SVMandClustering

November 30, 2021

SL Lab with Python 7: SVM and Clustering

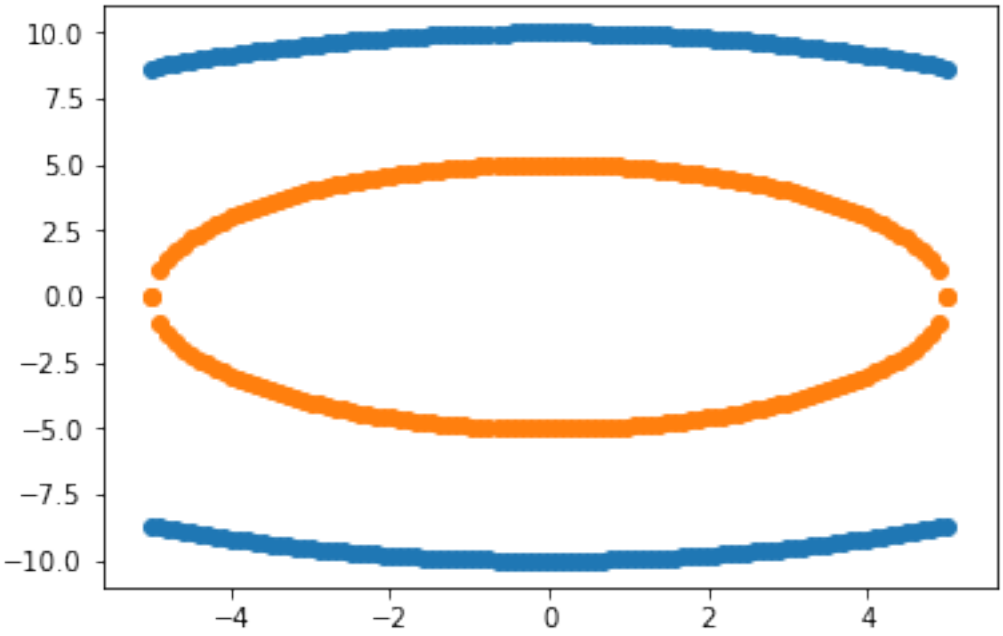
Statistical Learning (Sejong University)

Date: 2021.11.30 (By: S. M. Riazul Islam)

1. SVM

In [7]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score


#Outer Circle Data
x1 = np.linspace(-5.0, 5.0, 100)
y1 = np.sqrt(10**2 - x1**2) # Circle Equation
x1=np.hstack([x1,-x1])
y1=np.hstack([y1,-y1])

#Inner Circle Data
x2 = np.linspace(-5.0, 5.0, 100)
y2 = np.sqrt(5**2 - x2**2)

x2=np.hstack([x2,-x2])
y2=np.hstack([y2,-y2])


plt.scatter(x1,y1)
plt.scatter(x2,y2)
```

Out[7]:
```
<matplotlib.collections.PathCollection at 0x25029038f28>
```



In [2]:
```
# Create Dataset
df1 =pd.DataFrame(np.vstack([y1,x1]).T,columns=['X1','X2'])
df1['Y']=0

df2 =pd.DataFrame(np.vstack([y2,x2]).T,columns=['X1','X2'])
df2['Y']=1
df = df1.append(df2)
df.head()
```

Out[2]:

	X1	X2	Y
0	8.660254	-5.00000	0
1	8.717792	-4.89899	0
2	8.773790	-4.79798	0
3	8.828277	-4.69697	0
4	8.881281	-4.59596	0

In [3]:
```
# Predictors and Response Variable
X = df.iloc[:, :2]
y = df.Y

# Split the dataset into train and test
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.25,random_state=0)
```

In [4]:
```
#kernel{linear, poly, rbf, sigmoid, precomputed}, default=rbf

myclass=SVC(kernel="linear")
```

```
        myclass.fit(X_train,y_train)

        y_pred = myclass.predict(X_test)
        accuracy_score(y_test, y_pred)
```

```
Out[4]: 0.45
```

```
In [5]: myclass=SVC(kernel="poly")
        myclass.fit(X_train,y_train)

        y_pred = myclass.predict(X_test)
        accuracy_score(y_test, y_pred)
```

```
C:\Users\Preload\Anaconda3\lib\site-packages\sklearn\svm\base.py:196: FutureWarning: The defaul
  "avoid this warning.", FutureWarning)
```

```
Out[5]: 0.6
```

```
In [6]: myclass=SVC(kernel="rbf")
        myclass.fit(X_train,y_train)

        y_pred = myclass.predict(X_test)
        accuracy_score(y_test, y_pred)
```

```
C:\Users\Preload\Anaconda3\lib\site-packages\sklearn\svm\base.py:196: FutureWarning: The defaul
  "avoid this warning.", FutureWarning)
```

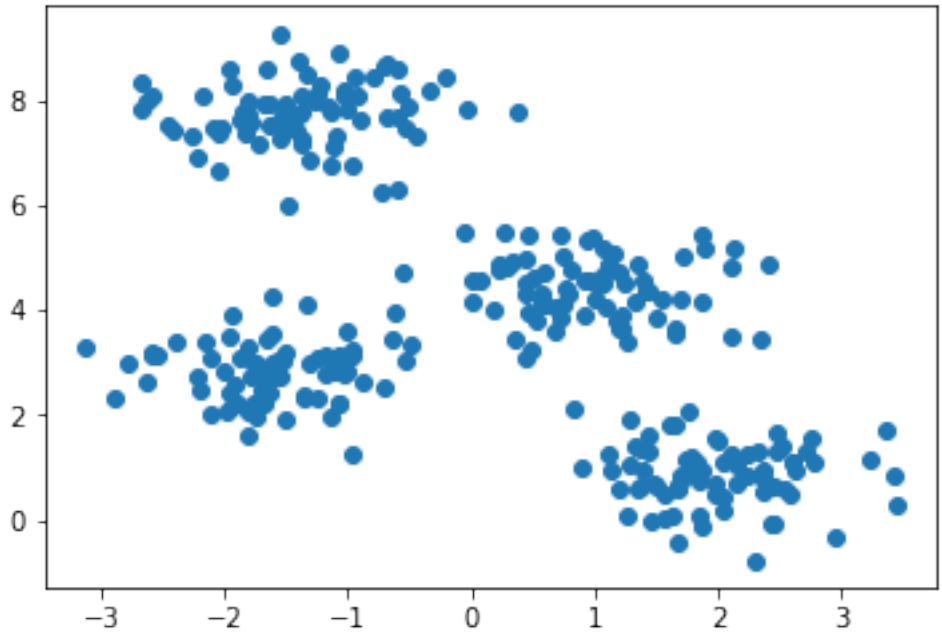
```
Out[6]: 1.0
```

2. K-means Clustering

```
In [8]: import numpy as np
        import pandas as pd
        from matplotlib import pyplot as plt

        from sklearn.datasets.samples_generator import make_blobs
        from sklearn.cluster import KMeans
```

```
In [12]: X, y = make_blobs(n_samples=300, n_features=2, centers=4, cluster_std=0.60, random_sta
         plt.scatter(X[:,0], X[:,1])
```

```
Out[12]: <matplotlib.collections.PathCollection at 0x250295d9f98>
```




```
In [34]: myCluster=KMeans(n_clusters=4)
         myCluster.fit_predict(X)
         centers=myCluster.cluster_centers_
         centers
```

```
Out[34]: array([[ 1.98258281,  0.86771314],
                [ 0.94973532,  4.41906906],
                [-1.58438467,  2.83081263],
                [-1.37324398,  7.75368871]])
```

