TASK 1 - IRIS FLOWER CLASSIFICATION

Data Science Internship

Batch - February Phase 2 OIBSIP.

Author - **HIRAL CHOKSI**

1) Importing Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

2) Importing Dataset

```
irisdataset=pd.read_csv("Iris.csv")
```

3)Inspecting the Data

```
irisdataset
```

	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
...	...	...	...	...	...	...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

```
irisdataset.shape
```

```
(150, 6)
```

```
irisdataset.size
```

```
900
```

```
irisdataset.head()
```

	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

```
irisdataset.tail()
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica

```
irisdataset.Species.value_counts()
```

```
Iris-setosa        50
Iris-versicolor    50
Iris-virginica     50
Name: Species, dtype: int64
```

```
irisdataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Id             150 non-null    int64
 1   SepalLengthCm  150 non-null    float64
 2   SepalWidthCm   150 non-null    float64
 3   PetalLengthCm  150 non-null    float64
 4   PetalWidthCm   150 non-null    float64
 5   Species        150 non-null    object
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB
```

```
irisdataset.describe()
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

4) Data Cleaning

a) Dropping ID Column

```
irisdataset = irisdataset.drop('Id', axis = 1)
irisdataset
```

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	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

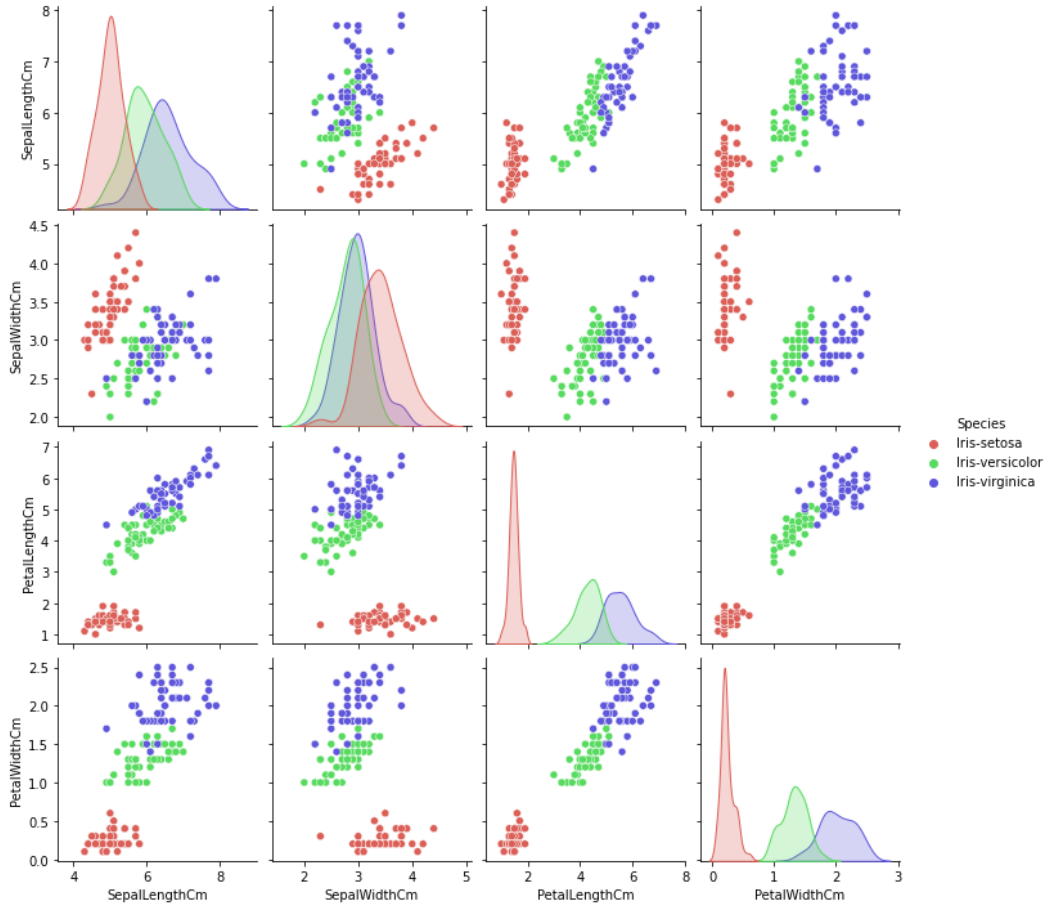
```
irisdataset.isnull().sum()
```

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

5) Visualization of the Dataset

```
sns.pairplot(irisdataset, hue ="Species" , palette = 'hls')
```

```
<seaborn.axisgrid.PairGrid at 0x7f7e89073b50>
```



