**PRACTICAL NO : 06**

**DATA ANALYICS 3**

**CODE :**

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

data=pd.read\_csv(r"E:\DSBDA\DSBDA Datasets\iris.csv")

print(data)

x=data.iloc[:,:4].values

y=data['species'].values

data.head(5)

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)

y\_pred=classifier.predict(x\_test)

y\_pred

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))

df=pd.DataFrame({'Real Values':y\_test, 'Predicted Values':y\_pred})

print(df)

OUTPUT :

runfile('E:/DSBDA/dsbdapr6.py', wdir='E:/DSBDA')

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

.. ... ... ... ... ...

145 6.7 3.0 5.2 2.3 virginica

146 6.3 2.5 5.0 1.9 virginica

147 6.5 3.0 5.2 2.0 virginica

148 6.2 3.4 5.4 2.3 virginica

149 5.9 3.0 5.1 1.8 virginica

[150 rows x 5 columns]

Accuracy : 0.9

Real Values Predicted Values

0 virginica virginica

1 versicolor versicolor

2 virginica virginica

3 virginica virginica

4 virginica virginica

5 setosa setosa

6 setosa setosa

7 virginica virginica

8 setosa setosa

9 setosa setosa

10 virginica virginica

11 virginica virginica

12 versicolor versicolor

13 versicolor virginica

14 setosa setosa

15 versicolor virginica

16 versicolor versicolor

17 setosa setosa

18 versicolor versicolor

19 versicolor versicolor

20 versicolor versicolor

21 versicolor virginica

22 virginica virginica

23 setosa setosa

24 versicolor versicolor

25 virginica virginica

26 versicolor versicolor

27 virginica virginica

28 virginica virginica

29 versicolor versicolor

data.head(5)

Out[2]:

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

y\_pred

Out[3]:

array(['virginica', 'versicolor', 'virginica', 'virginica', 'virginica',

'setosa', 'setosa', 'virginica', 'setosa', 'setosa', 'virginica',

'virginica', 'versicolor', 'virginica', 'setosa', 'virginica',

'versicolor', 'setosa', 'versicolor', 'versicolor', 'versicolor',

'virginica', 'virginica', 'setosa', 'versicolor', 'virginica',

'versicolor', 'virginica', 'virginica', 'versicolor'], dtype='<U10')

print("Accuracy :",accuracy\_score(y\_test,y\_pred))

Accuracy : 0.9

df=pd.DataFrame({'Real Values':y\_test, 'Predicted Values':y\_pred})

print(df)

Real Values Predicted Values

0 virginica virginica

1 versicolor versicolor

2 virginica virginica

3 virginica virginica

4 virginica virginica

5 setosa setosa

6 setosa setosa

7 virginica virginica

8 setosa setosa

9 setosa setosa

10 virginica virginica

11 virginica virginica

12 versicolor versicolor

13 versicolor virginica

14 setosa setosa

15 versicolor virginica

16 versicolor versicolor

17 setosa setosa

18 versicolor versicolor

19 versicolor versicolor

20 versicolor versicolor

21 versicolor virginica

22 virginica virginica

23 setosa setosa

24 versicolor versicolor

25 virginica virginica

26 versicolor versicolor

27 virginica virginica

28 virginica virginica

29 versicolor versicolor

print("Accuracy :",accuracy\_score(y\_test,y\_pred))

Accuracy : 0.9

**OUTPUT :**

