```
import numpy as np
import pandas as pd
from sklearn.datasets import load_iris
from sklearn.metrics import confusion_matrix
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
```

```
iris = load_iris()
```

```
X = pd.DataFrame(iris.data, columns=iris.feature_names)
```

```
X
```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
...	...	...	...	...
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3
149	5.9	3.0	5.1	1.8

150 rows × 4 columns

```
X.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 4 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   sepal length (cm)  150 non-null    float64
 1   sepal width (cm)   150 non-null    float64
 2   petal length (cm)  150 non-null    float64
 3   petal width (cm)   150 non-null    float64
dtypes: float64(4)
memory usage: 4.8 KB
```

```
X = X[['sepal length (cm)', 'sepal width (cm)']]
```

```
X
```

	sepal length (cm)	sepal width (cm)
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1
4	5.0	3.6
...	...	...
145	6.7	3.0
146	6.3	2.5
147	6.5	3.0
148	6.2	3.4
149	5.9	3.0

150 rows × 2 columns

```
Y = (iris.target)
```

```
Y
```

```
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, 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, 2, 2, 2, 2])
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, random_state=1)
```

```
X_train
```

	sepal length (cm)	sepal width (cm)
54	6.5	2.8
108	6.7	2.5
112	6.8	3.0
17	5.1	3.5
119	6.0	2.2
...	...	...
133	6.3	2.8
137	6.4	3.1
72	6.3	2.5
140	6.7	3.1
37	4.9	3.6

112 rows × 2 columns

```
X_test
```

	sepal length (cm)	sepal width (cm)
14	5.8	4.0
98	5.1	2.5
75	6.6	3.0
16	5.4	3.9
131	7.9	3.8
56	6.3	3.3
141	6.9	3.1
44	5.1	3.8
29	4.7	3.2
120	6.9	3.2
94	5.6	2.7
5	5.4	3.9
102	7.1	3.0
51	6.4	3.2
78	6.0	2.9
42	4.4	3.2
92	5.8	2.6
66	5.6	3.0
31	5.4	3.4
35	5.0	3.2
90	5.5	2.6
84	5.4	3.0
77	6.7	3.0
40	5.0	3.5
125	7.2	3.2
99	5.7	2.8
33	5.5	4.2
19	5.1	3.8
73	6.1	2.8
146	6.3	2.5
91	6.1	3.0

```
Y_train
```

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

53	5.5	2.3

```
Y_test
```

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

```
knn = KNeighborsClassifier(n_neighbors=5, metric='euclidean')
knn.fit(X_train, Y_train)
```

```
KNeighborsClassifier
KNeighborsClassifier(metric='euclidean')
```

```
Y_pred = knn.predict(X_test)
```

```
Y_pred
```

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

```
from sklearn.metrics import classification_report, confusion_matrix
cm=np.array(confusion_matrix(Y_test,Y_pred))
confusion=pd.DataFrame(cm, index=['é 0','é 1','é 2'], columns=['previu 0','previu 1',' previu 2
confusion
```

	previu 0	previu 1	previu 2