Machine Learning Models to Predict the Static Failure of Double-Lap Shear Bolted Connections

ALMUHANNA, H. <http://orcid.org/0000-0001-7486-6307>, TORELLI, G. and SUSMEL, Luca <http://orcid.org/0000-0001-7753-9176>

Fatigue & Fracture of Engineering Materials & Structures

WILEY

FFEMS Fatigue & Fracture of Engineering Materials & Structures

ORIGINAL ARTICLE OPEN ACCESS

Machine Learning Models to Predict the Static Failure of Double-Lap Shear Bolted Connections

H. Almuhanna[1] | G. Torelli[1] | L. Susmel[2]

[1]School of Mechanical, Aerospace and Civil Engineering, The University of Sheffield, Sheffield, UK | [2]School of Engineering and Built Environment, Sheffield Hallam University, Sheffield, UK

Correspondence: L. Susmel (l.susmel@shu.ac.uk)

Received: 4 April 2025 | **Revised:** 10 June 2025 | **Accepted:** 17 June 2025

Funding: The authors received no specific funding for this work.

Keywords: adaptive boosting | artificial neural network | bolted connections | decision tree | K-nearest neighbors | machine learning | support vector machine

ABSTRACT

This study investigates the potential of machine learning models to predict the failure load and mode of double-lap shear bolted connections. Five algorithms were evaluated: adaptive boosting, artificial neural network, decision trees, support vector machines with radial basis function kernel, and k-nearest neighbors. A dataset comprising 221 experimental and numerical tests with varying input parameters, including different grades of stainless and carbon steel, was used to train the models. Unlike previous studies, the inclusion of diverse materials enabled the development of more generalizable models. To address data limitations, reduce biases associated with data split, and mitigate overfitting, k-fold cross-validation was adopted instead of the conventional 80/20 split. Results show that both regression and classification models achieved high coefficients of determination across most algorithms. Adaptive boosting delivered the most accurate failure load predictions, while artificial neural network achieved the highest accuracy in classifying failure modes. The findings highlight the potential of well-trained machine learning models to outperform traditional codified methods in accurately predicting the structural response of bolted connections, especially when trained on diverse datasets.

1 | Introduction

Joining steel frames in structures using bolted connections requires a careful design procedure illustrated in relevant design standards such as BS EN 1993-1-8 [1] or ANSI/AISC 360-22 [2]. The codified design equations considered in these standard codes estimate the failure loads of bolted connections by adopting safety factors that ensure a robust structural performance of the designed joints. The procedure consists of estimating the final design failure load by using simplistic equations to calculate the failure corresponding to several specific modes and then taking the lowest of the obtained values. The methodology recommended by these standard codes results in conservative estimations for the failure mechanism, which may not reflect the failure mechanism accurately [3, 4]. While past studies generally indicate conservative estimations, some studies have demonstrated that standard approaches can be non-conservative under specific circumstances [5]. Furthermore, the increasing use of bolted connections with different materials including different grades of steel and stainless steel broadens the range of possible failure mechanisms, thus making codified equations more prone to errors. This issue was observed in past studies where authors noticed cases of underestimations and proposed and validated adjusted design equations that yield a better estimation for the specific material used in the test [6–11]. In order to overcome

Summary

- The static strength of bolted joints with varying configurations is investigated.
- Machine learning algorithms are trained and utilized.
- The effect of design variables is assessed using machine learning.
- Machine learning algorithms demonstrate higher accuracy levels than standard approaches.

all these issues, innovative techniques such as machine learning models can be trained and turned into reliable design tools suitable for accurately estimating the failure mechanism of such a complex structural part.

Machine learning is a specialized tool under the artificial intelligence umbrella that leverages collected data to analyze complex situations or problems. It learns from the data to build accurate models that can predict outcomes or generate meaningful insights [12]. Jiang et al. [13] and Jiang and Zhao [14] utilized eight machine learning models (see list in Table 1) to predict the design failure loads and modes of high-strength and stainless steel bolted connections, respectively. These studies collected numerical and experimental design data from past papers, trained the models, and tested/validated them to find the one with the best performance. The overall framework adopted by the authors is reported in the flowchart of Figure 1. It is worth noting that these authors described adaptive boosting (AdaBoost), light gradient boosting machine (LightGBM), extreme gradient boosting (XGBoost), and cat boosting (CatBoost) as sub-models developed from the decision tree algorithm (DT). Jiang et al. [13] found that all the models gave a reliable predictions of failure mechanisms, with support vector machine learning model (SVM-RBF) yielding the highest accuracy of predictions for both failure loads and modes. Moreover, the study concluded that as far as failure modes prediction is concerned, the accuracy of the models surpassed the accuracy of estimations obtained through standard approaches, which ranged from 67.9 to 85.3%. Similarly, Jiang and Zhao [14] found that all machine learning models yield high accuracy, with SVM giving the best results.

In a similar study, Zakir Sarothi et al. [15] evaluated the performance of 11 different algorithms to predict the failure of double-lap shear bolted connections. In contrast to Jiang et al. [13] and Jiang and Zhao [14] the results reported in Zakir Sarothi et al. study found that the random forest (RF) model gave the highest accuracy of predictions. This suggests that the model accuracy may depend on the input variables and quality of collected data, as well as the variation of sources used for the dataset.

This paper collects a dataset of double-lap shear bolted connections and uses it to train five different machine learning models to predict both failure loads and modes using MATLAB software [16]. In contrast to the studies mentioned above, this work incorporates an expanded set of input variables, including material type and yield ratio. These additional variables are added to better describe the dataset, which covers various material types, including different grades of steel and stainless steel. This adjustment ensures that the models effectively capture the complexities associated with these diverse materials and material grades.

2 | Dataset

A combination of 221 experimental and numerical tests data were collected from the technical literature (Table 2). The collected data consisted of various types of carbon steel, including mild and high-strength grades, and different types of stainless steel. Additionally, three types of bolted connections were included, based on one, two, and four bolts, respectively, as schematized in Figure 2. The number of bolts in a connection is here denoted as n_b. The considered geometrical variables included (see Figure 2) the thickness, t; length, l; breadth, b; end distance, e_1; edge distance, e_2; horizontal pitch distance, p_h; and vertical pitch distance, p_v. Note that, for illustrative purposes, geometrical variables, including the edge distance e_2, are shown symmetric about the centerline; however, the actual test specimens differ from this idealized symmetry. The arrangement of plates with two bolts is not always horizontal. Several specimens have a vertical bolt configuration. Consequently, either p_h or p_v is equal to zero depending on the specific arrangement. Furthermore, mechanical properties obtained from both the parent material characterization and the tests run using bolted connections were considered, namely failure load, F_u; yield ratio, F_y / F_u; and failure mode. The use of the yield ratio allowed the models' ability to distinguish between different material types and grades to be enhanced. Other strength-related properties, such as fracture strain or elongation at fracture, were excluded due to inconsistent reporting across sources. Similarly, elastic modulus was omitted given its limited variation and negligible influence on model performance. Figure 3 illustrates a general test setup for double-lap shear bolted connections. Since in a test the two inner plates are virtually identical, it is common practice

TABLE 1 | Machine learning models used in previous double-lap shear bolted connections failure predictions.

Machine learning model	Designation	Reference
Decision tree	DT	[13–15]
Random forest	RF	
Support vector machine with radial basis function kernel	SVM-RBF	
K-nearest neighbors	K-NN	
Adaptive boosting	AdaBoost	
Light gradient boosting machine	LightGBM	
Extreme gradient boosting	XGBoost	
Linear regression	LR	[15]
Ridge regression	RR	
Lasso regression	LR	
Artificial neural network	ANN	

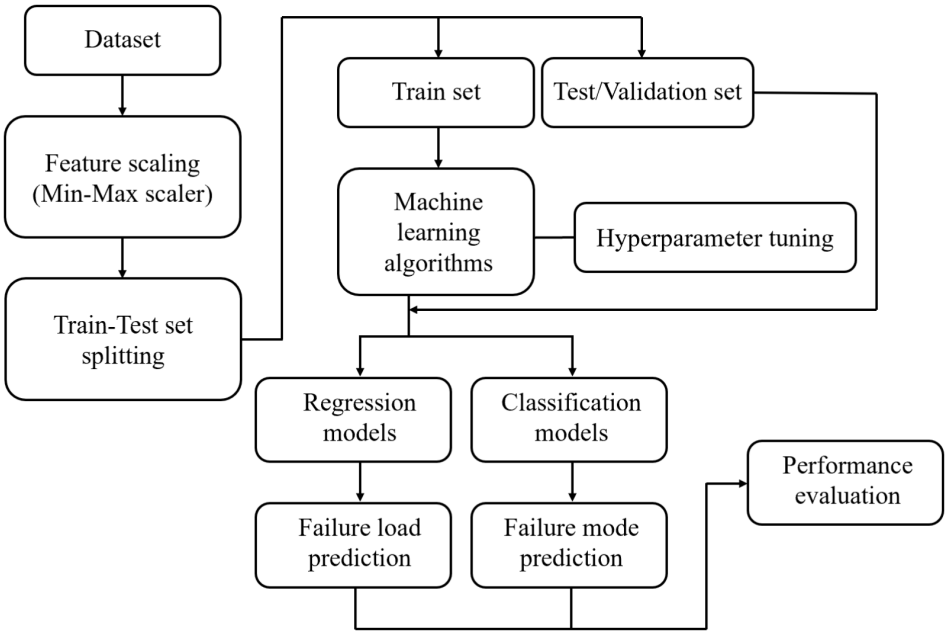

FIGURE 1 | Flowchart for constructing the machine learning models. Adapted from [13].

TABLE 2 | Summary of the literature data on double-lap shear bolted connections.

Source of data	Reference	Material	Number of tests (n)	Bolted connection type	Primary Failure mode
Experiment	[3]	Grade S235 Mild steel	19	1-bolt and 2-bolts	BT, EB, and NST
Experiment	[6]	S690QL ultra high-strength steel	38	1-bolt and 2-bolts	EB and NST
Experiment	[7]	S550Q high-strength steel	8	1-bolt	EB and NST
		S690Q high-strength steel	8	1-bolt	EB and NST
		S890Q high-strength steel	8	1-bolt	EB and NST
Experiment	[8]	Grade EN 1.4301 austenitic stainless steel	3	1-bolt and 2-bolts	EB
		Grade EN 1.4162 lean duplex stainless steel	2	1-bolt and 2-bolts	EB and NST
Experiment	[9]	S550Q high-strength steel	12	2-bolts	EB
		S690Q high-strength steel	12	2-bolts	EB
		S890Q high-strength steel	12	2-bolts	EB
Numerical modeling	[10]	Grade EN 1.4162 duplex stainless steel	84	1-bolt, 2-bolts, and 4-bolts	EB and NST
Experiment	[11]	Grade EN 1.4512 ferritic stainless steel	15	1-bolt and 2-bolts	EB

to replace one of them with an auxiliary plate with increased capacity to ensure one single plate is damaged, with this strategy reducing material waste and experiment costs.

The collected data covered three general failure modes including block tearing (BT), edge bearing (EB) of the hole, and net-section tension (NTS)—see Figure 4. These are the primary failure modes as defined in design standards EN 1993-1-8 and ANSI/AISC 360. While shear-out is also addressed in ANSI/AISC 360, it can be generally considered a form of bearing failure that occurs when the end distance is insufficient to prevent tearing along the shear planes. Normally, the plates suffer these failure modes at the ultimate load, F_u. However, past studies indicate that additional failure modes can occur, such as end splitting and out-of-plane deformation (curling), particularly during the post-necking stage of testing [3, 6–11]. The curling failure was especially observed when the weaker plate was placed as an outer rather than an inner plate in double-lap bolted connection tests [11]. However, for the

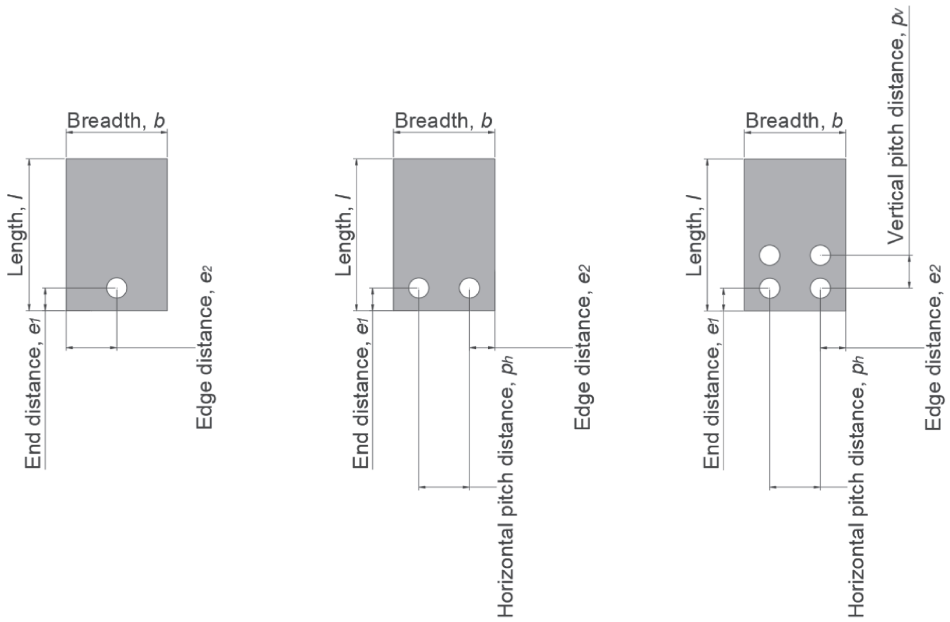

FIGURE 2 | Notation for geometrical parameters of the inner plate based on (a) 1-bolt, (b) 2-bolts, and (c) 4-bolts.

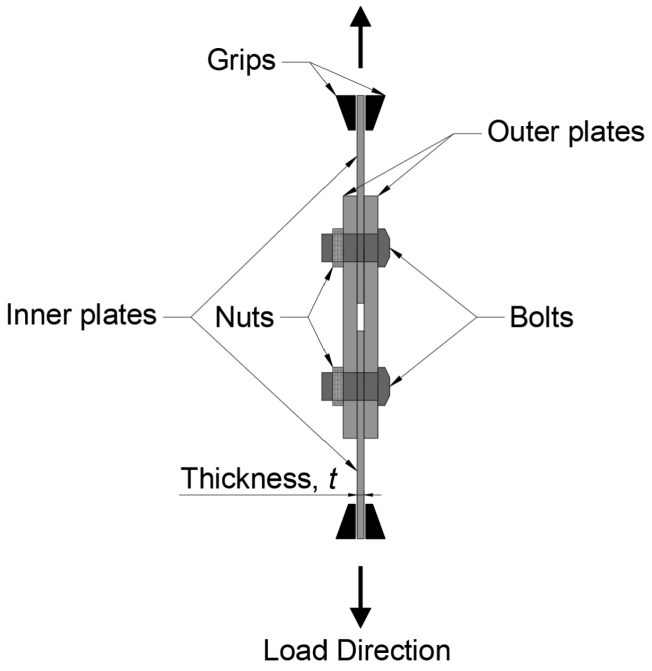

FIGURE 3 | Test setup for double-lap shear bolted connections.

uniformity of the dataset and to be able to use learning models consistently, only the primary failure modes BT, EB, and NST were included. Cases of shear-out were accordingly grouped under EB.

To summarize, the input variables comprised both categorical variables (type of material and failure mode) and numerical variables (n_b, t, l, b, e_1, e_2, P_h, P_v, F_u, and F_y/F_u). The categorical variables were converted into categorical types in the model. This conversion was necessary because machine learning models cannot process text directly. By encoding these variables as categorical types, the model could correctly interpret and distinguish between different categories during prediction.

3 | Machine Learning Models

3.1 | Review and Pre-Processing of the Dataset

To characterize the correlation between all input and output variables of the dataset used in the machine learning models, a correlation matrix was created (Figure 5). The matrix illustrates the relationships between the variables, with correlation coefficients ranging from +1 to −1. A coefficient of +1 indicates a perfect positive correlation, meaning that as one variable increases, the other increases proportionally. Conversely, a coefficient of −1 indicates a perfect negative correlation, meaning that as one variable increases, the other variable decreases. A coefficient of zero indicates no correlation, meaning that the two variables have no relation. In the collected dataset, there were no coefficients equal to zero, indicating that all variables affect each other. An example of a strong positive relation can be seen between the thickness, t and the diameter of the hole, d_o with a coefficient equal to 0.85. This reflects the fact that larger applied loads require both thicker plates and larger diameter bolts (and thus larger holes) to safely transfer the load. As a result, plate thickness, t, is also highly correlated with the failure force, F_u, as shown in the matrix. This correlation aligns with the codified design equations in Section 4, where the failure force increases proportionally with increasing thickness.

Table 3 presents the statistics for the geometrical variables, including the percentage coefficient of variation (COV), which is calculated by dividing the ratio of the standard deviation to the mean. The COV provides a measure of relative variability

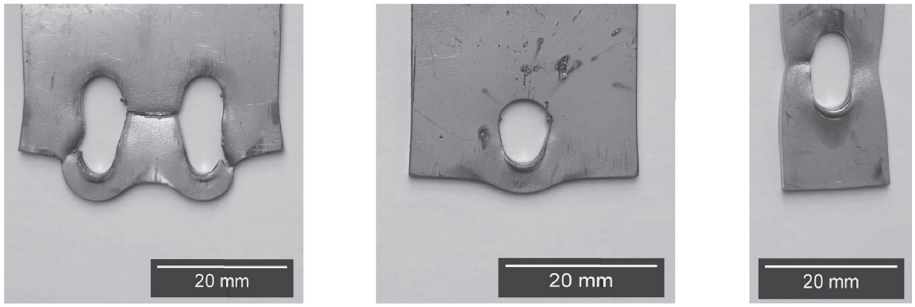

FIGURE 4 | Failure modes of double-lap shear bolted connections: (a) block tearing, (b) edge bearing, and (c) net-section tension.

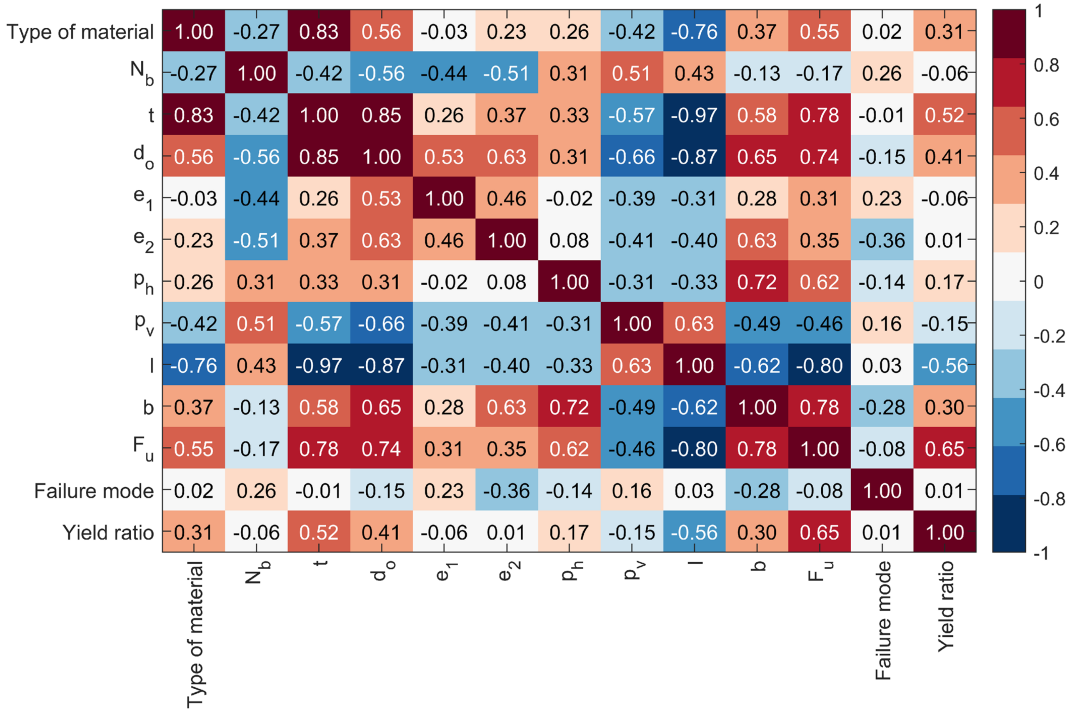

FIGURE 5 | Correlation matrix of dataset variables. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 | Geometrical variables statistics.

Geometrical variables	Minimum (mm)	Maximum (mm)	Mean (mm)	Median (mm)	COV (%)
Thickness, t	0.77	12.00	6.39	9.90	0.68
Diameter of hole, d_o	9.00	30.00	19.69	22.00	0.38
End distance, e_1	9.00	106.20	40.40	38.74	0.55
Edge distance, e_2	14.00	78.04	36.51	31.08	0.48
Horizontal pitch distance, p_h	0	104.00	23.81	0	1.41
Vertical pitch distance, p_v	0	70.00	10.58	0	1.73
Length, l	200.00	440.70	285.80	200.00	0.32
Breadth, b	41.44	216.00	98.24	72.00	0.53

in the data, with higher values indicating greater variation. It is evident that the vertical pitch distance p_v and the horizontal pitch distance p_h show the highest COV. This is largely because, in the dataset, these distances are considered zero for configurations where they do not exist. Specifically, for single-bolted connections, both p_v and p_h are zero, and for double-bolted connections, only p_v is zero. The length of the plates showed the lowest COV since it is almost consistent

for the collected dataset. While the length of the plates shows low variability, it does not have a significant influence on the failure mechanisms, as the primary factors driving failure are strongly related to the other geometrical variables and material properties. It is evident that since the thickness in the dataset ranges from 0.77 to 12 mm, making it have the highest COV reaching 0.68%.

Figure 6 displays the frequency (number of occurrences) of different variables of the machine learning dataset. Figure 6a indicates a high frequency of edge bearing failure making it the dominant failure mode. In comparison, the net-section tension frequency is approximately half that of the edge bearing. The low occurrence of block tearing is attributed to the relatively low number of 4-bolt configurations, as shown in Figure 6b. Additionally, Figure 6c represents the types of materials used, with plates produced from duplex stainless steel being the most prevalent within the dataset.

3.2 | Overview of the Used Machine Learning Algorithms

This study explored the predictive capabilities of five different machine learning algorithms: adaptive boosting (AdaBoost), decision tree (DT), support vector machine (SVM), k-nearest neighbors (K-NNs), and artificial neural network (ANN). The following sections provide an overview of these algorithms, together with a discussion on the implemented cross-validation method.

3.2.1 | Adaptive Boosting (AdaBoost)

AdaBoost is a robust ensemble learning technique that improves the accuracy of weak learners, which are simple models that perform slightly better than random guessing. As illustrated in Figure 7, AdaBoost ensembling begins by training a

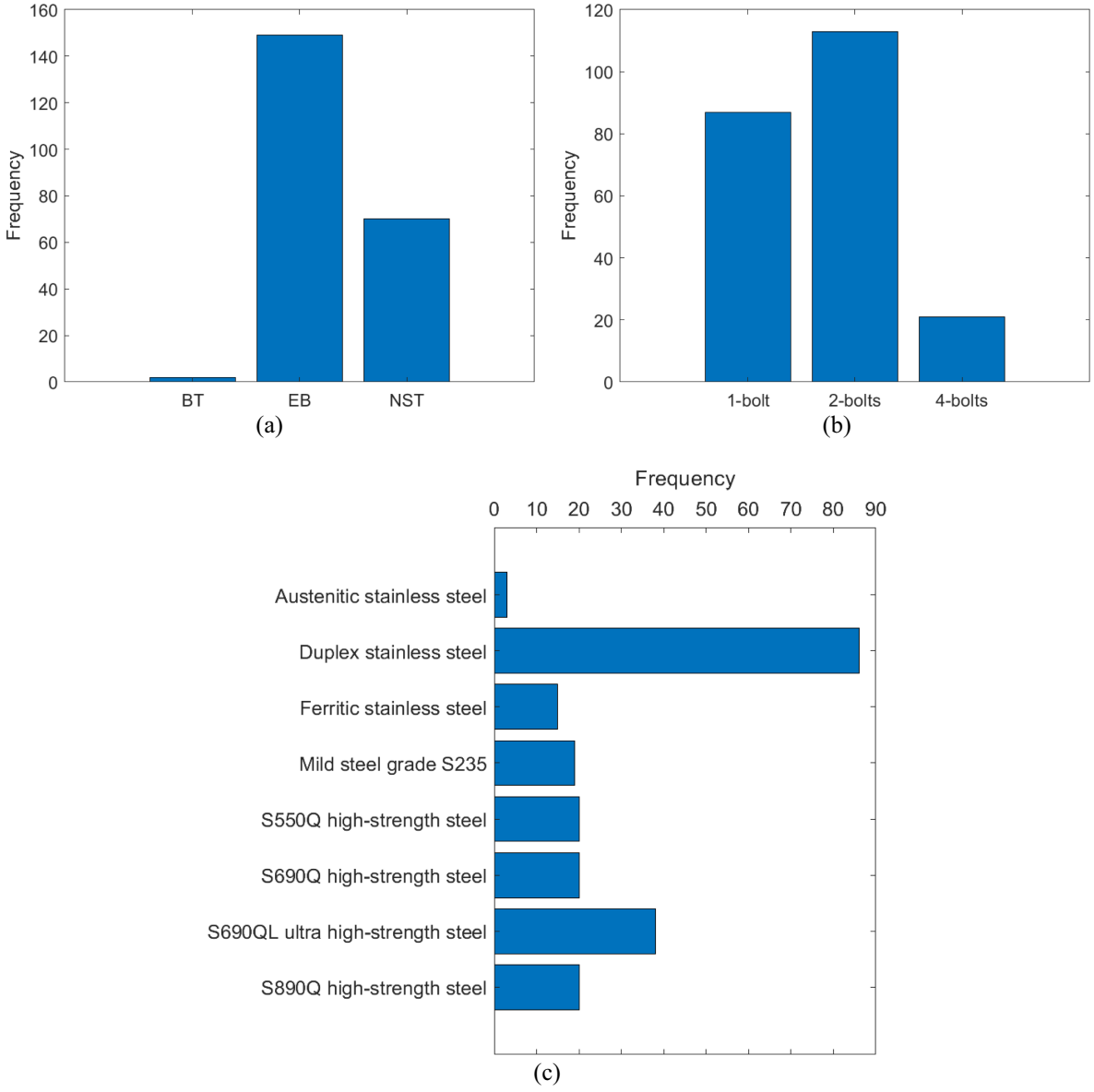

FIGURE 6 | Frequency percentage (%) of input data; (a) failure mode, (b) number of bolts, and (c) material type. [Colour figure can be viewed at wileyonlinelibrary.com]

