

# PORTFOLIO

김도연

연락처: 010-7506-1978

이메일: [kdy991201@naver.com](mailto:kdy991201@naver.com)

깃허브: [github.com/doyeon158](https://github.com/doyeon158)

# 목차

1.About Me

2. Summary

3. 주요프로젝트 상세

4. 수상 및 활동

5. 기술 역량

6. 자격증 및 교육과정

7. 깃허브 및 참고자료

# About Me



김도연 (KIM DoYeon)

kdy991201@naver.com  
github.com/doyeon158

## • Graduation

- 2018 아주대학교 산업공학과 입학
- 2024 아주대학교 산업공학과 졸업

## • Awarded

- 2023 산업 해커톤, 벨류체인 제조공정 최적화 - 우수상
- 2024 통계데이터 활용대회 - 우수상

## • Skills

- |                       |                        |    |
|-----------------------|------------------------|----|
| • Pandas, Numpy       | <div><div></div></div> | 80 |
| • scikit-learn        | <div><div></div></div> | 70 |
| • SQL                 | <div><div></div></div> | 80 |
| • matplotlib, seaborn | <div><div></div></div> | 70 |
| • Google Drive, Slack | <div><div></div></div> | 90 |

## • Project

- 2019 공급망 운영 시뮬레이션
- 2022 ML 모델 성능 최적화 프로젝트
- 2023 추천 시스템 DB 개발
- 2023 기상 데이터를 이용한 천문대 입지 솔루션

