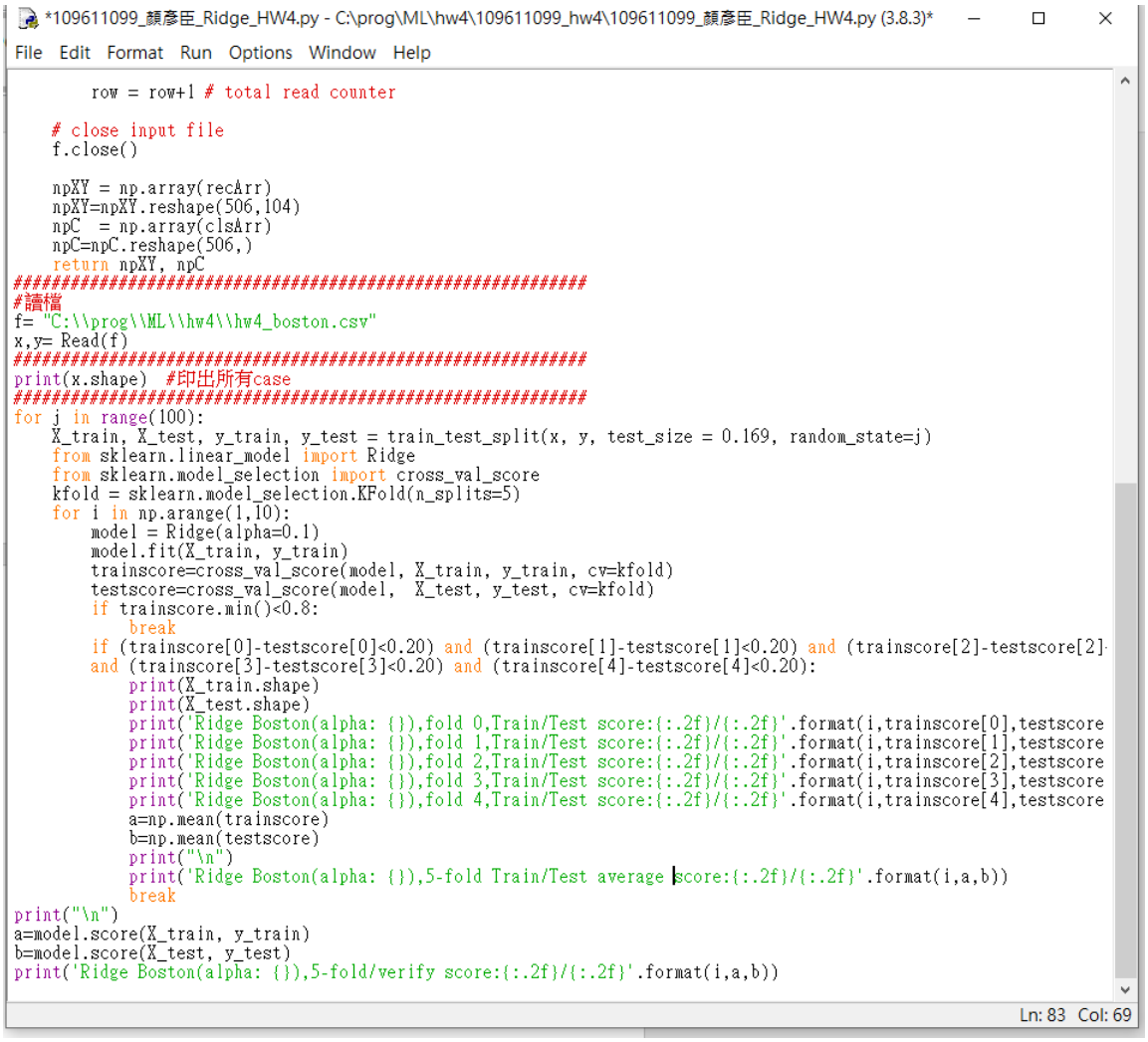
Code:

109611099_顏彥臣_Ridge_HW4.py - C:\prog\ML\hw4\109611099_hw4\109611099_顏彥臣_Ridge_HW4.py (3.8.3)

