Machine Learning-Based Prediction and Optimization of Density and Strength in Structural Lightweight Concrete

Odiljon Fozilov[1], Abdurasul Martazaev[1], Zokirjon Kholboev[1], Ravshanbek Mavlonov[1], Mukharram Abelkosimova[2] and Venkatesh Kodur[3]

[1]*Department of Construction of Buildings and Structures, Namangan State Technical University, Islam Karimov Ave. 12, 160720 Namangan, Uzbekistan*

[2]*Department of Construction and Installation of Engineering Communications, Namangan State Technical University, Islam Karimov Ave. 12, 160720 Namangan, Uzbekistan*

[3]*Department of Civil and Environmental Engineering, Michigan State University, 3574 Engineering Building, 428 S. Shaw Lane, 48824 East Lansing, MI, USA*

odiljon690.21.03@gmail.com, abdurasul.mas@gmail.com, zxx7712407@gmail.com, ravshanbek.mavlonov@gmail.com, abelkosimovamuharram4@gmail.com, kodur@egr.msu.edu

Keywords: Polynomial Regression, Lightweight Concrete, Mix Optimization, Compressive Strength, LOOCV, Machine Learning.

Abstract: This study combines an experimental lightweight-concrete program with a small-sample machine-learning benchmark. Eleven mixtures were produced by replacing normal-weight fine and coarse aggregates with expanded clay at volumetric ratios from 0 to 100%. Density and compressive strength were measured at 7, 14, and 28 days. Under the adopted requirement of 28-day strength not lower than 40 MPa, the mixture containing 60% lightweight and 40% normal-weight aggregate provided the lowest measured density (1925 kg/m³) while retaining 42.7 MPa compressive strength. Python-based degree-2 polynomial models were fitted to the 11 series averages and compared by leave-one-out cross-validation (LOOCV) with Random Forest, XGBoost, and a compact multilayer perceptron. The polynomial model achieved the best LOOCV density performance ($R^2 = 0.9883$; RMSE = 31.33 kg/m³) and remained competitive for strength ($R^2 = 0.9700$; RMSE = 0.91 MPa), whereas the MLP produced the best strength score. The results support an interpretable polynomial model for preliminary screening of this small, single-factor dataset, while broader predictive claims require larger multi-factor data.

1 INTRODUCTION

Structural lightweight concrete reduces structural self-weight, but its performance depends strongly on lightweight-aggregate type, grading, moisture condition, and replacement ratio. Studies using expanded clay and other porous aggregates show a recurring density-strength trade-off: moderate replacement can reduce unit weight while retaining structural performance, whereas excessive porosity or aggregate replacement can cause substantial strength loss [1]-[6].

Experimental optimization remains essential, but computational analysis can make mixture screening more systematic. Lightweight-concrete response has been examined through numerical structural models [7], dispersion-based theoretical assessment of concrete properties [8], and combined analytical-experimental interpretation of reinforced members [9]. More broadly, computational modeling can transform measured operational variables into quantitative engineering indicators [10], while data-driven workflows can convert heterogeneous observations into structured decision-support priorities [11]. Neural networks and other machine-learning methods have also been applied to concrete-strength prediction and lightweight-concrete composition design [12]-[16]. These studies show that computational model value depends on dataset size, input dimensionality, and the validation protocol.

The present dataset is intentionally small and structured: eleven mixture series vary only the total lightweight-aggregate replacement ratio, whereas the water-to-cement ratio and superplasticizer dosage remain constant. In this setting, a flexible black-box model may overfit and may perform poorly when a left-out endpoint must be predicted. A transparent benchmark is therefore needed to determine whether a low-degree polynomial is sufficient or whether Random Forest, XGBoost, or a neural network provides better leave-one-out generalization.

This study experimentally evaluates density and compressive strength of expanded-clay concrete produced under hot and dry conditions in Uzbekistan and identifies a strength-constrained mixture. A reproducible Python workflow then fits degree-2 polynomial models and benchmarks them against three machine-learning regressors using LOOCV. The contribution is a combined experimental and computational procedure for preliminary mix screening, with fitted-data and cross-validated metrics reported separately.

2 MATERIALS AND METHODS

2.1 Materials

An experimental program was conducted to evaluate structural lightweight concrete under the hot and dry climatic conditions of Uzbekistan. The reference mixture was designed for concrete class B30 and used a combination of normal-weight and expanded-clay aggregates.

Figure 1: The process of determining the strength of expanded clay using standard cylinders.

PS450K32.5 Portland cement from Namangantsement was used as the binder. River-gravel sand from the Uychi district served as the normal-weight fine aggregate, and granite gravel with a density of 2665 kg/m³ and a particle size of 5-20 mm served as the conventional coarse aggregate. Expanded clay produced by Junior Business-Trade LLC was used as the lightweight aggregate. Figure 1 illustrates aggregate-strength testing, and Table 1 summarizes the measured properties of the expanded-clay fractions.

Table 1: Main characteristics of lightweight aggregates.

№	Properties	Unit	Values by particle size, mm		
			5-10	10-20	5-20
1	Bulk density	kg/m³	880	720	730
2	Compressive strength of cylindrical specimens	MPa	2.6	2.3	2.4
3	Water absorption	%	14	17	16
4	Frost resistance	cycles	25	25	25

The reference B30 mixture composition is given in Table 2.

Table 2: Material for 1 m³ of concrete mix.

Material	Quantity
Cement (kg)	440
Fine aggregate (kg)	1072
Coarse aggregate (kg)	800
Water (kg)	169

2.2 Experimental Program and Specimen Preparation

Expanded clay was pre-soaked for 24 h and brought to a saturated-surface-dry condition before batching to limit unintended changes in effective water content and workability.

Initial trials replaced only the conventional coarse aggregate. As shown in Figure 2, this approach did not reduce concrete density sufficiently; therefore, the second experimental stage replaced both fine and coarse normal-weight aggregates with expanded clay.

Eleven mixtures with total lightweight-aggregate replacement ratios from 0 to 100% were then prepared. The replacement schedule, constant w/c ratio of 0.38, and number of specimens per age are given in Table 3.

A universal superplasticizer was added at 1% of cement mass. Specimens were cured at 20 ± 2°C and approximately 95% relative humidity. Three 100 × 100 × 100 mm cubes from each series were tested at 7, 14, and 28 days.

Figure 3 shows the prepared cube specimens and molds.

Compressive-strength testing followed GOST 10180-2012 [17]. Specimen dimensions and mass were recorded before loading, and cubes were centered on the compression-machine platen. Load was increased continuously to failure at 0.3-0.5 MPa/s (Fig. 4).

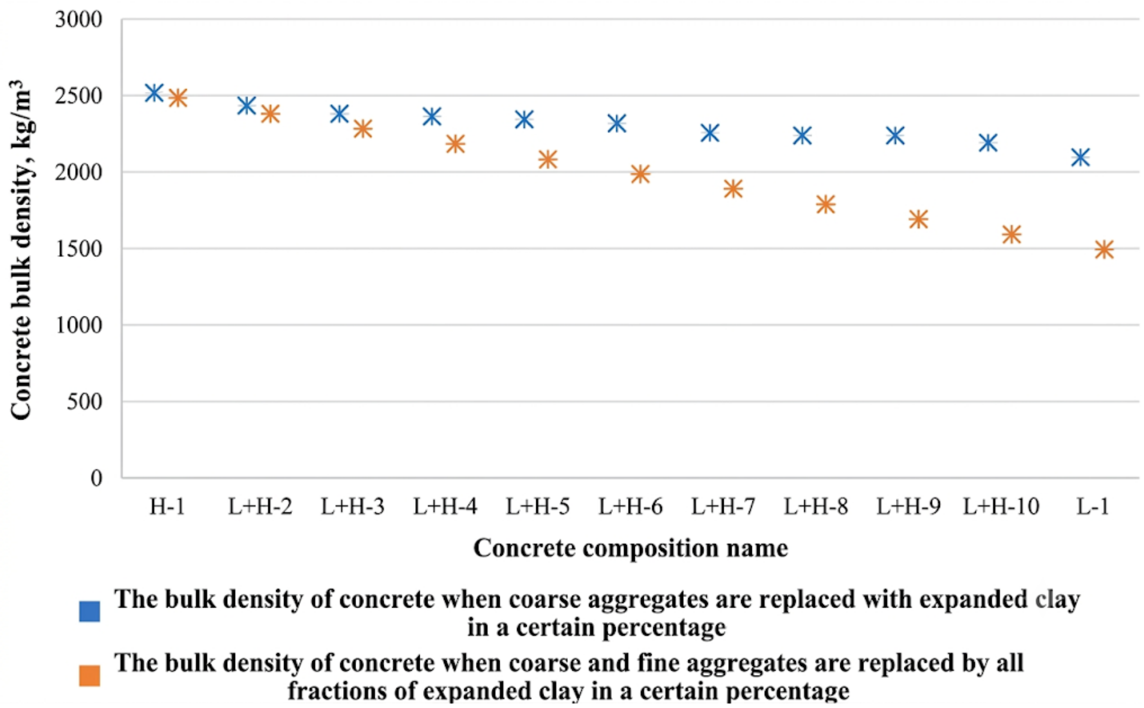

Figure 2: Bulk density of concrete.

Table 3: Mix design of lightweight and normal-weight aggregates.

Mix ID	Lightweight, agg. (%)	Normal weight agg. (%)	w/c	Number of specimens per age
H-1	0	100	0.38	3
L+H-2	10	90	0.38	3
L+H-3	20	80	0.38	3
L+H-4	30	70	0.38	3
L+H-5	40	60	0.38	3
L+H-6	50	50	0.38	3
L+H-7	60	40	0.38	3
L+H-8	70	30	0.38	3
L+H-9	80	20	0.38	3
L+H-10	90	10	0.38	3
L-1	100	0	0.38	3

Figure 3: Prepared cube specimens and molds.

Figure 4: Compression testing process of cube specimens.

2.3 Computational Analysis and Predictive Modeling

The computational analysis used the 11 series averages reported in Table 4. The sole varying input was the total lightweight-aggregate replacement ratio (%); w/c = 0.38 and superplasticizer dosage = 1% were constant. Targets were hardened density and 28-day compressive strength. Degree-2 and degree-3 polynomial regressions were evaluated, and R^2, root mean squared error (RMSE), and mean absolute percentage error (MAPE) were calculated.

Generalization was estimated by leave-one-out cross-validation (LOOCV): each mixture series was withheld once, the model was fitted to the remaining ten series, and the withheld response was predicted. Polynomial degree was selected from {2, 3} by LOOCV, retaining degree 2. The full-data degree-2 fit was used only to describe trends and generate Figures 5 and 6; LOOCV metrics in Table 5 were used for model comparison.

Three additional regressors were benchmarked under the same LOOCV protocol: Random Forest [18], XGBoost [19], and an 8×8 multilayer perceptron. Random Forest used 100 trees; XGBoost used 100 boosting rounds with learning rate 0.1; and the MLP used the L-BFGS solver on standardized inputs and targets. Random Forest and MLP results were averaged over ten random seeds. Computations were implemented in Python with scikit-learn [20] and XGBoost [19].

3 RESULTS AND DISCUSSION

Table 4 summarizes the measured density and compressive strength of the 11 mixture series. At 7 days, compressive strength reached approximately 64-66% of the 28-day value for most mixtures; at 14 days, most mixtures reached approximately 90-98%.

The expanded-clay aggregate properties were assessed against structural-lightweight-concrete criteria in ACI 213R [21] and ASTM C330/ C330M [22]. The measured bulk densities in Table 1 are consistent with lightweight-aggregate classification for the investigated fractions.

Mixture screening minimized density subject to a 28-day compressive-strength constraint of at least 40 MPa. Among the measured series, L+H-7 (60% lightweight and 40% normal-weight aggregate) provided the lowest density satisfying this constraint: 1925 kg/m³ at 42.7 MPa. The polynomial model was used as an interpretable trend model rather than as a substitute for the experimental constraint check.

The degree-2 polynomial models were fitted directly to the 11 series averages in Table 4. For 28-day compressive strength, the fitted model achieved $R^2 = 0.9885$, RMSE = 0.56 MPa, and MAPE = 1.27%. For density, the corresponding values were $R^2 = 0.9943$, RMSE = 21.83 kg/m³, and MAPE = 0.87%. These are in-sample fit metrics; generalization is assessed separately by LOOCV in Table 5.

The degree-2 polynomial fit for 28-day compressive strength is shown in Figure 5.

For mixture L+H-7 at a 60% lightweight-aggregate replacement ratio, the degree-2 polynomial models predicted a 28-day compressive strength of 42.21 MPa and a density of 1962.10 kg/m³. The corresponding experimental values were 42.7 MPa and 1925 kg/m³, respectively. Thus, the fitted model underestimated compressive strength by 0.49 MPa and overestimated density by 37.10 kg/m³.

The degree-2 polynomial fit for concrete density is shown in Figure 6.

Table 4: Measured density and compressive strength of cube specimens.

№	Mixture	Density, kg/m³	Compressive strength, (MPa)		
			7 days	14 days	28 days
1	H-1	2515	31.8	47.2	49.23
2	L+H-2	2461	31.4	46.3	47.5
3	L+H-3	2335	31.0	44.7	46.9
4	L+H-4	2258	30.6	44.1	46.4
5	L+H-5	2123	30.4	43.4	45.8
6	L+H-6	2092	29.0	42.4	44.5
7	L+H-7	1925	27.2	39.3	42.7
8	L+H-8	1861	26.25	36.3	40.38
9	L+H-9	1798	24.47	33.4	37.07
10	L+H-10	1702	22.94	31.3	34.76
11	L-1	1632	21.85	29.8	33.1

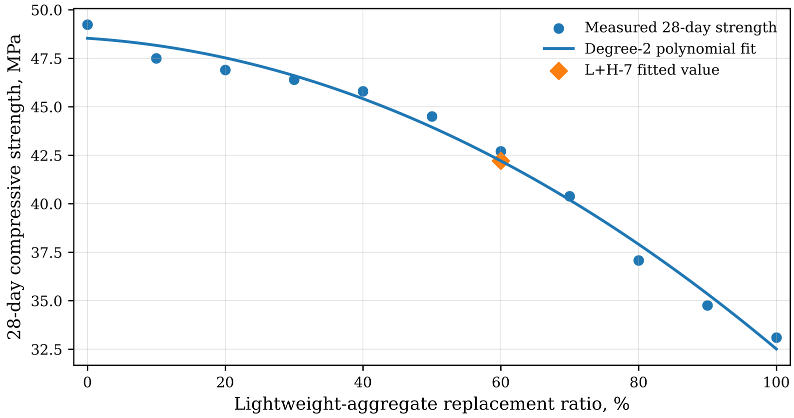

Figure 5: Degree-2 polynomial fit for 28-day compressive strength and the L+H-7 fitted value.

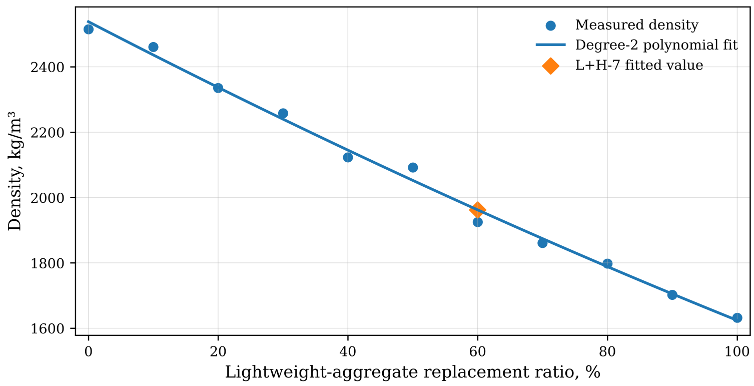

Figure 6: Degree-2 polynomial fit for density and the L+H-7 fitted value.

Table 5: LOOCV benchmark of polynomial and machine-learning regressors (RMSE in target units; stochastic models averaged over ten seeds).

Target	Model	R^2	RMSE	MAPE, %
Density, kg/m^3	Polynomial (deg. 2)	0.9883	31.33	1.24
	Random Forest	0.9343	74.32	3.19
	XGBoost	0.8949	94.01	4.24
	MLP (8×8)	0.9775	43.13	1.73
Strength, MPa	Polynomial (deg. 2)	0.9700	0.91	1.94
	Random Forest	0.9272	1.41	2.89
	XGBoost	0.8790	1.82	4.04
	MLP (8×8)	0.9826	0.66	1.27

The fitted curves reproduce the observed trends closely within the 11 investigated series and provide a compact description of the replacement-ratio effect. Broader predictive applicability is not inferred from these fitted-data metrics; the LOOCV benchmark in Table 5 is used for generalization comparison.

Table 5 compares the degree-2 polynomial with Random Forest, XGBoost, and the MLP under the common LOOCV protocol.

For density, the degree-2 polynomial achieved the best LOOCV result (R^2 = 0.9883; RMSE = 31.33 kg/m³), followed by the MLP (R^2 = 0.9775; RMSE = 43.13 kg/m³). For 28-day strength, the MLP performed best (R^2 = 0.9826; RMSE = 0.66 MPa), while the polynomial remained competitive (R^2 = 0.9700; RMSE = 0.91 MPa). Random Forest and XGBoost ranked lower for both targets.

The result is consistent with the structure of this dataset: one varying factor, eleven ordered series, and smooth monotone responses. Tree ensembles cannot extrapolate beyond observed response levels at withheld endpoints, whereas a low-degree polynomial represents the dominant trend with few parameters and remains interpretable for strength-constrained screening. On larger multi-factor concrete datasets, ensemble and neural models can become more advantageous [23], [24].

4 CONCLUSIONS

The experimental program showed that replacing only coarse aggregate with expanded clay did not provide the required density reduction. Combined replacement of fine and coarse aggregates was more effective, while 24 h pre-soaking helped control aggregate moisture before batching.

Among the 11 mixtures, L+H-7 with 60% lightweight and 40% normal-weight aggregate provided the best measured balance under the adopted constraint of 28-day compressive strength ≥ 40 MPa, achieving 1925 kg/m³ and 42.7 MPa.

Degree-2 polynomial regression provided an interpretable description of the experimental trend. Full-data fits yielded $R^2 = 0.9885$ for strength and 0.9943 for density, while LOOCV showed the polynomial to be best for density and competitive with the MLP for strength. These results support polynomial regression as a preliminary screening model for this small, single-factor dataset; larger multi-factor datasets are required for broader predictive optimization.

REFERENCES

[1] M. B. Khan, A. O. Baarimah, M. S. M. Banaimoon, M. B. Ramu, A. F. Ebrahim, B. T. Alsulami, and H. Almujibah, "Exploring the synergy between aluminum powder and silica fume in enhancing lightweight concrete characteristics," Scientific Reports, vol. 15, Art. no. 26349, 2025, [Online]. Available: https://doi.org/10.1038/s41598-025-26349-7.

[2] M. R. Ahmad and B. Chen, "Experimental research on the performance of lightweight concrete containing foam and expanded clay aggregate," Composites Part B: Engineering, vol. 171, pp. 46-60, 2019, [Online]. Available: https://doi.org/10.1016/j.compositesb.2019.04.025.

[3] T. J. Rajeeth, A. Sharma, R. Honnalli, S. Yantrapalli, S. N. Shinde, and S. Thenmozhial, "Experimental investigation and optimization of cellular lightweight concrete using foam content and prediction using response surface methodology," Discover Sustainability, vol. 6, 2025, [Online]. Available: https://doi.org/10.1007/s43621-025-01829-y.

[4] A. A. Aziz and I. F. Al Mulla, "An experimental investigation of the mechanical properties of lightweight concrete with pumice aggregate and polypropylene fibers," Engineering, Technology & Applied Science Research, vol. 15, no. 5, pp. 26344-26350, 2025, [Online]. Available: https://doi.org/10.48084/etasr.12495.

[5] X. Zhang, X. Kuang, J. H. Yang, and S. R. Wang, "Experimental study on mechanical properties of lightweight concrete with shale aggregate replaced part by natural sand," Electronic Journal of Structural Engineering, vol. 17, pp. 85-94, 2017, [Online]. Available: https://doi.org/10.56748/ejse.17222.

[6] S. Chandra and L. Berntsson, Lightweight Aggregate Concrete: Science, Technology and Applications. Norwich, NY, USA: Noyes Publications, 2002.

[7] O. Fozilov, A. Martazaev, S. Razzakov, R. Mavlonov, and B. Maksud ugli, "Numerical modeling of reinforced concrete structures made of lightweight concrete using ANSYS," Vibroengineering Procedia, vol. 60, pp. 683-689, 2025, [Online]. Available: https://doi.org/10.21595/vp.2025.25511.

[8] A. Martazaev, S. Razzakov, R. Mavlonov, O. Fozilov, and B. Maksud ugli, "Theoretical assessment of the mechanical properties of fiber concrete using the dispersion analysis method," Vibroengineering Procedia, vol. 60, pp. 788-794, 2025, [Online]. Available: https://doi.org/10.21595/vp.2025.25362.

[9] R. S. Mehany, H. M. Mohamed, and B. Benmokrane, "Flexural Strength and Serviceability of GFRP-Reinforced Lightweight Self-Consolidating Concrete Beams," Journal of Composites for Construction, vol. 26, no. 3, Art. no. 04022020, Jun. 2022, doi: 10.1061/(ASCE)CC.1943-5614.0001208.

[10] A. A. Yadav et al., "AI-driven sustainable concrete mix design: Hybrid deep Q-Learning and genetic algorithms-based multi-objective machine learning optimizations for high structural strength, low cost, and low carbon footprints," Structures, vol. 82, 2025, [Online]. Available: https://doi.org/10.1016/j.istruc.2025.110443.

[11] J. Wang, J. Deng, S. Li, W. Du, Z. Zhang, and X. Liu, "Explainable Machine Learning for Multicomponent Concrete: Predictive Modeling and Feature Interaction Insights," Materials, vol. 18, no. 19, Art. no. 4456, 2025, [Online]. Available: https://doi.org/10.3390/ma18194456.

[12] I. C. Yeh, "Modeling of strength of high-performance concrete using artificial neural networks," Cement and Concrete Research, vol. 28, no. 12, pp. 1797-1808, 1998, [Online]. Available: https://doi.org/10.1016/S0008-8846(98)00165-3.

[13] A. S. Balykov, E. A. Kaledina, and S. V. Volodin, "Compressive strength prediction and composition design of structural lightweight concretes using machine learning methods," Nanotechnologies in Construction, vol. 15, pp. 171-186, 2023, [Online]. Available: https://doi.org/10.15828/2075-8545-2023-15-2-171-186.

[14] F. Hussain, S. A. Khan, R. A. Khushnood, A. Hamza, and F. Rehman, "Machine learning-based predictive modeling of sustainable lightweight aggregate concrete," Sustainability, vol. 15, no. 1, Art. no. 641, 2023, [Online]. Available: https://doi.org/10.3390/su15010641.

[15] S. H. Han, K. H. Khayat, S. Park, and J. Yoon, "Machine learning-based approach for optimizing mixture proportion of recycled plastic aggregate concrete considering compressive strength, dry density, and production cost," Journal of Building Engineering, vol. 83, Art. no. 108393, 2024, [Online]. Available: https://doi.org/10.1016/j.jobe.2023.108393.

[16] P. Ziolkowski, "Influence of optimization algorithms and computational complexity on concrete compressive strength prediction machine learning models for concrete mix design," Materials, vol. 18, no. 6, Art. no. 1386, 2025, [Online]. Available: https://doi.org/10.3390/ma18061386.

[17] GOST 10180-2012: Concrete. Methods for Determining Strength Using Control Specimens. Moscow, Russia: Standartinform, 2013.

[18] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001, [Online]. Available: https://doi.org/10.1023/A:1010933404324.

[19] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 785-794, 2016, [Online]. Available: https://doi.org/10.1145/2939672.2939785.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.

[21] American Concrete Institute Committee 213, ACI 213R-03: Guide for Structural Lightweight-Aggregate Concrete. Farmington Hills, MI, USA, 2003.

[22] ASTM International, ASTM C330/C330M-17a: Standard Specification for Lightweight Aggregates for Structural Concrete. West Conshohocken, PA, USA, 2017.

[23] S. A. Stel'makh, E. M. Shcherban', A. N. Beskopylny, L. R. Mailyan, B. Meskhi, I. Razveeva, and N. Beskopylny, "Prediction of mechanical properties of highly functional lightweight fiber-reinforced concrete based on deep neural network and ensemble regression trees methods," Materials, vol. 15, no. 19, Art. no. 6740, 2022, [Online]. Available: https://doi.org/10.3390/ma15196740.

[24] Y. T. Altuncı, "A comprehensive study on the estimation of concrete compressive strength using machine learning models," Buildings, vol. 14, no. 12, Art. no. 3851, 2024, [Online]. Available: https://doi.org/10.3390/buildings14123851.