|  |  |  |  |
| --- | --- | --- | --- |
| 딥러닝 기반 핵심 산업별 빅데이터 분석  <머신러닝&딥러닝 파일럿 프로젝트> | | | |
| 주 제 | Sanfrancisco Crime Prediction | **링 크** | https://www.kaggle.com/c/sf-crime |
| 팀 명 | 듀란듀란 | **일 자** | 2018년 11월 24일 |
| 팀 장 | 민유진<ujmin0417@naver.com>  함윤선<dotstar801@gmail.com> | **팀 원** | 유선우<ysw900524@gmail.com>  장정호<jangclod92@gmail.com> |

1. **과제 개요**

|  |  |
| --- | --- |
| **주제** | **샌프란시스코 범죄 예측** |
| **목표** | 12년 동안 샌프란시스코의 모든 지역에서 발생한 범죄정보를 기반  시간과 장소가 주어지면 발생한 범죄 범주를 예측 |
| **과제수행효과** | 1. 실전데이터를 활용해, 데이터 전처리 및 데이터 분석 기법을 복습 2. 주어진 데이터에 TensorFlow를 활용해, 머신러닝 실전 연습 |
| **왜 이 주제인가?** | 1. 한국 사회 범죄 증가로 인해, 범죄우발지역 예방강화 필요성 증가 2. 본 과제를 통해, 추후에 한국 도시의 범죄 우발 지역 예측으로 활용하기 위함 3. TensorFlow 실제 활용 방법 및 경험 습득 4. 데이터 시각화에 용이한 데이터셋을 활용하여 데이터 시각화 복습 |

1. **데이터 설명**

|  |  |
| --- | --- |
| **주어진 데이터 파일** | |
| Train.csv | 미국 샌프란시스코 시의 범죄 학습 데이터 |
| Test.csv | 샌프란시스코 범죄 테스트 데이터 |
| **데이터 필드 상세 사항** | |
| Dates | timestamp of the crime incident |
| Category | category of the crime incident (only in train.csv).  This is the target variable you are going to predict. |
| Descript | detailed description of the crime incident (only in train.csv) |
| DayOfWeek | the day of the week |
| PdDistrict | name of the Police Department District |
| Resolution | how the crime incident was resolved (only in train.csv) |
| Address | the approximate street address of the crime incident |
| X | Longitude |
| Y | Latitude |

**3. 과제 수행 내역 4. 결과 보고**