COMPLEXITY OF LEARNING BITHRESHOLD NEURAL UNITS

Kotsovsky V. M.[1], Geche F. E.[2], Batyuk A. Ye.[3], Yurchenko M. V.[1], Mykoryak I. I.[1]
[1] IMST Department, Uzhhorod National University, Uzhhorod, 88000, Ukraine, kotsavlad@gmail.com
[2] CAM Department, Uzhhorod National University, Uzhhorod, 88000, Ukraine, fgeche@hotmail.com
[3] ACS Department, Lviv Polytechnic National University, Lviv, 79013, Ukraine, abatyuk@gmail.com

Our work is motivated by the lack of knowledge about bithreshold units [1]. We study both the conditions ensuring separability by means of bithreshold devices [2] and the complexity [3] of learning such units and networks consisting of them.

Let $\mathbf{w} = (w_1, \ldots, w_n)$ be a *n*-dimensional real vector. For an arbitrary vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathrm{R}^n$ the value of the dot product $\mathbf{w} \cdot \mathbf{x}$ is said to be a weighted sum corresponding to the vector **x**.

A computation unit with *n* inputs $x_1, \ldots, x_n$ and single output *y* is said to be a linear bithreshold unit (LBU) with the weight vector $\mathbf{w} = (w_1, \ldots, w_n)$ and thresholds $t_1, t_2 \in \mathrm{R}$ $(t_1 < t_2)$ if $y = f_{t_1,t_2}(\mathbf{w} \cdot \mathbf{x})$, where the bithreshold activation function f_{t_1,t_2} is defined in the following way:

$$f_{t_1,t_2}(z) = \begin{cases} -1, & \text{if } t_1 < z < t_2, \\ 1, & \text{otherwise} \end{cases}$$

The LBU is completely defined by the ordered triplet $(\mathbf{w}, t_1, t_2)$. We call this triplet the structure of LBU. The bithreshold unit with the structure $(\mathbf{w}, t_1, t_2)$ divides the *n*-dimensional real space into two subsets

$$\mathrm{R}^n_- = \{\mathbf{x} \in \mathrm{R}^n \mid t_1 < \mathbf{w} \cdot \mathbf{x} < t_2\}, \quad \mathrm{R}^n_+ = \mathrm{R}^n \setminus \mathrm{R}^n_- .$$

We call two sets A^+, A^- *bithreshold separable* (B-separable) in R^n if there exists LBU such that $A^+ \subseteq \mathrm{R}^n_+$ and $A^- \subseteq \mathrm{R}^n_-$ [1]. We denote by $\mathrm{Aff}(X)$ the affine hull [2] of the set *X* (the set of all possible affine combination of some elements of *X*).

Proposition 1. *Let A^+ be a finite or countable compact in the n-dimensional real space and $A^- = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ is the set of linearly independent vectors, $(k \le n)$. If there exists such index l $(1 \le l \le k)$ that for every $\mathbf{y} \in A^+ \cap \mathrm{Aff}(A^-)$ its coefficient α_l in the affine expansion $\mathbf{y} = \alpha_1 \mathbf{x}^1 + \ldots + \alpha_l \mathbf{x}^l + \ldots + \alpha_k \mathbf{x}^k$ ($\alpha_i \in \mathrm{R}$, $\alpha_1 + \ldots + \alpha_l + \ldots \alpha_k = 1$, $i = 1, \ldots, k$) satisfies the following condition: $\alpha_l \in (-\infty, 0) \cup (1, +\infty)$, then the sets A^+ and A^- are B-separable.*

Proposition 2. *The problem of checking the B-separability of finite sets A^+ and A^- is NP-complete even in the case $A^+ \cup A^- \subseteq \{a, b\}^n$, where $a \in \mathrm{R}$, $b \in \mathrm{R}$, $a \ne b$, $n \in \mathrm{N}$, and the absolute values of the weight coefficients of LBU can possess only two different values.*

Consequence. *Learning a neural network consisting of bithreshold neuron is NP-complete.*

References

1. Deolalikar, V.: A two-layer paradigm capable of forming arbitrary decision regions in input space. IEEE Transactions on Neural Networks 13(1), 15–21 (2002).
2. Geche, F., Kotsovsky, V.: Representation of finite domain predicates using multithreshold neural elements (in Ukrainian). Herald of the Uzhhorod University, Series of Mathematics and Informatics 6, 32-37 (2001).
3. Blum, A., Rivest, R.: Training a 3-node neural network is NP-Complete. Neural Networks 5(1), 117–127 (1992).