

## فایل توضیحات پوشه 1 از ضمیمه Q2S4

دستورالعمل تکمیلی سوال : در حالتی که رگرسیون دو جمله ای باشد ( و روی آن Elasticnet یا رگرسیون ساده اعمال شود . ) و در حالت ساده تر (درجه یک ) آزمایش تکرار شود . مقدار Alpha برابر ۱ در نظر گرفته شود.

نحوه ساختن مجموعه داده : از مجموعه داده “Diabetes” یکی از ویژگی ها انتخاب شده است و سپس داده پرت به آن اضافه شده است و فایل های ضمیمه به دست آمده است.

```
from sklearn import datasets
diabetes = datasets.load_diabetes()
# Using one feature
diabetes_X = diabetes.data[:, np.newaxis, 2]

# Split the data into train and test sets
from sklearn.model_selection import train_test_split

dataTr_x, dataTe_x, dataTr_t, dataTe_t = train_test_split(diabetes_X, diabetes.target,
    test_size=0.1, random_state=42)
dataTr_t = dataTr_t.reshape(-1,1)
dataTe_t = dataTe_t.reshape(-1,1)

#add outlier Points
for j in range(10):
    dataTr_x=np.vstack((dataTr_x,np.array([0.3+ 0.02*np.random.randn()]).reshape(-1,1)))
    dataTr_x=np.vstack((dataTr_x,np.array([0.3+ 0.02*np.random.randn()]).reshape(-1,1)))
    dataTr_t=np.vstack((dataTr_t,np.array([600+ 10*np.random.randn()]).reshape(-1,1)))
    dataTr_t=np.vstack((dataTr_t,np.array([-600+ 10*np.random.randn()]).reshape(-1,1)))

from numpy import savetxt

savetxt('X_train.csv', dataTr_x, delimiter=',')
savetxt('X_test.csv', dataTe_x, delimiter=',')
savetxt('Y_train.csv', dataTr_t, delimiter=',')
savetxt('Y_test.csv', dataTe_t, delimiter=',')
```

مثالی از نقاط آموزشی (سیاه) و آزمایشی (نارنجی) در کنار خط رگرسیونی آموزش دیده (سبز) در یکی از مراحل :

