DAY 2 28-12-2016

K-Means

**1.**

**import** **pandas** **as** **pd**

**from** **sklearn.cluster** **import** KMeans

2.

kmeans\_data1 = pd.read\_csv('K\_meansHW\_TomMitchell\_data1.csv',sep = ",",header=None,names=['x','y'])

3.

kmeans\_data1.head()

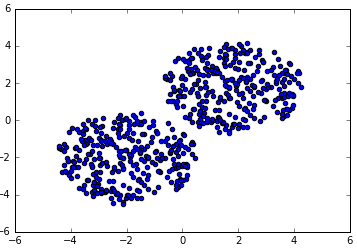
4.

%**matplotlib** inline

5.

**import** **matplotlib** **as** **plt**

6.  
plt.pyplot.scatter(kmeans\_data1['x'],kmeans\_data1['y'])



7.

kmeans\_result = KMeans(n\_clusters = 2,random\_state=0).fit(kmeans\_data1)

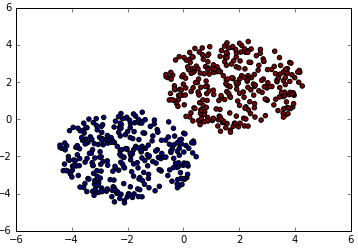
8.

kmeans\_result.labels\_

9.  
kmeans\_result

10.

plt.pyplot.scatter(kmeans['x'],kmeans['y'],c=kmeans\_result.labels\_)



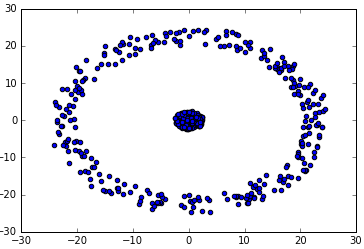
11.

kmeans\_data2 = pd.read\_csv('K\_meansHW\_TomMitchell\_data2.csv',sep = ",",header=None,names=['x','y'])

12.  
kmeans\_data2.head()

13.

plt.pyplot.scatter(kmeans\_data2['x'],kmeans\_data2['y'])



14.

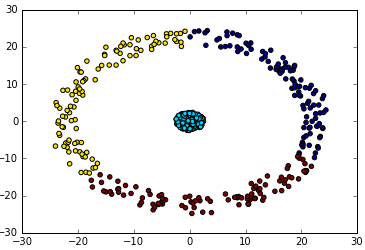
kmeans\_result2 = KMeans(n\_clusters = 4,random\_state=0).fit(kmeans\_data2)

15.

kmeans\_result2.labels\_

16.

plt.pyplot.scatter(kmeans\_data2['x'],kmeans\_data2['y'],c=kmeans\_result2.labels\_)



----------------------------------------------------------------------------------------------------------------------

Regression

-----------------------------------------

**1.**

**import** **pandas** **as** **pd**

**import** **matplotlib.pyplot** **as** **plt**

2.

train\_df = pd.read\_csv('girls\_train.csv',header = None, names=['Age','Height'])

3.

train\_df.head()

4.  
%**matplotlib** inline

5.

plt.scatter(train\_df['Age'],train\_df['Height'])

6.

train\_df['intercept'] = 1

7.

train\_df.head()