dim=4

def calcdist(pts,oldcentroid):

ans=[]

print("ponts")

print(pts)

print("old centrpid")

print(oldcentroid)

for i in range(0,len(oldcentroid)):

dist=0;

for j in range(0,dim):

if (pts[j]!=oldcentroid[i][j]):

dist=dist+1

ans.append(dist)

print("distance")

print(ans)

return ans

def calccentroid(oldcentroid,ptscluster,pts):

ans=[]

for i in range(0,len(oldcentroid)):

ans1=[]

modea=[]

modeb=[]

modec=[]

moded=[]

last=[]

for temp in range(0,dim):

ans1.append('a')

for temp in range(0,dim):

modea.append(0)

modeb.append(0)

modec.append(0)

moded.append(0)

for j in range(0,len(ptscluster)):

if (ptscluster[j]==i):

for k in range(0,dim):

if (pts[j][k]=='a'):

modea[k]+=1

elif (pts[j][k]=='b'):

modeb[k]+=1

elif (pts[j][k]=='c'):

modec[k]+=1

elif (pts[j][k]=='d'):

moded[k]+=1

last=pts[j]

for p in range(0,dim):

set=0

if ( (modea[p]>modeb[p]) and (modea[p]>modec[p]) and (modea[p]>moded[p]) ):

ans1[p]='a'

set=1

elif ( (modeb[p]>modea[p]) and (modeb[p]>modec[p]) and (modeb[p]>moded[p]) ):

ans1[p]='b'

set=1

elif ( (modec[p]>modea[p]) and (modec[p]>modeb[p]) and (modec[p]>moded[p]) ):

ans1[p]='c'

set=1

elif ( (moded[p]>modea[p]) and (moded[p]>modeb[p]) and (moded[p]>modec[p]) ):

ans1[p]='d'

set=1

if(set==0):

ans1[p]=last[p]

ans.append(ans1)

return ans

pts=[['a','b','c','d'],

['c','d','a','a'],

['b','a','b','c'],

['a','c','d','b'],

['b','d','a','d'],

['c','c','d','b'],

# ['b','b','b','c']

]

n\_cluster=2

centroids\_old=[];

pts\_cluster=[];

for i in range(0,len(pts)):

pts\_cluster.append([0])

for i in range(0,n\_cluster):

centroids\_old.append(pts[i])

#print(centroids\_old)

centroids\_new=[]

flag=1

while (flag):

for i in range(0,len(pts)):

print("point number:")

print(i)

dist=calcdist(pts[i],centroids\_old)

index=dist.index(min(dist))

print("cluster assigned")

print(index)

pts\_cluster[i]=index

centroids\_new=calccentroid(centroids\_old,pts\_cluster,pts)

print("new centroid")

print(centroids\_new)

if (centroids\_old==centroids\_new):

flag=0

else:

centroids\_old=centroids\_new

def calcdist(pts,oldcentroid):

ans=[]

print("ponts")

print(pts)

print("old centrpid")

print(oldcentroid)

for i in range(0,len(oldcentroid)):

dist=0;

for j in range(0,4):

if (pts[j]!=oldcentroid[i][j]):

dist=dist+1

ans.append(dist)

print("distance")

print(ans)

return ans

def calccentroid(oldcentroid,ptscluster,pts):

ans=[]

for i in range(0,len(oldcentroid)):

ans1=['a','a','a','a']

modea=[0,0,0,0]

modeb=[0,0,0,0]

modec=[0,0,0,0]

moded=[0,0,0,0]

for j in range(0,len(ptscluster)):

if (ptscluster[j]==i):

for k in range(0,4):

if (pts[j][k]=='a'):

modea[k]+=1

elif (pts[j][k]=='b'):

modeb[k]+=1

elif (pts[j][k]=='c'):

modec[k]+=1

elif (pts[j][k]=='d'):

moded[k]+=1

for p in range(0,4):

if ( (modea[p]>modeb[p]) and (modea[p]>modec[p]) and (modea[p]>moded[p]) ):

ans1[p]='a'

set=1

elif ( (modeb[p]>modea[p]) and (modeb[p]>modec[p]) and (modeb[p]>moded[p]) ):

ans1[p]='b'

set=1

elif ( (modec[p]>modea[p]) and (modec[p]>modeb[p]) and (modec[p]>moded[p]) ):

ans1[p]='c'

set=1

elif ( (moded[p]>modea[p]) and (moded[p]>modeb[p]) and (moded[p]>modec[p]) ):

ans1[p]='d'

set=1

ans.append(ans1)

return ans

pts=[['a','b','c','d'],

['c','d','a','a'],

['b','a','b','c'],

['a','c','d','b'],]

# ['b','d','a','d'],

#['c','c','d','b'],

# ['b','b','b','c']]

n\_cluster=2

centroids\_old=[];

pts\_cluster=[];

for i in range(0,len(pts)):

pts\_cluster.append([0])

for i in range(0,n\_cluster):

centroids\_old.append(pts[i])

#print(centroids\_old)

centroids\_new=[]

flag=1

while (flag):

for i in range(0,len(pts)):

print("point number:")

print(i)

dist=calcdist(pts[i],centroids\_old)

index=dist.index(min(dist))

print("cluster assigned")

print(index)

pts\_cluster[i]=index

centroids\_new=calccentroid(centroids\_old,pts\_cluster,pts)

print("new centroid")

print(centroids\_new)

if (centroids\_old==centroids\_new):

flag=0

else:

centroids\_old=centroids\_new