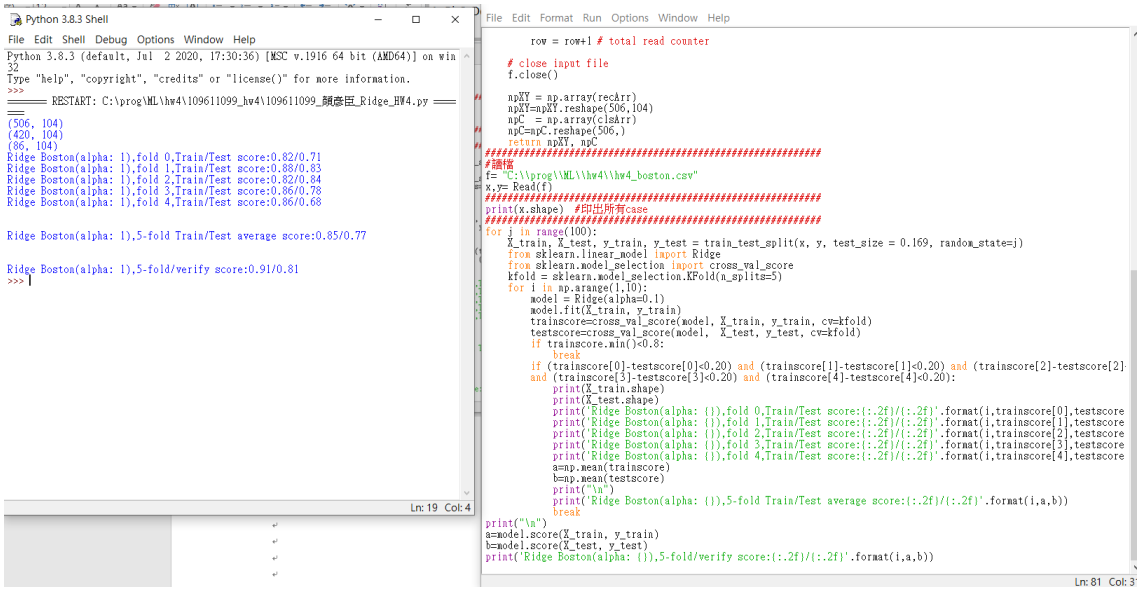
File Edit Format Run Options Window Help

```
import numpy as np
def Read(indata):
    #init
    recArr = []
    clsArr = []

    #read file
    f = open(indata, 'r')
    row = 0
    while (row < 506):
        s = f.readline()
        data1 = s.strip() # remove leading and ending blanks
        if (len(data1) <= 0):
            break

        # since we use append, value must be created in the loop
        val=[0.]
        size=104
        value=[val]*104
        target=[0.]

        data1 = data1.replace('[', '') # remove [
        data1 = data1.replace(']', '') # remove ]
        strs105 = data1.split()  # array of 2 str

        # convert to real
        for i in range(104):#0-103 共104各
            value[i] = eval(strs105[i])
        target[0]=eval(strs105[104])

        #print("row = {}".format(row) + ", {}\n".format(value), end='')

        recArr.append(value) ; # add 1 record at ending
        clsArr.append(target)

        row = row+1 # total read counter

    # close input file
    f.close()

    npXY = np.array(recArr)
    npXY=npXY.reshape(506,104)
    npC  = np.array(clsArr)
    npC=npC.reshape(506,)
    return npXY, npC
############################################################
#讀檔
lf= "C:\\prog\\ML\\hw4\\hw4_boston.csv"
```

Ln: 54 Col: 0


109611099_顏彥臣_Ridge_HW4.py - C:\prog\ML\hw4\109611099_hw4\109611099_顏彥臣_Ridge_HW4.py (3.8.3)
File Edit Format Run Options Window Help
row = row+1 # total read counter
# close input file
f.close()
npXY = np.array(recArr)
npXY=npXY.reshape(506,104)
npC = np.array(clsArr)
npC=npC.reshape(506,)
return npXY, npC
###
#讀檔
f= "C:\\prog\\ML\\hw4\\hw4_boston.csv"
x,y= Read(f)
###
print(x.shape) #印出所有case
###
for j in range(100):
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size = 0.169, random_state=j)
from sklearn.linear_model import Ridge
from sklearn.model_selection import cross_val_score
kfold = sklearn.model_selection.KFold(n_splits=5)
for i in np.arange(1,10):
model = Ridge(alpha=0.1)
model.fit(X_train, y_train)
trainscore=cross_val_score(model, X_train, y_train, cv=kfold)
testscore=cross_val_score(model, X_test, y_test, cv=kfold)
if trainscore.min()<0.8:
break
if (trainscore[0]-testscore[0]<0.20) and (trainscore[1]-testscore[1]<0.20) and (trainscore[2]-testscore[2]
and (trainscore[3]-testscore[3]<0.20) and (trainscore[4]-testscore[4]<0.20):
print(X_train.shape)
print(X_test.shape)
print('Ridge Boston(alpha: {}),fold 0,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[0],testscore
print('Ridge Boston(alpha: {}),fold 1,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[1],testscore
print('Ridge Boston(alpha: {}),fold 2,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[2],testscore
print('Ridge Boston(alpha: {}),fold 3,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[3],testscore
print('Ridge Boston(alpha: {}),fold 4,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[4],testscore
a=np.mean(trainscore)
b=np.mean(testscore)
print("\n")
print('Ridge Boston(alpha: {}),5-fold Train/Test average score:{:.2f}/{:.2f}'.format(i,a,b))
break
print("\n")
a=model.score(X_train, y_train)
b=model.score(X_test, y_test)
print('Ridge Boston(alpha: {}),5-fold/verify score:{:.2f}/{:.2f}'.format(i,a,b))
Ln: 83 Col: 69

