```
import pandas as pd
import numpy as np
```

```
dataset=pd.read_csv('/content/drive/MyDrive/Personal/Studies/MSC Data Science Material/SEM2/ML/Practical/data_set/Fish.csv')
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
    Mounted at /content/drive
```

```
dataset
```

| | Species | Weight | Length1 | Length2 | Length3 | Height | Width |
|---|---|---|---|---|---|---|---|
| **0** | Bream | 242.0 | 23.2 | 25.4 | 30.0 | 11.5200 | 4.0200 |
| **1** | Bream | 290.0 | 24.0 | 26.3 | 31.2 | 12.4800 | 4.3056 |
| **2** | Bream | 340.0 | 23.9 | 26.5 | 31.1 | 12.3778 | 4.6961 |
| **3** | Bream | 363.0 | 26.3 | 29.0 | 33.5 | 12.7300 | 4.4555 |
| **4** | Bream | 430.0 | 26.5 | 29.0 | 34.0 | 12.4440 | 5.1340 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **154** | Smelt | 12.2 | 11.5 | 12.2 | 13.4 | 2.0904 | 1.3936 |
| **155** | Smelt | 13.4 | 11.7 | 12.4 | 13.5 | 2.4300 | 1.2690 |
| **156** | Smelt | 12.2 | 12.1 | 13.0 | 13.8 | 2.2770 | 1.2558 |
| **157** | Smelt | 19.7 | 13.2 | 14.3 | 15.2 | 2.8728 | 2.0672 |
| **158** | Smelt | 19.9 | 13.8 | 15.0 | 16.2 | 2.9322 | 1.8792 |

159 rows × 7 columns

```
print("The different species are: ", list(dataset['Species'].unique()))
```

```
    The different species are:  ['Bream', 'Roach', 'Whitefish', 'Parkki', 'Perch', 'Pike', 'Smelt']
```

```
print("The data for the species Bream and Perch are: ")
dataframe=pd.DataFrame(dataset[dataset['Species'].isin(['Bream','Perch'])])
dataframe.index=range(len(dataframe))
dataframe
```

```
    The data for the species Bream and Perch are:
```

| | Species | Weight | Length1 | Length2 | Length3 | Height | Width |
|---|---|---|---|---|---|---|---|
| **0** | Bream | 242.0 | 23.2 | 25.4 | 30.0 | 11.5200 | 4.0200 |
| **1** | Bream | 290.0 | 24.0 | 26.3 | 31.2 | 12.4800 | 4.3056 |
| **2** | Bream | 340.0 | 23.9 | 26.5 | 31.1 | 12.3778 | 4.6961 |
| **3** | Bream | 363.0 | 26.3 | 29.0 | 33.5 | 12.7300 | 4.4555 |
| **4** | Bream | 430.0 | 26.5 | 29.0 | 34.0 | 12.4440 | 5.1340 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **86** | Perch | 1100.0 | 39.0 | 42.0 | 44.6 | 12.8002 | 6.8684 |
| **87** | Perch | 1000.0 | 39.8 | 43.0 | 45.2 | 11.9328 | 7.2772 |
| **88** | Perch | 1100.0 | 40.1 | 43.0 | 45.5 | 12.5125 | 7.4165 |
| **89** | Perch | 1000.0 | 40.2 | 43.5 | 46.0 | 12.6040 | 8.1420 |
| **90** | Perch | 1000.0 | 41.1 | 44.0 | 46.6 | 12.4888 | 7.5958 |

