|  |  |
| --- | --- |
| 교육 제목 | **머신러닝** |
| 교육 일시 | 2021-10-13 |
| 교육 장소 | 집 |
| **교육 내용** | |
| 오전 | 최소제곱추정량    손실함수      중회귀    다항회귀          파이썬으로 추정량 나타내기 |
| 오후 | 학습이란  지식 정보 🡪 정보처리  머신러닝 기계학습 분류 종류  학습데이터에 레이블 포함  분류 : classification -> kNN  회유 : Regression  학습데이터에 레이블 없음  Clustering :  Visualization & Dimensionality reduction :  Association Rule learning :  레이블이 있는 데이터/ 없는 데이터  Ml 전처리  Feature 들에 대한 적절한 전처리  초매개변수 조절  모델 성능평가  Ml 알고리즘  Ml 알고리즘 종류  모수(매개변수) 학습  파이썬 이용한 실습  모델링과정  Traininf set and test set  : Ml 모델의 성능평가를 위해서 자료를 분할    일반적인 ML 예측과정  학습세트 : 머신러닝 모델을 학습할때사용  검증세트 : 하이퍼 파라메터 결정할 때 사용  트레이닝세트 : 학습된 모델을 평가할 때 사용    ML모델의 치우침과 분산  치우침 : Bias 모델의 실제값(또는 평균) 과 예측치간의 차이를 의미   * 과소적합 : 치우침이 높은 모델이 발생 쉬움   분산 : variance 주어진 데이터에서 모델 예측은 변이로 정의   * 과적합 : 분산이 높은 모델이 발생하기 쉬움       Training set  Testing set  k-fold 교차검증  : k-fold CV 훈련 데이터를 동일한 크기의 k 그룹으로 무작위로 나누는 리샘플링 방법  추정치를 k 테스트 오류를 평균화 하여 계산    단계별로 차례대로 진행 위에서부터 |
|  |  |