

ANOMALY DETECTION FROM GLOBAL SEISMOLOGY DATA

Abstract

Earthquakes leave characteristic signatures in ground motion, but these are embedded in a sea of human and environmental noise. This project investigated whether low-cost Raspberry Shake stations can act as unsupervised earthquake detectors by training autoencoders (AEs) solely on non-earthquake (non-EQ) segments and then treating earthquakes (EQ) as reconstruction anomalies when the reconstruction fails.

A dataset spanning 20 Raspberry Shake stations across rural, urban, and industrial environments was compiled. Station metadata was obtained from the Raspberry Shake network and the EarthScope/IRIS MetaData Aggregator, while earthquake events of varying magnitudes were sourced from the USGS ComCat catalog. Data was segmented into fixed-length windows (25-s windows of 2,500 samples binned to 500 using Standard deviation (STD-binning), per-segment normalisation, and processed through AE training and threshold selection (95th percentile (q95) vs. ROC-optimal (Youden's J) criterion). Evaluation used both EQ and Non-EQ segments, with further analysis of latent 64-dimensional bottleneck features using dimensionality reduction (t-SNE/UMAP) and Random Forest (RF) classification.

Results showed that autoencoders assigned higher reconstruction errors to EQ than non-EQ segments, though overlap between distributions limited separability. A baseline model achieved ROC-AUC = 0.543 and F1 = 0.717, while an improved AE with a 64-dimensional bottleneck reached ROC-AUC = 0.717, Precision = 0.899, Recall = 0.704, and F1 = 0.790. Tests on larger external earthquakes (Kamchatka M8.8; Drake Passage M7.5) yielded ROC-AUC = 0.691 for the AE and 0.749 for the Random Forest classifier.

In addition, a Human Activity Trend Analysis framework was explored, demonstrating how seismic noise might reflect daily and weekly urban or

industrial cycles, though limited sampling prevented conclusive findings. Overall, the study confirms that AEs detect earthquakes as anomalies, but performance is constrained by station heterogeneity, non-EQ transients, and label granularity. Future work should explore CNN/LSTM encoders, station-specific adaptation, and larger earthquake corpora.

1. Introduction

Earthquakes remain one of the most destructive natural hazards, threatening lives, infrastructure, and economies worldwide. The key challenge in seismology is not only recording seismic events, but also reliably identifying them against a background of human and environmental noise such as traffic, industrial, and weather. Traditional detection pipelines often rely on manual picking of seismic arrivals or heuristic algorithms (Allen, 1982; Withers *et al.*, 1998). While effective at smaller scales, these approaches struggle with the vast and complex continuous waveform archives now available. This limitation has motivated the integration of machine learning, particularly unsupervised approaches such as anomaly detection, where earthquakes must be separated from abundant background activity.

Autoencoders (AEs) are neural networks trained to reproduce their input through a compressed bottleneck representation. Trained on normal data, they capture the regular structure of that data; inputs that deviate from what the model has seen incur larger reconstruction errors. This makes AEs a natural fit for unsupervised anomaly detection in time series, where anomalous segments can be flagged by reconstruction loss (Sakurada and Yairi, 2014; Chalapathy and Chawla, 2019).

Hinton and Salakhutdinov (2006) first demonstrated that deep autoencoders could preserve salient structure in a low-dimensional code, paving the way for reconstruction-based anomaly pipelines. Subsequent research applied to sequential data used recurrent encoder–decoder models to detect unusual patterns in time series (Malhotra *et al.*, 2016), while other work introduced robust deep autoencoders (Zhou and Paffenroth, 2017), deep one-class classification (Ruff *et al.*, 2018), and mixture-based decoders (Zong *et al.*, 2018). A consistent result is that reconstruction error is an effective anomaly

score when data are heterogeneous or labels are scarce, both common conditions in seismology.

Machine learning applications in seismology have expanded rapidly. Supervised deep learning has been widely adopted for event detection, phase picking, and characterisation of seismic signals. Bergen *et al.*, (2019) catalogued how convolutional (CNNs) and recurrent networks outperform traditional methods in picking P- and S-wave arrivals and events identification. Kong *et al.*, (2018) reviewed opportunities and challenges of machine learning in seismology, emphasising the under-explored role of unsupervised methods despite their potential for global monitoring. They highlighted three core obstacles: (i) limited generalisability across regions, (ii) data imbalance between rare earthquakes and abundant noise, and (iii) contamination of the non-EQ class by human and environmental signals.

Several landmark supervised models demonstrate the power of labelled data. PhaseNet, introduced by (Zhu and Beroza, 2019) A convolutional network trained on millions of labelled waveforms achieved over 95% accuracy in picking seismic phases, greatly accelerating cataloguing. Earthquake Transformers by (Mousavi *et al.*, 2020) extended this to continuous seismic data, detecting events and phases with precision and recall both above 95% in Southern California. While these models scale well when supplied with annotated datasets, their reliance on labels is a key limitation. Labels are expensive, region-specific, and scarce in global monitoring. This highlights the value of unsupervised approaches, which can exploit abundant unlabelled noise without requiring exhaustive manual labelling.

Autoencoder-based anomaly detection directly addresses this gap. By training exclusively on non-EQ waveforms, models learn to encode background seismicity. When tested on earthquake signals, reconstruction errors increase, enabling event identification. Beyond reconstruction error, bottleneck representation provides additional insight. The latent 64-dimensional bottleneck representation was analysed using t-SNE and UMAP for dimensionality reduction, and a Random Forest (RF) classifier was applied to probe discriminative structure beyond reconstruction loss. This dual perspective error-based anomaly detection and feature-based classification offers a robust framework for earthquake detection in a noisy, real-world seismic environment.

2. Methods

2.1 Data Collection and Station Characterisation

Seismic waveform data were sourced from the Raspberry Shake network, a citizen science initiative deploying affordable seismographs globally. This study used 20 stations (16 three-component [3D], 4 four-component [4D]). Each [3D] station recorded EHZ, EHE, EHN; each [4D] station recorded EHZ, ENE, ENN, ENZ. Waveforms were downloaded as 10-minute windows at 100 Hz, yielding 60,000 samples per channel.

In total, this yielded 22 waveform datasets:

- 10 corresponding to earthquake (EQ) events (all $< M5.0$, i.e., small to moderate magnitudes).
- 12 corresponding to non-earthquake (non-EQ) background noise.

This slight mismatch arises because one non-EQ station contributed two independent background noise windows, and one EQ station contributed two distinct earthquake events.

To ensure environmental diversity, each station was contextualised using metadata from the IRIS/SAGE MetaData Aggregator (MDA). Coordinates, elevation, and instrument type were extracted, while satellite map views from MDA were used to check surrounding land use and infrastructure. Based on this information, stations were grouped into:

- Rural environments: surrounded by vegetation, farmland, and sparse settlement.
- Urban environments: stations embedded in dense cityscapes with hospitals, schools, high-traffic road networks, residential blocks, and rail transit.
- Industrial environments: sites near factories, depots, and warehouses.

To minimise subjective bias in this classification, environmental labels were treated as working hypotheses rather than fixed assignments. They were later cross-checked against objective seismic features, such as

background amplitude levels and frequency-domain characteristics, to confirm whether the expected anthropogenic patterns were present. This ensures that classification was informed not only by visual inspection but also by measurable data.

The station characteristics, including codes, locations, components, environments, and coordinates, are summarised in Table 1. Their corresponding geographical distribution across multiple countries is illustrated in Figure 1.

Table 1. Metadata of Raspberry Shake stations used in this study, including geographic location, components, environments, and associated earthquake details where applicable

Station	Location	Components	Environment/Notes	Latitude	Longitude	Earthquake Info (distance in Km, Magnitude in M)
RCC8D	France	EHZ, EHE, EHN	Rural	49.8738	4.7394	None
RD9D6	United Kingdom	EHZ, EHE, EHN	Urban	53.2612	-2.1541	None
R5C62	Indonesia	EHZ, EHE, EHN	Urban	-6.3063	106.6740	29.6Km, M4.3
R76F0	United Kingdom	EHZ, EHE, EHN	Urban	53.8018	-1.5408	None
R0CEC	United Kingdom	EHZ, EHE, EHN	Rural	53.6396	-2.4921	None
R5B93	Pennsylvania, USA	EHZ, EHE, EHN	Urban	40.6936	-75.2058	None
R77F9	New York, USA	EHZ, EHE, EHN	Rural/Industrial	43.5495	-76.1124	25.36Km, M3.3
RE992_a	Ohio, USA	EHZ, EHE, EHN	Industrial	41.5675	-81.4972	14.6Km, M1.8

