



The following homework will be discussed on June 26th and June 29th:

H10.1. (Pseudo dimension of neural networks) Let N be any neural network with a single real-valued output unit, and form a neural network N' as follow. The network N' has one extra input unit and one extra computation unit. This additional computation unit is a linear threshold unit receiving input only from the output unit of N and from the new input unit, and it is the output unit of N' . If H' is the set of $\{0, 1\}$ -valued functions computed by N' and F the set of real-valued functions computed by N what is the relation between $\text{Pdim}(F)$ and $\text{VCdim}(H')$?

H10.2. (Memorization capacity)

- a) For the AlexNet 2012 (see Exercise H8.2b) represented by $f_\omega : \mathcal{X} \rightarrow \mathcal{Y}$, where ω is the vector of parameters, provide an upper bound on the number N such that for all $(x_i, y_i)_{i=1}^N \in (\mathcal{X} \times \mathcal{Y})^N$ there is a choice of parameters $\tilde{\omega}$ such that $f_{\tilde{\omega}}(x_i) = y_i$ for all $i \in \{1, \dots, N\}$.
- b) Compare this number with the number of training data points used.

H10.3. (Approximation by Neural networks)

- a) Show by explicit construction that for every f in the unit ball of $W^{1,\infty}([0, 1])$ and every $\epsilon > 0$ there is a \tilde{f} in this unit ball so that:
 - (i) the function \tilde{f} can be implemented by a feedforward neural network with a single hidden layer of $\lceil \frac{1}{\epsilon} \rceil$ neurons with ReLU-activation function and identity-activation at the output,
 - (ii) $\|f - \tilde{f}\|_\infty \leq \epsilon$.
- b) Do the parameters in your construction of \tilde{f} depend continuously on f ? What does this imply concerning the quality of the approximation?