Homework1

January 14, 2021

0.0.1 Qinhua Sun

1 Problem 1: Python & Data Exploration

[6]:
```
import numpy as np
import matplotlib.pyplot as plt
```

[7]:
```
iris = np.genfromtxt("data/iris.txt",delimiter=None)   # load the text file
Y = iris[:,-1]      # target value is the last column
X = iris[:,0:-1]    # features are the other columns
```

Question 1

[8]:
```
print(X.shape)
```

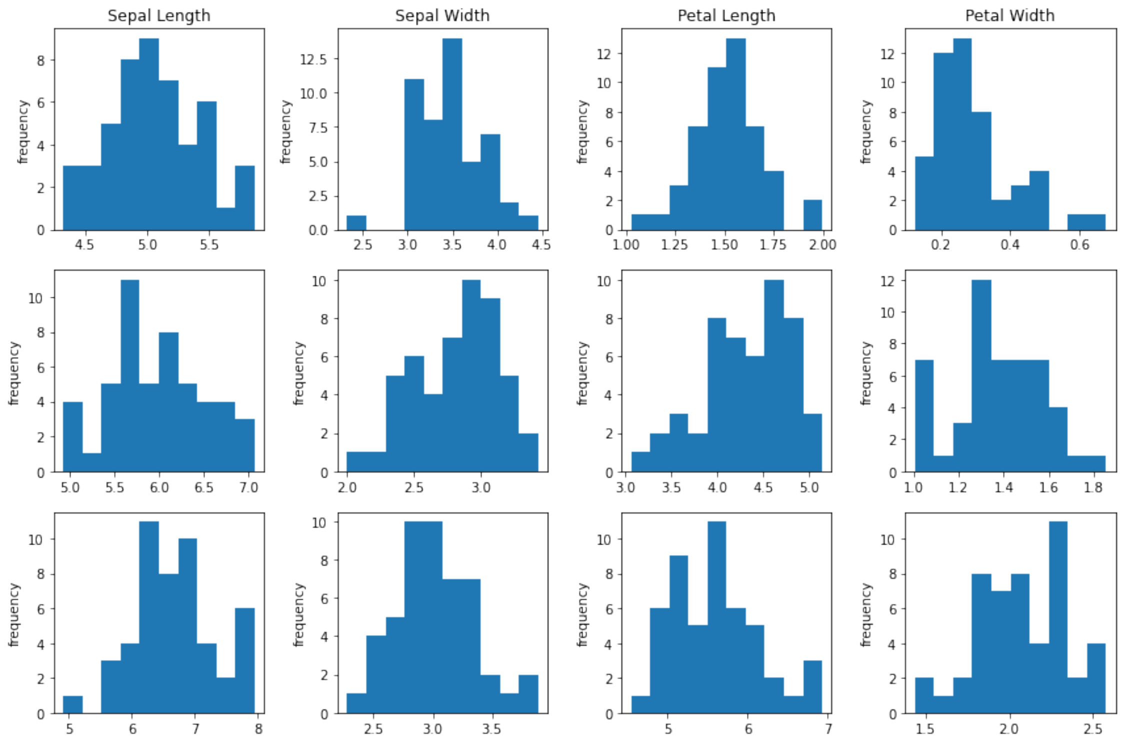
```
(148, 4)
```

There are 148 data points and 4 features.

1.1 Question 2

[9]:
```
# slicing the data into 3 groups
X1 = X[Y==0]
X2 = X[Y==1]
X3 = X[Y==2]

# # plot the data
fig, ax = plt.subplots(3, 4, tight_layout=True, figsize=(12,8))
features = ['Sepal Length','Sepal Width','Petal Length','Petal Width']

for i in range(4):
    ax[0,i].hist(X1[:,i])
    ax[0,i].set_title(features[i])

for i in range(4):
    ax[1,i].hist(X2[:,i])

for i in range(4):
    ax[2,i].hist(X3[:,i])
```

```python
for ax in ax.flat:
    ax.set(ylabel='frequency')
```



1.2 Question 3

[10]:
```python
print('Mean:',np.mean(X, axis=0))
print('Median:',np.median(X, axis=0))
print('Variance', np.var(X, axis=0))
print('SD', np.std(X, axis=0))
```

```
Mean: [5.90010376 3.09893092 3.81955484 1.25255548]
Median: [5.84664255 3.05980605 4.3998377  1.361768  ]
Variance [0.694559   0.19035057 3.07671634 0.57573564]
SD [0.83340207 0.43629184 1.75405711 0.75877246]
```

1.3 Question 4

[11]:
```python
import more_itertools
fig, ax = plt.subplots(1, 3, tight_layout=True, figsize=(15,5))
colors = ["b", "g", "r"]

for i in range(3):
```

```
    ax[0].scatter(X[Y==i][:,0],X[Y==i][:,1], color = colors[i], label = 'group␣
↪%d' % int(i+1))
    ax[0].set_title('Feature 1 vs. Features 2 ')
    ax[0].legend()
    ax[1].scatter(X[Y==i][:,1],X[Y==i][:,2], color = colors[i], label = 'group␣
↪%d' % int(i+1))
    ax[1].set_title('Feature 2 vs. Features 3 ')
    ax[1].legend()
    ax[2].scatter(X[Y==i][:,2],X[Y==i][:,3], color = colors[i], label = 'group␣
↪%d' % int(i+1))
    ax[2].set_title('Feature 3 vs. Features 4 ')
    ax[2].legend()
```

2 Problem 2

2.1 Question 1

When the matrix is square and the determinant of the matrix is not zero.

2.2 Question 2 - 5

#2 Det(A), Det(B)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -8 & -2 \\ 1 & -4 & -2 \\ -4 & 4 & 1 \end{bmatrix}$$

$$|A| = 1 \times ((-1)\times 2 - 3\times 1) - 2\times(2\times 2 - 1\times 1) + 3\times(\overset{6}{2\times 3} - 1\times(-1))$$
$$= 1\times(-5) - 2\times 3 + 3\times 7$$
$$= -5 - 6 + 21$$
$$= \boxed{10}$$

$$|B| = 0 - (-8)\times(1\times 1 - \overset{8}{(-4)\times(-2)}) + (-2)\times(1\times 4 - \overset{16}{(-4)\times(-4)})$$
$$= 0 - (-8)\times(-7) + (-2)\times(-12)$$
$$= 0 - 56 + 24$$
$$= \boxed{-32}$$

#3. Inverse of A, B.

Matrix of minors
$$\begin{bmatrix} (-1)\times 2 - 3\times 1 & 2\times 2 - 1\times 1 & 2\times(3) - 1\times(-1) \\ (2\times 2 - 3\times 3) & 1\times 2 - 1\times 3 & 1\times 3 - 2\times 1 \\ 2\times 1 - (-1)\times 3 & 1\times 1 - 2\times 3 & 1\times(-1) - 2\times 2 \end{bmatrix}$$

Cofactor
$$\Rightarrow \begin{bmatrix} -5 & 3 & 7 \\ -5 & -1 & 1 \\ 5 & -5 & -5 \end{bmatrix} \Rightarrow \begin{bmatrix} -5 & -3 & 7 \\ 5 & -1 & -1 \\ 5 & 5 & -5 \end{bmatrix}$$

$$\Rightarrow A^{-1} = \frac{1}{10}\begin{bmatrix} -5 & 5 & 5 \\ -3 & -1 & 5 \\ 7 & -1 & -5 \end{bmatrix} = \begin{bmatrix} -0.5 & 0.5 & 0.5 \\ -0.3 & -0.1 & 0.5 \\ 0.7 & -0.1 & -0.5 \end{bmatrix}$$

Matrix of minors of B.

$$\begin{bmatrix} (-4)\times 1 - 4\times(-2) & 1\times 1 - (-4)\times(-2) & 1\times 4 - (-4)\times(-4) \\ (-8)\times(-1) - 4\times(-2) & 0\times 1 - (-4)\times(-2) & 0\times 4 - (-4)\times(-8) \\ (-8)\times(-2) - (-4)\times(-2) & 0\times(-2) - 1\times(-2) & 0\times(-4) - 1\times(-8) \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -7 & -8 \\ 16 & -8 & -32 \\ 8 & 2 & 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 7 & -8 \\ -16 & -8 & 32 \\ 8 & -2 & 8 \end{bmatrix}$$

$$B^{-1} = \frac{1}{-32}\begin{bmatrix} 4 & 0 & 8 \\ 7 & -8 & -2 \\ -8 & 32 & 8 \end{bmatrix} = \begin{bmatrix} -\frac{1}{8} & 0 & -\frac{1}{4} \\ -\frac{7}{32} & \frac{1}{4} & -\frac{1}{16} \\ \frac{1}{4} & -1 & -\frac{1}{4} \end{bmatrix}$$

#4.

$$(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} -0.5 & -0.3 & 0.7 \\ 0.5 & -0.1 & -0.1 \\ 0.5 & 0.5 & -0.5 \end{bmatrix}$$

$$(B^T)^{-1} = (B^{-1})^T = \begin{bmatrix} -\frac{1}{8} & -\frac{7}{32} & \frac{1}{4} \\ 0 & \frac{1}{4} & -1 \\ -\frac{1}{4} & -\frac{1}{16} & -\frac{1}{4} \end{bmatrix}$$

#5.

$$C = AB \Rightarrow C^{-1} = (AB)^{-1} = B^{-1}A^{-1}$$

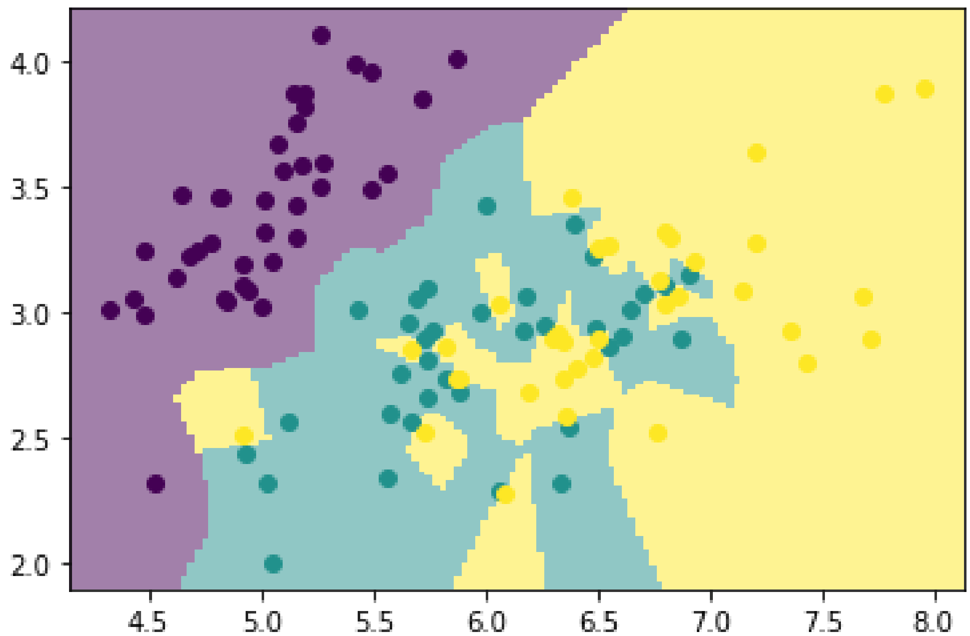
$$\begin{bmatrix} -\frac{1}{8} & 0 & -\frac{1}{4} \\ -\frac{7}{32} & \frac{1}{4} & -\frac{1}{16} \\ \frac{1}{4} & -1 & -\frac{1}{4} \end{bmatrix} \times \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{10} & -\frac{1}{10} & -\frac{1}{2} \\ \frac{1}{10} & -\frac{1}{10} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{9}{80} & -\frac{3}{80} & \frac{1}{16} \\ \frac{5}{64} & -\frac{9}{64} & -\frac{1}{64} \\ -\frac{1}{16} & \frac{5}{16} & -\frac{3}{16} \end{bmatrix}$$

