practical-5-1

March 7, 2025

[9]:
```
# Name:-Henil Purohit
# Roll No:-24BEE080
import pandas as pd
import numpy as np
import math
data=pd.read_csv('BostonHousing.csv',header='infer').values
print(data)

X=data[:,0:-1]
y=data[:,-1]

nrows=data.shape[0]
print("Total Rows:",nrows)

test_split=float(input("Enter a number between 0 and 1 to specify the test␣
 ↪split:"))

nrows_train=math.floor((1-test_split)*nrows)
all_indices=np.random.permutation(nrows)

X_train=X[all_indices[0:nrows_train],:]

y_train=y[all_indices[0:nrows_train:]]

X_test=X[all_indices[nrows_train:],:]

y_test=y[all_indices[nrows_train:]]

print("Shapes:",X_train.shape,y_train.shape,X_test.shape,y_test.shape)

print("Union:",len(set(all_indices[0:nrows_train]).
 ↪union(all_indices[nrows_train:])))

print("interaction:",len(set(all_indices[0:nrows_train]).
 ↪intersection(all_indices[nrows_train:])))
```

```
[[6.3200e-03 1.8000e+01 2.3100e+00 … 3.9690e+02 4.9800e+00 2.4000e+01]
 [2.7310e-02 0.0000e+00 7.0700e+00 … 3.9690e+02 9.1400e+00 2.1600e+01]
```

```
 [2.7290e-02 0.0000e+00 7.0700e+00 … 3.9283e+02 4.0300e+00 3.4700e+01]
 …
 [6.0760e-02 0.0000e+00 1.1930e+01 … 3.9690e+02 5.6400e+00 2.3900e+01]
 [1.0959e-01 0.0000e+00 1.1930e+01 … 3.9345e+02 6.4800e+00 2.2000e+01]
 [4.7410e-02 0.0000e+00 1.1930e+01 … 3.9690e+02 7.8800e+00 1.1900e+01]]
Total Rows: 506

Enter a number between 0 and 1 to specify the test split: 0.5

Shapes: (253, 13) (253,) (253, 13) (253,)
Union: 506
interaction: 0
```

[2]:
```
import numpy as np
from sklearn.model_selection import train_test_split
X=np.random.randint(100,size=20).reshape((10,2))
y=np.random.randint(100,size=10)
X_train,X_test,y_train,y_test=train_test_split(X,y,random_state=0)
print(X_train)
print(y_train)
```

```
[[85  6]
 [70  8]
 [70 67]
 [44 29]
 [97 60]
 [88 25]
 [12 47]]
[72 85  3 93 76 20 26]
```

[4]:
```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
data=pd.read_csv("BostonHousing.csv",header='infer').values
x=data[:,0:-1]
y=data[:,-1]

nrows=data.shape[0]
print("Total rows:",nrows)
test_split=float(input("Enter a numbewr between 0 and 1 to specify the test␣
 ↪split:"))
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=test_split)
print("Shapes:",x_train.shape,y_train.shape,x_test.shape,y_test.shape)
```

```
Total rows: 506

Enter a numbewr between 0 and 1 to specify the test split: 0.2

Shapes: (404, 13) (404,) (102, 13) (102,)
```

[ ]: