Bayesian-optimized CatBoost for ground-state nuclear charge-radius prediction

Mudassar Ahmed[1]   Abdul Kabir[1†] (ID)   Jameel-Un Nabi[2]   Laiba Hamid[1]   Manzoor Ahmad[3]

[1]Department of Space Science, Institute of Space Technology, Islamabad 44000, Pakistan
[2]University of Wah, Quaid Avenue, Wah Cantt 47040, Punjab, Pakistan
[3]Department of Physics, Faculty of Sciences, International Islamic University, Islamabad 44000, Pakistan

Abstract: Understanding nuclear shape, behavior, and stability, as well as improving nuclear models, depends on the precise determination of ground-state nuclear charge radii. Existing experimental techniques are limited to extremely narrow regions of the nuclear chart; theoretical models, including relativistic Hartree-Bogoliubov (RHB) and Hartree-Fock-Bogoliubov (HFB), predict broad trends of nuclear properties but miss fine isotopic features such as odd-even staggering effects and shell-closure kinks. High computational time and cost are other obstacles to theoretical approaches. Although machine-learning algorithms have made significant progress in predicting charge radii, they are still hindered by a lack of balanced data and characteristics, primarily centered around $A \geq 40$ and $Z \geq 20$. In the present study, we present the first application of CatBoost regression to compute nuclear charge radii. We integrated two experimental datasets with RHB-calculated point-coupling interaction (PC-X) theoretical features and extended our study range to $A \geq 17$, $Z \geq 8$. We found the best hyperparameters using Optuna's Tree-structured Parzen Estimator (TPE) sampler with 10-fold cross-validation (CV), achieving a CV root-mean-square error (RMSE) of 0.0106 fm and hold-out RMSE of 0.0102 fm, with only three features, i.e., neutron number (N), proton number (Z), and RHB theoretical binding energy (BE), outperforming nine other ML models: random forest (RF), quantile RF (QRF), Cubist, Gaussian process regression with polynomial kernel (GPPK), multivariate adaptive regression splines (MARS), SVR, ANN, convolutional neural network (CNN), and Brussels-Skyrme-on-a-grid 3 (BSkG3). SHapley Additive exPlanations (SHAP) analysis confirms the highest global influence of BE in the model's predictions, followed by proton number and neutron number. The proposed model can accurately reproduce the $N = 50$ kink and odd-even staggering effects in krypton and strontium chains. These results establish CatBoost as a robust and notably promising model for charge-radius prediction and beyond, with the potential to impact r-process modeling and future theoretical development.

Keywords: CatBoost, charge radius, machine learning, XGBoost

DOI: 10.1088/1674-1137/ae19dc   **CSTR:** 32044.14.ChinesePhysicsC.50024109

I. INTRODUCTION

Nuclear charge radius (R_c) is one of the most fundamental features of atomic nuclei and has been extensively investigated in nuclear physics [1]. It provides insights into the charge distribution within the nucleus, overall nuclear shape, and nuclear stability and behavior [2]. Ford *et al.* [3] emphasized the importance of nuclear charge radius in understanding nuclear shape and behavior. There are a number of experimental techniques to calculate charge radii, some of which have been in use since the mid-20th century. For example, one of the most precise experimental techniques is based on electron scattering experiments for the determination of charge distribution of nuclei, or more specifically, electron scattering ex-

periments of ^{6}Li, ^{7}Li, ^{12}C, and ^{16}O [4, 5]. Muonic atom spectroscopy and Kα X-ray isotope shift (KIS) are also extensively used [6]. However, limitations in experimental techniques restrict their application to specific regions of the nuclear chart.

For this reason, a wide variety of nuclear models have been developed. They can analyze ground state properties, including R_c, with much greater accuracy. Macroscopic models, including the liquid drop model [7, 8] and Garvey-Kelson relation [9] are the simplest approaches. They provide preliminary insights into nuclear structure and its properties. For better predictions, phenomena such as isospin dependence and shell effects have also been incorporated [10]. Microscopic models further enhance theoretical accuracy by delving into the nucleus at the nucle-

on level, focusing on individual interactions between protons and neutrons. The Skyrme-Hartree-Fock-Bogoliubov (HFB) models use the Skyrme interaction to model nucleon-nucleon forces, approximate the average potential experienced by each nucleon by using the Hartree-Fock approximation, and handle pairing correlations by the Bogoliubov transformation for accurate calculations [11, 12]. The relativistic effects on Skyrme-Hartree-Fock are studied in the RMF model [13]. Apart from these approaches, empirical formulae have also been derived by capturing the local relations, for example, δR, *i.e.*, the charge radii difference of two isotopes of a single element [14]. Advanced mean field models, *e.g.*, deformed relativistic Hartree-Bogoliubov, offer a further deep description by adding properties such as deformation and superfluidity [15−17]. Furthermore, ab initio no-core shell models consider direct interaction of nucleons without core approximations [18, 19]. Lastly, the Weizsäcker-Skyrme (WS*) model integrates both microscopic and macroscopic parameters for a better description of nuclear properties of exotic as well as stable nuclei [10].

Machine learning (ML) and deep learning (DL) offer high potential in their application to both theoretical and experimental nuclear physics [20]. For example, predictions of nuclear mass using support vector and Gaussian process regressors exhibit good performance both within and beyond the training set [21]. ML algorithms applied over AME2020 data [22, 23], outperformed the semi-empirical Bethe-Weisäker formula [24]. Similarly, the impact parameter in heavy-ion collisions can be determined using neural networks (NNs) [25]. Predictive modeling of ground state energies of even-even nuclei can be implemented using ANNs over a theoretical dataset calculated from the Hartree-Fock-Bogoliubov (HFB) model [26]. The study of unstable nuclei relies heavily on the α decay. An approach integrating the sophisticated α-decay model with a Bayesian neural network (BNN) is employed to enhance the prediction accuracy of α-decay half-lives. According to global and extrapolated analyses, model-based predictions of α decay can be more effectively described by the BNN approach [27]. The extreme gradient Boosting (XGBoost) framework is used to study the α-decay energy and half-life of superheavy nuclei, which are optimized using Bayesian hyperparameters. By incorporating key nuclear structural features, including mass number, proton-to-neutron ratio, magic number proximity, and angular momentum transfer, the optimized model captures essential physical mechanisms governing the α decay [28]. On the basis of WKB barrier penetrability, an improved formula considering the deformation effect for the α-decay half-lives was proposed. The improved empirical formula and XGBoost models were used to predict the α-decay half-lives of nuclei with $Z = 117, 118, 119$, and 120. You *et al*. [29], combined the

empirical formulas with ML techniques to explore the effect of nuclear deformation on α decay half-lives, enhancing predictive performance within deformed regions. The α-decay energies and half-lives have also been studied using a Gaussian process algorithm [30].

A recent and most comprehensive exploration of modern ML models applied to charge radii predictions, Bayram *et al*. [31], showed high potential of both modern tree-based algorithms (XgBoost, Cubist) along with conventional ML. These models have a relatively low RMSE (0.03992-0.0119) fm. Jalili *et al*. [32] have extensively explored Support Vector Regression (SVR) over a diverse range of Z. They pointed out the impact of including A as a feature on the model performance. Their most prominent contribution includes the novel application of RBF (Radial Basis Function) with SVR, a decrease in RMSE from 0.045 fm to 0.016 fm for over 1000 isotopes, as well as improved extrapolation ability of the model. Wu *et al*. [33] applied feed-forward networks with Z, N, and $B(E2)$ inputs to capture magic-number kinks and study symmetry-energy links. Utama *et al*. [34] bridged BNNs with density functional theory to substantially reduce prediction errors by a factor of three. Ma *et al*. [35] used naive Bayes classifiers to refine HFB and semi-empirical radii, with notable gains in terms of accuracy. Dong *et al*. [36] reduced errors significantly by combining Bayesian networks with the NP formula, and Dong *et al*. [37] later included physics-driven features to overcome overfitting. Recently, Jian *et al*. [38] implemented a continuous Bayesian probability (CBP) estimator and Bayesian model averaging (BMA) to enhance the predictions of charge radii from sophisticated theoretical nuclear structure models. They combined the CBP estimator and BMA to refine charge radii predictions from the HFB, RHB, and semi-empirical liquid-drop models. Their method pushed the standard deviation in predictions of R_C below 0.02 fm, demonstrating a strong potential for CBP and BMA in accurate predictions of nuclear charge radii.

Despite the broad scope of theoretical and ML-based studies, two prominent gaps persist. First, global models such as RHB reproduce broad charge-radius trends but fail to capture fine features at the isotopic chain level such as odd-even staggering and exotic-isotope kinks across the nuclear chart. Second, ML approaches (ANNs, Bayesian NNs, SVR) achieve competitive RMSEs but remain constrained by limited data and imbalanced experimental data, narrow feature sets, and a lack of interpretability. Moreover, most of the ML studies are limited to $A \geq 40$ and $Z \geq 20$. To the best of our knowledge, no previous study has leveraged CatBoost's ordered boosting with Optuna-TPE optimization, combined complementary experimental datasets and RHB-based theoretical features of PC-X data, and used SHAP analysis to provide feature-level insights both globally and at the

chain level around N=50.

In this study, CatBoost regression was employed for predicting R_C . The model was trained on merged and processed data extracted from three data tables: (Li *et al.* [39], Angeli *et al.* [40], and PC-X [41]. These tables systematically pushed our study range to $A \geq 17$ and $Z \geq 8$ below the previously reported range, *i.e.*, $A \geq 40$ and $Z \geq 20$. For the best possible performance, the hyperparameters were efficiently searched with Optuna's TPE sampler, and model's interpretation was conducted with SHAP analysis. The Kr and Sr chains were reproduced to assess the model's ability to capture the characteristic "kink" at $N = 50$ and odd-even staggering effects. We demonstrate that our approach outperforms the theoretical model (Skyrme-Hartree-Fock-Bogoliubov model) and established ML techniques in rigorously cross-validated tests. The main contribution of our study is the higher predictive performance achieved by the first implementation of CatBoost, which is systematically combined with Optuna-driven Bayesian hyperparameter adjustment.

The paper is organized as follows. Section II presents a quick overview of the theoretical background of the CatBoost algorithm, Optuna's Bayesian optimization with TPE sampler, SHAP values, and RHB-derived physics features. Section III describes dataset merging and preprocessing, feature selection, hyperparameter tuning, and model evaluation protocol. Section IV reports cross-validation and hold-out performance, comparative benchmarking, optimization history, SHAP insights, and isotopic-chain kink reproduction. Section V concludes the manuscript by highlighting the main findings of this study and outlining directions for future research.

II. THEORY

A. CatBoost

The CatBoost algorithm works on the gradient boosting framework using functional gradient descent to iteratively improve an ensemble predictor [42]. After every iteration m, the current approximation F_{m-1} is updated simply as

$$F_m(x) = F_{m-1}(x) + p_m h(x; a_m), \tag{1}$$

where p_m is the step size and $h(x; a_m)$ is a chosen binary decision tree parameterized by a_m, which is used to estimate the negative gradient of the loss over the training data [42]. Each tree segregates the feature space into disjoint regions and fits constant values on the leaves, which builds its structure greedily to minimize residual error.

CatBoost fits gradient-boosted decision trees using a second-order expansion of a regularized loss [43], which is itself an implementation of Prokhorenkova *et al.* [42]

and the original gradient boosting framework of Friedman [44] on the CatBoost model. At iteration t, new tree w_t is chosen to minimize

$$L = \sum_{i=1}^{n} \left[g_{i,t} w_t(x_i) + \tfrac{1}{2} h_{i,t} w_t(x_i)^2 \right] + u(w_t), \tag{2}$$

where $g_{i,t} = \partial_{W_{t-1}} \ell\left(y_i, W_{t-1}(x_i)\right)$, $h_{i,t} = \partial^2_{W_{t-1}} \ell\left(y_i, W_{t-1}(x_i)\right)$, and $u(w_t)$ denotes a regularization term to penalize the complexity of the tree. The distinct and powerful feature of CatBoost is ordered boosting. The standard gradient boosting approaches, such as XGBoost, LightGBM, and AdaBoost [45−47], are prone to overfitting on small or skewed datasets. These models are prone to overfitting the training set by learning from their residuals. This issue is addressed in the CatBoost using ordered boosting. Rather than building each tree using the entire training data, ordered boosting draws on progressively larger prefixes of data. For each new split, only previous observations (not those that are being split) are used to calculate gradients, preserving unbiasedness and independence (for further details, see Ref. [43]). CatBoost's ability to perform unbiased ordered boosting and effective categorical feature handling accelerates training and improves its generalization ability in relation to that of other boosting models.

B. Bayesian hyperparameter optimization with Optuna

Hyperparameter tuning is crucial for achieving high model accuracy, but all the possible configurations can be extremely expensive given that each combination of parameters requires full model training and validation. To address time and cost issues, we applied Optuna's Bayesian framework [48], which leverages past trial results to direct subsequent search and lower the total number of evaluations. Hyperparameter tuning is formulated as

$$\min_{x \in \mathcal{X}} f(x), \tag{3}$$

where $x \in \mathcal{X}$ denotes a hyperparameter configuration and $f(x)$ is the K-fold cross-validated loss. Optuna [48] constructs the search space dynamically within the objective using calls and allows conditional and hierarchical parameter definitions. It employs the tree-structured Parzen estimator (TPE) sampler [49], which models two densities:

$$p(x \mid y) = \begin{cases} \ell(x), & y < y^*, \\ g(x), & y \geq y^*, \end{cases} \tag{4}$$

where $\ell(x)$ is the density formed by observations $\{x^{(i)}\}$ with $f(x^{(i)}) < y^*$, and $g(x)$ is formed by the remaining observations. The TPE algorithm depends on y^* to the quantile γ of the observed y-values, so that $P(y < y^*) = \gamma$. New potential best combinations of hyperparameters are then drawn by maximizing the ratio $\ell(x)/g(x)$, focusing the search only on the promising regions of the hyperparameter space. This complete package of dynamic search-space definition, Bayesian TPE sampling, and effective metadata storage provides efficient exploration of complex, high-dimensional hyperparameter space in a short time and reduced computational cost.

C. SHAP value analysis

SHAP (SHapley Additive exPlanations) [50] is a technique based on game theory and used to interpret the output of ML models by determining the contribution of each feature to the model's performance. SHAP values use Lloyd Shapley's original cooperative game theoretic solution [51] to feature attribution in ML. The explanation model is given by a linear function of binary variables,

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i', \qquad (5)$$

which uniquely satisfies the axioms of local consistency, accuracy, and missingness [50]. Then, KernelSHAP approximates $\{\phi_i\}$ using a single weighted regression; on the other hand, TreeSHAP calculates them exactly for tree ensembles in polynomial time [52].

D. Physics-driven features from RHB (PC-X)

In the present analysis, we employed the physics-driven features computed in [41] using the relativistic Hartree-Bogoliubov framework with separable pairing force incorporating the latest PC-X parameterization. The nonlinear point-coupling PC-X functional within the covariant density functional theory of nuclear structure provides a self-consistent description of bulk saturation, shell effects, and pairing effects. Instead of recomputing theoretical features, we used S_n, S_{2n}, and BE directly from [41]. The reason behind our choice of features computed using PC-X functional was the large data volume, *i.e.*, 9162 nuclei (ensuring BE values for virtually all of our experimental entries), and its demonstrated high fidelity in reproducing the experimental ground state properties, including BE and nuclear charge radii across the nuclear chart. These properties make it a reliable physics-driven feature source, encoding saturation properties, symmetry-energy effects, shell-closure behavior, and pairing correlations, establishing a theoretical foundation for our Cat-Boost model.

III. METHODOLOGY

A. Data acquisition and preprocessing

We used three state-of-the-art datasets [39−41] to build this CatBoost model of nuclear charge radii. The first data source [40] is an updated compilation of experimental charge radii that contains experimentally measured charge radii of 909 isotopes ranging from hydrogen ($Z = 1$) to curium ($Z = 96$). The second data source [10] contains 257 samples of nuclear charge radii from beryllium ($Z = 4$) to radium ($Z = 88$), where 236 of them were measured using laser spectroscopy. The third data source is a set of theoretical ground state properties of 9162 isotopes calculated using the relativistic Hartree-Bogoliubov framework with separable pairing forces along with the latest point coupling density functional PC-X [41]. First, we integrated Angeli and Marinova's with the latest data table compiled by Li, Luo, and Wang. In the case of clashes, *i.e.*, repetition of the charge radius of a single isotope in both datasets, we retained the latest charge radius compiled by Li, Luo, and Wang. This integrated dataset was then introduced with new columns of theoretical ground state properties from the third dataset. A, N, and Z were used as identifiers for adding new features from the theoretical dataset. After the integration, Tukey's "1.5 × IQR" rule [53] was used to find and eliminate potential outliers. The Tukey's Interquartile Range (IQR) method is a robust, distribution-agnostic technique that is widely used in exploratory data analysis and outlier removal. In this method, first quartile (25th percentile) Q1 and third quartile (75th percentile) Q3 values were calculated using the pandas library, and then the interquartile range was established, followed by the calculation of lower and upper fences. The complete calculation is expressed as follows:

$$Q_1 = 4.2144 \text{ fm}, \qquad (6)$$

$$Q_3 = 5.3640 \text{ fm}, \qquad (7)$$

$$\text{IQR} = Q_3 - Q_1 = 1.1496 \text{ fm}, \qquad (8)$$

$$\text{Lower fence} = Q_1 - 1.5\,\text{IQR} = 2.4901 \text{ fm}, \qquad (9)$$

$$\text{Upper fence} = Q_3 + 1.5\,\text{IQR} = 7.0884 \text{ fm}. \qquad (10)$$

Based on this technique, 17 out of 1040 nuclei fell outside the lower fence. Figure 1 shows the resulting box plot with outliers highlighted in red, and Table 1 lists the corresponding nuclides.

The theoretical ground state properties were only available for the nuclides ranging from oxygen ($Z = 8$) to darmstadtium ($Z = 110$), and 26 nuclides had missing val-
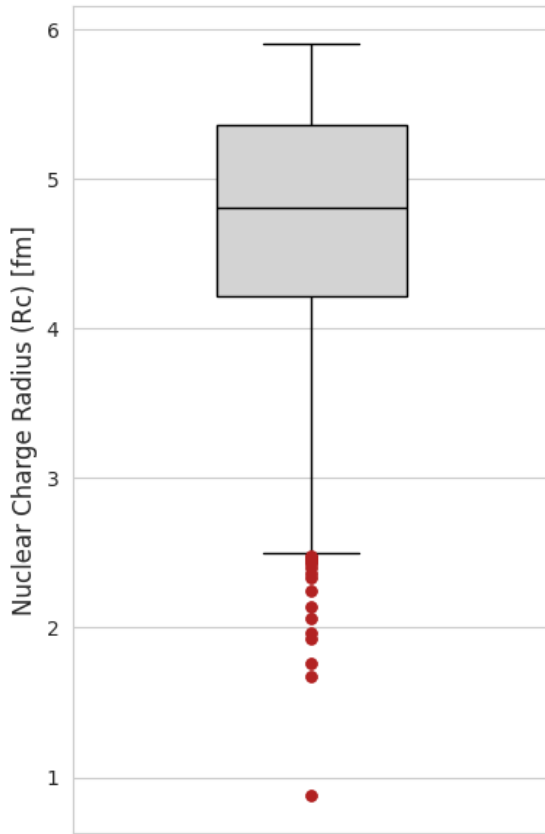
Fig. 1. (color online) Box-plot of R_c using 1.5×IQR rule.

Table 1. Nuclides flagged as outliers by Tukey's $1.5 \times$ IQR rule on R_c.

Sr.#	A	N	Z	R_c/fm
1	1	0	1	0.8783
2	2	1	1	2.1421
3	3	2	1	1.7591
4	3	1	2	1.9661
5	4	2	2	1.6755
6	6	4	2	2.0660
7	8	6	2	1.9239
8	7	4	3	2.4440
9	8	5	3	2.3390
10	9	6	3	2.2450
11	11	8	3	2.4820
12	10	6	4	2.3612
13	11	7	4	2.4669
14	10	5	5	2.4277
15	11	6	5	2.4060
16	12	6	6	2.4702
17	13	7	6	2.4614

ues of theoretical ground state properties. We simply excluded these isotopes from our study, which became restricted to the range between oxygen ($Z = 8$) and curium ($Z = 96$). The reasons behind exclusion are that 17 nuclei of them were flagged as outliers, and the remaining 9 isotopes were also removed to keep our data complete (without missing entries). Finally, we ended up with the consolidated data table of charge radii of 1014 isotopes with predictors N, Z, theoretical BE, S_{2n} (two-neutron separation energy), and S_n (one-neutron separation energy). Including mass number (A) would add perfect multicollinearity with neutron number and proton number (given that $A = N + Z$); moreover, it would provide no independent information. Our preliminary tests showed no improvement in predictive performance. Thus, its exclusion can enhance the model stability and preserve the clarity for SHAP-based feature importance analysis. However, as detailed in Section III.B, S_n and S_{2n} were subsequently removed from the feature set for the final model.

B. Feature selection and model implementation

In the feature selection phase, a significantly lightweight CatBoost model was trained with iterations = 500, learning_rate = 0.1 for quick insight into the candidate features and selection of the most promising ones. A random seed of 42 was maintained to ensure the reproducibility of the feature selection process. Rather than training on a single random test/train split, we k-fold cross-validated (described in Section III.D) on shuffled 10-folds but with a fixed random seed of 42. In this exploratory feature score of a lightweight model (illustrated in Fig. 2), neutron number, BE, and proton number were the most influential features, with mean importances of 36.8%, 28.1%, and 26.7,%. Based on these insights, we dropped less important features, *i.e.*, S_n and S_{2n}, corresponding to theoretical S_n and S_{2n}, which demonstrated notably low mean importances (4.6% and 3.8%, respectively). Therefore, they were dropped from the final model. Using only the three features mentioned above, we achieved promising predictive performance. Consequently, we reduced our feature set to these three features to demonstrate that CatBoost performs well even with a limited feature set.

C. Hyperparameter optimization

The model was then tuned for the best possible predictive performance by searching for optimal hyperparameters. Hyperparameters were selected via Optuna's TPE sampler. The search was performed with 10-fold cross-validation by targeting the minimum average RMSE across all folds. In total, 200 Optuna trials were executed, and the metadata of each trial was stored in a relational database file, which we queried later to analyze search history using the Optuna "best-so-far" plot (Fig. 3). Hyperparameter interactions were studied using a parallel-
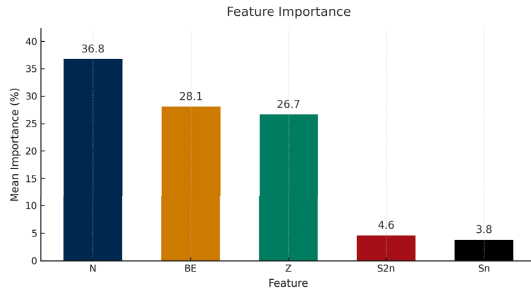
Fig. 2. (color online) Preliminary feature-importance bar plot of selected candidates.
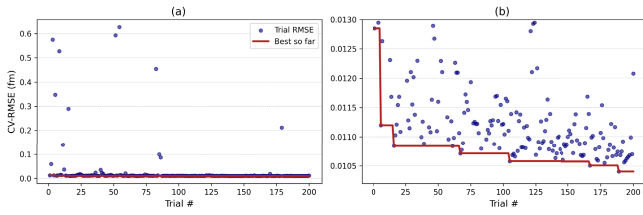

Fig. 3. (color online) Optuna CV-RMSE progress over 200 trials. (a) Full-range CV-RMSE history for all trials. (b) Zoomed-in view (outliers clipped).

coordinate plot (Fig. 4). The full search space is summarized in Table 2.

D. Model evaluation protocol

To assess the predictive performance and reliability of our model, we rigorously evaluated our model in a systematic protocol.

Error Metrics: Three standard regression metrics for performance evaluation were used on both 10-fold cross-validation and unseen hold-out sets:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| y_i - \hat{y}_i \right|, \tag{11}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}, \tag{12}$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \tag{13}$$

where N denotes the number of samples (isotopes in our case), y_i is the true (experimental) charge radius, and $\hat{y}_i$ is the predicted charge radius by Catboost.

● Mean absolute error (MAE): It represents the simple average absolute deviation. Given that it weighs all values equally, it is robust and relatively less sensitive
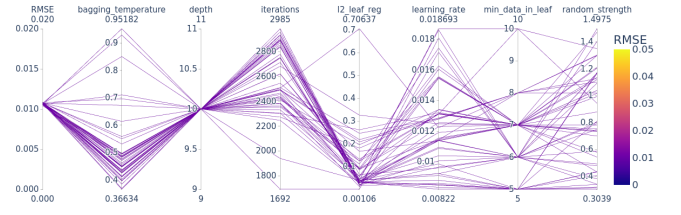

Fig. 4. (color online) Parallel-coordinates plot of the top 50 Optuna trials, showing the sampled hyperparameter values and their corresponding 10-fold CV-RMSE (color-coded).

Table 2. Search space for CatBoost hyperparameters.

Hyperparameter	Type	Search Range	Scale
iterations	integer	$1000-3000$	linear
learning_rate	float	$10^{-4}-10^{-1}$	log
depth	integer	$6-12$	linear
l2_leaf_reg	float	$10^{-3}-10$	log
random_strength	float	$0.0-2.0$	linear
bagging_temperature	float	$0.0-1.0$	linear
min_data_in_leaf	integer	$5-20$	linear

to large errors.

● Root mean squared error (RMSE): It is the most commonly used metric in regression studies. It squares errors, computing averages. This penalizes larger errors more heavily, thus making it robust to highlight significant prediction failures.

● Mean absolute percentage error (MAPE) : It describes errors as percentages of real values, allowing scale-independent comparison between different measurement ranges. Moreover, percentage error is always more intuitive than all other metrics.

Cross-Validation: The whole training dataset (80% of the total, *i.e.*, 811 nuclides) was segregated into ten equal samples via random sampling, and then training and testing were performed on these ten samples. Each time, nine samples were used to train, and the remaining sample was used as part of the validation set. This 10-fold CV was used both for hyperparameter tuning and to report error metrics (with standard deviation) over all folds (see Table 3 in Section (4) for fold-wise RMSE, MAE, and MAPE, plus mean ± SD and hold-out results).

Hold-Out Test: To ensure the robustness and reliability, the final hyperparameter-tuned model was tested on a fully unseen holdout test set of 203 nuclei. The final performance was reported on the 203-nuclide test set. This procedure diagnoses overfitting and allows robust assessment of model accuracy for real-world predictions.

Comparative Benchmarking: The results were compared against nine recently reported ML methods (RF,

Table 3. Fold-wise metrics (RMSE, MAE, MAPE) for 10-fold CV on the training set, along with mean ± SD and hold-out test results.

Fold	RMSE	MAE	MAPE (%)
1	0.0079	0.0058	0.135
2	0.0095	0.0053	0.124
3	0.0094	0.0060	0.139
4	0.0182	0.0089	0.228
5	0.0099	0.0065	0.153
6	0.0126	0.0079	0.198
7	0.0070	0.0056	0.127
8	0.0108	0.0068	0.155
9	0.0119	0.0078	0.180
10	0.0091	0.0058	0.131
Mean CV	0.0106 ± 0.0030	0.0067 ± 0.0011	0.157 ± 0.033
Test	0.0102	0.0067	0.154

QRF, Cubist, XGBoost, GPPK, MARS, SVR, ANN, and CNN) [31, 54] and the BRUSLIB BSkG3 [55] theoretical data library for ground state properties based on Hartree-Fock-Bogoliubov (HFB) with Skyrme Energy-Density Functional. All the models were compared assuming experimental measurements as the gold standard.

Interpretability: The performance of CatBoost is interpreted using residual analysis, parallel-coordinate plot, and SHAP summary analysis.

● **Residual Analysis:** Discrepancy plots and predicted-vs-true scatter (Fig. 5) were used for residual analysis and to highlight any systematic biases.
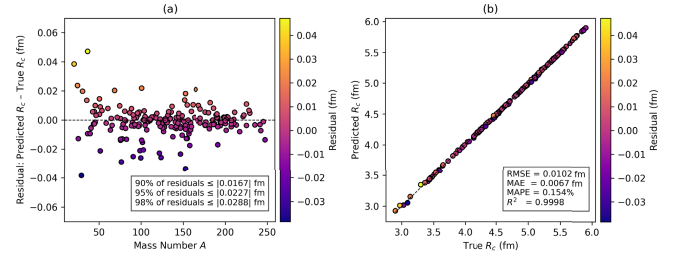
● **Hyperparameter Insights:** The Optuna "best-so-far" plot (Fig. 3) was used to track convergence history, while parallel-coordinate plots (Fig. 4) revealed key insights into hyperparameter interactions.

● **Feature Contributions:** SHAP values were computed on the training data, and a beeswarm SHAP summary plot was used to rank feature importances and to interpret our model.

This rigorous, systematic, and comprehensive evaluation protocol ensures that our conclusions about model performance and our claim of superiority over other models are both physically meaningful and statistically sound.

IV. RESULTS AND DISCUSSION

In this section, we analyze the accuracy of our Cat-Boost model and its performance on charge radii predictions. First, we present a summary of 10-fold cross-valid-

Fig. 5. (color online) (a) Discrepancy Plot of Residuals (Predicted R_c - True R_c) vs mass number A. (b) Predicted versus true R_c.

ation results on the training set by using RMSE, MAE, and MAPE as performance metrics to examine the overall stability across 10-fold isotopic splits. Then, we analyze the model's performance on hold-out unseen tests and quantitatively compare the predictive abilities of our model with recently reported ML algorithms. Then, we analyze the outcome of hyperparameter tuning, followed by a study of feature interactions and SHAP analysis. Finally, we use our trained model to predict krypton and strontium isotopic chains to compare our CatBoost predictions with theoretical models (BSkG3 and PC-X) and previously reported ML algorithms.

A. 10-fold cross-validation and hold-out test

Table 4 summarizes error metrics of the CatBoost model, while Table 3 presents detailed fold-wise error metrics. The hold-out RMSE test was 0.0102 fm, which is notably low, demonstrating a considerably high predictive performance on unseen real-world predictions. The mean RMSE of 10 folds was 0.0106 fm with a standard deviation of 0.0030 fm. The fold-wise RMSE ranged between 0.0070 fm and 0.0182 fm, where the upper value is still lower than the average of 10-fold error for many models (Table 5). The 10-fold average of MAE and MAPE was 0.0067 fm with a standard deviation of 0.0011 fm and 0.157% with a standard deviation of 0.033%, respectively, demostrating remarkable predictive accuracy. In K-fold CV, we found that eight out of 10 folds demonstrated RMSE below 0.012 fm. This consistently low error suggests that the model is stable and not sensitive to variations in training-validation splitting. In summary, consistently low error is evidence that our model captures the underlying relations between the charge radius and our features (N, Z, BE).

In Fig. 5, we present a two-way perspective on the

Table 4. 10-fold CV and hold-out test errors.

Metric	Mean (CV)	SD (CV)	Test
RMSE/fm	0.0106	0.0030	0.0102
MAE/fm	0.0067	0.0011	0.0067
MAPE (%)	0.157	0.033	0.154

Table 5. Cross-validation and RMSE/MAE (in fm) test for our CatBoost framework compared with other data-driven models on the charge-radius dataset.

Model	CV-RMSE	CV MAE	Test RMSE	Test MAE
CatBoost (this study)	0.0106	0.0067	0.0102	0.0067
Cubist [31]	0.01199	0.00770	0.0102	0.0075
Extreme Gradient Boosting (XGBoost) [31]	0.01533	0.01114	0.0125	0.0093
Random Forest (RF) [31]	0.01746	0.01173	0.0138	0.0102
Quantile Regression Forest (QRF) [31]	0.01791	0.01232	0.0140	0.0104
Gaussian Process Predictive Kernel (GPPK) [31]	0.03715	0.02901	0.0346	0.0273
Multivariate Adaptive Regression Splines (MARS) [31]	0.03462	0.02745	0.0346	0.0277
Support Vector Regression (SVR) [31]	0.04286	0.03300	0.0439	0.0336
Artificial Neural Network (ANN) [31]	0.03992	0.03174	0.0564	0.0494
Convolutional Neural Network (CNN) [54]	–	–	0.0156	–

generalization performance of the model. In panel (a), residuals (Predicted R_c – Actual R_c) are represented against atomic mass number A, and color encoding indicates the magnitude of residuals. This plot reveals that most of the points are tightly concentrated at the central zero error line, with very few outliers. The inset quantifying the residual distribution is added, which shows that 90%, 95%, and 98% of the absolute errors are below 0.0167 fm, 0.0227 fm, and 0.0288 fm, respectively. This proves that the errors of the model are consistently low, and very few predictions have errors beyond 0.03 fm. Importantly, no systematic bias over the whole range of mass numbers is observed. A few outliers can be seen with relatively higher residuals, suggesting the limitation of the model in the scarce light-nuclei region of the nuclear chart. Panel (b) illustrates the correlation between predicted and true charge radii, which presents an excellent agreement along the ideal $y = x$ line. The inset in panel (b) shows the performance metrics: RMSE = 0.0102 fm, MAE = 0.0067 fm, MAPE = 0.154%, and R^2 = 0.9998, which means that our model can explain approximately 99.98% of the variance in the nuclear charge radius.

1. *Comparison with Previous ML Studies*

In Table 5, we compare the 10-fold cross-validation of our model as well as hold-out test errors with a range of alternative data-driven algorithms [31, 54]. Notably, CatBoost outperforms all other algorithms by delivering the smallest 10-fold CV with RMSE equal to 0.0106 fm and MAE equal to 0.0067 fm, as well as the lowest test RMSE (0.0102 fm) and MAE (0.0067 fm). The closest competitor of CatBoost is Cubist, which achieves a CV-RMSE of 0.0120 fm with the same test RMSE. XGBoost, Random Forest, and Quantile Random Forest rank third, fourth, and fifth in the competition, respectively, with CV errors that exceed those of CatBoost by approximately 15%−30%. Gaussian process polynomial kernel (GPPK),

multivariate adaptive regression splines (MARS), support vector regression (SVR), and artificial neural networks (ANN) exhibit relatively larger errors (CV-RMSE up to 0.04286 fm and RMSE test up to 0.0564 fm), while the CNN exhibits a RMSE test of 0.0156 fm. In other words, GPPK, MARS, SVR, and ANN are approximately three to four times less accurate than CatBoost (see Table 5). These results establish the fact that CatBoost excels in nuclear charge-radii modeling, combining the lowest errors in both cross-validation and hold-out testing. Its consistently superior performance, especially against modern algorithms such as Cubist and XGBoost, demonstrates the robust generalization ability of Cat-Boost. This highlights the potential of CatBoost in nuclear science, particularly in ground state properties and nuclear structure, leading to efficient and more accurate modeling of nuclear structure and even nucleosynthesis.

B. Hyperparameter tuning history and interactions

In Fig. 3, we present a full-scale optimization history (panel a) and a zoomed-in view of the best-so-far trajectory (panel b) of Optuna-driven 10-fold CV hyperparameter search over 200 trials. If we discuss the first 50 trials in panel (a), trial 0 begins at 0.012849 fm, then for exploration it spikes at trials 1 (0.060102 fm) and 2 (0.574825 fm), and then settles back to 0.012948 fm at trial 3. We observe a prominent improvement at trial 5, where CV-RMSE falls to 0.011198 fm. In subsequent trials (6-49), CV-RMSE oscillates mostly between 0.010880 fm (trial 38) and 0.015588 fm (trial 24), with occasional higher values at trials 10 (0.139897 fm), 14 (0.289320 fm), 35 (0.024157 fm), and 39 (0.035871 fm) as the sampler initially explores various hyperparameter regions.

Panel (b), downscaled on the best-so-far CV-RMSE trajectory, which is also present in panel (a), shows extreme outliers clipped off to illustrate the evolution of optimization effectively. The best-so-far curve descends as

the trials proceed until trial 188, which achieves a minimum CV-RMSE of approximately 0.0104 fm. This trial remains the best through trial 200. This evolution of rapid early descent followed by continuous fine-tuning confirms that our 200-trial search was both efficient and sufficient to search the optimal CatBoost hyperparameters for the prediction of nuclear charge radii. We observe that additional trials might slightly reduce CV-RMSE, but given that we achieved the desired performance, and it stabilized by trial 200, we limited our search to 200 trials to balance accuracy and computational cost. To study the interactions among the hyperparameters, we filtered out the top 50 trials (with the lowest CV-RMSE) and visualized them using the parallel-coordinates plot among the CV-RMSE and all the hyperparameters. Fig. 4 shows a parallel-coordinates plot of the 50 Optuna trials with the lowest 10-fold CV-RMSE. Each line tracks the hyperparameter settings of one trial, *i.e.*, the number of iterations, learning rate, tree depth, l_2-leaf regularization, bagging temperature, minimum data in leaf, and random strength; note that it is color-encoded by its corresponding CV-RMSE. Interestingly, all the best configurations use a single depth of 10 and learning rates between 0.008 and 0.018. Moreover, the iteration counts evolve toward the upper end of the search range, and mostly spread between 2300 and 2985. In addition, l_2- leaf regularization mostly remains below 0.15, the bagging temperature falls in a moderate range, *i.e.*, between 0.4 and 0.55, and the minimum data in leaf is clustered on 5, 6, 7, and 8. The random strength is almost equally spread between 0.8 and 1.4. In Table 6, we report the best hyperparameter values obtained from trial 188 in Optuna. Using these parameters on our processed data, one can reproduce the CatBoost model with exact predictions and evaluation metrics. This highlights the full reproducibility of this study.

C.   SHAP analysis

In Fig. 6, we present a SHapley Additive exPlanations (SHAP) beeswarm plot that quantitatively analyzes how each input feature impacts the predictions of nuclear charge radii by the CatBoost model. The features are ordered from top to bottom based on the mean SHAP value. This highlights that PC-X BE is the most influential feature, followed by Z and N. Each point in the plot corresponds to an individual nuclide from the training set, horizontally positioned based on its SHAP value, which reflects its specific impact on the predicted radius. It is color-encoded to show the real value or magnitude of the corresponding feature value from blue to red (low to high). Note that all three features demonstrate both positive and negative SHAP contributions: nuclei with higher BEs, proton numbers, or neutron numbers systematically exhibit positive SHAP values, corresponding to larger predicted radii. This outcome is obvious and aligned with

Table 6.   Optimal CatBoost hyperparameters.

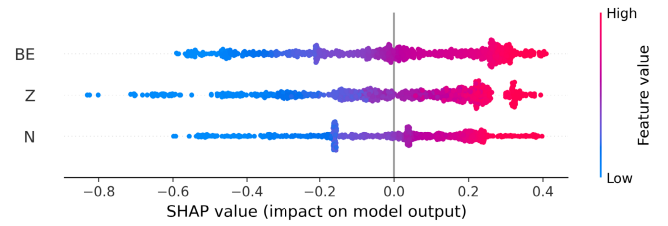| Hyperparameter	Value
iterations	2916
learning_rate	0.00912
depth	10
l2_leaf_reg	0.04002
bagging_temperature	0.48469
min_data_in_leaf	5
random_strength	1.16468

Fig. 6.   (color online) Beeswarm plot of SHAP values for the CatBoost charge-radius model.

the established nuclear theory, where greater proton and neutron numbers naturally lead to larger nuclear sizes owing to increased nucleon occupancy in nuclear shells. Likewise, total BE grows with mass number A; and given that the charge radius scales as $R \approx A^{1/3}$, nuclei with higher BE usually produce positive SHAP contributions. Therefore, in general (for local chain-level, see Section IV.C.1), charge radius increases with an increase in BE. Conversely, lower BE corresponds to a negative SHAP value, reflecting smaller charge radii.

Quantitatively (see Table 7), BE achieves the highest mean absolute SHAP (0.223), followed by Z (0.213) and N (0.176), demonstrating that it drives the model output more strongly on average. Z exhibits the largest SHAP range (−0.83 to +0.39), showing that extreme atomic numbers can both strongly increase and decrease predicted radii. The near-zero mean SHAP values ($Z \approx 3 \times 10^{-5}$, $N \approx 1.4 \times 10^{-3}$, BE $\approx -1.5 \times 10^{-3}$) indicate that positive and negative effects balance each other across the training set. This implies that the model is well balanced and does not overly overestimate or underestimate the predictions. In conclusion, this SHAP analysis confirms that BE is the most influential feature overall, consistent with its fundamental role in nuclear structure, followed by neutron and proton numbers. This highlights that the inclusion of theoretical BE in the dataset as a feature has helped the model to generalize well over the data and increase prediction accuracy.

1.   *Localized SHAP near N = 50 shell closure*

We also explored the localized SHAP for Sr ($Z = 38$) and Kr ($Z = 36$) chains, computed using an interventional

Table 7.   SHAP value summary: mean SHAP, mean absolute SHAP, span, and min-max range, ordered by descending mean absolute impact.

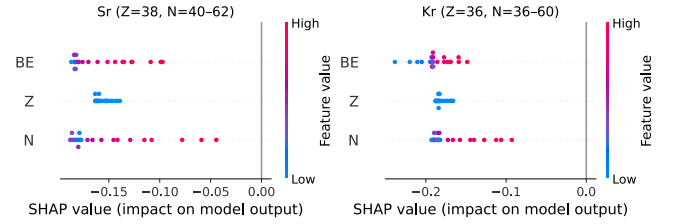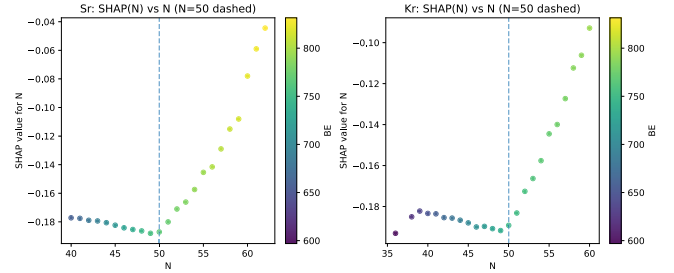| Feature	Mean SHAP	Mean |SHAP|	Span (max−min)	Range (min to max)
Binding energy (BE)	−0.00145	0.2234	0.9990	−0.5898 to +0.4092
Proton number (Z)	0.00003	0.2132	1.2248	−0.8305 to +0.3942
Neutron number (N)	0.00142	0.1759	0.9955	−0.5977 to +0.3978

TreeExplainer on a global background sampled from the whole training set. This common baseline stabilizes the sign and scale of attributions, and we interpret results within each chain.

The chain-level beeswarms (Fig. 7) show that at the chain level, neutron number N is the dominant feature (largest horizontal spread), i is negligible inside each chain (constant value), and BE is a secondary, often negative, contributor in this local neutron window. The dependence plot of SHAP(N) on N (Fig. 8) shows a sharp minimum at $N \approx 50$. For $N < 50$, adding more neutrons yields negative contributions (model predicts charge radius below the global expectation), whereas for $N > 50$, the contributions evolve toward zero, indicating a trend reversal across the shell closure. When *BE* is included as a feature, color-coding shows that although *BE* rises with N, the most negative SHAP(N) appears near $N \approx 50$, demonstrating that at magic number, the local shell effects can override the global trends.

D.   Isotopic chains prediction

Variations in the charge radii within isotopic chains encode the deformation, shell effects, and pairing properties, making isotopic chain patterns a cornerstone of nuclear structure investigations. Reproducing these isotopic chain trends with high accuracy can be used as a litmus test to assess any predictive model, ranging from a basic empirical model to a complex mean field model and even data-driven algorithms. To further claim the superiority of our model over others, we focused on krypton ($^{72\text{-}96}$Kr) and strontium ($^{78\text{-}100}$Sr), which exhibit the shell-closure "kink" at $N = 50$ and odd-even staggering due to pairing effects. This comparison allows us to critically assess how our model captures the underlying physics behind nuclear charge-radius variations. Moreover, these isotopic chains lie on or near potential astrophysical r-process pathways. Therefore, an accurate description of these chains can improve the nucleosynthesis models and our understanding of heavy element formation in the universe.

In Fig. 9(a), we present a comparison of CatBoost with Cubist, Random Forest (RF), and Quantile Regression Forest (QRF), using experimental data as a standard benchmark. The prediction data points of CatBoost not only follow the overall trend but also coincide with the experimental points (except $A = 75$) across $A = 72{-}96$,

Fig. 7.   (color online) Localized SHAP (beeswarm) for Sr (Z=38) and Kr (Z=36) with a single global background. Within each chain, N dominates; Z is near zero (constant locally).

Fig. 8.   (color online) SHAP(N) versus N with $N = 50$ marked. Both chains show a minimum near $N \approx 50$ (negative for $N < 50$, recovery for $N > 50$).

accurately reproducing the $N = 50$ shell-closure kink at $A = 86$. RF and QRF also follow the overall trend but consistently overestimate the charge-radius values, with several larger deviations. Cubist shows a prominent error at $A = 78$, and after the shell-closure kink point, it becomes excessively steep and overshoots the charge-radius values. By contrast, CatBoost provides predictions of both even and odd masses, and its odd-A predictions faithfully align with the experimental odd-A charge radii values.

