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

```
In [2]:  df=pd.read_csv("f:/dataset/classification/fruits.csv")
         df
```

Out[2]:

	diameter	weight	FruitName
0	3.0	30	Banana
1	6.0	100	Apple
2	6.1	95	Apple
3	3.2	35	Banana
4	5.5	80	Apple
5	7.1	120	Banana
6	2.5	60	Banana
7	2.3	100	Banana
8	4.8	70	Apple
9	4.8	79	Apple
10	5.8	120	Apple
11	2.6	85	Banana
12	6.0	110	Apple
13	6.3	95	Apple
14	3.0	40	Banana
15	3.5	25	Banana
16	5.5	100	Apple
17	7.5	120	Apple
18	2.5	50	Banana
19	2.7	40	Banana
20	4.8	90	Apple
21	5.8	90	Apple

```
In [4]:  df_features=df.iloc[:,:-1]
```

```
In [7]:  X=df_features.values
```

```
In [9]:  series_target=df.iloc[:,-1]
```

```
In [11]:  y=series_target.values
```

```
In [13]:  from sklearn.neighbors import KNeighborsClassifier
```

```
In [14]:  model=KNeighborsClassifier()    #untrained
```

```
In [15]:  model.fit(X,y)    #model training
```

```
Out[15]:   ▾ KNeighborsClassifier
           KNeighborsClassifier()
```

```
In [19]:  new_sample1=[3,40]
          new_sample2=[4,60]
```

```
In [17]:  preds=model.predict([new_sample1,new_sample2])
```

```
In [18]:  preds
```

```
Out[18]:  array(['Banana', 'Banana'], dtype=object)
```

General steps in ML Coding

- collect & load data
- prepare data
- select an algo
- create model
- fit(train) the model
- make prediction

```
In [20]:  from sklearn.neighbors import KNeighborsClassifier
          df=pd.read_csv("f:/dataset/classification/fruits.csv")
          X=df.iloc[:,:-1].values
          y=df.iloc[:,-1].values
          model=KNeighborsClassifier()
          model.fit(X,y)
```

```
Out[20]:   ▾ KNeighborsClassifier
           KNeighborsClassifier()
```

```
In [22]:  model.predict([[5,80]])
```

```
Out[22]:  array(['Apple'], dtype=object)
```

```
In [ ]:
```