Introduction To Machine Learning **Spring 2024**

Exercise 1

Sharon Rotgaizer srotgaizer@gmail.com

Due: 28.05.2024

Questions 1-3 - Programming assignment

See attached colab file in Moodle.

Question 4 - Parzen Windows

In Parzen windows we estimated only the likelihood of each class. However, given prior, we can fully perform classification. Such classification relies on MAP (which is the Bayesian decision rule under zero-one risk-loss function).

Note: Both sections of this question are not related to each other.

1. Show that the full classification can be reduced to the following rule: "To a given sample, assign the category which gets majority vote of neighbors". Recall that using window functions, majority vote of neighbors can be written as:

$$\forall j \neq i \in \{1, \ldots, C\} : \sum_{k=1}^{n_i} \phi\left(\frac{x_k^i - x}{h}\right) \geq \sum_{k=1}^{n_j} \phi\left(\frac{x_k^j - x}{h}\right) \tag{1}$$

   Where x is a test sample, C is the number of classes and n_i is the number of samples in class. The notation x_k^i means the k-th sample from the i-th class.

   Hint: Start from $\forall j \neq i : \mathbb{P}(w_i|x) \geq \mathbb{P}(w_j|x)$ and use Bayes rule to derive the rule in eq. 1.

2. You are given some samples from unknown distribution, $D = \{1, -3, 2, 4, 5, -8, 0, -1, -2, -4\}$.

   Given $h = 4$, use the 0-1 window function and build the full distribution. Attach only the drawn graph (the pdf function). No coding required.

   (a) According to the graph, what is the name of the real distribution of the data?

   (b) Use MLE to estimate the distribution parameters and attach the drawing of the real distribution. Show the similarity between both graphs.

   Note: You can use desmos to draw the pdf in last section.