Creado por:

Isabel Maniega

Naives Bayes

In [1]:
```
import warnings
warnings.filterwarnings("ignore")
```

In [2]:
```
from sklearn import datasets

wine = datasets.load_wine()
```

In [3]:
```
print("Features: ", wine.feature_names)
print("Labels: ", wine.target_names)
```

```
Features:  ['alcohol', 'malic_acid', 'ash', 'alcalinity_of_ash', 'magnesium', 'total_phenols', 'flavanoids', 'n
onflavanoid_phenols', 'proanthocyanins', 'color_intensity', 'hue', 'od280/od315_of_diluted_wines', 'proline']
Labels:  ['class_0' 'class_1' 'class_2']
```

In [4]:
```
wine.data.shape
```

Out[4]:
```
(178, 13)
```

In [6]:
```
import pandas as pd

df = pd.DataFrame(wine.data, columns=wine.feature_names)
df
```

Out[6]:

	alcohol	malic_acid	ash	alcalinity_of_ash	magnesium	total_phenols	flavanoids	nonflavanoid_phenols	proanthocyanins	color_intens
0	14.23	1.71	2.43	15.6	127.0	2.80	3.06	0.28	2.29	5
1	13.20	1.78	2.14	11.2	100.0	2.65	2.76	0.26	1.28	4
2	13.16	2.36	2.67	18.6	101.0	2.80	3.24	0.30	2.81	5
3	14.37	1.95	2.50	16.8	113.0	3.85	3.49	0.24	2.18	7
4	13.24	2.59	2.87	21.0	118.0	2.80	2.69	0.39	1.82	4
...	...	...	...	...	...	...	...	...	...	
173	13.71	5.65	2.45	20.5	95.0	1.68	0.61	0.52	1.06	7
174	13.40	3.91	2.48	23.0	102.0	1.80	0.75	0.43	1.41	7
175	13.27	4.28	2.26	20.0	120.0	1.59	0.69	0.43	1.35	10
176	13.17	2.59	2.37	20.0	120.0	1.65	0.68	0.53	1.46	9
177	14.13	4.10	2.74	24.5	96.0	2.05	0.76	0.56	1.35	9

178 rows × 13 columns

In [7]:
```
wine.target
```

Out[7]:
```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2])
```

In [8]:
```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(wine.data, wine.target, test_size=0.3, random_state=109)
```

In [12]:
```
from sklearn.naive_bayes import GaussianNB

gnb = GaussianNB()

gnb.fit(X_train, y_train)
```

Out[12]:
```
▼ GaussianNB
GaussianNB()
```

In [13]:
```
y_pred = gnb.predict(X_test)
y_pred
```

Out[13]:
```
array([0, 0, 1, 2, 0, 1, 0, 0, 1, 0, 2, 2, 2, 2, 0, 1, 1, 0, 0, 1, 2, 1,
       0, 2, 0, 0, 1, 2, 0, 1, 2, 1, 1, 0, 1, 1, 0, 2, 2, 0, 2, 1, 0, 0,
       0, 2, 2, 0, 1, 1, 2, 0, 0, 2])
```

In [14]:
```
y_test
```

Out[14]:
```
array([0, 0, 1, 2, 0, 1, 0, 1, 1, 0, 1, 1, 2, 2, 0, 1, 1, 0, 0, 1, 2, 1,
       0, 2, 0, 0, 1, 2, 0, 1, 2, 1, 1, 0, 1, 1, 0, 2, 2, 0, 2, 0, 0, 0,
       0, 2, 2, 0, 1, 1, 2, 1, 0, 2])
```

In [15]:
```
from sklearn import metrics

print("Accuracy: ", metrics.accuracy_score(y_test, y_pred))
```

```
Accuracy:  0.9074074074074074
```

Creado por:

Isabel Maniega