

구성

- **files (폴더):** 설치 자료 및 외부데이터 포함
 - malgun.ttf (폰트 파일)
 - requirements.txt (패키지 requirements)
 - 어린이(만 13세 미만)보행자 교통사고.xls.xlsx ([TAAS](#) 제공 어린이 보행자 교통사고 데이터)
- **코드 파일 9개 : 분석 전과정 소스코드**
 - 1.Setups.ipynb
 - 2.Data Load& Preprocessing.ipynb
 - 3.EDA.ipynb
 - 4.Training set generation [Task1 , Task2].ipynb
 - 5-1.Modeling [Task1].ipynb
 - 5-2.Modeling [Task1].ipynb
 - 6.Scoring [Task1].ipynb
 - 6.Scoring [Task2].ipynb
 - 7.Result [Task1 , Task2].ipynb
- **코드 실행 설명서 : .md , .pdf**

소스 코드 실행 순서 및 동작 내용

실행 순서 :

1. 1.Setups.ipynb
2. 2.Data Load& Preprocessing.ipynb
3. 3.EDA.ipynb
4. 4.Training set generation [Task1 , Task2].ipynb
5. 5-1.Modeling [Task1].ipynb
6. 5-2.Modeling [Task1].ipynb
7. 6.Scoring [Task1].ipynb
8. 6.Scoring [Task2].ipynb
9. 7.Result [Task1 , Task2].ipynb

동작 내용:

1.Setups.ipynb

- 분석에 필요한 패키지 설치
- 시각화에 필요한 폰트 설치 및 적용

2.Data Load& Preprocessing.ipynb

- Compas 제공 Geoband API로부터 데이터 수집
- 데이터 별 세부 전처리 후 저장

3.EDA.ipynb

- 분석배경에 관한 EDA 및 시각화
- 오산시에 대한 전반적 EDA 및 시각화
- 과제별 세부 주제에 관한 EDA 및 시각화

4.Training set generation [Task1 , Task2].ipynb

- 예측 및 스코어링 모델링을 위해 과제별 Feature 칼럼 계산 및 저장

5-1.Modeling [Task1].ipynb

- 과제1 예측 모델링을 위한 머신러닝 모델 학습 및 성능 비교 (R)

5-2.Modeling [Task1].ipynb

- 과제1 예측 모델링을 위한 머신러닝 모델 학습 및 성능 비교 (Python)

6.Scoring [Task1].ipynb

- 모델링 결과에 따른 과제1 어린이 교통사고 위험지역 20곳 선정

6.Scoring [Task2].ipynb

- 스코어 모델링을 통해 교통안전시설물 설치 우선순위 20곳 선정

7.Result [Task1 , Task2].ipynb

- 선정 지역 세부 좌표 설정 및 시각화