In [1]:
```
import numpy as plt
import pandas as pd
```

In [2]:
```
df=pd.read_csv('iris.csv')
```

In [3]:
```
df
```

Out[3]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [4]:
```
df.isnull().sum()
```

Out[4]:
```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
species         0
dtype: int64
```

In [5]:
```
df['species'].value_counts()
```

Out[5]:
```
Iris-setosa        50
Iris-versicolor    50
Iris-virginica     50
Name: species, dtype: int64
```

In [7]:
```
len(df['species'].value_counts())
```

Out[7]:
```
3
```

In [8]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   sepal_length  150 non-null    float64
 1   sepal_width   150 non-null    float64
 2   petal_length  150 non-null    float64
 3   petal_width   150 non-null    float64
 4   species       150 non-null    object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

In [9]:
```
X1 = df.drop(['species'], axis = 1)
Y1 = df['species']
```

In [11]:
```
X1.head(2)
```

Out[11]:

	sepal_length	sepal_width	petal_length	petal_width
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2

In [13]:
```
from sklearn.model_selection import train_test_split
```

In [14]:
```
X_train, X_test, Y_train, Y_test = train_test_split(X1,Y1,test_size=0.2,random_sta
```

In [15]:
```
X_train.shape, X_test.shape
```

Out[15]:
```
((120, 4), (30, 4))
```

In [16]:
```
from sklearn.neighbors import KNeighborsClassifier
```

In [18]:
```
knn = KNeighborsClassifier(n_neighbors=1)
knn = KNeighborsClassifier(n_neighbors=2)
knn = KNeighborsClassifier(n_neighbors=3)
knn = KNeighborsClassifier(n_neighbors=4)
knn = KNeighborsClassifier(n_neighbors=5)
```

In [21]:
```
knn = KNeighborsClassifier(n_neighbors=2)
knn.fit(X_train, Y_train)
```

Out[21]:
```
KNeighborsClassifier(n_neighbors=2)
```

In [22]:
```
X_test
```

Out[22]:

	sepal_length	sepal_width	petal_length	petal_width
10	5.4	3.7	1.5	0.2
115	6.4	3.2	5.3	2.3
54	6.5	2.8	4.6	1.5
146	6.3	2.5	5.0	1.9
63	6.1	2.9	4.7	1.4
76	6.8	2.8	4.8	1.4
86	6.7	3.1	4.7	1.5
138	6.0	3.0	4.8	1.8
64	5.6	2.9	3.6	1.3
35	5.0	3.2	1.2	0.2
120	6.9	3.2	5.7	2.3
95	5.7	3.0	4.2	1.2
130	7.4	2.8	6.1	1.9
109	7.2	3.6	6.1	2.5
43	5.0	3.5	1.6	0.6
131	7.9	3.8	6.4	2.0
69	5.6	2.5	3.9	1.1
99	5.7	2.8	4.1	1.3
119	6.0	2.2	5.0	1.5
96	5.7	2.9	4.2	1.3
46	5.1	3.8	1.6	0.2
141	6.9	3.1	5.1	2.3
36	5.5	3.5	1.3	0.2
92	5.8	2.6	4.0	1.2
142	5.8	2.7	5.1	1.9
2	4.7	3.2	1.3	0.2
129	7.2	3.0	5.8	1.6
147	6.5	3.0	5.2	2.0
124	6.7	3.3	5.7	2.1
126	6.2	2.8	4.8	1.8

In [23]:
```
y_predict1 = knn.predict(X_test)
y_predict1
```

Out[23]:
```
array(['Iris-setosa', 'Iris-virginica', 'Iris-versicolor',
       'Iris-virginica', 'Iris-versicolor', 'Iris-versicolor',
       'Iris-versicolor', 'Iris-versicolor', 'Iris-versicolor',
       'Iris-setosa', 'Iris-virginica', 'Iris-versicolor',
       'Iris-virginica', 'Iris-virginica', 'Iris-setosa',
       'Iris-virginica', 'Iris-versicolor', 'Iris-versicolor',
       'Iris-versicolor', 'Iris-versicolor', 'Iris-setosa',
       'Iris-virginica', 'Iris-setosa', 'Iris-versicolor',
       'Iris-virginica', 'Iris-setosa', 'Iris-virginica',
       'Iris-virginica', 'Iris-virginica', 'Iris-virginica'], dtype=object)
```

In [24]:
```
X_test.head(2)
```

Out[24]:

	sepal_length	sepal_width	petal_length	petal_width
10	5.4	3.7	1.5	0.2
115	6.4	3.2	5.3	2.3

In [35]:
```
knn.predict([[8,4,7,2]])
```

```
C:\Users\hi\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does n
ot have valid feature names, but KNeighborsClassifier was fitted with feature name
s
  warnings.warn(
```

Out[35]:
```
array(['Iris-virginica'], dtype=object)
```

In [29]:
```
Y_test[:2]
```

Out[29]:
```
10        Iris-setosa
115    Iris-virginica
Name: species, dtype: object
```

In [30]:
```
df.tail()
```

Out[30]:

	sepal_length	sepal_width	petal_length	petal_width	species
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

In [ ]:
```

```