ЗАЩИТА ИНФОРМАЦИИ И НАДЕЖНОСТЬ СИСТЕМ
INFORMATION PROTECTION AND SYSTEM RELIABILITY

УДК 004.056.5
DOI: 10.37661/1816-0301-2025-22-3-72-82

Оригинальная статья
Original Article

Аппроксимация двоичных функций на основе двухслойной искусственной нейронной сети

К. В. Латушкин, Ю. С. Харин✉

НИИ прикладных проблем математики и информатики
Белорусского государственного университета,
пр. Независимости, 4, Минск, 220030, Беларусь
✉*E-mail: Kharin@bsu.by*

Аннотация
Цели. Рассматриваются особенности применения двухслойных искусственных нейронных сетей в задачах аппроксимации двоичных функций многих двоичных переменных. Изучаются вопросы выбора начальных значений весов модели и количества нейронов на скрытом слое.
Методы. Задача аппроксимации двоичной функции с помощью искусственной нейронной сети сводится к геометрической задаче разделения вершин многомерного куба гиперплоскостями. Комбинаторными методами доказываются леммы о способах разбиения гиперкуба гиперплоскостью и строится оценка снизу количества двоичных функций, для аппроксимации которых достаточен один нейрон на скрытом слое.
Результаты. Рассмотрены особенности задания начальных значений весов искусственной нейронной сети. Построена оценка снизу числа двоичных функций, для аппроксимации которых достаточно искусственной нейронной сети с одним нейроном на скрытом слое. Найдена алгоритмическая сложность вычисления такой оценки. Представлены численные результаты применения двухслойных искусственных нейронных сетей для аппроксимации двоичных функций в задачах защиты информации.
Заключение. Результаты статьи позволяют выбирать параметры искусственной нейронной сети для повышения точности аппроксимации двоичных функций многих переменных.

Ключевые слова: двоичная функция, комбинаторика, искусственная нейронная сеть, аппроксимация функций, генераторы псевдослучайных последовательностей

Для цитирования. Латушкин, К. В. Аппроксимация двоичных функций на основе двухслойной искусственной нейронной сети / К. В. Латушкин, Ю. С. Харин // Информатика. – 2025. − Т. 22, № 3. – С. 72–82. – DOI: 10.37661/1816-0301-2025-22-3-72-82.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 16.07.2025
Подписана в печать | Accepted 18.08.2025
Опубликована | Published 30.09.2025

Approximation of binary functions based on two-layer artificial neural network

Konstantin V. Latushkin, Yuriy S. Kharin✉

Research Institute of Applied Problems of Mathematics and Informatics of the Belarusian State University, av. Nezavisimosti, 4, Minsk, 220030, Belarus
✉E-mail: Kharin@bsu.by

Abstract
Objectives. The article examines the features of using two-layer artificial neural network in problems of approximating binary functions of many binary variables. The issues of choosing the initial values of the model weights and choosing the number of neurons on the hidden layer are studied.
Methods. The problem of approximating a binary function using an artificial neural network is reduced to the geometric problem of dividing the vertices of a multidimensional cube by hyperplanes. Combinatorial methods are used to prove lemmas on ways of dividing a hypercube by a hyperplane and to construct a lower estimate for the number of binary functions that can be approximated using one neuron on the hidden layer.
Results. The features of setting the initial values of weights of an artificial neural network are considered. A lower bound is constructed for the number of binary functions that can be approximated using an artificial neural network with one neuron on the hidden layer. The algorithmic complexity of calculating such an estimate is found. Numerical results are presented for using two-layer artificial neural networks to approximate binary functions in information security problems.
Conclusion. The results of the article allow choosing the parameters of an artificial neural network to improve the accuracy of approximation of binary functions of many variables.

Keywords: binary functions, combinatorics, artificial neural network, function approximation, pseudorandom sequence generators

For citation. Latushkin K. V., Kharin Yu. S. *Approximation of binary functions based on two-layer artificial neural network.* Informatika [*Informatics*], 2025, vol. 22, no. 3, pp. 72−82 (In Russ.). DOI: 10.37661/1816-0301-2025-22-3-72-82.

Conflict of interest. The authors declare of no conflict of interest.

Введение. В последние годы искусственные нейронные сети (ИНС) начинают широко применяться в задачах анализа дискретных, в том числе двоичных, данных в криптологии и кибербезопасности [1–4]. Примерами таких задач являются: двоичная классификация по наблюдениям, представляемым двоичными векторами; аппроксимация двоичных функций в программных датчиках псевдослучайных последовательностей; оценка сложности s-блоков в криптографических системах; распознавание наличия компьютерной атаки на информационные системы. Математически эти задачи сводятся к задаче аппроксимации двоичных функций от многих двоичных переменных. Исследованию особенностей этой актуальной задачи посвящена данная статья.

Математическая модель и постановка задачи. Введем обозначения: $V = \{0,1\}$ – двоичный алфавит; s – натуральное число; V^s – s-мерный двоичный гиперкуб; $x = (x_1, x_2, \dots, x_s)' \in V^S$ – двоичный вектор-столбец, или, в геометрической интерпретации, некоторая вершина гиперкуба V^s; $\mathbb{1}\{B\} \in V$ – индикатор события B, $\mathbb{1}\{B\} = \{1,$ если справедливо $B; 0$ в противном случае$\}$.

Пусть на множестве V^s определена некоторая неизвестная двоичная функция s двоичных переменных:

$$y = f(x) = f(x_1, x_2, \dots, x_s), x \in V^s, y \in V. \quad (1)$$

В геометрической интерпретации функция (1) задает классификацию (разбиение) вершин гиперкуба V^s на два непересекающихся класса: $\Omega_0 = \{x \in V^s : f(x) = 0\}$ и $\Omega_1 = \{x \in V^s : f(x) = 1\}$: $V^s = \Omega_0 \cup \Omega_1$, $\Omega_0 \cap \Omega_1 = \emptyset$. Функцию (1) можно также интерпретировать как раскраску каждой вершины $x \in V^s$ гиперкуба в один из двух цветов: $y = 0$ или $y = 1$. Обозначим $\mathcal{F}$ – множество всевозможных двоичных функций от s двоичных переменных, $|\mathcal{F}| = K = 2^{2^s}$.

Рассмотрим задачу аппроксимации (восстановления) функции (1) по случайной выборке объемом n из V^s: $X = \{x^{(1)}, x^{(2)}, \dots, x^{(n)}\} \subseteq V^s$ и известным значениям $y^{(t)} = f(x^{(t)})$, $t = 1, \dots, n$. Для решения этой задачи используется двухслойная ИНС (с s входами, одним скрытым слоем с m нейронами и одним выходом), проиллюстрированная на рис. 1, где $H_l^{(1)}$ – l-й нейрон скрытого слоя, m – число нейронов, $H^{(2)}$ – выходной нейрон ИНС.

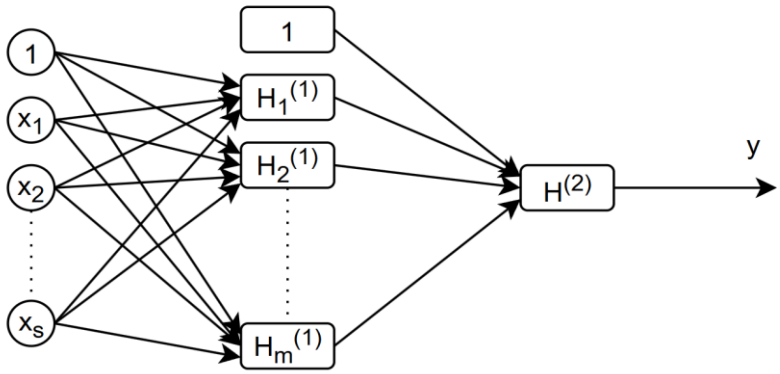

Рис. 1. Схема ИНС
Fig. 1. ANN diagram

ИНС на рис. 1 определяет аппроксимирующую для $f(\cdot)$ функцию $\hat{f}(x) = \hat{f}(x_1, x_2, \dots, x_s) \in V$ с $m(s+2)+1$ оцениваемыми по выборке X параметрами:

$$\hat{f}(x_1, \dots x_s) = \sigma\left(b_0 + \sum_{l=1}^{m} b_l \cdot ReLU\left(a_{0l} + \sum_{i=1}^{s} a_{il} x_i\right)\right), \tag{2}$$

где m – натуральное число (число нейронов на скрытом слое); $\{a_{il}\}$, $\{b_l\}$ – параметры (коэффициенты, веса) модели; $ReLU(z) = \max\{0, z\}$, $\sigma(z) = (1 + e^{-z})^{-1}$ – так называемые функции активации [5].

Отметим, что ИНС с одним скрытым слоем выбирается авторами с целью выявления математических особенностей аппроксимации двоичных функций, которые в дальнейшем могут обобщаться на многослойные ИНС.

Аппроксимация $f(\cdot)$ с помощью $\hat{f}(\cdot)$ получается как результат процесса обучения ИНС (2), заключающегося в минимизации функции потерь по $\{a_{il}\}$, $\{b_l\}$:

$$h(\hat{y}) = -\frac{1}{n}\sum_{t=1}^{n}\left(y^{(t)} \log\left(\hat{y}^{(t)}\right) + \left(1 - y^{(t)}\right)\log\left(1 - \hat{y}^{(t)}\right)\right) \to \min_{\{a_{il}\},\{b_l\}}. \tag{3}$$

Функцию потерь $h(\hat{y})$ в (3) принято называть «бинарной перекрестной энтропией» [5]. В формуле (3) величина $\hat{y}^{(t)} = \hat{f}(x^{(t)})$ – это полученная в процессе обучения оценка для $y^{(t)}$, а для оценивания точности обучения используется accuracy – доля правильно классифицированных вершин:

$$\alpha = accuracy = \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{\hat{y}^{(t)} = y^{(t)}\} \in [0,1].$$

Особенности использования ИНС для аппроксимации двоичных функций. При аппроксимации двоичных функций многих двоичных переменных на основе ИНС возникают две важные особенности, связанные:

– с наличием областей кусочного постоянства и многоэкстремальности целевой функции (3);

– выбором числа нейронов m.

Кусочное постоянство и многоэкстремальность функции потерь. Докажем лемму о свойствах коэффициентов ИНС (2).

Лемма 1. *Если коэффициенты $\{a_{il}\}$ некоторого l-го нейрона $H_l^{(1)}$ ИНС (2) на множестве вершин X удовлетворяют условию*

$$a_{0l} + \sum_{i=1}^{s} a_{il} x_i < 0, \forall x \in X, \tag{4}$$

то производная функции $h(\cdot)$, определяемой (3), по коэффициентам $\{a_{il}, b_l\}$ на множестве X равна 0.

Доказательство. Продифференцируем $h(\hat{y})$ по переменным $\{a_{il}, b_l\}$, используя уравнения (2) и (3):

$$\frac{\partial h(\hat{y})}{\partial a_{kl}} = -\frac{1}{n}\sum_{t=1}^{n}\left(y^{(t)}\left(1-\hat{y}^{(t)}\right) - \left(1-y^{(t)}\right)\hat{y}^{(t)}\right) b_l \mathbb{1}\left\{a_{0l} + \sum_{i=1}^{s} a_{il} x_i^{(t)} > 0\right\} \cdot \begin{cases} x_k^{(t)}, k \neq 0, \\ 1, k = 0, \end{cases}$$

$$\frac{\partial h(\hat{y})}{\partial b_l} = -\frac{1}{n}\sum_{t=1}^{n}\left(y^{(t)}\left(1-\hat{y}^{(t)}\right) - \left(1-y^{(t)}\right)\hat{y}^{(t)}\right) \cdot \begin{cases} ReLU\left(a_{0l} + \sum_{i=1}^{s} a_{il} x_i^{(t)}\right), l \neq 0, \\ 1, l = 0. \end{cases}$$

Учитывая условие (4), $l > 0$ и вид функции $ReLU(\cdot)$, получаем, что производные обращаются в ноль.

Лемма доказана.

Из леммы 1 следует, что на множестве параметров ИНС

$$A_l = \left\{(a_{0l}, a_{1l}, \ldots, a_{sl}) : a_{0l} + \sum_{i=1}^{s} a_{il} x_i < 0, \forall x \in X\right\} \subset R^{s+1}$$

целевая функция кусочно постоянна по параметрам l-го нейрона. На таких множествах кусочного постоянства применение градиентного спуска для решения задачи минимизации (3) приводит к ухудшению сходимости. Еще одна трудность в решении задачи (3) состоит в многоэкстремальности целевой функции в (3).

Для преодоления трудностей, связанных с наличием областей $\{A_l\}$ кусочного постоянства, предлагается специально подбирать начальные значения параметров $\{a_{il}, b_l\}$ следующим образом. Вначале генерируем их как случайные величины согласно работам [6, 7] из распределений вероятностей:

$$a_{il} \sim N\left(0, \frac{2}{s}\right), b_l \sim U\left[-\frac{\sqrt{6}}{\sqrt{m+1}}, \frac{\sqrt{6}}{\sqrt{m+1}}\right],$$

где $N(\mu, \sigma^2)$ – нормальное распределение с математическим ожиданием μ и дисперсией σ^2; $U[a, b]$ – равномерное распределение, заданное на промежутке $[a, b]$. Если $(a_{0l}, a_{1l}, \ldots, a_{sl}) \in A_l$, то для параметров $\{a_{0l}, \ldots, a_{sl}\}$ l-го нейрона ($l \in \{1, \ldots, m\}$) повторяем генерацию начальных значений до тех пор, пока не получим начальные значения вне области кусочного постоянства A_l.

О выборе числа нейронов. Будем использовать указанную выше геометрическую интерпретацию аппроксимации ИНС, связанную с раскраской вершин гиперкуба V^s. Каждый нейрон $H_l^{(1)}(x) = ReLU(a_{l0} + \sum_{i=1}^{s} a_{il}x_i)$ $(l \in \{1, \dots, m\})$, изображенный на рис. 1, порождает некоторую гиперплоскость, разделяющую множество вершин гиперкуба V^s на два подмножества $\Omega_{0l} = \left\{x \in V^s : H_l^{(1)}(x) = 0\right\}$ и $\Omega_{1l} = \left\{x \in V^s : H_l^{(1)}(x) > 0\right\}$. Исходя из такого способа разделения вершин, можно построить ИНС следующим образом: каждую вершину $x^{(l)} \in V^s$ с цветом 1 $\left(f\left(x^{(l)}\right) = 1\right)$ отделить гиперплоскостью от остальных так, чтобы на этой вершине значение l-го нейрона $H_l^{(1)}\left(x^{(l)}\right)$ стало положительным числом, а на остальных равнялось 0; $\{b_l\}$ задать из условий

$$\sigma(b_0) = \varepsilon, \sigma\left(b_0 + b_l H_l^{(1)}\left(x^{(l)}\right)\right) = 1 - \varepsilon \ (l \in \{1, \dots, m\}),$$

где $\varepsilon < 1/2$ – достаточно малое положительное число.

Таким образом, верхняя граница числа нейронов существует, конечна и зависит от размерности гиперкуба:

$$1 \le m \le 2^s.$$

При этом число всевозможных различных раскрасок гиперкуба ограничено и равняется $|\mathcal{F}| = K = 2^{2^s}$. Обозначим $W(s, m)$ число различных раскрасок вершин гиперкуба V^s, для аппроксимации которых необходимо и достаточно m нейронов на скрытом слое ИНС (2). Отсюда следует

$$\sum_{m=1}^{2^s} W(s, m) = 2^{2^s}.$$

Построим оценку снизу для $W(s, 1)$, т. е. для количества раскрасок вершин s-мерного гиперкуба, полученных разделением этого множества вершин одной гиперплоскостью $(m = 1)$.

Оценка снизу для $W(s, 1)$. Для построения оценки снизу докажем несколько лемм.

Лемма 2. *Любой s-мерный гиперкуб V^s содержит в себе $2^{s-l}C_s^l$ граней размерности l.*

Доказательство. Рассмотрим гиперкуб V^s размерности s. Грань размерности l представляет собой подмножество вершин гиперкуба, $s - l$ координат которых фиксированы. Число различных способов выбрать $s - l$ фиксированных координат из s возможных $C_s^{s-l} = C_s^l$. При этом каждой фиксированной координате нужно задать одно из двух значений 0 или 1. Количество различных наборов значений фиксированных координат равно 2^{s-l}.

Лемма доказана.

Лемма 3. *Пусть A, B – два конечных множества, k – некоторое неотрицательное целое число, $\mathrm{C}_k(*)$ – число различных способов выбрать k элементов из множества $*$. Тогда справедливо неравенство $C_k(A \cup B) \ge C_k(A) + C_k(B) - C_k(A \cap B)$.*

Доказательство. Пусть $|A| = m_1$, $|B| = m_2$, $|A \cap B| = m_3$. Тогда число способов выбрать k элементов из множеств A, B, $A \cup B$, $A \cap B$ определяется формулами

$$C_k(A) = \sum_{0 \le k_1 \le k} C_{m_1 - m_3}^{k_1} C_{m_3}^{k-k_1}, C_k(B) = \sum_{0 \le k_2 \le k} C_{m_2 - m_3}^{k_2} C_{m_3}^{k-k_2},$$

$$C_k(A \cup B) = \sum_{0 \le k_1 + k_2 \le k} C_{m_1 - m_3}^{k_1} C_{m_2 - m_3}^{k_2} C_{m_3}^{k-k_1-k_2}, C_k(A \cap B) = C_{m_3}^{k}.$$

Из этих формул следует, что числа $C_k(A)$, $C_k(B)$ и $C_k(A \cap B)$ входят в состав $C_k(A \cup B)$ (если подставить в сумме $k_2 = 0$, $k_1 = 0$, $k_1 = k_2 = 0$), откуда получается требуемое неравенство. Равенство достигается при $k = 0$ и $k = 1$.

Лемма доказана.

Условимся говорить, что k вершин $(1 \le k \le 2^s - 1)$ линейно выделимы в гиперкубе V^s, если существует такая $(s-1)$-мерная гиперплоскость, которая разбивает V^s на два непересекающихся подмножества k и $2^s - k$ вершин.

Лемма 4. *Если на некоторой грани гиперкуба k вершин линейно выделимы, то эти же k вершин можно линейно выделить во всем гиперкубе.*

Доказательство. Для доказательства леммы достаточно показать возможность перехода к $(t+1)$-мерному гиперкубу от любой его t-мерной грани. Без ограничения общности будем считать, что t-мерная грань получена из гиперкуба путем фиксирования переменной x_{t+1}. В зависимости от значения x_{t+1} таких граней можно получить две: $V_0 = \{(x_1, \dots, x_t, 0) : x_i \in V\}$, $V_1 = \{(x_1, \dots, x_t, 1) : x_i \in V\}$, $V_0 \cup V_1 = V^{t+1}$, $V_0 \cap V_1 = \emptyset$. Обозначим T множество вершин гиперкуба, определяющих t-мерную грань ($T = V_0$ либо $T = V_1$). По условию леммы k вершин грани T линейно выделимы, т. е. существует гиперплоскость, задающаяся уравнением

$$L_t(x_1, \dots, x_t) = a_0 + \sum_{i=1}^{t} a_i x_i = 0, \tag{5}$$

которая разделяет грань T на два непересекающихся подмножества $T_+ = \{x \in T : L_t(x_1, \dots, x_t) > 0\}$ и $T_- = \{x \in T : L_t(x_1, \dots, x_t) < 0\}$: $T = T_+ \cup T_-$, $T_+ \cap T_- = \emptyset$, $|T_+| = k$, $|T_-| = 2^t - k$. Достроим гиперплоскость (5) так, чтобы на оставшихся вершинах $V^{t+1} \setminus T$ гиперкуба V^{t+1} левая часть нового уравнения давала отрицательное значение. Уравнение t-мерной гиперплоскости при этом будет иметь вид

$$L_{t+1}(x_1, \dots, x_t, x_{t+1}) = b_0 + b_1 x_{t+1} + L_t(x_1, \dots, x_t) = 0. \tag{6}$$

Коэффициенты b_0, b_1 уравнения гиперплоскости (6) зададим из условий

$$\begin{cases} L_{t+1}(x_1, \dots, x_t, x_{t+1})|_{x \in T} = L_t(x_1, \dots, x_t), \\ L_{t+1}(x_1, \dots, x_t, x_{t+1})|_{x \in V^{t+1} \setminus T} < 0. \end{cases}$$

В зависимости от расположения грани T в гиперкубе V^{t+1} коэффициенты b_0, b_1 уравнения гиперплоскости (6) будут принимать следующие значения:

– если $T = V_0$, то $b_0 = 0$, $b_1 = -A$, $A > \max\limits_{x \in T} L_t(x_1, \dots, x_t)$;

– если $T = V_1$, то $b_0 = -A$, $b_1 = A$, $A > \max\limits_{x \in T} L_t(x_1, \dots, x_t)$.

Построенная гиперплоскость разделяет гиперкуб V^{t+1} на два непересекающихся подмножества T_+ и $V^{t+1} \setminus T_+$.

Лемма доказана.

Исходя из лемм 2–4, построим дискретную функция $Q(s,k)$ $(s, k \in \{0, 1, 2, \dots\})$, заданную следующим рекуррентным соотношением:

$$\begin{cases} Q(s,k) = \sum\limits_{i=1}^{s} (-1)^{i+1} 2^i C_s^{s-i} Q(s-i,k), \text{ если } k \le 2^{s-1}, \\ Q(s,k) = 0, \text{ если } k > 2^s, \\ Q(s,k) = Q(s, 2^s - k), \text{ если } k > 2^{s-1}, \\ Q(s,0) = Q(s, 2^s) = 1. \end{cases} \tag{7}$$

Теорема. *Значение функции $Q(s,k)$, полученное из рекуррентного соотношения (7), является оценкой снизу числа способов линейно выделить k вершин в s-мерном гиперкубе.*

Доказательство. Доказательство следует из лемм 2–4. О том, что значение функции $Q(s,k)$ является оценкой снизу, говорит лемма 3. Второе и четвертое условия рекуррентного соотношения (7) задают граничные условия. Третье условие следует из того, что гиперплоскость делит множество вершин гиперкуба V^s на два подмножества с k и 2^s-k вершинами.

Теорема доказана.

Следствие. *Чтобы найти оценку снизу для $W(s,1)$, необходимо просуммировать по k значения функции $Q(s,k)$:*

$$W(s,1) \geq \sum_{k=0}^{2^s} Q(s,k).$$

В табл. 1 для примера представлены все значения функции $\{Q(s,k)\}$ для $s \leq 3$, $k \leq 8$. Жирным шрифтом выделены значения, задаваемые условиями 2 и 4 соотношения (7). Таблица значений функции $Q(s,k)$ заполняется построчно слева направо, снизу вверх.

Таблица 1
Пример заполнения таблицы значений функции $Q(s,k)$
Table 1
Example of filling in the table of the function values $Q(s,k)$

3	**1**	8	12	24	6	24	12	8	**1**
2	**1**	4	4	4	**1**	**0**	**0**	**0**	**0**
1	**1**	2	**1**	**0**	**0**	**0**	**0**	**0**	**0**
0	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
s / k	0	1	2	3	4	5	6	7	8

Оценим затраты времени на заполнение таблицы. Условия 2 и 4 дают сложность $O(s2^s)$. Время, требуемое на заполнение строки под номером s, обозначим через $T(s)$. Эту строку можно разбить по k на промежутки $\{[1,1], [2,3], [4,7], [8,15], \dots, [2^{i-1}, 2^i-1], \dots, [2^{s-1}, 2^s-1]\}$, где каждое значение $Q(s,k)$ в i-м промежутке $(i = 0 \dots, s-1)$ требует $s-i$ значений из таблицы. Следовательно,

$$T(s) = \sum_{i=0}^{s-1} (s-i)2^i = 2^{s+1} - s - 2,$$

$$\sum_{t=1}^{s} T(t) = \sum_{t=1}^{s} (2^{t+1} - t - 2) = 2^{s+2} - \frac{s(s+1)}{2} - 2s - 4 = O(2^s).$$

Таким образом, общая сложность заполнения таблицы значений функции $Q(s,k)$ имеет порядок $O(s2^s)$. В табл. 2 представлено сравнение оценки $Q(s,k)$ с точным значением количества различных раскрасок для соответствующего значения k при $s \leq 3$.

Символом «-» в табл. 2 обозначены случаи $k > 2^s$. В последней строке представлена вероятность события, состоящего в том, что случайно выбранная раскраска вершин гиперкуба требует для аппроксимации один нейрон на скрытом слое (оценка и точное значение).

Как видно из табл. 2, расхождение оценки и точного значения произошло при $k = 4$, $s = 3$. Это объясняется тем, что оценка $Q(s,k)$ учитывает только варианты, когда раскраска вершин гиперкуба сводится к раскраске вершин его грани (т. е. гиперкуба меньшей размерности). При этом не учитываются случаи, когда раскрашенные вершины принадлежат сразу нескольким таким граням. Например, вершины трехмерного куба $(s = 3)$ можно разделить плоскостью $x_1 + x_2 + x_3 - \frac{3}{2} = 0$ (рис. 2).

Таблица 2
Сравнение оценки $Q(s,k)$ и точного значения для $s \le 3$
Table 2
Comparison of the estimate $Q(s,k)$ and exact value for $s \le 3$

Число вершин k цвета 1 *Number of vertices k of color 1*	Отрезок $s=1$ *Section $s=1$*		Квадрат $s=2$ *Square $s=2$*		Куб $s=3$ *Cube $s=3$*	
	Оценка $Q(s,k)$ *Grade $Q(s,k)$*	Точное значение *Exact meaning*	Оценка $Q(s,k)$ *Grade $Q(s,k)$*	Точное значение *Exact meaning*	Оценка $Q(s,k)$ *Grade $Q(s,k)$*	Точное значение *Exact meaning*
0	1	1	1	1	1	1
1	2	2	4	4	8	8
2	1	1	4	4	12	12
3	-	-	4	4	24	24
4	-	-	1	1	6	14
5	-	-	-	-	24	24
6	-	-	-	-	12	12
7	-	-	-	-	8	8
8	-	-	-	-	1	1
Сумма	4	4	14	14	96	104
Вероятность	1	1	0,875	0,875	0,375	0,40625

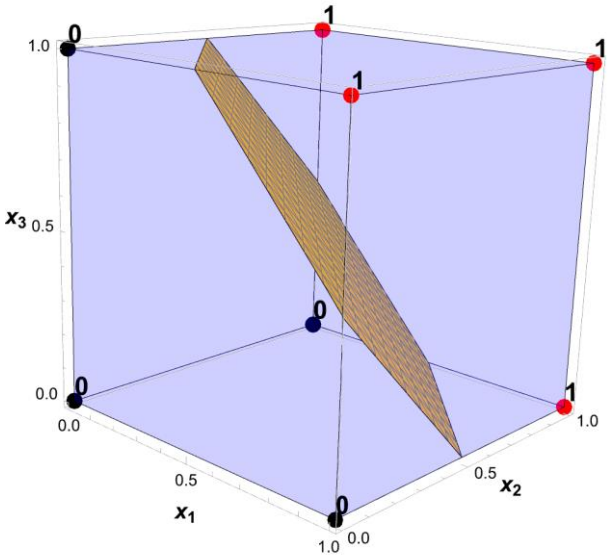

Рис. 2. Разделение вершин куба плоскостью $x_1 + x_2 + x_3 - \frac{3}{2} = 0$
Fig. 2. Separation of the cube vertices by the plane $x_1 + x_2 + x_3 - \frac{3}{2} = 0$

Такое разделение дает новую раскраску, которая не учитывается соотношением (7). Поворотами куба можно получить еще семь подобных раскрасок, что и объясняет различие между оценкой и точным значением в табл. 2 при $k = 4$, $s = 3$.

Применение ИНС для аппроксимации двоичных функций в задачах защиты информации

Аппроксимация порождающей функции генераторов псевдослучайных последовательностей. Рассмотрим задачу аппроксимации порождающих функций генераторов псевдослучайных последовательностей, основанных на регистрах сдвига с линейной обратной связью (РСЛОС) и регистрах сдвига с нелинейной обратной связью (РСНОС), задаваемых следующим рекуррентным соотношением общего вида:

$$x_\tau = f(x_{\tau-1}, x_{\tau-2}, \dots, x_{\tau-s}), \tau = s+1, s+2, \dots .$$

Для применения ИНС (2) обучающая выборка формировалась следующим образом:

$$x^{(t)} ::= (x_{t-1}, \dots, x_{t-s}), y^{(t)} ::= x_t, t \geq s+1.$$

Исследовались два РСЛОС [8] и шесть РСНОС [9]. Для каждого из них найдено наименьшее число нейронов m_{min}, необходимых для безошибочной аппроксимации ($\alpha = 1$). Результаты представлены в табл. 3.

Таблица 3
Аппроксимация порождающих функций РСЛОС и РСНОС
Table 3
Approximation of generating functions of LFSR and NLFSR

Число переменных s *Number of variables s*	Вид функции f *Function type f*	Число нейронов m_{min} *Number of neurons* m_{min}
7	$f = x_1 \oplus x_5$	2
15	$f = x_1 \oplus x_9$	2
17	$f = x_1 \oplus x_2 \oplus x_8 x_{11} \oplus x_{10} x_{16}$	6
17	$f = x_1 \oplus x_7 \oplus x_3 x_{10} \oplus x_8 x_{13}$	6
17	$f = x_1 \oplus x_2 \oplus x_4 \oplus x_{10} \oplus x_{13} \oplus x_8 x_{14}$	7
17	$f = x_1 \oplus x_2 \oplus x_8 \oplus x_{12} \oplus x_{14} \oplus x_7 x_{15}$	7
17	$f = x_1 \oplus x_4 \oplus x_9 \oplus x_{12} \oplus x_{13} \oplus x_4 x_{12}$	7
24	$f = x_1 \oplus x_2 \oplus x_9 \oplus x_{10} \oplus x_{16} \oplus x_8 x_{19}$	7

В ходе обучения ИНС при решении задачи аппроксимации РСЛОС и РСНОС было замечено, что веса, исходящие на рис. 1 из фиктивных входных переменных (т. е. переменных, не влияющих на значение функции f), стремились к нулевому значению. Это показывает, что двухслойная ИНС (2) способна находить фиктивные переменные и исключать их влияние на результат аппроксимации.

Аппроксимация блоков подстановки. Рассмотрим задачу аппроксимации блока подстановки известного стандартного алгоритма шифрования ГОСТ 28147-89 [10] с точностью $\alpha = 1$. Этот блок подстановки представлен в шестнадцатеричной записи в табл. 4.

Таблица 4
Блок подстановки ГОСТ 28147-89
Table 4
Substitution block GOST 28147-89

x	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
K_1	4	2	F	5	9	1	0	8	E	3	B	C	D	7	A	6
K_2	C	9	F	E	8	1	3	A	2	7	4	D	6	0	B	5
K_3	D	8	E	C	7	3	9	A	1	5	2	4	6	F	0	B
K_4	E	9	B	2	5	F	7	1	0	D	C	6	A	4	3	8
K_5	3	E	5	9	6	8	0	D	A	B	7	C	2	1	F	4
K_6	8	F	6	B	1	9	C	5	D	3	7	A	0	E	2	4
K_7	9	B	C	0	3	6	7	5	4	8	E	F	1	A	2	D
K_8	C	6	5	2	B	0	9	D	3	E	7	A	F	4	1	8

Блок подстановки представляет собой восемь функций $K_1, \dots, K_8$, действующих из V^4 в V^4. Каждую из этих векторных функций K_i можно разделить на четыре координатные функции, определяющие соответствующий бит значений функции K_i:

$$K_i(x) = y = (y_1, y_2, y_3, y_4) = \big(K_{i1}(x), K_{i2}(x), K_{i3}(x), K_{i4}(x)\big), x \in V^4, y \in V^4,$$
$$K_{il}(\cdot): V^4 \to V, i \in \{1, \dots, 8\}, l \in \{1, 2, 3, 4\}.$$

Для каждой координатной функции $K_{il}(\cdot)$ найдено наименьшее значение числа нейронов m_{min}, необходимых для ее аппроксимации. Результаты представлены в табл. 5.

Таблица 5
Минимальное число нейронов m_{min} для аппроксимации $K_{il}(\cdot)$
Table 5
Minimum number of neurons m_{min} for approximation of $K_{il}(\cdot)$

l	i							
	1	2	3	4	5	6	7	8
1	3	3	2	3	2	3	2	3
2	2	3	2	2	3	3	2	3
3	2	2	2	3	2	2	2	2
4	2	3	2	2	3	2	3	2

Заключение. В работе исследованы математические особенности применения ИНС для решения задач аппроксимации двоичных функций многих переменных, заключающиеся в наличии областей кусочного постоянства и многоэкстремальности целевой функции. Результаты статьи позволяют выбирать параметры двухслойной ИНС для повышения точности аппроксимации и иллюстрируются примерами аппроксимации порождающих функций генераторов псевдослучайных последовательностей.

Вклад авторов. *К. В. Латушкин* – аналитическое исследование, написание текста статьи, программная реализация модели, проведение численных экспериментов. *Ю. С. Харин* – построение модели, аналитическое исследование, подготовка текста статьи, анализ и интерпретация результатов исследования.

Список использованных источников

1. Gohr, A. Improving attacks on round-reduced speck32/64 using deep learning / A. Gohr // Advances in Cryptology – CRYPTO 2019: 39th Annual Intern. Cryptology Conf., Santa Barbara, CA, USA, 18–22 Aug. 2019. – Santa Barbara, 2019. – Pt. II. – P. 150–179.

2. Deep learning-based physical side-channel analysis / S. Picek, G. Perin, L. Mariot [et al.] // ACM Computing Surveys. – 2023. – Vol. 55(11). – P. 1–35.

3. Boanca, S. Exploring patterns and assessing the security of pseudorandom number generators with machine learning / S. Boanca // 16th Intern. Conf. on Agents and Artificial Intelligence, Rome, Italy, 24–26 Febr. 2024. – Rome, 2024. – Vol. 3. – P. 186–193.

4. Бетелин, В. Б. Математические задачи, связанные с искусственным интеллектом и искусственными нейронными сетями / В. Б. Бетелин, В. А. Галкин // Успехи кибернетики. – 2021. – Т. 2, № 4. – С. 6–14. – DOI: 10.51790/2712-9942-2021-2-4-1.

5. Николенко, С. Глубокое обучение. Погружение в мир нейронных сетей / С. Николенко, А. Кадурин, Е. Архангельская. – СПб. : Питер, 2018. – 480 с.

6. Glorot, X. Understanding the difficulty of training deep feedforward neural networks / X. Glorot, Y. Bengio // Proc. of the Thirteenth Intern. Conf. on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010. – Sardinia, 2010. – Vol. 9. – P. 249–256.

7. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification / K. He, X. Zhang, S. Ren, J. Sun // Proc. of the 2015 IEEE Intern. Conf. on Computer Vision, Santiago, Chile, 7–13 Dec. 2015. – Santiago, 2015. – P. 1026–1034.

8. Криптология / Ю. С. Харин, С. В. Агиевич, Д. В. Васильев, Г. В. Матвеев. – Минск : БГУ, 2023. – 511 с.

9. Dubrova, E. A list of maximum period NLFSRs / E. Dubrova // Cryptology ePrint Archive. – 2012. – URL: https://eprint.iacr.org/2012/166 (date of access: 18.04.2025).

10. Программирование алгоритмов защиты информации / А. В. Домашев, М. М. Грунтович, В. О. Попов [и др.]. – М. : Нолидж, 2002. – 416 с.

References

1. Gohr A. Improving attacks on round-reduced speck32/64 using deep learning. *Advances in Cryptology – CRYPTO 2019: 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, 18–22 August 2019*, pt. II, pp. 150–179.

2. Picek S., Perin G., Mariot L., Wu L., Batina L. *Deep learning-based physical side-channel analysis. ACM Computing Surveys*, 2023, vol. 55(11), pp. 1–35.

3. Boanca S. Exploring patterns and assessing the security of pseudorandom number generators with machine learning. *16th International Conference on Agents and Artificial Intelligence, Rome, Italy, 24–26 February 2024*, vol. 3, pp. 186–193.

4. Betelin V. B., Galkin V. A. *Mathematical problems of artificial intelligence and artificial neural networks*. Uspekhi kibernetiki [*Russian Journal of Cybernetics*], 2021, vol. 2, no. 4, pp. 6–14 (In Russ.). DOI: 10.51790/2712-9942-2021-2-4-1.

5. Nikolenko S., Kadurin A., Arkhangelskaya E. Glubokoe obuchenie. Pogruzhenie v mir nejronnyh setej. *Deep Learning. Dive into the World of Neural Networks.* Saint Petersburg, Piter, 2018, 480 p. (In Russ.).

6. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010*, vol. 9, pp. 249–256.

7. He K., Zhang X., Ren S., Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. *Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015*, pp. 1026–1034.

8. Kharin Yu. S., Agievich S. V., Vasilyev D. V., Matveev G. V. Kriptologiya. *Cryptology*. Minsk, Belorusskij gosudarstvennyj universitet, 2023, 511 p. (In Russ.).

9. Dubrova E. A list of maximum period NLFSRs. *Cryptology ePrint Archive*, 2012. Available at: https://eprint.iacr.org/2012/166 (accessed 18.04.2025).

10. Domashev A. V., Gruntovich M. M., Popov V. O., Pravikov D. I., Shcherbakov A. Y., Prokofyev I. V. Programmirovanie algoritmov zashhity informacii. *Programming of Information Security Algorithms*. Moscow, Nolidzh, 2002, 416 p. (In Russ.).

Информация об авторах

Латушкин Константин Вадимович, младший научный сотрудник, НИИ прикладных проблем математики и информатики Белорусского государственного университета.
E-mail: LatushkinKV@bsu.by

Харин Юрий Семенович, доктор физико-математических наук, академик НАН Беларуси, профессор, НИИ прикладных проблем математики и информатики Белорусского государственного университета.
E-mail: Kharin@bsu.by

Information about the authors

Konstantin V. Latushkin, Junior Researcher, Research Institute of Applied Problems of Mathematics and Informatics of the Belarusian State University.
E-mail: LatushkinKV@bsu.by

Yuriy S. Kharin, D. Sc. (Phys.-Math.), Acad. of the National Academy of Science of Belarus, Prof., Research Institute of Applied Problems of Mathematics and Informatics of the Belarusian State University.
E-mail: Kharin@bsu.by