# Reliability-based design shear resistance of headed studs in solid slabs predicted by machine learning models

Vitaliy V. Degtyarev[1*] and Stephen J. Hicks[2]

[1*]New Millennium Building Systems, LLC, 3700 Forest Dr. Suite 501, Columbia, 29204, SC, United States of America.
[2]School of Engineering, University of Warwick, Coventry, CV4 7AL, United Kingdom.

*Corresponding author(s). E-mail(s): vitaliy.degtyarev@newmill.com, vitdegtyarev@yahoo.com;
Contributing authors: stephen.j.hicks@warwick.ac.uk;

**Abstract**

The economical and reliable design of steel-concrete composite structures relies on accurate predictions of the resistance of headed studs transferring the longitudinal shear forces between the two materials. The existing mechanics-based or empirical design equations do not always produce accurate and safe predictions of the stud shear resistance. This study presents the evaluation of nine machine learning (ML) algorithms and the development of optimized ML models for predicting the stud resistance. The ML models were trained and tested using databases of push-out test results for studs in both normal weight and lightweight concrete. The reliability of ML model predictions was evaluated in accordance with European and US design practices. Reduction coefficients required for the ML models to satisfy the Eurocode reliability requirements for the design shear resistance were determined. Resistance factors used in US design practice were also obtained. The developed ML models were interpreted using the SHapley Additive exPlanations (SHAP) method. Predictions by the ML models were compared with those by the existing descriptive equations, which demonstrated a higher accuracy for the ML models. A web application that conveniently provides predictions of the nominal and design stud shear resistances by the developed ML models in accordance with both European and US design practices was created and deployed to the cloud.

**Keywords:** Headed studs, Shear resistance, Steel-concrete composite structures, Reliability, Machine learning

# 1 Introduction

The performance of steel-concrete composite structures depends on mechanical connectors transferring the longitudinal shear and tension forces between the steel and concrete. Headed studs welded to the steel components and embedded within the concrete are the most popular type of shear transfer devices used in construction due to their installation speed and performance reliability. Many researchers have experimentally studied the resistance of headed studs [1, 2] and proposed several design models based on regression analyses of the experimental data available at the time. Some models have been adopted in national and international design standards. The Ollgaard et al. [3] model based on the analysis of 48 push-out test results has been adapted by AISC 360 [4] and Eurocode 4 (EC4) [5, 6]. Later comparisons with more extensive test data showed that the design models presented in the standards and the literature lack the accuracy and/or do not satisfy the target reliability requirements [7–9].

Degtyarev et al. [10] applied the symbolic regression with genetic programming (GPSR) algorithm to the test data published in [1, 2] to derive improved design models. New design models in the form of simple descriptive equations for computing the nominal shear resistance of headed studs in normal weight (NWC) and lightweight concrete (LWC) slabs were developed. The nominal resistance models were subsequently calibrated to meet the reliability requirements of the Eurocodes [5, 6, 11]. The nominal and calibrated design resistance models demonstrated an improved prediction accuracy compared with the existing design models.

Numerous studies describing successful applications of rapidly developing machine learning (ML) techniques to structural engineering problems [12–15] indicate that the ML approach might improve the accuracy of the stud resistance predictions while also satisfying the reliability requirements of building codes. ML models represent pre-trained computer algorithms that humans cannot easily comprehend. They are often criticized for lacking transparency. However, the ML methods are based on well-established mathematical algorithms described in the literature and have been successfully tested on many problems in different industries [16–18].

Several studies describe the application of ML techniques for estimating the shear resistance of headed studs in solid concrete slabs [19–22]. Abambres and He [19] developed an artificial neural network (ANN) using a database of 234 push-out test results. The tensile strength and diameter of the studs and the concrete compressive strength were input parameters of the ANN, which outperformed the existing code-based equations. Avci-Karatas [20] proposed models for predicting the stud shear resistance based on the concepts of minimax probability machine regression (MPMR) and extreme machine learning (EML). The same test database and input parameters as those described in [19] were used for the model development. The MPMR and EML models demonstrated excellent prediction accuracy, which exceeded the accuracy of the models from design standards.

More recently, Wang et al. [22] presented a light gradient boosting machine (LightGBM) model to predict the shear resistance of headed studs in solid concrete slabs. The models trained on an extensive database of test results with 1092 samples outperformed the existing descriptive equations. The model hyperparameters were automatically optimized by employing the sequential model-based optimization method. The model's relative feature importance and feature dependence were evaluated using the SHapley Additive exPlanations (SHAP) method. The authors also created and deployed a web application based on the developed model to the cloud. It should be noted that many of the tests included in the database considered by Wang et al. [22] did not satisfy the rules for the standard push-out test specimen in Eurocode 4 [5, 6], and, therefore, were not considered in the presented study.

Setvati and Hicks [21] evaluated the performance of six ML algorithms for forecasting the stud shear resistance in NWC slabs. The algorithms implemented in MATLAB included linear regression, decision tree (DT), support vector machine (SVM), Gaussian Process Regression (GPR), ANN, and bagged ensemble trees (BET). The models, trained and optimized using 242 test results, outperformed the existing design models, with the SVM model being the most accurate.

The present paper extends the Setvati and Hicks' study [21] as follows: 1) ML models were developed and optimized for LWC slabs, in addition to NWC slabs; 2) six additional popular ML algorithms were evaluated, including k-nearest neighbors (KNN), random forest (RF), gradient boosting regressor (GBR), extreme gradient boosting (XGBoost), LightGBM, and gradient boosting with categorical features support (CatBoost); overall, nine ML models were considered, including DT, SVM, and ANN previously evaluated in [21]; 3) from comparing the performance of the developed ML models with test data, the models were updated according to European and US design practice, to ensure that the target reliability index was delivered; and 4) a user-friendly web application for predicting the stud shear resistance with the developed models was created and deployed to the cloud. The main objective of this work was to propose ML models that accurately predict the shear resistance of headed studs in NWC and LWC slabs, simultaneously satisfying the Eurocodes' reliability requirements and having resistance factors established according to the US design rules.

The novelty of the present study consists in accurate ML models for predicting the stud shear resistance in NWC and LWC slabs, which were calibrated to meet the reliability requirements of Eurocode 4 [5, 6] and US design practice and interpreted using the SHAP method. The models were trained using the databases with the results of the tests carefully selected to comply with the Eurocode 4 push-out test requirements. The created web application based on the calibrated ML models gives practitioners and researchers a new tool for rapid evaluations of the nominal and design values of the stud shear resistance and performing parametric studies.

To the authors' knowledge, this is the first research paper describing the reliability evaluation of ML models for the stud shear resistance according to the Eurocodes. Hitherto, this reliability evaluation has been confined to descriptive equations, such as those adopted by EC4 [5, 6]. Previously, the reliability of an ANN model for predicting the buckling resistance of steel hollow sections in accordance with Annex D of EN 1990 [11] was evaluated by Toffolon et al. [23]. Zarringol et al. [24] and Wakjira et al. [25] employed Monte Carlo simulation to establish resistance factors for ML models for predicting the axial compression capacity of concrete-filled steel tubes and the flexural capacity of reinforced concrete beams strengthened with a fabric reinforced cementitious matrix composites. Resistance factors for ML models were also evaluated according to various design standards other than EN 1990 in [26–29].

## 2 Research significance

The economical design, safety, and performance of many steel-concrete composite structures rely on the accurate predictions of the shear resistance of headed studs transferring the interface forces between the two materials. The existing mechanics-based or empirical design equations do not always produce accurate and safe predictions of the stud shear resistance. This paper proposes new ML models to compute the shear resistance of headed studs in solid NWC and LWC slabs, which outperform the existing descriptive equations. The proposed ML models, applicable to composite bridge beams and concrete-encased and concrete-filled steel columns, have been calibrated to meet the reliability requirements of Eurocodes and US design practice. A web application was created and made publicly available to facilitate rapid and accurate predictions of the stud shear resistance by the developed ML models.

## 3 Test databases

The ML models presented in this study were developed using 242 and 90 push-out test results for shear studs embedded in solid NWC and LWC slabs, respectively [1, 2]. A schematic drawing of the standard push-out test according to Eurocode 4 [5, 6] is shown in Fig. 1. The specimen dimensions in the standard push-out test prevent concrete edge failure (also known as concrete breakout failure mode [30]), which is not typical for steel-concrete composite structures [7]. All specimens in the databases failed by shear due to the studs reaching their shear capacity or concrete failure near the studs (also known as concrete pryout failure mode [30]).

The databases include the mean measured shear resistance per stud, $P_{\mathrm{em}}$, together with the mean measured and nominal values of the stud and concrete properties for each tested specimen, including compressive strength of concrete, $f_{\mathrm{cm}}$; concrete secant modulus of elasticity, $E_{\mathrm{cm}}$; ultimate tensile strength of studs, $f_{\mathrm{um}}$; diameter of stud shank, $d_{\mathrm{m}}$; weld collar diameter, $d_{\mathrm{dom}}$; weld collar height, $h_{\mathrm{wm}}$; stud height after welding, $h_{\mathrm{m}}$; stud height-to-diameter ratio, $h_{\mathrm{m}}/d_{\mathrm{m}}$; and concrete density (in the LWC database only).

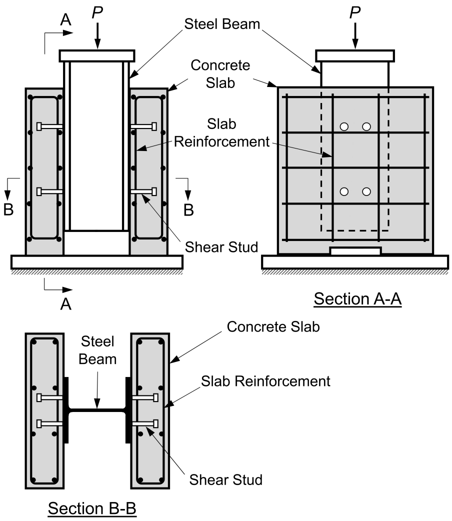



**Fig. 1**: Standard push-out test according to Eurocode 4

Statistical parameters of the database variables are given in Table 1, which indicate that the databases cover a wide range of variables typically found in real structures. More detailed information about the database variables, including their distributions and correlations can be found in [10].

**Table 1**: Statistical parameters of the database variables

| Variables | Minimum | | Maximum | | Mean | | Standard Deviation | | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NWC | LWC | NWC | LWC | NWC | LWC | NWC | LWC | NWC | LWC | NWC | LWC |
| $P_{\mathrm{em}}$ (kN) | 61.8 | 40.9 | 318.9 | 123.7 | 174.0 | 84.3 | 55.9 | 19.6 | 0.24 | -0.23 | -0.66 | -0.58 |
| $f_{\mathrm{cm}}$ (MPa) | 16.6 | 20.5 | 115.8 | 55.7 | 59.9 | 30.4 | 31.0 | 6.9 | 0.53 | 1.51 | -1.22 | 3.65 |
| $E_{\mathrm{cm}}$ (GPa) | 15.1 | 10.4 | 46.5 | 19.4 | 34.9 | 14.7 | 6.3 | 2.3 | -0.50 | 0.07 | 0.18 | -0.67 |
| $f_{\mathrm{um}}$ (MPa) | 426.0 | 406.8 | 675.0 | 600.0 | 518.8 | 484.6 | 49.5 | 44.7 | 0.50 | 0.28 | -0.14 | -0.31 |
| $d_{\mathrm{m}}$ (mm) | 12.7 | 12.7 | 31.8 | 22.2 | 21.0 | 17.9 | 2.8 | 2.7 | 0.43 | -0.88 | 0.92 | 0.02 |
| $d_{\mathrm{dom}}$ (mm) | 21.0 | 17.0 | 44.5 | 29.0 | 27.1 | 22.3 | 3.9 | 2.9 | 0.72 | 0.02 | 1.66 | 1.11 |
| $h_{\mathrm{wm}}$ (mm) | 3.0 | 3.0 | 8.6 | 6.0 | 5.8 | 5.4 | 0.9 | 1.1 | 0.06 | -1.38 | 0.66 | 0.23 |
| $h_{\mathrm{m}}/d_{\mathrm{m}}$ | 3.0 | 2.7 | 9.1 | 8.0 | 5.1 | 4.8 | 1.2 | 1.2 | 1.66 | 1.29 | 3.41 | 0.94 |
| $h_{\mathrm{m}}$ (mm) | 69.9 | 50.8 | 200.0 | 114.3 | 107.6 | 84.1 | 29.3 | 14.6 | 1.60 | 0.51 | 3.26 | -0.54 |
| density (kg/m$^3$) | – | 1409.6 | – | 1970.3 | – | 1688.5 | – | 116.6 | – | -0.11 | – | -0.13 |

# 4 Machine learning models

The performance of nine supervised ML regression algorithms for predicting the shear resistance of headed studs in solid concrete slabs was evaluated. The algorithms included k-nearest neighbors (KNN), decision tree (DT), random forest (RF), gradient boosting regressor (GBR), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), gradient boosting with categorical features support (CatBoost), support vector machine regression (SVR), and artificial neural network (ANN).

The six input variables (features) of the models included $f_{\mathrm{cm}}$, $f_{\mathrm{um}}$, $d_{\mathrm{m}}$, $d_{\mathrm{dom}}$, $h_{\mathrm{wm}}$, and $h_{\mathrm{m}}$. ML models with this feature combination produced better prediction accuracy compared with ML models that employed different combinations of features. $E_{\mathrm{cm}}$, which is considered by many existing design models, was excluded from the input variables because the ML models with $E_{\mathrm{cm}}$ as a feature demonstrated poorer performance than similar models without $E_{\mathrm{cm}}$. As was shown in [10] and will be confirmed hereafter, $E_{\mathrm{cm}}$ does not affect the stud resistance predictions when models properly account for the effect of $f_{\mathrm{cm}}$. Similar to the earlier reliability studies that formed the basis for Eurocode 4 [31, 32], $E_{\mathrm{cm}}$ was not measured in many tests included in the databases. For those cases, the $E_{\mathrm{cm}}$ values were computed from the relationship given in the fib Model Code [33], where $E_{\mathrm{cm}}$ is a function of $f_{\mathrm{cm}}$, aggregate type, and concrete type. Therefore, $E_{\mathrm{cm}}$ may only be needed in ML models or predictive equations to capture the effect of concrete type on the stud shear resistance, which was accomplished in the presented study by training separate models for NWC and LWC. The mean shear resistance per stud, $P_{\mathrm{em}}$, was the output variable (target) of the ML models.

Separate ML models were developed for NWC and LWC slabs. Single ML models for both concrete types based on the combined NWC and LWC databases were also evaluated. They performed worse than the separate models and required larger reduction factors to meet the reliability requirements, which will be discussed later in the present paper. The following subsection presents a brief overview of the considered ML algorithms. Detailed information about them can be found in [34–36].

ML models have parameters and hyperparameters. The former are learned by models during training. The latter are specified by the person developing ML models to control the learning process and define the model structure. It is essential to find optimal hyperparameters that produce accurate predictions and good generalization ability of the models, characterized by accurate predictions for new data. Several approaches are available for hyperparameter tuning, including grid search, random search, and various optimization techniques.

## 4.1 Review of machine learning algorithms

In KNN regression [34], targets are predicted by interpolating the outputs for $k$ nearest neighbors with similar features in the training set. The KNN regression hyperparameters include the number of neighbors ($k$), the weight function (uniform or inverse distance weighted), and the distance metric. When the uniform weight function is specified, all $k$ neighbors have the same effect on the predictions, meaning that the output values of the neighbors are averaged to make a new prediction. For the inverse distance weighted weight function, closer neighbors affect the prediction more than distant neighbors. A new prediction is made by taking an inverse distance weighted average of the output values for the $k$ nearest neighbors. The main advantage of KNN over other ML algorithms is a short calculation time because it does not have a training period. The predictions are made based on the training dataset directly. However, KNN often provides less accurate predictions than other more robust models, especially for large sets of noisy data.

DT models have a tree structure with the root node, decision nodes, and terminal nodes (leaves) [34]. A DT model incrementally develops by partitioning the dataset into smaller subsets. The learning process starts at the root node, which includes all training samples. The root node divides into decision nodes based on the algorithm splitting criteria. The splitting continues for the subsequent levels until the nodes have only one training data sample or when a predefined maximum tree depth is reached. The DT hyperparameters are the maximum tree depth, the minimum number of samples required for node splitting, the minimum number of samples at a leaf node, and others. DT models can be more easily understood than many other ML models. The algorithm is robust against missing values but prone to overfitting [37, 38], which can be avoided by limiting the DT size.

RF is an ensemble of decision trees trained via *bootstrap aggregating* (*bagging*) [34], where the DT algorithm is trained many times on different random subsets of the training set. The training set subdivision is done with replacement, where one sample may appear in various subsets. RF makes predictions by averaging predictions of multiple randomly generated decision trees, which improves the algorithm generalization ability and makes it robust against overfitting compared with DT. However, the need for building many trees and combining their outputs requires greater computational power. The RF hyperparameters are the number of trees, maximum tree depth, the minimum number of samples required to split at an internal node, the minimum number of samples at a leaf node, the number of features to consider when looking for the best split, maximum number of leaves, and others.

Gradient boosting algorithms, represented in this study by GBR, XGBoost, LightGBM, and CatBoost, are ensembles of decision trees trained via boosting, where each subsequent DT improves predictions of its predecessor by fitting to the residual errors from the previous predictors [39]. The advantages of the gradient boosting algorithms include overfitting resistance, high accuracy, flexibility, and insensitivity to missing data. However, the gradient boosting algorithms require high computational resources and tuning many hyperparameters. Their interpretability is limited. The hyperparameters of gradient boosting algorithms are learning rate, the number of boosting iterations, maximum depth of the individual regression estimators, the minimum number of samples required to split an internal node, the minimum number of samples at a leaf node, and others. XGBoost [40], LightGBM [41], and CatBoost [42] are improved modifications of GBR. XGBoost improvements include regularization for more accurate and faster predictions, custom loss functions, and parallel processing. LightGBM possesses lower memory usage, better accuracy, higher efficiency, improved training speed, and the ability to process large datasets by applying the Gradient-based One-Side Sampling (GOSS) method and parallel learning. CatBoost stands out by its ability to process categorical features with improved accuracy, ordered boosting to fight overfitting, and missing value support.

SVR is based on the structural risk minimization principle [43–45]. A regression function (hyperplane) is found to ensure that a tube with radius $\epsilon$ contains most data points. Data points outside the tube are penalized by $\xi$, which is a regularization parameter or a soft margin. It represents a degree of importance that is given to outliers. Data points near the decision boundaries are called support vectors. Various kernel functions are used in the algorithm to handle nonlinear data. Kernel functions transform the original data into high-dimensional kernel space where a linear hyperplane function can be found. The advantages of SVR compared with other ML algorithms include the high efficiency of handling high-dimensional data with balancing the model complexity and prediction error, insensibility to outliers, ability to handle nonlinear data, good generalization ability, and good performance on small datasets. However, SVR requires extensive memory, which results in a long training time, especially for large datasets. Finding an appropriate kernel function might be challenging. The SVR hyperparameters are the kernel function type and its parameters, the "soft margin" constant, and the margin of tolerance $\epsilon$.

Feedforward multilayer perceptron ANNs [35] were evaluated in this study. ANNs of this type consist of input, hidden, and output layers of multiple neurons; neuron connections (weights); and neuron-attached biases. In a feedforward ANN, the information flows from the input layer to the output layer through the hidden layer(s) without forming a cycle. Activation functions of various types are used in ANNs to transform values passed from one layer onto the subsequent layer. The weights and biases, the initial values of which are assumed at the beginning of the training, are learned by the network during the training process via backpropagation, which is based on the gradient descent method. The benefits of ANNs include a high efficiency in finding complex relationships between features and targets and all possible interactions between features. On the negative side, ANNs usually require large datasets for training an accurate model for a complex problem and high computational resources. ANNs are prone to overfitting and are challenging to interpret. The ANN hyperparameters are learning rate, number of layers, number of hidden units in each layer, activation function, optimizer, and others.

## 4.2 Implementation of machine learning algorithms

The following open-source Python libraries were used in the development of the ML models: *scikit-learn* (KNN, DT, RF, GBR, and SVR) [46], *XGBoost* [40], *LightGBM* [41], *CatBoost* [42], and *Keras* (ANN) [47].

The models were validated and tested via the ten-fold cross-validation method as follows. The NWC and LWC databases were randomly split into training and test sets, with 80% of samples assigned to the training set and 20% of samples left for testing. The training sets were divided into ten groups, nine of which were used for model training, with one group kept for model validation. The process was repeated ten times until each group served as the validation set. The test set, unseen by the models during the training, was used for the final test of the model performance and generalization abilities.

To improve the training process, the model features in the training set were standardized using Eq. (1) to make the feature scales uniform. The test set features were also standardized by applying the mean and standard deviation values of the features from the training set.

$$x' = \frac{x - \mu}{\sigma} \tag{1}$$

where $x'$ is the standardized value of the feature, $x$ is the original value of the feature, $\mu$ is the mean of the feature original values, and $\sigma$ is the standard deviation of the feature original values.

The following performance metrics commonly used in ML [48] were monitored:

- Root-mean-square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y-x)^2} \tag{2}$$

- Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y-x| \tag{3}$$

- Mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}|\frac{y-x}{y}| \tag{4}$$

- Coefficient of determination ($R^2$):

$$R^2 = \left[\frac{\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x-\bar{x})^2\sum_{i=1}^{n}(y-\bar{y})^2}}\right]^2 \tag{5}$$

where $n$ is the number of observations, $y$ is the stud resistance from tests, $x$ is the stud resistance predicted by models, $\bar{y}$ and $\bar{x}$ are the mean values of $y$ and $x$.

Each considered performance metric has specific features. RMSE penalizes large prediction errors more than MAE, making MAE more robust to outliers than RMSE. However, the use of RMSE for selecting the best-performing model allows for reducing large errors. MAPE is a convenient metric that shows the average percentage difference between the observations and predictions.

$R^2$ indicates how well the model can predict target variability. Chicco et al. [49] recommended $R^2$ as a standard metric for regression models in any scientific domain because it is more informative than RMSE, MAE, and MAPE. All these metrics are presented in this paper because publications on ML have traditionally reported them. As will be shown further in the paper, the model rankings based on each considered metric were identical.

The optimal values of the hyperparameters were found via extensive grid and random searches implemented in *scikit-learn* for all considered ML algorithms, except SVR, for which the *Optunity* [50] library with particle swarm optimization [51] was used. The following subsection presents the optimal hyperparameters for each evaluated model and the performance of the optimized models.

### 4.3 Developed machine learning models

The developed ML models are characterized by the hyperparameters given in Table 2, which shows the hyperparameter names used in the Python libraries. The hyperparameters that are not listed have the default values. More information on how each hyperparameter affects model performance can be found on the web pages of the libraries.

The performance of the developed ML models with the optimized hyperparameters on the training and test sets of the NWC and LWC databases is illustrated in Figs. 2 and 3. The performance metrics are presented in Tables 3 and 4.

In general, all evaluated models demonstrated an excellent prediction accuracy and reasonable generalization ability, characterized by a relatively small difference in the performance metrics for the training and test sets. For the

**Table 2**: Optimal hyperparameters of ML models

| Model | Hyperparameter | Optimal hyperparameters | |
|---|---|---|---|
| | | NWC | LWC |
| KNN | n_neighbors | 2 | 2 |
| | weights | uniform | uniform |
| | p | 2 | 3 |
| | leaf_size | 10 | 5 |
| DT | max_depth | 11 | 8 |
| | min_samples_split | 2 | 2 |
| | min_samples_leaf | 1 | 2 |
| | ccp_alpha | 2.4 | 0 |
| RF | n_estimators | 320 | 125 |
| | max_depth | 6 | None |
| | min_samples_split | 2 | 2 |
| | min_samples_leaf | 1 | 1 |
| GBR | learning_rate | 0.025 | 0.025 |
| | n_estimators | 430 | 640 |
| | max_depth | 3 | 3 |
| | min_samples_split | 2 | 3 |
| | min_samples_leaf | 1 | 1 |
| | subsample | 0.305 | 0.16 |
| XGBoost | eta | 0.025 | 0.025 |
| | n_estimators | 550 | 500 |
| | gamma | 0.5 | 0.25 |
| | max_depth | 3 | 3 |
| | min_child_weight | 3 | 2 |
| | colsample_bytree | 0.6 | 0.6 |
| LightGBM | learning_rate | 0.025 | 0.025 |
| | num_iterations | 680 | 580 |
| | num_leaves | 8 | 8 |
| | min_data_in_leaf | 2 | 2 |
| | max_depth | 3 | 3 |
| | bagging fraction | 0.5 | 0.8 |
| | bagging_freq | 4 | 10 |
| CatBoost | learning_rate | 0.03 | 0.03 |
| | iterations | 540 | 670 |
| | depth | 3 | 3 |
| | l2_leaf_reg | 1 | 3 |
| | random_strength | 1 | 5 |
| | bagging_temperature | 1 | 1 |
| SVR | kernel | RBF | RBF |
| | C | 708 | 203 |
| | gamma | 0.19 | 0.5 |
| | epsilon | 2.81 | 0.07 |
| ANN | Number of hidden layers | 1 | 1 |
| | Number of hidden layer nodes | 12 | 10 |
| | Learning rate | 0.1 | 0.1 |
| | Activation function in both layers | ReLU | ReLU |
| | Loss function | MSE | MSE |
| | Optimizer | Adam | Adam |
| | Mini-batch size | 8 | 8 |

test set, the gradient boosting models (GBR, XGBoost, LightGBM, and CatBoost) produced the best performance metrics for the stud shear resistance in NWC slabs, with LightGBM being the best model, followed by CatBoost and

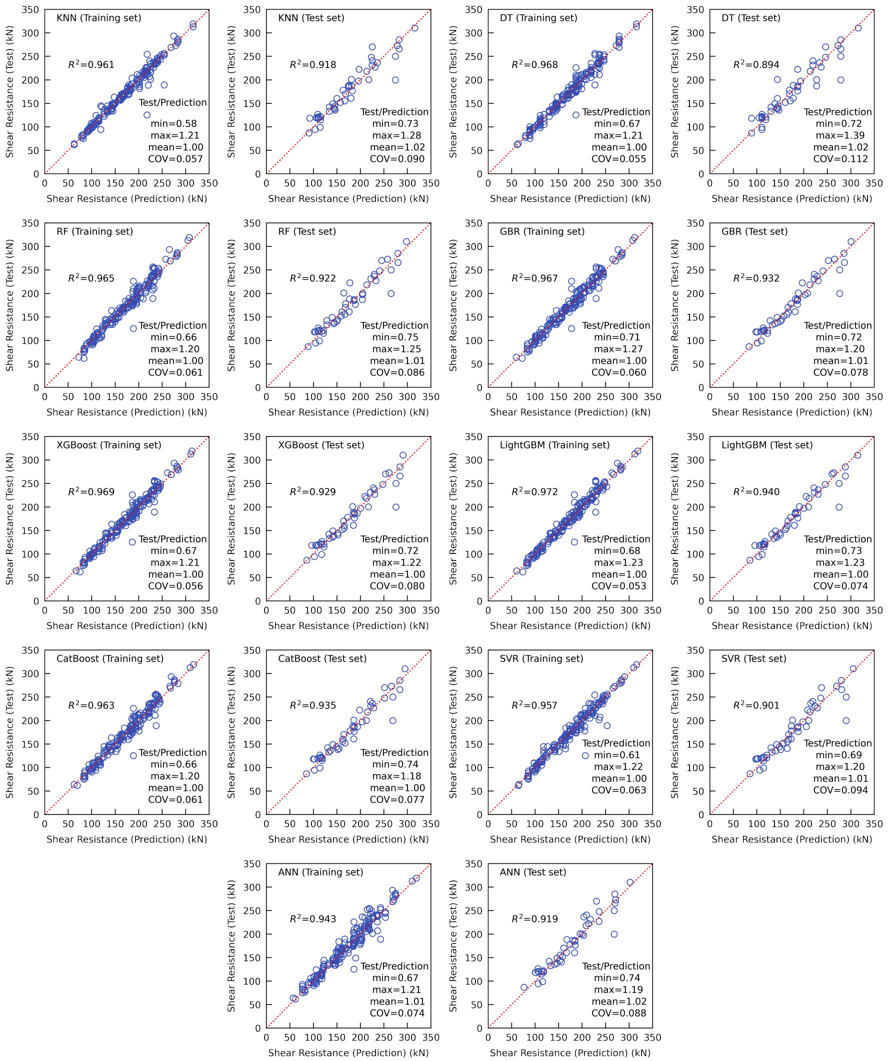



**Fig. 2**: Performance of ML models for predicting the nominal (mean) shear resistance of studs in NWC slabs

GBR. For LWC slabs, SVR and ANN produced the best performance metrics for the test set. They were followed by LightGBM, RF, CatBoost, and XGBoost. Overall, these six models produced comparable performance metrics for LWC slabs, with minor absolute differences in the metrics values. Tables

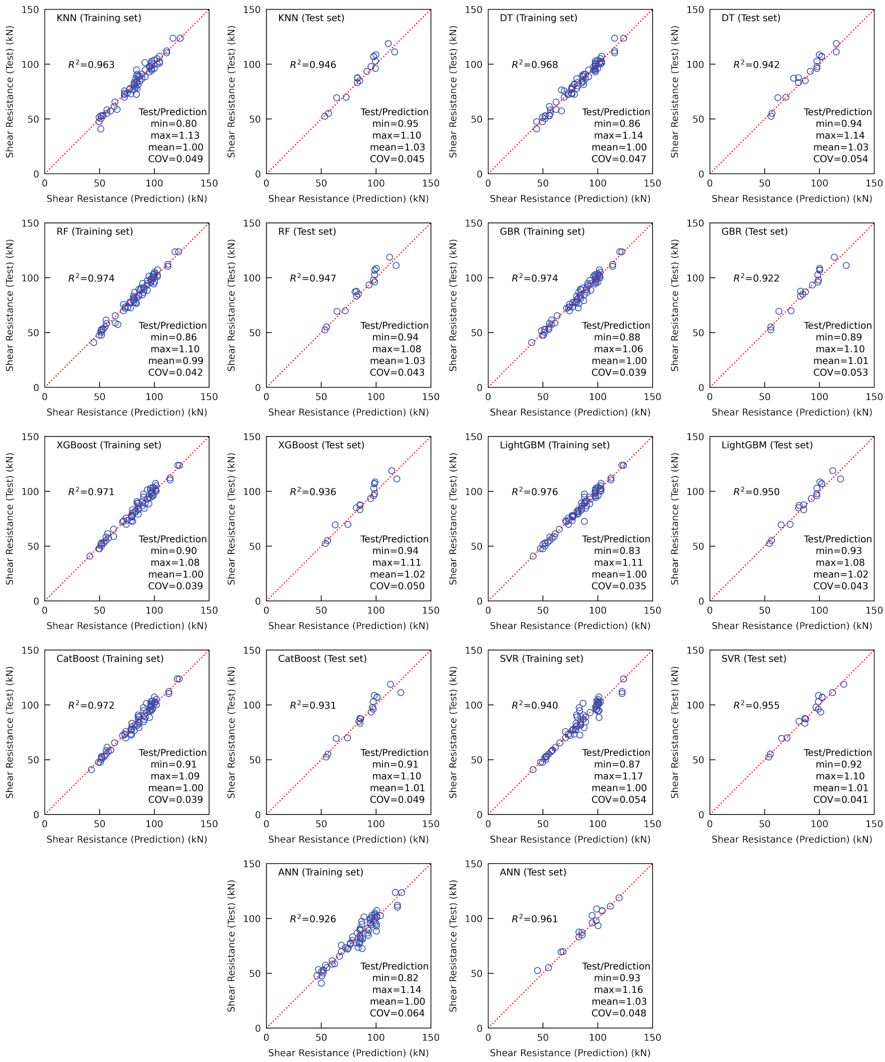



**Fig. 3**: Performance of ML models for predicting the nominal (mean) shear resistance of studs in LWC slabs

3 and 4 also demonstrate that model rankings based on each considered performance metric were identical. All developed ML models can be accessed at https://github.com/vitdegtyarev/Streamlit_Studs_Solid.

**Table 3**: Performance metrics of ML models for predicting the nominal (mean) shear resistance of studs in NWC slabs

| Model | RMSE (kN) | | MAE (kN) | | MAPE (%) | | $R^2$ | |
|---|---|---|---|---|---|---|---|---|
| | Train | Test | Train | Test | Train | Test | Train | Test |
| KNN | 11.0 | 17.2 | 5.4 | 12.0 | 3.3 | 7.0 | 0.961 | 0.918 |
| DT | 9.8 | 19.9 | 6.3 | 13.4 | 3.8 | 7.9 | 0.968 | 0.894 |
| RF | 10.5 | 16.1 | 7.2 | 11.5 | 4.5 | 6.6 | 0.965 | 0.922 |
| GBR | 10.1 | 15.3 | 7.1 | 10.1 | 4.5 | 5.9 | 0.967 | 0.932 |
| XGBoost | 9.8 | 15.5 | 6.5 | 10.6 | 4.0 | 6.2 | 0.969 | 0.929 |
| LightGBM | 9.3 | 14.7 | 6.0 | 9.6 | 3.7 | 5.6 | 0.972 | 0.940 |
| CatBoost | 10.6 | 14.8 | 7.4 | 10.0 | 4.6 | 5.9 | 0.963 | 0.935 |
| SVR | 11.5 | 18.7 | 6.8 | 12.2 | 4.3 | 7.2 | 0.957 | 0.901 |
| ANN | 13.4 | 16.5 | 9.4 | 11.4 | 5.6 | 6.8 | 0.943 | 0.919 |

**Table 4**: Performance metrics of ML models for predicting the nominal (mean) shear resistance of studs in LWC slabs

| Model | RMSE (kN) | | MAE (kN) | | MAPE (%) | | $R^2$ | |
|---|---|---|---|---|---|---|---|---|
| | Train | Test | Train | Test | Train | Test | Train | Test |
| KNN | 3.8 | 5.0 | 2.8 | 4.1 | 3.6 | 4.3 | 0.963 | 0.946 |
| DT | 3.5 | 5.3 | 2.9 | 4.4 | 3.7 | 4.9 | 0.968 | 0.942 |
| RF | 3.2 | 4.9 | 2.5 | 4.0 | 3.2 | 4.2 | 0.974 | 0.947 |
| GBR | 3.2 | 5.2 | 2.6 | 3.9 | 3.2 | 4.2 | 0.974 | 0.922 |
| XGBoost | 3.3 | 4.9 | 2.6 | 3.9 | 3.2 | 4.2 | 0.971 | 0.936 |
| LightGBM | 3.0 | 4.5 | 2.0 | 3.8 | 2.4 | 4.2 | 0.976 | 0.950 |
| CatBoost | 3.3 | 4.9 | 2.6 | 3.8 | 3.2 | 4.1 | 0.972 | 0.931 |
| SVR | 4.8 | 3.9 | 2.8 | 3.0 | 3.2 | 3.3 | 0.940 | 0.955 |
| ANN | 5.3 | 4.2 | 4.2 | 2.9 | 5.2 | 3.4 | 0.926 | 0.961 |

## 4.4 Interpretation of the developed machine learning models

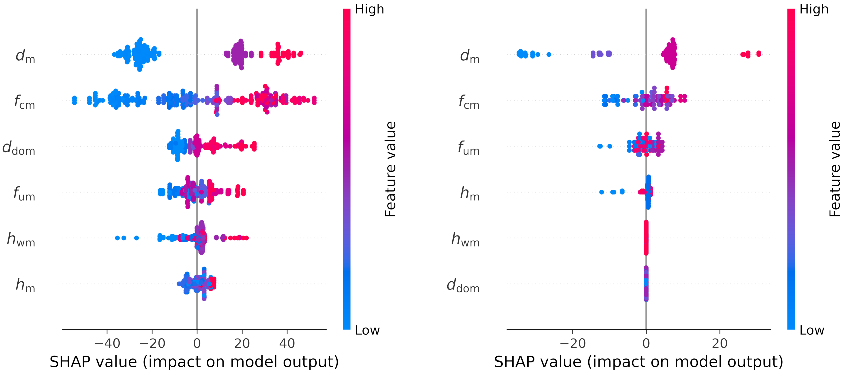
The developed models are based on ML algorithms that are often criticized for the lack of transparency. It is challenging for humans to understand how and why the models made specific predictions. This criticism can be addressed by applying special interpretability and explainability techniques to the trained ML models [52].

In the present study, relative feature importance and partial dependence were evaluated using the SHapley Additive exPlanations (SHAP) method [53]. This method estimates the contribution of each feature to the ML model prediction based on the cooperative game theory's Shapley values [54]. The Shapley values indicate the average contribution of each feature (which serves as a player in the context of the cooperative game theory) to the ML predictions. Features with larger absolute average Shapley values are more important for the model predictions than others.

The SHAP method estimates the effect of features in the trained model, not in the dataset used for the model training. Therefore, the SHAP feature importance and dependence for an inaccurate model are not accurate either. In the present study, all optimized models demonstrated good accuracy. Thus, the SHAP feature importance and dependence evaluated for the optimized models can shed light on the actual effects of the independent variables of the database on the stud shear resistance in NWC and LWC slabs.

Fig. 4 shows SHAP summary plots for the stud shear resistance in NWC and LWC slabs predicted by the LightGBM models, which were among the most accurate models for both concrete types. A Shapley value for each dataset sample is shown with a point on the SHAP summary plot. The color of the points indicates the feature value ranging from low (shown in blue) to high (shown in red). Points with the same Shapley values for each feature are scattered vertically to demonstrate their distribution. The feature order from top to bottom follows their importance.

It can be noticed from Fig.4 that the absolute SHAP values for NWC are higher than those for LWC, which is caused by the higher shear resistances of studs in NWC compared with LWC in the considered databases (see Table 1). The absolute SHAP values representing a contribution of each feature to




(a) NWC (b) LWC

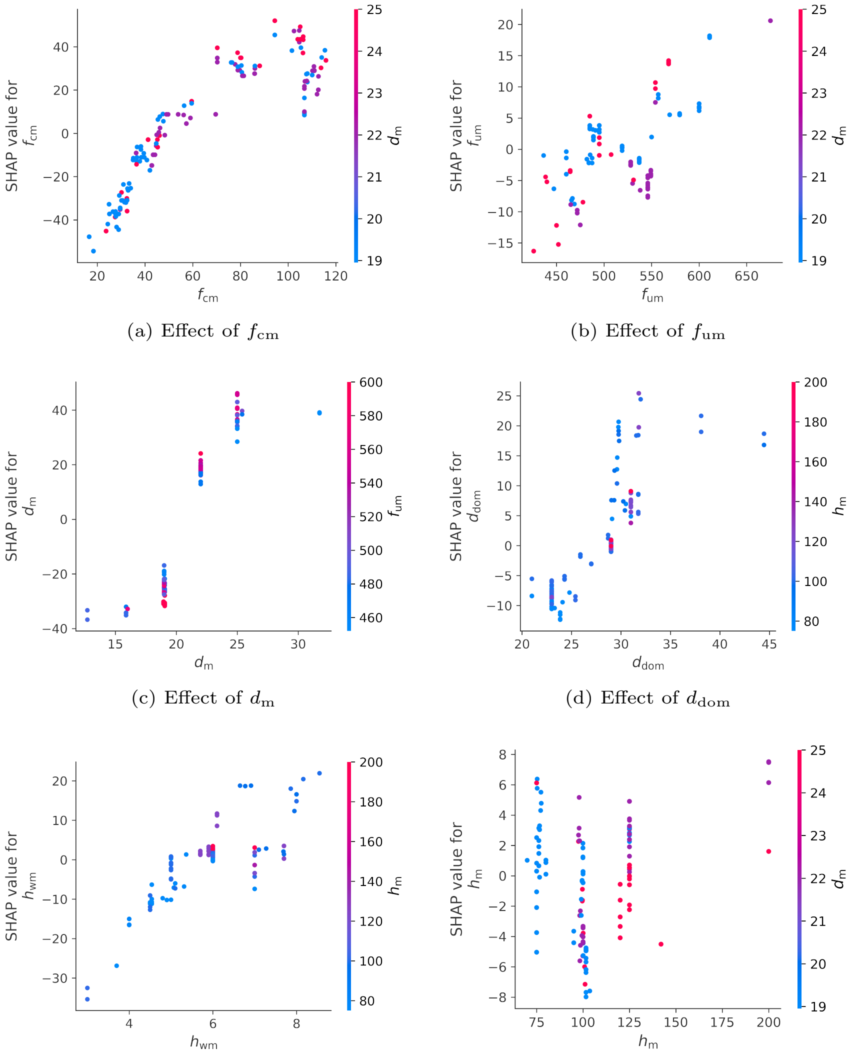
**Fig. 4**: SHAP summary plots for LightGBM models

the ML model predictions tend to be higher in the models trained on datasets containing targets with higher values. The SHAP values also depend on the model performance. Therefore, comparing absolute SHAP values for different models is not recommended even when the models were trained on the same dataset.

Fig. 4(a) indicates that the stud diameter, $d_{\mathrm{m}}$, and the concrete compressive strength $f_{\mathrm{cm}}$ have the most significant importance on the stud shear resistance in NWC slabs. The weld collar diameter, $d_{\mathrm{dom}}$, the stud tensile strength, $f_{\mathrm{um}}$, and the weld collar height, $h_{\mathrm{wm}}$, have less significant impacts on the stud shear resistance, with the stud height, $h_{\mathrm{m}}$, barely affecting the stud resistance.

