읍 천내리 118)

③ 삭제 행

- 오기 여부 확인에 어려움이 있는 경우 (예 : 대구광역시 동구 팔공산로199길 12 - 대구광역시 팔공산로 119길 12)*

⇒ 수정한 파일을 데이터 프레임으로 만들고 `df = df[df['방사포획일치']==True]`로 필터링 후 '방사포획일치'행 삭제.

(6) 성별 범주화 (0: 수컷, 1: 암컷)

⑤ 결과 및 기대효과

[길고양이 특성들을 활용하여 대구 권역을 유추하는 분류 모델 생성]

1. 전처리

크롤링한 데이터를 활용 중 본 목적에 사용하기 위해서는 데이터 부족하다는 것을 판단하였다.

이를 보완하기 위해 크롤링을 더 시도해보려고 했으나 조원의 데이터를 활용하였다.

전처리를 다음과 같이 진행하였다.

- '연령' 열의 데이터 정리

아래와 같이 무질서하게 제시된 연령대를 통일성을 갖추어 제시하였다.

모두 '~살'의 형태로 통일하였고, 정제 시도를 하기 힘든 경우 '추정 불가'로 처리하였다

```
1 data['연령'].replace(to_replace=['03월 04일', '5-6개월', '5개월령', '10개월추정', '5개월추정', '6개월', '6달추정', '6개월추정',
2 data['연령'].replace(to_replace=['1살미만', '1세 추정', '7개월 추정', '8개월 추정', '8개월', '9개월 추정', '10개월', '7개월령',
3 data['연령'].replace(to_replace=['1살추정', '1년추정', '1세령', '1년', '1살 추정', '1', '1세추정', '1년 추정', '1년령', '1세',
4 data['연령'].replace(to_replace=['2살추정', '2세 추정', '2세추정', '1-2살추정', '2년추정', '2년령', '2살 추정', '2살', '2세', '2살',
5 data['연령'].replace(to_replace=['2-3년추정', '3살추정', '3세추정', '3년추정', '3살 추정', '3세 추정', '3년', '3', '3살', '3년',
6 data['연령'].replace(to_replace=['4살 추정', '6-8달추정', '4세 추정', '3-5살추정', '4년', '4살추정', '4', '4년 추정', '4년추정',
7 data['연령'].replace(to_replace=['5살이상', '5년', '5세추정', '5년령', '5', '5세 추정', '5살추정', '2-5년추정', '5살 추정'],
8 data['연령'].replace(to_replace=['6세추정', '6살추정', '6살 추정', '6세 추정'], value='6살', inplace=True)
9 data['연령'].replace(to_replace=['7살 추정', '7년령', '7살추정', '7세추정', '7세 추정'], value='7살', inplace=True)
10 data['연령'].replace(to_replace=['8살 추정', '8살추정', '8세추정', '8세 추정'], value='8살', inplace=True)
11 data['연령'].replace(to_replace=['9세 추정', '9살추정'], value='9살', inplace=True)
12 data['연령'].replace(to_replace=['10세추정', '10살 추정'], value='10살', inplace=True)
13 data['연령'].replace(to_replace=['11', '10', '06월 07일', '8', '04월 05일', '05월 06일'], value='추정불가', inplace=True)
```

- '품종' 열의 데이터 정리

위의 '연령' 열과 같이 통일성 있는 자료 형식으로 제시하였다.

- '종결정보' 열의 데이터 정리

위의 '연령', '품종' 열과 같이 통일성 있는 자료 형식으로 제시하였다.

- 형변환

'category'로 형변환이 필요한 경우 형변환을 실시하였다.

2. 시각화

머신러닝 프로젝트에서 데이터의 시각화는 매우 중요한 단계이다. 시각화를 통해 데이터의 패턴, 경향, 이상치 등을 파악할 수 있기 때문이다. 이렇게 데이터를 탐색하고 이해하는 과정은 모델 구축 및 하이퍼파라미터 튜닝 등의 후속 작업에 큰 영향을 미친다.

먼저 데이터의 특성에 따라 적절한 시각화 기법을 선택한다.

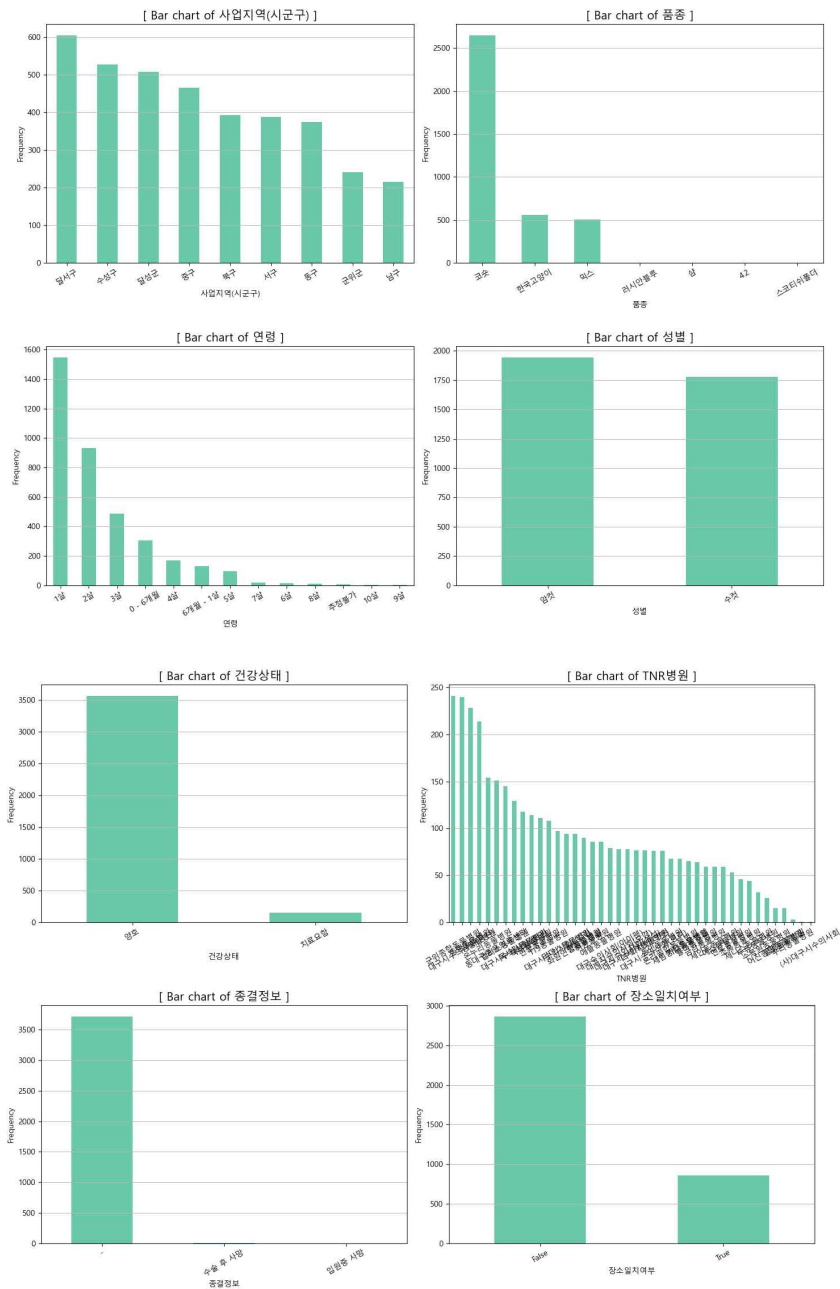
• 연속형 변수라면 히스토그램, 산점도 등이, 범주형 변수라면 파이차트, 바차트 등이 유용하다.

• 다변수 데이터의 경우 산점도 행렬, 평행 좌표계 등의 기법이 유용하다.

