Place Recognition / Deep Learning

1 Numerical Exercises

1. Consider the clustering of the following points in $\mathbb{R}^2$ using the k-means clustering, where $k = 2$.

x_1	0	0
x_2	0	1
x_3	-1	2
x_4	2	0
x_5	3	0
x_6	4	-1

Table 1: Datapoints

(a) In a first step, compute the squared distance matrix $D_{ij} = dist_{eucl.}(\mu_i, x_j)^2$ between the datatpoints x_j and the initial cluster centers μ_i. Assume that the first and last datapoint are the initial centers.
Solution

	x_1	x_2	x_3	x_4	x_5	x_6
$\mu_1 = x_1$	0	1	5	4	9	17
$\mu_2 = x_6$	17	4	34	5	2	0

(b) Based on the distance matrix D_{ij}, perform one iteration (cluster assignment and center update) of the k-mean clustering algorithm. **Solution**
Based on the distance matrix D_{ij}, the two cluster centers μ_1 and μ_2 have the following assignment $P_1 = \{x_1, x_2, x_3, x_4\}$ and $P_2 = \{x_5, x_6\}$.
The updated cluster centers μ'_1 and μ'_2 are computed by taking the mean of the corresponding assignment.

$$\mu'_1 = \frac{1}{|P_1|} \sum_{x \in P_1} = \frac{1}{4}(\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}) = \frac{1}{4}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\mu'_2 = \frac{1}{|P_2|} \sum_{x \in P_2} = \frac{1}{2}(\begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \end{bmatrix}) = \frac{1}{2}\begin{bmatrix} 7 \\ -1 \end{bmatrix}$$

2. Consider the following query image with the set of visual words Q and the image vocabulary V. Using the image retrival method presented in the lecture, construct the voting array and state which image (A, B, C or D) is the closest to the query image.

$$Q = \{1, 2, 3, 4\}$$

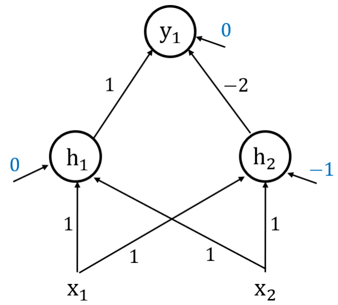
$$V = \{1 = \{A, B\},\ 2 = \{A, B, C\},\ 3 = \{C\},\ 4 = \{A, B, C, D\}\}$$

Solution
To construct the voting array, it is important to weight the votes by the inverse of the frequency of a visual word. The resulting voting array looks as follow:

A	B	C	D
$\frac{1}{2}+\frac{1}{3}+\frac{1}{4}$	$\frac{1}{2}+\frac{1}{3}+\frac{1}{4}$	$\frac{1}{3}+1+\frac{1}{4}$	$\frac{1}{4}$

Thus, image C is the image closest to the query image.

3. Consider the following MLP with the black numbers above the edges representing the weights and the blue numbers above the arrows the biases. All activations are ReLU function, i.e., $f(x) = max(0, x)$. Compute the hidden activations $h_1$and h_2 and output y_1 for the following inputs to the network.



(a) $x_1 = 0$ and $x_2 = 0$
Solution
Based on the structure, the network activations and output can be computed using the following formulas
$h_1 = max(0, 1 \cdot x_1 + 1 \cdot x_2 + 0)$
$h_2 = max(0, 1 \cdot x_1 + 1 \cdot x_2 - 1)$
$y_1 = max(0, 1 \cdot h_1 - 2 \cdot h_2 + 0)$
Thus, for $x_1 = 0$ and $x_2 = 0$ we get
$h_1 = max(0, 1 \cdot 0 + 1 \cdot 0 + 0) = 0$
$h_2 = max(0, 1 \cdot 0 + 1 \cdot 0 - 1) = 0$
$y_1 = max(0, 1 \cdot 0 - 2 \cdot 0 + 0) = 0$

(b) $x_1 = 1$ and $x_2 = 0$
Solution
$h_1 = max(0, 1 \cdot 1 + 1 \cdot 0 + 0) = 1$
$h_2 = max(0, 1 \cdot 0 + 1 \cdot 1 - 1) = 0$
$y_1 = max(0, 1 \cdot 1 - 2 \cdot 0 + 0) = 1$

(c) $x_1 = 0$ and $x_2 = 1$
Solution
$h_1 = max(0, 1 \cdot 0 + 1 \cdot 1 + 0) = 1$
$h_2 = max(0, 1 \cdot 1 + 1 \cdot 0 - 1) = 0$
$y_1 = max(0, 1 \cdot 1 - 2 \cdot 0 + 0) = 1$

(d) $x_1 = 1$ and $x_2 = 1$
Solution
$h_1 = max(0, 1 \cdot 1 + 1 \cdot 1 + 0) = 2$
$h_2 = max(0, 1 \cdot 1 + 1 \cdot 1 - 1) = 1$
$y_1 = max(0, 1 \cdot 2 - 2 \cdot 1 + 0) = 0$

(e) For the above binary inputs, what function does this MLP aproximate?
Solution
The MLT approximates the XOR (Exclusive or) operation.