Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 4, Issue 2, 2026

MACHINE LEARNING-BASED PREDICTION OF CONCRETE STRUCTURAL PERFORMANCE: A CRITICAL REVIEW OF MODELS, APPLICATIONS, AND CHALLENGES

Dr. M. Adil Khan[*1], Waqas Aziz[2], Hafiz Muhammad Shahzad Aslam[3], Dilshad Ahmed[4]

[*1]*Resident Engineer, NESPAK*
[2]*Assistant Professor, University of southern Punjab*
[3]*Lecturer Department of Technology, The University of Lahore, Pakistan*
[4]*Deputy Director Punjab Mass Transit Authority*

[*1]adee.uol@gmail.com, [2]**waqasaziz30@gmail.com,** [3]engr.hafizshahzad06@gmail.com,
[4]dilshaddil@gmail.com/ddeng.piu.pmu@pma.punjab.gov.pk

DOI: https://doi.org/10.5281/zenodo.21278983

Keywords
machine learning, concrete structures, structural performance prediction, deep learning, neural networks, civil engineering, performance modeling, predictive algorithms

Article History
Received: 07 December 2025
Accepted: 22 January 2026
Published: 06 February 2026

Corresponding Author: *
Dr. M. Adil Khan

Abstract
Concrete remains the world's most widely used construction material, yet predicting its structural performance under various conditions presents significant challenges due to complex nonlinear relationships between material composition, environmental factors, and mechanical properties. Traditional experimental and empirical methods for assessing concrete strength, durability, and structural integrity are time-consuming, costly, and labor-intensive. Recent advances in machine learning (ML) and deep learning (DL) have emerged as powerful alternatives for modeling concrete structural performance with unprecedented accuracy and efficiency. This review synthesizes developments in ML applications for concrete structural prediction, examining neural network architectures, ensemble methods, hybrid optimization frameworks, and explainable artificial intelligence techniques. The paper critically analyzes 60+ peer-reviewed studies covering compressive strength prediction, tensile strength estimation, fracture energy prediction, durability modeling, crack detection, seismic damage assessment, and structural health monitoring. Key findings demonstrate that ensemble methods—particularly XGBoost, CatBoost, and Random Forest—achieve prediction accuracies exceeding $R^2 = 0.96$ across diverse concrete formulations. However, significant challenges persist, including limited field-based datasets, poor generalization across material systems, data imbalance issues, and insufficient model explainability. This review identifies critical research gaps and recommends future directions toward physics-informed machine learning, multi-task learning frameworks, integration of non-destructive testing data, and development of open-access benchmark datasets to advance the field toward practical deployment in civil engineering design and infrastructure management.

1. Introduction
Concrete compressive strength (CS) is a crucial performance parameter in concrete structure design. Reliable strength prediction reduces costs and time in design and prevents material waste from extensive mixture trials. (Elshaarawy,

Alsaadawi and Hamed, 2024) The global concrete industry faces mounting pressure to optimize material usage, reduce environmental impact, and accelerate construction timelines while maintaining structural safety and durability. Traditional approaches rely on destructive testing methods that require costly laboratory facilities, trained personnel, and extended curing periods. Concrete properties, including its compressive strength, are in general highly nonlinear functions of its components. Concrete mix design methods are basically simulations that require costly and time consuming adjustments in laboratory. (Neira *et al.*, 2020) (Shah, Ullah, Khoso, & Umali, 2026; Shah, Ullah, Khoso, Iqbal, et al., 2026)

The emergence of machine learning and artificial intelligence in civil engineering represents a paradigm shift toward data-driven decision-making. Concrete, a ubiquitous material in construction, is known for its strength and durability, but its performance can be significantly improved through the optimization of material properties such as strength, fire resistance, and impact protection. Traditional methods of optimizing these properties rely heavily on empirical testing and expert intuition, which can be time-consuming and may not fully capture the complex interactions between different material components. (Ogunsaya and Taiwo, 2024)

This review critically examines machine learning applications in predicting concrete structural performance, encompassing both conventional and specialized concrete types. The scope includes neural networks, ensemble methods, deep learning architectures, optimization techniques, and emerging approaches integrating physics-informed learning. Unlike previous reviews focusing on single applications or model types, this comprehensive analysis synthesizes the landscape across multiple prediction tasks including compressive strength, tensile properties, durability, crack detection, and structural damage assessment. The review identifies methodological advances, validates performance metrics, critiques limitations, and proposes strategic research directions for advancing ML-based concrete structural prediction toward practical civil engineering implementation. Six key sections structure the analysis: (1) Introduction framing the problem context; (2) Methodology examining diverse ML approaches; (3) Results synthesizing performance findings; (4) Discussion analyzing challenges and limitations; (5) Research gaps identifying unexplored areas; and (6) Conclusion with recommendations.

2. Methodology and Machine Learning Models in Concrete Performance Prediction

2.1 Neural Network Architectures

Artificial Neural Networks (ANNs) constitute the foundational approach for concrete property prediction. Concrete compressive strength is a critical parameter in construction and structural engineering. Destructive experimental methods that offer a reliable approach to obtaining this property involve time-consuming procedures. Recent advancements in artificial neural networks (ANNs) have shown promise in simplifying this task by estimating it with high accuracy. (Alibrahim, Habib and Habib, 2025) Multi-layer perceptron networks with backpropagation algorithms have demonstrated consistent performance across diverse datasets. The work presents the results of an experimental campaign carried out on concrete elements in order to investigate the potential of using artificial neural networks (ANNs) to estimate the compressive strength based on relevant parameters, such as the water–cement ratio, aggregate–cement ratio, age of testing, and percentage cement/metakaolin ratios (5% and 10%). An ANN model was developed, different ANN configurations were tested and compared to identify the best ANN model. Using this model, it was possible to assess the contribution of each input variable to the compressive strength of the tested concretes. The results indicate an excellent performance of the ANN model developed to predict compressive strength from the input parameters studied, with an average error less than 5%. (Silva *et al.*, 2021) (Garzali Gali et al., 2023)

Convolutional Neural Networks (CNNs) have emerged as specialized architectures for image-based predictions and crack detection. This paper investigates the potential of deep learning in predicting the compressive strength of ultra-high-performance concrete (UHPC) by developing a convolutional neural network (CNN) model with two convolutional layers. The proposed CNN architecture demonstrates capability in accurately predicting UHPC compressive strength from tabular data encompassing various material compositions. (Xu *et al.*, 2025) Long Short-Term Memory (LSTM) networks capture temporal dependencies in concrete degradation processes. This research employs a deep learning methodology to examine the influence of fly ash, slag, superplasticizer, cement, water, coarse aggregate, and fine aggregate on the strength of concrete. This review utilises data mining techniques and the ABC algorithm for feature selection. Following the application of min-max normalisation to the dataset, a Vanilla LSTM network was employed for modelling. With a prediction accuracy of 96.48%, VannilaLSTM demonstrated superior efficacy among the four deep learning models assessed for discerning complex patterns influencing concrete strength. (Gurrala *et al.*, 2025)

2.2 Ensemble Methods and Hybrid Approaches

Ensemble learning combines multiple base learners to achieve superior predictive performance. The ML models included both non-ensemble and ensemble types. The non-ensemble models were regression-based, evolutionary, neural network, and fuzzy-inference-system. Meanwhile, the ensemble models consisted of adaptive boosting, random forest, and gradient boosting. A sensitivity analysis using Shapley-Additive-exPlanations (SHAP) was conducted to understand better how each input variable affects CS. The findings showed that the Categorical-Gradient-Boosting (CatBoost) model was the most accurate prediction during the testing stage. It had the highest determination-coefficient (R^2) of 0.966 and the lowest Root-Mean-Square-Error (RMSE) of 3.06 MPa. (Elshaarawy, Alsaadawi and Hamed, 2024)

Four predictive algorithms, XGBoost, ADABoost, CATBoost, and GBM, are employed to predict the tensile strengths of HPC. The computational efficiency of the proposed models was assessed using several performance metrics. The XGBoost model has the highest performance in training, testing, and validation, followed by ADABoost, GBM, and CATBoost. (Kumar *et al.*, 2025) Random Forest methods have proven effective for handling non-linear relationships. Comparative analysis of prediction accuracy, residuals, and feature importance revealed that while both models performed well, random forest regression outperformed AdaBoost regression, achieving MAE = 0.69, MAPE = 0.01, MSE = 0.87, RMSE = 0.93 MPa, SI = 0.02, MBE = -0.05, R2 = 0.99 and a20_index = 1. (Punitha *et al.*, 2025)

2.3 Optimization Algorithms and Hyperparameter Tuning

The DMP-PSO algorithm dynamically optimizes the hyperparameters of the BPNN, achieving superior prediction accuracy. The predictive performance significantly outperforms that of BPNNs optimized using PSO, Bayesian optimization (BO), and grid search methods. (Zhang *et al.*, 2025) Bayesian optimization and genetic algorithms enhance model performance through systematic hyperparameter exploration. Utilizing the Keras Tuner library, the most effective hyperparameters for the DNN models were identified. These models are then trained and evaluated using an experimental dataset of 1030 instances. The optimal configuration for the hidden layers in the model was identified as five layers, containing 12, 16, 16, 40, and 26 neurons, respectively. The optimal learning rate for the model was 0.001. (Naved, Asim and Ahmad, 2024)

2.4 Data-Driven Surrogate Models and Physics-Informed Learning

Emerging approaches integrate physical knowledge with data-driven learning. Machine learning (ML) models such as deep neural networks, Gaussian processes, and ensemble learners have been trained to approximate finite element analyses within iterative optimization loops. (Azanaw, 2025) This study proposes a physics-guided probabilistic deep learning framework integrating Bayesian inference and physics-informed neural network (PINN) for modeling corrosion-induced cracking in reinforced concrete. Unlike existing PINN-based approaches that yield only deterministic predictions, the proposed framework embeds the governing physical equations directly into the Bayesian neural network training process, simultaneously enforcing physical consistency and quantifying prediction uncertainty. (Zhang, Wang and Li, 2026)