RE992_b	Ohio, USA	EHZ, EHE, EHN	Industrial	41.5675	-81.4972	44.63Km, M1.7
R853C	Argentina	EHZ, EHE, EHN	Rural/ Industrial	- 32.8918	-68.8708	16.8Km, M3.9
R119F_a	Germany	EHZ, EHE, EHN	Urban	49.0180	8.4208	None
R119F_b	Germany	EHZ, EHE, EHN	Urban	49.0180	8.4208	None
R1906	Germany	EHZ, EHE, EHN	Urban/ Industrial	52.4324	9.8422	None
R74A7	Japan	EHZ, EHE, EHN	Urban	34.9099	135.7970	14.5Km, M4.2
R3949	Mexico	EHZ, EHE, EHN	Rural	19.5135	-102.4183	None
R30A2	Texas, USA	EHZ, EHE, EHN	Urban	32.0270	-102.1010	42.9Km, M4.8
R9F13	Poland	EHZ, EHE, EHN	Urban/ Industrial	52.2162	21.1319	None
R5156	Lebanon	EHZ, EHE, EHN	Urban	33.8648	35.4917	82.818Km, M4.4
R7000	Egypt	ENE, ENN, ENZ, EHZ	Rural	26.7027	27.2290	None
R47BB	South Africa	ENE, ENN, ENZ, EHZ	Industrial/ Urban	-25.6576	27.2454	85.3Km, M4.2
R57CE	Sri Lanka	ENE, ENN, ENZ, EHZ	Industrial/ Urban	6.7927	79.8949	None
R5E3B	Armenia	ENE, ENN, ENZ, EHZ	Urban	40.2252	44.5709	13.19Km, M2.3

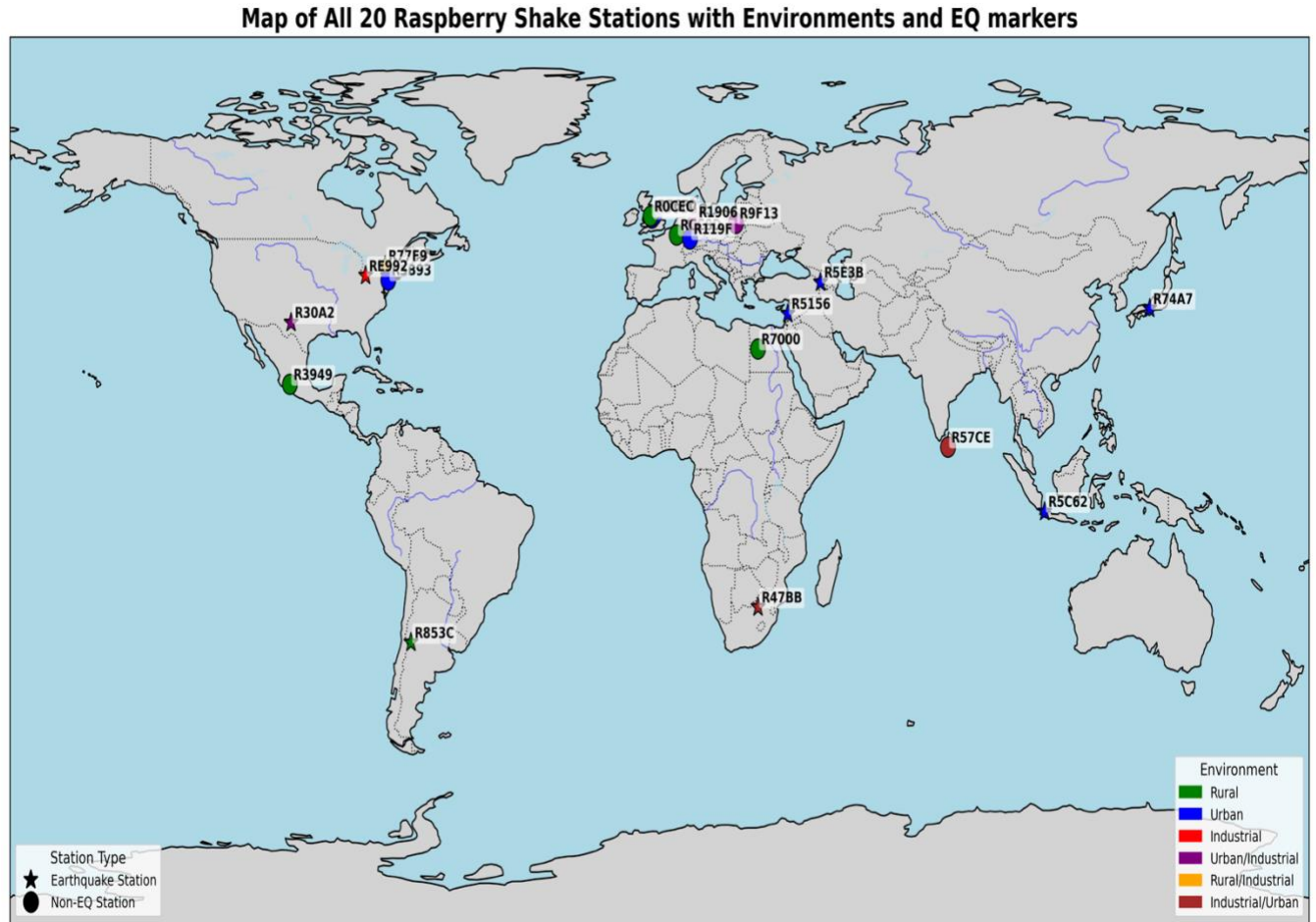


Figure 1. Global distribution of Raspberry Shake stations used in this study. The stations listed in Table 1 are shown by their country locations, illustrating the geographical spread across Europe, North America, South America, Asia, and Africa.

2.2 Earthquake and Non-Earthquake Waveform Selection

Earthquake data were obtained from the USGS Earthquake Catalog (ComCat), queried with station coordinates and time windows. For each event, 10-minute windows were extracted: 5 minutes before to 5 minutes after the reported origin time.

Non-EQ background data required additional care. For each chosen non-EQ station, the catalog was searched across weeks to months to identify periods devoid of local seismic events. Matching 10-minute windows were then downloaded at 100 Hz, providing true background noise comparable in length to EQ windows. This careful manual curation ensured that undetected small earthquakes did not inadvertently contaminate non-EQ data.

This produced 912 non-EQ segments and 768 EQ segments, roughly balanced. Additionally, to complement the main dataset ($<M5.0$), two recent large-magnitude earthquakes were downloaded:

- M8.8 Kamchatka (2025) and
- M7.5 Drake Passage (2025)

For these, 25-minute windows were extracted (-1 minute to $+24$ minutes around origin time), yielding 150,000 samples per channel at 100Hz (e.g., RF6DC, 3D station) and segmented identically. These served as a stress test to evaluate whether large events produced clearer anomalies. Their segment count (900, 500-point segments) was close to the main non-EQ set (912, 500).

2.3 Pre-processing and Segmentation

Each raw traces (10 or 25 minutes) were segmented into 25-second windows (2,500 samples). To reduce dimensionality while retaining structural information, a standard-deviation binning (STD-binning) procedure was applied:

- Each 25-second window was binned into 5-second blocks (reduced to a 500-sample input per segment).
- For each block, the standard deviation was computed.
- This yielded a compressed 500-sample vector per segment.

The rationale was that short-term standard deviation captures the essential envelope and variability of the waveform while discarding high-frequency noise, which is less relevant for anomaly detection. This is consistent with prior work where variance statistics summarise local activity levels in noisy time series.

Each 500-sample segment was then z-normalised (mean subtracted and divided by its standard deviation). This per-segment normalisation ensured that learning focused on waveform shape rather than raw amplitude or station gain differences.

For reproducibility, the non-EQ dataset was split once into `non_eq_X_train.npy` (729 segments) and `non_eq_X_test.npy` (183 segments),

stored on disk and reused consistently across experiments to avoid leakage. EQ datasets were stratified into 80/20 train/test splits when required.

2.4 Autoencoder Architectures

Two autoencoder families were implemented:

Baseline + Regularised Variant (Notebook 1)

The first autoencoder served as a baseline design. It consisted of a fully connected encoder-decoder pipeline with progressively reduced dimensionality ($500 \rightarrow 256 \rightarrow 128 \rightarrow 32 \rightarrow 128 \rightarrow 256 \rightarrow 500$). **Figure 2** shows this baseline architecture, where input waveforms are compressed into 32-dimensional bottleneck. This design was trained solely on non-earthquake (non-EQ) data to learn a representation of background seismic noise. Training used Adam optimiser with a learning rate of 0.001 (1e-3), activation functions were ReLU in hidden layers and linear at the output, batch size 64, and mean squared error (MSE) loss.

A regularised variant was also trialled with dropout (0.4) in each hidden layer and L2 weight regularisation (1e-3). Early stopping (patience = 10, max = 200 epochs) prevented overfitting.



Figure 2 Baseline autoencoder architecture ($500 \rightarrow 256 \rightarrow 128 \rightarrow 32 \rightarrow 128 \rightarrow 256 \rightarrow 500$).

Redesigned / Improved Model (Notebook 2)

The second autoencoder used a bottleneck of 64 ($500 \rightarrow 128 \rightarrow 64 \rightarrow 128 \rightarrow 500$), determined by a parameter sweep comparing reconstruction ROC-AUC across bottleneck/dropout settings. This architecture aimed at improving both stability and detection performance. This enabled richer latent representation of seismic variability while still enforcing compression.

Regularisation was lighter: L2 = 1e-4, no dropout, reflecting the need to balance robustness against underfitting. Training used Adam optimiser

(learning rate = 0.001, batch size = 128) with early stopping (patience = 5) and converged after ~80 epochs out of the 200 budget.

