KNN

February 17, 2022

[7]:
```
import pandas as pd
from sklearn.datasets import load_boston
import matplotlib.pyplot as plt
from sklearn.neighbors import KNeighborsRegressor
from sklearn.model_selection import cross_val_score
import numpy as np
from sklearn.metrics import mean_squared_error
import math
```

[4]:
```
Boston=load_boston()
```

[5]:
```
f = pd.DataFrame(Boston.data)
li_ftrs = list(Boston.feature_names)
f.columns = Boston.feature_names
f["PRICES"] = Boston.target
```

[6]:
```
f.head()
```

[6]:

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	\
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	

	PTRATIO	B	LSTAT	PRICES
0	15.3	396.90	4.98	24.0
1	17.8	396.90	9.14	21.6
2	17.8	392.83	4.03	34.7
3	18.7	394.63	2.94	33.4
4	18.7	396.90	5.33	36.2

[8]:
```
def knn(x):
    rng = np.random.default_rng(x)
    idx_feat = (np.floor(13*rng.uniform(size=4))).astype(int)
    X = Boston["data"][:,idx_feat]
    label = Boston["feature_names"][idx_feat]
    Y = Boston["target"]
```

```
    model = KNeighborsRegressor(n_neighbors=4).fit(X,Y)
    selected_features = [li_ftrs[index] for index in idx_feat]
    print(selected_features)
    score = model.score(X, Y)
    print(score)
```

[10]:
```
lst = []
n = int(input("Enter number of elements : "))
for i in range(0, n):
    ele = int(input())

    lst.append(ele)

print(lst)
for i in range(len(lst)):
    knn(i)
```

```
Enter number of elements : 5
122
56728
17364
24641
23648
[122, 56728, 17364, 24641, 23648]
['RAD', 'CHAS', 'CRIM', 'CRIM']
0.5751234289240716
['AGE', 'LSTAT', 'ZN', 'LSTAT']
0.7545728621093453
['CHAS', 'CHAS', 'PTRATIO', 'ZN']
0.41826464271846775
['ZN', 'CHAS', 'PTRATIO', 'DIS']
0.7236724765029688
['LSTAT', 'AGE', 'LSTAT', 'ZN']
0.7545728621093453
```

[ ]: