

서울시립대학교 통계학과 데이터 분석 경진대회

Yoon-Yeong Kim, Seong-ho Lee

Dept. of Statistics, UOS

대회 개요

- 대회 주제
 - 대기오염 데이터 분석을 통한 대기오염지수 예측모형 개발 및 예측
- 대회 사이트
 - Kaggle 사이트를 활용하여 데이터 다운로드 및 예측결과물 제출
 - <https://www.kaggle.com/t/026a38ece87747fead956e6a9903b9c0>
- 시상 내역
 - 대상 1팀 100만원
 - 최우수상 1팀 60만원
 - 우수상 2팀 각 30만원
 - 장려상 2팀 각 15만원

데이터 설명 (1/2)

- 학습 데이터
 - train_data.csv
 - 2019년 1월 ~ 2022년 12월 데이터
 - 총 6개의 오염물질 측정치(no2, o3, co, so2, pm10, pm2.5) 및 측정소 위치
 - 대기 지수 등급 (오염 정도에 따라 1~4등급으로 분류)
- 테스트 데이터
 - test_data.csv
 - 2023년 1월 ~ 2023년 12월 데이터
 - 총 6개의 오염물질 측정치 및 측정소 위치
- 과제
 - 테스트 데이터(각 오염물질 측정치 및 측정소 위치)의 대기 지수 등급을 예측

데이터 설명 (2/2)

- 보조 데이터
 - 대기오염과 관련된 2019~2023년 사이에 수집된 데이터
 - 학습 및 테스트 데이터와 함께 모형 개발에 사용 가능

파일명	데이터 설명
aux_data_DailyPopulaiton.csv	일별 행정동 단위 서울 생활인구 정보
aux_data_DailyRain.csv	일별 자치구 단위 강수량 정보
aux_data_MonthlyBike.csv	월별 대여소 단위 자전거 이용정보
aux_data_MonthlyCar.csv	월별 등록소 단위 자동차 등록정보
aux_data_YearlyMountainFire.csv	연도별 자치구 단위 임야화재 정보
aux_data_StationLocation.csv	행정구역 공간정보

- 참고
 - 제공하는 데이터에 대한 자세한 정보는 Kaggle 사이트 내 Data 탭을 참고
 - 제공하는 데이터 외에도 자체적으로 데이터 수집 및 사용 가능

주요 일정

- 중간 점검
 - 기간: 9월 30일 ~ 10월 5일
 - 팀별 담당 조교와의 미팅을 통해 진행 상황 점검
- 발표 및 심사
 - 일시: 10월 29일 오후 5시
 - 장소: 추후 공지
 - 팀별로 15분 이내로 개발 모형에 대한 발표 진행
 - 평가 기준: 제출 결과물의 예측 정확도 및 모형의 적합성
- 시상
 - 날짜: 11월 1일
 - <통계인의 밤> 행사에서 시상 진행

참여 팀 및 담당 조교

팀명	팀원	담당 조교 (연락처)
교수님 우리 잘했조	오제건, 정현우, 서해훈	안종혁 (jh981017@naver.com)
HEI	황보겸, 정지혁	
Easy	박태환, 김은호	
StatUs	송화원, 김민아	이시화 (129_hwa@naver.com)
최우수상입니다	최연주, 오다빈, 정연우, 정희진	
생경명영	황신영, 성경운	
최강스과	김태현, 박준성, 김건화	
통계<3	구기현, 김진웅, 김진혁	한병찬 (hanbc12@naver.com)
CODE:707	최동준, 김선재, 선유진	
통계RahL	오성우, 황서연, 김예령	
btcon	김동은, 조윤기	

예측결과 및 발표자료 제출 안내

- 예측결과 제출
 - 제출 경로: Kaggle 사이트 내 submissions 탭 활용
 - 제출 마감: 10월 27일 밤 23시 59분까지
 - 참고 사항
 - 각 팀별 제출 담당 학생 1명이 제출
 - 1일 최대 제출 횟수는 2회로 제한
 - 최소 정확도 약 85% 이상을 달성해야 유효한 제출로 인정됨
- 발표자료 제출
 - 제출 경로: 팀별 담당 조교에게 제출
 - 제출 마감: 10월 27일 밤 23시 59분까지

문의 사항

- 문의 사항은 아래 연락처를 통해 연락
 - 김윤영 교수: yykim@uos.ac.kr
 - 이성호 교수: seongho@uos.ac.kr
 - 안종혁 조교: jh981017@naver.com
 - 이시화 조교: 129_hwa@naver.com
 - 한병찬 조교: hanbc12@naver.com