INTERPRETABLE MACHINE LEARNING FOR PROTON-INDUCED NEUTRON REACTION CROSS-SECTIONS PREDICTION

YanBang Tang

University of Chinese Academy of Sciences, Beijing, China
and
European Organization for Nuclear Research (CERN), Geneva, Switzerland
tangyanbang2022@163.com, yanbang.tang@cern.ch

Received 6 May 2025, accepted 18 November 2025, published online 8 December 2025

The accurate prediction of proton-induced neutron (p, n) reaction cross sections is critical for applications in nuclear engineering, medical isotope production, and astrophysics. This study introduces a machine learning framework that combines high-predictive accuracy with interpretability and uncertainty quantification. We integrate the Hiking Optimization Algorithm (HOA) with the eXtreme Gradient Boosting (XGBoost) model and employ Monte Carlo (MC) Dropout for uncertainty estimation. The framework is interpreted using SHapley Additive exPlanations (SHAP). First, HOA is utilized to navigate the high-dimensional hyperparameter space, optimizing the XGBoost model for performance. Subsequently, the trained model's predictions are benchmarked against experimental data from the EXFOR database and theoretical calculations from TALYS 2.0. Our HOA–XGBoost model demonstrates better predictive accuracy over other machine learning models and provides predictions closer to experimental values than TALYS 2.0. The inclusion of MC Dropout provides uncertainty bounds for the model's predictions. A detailed SHAP analysis reveals the underlying physical drivers of the model's decisions: the incident proton energy (EN) is identified as the most influential feature, with its strong interaction with the reaction Q-value and product proton number (Z_2) highlighting the model's ability to learn fundamental concepts such as reaction thresholds and the Coulomb barrier. The product nuclide's neutron (N_2) and proton (Z_2) numbers also show influence related to nuclear stability, while the product mass number (A_2) has a lesser impact. This work presents a complementary methodology for nuclear data evaluation, paving the way for more reliable predictions and targeted experimental design.

DOI:10.5506/APhysPolB.56.12-A5

1. Introduction

The precise determination of nuclear reaction cross sections is fundamental for advancements in nuclear engineering, astrophysics, and medical physics. Among these, proton-induced neutron emission ((p, n) reactions) are of particular importance for their role in neutron source design, the production of radioisotopes for medicine, and understanding stellar nucleosynthesis [1–5]. Consequently, the development of a generalizable, high-accuracy, and physically interpretable model for predicting (p, n) reaction cross sections is essential for augmenting nuclear reaction databases and enhancing the reliability of their applications.

Traditional methods for cross-section evaluation have centered on experimental measurements and theoretical models. Theoretical approaches, such as those based on the Feshbach–Kerman–Koonin or optical models, have successfully described reaction mechanisms within specific energy and mass ranges [6–8]. Concurrently, semi-empirical systematics fit parameterized equations to extensive experimental data, leveraging physical principles to improve predictive accuracy [9, 10]. However, both approaches face limitations, particularly when dealing with unstable target nuclei or energy regions where experimental data are sparse, thereby restricting their predictive power and generalizability.

In recent years, machine learning (ML) has emerged as a powerful data-driven paradigm in nuclear physics, capable of capturing complex, nonlinear relationships in high-dimensional feature spaces that often elude conventional models [11–17]. Despite their predictive success, many ML models operate as "black boxes", a significant drawback in a field where understanding the underlying physical mechanisms is paramount. For a model's predictions to be truly valuable, its decision-making process must be transparent and align with the established physical principles.

To address these challenges, this study proposes an interpretable machine learning framework for predicting (p, n) reaction cross sections using data from the EXFOR database [18]. Our framework synergizes three key components: the eXtreme Gradient Boosting (XGBoost) model for its state-of-the-art performance, the Hiking Optimization Algorithm (HOA) [19] for automated hyperparameter tuning, and SHapley Additive exPlanations (SHAP) [20] for deep model interpretability. Furthermore, we address the critical issue of predictive uncertainty by incorporating the Monte Carlo (MC) Dropout technique, providing confidence intervals for our predictions. By comparing our model with the widely-used TALYS 2.0 nuclear reaction code [21], we not only demonstrate enhanced predictive accuracy but also use SHAP to deconstruct the model's logic, quantifying the influence of physical inputs such as incident energy, nuclide properties, and reaction Q-value. This work aims at bridging the gap between data-driven accuracy and physical insight, offering an efficient and reliable pathway for nuclear data evaluation.

2. Materials and methods

The workflow of this study is structured into four stages: data preparation, model development and optimization, uncertainty quantification, and interpretability analysis, as depicted in Fig. 1.

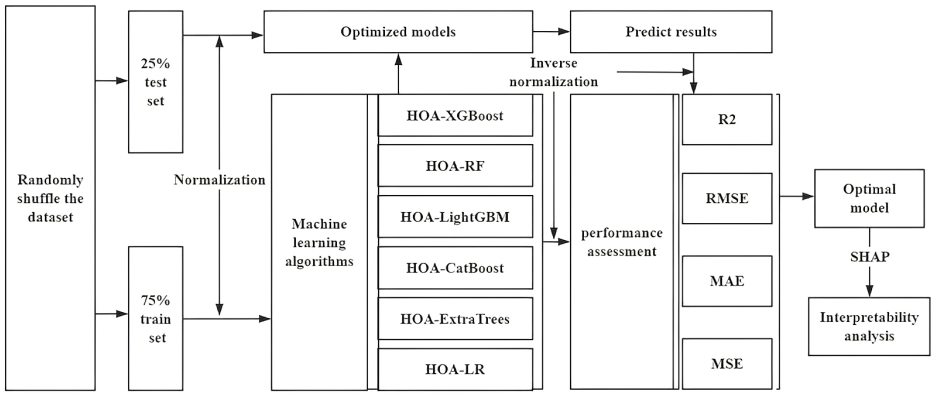

Fig. 1. Overall workflow of HOA–XGBoost–SHAP for the (p, n) nuclear reaction cross-section prediction.

2.1. Data and features

A dataset of 3416 experimental (p, n) reaction cross-section measurements was compiled from the EXFOR database. The dataset covers incident proton energies from 2 MeV to 88 MeV and a wide range of target nuclei (proton number $Z = 4$ to 92; mass number $A = 9$ to 235). Eight physically motivated features were selected as inputs for the models:

— EN: Incident proton energy [MeV].

— A_1, N_1, Z_1: Mass number, neutron number, and proton number of the target nucleus.

— A_2, N_2, Z_2: Mass number, neutron number, and proton number of the product nuclide.

— Q: The reaction Q-value [MeV].

The dataset was randomly shuffled and split into a training set (75%) and a testing set (25%). All features were standardized to have zero mean and unit variance to ensure stable model training.

2.2. Machine learning models

Six established machine learning models were selected to benchmark performance on the (p, n) cross-section prediction task.

XGBoost (eXtreme Gradient Boosting): An optimized gradient boosting library known for its high performance, scalability, and built-in regularization to prevent overfitting [22].

Random Forest (RF): An ensemble method that constructs a multitude of decision trees and outputs the average prediction, reducing variance and overfitting [23].

LightGBM: A fast, high-performance gradient boosting framework that uses novel techniques like Gradient-based One-Side Sampling (GOSS) [24].

CatBoost: A gradient boosting algorithm optimized for handling categorical features, which also incorporates mechanisms to reduce gradient bias [25].

ExtraTrees (Extremely Randomized Trees): An ensemble method similar to Random Forest but introducing more randomness in the selection of feature thresholds, often enhancing robustness [26].

Linear Regression (LR): A fundamental statistical model used as a baseline to assess the performance gain from more complex, nonlinear models [27].

