Homework #5 Due: June. 1, 2021

2021310001 이성우

Implement three different clustering methods (i.e., K-means, DBSCAN, Hierarchical) and apply it to the half-moons dataset. Discuss about the results. You may use existing frameworks or packages or Matlab functions (e.g., kmeans or cluster) for clustering.

파이썬이 더 익숙해 파이썬으로 진행하였습니다.

텍스트, 검은색, 명판이(가) 표시된 사진

자동 생성된 설명텍스트이(가) 표시된 사진

자동 생성된 설명

Data shape

Raw Data

df = pd.read\_table('half\_moons.txt', header=None)

plt.figure(figsize=(10,10))

plt.title("half\_moons\_Before", fontsize=15)

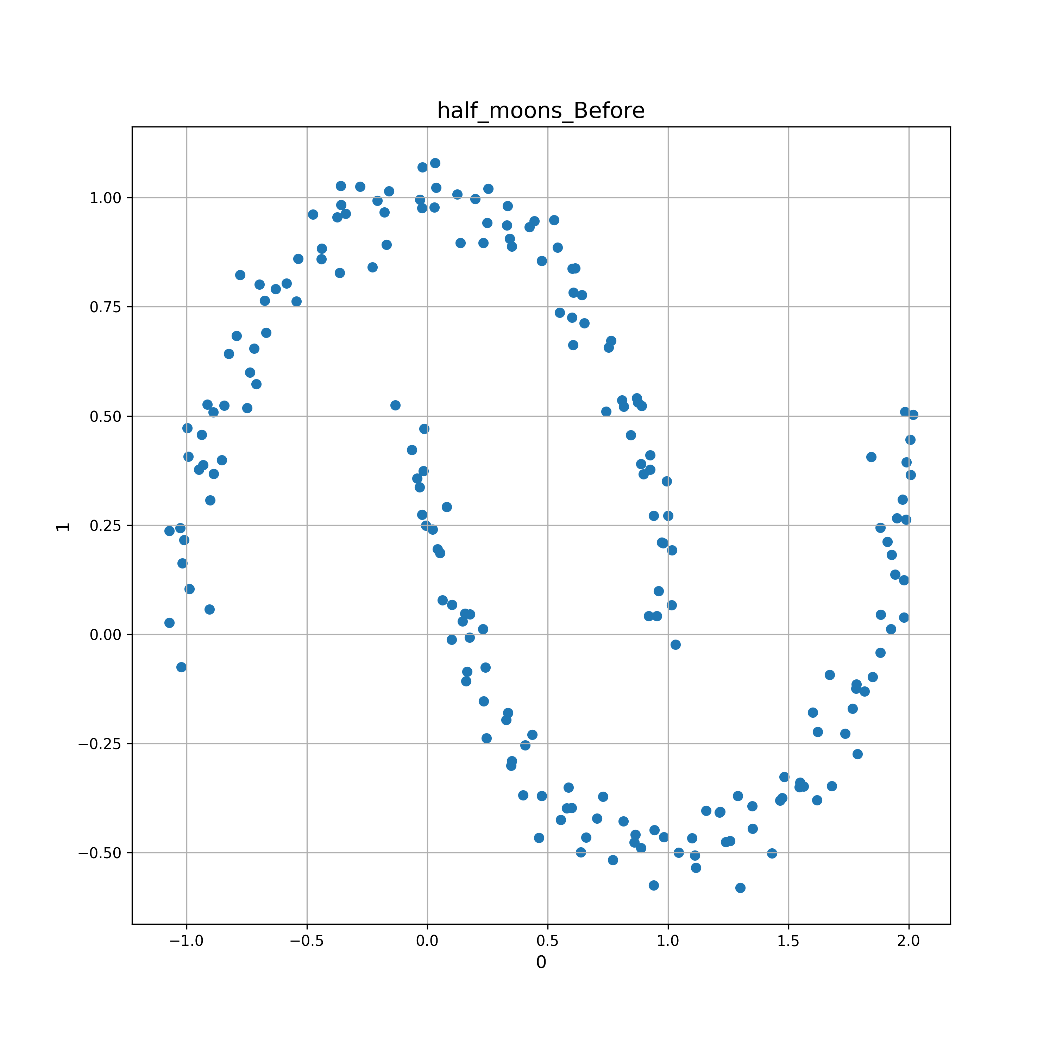
plt.scatter(df[0], df[1])

plt.xlabel("0", fontsize=12)

plt.ylabel("1", fontsize=12)

plt.grid()

plt.savefig('half\_moons\_Before', dpi=300)



