

Machine Learning Classification for Irradiated Material Gamma Ray Spectra

A. Robb Taylor

arobb8@gatech.edu

Georgia Institute of Technology

Steven R. Biegalski

Georgia Institute of Technology

Derek Haas

The University of Texas at Austin

Khiloni Shah

The University of Texas at Austin

Kohlton Mathis

The University of Texas at Austin

Joseph Lapka

The University of Texas at Austin

Sarah Castro

The University of Texas at Austin

Research Article

Keywords: Machine Learning, Gamma-Ray Spectroscopy, Radiation Detection, Neutron Activation Analysis, Random Forest (RF), Support Vector Machine (SVM)

Posted Date: July 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9771858/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Radioanalytical and Nuclear Chemistry on October 23rd, 2026. See the published version at <https://doi.org/10.1007/s10967-026->

Machine Learning Classification for Irradiated Material Gamma Ray Spectra

A. Robb Taylor^{1*}, Steven R. Biegalski¹, Derek Haas², Khiloni Shah², Kohlton Mathis²,
Joseph Lapka², Sarah Castro²

¹*Nuclear and Radiological Engineering Program, Georgia Institute of Technology, 770
State Street, Atlanta, GA 30332, USA*

²*Nuclear and Radiation Engineering Program, University of Texas at Austin, 110 Inner
Campus Drive, Austin, TX 78705, USA*

**arobbb8@gatech.edu*

Abstract

This work presents a preliminary evaluation of Machine Learning (ML) methods for automated gamma-ray spectroscopy classification across two levels of spectral complexity and three levels of detector resolution. Samples analyzed include both single isotope sources and materials irradiated at the University of Texas at Austin Research Reactor. Sample collection was conducted with Sodium Iodide (NaI(Tl)), Cadmium Zinc Telluride (CZT), and High Purity Germanium (HPGe) detectors. A collection of ML techniques was used and compared to classify spectra according to their correct identity; contrastive techniques were first applied to the single isotope sources before application to the complex irradiated samples with changing activities. Activated samples included NIST-certified coal, soil, tomato leaves, and natural uranium oxide (U_3O_8) irradiated at the UT Austin TRIGA reactor. A five-step preprocessing pipeline consisting of counts-per-second normalization, interpolation onto a common energy grid, total-area normalization, $\log_{1p}$ transformation, and Poisson noise augmentation was applied to place all spectra on a physically meaningful, detector-agnostic representation. Principal

Component Analysis (PCA) reduced the 3001-dimensional feature vectors to between 4 and 13 components prior to classification. Three classifiers were evaluated under Leave-One-Out Cross-Validation (LOO-CV): Random Forest (RF), Support Vector Machine with RBF kernel (SVM), and k-Nearest Neighbors (KNN). For single-isotope sources, all three classifiers achieved no observed failures across all detector types. For the complex irradiated samples, RF and KNN achieved no observed failures (accuracy 1.000, Matthews Correlation Coefficient 1.000) on all three detectors, with SVM reaching 0.974 on HPGe due to cooling-time-dependent spectral drift in the tomato leaf class. Clopper-Pearson 95% confidence intervals and per-class prediction confidence analysis indicate that observed accuracy and prediction reliability diverge for low-statistics measurements, most notably for short-count CZT spectra. These results establish a validated preprocessing and classification baseline for experimental gamma spectroscopy datasets with variable acquisition times, detector types, and post-irradiation histories, and provide a framework for extension into more complex identification.

Keywords

Machine Learning, Gamma-Ray Spectroscopy, Radiation Detection, Neutron Activation Analysis, Random Forest (RF), Support Vector Machine (SVM)

Introduction

Accurate and timely identification of irradiated nuclear materials is a fundamental challenge in nuclear security. When nuclear materials undergo neutron activation, through reactor irradiation or fission, the resulting activation products and/or fission products emit characteristic gamma-ray signatures. Distinguishing these signatures automatically and reliably is critical for applications ranging from environmental monitoring, safeguards verification, non-proliferation, and post-detonation nuclear forensics [1]. Traditional gamma-ray analysis relies on peak-search methods that require subject matter expertise, accurate efficiency calibration, and are time-intensive [2]. These conventional methods do not leverage the full spectral shape of the data, instead focusing

on the relative heights of select photopeaks. Machine Learning (ML) approaches offer the potential to reduce the computational expense of traditional peak-picking, exploit the full spectral shape, and provide rapid automated classification even in low-resolution or dynamic spectral environments [3].

Recent ML advances have demonstrated significant potential for automated gamma-ray spectroscopy analysis [4], [5]. Neural networks operating on full spectral shapes rather than extracted peaks eliminate the need for expert-driven feature selection [6]. Convolutional Neural Networks (CNNs) have shown promise for spectroscopy applications through their ability to capture local photopeak and Compton continuum features [7], [8], [9]. Multi-Layer Perceptrons (MLP) and Naive Bayes classifiers have demonstrated high accuracy even with limited training data [10]. Beyond classification, ML techniques have been applied to detector calibration, anomaly detection, and nuclear forensic attribution tasks including reactor type determination from isotope ratio signatures [10], [11]. Despite this growing body of work, several areas have not been addressed. The majority of published studies focus on single-isotope identification tasks or rely on simulated spectral datasets that do not capture the full complexity of real experimental measurements [12]. This research aims to explore: cross-detector generalization, where a classifier trained on one detector type can perform accurately on another. And classifying thermally neutron-activated environmental matrices, whose spectral character evolves continuously with post-irradiation cooling time due to the decay of short-lived activation products. These goals are addressed by developing and evaluating a complete ML classification pipeline applied to the experimental spectroscopy dataset. This is foundational for a more complex classification task involving post-detonation fission product spectra from fissile materials.

The experimental dataset spans nine source classes across three spectral complexity levels. The simplest level consists of single-isotope sources (^{133}Ba , ^{60}Co , ^{137}Cs , ^{152}Eu , ^{22}Na). The dynamic spectra consist of NIST-certified environmental reference materials (coal, soil, tomato leaves, and U_3O_8) irradiated at the University of Texas at Austin TRIGA reactor [13] and measured across various post-irradiation time windows. Natural uranium oxide exhibits the most temporally dynamic evolution due to

activation of uranium and its progeny. Measurements were collected simultaneously on CZT, HPGe, and NaI(Tl) detectors, enabling direct intra-site cross-detector comparison.

This study establishes generalizable preprocessing and classification strategies for experimental gamma spectroscopy datasets with variable channel counts, variable acquisition times, and variable post-irradiation histories. The results provide a validated baseline for future extension to the more complex spectral discrimination problem of post-detonation fission product identification, where correct identification is critical and available measurement windows are significantly shorter [1], [14].

The present work makes several contributions that distinguish it from the existing literature. First, while prior ML studies in gamma spectroscopy have predominantly relied on single-isotope identification tasks [4], [7] or simulated training datasets [15], [16], this work evaluates classification performance on fully experimental spectra from neutron-activated environmental matrices with no simulated data in the pipeline. Second, the classification of temporally evolving activation product spectra, where spectra change continuously with post-irradiation cooling time, has not been addressed in prior ML spectroscopy literature to the authors' knowledge. Third, simultaneous cross-detector evaluation on CZT, HPGe, and NaI detectors using identical preprocessing and classification pipelines enables direct comparison of detector utility for this task. The application of Poisson noise augmentation to raw counts prior to preprocessing, rather than to processed feature vectors, preserves the correct counting statistics of the measurement process and represents a physically motivated augmentation strategy distinct from generic noise injection approaches used in other spectroscopy ML studies [17].

