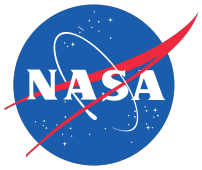
NASA/TM–20250011359



Interpretable Machine Learning for Acoustic Classification of Incipient Boiling Regimes

Krishnanshu Gupta
California Polytechnic State University, San Luis Obispo, California

James Lamkin
California Polytechnic State University, San Luis Obispo, California

Andrew Martinez
California Polytechnic State University, San Luis Obispo, California

Zachary Weinfeld
California Polytechnic State University, San Luis Obispo, California

Kelly Bodwin
California Polytechnic State University, San Luis Obispo, California

Alex Dekhtyar
California Polytechnic State University, San Luis Obispo, California

Michael Khasin
Ames Research Center, Moffett Field, California

December 2025

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

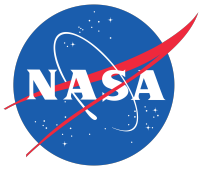
- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.
- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI Program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*
- E-mail your question to *help@sti.nasa.gov*
- Fax your question to the NASA STI Information Desk at 443-757-5803
- Phone the NASA STI Information Desk at 443-757-5802
- Write to:
  STI Information Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076–1320

NASA/TM–20250011359



Interpretable Machine Learning for Acoustic Classification of Incipient Boiling Regimes

Krishnanshu Gupta
California Polytechnic State University, San Luis Obispo, California

James Lamkin
California Polytechnic State University, San Luis Obispo, California

Andrew Martinez
California Polytechnic State University, San Luis Obispo, California

Zachary Weinfeld
California Polytechnic State University, San Luis Obispo, California

Kelly Bodwin
California Polytechnic State University, San Luis Obispo, California

Alex Dekhtyar
California Polytechnic State University, San Luis Obispo, California

Michael Khasin
Ames Research Center, Moffett Field, California

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

December 2025

Acknowledgments

Funding for this effort was provided by the NASA Ames Research Center FY24 Center Innovation Fund.

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320

National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Available electronically at http://www.sti.nasa.gov

Interpretable Machine Learning for Acoustic Classification of Incipient Boiling Regimes

Krishnanshu Gupta
Computer Science Dept.
Cal Poly, SLO
San Luis Obispo, CA
kgupta15@calpoly.edu

James Lamkin
Computer Science Dept.
Cal Poly, SLO
San Luis Obispo, CA
lamkin@calpoly.edu

Andrew Martinez
Statistics Dept.
Cal Poly, SLO
San Luis Obispo, CA
amart531@calpoly.edu

Zachary Weinfeld
Mathematics Dept.
Cal Poly, SLO
San Luis Obispo, CA
zweinfel@calpoly.edu

Kelly Bodwin
Computer Science Dept.
Cal Poly, SLO
San Luis Obispo, CA
kbodwin@calpoly.edu

Alex Dekhtyar
Statistics Dept.
Cal Poly, SLO
San Luis Obispo, CA
dekhtyar@calpoly.edu

Michael Khasin
Ames Research Center
NASA Ames
Moffett Field, CA
michael.khasin@nasa.gov

Summary

Acoustic emissions provide a non-intrusive window into the mechanisms and regimes of incipient boiling triggered by localized heat leaks—a phenomenon relevant to Cryogenic Fuel Management (CFM). This Technical Memorandum (TM) is an explicit machine-learning (ML) companion to NASA/TM–20250009668, which establishes the experimental foundation and physics-based interpretation of accelerometer signals synchronized with high-speed video in a surrogate benchtop setup. Here we focus on operationalizing that regime understanding into an interpretable classification workflow using 441 short-duration boiling runs labeled by human annotators. We extract time- and frequency-domain features designed to capture event rate, rhythmic structure, and spectral content, and we use (i) unsupervised clustering for regime discovery and (ii) a decision-tree classifier selected for transparency and auditability. We also provide a web-based application that reproduces the same feature pipeline for interactive exploration and consistent classification of new runs. The resulting framework supports scientifically grounded, interpretable mapping from acoustic signatures to boiling regimes, complementing the physics-first narrative and enabling future physics-informed ML diagnostics.

1 Introduction

Incipient boiling at localized heat leaks can generate strong, structured mechanical disturbances that couple into a tank wall. If these disturbances contain regime-specific information, then acoustic emissions measured by externally mounted accelerometers may serve as a scientific diagnostic of boiling mechanisms, not merely a threshold detector. Earlier work (Ref. 1) establishes the experimental foundation and a physics-based interpretation of such signals using accelerometer measurements synchronized with high-speed video in a surrogate benchtop configuration.

A key implication of earlier work (Ref. 1) is that interpretability is a scientific requirement: without a clear link between acoustic signatures and physical processes, data-driven models become difficult to trust, difficult to debug, and difficult to use for hypothesis-building. This TM adopts that constraint explicitly. Our objective is an interpretable, feature-transparent mapping that operationalizes regime concepts while remaining consistent with the physics-first narrative of (Ref. 1). The companion TM (Ref. 1) reports two repeatable boiling regimes in this experiment: a stochastic *random* regime and a periodic *rhythmic* regime. In this ML-focused companion, we keep that two-regime interpretation as the top-level taxonomy, but use a finer set of annotation labels for training and diagnostics. Specifically, *Single Rhythmic*, *Double Rhythmic*, and *Rhythmic Climax* are treated as subtypes of the companion TM's *rhythmic* regime; *Random* corresponds to the companion TM's *random* regime; *Rhythmic & Random* denotes mixed or transitional behavior; and *Noise* denotes runs (or windows) in which the current preprocessing does not yield confident boiling-event structure rather than a distinct physical boiling mechanism.

Accordingly, this memorandum develops an ML workflow using 441 short runs with regime labels assigned by human annotators using video, acoustic time series, and experimental metadata. We (i) engineer time- and frequency-domain features intended to capture event rate, rhythmicity, and spectral structure, (ii) apply unsupervised clustering as an exploratory check on regime separability in feature space, and (iii) train a decision-tree classifier chosen specifically for transparency and auditability. In particular, the rhythm-detection features are designed to be sensitive to the recovery-controlled periodicity reported in the companion TM (cycle time $\tau \approx 7$ s, i.e., minimum frequency ~ 0.15 Hz). Finally, we package the workflow into a web application for interactive exploration and consistent classification of newly uploaded runs using the same feature pipeline. In this way, ML is positioned as a complementary, interpretable layer that helps translate mechanistic regime understanding into reproducible analysis tools.

2 Dataset Description

This study utilizes data (Ref. 2) from 441 boiling experiments, each lasting between 1 and 30 seconds. The experiments were conducted using a surrogate spherical vessel (water) used to emulate localized heat leaks relevant to cryogenic tanks, instrumented with acoustic sensors under controlled laboratory conditions.

2.1 Experimental Setup

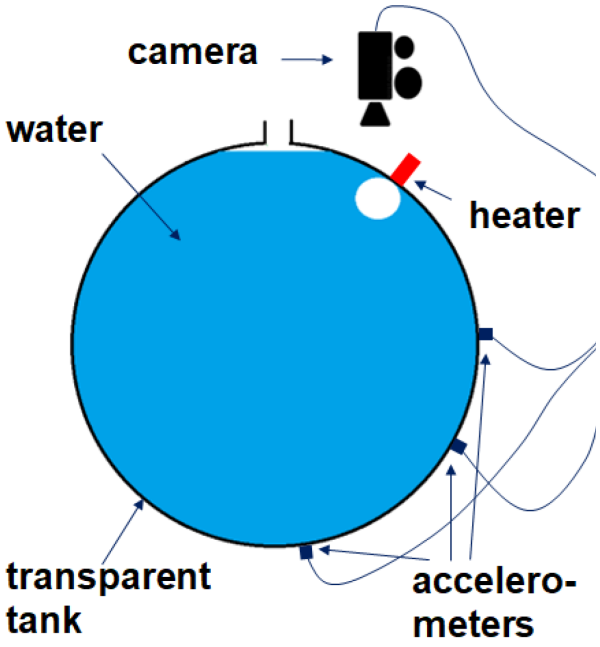


Figure 1.—Experimental setup: surrogate spherical vessel (water) used to emulate localized heat leaks relevant to cryogenic tanks, with externally mounted accelerometers.

Each experiment consists of a spherical tank filled with liquid and equipped with accelerometers mounted on the tank walls, as illustrated in Figure 1. During each run, a localized heat source induces boiling near the wall surface, and acoustic emissions are recorded via the accelerometers. Although two sensor channels are available, the analysis presented here focuses on a single channel due to signal redundancy. The sampling rate for each sensor is 10,000 Hz.

2.2 Data Format and Visualization

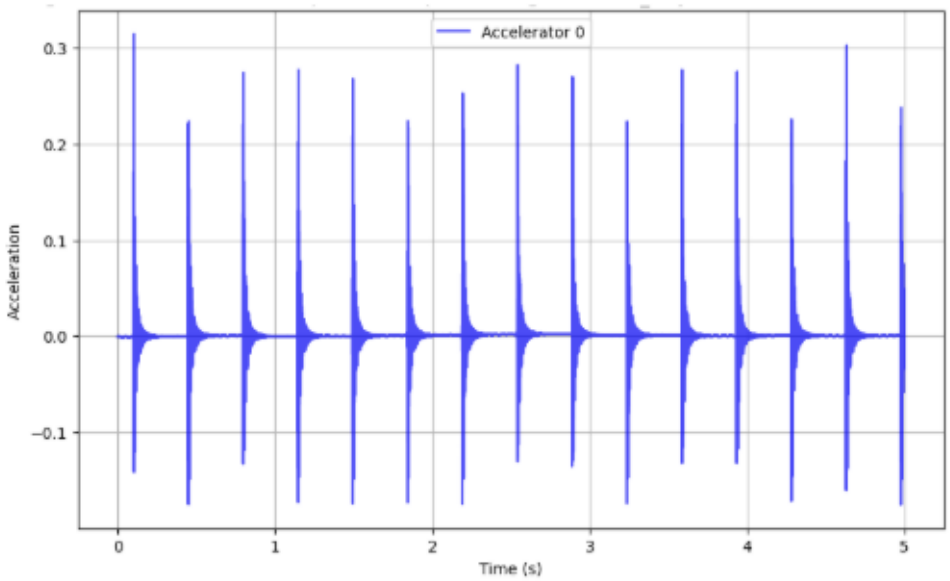


Figure 2.—Raw time-domain signal from a representative boiling experiment.

Each run is stored in a comma-separated values (CSV) file containing time-stamped accelerometer data. The raw time-domain signal is visualized in Figure 2 as amplitude versus time, where prominent peaks correspond to bubble nucleation and collapse events.

2.3 Physical Interpretation

Acoustic emissions in these experiments arise from rapid, localized fluid–structure loading associated with bubble nucleation, growth, detachment, and subsequent fluid motion near the heated wall. In the companion TM (Ref. 1), two repeatable boiling regimes were observed and interpreted: *random* boiling, characterized by irregular nucleation timing and more variable acoustic amplitudes, and *rhythmic* boiling, characterized by regular periodicity with comparatively uniform acoustic pulses. A key physical interpretation in that work is that the rhythmic regime is governed by a thermal recovery (reheating) cycle that sets a characteristic timescale τ (reported as $\tau \approx 7$ s, corresponding to a minimum frequency ~ 0.15 Hz) (Ref. 1). The feature set used here is designed to operationalize these distinctions in a transparent way: time-domain peak statistics capture event rate and regularity, while frequency-domain summaries capture whether energy concentrates near a dominant periodic component.

Although reduced-gravity environments can alter buoyancy-driven convection, bubble departure, and heat transfer pathways, the present dataset is collected in a 1-g surrogate setup (water) and is used here to build an interpretable mapping between measured acoustic structure and regime labels grounded in synchronized video. Any claim that microgravity will necessarily make boiling "more violent" or "more structured" is too strong without dedicated microgravity data; instead, the appropriate statement is that the mechanisms and regime concepts are *motivated by* scenarios relevant to Cryogenic Fuel Management (CFM) and that the mapping should be revalidated and, if needed, retrained for the target environment (Ref. 3, 4).

3 Feature Engineering

Features are extracted from both time and frequency domains to capture multiple dimensions of the boiling signal and enable characterization of different boiling regimes.

3.1 Time Domain Features

3.1.1 Peak Detection

SciPy's `find_peaks` is used with adaptive thresholds. The vertical threshold is based on the greater of the 99.5th percentile or 10% of the max value, clamped between 0.015 and 0.1. The horizontal threshold scales inversely with peak amplitude to avoid false detections.

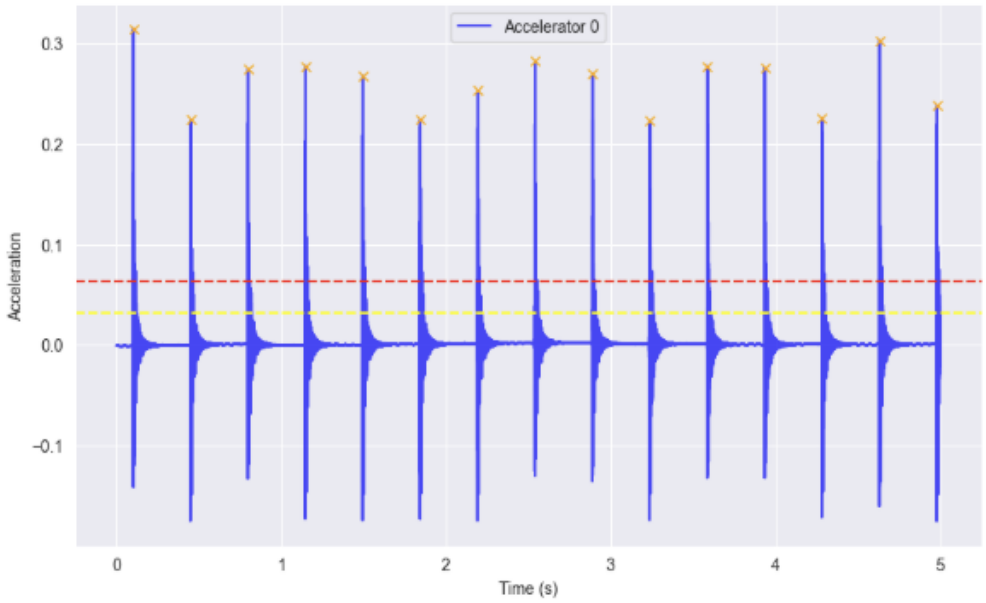


Figure 3.—Detected peaks using adaptive thresholding.

3.1.2 Derived Peak Features

From each run, the following are calculated:

- Number of statistically significant dominant rhythms
- Mean, median, and standard deviation of peak-to-peak intervals
- Peak magnitude statistics (max, median, std)
- Peaks per second
- Sum of peak magnitudes per second

The algorithm for detecting the number of statistically significant dominant rhythms is detailed in Appendix A.

3.1.3 Additional Metrics

The following metrics are included:

- Percent of time above threshold
- Global signal standard deviation

3.2 Frequency Domain Features

3.2.1 Spectral Transformation

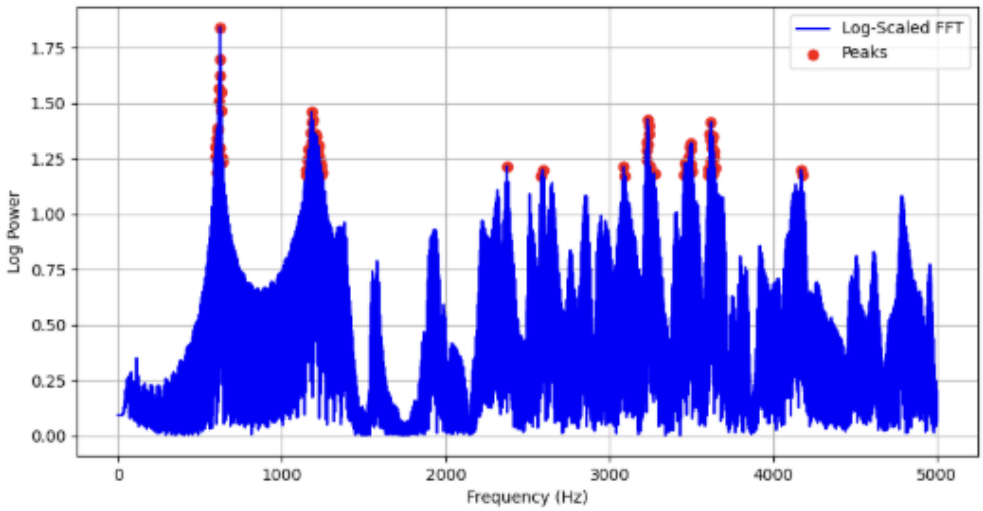


Figure 4.—FFT of filtered accelerometer signal.

A high-pass Butterworth filter is applied, followed by FFT and Welch's method to compute spectral features, as seen in Figure 4.

3.2.2 Extracted Features

- Dominant frequency
- Number of spectral peaks
- Weighted average frequency
- Mean and standard deviation of spectral power
- Spectral entropy

4 Unsupervised Learning

4.1 Pipeline

An unsupervised learning pipeline is implemented to analyze acoustic boiling regime data. The pipeline consists of the following steps:

- **Feature Standardization:** All extracted features are standardized using `StandardScaler` to ensure each feature contributes equally to the analysis.

- **Dimensionality Reduction:** Principal Component Analysis (PCA) is applied to the standardized features. The first 10 principal components are retained, capturing approximately 97% of the total variance in the dataset. This reduces noise and highlights the most informative patterns. *(See Figure 5 for the scree plot, which shows the cumulative variance explained by varying numbers of principal components.)*

- **Clustering:** KMeans clustering is performed in the reduced PCA space, enabling the identification of natural groupings and patterns within the data without relying on labels.

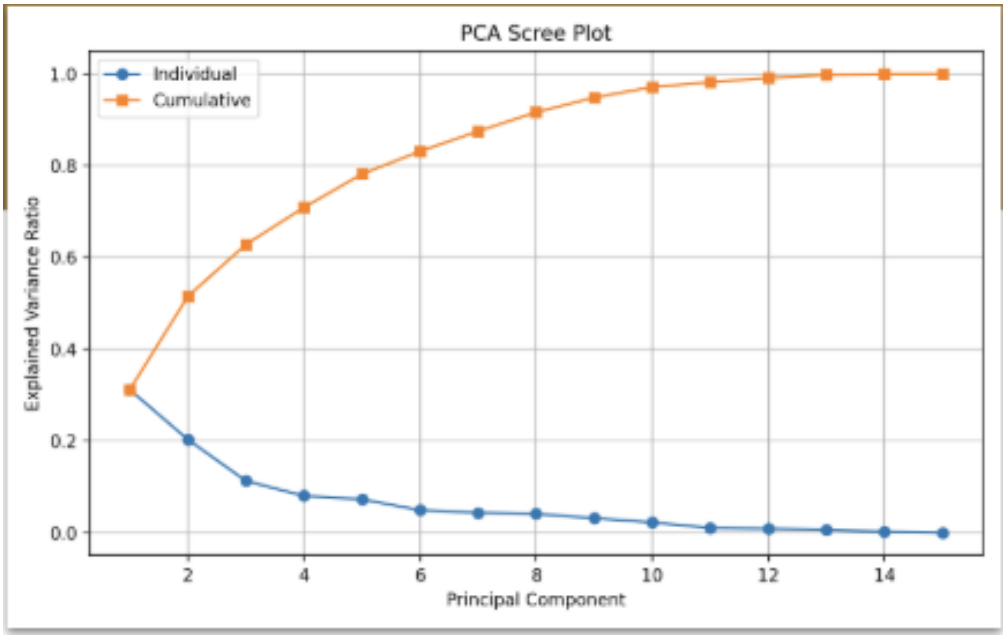


Figure 5.—Scree plot of PCA variance.

4.2 Cluster Analysis

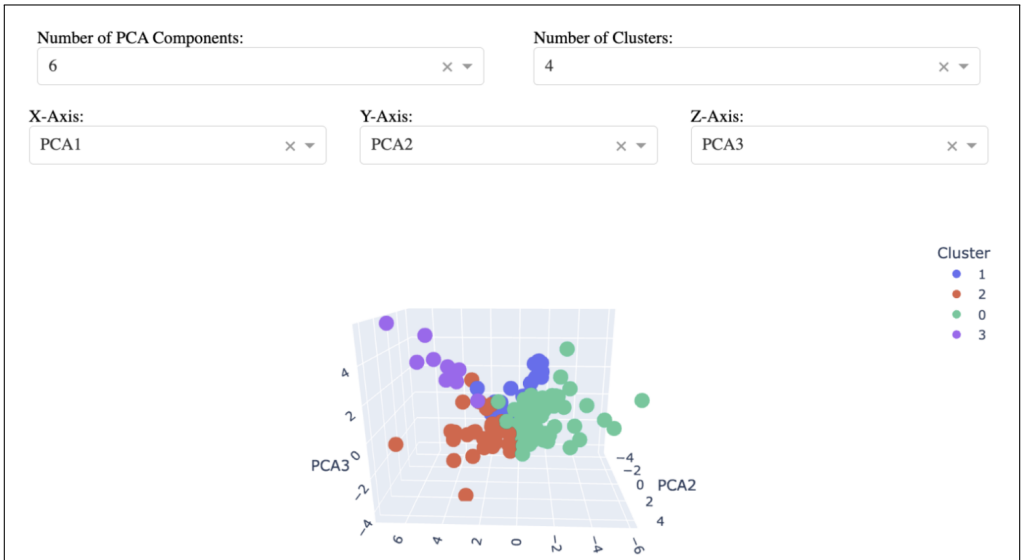


Figure 6.—PCA projection of clustered experiments.

Varying k revealed distinct groupings, interpreted using PCA loadings and rhythm patterns. Figure 6 displays the data plotted in the PCA space, with k-means clustering applied with $k = 4$.

4.3 Cluster Validation

To assess clustering quality and interpret cluster meanings, the Szymkiewicz–Simpson overlap is computed between each cluster and true class:

$$\mathrm{Sim}(A, B) = \frac{|A \cap B|}{\min(|A|, |B|)}$$

Values close to 1 indicate a strong match between a cluster and a class, whereas values near 0 suggest little to no overlap.

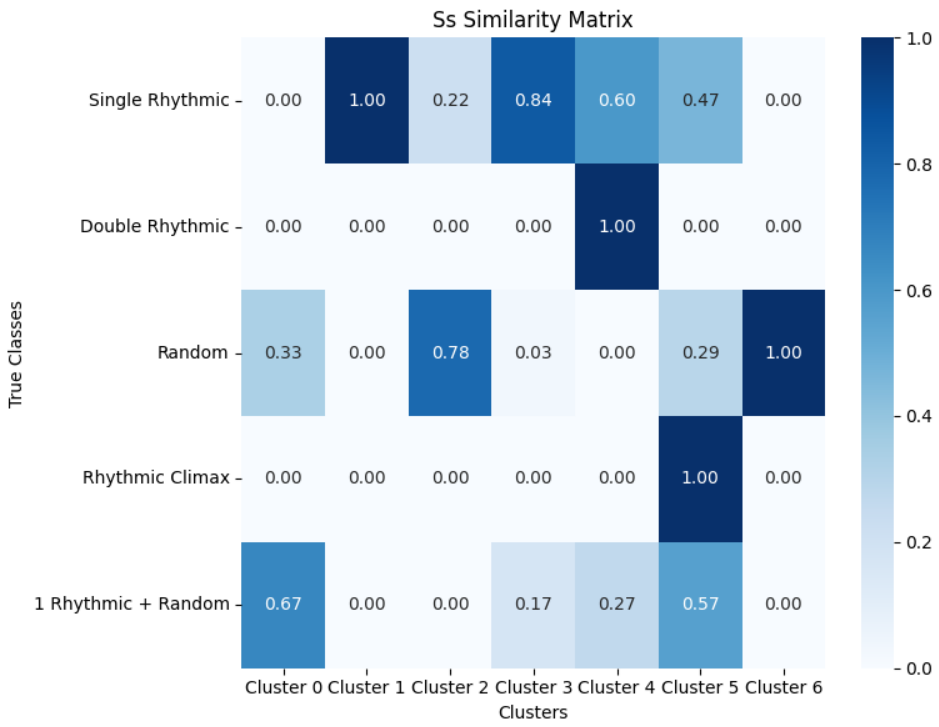


Figure 7.—Cluster validation using Szymkiewicz–Simpson overlap coefficient.

Figure 7 presents the resulting overlap matrix with 7 clusters.

5 Supervised Learning

5.1 Training and Evaluation

The supervised learning stage of the pipeline leverages labeled data to train a Decision Tree classifier for boiling regime classification. After feature extraction, categorical labels are encoded and the data is split into training and validation folds. Cross-validation is used to provide robust estimates of model performance.

Model training and evaluation are fully automated: the decision tree is fit on the training data, and predictions are generated for each fold. Performance metrics, including per-class F1-scores and a weighted F1-score, are computed to assess classification accuracy across both dominant and minority classes. The pipeline also generates visualizations such as feature importance plots and confusion matrices, which are saved for downstream analysis and reporting.

This approach ensures that model evaluation is consistent, transparent, and easily extensible to new data.

5.2 Results

Table 1.—F1-scores by class from Decision Tree classifier

Class	F1-score (%)
Single Rhythmic	85.2
Random	83.4
Noise	87.4
Rare Classes	Low due to insufficient support
Weighted F1-score	**79.1**

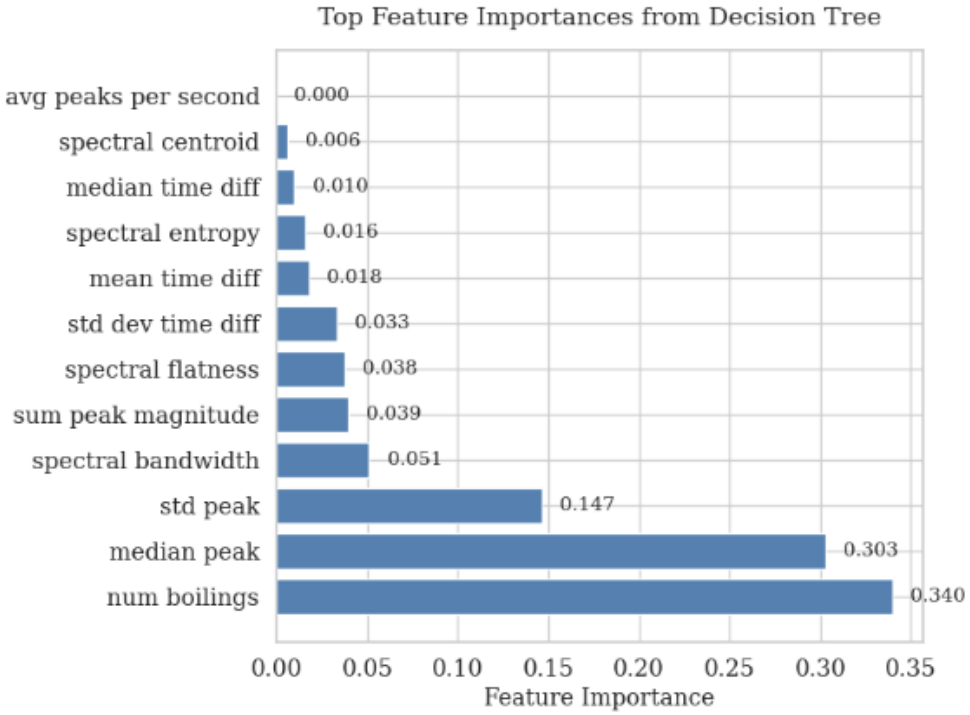


Figure 8.—Feature importance from decision tree.

As shown in Figure 8, the most influential features identified by the decision tree include the number of boilings, which captures the frequency of boiling events; the median and standard deviation of peak magnitudes, which reflect the central tendency and variability of acoustic signal intensity and the spectral bandwidth, which characterizes the distribution of energy across the frequency domain. These features collectively provide a robust representation of both temporal and spectral characteristics relevant to boiling regime classification.

5.3 Comparison with Other Methods

Alternative modeling approaches, such as XGBoost and neural networks, were also explored for boiling regime classification. However, these models consistently underperformed when compared to the Decision Tree classifier. Across both dominant and minority classes, they yielded lower F1-scores and failed to identify meaningful distinctions among rhythmic patterns. Moreover, their lack of interpretability limited the ability to extract actionable insights or visualize decision boundaries—a key requirement for this domain. In contrast, the Decision Tree offered a reliable and transparent solution, aligning well with the interpretability demands and data limitations of this study.

5.4 Handling Class Imbalance

Several techniques were evaluated to address the imbalance in class distribution, including class weighting, random oversampling, and Synthetic Minority Over-sampling Technique (SMOTE). Despite these efforts, improvements remained modest due to the severe sparsity in certain labels. Minority classes such as *Double Rhythmic* and *Rhythmic Climax* were especially affected, with insufficient samples to enable generalizable learning. These limitations suggest that future progress may depend more on collecting targeted additional data or employing domain-informed data augmentation techniques to enrich underrepresented classes.

5.5 Confusion Matrix Analysis

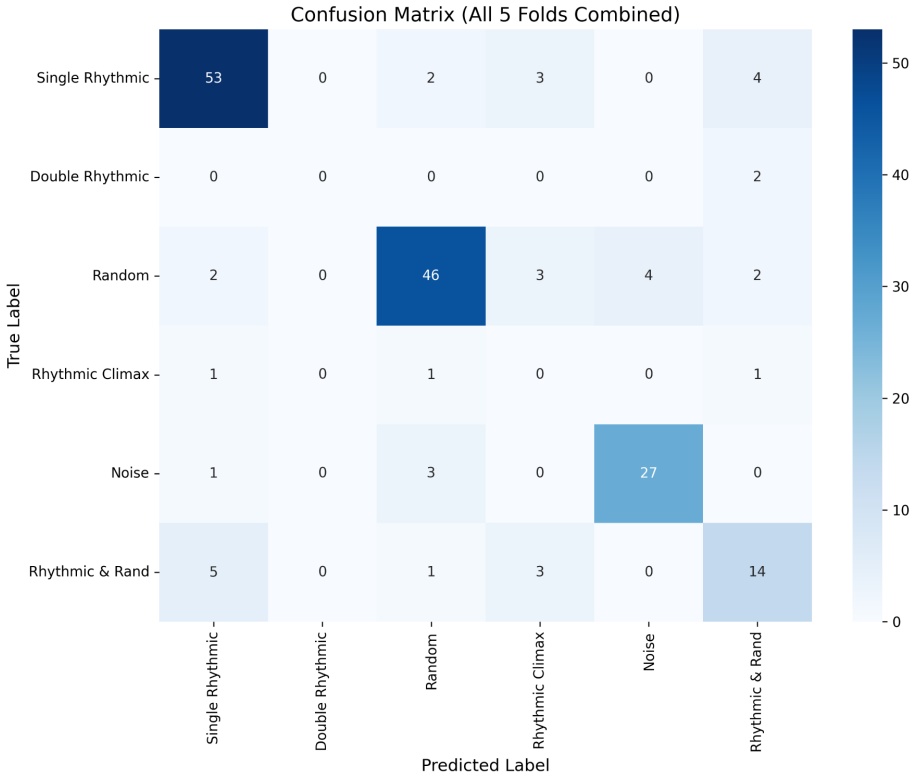


Figure 9.—Confusion matrix.

Figure 9 presents the confusion matrix aggregated across all five folds of cross-validation. Most misclassifications occurred between acoustically or structurally similar boiling regimes. Notably:

- *Single Rhythmic* was occasionally misclassified as *Rhythmic Climax* or *Rhythmic & Random*.
- *Random* sometimes overlapped with both *Noise* and *Rhythmic Climax*.
- *Double Rhythmic* and *Rhythmic Climax* classes yielded very low true positive rates, reinforcing the issue of class sparsity.

These trends suggest that while the model is effective at detecting dominant rhythmic structures, it struggles to separate closely related or infrequent patterns. This difficulty likely stems from overlapping feature representations and limited training data for the minority classes.

6 Web Application for Interactive Exploration and Classification

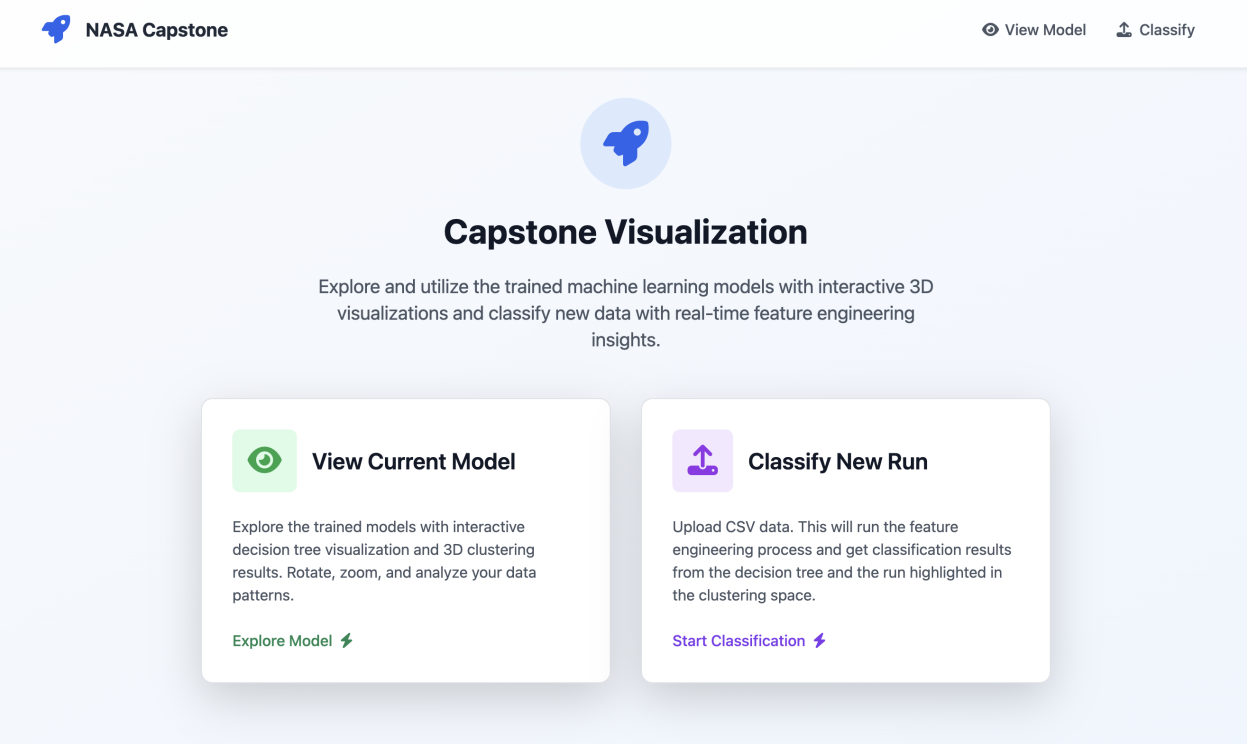


Figure 10.—Web app landing page presenting two primary modes: model exploration and new run classification.