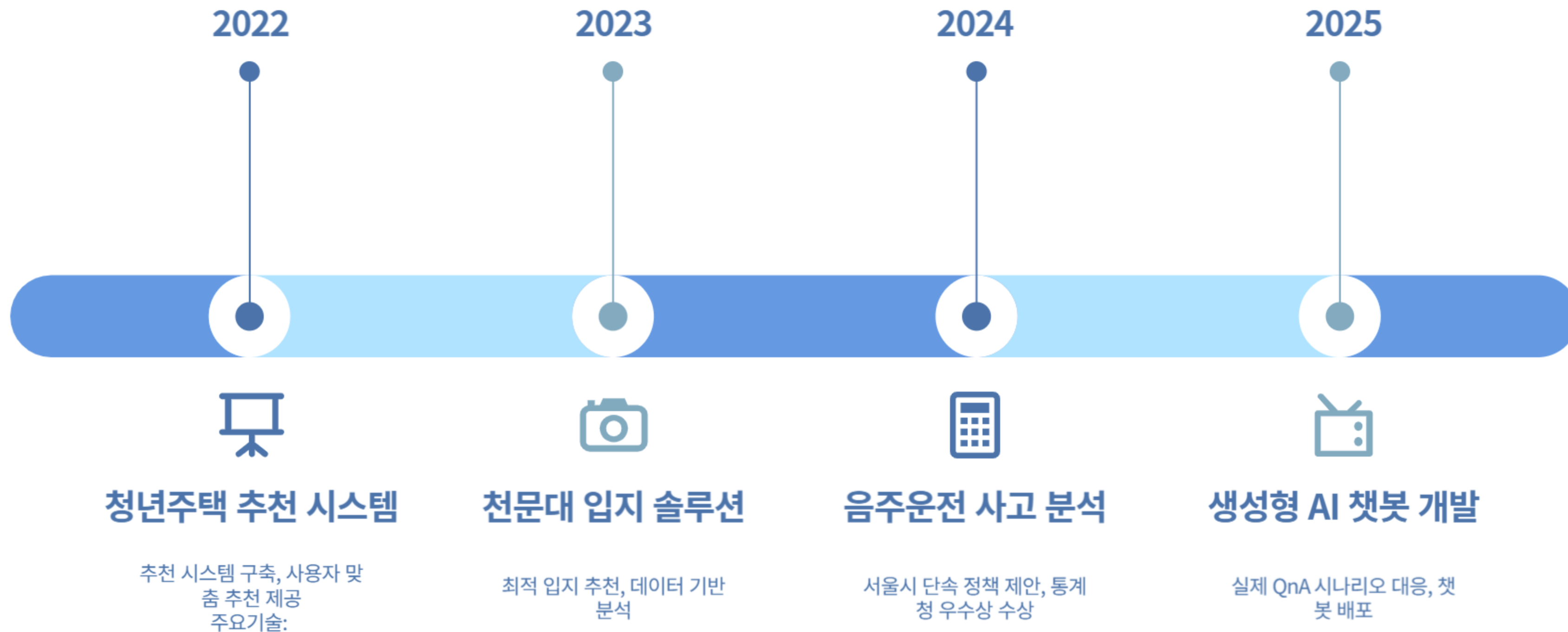
## • About

- 2021 온라인 스토어 창업 운영

## • Licence

- 2024 빅데이터 분석기사 (필기)
- 2024 데이터분석 준전문가
- 2023 SQL개발자
- 2023 6시그마 GB
- 2023 SNS마케팅전문가

## 2. Summary : 프로젝트 총 요약



# 천문대 입지 선정 최적화

2023.10 - 2023.12