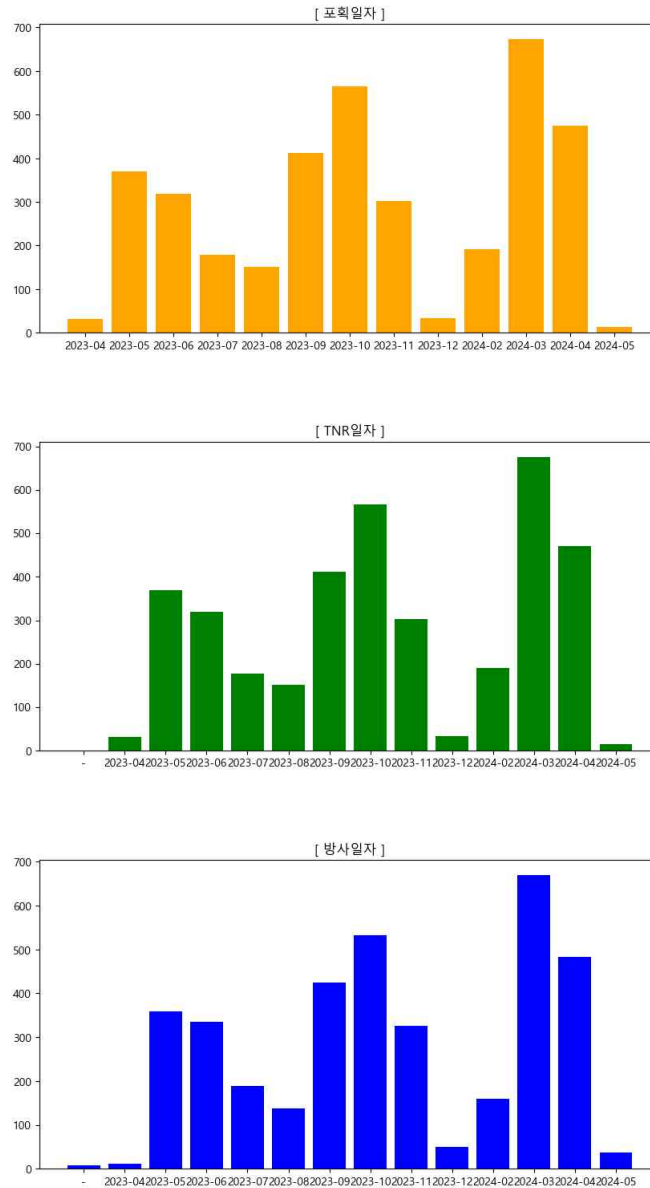
시각화를 통해 특징 간의 상관관계, 클래스 분포, 이상치 등을 파악하면 전처리 과정에서 어떤 작업이 필요한지 가이드라인을 얻을 수 있다.

예를 들어 특징과 타겟 간의 상관관계가 약하다면 해당 특징을 제거하는 것이 모델 성능 향상에 도움이 될 수 있다. 또한 시각화 결과를 바탕으로 데이터 증강, 오버/언더 샘플링 등의 전략을 세울 수 있다.

클래스 분포가 불균형적이라면 리샘플링이 필요할 것이다. 요컨대 시각화를 통한 데이터 탐색은 전체 머신러닝 파이프라인에서 중요한 역할을 하며, 프로젝트의 목표와 데이터 특성에 맞는 시각화 기법을 선택하는 것이 관건이다. 간단히 데이터의 분포를 확인하고 머신러닝 주제의 타겟팅을 위해 시각화를 실시하였다.



또한 월별로 진행 건수에 대해 간단히 시각화 해보았다.



3. 머신러닝 모델

데이터의 몇가지 정보를 가지고 사업 지역을 구분하는 머신러닝을 제작하였다.

- ① 먼저 위에서 진행된 전처리한 파일을 새로운 csv 파일로 저장하였다.
- ② 저장된 csv 파일이 정상적으로 읽히는지 확인하였다.
- ③ 그 후 사용할 데이터만 남기며 데이터프레임을 형성하였다.
- ④ 머신러닝을 위해 '연령'의 컬럼을 숫자로 변환하는 원핫인코딩을 진행하였다.
- ⑤ 그 후 상관관계 확인이 가능한 숫자 관련 데이터를 이용하였다.
- ⑥ 머신러닝을 위해 데이터셋을 분리하였다. 이때 train : test = 0.8 : 0.2 로 설정하였다.
- ⑦ 데이터 스케일링을 StandardScaler를 이용하여 진행하였다.
- ⑧ LogisticRegression, KneighborsClassifier, RandomForestClassifier을 이용하여 머신러닝을 진행한 후 결과

를 비교하였다.

⑨ 추후 시계열분석을 위하여 각 일자별로 저장된 데이터들을 월을 기준으로 분류하여 데이터를 분류하였다.

⑩ 관련 코드는 github에 업로드하였다.

https://github.com/Livinglab9997/livinglab_9997

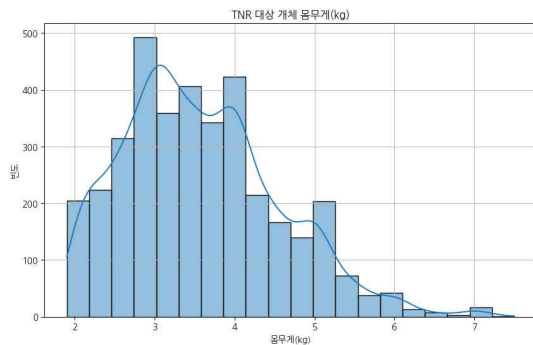
모델 종류	결과 및 성능 비교				
Logistic Regression	Accuracy: 0.30				
	Classification Report:				
		precision	recall	f1-score	support
	군위군	0.48	0.88	0.62	32
	남구	0.00	0.00	0.00	46
	달서구	0.20	0.28	0.23	131
	달성군	0.38	0.30	0.34	105
	동구	0.19	0.41	0.25	64
	북구	0.00	0.00	0.00	90
	서구	0.29	0.06	0.10	79
	수성구	0.28	0.43	0.34	101
	중구	0.48	0.52	0.50	96
	accuracy			0.30	744
	macro avg	0.26	0.32	0.27	744
	weighted avg	0.26	0.30	0.26	744
Kneighbors Classifier	Confusion Matrix:				
	[[28 0 1 0 1 0 0 2 0]				
	[2 0 19 0 10 0 0 15 0]				
	[5 0 37 20 24 0 0 44 1]				
	[0 0 24 32 21 0 9 6 13]				
	[0 0 19 2 26 0 0 17 0]				
	[4 0 31 10 33 0 0 10 2]				
	[8 0 15 0 7 0 5 16 28]				
	[11 0 30 0 6 0 1 43 10]				
	[0 0 12 20 12 0 2 0 50]]				
	Accuracy: 0.36				
	Classification Report:				
		precision	recall	f1-score	support
	군위군	0.75	0.84	0.79	32
	남구	0.10	0.11	0.11	46
	달서구	0.35	0.36	0.35	131

	달성군	0.42	0.46	0.44	105
	동구	0.21	0.22	0.21	64
	북구	0.31	0.28	0.29	90
	서구	0.24	0.20	0.22	79
	수성구	0.33	0.33	0.33	101
	중구	0.56	0.54	0.55	96
	accuracy			0.36	744
	macro avg	0.36	0.37	0.37	744
	weighted avg	0.36	0.36	0.36	744
	Confusion Matrix:				
	[[27 0 2 0 0 1 1 1 0]				
	[0 5 12 4 5 5 2 11 2]				
	[3 8 47 15 12 11 8 23 4]				
	[0 4 16 48 11 9 5 2 10]				
	[2 4 16 13 12 25 7 8 3]				
	[2 4 13 10 5 5 16 11 13]				
	[2 10 10 7 7 8 18 33 6]				
	[0 6 8 13 2 8 7 0 52]]				
Random Forest Classifier	Accuracy: 0.37				
	Classification Report:				
	precision	recall	f1-score	support	
	군위군	0.68	0.84	0.75	32
	남구	0.17	0.17	0.17	46
	달서구	0.33	0.31	0.32	131
	달성군	0.47	0.48	0.47	105
	동구	0.16	0.19	0.17	64
	북구	0.28	0.31	0.29	90
	서구	0.32	0.27	0.29	79
	수성구	0.38	0.34	0.36	101
	중구	0.54	0.57	0.56	96
	accuracy			0.37	744
	macro avg	0.37	0.39	0.38	744
	weighted avg	0.37	0.37	0.37	744
	Confusion Matrix:				
	[[27 0 2 0 0 1 1 1 0]				
	[0 8 11 1 7 7 3 7 2]				
	[3 6 40 18 18 12 5 24 5]				

	[0 1 11 50 8 20 3 2 10]
	[2 3 19 12 9 28 8 5 4]
	[5 4 7 7 4 9 21 10 12]
	[2 12 13 6 9 10 11 34 4]
	[0 6 8 10 8 2 6 1 55]]

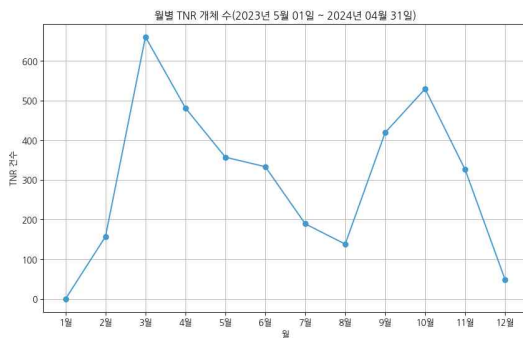
[길고양이 개체 수 관련 경향 분석]

1. TNR 대상 개체 몸무게(kg) :



TNR 대상은 몸무게 2kg 이상의 건강하고 임신하지 않은 개체지만 실제로는 2kg에 미치지 못하더라도 TNR을 시행하고 있음을 알 수 있다.