```
In [35]: mykm=myCluster.fit_predict(X)
         mykm
```

```
Out[35]: array([1, 0, 3, 0, 1, 1, 2, 3, 0, 0, 2, 0, 3, 0, 1, 3, 3, 1, 2, 2, 1, 1,
                3, 2, 2, 3, 1, 3, 2, 3, 0, 0, 3, 0, 0, 0, 0, 0, 2, 1, 3, 2, 3, 3,
                2, 2, 0, 2, 0, 1, 2, 1, 0, 1, 1, 2, 0, 2, 0, 1, 0, 3, 0, 2, 2, 2,
                0, 1, 0, 2, 3, 2, 0, 2, 2, 0, 2, 3, 1, 0, 1, 3, 1, 1, 0, 3, 1, 3,
                0, 0, 3, 1, 0, 2, 2, 3, 1, 1, 3, 2, 0, 1, 0, 1, 3, 1, 1, 3, 0, 3,
                2, 2, 1, 0, 1, 3, 0, 1, 1, 3, 2, 1, 2, 1, 1, 1, 1, 2, 1, 2, 0, 2,
                2, 1, 0, 2, 2, 0, 3, 0, 0, 2, 3, 2, 3, 2, 0, 3, 0, 0, 0, 3, 0, 3,
                1, 2, 0, 2, 1, 3, 0, 3, 3, 1, 3, 2, 2, 3, 1, 3, 3, 0, 1, 3, 2, 0,
                1, 1, 3, 2, 1, 3, 2, 2, 3, 3, 3, 3, 1, 0, 3, 2, 3, 3, 2, 2, 2, 3,
                2, 0, 3, 2, 1, 2, 3, 0, 2, 0, 3, 0, 3, 2, 3, 3, 0, 2, 2, 1, 1, 3,
                0, 1, 1, 2, 1, 2, 3, 0, 0, 3, 3, 0, 3, 1, 2, 3, 1, 2, 0, 2, 1, 3,
                1, 0, 0, 0, 0, 2, 2, 0, 3, 2, 1, 3, 2, 2, 2, 1, 1, 0, 3, 3, 2, 1,
                0, 2, 3, 0, 3, 1, 1, 2, 2, 3, 1, 1, 1, 3, 0, 0, 1, 1, 3, 1, 1, 1,
                0, 2, 0, 3, 1, 1, 0, 0, 0, 1, 1, 3, 0, 2])
```

```
In [36]: X

Out[36]: array([[ 8.36856841e-01,  2.13635938e+00],
                [-1.41365810e+00,  7.40962324e+00],
                [ 1.15521298e+00,  5.09961887e+00],
                [-1.01861632e+00,  7.81491465e+00],
                [ 1.27135141e+00,  1.89254207e+00],
                [ 3.43761754e+00,  2.61654166e-01],
                [-1.80822253e+00,  1.59701749e+00],
                [ 1.41372442e+00,  4.38117707e+00],
                [-2.04932168e-01,  8.43209665e+00],
                [-7.11099611e-01,  8.66043846e+00],
                [-1.71237268e+00,  2.77780226e+00],
                [-2.67000792e+00,  8.35389140e+00],
                [ 1.24258802e+00,  4.50399192e+00],
                [-2.22783649e+00,  6.89479938e+00],
                [ 1.45513831e+00, -2.91989981e-02],
                [ 4.53791789e-01,  3.95647753e+00],
                [ 1.06923853e+00,  4.53068484e+00],
                [ 2.56936589e+00,  5.07048304e-01],
                [-1.06690610e+00,  3.13165795e+00],
                [-1.07859101e+00,  2.20451529e+00],
                [ 2.71506328e+00,  1.29082190e+00],
                [ 1.77710994e+00,  1.18655254e+00],
                [ 7.34363910e-01,  5.03725437e+00],
                [-1.99648151e+00,  2.85164868e+00],
                [-1.91828017e+00,  2.60516867e+00],
                [-5.55523811e-01,  4.69595848e+00],
                [ 1.69747910e+00,  8.66123282e-01],
                [ 5.94762432e-01,  4.70964730e+00],
                [-2.88024255e+00,  2.30437816e+00],
                [ 1.86725632e-01,  4.02683656e+00],
                [-5.13333436e-01,  7.87437368e+00],
                [-2.05701103e+00,  7.37851598e+00],
                [ 1.87271752e+00,  4.18069237e+00],
                [-1.13121396e+00,  6.76652230e+00],
                [-1.64428957e+00,  7.94095139e+00],
                [-2.41933824e+00,  7.43952498e+00],
                [-2.01606928e+00,  7.48420129e+00],
                [-2.62142780e+00,  7.98635066e+00],
                [-2.20299950e+00,  2.47947561e+00],
                [ 1.66382237e+00,  6.63091635e-01],
                [ 6.69786996e-01,  3.59540802e+00],
                [-1.98539037e+00,  2.05520738e+00],
                [-4.74920358e-02,  5.47425256e+00],
                [ 1.54462126e+00,  4.21078127e+00],
                [-1.70200643e+00,  2.46098693e+00],
                [-1.02192525e+00,  2.76820711e+00],
```

```
[-1.37842552e+00,  8.10269597e+00],
[-1.55220688e+00,  2.74574995e+00],
[-1.47904303e+00,  7.56915836e+00],
[ 1.98894313e+00,  1.50976729e+00],
[-1.94972418e+00,  3.48383870e+00],
[ 2.45509737e+00,  6.21231788e-01],
[-8.94327882e-01,  7.61714473e+00],
[ 1.69687788e+00,  7.54910622e-01],
[ 1.75644805e+00,  2.05538289e+00],
[-1.11064012e+00,  2.82213820e+00],
[-4.15017659e-02,  7.80870276e+00],
[-1.14091533e+00,  1.97550822e+00],
[-1.80634968e+00,  7.72830656e+00],
[ 1.39263752e+00,  9.28962707e-01],
[-2.25698670e+00,  7.30207720e+00],
[ 5.71670482e-01,  4.32288566e+00],
[-1.54994580e+00,  9.28293222e+00],
[-1.03819212e+00,  2.95323761e+00],
[-2.10956676e+00,  3.10677258e+00],
[-1.18652985e+00,  2.78427720e+00],
[-2.45809975e+00,  7.51178724e+00],
[ 2.36960214e+00,  9.50716912e-01],
[-2.66676007e+00,  7.84766052e+00],
[-1.49720702e+00,  3.21418433e+00],
[ 1.32222457e+00,  4.17880807e+00],
[-4.87271301e-01,  3.32858293e+00],
[-1.03718771e+00,  8.06300134e+00],
[-1.60496819e+00,  2.97425420e+00],
[-1.50372568e+00,  1.92385320e+00],
[-7.85412206e-01,  8.45312331e+00],
[-1.75783190e+00,  2.97449321e+00],
[ 1.19008992e+00,  4.72773123e+00],
[ 2.14043942e+00,  7.06066610e-01],
[-1.03540116e+00,  8.20559195e+00],
[ 1.25471244e+00,  8.96331565e-02],
[ 5.95676822e-01,  4.08614263e+00],
[ 1.27955338e+00,  1.05789418e+00],
[ 1.72955064e+00,  1.14729369e+00],
[-9.49041610e-01,  8.46424331e+00],
[ 9.35325115e-01,  5.33179296e+00],
[ 2.36923352e+00,  7.94735861e-01],
[ 4.28576743e-01,  4.97399710e+00],
[-2.04758277e+00,  6.65428520e+00],
[-1.45672617e+00,  7.48669893e+00],
[ 5.29417185e-01,  3.80792420e+00],
[ 1.78194802e+00,  9.08151155e-01],
[-1.95575053e+00,  8.61631686e+00],
[-1.74572014e+00,  3.01190457e+00],
```

```
[-1.17979111e+00,  3.12767494e+00],
[ 1.16411070e+00,  3.79132988e+00],
[ 1.36155806e+00,  1.36638252e+00],
[ 2.60137487e+00,  1.08799459e+00],
[ 2.72396035e-01,  5.46996004e+00],
[-3.12240736e+00,  3.28167398e+00],
[-5.88085172e-01,  8.61427320e+00],
[ 1.66909648e+00, -4.36378231e-01],
[-6.82774903e-01,  7.67490317e+00],
[ 2.36790645e+00,  5.52190878e-01],
[ 1.05241733e+00,  4.54498095e+00],
[ 2.22707373e+00,  1.26296996e+00],
[ 2.43934644e+00, -7.25099666e-02],
[ 1.34471770e+00,  4.85711133e+00],
[-1.31454942e+00,  6.83904013e+00],
[ 9.82570091e-01,  5.37530962e+00],
[-1.06295223e+00,  2.20755388e+00],
[-1.60712495e+00,  3.56452854e+00],
[ 1.57322172e+00,  4.83933793e-01],
[-2.17851338e+00,  8.08585131e+00],
[ 1.83375842e+00,  7.54036153e-01],
[ 2.10616050e+00,  3.49513189e+00],
[-1.64338755e+00,  7.52691964e+00],
[ 1.10550448e+00,  1.26389129e+00],
[ 1.61152972e+00,  1.82347242e+00],
[ 4.59534668e-01,  5.44982630e+00],
[-5.38023054e-01,  3.01641891e+00],
[ 1.67841499e+00,  6.09047389e-01],
[-1.01214966e+00,  3.60254338e+00],
[ 1.34195197e+00,  5.93573847e-01],
[ 1.42811550e+00,  1.62435083e+00],
[ 2.04505527e+00,  1.12515470e+00],
[ 1.67280531e+00,  6.59300571e-01],
[-1.35863899e+00,  2.32200809e+00],
[ 1.13078931e+00,  9.35620856e-01],
[-1.73896306e+00,  1.94799775e+00],
[-3.40258925e-01,  8.16733925e+00],
[-1.63793622e+00,  2.43329317e+00],
[-1.68754414e+00,  2.24107546e+00],
[ 2.43040639e+00, -6.35709334e-02],
[-1.37998039e+00,  7.18503804e+00],
[-1.25245465e+00,  2.33911505e+00],
[-2.39482483e+00,  3.39806103e+00],
[-2.09183352e+00,  7.48142543e+00],
[ 4.88382309e-01,  3.26801777e+00],
[-5.39428614e-01,  7.45631776e+00],
[-2.59201821e+00,  8.07635945e+00],
[-1.04656513e+00,  2.96483837e+00],
```

```
[ 1.25566754e+00,  3.38204112e+00],
[-1.62150422e+00,  4.27191636e+00],
[ 1.86922139e+00,  5.44132083e+00],
[-1.76404140e+00,  2.22222995e+00],
[-1.38181918e+00,  7.28761248e+00],
[ 7.93137001e-03,  4.17614316e+00],
[-1.10298500e+00,  7.30222786e+00],
[-1.79351372e+00,  7.58086944e+00],
[-1.51176949e+00,  7.94426307e+00],
[ 9.59360742e-01,  4.56078645e+00],
[-6.01337393e-01,  6.29994922e+00],
[ 2.24592863e-01,  4.77028154e+00],
[ 1.56724897e+00,  1.78090633e-02],
[-1.03395259e+00,  2.92105785e+00],
[-9.22268640e-01,  8.09888579e+00],
[-1.88609638e+00,  2.24834407e+00],
[ 1.86873582e+00,  9.56103760e-01],
[ 1.10123507e+00,  4.88977075e+00],
[-1.93213725e+00,  8.30643453e+00],
[ 6.70478769e-01,  4.04094275e+00],
[ 7.43873988e-01,  4.12240568e+00],
[ 1.64003761e+00,  1.81948989e+00],
[ 8.15468056e-01,  4.78526116e+00],
[-2.63274574e+00,  2.63109786e+00],
[-9.61076361e-01,  1.27407596e+00],
[ 2.13979079e-01,  4.88542535e+00],
[ 1.43472182e+00,  1.30662037e+00],
[ 1.21387411e+00,  3.64795042e+00],
[ 1.08272576e+00,  4.06271877e+00],
[-1.22605234e+00,  8.29620146e+00],
[ 1.48170052e+00,  6.90074595e-01],
[ 1.89593761e+00,  5.18540259e+00],
[-1.32356154e+00,  4.13050840e+00],
[-1.14989863e+00,  7.89325283e+00],
[ 2.46915842e+00,  1.67877140e+00],
[ 2.31102276e+00,  1.30380848e+00],
[ 5.72793810e-01,  4.08805543e+00],
[-9.67794989e-01,  3.12186125e+00],
[ 2.62492001e+00,  9.50194405e-01],
[ 1.68353782e+00,  4.19583243e+00],
[-2.22131717e+00,  2.73050691e+00],
[-1.57846247e+00,  3.03445802e+00],
[ 8.15155229e-02,  4.56742235e+00],
[ 1.43289271e+00,  4.37679234e+00],
[ 1.06269622e+00,  5.17635143e+00],
[ 7.67522789e-01,  4.39759671e+00],
[ 2.47019077e+00,  1.31451315e+00],
[-1.73163168e+00,  7.16447286e+00],
```