In Fig. 9(b), we compare CatBoost with extreme gradient boosting (XGBoost), Gaussian process predictive kernel (GPPK), and multivariate adaptive regression splines (MARS). Similar to RF and QRF, XGBoost also consistently overpredicts the charge radii values and shows a steep slope after $A = 50$, producing a larger deviation from the experimental values. GPPK underpredicts with notably large errors below $A = 80$ and then oversmoothes the data points, producing a too flat curve. This shows its inability to reproduce actual chain trends and capture odd-even staggering and kinks at shell closure. MARS, while closer to the experimental slope, underes-
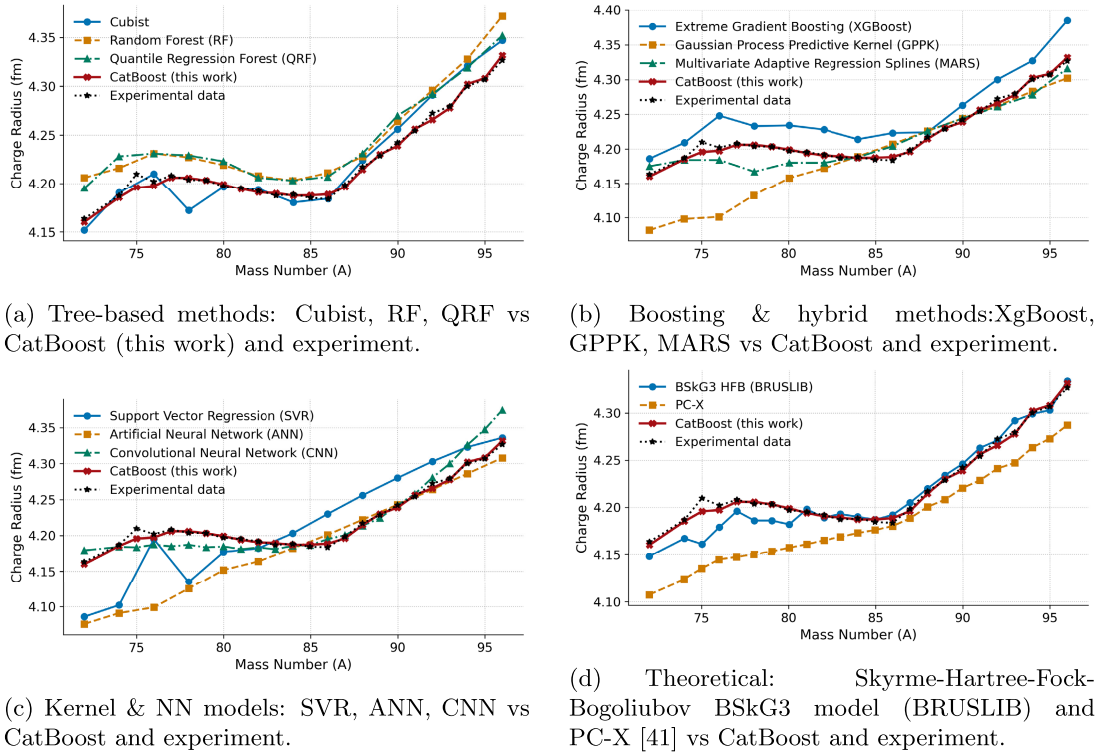
(a) Tree-based methods: Cubist, RF, QRF vs CatBoost (this work) and experiment.

(b) Boosting & hybrid methods:XgBoost, GPPK, MARS vs CatBoost and experiment.

(c) Kernel & NN models: SVR, ANN, CNN vs CatBoost and experiment.

(d) Theoretical: Skyrme-Hartree-Fock-Bogoliubov BSkG3 model (BRUSLIB) and PC-X [41] vs CatBoost and experiment.

Fig. 9. (color online) Charge-radius predictions for the $^{72-96}$Kr chain. CNN and experimental data cover both even and odd-A (note $A = 73$ is unavailable experimentally). Therefore, $A = 73$ data point was dropped. Other models (RF, QRF, Cubist,XgBoost, GPPK, MARS, SVR, ANN and PC-X) report only even-A values, hence gaps visible at odd-A.

timates several points between $A = 74$ to 82. Overall, it follows the trend and captures the kink at shell closure relatively better than tree-based models. Despite this, CatBoost predictions are still superior to those of MARS, as demonstrated by precise overlapping of experimental data points and capture of the chain trend.

In Fig. 9(c), support vector regression (SVR) underestimates the radii predictions at the low-A end. At $A = 82$, it shows an almost constant slope, deviating significantly from experimental values, demonstrating its inability to capture both trend and shell-closure kink. Likewise, the ANN underpredicts the overall trend but accurately predicts in the mid-A range, as shown in the figure. The CNN produces both even and odd A predictions, accurately captures $N = 50$ kinks, and slightly overpredicts radii in the neutron-rich end.

In Fig. 9(d), we compare our model with the theoretical BSkG3 model (Skyrme-Hartree-Fock-Bogoliubov). We extracted these data from the Brussels Nuclear Library for Astrophysical Applications (BRUSLIB). It captures the chain trend effectively and precisely predicts the odd-even oscillations as well as the kink at $N = 50$. However, it underestimates the charge radii values in the lighter half of the chain until $A = 82$, which highlights its limited performance for the neutron-deficient range.

In Fig. 10, we present the same comparison for the strontium chain. In the strontium chain ($A = 78 - 100$),

classical tree models (Cubist, RF, QRF) show a relatively good fit with shell-closure kink but also exhibit localized deviations at multiple points larger than Catboost's predictions. XGBoost follows the overall trend, captures the kink at shell closure, but yet overestimates beyond $A = 96$, performing worse than CatBoost concerning predictions. GPPK and MARS underestimate at both the light- and heavy-mass ends and produce overly smooth, too-flat profiles that diverge from experimental values. SVR also demonstrates a too-flat line and misses the $N = 50$ kink and overall trend pattern, while ANN underestimates both heavy-mass and light-mass end radii and only fits relatively well in the mid-range ($A = 86$ to 97). Among all approaches (except CatBoost), CNN exhibits the closest fit but still deviates from A=95 to 97. The BSkG3 model reproduces the kink and accurate charge radii between A=85 and 95 but underestimates radii below $A = 80$.

The CatBoost model consistently outperforms all the models in the prediction of both krypton and strontium chains, simultaneously capturing the shell-closure kink and the subtle odd-even oscillations while accurately following the overall trend. Although individual models (tree models, boosting, ANN, CNN, and BSkG3) [31, 54, 55] excel in one or two aspects, none is best in all at once. CatBoost excels at all these qualities simultaneously, which proves its predictive superiority against these models and suggests that it can be a highly promising ap-
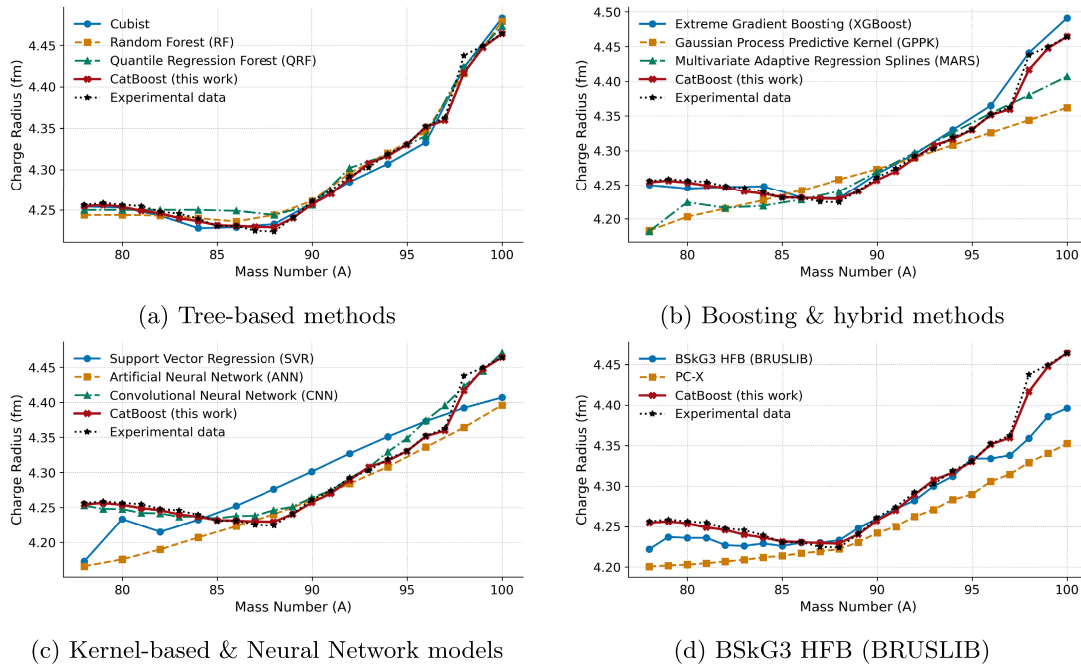
(a) Tree-based methods

(b) Boosting & hybrid methods

(c) Kernel-based & Neural Network models

(d) BSkG3 HFB (BRUSLIB)

Fig. 10.　(color online) Same as Fig. 9, but for the $^{78\text{-}100}$Sr chain.

proach for ground state predictions of properties, including charge radii, error quantification, and refinement of nuclear models.