Experimental

Measurement Facilities and Detector Systems

Measurements were conducted at two independent facilities. At Georgia Institute of Technology, gamma spectra were acquired using three co-located detector systems: a cadmium zinc telluride (CZT) detector, a high-purity germanium (HPGe) detector, and a

sodium iodide thallium-doped NaI(Tl) scintillation detector. All single isotope measurements used a uniform acquisition time of 1800 seconds (30 minutes) per spectrum.

At the University of Texas at Austin Nuclear Engineering Teaching Laboratory (NETL) [13], gamma spectra were acquired using the same three detector types (CZT, HPGe, and NaI(Tl)), deployed simultaneously during each measurement session. Acquisition times varied by source type and detector. Background and ^{60}Co spectra were collected at 30 minutes per spectra. ^{137}Cs and ^{152}Eu were acquired at 5-minute intervals for CZT and NaI, and 30 minutes for HPGe. Activated tomato leaves and soil were counted for 30 minutes. Activated uranium oxide was counted twice, first in 5-minute intervals 2 hours after irradiation, and again in 2-hour intervals 10 hours after irradiation. Coal was counted in 2-hour intervals.

Materials Irradiation

Thermal neutron activation of the environmental reference materials was performed at the University of Texas at Austin TRIGA Mark II research reactor. Four samples were irradiated with their parameters are summarized in Table 1. All irradiations were conducted on August 6–7, 2025.

Table 1 Sample and Irradiation Characteristic. * = Sample irradiated in multiple intervals.

Sample	Mass (g)	Irradiation Time (s)	Power (kW)
Coal (NIST 1632c)	0.13398	300	900
Soil (NIST 2709)	0.18145	30	900
Tomato Leaves (NIST 1573a)	0.10792	10	900
* Natural U_3O_8	0.01635	30	50
		30	1
		120	1

Following irradiation, samples were measured simultaneously on the three detector systems. The post-irradiation cooling time is defined as the elapsed time between the end of irradiation and the middle of each individual measurement file. Cooling time varied because only one sample could be measured at a time. Table 2 summarizes the post-irradiation cooling time ranges for each activated sample and the number of sequential measurement files collected per sample.

Table 2 Spectral Cool Times After Irradiation

Sample	Spectra 1 time, min	Spectra 2 time, min	Spectra 3 time, min	Spectra 4 time, min	Spectra 5 time, min	Spectra 6 time, min	Spectra 7 time, min	Spectra 8 time, min
Coal	507	627	747	867	987	1107	1227	1347
U ₃ O ₈	120	125	130	135	140	145		
U ₃ O ₈	648	768	888	1008	1128	1248	1368	
Leaves	1925	1955	1985	2015	2045	2075		
Soil	1690	1720	1750	1780	1810			

All spectra were stored as two-column tab-delimited text files containing calibrated energy (keV) in the first column and raw integer counts in the second column. The CZT and NaI detectors had 1024-channel spectra with a channel width of approximately 2–3 keV, while the HPGe detector had 16,384-channel spectra with a channel width of approximately 0.12 keV per channel. This variability in channel count and energy scale across files necessitated the interpolation preprocessing step described in the following section. Figures 1 through 5 show the raw energy versus counts of all activation product spectra across the three detector resolutions.

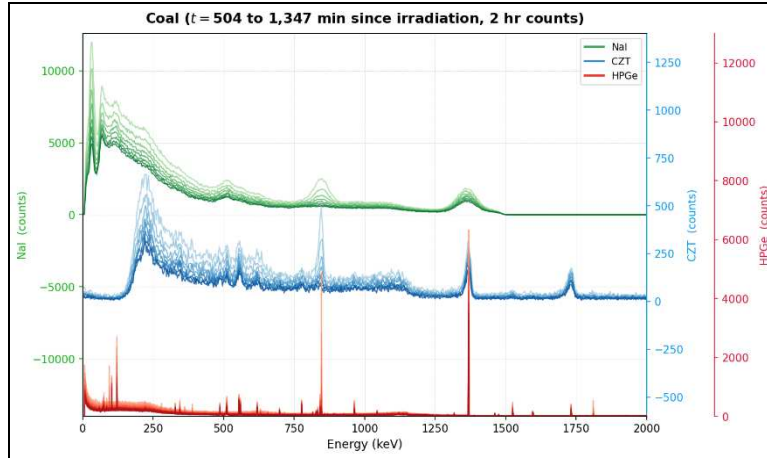


Fig. 1 Coal Activation Spectra

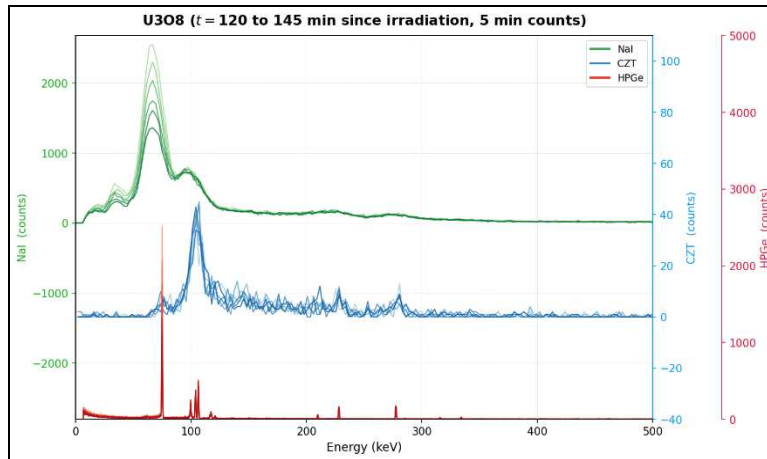


Fig. 2 Early U_3O_8 Activation Spectra

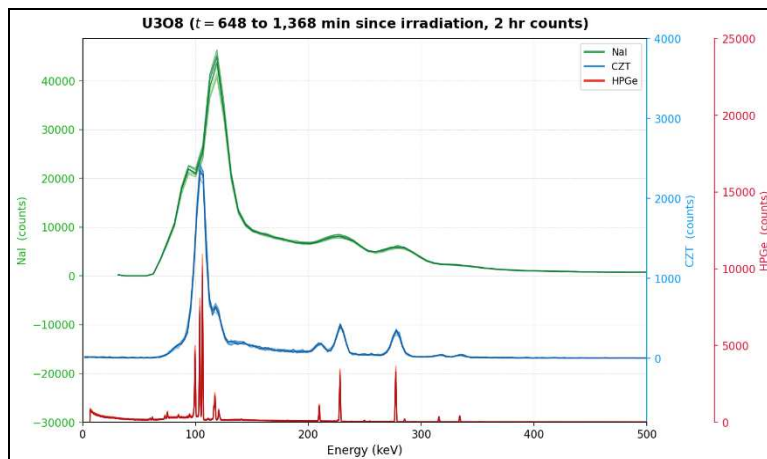


Fig. 3 Late U_3O_8 Activation Spectra

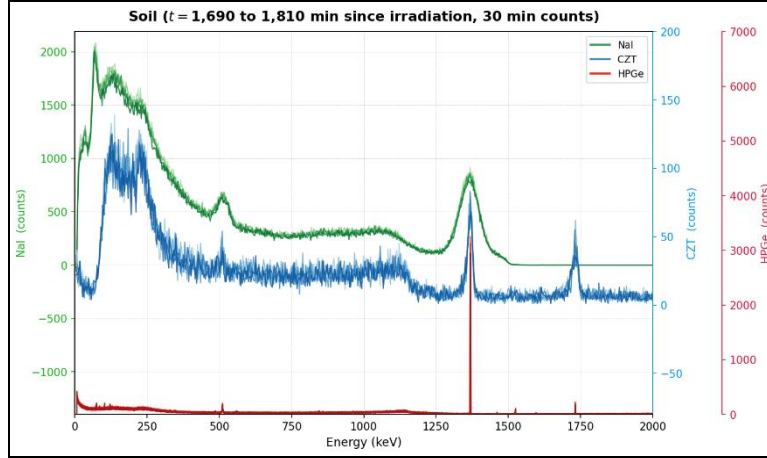


Fig. 4 Soil Activation Spectra

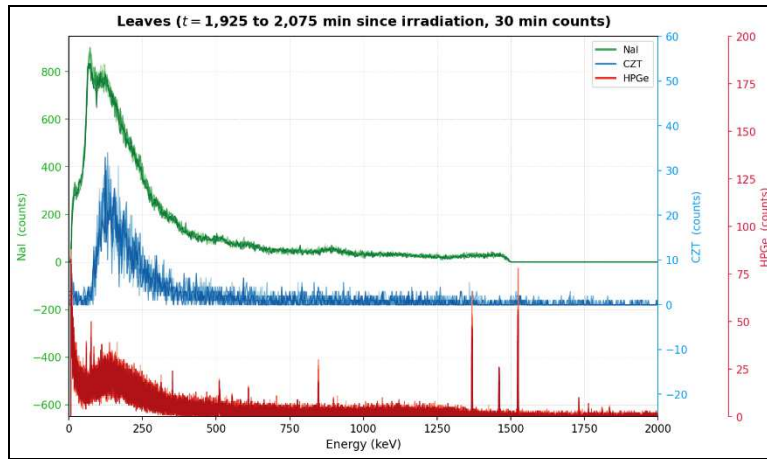


Fig. 5 Tomato Leaves Activation Spectra

Methods

Preprocessing Pipeline

A uniform five-step preprocessing pipeline was applied to every spectrum file prior to feature extraction or model training. The pipeline was designed to place all spectra on a common, physically meaningful representation that is independent of acquisition time, detector channel count, source activity, and source-to-detector distance. This included: Counts-per-second (CPS) normalization, interpolation onto a common energy grid, total-area normalization, Log_{1p} transformation, and feature vector assembly.

Counts-per-second (CPS) normalization involved dividing the raw integer counts in each channel by the acquisition time in seconds, thus converting the spectrum to a count rate representation. This step is important for the activation product spectra where acquisition times differ between samples. CPS normalization allows spectral shape, rather than total accumulated counts, to drive the subsequent analysis.

Interpolation onto a common energy grid was important because the channel quantity varied by detector. Each CPS spectrum was resampled onto a fixed common energy axis spanning 0 to 3000 keV in 1 keV steps, yielding a 3001-element feature vector per spectrum. Resampling was performed by linear interpolation using the calibrated energy values. Energy values outside the measured range of a given file were assigned a CPS value of zero rather than extrapolating, which is physically appropriate since a detector that did not measure a given energy range contributes no counts there. The 3000 keV upper limit was chosen to capture all gamma lines of interest for the sources present but excluded the high-energy tail of the HPGe spectra.

Total-area normalization is important so the ML techniques train on spectral shape rather than intensity. Each interpolated CPS spectrum was divided by its total integrated area, converting the spectrum to a shape descriptor that is independent of absolute count rate.

The Log_{1p} transformation is a compression step often used in gamma spectroscopy ML applications [7], [10] because it reduces the dynamic range between tall, narrow photopeaks and the broad, low-level Compton continuum. This $\log(1 + x)$ transform was applied element-wise to the normalized spectrum. Without this step, classifiers trained on raw normalized spectra would be dominated by the few channels containing high-count photopeaks and would largely ignore the continuum structure that also carries discriminating information. The $\log_{1p}$ form rather than $\log$ is used to handle zero-count channels without producing negative-infinity values.

The result of these steps is a 3001-dimensional floating-point feature vector representing the log-compressed, normalized spectral shape of a single measurement. This vector is the sole input to the dimensionality reduction and classification stages. No peak picking or regions of interest were defined, and thus all discriminating information is derived from the full spectral shape.

196 *Poisson Noise Augmentation*

197 Given the small number of original measurement files per sample type (between 5–8 files
 198 per sample per detector), Poisson noise augmentation was used to increase the effective
 199 training set size [17], [18]. The physical motivation for this approach is that gamma-ray
 200 counting is a Poisson process [19], [20]: the number of counts registered in any channel
 201 during a fixed acquisition time follows a Poisson distribution with mean equal to the true
 202 expected count rate multiplied by the acquisition time. This means that physically
 203 plausible synthetic realizations of any measured spectrum can be generated by drawing
 204 new Poisson-distributed count vectors from the observed counts as the Poisson mean.

205 Augmentation was applied to raw integer counts before any preprocessing step, to
 206 preserve the correct Poisson statistical structure. For each original spectrum, N synthetic
 207 count vectors were generated by drawing independently from $\text{Poisson}(c)$ for each
 208 channel, where c is the observed raw count in that channel. Each synthetic count vector
 209 was then passed through the complete five-step preprocessing pipeline to produce a
 210 corresponding synthetic feature vector. The target augmentation level was set to 30
 211 training samples per class per cross-validation fold. Augmentation was only applied to
 212 the training portion of each cross-validation fold, not the test set. All test evaluations
 213 completed on the unaugmented spectra, ensuring that reported classification accuracy
 214 reflects physical measurements rather than performance on correlated synthetic data. A
 215 sensitivity analysis evaluating augmentation targets of 20, 30, and 50 samples per class
 216 was conducted, but found no meaningful difference in classification accuracy for the
 217 single isotope or irradiated samples. All three values produced identical overall accuracy
 218 results. This confirmed that the classifier performance reported at $\text{aug_target} = 30$ is
 219 stable and not sensitive to the specific augmentation level within the tested range.

220 *Dimensionality Reduction via PCA*

221 The 3001-dimensional preprocessed feature vectors were reduced to a low-dimensional
 222 representation using Principal Component Analysis (PCA) before classifier training. With
 223 approximately 30–50 training samples per class following augmentation, direct

classification in 3001-dimensional space would lead to severe overfitting due to the large number of channels. PCA was therefore applied as a dimensionality reduction step.

PCA was fit on the augmented training data within each cross-validation fold and then applied to transform both training and test data. Fitting PCA on all data before splitting would result in data leakage, as the test spectrum would influence the principal component directions. The number of retained components was selected 95% explainability of the spectrum. In practice this resulted in 4 retained components for the single isotope spectra and 4–13 components for the activation analysis spectra, reflecting the greater spectral diversity of the irradiated material.

The physical interpretability of the retained PCA components was examined through loadings analysis, in which the component weight vectors are plotted as a function of energy and annotated with the known gamma line energies of sources present in the dataset, allowing the principal directions of spectral variation to be related to specific physical emission features.

Classification Algorithms

Three classification algorithms were evaluated: Random Forest (RF), Support Vector Machine with RBF kernel (SVM), and k-Nearest Neighbors (KNN). The selection was motivated by the small training set size, which rules out deep learning approaches but is well-suited to these methods [21].

Random Forest (RF) is an ensemble learning method that constructs multiple decision trees during training using bootstrap aggregation (bagging), with each tree trained on a random subsample of the training data and each node split constrained to a random subset of features [21], and classification is determined by majority vote across all trees. RF was selected as the primary classifier because of its robustness to overfitting in small-sample regimes, its native multi-class support, and its ability to provide feature importance estimates that can be related back to PCA components and ultimately to spectral energy regions. The fixed hyperparameters used were: 200 trees, maximum depth of 10, and minimum samples per leaf of 2.

