

```
data = pd.read_csv('dataset/tennisdata.csv')
data
```

	Outlook	Temperature	Humidity	Wind	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
5	Rain	Cool	Normal	Strong	No
6	Overcast	Cool	Normal	Strong	Yes
7	Sunny	Mild	High	Weak	No
8	Sunny	Cool	Normal	Weak	Yes
9	Rain	Mild	Normal	Weak	yes
10	Sunny	Mild	Normal	Strong	Yes
11	Overcast	Mild	High	Strong	Yes
12	Overcast	Hot	Normal	Weak.Yes	NaN
13	Rain	mild	High	Strong	no

The first 5 values of data is :

	Outlook	Temperature	Humidity	Wind	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

The First 5 values of train data is

	Outlook	Temperature	Humidity	Wind
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

The first 5 values of Train output is

```
0    No
1    No
2   Yes
3   Yes
4   Yes
Name: PlayTennis, dtype: object
```

Now the Train data is :

	Outlook	Temperature	Humidity	Wind
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

```
le_PlayTennis = LabelEncoder()
y = le_PlayTennis.fit_transform(y)

print("Now the Train output is\n")
y
```

Now the Train output is

```
array([0, 0, 1, 1, 1, 0, 1, 0, 1, 3, 1, 1, 4, 2])
```

```
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)
```

▼ GaussianNB

GaussianNB()

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

Accuracy is: 0.6666666666666666