

2. Supervised Learning Challenge 1

Supervised Learning Challenge : Regression

- due date : **Sat. 10/5 11:59 pm**
- 이 과제는 2학기 프로젝트 조 편성 시 반영됩니다.

아래 링크를 많이 참고하였습니다.

<http://www.sthda.com/english/articles/37-model-selection-essentials-in-r/153-penalized-regression-essentials-ridge-lasso-elastic-net/> (<http://www.sthda.com/english/articles/37-model-selection-essentials-in-r/153-penalized-regression-essentials-ridge-lasso-elastic-net/>)

- 다음의 패키지들을 이용하세요. : MASS, glmnet, caret

#코드를 입력하세요.

- MASS 패키지를 이용하여 **보스턴 하우스 데이터**를 불러오세요.
- 7:3의 비율로 train data와 test data를 분리하세요.
- MEDV 변수를 y로 하고, 나머지 변수들을 x로 하여, x_train, y_train, x_test, y_test로 할당하세요.

#코드를 입력하세요.

```
set.seed(10)
```

Linear Regression

- training data에 대하여, MEDV를 y로 하고, 나머지 모든 변수들을 x로 하는 Linear Regression을 적합하세요.
- 적합한 결과를 보이고, plot으로 나타내세요.

#코드를 입력하세요.

Ridge Regression

- 다음의 코드를 실행시키세요. cross validation을 이용해서 구한 람다 값을 확인하세요.

```
cv.ridge<-cv.glmnet(x_train, y_train, alpha=0)
par(mfrow=c(1,1))
plot(cv.ridge)
cv.ridge$lambda.min
```

- 위에서 구한 lambda 값을 이용하여 ridge regression을 적합시키세요.
- test data에 대하여 RMSE, R2 값을 구하세요.

#코드를 입력하세요.

Lasso Regression

- 다음의 코드를 실행시키세요. cross validation을 이용해서 구한 람다 값을 확인하세요.

```
cv.lasso<-cv.glmnet(x_train, y_train, alpha=1)
plot(cv.lasso)
cv.lasso$lambda.min
```

- 위에서 구한 lambda 값을 이용하여 lasso regression을 적합시키세요.
- test data에 대하여 RMSE, R2 값을 구하세요.

#코드를 입력하세요.

수고하셨습니다.