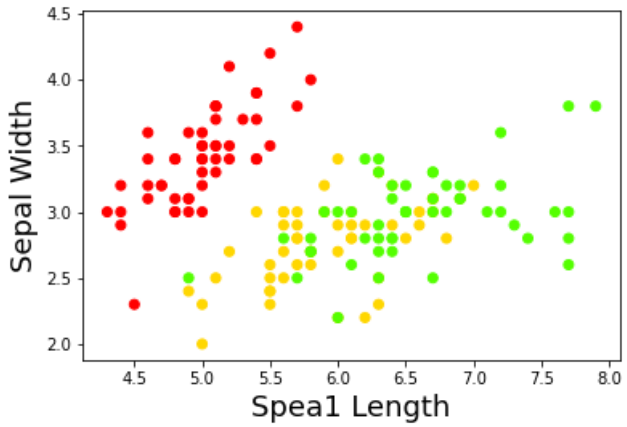
In [1]:
```
from sklearn import datasets
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.cluster import KMeans
```

In [2]:
```
iris = datasets.load_iris()
X = iris.data[:, :2]
y = iris.target
```

In [3]:
```
plt.scatter(X[:,0], X[:,1], c=y, cmap='prism')
plt.xlabel('Spea1 Length', fontsize=18)
plt.ylabel('Sepal Width', fontsize=18)
```

Out[3]: Text(0, 0.5, 'Sepal Width')

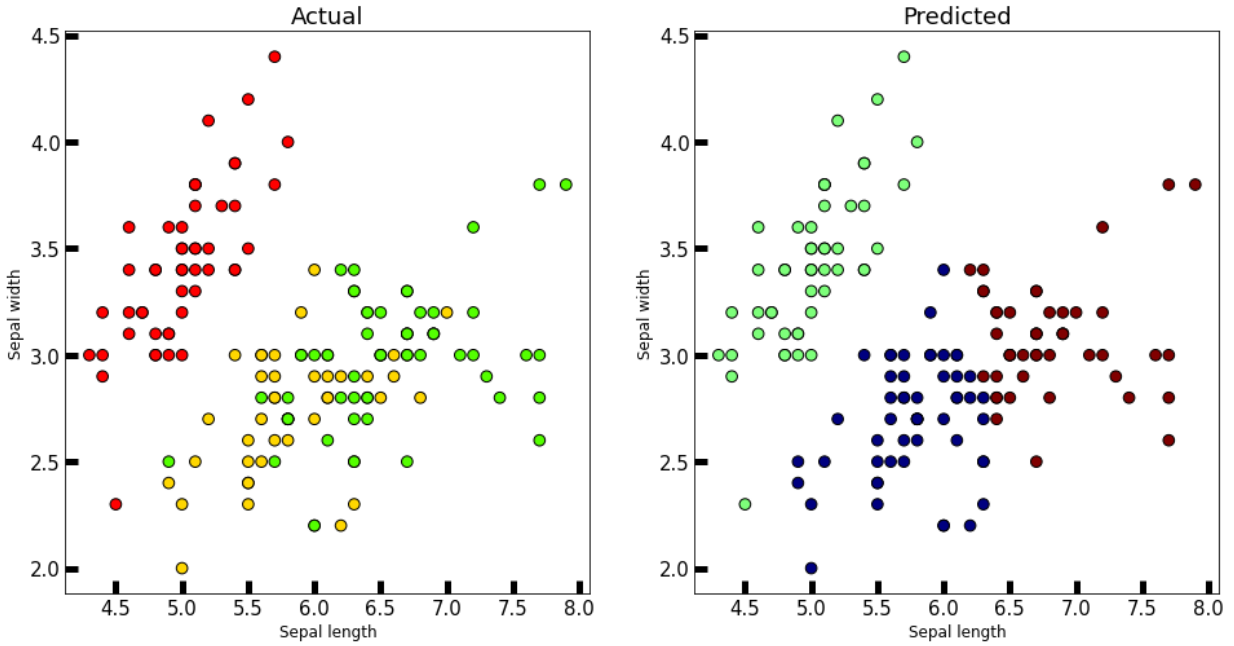


In [4]:
```
km = KMeans(n_clusters = 3, init='k-means++', n_init=10, max_iter=300, tol=0.0001, verbose=0, random_
km.fit(X)
centers = km.cluster_centers_
print(centers)
new_labels = km.labels_
print(new_labels)
print(y)
```

```
[[5.77358491 2.69245283]
 [5.006      3.428     ]
 [6.81276596 3.07446809]]
[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 0 2 0 2 0 2 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0
 2 2 2 2 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 2 0 2 2 2 2 0 2 2 2 2
 2 2 0 0 2 2 2 2 0 2 0 2 0 2 2 0 0 2 2 2 2 2 0 0 2 2 2 0 2 2 2 0 2 2 2 0 2
 2 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 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 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
 2 2]
```

In [5]:
```
fig, axes = plt.subplots( 1,2, figsize=(16,8))
axes[0].scatter(X[:, 0], X[:, 1], c=y, cmap='prism',edgecolor='k', s=75)
axes[1].scatter(X[:, 0], X[:, 1], c=new_labels, cmap='jet',edgecolor='k', s=75)
axes[0].set_xlabel('Sepal length', fontsize=12)
axes[0].set_ylabel('Sepal width', fontsize=12)
axes[1].set_xlabel('Sepal length', fontsize=12)
axes[1].set_ylabel('Sepal width', fontsize=12)
axes[0].tick_params(direction='in', length=10, width=5, colors='k', labelsize=15)
axes[1].tick_params(direction='in', length=10, width=5, colors='k', labelsize=15)
axes[0].set_title('Actual', fontsize=18)
axes[1].set_title('Predicted', fontsize=18)
```

Out[5]:
```
Text(0.5, 1.0, 'Predicted')
```



In [6]:
```
km = KMeans(n_clusters = 3, init='k-means++', n_init=10, max_iter=300, tol=0.0001, verbose=0, random_
km.fit(X)
km.fit_predict(X)
km.fit_transform(X)
km.get_params()
km.predict(X)
km.score(X)
km.transform(X)
```

Out[6]:
```
array([[1.05159358, 0.11840608, 1.76483558],
       [0.9261403 , 0.44093083, 1.91421501],
       [1.18751365, 0.38160189, 2.11649197],
       [1.24233499, 0.52193869, 2.21291325],
       [1.19250806, 0.17210462, 1.88740674],
       [1.26401576, 0.61483331, 1.6362795 ],
       [1.37037386, 0.40696437, 2.23658316],
       [1.04835901, 0.02863564, 1.84176313],
       [1.38917649, 0.80375369, 2.41906566],
       [0.96397369, 0.34470277, 1.91293635],
       [1.07457758, 0.47876926, 1.54505593],
       [1.20353254, 0.2078942 , 2.03892075],
       [1.02100579, 0.47499474, 2.01414307],
       [1.50533642, 0.82560281, 2.51386918],
       [1.30781396, 0.9785806 , 1.37197092],
       [1.70913197, 1.19432826, 1.73068857],
       [1.26401576, 0.61483331, 1.6362795 ],
       [1.05159358, 0.11840608, 1.76483558],
       [1.10998895, 0.78741349, 1.32839928],
```