MDL - TD4

13/02/2024

Exercise 1: Sawtooth and the power of depth

For L, W integers and σ the ReLU function, we denote by $\mathcal{N}(\sigma, L, W)$ the set of MLPs with depth at most L and width at most W. The classification power of a class of functions $\mathbb{R}^d \to \mathbb{R}$ is measured by the classification error, given a set of n points $z = \big(x_i, y_i\big)_{1\leq i\leq n}$:

$$\mathcal{R}_z(f) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{y_i \cdot f(x_i) > 0\}\,,$$

where $y_i \in \{-1, 1\}$. The classification power of a class of functions $\mathcal{F}$ is then defined as $\min_{f\in\mathcal{F}} \mathcal{R}_z(f)$.

Q1: For $p \in \mathbb{N}$, construct a function $f_p : [0, 1] \to \mathbb{R}$ that is piecewise linear with 2^p affine regions, that takes $2^{p-1} + 1$ times the value 0 and 2^{p-1} times the value 1 in $[0, 1]$, and that can be expressed as a neural net with depth $p + 1$ and width 2.

Q2: Use the above result to show that, for $n = 2^p$ and $z = \big(\frac{i}{n}, (-1)^i\big)_{0\leq i\leq n} \in \big([0, 1] \times \{-1, 1\}\big)^n$,

$$\min_{f\in\mathcal{N}(\sigma,p+1,2)} \mathcal{R}_z(f) = 0\,.$$

Q3: Let $f \in \mathcal{N}(\sigma, L, W)$ for $L, W \in \mathbb{N}^*$. Upper bound the number of affine regions of f.

Q4: Show that, for any L, W,

$$\min_{f\in\mathcal{N}(\sigma,L,W)} \mathcal{R}_z(f) \geq \frac{1}{2} - \frac{(W + 1)^{L-1}}{2n}\,.$$

Q5: Conclude in terms of power of depth for neural networks.

Exercise 2: Counting triangles with graph neural networks

We recall that a graph neural network (GNN) is a model that takes an undirected graph $G = (V, E)$ with node set V and edge set E as input and maps it to a vector $u_G \in \mathbb{R}^d$ using an iterative procedure composed of three steps:

1. **Initialization:** For each node $i \in V$, we initialize the node attributes $u_{i,0} = 0 \in \mathbb{R}^d$.
2. **Aggregation:** For each node $i \in V$ and model layer $l \in \{1, \ldots, L - 1\}$,

$$u_{i,l+1} = \phi_l\left(u_{i,l}, \{u_{j,l}\}_{\{i,j\}\in E}\right)\,.$$

3. **Readout:** $u_G = \psi\left(\{u_{i,L}\}_{i\in V}\right)$.

where the functions $\phi_1, \ldots, \phi_L$ and ψ are permutation invariant neural networks.

Q6: Give a simple example for the functions $\phi_1, \ldots, \phi_L$ and ψ that count the number of nodes of the input graph.

Q7: Give a simple example for the functions $\phi_1, \ldots, \phi_L$ and ψ that count the number of edges of the input graph.

Q8: Show that a GNN (as defined above) cannot count the number of triangles of a graph, i.e. there exists graphs G, G' such that their number of triangles differ and, for any GNN f, we have $f(G) = f(G')$.

Q9: Conclude on the universality of GNNs.