

151-15-5116_Lab-4

February 23, 2018

```
In [6]: import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from scipy.spatial import distance
from sklearn.metrics import accuracy_score
from sklearn.neighbors import KNeighborsClassifier
import matplotlib.pyplot as plt

data = pd.read_csv('iris.csv')

a = data.values[:, :4]

b = np.zeros(150)

for i in range(len(b)):
    if data.values[i, 4]=='setosa':
        b[i] = 0
    elif data.values[i, 4]=='versicolor':
        b[i] = 1
    elif data.values[i, 4]=='virginica':
        b[i] = 2

a_train, a_test, b_train, b_test = train_test_split(a, b, test_size=0.33, random_state=4)
d = distance.cdist(a_test, a_train, 'euclidean')

for i in range(50):
    b1=np.argsort(d,axis=1)
    j=int(7)
    k=int(0)
    b_predict=[]
    for j in range(0,50):
```

```

b2 = np.zeros(3)
for i in range (0,j):
    inde=int(b1[k][i])
    value=int(b_train[inde])
    b2[value]+=1

b_predict.append(np.argmax(b2, axis=0))
accuracy_score(b_test, b_predict)

```

Out[6]: 0.28000000000000003

```

In [18]: k2 = np.array([1,2,5,7,9,11,13,15])
result=[]
r=int(0)
for k1 in k2:
    b_predict=[]
    for j in range(0,50):
        b2 = np.zeros(3)
        for i in range (0,k1):
            inde=int(b1[j][i])
            value=int(b_train[inde])
            b2[value]+=1

        b_predict.append(np.argmax(b2, axis=0))
    result.append(accuracy_score(b_test,b_predict))
final=np.array(result)
plt.plot(final,k2)
plt.show()

```



