A robust approach to shear strength prediction of reinforced concrete deep beams using ensemble learning with SHAP interpretability

A Robust Approach to Shear Strength Prediction of Reinforced Concrete Deep Beams using Ensemble Learning with SHAP Interpretability

Achyut Tiwari[a,*], Ashok Kumar Gupta[b], Tanmay Gupta[c]

[a]*Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat, Solan, 173234, Himachal Pradesh, India*
[b]*Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat, Solan, 173234, Himachal Pradesh, India*
[c]*Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat, Solan, 173234, Himachal Pradesh, India*

Abstract

The behavior of reinforced concrete (RC) deep beams is complex and difficult to predict due to factors such as compressive and shear stress and beam geometry. To address this challenge, researchers have proposed various machine learning models such as Artificial Neural Network, Decision Tree, Support Vector Machine, Adaptive Boosting, Extreme Gradient Boosting, Random Forest, Gradient Boosting, and Voting Regressor. In this study, the authors evaluated the performance of these models in predicting shear strength of RC deep beams by using metrics such as R2, Mean Squared Error, Root Mean Squared Error, Mean Absolute Percentage Error and Mean Absolute Error. Furthermore, the authors optimize the ensemble learning models using customized hyperparameters. The XGBoost model exhibited the highest accuracy with an R2 value of 0.92 and the least model error, with MAE of 29.65 and RMSE of 47.76 & MAPE of 9.79.The authors compared these models with mechanics-driven models from different country codes including the United States, China, Europe, British (CIRIA), Canada and found that ensemble learning models, specifically XGBoost, outperformed mechanics-driven models. The authors used an explainable machine learning (EML) technique called SHapley Additive exPlanations (SHAP) to gain additional insights into the developed XG-Boost model. The outcomes of feature selection and SHAP analysis suggest that the grade of concrete and beam geometry predominantly influence the prediction of shear strength in RC deep beams, whereas steel properties exert minimal impact in this regard.

Keywords: Ensemble Learning, Machine Learning, Boosting, Reinforced Concrete Deep Beams, Shear Strength, SHapley Additive exPlanations (SHAP)

*Corresponding Author
Email addresses: achyuttiwari@icloud.com (Achyut Tiwari), ashok.gupta@juit.ac.in (Ashok Kumar Gupta), tanmay.gupta@juit.ac.in (Tanmay Gupta)

1. Introduction

The exponential growth and advancements in computational sciences have brought new technologies such as artificial intelligence (AI) and machine learning (ML) [1, 2, 3]. In various fields of engineering As computational power has become more democratised and widely available for research purposes, new paradigms are evolving some of which are structural health monitoring, computer-aided design development, longevity of structures etc. [4, 5, 6, 7, 8, 9, 10, 11]. These computational sciences have permeated various sectors of the economy, including the real estate industry. Modern structures focuses on various factors such as resilience, safety, sustainability, reliability, economy and aesthetics. Most fundamental building block of any infrastructure is reinforced concrete beams which is widely used in construction to support loads and distribute them to the columns or walls [12, 13]. The use of RC Deep Beams has increased exponentially since the start of building taller structures [14, 15]. However, these tall structures face various failures like tensile, bending, or shear failures, which can be prevented by embedding steel reinforcing bars in concrete beams [16]. Shear failure, caused by shear force combined with axial loads and moments, is one of the most dangerous failure types as it can occur without warning [17]. In contrast, flexural failure develops gradually due to the yielding of rebars [18]. The shear transmission process becomes random after shear fractures begin [17].

Studies have shown that computational science plays a vital role in engineering, specifically in Civil Engineering. Ensemble learning, a computer science field, is widely used in various disciplines such as biology, engineering, and sociology [19, 20, 21, 22, 23]. Machine Learning (ML) is commonly used in building structural design and performance assessment, enhancing concrete properties predictions, and improving the finite element modeling of structures [24, 25, 26, 27, 28, 29]. Ensemble learning is a powerful ML technique that improves the accuracy of predictions made by a model [30, 19]. It is particularly useful for large datasets with many features, as it trains a group of models on different subsets of the data and combines their predictions to make a final prediction. This technique can be used to reduce variance in the model's predictions [30].

Predicting reinforced concrete beam shear strength is a complex problem due to the nature of the materials involved. Using ensemble learning can be beneficial in reducing variance in predictions made by the model [31]. The most common type of shear reinforcement in concrete beams is stirrups, which transfer shear forces between the concrete and steel. There are several ways to predict the shear strength of reinforced concrete beams, but empirical codes are the most common method. Every country has its own empirical codes to find shear strength. In the US, the American Concrete Institute (ACI) 318 code governs the design of reinforced concrete beams Committee [32]. Similarly, in Europe, the use of Eurocode 2 Bethlehem [33] for designing concrete structures is prominent.

This study aims to formulate and compare various boosting machine learning algorithms to predict the shear strength of Reinforced Concrete (RC) deep beams, which is a complex task due to uncertain factors. The study explores ensemble learning methods like Adaptive Boosting, Extreme Gradient Boosting, Random Forest, Gradient Boosting, and Voting Regressor. In addition conventional ML algorithms including ANN, DT & SVM are also compared. The objective is to identify the most effective machine learning approach that outperforms traditional mechanics-driven models based on country codes, including the United States [32], China [34], Europe [33], British (CIRIA) [35], and Canada [36] by evaluating their performance using metrics like R^2, MSE, RMSE, and MAE. This comparison aims to assess whether the machine learning-based approach, particularly the XGBoost model, can surpass the accuracy and performance of traditional mechanics-driven models in predicting the shear strength of RC deep beams across various regions.

The study also incorporates the use of the explainable machine learning (EML) technique, SHapley Additive exPlanations (SHAP) [37], to gain interpretability and insights into the developed best-performing model. This step is crucial for understanding the factors that contribute to the predictions and enhancing transparency in the decision-making process. The authors have also performed feature selection analysis to understand how varied parameters affect the prediction of shear strength while using ML Models.

In summary, the objectives of this study are to advance the understanding and prediction capabilities of RC deep beam shear strength through the application of state-of-the-art machine learning methods. By comparing these models against established mechanics-driven models, the study aims to provide engineers and researchers with a more accurate and reliable tool for designing and assessing the structural behaviour of RC deep beams, ultimately contributing to advancements in the field of civil engineering and construction.

The simulation conducted uses Python language to code and build the relevant models on colab.research.google.com. For comparison and calculations of mechanics-driven models, Microsoft excel was used.

The remainder of this paper will describe the literature review in Section 2, methodology of ML models in Section 3, model structure, dataset collection, dataset limitation, model selection, model evaluation and hyper-parameter optimisation in Section 4 and comparison between conventional and ensemble models, comparison between ML models and mechanics driven models, SHAP of XGboost model, feature importance analysis in Section 5.

2. Literature Review

The use of artificial intelligence (AI) and machine learning (ML) techniques in structural engineering has been in play since the 1980s as the researchers realised the conventional approaches e.g., finite element models and analytical models have difficulties in accurately and efficiently predicting the structural behaviours [1]. With the rapid development and democratisation of computer science, new paradigms came into play. As the computation power was widely available, more powerful algorithms were proposed which turned out to widen the scope of structural engineering. There are 3 major fields that showed significant progress i.e. Structural Health Monitoring (SHM), performance evaluation and modelling of mechanical properties [1].

SHM uses machine learning as it collects huge amounts of data using various sensors and later processes these large amounts of collected data making data-driven models. In addition, unsupervised algorithms or clustering techniques can also be used [38, 39, 40].

Performance evaluation is another major area that has improved with the implementation of ML. Conventional performance evaluation methods including fragility and reliability assessment require huge amounts of data in order to take into account the uncertainty and randomness in the structure. As seen in SHM, the data is collected using sensors which can even collect real-time data. Applying ML to formulate dynamic models in accordance with situational data results in low use of computation as seen in various studies [41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51].

In the last two decades, there has been a prevalence of modelling the mechanical behavior of structures, while also exploring the diverse usage of concrete and its dynamic behaviors in various structures. Concrete has been widely employed in the construction of structures due to its advantageous engineering characteristics, including rich raw materials, low cost, strong compressive strength, and exceptional durability. Long spans and structures without intermediary columns both benefit from the use of deep beams. Deep beams are employed as girders to support the carriageway in bridges. Deep Beams are also employed as side walls in RCC water tanks and as connections for the pile caps in pile foundations. The shear span depth ratio is used to classify deep beams, which are ones with a greater depth than commonly utilised beams. The deep beams are defined differently by different codes. The beams having a depth greater with respect to its span are generally referred to as deep beams [35, 32, 34, 36, 33]. The ratio of effective span to overall depth when considered less than 3 the beam is called as deep as per Eurocode [33]. , As per ACI Code [32] shear design is specially done when clear span to effective depth ratio is less than 5 [32]. Leonhardt and Walter 1966 [52], experimentally proved that elastic design for such deep beam is not valid. Their investigation further highlighted the significance of accurate steel details in deep beams. The distribution of strain in a section of a deep beam is not linear and cannot be determined by elasticity theory. In general impact of shear in beam design is taken care of by longitudinal reinforcement provisions. However, In case of excess shear transverse reinforcement is separately designed [53]. For the case of deep beams such simplifying assumptions are not adequate and various approaches such as compression field theory, tension field theory etc. are proposed by various researchers based on which different country codes have proposed their procedures for shear design [35, 32, 34, 36, 33]. The amount of reinforcement to be used and concrete directly depends on accurate prediction of the shear capacity of a section [54].

Machine learning (ML) is one of the widely accepted methods to tackle structural problems [55, 56, 57, 58, 59, 60, 61, 62, 63, 64]. In Sandeep et al. [65], the authors thoroughly discuss the implementation of ML approaches for predicting the shear strength of RC deep beams, covering in-depth procedures, various algorithms, and the basics of modeling, training, testing, underfitting, and overfitting. However, the authors show no real-time implementation and results. In Fu and Feng [66], the authors formulate ML algorithms to predict the shear strength of Corroded reinforced concrete beams using a gradient gradient-boosting regression tree. Authors use 158 shear tests for the corroded reinforced concrete beam meanwhile showing how empirical models cannot take into account how corrosion influences calculations. The authors also calculate Time-dependent corrosion extent and lifetime shear strength prediction. However, the issue of interpretability isn't discussed. Similarly, the dataset taken into account is very small. In Chou et al. [67], the authors integrate the smart fly algorithm and least square support vector regression to build a hybrid model into a multi-source dataset sourced from North America, Australia & America. The hybrid model shows promising results of MAPE 18.95%. However, the authors show no correlations of features and any method for interpretability of models employed. In Naik and Kute [68], the authors have implemented an artificial neural net for predicting the shear strength of high-strength steel fibre-reinforced concrete deep beams. The validation method used is the residual sum of squares. Authors also establish a relationship between various features using ANN. The developed ANN8 establishes the relations between various parameters affecting the complex behaviour of steel fibre-reinforced concrete deep beams. In Concha et al. [69], authors develop a hybrid Neuro-Swarn model to predict the shear strength of steel fibre-reinforced concrete deep beams. The model was developed using 116 experimental datasets. The analysis of the variance test showed prominent results. Authors also present various models used for shear strength calculation and prediction in conventional approaches Committee [32], Vamdewalle and Mortelmans [70], Al-Ta'an and Al-Feel [71], Sharma [72], Khuntia et al. [73], Cho and Kim [74]. However the experimental data size is small which may result in overfitting. In Pak et al. [75], the authors have proposed a novel approach named the transfer ensemble neural network (TENN) model to increase the performance of the model while predicting shear capacity on small datasets. In the models, authors have incorporated both ensemble learning and transfer learning in order to control the high variability of ML models. However the results are impressive, similarly the the issue of the black box approach and overfitting remains an open issue. In Almasabha et al. [76], the authors have worked on a new dataset of 102 instances of synthetic fibre-reinforced concrete (SyFRC) for reinforced concrete structures. Authors predict the shear strength of SyFRC beams without stirrups using ACI code and ML algorithms- LightGBM, XGBoost and Gene Expression. The study shows that, apart from the ACI equation, all considered models effectively predict the effects of the shear span-to-depth ratio. In Ly et al. [77], the authors have implemented real-code genetic algorithms and animal-based firefly algorithms in order to predict the shear strength of reinforced concrete deep beams. The dataset contains 463 instances. Later in the study, the authors compare the obtained results with neural nets which shows promising results. In Olalusi and Awoyera [78], the authors implement Gaussian Process regression (GPR) and the Random Forest (RF) to predict the shear resistance of steel fibre-reinforced concrete slender beams without stirrups. The results obtained during the study were compared with statistical and German guidelines. The authors also present the inconsistencies in prediction observed during the study. In Hossain et al. [79], the authors have formulated an ANN approach to predict shear strength on the experimental database containing 173 steel fibre-reinforced concrete (SFRC) beams without stirrups. Additionally, the approach is tested with data from 36 experimental beams. The authors show how ANN is better when it comes to empirical equations for high and ultra-high strength of SFRC beams. However other possible techniques are not explored in this scenario. In Tapeh and Naser [1], the authors have conducted state of a state-of-the-art review for AI, ML & Deep Learning (DL) implementations in structural engineering, particularly earthquake, wind, and fire engineering. The authors introduce a wide range of techniques and their varied implications and benefits in the field of structural engineering. Authors cover more than 4000 scholarly works in order to identify best practices. The authors also cover shear strength prediction for RC deep beams, however, the scholarly works are limited to only two on the specific issue. Overall, the paper gives an overview of the last decades of how AI, ML & DL have shaped structural engineering. In Marie et al. [80], the authors present a framework predicting the shear strength of reinforced concrete beam-column connections which is subjected to cyclic loading. The authors use classical prediction models such as K-nearest neighbour regression (KNN), Multivariate Adaptive Regression Splines (MARS), Ordinary least Squares (OLS), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and kernel regression with mixed data types (Kernel regression) which are implemented on a dataset of 98 instances. The authors show kernel regression predicted the joint shear strength with the highest accuracy. However, neither model interpretability nor feature importance is present. In Wakjira et al. [81] authors have implemented Existing predictive models which have shown unsatisfactory results. In response, the research proposed machine learning (ML) based models, considering all important variables, for predicting shear capacity. The analyses demonstrated successful predictions using the ML models, with extreme gradient boosting (XGBoost) showing the highest capability. Comparisons with existing models revealed the superiority of XGBoost in terms of accuracy, safety, and economic aspects. However, limitations concerning model interpretability were not addressed. Finally, reliability analysis was performed to calibrate resistance reduction factors, improving the confidence and applicability of the proposed model. Further research is needed to address this issue and explore additional avenues for enhancing ML techniques in structural engineering. In Liu et al. [82], the authors aimed to establish an accurate prediction model for precast concrete joints (PCJ) direct shear strength (DSS) using support vector regression (SVR), a machine learning algorithm. They assembled a comprehensive database of 304 test results with 23 input parameters and employed a novel correlation matrix-based feature selection method for improving the SVR model's performance. The experimental validation showed that the SVR model outperformed traditional mechanical models in predicting DSS for PCJs. Additionally, the study provided insights into the SVR model's results using partial dependence and individual conditional expectation plots. Another study addressed the challenges in accurately predicting the shear strength of fiber-reinforced steel (FRS) due to the complex soil-fiber interaction mechanism. To tackle this, they compiled a high-quality database of triaxial and direct shear tests on FRS from 1983 to 2015, including crucial information on sand properties, fiber characteristics, soil-fiber interface properties, and stress parameters. This database served as a solid foundation for further analysis and future developments of improved mechanical models for predicting FRS shear strength.

