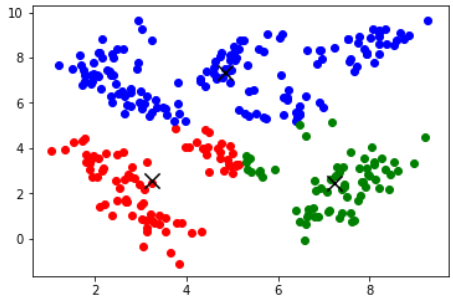
PROJECT 2 REPORT

Islam Hamed

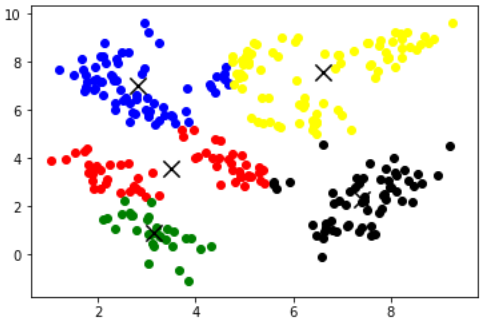
1221217088

1- K-means-Strategy1:

- For k=3
  - Initial points are:
    - [3.24516611, 0.8218365]
    - [2.61234619, 8.39116666]
    - [8.9702889, 3.32150578]
  - Final centroid points are:
    - [3.24896423, 2.58027691]
    - [4.83375318, 7.31605824]
    - [7.23975119, 2.48208269]
  - Loss is 1338.1059838029246

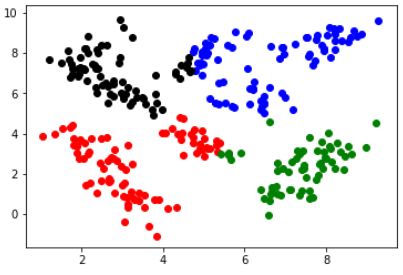


- For k=5:
  - Initial points are:
    - [1.87131855, 3.43365823]
    - [3.9649361, 5.20027567]
    - [3.40504475, 1.04980673]
    - [4.74625798, 3.54661053]
    - [6.12393256, 5.49223251]
  - Final centroid points are:
    - [3.49556658, 3.56611232]
    - [2.81706606, 7.010913]
    - [3.14506148, 0.90770655]
    - [7.41419243, 2.32169114]
    - [6.60345839, 7.57042104]
  - Loss is 649.926657048097

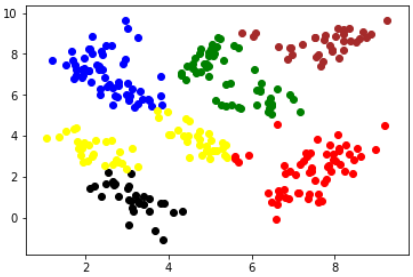


2- K-means-Strategy2:

- For k=4
  - Initial point is [2.04945194, 2.75937105]
  - Final centroid points are:
    - [3.339957483138508, 2.5921522375769444]
    - [6.603458393504191, 7.570421042158782]
    - [7.380762638700798, 2.332455315679148]
    - [2.8585923471789103, 6.931365250947319]
  - Loss is 788.2693490065562



- For k=6
  - Initial point is [6.39627447, 1.24125663]
  - Final centroid points are:
    - [7.414192434680615, 2.3216911383868664]
    - [2.5633381461259046, 6.978224800606624]
    - [5.464277356727894, 6.837713536435891]
    - [3.145061482959145, 0.9077065486588153]
    - [3.4955665791995627, 3.5661123157286907]
    - [7.756483249146484, 8.556689279063415]
  - Loss is 476.11875167635293



The compare can be between number of clusters, so in both strategy number of cluster affects the loss function or the cost function. However, the comparison between two strategy is not accurate unless we have the same number of clusters. So, after running strategy 2 on k = 5 the loss function is 653.956790 8403947 which means strategy 1 works better here in k = 5.