## ● 문제인식

관련 논문에서 천문대 입지 선정 시 교통, 빛 공해, 기상 조건 등을 복합적 요인을 고려하여 최적 지역 선정해야하는 어려움을 인식

## ● 목표

기상, 교통 접근성, 빛 공해를 고려해 천문대 최적 입지를 추천함

## ● 프로세스

데이터 수집 및 전처리

: 빛 공해, 기상, 교통 데이터를 수집 후 결측값과 이상치를 보정. 데이터 정규화 수행

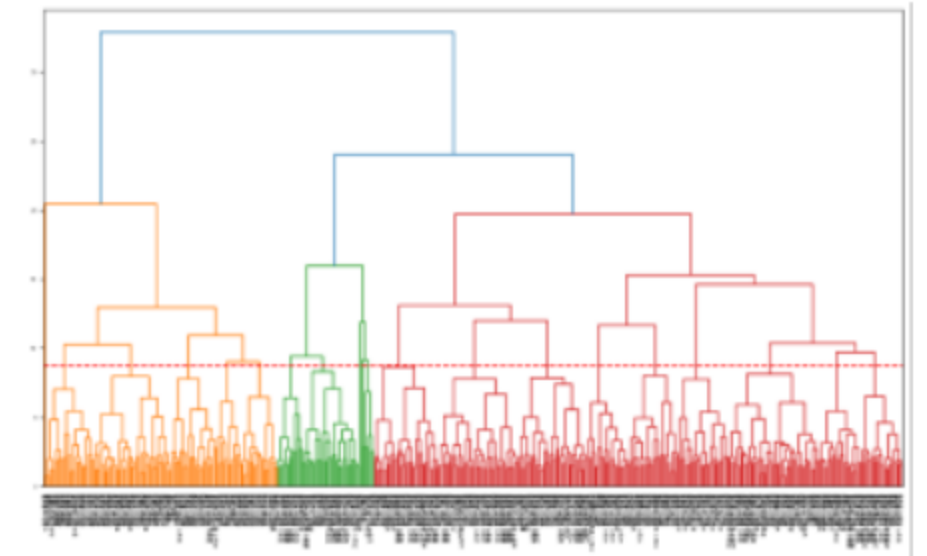
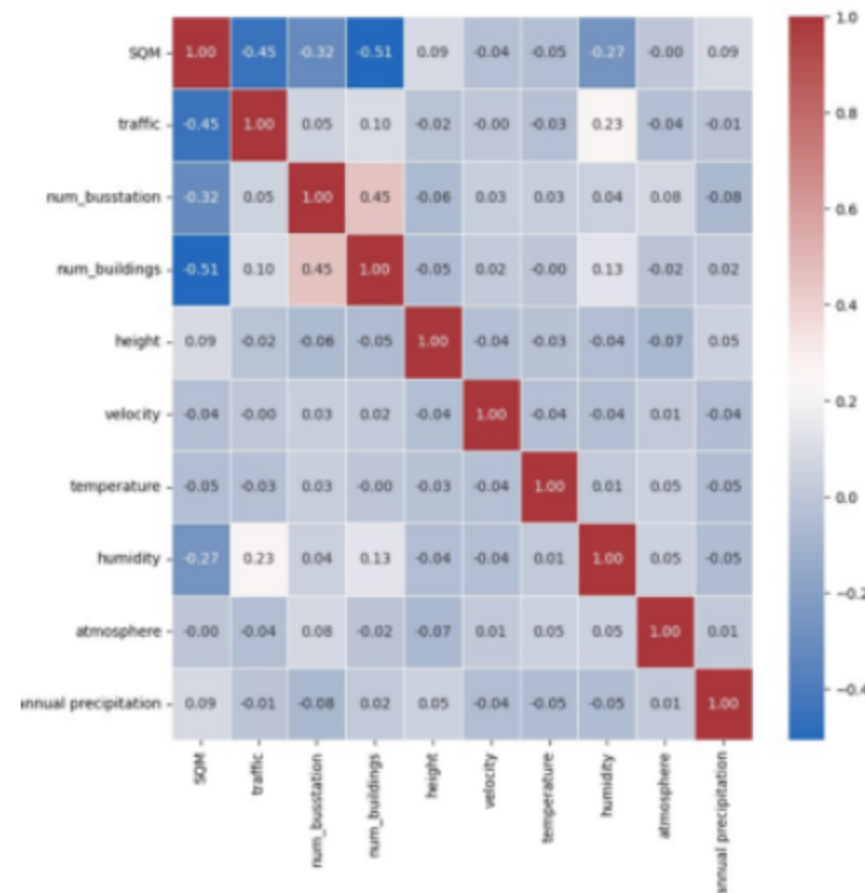
분석 모델 구축

