

Tnt 2주차 과제

1. 드라이브에 제시된 MPG 파일을 이용하여 다음 문제를 해결하세요.

mpg 데이터는 여러 차량의 mpg(연비)와 이와 관련된 여러 자료들이 있는 데이터프레임입니다.

각 열이 나타내는 정보는 다음과 같습니다.

mpg : 차량의 연비

cylinders : 차량의 엔진 분류(실린더 수)

displacement : 배기량

horsepower : 마력

weight : 차량무게

accerleration : 차량 가속

model_year : 차량 모델 제작년도

origin : 차량 모델 제작 국가(회사가 속한 국가)

name : 차량 이름

1) 자동차 모델의 여러 정보들을 이용하여 차량의 mpg를 예측하는 모델을 만들 것입니다. mpg를 예측하는 최적의 모델을 만드세요.

2) 1번에서 만든 모델을 이용하여 mpg와 각각의 변수 간의 관계를 설명하세요.

가이드라인

① 위 데이터는 1주차 시각화 과제에서 사용한 MPG 데이터와 동일합니다.

② 1)에서 만드는 최적의 모델을 알 수 있는 지표는 Cp statistic, Adjusted R_squared, BIC, AIC 중 하나를 사용해주세요.

③ 2)를 해결하기 위해서는 Algorithmic model이 아닌 statistical modeling이 필요합니다. (random forest, boosting 등등 사용하면 2) 해결을 할 수 없음.)

2. 드라이브에 제시된 concrete_train, concrete_test 데이터를 이용하여 다음 문제를 해결하세요.

concrete_train, concrete_test 데이터는 콘크리트의 내구성과 관련된 여러 자료들이 있는 데이터 프레임입니다.

각 열이 나타내는 정보는 다음과 같습니다.

Cement (component 1) -- quantitative -- kg in a m3 mixture -- Input Variable
Blast Furnace Slag (component 2) -- quantitative -- kg in a m3 mixture -- Input Variable
Fly Ash (component 3) -- quantitative -- kg in a m3 mixture -- Input Variable
Water (component 4) -- quantitative -- kg in a m3 mixture -- Input Variable
Superplasticizer (component 5) -- quantitative -- kg in a m3 mixture -- Input Variable
Coarse Aggregate (component 6) -- quantitative -- kg in a m3 mixture -- Input Variable
Fine Aggregate (component 7) -- quantitative -- kg in a m3 mixture -- Input Variable
Age -- quantitative -- Day (1~365) -- Input Variable
Concrete compressive strength -- quantitative -- MPa -- Output Variable

1) 주어진 데이터를 이용하여 최적의 Concrete Compressive strength를 예측하는 모델을 만들 것입니다. Concrete compressive strength(MPa)를 예측하는 최적의 모델을 만드세요.

2) 1)에서 만든 모델을 이용하여 Mpa와 각각 변수간의 관계를 설명하세요.

3) 1)에서 만든 모델의 test_mse를 구하세요.

가이드라인

① concrete_test는 test_mse를 구할 때만 사용하세요. (학습시킬 때 concrete_test를 사용하면 안 되고, concrete_train만 사용해야 합니다.)

② 2)를 해결하기 위해서는 Algorithmic model이 아닌 statistical modeling이 필요합니다. (random forest, boosting 등등 사용하면 2) 해결을 할 수 없음.)