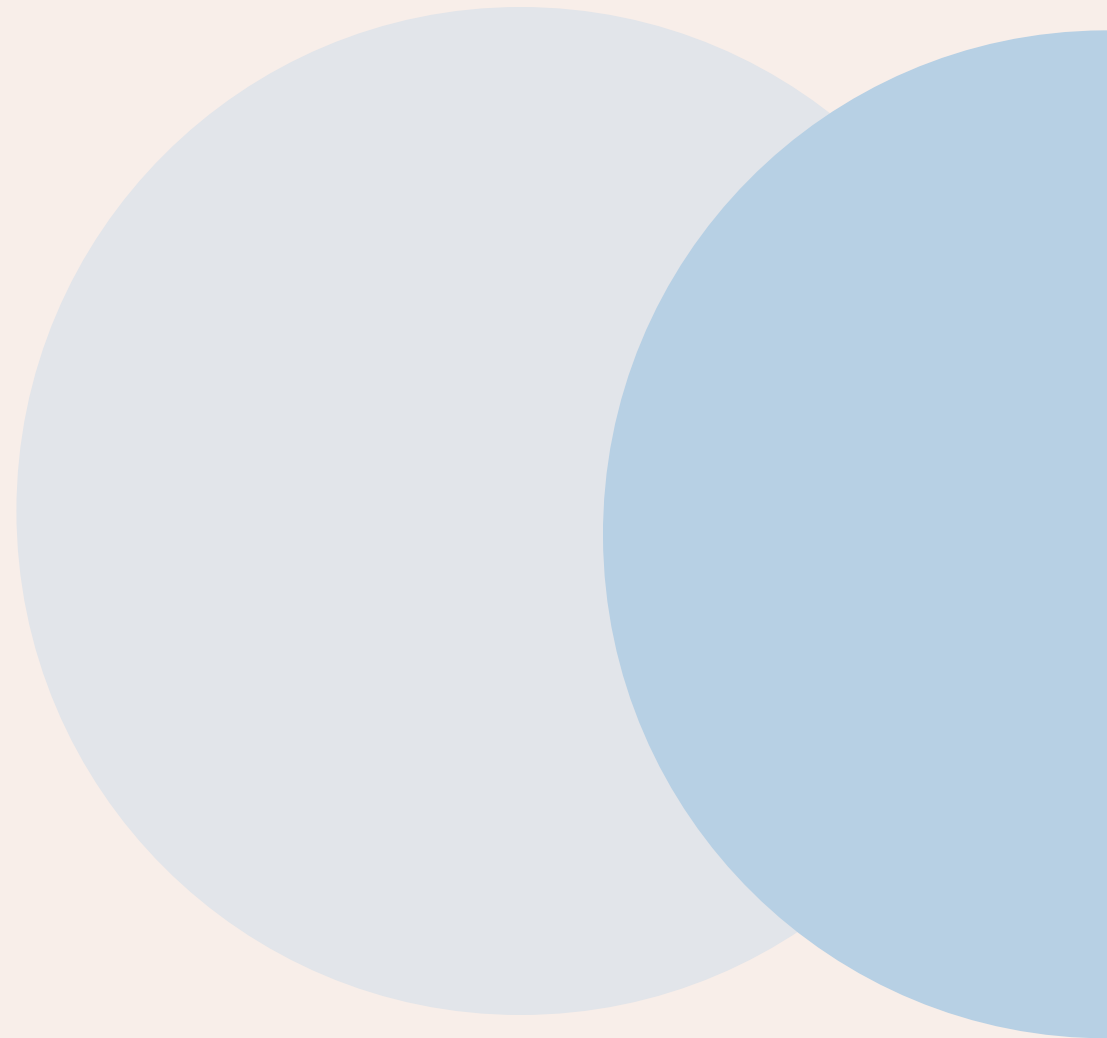


유기견 매칭 서비스 프로젝트 기획안

프로젝트 1조

Data Scientist	김경현 김태훈 이지훈
Data Engineer	손지수 이상엽 이서정



프로젝트 주제 및 목적

주제

입양 희망자의 성향과 니즈를 파악해 잘 맞는 유기견을 매칭 시켜주는 서비스

목적

- 유기견 파양률 감소
- 유기견 입양 문화 활성화
- 입양희망자가 유기견의 성향을 잘 파악할 수 있도록 정보 제공

프로젝트 수행 방향

사이언스

유기동물을 군집화 하여 사용자가 원하는 항목을 고르면 그에 가까운 유형의 유기동물을 매칭시켜 줌
로그인 시 작은 설문 항목을 만들어 이용자의 주거환경, 경제상태, 입양 시 중점적으로 보는 부분(견종크기, 털빠짐...)등을 조사할 수 있도록 함

엔지니어링

유기견 정보와 견종별 특성 데이터를 수집하고 구현될 웹 서비스의 로그 데이터를 수집하여 분석 모델에 활용할 수 있도록 데이터 파이프라인을 구축함

프로젝트 수행 도구

사이언스

- Jupyter Notebook
- PyCharm

엔지니어링

- VS Code
- Jupyter Notebook
- Docker

공통

- GIT : 병합 및 버전 관리, 기본 템플릿 공유
- Trello : 일정, 업무 진행 상황 파악, 기획서, 설계도, 회의록 정리
- Slack : 채널을 통한 팀 별 소통, 화면 공유 및 음성 대화, 개인 메시지 대화
- Data

농림축산식품부 동물보호관리시스템 유기동물 정보 API, 네이버 지식 백과 견종별 특성

Data - 유기동물 정보 API

```
<filename>http://www.animal.go.kr/files/shelter/2021/12/202201031501804_s.jpg</filename>
<happenDt> 20211231</happenDt>
<happenPlace>장흥면 유원지로 89 번길 90 인근</happenPlace>
<kindCd>[개] 말티즈</kindCd>
<colorCd>흰</colorCd>
<age>2020(년생)</age>
<weight>3.2(Kg)</weight>
<noticeNo>경기-양주-2022-00005</noticeNo>
<noticeSdt>20220104</noticeSdt>
<noticeEdt>20220114</noticeEdt>
<popfile>http://www.animal.go.kr/files/shelter/2021/12/202201031501804.jpg</popfile>
<processState>보호중</processState>
<sexCd>M</sexCd>
<neuterYn>Y</neuterYn>
<specialMark>부정교합. 코갈색. 킁킁댐/사람좋아함. 꼬리단미안됨.</specialMark>
<careNm>한국동물구조관리협회</careNm>
<careTel>031-867-9119</careTel>
<careAddr>경기도 양주시 남면 감악산로 63-37 (남면) </careAddr>
<orgNm>경기도 양주시</orgNm>
<chargeNm>양주시 동물보호팀</chargeNm>
```

```
<filename>http://www.animal.go.kr/files/shelter/2021/12/202201041101258_s.jpg</filename>
<happenDt> 20211231</happenDt>
<happenPlace>공주시 우금티터널</happenPlace>
<kindCd>[개] 믹스견</kindCd>
<colorCd>갈색</colorCd>
<age>2019(년생)</age>
<weight>5(Kg)</weight>
<noticeNo>충남-공주-2022-00001</noticeNo>
<noticeSdt>20220104</noticeSdt>
<noticeEdt>20220117</noticeEdt>
<popfile>http://www.animal.go.kr/files/shelter/2021/12/202201041101258.png</popfile>
<processState>보호중</processState>
<sexCd>M</sexCd>
<neuterYn>N</neuterYn>
<specialMark>파란색목줄착용, 겁이많음,진드기, 털영킴, 치석증</specialMark>
<careNm>이기영수의과병원</careNm>
<careTel>041-853-7575</careTel>
<careAddr>충청남도 공주시 감영길 7 (반죽동) </careAddr>
<orgNm>충청남도 공주시</orgNm>
<chargeNm>축산유통팀</chargeNm>
<officetel>041-840-8877</officetel>
```

Data - 네이버 지식 백과 견종별 특성 크롤링

원산지	이탈리아
체고	26cm이하
체중	2~3kg
크기	소형
외모	코는 점점 까매지고 순백의 비단같은 털이 온몸에 덮여 있음
성격	몹시 까다롭고 활발하며 질투와 응석이 있고 사람들과 친밀하게 지냄
추천성향	아파트/단독주택/전원주택, 어린이/젊은층/노년층, 운동량 보통
주요유의질병	심장판막증, 심장마비, 결막염, 각막염, 비만
색상	흰색
그룹구분	토이 (일반적으로 애완견을 의미)
친화성	높음
털빠짐	보통
집지키기	보통
실내외구분	실내