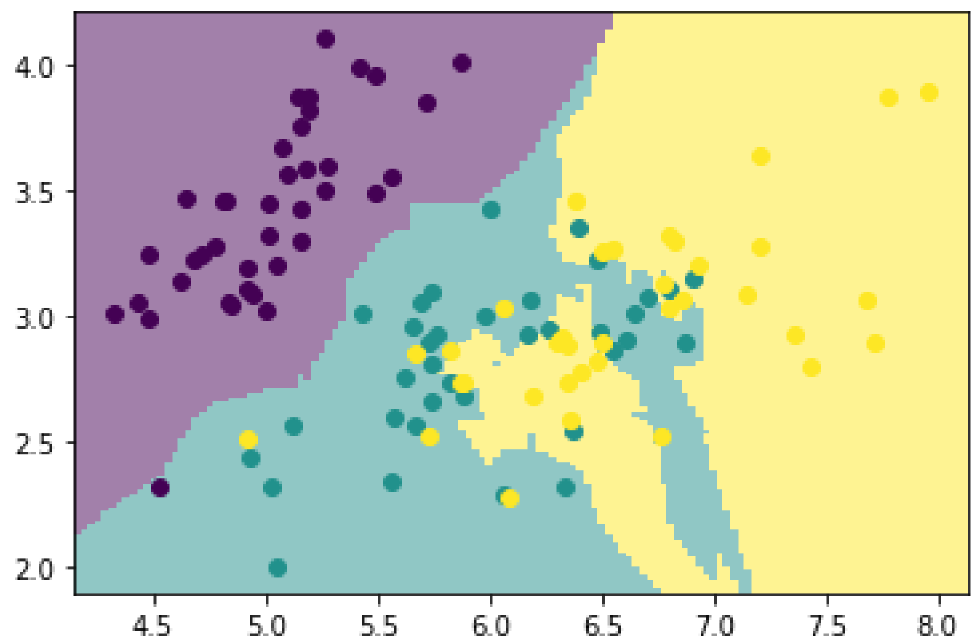
3 Probelm 3

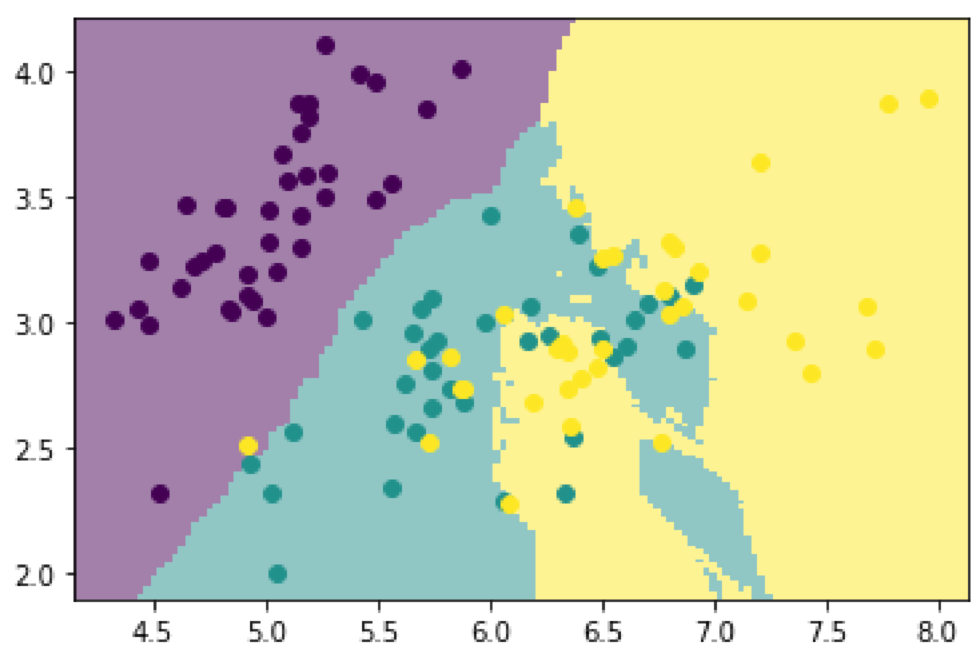
[52]:
```
import mltools as ml
np.random.seed(0)

iris = np.genfromtxt("data/iris.txt", delimiter=None)
Y = iris[:,-1]
X = iris[:,0:2]

# shuffle the data
X,Y = ml.shuffleData(X,Y)

# split the data into 80/20 train/validation set
Xtr,Xva,Ytr,Yva = ml.splitData(X,Y, 0.75);
```