**1. Kmean**

from sklearn.cluster import KMeans

kmeans\_ = KMeans(n\_clusters=2).fit(df.values)

df['cluster\_km'] = kmeans\_.labels\_

sklearn.cluster의 KMeans 패키지 사용

* 시각화

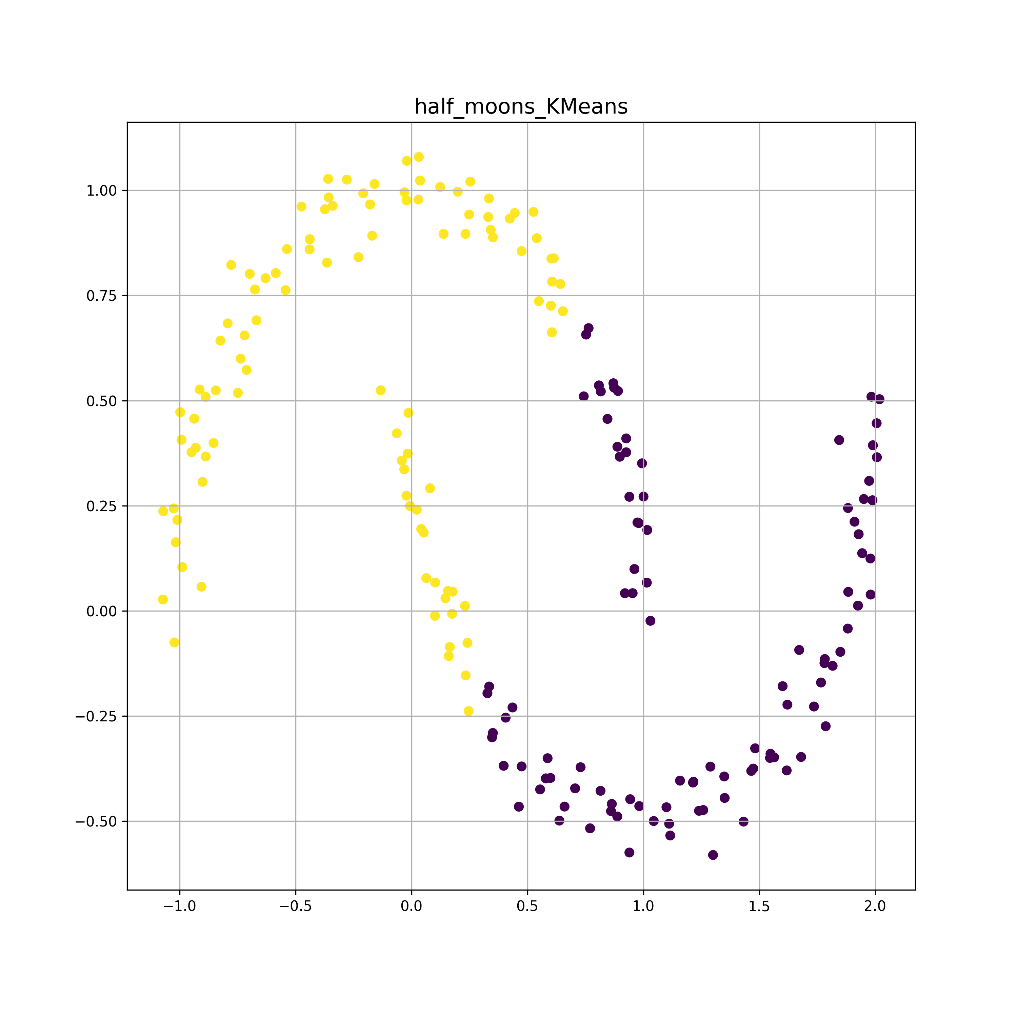
plt.figure(figsize=(10,10))

plt.title("half\_moons\_KMeans", fontsize=15)

plt.scatter(df[0],df[1],c=df['cluster\_km'])

plt.grid()

plt.savefig('half\_moons\_KMeans', dpi=300)