몰티즈

원산지	한국
체고	45~53cm
체중	15~20kg
크기	중형
외모	역삼각형의 머리와 곧게 선 귀가 특징
성격	수렵에 대한 강한 본능과 대담, 용감하며 신중하고 맹렬함
추천성향	단독주택/전원주택, 어린이/젊은층/노년층, 운동량 보통
색상	황색 흰색 검은색 브린들
그룹구분	한국개
친화성	낮음
털빠짐	보통
집지키기	뛰어남
실내외구분	실외

진돗개

서비스 시나리오

1. 회원가입

- 1) 사용자는 웹페이지를 통해 회원가입 버튼을 클릭한다.
- 2) 기본정보(이름, 아이디, 비밀번호 등)을 입력한다.

2. 로그인

- 1) 사용자는 웹페이지를 통해 로그인 버튼을 클릭한다.

서비스 시나리오

3. 사용자가 원하는 유기동물의 특성을 선택하는 설문조사1 화면이 나온다 [필수]

1) 설문조사1 (선호하는 견종크기, 털빠짐, 친화성 등)을 작성한다.

4. 사용자의 추가정보를 기입하는 설문조사 2 화면이 나온다 [선택]

1) 설문조사 2(주거환경, 경제상태, 여가시간 등)을 작성한다.

2) '다음에 할래요 ' 버튼을 눌러 스킵할 수 있다.

서비스 시나리오

5. 입양하고 싶은 유기동물 탐색

- 1) 홈화면에 사용자의 설문조사 1의 조건에 부합하는 유기동물 리스트가 보여진다. (조건문으로 분류)
- 2) 필터를 통해 지역/ 축종을 선택할 수 있다.
- 3) 리스트에서 사용자가 관심있는 유기동물을 클릭하면 해당 유기동물의 상세 페이지로 들어간다.
- 4) 유기동물의 상세 내용을 확인할 수 있고 하단에 같은 군집의 다른 유기동물을 보여준다. (이런 아이들도 있어요)
- 5) 해당 유기동물을 관심동물로 지정할 수 있다.

서비스 시나리오

6. 추후 발전된 유기동물 매칭 서비스

(서비스 이용자 데이터가 충분히 축적되었을 경우)

- 설문조사2(주거환경, 경제상태, 여가시간 등)와 사용자 로그데이터(어떤 동물들의 상세페이지를 봤었는지, 어떤 동물들을 관심동물로 지정했는지, 실제 입양했는지) 통해 입양 희망자 간의 유사도를 분석하여 비슷한 입양희망자가 관심동물로 지정하거나, 입양을 신청했던 유기동물들을 화면에 제시한다. (군집화 / 협업기반추천시스템)

필수 기능

사이언스

① 군집화

가장 군집화를 잘할 수 있는 알고리즘 선정

② 실루엣 분석

군집화 평가지표

각 군집 간의 거리가 얼마나 효율적으로 분리되어 있는지 분석

③ 추천시스템

추후 설문이나 로그 등을 통해 데이터가 확보되면 그것을 바탕으로
이용자에게 유기견을 추천

필수 기능

엔지니어링

- ① 유기견 정보, 견종별 특성 데이터 수집 - HDFS
유기동물 정보 API 활용
네이버 지식 백과 견종별 특성 크롤링
- ② Data Warehouse, Data Mart 설계 및 구축 - Oracle ATP
ERD Cloud 활용 DB 테이블 설계
Oracle ATP 계정에 DW/DM 구축

필수 기능

엔지니어링

③ 데이터 파이프라인 구축

각 ETL 과정을 클래스화 하여 app으로 구현

Airflow 배치관리로 파이프라인 자동화

④ 웹 서비스 개발

Rest API 서버 개발- Django Rest Framework 활용

⑤ 실시간 로그 데이터 수집

Kafka 활용

포함 기술

사이언스

- Python
- Scikit-learn

엔지니어링

- Python
- Django
- Spark
- Hadoop
- Oracle
- Airflow
- Kafka

감사합니다.