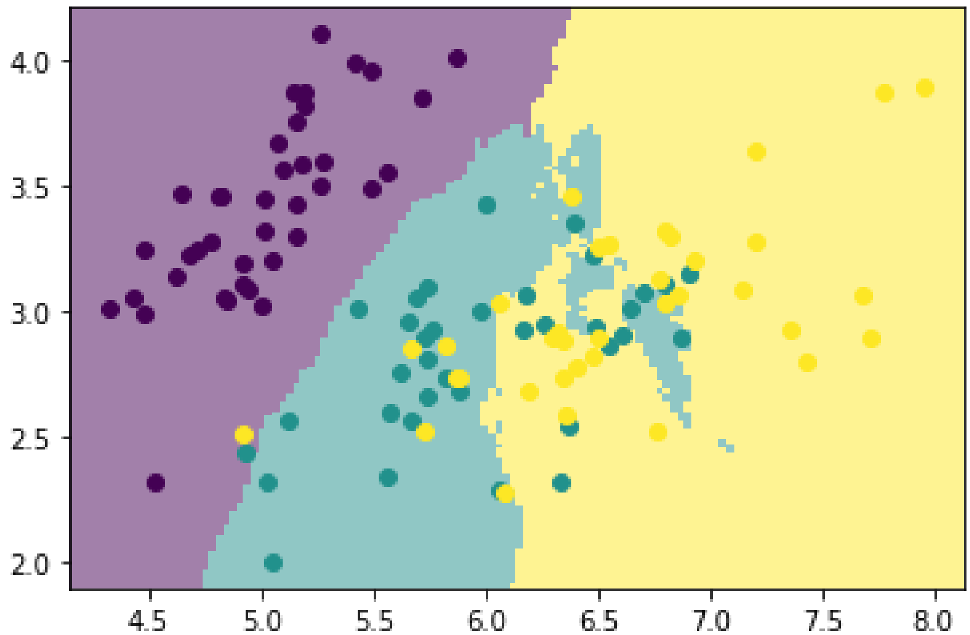
3.1 Question 1

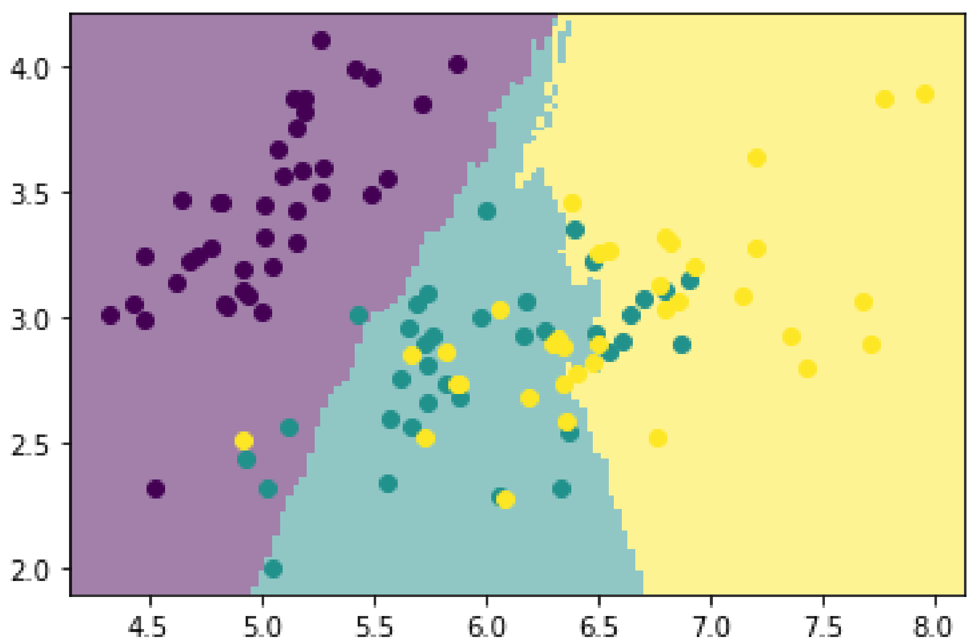
[67]:
```
k= [1,5,10,15,50]

for i,k in enumerate(k):
    knn = ml.knn.knnClassify()
    knn.train(Xtr, Ytr, k)
    YvaHat = knn.predict(Xva)
    ml.plotClassify2D(knn, Xtr, Ytr)
```




4.0
3.5
3.0
2.5
2.0
