In [1]:

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

In [4]:

```
iris = pd.read_csv("F:\\DSE\\3rd year engineering\\5th sem\\6th sem\\DSBDA\\dataset\\iri
```

In [5]:

```
iris
```

Out[5]:

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

150 rows × 5 columns

In [6]:

```
iris.head()
```

Out[6]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

In [7]:

```
iris.describe()
```

Out[7]:

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [8]:

```
iris.dtypes
```

Out[8]:

```
sepal_length    float64
sepal_width     float64
petal_length    float64
petal_width     float64
species          object
dtype: object
```

In [9]:

```
iris.ndim
```

Out[9]:

```
2
```

In [10]:

```
iris.size
```

Out[10]:

```
750
```

In [11]:

```
iris.shape
```

Out[11]:

```
(150, 5)
```

In [12]:

```
from sklearn import preprocessing
```

In [13]:

```
iris['species'].unique()
```

Out[13]:

```
array(['setosa', 'versicolor', 'virginica'], dtype=object)
```

In [16]:

```
le = preprocessing.LabelEncoder()
```

In [17]:

```
iris['species']=le.fit_transform(iris['species'])
```

In [18]:

```
iris['species'].unique()
```

Out[18]:

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

In [ ]:

```

```