MDL - TD3

06/02/2024

Exercise 1: Depth vs. expressivity of ReLU networks

We want to show that any MLP $\mathbb{R}^d \to \mathbb{R}$ with ReLU activation function represents a piece-wise linear continuous function, and that any piece-wise linear function $\mathbb{R}^d \to \mathbb{R}$ can be represented by an MLP with depth at most $\lceil \log(d+1) \rceil + 1$.

Q1: Construct a 2-layered NN f such that $f(x, y) = \max(x, y)$.

Q2: For $f_1, ..., f_K$ functions that can each be represented by w_i-layered neural networks, show that $f = \max(f_1, \dots, f_K)$ can be represented as a neural network with depth at most $\max(w_1, ..., w_K) + \lceil \log_2(K) \rceil$.

Q3: Conclude that any continuous piece-wise linear function f on $\mathbb{R}$ can be expressed as an MLP with depth at most $\lceil \log_2(d+1) \rceil + 1$. *Hint: there exists* $s_1, \dots, s_N \in \{-1, 1\}$, $S_1, \dots, S_N \subset [\![1, K]\!]$ *and* $(\ell_i)_{i \in [\![1,K]\!]}$ *affine functions such that* $|S_j| \leq d + 1$ *and* $f(x) = \sum_{j=1}^{N} s_j \max_{i \in S_j} \ell_i(x)$.

Exercise 2: Approximation of real functions

Let $\sigma(x) = \max(0, x)$ for any $x \in \mathbb{R}$. We want to prove that any continuous function $f : \mathbb{R} \to \mathbb{R}$ can be approximated on any compact set, at any given precision, by a 2-layered neural net (1 hidden layer) with activation function σ.

Q4: Approximate with a precision $\varepsilon > 0$ the function $f(x) = x^2/2$ on $[0, 1]$.

Q5: Approximate with a precision $\varepsilon > 0$ the function $f(x) = x^n/2$ on $[0, 1]$ for any $n \in \mathbb{N}$.

Q6: Approximate any continuous function on $[0, 1]$ by an MLP.

Q7: If $f : [0, 1] \to \mathbb{R}$ is L-Lisphitz, how many neurons are required for a given precision $\varepsilon > 0$?