In [1]:
```
import pandas as pd
data = pd.read_csv("Iris.csv")
data.head()
```

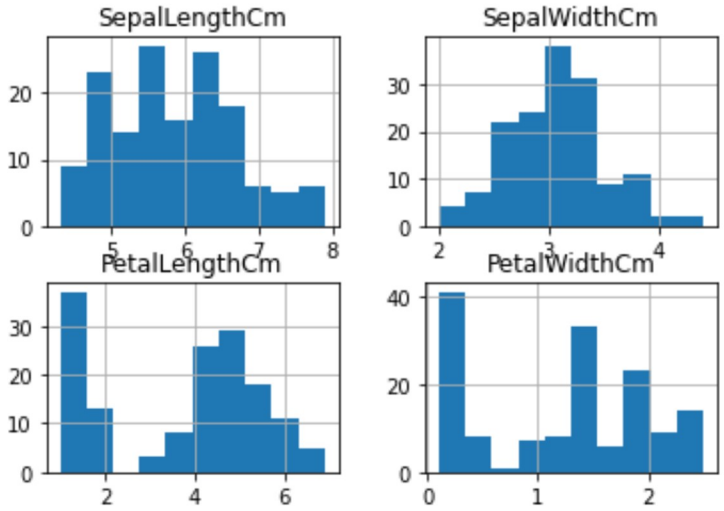
Out[1]:

	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

Visualisation

In [2]:
```
import matplotlib.pyplot as plt
data.hist()
```

Out[2]:
```
array([[<AxesSubplot:title={'center':'SepalLengthCm'}>,
        <AxesSubplot:title={'center':'SepalWidthCm'}>],
       [<AxesSubplot:title={'center':'PetalLengthCm'}>,
        <AxesSubplot:title={'center':'PetalWidthCm'}>]], dtype=object)
```



Train Test Splitting

In [3]:
```
X = data.drop(columns= "Species", axis=1)
Y = data["Species"]
```

In [4]:
```
from sklearn.model_selection import train_test_split
train_set_x, test_set_x, train_set_y, test_set_y = train_test_split(X, Y, test_size = 0.2, random_state = 0)
print(f"train_set_x: {len(train_set_x)} \ntrain_set_y: {len(train_set_y)} \ntest_set_x: {len(test_set_x)} \ntest_set_y: {len(t
```

```
train_set_x: 120
train_set_y: 120
test_set_x: 30
test_set_y: 30
```

In [5]:
```
train_set_x.head()
train_set_y.head()
test_set_x.head()
test_set_y.head()
```

Out[5]:
```
114     Iris-virginica
62     Iris-versicolor
33         Iris-setosa
107     Iris-virginica
7          Iris-setosa
Name: Species, dtype: object
```

Classification

In [6]:
```
from sklearn.neighbors import KNeighborsClassifier
k=3
model = KNeighborsClassifier(n_neighbors=k).fit(train_set_x, train_set_y)
model
```

Out[6]:
```
KNeighborsClassifier(n_neighbors=3)
```

Prediction

In [7]:
```
test_set_y.head(30)
```

Out[7]:
```
114     Iris-virginica
62     Iris-versicolor
33         Iris-setosa
107     Iris-virginica
7          Iris-setosa
100     Iris-virginica
40         Iris-setosa
86     Iris-versicolor
76     Iris-versicolor
71     Iris-versicolor
134     Iris-virginica
51     Iris-versicolor
73     Iris-versicolor
54     Iris-versicolor
63     Iris-versicolor
37         Iris-setosa
78     Iris-versicolor
90     Iris-versicolor
45         Iris-setosa
16         Iris-setosa
121     Iris-virginica
66     Iris-versicolor
24         Iris-setosa
8          Iris-setosa
126     Iris-virginica
22         Iris-setosa
44         Iris-setosa
97     Iris-versicolor
93     Iris-versicolor
26         Iris-setosa
Name: Species, dtype: object
```

In [8]:
```
y_pred = model.predict(test_set_x)
y_pred
```

Out[8]:
```
array(['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-virginica', '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'], dtype=object)
```

Accuracy

In [9]:
```
from sklearn import metrics
accu = metrics.accuracy_score(test_set_y, y_pred)
accu
```

Out[9]:
```
0.9666666666666667
```

In [ ]: