

Kmeans

October 12, 2023

```
[1]: # import required libraries
import numpy as np
import matplotlib.pyplot as plt
from sklearn.datasets import make_blobs
```

```
[3]: # Generate our dataset
dataset = make_blobs(n_samples=200,
                    centers=4,
                    n_features=2,
                    cluster_std=1.5,
                    random_state=50)
```

```
[8]: # print (dataset)
points = dataset[0]
```

```
[9]: # import kmeans
from sklearn.cluster import KMeans
```

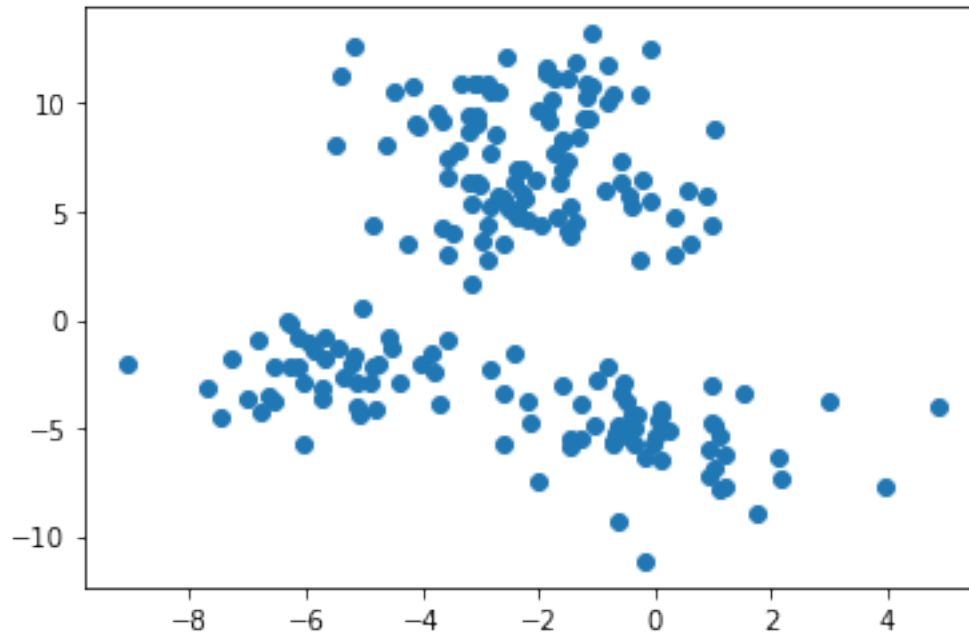
```
[11]: # Create a kmeans objects
kmeans = KMeans(n_clusters=4)
```

```
[12]: # fit the kmeans to the dataset
kmeans.fit(points)
```

```
[12]: KMeans(n_clusters=4)
```

```
[14]: plt.scatter(dataset[0][:,0],dataset[0][:,1])
```

```
[14]: <matplotlib.collections.PathCollection at 0x7fd612bca9d0>
```



```
[15]: clusters = kmeans.cluster_centers_
```

```
[16]: # print out the clusters
      print (clusters)
```

```
[[ -5.45922685  -2.35556894]
 [ -1.90808931   5.14791029]
 [  0.06912515  -5.38143605]
 [ -2.39930521  10.05258314]]
```

```
[17]: y_km = kmeans.fit_predict(points)
```