2.3. Hyperparameter optimization with HOA

The performance of ML models is highly dependent on their hyperparameters. We employed the Hiking Optimization Algorithm (HOA), a novel metaheuristic inspired by the exploration strategies of mountain hikers [19]. HOA efficiently balances global exploration and local exploitation of the hyperparameter space. It simulates a group of "hikers" searching for an optimal solution by iteratively performing initialization, low-altitude exploration, climbing ascent, and segmented evaluation, before terminating when a stopping criterion is met. This automated process, illustrated for XGBoost in Fig. 2, significantly enhances model performance compared to manual tuning or exhaustive grid searches.

2.4. Uncertainty quantification with Monte Carlo dropout

A critical limitation of many deterministic ML models is the lack of predictive uncertainty. To address this, we integrated Monte Carlo (MC) Dropout into our framework. Dropout is a regularization technique typically used during training, where neurons are randomly "dropped" to prevent co-adaptation. By keeping Dropout active during the inference phase, the model becomes stochastic. For a single input, we perform multiple forward

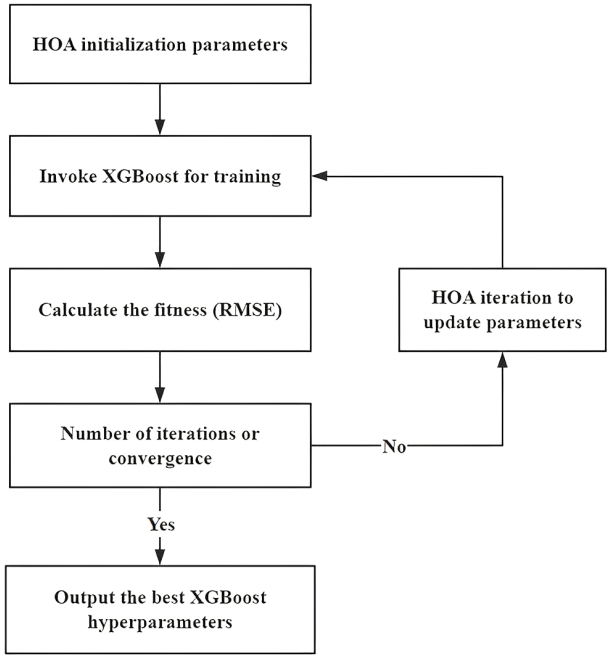

Fig. 2. Flowchart of the HOA-based XGBoost Hyperparameter Optimization.

passes, each with a different randomly dropped set of neurons. This generates a distribution of predictions, from which we can compute a mean value (the final prediction) and a standard deviation, which serves as a measure of the model's uncertainty. This approach provides a robust method for quantifying confidence in each prediction.

2.5. SHAP-based interpretability analysis

To move beyond a "black-box" model, we employed SHAP (SHapley Additive exPlanations) to interpret the predictions of our optimized HOA–XGBoost model [20]. Grounded in cooperative game theory, SHAP assigns each feature an importance value — a Shapley value — for each individual prediction. This value represents the marginal contribution of that feature to pushing the model's output away from the baseline (average) prediction. SHAP offers consistent and locally accurate feature attributions, enabling both global analysis (which features are most important overall) and local analysis (why a specific prediction was made), thereby connecting the data-driven model back to physical intuition.

3. Results and discussion

3.1. Evaluation of HOA-optimized machine learning models

To select the best-performing model architecture, the six ML models, optimized with HOA, were evaluated on the held-out test set. We used four standard regression metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2).

As shown in Table 1, the HOA–XGBoost model achieved the best performance across all metrics, with the highest R^2 of 0.88 and the lowest MSE, RMSE, and MAE. This indicates its better ability to capture the complex nonlinear dependencies in the data. The poor performance of HOA–LR ($R^2 = 0.29$) confirms that a simple linear model is insufficient for this task. Based on these results, HOA–XGBoost was selected for further analysis and comparison with TALYS 2.0.

Table 1. Performance comparison of six HOA-optimized machine learning models in the (p, n) nuclear reaction cross-section prediction.