To bridge the gap between model development and researcher usability, an interactive web application was built using Dash, as seen in Figure 10. This web interface serves as both an analytical and interpretive tool, enabling researchers at NASA to explore clustering results, visualize decision boundaries, and classify new experimental runs with minimal technical overhead.

The application supports two primary modes of use:

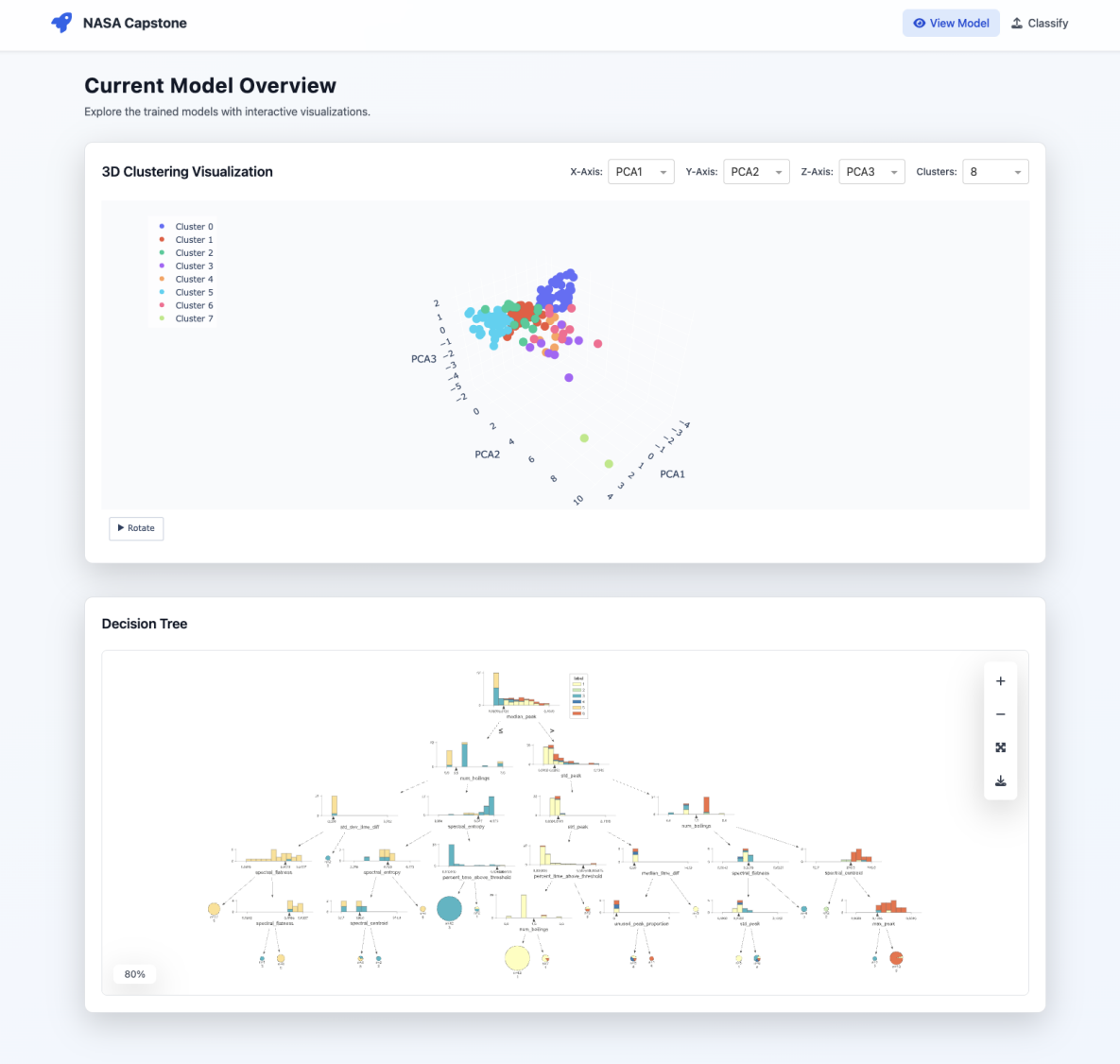


Figure 11.—Model Exploration mode.

1. **Model Exploration Mode (Figure 11):** The behavior of the current trained unsupervised and supervised models can be inspected. This includes an interpretable decision tree visualization rendered as a zoomable, pannable SVG, allowing users to trace the logical flow of predictions and understand how specific features influence model decisions. In parallel, clustering results can be explored in a 3D PCA-reduced space. The number of clusters and selected PCA axes can be dynamically adjusted to reveal latent structure in the data and identify potential regime separability.

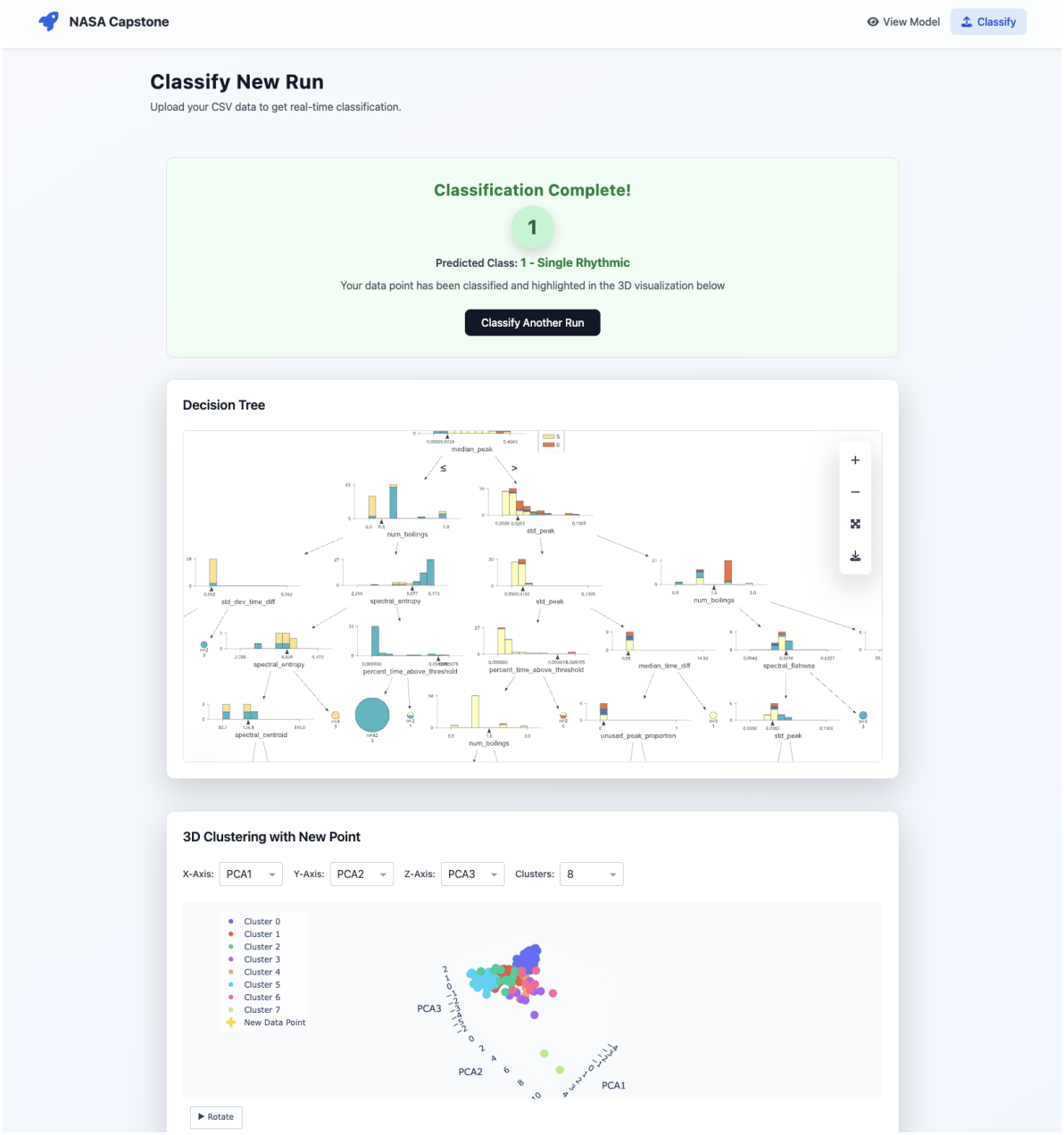


Figure 12.—Classification mode output: predicted boiling regime and projection of the new data point in PCA cluster space.

