Quiz 1

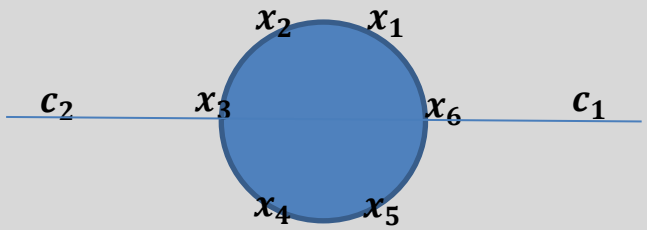
(Write answers either in Korean or English)

Problem 1) Please describe the key difference between ‘Supervised learning’ and ‘Unsupervised learning’.

(Problem 2-4) There are six data in the 2-dimensional space: $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \dots \boldsymbol{x_6}\}$. For the *k*-th data $\boldsymbol{x_k}$, its coordinate is assigned as follows:

$$\boldsymbol{x_k} = \begin{bmatrix} \cos\left(\frac{\pi}{3}k\right) \\ \sin\left(\frac{\pi}{3}k\right) \end{bmatrix}$$

We applied K-means clustering algorithm using K=2 and Euclidean distance measure on these six data by initializing two cluster centroids as $\boldsymbol{c_1} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\boldsymbol{c_2} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$ ($\boldsymbol{c_1}$ is the centroid of cluster 1, $\boldsymbol{c_2}$ is the centroid of cluster 2). The figure below describes our 2D feature space:



Problem 2) Please discuss after the first iteration of K-means clustering, how six data $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \dots \boldsymbol{x_6}\}$ are assigned to clusters?
Problem 3) What are the updated cluster centroids $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$, after the first iteration?
Problem 4) The decision rule for the SVM is: $\mathrm{y} = sign(\boldsymbol{w^T x} + b)$, where sign() is the function that returns the sign (+/-) of the input. When $\boldsymbol{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $b = 0$ are obtained from the training stage of SVM, what is the binary label ($y_i = +1$ or -1) for $\boldsymbol{x_3}$ and $\boldsymbol{x_6}$?

Name: ID: