Neural Networks in Lean4

mkaratarakis

May 13, 2025

Chapter 1

NN

Definition 1. An (artificial) neural network is a directed graph $G = (U, C)$, where neurons $u \in U$ are connected by directed edges $c \in C$ (connections). The neuron set is partitioned as $U = U_{\text{in}} \cup U_{\text{out}} \cup U_{\text{hidden}}$, with $U_{\text{in}}, U_{\text{out}} \neq \emptyset$ and $U_{\text{hidden}} \cap (U_{\text{in}} \cup U_{\text{out}}) = \emptyset$. Each connection $(v, u) \in C$ has a weight w_{uv}, and each neuron u has real-valued quantities: network input net_u, activation act_u, and output out_u. Input neurons $u \in U_{\text{in}}$ also have a fourth quantity, the external input ext_u. The predecessors and successors of a vertex u in a directed graph $G = (U, C)$ are defined as $\text{pred}(u) = \{v \in V \mid (v, u) \in C\}$ and $\text{succ}(u) = \{v \in V \mid (u, v) \in C\}$ respectively. Each neuron u is associated with the following functions:

$$f_{\text{net}}^{(u)} : \mathbb{R}^{2|\text{pred}(u)|+\kappa_1(u)} \to \mathbb{R}, \quad f_{\text{act}}^{(u)} : \mathbb{R}^{1+\kappa_2(u)} \to \mathbb{R}, \quad f_{\text{out}}^{(u)} : \mathbb{R} \to \mathbb{R}.$$

These functions compute net_u, act_u, and out_u, where $\kappa_1(u)$ and $\kappa_2(u)$ count the number of parameters of those functions, which can depend on the neurons. Specifically, the new activation act'_u of a neuron u is computed as follows:

$$\text{act}'_u = f_{\text{act}}^{(u)}(f_{\text{net}}^{(u)}(w_{uv_1}, \ldots, w_{uv_{\text{pred}(u)}}, f_{\text{out}}^{(v_1)}(\text{act}_{v_1}), \ldots, f_{\text{out}}^{(v_{\text{pred}(u)})}(\text{act}_{v_{\text{pred}(u)}}), \sigma^{(u)}), \theta^{(u)})$$

where $\sigma^{(u)} = (\sigma_1^{(u)}, \ldots, \sigma_{\kappa_1(u)}^{(u)})$ and $\theta = (\theta_1^{(u)}, \ldots, \theta_{\kappa_2(u)}^{(u)})$ are the input parameter vectors.

Bibliography