

In []:

```
1 Program-5:
2 Write a program to implement the naïve Bayesian classifier for a sample training data
3 set stored as a .CSV file. Compute the accuracy of the classifier, considering few
4 test data sets.
```

In [1]:

```
1 # import necessary libarities
2 import pandas as pd
3 from sklearn import tree
4 from sklearn.preprocessing import LabelEncoder
5 from sklearn.naive_bayes import GaussianNB
6
7 # Load data from CSV
8 data = pd.read_csv('tennisdata.csv')
9 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	False	No
1	Sunny	Hot	High	True	No
2	Overcast	Hot	High	False	Yes
3	Rainy	Mild	High	False	Yes
4	Rainy	Cool	Normal	False	Yes

In [2]:

```
1 # obtain Train data and Train output
2 X = data.iloc[:, :-1]
3 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	False
1	Sunny	Hot	High	True
2	Overcast	Hot	High	False
3	Rainy	Mild	High	False
4	Rainy	Cool	Normal	False

In [3]:

```
1 y = data.iloc[:, -1]
2 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 [4]:

```
1 # Convert then in numbers
2 le_outlook = LabelEncoder()
3 X.Outlook = le_outlook.fit_transform(X.Outlook)
4
5 le_Temperature = LabelEncoder()
6 X.Temperature = le_Temperature.fit_transform(X.Temperature)
7
8 le_Humidity = LabelEncoder()
9 X.Humidity = le_Humidity.fit_transform(X.Humidity)
10
11 le_Windy = LabelEncoder()
12 X.Windy = le_Windy.fit_transform(X.Windy)
13
14 print("\nNow the Train data is :\n",X.head())
```

Now the Train data is :

	Outlook	Temperature	Humidity	Windy
0	2	1	0	0
1	2	1	0	1
2	0	1	0	0
3	1	2	0	0
4	1	0	1	0

In [5]:

```
1 le_PlayTennis = LabelEncoder()
2 y = le_PlayTennis.fit_transform(y)
3 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 1 0]
```

In [6]:

```
1 from sklearn.model_selection import train_test_split
2 X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20)
3
4 classifier = GaussianNB()
5 classifier.fit(X_train,y_train)
6
7 from sklearn.metrics import accuracy_score
8 print("Accuracy is:",accuracy_score(classifier.predict(X_test),y_test))
```

Accuracy is: 0.6666666666666666

In []:

```
1
```