The Support Vector Machine (SVM) with RBF kernel finds the maximum-margin hyperplane in the PCA-reduced feature space that separates training class boundaries. SVMs are common in gamma spectroscopy classification [7] and can outperform RF in low-sample-count regimes due to their large-margin optimization. Fixed hyperparameters were $C = 10$ and $\gamma = \text{'scale'}$ (data-adaptive kernel bandwidth). The KNN classifier ($k = 5$) was included as a baseline that classifies each test spectrum by plurality vote among its five nearest neighbors in PCA space, allowing direct inspection of which training spectra are most similar to any given test spectrum.

Fixed hyperparameters were used for all three classifiers rather than inner cross-validation tuning. This decision was made because of the small dataset size and the large number of models being evaluated ($6 \text{ site/detector combinations} \times 3 \text{ classifiers} \times 3 \text{ augmentation targets}$).

Cross Validation Strategy

Leave-One-Out Cross-Validation (LOO-CV) was used as the main evaluation strategy because it is well suited for this small dataset. A single 80/20 train/test split would leave only approximately one test sample per class, making the accuracy estimate highly sensitive to which specific file was held out. LOO-CV provides the maximum use of available training data in each fold while ensuring that every measured file is evaluated as a test case. In each outer fold, one original measured spectrum per class was held out as the test sample. Augmentation was then applied to the remaining original training spectra, PCA was fit on the augmented training set, and each classifier was trained and used to predict the class of the held-out real spectrum. This procedure was repeated so every experimental file served as a test sample exactly once. Final accuracy metrics were aggregated across all LOO folds.

Two classification objectives were evaluated simultaneously: source identity classification (e.g., predicting 'Coal', 'Cs-137', 'U₃O₈') and category classification (predicting 'Activation', 'PointSource', or 'Background'). The category label was derived deterministically from the source identity prediction using a fixed lookup table, so no

separate category classifier was trained. Per-class accuracy, overall accuracy, and confusion matrices were computed for both classification levels.

It is important to note that the dataset used in this study is small by machine learning standards, with 5–8 original measured spectra per class per detector. While LOO-CV maximizes the use of available data and Poisson augmentation increases the effective training set size, classification accuracy estimates derived from such small sample counts should be interpreted with appropriate caution. With as few as 5–8 test evaluations per class, a single misclassification corresponds to an error rate of 14–20%, meaning that the reported per-class accuracy values represent coarse estimates of true population performance rather than statistically precise quantities. The results are best interpreted as proof-of-concept demonstrations of pipeline viability rather than definitive performance benchmarks.

Preliminary Analysis: PCA + Clustering and Random Forest

Prior to the primary supervised classification analysis as defined above, a preliminary unsupervised analysis was conducted using PCA combined with K-means clustering. This analysis verified that the sample classes would form geometrically distinct clusters in PCA space without any label information, and also established an interpretable visual baseline for the more complex ML techniques. The unsupervised analysis is methodologically independent of the LOO-CV pipeline described above and uses all available spectra without a train/test split, since no ground-truth labels are used in the clustering step itself.

PCA was applied to the full preprocessed dataset for each detector type, retaining the number of components required to explain at least 60–80% of total variance. K-means clustering was then applied to the PCA score matrix, with the number of clusters set equal to the number of sample types. Cluster assignment quality was evaluated using the Adjusted Rand Index (ARI), which measures the agreement between the K-means cluster assignments and the known ground-truth class labels on a scale of 0 (random) to 1.0 (perfect agreement). This analysis was conducted per detector type.

A preliminary 80/20 stratified train/test split Random Forest classifier was also applied to the dataset using StandardScaler normalization and a 512-channel interpolated energy grid, providing a rapid estimate of classification performance prior to the full LOO-CV pipeline. This preliminary RF analysis used 200 trees and maximum depth 20 with stratified K-fold cross-validation, serving as a sanity check on the overall approach before the more methodologically rigorous LOO-CV framework was applied to the full multi-detector, multi-site dataset. Additionally this approach was used to determine if the more complex preprocessing and ML techniques would be necessary for this dataset.

Statistical Evaluations

The evaluation of classification models relies on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values. Accuracy (eq. 1) measures the overall proportion of correct predictions. But it is an inadequate metric for imbalanced datasets, as majority class predictions disproportionately skew the result. Precision and recall (eq. 2 and eq. 3) provide more granular assessments. Precision quantifies the exactness of positive predictions and is optimized when false positives carry a significant penalty. Conversely, recall, or sensitivity measures the proportion of actual positive instances correctly identified. Recall optimization is critical when false negatives represent a severe risk. To balance these competing objectives, the F1 score (eq. 4) utilizes the harmonic mean of precision and recall. The harmonic mean inherently penalizes disparate values, making the F1 score mathematically robust for imbalanced datasets. However, because this formula excludes True Negatives, it strictly evaluates positive class performance, thereby lacking the comprehensive bidirectional assessment provided by the Matthews Correlation Coefficient [21].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN} \quad (4)$$

331

332

333 The macro-averaged Matthews Correlation Coefficient (MCC_i) is an evaluation
 334 metric for multi-class classification models, see eq. 5. It is calculated by determining the
 335 standard binary MCC for each class independently and then taking the unweighted
 336 arithmetic mean of those individual scores. Because the macro-average treats every class
 337 equally regardless of its frequency in the dataset, it is particularly rigorous when
 338 evaluating models on highly imbalanced data [22].

$$MCC_i = \frac{1}{C} \sum_{i=1}^C \frac{TP \times TN - FP \times FN}{(TP + FP)(TP + FN)(TN + FP)(TN + FN)} \quad (5)$$

339

Results and Discussion

340

Preliminary Analysis Results

341 A preliminary unsupervised and supervised analysis was conducted to determine whether
 342 the full preprocessing and LOO-CV pipeline was warranted, or whether simpler
 343 approaches would suffice for the classification task. The unsupervised stage evaluated
 344 four clustering algorithms on the PCA reduced spectra: K-means, Gaussian Mixture
 345 Models (GMM), DBSCAN, and HDBSCAN. The cluster assignment quality score
 346 against the true label used the Adjusted Rand Index (ARI). The supervised stage applied
 347 an 80/20 stratified-split Random Forest classifier (200 trees, max depth 20) using a
 348 simplified preprocessing pipeline consisting of StandardScaler normalization and a 512-
 349 channel interpolated energy grid, without $\log_{1p}$ compression, total-area normalization, or
 350 Poisson augmentation.

For the single-isotope sources, all clustering algorithms achieved $ARI = 1.000$ on all three detector types using 2 principal components, and the simplified RF reached 100% training, test, and validation accuracy. These results confirm that single-isotope spectra are easily separable, with the photopeak structure dominating the feature space, no sophisticated preprocessing is required when each class is defined by one or two well-resolved emission lines. The HPGe PCA score plot (Figure 6B, 75.4% explained variance) shows six fully isolated clusters with Background separated along PC2 and individual sources distributed across PC1 in an order reflecting their characteristic photopeak energies. CZT produced larger within-cluster scatter due to lower counting statistics from the smaller crystal volume, and HPGe produced the tightest clusters as expected from its superior energy resolution.

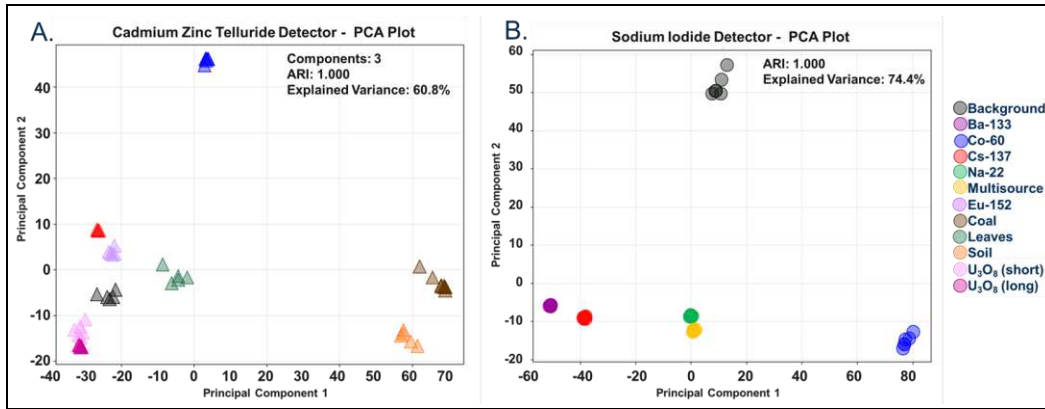


Fig. 6 Preliminary Analysis Clustering Results from NaI and CZT Representative Data

For the reactor-irradiated environmental matrices, however, the preliminary methods exposed several failure modes that motivated the full pipeline. The clustering algorithms diverged in their behavior on this dataset. K-means and GMM reached $ARI = 1.000$ only after the component count was increased to 3 and explained-variance thresholds were relaxed (NaI 78.8%, CZT 60.8%, HPGe 69.7%), and even then the cluster boundaries depended on initialization seed for K-means and on covariance-type selection for GMM. Other clustering techniques were explored, but they were found to classify longer cooling time samples as noise or not sufficiently resolve the sample types. Thus for the more in depth ML analysis, unsupervised techniques alone could not be relied on for classification of the complex irradiated materials.

The simplified supervised RF results reinforced this conclusion. On the irradiated-materials dataset the preliminary RF achieved 96.8% test accuracy with cross-validation accuracy of 96.7% ($\pm 6.24\%$), corresponding to multiple misclassifications under the 80/20 split — predominantly Coal predicted as Soil and the reverse. This pattern is consistent with the compositional similarity of these two geologically related aluminosilicate matrices, which share overlapping trace-element activation product inventories, and was most pronounced on CZT where lower energy resolution prevented separation of the closely spaced activation lines that distinguish them. Misclassifications were also concentrated at longer post-irradiation cooling times, where short-lived activation products had decayed and the residual spectral shape carried less class-discriminating information.

Together, the preliminary analyses established three findings that directly motivated the methodological choices in the full pipeline. First, the unsupervised clustering results showed that activation spectra do not form well-separated compact clusters under simple preprocessing, particularly when post-irradiation cooling-time evolution is present, motivating the $\log_{1p}$ compression and total-area normalization steps that emphasize spectral shape over intensity. Second, the Coal–Soil confusion under simplified preprocessing motivated the move from a single 80/20 split to LOO-CV, since the small dataset size produced highly variable error estimates depending on which spectra fell into the test fold. Third, the cross-validation standard deviation of $\pm 6.24\%$ under the simplified pipeline indicated that the preliminary point estimate of 96.8% carried substantial uncertainty, and that a more rigorous statistical evaluation framework would be necessary. These findings provided the justification for the full preprocessing pipeline, augmentation strategy, and LOO-CV framework described in the Methods section, and confirmed the selection of RF as one of the three primary classifiers in the final analysis.

Single Isotope ML Classification Results

All three classifiers (RF, SVM, and KNN) achieved 100% overall accuracy on the single isotope dataset of ^{133}Ba , ^{60}Co , ^{137}Cs , and ^{22}Na across all three detector types (CZT,

HPGe, NaI). Per-class accuracy was 1.000 for every source class across all models. Because LOO-CV on six original files per class produces six binary trials per class, these point estimates reflect consistent correct classification on every available measured spectrum rather than a statistically precise estimate of population accuracy. Classification performance was also robust to the specific augmentation target within the range tested, confirming that the reported results are not sensitive to that hyperparameter choice.

These results are consistent with physical expectations and serve primarily as a validation of the pipeline before applying it to the activation spectra. Single-isotope sources measured in a controlled laboratory setting on well-calibrated detectors with uniform 30-minute acquisition times produce highly reproducible spectral shapes with prominent, well-resolved photopeaks. The separability of these classes in a 4-component PCA space reflects the fundamental distinctiveness of characteristic gamma emission lines at the energies of ^{133}Ba (356 keV), ^{60}Co (1173, 1332 keV), ^{137}Cs (662 keV), ^{152}Eu (122, 344, 779, 1408 keV), and ^{22}Na (511, 1275 keV). The multi-source configuration, which contains a simultaneous mixture of ^{133}Ba , ^{60}Co , ^{137}Cs , and ^{22}Na , is correctly distinguished from all single-source spectra by all three classifiers at 100% accuracy, demonstrating that the pipeline learns the composite spectral shape of a multi-isotope mixture rather than responding only to the most prominent individual peak.

Activation Analysis ML Classification Results

Classification of the thermally neutron-activated environmental matrices against background presented a substantially more challenging task than the single-isotope sources, owing to the spectral complexity of multi-isotope activation product mixtures, variable post-irradiation cooling times, and the compositional similarity of several natural matrix samples. Despite these challenges, the pipeline achieved high to perfect classification accuracy across all three detector types on the complex irradiated sample subset (Coal, Leaves, Soil, U_3O_8 , $\text{U}_3\text{O}_8_5\text{m}$, and BG), see Figure 7 for F1 scores, and see the Appendix for additional metrics per sample and per detector.



Fig. 7 F1 Scores per Detector

For NaI, all three classifiers achieved perfect classification on the complex irradiated materials (accuracy = 1.000, MCC = 1.000, n = 37). Despite NaI's lower energy resolution, the dominant ^{24}Na lines, ^{40}K feature, and differing Compton continuum profiles of each activated matrix are captured at sufficient fidelity in the log1p-compressed full-spectrum feature vector for perfect separation in PCA space.

For CZT, RF, SVM, and KNN each attained 1.000 overall LOO-CV accuracy and MCC on the complex irradiated samples (n = 37). This is notable given CZT's lower energy resolution. The five activation classes differ primarily in the relative intensities of ^{24}Na (1369, 2754 keV), ^{40}K (1460 keV), and trace element activation lines. Although these features are not sharply resolved on CZT, the full-spectrum preprocessing approach captures their relative intensities as spectral shape features sufficient for perfect separation.

The HPGe Results for RF and KNN achieved perfect accuracy on the complex irradiated samples (accuracy = 1.000, MCC = 1.000, n = 38), while SVM reached 0.974 (MCC = 0.969). The single SVM misclassification involved a tomato leaf spectrum acquired near 2,000 minutes post-irradiation, where the short-lived ^{24}Na component had largely decayed and the spectral shape increasingly resembled background [23]. The SVM margin-based boundary placed this spectrum on the wrong side of the Background/Activation boundary, while RF and KNN were more tolerant of this gradual spectral drift. This suggests that for activation spectra with significant cooling-time-dependent evolution, ensemble and neighbor-based methods are preferable to margin-based classifiers.

RF and KNN achieved 1.000 accuracy on every detector, with SVM degrading only on HPGe due to the cooling-time spectral evolution of the Leaves class. The consistency of RF and KNN across all three detectors confirms that ensemble and neighbor-based methods are the more robust choice for activation classification tasks where spectral shape evolves temporally within a class.

