```
1. Notebook의 사용법 입력과 결과의 확인
                                                                 2. Markdown 사용법
                                                                           1. 리스트 데이터 형의 특징 (numpy를 사용하지 않는) 2. 수치 연산이 불편
                                                                                          배열(array)을 활용
1. 편리한 수치연산 arr데이터형의 속성을 확인 dtype, astype, shape, ndim...
                                                                 3. Numpy
                                                                                         2. 색인 2. 팬시색인
                                                                          2. numpy의 사용
                                                                                               3. 부울색인
                                                                                         1. 원소별함수
2. 축별함수
                                           B-01.Python 필수 라이브러리
                                                                 4. matolotlib 1. 그래프 그리기 2. 산포도(scatter plot) 그리기
                                                                            1. xlsx 파일 읽기
                                                                  5. pandas 2. 컬럼이름 부여
                                                                          3. DataFrame의 속성정보 활용
                                                                                       1. 모델선택
                                                                                       2. 모델설정 (파라미터 설정)

      6. Machine Learning 맛보기
      3. 지도학습 수행

      4. 새로운 데이터로 분류 예측

                                                                                     5. 여기서 문제정 모델의 정확도를 알 수 없다 모델의 평가 곤란
                                                  1. 주요 라이브러리 import
                                                              1. 엑셀문서 읽기 iris
                                                  2. 데이터 준비 2. 데이터를 확인 2. 자료형의 변형 3. 유형(Class)확인
                                                                           4. 데이터 시각화
                                                                    1. 라이브러리 import
                                                                   2. 모델 선택
                                                  3. 모델 훈련
3. 기계확습 수행 (KNN) 4. 예측
                                                                    5. 성능평가
                                                                    주의사항 train/test 구분 없다 과적합
                                                                  ✔ 1. 훈련/시험데이터 구분
                                                  4. 기계학습 수행 2 2. 모델 훈련
3. 예측
                                                                 1. 엑셀문서 읽기 boston
                                                                 2. 데이터를 확인 1. 목표변수의 선택
                                                                 3. 시각화 Skip
                                                  5. 지도학습 (회귀) 4. 탐색적 데이터 분석 히스토그램
                                                                 5. 모델 선택 KNeighborsRegressor
                                                                 6. 모델 훈련
                                                                7. 성능평가 성능이 좋지 못하다

        1. 엑셀문서 일기
        cancer

        2. 데이터를 확인
        1. 목표변수의 선택

                                                                 3. 탐색적 데이터 분석
                                                                 4. 데이터 준비 환경 데이터
                                                  6. 지도학습 (분류)
                                                                 5. 모델 선택 KNeighborsClassifier
                                                                 6. 모델 훈련
                                                                           1. 파라미터의 선택
                                                                 7. 성능평가 2. 파라미터 조합의 시각화
                                                                          3 그리드 서치
                                                                            1. 데이터 사용 boston
                                                                           2. 데이터 준비
                                                  7. 지도학습 (회귀) 그리드 서치 적용 📗 3. 모델 선택 KNeighborsRegressor
                                                                          4. 그리드서치 실행
2. 결과 시각화
                                                  과소적함
과대적함의 개념을 익히기 p57 그래프 참고
                                                  9. 선형모델 💢 MC 뉴런
                                                  1. 선형모델 계속 💢 1. 선형모델 개요
                                                                 1. 엑셀문서 읽기 boston
                                                                2. 데이터 준비 종속, 독립변수 나누기
훈련/테스트 데이터 구분
                                                  2. 선형모델 (회귀) 3. 모델선택 LinearRegression
                                                               4.성능평가 score
                                                                ✔ 5. 포인트 과소적합의 문제를 해결해야 한다
                                                             1. 차원의 확대 PolynomialFeatures
                                                            2. 모델선택 LinearRegression
                                                             3. 모델훈련
                                                  3. 특성공학 4. 성능평가
                                                             과소적함 문제를 해결했으나
과적함효과가 있다
                                                            6. 개선포인트 선형모델은 튜닝의 여지가 없으므로
모델을 교체한다
                                                                                   1. Data boston
                                                                                  2. 모델선택 Ridge
                                                  4. 선형모델 + 표현력 조절이 가능한 모델 (회귀)
                                                                                 1. score를 본 뒤
4. 성능평가 2. 파라미터를 조절한다
                    데이터형
                                                                                            3. 파라미터 조절의 결과를 시각화
                     속성