6) Seperating INPUT COLUMNS and OUTPUT COLUMS

```
X = irisdataset.drop('Species', axis=1)
y = irisdataset['Species']
```

```
print(X)
```

```
     SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
0              5.1           3.5            1.4           0.2
1              4.9           3.0            1.4           0.2
2              4.7           3.2            1.3           0.2
3              4.6           3.1            1.5           0.2
4              5.0           3.6            1.4           0.2
..             ...           ...            ...           ...
145            6.7           3.0            5.2           2.3
146            6.3           2.5            5.0           1.9
147            6.5           3.0            5.2           2.0
148            6.2           3.4            5.4           2.3
149            5.9           3.0            5.1           1.8

[150 rows x 4 columns]
```

```
print(y)
```

```
0         Iris-setosa
1         Iris-setosa
2         Iris-setosa
3         Iris-setosa
4         Iris-setosa
            ...
145    Iris-virginica
146    Iris-virginica
147    Iris-virginica
148    Iris-virginica
149    Iris-virginica
Name: Species, Length: 150, dtype: object
```

a) Seperate the features and target

```
data = irisdataset.values

X = data[ :,0:4]
y = data[:,4]
print(X)
```

```
[[5.1 3.5 1.4 0.2]
 [4.9 3.0 1.4 0.2]
 [4.7 3.2 1.3 0.2]
 [4.6 3.1 1.5 0.2]
 [5.0 3.6 1.4 0.2]
 [5.4 3.9 1.7 0.4]
 [4.6 3.4 1.4 0.3]
 [5.0 3.4 1.5 0.2]
 [4.4 2.9 1.4 0.2]
 [4.9 3.1 1.5 0.1]
 [5.4 3.7 1.5 0.2]
 [4.8 3.4 1.6 0.2]
 [4.8 3.0 1.4 0.1]
 [4.3 3.0 1.1 0.1]
 [5.8 4.0 1.2 0.2]
 [5.7 4.4 1.5 0.4]
 [5.4 3.9 1.3 0.4]
 [5.1 3.5 1.4 0.3]
 [5.7 3.8 1.7 0.3]
 [5.1 3.8 1.5 0.3]
 [5.4 3.4 1.7 0.2]
 [5.1 3.7 1.5 0.4]
 [4.6 3.6 1.0 0.2]
 [5.1 3.3 1.7 0.5]
 [4.8 3.4 1.9 0.2]
 [5.0 3.0 1.6 0.2]
 [5.0 3.4 1.6 0.4]
 [5.2 3.5 1.5 0.2]
 [5.2 3.4 1.4 0.2]
 [4.7 3.2 1.6 0.2]
 [4.8 3.1 1.6 0.2]
 [5.4 3.4 1.5 0.4]
 [5.2 4.1 1.5 0.1]
 [5.5 4.2 1.4 0.2]
 [4.9 3.1 1.5 0.1]
 [5.0 3.2 1.2 0.2]
 [5.5 3.5 1.3 0.2]
 [4.9 3.1 1.5 0.1]
 [4.4 3.0 1.3 0.2]
 [5.1 3.4 1.5 0.2]
 [5.0 3.5 1.3 0.3]
 [4.5 2.3 1.3 0.3]
 [4.4 3.2 1.3 0.2]
 [5.0 3.5 1.6 0.6]
 [5.1 3.8 1.9 0.4]
 [4.8 3.0 1.4 0.3]
 [5.1 3.8 1.6 0.2]
 [4.6 3.2 1.4 0.2]
 [5.3 3.7 1.5 0.2]
 [5.0 3.3 1.4 0.2]
 [7.0 3.2 4.7 1.4]
 [6.4 3.2 4.5 1.5]
 [6.9 3.1 4.9 1.5]
 [5.5 2.3 4.0 1.3]
 [6.5 2.8 4.6 1.5]
 [5.7 2.8 4.5 1.3]
 [6.3 3.3 4.7 1.6]
 [4.9 2.4 3.3 1.0]
```

```
print(y)
```

```
['Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
```

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

6) Splitting the data into Training and Testing

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
print(X_train)
```

```
[[6.4 3.1 5.5 1.8]
 [5.4 3.0 4.5 1.5]
 [5.2 3.5 1.5 0.2]
 [6.1 3.0 4.9 1.8]
 [6.4 2.8 5.6 2.2]
 [5.2 2.7 3.9 1.4]
 [5.7 3.8 1.7 0.3]
 [6.0 2.7 5.1 1.6]
 [5.9 3.0 4.2 1.5]
 [5.8 2.6 4.0 1.2]
 [6.8 3.0 5.5 2.1]
 [4.7 3.2 1.3 0.2]
 [6.9 3.1 5.1 2.3]
 [5.0 3.5 1.6 0.6]
 [5.4 3.7 1.5 0.2]
 [5.0 2.0 3.5 1.0]
 [6.5 3.0 5.5 1.8]
 [6.7 3.3 5.7 2.5]
 [6.0 2.2 5.0 1.5]
 [6.7 2.5 5.8 1.8]
 [5.6 2.5 3.9 1.1]
 [7.7 3.0 6.1 2.3]
 [6.3 3.3 4.7 1.6]
 [5.5 2.4 3.8 1.1]
 [6.3 2.7 4.9 1.8]
 [6.3 2.8 5.1 1.5]
 [4.9 2.5 4.5 1.7]
 [6.3 2.5 5.0 1.9]
 [7.0 3.2 4.7 1.4]
 [6.5 3.0 5.2 2.0]
 [6.0 3.4 4.5 1.6]
 [4.8 3.1 1.6 0.2]
 [5.8 2.7 5.1 1.9]
 [5.6 2.7 4.2 1.3]
 [5.6 2.9 3.6 1.3]
 [5.5 2.5 4.0 1.3]
 [6.1 3.0 4.6 1.4]
 [7.2 3.2 6.0 1.8]
 [5.3 3.7 1.5 0.2]
 [4.3 3.0 1.1 0.1]
 [6.4 2.7 5.3 1.9]
 [5.7 3.0 4.2 1.2]
 [5.4 3.4 1.7 0.2]
 [5.7 4.4 1.5 0.4]
 [6.9 3.1 4.9 1.5]
 [4.6 3.1 1.5 0.2]
 [5.9 3.0 5.1 1.8]
 [5.1 2.5 3.0 1.1]
 [4.6 3.4 1.4 0.3]
```

