Naive Bayes

Getting Data

In [13]:
```python
import pandas
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression

data_frame = pandas.read_csv("https://raw.githubusercontent.com/whitehatjr/datasets/master/C120/income.csv")

x = data_frame[["age", "hours-per-week", "education-num", "capital-gain", "capital-loss"]]
y = data_frame["income"]
```

Train Test Split

In [14]:
```python
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25, random_state=42)
```

Standard Scaler

In [15]:
```python
stan_scal = StandardScaler()
x_train = stan_scal.fit_transform(x_train)
x_test = stan_scal.fit_transform(x_test)
```

Gaussian Naive Bayes

In [16]:
```python
gauss = GaussianNB()
gauss.fit(x_train, y_train)
```

Out[16]: GaussianNB()

Getting Accuracy Score

In [17]:
```python
y_pred = gauss.predict(x_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"The Accuracy is: {accuracy}")
```

The Accuracy is: 0.7896692021935255