2. **Classification Mode (Figure 12):** New experimental runs can be uploaded in a CSV format. Once uploaded, the application automatically applies the full feature engineering pipeline, including time-domain and frequency-domain transformations. The extracted features are passed through the trained decision tree classifier, which outputs a predicted boiling regime, class probabilities, and a list of influential features. Simultaneously, the new run is projected into the PCA space and overlaid on the existing cluster visualization. This

dual approach allows classification outcomes to be assessed in both a supervised context (via the decision tree) and an unsupervised context (via clustering), offering a more holistic understanding of the run's regime.

Together, these capabilities transform the machine learning pipeline into a robust, real-time decision support tool. The web application enables interactive exploration of both the unsupervised (clustering) and supervised (decision tree) models, validation of predictions, and interpretation of key signal characteristics driving boiling regime classification. Additionally, the modular architecture of the codebase supports future extensions and modifications, whether to the regimes, experimental runs, decision tree logic, or clustering pipeline. This flexibility enables potential integrations such as onboard deployment with spacecraft sensor streams or adaptive model retraining as new data becomes available.

7 Conclusion

A machine learning pipeline for detecting early-stage boiling in cryogenic fuel systems using high-resolution acoustic emissions is presented. The approach leverages both time- and frequency-domain signal processing to extract responsive boiling indicators. These features feed into a hybrid modeling framework consisting of unsupervised clustering for regime discovery and supervised decision tree classification for real-time interpretability and classification.

To support deployment and usability in practical research settings, a custom web application was developed to visualize model behavior, inspect boiling regime boundaries, and classify new experimental runs. This interface bridges the gap between data science and application, making advanced analytics accessible and actionable.

The results demonstrate that acoustic-based detection, when paired with machine learning, offers a viable path forward for enhancing safety and situational awareness in cryogenic systems. The modular and extensible design of the pipeline also enables ongoing adaptation, supporting future extensions such as retraining with new data or expansion to other discovered regimes. Ultimately, this work lays the foundation for building safer, smarter, and more autonomous cryogenic fuel monitoring systems for future space missions.

Appendix A Detection of Dominant Rhythmic Patterns Algorithm

Purpose

This algorithm aims to detect recurring rhythmic patterns in a sequence of detected acoustic peaks. It evaluates every possible candidate interval derived from peak differences, tests each rhythm statistically, and consolidates similar patterns while removing weak or redundant ones. The result is a clean list of dominant intervals that likely represent true underlying rhythmic structure in the data.

Plain English Overview of Algorithm

0. **Estimate Noise Level**
   Before any rhythm analysis, the algorithm estimates the natural variability in peak timing and amplitude. This standard deviation is used to compute a margin of tolerance when evaluating rhythmic alignment and to calibrate statistical tests later in the process.

1. **Generate Rhythm Candidates**
   Every pair of peaks is used to define a potential rhythm interval $d = x_j - x_i$, anchored at x_i. Only differences that allow a sufficient number of repetitions (typically based on $\sqrt{n}$) within the signal duration are retained.

2. **Count Hits Per Candidate**
   For each candidate, the algorithm walks forward in time from the anchor using $t_k = \text{anchor} + k \cdot d$, checking whether predicted steps align closely with real peaks, both in the x and y direction. Each match is a "hit," and both the number of hits and total prediction attempts (tries) are recorded.

3. **Prune Candidates via Binomial Test**
   The algorithm tests each candidate using a one-sided binomial hypothesis test. It asks: "Is the observed hit rate significantly better than random chance?" Candidates that fail this test (after adjusting for multiple comparisons) are removed.

4. **Group Similar Candidates**
   Candidates that share a large number of overlapping hits are grouped using a similarity metric (Szymkiewicz–Simpson coefficient). Within each group, the strongest candidate is retained and others are marked as absorbed.

5. **Final Filtering by Support**
   As a final step, candidates that don't have enough total support (from both their own hits and absorbed ones) are removed. This ensures the algorithm outputs only the most robust rhythmic patterns.

The pseudocode for the algorithm is presented in Algorithm 1.

Algorithm 1: Detection of Dominant Rhythmic Patterns in Acoustic Peaks

Input: sorted peak times $x = [x_1, \ldots, x_n]$, amplitudes $y = [y_1, \ldots, y_n]$, total duration T
Output: filtered list of dominant rhythmic candidates

Step 0: Estimate Noise Level;
$\sigma_x \leftarrow$ `estimate_std`(x); `// timing variability`
$\sigma_y \leftarrow$ `estimate_std`(y); `// amplitude variability`
$m_x \leftarrow z_\alpha \sigma_x$,;
$m_y \leftarrow z_\alpha \sigma_y$;
$p_{\text{null}} \leftarrow \min\left(1, \frac{2m_x n}{T}\right)$;
Step 1: Generate Rhythm Candidates;
for $i \leftarrow 1$ **to** $n-1$ **do** `// pairwise peak spacings`
  for $j \leftarrow i+1$ **to** n **do**
    $d \leftarrow x_j - x_i$;
    if $x_j + \sqrt{n}\, d > T + m_x$ **then**
      break
    end
    else
      add candidate (d, x_i) to list;
    end
  end
end
Step 2: Count Hits Per Candidate;
for *each candidate* (d, a) *in list* **do**
  for $k \leftarrow 0$ **to** $\lfloor (T-a)/d \rfloor$ **do**
    $t_k \leftarrow a + kd$;
    estimate $\bar{y}_k$;
    if $\exists\, x_i \in [t_k \pm m_x]$ ***and*** $y_i \in [\bar{y}_k \pm m_y]$ **then**
      hit++;
    end
    tries++;
  end
end
Step 3: Prune via Binomial Test;
for *each candidate with* h *hits and* t *tries* **do**
  $\alpha' \leftarrow \alpha / R$;
  $p \leftarrow$ `BinomialTest`$(h-2, t-2, p_{\text{null}})$;
  if $h < \sqrt{n}$ ***or*** $p \geq \alpha'$ **then**
    remove candidate;
  end
end
Step 4: Group Similar Candidates;
build similarity graph (edge if Szymkiewicz–Simpson $\geq \tau$);
for *each connected component* **do**
  keep candidate with highest centrality;
  absorb unique hits from others;
end
Step 5: Final Filtering by Support;
for *each remaining candidate* **do**
  if *hits + absorbed* $< \sqrt{n}$ **then**
    remove candidate;
  end
end
return remaining candidates as dominant rhythmic intervals;

Example Walkthrough Using Peak Time Array

We demonstrate the full algorithm using the following observed peak times from an acoustic signal:

$$\text{peaks} = \left\{ \begin{array}{llll} 0.1035, & 0.3391, & 0.5746, & 0.8100, \\ 1.0454, & 1.2791, & 1.5160, & 1.7496, \\ 1.9866, & 2.2220, & 2.4574, & 2.6930, \\ 2.9286, & 3.1641, & 3.3999, & 3.6359, \\ 3.8722, & 4.1088, & 4.3458, & 4.5828, \\ 4.8197 & & & \end{array} \right\}$$

The following plot shows the original acoustic signal, with peaks marked:

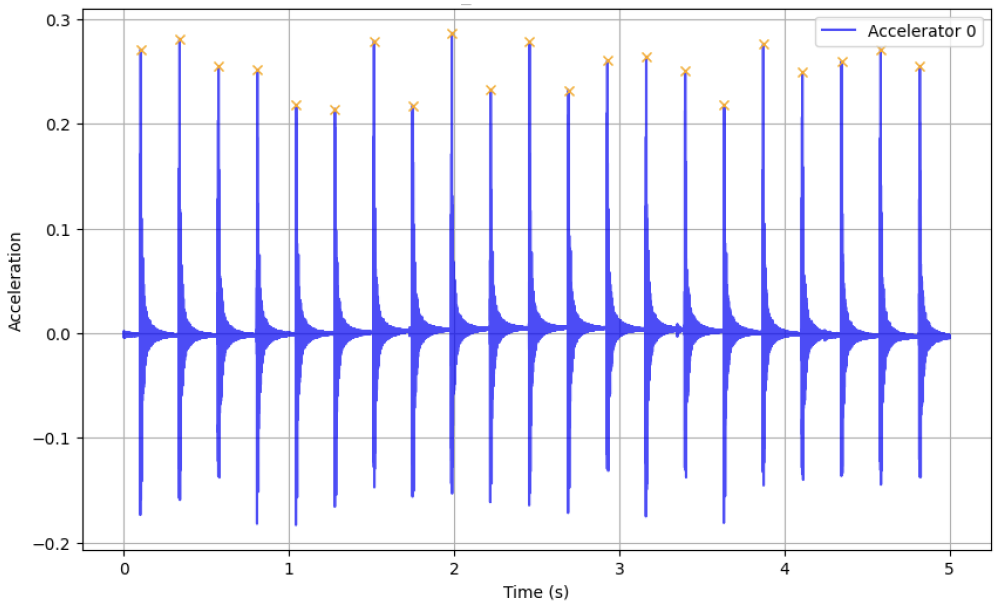


Figure 13.—Acoustic signal with extracted peaks.

