Complexity of Neural Networks

Vasquez Alessandro

Université Nice-Sophia-Antipolis

26 février 2018
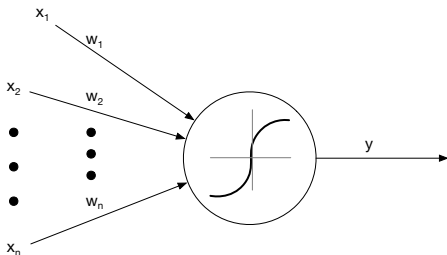
Neuron's representation :
$$y = \sigma(\sum_{j=1}^{n} w_j x_j - w_0)$$
Transfer functions :

- $sign(t) = \begin{cases} 0 & \text{if } t \leq 0 \\ 1 & \text{if } t > 0 \end{cases}$

- $\sigma(t) = (1 + e^{-t})^{-1}$

Vasquez Alessandro   Complexity of Neural Networks

Preliminaries

Neural nets can be

- cyclic or acyclic
- symmetric or asymmetric w.r.t. interconnections
- continuous or discrete w.r.t. updates
- synchronous or asynchronous w.r.t. updates in discrete neural nets

We're going to consider

- back-propagation nets
- Hopfield nets

Preliminaries

Type of neurons

- perceptron (with transfer function)
- threshold gate (with binary valued-input)
- majority gate (all $w = 1$, $b = \frac{|X|}{2}$)

Parameters for evaluating neural nets' complexity

- size
- depth
- weight

Definition (Threshold function)

A boolean function $t : \{0, 1\} \rightarrow 0, 1$ is a *threshold function* if it is computable by a linear threshold unit.

Theorem

Any threshold function on n variables can be computed by a threshold gate with integer weights w_i st :

$$|w_i| \leq \frac{(n+1)^{(n+1)/2}}{2^n}, \forall i = 0, ..., n.$$

This theorem defines a class of threshold gates of polynomially bounded weights.

Theorem

For infinitely many n, there are threshold functions on n variables whose computation by a single threshold gate requires weights as large as $\frac{n^{n/2}}{2^n}$.

These threshold gates have not polynomially bounded gates which, on the other hand, are representable by $O(s \log s)$ bits.

Vasquez Alessandro   Complexity of Neural Networks

Circuit Classes
Definitions

Definition (Complexity classes)

We define five class of functions computable by specifically structured circuits.

- NC^k : polynomial-size circuits of depth $O(\log^k(n))$, using bounded fan-in AND, OR, and NOT gates.
- AC^k : polynomial-size circuits of depth $O(\log^k(n))$, using unbounded AND, OR, and NOT gates.
- TC^k : threshold circuits of polynomial size and depth $O(\log^k n)$,
- TC_d^0 threshold circuits of polynomial size and depth d,
- $\widehat{TC}_d^0$ majority circuits of polynomial size and depth d.

Vasquez Alessandro   Complexity of Neural Networks

Theorem (Complexity classes' hierarchy)

$$AC^k \subseteq TC^k \subseteq NC^{k+1}.$$

$$\widehat{TC}_d^0 \subseteq TC_d^0 \subseteq \widehat{TC}_{d+1}^0.$$

We only know that $AC^0 \subset TC^0$, e.g. because of the majority function, and $TC_1^0 \subset \widehat{TC}_2^0$ for the parity function.

Definition (Threshold circuit loading)

Given $\{(\vec{x_1}, b_1), ..., (\vec{x_m}, b_m)\}$, where each $\vec{x_i} \in \{0,1\}^n$ and each $b_i \in \{0,1\}$, and a directed acyclic graph, can we find a weight assignment st given f the function computed by the resulting threshold circuit : $f(\vec{x_i}) = b_i, \forall i = 1, ..., m$?

Definition (Threshold circuit minimization)

Given $\{(\vec{x_1}, b_1), ..., (\vec{x_m}, b_m)\}$, where each $\vec{x_i} \in \{0, 1\}^n$ and each $b_i \in \{0, 1\}$, and $K \in \mathbb{N}$, does exists a threshold circuit of at most K neurones st given f the function computed by the resulting threshold circuit : $f(\vec{x_i}) = b_i, \forall i = 1, ..., m$?

Vasquez Alessandro   Complexity of Neural Networks

Definition (Hopfield network)

Neural net in which

- all arcs are bidirectional
- each neuron is interconnected to all other neurons
- each neuron has a state
- each neuron's state is updated asynchronously
- each state has an "energy function" assigned :

$$E = -\frac{1}{2} \sum_{i,j} w_{ij} s_i s_j - \sum_i b y_i s_i$$

Theorem

In a symmetric Hopfield net with w_{ij} integer weights and n neurons, the convergence is obtain in

$$3 \sum_{j<i} |w_{ij}| = O(n^2 \cdot \max_{i,j} |w_{ij}|)$$

asynchronous changes of neuron states.

Vasquez Alessandro   Complexity of Neural Networks

Definition (Attraction domain)

Let $\vec{x_i}$ be a stable global state for a given Hopfield net. The *attraction domain* of $\vec{x_i}$ is the set of all vectors which are guaranteed to converge to $\vec{x_i}$.

Definition (Attraction radius)

The *attraction radius* of x_i is the largest Hamming distance from which all other vectors are guaranteed to converge to $\vec{x_i}$.

- Decide if a given symmetric, simple network has more than one stable state.
- Determine if a symmetric simple network converges to a given stable state from any other initial state.

Vasquez Alessandro   Complexity of Neural Networks

- Decide if a given asymmetric net under synchronous updates will converge from any (or all) initial state.
- Computing the attraction radius of a given stable state in a symmetric simple net.

Definition (Hopfield Net Loading Problem)

Given a set of vectors $X = \{\vec{x_1}, .., \vec{x_m}\}$, where each $\vec{x_i} \in \{-1, 1\}^n$, and a constant ρ, is there a symmetric net of n neurons that has X as stable states with attraction radii $\geq \rho$?

The following questions are yet to be answered.

- Is the Hopfield Net Loading Problem in P or in NP ?
- Are the separations $TC^0_d \subseteq TC^0_{d+1}, \forall d \geq 2$ proper ?

Thank you for your attention.