|  |  |
| --- | --- |
| 교육 제목 | 머신러닝 |
| 교육 일시 | 2021년 10월 13일 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | \* 최소제곱추정량 linear model y'(xi) = θ^0 + θ^1xi  \* 단순회귀와 중회귀  \* 머신러닝이란 개요에 대하여  \*artificial intelligence >> machine learnig >> deep learnig 순으로 발전함  \* 의학 자료로 RSNA 이미지가 2017년 추가되었다  \* 기계학습의 분류: 학습 데이터의 레이블 포함, 학습데이터의 레이블없음  \* 레이블이 있는 데이터와 없는 데이터 조합으로 이루어짐 |
| 오후 | ML 전처리 : 1, Feature들에 대한 적절한 전처리 (pre-processing)  2,초매개변수 조절 (Tuning hyperparameters) 3,모델 성능 평가  ML 알고리즘 : 1,ML 알고리즘 종류 2,모수(매개변수) 학습 (Parameter Learning)  3, 파이썬을 이용한 실습  Trainig set and test set  Training set: 모델의 알고리즘 learnig, 모델에 사용될 feature 들을 결정, 초매개변수 조절 (약 전체 자료수의 70% 로 설정) (7:3) 비율로 조절  test set : 최종 선택된 모델의 성능 평가 (7:3) 비율로 조절  ML 예측 과정>>: 1 학습세트: 머신러닝 모델을 학습할 때 사용  2: 검증세트: 하이퍼 파라메터 결정할 때, 3: 테스트세트: 학습된 모델을 평가할때  \*치우침 (bias) - 분산( (variance) trade-off  \*치우침은 모델의 실제값 (또는 평균) 과 예측치 간의 차이를 의미.  \*과소적합 (underfitting)은 치우침이 높은 모델이 발생하기 쉬움.  \*분산 (Variance)은 주어진 데이터에서 모델 예측은 변이로 정의.  \*과적합 (overfitting)은 분산이 높은 모델이 발생하기 쉬움. |