```
from sklearn import datasets
iris=datasets.load_iris()
```

```
iris
```

```
       [6.3, 2.8, 5.1, 1.5],
       [6.1, 2.6, 5.6, 1.4],
       [7.7, 3. , 6.1, 2.3],
       [6.3, 3.4, 5.6, 2.4],
       [6.4, 3.1, 5.5, 1.8],
       [6. , 3. , 4.8, 1.8],
       [6.9, 3.1, 5.4, 2.1],
       [6.7, 3.1, 5.6, 2.4],
       [6.9, 3.1, 5.1, 2.3],
       [5.8, 2.7, 5.1, 1.9],
       [6.8, 3.2, 5.9, 2.3],
       [6.7, 3.3, 5.7, 2.5],
       [6.7, 3. , 5.2, 2.3],
       [6.3, 2.5, 5. , 1.9],
       [6.5, 3. , 5.2, 2. ],
       [6.2, 3.4, 5.4, 2.3],
       [5.9, 3. , 5.1, 1.8]]),
 'target': 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]),
 'frame': None,
 'target_names': array(['setosa', 'versicolor', 'virginica'], dtype='<U10'),
 'DESCR': '.. _iris_dataset:\n\nIris plants dataset\n--------------------\n\n**Data Set
Characteristics:**\n\n:Number of Instances: 150 (50 in each of three classes)\n:Number of Attributes: 4
numeric, predictive attributes and the class\n:Attribute Information:\n    - sepal length in cm\n    - sepal
width in cm\n    - petal length in cm\n    - petal width in cm\n    - class:\n            - Iris-Setosa\n
- Iris-Versicolour\n            - Iris-Virginica\n\n:Summary Statistics:\n\n============== ==== ==== =======
===== ====================\n                Min  Max   Mean    SD   Class Correlation\n============== ==== ====
======= ===== ====================\nsepal length:   4.3  7.9   5.84   0.83    0.7826\nsepal width:    2.0  4.4
3.05   0.43   -0.4194\npetal length:   1.0  6.9   3.76   1.76    0.9490  (high!)\npetal width:    0.1  2.5
1.20   0.76    0.9565  (high!)\n============== ==== ==== ======= ===== ====================\n\n:Missing
Attribute Values: None\n:Class Distribution: 33.3% for each of 3 classes.\n:Creator: R.A. Fisher\n:Donor:
Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)\n:Date: July, 1988\n\nThe famous Iris database, first used by
Sir R.A. Fisher. The dataset is taken\nfrom Fisher\'s paper. Note that it\'s the same as in R, but not as in
the UCI\nMachine Learning Repository, which has two wrong data points.\n\nThis is perhaps the best known
database to be found in the\npattern recognition literature.  Fisher\'s paper is a classic in the field and\nis
referenced frequently to this day.  (See Duda & Hart, for example.)  The\ndata set contains 3 classes of 50
instances each, where each class refers to a\ntype of iris plant.  One class is linearly separable from the
other 2; the\nlatter are NOT linearly separable from each other.\n\n.. dropdown:: References\n\n  - Fisher,
R.A. "The use of multiple measurements in taxonomic problems"\n    Annual Eugenics, 7, Part II, 179-188 (1936);
also in "Contributions to\n    Mathematical Statistics" (John Wiley, NY, 1950).\n  - Duda, R.O., & Hart, P.E.
(1973) Pattern Classification and Scene Analysis.\n    (Q327.D83) John Wiley & Sons.  ISBN 0-471-22361-1.  See
page 218.\n  - Dasarathy, B.V. (1980) "Nosing Around the Neighborhood: A New System\n    Structure and
Classification Rule for Recognition in Partially Exposed\n    Environments".  IEEE Transactions on Pattern
Analysis and Machine\n    Intelligence, Vol. PAMI-2, No. 1, 67-71.\n  - Gates, G.W. (1972) "The Reduced Nearest
Neighbor Rule".  IEEE Transactions\n    on Information Theory, May 1972, 431-433.\n  - See also: 1988 MLC
Proceedings, 54-64.  Cheeseman et al"s AUTOCLASS II\n    conceptual clustering system finds 3 classes in the
data.\n  - Many, many more ...\n',
 'feature_names': ['sepal length (cm)',
  'sepal width (cm)',
  'petal length (cm)',
  'petal width (cm)'],
 'filename': 'iris.csv',
 'data_module': 'sklearn.datasets.data'}
```

```
import pandas as pd
data=pd.DataFrame({"sepal length": iris.data[:,0],
                   "sepal width": iris.data[:,1],
                   "petal length": iris.data[:,2],
                   "petal width": iris.data[:,3],
                   "species": iris.target})
```

```
data
```

	sepal length	sepal width	petal length	petal width	species
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

```
#model traing: train test splitting
from sklearn.model_selection import train_test_split
x=data[["sepal length","sepal width","petal length","petal width"]]
y=data["species"]
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.20,random_state=shuffle=tate=42)
```

148	6.2	3.4	5.4	2.3	2
149	5.9	3.0	5.1	1.8	2

```
  File "/tmp/ipython-input-3996754636.py", line 5
    x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.20,random_state=-state=42)
                                                                                   ^
SyntaxError: invalid syntax
```

```
from sklearn.ensemble import RandomForestClassifier
model=RandomForestClassifier(n_estimators=100)
```

```
model.fit(x_train,y_train)
```

```
▾ RandomForestClassifier
RandomForestClassifier()
```

```
model.predict(x_test)
```

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

```
iris.target_names[0]
```

```
np.str_('setosa')
```

```
from sklearn import metrics
metrics.accuracy_score(y_pred,y_test)
```

```
1.0
```

```
y_pred = model.predict(x_test)
```