УДК 681.14

Гече Ф. Е., Коцовський В. М.
Ужгородський національний університет

Двопорогово сепарабельні множини у *n*-вимірному дійсному просторі.

Нейронні елементи та нейромережі, побудовані з них знаходять застосування при розв'язуванні широкого кола практичних задач [1-2]. У зв'язку з обмеженими можливостями класичних порогових елементів інтенсивно вивчалися їх різноманітні узагальнення [3]. Двопорогові нейронні елементи (ДНЕ) з дійсними вагами та порогами (bithreshold neuron) були введені у розгляд багатьма авторами (наприклад у роботах [4-7]). У цих роботах були встановлені оцінки кількості двопорогово сепарабельних дихотомій скінченних множин у просторі $\mathbb{R}^n$ [4-5] та умови, яким повинні задовольняти множини для того, щоб їх можна було відокремити за допомогою пари паралельних гіперплощин [6]. У даній роботі наводяться нові достатні умови двопорогової сепарабельності підмножин простору $\mathbb{R}^n$ та покращується верхня оцінка кількості двопорогово сепарабельних дихотомій скінченних множин, встановлена у роботі [5] для нейронних елементів більш загального вигляду.

Двопороговим дійсним нейронним елементом з ваговим вектором $\mathbf{w} \in \mathbb{R}^n$, порогами $t_1, t_2 \in \mathbb{R}$ $\left(t_1 < t_2\right)$ будемо називати функціональний елемент з *n* дійсними входами $x_1, x_2, \ldots, x_n$ та одним виходом $y \in \{-1, 1\}$, поведінка якого описується співвідношеннями:

$$y = \begin{cases} -1, & \text{якщо } t_1 < (\mathbf{w}, \mathbf{x}) < t_2, \\ 1, & \text{якщо } (\mathbf{w}, \mathbf{x}) \le t_1 \text{ або } (\mathbf{w}, \mathbf{x}) \ge t_2. \end{cases}$$

ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює розбиття простору $\mathbb{R}^n$ на дві підмножини наступним чином:

$$\mathbb{R}^{n+} = \left\{x \in \mathbb{R}^n \mid (\mathbf{w}, \mathbf{x}) \le t_1\right\} \cup \left\{x \in \mathbb{R}^n \mid (\mathbf{w}, \mathbf{x}) \ge t_2\right\}, \quad \mathbb{R}^{n-} = \left\{x \in \mathbb{R}^n \mid t_1 < (\mathbf{w}, \mathbf{x}) < t_2\right\}.$$

Дві множини $A^+ \subset \mathbb{R}^n$ і $A^- \subset \mathbb{R}^n$ назвемо двопорогово-сепарабельними (д-розділимими), якщо знайдеться такий ДНЕ із вектором структури $(\mathbf{w}, t_1, t_2)$, що $A^+ \subset \mathbb{R}^{n+}$ і $A^- \subset \mathbb{R}^{n-}$.

Множину всіх точок, кожна з яких є афінною комбінацією деяких точок множини *X* позначимо через Aff(*X*).

Теорема 1. *Нехай A^+ – не більш, ніж зліченний компакт у просторі $\mathbb{R}^n$ і $A^- = \left\{\mathbf{x}^1, \ldots, \mathbf{x}^k\right\}$ – множина лінійно незалежних векторів, $k \le n$. Тоді якщо знайдеться такий вектор $\mathbf{x}^l \in A^-$, що для довільного $\mathbf{y} \in A^+ \cap \operatorname{Aff}\left(A^-\right)$ коефіцієнт α_l у афінному розкладі* **y** *за елементами множини A^- задовольняє умову*

$$\alpha_l \in (-\infty, 0) \cup (1, +\infty) \tag{1}$$

то множини A^+ і A^- є д-розділимими.

Наслідок. В умовах теореми 1 д-розділимим є множини A^+ і $\operatorname{conv} A^-$.

Теорема 2. *Нехай A^+ – компакт у просторі $\mathbb{R}^n$ і $A^- = \left\{\mathbf{x}^1, \ldots, \mathbf{x}^k\right\}$ – множина лінійно незалежних векторів, $k \le n$. Тоді якщо знайдеться такий вектор $\mathbf{x}^l \in A^-$ і таке поповнення $\left\{\mathbf{x}^1, \ldots, \mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\right\}$ множини A^- до базису простору $\mathbb{R}^n$, що для довільного $\mathbf{y} \in A^+$ його координата α_l у цьому базисі задовольняє умову* (1), *то множини A^+ і $\operatorname{conv} A^-$ є д-розділимими.*

Нехай $D_2(A)$ — кількість різних д-розбитів (A^+, A^-) скінченної m-елементної множини $A \subset \mathbb{R}^n$, $D_2(m,n) = \max\left\{D_2(A) \middle| A \subset \mathbb{R}^n, \operatorname{Card} A = m\right\}$. Надалі без додаткових застережень будемо вважати, що $A \subset \mathbb{R}^n$ і $\operatorname{Card} A = m$.

Теорема 3. *При $m > n$ має місце нерівність*

$$D_2(m,n) < 2\sum_{i=0}^{n+1} C_{2m-1}^{i} . \tag{2}$$

Якщо елементи множини A знаходяться у загальному положенні, то

$$D_2(A) \geq 2\sum_{i=0}^{n+1} C_{m-1}^{i} - 1 . \tag{3}$$

Верхня оцінка (2) значно покращує верхню оцінку, отриману в роботі [5]. При $m >> n$ з (2) і (3) можна зробити висновок, що $D_2(A) = O\left(m^{n+1}\right)$.

Теорема 4. *Нехай $\mathbb{F}$ – клас усіх ДНЕ з n дійсними входами. Тоді*

$$2n \leq \operatorname{VCDim}\mathbb{F} \leq 5n,$$

де $\operatorname{VCDim}\mathbb{F}$ *– розмірність Вапніка-Червоненкіса* [1] *класу* $\mathbb{F}$.

Нехай $\mathbb{FN}$ – клас нейромереж прямого поширення, побудованих із ДНЕ. Тоді

$$\operatorname{VCDim}\mathbb{FN} = O\left(m \ln m\right),$$

де m – кількість вільних параметрів мережі (кількість вагових коефіцієнтів та порогів усіх нейронів мережі).

1. *Haykin S.* Neural Networks. A Comprehensive Foundation. – N.Y.: Prentice Hall, Inc. 1999. – 806 с.
2. *Руденко О. Г., Бодянський Є. В.* Штучні нейронні мережі. – Харків: ТОВ "Компанія СМІТ', 2006. – 404 с.
3. *M. H. Hassoun.* Fundamentals of Artificial Neural Networks. – MIT Press, 1995. – 410 с.
4. *R. Takiyama.* The separating capacity of multi-threshold element, IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-7, pp. 112-116, Jan. 1985.
5. *S. Olafsson and Y. S. Abu-Mostafa.* The capacity of multilevel threshold function, IEEE Trans. Pattern Anal. Machine Intell., vol. 10, NO. 2, pp. 277-281, March 1988.
6. *V. Deolalikar.* A Two-Layer Paradigme Capable of Forming Arbitrary Decision Regions in Input Space, IEEE Trans. on Neural Networks, vol. TNN-4, No. 2 pp. 343-347, March 2001.
7. *Ф. Гече, А. Батюк, В. Коцовський.* Властивості бульових функцій реалізовних на двопорогових елементах // Вісник Національного університету "Львівська політехніка". Комп'ютерна інженерія та інформаційні технології. -- Львів, 2001. – №438. – С. 22-25.