МОДЕЛИРОВАНИЕ ПОЛНОЙ ЭНЕРГИИ МОЛЕКУЛЫ УГЛЕРОДА В ПЕРВОМ ПОРЯДКЕ ТЕОРИИ ВОЗМУЩЕНИЙ С ПОМОЩЬЮ АЛГОРИТМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Ю. Н. Штанов[1,а], **В. П. Кощеев**[2,б]

[1] *Тюменский индустриальный университет, филиал «Сургутский институт нефти и газа», г. Сургут, Российская Федерация*

[2] *Московский авиационный институт (национальный исследовательский университет), филиал «Стрела», г. Жуковский, Московская область, Российская Федерация*

[а] *ORCID: https://orcid.org/0000-0002-5822-3368,* ✍ *yuran1987@mail.ru*

[б] *ORCID: https://orcid.org/0000-0002-0724-9760, koshcheev1@yandex.ru*

Аннотация: в работе рассматривается задача вычисления полной энергии молекулы углерода с помощью алгоритма искусственного интеллекта. Разработанный алгоритм искусственного интеллекта сравнивается с методом численного интегрирования и известными алгоритмами искусственного интеллекта: нейросеть Колмогорова–Арнольда и случайный лес. Оценка качества разработанных алгоритмов производится с использованием метрик R^2, MAE, RMSE. Особое внимание уделяется сравнительному анализу точности, вычислительной эффективности и области применимости каждого подхода.

Ключевые слова: теория возмущений, полная энергия, машинное обучение, Python, случайный лес, многослойный перцептрон, искусственный интеллект.

Для цитирования: Штанов Ю. Н., Кощеев В. П. Моделирование полной энергии молекулы углерода в первом порядке теории возмущений с помощью алгоритма искусственного интеллекта. *Успехи кибернетики.* 2026;7(3):139–147.

Поступила в редакцию: 29.06.2026. *В окончательном варианте: 20.09.2026.*

CALCULATION OF THE TOTAL ENERGY OF A CARBON MOLECULE USING FIRST-ORDER PERTURBATION THEORY AND MACHINE LEARNING

Yu. N. Shtanov[1,a], **V. P. Koshcheev**[2,b]

[1] *Industrial University of Tyumen, Surgut Branch, Surgut, Russian Federation*

[2] *Moscow Aviation Institute (National Research University), Strela Branch, Zhukovsky, Moscow Region, Russian Federation*

[а] *ORCID: https://orcid.org/0000-0002-5822-3368,* ✍ *yuran1987@mail.ru*

[б] *ORCID: https://orcid.org/0000-0002-0724-9760, koshcheev1@yandex.ru*

Abstract: we developed a machine-learning approach for calculating the total energy of a carbon molecule. We compared its performance with numerical integration and two established machine-learning methods: a Kolmogorov–Arnold network and a random forest. We evaluated the accuracy of the methods using the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). We compared the methods in terms of accuracy, computational efficiency, and applicability to different calculation tasks.

Keywords: perturbation theory, total energy, machine learning, Python, random forest, multilayer perceptron, artificial intelligence.

Cite this article: Shtanov Yu. N., Koshcheev V. P. Calculation of the Total Energy of a Carbon Molecule Using First-Order Perturbation Theory and Machine Learning. *Russian Journal of Cybernetics.* 2026;7(3):139–147.

Original article submitted: 29.06.2026. *Revision submitted: 20.09.2026.*

Введение

С помощью компьютерного эксперимента в [1, 2] исследуются новые типы упорядоченных структур, состоящих из атомарного углерода. Потенциальная энергия взаимодействия между атомами вычисляется с помощью уравнения теории функционала плотности, решение которого для основного состояния совпадает с решением уравнения Хартри–Фока [3]. С другой стороны, достоверность вычисления с помощью многочисленных версий теории функционала плотности

достигается тогда, когда результаты совпадают между собой [4]. В работе [5] показано, что полная энергия трехатомной молекулы может быть вычислена как сумма полных энергий двухатомных молекул с помощью метода компьютерного моделирования в первом порядке теории возмущений.

Аппроксимация многомерных функций является весьма актуальной задачей машинного обучения [6, 7]. В настоящей работе предлагается алгоритм аппроксимации сложной многомерной функции с использованием глубокого многослойного перцептрона [8] с предварительной активацией слоя, адаптивным интервальным разбиением области определения для локализации сложности функции. Алгоритмы реализованы на языке Python с использованием фреймворка PyTorch и scikit-learn [9]. Полученный алгоритм сравнивается с алгоритмом машинного обучения случайный лес (Random Forest) [9] и нейросетью Колмогорова–Арнольда (Kolmogorov–Arnold Networks, KAN) [10]. Оценка качества разработанных алгоритмов производится с использованием метрик MAE, RMSE и R^2.

Полная энергия системы

Трехатомную молекулу будем описывать с помощью стационарного уравнения Шредингера:

$$H\psi = E\psi. \tag{1}$$

Гамильтониан уравнения (1) представим в виде:

$$H = H^0 + U. \tag{2}$$

$$\begin{aligned} U = {} & \tfrac{Z_1 Z_2 e^2}{|r_1 - r_2|} + \sum_{j_1=1}^{Z_1} \sum_{j_2=1}^{Z_2} \tfrac{e^2}{|r_1 + r_{1j_1} - r_2 - r_{2j_2}|} - \sum_{j_1=1}^{Z_1} \tfrac{Z_2 e^2}{|r_1 + r_{1j_1} - r_2|} - \sum_{j_2=1}^{Z_2} \tfrac{Z_1 e^2}{|r_1 - r_2 - r_{2j_2}|} + \\ & + \tfrac{Z_1 Z_3 e^2}{|r_1 - r_3|} + \sum_{j_1=1}^{Z_1} \sum_{j_3=1}^{Z_3} \tfrac{e^2}{|r_1 + r_{1j_1} - r_3 - r_{3j_3}|} - \sum_{j_1=1}^{Z_1} \tfrac{Z_3 e^2}{|r_1 + r_{1j_1} - r_3|} - \sum_{j_3=1}^{Z_3} \tfrac{Z_1 e^2}{|r_1 - r_3 - r_{3j_3}|} + \\ & + \tfrac{Z_3 Z_2 e^2}{|r_3 - r_2|} + \sum_{j_3=1}^{Z_3} \sum_{j_2=1}^{Z_2} \tfrac{e^2}{|r_3 + r_{3j_3} - r_2 - r_{2j_2}|} - \sum_{j_3=1}^{Z_3} \tfrac{Z_2 e^2}{|r_3 + r_{3j_3} - r_2|} - \sum_{j_2=1}^{Z_2} \tfrac{Z_3 e^2}{|r_3 - r_2 - r_{2j_2}|}, \end{aligned} \tag{3}$$

где $U = U_{12} + U_{13} + U_{23}$— потенциальная энергия взаимодействия трех атомов; r_1, r_2и r_3 — координаты первого, второго и третьего атомного ядра; $r_1 + r_{1j_1}$, $r_2 + r_{2j_2}$ и $r_3 + r_{3j_3}$ — координаты j_1-го, j_2-го и j_3-го электронов первого, второго и третьего атома соответственно; $r_1 - r_2 = r_{12} + \delta r_{12}$, $r_1 - r_3 = r_{13} + \delta r_{13}$, $r_3 - r_2 = r_{32} + \delta r_{32}$; коллективные колебания атомных электронов и ядер описываются вектором $\delta r_{mn} = \delta r_{mn}^{\text{nucl}} + \delta r_{mn}^{\text{el}}$; $r_{mn} = |r_{mn}|$ — расстояние между n-ым и m-ым атомами в кластере.

Решение уравнения (1) с гамильтонианом (2) будем искать с помощью теории возмущений $\psi = \psi^0 + \psi^1 + \ldots$ и $E = E^0 + E^1 + \ldots$. Потенциальную энергию (электронные термы) трехатомного кластера будем искать в первом порядке теории возмущений:

$$E^1 = \left\langle \psi^0 \right| U \left| \psi^0 \right\rangle, \tag{4}$$

где угловые скобки $\langle \ldots \rangle$ были введены Дираком [11].

Гамильтониан H^0 представим в виде $H^0 = H_1^0 + H_2^0 + H_3^0$, где H_i^0 — гамильтониан i-го атома; i=1,2,3.

Решение уравнения Шредингера $H^0\psi^0 = E^0\psi^0$ будем искать в виде $\psi^0 = \psi_1^0\psi_2^0\psi_3^0$ и $E^0 = E_1^0 + E_2^0 + E_3^0$, где уравнение Шредингера для i-го изолированного атома имеет вид:

$$H_i^0\psi_i^0 = E_i^0\psi_i^0, \tag{5}$$

где $\psi_i^0 = \psi_i^0\,(r_{i1}, r_{i2}, \ldots, r_{iZ_i})$.

Усреднение по квантовым флуктуациям местоположения атомных электронов будем осуществлять с помощью метода [12], который Бете использовал для вычисления атомного формфактора, а усреднение по коллективным колебаниям атомных электронов и ядер выполним по квадрату модуля волновой функции гармонического осциллятора в основном состоянии. Средние величины будем обозначать $\left\langle \psi_1^0 \right| U \left| \psi_1^0 \right\rangle = \langle \ldots \rangle_{e1}$, $\left\langle \psi_2^0 \right| U \left| \psi_2^0 \right\rangle = \langle \ldots \rangle_{e2}$, $\left\langle \psi_3^0 \right| U \left| \psi_3^0 \right\rangle = \langle \ldots \rangle_{e3}$ и $\left\langle \psi^0 \right| U \left| \psi^0 \right\rangle_{\text{pl}} = \langle \ldots \rangle_{\text{pl}}$.

Разложим потенциальную энергию взаимодействия (3) в интеграл Фурье:

$$U = \int \frac{d^3k}{(2\pi)^3} \frac{4\pi e^2}{k^2} \left[Z_1 Z_2 + \frac{1}{2} \sum_{m \neq n=1}^{3} V_{mn}(k) \exp\left(ik\delta r_{mn}\right) \right] \exp\left(ikr_{mn}\right), \tag{6}$$

где $V_{12}(k) = \sum_{j_1=1}^{Z_1} \sum_{j_2=1}^{Z_2} \exp\left(ik\left(r_{1j_1} - r_{2j_2}\right)\right) - Z_2 \sum_{j_1=1}^{Z_1} \exp\left(ikr_{1j_1}\right) - Z_1 \sum_{j_2=1}^{Z_2} \exp\left(ikr_{2j_2}\right)$.

Усредним потенциальную энергию взаимодействия трех атомов (6):

$$\langle U \rangle = \frac{1}{2} \sum_{m \neq n=1}^{3} \int \frac{d^3k}{(2\pi)^3} \frac{4\pi e^2}{k^2} \left[Z_m - F_m(k) \right] \left[Z_n - F_n(k) \right] \exp\left[ikr_{mn} - k^2 \Delta_{mn} \right], \tag{7}$$

где $\left\langle \sum_{j_m=1}^{Z_m} \exp\left(ikr_{mj_m}\right) \right\rangle_{em} = F_m(k)$; $F_m(0) = Z_m$ — атомные форм-факторы первого, второго и третьего атомов; $\left\langle \exp\left(ik \sum_{m \neq n=1}^{3} \delta r_{mn} \right) \right\rangle_{\text{pl}} = \exp\left[-k^2 \sum_{m \neq n=1}^{3} \Delta_{mn} \right]$;$\Delta_{mn} = \Delta_{nm}$; $2\Delta_{mn} = \sigma_{\text{nucl}}^2 + \sigma_{\text{el}}^2$ — сумма средних квадратов амплитуды коллективных атомных ядерных $\sigma_{\text{nucl}}^2 = \frac{\hbar}{2\omega_{\text{nucl}}\mu_{\text{nucl}}}$ и электронных $\sigma_{\text{el}}^2 = \frac{\hbar}{2\omega_{\text{el}}\mu_{\text{el}}}$ колебаний в расчете на одну степень свободы; упругая постоянная $\omega_{\text{el}}^2\mu_{\text{el}} = \omega_{\text{nucl}}^2\mu_{\text{nucl}} = U''_{mn}(r_{mn}^{\min})$ есть значение второй производной в минимуме потенциальной энергии между n-ым и m-ым атомами в кластере.

Так как $\omega_{\text{el}}^2/\omega_{\text{nucl}}^2 = \mu_{\text{nucl}}/\mu_{\text{el}} \gg 1$, то квант энергии коллективных колебаний электронов (энергия плазмона) равен:

$$\hbar\omega_{\text{el}} \approx 4\Delta_{mn} U''_{mn}\left(r_{mn}^{\min}\right), \tag{8}$$

где $r_{mn}^{\min}$ находим из решения уравнения $\partial U_{mn}\left(r_{mn}^{\min}, \Delta\right)/\partial r = 0$.

В результате поиска максимального значения уравнения (8) получим единственное решение, содержащее непосредственно измеряемые величины $\hbar\omega_{el}$ и $r_{\min}$.

Известно, что волновая функция трехатомного кластера должна быть антисимметричной относительно перестановки координат, определяющих местоположения электронов. Если волновая функция трехатомного кластера выбрана в виде произведения волновых функций изолированных атомов $\psi^0 = \psi_1^0\psi_2^0\psi_3^0$, то обменные силы будут учитываться только для электронов изолированных атомов.

Выражение (7) для потенциальной энергии взаимодействия без учета обменных сил между электронами различных изолированных атомов запишем в виде:

$$\begin{gathered} \langle U \rangle = \tfrac{1}{2} \sum_{m \neq n=1}^{3} U\left(r_{mn}, \Delta_{mn}\right); \\ U\left(r_{mn}, \Delta_{mn}\right) = Z_m Z_n e^2 \Phi\left(r_{mn}, \Delta_{mn}\right)/r_{mn}, \end{gathered} \tag{9}$$

где $\Phi\left(r_{mn}, \Delta_{mn}\right)$ — функция экранирования потенциальной энергии двухатомного кластера (молекулы) [5] имеет вид:

$$\Phi\left(r_{mn}, \Delta_{mn}\right) = \frac{2}{\pi} \int_0^\infty \frac{1}{k} \left[1 - \frac{F_m(k)}{Z_m} \right] \left[1 - \frac{F_n(k)}{Z_n} \right] \sin\left(kr_{mn}\right) \exp\left[-k^2 \Delta_{mn} \right] dk. \tag{10}$$

Известно [3], что вместо многочастичного уравнения Шредингера (5) строятся одночастичные приближения на основе уравнений Хартри–Фока или формализма теории функционала плотности. Атомный форм-фактор является Фурье-компонентой плотности распределения атомных электронов:

$$F_m(k) = \int n_m(r) \exp\left(-ikr\right) d^3r, \tag{11}$$

где $n_{1,2}(r) = N_1 \left|\psi_{1\text{s}}(r)\right|^2 + N_2 \left|\psi_{2\text{s}}(r)\right|^2 + N_3 \left|\psi_{2\text{p}}(r)\right|^2$ — электронная плотность для атома, N_1, N_2, N_3 — количество электронов на оболочках 1s, 2s, 2p.

Волновые функции, которые аппроксимируют решение уравнения Хартри–Фока для изолированных атомов и энергии E_i^0, представлены в [13], тогда выражение (11) после вычисления интеграла для $Z = 2 \dots 10$ примет вид

$$F_m(k) = N_1 F_{1s}(k,\alpha) + N_2 F_{2s}(k,\beta) + N_3 F_{2p}(k,\gamma), \tag{12}$$

$$F_{1s}(k,A) = F_{2s}(k,c) = \sum_{i=1}^{6}\sum_{j=1}^{6} A_i A_j \times \begin{cases} \frac{8\xi_i^{3/2}\xi_j^{3/2}\xi_{ij}}{\left(\xi_{ij}^2+k^2\right)^2}, & i<2\ 8\ j<2 \\ \frac{8\xi_i^{3/2}\xi_j^{5/2}\left(3\xi_{ij}^2-k^2\right)}{\sqrt{3}\left(\xi_{ij}^2+k^2\right)^3}, & i<2\neq j<2 \\ \frac{32\xi_i^{5/2}\xi_j^{5/2}\xi_{ij}\left(\xi_{ij}^2-k^2\right)}{\left(\xi_{ij}^2+k^2\right)^4}, & i>2,j>2 \end{cases}$$

$$F_{2p}(k,\gamma) = 32\sum_{i=7}^{10}\sum_{j=7}^{10}\gamma_{i-6}\gamma_{j-6}(\xi_i\xi_j)^{5/2}\frac{\xi_{ij}\left(\xi_{ij}^2-5k^2\right)}{\left(k^2+\xi_{ij}^2\right)^4},$$

где $\xi_{ij} = \xi_i + \xi_j$, α,β,γ,ξ — константы приближения из работы [13].

Выражение (11) для $Z = 11 \dots 12$ запишется в виде:

$$F_m(k) = N_1 F_{1s}(k,\alpha) + N_2 F_{2s}(k,\beta) + N_3 F_{3s}(k,\gamma) + N_4 F_{2p}(k,\eta), \tag{13}$$

где $F_{1s}(k,c) = F_{2s}(k,c) = F_{3s}(k,c) = \sum_{i=1}^{8}\sum_{j=1}^{8} c_i c_j \times I_{ij}(k)$, $F_{2p}(k,c) = \sum_{i=9}^{13}\sum_{j=9}^{13} c_{i-8}c_{j-8}Q_{ij}(k)$,

$$I_{ij}(k) = \begin{cases} \frac{16\xi_1^4}{\left[4\xi_1^2+k^2\right]^2}, & i=j=1 \\ \frac{32\sqrt{10}}{5}\xi_i^{3/2}\xi_j^{7/2}\frac{\xi_{ij}(\xi_{ij}^2-k^2)}{\left[\xi_{ij}^2+k^2\right]^4}, & i=1,j\geq 2 \text{ или } i\geq 2,j=1 \\ \frac{128}{3}(\xi_i\xi_j)^{7/2}\frac{\xi_{ij}(3\xi_{ij}^4-10\xi_{ij}^2k^2+3k^4)}{\left[\xi_{1j}^2+k^2\right]^6}, & i,j>2 \end{cases}$$

$$Q_{ij}(k) = \begin{cases} \frac{64\xi_9^6(4\xi_9^2-5k^2)}{\left[4\xi_9^2+k^2\right]^4}, & i=j=9 \\ \frac{192}{\sqrt{105}}\xi_9^{5/2}\xi_j^{9/2}\frac{\xi_{ij}\left(21k^4-54k^2\xi_{ij}^2+5\xi_{ij}^4\right)}{\left(k^2+\xi_{ij}^2\right)^6}, & \text{иначе} \\ \frac{512}{7}(\xi_i\xi_j)^{9/2}\frac{\xi_{ij}(7\xi_{ij}^6-133\xi_{ij}^4k^2+169\xi_{ij}^2k^4-27k^6)}{\left[\xi_{1j}^2+k^2\right]^8}, & i,j>9. \end{cases}$$

Выражение (11) для $Z = 13 \dots 18$ запишется в виде:

$$F_m(k) = N_1 F_{1s}(k,\alpha) + N_2 F_{2s}(k,\beta) + N_3 F_{3s}(k,\gamma) + N_4 F_{2p}(k,\eta) + N_5 F_{3p}(k,\zeta), \tag{14}$$

где $F_{1s}(k,c) = F_{2s}(k,c) = F_{3s}(k,c) = 16\sum_{i=1}^{8}\sum_{j=1}^{8} c_i c_j I_{ij}(k)$, $\xi_{ij8} = \xi_{i+8} + \xi_{j+8}$

$$I_{ij}(k) = \begin{cases} \frac{\xi_1^4}{(4\xi_1^2+k^2)^2}, & i=j=1 \\ \frac{6\sqrt{10}}{15}\xi_{\min(i,j)}^{3/2}\xi_{\max(i,j)}^{7/2}\cdot\frac{\xi_{ij}[\xi_{ij}^2-k^2]}{[\xi_{ij}^2+k^2]^4}, & \min(i,j)=1, \max(i,j)\neq 1 \\ \frac{8}{3}\xi_i^{7/2}\xi_j^{7/2}\cdot\frac{\xi_{ij}[3\xi_{ij}^4-10\xi_{ij}^2k^2+3k^4]}{[\xi_{ij}^2+k^2]^6}, & i>1,j>1 \end{cases}$$

$$F_{2p}(k,c) = F_{3p}(k,c) = \frac{64c_1^2\xi_9^6\left(4\xi_9^2-5k^2\right)}{\left(k^2+4\xi_9^2\right)^4} + \\ + \frac{384c_1\xi_9^{5/2}}{\sqrt{105}}\sum_{j=2}^{8}\frac{c_j\xi_{j+8}^{9/2}(\xi_9+\xi_{j+8})\left[21k^4-54k^2(\xi_9+\xi_{j+8})^2+5(\xi_9+\xi_{j+8})^4\right]}{(k^2+(\xi_9+\xi_{j+8})^2)^6} - \\ - \frac{512}{7}\sum_{i=2}^{8}\sum_{j=2}^{8}\frac{c_ic_j\xi_{i+8}^{9/2}\xi_{j+8}^{9/2}[27k^6\xi_{ij8}-169k^4\xi_{ij8}^3+133k^2\xi_{ij8}^5-7\xi_{ij8}^7]}{\left[\xi_{ij8}^2+k^2\right]^8}.$$

Программная реализация

Для решения задачи компьютерного моделирования полной энергии взаимодействия двух атомов была разработана программа TotalEnergy [14, 15]. Программа написана на языке программирования C++, а графический интерфейс строится с использованием библиотеки Qt 6 [16]. Для вычисления выражения (9) используется численное интегрирование с применением библиотеки GNU Scientific Library (GSL) [17].

Аппроксимация функции экранирования $\Phi(r,\Delta)$ представляет собой довольно сложную задачу, так как при малых значениях Δ возникают быстрые осцилляции, и функция смещается в область малых r/a_0, а при больших значениях Δ функция становится гладкой и смещается в область больших r/a_0. При решении задачи аппроксимации $\Phi(r,\Delta)$ было выявлено, что классические подходы имеют ограничения:

1. Кубическая сплайн-аппроксимация требует правильного выбора сетки и не масштабируется на случай функций многих переменных.

2. Обученные на всем интервале алгоритмы машинного обучения не позволяют получить гладкую функцию, нейросетевые алгоритмы плохо обучаются из-за осцилляций функции.

Поэтому предлагается подход, сочетающий в себе преимущества сплайнов и нейросетей. Интервальное разбиение области по r/a_0 с перекрытием выполнено с помощью метода k-средних [9, 18]. В качестве признаков, характеризующих функцию на каждом интервале, добавлялись базисы B-сплайнов порядка $p=3$ (кубические сплайны) с G внутренними узлами и количеством $M=G+p-1$, вычисленные с помощью формулы Кокса – де Бура [19]:

$$B_{i,p}(x) = \frac{x - t_i}{t_{i+p} - t_i} \cdot B_{i,p-1}(x) + \frac{t_{i+p+1} - x}{t_{i+p+1} - t_{i+1}} \cdot B_{i+1,p-1}(x), \tag{15}$$

$$B_{i,0}(x) = \begin{cases} 1, & t_i \le x \le t_{i+1} \\ 0, & \text{иначе} \end{cases},$$

и Фурье-признаки, отвечающие за осцилляции функции [20]:

$$C(x) = [\sin(x \cdot c_1), \cos(x \cdot c_1); \ldots \sin(x \cdot c_m), \cos(x \cdot c_m)], \tag{16}$$

где m — количество признаков, c_i — независимое случайное число, распределенное по нормальному закону с нулевым математическим ожиданием и дисперсией, отвечающей за масштаб частот.

Нейросетевой алгоритм будет строиться на основе многослойного персептрона, каждый слой которого представляет собой Residual блок [21], в котором реализуется принцип «активация на ребрах» [22]. В качестве функции активации была выбрана SiLU [23]. В результате алгоритм построения Residual блока имеет вид:

$$\begin{array}{l} z = LN(x) \\ z = SiLU(z) \\ z = z \cdot \mathrm{A}^{\mathrm{T}} + b_1 \\ z = LN(z) \\ z = SiLU(z) \\ y = z \cdot \mathrm{B}^{\mathrm{T}} + b_2 + s(x) = F(z) + s(x) \end{array}, \tag{17}$$

где A,B,C — матрицы весов, которые подбирают в процессе обучения нейросети; $LN(x_i) = \gamma_i \cdot (x_i - \mu) \Big/ \sqrt{\sigma^2 + \varepsilon} + \beta_i$ — функция для нормализации слоя [24] для i-го нейрона в слое; $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$— математическое ожидание; дисперсия $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$; $\varepsilon > 0$ — малая константа, необходимая для исключения деления на ноль; γ_i,β_i — параметры, настраиваемые вместе с остальными весами нейросети во время обучения, остающиеся фиксированными во время использования; $s(x)$ отвечает за проекционный слой и shortcut-путь:

$$s(x) = \begin{cases} \mathrm{C}^{\mathrm{T}} \cdot x + b_3, & L_{\text{in}} \ne L_{\text{out}} \\ x, & L_{\text{in}} = L_{\text{out}} \end{cases}, \tag{18}$$

где L_{in},L_{out} — размерность входного и выходного вектора.

В алгоритме (17) степень смешивания информации никак не регулируется, что приводит к нежелательным эффектам. Для изменения степени нелинейности воспользуемся механизмом Highway Networks [25], тогда получим для k-го интервала:

$$\hat{\Phi}_k = \alpha\left(Z_1,Z_2,\Delta\right)\cdot\left[g\left(\Delta\right)\cdot F\left(x\right)+\left(1-g\left(\Delta\right)\right)\cdot s\left(x\right)\right], \tag{19}$$

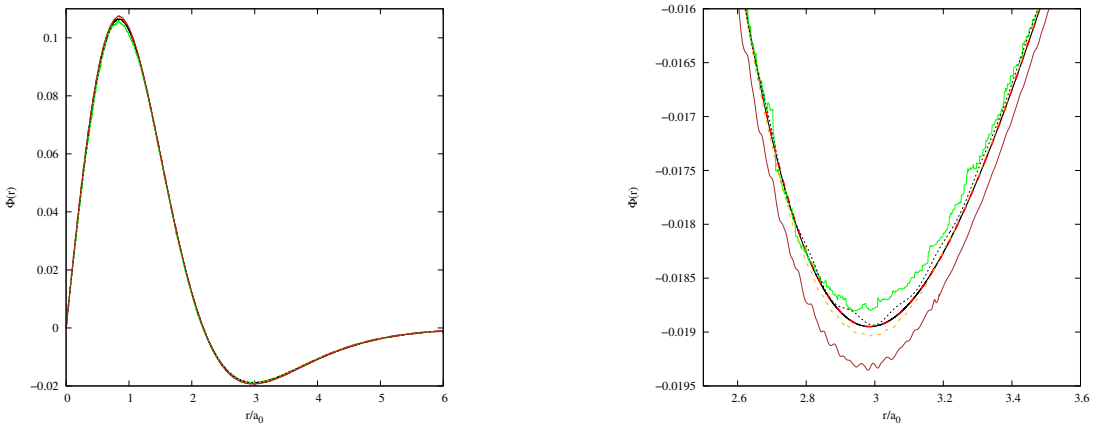
а результирующее предсказание

$$\hat{\Phi}\left(Z_1,Z_2,\Delta,r\right)=\frac{\sum_{k=1}^{K}\omega_k\cdot\hat{\Phi}_k}{\sum_{k=1}^{K}\omega_k}, \tag{20}$$

где $\omega_k=\exp\left(-\varepsilon_k\right)$ – вес при ε_k ошибке на k-м интервале; $g\left(\Delta\right)=MLP_{gate}\left(\Delta\right)$ — трехслойная нейросеть предсказывает значения $g\in[0;1]$, $\alpha(Z_1,Z_2,\,\Delta)=MLP_{amp}(Z_1,Z_2,\Delta)$ — четырехслойная нейросеть предсказывает значения $\alpha\in[0;\infty]$.

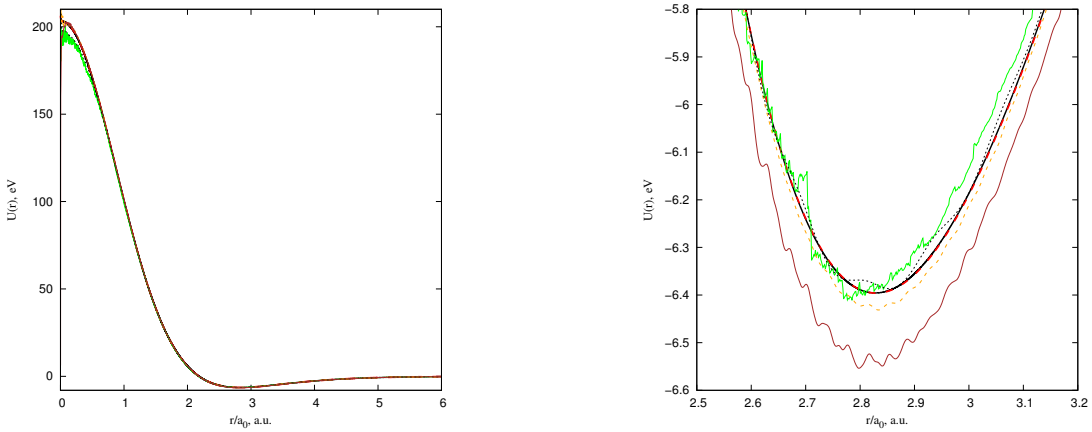
Параметр $g\left(\Delta\right)$ отвечает за степень смешивания и определяет, сколько информации берем от $F\left(x\right)$ и $s\left(x\right)$, а параметр $\alpha(Z_1,Z_2,\,\Delta)$ управляет масштабом. Данный подход можно назвать физически обусловленной вентильной (gate) гиперсвязью, или коротко **Gated Hyper-Connection (GHD)**, он сочетает механизмы Highway Networks, Manifold-Constrained Hyper-Connections (mHC) [26].

С помощью модифицированной программы TotalEnergy был создан датасет для выражения (10) для двухатомных молекул с Z_1=2...21, Z_2=2...21 и $\Delta/a_0^2\in[0.001;1]$, $r/a_0\in[0;6]$. Таким образом, был сформирован набор данных, содержащий $4\cdot10^6$ образцов, каждый из которых представлен вектором базовых признаков $x=(Z_1,Z_2,\Delta,r)$ и соответствующим значением целевой функции $\Phi(r,\Delta)$. Для обучения и тестирования нейросетей использовался язык программирования Python и библиотеки PyTorch [9], выполнялось разделение выборки на обучающую (65 %), валидационную (15 %) и тестовую (20 %). Количество признаков на каждом интервале обучения было равным 40 сплайн- и 20 Фурье-признакам, которые вместе с базовыми признаками были нормированы с помощью нормализации [9], а целевая функция нормирована с помощью масштабирования в диапазон [-1;1]. Архитектура нейросети состояла из 4 Residual блоков с размерностью слоев [512,512,384,256,128] с учетом механизма (20) при $K=5$. Обучение проводилось в течение 500 эпох с механизмом ранней остановки для предотвращения переобучения. При обучении нейросетей использовалась видеокарта Nvidia RTX 5070 с 12 Гб видеопамяти. На рис. 1 а, б выполнено сравнение результатов расчета функции экранирования $\Phi\left(r,\Delta\right)$, на рис. 2 а, б — потенциальной энергии взаимодействия для молекулы углерода C(3p)+C(3p) при $\Delta/a_0^2\approx0.278$, полученных с помощью нейросетей и численного интегрирования.



а *б*

Рис. 1. *а — Сравнение результатов расчета функции экранирования* $\Phi\left(r,\Delta\right)$ *для молекулы углерода C(3p)+C(3p) при* $\Delta/a_0^2=0.278$, $r/a_0\in[0;6]$; *б — Сравнение результатов расчета функции экранирования* $\Phi\left(r,\Delta\right)$ *для молекулы углерода C(3p)+C(3p) при* $\Delta/a_0^2=0.278$, $r/a_0\in[2.5;3.6]$



а *б*

Рис. 2. *а — Сравнение результатов расчета потенциальной энергии* $U\,(r,\Delta)$ *для молекулы углерода C(3p)+C(3p) при* $\Delta/a_0^2 = 0.278$, $r/a_0 \in [0;6]$*; б — Сравнение результатов расчета потенциальной энергии* $U\,(r,\Delta)$ *для молекулы углерода C(3p)+C(3p) при* $\Delta/a_0^2 = 0.278$, $r/a_0 \in [2.5;3.2]$

Кроме нейросети, обученной с учетом механизма (20) при $K = 5$ и $K = 1$, были обучены нейросеть Колмогорова–Арнольда, интервальная модель случайного леса и классический многослойный персептрон (MLP). На рисунках введены следующие обозначения: черная сплошная линия — численное интегрирование; красная пунктирная — нейросеть Колмогорова–Арнольда [10]; зеленая сплошная — алгоритм случайного леса [9]; оранжевая пунктирная — нейросеть с учетом механизма (20) при $K = 5$; черная пунктирная — нейросеть, обученная с учетом механизма (20) при $K = 1$; коричневая сплошная — классический многослойный персептрон. После обучения нейросетей на тестовой выборке было выполнено вычисление метрик MAE, RMSE и R^2, которые представлены в таблице.

Таблица

Показатели обученных моделей (Random Forest, KAN, классический MLP) и нейросетей, обученных с учетом механизма (20) *при K=5, K=1 на тестовой выборке*

Показатели модели	Random Forest	KAN	Классический MLP	Нейросеть, обученная с учетом механизма (20) при $K = 5$	Нейросеть, обученная с учетом механизма (20) при $K = 1$
Время обучения, минут	3	195	179	276	193
R^2	0.997	0.9995	0.999469	0.9994	0.9989
MAE, 10^{-4}	2.54	1.44	3.39	3.16	4
RMSE $\cdot 10^{-4}$	6.58	8.7	9.19	9.95	13.04
MSE $\cdot 10^{-7}$	4376	7.576	8.439	9.906	16.992
SMAPE, %	3.1	2.87	10.8	5.7	8.44

Вывод

Видно, что обученная нейросеть Колмогорова–Арнольда (KAN) показывает результаты, сопоставимые с результатами численного интегрирования. Интервальная модель случайного леса, несмотря на высокое значение R^2, имеет существенный недостаток: отсутствие гладкости в областях экстремумов функции, поэтому можно сделать вывод о ее непригодности для аппроксимации функций. Разработанная нейросеть с механизмом Gated Hyper-Connection при $K = 5$ показывает результаты, наиболее близкие к KAN, но требует дополнительного исследования для

улучшения модели. Нейросетевой подход, апробированный на двухатомной молекуле, показывает перспективные результаты для многоатомных структур и позволяет использовать обученную нейросетевую модель для решения более сложных задач молекулярной физики.

ЛИТЕРАТУРА

1. Borsovszky J. et al. Photodissociation of Dicarbon: How Nature Breaks an Unusual Multiple Bond. *PNAS.* 2021;118(52):e2113315118. DOI: 10.1073/pnas.2113315118.
2. Yurchenko S. N., Szabó I., Pyatenko E., Tennyson J. ExoMol Line Lists XXXI: Spectroscopy of Lowest Eights Electronic States of C_2. *MNRAS.* 2018;480(3):3397–3411. DOI: 10.1093/mnras/sty2050.
3. Sarry A. M., Sarry M. F. On the Density Functional Theory. *Physics of the Solid State.* 2012;54:1315–1322. DOI: 10.1134/S1063783412060297.
4. Teale A. M. et al. DFT Exchange: Sharing Perspectives on the Workhorse of Quantum Chemistry and Materials Science. *Phys. Chem. Chem. Phys.* 2022;24(47):28700-28781. DOI: 10.1039/D2CP02827A.
5. Koshcheev V. P., Shtanov Y. N. Simulation of the Total Energy of a Triatomic Silicon Cluster in the First Order of Perturbation Theory. 2025. arXiv:2507.00053. DOI: 10.48550/arXiv.2507.00053.
6. Галкин В. А., Гавриленко Т. В., Смородинов А. Д. Некоторые аспекты аппроксимации и интерполяции функций искусственными нейронными сетями. *Вестник КРАУНЦ. Физ.-мат. науки.* 2022;38(1):54–73. DOI: 10.26117/2079-6641-2022-38-1-54-73.
7. Галабурдин А. В. Аппроксимация функций нейронными сетями. *Computational Mathematics and Information Technologies.* 2026;10(1):50–57. DOI: 10.23947/2587-8999-2026-10-1-50-57.
8. Rumelhart D. E., McClelland J. L. Learning Internal Representations by Error Propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations.* MIT Press. 1987:318–362.
9. Рашка С. *Машинное обучение с PyTorch и Scikit-Learn.* Астана: Фолиант; 2024. 688 с.
10. Liu Z., Wang Y., Vaidya S. et al. KAN: Kolmogorov–Arnold Networks. 2024. arXiv:2404.19756. DOI: 10.48550/arXiv.2404.19756.
11. Дирак П. А. М. *Принципы квантовой механики.* М.: Наука; 1979. 479 с.
12. Бете Г. *Квантовая механика.* М.: Мир; 1965. 333 с.
13. Clementi E., Roetti C. Roothaan-Hartree-Fock Atomic Wavefunctions: Basis Functions and Their Coefficients for Ground and Certain Excited States of Neutral and Ionized Atoms, $Z \leq 54$. *Atomic Data and Nuclear Data Tables.* 1974;14(3):177–478. DOI: 10.1016/S0092-640X(74)80016-1.
14. Штанов Ю. Н., Кощеев В. П. Программа моделирования полной энергии взаимодействия двух атомов в первом порядке теории возмущений (Моделирование полной энергии взаимодействия атомов в первом порядке теории возмущений). Свидетельство о государственной регистрации компьютерных программ № 2024660504. М.: Роспатент; 2024.
15. Штанов Ю. Н., Кощеев В. П., Моргун Д. А. *Библиотека программ «JINRLIB».* Режим доступа: http://wwwinfo.jinr.ru/programs/jinrlib/tropics/index.html.
16. Прохоренок Н. А. *Qt 6. Разработка оконных приложений на C++.* СПб.: БХВ-Петербург; 2022. 512 с.
17. GNU Project. *GNU Scientific Library.* Режим доступа: https://www.gnu.org/software/gsl/.
18. Arthur D., Vassilvitskii S. k-Means++: The Advantages of Careful Seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms.* 2007;8:1027–1035.
19. De Boor C. Subroutine Package for Calculating with B-splines. *Techn. Rep.* 1971:109–121.
20. Tancik M., Srinivasan P. P., Mildenhall B. et al. Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. 2020. arXiv:2006.10739. DOI: 10.48550/arXiv.2006.10739.
21. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. 2015. arXiv:1512.03385. DOI: 10.48550/arXiv.1512.03385.
22. He K., Zhang X., Ren S., Sun J. Identity Mappings in Deep Residual Networks. *Computer Vision – ECCV 2016. Lecture Notes in Computer Science.* Eds: Leibe B., Matas J., Sebe N., Welling M. 2016;9908. Springer, Cham. DOI: 10.1007/978-3-319-46493-0_38.

23. Elfwing S., Uchibe E., Doya K. Sigmoid-Weighted Linear Units for Neural Network Function Approximation in Reinforcement Learning. 2017. arXiv:1702.03118. DOI: 10.48550/arXiv.1702.03118.
24. Ba J. L., Kiros J. R., Hinton G. E. Layer Normalization. 2016. arXiv:1607.06450. DOI: 10.48550/arXiv.1607.06450.
25. Srivastava R. K., Greff K., Schmidhuber J. Highway Networks. 2015. arXiv:1505.00387. DOI: 10.48550/arXiv.1505.00387.
26. Xie Z., Liang W. et al. mHC: Manifold-Constrained Hyper-Connections. 2025. arXiv:2512.24880. DOI: 10.48550/arXiv.2512.24880.