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


tennis=pd.read_csv('/content/tennis.csv')
tennis.head()
tennis.tail()
```

	day	outlook	temp	humidity	wind	play
9	D10	Rain	Mild	Normal	Weak	Yes
10	D11	Sunny	Mild	Normal	Strong	Yes
11	D12	Overcast	Mild	High	Strong	Yes
12	D13	Overcast	Hot	Normal	Weak	Yes
13	D14	Rain	Mild	High	Strong	No

```
X=tennis.iloc[:,1:5].values
y=tennis.iloc[:,-1].values
print(X)
print(y)
```

```
[['Sunny' 'Hot' 'High' 'Weak']
 ['Sunny' 'Hot' 'High' 'Strong']
 ['Overcast' 'Hot' 'High' 'Weak']
 ['Rain' 'Mild' 'High' 'Weak']
 ['Rain' 'Cool' 'Normal' 'Weak']
 ['Rain' 'Cool' 'Normal' 'Strong']
 ['Overcast' 'Cool' 'Normal' 'Strong']
 ['Sunny' 'Mild' 'High' 'Weak']
 ['Sunny' 'Cool' 'Normal' 'Weak']
 ['Rain' 'Mild' 'Normal' 'Weak']
 ['Sunny' 'Mild' 'Normal' 'Strong']
 ['Overcast' 'Mild' 'High' 'Strong']
 ['Overcast' 'Hot' 'Normal' 'Weak']
 ['Rain' 'Mild' 'High' 'Strong']]
['No' 'No' 'Yes' 'Yes' 'Yes' 'No' 'Yes' 'No' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes'
 'No']
```

```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
y=le.fit_transform(y)
print(y)
```

```
[0 0 1 1 1 0 1 0 1 1 1 1 1 0]
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.metrics import make_scorer, accuracy_score,precision_score,classification_report
from sklearn.naive_bayes import GaussianNB
from sklearn.naive_bayes import CategoricalNB


X_train,X_test,Y_train,Y_test=train_test_split(X,y,test_size=0.5,random_state=0)
print(X_train)
```

```
[['Sunny' 'Hot' 'High' 'Strong']
 ['Sunny' 'Mild' 'High' 'Weak']
 ['Sunny' 'Mild' 'Normal' 'Strong']
 ['Rain' 'Mild' 'High' 'Weak']
 ['Sunny' 'Hot' 'High' 'Weak']
 ['Rain' 'Cool' 'Normal' 'Strong']
 ['Overcast' 'Hot' 'Normal' 'Weak']]
```

```
tennis=pd.read_csv('/content/tennis.csv')
tennis.head()
tennis.tail()
X=tennis.iloc[:,1:5].values
y=tennis.iloc[:,-1].values
print(X)
print(y)
outlook1={'Sunny':1, 'Overcast':2, 'Rain':3}
tennis.outlook=tennis.outlook.map(outlook1)

temp1={'Hot':1, 'Mild':2, 'Cool':3}
tennis.temp=tennis.temp.map(temp1)

humid1={'Normal':1, 'High':2}
tennis.humidity=tennis.humidity.map(humid1)

wind1={'Weak':1, 'Strong':0}
tennis.wind=tennis.wind.map(wind1)

play1={'Yes':1, 'No':0}

tennis.play=tennis.play.map(play1)
x=tennis.iloc[:,1:5].values
y=tennis.iloc[:,-1].values
print(x)
print(y)
x_train,x_test,y_train,y_test=train_test_split(x,y,train_size=0.3,random_state=0)
print('xtrain = ',x_train)
print('xtest = ',x_test)
print('ytrain = ',y_train)
print('ytest = ',y_test)
gnb=GaussianNB()
gnb.fit(x_train,y_train)
y_pred=gnb.predict(x_test)
print('Confusion matrix = ',confusion_matrix(y_test,y_pred))
```

```
accuracy_nb=(accuracy_score(y_test,y_pred)*100,2)
acc=(gnb.score(x_train,y_train)*100,2)
accuracy=accuracy_score(y_pred,y_test)
precision=precision_score(y_test,y_pred)
from sklearn import metrics
print("Gaussian Naive Bayes model accuracy(in %):", acc[0])
print(accuracy_score(y_test,y_pred)*100)
print(recall_score(y_test,y_pred))
```

```
[['Sunny' 'Hot' 'High' 'Weak']
 ['Sunny' 'Hot' 'High' 'Strong']
 ['Overcast' 'Hot' 'High' 'Weak']
 ['Rain' 'Mild' 'High' 'Weak']
 ['Rain' 'Cool' 'Normal' 'Weak']
 ['Rain' 'Cool' 'Normal' 'Strong']
 ['Overcast' 'Cool' 'Normal' 'Strong']
 ['Sunny' 'Mild' 'High' 'Weak']
 ['Sunny' 'Cool' 'Normal' 'Weak']
 ['Rain' 'Mild' 'Normal' 'Weak']
 ['Sunny' 'Mild' 'Normal' 'Strong']
 ['Overcast' 'Mild' 'High' 'Strong']
 ['Overcast' 'Hot' 'Normal' 'Weak']
 ['Rain' 'Mild' 'High' 'Strong']]
['No' 'No' 'Yes' 'Yes' 'Yes' 'No' 'Yes' 'No' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes'
 'No']
[[1 1 2 1]
 [1 1 2 0]
 [2 1 2 1]
 [3 2 2 1]
 [3 3 1 1]
 [3 3 1 0]
 [2 3 1 0]
 [1 2 2 1]
 [1 3 1 1]
 [3 2 1 1]
 [1 2 1 0]
 [2 2 2 0]
 [2 1 1 1]
 [3 2 2 0]]
[0 0 1 1 1 0 1 0 1 1 1 1 1 0]
xtrain =  [[3 2 2 1]
 [1 1 2 1]
 [3 3 1 0]
 [2 1 1 1]]
xtest =  [[1 3 1 1]
 [2 3 1 0]
 [3 3 1 1]
 [2 2 2 0]
 [2 1 2 1]
 [3 2 2 0]
 [3 2 1 1]
 [1 1 2 0]
 [1 2 2 1]
 [1 2 1 0]]
ytrain =  [1 0 0 1]
```

```
ytest =  [1 1 1 1 1 0 1 0 0 1]
Confusion matrix =  [[2 1]
 [3 4]]
Gaussian Naive Bayes model accuracy(in %): 75.0
60.0
0.5714285714285714
```