Građevinski materijali i konstrukcije
Building Materials and Structures
journal homepage: www.dimk.rs

https://doi.org/10.5937/GRMK2400012K

Review paper

Application of machine learning in asphalt and concrete material testing: a comprehensive review

Meisam Khorshidi[*1)] Eshan Dave[1)] Jo Sias[1)]

[1)] *Department of Civil and Environmental Engineering, University of New Hampshire, 33 Academic Way, Durham, NH 03824, USA*

Article history

Received: 21 October 2024
Received in revised form: 27 November 2024
Accepted: 30 November 2024
Available online: 20 December 2024

Keywords

predictive modeling,
material performance prediction,
pavement distress classification,
unsupervised learning,
ensemble methods,
hybrid models,
artificial neural networks (ANN),
gaussian process regression (GPR)

ABSTRACT

This literature review explores the application of machine learning (ML) techniques in civil engineering material testing, with a focus on asphalt mixtures, concrete properties, and pavement system classification. The review provides a comprehensive comparison of various ML models, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forest (RF), Gradient Boosting (GB), and Gaussian Process Regression (GPR), assessing their strengths and limitations in predicting material performance. Key findings indicate that ensemble methods, such as Gradient Boosting and XGBoost, consistently outperformed other models in terms of prediction accuracy and handling nonlinear relationships, although they require significant computational power. In contrast, simpler models like SVM and ANN demonstrated strong predictive capabilities with smaller datasets but were prone to overfitting and computational challenges. Additionally, unsupervised learning methods, such as K-means clustering and Principal Component Analysis (PCA), proved effective in classifying pavement conditions and detecting anomalies, with K-means offering simplicity and efficiency at the cost of sensitivity to initialization and cluster definitions. The review concludes by emphasizing the potential of hybrid and ensemble models to improve prediction accuracy and reduce computational costs, highlighting the need for further research to address data availability, model interpretability, and practical implementation challenges in real-world applications.

1 Introduction

Civil engineering has long relied on empirical methods and extensive experimental testing to evaluate the performance of materials, structures, and infrastructure systems. However, the increasing complexity of modern construction projects and the growing need for more accurate predictions of material behavior under varying conditions have led to a shift towards more data-driven approaches. In this context, machine learning (ML) has emerged as a powerful tool for advancing civil engineering, particularly in the field of material testing and performance prediction [1], [2].

Machine learning enables civil engineers to analyze vast amounts of experimental data, detect patterns, and build predictive models that can forecast material behavior under different loading conditions, environmental factors, and time frames. With the ability to model nonlinear relationships and optimize multiple variables simultaneously, ML offers significant advantages over traditional statistical and empirical models [3], [4]. It can enhance decision-making processes in areas such as material design, optimization, and failure prediction, thereby reducing costs, increasing efficiency, and improving overall performance [2], [5].

1.1 Machine learning in material testing

In civil engineering, material testing is critical for determining the properties of construction materials such as asphalt, concrete, and fiber-reinforced composites. These materials exhibit complex behaviors when subjected to stress, temperature changes, and aging. Machine learning models can simulate these behaviors and offer insights that would otherwise require costly and time-consuming physical tests [6]–[8].

For example, ML algorithms are used to predict key material properties such as compressive strength, modulus of elasticity, tensile strength, rut depth, fracture energy, and more. Techniques such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RF), and Gradient Boosting (GB) have demonstrated strong predictive capabilities in areas like asphalt mixture performance and concrete strength estimation. These models not only improve the accuracy of predictions but also allow for the integration of a wide range of input parameters,

* Corresponding author:
E-mail address: meisam.khorshidi@unh.edu

such as material composition, environmental conditions, and load types [6], [9]–[11].

1.2 Advantages and challenges

The application of machine learning in material testing offers several advantages. First, it reduces the reliance on extensive experimental testing by providing accurate predictions based on historical data. This is especially beneficial in large-scale infrastructure projects with limited time and resources [12]–[14]. Second, ML models are highly flexible, and able to account for nonlinear interactions between multiple variables, thus offering deeper insights into how different factors influence material behavior [15], [16]. Finally, these models can be continually improved as more data becomes available, leading to more refined predictions over time [17]–[19].

However, the adoption of machine learning in civil engineering also presents challenges. One of the primary concerns is the availability and quality of data. ML models require large datasets to function effectively, and inconsistencies in data collection can lead to inaccurate predictions [20], [21]. Moreover, the "black box" nature of some machine learning algorithms, particularly deep learning models, may hinder the interpretability of results, making it difficult for engineers to trust the outcomes without a clear understanding of how predictions were generated [22].

1.3 Current trends in research

Recent research in civil engineering has explored the use of machine learning models to solve complex material testing problems, including asphalt mixture performance and optimization, concrete property prediction and structural performance, and classification and pattern recognition in pavement systems. Many studies have demonstrated the effectiveness of machine learning in improving accuracy, reducing experimental costs, and providing actionable insights for material design and testing. For example, Artificial Neural Networks (ANNs) have been widely used to predict the compressive strength of concrete, while Support Vector Machines (SVMs) have shown strong performance in predicting fracture energy and elastic modulus in various materials [9], [10], [15]. Figure 1 illustrates key machine learning methods in civil engineering material testing, organized by learning type (supervised, unsupervised, ensemble, hybrid) and their applications in asphalt, concrete, and pavement analysis.

1.4 Scope of this review

This literature review aims to provide a comprehensive overview of the applications of machine learning in material testing within civil engineering, focusing on the following three key areas:

- Asphalt Mixture Performance and Optimization: Includes studies predicting the properties of asphalt mixtures, such as dynamic modulus, rut depth, and binder content, as well as optimizing asphalt mix designs.
- Concrete Property Prediction and Structural Performance: Covers the prediction of concrete properties like compressive strength, elasticity, and shear strength, as well as the performance of fiber-reinforced concrete.
- Classification and Pattern Recognition in Pavement Systems: Discusses studies that use machine learning to classify pavement distress, predict cracking patterns, and identify structural issues within pavement systems.

The review will discuss the different machine learning models used in literature, the key performance metrics they predict, and the pros and cons of each approach. Special emphasis will be placed on comparing multiple models applied simultaneously in material testing, as researchers increasingly use ensemble methods and comparative analysis to identify the best-performing models for specific engineering problems.

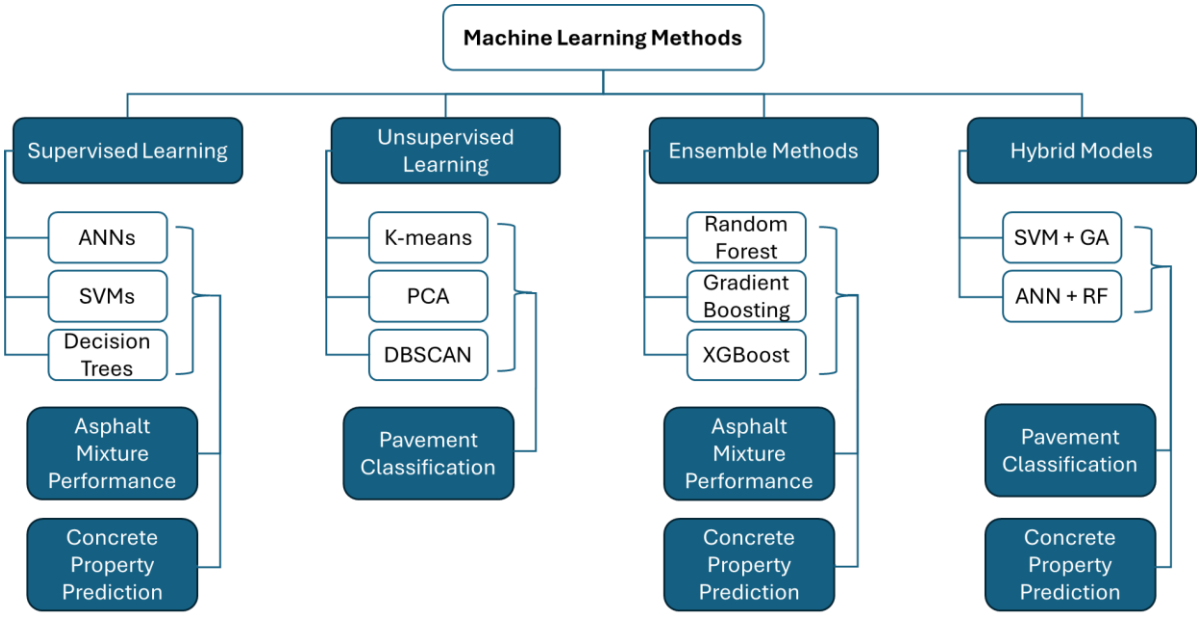

Figure 1. Machine learning methods in civil engineering material testing

2 Asphalt mixture performance and optimization

Machine learning has played a transformative role in the prediction and optimization of asphalt mixture performance, with various models being applied to forecast essential performance metrics such as dynamic modulus, rut depth, and Marshall stability.

Fan et al. (2024) utilized a Back-Propagation Neural Network (BPNN) and Support Vector Machine (SVM) to predict the strength of asphalt mixtures across diverse service conditions. They considered inputs such as stress states (direct tensile, uniaxial compression, indirect tensile, and four-point bending), temperature ranges from −25°C to 35°C, and loading rates between 0.02 MPa/s and 0.5 MPa/s. While SVM achieved slightly better accuracy (R^2 of 0.9983) than BPNN (R^2 of 0.9979), BPNN performed better in terms of minimizing small errors (Mean Absolute Percentage Deviation: 0.067 vs. 0.145). This demonstrated that although SVM excels in accuracy, BPNN could offer more consistent performance in specific scenarios by reducing smaller errors [2].

Upadhya et al. (2022) applied ANN, SVM, Gaussian Processes (GP), and Random Forest (RF) to predict the Marshall stability of glass fiber-reinforced asphalt mixes. The input variables included bitumen content, glass fiber content, fiber length, and bitumen grade, while the output variable was Marshall stability. The study showed that SVM with a Pearson Universal Kernel (PUK) achieved the best results (CC = 0.8776 and RMSE = 1.9653), making it the most accurate model for this application. ANN demonstrated reliable performance but showed higher errors during the testing phase, while GP and RF performed competitively but were outperformed by SVM. The results emphasized SVM's strength in managing nonlinear relationships, though its effectiveness depended heavily on kernel tuning, which posed a challenge in some applications [6].

In the study by Rondinella et al. (2023), SVM and Categorical Boosting (CatBoost) were applied to predict the mechanical and volumetric properties of road pavement asphalt mixtures incorporating recycled materials such as construction and demolition waste (C&DW) and reclaimed asphalt pavement (RAP). The input variables included gyratory revolutions, RAP content, water content, and bitumen properties, with the output variables being indirect tensile strength (ITS) and saturated surface dry voids (SSDV). CatBoost demonstrated superior accuracy (R^2 = 0.9916 for ITS) compared to SVM (R^2 = 0.8828), particularly excelling in handling categorical data, which made it more efficient for complex material datasets [4].

The study by Khorshidi et al. (2023) investigated the effects of different proportions of alternative materials, including Reclaimed Asphalt Pavement (RAP), crumb rubber (CR), steel slag (SS), and waste engine oil (WEO), on the performance of asphalt mixtures. Using 44 mixtures with varying RAP (0–75%), WEO (0–15%), CR (0–15%), and SS (0% or 20%) contents, the study evaluated cracking resistance, rutting resistance, and moisture damage. Machine learning models, including feed-forward neural networks (FNN), generalized linear models (GLM), support vector regression (SVR), and Gaussian process regression (GPR), were applied to predict the optimal content combinations. GPR performed the best, accurately identifying the most suitable material ratios for different high-traffic conditions. While GPR effectively modeled complex relationships, its computational demands and tuning requirements were noted as challenges. The study concluded that GPR provided reliable predictions for optimizing the balance between cracking resistance, rutting resistance, and moisture damage in asphalt mixtures with recycled materials [23].

Another study conducted by Khorshidi et al. (2023) assessed the effects of RAP, CR, SS, and WEO on the cracking resistance of asphalt mixtures. Using performance indices from the Illinois Flexibility Index Test (I-FIT), a deep neural network (DNN) model was applied to predict mixture performance and was compared with linear and polynomial regression models. The DNN outperformed the other models, achieving a coefficient of determination (R^2) of 0.84, compared to 0.60 for linear and 0.66 for polynomial regression. DNN's advantages included its ability to capture complex nonlinear relationships, providing more accurate predictions. However, it required more data and computational resources. Overall, DNN proved to be a reliable model for predicting cracking resistance in asphalt mixtures with recycled materials [5].

Liu et al. (2023) applied multiple models, including SVR, KRR, ANN, Gradient Boosting (GB), and XGBoost, to predict the dynamic modulus ($|E^*|$) of asphalt mixtures. The input variables consisted of temperature, loading frequency, binder properties (such as viscosity and phase angle), and aggregate gradation. XGBoost delivered the highest accuracy (R^2 = 0.9867, RMSE = 2.7422) due to its ability to handle nonlinear interactions and prevent overfitting through regularization techniques. However, it required considerable computational resources, which posed a limitation for its scalability in large-scale applications. Other models like ANN, while effective, were prone to overfitting and required substantial hyperparameter tuning, which made them less practical for routine use [1].

Liu et al. (2022) further explored the prediction of rut depth using SVR, RF, ANN, and GB models. The input variables included traffic data (e.g., Equivalent Single Axle Loads, ESALs), climate conditions, pavement material properties (e.g., binder content, air voids), and structural attributes (layer thicknesses). GB was identified as the best-performing model, achieving an R^2 of 0.9236, showcasing its effectiveness in capturing nonlinear interactions within the dataset. While ANN also performed well (R^2 = 0.9021), it required more computational power and tuning. RF lagged in performance with lower accuracy, while SVR showed significant variance in predictions due to its sensitivity to parameter selection [24].

In a separate study, Liu et al. (2022) used machine learning models to predict effective asphalt content (Pbe) and absorbed asphalt content (Pba) in asphalt mixtures. Gradient Boosting was the top performer, with R^2 values of 0.9479 and 0.9459 for Pbe and Pba, respectively, excelling in managing nonlinear relationships. RF performed adequately but was less accurate compared to Gradient Boosting. SVR showed moderate accuracy but was more prone to performance drops when handling larger datasets [25].

Liu et al. (2022) also explored the prediction of the International Roughness Index (IRI) of asphalt pavements using Support Vector Regression (SVR), Random Forest (RF), Artificial Neural Networks (ANN), Gaussian Process Regression (GPR), Extra-Trees, and Gradient Boosting (GB), combined with dimensionality reduction techniques like Autoencoders (AE), Principal Component Analysis (PCA), and Recursive Feature Elimination (RFE). The input variables included temperature, Equivalent Single Axle Loads (ESALs), layer thickness, binder content, air voids, and aggregate gradation, while the output variable was IRI. The AE-GPR model demonstrated the highest accuracy

(R^2 = 0.939), outperforming other models by efficiently managing high-dimensional data. Autoencoders significantly improved model performance by reducing input noise and computational load. In contrast, models like SVR and RF showed lower performance due to overfitting and sensitivity to hyperparameter tuning, while ANN performed well but was computationally expensive [26].

Majidifard et al. (2020) employed Gene Expression Programming (GEP) to predict rut depth in asphalt mixtures subjected to the Hamburg Wheel-Tracking Test (HWTT). The model inputs included asphalt binder properties, aggregate size, and reclaimed asphalt content, while the output was rut depth. GEP outperformed ANN by providing explicit mathematical expressions, making the model more interpretable and offering engineers insights into the factors driving rutting behavior. However, GEP required careful tuning of parameters like chromosome length, limiting its practicality for complex datasets [27].

Rahman et al. (2021) explored various ensemble methods, including Extra-Trees, GB, and SVR, to predict rut depth and indirect tensile (IDT) strength in asphalt mixtures. Extra-Trees demonstrated the highest prediction accuracy (R^2 = 0.922 for rut depth), but it was sensitive to imbalanced data, which affected its generalization. GB and SVR provided more robust predictions across diverse datasets but required more extensive computational resources to minimize bias and ensure balanced predictions [9].

Tiwari et al. (2022) applied ANN with various activation functions (Exponential Linear Unit, ELU, and Hyperbolic Tangent, TanH) to predict the mechanical properties of asphalt mixes with industrial waste fillers. The input variables included air void content, filler type, and filler content, while the output variables included Marshall stability and indirect tensile strength. The TanH activation function performed better, achieving R^2 = 0.9967, though it required higher computational power due to increased complexity in capturing nonlinear relationships [28].

In another study, Tiwari et al. (2023) applied ANN with different activation functions, including TanH and ELU, to predict mechanical properties of asphalt mixtures with silica fume fillers. The TanH-SNN model achieved the highest accuracy (R^2 = 0.9988), outperforming other models in terms of capturing nonlinear relationships between inputs and outputs, though the increased complexity required more computational power [29].

Ali et al. (2021) used XGBoost to predict dynamic modulus in asphalt mixtures, outperforming traditional models such as the Witczak and Hirsch models (R^2 = 0.961). XGBoost's strength lay in its ability to handle complex nonlinear relationships and avoid overfitting, though its computational demands limited its practicality in smaller-scale applications. ANN models, while competitive, lacked interpretability and required more extensive resources to train [30].

Mirzaiyanrajeh et al. (2022) used ANN, Self-Validated Ensemble Modeling (SVEM), and Augmented Full Quadratic Model (AFQM) to predict low-temperature fracture energy of asphalt mixtures. ANN provided the highest accuracy but was computationally expensive, whereas SVEM, although slightly less accurate, was more efficient with small datasets, striking a balance between accuracy and computational efficiency [31].

Liu et al. (2023) utilized recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent units (GRU) for time series modeling to predict rutting depth. Input variables included historical rutting depth, temperature, and pavement properties. GRU outperformed both RNN and LSTM, achieving an R^2 value of 0.90. GRU's ability to retain long-term memory with fewer parameters made it more computationally efficient, though LSTM still performed well in capturing seasonal trends in the data [32].

Finally, Al-Sabaeei et al. (2023) employed XGBoost and Random Forest Regression (RFR) to predict mixing and compaction temperatures for bio-modified asphalt using crude palm oil (CPO) and tire pyrolysis oil (TPO) as modifiers. XGBoost outperformed RFR in predicting shear viscosity, but RFR demonstrated better accuracy for temperature predictions, with R^2 values of 0.96583 for mixing temperature and 0.96281 for compaction temperature. Both models excelled in accuracy but were limited by their high computational requirements [33].

2.1 Summary of methods

Table 1 provides a detailed summary of studies focusing on machine learning approaches and their applications in asphalt mixture performance prediction. Across the studies reviewed, several machine learning methods were employed to predict key asphalt mixture performance metrics, including dynamic modulus, rut depth, Marshall stability, International Roughness Index (IRI), and crack resistance. Each method presented unique strengths and weaknesses, as highlighted below:

- Artificial Neural Networks (ANNs): Frequently used for predicting complex performance metrics, ANNs demonstrated strong accuracy in predicting properties such as Marshall stability, dynamic modulus, and fracture energy. ANNs excel in capturing intricate nonlinear relationships between variables, especially when paired with activation functions like TanH and ReLU. However, their major drawbacks include computational expense, the need for large datasets, and a propensity for overfitting without careful tuning of hyperparameters. In studies by Upadhya et al. (2022) Tiwari et al. (2022), and Khorshidi et al. (2023), ANNs and DNNs performed well but required significant computational resources and hyperparameter optimization [5], [6], [28].
- Support Vector Machines (SVMs): SVM models, particularly when paired with kernel methods like the Pearson Universal Kernel (PUK), were highly accurate in predicting metrics such as strength and Marshall stability. SVMs excel in handling nonlinear relationships and are particularly effective with small- to medium-sized datasets. However, as seen in studies like Fan et al. (2024) and Upadhya et al. (2022), SVMs require careful kernel tuning and can struggle with large datasets due to high computational costs and sensitivity to hyperparameters [2], [6].
- Gradient Boosting (GB) and XGBoost: These ensemble learning methods consistently outperformed other models in predicting dynamic modulus, rut depth, and other asphalt mixture properties. XGBoost, in particular, has proven to be highly effective at managing nonlinear interactions, regularizing models to avoid overfitting, and delivering superior prediction accuracy. This method was widely used in studies such as Liu et al. (2023) and Ali et al. (2021), where XGBoost delivered top results in predicting dynamic modulus and shear viscosity [1], [30]. However, XGBoost requires significant computational power and tuning, which can limit its practicality in certain applications.
- Gaussian Process Regression (GPR): GPR excels at modeling complex nonlinear relationships and provides both predictions and uncertainty estimates. It is particularly effective for small to medium datasets but can be

computationally demanding and requires careful tuning of hyperparameters. In studies by Khorshidi et al. (2023), GPR outperformed other models in predicting the optimal combinations of alternative materials in asphalt mixtures, accurately balancing cracking resistance, rutting resistance, and moisture damage, though its high computational requirements were noted [23].

- Random Forest (RF): While RF models provided solid predictions, particularly in larger datasets, they generally lagged behind ensemble methods like Gradient Boosting in terms of accuracy. Studies such as Liu et al. (2022) and Rahman et al. (2021) showed that RF models, though effective in certain scenarios, were prone to higher error rates when handling complex datasets and large feature spaces [9], [26]. RF's strength lies in its ability to handle overfitting better than simpler models, but it can underperform when compared to more advanced techniques like XGBoost.
- Gene Expression Programming (GEP): GEP, as applied by Majidifard et al. (2020), provided interpretable models that elucidate the relationships between input variables and performance metrics, such as rut depth. This transparency made GEP attractive for engineers who require interpretable results [27]. However, GEP required precise parameter tuning, making it less effective for highly complex datasets or situations where rapid model development was needed.
- Autoencoders (AE) and Dimensionality Reduction Techniques: In Liu et al. (2022), the combination of Autoencoders (AE) with Gaussian Process Regression (GPR) showed how dimensionality reduction can improve machine learning models by reducing input noise and computational complexity. AE-GPR outperformed models like SVR and RF by effectively managing high-dimensional data in predicting IRI, proving that reducing input space can lead to improved accuracy and efficiency [26].
- Self-Validated Ensemble Modeling (SVEM): While less commonly used, SVEM provided a balanced approach between accuracy and computational efficiency, especially for smaller datasets. In Mirzaiyanrajeh et al. (2023), SVEM was found to be more practical than ANN in predicting fracture energy for smaller datasets, offering reliable results with fewer computational resources [31]. However, its predictive capacity could be slightly lower than ANN in more complex scenarios.

Table 1. Summary of machine learning applications in predicting asphalt mixture performance: overview of data collection methods, model types, and justifications for model selection across studies

Reference	Method of Data Collection	Database Size	Input Variables	Output Variables	Type of Model	Justification for Model Selection	Availability of Code/Database
Fan et al., 2024[2]	Experimental tests under various conditions	Not specified	Stress state, temperature, loading rate	Asphalt mixture strength	BPNN, SVM	Chosen for multi-field interaction prediction in asphalt strength; SVM showed better accuracy	Data available upon request
Upadhya et al., 2022[6]	Experimental data from lab tests and literature	110 observations	Bitumen content, glass fiber, bitumen grade, fiber length, fiber diameter	Marshall stability	ANN, SVM, GP, RF	SVM_PUK was optimal for accuracy in predicting Marshall stability	Not disclosed
Rondinella et al., 2023[4]	Experimental data with hot and cold asphalt mixtures, using C&DW and RAP	70 observations	RAP, construction and demolition waste content, water, cement, emulsion bitumen, total bitumen, gyratory revolutions	SSDV, ITS	SVM, CatBoost	CatBoost was chosen for high predictive accuracy, with R^2 = 0.99 in ITS predictions	Data available upon request
Khorshidi et al., 2023[23]	Experimental data with varying RAP, CR, WEO, and SS contents	44 mixtures	RAP content, crumb rubber, waste engine oil, steel slag	Rutting, cracking, moisture resistance	FNN, GLM, SVM, GPR	GPR chosen for accuracy in identifying optimal material ratios for performance	Data unavailable due to privacy restrictions
Khorshidi et al., 2023[5]	Experimental data from I-FIT tests with recycled materials in RAP mixtures	Not specified	RAP, crumb rubber, waste engine oil, steel slag aggregates	Cracking performance indices (RDCI, FI, CT-index, FN, TSR)	DNN, Linear, Polynomial regression	DNN selected for best prediction accuracy; MAE, RMSE, and R^2 validation metrics support model choice	Not disclosed

Ali et al., 2021[30]	Tiwari et al., 2023[29]	Tiwari et al., 2022[9]	Rahman et al., 2021[9]	Majidifard et al., 2021[27]	Jian Liu et al., 2022[26]	Jian Liu et al., 2022[25]	Liu et al., 2022[24]	Liu et al., 2023[1]
Experimental data from dynamic modulus tests on asphalt mixtures	Experimental data comparing OPC with silica fume as a filler	Experimental data from lab tests using industrial waste fillers	Data from Texas Department of Transportation (TxDOT) SMGR database	Comprehensive collection of Hamburg test results for asphalt mixtures	LTPP database for IRI prediction with ML models	Data from LTPP program focusing on rut depth for surface AC layer design	Literature and lab data from Superpave mix design	Combined experimental data and NCHRP 9–19 database
1152 observations	24 mixtures	32 samples (expanded to 56 through data	3,139 samples (HWT), 2,805 samples (IDT)	96 tests (288 samples)	194 sections	356 pavement sections	512 entries (cleaned to 441)	7400 samples
Testing conditions, volumetric properties, gradation	Filler content, bitumen content, material type	Filler type, filler content, air voids	Aggregate gradation, absorption, binder content, PG, warm mix additive, RAP/RAS, density, wheel-passes	Mix type, temperature, high-temperature PG grade, AC, NMAS, ABR, RAP, RAS, gradation type, aggregate type, CRC, wheel passes	Climate, subgrade, base, binder, asphalt layer properties	Traffic, climate, pavement material properties, pavement structure	Gradation, binder properties, bulk specific gravity, blend absorption, air void, PG grade, gyrations (Ndes)	Test conditions, binder properties, gradation, volumetric properties
Dynamic modulus	Marshall stability, ITS, water sensitivity	Marshall stability, Marshall quotient, ITS, ITSR	HWT rut depth, IDT strength	Rut depth	International roughness index (IRI)	Rut depth	Effective asphalt content (Pbe), absorbed asphalt content (Pba)	Dynamic modulus (|E*|)
XGBoost	ANN (Shallow Neural Network)	SNN (ANN-based model)	SVR, Bagging, RF, Extra-trees, GB	GEP, ANN	SVR, RF, ANN, GPR, Extra-trees, GB	SVR, RF, ANN, Gradient Boosting	SVR, Ridge, RF, AdaBoost, GB	SVR, KRR, ANN, GPR, GB, XGBoost
XGBoost chosen for high accuracy in nonlinear relationships; robust performance across varying temperatures and frequencies	ANN was chosen for small data modeling with high accuracy, enhanced by leave-one-out cross-validation	SNN chosen for reliability in small datasets, enhanced through k-fold cross-validation and MAKIMA data augmentation	Extra-trees achieved highest accuracy (R^2 = 0.922 for HWT and R^2 = 0.904 for IDT)	GEP was chosen for transparency and interpretability, with superior accuracy over ANN on validation	AE-GPR model chosen for optimal prediction accuracy (R^2 = 0.94) when combined with autoencoders	Gradient Boosting selected for highest performance (R^2 = 0.92) compared to other models	Gradient Boosting chosen for best prediction accuracy (R^2 = 0.948)	XGBoost chosen for highest predictive accuracy (R^2 = 0.987); feature importance shows binder and test conditions as key factors
Not disclosed	Data available upon request	Not disclosed	Data available upon request	Data available upon request	Not disclosed	Not disclosed	Data unavailable due to permission restrictions	Not disclosed

Mirzaiyanrajeh et al., 2023 [31]	Experimental data from DCT tests for fracture energy	852 samples	Binder grade, aggregate type, RAP content, ESALs, VMA, gradation	Fracture energy	ANN, SVEM, AFQM	SVEM chosen for high reliability and efficiency; ANN for non-linear accuracy	Web-based model provided
Liu et al., 2023[32]	Time-series data from RIOHTrack on rutting depth of 19 pavement sections	19 sections (44 months)	Climate, traffic, pavement structure, material properties	Rutting depth	RNN, LSTM, GRU, ARIMAX, GP, K-means	GRU with time series architecture chosen for high accuracy in capturing seasonal patterns	Code available on Google Colab
Al-Sabaeei et al., 2023[33]	Experimental lab data from dynamic shear viscosity tests on bio-modified asphalt	16 samples (various CPO and TPO levels)	CPO%, TPO%, test temperature	Shear viscosity, mixing and compaction temperatures	XGBoost, RFR, RSM	XGBoost chosen for high accuracy in shear viscosity; RFR for optimal mixing and compaction prediction accuracy	Not disclosed

2.2 Overall trends

The studies reviewed consistently demonstrated that ensemble methods such as Gradient Boosting and XGBoost were the most effective in terms of both accuracy and robustness. These models were particularly useful in handling large datasets and complex, nonlinear relationships within asphalt mixture data. However, their high computational costs and complexity in hyperparameter tuning limited their practicality in some real-world scenarios. On the other hand, simpler models like SVM and ANN, while still effective in certain cases, struggled with overfitting and computational demands when faced with large, high-dimensional datasets. GPR also proved highly effective, particularly for small to medium datasets, though it required substantial computational resources and careful tuning. Dimensionality reduction techniques such as Autoencoders (AE) and Principal Component Analysis (PCA) helped mitigate these issues by streamlining input features, improving the efficiency and accuracy of models like GPR and SVR.

Finally, interpretability remains a key consideration, with methods like Gene Expression Programming (GEP) offering more transparent models than black-box approaches like ANN and XGBoost. This interpretability can be critical for engineers looking to understand the underlying relationships between variables and performance outcomes.

3 Concrete property prediction and structural performance

Machine learning (ML) models have become an essential tool in predicting concrete properties and optimizing structural performance, addressing the limitations of traditional empirical methods. This section explores various machine learning techniques applied to predict key concrete properties such as compressive strength, tensile strength, modulus of elasticity, and fracture energy. These studies demonstrate the advantages and disadvantages of different ML approaches in terms of prediction accuracy, computational complexity, and model interpretability.

Song et al. (2022) applied machine learning models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees (DT), Random Forest (RF), and Gradient Boosted Regression Trees (GBRT) to optimize cementitious material mixtures. Input variables included water content, cement content, supplementary cementitious materials (SCMs), and aggregate content, while the outputs were uniaxial compressive strength (UCS) and durability. ANN excelled in capturing nonlinear relationships but required significant computational resources and careful tuning to avoid local minima. SVM performed well in generalization but was highly sensitive to hyperparameter tuning, and RF improved accuracy by reducing variance, though it came with higher computational costs. GBRT offered the highest accuracy in UCS prediction but increased computational complexity. Metaheuristic algorithms such as Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) were used to optimize the model parameters and enhance the performance of the ML models [34].

Hafez et al. (2022) developed a machine learning regression model, Pre-bcc, to predict slump, compressive strength, carbonation, and chloride ingress resistance for blended cement concrete (BCC) using supplementary cementitious materials (SCMs) such as fly ash, ground granulated blast-furnace slag, silica fume, lime powder, and calcined clay. Input variables included SCM types and proportions. ANN, RF, and SVM models were tested, with RF showing better accuracy and interpretability, though computationally intense. SVM required careful tuning but handled generalization well. Pre-bcc offers high prediction accuracy for slump and strength but is computationally complex when handling multiple SCMs, improving the understanding of SCM effects in BCC [16].

Hafez et al. (2023) then introduced Opt-bcc, an optimization tool using Genetic Algorithms (GA) with Pre-bcc to optimize sustainability scores of blended cement concrete mixes. Input variables included various SCM types and proportions, while output variables were strength, slump, and durability indices. GA effectively minimized environmental and cost impacts but required complex tuning. Opt-bcc achieved significant cost and environmental reductions compared to existing models, though functional parameter prediction models were nonlinear, demanding higher computational resources. This study highlighted GA's potential in eco-friendly concrete optimization while balancing functional and economic criteria [35].

Pfeiffer et al. (2024) utilized an amortized Gaussian Process (GP) model integrated with an inverse optimization framework to design concrete mixes minimizing climate impact and cost. Input variables were SCM proportions, water/cementitious material ratio, and aggregate composition, while the output variable was compressive

strength at 28 days. The GP model provided mean predictions and uncertainty estimates, making it more robust than traditional models like ANN and RF, which lack uncertainty measures. The GP's flexibility for industrial-scale datasets added accuracy, but computational demands were significant. This study demonstrated GP's effectiveness for mix design, balancing environmental and economic objectives with structural performance requirements [36].

Moein et al. (2023) reviewed several machine learning and deep learning models for predicting concrete properties, including SVM, ANN, Random Forest, and Extreme Learning Machines (ELM). The input variables included cement content, aggregate composition, water-cement ratio, and curing age. ANN showed high accuracy but was prone to overfitting without proper tuning, while SVM was more effective for smaller datasets but struggled with high-dimensional data. ELM provided faster training times compared to ANN but at the cost of prediction accuracy. Genetic Algorithms (GA) were used to enhance model optimization when combined with other ML models. Random Forest and ANN were identified as the most reliable models for concrete property prediction, with RF offering better interpretability and ANN excelling in predictive performance [11].

Yu et al. (2018) compared an Enhanced Cat Swarm Optimization (ECSO)-optimized SVM model with traditional models like ANN and Extreme Learning Machines (ELM) for predicting the compressive strength of high-performance concrete (HPC). Input variables included water content, cement content, and supplementary materials. The ECSO-optimized SVM model achieved superior accuracy (R^2 = 0.9526), outperforming ANN (R^2 = 0.8716). While SVM required significant parameter tuning, ECSO improved the convergence rate and avoided local minima, making it more efficient. ANN, though effective, suffered from overfitting and was computationally expensive [37].

Pham et al. (2016) used a Least Squares Support Vector Regression (LS-SVR) model optimized by the Firefly Algorithm (FA) to predict the compressive strength of high-performance concrete (HPC). Input variables included cement, aggregates, and curing conditions. FA-LS-SVR achieved the highest accuracy (R^2 = 0.89) compared to ANN and traditional SVM models. The optimized SVM model outperformed ANN by providing better generalization and reducing prediction errors. However, the model required careful tuning of parameters like penalty factors, making it computationally demanding [38].

Yaseen et al. (2018) used Extreme Learning Machines (ELM) to predict the compressive strength of lightweight foamed concrete, outperforming other models like Multivariate Adaptive Regression Splines (MARS) and M5 Tree. Input variables included cement content, oven dry density, and foam volume. ELM achieved an R^2 of 0.875, making it the fastest model in terms of training speed, though it was less accurate for highly complex data. MARS and M5 Tree provided reasonable accuracy but failed to capture complex relationships, while ELM's fast training and simplicity made it an efficient option for lightweight concrete strength prediction [39].

Omran et al. (2016) compared Gaussian Process Regression (GPR), Multilayer Perceptron (MLP), and Support Vector Machines (SVM) for predicting the compressive strength of environmentally friendly concrete. GPR outperformed the other models, achieving the highest accuracy (R^2 = 0.9842) and offering better generalization through its probabilistic approach. However, GPR was computationally intensive. Ensemble methods like Additive Regression and Bagging with GPR also provided high accuracy, while SVM and MLP required extensive parameter tuning to avoid overfitting. GPR was highlighted for its balance between accuracy and computational efficiency, making it a strong choice for concrete strength prediction [40].

Bonifácio et al. (2019) applied Support Vector Regression (SVR) and the Finite Element Method (FEM) to predict the compressive strength and Young's modulus of lightweight aggregate concrete (LWAC). SVR outperformed FEM slightly, achieving a lower deviation from experimental results (5.46% for compressive strength), with the key advantage being SVR's reusability with new data and speed. FEM, although slightly less accurate, required fewer inputs and was advantageous in cases where experimental results were scarce. SVR required a larger training dataset, making it more computationally intensive [10].

Tanyildizi (2018) applied ANN and SVM to predict the strength properties of carbon fiber-reinforced lightweight concrete exposed to high temperatures. Input variables included silica fume, carbon fiber content, and temperature. ANN achieved the highest accuracy (R^2 = 0.9902 for compressive strength), outperforming SVM (R^2 = 0.9701). While ANN offered superior predictive accuracy, it required more computational resources and careful optimization of hidden neurons and learning algorithms. SVM was simpler to use but less accurate, making it a better choice for smaller datasets [15].

Mozumder et al. (2017) used Support Vector Regression (SVR) to predict the uniaxial compressive strength of fiber-reinforced polymer (FRP) confined concrete, achieving higher accuracy (R^2 = 0.9832 for CFRP) than ANN models and empirical methods. SVR's ability to avoid local minima and provide better generalization made it a more reliable method, though it required substantial computational effort and parameter tuning compared to ANN, which suffered from slower convergence and higher prediction errors [8].

Keshtegar et al. (2019) applied a hybrid RSM-SVR model to predict the shear strength of steel fiber-reinforced concrete beams (SFRCBs). The hybrid model outperformed ANN and other traditional methods, achieving an R^2 of 0.9508, thanks to its ability to capture nonlinear relationships and cross-correlations between input variables. Although the hybrid model required significant computational power, it proved to be the most accurate for predicting SFRCBs shear strength, demonstrating the advantage of combining multiple modeling approaches [7].

Aiyer et al. (2014) compared Least Square Support Vector Machines (LSSVM) and Relevance Vector Machines (RVM) for predicting the compressive strength of self-compacting concrete. RVM outperformed LSSVM and ANN, offering additional benefits such as handling variance and uncertainty. While LSSVM was accurate, RVM's ability to calculate variance made it a better tool for assessing uncertainty in predictions, especially in civil engineering applications [41].

Yuvaraj et al. (2013) applied SVR to predict fracture characteristics, such as fracture energy and failure load, of high-strength and ultra-high-strength concrete beams. The SVR model achieved high prediction accuracy (R^2 close to 1 for all parameters), outperforming traditional empirical models. The SVR model's strength lay in its ability to handle nonlinear relationships even with limited datasets, though it required careful parameter tuning to optimize its predictive performance [42].

Yan & Shi (2010) used SVM to predict the elastic modulus of normal and high-strength concrete,

outperforming traditional empirical models. SVM achieved better accuracy with fewer parameters compared to empirical models, though it required careful tuning of kernel parameters. ANN, while effective, was more complex to tune and prone to local minima, making SVM the preferred model for this application [43].

Nazari & Sanjayan (2015) optimized SVM using Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and other metaheuristic algorithms to predict the compressive strength of geopolymer concrete. The hybrid models, particularly the ICOA-SVM model, achieved superior prediction accuracy (R^2 = 0.8993), though they were computationally intensive due to the optimization process [44].

Deng et al. (2018) used Convolutional Neural Networks (CNN) to predict the compressive strength of recycled aggregate concrete (RAC), outperforming both Backpropagation Neural Networks (BPNN) and SVM in terms of accuracy and efficiency. CNN's advantage was its ability to automatically extract deep features from input data without requiring manual preprocessing, though it was more computationally intensive [45].

Kaloop et al. (2019) compared LSSVM, ANN, and regression models to predict the resilient modulus (Mr) of recycled concrete aggregate blends. LSSVM achieved the highest accuracy (R^2 = 0.982), outperforming both ANN and regression models, particularly with smaller datasets, though it required careful tuning of regularization parameters [46].

Cheng et al. (2014) applied the Genetic Weighted Pyramid Operation Tree (GWPOT) to predict the compressive strength of high-performance concrete, outperforming ANN and SVM models. GWPOT provided interpretable mathematical formulas, offering better transparency, though it required higher computational resources for optimization [47].

Zhang et al. (2019) used Random Forest (RF) optimized with Beetle Antennae Search (BAS) to predict the uniaxial compressive strength of lightweight self-compacting concrete. BAS improved the hyperparameter tuning process, resulting in an R^2 value of 0.97, significantly outperforming traditional regression models. However, the computational complexity was higher due to the optimization process [48].

3.1 Summary of methods

Table 2 provides a detailed summary of studies focusing on machine learning approaches and their applications in predicting concrete properties. In the reviewed studies, machine learning techniques were applied to predict concrete properties such as compressive strength, tensile strength, modulus of elasticity, and fracture energy, offering improvements in accuracy and efficiency over traditional empirical methods. The following methods were highlighted for their strengths and weaknesses:

- Artificial Neural Networks (ANNs): ANNs were frequently applied in predicting nonlinear relationships in concrete properties, such as compressive strength and fracture energy. Studies like Tanyildizi (2018) and Song et al. (2022) demonstrated that ANNs performed well in capturing complex data patterns [15], [34]. However, ANNs often faced challenges such as overfitting and the need for large datasets, which made them computationally expensive. Yu et al. (2018) and Mozumder et al. (2017) further emphasized that proper tuning of hyperparameters, such as the number of hidden neurons and learning rates, is crucial to achieving high accuracy without overfitting [8], [37].
- Support Vector Machines (SVMs): SVMs were consistently highlighted as strong performers, especially when dealing with smaller datasets, as shown in Yu et al. (2018), Mozumder et al. (2017), and Yan & Shi (2010) [8], [37], [43]. SVM models excelled at predicting compressive strength, fracture characteristics, and elastic modulus, particularly when optimized using techniques such as Enhanced Cat Swarm Optimization (ECSO) and the Firefly Algorithm (FA) [37], [38]. These optimizations significantly improved convergence and accuracy. However, SVMs can be computationally intensive and sensitive to hyperparameter tuning, requiring careful selection of kernel functions.
- Random Forest (RF): Random Forest models, applied in studies such as Song et al. (2022) and Zhang et al. (2019), were particularly effective in handling complex, high-dimensional datasets [34], [48]. RF's ability to reduce overfitting by averaging multiple decision trees made it a popular choice for predicting properties like compressive strength. Despite its robustness, RF models are computationally demanding and require tuning of hyperparameters such as the number of trees and depth to achieve optimal results.
- Gradient Boosting and Boosted Regression Trees (GBRT): Gradient Boosting models were often the most accurate in predicting concrete properties, particularly in Song et al. (2022) where they excelled at predicting uniaxial compressive strength (UCS) [34]. These models effectively captured nonlinear relationships between variables but came at a high computational cost due to their iterative learning process. Gradient Boosting methods like XGBoost are powerful but require significant tuning to prevent overfitting, especially when dealing with large datasets.
- Extreme Learning Machines (ELM): Yaseen et al. (2018) demonstrated that ELM models provided a fast and computationally efficient method for predicting concrete properties, particularly lightweight foamed concrete [39]. ELM's ability to train quickly made it useful for simpler datasets, but it lacked the accuracy of more complex models like RF and Gradient Boosting when dealing with high-dimensional or intricate data.
- Gaussian Process Regression (GPR): Omran et al. (2016) highlighted that GPR was highly accurate in predicting concrete compressive strength [40]. GPR's probabilistic approach offered the added benefit of estimating uncertainty, which made it suitable for cases where confidence in the predictions was critical. However, GPR's computational demands increase significantly with larger datasets, limiting its practicality for large-scale applications.
- Least Squares Support Vector Machines (LSSVM): Enhanced versions of SVM, such as LSSVM, were applied in Pham et al. (2016) and Kaloop et al. (2019) to improve predictive performance and computational efficiency [38], [46]. LSSVM, optimized by metaheuristic algorithms like the Firefly Algorithm (FA), outperformed standard SVM and ANN models, especially in smaller datasets. However, LSSVM still required careful tuning of parameters like the regularization factor to achieve high accuracy.
- Convolutional Neural Networks (CNNs): In Deng et al. (2018), CNNs were shown to outperform traditional models like SVM and Backpropagation Neural Networks (BPNN) when predicting compressive strength in recycled aggregate concrete [45]. CNNs excelled at automatically extracting deep features from raw data, which improved accuracy and reduced the need for manual feature engineering. However,

CNNs are computationally intensive and require large datasets to fully leverage their potential.

- Hybrid Models (e.g., RSM-SVR, ANN-MOGWO): Hybrid models combining machine learning algorithms with optimization techniques, such as Keshtegar et al. (2019)'s RSM-SVR model, showed superior performance in predicting complex properties like shear strength [7]. These models combine the strengths of multiple techniques, improving accuracy by capturing nonlinearities and complex relationships between variables. However, hybrid models are computationally expensive due to the complexity of integrating multiple approaches.

Table 2. Summary of Machine Learning Applications in Predicting Concrete Properties: Overview of Data Collection Methods, Model Types, and Justifications for Model Selection Across Studies

Reference	Method of Data Collection	Database Size	Input Variables	Output Variables	Type of Model	Justification for Model Selection	Availability of Code/Database
Hafez et al., 2022[16]	Experimental data from >150 sources	1650+ data points	SCM types, replacement percentages, mix design parameters (e.g., slump, strength, carbonation resistance)	Slump, strength, carbonation resistance, chloride ingress	Multi-layer regression	Captures wide variety of SCMs and functional properties with high statistical accuracy (R = 0.94-0.97)	Available online via Pre-bcc tool
Pfeiffer et al., 2024[36]	Industrial dataset (job site data from 2017-2020)	9296 mixes	Water/cement ratio, SCM quantities, 12 constituent features	Compressive strength	Amortized Gaussian Process	Captures strength evolution with uncertainty estimates and supports optimization for cost and climate impact	Not disclosed
Hafez et al., 2023[35]	Regression model predictions via Pre-bcc	Derived from Pre-bcc model	SCM types, mix proportions, sustainability indicators	Strength, slump, chloride resistivity, carbonation	Genetic Algorithm (Opt-bcc tool)	Optimizes sustainability score (economic, environmental, and functional properties) with superior results	Available online via Opt-bcc tool
Yu et al., 2018[37]	Data collected from 183 studies (1998–2015)	1761 groups of data	Water, cement, blast furnace slag, fly ash, superplasticizer, coarse and fine aggregates, curing age	Compressive strength	Enhanced Cat Swarm Optimization-SVM	Superior optimization of SVM parameters for high accuracy and efficiency	Not disclosed
Pham et al., 2015[38]	Experimental data from the Nga Ba Hue infrastructure project	239 tests	Cement, sand, small and medium coarse aggregate, water, superplasticizer, curing time	Compressive strength	Firefly Algorithm-Optimized LS-SVR	High accuracy and generalization for HPC strength prediction; RMSE validation	Not disclosed
Yaseen et al., 2018[39]	Experimental database retrieved from literature	91 data points	Cement content, oven dry density, water/binder ratio, foam volume	Compressive strength	Extreme Learning Machine (ELM), MARS, M5 Tree, SVR	ELM chosen for superior accuracy and reliability over other methods in predicting compressive strength of foamed concrete	Not disclosed
Omran et al., 2016[40]	Experimental data for compressive strength prediction	144 data points	Cement type, curing age, water, cementitious material, fly ash, sand, pea gravel, Haydite lightweight aggregate, Micro Air	Compressive strength	Gaussian Processes Regression, Additive Regression, Bagging, M5P, REPTree, Multilayer Perceptron	GPR achieved the highest prediction accuracy (R^2 = 0.9842); additive regression and bagging enhanced predictive performance	Not disclosed
Bonifácio et al., 2019[10]	Experimental data from Lightweight Aggregate Concrete (LWAC)	180 mixtures	Water/cement ratio, quantity of cement, aggregate volume, aggregate density	Compressive strength, Young's modulus	Support Vector Regression (SVR), Finite Element Method (FEM)	SVR achieved lower average error compared to FEM; both methods validated against experimental results	Not disclosed

Zhang et al., 2019[48]	Cheng et al., 2014[47]	Kaloop et al., 2019[46]	Deng et al., 2018[45]	Nazari and Sanjayan, 2015[44]	Yan and Shi, 2010[43]	Yuvaraj et al., 2013[42]	Aiyer et al., 2014[41]	Keshtegar et al., 2019[7]	Mozumder et al., 2017[8]	Tanyildizi, 2018[15]
Laboratory-based dataset for lightweight self-compacting concrete	Experimental data from UCI repository	Experimental data from recycled aggregate/ concrete blends	Experimental data on recycled concrete	Experimental data from literature	Experimental data for NSC and HSC	Experimental data from fracture tests on HSC and UHSC	Experimental data from SCC with database by Siddique et al. (2011)	Experimental data for steel fiber-reinforced concrete beams	Experimental data from various FRP-wrapped concrete tests	Experimental data for lightweight concrete
131 samples	1030 samples	128 datasets	74 datasets	1347 datasets	89 cases (HSC), 70 cases (NSC)	87 datasets	80 samples	139 samples	238 samples	144 data points
w/b ratio, PP and steel fiber content, scoria and CR content, NFA and NCA content, temperature	Cement, fly ash, slag, water, superplasticizer, coarse and fine aggregates, age	RCM ratio, stress states, bulk stress, shear stress	Water-cement ratio, recycled coarse/fine aggregate replacement ratio, fly ash replacement ratio	Fly ash, slag, coarse and fine aggregate, water, superplasticizer, NaOH, KOH, curing conditions	Compressive strength (fc)	Beam geometry, water/cement ratio, compressive strength, tensile strength, modulus of elasticity	Cement, fly ash, water/powder ratio, superplasticizer, sand, coarse aggregate	Compressive strength, longitudinal steel strength, shear span-to-depth ratio, steel fiber properties	Cylinder diameter, height/diameter ratio, compressive strength, FRP thickness, tensile strength	Cement, silica fume, carbon fiber, aggregate content, temperature
Uniaxial compressive strength	Compressive strength	Resilient modulus	Compressive strength	Compressive strength	Elastic modulus (Ec)	Fracture energy, critical stress intensity factor, critical crack tip opening isplacement, failure load	Compressive strength	Shear capacity of SFRC beams	Compressive strength of FRP-confined concrete	Compressive strength, flexural strength
Beetle Antennae Search-Random Forest (BAS-RF)	Genetic Weighted Pyramid Operation Tree (GWPOT)	LSSVM, ANN	CNN-based deep learning model	SVM optimized by GA, PSOA, ACOA, ABCOA, ICOA	Support Vector Machine (SVM)	SVR	LSSVM, RVM, ANN	Hybrid RSM-SVR	SVR (RBF, Polynomial, Exponential RBF), ANN	ANN, SVM
High accuracy (R^2 = 0.97); BAS tuned RF hyperparameters; identified temperature and w/b ratio as key variables	Outperformed ANN, SVM, and ESIM; generated explicit formulas for practical applications	LSSVM outperformed ANN and regression models in accuracy and computational efficiency	Achieved higher precision, efficiency, and generalization ability compared to traditional models	Hybrid models achieved superior accuracy in predicting compressive strength of geopolymer concrete	SVM outperformed ANN, fuzzy logic, and regression methods in RMSE and MAPE for HSC and NSC prediction	SVR achieved excellent predictive performance (R^2 > 0.99) for all fracture parameters and failure load	RVM achieved the best accuracy with variance prediction, outperforming LSSVM and ANN on RMSE and MAE criteria	RSM-SVR demonstrated superior accuracy over standalone RSM, SVR, ANN, and empirical formulations in predicting shear capacity	SVR (RBF) demonstrated the highest prediction accuracy (R = 0.9908, RMSE = 6.03, MAPE = 2.61) compared to ANN and empirical models	ANN showed the best predictive accuracy (R^2 = 0.99 for compressive, 0.968 for flexural), followed by SVM (R^2 = 0.9701)
Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed	Not disclosed

3.2 Overall trends

The studies consistently demonstrated that ensemble methods like Random Forest and Gradient Boosting delivered the best performance in predicting concrete properties, excelling in handling high-dimensional datasets and capturing complex nonlinear relationships. However, these methods were computationally demanding and required careful tuning.

Support Vector Machines (SVMs), particularly when enhanced with optimization algorithms, were effective for smaller datasets but required significant computational resources and careful parameter tuning. Artificial Neural Networks (ANNs) were highly accurate in capturing complex relationships but often suffered from overfitting and required large datasets and computational resources.

Hybrid models, such as RSM-SVR and ANN-MOGWO, offered the highest accuracy by combining the strengths of multiple approaches, but their complexity and computational requirements made them more suitable for research applications.

Convolutional Neural Networks (CNNs) and Gaussian Process Regression (GPR) were also strong performers, with CNNs excelling at feature extraction and GPR providing uncertainty estimates. However, both models required substantial computational power.

In summary, ensemble methods and hybrid models proved to be the most accurate, while SVMs and ANNs were useful but required extensive tuning. The model choice ultimately depended on the dataset size, complexity, and available computational resources.

4 Classification and pattern recognition in pavement systems

Machine learning models, particularly unsupervised learning methods, have been instrumental in classifying pavement conditions, detecting cracks, and identifying patterns that are crucial for effective pavement management. Below is a detailed analysis of several studies using unsupervised models to address pavement-related problems.

Shao et al. (2022) applied K-means clustering to classify pavement performance patterns based on long-term Pavement Condition Index (PCI) and Riding Quality Index (RQI) data. Their model classified pavement performance into five distinct patterns, facilitating the evaluation of road maintenance strategies. The key strengths of K-means were its simplicity and ability to process large datasets effectively. However, its requirement to predefine the number of clusters and its sensitivity to data with varying densities or missing values were major limitations, suggesting the need for more adaptive models to achieve greater accuracy [49].

Mathavan et al. (2014) used a Self-Organizing Map (SOM), an unsupervised neural network, to classify doweled concrete pavement joints based on Falling Weight Deflectometer (FWD) data. Input parameters included load transfer efficiency (LTE), void intercepts (VI), and absolute deflection (D). SOM classified the joints into three categories: good, marginal, and poor. The model achieved an accuracy of 65-70%, improving to 87.5% when only LTE and D were used, demonstrating the potential to reduce human inconsistencies in manual assessments. The pros of SOM include its ability to capture complex patterns and automate classification, though its slow training process and sensitivity to unbalanced data were notable drawbacks [50].

Mubashshira et al. (2020) employed K-means clustering to detect road surface cracks by segmenting 2D road surface images. After preprocessing to reduce noise, K-means was used for image segmentation, followed by Otsu thresholding and morphological operations to refine the detected cracks. The model achieved an average detection accuracy of 97.75%, outperforming traditional edge detection methods by reducing false negatives. While K-means clustering was effective in handling noise and irregularities, its efficiency was limited by the reliance on post-processing steps, particularly for large-scale real-time applications [51].

Li et al. (2021) proposed a novel model that fused Convolutional Neural Networks (CNN) with K-means clustering for road crack classification. The input data consisted of crack images collected via automated vehicles and smartphones, and the model classified crack types—transverse, longitudinal, and alligator cracks—with accuracies of 80.6%, 79.2%, and 91.3%, respectively. The fusion of CNN and K-means allowed for iterative refinement of clustering assignments, reducing the need for manually labeled data. However, the model faced challenges due to its high computational cost during training and the need for extensive optimization [52].

Golmohammadi et al. (2024) combined PCA and DBSCAN for anomaly detection in pavement health monitoring using Fiber Bragg Grating (FBG) sensors. The system processed strain and temperature data to detect structural anomalies in pavement layers. DBSCAN effectively distinguished between normal and abnormal patterns without labeled data, demonstrating high accuracy in anomaly detection. The system, however, was sensitive to sensor placement and required considerable computational resources for continuous monitoring [53].

Abdelmawla et al. (2021) utilized PCA and K-means clustering to classify pavement cracks from 1,125 road surface images. The input images were preprocessed with edge detection and morphological operations, followed by dimensionality reduction using PCA and clustering using K-means. The study identified three clusters: multi-directional cracks, longitudinal cracks, and images without cracks. PCA reduced dimensionality, improving computational efficiency, but struggled with nonlinear relationships in the data. K-means effectively classified the cracks, although it was sensitive to initial cluster assignments [54].

Dong et al. (2021) classified climatic regions for pavement systems using PCA and K-means clustering. Input data from the Long-Term Pavement Performance (LTPP) database included 16 climate variables. Four primary clusters—wet no freeze, dry no freeze, dry freeze, and snow freeze—were identified, and results from Artificial Neural Networks (ANN) and Fisher's linear discriminant analysis were compared. ANN achieved higher prediction accuracy than discriminant analysis, though ANN required more tuning and computational power. K-means clustering proved efficient for handling large datasets, though the need to predefine clusters remained a limitation [55].

Shi et al. (2024) applied K-means clustering to analyze Acoustic Emission (AE) data from epoxy asphalt mixtures with varying crumb rubber (CR) content. Four damage modes were identified: cohesive cracking, aggregate-asphalt interface cracking, aggregate fracture, and aggregate friction. The model effectively classified AE signals and demonstrated that 4% CR content was optimal for balancing toughness and strength. However, the model's sensitivity to initial cluster selection and overlapping clusters presented challenges, especially when dealing with highly correlated data [56].

Akhtar et al. (2020) implemented a parallel K-means clustering model to assess adhesion failure in Warm Mix Asphalt (WMA) through high-resolution image processing. The model reduced execution time by 30-46% compared to sequential K-means clustering, improving the detection of adhesion failure. Although the parallel model handled large image datasets more efficiently, its complexity and sensitivity to initial cluster centers posed challenges during the implementation [57].

Sahari Moghadam et al. used K-means clustering in conjunction with k-nearest neighbors (KNN) and support vector machines (SVM) to classify asphalt coating conditions in loose mixtures. The input images from static immersion tests were segmented using K-means, followed by classification using KNN and SVM. The model improved classification accuracy by reducing human bias, though it was sensitive to image quality and relied on robust preprocessing steps to ensure accuracy [58].

4.1 Summary of methods

Table 3 provides a detailed summary of studies focusing on machine learning approaches and their applications in classification and pattern recognition for pavement systems. Unsupervised learning techniques, particularly clustering and pattern recognition methods, have been extensively applied in pavement system classification and anomaly detection. Here's an overview of the key methods:

- K-means Clustering: Widely used for classifying pavement performance patterns, crack detection, and assessing asphalt coating conditions (Shao et al., 2022; Akhtar et al., 2020). Its simplicity and efficiency in handling large datasets are advantageous, but it often struggles with predefined cluster requirements and sensitivity to initial conditions [49], [57].
- Principal Component Analysis (PCA): Commonly paired with K-means for dimensionality reduction, PCA improved computational efficiency in multi-dimensional data applications like crack classification and climate region analysis (Abdelmawla et al., 2021; Dong et al., 2021) [54], [55]. However, PCA's linearity limits its ability to capture complex relationships.
- Self-Organizing Maps (SOM): Used for classifying pavement joint conditions (Mathavan et al., 2014). SOM excels at handling high-dimensional data but has a slow training process and reduced accuracy with unbalanced datasets [50].
- Convolutional Neural Networks (CNN): When combined with K-means, CNN was effective for road crack classification, automating feature extraction and improving classification performance (Li et al., 2021). The downside is its high computational cost and complexity in optimization [52].
- DBSCAN: Applied for anomaly detection in pavement monitoring systems (Golmohammadi et al., 2024), DBSCAN is effective in handling unlabelled data, though sensitive to sensor placement and computationally intensive in continuous monitoring [53].
- Parallel K-means Clustering: Enhanced efficiency in large-scale image processing tasks by reducing execution time (Akhtar et al., 2020). However, its implementation is complex, particularly when managing communication between computing nodes [57].
- K-means with Supervised Classifiers (KNN, SVM): Combining K-means with KNN or SVM improved accuracy in tasks like asphalt coating assessment but required high-quality image data for optimal performance (Sahari Moghadam et al.) [58].

Table 3. Summary of machine learning applications in classification and pattern recognition in pavement systems: overview of data collection methods, model types, and justifications for model selection across studies

Reference	Method of Data Collection	Database Size	Input Variables	Output Variables	Type of Model	Justification for Model Selection	Availability of Code/Database
Shao et al., 2022[49]	Long-term historical data from expressways	762 road sections (439 valid)	PCI and RQI time series	Pavement performance evolution patterns	K-means clustering (unsupervised learning)	Chosen for its ability to effectively categorize long-term evolution patterns; identified five distinct performance patterns	Not disclosed
Mathavan et al., 2014[50]	Falling Weight Deflectometer (FWD) data	1,409 deflection bowls	Absolute deflection, Load transfer efficiency (LTE), Void intercepts	Concrete pavement joint condition (Good, Marginal, Poor)	Self-Organizing Map (SOM)	SOM achieved 87.5% classification accuracy for joint conditions; captures engineering expertise for non-subjective decisions	Not disclosed
Mubashshira et al., 2020[51]	Road surface images from multiple sources	200 images (120 for road detection)	RGB values, cracks morphology	Crack detection accuracy (Precision, Recall, IoU)	K-means clustering with Otsu binarization	Efficiently segmented crack regions with high accuracy; achieved 97.75% overall accuracy; outperformed edge detection methods	Not disclosed

Sahari Moghaddam et al., 2019[58]	Akhtar et al., 2020[57]	Shi et al., 2024[56]	Dong et al., 2021[55]	Abdelmawla et al., 2021[54]	Golmohammadi et al., 2024[53]	Li et al., 2021[52]
Image processing of loose asphalt mixtures	Image processing of WMA and HMA samples	Acoustic Emission (AE) data from UC and IDT tests	Historical climatic data from LTPP program	Pavement surface images	FBG sensor data collected from a test track	Automated vehicle and smartphone crack images
159 samples	High-resolution images (RGB, HSV)	4 rubberized epoxy asphalt mixtures with varying CR contents	21,666 annual datasets from 800 weather stations	Not specified	4 months of data from 12 FBG sensors	2,995 images (various types)
Image pixel intensities, specular highlights, RGB histograms of coated/uncoated areas	Pixel intensity values, adhesion failure, moisture sensitivity test results	Nine AE parameters, including rise time, ringing count, energy count, peak frequency, and amplitude	16 climatic variables: temperature, precipitation, humidity, snowfall, freezing conditions	Processed crack patterns via image processing and PCA	Strain and temperature readings processed using PCA	Crack image features extracted by AlexNet CNN
Retained asphalt coating (%)	Adhesion failure quantification	Damage evolution patterns, damage modes, toughness	Climatic region classification	Clusters of crack patterns	Anomaly detection in strain data	Crack classification (Transverse, Longitudinal, Alligator)
K-means clustering with SVM and KNN classifiers	Parallel K-means clustering (PKIP algorithm)	K-means clustering (unsupervised learning)	PCA, Factor Analysis, K-means Clustering, Fisher's Discriminant Analysis, ANN	K-means clustering (unsupervised learning)	DBSCAN clustering (unsupervised learning)	AlexNet CNN with K-means clustering
Automated stripping assessment system; demonstrated a mean difference of 4.91% compared to manual technician assessment	PKIP algorithm optimized for execution time (30–46% faster than sequential K-means); effectively quantified adhesion failure in different conditions	Efficiently classified AE signals into damage patterns, highlighting CR's effects on toughness and cracking mechanisms; identified four distinct damage modes	PCA and Factor Analysis identified key components (temperature and humidity) for dimensionality reduction; ANN achieved high accuracy in climatic region classification	Applied classic image processing and PCA to delineate and cluster crack patterns; revealed meaningful correlations between crack patterns and derived clusters	Efficiently detected anomalies using PCA feature fusion and clustering, aiding in proactive maintenance strategies	Fused CNN and K-means achieved unsupervised classification; superior accuracy (AP = 0.806, 0.792, and 0.913) for crack types
Not disclosed	Not disclosed	Not disclosed	TPP dataset available: InfoPave	Not disclosed	Not disclosed	Dataset available on GitHub: CrackDataset_DL_HY

4.2 Overall trends

K-means clustering emerged as a dominant method for its simplicity and effectiveness in pavement performance analysis, but it often required careful parameter tuning and was limited by predefined clusters. PCA was valuable for dimensionality reduction but struggled with nonlinearity. SOM and DBSCAN offered robust classification and anomaly detection in high-dimensional and sensor data, respectively, though at the cost of computational efficiency.

The combination of CNN with K-means showed promise in handling complex, image-based data, albeit with high computational demands. Parallel implementations, like Parallel K-means, improved processing times but introduced complexity in deployment. Hybrid models combining unsupervised and supervised techniques, such as K-means with KNN and SVM, offered higher accuracy but depended on data quality and preprocessing efforts.

Table 4 below provides a concise summary of the machine learning methods applied across various studies in asphalt mixture performance, concrete property prediction, and pavement classification. It outlines the key advantages and disadvantages of each method, complementing the trends discussed in the preceding sections.