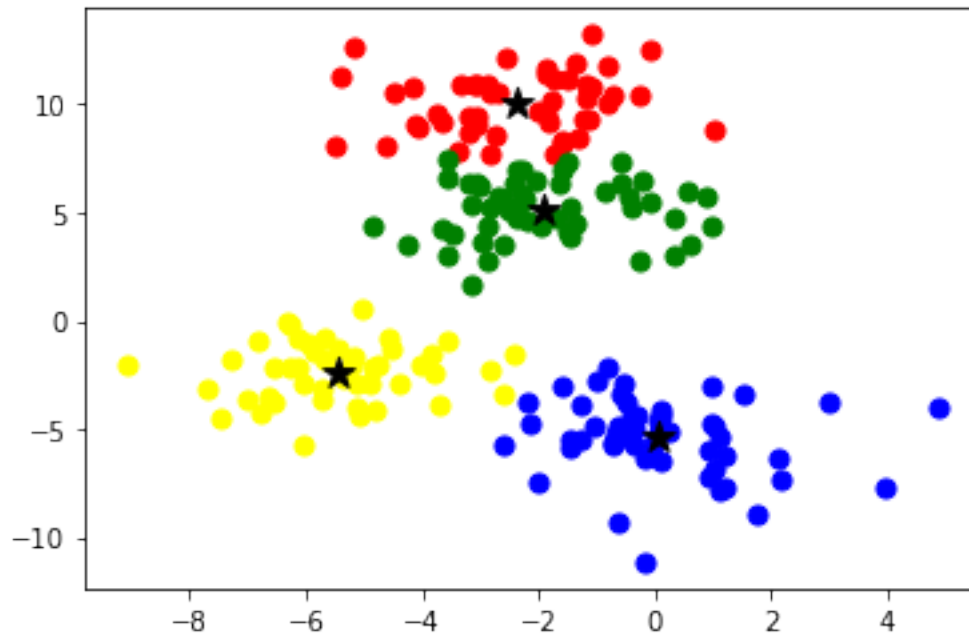
```
[18]: y_km
```

```
[18]: array([0, 1, 2, 0, 3, 1, 0, 1, 1, 3, 3, 0, 3, 2, 3, 2, 0, 0, 3, 3, 3, 2,
           0, 0, 0, 2, 0, 1, 2, 0, 0, 3, 3, 1, 1, 3, 3, 3, 2, 2, 0, 3, 2, 2,
           1, 1, 0, 2, 2, 0, 1, 3, 2, 0, 2, 2, 1, 3, 3, 2, 1, 1, 0, 1, 1, 0,
           1, 2, 0, 0, 1, 3, 0, 1, 2, 2, 3, 2, 2, 1, 3, 2, 3, 3, 3, 3, 0, 1,
           0, 3, 1, 0, 1, 2, 2, 3, 1, 2, 1, 2, 3, 1, 2, 0, 2, 1, 2, 3, 3, 0,
           1, 3, 0, 0, 3, 2, 0, 0, 2, 1, 0, 2, 0, 0, 1, 0, 0, 2, 0, 1, 3, 1,
           1, 2, 0, 2, 1, 0, 0, 3, 0, 2, 2, 1, 3, 0, 1, 2, 1, 3, 0, 3, 2, 1,
           1, 2, 0, 1, 1, 1, 2, 2, 1, 1, 3, 3, 3, 2, 1, 1, 1, 1, 3, 2, 1, 0,
           3, 1, 0, 2, 1, 2, 2, 0, 3, 1, 3, 0, 2, 1, 2, 2, 3, 2, 1, 0, 3, 2,
           3, 0], dtype=int32)
```

```
[29]: plt.scatter(points[y_km==0,0],points[y_km ==0,1],s=50,color='red')
plt.scatter(points[y_km==1,0],points[y_km ==1,1],s=50,color='blue')
plt.scatter(points[y_km==2,0],points[y_km ==2,1],s=50,color='green')
plt.scatter(points[y_km==3,0],points[y_km ==3,1],s=50,color='yellow')

plt.scatter(clusters[0][0],clusters[0][1],marker='*',s=150,color='black')
plt.scatter(clusters[1][0],clusters[1][1],marker='*',s=150,color='black')
plt.scatter(clusters[2][0],clusters[2][1],marker='*',s=150,color='black')
plt.scatter(clusters[3][0],clusters[3][1],marker='*',s=150,color='black')
```

[29]: <matplotlib.collections.PathCollection at 0x7fd6140d8640>



[]: