УДК 519.6

**Экстраполяция вариационных расчетов с помощью искусственных нейронных сетей**

**А. О. Белозеров1, А. И. Мазур1, Р. Э. Шарыпов1, А. М. Широков2**

*1Тихоокеанский государственный университет (г. Хабаровск)*

*2НИИ ядерной физики им. Д. В. Скобельцына МГУ (г. Москва)*

*С помощью машинного обучения разработан метод* *экстраполяции наблюдаемых в ядрах – энергии основного состояния и среднеквадратичного радиуса – на основе входных данных, полученных в модели оболочек без инертного кора. Основная идея состоит в том, чтобы построить ансамбль искусственных нейронных сетей для определения среднестатистического значения наблюдаемой характеристики для произвольных ядер и NN-взаимодействий. После получения предсказаний ансамбля следует их анализ и, как следствие, отбор подмножества нейронных сетей, предсказания которых удовлетворяют некоторым критериям. Предложенная топология нейронной сети, позволяют получать устойчивые результаты и обеспечивают высокую сходимость предсказаний по мере включения в обучающую выборку данных из более высоких модельных пространств. Метод апробирован на вариационных расчетах ядра 2H.*

**EXTRAPOLATION OF VARIATIONAL CALCULATIONS with ARTIFICIAL NEURAL NETWORKS**

1. **O. Belozerov1, A. I. Mazur1, R. E. Sharypov1, A. M. Shirokov2,**

*1Pacific National University (Khabarovsk)*

*2Skobeltsyn Institute of Nuclear Physics (Moscow)*

*Using machine learning we develop a tool for extrapolating nuclear observables — the ground-state energy and the root-mean-square radius obtained within the No-core shell model. The idea is to build an ensemble of artificial neural networks to determine the average value of the observable for arbitrary nuclei and NN interactions. After retrieving the predictions of the ensemble, their analysis follows and the subset of ensemble is built, in which all neural networks satisfy certain criteria. A neural network topology is proposed that provides stable results and ensuring convergence of predictions as data from larger model spaces are included in the training set. The proposed approach was tested in variational calculations of the 2H nucleus.*

Модель оболочек без инертного кора (МОБИК) [1] является одним из основных методов описания свойств атомных ядер. Такой подход называется *ab initio* (из первых принципов), поскольку в качестве входной информации используется только реалистическое нуклон-нуклонное взаимодействие.

Результаты расчетов в оболочечных моделях, включая МОБИК, зависят от двух параметров: размера базисного пространства модели, определяемого максимальным числом осцилляторных квантов возбуждения , и величины осцилляторной энергии .

Основная проблема МОБИК – резкое увеличение требуемых вычислительных ресурсов с увеличением , что связано с экспоненциальным ростом числа базисных функций. Современные суперкомпьютеры позволяют проводить расчеты в МОБИК для легких ядер с массовым числом вплоть до , причем если для легчайших ядер *s-* и начала *p-*оболочки возможны расчеты с , то для ядер середины *p-*оболочки доступны расчеты с Для предсказания результатов, соответствующих бесконечному базису, разработаны различные методы экстраполяции (см., например, [2]). Однако все они не имеют строгого обоснования, поэтому вопрос поиска новых методов экстраполяции остается актуальным.

На сегодняшний день методы машинного обучения прочно вошли в арсенал как теоретиков, так и экспериментаторов в области ядерной физики [3]. Они применяются как в области низких энергий, так и при более высоких энергиях. В экспериментальной области машинное обучение применяется в том числе для реконструкции событий, их анализа, фильтрации экспериментальных данных, а также для создания эмуляций.

В данной работе предлагается и исследуется метод экстраполяции результатов вариационных расчетов на случай бесконечно больших модельных пространств, основанный на машинном обучении. Следуя идеям работы [4], мы обучаем ансамбль искусственных нейронных сетей на результатах расчетов в МОБИК и экстраполируем предсказания на случай бесконечно больших модельных пространств, подавая на вход в каждую нейронную сеть достаточно большое значение .

Алгоритм экстраполяции с помощью ансамбля нейронных сетей в данной работе тестируется на модельной вариационной задаче – расчете энергии основного состояния и среднеквадратичного радиуса дейтрона с реалистическим нуклон-нуклонным потенциалом Nijmegen-II, точные значения которых известны: , . Выбор задачи обусловлен тем, что в ней наблюдается медленная сходимость. Кроме того, отчетливо проявляется так называемый четно-нечетный эффект относительно значений (рис. 1 и 2), что позволяет дополнительно тестировать корректность предсказаний и эффективность алгоритма.

В качестве модели, с помощью которой аппроксимируются исходные данные, был выбрана искусственная нейронная сеть – многослойный перцептрон с тремя скрытыми слоями, в каждом из которых содержится 10 нейронов. На первом скрытом и на выходном слое используется линейная активирующая функция , на остальных – сигмоидальная . На входы нейронной сети поступает пара значений , а на выходе получается значение энергии основного состояния или среднеквадратичного радиуса .

Алгоритм экстраполяции заключается в следующем. Проводится обучение ансамбля из 1024 нейронных сетей. Каждая нейронная сеть из ансамбля отличается от другой лишь начальными значениями весовых коэффициентов связей. Обучение производится методом градиентного спуска с помощью оптимизатора Adam [5]. Для обучения используется модуль TensorFlow [6] для Python.

*Рис. 1.* Зависимость от энергии основного состояния ядра 2H в расчетах с NN-потенциалом   
Nijmegen-II при с . Черной пунктирной линией на графике показаны результаты расчетов с . Шаг по равен , отдельные точки не показаны

*Рис. 2.* Зависимость от энергии среднеквадратичного радиуса ядра 2H в расчетах с   
NN-потенциалом Nijmegen-II при . Обозначения как на рис. 1.

В обучении используются не все исходные данные. Берутся данные модельных пространств с **,** где используются различные . Для обучения нейронных сетей, предсказывающих энергию, берутся данные, лежащие правее минимума при каждом вплоть до МэВ, а в задаче предсказания радиуса данные берутся из интервала .

Следующий этап заключается в отборе нейронных сетей, дающих достоверные предсказания. Для этого мы вводим следующие критерии: независимость предсказаний от при ; малость функции ошибки на обучающей выборке. Кроме этого, нейронные сети, предсказывающие энергию, должны удовлетворять вариационному принципу. Далее формируется подмножество ансамбля нейронных сетей, распределение предсказаний которых оказывается близким к нормальному, и мы проводим дополнительный отбор по правилу . Последним этапом является вычисление по сформированному множеству среднеарифметического значения и среднеквадратичного отклонения .

В табл. 1 и 2 представлены результаты экстраполяции энергии основного состояния и среднеквадратичного радиуса ядра 2H. Обучение проводилось с использованием данных вплоть до заданного . В таблицах также указано количество нейронных сетей , прошедших отбор. Также приведены величины энергии, полученные с помощью экстраполяции *B* [2] с погрешностью — и .

**Таблица 1.****Экстраполированная энергия основного состояния**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Все | | | Только при четных | | | ,  МэВ | ,  МэВ |
| , МэВ | , МэВ |  | , МэВ | , МэВ |  |
| 20 | -2,08 | 0,50 | 49 | -2,32 | 0,24 | 130 | -1,73 | 0,194 |
| 22 | -4,98 | 2,54 | 106 | – | – | – | – | – |
| 24 | -2,85 | 0,99 | 211 | -2,16 | 0,15 | 46 | -1,92 | 0,071 |
| 26 | -2,82 | 1,00 | 203 | – | – | – | – | – |
| 28 | -2,18 | 0,13 | 301 | -2,20 | 0,12 | 417 | -2,02 | 0,033 |
| 30 | -2,24 | 0,11 | 170 | – | – | – | – | – |
| 32 | -2,17 | 0,06 | 356 | -2,16 | 0,05 | 601 | -2,07 | 0,017 |
| 34 | -2,22 | 0,08 | 177 | – | – | – | – | – |
| 36 | -2,17 | 0,04 | 394 | -2,18 | 0,04 | 830 | -2,11 | 0,007 |
| 38 | -2,18 | 0,04 | 158 | – | – | – | – | – |
| 40 | -2,17 | 0,03 | 474 | -2,20 | 0,04 | 899 | -2,12 | 0,005 |
| 42 | -2,17 | 0,03 | 324 | – | – | – | – | – |
| 44 | -2,17 | 0,02 | 592 | -2,21 | 0,05 | 885 | -2,14 | 0,003 |
| 46 | -2,17 | 0,01 | 563 | – | – | – | – | – |
| 48 | -2,17 | 0,02 | 750 | -2,20 | 0,04 | 850 | -2,143 | 0,002 |
| 50 | -2,16 | 0,01 | 747 | – | – | – | – | – |
| 52 | -2,17 | 0,01 | 764 | -2,21 | 0,04 | 821 | -2,148 | 0,001 |

**Таблица 2.****Экстраполированный среднеквадратичный радиус**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Все | | | Только при четных | | |
| , фм | , фм |  | , фм | , фм |  |
| 20 | 2,09 | 0,73 | 97 | 2,84 | 0,29 | 45 |
| 22 | 3,88 | 0,68 | 230 | – | – | – |
| 24 | 1,72 | 0,24 | 203 | 1,84 | 0,14 | 50 |
| 26 | 3,16 | 1,15 | 160 | – | – | – |
| 28 | 1,88 | 0,17 | 351 | 1,77 | 0,13 | 184 |
| 30 | 4,57 | 2,04 | 482 | – | – | – |
| 32 | 1,89 | 0,31 | 524 | 1,80 | 0,12 | 381 |
| 34 | 3,35 | 1,74 | 509 | – | – | – |
| 36 | 1,90 | 0,23 | 631 | 1,80 | 0,12 | 381 |
| 38 | 2,15 | 0,35 | 302 | – | – | – |
| 40 | 1,93 | 0,22 | 703 | 1,86 | 0,09 | 513 |
| 42 | 2,05 | 0,23 | 541 | – | – | – |
| 44 | 1,91 | 0,15 | 666 | 1,90 | 0,07 | 536 |
| 46 | 1,99 | 0,08 | 532 | – | – | – |
| 48 | 1,94 | 0,05 | 611 | 1,93 | 0,05 | 539 |
| 50 | 1,99 | 0,04 | 432 | – | – | – |
| 52 | 1,97 | 0,04 | 634 | 1,95 | 0,04 | 441 |

Из таблиц видно, что использование четных для обучения, во-первых, дает достоверные результаты при меньшем значении ; во-вторых, сходимость гораздо более равномерная по сравнению с использованием всех ; в-третьих, после отбора остается гораздо больше нейронных сетей. Следует отметить, что при использовании четных результаты экстраполяции энергии основного состояния при согласуются с точным значением, тогда как при использовании всех согласованности не наблюдается, хотя между собой результаты экстраполяции с помощью нейронных сетей находятся в согласии.

В случае экстраполяции радиуса получаемые значения находятся в хорошем согласии как при использовании всех для обучения, так и при использовании только четных , однако в случае последних неопределенность результата меньше; при результаты экстраполяции согласуется с точным значением.

Таким образом, рассматриваемый метод экстраполяции позволяет получать устойчивые результаты для двух различных наблюдаемых – энергии основного состояния и среднеквадратичного радиуса при использовании одной и той же топологии нейронной сети. Следовательно, этот метод можно применить и к другим наблюдаемым, таким как квадрупольный момент, а также использовать для изучения более тяжелых ядер.

**л и т е р а т у р а**

1. Barrett, B.R., Navrátil, P., Vary, J. P. Ab initio no core shell model // Progress in Particle and Nuclear Physics – 2013. – V. 69(1). – P.131–181.

2. Maris, P., Vary, J.P., Shirokov, A.M. Ab initio no-core full configuration calculations of light nuclei // Phys. Rev. C – 2009. – V. 79. – 014308.

3. Boehnlein A et al. Machine Learning in Nuclear Physics // arXiv:2112.02309v2 [nucl-th] – 2022.

4. Negoita, G.A., Vary, J.P., Luecke, G.R., Maris, P., Shirokov, A.M., Shin, I.J., Kim, Y., Ng, E.G., Yang, C., Lockner, M., Prabhu, G.M. Deep learning: Extrapolation tool for ab initio nuclear theory // Phys. Rev. C – 2019. – V. 99 – 054308.

5. Kingma, D. P., Ba, J. Adam: A Method for Stochastic Optimization. arXiv:1412.6980v9 [cs.LG] – 2017.

6. Abadi, M. et al. TensorFlow: Large-scale machine learning on heterogeneous systems // Software available from tensorflow.org – 2015.