

Assignment 2

3종류의 데이터셋 중 자유 선택

Dataset 1 - feature (x) : 1열 ~ 7열, target (y) : 8열

<https://drive.google.com/open?id=1NPr1pDtwPb9QlzRXf1fcfMuTRGGSpfsA>

Dataset 2 - feature (x) : 1열 ~ 10열, target (y) : 11열

https://drive.google.com/open?id=1VTACet0_HsmWY_w4gOpjqckiW1_M1jlq

Dataset 3 - feature (x) : 1열 ~ 13열, target (y) : 14열

<https://drive.google.com/open?id=1W4E7ECM0ycU0xcm8K714uaLrtha3OwEf>

1. 선택한 데이터셋의 target 값을 보고 어떤 문제인지 파악 (regression, binary classification, multi-class classification 중 하나)
2. 데이터 로드 후 선택한 데이터셋에 맞춰 필요에 따라 data preprocessing
 - data scaling, training/ test set split, one-hot encoding
3. 수업시간에 배운 Learning algorithm 중 선택한 데이터셋에 맞춰 자유선택 (K-Nearest Neighbors 제외)
 - Linear Regression (Ridge/Lasso), Logistic Regression, Decision Tree, Random Forest, SVM, SVR, Neural Network
4. 선택한 알고리즘에 따라 Hyperparameter 조정하며 결과 비교
5. 평가지표는 선택한 데이터셋이 어떤 문제인지에 따라서 사용
 - MAE, RMSE (regression 문제), Accuracy (classification 문제)

A4 한페이지 이내로 진행한 내용에 대해 설명, 형식 자유, **분량 초과시 0점**
뒷장에 코드 첨부 (페이지 제한 없음)

제출 기한 : 10월 13일 23시 59분

파일명 : 학번_이름_과제2

이메일 제출(jdm486@hanmail.net), 이메일 제목 : [데이터과학] 학번_이름_과제2

기한 초과시 0점