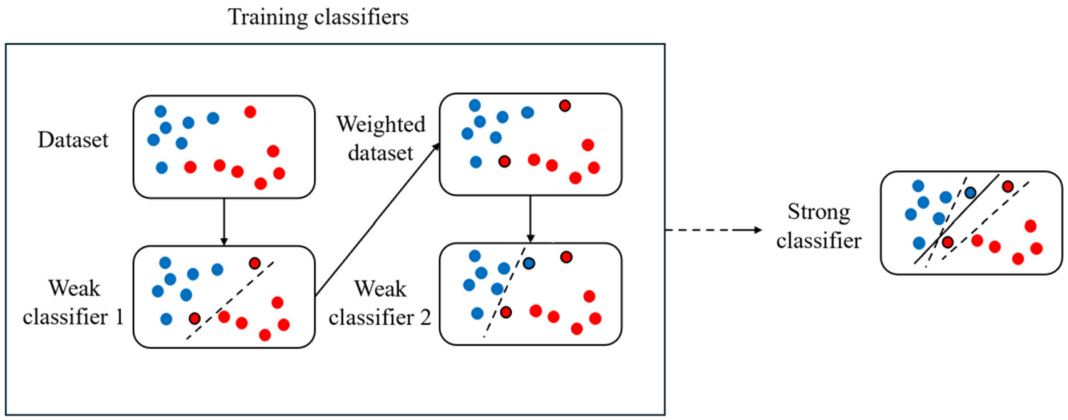

FIGURE 7 | Graphical illustration of adaptive boosting (AdaBoost) machine learning model adapted from [17]. [Colour figure can be viewed at wileyonlinelibrary.com]

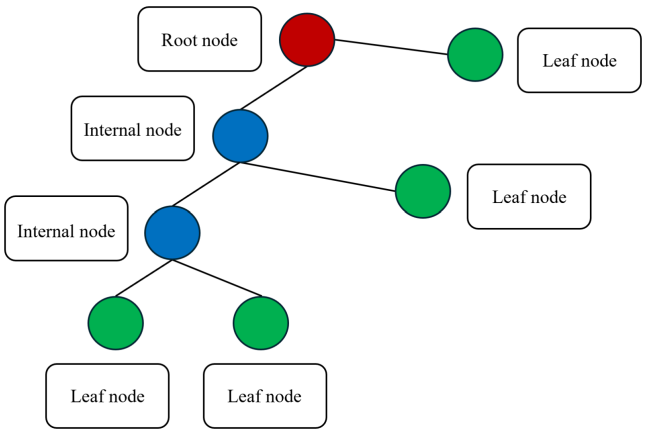

FIGURE 8 | Simplified schematic of decision tree (DT) machine learning model. [Colour figure can be viewed at wileyonlinelibrary.com]

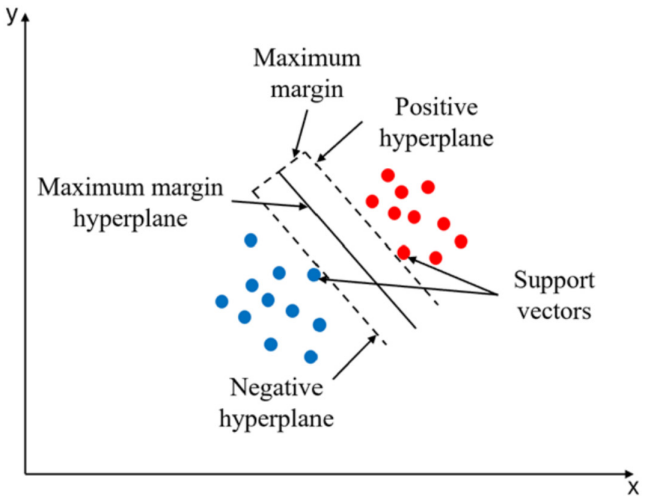

FIGURE 9 | Schematic visualization of one-dimensional support vector machine (SVM) model. [Colour figure can be viewed at wileyonlinelibrary.com]

weak classifier and then adjusting the distribution of the data, where the classifier with the highest errors is given larger date to enhance the performance of the next weak classifier in subsequent rounds. Then, the classifiers are combined, and weights based on their performance are assigned to create a strong classifier [18]. Because of its demonstrated robustness, AdaBoost is here chosen over other machine learning models, including RF, LightGBM, XGBoost, and CatBoost, as they take similar approach [19]. In MATLAB, Adaboost sub-models LSBoost and RUSBoost are used as they are specified for regression, and classification models, respectively.

3.2.2 | Decision Tree (DT)

The DT model (Figure 8) works as a tree-like flowchart that starts with a root node connected with internal nodes by branches. Each node represents an examination of a particular variable, and its corresponding leaf node shows the classification. Based on the outcome of these examinations, the process moves through the branches to reach the leaf node that provides final classification [15]. Although DT is not considered the most effective machine learning algorithm, it is selected in this study because of its simplicity and interpretability, given by the fact that each path from root to leaf can be seen as a clear decision-making process [20].

3.2.3 | Support Vector Machine With Radial Basis Function Kernel (SVM-RBF)

SVM is a type of machine learning model that separates different classes by hyperplanes to create an optimal classifier (Figure 9). The boundaries, known as support vectors, are selected to maximize the margin between the hyperplanes and the nearest training data points from each class to reduce the classification error [21]. As seen in Figure 10, the purpose of the radial basis function kernel (RBF) is to separate classes for a nonlinear problem when a linear boundary is not possible by transforming the data into higher-dimensional space [22]. SVM is selected in this work owing to its proven effectiveness in high-dimensional spaces and its robustness against overfitting [23, 24].

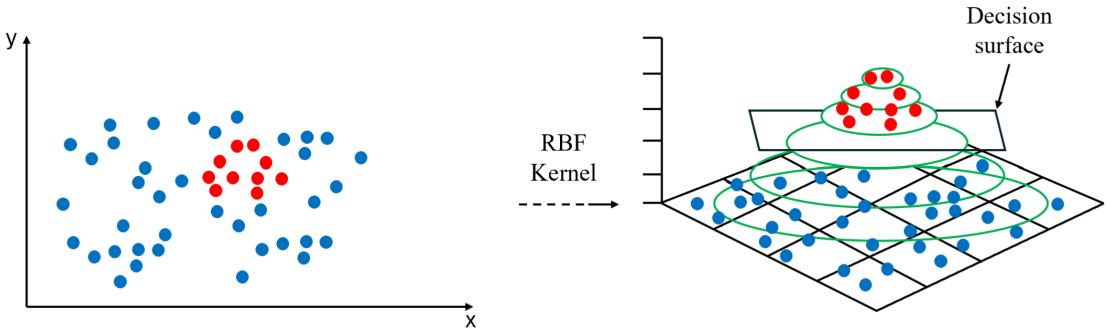

FIGURE 10 | Schematics of the use of radial basis function kernel (RBF) for nonlinear problem adapted from [22]. [Colour figure can be viewed at wileyonlinelibrary.com]

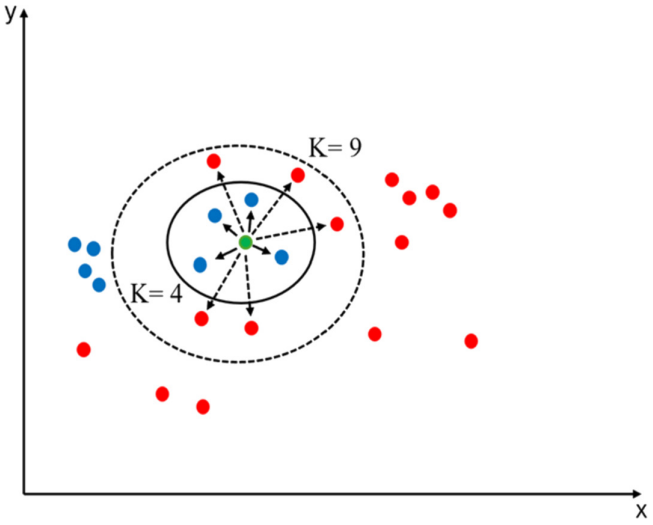

FIGURE 11 | Representation of the k-nearest neighbors (K-NN) model adapted from [26]. [Colour figure can be viewed at wileyonlinelibrary.com]

3.2.4 | K-Nearest Neighbors (K-NN)

The K-NN machine learning model is categorized as a non-parametric since its technique is based on identifying patterns for classifications and predictions, in contrast to other models that depend on applying equations to data [25]. K-NN classifies the unknown variable based on the characteristics of its nearest neighbors, where the neighborhood closeness is defined by a measure such as the Euclidean distance (Figure 11). Noting that major setbacks of this model include the fact that the K value is a user-defined input, which can lead to lower accuracy, and the high computational cost associated with its use [26].

3.2.5 | Artificial Neural Network (ANN)

ANN is a computational model inspired by the functioning of the human brain (Figure 12). It consist of layers of interconnected nodes, analogous to neurons, that work together to process information and solve complex tasks. The network takes in input data, process it through hidden layers, and adjust the strength of connections based on detecting patterns in the input data [28]. ANN has proven to be reliable tools in various applications including structural engineering, as evidenced by several studies, see for instance [27, 29, 30].

3.3 | Training the Data With Cross-Validation and Setting Hyperparameters

MATLAB was the software used to train the models. The common practice of training data is using the hold-out method, splitting them by 80% for training and 20% for testing, with this split remaining fixed. However, k-fold cross-validation is often adopted as a more robust approach that uses all the data for training and validation in each iteration, resulting in improving the accuracy of results for both prediction and classification models [31]. In this study, k-fold cross-validation was implemented for all models by partitioning the data into five k-folds. This means that in each iteration, one-fold (equivalent to 20% of the dataset) was used as the validation set, while the remaining four folds were used for training. It is important to note that increasing the number of folds to higher than five in cross-validation did not lead to significant improvements in models' performance. Therefore, the five folds were considered the optimal configuration. Each k-fold was evaluated separately, and the final performance of the regression and classification models was determined based on the validation results from all k-fold trials, specifically reporting the best overall performance across multiple cross-validation runs. This approach helps ensure an unbiased and robust assessment of model performance by reducing the influence of any single fold that may overperform or underperform relative to the others. The hyperparameters used to tune the models were selected after multiple trials to ensure optimal prediction performance while minimizing the risk of overfitting. Table 4 displays the most important hyperparameters for each model.

4 | Codified Design Estimations

This section presents the equations to estimate the design loads of double-lap shear bolted connections obtained from the most-commonly adopted standard codes, including EN 1993-1-8 and ANSI/AISC 360 for carbon steel, and ANSI/AISC 370-21 [32] for stainless steel. This allows to highlight the differences and similarities between the standards and give a comprehensive understanding of their approaches.

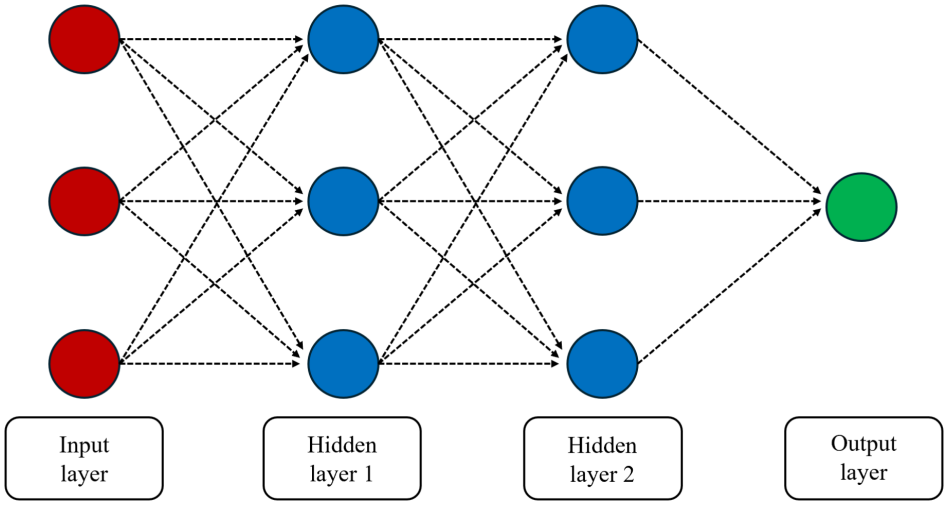

FIGURE 12 | Schematic of artificial neural network (ANN). Adapted from [27]. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 4 | Hyperparameters for each model.

Machine learning algorithm	Hyperparameter	Value
SVM-RBF	Regularization parameter	1000
	Kernel function coefficient	0.8
DT	Maximum number of splits	35
K-NN	Number of neighbors	3
AdaBoost	Number of learning cycles	30
	Learning rate	0.1
ANN	Number of hidden layers	20
	Number of epochs	50

4.1 | EN 1993-1-8

The reference European standard for the design of steel bolted connection is code EN 1993-1-8. According to this code, the design bearing resistance, $F_{b,EC3}$ is calculated using Equation (1):

$$F_{b,EC3} = n_b \times \frac{k_m \times a_b \times f_u \times d \times t}{\gamma_{M2}} \quad (1)$$

where k_m is a reduction factor that depends on the steel grade of the plate, f_u is the nominal ultimate tensile strength of the plate, d is the nominal diameter of the bolt, and a_b defined as the minimum value between e_1 / d_o (edge bearing limit), $3 \times f_{ub} / f_u$ (bolt bearing limit), and 3 (plate bearing/shear-out limit). It is evident that the lowest value of a_b is the value associated with the bearing type that occurs in the connections. The notation γ_{M2} is the partial safety factor for the bolt's resistance.

For calculating the design net-section tension resistance, $F_{t,Rd}$, EN 1993-1-8 refers to Equations (2), (3), (4), (5), and (6):

$$F_{t,Rd} = min\left[F_{pl,EC3}; F_{n,EC3}\right] \quad (2)$$

$$F_{pl,EC3} = \frac{A_g \times f_y}{\gamma_{M0}} \quad (3)$$

$$F_{n,EC3} = \frac{k \times A_{net} \times f_u}{\gamma_{M2}} \quad (4)$$

$$A_{net} = A_g - \Delta A_{net} \quad (5)$$

$$\Delta A_{net} = n_b \times d_o \times t \quad (6)$$

As seen in Equation (2), the design net-section tension resistance, $F_{t,EC3}$ of the plate is the minimum value of the design plastic resistance of the gross cross-section, $F_{pl,EC3}$ or the design ultimate resistance of the net cross-section, $F_{n,EC3}$. Equation (3) is used to calculate $F_{pl,EC3}$, noting that A_g is the cross-sectional area of the plate and γ_{M0} is the partial safety factor for the resistance of cross sections. For calculating $F_{u,EC3}$, Equation (4) is used considering that k is a value that depends on the method implemented for fabricating the hole of the plate, and A_{net} is the net area of the plate cross-section, which is calculated by Equation (5) where, ΔA_{net} is the section area deducted considering the presence of the hole/s. The value of ΔA_{net} can be calculated by using Equation (6).

The standard also recommends that Equation (7) is used to estimate the design block tearing resistance of a plate with two or more bolts, $F_{eff,EC3}$.

$$F_{eff,EC3} = \frac{\left(A_{nt} \times f_u\right) + min\left[\frac{A_{gv} \times f_y}{\sqrt{3}}; \frac{A_{nv} \times f_u}{\sqrt{3}}\right]}{\gamma_{M2}} \quad (7)$$

The parameter A_{nt} is the net area subjected to tension, A_{gv} is the gross area subjected to shear, A_{nv} is the net area subjected to shear, and f_y is the yield strength of the plate.

4.2 | ANSI/AISC 360 and 370

ANSI/AISC 360 and ANSI/AISC 370 are the reference standards for the design of carbon steel and stainless steel bolted connections,

respectively. Similar to EN 1993-1-8, ANSI/AISC 360 gives a set of equations to estimate the failure estimations shear bolted connection, though including additional considerations and different partial safety factors. For instance, Equation (8) is used to estimate the design bearing capacity, $F_{b,AISC360}$ obtained by either including a load and resistance factor design (LRFD) in the numerator or an allowable strength design (ASD) partial safety factor in the denominator, with LRFD and ASD having values of 0.75 and 2, respectively. Noting the values of ASD and LRFD are same for all design estimation equations presented for carbon steel.

$$F_{b,AISC360} = n_b \times C_b \times f_u \times d \times t \times \left(LRFD\ or \frac{1}{ASD}\right) \tag{8}$$

In Equation (8), the coefficient C_b is to be taken as equal to 3 if the bolt-hole deformation at the serviceability limit state is not considered, or 2.4 if it is a design consideration.

Equation (9) represents the net-section tension design resistance, $F_{t,ASCI360}$ according to ANSI/AISC 360. Note that this is the same equations represented in EN 1993-1-8 but with different partial safety factors. The values of $N_{pl,ASCI360}$ and $N_{n,ASCI360}$ are calculated using Equations (10), and (11), respectively.

$$F_{t,ASCI360} = min\left[F_{pl,AISC360}; F_{n,AISC360}\right] \tag{9}$$

$$F_{pl,ASCI360} = A_g \times f_y \times \left(LRFD\ or \frac{1}{ASD}\right) \tag{10}$$

$$F_{n,ASCI360} = A_{net} \times f_u \times \left(LRFD\ or \frac{1}{ASD}\right) \tag{11}$$

For block tearing design resistance, $F_{eff,AISC360}$, as shown in Equation (12) that ANSI/AISC 360 uses the same equation used in EN 1993-1-8 but with the different partial safety factors. Note that the constant factor, which is equal to 0.6, slightly differs from EN 1993-1-8.

$$F_{eff,AISC360} = \left[\left(A_{nt} \times f_u\right) + min\left(0.6 \times A_{gv} \times f_y; 0.6 \times A_{nv} \times f_u\right)\right] \times \left(LRFD\ or \frac{1}{ASD}\right) \tag{12}$$

For stainless steel bolted connections, ANSI/AISC 370 provides similar equations. The only slight adjustment, in relation to the provisions of ANSI/AISC 360 for carbon steel, is for the bearing resistance equation where C_b value is 2.5 fo r $l_2 / d_o > 1.5$ or 2 for $l_2 / d_o \leq 1.5$. Note that the parameter l_2 is the distance from the centre of the hole to the centre of the adjacent hole or the edge of the plate.

Overall, EN 1993-1-8 and ANSI/AISC 360&370 adopt broadly similar approaches for calculating design resistance, although they differ in the formulation and value of the partial safety factors applied. Both standards are intentionally conservative, incorporating these safety factors to account for uncertainties and to ensure the safe design of structural components. For certain connection configurations where design predictions were not reported in the literature, the missing values were manually calculated using the respective code equations without applying partial safety factors, in order to reflect the nominal resistance and allow for consistent comparison with experimental results.

As discussed in Section 5, EN 1993-1-8 generally produced more conservative estimates than ANSI/AISC 360&370. However, due to the omission of safety factors in the manual calculations, the predictions based on ANSI/AISC 360&370 in particular tended to overestimate the failure loads, especially in the mid-to-high load range.

5 | Performance Analysis

5.1 | Regression Models

The regression models and the codes' estimations were evaluated by three general metrics, namely the mean square error (MAE), the root mean square error (RSME), and the coefficient of determination (R^2), using Equations (13), (14), and (15), respectively. Noting that notation n in the equations represents the amount of data. All reported training metrics for the regression models represent the average performance across the training folds in the k-fold cross-validation. Validation metrics are calculated based on the aggregated predictions from all validation folds, providing an overall assessment of each model's generalization performance. The regression plots are also based on these aggregated validation predictions.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|F_u - F_{est}\right| \tag{13}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(F_u - F_{est}\right)^2} \tag{14}$$

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^{n}\left(F_u - F_{est}\right)^2}{\sum_{i=1}^{n}\left(F_u - \frac{1}{n}\sum F_u\right)^2} \tag{15}$$

As shown in Table 5, the performance of AdaBoost surpassed that of the other models and codes across all three metrics for both training and validation sets. Notably, AdaBoost achieved a high validation R^2 value of 0.96, indicating strong agreement with experimental results. Figure 13a shows that AdaBoost produced a well-fitted regression model when comparing the experimental failure loads, F_u with the predicted values, F_{est}. ANN and DT also performed well, with relatively low prediction errors and R^2 values of 0.95, indicating good model fitness (Figure 13b,c). Although SVM-RBF and K-NN both achieved an R^2 of 0.90, their error patterns differed. The SVM-RBF model exhibited the highest MAE among all models, indicating a considerable average prediction error. As shown in Figure 13d, the SVM-RBF predictions are widely scattered around the ideal line, with noticeable deviations in both directions across the mid and upper load ranges. This dispersion reflects a lack of consistency in prediction, despite reasonable correlation with experimental values. Similarly, the K-NN model reached the highest RMSE, possibly due to poor performance on extreme values. As illustrated in Figure 13e, K-NN tended to systematically underestimate the failure loads, particularly at higher values. This bias, along with its high error, demonstrates the model's limited

TABLE 5 | Evaluation metrics for the models and codes performance.

	Train set			Validation set		
Machine learning model designation	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2
AdaBoost	28.82	18.61	0.99	46.21	27.71	0.96
ANN	32.52	18.88	0.98	55.98	33.36	0.95
DT	36.04	19.13	0.98	56.31	32.58	0.95
SVM-RBF	33.57	28.28	0.98	76.30	56.12	0.90
K-NN	69.83	36.97	0.92	77.27	48.21	0.90
Codified design estimations	RMSE	MAE	R^2			
EN 1993-1-8	56.64	37.70	0.95			
ANSI/AISC 360&370	95.18	55.61	0.85			

generalization capability compared to the other machine learning approaches.

The codified predictions in EN 1993-1-8 consistently underestimated the failure loads, as shown in Figure 13f. In contrast, the ANSI/AISC 360&370 predictions exhibited greater scatter and tended to overestimate several values, especially in the mid-to-high load range (Figure 13g). This variability contributed to its relatively lower R^2 and higher error values. Overall, Table 5 and Figure 13 demonstrate that AdaBoost outperformed both code-based approaches in terms of accuracy and consistency, with fewer instances of failure load misprediction. The ANN and DT models, although comparable to EN 1993-1-8 in overall error metrics, provided more balanced predictions with reduced underestimation bias.

5.2 | Classification Models

Figure 14 presents the aggregated performance across the validation sets of k-fold cross-validation for the classification models, shown in the form of confusion matrices. The diagonal blue brackets indicate the correctly predicted failure modes, while the light orange brackets highlight the misclassified instances. As described in [33], recall refers to the percentage of actual failure modes that the model correctly identifies, while precision represents the percentage of predicted failure modes that are correct. Accuracy reflects the overall proportion of correct predictions across all failure modes. Among the models, the ANN demonstrated the best overall performance, achieving the highest accuracy of 91.0%, with partial but improved classification of the rare BT failure mode and strong recall and precision scores for EB and NST. The DT and SVM-RBF models also performed well, with accuracies of 90.0% and 89.1%, respectively, showing consistent results for the dominant failure modes EB and NST, but struggling with the rare BT failure mode. The K-NN model achieved an accuracy of 83.7%, correctly identifying one BT case but showing slightly lower recall for the NST failure mode. In contrast, the AdaBoost model underperformed, with a reduced accuracy of 67.9%, primarily due to frequent misclassifications between EB and NST and a complete failure to correctly classify any BT instances. Overall, the results indicate that ANN, followed by DT and SVM-RBF, offered the most reliable classification of failure modes, while AdaBoost's performance was notably weaker in this task.

6 | Discussion of Results

The results presented in the previous section show that machine learning approaches offer a robust alternative to traditional codified methods for predicting failure loads and modes in bolted connections. While EN 1993-1-8 generally produced conservative estimates, tending to underestimate failure loads, the ANSI/AISC 360&370 predictions, particularly when calculated without safety factors, were observed to overestimate the failure loads in several cases. In contrast, regression models such as AdaBoost, ANN, and DT, as well as classification models including ANN, SVM-RBF, and DT, yielded more accurate and balanced predictions. Hence, these machine learning tools have a significant potential in the design of bolted connections, where accuracy is critical for achieving a design that is both material-efficient and safe.

The regression and classification analyses provided a clear understanding of the performance of the five models used for the design predictions of double-lap shear bolted connections. AdaBoost stood out as the best-performing model for failure load predictions, achieving the lowest prediction errors and highest R^2 value. The regression analysis highlighted a significant variation in the models' performance for predicting failure loads, reflecting the complexity of estimations for such a large quantity of input variables. For classification, the ANN model achieved the highest accuracy, correctly identifying one of the few instances of the BT failure mode and performing strongly across the other failure modes. The DT and SVM-RBF models also performed reliably, while AdaBoost showed the weakest classification performance. The difference in model rankings between regression and classification tasks suggests that some algorithms are better suited to numerical prediction than categorical classification, depending on how the learning algorithm interacts with the data structure. These findings emphasize the importance of selecting machine learning models that are

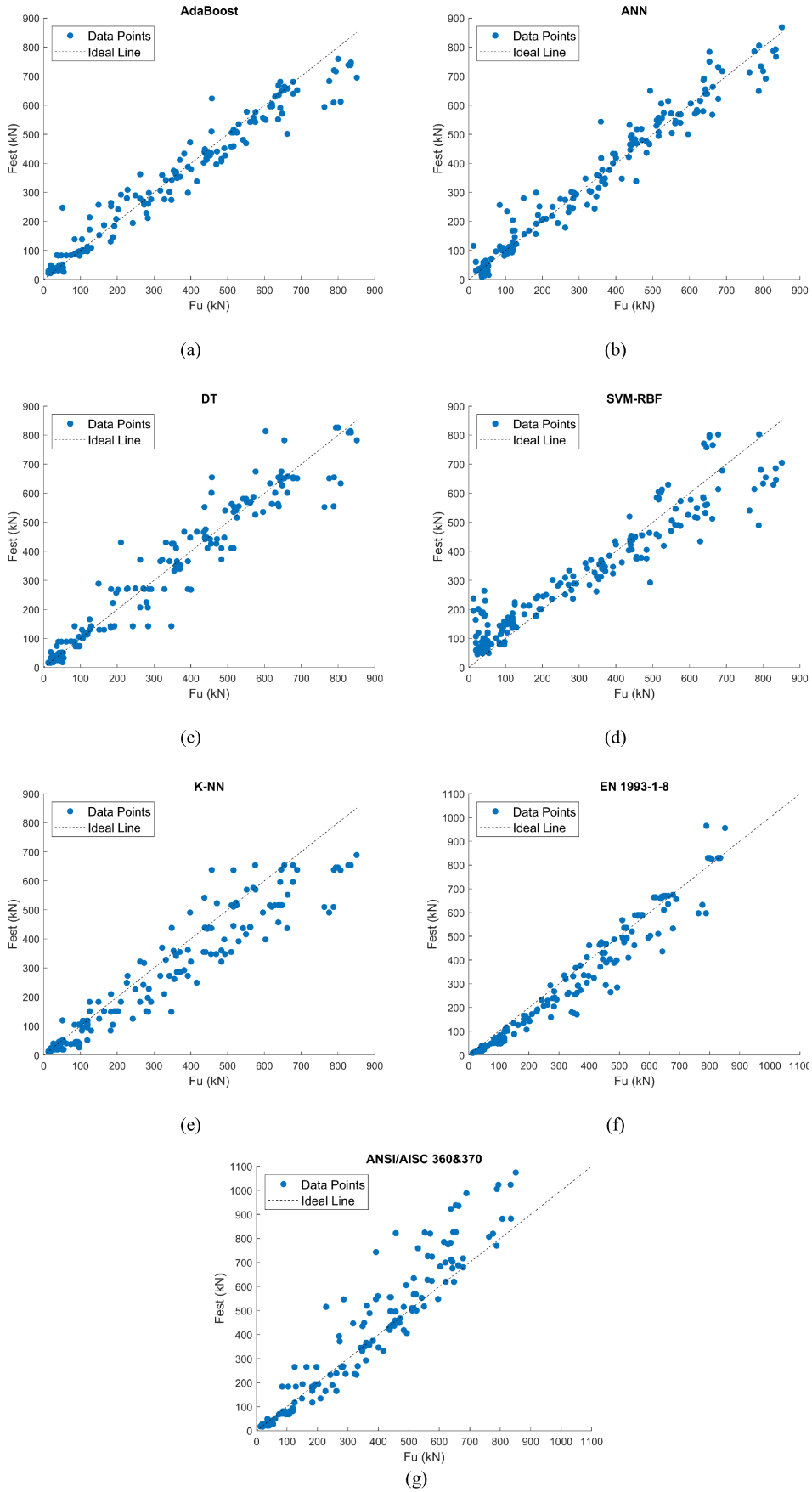

FIGURE 13 | Comparison between actual failure loads, F_u and predicted values, F_{est} by machine learning models and design standards. [Colour figure can be viewed at wileyonlinelibrary.com]

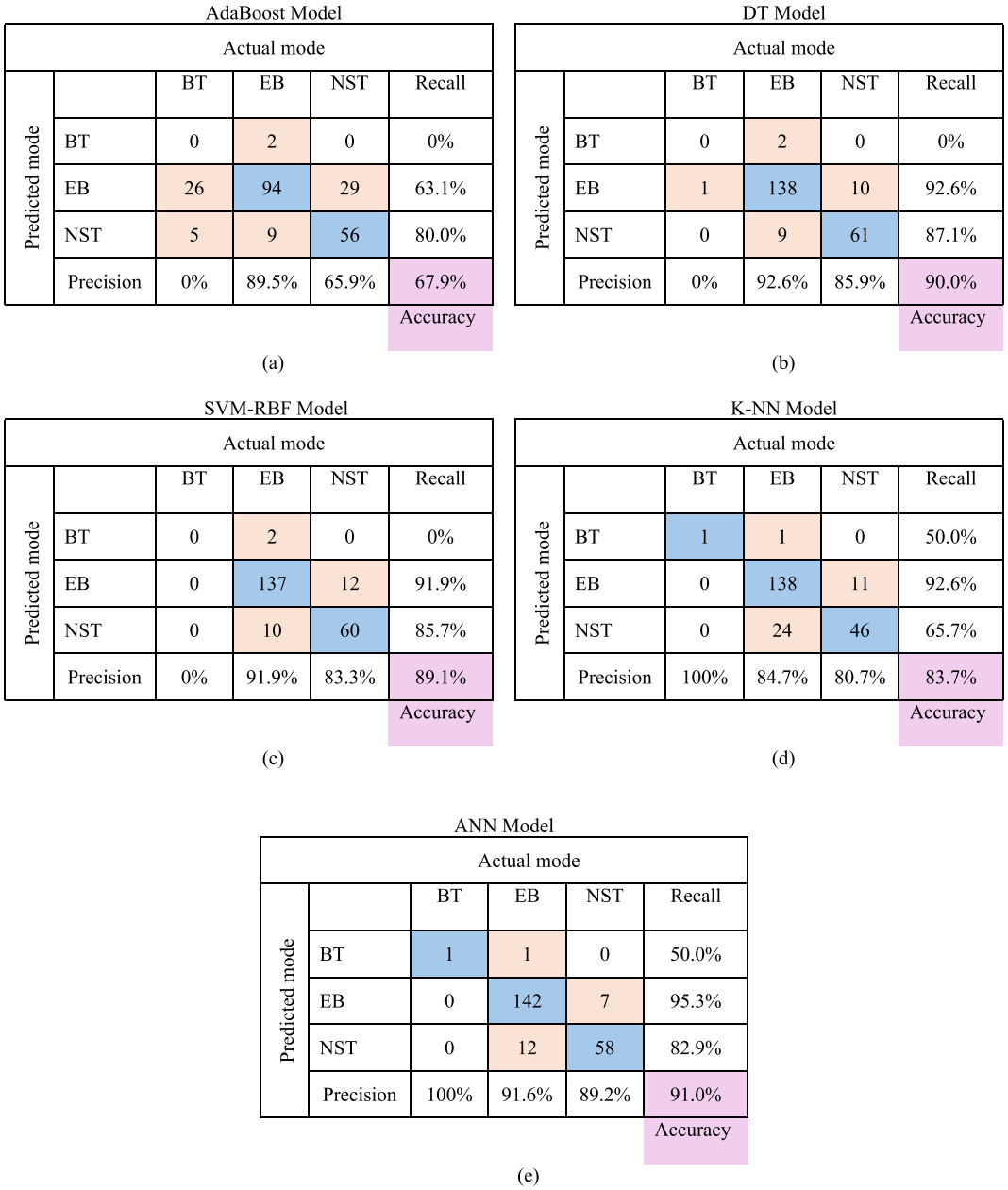

FIGURE 14 | Confusion matrices summarizing the performance of the classification models for failure mode predictions. [Colour figure can be viewed at wileyonlinelibrary.com]

appropriate to the nature of the prediction task to maximize reliability.

7 | Conclusions

This study evaluated the effectiveness of machine learning approaches as an alternative to traditional codified design methods for predicting the failure loads and modes of double-lap shear bolted connections across various types of steel and stainless steel materials. Data from experimental results and numerical analyses were collected from past papers and used to train five different models, including adaptive boosting (AdaBoost), decision tree (DT), support vector machine with radial basis function kernel (SVM-RBF), k-nearest neighbors (K-NNs), and artificial neural network (ANN). k-fold cross-validation was used, with five folds identified as optimal. Based on the results and discussions, the following conclusions are drawn:

- Most of the evaluated machine learning models outperformed traditional codified approaches in accuracy for failure load predictions. In particular, AdaBoost, ANN, and DT demonstrated strong performance in this task.
- AdaBoost achieved the highest R^2 value of 0.96 and the lowest prediction errors across both training and validation sets, indicating excellent fit and generalization.
- The ANN and DT models showed comparable performance to EN 1993-1-8 in regression, with similar R^2 values (0.95) and slightly lower MAE. Although not the top-performing machine learning models, both ANN and DT outperformed the ANSI/AISC 360&370 predictions across all metrics.

Both models also provided more balanced predictions and exhibited reduced underestimation bias compared to the conservative trend observed in EN 1993-1-8.

- Despite its superiority in regression, AdaBoost was the weakest among the machine learning models in classifying failure modes. It misclassified nearly all BT cases and confused EB and NST modes, resulting in the lowest classification accuracy (67.9%). This contrast highlights its limitation in handling imbalanced categorical data and suggests that model suitability varies significantly by prediction type.
- ANN provided a strong balance between accuracy and robustness. It not only achieved one of the lowest regression errors, rivaling AdaBoost, but also achieved the highest classification accuracy at 91.0%. This highlights its adaptability for both numerical and categorical predictions in structural engineering applications, such as double-lap shear bolted connections.
- Although hyperparameter tuning values remained consistent across both regression and classification analyses, the models exhibited notable differences in performance. This shows that model performance is strongly influenced by the type of prediction problem.
- Overall, the study revealed that AdaBoost is the most reliable model for predicting failure loads, while the ANN model demonstrated the best collective performance across both regression and classification tasks. Although this study is limited to static loading, the machine learning framework applied is adaptable and could be extended to predict fatigue-related failure, provided suitable cyclic loading data is available.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. "BS EN 1993-1-8. Eurocode 3. Design of Steel Structures. Part 1–8: Joints". British Standards Institution; 2024.

2. "ANSI/AISC 360–22. Specification for Structural Steel Buildings". American Institute of Steel Construction; 2022.

3. P. Može and D. Beg, "A Complete Study of Bearing Stress in Single Bolt Connections," *Journal of Constructional Steel Research* 95 (2014): 126–140.

4. L. H. Teh and M. E. Uz, "Block Shear Failure Planes of Bolted Connections — Direct Experimental Verifications," *Journal of Constructional Steel Research* 111 (2015): 70–74.

5. L.-T. Hai, G.-Q. Li, Y.-B. Wang, F.-F. Sun, and H.-J. Jin, "Experimental Investigation on Cyclic Behavior of Q690D High Strength Steel H-Section Beam-Columns About Strong axis," *Engineering Structures* 189 (2019): 157–173.

6. P. Može and D. Beg, "High Strength Steel Tension Splices With one or two Bolts," *Journal of Constructional Steel Research* 66 (2010): 1000–1010.

7. Y.-B. Wang, Y.-F. Lyu, G.-Q. Li, and J. Y. R. Liew, "Behavior of Single Bolt Bearing on High Strength Steel Plate," *Journal of Constructional Steel Research* 137 (2017): 19–30.

8. Y. Cai and B. Young, "Carbon Steel and Stainless Steel Bolted Connections Undergoing Unloading and re-Loading Processes," *Journal of Constructional Steel Research* 157 (2019): 337–346.

9. Y.-B. Wang, Y.-F. Lyu, G.-Q. Li, and J. Y. R. Liew, "Bearing-Strength of High Strength Steel Plates in Two-Bolt Connections," *Journal of Constructional Steel Research* 155 (2019): 205–218.

10. J. Dobrić, Y. Cai, B. Young, and B. Rossi, "Behaviour of Duplex Stainless Steel Bolted Connections," *Thin-Walled Structures* 169 (2021): 108380.

11. K. Jiang and O. Zhao, "Ferritic Stainless Steel Thin Sheet Bolted Connections Failing by Bearing–Curling Interaction: Testing, Modelling and Design," *Engineering Structures* 283 (2023): 115919.

12. E. Alpaydin, *Introduction to Machine Learning* (MIT Press, 2014).

13. K. Jiang, Y. Liang, and O. Zhao, "Machine-Learning-Based Design of High Strength Steel Bolted Connections," *Thin-Walled Structures* 179 (2022): 109575.

14. K. Jiang and O. Zhao, "Unified Machine-Learning-Assisted Design of Stainless Steel Bolted Connections," *Journal of Constructional Steel Research* 211 (2023): 108155.

15. S. Zakir Sarothi, K. Sakil Ahmed, N. Imtiaz Khan, A. Ahmed, and M. L. Nehdi, "Predicting Bearing Capacity of Double Shear Bolted Connections Using Machine Learning," *Engineering Structures* 251 (2022): 113497.

16. MathWorks, *MATLAB R2023a* (MathWorks Inc., 2023).

17. P.-H. Kuo, M.-J. Huang, P.-C. Luan, and H.-T. Yau, "Study on Bandwidth Analyzed Adaptive Boosting Machine Tool Chatter Diagnosis System," *IEEE Sensors Journal* 22 (2022): 8449–8459.

18. D.-C. Feng, Z.-T. Liu, X.-D. Wang, et al., "Machine Learning-Based Compressive Strength Prediction for Concrete: An Adaptive Boosting Approach," *Construction and Building Materials* 230 (2020): 117000.

19. M. Adnan, A. A. S. Alarood, M. I. Uddin, and I. u. Rehman, "Utilizing Grid Search Cross-Validation with Adaptive Boosting for Augmenting Performance of Machine Learning Models," *PeerJ Computer Science* 8 (2022): e803.

20. I. Karsznia and K. Sielicka, "When Traditional Selection Fails: How to Improve Settlement Selection for Small-Scale Maps Using Machine Learning," *ISPRS International Journal of Geo-Information* 9 (2020): 230.

21. B. H. Boyle, *Support Vector Machines Data Analysis, Machine Learning and Applications* (Nova Science Publishers, 2011).

22. D. Yu, A. Zhang, and Z. Gao, "Fault Diagnosis Using Redundant Data in Analog Circuits via Slime Module Algorithm for Support Vector Machine," *Journal of Ambient Intelligence and Humanized Computing* 14 (2023): 14261–14276.

23. J. Liu and R. Danait, "Instance Selection in the Projected High Dimensional Feature Space for SVM," in *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)* (IEEE, 2018), 575–579.

24. X.-Y. Wang, J.-W. Chen, and H.-Y. Yang, "A new Integrated SVM Classifiers for Relevance Feedback Content-Based Image Retrieval Using EM Parameter Estimation," *Applied Soft Computing* 11 (2011): 2787–2804.

25. A. Nemes, R. T. Roberts, W. J. Rawls, Y. A. Pachepsky, and M. T. van Genuchten, "Software to Estimate −33 and −1500kPa Soil Water Retention Using the Non-Parametric k-Nearest Neighbor Technique," *Environmental Modelling and Software* 23 (2008): 254–255.

26. A. M. Musolf, E. R. Holzinger, J. D. Malley, and J. E. Bailey-Wilson, "What Makes a Good Prediction? Feature Importance and Beginning to Open the Black box of Machine Learning in Genetics," *Human Genetics* 141 (2022): 1515–1528.

27. V. Aguilar, C. Sandoval, J. M. Adam, J. Garzón-Roca, and G. Valdebenito, "Prediction of the Shear Strength of Reinforced Masonry Walls Using a Large Experimental Database and Artificial Neural Networks," *Structure and Infrastructure Engineering* 12 (2016): 1661–1674.

28. J. Kufel, K. Bargieł-Łączek, S. Kocot, et al., "What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine," *Diagnostics* 13 (2023): 13.

29. S. Das and S. Choudhury, "Influence of Effective Stiffness on the Performance of RC Frame Buildings Designed Using Displacement-Based Method and Evaluation of Column Effective Stiffness Using ANN," *Engineering Structures* 197 (2019): 109354.

30. M. J. Moradi, G. Mehrpour, M. Adelzadeh, and H. Hajiloo, "Structural Damage Levels of Bridges in Vehicular Collision Fires: Predictions Using an Artificial Neural Network (ANN) Model," *Engineering Structures* 295 (2023): 116840.

31. Z. Lyu, Y. Yu, B. Samali, et al., "Back-Propagation Neural Network Optimized by k-Fold Cross-Validation for Prediction of Torsional Strength of Reinforced Concrete Beam," *Materials* 15 (2022): 1477.

32. "ANSI/AISC 370–21. Specification for Structural Stainless Steel Buildings" 2021. American Institute of Steel Construction.

33. L. Rokach, *Pattern Classification Using Ensemble Methods*, 1st ed. (World Scientific, 2010).