: PCA로 주요 특징을 추출하고, 덴드로그램으로 분리되기 적절한 군집 수를 찾음

시각화 및 결과 분석

: 히트맵과 덴드로그램을 통해 지역별 특성을 시각화

우선순위 방식으로 천문대 설치에 적합한 지역 선정



Portfolio

## ● 결과

교통 편의성과 빛 공해가 낮은 최적 지역 10곳을 선정하여 입지 후보지를 제안  
기상 조건이 유리한 지역을 우선 배치하여 후보지를 선정함  
실제 이 지역에 천문대가 설치되어 있음을 확인.

## ● 배운 점

복잡한 요인을 종합적으로 고려한 입지 선정 문제 해결 능력 배양  
교통 요인과 빛 공해 요인처럼 서로 상반되는 요인을 절충할 수 있는 지역 선정 가능

# 청년주택 추천 시스템 DB 설계

2023.05 - 2023.06

## 문제인식

철도 유희부지의 활용 대안으로, 교통 편의를 고려한 청년 주택을 사용자에게 추천하는 방안을 구체화

## 목표

청년의 선호에 맞는 맞춤형 주택 추천 시스템 DB를 구축하여 사용자 만족도를 높임

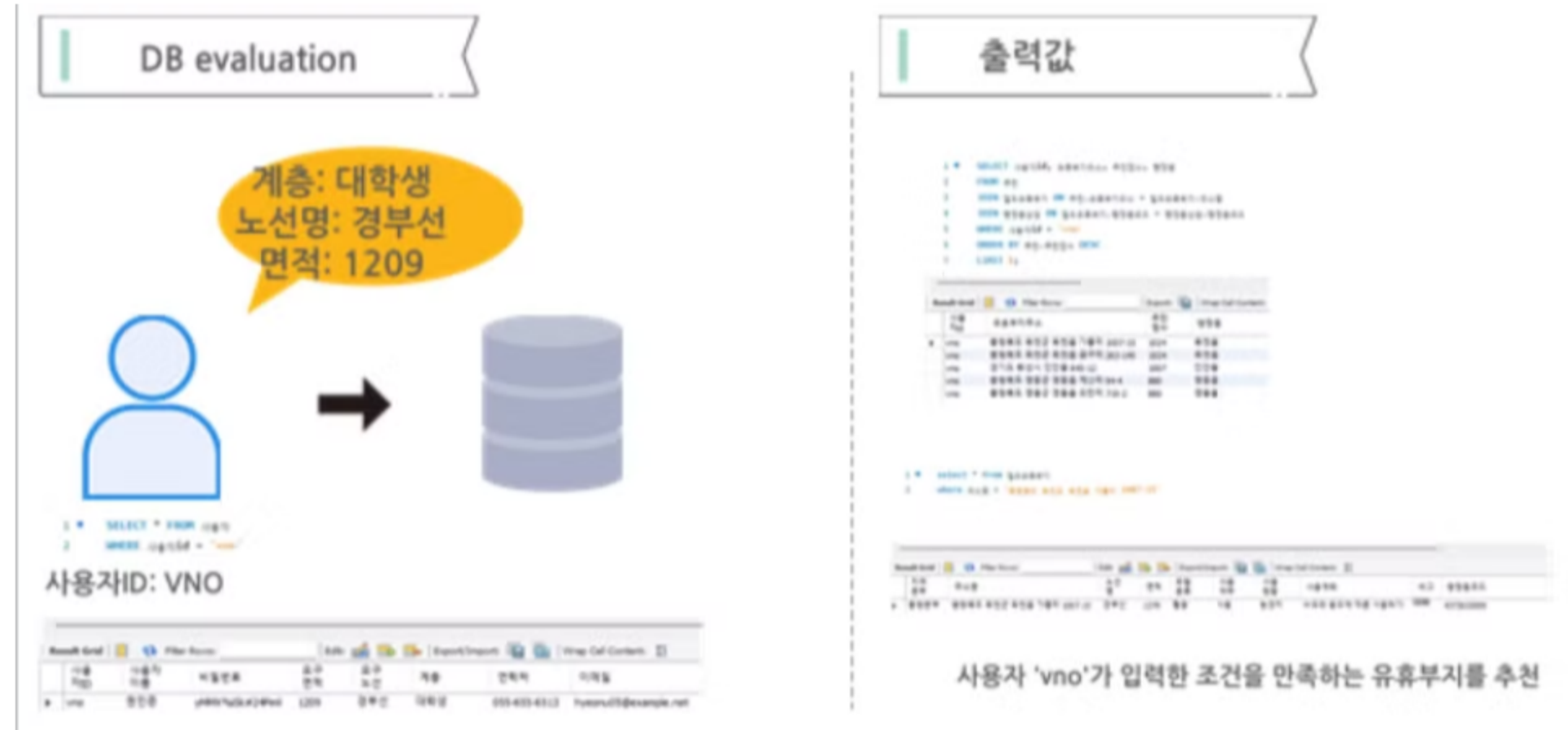
## 프로세스

요구사항 분석 및 ERD 설계:

PCA를 통해 주요 요소를 선정하고, 데이터베이스 구조를 정규화하여 이중 기본키 테이블로 정규화를 만족하는 테이블 설계

데이터 모델링 및 쿼리 구현:

청년주택 데이터와 사용자의 요구사항을 결합하여 추천 점수를 계산하고, 최적의 주택 5곳을 출력하도록 SQL 쿼리를 구현



## 결과

사용자 특성을 반영하여, 주택 추천 시스템을 통해 주거 정보를 제공

## 배운 점

데이터 구조화와 요구 사항에 맞춘 SQL 활용 능력 향상

사용자에게 맞춤형 주거 정보 제공하는 추천 DB 구조 설계 가능



# 음주운전 발생 분석 및 맞춤형 단속 솔루션 제안

2024.06 - 2024.07

- **문제인식**     단속 인력과 예산의 한계로 효율적인 단속이 어려운 상황
- **목표**             음주운전 발생 가능성을 예측하여 효율적이고 효과적인 단속 정책을 제안
- **프로세스**


모델링 및 분석: k-means 군집 분석으로 지역별 음주운전 패턴을 분석.  
랜덤포레스트로 피쳐 중요도 산정 → 회귀모델을 통해 음주운전 예측  
결과: 예측 성능 ( $R^2 = 0.77$ , MAE = 9건) 달성  
시뮬레이션 및 시각화: Folium로 위경도 데이터를 기반으로 음주운전 다발 지역 시각화  
강동구를 예시로 토요일 20시 이후 단속 시, 50건 감소 및 25억원 사고처리비용 절감.

