MP3 Hanna Puhachova

Q1. Mean shift and K-Means algorithm are two similar clustering algorithms; both of them extract information from data with some kind of mean vector operations. The K-means algorithm Iteratively aims to group data samples into K clusters, where. each sample belongs to the cluster with the nearest mean. The mean shift algorithm is a non-parametric algorithm that clusters data iteratively by finding the densest regions (clusters) in a feature space. The difference between K-Means algorithm and Mean-Shift is that later one does not need to specify the number of clusters in advance because the number of clusters will be determined by the algorithm w.r.t data.

Q2. Iteration 1. Initial centroids 3.0, 4.0. Since we have 1D we need to simply |x-y.

Data Points	Dist Mean1 3.0	Dist Mean 2 4.0	Cluster
0	3.0	4.0	1
1	2.0	3.0	1
2	1.0	2.0	1
3	0	1.0	1
4	1.0	0	2

Cluster 1:{0,1,2,3} and Cluster 2:{4}

New centroids: Cluster 1 = (0+1+2+3)/4=1.5; Cluster 2= 4.0

Iteration 2:

Data Points	Dist Mean1 1.5	Dist Mean 2 4.0	Cluster
0	1.5	4.0	1
1	0.5	3.0	1
2	0.5	2.0	1
3	1.5	1.0	2
4	2.5	2.0	2

Cluster 1:{0,1,2}

Cluster 2:{3,4}

New centroids: Cluster 1 = (0+1+2)/3=1.0; Cluster 2= (3+4)/2=3.5

Q3.