The following homework will be discussed on June 8th to 10th:

I. Neural networks - representation lower bounds

1. In the lecture we proved that every Boolean function $f : \{0,1\}^d \to \{0,1\}$ can be exactly represented by a neural network with 2^d neurons in a single hidden layer and $\sigma(z) = \mathbb{1}_{z\geq 0}$. Let $\nu(d)$ be the minimal number of neurons that is required to achieve such a representation by any multilayerd feedforward neural network using this activation function. Use the VC-dimension to show that ν has to grow exponentially with d.

2. Give an example of a Boolean function $f : \{0,1\}^d \to \{0,1\}$ that requires exponentially many neurons to be represented using a multilayerd feedforward neural network.

II. Rademacher complexity and neural networks

1. Prove Lemma 1 below.

2. Prove Lemma 2 below.

3. Consider the class of feedforward neural networks with d inputs, a single layer of m neurons, a fixed activation function ϕ that is L-Lipschitz and satisfies $\phi(0) = 0$, weights $w_1, \ldots, w_m \in \mathbb{R}^d$ for the hidden neurons that satisfy $||w_i||_1 \leq a$ and weights $u_1, \ldots, u_m \in \mathbb{R}$ of the output neuron that satisfy $||u||_1 \leq b$. Assume further that the inputs $x \in \mathbb{R}^d$ are bounded by $||x||_\infty \leq c$.

   Compute an upper bound on the corresponding Rademacher complexities $\mathcal{R}_n$ in terms of n, L, a, b and c. (Hint: use Lemma 1 and Lemma 2)

4. Both, the Rademacher complexities estimated here and the VC-dimension estimated in the lecture lead to PAC bounds for neural networks. Compare them!

Lemma 1: Let A be a finite subset of $\mathbb{R}^m$ that is contained in a Euclidean ball of radius r. Then it holds that

$$\mathbb{E}_\sigma \left[\max_{a \in A} \sum_{i=1}^m \sigma_i a_i \right] \leq r\sqrt{2 \ln |A|}, \tag{1}$$

where $\sigma \in \{-1,1\}^m$ uniformly.

Lemma 2: Let $\phi : \mathbb{R} \to R$ be L-Lipschitz and satisfy $\phi(0) = 0$. Then for any $A \subseteq \mathbb{R}^m$:

$$\mathbb{E}_\sigma \left[\frac{1}{2} \sup_{a \in A} \left| \sum_{i=1}^m \sigma_i \phi(a_i) \right| \right] \leq \mathbb{E}_\sigma \left[L \sup_{a \in A} \left| \sum_{i=1}^m \sigma_i a_i \right| \right], \tag{2}$$

where $\sigma \in \{-1,1\}^m$ uniformly.