Reliability and Statistical Considerations

95% Confidence Intervals (CIs) (Table 3) and classification accuracy were computed using the Clopper-Pearson exact binomial method, which is appropriate for small samples where the normal approximation breaks down. On CZT and NaI, all classifiers achieved perfect observed accuracy (1.000) with a CI lower bound of 0.905, so true population accuracy could plausibly be as low as ~90%. RF and KNN on HPGe matched this (CI [0.908, 1.000]). The one exception was HPGe SVM at 0.974 (CI [0.862, 0.999]), driven by a single tomato leaf misclassification at ~2,000 min cooling time. Even its lower bound of 0.862 reflects strong performance on a physically difficult task. These intervals confirm that irradiated sample discrimination is robust across all three detectors within the limits of the available data; larger datasets would tighten the bounds and yield more precise population-level estimates.

Table 3 Aggregate LOO-CV Classification Metrics for Background and Irradiated Samples. 95% confidence intervals computed using the Clopper-Pearson exact binomial method.

Detector	Classifier	n	Accuracy	95% CI	MacroF1	MCC
CZT	RF	37	1	[0.905, 1.000]	1	1
CZT	SVM	37	1	[0.905, 1.000]	1	1
CZT	KNN	37	1	[0.905, 1.000]	1	1

				1.000]		
HPGe	RF	38	1	[0.908, 1.000]	1	1
HPGe	SVM	38	0.974	[0.862, 0.999]	0.97	0.969
HPGe	KNN	38	1	[0.908, 1.000]	1	1
NaI	RF	37	1	[0.905, 1.000]	1	1
NaI	SVM	37	1	[0.905, 1.000]	1	1
NaI	KNN	37	1	[0.905, 1.000]	1	1

475

476

477 The confidence metrics indicate fragility within the systems that cannot be

478 explicitly expressed through accuracy. The CZT U₃O₈ 5 minute group classified correctly479 but at RF mean confidence 0.642 ± 0.092 (min 0.548). That is a correct but unreliable

480 prediction: in a deployed system a small perturbation could flip it. Accuracy alone would

481 never surface this; the confidence column does. This indicates that observed accuracy and

482 prediction reliability diverge for low statistic measurements, as seen with the small CZT

483 semiconductor crystal, it has much lower counting efficiency than a scintillator like NaI,

484 and with a low count time of 5 minutes it results in poor counting statistics.

485 **Table 4** Mean Prediction Coefficients per Sample and Detector for RF and SVM

486 Classifiers. Values reported as mean +/- standard deviation across LOO-CV folds.

487 Accuracy (Acc) column shows per class LOO-CV accuracy for reference.

Detector	Class	RF Mean Conf $\pm$ Std	RF Min	RF Max	RF Acc	SVM Mean Conf $\pm$ Std	SVM Min	SVM Max	SVM Acc
CZT	BG	0.737 ± 0.056	0.668	0.826	1	0.884 ± 0.060	0.776	0.94	1
CZT	Coal	$0.992 \pm$	0.961	1	1	$0.927 \pm$	0.925	0.929	1

		0.013				0.002			
CZT	Leaves	0.864 ± 0.076	0.776	0.971	1	0.908 ± 0.019	0.87	0.929	1
CZT	Soil	0.983 ± 0.019	0.957	1	1	0.913 ± 0.001	0.911	0.915	1
CZT	U ₃ O ₈	0.951 ± 0.068	0.802	1	1	0.899 ± 0.004	0.894	0.904	1
CZT	U ₃ O ₈ _5m	0.642 ± 0.092	0.548	0.762	1	0.914 ± 0.013	0.892	0.929	1
HPGe	BG	0.876 ± 0.083	0.72	0.944	1	0.515 ± 0.017	0.492	0.538	1
HPGe	Coal	0.999 ± 0.003	0.992	1	1	0.904 ± 0.008	0.886	0.913	1
HPGe	Leaves	0.891 ± 0.124	0.69	0.993	0.833	0.464 ± 0.019	0.438	0.493	0.833
HPGe	Soil	0.999 ± 0.001	0.997	1	1	0.944 ± 0.003	0.938	0.947	1
HPGe	U ₃ O ₈	1.000 ± 0.000	1	1	1	0.898 ± 0.002	0.895	0.901	1
HPGe	U ₃ O ₈ _5m	0.953 ± 0.066	0.859	1	1	0.866 ± 0.030	0.803	0.89	1
NaI	BG	1.000 ± 0.000	1	1	1	0.914 ± 0.002	0.912	0.917	1
NaI	Coal	0.980 ± 0.025	0.92	1	1	0.918 ± 0.015	0.892	0.936	1
NaI	Leaves	0.960 ± 0.089	0.76	1	1	0.909 ± 0.003	0.903	0.913	1
NaI	Soil	0.993 ± 0.010	0.974	1	1	0.924 ± 0.003	0.92	0.928	1
NaI	U ₃ O ₈	0.989 ± 0.010	0.974	1	1	0.885 ± 0.007	0.88	0.902	1
NaI	U ₃ O ₈ _5m	0.999 ± 0.002	0.995	1	1	0.863 ± 0.027	0.82	0.89	1

LOO-CV can slightly inflate apparent performance because augmented copies of the held-out class's neighbors remain in training; PCA is correctly fit inside the fold, but with only 4–13 components and ~30 augmented samples, the components are estimated from few independent measurements; and all "augmentation" derives from the same handful of physical files, so it expands effective sample size for variance estimation but cannot introduce new physical diversity (cooling time or additional unique samples).

The small sample sizes found in this dataset require additional analyses to answer if these results are reliable. With 5-8 spectra per class, a single misclassification is a 14-20% error. When something provides 100% accuracy that indicates zero errors out of ~37 trials, this can be better reported as a ML technique resulting in 'no observed failures' rather than 100% accuracy. With this, it is important to have confidence that the positive scores seen from these results are real and not artifacts. For the decision trees using a majority vote in RF and a plurality vote among neighbors in KNN, these techniques are non-parametric and make few distributional assumptions, so the fact that they both result in no observed failures across all three detectors is difficult to get by chance. Alternatively, with SVM and the RBF kernel, this technique uses a margin based boundary that is sensitive to points drifting towards another class. An example of this was seen for the tomato leaf sample with a long cooling time that was misclassified as background. This error can be explained by how SVM works and is not random. These three independent algorithms converging on a solution support the claim that these results are reliable.

Conclusion

This work presented a complete machine learning pipeline for the automated classification of experimental gamma-ray spectra, evaluated across two levels of spectral complexity and three detector types (CZT, HPGe, and NaI(Tl)). A five-step preprocessing chain allowed the spectra to be analyzed independent of their collection time. Preprocessing included counts-per-second normalization, interpolation onto a common 3001-point energy grid, total-area normalization, $\log_{10}$ compression, and Poisson noise augmentation applied to raw counts. Following PCA dimensionality

reduction, three classifiers (RF, SVM, and KNN) were evaluated under Leave-One-Out Cross-Validation on both single-isotope sources and thermal neutron-activated samples. For single-isotope sources, all three classifiers achieved perfect observed accuracy (no observed failures) on every detector. For the more complex irradiated materials, RF and KNN achieved perfect accuracy and Matthews Correlation Coefficient on all three detectors, while SVM matched this on CZT and NaI and reached 0.974 on HPGe. The single SVM error was a tomato leaf spectrum measured near 2,000 minutes post-irradiation, where decay of the short-lived ^{24}Na component left a spectral shape resembling background. This failure is consistent with the margin-based nature of the SVM decision boundary and indicates that ensemble and neighbor-based classifiers are the more robust choice for activation spectra whose shape evolves with cooling time.

