

In [1]:

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

In [2]:

```
df=pd.read_csv('F:\\DSE\\3rd year engineering\\5th sem\\6th sem\\DSBDA\\dataset\\iris.csv')
```

In [3]:

```
df.head()
```

Out[3]:

| | 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 |

In [5]:

```
x=df.iloc[:, :4].values
y=df['species'].values
```

In [6]:

```
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)
```

In [9]:

```
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
x_train=sc.fit_transform(x_train)
x_test=sc.transform(x_test)
```

In [12]:

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

Out[12]:

```
GaussianNB()
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.

On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [14]:

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

Out[14]:

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

In [18]:

```
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: 0.9333333333333333

Out[18]:

```
array([[ 8,  0,  0],
       [ 0,  8,  2],
       [ 0,  0, 12]], dtype=int64)
```

In [21]:

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

Out[21]:

```
'          precision    recall  f1-score   support\n\n
1.00          1.00          1.00         8\n versicolor          1.00          0.80
0.89          10\n  virginica          0.86          1.00          0.92         12\n\n
accuracy          0.93          30\n macro avg          0.94          0.93          0.95
0.93          0.94          30\nweighted avg          0.94          0.93          0.93
30\n'
```

In []: