Name : Mohit Manish Bhavsar

Roll No : 20U437

Div : 4

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

```
dataset = pd.read_csv('https://raw.githubusercontent.com/mk-gurucharan/Classification/master/IrisDataset.csv')
X = dataset.iloc[:,:4].values
y = dataset['species'].values
dataset.head(5)
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)
```

```
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

```
from sklearn.naive_bayes import GaussianNB
classifier = GaussianNB()
classifier.fit(X_train, y_train)
```

```
GaussianNB()
```

```
y_pred = classifier.predict(X_test)
y_pred
```

```
array(['virginica', 'setosa', 'virginica', 'versicolor', 'virginica',
       'setosa', 'versicolor', 'virginica', 'virginica', 'setosa',
       'versicolor', 'virginica', 'versicolor', 'virginica', 'setosa',
       'setosa', 'setosa', 'versicolor', 'versicolor', 'virginica',
       'versicolor', 'versicolor', 'setosa', 'setosa', 'virginica',
       'virginica', 'setosa', 'versicolor', 'versicolor', 'versicolor'],
      dtype='<U10')
```

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
```

```
from sklearn.metrics import accuracy_score
print ("Accuracy : ", accuracy_score(y_test, y_pred))
cm
```

```
Accuracy :  0.9666666666666667
```

```
array([[ 9,  0,  0],
       [ 0, 10,  0],
       [ 0,  1, 10]])
```

```
df = pd.DataFrame({'Real Values':y_test, 'Predicted Values':y_pred})
df
```

	Real Values	Predicted Values
0	virginica	virginica
1	setosa	setosa
2	virginica	virginica
3	versicolor	versicolor
4	virginica	virginica
5	setosa	setosa
6	versicolor	versicolor
7	virginica	virginica
8	virginica	virginica
9	setosa	setosa
10	versicolor	versicolor
11	virginica	virginica
12	versicolor	versicolor
13	virginica	virginica
14	setosa	setosa
15	setosa	setosa
16	setosa	setosa
17	versicolor	versicolor
18	versicolor	versicolor
19	virginica	virginica
20	versicolor	versicolor
21	versicolor	versicolor
22	setosa	setosa
23	setosa	setosa
24	virginica	virginica
25	virginica	virginica
26	setosa	setosa
27	versicolor	versicolor
28	versicolor	versicolor
29	virginica	versicolor

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
df =
pd.read_csv('https://gist.githubusercontent.com/netj/8836201/raw/6f9306ad2139
8ea43cba4f7d537619d0e07d5ae3/iris.csv')

df.head()
df.dtypes
```

```
sepal.length    float64
sepal.width     float64
petal.length    float64
petal.width     float64
variety          object
dtype: object
```

```
df.head()
#To know the data types of the variables.
df.dtypes
```

```
sepal.length    float64
sepal.width     float64
petal.length    float64
petal.width     float64
variety          object
dtype: object
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
random_state=0)