These results should be interpreted as a validated proof-of-concept rather than a definitive performance benchmark. With only 5–8 measured spectra per class, the perfect accuracy values represent zero observed errors over a small number of trials, not statistically precise population estimates. Because the three ML approaches achieved similar results, the observed performance reflects genuine separation of sample classes rather than an artifact of the small dataset. The confidence analysis revealed that even with high to perfect prediction accuracy, the confidence may be lower. The CZT uranium oxide samples counted for only 5 minutes were classified correctly but at low confidence (RF mean 0.642), demonstrating that short counting times on small-volume detectors may yield correct predictions, but they are not stable. A confidence threshold near 0.70 is therefore recommended as a practical quality-control criterion for any deployed system.

The present pipeline was trained and evaluated within each detector type independently, so each classifier was trained exclusively on spectra from the same detector it was tested on. Training on one detector type and predicting another was not evaluated in this study and represents an important open question. Physical intuition suggests that the total area normalized, $\log_{1p}$ compressed full spectrum representation may carry detector shape to support some amount of transferability between predictions, especially for detectors of similar resolution class (e.g., CZT to NaI). However, the substantially different energy resolution, channel width, and Compton continuum profiles between HPGe and scintillator detectors would likely require either domain adaptation

techniques, detector-response normalization, or retraining on mixed detector datasets to achieve reliable cross-detector performance. This is a natural and important extension of the present work and is planned as a future study.

The pipeline establishes a generalizable preprocessing and classification baseline for experimental gamma spectroscopy datasets with variable channel counts, acquisition times, and post-irradiation histories. Future work will address cross-detector training, expand the dataset to narrow the confidence intervals, expand the dataset through real or simulated spectra, and extend the framework toward fission product identification.

Acknowledgements

Thank you to the staff of the University of Texas at Austin Nuclear Engineering Teaching Laboratory for facilitating access to the TRIGA Mark II research reactor and supporting this irradiation and counting campaign.

Thank you also to Dr. Sharmistha Mukhopadhyay at the Georgia Institute of Technology for her help in setting up and running HPGe experiments in the Radiological Science and Engineering Laboratory.

Funding Statement

This material is based upon work supported by the Department of Energy / National Nuclear Security Administration under Award Number(s) DE-NA0004196.

References

- [1] E. T. Journey, P. J. Bendt, and T. R. England, "Fission product gamma spectra," Los Alamos Scientific Lab., NM (USA), LA-7620-MS, Jan. 1979. doi: 10.2172/6414413.
- [2] J. R. Curtis, "Special Nuclear Material Classification and Mass Content Estimation Using Temporal Gamma-Ray Spectroscopy and Machine Learning Methods."
- [3] R. S. Williford, "Temporal gamma-ray spectrometry to quantify relative fissile material content."

- 574 [4] S. M. Galib, P. K. Bhowmik, A. V. Avachat, and H. K. Lee, “A comparative
575 study of machine learning methods for automated identification of radioisotopes using
576 NaI gamma-ray spectra,” *Nucl. Eng. Technol.*, vol. 53, no. 12, pp. 4072–4079, Dec. 2021,
577 doi: 10.1016/j.net.2021.06.020.
- 578 [5] J. Mathew *et al.*, “A comparison of machine learning methods to classify
579 radioactive elements using prompt-gamma-ray neutron activation data,” *Sci. Rep.*, vol.
580 13, no. 1, p. 9948, Jun. 2023, doi: 10.1038/s41598-023-36832-8.
- 581 [6] P. Olmos, J. C. Diaz, J. M. Perez, P. Aguayo, P. Gomez, and V. Rodellar, “Drift
582 problems in the automatic analysis of gamma-ray spectra using associative memory
583 algorithms,” *IEEE Trans. Nucl. Sci. Inst. Electr. Electron. Eng. U. S.*, vol. 41:3 Pt2, Jun.
584 1994, doi: 10.1109/23.299814.
- 585 [7] M. Kamuda, J. Zhao, and K. Huff, “A comparison of machine learning methods
586 for automated gamma-ray spectroscopy,” *Nucl. Instrum. Methods Phys. Res. Sect. Accel.*
587 *Spectrometers Detect. Assoc. Equip.*, vol. 954, p. 161385, Feb. 2020, doi:
588 10.1016/j.nima.2018.10.063.
- 589 [8] Y. Sun, F. Tuo, W. Lin, Q. Zhou, and B. Yang, “Machine learning application in
590 NaI(Tl) gamma-ray spectroscopy for radionuclide identification: A systematic review,”
591 *Radiat. Med. Prot.*, vol. 06, no. 05, pp. 251–258, Oct. 2025, doi:
592 10.1016/j.radmp.2025.09.009.
- 593 [9] K. J. Bilton, T. H. Y. Joshi, M. S. Bandstra, J. C. Curtis, D. Hellfeld, and K.
594 Vetter, “Neural Network Approaches for Mobile Spectroscopic Gamma-Ray Source
595 Detection,” *J. Nucl. Eng.*, vol. 2, no. 2, pp. 190–206, Jun. 2021, doi: 10.3390/jne2020018.
- 596 [10] M. Zehtabvar, K. Taghandiki, N. Madani, D. Sardari, and B. Bashiri, “A Review
597 on the Application of Machine Learning in Gamma Spectroscopy: Challenges and
598 Opportunities,” *Spectrosc. J.*, vol. 2, no. 3, pp. 123–144, Sep. 2024, doi:
599 10.3390/spectroscj2030008.
- 600 [11] P. J. O’Neal, S. P. Martinson, and S. S. Chirayath, “Validation of a Machine
601 Learning–Based Nuclear Forensics Methodology for the Discrimination of a Chemically
602 Separated Plutonium Sample from Low-Enriched Uranium,” *Nuclear Science and*
603 *Engineering*, vol. 198, no. No. 9, pp. 1817–1829, Jul. 2024, doi: doi:
604 10.3390/spectroscj2030008.

- [12] L. Lee-Brewin, R. Holden, and C. Shenton-Taylor, "Validation of simulated training sets using a convolutional neural network for isotope identification in urban environments," *PLOS ONE*, vol. 20, no. 6, p. e0323053, Jun. 2025, doi: 10.1371/journal.pone.0323053.
- [13] NRC, "University of Texas at Austin Nuclear Engineering Teaching Laboratory TRIGA Research Reactor." License No. R-129 Docket No. 50-602, Austin, TX, Aug. 2023.
- [14] M. S. Bandstra, J. C. Curtis, J. M. G. Jr, A. C. Jones, and T. H. Y. Joshi, "Explaining machine-learning models for gamma-ray detection and identification," *PLOS ONE*, vol. 18, no. 6, p. e0286829, Jun. 2023, doi: 10.1371/journal.pone.0286829.
- [15] M. Gomez-Fernandez *et al.*, "Isotope identification using deep learning: An explanation," *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.*, vol. 988, p. 164925, Feb. 2021, doi: 10.1016/j.nima.2020.164925.
- [16] L. Lee-Brewin, R. Holden, and C. Shenton-Taylor, "Validation of simulated training sets using a convolutional neural network for isotope identification in urban environments," *PLOS ONE*, vol. 20, no. 6, p. e0323053, Jun. 2025, doi: 10.1371/journal.pone.0323053.
- [17] A. Gracia Moisés, I. Vitoria Pascual, J. J. Imas González, and C. Ruiz Zamarreño, "Data Augmentation Techniques for Machine Learning Applied to Optical Spectroscopy Datasets in Agrifood Applications: A Comprehensive Review," *Sensors*, vol. 23, no. 20, p. 8562, Jan. 2023, doi: 10.3390/s23208562.
- [18] J. Katakura and T. R. England, "Augmentation of ENDF/B fission product gamma-ray spectra by calculated spectra," Art. no. LA--12125-MS, Nov. 1991, Accessed: May 18, 2026. [Online]. Available: <https://inis.iaea.org/records/mb8w5-9fv93>
- [19] C. K. Wang, *Atoms, Nuclei, and Interactions of Ionizing Radiation with Matter*. Atlanta, GA: Georgia Institute of Technology, 2016.
- [20] G. Knoll, *Radiation Detection and Measurement*, 4th ed. Ann Arbor, Michigan: John Wiley & Sons, Inc., 2010.
- [21] M. Kuhn and K. Johnson, "Applied Predictive Modeling | Springer Nature Link." Accessed: May 18, 2026. [Online]. Available: <https://link.springer.com/book/10.1007/978-1-4614-6849-3>