- **결과**             음주운전 발생 패턴을 분석하여, 빈발 지역별 음주단속 솔루션 제안
- **배운 점**             예측 모델링과 솔루션 도출로 데이터를 기반으로 한 정책 제안 및 의사결정 지원  
발생 특성을 활용하여 효율적으로 자원 배분, 단속 효과 극대화 가능

**(1) 각 자치구별 음주운전 집중 단속 시간대**  
군집 분석으로, 서울 25개의 각 자치구는 표와 같이 집중 단속을 해야하는 **요일과, 시간대, 월별**로 결과를 확인할 수 있음.

자치구	요일	월	시간	연령
강남구	토	4,5,9월	0-4시	30대
강동구	토	6,12월	20-24시	30-40대
강북구	토,일	7,9,10월	20-24시	20-30대
강서구	목,금,토	4,5,9월	20-24시	30대
:	:	:	:	:

**(2) 음주운전 빈발 지역 선정**  
예측모델을 이용해, 아래와 같이 **50건 이상 음주운전 건수가 높은 지역(총 92개)**을 선정함.



# 생성형 AI 기반 보험 챗봇 개발

2025.04 ~ 2025.06

## • 사용기술

백엔드 : Django 5.2, Gunicorn, Nginx, MySQL 8.0

프론트엔드 : HTML/CSS/JS (템플릿 기반)

데이터 처리 : Python, BeautifulSoup, Requests, Pandas

AI 모델 : Qwen3-8B, text-embedding-3-small (벡터DB), RAG 기반 검색 시스템

색 시스템

배포 및 운영 : Docker, FAISS

기타 : Google Drive 버전 관리, BLEU Score 평가

## • 프로젝트 개요

반려견의 행동 문제를 보호자가 쉽게 이해하고 개선할 수 있도록, 자연어 기반 상담 챗봇을 개발.

기존 텍스트 고정형 상담의 한계를 극복하고, 맞춤형 대화 및 정보 추천을 통해 반려견과 보호자의 관계 개선을 지원.