4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0


4.0
3.5
3.0
2.5
2.0
4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0



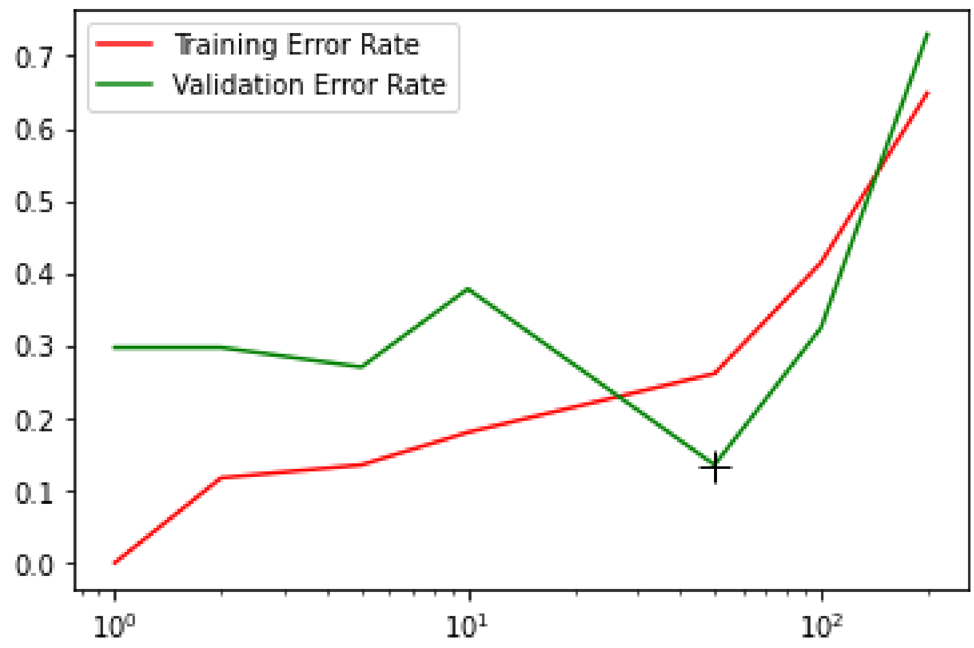


As we increase the value of K, wea're averaging more neighbors and therefore will have smoother decision boundary. However, the error will also be increasing.