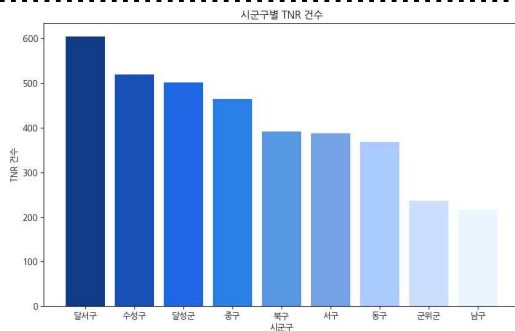
2. 월별 TNR 개체 수 비교(월별 TNR 개체 수(2023년 5월 01일 ~ 2024년 04월 31일))



동절기 (12월, 1월, 2월) 및 장마철과 하절기 (7월, 8월)의 경우 수술 후 회복의 어려움을 고려하여 수술을 지양하고 있어 그 TNR 건수가 적은 것을 알 수 있다.

대부분 2월 중에 신청을 받아 3월부터 TNR을 시행하기 때문에 3월 TNR 건수가 가장 많다. 또한 하절기가 지난 이후 9월부터 10월까지 TNR 건수가 늘다가 TNR 사업이 종료되는 12월까지 꾸준히 감소하고 있다.

3. 각 지역(시군구별) TNR 개체 수 비교



TNR을 가장 적극적으로 실시하고 있는 시군구는 '달서구'이다. 달서구는 지역 기초자치단체 최초로 '길고양이 급식소'를 시범 설치해 운영하고 있다. '대구신문'의 2023년 01월 09일 기사 '달서구 '길고양이 급식소' 확대 운영'에 따르면 '2019년 달서구청의 길고양이 급식소 설치 사업이 시작된 이후 2022년까지 도원·경원고, 월성공원, 도원근린공원 등 지역 12곳에 총 14개소의 급식소가 설치했고 급식소에 상주하는 길고양이 130여 마리 중 절반 정도인 60여 마리에 대한 TNR사업을 완료하였다. 반면 TNR을 가장 소극적으로 실시하고 있는 시군구는 '남구'이다.

4. 시군구별 TNR 건수와 인구 수 비교

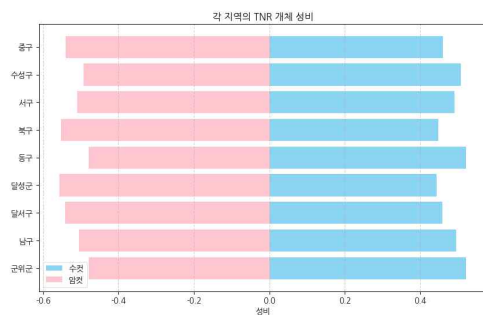


TNR 건수와 각 지역의 인구 수 간의 상관계수는 0.639664이다.

달서구는 대구 내에서 TNR 건수와 인구 수가 모두 가장 높은 시군구이다.

달서구, 수성구, 남구의 경우 TNR건수와 인구 수가 비례하지만 대부분의 시군구에서 그 관계가 비례하고 있지 않다. 따라서 TNR 건수와 인구 수는 큰 상관관계를 갖지 않는다고 판단할 수 있다.

6. 각 지역별 TNR 개체 수 성비 비교



TNR 개체의 성비는 큰 차이를 보이지 않는다.

7. 한계점

- TNR은 75%이상 완료되어야 효과가 있지만 길고양이 특성 상 각 시군구에 있는 개체 수 추측에 어려움이 있어 TNR의 효과를 예측하기 어렵다.

8. 결론 :

- TNR 대상이 되는 개체는 2kg 후반대부터 3~4kg대의 개체가 가장 많다.
 - TNR은 3월에 가장 많이 이루어지고 있으며 수술 후 회복의 어려움을 고려하여 TNR 지양하는 장마철과 하절기 이후인 10월이 그 뒤를 잇고 있다.
 - 대구 내에서 TNR을 가장 적극적으로 실시하고 있는 시군구는 '달서구'이며, 가장 소극적으로 실시하고 있는 시군구는 '남구'이다.
 - TNR 건수와 인구 수는 큰 상관 관계를 갖지 않는다.
 - TNR 개체의 성비는 큰 차이를 보이지 않는다.
-

3. 추진내용

① 리빙랩 지원 성과 상세내용

[잘한 점]

사회적으로 드물지 않게 발생하는 문제에 대해 근본적인 해결책을 고민하고 지역사회 차원에서부터 개인적 차원까지 어떠한 노력을 기울일 수 있을지 구체적으로 고민하여 심도 있는 연구를 할 수 있었다.

[부족한 점]

가설을 바탕으로 세워진 프로젝트이니만큼 한계점이 매우 명확하다고 생각한다. 실질적 길고양이의 개체수를 알 수 없다는 점과, 지자체에서 시행하는 사업은 예산 내에서만 집행된다는 점을 고려하지 못했다.

[달라진 점 (보람된 점과 알고 싶은 점)]

배웠던 기술들을 사용하여 보다 직관적으로 문제를 바라보고 타인을 설득할 수 있을 만큼의 자료를 제작할 수 있게 되어 보람되었다.

'공생하는 법'에 대한 사회 전반적으로 꾸준한 토의가 필요하다 느꼈다. 눈에 거슬리고 불편하다고 치워버리고 화내기만 할 생각보다는 어떻게 하면 불편을 줄여나갈 수 있을 것인지 고민하는 자세가 사회 전반적으로 필요하다고 느꼈다.

[결과물 구체적 기술]

TNR 사업 결과물에 대한 통계학적 분석을 실시했더니 다음과 같았다.

- TNR 대상이 되는 개체는 2kg 후반대부터 3~4kg대의 개체가 가장 많다.
- TNR은 3월에 가장 많이 이루어지고 있으며 수술 후 회복의 어려움을 고려하여 TNR 지양하는 장마철과 하절기 이후인 10월이 그 뒤를 잇고 있다.
- 대구 내에서 TNR을 가장 적극적으로 실시하고 있는 시군구는 '달서구'이며, 가장 소극적으로 실시하고 있는 시군구는 '남구'이다.
- TNR 건수와 인구 수는 큰 상관 관계를 갖지 않는다.
- TNR 개체의 성비는 큰 차이를 보이지 않는다.

머신러닝 모델의 성능 평가 지표 중 하나인 정확도(Accuracy)가 0.3으로 나온 경우, 이는 실제 서비스를 위해서는 다소 낮은 수준의 성능이라고 판단할 수 있다. 정확도 0.3은 모델이 예측한 결과 중 30%만이 실제 정답과 일치한다는 의미이기 때문이다. 일반적으로 실제 서비스에 적용되기 위해서는 더 높은 수준의 정확도가 요구된다.

따라서 이를 보완하기 위한 다음과 같은 노력이 필요할 것으로 보인다:

- 데이터 전처리 및 augmentation: 학습에 사용된 데이터의 품질과 양을 개선하여 모델의 일반화 능력을 높일 수 있다.
- 하이퍼파라미터 튜닝: 모델의 하이퍼파라미터(학습률, 규제 강도 등)를 최적화하여 성능을 개선할 수 있다.
- 모델 아키텍처 변경: 다른 종류의 모델 아키텍처를 적용하거나 앙상블 기법을 사용하여 성능 향상을 도모할 수 있다.
- 전이학습(Transfer Learning) 적용: 대규모 데이터로 사전 학습된 모델의 지식을 전이받아 활용하면 성능 개선에 도움이 될 수 있다.

또한, 이 모델을 실제 서비스에 적용하기 위해서는 다음과 같은 기술적 조건들을 고려해야 한다:

- 모델 최적화 및 경량화: 제한된 컴퓨팅 자원에서 효율적으로 동작하도록 모델을 최적화하고 경량화할 필요가 있다.
- 배치 및 운영 체계 구축: 모델을 배치하고 지속적으로 모니터링하며 업데이트할 수 있는 운영 체계를 마련해야 한다.
- 데이터 파이프라인 구축: 새로운 데이터를 지속적으로 수집하고 전처리하여 모델 재학습에 활용할 수 있는 파이프라인이 필요하다.
- 모니터링 및 피드백 체계: 실제 서비스에서 발생하는 에러나 edge case를 모니터링하고 분석하여 모델 개선에 반영할 수 있는 체계가 필요하다.

이와 같이 모델 성능 개선과 실제 서비스 적용을 위한 기술적 고려 사항들을 종합적으로 검토하고 대응할 필요가 있다.

② 향후추진계획

본 활동은 TNR 사업의 효과성을 데이터 기반으로 평가하는 것을 목표로 두 가지 주요 분석을 진행하였으나 초기 결과에서 한계점이 발생하였다. 이에 각 분석 영역에서 내용을 보완하여 정확성과 유의미성을 높이하고자 한다.

먼저 지역별 길고양이 특성 분류 모델의 성능을 개선하고자 한다. 현재 모델의 정확도가 0.3에 머물러 실용성이 낮다는 점이 확인되었다. 첫 번째로 이 문제를 해결하기 위해 국가동물보호정보시스템을 통해 지속적으로 추가 데이터를 수집할 예정이며 대구 인근 도시나 유사한 도시 환경의 데이터를 포함시켜 데이터의 양과 다양성을 높이하고자 한다. 두 번째로 문제를 재정의하고자 한다. 현재 모델은 정확한 행정구역을 예측하려 하였으나 너무 세분화된 탓에 어려움을 겪었다. 이에 범주를 넓혀 분류를 시도하는 것으로 각 환경 유형에 따른 길고양이 특성 차이를 더 잘 반영하려 한다. 마지막으로 데이터에 계절이나 연도 정보를 추가하여 시계열 특성을 반영하고자 한다.

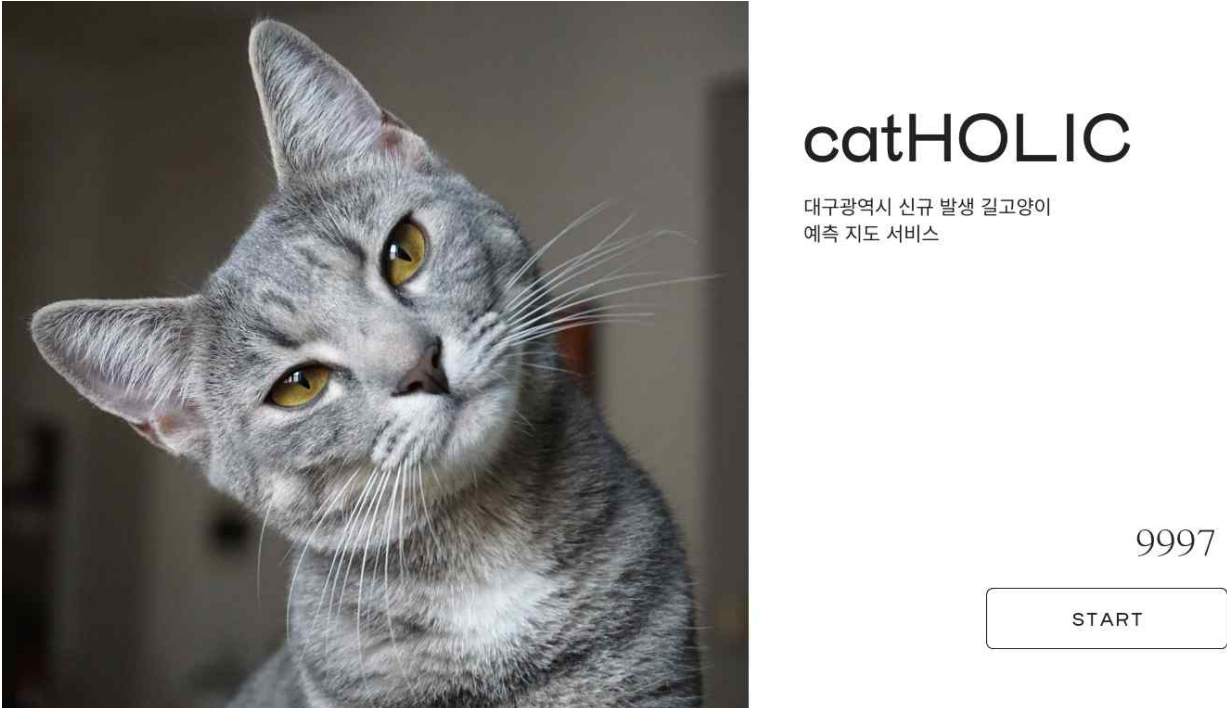
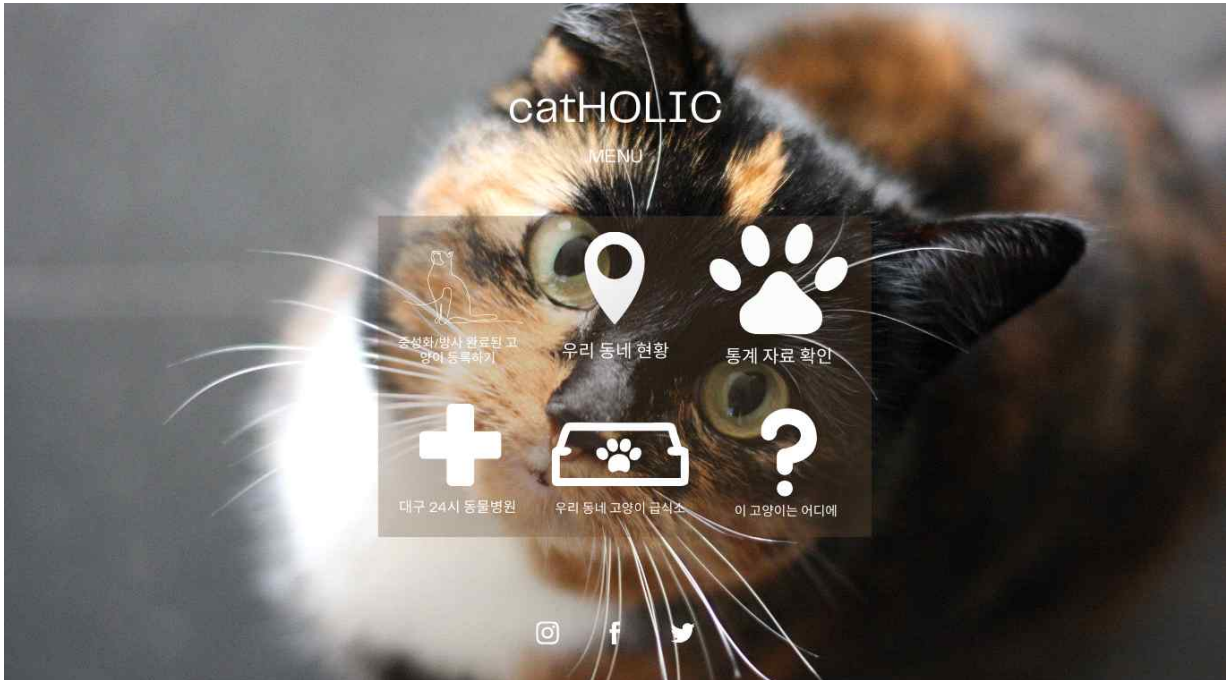
또한, 길고양이 개체 수 관련 분석을 강화하고자 한다. 현재까지의 분석은 TNR 사업 대상 개체 수를 단순 비교하는 수준에 가깝다. 이에 데이터를 보강하고 시계열 분석을 적용하며 외부 요인을 고려하여 분석을 진행하고자 한다. 또한 지역 간 비교 연구를 진행하여 TNR 사업이 활발한 지역과 그렇지 않은 지역을 비교 분석하고자 한다.

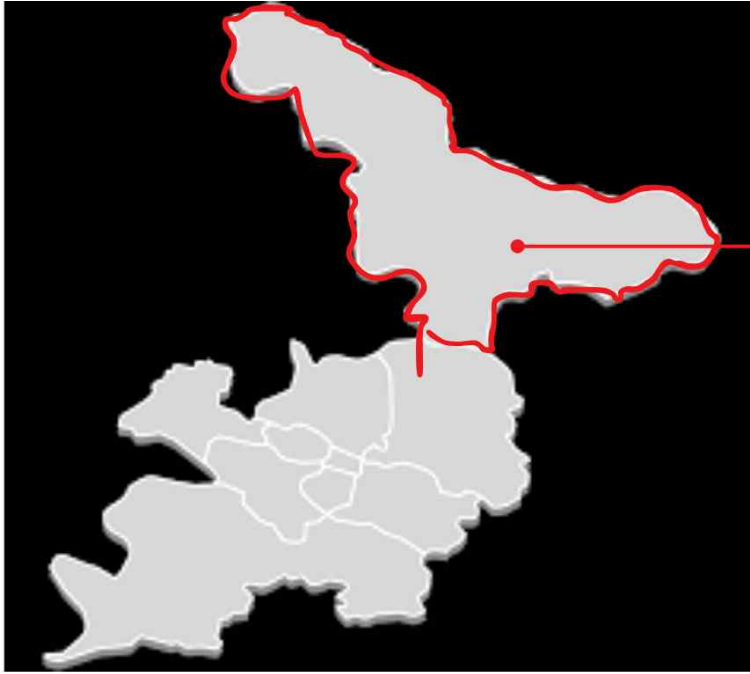
첨 부 4

결과보고서 증빙자료

1) 증빙자료

- 활동사진 및 결과물

1	웹페이지가 될 메인 페이지
	
2	각 기능으로 연결되는 메뉴 페이지
	



군위군

현재 중성화된 고양이 n개체
□ □ 특징을 가진 nn한 고양이
다수 이 곳에서 발견 예상

첨 부 4

소요예산 내역

○ 지출세부내역

구분	항목	사업비 산출내역	금액
지역주민 리빙랩 지원	회의비	1회차(29,500원) + 2회차(30,000원) + 3회차(25,500원) + 4회차(30,000원)	115,000원
	실습재료비	노션(Notion, 41,120원) 클로드(Claude, 90,600원)	172,590원
	소모품비		
	여비		
합계			287,590원