## • 주요 내용

✓ 데이터 수집 및 전처리

반려견 행동·훈련 관련 콘텐츠 약 1만 건 크롤링

광고·불필요 문구·특수문자 제거 및 문장 정제

300여 종 견종 데이터 및 전문가 QnA 기반 고품질 데이터셋 구축

✓ LLM 및 RAG 시스템 구현

## • 성과 및 특징

타사 대비 차별화된 행동 상담 특화 챗봇 개발

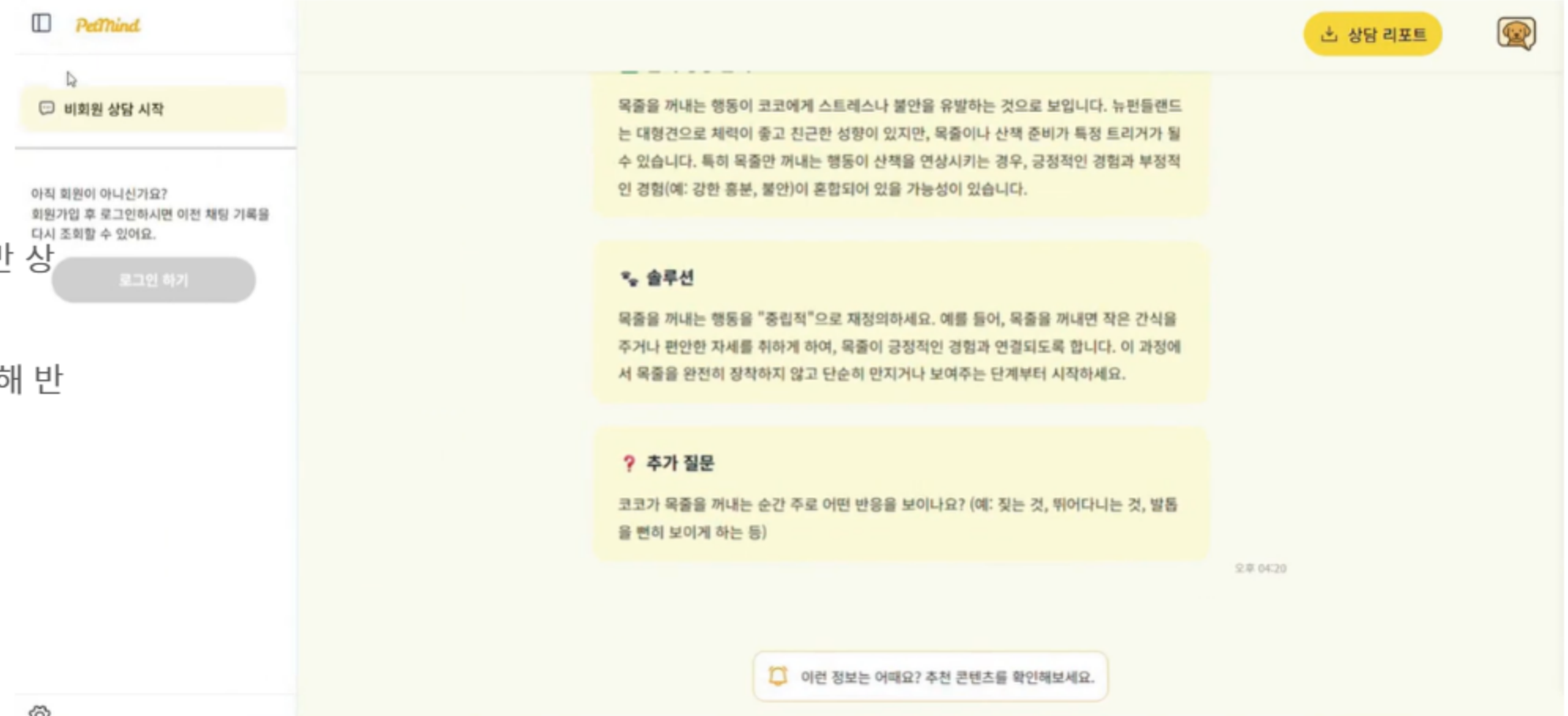
RAG 기반 개인화 콘텐츠 추천 → 상담→실행→정보 제공의 자연스러운 흐름 구축

데이터 기반 맞춤형 응답 및 후속 질문 설계로 대화 연속성 확보

프롬프트 고도화로 상담 유형별 응답 품질 개선

시스템 테스트 통해 90% 이상 핵심 기능 안정성 검증

초보 보호자도 반복 질문·훈련법을 쉽게 습득 가능





## 4. 수상 및 활동

- 2023 산업 해커톤, 벨류체인 제조공정 최적화 - 우수상
- 2024 통계데이터 활용대회 - 우수상

## 5. 기술 역량

- Python (pandas, numpy, matplotlib, seaborn)
- scikit-learn, SQL (MySQL)
- 시뮬레이션 (ARENA 등)
- 협업 툴: Google Drive, Slack

## 5. 자격증 및 교육과정

- 2024 빅데이터 분석기사 (필기 합격, 실기 준비 중)
- 2024 데이터분석 준전문가 (ADsP)
- 2023 SQL 개발자 (SQLD)
- 2023 6시그마 GB
- 2023 SNS 마케팅 전문가
- 2025 SK네트웍스 AI 교육 수료
  - 데이터 분석 및 생성형 AI 모델 기반 프로젝트 수행