REPRESENTATIONAL CAPABILITIES OF BITHRESHOLD NEURONS

Kotsovsky V. M.[1], Batyuk A. Ye.[2], Yurchenko M. V.[1], Mykoriak I. I.

[1] *IMST Department, Uzhhorod National University, Uzhhorod, 88000, Ukraine, kotsavlad@gmail.com*
[2] *ACS Department, Lviv Polytechnic National University, Lviv, 79013, Ukraine, abatyuk@gmail.com*

Multithreshold neural units were introduced in the 1960s [1]. The main inspiration for dealing with multithreshold devices is their better expressive power compared to ordinary single threshold units [2]. We treat questions related to the ability of bithreshold neurons and neural networks to solve the classification tasks. We study how many partitions of a finite set in *n*-dimensional vector space can be computed by using bithreshold neurons. Our upper bound considerably improve the results of papers [3, 4].

The *bithreshold neuron* (BN) is completely defined by the ordered triplet $(\mathbf{w}, t_1, t_2)$, where *n* is the number of inputs, $\mathbf{w} = (w_1, \ldots, w_n) \in \mathrm{R}^n$ is a weight vector, $t_1, t_2 \in \mathrm{R}$ $(t_1 < t_2)$ are thresholds. We call this triplet the *structure* of BN. The output of BN is obtained by the application of the bithreshold activation function f_{t_1,t_2} to the weighted sum $\mathbf{w} \cdot \mathbf{x} = w_1 x_1 + \ldots + w_n x_n$ [2], where the bithreshold activation function is defined in the following way:

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

A bithreshold neuron can be conveniently represented graphically as shown in Figure 1.

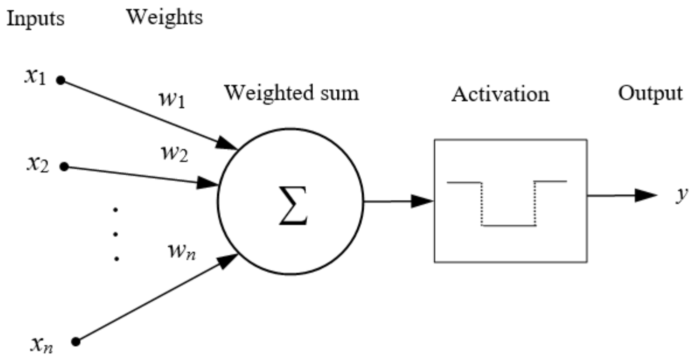

Fig. 1. A bithreshold neuron

The BN 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 BN such that $A^+ \subseteq \mathrm{R}^n_+$ and $A^- \subseteq \mathrm{R}^n_-$ [2]. Let $D_2(A)$ be the number of dichotomies (A^+, A^-) of the finite set *A* in the space R^n, which can be produced by some BN, and $D_2(m,n) = \max\{D_2(A) \mid A \subset \mathrm{R}^n, |A| = m\}$.

Proposition 1. *If* $n \le 2m - 2$, *then the following inequality is true*

$$D_2(m,n) \le \sum_{i=0}^{n+1} \binom{2m-1}{i},$$

where $\binom{2m-1}{i}$ *are binomial coefficients,* $i = 0,\dots,n+1$.

If all points of the set A are in general position (i.e., every subset of n or fewer vectors are linearly independent) and $m > n+1$*, then*

$$D_2(A) \ge \sum_{i=0}^{n+1} \binom{m-1}{i}.$$

Corollary.

$$2\binom{m-1}{n+1} \le D_2(m,n) \le \frac{3(2m-1)^{n+1}}{2(n+1)!} = O\left(m^{n+1}\right).$$

Therefore,

$$D_2(m,n) = \Theta\left(m^{n+1}\right). \quad (1)$$

It should be mentioned that the number of linearly separable dichotomies is $\Theta\left(m^n\right)$ [5]. Thus, (1) expresses the measure of the advantage of bithreshold neurons over threshold ones.

Let BN_n denotes the set of all bithreshold neurons with *n* inputs. We say that the set *A* is shattered by BN_n if all its dichotomies are b-separable. The maximum cardinality of a set *A* shattered by BN_n is the Vapnik-Chervonenkis dimension (VCdim) of BN_n [5]. This is one of the most important measure in the learning theory. We shall be interested in the estimating VCdim BN_n.

Proposition 2. *If* BN_n *is the set all BNs with n inputs and n > 5, then*

$$2n \le \text{VCDim } BN_n < 5n.$$

For comparison, $\text{VCDim } T_n = n+1$, where T_n is the set of all threshold neurons with *n* inputs [5].

We estimated the upper and lower bounds on the number of bithreshold separable dichotomies. Our lower bound is the same as in [3], whereas the upper bound is considerably better than the one stated in [4]. It is very important that for the first time this bounds have the same growth order. It should be noted that our approach can be generalized on the case of common multithreshold units.

REFERENCES

1. Haring, D.: Multiple threshold perceptron. Pattern Recognition, EC-15, 45-65 (1965)
2. Kotsovsky, V., Geche, F., Batyuk, A.: On the computational complexity of learning bithreshold neural units and networks. In: Lytvynenko, V. et al. (eds.) Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2019. Advances in Intelligent Systems and Computing, vol. 1020, pp. 189-202. Springer, Cham (2020)
3. Takiyama, R.: The separating capacity of multithreshold threshold element. IEEE Transactions on Pattern Analysis and Machine Intelligence vol. PAMI-7 (1), 112–116 (1985)
4. Olafsson, S., Abu-Mostafa, Y. S: The capacity of multilevel threshold function. IEEE Transactions on Pattern Analysis and Machine Intelligence 10(2), 277–281 (1988)
5. Anthony, M.: Discrete Mathematics of Neural Networks: Selected Topics. SIAM, Philadelphia (2001)