ML models	R^2	MSE	RMSE	MAE
HOA–XGBoost	0.88	3360.64	57.97	33.31
HOA–RF	0.80	5679.58	75.36	40.52
HOA–LightGBM	0.86	3852.24	62.06	37.17
HOA–CatBoost	0.79	5960.41	77.20	39.83
HOA–ExtraTrees	0.82	5023.79	70.87	35.36
HOA–LR	0.29	20468.89	143.06	99.02

3.2. Comparison with TALYS 2.0 and uncertainty analysis

We compared the predictions of our HOA–XGBoost model with theoretical calculations from TALYS 2.0 and the experimental data from EXFOR. Crucially, we include the predictive uncertainty of our model, estimated via MC Dropout. A representative subset of these comparisons is shown in Table 2.

The analysis of Table 2 reveals several key points. In a majority of cases, the mean prediction from HOA–XGBoost is closer to the experimental value than the TALYS 2.0 calculation. Furthermore, the uncertainty bounds provided by MC Dropout often encompass the experimental value, lending credibility to the model's predictions. However, TALYS demonstrates higher accuracy in some instances. Such discrepancies may arise where the training data for our ML model is sparse, while TALYS benefits from its underlying

physics models which are well-calibrated in that specific region. This highlights that both approaches have their strengths, and a hybrid methodology could offer a path toward more robust nuclear data evaluations.

Table 2. Comparison of predicted cross sections (in mb) from HOA–XGBoost and TALYS 2.0 with experimental EXFOR data.

Energy [MeV]	Target Z	Target A	EXFOR	HOA–XGBoost	TALYS
10.00	58	142	91.00 ± 10	58.83 ± 12.1	83.48
13.78	28	64	528.60 ± 113	462.98 ± 35.5	257.62
10.30	68	168	249.00 ± 30	237.99 ± 21.8	152.26
10.00	45	103	703.00 ± 73	421.97 ± 41.2	510.31
19.80	51	121	80.00 ± 13	75.28 ± 9.8	65.30
8.60	26	57	633.00 ± 114	486.41 ± 45.1	546.49
6.00	23	51	511.00 ± 58	500.27 ± 33.6	512.36
52.00	51	121	6.90 ± 1.5	11.32 ± 3.1	19.48
35.40	64	160	13.80 ± 0.8	12.64 ± 4.5	26.82
13.10	34	80	253.00 ± 38	250.99 ± 20.7	337.46
16.08	92	235	5.80 ± 0.3	19.89 ± 6.2	8.85
7.06	12	26	47.40 ± 7.1	17.41 ± 8.9	187.82
26.10	78	198	19.90 ± 3.6	22.57 ± 5.6	41.54
16.43	74	186	23.42 ± 1.06	40.12 ± 7.1	44.02
6.24	39	89	102.00 ± 10.1	98.64 ± 14.2	198.43
8.64	18	38	23.40 ± 3	19.04 ± 6.7	103.66
15.00	5	11	145 ± 20	172.03 ± 18.3	161.09
6.70	20	43	111 ± 10	156.45 ± 16.2	369.13

3.3. SHAP-based physical interpretation of the model

To ensure our model's predictions are not just accurate but also physically meaningful, we performed a SHAP analysis.

Figure 3 shows a SHAP waterfall plot for a single prediction. Taking the left subfigure as an example, we observe each feature's influence on this sample's (p, n) cross-section prediction. First, the product proton number $Z_2 = 29$ contributes a $+149.35$ upward shift; next, the product neutron number $N_2 = 35$ adds a further $+104.02$. By summing these positive and negative contributions around the baseline $E[f(x)] = 132.183$, we arrive at the model's final predicted (p, n) cross section of 462.978.

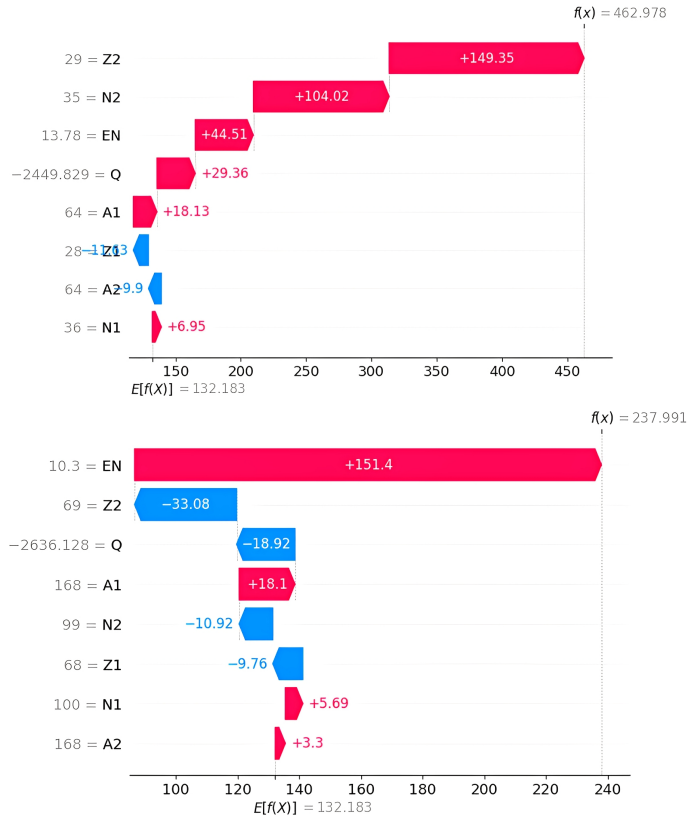

Fig. 3. The SHAP contribution waterfall plot for the (p, n) reaction cross-section prediction. The plot decomposes a single prediction, starting from the baseline average prediction $E[f(x)]$ and adding or subtracting the contribution from each feature to arrive at the final output $f(x)$.

Figure 4 provides a global summary of feature impacts across the entire test set. Several physically intuitive patterns emerge:

EN (Incident Energy): It is clearly the most important feature. High EN values (red points) have negative SHAP values, indicating they tend to decrease the (p, n) cross section. This is physically sound, as higher energies open up competing reaction channels (*e.g.*, $(p, 2n)$, (p, pn)), thus reducing the probability of the exclusive (p, n) channel.

N_2 and Z_2 (Product Neutron/Proton Number): These are the next most important features, indicating that the stability and structure of the product nucleus are critical.

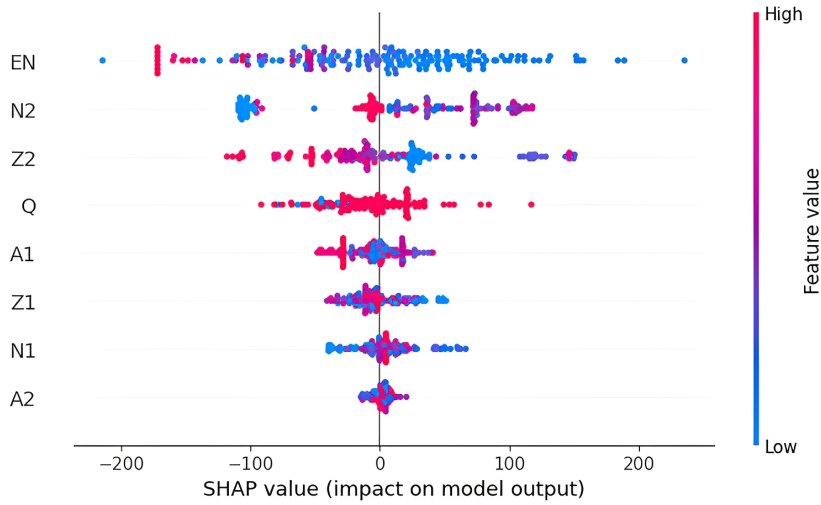

Fig. 4. The global SHAP feature importance summary plot for the (p, n) reaction cross-section prediction. Each point is a Shapley value for a feature and an instance. The color indicates the feature's value (red for high, blue for low).

