Machine Learning Classification for Acoustic Mix Identification in Styrofoam Lightweight Concrete: A Feature Sensitivity Study

Article history

Received:
2 April 2026

Received in revised form:
12 April 2026

Accepted:
5 June 2026

Published online:
15 June 2026

*Corresponding author
ngudi_haricrista@usm.ac.id

Ngudi Hari Crista[1], Sudharto P. Hadi[2], Erni Setyowati[3], Aimmatul Husna[4]

[1,2,3] *Faculty of Environmental Science and Engineering, Universitas Diponegoro, Indonesia*
[4]*Faculty of Engineering, Universitas Semarang, Indonesia*

ngudi_haricrista@usm.ac.id[1], sudhartophadi@yahoo.co.id[2], ernisyahdu@gmail.com[3], aimmatul@usm.ac.id[4]

Abstract

Concrete has long been the backbone of modern construction, yet the demands placed on building materials have changed considerably. Strength and structural integrity remain essential, but acoustic performance, how effectively a structure manages sound, has become a serious design consideration. This study takes that challenge as its starting point, examining Styrofoam-based lightweight concrete in which expanded polystyrene beads partially replace fine aggregate to improve sound absorption. Four concrete mixtures (FS, ST1, ST2, and ST3), were prepared with progressively increasing Styrofoam content and characterised for compressive strength, flexural strength, and sound absorption coefficient (α) using the impedance tube method (ASTM E1050-12). A dataset of 312 records was compiled from these four unique material compositions, each tested across 13 frequency levels (100–1600 Hz) and six acoustic panel configurations.

Rather than predicting continuous acoustic absorption values, this study focuses on identifying between concrete mix types based on their measured material and acoustic characteristics using machine learning (ML) classification models. The acoustic behaviour of Styrofoam-based lightweight concrete is shaped by the nonlinear interaction of pore geometry, bead size distribution, and frequency-dependent viscous losses relationships that conventional empirical models cannot easily capture. ML offers a data-driven route to material identification that sidesteps those modelling constraints. Four algorithms k-Nearest Neighbours (k-NN), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Logistic Regression (LR), were evaluated using 10-fold cross-validation in Orange Data Mining 3.40.0, through two complementary analyses: Analysis 1 with the full feature set, and Analysis 2 restricted to acoustic features only.

In Analysis 1, ANN achieves perfect CA = 1.000 and AUC = 1.000, while SVM reaches CA = 0.923. However, both k-NN and Logistic Regression return CA = 0.250, identical to random chance, due to the absence of internal feature normalisation in Orange. A feature sensitivity analysis confirms that ANN's perfect result is mechanistically explained by class-unique compositional features, not acoustic learning. In Analysis 2, removing those identifying features drops all models to or near random chance (CA = 0.21–0.25; balanced baseline = 0.25). A regression baseline confirms that continuous α prediction is not currently feasible ($R^2 \leq -0.32$). The main contribution lies not in the accuracy numbers but in the dual-analysis evaluation framework, a transparent approach that clearly separates what ML can and cannot do with small experimental datasets. Experimentally, ST2 (1:1.5:2.5) emerged as the most balanced mix.

Keywords: *Styrofoam, Acoustic concrete, Machine learning classification, Feature sensitivity, Orange Data Mining, Exploratory study*

* *Corresponding author. ngudi_haricrista@usm.ac.id*

1. Introduction

Concrete is not going anywhere, it remains the material that modern construction is built on, and that is unlikely to change. But what buildings need from their materials has shifted. Loading capacity and durability are still the foundation, yet today's buildings are also expected to manage how sound behaves inside them, filtering what moves through walls, floors, and ceilings in ways that affect the people living and working there. This has made acoustic performance a genuine design consideration, sitting alongside the traditional concerns about strength and cost [1,2]. Lightweight cementitious composites have emerged as one promising response to these dual demands, offering the possibility of reducing material weight while repurposing industrial and consumer waste [3].

The case for better acoustic insulation goes beyond comfort. Sustained exposure to noise in residential and working environments has been linked to elevated stress, disrupted sleep, and measurable cognitive effects [4]. Against this backdrop, materials that can absorb sound effectively without sacrificing structural performance hold real practical appeal. Lightweight concretes, have attracted steady research interest for their ability to combine these properties, thanks largely to the pore networks created by low-density aggregate substitutes [5,6]. Expanded polystyrene (Styrofoam) is one such substitute lightweight, widely available as waste, and capable of modifying the internal void structure of concrete in ways that tend to improve acoustic behaviour [2,7].

Machine learning has become a well-established tool for concrete property prediction, with most existing work framed as regression problems targeting mechanical strength [8,9]. The present study takes a different approach. The goal here is not to predict α as a continuous value, but to determine whether ML classifiers can identify which mix type a given set of measurements belongs to a material identification task, not a property prediction one. In this framing, α is an input feature rather than an output target, and the practical application is something closer to quality control or mix verification than design optimisation. This formulation is consistent with classification-based approaches in construction materials research [8,10].

The motivation for using ML specifically rests on a characteristic of the problem that makes conventional approaches difficult. The acoustic behaviour of Styrofoam-based lightweight concrete is governed by the complex interaction of pore geometry, bead size distribution, and frequency-dependent viscous losses relationships that are inherently nonlinear and hard to capture through simple empirical equations or linear regression [10]. Traditional acoustic characterization relies on repeated physical testing across many configurations, which is both time-consuming and resource intensive. Machine learning offers a data-driven alternative: rather than assuming a functional form for the acoustic-composition relationship, ML models learn directly from experimental measurements. In this study, the classification formulation is specifically motivated by a practical need verifying mix identity from a combination of routinely measured properties without requiring a new impedance tube test for each specimen. This is a useful capability for quality assurance in lightweight concrete production, where rapid material identification could reduce dependence on specialized acoustic instrumentation [8,10].

Four algorithms k-NN, SVM, ANN, and Logistic Regression are applied within Orange Data Mining [11]to give a structured comparison across distance-based, margin-based, neural-network, and linear modelling paradigms [8], [10]. To keep the evaluation honest, two analyses run in parallel: one using the full feature set, one limited to acoustic inputs only. A feature sensitivity analysis and a regression baseline are included to make the results interpretable rather than simply impressive. The study's conceptual framework is shown in Figure 1.

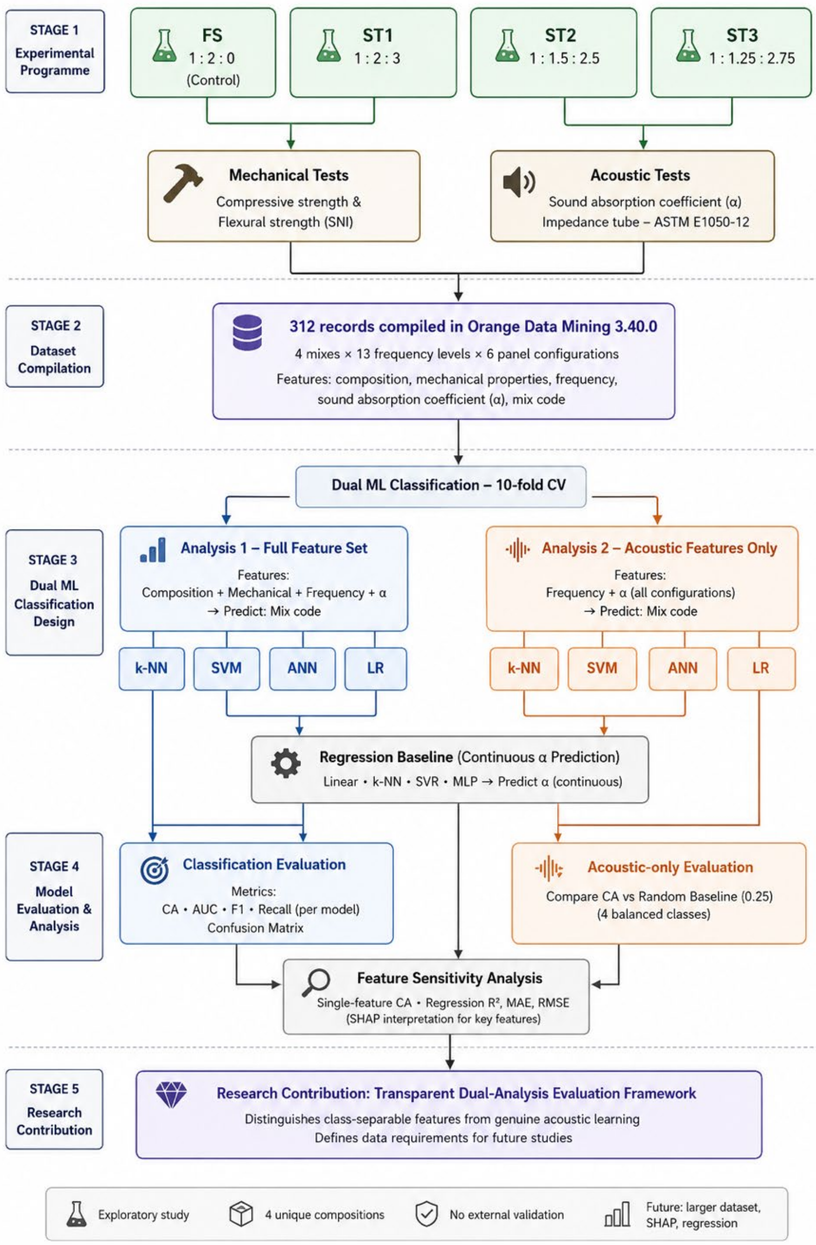

Figure 1. Conceptual framework of the ML classification study for Styrofoam-based lightweight concrete

The study has three specific aims: to experimentally characterize the acoustic and mechanical performance of four Styrofoam-based concrete mixtures, to evaluate whether ML classification can reliably distinguish between mix types under two feature conditions full features and acoustic-only, and to assess the current feasibility of continuous acoustic property prediction through a regression baseline. Together, these aims are designed to produce findings that are honest about both what ML can achieve and where the limits of a small experimental dataset begin.

2. Related Work

2.1 Acoustic Performance of Cement-Based and Lightweight Concrete

Interest in the acoustic properties of cement-based materials has grown alongside rising urban noise levels and increasingly demanding comfort standards for buildings. The general finding across the literature is that lightweight concretes by virtue of their lower density and more porous internal structure tend to outperform conventional dense concrete in sound absorption [12,13]. The physical explanation is relatively straightforward: sound waves entering a porous material lose energy through viscous friction as they travel through air-filled pore channels, and through resonance effects that depend on pore geometry [6]. What is harder to predict is how these mechanisms interact across a range of frequencies and mix configurations, which is part of what makes ML an attractive modelling tool for this domain [14]. Acoustic performance is strongly influenced by multiple mix-related parameters, creating nonlinear material behaviour that conventional empirical approaches find difficult to capture [14].

2.2 Styrofoam as a Sustainable Acoustic Aggregate

The case for incorporating Styrofoam into concrete rests on two arguments. The first is practical: EPS beads are lightweight and produce a network of internal voids that modifies the acoustic and thermal behaviour of the resulting composite [15,16]. The second is environmental: Styrofoam is generated in large volumes as packaging waste and incorporating it into construction materials offers a route to diverting it from landfill [16]. Studies consistently find that Styrofoam content is a sensitive parameter pushing it too high tends to compromise mechanical strength and can produce an irregular pore network that is acoustically less stable than intended [15]. Finding the right substitution level is therefore a balancing act between acoustic gain and structural penalty [17].

2.3 Machine Learning in Concrete Property Prediction

ML has become a standard analytical tool in concrete research, with applications ranging from predicting compressive strength to optimizing mix design [8,9]. Among comparative studies, Sun et al. [18] provide a useful benchmark for multi-property prediction using ANN, SVM, and Random Forest on UHPC, confirming that nonlinear models consistently outperform linear ones. Mohtasham Moein et al. [10] offer a comprehensive review of ML and deep learning approaches for concrete properties, noting that model performance is highly sensitive to dataset quality and feature selection. For acoustic property prediction specifically, ML applications in cementitious composites remain limited [14], leaving a clear opening for studies that combine acoustic characterization with transparent ML evaluation.

2.4 Where This Study Fits

What is notably absent from the existing literature is a study that combines Styrofoam-based lightweight concrete with a comparative ML classification framework, a feature sensitivity analysis, and a regression baseline all within a single reproducible workflow. The present study is designed to fill that gap. A summary of related studies is given in Table 1.

Table 1. Comparison of recent ML-based studies on concrete property prediction

Study	Material	Target	ML Methods	Acoustic	Comparative
Galip et al. [12]	Porous or lightweight concrete	Sound absorption	-	Yes	No
Kang et al. [13]	Lightweight concrete	Acoustic absorption behaviour	-	Yes	No
Amran et al. [14]	Cement-based and lightweight concrete	Acoustic performance	-	Yes	No
Mohtasham Moein et al. [10]	Cement-based materials	Acoustic performance prediction	ML-based modelling	Yes	No
Sun et al. [18]	Cementitious materials	Multi-property prediction	ANN, SVM, RF	No	Yes
Crista et al. [15]	Styrofoam lightweight concrete	Acoustic performance	-	Yes	No
Elghomari & Tilioua [16]	EPS-based lightweight concrete	Sustainability	-	No	No
Patrisia et al. [19]	Waste - based concrete composites	Thermo-acoustic performance	-	Yes	No
This study	Styrofoam-lightweight concrete	Acoustic classification	k-NN, SVM, ANN, LR	Yes	Yes

3. Methodology

3.1 Material Preparation and Testing

The dataset used in this study was built from laboratory experiments on four Styrofoam-based lightweight concrete mixtures. The mixture (FS) used no Styrofoam, while ST1 (1:2:3), ST2 (1:1.5:2.5), and ST3 (1:1.25:2.75) substituted progressively larger proportions of fine aggregate with EPS beads ratios expressed as Portland cement, sand, and styrofoam by volume. Figures 2 and 3 show the mixing and mechanical testing setups. Compressive strength tests followed SNI 1974:2011 on cylindrical specimens and flexural strength tests followed SNI 4431:2011 on beam specimens. Acoustic performance was measured using the impedance tube method per ASTM E1050-12 [20].

Figure 2. Concrete mixing procedure for Styrofoam-based lightweight concrete specimens