print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)
```

```
(105, 4)
(105,)
(45, 4)
(45,)
```

```
from sklearn.svm import SVC
from sklearn.metrics import confusion_matrix

clf = SVC(kernel = 'linear').fit(X_train,y_train)
clf.predict(X_train)
```

```
array(['versicolor', 'virginica', 'virginica', 'virginica', 'virginica',
       'versicolor', 'virginica', 'versicolor', 'versicolor', 'virginica',
       'virginica', 'virginica', 'virginica', 'versicolor', 'virginica',
       'versicolor', 'setosa', 'virginica', 'versicolor', 'versicolor',
       'versicolor', 'versicolor', 'virginica', 'setosa', 'setosa',
       'virginica', 'versicolor', 'setosa', 'setosa', 'versicolor',
       'setosa', 'virginica', 'versicolor', 'setosa', 'versicolor',
```

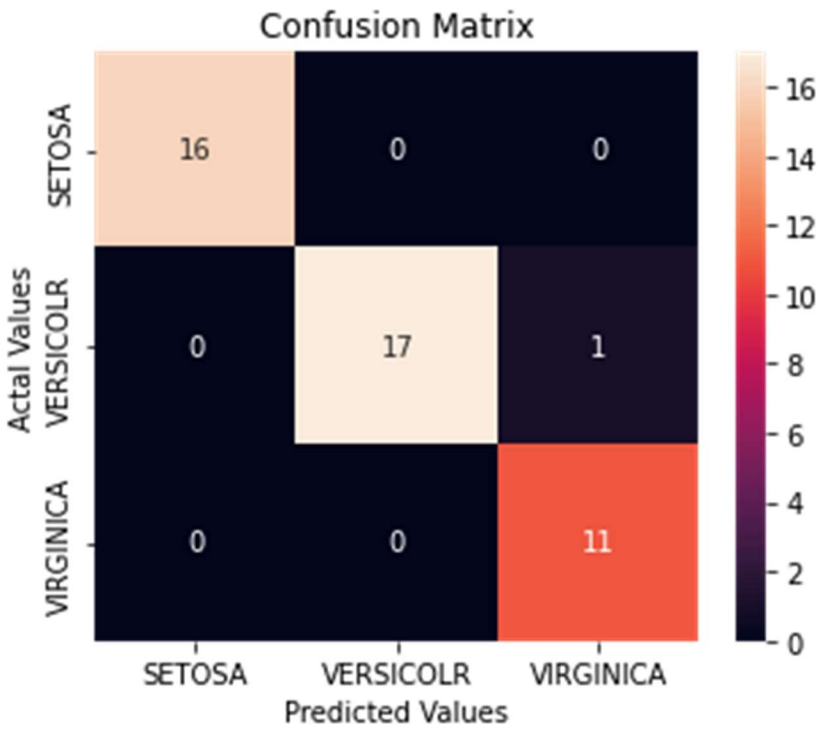
```
       'virginica', 'versicolor', 'setosa', 'virginica', 'virginica',
       'virginica', 'virginica', 'setosa', 'setosa', 'virginica',
       'virginica', 'setosa', 'virginica', 'setosa', 'virginica',
       'virginica', 'setosa', 'setosa', 'virginica', 'setosa', 'setosa',
       'setosa', 'versicolor', 'virginica', 'virginica', 'setosa',
       'setosa', 'setosa', 'versicolor', 'versicolor', 'setosa', 'setosa',
       'versicolor', 'setosa', 'virginica', 'versicolor', 'virginica',
       'versicolor', 'setosa', 'virginica', 'setosa', 'virginica',
       'setosa', 'setosa', 'virginica', 'setosa', 'virginica',
       'versicolor', 'versicolor', 'versicolor', 'virginica', 'virginica',
       'virginica', 'versicolor', 'setosa', 'versicolor', 'virginica',
       'virginica', 'setosa', 'versicolor', 'versicolor', 'virginica',
       'versicolor', 'setosa', 'setosa', 'setosa', 'virginica',
       'versicolor', 'virginica', 'setosa'], dtype=object)

y_pred = clf.predict(X_test)

cm = confusion_matrix(y_test, y_pred)

cm_df = pd.DataFrame(cm,
                     index = ['SETOSA','VERSICOLR','VIRGINICA'],
                     columns = ['SETOSA','VERSICOLR','VIRGINICA'])

plt.figure(figsize=(5,4))
sns.heatmap(cm_df, annot=True)
plt.title('Confusion Matrix')
plt.ylabel('Actal Values')
plt.xlabel('Predicted Values')
plt.show()
```


Confusion Matrix
Actal Values
SETOSA
VERSICOLR
VIRGINICA
16
0
0
0
17
1
0
0
11
SETOSA
VERSICOLR
VIRGINICA
Predicted Values
16
14
12
10
8
6
4
2
0