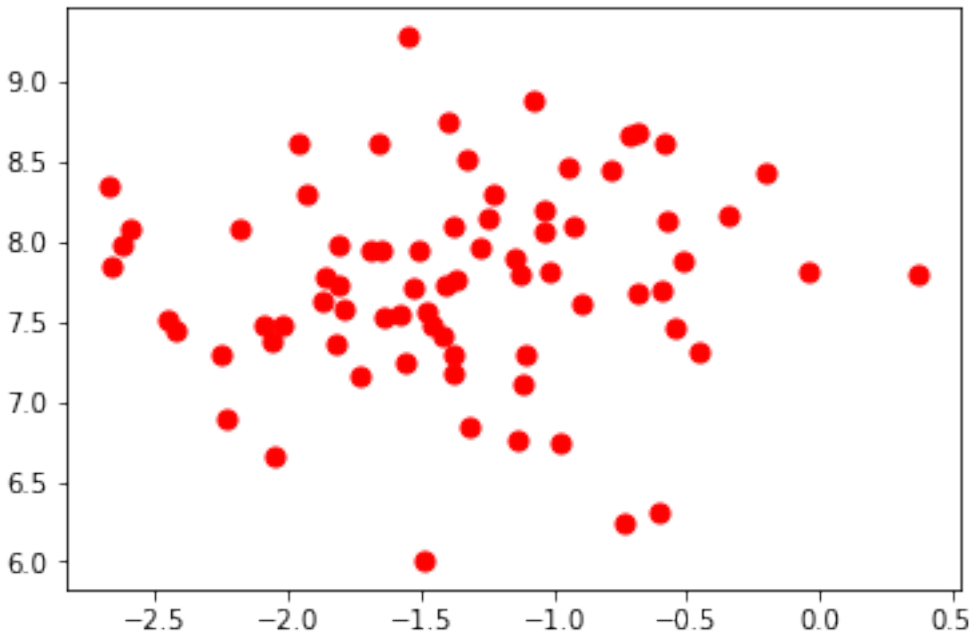
```
[ 3.47138300e-01,  3.45177657e+00],
[-1.00140436e+00,  2.84881778e+00],
[ 1.01618041e+00,  4.48527047e+00],
[ 5.59529363e-01,  4.21400660e+00],
[-2.11821046e+00,  2.03478126e+00],
[-1.36219420e+00,  2.38333321e+00],
[-2.78366586e+00,  2.99211541e+00],
[ 1.65225163e+00,  3.65582842e+00],
[-1.93960658e+00,  2.18943582e+00],
[-1.81487687e+00,  7.97832190e+00],
[ 1.20212540e+00,  3.64414685e+00],
[-9.69200946e-01,  3.26721702e+00],
[ 1.86985974e+00, -1.07938624e-01],
[-1.80701134e+00,  2.06773859e+00],
[ 1.21767506e+00,  3.89290127e+00],
[-1.48368917e+00,  6.00777418e+00],
[-1.56387985e+00,  2.85349910e+00],
[-6.86210297e-01,  8.68285664e+00],
[ 1.07627418e+00,  4.68480619e+00],
[-9.75986627e-01,  6.73761577e+00],
[ 1.37964693e+00,  4.54826443e+00],
[-1.64129611e+00,  2.68097255e+00],
[-1.84892963e-03,  4.58145668e+00],
[ 1.71444449e+00,  5.02521524e+00],
[-1.40524304e+00,  7.72611315e+00],
[-7.08184904e-01,  2.50421275e+00],
[-8.85798374e-01,  2.64585078e+00],
[ 1.98436909e+00,  4.89556738e-01],
[ 2.95195825e+00, -3.44327355e-01],
[ 4.31891060e-01,  4.33495456e+00],
[-1.86621694e+00,  7.62549400e+00],
[ 2.52706430e+00,  6.17812202e-01],
[ 2.04067185e+00,  4.54845114e-01],
[-2.58043836e+00,  3.18844294e+00],
[ 1.61990909e+00,  6.76452867e-02],
[-2.58802708e+00,  3.13117134e+00],
[ 4.43598630e-01,  3.11530945e+00],
[-4.56773649e-01,  7.30600872e+00],
[-1.12867566e+00,  7.80509789e+00],
[ 2.13003529e+00,  5.19209620e+00],
[ 1.00372519e+00,  4.19147702e+00],
[-1.39346767e+00,  8.74641672e+00],
[ 7.28098690e-01,  3.85531444e+00],
[ 8.93499638e-01,  1.01093082e+00],
[-1.10782972e+00,  2.92014479e+00],
[ 7.89338559e-01,  4.33748653e+00],
[ 1.97553917e+00,  7.18989132e-01],
[-1.24902582e+00,  3.08533972e+00],
```

```
[-1.07768797e+00,  8.88106128e+00],
[-1.86849125e+00,  3.07982487e+00],
[ 2.76808540e+00,  1.08782923e+00],
[ 2.77180174e-01,  4.84428322e+00],
[ 3.41085289e+00,  8.72309369e-01],
[-1.58084155e+00,  7.55307742e+00],
[-1.53027550e+00,  7.70542113e+00],
[-1.82455840e+00,  7.35958960e+00],
[-1.68568257e+00,  7.95339446e+00],
[-1.65104622e+00,  3.44598961e+00],
[-1.30440901e+00,  3.00311934e+00],
[-7.30698168e-01,  6.24208856e+00],
[ 2.40615694e+00,  4.87047502e+00],
[-1.53631328e+00,  3.01443916e+00],
[ 1.48859977e+00,  6.51633844e-01],
[ 5.14320434e-01,  4.62733684e+00],
[-1.81469750e+00,  3.29009724e+00],
[-1.93731055e+00,  3.91361274e+00],
[-6.14907097e-01,  3.94963585e+00],
[ 2.03169783e+00,  1.96807561e-01],
[ 2.14917144e+00,  1.03697228e+00],
[-1.36999388e+00,  7.76953035e+00],
[ 9.14338767e-01,  4.55014643e+00],
[ 3.33818506e-01,  4.93645836e+00],
[-2.15968109e+00,  3.41003096e+00],
[ 1.36678633e+00,  6.34971633e-01],
[-5.71099336e-01,  8.13306058e+00],
[-1.00552592e+00,  3.08400540e+00],
[ 1.49493180e+00,  3.85848832e+00],
[-5.90447667e-01,  7.69493053e+00],
[ 7.15177948e-01,  5.41334556e+00],
[ 2.11390250e+00,  1.24743587e+00],
[ 1.20083098e+00,  6.01671730e-01],
[-2.54576750e+00,  3.15025055e+00],
[-1.95866665e+00,  2.43008647e+00],
[ 2.33812285e+00,  3.43116792e+00],
[ 3.35320909e+00,  1.69958043e+00],
[ 1.84287117e+00,  7.26928839e-02],
[ 1.32000621e+00,  1.40428145e+00],
[ 2.09680545e+00,  4.84741412e+00],
[-1.24307904e+00,  8.15166254e+00],
[-1.85908090e+00,  7.78874716e+00],
[ 2.74666646e+00,  1.54543482e+00],
[ 2.60778282e+00,  1.08890025e+00],
[ 1.65991049e+00,  3.56289184e+00],
[ 2.35151259e+00,  8.28001297e-01],
[ 2.22322228e+00,  8.38773426e-01],
[ 3.22881491e+00,  1.13171965e+00],
```

