Побудова класифікаторів на основі двопорогових нейронних елементів

Доц., к.т.н. Коцовський В. М., проф., д.т.н. Гече Ф. Е., доц., к.т.н. Левчук О. М., Федорка П. П., Шикула Г. В.
ДВНЗ "Ужгородський національний університет", м. Ужгород, Україна

Штучні нейромережі на базі нейронних елементів (НЕ) із дискретними функціями активації успішно використовуються для розв'язування широкого кола задач розпізнавання образів та класифікації об'єктів [1]. Обмеженість можливостей класичних НЕ з пороговою функцією активації [2] зумовила інтерес до моделей нейроподібних пристроїв більш загального вигляду. Одним із таких пристроїв є двопороговий нейроелемент.

Двопороговим нейронним елементом (ДНЕ) з ваговим вектором $\mathbf{w} = (w_1, \dots, w_n) \in \mathrm{R}^n$, порогами $t_1, t_2 \in \mathrm{R}\ (t_1 < t_2)$ називається функціональний елемент з n дійсними входами $x_1, \dots, x_n$ та одним виходом $y \in \{-1, 1\}$, поведінка якого описується співвідношеннями:

1) якщо $t_1 < (\mathbf{w}, \mathbf{x}) < t_2$, то $y = -1$,
2) в усіх інших випадках $y = 1$,

де $\mathbf{x} = (x_1, \dots, x_n)$ — вектор входів, $(\mathbf{w}, \mathbf{x})$ — скалярний добуток векторів $\mathbf{x}$ та $\mathbf{w}$.

За допомогою ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ можна провести розбиття n-вимірного дійсного простору на дві підмножини, віднісши до однієї з них усі точки, розташовані між паралельними гіперплощинами $(\mathbf{w}, \mathbf{x}) = t_1$ та $(\mathbf{w}, \mathbf{x}) = t_2$, а до іншої — усі інші точки. У зв'язку із цим ДНЕ можна використовувати для побудови класифікаторів.

Дві множини A^+ та A^- називають двопорогово сепарабельними [3] (д-сепарабельними), якщо знайдеться такий ДНЕ із вектором структури $(\mathbf{w}, t_1, t_2)$, що для усіх $\mathbf{x} \in A^-$ $t_1 < (\mathbf{w}, \mathbf{x}) < t_2$, а для усіх $\mathbf{x} \in A^+$ або $(\mathbf{w}, \mathbf{x}) \le t_1$, або $(\mathbf{w}, \mathbf{x}) \ge t_2$. При цьому пару (A^+, A^-) називають д-розбиттям множини $A = A^+ \cup A^-$.

Двопорогові нейронні елементи вивчалися у роботах [3–5].

У зв'язку із дослідженням перспектив застосування ДНЕ для розв'язування практичних задач теорії розпізнавання образів постають три основні задачі:

1) встановлення умов д-сепарабельності;
2) знаходження кількісних оцінок, які характеризують переваги застосування ДНЕ порівняно із звичайними пороговими елементами;
3) розробка алгоритмів навчання ДНЕ та нейромереж на їх основі та дослідження алгоритмічної складності задачі навчання.

Часткова відповідь на перше питання знайдена у [4], на друге питання — в [3, 5]. У доповіді сформульовано нові умови, виконання яких гарантує д-сепарабельність підмножин у n-вимірному дійсному просторі, та наведено верхню оцінку кількості д-сепарабельних дихотомій скінченних множин, яка покращує аналогічну оцінку роботи [5]. Усі наведені нижче результати стосуються множин, елементами яких є вектори простору R^n.

Перейдемо до першої задачі. Наступні два твердження містять достатні умови, виконання яких забезпечує можливість відокремлення множин за допомогою ДНЕ.

Твердження 1. *Нехай* $A^- = \{\mathbf{x}^1, \dots, \mathbf{x}^k\}$ *— множина лінійно незалежних векторів,* $k \le n$, A^+ *— скінченна множина векторів. Тоді якщо знайдеться такий індекс* $l\ (1 \le l \le k)$, *що для довільного* $\mathbf{y} \in A^+$ *, який може бути записаний у вигляді афінної комбінації*

$$\mathbf{y} = \alpha_1 \mathbf{x}^1 + \dots + \alpha_n \mathbf{x}^k \left(\alpha_i \ge 0,\ i = \overline{1,k},\ \sum_{i=1}^{k} \alpha_i = 1 \right),$$

коефіцієнт α_l *задовольняє умову* $\alpha_l \notin [0,1]$, *то множини* A^+ *і* A^- *є д-сепарабельними.*

Твердження 2. *Нехай* $A^- = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ *— множина лінійно незалежних векторів,* $k \le n$, A^+ *— компакт. Тоді якщо існує таке поповнення* $\{\mathbf{x}^1, \ldots, \mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\}$ *множини* A^- *до базису простору* R^n *і такий індекс* $l \ (1 \le l \le k)$, *що для довільного* $\mathbf{y} \in A^+$ *його координата* $\alpha_l \notin [0,1]$, *то множини* A^+ *і* $\operatorname{conv} A^-$ *є д-сепарабельними.*

Перейдемо до другої задачі. Нехай $D_2(A)$ — кількість різних д-розбитів (A^+, A^-) скінченної *m*-елементної множини, $D_2(m,n) = \max\{D_2(A) | A = \{a_1, \ldots, a_m\} \subset \mathrm{R}^n\}$.

Наступне твердження містить оцінки величини $D_2(m,n)$.

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

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

Отримана оцінка є покращенням верхньої оцінки, отриманої в роботі [5]. Якщо число *m* набагато більше за *n*, то попередньої оцінки та верхньої оцінки [5] випливає, що $D_2(A) = O(m^{n+1})$. Для порівняння $D_1(m,n) = O(m^n)$, де $D_1(m,n)$ — максимальна можлива кількість різних лінійних розбиттів *m*-елементної множини у *n*-вимірному дійсному просторі.

Розглянемо третю задачу. Наступний результат показує, що існування ефективного алгоритму навчання ДНЕ є малоймовірним.

Твердження 4. *Задача перевірки д-сепарабельності скінченних множин* A^+ *та* A^- *є NP-повною навіть за умови, що* $A^+ \cup A^- \subset \{a,b\}^n$, *де* $a,b \in \mathrm{R}\,(a \ne b)$.

Обґрунтування деяких з наведених результатів наведено у [6].

Перелік посилань

1. Руденко, О. Г. Штучні нейронні мережі / О. Г. Руденко, Є. В. Бодянський. — Харків: ТОВ «Компанія СМІТ», 2006. — 404 с.
2. Хайкин С. Нейронные сети: полный курс / С. Хайкин. — 2-е изд. — М. : Вильямс-Телеком, 2006. — 1104 с.
3. Takiyama, R. Multiple threshold perceptron / R. Takiyama // Pattern recognition. – 1978, vol. 10. – PP. 27-30.
4. Deolalikar, V. A Two-Layer Paradigme Capable of Forming Arbitrary Decision Regions in Input Space / V. Deolalikar // IEEE Transactions on Neural Networks. – 2001, vol. TNN-4, # 2. – PP. 343-347.
5. Olafsson, S. The capacity of multilevel threshold function / S. Olafsson and Y. S. Abu-Mostafa // IEEE Trans. Pattern Anal. Machine Intell. – 1988, vol. 10, No 2. – PP. 277-281.
6. Коцовський, В. М. Властивості дихотомій, породжених двопороговими нейронними елементами / В. М. Коцовський // Науковий вісник Ужгородського університету. Сер. матем. і інформ. – Ужгород, 2010. – С. 67-78.