3.2 Question 2

[128]:
```
K = [1,2,5,10,50,100,200];

errTrain =[]
errVa = []

for i,k in enumerate(K):
    learner = ml.knn.knnClassify()
    k = K[i]
    learner.train(Xtr, Ytr, k)
    Yhat = learner.predict(Xtr)
    errTrain.append(np.mean(Yhat != Ytr))
    YvaHat = learner.predict(Xva)
    errVa.append(np.mean(YvaHat != Yva))

print(errTrain)
print(errVa)

plt.semilogx(K, errTrain, "r", label = "Training Error Rate")
plt.semilogx(K, errVa, "g", label = "Validation Error Rate" )
plt.legend()

plt.semilogx(K[np.argmin(errVa)],min(errVa),'k+',markersize=12)
print('\n')
print(K[np.argmin(errVa)],'is when validation error is the lowest')
```

```
[0.0, 0.11711711711711711, 0.13513513513513514, 0.18018018018018017,
0.26126126126126126, 0.4144144144144144, 0.6486486486486487]
[0.2972972972972973, 0.2972972972972973, 0.2702702702702703, 0.3783783783783784,
0.13513513513513514, 0.32432432432432434, 0.7297297297297297]


50 is when validation error is the lowest
```



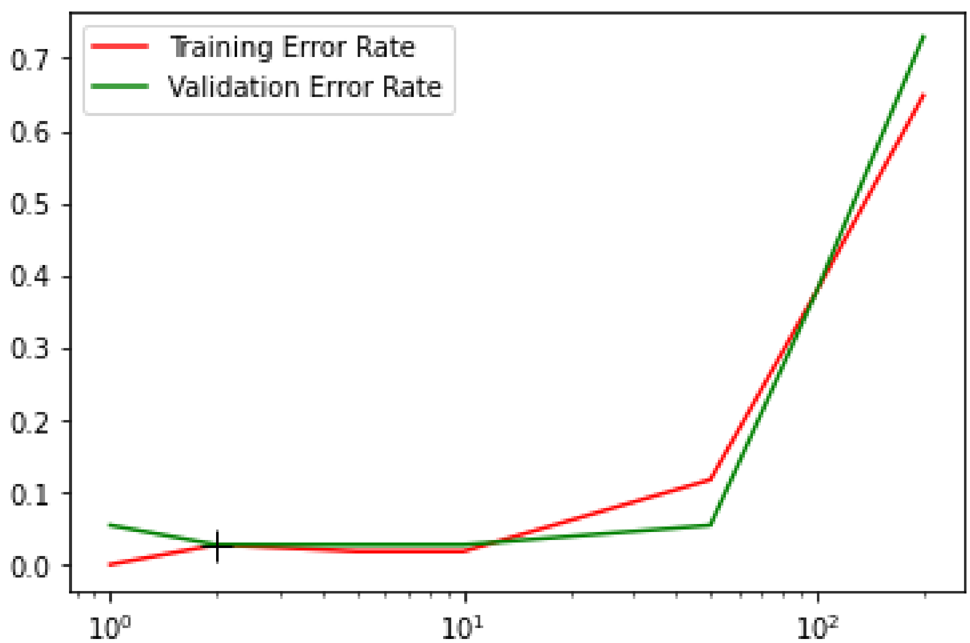
I will choose k=50 as it has the lowest validation error rate.

