Data Mining assignment #1

첨부된 DataMining\_HW1.ipynb 파일 내 지시사항대로 미완성된 코드를 작성하여 제출할 것.

11가지의 분류 모델(Decision Tree, Naïve Bayes, Random Forest, LightGBM, XGBoost, SVM, MLP, Bagging, Boosting, Voting, 자신만의 방법으로 구현한 모델)을 사용하여 주어진 데이터셋에 대한 classification accuracy를 구하고 이를 표로 작성하여 이미지 파일로 제출할 것.

자신만의 방법으로 구현한 분류 모델의 경우, 다양한 방법을 사용하여 (Feature Engineering, Ensemble, etc) 가장 높은 Accuracy를 달성한 코드를 제출할 것. 만약 본인이 구현한 모델의 출력이 random한 결과를 보이는 경우, 채점 시 불이익이 없도록 높은 Accuracy를 보이는 random state를 코드 내에 미리 지정할 것. 또한 본인이 구현한 모델은 채점 시 주어진 데이터셋에 대하여 K-Fold Cross Validation을 사용하여 산출한 정확도의 평균을 사용하여 평가할 예정임.

- Accuracy 결과 표 예시

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Decision Tree | Naïve Bayes | Random Forest | XGBoost | LightGBM | SVM | MLP | Bagging | Boosting | Voting | My Method |
| Acc | 0.7760416666666666 | 0.7135416666666666 | 0.734375 | 0.7291666666666666 | 0.7395833333333334 | 0.7135416666666666 | 0.7083333333333334 | 0.7239583333333334 | 0.75 | 0.7239583333333334 | 0.7708333333333334 |

Due: 4/15 23:59 (Late를 허용하지 않음)

제출 파일 :

ML\_HW1\_학번.ipynb : 첨부된 DataMining\_HW1.ipynb 숙제 파일

ML\_HW1\_학번\_table.(jpeg or png) : accuracy 결과 정리 표 캡쳐 이미지