

«به نام خدا»



هوش مصنوعی و محاسبات زیستی – دکتر حاجی پور
ترم 1402_01

پرnian طاهری 99106352

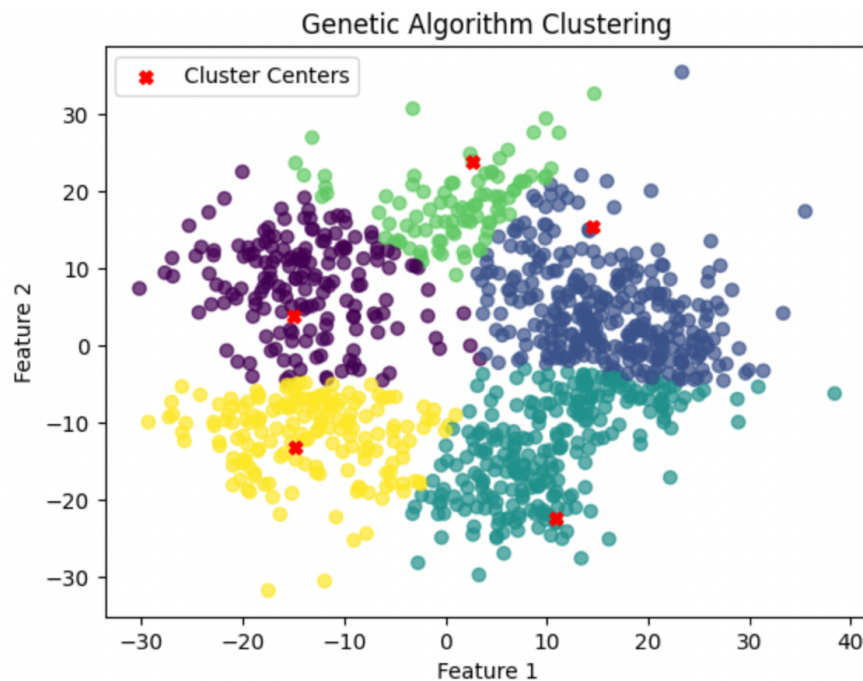
تمرین کامپیوتری 3 – سوال 3

Q1, Q2:

In these parts I used “deap” library to write an GA. In the GA, genes are the location of centroids so each chromosome has 10 genes.

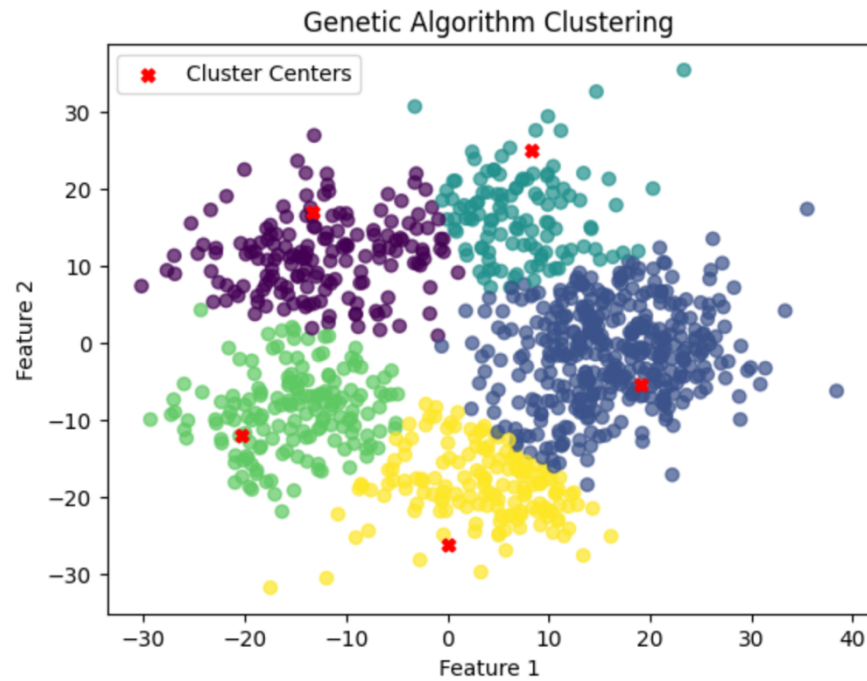
Q1:

In this part, in order to minimize the intra-cluster-distance (distance between data inside each cluster) I calculated pairwise distance. However, the result would be 5 centroids in the middle, therefore, I included the inter-cluster-distance (between centroids) to maximize the distance between each cluster. And also, there is no constraints on the number of clusters.



Q2:

In this part, in order to satisfy the constraint, which is exactly 5 clusters, I used a correction term, which when the length of the chromosome is less than 5, it picks random numbers of data from the cluster with the maximum size of data and gives it a label, which does not exist.

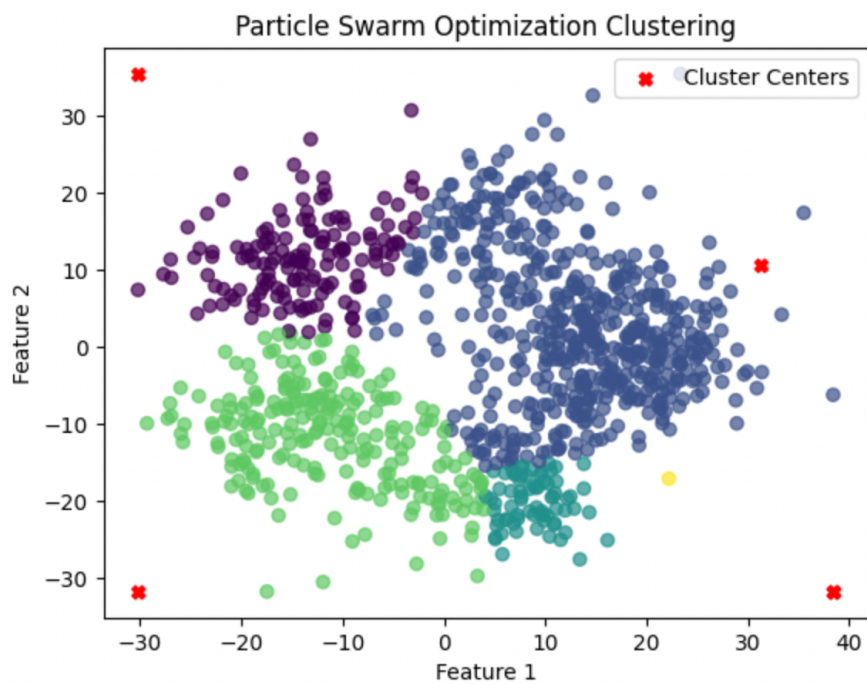


Q3, Q4:

For these parts I used “pyswarm” library, which runs a PSO algorithm.

Q3:

For this part, the fitness function is the same as part Q1 and it does not necessarily generate 5 clusters.



Q4:

In this part, I used a different fitness function. Not only I used a correction term as said in Q2, but also, I used a punishment term, which adds a large number to the fitness function whenever the length of the chromosome is less than 5 or any of clusters have data less than a threshold.