- 636 [22] D. Chicco and G. Jurman, “The advantages of the Matthews correlation
 637 coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC*
 638 *Genomics*, vol. 21, no. 1, p. 6, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
- 639 [23] R. E. Abdel-Aal and M. N. Al-Haddad, “Determination of radioisotopes in
 640 gamma-ray spectroscopy using abductive machine learning,” *Nucl. Instrum. Methods*
 641 *Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.*, vol. 391, no. 2, pp. 275–288,
 642 Jun. 1997, doi: 10.1016/S0168-9002(97)00391-4.

643 Appendix

644 *Table 5. ML Per-Class Metrics*

Detector	Classifier	Class	Category	n	Sparse	Recall	Precision	F1
CZT	RF	BG	Background	5	FALSE	1	1	1
CZT	RF	Co60	PointSource	6	FALSE	1	1	1
CZT	RF	Coal	Activation	8	FALSE	1	1	1
CZT	RF	Cs137	PointSource	7	FALSE	1	1	1
CZT	RF	Eu152	PointSource	6	FALSE	1	1	1
CZT	RF	Leaves	Activation	6	FALSE	1	1	1
CZT	RF	Soil	Activation	5	FALSE	1	1	1
CZT	RF	U3O8	Activation	7	FALSE	1	1	1
CZT	RF	U3O8_5m	Activation	6	FALSE	1	1	1
CZT	SVM	BG	Background	5	FALSE	1	1	1
CZT	SVM	Co60	PointSource	6	FALSE	1	1	1
CZT	SVM	Coal	Activation	8	FALSE	1	1	1
CZT	SVM	Cs137	PointSource	7	FALSE	1	1	1
CZT	SVM	Eu152	PointSource	6	FALSE	1	1	1

			e		E			
CZT	SVM	Leaves	Activation	6	FALS E	1	1	1
CZT	SVM	Soil	Activation	5	FALS E	1	1	1
CZT	SVM	U3O8	Activation	7	FALS E	1	1	1
CZT	SVM	U3O8_5 m	Activation	6	FALS E	1	1	1
CZT	KNN	BG	Backgroun d	5	FALS E	1	1	1
CZT	KNN	Co60	PointSourc e	6	FALS E	1	1	1
CZT	KNN	Coal	Activation	8	FALS E	1	1	1
CZT	KNN	Cs137	PointSourc e	7	FALS E	1	1	1
CZT	KNN	Eu152	PointSourc e	6	FALS E	1	1	1
CZT	KNN	Leaves	Activation	6	FALS E	1	1	1
CZT	KNN	Soil	Activation	5	FALS E	1	1	1
CZT	KNN	U3O8	Activation	7	FALS E	1	1	1
CZT	KNN	U3O8_5 m	Activation	6	FALS E	1	1	1
HPGe	RF	BG	Backgroun d	5	FALS E	1	0.8333	0.909 1
HPGe	RF	Co60	PointSourc e	1	TRUE	0		
HPGe	RF	Coal	Activation	8	FALS E	1	1	1
HPGe	RF	Cs137	PointSourc e	2	TRUE	1	1	1
HPGe	RF	Eu152	PointSourc e	1	TRUE	0		
HPGe	RF	Leaves	Activation	6	FALS E	1	0.8571	0.923 1
HPGe	RF	Soil	Activation	6	FALS E	1	1	1
HPGe	RF	U3O8	Activation	7	FALS E	1	1	1
HPGe	RF	U3O8_5 m	Activation	6	FALS E	1	1	1

HPGe	SVM	BG	Background	5	FALSE	1	0.8333	0.9091
HPGe	SVM	Co60	PointSource	1	TRUE	0		
HPGe	SVM	Coal	Activation	8	FALSE	1	1	1
HPGe	SVM	Cs137	PointSource	2	TRUE	1	1	1
HPGe	SVM	Eu152	PointSource	1	TRUE	0		
HPGe	SVM	Leaves	Activation	6	FALSE	0.8333	0.7143	0.7692
HPGe	SVM	Soil	Activation	6	FALSE	1	1	1
HPGe	SVM	U3O8	Activation	7	FALSE	1	1	1
HPGe	SVM	U3O8_5m	Activation	6	FALSE	1	1	1
HPGe	KNN	BG	Background	5	FALSE	1	1	1
HPGe	KNN	Co60	PointSource	1	TRUE	0		
HPGe	KNN	Coal	Activation	8	FALSE	1	1	1
HPGe	KNN	Cs137	PointSource	2	TRUE	1	1	1
HPGe	KNN	Eu152	PointSource	1	TRUE	0		
HPGe	KNN	Leaves	Activation	6	FALSE	1	0.75	0.8571
HPGe	KNN	Soil	Activation	6	FALSE	1	1	1
HPGe	KNN	U3O8	Activation	7	FALSE	1	1	1
HPGe	KNN	U3O8_5m	Activation	6	FALSE	1	1	1
NaI	RF	BG	Background	5	FALSE	1	1	1
NaI	RF	Co60	PointSource	6	FALSE	1	1	1
NaI	RF	Coal	Activation	8	FALSE	1	0.8889	0.9412
NaI	RF	Cs137	PointSource	7	FALSE	0.8571	1	0.9231
NaI	RF	Eu152	PointSource	6	FALSE	1	1	1

NaI	RF	Leaves	Activation	6	FALS E	1	1	1
NaI	RF	Soil	Activation	5	FALS E	1	1	1
NaI	RF	U3O8	Activation	7	FALS E	1	1	1
NaI	RF	U3O8_5 m	Activation	6	FALS E	1	1	1
NaI	SVM	BG	Backgroun d	5	FALS E	1	1	1
NaI	SVM	Co60	PointSourc e	6	FALS E	1	1	1
NaI	SVM	Coal	Activation	8	FALS E	1	1	1
NaI	SVM	Cs137	PointSourc e	7	FALS E	0.857 1	1	0.923 1
NaI	SVM	Eu152	PointSourc e	6	FALS E	1	1	1
NaI	SVM	Leaves	Activation	6	FALS E	1	1	1
NaI	SVM	Soil	Activation	5	FALS E	1	0.8333	0.909 1
NaI	SVM	U3O8	Activation	7	FALS E	1	1	1
NaI	SVM	U3O8_5 m	Activation	6	FALS E	1	1	1
NaI	KNN	BG	Backgroun d	5	FALS E	1	1	1
NaI	KNN	Co60	PointSourc e	6	FALS E	1	1	1
NaI	KNN	Coal	Activation	8	FALS E	1	1	1
NaI	KNN	Cs137	PointSourc e	7	FALS E	0.857 1	1	0.923 1
NaI	KNN	Eu152	PointSourc e	6	FALS E	1	1	1
NaI	KNN	Leaves	Activation	6	FALS E	1	1	1
NaI	KNN	Soil	Activation	5	FALS E	1	0.8333	0.909 1
NaI	KNN	U3O8	Activation	7	FALS E	1	1	1
NaI	KNN	U3O8_5 m	Activation	6	FALS E	1	1	1

