```
import pandas as pd
from sklearn import preprocessing
df = pd.read_csv('iris.csv')

print(df.head(10))
```

```
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1            5.1           3.5            1.4           0.2  Iris-setosa
1   2            4.9           3.0            1.4           0.2  Iris-setosa
2   3            4.7           3.2            1.3           0.2  Iris-setosa
3   4            4.6           3.1            1.5           0.2  Iris-setosa
4   5            5.0           3.6            1.4           0.2  Iris-setosa
5   6            5.4           3.9            1.7           0.4  Iris-setosa
6   7            4.6           3.4            1.4           0.3  Iris-setosa
7   8            5.0           3.4            1.5           0.2  Iris-setosa
8   9            4.4           2.9            1.4           0.2  Iris-setosa
9  10            4.9           3.1            1.5           0.1  Iris-setosa
```

```
print(df.isna().sum())
```

```
Id               0
SepalLengthCm    0
SepalWidthCm     0
PetalLengthCm    0
PetalWidthCm     0
Species          0
dtype: int64
```

```
print(df.dtypes)
```

```
Id                int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species           object
dtype: object
```

```
df['PetalLengthCm']= df['PetalLengthCm'].astype("int")
print(df.dtypes)
```

```
Id                int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm      int32
PetalWidthCm     float64
Species           object
dtype: object
```

```
x = df[['SepalLengthCm']].values.astype(float)
min_max_scaler = preprocessing.MinMaxScaler()
x_scaled = min_max_scaler.fit_transform(x)
df_normalized = pd.DataFrame(x_scaled)
print(df_normalized.head(10))
```

```
          0
0  0.222222
1  0.166667
2  0.111111
3  0.083333
4  0.194444
5  0.305556
6  0.083333
7  0.194444
8  0.027778
9  0.166667
```

```
print(df['Species'].unique())
```

```
['Iris-setosa' 'Iris-versicolor' 'Iris-virginica']
```

```
label_encoder = preprocessing.LabelEncoder()
df['Species']= label_encoder.fit_transform(df['Species'])
print(df)
print(df['Species'].unique())
```

```
      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  Species
0      1            5.1           3.5              1           0.2        0
1      2            4.9           3.0              1           0.2        0
2      3            4.7           3.2              1           0.2        0
3      4            4.6           3.1              1           0.2        0
4      5            5.0           3.6              1           0.2        0
..   ...            ...           ...            ...           ...      ...
145  146            6.7           3.0              5           2.3        2
146  147            6.3           2.5              5           1.9        2
147  148            6.5           3.0              5           2.0        2
148  149            6.2           3.4              5           2.3        2
149  150            5.9           3.0              5           1.8        2

[150 rows x 6 columns]
[0 1 2]
```

```
features_df=df.drop(columns=['Species'])
enc = preprocessing.OneHotEncoder()
enc_df=pd.DataFrame(enc.fit_transform(df[['Species']]))
df_encode = features_df.join(enc_df)
df_encode.rename(columns = {0:'Iris-Setosa',
1:'Iris-Versicolor',2:'Iris-virginica'}, inplace = True)
print(df_encode.tail(40))
```

```
      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
110  111            6.5           3.2              5           2.0
111  112            6.4           2.7              5           1.9
112  113            6.8           3.0              5           2.1
113  114            5.7           2.5              5           2.0
114  115            5.8           2.8              5           2.4
115  116            6.4           3.2              5           2.3
116  117            6.5           3.0              5           1.8
117  118            7.7           3.8              6           2.2
118  119            7.7           2.6              6           2.3
119  120            6.0           2.2              5           1.5
120  121            6.9           3.2              5           2.3
121  122            5.6           2.8              4           2.0
122  123            7.7           2.8              6           2.0
123  124            6.3           2.7              4           1.8
123  124            6.3           2.7              4           1.8
124  125            6.7           3.3              5           2.1
125  126            7.2           3.2              6           1.8
126  127            6.2           2.8              4           1.8
127  128            6.1           3.0              4           1.8
128  129            6.4           2.8              5           2.1
129  130            7.2           3.0              5           1.6
130  131            7.4           2.8              6           1.9
131  132            7.9           3.8              6           2.0
132  133            6.4           2.8              5           2.2
133  134            6.3           2.8              5           1.5
134  135            6.1           2.6              5           1.4
135  136            7.7           3.0              6           2.3
136  137            6.3           3.4              5           2.4
137  138            6.4           3.1              5           1.8
138  139            6.0           3.0              4           1.8
139  140            6.9           3.1              5           2.1
140  141            6.7           3.1              5           2.4
141  142            6.9           3.1              5           2.3
142  143            5.8           2.7              5           1.9
143  144            6.8           3.2              5           2.3
144  145            6.7           3.3              5           2.5
145  146            6.7           3.0              5           2.3
146  147            6.3           2.5              5           1.9
147  148            6.5           3.0              5           2.0
148  149            6.2           3.4              5           2.3
149  150            5.9           3.0              5           1.8
```

.

	Iris-Setosa
110	(0, 2)\t1.0
111	(0, 2)\t1.0
112	(0, 2)\t1.0
113	(0, 2)\t1.0
114	(0, 2)\t1.0
115	(0, 2)\t1.0
116	(0, 2)\t1.0
117	(0, 2)\t1.0
118	(0, 2)\t1.0
119	(0, 2)\t1.0
120	(0, 2)\t1.0
121	(0, 2)\t1.0
122	(0, 2)\t1.0
123	(0, 2)\t1.0
124	(0, 2)\t1.0
125	(0, 2)\t1.0
126	(0, 2)\t1.0
127	(0, 2)\t1.0
128	(0, 2)\t1.0
129	(0, 2)\t1.0
130	(0, 2)\t1.0
131	(0, 2)\t1.0
132	(0, 2)\t1.0
133	(0, 2)\t1.0
134	(0, 2)\t1.0
135	(0, 2)\t1.0
136	(0, 2)\t1.0
137	(0, 2)\t1.0
138	(0, 2)\t1.0
139	(0, 2)\t1.0
140	(0, 2)\t1.0
141	(0, 2)\t1.0
142	(0, 2)\t1.0
143	(0, 2)\t1.0
144	(0, 2)\t1.0
145	(0, 2)\t1.0
146	(0, 2)\t1.0
147	(0, 2)\t1.0
148	(0, 2)\t1.0
149	(0, 2)\t1.0

```
one_hot_df = pd.get_dummies(df, prefix="Species",
columns=['Species'], drop_first=False)
print(one_hot_df.head(20),one_hot_df.tail(20))
```

```
    Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  Species_0  \
0    1            5.1           3.5              1           0.2          1
1    2            4.9           3.0              1           0.2          1
2    3            4.7           3.2              1           0.2          1
3    4            4.6           3.1              1           0.2          1
4    5            5.0           3.6              1           0.2          1
5    6            5.4           3.9              1           0.4          1
6    7            4.6           3.4              1           0.3          1
7    8            5.0           3.4              1           0.2          1
8    9            4.4           2.9              1           0.2          1
9   10            4.9           3.1              1           0.1          1
10  11            5.4           3.7              1           0.2          1
11  12            4.8           3.4              1           0.2          1
12  13            4.8           3.0              1           0.1          1
13  14            4.3           3.0              1           0.1          1
14  15            5.8           4.0              1           0.2          1
15  16            5.7           4.4              1           0.4          1
16  17            5.4           3.9              1           0.4          1
17  18            5.1           3.5              1           0.3          1
18  19            5.7           3.8              1           0.3          1
19  20            5.1           3.8              1           0.3          1

    Species_1  Species_2
0           0          0
1           0          0
2           0          0
3           0          0
4           0          0
5           0          0
6           0          0
7           0          0
8           0          0
9           0          0
10          0          0
11          0          0
12          0          0
13          0          0
14          0          0
15          0          0
16          0          0
17          0          0
18          0          0
19          0          0        Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  Species_0  \
130  131            7.4           2.8              6           1.9          0
131  132            7.9           3.8              6           2.0          0
132  133            6.4           2.8              5           2.2          0
133  134            6.3           2.8              5           1.5          0
134  135            6.1           2.6              5           1.4          0
```

.

135	136	7.7	3.0	6	2.3	0
136	137	6.3	3.4	5	2.4	0
137	138	6.4	3.1	5	1.8	0
138	139	6.0	3.0	4	1.8	0
139	140	6.9	3.1	5	2.1	0
140	141	6.7	3.1	5	2.4	0
141	142	6.9	3.1	5	2.3	0
142	143	5.8	2.7	5	1.9	0
143	144	6.8	3.2	5	2.3	0
144	145	6.7	3.3	5	2.5	0
145	146	6.7	3.0	5	2.3	0
146	147	6.3	2.5	5	1.9	0
147	148	6.5	3.0	5	2.0	0
148	149	6.2	3.4	5	2.3	0
149	150	5.9	3.0	5	1.8	0

	Species_1	Species_2
130	0	1
131	0	1
132	0	1
133	0	1
134	0	1
135	0	1
136	0	1
137	0	1
138	0	1
139	0	1
140	0	1
141	0	1
142	0	1
143	0	1
144	0	1
145	0	1
146	0	1
147	0	1
148	0	1
149	0	1