Figure 5 confirms the ranking observed in Fig. 4, quantifying the average impact of each feature. EN is dominant, followed by N_2, Z_2, and Q. Interestingly, the product mass number A_2 has the least impact. This is a sensible finding, as the model has learned that the specific composition (N_2 and Z_2), which relates to shell structure and the N/Z ratio, is more predictive than the total nucleon count.

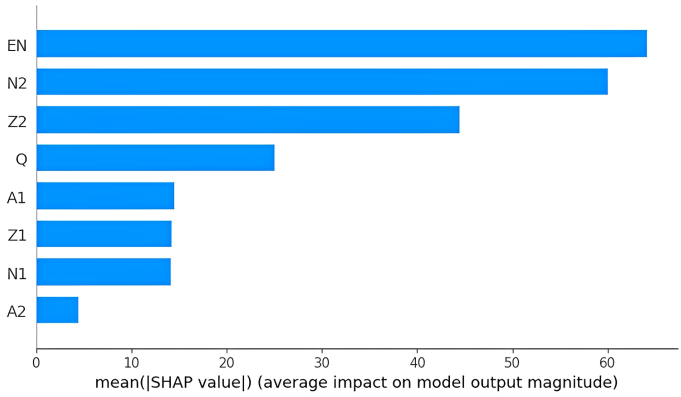

Fig. 5. Mean absolute SHAP values ranking feature importance.

To probe deeper, we examined feature interactions (Fig. 6). The most significant interactions are observed with EN:

Interaction between EN and Q: There is a strong interaction, which is physically expected. The Q-value determines the reaction's energy threshold. The cross-section's behavior is dependent on how the incident energy EN surpasses this threshold. The model has clearly learned this crucial relationship.

Interaction between EN and Z_1/Z_2: This interaction likely captures the effect of the Coulomb barrier. For a target nucleus with a higher Z_1 (and thus a product with higher Z_2), a greater incident energy EN is required to overcome the electrostatic repulsion. The model has identified this coupled effect.

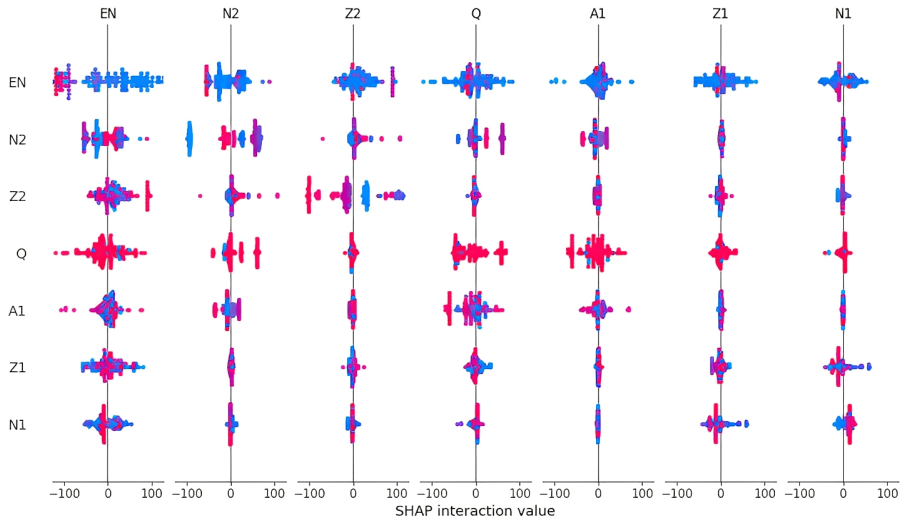

Fig. 6. The SHAP interaction matrix for key features. Diagonal plots show the marginal effect of each feature, while off-diagonal plots show the interaction effects between pairs of features.