3. Methodology

3.1. Ensemble Learning

Ensemble learning is a machine learning technique that combines multiple models to improve the accuracy and robustness of predictions. Ensemble learning is a very advanced and significant machine learning technique within the academic domain. The core principle of this approach is centred on combining various foundational models or "learners" to generate a more powerful prediction model that exhibits improved accuracy and robustness. This technique is based on the long-standing belief that the combined knowledge and insights of a group frequently exceed those of an individual. Within the domain of machine learning, ensemble learning encompasses the use of this principle to algorithms, hence showcasing the potential for enhanced predictive results through the collaborative integration of several models.

Ensemble learning, at its fundamental essence, aims to enhance predictive accuracy, strengthen generalisation skills, and reinforce model stability. The objective is to mitigate the inherent constraints of individual models through the use of variety and collaboration among the constituent learners. In the realm of academic discourse pertaining to ensemble learning, a number of crucial notions emerge as prominent.

The first consideration pertains to the concept of diversity within the foundational models. Diversity plays a fundamental role in ensemble learning, which is accomplished through a range of strategies including the utilisation of diverse algorithms, the incorporation of distinct subsets of data, and the introduction of variances in hyperparameters throughout the training process. The underlying concept posits that the presence of diverse models results in distinct errors being made on various portions of the data. This collective diversity ultimately enhances the probability of making accurate predictions.

Another crucial factor to consider is the consolidation of forecasts generated by individual models. Ensemble methods utilise many aggregation approaches, such as majority voting, weighted averaging, and stacking, each of which is based on distinct mathematical concepts and possesses distinct features.

The selection of base learners is a critical aspect in the ensemble learning procedure. The category of basic learners includes both elementary models, such as decision trees, as well as more intricate ones, such as neural networks. The choice of suitable base learners is contingent upon the distinct attributes of the data and the inherent nature of the problem under consideration.

Ensemble learning comprises a range of ensemble forms, including bagging, boosting, and stacking, each characterised by unique methodologies for aggregating base models. The scholarly literature has exhaustively examined these many sorts of ensembles, providing insights into their individual merits and limitations.

Ensemble learning offers a structured approach to effectively manage the bias-variance trade-off, a crucial consideration within the field of machine learning. Ensembles has the ability to address the issue of overfitting, characterised by large variance, by integrating various models. Simultaneously, ensembles are capable of capturing detailed patterns in the data, hence minimising bias.

The topic of model interpretability is a subject of scholarly inquiry in the field of ensemble learning. Ensemble approaches frequently augment prediction performance, but concomitantly bring complexity to the overarching model. Scholars are currently engaged in the investigation of methods that aim to achieve a harmonious equilibrium between precision and interpretability of models, so guaranteeing that the knowledge obtained from the model remains lucid and comprehensible.

Finally, scholarly discourse surrounding ensemble learning encompasses its practical implementation in various fields, such as banking, healthcare, image identification, and natural language processing. Researchers continually strive to illustrate the capacity of ensemble methodologies to offer more effective solutions to real-world situations, thus emphasising the practical significance of ensemble learning.

The mathematical notation for ensemble learning involves defining a set of base models, and then combining them to produce a final prediction [83].

Let X be the input data, and Y be the target variable we wish to predict. We define a set of N base models, denoted by M1, M2, ..., MN. Each base model takes X as input and produces a predicted output, denoted by Mi(X).

The ensemble model then combines the predictions of the base models to produce a final prediction, denoted by F(X). There are many ways to combine the predictions of the base models, but one common approach is to use a weighted average is defined by equation (1).

$$F(X) = w1 * M1(X) + w2 * M2(X) + ... + wN * MN(X) \tag{1}$$

where w1, w2, ..., wN are the weights assigned to each base model. The weights can be learned from the data or set manually based on prior knowledge.

Ensemble learning is a popular machine learning technique that combines multiple models to achieve better accuracy and generalization performance than using a single model. In the context of classification, ensemble learning involves constructing a set of base classifiers that make predictions on a given dataset , and then combining these predictions using a specified aggregation method to obtain the final classification result [19].

3.1.1. Boosting

Boosting is a common ensemble learning method that sequentially trains a set of weak classifiers on re-weighted versions of the training data, such that the misclassified samples in each iteration receive higher weights in the subsequent iterations. The final classification is then obtained by weighted voting of the individual classifier outputs [84, 85]. Mathematically, the boosting algorithm can be formulated as follows:

Given a training dataset $D = (x_i, y_i)_{i=1}^n$, where x_i denotes the feature vector of the i-th sample and $y_i \in -1, +1$ represents its class label, and a set of weak classifiers $h_m(x)$, $m = 1, \ldots, M$, the boosting algorithm aims to learn a strong classifier $H(x)$ as follows:

1. Initialize sample weights $w_i = 1/n$, $i = 1, \ldots, n$.
2. For each iteration $m = 1, \ldots, M$:
   - Train the m-th weak classifier $h_m(x)$ on the weighted training dataset $D_m = (x_i, y_i, w_i)_{i=1}^n$.
   - Compute the error rate $\epsilon_m = \sum_{i=1}^n w_i I(y_i \neq h_m(x_i))$, where $I(\cdot)$ is the indicator function.
   - Compute the classifier weight $\alpha_m = \frac{1}{2} \log \frac{1-\epsilon_m}{\epsilon_m}$.
   - Update the sample weights as $w_i \leftarrow w_i \exp(-\alpha_m y_i h_m(x_i))$.
3. Output the final classifier is defined by equation (2).

$$H(x) = \text{sign}\left(\sum_{m=1}^{M} \alpha_m h_m(x)\right) \tag{2}$$

3.1.2. Stacking

Stacking is another popular ensemble learning technique that combines the outputs of multiple base classifiers using a higher-level meta-classifier, which is trained on the predictions of the base classifiers . Specifically, stacking consists of the following steps:

1. Split the training dataset D into k disjoint subsets, or folds, $D_1, \ldots, D_k$.

2. For each fold $i = 1, \ldots, k$: Train the M base classifiers on the $k-1$ folds other than D_i. Obtain the predicted class probabilities for the samples in D_i from each base classifier. Concatenate the predicted probabilities from all base classifiers to form a new feature vector for each sample in D_i. Store the new feature vectors and the corresponding true class labels as a new training dataset D'_i.
3. Train a meta-classifier, such as logistic regression or SVM, on the augmented training dataset $D'_1, \ldots, D'_k$.
4. Combine the base classifiers and the meta-classifier to form the final stacked classifier.

Mathematically, the stacking algorithm can be represented as follows:

Given a training dataset $D = (x_i, y_i)_{i=1}^n$, where x_i denotes the feature vector of the i-th sample and $y_i \in -1, +1$ represents its class label, and a set of base classifiers

3.1.3. Bootstrap Aggregating Algorithm

Bagging, short for Bootstrap Aggregating, is another popular ensemble learning method that trains multiple base classifiers on different bootstrap samples of the training data, and combines their outputs by majority voting to obtain the final classification. The bagging algorithm can be mathematically represented as follows:

Given a training dataset $D = (x_i, y_i)_{i=1}^n$, where x_i denotes the feature vector of the i-th sample and $y_i \in -1, +1$ represents its class label, and a set of weak classifiers $h_m(x)$, $m = 1, \ldots, M$, the bagging algorithm aims to learn a strong classifier $H(x)$ as follows:

1. For each iteration $m = 1, \ldots, M$: Generate a bootstrap sample D_m of size n by randomly sampling n samples from $Dwith$ replacement. Train the m-th weak classifier $h_m(x)$ on the bootstrap sample D_m.
2. Output the final classifier is defined by equation (3).

$$H(x) = \text{sign}\left(\sum_{m=1}^{M} h_m(x)\right). \tag{3}$$

3.2. Overview of the ML Models

3.2.1. Artificial Neural Network

An Artificial Neural Network is a computational model inspired by the structure and functionality of biological neural networks in the human brain. It is a type of machine learning algorithm designed to recognize patterns, solve complex problems, and make decisions based on input data [86].

3.2.2. Decision Tree

A decision tree is a non-linear predictive model and a popular supervised learning algorithm used for classification and regression tasks. It is a graphical representation of a set of rules and decisions based on input features that recursively partition the data into subsets, leading to a hierarchical tree-like structure [87].

3.2.3. Support Vector Machine

The concept behind Support Vector Machine(SVM) is to find the best decision boundary (hyperplane) that separates the data points of different classes with the largest margin possible. The data points closest to the hyperplane, known as support vectors, play a crucial role in defining the optimal hyperplane. These support vectors are used to determine the margin and influence the overall performance of the SVM [88].

3.2.4. Random Forest

Random forest is a supervised learning algorithm. It can be used for both classification and regression. The algorithm works by building multiple decision trees (hence the ”forest”) and then selecting the tree that predicts the label for a new data point which is the best. The decision trees are built using a random subset of the features, and the predictions are made by averaging the predictions of all the trees [89, 90].

3.2.5. Gradient Boosting

Gradient boosting is a machine learning technique that can be used for both regression and classification problems. It creates a prediction model as an ensemble of weak prediction models, often decision trees. Like other boosting methods, it builds the model incrementally in a stage-wise fashion. It also allows for the optimization of an arbitrary differentiable loss function, which helps to generalize the model. [91].

3.2.6. Adaptive Boosting

The Adaptive Boosting Algorithm is a classification technique that is used to improve the accuracy of a model by combining a set of weak models. The algorithm adaptively changes the weights of the models in the ensemble so that the model with the highest error rate is given more weight. The algorithm then continues to iteratively train the model and update the weights until the desired accuracy is achieved [92].

3.2.7. Extreme gradient Boosting(XGBoost)

The extreme gradient boosting algorithm is a powerful machine learning algorithm that is often used for classification tasks. This algorithm is a modification of the gradient boosting algorithm that is designed to be more efficient and to better handle data with a large number of features. The extreme gradient boosting algorithm works by building a model in a stage-wise fashion. In each stage, a new tree is added to the model and the predictions of the new tree are combined with the predictions of the existing trees in the model. The trees are added in a way that minimizes the loss function of the model. The elaborated model is shown in the Figure 1 [93].

The extreme gradient boosting algorithm is very effective at handling data with a large number of features. This is because the algorithm can choose which features to use in each stage of the model. This allows the algorithm to focus on the most important features and to ignore the less important features.The extreme gradient boosting algorithm is also effective at handling data that is imbalanced. This is because the algorithm can learn from the mistakes that it makes on the minority class and use this knowledge to improve the predictions on the minority class. XGBoost is also used in classifying image [94], malware detection [95], predicting the death of patient during COVID-19 treatment [96] and detecting fraudulent activities [97].

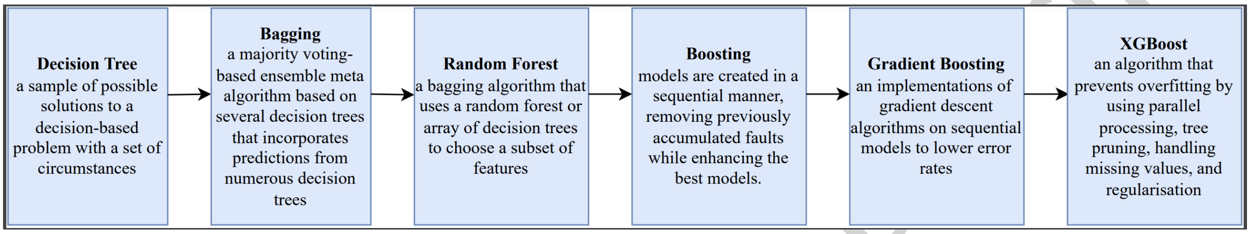

Figure 1: Evolution of XGBoost

3.2.8. Voting Regressor

A voting regressor is an ensemble learning method for regression that works by combining the predictions of multiple individual regressors. The individual regressors can be any type of regression algorithm, such as linear regression, support vector regression, or decision tree regression. The predictions from the individual regressors are combined using a simple majority vote. The voting regressor is a powerful tool because it can reduce the variance of the predictions, making the predictions more robust and accurate. In addition, the voting regressor can help to avoid overfitting because it is less likely to overfit to the training data than a single regressor [98].

4. Model Structure

4.1. Data Collection

The data set of RC beams is compiled from the published literature [99]. In this study, a total of 271 test data samples of RC beams from the literature were collected and used. These test data samples were related to RC deep beams of which 52 samples were from [100], 25 samples were from [101], 37 samples were from [102], 53 samples were from [103], 4 samples were from [104], 12 samples were from [105], 19 specimens are from [106]. 12 samples were from [107] and 39 samples were from [108] .

The database includes a wide range of RC deep beams so that the model can generate data more effectively. The database contains four different types of deep beams, including beam without web reinforcements (WOR), beams

12

with horizontal web reinforcements (WHR), beams with vertical web reinforcements (WVR), and beams with both horizontal and vertical reinforcements (WHVR). Four distinct deep beam types— beams without web reinforcements, beams with horizontal web reinforcements, beams with vertical web reinforcements, and beams with both horizontal and vertical web reinforcements—are represented in the dataset used for this work. In the data set, this classification is marked with the help of parameters such as area/spacing of vertical web reinforcement and area/spacing of horizontal web reinforcement.

Table 1: Statistical Information of parameters in deep beam database

Category	Variable	Unit	Min.	Max.	Mean	St.D.	Type
Geometrical dimensions	l_0	mm	500.00	4065.00	1484.45	643.56	Input
	h	mm	254.00	915.00	523.48	147.94	Input
	h_0	mm	216.00	844.00	469.01	145.44	Input
	b	mm	76.00	305.00	122.81	44.33	Input
	a	mm	125.00	1290.00	467.73	238.75	Input
	l_0/h	–	0.91	5.00	2.93	1.07	Input
	a/h_0	–	0.22	2.70	1.06	0.52	Input
Longitudinal reinforcement	ρ_l	%	0.12	4.08	1.62	0.71	Input
	f_y	MPa	210.00	504.80	361.76	86.86	Input
Horizontal web reinforcement	ρ_h	%	0.00	2.45	0.38	0.44	Input
	s_h	mm	50.18	801.00	127.63	98.95	Input
	f_{yh}	MPa	210.00	586.00	402.17	71.75	Input
Vertical web reinforcement	ρ_v	%	0.00	2.45	0.41	0.41	Input
	s_v	mm	50.80	457.50	204.48	99.03	Input
	f_{yv}	MPa	210.00	586.00	387.23	66.19	Input
Concrete property	f_c'	MPa	12.26	73.60	30.89	14.86	Input
Shear strength	V_u	kN	67.62	1357.00	287.20	180.67	Output

The input variables for these beams are 16 design features that fall into four groups, (1) geometric dimensions: beam span l_0, height h,effective height h_0, width b, shear span a.; (2) longitudinal reinforcement information: reinforcement ratio ρ_l and strength f_{yl}; (3) web reinforcement information: horizontal reinforcement ratio ρ_h, spacing s_h and strength f_{yh}, vertical reinforcement ratio ρ_v, spacing s_v and strength f_{yv}; (4) concrete property: concrete strength f_c'. The output is the beam's shear strength, denoted by V_u. The value ranges for these variables, as well as the statistical information (mean and standard derivation, etc.), are listed in Table 1. Meanwhile, Figure 2 also plots the distributions of the deep beam parameters frequencies.

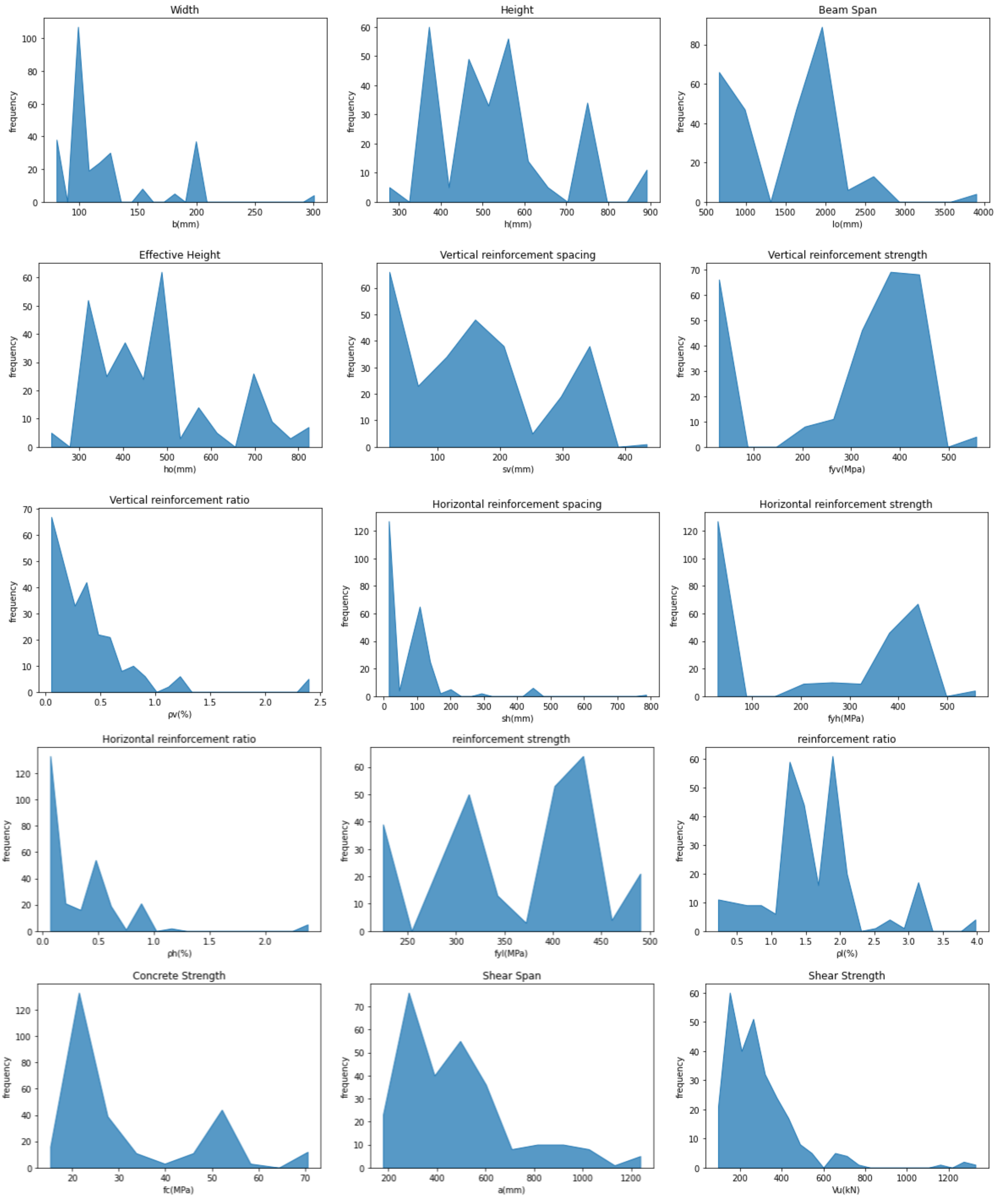

Figure 2: Deep Beam Parameters Frequencies from the database [99]

4.1.1. Limitations

The dataset used for the analysis of shear strength in RC (Reinforced Concrete) deep beams poses certain limitations that need to be considered when implementing machine learning algorithms. Firstly, the dataset contains only 271 samples, which might not be sufficient to fully capture the wide variability of RC deep beams in practice. A small sample size could lead to reduced statistical significance and limit the generalizability of the machine learning models.

Secondly, the data is retrieved from old construction sites, potentially introducing bias and representativeness issues. Construction practices, materials, and design standards may have evolved, making the dataset less relevant to current scenarios. This temporal difference might affect the accuracy of the predictions.

Thirdly, the limited sample size can result in a lack of diversity within the dataset. As a result, the machine learning algorithms might not adequately capture the variations in beam configurations, reinforcement details, and loading conditions, which are crucial factors influencing shear strength.

To address some of these limitations, researchers should interpret the results with caution.

4.2. Model Selection

In this study, the authors have utilised Random forest, Adaptive boosting, Gradient Boosting, XGBoost, Support Vector Machine (SVM) and ANN. The authors have also implemented Voting Regressor over the top best performing algorithms to ensure better and generalised results. Figure 3 shows the step-by-step model approach taken in this study.

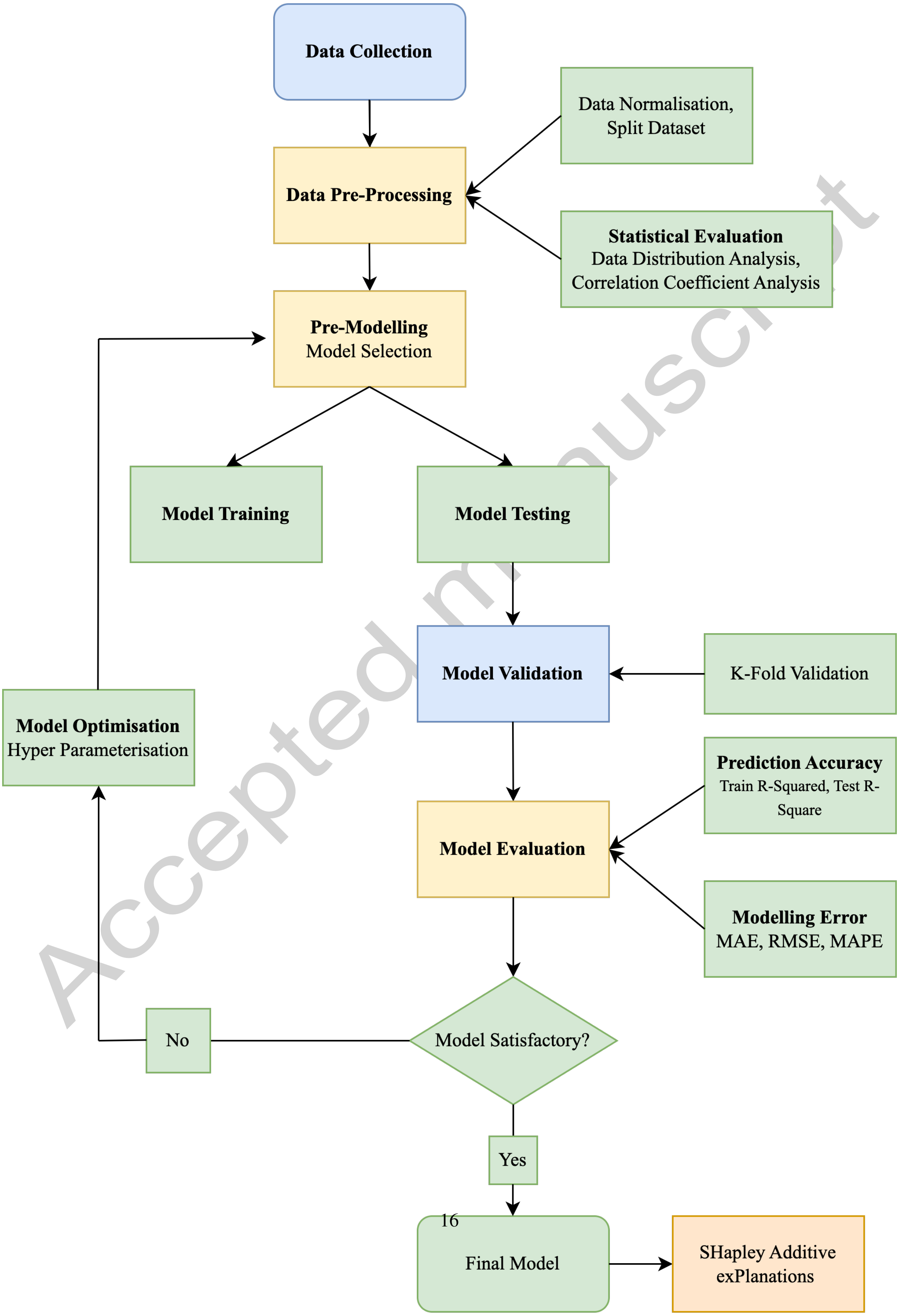

Figure 3: Step by step ensemble learning modelling approach with Interpretations

4.3. Hyper-parameter Optimisation

Once the data preprocessing is complete, the next task is to tune the hyperparameters in accordance with correlations and multiple other factors. In order to discover the hyperparameters, the grid search approach is paired with k-fold cross-validation (CV) as shown in Figure 4. The optimisation of model parameters is a crucial phase in ensemble learning, which involves making decisions on many factors such as the quantity of weak learners, learning rates, and maximum tree depths. In order to facilitate this procedure, a methodical methodology is employed, commencing with the determination of parameter boundaries derived from previous research and scholarly sources. This frequently involves constructing a parameter grid that encompasses potential values for every hyperparameter.

The succeeding stage encompasses numerous iterations of model training, wherein different combinations of hyperparameters inside the specified grid are examined. Nevertheless, the effectiveness and dependability of this procedure are contingent upon the manner in which we assess the performance of the model. K-fold cross-validation (CV) assumes a crucial function in this context.

The K-fold cross-validation technique involves dividing the dataset into 'k' folds of equal size. The model is subsequently trained 'k' times, where each fold is utilised as the validation set once, while the remaining 'k-1' folds are employed as training data. K-fold cross-validation (CV) is considered to be of utmost importance for various reasons.

Firstly, the practise of evaluating the model on several data subsets helps mitigate bias in performance estimates. This approach enhances the robustness of the results and reduces their dependence on specific data divisions. Additionally, the utilisation of k-fold cross-validation (CV) offers a more accurate estimation of the variability in the performance of the model. This aids in evaluating the consistency and reliability of the model when applied to diverse subsets of data.

Furthermore, the selection of the value 'k' in k-fold cross-validation has an impact on the determination of the optimal hyperparameters. A higher value of 'k' (for example, 10) provides a more extensive investigation of hyperparameters, but at the expense of increased processing burdens. On the other hand, a reduced value of 'k' (such as 5) exhibits computational efficiency, although it may result in estimations that are comparatively less reliable. Therefore, the selection of 'k' is determined by balancing the available computational resources with the desired level of reliability.

The average of the 'k' rounds of training and validation is commonly used to describe the overall model performance. This metric offers a thorough evaluation of the model's ability to generalise across various subsets of data.

It is generally advised to choose a value of 'k=10' in most search scenarios, since this choice achieves a suitable compromise between computational feasibility and accurate performance estimation. However, the precise value of

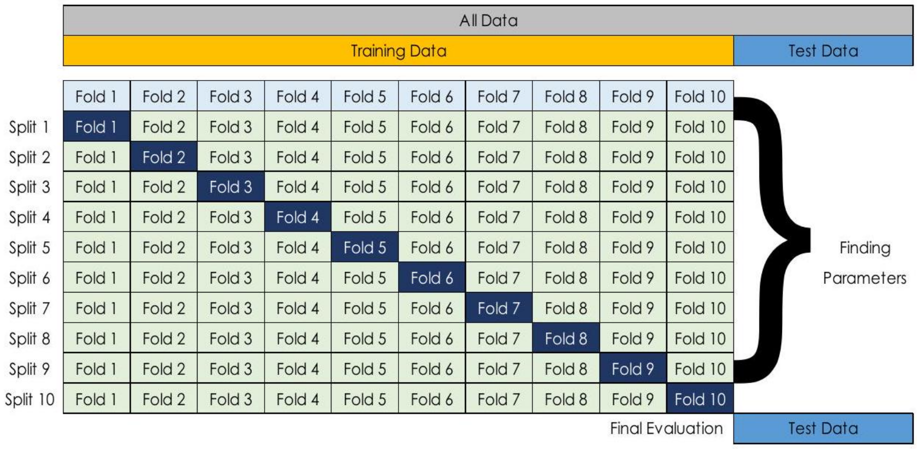

Figure 4: K-fold cross-validation method [109]

the 'k' parameter may differ based on factors like as the size of the dataset, the computational resources at hand, and the desired level of confidence in the obtained results.

A better technique for dealing the bias of the training set's random selection is the k-fold CV. A loop of k rounds is conducted, where the training set is divided into k equal-sized subsets. In each round, one subset is used to test the model and the remaining k-1 subsets are used to train the model. The Random forest algorithm involves optimizing three parameters, which include the total number of trees, the total number of features chosen randomly, and the maximum tree depth.

For XGBoost, There are separate value ranges specified using grids for the number of trees, learning rate, and maximum tree depth. [0: 20: 600], [0.02, 0.05, 0.1, 0.2], and [2, 4, 8, 12, 14]. When the tree number is low, the R^2 score rises fast with it, and once it reaches a specific value, the trend becomes progressively steady. The learning rate has a big impact on performance. In order to achieve the same R^2 score for a training set, a model trained with a smaller learning rate will require more trees than a model trained with a larger learning rate. Increasing the number of trees, however, is not essential to improve the R^2 score for a high learning rate. For instance, when the learning rate is between 0.1 and 0.2, the score drops as the number of trees exceeds between 100 & 200. For learning rates of 0.02 and 0.05, however, the score does not peak until the number of trees exceeds 400, at about 0.8. The greatest tree depth of 8 and 16 yields scores that are quite close. Based on the analysis, the optimal values for the number of trees, learning rate, and maximum depth are 600, 0.1, and 10, respectively.

4.4. Model Evaluation

This study used four different statistical measurement parameters to assess the prediction accuracy of various ensemble learning models. These evaluation parameters compare the accumulated error in the predictions with the

18

actual observations. The statistical parameters used are the coefficient of determination (R-squared), mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). These metrics provide information about the accuracy and precision of the predictions made by the ensemble learning models. These mathematical formulations are defined as follows:

- Coefficient of determination R^2 [110]

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{m} (P_i - T_i)^2}{\sum_{i=1}^{m} \left(P_i - \bar{T}\right)^2} \tag{4}$$

- Mean Absolute Error (MAE) [111]

$$\mathrm{MAE} = \frac{\sum_{i=1}^{m} |P_i - T_i|}{m} \tag{5}$$

- Root Mean Squared Error (RMSE) [112]

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} (P_i - T_i)^2}{m}} \tag{6}$$

- Mean Absolute Percentage Error (MAPE) [113]

$$MAPE = \frac{100\%}{m} \sum_{i=1}^{m} \left| \frac{P_i - T_i}{T_i} \right| \tag{7}$$

where $P_i$and T_i are the predicted and tested values respectively; $\bar{T}$ is the mean value of all the samples in the database.

Clearly, the four metrics provide for a thorough assessment of the model's performance. R^2 , which is better if closer to 1 [114], assesses the linear relationship between predicted values and actual values. The first-order and second-order relative errors (measured by RMSE, MAE, and MAPE) between the predicted value and actual value are better when smaller [115].

5. Results & Discussion

5.1. Comparison between ML Algorithms

Traditional single learning techniques like decision trees (DT), support vector machines (SVM), and artificial neural networks (ANN) are contrasted with the performance of ensemble learning techniques. To ensure a fair comparison, the hyper parameters of the single learning methods are also established through grid search and 10-fold cross-validation.

The authors have compared 7 machine learning models on the testing dataset, i.e. 3 conventional ML models and 4 ensemble learning model. Figure 5 compares the performance of the four ensemble models on the testing dataset. It is clear that compared to single learning models, ensemble learning models exhibit significant improvements. For instance, the worst ensemble learning random forest (RF) model has an R^2 value of 0.906 whereas the greatest single learning DT model has an R-squared value of 0.887. As shown in Table 2 & Table 3, the root mean squared error (RMSE) in prediction of shear strength of single learning models ranges from 63 to 72 kN, but that of the four ensemble models is around 55 kN. The MAE of the single learning model is greater than 40 kN, whereas the MAE of the ensemble models is less than 38 kN. The mean absolute percentage error (MAPE) of the ANN model is higher than 18%, whereas that of the ensemble models is lower than 14%, and that of the XGBoost model is only about 10%.

Table 2: Comparison of R^2, MAE, RMSE & MAPE values between conventional ML models

Models	Sets	R^2	MAE (kN)	RMSE (kN)	MAPE (%)
DT	Training	0.958	22.912	36.416	8.43
	Testing	0.887	42.559	63.145	14.41
SVM	Training	0.980	11.916	24.672	3.64
	Testing	0.852	40.260	72.020	11.76
ANN	Training	0.984	16.526	22.612	6.51
	Testing	0.856	52.050	71.111	18.13

As shown in Table 3, the root mean squared error (RMSE) in prediction of shear strength of ensemble learning models ranges from 47 to 57 kN. The MAE of the ensemble models is less than 38 kN. The mean absolute percentage error (MAPE) of the ensemble models is lower than 14%, and that of the XGBoost model is only about 10%. Overall, the ensemble models and the XGBoost model in particular—perform better than conventional machine learning models.

20

Table 3: Comparison of R^2, MAE, RMSE & MAPE values in ensemble learning models

Models	Sets	R^2	MAE (kN)	RMSE (kN)	MAPE (%)
Random Forest	Training	0.956	20.887	37.327	7.93
	Testing	0.906	38.302	57.477	12.35
AdaBoost	Training	0.970	25.889	30.594	12.38
	Testing	0.919	36.659	53.274	13.16
GBRT	Training	0.999	2.211	3.298	0.85
	Testing	0.910	36.294	56.158	12.47
XGBoost	Training	0.999	0.240	1.450	0.78
	Testing	0.928	29.65	47.76	9.79

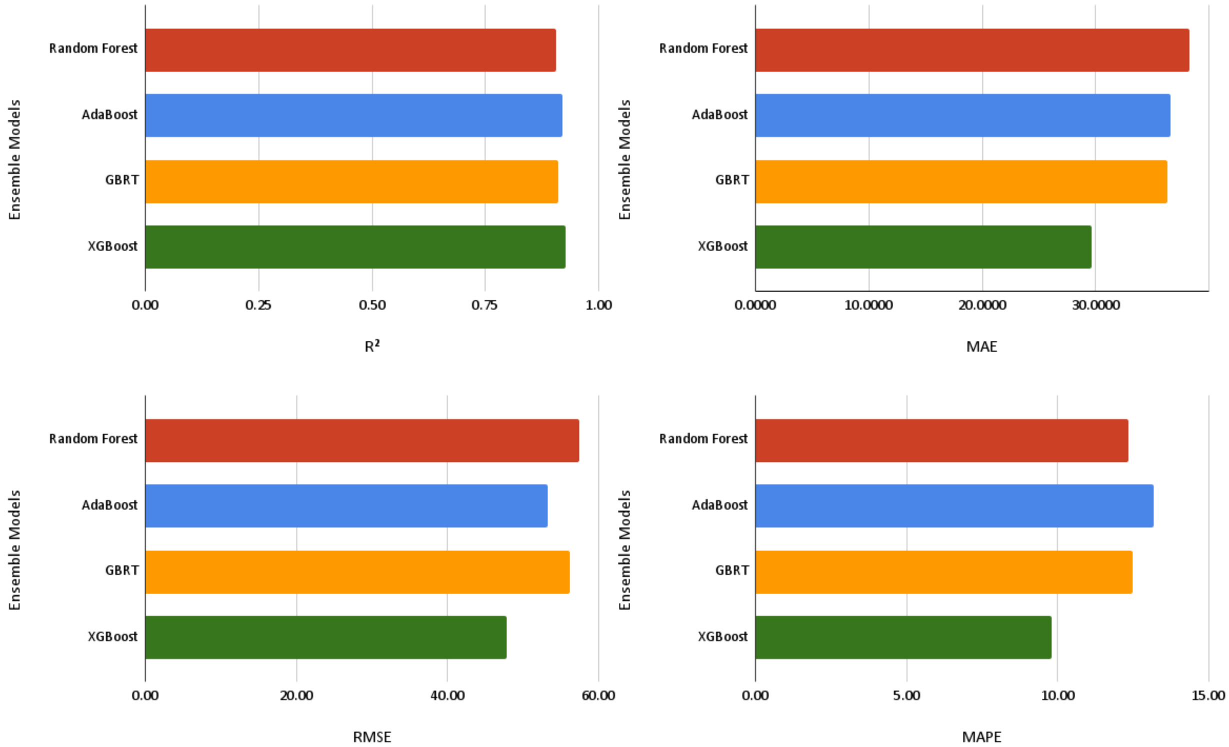

Figure 5: Comparison of R^2, MAE, RMSE & MAPE values between Ensemble ML algorithms.

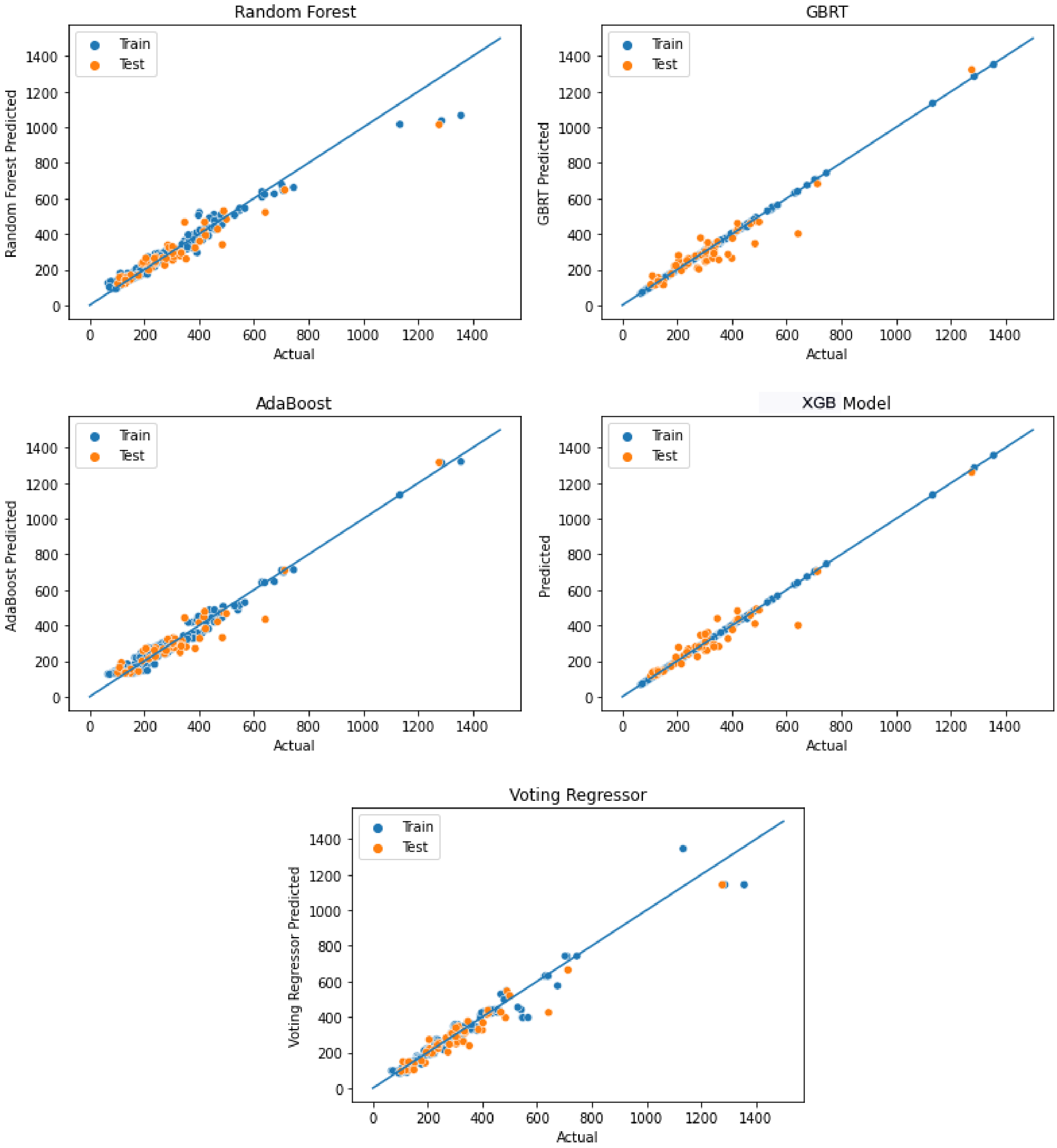

Figure 6: Comparison between ML algorithms with Training and testing data.

The dataset is split in two parts i.e. 80% training set and 20% testing set. The performance of 4 ensemble learning models and voting regressor is shown in Figure 6, where the models are evaluated on the basis of testing dataset compared to given experimental data. The experimental data and the prediction are identical, as shown by the diagonal line (y = x). As can be observed from the scatter plots' near proximity to the diagonal, all four ensemble

22

models generally obtained good results. In case of Voting Regressor, the regressor shows much generalised results compared to all other models.

5.2. *Overview of Mechanics Driven Models*

As opposed to normal beams, deep beams structural analysis is more complex, hence the assumption that the plane section will remain plane before and after bending is invalid because the strain is not distributed linearly. The pressure that is applied will have a greater impact on the stress than the strain. Shear deformation can also be ignored in normal beams, but it cannot be ignored in deep beams where shear is a major factor in failure. Larger depths, when applied in the conventional procedure, cause stress to not be linear in the elastic stage and prevent the ultimate stress from becoming the parabolic shape, which is another important factor in the shear failure of deep beams. European guideline states that a beam is considered to be deep if its effective span to overall depth ratio is less than 3.0 beam [33].

Deep beams are members that are loaded on one face and supported on the other face in accordance with ACI-318 clause 10.7.1 so that compression struts can form between the loads and the supports. In four times the overall member depth or less, or areas where loads are concentrated within a member's depth of twice the support's face [32].

Five expressions for determining the shear strength of RC deep beams are taken from the design codes of China, British (CIRIA), the United States, Canada, and Europe. While the other three are determined based on the strut-and-tie model, the expression of China in Chinese code and British (CIRIA) are semi-empirical semi-analytical equation. The following is a list of the detailed expressions:

- British (CIRIA Guide) [35]

$$V_{u,\text{CIRIA}}, = C_1\left(1 - 0.35\frac{a}{h_o}\right) f_t b h_o + C_2 \sum_{1}^{n} A_1 \frac{y_1}{h_o} \sin^2\alpha$$

where C_1 and C_2 are constants depending on grade of concrete and steel; $f_t = 0.5\sqrt{f_c}$; A_1 = Area of reinforcement; y_1 = depth from the top of the beam to the point where the bar intersects the critical diagonal crack line α= angle between the bar considered and the critical diagonal crack.

- US code: ACI 318 [32]

$$V_{u,\text{ACI}} = 0.85\beta_s f_c' b w_s \sin\theta$$

with

$$w_s = \left[1.85 w_t \cos\theta + \left(l_{pE} + l_{pP}\right)\sin\theta\right]/2$$

$$\theta = \arctan\frac{d_b}{a} \geq 25^\circ$$

where β_s is strut coefficient; θ is the angel between the strut and the longitudinal axis; w_s is the width of the strut; w_t is the height of the nodal region; l_{pE} and l_{pP} are the width of the top loading and bottom supporting plates, respectively; d_b is the distance between the top and bottom nodal region.

- Chinese code: GB50010-2010 [34]

$$V_{u,GB} = \frac{1.75}{\lambda+1} f_t b h_0 + \frac{l_0/h - 2}{3} f_{yv}\frac{A_{sv}}{s_h} h_0 + \frac{5 - l_0/h}{6} f_{yh}\frac{A_{sh}}{s_v} h_0$$

where f_t is the concrete tensile strength; $\lambda = a/h_0$ is the shear spanto-depth ratio. Other variables are the same as defined in Table 1.

- Canadian code: CSA A23.3-04 [36]

$$V_{u,\mathrm{CSA}} = \frac{f_c'}{0.8 + 170\epsilon_1} b w_s \sin\theta$$

with

$$w_s = \left[1.88 w_t \cos\theta + \left(l_{pE} + l_{pP}\right)\sin\theta\right]/2\epsilon_1 = \epsilon_s + (\epsilon_s + 0.002)\cot^2\theta$$

where $\in_s = 0.75\lambda f_c' w_t b / E_s A_s$ is the tensile strain of the tie.

- European code: EN 1-1-1992:2004 [33]

$$V_{u,\mathrm{EU}} = 0.85\beta_s f_c' b w_s \sin\theta$$

5.3. Comparison between Ml algorithms and mechanics-driven models

In this section, a comparison of various statistical metrics with Ml algorithms and mechanics-driven models is drawn.

Table 4, compares the ratio of predicted shear strength to experimentally tested shear strength datasets mean, maximum, minimum, standard deviation & covariance values from 5 codal provisions and the best-performing ensemble learning model i.e. XGBoost algorithm. In this case, if the standard deviation of the dataset is low, it might indicate that the data is consistent and reliable and that any predictions or conclusions drawn from the data are likely to be accurate. On the other hand, a high standard deviation would suggest that the data is more variable and less predictable and that any predictions or conclusions based on the data should be interpreted with caution.

Table 4: Performance comparison between mechanics-driven models and best performing ensemble model

Models	**Predicted-to-test-ratio**				
	Min.	**Max.**	**St.D.**	**Mean**	**COV (%)**
CIRIA	0.29	2.79	0.47	1.23	38.38
ACI 318	0.3	5.27	0.69	1.57	44.25
GB50010-2010	0.36	3.24	0.39	1.43	27.01
CSA 23.3-04	0.59	4.50	0.56	1.56	35.71
EC2	0.44	3.47	0.54	1.42	38.05
XGBoost	0.74	1.60	0.06	1.00	6.38

In general, a low standard deviation is desirable in many applications because it indicates that the data is well-behaved and can be easily analyzed and understood. XGBoost evolution is depicted in Figure 1. It is one of the most superior boosting ensemble learning models because it has both linear model solver and tree learning algorithms. As also shown in Table 4, predicted to test ratios dataset mean value is coming nearly about 1 and the standard deviation is also very low.

25

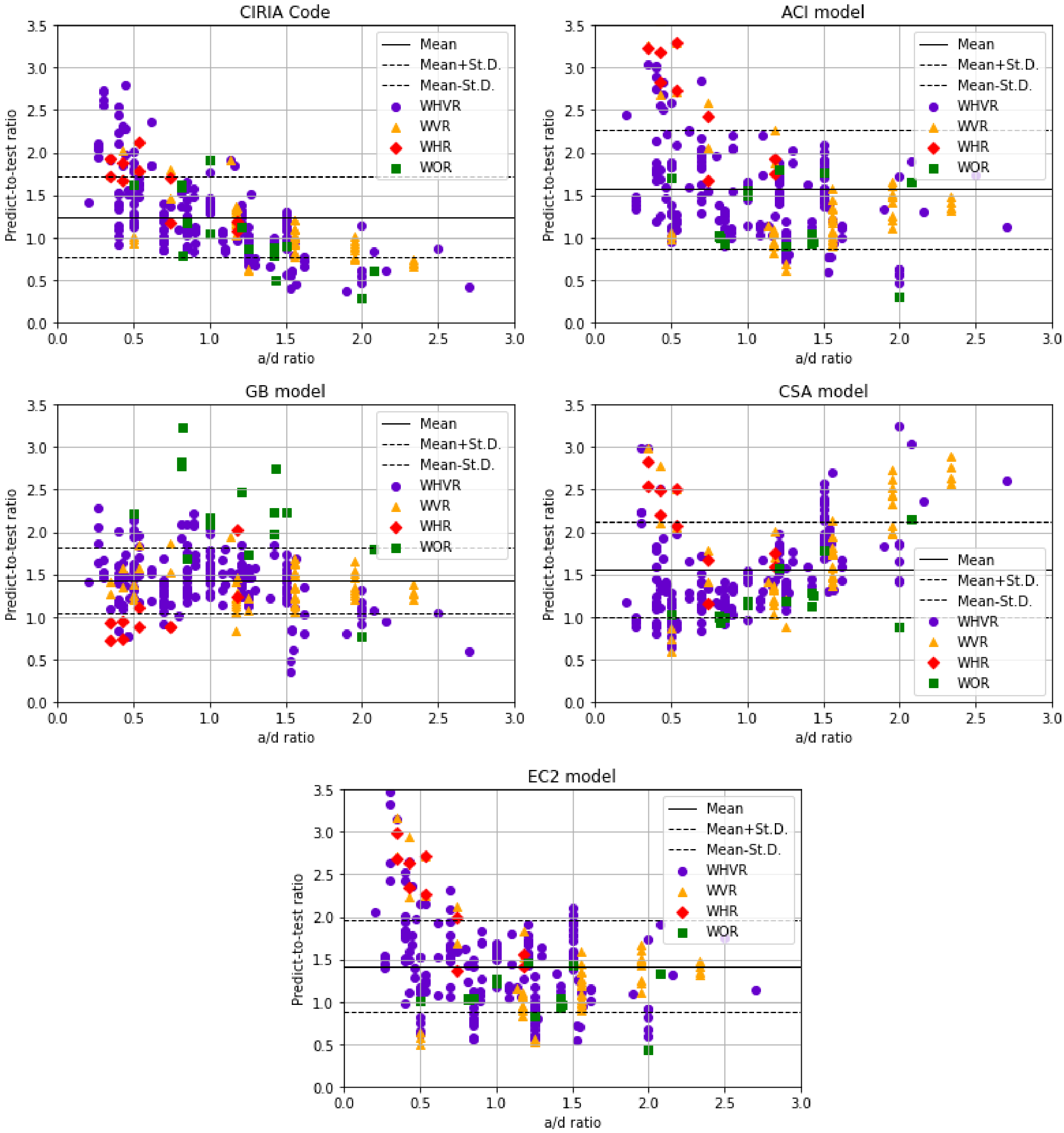

Figure 7: Predicted to test shear strength ratio for different RC beams by mechanics-driven models.

A comparison between predicted to test shear strength ratio plotted against various a/d ratios for mechanics-driven model results (Figure 7) vs. ensemble learning models (Figure 8) clearly depicts better prediction of shear strength on all types of RC deep beams with the XGBoost algorithm. In this study, the authors have also implemented voting regressor over top-performing boosting algorithms to get a better generalised view of ML models as shown in Figure

26

8. Unlike black box ML algorithms, Voting regressors are prominent when it comes to transparency.

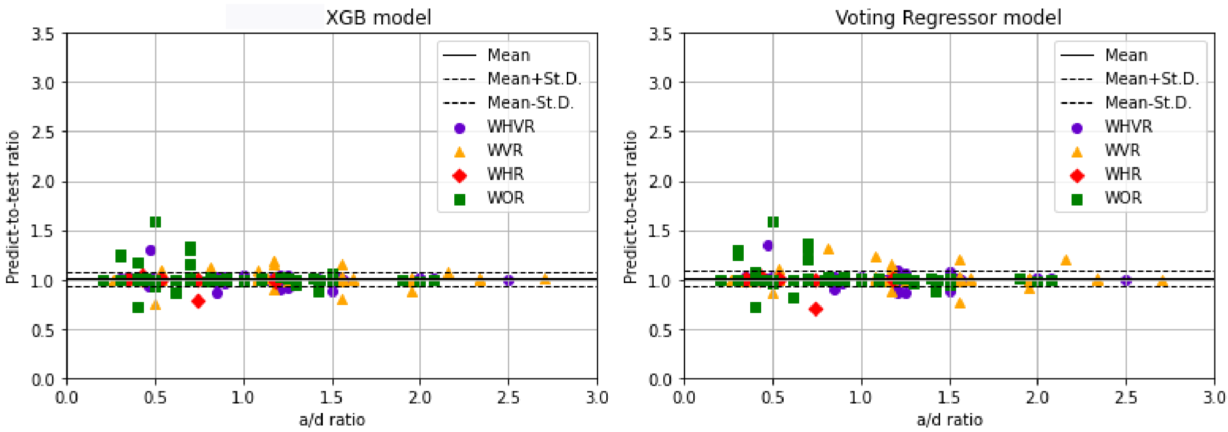

Figure 8: Predicted to test shear strength ratio for different RC beams by XGBoost model and Voting Regressor model.

5.4. SHapley Additive exPlanations for XGBoost

The key idea behind SHapley Additive exPlanations (SHAP) is to evaluate the contribution of each feature in a prediction by considering all possible combinations of features and how they affect the XGBoost model's output. It calculates the average marginal contribution of each feature across all possible feature permutations. This process provides a more robust and balanced measure of feature importance compared to other methods that might suffer from issues like feature interdependence or lack of consistency Lundberg and Lee [37]. SHAP overcomes the major drawback of using ML models which is its black box nature.

The authors have interpreted the SHAP values for all the features (l_0, h , h_0 ,b ,a ,ρ_l ,f_{yl} ,ρ_h , s_h ,f_{yh},ρ_v , s_v, f_{yv} , f'_c , V_u) as shown in Figure 9. Concrete compressive strength f'_c affects the model the highest and horizontal reinforcement strength f_{yv} affects the model the least. Concrete compressive strength f'_c, shear span a, width b & height h affects the models prediction majorly.

In summary, SHAP values provide an interpretable way to understand how each feature affects the model's output. They can help identify which features are driving the model predictions and the direction of their impact. Understanding these feature contributions can be valuable in gaining insights into the XGBoost model's behavior and making data-driven decisions.

5.5. Feature Importance Analysis

Feature importance is important in machine learning models because it helps identify which features are most important for making predictions. This is useful for a number of reasons. First, understanding the relative importance of each feature can help build simpler, more interpretable models. By only using the most important features, it is

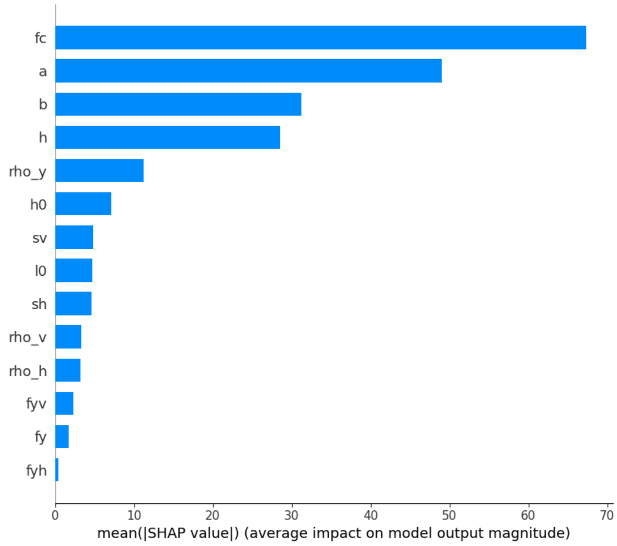

Figure 9: SHAP feature importance

possible to build a model that is easier to understand and explain to others. This can be especially useful in domains where interpretability is important, such as in healthcare or finance. Second, feature importance can help identify features that are redundant or irrelevant. These features can be removed from the model, which can improve its performance by reducing overfitting and increasing generalization. Third, understanding feature importance can help guide feature engineering efforts. By focusing on the most important features, it is possible to create new features that are more predictive and improve the performance of the model [116].

Overall, feature importance is an important tool for understanding and improving machine learning models. It can help identify the most important features, remove redundant or irrelevant features, and build simpler, more interpretable models.

The concrete compressive strength (f_c), standardised to a relative relevance of 100%, was discovered to be the most crucial factor for forecasting the shear strength of RC deep beams, as shown in Figure 10. Shear span (a) and vertical web reinforcement spacing, which have importance values between one-fourth and one-third of the concrete strength, are the second and third most crucial properties, respectively. This makes sense given that these characteristics have a direct impact on the shear mechanism of deep beams. Other characteristics, which account for around 18% of the relevance of shear strength, include section width, shear span-to-depth ratio, and horizontal web reinforcement spacing. Web and longitudinal reinforcement ratios are less important characteristics, with importance values of only about 10% of concrete strength. Other features were found to be of minor significance, with their combined influences being less than 10% of the most significant ones.

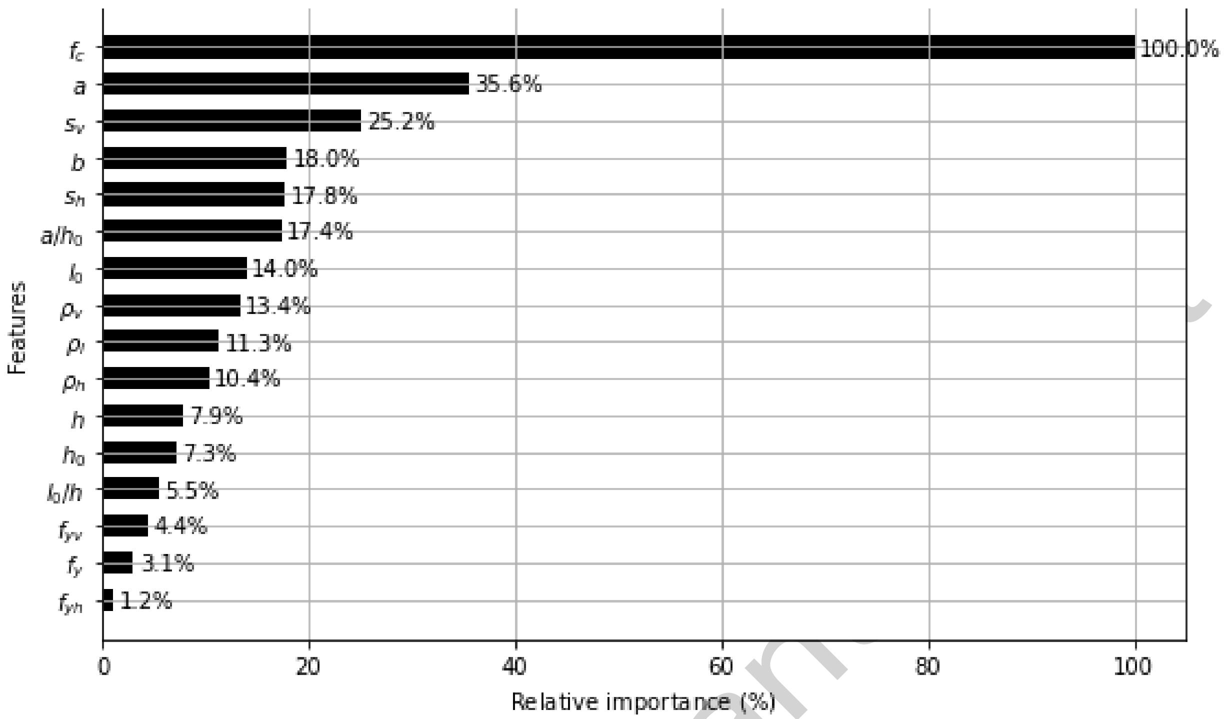

Figure 10: Feature Importance analysis result.

5.6. Conclusion

This paper presents an ML technique-based approach with SHAP to predict the shear strength of RC deep beams. A total of 271 test data samples of RC beams were divided into four groups namely beams without web reinforcements (WOR), beams with horizontal web reinforcements (WHR), beams with vertical web reinforcements (WVR), and beams with both horizontal and vertical reinforcements (WHVR) from the literature were collected and used to train and test the models. The models were trained upon 16 parameters using 3 machine learning and 4 ensemble learning algorithms which were evaluated with each other on parameters coefficient of determination, mean absolute error, root mean squared error and mean absolute percentage error in which XGboost algorithms performed the best. XGboost algorithm was then compared with the mechanics-driven model of CIRIA, United States, Euro Code, Chinese code and Canadian code. According to the results, the following conclusion can be drawn:

- The ML models provide a superior approach to predicting the shear strength of RC deep beams. The approach is robust in nature and can be replicated easily. The approach can be understood with ease rather than the numerical and theoretical derivations of mechanics-driven modelling. The only fundamental requirement is the dataset which can be easily collected and used for long-term structural health monitoring systems.

- The XGBoost algorithm performance the best among ANN, Decision Tree, Support Vector Machine, Random

29

Forest, Gradient Boosting Algorithm and Adaptive boosting algorithm with a coefficient of determination of 0.92 (testing), 0.99 (training) , mean absolute error of 29.65 (testing), 2.47 (training), root mean squared error of 47.76 (training), 1.45(testing) and mean absolute percentage error of 9.79(training), 0.78(testing) which are far superior to the mechanics driven models.

- The hyperparameters for all the models are selected based on their performance in producing the best k-fold cross-validation results. The XGBoost model is found to perform optimally based on multiple iterations in learning rate, number of trees, and maximum depth, with the most suitable parameters being 600 trees, 0.1 learning rate, and a maximum depth of 10.

- The standard deviation, mean and covariance value of predicted to test ratio for XGBoost model were found 0.06, 1.00 and 6.38 respectively in comparison to mechanics driven models British (CIRIA Guide)- 0.47, 1.23, 38.38; United States code: 0.69, 1.57, 44.25; Chinese code: 0.39, 1.43, 27.01, Canadian code: 0.56, 1.56 35.71; and European code: 0.54, 1.42, 38.05. This validates the superiority of the ensemble learning approach, particularly the XGBoost model, over traditional mechanics-driven models, highlighting its potential for accurate shear strength prediction.

- SHapley Additive exPlanations is proposed for XGBoost algorithms results in order to interpret the inner working of the model removing the black box nature of these ML algorithms and feature importance is shown to deduce the parameters which affects the shear strength of RC deep beams the most.

- From SHapley Additive exPlanations and feature importance analysis, the Study concludes that compressive strength of concrete and geometry of the beam are the most influential parameters while properties of steel affects the least while predicting the shear strength of RC deep beams.

5.7. Discussion

This study deduces that the ensemble learning models specifically the XGBoost model is the best choice to predict the shear strength of RC deep beams that predicted to experimentally tested shear strength ratio data has the best mean and least standard deviation as compared to other codal methods. The XGBoost model's predictions of the shear strength ratio for different RC beams indicate that the WHR prediction value is closest to the mean, followed by WVR, WHVR, and WOR, which also show proximity to the mean value in that order. In general the use of ensemble learning for shear strength prediction may lead to a reliance on black box algorithms that are difficult to interpret and understand. This could potentially pose challenges for engineers in comprehending the rationales behind ensemble predictions and evaluating their reliability. Consequently, a lack of trust in the ensemble's predictions might impede its widespread adoption within the construction industry. Meanwhile, the authors have utilized SHapley Additive Explanations (SHAP) to interpret the internal mechanisms of the model and identify correlations among parameters that influence the model predictions. This approach effectively addresses the challenges associated with black box algorithms.

Predicting the shear strength of reinforced concrete (RC) deep beams using ensemble learning can have several implications and potential problems. One potential implication is that the use of ensemble learning for shear strength prediction could improve the accuracy of structural design in the construction industry. By combining the predictions of multiple models, ensemble learning can provide more reliable estimates of shear strength, which can help engineers design safer and more efficient structures. This could ultimately lead to a reduction in structural failures and improve the safety of buildings and other infrastructure [117].

However, there are also potential problems associated with the use of ensemble learning for shear strength prediction. One potential problem is that the accuracy of ensemble learning models depends on the quality and diversity of the individual models that are combined. If the models used in the ensemble are not sufficiently diverse or are based on limited or biased data, the predictions of the ensemble may not be accurate. This could lead to incorrect design decisions and potentially unsafe structures [118, 119].

When it comes to drawing direct comparisons between different studies in the literature on the prediction of shear strength in RC deep beams can be challenging for several reasons. One major obstacle is the variation in the datasets used across different studies. Each study may utilize different experimental data or numerical simulations, resulting in disparities in the dataset size, composition, and quality. This variation can significantly impact the performance and reliability of the predictive models. Moreover, the studies often involve a wide range of parameters affecting shear strength prediction, such as the concrete mix design, steel reinforcement, beam geometry, loading conditions, and boundary conditions. The differences in these parameters among studies can lead to divergent outcomes and hinder the establishment of a consistent comparison framework [67, 66, 78, 77, 65, 103, 120, 121, 122, 123, 124, 125, 126, 127]. Furthermore, researchers adopt various methodologies to solve the problem of shear strength prediction in RC deep beams. These methods may include analytical approaches, experimental investigations, empirical equations, and machine learning techniques. Each method possesses its unique assumptions, limitations, and uncertainties, making it challenging to directly compare their outcomes. Given these variations in datasets, parameters, and methodologies, it becomes impractical to draw straightforward and reasonable comparisons between the literature.

Overall, the use of ensemble learning for predicting the shear strength of RC deep beams has the potential to improve the accuracy and efficiency of structural design. However, it is important to carefully consider the potential problems and challenges associated with this approach and to address them in order to ensure that it is used safely and effectively in the future.

Statements & Declarations

Acknowledgement

The authors acknowledge the invaluable contributions of the reviewers to the manuscript, as their feedback and comments played a crucial role in improving the content of the paper.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Data Availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Achyut Tiwari: conceptualisation, methodology, software, validation, investigation, writing the original draft. Ashok Kumar Gupta: conceptualisation, methodology, Formal Analysis, supervision, review and editing. Tanmay Gupta: Comparison with mechanics driven models of CIRIA, the US, Canada, China & European Union, review and editing.

Abbreviation

AI	Artificial Intelligence
ML	Machine learning
RC	Reinforced concrete
ACI	American Concrete Institute
DT	Decision Tree
SVM	Support Vector Machines
ANN	Artificial Neural Networks
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
GBRT	Gradient Boosting Regression Tree
RF	Random Forest
SHAP	Shapley Additive Explanations
SHM	Structural Health Monitoring
EML	Explainable Machine Learning
ACI	American Concrete Institute
WOR	Without Web Reinforcements
WHR	Horizontal Web Reinforcement
WVR	Vertical Web Reinforcements
WHVR	Both Horizontal and Vertical Web Reinforcement
TENN	Transfer Ensemble Neural Network
l_0	beam span
h	height
h_0	effective height
b	width
a	span
ρ_l	reinforcement ratio
f_{yl}	reinforcement strength
ρ_h	horizontal reinforcement ratio
s_h	horizontal reinforcement spacing
f_{yh}	horizontal reinforcement strength
ρ_v	vertical reinforcement ratio
s_v	vertical reinforcement spacing
f_{yv}	vertical reinforcement strength
f'_c	Concrete Strength
V_u	Shear Strength

33

References

[1] A. T. G. Tapeh, M. Naser, Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices, Archives of Computational Methods in Engineering 30 (2023) 115–159.

[2] B. Ma, K. Terayama, S. Matsumoto, Y. Isaka, Y. Sasakura, H. Iwata, M. Araki, Y. Okuno, Structure-based de novo molecular generator combined with artificial intelligence and docking simulations, Journal of Chemical Information and Modeling 61 (2021) 3304–3313.

[3] A. Tiwari, A. Chugh, A. Sharma, Chapter 7 - uses of artificial intelligence with human-computer interaction in psychology, in: Innovations in Artificial Intelligence and Human-Computer Interaction in the Digital Era, Intelligent Data-Centric Systems, Academic Press, 2023, pp. 173–205.

[4] A. T. G. Tapeh, M. Naser, Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices, Archives of Computational Methods in Engineering 30 (2023) 115–159.

[5] D. M. Dimiduk, E. A. Holm, S. R. Niezgoda, Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering, Integrating Materials and Manufacturing Innovation 7 (2018) 157–172.

[6] H. Salehi, R. Burgueño, Emerging artificial intelligence methods in structural engineering, Engineering structures 171 (2018) 170–189.

[7] S. K. Baduge, S. Thilakarathna, J. S. Perera, M. Arashpour, P. Sharafi, B. Teodosio, A. Shringi, P. Mendis, Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications, Automation in Construction 141 (2022) 104440.

[8] R. Sacks, T. Bloch, M. Katz, R. Yosef, Automating design review with artificial intelligence and bim: State of the art and research framework, in: ASCE International Conference on Computing in Civil Engineering 2019, American Society of Civil Engineers Reston, VA, 2019, pp. 353–360.

[9] M. Hatami, B. Franz, S. Paneru, I. Flood, Using deep learning artificial intelligence to improve foresight method in the optimization of planning and scheduling of construction processes, in: Computing in Civil Engineering 2021, 2022, pp. 1171–1178.

[10] Y. Pan, L. Zhang, Roles of artificial intelligence in construction engineering and management: A critical review and future trends, Automation in Construction 122 (2021) 103517.

[11] C. Palsara, V. Kumar, J. Pal, M. Naresh, Structural health monitoring of asce benchmark building using machine learning algorithms, Asian Journal of Civil Engineering (2023) 1–14.

[12] J. Rahman, K. S. Ahmed, N. I. Khan, K. Islam, S. Mangalathu, Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach, Engineering Structures 233 (2021) 111743.

[13] L. Truong-Hong, R. Lindenbergh, Automatically extracting surfaces of reinforced concrete bridges from terrestrial laser scanning point clouds, Automation in Construction 135 (2022) 104127.

[14] Y. Fujino, D. M. Siringoringo, Y. Ikeda, T. Nagayama, T. Mizutani, Research and implementations of structural monitoring for bridges and buildings in japan, Engineering 5 (2019) 1093–1119.

[15] H. Hao, K. Bi, W. Chen, T. M. Pham, J. Li, Towards next generation design of sustainable, durable, multi-hazard resistant, resilient, and smart civil engineering structures, Engineering Structures 277 (2023) 115477.

[16] L. Jin, T. Wang, X.-a. Jiang, X. Du, Size effect in shear failure of rc beams with stirrups: Simulation and formulation, Engineering Structures 199 (2019) 109573.

[17] W. Zhang, X. Liu, Y. Huang, M.-N. Tong, Reliability-based analysis of the flexural strength of concrete beams reinforced with hybrid bfrp and steel rebars, Archives of Civil and Mechanical Engineering 22 (2022) 1–20.

[18] M. Al-Osta, M. Isa, M. Baluch, M. Rahman, Flexural behavior of reinforced concrete beams strengthened with ultra-high performance fiber reinforced concrete, Construction and Building Materials 134 (2017) 279–296.

[19] A. Tiwari, A. Chugh, A. Sharma, Ensemble framework for cardiovascular disease prediction, Computers in Biology and Medicine (2022) 105624.

[20] C. Zhang, Y. Lu, Study on artificial intelligence: The state of the art and future prospects, Journal of Industrial Information Integration 23 (2021) 100224.

[21] Y. Xu, X. Liu, X. Cao, C. Huang, E. Liu, S. Qian, X. Liu, Y. Wu, F. Dong, C.-W. Qiu, et al., Artificial intelligence: A powerful paradigm for scientific research, The Innovation 2 (2021).

[22] O. E. Gundersen, S. Kjensmo, State of the art: Reproducibility in artificial intelligence, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 32, 2018.

[23] L. Aristodemou, F. Tietze, The state-of-the-art on intellectual property analytics (ipa): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (ip) data, World Patent Information 55 (2018) 37–51.

[24] S. Selvaraj, S. Sivaraman, Prediction model for optimized self-compacting concrete with fly ash using response surface method based on fuzzy classification, Neural Computing and Applications 31 (2019) 1365–1373.

[25] M. Castelli, L. Vanneschi, S. Silva, Prediction of high performance concrete strength using genetic programming with geometric semantic genetic operators, Expert Systems with Applications 40 (2013) 6856–6862.

[26] H. Sun, H. V. Burton, H. Huang, Machine learning applications for building structural design and performance assessment: State-of-the-art review, Journal of Building Engineering 33 (2021) 101816.

[27] J. Naranjo-Pérez, M. Infantes, J. F. Jiménez-Alonso, A. Sáez, A collaborative machine learning-optimization algorithm to improve the finite element model updating of civil engineering structures, Engineering Structures 225 (2020) 111327.

[28] O. R. Abuodeh, J. A. Abdalla, R. A. Hawileh, Assessment of compressive strength of ultra-high performance concrete using deep machine learning techniques, Applied Soft Computing 95 (2020) 106552.

[29] H. Sun, H. V. Burton, H. Huang, Machine learning applications for building structural design and performance assessment: State-of-the-art review, Journal of Building Engineering 33 (2021) 101816.

[30] X. Dong, Z. Yu, W. Cao, Y. Shi, Q. Ma, A survey on ensemble learning, Frontiers of Computer Science 14 (2020) 241–258.

[31] J. Zhang, Y. Sun, G. Li, Y. Wang, J. Sun, J. Li, Machine-learning-assisted shear strength prediction of reinforced concrete beams with and without stirrups, Engineering with Computers (2020) 1–15.

[32] A. Committee, Building code requirements for structural concrete (aci 318-08) and commentary, American Concrete Institute, 2008.

[33] D. Bethlehem, The european union, in: National Implementation of United Nations Sanctions, Brill Nijhoff, 2004, pp. 123–165.

[34] G. C. Standard, 50010-2010, Code for design of concrete structures (2002).

[35] O. Arup, Partners, The design of deep beams in reinforced concrete, CIRIA Guide 144 (1977).

[36] D. Darwin, C. W. Dolan, A. H. Nilson, Design of concrete structures, volume 2, McGraw-Hill Education New York, NY, USA:, 2016.

[37] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Advances in neural information processing systems 30 (2017).

[38] Z. He, W. Li, H. Salehi, H. Zhang, H. Zhou, P. Jiao, Integrated structural health monitoring in bridge engineering, Automation in Construction 136 (2022) 104168.

[39] M. Gordan, S.-R. Sabbagh-Yazdi, Z. Ismail, K. Ghaedi, P. Carroll, D. McCrum, B. Samali, State-of-the-art review on advancements of data mining in structural health monitoring, Measurement 193 (2022) 110939.

[40] H. Sarmadi, K.-V. Yuen, Structural health monitoring by a novel probabilistic machine learning method based on extreme value theory and mixture quantile modeling, Mechanical Systems and Signal Processing 173 (2022) 109049.

[41] V. Dubourg, B. Sudret, J.-M. Bourinet, Reliability-based design optimization using kriging surrogates and subset simulation, Structural and

35

Multidisciplinary Optimization 44 (2011) 673–690.

[42] V. Dubourg, B. Sudret, F. Deheeger, Metamodel-based importance sampling for structural reliability analysis, Probabilistic Engineering Mechanics 33 (2013) 47–57.

[43] P. Kersaudy, B. Sudret, N. Varsier, O. Picon, J. Wiart, A new surrogate modeling technique combining kriging and polynomial chaos expansions–application to uncertainty analysis in computational dosimetry, Journal of Computational Physics 286 (2015) 103–117.

[44] H. Dai, Z. Cao, A wavelet support vector machine-based neural network metamodel for structural reliability assessment, Computer-Aided Civil and Infrastructure Engineering 32 (2017) 344–357.

[45] N. D. Lagaros, M. Fragiadakis, Fragility assessment of steel frames using neural networks, Earthquake Spectra 23 (2007) 735–752.

[46] N. D. Lagaros, Y. Tsompanakis, P. N. Psarropoulos, E. C. Georgopoulos, Computationally efficient seismic fragility analysis of geostructures, Computers & Structures 87 (2009) 1195–1203.

[47] S. Mangalathu, G. Heo, J.-S. Jeon, Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes, Engineering Structures 162 (2018) 166–176.

[48] Z. Wang, N. Pedroni, I. Zentner, E. Zio, Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment, Engineering Structures 162 (2018) 213–225.

[49] F. Jalayer, H. Ebrahimian, A. Miano, G. Manfredi, H. Sezen, Analytical fragility assessment using unscaled ground motion records, Earthquake Engineering & Structural Dynamics 46 (2017) 2639–2663.

[50] S. Mangalathu, J.-S. Jeon, R. DesRoches, Critical uncertainty parameters influencing seismic performance of bridges using lasso regression, Earthquake Engineering & Structural Dynamics 47 (2018) 784–801.

[51] S. Mahmoudi, L. Chouinard, Seismic fragility assessment of highway bridges using support vector machines, Bulletin of Earthquake Engineering 14 (2016) 1571–1587.

[52] F. Leonhardt, R. Walter, Deep beams, in: Bulletin 178, Deutcher Ausschuss fur Stahlbeton, Berlin, 1966.

[53] K. S. Ismail, M. Guadagnini, K. Pilakoutas, Strut-and-tie modeling of reinforced concrete deep beams, Journal of Structural Engineering 144 (2018).

[54] B. I. Mihaylov, E. C. Bentz, M. P. Collins, Behavior of large deep beams subjected to monotonic and reversed cyclic shear., ACI Structural Journal 107 (2010).

[55] Z. M. Yaseen, R. C. Deo, A. Hilal, A. M. Abd, L. C. Bueno, S. Salcedo-Sanz, M. L. Nehdi, Predicting compressive strength of lightweight foamed concrete using extreme learning machine model, Advances in Engineering Software 115 (2018) 112–125.

[56] G. Zhang, Z. H. Ali, M. S. Aldlemy, M. H. Mussa, S. Q. Salih, M. M. Hameed, Z. S. Al-Khafaji, Z. M. Yaseen, Reinforced concrete deep beam shear strength capacity modelling using an integrative bio-inspired algorithm with an artificial intelligence model, Engineering with Computers (2020) 1–14.

[57] R. Biswas, E. Li, N. Zhang, S. Kumar, B. Rai, J. Zhou, Development of hybrid models using metaheuristic optimization techniques to predict the carbonation depth of fly ash concrete, Construction and Building Materials 346 (2022) 128483.

[58] P. G. Asteris, D. J. Armaghani, G. D. Hatzigeorgiou, C. G. Karayannis, K. Pilakoutas, Predicting the shear strength of reinforced concrete beams using artificial neural networks, Computers and Concrete, An International Journal 24 (2019) 469–488.

[59] D. J. Armaghani, G. D. Hatzigeorgiou, C. Karamani, A. Skentou, I. Zoumpoulaki, P. G. Asteris, Soft computing-based techniques for concrete beams shear strength, Procedia Structural Integrity 17 (2019) 924–933.

[60] S. A. Basilio, L. Goliatt, Gradient boosting hybridized with exponential natural evolution strategies for estimating the strength of geopolymer self-compacting concrete, Knowledge-Based Engineering and Sciences 3 (2022) 1–16.

[61] M. Z. Esteghamati, M. M. Flint, Developing data-driven surrogate models for holistic performance-based assessment of mid-rise rc frame

buildings at early design, Engineering Structures 245 (2021) 112971.

[62] T. G. Wakjira, A. Al-Hamrani, U. Ebead, W. Alnahhal, Shear capacity prediction of frp-rc beams using single and ensenble explainable machine learning models, Composite Structures 287 (2022) 115381.

[63] B. Salman, M. M. Kadhum, Predicting of load carrying capacity of reactive powder concrete and normal strength concrete column specimens using artificial neural network, Knowledge-Based Engineering and Sciences 3 (2022) 45–53.

[64] C. R. Farrar, K. Worden, Structural health monitoring: a machine learning perspective, John Wiley & Sons, 2012.

[65] M. S. Sandeep, K. Tiprak, S. Kaewunruen, P. Pheinsusom, W. Pansuk, Shear strength prediction of reinforced concrete beams using machine learning, in: Structures, volume 47, Elsevier, 2023, pp. 1196–1211.

[66] B. Fu, D.-C. Feng, A machine learning-based time-dependent shear strength model for corroded reinforced concrete beams, Journal of Building Engineering 36 (2021) 102118.

[67] J.-S. Chou, T.-P.-T. Pham, T.-K. Nguyen, A.-D. Pham, N.-T. Ngo, Shear strength prediction of reinforced concrete beams by baseline, ensemble, and hybrid machine learning models, Soft Computing 24 (2020) 3393–3411.

[68] U. Naik, S. Kute, Span-to-depth ratio effect on shear strength of steel fiber-reinforced high-strength concrete deep beams using ann model, International Journal of Advanced Structural Engineering 5 (2013) 29.

[69] N. Concha, J. R. Aratan, E. M. Derigay, J. M. Martin, R. E. Taneo, A hybrid neuro-swarm model for shear strength of steel fiber reinforced concrete deep beams, Journal of Building Engineering (2023) 107340.

[70] M. I. Vamdewalle, F. Mortelmans, Shear capacity of steel fiber high-strength concrete beams, Special Publication 149 (1994) 227–242.

[71] S. Al-Ta'an, J. Al-Feel, Evaluation of shear strength of fibre-reinforced concrete beams, Cement and Concrete Composites 12 (1990) 87–94.

[72] A. Sharma, Shear strength of steel fiber reinforced concrete beams, in: Journal Proceedings, volume 83, 1986, pp. 624–628.

[73] M. Khuntia, B. Stojadinovic, S. C. Goel, Shear strength of normal and high-strength fiber reinforced concrete beams without stirrups, Structural Journal 96 (1999) 282–289.

[74] S.-H. Cho, Y.-I. Kim, Effects of steel fibers on short beams loaded in shear, Structural journal 100 (2003) 765–774.

[75] H. Pak, S. Leach, S. H. Yoon, S. G. Paal, A knowledge transfer enhanced ensemble approach to predict the shear capacity of reinforced concrete deep beams without stirrups, Computer-Aided Civil and Infrastructure Engineering (2023).

[76] G. Almasabha, K. F. Al-Shboul, A. Shehadeh, O. Alshboul, Machine learning-based models for predicting the shear strength of synthetic fiber reinforced concrete beams without stirrups, in: Structures, volume 52, Elsevier, 2023, pp. 299–311.

[77] H.-B. Ly, T.-T. Le, H.-L. T. Vu, V. Q. Tran, L. M. Le, B. T. Pham, Computational hybrid machine learning based prediction of shear capacity for steel fiber reinforced concrete beams, Sustainability 12 (2020) 2709.

[78] O. B. Olalusi, P. O. Awoyera, Shear capacity prediction of slender reinforced concrete structures with steel fibers using machine learning, Engineering Structures 227 (2021) 111470.

[79] K. M. Hossain, L. R. Gladson, M. S. Anwar, Modeling shear strength of medium-to ultra-high-strength steel fiber-reinforced concrete beams using artificial neural network, Neural Computing and Applications 28 (2017) 1119–1130.

[80] H. S. Marie, K. Abu El-hassan, E. M. Almetwally, M. A. El-Mandouh, Joint shear strength prediction of beam-column connections using machine learning via experimental results, Case Studies in Construction Materials 17 (2022) e01463.

[81] T. G. Wakjira, U. Ebead, M. S. Alam, Machine learning-based shear capacity prediction and reliability analysis of shear-critical rc beams strengthened with inorganic composites, Case Studies in Construction Materials 16 (2022) e01008.

[82] T. Liu, Z. Wang, Z. Long, J. Zeng, J. Wang, J. Zhang, Direct shear strength prediction for precast concrete joints using the machine learning method, Journal of Bridge Engineering 27 (2022) 04022026.

[83] T. G. Dietterich, Ensemble methods in machine learning, in: Multiple Classifier Systems: First International Workshop, MCS 2000 Cagliari,

37

Italy, June 21–23, 2000 Proceedings 1, Springer, 2000, pp. 1–15.

[84] Y. Freund, R. Iyer, R. E. Schapire, Y. Singer, An efficient boosting algorithm for combining preferences, Journal of machine learning research 4 (2003) 933–969.

[85] Y. Freund, R. E. Schapire, et al., Experiments with a new boosting algorithm, in: icml, volume 96, Citeseer, 1996, pp. 148–156.

[86] B. Yegnanarayana, Artificial neural networks, PHI Learning Pvt. Ltd., 2009.

[87] O. Z. Maimon, L. Rokach, Data mining with decision trees: theory and applications, volume 81, World scientific, 2014.

[88] L. H. Hamel, Knowledge discovery with support vector machines, John Wiley & Sons, 2011.

[89] A. S. Bakouregui, H. M. Mohamed, A. Yahia, B. Benmokrane, Explainable extreme gradient boosting tree-based prediction of load-carrying capacity of frp-rc columns, Engineering Structures 245 (2021) 112836.

[90] M. Belgiu, L. Drăguţ, Random forest in remote sensing: A review of applications and future directions, ISPRS journal of photogrammetry and remote sensing 114 (2016) 24–31.

[91] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, Frontiers in neurorobotics 7 (2013) 21.

[92] Q. Wu, C. J. Burges, K. M. Svore, J. Gao, Adapting boosting for information retrieval measures, Information Retrieval 13 (2010) 254–270.

[93] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, et al., Xgboost: extreme gradient boosting, R package version 0.4-2 1 (2015) 1–4.

[94] Y. Jiang, G. Tong, H. Yin, N. Xiong, A pedestrian detection method based on genetic algorithm for optimize xgboost training parameters, IEEE Access 7 (2019) 118310–118321.

[95] D. Wu, P. Guo, P. Wang, Malware detection based on cascading xgboost and cost sensitive, in: 2020 International Conference on Computer Communication and Network Security (CCNS), IEEE, 2020, pp. 201–205.

[96] M. Kivrak, E. Guldogan, C. Colak, Prediction of death status on the course of treatment in sars-cov-2 patients with deep learning and machine learning methods, Computer methods and programs in biomedicine 201 (2021) 105951.

[97] J. Hancock, T. M. Khoshgoftaar, Performance of catboost and xgboost in medicare fraud detection, in: 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2020, pp. 572–579.

[98] R. Chen, Y. Ma, N. Chen, D. Lee, W. Wang, Cephalometric landmark detection by attentive feature pyramid fusion and regression-voting, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 873–881.

[99] D.-C. Feng, W.-J. Wang, S. Mangalathu, G. Hu, T. Wu, Implementing ensemble learning methods to predict the shear strength of rc deep beams with/without web reinforcements, Engineering Structures 235 (2021) 111979.

[100] K. Smith, A. Vantsiotis, Shear strength of deep beams, in: Journal Proceedings, volume 79, 1982, pp. 201–213.

[101] F.-K. Kong, P. J. Robins, D. F. Cole, Web reinforcement effects on deep beams, in: Journal Proceedings, volume 67, 1970, pp. 1010–1018.

[102] A. P. Clark, Diagonal tension in reinforced concrete beams, in: Journal Proceedings, volume 48, 1951, pp. 145–156.

[103] J.-K. Oh, S.-W. Shin, Shear strength of reinforced high-strength concrete deep beams, Structural Journal 98 (2001) 164–173.

[104] G. Aguilar, A. B. Matamoros, G. Parra-Montesinos, J. A. Ramírez, J. K. Wight, Experimental evaluation of design procedures for shear strength of deep reinfoced concrete beams, American Concrete Institute, 2002.

[105] C. G. Quintero-Febres, G. Parra-Montesinos, J. K. Wight, Strength of struts in deep concrete members designed using strut-and-tie method, ACI Materials Journal 103 (2006) 577.

[106] K.-H. Tan, F.-K. Kong, S. Teng, L. Guan, High-strength concrete deep beams with effective span and shear span variations, Structural Journal 92 (1995) 395–405.

[107] V. Ramakrishnan, Y. Ananthanarayana, Ultimate strength of deep beams in shear, in: Journal Proceedings, volume 65, 1968, pp. 87–98.

[108] G. Shaoxi, The shear strength capability of reinforced concrete deep beam under symmetric concentrated loads, Journal of Zhengzhou

Technology Institute 1 (1982) 52–68.

[109] V. Rathakrishnan, S. Bt. Beddu, A. N. Ahmed, Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms, Scientific Reports 12 (2022) 9539.

[110] A. Di Bucchianico, Coefficient of determination (r 2), Encyclopedia of statistics in quality and reliability 1 (2008).

[111] P. Maragos, Morphological correlation and mean absolute error criteria, in: International Conference on Acoustics, Speech, and Signal Processing,, IEEE, 1989, pp. 1568–1571.

[112] T. Chai, R. R. Draxler, Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature, Geoscientific model development 7 (2014) 1247–1250.

[113] A. De Myttenaere, B. Golden, B. Le Grand, F. Rossi, Using the mean absolute percentage error for regression models, in: Proceedings, Presses universitaires de Louvain, 2015, p. 113.

[114] O. Renaud, M.-P. Victoria-Feser, A robust coefficient of determination for regression, Journal of Statistical Planning and Inference 140 (2010) 1852–1862.

[115] G. Hackeling, Mastering Machine Learning with scikit-learn, Packt Publishing Ltd, 2017.

[116] G. König, C. Molnar, B. Bischl, M. Grosse-Wentrup, Relative feature importance, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021, pp. 9318–9325.

[117] H. Fathipour-Azar, Stacking ensemble machine learning-based shear strength model for rock discontinuity, Geotechnical and Geological Engineering 40 (2022) 3091–3106.

[118] M. H. Alobaidi, F. Chebana, M. A. Meguid, Robust ensemble learning framework for day-ahead forecasting of household based energy consumption, Applied energy 212 (2018) 997–1012.

[119] G. Seni, J. F. Elder, Ensemble methods in data mining: improving accuracy through combining predictions, Synthesis lectures on data mining and knowledge discovery 2 (2010) 1–126.

[120] Y.-K. Kwak, M. O. Eberhard, W.-S. Kim, J. Kim, Shear strength of steel fiber-reinforced concrete beams without stirrups, ACI Structural journal 99 (2002) 530–538.

[121] J. Rahman, K. S. Ahmed, N. I. Khan, K. Islam, S. Mangalathu, Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach, Engineering Structures 233 (2021) 111743.

[122] J. Zhang, Y. Sun, G. Li, Y. Wang, J. Sun, J. Li, Machine-learning-assisted shear strength prediction of reinforced concrete beams with and without stirrups, Engineering with Computers (2020) 1–15.

[123] T. G. Wakjira, A. Al-Hamrani, U. Ebead, W. Alnahhal, Shear capacity prediction of frp-rc beams using single and ensenble explainable machine learning models, Composite Structures 287 (2022) 115381.

[124] O. R. Abuodeh, J. A. Abdalla, R. A. Hawileh, Prediction of shear strength and behavior of rc beams strengthened with externally bonded frp sheets using machine learning techniques, Composite Structures 234 (2020) 111698.

[125] S. Mangalathu, J.-S. Jeon, Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques, Engineering Structures 160 (2018) 85–94.

[126] J.-S. Chou, T.-P.-T. Pham, T.-K. Nguyen, A.-D. Pham, N.-T. Ngo, Shear strength prediction of reinforced concrete beams by baseline, ensemble, and hybrid machine learning models, Soft Computing 24 (2020) 3393–3411.

[127] D. Prayogo, M.-Y. Cheng, Y.-W. Wu, D.-H. Tran, Combining machine learning models via adaptive ensemble weighting for prediction of shear capacity of reinforced-concrete deep beams, Engineering with Computers 36 (2020) 1135–1153.