結果:



```
Python 3.8.3 Shell
File  Edit  Shell  Debug  Options  Window  Help
Python 3.8.3 (default, Jul  2 2020, 17:30:36) [MSC v.1916 64 bit (AMD64)] on win
32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
========== RESTART: C:\prog\ML\hw4\109611099_hw4\109611099_顏彥臣_Ridge_HW4.py ==
==
(506, 104)
(420, 104)
(86, 104)
Ridge Boston(alpha: 1),fold 0,Train/Test score:0.82/0.71
Ridge Boston(alpha: 1),fold 1,Train/Test score:0.88/0.83
Ridge Boston(alpha: 1),fold 2,Train/Test score:0.82/0.84
Ridge Boston(alpha: 1),fold 3,Train/Test score:0.86/0.78
Ridge Boston(alpha: 1),fold 4,Train/Test score:0.86/0.68


Ridge Boston(alpha: 1),5-fold Train/Test average score:0.85/0.77


Ridge Boston(alpha: 1),5-fold/verify score:0.91/0.81
>>>
```

```
File  Edit  Format  Run  Options  Window  Help
        row = row+1 # total read counter

    # close input file
    f.close()

    npXY = np.array(recArr)
    npXY=npXY.reshape(506,104)
    npC  = np.array(clsArr)
    npC=npC.reshape(506,)
    return npXY, npC
##########################################################
#讀檔
f= "C:\\prog\\ML\\hw4\\hw4_boston.csv"
x,y= Read(f)
##########################################################
print(x.shape) #印出所有case
##########################################################
for j in range(100):
    X_train, X_test, y_train, y_test = train_test_split(x, y, test_size = 0.169, random_state=j)
    from sklearn.linear_model import Ridge
    from sklearn.model_selection import cross_val_score
    kfold = sklearn.model_selection.KFold(n_splits=5)
    for i in np.arange(1,10):
        model = Ridge(alpha=0.1)
        model.fit(X_train, y_train)
        trainscore=cross_val_score(model, X_train, y_train, cv=kfold)
        testscore=cross_val_score(model,  X_test, y_test, cv=kfold)
        if trainscore.min()<0.8:
            break
        if (trainscore[0]-testscore[0]<0.20) and (trainscore[1]-testscore[1]<0.20) and (trainscore[2]-testscore[2]
        and (trainscore[3]-testscore[3]<0.20) and (trainscore[4]-testscore[4]<0.20):
            print(X_train.shape)
            print(X_test.shape)
            print('Ridge Boston(alpha: {}),fold 0,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[0],testscore
            print('Ridge Boston(alpha: {}),fold 1,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[1],testscore
            print('Ridge Boston(alpha: {}),fold 2,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[2],testscore
            print('Ridge Boston(alpha: {}),fold 3,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[3],testscore
            print('Ridge Boston(alpha: {}),fold 4,Train/Test score:{:.2f}/{:.2f}'.format(i,trainscore[4],testscore
            a=np.mean(trainscore)
            b=np.mean(testscore)
            print("\n")
            print('Ridge Boston(alpha: {}),5-fold Train/Test average score:{:.2f}/{:.2f}'.format(i,a,b))
            break
print("\n")
a=model.score(X_train, y_train)
b=model.score(X_test, y_test)
print('Ridge Boston(alpha: {}),5-fold/verify score:{:.2f}/{:.2f}'.format(i,a,b))
```