Step 0: Estimate Noise Level

We calculate the global standard deviation of peak-to-peak spacing:

$$\text{SD} = 0.00091 \quad \Rightarrow \quad \text{Margin} = z \cdot \text{SD} = 0.0015$$

The probability of a random hit in a margin window is estimated as:

$$p_{\text{null}} = \frac{2 \cdot \text{margin} \cdot \#\text{peaks}}{\text{run length}} = \frac{2 \cdot 0.0015 \cdot 21}{5} \approx 0.0126$$

Step 1: Generate Rhythm Candidates

For each peak pair (i, j), we compute the difference $d = x_j - x_i$ and retain it as a candidate if it allows at least $\sqrt{n}$ steps within the duration.

Example early candidates:

- Candidate 0: $d = 0.2356$, anchor $= 0.1035$

- Candidate 1: $d = 0.4711$, anchor $= 0.1035$

- Candidate 2: $d = 0.7065$, anchor $= 0.1035$

This process generates a total of 41 candidates.

Step 2: Count Hits Per Candidate

We count how many predicted steps align with real peaks (within the x and y margin), starting from each candidate's anchor and incrementing by d. The first two hits (anchor and anchor+d) are assumed, and only later steps are evaluated.

Candidate ID	d **(sec)**	**Anchor**	**Hits**	**Tries**
0	0.2356	0.1035	20	20
1	0.4711	0.1035	10	10
2	0.7065	0.1035	6	6
3	0.9429	0.1035	5	5
4	1.1785	0.1035	3	4

Step 3: Prune Candidates via Binomial Test

Each candidate is evaluated using a one-sided binomial test under:

$$H_0 : \text{hit rate} = p_{\text{null}} \qquad H_A : \text{hit rate} > p_{\text{null}}$$

A Bonferroni adjustment is applied using α/R for R candidates. Clusters failing the test are removed.

Step 4: Group Similar Candidates

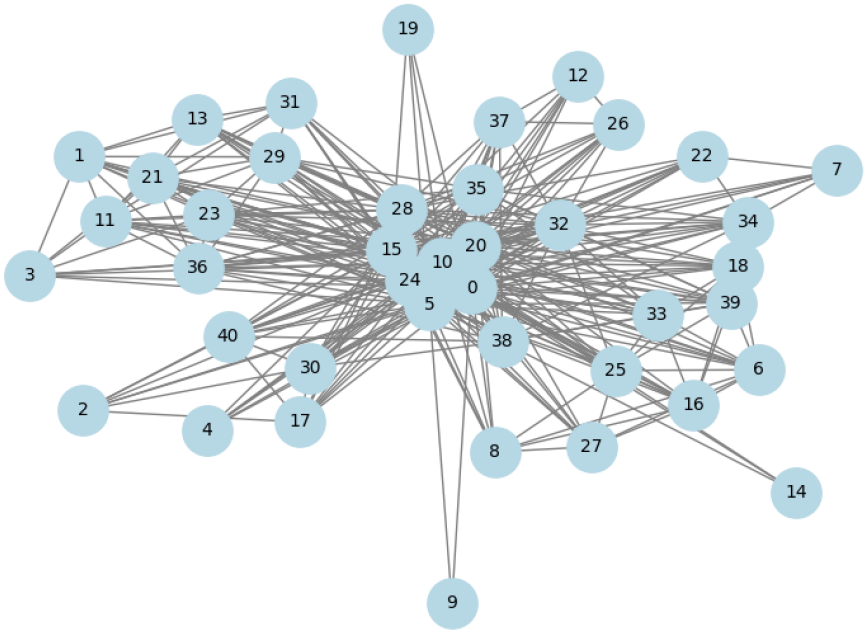


Figure 14.—Candidate similarity graph.

Remaining candidates are grouped based on the Szymkiewicz–Simpson overlap of their hit indices:

$$\mathrm{Sim}(A, B) = \frac{|A \cap B|}{\min(|A|, |B|)}$$

Note that $0 \leq \mathrm{Sim}(A, B) \leq 1$, where a value of 0 indicates that the sets are disjoint, and a value of 1 indicates that the smaller set is entirely contained within the larger. To identify and consolidate redundant candidates, a similarity graph is constructed by connecting pairs of candidates whose overlap exceeds a predefined threshold (e.g., 0.8). Each edge represents significant shared alignment between two candidates. Within each connected component of this graph, the candidate with the highest degree centrality is selected as the representative rhythm. All other candidates in the component are considered redundant and are absorbed into the dominant one, contributing their supporting peak hits.

Step 5: Final Filtering by Support

Candidates with too few total hits (direct + absorbed) are removed. In this example, only rhythms with $\geq \sqrt{21} \approx 5$ hits are retained.

Final Result

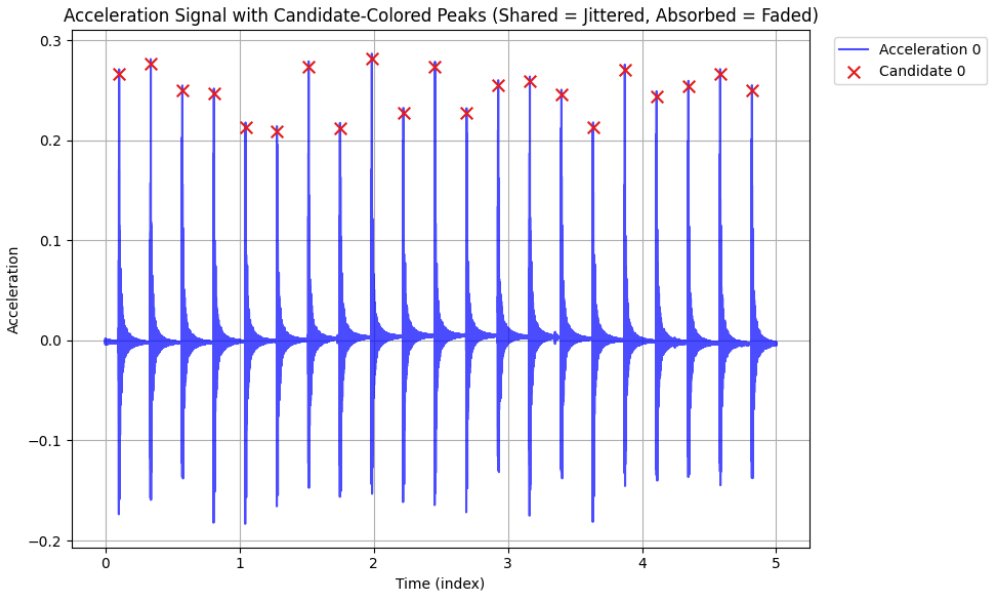


Figure 15.—Labeled acoustic signal with candidate assignments.

The final output is the strongest rhythmic regime:

Detected Rhythm: ID = 0, $\mu = 0.2356$ seconds, aligned with 20 out of 20 peaks

This rhythm is statistically significant, and it absorbs several weaker candidates into a unified regime.

The figure below shows the acoustic signal with peaks labeled according to their assigned rhythmic candidate. Each candidate is represented by a distinct color—if multiple rhythmic candidates are present, the peaks associated with each are shown in different colors.

Example on Double Rhythmic Signal

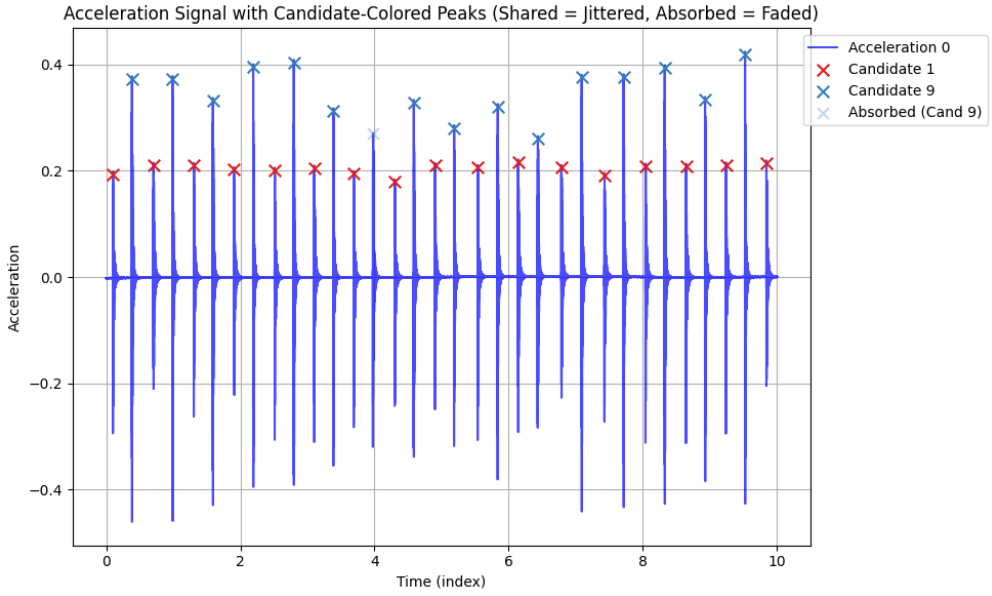


Figure 16.—Labeled double rhythmic acoustic signal with candidate assignments.