Table 4. Summary of machine learning methods, advantages, and disadvantages across pavement and concrete applications

Method	Category			Advantages	Disadvantages
	A	C	P		
Artificial Neural Networks (ANNs)	×	×		Captures complex nonlinear relationships High accuracy in predicting performance metrics	Computationally intensive Prone to overfitting Requires large datasets
Support Vector Machines (SVMs)	×	×		Handles nonlinear relationships well Effective with smaller datasets	Sensitive to parameter tuning Computationally expensive for large datasets
Gradient Boosting/XGBoost	×	×		High accuracy and handles complex nonlinear relationships Regularization avoids overfitting	Requires significant computational power Complex tuning required
Random Forest (RF)	×	×		Good interpretability and robust handling of overfitting	Less accurate compared to ensemble methods Computationally expensive
Gaussian Process Regression (GPR)	×	×		Provides accurate predictions and uncertainty estimates. Handles complex nonlinear relationships well	Computationally demanding. Requires careful tuning of hyperparameters
Gene Expression Programming (GEP)	×			Offers transparent, interpretable models	Requires precise tuning Limited for large and complex datasets
Hybrid Models (e.g., SVM with optimization)	×	×		Optimizes parameter selection Improves prediction accuracy and model convergence	Computationally intensive Requires additional resources for optimization algorithms
Self-Organizing Maps (SOM)			×	Captures complex patterns in high-dimensional data Reduces inconsistencies in manual classification	Slow training process Struggles with unbalanced datasets
DBSCAN			×	Effective in anomaly detection with unlabelled data Handles noisy data	Sensitive to sensor placement Computationally intensive for continuous monitoring
Convolutional Neural Networks (CNN)			×	Automatically extracts deep features from images Highly effective for image-based crack detection and classification	High computational cost Requires extensive optimization and tuning of hyperparameters
Parallel K-means			×	Reduces execution time by 30-46% Efficient for large-scale image processing	Complex implementation Communication overhead between nodes
K-means Clustering			×	Simple and efficient for large datasets Effective in crack detection and pavement classification	Requires predefined cluster number Sensitive to initial conditions and varying densities
Principal Component Analysis (PCA)			×	Reduces dimensionality, improving computational efficiency	Limited by linearity, struggles with complex relationships

Note: A is Asphalt Mixture Performance and Optimization, C is Concrete Property Prediction and Structural Performance, P is Classification and Pattern Recognition in Pavement Systems

5 Conclusion

This literature review examines the application of machine learning techniques in material testing across three key areas: asphalt mixture performance, concrete property prediction, and classification and pattern recognition in pavement systems. The studies demonstrate the significant potential of machine learning to improve prediction accuracy, optimize material design, and reduce reliance on costly experimental testing.

Across all categories, models such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RF), Gradient Boosting (GB), Gaussian Process Regression (GPR), and Convolutional Neural Networks (CNN) show strong predictive capabilities for complex material behaviors. While ANN and SVM models are effective for smaller datasets, ensemble methods like GB and RF excel in handling larger, nonlinear data sets but are computationally expensive. GPR stands out for its ability to model complex relationships and provide uncertainty estimates, particularly in small to medium datasets, though it requires extensive tuning and computational power. Deep learning methods, such as CNN, extract deep features without manual preprocessing but require significant computational resources.

In classification and pattern recognition, unsupervised models, especially K-means clustering, are frequently applied for pavement condition classification and crack detection. Principal Component Analysis (PCA) is often used for dimensionality reduction, improving model efficiency but facing challenges with nonlinearity. Advanced techniques such as DBSCAN and CNNs are gaining traction for anomaly

detection and feature extraction but require extensive computational power and proper parameter tuning.

In general, machine learning models are effective tools in material testing and performance prediction, offering flexibility and accuracy, though computational complexity and data quality remain key challenges.

Credit Authorship Contribution Statement

Meisam Khorshidi: Conceptualization, Methodology, Literature review, Investigation, Writing - original draft, Visualization.

Eshan Dave: Conceptualization, Supervision, Writing - review & editing, Project administration.

Jo Sias: Conceptualization, Supervision, Writing - review & editing, Project administration.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments:

No external funding or assistance was received for this research.

References

[1] J. Liu, F. Liu, Z. Wang, E. O. Fanijo, and L. Wang, "Involving prediction of dynamic modulus in asphalt mix design with machine learning and mechanical-empirical analysis," *Constr. Build. Mater.*, vol. 407, no. October, p. 133610, 2023, doi: 10.1016/j.conbuildmat.2023.133610.

[2] X. Fan, S. Lv, C. Xia, D. Ge, C. Liu, and W. Lu, "Strength prediction of asphalt mixture under interactive conditions based on BPNN and SVM," *Case Stud. Constr. Mater.*, vol. 21, no. July, p. e03489, 2024, doi: 10.1016/j.cscm.2024.e03489.

[3] J. Zhang, Y. Huang, Y. Wang, and G. Ma, "Multi-objective optimization of concrete mixture proportions using machine learning and metaheuristic algorithms," *Constr. Build. Mater.*, vol. 253, p. 119208, 2020, doi: 10.1016/j.conbuildmat.2020.119208.

[4] F. Rondinella, C. Oreto, F. Abbondati, and N. Baldo, "Laboratory Investigation and Machine Learning Modeling of Road Pavement Asphalt Mixtures Prepared with Construction and Demolition Waste and RAP," *Sustain.*, vol. 15, no. 23, 2023, doi: 10.3390/su152316337.

[5] M. Khorshidi, M. Ameri, and A. Goli, "Cracking performance evaluation and modelling of RAP mixtures containing different recycled materials using deep neural network model," *Road Mater. Pavement Des.*, 2023, doi: 10.1080/14680629.2023.2222835.

[6] A. Upadhya, M. S. Thakur, A. Mashat, G. Gupta, and M. S. Abdo, "Prediction of Binder Content in Glass Fiber Reinforced Asphalt Mix Using Machine Learning Techniques," *IEEE Access*, vol. 10, pp. 33866–33881, 2022, doi: 10.1109/ACCESS.2022.3157639.

[7] B. Keshtegar, M. Bagheri, and Z. M. Yaseen, "Shear strength of steel fiber-unconfined reinforced concrete beam simulation: Application of novel intelligent model," *Compos. Struct.*, vol. 212, no. January, pp. 230–242, 2019, doi: 10.1016/j.compstruct.2019.01.004.

[8] R. A. Mozumder, B. Roy, and A. I. Laskar, "Support Vector Regression Approach to Predict the Strength of FRP Confined Concrete," *Arab. J. Sci. Eng.*, vol. 42, no. 3, pp. 1129–1146, 2017, doi: 10.1007/s13369-016-2340-y.

[9] S. Rahman, A. Bhasin, and A. Smit, "Exploring the use of machine learning to predict metrics related to asphalt mixture performance," *Constr. Build. Mater.*, vol. 295, p. 123585, 2021, doi: 10.1016/j.conbuildmat.2021.123585.

[10] A. L. Bonifácio, J. C. Mendes, M. C. R. Farage, F. S. Barbosa, C. B. Barbosa, and A. L. Beaucour, "Application of support vector machine and finite element method to predict the mechanical properties of concrete," *Lat. Am. J. Solids Struct.*, vol. 16, no. 7 CILAMCE 2018, pp. 1–11, 2019, doi: 10.1590/1679-78255297.

[11] M. Mohtasham Moein *et al.*, "Predictive models for concrete properties using machine learning and deep learning approaches: A review," *J. Build. Eng.*, vol. 63, no. August 2022, 2023, doi: 10.1016/j.jobe.2022.105444.

[12] Ba-Nhan Phung, Thanh-Hai Le, Minh-Khoa Nguyen, Thuy-Anh Nguyen, and Hai-Bang Ly, "Practical Numerical Tool for Marshall Stability Prediction Based On Machine Learning: An Application for Asphalt Concrete Containing Basalt Fiber," *J. Sci. Transp. Technol.*, vol. 3, no. 3, pp. 27–45, 2023, doi: 10.58845/jstt.utt.2023.en.3.3.27-45.

[13] D. C. Feng *et al.*, "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," *Constr. Build. Mater.*, vol. 230, p. 117000, 2020, doi: 10.1016/j.conbuildmat.2019.117000.

[14] N. Baldo, M. Miani, F. Rondinella, J. Valentin, P. Vackcová, and E. Manthos, "Stiffness Data of High-Modulus Asphalt Concretes for Road Pavements: Predictive Modeling by Machine-Learning," *Coatings*, vol. 12, no. 1, pp. 1–20, 2022, doi: 10.3390/coatings12010054.

[15] H. Tanyildizi, "Prediction of the strength properties of carbon fiber-reinforced lightweight concrete exposed to the high temperature using artificial neural network and support vector machine," *Adv. Civ. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/5140610.

[16] H. Hafez, A. Teirelbar, R. Kurda, N. To, and A. De, "Pre-bcc : A novel integrated machine learning framework for predicting mechanical and durability properties of blended cement concrete," vol. 352, no. August, 2022, doi: 10.1016/j.conbuildmat.2022.129019.

[17] P. Marcelino, M. de Lurdes Antunes, E. Fortunato, and M. C. Gomes, "Machine learning approach for pavement performance prediction," *Int. J. Pavement Eng.*, vol. 22, no. 3, pp. 341–354, 2021.

[18] H. Nguyen, T. Vu, T. P. Vo, and H. T. Thai, "Efficient machine learning models for prediction of concrete strengths," *Constr. Build. Mater.*, vol. 266, p. 120950, 2021, doi: 10.1016/j.conbuildmat.2020.120950.

[19] A. Behnood and E. M. Golafshani, "Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves," *J. Clean. Prod.*, vol. 202, pp. 54–64, 2018, doi: 10.1016/j.jclepro.2018.08.065.

[20] S. E. Whang and J. G. Lee, "Data Collection and Quality Challenges for Deep Learning," *Proc. VLDB Endow.*, vol. 13, no. 12, pp. 3429–3432, 2020, doi: 10.14778/3415478.3415562.

[21] H. G. Melhem and S. Nagaraja, "Machine learning and its application to civil engineering systems," *Civ. Eng. Syst.*, vol. 13, no. 4, pp. 259–279, 1996, doi: 10.1080/02630259608970203.

[22] C. Rudin, “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,” *Nat. Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019, doi: 10.1038/s42256-019-0048-x.

[23] M. Khorshidi, A. Goli, M. Orešković, K. Khayambashi, and M. Ameri, “Performance Evaluation of Asphalt Mixtures Containing Different Proportions of Alternative Materials,” *Sustainability*, vol. 15, no. 18, p. 13314, 2023, doi: 10.3390/su151813314.

[24] J. Liu, F. Liu, C. Zheng, D. Zhou, and L. Wang, “Optimizing asphalt mix design through predicting the rut depth of asphalt pavement using machine learning,” *Constr. Build. Mater.*, vol. 356, no. September, p. 129211, 2022, doi: 10.1016/j.conbuildmat.2022.129211.

[25] J. Liu, F. Liu, C. Zheng, D. Zhou, and L. Wang, “Optimizing asphalt mix design through predicting effective asphalt content and absorbed asphalt content using machine learning,” *Constr. Build. Mater.*, vol. 325, no. November 2021, p. 126607, 2022, doi: 10.1016/j.conbuildmat.2022.126607.

[26] J. Liu, F. Liu, C. Zheng, E. O. Fanijo, and L. Wang, “Improving asphalt mix design considering international roughness index of asphalt pavement predicted using autoencoders and machine learning,” *Constr. Build. Mater.*, vol. 360, no. April, p. 129439, 2022, doi: 10.1016/j.conbuildmat.2022.129439.

[27] H. Majidifard, B. Jahangiri, P. Rath, L. Urra Contreras, W. G. Buttlar, and A. H. Alavi, “Developing a prediction model for rutting depth of asphalt mixtures using gene expression programming,” *Constr. Build. Mater.*, vol. 267, p. 120543, 2021, doi: 10.1016/j.conbuildmat.2020.120543.

[28] N. Tiwari, N. Baldo, N. Satyam, and M. Miani, “Mechanical Characterization of Industrial Waste Materials as Mineral Fillers in Asphalt Mixes: Integrated Experimental and Machine Learning Analysis,” *Sustain.*, vol. 14, no. 10, pp. 1–25, 2022, doi: 10.3390/su14105946.

[29] N. Tiwari, F. Rondinella, N. Satyam, and N. Baldo, “Experimental and Machine Learning Approach to Investigate the Mechanical Performance of Asphalt Mixtures with Silica Fume Filler,” *Appl. Sci.*, vol. 13, no. 11, 2023, doi: 10.3390/app13116664.

[30] Y. Ali, F. Hussain, M. Irfan, and A. S. Buller, “An eXtreme Gradient Boosting model for predicting dynamic modulus of asphalt concrete mixtures,” *Constr. Build. Mater.*, vol. 295, p. 123642, 2021, doi: 10.1016/j.conbuildmat.2021.123642.

[31] D. Mirzaiyanrajeh, E. V. Dave, J. E. Sias, and P. Ramsey, “Developing a prediction model for low-temperature fracture energy of asphalt mixtures using machine learning approach,” *Int. J. Pavement Eng.*, vol. 24, no. 2, 2022, doi: 10.1080/10298436.2021.2024185.

[32] J. Liu, C. Cheng, C. Zheng, X. Wang, and L. Wang, “Rutting prediction using deep learning for time series modeling and K-means clustering based on RIOHTrack data,” *Constr. Build. Mater.*, vol. 385, no. April, p. 131515, 2023, doi: 10.1016/j.conbuildmat.2023.131515.

[33] A. M. Al-Sabaeei *et al.*, “Utilization of response surface methodology and machine learning for predicting and optimizing mixing and compaction temperatures of bio-modified asphalt,” *Case Stud. Constr. Mater.*, vol. 18, no. April, p. e02073, 2023, doi: 10.1016/j.cscm.2023.e02073.

[34] Y. Song, X. Wang, H. Li, Y. He, Z. Zhang, and J. Huang, “Mixture Optimization of Cementitious Materials Using Machine Learning and Metaheuristic Algorithms: State of the Art and Future Prospects,” *Materials (Basel).*, vol. 15, no. 21, 2022, doi: 10.3390/ma15217830.

[35] H. Hafez, A. Teirelbar, N. To, and A. De Fuente, “Data-driven optimization tool for the functional , economic , and environmental properties of blended cement concrete using supplementary cementitious materials,” vol. 67, no. February, 2023, doi: 10.1016/j.jobe.2023.106022.

[36] O. P. Pfeiffer *et al.*, “Cement and Concrete Research Bayesian design of concrete with amortized Gaussian processes and multi-objective optimization,” *Cem. Concr. Res.*, vol. 177, no. October 2022, p. 107406, 2024, doi: 10.1016/j.cemconres.2023.107406.

[37] Y. Yu, W. Li, J. Li, and T. N. Nguyen, “A novel optimised self-learning method for compressive strength prediction of high performance concrete,” *Constr. Build. Mater.*, vol. 184, pp. 229–247, 2018, doi: 10.1016/j.conbuildmat.2018.06.219.

[38] A.-D. Pham, N.-D. Hoang, and Q.-T. Nguyen, “Predicting Compressive Strength of High-Performance Concrete Using Metaheuristic-Optimized Least Squares Support Vector Regression,” *J. Comput. Civ. Eng.*, vol. 30, no. 3, pp. 28–31, 2016, doi: 10.1061/(asce)cp.1943-5487.0000506.

[39] Z. M. Yaseen *et al.*, “Predicting compressive strength of lightweight foamed concrete using extreme learning machine model,” *Adv. Eng. Softw.*, vol. 115, no. September 2017, pp. 112–125, 2018, doi: 10.1016/j.advengsoft.2017.09.004.

[40] B. A. Omran, Q. Chen, and R. Jin, “Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete,” *J. Comput. Civ. Eng.*, vol. 30, no. 6, pp. 1–13, 2016, doi: 10.1061/(asce)cp.1943-5487.0000596.

[41] B. G. Aiyer, D. Kim, N. Karingattikkal, P. Samui, and P. R. Rao, “Prediction of compressive strength of self-compacting concrete using least square support vector machine and relevance vector machine,” *KSCE J. Civ. Eng.*, vol. 18, no. 6, pp. 1753–1758, 2014, doi: 10.1007/s12205-014-0524-0.

[42] P. Yuvaraj, A. Ramachandra Murthy, N. R. Iyer, S. K. Sekar, and P. Samui, “Support vector regression based models to predict fracture characteristics of high strength and ultra high strength concrete beams,” *Eng. Fract. Mech.*, vol. 98, no. 1, pp. 29–43, 2013, doi: 10.1016/j.engfracmech.2012.11.014.

[43] K. Yan and C. Shi, “Prediction of elastic modulus of normal and high strength concrete by support vector machine,” *Constr. Build. Mater.*, vol. 24, no. 8, pp. 1479–1485, 2010, doi: 10.1016/j.conbuildmat.2010.01.006.

[44] A. Nazari and J. G. Sanjayan, “Modelling of compressive strength of geopolymer paste, mortar and concrete by optimized support vector machine,” *Ceram. Int.*, vol. 41, no. 9PartB, pp. 12164–12177, 2015, doi: 10.1016/j.ceramint.2015.06.037.

[45] F. Deng, Y. He, S. Zhou, Y. Yu, H. Cheng, and X. Wu, “Compressive strength prediction of recycled concrete based on deep learning,” *Constr. Build. Mater.*, vol. 175, pp. 562–569, 2018, doi: 10.1016/j.conbuildmat.2018.04.169.

[46] M. R. Kaloop, A. R. Gabr, S. M. El-Badawy, A. Arisha, S. Shwally, and J. W. Hu, “Predicting resilient modulus of recycled concrete and clay masonry blends for

pavement applications using soft computing techniques," *Front. Struct. Civ. Eng.*, vol. 13, no. 6, pp. 1379–1392, 2019, doi: 10.1007/s11709-019-0562-2.

[47] M. Y. Cheng, P. M. Firdausi, and D. Prayogo, "High-performance concrete compressive strength prediction using Genetic Weighted Pyramid Operation Tree (GWPOT)," *Eng. Appl. Artif. Intell.*, vol. 29, pp. 104–113, 2014, doi: 10.1016/j.engappai.2013.11.014.

[48] J. Zhang, G. Ma, Y. Huang, J. sun, F. Aslani, and B. Nener, "Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression," *Constr. Build. Mater.*, vol. 210, pp. 713–719, 2019, doi: 10.1016/j.conbuildmat.2019.03.189.

[49] M. Shao, Y. Sun, and X. Yun, "Long-term Evolution Recognition and Management Level Evaluation of Pavement Performance Based on Clustering Analysis," *Int. Conf. Geoinformatics*, vol. 2022-Augus, pp. 1–6, 2022, doi: 10.1109/Geoinformatics57846.2022.9963880.

[50] S. Mathavan, M. M. Rahman, and M. Stonecliffe-Jones, "Unsupervised artificial neural network for efficient mapping of doweled concrete pavement joints condition," *Int. J. Pavement Res. Technol.*, vol. 7, no. 4, pp. 287–296, 2014, doi: 10.6135/ijprt.org.tw/2014.7(4).287.

[51] S. Mubashshira, M. M. Azam, and S. M. Masudul Ahsan, "An Unsupervised Approach for Road Surface Crack Detection," *2020 IEEE Reg. 10 Symp. TENSYMP 2020*, no. June, pp. 1596–1599, 2020, doi: 10.1109/TENSYMP50017.2020.9231023.

[52] W. Li, J. Huyan, R. Gao, X. Hao, Y. Hu, and Y. Zhang, "Unsupervised Deep Learning for Road Crack Classification by Fusing Convolutional Neural Network and K_Means Clustering," *J. Transp. Eng. Part B Pavements*, vol. 147, no. 4, 2021, doi: 10.1061/jpeodx.0000322.

[53] A. Golmohammadi, D. Hernando, and W. Van Den Bergh, *Advanced data-driven FBG sensor-based pavement monitoring system using multi-sensor data fusion and an unsupervised learning approach*. Elsevier Ltd, 2024. doi: 10.1016/j.measurement.2024.115821.

[54] A. Abdelmawla, J. J. Yang, and S. S. Kim, "Unsupervised Learning of Pavement Distresses from Surface Images," in *Advances in Innovative Geotechnical Engineering*, Y. Liu, S. Cuomo, and J. Yang, Eds., Cham: Springer International Publishing, 2021, pp. 1–8.

[55] Q. Dong, X. Chen, S. Dong, and J. Zhang, "Classification of pavement climatic regions through unsupervised and supervised machine learnings," *J. Infrastruct. Preserv. Resil.*, vol. 2, no. 1, 2021, doi: 10.1186/s43065-021-00020-7.

[56] C. Shi *et al.*, "Analysis of crumb rubber content influence on damage evolution and pattern recognition of rubberised epoxy asphalt mixture using acoustic emission techniques," *Int. J. Pavement Eng.*, vol. 25, no. 1, pp. 1–15, 2024, doi: 10.1080/10298436.2024.2356762.

[57] M. N. Akhtar, W. Ahmed, M. R. Kakar, E. A. Bakar, A. R. Othman, and M. Bueno, "Implementation of Parallel K-Means Algorithm to Estimate Adhesion Failure in Warm Mix Asphalt," *Adv. Civ. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/8848945.

[58] A. Sahari Moghaddam, E. Rezazadeh Azar, Y. Mejias, and H. Bell, "Estimating Stripping of Asphalt Coating Using k-Means Clustering and Machine Learning–Based Classification," *J. Comput. Civ. Eng.*, vol. 34, no. 1, pp. 1–11, 2020, doi: 10.1061/(asce)cp.1943-5487.0000864.