**2. DBSCAN**

from sklearn.cluster import DBSCAN

db\_scan = DBSCAN(eps=0.3, min\_samples=5).fit(df.values)

df['cluster\_db'] = db\_scan.labels\_

sklearn.cluster의 DBSCAN 패키지 사용

* 시각화

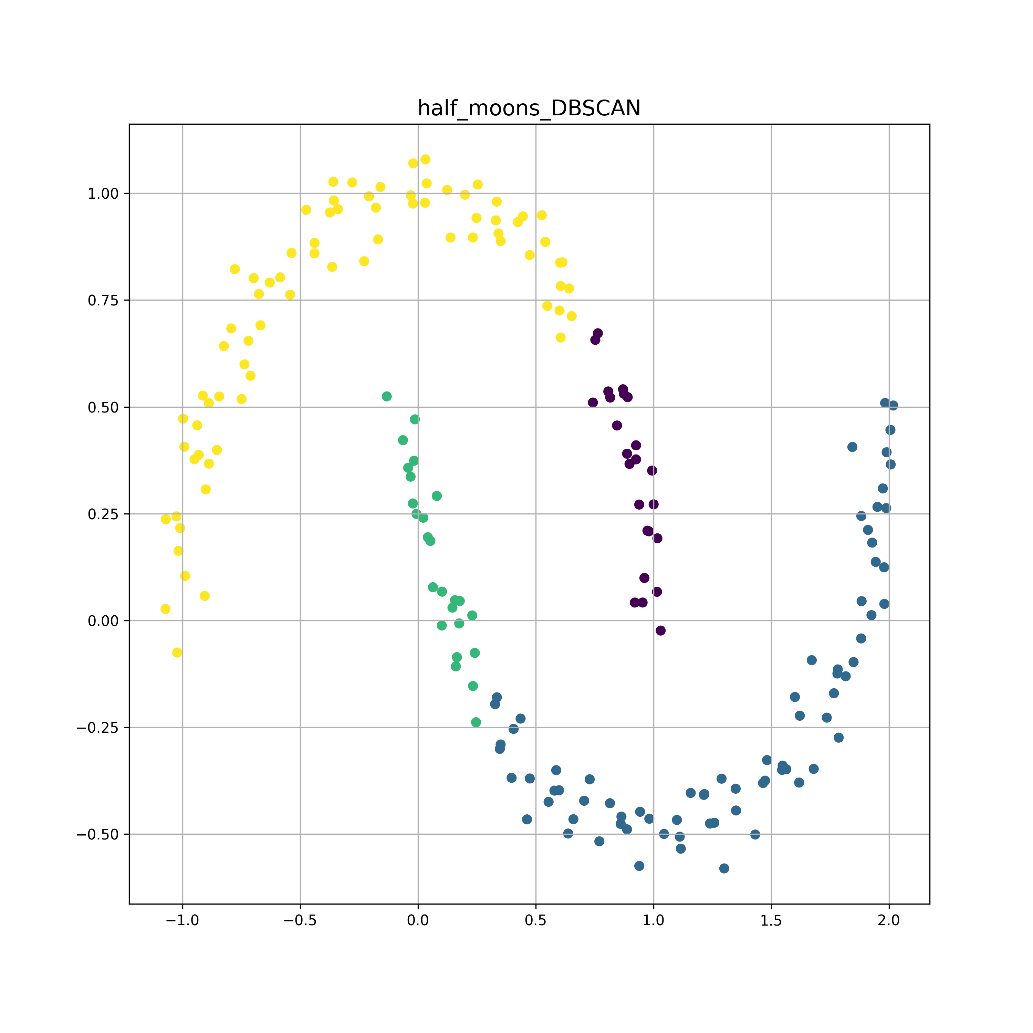
plt.figure(figsize=(10,10))

plt.title("half\_moons\_DBSCAN", fontsize=15)

plt.scatter(df[0],df[1],c=df['cluster\_db'])

plt.grid()

plt.savefig('half\_moons\_DBSCAN', dpi=300)



3. hierarchy

import scipy.cluster.hierarchy as sch

plt.figure(figsize=(10,10))

plt.title('half\_moons\_hierarchy\_순서도')

dend=shc.dendrogram(shc.linkage(df[[0, 1]],method='ward'))

df[[0, 1]].iloc[:,0]

plt.savefig('half\_moons\_DBSCAN\_hierarchy', dpi=300)

# n개 군집까지 클러스터링 하기

predict=pd.DataFrame(shc.fcluster(shc.linkage(df[[0, 1]],method='ward'),13,criterion='distance'))

predict.columns=['predict']

predict

df = pd.concat([df, predict], axis=1)

plt.figure(figsize=(10,10))

plt.title("half\_moons\_hierarchy", fontsize=15)

plt.scatter(df[0],df[1],c=df['predict'])

plt.grid()

plt.savefig('half\_moons\_hierarchy', dpi=300)