                                                                        1. Data cancer
                    method

    1. 퍼셀트론
    2. 모델선택
    Perceptron

    3. 모델훈련
    fit

명목형 변수유형
                             PyML요약정리
Class
주요 용어
                                                                      4.성능평가 score
                                                                           1. Data cancer
과대적합, 과소적합
                                                                          2. 모델선택 Log
                                                           2. 로지스틱 회귀 3. 모델훈련 fit
                                                                          4.성능평가 score
                                                                        1. Data cancer
                                                                      2. 모델선택 LogisticRegression
                                                  6. 로지스틱 회귀 모델의 튜닝 3. 모델훈련 fit
                                                                     파라미터 조절
4. 성능평가
결과 시각화
                                                             특징 비선형적 데이터도 문제없다
                                                                           1. Data cancer
                                                                          2. 데이터 준비 종속변수, 독립변수 준비
트레인, 테스트 데이터 나누기
                                                 7. 결정트리 1. Decision Tree
                                                                           3. 모델선택 DecisionTreeClassifier
                                                                           4. 모델훈련 fit
                                                                          score

5. 성능평가 과소적합은 고려할 필요없고, 과대적합을 주의해야 함
                                                                                    깊이를 제한하여 과적함을 개선한다

        8. 파라미터를 적용한 결정트리
        1. 파라미터 적용
        DecisionTreeClassifier(max_depth = 4)

        2. 성능평가
        파라미터를 조절하면서 score를 검토

                                                  9. 결정트리의 특징 
2. 특성 중요도 분석
                                                  1. 부트스트립 개요
                                                                 1. 라이브러리 import
                                                                 2. Data cancer
                                                                 3. 데이터 준비 종속변수, 독립변수 준비
트레인, 테스트 데이터 나누기
                                                                비교함모델 DecisionTreeClassifier
4. 모델선택 《개선함모델 RandomForestClassifier
                                                  2. RandomForest
                                                                 5. 모델훈련 fit
                                                                 6.성능평가 score 파라미터의 조절

▼ 7. 장점 특성 중요도 확인

                                                                ✔ 8. 단점 외삽을 할 수 없다
                                                                      1. Data boston

    2. 데이터 준비

    트레인, 테스트 데이터 나누기

                                                  3. RandomForest의 단점 3. 모델선택 RandomForestRegressor
                                                                     4. 모델훈련
                                                                    5.성능평가 score 과대적합의 확인 파라미터의 조절

√ 7. 외삽 문제의 시각화 외삽 문제의 해결 선형모델의 적용

                                                                  1. Data boston

    4. 외상문제의 해결
    2. 모델선택
    Ridge

    3. 모델훈련
    fit

                                                                4. 성능평가 score
                                                  💢 5. 훈련모델을 파일로 내보내기
                                                           1. 라이브러리 import from sklearn ensemble import GradientBoostingClassifier
                                                          2. Data cancer 
분류문제
                                                  6. GBRT 3 데이터 준비 종속변수, 독립변수 준비 트레인, 테스트 데이터 나누기
                                                           4. 모델선택 GradientBoostingClassifier
                                                         5. 모델훈련 fit
                                                         6.성능평가 score
                                                              모델 최적 설정 찾기
                                                  7. 그리드 서치 2행 fit 1. 라이브러리 import From sklearn.model_selection import GridSearchCV
                                                             3. 그리드서치 실행결과 score

    8. 교차검증
    2. 교차검증 실행

    1. 라이브러리 import
    from sklearn.model_selection import cross_val_score

                                                  ✔ 9. 그리드서치와 교차검증의 결함 grid_search = GridSearchCV(estimator=gbrt, param_grid= set_range, cv = 10)
                                                           1. 라이브러리 import 부분 분류와 회귀 보다 보다 되었다. SVC, SVR # SVC 분류기, SVR 회귀
                                                          2. Data cancer
                                                  10. SVM 3. 모델선택 SVM
                                                         4. 모델훈련 fit
                                                         랜텀포리스트는 필요없었다
                                                                                      1. 라이브러리 import from sklearn preprocessing import MinMaxScaler # 0과 1사이에 값이 분포한다

    2. 데이터 준비
    스케일 적용

                                                                                       3. 모델훈련 fit
                                                                                      4.성능평가 score
                                                  11. 데이터 전처리와 스케일 조정
                                                                                         1. Data boston
                                                                                        2. 데이터 준비 스케일 적용
                                                                          2. 회귀데이터 활용 3. 그리드서치 + 교차검증
                                                                                         ↓ 4.모델훈련
                                                                                         5.성능평가
                                                  12. 실제 데이터
```