91 rows × 7 columns

```
dataframe
```

| | Species | Weight | Length1 | Length2 | Length3 | Height | Width |
|---|---|---|---|---|---|---|---|
| 0 | Bream | 242.0 | 23.2 | 25.4 | 30.0 | 11.5200 | 4.0200 |
| 1 | Bream | 290.0 | 24.0 | 26.3 | 31.2 | 12.4800 | 4.3056 |
| 2 | Bream | 340.0 | 23.9 | 26.5 | 31.1 | 12.3778 | 4.6961 |
| 3 | Bream | 363.0 | 26.3 | 29.0 | 33.5 | 12.7300 | 4.4555 |
| 4 | Bream | 430.0 | 26.5 | 29.0 | 34.0 | 12.4440 | 5.1340 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 86 | Perch | 1100.0 | 39.0 | 42.0 | 44.6 | 12.8002 | 6.8684 |

```
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
```

```
     90    Perch  1000.0     41.1     44.0     46.6  12.4888  7.5958
```

```
scaler=StandardScaler()
dataframe.iloc[:,1:]=scaler.fit_transform(dataframe.iloc[:,1:])
xtrain,xtest,ytrain,ytest=train_test_split(dataframe.iloc[:,1:].values,dataframe['Species'].values, test_size=0.15)
```

```
ytest
```

```
array(['Perch', 'Bream', 'Perch', 'Perch', 'Bream', 'Perch', 'Perch',
       'Bream', 'Perch', 'Perch', 'Bream', 'Perch', 'Bream', 'Bream'],
      dtype=object)
```

```
xtrain
```

```
array([[-1.08637985, -1.15504774, -1.13614705, -1.17119417, -1.13963241,
        -0.90212373],
       [ 1.49019614,  1.42882797,  1.41730767,  1.5187712 ,  1.58779326,
         0.91694411],
       [ 0.08479106,  0.13689012,  0.10227849,  0.36432773,  0.81126219,
        -0.03269948],
       [-1.21130474, -1.31823989, -1.31488888, -1.36173339, -1.26648251,
        -1.32657289],
       [-0.69598955, -0.2846896 , -0.30627427, -0.45387007, -0.77608995,
        -0.29731564],
       [ 0.7874936 ,  0.58566853,  0.65127125,  0.89111261,  1.29895726,
         0.70340721],
       [ 0.6625687 ,  0.53127115,  0.52359852,  0.7005734 ,  1.21519825,
         0.24389574],
       [ 0.70941554,  0.59926788,  0.65127125,  0.84627986,  1.27163957,
         0.3910094 ],
       [ 0.3971033 ,  0.25928423,  0.26825305,  0.47640962,  1.08818813,
         0.38495257],
       [-1.10199546, -1.01905428, -1.00847432, -1.05911228, -1.04419194,
        -1.07709881],
       [ 0.08479106, -0.09429876, -0.02539425,  0.17378851,  0.80049513,
         0.18204878],
       [ 1.08419023,  1.30643386,  1.28963494,  1.07044364,  0.10452267,
         1.09158269],
       [-0.41490853,  0.27288358,  0.26825305,  0.48761781,  0.73955039,
         0.04361657],
       [-0.86151503, -0.54307717, -0.54885247, -0.66682567, -0.94109338,
        -0.51751505],
       [-1.0239174 , -0.92385886, -0.9191034 , -0.98065496, -1.08003641,
        -0.92736052],
       [ 1.5682742 ,  1.3472319 ,  1.41730767,  1.45152207,  1.81911352,
         1.17052337],
       [-0.47737098,  0.0416947 ,  0.01290757, -0.15124897, -0.7000119 ,
        -0.15908533],
       [-1.00830179, -0.95105755, -0.94463795, -1.00307134, -0.8882868 ,
        -0.93139841],
       [-0.91460811, -0.6110739 , -0.62545611, -0.72286661, -0.96886827,
        -0.89135603],
       [-0.57106465, -0.47508044, -0.45948155, -0.19608173,  0.41198956,
        -0.47269451],
       [ 0.006713  ,  0.12329077,  0.14058031,  0.36432773,  0.8195375 ,
         0.06474817],
       [-0.88962313, -0.66547129, -0.6765252 , -0.75649118, -0.90168275,
        -0.56677726],
       [ 0.55325942,  1.22483778,  1.1619622 ,  0.94715356,  0.10488843,
         0.66962356],
       [-1.07076423, -1.1142497 , -1.09784523, -1.13756961, -0.98155556,
        -0.9919667 ],
       [-0.41490853, -0.48867979, -0.433947  , -0.20728991,  0.38862665,
        -0.2098954 ],
       [-0.72097452, -0.58387521, -0.57438701, -0.33057999,  0.19253364,
        -0.6648979 ],
       [-0.07136506,  0.01449601,  0.01290757,  0.24103765,  0.76058158,
