Mathematical Foundations of Machine Learning, Spring 2025: Assignment I

1. [30%] Let $f(x) := \sum_{m=1}^{M} c_m \mathbf{1}_{[2m,2m+1]}(x)$. Compute the modulus $\omega_1(f,t)_p$, for all $0 < t < 1/2$, and $0 < p \le \infty$.

Remark: There are two cases: $0 < p < \infty$, $p = \infty$.

2. [30%] Prove the following equality for any $N \ge 1$, $x, h \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$,

$$\Delta_{Nh}^{r}(f,x) = \sum_{k_1=0}^{N-1} \cdots \sum_{k_r=0}^{N-1} \Delta_h^r(f, x + k_1 h + \ldots + k_r h).$$

Hint: recall we proved in class for $r = 1$. Now apply induction on r. Make sure the notations are correct.

3. [30%] Support Vector Machines – Write the details of the explanation and the derivation of the functional L_D in (12.13) from the functional L_p in (12.9).

4. [10%] How would you speed up the construction during training of a Random Forest composed of 5 trees using 20 parallel processors? Try to describe an optimal scenario where the 20 processors are fully utilized all the time.