|  |  |  |  |
| --- | --- | --- | --- |
| **팀명** | AI제조 팀 | **이름** | 남도형 |
| **일시** | 7/1 | **장소** | 온라인 |
| **프로젝트명** | 머신러닝을 활용한 스마트폰 불량품탐지 | | |
| **개별활동**  **내용** | . from sklearn.model\_selection import train\_test\_split  # 데이터 준비  X = data.iloc[:, 1:]  y = data['TARGET']  # 데이터 분할  X\_train\_val, X\_test, y\_train\_val, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=42)  X\_train, X\_val, y\_train, y\_val = train\_test\_split(X\_train\_val, y\_train\_val, test\_size=0.25, random\_state=42)  # 각 세트의 크기 확인  print("훈련 세트 크기:", X\_train.shape)  print("검증 세트 크기:", X\_val.shape)  print("테스트 세트 크기:", X\_test.shape)  위 코드를 통해 데이터를 훈련, 검증, 테스트 데이터로 분할했다.    나온 결과값이다.  훈련 + 검증 세트 중 검증 세트의 비율을 25%로 설정했다. 따라서 훈련 세트는 전체 데이터의 52.5%가 됐다.    StandardScaler를 통해 훈련 세트를 사용하여 스케일링 학습을 진행했고 훈련, 검증, 테스트 세트에 대해 동일한 스케일링을 적용했다.    MinMaxScaler을 통해서도 스케일링 학습을 진행했다. | | |
| **향후추진 계획** | StratifiedKFold, KFold를 통해 교차 검증을 진행할 예정이다. | | |