2.5 Thresholding Strategy

Since the autoencoder was trained exclusively on non-earthquake (non-EQ) segments, its reconstruction error distribution provided the natural reference point for threshold setting.

Statistical thresholding: $\mu + 3\sigma$ rule

Following (Shewhart, 1931), anomalies can be defined as points exceeding the mean plus three standard deviations ($\mu + 3\sigma$) of the non-EQ reconstruction error distribution. Under Gaussian assumptions, this flags the rarest 0.3% of errors as anomalies.

This idea has since been widely adopted in outlier detection and anomaly detection research. (Chandola, Banerjee and Kumar, 2009) warns that if errors are non-Gaussian or heavy-tailed, $\mu + 3\sigma$ can be too strict.

The $\mu + 3\sigma$ threshold was therefore computed as:

Equation 1

$$\text{Threshold} = \mu_{\text{non-EQ}} + 3\sigma_{\text{non-EQ}}$$

Where $\mu_{\text{Non-EQ}}$ is the mean reconstruction error of the non-EQ training set, $\sigma_{\text{non-EQ}}$ is the standard deviation of the reconstruction error distribution in the non-EQ training set.

Data-driven thresholds

To address these limitations, we also explored data-driven thresholds, including:

- The 95th percentile (q95) of non-EQ training errors, which relaxes the strictness of $\mu + 3\sigma$ by flagging the top 5% of highest-error non-EQ segments as anomalous.

- A supervised threshold derived from the labelled test set using the ROC-optimal (Youden's J) criterion, which maximises the trade-off between true positive rate (TPR) and false positive rate (FPR).

This dual approach enabled evaluation under both unsupervised (q95) and supervised (ROC-optimal) settings, reflecting performance in practical vs analysis-oriented scenarios.

2.4 Evaluation Protocol

Evaluation combined both reconstruction-based anomaly scoring and latent-space analysis:

- **Reconstruction error metrics:** Precision, recall, F1, ROC-AUC, confusion matrices; histograms of EQ vs non-EQ error distributions with threshold marked; bootstrap confidence intervals for reconstruction error means to quantify uncertainty.
- **Bottleneck feature analysis:** Mean feature difference (EQ vs non-EQ), statistical tests, Random Forest (RF) feature importance, and waveform visualisations linked to top bottleneck neurons.
- **Dimensionality reduction:** t-SNE and UMAP projections of bottleneck features, with silhouette scores for separation.
- **Supervised probe:** RF classifier trained on bottleneck features with balanced class weights, testing whether supervised separation exists even if unsupervised scatter plots look mixed
- **Human activity trend analysis:** RF probability of EQ-like background segments aggregated across station types (rural, urban, industrial), analysed for hourly/weekday trends, station ECDFs, and anomaly mapping.

Computation of EQ-likeness ($p(\text{EQ})$)

The Random Forest classifier outputs a probability score for each segment, denoted $p(\text{EQ})$, representing the likelihood that the segment resembles earthquake data in the 64-dimensional bottleneck feature space. This

probability is obtained by averaging votes across the ensemble: each decision tree assigns a segment to a terminal leaf, where the proportion of EQ-labelled training examples defines the tree’s vote. The final $p(\text{EQ})$ score is the mean of these votes across all trees, yielding a continuous measure of “EQ-likeness” between 0 and 1.

3. Results

3.1 Baseline Autoencoder

To evaluate the anomaly detection pipeline, we first measured the distribution of reconstruction errors on non-EQ background segments. The mean error was 0.9552 (standard deviation = 0.0668). Applying the conventional $\mu + 3\sigma$ thresholding rule gave a cut-off of 1.1556. This value was so conservative that no anomalies were detected (**see Figure 2**), illustrating the limitation of fixed statistical thresholds for earthquake detection.

A precision-recall sweep was therefore performed to identify the optimal threshold. This sweep determined that 0.9841 was the cut-off that maximised the F1-score.

Performance at this threshold is summarised in **Table 2**.

Table 2. Performance metrics of the baseline autoencoder at the optimal threshold (0.9841)

Metric	Value
ROC-AUC	0.5434
Precision	0.842
Recall	0.624
F1-score	0.717

Although precision was relatively high, recall was modest, and the overall ROC-AUC of 0.5434 indicates near-chance discrimination between EQ and non-EQ segments. This motivated a redesigned architecture with a more expressive bottleneck.

Figure 3 shows the baseline reconstruction error distribution for EQ and non-EQ segments. While EQ samples tended to incur higher errors on average, the distribution overlaps substantially, which explains the limited separability achieved by the autoencoder.

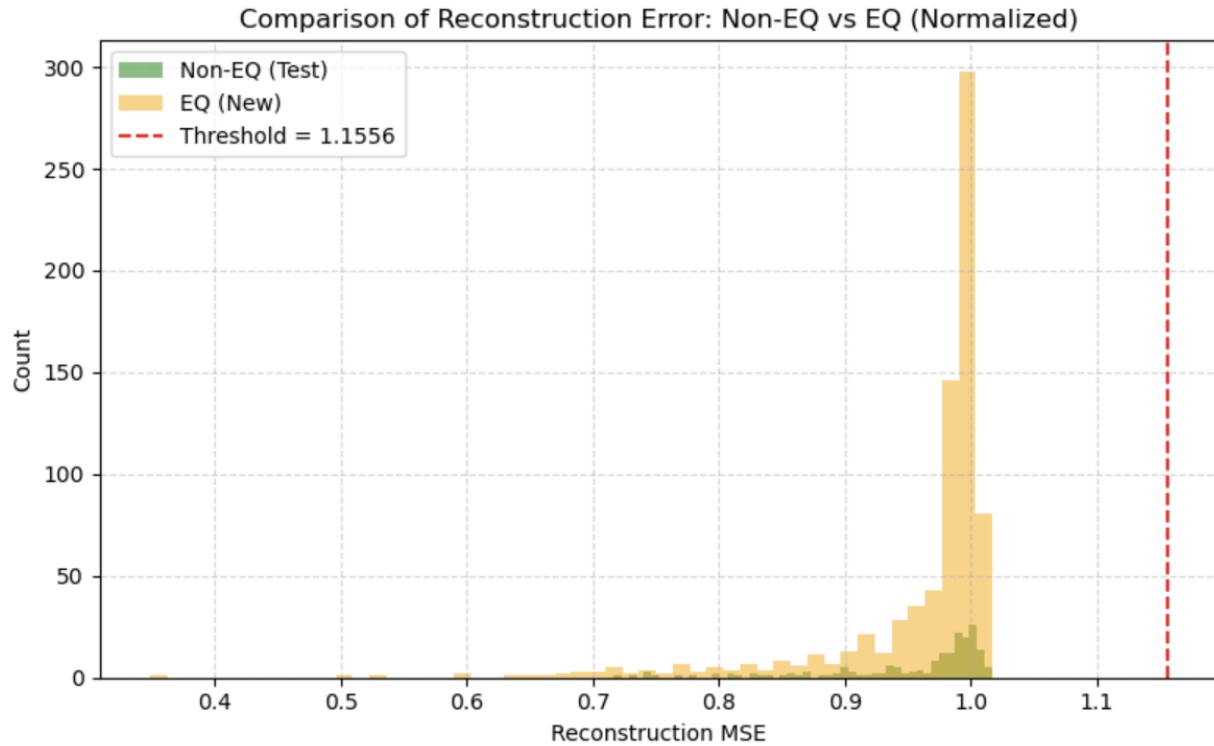


Figure 3. Baseline autoencoder reconstruction errors (non-EQ vs EQ) using the strict thresholding rule $\mu + 3\sigma = 1.1556$.

3.2 Redesigned Autoencoder and Extended Analysis

The redesigned autoencoder was implemented with a 64-dimensional bottleneck and trained until convergence at approximately 80 epochs. This model formed the main framework for all subsequent analyses, with its latent features (non_eq_X_train and non_eq_X_test) reused throughout the project.

3.2.1 Reconstruction Errors

To characterise the behaviour of reconstruction errors, we measured the mean squared error (MSE) for earthquake (EQ) and non-earthquake (non-EQ) test segments. The results are shown in **Table 3**.

Table 3. Reconstruction error comparison between EQ and non-EQ test data

Segment Type	Mean MSE	Std.Dev.
Non-EQ	0.7496	0.1690
EQ	0.8709	0.1225

These results show that EQ segments consistently had higher reconstruction errors than non-EQ segments, though there was substantial overlap between the two distributions.

3.2.2 Thresholding Outcomes

We evaluated detection performance under two thresholding strategies:

- **Unsupervised threshold (q95 = 0.9404).**

The threshold was set to the 95th percentile of non-EQ training errors and applied to the held-out non-EQ test set (n=183). Twenty-seven segments were flagged as anomalous (27/183 = 14.75% false alarms; specificity = 85.25%). This quantifies the background false-alarm rate under a purely label-free setting.

- **Supervised optimal threshold (ROC-optimal = 0.8497).**

Using the labelled mixed test set, the threshold was selected according to the ROC-optimal criterion (Youden's J), which maximises the difference between true positive rate (TPR) and false positive rate (FPR). This yielded an operating point at 0.8497.

