**1. The Clustering Part**

**Input:** Data matrix A of dimension n x m, where n is the number of genes and m is the number of samples or experiments, the value of number of clusters K, algorithm for K-means clustering and maximum number of iterations for K-means.

**Algorithm:**

1. Cluster the dataset into K different clusters using the given algorithm.

2. Generate a list of K elements L, where each element contains all the data points assigned to a cluster. This list is used in the Cluster Merging and ERT steps.

3. Calculate a distance matrix D of dimension K x K where distance between each pair of cluster centers is stored.

**Output:** L, list of K elements and distance matrix D

**2. Entropy Reduction Technique (ERT) Part**

**Input**: Cluster ID

**Algorithm:**

1. Collect data points n of the given cluster ID from the list L generated in Clustering Part

2. Generate data matrix TD of the genes from the main data matrix A where columns contain genes and rows contain samples.

3. Discretize TD.

4. Calculate Mutual Information (MI) matrix M of dimension n x n and normalize the mutual information in an n x n dimensional matrix NMI using Linfoot definition of normalization to have the mutual information values in the range of 0 to 1 [2].

**M [i,j] = MutualInformation(TD[,i], TD[,j])**

5. Calculate the single entropy matrix E of dimension n.

**E[i] = Entropy (TD [,i])**

6. Calculate an n x n dimensional conditional entropy matrix CE between all pair of variables.

**CE [i,j] <- ConditionalEntropy(TD[,i],TD[,j])**

7. Calculate an n x n dimensional Reduced Entropy matrix RE between each pair of variables i,j using mutual information matrix M and single entropy matrix E using the equation,

**RE [i,j] = (M[i,j]) / E[i]**

8. Generate an n x n dimensional ERT matrix ERTM using the following condition,

**If CE [i,j] = E[i] , then ERTM [i,j] = 1**

**Else ERTM [i,j] = 0**

9. Generate a connection matrix C of n x n dimension using the following condition,

**If ERTM [i,j] == 1, then** **C [i,j] = RE [i,j]**

**Else C [i,j] = 0**

**Output**: Connection matrix C of reduced entropy values.

**3. Calculate Nearby clusters**

**Input:** Distance matrix of cluster centers

**Algorithm:**

For each cluster

1. Find the distance of closest cluster (distance between two centers)

2. Multiply the distance by 2

3. Find the clusters whose centers are closer than the multiplied distance.

**Output: Vector of nearby clusters for each cluster**

**4. Apply ERT on all Merged Clusters**

**5. Merge all the connection matrices returned from ERT steps to generate Final Connection Matrix.**