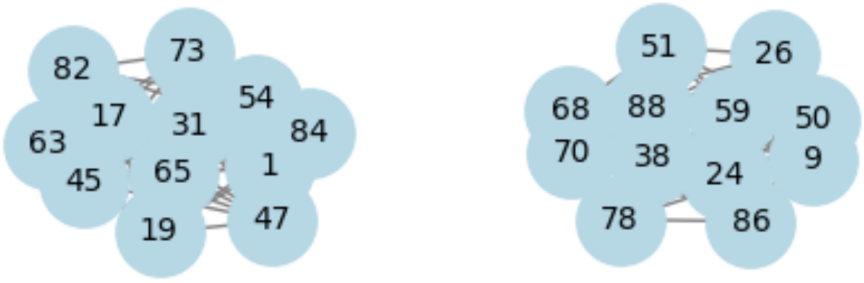


Figure 17.—Double rhythmic similarity graph.

We also illustrate the algorithm on a run classified as double rhythmic in Figure 16. The similarity graph is shown in Figure 17. Each of the two distinct connected components corresponds to one of the two rhythms, as expected.

Concurrent bubble nucleation sequences were found to have significantly larger variance in their inter-nucleation times than single rhythm sequences, suggesting that concurrent nucleations affect one another. A detailed statistical analysis of this phenomenon will be published elsewhere.*

*Martinez, Khasin, in preparation.

Notes

- Peaks that align with a candidate's predicted steps—but were not part of its original cluster—are marked as absorbed, as seen in Figure 16.
- Absorbed peaks are visualized using a lighter shade of the candidate's color to distinguish them from originally clustered peaks.
- Peaks not matched to any candidate rhythm are considered unused and are displayed in gray and likely correspond to random boiling. A higher proportion of unused peaks may indicate random boilings.

A.1 PCA Loadings and Cluster Interpretation

Principal Component Analysis (PCA) was used to reduce the complex, multi-dimensional feature space into a few principal axes that capture the dominant patterns in the boiling regime data. By examining the contributions of each original feature to the principal components, we can interpret the underlying physical and statistical characteristics that differentiate the data.

Principal Component 1 (PCA1):

The first principal component primarily reflects the overall intensity and activity of the boiling signal. High values along this axis correspond to runs with more frequent and larger acoustic events, as well as signals that are more complex and less uniform. In contrast, lower values indicate quieter runs with fewer, smaller, or more uniform events, and a more regular or "flat" signal profile.

Principal Component 2 (PCA2):

The second principal component is most influenced by the temporal structure of the boiling events and the presence of unused or less prominent acoustic features. High values along this axis are associated with runs where the timing between events is more variable or extended, and where a greater proportion of potential events are not strongly expressed in the signal. This suggests a regime with more irregular or sporadic boiling activity.

Principal Component 3 (PCA3):

The third principal component captures variation related to the complexity and distribution of energy within the signal. High values along this axis are associated with runs where the acoustic energy is more broadly distributed across different frequencies and where the signal exhibits greater unpredictability or irregularity. This suggests that PCA3 distinguishes between boiling regimes with more complex, less periodic acoustic patterns and those with simpler, more predictable structures.

Cluster Interpretation in PCA Space:

- **Cluster 0 (green):** This group is distinguished by higher values along the second principal component, indicating runs with more irregular timing between events and a greater presence of subtle or less pronounced acoustic activity.
- **Cluster 1 (blue):** This cluster stands out for its higher values along the first principal component, representing runs with more intense, frequent, and complex boiling activity.
- **Cluster 2 (red):** This group is characterized by lower values along the first principal component, corresponding to runs with less intense, more uniform, and quieter boiling signals.

This analysis shows that the main axes of variation in the data correspond to differences in both the intensity and regularity of boiling activity, as well as the presence of subtle or sporadic events. The clustering in PCA space thus provides a physically meaningful separation of boiling regimes, reflecting both the strength and the temporal structure of the acoustic signals.

Appendix B Decision Tree Summary and Metrics

B.1 Trained Decision Tree

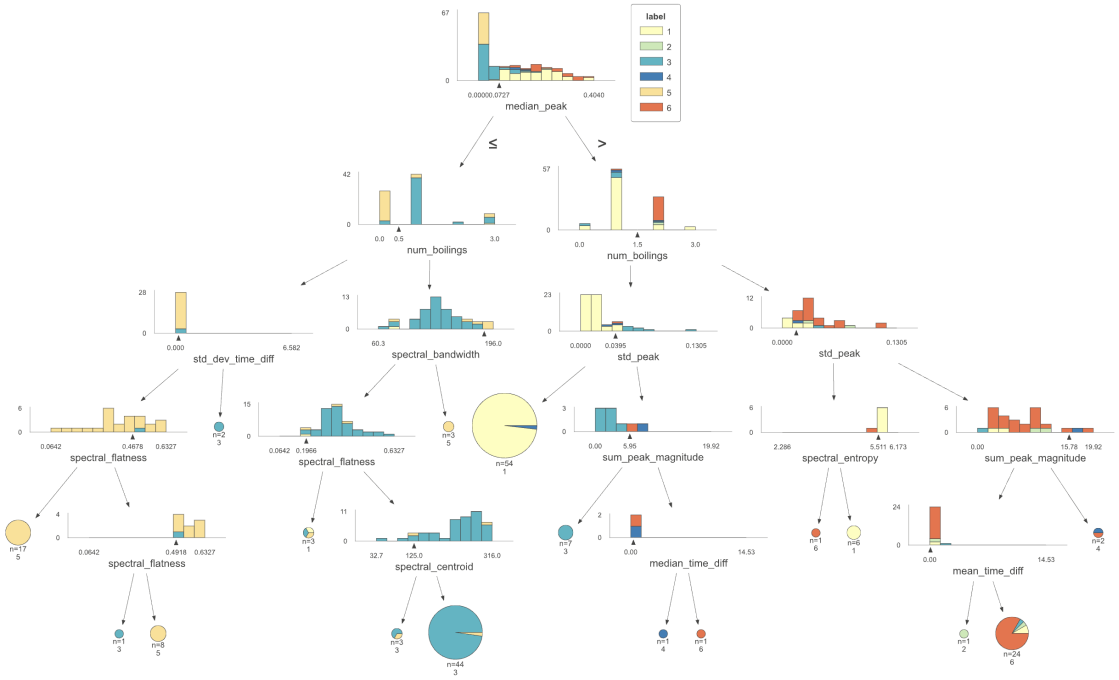


Figure 18.—Final decision tree trained on acoustic feature set.

Figure 18 shows our final trained decision tree. This model was constrained to a maximum depth of five to improve interpretability and prevent overfitting.

Table 2.—Decision Tree Classification Report

Class	Support	Precision	Recall	F1-Score
Single Rhythmic	62	86.56	85.48	85.22
Double Rhythmic	2	0.00	0.00	0.00
Random	57	87.54	80.70	83.36
Rhythmic Climax	3	0.00	0.00	0.00
Noise	31	90.16	87.09	87.38
Rhythmic & Rand	23	68.98	60.86	58.11
Weighted Avg	**178**	**82.80**	**78.65**	**79.11**

B.2 Interpretation

The decision tree performed well for high-support classes, achieving F1-scores above 83% for Single Rhythmic, Random, and Noise. However, the model struggled with underrepresented classes (Double Rhythmic, Rhythmic Climax), resulting in F1-scores of zero due to insufficient training data. These results underscore the need for class-balancing techniques and possibly more data or ensemble methods to boost performance on rare classes.

Data availability. The acoustic, visual, and metadata files used in this study are publicly available through NASA's Intelligent Systems Division datasets page under the entry "Acoustic and Visual Data for Incipient Boiling at Local Heat Leaks in a Water Tank" (Ref. 2).

References

1. M. Khasin, J. Rogers, and V. Osipov, "Acoustic Data-based Characterization of Incipient Boiling for Space Applications," NASA Technical Memorandum NASA/TM–20250009668, September 2025.

2. M. Khasin, "Acoustic and Visual Data for Incipient Boiling at Local Heat Leaks in a Water Tank (v1.0)," NASA Intelligent Systems Division Datasets, NASA Ames Research Center, last updated Nov. 25, 2025. Available: `https://www.nasa.gov/intelligent-systems-division/datasets/`. Accessed: December 13, 2025.

3. V. K. Dhir, "Boiling under microgravity conditions," in *Proc. 12th Int. Heat Transfer Conf.*, Grenoble, France, 2002.

4. C. B. Muratov, V. V. Osipov, and V. N. Smelyanskiy, "Issues of long-term cryogenic propellant storage in microgravity," NASA Technical Memorandum TM-2011-215988, ARC-E-DAA-TN4295, 2011.