```
       [6.2 2.2 4.5 1.5]
       [7.2 3.6 6.1 2.5]
       [5.7 2.9 4.2 1.3]
       [4.8 3.0 1.4 0.1]
       [7.1 3.0 5.9 2.1]
       [6.9 3.2 5.7 2.3]
       [6.5 3.0 5.8 2.2]
       [6.4 2.8 5.6 2.1]
       [5.1 3.8 1.6 0.2]
```

```
print(y_train)
```

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

```
print(X_test)
```

```
     [[5.8 2.8 5.1 2.4]
      [6.0 2.2 4.0 1.0]
      [5.5 4.2 1.4 0.2]
      [7.3 2.9 6.3 1.8]
      [5.0 3.4 1.5 0.2]
      [6.3 3.3 6.0 2.5]
      [5.0 3.5 1.3 0.3]
      [6.7 3.1 4.7 1.5]
      [6.8 2.8 4.8 1.4]
      [6.1 2.8 4.0 1.3]
      [6.1 2.6 5.6 1.4]
      [6.4 3.2 4.5 1.5]
      [6.1 2.8 4.7 1.2]
      [6.5 2.8 4.6 1.5]
      [6.1 2.9 4.7 1.4]
      [4.9 3.1 1.5 0.1]
      [6.0 2.9 4.5 1.5]
      [5.5 2.6 4.4 1.2]
      [4.8 3.0 1.4 0.3]
      [5.4 3.9 1.3 0.4]
      [5.6 2.8 4.9 2.0]
      [5.6 3.0 4.5 1.5]
      [4.8 3.4 1.9 0.2]
      [4.4 2.9 1.4 0.2]
      [6.2 2.8 4.8 1.8]
      [4.6 3.6 1.0 0.2]
      [5.1 3.8 1.9 0.4]
      [6.2 2.9 4.3 1.3]
      [5.0 2.3 3.3 1.0]
      [5.0 3.4 1.6 0.4]]
```

```
print(y_test)
```

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

MODEL 1 : Support Vecotor Machine Algorithm

```
from sklearn.svm import SVC
model_svc = SVC()
model_svc.fit(X_train, y_train)
```

```
SVC()
```

```
Prediction1 = model_svc.predict(X_test)
from sklearn.metrics import accuracy_score
print("Accuracy of the Decision Tree Classifier Model :", accuracy_score(y_test,Prediction1)*100)
for i in range(len(Prediction1)):
  print(y_test[i],Prediction1[i])
```

```
Accuracy of the Decision Tree Classifier Model : 100.0
Iris-virginica Iris-virginica
Iris-versicolor Iris-versicolor
Iris-setosa Iris-setosa
Iris-virginica Iris-virginica
Iris-setosa Iris-setosa
Iris-virginica Iris-virginica
Iris-setosa Iris-setosa
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-virginica Iris-virginica
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-setosa Iris-setosa
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-setosa Iris-setosa
Iris-setosa Iris-setosa
Iris-virginica Iris-virginica
Iris-versicolor Iris-versicolor
Iris-setosa Iris-setosa
Iris-setosa Iris-setosa
Iris-virginica Iris-virginica
Iris-setosa Iris-setosa
Iris-setosa Iris-setosa
Iris-versicolor Iris-versicolor
Iris-versicolor Iris-versicolor
Iris-setosa Iris-setosa
```

MODEL 2 : KNN (K Nearest Neighbor)

```
from sklearn.neighbors import KNeighborsClassifier
model_knn = KNeighborsClassifier()
```

```
model_knn.fit(X_train, y_train)
```

```
KNeighborsClassifier()
```

```
y_pred=knclr.predict(X_test)
```

```
from sklearn import metrics
print("Accuracy of the K Nearest Neighbor Model :", metrics.accuracy_score(y_test,y_pred)*100)
```

```
Accuracy of the Decision Tree Classifier Model : 100.0
```

MODEL 3 : Decision Tree Classifier

```
from sklearn.tree import DecisionTreeClassifier
model_dtc = DecisionTreeClassifier()
model_dtc.fit(X_train,y_train)
```

```
DecisionTreeClassifier()
```

```
print("Accuracy of the Decision Tree Classifier Model :", model_dtc.score(X_test, y_test)*100)
```

```
Accuracy of the Decision Tree Classifier Model : 100.0
```

THANK YOU!