|  |  |
| --- | --- |
| 교육 제목 | ML 모델링 기초 |
| 교육 일시 | 10월 13일(수) |
| 교육 장소 | 강의실 |
| **교육 내용** | |
|  | * ML Modelling process   AI > ML > DL 의 범주로 포함되어 진다.   * 기계학습의 분류 * 지도 학습(학습데이터에(label) 포함) : 분류(KNN, Logistic Regression, SVM, Decision Tree.....),   회귀(Linear Regression, SVR, Neural Networks….)   * 비지도 학습(학습데이터에 레이블(label)없음) : 군집(KMeans,HCA...),   영상분석(PCA, Kernel PCA, LLE......), 연관학습(Apriori, Eclat,....)     * 준 지도 학습 : 지도학습 , 비지도학습의 조합으로 이루어짐 * 강화학습 : 에이전트가 상태의 보상을 최대화 하기 위해서 행동함 * Training set and test set   ML 모델의 성능평가를 위해서 자료를 분할  Training set : 모델의 알고리즘 learning, 모델에 사용될 feature들을 결정, 초매개변수 조절  Test set : 최종 선택된 모델의 성능평가(약 전체 자료수의 30%로 설정), 자료의 수가 적을 경우 생략 가능   * 일반적인 ML 예측과정   학습세트 : 머신러닝 모델을 학습할 때 사용  검증세트 : 하이퍼 파라메터 결정할 때 사용  테스트 세트 : 학습된 모델을 평가할 때 사용   * 편향(Bias)과 분산(variance)   치우침(Bias) : 치우침은 실제값(또는 평균) 과 예측치 간의 차이를 의미  과소적합(underfitting)은 치우침이 높은 모델이 발생하기 쉬움  분산(Variance) : 분산은 주어진 데이터에서 모델 예측의 변이로 정의  과적합(overfitting)은 분산이 높은 모델이 발생하기 쉬움.  편향과 분산은 항상 Treade-off 관계가 발생 |
| 전체 내용 |