

АВТОМАТИЗИРОВАННОЕ МАШИННОЕ ОБУЧЕНИЕ ДЛЯ НЕЙТРОННОЙ СПЕКТРОМЕТРИИ

К. А. Чижов^{1, 2, *}, *А. А. Белый*^{2, **}, *М. Д. Стариковская*^{2, ***}

¹ Объединенный институт ядерных исследований, Дубна, Россия

² Государственный университет «Дубна», Дубна, Россия

Радиационный мониторинг за защитами ускорителей проводится многошаровым спектрометром Боннера, который детектирует нейтроны в широком диапазоне энергий. Поскольку нейтронные спектры не могут быть измерены напрямую, для восстановления применяются специальные математические методы. В данной работе восстановление проведено с использованием фреймворков машинного обучения LightAutoML и FEDOT, автоматически оптимизирующих гиперпараметры и ансамблевые комбинации нескольких алгоритмов. Для обучения и валидации сгенерирован набор спектров, представленный как взвешенная комбинация тепловой, надтепловой, быстрой и высокоэнергетической компонент. Модели были проверены на 375 реальных спектрах. Неопределенность восстановления оценена методом Монте-Карло путем добавления случайной ошибки в исходные данные. Качество восстановления спектров оценено по корреляции Спирмена и Пирсона, косинусному сходству, перекрестной энтропии, расстоянию Вассерштейна, дивергенции Кульбака–Лейблера, максимальному среднему расхождению и R^2 .

Radiation monitoring outside accelerator shielding is performed using the Bonner multi-sphere spectrometer, which detects neutrons across a wide energy range. Since neutron spectra cannot be directly measured, computational unfolding methods are required. In this study, neutron spectrum deconvolution models were developed using the LightAutoML and FEDOT automated machine learning frameworks, optimizing hyperparameters and ensemble combinations of multiple algorithms. For training and validation a synthetic spectra were generated as weighted combinations of thermal, epithermal, fast and high-energy component. The models were evaluated on a test set of 375 real spectra, with uncertainty assessed via Monte Carlo simulations incorporating random input errors. Reconstruction quality was measured using Spearman and Pearson correlation, cosine similarity, cross-entropy, Wasserstein distance, Kullback–Leibler divergence, maximum mean discrepancy and R^2 .

PACS: 02.30.Zz; 07.05.Mh; 29.30.Hs

* E-mail: kchizhov@jinr.ru

** E-mail: artmbely@yandex.ru

*** E-mail: mariastar3105@gmail.com

ВВЕДЕНИЕ

Нейтроны высоких энергий на объектах ядерной энергетики, ускорителях и в космическом пространстве представляют определенную угрозу из-за потенциального риска для здоровья персонала. Точная оценка спектра нейтронов и эффективной мощности дозы крайне важна для радиационной защиты. Однако прямое измерение спектра нейтронов, особенно с энергиями более 20 МэВ, является сложной экспериментальной задачей, что привело к разработке методов непрямой спектрометрии, среди которых наиболее широко используется многосферовой спектрометр Боннера (МСБ) [1]. МСБ состоит из детектора тепловых нейтронов и набора полиэтиленовых шаров-замедлителей различного диаметра. Измерения проводятся последовательно для каждого шара-замедлителя. Спектр нейтронов $\varphi(E)$ связан с измерениями МСБ системой из M интегральных уравнений Фредгольма 1-го рода [2, 3]:

$$\int_{E_{\min}}^{E_{\max}} K_j(E) \varphi(E) dE = Q_j, \quad j = 1, \dots, M, \quad (1)$$

где Q_j — показание МСБ для j -го шара; $K_j(E)$ — ядро j -го уравнения, представляющее собой функцию чувствительности (ФЧ) детектора к нейтронам различных энергий [1]; M — количество шаров-замедлителей, используемых в измерениях. Пределы интегрирования $E_{\min}$ и $E_{\max}$ задаются энергиями нейтронов (E) и набором используемых замедлителей.

Сравнение методов решения (1) приведено в [4]. Алгоритмы машинного обучения (МО) и нейронные сети (НС) [5], обученные на смоделированных или аугментированных [1] данных, также были применены для решения этой задачи. Для МО измерения МСБ являются входными признаками, а спектры, дискретизированные по сетке вдоль энергетической оси, — искомой целью. В нашей работе мы дополнили обучающий набор данных высокоэнергетическими спектрами для оценки спектров ускорителей и применили автоматизированное МО (AutoML).

МАТЕРИАЛЫ И МЕТОДЫ

Нами была собрана база данных из 375 нейтронных спектров с различных ядерных установок из компендиума МАГАТЭ [1] и открытых источников. Все спектры были нормированы на единицу, и затем были рассчитаны показания МСБ GSF для 10 сфер-замедлителей (0''–18'') [1]. Этот набор данных был использован для тестирования. Ввиду ограниченного количества реальных спектров для обучения был сгенерирован набор из $5 \cdot 10^5$ спектров по методу FRUIT [1], который моделирует спектры как суперпозицию компонент деления, испарения, гауссовой

и высокоэнергетических нейтронов. Генерация синтетических данных включала: создание исходных спектров через параметрические функции FRUIT, их нормализацию и расчет показаний МСБ с использованием функций отклика (1). Набор данных содержит $M = 10$ входных признаков (показания МСБ) и $N = 60$ целевых значений (энергетические бины в логарифмическом масштабе от 10^{-9} до $6,3 \cdot 10^2$ МэВ). Данные были разделены в пропорции 80/20 % для обучения и валидации. Метод Монте-Карло был применен для оценки влияния ошибки: в показания детекторов вносился шум, распределенный по нормальному закону в пределах $\zeta_Q = 1\%$.

МЕТРИКИ СРАВНЕНИЯ

Для оценки сходства истинного и восстановленного спектров необходимо было определить метрики: среднюю абсолютную процентную ошибку (MAPE) мощности эффективной дозы для изотропного облучения $\dot{H}_{\text{ISO}}$, рассчитанную по спектру [5], среднюю квадратичную ошибку (MSE), среднюю абсолютную ошибку (MAE), максимальное среднее расхождение (MMD), косинусное сходство (CS), корреляцию Пирсона (PC), корреляцию Спирмена (SC), расстояние Вассерштейна (WD), дивергенцию Кульбака–Лейблера (KL), перекрестную энтропию (CE) и коэффициент детерминации (R^2).

AutoML-МОДЕЛИ

AutoML обеспечивает автоматическую тонкую настройку гиперпараметров модели и смешивание различных моделей. В этой работе мы использовали фреймворки Python: LightAutoML [6] и FEDOT [7]. LightAutoML в режиме многомерной регрессии подбирает оптимальную модель среди линейной регрессии с L2-регуляризацией, LightGBM, CatBoost и случайного леса. FEDOT основан на эволюционном подходе: этот фреймворк итерирует, смешивает и выбирает лучший метод среди регрессоров [7]. Оба AutoML определили случайный лес в качестве оптимального регрессора после 24 ч оптимизации пайплайна.

РЕЗУЛЬТАТЫ

Были рассчитаны кривые обучения для моделей. Кривая обучения для R^2 выходит на плато для обучающих данных на уровне 0,75, а для валидационных — на 0,80. В таблице приведены метрики по всем тестовым спектрам. FEDOT показывает лучшие результаты по сравнению с LightAutoML благодаря более обширному пространству поиска среди моделей и эволюционному подходу к оптимизации. Пример восстановленных спектров приведен на рис. 1. LightAutoML включает метод оценки важности признаков. Для МСБ GSF шар-замедлитель 6'' является наиболее важным (17%). Напротив, шар 8'' вносит только 3%-й вклад

Сравнение метрик для тестовых данных

Название модели	R^2	MSE ($\cdot 10^{-4}$)	MAE ($\cdot 10^{-3}$)	CS	SC	PC
LightAutoML	0,74	3,95	6,76	0,916	0,74	0,866
FEDOT	0,75	3,55	6,34	0,920	0,77	0,873

Название модели	MMD	WD ($\cdot 10^{-3}$)	Энтропия	CE	KL	MAPE, %
LightAutoML	0,09	4,5	2,8	3,0	0,19	14,0
FEDOT	0,06	4,1	2,8	3,2	0,18	9,2

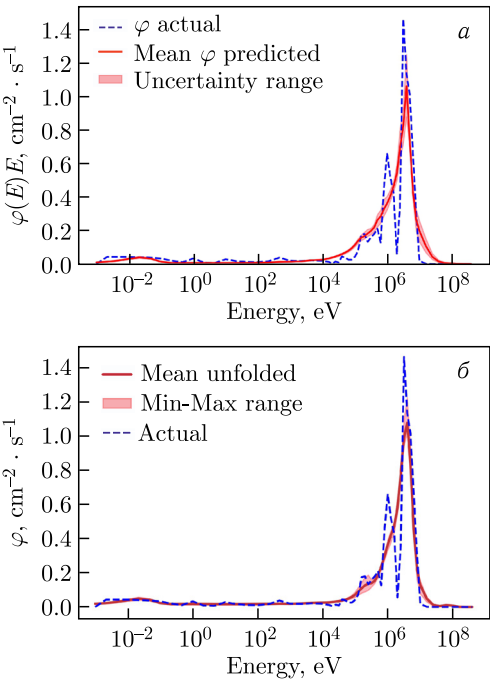


Рис. 1. Примеры восстановленных спектров для фреймворков LightAutoML (а) и FEDOT (б)

в конечный результат. Наряду с другими методами оптимизации [8] оценка важности признаков может быть использована для сокращения количества измерений и минимизации доз персонала.

ОБСУЖДЕНИЕ

Результаты демонстрируют, что фреймворки AutoML могут эффективно решать задачу восстановления спектра нейтронов, что позволяет их применять наряду с другими зарекомендовавшими себя методами [4].

AutoML обеспечивает автоматический выбор наилучшего алгоритма и его гиперпараметров по заданной метрике, а также предоставляет интерпретируемость модели через ранжирование признаков по их значимости. В будущих проектах планируется разработать гибридные методы, сочетающие детерминированные методы и МО. Метод применим для стационарных нейтронных полей и требует обучения отдельной модели для каждого МСБ.

ЗАКЛЮЧЕНИЕ

В исследовании реализованы фреймворки LightAutoML и FEDOT для восстановления нейтронного спектра по данным МСБ. Примененные к 10 измерениям, дискретизированным в 60 энергетических бинах, оба фреймворка показали хорошие результаты в восстановлении спектра, причем FEDOT превзошел LightAutoML по большинству метрик, за исключением перекрестной энтропии. $\dot{H}_{\text{ISO}}$, рассчитанная по восстановленным спектрам, находится в пределах 14 % от фактических значений. В работе предложен подход к выбору оптимального набора шаров-замедлителей на основе важности признаков модели.

Благодарности. Авторы выражают благодарность Многофункциональному информационно-вычислительному комплексу ОИЯИ за предоставленные вычислительные ресурсы.

Финансирование. Исследование выполнено в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № 124112200072-2).

СПИСОК ЛИТЕРАТУРЫ

1. *McGreivy J., Manfredi J.J., Siefman D.* Data Augmentation for Neutron Spectrum Unfolding with Neural Networks // JNE. 2023. V. 4. P. 77–95; doi: 10.3390/jne4010006.
2. *Chizhov K., Beskrovnaya L., Chizhov A.* Neutron Spectra Unfolding from Bonner Spectrometer Readings by the Regularization Method Using the Legendre Polynomials // Phys. Part. Nucl. 2024. V. 55. P. 532–534; doi: 10.1134/S1063779624030298.
3. *Chizhov K., Beskrovnaya L., Chizhov A.* Neutron Spectrum Unfolding Method Based on Shifted Legendre Polynomials, Its Application to the IREN Facility // Phys. Part. Nucl. Lett. 2025. V. 22, No. 2. P. 337–340; doi: 10.1134/S154747712470239X.
4. *Gómez-Ros J.M. et al.* Results of the EURADOS International Comparison Exercise on Neutron Spectra Unfolding in Bonner Spheres Spectrometry // Radiat. Meas. 2022. V. 153. P. 106755; doi: 10.1016/j.radmeas.2022.106755.
5. *Chizhov K. et al.* Reconstruction of the Energy Spectrum of the Neutron Radiation Flux Using the Random Forest Machine Learning Algorithm // Mod. Inf. Technol. IT-Educ. 2024. V. 20, No. 4.

6. *Vakhrushev A. et al.* LightAutoML: AutoML Solution for a Large Financial Services Ecosystem. arXiv:2109.01528. 2021.
7. *Nikitin N. O. et al.* Automated Evolutionary Approach for the Design of Composite Machine Learning Pipelines // *Future Gener. Comput. Syst.* 2022. V. 127. P. 109–125; doi: 10.1016/j.future.2021.08.022.
8. *Chizhov K., Chizhov A.* Optimization Approach to Neutron Spectra Unfolding with Bonner Multi-Sphere Spectrometer // *Math. Model.* 2024. V. 7. P. 89–90.