

PORTFOLIO

김도연

연락처: 010-7506-1978

이메일: kdy991201@naver.com

깃허브: github.com/doyeon158

자기소개



김도연
(KIM DoYeon)

● 학력

2024 아주대학교 산업공학과 졸업

● 활동

- 생성형 AI 챗봇 개발
- 공공 데이터 기반 정책 분석 프로젝트
- 제조 공정 시뮬레이션 최적화
- 데이터 기반 천문대 입지 솔루션
- 추천 시스템 DB 개발

● 자격증

- 빅데이터 분석기사 (필기)
- 데이터분석 준전문가 (ADsP)
- SQL개발자 (SQLD)

● 보유기술

- Python
- MySQL
- pandas, scikit-learn, Tensorflow, LangChain, FastAPI Django, Streamlit

● 수상 이력

2024 통계데이터 활용대회 우수상
2023 산업 해커톤, 공정 최적화 우수상

● 최근 교육

SK네트웍스 AI·데이터분석 부트캠프
수료 (2024.12~2025.06)

생성형 AI 기반 반려견 상담 챗봇 개발

기간 : 2025.04 ~ 2025.06

4인 팀프로젝트

- 프로젝트 목표
 - 생성형 AI를 활용한 반려견 행동 교정 상담 챗봇 서비스 개발
 - 실사용 가능한 서비스 구현·테스트·배포까지 수행
- 사용기술
 - OpenAI GPT, Streamlit, Django, AWS, VectorDB, Python
- 주요 역할
 - ✓ 데이터 수집 및 전처리 : 반려견 훈련 콘텐츠 약 2천 건 크롤링 및 문장 전처리
 - ✓ 응답 테스트 및 품질 개선 : 300회 모델 응답 테스트 자동화 진행
 - ✓ 웹 플랫폼 개발 : Django 기반 반려견 프로필 관리, 피드백 기능 구현
 - ✓ Docker 환경 구성 및 배포



음주운전 발생 분석 및 맞춤형 단속 솔루션 제안

기간 : 2024.06 - 2024.07

3인 팀프로젝트

● 프로젝트
목표

- 서울시 자치구별 음주운전 사고 특성을 데이터로 분석
- 군집분석과 상관분석을 통해 지역별 위험 요인 도출
- 주요 요인을 기반으로 음주운전 발생 건수 예측 모델 개발

● 사용기술

Python, pandas, scikit-learn, k-means, matplotlib, seaborn

● 핵심 성과

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 솔루션 적용 시, 50건 감소 및 25억원 사고처리비용 절감


● 주요 역할

음주운전 요인 파악하여 지역 행정동별 세분화한 예측 모델 개발
단속 확률 계산시, 사고 감소량 및 처리비용 계산

(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.



천문대 입지 선정 최적화

기간 : 2023.10 - 2023.12

5인 팀프로젝트

- 프로젝트 목표

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

기상, 교통 접근성, 빛 공해를 고려해 천문대 최적 입지를 추천함

- 사용기술

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

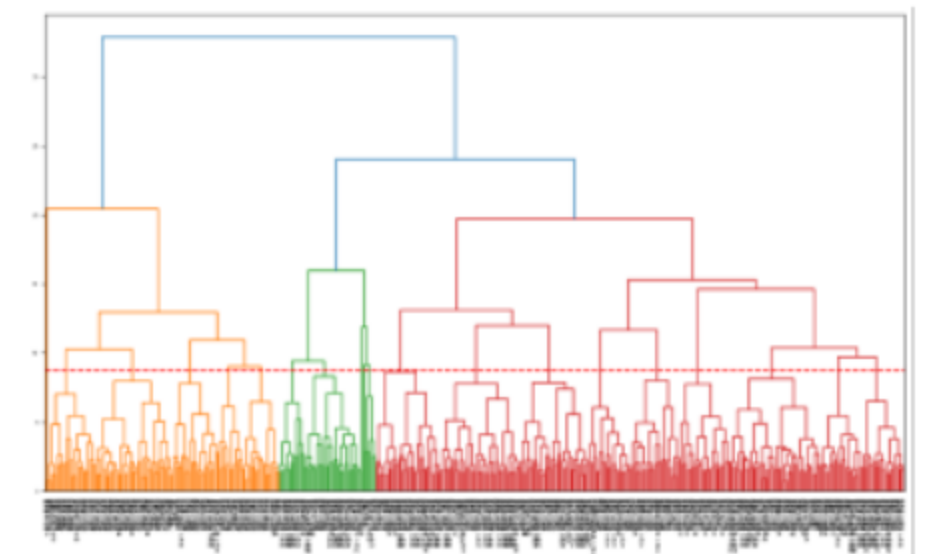
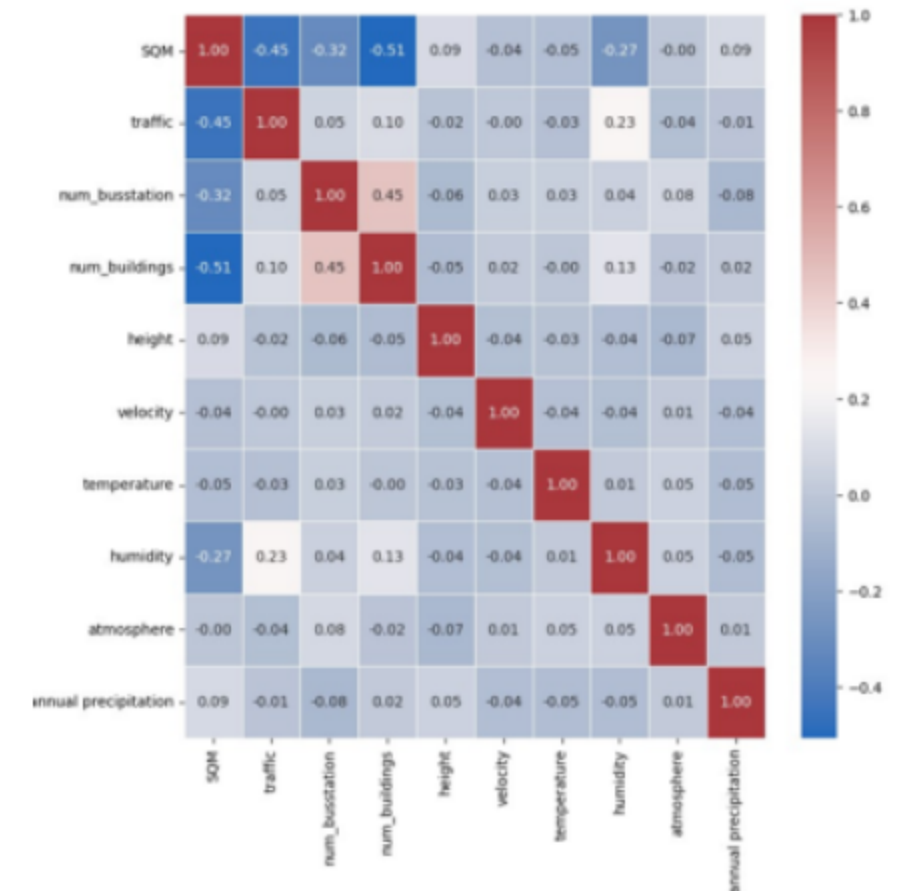
- 주요 역할

교통 편의성과 빛 공해가 낮은 최적 지역 10곳을 선정하여 입지 후보지를 제안

기상 조건이 유리한 지역을 우선 배치하여 후보지를 선정함

실제 이 지역에 2곳에 천문대가 설치되어 있음을 확인.

서로 상반되는 요인을 절충할 때, 우선순위를 통한 후보지 선정



청년주택 추천 시스템 DB 설계

기간 : 2023.05 - 2023.06

4인 팀프로젝트

- 프로젝트 목표

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

- 사용기술

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

- 핵심 성과

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

- 주요 역할

사용자 특성을 반영하여, 주택 추천 시스템을 통해 주거 정보를 제공
데이터 구조화와 요구 사항에 맞춘 SQL 활용 능력 향상
사용자에게 맞춤형 주거 정보 제공하는 추천 DB 구조