V. CONCLUSIONS

In the present analysis, we reported the first utilization of the CatBoost model for more accurate computations of ground-state nuclear charge radii across 1014 nuclides. To create a high-quality dataset that is larger than prior research, we combined two extensive experimental compilations (Angeli & Marinova; Li *et al.*). We also incorporated physics-informed features from PC-X computations. An Optuna driven hyperparameter search was conducted to minimize the average 10-fold cross-validation RMSE to 0.0106 fm, and (MAE = 0.0067 fm, MAPE = 0.154%) yielded a hold-out RMSE test of

0.0102 fm (MAE = 0.0067 fm, MAPE = 0.154%). SHAP analysis indicated that BE has an advantageous influence on the model's performance and appeared as the leading predictor, which is consistent with nuclear theory. The residual analysis revealed no significant consistent bias throughout nearly entire atomic mass range. When compared to nine established ML methods and the BRUSLIB library (BSkG3), CatBoost exhibited higher global predictive power and stability. Notably, the proposed model accurately modeled isotopic properties including the $N = 50$ shell-closure kink and odd even staggering in Kr and Sr chains, demonstrating its excellent generalizability. These findings establish CatBoost as a powerful, interpretable tool for nuclear radius prediction, paving the path for future enhancements such as enlarged feature sets, uncertainty quantification, and hybrid modeling using ab initio techniques.

References

[1]　P. Ring, and P. Schuck, *The nuclear many-body problem* (New York: Springer Science & Business Media, 2004)

[2]　K. S. Krane, *Introductory nuclear physics* (New York: John Wiley & Sons, 1991)

[3]　K. W. Ford and D. L. Hill, Ann. Rev. Nucl. Part. Sci. **5**, 25 (1955)

[4]　H. F. Ehrenberg, R. Hofstadter, U. Meyer-Berkhout *et al.*, Phys. Rev. **113**, 666 (1959)

[5]　L. R. Suelzle, M. R. Yearian, H. Crannell, Phys. Rev. **162**, 992 (1967)

[6]　I. Angeli, Y. P. Gangrsky, K. P. Marinova *et al.*, J. Phys. G Nucl. Part. Phys. **36**, 085102 (2009)

[7]　C. V. Weizsäcker, Zeitschrift für Physik. **96**, 431 (1935)

[8]　J. Duflo, Nucl. Phys. A **29**, 576 (1994)

[9]　J. Piekarewicz, M. Centelles, X. Roca-Maza *et al.*, Eur. Phys. J. A **46**, 379 (2010)

[10]　N. Wang, T. Li, Phys. Rev. C **88**, 011301 (2013)

[11]　M. V. Stoitsov, J. Dobaczewski, W. Nazarewicz *et al.*, Phys. Rev. C **68**, 054312 (2003)

[12]　S. Goriely, N. Chamel, and J. Pearson, Phys. Rev. C. **88**, 024308 (2013)

[13]　P. G. Reinhard, Rep. Prog. Phys. **52**, 439 (1989)

[14]　C. Ma, Y. Y. Zong, Y. M. Zhao *et al.*, Phys. Rev. C **104**, 014303 (2021)

[15]　R.-Y. Zheng, X.-X. Sun, G. F. Shen *et al.*, Chin. Phys. C **48**, 014107 (2024)

[16] X. Zhang, Z. Niu, W. Sun *et al*., Phys. Rev. C **108**, 024310 (2023)

[17] K. Zhang *et al*., At. Data Nucl. Data Tables **144**, 101488 (2022)

[18] P. Choudhary, P. C. Srivastava, and P. Navrátil, Phys. Rev. C **102**, 044309 (2020)

[19] C. Forssén, E. Caurier, and P. Navrátil, Phys. Rev. C **79**, 021303 (2009)

[20] A. Boehnlein *et al*., arXiv: 2112.02309

[21] E. Yüksel, D. Soydaner, and H. Bahtiyar, Phys. Rev. C **109**, 064322 (2024)

[22] M. Wang, W. Huang, and F. G. Kondev, Chin. Phys. C **45**, 030003 (2021)

[23] W. J. Huang, M. Wang, F. G. Kondev *et al*., Chin. Phys. C **45**, 030002 (2021)

[24] B. Pandey and S. Giri, AIP Adv. **14**, 10 (2024)

[25] C. Y. Tsang *et al*., arXiv: 2107.13985

[26] T. Bayram, S. Akkoyun, S. Okan Kara *et al*., Ann. Nucl. Energy. **63**, 172 (2014)

[27] Z. Jin and M. Yan, Phys. Rev. C **108**, 014326 (2023)

[28] Q. Yuan and P. Qi, arXiv: 2508.03155

[29] H. Q. You, X. T. He *et al*., Nucl. Sci. Tech. **36**, 1 (2025)

[30] Z. Yuan, D. Bai *et al*., Chin. Phys. C **46**, 024101 (2022)

[31] T. Bayram, C. M. Yeşilkanat, and S. Akkoyun, Phys. Scr. **98**, 125310 (2023)

[32] A. Jalili and A.-X. Chen, New J. Phys. **26**, 103017 (2024)

[33] D. Wu, C. Bai, H. Sagawa *et al*., Phys. Rev. C **102**, 054323 (2020)

[34] R. Utama, W. C. Chen, and J. Piekarewicz, J. Phys. G: Nucl. Part. Phys. **43**, 114002 (2016)

[35] Y. Ma, C. Su, J. Liu *et al*., Phys. Rev. C **101**, 014304 (2020)

[36] X. X. Dong, R. An, J. X. Lu *et al*., Phys. Rev. C **105**, 014308 (2022)

[37] X. X. Dong, R. An, J. X. Lu *et al*., Phys. Lett. B. **838**, 137726 (2023)

[38] Jian Liu, Kai-Zhong Tan, Lei Wang *et al*., Nucl Sci Tech. **36**, 215 (2025)

[39] T. Li, Y. Luo, and N. Wang, Atomic Data and Nuclear Data Tables **140**, 101440 (2021)

[40] I. Angeli and K. Marinova, Atom. Data. Nucl. Data Tables **99**, 69 (2013)

[41] Z. X. Liu *et al*., At. Data Nucl. Data Tables **156**, 101635 (2024)

[42] L. Prokhorenkova, G. Gusev *et al*., Adv Neural Inf Process Syst. **31** (2018)

[43] B. So and E. A. Valdez, arXiv: 2406.16206

[44] J. H. Friedman, Ann. Stat. **29**, 1189 (2001)

[45] T. Chen and C. Guestrin, XgBoost: A scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA, 2016), pp. 785.

[46] G. Ke, Q. Meng *et al*., Adv. Neural Inf. Process **30** (2017)

[47] Y. Freund, R. E. Schapire, ICML. **96**, 148 (1996)

[48] T. Akiba, S. Sano, *et al*.: A next-generation hyperparameter optimization framework, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA, 2019)

[49] J. Bergstra, R. Bardenet, Y. Bengio *et al*. Algorithms for hyper-parameter optimization, in *Proceedings of the 25th International Conference on Neural Information Processing Systems* (Granada, Spain, 2011)

[50] S. M. Lundberg and S. I. Lee, A unified approach to interpreting model predictions, in *Proceedings of the 31st International Conference on Neural Information Processing Systems* , 2017, p. 4768

[51] Shapley, L. S., A value for n-person games.

[52] S. M. Lundberg, G. G. Erion, and S. I. Lee, arXiv: 1802.03888

[53] J. W. Tukey, *Exploratory data analysis* (New York: Springer, 1997)

[54] Y. Y. Cao, J. Y. Guo, and B. Zhou, Nucl. Sci. Tech. **34**, 152 (2023)

[55] Y. Xu, S. Goriely, and A. Jorissen, Astron. Astrophys. **549**, A106 (2013)