In summary, the SHAP analysis provides evidence that the HOA–XGBoost model has not simply memorized data but has learned some representations of key physical principles governing (p, n) reactions.

4. Conclusion and outlook

This study introduced an interpretable machine learning framework for the prediction of (p, n) reaction cross sections. By combining the HOA-optimized XGBoost model with MC Dropout for uncertainty quantification and SHAP for interpretability, we have developed a tool that is both accurate and transparent.

Our key findings are:

— The HOA–XGBoost model demonstrates better predictive accuracy compared to other ML models and, in many cases, provides predictions closer to experimental data than the theoretical code TALYS 2.0.

— The integration of MC Dropout provides a vital measure of predictive uncertainty, a critical feature for any model intended for use in nuclear data evaluation.

— SHAP analysis confirms the physical soundness of our model. It identifies incident energy as the primary driver of cross-section values and captures physically meaningful interactions, such as the interplay between incident energy, the Q-value threshold, and the Coulomb barrier.

Looking forward, this work opens several avenues for future research. A crucial validation step would be to perform a closure test: training the ML model on a large, clean dataset generated by TALYS to verify its ability to reproduce a known physical model. This would further build confidence in the ML architecture itself. Moreover, the framework presented here can be used to augment theoretical models. By identifying regions where the ML model's predictions are confident (low uncertainty) but deviate significantly from TALYS, we can provide targeted guidance for refining the underlying physics models within TALYS. In conclusion, the HOA–XGBoost–SHAP framework, enhanced with uncertainty quantification, represents a step towards building reliable and physically intuitive machine learning tools for the nuclear physics community.

REFERENCES

[1] J. Jarošík *et al.*, «Comparison of fast-neutron-induced reaction cross-sections in ^{27}Al, ^{197}Au, ^{209}Bi, ^{59}Co, ^{19}F, ^{23}Na, and ^{89}Y in quasi-monoenergetic fields of p + Li and d + TiT neutron sources», *Nucl. Instrum. Methods Phys. Res. B* **559**, 165566 (2025).

[2] C.H.M. Broeders, A.Y. Konobeyev, «Systematics of (p, n) reaction cross-section», *Radiochim. Acta* **96**, 387 (2008).

[3] J. Wing, J.R. Huizenga, «(p,n) Cross Sections of V^{51}, Cr^{52}, Cu^{63}, Cu^{65}, Ag^{107}, Ag^{109}, Cd^{111}, Cd^{114}, and La^{139} from 5 to 10.5 MeV», *Phys. Rev.* **128**, 280 (1962).

[4] S. Kailas *et al.*, «Total (p,n) reaction cross-section measurements on ^{50}Ti, ^{54}Cr, and ^{59}Co», *Phys. Rev. C* **12**, 1789 (1975).

[5] P. Reimer, S.M. Qaim, «Excitation Functions of Proton Induced Reactions on Highly Enriched ^{58}Ni with Special Relevance to the Production of ^{55}Co and ^{57}Co», *Radiochim. Acta* **80**, 113 (1998).

[6] Y. Watanabe *et al.*, «Feshbach–Kerman–Koonin model analysis of preequilibrium (p,p') and (p,n) reactions at 12 to 26 MeV», *Phys. Rev. C* **51**, 1891 (1995).

[7] C.H. Johnson, A. Galonsky, R.L. Kernell, «(p,n) reaction for $89 < A < 130$ and an anomalous optical model potential for sub-Coulomb protons», *Phys. Rev. C* **20**, 2052 (1979).

[8] R. Bonetti, M. Camnasio, L. Colli Milazzo, P.E. Hodgson, «Analysis of precompound processes in (p,n) reactions with the statistical multistep direct emission theory», *Phys. Rev. C* **24**, 71 (1981).

[9] O. Lyes, N. Amrani, «New semi-empirical systematic of (p,n) reaction cross section at 7.5 MeV», *Kerntechnik* **88**, 279 (2023).

