

Homework 2.

<Python Study>

1. Pandas를 이용하여 'waterquality.csv'를 본인의 파이썬에 불러오시오.
2. F.Coli(독립변수)와 T.Coli(종속변수)의 선형 관계($y = w1 * X + W2 * 1$)를 분석할 것이다.
 - 2-1) 데이터 프레임에서 2000이후, F.Coli, T.Coli의 0이상인 데이터만 추출해보시오.
 - 2-2) Matplotlib과 Numpy를 이용하여 간단한 산점도를 통해 이 둘의 선형관계를 그림으로 나타내시오.
 - 2-3) 이 두 변수의 관계를 알아볼 것이다. Scikit-learn을 통해 Training data set과 Test data set을 7:3으로 나누어 보시오.
 - 2-4) Training data set을 통해 단순 선형 회귀를 진행해 보시오.
(목표함수는 RMSE를 사용하고 최적화 방법은 경사하강법을 이용한 코드를 사용하시오.)
 - 2-5) Test data set을 이용하여 단순 선형 회귀 식의 성능을 평가해 보시오
3. 위 데이터 프레임의 Chla의 농도를 예측하고 싶다. 이 때 Chla의 농도가 TP, DO에 영향을 받는다고 가정 했을 때, 위의 과정과 비슷하게 다중회귀($Y = W1 * X1 + W2 * X2 + W3$)를 실시해보시오.

Made by Theo Kim (2014890021)

Information and Knowledge should be shared by Everyone