naïve Bayesian classifier

In [1]:
```
import pandas as pd
from sklearn import tree
from sklearn.preprocessing import LabelEncoder
from sklearn.naive_bayes import GaussianNB
```

In [4]:
```
data = pd.read_csv('program4.csv')
print("The first 5 values of data is :\n",data.head())
```

```
The first 5 values of data is :
     Outlook Temperature Humidity   Windy PlayTennis
0      Sunny         Hot     High    Weak         No
1      Sunny         Hot     High  Strong         No
2   Overcast         Hot     High    Weak        Yes
3       Rain        Mild     High    Weak        Yes
4       Rain        Cool   Normal    Weak        Yes
```

In [5]:
```
x = data.iloc[:,:-1]
print("\nThe first 5 values of train data is\n",x.head())
```

```
The first 5 values of train data is
     Outlook Temperature Humidity   Windy
0      Sunny         Hot     High    Weak
1      Sunny         Hot     High  Strong
2   Overcast         Hot     High    Weak
3       Rain        Mild     High    Weak
4       Rain        Cool   Normal    Weak
```

In [6]:
```
y = data.iloc[:,-1]
print("\nThe first 5 values of train output is\n",y.head())
```

```
The first 5 values of train output is
 0     No
1     No
2    Yes
3    Yes
4    Yes
Name: PlayTennis, dtype: object
```

In [8]:
```
le_Outlook = LabelEncoder()
x.Outlook = le_Outlook.fit_transform(x.Outlook)

le_Temperature=LabelEncoder()
x.Temperature=le_Temperature.fit_transform(x.Temperature)

le_Humidity=LabelEncoder()
x.Humidity=le_Humidity.fit_transform(x.Humidity)

le_Windy=LabelEncoder()
x.Windy=le_Windy.fit_transform(x.Windy)

print("\nNow the train data is :\n",x.head())
```

```
Now the train data is :
    Outlook  Temperature  Humidity  Windy
0        2            1         0      1
1        2            1         0      0
2        0            1         0      1
```

```
3        1           2         0      1
4        1           0         1      1
```

In [9]:

```
le_PlayTennis=LabelEncoder()
y=le_PlayTennis.fit_transform(y)
print("\nNow the train output is \n",y)
```

```
Now the train output is
 [0 0 1 1 1 0 1 0 1 1 1 1 2]
```

In [10]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.20)

classifier=GaussianNB()
classifier.fit(x_train,y_train)

from sklearn.metrics import accuracy_score
print("Accuracy is:",accuracy_score(classifier.predict(x_test),y_test))
```

```
Accuracy is: 0.3333333333333333
```