Model Type	Base Algorithm	Key Advantages	Typical R^2 Range	Applications
ANN	Backpropagation	Fast training, flexible architecture	0.88-0.96	Strength prediction, fresh properties
CNNs	Convolutional layers	Image feature extraction, crack detection	0.85-0.98	Visual inspection, defect detection
Random Forest	Ensemble decision trees	Robustness, feature importance analysis	0.92-0.99	Mix design optimization
XGBoost/CatBoost	Gradient boosting	High accuracy, handles non-linearity	0.94-0.97	Multi-property prediction
LSTM Networks	Recurrent layers	Temporal dependency capture	0.90-0.96	Durability modeling, time-series
Physics-Informed NNs	PINNs	Physical constraint enforcement	0.90-0.98	Degradation mechanisms

Table 1: Comparison of Machine Learning Models for Concrete Performance Prediction. This table summarizes key ML approaches, their algorithmic basis, comparative advantages, typical prediction accuracy ranges (R^2 coefficients of determination), and primary civil engineering applications based on reviewed literature.

3. Results: Performance Across Concrete Prediction Tasks

3.1 Compressive Strength Prediction

Compressive strength remains the most extensively studied concrete property with multiple datasets now publicly available. Super learner models, which are based on several ensemble methods including random forest regression (RFR), an adaptive boosting (AdaBoost), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), categorical gradient Boosting (CatBoost), are used to solve this complicated problem. The results on four popular datasets show significant improvement of the proposed super learner models in terms of prediction accuracy. It also reveals that their trained models always perform better than other methods since their errors (MAE, MSE, RMSE) are always much lower and values of R2 are higher than those of the previous studies. (Lee *et al.*, 2022)

In this study, 10 different machine learning models were evaluated for their capacity to predict the elastic modulus of UHPC. The results showed that XGBoost demonstrated the highest accuracy in predictions with large training datasets, followed by KNNs. For smaller training datasets, Decision Tree exhibited the greatest accuracy, while XGBoost was the second-best performing model. (Zhang *et al.*, 2024) Specialized concrete formulations have received targeted attention. A dataset comprising 308

UHPC compressive strength tests with varying amounts of cement, silica fumes, superplasticizers, ground granulated blast furnace slag, fine aggregates, coarse aggregates, age, and temperature was used to train these models. Among the models, the RF model achieved the highest prediction accuracy with an R2 of 0.9877 and an RMSE of 0.0323. (Kumar *et al.*, 2024)

3.2 Tensile and Shear Strength Prediction

Results demonstrate that SGB achieved the best performance for predicting strength enhancement, with an R2 of 0.850, the lowest RMSE (0.190), and the highest generalization capability, making it the most reliable model for strength enhancement predictions. For strain enhancement prediction, XGB outperformed other models, achieving an R2 of 0.779 with the lowest RMSE (2.162). (Shang *et al.*, 2025) The results show that Levy-DT outperforms all other models, achieving the highest prediction accuracy with strong generalization. SHAP analysis is carried out, identifying axial force and reinforcement depth as key contributors to the shear strength estimation. (Çiftçioğlu *et al.*, 2025) This study pioneers the application of an Artificial Neural Network (ANN) to model these strengths and to classify failure modes in beams. Leveraging a dataset of 116 experimental tests on ultimate strengths from extensive literature, the ANN was meticulously trained, tested, and validated, revealing that the optimal neuron count for the modeling task was 15. This configuration achieved a root mean square error (RMSE) of 0.096 MPa and a coefficient of determination (R^2) of 0.95, outperforming traditional design models. (Nassif *et al.*, 2024)

3.3 Fracture Energy and Durability Modeling

The fracture energy of fiber-reinforced concrete (FRC) affects the durability and structural performance of concrete elements. A dataset of 500 data points was employed to train and evaluate the machine learning techniques. The findings showed that Gaussian process regression (GPR) outperforms all other models in terms of predictive accuracy, achieving the highest R 2 of 0.93 and the lowest RMSE of 13.91 during holdout cross-validation. (Al-Shamasneh *et al.*, 2025)

This study develops three hybrid predictive models: artificial neural network (ANN) optimized with genetic algorithm (GA), particle swarm optimization (PSO), and the Levenberg–Marquardt (LM) algorithm. The results demonstrated that all three hybrid models outperformed the reference study, with the ANN–LM model yielding the highest prediction accuracy. (Khosravi *et al.*, 2025) Self-healing concrete performance has gained research attention. Nine advanced machine learning algorithms were trained and evaluated using fivefold cross-validation. TabPFN delivered exceptional accuracy (R2 = 0.942, RMSE < 0.072 mm), surpassing ensemble methods (XGBoost, RFR, GBR) and kernel-based approaches (SVR, NuSVR, GPR). (Mahmoodzadeh *et al.*, 2025)

3.4 Structural Health Monitoring and Damage Assessment

Seismic performance prediction combines multiple input variables with complex nonlinear relationships. Various supervised machine learning models were compared. Results indicate that the Random Forest model predicts seismic performance scores with high accuracy. These findings suggest that such approaches can serve as fast, effective, and reliable decision-support tools for post-earthquake building risk prioritization. (Sağlam *et al.*, 2026) Reinforced concrete infrastructure condition monitoring requires ensemble approaches. Stacking Ensemble demonstrated superior overall performance with accuracy of 72.8%, precision of 82.8%, recall of 72.8%, F1-score of 75.9%, and RMSE of 0.5334. Feature importance analysis revealed that the top four predictors—age, length, pipe diameter, and depth—collectively contributed 81.6% influence on condition prediction. (Mohammadagha *et al.*, 2025)

The overall accuracy score for rectangular RC columns is 94% and for circular RC columns is 86%. The latter performances are influenced by the size of the testing and training sets of data and are independent of the number of decision trees in the forest employed in the random forest algorithm. (Megalooikonomou and Beligiannis, 2023)

Prediction Task	Best-Performing Model	Dataset Size	R^2 / Accuracy	RMSE/MAE	Key Features
Compressive Strength	CatBoost	1,030	0.966	3.06 MPa	Cement, water, aggregates, age
Tensile Strength (HPC)	XGBoost	714	0.912	Reported	Mix composition, curing
Shear Strength (RC Beams)	Levy-DT	195	Not reported	Lowest RMSE	Depth, reinforcement ratio
Fracture Energy (FRC)	GPR	500	0.93	13.91	Fiber volume, length, matrix
Elastic Modulus (UHPC)	XGBoost	162+	0.961	Reported	Silica fume, water/cement ratio
Seismic Damage Score	Random Forest	3,543	High accuracy	< 0.15 MSE	Geometric, structural parameters
CMOD (UHPC)	TabPFN	Multi-source	0.942	<0.072 mm	Fiber properties, w/b ratio

Table 2: Summary of Best-Performing Models Across Concrete Prediction Tasks. This table aggregates results from studies examining different concrete structural performance metrics, identifying the optimal algorithm for each application, reporting dataset sizes, and noting achieved accuracy metrics and primary input variables.

4. Crack Detection and Visual Inspection Applications

Advanced computer vision integrated with deep learning has transformed crack detection from manual inspection to automated assessment. This paper systematically discusses the deep learning in the identification of concrete structure cracks, expounds the deep learning technology, studies the SqueezeNet network model and YOLO (You Only Look Once) network model in the field of concrete structure crack identification. The results show that the classification accuracy of the improved SqueezeNet model is more than 82% and the recognition accuracy of YOLO model is more than 92%. (Bian, 2024)

Transfer learning strategies enable efficient model development with limited labeled data. The framework utilizes pre-trained Convolutional Neural Networks (CNNs), such as ResNet50, Xception, and InceptionV3, which are fine-tuned to address the unique challenges of concrete crack detection. The findings reveal that the ResNet50 model achieves remarkable accuracy, reaching 99.95% on the METU dataset and consistently exceeding 97% on the balanced SDNET2018 dataset. (Alkannad *et al.*, 2025)

Monitoring the structural health of infrastructure such as buildings, bridges, and roads requires efficient and accurate crack detection methods. This research proposes the integration of MobileNet V2 for concrete crack image feature extraction with the Improved Grasshopper Optimization Algorithm (IGOA) and XGBoost for classification. The experimental results show that the IGOA-XGBoost method provides the best performance with 99.75% accuracy, 99.80%

precision, 99.70% recall, and 99.75% F1-score. (Barkiah *et al.*, 2026)

Advanced techniques address real-world challenges. Detection and assessment of cracks in civil engineering structures such as roads, bridges, dams and pipelines are crucial tasks for maintaining the safety and cost-effectiveness of those concrete structures. However, most of the proposed models are trained on images taken in ideal conditions and are only capable of achieving high accuracy when applied to the concrete images devoid of irregular illumination conditions, shadows, shading, blemishes, etc. (Pal *et al.*, 2021) This study proposes a two-step rapid inspection method for underwater concrete bridge piers by combining acoustics and optics. The first step combines macroscopic sonar scanning with an enhanced YOLOv7 to detect and locate piers and defects. The results demonstrate an average mean average precision of 95.10% for defect and pier detection. (Sun *et al.*, 2024)

5. Discussion: Challenges, Limitations, and Critical Analysis

5.1 Data-Related Challenges

A critical limitation across ML applications remains heavy dependence on laboratory-generated datasets. The review reveals that most studies utilize existing literature-based datasets, with few contributing novel data and limited open access to these databases, which hampers broader validation and application. The review classifies the features analyzed in studies into categories such as mixture proportions, engineering properties, exposure conditions, test parameters, and chemical compositions, highlighting a growing emphasis on chemical compositions. (Taffese *et al.*, 2025)

Dataset size and diversity significantly impact generalization capacity. This dataset outlines the outcomes of an extensive 10-year laboratory investigation into concrete materials involving mechanical tests and non-destructive assessments within a comprehensive dataset denoted as ConcreteXAI. The dataset encompasses 18,480 data points, establishing itself as a cutting-edge resource for concrete analysis. (Guzmán-Torres *et al.*, 2024) Limited field-based validation remains a persistent gap. Studies predominantly utilize controlled laboratory conditions that may not capture real-world variability in material aging, environmental exposure, and complex loading scenarios.

5.2 Model Generalization and Transfer Learning

Generalization across diverse concrete formulations remains challenging. While ensemble methods demonstrate robust performance within specific datasets, their predictive capability often degrades when applied to concrete mixes substantially different from training data. This research utilizes a novel step transfer learning (STL) added extreme learning machine (ELM) approach to develop an automatic assessment strategy for surface cracks in concrete structures. The proposed model achieved at least 2.5%, 4.8%, and 0.8% improvement in accuracy, recall, and precision, respectively, in comparison to the other studies, indicating that the proposed model could aid in the automated inspection of concrete structures, ensuring high generalization ability. (Sohaib *et al.*, 2024)

5.3 Model Interpretability and Explainability

Recent emphasis on explainable AI through SHAP and similar techniques provides insight into feature importance but masks potential limitations in understanding failure modes. The review stresses the growing importance of model explainability, noting that model-agnostic methods like SHAP are frequently used and that the focus on explainability has increased. (Taffese *et al.*, 2025) However, explainability does not guarantee causal understanding or robustness to distribution shifts.

5.4 Computational Efficiency and Deployment

While many studies achieve excellent accuracy, practical deployment constraints including computational cost, real-time inference requirements, and embedded systems limitations receive insufficient attention. We present an embedded AI model for real-time concrete crack detection, considering the trade-offs between model complexity and accuracy. The research provides a quantitative comparison between MobileNet versions 1 and 2, and two CNN models with different numbers of layers, aiming to create an efficient and accurate solution for crack detection on surfaces of concrete structural members. (Dantas *et al.*, 2024)

Challenge Category	Specific Issues	Impact Level	Proposed Mitigation
Data Scarcity	Limited field datasets, lab-centric focus	Critical	Develop standardized field testing protocols
Generalization	Poor transfer across formulations	High	Multi-task learning, domain adaptation
Interpretability	Black-box predictions, limited causality	Medium	Physics-informed models, feature visualization
Computational Cost	High inference latency, memory requirements	Medium	Model compression, edge AI deployment
Class Imbalance	Rare failure modes underrepresented	High	Weighted loss functions, synthetic data generation
Validation Gap	Limited full-scale structural testing	Critical	Long-term field monitoring programs
Uncertainty Quantification	Confidence intervals often missing	High	Bayesian approaches, ensemble methods
Standards Integration	Insufficient alignment with design codes	High	Code-level validation studies

Table 3: Critical Challenges in ML-Based Concrete Performance Prediction. This table systematically categorizes implementation challenges, assesses their impact level on practical deployment, and proposes evidence-based mitigation strategies identified from reviewed literature.

6. Research Gaps and Future Directions

6.1 Multi-Task and Multi-Output Learning

Current approaches typically predict single properties independently, missing opportunities for joint optimization. A multi-output artificial neural network (ANN) with a 7-14-9-14-4 architecture was trained on an N = 180 dataset compiled from this study and published literature, achieving R^2 values of 0.989, 0.985, 0.977, and 0.981 for flexural strength, ultimate load, mid-span deflection, and compressive strength respectively—outperforming XGBoost, Random Forest, and SVR. (Maidase *et al.*, 2026) Future work should systematically develop frameworks predicting multiple performance attributes simultaneously while capturing interdependencies.

6.2 Physics-Informed and Hybrid Frameworks

The durability assessment of fly ash-based geopolymer high-performance concrete (HPC) under extreme environment treatment remains a subject of critical concern. To enhance this performance, the present research develops a unique integrated framework based on physics informed and deep learning type analytical designs built to predict the durability and residual service life of geopolymer HPC subjected to freezing-and-thawing, sulfate, and thermal degradation. (Vairagade, 2025) Integration of domain knowledge, mechanistic models, and data-driven learning represents critical future direction.

6.3 Real-Time Monitoring and Digital Twins

Accurate long-term deformation prediction is essential to ensure the structural security and ongoing stability of large water-related concrete structures like ultra-high arch dams. This study proposes a multi-point deformation forecasting model that incorporates both spatial and temporal correlations between environmental factors and deformation, utilizing advanced deep learning (DL) techniques. The Transformer-based convolutional long short-term memory (ConvLSTM) model captures the spatiotemporal dependencies across numerous temperature and deformation monitoring sequences. (Gu *et al.*, 2025)

6.4 Integration with Non-Destructive Testing

This study developed a deep learning-based framework utilizing a convolutional neural network (CNN) trained on 284 samples with 11 critical features related to material composition and environmental conditions. The CNN's performance was benchmarked against four machine learning (ML) models, achieving a coefficient of determination (R2) = 0.849 and a lower root mean square error (RMSE) = 0.18%, outperforming conventional models. (Abdellatief *et al.*, 2025) Future research should systematically incorporate ultrasonic, impact-echo, ground-penetrating radar, and acoustic emission data into predictive frameworks.

6.5 Uncertainty Quantification and Probabilistic Predictions

This study explored the use of artificial neural networks to predict UHPC compressive strengths given thermal history and key mix components. The model developed herein employs Bayesian variational inference using Monte Carlo dropout to convey prediction uncertainty. On average, 8.2% of experimental results in the final model fell outside of the predicted range with 67.9%of these cases conservatively underpredicting. (Roberson *et al.*, 2022) Probabilistic frameworks providing confidence intervals and risk quantification require expansion beyond current deterministic approaches.

Research Gap	Current State	Recommended Direction	Expected Impact
Multi-Output Prediction	Single-property focus	Joint optimization with interdependency modeling	Reduced total model ensemble
Physics Integration	Data-driven dominated	PINN + mechanistic model hybrid	Enhanced interpretability, robustness
Field Validation	Lab-centric	Long-term in-situ monitoring networks	Practical deployment readiness
Uncertainty Quantification	Mostly absent	Bayesian, ensemble probabilistic	Risk-aware decision support
Temporal Dynamics	Limited exploration	LSTM, Transformer for degradation paths	Lifecycle prediction capability
Material Heterogeneity	Insufficient modeling	Microstructure-informed deep learning	Mechanistic understanding
Digital Twins	Emerging concept	Real-time coupled prediction-monitoring	Continuous optimization
Code Integration	Negligible	Validation against design standards	Regulatory acceptance

Table 4: Identified Research Gaps and Recommended Future Directions. This table documents areas where current ML literature shows limitations, recommends specific research directions based on identified needs, and projects expected impact on advancing practical implementation.

7. Conclusion and Recommendations

7.1 Synthesis of Key Findings

Machine learning techniques have demonstrated remarkable capacity to predict diverse concrete structural performance metrics with accuracies frequently exceeding 95% R^2 when applied to appropriate datasets. Ensemble methods—particularly XGBoost, CatBoost, and Random Forest—emerge as optimal general-purpose approaches, while specialized architectures (CNNs for imaging, LSTMs for temporal data, physics-informed NNs for mechanism-driven tasks) excel within their respective domains. Civil engineering is experiencing a paradigm shift driven by artificial intelligence (AI), the Internet of Things (IoT), automation, and sustainable practices. Machine learning models accurately predict compressive strength and embodied carbon of eco-friendly concrete, while Artificial Intelligence based life cycle analysis improves the sustainability assessment of geopolymer concretes. (M. and N, 2025)

The field demonstrates maturation through increasing emphasis on model explainability, uncertainty quantification, and physics-informed approaches. However, critical gaps persist between research capabilities and practical deployment. Laboratory-centric datasets limit generalization; limited integration with design codes restricts regulatory acceptance; and insufficient full-scale validation hampers confidence in field applications. The disconnect between academic excellence in prediction accuracy and practical implementation challenges represents the foremost obstacle to widespread adoption.

7.2 Strategic Recommendations

For Researchers: (1) Prioritize development of diverse, field-based datasets with open-access architecture encouraging validation across multiple research groups; (2) systematically investigate transfer learning and domain adaptation techniques to enable generalization across formulations; (3) integrate physics-based knowledge constraints into data-driven models to improve robustness and interpretability; (4) develop probabilistic frameworks quantifying prediction uncertainty for risk-aware engineering decisions; (5) conduct comparative studies benchmarking multiple algorithms on identical datasets to establish performance baselines.

For Industry and Practice: (1) Establish guidelines for responsible deployment of ML models in design and construction workflows, including validation protocols and performance monitoring; (2) integrate ML-based optimization into concrete mix design processes, leveraging prediction models to accelerate innovation; (3) invest in structural health monitoring systems that collect comprehensive datasets for model training and real-world validation; (4) adopt digital twin frameworks coupling real-time monitoring with predictive models to enable proactive maintenance.

For Code and Standards Development: (1) Convene working groups examining integration of data-driven prediction methods into design codes; (2) establish acceptance criteria and validation requirements for ML-based predictions in structural design; (3) develop harmonized datasets and benchmark problems enabling cross-disciplinary model comparison; (4) create guidelines for uncertainty quantification and probability-based safety factors reflecting ML prediction confidence.

7.3 Concluding Statement

Machine learning represents a transformative technology with unprecedented potential to accelerate concrete structural design, optimize material usage, enhance durability predictions, and improve infrastructure safety. The technical foundation proves robust, with demonstrated prediction capabilities across diverse applications. Yet realizing this potential requires bridging gaps between algorithmic sophistication and practical utility through concerted effort toward standardization, field validation, code integration, and collaborative knowledge sharing. The convergence of advanced computational methods, expanded computational resources, and growing datasets positions the field at an inflection point. Strategic investment in translating laboratory advances into field-deployed systems will determine whether ML-based concrete prediction evolves from academic achievement to engineering practice essential to building resilient, sustainable infrastructure for future generations.

REFERENCES

Abdellatief, M. *et al.* (2025) 'A convolutional neural network-based deep learning approach for predicting surface chloride concentration of concrete in marine tidal zones', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-12035-1.

Alibrahim, B., Habib, A. and Habib, M. (2025) 'Developing a brain inspired multilobar neural networks architecture for rapidly and accurately estimating concrete compressive strength', *Scientific Reports* [Preprint]. doi:10.1038/s41598-024-84325-z.

Alkannad, A.A. *et al.* (2025) 'CrackVision: Effective concrete crack detection with deep learning and transfer learning', *IEEE Access* [Preprint]. doi:10.1109/ACCESS.2025.3540841.

Al-Shamasneh, A.R. *et al.* (2025) 'High-fidelity machine learning framework for fracture energy prediction in fiber-reinforced concrete', *Computer Modeling in Engineering & Sciences* [Preprint]. doi:10.32604/cmes.2025.068887.

Azanaw, G. (2025) 'Blending data-driven surrogates with physics - based topology optimization: A critical review of machine learning - accelerated design in fibre - reinforced polymer and concrete structures', *American Journal of Science, Engineering and Technology* [Preprint]. doi:10.11648/j.ajset.20251003.11.

Barkiah, I. *et al.* (2026) 'A machine learning model based on MobileNet V2 and IGOA-XGBoost for structural health monitoring', *International Journal of Civil Engineering* [Preprint]. doi:10.14445/23488352/ijce-v13i5p117.

Bian, Z. (2024) 'Identification of concrete crack using deep learning based approach', *Advances in Research* [Preprint]. doi:10.9734/air/2024/v25i51160.

Çiftçioğlu, A.Ö. *et al.* (2025) 'Machine learning based shear strength prediction in reinforced concrete beams using levy flight enhanced decision trees', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-12359-y.

Dantas, H.C. *et al.* (2024) 'Concrete crack detection using embedded machine learning', *2024 8th International Symposium on Instrumentation Systems, Circuits and Transducers (INSCIT)* [Preprint]. doi:10.1109/INSCIT62583.2024.10693377.

Elshaarawy, M.K., Alsaadawi, M.M. and Hamed, A.K. (2024) 'Machine learning and interactive GUI for concrete compressive strength prediction', *Scientific Reports* [Preprint]. doi:10.1038/s41598-024-66957-3.

Gu, H. *et al.* (2025) 'Environmental-aware deformation prediction of water-related concrete structures using deep learning', *Computer-Aided Civil and Infrastructure Engineering* [Preprint]. doi:10.1111/mice.13513.

Gurrala, R.R. *et al.* (2025) 'Deep learning-based hybrid framework for strength estimation of blended cement concrete', *2025 3rd International Conference on Data Science and Information System (ICDSIS)* [Preprint]. doi:10.1109/ICDSIS65355.2025.11071206.

Guzmán-Torres, J.A. *et al.* (2024) 'ConcreteXAI: A multivariate dataset for concrete strength prediction via deep-learning-based methods', *Data in Brief* [Preprint]. doi:10.1016/j.dib.2024.110218.

Khosravi, H. *et al.* (2025) 'Prediction of crack repair percentage in self-healing concrete using machine learning', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-30158-3.

Kumar, D.R. *et al.* (2025) 'Optimized machine learning models for predicting the tensile strength of high-performance concrete', *Journal of Structural Integrity and Maintenance* [Preprint]. doi:10.1080/24705314.2025.2471682.

Kumar, R. *et al.* (2024) 'Development of a prediction tool for the compressive strength of ternary blended ultra-high performance concrete using machine learning techniques', *Journal of Structural Integrity and Maintenance* [Preprint]. doi:10.1080/24705314.2024.2385206.

Lee, S. *et al.* (2022) 'Super learner machine-learning algorithms for compressive strength prediction of high performance concrete', *Structural Concrete* [Preprint]. doi:10.1002/suco.202200424.

M., M. and N, U. (2025) 'Artificial intelligence and smart technologies for sustainable civil engineering', *International Journal of Advanced Engineering Research and Science* [Preprint]. doi:10.22161/ijaers.1210.5.

Mahmoodzadeh, A. *et al.* (2025) 'Machine learning-based prediction of crack mouth opening displacement in ultra-high-performance concrete', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-23610-x.

Maidase, P.D. *et al.* (2026) 'Structural performance of recycled aggregate concrete beams under flexural loading: Experimental investigation, mix design optimisation, and machine learning-based multi-output prediction', *International Journal of Drug Delivery Technology* [Preprint]. doi:10.25258/ijddt.16.43s.106.

Megalooikonomou, K. and Beligiannis, G. (2023) 'Random forests machine learning applied to PEER structural performance experimental columns database', *Applied Sciences* [Preprint]. doi:10.3390/app132312821.

Mohammadagha, M. *et al.* (2025) 'Machine learning models for reinforced concrete pipes condition prediction: The state-of-the-art using artificial neural networks and multiple linear regression in a wisconsin case study', *Innovative Infrastructure Solutions* [Preprint]. doi:10.1007/s41062-025-02459-z.

Nassif, N. *et al.* (2024) 'Performance-based prediction of shear and flexural strengths in fiber-reinforced concrete beams via machine learning', *Structural Engineering International* [Preprint]. doi:10.1080/10168664.2024.2310520.

Naved, M., Asim, M. and Ahmad, T. (2024) 'Prediction of concrete compressive strength using deep neural networks based on hyperparameter optimization', *Cogent Engineering* [Preprint]. doi:10.1080/23311916.2023.2297491.

Neira, P. *et al.* (2020) 'Prediction of concrete compressive strength through artificial neural networks', *Journal of the Croatian Association of Civil Engineers* [Preprint]. doi:10.14256/jce.2438.2018.

Ogunsaya, O.A. and Taiwo, R. (2024) 'Enhancing concrete structures: Integrating machine learning and deep learning for optimizing material strength, fire resistance, and impact protection', *World Journal of Advanced Research and Reviews* [Preprint]. doi:10.30574/wjarr.2024.23.3.2697.

Pal, M. *et al.* (2021) 'An overview of challenges associated with automatic detection of concrete cracks in the presence of shadows', *Applied Sciences* [Preprint]. doi:10.3390/app112311396.

Punitha, A. *et al.* (2025) 'Prediction of compressive strength of blended concrete with alccofine and GGBFS by applying ensemble machine learning algorithms', *Journal of Structural Integrity and Maintenance* [Preprint]. doi:10.1080/24705314.2025.2489858.

Roberson, M.M. *et al.* (2022) 'Probabilistic neural networks that predict compressive strength of high strength concrete in mass placements using thermal history', *Computers & structures* [Preprint]. doi:10.1016/j.compstruc.2021.106707.

Sağlam, R. *et al.* (2026) 'Seismic performance prediction of reinforced concrete buildings using the RYTEIE method: A comparison of machine learning models', *International Journal of Pure and Applied Sciences* [Preprint]. doi:10.29132/ijpas.1928087.

Shang, L. *et al.* (2025) 'Prediction of ultimate strength and strain in FRP wrapped oval shaped concrete columns using machine learning', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-95272-8.

Silva, F.A.N. *et al.* (2021) 'Use of nondestructive testing of ultrasound and artificial neural networks to estimate compressive strength of concrete', *Buildings* [Preprint]. doi:10.3390/BUILDINGS11020044.

Sohaib, M. *et al.* (2024) 'Generalizing infrastructure inspection: Step transfer learning aided extreme learning machine for automated crack detection in concrete structures', *Measurement Science and Technology* [Preprint]. doi:10.1088/1361-6501/ad296c.

Sun, W. *et al.* (2024) 'Two-step rapid inspection of underwater concrete bridge structures combining sonar, camera, and deep learning', *Computer-Aided Civil and Infrastructure Engineering* [Preprint]. doi:10.1111/mice.13401.

Taffese, W.Z. *et al.* (2025) 'Machine learning in concrete durability: Challenges and pathways identified by RILEM TC 315-DCS towards enhanced predictive models', *Materials and Structures* [Preprint]. doi:10.1617/s11527-025-02664-3.

Vairagade, V.S. (2025) 'Durability and service life prediction of fly ash based geopolymer high performance concrete under extreme environmental conditions', *Scientific Reports* [Preprint]. doi:10.1038/s41598-025-27229-w.

Xu, L. *et al.* (2025) 'Prediction of ultra-high-performance concrete (UHPC) compressive strength based on convolutional neural networks', *Materials* [Preprint]. doi:10.3390/ma18122851.

Zhang, C. *et al.* (2024) 'Elastic modulus prediction of ultra-high-performance concrete with different machine learning models', *Buildings* [Preprint]. doi:10.3390/buildings14103184.

Zhang, Y. *et al.* (2025) 'Intelligent structural design of composite concrete-encased steel columns based on hybrid machine learning and multiobjective optimization', *Structural Concrete* [Preprint]. doi:10.1002/suco.70292.

Zhang, Y., Wang, K. and Li, L. (2026) 'Physics-guided probabilistic deep learning for modeling the cracking process of corroded reinforced concrete', *Advances in Structural Engineering* [Preprint]. doi:10.1177/13694332261459980.

Garzali Gali, Adnan Khan, Salman khan, Shahzad khalil, Qaim Shah, & Murad Khan. (2023). Experimental Study on the Effects of Freeze-Thaw Progressions and Performance of Soil with Non-Toxic Bio-Enzyme. Journal of ICT, Design, Engineering and Technological Science, 7(1), 12–20. https://doi.org/10.33150/JITDETS-7.1.3

Shah, Q., Ullah, I., Khoso, W. A., Iqbal, F., & Umali, M. (2026). Explainable and comparative machine learning models for punching shear prediction in reinforced concrete flat slabs. Multiscale and Multidisciplinary Modeling, Experiments and Design, 9(1), 146. https://doi.org/10.1007/s41939-026-01219-5

Shah, Q., Ullah, I., Khoso, W. A., & Umali, M. (2026). Interpretable machine learning for predicting reinforced concrete beam column joint shear strength. Discover Civil Engineering, 3(1), 153. https://doi.org/10.1007/s44290-026-00539-2