        -0.1104961 ],
       [-1.05514862, -1.01905428, -1.00847432, -1.05911228, -1.09791292,
        -0.99802353],
       [ 0.08479106,  0.21848619,  0.20441668,  0.38674411,  0.70441458,
        -0.43070049],
```

```
xtest
```

```
array([[-1.0239174 , -0.74706736, -0.75312884, -0.8349485 , -0.98361296,
        -0.79612921],
       [ 1.39650247,  1.0208476 ,  1.09812583,  1.24977466,  1.68229647,
         0.87373872],
       [-1.16445791, -1.53582942, -1.51916526, -1.54106441, -1.24709723,
        -1.13490788],
       [-0.64914271, -0.2846896 , -0.30627427, -0.45387007, -0.82894226,
        -0.45290887],
       [-0.13382751, -0.1350968 , -0.11476516,  0.11774757,  0.40375997,
         0.08480301],
       [ 0.67818432,  0.96645021,  0.90661673,  0.71178159, -0.02424767,
         0.91431948],
       [-1.45834373, -2.71897251, -2.74482353, -2.70671607, -1.95813443,
        -2.42272446],
       [-0.07136506, -0.09429876, -0.02539425,  0.19620489,  0.66857012,
        -0.05423488],
       [ 1.9586645 ,  1.71441423,  1.67265314,  1.40668931,  0.41941906,
         1.62088231],
       [ 1.08419023,  1.23843713,  1.1619622 ,  0.93594537,  0.40085675,
         1.57707124],
       [ 0.75313925,  0.70806264,  0.77894399,  0.95836175,  1.33484745,
         0.56766693],
       [ 1.64635226,  1.72801358,  1.73648951,  1.46273026,  0.44033595,
         2.10913008],
       [ 1.41211808,  1.18403975,  1.22579857,  1.38427293,  1.8462712 ,
         1.17213853],
       [ 0.70941554,  0.39527769,  0.39592578,  0.62211608,  0.97388817,
         0.12753731]])
```

```
ytrain
```

```
array(['Perch', 'Bream', 'Bream', 'Perch', 'Perch', 'Bream', 'Bream',
       'Bream', 'Bream', 'Perch', 'Bream', 'Perch', 'Bream', 'Perch',
       'Perch', 'Bream', 'Perch', 'Perch', 'Perch', 'Bream', 'Bream',
       'Perch', 'Perch', 'Perch', 'Bream', 'Bream', 'Bream', 'Perch',
       'Bream', 'Bream', 'Perch', 'Perch', 'Perch', 'Bream', 'Perch',
       'Bream', 'Perch', 'Bream', 'Perch', 'Perch', 'Bream', 'Bream',
       'Perch', 'Perch', 'Perch', 'Perch', 'Perch', 'Bream', 'Bream',
       'Perch', 'Bream', 'Perch', 'Perch', 'Perch', 'Perch', 'Perch',
       'Perch', 'Perch', 'Perch', 'Perch', 'Perch', 'Bream', 'Bream',
       'Bream', 'Perch', 'Perch', 'Perch', 'Perch', 'Perch', 'Perch',
       'Perch', 'Perch', 'Bream', 'Perch', 'Bream', 'Perch', 'Perch'],
      dtype=object)
```

```
from sklearn.naive_bayes import GaussianNB
```

```
NB_model=GaussianNB()
NB_model.fit(xtrain,ytrain)
```

```
▾ GaussianNB
GaussianNB()
```

```
print("Training R2 score: ",NB_model.score(xtrain,ytrain))
```

```
Training R2 score:  0.8311688311688312
```

```
ypred=NB_model.predict(xtest)
print("Predictions \n", ypred)
```

```
Predictions
 ['Perch' 'Bream' 'Perch' 'Perch' 'Bream' 'Bream' 'Perch' 'Bream' 'Perch'
 'Bream' 'Bream' 'Perch' 'Bream' 'Bream']
```

```
print("True Values \n", ytest)
```

```
True Values
 ['Perch' 'Bream' 'Perch' 'Perch' 'Bream' 'Perch' 'Perch' 'Bream' 'Perch'
 'Perch' 'Bream' 'Perch' 'Bream' 'Bream']
```

```
conf_mat=confusion_matrix(ytest, ypred)
print("Confusion Matrix \n", conf_mat)
print(classification_report(ytest,ypred))
print("Accuracy: ",(conf_mat[0][0]+conf_mat[1][1])/len(ytest))
```

```
Confusion Matrix
 [[6 0]
 [2 6]]
              precision    recall  f1-score   support

       Bream       0.75      1.00      0.86         6
       Perch       1.00      0.75      0.86         8

    accuracy                           0.86        14
   macro avg       0.88      0.88      0.86        14
weighted avg       0.89      0.86      0.86        14

Accuracy:  0.8571428571428571
```

```
probs=NB_model.predict_proba(xtest)
probs=probs[:,1]
```