From this, performance metrics were obtained: ROC-AUC = 0.7173, precision = 0.899, Recall = 0.704, and F1 = 0.790. On the non-EQ subset, this corresponds to a 33.3% false-alarm rate (61/183), higher than the q95 outcome, reflecting the trade-off whereby a supervised threshold improves recall of EQs but at the cost of increased false positives.

Table 4. Thresholding outcomes for EQ vs non-EQ test detection.

Threshold Strategy	Threshold Value	EQ Detection Rate	Non-EQ False Alarm Rate
q95 (unsupervised)	0.9404	34.5%	14.8%
ROC-optimal	0.8497	70.4%	33.3%

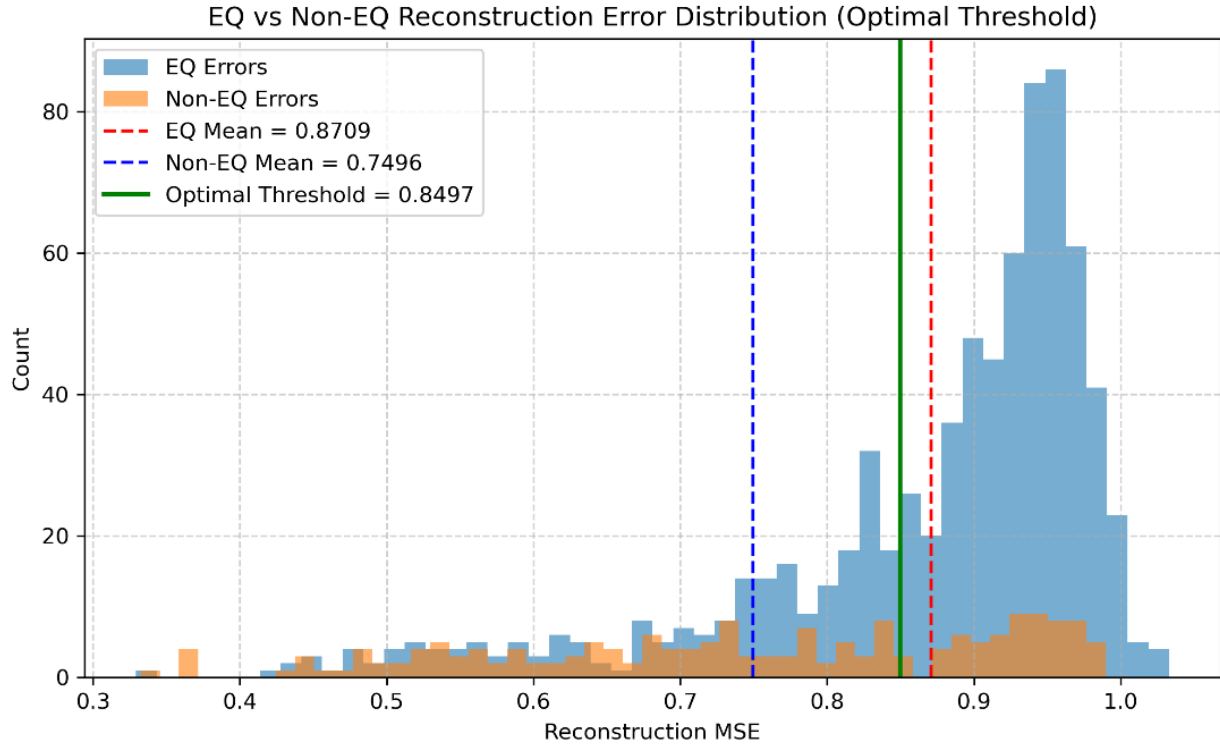


Figure 4. Improved autoencoder reconstruction error distributions under the supervised optimal threshold (0.8497). EQ errors shifted higher than non-EQ, yielding ROC-AUC = 0.7173, Precision = 0.899, Recall = 0.704

3.2.3 Bottleneck Feature Analysis

The 64-D latent space was interrogated to see whether EQ vs non-EQ separation emerged.

- Mean differences:** We compared the average activation of each of the 64 bottleneck dimensions. **Figure 5** shows the mean activation difference (EQ minus non-EQ) per feature. Some units were biased toward EQ (positive/red bars), while others skewed toward non-EQ (negative/blue bars). Importantly, the difference spanned a relatively narrow range (approximately -0.7 to +0.3) and formed a smooth gradient across features, rather than showing a few isolated spikes. This indicates that separation is distributed across many latent dimensions rather than concentrated in a small subset. In other words, the autoencoder encodes discriminative structure in the bottleneck, but EQ information is diffuse and intermixed with background variance rather than being cleanly isolated.

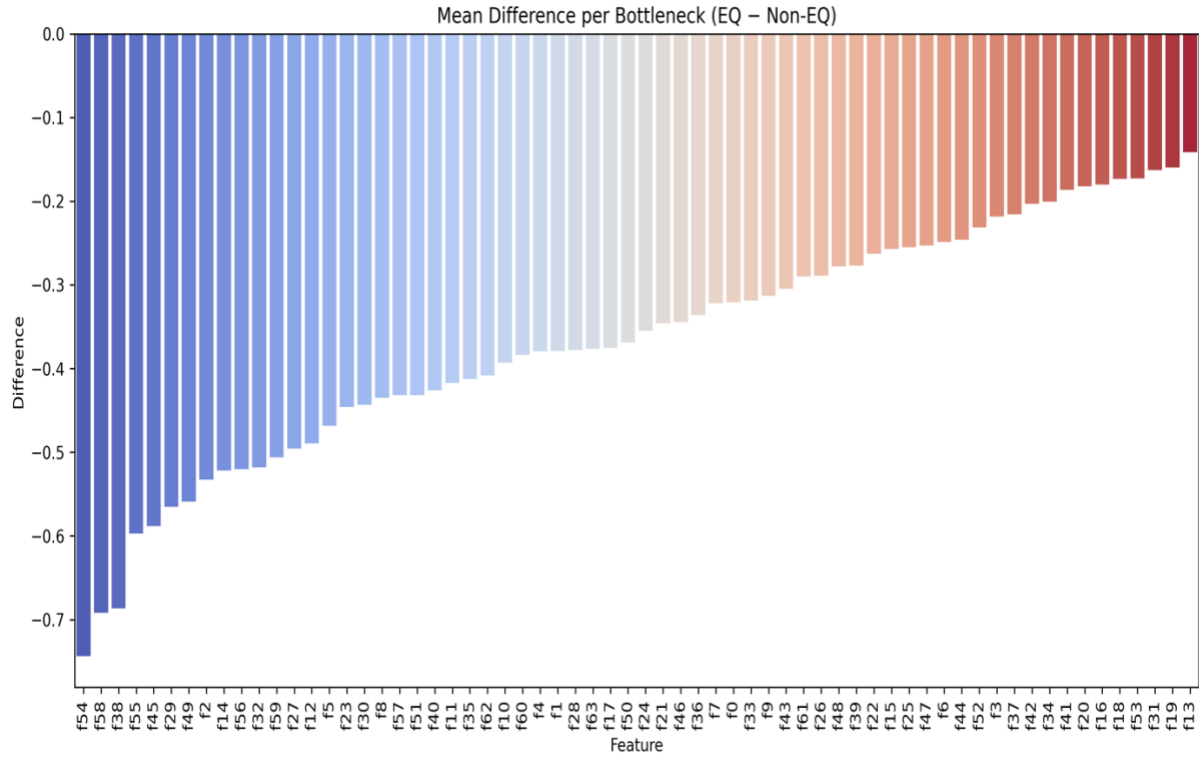


Figure 5. Mean activation difference (EQ minus non-EQ) across the 64 bottleneck dimensions. Negative values (blue) indicate higher non-EQ activations, while positive values (red) indicate higher EQ activations

Statistical Tests and Classifier Analysis: We combined statistical testing with supervised Random Forest (RF) classification. For each of the 64 bottleneck neurons, Welch’s t-test compared EQ vs non-EQ activations. Neurons with $p < 0.05$ were considered significantly different and marked in the plot as ▲ (EQ > non-EQ) or ▼ (non-EQ > EQ).

Figure 6 shows the autoencoder (AE) perspective: virtually all significant markers pointed downward (▼ blue), meaning non-EQ activations were consistently higher. No EQ-dominant neurons (▲ red) were observed. This indicated that earthquake information was not isolated in specific units but dispersed across the latent space, consistent with overlapping in the t-SNE plot (see **Figure 8**).

The RF classifier achieved Accuracy = 0.659, Precision = 0.549, Recall = 0.681, F1 = 0.608, and ROC-AUC = 0.706. RF feature inspection showed some neurons had ▼ blue markers (non-EQ > EQ), many showed no significance

(**Figure 7**), and none showed EQ dominance. Even when restricted to the top 20 most important features, separation remained diffuse.

AE and RF confirm bottleneck activations encode a weak discriminative signal. RF leveraged certain neurons for modest performance, but the absence of EQ-dominant activations and overlap indicates that EQ vs non-EQ separation remains entangled in the latent space.

The conventional significance level $\alpha = 0.05$ was applied, meaning results with $p < 0.05$ were marked significant. $\alpha = 0.05$ corresponds to accepting a 5% risk of falsely declaring a difference when none exists. Here, p denotes the probability of observing a difference at least as extreme if EQ and non-EQ means were equal. Since 64 neuron tests were run, some may be false positives; however, tests were used descriptively rather than for strict inference. Welch's t-test was chosen because it does not assume equal variances or sample sizes, making it more robust for heterogeneous seismic activations.

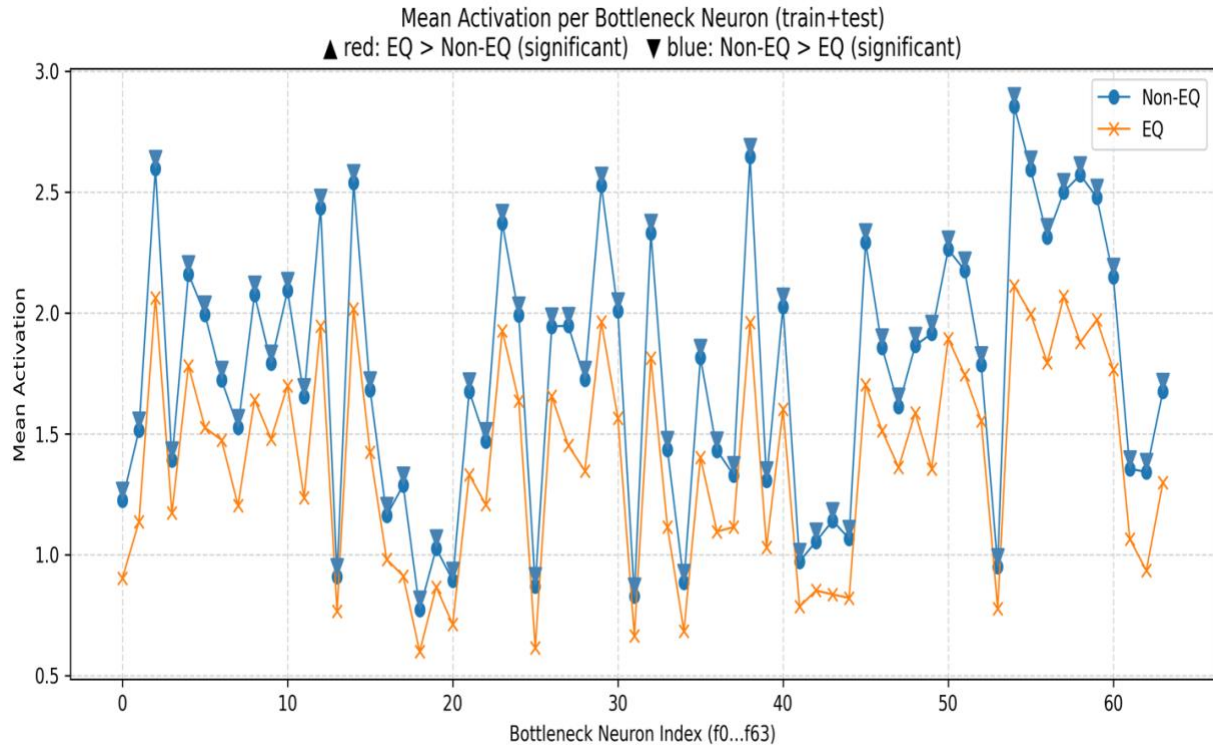


Figure 6. Mean bottleneck activations across the 64-dimensional latent space, comparing EQ and non-EQ segments. Downward arrows denote dimensions significantly skewed toward non-EQ showing that variance was spread rather than concentrated in specific latent unit.

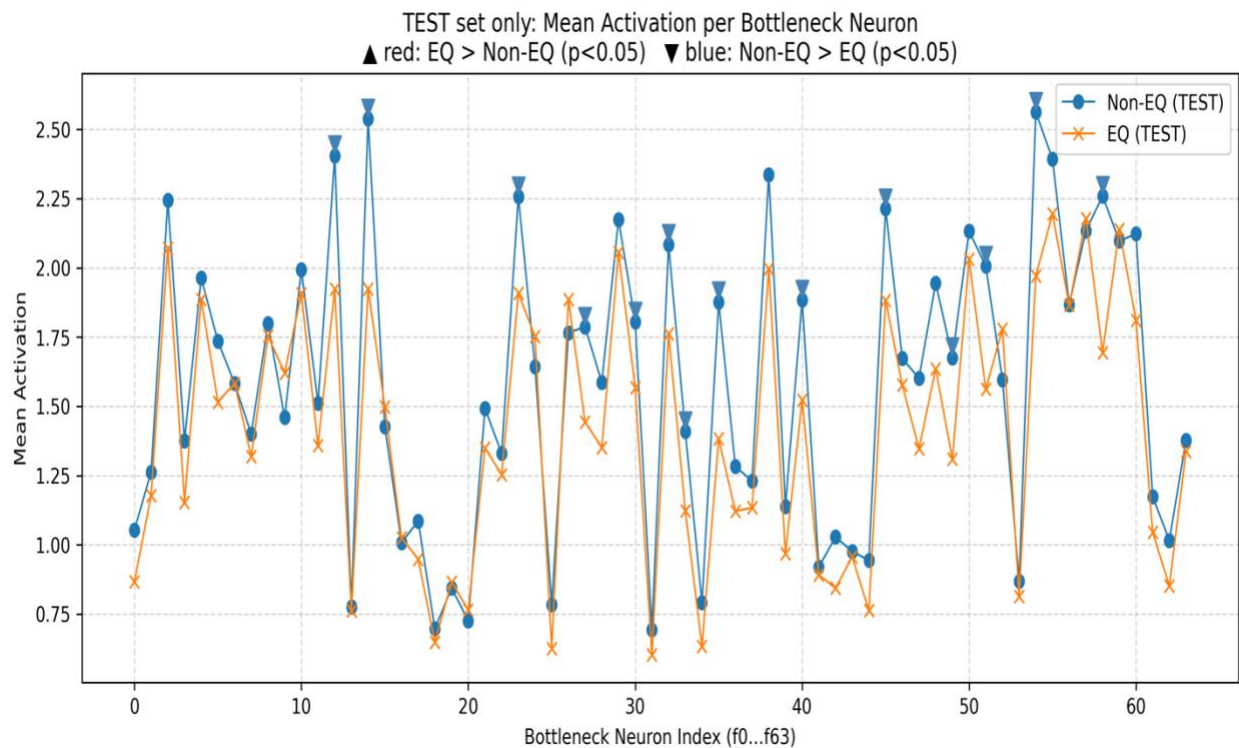


Figure 7. Mean bottleneck activations on the Test set only. Unlike Figure 6, not all latent neurons showed significant EQ vs non-EQ separation, indicating reduced variance and less skew in test-only activations.

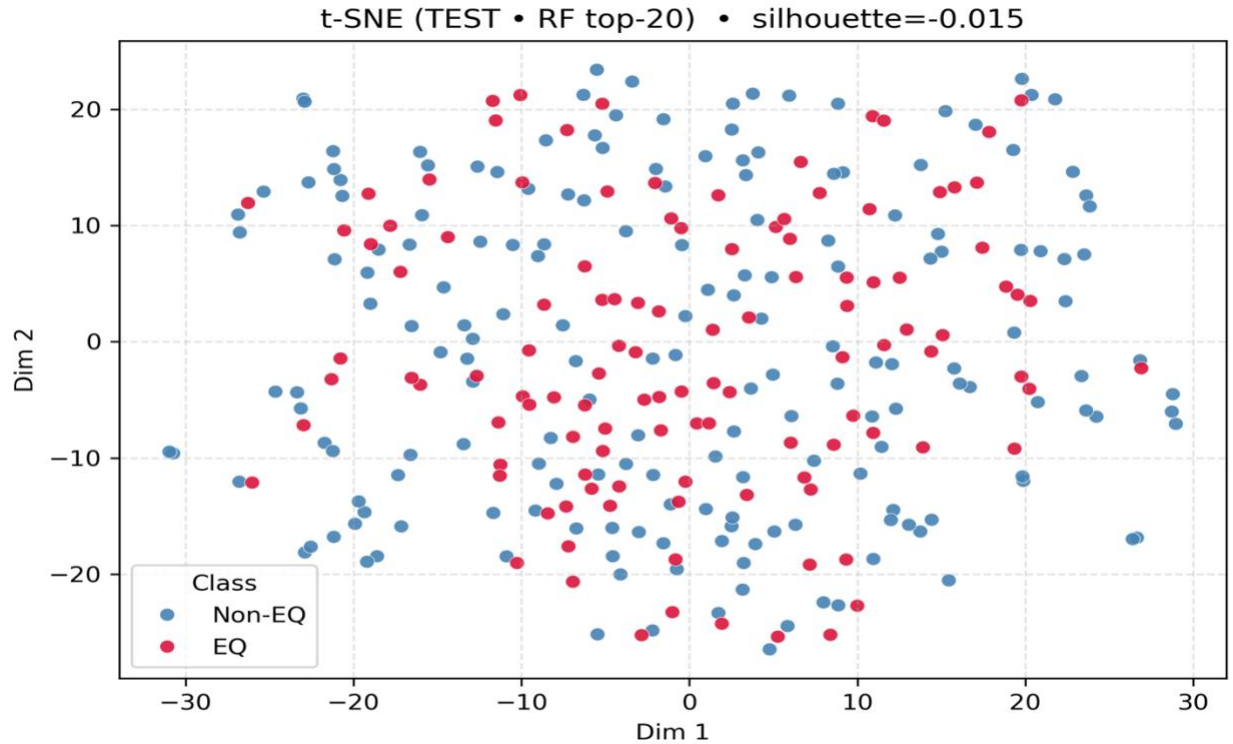


Figure 8. *t-SNE* plot showing the consistent overlapping structure in the latent space.

Waveform visualisations

To interpret what the Random Forest was learning, raw 5-second seismic waveforms were inspected for the neurons it ranked most important. For each top bottleneck neuron, the test samples with the highest activations were selected separately for non-EQ and EQ classes, and their raw traces were plotted (**Figure 9**).

Neuron f56 • RF importance=0.0330

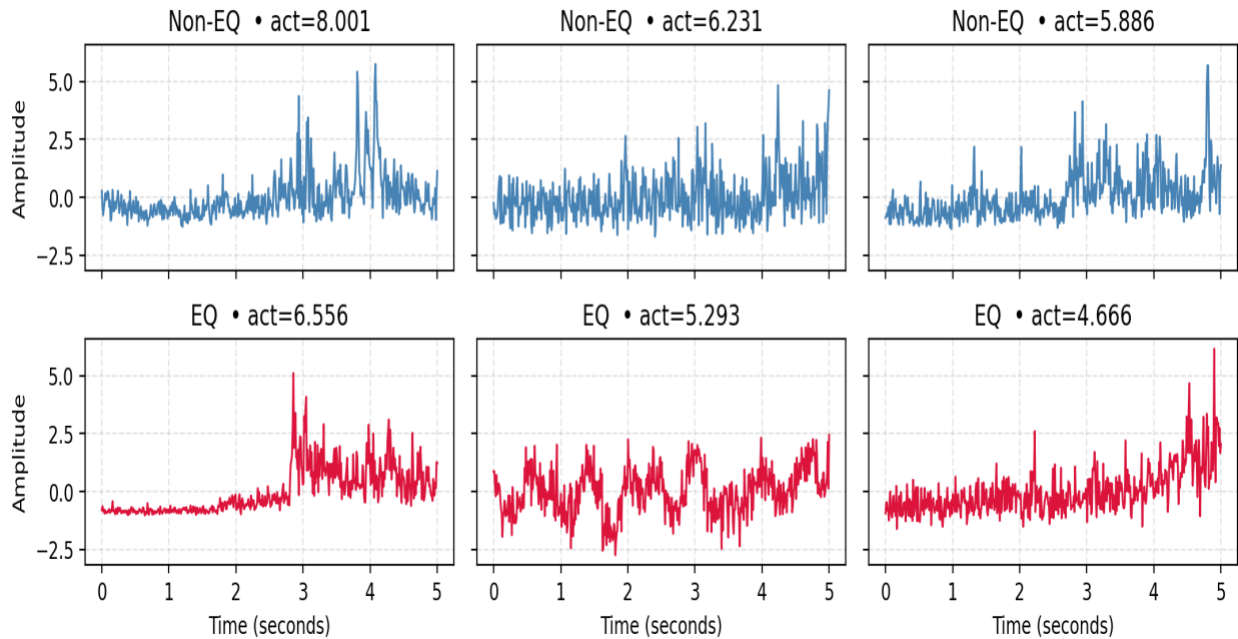


Figure 9. Waveform visualisation of top important features

These panels confirmed that the most influential neurons responded to transient bursts of energy and envelope fluctuations in the signal. However, such bursts were visible in both EQ and non-EQ traces, e.g., neurons f56 (Figure 8) showed strong activations for short-lived spikes regardless of class, suggesting that the RF was keying onto generic high-energy transients rather than earthquake-specific structure.

This analysis demonstrates that while the bottleneck features captured physically meaningful fluctuations, they were not uniquely earthquake signatures, which helps explain why classification performance remained moderate (ROC-AUC \approx 0.71).

3.2.5 Human Activity Trend Analysis + Methodological Justification

Exploratory analyses aggregated RF $p(\text{EQ})$ scores across non-EQ stations to probe potential anthropogenic patterns.

For each non-EQ segment, RF probabilities of being earthquake-like ($p(\text{EQ})$) were computed; in this section, EQ-like means higher $p(\text{EQ})$. These scores

were then aggregated to examine whether non-seismic background activity exhibited systematic anthropogenic patterns across stations and environments.

Methodological justification

To avoid relying on a single summary statistic, several complementary strategies were used:

- Medians and ECDF fractions were preferred over means to capture robust central tendencies and distributional spread, reducing the impact of outlier segments.
- Outlier detection (top $p(\text{EQ})$ segments) was necessary because anthropogenic signals are often sporadic (e.g., traffic bursts) and might only appear in extreme cases.
- Autoencoder (AE) reconstruction errors provided an orthogonal validation: while RF probed EQ-likeness, AE captured waveform complexity and reconstruction difficulty, highlighting environments with persistently hard-to-model anthropogenic signatures.

Station-level medians (computed from station-wise $p(\text{EQ})$ distributions)

Medians and ECDF fractions were calculated from all RF scores per station.

- Medians summarised the central tendency, robust to outliers.
- ECDF fractions quantified the proportion of segments with $p(\text{EQ}) \leq 0.10$ and ≤ 0.20 , indicating how much of the background was confidently classified as “non-EQ-like.”

Table 5. Exploratory analyses aggregated RF $p(\text{EQ})$ scores across non-EQ stations to probe potential anthropogenic patterns

Station	Environment	Median $p(\text{EQ})$	ECDF ≤ 0.10	ECDF ≤ 0.20
R57CE	Industrial/Urban	0.156	0.146	0.688
R3949	Rural	0.129	0.222	0.792
R7000	Rural	0.114	0.417	0.698
RCC8D	Rural	0.106	0.472	0.708
R0CEC	Rural	0.084	0.708	0.889

R76F0	Urban	0.139	0.264	0.736
R119F_a	Urban	0.115	0.319	0.819
RD9D6	Urban	0.100	0.528	0.875
R119F_b	Urban	0.085	0.625	0.833
R5B93	Urban	0.085	0.694	0.833
R1906	Urban/Industrial	0.134	0.292	0.722
R9F13	Urban/Industrial	0.107	0.430	0.875

R57CE emerged as the standout station with the highest systematic EQ-likeness (median = 0.156), consistent with its mixed industrial/urban setting. Urban stations such as R76F0 and R5B93 had lower medians but broader distributions, indicating sporadic bursts of highly EQ-like signals.

Outlier Top-scoring segments (obtained by sorting all non-EQ segments by $p(\text{EQ})$).

Segments were ranked by $p(\text{EQ})$ to identify outliers ($p(\text{EQ}) \sim 0.99$).

Table 6. List of top EQ-like non-EQ segments

p_eq_rf	station2	channel	source_file	seg_idx	seg_time
0.995	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	17	2023-05-24 22:29:45
0.995	R57CE	ENZ	R57CE_EQ_20240918T101939_10min_ENZ_segments.npy	11	2024-09-18 10:24:14
0.99	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	22	2024-09-18 10:28:49
0.985	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	6	2023-05-24 22:25:10
0.985	R7000	ENN	R7000_EQ_20250528T124439_10min_ENN_segments.npy	9	2025-05-28 12:48:24
0.9825	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	21	2023-05-24 22:31:25
0.975	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	16	2023-05-24 22:29:20
0.975	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	0	2023-05-24 22:22:40
0.97	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	1	2023-07-21 03:16:24
0.97	R57CE	EHZ	R57CE_EQ_20240918T101939_10min_EHZ_segments.npy	10	2024-09-18 10:23:49

0.9675	R1906	EHN	R1906_EQ_20230924T205322_10min_EHN_segments.npy	5	2023-09-24 20:55:27
0.955	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	23	2023-05-24 22:32:15
0.95	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	0	2024-09-18 10:19:39
0.9475	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	20	2024-09-18 10:27:59
0.935	R76F0	EHE	R76F0_EQ_trimmed_10min_EHE_segments.npy	18	2023-05-24 22:30:10
0.9325	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	14	2024-09-18 10:25:29
0.925	R57CE	ENZ	R57CE_EQ_20240918T101939_10min_ENZ_segments.npy	14	2024-09-18 10:25:29
0.9225	R76F0	EHZ	R76F0_EQ_trimmed_10min_EHZ_segments.npy	21	2023-05-24 22:31:25
0.9125	R7000	ENE	R7000_EQ_20250528T124439_10min_ENE_segments.npy	14	2025-05-28 12:50:29
0.9125	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	16	2024-09-18 10:26:19
0.91	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	13	2023-07-21 03:21:24
0.885	R57CE	ENZ	R57CE_EQ_20240918T101939_10min_ENZ_segments.npy	13	2024-09-18 10:25:04
0.8775	R76F0	EHN	R76F0_EQ_trimmed_10min_EHN_segments.npy	4	2023-05-24 22:24:20
0.8725	R9F13	EHN	R9F13_EQ_20190410T161508_10min_EHN_segments.npy	7	2019-04-10 16:18:03
0.855	R7000	ENN	R7000_EQ_20250528T124439_10min_ENN_segments.npy	19	2025-05-28 12:52:34
0.8525	R57CE	ENN	R57CE_EQ_20240918T101939_10min_ENN_segments.npy	19	2024-09-18 10:27:34
0.8525	R119F_a	EHE	R119F_EQ_20230921T132653_10min_EHE_segments.npy	18	2023-09-21 13:34:23
0.8325	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	10	2023-07-21 03:20:09
0.82	R7000	ENE	R7000_EQ_20250528T124439_10min_ENE_segments.npy	20	2025-05-28 12:52:59
0.7975	R1906	EHN	R1906_EQ_20230924T205322_10min_EHN_segments.npy	18	2023-09-24 21:00:52
0.7975	RD9D6	EHE	RD9D6_EQ_trimmed_10min_EHE_segments.npy	19	2023-10-13 09:04:35
0.7875	R1906	EHN	R1906_EQ_20230924T205322_10min_EHN_segments.npy	4	2023-09-24 20:55:02

0.7825	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	20	2023-07-21 03:24:19
0.78	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	2	2023-07-21 03:16:49
0.77	R3949	EHZ	R3949_EQ_20230721T031559_10min_EHZ_segments.npy	3	2023-07-21 03:17:14
0.7525	R7000	ENN	R7000_EQ_20250528T124439_10min_ENN_segments.npy	16	2025-05-28 12:51:19
0.735	R3949	EHE	R3949_EQ_20230721T031559_10min_EHE_segments.npy	5	2023-07-21 03:18:04
0.7325	R57CE	ENZ	R57CE_EQ_20240918T101939_10min_ENZ_segments.npy	0	2024-09-18 10:19:39
0.73	R3949	EHZ	R3949_EQ_20230721T031559_10min_EHZ_segments.npy	20	2023-07-21 03:24:19
0.7275	R9F13	EHE	R9F13_EQ_20190410T161508_10min_EHE_segments.npy	7	2019-04-10 16:18:03

These extreme outliers link the RF distributions to tangible waveform cases, often activating RF-important bottleneck units (e.g., neuron f56, RF importance = 0.0330). The associated waveform panels showed transient bursts/envelope spikes, and the same motif appears in high-p(EQ) non-EQ outliers. This demonstrates that short traffic-like bursts (Urban) and persistent machinery-like activity (Industrial/Urban) both produce false EQ-like patterns.

Autoencoder (AE) reconstruction errors

To cross-validate the RF results, mean reconstruction errors were computed per file (38 files total). AE errors capture waveform complexity: higher values suggest signals harder to reconstruct, potentially reflecting anthropogenic noise.

Table 7. Selected AE mean errors (top vs bottom performers).

source_file	Station	Channel	N_segments	Mean_AE_Error	Std_AE_Error
R57CE_EQ_20240918T101939_10min_ENN_segments.npy	R57CE	ENN	24	0.7900223731994630	0.12552009522914900
R57CE_EQ_20240918T101939_10min_ENZ_segments.npy	R57CE	ENZ	24	0.7630646824836730	0.11076847463846200

R76F0_EQ_trimmed_10min_EHE_segments.npy	R76F0	EHE	24	0.7496693134307860	0.14971689879894300
R57CE_EQ_20240918T101939_10min_ENE_segments.npy	R57CE	ENE	24	0.7285366058349610	0.08103068917989730
R76F0_EQ_trimmed_10min_EHZ_segments.npy	R76F0	EHZ	24	0.7213208675384520	0.06917767971754070
R76F0_EQ_trimmed_10min_EHN_segments.npy	R76F0	EHN	24	0.7157058119773870	0.08348492532968520
R3949_EQ_20230721T031559_10min_EHE_segments.npy	R3949	EHE	24	0.713252067565918	0.17530544102191900
R1906_EQ_20230924T205322_10min_EHN_segments.npy	R1906	EHN	24	0.6831576824188230	0.16280321776866900
R57CE_EQ_20240918T101939_10min_EHZ_segments.npy	R57CE	EHZ	24	0.6457803249359130	0.0997786894440651
R7000_EQ_20250528T124439_10min_ENN_segments.npy	R7000	ENN	24	0.6424209475517270	0.19282570481300400
R119F_EQ_20230921T132653_10min_EHE_segments.npy	R119F	EHE	24	0.6208991408348080	0.15157818794250500
R3949_EQ_20230721T031559_10min_EHN_segments.npy	R3949	EHN	24	0.6168969869613650	0.12231038510799400
R1906_EQ_20230924T205322_10min_EHZ_segments.npy	R1906	EHZ	24	0.5987613797187810	0.0946052297949791
R3949_EQ_20230721T031559_10min_EHZ_segments.npy	R3949	EHZ	24	0.584376871585846	0.14817632734775500
R9F13_EQ_20190410T161508_10min_EHN_segments.npy	R9F13	EHN	24	0.5809403657913210	0.1296617090702060
RD9D6_EQ_trimmed_10min_EHN_segments.npy	RD9D6	EHN	24	0.5766527056694030	0.11278531700372700
RD9D6_EQ_trimmed_10min_EHE_segments.npy	RD9D6	EHE	24	0.5712005496025090	0.10660089552402500
R5B93_EQ_trimmed_10min_EHZ_segments.npy	R5B93	EHZ	24	0.5710468888282780	0.13843971490860000
R7000_EQ_20250528T124439_10min_ENE_segments.npy	R7000	ENE	24	0.5699559450149540	0.20173479616642000
R9F13_EQ_20190410T161508_10min_EHE_segments.npy	R9F13	EHE	24	0.568406879901886	0.11996280401945100
RD9D6_EQ_trimmed_10min_EHZ_segments.npy	RD9D6	EHZ	24	0.5587136745452880	0.09962152689695360
R5B93_EQ_trimmed_10min_EHE_segments.npy	R5B93	EHE	24	0.5484936833381650	0.13668058812618300

R119F_EQ_20230921T132653_10min_EHN_segments.npy	R119F	EHN	24	0.5436755418777470	0.15761184692382800
R7000_EQ_20250528T124439_10min_ENZ_segments.npy	R7000	ENZ	24	0.5329992175102230	0.16650143265724200
R1906_EQ_20230924T205322_10min_EHE_segments.npy	R1906	EHE	24	0.5329487919807430	0.0979420468211174
R9F13_EQ_20190410T161508_10min_EHZ_segments.npy	R9F13	EHZ	24	0.5283544659614560	0.10848435014486300
RCC8D_EQ_trimmed_10min_EHZ_segments.npy	RCC8D	EHZ	24	0.5110543370246890	0.12356926500797300
RCC8D_EQ_trimmed_10min_EHN_segments.npy	RCC8D	EHN	24	0.4994727075099950	0.16694967448711400
RCC8D_EQ_trimmed_10min_EHE_segments.npy	RCC8D	EHE	24	0.4970572888851170	0.12945610284805300
R119F_EQ_20231017T080517_10min_EHE_segments.npy	R119F	EHE	24	0.4960600435733800	0.1529081165790560
R7000_EQ_20250528T124439_10min_EHZ_segments.npy	R7000	EHZ	24	0.4836767613887790	0.17709927260875700
R119F_EQ_20230921T132653_10min_EHZ_segments.npy	R119F	EHZ	24	0.4687519073486330	0.13727118074894000
R119F_EQ_20231017T080517_10min_EHZ_segments.npy	R119F	EHZ	24	0.4518754482269290	0.13512133061885800
R119F_EQ_20231017T080517_10min_EHN_segments.npy	R119F	EHN	24	0.43855020403862	0.13460783660411800
R5B93_EQ_trimmed_10min_EHN_segments.npy	R5B93	EHN	24	0.41875991225242600	0.11812558025121700
R0CEC_EQ_trimmed_10min_EHE_segments.npy	R0CEC	EHE	24	0.3548556864261630	0.1055373027920720
R0CEC_EQ_trimmed_10min_EHZ_segments.npy	R0CEC	EHZ	24	0.35337892174720800	0.08853407949209210
R0CEC_EQ_trimmed_10min_EHN_segments.npy	R0CEC	EHN	24	0.3481728136539460	0.12438604235649100

R57CE consistently produced the highest AE errors across multiple channels, reinforcing its RF profile as the most systematically EQ-like non-EQ site. By contrast, rural/industrial sites such as R0CEC showed much lower errors, confirming their waveforms are easier to reconstruct and less confusable with earthquakes.

Both models converge on a nuanced picture: anthropogenic activity leaves distinct seismic fingerprints that vary by station environment, detectable through anomaly-detection pipelines.

3.3 Kamchatka (M8.8) and Drake Passage (M7.5)

To test whether extreme earthquakes improved separability, ~900 EQ segments (Kamchatka = 540, Drake = 360) were compared against non-EQ. Using the trained AE, two thresholds were applied: q95 = 0.7345, corresponding to the 95th percentile of non-EQ training errors (a fixed baseline to reject normal activity), and best-F1 = 0.7379, identified by sweeping thresholds on the test data to maximize F1 (balancing false positives and recall). EQ errors were slightly higher (mean ratio = 1.14) but overlapped with non-EQ. AE reached ROC-AUC = 0.691; Random Forest improved ROC-AUC to 0.749 (accuracy ~ 0.70). t-SNE showed no distinct clusters.

Table 8 shows the confusion matrices for EQ vs. non-EQ classification.

Method	Threshold	Non-Eq Correct	Non-Eq Wrong	EQ Correct	EQ Wrong	Accuracy
Autoencoder(q95)	0.7345	80	103	151	29	0.636
Autoencoder (Best-F1)	0.7379	82	101	151	29	0.642

Interpretation

Even with stronger events, the AE bottleneck did not yield a “day vs night” separation. The system detected EQs more reliably than chance but failed to provide the clear margin expected from such high magnitudes. This reinforced the conclusion: AE anomaly detection captures statistical differences but is fundamentally limited by overlap with noisy non-EQ data.

4. Discussion

4.1 What the autoencoder achieved

The aim was to test whether an autoencoder (AE) trained on background noise could detect earthquake (EQ) waveforms as anomalies. EQ segments had higher reconstruction errors (mean = 0.8709, σ = 0.1225) compared with non-EQ (mean = 0.7496, σ = 0.1690). The supervised ROC-optimal threshold (0.8497) delivered ROC-AUC = 0.7173, Precision = 0.899, Recall = 0.704, F1 = 0.790. On clean non-EQ test data, the unsupervised q95 baseline (0.9404) flagged ~15% false alarms (specificity = 85.25%). These results confirm AE separates the two distributions but overlap remains.

Notebook 3 extended the evaluation to extreme events (M8.8 Kamchatka, M7.5 Drake Passage). Even at these magnitudes, AE performance plateaued (ROC-AUC ~ 0.691), with marginal improvements using a best-F1 threshold (0.7379). Random Forest (RF) classification on bottlenecks performed better (ROC-AUC = 0.749, Accuracy ~ 0.70). Event size did not yield perfect separability: many EQ windows contained quiet or coda phases, while anthropogenic noise often resembled arrivals. This aligns with prior findings that unsupervised reconstructions alone rarely achieve phase-level discrimination (Perol et al., 2018; Zhu & Beroza, 2019).

4.2 Why separation remained imperfect

Overlap is explained by:

1. **Station heterogeneity.** Urban/industrial sites generated transient bursts that mimic EQ onsets. Rural sites produced cleaner separation.
2. **Segment-level labels.** Each 5-second slice was labelled “EQ” or “Non-EQ” even if only part of it contained seismic energy.
3. **Model family.** Dense AE on binned amplitudes captures envelopes but not spectral/temporal patterns. CNN or spectrogram encoders separate better because they exploit shift and frequency invariances.
4. **Thresholding limits.** A global threshold collapses error distributions into a single decision boundary, trading recall against specificity.

These are not mistakes in design but expected consequences of adopting an unsupervised, reconstruction-based approach on heterogeneous seismic data.

4.3 Insights from the latent space and RF probe

Bottleneck analysis and RF classification revealed that “EQ-ness” is distributed across multiple latent units. Mean-difference plots showed neurons skewed toward non-EQ, while RF importance highlighted features tied to transient energy. This indicates that the AE latent space encodes seismic information, though not linearly separable (**see Figure 8**).

t-SNE projections reinforced this: clusters were visible but overlapped.

4.4 Demonstration of methodology vs physical conclusions

RF distributions and AE cross-checks show background waveforms at some stations exhibit “EQ-like” properties. Industrial/urban sites sustained elevated median $p(\text{EQ})$ values and higher reconstruction errors, while rural sites were easier to reconstruct. Outlier from composition checks confirmed spikes were genuine patterns, not sampling bias.

What this demonstrates is the methodology: distributions, outlier linkage to waveform panels, and AE validation converge on the same narrative that anthropogenic environments leave EQ-like imprints. However, because the non-EQ set came from short, catalog-filtered windows rather than continuous monitoring, these results reflect the model's perception rather than true diurnal or weekly human cycles. Robust physical conclusions would require long-term, uninterrupted background recordings.

4.5 Strengths, limitations, and relation to prior work

Strengths

- **Reproducible pipeline:** consistent segmentation, z-normalisation, and strict train/test splits avoided leakage.
- **Transparent evaluation:** ROC/PR curves, threshold sweeps, and latent-space probes provided multiple views of performance.
- **Proof-of-concept:** Demonstrated that even a simple dense AE can separate EQ and Non-EQ distributions beyond chance.

Limitations

- **Data imbalance.** 768 EQ vs 912 non-EQ in main set; imbalance influenced thresholds.

- **Temporal bias.** Randomly chosen non-EQ windows do not cover all hours or weekdays.
- **Model expressivity.** Dense AE underuses spectral/phase information compared to CNN/LSTM encoders.
- **Single-threshold assumption.** A global cutoff is unlikely to work equally well across diverse stations.

Compared with prior literature, the ROC-AUC ~ 0.72 is competitive for a purely unsupervised setting, though still below supervised CNN/Transformer detectors (AUC > 0.9 in Perol et al., 2018; Zhu & Beroza, 2019). This positions the project as a proof-of-concept demonstration of feasibility under constrained design choices, rather than a fully deployable detector.

4.6 Future directions

1. **Continuous non-EQ data acquisition.** Weeks to months of uninterrupted background recordings are needed to eliminate temporal bias.
2. **Temporal aggregation.** Hourly/daily/weekly features, analysed with decomposition methods (e.g., STL), to reveal anthropogenic cycles.
3. **Model refinement.** CNN/LSTM or variational autoencoders could better capture phase morphology and spectral cues.
4. **Station-specific calibration.** Local thresholds or domain-adapted models would account for site noise differences.
5. **Dual objective.** Combining reconstruction with supervised classification (RF, CNN) may yield sharper EQ/Non-EQ boundaries.

5. Conclusion

This project tested whether an autoencoder (AE) trained on non-earthquake (non-EQ) Raspberry Shake data could detect earthquakes as anomalies. Using 20 stations across varied environments, 5-s segments were z-normalised and evaluated with AEs and Random Forests (RF) on bottleneck features. EQs showed higher reconstruction errors (mean ratio ≈ 1.14) but overlap persisted. The improved AE achieved ROC-AUC ~ 0.72 (F1 ~ 0.79), while RF reached 0.70–0.75. Large events (Kamchatka M8.8, Drake M7.5)

showed only modest gains. Future work should expand labelled data, explore CNN/LSTM/Transformer models, apply station-specific thresholds, and use Hidden Markov Models (HMMs) as probabilistic models for sequential patterns to reduce false positives.

Reference

- Allen, R. (1982) 'Automatic phase pickers: Their present use and future prospects', *Bulletin of the Seismological Society of America*, 72(6B), pp. S225-S242.
- Bergen, K. J., Johnson, P. A., de Hoop, M. V. and Beroza, G. C. (2019) 'Machine learning for data-driven discovery in solid Earth geoscience', *Science*, 363(6433), pp. eaau0323.
- Chalapathy, R. and Chawla, S. (2019) 'Deep learning for anomaly detection: A survey', *arXiv preprint arXiv:1901.03407*.
- Chandola, V., Banerjee, A. and Kumar, V. (2009) 'Anomaly detection: A survey', *ACM computing surveys (CSUR)*, 41(3), pp. 1-58.
- Hinton, G. E. and Salakhutdinov, R. R. (2006) 'Reducing the dimensionality of data with neural networks', *science*, 313(5786), pp. 504-507.
- Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J. and Gerstoft, P. (2018) 'Machine Learning in Seismology: Turning Data into Insights', *Seismological Research Letters*, 90(1), pp. 3-14.
- Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P. and Shroff, G. (2016) 'LSTM-based encoder-decoder for multi-sensor anomaly detection', *arXiv preprint arXiv:1607.00148*.
- Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. and Beroza, G. C. (2020) 'Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking', *Nature communications*, 11(1), pp. 3952.
- Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E. and Kloft, M. 'Deep one-class classification'. *International conference on machine learning*: PMLR, 4393-4402.
- Sakurada, M. and Yairi, T. 'Anomaly detection using autoencoders with nonlinear dimensionality reduction'. *Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis*, 4-11.
- Shewhart, W. A. (1931) 'Economic control of quality of manufactured product'.
- Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S. and Trujillo, J. (1998) 'A comparison of select trigger algorithms for automated global seismic phase and event detection', *Bulletin of the Seismological Society of America*, 88(1), pp. 95-106.
- Zhou, C. and Paffenroth, R. C. 'Anomaly detection with robust deep autoencoders'. *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 665-674.
- Zhu, W. and Beroza, G. C. (2019) 'PhaseNet: a deep-neural-network-based seismic arrival-time picking method', *Geophysical Journal International*, 216(1), pp. 261-273.
- Zong, B., Song, Q., Min, M. R., Cheng, W., Lumezanu, C., Cho, D. and Chen, H. 'Deep autoencoding gaussian mixture model for unsupervised anomaly detection'. *International conference on learning representations*.
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3), 530–533.

Raspberry Shake (2023). Raspberry Shake Developers Corner – FDSN Web Services API.
Retrieved from <https://manual.raspberryshake.org/developer.html>