3.3 Question 3

[129]:
```
import mltools as ml
np.random.seed(0)

iris = np.genfromtxt("data/iris.txt", delimiter=None)
Y = iris[:,-1]
X = iris[:,0:-1]

# shuffle the data
X,Y = ml.shuffleData(X,Y)

# split the data into 80/20 train/validation set
Xtr,Xva,Ytr,Yva = ml.splitData(X,Y, 0.75);

K = [1,2,5,10,50,100,200];

errTrain =[]
errVa = []

for i,k in enumerate(K):
    learner = ml.knn.knnClassify()
    k = K[i]
    learner.train(Xtr, Ytr, k)
```

```
    Yhat = learner.predict(Xtr)
    errTrain.append(np.mean(Yhat != Ytr))
    YvaHat = learner.predict(Xva)
    errVa.append(np.mean(YvaHat != Yva))

print(errTrain)
print(errVa)

plt.semilogx(K, errTrain, "r", label = "Training Error Rate")
plt.semilogx(K, errVa, "g", label = "Validation Error Rate" )
plt.legend()

plt.semilogx(K[np.argmin(errVa)],min(errVa),'k+',markersize=12)
print('\n')
print(K[np.argmin(errVa)],'is when validation error is the lowest')
```

```
[0.0, 0.02702702702702703, 0.018018018018018018, 0.018018018018018018,
0.11711711711711711, 0.3783783783783784, 0.6486486486486487]
[0.05405405405405406, 0.02702702702702703, 0.02702702702702703,
0.02702702702702703, 0.05405405405405406, 0.3783783783783784,
0.7297297297297297]


2 is when validation error is the lowest
```



The plot generated by using all four features looks very different. Now I would recommand using k=2.

4 Problem 4

4.1 Question 1

Class probability:

$$p(y=+1) = \frac{4}{10} = \frac{2}{5}$$
$$p(y=-1) = \frac{6}{10} = \frac{3}{5}$$

Individual feature probability:

$$P(X_1 = 0|y=+1) = \frac{1}{4}$$
$$P(X_1 = 1|y=+1) = \frac{3}{4}$$
$$P(X_1 = 0|y=-1) = \frac{1}{2}$$
$$P(X_1 = 1|y=-1) = \frac{1}{2}$$

$$P(X_2 = 0|y=+1) = 1$$
$$P(X_2 = 1|y=+1) = 0$$
$$P(X_2 = 0|y=-1) = \frac{1}{6}$$
$$P(X_2 = 1|y=-1) = \frac{5}{6}$$

$$P(X_3 = 0|y=+1) = \frac{1}{4}$$
$$P(X_3 = 1|y=+1) = \frac{3}{4}$$
$$P(X_3 = 0|y=-1) = \frac{1}{3}$$
$$P(X_3 = 1|y=-1) = \frac{2}{3}$$

$$P(X_4 = 0|y=+1) = \frac{1}{2}$$
$$P(X_4 = 1|y=+1) = \frac{1}{2}$$
$$P(X_4 = 0|y=-1) = \frac{1}{6}$$
$$P(X_4 = 1|y=-1) = \frac{5}{6}$$

$$P(X_5 = 0|y = +1) = \frac{3}{4}$$

$$P(X_5 = 1|y = +1) = \frac{1}{4}$$

$$P(X_5 = 0|y = -1) = \frac{2}{3}$$

$$P(X_5 = 1|y = -1) = \frac{1}{3}$$

4.2 Question 2

$$p(y = 1|x = (00000) = p(y = +1) \times \frac{1}{4} \times 1 \times \frac{1}{4} \times \frac{1}{2} \times \frac{3}{4} = 0.0093$$

$$p(y = -1|x = (00000) = p(y = -1) \times \frac{1}{2} \times 1 \times \frac{1}{6} \times \frac{1}{3} \times \frac{1}{6} \times \frac{2}{3} = 0.0019$$

x = (0 0 0 0 0) would be predicted for as +1, and x = (1 1 0 1 0) would be -1.

4.3 Question 3

$$p(y = +1|x = (11010) = \frac{p(x = (11010)|y = +1)) \times p(y = +1)}{p(x = (11010))}$$

since

$$p(x = (11010)|y = +1) = \frac{3}{4} \times 0 \times ...$$

The posterior probability is virtually 0.

4.4 The remaining questions will be submitted through next submission.

Statement of Collaboration

This is the Statement of Collaboration,meaning that I have followed the academic honesty guidelines and did not discussed this assignment with anyone.

Qinhua Sun