For LWC slabs, the stud diameter, $d_{\mathrm{m}}$, affects the stud resistance more significantly than other variables for the samples in the considered database (see Fig. 4(b)). The weld collar diameter and height, $d_{\mathrm{dom}}$ and $h_{\mathrm{wm}}$, have no effect on the stud shear resistance in LWC slabs.

Figs. 5 and 6 present SHAP dependence plots for NWC and LWC slabs, respectively, for the LightGBM models. The dependence plots demonstrate how each feature affects the stud shear resistance for the entire range of the feature values. The evaluated features are shown on the horizontal axes. The dependence plot points represent database samples. The point color corresponds to the second feature, which has the highest interaction with the evaluated feature, as was determined by the algorithm.




(a) Effect of $f_{\mathrm{cm}}$

(b) Effect of $f_{\mathrm{um}}$

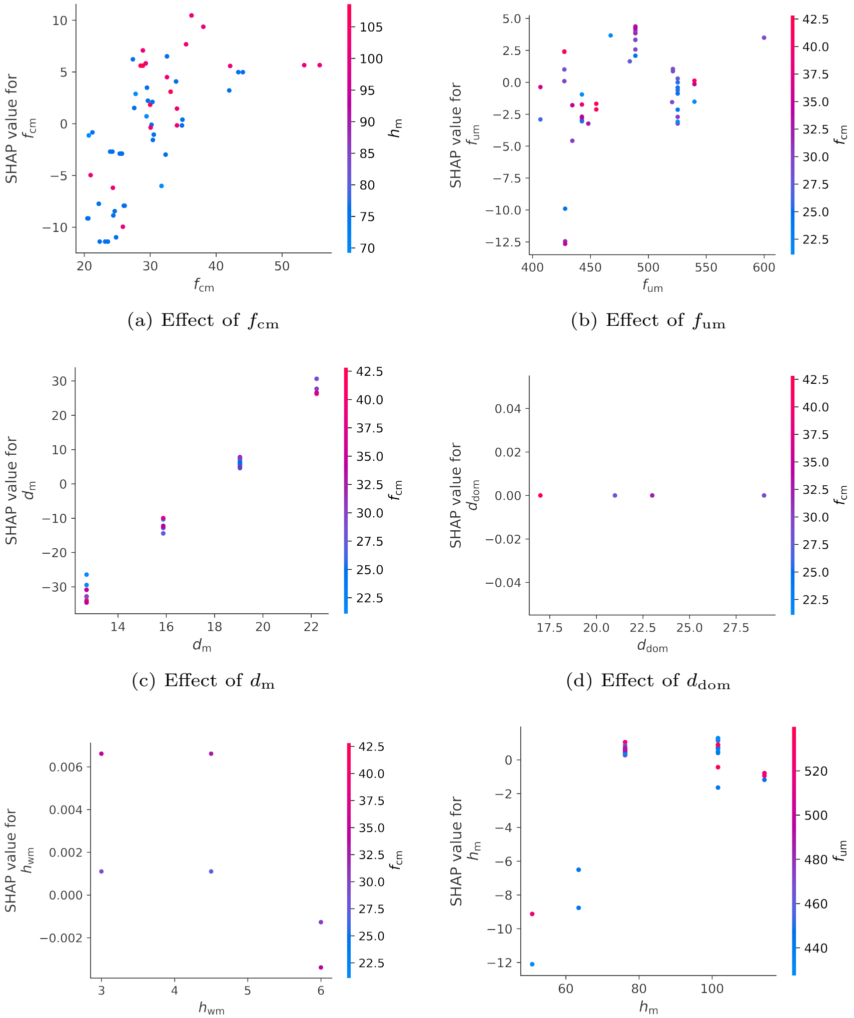
(c) Effect of $d_{\mathrm{m}}$

(d) Effect of $d_{\mathrm{dom}}$

(e) Effect of $h_{\mathrm{wm}}$

(f) Effect of $h_{\mathrm{m}}$

**Fig. 5**: SHAP dependence plots for LightGBM NWC model




(a) Effect of $f_{\mathrm{cm}}$ (b) Effect of $f_{\mathrm{um}}$

(c) Effect of $d_{\mathrm{m}}$ (d) Effect of $d_{\mathrm{dom}}$

(e) Effect of $h_{\mathrm{wm}}$ (f) Effect of $h_{\mathrm{m}}$

**Fig. 6**: SHAP dependence plots for LightGBM LWC model

Fig. 5(a) indicates that there is a nonlinear relationship between the compressive strength of NWC and the stud shear resistance. The stud resistance increases more significantly when the concrete strength goes up from 20 to 50 MPa compared with the concrete strength increase from 50 to 120 MPa.

It is especially obvious for the studs of small and medium diameters (represented by the blue and purple points), which resistance is governed by the stud material strength. The stud shear resistance in NWC slabs generally increases when the stud tensile strength increases (see Fig. 5(b)). However, a noticeable scatter in the effect of $f_{\mathrm{um}}$ on the stud resistance can be observed because the latter can be governed by the concrete strength when it is low. The stud resistance goes up when the stud diameter increases, especially in the range from 16 to 25 mm, and when the $f_{\mathrm{um}}$ values are high (see Fig. 5(c)). The stud resistance increase does not occur when $d_{\mathrm{m}}$ increases from 25 to 32 mm. That is likely because the NWC test database includes only two test results for $d_{\mathrm{m}}$=32 mm, with $f_{\mathrm{cm}}$ of 22 and 24 MPa, making the stud resistance governed by the concrete strength. The effect of the weld collar diameter on the stud resistance is somewhat similar to that of $d_{\mathrm{m}}$ (see Fig. 5(d)) because $d_{\mathrm{dom}}$ and $d_{\mathrm{m}}$ are correlated. The stud shear resistance increases when the weld collar height increases (see Fig. 5(e)). The effect of the stud height on the stud shear resistance presented in Fig. 5(f) is not obvious.

For LWC slabs, there is a tendency for the stud resistance to increase when the concrete strength increases, similar to that observed for NWC (see Fig. 6(a)), but the scatter is more pronounced for LWC. The $f_{\mathrm{um}}$ effect on the stud resistance also demonstrates a large scatter (see Fig. 6(b)). For the low and medium concrete strengths (blue and purple points), the stud resistance increases when $f_{\mathrm{um}}$ increases up to approximately 500 MPa, indicating that the stud material strength governs the stud resistance. Further increase of $f_{\mathrm{um}}$ does not affect the stud resistance (indicating that concrete strength governs the stud resistance) or causes its reduction, which is likely due to the small number of tests in the LWC database and other factors affecting the stud resistance. Fig. 6(c) indicates a strong linear correlation between the stud diameter and stud shear resistance in LWC slabs. The weld collar diameter and height practically have no effect on the stud shear resistance (see Fig. 6(d) and 6(e)). The stud resistance in LWC goes up when the stud height increases up to approximately 80 mm and does not change with the further increase of $h_\mathrm{m}$ (see Fig. 6(f)).

Generally, it can be concluded that the SHAP partial importance and dependence plots for the LightGBM models align well with the mechanics-based knowledge, indicating that the proposed ML models can capture feature importance and dependence from the test data.

# 5 Reliability evaluation of machine learning models and design shear resistance

The developed optimized ML models can predict the nominal (mean) stud shear resistance with outstanding accuracy, but it is not evident that the ML predictions meet the reliability requirements of design standards. To address this concern, the reliability of the developed ML models was evaluated in accordance with European and US design practices, in a similar way as presented in [10].

## 5.1 European design practice

Eurocode 4 (EC4) [5, 6] governs the design of steel-concrete composite structures in Europe. It recommends the partial factor for the stud shear resistance $\gamma_\mathrm{V}$ of 1.25, which aims to ensure a probability of failure not greater than $P_\mathrm{f} = 1.2 \times 10^{-3}$ [8, 11]. This probability of failure corresponds to the adjusted target reliability index $\alpha_\mathrm{R}\beta$ of 3.04, which is obtained by multiplying the target reliability index $\beta$ of 3.8 for a 50-year reference period by the First Order Reliability Method (FORM) sensitivity factor for resistance $\alpha_\mathrm{R}$ of 0.8.

The reliability analyses for evaluating the design resistance should be in accordance with the method presented in Annex D of EN 1990 [11]. The specific steps of the method are briefly described below. More information about the reliability analysis of the stud shear resistance, including all corresponding equations, can be found in [8, 10].

- Develop a design model for the theoretical resistance, which is an ML model in the context of this paper;
- Determine the correction factor $b$ from the comparison of the theoretical and experimental resistances;
- Estimate the coefficients of variation of the error terms $V_\delta$ and the theoretical resistance $V_{\mathrm{rt}}$, from the combination of which the coefficient of variation $V_{\mathrm{r}}$ is obtained. The coefficient of variation of the theoretical resistance $V_{\mathrm{rt}}$ was estimated in this study through Monte Carlo simulation, using the same parameters as described in [10];
- Determine the characteristic and design resistances $R_{\mathrm{k}}$ and $R_{\mathrm{d}}$;
- Compute the partial factor $\gamma_{\mathrm{M}} = R_{\mathrm{k}}/R_{\mathrm{d}}$;
- Compute the corrected partial factor $\gamma_{\mathrm{M}}^{*} = \gamma_{\mathrm{V}} = R_{\mathrm{n}}/R_{\mathrm{d}} = k_{\mathrm{c}}\gamma_{\mathrm{M}}$, where $R_{\mathrm{n}}$ is the nominal resistance determined from the design model using nominal values of the variables and $k_{\mathrm{c}} = R_{\mathrm{n}}/R_{\mathrm{k}}$ [8].

The reliability analyses of the ML models for the nominal (mean) stud shear resistance indicated that a reduction coefficient $k_{\mathrm{red}}$ should be applied to the model predictions to satisfy the EC4 reliability requirements with $\gamma_{\mathrm{V}} = 1.25$. A summary of the reliability analysis results, including the required reduction coefficient for each considered ML model, is presented in Table 5. The reliability analyses were performed assuming the case "$V_{\mathrm{X}}$ unknown", which is more conservative than the case "$V_{\mathrm{X}}$ known". EN 1990 requires the coefficient of

468 variation to be no smaller than 0.10 for the case "$V_X$ unknown", which was
469 considered in the reliability analyses.

**Table 5**: Reliability analysis results for ML models per EN 1990

| Model | $k_{red}$ | $b$ | $V_\delta$ | $V_{rt}$ | $V_r$ | $\gamma_M$ | $k_c$ | $\gamma_M^* = \gamma_V$ |
|---|---|---|---|---|---|---|---|---|
| | | | | NWC ($n$=242) | | | | |
| KNN | 0.89 | 1.119 | 0.070 | 0.111 | 0.132 | 1.20 | 1.04 | 1.25 |
| DT | 0.75 | 1.331 | 0.072 | 0.150 | 0.167 | 1.26 | 0.99 | 1.25 |
| RF | 0.94 | 1.067 | 0.069 | 0.077 | 0.103 | 1.15 | 1.07 | 1.24 |
| GBR | 0.97 | 1.030 | 0.066 | 0.091 | 0.112 | 1.17 | 1.07 | 1.25 |
| XGBoost | 0.89 | 1.122 | 0.064 | 0.122 | 0.138 | 1.21 | 1.03 | 1.25 |
| LightGBM | 0.96 | 1.041 | 0.061 | 0.095 | 0.113 | 1.17 | 1.07 | 1.25 |
| CatBoost | 1.00 | 1.000 | 0.067 | 0.081 | 0.105 | 1.16 | 1.08 | 1.25 |
| SVR | 0.82 | 1.212 | 0.075 | 0.142 | 0.160 | 1.25 | 0.99 | 1.23 |
| ANN | 0.80 | 1.262 | 0.079 | 0.150 | 0.170 | 1.27 | 0.98 | 1.25 |
| | | | | LWC ($n$=90) | | | | |
| KNN | 0.96 | 1.048 | 0.051 | 0.095 | 0.108 | 1.17 | 1.07 | 1.25 |
| DT | 0.86 | 1.171 | 0.050 | 0.114 | 0.124 | 1.19 | 0.99 | 1.17 |
| RF | 1.00 | 1.005 | 0.045 | 0.074 | 0.100 | 1.11 | 1.05 | 1.17 |
| GBR | 0.98 | 1.022 | 0.043 | 0.101 | 0.110 | 1.17 | 1.07 | 1.25 |
| XGBoost | 0.99 | 1.015 | 0.042 | 0.089 | 0.100 | 1.15 | 1.06 | 1.21 |
| LightGBM | 1.00 | 1.005 | 0.039 | 0.085 | 0.100 | 1.13 | 1.07 | 1.21 |
| CatBoost | 1.00 | 1.004 | 0.042 | 0.078 | 0.100 | 1.12 | 1.04 | 1.16 |
| SVR | 0.90 | 1.110 | 0.051 | 0.107 | 0.118 | 1.18 | 1.06 | 1.25 |
| ANN | 0.91 | 1.099 | 0.064 | 0.105 | 0.123 | 1.19 | 0.99 | 1.17 |
| | | | | NWC & LWC ($n$=332) | | | | |
| KNN | 0.89/0.96 | 1.113 | 0.074 | 0.107 | 0.130 | 1.20 | 1.04 | 1.25 |
| DT | 0.75/0.86 | 1.315 | 0.089 | 0.141 | 0.167 | 1.26 | 0.99 | 1.25 |
| RF | 0.94/1.00 | 1.062 | 0.069 | 0.077 | 0.103 | 1.16 | 1.07 | 1.24 |
| GBR | 0.97/0.98 | 1.030 | 0.060 | 0.094 | 0.112 | 1.17 | 1.07 | 1.25 |
| XGBoost | 0.89/0.99 | 1.112 | 0.075 | 0.114 | 0.136 | 1.21 | 1.04 | 1.25 |
| LightGBM | 0.96/1.00 | 1.038 | 0.058 | 0.092 | 0.109 | 1.17 | 1.07 | 1.25 |
| CatBoost | 1.00/1.00 | 1.000 | 0.062 | 0.080 | 0.101 | 1.15 | 1.08 | 1.24 |
| SVR | 0.82/0.90 | 1.203 | 0.083 | 0.109 | 0.137 | 1.21 | 1.03 | 1.25 |
| ANN | 0.80/0.91 | 1.246 | 0.098 | 0.140 | 0.171 | 1.27 | 0.98 | 1.25 |

Note: For the NWC & LWC data, the first and second values of $k_{red}$ are for NWC and LWC, respectively.

470 Table 5 shows that the reliability level required by Eurocodes can be
471 achieved for all considered ML models when the presented reduction factors
472 are applied. Therefore, the design shear resistance of studs in NWC and LWC
473 slabs can be taken as the ML model prediction multiplied by the reduction
474 coefficient.

Due to the high accuracy and good generalization ability, the CatBoost models for NWC and LWC did not require reduction factors at all, while the GBR and LightGBM models needed minor reductions, not exceeding 4%. Surprisingly, the SVR and ANN models, the most accurate models for predicting the nominal (mean) stud resistance in LWC, required relatively large reduction coefficients, adversely affecting the design stud resistance predicted by these models. The large resistance reductions for the SVR and ANN models were likely caused by the poor generalization abilities of the models. This finding indicates that reliability analyses can be used as an additional test of the ML model generalization ability.

## 5.2 US design practice

AISC 360 [4], which governs the design of steel-concrete composite structures in the US, includes shear strength provisions for studs in composite beams and other composite components. The stud shear strength equation for composite beams does not include a resistance factor. The required reliability of the composite beams is achieved via the resistance factor applied to the overall composite beam strength [7, 55]. Therefore, a new stud resistance model theoretically requires a calibration of the composite beam resistance factor, which was beyond the scope of this study. The stud shear strength equations for other composite components include a resistance factor of 0.65.

Following the Pallarés and Hajjar study [7], the resistance factors for the developed ML models were determined from Eq. (6) [56].

$$\phi_{\mathrm{v}} = \frac{R_{\mathrm{m}}}{R_{\mathrm{n}}} e^{(-\alpha \beta V_{\mathrm{R}})} \tag{6}$$

$$V_{\mathrm{R}} = \sqrt{V_{\mathrm{F}}^2 + V_{\mathrm{P}}^2 + V_{\mathrm{M}}^2} \tag{7}$$

where $R_{\mathrm{m}}/R_{\mathrm{n}}$ is the average ratio between the experimental and predicted values, $\alpha = 0.55$, $\beta$ is the target reliability index, $V_{\mathrm{F}} = 0.05$ is the coefficient of variation on fabrication (stud dimensions), $V_{\mathrm{P}}$ is the coefficient of variation of $R_{\mathrm{m}}/R_{\mathrm{n}}$, and $V_{\mathrm{M}} = 0.09$ is the coefficient of variation of the material properties.

The resistance factors were computed for the target reliability indices of 3.0 for composite beams and 4.0 for other applications requiring a higher level of reliability [7]. The following two approaches were used to determine the $R_{\mathrm{m}}/R_{\mathrm{n}}$ and $V_{\mathrm{P}}$ values: 1) based on the mean measured concrete strength, $f'_{\mathrm{cr}} = f_{\mathrm{cm}}$, and the mean stud tensile strength, $f_{\mathrm{um}}$; and 2) based on the specified concrete strength, $f'_{\mathrm{c}}$, and the stud nominal tensile strength of $f_{\mathrm{u}}$=450 MPa (65 ksi). The ACI 301 [57] equations were used to determine $f'_{\mathrm{c}}$ from $f'_{\mathrm{cr}}$. The stud tensile strength of 450 MPa (65 ksi) was selected as the most common tensile strength for studs used in the US [58]. The test results for the stud tensile strength smaller than 450 MPa (9 and 29 in the NWC and LWC databases, respectively) were excluded from the resistance factor calculations according to the second approach.

Tables 6 and 7 present the computed resistance factors. For the best ML models, relatively high resistance factors and small differences between the resistance factors for $\beta$=3.0 and 4.0 can be observed from Tables 6 and 7 due to the high accuracy of the models and the low values of the coefficients of variation. The resistance factors based on the target reliability index of 4.0 determined using both approaches exceed the resistance factor of 0.65 specified by AISC 360 [4] for other composite components.

**Table 6**: Resistance factors for the nominal strength ML models per US design practice based on concrete strength of $f'_{\mathrm{cr}}$=$f_{\mathrm{cm}}$ and stud strength of $f_{\mathrm{um}}$

| Model | $R_{\mathrm{m}}/R_{\mathrm{n}}$ | $V_{\mathrm{p}}$ | $\phi_{\mathrm{v}}$ $\beta$=3.0 | $\phi_{\mathrm{v}}$ $\beta$=4.0 |
|---|---|---|---|---|
| NWC ($n$=242) | | | | |
| KNN | 1.003 | 0.066 | 0.82 | 0.77 |
| DT | 1.003 | 0.071 | 0.82 | 0.76 |
| RF | 1.000 | 0.068 | 0.82 | 0.76 |
| GBR | 1.000 | 0.064 | 0.82 | 0.77 |
| XGBoost | 0.999 | 0.062 | 0.82 | 0.77 |
| LightGBM | 1.001 | 0.058 | 0.82 | 0.77 |
| CatBoost | 0.999 | 0.065 | 0.82 | 0.76 |
| SVR | 0.999 | 0.070 | 0.81 | 0.76 |
| ANN | 1.013 | 0.077 | 0.82 | 0.76 |
| LWC ($n$=90) | | | | |
| KNN | 1.002 | 0.050 | 0.83 | 0.78 |
| DT | 1.007 | 0.050 | 0.83 | 0.78 |
| RF | 1.001 | 0.044 | 0.83 | 0.78 |
| GBR | 1.000 | 0.042 | 0.83 | 0.78 |
| XGBoost | 1.003 | 0.042 | 0.83 | 0.78 |
| LightGBM | 1.004 | 0.038 | 0.84 | 0.79 |
| CatBoost | 1.001 | 0.042 | 0.83 | 0.78 |
| SVR | 1.001 | 0.052 | 0.83 | 0.78 |
| ANN | 1.002 | 0.062 | 0.82 | 0.77 |
| NWC & LWC ($n$=332) | | | | |
| KNN | 1.003 | 0.062 | 0.82 | 0.77 |
| DT | 1.004 | 0.066 | 0.82 | 0.77 |
| RF | 1.000 | 0.062 | 0.82 | 0.77 |
| GBR | 1.000 | 0.059 | 0.82 | 0.77 |
| XGBoost | 1.000 | 0.057 | 0.82 | 0.77 |
| LightGBM | 1.001 | 0.053 | 0.83 | 0.78 |
| CatBoost | 1.000 | 0.059 | 0.82 | 0.77 |
| SVR | 1.000 | 0.066 | 0.82 | 0.76 |
| ANN | 1.010 | 0.074 | 0.82 | 0.76 |

**Table 7**: Resistance factors for the nominal strength ML models per US design practice based on concrete strength of $f'_{\mathrm{c}}$ and stud strength of $f_{\mathrm{u}}$=450 MPa

| Model | $R_{\mathrm{m}}/R_{\mathrm{n}}$ | $V_{\mathrm{p}}$ | $\phi_{\mathrm{v}}$ $\beta$=3.0 | $\phi_{\mathrm{v}}$ $\beta$=4.0 |
|---|---|---|---|---|
| NWC ($n$=233) | | | | |
| KNN | 1.073 | 0.159 | 0.79 | 0.71 |
| DT | 1.120 | 0.178 | 0.80 | 0.71 |
| RF | 1.072 | 0.128 | 0.82 | 0.75 |
| GBR | 1.131 | 0.130 | 0.86 | 0.79 |
| XGBoost | 1.137 | 0.138 | 0.86 | 0.78 |
| LightGBM | 1.114 | 0.126 | 0.85 | 0.78 |
| CatBoost | 1.129 | 0.121 | 0.87 | 0.80 |
| SVR | 1.143 | 0.191 | 0.80 | 0.71 |
| ANN | 1.150 | 0.164 | 0.84 | 0.75 |
| LWC ($n$=61) | | | | |
| KNN | 1.143 | 0.156 | 0.84 | 0.76 |
| DT | 1.115 | 0.134 | 0.84 | 0.77 |
| RF | 1.136 | 0.117 | 0.88 | 0.81 |
| GBR | 1.175 | 0.116 | 0.91 | 0.84 |
| XGBoost | 1.152 | 0.113 | 0.90 | 0.82 |
| LightGBM | 1.110 | 0.117 | 0.86 | 0.79 |
| CatBoost | 1.133 | 0.124 | 0.87 | 0.80 |
| SVR | 1.073 | 0.172 | 0.77 | 0.69 |
| ANN | 1.219 | 0.129 | 0.93 | 0.85 |
| NWC & LWC ($n$=294) | | | | |
| KNN | 1.088 | 0.160 | 0.79 | 0.72 |
| DT | 1.119 | 0.170 | 0.81 | 0.72 |
| RF | 1.085 | 0.128 | 0.83 | 0.76 |
| GBR | 1.140 | 0.128 | 0.87 | 0.79 |
| XGBoost | 1.140 | 0.134 | 0.86 | 0.79 |
| LightGBM | 1.113 | 0.124 | 0.85 | 0.78 |
| CatBoost | 1.130 | 0.122 | 0.87 | 0.80 |
| SVR | 1.128 | 0.189 | 0.79 | 0.70 |
| ANN | 1.165 | 0.158 | 0.85 | 0.77 |

## 5.3 Scope of application for the developed ML models

The developed ML models may produce inaccurate results for feature ranges outside those given in the test databases used for the model training. Therefore, the applicability of the developed ML models is limited by the following values:

- NWC:
  - 20 MPa $\leq f_{\mathrm{cm}} \leq$ 115 MPa (12 MPa $\leq f_{\mathrm{ck}} \leq$ 90 MPa);
  - 450 MPa $\leq f_{\mathrm{u}} \leq$ 600 MPa;
  - 16 mm $\leq d \leq$ 25 mm;

- $3 \leq h/d \leq 9$.

- LWC:
  - $24 \text{ MPa} \leq f_{\text{cm}} \leq 58 \text{ MPa}$ ($16 \text{ MPa} \leq f_{\text{ck}} \leq 50 \text{ MPa}$);
  - $450 \text{ MPa} \leq f_{\text{u}} \leq 600 \text{ MPa}$;
  - $13 \text{ mm} \leq d \leq 22 \text{ mm}$;
  - $3 \leq h/d \leq 8$.

# 6 Comparisons of developed machine learning models with existing design models

Predictions by the developed ML models for the nominal and design stud shear resistance were compared with those given by the following existing design models:

- Eurocode 4 (EC4) [5, 6]:

$$P_{\text{Rd}} = \frac{min\left\{0.8 f_{\text{u}} \frac{\pi d^2}{4};\ 0.29\alpha d^2 \sqrt{f_{\text{ck}} E_{\text{cm}}}\right\}}{\gamma_{\text{V}}} \tag{8}$$

where $\alpha = 0.2\,(h/d+1)$ for $3 \leq h/d \leq 4$ and $\alpha = 1$ for $h/d > 4$.

- AISC 360 [4]:

$$P_{\text{n}} = 0.5\frac{\pi d^2}{4}\sqrt{f_{\text{c}}' E_{\text{c}}} \leq 0.75\frac{\pi d^2}{4} f_{\text{u}} \tag{9}$$

- JSCE [59]:

$$P_{\text{Rd}} = \frac{min\left\{f_{\text{u}} \frac{\pi d^2}{4};\ 31\frac{\pi d^2}{4}\sqrt{\frac{h}{d}\frac{f_{\text{ck}}}{1.3}} + 10000\right\}}{\gamma_{\text{b}}} \tag{10}$$

where $\gamma_{\text{b}} = 1.3$.

- Pallarés and Hajjar No. 4 (PH4) [7]:

$$P_{\mathrm{n}} = 9\lambda \left(f'_{\mathrm{c}}\right)^{0.5} (d)^{1.4} (h)^{0.6} \leq \frac{\pi d^2}{4} f_{\mathrm{u}} \tag{11}$$

where $\lambda$ is a factor taken as 0.75, 0.85, and 1 for all-lightweight, sand-lightweight, and normal weight concrete, respectively.

- SRN1 [10]:

$$P_{\mathrm{n}} = 1.1\lambda \sqrt[4]{f_{\mathrm{cm}} f_{\mathrm{u}}^3 \frac{h}{d}} \frac{\pi d^2}{4} \tag{12}$$

where $\lambda$ is the concrete type factor, taken as 1.00 for NWC and 0.84 for LWC.

- SRN2 [10]:

$$P_{\mathrm{n}} = (1.1 - 0.1\eta) \sqrt[4]{f_{\mathrm{cm}} f_{\mathrm{u}}^3 \left(\frac{h}{d} - \eta\right)} \frac{\pi d^2}{4} \tag{13}$$

where $\eta$ is the concrete type coefficient taken as 0 for NWC and 1 for LWC.

- SRD1 [10]:

$$P_{\mathrm{Rd}} = \lambda \sqrt[4]{f_{\mathrm{ck}} f_{\mathrm{u}}^3 \frac{h}{d}} \frac{\pi d^2}{4} \frac{1}{\gamma_{\mathrm{V}}} \tag{14}$$

- SRD2 [10]:

$$P_{\mathrm{Rd}} = (1 - 0.1\eta) \sqrt[4]{f_{\mathrm{ck}} f_{\mathrm{u}}^3 \left(\frac{h}{d} - \eta\right)} \frac{\pi d^2}{4} \frac{1}{\gamma_{\mathrm{V}}} \tag{15}$$

It should be noted that the EC4, JSCE, SRD1, and SRD2 models predict the design shear resistance with the partial factor applied, while the AISC 360, PH4, SRN1, and SRN2 models predict the nominal shear strength with no resistance factor. Moreover, the EC4, JSCE, SRD1, SRD2, SRN1, and SRN2 models are based on SI units, whereas the AISC 360 and PH4 models require USCS units.

The EC4 and AISC 360 equations were selected to represent the well-known stud shear resistance models adopted in design standards. Whilst the JSCE model is less well known, it was included in the comparison because it previously demonstrated the best prediction accuracy than other existing models for the design shear resistance of headed studs in NWC slabs [10]. The PH4 model was one of the best for predicting the nominal shear strength of studs in NWC and LWC slabs [10]. SRN1, SRN2, SRD1, and SRD2 are relatively new models derived from employing symbolic regression with genetic programming (GPSR), which showed improved predictions compared with other existing descriptive equations for studs in NWC and LWC slabs. Other existing design models proposed by Hicks [8], Konrad et al. [60], and Hanswille and Porsch [61], as well as those adopted in the AS/NZS standards [62, 63], were previously evaluated against the test databases used in the present study [10].

The design and nominal stud shear resistances predicted by the existing descriptive equations and the developed ML models were compared with the mean measured shear resistance per stud from the NWC and LWC databases. Only GBR, XGBoost, LightGBM, and CatBoost models were considered in the comparisons because they demonstrated better performance than other ML models. The comparisons are summarized in Tables 8 and 9 for the design and nominal resistances, respectively.

Tables 8 and 9 demonstrate that the developed ML models provide considerably more accurate predictions of the design and nominal shear resistances of studs than the existing descriptive equations. For the design shear resistance, the GBR, LightGBM, and CatBoost models produce comparable performance metrics, while the XGBoost model predictions for NWC suffered from a larger reduction factor than those required for other models. All four ML models demonstrate a similar accuracy for the nominal shear resistance. Tables 8 and 9 also show that the test-to-prediction ratios for the proposed ML models are lower than those for the existing models. It indicates that the ML models

**Table 8**: Performance metrics of the existing models and proposed ML models for the design resistance

| Model | RMSE (kN) | MAE (kN) | MAPE (%) | $R^2$ | Test-to-Prediction Ratio min | max | mean | CoV |
|---|---|---|---|---|---|---|---|---|
| NWC | | | | | | | | |
| EC4 | 51.8 | 43.7 | 23.8 | 0.765 | 0.832 | 1.862 | 1.329 | 0.148 |
| JSCE | 24.1 | 19.1 | 11.1 | 0.880 | 0.802 | 1.593 | 1.101 | 0.116 |
| SRD1 | 27.3 | 22.2 | 12.8 | 0.856 | 0.774 | 1.557 | 1.119 | 0.130 |
| SRD2 | 27.3 | 22.2 | 12.8 | 0.856 | 0.774 | 1.557 | 1.119 | 0.130 |
| GBR | 12.6 | 9.1 | 5.3 | 0.958 | 0.727 | 1.310 | 1.031 | 0.064 |
| XGBoost | 22.8 | 20.0 | 11.2 | 0.960 | 0.754 | 1.374 | 1.122 | 0.062 |
| LightGBM | 12.8 | 9.5 | 5.5 | 0.964 | 0.710 | 1.284 | 1.042 | 0.058 |
| CatBoost | 11.6 | 8.0 | 4.9 | 0.957 | 0.664 | 1.197 | 0.999 | 0.065 |
| LWC | | | | | | | | |
| EC4 | 23.5 | 22.1 | 26.7 | 0.812 | 0.946 | 2.571 | 1.389 | 0.152 |
| JSCE | 12.2 | 9.2 | 11.6 | 0.762 | 0.705 | 1.256 | 0.938 | 0.116 |
| SRD1 | 12.0 | 9.9 | 11.8 | 0.817 | 0.840 | 1.708 | 1.128 | 0.120 |
| SRD2 | 11.7 | 9.6 | 11.4 | 0.815 | 0.869 | 1.668 | 1.120 | 0.118 |
| GBR | 4.1 | 3.2 | 3.7 | 0.965 | 0.901 | 1.125 | 1.020 | 0.042 |
| XGBoost | 3.7 | 2.9 | 3.4 | 0.965 | 0.904 | 1.106 | 1.003 | 0.042 |
| LightGBM | 3.4 | 2.4 | 2.8 | 0.970 | 0.826 | 1.109 | 1.004 | 0.038 |
| CatBoost | 3.7 | 2.9 | 3.4 | 0.965 | 0.907 | 1.104 | 1.001 | 0.042 |
| NWC & LWC | | | | | | | | |
| EC4 | 45.9 | 37.9 | 24.6 | 0.858 | 0.832 | 2.571 | 1.345 | 0.150 |
| JSCE | 21.5 | 16.5 | 11.2 | 0.907 | 0.705 | 1.593 | 1.057 | 0.135 |
| SRD1 | 24.1 | 18.9 | 12.6 | 0.911 | 0.774 | 1.708 | 1.122 | 0.127 |
| SRD2 | 24.1 | 18.8 | 12.4 | 0.911 | 0.774 | 1.668 | 1.120 | 0.127 |
| GBR | 10.9 | 7.5 | 4.8 | 0.975 | 0.727 | 1.310 | 1.028 | 0.059 |
| XGBoost | 19.6 | 15.4 | 9.1 | 0.973 | 0.754 | 1.374 | 1.093 | 0.073 |
| LightGBM | 11.1 | 7.6 | 4.7 | 0.978 | 0.710 | 1.284 | 1.032 | 0.056 |
| CatBoost | 10.1 | 6.6 | 4.5 | 0.974 | 0.664 | 1.197 | 1.000 | 0.059 |

produce less conservative stud resistance predictions than the existing models whilst still providing the required reliability level (see Section 5), which was possible due to the smaller scatter of the ML model predictions when compared with the tests.

Fig. 7 shows test-to-prediction ratio distributions for the considered design resistance models based on the combined NWC and LWC database, with the design shear resistance determined using the nominal values of the test database variables. The presented distributions illustrate how safe and conservative the models are.

**Table 9**: Performance metrics of the existing models and proposed ML models for the nominal resistance

| Model | RMSE (kN) | MAE (kN) | MAPE (%) | $R^2$ | Test-to-Prediction Ratio min | max | mean | CoV |
|---|---|---|---|---|---|---|---|---|
| NWC | | | | | | | | |
| AISC | 50.4 | 41.9 | 22.2 | 0.689 | 0.687 | 1.934 | 1.281 | 0.174 |
| PH4 | 21.6 | 16.0 | 9.3 | 0.855 | 0.713 | 1.451 | 1.023 | 0.123 |
| SRN1 | 21.6 | 17.1 | 10.4 | 0.856 | 0.703 | 1.416 | 1.018 | 0.130 |
| SRN2 | 21.6 | 17.1 | 10.4 | 0.856 | 0.703 | 1.416 | 1.018 | 0.130 |
| GBR | 11.4 | 7.7 | 4.8 | 0.958 | 0.705 | 1.271 | 1.000 | 0.064 |
| XGBoost | 11.2 | 7.3 | 4.5 | 0.960 | 0.671 | 1.223 | 0.999 | 0.062 |
| LightGBM | 10.6 | 6.7 | 4.1 | 0.964 | 0.682 | 1.233 | 1.001 | 0.058 |
| CatBoost | 11.6 | 8.0 | 4.9 | 0.957 | 0.664 | 1.197 | 0.999 | 0.065 |
| LWC | | | | | | | | |
| AISC | 9.8 | 8.0 | 10.5 | 0.847 | 0.646 | 1.898 | 1.071 | 0.144 |
| PH4 | 12.8 | 10.8 | 12.8 | 0.687 | 0.786 | 1.590 | 1.085 | 0.144 |
| SRN1 | 8.8 | 7.5 | 9.4 | 0.817 | 0.763 | 1.553 | 1.025 | 0.120 |
| SRN2 | 8.9 | 7.6 | 9.5 | 0.815 | 0.782 | 1.502 | 1.008 | 0.118 |
| GBR | 3.7 | 2.9 | 3.4 | 0.965 | 0.883 | 1.102 | 1.000 | 0.042 |
| XGBoost | 3.7 | 2.9 | 3.4 | 0.965 | 0.904 | 1.106 | 1.003 | 0.042 |
| LightGBM | 3.4 | 2.4 | 2.8 | 0.970 | 0.826 | 1.109 | 1.004 | 0.038 |
| CatBoost | 3.7 | 2.9 | 3.4 | 0.965 | 0.907 | 1.104 | 1.001 | 0.042 |
| NWC & LWC | | | | | | | | |
| AISC | 43.4 | 32.7 | 19.0 | 0.792 | 0.646 | 1.934 | 1.224 | 0.186 |
| PH4 | 19.6 | 14.6 | 10.2 | 0.908 | 0.713 | 1.590 | 1.040 | 0.133 |
| SRN1 | 19.0 | 14.5 | 10.1 | 0.911 | 0.703 | 1.553 | 1.020 | 0.127 |
| SRN2 | 19.0 | 14.6 | 10.2 | 0.910 | 0.703 | 1.502 | 1.015 | 0.127 |
| GBR | 9.9 | 6.4 | 4.4 | 0.975 | 0.705 | 1.271 | 1.000 | 0.059 |
| XGBoost | 9.7 | 6.1 | 4.2 | 0.976 | 0.671 | 1.223 | 1.000 | 0.057 |
| LightGBM | 9.2 | 5.5 | 3.8 | 0.979 | 0.682 | 1.233 | 1.001 | 0.053 |
| CatBoost | 10.1 | 6.6 | 4.5 | 0.974 | 0.664 | 1.197 | 1.000 | 0.059 |

The mean values of the test-to-prediction ratios for the developed ML models are smaller than those for the existing descriptive equations, which indicates that the ML models produce higher shear resistance predictions on average whilst still satisfying the reliability requirements. The coefficients of variation of the test-to-prediction ratios for the developed ML models are comparable with those for the existing models.

Fig. 8 shows the test-to-prediction ratios as functions of the combined NWC and LWC database variables for the nominal resistance LightGBM model and equations from the design standards.

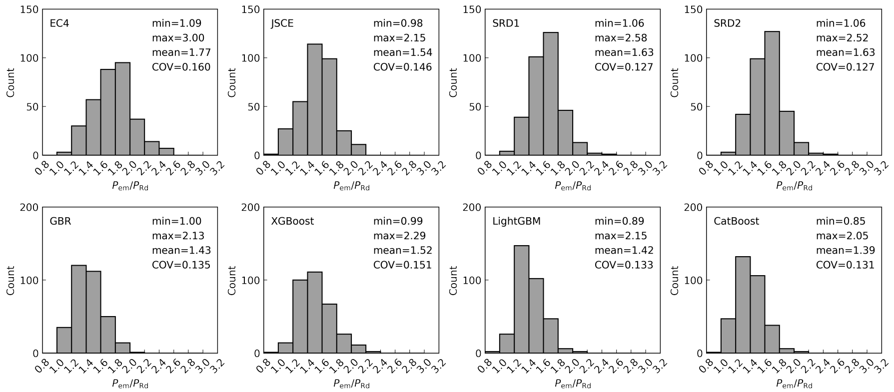



**Fig. 7**: Test-to-prediction ratio distributions per the existing design models and ML models for the combined NWC and LWC database

It can be seen that the LightGBM model produces consistently accurate predictions of the stud shear resistance for the entire range of the variables. The design standard equations show either an increase or a reduction of the test-to-prediction ratios when some of the variables increase, indicating that the design standard equations do not accurately reflect the effect of some variables on the stud shear resistance. A considerably smaller scatter of the test-to-prediction ratios for the LightGBM model compared with the design standard equations can be observed from Fig. 8.

The developed ML models do not include the concrete modulus of elasticity, $E_{\mathrm{cm}}$, as an input parameter. The plot of the test-to-prediction ratio as a function of $E_{\mathrm{cm}}$ for the LightGBM model in Fig. 8 suggests that $E_{\mathrm{cm}}$ does not affect the stud shear resistance if the effect of the concrete strength, which is correlated with $E_{\mathrm{cm}}$, is appropriately considered. This finding was also observed elsewhere [10].

Figs. 9 and 10 present the design and nominal stud resistances normalized by $d^2$ for NWC as functions of the concrete strength for various values of $d$, $f_{\mathrm{u}}$,and $h/d$. It can be noted that the boosting ML models do not produce

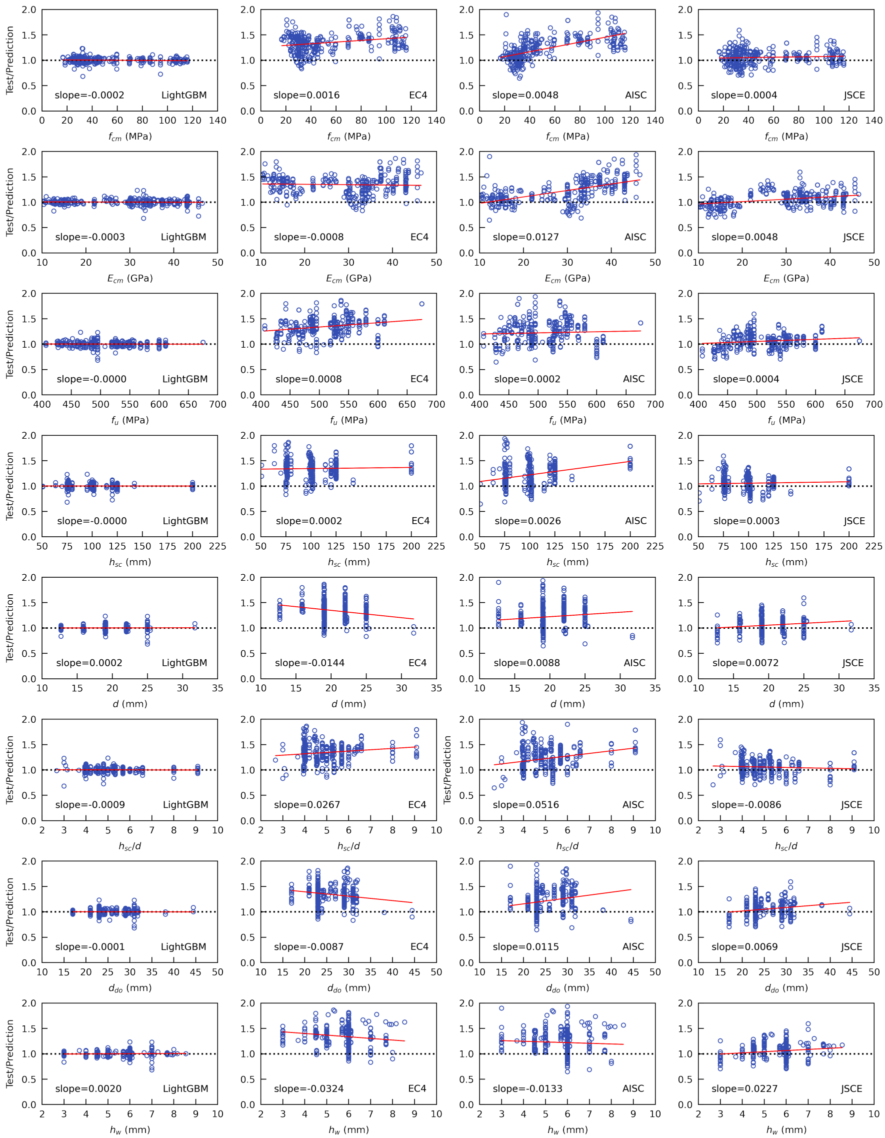


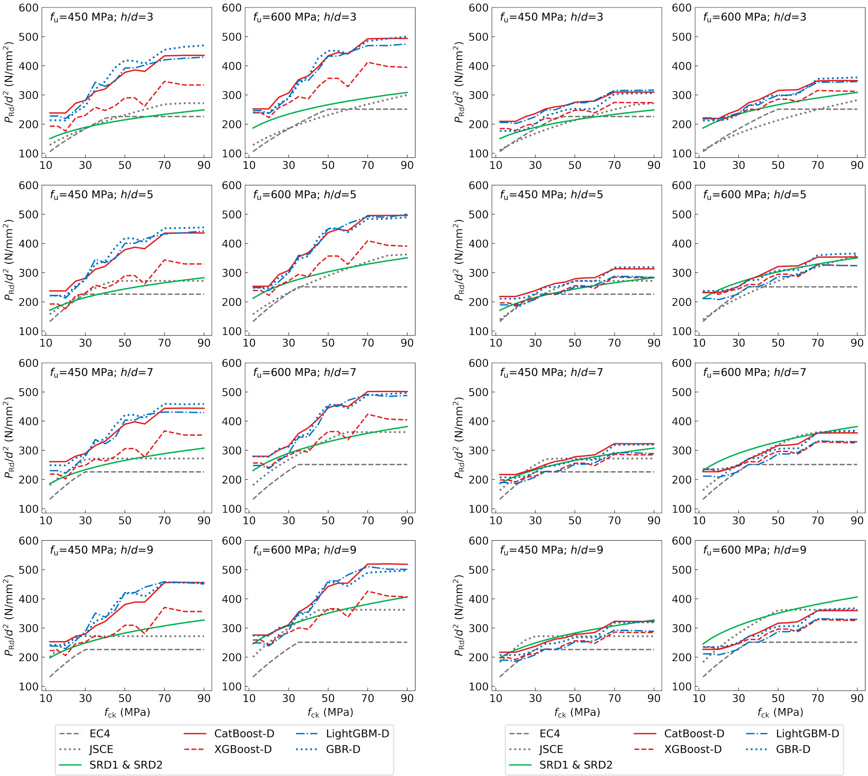

**Fig. 8**: Test-to-prediction ratios versus variables for the LightGBM model and existing design models for the combined NWC and LWC database ($n = 322$ tests)

616 smooth curves due to the nature of the tree-based algorithms and the rela-
617 tively small number of samples in the database. Normalized stud resistance
618 reductions when the concrete strength increases can be observed for some ML

models, especially those based on the XGBoost algorithm, which is caused by the limited number of samples available for the model training. For $d$=16 mm, $h/d$=3, and higher NWC strengths, the ML models predict higher normalized stud resistances than the existing descriptive equations. The predictions of the ML models and the existing equations become closer to each other when the $h/d$ values increase, with the ML models still producing higher normalized stud resistances. For $d$=25 mm, the predictions of the ML models and descriptive equations are closer to each other than those for $d$=16 mm, with the ML models producing higher normalized stud resistances for smaller $h/d$ ratios and smaller normalized stud resistances for larger $h/d$ ratios than the SRD1 and SRD2 models developed in [10]. This observation suggests that the developed ML models are less sensitive to changes in the value of $h/d$ than the SRD1 and SRD2 models.

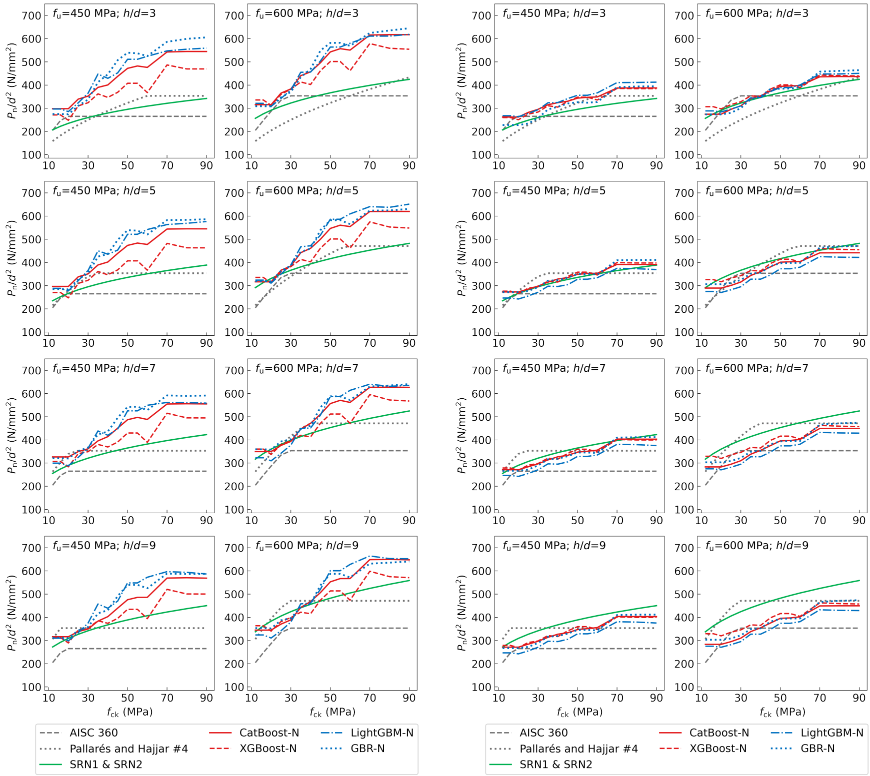
Figs. 11 and 12 show the normalized design and nominal stud resistances as functions of the concrete strength for LWC. The LWC database includes a smaller number of samples than the NWC database, resulting in more inconsistent predictions of the ML models than those for NWC; this can be clearly seen for the stud diameter of 13 mm, where a steady increase in concrete strength first causes the normalized stud resistance to reduce before it increases again. The curves are more smooth for $d$=22 mm. Making more data available for model training should alleviate this issue.

Values of the independent variables significantly affect how the normalized stud resistance predicted by ML models compares with that given by the descriptive equations for LWC. Generally, the ML models produce higher normalized stud resistance for the small values of $d$, $h/d$, and $f_{\mathrm{u}}$, whereas the descriptive equations give higher normalized stud resistances for large values of $d$, $h/d$, and $f_{\mathrm{u}}$.



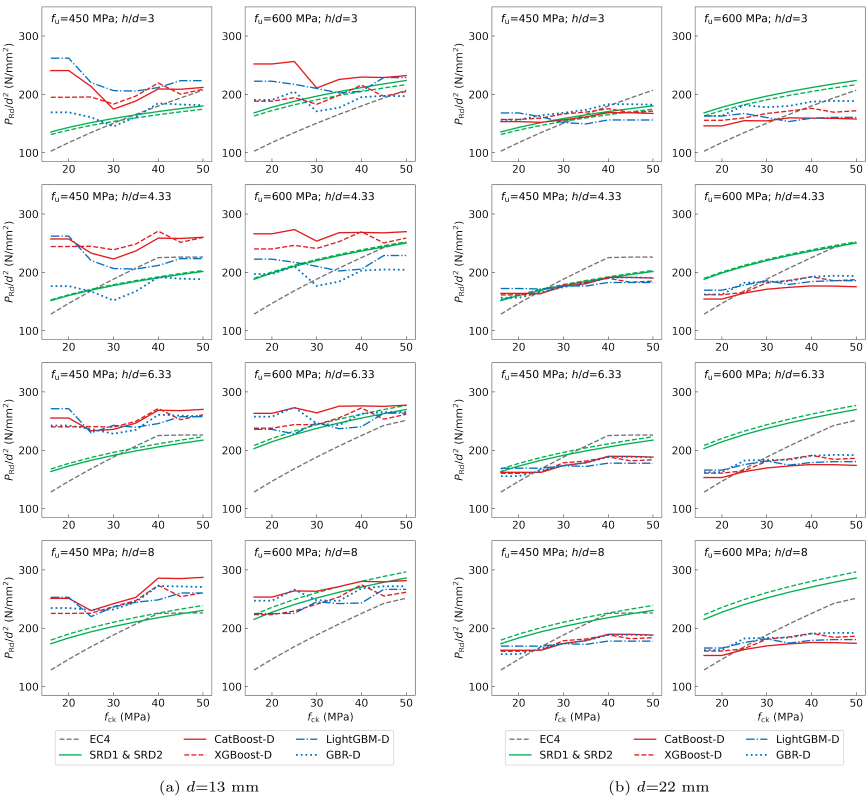

(a) $d$=16 mm (b) $d$=25 mm

**Fig. 9**: Comparisons of design shear resistances of studs in NWC predicted by the existing and proposed models



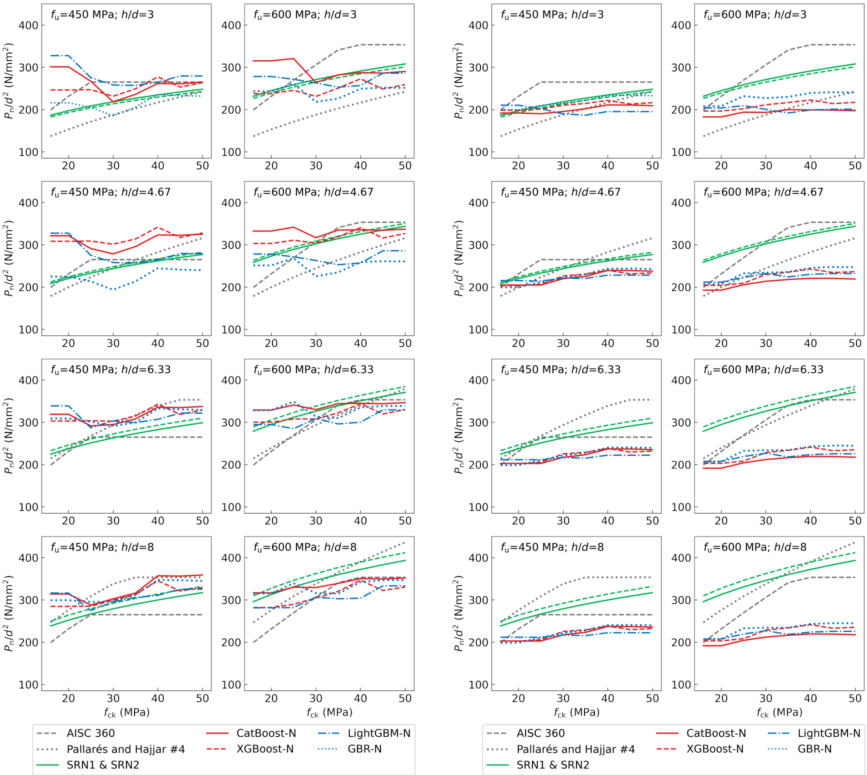

(a) $d$=16 mm (b) $d$=25 mm

**Fig. 10**: Comparisons of nominal shear resistances of studs in NWC predicted by the existing and proposed models




(a) $d$=13 mm (b) $d$=22 mm

**Fig. 11**: Comparisons of design shear resistances of studs in LWC predicted by the existing and proposed models




(a) $d$=13 mm (b) $d$=22 mm

**Fig. 12**: Comparisons of nominal shear resistances of studs in LWC predicted by the existing and proposed models

# 7 Interactive web application

The optimized ML models based on the GBR, LightGBM, and CatBoost algorithms, which demonstrated the best performance, were used for developing an interactive web application in Streamlit (https://streamlit.io). Fig. 13 presents the GUI of the application. It facilitates rapid predictions of the nominal and design stud shear resistance by the ML models based on design practice (Europe or United States), concrete type and strength, stud diameter, stud height-to-diameter ratio, and stud tensile strength specified by the user. Ranges of the input variables available for the selection correspond to those used for the model training (see Section 5.3).

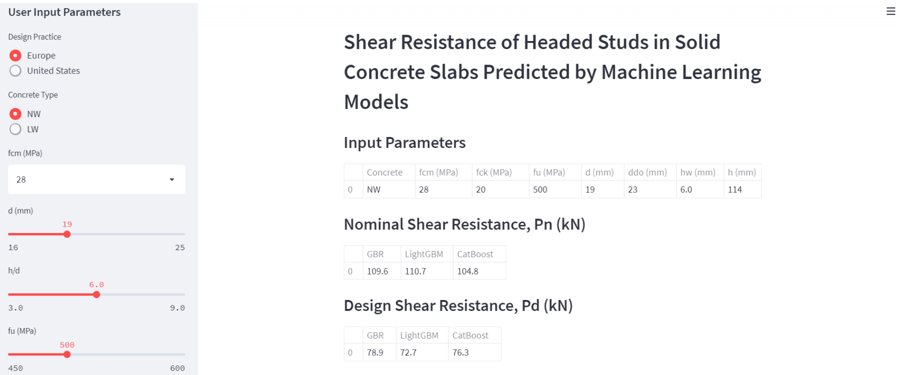



**Fig. 13**: GUI of the interactive web application

The web application also creates and displays stud resistance plots as functions of design variables, with the user-selected input parameters indicated by circle markers. The plots provide additional insight into the effects of design variables on the nominal and design stud resistance. In particular, the plots demonstrate that the stud resistance is insensitive to the $h/d$ changes between 3 and 5, as opposed to the Eurocode 4 provisions requiring stud resistance reductions when $h/d$ is less than 4.

The web application has been deployed to the cloud at http://studs-solid.herokuapp.com/. It can be opened and run in any internet browser on any device, including mobile. The source code of the application available at https://github.com/vitdegtyarev/Streamlit_Studs_Solid can be used for running the app locally.

# 8 Conclusions

This paper presents the development and reliability analysis of nine ML models for predicting the shear resistance of headed studs in solid NWC and LWC slabs. The considered models included KNN, DT, RF, GBR, XGBoost, LightGBM, CatBoost, SVR, and ANN. Databases of push-out test results for studs in NWC and LWC slabs with 242 and 90 samples, respectively, were used for the model development. The input parameters for the models included concrete compressive strength; stud tensile strength, diameter, and height; and weld collar diameter and height.

Optimal hyperparameters of the models were found via an extensive tuning process. The ML models were validated through the ten-fold cross-validation method. The nominal (mean) stud shear resistance predicted by all developed models compared favourably with the test results.

The developed ML models were interpreted by evaluating the SHAP partial importance and dependence, which showed that the stud diameter and concrete compressive strength are the most important features for predicting the shear resistance of studs in NWC and LWC slabs. The SHAP partial importance and dependence plots align well with the mechanics-based knowledge, indicating that the proposed ML models can capture feature importance and dependence from the test data.

The reliability of the predictions by the ML models was subsequently evaluated in accordance with European and US design practices. Reduction coefficients for the ML model predictions required to satisfy the reliability requirements by the Eurocodes were determined. The ML model predictions multiplied by the reduction coefficients produce the design shear resistance of studs. Following US design practice, resistance factors for the ML models based on the target reliability indices of 3.0 and 4.0 were established. The presented study also demonstrated that reliability analyses can serve as an additional test of the ML generalization ability.

The nominal and design resistances obtained with the developed ML models were compared with those computed using existing descriptive equations. The ML models demonstrated considerably better prediction accuracy than the descriptive equations. It was also found that the concrete modulus of elasticity does not affect the stud resistance predictions when the concrete compressive strength is appropriately accounted for in the models.

An interactive web application for predicting the nominal and design shear resistances by the most accurate ML models was created and deployed to the cloud. The application can be run in any internet browser on any device, including mobile. The application's source code, which can be used for running the application locally, has also been made publicly available.

# References

[1] Hicks SJ (2021) Push test database of headed stud connectors embedded in solid concrete slabs. Mendeley Data, https://doi.org/doi.org/10.17632/rfrw3z4hs7.2

[2] Hicks SJ (2021) Database of push tests on headed stud shear connectors embedded in solid slabs using lightweight concrete. Mendeley Data, https:

//doi.org/doi.org/10.17632/xtg3w85hdr.1

[3] Ollgaard JG, Slutter RG, Fisher JW (1971) Shear strength of stud connectors in lightweight and normal-weight concrete. AISC Engineering Journal 8:55–64

[4] ANSI/AISC 360-16 (2016) Specification for structural steel buildings. American Institute of Steel Construction, Chicago, Illinois, USA

[5] EN 1994-1-1 (2004) Eurocode 4: Design of composite steel and concrete structures - Part 1-1: General rules and rules for buildings. European Committee for Standardization, Brussels, Belgium.

[6] EN 1994-2 (2005) Eurocode 4: Design of composite steel and concrete structures – Part 2: General rules and rules for bridges. European Committee for Standardization, Brussels, Belgium.

[7] Pallarés L, Hajjar JF (2010) Headed steel stud anchors in composite structures, part i: Shear. Journal of Constructional Steel Research 66(2):198–212. https://doi.org/10.1016/j.jcsr.2009.08.009

[8] Hicks SJ (2017) Design shear resistance of headed studs embedded in solid slabs and encasements. Journal of Constructional Steel Research 139:339–352. https://doi.org/10.1016/j.jcsr.2017.09.018

[9] Bonilla J, Bezerra LM, Mirambell E, et al (2018) Review of stud shear resistance prediction in steel-concrete composite beams. Steel and Composite Structures 27(3):355–370. https://doi.org/10.12989/scs.2018.27.3.355

[10] Degtyarev VV, Hicks SJ, Hajjar JF (2022) Design models for predicting shear resistance of studs in solid concrete slabs based on symbolic regression with genetic programming. Steel and Composite Structures 43(3):293–309. https://doi.org/10.12989/scs.2022.43.3.293

[11] EN 1990:2002+A1 (2005) Eurocode: Basis of structural design. European Committee for Standardization, Brussels, Belgium.

[12] Kartam N, Flood I, Garrett JH (1997) Artificial Neural Networks for Civil Engineers: Fundamentals and Applications. American Society of Civil Engineers

[13] Salehi H, Burgueño R (2018) Emerging artificial intelligence methods in structural engineering. Engineering Structures 171:170–189. https://doi.org/10.1016/j.engstruct.2018.05.084

[14] Sun H, Burton HV, Huang H (2020) Machine learning applications for building structural design and performance assessment: state-of-the-art review. Journal of Building Engineering p 101816. https://doi.org/10.1016/j.jobe.2020.101816

[15] Thai HT (2022) Machine learning for structural engineering: A state-of-the-art review. Structures pp 448–491. https://doi.org/10.1016/j.istruc.2022.02.003

[16] Jiang F, Jiang Y, Zhi H, et al (2017) Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology 2(4). https://doi.org/dx.doi.org/10.1136/svn-2017-000101

[17] Tadapaneni NR (2019) Artificial intelligence in finance and investments. International journal of innovative research in science, engineering and

technology 9(5)

[18] Ma Y, Wang Z, Yang H, et al (2020) Artificial intelligence applications in the development of autonomous vehicles: a survey. IEEE/CAA Journal of Automatica Sinica 7(2):315–329. https://doi.org/10.1109/JAS.2020.1003021

[19] Abambres M, He J (2019) Shear Capacity of Headed Studs in Steel-Concrete Structures: Analytical Prediction via Soft Computing, URL https://hal.archives-ouvertes.fr/hal-02074833, working paper or preprint

[20] Avci-Karatas C (2022) Application of machine learning in prediction of shear capacity of headed steel studs in steel–concrete composite structures. International Journal of Steel Structures pp 1–18. https://doi.org/10.1007/s13296-022-00589-z

[21] Setvati MR, Hicks SJ (2022) Machine learning models for predicting resistance of headed studs embedded in concrete. Engineering Structures 254:113,803. https://doi.org/10.1016/j.engstruct.2021.113803

[22] Wang X, Liu Y, Chen A, et al (2022) Auto-tuning ensemble models for estimating shear resistance of headed studs in concrete. Journal of Building Engineering 52:104,470. https://doi.org/10.1016/j.jobe.2022.104470

[23] Toffolon A, Kraus MA, Taras A (2021) Deep learning based method for the prediction of the buckling resistance of SHS and RHS. ce/papers 4(2-4):1076–1084. https://doi.org/10.1002/cepa.1398

[24] Zarringol M, Thai HT, Naser M (2021) Application of machine learning models for designing CFCFST columns. Journal of Constructional Steel

Research 185:106,856. https://doi.org/10.1016/j.jcsr.2021.106856

[25] Wakjira TG, Ibrahim M, Ebead U, et al (2022) Explainable machine learning model and reliability analysis for flexural capacity prediction of RC beams strengthened in flexure with FRCM. Engineering Structures 255:113,903. https://doi.org/10.1016/j.engstruct.2022.113903

[26] Xu Y, Zheng B, Zhang M (2021) Capacity prediction of cold-formed stainless steel tubular columns using machine learning methods. Journal of Constructional Steel Research 182:106,682. https://doi.org/10.1016/j.jcsr.2021.106682

[27] Xu Y, Zhang M, Zheng B (2021) Design of cold-formed stainless steel circular hollow section columns using machine learning methods. Structures 33:2755–2770. https://doi.org/10.1016/j.istruc.2021.06.030

[28] Fang Z, Roy K, Mares J, et al (2021) Deep learning-based axial capacity prediction for cold-formed steel channel sections using Deep Belief Network. Structures 33:2792–2802. https://doi.org/10.1016/j.istruc.2021.05.096

[29] Fang Z, Roy K, Chen B, et al (2021) Deep learning-based procedure for structural design of cold-formed steel channel sections with edge-stiffened and un-stiffened holes under axial compression. Thin-Walled Structures 166:108,076. https://doi.org/10.1016/j.tws.2021.108076

[30] ACI 318-19 (2019) Building code requirements for structural concrete. American Concrete Institute, Farmington Hills, Michigan, USA

[31] Roik K, Hanswille G, Cunze-O Lanna A (1989) Harmonisation of the European construction codes – Eurocode 2, 4 and 8/Part 3 – Report on

Eurocode 4 Clause 6.3.2 Stud connectors, Report EC4/8/88, Institut für Konstruktiven Ingenieurbau, Ruhr-Universitat-Bochum

[32] Stark J, van Hove B (1991) Statistical analysis of push-out tests on stud connectors in composite steel and concrete structures, Part 2: Solid Concrete Slabs, TNO report BI-91-163, Delft

[33] fib (2013) fib Model Code for Concrete Structures 2010. federation internationale du beton, Lausanne, Switzerland.

[34] Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media

[35] Aggarwal CC (2018) Neural networks and deep learning: A textbook. Springer

[36] Géron A (2019) Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media

[37] Breiman L (2001) Random forests. Machine learning 45(1):5–32. https://doi.org/10.1023/A:1010933404324

[38] Ho TK (1995) Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, IEEE, pp 278–282, https://doi.org/10.1109/ICDAR.1995.598994

[39] Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Annals of Statistics pp 1189–1232

[40] Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 785–794

[41] Ke G, Meng Q, Finley T, et al (2017) LightGBM: A highly efficient gradient boosting decision tree. Advances in neural information processing systems 30:3146–3154

[42] Dorogush AV, Ershov V, Gulin A (2018) CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:181011363

[43] Vapnik V (1995) The nature of statistical learning theory. Springer, New York, NY

[44] Cortes C, Vapnik V (1995) Support-vector networks. Machine learning 20(3):273–297

[45] Vapnik V, Golowich SE, Smola A, et al (1997) Support vector method for function approximation, regression estimation, and signal processing. Advances in neural information processing systems pp 281–287

[46] Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine learning in python. the Journal of machine Learning research 12:2825–2830

[47] Chollet F, et al (2015) Keras. https://github.com/fchollet/keras

[48] Naser M, Alavi AH (2021) Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences. Architecture, Structures and Construction https://doi.org/10.1007/s44150-021-00015-8

[49] Chicco D, Warrens MJ, Jurman G (2021) The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. PeerJ Computer Science 7:e623

[50] Claesen M, Simm J, Popovic D, et al (2014) Easy hyperparameter search using Optunity. arXiv preprint arXiv:14121114

[51] Clerc M, Kennedy J (2002) The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE transactions on Evolutionary Computation 6(1):58–73

[52] Naser MZ (2021) An engineer's guide to eXplainable Artificial Intelligence and Interpretable Machine Learning: Navigating causality, forced goodness, and the false perception of inference. Automation in Construction 129:103,821. https://doi.org/10.1016/j.autcon.2021.103821

[53] Lundberg S, Lee SI (2017) A unified approach to interpreting model predictions. arXiv preprint arXiv:170507874

[54] Peleg B, Sudhölter P (2007) Introduction to the theory of cooperative games, vol 34. Springer Science & Business Media

[55] Mujagić J, Easterling W (2009) Reliability assessment of composite beams. Journal of Constructional Steel Research 65(12):2111–2128. https://doi.org/10.1016/j.jcsr.2009.07.007

[56] Ravindra M, Galambos T (1978) Load and resistance factor design for steel. Journal of the Structural Division 104(9):1337–1353. https://doi.org/10.1061/JSDEAG.0004981

[57] ACI 301-16 (2016) Specifications for structural concrete. American Concrete Institute, Farmington Hills, Michigan, USA

[58] AISC (2016) Steel Construction Manual, 15th edn. American Institute of Steel Construction, Chicago, Illinois, USA

[59] JSCE (2017) Standard Specifications for Steel and Composite Structures: I General Provision, II Structural Planning, III Design . Japan Society of Civil Engineers, Tokyo, Japan

[60] Konrad M, Eggert F, Kuhlmann U, et al (2020) New approach for the design shear resistance of headed studs in profiled steel sheeting with ribs transverse to supporting beam. Steel Construction 13(4):252–263. https://doi.org/10.1002/stco.202000018

[61] Hanswille G, Porsch M (2007) Zur festlegung der tragfähigkeit von kopfbolzendübeln in vollbetonplatten in din 18000-5 und en 1994-1-1. Schriftenreihe des Instituts für konstruktiven Ingenieurbau, Ruhr-Universität Bochum, Festschrift Prof Kindmann, Bochum 6

[62] AS/NZS 2327 (2017) Australian/New Zealand Standard. Composite structures - Composite steel-concrete construction in buildings. Standards Australia/Standards New Zealand, Sydney, Australia/Wellington, New Zealand

[63] AS/NZS 5100.6 (2017) Australian/New Zealand Standard. Bridge design. Part 6: Steel and composite construction. Standards Australia/Standards New Zealand, Sydney, Australia/Wellington, New Zealand