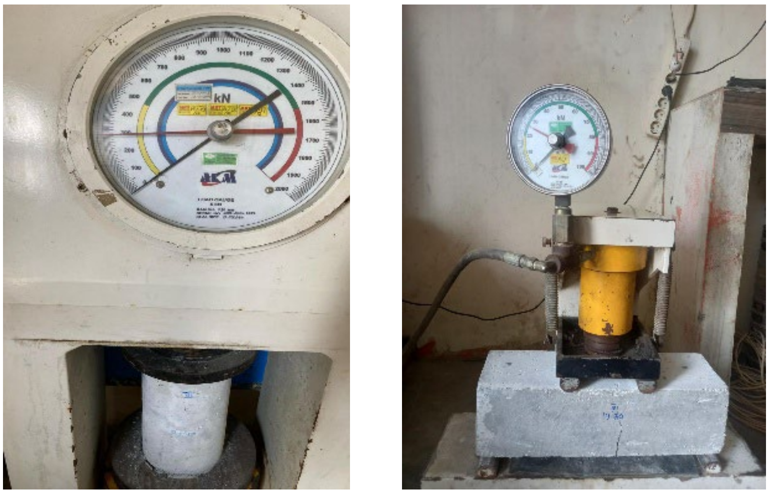

Figure 3. Compressive and flexural strength testing setup

3.2 Dataset Structure

Each of the four mix compositions was tested at 13 standardized frequency levels (100–1600 Hz) and across six acoustic panel configurations (Table 2), generating $4 \times 13 \times 6 = 312$ records in total. Before taking that number at face value, it is worth pausing: 312 rows sounds substantial, but every one of those records comes from just four distinct material compositions. Each mix contributes 78 entries, all of them representing different frequency–configuration pairings rather than independent observations. Compressive and flexural strength values are constant within each mix class, which creates a structural pattern that classifiers can exploit without learning anything about acoustics. The implications of this are examined in Section 4.4. A complete list of dataset variables is given in Table 3.

Table 2. Acoustic panel configurations used in the experimental program

Code	Description
A	Plain concrete panel (no cavity)
A CAV 10	Panel with 10 mm cavity
A 10	10 mm-thick solid panel
A 10 CAV 10	10 mm-thick panel with 10 mm cavity
A 15	15 mm-thick solid panel
A 15 CAV 15	15 mm-thick panel with 15 mm cavity

Table 3. Dataset variables, types, and measurement units

Variable	Type	Unit	Description
Sand	Independent	Mix ratio	Fine aggregate portion
Styrofoam	Independent	Mix ratio	Partial sand replacement level
Portland Cement (PC)	Independent	Mix ratio	Binder (constant at 1 part)
Compressive strength	Derived	kg/cm²	28-day test, SNI 1974:2011
Flexural strength	Derived	kg/cm²	28-day test, SNI 4431:2011
Test frequency	Independent	Hz	13 levels: 100–1600 Hz
Panel configuration	Independent	Code	Six thickness/cavity combos (Table 2)
α (sound absorption)	Dependent (continuous)	0–1	Measured per ASTM E1050-12
Mix code (FS, ST1, ST2, ST3)	Classification target	-	Class label in Analyses 1 & 2

3.3 Machine Learning Setup

All ML modelling was carried out in Orange Data Mining 3.40.0 [11], a visual platform that allows multiple algorithms to be trained and evaluated within a consistent, transparent workflow. Four algorithms with fixed parameter settings (Table 4) represent distinct modelling paradigms: k-NN (distance-based), SVM (margin-based, with internal normalisation), ANN (neural-network, with internal normalisation), and Logistic Regression (linear baseline, no internal normalisation) [8], [10]. No automated hyperparameter tuning was applied. An important technical observation: Orange's k-NN and Logistic Regression nodes do not apply internal feature normalisation. This means that the frequency variable (range 100–1600 Hz) overwhelms the distance calculation for k-NN, and the large-scale differences between features cause Logistic Regression to converge to a degenerate solution. Both models consequently return CA = 0.250, the random chance baseline despite being legitimate algorithms. SVM and ANN avoid this problem through their internal normalisation mechanisms. This pattern is a core finding of the study and is discussed in Sections 4.3 and 4.4.

Table 4. Machine learning algorithms and parameter settings

Algorithm	Parameter Settings
k-NN	k = 3, Euclidean distance. No internal normalisation in Orange k-NN.
SVM	RBF kernel, C = 1, γ = auto (internal normalisation applied)
ANN	1 hidden layer, 5 neurons, sigmoid, learning rate = 0.1, max iter = 200 (internal normalisation applied)
Logistic Regression	Solver = liblinear, C = 1, one-vs-rest (no internal normalisation)

3.4 Classification Problem Formulation

The models are not predicting how much sound a panel will absorb, that is a regression problem, and the regression baseline in Section 4.7 addresses it directly. Instead, the models are asked: given this combination of measured properties, which of the four mix types does this record most likely belong to? The practical value of this framing is in quality control and material identification contexts, where quickly confirming mix type from measured properties could reduce the need for specialized acoustic testing [8]. A regression baseline predicting α from compositional inputs was added to directly compare the feasibility of the two formulations [10,18].

3.5 Analysis Design and Limitations

Analysis 1-Full Feature Set. Styrofoam content, compressive strength, flexural strength, frequency, and α (configuration A) as features; mix code as target.

Analysis 2-Acoustic Features Only**.** Compositional and mechanical features removed; frequency and α across all six configurations as inputs. Tests whether acoustic patterns alone can distinguish mix types.

Regression Baseline**.** Linear Regression, k-NN Regressor, SVR, and MLP Regressor applied to predict α from compositional inputs, with 10-fold cross-validation [10,18].

All classification analyses used 10-fold stratified cross-validation, evaluated on CA, F1-score, and recall. Regression performance was assessed by R^2, MAE, and RMSE. Four explicit limitations apply throughout only four unique compositions were tested, no external validation set was used, no independent samples from other sources were available, and hyperparameters were not optimized. All findings are exploratory.

4. Results and Discussion

4.1 Mechanical Performance

The mechanical test results in Table 5 tell a consistent story. As Styrofoam content increases from FS through to ST3, both compressive and flexural strength decline systematically. FS, the reference mix, recorded the highest values: 149.36 kg/cm² in compression and 89.94 kg/cm² in flexure. The drop to ST1 is substantial: EPS beads create weak points in the cementitious matrix where the low-stiffness beads disrupt load transfer [21]. ST2 retains moderate performance on both measures, while ST3, with the highest Styrofoam fraction, shows the steepest decline. Crucially, each mix produces a unique pair of strength values, a detail that becomes important when interpreting the ML results. Figure 4 shows the comparison visually.

Table 5. Compressive and flexural test results

Mix	Proportion (PC:Sand:Styrofoam)	Compressive Strength (kg/cm²)	Flexural Strength (kg/cm²)
FS	1:2:0	149.36	89.94
ST1	1:2:3	98.99	61.09
ST2	1:1.5:2.5	78.65	55.15
ST3	1:1.25:2.75	66.77	40.73

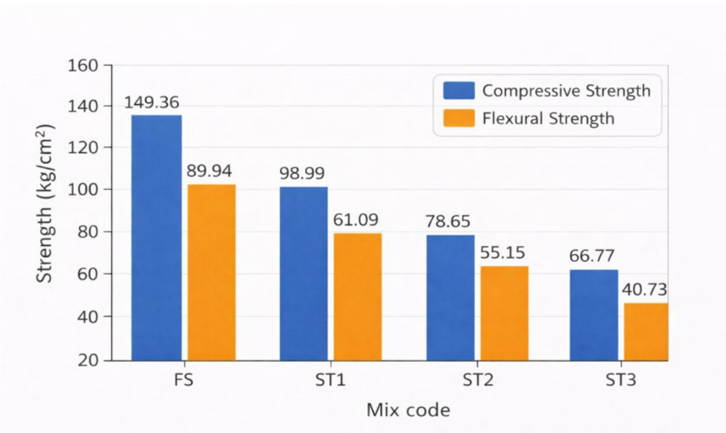

Figure 4. Comparison of compressive and flexural strength across mix variations

4.2 Acoustic Performance (Configuration A)

Figure 5 shows the impedance tube setup used in the laboratory; Figure 6 plots the sound absorption curves for configuration A across the full frequency range tested.

Figure 5. Laboratory setup for acoustic testing using the impedance tube method (ASTM E1050-12), configuration A

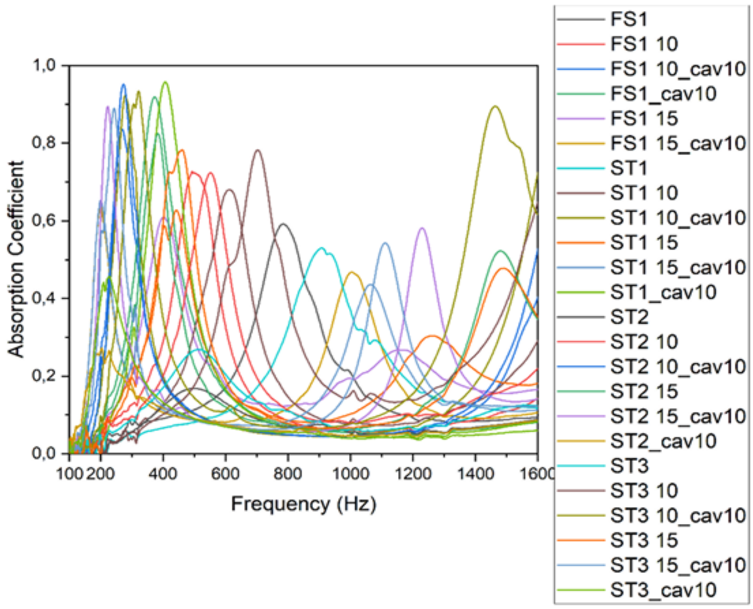

Figure 6. Sound absorption coefficient (α) versus frequency

The reference mix, FS, barely absorbs sound at all α remains below 0.20 from one end of the spectrum to the other, which is exactly what one would expect from a dense, tightly packed material with few internal voids. The Styrofoam-modified mixes paint a more interesting picture. ST1 builds to a peak of around α = 0.58 near 800 Hz, while ST2 maintains more consistent absorption across a broader frequency range without the same sharp peak. ST3, with the most Styrofoam, reaches high values at certain frequencies but fluctuates considerably, a sign that the pore network at that substitution level is somewhat irregular in structure.

The physical explanation is the one common to all porous sound absorbers: sound waves entering the void network lose energy through viscous friction and internal reflections at the air–solid interfaces created by the EPS beads [6,14]. ST2's more uniform behaviour suggests a pore structure that distributes this dissipation more evenly, whereas ST3's irregularity may reflect isolated or partially collapsed voids that form when bead fraction is pushed too high [17]. Taken together, the acoustic and mechanical data point to ST2 as the most balanced composition, meaningful absorption without the structural compromise of ST3.

4.3 Analysis 1: Full-Feature Classification

The Analysis 1 results are shown in Table 6. ANN comes out on top with a perfect CA = 1.000 and AUC = 1.000, classifying every one of the 52 records correctly. SVM follows at CA = 0.923. Both k-NN and Logistic Regression return CA = 0.250, identical to random chance for a balanced four-class problem. The asterisk against the ANN figure and the dagger against Logistic Regression are reminders that neither number, high nor low, tells the full story on its own.

Table 6. Analysis 1-classification performance with full feature set (Orange Data Mining 3.40.0, 10-fold CV stratified, n = 52)

Model	AUC	CA	F1	Recall
k-NN (k=3, no normalisation)	0.469	0.250	0.174	0.250
SVM (RBF, internal norm.)	0.949	0.923	0.924	0.923
ANN * (internal norm.)	1.000	1.000	1.000	1.000
Logistic Regression ** (no norm.)	0.517	0.250	0.100	0.250

Notes:
** See Section 4.4.*
*** Collapses to predicting all records as class FS*

The k-NN and Logistic Regression failures both trace back to the same root cause: the absence of feature normalisation. Orange's k-NN node does not apply internal feature scaling, so the frequency variable (range 100–1600 Hz) dominates the Euclidean distance calculation and overwhelms all other features, the model ends up classifying almost entirely on frequency, which carries no class-specific information (Table 7). Logistic Regression collapses more dramatically: without normalisation, the compressive strength values (66–149 kg/cm²) and frequency values (100–1600 Hz) create such large coefficient imbalances that the solver converges to a degenerate solution, predicting every single record as class FS. The confusion matrix confirms this: 13 correct predictions for FS, zero for every other class. SVM and ANN avoid these problems because both learners apply internal normalisation. This pattern, two of four models failing completely due to a preprocessing gap, reinforces the importance of verifying normalisation settings when using visual ML platforms [8].

4.4 Feature Sensitivity Analysis

The results in Table 7 explain the ANN figure efficiently. When tested individually with a normalized k-NN classifier, Styrofoam content, compressive strength, and flexural strength each achieve CA = 1.000, a perfect score, on their own, without any information about frequency or sound absorption.

Table 7. Feature sensitivity analysis: single-feature classification accuracy (k-NN, k = 3, normalized, 10-fold CV)

Feature	Feature Type	Single feat. CA (k-NN, norm.)	Role in Classification
Styrofoam ratio	Compositional	1.000	Unique per class class identifier
Compressive strength	Mechanical	1.000	Unique per class class identifier
Flexural strength	Mechanical	1.000	Unique per class class identifier
Frequency	Test variable	0.000	No standalone discriminability
α (Config A)	Acoustic response	0.310	Moderate discriminability
α (Config A CAV 10)	Acoustic response	0.150	Limited discriminability
α (Config A 15 CAV 15)	Acoustic response	0.233	Limited discriminability

The explanation is not complicated. Styrofoam ratios across the four classes are 0, 3.0, 2.5, and 2.75 distinct values, no ambiguity. Compressive strengths follow the same pattern: 149.36, 98.99, 78.65, and 66.77 kg/cm², one figure per class, no crossover. Any classifier handed these numbers can name the mix type correctly every time, and it does not need to know a single thing about how that mix absorbs sound to do it. These features are not predictors of class membership in the usual sense; they are definitional properties of each class, embedded directly in the dataset.

What this means for the ANN result is important. The perfect CA = 1.000 and AUC = 1.000 should be read as confirmation that ANN successfully recognized these class-defining patterns not as evidence that the model has learned to predict acoustic behaviour. ANN's internal normalisation and nonlinear decision boundary allow it to exploit all three class-identifier features simultaneously, making perfect classification mathematically guaranteed when those features are present. This is not a demonstration of acoustic learning; it is a demonstration of what happens when a sufficiently expressive model is given features that perfectly label its target classes

[10]. Acoustic features, by contrast, top out at CA = 0.31 for a single α configuration and offer much less consistent discrimination across the others.

4.5 Analysis 2: Acoustic-Only Classification

Stripping out the compositional and mechanical features and running the classifiers on frequency and α alone produces the results in Table 8. The dataset is balanced at 13 records per class per configuration, so random classification corresponds to CA = 0.25. Every model falls at or below that threshold.

Table 8. Analysis 2 - acoustic-only classification

(frequency + 6 × α, Orange Data Mining, 10-fold CV, n = 52)

Model	CA	F1	vs. Random Baseline (CA = 0.25)
k-NN	0.240	0.039	Below random
SVM	0.210	0.064	Below random
ANN	0.247	0.228	Near random
Logistic Regression	0.247	0.100	Near random

No model exceeds CA = 0.247, with the best results going to ANN and Logistic Regression (both at CA = 0.247), barely touching the random baseline. This is not a model failure; it is a data limitation. With only four distinct compositions and 13 measurements per class, the acoustic patterns of different mix types overlap considerably across the frequency range. There is not enough compositional diversity for any classifier to learn stable discriminative rules from acoustic response alone. The implication for future work is clear: meaningful acoustic classification requires testing a much wider range of mix compositions under comparable conditions before ML models can identify robust acoustic patterns.

4.6 Confusion Matrix (SVM, Analysis 1)

Table 9 presents the SVM confusion matrix from 10-fold cross-validation, where overall accuracy settled at 0.923. The ANN matrix is not shown separately because it is perfectly diagonal, all 52 records classified correctly and carries no additional diagnostic information. The SVM matrix is more instructive because its four misclassifications reveal a meaningful pattern.

Table 9. SVM confusion matrix

(Analysis 1, 10-fold CV; rows = actual, columns = predicted)

Actual \ Predicted	FS	ST1	ST2	ST3
FS	13	0	0	0
ST1	0	11	2	0
ST2	0	1	12	0
ST3	0	0	1	12

FS is classified perfectly by SVM (13/13). The four errors concentrate between adjacent Styrofoam content levels: ST1 vs ST2 (2 misclassifications), ST2 vs ST1 (1), and ST3 vs ST2 (1). This pattern is informative the misclassifications occur precisely at the boundaries where Styrofoam ratios are closest together (2.5 for ST2 and 2.75 for ST3), and where the difference in compressive strength is smallest. Even a capable kernel-based model finds these boundaries ambiguous when separating features differ by only 0.25 ratio units [21].

4.7 Regression Baseline

Table 10 shows what happens when the problem is reframed as regression predicting α directly from compositional and frequency inputs.

Table 10. Regression baseline: prediction of α (configuration A) from compositional and frequency inputs (10-fold CV, n = 52)

Model	R^2	MAE	RMSE
Linear Regression	-0.898	0.089	0.131
k-NN Regressor	-1.865	0.068	0.127
SVR (RBF)	-1.573	0.100	0.135
MLP Regressor	-0.319	0.077	0.135
Baseline (predict mean)	-0.733	0.086	0.135

All four regression models finish with negative R^2 values. The least poor result belongs to MLP Regressor at $R^2 = -0.319$, which still means it performs worse than the trivial strategy of predicting the mean α for every record. Within each mix class, α varies substantially across the 13 frequency levels, and the four classes do not differ consistently enough in their average α profiles to give the models anything stable to learn from [10,18]. A more productive regression study would need a dataset covering at minimum 10-15 distinct mix proportions rather than four.

4.8 What the Results Mean Together

The full picture is more revealing than any single model's accuracy number would suggest. ST2 performs best experimentally, acoustically consistent, mechanically adequate. In Analysis 1, two of the four models (ANN and SVM) achieve high accuracy, but through completely different mechanisms: ANN exploits class-identical features to achieve perfection; SVM uses a kernel that normalizes implicitly and achieves near-perfect separation with four boundary errors. The other two models fail entirely due to the absence of normalisation, a preprocessing issue that has nothing to do with acoustic learning. Strip out the compositional features and restrict to acoustics only, and all four models collapse to or below random chance. This clean separation, high accuracy when class-identifier features are present, complete failure without them, is precisely the evidence the dual-analysis framework was designed to surface.

5. Conclusion

This study examined Styrofoam-based lightweight concrete through a comparative ML classification framework and found that ST2 (PC:Sand:Styrofoam = 1:1.5:2.5) offers the most balanced combination of acoustic and mechanical performance.

Three findings stand out. First, when given the full feature set, ANN achieves perfect CA = 1.000 and SVM reaches CA = 0.923, but feature sensitivity analysis shows these reflect compositional features that uniquely identify each class rather than any acoustic learning. Logistic Regression and k-NN fail completely (CA = 0.250) due to missing normalisation, which is itself a finding about the importance of preprocessing in visual ML platforms. Second, restricting classifiers to acoustic features only drops all models to or near random chance (CA = 0.21–0.25), confirming that acoustic fingerprinting is not feasible with a four-composition dataset. Third, regression models for continuous α prediction achieve $R^2 \leq -0.319$, confirming the dataset does not support reliable acoustic property forecasting.

When the dataset is small, results that look impressive on paper are precisely the ones that need the most scrutiny, a high accuracy figure on four material compositions can reflect the structure of the data far more than the capability of the model. Acknowledging that openly is more useful than reporting the number without context. As for where the work goes from here, the path is not difficult to map out: more mix variations need to be tested, replicates should be collected in a way that produces genuinely independent data points, and tools like SHAP values should be brought in to make visible which acoustic features are actually doing the discriminative work. Regression-based prediction is a natural next step, but only once the dataset has enough breadth to make it a meaningful exercise.

Acknowledgments

This work was conducted as part of academic research activities and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References

[1] M. J. Islam, M. S. Meherier, and A. K. M. R. Islam, "Effects of waste PET as coarse aggregate on the fresh and harden properties of concrete," *Constr. Build. Mater.*, vol. 125, pp. 946–951, Oct. 2016, doi: 10.1016/j.conbuildmat.2016.08.128.

[2] B. Rosca, "Eco-Friendly Lightweight Aggregate Concrete of Structural Grade Made with Recycled Brick Aggregate Containing Expanded Polystyrene Beads," *Sustainability (Switzerland)*, vol. 17, no. 7, Apr. 2025, doi: 10.3390/su17073050.

[3] C. Sposato, T. Cardinale, M. B. Alba, A. Feo, L. Pala, and P. De Fazio, "Innovative Lightweight and Sustainable Composite Material for Building Applications," *Sustainability (Switzerland)*, vol. 17, no. 16, Aug. 2025, doi: 10.3390/su17167319.

[4] U. Berardi and G. Iannace, "Acoustic characterization of natural fibers for sound absorption applications," *Build. Environ.*, vol. 94, pp. 840–852, Dec. 2015, doi: 10.1016/j.buildenv.2015.05.029.

[5] S. N. Shah *et al.*, "Physical, strength and acoustic properties of lightweight cement composite with preplaced chemically-treated crumb rubber," *Case Studies in Construction Materials*, vol. 20, Jul. 2024, doi: 10.1016/j.cscm.2023.e02821.

[6] T. S. Tie, K. H. Mo, A. Putra, S. C. Loo, U. J. Alengaram, and T. C. Ling, "Sound absorption performance of modified concrete: A review," Jul. 01, 2020, *Elsevier Ltd*. doi: 10.1016/j.jobe.2020.101219.

[7] K. A. P. Wijesinghe, G. Lanarolle, C. Gunasekara, D. W. Law, H. D. Hidallana-Gamage, and L. Wang, “Thermal and acoustic performance of solid waste incorporated cement based composites: an analytical review,” Feb. 01, 2025, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s43452-025-01160-3.
[8] Z. Li *et al.*, “Machine learning in concrete science: applications, challenges, and best practices,” Dec. 01, 2022, *Nature Research*. doi: 10.1038/s41524-022-00810-x.
[9] B. Ni, M. Z. Rahman, S. Guo, and D. Zhu, “A review on properties and multi-objective performance predictions of concrete based on machine learning models,” Mar. 01, 2025, *Elsevier Ltd*. doi: 10.1016/j.mtcomm.2025.112017.
[10] M. Mohtasham Moein *et al.*, “Predictive models for concrete properties using machine learning and deep learning approaches: A review,” Jan. 01, 2023, *Elsevier Ltd*. doi: 10.1016/j.jobe.2022.105444.
[11] J. Demšar *et al.*, “Orange: Data Mining Toolbox in Python Tomaž Curk Matija Polajnar Laň Zagar,” 2013.
[12] N. S. Galip *et al.*, “Identification of Mechanical and Sound Absorption Properties of Porous Concrete Containing Different Amounts of Palm Oil Clinker,” *Journal of Mechanical Engineering*, vol. 13, pp. 101–120, 2024, doi: 10.24191/jmeche.v13i1.3759.
[13] L. S. Kang *et al.*, “Acoustic properties of lightweight foamed concrete with eggshell waste as partial cement replacement material,” *Sains Malays.*, vol. 50, no. 2, pp. 537–547, Feb. 2021, doi: 10.17576/jsm-2021-5002-24.
[14] M. Amran, R. Fediuk, G. Murali, N. Vatin, and A. Al-Fakih, “Sound-absorbing acoustic concretes: A review,” Oct. 01, 2021, *MDPI*. doi: 10.3390/su131910712.
[15] N. H. Crista, S. P. Hadi, and E. Setyowati, “The Effectiveness of Styrofoam Mixtures in Lightweight Concrete Walls,” 2024.
[16] A. Elghomari and A. Tilioua, “Investigation and valorization of expanded polystyrene waste in building materials: thermal and mechanical characterization,” *Results in Materials*, vol. 28, p. 100782, Dec. 2025, doi: 10.1016/j.rinma.2025.100782.
[17] L. Prasittisopin, P. Termkhajornkit, and Y. H. Kim, “Review of concrete with expanded polystyrene (EPS): Performance and environmental aspects,” Sep. 15, 2022, *Elsevier Ltd*. doi: 10.1016/j.jclepro.2022.132919.
[18] C. Sun, K. Wang, Q. Liu, P. Wang, and F. Pan, “Machine-Learning-Based Comprehensive Properties Prediction and Mixture Design Optimization of Ultra-High-Performance Concrete,” *Sustainability (Switzerland)*, vol. 15, no. 21, Nov. 2023, doi: 10.3390/su152115338.
[19] Y. Patrisia, C. Gunasekara, D. W. Law, S. Setunge, and B. Kaminsky, “Engineering and thermo-acoustic insulation performance of recycled waste concrete composites,” *J. Sustain. Cem. Based. Mater.*, vol. 14, no. 11, pp. 2441–2459, 2025, doi: 10.1080/21650373.2025.2533996.
[20] K. A. Moges, N. Dalila, P. Plaskota, and S. Pyo, “Evaluation methods, testing standards, and simulation techniques of sound absorption capabilities of cementitious materials: A review,” Nov. 01, 2024, *Elsevier Ltd*. doi: 10.1016/j.jobe.2024.110468.
[21] D. He, W. Zheng, Z. Chen, Y. Qi, D. Zhang, and H. Li, “Influence of Paste Strength on the Strength of Expanded Polystyrene (EPS) Concrete with Different Densities,” *Polymers (Basel).*, vol. 14, no. 13, Jul. 2022, doi: 10.3390/polym14132529.