

TO DO

우리가 할 일!

- ▶ 당뇨 데이터를 분석하고 회귀모델을 만들어 보자!
- ▶ 데이터 로드 - 깃허브 해당 레포지토리 참고

문제 1

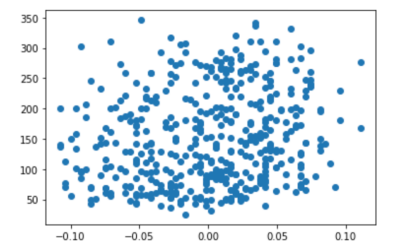
데이터의 크기를 구하시오. 출력 예) (50, 12) hint : 일단 특성은 총 10개 이다.

문제 2

10개의 특성에 대한 산점도를 그리시오.(for문을 이용할 것)

(matplotlib을 이용한다.)

출력 예)



문제 3

인덱스 2번, 8번, 9번 feature 에 대한 선형회귀 모델을 만드시오.(각각)

데이터를 train_input, train_target, test_input, test_target으로 쪼개서 진행할 것
sklearn을 이용!

산점도 위에 모델이 예측한 선을 그릴것

과적합/과소적합 여부를 확인할 것

모델의 score점수를 알아낼 것

coef_와 intercept_ 를 알아낼 것

문제 4

인덱스 6번 feature에 대한 다항 회귀모델을 만드시오. (2차 방정식)

데이터를 train_input, train_target, test_input, test_target으로 쪼개서 진행할 것
sklearn을 이용!

산점도 위에 모델이 예측한 선을 그릴것

과적합/과소적합여부를 판단할 것

모델의 score점수를 알아낼 것

coef_와 intercept_ 를 알아낼 것

문제 5

인덱스 2번, 8번, 9번 feature을 모두 이용하여 1개의 다중 회귀 모델을 만드시오.

데이터를 train_input, train_target, test_input, test_target으로 쪼개서 진행할 것
sklearn을 이용!

과적합/과소적합여부를 판단할 것

모델의 score점수를 알아낼 것

coef_와 intercept_ 를 알아낼 것

최종 마감 : 7월 28일

우리 팀 노션으로 주피터 노트북 file을 올려주세요!