[10] C. Broeders, A.Yu. Konobeyev, «Systematics of (p,α), $(p,n\alpha)$, and (p,np) reaction cross-sections», *Appl. Radiat. Isot.* **65**, 1249 (2007).

[11] R. Teelock-Gaya, V. Raffuzzi, E. Shwageraus, L. Morgan, «Prediction of Exotic $(n,2n)$ Cross Sections Using a Regression Tree Machine Learning Algorithm», *Nucl. Sci. Eng.* **1**, PHYSOR 2024 (2025).

[12] F. Akbari *et al.*, «Predicting electronic stopping powers using stacking ensemble machine learning method», *Nucl. Instrum. Methods Phys. Res. B* **538**, 8 (2023).

[13] A. Al Hammal *et al.*, «Neural network predictions of inclusive electron–nucleus cross sections», *Phys. Rev. C* **107**, 065501 (2023).

[14] B.E. Kowal *et al.*, «Empirical fits to inclusive electron–carbon scattering data obtained by deep-learning methods», *Phys. Rev. C* **110**, 025501 (2024).

[15] K. Xing *et al.*, «Phase shift deep neural network approach for studying resonance cross sections for the ^{235}U (n,f) reaction», *Phys. Lett. B* **855**, 138825 (2024).

[16] H. Özdoğan, Y. Ali Üncü, M. Şekerci, A. Kaplan, «Estimations for (n,a) reaction cross sections at around 14.5 MeV using Levenberg–Marquardt algorithm-based artificial neural network», *Appl. Radiat. Isot.* **192**, 110609 (2023).

[17] J. Grigsby *et al.*, «Deep learning analysis of deeply virtual exclusive photoproduction», *Phys. Rev. D* **104**, 016001 (2021).

[18] V.V. Zerkin, B. Pritychenko, «The experimental nuclear reaction data (EXFOR): Extended computer database and Web retrieval system», *Nucl. Instrum. Methods Phys. Res. A* **888**, 31 (2018).

[19] S.O. Oladejo, S.O. Ekwe, S. Mirjalili, «The Hiking Optimization Algorithm: A novel human-based metaheuristic approach», *Knowl.-Based Syst.* **296**, 111880 (2024).

[20] S.M. Lundberg, Su-In Lee, «A Unified Approach To Interpreting Model Predictions», in: U. von Luxburg *et al.* (Eds.) «NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing System», Long Beach, CA, USA, 4–9 December, 2017, pp. 4768–4777.

[21] A. Koning, S. Hilaire, S. Goriely, «TALYS: modeling of nuclear reactions», *Eur. Phys. J. A* **59**, 131 (2023); *Erratum ibid.* **59**, 146 (2023).

[22] T. Chen, C. Guestrin, «XGBoost: A Scalable Tree Boosting System», in: «KDD'16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining», San Francisco, California, USA, 13–17 August, 2016, pp. 785–794, arXiv:1603.02754 [cs.LG].

[23] L. Breiman, «Random forests», *Mach. Learn.* **45**, 5 (2001).

[24] G. Ke *et al.*, «LightGBM: A Highly Efficient Gradient Boosting Decision Tree», in: U. von Luxburg *et al.* (Eds.) «NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing System», Long Beach, CA, USA, 4–9 December, 2017, pp. 3149–3157.

[25] L. Prokhorenkova *et al.*, «CatBoost: unbiased boosting with categorical features», in: S. Bengio *et al.* (Eds.) «32nd Conference on Neural Information Processing Systems (NeurIPS 2018)», Montreal, Canada, 3–8 December, 2018, pp. 6639–6649. arXiv:1706.09516 [cs.LG].

[26] P. Geurts, D. Erns, L. Wehenkel, «Extremely randomized trees», *Mach. Learn.* **63**, 3 (2006).

[27] M. Huang, «Theory and Implementation of linear regression», in: «Proceedings of the 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)», Chongqing, China, 10–12 July, 2020, pp. 210–217.