

## Theoretical Explanation

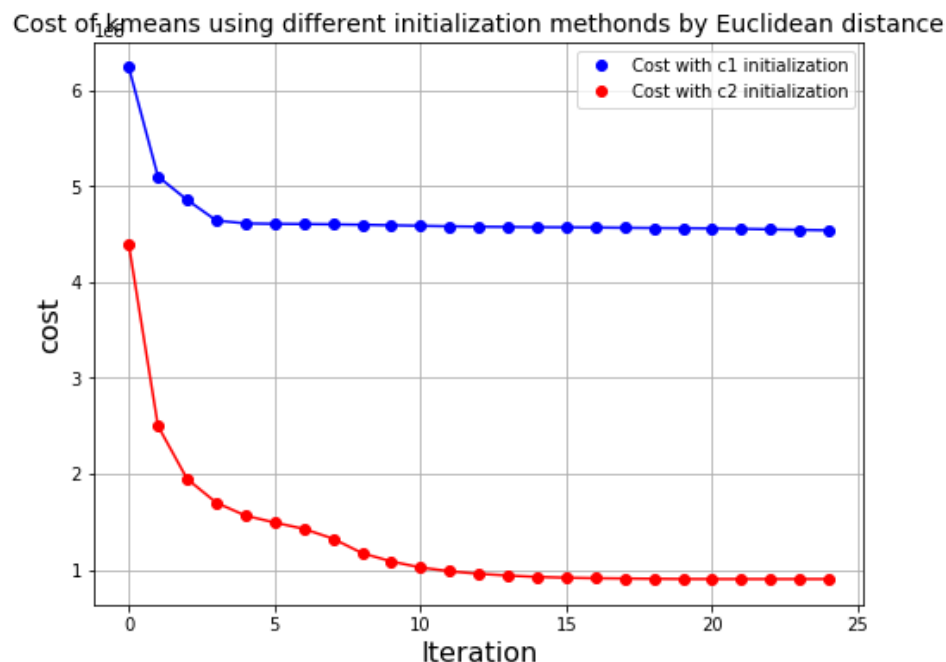
K-means clustering is an iterative method that minimizes the cost function for each iteration. (Like a gradient descent algorithm)

The cost function is;

$$J(c_k) = \sum_{x_i \in c_k} ||x_i - \mu_k||^2$$

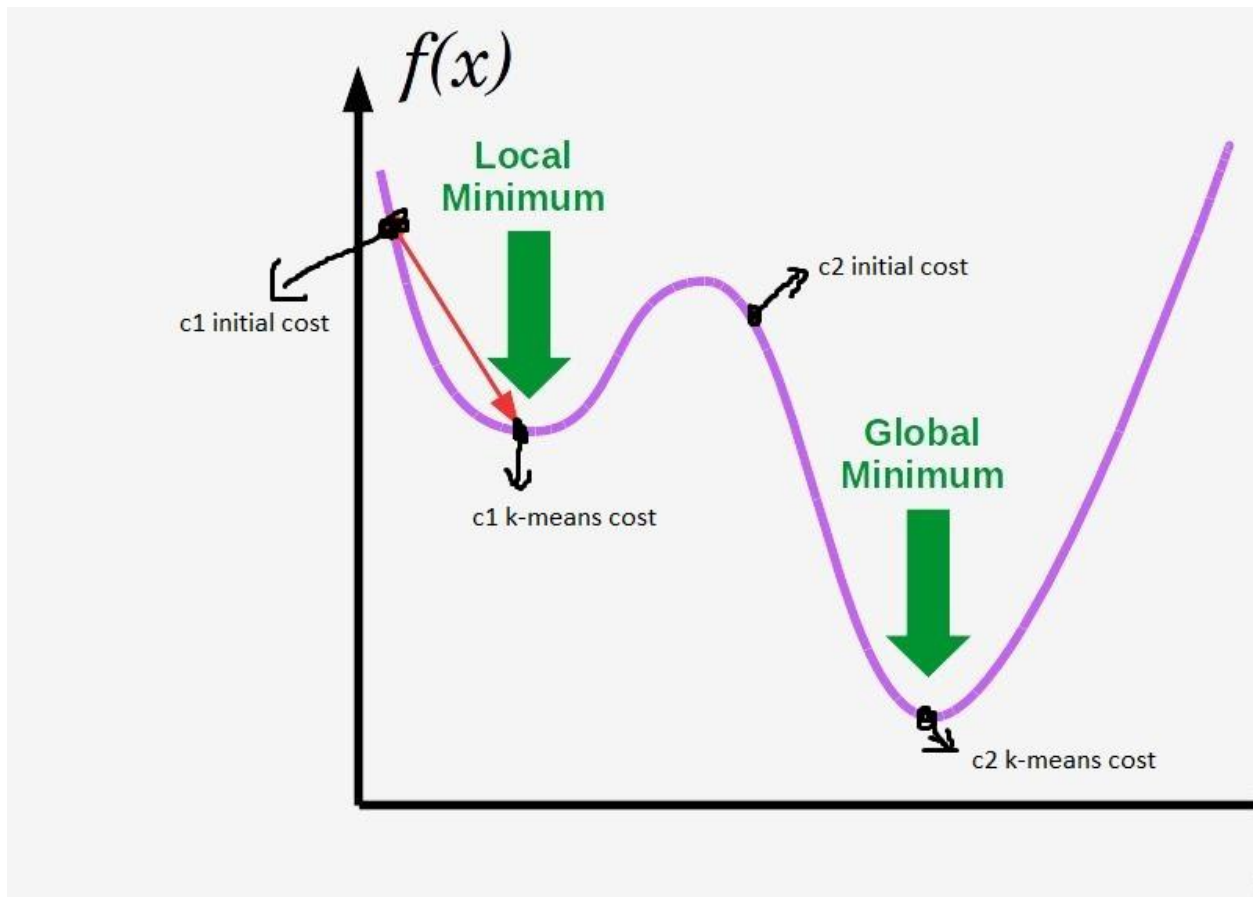
So, after the initialization of the cluster centers, the algorithm will perform a greedy approach to reduce the cost. This greedy approach makes the algorithm dependent on the cost of initial centers and algorithm sticks at the local minimum. In other words, for each different set of initial centers, there will a different cost.

Like in the graph bellow,



As seen, depending on the initial centers, the k-means algorithm can return different costs. So, a bad choice of initial centers will affect the performance of the algorithm.

The graph below shows how k-means clustering can get stuck in local optimal.



To avoid this problem, metaheuristic algorithms like simulated annealing or Genetic algorithm can be used to select the initial centers. But to find optimal solution, the cost of all possible initial centers should be calculated. Which will take exponential amount of time and in most of the cases, the set of nodes (centers) is a continuous set. So it is impossible to try all possible points.

