УДК 004.85
[1] **В. М. Коцовський**
к. т. н., доцент
[1] **І. І. Микоряк**
аспірант
[1] **М. В. Юрченко**
аспірант
[1]*ДВНЗ «Ужгородський національний університет», Ужгород*

Спектральний алгоритм навчання поліноміальних порогових елементів

Вступ. Нейромережі на базі поліноміальних нейронних елементів (ПНЕ) використовуються для розв'язування задач класифікації об'єктів, розпізнавання образів, стиску інформації, наближення функцій, прогнозування та прийняття рішень [1]. Одним із основних напрямків досліджень у межах конекціоністського підходу до штучного інтелекту є розробка та аналіз алгоритмів навчання нейромереж.

Нехай $E_2 = \{-1,1\}$, $E_2^n = E_2 \times \ldots \times E_2$. Функції вигляду $f : E_2^n \to E_2$ будемо називати булевими функціями (БФ) у біполярному алфавіті E_2. Визначимо мономи $P_j : E_2^n \to E_2$ наступним чином: $P_j(\mathbf{a}) = a_1^{j_1} \cdot \ldots \cdot a_n^{j_n}$, де $\mathbf{a} = (a_1,\ldots,a_n) \in E_2^n$, $j = j_1 \cdot 2^{n-1} + j_2 \cdot 2^{n-2} + \ldots + j_{n-1} \cdot 2^1 + j_n$ $\left(j_i \in \{0,1\},\ i = \overline{1,n} \right)$.

Нехай $\Pi = \{P_1,\ldots,P_m\}$ — деяка система поліномів, побудованих із змінних $x_1,\ldots,x_n$, $\mathbf{w} = (w_1,\ldots,w_m) \in \mathrm{R}^m$. Якщо для БФ $f : E_2^n \to E_2$ виконується умова

$$\text{для всіх } \mathbf{a} \in E_2^n \quad f(\mathbf{a}) = \operatorname{sgn} \sum_{j=1}^{m} w_j P_j(\mathbf{a}), \qquad (1)$$

то кажуть, що відносно системи поліномів Π функція $f(\mathbf{x})$ реалізується на ПНЕ з ваговим вектором **w**. Без втрати загальності міркувань надалі будемо вважати, що усі поліноми системи Π є мономами. Якщо для $f(\mathbf{x})$ знайдеться ваговий вектор **w**, для якого виконується умова (1), то функцію $f(\mathbf{x})$ назвемо Π-пороговою БФ. При цьому для вагового вектора **w** має виконуватися умова: для всіх $\mathbf{a} \in E_2^n$ скалярний добуток $(\mathbf{w}, P(\mathbf{a}))$ відмінний від нуля, де $(\mathbf{w}, P(\mathbf{a})) = w_1 P_1(\mathbf{a}) + \ldots + w_m P_m(\mathbf{a})$. Такі вагові вектори будемо називати Π-допустимими.

Під навчанням ПНЕ реалізовувати Π-порогову булеву функцію $f(\mathbf{x})$ будемо розуміти процес побудови скінченної послідовності Π-допустимих вагових векторів

$$\mathbf{w}^0, \mathbf{w}^1, \ldots, \mathbf{w}^t, \qquad (2)$$

такої, що функція $f(\mathbf{x})$ реалізується на ПНЕ з ваговим вектором $\mathbf{w}^t$.

Визначимо характеристичний (спектральний) вектор $\mathbf{s}_\Pi(f) = (s_1, ..., s_m)$ функції $f(\mathbf{x})$ відносно системи П наступним чином [2]:

$$s_j = \sum_{\mathbf{a} \in E_2^n} P_j(\mathbf{a}) f(\mathbf{a}), \quad j = \overline{1, m}.$$

У процесі навчання ми будемо використовувати правило вигляду

$$\mathbf{w}^{k+1} = \mathbf{w}^k + t_k \frac{\left(\mathbf{w}^k, \mathbf{s}_\Pi\left(f^k\right) - \mathbf{s}_\Pi(f)\right)}{\left\|\mathbf{s}_\Pi(f) - \mathbf{s}_\Pi\left(f^k\right)\right\|^2} \left(\mathbf{s}_\Pi(f) - \mathbf{s}_\Pi\left(f^k\right)\right), \tag{3}$$

де t_k — нормуючий множник приросту.

Твердження 1. Якщо БФ *f* — П-порогова, послідовність (2) П-допустимих вагових векторів $\left\{\mathbf{w}^k\right\}$ будується з використанням корекцій вигляду (3) і нормуючий множник приросту має вигляд

$$t_k = \sigma_k + \frac{\tau_k}{\left(\mathbf{w}^k, \mathbf{s}_\Pi\left(f^k\right) - \mathbf{s}_\Pi(f)\right)},$$

де $0 \le \sigma_k \le 2$ і $0 < C \le \sigma_k \le D$ для деяких сталих *C* та *D*, $(k = 0, 1, ...)$, то навчання успішно завершується після скінченної кількості кроків.

Схема доведення, результати комп'ютерного моделювання алгоритму навчання та їх аналіз наведені в [2].

Висновки. У роботі розглянуто алгоритм offline навчання ПНЕ, заснований на використанні спектральних коефіцієнтів булевих функцій. Наведено умови, дотримання яких гарантує скінченність та результативність процесу навчання. Тим самим підтверджено припущення роботи [3] про покращення ефективності алгоритму навчання за рахунок використання більш широкого діапазону значень нормуючих множників приростів.

Список використаних джерел

1. Хайкин С. Нейронные сети: полный курс. — М.: Вильямс-Телеком, 2006. — 1104 с.
2. Kotsovsky V., Geche F., Batyuk A. Finite Generalization of the Offline Spectral Learning. Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining & Processing, DSMP 2018. — 21-25 August 2018, Lviv, Ukraine. — pp. 356-360.
3. Hampson S., Kibler D. Minimum Generalization via Reflection: A Fast Linear Threshold Learner. Machine Learning. — 1999. — 37(1). — P. 51–73.