```
       [-1.55876720e+00,  7.24816210e+00],
       [-6.46956784e-01,  3.42941343e+00],
       [-1.32688818e+00,  8.51530794e+00],
       [ 9.17198564e-01,  3.90570036e+00],
       [ 2.29469533e+00, -7.65891994e-01],
       [ 1.81559810e+00,  1.11969719e+00],
       [-1.12016775e+00,  7.11031582e+00],
       [-1.65507124e+00,  8.61416749e+00],
       [-1.27567815e+00,  7.96776461e+00],
       [ 1.97369770e+00,  1.57979848e+00],
       [ 2.51834185e+00,  1.39176615e+00],
       [ 4.38990142e-01,  4.53592883e+00],
       [ 3.69478657e-01,  7.79110522e+00],
       [-1.79145759e+00,  2.74966896e+00]])
```

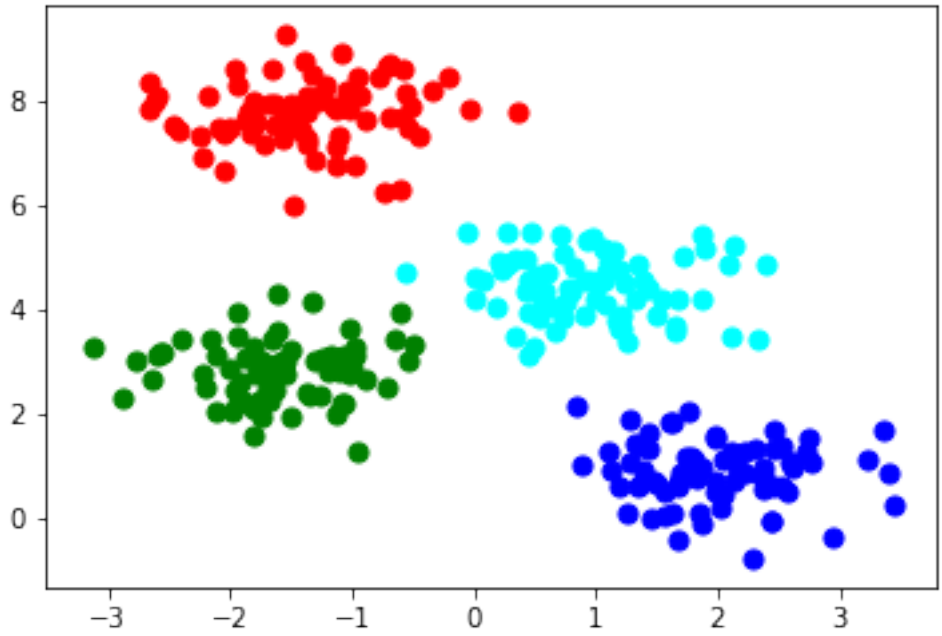
In [37]:
```
plt.scatter(X[mykm==0,0],X[mykm==0,1],s=50,color='red')
```

Out[37]:
```
<matplotlib.collections.PathCollection at 0x250297b9400>
```



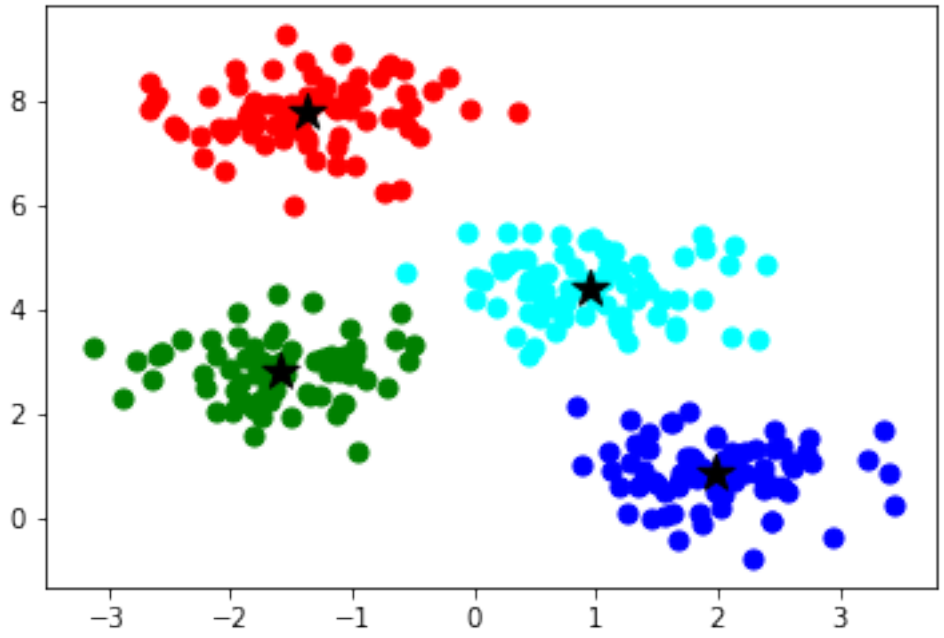
In [38]:
```
plt.scatter(X[mykm==0,0],X[mykm==0,1],s=50,color='red')
plt.scatter(X[mykm==1,0],X[mykm==1,1],s=50,color='blue')
plt.scatter(X[mykm==2,0],X[mykm==2,1],s=50,color='green')
plt.scatter(X[mykm==3,0],X[mykm==3,1],s=50,color='cyan')
```

Out[38]:
```
<matplotlib.collections.PathCollection at 0x2502981d588>
```



In [45]:
```
plt.scatter(X[mykm==0,0],X[mykm==0,1],s=50,color='red')
plt.scatter(X[mykm==1,0],X[mykm==1,1],s=50,color='blue')
plt.scatter(X[mykm==2,0],X[mykm==2,1],s=50,color='green')
plt.scatter(X[mykm==3,0],X[mykm==3,1],s=50,color='cyan')

plt.scatter(centers[0][0],centers[0][1], marker='*', s=200, color='black')
plt.scatter(centers[1][0],centers[1][1], marker='*', s=200, color='black')
plt.scatter(centers[2][0],centers[2][1], marker='*', s=200, color='black')
plt.scatter(centers[3][0],centers[3][1], marker='*', s=200, color='black')
```

Out[45]:
```
<matplotlib.collections.PathCollection at 0x2502aa57c18>
```



In [46]:
```
myCluster = KMeans(n_clusters=4, init='k-means++', max_iter=300, n_init=10, random_sta
pred_y = myCluster.fit_predict(X)

plt.scatter(X[:,0], X[:,1])
plt.scatter(myCluster.cluster_centers_[:, 0], myCluster.cluster_centers_[:, 1], s=300
plt.show()
```