# ECE421 - Winter 2022
# Homework Problems - Tutorial #7

*Theme: K-means algorithm*

Due: March 13, 2022 11:59 PM

## Question 1 (Problem 6 - Final Exam 2018)

Consider the $K$-means algorithm. Let $K = 2$ and let $\mathcal{D}$ be a dataset consisting of four data points with $\mathcal{D} = \{0, 0.5, 0.5 + \Delta, 1.5 + \Delta\}$, where $\Delta \geq 0$ is a problem parameter. All data points lie on the real line.

(a) For this part, let $\Delta = 0.5$ and initialize $K$-means by initializing the two cluster centers at $\mu_1[0] = 1$ and $\mu_2[0] = 2$. Run $K$-means till convergence. For each iteration $l$ until convergence, describe your set membership $\{\mathcal{B}_1[l], \mathcal{B}_2[l]\}$ and cluster centers $\{\mu_1[l], \mu_2[l]\}$. Make sure you identify the final values of the cluster centers and set membership at convergence.

(b) For this part, find a condition that $\Delta$ <u>must satisfy</u>, such that $\Delta$ has a <u>small</u> positive value, and $K$-means (initialized in the same manner as in part (a), i.e., $\mu_1[0] = 1$ and $\mu_2[0] = 2$) converges to a different solution from that obtained in part (a). In your solution, describe,

   (i) What is this condition on $\Delta$ and explain your reasoning/derivation.

   (ii) As in part (a), run the cluster algorithm, describe the values of cluster centers and set membership for each iteration until convergence.